Lecture Notes on Data Engineering and Communications Technologies 141

S. Smys
Pavel Lafata
Ram Palanisamy
Khaled A. Kamel *Editors*

Computer Networks and Inventive Communication Technologies

Proceedings of Fifth ICCNCT 2022

Springer

Lecture Notes on Data Engineering and Communications Technologies

Volume 141

The aim of the book series is to present cutting edge engineering approaches to data technologies and communications. It will publish latest advances on the engineering task of building and deploying distributed, scalable and reliable data infrastructures and communication systems.

The series will have a prominent applied focus on data technologies and communications with aim to promote the bridging from fundamental research on data science and networking to data engineering and communications that lead to industry products, business knowledge and standardisation.

Indexed by SCOPUS, INSPEC, EI Compendex.

All books published in the series are submitted for consideration in Web of Science.

S. Smys · Pavel Lafata · Ram Palanisamy ·
Khaled A. Kamel
Editors

Computer Networks and Inventive Communication Technologies

Proceedings of Fifth ICCNCT 2022

Springer

Editors
S. Smys
Department of CSE
RVS Technical Campus
Coimbatore, Tamil Nadu, India

Ram Palanisamy
Gerald Schwartz School of Business
St. Francis Xavier University
Antigonish, NS, Canada

Pavel Lafata
Department of Telecommunication Engineering
Czech Technical University in Prague
Prague, Czech Republic

Khaled A. Kamel
Department of Computer Science
Texas Southern University
Houston, TX, USA

ISSN 2367-4512 ISSN 2367-4520 (electronic)
Lecture Notes on Data Engineering and Communications Technologies
ISBN 978-981-19-3034-8 ISBN 978-981-19-3035-5 (eBook)
https://doi.org/10.1007/978-981-19-3035-5

We are honored to dedicate the proceedings of ICCNCT 2022 to all the participants, reviewers, and editors of ICCNCT 2022.

Preface

With a deep satisfaction, we write this preface to the proceedings of ICCNCT 2022 held in RVS Technical Campus, Coimbatore, Tamil Nadu, India, on April 1–2, 2022.

This conference proceedings volume contains the written versions of most of the contributions presented during the conference of ICCNCT 2022. The conference has provided a setting for discussing the recent developments in a wide variety of topics including network operations and management, QoS and resource management, wireless communications, and delay-tolerant networks. The conference has been a good opportunity for participants who came from different destinations across the globe to present and discuss the topics in their respective research areas.

ICCNCT 2022 tends to collect the latest research results and applications on computer networks and next-generation communication technologies. It includes a selection of 67 papers from 309 papers submitted to the conference from universities and industries all over the world. All accepted papers were subjected to strict peer review by 2–4 expert referees. The papers have been selected for this volume because of their quality and relevance to the conference.

ICCNCT 2022 would like to express our sincere appreciation to all authors for their contributions to this book. We would like to extend our special thanks to all the referees for their constructive comments on all papers, and moreover, we would also like to thank the organizing committee for their hard work. We also wish to

acknowledge the speakers and participants who attended this conference. Finally, we would like to thank Springer publications for producing this volume.

Dr. S. Smys
Professor, Department of CSE
RVS Technical Campus
Coimbatore, India

Dr. Pavel Lafata
Faculty of Electrical Engineering
Czech Technical University
Prague, Czech Republic

Dr. Ram Palanisamy
Professor, Gerald Schwartz School
of Business
St. Francis Xavier University
Antigonish, NS, Canada

Dr. Khaled A. Kamel
Professor, Department of Computer
Science
Texas Southern University
Houston, TX, USA

Contents

About the Editors

Dr. S. Smys received his M.E. and Ph.D. degrees in Wireless Communication and Networking from Anna University and Karunya University, India. His main area of research activity is localization and routing architecture in wireless networks. He serves as Associate Editor of *Computers and Electrical Engineering (C&EE) Journal*, Elsevier, and Guest Editor of *MONET Journal*, Springer. He served as a Reviewer for IET, Springer, Inderscience, and Elsevier journals. He has published many research articles in refereed journals and IEEE conferences. He has been General chair, Session Chair, TPC Chair, and Panelist in several conferences. He is Member of IEEE and Senior Member of IACSIT wireless research group. He has been serving as Organizing Chair and Program Chair of several international conferences and in the program committees of several international conferences. Currently, he is working as Professor in the Department of Computer Science and Technology at RVS technical Campus, Coimbatore, India.

Dr. Pavel Lafata obtained his M.Sc. from the Faculty of Electrical Engineering of the CTU in Prague in 2007 and his Ph.D. from the Faculty of Electrical Engineering of the CTU in Prague in 2011. Since 2011, he is Assistant Professor at the Department of Telecommunication Engineering, CTU in Prague. Od roku 2011 působí jako odborný asistent na Katedře telekomunikační techniky, Fakulty elektrotechnické, ČVUT v Praze. His specialization includes fixed access networks and technologies, VDSL2/G.fast, transmission systems modeling, topology modeling, transmission media and topologies modeling, passive optical networks (PON), PON topology optimization, FTTx, VHDL, and FPGA. Since 2011, he has been Supervisor of more than 20 bachelor and diploma thesis and student projects, since 2014 he has been Supervisor of three Ph.D. students. Since 2011, he has been Teacher and Lecturer of various courses, especially in the field of telecommunication networks and systems, optical systems, digital systems, fixed access networks, etc. He is author or co-author of numerous scientific papers published in reviewed international journals or conferences, he is also Member of several scientific committees of international conferences and editorial boards of international journals, and he also acts as Fellow Reviewer for numerous impact journals.

Dr. Ram Palanisamy is Professor of Enterprise Systems at the Gerald Schwartz School of Business, St. Francis Xavier University, Canada. He obtained his Ph.D. in information systems management from Indian Institute of Technology (IIT), New Delhi, India. He had academic positions at Wayne State University, Detroit, USA; University Telekom Malaysia; and National Institute of Technology, Tiruchirappalli, India. Palanisamy's research has appeared in several peer-reviewed articles in several journals, edited books, and conference proceedings.

Dr. Khaled A. Kamel is currently Professor at Texas Southern university, College of Science and Technology, Department of Computer Science, Houston, TX. He has published many research articles in refereed journals and IEEE conferences. He has more than 30 years of teaching and research experience. He has been General chair, Session Chair, TPC Chair, and Panelist in several conferences and acted as Reviewer and Guest Editor in referred journals. His research interest includes networks, computing, and communication systems.

Cluster Head Selection Based on Mapping-based Cuttlefish Optimization Algorithm for Multipath Routing in MANET

S. Venkatasubramanian and A. Suhasini

Abstract The study develops the mapping-based cuttlefish optimization algorithm (MCFA) to select the cluster head (CH). A local and global search integration is occurred by using this algorithm, and chaos mapping is introduced to solve the issues of trapping in local optimum. NS-2 tools are used and test the efficiency of the MCFA algorithm with ant colony optimization (ACO) in terms of path optimality, delay time, etc. The results prove that the proposed MCFA achieves better performance than the existing ACO algorithm.

Keywords Multipath routing · Network lifetime · Cluster head selection · Mobile ad hoc network · Mapping-based cuttlefish optimization algorithm · Routing overhead

1 Introduction

MANET is a wireless network structure that will not depend on any infrastructure, and it has various self-characteristics such as organizing, sustaining, and regulating mobile hosts [1, 2]. Nowadays, MANET has become a research center because of its incredible potential in various fields such as military battlefields, commercial, emergency aid, and civilian applications [3, 4]. Features of MANET include dynamic terrain, unstable connections, limited power, and a lack of stable infrastructure compared with wired networks. The battery power, storage capacity, and computing power of nodes are terminated in MANET, where these nodes are considered as mobiles devices that are used for communication processes including laptops, mobile phones, iPad, and personal computers. Therefore, these terminals provide a

S. Venkatasubramanian (✉)
Department of Computer Science and Engineering, Saranathan College of Engineering, Trichy, Tamil Nadu 620012, India
e-mail: veeyes@saranathan.ac.in

A. Suhasini
Department of Computer Science and Engineering, Annamalai University, Annamalainagar, Chidambaram, Tamil Nadu 608002, India

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_1

major challenge for designing effective routing protocols [5, 6]. Routing has two functions such as finding the best route and another is transferring the packets. Hence, in such communication, routing is considered as a major part [7, 8], and it is not possible for communication directly without any kind of this infrastructure. Coordination between these mobile nodes is finally required, and routing in multi-hop is adopted, where the distribution of data packets from source to destination is carried out by intermediate node, which acts as a relay in the communication [9, 10].

In MANET, the connection may get failed due to the battery being out of power. Therefore, the power of the mobile device must be taken into account in the routing protocol for extending the network lifetime and ensuring network connectivity [11]. In the last few years, researchers implemented the energy-awareness protocol to minimize the high energy consumption of nodes. Using this above protocol, different path selection schemes are explored for improving the MANET's power efficiency [12]. The network load is balanced, and utilization of the network is improved by diverting traffic to other paths using multipath routing and also minimizing the network congestion while comparing with traditional single-path routing. To some extent, it has effectively improved the credibility of the network [13].

Therefore, the lifetime of various applications such as reliability, network performance, and less energy consumption is highly adjusted by the clustering technique [14]. Selection of CH is the process of choosing a node as the leader node in the cluster, and this block header stores information about its group that includes node list in the cluster and the path to each node. Communication must be done with all the nodes of its group which is the main responsibility of the CH. This communication can take place either directly with the respective CH or through gateways. There are three steps in the communication, where initially, data sent by group members is received by CH, compressing this data as the second step, and at last, the transmission of data is carried out to the final destination or other CH. A suitable group head reduces power consumption and extends network life [15]. Choosing a particular node as the header of a block is a very important but complex one. Several factors can be taken into consideration when choosing the best node as a CH [16]. Some of these factors include node position, movement, energy, confidence, and performance of a node compared to other nodes. WSN and MANET nodes are running out of battery power and resources. The selection process enhances the overall processing efficiency of the network. Therefore, the processing and power limits must be considered to hold the selection process [17].

Due to dynamic traffic and controlled resources, MANETs must maintain a professional dynamic architecture that can detect multipath with minimal overhead traffic. Due to a large amount of node movement, the tip increases the density of multipath routing, which leads to a breakdown of the active connection. Moreover, MANET problems are minimized by many recent applications such as computational intelligence, genetic algorithm-based methodologies, and artificial intelligence. In this research work, MCFA-based CH selection technique is implemented in multipath routing for improving the link scalability and energy efficiency. This MCFA uses the two processes such as reflection and visibility for finding the optimal solution to choose CH in a network. In addition, the problem of local trap is solved by

chaos mapping, and therefore, failure of a link is minimized by increasing the packet delivery ratio.

1.1 Structure of the Research Paper

It is constructed based on the following information: A basic introduction about MANET, multipath routing, and clustering process is given in the introduction. The study of existing techniques that are used in multipath routing is described in Sect. 2. A brief explanation of CH selection using MCFA with cluster formation is provided in Sect. 3. The simulation analysis on the NS-2 tool to validate the effectiveness of the proposed MCFA model with the existing technique is presented in Sect. 4. The scientific contribution of the research methodology is provided in Sect. 5.

2 Related Works

In MANET, energy-efficient navigation is ensured by offering a trust-based approach from the researcher Veeraiah et al. [18], where this paper developed a hybrid algorithm called slap single player with cat algorithm for identifying the optimum paths and routes. In this work, CH is elected by using a fuzzy clustering model that is used based on relevance values of direct and indirect trust values. The calculation is based on the confidence limit of the additional nodes detected. This means that different routes that are planned by all call centers choose the best route and integrate features such as overall route performance, response time, and connectivity. According to the trust value only, the CH is elected, and if any one of the values is mismatched, it is difficult to select the CH.

Prasad and Shankar [19] developed the Energy-Aware Routing on Demand Protocol (EARDP) for avoiding energy wastage in the network model. The protocol keeps MANET as active as possible and acts as an effective routing method when the mode of routing is changed. The best way for MANET routes is to set up the routes so that there is power in the mobile nodes and protocol functions on the network. When the mobile note is idle/disabled, the transceiver's power consumption protocol is disabled. Any communication requests over the network are taken care of by this model. Two more routing processes such as conditional max–min battery capacity and dynamic source routing are considered in this work to test the effectiveness of the EARDP model. However, the scalability of nodes is minimized in this work and occurred as a major issue in the EARDP model. In addition, common problems are raised for both environments when implementing EARDP with two different nodes.

To solve the problem of pathway elongation and energy consumption, the energy-efficient route discovery as EE-OHRA is developed by Kumar and Anuratha [20] in MANET. Routing failures are avoided by taking into account the life span of the path when determining the path. When the total computational load on each node increases

due to route detection techniques, one of the possible route detection methods cancels out this effect. Without malicious nodes, finding the longest longevity path, this model should focus on safety.

For network nodes, fault tolerance and less energy consumption are achieved by Wang [21] using a model based on a genetic algorithm (GA). Different distance parameters between various nodes are considered for obtaining efficient fitness functions. Some of the major characteristics are included in this work such as the base station receiving several hops from the devices, these base stations are sending the available hops to the next hop for process, and finally, the distance is calculated between sending device and base station. The simulation study confirmed the efficiency of the model; however, exponential increases are occurred in the search space of the GA, when the number of elements in the mutation is large.

Ran et al. [22] designed a multipath quality of service-based routing security algorithm on blockchain (MQS) by improving the traditional AODV protocol. The first step is to create a network of nodes in the network, starting with intermediate nodes and storing the location of all nodes in the network. When connected with the blockchain, the model ensured that only trusted nodes are connected and developed a smart contract code, which is used to present the primary communication route in the blockchain network. To be most effective, it should be focused on a large number of attacks.

As suggested by Sharma and Tharani [23], a multipath routing protocol based on ACO with node disjoint is developed by using traditional AODV protocol. In this particular system, one-way detection detects all node split paths from source to destination by eliminating the other routing control overrides. After the first path identification, the movement of data packets begins. All other secondary pathways are detected simultaneously. Toward the destination, the traffic is allowed to flow by modifying all the existing routes from a damaged link, and this process is carried out by the route repair technique, which is focused on in this study. However, the probability distribution of ACO is changed by iterations and high convergence is obtained, which are the major limitations in this model. Therefore, finding the CH for improving the network lifetime is the most challenging task in the environment.

3 Proposed Work

This section presents the overall description of the proposed MCFA-based CH election in the multipath routing protocol.

3.1 Cluster Formation

The nodes are positioned at random locations in MANET, and they are grouped to form different clusters. The nodes are formed into a cluster based on some criteria, and

the nodes that do not fall into the criteria are included in another cluster. Each node in the cluster is assigned with any one of the functions such as member node, CH, and gateway node. The CH coordinates the intra-cluster communication, whereas the gateway coordinates the inter-cluster communication. The CHs communicate with each other through the gateway node to reduce the traffic overhead. Generally, the clusters are formed by broadcasting the beacon messages to the neighbors. This generates more traffic and leads to congestion. The proposed method utilized the characteristics of fish, and the status of the nodes is gathered from their neighbors to reduce the number of beacon transmissions. A threshold RSSI value is estimated, and the nodes lie in a certain cluster based on their transmission range.

By using a Unique Identification Number (UID), the network allows the nodes to join into it, where HELLO message is transmitted by the nodes with its transmission range. According to the neighbor node's highest degree value, the node is elected as CH, which has some important parameters that are described as follows:

- The maximum number of nodes in the cluster is equal to the maximum size of the cluster.
- The maximum number of levels in a cluster is equal to the maximum hop limit.
- The UID of the CH is its ID.
- The CH acts as the advertising node.

The advertising node advertises the ID in its neighborhood range. If any neighbor node wants to join the cluster, a joining request is sent by the node to the CH through the advertising node. The request is received by the CH and accepted, only if:

- The maximum size of the cluster is less than the total number of nodes in the cluster.
- The maximum hop limit is less than the level of requesting nodes.

If the requesting node satisfies the above criteria, the CH accepts it as a cluster client (CC) and allocates an ID. The routing table is updated by registering the CC's ID by CH. The CC is considered as a gateway or border node when it has more than one ID of various clusters. The advertising node is modified into newly registered CC by CH, where in the same cluster, if the CC has more than one ID, then the smallest length ID is advertised by CC. The steps are repeated until the cluster has reached the maximum size.

At the time of cluster formation, UID is assigned only after the node joined the network. To maintain the neighbor table, the HELLO message is broadcast by the nodes. CH is elected by considering which node has the highest number of neighboring values, and it remains as CH for a predefined period, until it dies [24].

3.2 Selection of CH by Considering the Objective Function

In this work, the objective function is achieved by using the distance of the intra-cluster between sensors and the distance from the target node. Consider the average

intra-group activity and target distance of CH as f_1, and it should be reduced for the best CH selection. Suppose, the inverse function of the total strength of all selected CHs is f_2. Both placed functions should be minimized to be within (f_1, f_2) [0,1] to get the best results. The objective function f_1 is depicted as follows:

$$\min f_1 = \sum_{i=1}^{m} \frac{1}{n_i} \left(\sum_{i=1}^{n_i} \theta\left(T_j, \mathrm{CH}_i\right) + \left(\mathrm{CH}_i, \mathrm{BS}\right) \right) \tag{1}$$

where the distance between CH of *i,* target node of *j* is depicted as $\theta(T_j, \llbracket \mathrm{CH} \rrbracket_i)$, the distance between the node, CH is represented as ($\llbracket \mathrm{CH} \rrbracket_i$, BS); finally, a total number of the target node is defined as *n,* and CH is presented as *m.*

The mathematical expression for objective function f_2 is given as:

$$\min f_2 = \frac{1}{\sum_{i=1}^{m} E_{\mathrm{CH}_i}} \tag{2}$$

where E_{CH_i} refers to the mass of the remaining energy from the head. To minimize the linear function and objective function, the optimal CH is identified by using MCFA that is designed in this research work. Mathematically, it represents the integral objective function as (3).

$$F = \mu \times f_1 + (1 - \mu) f_2, \quad 0 < \mu < 1 \tag{3}$$

where μ is the weight parameter in the range of [0,1]. The search operator with the lowest objective value is CH.

In most of the existing techniques for multipath routing, an energy-efficient clustering algorithm is designed to select the optimal CH and increases the network stability by developing ticket ID-based clustering manager (TID-CM) [25] and ticket ID zone manager routing protocol (TID-ZMGR) [26]. The issues of energy consumption and workload of CH are reduced by (TID-CM) with three mechanisms such as ticketing pool, route planning, and ID controller. In [25], TID-CM uses the cluster agent (CA) for monitoring and updating the information about CH in the groups with high energy levels and low distance. In [26], the load is balanced and traffic is controlled by zone leader (ZL) in TID-ZMGR, where the node is elected as ZL by considering various parameters such as link quality, distances, connectivity, and energy. But, in this research work, CH is elected based on finding the optimal solution of MCFA using only two processes such as reflection and visibility and using limited parameters. Moreover, the existing techniques [25, 26] use only 100 nodes with 200 s of simulation time for experimentation, but the proposed MCFA model uses 150 nodes with 160 s of simulation time, which proves that the proposed model is more effective than existing techniques. To get the lowest objective value (optimal solution), the proposed model uses the MCFA that is described as follows. Figure 1

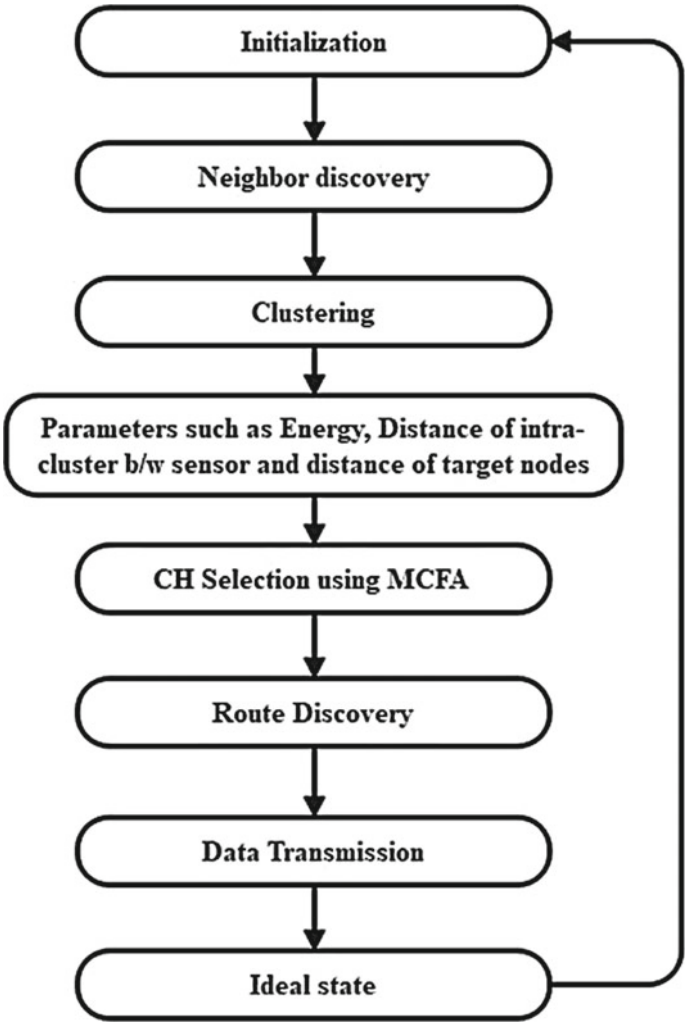


Fig. 1 Flow diagram of proposed methodology

shows the working flow of the proposed methodology.

3.2.1 MCFA

The selection of CH among the cluster groups is determined using a MCFA [27], which is a modified version of CFA that reflects the cuttlefish's skin color-changing mechanisms. CFA uses two main processes to search for the optimal solution: visualization and reflection. The reflection process simulates the reflections of light, and

the vision process simulates seeing the matching patterns. The major feature of CFA is that it can able to integrate the search of local and global values, where during the searching process, the contribution of this integration is adjusted in the initial step. But, the major problem of CFA is that it will get trapped into the local optimum. By enabling the random search, chaos mapping is applied to CFA for solving this above-mentioned problem that brings dynamism to the algorithm. This solution helps the algorithm to overcome the local minima. Therefore, the proposed MCFA suggested following the chaos mapping of CFA which improves the performance of the algorithm in dealing with various optimization problems [28]. Here, MCFA is used as a feature selection technique for predicting heart diseases in [28], where this research study uses the MCFA for finding the optimal solution to elect the CH in MANET's multipath routing technique.

Therefore, a chaos mapping-based population initialization is done in this MCFA. Search supported by chaos designations has the potential to reach most states within a given region without redundancy. Specific chaos assignment equations have been written to generate sequences of chaos in the MCFA proposed in (4) and (5).

$$\mathrm{Cr}_{n+1} = \delta * \mathrm{Cr}_n * (1 - \mathrm{Cr}_n) \quad \text{for } 0 < \delta \leq 4 \tag{4}$$

$$\mathrm{Br}_{n+1} = \delta * \mathrm{Br}_n * (1 - \mathrm{Br}_n) \quad \text{for } 0 < \delta \leq 4 \tag{5}$$

where function's initial value is depicted by δ, Cr and Br denote a function based on the logistic mapping functions, which varies from 0 to 1, and n is the number of nodes present in the network. Using this framework improves the diversity and coverage of solutions. The parameters Cr and Br are used in MCFA, instead of the random parameter used in the original CFA. After random starting, a new solution was found using visualization and reflection as shown in Eq. (6).

$$N_s = R_e + V_s \tag{6}$$

N_s stands for the new solution, R_e stands for reflection, and V_s stands for vision. This algorithm divides the population into four classes. For type 1, the algorithm uses state 1 and state 2 (the association between erythrocytes and chromatophores) to create a new solution. Both events are comprehensive research using each point's value to identify new areas around the best solution in a given period. In type 2, the algorithm uses state 3 (the erythrocyte inversion factor) and state 4 (the interaction between erythrocytes and chromatophores), which act as a local search for a new solution. For type 3, state 5 (interaction between chromatophores and lucopores) is used to create solutions close to the best solution (local search). Finally, for type 3, condition 6 (Lucophorus reflection factor) implements a global search that reflects incoming light without any changes. The equations used to calculate reflection and visibility for the four types are described below.

(a) States 1 and 2 for Type 1

$$R_e[j] = \mathrm{Cr} * G_1[i].\mathrm{points}[j] \tag{7}$$

$$V_s[j] = \mathrm{Br} * (b.\mathrm{point}[j] - G_1[i].\mathrm{points}[j]) \tag{8}$$

where G_1 is type 1, which shows an element/cluster as I in group as G_1; points [j] represent point j of element i in the set; and b is the best solution or optimal solution for electing the CH, where the value of the best points is averaged. In addition, Cr and Br are chaos mapping numbers, Re represents the magnitude of the reflection, and V_s represents the size of the final shape.

(a) States 3 and 4 for Type 2

$$R_e[j] = \mathrm{Cr} * G_1[i].\mathrm{points}[j] \tag{9}$$

$$V_s[j] = \mathrm{Br} * (b.\mathrm{point}[j] - G_2[i].\mathrm{points}[j]) \tag{10}$$

(b) State 5 for Type 3

$$R_e[j] = \mathrm{Cr} * G_1[i].\mathrm{points}[j] \tag{11}$$

$$V_s[j] = \mathrm{Br} * (b.\mathrm{point}[j] - \mathrm{AV_B}) \tag{12}$$

where $\mathrm{AV_B}$ specifies the average value of the best points.

(a) State 6 for Type 4

$$p[i].\mathrm{points}[j] = \mathrm{random} * (U_1 - L_I) + L_I \tag{13}$$

where $i, j = 1, 2, \ldots, n$ and U_1 and L_I indicate the upper and lower limits of the problem domain.

4 Results and Discussion

The selected algorithm simulates an NS-2.33 simulation using the network topology as shown in Table 1. The binary dialog model uses TCP packets to enable the movement of a full 150 nodes in all directions. The 30 dynamic nodes are linked together from the whole nodes and run the simulations for 160 s. For the first fifteen decades, no conversation started. The initial track started moving them after 10 s, and the next track started moving them after 20 s. The conversation ended in 160 s. In MANET, therefore, only packets are selected in connections till they reach their final destination, where no new TCP packets are generated.

Table 1 Simulation setup

Parameter description	Value
Types of antenna	OMNI
Simulation beginning time	11.5 s
Delay propagation	165 ms
Opened connections simultaneously	12 connections of TCP (i.e., 24nodes)
Network type and channels	Wireless
Speed of the nodes	15–50 m/s
Topological dimension	(1600 × 1600)
Types of MAC	IEEE 802.11e
Propagation	2 way
Model of mobility	Random waypoint
Network' pause time	55–355 s

4.1 Performance Parameters

End-to-end Delay (ETED): The average time between packet generation at the source node and the effective transport of the packet at the target node is measured. It calculates all possible delays in the source, including sequence time, node transfer, transmission and restarts on the MAC layer. Class time can be created by network blocking or the inability of appropriate methods.

Number of Packets Dropped: During communication, the data packets are not effectively transmitted to the target destination.

Packet Delivery Ratio (PDR): It effectively measures the number of data packets delivered to each target node and information packets generated by full source nodes.

Throughput: The transfer rate parameter calculates the amount of information that synchronization can provide to the mobile network. Productivity is the pocket size that comes in the sink per MS.

Routing Load: It is a comparative ratio of the number of routing connections reported at each station on the MANET and the total number of data packets effectively delivered to all target stations.

Jitter: Between all other nodes, the packet delay time in standard deviation is mentioned by the metric called jitter.

4.2 Validation Analysis of Proposed Model

In this section, the performance of the proposed MCFA algorithm is tested and its efficiency is compared with existing techniques, namely EE-OHRA [20] and ACO [23] in terms of various parameters. Figure 2 shows the comparative analysis of the proposed model with existing techniques in terms of ETED.

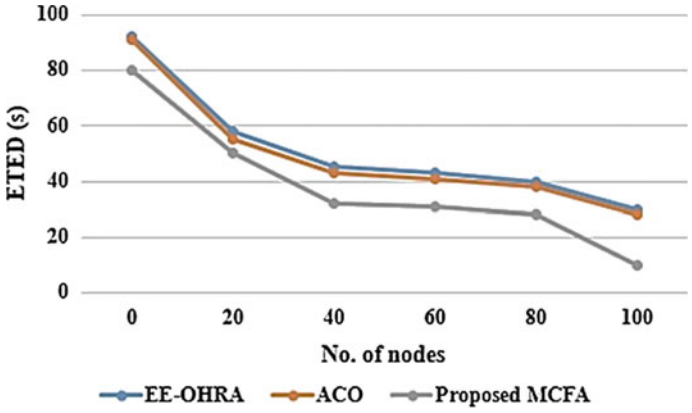


Fig. 2 Graphical representation of MCFA with existing models in terms of ETED

Here, initially, 100 nodes are considered and tested for ETED, where initially, the delay time is measured according to the node pause time that increases the impact of the total number of nodes in the network. When the node is 100, the node pause time of MCFA is only 10 s, where the ACO and EE-OHRA have 28 and 30 s. According to the total number of nodes, it may vary, for instance, when the node is 40, ACO has 43 s, EE-OHRA has 45 s, and the proposed MCFA has only 32 s. From the graph analysis, it is proven that the MCFA has less node pause time than other existing techniques. The reason is that mapping criteria are used in the proposed model for the objective function effectively. Figure 3 provides the experimental analysis of MCFA on average jitter.

When the node increases, the average jitter is also increased, for instance, the MCFA has 0.42 s when the node is 20, and the same method achieved 0.58 s of

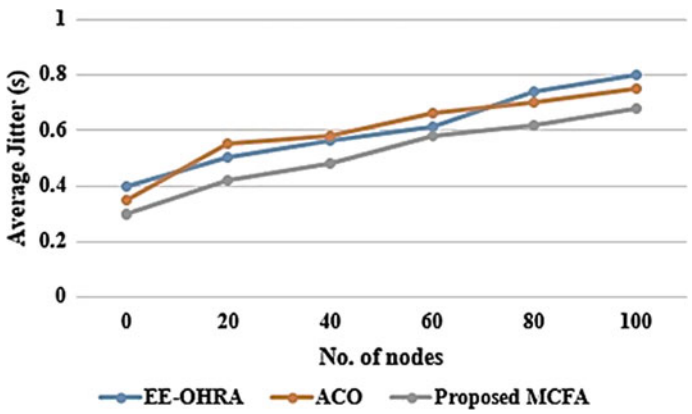


Fig. 3 Graphical representation of MCFA with existing models employing Jitter

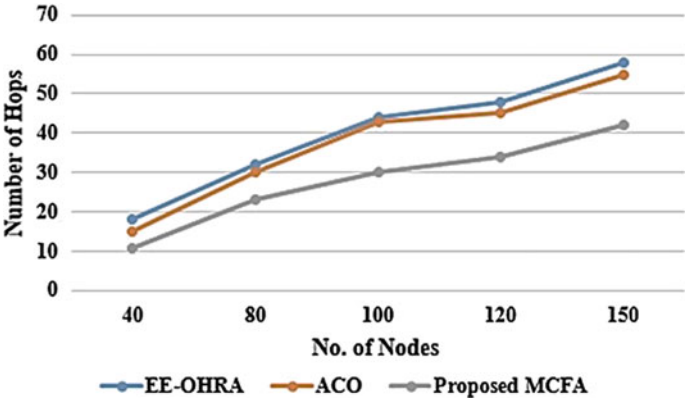


Fig. 4 Graphical representation of MCFA with existing models by considering path optimality

average jitter for the node 60. However, the proposed MCFA model achieved less average jitter, when compared with existing techniques. For instance, when the node is 40, EE-OHRA has 0.56 s, ACO has 0.58 s, and the proposed MCFA model has only 0.48 s. When the node reaches 80, the existing techniques achieved 0.74 and 0.70 s, where the MCFA achieved 0.62 s. This analysis shows that the MCFA provides better performance than existing techniques, even the impact of node size is high. Figure 4 depicts the analysis of MCFA with ACO and EE-OHRA in terms of path optimality.

From the analysis, the reduced number of hops is used only by the proposed MCFA model for processing the packets in the network than existing techniques. The reason is that the multipath routing of MCFA uses only two main processes such as reflection and visibility as a search strategy for finding the optimal solution of CH, where ACO uses the searching behavior of ants to find the optimal path, and once the path has deviated, it is difficult to find the solution. Here, the EE-OHRA method achieved 18, 44, and 58%, ACO achieved 15, 43, and 55%, and the proposed model achieved 11, 30, and 42% when the node is 40, 100, and 150. When the node size increases gradually, the number of hops for path optimality is also gradually increased by the algorithms. For instance, when the node is 120, the MCFA achieved 34%, and the same method achieved 30% of path optimality when the node is 100. Figure 5 gives the experimental results of the proposed model with existing techniques in terms of load balance.

The load balance implies the capacity of an algorithm for handling the mobility speed. For instance, the mobility speed of MCFA is 215 kmph for node 2, and the same method achieved 220 kmph for node 8. The existing ACO method achieved 125 kmph for node 2, and it has only 140 kmph for node 8. While comparing with all techniques, EE-OHRA achieved less mobility speed such as 195 and 198 kmph, ACO achieved 150 and 130 kmph, and MCFA achieved 214 and 218 kmph for nodes 10 and 12. Figure 6 provides the analysis of the proposed MCFA with ACO and EE-OHRA in terms of PDR.

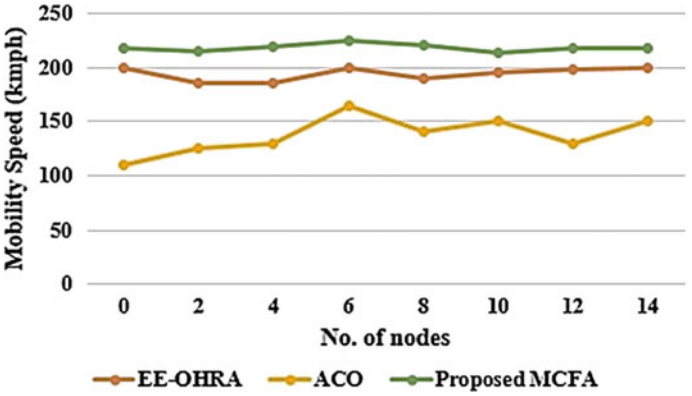


Fig. 5 Graphical illustration of the proposed model with existing techniques using load balancing

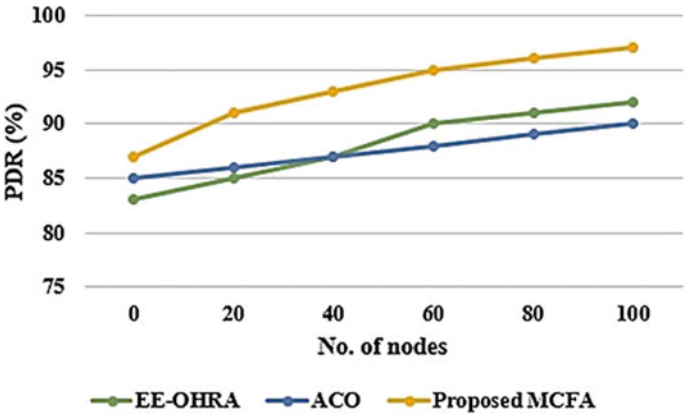


Fig. 6 Graphical illustration of the proposed model with existing techniques by considering PDR

To identify the PDR, the three parameters such as mobility, packet size, and network size are considered in the simulation part. When the packet sizes improve, the proposed model increased nearly 5–6% of PDR than existing techniques. When the node is 100, the EE-OHRA achieved 92%, ACO achieved 90%, and MCFA achieved 97%, where these existing techniques achieved 87% of PDR and proposed MCFA achieved 93% of PDR, when the node is 40. PDR is gradually increased for all techniques when the node size is increased. For instance, MCFA achieved 91% of PDR for node 20, and it achieved 96% of PDR for node 80. From these various experimental analyses, it is proven that the proposed MCFA achieved better performance than other techniques.

5 Conclusion

In this research work, multipath routing is developed by MCFA-based CH selection technique for improving the network lifetime and stability of the MANET. From the group of clusters, the CH is elected based on the optimal solution of MCFA. Here, the main feature of the proposed algorithm is that local and global searches are integrated and the local trap is solved by the chaos mapping function. According to the visibility and reflection process, an optimal solution is obtained for electing the CH and multipath routing is carried out. The experiments are conducted to verify the performance of MCFA with existing techniques called ACO and OHRA model in terms of major parameters such as ETED, PDR, mobility speed, average jitter, and path optimality. The results proved that the MCFA achieved 97% of PDR, 30% of path optimality, 0.68 s of average jitter, 10 s of ETED for node 100, where the ACO model achieved 90% of PDR, 43% of path optimality, 0.75 s of average jitter, 28 s of ETED for the same node 100. This analysis proves that the MCFA achieved better performance than existing models for electing the CH that provides better multipath routing and improved the network lifetime.

References

1. Liu J, Xu Y, Shen Y, Jiang X, Taleb T (2017) On performance modelling for MANETs under general limited buffer constraint. IEEE Trans Veh Technol 66(10):9483–9497
2. Chen J, Mao G, Li C, Liang W, Zhang D-G (2018) Capacity of cooperative vehicular networks with infrastructure support: multiuser case. IEEE Trans Veh Technol 67(2):1546–1560
3. Zhang D-G, Qiu J-N, Zhang T, Wu H (2019) New energy-efficient hierarchical clustering approach based on neighbor rotation for edge computing of IOT. Proc ICCCN, Valencia, Spain 8(1):291–295
4. Sushma T (2021) A review of the cluster based mobile adhoc network intrusion detection system. Turkish J Comp Math Educ (TURCOMAT) 12(2):2070–2076
5. Zhang DG, Zhou S, Chen J (2017) New dv-distance method based on path for wireless sensor network. Intell Autom Soft Comput 23(2):219–225
6. Zhang D-G, Niu H-L, Liu S, Ming X-C (2017) Novel positioning service computing method for WSN. Wirel Pers Commun 92(4):1747–1769
7. Saseekala M, Kamalakannan P (2018) Resource optimized kernel support vector for scalable and reliable multipath multicast routing in MANET. J Theor Appl Inf Technol 96(5):1225–1237
8. Liu X-H, Zhang D-G, Yan H-R, Cui Y-Y, Chen L (2019) A new algorithm of the best path selection based on machine learning. IEEE Access 7:126913–126928. https://doi.org/10.1109/ACCESS.2019.2939423
9. Bhalaji N (2021) Cluster formation using fuzzy logic in wireless sensor networks. IRO J Sustain Wirel Syst 3:31–39
10. Haoxiang W, Smys S (2020) Soft computing strategies for optimized route selection in wireless sensor network. J Soft Comput Paradigm (JSCP) 2(01):1–12
11. Liu S, Zhang D-G, Liu X-H, Zhang T, Gao J-X, Gong C-L, Cui Y-Y (2019) Dynamic analysis for the average shortest path length of mobile ad hoc networks under random failure scenarios. IEEE Access 7:21343–21358. https://doi.org/10.1109/ACCESS.2019.2896699
12. Taha A, Alsaqour R, Uddin M, Abdelhaq M, Saba T (2017) Energy efficient multipath routing protocol for mobile ad-hoc network using the fitness function. IEEE Access 5:10369–10381

13. Geng H, Shi X, Yin X (2017) Algebra and algorithms for multipath QoS routing in link state networks. J Commun Netw 19(2):189–200
14. Prakash V, Pandey S, Singh D (2021) Best cluster head selection and route optimization for cluster based sensor network using (M-pso) and Ga algorithms. https://doi.org/10.21203/rs.3.rs-173355/v1
15. Agarwal R, Motwani D. Survey of clustering algorithms for MANET. http://arxiv.org/abs/0912.2303
16. Khatoon N (2017) A node stability based multi-metric weighted clustering algorithm for mobile ad hoc networks. In: Proceeding of the second international conference on microelectronics, computing and communication systems (MCCS 2017), pp 63–77. Springer, Singapore (2017)
17. Muneeswari B, Manikandan MSK (2019) Energy efficient clustering and secure routing using reinforcement learning for three-dimensional mobile ad hoc networks. IET Commun 13(12):1828–1839
18. Veeraiah N, Khalaf OI, Prasad CV, Alotaibi Y, Alsufyani A, Alghamdi SA, Alsufyani N (2021) Trust aware secure energy efficient hybrid protocol for MANET. IEEE Access 9:120996–121005. https://doi.org/10.1109/ACCESS.2021.3108807
19. Prasad R, Shankar PS (2020) Efficient performance analysis of energy aware on demand routing protocol in mobile ad-hoc network. Eng Rep 2(3), Art. no. e12116
20. Kumar SV, Anuratha V (2020) Energy efficient routing for MANET using optimized hierarchical routing algorithm (Ee-Ohra). Int J Sci Technol Res 9(2):2157–2162
21. Wang S (2021) Multipath routing based on genetic algorithm in wireless sensor networks. Math Probl Eng 2021, Art. no. 4815711
22. Ran C, Yan S, Huang L, Zhang L (2021) An improved AODV routing security algorithm based on blockchain technology in ad hoc network. EURASIP J Wirel Commun Netw 2021(1):1–16
23. Sharma A, Tharani L (2021) Ant colony based node disjoint local repair in multipath routing in MANET network. Wirel Pers Commun:1–28
24. Rao DS, Rao PN (2016) A cluster based scalable and energy efficient multi path routing protocol for MANET. Int J Eng Sci 6(3):3093–3099
25. Venkatasubramanian S, Suhasini A, Vennila C (2021) An energy efficient clustering algorithm in mobile Adhoc network using ticket id based clustering manager. Int J Comp Sci Netw Secur 21(7):341–349
26. Venkatasubramanian S, Suhasini A, Vennila C (2021) Efficient multipath zone-based routing in MANET using (TID-ZMGR) ticked-ID based zone manager. Int J Comp Netw Appl 8(4):435–443
27. Eesa AS, Brifcani AMA, Orman Z (2013) Cuttlefish algorithm-a novel bio-inspired optimization algorithm. Int J Sci Eng Res 4(9):1978–1986
28. Khan MA (2020) An IoT framework for heart disease prediction based on MDCNN classifier. IEEE Access 8:34717–34727

Opinion Mining of Movie Reviews Using Hybrid Deep Learning Technique

Yash Patel, Jaimeel Shah, and Shital Pathar

Abstract Due to Internet, vast amount of data is generated day by day; from those data to find useful insights, there is need to identify and extract the subjective information. Today's trends show that people are buying any products or watching any movie on Web sites, and they write the feedbacks related to that product or movie, which will be helpful to business in terms of profit. For that, the need is analysis of written reviews which will be done by sentiment analysis. It is a method which is used to gauge opinions of individuals or groups of persons related to their products or movies. This method will extract the meaningful insights from the written reviews in the form of positive, negative, or neutral. Analysis of sentiment is also known as opinion mining. In this paper, hybrid deep learning model (CNN + LSTM) is applied on IMDB movie review dataset and performs a comparison with CNN model.

Keywords Hybrid model · Convolutional neural network (CNN) · LSTM · CNN-LSTM · Glove word embedding · word2vec word embedding · Opinion mining

1 Introduction

Due to the corona pandemic, people are not going to theaters. They are watching movies from online digital platforms like Amazon Prime, Netflix, Hotstar, and some other Web sites. These online sites provide the platform on which customers are able to write and share their feedbacks about movie they have watched. Analysis of their positive opinions tells us particular movie's box office collection success.

Y. Patel (✉) · J. Shah · S. Pathar
Parul Institute of Engineering and Technology, Vadodara, India
e-mail: yashpatel6560@gmail.com

J. Shah
e-mail: Jaimeel.shah@paruluniversity.ac.in

S. Pathar
e-mail: shital.pathar2882@paruluniversity.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_2

Analysis of negative opinions tells us the strong and weak points of the movie. So, the analysis of customer opinions helps movie business to connect with their customers. That helps them to meet their customer expectations after doing the analysis. After the analysis, their business will grow. So analysis is important for business growth and benefits. Since movies play an important role in the entertainment industry, not only do they entertain people, but they also rely on ratings and profits based on reviews from all over the world [1]. Producers take better decisions from the online reviews, like how to satisfy their viewers and make more sales taking an advantage. It also helps viewers to quickly determine which movie to watch. The analysis of opinions is done using the sentiment analysis [2]. It is a technique which extracts and understands the emotions in form of positive or negative present in the text [3]. Extracting sentiments from the large dataset is a challenging task. For the processing of these large datasets, various deep learning algorithms are required. The deep learning algorithms are different from the traditional machine learning algorithms. Particularly, deep learning algorithms are not dependent on the different feature extraction techniques. They extract the features during the training time.

2 Literature Review

Due to corona pandemic, most of the businesses are running on online platforms. Expressing opinions and posting reviews about the product purchased or movies that are seen have become trending nowadays. Sentiment analysis performed on movie review dataset using deep learning model CNN also used machine learning algorithms like Naïve Bayes, SVM, logistic regression, KNN, and ensemble methods. As a result, they found that the deep learning model CNN gives better performance than the other models [1]. Sentiment analysis is performed on mobile phone reviews dataset. Reviews are categorized using different machine learning techniques such as Naïve Bayes, SVM, and decision tree. Model performances of classifiers are compared and get result as SVM classifier gave better performance than the other machine learning methods [2]. Researchers did the sentiment analysis on the product review dataset, used different machine learning methods: Naïve Bayes and support vector machine (SVM), and found that the SVM technique gives better performance [3]. Deep learning technique LSTM is applied on custom review dataset which gives efficient result [4]. Sentiment classification of movie reviews is done using different machine learning classifiers such as decision tree, SVM, and maximum entropy. We compared the model performance and found that maximum entropy techniques give better accuracy [5]. Deep learning technique is applied for sentiment classification of smartphone review dataset. We used CNN as a deep learning method and compared their model performance with machine learning methods such as Naïve Bayes and support vector machine (SVM). A deep learning method is better performed than machine learning techniques [6]. Researchers performed sentiment analysis on movie reviews. Machine learning classifiers are applied on dataset. Naïve Bayes, KNN, and random forest classifiers are applied and achieved 81.4% accuracy [7]. Opinion

mining is performed on movie reviews, and machine learning and deep learning methods are applied on the review dataset. For machine learning, SVM and Naïve Bayes techniques are used, and for deep learning, neural network technique is used. A deep learning technique gave better performance than the other methods [8]. Deep learning methods are applied on hotel's review dataset for classification of sentiments. We applied convolutional neural networks (CNN) on hotel review data and classified into sentiments. Comparing the model performance of the CNN method with other methods and achieving the result as convolutional neural networks give better performance than the other methods [9]. Sentiment classification of restaurant reviews is done. We applied various machine learning models on the dataset and got highest accuracy by SVM classifier [10]. Sentiment analysis is applied on persian movie reviews. We performed deep learning techniques CNN and LSTM on review dataset and compared their model performance with machine learning techniques like SVM, MLP, and logistic regression. The result found that LSTM model performance was better than the other techniques [11]. Machine learning classifiers such as decision tree, Naïve Bayes, and KNN are applied for sentiment analysis. We compared the performance of the models. Found result as Naïve Bayes model performed the better than the other models [12]. We performed sentiment analysis on Twitter dataset using machine learning classifiers such as decision tree and Naïve Bayes and found that decision tree machine learning classifiers gave better results [13]. We applied machine learning models Naïve Bayes and SVM on reviews dataset for the prediction of sentiments, compared their performances, and found that Naïve Bayes performance is better than the SVM. Different machine learning methods such as Naïve Bayes, maximum entropy, and SVM are applied on review dataset, and the performance using different evaluation parameters is measured [14]. We applied various machine learning and deep learning techniques on amazon review dataset for predicting their outcomes in form of positive and negative and achieved good accuracy [15].

3 Proposed Work

The proposed diagram shows that the different steps are involved to predict the sentiment as positive or negative from the movie reviews (Fig. 1).

To perform opinion mining or sentiment analysis on reviews, the need is to collect the data from standard sources.

3.1 Collection of Data

IMDB movie review dataset is used in this work. This dataset is available on Kaggle Web site. This dataset contains 50 K movie reviews.

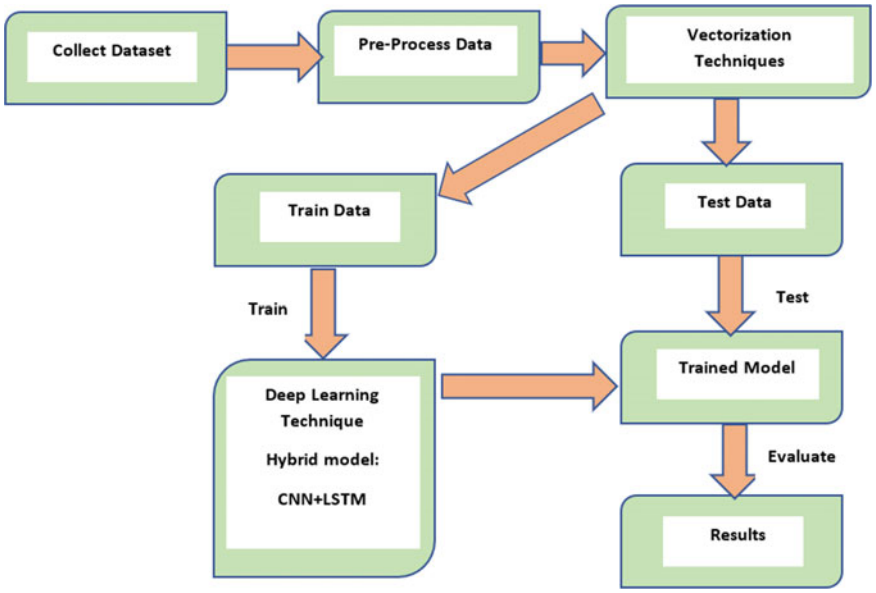


Fig. 1 Proposed model

3.2 Pre-processing of Data

Dataset is used in the form of CSV format. Pre-processing is the most important step, which helps for determining sentiments from the data. This step performs a major role in opinion mining tasks. It will remove the unnecessary and incomplete data and convert it into clean data. Different pre-processing steps involve tokenization, stop word removal, stemming, lemmatization, and punctuation removal.

3.3 Text to Numeric Conversion

Process of converting text information into numeric is called as vectorization process, which will be understood by the deep learning algorithms. In this work for the vectorization process, word2vec word embedding is used. In word embedding, text data are converted into vector format. I have used pre-trained word2vec word embedding to convert each word into a 300-dimensional vector size. In word2vec embedding technique, it represents the different words, which are used in the same context and have very close vector representation which will be useful to the model to understand meaning of the word correctly and predict the accurate outcomes. In this work, pre-trained glove word embedding technique is also used to check the model performance of CNN-LSTM with word2vec word embedding with hybrid model CNN-LSTM.

3.4 Segregation of Training and Test Data

The most important step in machine learning and deep learning is to feed some amount of data into the model for the training purpose so that model can understand the meaning of that data, which will help the model for predicting the unseen data. From that, we will check the level of understanding of model how well it understands the data. So, to predict the unseen data by the model, necessary step is to first train the model, and for that, separation of data is required. In this work, we split the data into 80:20 ratio. We have taken 80% data for training the model and 20% for testing the model understanding.

3.5 Classification

After dividing the dataset into two parts, I feed the training data into the deep learning hybrid approach CNN-LSTM model. Keras library's sequential layer is used for the model. The first layer of the hybrid model is the embedding layer. The embedding layer will convert the data into a dense vector size of 300, and this review matrix will be fed into the convolutional layer. Each filter in the CNN will extract specific features. It will pass to the max pooling layer. This layer selects max value from the parameters and reduces the dimensionality which reduces the complexity of computation. These spatial features are learned sequentially by an LSTM layer of 150 units. Outputs from the LSTM model are passed in to the dense layer. As there are two labels that the model will predict for that purpose. Sigmoid activation function is used at the last layer. Model is trained using 50 epochs and 256 batch size. Now model is ready to predict the unseen data. In this work, 7080 training samples and 3450 testing samples are taken (Fig. 2).

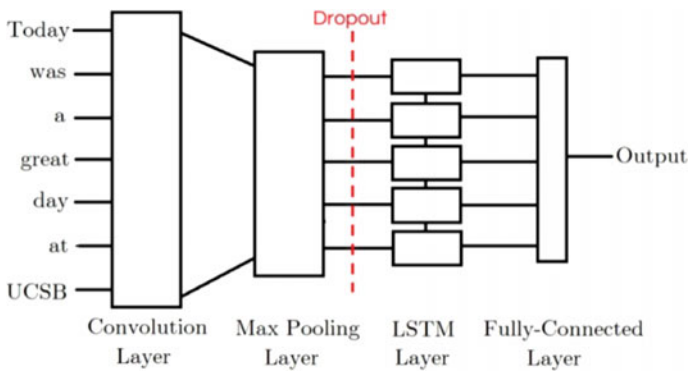


Fig. 2 Hybrid model architecture

3.5.1 Pseudocode for the CNN_LSTM Model

Input: Model, train, test.

// Model = keras sequential Model, train x = training data, test y = testing data.

AF = RELU (), **OAF** = sigmoid (),

Hyperparameters: filters, kernels, pool size, max_ sequence length, batch size, epochs.

seq_in = Input(shape = (MSL)) // MSL = max_sequence_length of sentence.

embedded_seq = embedding_layer(seq_in).

X = CONV1D (NCF, WINSIZE, AF) (embedded_seq) //**NCF** = no of convolution filter, **WINSIZE** = Kernel size.

X = MP1D(PS)(X) // **MP1D** = Max pooling 1D layer, **PS** = pooling size,

X = LSTM (NOU, Dropout = 0.1, Recurrent_Dropout = 0.2) (X) //**NOU** = no of Units.

Pred = FC (NC, OAF) (X) // **NC** = no of class to predict, **FC** = fully connected layer.

Model = Model (seq_in, pred).

Opt = Adam, Loss = binary cross entropy.

model. Compile (Opt, Loss, AC) //**opt** = optimizer, **AC** = accuracy Measures.

Model.fit (train x, train y, epochs, batch size, test x, test y).

4 Result

The model's performance is measured using different evaluation parameters which are as follows: accuracy, precision, recall, etc.

Table 1 shows that the glove word embedding with CNN-LSTM hybrid model achieved better accuracy than the word2vec word embedding with CNN-LSTM model. In this work, the reason behind to use hybrid model is CNN model can extract the important feature without seeing its position from the sentence; it means CNN cannot remember the position of the capture feature from the sentence so that model cannot understand the meaning of the whole sentence clearly. If those extracted features are passed sequence-wise to the LSTM model, then this LSTM model can remember the previous data of the input sentence. It means LSTM model will remember the sequence of extracted features, and from that, it will find the contextual meaning of the word from the surrounding words of the sentence. So

Table 1 Model performance

No.	Model	Accuracy	Precision	Recall
1	Word2VEC + CNN_LSTM	0.9071	0.89	1.00
2	GLOVE + CNN_LSTM	0.9176	0.95	0.94
3	Word2VEC + CNN	0.8050	0.75	0.83

model will understand the meaning of whole sentence precisely and also gives better performance.

Figure 3 shows the receiver operating characteristic (ROC) curve for glove word embedding with CNN-LSTM hybrid classifier.

The hybrid word2vec word embedding with CNN-LSTM model and glove word embedding with CNN-LSTM model performance comparison results are represented graphically which is shown in Fig. 4.

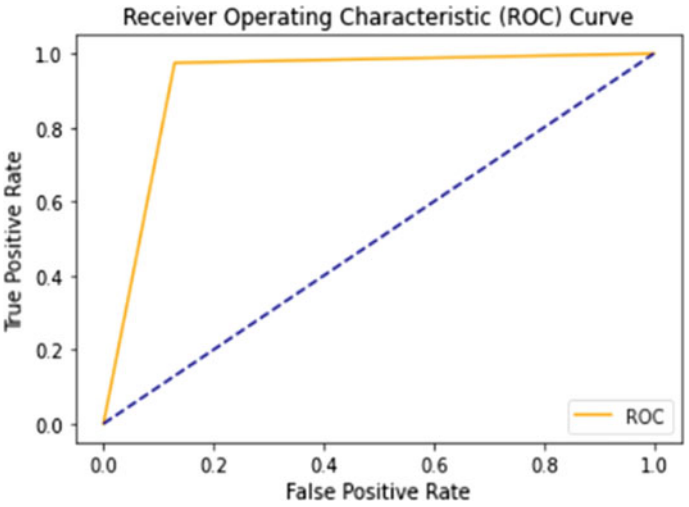


Fig. 3 ROC curve

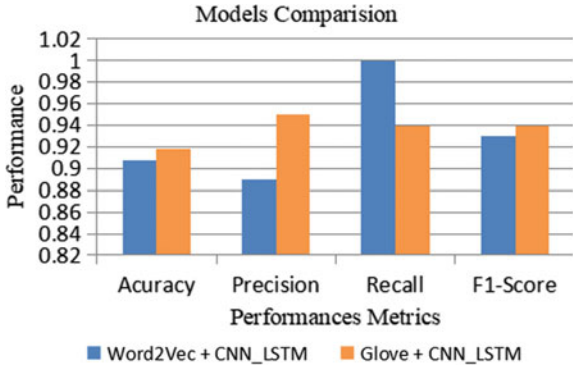


Fig. 4 Performance measures

5 Conclusion

Opinion mining of movie reviews is done using hybrid deep learning approach. For that, word2vec word embedding with CNN-LSTM hybrid model is applied on IMDB movie review dataset. And its performance is compared with glove embedding with CNN-LSTM hybrid model. In this work, glove embedding with CNN-LSTM hybrid model performs better than the word2vec embedding with CNN-LSTM. So, by using different types of vectorization technique, it will affect the performance of the model. With an increase in size of data, model performance also increases. Batch size, No of epoch, optimizer, and activation functions affect the performance of the deep learning model.

References

1. Dholpuria T, Rana YK, Agrawal C (2018) A sentiment analysis approach through deep learning for a movie review. IEEE
2. Panthati J, Bhaskar J, Ranga TK, Challa MR (2018) Sentiment analysis of product reviews using deep learning. IEEE
3. Ribeiro de Souza JG, de Paiva Oliveira A (2018) Deep learning approach for sentiment analysis applied to hotel's reviews. Springer International Publishing
4. Tholusuri A, Anumala M, Malapolu B, Jaya Lakshmi G (2019) Sentiment analysis using LSTM. IJEAT
5. Vanaja S, Belwal M (2018) Aspect-level sentiment analysis on e-commerce data. IEEE
6. Malini R, Sunitha MR (2020) Opinion mining on movie reviews. IEEE
7. Rahman A, Hossen MS (2020) Sentiment analysis on movie review data using machine learning approach. IEEE
8. Baid P, Gupta A, Chaplot N (2017) Sentiment analysis of movie reviews using machine learning techniques. IJCE
9. Ouyang X, Zhou P, Li CH (2015) Sentiment analysis using convolutional neural network. IEEE
10. Chawla S, Dubey G, Rana A (2017) Product opinion mining using sentiment analysis on smartphone reviews. IEEE
11. Dashtipour K, Gogate M, Larijani H, Hussain A (2021) Sentiment analysis of Persian movie reviews using deep learning. MDPI
12. Krishna A, Akhilesh V, Aich A, Hegde C (2019) Sentiment analysis of restaurant reviews using machine learning techniques. Springer
13. Bayhaqy A, Sfenrianto S, Nainggolan K, Kaburuan ER. Sentiment analysis about e-commerce from Tweets using Decision Tree, K-Nearest Neighbor, and Naïve Bayes. IEEE
14. Shah A (2021) Sentiment analysis of product reviews using supervised learning, RT&A, Special Issue № 1 (60), Volume 16
15. Sisodia DS, Reddy NR (2017) Sentiment analysis of prospective buyers of mega online sale using Tweets. IEEE

Hybrid Approach to Predict the Death Rate of COVID-19 Patients

P. Keerthika, P. Suresh, R. Manjula Devi, S. Vaishnavi, C. Shanmathi, and V. Surendar

Abstract From closedown of December 2019, coronavirus has directly exhibited a lofty rate of transmission, coercing the World Health Organization to contend in the month of March 2020 that this unbeknownst coronavirus can be depicted as a pandemic. COVID-19 epidemic has guided to an operatic misplacement of deathly life over the public and presents an unbeknownst complaint to public fitness. It also affects the food systems of the person and the world of work. Once the person is infected by COVID, the metabolic exertion of vulnerable cells in his or her body is enhanced, similar as the one driven by COVID-19. The country's dietary habits are analyzed to predict the particular person's death rate. By using KNN algorithm, the performance metrics such as accuracy, precision, recall, and F1 score are evaluated for the country's dietary habits. In this research, both clustering and classification are combined to increase the accuracy of the prediction of death rate of the person. K-means is used for the clustering of the countries, and KNN is used for classifying the countries. The 170 countries are clustered based on the country's dietary habits, and other disease affected rate using K-means clustering algorithm. Countries are

P. Keerthika (✉) · P. Suresh
Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: keerthikame@gmail.com

P. Suresh
e-mail: sureshme@gmail.com

R. M. Devi · S. Vaishnavi · C. Shanmathi · V. Surendar
Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, Erode, Tamil Nadu, India
e-mail: rmanjuladevi.gem@gmail.com

S. Vaishnavi
e-mail: vaishnavis.18cse@kongu.edu

C. Shanmathi
e-mail: shanmathic.18cse@kongu.edu

V. Surendar
e-mail: surendarv.18cse@kongu.edu

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_3

clustered into high and normal death rate countries based on the country's dietary habits and another cluster into high and normal death rate based on the other disease affected rate rather than COVID-19. Using the country's dietary habits and other disease affected clusters, the death rate of the person is predicted. After clustering the data based on the country's dietary habits and other disease affected rate, the KNN algorithm is used to classify and identify the person's death rate. Using clustering and classification algorithms in a combined way, an accuracy of 79% is achieved.

Keywords Machine learning · COVID-19 · Classification · Country's dietary habits · Other disease affected · Survival rate

1 Introduction

The severe acute respiratory motive coronavirus 2 contagion causes coronavirus illness, a transmittable disease. Kwekha-Rashid et al. [1] describes the adulthood of those infected with the contagion will hold genial to moderate respiratory symptoms and will get back without the lack of medical concentration. Some, on a different angle, will pick up critically crummy and deliver medical assistance. Sedate illness is additionally probable to walk out the elderly and those with bolstering medical conditions analogous as cardiovascular indictment, diabetes, chronic respiratory indictment, or cancer. De Felice and Polimeni [2] says that COVID-19 may frame anyone ill and bring them to pick up truly ill or dead-end at any age. Stay at least 1 cadence downward from people, burn out a fluently decent vizard, and splash your hands or apply an alcohol-predicated annoyance periodically to pinch-hit yourself and others from infection. Salam et al. [3] says that when it is your wander, get vaccinated and supervise foremost counsel. Larger respiratory driblets to lower aerosols are among the particles. Rustam et al. [4] refers that however, it is critical to borrow respiratory form, similar to coughing into a flexed elbow, if you are sick.

2 Related Works

2.1 Effects of COVID-19 on Patients with Other Disease

The nimbus contagion complaint 2019 (COVID-19) epidemic has accelerated the dangers for cases with different conditions and has commanded an expressive collision on conventional healthcare capitals. Patients with lung disease, respiratory challenges or cancer who likewise hold COVID-19 held a significantly lesser mortality rate. Older era, manly coitus, a record of smoking, the presence of multitudinous functional comorbidities, interpretation deal, and developing malice are presently verified threat deputies for mortality in people with COVID-19. Pinter et al. [5] describes that rotundity and some attendant ails hold likewise existed recognized as probable clinical threat attorneys for mortality in people with COVID-19. Furthermore, they are at a higher risk of contracting opportunistic infections and having

serious complications as a result of these illnesses. Postponing or changing the treatment schedule may result in a negative consequence, with a well-documented impact on clinical outcomes. Furthermore, patients may be at a higher risk of infection as a result of their frequent hospital visits. Cases must peregrinate to establishments for treatment, despite considerable trip challenges, and must-have line to be filtered for COVID-19 infection gesticulations. Fu et al. [6] says that in a COVID-19-prioritized fitness network, we must equalize the COVID-19 difficulty and the precautionary conduct we are straining to restrict this epidemic with the irrefutable peril of enlarged morbidity and fate for cases. Punn et al. [7] says that patients, as well as healthcare practitioners, are unsure and concerned about the current epidemic and the amount of available information, which includes disinformation and rumors. Yadav [8] refers to whether individuals with cirrhosis and COVID-19 are more likely to decompensate or suffer acute-on-chronic liver failure (ACLF).

2.2 Identification of COVID-19 Applying X-Rays and CT Scans

The use of various machine learning methods facilitates the use of X-ray images and CT scans to diagnose COVID-19 more effectively. The accuracy rate of predicting COVID-19 using X-rays and CT scans range from 78% to more than 99%. Rasheed et al. [9] says that this method not only helps in identifying COVID-19 but also helps in predicting the severity of the disease and chance of early mortality.

2.3 Diagnosis Based on Blood and Urine Tests

COVID-19 is diagnosed using epidemiological factors, clinical characteristics, imaging results, and nucleic acid screening, among other things. These technologies took a long time to produce the diagnosis results and were prone to errors. Soares [10] says that for a patient with COVID-19 infection, many types of clinical data were obtained and manually merged by doctors to make diagnostic judgments. Using clinical information and blood/urine test data, this study looked at detecting critically unwell COVID-19 individuals from those with minor symptoms.

2.4 Vaccination

Enough vaccine boluses have nowadays been contributed to completely vaccinate 44.6% of the world's population, but the admeasurement has existed irregularly. The best vaccinations are very successful at reducing illness and death, stopping a

pandemic requires a concerted approach. Vaccinating 70–85% of the US population, according to infectious-disease experts, would allow for a return to normalcy. That is a frightening degree of immunization on a worldwide basis. The current global immunization rate averages 31,953,362 doses per day. It will take another 5 months to cover 75% of the population at the current rate. This progress is under jeopardy. Ong et al. [11] says that new strains have resurrected epidemics, headed by the highly transmissible delta variant. Vaccines and viruses are now in a life-or-death battle. Abdulkareem et al. [12] refers that unvaccinated person are now more vulnerable than ever before. According to the most recent statistics, even among people who have been vaccinated, the delta variation can cause minor cases, and those who become ill can spread the disease to others.

2.5 New Delegates Companies with COVID-19 Transmission

Profitable differences, among numerous proxies associated to COVID-19 threat, accelerated the accident of COVID-19 transmission. Sear et al. [13] refers that COVID-19 losses were negatively associated with per capita sanitarium bunks. Blood varieties B and AB were displayed to be defensive against COVID-19, but blood variety A was planted to be a threat agent. Samui et al. [14] refers that COVID-19 transmission is hampered additionally by lofty temperatures than by equatorial temperatures.

2.6 Reduction in Mobility and COVID-19 Transmission

Commonwealths have tried to degrade severe acute respiratory motive coronavirus 2 transmission in reaction to the COVID-19 epidemic by confining population shifting through gregarious distancing ways, away lowering the composition of interactions. Mele and Magazzino [15] refers to minimizing neighborhood-position transmission of severe acute respiratory motive coronavirus 2 and ammunition endangered populations, gregarious spacing, case insulation, and covering possessed breaths routinely adopted.

2.7 Environmental Perspective on COVID-19

This outbreak has also impacted people's lifestyles; it has resulted in widespread job losses and put millions of people's lives in jeopardy as companies shut down to stop the virus from spreading, inferring that lights have been canceled and transportation networks have been shut down all around the world [16]. However, the COVID-19

pandemic's containment resulted in improved air quality in many cities throughout the world, as well as a reduction in water pollution in some areas.

3 Proposed Work

In this research, a hybrid of classification and clustering is used to predict the possibility of survival of the particular person. The country's dietary habits and other disease affected person data are analyzed to predict the particular person's survival rate. The country's dietary habits and the other disease affected person dataset are clustered into four types of clusters by using K-means. PCA algorithm is used to normalize the data. After clustering, the data in the dataset is given as the training data to the KNN algorithm. After training the model, a particular person's details are given as input, to predict the particular patient's survival rate based on the country the person belongs to, the dietary habits cluster and also with the help of other disease affected clusters. The country's dietary habits and other disease affected person data are analyzed to predict the particular person's survival rate. By using KNN algorithm on the country's dietary habits and the other disease affected person dataset, the performance metrics such as accuracy, F1 score, recall, and precision are evaluated.

3.1 PCA

PCA is used to reduce the dataset's dimensionality, minimizing the information loss and increasing the interpretability. PCA is the primary proportion reduction fashion, which facilitates us to decrease the size of the dataset and count undesirable connections. PCA is hypercritical while the dataset has millions of dimensions, we can use it to extract functions having redundant data. As we cannot fantasize redundant than three dimensions, we can use PCA to lessen the size to visualize the data. Dimensions of the dataset are to be reduced when there are many characteristics in the dataset, which makes distinguishing between relevant and redundant data difficult. In Fig. 1, the original data is rounded in between 0 to 1.

The values in food supply in Kcal dataset, food supply in Kg dataset, fat supply dataset, protein dataset, and other disease affected dataset are dimensionally reduced using PCA. It does so via means of compressing the characteristic area via means of figuring out a subspace that captures maximum of the facts with inside the entire characteristic matrix. By this way, it reduces the dimensionality of the dataset.

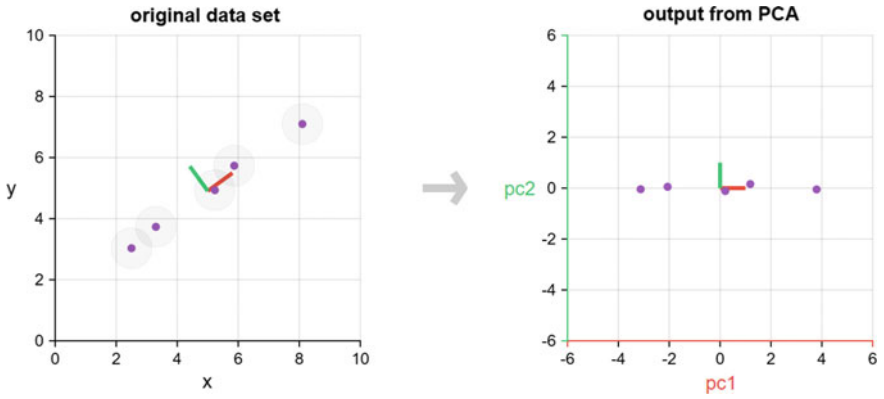


Fig. 1 Principal component analysis

3.2 K-Means

The K-means is used to find groups which have not been explicitly labeled in the data. It is used to find similarities in the data.

- Intake for K-means clustering: Dataset K number of desired clusters.
- Affair for K-means clustering: K set of clusters.

In Fig. 2, each centroid point is clustered into *K* clusters. Initialize the composition of cluster *k* and gather the original centroid aimlessly. Also, the squared Euclidean length will be computed from each valuation to each cluster is reckoned, and each thing is charged to the closest cluster. After that, for each cluster, the new centroid

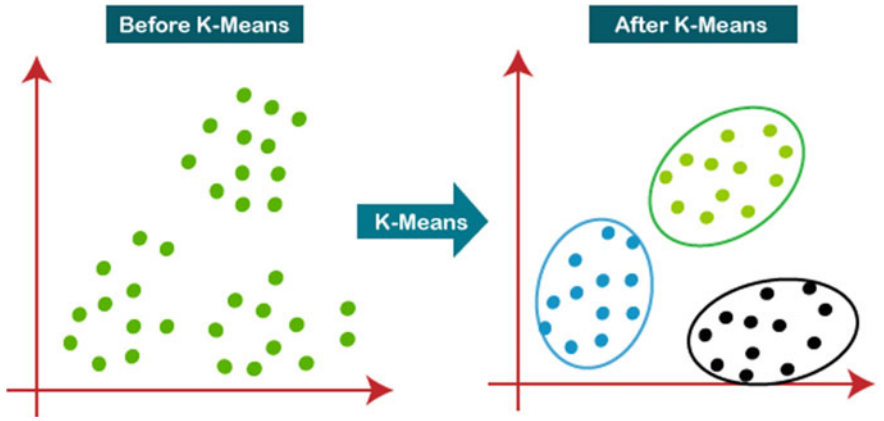


Fig. 2 K-means

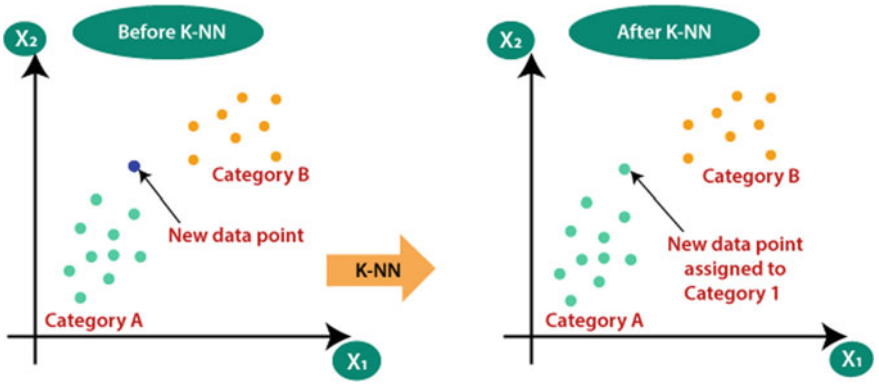


Fig. 3 KNN

is reckoned and each value is now replaced by the separate cluster centroid. K-means algorithm is easily adaptable to new dataset while training and testing data. Compared to other clustering algorithms, k-means gives fast and efficient results. Since the dataset supplied is generalized, the k-means algorithm produces the best result. So that k-means algorithm is chosen to cluster the dataset.

3.3 KNN

KNN algorithm finds the parallelism between the latest data and formerly accessible data or case and adds the latest data into the order which is additionally analogous to the data. KNN stores already available categories and classifies the new data or case based on the similarity between data. So whenever new data comes, it classifies easily based on a well-suited category. In Fig. 3, there are two categories, category A and category B. A new data point arrived. The new data point belongs to category A or category B is decided by KNN algorithm based on Euclidean distance. Actually, KNN has quick calculation time and also it has high accuracy. KNN algorithm is easily adaptable to new data. So that KNN algorithm is chosen to classify the dataset. In this research, KNN is used to classify the data based on the similarity.

3.4 Dataset

In this exploration, salutary fashions of the country and disparate complaint affected dataset are employed to prognosticate the person's survival grade. It combines data of dissimilar classes of food and COVID-19 cases and deaths each around the people. Each dataset contains data about 170 county's food practices. The dataset includes

Table 1 Dataset description

Dataset	Rows (countries)	Columns (food habits)
Food supply in Kcal	170	32
Food supply in Kg	170	32
Protein dataset	170	32
Fat dataset	170	32
Other disease dataset	170	5

- % of fat consumed
- % of food supply (in kilogram) consumed
- % of energy (in kilocalories) consumed
- % of protein consumed.

The fat consumed dataset, food supply in kg dataset, food supply in kcal dataset, and protein consumed dataset contain animal products, aquatic products, fish, seafood, fruits, vegetables, meat, oil crops, pulses, cereals, species, and also COVID-19 active cases, death cases, recovered cases, and population. The dataset also contains the details of the person affected with any other disease other than COVID-19. It includes diseases like diabetes, chronic kidney disease (CKD), chronic obstructive pulmonary disease (COPD), obesity, and undernourishment. And also, it contains the active cases, confirmed cases, death, and recovered cases. Countries are clustered into high death rate and normal death rate countries based on other disease affected person dataset. Table 1 describes the number of rows and columns in each and every dataset.

3.5 Hybrid Approach

In this research, both clustering and classification are used to increase the accuracy. For clustering, the K-means algorithm is used. PCA is used for normalizing the dataset. For classification, KNN algorithm is used. PCA is the important method for reducing the dimensionality of the dataset. The K-means is used to find groups which have not been explicitly labeled in the data. It is used to find similarities in the country's dietary dataset and other disease dataset. By using both k-means and KNN, the survival rate of the particular person was predicted. In this hybrid approach, clustering and classification algorithms are combined. Only by classifying the dataset, the accurate prediction cannot be achieved. Whereas by implementing clustering and classification algorithms in a combined way, the accurate prediction is achieved. Compared with performance metrics achieved only by classifying the dataset, the hybrid approach's performance metrics like accuracy, F1 score, precision, and recall were increased.

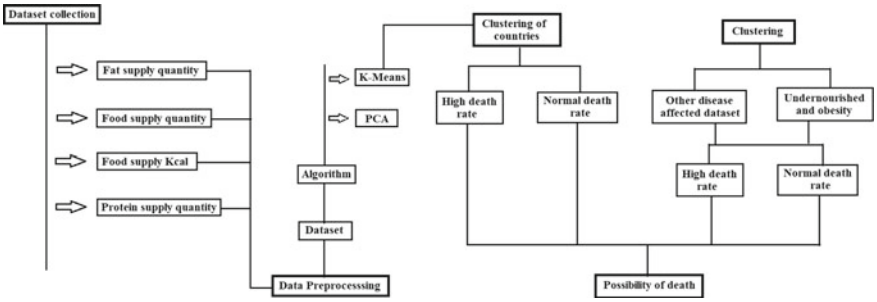


Fig. 4 Workflow

3.6 Workflow

First of all, all the required dataset is collected for clustering and classification. After collecting all the necessary data, the data in the dataset is preprocessed. The missing data in the dataset is replaced with −1. There is no categorical data in the dataset. For feature scaling, the PCA algorithm is used. The dietary habit dataset and other disease dataset are taken for training the model. By using K-means clustering, the country's dietary habits and other disease affected person data are clustered into four clusters, namely high death rate countries and normal death rate countries. By using KNN algorithm, the dataset is classified. After classification, a particular person's details are given as input, to predict the particular patient's survival rate based on the country the person belongs to, the dietary habits dataset and also with the help of other disease affected dataset.

Figure 4 describes the workflow of this research work. In this research, the country's dietary habits and other disease affected person data are analyzed to predict the particular person's survival rate. The country's dietary habits and the other disease affected cluster are classified by using KNN algorithm.

4 Result

4.1 Performance Metrics

Accuracy, Precision

Delicacy (Accuracy) is a metric for assessing bracket miniatures. Informally, delicacy is the bit of prognostications the miniature was accurate. Formally, delicacy is the number of accurate prognostications per total number of prognostications. Precision is a metric that quantifies the number of correct positive prognostications made. Equation (2) describes the distinction between the caliber of true positive and the extension of true positive and false positive.

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{2}$$

Recall, F1 Score

Recall is a metric that quantifies the number of correct positive prognostications made out of all positive prognostications that could have been made. Equation (3) describes the anamnesis caliber of true positive and the extension of true positive and false negative. The F1 score is a measure of a model's delicacy on a dataset. Equation (4) describes the F1 score as twice the rate of addition of perfection and recall to the addition of perfection and recall.

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{3}$$

$$\text{F1 Score} = 2 * \left(\frac{(\text{precision} * \text{recall})}{(\text{precision} + \text{recall})} \right) \tag{4}$$

4.2 Performance Evaluation

Figure 5 describes the accuracy and precision for KNN classification and the hybrid approach. In the *y*-axis, the percentage is taken, and in the *x*-axis, accuracy and precision for KNN classification and hybrid approach were taken. Figure 6 describes

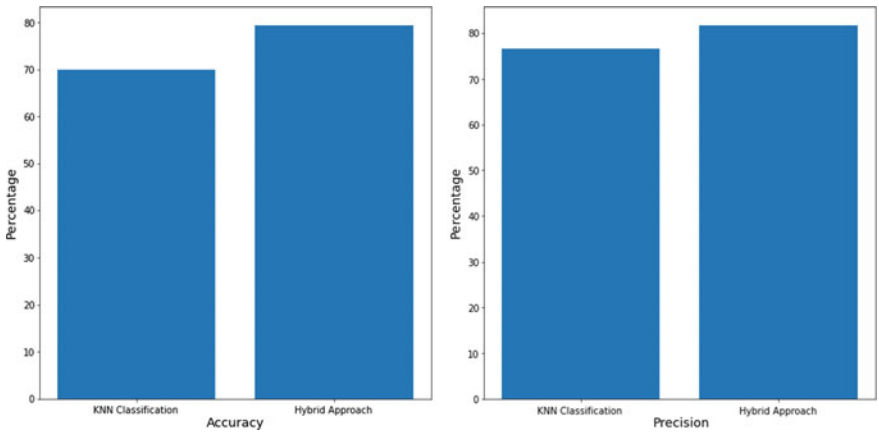


Fig. 5 Accuracy and precision

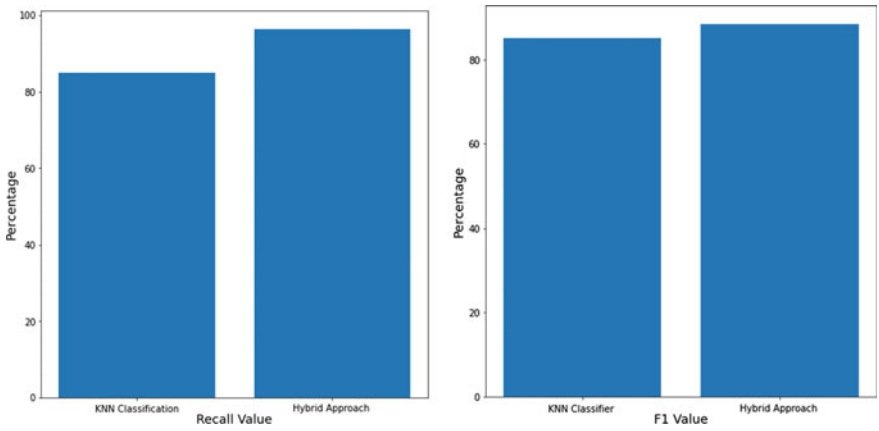


Fig. 6 Recall value and F1 score

the recall and F1 score value for KNN classification and hybrid approach. In the *y*-axis, the percentage is taken, and in the *x*-axis, the recall and F1 score value for KNN classification and hybrid approach were taken. From the below graphs, it is clear that the proposed system performs more efficiently than the existing system. The accuracy of the proposed system is 79%.

5 Conclusion and Future Work

The goal of this project is to predict the person's possibility of survival based on the country they belong to and if the person is affected by any other diseases. Punn et al. [7] refers that COVID-19 is being spread rapidly and many technologies have been identified for prediction of survival rate. KNN is the most commonly used classification algorithm. Though many inventions have been made, using KNN algorithms is the best and fastest way to predict the survival rate of people. It stores all the data that is fed through clustering of the data based on country dietary habits and other diseases and classifies the new data, based on similarity in the test data and training data. The purpose of this project is to predict the person's possibility of survival [17]. The future work of the project is to develop a Web-based application which will be useful for the hospital management system and make it free of cost and have planned to develop two additional modules, assessment using RNN and IoT. The analysis using RNN module will act as a voice-enabled chat bot and ask some questions to the users. It is intelligent to change the questions from the user's mental state and their previous answers [18]. The final module is analysis using IoT which makes use of pulse rate monitors and EEG sensors to detect depression at a very accurate level. This will be very effective and useful for the doctors and also the public.

References

1. Kwekha-Rashid AS, Abduljabbar HN, Alhayani B (2021) Coronavirus disease (COVID-19) cases analysis using machine-learning applications. Appl Nanosci. https://doi.org/10.1007/s13204-021-01868-7
2. De Felice F, Polimeni A (2020) Coronavirus disease (COVID-19): a machine learning bibliometric analysis. In vivo 34: 1613–1617
3. Salam MA, Taha S, Ramadan M (2021) COVID-19 detection using federated machine learning. PLoS ONE. https://doi.org/10.1371/journal.pone.0252573
4. Rustam F, Reshi AA, Mehmood A, Ullah S, On B-Y, Aslam W, Choi GS (2020) COVID-19 future forecasting using supervised machine learning models. IEEE Access 8:101489–101499. https://doi.org/10.1109/ACCESS.2020.2997311
5. Pinter G, Felde I, Mosavi A, Ghamisi P, Gloaguen R (2020) COVID-19 pandemic prediction for hungary; a hybrid machine learning approach. MDPI (6):890. https://doi.org/10.3390/math806089
6. Fu L, Li Y, Cheng A, Pang PP, Shu Z (2020) A novel machine learning-derived radiomic signature of the whole lung differentiates stable from progressive COVID-19 infection. J Thoracic Imag 35(6):361–368. https://doi.org/10.1097/RTI.0000000000000544
7. Punn NS, Sonbhadra SK, Agarwal S (2020) COVID-19 epidemic analysis using machine learning and deep learning algorithms. medRxiv. https://doi.org/10.1101/2020.04.08.20057679
8. Yadav RS (2020) Data analysis of COVID-2019 epidemic using machine learning methods: a case study of India. Int J Inf Technol 12(4):1321–1330. https://doi.org/10.1007/s41870-020-00484-y
9. Rasheed J, Hameed AA, Djeddi C, Jamil A, Al-Turjman F (2021) A machine learning-based framework for diagnosis of COVID-19 from chest X-ray images. Interdiscip Sci Comput Life Sci 13:103–117. https://doi.org/10.1007/s12539-020-00403-6
10. Soares F (2020) A novel specific artificial intelligence-based method to identify COVID-19 cases using simple blood exams. medRxiv. https://doi.org/10.1101/2020.04.10.20061036
11. Ong E, Wong MU, Huffman A, He Y (2020) COVID-19 coronavirus vaccine design using reverse vaccinology and machine learning. Front Immunol.https://doi.org/10.3389/fimmu.2020.01581
12. Abdulkareem NM, Abdulazeez AM, Zeebaree DQ, Hasan DA (2021) COVID-19 world vaccination progress using machine learning classification algorithms. Qubahan Acad J 1(2):100–105. https://doi.org/10.48161/qaj.v1n2a53
13. Sear RF, Velásquez N, Leahy R, Restrepo NJ, El Oud S, Gabriel N, Lupu Y, Johnson NF (2020) Quantifying COVID-19 content in the online health opinion war using machine learning, vol 8. IEEE, pp 91886–91893. https://doi.org/10.1109/ACCESS.2020.2993967
14. Samui P, Mondal J, Khajanchi S (2020) A mathematical model for COVID-19 transmission dynamics with a case study of India. Chaos, Solitons & Fractals, vol 140.https://doi.org/10.1016/j.chaos.2020.110173
15. Mele M, Magazzino C (2020) Pollution, economic growth, and COVID-19 deaths in India: a machine learning evidence, vol 28. Springer, pp 2669–2677. https://doi.org/10.1007/s11356-020-10689-0
16. Chen JIZ, Hengjinda P (2021) Early prediction of coronary artery disease (CAD) by machine learning method-a comparative study. J Artif Intell 3(1):17–33
17. Vivekanandam B (2021) Design an adaptive hybrid approach for genetic algorithm to detect effective malware detection in android division. J Ubiquitous Comput Commun Technol 3(2):135–149
18. Ramanujam E, Chandrakumar T, Nandhana K, Laaxmi NT (2019) Prediction of fetal distress using linear and non-linear features of CTG signals. In: International conference on computational vision and bio inspired computing. Springer, Cham, pp 40–47

A Study on Reinforcement Learning-Based Traffic Engineering in Software-Defined Networks

A. Bhavani, Y. Ekshitha, A. Mounika, and U. Prabu

Abstract Modern communication networks have grown highly complex and dynamic, making them difficult to describe, forecast, and govern. So, the software-defined networks (SDNs) have emerged. It is a centralized network, and it is flexible to route network flows. Traffic engineering (TE) technologies are used with deep reinforcement learning (RL) in SDN to make networks more agile. Different strategies for network balance, improvement, and minimizing maximum link usage in the overall network were considered. In this article, recent work on routing as well as TE in SDN and hybrid SDN is analyzed. The mathematical model and algorithm used in each method are interpreted, and an in-depth analysis has been done.

Keywords Software-defined networks · Hybrid software-defined networks · Traffic engineering · Routing · Reinforcement learning

1 Introduction

SDN aims to separate the control plane, and data plane to make network management easier, cut operating costs, and encourage innovation and development [1]. It is a networking design that allows for a dynamically, efficient network arrangement to boost overall network ability while also making networks more agile and adaptable [2]. It enables software applications to govern the network centrally. This allows operators to control the whole network. It allows a controller to make centralized decisions and adaptively design packet-switching nodes [3]. The SDN architecture is shown in Fig. 1.

A. Bhavani · Y. Ekshitha · A. Mounika · U. Prabu (✉)
Department of Computer Science and Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijaywada, India
e-mail: uprabu28@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_4

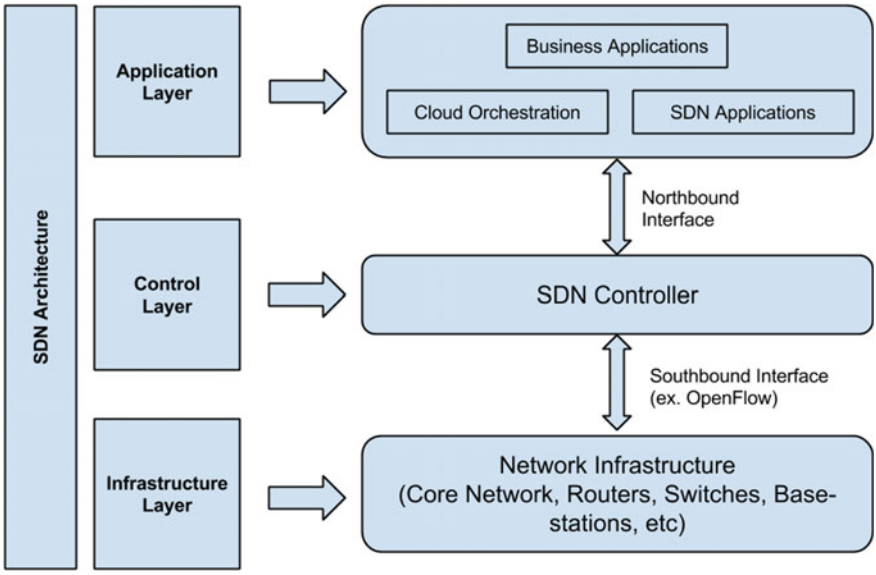


Fig. 1 SDN architecture [4]

1.1 Hybrid SDN

A hybrid SDN may be a networking strategy that mixes ancient networking with software-defined network manner within the same atmosphere [5]. Engineers will operate SDN technologies and old shift rules on constant physical hardware in a very hybrid SDN [3]. Whereas traditional distributed networking protocols still steer the bulk of the traffic on the network, a network manager will style the SDN management plane to seek out and govern sure traffic flows.

RL-Routing is a reinforcement learning routing technique that solves a traffic engineering (TE) problem in terms of throughput and delay in an SDN. Instead of constructing an exact mathematical model, RL-Routing tackles the TE problem through experience. One-to-many network setup for routing choices is used and extensive network information for state representation is considered. The reward mechanism, which uses network throughput and delay to optimize network throughput, can be adjusted to optimize upward or downward network throughput. The agent develops a policy that anticipates future behavior of the underlying network and offers improved routing paths between switches after receiving suitable training.

1.2 Traffic Engineering

TE is a network management technology that enhances the performance of networks by optimizing traffic routing approaches and communicating by anticipating and controlling the behavior of transmitted data [1]. Using data flow to determine link status paths within a network balances the load on multiple connections, routers, and switches [6]. This is particularly important in networks with several parallel paths.

1.3 Routing

Routing is the method of choosing the most convenient path through the associate network to send packets to a destination host or hosts so that the router forwards the packets to those hosts [6]. It is the method of creating a traffic path among a network, likewise as across and across many networks. The neural network's purpose is to reduce network time delay while optimizing the packet pathways that are being addressed. The shortest path is regarded the most important issue in any routing method that may be carried out in real time.

1.4 Reinforcement Learning

RL could be an ML coaching strategy that rewards fascinating behaviors whereas laborious undesirable ones [7]. A reinforcement learning agent will understand and comprehend its surroundings, act, and learn through trial and error usually. It is all regarding working out a way to behave optimally during a given scenario to maximize reward. Associate degree agent explores associate degree unknown atmosphere to attain a goal within the reinforcement learning downside [8]. RL is made on the concept that the increasing of expected accumulative reward could also be wont to represent any goal. To maximize reward, the agent should learn to sense and disturb the state of the atmosphere through its activities.

Both deep learning and reinforcement learning are self-learning systems. Deep learning involves learning from a training set and then applies that knowledge to fresh data, whereas reinforcement learning involves dynamically learning by altering actions depending on continuous response to enlarge a reward. Reinforcement learning and deep learning are not mutually inconsistent.

1.5 Deep Reinforcement Learning

Deep reinforcement learning may be a variety in ML that permits intelligent robots to find out from their actions within the same from that folk do [9]. The fact that associate agent is rewarded or penalized to support their actions is inherent during this variety of machine learning [10]. Deep RL may be an answer that features deep learning permitting agents to form selections supported by unstructured computer files while not manually constructing the state area.

This paper's primary contributions are summarized as follows:

- The article shows different approaches to improve network performance using TE strategies with reinforcement learning in SDN and hybrid SDN.
- Various routing methods in the software-defined network to improve network efficiency.
- The paper shows a comparison of various traffic engineering and routing algorithms in SDN and hybrid SDN.

The remainder of this paper is organized as follows: Sect. 2 represents the various TE methods in SDN and hybrid SDN using deep RL. Section 3 describes the different routing approaches in SDN, and Sect. 4 represents the comparison of various traffic engineering and routing algorithms in SDN and hybrid SDN.

2 TE in SDN and Hybrid SDN Using Reinforcement Learning

2.1 TE in Hybrid SDN

Guo et al. [6] proposed a novel method on the TE in hybrid SDN. With the rise of software-defined networks (SDNs), network routing is becoming more centralized and flexible. Traffic engineering in hybrid SDNs is a topic that is attracting wide interest from academia and industry. RL-based traffic splitting method that learns to balance the dynamically changing traffic and address it through a traffic splitting agent in hybrid SDN. To generate a routing strategy for new traffic needs quickly and intelligently, a traffic splitting agent is created and learned offline using the RL algorithm to build a direct link between traffic demands and traffic splitting rules. The powerful traffic splitting guidelines provides the policies that can be used to set up the traffic splitting ratios on SDN switches. These switches can be advanced quickly and can expand rapidly once the traffic splitting agent has been learned [11]. It is usually recommended to create a suitable simulation environment to avoid routing loops. The traffic splitting guidelines are used to fulfill the interactive standards.

An RL approach for tackling the TE problem of the hybrid SDN consists of two steps: offline learning and online routing. Draw the directed acyclic graph (DAG) [12] based on hybrid SDN architecture and traffic statistics to ensure a hybrid SDN

environment with loop-free routing for the RL agent's interaction during the offline learning stage. A traffic-partitioning agent is taught to make a direct link between the network environment and the routing techniques using the RL methods and the constructed DAG. When the demand of the traffic changes, the trained traffic splitting agent can quickly establish an acceptable routing scheme at the online routing stage.

The TE problem's network model is a hybrid SDN with dynamic traffics. The network topology is represented as an undirected graph $H = (X, F)$, where X is the set of forwarding devices, which is made up of SDN switch Xs and router $X1$, and F is the link set, with $C(f)$ denoting the capacity of connection $f \in F$.

The group of (TMs) may be indicated by $b = \{B_1, B_2, \ldots, B_n\}$ also r_i may be the weight constant of TM b_i. Those components $B_i(p, q)$ to b_i speak the traffic demand starting from gadget p to gadget $q.p,q \in X$, W is the link weight setup under the OSPF protocol. Variable e means the partitioning of traffic flows.

A collection of TMs B, the purpose of TE on the hybrid SDN is to increase the network performance by minimizing the maximum link usage (MLU) [13] for every single TM B_i. In the hybrid SDN, both the link weight settings w and the partitioning traffic movement f on each SDN switch decide the result of MLU. TE problems on hybrid SDN are defined as

$$\textbf{Minimize} \sum_{i=1}^{n} r U_i^{\max}$$

$$\sum_{j=1}^{n} r_i = 1, \quad 0 \le r_j \le 1$$

$$w(f) \in N, \ \forall f \in F$$

$$0 \le U_i^{\max} \le 1$$

A node $q \in X$ and TM $B_i \in B$, we calculate node q's MLK as below:

$$\text{MLK } B(v) = \max U_i(e)$$

The value of MLK of node q in different TMs is calculated based on MLK(q) by computing equation

$$\text{score}(v) = \sum_{B_i \in B} \text{MLK } B_i(v)$$

The TE's goal in hybrid S reduces the MLU as much as possible. As a result, MLU is included in constructing reward r_t as defined to enable the traffic-partitioning agent to effectively learn optimal rules with small MLU

$$r_t = \begin{cases} -e^{2(\frac{1}{a}-1)} \\ 0 \\ e^{2(a-1)} \end{cases}$$

$$a = U_i^{\max} / U_t^{\max}$$

The index a represents the level of performance improvement after traffic-partitioning rules are implemented at t time.

RL approach is compared to other methods such as open shortest path first and WA-SRT. RL technique outperforms the OSPF method in terms of MLU reduction and comes close to the WA-SRTE method [14]. The RL approach demonstrates the ability to relate network status to traffic-partitioning rules, assisting in the link capacity balancing process.

2.2 TE in Hybrid SDN

Zhang et al. [5] proposed a novel method for TE issues with RL. By rerouting as several flows as potential, traffic engineering approaches area units capable of achieving the best performance [13]. One TE methodology for mitigating the impact of a network outage is to balance link usage by forwarding the bulk of traffic matrices that area unit wedged. The networking unit provides two examples: ECMP [15] and resending form of essential flows utilizing a code package. Extremely, crucial flows rerouting RL is another RL technique, mechanically learn the rules for choosing needed flow for every TM. CFR-RL formulates and solves an easy arithmetic drawback so as to direct these essential flows so as to balance network association usage.

An RL-based technique for balancing network link utilization by learning critical flows selecting strategy and rerouting relevant essential flow. CFR-RL uses the reinforce [16] technique to train this neural network, with minor tweaks. Reward and state space of a key flow selection strategy utilizing a customized RL technique are all included. State space inputs are the traffic matrix, and RL issue necessitates a huge action space containing massive network nodes, with the LP function as a reward. ECMP routing is used to disperse the traffic.

State: An agent is given state $c_t = T_t$, where T_t is TM at time t that provides information about each flow's traffic demand.

Action Space: Action spaces are of two types. They are discrete action space and continuous action space. Here, the agent chooses from a finite action set which distinct action to do using a discrete action space. Actions are conveyed as a single real-valued vector in a continuous action space. CFR-RL would choose P essential flows for each state st. Given that a network with K nodes has a total of $K * (K - 1)$ flows, RL issue will necessitate huge action space the value of the $C^p_{k*(k-1)}$. Permit agent to some P distinct actions in every time t by setting the action space to $(0, 1, \ldots, (K * (K - 1)))$ [17].

Reward: CFR-RL resends these key flows and gets the highest link usage U by resolving the problem of routing optimization after sampling P distinct essential flows for state c_t. Reward q is set to $1/U$ that reflects network performance after crucial traffic is rerouted to balance link use.

The following is a description of the crucial flow rerouting problem. The goal is to acquire optimal routing ratios $\sigma_{s,d}^{i,j}$ for each essential flow, so in MLU, U is minimized, given a network $H(X, F)$ with the group of demands of traffic Bs, d for the set of essential flows (fk) and link load $L_{i,j}$ supplied by the rest of the flows that are utilizing the default settings. We construct the problem of rerouting as an optimization to find all viable under-utilized pathways for the specified important flows

$$\textbf{minimize } U + \sum_{(i,j)\in F} \sum_{(s,d)\in f_k} \sigma_{i,j}^{s,d}$$

The optimal routing result for chosen important flow is derived by addressing the aforementioned LP problem with LP solvers. The SDN controller then installs and changes flow entries at the switches in the appropriate order.

The CFR-RL scheme is presented with the goal of lowering maximum link utilization in a network and minimizing network disturbance that causes service disruption. By resending just 11–20.3% of the entire traffic, CFR-RL delivers near-optimal performance. CFR-RL achieves optimal load balancing performance in excess of 95% of the time.

2.3 ScaleDRL Scheme for TE in SDN Using Pinning Control

Sun et al. [10] proposed a method for SDN issues using TE. Deep reinforcement learning and software-defined learning able to develop a model-free TE system with the assistance of networking technologies. Existing DRL-based TE results, on the opposite hand, all have a quantifiability issue that forestalls them from getting used in giant networks. A method that mixes management theory associate degreed DRL technology is planned to develop an economical network management approach for TE [18]. ScaleDRL could be a planned approach that employs the construct of promise management theory to spot and label the group of network links as essential. Supported traffic distribution info received by SDN controller, DRL approach is utilized to effectively change the collection of link loads for vital links. The forwarding pathways of network flows are often dynamically modified employing a weighted shortest path technique.

ScaleDRL presents a mechanism for evaluating the importance of network links in routing path development. A control theory-based flow selection algorithms is developed on this assessment approach [19]. ScaleDRL adapts the DRL method to manage communication network traffic allotment statistics and dynamically construct TE

regulations. ScaleDRL is validated using OMNet++ in a fine-grained simulation with several network topologies of various sizes.

The link weights are the goal of DRL algorithm in ScaleDRL. As a result, in order to pick a fraction of data plane links with pinning control, we basically introduce the concept of equilibrium to define the flow relevance in overall network's routing patterns. Link centrality, in particular, refers to the correlations that exist between links as a result of routing pathways.

For network H, R is to represent the group of node r and F is to represent links f between nodes, having $H = (R, F)$. For a couple of node r_i, $r_j \in R$, that shall contain at least a single readdressing path $p_{i,j} = \{f1, f2, \ldots, f|o|\}$ that will send traffic from r_i to r_j nodes, and represent the shortest path between them as $o^*_{i,j}$.

With weighted shortest path algorithms, computed the shortest path on H, where value of weight for the link fm is represented by w_m $(0 \leq m \leq |F|)$. Use the indicator $y^m_{i,j}$ to denote if link fm is acquire in $o^*_{i,j}$ that is $y^m_{i,j} = 1$ if $o^*_{i,j}$ carry fm, $y^m_{i,j} = 0$. The equilibrium of the link fm is defined as

$$f_m = \frac{\sum_{i=1}^{|R|} \sum_{j=1}^{|R|} x^m_{i,j}}{|R| \times |R|}$$

A MDP is used to define model DRL's working process (MDP) [20]. The DRL algorithms in the MDP interconnect with target environment. MDP is characterized as follows:

$$N = (T, A, Q, S, Z),$$

with T represents state space s, A represents action space a, Q represents reward space q, S represents the group of probability P, and Z denoting a discount factor.

The entire training goal of DRL algorithm is to increase the cumulative rewards

$$Q = \sum_{t=0}^{T} \gamma^t q_t$$

In numerous network topologies, the packets simulation reveals the ScaleDRL decreases typical end coordinated universal time up to 40% in comparison with progressive DRL-based TE system. The link centrality-based choice theme had the best performance of all the schemes studied that confirms the link centrality-based choice strategy.

2.4 RL Approach for TE Based on Link Control

Xu et al. [3] proposed a method for TE issues using RL. Deep reinforcement learning (DRL) permits to use of machine learning to make a model-free TE theme.

Existing DRL-based TE solutions, on the opposite hand, cannot be utilized in massive networks. To develop a TE theme, a theme that mixes management theory and DRL is bestowed. The prompt arrange ScaleDRL selects link range in network and names the essential flows employing a construct from promise management theory [21]. A DRL technique is employed to dynamically alter the weights of links for essential flows supported traffic distribution statistics. The forwarding pathways of the flows will be dynamically changed employing a weighted shortest path technique.

The basic idea for pinning management is during advanced and dominant whole network elements to attain network state consumes a lot of resources within the management algorithmic rule and so does not happen; instead, bound management signals in mere a part of the network may be removed to attain the supposed synchronization mode. ScaleDRL is enforced supported SDN. There is a vital link algorithmic rule and a DRL algorithmic rule [22] that resides within the ScaleDRL management. With SDN, traffic distribution may be collected from time to time by the controller, and TE rules may be upgraded sporadically. Supported this technology, ScaleDRL operates in 2 categories: offline and online. Within the offline section, the link choice algorithmic rule analyzes constellation and a select group of flows as important network flows supported pin-control perspective. Within the online section, the DRL algorithmic rule manages flow weights to direct traffic networks.

In link weight algorithm [19], for network H, R is to represent the group of node r, f is to represent links f between nodes, having $H = (R, F)$. For a couple of nodes $r_i, r_j \in R$, that shall contain at least a single readdressing path $p_{i,j} = \{f1, f2, \ldots, f|o|\}$ that will send traffic from r_i to r_j nodes, and represent shortest path between them as $o^*_{i,j}$.

With weighted shortest path algorithms, computed the shortest path on H, where value of weight for the link f_m is represented by w_m $(0 \leq m \leq |\mathrm{F}|)$. Use the indicator $y^m_{i,j}$ to denote if link f_m is acquire in $o^*_{i,j}$ that is $y^m_{i,j} = 1$ if $o^*_{i,j}$ carry f_m, $y^m_{i,j} = 0$. The equilibrium β of the link f_m is defined as

$$\beta(b_m) = \frac{\sum_{i=1}^{|R|} \sum_{j=1}^{|R|} y^m_{i,j}}{|R| \times |R|}$$

DRL algorithm develop an action with in neural network to the surroundings at time t of MDP supported the observation of the state k_t. The DRL algorithm obtains reward r_t after a_t is performed in the environment, which assesses a_t's performance in the environment. There are currently several forms of DRL algorithms, with the key distinction being the mechanism used to update the neural network parameters in DRL. The DRL framework we are using is ACKTR [23].

The packets simulation reveals the ScaleDRL decreases typical end coordinated universal time up to 40% in comparison with progressive DRL-based TE system.

3 Routing in SDN Using Deep Reinforcement Learning

3.1 RL-Routing: An SDN Routing Algorithm Based on DRL

Chen et al. [7] proposed a method for routing in SDN based on RL. Because communication networks have become so complex and dynamic, they are challenging to describe and anticipate. To overcome traffic engineering challenge of SDN with respect to throughput, latency to create a reinforcement learning routing method. Instead of constructing an exact mathematical model, RL-Routing tackles TE problems through training. To employ a one-to-many network setup for routing options and use extensive network information for state representation. The reward method, which will use networks work rate and delay to optimize network throughput, may be adjusted to optimize up or down network work rate. The algorithm develops a rule which anticipates further behavior of basic networks and offers improved routing pathways among switches after receiving suitable training.

The RL-Routing is located, how it interconnects with remaining components in SDN architecture. For message exchange, the controller attaches to switch over the OpenFlow path. There are two main modules in the RL-Routing application. They are network monitoring modules, which obtain network information through passive and active network measurements [24]. Network information relates to the position of network tool, such as flow delay and flow work rate. Another module is the action translator module, which converts the algorithm chosen action into the suitable group of OpenFlow messages to modify switch link tables.

Assume network as DGH(X, F) where $X = \{s_1, s_2, \ldots s_n\}$ is set of switches and $F \subseteq X \times X$ is the group of the flows in a network, $|F| = m$. Consider the network flows are two-way directional means $f_{i,j}$ and f_j, I are upward and downward flows connect to s_i. Neighbors of switches s_i are $N(s_i) = \{s_j \in X\}.F(s_j) = \{f_i, k \in F\}$ is group of adjacent edge to s_i. s_{src} is the source switch.

Let Dsrc $\subset X - \{s_{\text{src}}\}$ be collection of entire destination switch from s_{src}. Path $p_{\text{src,des}}$ is a path in network $H(X, F)$ that interconnects s_{src} to s_{des} to the order of switch $(s_{\text{src}}, s_i, s_j, s_k, s_{\text{des}})$, where the besides switches in order form edges in F, and every switch will be visited once. $b_t(f_{i,j})$ is link bandwidth $f_{i,j}$ that interconnects s_i to s_j at period intervals Δt. Delay$(f_{i,j})$ anderror $t(f_{i,j})$ represent delay of link and indicator for the fault occurred in $f_{i,j}$ for time intervals Δt. Bandwidth of path $b_t(p_{\text{src,des}}) = \min b_t(f_{i,j})$ is min bandwidth of the link at period intervals Δt. Delay of the path $\text{delay}_t(p_{\text{src,des}}) = \sum_{f_{i,j}} \text{delay}(f_{i,j})$ is addition of flow delay in path at period intervals Δt.

The problem of TE is represented as given $H(X, F)$, Ssrc, Dsrc and finds a group of flows for the succeeding ssrc's data to switch in Dsrc. The aim is to increase swsrc's work rate and reduces communications delay.

The description of RL-Routing is given here. Then, provided a Q learning algorithm [25] to resolve the traffic engineering problem.

Description of RL-Routing

1. The template for routing is designated as $O = (L, B, S, U, N)$ where $L \subset Sz$ represents state space.
2. B represents action space.
3. $S : T \times B \rightarrow S$ represents reward function.
4. U represents transition probability.
5. $N \in [0, 1]$ represents discount rate.

On various network topologies, simulation output demonstrates the RL-Routing earns greater reward and allows host to send the big file quicker than OSPF and LL algorithm. On the NSF Net topology, for example, the total of RL-rewards routings is 119.30, where OSPF and LLs are 106.59 and 74.76. The average RL-Routing transmission period for 40 GB file is 25.2 s. OSPF and LL have 63 and 53.4 s.

3.2 TIDE: Time-Relevant Deep Reinforcement Learning for Routing Optimization

Sun et al. [8] proposed a novel method for routing issues. TIDE is a smart network control architecture in view of DRL that can progressively advance routing algorithms in an SDN network without requiring human cooperation. TIDE has been completely tried and executed in a real-world network setting. The discoveries of the test show that TIDE can change the directing system powerfully founded on the network and can limit in general network sending delay by about 9% contrasted with standard calculations. The optimization has been studied for a protracted time in network style, and several other optimization ways are given by each lecturers and business. However, such systems are either too troublesome to use in applications or perform poorly. AI-based routing ways are planned in early years, with an emergence of SDN and computing. TIDE, associate degree intelligent network management design supported DRL that may effectively improve routings ways in associate degree network while not requiring intervention of human, is planned during this study. TIDE is tested and enforced during a real-world network surroundings. The results of the experiment show that TIDE will dynamically adapt the routing strategy supported the network state of affairs and improve the full network transmittal delay.

To implement the automated routing strategy in SDN, an initial have to be compelled to build an Associate in Nursing intelligent network management design known as TIDE [26]. 3 logic planes form up the recommended design for intelligent network control: information plane, management plane, and AI plane. There are 3 components to the intelligent call loop: reward, state and assortment, rule development, and policy preparation. The most contributions are to execute intelligent routing management of a sending network, "collections-decision-adjustment" loop is bestowed, and RNN-based DRL system is rigorously created for abstracting traffic properties and might effectively develop a closure optimal routing set up counting on the ever-changing traffic distribution.

The fundamental method of TIDE is DDPG [27], a DRL framework for constant control. DDPG's result is not distinct as a narrow group of some actions, unlike the bulk of reinforcement learning models like DQN. In order to elegantly regulate entire network traffic flows, routing optimization usually requires adjusting the link capacity for every link. As a result, the link capacity space value in the network should be large, making constant algorithms like DDPG is a good choice for creating routing strategies.

The interconnection process between agent and environment is viewed as a MDP in RL. The element tuple of MDP is $O = (V, B, K, D, Z)$, where V represents state space v, B represents action space b, K represents reward space k, D represents transition probability method, and $z[0, 1]$ represents discount factor. An agent selects action b under state v according to rule, which is represented as $(b|v)$ in normal rules and $b = (v)$ in deterministic rules.

Value functions are used to determine if a policy is beneficial or not. The value of C is a prominent value function in reinforcement learning. When choosing action b in state v, value of policy C is defined as [28]

$$C(st, b) = E\left[\sum_{k=0}^{\infty} y^k K(v_{t+k}, b_{t+k})\right]$$

TIDE decreases the overall transmission latency of entire traffic by around 9%. This is due to the growing unpredictability of noise traffic, which makes it more difficult for TIDE categorize network traffic, limiting TIDE's capacity to make perfect decisions.

3.3 QR-SDN: Toward Reinforcement Learning States, Actions, and Rewards for Direct Flow Routing in Software-Defined Networks

Rischke et al. [9] proposed a novel method for SDN issues. QR-SDN could be an ancient tabular reinforcement learning system that builds and evaluates routing patterns of single flows in an action statehouse. The findings are accustomed produce a model-free reinforcement learning strategy. Owing to direct illustration of link routes within QR-SDN action statehouse, QR-SDN is the initial RL-Routing technique that changes many routing ways in which among given offer switch destination try whereas holding flow integrity. In alternative words, with QR-SDN, packets from an eternal flow follow a set routing path, however, flows from a continuing source destinations switches might take a spread of routes. QR-SDN tends to be enforced in an exceedingly extremely SDN compete for the testbed.

The presentation of SDN link routing drawback to the economical higher cognitive process by RL agent is not totally been investigated. The planning of the states and actions, particularly, should be self-addressed as to adequately represent link routing

drawback for a process by RL algorithm, to see, however, with success an action solves the flow routing drawback, we tend to utilize the reward. The total of latencies on these pathways of the flows is the planned incentive.

Assume the network $H(W, X)$, where X is a collection of edges that connects a group of vertices W. We prefer to focus on a single communication flow, that is, the flows that convey data from a single sender to a single receiver. Flow e represents information transfer from a given sender se to a given receiver de for a certain application or transport layer context, such as a given TCP flow. E is commonly used to represent the group of all flows. We usually assume that flow e transfers a certain traffic rate Q_e into the network from supply host.

The path $V_{s,d}$ is an order of vertices $V = (p_1, \ldots, p_n)$ from a group of every possible path $V \in V_{s,d}$ interconnecting s to d, where group $V_{s,d}$ might be defined by search algorithms like DFS [29–33].

The SDN controller's RL agent monitors the environment by monitoring the required main performance indicator, as bandwidth, at different times $t = 0, 1, 2 \ldots$. The observation contains reward $S_t \in S \subset S$ and the environment's state R_t from the group of states $R = \{R_1, R_2, \ldots\}$.

The state R_t should be made up of a table with the presently chosen paths Q for every flow e. An action $B_t \in B$ is chosen based on the state R and its accompanying reward S. The group of alternative paths including the present path determines the set of actions $B = \{B_{t,1}, B_{t,2}, \ldots\}$. The total latencies L_e along the present routes $V_{s,d}$ of flows $e \in E$ is reward S_t.

For moderate to high loads, the link-preserving different routing of paths QR-SDN provides much lesser latencies than conventional unicast path routing systems, according to the tests. Shifts, like load changes owing to additional flows that end, are successfully accommodated by QR-SDN.

4 Comparison of Various Traffic Engineering and Routing Algorithms in SDN and Hybrid SDN

In the section, various traffic engineering techniques and routing algorithms used in SDN and hybrid SDN are analyzed and given in Table 1.

The mentioned algorithms or techniques (Table 2) are used to improve the performance of the network and to increase the efficiency of routing in software-defined networking. Deep reinforcement algorithms and traffic engineering techniques that mentioned in the previous approaches are to optimize the maximum link utilization and to improve the flow routing in the network. Link selection algorithms are used to optimize distributed estimation and increase network performance. DRL algorithms are used for traffic control and channel rerouting in the network. RL framework helps to improve dynamically routing of flows in the network. So, to improve the efficiency of networks and to increase the overall performance of the SDN, it is needed to use traffic engineering schemes and reinforcement learning methods in the model.

Table 1 An overview on various traffic engineering techniques and routing algorithms

Research work	Addressed issue	Compared with	Result
Xu et al. [1]	Traffic engineering	Shortest path (SP), load balance, DDPG	DRL-TE consistently outperforms DDPG
Wu et al. [2]	Deep reinforcement learning	Convolution neural network (CNN), deep Q learning networking (DQN)	Existing methods are outperformed by the proposed algorithm
Sun et al. [3]	Optimal traffic scheduling	TIDE, DRL-TE	ScaleDRL has better control performance than other DRL solutions
Zhang et al. [5]	Traffic engineering	ECMP	CFR-RL, able to derive to unknown traffic matrices, according to the evaluation findings
Guo et al. [6]	Traffic engineering	OSPF, WA-SRTE, MCF	The proposed ROAR method achieves near-optimal network performance in the hybrid SDN
Chen et al. [7]	Deep reinforcement learning	LL, open shortest path first (OSPF), RL-Routing	As a result of RL-Routing, a host can transfer the large data effectively than OSPF and receive higher rewards as a result
Sun et al. [8]	Routing optimization	Quality of service (QoS)	The effectiveness of TIDE is validated
Rischke et al. [9]	Reinforcement learning	Deep reinforcement learning-traffic engineering (DRL-TE)	A traditional tabular RL technique for the link routing in SDN was developed and assessed
Sun et al. [10]	Deep reinforcement learning	Pinning control	Validate the effectiveness of ScaleDRL

5 Conclusion

The state-of-the-art TE techniques and routing algorithms in SDN and hybrid SDN were analyzed in-depth. For each of the article, the issues addressed, mathematical model or algorithm used along with its core classification is tabulated very clearly for the researchers to have an idea on the literature. The study gives a very elaborated insight.

Table 2 Mathematical model/algorithm used

Mathematical model/algorithm	Purpose of usage	Classification
DRL-TE	The DRL-TE system is greatly reducing end-to-end delay as well as improving total utility	Reinforcement learning algorithms
TCCA-MADDPG algorithms	For traffic control and channel rerouting, the objection function must be optimized	Deep RL method
Link selection algorithm	To optimize the distributed estimation and improve the performance of network by changing topology	Sorting algorithms
Heuristic algorithm	Faster and more efficient approach to solving a problem	Reinforcement learning in SDN
Reinforcement learning algorithm	In computing, a method of determining what actions will be taken by software agents in a given circumstance	Reinforcement learning
Reinforcement learning routing algorithms	RL framework helps improve adaptive routing algorithms	Reinforcement learning
1. Deep deterministic policy gradient algorithms 2. Markov decisions process	The Q-function and the policy are simultaneously learned	Graph theory
Flow routing algorithm	To address the routing flows	Flows routing in SDN
DRL algorithm	Flows weights for selected flows are dynamically adjusted with the DRL algorithm	Pinning control

References

1. Xu Z, Tang J, Meng J, Zhang W, Wang Y, Liu CH, Yang D (2018) Experience-driven networking: a deep reinforcement learning based approach. In: IEEE INFOCOM 2018-IEEE conference on computer communications. IEEE, pp 1871–1879
2. Wu T, Zhou P, Wang B, Li A, Tang X, Xu Z, Ding X (2020) Joint traffic control and multi-channel reassignment for core backbone network in SDN-IoT: a multi-agent deep reinforcement learning approach. IEEE Trans Netw Sci Eng 8(1):231–245
3. Sun P, Lan J, Li J, Zhang J, Hu Y, Guo Z (2020) A scalable deep reinforcement learning approach for traffic engineering based on link control. IEEE Commun Lett 25(1):171–175
4. https://www.fiber-optic-cable-sale.com/will-sdn-change-future-network.html
5. Zhang J, Ye M, Guo Z, Yen CY, Chao HJ (2020) CFR-RL: Traffic engineering with reinforcement learning in SDN. IEEE J Sel Areas Commun 38(10):2249–2259
6. Guo Y, Wang W, Zhang H, Guo W, Wang Z, Tian Y, …, Wu J (2021) Traffic engineering in hybrid software defined network via reinforcement learning. J Netw Comp Appl 103116
7. Chen YR, Rezapour A, Tzeng WG, Tsai SC (2020) RL-routing: an SDN routing algorithm based on deep reinforcement learning. IEEE Trans Netw Sci Eng 7(4):3185–3199
8. Sun P, Hu Y, Lan J, Tian L, Chen M (2019) TIDE: time-relevant deep reinforcement learning for routing optimization. Futur Gener Comput Syst 99:401–409

9. Rischke J, Sossalla P, Salah H, Fitzek FH, Reisslein M (2020) Qr-sdn: towards reinforcement learning states, actions, and rewards for direct flow routing in software-defined networks. IEEE Access 8:174773–174791
10. Sun P, Guo Z, Lan J, Li J, Hu Y, Baker T (2021) ScaleDRL: a scalable deep reinforcement learning approach for traffic engineering in SDN with pinning control. Comput Netw 190:107891
11. Fortz B, Thorup M (2000) Internet traffic engineering by optimizing OSPF weights. In: Proceedings IEEE INFOCOM 2000 conference on computer communications. Nineteenth annual joint conference of the IEEE computer and communications societies (Cat. No. 00CH37064), vol 2. IEEE, pp 519–528
12. Tian Y et al (2020) Traffic engineering in partially deployed segment routing over IPv6 network with deep reinforcement learning. IEEE/ACM Trans Network 28(4):1573–1586. https://doi.org/10.1109/TNET.2020.2987866
13. Guo Y, Wang Z, Yin X, Shi X, Wu J (2014) Traffic engineering in SDN/OSPF hybrid network. In: 2014 IEEE 22nd international conference on network protocols, pp 563–568. https://doi.org/10.1109/ICNP.2014.90
14. Zhang J, Xi K, Luo M, Chao HJ (2014) Dynamic hybrid routing: achieve load balancing for changing traffic demands. In: IEEE international symposium on quality of service'14, pp 105–110
15. Villamizar C (1999) Ospf optimized multipath (ospf-omp). IETF Internet-Draft, draft-ietf-ospf-omp-03.txt
16. Williams RJ (1992) Simple statistical gradient-following algorithms for connectionist reinforcement learning. Mach Learn 8(3):229–256
17. Mao H, Alizadeh M, Menache I, Kandula S. Resource management with deep reinforcement learning. In: ACM workshop on hot topics in networks'16, pp 50–56
18. Xu Z, Tang J, Meng J, Zhang W, Wang Y, Liu CH, Yang D (2018) Experience driven networking: a deep reinforcement learning based approach. In: IEEE INFOCOM 2018-IEEE conference on computer communications. IEEE, pp 1871–1879
19. Liu Y-Y, Slotine J-J, Barabási A-L (2011) Controllability of complex networks. Nature 473(7346):167
20. Wu Y, Mansimov E, Grosse RB, Liao S, Ba J (2017) Scalable trust-region method for deep reinforcement learning using kronecker-factored approximation. In: Advances in neural information processing systems, pp 5279–5288
21. Lin S-C, Akyildiz IF et al (2016) Qos-aware adaptive routing in multi-layer hierarchical software defined networks: a reinforcement learning approach. In: 2016 IEEE international conference on services computing (SCC). IEEE, pp 25–33.
22. Xu Z, Tang J et al (2018) Experience-driven networking: a deep reinforcement learning based approach. In: IEEE INFOCOM 2018-IEEE conference on computer communications. IEEE, pp 1871
23. Wu Y, Mansimov E, Grosse RB, Liao S, Ba J (2017) Scalable trustregion method for deep reinforcement learning using kronecker-factored approximation. In: Advances in neural information processing systems, pp 5279–5288
24. Boyan JA, Littman ML (1993) Packet routing in dynamically changing networks: a reinforcement learning approach. In: Proceedings of 6th international conference on neural information processing systems, NIPS'93. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp 671–678. [Online]. Available: http://dl.acm.org/citation.cfm?id=2987189.298727
25. Openflow switch specification version 1.3.5. open networking foundation. [Online]. Available: https://www.opennetworking.org/wpcontent/uploads/2014/10/openflow-switch-v1.3
26. Karakus M, Durresi A (2017) Quality of service (QoS) in software defined networking (SDN): a survey. J Netw Comput Appl 80:200–218
27. Lillicrap TP, Hunt JJ, Pritzel A et al (2015) Continuous control with deep reinforcement learning. Computer Science 8(6):A187
28. Silver D, Lever G, Heess N et al (2014) Deterministic policy gradient algorithms. In: International conference on machine learning, pp 387–395. JMLR.org

29. Tarjan R (1972) Depth-first search and linear graph algorithms. SIAM J Comput 1(2):146–160
30. Guck JW, Van Bemten A, Reisslein M, Kellerer W (2018) Unicast QoS routing algorithms for SDN: a comprehensive survey and performance evaluation. IEEE CommunSurveys Tuts 20(1):388–415
31. Chen JIZ, Smys S (2020) Optimized dynamic routing in multimedia vehicular networks. J Inf Technol 2(3)
32. Bhalaji N (2021) Cluster formation using fuzzy logic in wireless sensor networks. IRO J Sustain Wirel Syst 3(1)
33. Ram GM, Ilavarsan E (2021) Review on energy-efficient routing protocols in WSN. In: Computer networks, big data and IoT

Enhancing the Performance of Heterogeneous Networks Using Optimized Cluster-Based Algorithm

Abdullah Mohammed Abdullah Al-Amodi, Amlan Datta, Abdulrahman Mohammed Hussain Obaid, and Arunabha Das

Abstract With an ever-increasing number of user equipment (UEs) and rising bandwidth demands of new applications, deployment dense heterogeneous cellular networks have been embraced in various network scenarios. The cells experience unloaded due to the random UEs mobility and cells deployment, which degrades the network performance such as handover success, throughput, and load distribution. We propose an enhanced cluster-based algorithm for small cell mobility load-balancing networks to address such a problem. The conventional mobility load-balancing (MLB) algorithms consider only the contiguous neighboring cells and do not expand enough performance of the network, while other MLB algorithms consider the neighboring cells of the total network experienced unneeded MLB actions. The proposed load-balancing algorithm studies overloaded cells and neighbors using the proposed efficiency parameter. To begin with and to identify the overloaded cells, we propose an efficiency factor *B* that compares a pre-defined threshold and the network threshold to control the algorithm triggering by modifying the CIO parameters of the cluster cells in both medium-loaded and overloaded cells triggering in both medium-loaded and overloaded cells. Then, to control the distribution, we propose a method to shift only a portion of the serving cells load, so the target cell load after handover always be equal to or less than *B*. In a low UE speed scenario, the simulation results showed a lower standard deviation (SD) by 19.99% and enhanced throughput by 7.355%.

A. M. A. Al-Amodi (✉) · A. Datta · A. Das
Electronics School, KIIT University, Bhubaneswar, Odisha, India
e-mail: Eng.aalamodi@gmail.com

A. Datta
e-mail: amlanfet@kiit.ac.in

A. Das
e-mail: 1904411@kiit.ac.in

A. M. H. Obaid
21 September University of Medical and Applied Sciences, Yemen, Yemen
e-mail: obaid.eng@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_5

Keywords Self-organizing network · Small cell · Mobility load balancing · HetNets · Throughput · Handover · Physical resource block

1 Introduction

Rising demands for mobile communication services with an increasing high data rate and quality of service (QoS) stimulate researchers to explore alternative solutions. In the year 2022, the anticipated demands of wireless data are 77 exabytes each month which represent 49 times of 2016 demands [1] [edge]. Small cells with low power and cost, and cover ten to several hundreds of meters were introduced to support those demands and increase the capacity to play the main role in the fifth generation (5G) of cellular communication, it is used to cover blind spots of the macrocells to raise the capacity of the mobile wireless network as well as the throughput which make the small cell density is the spine of the 5G networks [2, 3]. The design of the small cell was originally to extend the coverage of macrocell, the more utilization of small cells in a wireless network, the better throughput, and capacity the network gained. Small cells deployment increases rapidly for residential and non-residential areas, the deployment could be planned or unplanned depending on the policy of operators [cloude load est.] analysis [4]. In small cells, the mobility of user equipment (UE) increases due to the low power of small cells producing load imbalance of small cells across the network and the performance of the network degrades related to handover (HO) rate of success and capacity. The UEs may move to a small cell with a high data rate request which is higher than the cell capacity, the cell gets congested and leads to failures of HO or poor quality of service (QoS) [2, 5].

Diverse solutions have been proposed to decrease the problem of load balancing and enhance the network performance, mobility parameters for intra-LTE in comparison with the mobility of non-optimized/static cell characteristics, the auto-adjustment depending on the present load of the small cell improve the capacity of the system [6]. However, the forced load balancing must not have a detrimental impact on UE QoS. In [7], researchers were the first to demonstrate the effectiveness of basic load-balancing algorithms in decreasing call blockage rates and increasing cell edge throughput using simulation for auto-adjustment. In [8], overloaded cells arrange UEs according to the best nearby eNB based on measurements of RSRP and load of neighboring cells, per each handover offset value, it classifies the surrounding eNBs in decreasing order depending on the number of feasible handovers. If the anticipated load at the neighboring eNB does not exceed the allowed level, the entire group will be handed over. In the optimization area, the MLB approach introduced in [9, 10] took into account non-adjacent neighborhood cells in the optimization zone. When shifting UEs from source cells, the radio connection state of neighboring cells is considered. MLB enhancing and handover parameter optimization (HPO) algorithms affect UE handover decisions. This interaction minimizes the desirable outcomes of every function. The coordination among MLB and HPO is examined in [11]. The coordinator gives a solution that volume the capabilities of the separate

algorithms to improve performance. In [12], the researchers introduced a multi-traffic load-balancing algorithm that restores the load to enhance the capacity of the network. The algorithm improves the quality of service of the UE (QoS) by adopting a new cell reselection method, two parameters (handover and time to trigger (TTT)) are adaptively modified to eliminate the rates of call drop and enhance the load balancing by handing over the edge UE of serving cells to the under-load neighbor cells. The process considers two conditions: physical resource blocks (PRB) and signal strength to detect statues of overloaded cells, and a fixed threshold is used which cannot predict adopting the variation of the network.

In [13], the algorithm was presented based on the status of network load and load estimation, the algorithm adjusts the parameters of HO depending on to load status of the overloaded serving cell and neighbor cells. Overloaded cells are detected by the adaptive threshold, which is changed according to the average load of the network situation. To prevent the HO ping pong, the movable load from overloaded cell to light-loaded neighbor cell was considered with restricted conditions. To control the overloaded serving and light-loaded neighbor cells. Besides, the algorithm takes into account the estimated load of both overloaded serving and light-loaded neighbor cells. However, the algorithm did not consider the overloaded cell load status as well as the load estimation of the edge UE load after handover. In [14], an enhanced adaptive load-balancing algorithm for small cells was introduced, and the parameters of HO are adjusted over the network by proposing a "load-balancing efficiency factor." The algorithm estimates the edge UE loads after handover and the remaining available load of neighbor cells and specifies the operation sequence. However, offloading from an overloaded cell to light-loaded cell was restricted with conditions, one of these conditions was the difference load between the overloaded and light-loaded cells and was less than the gap threshold (0.1) which in some cases the offloading cannot be accomplished. In [15], the authors introduced an algorithm to create cluster MLB. The method takes into account any overloaded cells in the network and the adjacent cells to create *N* tiers. The technique provides local load balancing by modifying the CIO parameters of the cluster cells. Since load balancing is done inside clusters, the unrequired MLB activities are eliminated to improve the network more effectively.

This work is structured as: Section 2 explains the system model, it describes the adaptive threshold as well. Section 3 provides an overview of the load-balancing control factors. The architecture of the access network is shown in Fig. 1, and the proposed algorithms are presented in Sect. 4. Then, the performance assessment is introduced and followed by the paper's conclusions.

2 System Model

The main objective of this paper is to create a balanced network and take the advantage of the unused physical resource block (PRB) of the neighbor cell for those overloaded

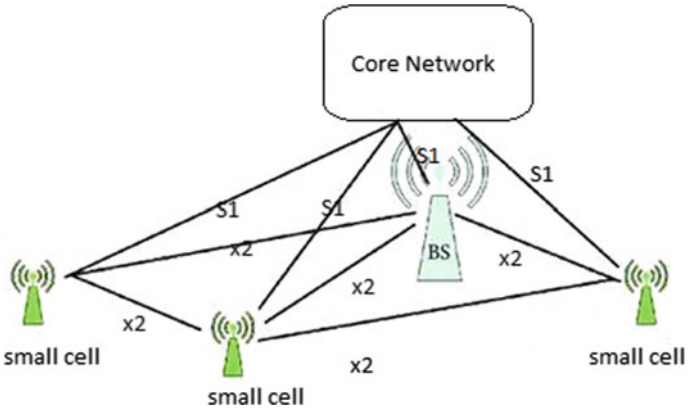


Fig. 1 Architecture of the network

serving cells which can be accomplished by shifting the UEs of overloaded cell edge to the light-loaded cell and applying some measurements.

2.1 Network Architecture

This work considers the ultra-dense LTE network which is made up of a collection of cells, S, and includes small and macrocells, and several UEs, E, as done in [16]. The cell set is denoted by $S = 1,\ldots, S$, while the UE subset is denoted by $E = 1, \ldots, U$. The macrocells of the network in the scenario were organized in a hexagonal design of three, upon each section operating as an individual cell. The small cells, on the other hand, are considered to be omnidirectional and uni-sectored; subsequently, they work in open access mode as depicted in Fig. 1. The cells were linked together by an (X2) interface, to directly interact with one another and execute functions like mobility optimization, load management, handovers, and so on [17]. Accordingly, UE handovers among small cells are feasible. UEs, be able to travel freely across network cells the dual connectivity capability of UEs in this research is excluded. A centralized self-organizing network (SON) subsystem is used for network parameter estimation [18]. The network's cells are linked to the SON unit through the S1 interface [19]. To accomplish load balancing, the SON gathers the relevant load information, and the parameters of handover are improved.

2.2 Cell Load Measurements

Balancing the cells' loads will improve the performance of the wireless network by finding a measurement to calculate the cell loads (number of users, processing load,

transparent load, and radio load) accurately. To do so, we must know the used PRBs of a cell to find PRBU or load ratio RB, which is called the ratio of the used physical RBs of the cell to the total PRBs in a certain time T. The average resource block utilization ratio of a small cell i at time t, $\overline{L}_i^t$ can be expressed as:

$$\overline{L}_i^t = \frac{\sum_{t\in(t-T,t)} RB_i^t}{T.S_{\text{PRB}}} \tag{1}$$

where RB_i^t is the number of the allocated blocks and S_{PRB} is the total number of PRBs at duration time t.

For a small cell, when the average $\overline{p}_i^t$ is high, it directs to a high load and low unused resources. However, if $\overline{p}_i^t$ reaches 1, the cell has no available resources and the moved UEs to the cell will be dropped or experienced low throughput. In this paper, we proposed an algorithm to decrease the load of the overloaded cells by shifting UEs to light-loaded neighbor cells.

2.3 Formulation of the Problem

The main objective of this work is to make balance and take the advantage of the available PRBs of the neighbor cell for those overloaded serving cells which can be accomplished by shifting the edge UEs of the overloaded cells to the light-loaded cells by applying some measurements [20].

For a network of an S small cells at time t during the T period of time the average network load [4, 13, 14].

$$\overline{L}_{\text{net}}^t = \frac{\sum_{i\in s} \overline{L}_i^t}{S}. \tag{2}$$

The standard deviation of p in the network among all cells is

$$\sigma\left(\overline{p}^t\right) = \frac{\sqrt{\left(\overline{L}_i^t - \overline{L}_{net}^t\right)^2}}{S} \tag{3}$$

The expression of the problem indicates minimizing the average load ratio (PBRs)

$$0 \le RB_i^t \le S_{\text{PRB}}, \quad \forall i \in S, \ \forall t \in (t-T, t) \tag{4}$$

To move a load of an overloaded cell to a light-loaded cell and perform a balanced network, the handover parameters are adjusted adaptively and monitor the impact of the shifted load on the neighbor cell.

2.4 Load Threshold Calculation

MLB algorithm aims to reduce a load of overloaded cells below the pre-defined threshold to balance the network load. To achieve that we introduce a method to balance the network of small cell loads by categorizing the network into three levels: light load, medium load, and overload cells.

The light load network is considered when the threshold is lower than the pre-threshold (Ppre), while when the network threshold is equal or higher than pre-threshold, the network is considered as medium loads, and the network reached the threshold, P_{Thr} then it is considered an overloaded determined as:

$$P_{Thr} = \frac{1}{N} \sum_{i \in N} \overline{L}_i \tag{5}$$

where P_{pre} is a selected value that represents a medium load network.

3 Control Parameters

3.1 Handover Parameters for Load Balancing

In LTE, eight events were specified to report the measurements, six of them for the measurements of intra-LTE, and two for inter-RAT measurements (B1 and B2). In our paper, we will consider load balancing for intra-LTE.

UEs trigger (A1, A2, A4, and A5) events to compare the events with a given threshold. A1event is triggered when the signal quality of a serving cell became greater than a specified threshold. UE triggers the A2 event when serving cell signal quality becomes worse than the pre-defined threshold. A4 event is activated when the signal quality of the neighbor cell became better than a given threshold. UE triggers the A5 event when the neighbor cell signal quality became better than that given threshold. However, (A3 and A6) events are triggered when the real-time signal quality is compared among cells. UE triggers the A3 event under carrier aggregation when the offset of the neighbor cell is better than the serving cell's offset. A6 event under carrier aggregation when the offset of the neighbor cell is better than secondary cells. Based on the UE demands, primary cells add or remove the secondary cells when considering carrier aggregation. A3 is used to find the best neighbor cell and triggers handover if the criteria of A3 triggering satisfy longer than the time to trigger (TTT). A3 event reporting measurement is:

$$\text{Mn} + \text{Ofn} + \text{Ocn} - \text{Hys} > \text{Mp} + \text{Ofp} + \text{Ocp} + \text{Off} \tag{6a}$$

where Mn and Mp are the average values of RSRP; Ofn and Ofp are the frequency-specific offsets; Ocn and Ocp are the target and serving cells individual offsets respectively, Hyst is hysteresis parameter, and Off is the offset of A3 between the main and neighbor cells. For inter-frequency, Ofn and Ofp are used so we forego them. Ocn, Ocp, and Off parameters are changed to shift the handover decision area. The load state of source and neighbor cells is playing a big role in delaying or hastening the handover. By increasing the neighbor cell Ocn value or decreasing the serving cell Ocp, the range of the cell is reduced and UE is handed over to the neighbor cell.

3.2 Information of Serving Cells Edge—UEs

SON is used to achieve load balance in the network, by shifting the UEs from the overloaded serving cell to under-loaded neighbor cells. The UEs of overloaded serving cells are being nominee to be shifted to under-loaded cells, UEs trigger an A4 event when the neighbor cell's RSRP is better than a given threshold. Each cell of the network collects the information from its UEs to share that information with SON including the A4 event.

Triggered A4 event expressed as:

$$\mathrm{Mn} + \mathrm{Ofn} + \mathrm{Ocn} - \mathrm{Hys} > \mathrm{Threshold} \tag{6b}$$

When A4 is accomplished, the measurements will be reported by the UEs. The small cell provides the information of edge UEs (when a reasonable threshold value is set and only UEs that report measurements are selected), and the candidate UEs are listed to be shifted from full loaded serving cell to the light-loaded neighbor cell as a result of network load balancing.

3.3 SNR Estimation

The PRBs of UEs in LTE are allocated for serving cells based on channel quality indicators (CQI) that report all UEs. To achieve load balance, the required PRBs of UEs needed to be found and estimated at the neighbor cell before HO. Unfortunately, in LTE networks the information of CQI of the neighbor cells is unsupported, which forces us to predict the required PRBs by Shannon formula. We use RSRQ measurement reports of the neighbor cell information to estimate the SINR (θ) and stated as:

$$\theta = \frac{p}{I + N} \tag{7}$$

where P indicates serving cell signal power, I is the average interference, and N is the noise.

Based on [21] by using RSRP, P can be formulated as:

$$P = x.12.S_{\text{PRB}}.\text{RSRP} \tag{8}$$

x Is pre antenna subcarrier factor.
RSSI based on [22],

$$\text{RSSI} = P + I + N \tag{9}$$

And based on [23], RSRQ is formulated as:

$$\text{RSRP} = S_{\text{PRB}}.\frac{\text{RSRP}}{\text{RSSI}} \tag{10}$$

Therefore, from the above equations, SINR becomes:

$$\theta_i = \frac{1}{\frac{1}{12.\text{RSRQ}} + x} \tag{11}$$

4 Algorithms Proposed

In this portion, we present our proposed algorithms regarding load balancing, which is executed in the subset of SON. Table 1 clarifies the algorithm notations.

The load information is collected from all the cells to compute the threshold P_{Thr}, and to that end, we propose a method to compare P_{Thr} with a pre-defined threshold P_{pre} to identify the higher value of them and set as a threshold which given a name of efficiency parameter B.

$$B \rightarrow \max\left(P_{\text{Thr}}, P_{\text{pre}}\right) \tag{12}$$

B function is to control the algorithm triggering and categorizing the light, medium, and overloaded cells by selecting the values of P_{Thr} and P_{pre}. For instance, if the P_{Thr} is lower than P_{pre}, the network load is under-loaded so the threshold becomes P_{pre} and the algorithm is triggered for the medium-loaded cells that reach P_{pre}. When P_{Thr} is higher than P_{pre} then the network cells are considered as over-loaded. The load balancing is created for the medium and overloaded cells. The average load of overloaded small cells over the value of B triggers the algorithm until the network becomes more balanced.

Table 1 Flowchart summary notation

Notations	Description
$\overline{L_i}$	Average resource block utilization ratio of a small cell i at time t
P_{Thr}	Threshold load of network
P_{pre}	The selected value represents the network medium load
B	Efficiency parameter
c_0	Cluster cells to perform load balancing with
$\overline{p}c_0$	Average RBUR of a cell in a cluster
mth	Maximum tires of neighboring to form a cluster
L	Cluster cells
$\overline{L}_{(i,u)}$	Movable load of an edge cell i
$\widetilde{L}_i$	Estimated cell load after the handover of the UE u
$\tilde{L}_{(i,u)}$	Serving cell i consumed load by UE u
$\tilde{L}_{(j,u)}$	Estimated load of a UE u at target cell j
R_f	Rescue factor
$\text{CIO}_{(i\to j)}$	Specific offset parameter of cell i for neighbor cell j

4.1 Cluster Forming

Formation of cluster idea is inspired from [15], it begins with collecting information of edge users and the suitable (best) neighbor cells for them as explained in Sect. 3.2, the movable UEs are considered by information of edge UEs, and the system is demonstrated as described in [15] by directed multigraph which is, in brief, indicated that the cell can join the cluster if it has an incoming edge for other neighbor cells.

The SON starts to gather the cluster information after generating such a graph. For that purpose, our proposed algorithm gathers all cells' load status in the graph and calculates the P_{Thr} using (5), then it discovers overloaded cells using (12) if their load is higher than the efficiency parameter (B). A new cluster is formed to include those overloaded cells is formed. After that, it selects the overloaded cells for the neighbors as first-tier, gradually, the cells are added to clusters, but only if the contiguous cells share as a minimum one incoming edge at least from any of the cluster members and it is not included into another cluster. Once the algorithm adds a contiguous cell to a particular cluster, it labels it as a cell cluster member. This method is repeated until the neighbor cells are complete. If the list of neighbors is complete and $\overline{P}C0$ remained greater than B, the process involves the cluster follow-tier of neighbors until the maximum allowable adjacent tier, $m_{\max}$, is fulfilled.

Clusters are produced for the overloaded cells in the graph using the procedure described above. The cluster evaluated load balancing is depicted in Fig. 2c, which was built from the network depicted in Fig. 2a. The technique constructs a cluster with Cell 1 as a member for the overloaded Cell 1. The program then takes Cells 2

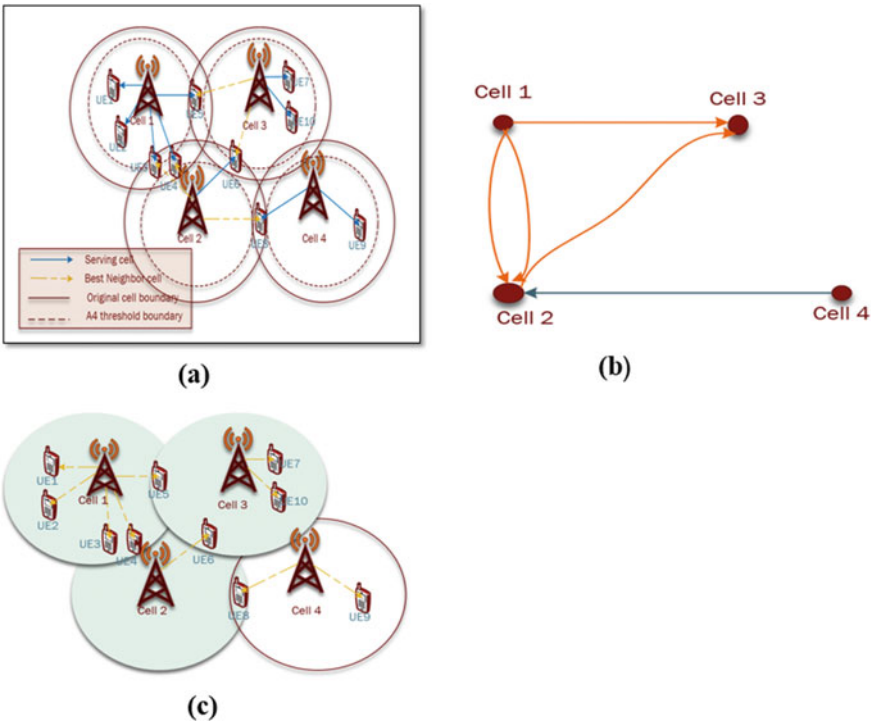


Fig. 2 Construction of directed multigraph network **a** network with 4 small cells and 10 UEs **b** modeled graph of the network **c** cluster for load balance

and 3 and adds them as well. However, because Cell 4 does not share any incoming edges from the other members, then it is excluded from the cluster. The flowchart of algorithm 1 is illustrated in Fig. 3.

4.2 Load Balancing

Shifting loads of medium/overloaded serving cells to light/medium neighbor cells is done using the proposed algorithms by adjusting the cell CIO parameters to detect the medium/overloaded cells of the cluster. In that end, any edge UEs of overloaded cells share the coverage with the neighbor cells report RSRPs of neighbor cells and then put in descending order, the edge UEs of the overloaded cells are transferred one after another so that the serving cell load becomes lower than or equal to B. However, the movable loads are estimated for serving and neighbor cells before the handover decision. In the serving cell, PRBs are used to calculate the accumulative load of the edge UEs. The movable load of an edge UE u in a serving cell i, $\overline{L}_{(i,u)}$ is determined as:

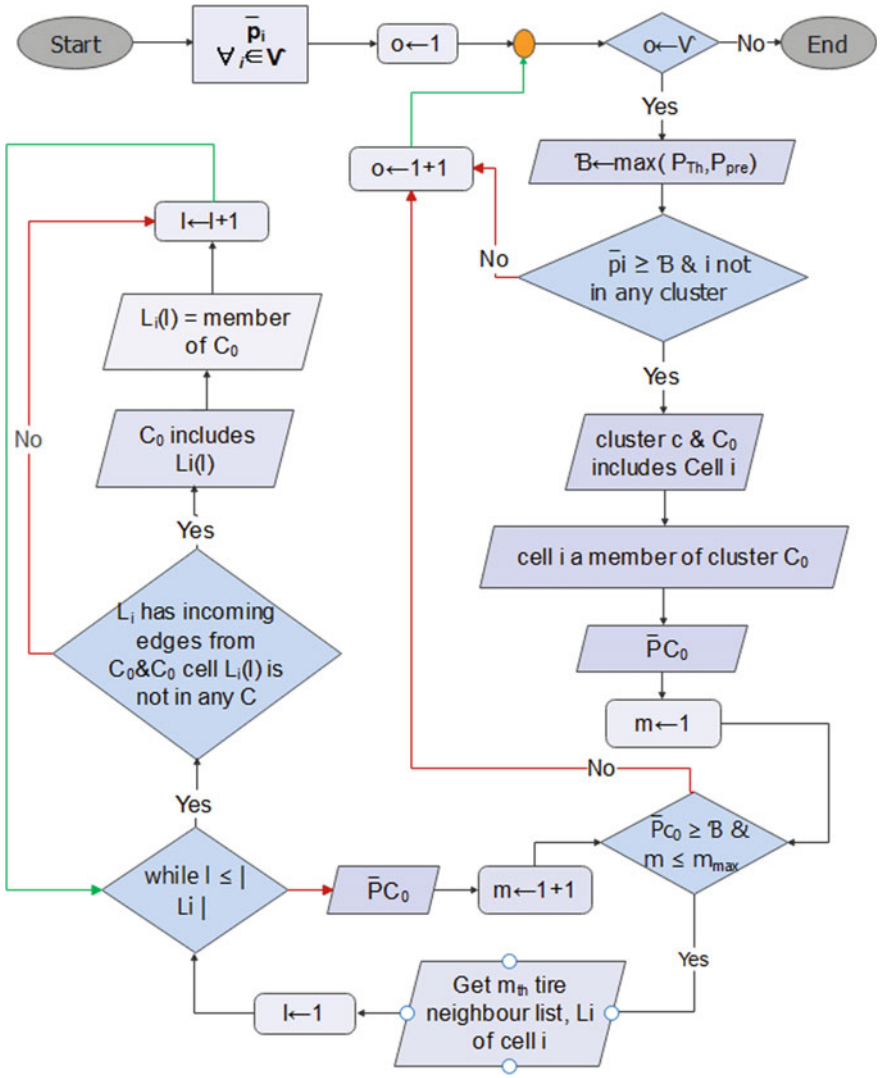


Fig. 3 Cluster forming flowchart

$$\overline{L}_{(i,u)} = \frac{\sum_{t\in(t-T,t)} RB^t_{(i,u)}}{T.\text{SPRB}} \tag{13}$$

where $RB^t_{(i,u)}$ is PRBs that are allocated by overloaded cell i to UE u.

The estimated cell load of a serving cell after handover of the UE u, $\widetilde{L}_i$ is calculated as

$$\widetilde{L_i} = \overline{L_i} - \overline{L}_{(i,u)} \tag{14}$$

The neighbor cell signal quality would be different for the shifted UE; therefore, the needed PRB will be different as well.

The estimation load of the moved (shifted) load to a target (neighbor) cell j becomes

$$\widetilde{L_j} = \overline{L_j} + \tilde{L}_{(j,u)} \tag{15}$$

where

$$\tilde{L}_{(j,u)} = \frac{RB_{(j,u)}}{1000.N_{\text{PRB}}} \tag{16}$$

If $\widetilde{L_j}$ is lower than B, the serving and target CIO values are informed by the algorithm so the candidate edge UE could be moved to the target neighbor cell as:

$$\text{CIO}(i \rightarrow j) = \text{Mp} - \text{Mn} + \text{Hyst} + \Delta \tag{17}$$

While 1

$$\text{CIO}(j \rightarrow i) = -\text{CIO}(i \rightarrow j) \tag{18}$$

Δ is a specified LTE cemented step [18].

If $\widetilde{L_j}$ is higher than B, we proposed a factor that transfers partition of the load that makes $\widetilde{L_j}$ lower than or equal to B which is named rescue factor R_f.

$$R_f = \frac{B_-}{\overline{L}_{(i,u)}} \widetilde{L_j} \tag{19}$$

$$\widetilde{L_{j0}} = \overline{L_j} + R_f \widetilde{L}_{(j,u)} \tag{20}$$

Figures 4 and 5 represent cluster forming, load balancing, and unloading process flowcharts, respectively.

5 Performance Assessment

5.1 Environment of Simulation

To investigate the proposed algorithm, we introduce the network of four macrocells and ten small cells, the users have been set as 200 UEs as in Fig. 6a, and the network

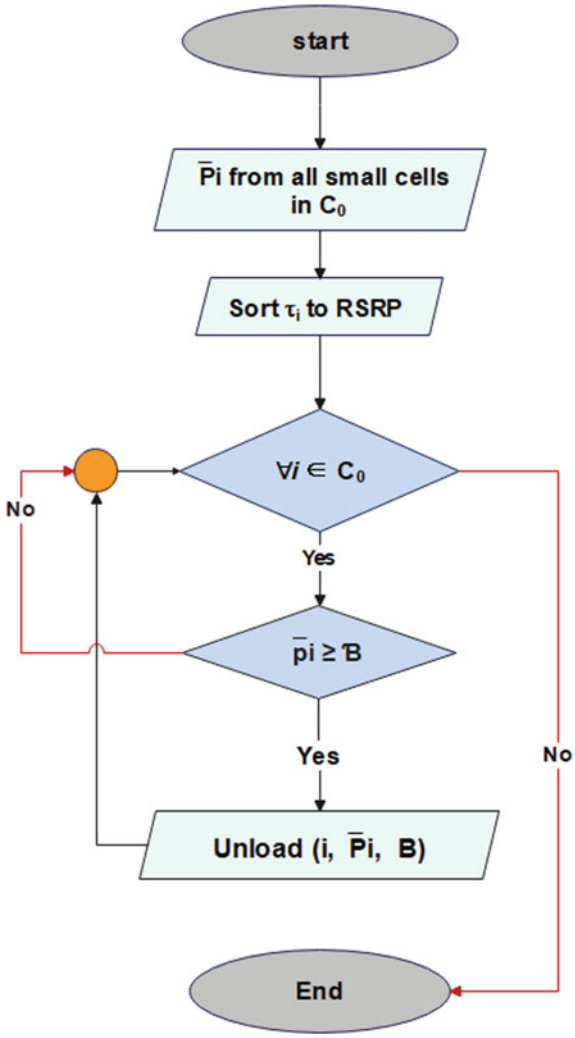


Fig. 4 Load-balancing flowchart

has been performed with small cells of a bandwidth of 20 MHz which enables 100 PRBs. The network *Tx* power was modeled to 24 dBm and propagation loss as non-line of sight. Channel and QoS aware (CQA) scheduler was used to allot resources to UEs. The pre-threshold was set to 0.35, and the remaining parameters are illustrated in Table 2.

50% of static and non-uniform UEs were spread over the network and 50% in a random circular mechanism (RCM).

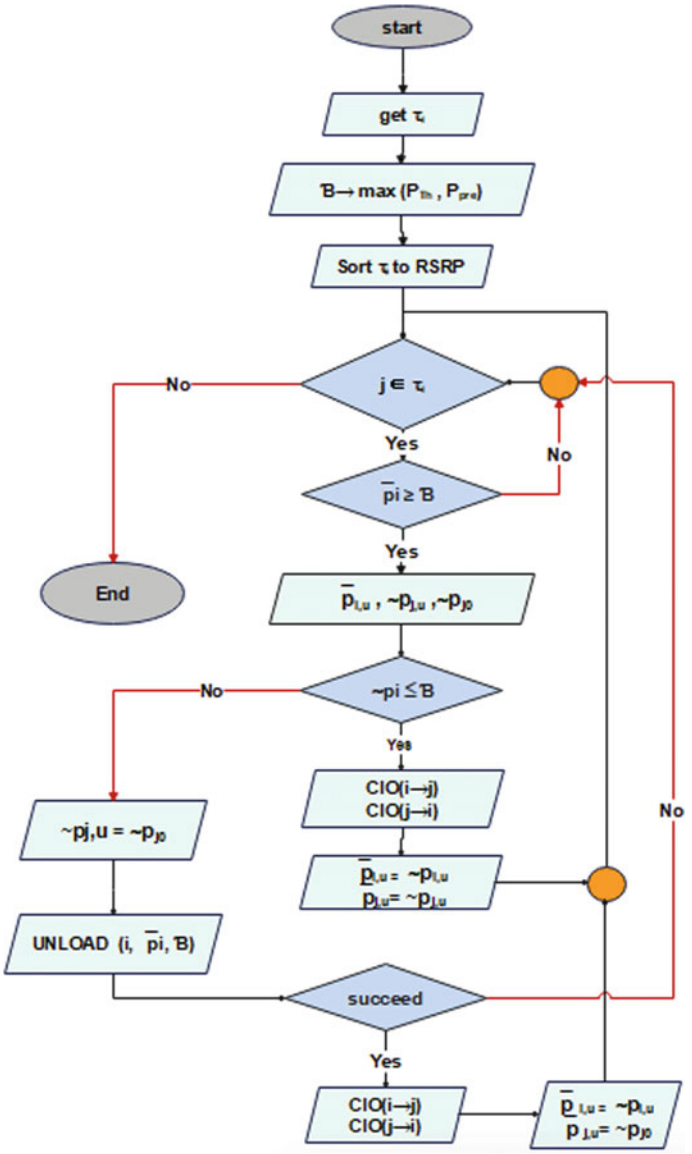


Fig. 5 Unloading process flowchart

6 Assessment of Performance

We investigated two major features to evaluate the suggested method. The first consideration is load distribution across cells. We used the standard deviation of RBURs among cells to calculate load distribution. The second consideration is network

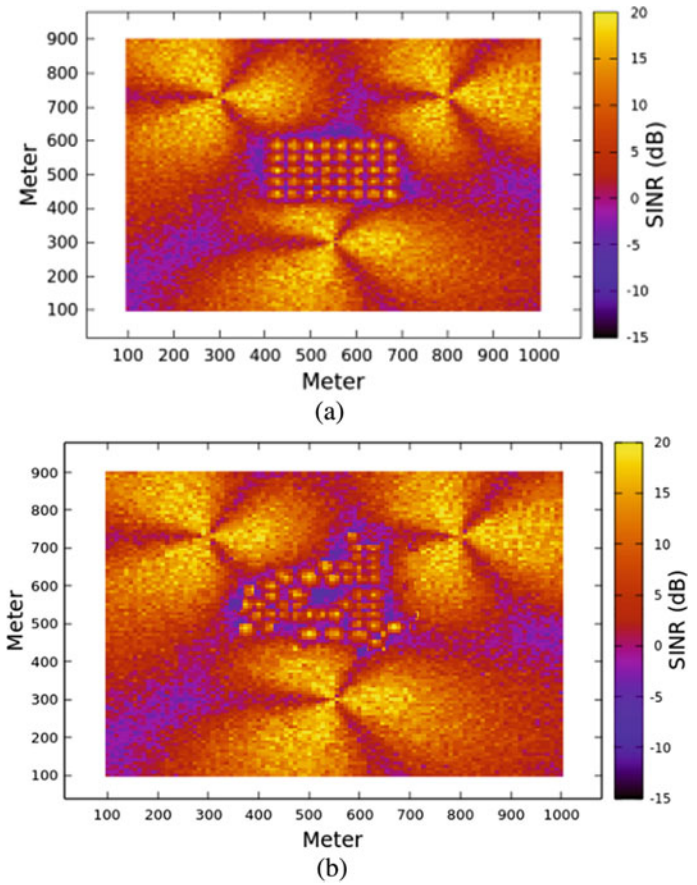


Fig. 6 **a** Environment of simulation, **b** radio atmosphere map

throughput. A load-balancing algorithm may balance network load by putting big CIO values in the unsuitable cells, resulting in network performance degradation. As a result, we investigated the impact of using MLB methods on network throughput.

$$\text{Throughput} = \sum_{i=1}^{u} \text{DR}_i.$$

where DR_i is the average received data rate by UE_i.

We compared the proposed algorithm with two earlier algorithm approaches for performance validation: the enhanced mobility load-balancing algorithm [14] and the clustered mobility load-balancing algorithm [15]. For clarity, we refer to [15] as proposed MLB, [14] as enhanced MLB, and the simulations not using MLB algorithm as No MLB in the figures.

Table 2 Simulation parameters

Parameters	Values	
	Macrocell	Small cell
Cells number	3	40
Cell layout	Hexagonal grid	Manhattan layout
Cell high	25 m	5 m
Tx power	46 dBi	24 dBi
Max. antenna	10 dBi	5 dBi
Antenna mode	Directional	Isotropic
Inter-site distance	500 m	35 m
System bandwidth	20 MHz	20 MHz
Path loss	$PL = 147.4 + 43.3 \log_{10}(D)$	
Fading	Standard deviation 4 dB, log normal	
Noise	5 dB	
Resource scheduling	CQA scheduler	
P_{pre}	0.38	
Hysteresis	2 dB	
TTT	256 ms	
User *Tx* power	2 dBi	
User antenna high	1 m	
Traffic model	2 Mbps at a guaranteed bit rate GBR	

A. Effect on Load Distribution

To investigate the effect of the proposed method on load distribution among cells, the three network scenarios were simulated for 10 min of real-time network operation. The standard deviation among the cells using three distinct MLB techniques in the scenarios is shown in Fig. 7. Since No MLB is taken into account, the overloaded cells do not deliver sufficient resources to the UEs due to a shortage of resources; however, the low-loaded cells are under-used. As a consequence, the standard deviation is increased in the figure once the network excludes the MLB algorithm. MLB algorithms, on the other hand, divide the load over the cell; as a result, the RBUR SD accumulates when the MLB is used.

As soon as the UEs' speed is low, the SD converges gradually and is constant. However, the higher speed of the UE the higher oscillations in SD. Because high-speed UEs regularly modify network topology, gigantic handover numbers cause fluctuated loads across cells. As an outcome, the load-balancing approach nominates fewer UEs to be shifted, leaving load balancing more difficult for MLB algorithms. Regardless of different user mobility, data discloses that the proposed method reduces the RBUR standard deviation among cells more than existing algorithms. Because the suggested approach takes into account when exactly to make load balancing by categorizing the network loads into levels and triggers the process of unloading when the network has a medium load to avoid the complexity calculations.

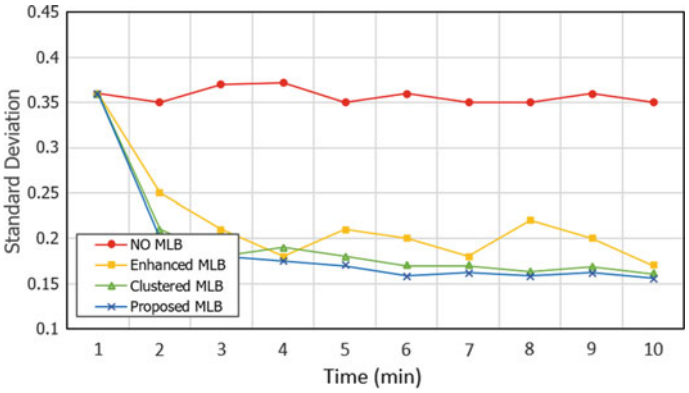


Fig. 7 RBUR cells standard deviation

B. Throughput effect

Figure 8 depicts a network throughput comparison using several MLB algorithms with three different UE mobility scenarios. Due to a shortage of available resources in the overloaded cells, they are unable to distribute the resources needed to the UEs, and the data rate of UE in such overloaded cells is compromised. The MLB algorithms assist overloaded cells to distribute the needed resources to UEs to sustain the data rates by equally distributing the load among the cells. As a result, when MLB techniques are used in the network, the throughput increases. Our proposed algorithm outperformed the existing MLB techniques because it distributes the load

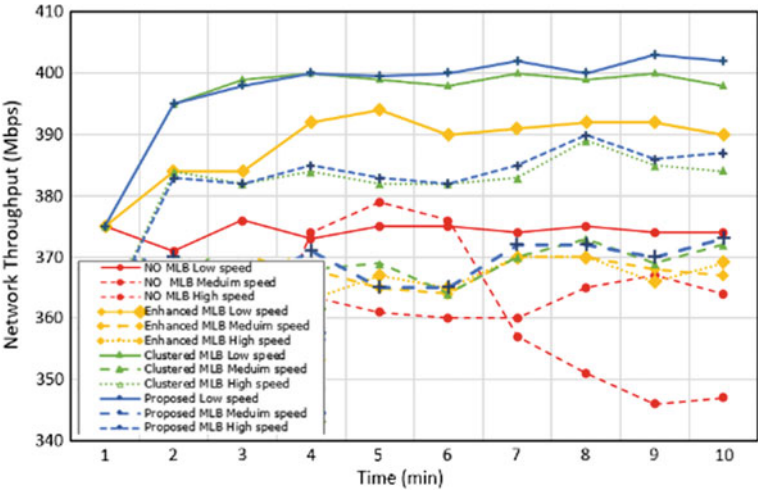


Fig. 8 Network throughput of multi speed scenarios

more equally throughout the cells by moving relevant edge UEs and the resources allocated to the UEs most effectively.

Figure 8 demonstrates that in low speeds of UE, our algorithm boosted the average network throughput by 7.355%. Due to the frequent changes in network structure, the network throughput at higher speeds decreased compared to lower speeds. However, in the three mobility scenarios, our proposed technique provided the maximum network performance. As a consequence, simulation results from many areas show that the proposed algorithm achieves more even load balancing. Hence, the throughput of the network is increased.

C. Influence of tiers on MLB

We investigated the impact of neighboring tiers on load balancing, and the network situation with low-velocity UEs is tested using the suggested approach in Algorithm 1 while taking four values as $m_{\max}$. Figure 9 compares the performance of standard deviation across the cells, and it shows that providing more adjacent tiers improves the performance of the suggested load-balancing algorithm since so many neighbors enhance the chance for load balancing. Nonetheless, the suggested four-tiered algorithm's performance gets saturation. Nonetheless, the suggested four-tiered algorithm's performance gets saturated. Figure 10 depicts the achieved throughput of network simulations. The network throughput improves as the number of surrounding layers grows in load balancing. However, there is a slight difference in throughput between the situations of third and fourth tiers.

D. Effect on real network

Small cells in reality are generally arranged casually. Hence, to demonstrate the stability of our algorithm, we investigated its effect on a network with randomly placed small cells. Small cell deployment data came from a scenario in [15]. For the aim of applying a heterogeneous network, we evaluated 40 small cells and installed

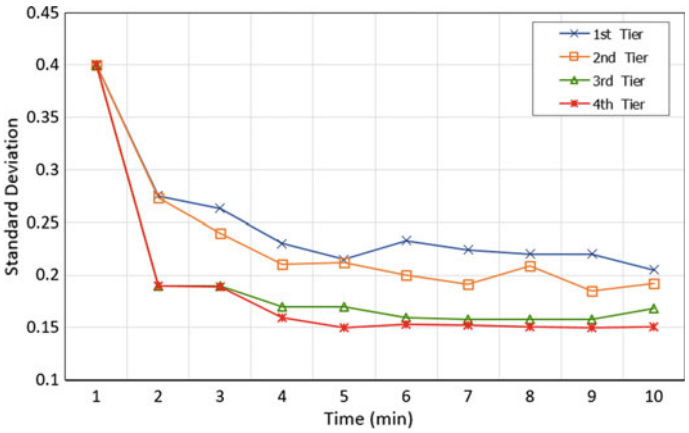


Fig. 9 Network tiers RBUR standard deviation

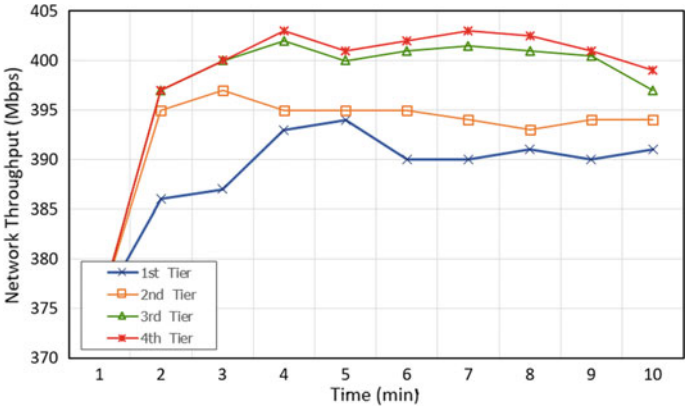


Fig. 10 Network tiers throughput

three macrocells over the networked system. Figure 6b depicts the radio atmosphere map of the network under consideration. In all, two hundred UEs are deployed in the network. Originally, 75% of them were installed within the small cell transmission range, and the remains were distributed randomly across the network. Automobile UEs and pedestrian UEs were modeled for UE mobility. The speeds of the pedestrian UEs were chosen at random between five and ten km/h. Furthermore, for automobile users, we fixed the speeds at random around ten km/h and thirty km/h. beyond the small cell network, UEs were represented as random walked UEs with a random speed set up between five to ten km/h.

The SD analysis of the MLB methods in the system with uneven small cell deployment is shown in Fig. 11.

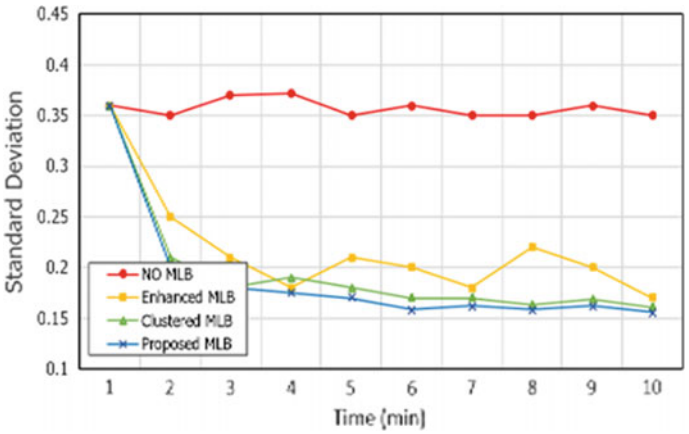


Fig. 11 Network standard deviation

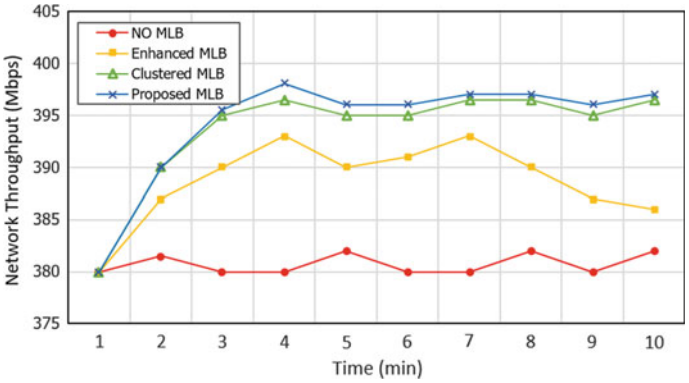


Fig. 12 Overall network throughput

The picture shows that, despite varying cell sizes, random small cell installation, and specified user mobility, our suggested algorithm performs better than the previous MLB algorithms. The overall performance in terms of the average throughput is depicted in Fig. 12. The graph shows that the proposed technique outperformed the other algorithms.

7 Conclusion

In this work, we presented an ultra-dense heterogeneous networks mobility load-balancing clustering algorithm that first, categorizes the load status of a small cell as under-loaded, light-loaded (medium), or overloaded using the proposed efficiency parameter which considers the maximum value between the pre-threshold and adaptive threshold. In the case of shifting loads from overloaded/light-loaded cells to light-loaded/under-loaded cells, the proposed rescue factor estimates the load statuses of the target cell after handover to assure its load is less than the adaptive threshold. The system simulation of the proposed algorithm showed a reduction of standard deviation by 19.99% and increased the throughput by 7.355%, and PRB utilization is increased as compared to No MLB.

References

1. Cisco visual networking index: global mobile data traffic forecast update 2017–2022. Cisco, San Jose, CA, USA, vol 1486680503328360
2. Al Amodi A, Datta A (2020) The impact of heterogenous ultra-dense network technologies on the performance of 4G and 5G networks. Int J Inno Technol Exp Eng 10:35–44
3. Andrews JG, Claussen H, Dohler M, Rangan S, Reed MC (2012) Femtocells: past, present, and future. IEEE J Sel Areas Commun 30:497–508

4. Addali KM, Melhem SYB, Khamayseh Y, Zhang Z, Kadoch M (2019) Dynamic mobility load balancing for 5G small-cell networks based on utility functions. IEEE Access 7:126998–127011
5. Ge X, Tu S, Mao G, Wang C-X, Han T (2016) 5G ultra-dense cellular networks. IEEE Wirel Commun 23:72–79
6. Feng S, Seidel E (2008) Self-organizing networks (SON) in 3GPP long term evolution. Nomor Research GmbH, Munich, Germany, vol 20
7. Kwan R, Arnott R, Paterson R, Trivisonno R, Kubota M (2010) On mobility load balancing for LTE systems. In: 2010 IEEE 72nd vehicular technology conference-fall, pp 1–5
8. Lobinger A, Stefanski S, Jansen T, Balan I (2010) Load balancing in downlink LTE self-optimizing networks. In: 2010 IEEE 71st vehicular technology conference, pp 1–5
9. Zia N, Mitschele-Thiel A (2013) Self-organized neighborhood mobility load balancing for LTE networks. In: 2013 IFIP Wireless Days (WD), pp 1–6
10. Damnjanovic A, Montojo J, Wei Y, Ji T, Luo T, Vajapeyam M (2011) A survey on 3gpp heterogeneous networks. Wireless Comm. IEEE
11. Lobinger A, Stefanski S, Jansen T, Balan I (2011) Coordinating handover parameter optimization and load balancing in LTE self-optimizing networks. In: 2011 IEEE 73rd vehicular technology conference (VTC Spring), pp 1–5
12. Huang Z, Liu J, Shen Q, Wu J, Gan X (2015) A threshold-based multi-traffic load balance mechanism in LTE-A networks. In: 2015 IEEE wireless communications and networking conference (WCNC), pp 1273–1278
13. Hasan MM, Kwon S, Na J-H (2018) Adaptive mobility load balancing algorithm for LTE small-cell networks. IEEE Trans Wireless Commun 17:2205–2217
14. Addali K, Kadoch M (2019) Enhanced mobility load balancing algorithm for 5G small cell networks, In: 2019 IEEE Canadian conference of electrical and computer engineering (CCECE), pp 1–5
15. Hasan MM, Kwon S (2019) Cluster-based load balancing algorithm for ultra-dense heterogeneous networks. IEEE Access 8:2153–2162
16. Hasan MM, Kwon S, Oh S (2018) Frequent-handover mitigation in ultra-dense heterogeneous networks. IEEE Trans Veh Technol 68:1035–1040
17. 3GPP (2016) Evolved universal terrestrial radio access network (E-UTRAN); X2 application protocol (X2AP). In: 3rd generation partnership project (3GPP), vol TS 36.423
18. Feng S, Seidel E (2008) Self-organizing networks (SON) in 3GPP long term evolution. Newsletter, Nomor Research GmbH, Munich, Germany, vol. Tech. Rep
19. 3GPP (2016) Evolved universal terrestrial radio access network (E-UTRAN); S1 application protocol (S1AP). In: 3rd generation partnership project (3GPP), vol TS 36.413
20. Bian YQ, Rao D (2014) Small cells big opportunities. Global Business Consulting. Huawei Technologies Co., Ltd
21. 3GPP (2017) LTE; evolved universal terrestrial radio access (E-UTRA); physical layer; measurements, document. 3GPP TS 36.214
22. Afroz F, Subramanian R, Heidary R, Sandrasegaran K, Ahmed S (2015) SINR, RSRP, RSSI and RSRQ measurements in long term evolution networks. Int J Wirel Mob Netw 7:113–123
23. Afroz F, Subramanian R, Heidary R, Sandrasegaran K, Ahmed S (2015) SINR, RSRP, RSSI and RSRQ measurements in long term evolution networks. Int J Wirel Mob Netw

An Efficient Machine Learning Classifier for Sarcasm Detection

P. Keerthika, R. Manjula Devi, P. Suresh, K. K. Indiraa, P. V. Jayasri, and N. Kishore

Abstract Irony is a sarcastic term that is used to actively criticize people. Irony is frequently expressed through indirect phrases. Politics, sports, business, and social media are the different places where ironic statements can be found. Some people can perceive irony in words, while others are clueless about how others have expressed themselves. To recognize these elements of speech, various machine learning and deep learning algorithms are applied. The main purpose of this study is to propose a broad summary of the ironic and non-ironic words found in subreddit data were used to categorize the text comments. This manuscript describes about the well-known datasets that are used to create the ironic detection networks by providing a major focus on classifying the comments. The considered dataset is partitioned to make the task easier. Finally, the process used for identifying irony phrases and the most successful classifiers and evaluation criteria are also discussed.

Keywords Sarcasm · Machine learning · Partitioning · Supervised pattern · Classification · Deep learning

P. Keerthika (✉) · P. Suresh
Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India
e-mail: keerthikame@gmail.com

R. M. Devi · K. K. Indiraa · P. V. Jayasri · N. Kishore
Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, Erode, Tamil Nadu, India
e-mail: indiraakk.18cse@kongu.edu

P. V. Jayasri
e-mail: jayasripv.18cse@kongu.edu

N. Kishore
e-mail: kishoren.18cse@kongu.edu

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_6

1 Introduction

Sarcasm is a feeling that people share although the precise term does not correspond to the underlying meaning. It is easy to comprehend and evaluate sarcastic statements made in front of others, but determining the true meaning of the context digitally is more difficult. Since social media is a free forum for everyone to express true emotions of people, they feel and write their opinions in the best possible way [1, 2]. For instance: In this tweet: "There's nothing like a crowded library with no empty seats to make me happy," even if no harsh terms are used in this statement; the digitally accurately designed sarcasm analyzer records the person's feelings, which will convey the sentence's exact meaning [3, 4]. People employ positive emotion (love) in this example; however, the entire tweet indicate negative views about the library.

According to the Random House dictionary, sarcasm is the act of speaking or writing anything that is the polar opposite of what you intend. The Macmillan English dictionary states that sarcasm is the act of expressing or writing anything that is the polar opposite of what you mean, a strongly ironical insult or of communicating in a way that is intended to make another person feel inadequate or demonstrate that you are upset. According to the Collins English dictionary, mocking, contemptuous, or sarcastic phrases are used to express insults or disdain.

This survey gives clear insights into many sarcasm detection algorithms proposed by various researchers. The remaining part of the paper will proceed as follows:

The following section is about to discuss the different types of sarcasm with a detailed literature review. Further, a research discussion has been initiated on feature set analysis. A variety of methods for detecting sarcasm are also addressed. In a table, certain study articles are contrasted. Finally, the proposed study is concluded with some final thoughts and possible future research directions.

Literature Survey

Many researchers have studied sarcasm's attributes, such as theories of sarcasm, syntactic characteristics, sarcasm's psycholinguistic aspects, lexical features, semantic properties, and so on [5, 6]. As a result, the goal of this study is to present the most prevalent strategy to detect sarcasm, which has gone as far as semantic highlights as well as the data analysis methods that applies a machine learning method in the preceding 15 years.

The research study proposed by Deep et al. [7] aims to automate the process of classifying sarcastic tweets into four categories: positive, negative, self-deprecating, and miscellaneous by using three types of features: lexical, pragmatic, and linguistic. For multiclass classification, four machine learning classifiers were used: logistic regression, Naive Bayes, random forest, and support vector machine. The results show that the binary and multiclass sarcasm detection may reach accuracy up to 89% and 39%, respectively.

According to Pawar and Bhingarkar [8], sentiment analysis refers to the detection and aggregation of expressed attitudes and opinions of Internet users in a certain

community. This research proposes an example-based strategy for detecting mockery using Twitter data. Mocking and non-snide tweets are differentiated using four sets of principles that cover a wide spectrum of mockery. The proposed method with different sets as well as supplementary expenditure classes is examined.

Most of the previous research study uses text data to identify sarcasm, according to Das and Clark [9]; however, a few recent studies propose using visual data. No research study has attempted to analyze the user interaction patterns as a source of sarcasm detection. This study describes about the supervised machine learning-based method that considers both the content of postings (e.g., text, picture) and how people interact with such posts on Facebook.

We investigate a new pattern-based methodology for detecting sarcasm, as well as a behavioral modeling strategy for successful sarcasm identification by analyzing the content of tweets while considering the activity attributes of people based on their prior actions, according to Deshmukh and Solanke [10]. Random forest, support vector machine (SVM), k-nearest neighbors (KNN), and maximum entropy were among the classifiers tested.

Ashwitha et al. [11] concluded that the observable element overcomes any issues, but it becomes hard to recognize its exacting and planned implications. Recognizing the text from the Internet-based informal communities like Facebook, Twitter, Instagram, studies, and different spots has become critical as it influences social and individual collaborations. Sarcasm detection is also remaining as a significant challenges for natural language processing [NLP] models.

The research study conducted by Bharti et al. [12] reveals that different methods are developed to extract the text aspects such as lexical, hyperbolic, behavioral, and universal facts. The machine learning classifiers that are considered in this work include support vector machines, Naive Bayes, maximum entropy, and decision trees. Finally, this research study has used an extracted feature set to train the aforementioned classifiers to detect sarcasm in Twitter data. In comparison to existing methodologies, this work has achieved a significant increase in accuracy.

Katyayan and Joshi [13] have proposed a sentiment analysis (opinion mining) attempts to enhance a computer's accuracy and enable the computers to use natural language processing (NLP) methods to discern human emotions in many types of online writing. Several attempts have been made by researchers to develop interpretable robots for recognizing sarcasm.

Haripriya and Patil [14] have overcome any barrier between the exact and expected outcome. Mockery recognition is a significant challenge in natural language processing (NLP) domain. Moreover, the mockery/sarcasm detection is required for providing better comprehension to fill in a point of interaction for providing a shared correspondence among machines and people.

Product reviews on e-commerce sites have become increasingly important in deciding which items to purchase. However, the conclusions of the product reviews are unreliable due to the extensive sarcastic remarks. By combining the methodologies of sentiment analysis and fuzzy, Felbo et al. [15] have proposed a novel strategy in the field of sarcasm identification by comparing the apparently sarcastic thoughts of reviews from other consumers of similar products. According to the conclusions

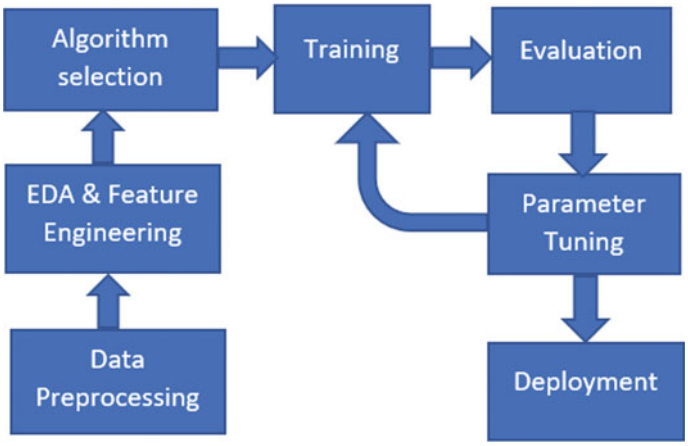


Fig. 1 Logistic regression is depicted as a block diagram

of this study, the greater the gap between positive and negative ratings, the more negative the review appears.

2 Existing Work

2.1 Logistic Regression

Logistic regression is a method used for classifying data in machine learning. Researchers commonly use logistic regression whenever it is required to group data into two or more classes. Binary logistic regression is the first, while multiclass logistic regression is the second. The binary class has two classes, as the name implies: Yes/No, True/False, 0/1, and so on. In multiclass classification, there are more than two classes for classifying data. A classification strategy for categorical variables such as Yes/No, True/False, 0/1, and so on is logistic regression (Fig. 1).

2.2 Support Vector Machine

SVM (supervised machine learning) is a type of machine learning that can be used to address both classification and regression challenges. However, it is mostly employed to solve categorization issues. Every piece of information is represented as a point in *n*-layered space (where *n* is the number of highlights), with the value of each element acting as the incentive for a certain location in the SVM calculation. It accomplishes this by identifying the hyperplane that clearly separates the two groups. Support

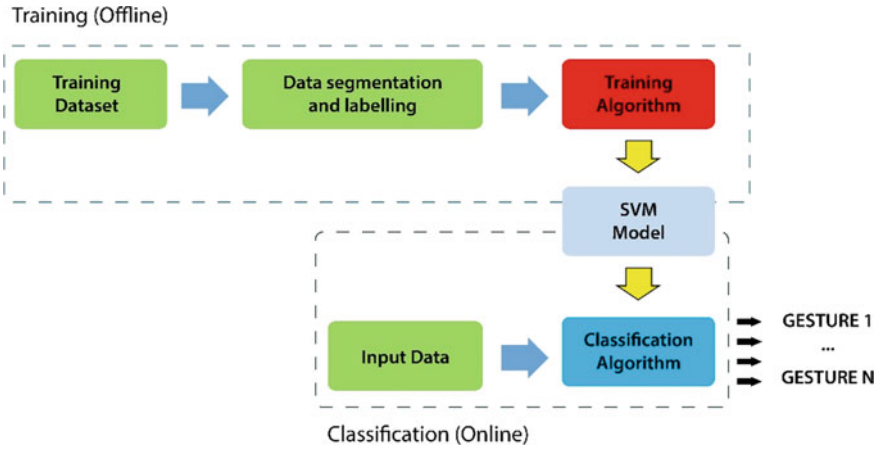


Fig. 2 Block diagram of SVM

vectors are simply the directions in which each individual perceives the world. The SVM classifier is an effective approach to separate the two classes (hyperplane or line).

The main goal of the support vector machine approach is to create a decision boundary for dividing the *n*-dimensional space into different categories, which can be readily transferred to a new data center in the ideal class in the near future. The optimum decision boundary may be considered as the hyperplane. Other applications of this support vector machine technology include text classification, image categorization, and face recognition (Fig. 2).

2.3 Dataset Description

Totally, 1.3 million sarcastic comments from the online review site Reddit are included in the dataset. This dataset was created by extracting (sarcasm) tags from Reddit comments (not by me:) This tag is frequently used by Redditors to signal that their comments are sarcastic and not meant to be taken seriously, and it is a good predictor of sarcastic comment content. The training and testing set accounts for 80% and 20% of the data collected.

3 Proposed Work

As the name implies, Bayes denotes the Bayes theorem, whereas Naive denotes simplicity.

The following assumptions are made for the proposed algorithm:

1. The order in which the words appear makes no impact; however, the repetition of terms does.
2. Words in one document class have no bearing on the other.

Working:

Step 1: Extract raw text with or without the #sarcastic label. Cleaning, stop word removal, stemming, and lemmatization are all important preprocessing stages.
Step 2: Using scikit learn TF-IDF vectorisers, convert the text into TF-IDF vectors.
Step 3: Each document should be represented as a word document with a TF-IDF score.
Step 4: Find $P(X|Y = y)$ for all the documents that fall under a specific label and use it to create a probabilistic model.
Step 5: It has y, which is more likely to produce X.

The NB classifier is referred to as a probabilistic learning model. It is used to solve order problems and is based on Bayes' hypothesis. A major underlying assumption of this technique is the autonomy of the highlights or qualities connected with a forecast. This technique builds expectations about every event to a specified class by processing the class earlier likelihood, the probability to a given class, the back likelihood, and the indicator earlier likelihood.

$$P(H|E) = P(H) * P(E|H)P(E)$$

where

- $P(H|E)$ is the chance of class c given a predictor's posterior probability.
- $P(H)$ is the probability of *class*.
- $P(E|H)$ is the probability of the predictor given class, which is the likelihood.
- $P(E)$ is the prior probability of a predictor.

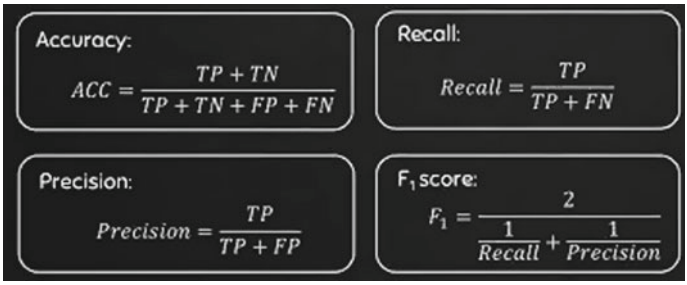
4 Result and Discussion

When evaluating the accuracy of these three machine learning algorithms for identifying irony in a dataset, it was observed that the Naive Bayes strategy outperforms the other two. The X-axis depicts the fraction of measures acquired, while the Y-axis depicts various situations under which distinct conclusions are obtained. The findings demonstrate that the Naive Bayes algorithm detects irony phrases/words in the given dataset with a success rate of 74%.

4.1 Performance Metrics

The cycle in which a prepared model is examined by using a testing informational index is referred to as model approval. The preparation set is established by using the testing informational index, which is a separate piece of comparable informative collection. The purpose of model approval is to assess the model's accuracy and execution in light of previous data for which we already have actual results.

Metrics considered for evaluation: accuracy, precision, recall, F1-score.



The following are the resultant graphs obtained under each considered metrics (Figs. 3, 4, 5 and 6).

For a more effective comparison, the aggregated findings have been structured into a table (Table 1).

Low false-positive rate is related to high accuracy. For the suggested task, we received an accuracy of 0.74. Recall (sensitivity) is the ratio of properly anticipated

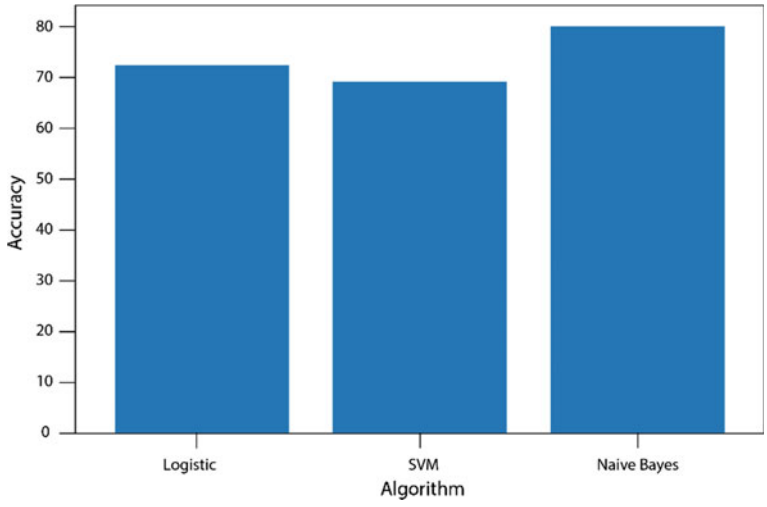


Fig. 3 Accuracy graph

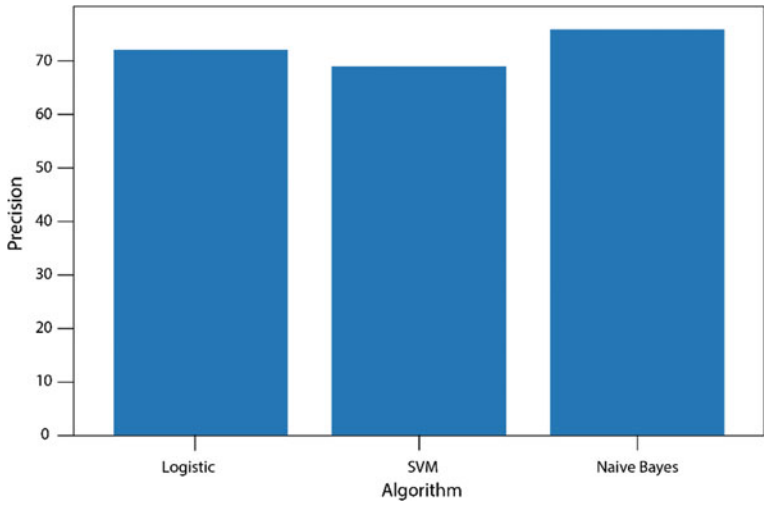


Fig. 4 Precision graph

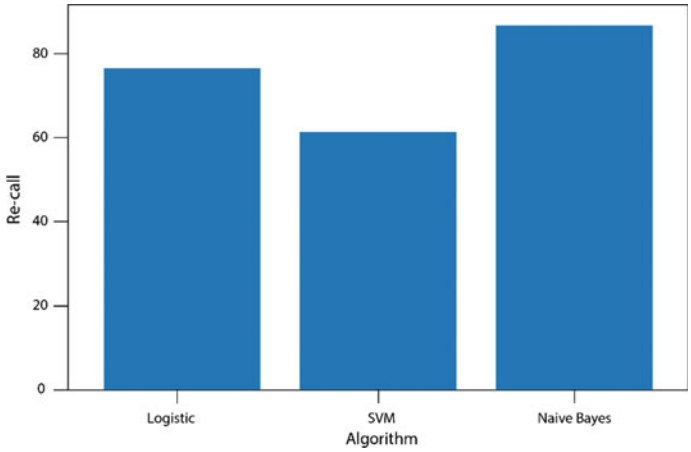


Fig. 5 Recall graph

positive feedback to all input in the actual class; moreover, recall is the ratio of correctly predicted positive feedback to all feedback in the actual class. When false positives and false negatives have the same cost, accuracy works best.

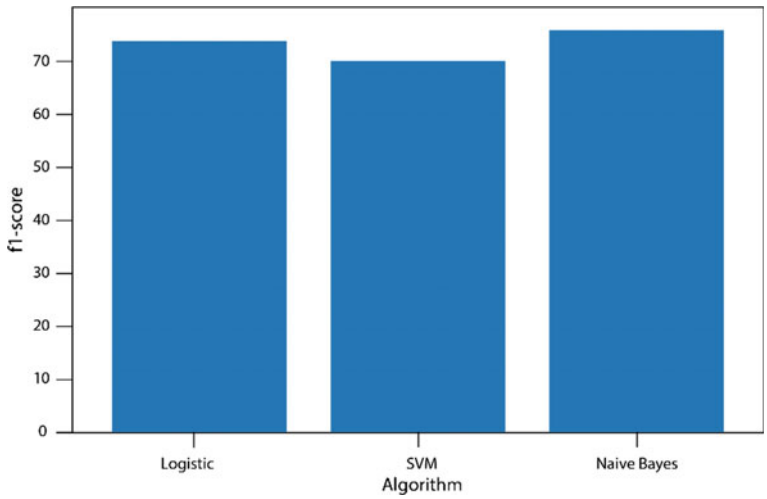


Fig. 6 F1-score graph

Table 1 Accuracy comparison table

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Logistic regression	65	70	72	72
Support vector machine	72	69	65	60
Naive Bayes	74	65	69	65

5 Conclusion and Future Work

Many preprocessing techniques are used to identify irony phrases/words in the dataset. The Reddit dataset was used to evaluate the logistic regression, support vector machine, and Naive Bayes techniques for detecting sarcasm, wherein the Naive Bayes strategy is proving to be the ideal option. Algorithms including logistic regression and SVM were used on the Reddit dataset, with the Naive Bayes approach delivering the best results with an accuracy of 74%. Furthermore, it is highly essential to develop more powerful modules based on other algorithms, such as a hybrid approach, which may be more reliable while offering more accurate and faster results. This hybrid method can not only be utilized for this, but it can also be used for many more relevant and even non-relatable scenarios that may be tackled by using this discovery and making it more valuable and opening up new platforms for solving the challenges using machine learning algorithm.

References

1. Kumar A, Narapareddy VT, Srikanth VA, Malapati A, Neti LBM (2020) Sarcasm detection using multi-head attention based bidirectional LSTM. IEEE Access 8:6388–6397
2. Joshi A (2017, Sept) Detecting sarcasm using different forms of incongruity. In: Proceedings of the 8th workshop on computational approaches to subjectivity, sentiment and social media analysis
3. Lunando E, Purwarianti A (2013, Sept) Indonesian social media sentiment analysis with sarcasm detection. In: 2013 International conference on advanced computer science and information systems (ICACSIS). IEEE, pp 195–198
4. Kunneman F, Liebrecht C, Van Mulken M, Van den Bosch A (2015) Signaling sarcasm: from hyperbole to hashtag. Inf Process Manage 51(4):500–509
5. Wicana SG, İbisoglu TY, Yavanoglu U (2017, Jan) A review on sarcasm detection from machine-learning perspective. In: 2017 IEEE 11th international conference on semantic computing (ICSC). IEEE, pp 469–476
6. Bharti SK, Babu KS (2018) Sarcasm as a contradiction between a tweet and its temporal facts: a pattern-based approach. Int J Nat Lang Comput (IJNLC) 7
7. Deep P, Mittal A, Pandey E, Agarwal S (2021) Lexical, pragmatic and linguistic feature based two-level sarcasm detection using machine learning techniques. In: Soft computing for problem solving. Springer, Singapore, pp 699–716
8. Pawar N, Bhingarkar S (2020, June) Machine learning based sarcasm detection on Twitter data. In: 2020 5th International conference on communication and electronics systems (ICCES). IEEE, pp 957–961
9. Das D, Clark AJ (2018, Oct) Sarcasm detection on facebook: a supervised learning approach. In: Proceedings of the 20th international conference on multimodal interaction: adjunct, pp 1–5
10. Deshmukh P, Solanke S (2017) Sarcasm detection and observing user behavioral. Int J Comput Appl 166(9):39–41
11. Ashwitha A, Shruthi G, Shruthi HR, Upadhyaya M, Ray AP, Manjunath TC (2021) Sarcasm detection in natural language processing. Mater Today: Proc 37:3324–3331
12. Bharti SK, Pradhan R, Babu KS, Jena SK (2017) Sarcasm analysis on twitter data using machine learning approaches. In: Trends in social network analysis, pp 51–76
13. Katyayan P, Joshi N (2019) Sarcasm detection approaches for English language. In: Smart techniques for a smarter planet. Springer, Cham, pp 167–183
14. Haripriya V, Patil PG (2017) A survey of sarcasm detection in social media. Int J Res Appl Sci Eng Technol (IJRASET)
15. Felbo B, Mislove A, Søgaard A, Rahwan I, Lehmann S (2017) Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm. arXiv preprint arXiv:1708.00524

A Delicate Authentication Mechanism for IoT Devices with Lower Overhead Issues

R. Raja and R. Saraswathi

Abstract IoT sensor/edge nodes are vulnerable to many well-known threats. Maintaining the absolute reliability of any IoT LAN requires validating edge nodes before entering a network, mainly post entering a sleep state. These IoT nodes possess limited resources and measurable limits, making this a complicated issue. IoT equipment is frequently exposed to the elements since many IoT installations occur in unregulated settings. Equipment cloning and the theft of private keys from edge nodes are among the most common threats on IoT networks because of their easy physical access. These issues motivate to design a highly authentic device that intends to connect with other devices to preserve the stored and incoming data. When the authentication information matches, the connection is established with requesting device. This work aims to create an ultra-thin verification mechanism for IoT LAN-connected end equipment. The gateway, which serves as edge computing equipment, authenticates the end-user equipment. Formal and informal safety checks are performed on the recommended verification process. The simulation is done in MATLAB 2020a environment, where the proposed delicate authentication scheme works well with superior outcomes.

Keywords IoT · Edge nodes · Authentication scheme · Secure nodes · Lower overhead

The original version of this chapter was revised, the affiliation of the author R. Saraswathi has been updated. The correction to this chapter can be available at https://doi.org/10.1007/978-981-19-3035-5_68

R. Raja (✉)
Department of CSIT, CVR College of Engineering, Hyderabad, India
e-mail: nsraja1984@gmail.com

R. Saraswathi
Assistant Professor, Computer Science and Technology, SNS College of Engineering, Coimbatore, India
e-mail: sarasrajagopal2@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_7

1 Introduction

Outdoor IoT deployments are difficult to secure due to easy access to sensor nodes. To be eligible for an IoT LAN, equipment must be able to detect and respond to environmental conditions via wireless communication. Most of these equipment are sensors that collect information wirelessly and send it to gateway/cloud application. IoT networks need selectors to perform appropriately. The actuators reply to service requests from the gateway or cloud application as per the measured attributes. Any environmental area can be monitored and managed in real-time using IoT. Their safety and information protection are critical because of this reliance on networked, intelligent equipment. Due to the nature of crucial real-time systems, frequent patching is not feasible. Encrypted private keys are sent to the platform to authenticate IoT edge nodes, followed by cross-validation. An authenticated requester has been pre-registered. Ordinarily, the device's non-volatile memory contains the equipment's hard-coded private key. That means, an opponent can use a machine to copy the private key shared by the client and server during verification [1]. Attackers can break trust in an IoT LAN using two methods given below:

- In the first approach, an attacker could use hardware tools such as AVR Dragon [1] to make a clone of the authentic equipment. The attacker would then convert the firmware into an image file. The private keys contained in the memory will be passed over to the cloned equipment together with the firmware during this retrieval operation. The cloned one is effectively authenticated to enter the network as a legitimate one [2].
- In the other technique, the attacker will retrieve the firmware and then reverse engineer the hexadecimal code to retrieve the private key. Once, the attacker has obtained the private key, they can create their HTTP request according to the platform's encoding format and then transmit the incorrect information to the forum. This methodology necessitates reverse-engineering proficiency on the attacker's part [3]. Complex verification techniques are required to counter such attacks. The primary issues in establishing a thin verification safety paradigm are two-fold:
- IoT nodes have limited resources. Reduced equipment memory ($\leq 2KBRAMand32KBmemory$) and ultra-low and minimum power availability introduce multiple fresh safety measures in the IoT domain.
- Typical safety models are nearly impossible to implement in such limited networks; protecting it from common threats such as eavesdropping, faking the IP address, node fabrication, and flooding is roughly impossible [4]. So, a reasonable option would be to keep expensive computations at the gateway and keep light analyses at the edge nodes [5].

Our proposed strategy achieves this goal and is summarized below:

(1) How to build reliability among participating equipment is always a fundamental and practical challenge in an IoT ecosystem. This study aims to clarify that IoT network is protected from possible attacks that arise from simple access to edge

nodes. (2) Before establishing a reliable connection, each edge node in a network can authenticate itself with the coordinator/gateway node. (3) Because the edge nodes lack resources, they cannot securely store keys, allowing trust attributes to be readily stolen and trust to be demolished. (4) Given the limits of the edge nodes, it is critical to maintaining the verification process as simple, requiring the fewest resources and consuming the least amount of battery power. (5) Else, keeping a static key to complete the verification phase is not ideal for IoT systems. (6) Here, a delicate authentication mechanism is proposed to establish a better scheme to measure the security among the edge nodes. The major research contribution is listed below:

(1) To model an efficient authentication mechanism to reduce the overhead among the IoT connected devices.
(2) The authentication mechanism performs authentication verification and acknowledgment among the shared keys for preserving the incoming and shared data.
(3) The simulation is done in MATLAB 2020a environment, and the performance is compared with various other approaches.

The work is structured as: In Sect. 2, an elaborate discussion is made regarding the various existing approaches and the pros and cons. The methodology is discussed in Sect. 3 and intends to provide a superior authentication mechanism. The numerical results are discussed in Sect. 4 with the graphical representation, where the summary is interpreted in Sect. 5.

2 Related Works

The set of edge nodes joins the AC supply and forms the recommended verification mechanism in [6] which is adequate for home automation equipment. Because calculation at the edge nodes is intensive, it is not suitable for sensor nodes with limited resources when the verification is needed frequently than which is provided in [6]. Aside from this shortcoming, MAC values are employed in the paradigm, which uses a similar private key to produce and validate the records transferred. The sender and recipient messages are exchanged with pre-agreed keys afore commencing communications, comparable to symmetric encryption. The MAC problem relies on the user who performs validation with MAC and spoofs MACs for other messages.

The verification technique [7] offers a new approach for establishing trust between sensor nodes, proxy servers, and the cloud. The computation is routed through the proxy to limit resource usage at the sensor nodes and connect to the cloud. Nodes can execute symmetric key encryption using this key. Expecting a secure connection between sensors and proxies in most installations where nodes are is not a feasible without the transmission of keys in encrypted form. The suggested approach's weakness is the long run of private key because it is retained in memory. The author in [8] discusses thin fundamental exchange and mutual verification approaches based on hash function. Edge nodes must execute rather intensive computing as part of the

handshaking process, which drains the batteries again. In [9], the author suggested a verification approach based on the equipment's unique fingerprints. As the environmental impacts surrounding items, the presented technique can distinguish between safety threats and regular changes in fingerprints. It recommends a RSS approach for proximity recognition using antenna. It is intended for smart home systems, where an attacker attempts to meddle with intelligent equipment. The differences in both scenarios are sufficient to detect potential threat buildups.

The author in [10] proposes a thin IoT verification protocol. It uses a two-phase verification process: static and continuous. To authenticate equipment at the start of a verification period T, a stationary verification phase is employed, while a constant verification method is used to establish all observed information transmissions within the current verification period T. During the verification window T, the static verification approach creates a verification token for the connecting equipment. After that, a gateway can validate the reliability of the sensor during T verification period and whenever fresh information is exchanged. This technique does not employ encryption. The system relies on a private key created by the gateway and sent unencrypted to the nodes. This key is easily sniffed. The keys are supposed to be held in safe memory, easily accessible via a memory dump. The major research challenges with the various existing IoT implementations are: (1) insecure network services, (2) insufficient authorization and authentication, (3) insecure web interface, (4) poor physical security, (5) insufficient security configuration, and (6) insecure mobile interface.

3 Methodology

In the IoT ecosystem, edge nodes, or "things", are primarily made up of actuators and sensors that sense to requests initiated by applications located on a gateway or in the cloud. Due to their limited resources, edge nodes are vulnerable to a broader range of threat vectors. The majority of sensor nodes spend most of their lives in a dormant state. Another sensor continuously senses its surroundings but transmits information only when values change beyond a specified threshold. Sensors that continually stream information are those found in essential infrastructure. The first two categories of sensor operation modes necessitate regular verification, whereas the third requires handshake before transferring the sensed values. In some cases, anytime equipment rouses and answers a request, it should authenticate itself to the gateway to ensure that the intended equipment is accessing the network. This approach is intended for equipment that operates in the first two modes, whereas constantly streaming edge nodes are verified once. In such cases, the session timeout length might be set to a large amount, reducing the frequency with which the re-verification phase is invoked. Alternately, the streaming edge nodes can be re-verified if they go silent for an extended period. Figure 1 depicts the overall centralized delicate authentication scheme.

IoT equipment would have to verify itself to transmit encrypted information using securely kept keys. It must always preserve a unique identification resistant

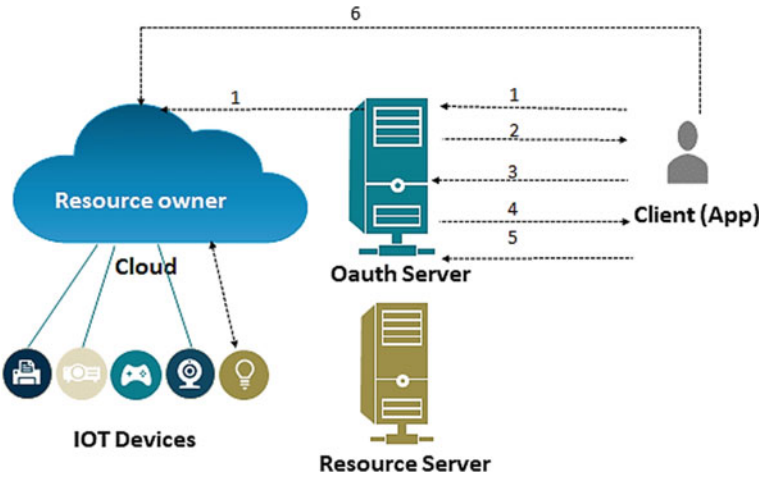


Fig. 1 Centralized delicate authentication scheme

to spoofing and the related encryption keys in the equipment for verification and encryption reasons. Unless these attributes are securely saved, the verification and encryption mechanisms would be useless. The software techniques are insecure, mainly when attackers use software vulnerabilities to enter an IoT LAN. The private is utilized for data decryption along with the counterpart, and the public key is used for encryption purpose. The key needs to safeguard as it is the only key to perform data encryption and decryption. Therefore, the encryption key is used for preserving the incoming data from the connected IoT devices. Hardware-based solutions for essential safety are more visible and tougher to breach. Adding encryption to the nodes is always expensive in constrained energy. Therefore, it is utilized with prudence afore widespread implementations.

(a) **Verification scheme**

With the initiation of authentication cycle between the edge and gateway, the gateway needs to maintain the hardware id (Hid) with initialized vector (V_i), nonce (N_n), and random number (R_n). It is securely exchanged among the corresponding edges, where the gateway evaluates the Key $= h(\text{hidden} \oplus N_n)$ and $\text{key}_{\text{next}} = h(\text{Hidden} \oplus R_n)$. Then, it sets $\text{Key}_{\text{previous}} = \text{Key}_{retreive}$ and $\text{Key}_{\text{previous}} = \text{Key}_{\text{next}}$ for the successive cycles. Gateway needs to initiate the encryption key $E_{\text{key}} = \text{hidden} \oplus V_i$. The preliminary packet is provided in the form of $E_{\text{key}}(\text{hidden}, \text{Key}_{\text{retrieve}}, \text{Key}_{\text{next}}, N_n)$. Then, the gateway needs to maintain a record with authentication table with a form of Hidden, V_i, R_n, $H(N_n)$, $\text{Key}_{\text{retrieve}}$, Key_{next}, E_{key}, SYN, $and ACK$. The gateway transfers the encrypted packets to the nodes, and the nodes have to transfer the receipt of the authentication packet by decrypting the message.

(b) **Verification-based acknowledge phase**

The nodes evaluate the decrypted key (D_{key}) as $D_{\text{key}} = \text{hidden} \oplus V_i$. With the D_{key}, the incoming packets are decrypted to the edge and the parameters are hauled out. Then, it preserves the memory location with the present Key_{next} value for decrypting the authentication packet for the successive cycles transferred by the gateway. The decrypted key is preserved for the successive cycle as stored as $D_{\text{key}} = \text{present KEY}_{\text{next}}$. The nodes set the key for every packet as the encryption key for authentication is transferred to the nodes via gateway. The encrypted key is set as $E_{\text{key}} = \text{key}_{\text{retrieval}}$. The authentication-based packet acknowledgment derived from the nodes is provided to the gateway $\ll h(N_n), E_{\text{key}}(\text{hidden}, \text{Key}_{\text{next}} \gg$. Pseudocode 1 depicts the connectivity among the heterogeneous network nodes. For instance, the connectivity is established among two gateways. The key is generated with 128 bits and establishes the authentication cycle. The authentication measure is to be evaluated between the two nodes, and the key is retrieved between the available and hidden nodes (id). The successive key values are generated based on the authentication cycle. The key is generated for preserving the stored information among the nodes. The retrieval key is used for authentication with retrieval key and decryption key.

Pseudocode 1:
1. Generate gateway using the successive parameters;
2. $H_{id} = 32 \rightarrow providedas128bit$
3. Initialize vector value to establish security and authentication cycle;
4. Set $randomno$;
5. Set $Nonce$;
6. Evaluate authentication with gateway measure;
7. $Key_{retreive} = h(Hidden_{id} \oplus N_n)$;
8. $Key_{successive} = h(Hidden_{id} \oplus R_n)$;
9. Store the successive key values with authentication cycle;
10. $Key_{preserve} = Key_{retreive;}$
11. $Key_{successive} = Key_{next;}$
12. Generate encrypted key;
13. $E_{key} = \text{Hidden} \oplus V_i$;
14. Generate authentication table for successive evaluation;
15. Use $Key_{retrieve}$ value for authenticating the acknowledged packet with decrypted key;
16. end process

(c) **Thin model**

The delicate approach offloads much of the critical calculation, nonce creation, and random number to the gateway to preserve entire computing overheads to a minimum value. The edge nodes may start to operate within the constraints of

the limited resources. The edge node decrypts the acknowledgment packet using fundamental values exchanged during the prior authentication phase already and encrypts the data along with the keys from the present process. An entry in the auth table represents each acknowledgment phase. An ACK is acquired from the node when the verification phase is tagged successfully. The SYN flag is set for entry, and gateway upgrades the ACK flag in the authentication table. If an authentication with SYN/ACK packet is lost, the process is repeated for a certain number of rounds to finish the authentication phase.

4 Numerical Results

The simulation is done in the MATLAB 2020a environment, where most edge nodes consider the battery power. As a result, each computation costs the node extra power. The goal of the experiments in this part is to demonstrate that the calculation suggested running on the edge nodes does not result in a significant rise in battery usage. The assessment set included a coordinator/gateway and end equipment that could function as either a router or an edge node. Table 1 depicts the voltage drop during the process of authentication.

Figure 2 depicts the comparison of authentication versus time (s). Based on the authentication, more number of devices are connected and the voltage drops when the node initiates the process with authentication. Further, the voltage drains with further process like transmission, reception, and so on. The proposed delicate authentication

Table 1 Voltage drop analysis after authentication

Authentication-based voltage–time drainage				
Time	Voltage	Current (mA)	Current (A)	Power (Watt)
1	3.1	50.3	0.0514	0.1540
1.05	2.94	51	0.0510	0.1450
1.09	2.93	51.1	0.0509	0.1485
1.13	2.92	51.2	0.0508	0.1476
1.17	2.92	51	0.0509	0.1475
1.3	2.89	51.6	0.0510	0.1457
1.25	2.90	51.8	0.0511	0.1457
1.29	2.90	51.8	0.0512	0.1458
1.4	2.86	51.7	0.0507	0.1452
1.45	2.86	51.5	0.0506	0.1449
1.49	2.84	51.4	0.0505	0.1437
1.53	2.82	50.4	0.0503	0.1414
1.57	2.82	50.4	0.0504	0.1415
2.1	2.80	50.3	0.0502	0.1401

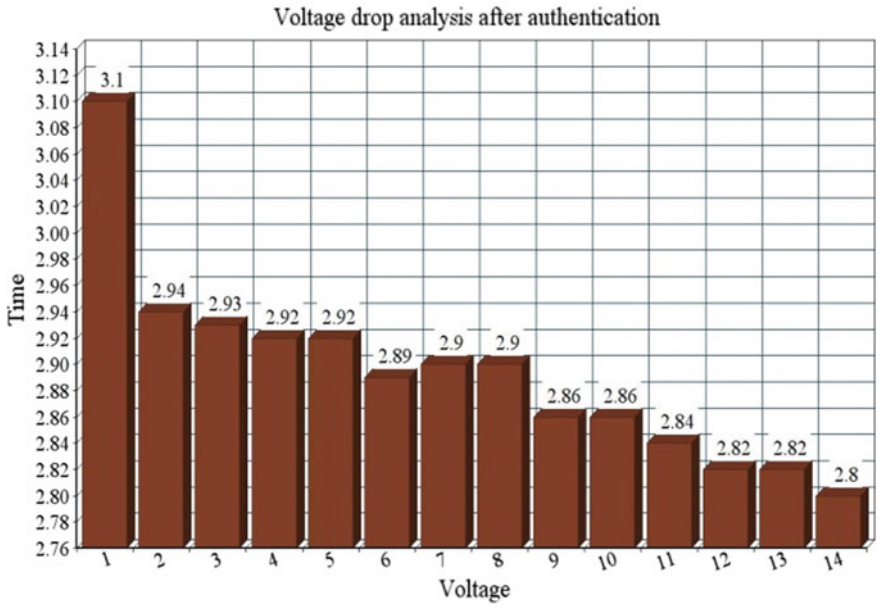


Fig. 2 Time-based voltage drain

scheme is tested with the available edge nodes and routers. The experimentation is done based on authentication and without an authentication mechanism. The authentication mechanism wants the router to be awake and respond to the query from the coordinate nodes. The nodes transfer the series or successive requests from every selective second and communicate the sensed values to the coordinates. The routers and edge nodes are maintained in the LOS with the coordinate nodes devoid of any obstacles from the intermediate nodes. The authentication with the end devices is continued and remains alive until the voltage drops to 2.50 V.

Tables 1 and 2 depict the current and voltage drainage in amps and milliamps of the anticipated model (without and with authentication). While the network model is configured with routers and coordinators, the routers must be awake throughout the entire cycle and perform continuous streaming with essential parameters. During the preliminary stage, the power drain is higher for enabling authentication; however, the voltage drainage shows some consistency. The *t*-test value shows the observed difference among the mean values with appropriate voltage drop significantly with and without authentication.

The nodes are configured as edges where the router acts as a medium and assists the nodes to hop the data to available gateway (nearby). It may acknowledge to the request from the gateway and remain ideal without active. Some routers remained in the awakening stage and configured to receive the request and response. Based on Figs. 2, 3, 4, and 5, it is evident that the power utilization and voltage drop (without and with authentication) are not so significant to achieve overhead. Therefore, the

Table 2 *t*-test analysis

t-test		
Variables	Mean value (router gateway)	Mean (edge gateway)
Drainage (V) without authentication	2.689	2.8528
Drainage (V) with authentication	2.6593	2.8356
Current (I) without authentication	49.9808	51.2365
Current (I) with authentication	49.8756	50.9797
Current drop without authentication	0.2372	0.1486
Current drop with authentication	0.1254	0.1468

authentication scheme consumes elemental power and pretends to offer a delicate authentication mechanism.

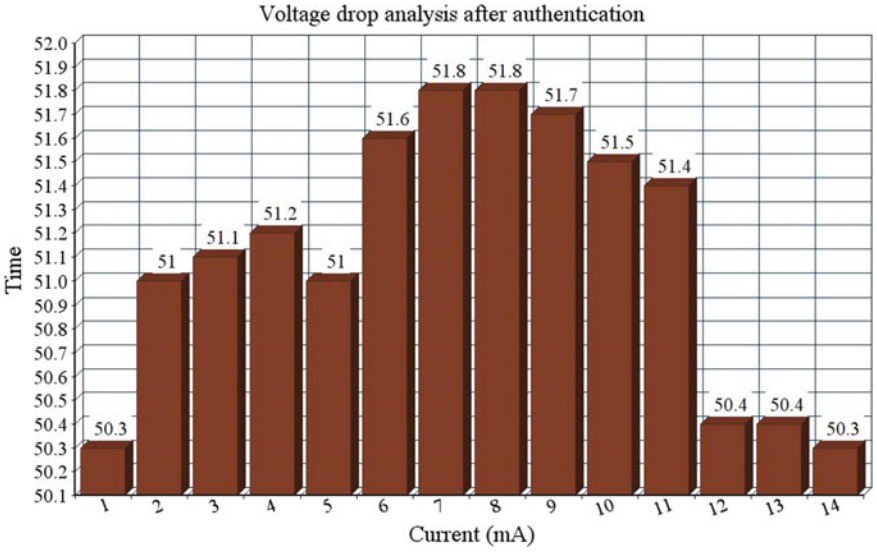


Fig. 3 Current drain analysis

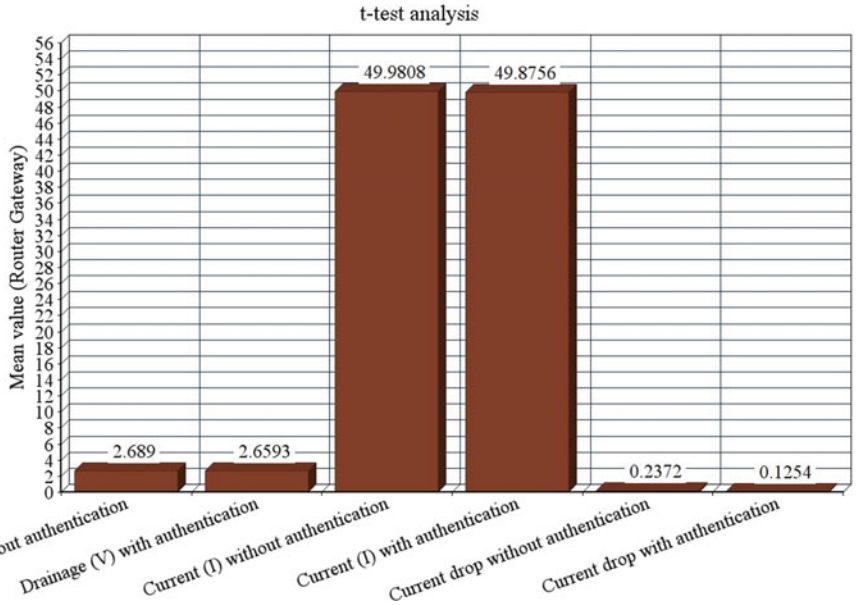


Fig. 4 *t*-test analysis (router to gateway)

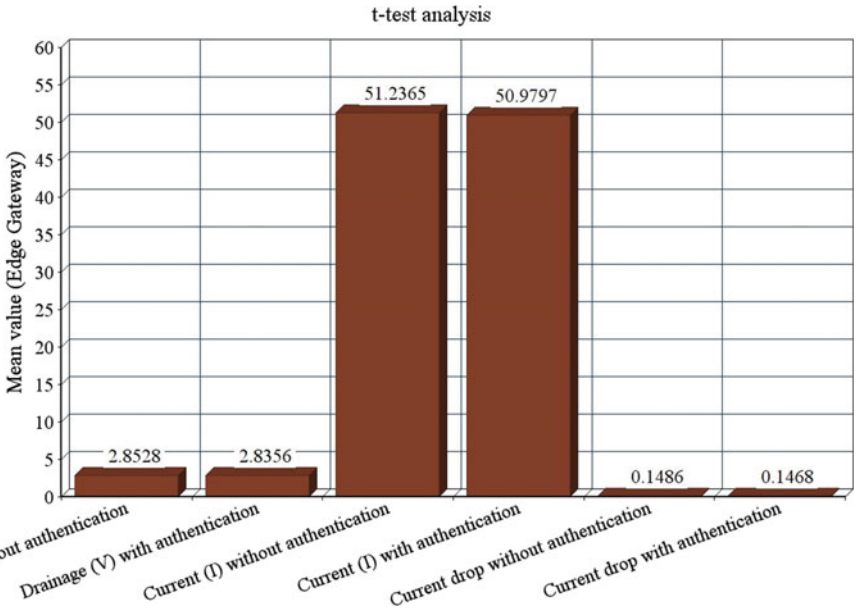


Fig. 5 *t*-test analysis (edge to gateway)

5 Conclusion

The edge nodes in the current IoT implementation rely on encryption keys for safe connections. Those keys are often held immediately or chosen from a pool of keys controlled by a reliable distribution center. The asymmetric encryption keys are created by a third party in mathematical procedures. These are measurably expensive with limited processing power. To solve these flaws, we designed and implemented an appropriate encryption generation system in which verification records are built and protectively shared among the edges and gateways with low measurable overheads. Like many other traditional verification approaches, our suggested method does not necessitate the memory of static key value over the equipment as keys are protectively transferred to avoid message and equipment replicate threats. Those two types of attacks are widespread on IoT equipment deployed in the field. Tests were carried out to validate the efficacy of our approach. Our simulation results indicated no considerable variation in power drainage on IoT equipment.

References

1. Singh J, Pasquier T, Bacon J, Ko H, Eyers D (2016) Twenty security considerations for cloud-supported Internet of Things. IEEE Internet Things J 3(3):269–284
2. He D, Zeadally S (2015) An analysis of RFID authentication schemes for Internet of Things in a healthcare environment using elliptic curve cryptography. IEEE Internet Things J 2(1):72–83
3. Wang X, Hao P, Hanzo L (2016) Physical-layer authentication for wireless security enhancement: current challenges and future development. IEEE Commun Mag 54(6):152–158
4. Dao NN, Kim Y, Jeong S, Park M, Cho S (2017) Achievable multi security levels for lightweight IoT-enabled devices in infrastructure-less peer-aware communications. IEEE Access 5:26743–26753
5. Chen D et al (2017) S2M: a lightweight acoustic fingerprint-based wireless device authentication protocol. IEEE Internet Things J 4(1):88–100
6. Feng W, Qin Y, Zhao S, Feng D (2018) AAoT: Lightweight attestation and authentication of low-resource things in IoT and CPS. Comput Netw 134:167–182
7. Delvaux J (2019) Machine-learning attacks on PolyPUFs, OB-PUFs, RPUFs, LHS-PUFs, and PUF–FSMs. IEEE Trans Inf Forensics Security 14(8):2043–2058
8. Bansal G, Chamola V, Sikdar B, Kumar N, Guizani M (2020) Lightweight mutual authentication protocol for V2G using the physically unclonable function. IEEE Trans Veh Technol 69(7):7234–7246
9. Nozaki Y, Yoshikawa M (2019) Secret sharing schemes based secure authentication for physical unclonable function. In: Proceedings of IEEE 4th international conference on computing and communication system (ICCCS), pp 445–449
10. Liang W, Xie S, Long J, Li K-C, Zhang D, Li K (2019) A double PUF based RFID identity authentication protocol in service-centric internet of things environments. Inf Sci 503:129–147

A Comprehensive Study on Crop Disease Prediction Using Learning Approaches

S. Sandeepkumar and K. Jagan Mohan

Abstract The detection of plant disease is an important problem that requires concentrating on the effective production of agriculture and the economy. The traditional techniques were used to detect this plant disease, but it is a difficult task that needs a lot of time, work, and expertise. The recent research gathered more attention recently between researchers, practitioners, and academicians called automatic detection of plant disease. This automation uses two techniques called machine learning and deep learning that helps for the identification of plant disease at the earlier stage as it finds in leaves of plants. A complete examination was done for the evaluation using the modern study on the possibility of acquiring the machine learning models to find the plant disease. There are four methods of diseases and infection on crops. Primarily, various possible diseases and infections on various crop types are investigated with the cause for the happening and feasible symptoms for the identifications. A thorough investigation on the various pace is needed that is evolved in the detection of plant disease and the categorization with the help of deep learning and machine learning that are given. Different datasets that are there in online to detect plant disease are also provided. A complete investigation regarding different machine learning and deep learning depending on the classification models is discussed, which are given previously and are suggested over the world by various researches for the four above-mentioned crops related to the evaluation on performance, the used dataset, and the method of feature extraction. Finally, different difficulties are listed and presented when using machine learning and deep learning techniques to identify the plant disease and present the future research scope.

Keywords Plant disease · Deep learning · Disease prediction · Performance evaluation · Feature representation

S. Sandeepkumar (✉) · K. J. Mohan
Department of Information Technology, Annamalai University, Chidambaram, Tamilnadu, India
e-mail: ssandeep794@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_8

1 Introduction

The world's population will increase up to 9.1 billion by 2050, according to the Food and Agriculture Organization (FAO) of the United Nations. Henceforth, the rate of growth needs to be risen to 70% by 2050 [1] to manage the huge number of people's nutrient needs. Moreover, various features are considered as the constraints to increasing the production of foods, such as the unavailability of cultivation and clean water in large areas. However, crop disease is the result of the considerable minimization in both the quality of crops and the number of crops.

There are few damages that created an intense impact on the economy that gives low income to the farmer and high prices to the customers in food. Further, this leads to wide shortages of food that follows serious hunger and deprivation, especially in the states which are not developed where the preventive measures are restricted. Generally, human resources physically identify crop diseases via visiting the crops regions. Moreover, this is impossible practically, and it is a time-consuming activity for humans to examine every plant [2]. Henceforth, there is a delay often while examining the steps of crops disease because of the available manpower with restriction. Henceforth, the recognition of crop disease is essential at the initial stage. It also improves the rate of food production and secures the farmers from expensive processes for spraying. The research community has shifted the concentration to present and analyse computer-oriented techniques to manage the difficulties for the detection of plant disease manually. These techniques are used to make this complex task easy and to enhance the automated techniques for the analysis of crops efficiently [3].

Primarily, suggested techniques are used at the earlier stage in immunology and molecular biology for the detection of diseases in crops [4, 5]. Moreover, huge resources, costs, and human experts are needed by these techniques that need to be implemented. The majority of regions for cultivation are small and implemented in non-developed nations by people with minimal income, and this is stated according to FOA [6]. Henceforth, this is not practically possible to them due to expensive solutions, and authors require suggesting effective and efficient techniques available to each farmer [7]. New machine learning techniques are used that are based on the hand-coded techniques in the field of agriculture to increase the power of decision-making [8]. A large amount of information is collected in the real world due to the enhancement of digital methods on which machine learning-based techniques are used for making the optimal decisions. The process of detecting the crop disease is tested with the methods such as decision trees (DT) [9], support vector machines (SVM) [10], K-nearest neighbours (KNN) [11], and Gaussian frameworks [12]. The computational methods that have hand-coded key points are easy to use, and there is no need for much training data.

Moreover, these approaches are time-consuming and require human expertise skills. Further, the traditional machine learning models having keypoint's computational techniques face a trade among complexity and powerful detection. Using the small feature set lowers to localize the system's effectiveness as the computational

high keypoint's vector enhances the economic burden [13]. Henceforth, the improvement on the performance should be needed, especially for the decision support model, which helps to convert a large number of data to the beneficial recommendation.

It is evident that the efficiency of deep learning-based techniques like CNN [14], recurrent neural networks (RNNs) [15], and deep belief networks [16, 17] is used in various applications, which includes segmentation of image [18], classification [19], detection of change [20], and agriculture. For instance CNN is used to empower under the deep learning-based technique, which perceives the optimal elements automatically from the input samples with no intervention of human expertise. The human brain's processing is emulated with the help of using deep learning-based approaches. The numerous object samples are observed to identify and localize the objects by people. Object recognition and pattern recognition follow the same steps by the techniques. Rather than customized machine learning-based approaches, deep learning architectures create more accurate outcomes that improve decision-making. Deep learning frameworks are inspected for resolving the difficult issues due to the extensive progression of the hardware equipment in a considerably low time period.

The deep learning-related techniques exhibit modern accuracy in the agriculture field and have also been general to various works. The significant performance has been achieved by the various forms of deep neural networks (DNNs) above hyperspectral analysis [21]. GoogLeNet [22], AlexNet [23], VGG [24], and ResNet [25] are a few common CNN models that perform better in the classification of crop-related tasks that is to predict the quantity of yield, detection of crop heads, counting the fruit, detection of plant disease, classification, and more. The minimum pre-processing and effort in calculation with powerful results are achieved by these frameworks. Henceforth, exploiting the topological data from the input images results from the better performance for localization of objects by the deep learning-based approaches and unvaried to the post-processing attacks such as translation, rotation. Also, this technique creates good results with respect to quality along with pre-trained approaches. Since a large volume of work is introduced to detect and classify crop disease, improvisation is needed in these frameworks.

Rather than the intelligence of the human brain, the capabilities of machine learning methods are significant to solve the difficulties in the problems of the real world. The low efficiency and high processing time are the important problems of machine learning-based approaches on the detection of crop disease. Also, these techniques create long code, which enhances the complexity computationally. Deep learning methods are introduced to manage these issues of long codes and, moreover, at the costs of improved complexity of code. However, these techniques are not useful for the identification of various diseases above a single sample, and existing techniques are not common to the real-world scenarios also to locate the various examples in a single image for a similar disease [26]. Table 1 depicts the comparison of various existing approaches.

Table 1 Comparative analysis of various existing approaches

Method	Performance	Advantages	Disadvantages
K-FLBPCM framework, along with a support vector machine, was used for the classification of crop disease [3]	98.63% of accuracy	The classification accuracy is improved in the method to the plants having the same morphological textures	The distorted samples face the accuracy degrades while detecting the plant disease
DLQP technique along with support vector machine classifier was presented for the classification of different plant diseases [7]	96.53% of accuracy	This method is powerful for the classification to detect the disease on plant leaves undergoing intense scale and angle variations in input samples	Further improvements are needed on the performance of classification
GLCM technique along with the Harris method was used to compute the features, while the SVM classifier was used to classify the tea plant disease [14]	98.5% of accuracy	This method can able to detect the infected segment of leaves from the complicated background	High computational cost is the problem in this method
SVM classifier along with the LBP algorithm was used to classify the plant disease [14]	95% of accuracy	Better generalization power is used in this method	Degradation on the performance of classification over noisy samples
The important points were calculated by using GLCM and LBP descriptors, while the SVM was used to classify the plant disease [16]	98.2% of accuracy	The framework can locate the diseased plant portion from the suspected samples under the presence of intense light variations	Results are described for the smaller datasets
Random forest classifier along with HOG approach was used for the classification on the samples of infected plants to different classes [19]	70.14% of accuracy	This method is effective computationally	Further improvements are needed on the performance of classification

(continued)

Table 1 (continued)

Method	Performance	Advantages	Disadvantages
GLCM methods, along with SVM classifier and K-means clustering, were used for the classification of turmeric leaves disease [20]	91% of accuracy	Infected plant leaves can be located in this method from the blurry samples	Infected plant leaves can be located in this method from the blurry samples
A deep learning-based technique called FSL was used for the detection and classification of plant disease [20]	91.4% of accuracy	Less training data is used in this approach	Results are described for the smaller datasets
The CNN-based architecture was used for the localization and classification of tomato crops diseases [24]	91.2% of accuracy	This method is effective computationally	The problem of over-fitting has occurred in this method in smaller classes
Mobile application-based approach which used the deep learning-based model called ResNet50 was used for the classification on the different maize crops disease [25]	99% of accuracy	Better generalization power is used in this method	Because of the processing power and the requirements on the consumption of battery, this method will not be suitable for all mobile phones
Faster-RCNN model was used to classify the tomato crops disease [26]	mAP 97.18%	This method is powerful when the distortions and noise are there in suspected samples	This approach is not suitable economically
A deep learning model called AlexNet, along with the K-nearest neighbour classifier, was employed for the classification of tomato leaves like being healthy or unhealthy [28]	76.1% of accuracy	Infected areas are identified from the low-intensity images in this method	This technique is time-consuming and works slowly

(continued)

Table 1 (continued)

Method	Performance	Advantages	Disadvantages
Deep Siamese network along with K-nearest neighbour classifier was employed to classify the plant disease [29]	96% of accuracy	Samples classification accuracy was improved in this method with complicated backgrounds	This method is affected by the issue of over-fitting
CNN-based technique along with a residual network was used to compute and classify the deep features [30]	98% of accuracy	This method is powerful for noisy samples	This technique is expensive computationally
Deep learning-based technique along with the LeNet approach was used for the extraction of important points and classification of the class samples into healthy and unhealthy [30]	94.8% of accuracy	Less training data is needed for this method	This method is not powerful for noisy images

2 Reviews on Plant Disease

Plant diseases are of two types. They are (i) abiotic and (ii) biotic. Abiotic diseases arise in plants due to various elements such as very high temperature, extreme moisture, insufficient vitamins, bad mild, greenhouses gases, and bad soil pH. Biotic diseases arise from dwelling organisms like bacteria, insects, viruses, and fungi [12]. The challenging tasks are identifying and classifying the diseases and the infections in plants, and these diseases are not visible just by visualizing them. Various plant diseases and infections also generally have symptoms, which are usually between various plant diseases that challenge recognizing the exact disease in plants. In certain scenarios, the uncommon signs or symptoms such as the unusual growth of leaves, leaves colour distortion, stunted growth degraded and shrunk pods are used to analyse the disease, and the process of identifying the disease is difficult. For example the blatant case of the infected leaves helps to predict the disease using these symptoms. Early identification and diagnosis of the plant disease are critical processes to safeguard the circumstance, yield of the crop, and the economy incorporated with the agriculture area. There are some datasets used for crop disease prediction. They are: (1) Kaggle dataset; (2) PlantDoc; (3) Village dataset; and so on. Figure 1 depicts four diverse types disease identified in rice leaves.

a) Rice blast b) Red blight c) Stripe blight d) Sheath blight

Fig. 1 Four types of disease in rice leaf

3 Reviews on Crop Disease Prediction

The identification of plant disease automatically denotes the infections that require the hour and need to lower the infections of the crop and loss the crops caused by the infections. The plant diseases are identified, which is an automated process that will minimize the cost of labour needed in the monitoring of crops continuously and closely for the occurred infection. Computer vision and machine learning are generalized techniques that automatically evolve consecutive steps to find possible diseases in crops and plants. This ranges from achieving and gathering the pictures of the plant with the help of various IoT sensors that deploys in the farming fields to the gathered images are processed to feed to the machine learning models to classify the plants as affected or not affected.

Digital image processing has the leading and critical phases, which helps for the conversion of a visible image to an array of binary records for extra processing on a computer. The process of detaining and achieving the suitable images of the objects that leads to machine learning models to learn and classify the purpose is known as image acquisition. The high-resolution digital cameras [9, 25] is used to capture the image in this step. The usage of smartphones helps to capture the samples of images in different assisted formats such as png, jpg, gif [26]. The images are transferred to the stage of pre-processing images after successfully acquiring the needed images based on the requirements. When the collected image is not relevant to the processing sage, the strategies are performed to enhance the image [27]. There are three steps included in the image acquisition technique [27]. They are: (i) using the optical system, the energy is accumulated, (ii) energy is reflected from the surface of the target object, and (iii) the amount of energy is measured by the sensors.

The success of automated classification of plant diseases has a significant role played by the complete step of image acquisition. The accuracy and correctness of

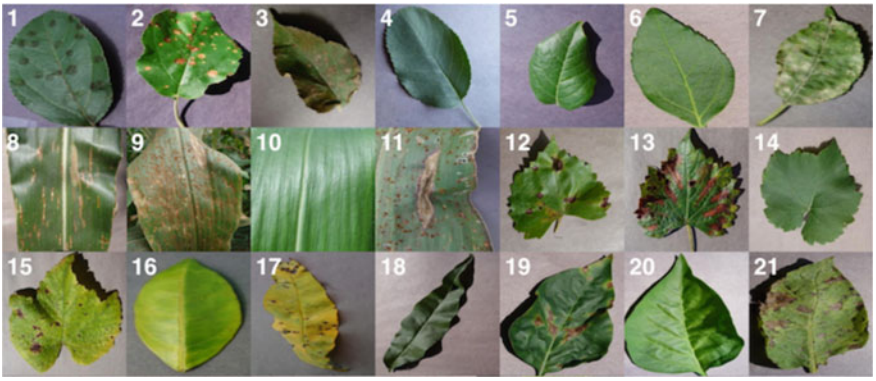


Fig. 2 Samples from Village dataset

the entire technique depend on the quantity and quality of the collected images, and the images are utilized for training the machine learning techniques [9]. Deploying the IoT-based collecting devices in the farming field helps to achieve the classification of plant images as healthy or unhealthy that can be configured periodically to capture the images of the crops. The machine learning techniques that are used to classify and identify the affected diseases in the plant can be trained with the help of previous image databases of the plants, which are available already in the study. Different popular datasets are there in the study, which contains both the healthy and unhealthy images of different plants (see Fig. 2). Few datasets are paid, but many of the datasets are available free of cost. They can be utilized to train the model and evaluate the performance of the suggested machine learning related classification technique.

The first stage of image acquisition is the image pre-processing that performs on the achieved images to get the images to the standard benchmark features, and different machine learning techniques are applied to these features. Image size regularization is the different pre-processing step evolved in this stage to adjust the images based on the used colour scale of the image and so on. The computational cost is reduced, and the make standards on resolutions of images to a particular standard benchmark are used by the pre-processing stage. The aim of image pre-processing is to modify the intensities of the images to highlight the targeted regions [28]. The background images are selected as a white base to reduce the complexity [8]. Pre-processing techniques are used widely, such as removal of distortion, removal of noise, conversion of colour space, cropping the image, enhancing and smoothing the image. The most commonly used pre-processing methods are saturation, hue, and value method (HSV), as it duplicates the humans sensing abilities. The process of removing the background and masking is utilized to fasten the processing and enhance accuracy [9]. However, this mirrors the human sensing features, the transformation of coloured image to popular hue, saturation, and intensity (HSV) representation of colour space. The components are attenuated using the low-pass high frequency. The negative weighting coefficients are used by the high pass filter for

the adjacent pixels, which improves the areas of the image adequately with more gradient intensity so as to highlight the information precisely [15]. The contours in the images are emphasized by the Laplacian filter, and it is utilized to modify the gradients distribution of an image [16]. Either maximal or minimal value from the group of most adjoining pixels is used for the maximal and minimal filter that works as analytical filters, which switches every pixel value [17]. Fast Fourier transform algorithm helps the spatial frequency domain of an image to transform the image from the Fourier transform (FT) [18]. Sigma likelihood of the Gaussian dispersion uses the straightforward and robust common smoothing channel, and the picture clamour is smoothed by taking the average of local pixels that includes the power with the fixed middle pixel sigma scope [19]. Traditional image processing approach involves interconnected steps such as segmentation, feature extraction, and classification. Hence, the prediction depends on given features in this method. Deep learning method operates directly on raw pixels and learns features by itself (Fig. 3) [9].

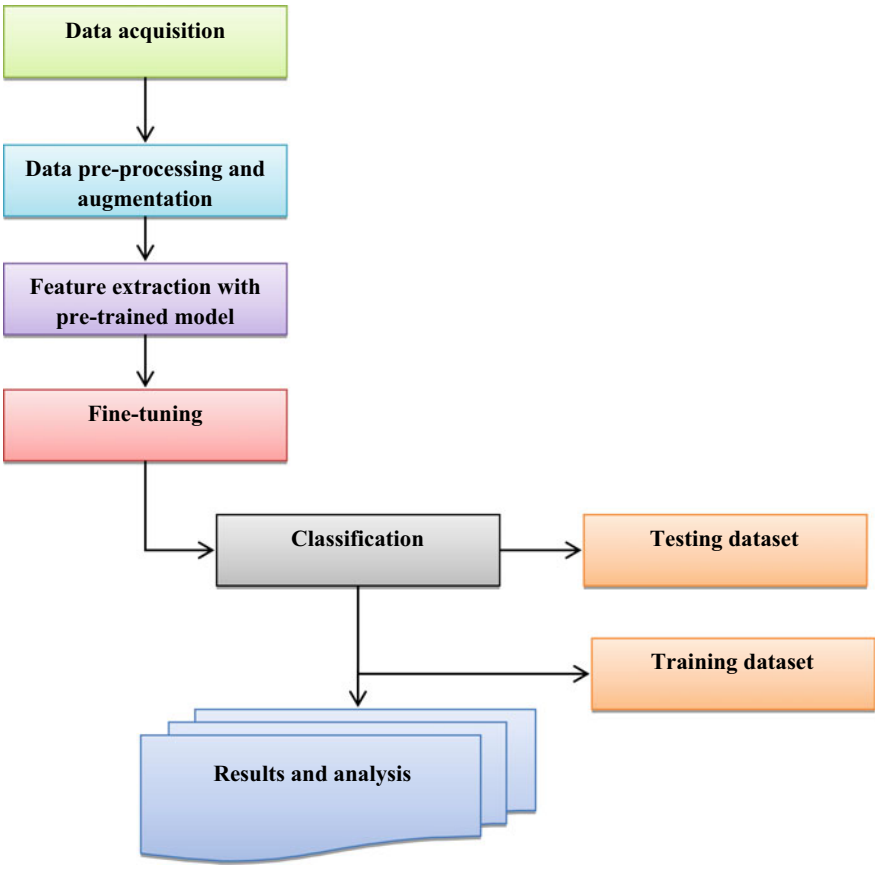


Fig. 3 Block diagram for disease prediction process

4 Reviews on Segmentation Process

The process of subdividing the image into its component objects is called segmentation. The major goal is to examine every object for the extraction of a few useful features. The differentiation of healthy and unhealthy segments of an image is done on the basis of the extracted features [2, 9]. The segmentation is performed by sending the achieved and pre-processed images for the extraction of different relevant, useful attributes needed in the classification and learning process of the infected leaves. Segmentation involves some steps. The first step of this technique works at the pixel level to select the seed points. Certain criteria are used for the region to start growing, like greyscale, the intensity of pixel, similarity, and adjacency.

- The complete image is segmented into the pattern of cells.
- Every cell pattern is relevant to the adjoined cells, and it is attached to the region that depends on the same intensity values.
- The region is developed continuously by analysing the nearest regions till there is no remaining combinative region.
- The above steps are repeated till all cells are attached to the related segment or region.

5 Reviews on Feature Extraction

The pre-trained model can be used as a separate feature extraction programme, in which case input can be pre-processed by the model or portion of the model to a given an output (e.g. vector of numbers) for each input image that can then use as input when training a new model. The colour, morphology, and texture are the basic features of an image. The better results are produced by the morphological features for the disclosure of affected regions of leaves than other features [2, 30]. The Gabor texture and colour moments have utilized the features of colours generally, which are achieved using a number of approaches such as colour histogram [14], colour correlogram [15], colour R moment [16]. Contrast, variance, homogenous, and entropy are the attributes that are associated with the textures. The results are better by the texture features for the problems in detecting the plant disease [9]. The properties of texture are entropy, energy, the moment of inertia, homogeneity, and so on of the needed portion can be calculated with the help of the grey-level co-occurrence matrix (GLCM) [11, 17]. The extraction of texture features with the help of other approaches such as difference operator, FT, discrete cosine transform (DCT), and wavelet packet decomposition techniques [9]. On the basis of colour features, the results for the classification of plant leaves diseases are desired and better [18].

6 Reviews on Classification

Machine learning means computers learning from data using algorithms to perform a task without being explicitly programmed. Deep learning uses a complex structure of algorithms modelled on the human brain. This enables the processing of unstructured data such as documents, images, and text. The process of identifying and classifying the input data into various related sample classes is known as classification. The exact detection and the input data are separated in this phase into various classes. This phase is the initial task in the vision of the computer. The plant images are classified and managed into various classifications in the phase of plant leaves classification that is related to the infected disease. In this phase, leaves are categorized as healthy and unhealthy. The classification of the infected leaves is classified further into possible disease classes, especially when the leaf is suffered. The detection of disease is easier than the classification of disease. Hence, the relevant classifiers are selected depending on the critical problems during the classification of diseases. Machine learning approaches are common of two types. They are unsupervised and supervised models. Generally, the classifier techniques are used such as support vector machine (SVM), K-means, random forest (RF), decision tree, genetic algorithm (GA), KNN, artificial neural network (ANN), fuzzy C-means, BPNN, convolutional neural network (CNN), and fuzzy logic. Moreover, BPNN, CNN, and ANN are taken for the different classification issues that include the detection of disease in plants. These techniques better classify the plant leaves with automatic extraction of features [20]. Applying fine-tuning allows us to utilize pre-trained networks to recognize classes they were not originally trained on. And furthermore, this method can lead to higher accuracy than transfer learning via feature extraction.

7 Performance Metrics

For classifying the crop images, researchers must assess the quality of the classification findings before evaluating the classifier's performance. The classification performance is computed to graphically illustrate the viability of the newly designed algorithm, followed by a comparative study of existing classification algorithms and additional enhancements depending on their inadequacies [5]. Kappa statistics, production accuracy, posterior probability, total classification accuracy, user accuracy, and other assessment indicators are commonly employed, and its execution principle will be momentarily introduced underneath.

a. **Confusion matrix**

   This is also known as an error matrix and is utilized to see if the classification result matches the original ground cover. It is the foundation for several additional assessment criteria. It is expressed in Eq. (1):

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1c} \\ \vdots & \ddots & \vdots \\ x_{c1} & \cdots & x_{cc} \end{bmatrix} \tag{1}$$

where n_{ij} = number of subcategories and x_{ij} = number of categories. The element on the diagonal indicates the number of successfully split data points, and $(i, j = 1, 2, \ldots, n)$ denotes the number of samples from the n-th category split into the (n-1)-th category. Where i is the absolute number, and j is the total number of sample elements.

$$n = \sum_{i=1}^{c} \sum_{j=1}^{c} x_{ij} \tag{2}$$

b. **Overall Accuracy**

The overall prediction accuracy is improved using the adoption of more data. The missing and outlier values need to be considered along with proper feature engineering and representation. The overall accuracy is calculated by dividing the number of accurately sampled data points by the entire sample. The formula for the computation is expressed in Eq. (3):

$$P_{OA} = \frac{\sum_{i=1}^{c} x_{ii}}{\sum_{i=1}^{c} \sum_{j=1}^{c} x_{ij}} \tag{3}$$

c. **Kappa Coefficient**

Kappa coefficient is commonly used to estimate interrater reliability and can be employed in the context of test–retest. In test–retest, the Kappa coefficient indicates the extent of agreement between frequencies of two sets of data collected on two different occasions. The Kappa factor may be used to analyse the division outcomes of images in a particular way, and it examines the number of successfully split sample points and the number of erroneous divisions when evaluating classification performance, which is highly compelling. The following formula can be utilized based just on the confusion matrix and expressed in Eq. (4):

$$P_{KC} = \frac{n \sum_{i=1}^{c} x_{ii} - n \sum_{i=1}^{c} (x_{i+} x_{i+})}{n^2 - \sum_{i=1}^{c} (x_{i+} x_{i+})} \tag{4}$$

where 'n' is the total number of items in each row; x_{i+} is the sum of the elements in each row. It is expressed in Eq. (5):

$$x_{i+} = \sum_{i=1}^{c} x_{ij} \tag{5}$$

Here, x_{i+} denotes the total of the items in each column. Then, precision, accuracy, recall, and F-measure are among the performance measures evaluated for prediction. The performance indicators are based on the confusion matrix, with TP denoting a lot of positive instances that are positive, TN representing the number of negative instances that are negative, FP indicating the series of adverse instances that are positive but are intended just to be positive, and FN indicating the number of positive instances that are negative but are designed to be positive. The model's capacity to handle a real example positively or negatively is solely dependent on the accuracy of the land cover/land use forecast. The fraction of correctly predicted positive instances out of all optimistic predictions supplied by the predictor model is precision. Equation (6) expresses it as follows:

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{6}$$

The recall is represented as the percentage of projected positive events that are always positive. It is written like this in Eq. (7):

$$\text{Recall} = \frac{TP}{(TP + FN)} \tag{7}$$

The $F1$-score is a commonly used metric for classification issues. The recall and accuracy rates are calculated using the average harmonic technique, with the maximum value set to 1 and the minimum value set to 0. Equation (8) is a mathematical expression for it:

$$F1\text{-score} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \tag{8}$$

Specificity is defined as the ability to correctly forecast samples that are not deemed valid, i.e. true negative samples. Equation (9) is a mathematical expression for it:

$$\text{Specificity} = \frac{TN}{FP + TN} * 100 \tag{9}$$

8 Research Challenges

The technique is needed for the extraction of the infected region from the image called segmentation. Moreover, the process of extracting the infected region of a leaf segment from the image is considered tedious when the background of the image has colour rather than black and white, or it has few elements such as leaves, plants, grass, soil [18]. When the farmer needs to find the crops diseases from the fields, the

image has various elements in the background in the real world. Henceforth, there is a capability by the system to remove the entire unwanted elements to acquire the needed segment alone from an image. Most researchers have utilized a few thousand pictures in the study for the training of models. Data augmentation is the provided solution in deep learning to enhance the learning dataset size. Data augmentation is the process of modifying or 'augmenting' a dataset with additional data. This additional data can be anything from images to text, and its use in machine learning algorithms helps improve their performance. The accuracy for the prediction model of disease will be increased using a large dataset to train the model, but the computational cost is high. The covariate shift arises from the difference among distributions of trained data to learn data and the model on which the model is used.

9 Future Enhancements

The minimization of crop loss is essential by taking the relevant precaution at the earlier stages in forecasting disease. Some mobile-based applications and web services are used for the forecasting of disease, and this information will be delivered to the public. In the literature, highly efficient and accurate models are presented. Moreover, in real-world mobile applications and websites, these models are needed to be deployed and tested. Hence, the farmers can utilize the modern automatic disease detection approaches directly by having the image of doubted plant leaves in their farming [19]. There is a restriction that is no real-time machine learning techniques accessible for the identification of disease, and these are taken from the existing studies.

10 Conclusion

The exact classification and recognition of plant diseases are important for farming crops. There are different difficulties considered in the manual detection of crops, such as substantial efforts are needed related to the labour, cost, and expertise to accurately detect the disease of crops. In addition, various attributes are considered, such as farming scale, various types of diseases, and the relevant symptoms for various diseases. The detection of diseases is difficult to work for a farmer, and the diseases need to be found accurately and on a timely basis. This will reflect on the yield and also on the quality of crops. Hence, automatic plant disease detection is introduced to consider the mentioned problems with the help of deep learning and machine learning models. Different deep learning techniques are used for the earlier prediction of diseases efficiently and in a robust manner by examining the leaves of plants. Moreover, it is essential to find different popular machine learning models to synthesize higher and suited best for these applications. This also involves steps that are basic to automatically detect the plant's disease from infection. The deep learning

methods are also required to identify the reason from which the plant is suffering. A detailed investigation needs to be taken to find different machine learning and deep learning models in the study. This study helps to understand which learning model can give better results in various difficult real-time farming conditions. Numerous studies were investigated by the manuscript, and suitable models are investigated to detect the plant's disease, its classification, and different infections which can occur in various types of plants. There are different difficulties while using the deep learning models that impacts are identified on the performance of the automated detection systems of plant disease.

References

1. Atoum Y, Afridi MJ, Liu X, McGrath JM, Hanson LE (2016) On developing and enhancing plant-level disease rating systems in real fields. Pattern Recogn 53:287–299
2. Chen L, Wang R, Yang J, Xue L, Hu M (2019) Multi-label image classification with recurrently learning semantic dependencies. Vis Comput 35(10):1361–1371
3. Ferentinos KP (2018) Deep learning models for plant disease detection and diagnosis. Comput Electron Agric 145:311–318
4. Guan W, Sun Y, Wang J (2017) Automatic image-based plant disease severity estimation using deep learning. Comput Intell Neurosci 2017:2917536
5. He N, Wang T, Chen P, Yan H, Jin Z (2018) An android malware detection method based on deep autoencoder. In: Proceedings of the 2018 artificial intelligence and cloud computing conference, pp 88–93
6. Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 7132–7141
7. Ji M, Zhang L, Wu Q (2019) Automatic grape leaf diseases identification via united model based on multiple convolutional neural networks. Inf Process Agric. https://doi.org/10.1016/j.inpa.2019.10.003
8. Lee J, Seo W, Park JH, Kim DW (2019) Compact feature subset-based multi-label music categorization for mobile devices. Multimed Tools Appl 78(4):4869–4883
9. Pouyanfar S, Wang T, Chen SC (2019) A multi-label multimodal deep learning framework for imbalanced data classification. In: 2019 IEEE conference on multimedia information processing and retrieval (MIPR), pp 199–204
10. Sun H, Wei J, Zhang J, Yang W (2014) A comparison of disease severity measurements using image analysis and visual estimates using a category scale for genetic analysis of resistance to bacterial spot in tomato. Eur J Plant Pathol 139(1):125–136
11. Zhang Z, Hong WC (2019) Electric load forecasting by complete ensemble empirical mode decomposition adaptive noise and support vector regression with quantum-based dragonfly algorithm. Nonlinear Dyn 98(2):1107–1136
12. Zoph B, Vasudevan V, Shlens J, Le QV (2018) Learning transferable architectures for scalable image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 8697–8710
13. Canizares MC, Rosas-Diaz T, Rodriguez-Negrete E, Hogenhout SA, Bedford ID, Bejarano ER, Navas-Castillo J, Moriones E (2015) Arabidopsis thaliana, an experimental host for tomato yellow leaf curl disease-associated begomoviruses by agroinoculation and whitefly transmission. Plant Pathol 64:265–271
14. Fuentes A, Yoon S, Youngki H, Lee Y, Park DS (2016) Characteristics of tomato plant diseases—a study for tomato plant disease identification. Proc Int Symp Inf Technol Converg 1:226–231

15. Chen H, Qi XJ, Cheng JZ, Heng PA (2016) Deep contextual networks for neuronal structure segmentation. In: Thirtieth AAAI conference on artificial intelligence. Association for the Advancement of Artificial Intelligence (AAAI) Publications
16. Fuentes AF, Yoon S, Lee J, Park DS (2018) High-performance deep neural network-based tomato plant diseases and pests diagnosis system with refinement filter bank. Front Plant Sci 9:1–15. https://doi.org/10.3389/fpls.2018.01162
17. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition, pp 770–778. https://doi.org/10.1109/CVPR.2016.90
18. Lin D, Dai J, Jia J, He K, Sun J (2016) Scribblesup: scribble-supervised convolutional networks for semantic segmentation, pp 3159–3167. https://doi.org/10.1109/CVPR.2016.344
19. Milioto A, Lottes P, Stachniss C (2018) Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs, pp 2229–2235. https://doi.org/10.1109/ICRA.2018.8460962
20. Mohanty SP, Hughes DP, Salathé M (2016) Using deep learning for image-based plant disease detection. Front Plant Sci 7:1419. https://doi.org/10.3389/fpls.2016.01419
21. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788. https://doi.org/10.1109/CVPR.2016.91. Ren S, He K, Girshick R, Sun J (2015) Faster r-cnn: towards real-time
22. Zhang Q, Chen MS, Li B (2017) A visual navigation algorithm for paddy field weeding robot based on image understanding. Comput Electron Agric 143:66–78. https://doi.org/10.1016/j.compag.2017.09.008
23. Yamamoto K, Togami T, Yamaguchi N (2017) Super-resolution of plant disease images for the acceleration of image-based phenotyping and vigor diagnosis in agriculture. Sensors 17(11):2557
24. Barbedo JGA (2018) Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Comput Electron Agric 153:46–53
25. Park K, Hong YK, Kim GH, Lee J (2018) Classification of apple leaf conditions in hyper-spectral images for diagnosis of Marssonina blotch using mrmr and deep neural network. Comput Electron Agric 148:179–187
26. Yang Z, Luo T, Wang D, Hu Z, Gao J, Wang L (2018) Learning to navigate for fine-grained classification. In: Proceedings of the European conference on computer vision (ECCV), pp 420–435
27. Ghazi MM, Yanikoglu B, Aptoula E (2017) Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing 235:228–235
28. Chen SW, Shivakumar SS, Dcunha S, Das J, Okon E, Qu C, Taylor CJ, Kumar V (2017) Counting apples and oranges with deep learning: a data-driven approach. IEEE Robotics and Automation Letters 2(2):781–788
29. He K, Zhang X, Ren S, Sun J (2016) Identity mappings in deep residual networks. European conference on computer vision. Springer, Belin, pp 630–645
30. Singh A, Ganapathysubramanian B, Singh AK, Sarkar S (2015) Machine learning for high throughput stress phenotyping in plants. Trends Plant Sci 21:110–124. https://doi.org/10.1016/j.tplants.2015.10.015

Design of Various SRAM Attainment for FINFET

K. Thiagarajan, Karrar Hussain, G. Suresh Kumar, C. Senthilkumar, and Rashmi Maniar

Abstract The FinFET generates a lot of effort, which leads to a higher level of control in the conditional channel. Despite its low performance, 6T static process of random-access memory is modified by circuit function design. This increases the speed of static random-access memory by reducing the bit line with loading effect. With a standard level of cell, 6T static random-access memory face challenges in terms of stability and degradation. It causes disruption in low-level power mode. With a reduced low level of threshold voltage, 6T SRAM has difficulties with output level voltage. In 6 T-type SRAM cell, the conditional transistor destroys the entire read operation. In 6 T SRAM cell, noises destroy the stored level data in nodes, which establishes a direct path between the bit line and storage nodes. We must overcome the 8T-level SRAM as a cell in their recommended read-level stability. To improve the 8T-level SRAM read level of stability, the FinFET with 6T and 8T for conditional achievement of SRAM in 8T level for static random-access memory, and 10T type for SRAM with cells has been increased to improve the need to compare different results by using micro wind as a tool.

K. Thiagarajan (✉)
Department of Mathematics, K. Ramakrishnan College of Technology, Samayapuram, Trichy, Tamilnadu, India
e-mail: arunraja21@gmail.com

K. Hussain
Department of ECE, CVR College of Engineering, Ranga Reddy District, Ibrahimpatnam, Telangana 501510, India

G. S. Kumar
Department of Electronics and Communication Engineering, Vinayaka Mission's Kirupananda Variyar Engineering College, Vinayaka Mission's Research Foundation (Deemed to Be University), Salem, Tamilnadu, India

C. Senthilkumar
Department of Electronics and Communication Engineering, Saveetha School of Engineering, Saveetha University, Chennai, Tamilnadu, India

R. Maniar
Department of English, Lingayas Vidyapeeth, Faridabad, Haryana, India

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_9

Keywords Fin field-effect transistor · Low power · SRAM · System on chip

1 Introduction

The SRAM area is projected to exceed 90% of the total chip area due to the greater convergence and efficiency. The SRAM's substantial contribution has a significant effect on chip yield and cost. The tiny systems technology is used in elevated SRAM cells to improve the data storage density, rendering them more susceptible to process variation. As per the result of process variance, the actual output deviates from the expected value and the device parameter differs from the actual return.

One more issue in scaled technology is increased energy consumption, moreover active power increases with scaling besides leakage power. Lower power consumption can be done by lower supply voltage, but it has a negative impact on cell stability. In SRAM cell, voltage-level scaling is limited by low reliability and improved process differences. In nm technology, attaining the power requirements with a certain level of system tolerance has become a significant challenge.

The SRAM cells access time and stabilization are also affected by process variables. The most common SRAM cell stability failure mechanisms are write, read, keep, and access time failures. It promotes energy usage and reduces cell activity. Further, the stability loss occurs when the voltage level is decreased from a design matrix perspective, and there is still insufficient research analysis available to compare different reported low-voltage process invariant cells in nm technology.

2 Literature Survey

A lot of work is done in field of SRAM cell design. In SRAM cells, several researchers have made significant efforts to reduce strength and improve the stability and design process [1]. Various cell designs based on loop cutting, size, linking bit lines to pseudo-storage nodes, to make reading and writing as stable as possible, as well as read–write assist approaches have been given in the literature [2]. Excessive leakage current is another drawback with SRAM cells. Lowering the supply voltage can help to decrease it. Stack transistor technique is one of the available circuit-level approaches to minimize leakage current [3–5]. Stability and leakage current are the important issues in nanometer region. To solve these issues, various process invariant SRAM cells based on Schmitt trigger, transmission gate, and read buffer have been proposed in the literature [6–8].

Stack transistor is a technique, which is used to minimize leakage current in many circuit-level approaches [9]. As technologist mentioned, stability and leakage current make the important issue in nanometer region. SRAM cells based on transmission gate, read buffer, and Schmitt trigger are proposed in the literature to address the above issues [10].

Loop cutting, connecting bit lines to pseudo-storage nodes, sizing are some of such techniques. Cutting the input between the inverters, while read process is the result of loop cutting technique, and the stored data is fully free from bit line noise. To reach desire stability, write and read ports have to split and configure independently, and one more strategy is explained in the literature [11, 12].

Currently, to solve stability and process parameter issues, a lot of research going on in invariant SRAM cells. Still now, few works are available effectively to compare low-voltage SRAM cell architecture, process variance, and stability in nm technology [13]. The main theme of this paper is to look into previously published SRAM cell designs for leakage, efficiency, stability, current, and process variation robustness. And, it will induce the researchers in solving the issue of nm technology SRAM design [14].

3 Invariant SRAM Cell for Low-Voltage Process

CMOS-based IOT (ST-3) SRAM cell with low-voltage invariant read–write is proposed. The stored data is not disturbed by bit lines. Thus, the ST-3 SRAM cell improves the read stability.

Through weak inverter pairs, the write stability gets increased. To achieve low leakage current in hold mode, the ST-3 SRAM cell uses stack transistor technique. Further, ST-3 cell is used for low-voltage operation. In proposed cell, three sigma (3σ = 30% variation) deviations in threshold voltage due to inter-dia process parameter are used to explain the process tolerance. The simultaneous use of Schmitt trigger and read buffer techniques shows higher process tolerance.

3.1 Proposed Condition for 10T-Type SRAM Cell

The ST-3 cell seems like a conventional 6T SRAM cell except that inverter M3-M1 and M6-M4 are connected to M7-M8 and M9-M10 network, respectively. When read word line (RWL) signal is asserted high, the lower sub-circuit is used to read data from internal storage node, and the upper sub-circuit is used to write data into the cell. In addition, M9 and M7 transistors are used to reduce leakage level current in standby mode. To achieve higher density, the proposed system has used minimum width and minimum length ($W = L = 32$ nm) transistors. During read mode, the proposed level of 10T with SRAM cell appear, which is explained in Fig. 1, and it makes VDD (logic '1'). The RWL transistor makes VDD with WWL kept at GND to start read operation. If logic is "1", store it to node Q, then BL makes discharge level to series connected with NMOS transistors level along M7 and M8.

RWL is turned to VDD, and 10T SRAM cell enhances the read operation in the proposed system. Also, node Q in high, turn on the read access transistor M8. As a result, by transistor M7 and M8, the read current (Ion) flows from BL to GND. Due

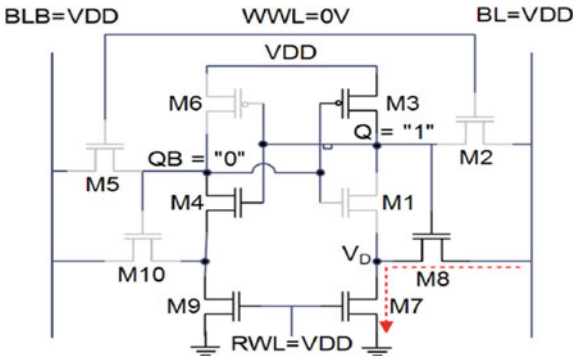


Fig. 1 Read operation—ST-3 SRAM cell

to voltage divider network, the intermediate node VD increases to positive value, which is done in bit line discharge path. Now, the threshold voltage of transistor M1 increases and source voltage (VGS) of transistor M1 turns to negative. At Q node, the increased threshold voltage helps to preserve logic "1" until M1 gate voltage is less than the increased threshold voltage. So, the inverter in ST-3 SRAM cell shows voltage transfer characteristic (VTC) like an ideal inverter. The node voltage Q = VDD is maintained till the positive feedback mechanism between Q and VD is valid. During write operation data is written as Q from BL and QB from BLB through transistors M2 and M5 to internal storage with their nodes.

Consider that, data at node QB is "0" and Q is "1". To commence, write operation RWL is maintained at 0 V signal WWL transitions to VDD. RWL read control signal interrupts the write access transistor (M5 and M2) node QB and Q also GND connection of the inverters (M4-M6 and M1-M3). Proposed cell shows better writing ability, however, the inverter does not have connectivity with GND, and cell data changes easily because the strength is reduced (Fig. 2).

By stack transistor M9 and M7, low leakage current is produced by the proposed system. Voltage of the data storage node is not changed during the read mode (not

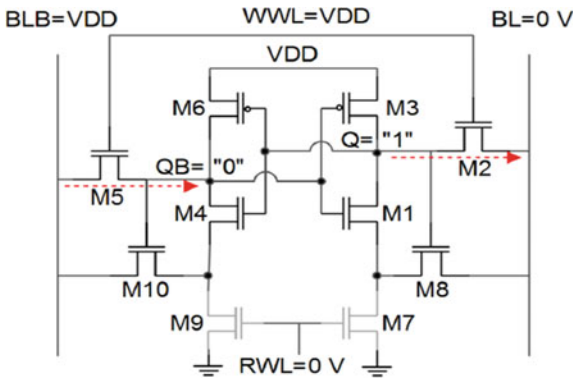


Fig. 2 Write operation with SRAM cell

changed to positive value) it is based on Schmitt trigger technique and buffer technique unlike conventional 6T SRAM cell. Because of read buffer (gate sensing) technique, LP-10T cell and 9T cell enhance read stability. Stored logic is sensed by gate sensing technique through the gate of the transistor. However, by Schmitt trigger technique, ST-1 and ST-2 SRAM cells improve read stability. In ST-3 cell both techniques are used simultaneously, so compared to all other SRAM cell the proposed cell gives more read stability.

Also during hold mode, the proposed cell is disconnected from GND requires less leakage current in hold mode. Design parameter analysis of the proposed 10T SRAM cell is explained in further section. The proposed ST-3 cell is based on read–write decoupled technique and Schmitt trigger. Henceforth, ST-3 cell results are compared with LP-10T cells.

3.2 Performance Parameter Analysis

The performance of the SRAM cell is characterized by bit line leakage current (Ioff) parameters. The Ion/Ioff parameter is used to calculate large size SRAM design. Similarly, for high performance memory design, the Ion is considered as the key parameter. To characterize ST-3 SRAM cells effectively, it is necessary to know these parameters.

The TRA (read access time or read delay) of a SRAM cell is defined as the time difference between the 50 m V falling edge of BLB or BL and 50% rising edge of signal RWL. Here, we calculated read access time only for possible data value (when Q is high). We assumed that RWL takes 1 ps to move from VDD to GND (slope) and GND (0 V) to VDD (0.9 V). Supply voltage of 0.4 to 0.9 V to all SRAM cells for read access time (Table 1).

Comparison of the read access time of all SRAM cells at the nominal supply voltage (0.9 V) is plotted in Fig. 3.

TRA is compared to ST-1 SRAM cell, ST-2 SRAM cell and 6T SRAM cell, respectively at 0.9 V. Because of larger body effect in read buffer transistor (M8 in

Table 1 0.4015 to 0.899 V ranges for read access time with different SRAM cells

VDD(V)	RAT (s)		
	Random-access memory for 6T-S	Random-access memory for 9T-S	Random-access memory for 10T-S
0.4015	0.845E−10	FAIL	FAIL
0.5235	2.201E−10	1.01E−10	1.1203E−10
0.6235	3.203E−11	2.508E−11	3.2102E−11
0.699	4.1021E−11	1.128E−11	1.1502E−11
0.799	2.056E−11	1.209E−11	1.1012E−11
0.899	1.3306E−11	5.5514E−12	5.215E−12

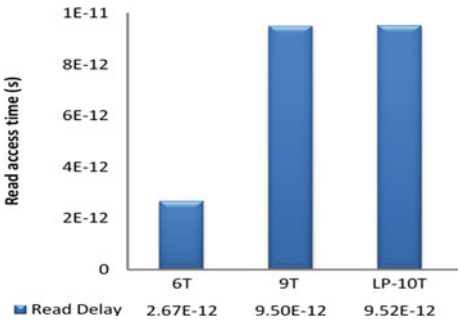


Fig. 3 SRAM cells different ranges from 0.899 V for read access time

this case) the penalty occurs. The "1" in Q (assumed) turns on the transistor M8 and "0" in QB (assumed) turns off the transistor M10, which transfers the node voltage VD to higher than GND potential, when 50mv difference is occur between BL and BLB. It is raised consequently, the reduction in voltage increases body effect, threshold voltage and strength of the transistor M8. So it takes more time to discharge BL. Consequently, M2 transistor (in 6T SRAM cell) body effect is relatively much smaller than transistor M8 (in ST-3 SRAM cell). In discharge path, the ST-1 SRAM cell has three transistors, so it takes long time for read access than ST-2 SRAM cell. ST-2 has two transistors in discharge path, so the same read delay occur as conventional 6T SRAM cell. Read delay for various SRAM cells at low supply voltage (0.5 V) is mentioned here. Lower read delay of ST-3 SRAM cells is compared with LP-10T and 9T cells, respectively. TWA (write access time) or write time of SRAM cell is defined as the time difference between the time when data is written into memory cell and write word line (WWL) (Fig. 4, Table 2).

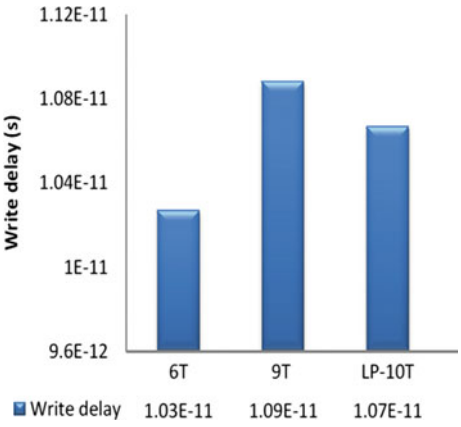


Fig. 4 SRAM cells ranges at 0.899 V with write access time

Table 2 From 0.399 V to 0.899 V ranges from SRAM cells for write access time

VDD(V)	Write access time (s)		
	6T-S random-access memory	9T-S random-access memory	10T-S random -ccess memory
0.399	FAIL	FAIL	FAIL
0.491	6.8809E−11	6.212E−11	6.12E−11
0.592	4.1122E−11	3.2213E−11	3.11E−11
0.693	3.0121E−11	2.223E−11	2.12E−11
0.799	2.130E−11	2.23E−11	2.11E−11
0.899	2.150E−11	2.22E−11	2.112E−11

4 Conclusion

The proposed research work represents the fast read–write with low leakage and low-voltage process invariant SRAM cell for future technology applications. The result from this paper suggests that, the process tolerance capability of SRAM cells dramatically increases by joining read buffer technique with Schmitt trigger technique. From the simulations results, we know that this technique not only provides high RSNM but also gives tight RSNM distribution when compared with other SRAM cells.

The first type of proposed cell shows the value 3.56×, 3.56× and 3.43× as the penalty indicates the read access time which is the value 6T and also ST-9-type SRAM cells value ranges from 0.9 V, and the indication of higher effect through the body effect in which the read value of buffer through the transistor. The second part of the value with the number as the control line most of the greater conventional through the 6T SRAM cell; SRAM architecture required for the condition different in SRAM architecture. A design through the working condition of framework as the variation in the design matrices in the SRAM cells as the application of nanotechnology nodes. Over all proposed cell shows does not sufficient higher enhancement were the condition is the most of the design parameters over the earlier SRAM cells. Therefore, it may be an attractive choice. It is an SRAM cell having a minimum size transistor; therefore, it is suitable for robust high-density sub-threshold SRAM designs. Hence, the proposed 10T SRAM cell is a valid choice for future SRAM design.

References

1. Huang CR, Chiou L-Y (2018) Single bit-line 8T SRAM cell with asynchronous dual word-line control for bit-interleaved ultra-low voltage operation. IET Circuits Devices Syst 12(6):713–719
2. Singh P, Vishvakarma SK (2018) Ultra-low power high stability 8T SRAM for application in object tracking system. IEEE Access 6:2279–2290

3. Nirmal D (2019) Design and efficiency analysis of nanocarbon interconnect structures. J Electron 1(01):12–23
4. Huang C, Wu K, Wu C, Chiou L (2018) Ultra-low stand by power SRAM with adaptive data-retention-voltage-regulating scheme. In: EEE international symposium on circuits and systems, pp 0–3
5. Mansore SR (2017) A single-ended read decoupled 9T SRAM cell for low power applications. In: IEEE international symposium on nanoelectronic and information systems, pp 220–223
6. Aswathy AK, Chalil A (2019) Performance analysis of 6T SRAM cell on planar and FinFET technology. In: 2019 international conference on communication and signal processing (ICCSP), IEEE, pp1–6
7. Kunaraj A, Joy Mathavan J, Jayasekara KGDR (2021) Portable multifunction tester design to check the continuity of wires and to measure the electrical parameters. In: Proceedings of international conference on intelligent computing, information and control systems. Springer, Singapore, pp 863–878
8. Huang C-R, Chiou L-Y (2018) Single bit-line 8T SRAM cell with asynchronous dual word-line control for bit-interleaved ultra-low voltage operation. IET Circuits Devices Syst 12(6):713–719
9. Pradeep K, Mohith B, Manjunath KP, Sunita MS (2020) Comparative analysis of FinFET and planar MOSFET SRAMs. In: 2020 international SoC design conference (ISOCC), IEEE, pp 1–16
10. Shin K, Choi W, Park J (2017) Half-select free and bit-line sharing 9T SRAM for reliable supply voltage scaling. IEEE Trans Circuits Syst I Regul Pap 64(8):2036–2048
11. Arunraja A, Begum RV, Dixit CK, Binny S (2021) POSTGRESQL based monitoring and controlling the parameters for patients. In: Proceedings of the 6th international conference on communication and electronics systems, ICCES 2021, pp 1484–1491, https://doi.org/10.1109/ICCES51350.2021.9489147
12. Ansari M et al (2015) A near-threshold 7T SRAM cell with high write and read margins and low write time for sub-20 nm FinFET technologies. Integration 50:91–106
13. Suzuki T (2010) 0.5V, 150 MHz, Bulk-CMOS SRAM with suspended bit-line read scheme. Proceedings of IEEE European solid-state circuits conference 102:354–357
14. Ramachandran G, Dixit CK, Kishore K, Arunraja A (2021) Performance analysis of mantissa multiplier and Dadda tree multiplier and implementing with DSP architecture. Proceedings of international conference on artificial intelligence and smart systems, ICAIS 2021:1583–1587. https://doi.org/10.1109/ICAIS50930.2021.9395883

A Novel Channel Estimation Technique for Intelligent Reflecting Surface (IRS)-Aided MmWave Systems Based on Sparse Kalman Filter (SKF)

S. Nandan and M. Abdul Rahiman

Abstract Intelligent reflecting surface (IRS) has been materialized as a new physical layer technology to reconfigure and intelligently control the wireless propagation environment. However, for improved system performance in the presence of IRS, knowledge of proper channel state information (CSI) is essential. In this paper, we present a novel and efficient channel estimation technique based on sparse Kalman filter (SKF) in intelligent reflecting surface-supported (IRS) multiple input multiple output (MIMO) millimeter wave (mmWave) wireless communication systems. In addition, we mitigate the training overhead by taking advantage of the sparse attributes and low-rank behavior of the IRS-mmWave MIMO-cascaded channel. Channel effect is evaluated with Rayleigh fading and QPSK as the modulation technique. Simulation results demonstrate that our proposed sparse Kalman filter-based channel estimation method outperforms the conventional methods in the matter of channel estimation accuracy, bit error rate performance, and reduced training overhead.

Keywords IRS-assisted mmWave systems · mmWave wireless communication · IRS channel estimation · Sparse channel · Kalman filter

1 Introduction

Future wireless communication systems mandate a superior data rate and increased spectrum and energy efficiency, better coverage, and reduced latency since they should support billions of interconnected devices. Unfortunately, existing methods like mmWave and massive-MIMO (mMIMO) are deficient in forgathering these requirements [1–3]. Therefore, intelligent reflecting surfaces (IRSs) have stood out as a savvy radio concept for reconfiguring and controlling the wireless communication channel. Scattering and reflection in wireless surroundings are modeled as system

S. Nandan (✉) · M. A. Rahiman
LBS Institute of Technology for Women Trivandrum, APJ Abdul Kalam Technological University Kerala, Thiruvananthapuram, India
e-mail: nandan@lbsitw.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_10

parameters since they are stochastic. IRS enables reconfiguration of the wireless environment with reduced power consumption. They are furnished with many low-cost passive reflective elements like phase shifters and dipoles. IRS is an exclusive hardware technology for enhancing signal coverage, lessen energy consumption, and to reduce the implementation costs. It in-holds many small, passive, and inexpensive reflecting elements. They only reflect the incident signal with an adjustable phase shift and doesn't require a devoted energy source for further handling and retransmission, and hence, IRS has now been included in recent wireless communication systems.

IRS can be employed in mmWave wireless communication systems because of its ability to extend coverage with reduced power consumption because of its passive reflections. Particularly, when the direct path between the transmitter and received is blocked, the IRS helps to deliver a virtual line-of-sight path through passive reflections. Furthermore, by suitably crafting the reflections in the elements, we can enhance the beamforming gain and overall system performance. A wide range of frequency spectrum is on tap in the mmWave frequency range (i.e., 30–300 GHz). But, specifically, for outdoor communication, these frequencies encounter degradation in performance mainly because of first-order reflections and scattering. In addition, the shadowing effect will also affect the received signal strength. The sparsity and low-rank nature of the mmWave channel is another challenge. But, this sparsity property is wielded properly to predict the channel in mmWave systems with hybrid architecture [4]. The introduction of IRS enhances the system performance if the direct link between the base station and the user is blocked [5]. To utilize the complete potential of IRS, the fidelity of guessed channel state information (CSI) is extremely important. The sparse attributes of the mmWave channel can be capitalized to perform various channel recovery techniques. There are two methods of channel estimation available in the literature, viz., semi-passive IRS channel estimation and fully passive channel estimation. Several semi-passive methods are explained in [6], in which low power but active elements like sensors interlaced over IRS elements are used for channel estimation. This will reduce the training overhead since the base station IRS channel and IRS-user channel are estimated separately. Unfortunately, this increases the overall power consumption.

For fully passive IRS estimation, a strategy evolved from the least square (LS) is engaged [7]. Since a fully passive scheme doesn't have active elements, instead of separate channel estimation, the cascaded base station-IRS-user channel is estimated. Even though the cascaded channel dimension is large, it exhibits sparse and low-rank characteristics, which can be put to use for lowering the training expenses considerably. One of the most fundamental iterative greedy algorithms for sparse channel estimation is orthogonal matching pursuit (OMP), applied in [8] and [9], which is a low complex method. For predicting channel in doubly selective systems, Kalman filter using the method of sparse Bayesian learning (SBL) is explained in [10], and a Kalman filter employed channel assessment, and tracking in MIMO-OFDM systems is explained in [11], but both methods are not extended for IRS-aided communication systems. Channel judgment with the help of sparse matrix factorization is one of the methods in which low-rank configuration of the channels are utilized for estimation. A group of passive elements is switched ON and OFF alternatively, but this increases

the complexity since separate amplitude management is required for the same. SBL [12] has proved to surpass various sparse techniques like FOCUSS and basis pursuit (BP). Other compressed sensing algorithms are adaptive grid matching pursuit [13], atomic norm minimization [14], and iterative reweighted method [15]. Even though various channel prediction methods are offered in literature, efficient techniques are yet to be hunted up in IRS-aided mmWave systems for faithful judgment of CSI.

In this paper, we have developed an efficient and accurate method built on sparse Kalman filter (SKF) for estimating the cascaded channel in IRS-assisted mmWave systems. Stepping on the sparsity inherent in mmWave channel, the burden over training is minimized. We have harnessed the sparsity traits in the channel to recover the cascaded state information. Simulation results verify that our proposed method of estimating the IRS channel outplays the established approaches in terms of channel estimation accuracy, bit error rate performance, and reduced training overhead.

2 System Model and Problem Formulation

Assume an mmWave communication system downlink transmission supported by IRS consisting of N number of passive reflecting elements mounted on a 2D plane as shown in Fig. 1. Initially, assume a narrowband system with only one antenna at the transmitter and the receiver units. This will be later extended to wideband multiple antenna case in this paper.

Let $x_b(t)$ represents the complex baseband transmitted signal and $\alpha_k e^{-j\varphi_k}$ be the baseband channel coefficient between transmitter and the kth reflecting element, where $k \in \{1, 2, \ldots N\}$. The received passband signal $\tilde{y}_k(t)$ at the kth element is given as follows:

$$\tilde{y}_k(t) = \mathrm{Re}\{\alpha_k e^{-j\varphi_k} x_b(t) e^{j2\pi f_c t}\} \tag{1}$$

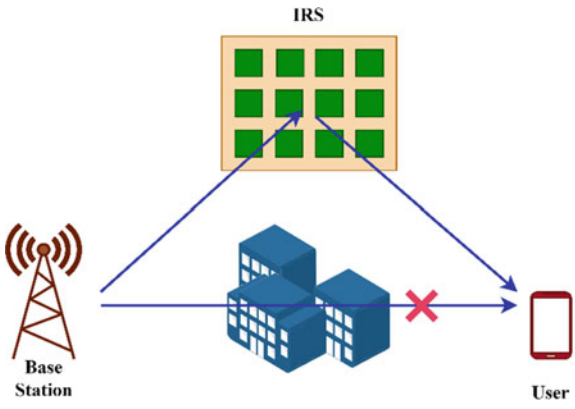


Fig. 1 Basic principle of IRS

Let β_k denotes the reflection amplitude in the kth element and t_k be the time delay introduced to the incident signal; then, the reflected signal from the IRS element, $y_k(t)$, is given as:

$$y_k(t) = \beta_k \tilde{y}_k(t - t_k) \tag{2}$$

$$= \mathrm{Re}\{\beta_k \alpha_k e^{-j\varphi_k} x_b(t - t_k) e^{j2\pi f_c(t-t_k)}\} \tag{3}$$

Considering the narrowband assumption, this can be approximated as:

$$y_k(t) \approx \mathrm{Re}\left\{\left[\beta_k e^{-j\theta_k} \alpha_k e^{-j\varphi_k} x_b(t)\right] e^{j2\pi f_c t}\right\} \tag{4}$$

where θ_k is the phase shift introduced by the kth reflecting element. Thus,

$$y_k(t) \approx \mathrm{Re}\{\beta_k e^{-j\theta_k} x_b^{'}(t) e^{j2\pi f_c t}\} \tag{5}$$

where $x_b^{'}(t) = \alpha_k e^{-j\varphi_k} x_b(t)$ is the equivalent baseband signal of $\tilde{y}_k(t)$.

Equation (5) can be re-written as

$$y_k(t) \approx \mathrm{Re}\{x_b^o(t) e^{j2\pi f_c t}\} \tag{6}$$

where $x_b^o(t) = \beta_k e^{-j\theta_k} x_b^{'}(t)$ is the baseband equivalent of $y_k(t)$. The IRS-reflected signal will experience a similar narrowband flat fading from the kth reflecting element to the user device, and hence, the baseband channel coefficient is given by $\alpha_k^r e^{-j\varphi_k^r}$. The received passband signal at the user device via the IRS, $y_k^r(t)$ is thus given as:

$$y_k^r(t) = \mathrm{Re}\left\{\left[\alpha_k e^{-j\varphi_k} \beta_k e^{j\theta_k} \alpha_k^r e^{-j\varphi_k^r} x_b(t)\right] e^{j2\pi f_c t}\right. \tag{7}$$

The channel coefficients in the cascaded channel model in (7) can be written as $f_k^* = \alpha_k e^{-j\varphi_k}$ and $g_k = \alpha_k^r e^{-j\varphi_k^r}$, where f_k^* corresponds to the channel coefficients between transmitter and kth reflecting element and g_k represents the element-user device link. Equation (12) can thus be re-written as

$$y_k^r(t) = \mathrm{Re}\{[f_k^* \beta_k e^{j\theta_k} g_k x_b(t)] e^{j2\pi f_c t}\} \tag{8}$$

$$= \mathrm{Re}\{y_k(t) e^{j2\pi f_c t}\} \tag{9}$$

where $y_k(t)$ is the baseband signal model from (8). The resultant baseband signal from all the N reflecting elements (neglecting coupling effects) can thus be expressed as:

$$y(t) = \left(\sum_{k=1}^{N} f_k^* \beta_k e^{j\theta_k} g_k\right) x_b(t) \tag{10}$$

$$= \mathrm{f}^{\mathrm{H}} \mathrm{Q} \mathrm{g} x_b(t) \tag{11}$$

where $\mathrm{f}^{\mathrm{H}} = [f_1^*, f_2^* \ldots, f_N^*]$, $\mathrm{g} = [g_1, g_2 \ldots, g_N]^{\mathrm{T}}$ and $\mathrm{Q} = \mathrm{diag}(\beta_1 e^{j\theta_1}, \beta_2 e^{j\theta_2} \ldots, \beta_N e^{j\theta_N})$. Q is called reflection matrix and is a complex diagonal matrix of size $N \times N$. Reflection matrix is diagonal since the elementwise reflections are independent, and there is no coupling between IRS elements. These channel coefficients depend on the path loss, large-scale shadowing, and small-scale fading effects. If there exists a direct linkage from the transmitter to the receiver, the received signal y derived from the reflection model explained in (11) can be written as

$$y = \left(f^H \mathrm{Q} g + h_d^*\right) \sqrt{P_t} x + n \tag{12}$$

where P_t is the transmit power allotted to the base station and n indicates the additive white Gaussian noise (AWGN). Extending this channel model to the MIMO case with N_t transmit antennas, N_r receive antennas and N IRS elements and supposing that there exists a narrowband flat-faded environment, the received signal is expressed as

$$y = \left(\mathrm{F}^{\mathrm{H}} \mathrm{QG} + \mathrm{H_d}\right) x + n \tag{13}$$

where $\mathrm{F} \in \mathbb{C}^{N \times N_t}$, $\mathrm{G} \in \mathbb{C}^{N \times N_r}$ and $\mathrm{H_d} \in \mathbb{C}^{N_t \times N_r}$.

Assuming that the direct channel between base station and user is absent, this can be equivalently written as

$$y = \left(e^H(t) \boldsymbol{H} d(t)\right) x + n \tag{14}$$

where $\boldsymbol{H}$ is the cascaded channel, $d(t)$ is the beamforming vector adopted at the base station and $d \in \mathbb{C}^{N_t}$ and $e = \left[e^{j\theta_1}, e^{j\theta_2} \ldots, e^{j\theta_N}\right]^{\mathrm{T}}$. Our aim is to attain perfect cognizance of the cascade channel matrix for th conjunct achievement of active as well as passive beamforming, i.e., optimize d and e to peak the signal power bagged at the receiver.

Our intention here is for predicting the cascaded channel matrix $\boldsymbol{H}$ arising out of the observations at the receiver $\{y(t)\}$. In our case, this operation is brought off at the user, and not at the IRS. Based on the parametric spatial model explained in [11], the cascaded channel through IRS can also be expressed as:

$$H_{IRS} = \sum_{l=1}^{L} \alpha_l^d \alpha_l^{IRS} \, \mathrm{a}_R\left(\theta_l^r, \varphi_l^r\right) \mathrm{a}_T^*\left(\theta_l^t, \varphi_l^t\right) \tag{15}$$

where α_l^d and α_l^{IRS} represent the path gain corresponding to lth direct path and lth IRS-assisted path. a_R and a_T indicate the array response vector at the receiver and transmitter, respectively. Here, θ corresponds to the azimuth angle, and φ indicates the elevation angle. The amplitude/phase parameter α_l^{IRS} is controllable by IRS. Moreover, a meta-surface-based IRS can control the angles of arrival θ_l^r and φ_l^r.

3 Proposed Channel Estimation Method

The cascaded channel between base station-IRS-user is estimated for uplink and downlink transmission for fully passive channel estimation. Figure 2 shows a fully passive estimation scheme. By exploiting the channel reciprocity property in TDD systems, estimation is done in one direction (uplink/downlink) and uses the same to compute the CSI in the other direction. Hence, we can use the fully passive scheme for TDD and FDD systems.

3.1 Sparse Kalman Filter (SKF)-Dependent IRS Channel Estimation

Kalman filter is a technique used to estimate the states of a linear system. Among the available potential filters, the Kalman filter outshines since it minimizes the variance confederated with the error happened during estimation. The conventional Kalman filter-based (KF) IRS channel estimation of time-varying signals does not deal with the sparsity of the channel and hence results in performance degradation. The proposed sparse Kalman filter estimator combines the advantages of Kalman filter and an improved version of sparse Bayesian learning (I-SBL). Let h_s denotes

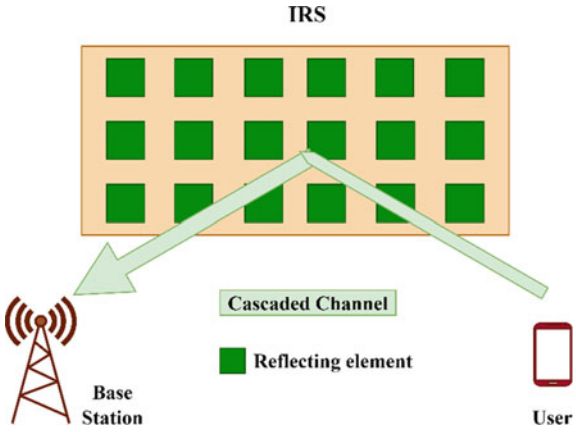


Fig. 2 Fully passive IRS channel

the sparse beam space vector and assumes a parameterized Gaussian prior to h_s.

$$p(h_s;\, P) = \prod_{l=1}^{L} \prod_{g=1}^{G^2} \frac{1}{\pi \rho_l} e^{-\frac{|h_{s,g}(k)|}{\rho_l}} \tag{16}$$

where ρ_l denotes the hyperparameters, P is the hyperparameter matrix, and G corresponds to the set size of quantized spatial angles. Let $\tilde{h}_{s,k|k-1}$ denotes the predicted estimate and $\mathrm{E}_{s,k|k-1}$ denotes the error covariance matrix of the sparse beam space vector at the kth filtering block. The estimate of hyperparameter matrix is denoted as $\tilde{P}_k^{[n]}$; filtered estimate and corresponding error covariance matrix are denoted as $\tilde{h}_{s,k|k}$ and $\mathrm{E}_{s,k|k}$, respectively, and are obtained from the standard Kalman filter prediction step explained as follows:

$$\text{Predictionstep} : \tilde{h}_{s,k|k-1} = \delta \tilde{h}_{s,k-1|k-1} \tag{17}$$

where $\tilde{h}_{s,k|k-1}$ is the prediction at time k, δ is the temporal correlation, and $\tilde{h}_{s,k-1|k-1}$ is the estimate at time $k-1$. If the measurement matrix is denoted by M, then the prediction of measurement or observation at time k is obtained as

$$\tilde{y}_{k|k-1} = M \tilde{h}_{s,k|k-1} \tag{18}$$

$$\mathrm{E}_{s,k|k-1} = \delta^2 \mathrm{E}_{s,k-1|k-1} + (1-\delta^2)\tilde{P}_k^{[n]} \tag{19}$$

Since sparse Kalman filter estimation proposed here involves an improved version of SBL, it is formed on the expectation–maximization (EM) algorithm. Using this, one shall be able to find the maximum likelihood estimates by making do with iterative steps. EM algorithm encompasses two courses of action—initial one is the expectation step (E-Step) which yields the mean value of log-likelihood function. In the maximization step (M-Step), we update the estimate of hyperparameter. The detailed explanation of modified EM algorithm is as follows:

Expectation step (E-Step). The Bayesian hierarchy of the system is modeled as follows:

$$p\big(h_s|y_k;\, P;\, \sigma^2\big) = \frac{p\big(y_k|h_s;\, \sigma^2\big)p(h_s;\, P)}{\sum_{h_s} p\big(y_k|h_s;\, \sigma^2\big)p(h_s;\, P)} \tag{20}$$

where the conditional likelihood $p\big(y_k|h_s;\, \sigma^2\big)$ and apriori probability $p(h_s;\, P)$ are complex Gaussian distributed parameters, i.e., $p\big(y_k|h_s;\, \sigma^2\big) \sim \mathcal{CN}(Mh_s, \sigma^2 I)$ and $p(h_s;\, P) \sim \mathcal{CN}(0,\, P)$. The hidden hyperparameter matrix P can be written as

$$P = Udiag\{\rho_0 V_0, \rho_1 V_1, \ldots \rho_{L-1} V_{L-1}\} \tag{21}$$

where $V_0, V_1, \ldots V_{L-1}$ are covariance sub-matrix elements, the channel tap count be L, and ρ_l is the hyperparameter that manipulates the channel sparsity. Thus, the aposteriori likelihood $p\left(h_s|y_k;\, P;\, \sigma^2\right)$ is also Gaussian and can be expressed as

$$p\left(h_s|y_k;\, P;\, \sigma^2\right) \sim \mathcal{CN}(\mu_{y,k}^n, \mathrm{E}_{y,k}^n) \tag{22}$$

with aposteriori mean value

$$\mu_{y,k}^n = \mathrm{E}_{y,k}^n M^H \left(\sigma^2\right)^{-1} y_k \tag{23}$$

and aposteriori error covariance

$$\mathrm{E}_{y,k}^n = \left[M^H R_w^{-1} M + \left(\tilde{P}_k^{[n]}\right)^{-1}\right]^{-1} \tag{24}$$

Maximization step (M-Step). M-step involves upsurge in the value of the joint likelihood function $p(y_k; \varphi)$, where $\varphi = \left\{\gamma_l, V, \sigma^2\right\}$. This is equivalent to minimizing the log-likelihood function $-log\, p(y_k; \varphi)$. Taking the derivative and equating to zero yields

$$V = \frac{1}{L_N} \sum_{l=0}^{L_N-1} \frac{\left(\mathrm{E}_{y,k,l}^n + h_{s,l} h_{s,l}^H\right)}{\gamma_l} \tag{25}$$

where L_N be the non-zero channel tap count at the current EM iteration and

$$\gamma_l = \frac{1}{N_r} \mathrm{trace}\left[V^{-1}\left(h_{s,l} h_{s,l}^H\right)\right] \tag{26}$$

From this, the residual noise power can be computed as

$$\sigma^2 = \frac{\|y_k - M h_s\| + \widehat{\sigma}^2 \left\{N_r L - Tr\left[\mathrm{E}_{y,k}^n \left(\tilde{P}_k^{[n]}\right)^{-1}\right]\right\}}{L_p - L + 1} \tag{27}$$

Thus, we compute and update the hyperparameters, and the prediction and hyperparameter updating steps are continued iteratively until the convergence is reached. The correction steps in sparse Kalman filter are similar to that of in standard Kalman filter algorithm and is explained as follows:

$$\text{Kalmangain, } G_k = \mathrm{E}_{s,k|k-1} M^H \left(\sigma^2 + M \mathrm{E}_{s,k|k-1} M^H\right)^{-1} \tag{28}$$

$$\text{Innovation, } y_{e,k} = y_k - \tilde{y}_{k|k-1} \tag{29}$$

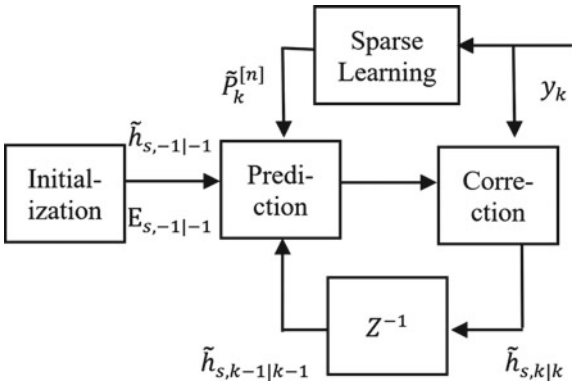


Fig. 3 Block schematic of sparse Kalman filter estimation

$$\tilde{h}_{s,k|k} = \tilde{h}_{s,k|k-1} + G_k y_{e,k} \tag{30}$$

$$\mathrm{E}_{s,k|k} = (\mathrm{I} - G_k M)\mathrm{E}_{s,k|k-1} \tag{31}$$

The termination criterion selected for our SKF algorithm is a given below.

$$\frac{\|\tilde{h}_s^k - \tilde{h}_s^{k-1}\|}{\|\tilde{h}_s^k\|} \leq \varepsilon \tag{32}$$

The block schematic of sparse Kalman filter estimator is as shown in Fig. 3.

where ε is the convergence threshold and its value determines the accuracy of the SKF filter. P_k^0 is initially obtained from the training sequence.

4 Simulation Results

We introduce the achieved simulation results in this section to validate the effectiveness of the proposed sparse Kalman filter-based channel estimation scheme. Furthermore, the results are compared with the traditional methods—orthogonal matching pursuit algorithm (OMP), the traditional Kalman filter relied method and SBL. Table 1 shows the parameters considered.

Figure 4 shows the comparative plot of normalized mean square error values obtained by implementing the orthogonal matching pursuit (OMP) algorithm, SBL algorithm, standard Kalman filter estimation, and the proposed sparse Kalman filter estimation.

Results show that the proposed method exceeds all other prevailing methods available in literature in the matter of the mean square error. Standard Kalman filter shows poor performance compared to OMP and SBL because of its inability to

Table 1 Parameters considered for simulation

Sl. No.	Parameter	Value
1	No. of transmit antennas	32
2	No. of receive antennas (N_r)	32
3	No. of IRS elements (N)	64
4	No. of RF chains	8
5	No. of pilot symbols (L_p)	32
6	No. of iterations	100
7	No. of channel taps (L)	4
8	Doppler shift (f_D)	300 Hz
9	Fading type	Rayleigh
10	Block length (L_B)	300 μs
11	Coherence time	3 ms
12	Channel estimation category	Cascaded (Fully passive)
13	Modulation scheme	QPSK
14	Convergence threshold (ε)	1×10^{-3}
15	Error Correction schemes	Nil
16	Normalized MSE calculation	$\frac{1}{N_t N_r} \|\tilde{H} - H\|^2$
17	Simulation software	Python 3.8

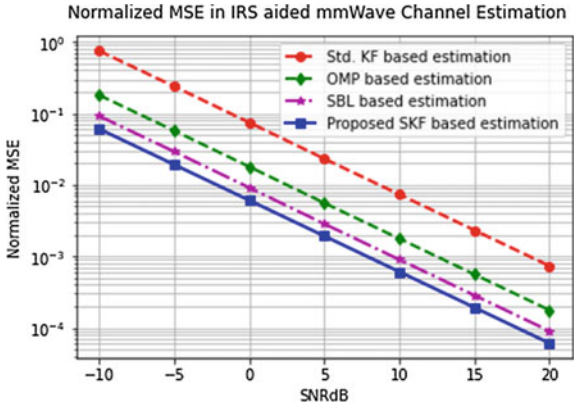


Fig. 4 Normalized MSE versus SNR plot

exploit the sparsity property of the high-dimension cascaded channel. Simulation results obtained here show that there is a significant increase in the performance of the proposed scheme at high SNR regions compared to OMP and SBL. This clearly indicates the improvement in channel estimation accuracy.

Figure 5 demonstrates the bit error rate efficacy of the system by incorporating the proposed and existing channel estimation method. Since channel estimation accuracy is high for our proposed method, the reliability of overall communication increases,

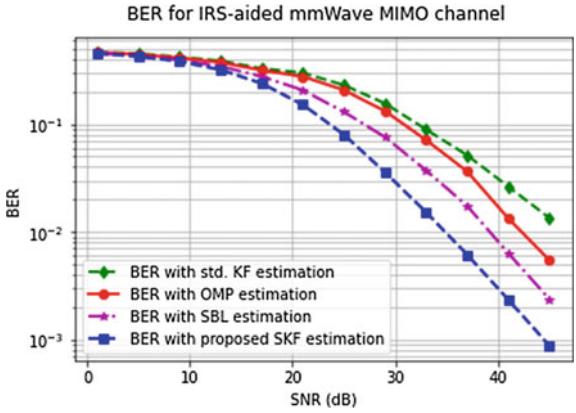


Fig. 5 BER versus SNR plot

i.e., the bit error rate of the proposed method significantly reduces at high SNR region. For example, for a bit error probability value of 1.5×10^{-3}, there shows an improvement of approximately 7 dB in the SNR value for our proposed scheme compared to the existing SBL-based estimation. Proper error correction coding schemes may be employed to reduce the BER range in practical scenarios.

The number of pilot symbols (L_p) assumed for simulation is 32. The proposed method outperforms all other methods for a given fixed pilot sequence length. As pilot size increases, the overall estimation performance improves, but at the cost of increased overhead. The pilot size is comparable to the number of transceiver antennas, and hence, the burden over channel training in the proposed scheme is much less, without degrading the estimation accuracy and reliability.

5 Conclusion

A novel and efficient method for the estimation of cascaded channel in IRS-aided mmWave wireless systems possessing fully passive IRS structure using sparse Kalman filter (SKF) is proposed in this paper. Instead of estimating the base station-IRS and IRS-user channels separately, the cascaded channel from the base station to the user is estimated upon the assumption that the direct connection between the base station and the user device is completely blocked. Since accurate CSI is an important parameter for optimizing the passive reflections in IRS, the beamforming gain can be maximized using our proposed method. Simulation results exhibit the superior potential of the proposed method in contrast with the existing prevalent sparse recovery techniques concerning the estimation accuracy, BER performance, and reduced training overhead.

References

1. Chen JIZ (2021) Modified backscatter communication model for wireless communication network applications. IRO Journal on Sustainable Wireless Systems 3(2):107–117
2. Bashar A (2020) Artificial intelligence based LTE MIMO antenna for 5th generation mobile networks. J Artif Intell 2(03):155–162
3. Dutta KP, Sundi S, Khalkho SS, Tirkey J, Hembrom KK (2021) Minimization of Sidelobe level of uniformly excited concentric circular array antenna using evolutionary technique. Innovative data communication technologies and application. Springer, Singapore, pp 185–194
4. Ayach OE, Rajagopal S, Abu-Surra S, Pi Z, Heath RW (2014) Spatially sparse precoding in millimeter wave MIMO systems. IEEE Trans Wireless Commun 13(3):1499–1513
5. Wang P, Fang J, Yuan X, Chen Z, Li H (2020) Intelligent reflecting surface-assisted millimeter wave communications: joint active and passive precoding design. IEEE Trans Veh Technol 69(12):14960–14973
6. Heath RW, González-Prelcic N, Rangan S, Roh W, Sayeed AM (2016) An overview of signal processing techniques for millimeter wave MIMO systems. IEEE Journal of Selected Topics in Signal Processing 10(3):436–453
7. Mishra D, Johansson H (2019) Channel estimation and low-complexity beamforming design for passive ıntelligent surface assisted MISO wireless energy transfer. In: ICASSP 2019–2019 IEEE ınternational conference on acoustics, speech and signal processing (ICASSP), pp 4659–4663
8. Wang P, Fang J, Duan H, Li H (2020) Compressed channel estimation for ıntelligent reflecting surface-assisted millimeter wave systems. IEEE Signal Process Lett 27:905–909
9. Ardah K, Gherekhloo S, de Almeida ALF, Haardt M (2021) TRICE: a channel estimation framework for RIS-aided millimeter-wave MIMO systems. IEEE Signal Process Lett 28:513–517
10. Srivastava S, Jagannatham AK (2019) Sparse Bayesian learning-based Kalman filtering (SBL-KF) for group-sparse channel estimation in doubly selective mmWave hybrid MIMO systems. In: 2019 IEEE 20th international workshop on signal processing advances in wireless communications (SPAWC), pp 1–5
11. Dalwadi DC , Soni HB (2017) A novel channel estimation technique of MIMO-OFDM system based on extended Kalman filter. In: 2017 4th ınternational conference on electronics and communication systems (ICECS), pp 158–163
12. Wipf DP, Rao BD (2004) Sparse Bayesian learning for basis selection. IEEE Trans Signal Process 52(8):2153–2164
13. Jia C, Cheng J, Gao H, Xu W (2020) High-resolution channel estimation for ıntelligent reflecting surface-assisted mmWave communications. In: 2020 IEEE 31st annual ınternational symposium on personal, indoor and mobile radio communications, pp 1–6
14. He J, Wymeersch H, Juntti M (2021) Channel estimation for RIS-aided mmWave MIMO systems via atomic norm minimization. IEEE Trans Wireless Commun 20(9):5786–5797
15. He J, Leinonen M, Wymeersch H, Juntti M (2020) Channel estimation for RIS-aided mmWave MIMO systems. In: GLOBECOM 2020—2020 IEEE global communications conference, pp 1–6

Low-Power Bit Pair Recoding Technique Using Pre-encoding Mechanism

K. B. Sowmya, Vishal G. Sarashetti, and Anil Nageshwar Rangapure

Abstract Multipliers are widely used in the arithmetic units of digital signal processors. The objective is to design Bit Pair Recoding technique using M-GDI, CMOS technology and to analyze the performance of Bit Pair Recoding technique in terms of area, power, and latency. The methodology of the project consists of a Bit Pair Recoding technique as a top module. In the first step, the pre-encoder is designed for Bit Pair Recoding. The second and third steps involve developing an encoder to generate encoded signals and developing a decoder that generates the partial products for the respective encoded signals, a multiplier register to hold the data from the data bus, and developing a carry propagation adder for the purpose of adding partial products. The Bit Pair Recoding technique is designed using CMOS 45 nm technology and M-GDI 45 nm technology. The power consumption of the design with M-GDI 45 nm technology using a pre-encoded mechanism is about 6.0197 mW. The Bit Pair Recoding technique with M-GDI 45 nm technology provides 17.82.

Keywords Bit Pair Recoding · M-GDI technology · ALU · Pre-encoder

1 Introduction

Multipliers are the most important blocks of multimedia, microprocessors, and digital signal processors. Multipliers are widely used in convolution neural networks (CNNs) for the purpose of multiply accumulate (MAC) operations [1, 2]. Multipliers play an important role in a variety of hardware platforms, including image processing, graphics, and robotic applications. It is necessary to design digital signal processing (DSP) systems and portable multimedia applications which typically require low-power consumption, reduced design cycle, and high processing ability. As the multimedia applications and DSP systems are highly multiplication intensive in nature, the power consumption and performance of these systems are dominated

K. B. Sowmya (✉) · V. G. Sarashetti · A. N. Rangapure
RV College of Engineering, Bengaluru, Karnataka, India
e-mail: sowmyakb@rvce.edu.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_11

by multiplication blocks. Hence, it is necessary to develop high-performance and power-efficient multipliers. The Bit Pair Recoding technique is popularly used to enhance the performance of multiplier because it helps to minimize the number of partial products by half [3].

1.1 Length of Manuscript

Bit Pair Recoding technique with a pre-encoded mechanism has been developed to minimize the power consumption of the Bit Pair Recoding technique. Along with signal processing applications, multimedia, and 3D graphics, performance in most cases strongly depends on the effectiveness of the hardware used for computing multiplications. It is needed to develop a power-efficient multiplier to compose high-performance and low-power applications in order to reduce a considerable amount of energy. Traditional array multipliers generate so many partial products, $N * N$ multiplication has N partial product rows. Accumulation of all partial products in order to generate a final product. The power consumption of array multipliers is very high, and hence, they are not power efficient. Booth algorithm with radix-2 Booth technique is used to enhance the performance of the multiplication [4], and compressor technique is used to enhance the working at low voltage supply [5]. The Booth algorithm with radix-4 technique is used to minimize the number of partial product rows by half in order to aid the multiplication [6, 7], and radix-4 serial–parallel multiplier technique is also used for accelerating applications such as artificial neural networks, digital filters, and other machine learning algorithms [8]. The same Bit Pair Recoding technology is proposed to multiply the two signed numbers as shown in [9]. The redundant binary multipliers are used to reduce the power consumption [7] which is same as Bit Pair Recoding mechanism. In the second phase of the multiplication, we can achieve a reduction in power consumption, cost, and improvement in the overall performance of the Bit Pair Recoding technique. The basic Bit Pair Recoding technique consists of simple encoding/decoding methods [10]. The traditional designs of Bit Pair Recoding techniques suffer from the high cost of encoders and high power consumption. The performance of the encoder and decoder is improved by the pre-encoder mechanism. The efficient adder is implemented in radix-4 booth multiplier to accelerate the multiplication process [11, 12].

2 Theory and Concepts

The Bit Pair Recoding technique as a top module consists of sub-blocks such as decoder, encoder, and pre-encoder. The encoder and pre-encoder are designed based on the Bit Pair Recoding technique. The Bit Pair Recoding technique is chosen to generate fewer partial products compared to the traditional multiplier. The 16×16 bit booth multiplier based on novel 4-2 compressor also works on the same principle of reducing the power and also making the multiplier to work at high frequencies [13].

Table 1 Encoded signals of the Bit Pair Recoding scheme [3]

Y_{2i+1}	Y_{2i}	Y_{2i-1}	Operation on X	ZERO_i	neg_i	ot_i	cor_i
0	0	0	$0X$	1	0	d	0
0	0	1	$+1X$	0	0	1	0
0	1	0	$+1X$	0	0	1	0
0	1	1	$+2X$	0	0	0	0
1	0	0	$-2X$	0	1	0	1
1	0	1	$-1X$	0	1	1	1
1	1	0	$-1X$	0	1	1	1
1	1	1	$0X$	1	1	d	1

2.1 Traditional Bit Pair Recoding Technique

The important blocks in the Bit Pair Recoding technique are the encoder and the decoder. Multiplication is the basic arithmetic operation. The performance and power consumption of the multiplier are very important. The conventional multiplier generates partial products in a huge number; i.e., $N * N$ multiplication has N partial product rows, and the final product is obtained by the accumulation of all the generated partial products. Since the basic multiplier consists of a larger number of generated partial products, it is considered to be less power efficient.

$$\begin{aligned} X \times Y &= X \times \left(-y_{n-1} 2^{n-1} + \sum_{i=0}^{n-2} y_i 2^i \right) \\ &= X \times \sum_{i=0}^{\frac{n}{2}-1} \left(-2y_{2i+1} + y_{2i} + y_{2i-1} \right) 2^{2i} \\ &= X \times \sum_{i=0}^{\frac{n}{2}-1} M_i 2^{2i} \end{aligned} \tag{2.1}$$

$$X \times Y = \sum_{i=0}^{\frac{n}{2}-1} M_i 2^{2i} X \tag{2.2}$$

To facilitate the multiplication, the Bit Pair Recoding technique can minimize the number of partial product rows by half. The powerful method to improve the performance of the multiplication is to use Bit Pair Recoding technique. Table 1 shows the Bit Pair Recoding scheme. The Bit Pair Recoding technique partitions the multiplier Y (Y_{n-1}, $Y_{n-2} \ldots Y_0$) into overlapping groups of continuous three bits. Encoding each group and then decoding them is performed in order to perform

multiplication operation with multiplicand X (X_{n-1}, $X_{n-2} \ldots X_0$) to generate partial products. The n-bit multiplication using Bit Pair Recoding technique can be expressed as Eq. (2.2).

Multiplier Y and Multiplicand X are n bit two's complement numbers. Based on the continuous three bits of multiplier Y (Y_{2i+1}, Y_{2i}, Y_{2i-1}), the corresponding coefficient M_i are generated by the Bit Pair Recoding technique. The coefficient M_i has five possible values (± 1, ± 2, and 0) as shown in Table 1. The Bit Pair Recoding technique can minimize the number of partial product rows by half as shown in Eq. 2.2, and one particular feature is that the corresponding partial products are 0's when the continuous three bits of multiplier Y (Y_{2i+1}, Y_{2i}, Y_{2i-1}) ($M_i = 0$) have the same value.

2.2 *System Architecture of Bit Pair Recoding Technique Using Pre-encoding Mechanism*

The architecture of the Bit Pair Recoding technique consists of encoders, decoders, pre-encoding mechanism, and multiplier registers in order to store the multiplicand and multiplier bits from the data bus as shown in Fig. 1.

In Fig. 2, it is shown that the Bit Pair Recoding technique with a pre-encoding mechanism consists of partial product rows that each row consists of one encoder and $(n + 1)$ decoders. The encoder and decoder generate the partial products. Generated partial products are added using full adders and carry propagation adders to get the final product.

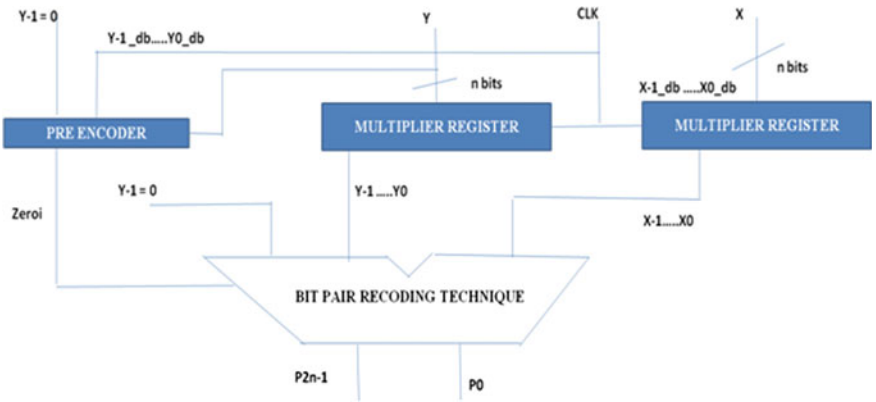


Fig. 1 Architecture of Bit Pair Recoding technique with pre-encoding mechanism

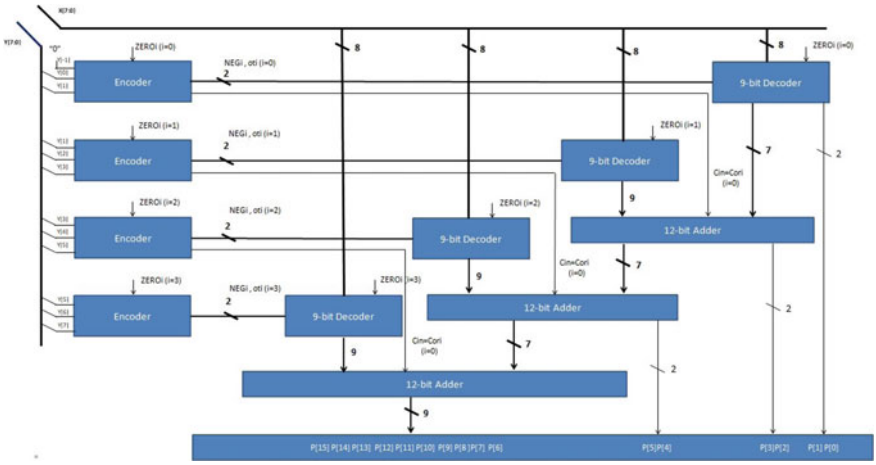


Fig. 2 Architecture of Bit Pair Recoding technique with pre-encoding mechanism

3 Design

The Bit Pair Recoding technique as a top module consists of sub-blocks such as decoder, encoder, pre-encoder, multiplier register, and carry propagation adder. The design of the encoder and pre-encoder is developed based on the concept of the Bit Pair Recoding technique. The decoder implementation is based on output signals from the pre-encoder and encoder that produce partial products. Generated partial products are accumulated through the carry propagation adder to get the final result.

3.1 Design Considerations

The timing chart of the Bit Pair Recoding technique using a pre-encoded mechanism is shown in Fig. 3. Before the multiplication, i.e., multiplication phase, during data setup time, the multiplier Y and multiplicand X are set on the data bus. Figure 3 illustrates that the negative level of the clock pulse is represented with a pre-encode phase, and the positive level of the clock pulse is represented with a multiplication phase by the encoder and decoder.

During the pre-encode phase, the multiplier is fed as input to the pre-encoder. Based on the generated pre-encoder output, the encoder and decoder perform the multiplication operation.

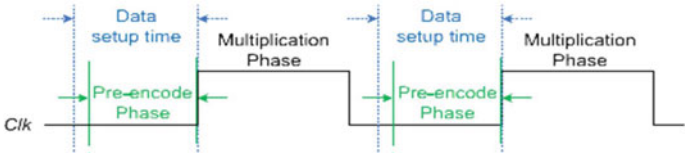


Fig. 3 Timing chart of the Bit Pair Recoding technique using pre-encoded mechanism [3]

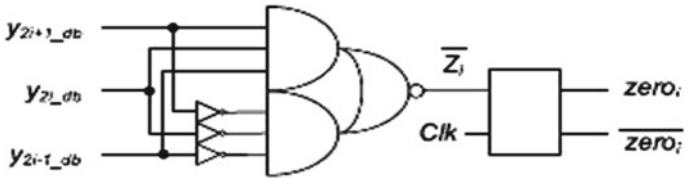


Fig. 4 Gate-level design of pre-encoder [3]

3.2 Design of Pre-encoder

In a multiplier redundant switching activities make the multiplier to consume large amount of power. To overcome this problem, pre-encoder is integrated to find the "0*X*" case. (As mentioned in Table 1, truth table Bit Pair Recoding scheme.) By integrating the pre-encoder, redundant switching activities of the encoder and decoder are reduced. On arrival of "0*X*" case, with the pre-encoded mechanism both the encoder and decoder are turned off.

The Bit Pair Recoding technique is integration of multiplier and multiplicand registers, pre-encoders, and multiplier (includes adders, the low-cost encoder and decoder).

Figure 4 shows the gate-level implementation of the pre-encoder. The pre-encoder mechanism can detect the "0*X*" case throughout the pre-encode phase. To detect the "0*X*" case beforehand of the multiplication, pre-encoders are inserted into the design. As shown in Fig. 1, i.e., the architecture of the Bit Pair Recoding technique, it is needed to integrate the pre-encoder for every encoder.

On arrival of the "0*X*" case, the pre-encoder output of the pre-encoder is set to HIGH. The arrival of the "0*X*" case immediately turns off the corresponding decoders and encoders, and the corresponding partial products are set to "0" to reduce the redundant switching activities. As the "0*X*" case is processed by pre-encoders beforehand, the decoders and encoders only need to perform the multiplication operation for "$\pm 1X$" and "$\pm 2X$" cases.

The output of the pre-encoder, i.e., ZERO_i, can be written as

$$\mathrm{ZERO}_i = ((Y_{2n+1})' \times (Y_{2i})' \times (Y_{2n-1})')\mathrm{db} + ((Y_{2n+1}) \times (Y_{2i}) \times (Y_{2n-1}))\mathrm{db} \tag{3.1}$$

In the pre-encode phase, the pre-encoder starts to detect the "0*X*" case on the arrival of multiplier Y on the data bus. The signal ZERO_i is generated to control

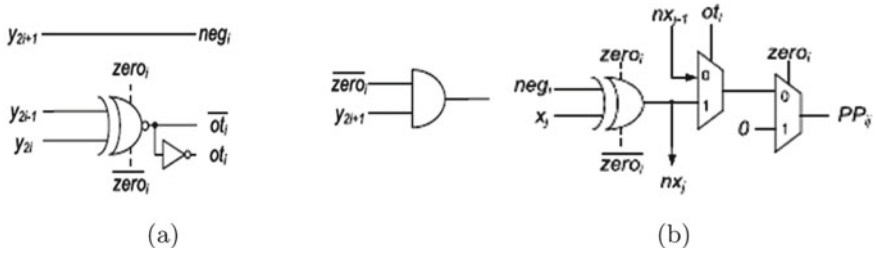


Fig. 5 **a** Gate-level design of encoder, **b** the gate-level design of decoder

the low-cost encoder and decoder. According to the signal ZERO_i, the encoder and decoder work in the multiplication phase.

3.3 Design of Encoder and Decoder

Figure 5a, b shows the gate-level design of the encoder and decoder, respectively.

In the pre-encode phase, if the bits of the multiplier are not the same, the encoder and decoder responsible to generate the corresponding partial products in the multiplication phase [3]. In the pre-encode phase, if the bits of the multiplier are the same, the signal ZERO_i will be set to value 1 by the pre-encoder. The encoder and decoder are turned off, and the partial product is set to 0 directly [3]. Table 1 shows the pre-encode Bit Pair Recoding scheme.

The inputs Y_{2i+1}, Y_{2i} and Y_{2i-1} to the encoder are the outputs of the corresponding multiplier registers (from architecture). The encoder generates the signals neg_i and ot_i, and they are given as,

$$\text{neg}_i = Y_{2n+1} \tag{3.2}$$

$$\text{ot}_i = Y_{2n+1} \wedge Y_{2i} \tag{3.3}$$

$$\text{cor}_i = Y_{2n+1} \times (\text{ZERO}_i)' \tag{3.4}$$

Signal neg_i is equal to Y_{2i+1}, and signal ot_i can be generated by an XOR operation between Y_{2i-1} and Y_{2i}; correction bits can be generated by an AND operation of Y_{2i+1} and ZERO_i'.

The output of the corresponding multiplicand register is connected to the input X_j of the corresponding decoder. The decoder can generate the corresponding partial products according to the encoded signals of the pre-encoder and encoder. As mentioned before, the partial products should be 0 when the case "0X" occurs. The partial product of "+1X" is X_j, and the partial product of "−1X" is X_j'. The partial product of "+2X" is X_{j-1}, and the partial product of "−2X" is X_{j-1}'.

$$\begin{aligned} \text{Partial product } (\pm 1X) &= \text{neg}_i \wedge X_j \\ &= nX_j \end{aligned} \tag{3.5}$$

$$\begin{aligned} \text{Partial product } (\pm 2X) &= \text{neg}_i \wedge X_{j-1} \\ &= nX_{j-1} \end{aligned} \tag{3.6}$$

$$\text{Partial product}(i, j) = \text{ZERO}_i \times 0 + \text{ZERO}_i^{'} \times (\text{ot}_i \times nX_j + \text{ot}_i^{'} \times nX_{j-1}) \tag{3.7}$$

4 Result Analysis

To verify the design of the Bit Pair Recoding technique, the Cadence Virtuoso tool is used. The multiplication result for various input combinations is verified functionally. The power analysis of the Bit Pair Recoding technique is verified for different input combinations. Comparative analysis of Bit Pair Recoding among CMOS 45 nm technology and MGDI 45 nm technology is done.

4.1 Inference of the Results Obtained

From Table 2, it is observed that

- According to the analysis, the number of transistors in CMOS technology is 4040 and 3320 in M-GDI technology. Hence, it is observed that the number of transistors using M-GDI is much less than CMOS technology.
- Implementation of the design using M-GDI technology results in reduction of complexity and area of the design.
- Power reduction is also achieved in M-GDI technology compared with CMOS technology.
- The overall minimum power reduction achieved in Bit Pair Recoding technique is about 6.0197 mW.

The implemented ALU has the following features:

- The design of ALU in M-GDI with 45 nm technology, it is observed that the total power consumption is about 18.919 mW, and the total number of transistors used to design the ALU is about 4358.
- Implementation of the ALU design using M-GDI technology results in reduction of complexity and area of the design.

Table 2 Power calculation for various input combination

		Traditional (CMOS) design with 180 nm technology	Traditional (CMOS) design with 45 nm technology	MGDI with 180 nm technology	MGDI with 45 nm technology
No of transistors used	Pre-encoder	34	34	34	34
	Encoder	36	36	30	30
	Decoder	23	23	17	17
Total power consumption (uW)	Pre-encoder	551.0046	29.41	392.755	19.609
	Encoder	909.199	27.165	449.291	27.683
	Decoder	126.133	9.108	483.113	25.262
Transistors used		4040	4040	3320	3320
Total delay (ns)		27.54	25.93	26.18	25.69

- The amount of time taken to compute multiplication operation has been reduced by replacing traditional multiplier by a Bit Pair Recoding technique using pre-encoding technique. The total delay achieved in ALU is about 31.01 ns.

The verification of the implemented Bit Pair Recoding technique is done in the ANALOG Design Environment (ADE) of the Cadence Virtuoso tool. The design with M-GDI 45 nm technology is implemented using fewer transistors, and hence, the cost of the design is much reduced using M-GDI 45 nm technology. The designed Bit Pair Recoding technique is integrated into the ALU as a multiplication part, and performance analysis of the ALU is done.

5 Conclusion

Implementation of the design is completed in Cadence Virtuoso in CMOS Technology and M-GDI Technology in 45 nm technology, and the results show that the design using M-GDI technology can provide a significant reduction in the number of transistors used, reduction in area, reduction in power consumption, and delay compared with the design using CMOS Technology. The number of transistors is reduced by nearly 720 in M-GDI technology, and also reduced delay is achieved compared to CMOS technology. It is observed that the Bit Pair Recoding technique with M-GDI 45 nm technology provides 17.82% more efficiency in transistor count

than CMOS 45 nm technology. The total delay of the design with M-GDI 45 nm technology is comparable to CMOS technology.

The ALU is designed. The multiplication part of the ALU is replaced with the Bit Pair Recoding technique, and it is observed that the transistor count, overall delay, and power consumption reduction are efficiently achieved.

6 Future Enhancement

The present work on the Bit Pair Recoding technique using pre-encoding architecture can be extended in various applications and in various directions. Some suggestions are given below:

- Digital signal processing (DSP) applications with computationally very high convolution require Bit Pair Recoding techniques in order to save power and reduce the complexity and area of the circuit.
- Machine learning and artificial intelligence applications with deep learning neutral networks require high-speed multipliers compared to conventional multipliers.

References

1. Moons TTB, Verhelst M (2017) An energy-efficient precision-scalable ConvNet processor in 40-nm CMOS. IEEE J Solid-State Circuits 52(4):903–914. https://doi.org/10.1109/JSSC.2016.2636225
2. Jo J, Kim S, Park I (2018) Energy-efficient convolution architecture based on rescheduled dataflow. IEEE Trans Circuits Syst I Reg Pap 65(12):4196–4207. https://doi.org/10.1109/TCSI.2018.2840092
3. Chang Y-J, Cheng Y-C, Liao S-C, Hsiao C-H (2020) A low power radix-4 booth multiplier with pre-encoded mechanism. IEEE Access. https://doi.org/10.1109/ACCESS.2020.3003684
4. Booth AD (1951) A signed binary multiplication technique. Q J Mech Appl Math 4(2):236–240. https://doi.org/10.1093/qjmam/4.2.236
5. Chang C-H, Gu J, Zhang M (2004) Ultra low-voltage low-power CMOS 4–2 and 5–2 compressors for fast arithmetic circuits. IEEE Trans Circuits Syst I Reg Pap 51(10):1985–1997. https://doi.org/10.1109/TCSI.2004.835683
6. Macsorley OL (1961) High-speed arithmetic in binary computers. Proc IRE 49(1):67–91. https://doi.org/10.1109/JRPROC.1961.287779
7. Katariya S, Singhal M (2016) A hybrid 4-bit radix-4 low power Booth multiplier with high performance. In: 2016 IEEE international conference on recent advances and innovation Eng (ICRAIE), pp 1–5. https://doi.org/10.1109/ICRAIE.2016.7939602
8. Moss DJM, Boland D, Leong PHW (2019) A two-speed, radix-4, serial-parallel multiplier. IEEE Trans Very Large Scale Integr (VLSI) Syst 27(4):769–777. https://doi.org/10.1109/TVLSI.2018.2883645
9. Kumar BVNT, Chitiprolu A, Reddy GHK, Agrawal S (2020) Analysis of high speed radix-4 serial multiplier. In: 2020 third international conference on smart systems and inventive technology (ICSSIT), pp 498–503. https://doi.org/10.1109/ICSSIT48917.2020.9214270

10. Abu-Khater IS, Bellaouar A, Elmasry MI (1996) Circuit techniques for CMOS low-power high performance multipliers. IEEE J Solid-State Circuits 31(10):1535–1546. https://doi.org/10.1109/4.540066
11. Akbar A, Adiono T, Harimurti S, Putra TAM (2018) Design of neuron net function using modified radix-4 Booth multiplier with a flipped logic parallel prefix adder. In: 2018 international symposium on electronics and smart devices (ISESD), pp 1–6. https://doi.org/10.1109/ISESD.2018.8605453
12. Arthanto YF, Ibad SI, Adiono T (2019) High speed full custom parallel multiplier based on radix-4 Booth. In: 2019 international symposium on electronics and smart devices (ISESD), pp 1–4. https://doi.org/10.1109/ISESD.2019.8909526
13. Rahnamaei A, Fatin GZ (2018) High speed 16×16 bit Booth multiplier based on novel 4-2 compressor structure. In: 2018 1st international conference on advanced research in engineering sciences (ARES), pp 1–5. https://doi.org/10.1109/ARESX.2018.8723266

A Survey of Deep Learning Region Proposal and Background Recognition Techniques for Moving Object Detection

T. G. Vibha and S. Sivaramakrishnan

Abstract Object detection is the technique of locating targets in an image scene. Moving object detection is to track the object in successive frames which forms the video. Moving object detection forms the primary step in surveillance applications and traffic monitoring systems. With the advent of deep learning remarkable performance is observed in object detection. But the following challenges of detecting small and low-resolution moving objects in the region of interest (ROI), objects with infrequent motion still remain. Deep learning background detection techniques aim at cancelling global camera motion and background movements. The review focuses on the strengths and weaknesses of various deep learning-based moving target region proposal networks and background separation methods.

Keywords Deep learning · Region proposal · Convolutional neural network · Object detection · Average precision

1 Introduction

Deep learning has led to a marked transformation in computer vision field by giving promising results. One such area where the rapid betterment seen with the deep learning networks is object and moving object detection in videos. Moving object detection forms the decisive part for video surveillance systems.

Object detection, tracking and localization form the primary step in video surveillance applications and also in scene interpretation. Scene understanding requires pedestrian detection, direction recognition, identifying moving vehicles in a scene,

T. G. Vibha (✉)
Department of Electronics and Communication Engineering, School of Engineering, Dayananda Sagar University, Bengaluru, Karnataka, India
e-mail: vibhatg@gmail.com

S. Sivaramakrishnan
Department of Electronics and Communication Engineering, Dayananda Sagar University, Bengaluru, Karnataka, India

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_12

tracking different objects. The steps involved for detection includes capturing the region of interest, distinguishing them as foreground and background, labelling them and tracking the same in video for surveillance purpose.

The continuous research is going on in the object detection field with the basis of deep learning networks for achieving enhanced detector performance. The challenges of recognizing small low-resolution objects, cancelling camera, jitter motion, detecting objects with infrequent motions are being addressed by refining deep learning networks. The paper focuses on various deep learning-based region localization networks and background separation techniques for moving object detection.

2 Region Proposal Networks

Few of the region localization networks and techniques used for moving object detection are discussed below. The various such networks focus on identifying the region of interest which forms of the boundary of the object to be detected under consideration.

2.1 RCNN and Fast RCNN—Object and Motion Detection

With the recent development in deep learning approaches, region-based convolutional neural networks (RCNN) and fast RCNN have shown a remarkable improvement in proposing objects in a scene. Figure 1 explains the RCNN [1] architecture which employs a selective methodology of searching of objects in an image and drawing the bounding box around it. The obtained region projections are passed through the convolution networks which extract the feature maps from the proposed regions. The co-ordinate offsets for each object are calculated by finding a good fitting bounding box over overlapping boxes. The offset calculation goes as given by Eqs. 1, 2, 3 and 4. The objects are classified as background with intersection over union (IOU) < 0.3 and as a class for IOU > 0.7.

$$\mathbf{O_x} = (\mathbf{x_g} - \mathbf{x_r})/\mathbf{w_r} \tag{1}$$

$$\mathbf{O_y} = (\mathbf{y_g} - \mathbf{y_r})/\mathbf{h_r} \tag{2}$$

$$\mathbf{O_g} = \log(\mathbf{w_g}/\mathbf{w_r}) \tag{3}$$

$$\mathbf{O_h} = \log(\mathbf{h_g}/\mathbf{h_r}) \tag{4}$$

where x_g, y_g, w_g, h_g are co-ordinates, width and height of ground truth box and xr, yr, wr, hr are co-ordinates, width and height of region. O_x, O_y, O_g, O_h are detected object offsets.

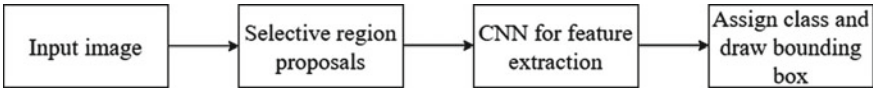


Fig. 1 RCNN functional blocks

RCNN-based moving vehicle recognition was proposed [2], wherein background subtraction is used for obtaining the area of interest against selective search to detect moving vehicles. The ROIs thus obtained are wrapped and applied to CNN for feature drawing from the region proposed. Finally, support vector machine (SVM) classifier recognizes the moving vehicle in the scene. Rigorous pixel movement happening due to faster motion is smoothened/averaged using box filtering technique. This approach showed an improved identification accuracy and precision as compared with SVM + Histogram (HOG), K-Nearest Neighbours (KNN) + HOG, Decision tree + HOG techniques.

The limitation of the increased training period and computation due to convolutional networks used for drawing the features from each identified regions proposed is addressed in fast RCNN [3] proposed by Girshick by combining the RCNN features along with some added layers to reduce training as well as computation time. Figure 2 explains fast RCNN which unlike RCNN performs the convolution on input image drawing out the feature vectors. Then it employs the selective search methodology which proposes regions of interest from projected feature maps. This reduces the computation time as compared to RCNN which does the selective search over the whole image. Later an ROI pooling is applied on all the featured maps in the cropped area downscaling to fixed size as given by Eqs. 5 and 6.

$$\text{New } (x, y) \text{ co-ordinates} = \text{Old } (x/n, y/n) \tag{5}$$

$$\text{New width and height } (w, h) = \text{Old } (w/n, h/n) \tag{6}$$

where n is the compression factor. After downsizing fully connected layers and classifier functions puts the regions into various identified object classes. Training and testing time are reduced by using both training classifier and regressor. The regressor offset calculation for drawing bounding boxes goes with RCNN.

2.2 Faster RCNN-Object Detection

Ren et al. [4] proposed this method as an improvement over RCNN and fast RCNN which adds a region proposal network (RPN) against for selective region projection method used. RPN employs a sliding window which moves over the entire feature map obtained from convolutional layers by downscaling the maps. This is shared across both classifier and regressor. The advantage is the classifier outputs the identified class with the scores assigning each region of interest (ROI) with the object and not object probability. The regressor outputs the boxes over regions by calcu-

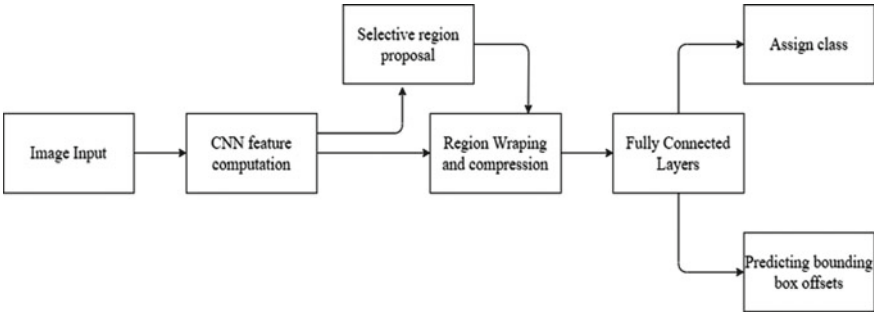


Fig. 2 Fast RCNN functional blocks

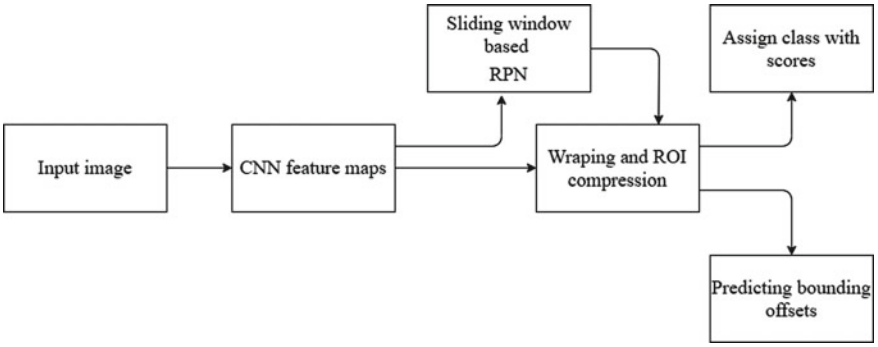


Fig. 3 Faster RCNN architecture

lating width, height and depth. The architecture in Fig. 3 describes the faster RCNN functions.

Dimou et al. [5] along with faster RCNN as detector picked up the challenges of detecting the region of interest in the presence of motion blur, occlusion, moving background and severe scale changes with the CCTV video surveillance clips. They proposed an idea of heterogeneous training and data augmentation as an improvement for detection of objects in such challenging scenarios. The fine-tuning of the deep learning tracking system is done through RPN + Spatial transformer network (SPN). Along with that data augmentation is included to increase the training dataset by means of heterogeneous data sources and adding blurred instances. Blurred instances are developed by convolving original images with motion blurred kernels as represented by Eq. 7.

$$B_i = i * k. \tag{7}$$

where i—input image and K—kernel used.

Faster RCNN is used as the detector and is modelled using Recurrent Neural Networks (RNN) + SPN [6] which aims for accurate position detection of targets

through scaling, rotation zooming so on as given in Eqs. 8 and 9.

$$A_f = \text{Spatial transformation}(\text{fconv}(i)) \tag{8}$$

$$h_t = \text{frnntrans}(A_f, h_{t-1}) \tag{9}$$

where A_f—current frame affine transformation, h_{t-1} is the RNN model hidden state in the preceding step. The performance improvement is seen in terms of average precision (AP). The AP without blur and with blurring effect is found to be 63.2 and 62.3% on VOC10 dataset.

2.3 *Mask RCNN and Its Extensions for Object/Moving Object Detections*

Mask RCNN [7] as an extension to faster RCNN outputs an objectness mask on each region of interest which results in a better separation of moving background from foreground. Mask is usually a binary mask created on the detected objects in the bounding boxes which provides a clear vision. Mask layer is implemented as a fully convolutional network (FCN) on each obtained region proposed which outputs a segmentation mask. This network is added in parallel with classifier and bounding box predictor. Mask RCNN architecture is described in Fig. 4.

Zhou et al. [8] have proposed a fusion of two masks, saliency mask through optical flow and objectness mask using Mask RCNN and arrived at improved contour accuracy and region similarity as given by Eq. 10.

$$F_m = D(S, r) \bigcap O \tag{10}$$

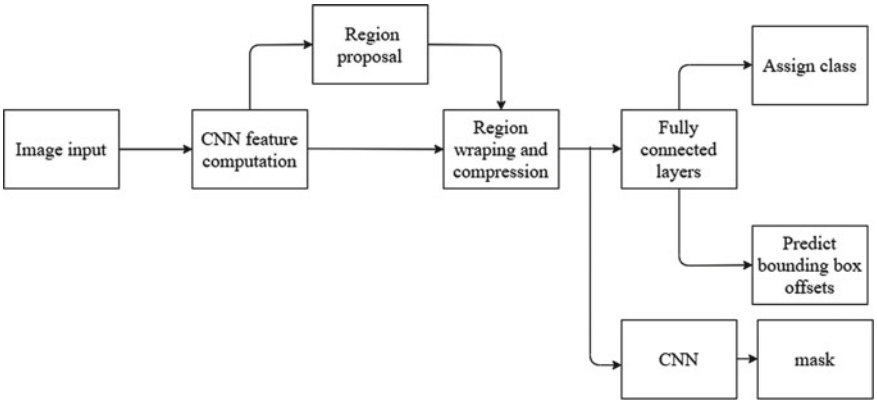


Fig. 4 Mask RCNN architecture

F_m represents the fused segmentation mask by performing dilation (D) on saliency mask S and then applying the intersections of dilated generic object mask with the objectness mask (O). Forward propagation is applied on the mask obtained by fusion to improve segmentation accuracy which considers the segmentation masks propagated through the previous frame to the current one and is applied on both the saliency and objectness masks to obtain the refined one as mentioned in Eq. 11.

$$\text{Refined mask} = \text{Dilation}(\text{refined saliency mask}, r) \bigcap \text{refined objectness mask}. \tag{11}$$

2.4 *Spatio-Temporal—Moving Region Proposal and Detection*

Target detection in an image frame has achieved an enormous development with the use of deep convolutional networks, among which various RCNN networks play a significant role in localizing the objects by proposing the bounding boxes over it. Survey by Wang et al. [9] identified that most of the area proposal networks fail in detecting low-resolution objects and henceforth proposed a MPNET framework which extracts the features of different resolutions from convolution layers by employing VGG16. The multi-dimension information about the moving objects is extracted from adjacent frames, and a fused spatial temporal information of the same is obtained. The implementation is depicted in Eqs. 12 and 13.

$$\text{Frame Difference: } \Delta f = F_m - F_{m-1} \text{ for each frame.} \tag{12}$$

$$\text{Normalize the feature difference for each frame } \Delta F = \sum_{i=0}^{n} \Delta f / n \tag{13}$$

where n represents no of channels for each convolution layer. After this morphological process is applied to reduce noise and detect the edges of the moving objects, and a bounding box is drawn over the moving targets; these steps are carried out for all the frames of the video and on all layers resulting in the detection of objects with different scales. The bounding box adjustment is done over the target to avoid the ambiguities among the detected targets which includes scaling adjustment, extracting feature blocks, normalizing and adding fully connected layers to grab global features and fed into classifier to separate objects from background. Non-maximum suppression (NMS) filtering is done to overcome any overlapping boxes for the same object. To overcome the problems faced when there are moving objects in the scene score detection method is employed which is intended to localize the object in next frame based on the previous frame co-ordinates. Here confidence scores are attached

to the objects as high when object's score is greater than threshold th and as low for low-resolution objects whenever the object's score falls within range *t*min-th where *t*min is the minimum threshold. Object location in say $(n + 1)$ frame is predicted based on the width, height, change in width and height of object from $n - 1$ to n frames. The relative motion of object along the x and y co-ordinates is used to predict the co-ordinates, width, height of object in $n + 1$ frame. The th = 0.5 and tmin = 0.3 give improved performance in terms of average precision on AVSS dataset.

Huang et al. [10] also proposed a attention-based convolutional model which concentrates on maintaining temporal consistency in a video by focussing on the moving regions. A flow-guided model is initially employed to extract the moving object dimensions and is passed through convolution unit to decide on revising and removing the amount of past information. The convolution unit is built as Bidirectional-Convolutional Gated Recurrent Unit (Bi-ConvGRU) to extract spatio-temporal information bidirectionally. Two such Bi-ConvGRUs are used to encode and refine the spatio-temporal information in both the directions. A convolution layer in between two Bi-ConvGRUs is used to produce a focussed weighted map.

Han et al. [11] have proposed the following end-to-end model for multi-target tracking (MTT) by maintaining spatial and temporal coherency. In MTT, the problem is to find associate resembling objects in the successive frames by tracking their paths. To address the problem resemblance matrix is created after extracting the features using VGG16. The matrix is calculated for similarity of object features between the present and previous frames. Then the accumulator matrix is obtained by adding the feature similarity matrix of present and previous frames. Hungarian algorithm is employed to update the path of targets by finding the most appropriate match to the target in the current frame. A residual-like network module (SRN) is combined with VGG16 to capture the resemblance in successive frames at pixel level. Octave convolutional network (OCN) is used to decrease the spatial redundancy. Tracking accuracy and suppression of false negatives are increased as compared to MOT15 and MOT17 datasets.

Kang et al. [12] have looked into improving the moving object detection results by propagating the temporal information to neighbouring frames in the form of tubelet scoring. It includes the still image object detection CNN networks DeepID and CRAFT to obtain the region proposals for the object and scoring is given for the proposed areas for separating them as object and background. The scoring is averaged to improve the detection score. The ranking for detection scores is given using threshold concept, thus separating the classes as high confidence and low confidence. This is followed by suppression of low-ranked classes reducing false positives and by propagating the motion information through the frames to recover false negatives.

2.5 *LSTM and Bidirectional LSTM—Moving Object Detection*

Chen et al. [13] proposed a pixel-wise segmentation approach for localizing the foreground moving objects which incorporated LSTM architecture. The encoder–decoder architecture proposes the foreground objects by depicting the probability in terms of pixel. The input segmented video frames are passed sequentially through the LSTM network which detects the moving objects by paying attention on moving areas and assigning weights. The attention-based LSTM architecture is shown in Fig. 5.

The LSTM networks with the advantage of having long memory as compared to RNNs are utilized to maintain temporal consistency by keeping the memory from previous frame sequences. Through hidden states it learns to localize the moving objects in the video based on the weightage as given by the perceptron network. The perceptron network concentrates on the moving objects by considering the motion property which explores the relative position of the objects from frame to frame to decide on the attention to be paid on each object. Equations 14, 15, 16, 17 and 18 describe the above operation.

$$\begin{bmatrix} f_t \\ i_t \\ n_t \\ o_t \end{bmatrix} = \begin{bmatrix} \sigma w_{fh} \\ \sigma w_{ih} \\ \tanh w_{nh} \\ \sigma w_{oh} \end{bmatrix} h(t-1) + \begin{bmatrix} \sigma w_{fx} \\ \sigma w_{ix} \\ \tanh w_{nx} \\ \sigma w_{ox} \end{bmatrix} x(t) + \begin{bmatrix} b_f \\ b_i \\ b_n \\ b_o \end{bmatrix} \tag{14}$$

$$c_t = c_{it} + c_{ft} \tag{15}$$

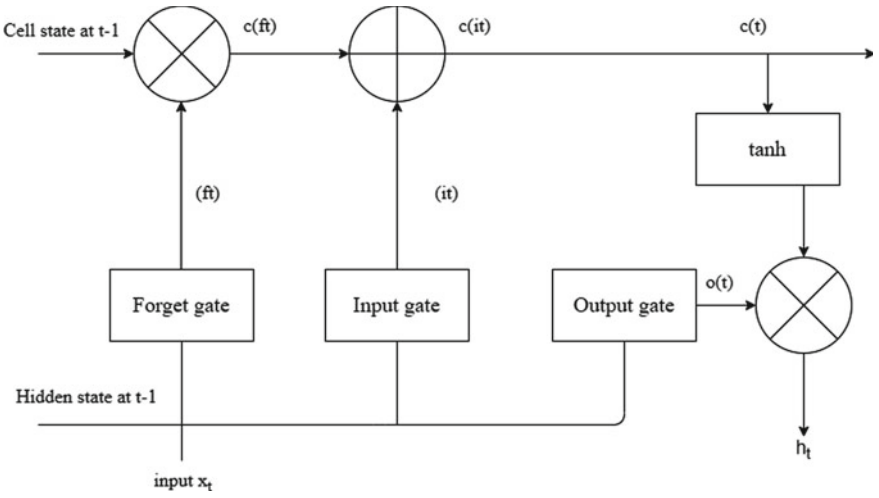


Fig. 5 LSTM architecture

$$c_{it} = i_t * n_t \tag{16}$$

$$c_{ft} = c_{t-1} * f_t \tag{17}$$

$$h_t = \tanh(c_t) * o_t \tag{18}$$

c_{it} = cell state output of input gate.
c_{ft} = cell state output of forget gate.
n_t = output of input node.
h_t = hidden state at time t.
$w_{fh}, w_{ih}, w_{nh}, w_{oh}$ are the weights attached to forget gate, input gate, input node, output gate with respect to hidden state input and $w_{fx}, w_{ix}, w_{nx}, w_{ox}$ with respect to input nodes. b_f, b_i, b_n, b_o are the added biases.

Detection of moving events in a complex background is another research area in object detection. Chen et al. [14] carried out the work by considering the complicated background by introducing the Gaussian noise. The approach includes Mask RCNN to mask the localized objects and VGG16 to draw out the detected features, followed by bidirectional LSTM network which considers both previous frame contents as well as future thus the network learns temporal information from both the directions, and the required features are highlighted using weighted attention method. The forward propagation method is employed which inputs the previous frame features to the next node through hidden layers along with the input frame. Background propagation is performed in the similar way, thus preserving the temporal information. The output features thus obtained from both directions are averaged out and provided as the detected object output. The bidirectional LSTM network is shown in Fig. 6.

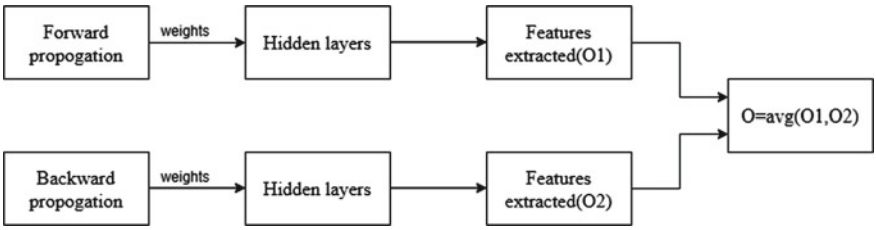


Fig. 6 Bidirectional LSTM architecture

2.6 Cascaded RCNN-High Quality Object Detection

To reduce the false positives is one more research area in object detection. Cascaded RCNN was proposed by Cai [15] which is built as series of detection, classification and bounding box blocks. The output of one stage is used for training the other stage thus IOU can be varied from stage to stage to improve similarity between ground truth and detected target which results in reduction of false positives. The proposed cascaded version of RCNN shows an improvement over the datasets such as COCO, pascal VOC, Kitti which are presumed to be challenging datasets.

2.7 Cuboid Region Proposal Network (CPN)—For Video Object Detection

Tang et al. [16] proposed a cuboid proposal for localizing object region in videos which takes n frames as input as against RPN and outputs cuboid boxes over each object region by computing width height and spatial co-ordinates over all the frames. This results in n boxes for each location. The videos are divided as small segments, and again fast RCNN is applied on all the frames to output refined boxes over the location and assign a classification score by calculating mean of the scores obtained over all the frames. The temporal coherency is maintained by assigning the highest score for the region class by picking the one from the pool of small video slices whose IOU crosses the threshold. The method shows an improvement in mean average precision (mAP) of 74.5%.

2.8 Comparison Table for Various Region Proposal Object Detection Methods

Table 1 shows the comparison among various detection techniques regarding the challenges taken, outcomes observed and future way to carry out the research.

3 Background Recognition Methods for Moving Target Detection

Accurate recognition of background is one of the important aspects in moving target detection in videos as there arise problems due to illumination changes, camera jitter motion and shadow formation. It also creates an ambiguity in the classification of moving objects as foreground and background objects. The other challenges are infrequent motion of objects, blurriness created due to fast motion and so on. Some

Table 1 Comparison of various detection techniques

Object/moving detection techniques	Problems addressed	Datasets used	Metrics evaluated	Results	Future scope
RCNN [1]	To project region of interest, to improve accuracy with less training data	Pascal VOC 2007 and 2012 for training and testing	Mean average precision	Improved mAP of 30% achieving 53.3% compared with best results achieved previously on VOC 2012. Improved recognition accuracy of around 91% as compared to KN, SVM, decision tree + SOG techniques	To reduce computation time and training time
RCNN + SVM classifier [2]	Detection of moving vehicles	Video frames including moving vehicles	Accuracy, precision, sensitivity, specificity		
Fast RCNN [3]	To enhance training, testing speed and detection accuracy	VOC 2007, VOC 2010, VOC 2012, CNN's Caffenet, VGGCNN, VGG16	Mean average precision	Shows best mAP improved results of around 68.5% on VOC 2012. Training period-reduced by 9 times as compared to RCNN, test time −7 times faster. FastRCNN + softmax shows more mAP	Despite training and testing speed. Fast RCNN lacked in reducing the time in selecting a region of interest. Provides scope for developing end-to-end networks
Faster RCNN [4]	To improve region projection map	VOC 2007, VOC 2012 (20 object classes)	Mean average precision	mAP increased to 73.2% on VOC 2007, 70.3% on VOC 2012	
Mask RCNN [7]	Segmenting individual instances and detecting object classes	COCO 2016	Average precision	Shows an increased AP of around 37.1, for instance, segmentation as compared to winners (MNC, FCIS) of COCO dataset	Optimization of speed and accuracy
Optical flow-based salient mask + mask RCNN-based objectness mask [8]	Unsupervised forward propagation-based motion detection	Training + testing on DAVIS 16 (challenging dataset with complex videos) and also on FBMS-59 (29-training, 30-testing)	Region similarity J, contour accuracy F and temporal stability T	Results shows improved region similarity J, contour accuracy F achieved are good. It shows a gain of 6.2% on DAVIS 2016 dataset compared with best performers on DAVIS 2016	To improve temporal stability, this method motivates more unsupervised online VOS

(continued)

Table 1 (continued)

Object/moving detection techniques	Problems addressed	Datasets used	Metrics evaluated	Results	Future scope
Spatio-temporal information extraction for moving region propsals [9]	Detecting moving objects + reduce computational cost	AVSS, CUHK datasets which consists of traffic surveillance videos	Average precision	The average precision of MPNET framework shows an increase of around 4.32% and shows a decrease of 0.02 s in terms of computation time with respect to faster RCNN on CUHK dataset	
LSTM [13]	Moving object detection	CDnet 2014–53 video sequences and LASIESTA	Average precision, F-score and recall	Enhanced robustness and spatial smoothness in terms of time cost and achieved good F-measure performance	Scope: to effectively enhance integrity of foreground objects. Increasing training data to increase performance
Bidirectional LSTM [14]	Fall movement detection in complex background	Standard UR fall dataset and self-built dataset	Average precision, F-score and recall	Average precision, recall and F-score with the values being 0.981, 0.921 and 0.948, respectively	
Cascaded RCNN [15]	High quality object detection	COCO, PASCAL VOC, KITTI	Average performance	Shows a gain of around 8% on COCO dataset for average precision	The cascaded version can be used in future detections as it shows a gain

of the techniques designed to address the issues are discussed here. The performance of the system is evaluated in terms of the reduction of false positives and negatives.

3.1 CNN-Based Background Detection

Identification of correct background is one of the crucial parts in object detection, so Patil and Murala [17] have proposed a one such novel network named as 'MSFGNet' which performs moving object detection by overcoming infrequent moving objects in a video. This network was developed in comparison with histogram-based background extraction which lacked in detecting the rare motions of the objects. The model is built as background approximation network which consists of combinational convolutional and pooling layers to estimate background by eliminating rarely moving objects. This is as depicted in Eqs. 19, 20 and 21.

$$O_t(\text{temporal pooling layer output}) = \text{average}(\text{all input pixels over } N \text{ frames}) \tag{19}$$

$$O_c(\text{convolution layer output}) = (K * O_t) + \text{bias} \tag{20}$$

where K = kernel $8 * 3 * 3$

$$O_m(\text{output of max pooling layer}) = \text{maximum}(O_c) \text{ with stride} = 8. \tag{21}$$

The obtained background is stitched to all n video frames, followed by coupling the stitched frames and original video stream frames from input side by maintaining the temporal consistency. So, it results in $2n$ video frames. The difference between original and coupled frames highlights the semantic features.

An end-to-end encoder–decoder VGG16 network is designed with reduced convolutional layers to extract the foreground objects which is capable of detecting infrequent moving objects as well. The network is trained with CD Net 2014, LASIESTA and PTIS datasets through stochastic gradient descent backpropagation method by adjusting weights. The improvement in results is shown in terms of F-measure, precision and recall parameters by comparing it with the state-of-the-art methods and shows improvement in execution time on testing set.

Zhou et al. [18] have proposed deep alignment network which incorporates base network, transformation network with an alignment estimation module for identifying the occluded objects detected in the region. Here the ResNet used as base networks captures the object features from the input image. The input image is the detected pedestrian as obtained from an automatic detector. The inaccurately detected input image is aligned through alignment model by employing affine transformation and is provided as input to transformation network which determines the effectiveness of the alignment of input image by computing loss function.

The affine transformation: (source co-ordinates) = (rotation and scaling) applied on (target co-ordinates) and bilinear sampling is applied on input features obtained from second ResNet block so as to obtain corrected version of misaligned input vectors.

This is followed by construction of association cost matrix and applying Hungarian algorithm for keeping the trajectories of the detected objects through the input frames which is described in Eqs. 22 and 23.

$$\text{Cost}(t(i)_{(n-1)}, t(i)_n) = \begin{bmatrix} a \cdot U_n^{(i)(j)} + (1-a) \cdot \text{cosine distance}(n), & U_n^{(i)(j)} > U_{\text{th}} \\ b \cdot U_n^{(i)(j)} + (1-b) \cdot \text{cosine distance}(n), & \text{otherwise} \end{bmatrix} \tag{22}$$

where

$$U_n^{(i)(j)}(\text{IOU}) = \frac{\text{area}(\text{target as detected in } (n-1)\text{th frame} \bigcap \text{target as detected in } n\text{th frame})}{\text{area}(\text{target as detected in } (n-1)\text{th frame} \bigcup \text{target as detected in } n\text{th frame})} \tag{23}$$

Kalman filter is employed for detecting occluded positions of the targets and such frames are again fed into the ResNet transformation model for reidentification. Model is trained using MOT16 (7 videos with challenging scenarios), ISSIA, PETS09, TUD. State-of-the-art comparison shows improved results for metrics. (Multiple objects tracking precision, False positives, False negatives, multi-target tracking accuracy and so on.) With faster RCNN as the detector, the proposed framework shows an increase in the accuracy of 19.7% and increased precision but doesn't show improvement over false positive reduction for which the reason was found to be the use of automatic trackers and the setting highest threshold value for occlusion. Other limitations mentioned are the model that was not able to follow the path of such targets, which resulted in false merging with two nearby targets in a detected box, in handling long-term occlusions and rapid camera movement. They proposed utilizing semantic information as a solution for handling abrupt changes in a scene.

Dominguez et al. [19] proposed a methodology of detecting pedestrians which includes linear SVM and HOG technique to first detect the pedestrians and is the detected pedestrian frames are input to optical flow method which performs the subtraction of the image sequences to obtain consecutive subtracted frames and are summed up which depicts the direction of the pedestrian movement which forms the output frame for their dataset, and this is applied to the CNN architecture which detects the person by drawing a bounding box by extracting the hidden features. The CNN architectures such as AlexNet (19), GoogleNet (19) and ResNet (20) are modified and trained as required for their application. Classification accuracy was found to be −74.8 to 83.8%. They mentioned the future scope is to train the pretrained CNN through online leraning and use of data augmentation.

Maczyta et al. [20] proposed a motion saliency approach which involves two main stages consisting first in cancelling the global motion due to the camera move-

ment, then in applying a deep learning classification framework. First stage is based on image warping and on residual flow, respectively. Experiments on real videos demonstration are carried out. The best method (RFS-Motion2D) reaches an overall accuracy of 87.5% on real dataset, and even 93.3% on the DAVIS 2016 dataset. It takes advantage of explicit motion information given by the residual flow. Future work is to improve the dominant motion, estimation for complex scenes, and to investigate the introduction of temporal regularization.

3.2 Contextual Recurrent Neural Networks (CRNN)-Based Background Estimation

Fan et al. [21] have proposed a method of encoding an image scene to overcome ambiguities in classifying background and foreground due to similarity in pixel values among the objects. The method tries to solve the same by considering three contexts local, global and topic (GIST) features to enhance the classification accuracy. Local context is implemented as acyclic directed graphs to model long range of an image. Global context is implemented by applying maxpooling technique on all 3 * 3 partitioned image which captures the entire image information in concatenated form. Image topic context is implemented by extracting GIST features which extract an image as low-level channels which includes intensity, colour, orientation and so on. All the models include RNN to extract the hidden features. The proposed model is tested on CamVid, KITTI and Cityscapes (traffic scene datasets) and Sift flow and Stanford-background (outdoor scene datasets), and the performance is verified using the metrics like pixel accuracy, class accuracy and mean intersection over union. An attention model is adopted to improve effectiveness. The method shows an increased accuracy in terms of pixel and class in comparison with previously identified methods. The limitations stated are misclassification of pixels among objects of similar appearance and missing of small objects due to pooling. This provides a scope for improving pixel and class accuracy further in terms of distinguishing similar and detecting small objects.

3.3 Bilateral Texturing Method

Nasaruddin et al. [22] proposed an attention-based method which aims at separating the background and foreground region to localize the moving vehicles in a scene by removing camera jitters, swaying movements and other unwanted noise. They created a new dataset by employing data augmentation, and bilateral texturing method is employed which includes hamming distance calculation between the adjacent frames to detect moving region which is the attention region. This is provided as an input to the YOLOv3 [deep CNN] which classifies the vehicles by drawing a bounding

box around the objects. Improved performance is seen in terms of False Positive rate (FPR), False Negative Rate (FNR), Precision and F-measure. The YOLOv3 model is trained using CDNet dataset.

3.4 Dual Background for Object Detection

Shadow detection and illumination changes that happen in the scene are one more challenge in obtaining the accurate background [23]. Dual background (DB) [24], wherein two separate backgrounds are created from incoming video frames. That is, the 'present background' is created from first frame and is upgraded recurrently at rapid rate between frames and a 'buffered background' is created by maintaining pixel equality in all those regions deprived of any changes. The process is repeated at slower rate as compared to 'present background'. The difference between the two gives the steady objects in the scene. Park et al. used this double background concept along with object masking to detect shadow regions. The segmented steady object obtained from dual background method is compared with the object mask obtained by mask RCNN for validation. If there exists a match between the background pixels obtained from DB and segmented object mask obtained from video frames, it indicates that region is steady and is regarded as shadow region.

4 Conclusion

The paper describes various deep learning-based object and moving object detection methodologies. The survey is carried out by concentrating on various region proposal techniques aimed at improving average precision to enhance the detection system performance. Referring to the comparison Table 1, the various region proposal networks focus on detecting region of interest aiming at reducing the computation time and training time. The evaluation is done through accuracy and mean average precision. Cascaded RCNN shows an improved precision of around 8% which can be considered against single stage RCNN. From the observation, spatio-temporal and LSTM-based region proposals give good results as they keep in track information from previous frames as well thus maintaining temporal coherency among the video frames. Spatio-temporal technique provides an improvement on average precision of around 4.32%. Various background detection techniques are also discussed which concentrate on overcoming the challenges like illumination changes, blurred motion, camera motion that occurs while detecting moving objects are also discussed. The paper also provides future perspective for research in moving object detection.

References

1. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 580–587. Simon MK, Hinedi SM, Lindsey WC (1995) Digital communication techniques: signal design and detection, vol 1. Prentice Hall
2. Murugan V, Vijaykumar VR, Nidhila A (2019) A deep learning RCNN approach for vehicle recognition in traffic surveillance system. In: 2019 international conference on communication and signal processing (ICCSP), Apr 2019. IEEE, pp 0157–0160
3. Girshick R (2015) Fast R-CNN. In: ICCV 2015
4. Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: towards real-time object detection with region proposal networks. Adv Neural Inf Process Syst 28:91–99
5. Dimou A, Medentzidou P, Garcia FA, Daras P (2016) Multi-target detection in CCTV footage for tracking applications using deep learning techniques. In: 2016 IEEE international conference on image processing (ICIP), Sept 2016. IEEE, pp 928–932
6. Sønderby SK, Sønderby CK, Maaløe L, Winther O (2015) Recurrent spatial transformer networks. arXiv preprint arXiv:1509.05329
7. He K, Gkioxari G, Dollár P, Girshick R (2017) Mask R-CNN. In: Proceedings of the IEEE international conference on computer vision, pp 2961–2969
8. Zhuo T, Cheng Z, Zhang P, Wong Y, Kankanhalli M (2019) Unsupervised online video object segmentation with motion property understanding. IEEE Trans Image Process 29:237–249
9. Wang H, Wang P, Qian X (2018) MPNET: an end-to-end deep neural network for object detection in surveillance video. IEEE Access 6:30296–30308
10. Huang L, Yan P, Li G, Wang Q, Lin L (2019) Attention embedded spatio-temporal network for video salient object detection. IEEE Access 7:166203–166213
11. Han G, Gao Y, Sun N (2019) Multi-target tracking based on high-order appearance feature fusion. IEEE Access 7:173393–173406
12. Kang K, Li H, Yan J, Zeng X, Yang B, Xiao T, Zhang C, Wang Z, Wang R, Wang X, Ouyang W (2017) T-CNN: tubelets with convolutional neural networks for object detection from videos. IEEE Trans Circuits Syst Video Technol 28(10):2896–2907
13. Chen Y, Wang J, Zhu B, Tang M, Lu H (2017) Pixelwise deep sequence learning for moving object detection. IEEE Trans Circuits Syst Video Technol 29(9):2567–2579
14. Chen Y, Li W, Wang L, Hu J, Ye M (2020) Vision-based fall event detection in complex background using attention guided bi-directional LSTM. IEEE Access 8:161337–161348
15. Cai Z, Vasconcelos N (2019) Cascade R-CNN: high quality object detection and instance segmentation. IEEE Trans Pattern Anal Mach Intell
16. Tang P, Wang C, Wang X, Liu W, Zeng W, Wang J (2019) Object detection in videos by high quality object linking. IEEE Trans Pattern Anal Mach Intell 42(5):1272–1278
17. Patil PW, Murala S (2018) MSFgNet: a novel compact end-to-end deep network for moving object detection. IEEE Trans Intell Transp Syst 20(11):4066–4077
18. Zhou Q, Zhong B, Zhang Y, Li J, Fu Y (2018) Deep alignment network based multi-person tracking with occlusion and motion reasoning. IEEE Trans Multimed 21(5):1183–1194
19. Dominguez-Sanchez A, Cazorla M, Orts-Escolano S (2017) Pedestrian movement direction recognition using convolutional neural networks. IEEE Trans Intell Transp Syst 18(12):3540–3548
20. Maczyta L, Bouthemy P, Le Meur O (2019) CNN-based temporal detection of motion saliency in videos. Pattern Recognit Lett 128:298–305
21. Fan H, Mei X, Prokhorov D, Ling H (2018) Multi-level contextual RNNs with attention model for scene labeling. IEEE Trans Intell Transp Syst 19(11):3475–3485
22. Nasaruddin N, Muchtar K, Afdhal A (2019) A lightweight moving vehicle classification system through attention-based method and deep learning. IEEE Access 7:157564–157573
23. Park H, Park S, Joo Y (2020) Detection of abandoned and stolen objects based on dual background model and mask R-CNN. IEEE Access 8:80010–80019

24. Singh A, Sawan S, Hanmandlu M, Madasu VK, Lovell BC (2009) An abandoned object detection system based on dual background segmentation. In: 2009 sixth IEEE international conference on advanced video and signal based surveillance, Sept 2009. IEEE, pp 352–357

Evaluation of Machine Learning Approaches for Prediction of Dengue Fever

Tasmiah Rahman and Md. Mahmudur Rahman

Abstract Dengue is a mosquito-borne, deadly viral disease that is a major threat to public health all over the world. Dengue and covid-19 symptoms are almost same, and sometimes, people are confused about which disease they are infected with. This year in Bangladesh dengue and covid-19 patients have been increasing at an alarming rate, and most of the time people didn't properly recognize the disease. A developing country like Bangladesh has faced many difficulties to handle this situation. The target of this research work is to analyze the symptoms and predict the chances to get infected with dengue fever. Machine learning techniques are widely utilized in the health industry to detect fraud in treatment at lower cost, predictive analysis, cure the disease. Four machine learning algorithms are used which are support vector machine, decision tree, *K*-nearest neighbor, random forest to predict dengue fever based on symptoms. The results were compared for percentage split and *K*-fold cross-validation method for before and after applying principal component analysis. The experimental result shows that the support vector machine algorithm provides the highest performance compared to others algorithms.

Keywords Bangladesh · Machine learning · Support vector machine · Principal component analysis · Dengue · Cross-validation · Percentage split

1 Introduction

Dengue is a mosquito-tainted viral illness that can spread rapidly under tropical climatic conditions. Dengue is primarily caused by the bite of female mosquito called "Aedes aegypti". Mainly the prevalence of dengue is increasing due to variations in temperature, rainfall and unplanned rapid urbanization. Dengue epidemic first

T. Rahman (✉) · Md. M. Rahman
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh
e-mail: tasmiah15-8805@diu.edu.bd

Md. M. Rahman
e-mail: mahmudur3720@diu.edu.bd

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_13

clinically identify in Asia, Africa, and North America occurred nearly concurrently in 1780s condition as "Break Bone Fever." 3.9 billion people all over the world are in danger of infection with dengue viruses [1]. Each year almost 50 million people are suffering from dengue fever worldwide. Today, dengue cases have multiplied all throughout the planet; regardless, the genuine numeral of dengue occurrence is either never or some of the time erroneously dispersed. Dengue virus is found in torrid and sub-torrid weather around the globe, especially in townish and semi-townish areas. Outrageous dengue is a primary wellspring of passing among kids and authentic sickness in some Latin American and Asian nations. Dengue fever has no specific treatment, but early, prediction can reduce the damage of health. Dengue is a mosquito-borne disease which is specially found in sub-tropical and tropical areas in the whole world.

This year dengue fever turned into a terribly situation in Bangladesh. The directorate general of health services (DGHSs) announced that 6450 patients are hospitalized this year and 28 people are death among them [2]. Most of the people in the country are affected in DEN-3 that is a dengue virus variant which can rise the mortality risk. Bangladesh outbreaks of DEN-1 and DEN-2 variants, previous year, but this year DEN-3 is more prevalent. However, due to the covid-19 epidemic this year, dengue has become riskier and more complex because dengue fever patients show the same symptoms like covid-19 pandemic, and it makes difficulties for diagnosis this disease [3].

Machine learning is one the important and motivating extent of research work with the legitimate of invention comprehensive knowledge from huge number of data sets. For identify hidden knowledge in health data, health care is most popular field application in machine learning. Machine learning is utilized in health industry to detect fraud in treatment at lower cost, health insurance, cure the disease. In this work, different machine learning models are utilized to acquisition the best model which predict dengue fever more accurately.

2 Related Work

Significant amount of deed was done by numerous researches all over the world on disease prediction utilizing machine learning and data mining tools. Mostly, they have used one of the two, either data mining or machine learning, to create a model, often times by utilizing several techniques of their choice. The following is a brief description of some of those works.

Permatasari et al. in their research work, they utilized quadratic discriminant analysis and mapping the occurrence within five areas in Palembang founded on different attributes, for example, gender, blood group, age, etc. 66.7% was the accurate percentage of the mapping outcome [4]. Saif et al. published an article where they utilized 1047 data. They applied support vector machine and random forest classifier and found the result is almost same which are 0.694 for support vector machine and 0.681 for random forest classifier [5]. Ramandeep et al. applied clustering in their

research work which is a part of datamining. They classified patients' information by gender, age between 25–60, and attribute. In this work, they find the maximum ratio of male gender. Dengue fever in most case rely on Rashes and Myalgia [6]. Kamran et al. did a work where they collected the dataset from DHQ hospital Jhelum, Pakistan. They utilized Naïve Bayesian, J48, random tree, and SMO algorithms and find the highest accuracy 88% from J48 and 92% from Naïve Bayesian algorithms [7]. Fatimah et al. developed a dengue fever prediction system by utilizing clinical symptoms. They utilized multilayer feed-forward neural network and find 90% accuracy of prediction [8]. Jorge et al. compared two machine learning algorithms which are ANN and SVM that help in medical diagnosis. ANN-MLP produced the best performance and achieved 96% accuracy, 97% specificity, and 96% sensitivity [9]. Abdiel et al. proposed a model to predict dengue fever outbreak where they used ANN algorithm that was trained with genetic algorithm. They discuss that ANN algorithm had an excellent predictive ability [10]. In another work, Abdul et al. utilized discriminant analysis (DA) and principal component analysis (PCA) to predict dengue possibility. DA method found the accuracy between 93 and 98%, specificity between 94 and 100%, and sensitivity between 75 and 89% [11]. Sanjudevi et al. used decision tree, support vector machine, and SMO for prediction of dengue fever in their research work. They achieved the highest accuracy from support vector machine which is 99% [12]. In another research work, permatasari et al. were used two different machine learning techniques that are artificial neural network (ANN) and random forest (RF) with ten-fold cross-validation method. They found the highest accuracy 58% from random forest classifier [13]. Shalini et al. applied decision tree, artificial neural network, and Naïve Bayes algorithms in their research work where the used 110 patient information for prediction of dengue fever. They find the highest accuracy from ANN algorithm which is 79.09% [14]. Shameem et al. utilized support vector machine for analysis dengue fever. They found 0.9042 accuracy in this research work [15]. Sungheeta et al. in their research work for detection and classification of diabetes retinopathy, they utilized image processing, SVM, Naive Bayes and proposed deep features CNN. They found 97% accuracy from proposed deep features CNN which is the highest among all methods [16]. In another research work, Milan Tripathi et al. utilized three machine learning methods which are Naive Bayes, SVM, and LSTM for detecting sentiment of covid-19 tweets analysis, where they utilized 4035 data. After analysis, they found the highest accuracy from LSTM classifier which is 79% [17].

3 Proposed Method

In this study, a several number of machine learning techniques are analyzed for dengue fever prediction. The first quest was to acquire a dataset with numerical value. Then, prepare the dataset for utilize. At the end, using percentage split, tenfold cross-validation, performance matrix, and various classifications are evaluated for finding the best model which can predict dengue fever more accurately. Figure 1 shows the

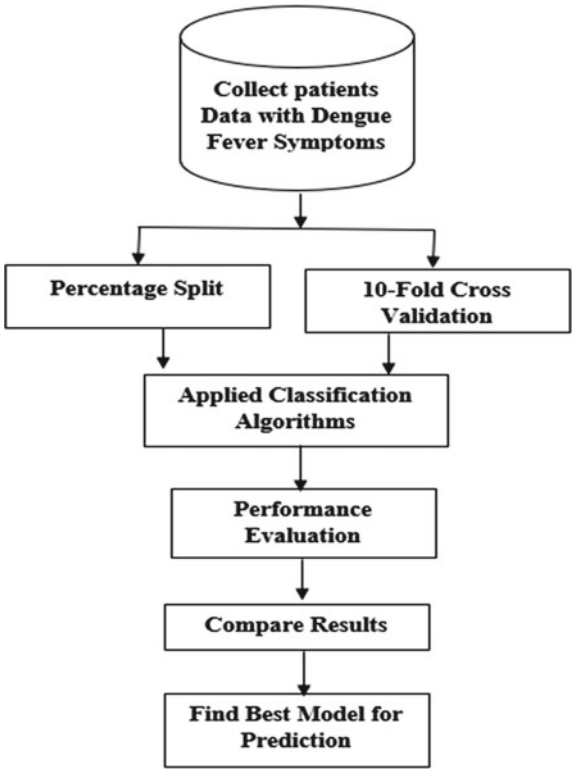


Fig. 1 Workflow of the proposed system

workflow of our system.

3.1 Dataset

Dataset is called the heart of any research work. This is one of the most important part. This dataset is collected by using survey and Internet from the people who are affected in dengue fever and who have dengue fever symptoms. Dengue fever symptoms are used for making this dataset. By surveying, we are collected 1116 people's information where 646 are not affected in dengue fever and 470 people are affected in dengue fever which is shown in Fig. 2. Here, "0" means not affected by dengue fever, and "1" means affected by dengue fever. Symptoms of dengue fever which we used to make our dataset are given in Table 1.

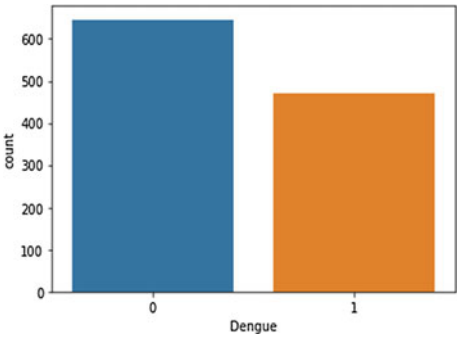


Fig. 2 Patient's distribution

Table 1 Symptoms of dengue fever

Most common	Moderate	Severe
Fever, cough, fatigue, muscle pain	Sore throat, rashes, severe headache, loss of taste, nausea, pain back of the eyes	Bleeding, low heart rate, platelet decrease, black closet

3.2 Data Preprocessing

Dataset preprocessing is necessary for making the dataset suitable to use. After preprocess, we are able to apply different machine learning methods. Real-life data are most of the time contain many errors, incomplete, and lack of many behaviors. We prepare this data by applying different preprocess method. Computer can only work with numeric value. In this work, we convert categorical values to numeric value by using level encoding. We also handle the missing value by using the mean values of that column.

3.3 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) which is a dimensionality wane technique that utilized orthogonal variation to reduce a huge set of instances to a tiny set of instances while retaining most of the information present. The procedure works by converting the highly correlated variables of the original dataset to a smaller number of uncorrelated linear instances called principal components. These principal components then account for most of the variance of the original dataset. PCA is very useful when the data have dimensions of 3 or higher since it becomes extremely hard to make predictions from such large amount of actuality. Furthermore, visualization of high dimensional data is also difficult. PCA solves this problem regarding data visualization as well. Number of principal components is less than or equal to the smaller of the number of original features or the number of observations. Initially,

the dataset comprised 1116 observations and 16 useful features, so the maximum number of principal components was 16.

3.4 Train-Test Split

This method is utilized for evaluating the performance of different machine learning algorithms; it is utilized in supervised machine learning algorithms and also utilized in classification and regression problems. In this method, the dataset is separated into two parts. The larger part is utilized for training dataset, and the smaller part is used for testing the model. In this work, we utilized 70% data for trained the model and 30% data for test the model.

3.5 K-*Fold Cross-Validation*

Cross-validation is a great way to evaluate performances estimators or machine learning techniques. The dataset is split in to two parts, one for training the estimators and the other to test its performance. After the model is trained, its cross_val_score is calculated, which is the measurement of how good of a learner that model is, K-fold cross-validation is a resampling procedure with only one parameter, K that assesses machine learning algorithms on a finite data. The data sample is to be divided in to K number of groups, hence the name K-fold cross-validation. The idea is to train the algorithm using a little from the data sample and approximate the performance of the model on unseen data, i.e., data not used in the training phase.

3.6 Algorithms

For implementation of the work, we use four different machine learning algorithms these are support vector machine (SVM), random forest (RF), decision tree (DT), k-nearest neighbor (KNN). We applied these algorithms in our dataset for both train–test split and cross-validation method. Then, compare the result by using different parameter.

Random Forest (RF). Random forest is a machine learning techniques that are utilized for performing both relapse and characterization issues. It is somewhat of an amazing and famous troupe classifier which is utilizing choice tree calculation in a randomize way. It is joined of various choice tree. To assemble each and every choice tree, it utilizes sacking method. At the point, when we group another item, we got order from each tree as tree vote. What's more, the significant decision in favor of grouping is acknowledged. That is the reason it rather gives more exact result than

single choice tree. What's more, on account of relapse takes the normal of the yields by various trees.

Support Vector Machine (SVM). It is a supervised machine learning techniques utilized in both classification and regression process. An assortment piece capacity is given by SVM to explicit choice capacities, with the chance of indicating custom portions, making them extremely adaptable. Likewise, they are highly retention productive as they utilize subsets of preparing focuses to the choice capacities. For linear classifier, SVM uses subset of training data, and it is used to represent decision boundary. The hyperplane depends of maximum width of margin, so you have to choose most optimal one hyperplane. It is not equally performed well on any unseen example.

Decision Tree (DT). One of the most broadly utilized and effective technique for classification and prediction is the decision tree. It is a tree-like development anyway changed, for instance its establishment is at the top, included different center points and branches. At the point, when an information comes in the model, decision tree shows the class of that specific info. The meaning of decision node is only test. Test is performed on the quality. To make a decision tree, we need a dataset. Then, at that point, we need to ascertain the entropy of target worth and indicator quality worth. Then, at that point, we will accumulate data of all trait lastly which characteristic has more data it would be root node. In this manner, information would be parted, and we will get our decision tree and might settle on decision.

***K* -Nearest Neighbor (KNN)**. KNN is a kind of supervised machine learning techniques that is normally utilized in classification and prediction problems. Nonetheless, this calculation could likewise be utilized as a predictive regression model. The principle thought of it is to appoint a worth to another information point dependent on the fact that it is so near different information focuses, previously existing in the dataset. By appointing the worth, KNN hence orders that specific information highlights a specific gathering closest to it. As, rather than having a particular preparing stage, this calculation utilizes the information for preparing, while at the same time arranging new information focuses, it is named as a lazy learner calculation. It is otherwise called a non-parametric calculation since it doesn't make any presumptions about the fundamental information during classification.

4 Experimental Result Analysis

For finding the expected output, I applied percentage split and *K*-fold cross-validation method. In percentage split method, I utilized 70% data for trained the model and 30% data for test the models. On the other hand, in *K*-fold cross-validation method, where, I use the value of *K* is 10. I also compare the result for before and after principal component analysis for both percentage split and cross-validation method where I use eight principal components. I applied four different machine learning algorithms to this method for finding the dengue disease more accurately. These algorithms are support vector machine (SVM), decision tree (DT), random forest (RF), *K*-nearest

neighbor (KNN). The performance of machine learning algorithms is evaluated using several performance metrics. They were accuracy, precision, recall, and f1-measure.

4.1 Evaluation Matrix

Accuracy = [(no of true positive + no of true negative)/total number]
Precision (p) = [no of true positive/predicted yes]
Recall (r) = [no of true positive/actual yes]
F1-Measure = [2 * (p * r)/(p + r)]

4.2 Compare Results

After utilizing several machine learning algorithms. We compared the result based on different parameter. In this section, we will compare the results and find the best model which is better for prediction of dengue fever disease.

In Table 2, we see that percentage method is used for analyzing the result. Here, 70% data are used for training the model, and 30% data are utilized for test the model. To compare the result, we applied four different machine learning algorithms which are support vector machine, decision tree, random forest, K-nearest neighbor. We applied these algorithms for before and after principal component analysis where we use eight principal components. Here, in the table, we see that support vector machine algorithm provides the highest accuracy which is 87.16% before applying PCA and 87.76% after principal component analysis. Precision, recall, and F1-score also justify that support vector machine is the best algorithm within all algorithms which we applied in our dataset (Fig. 3).

In Table 3, we see that k-fold cross-validation method is used for analyzing the result. Here, the value of k is 10. To compare the result, we applied four different machine learning algorithms which are support vector machine, decision tree, random

Table 2 Comparison of evaluation matrix for percentage split

		Precision	Recall	F1-score	Accuracy (%)
Without PCA	RF	0.812	0.760	0.785	82.38
	SVM	0.861	0.830	0.845	87.16
	DT	0.75	0.760	0.755	79.10
	KNN	0.852	0.774	0.811	84.77
With PCA	RF	0.804	0.781	0.792	82.68
	SVM	0.863	0.845	0.854	87.76
	DT	0.712	0.697	0.704	74.62
	KNN	0.846	0.774	0.808	84.47

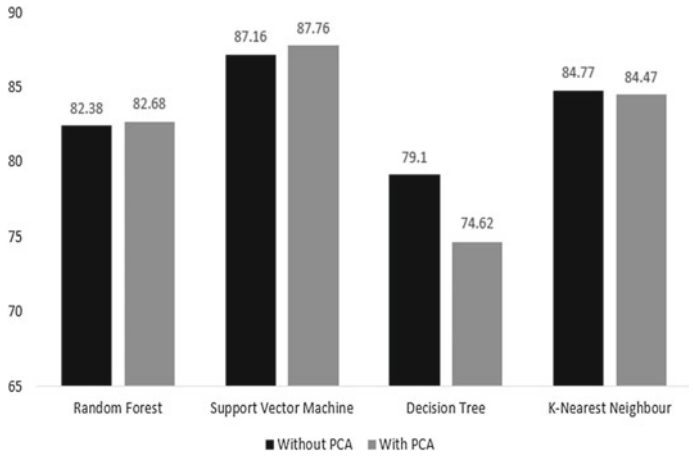


Fig. 3 Accuracy result for before and after applying PCA in percentage split method

Table 3 Comparison of evaluation matrix for tenfold cross-validation

		Precision	Recall	F1-score	Accuracy (%)
Without PCA	RF	0.809	0.768	0.788	82.31
	SVM	0.856	0.823	0.839	86.70
	DT	0.715	0.742	0.728	76.31
	KNN	0.823	0.812	0.817	84.73
With PCA	RT	0.784	0.768	0.776	80.70
	SVM	0.856	0.823	0.839	86.70
	DT	0.710	0.742	0.726	75.59
	KNN	0.823	0.812	0.817	84.73

forest, *K*-nearest neighbor. We applied these algorithms for before and after principal component analysis where we use eight principal components. Here, in the table, we see that support vector machine algorithm provides the highest accuracy which is 86.70% before applying PCA and also 86.70% after principal component analysis. Precision, recall, and F1-score also justify that support vector machine is the best algorithm among all algorithms which we applied in our dataset (Fig. 4).

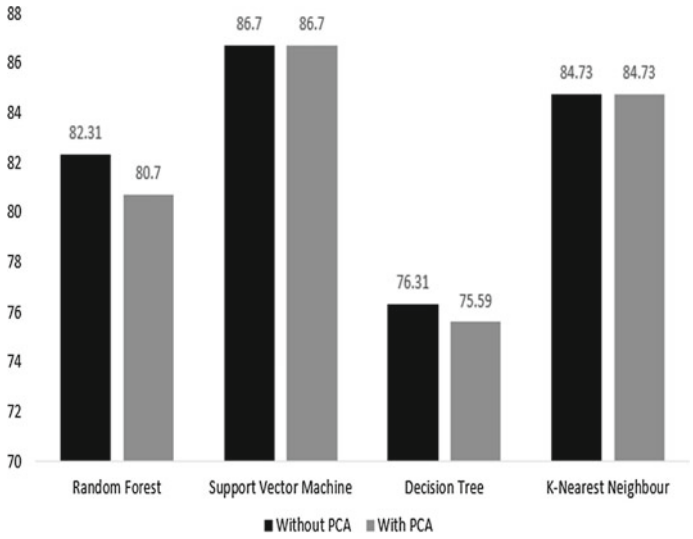


Fig. 4 Accuracy result for before and after applying PCA in tenfold cross-validation method

5 Conclusion

The target of this research work is to find out a model which can predict dengue fever at an early stage and can reduce the risk of health. If a person provides some symptoms of weakness to the model, it can predict that the person has dengue fever or not. For finding that four several machine learning algorithms are utilized which are support vector machine, decision tree, random forest, *K*-nearest neighbor. The results are compared by percentage split and cross-validation method where 70% data are utilized as trained data and 30% data are utilized for test the model in percentage split method and k-fold cross-validation method the value of k is 10. The experimental results show that support vector machine algorithms provide the highest accuracy which is 87.16% in percentage split and 86.70% for tenfold cross-validation before applying PCA. After applying eight principal components, support vector machine also provides the highest accuracy which is 87.76% in percentage split and 86.70% in tenfold cross-validation method. The precision, recall, and F1-measure values also performed better in SVM algorithm. In the future, we can use more related features. Advanced and combined algorithms to improve this model.

References

1. Dengue and severe dengue. https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue. Accessed 20 Sept 2021
2. 6,450 hospitalised this year, DGHS worried. https://en.prothomalo.com/bangladesh/6450-hospitalised-this-year-dghs-worried?fbclid=IwAR1oyBDQW7bna-bE37LF4s35wu4oQs970TC_Kg6Qy9jaNe4NDqHcpByVPz0. Accessed 20 Sept 2021
3. Dengue wreaking havoc in Bangladesh amid pandemic|Dhaka Tribune. https://www.dhakatribune.com/bangladesh/2021/07/31/dengue-wreaking-havoc-in-bangladesh-amid-pandemic?fbclid=IwAR30RcLO8ntkM1pu5svRxXPsnmujWJpMbEFnRd2lwXGlYNm_aW1CCM-zrdU. Accessed 8 Aug 2021
4. Silitonga P, Bustamam A, Muradi H, Mangunwardoyo W, Dewi BE (2021) Comparison of dengue predictive models developed using artificial neural network and discriminant analysis with small dataset. Appl Sci 11:1–16. https://doi.org/10.3390/app11030943
5. Dourjoy SMK, Rafi AMGRR, Tumpa ZN, Saifuzzaman M (2020) A comparative study on prediction of dengue fever using machine learning algorithm. Presented at the 2020. https://doi.org/10.1007/978-981-15-4218-3_49
6. Kaur R (2017) Demographic analysis of dengue fever using data mining. Int J Adv Res Comput Sci 8:1233–1237. https://doi.org/10.26483/ijarcs.v8i7.4565
7. Shaukat Dar K, Ulya Azmeen SM (2015) Dengue fever prediction: a data mining problem. J Data Mining Genomics Proteomics 06. https://doi.org/10.4172/2153-0602.1000181
8. Ibrahim F, Taib MN, Abas WABW, Guan CC, Sulaiman S (2005) A novel dengue fever (DF) and dengue haemorrhagic fever (DHF) analysis using artificial neural network (ANN). Comput Methods Programs Biomed 79:273–281. https://doi.org/10.1016/j.cmpb.2005.04.002
9. Mello-Román JD, Mello-Román JC, Gómez-Guerrero S, García-Torres M (2019) Predictive models for the medical diagnosis of dengue: a case study in Paraguay. Comput Math Methods Med. https://doi.org/10.1155/2019/7307803
10. Laureano-Rosario A, Duncan A, Mendez-Lazaro P, Garcia-Rejon J, Gomez-Carro S, Farfan-Ale J, Savic D, Muller-Karger F (2018) Application of artificial neural networks for dengue fever outbreak predictions in the Northwest Coast of Yucatan, Mexico and San Juan Puerto Rico. Trop Med Infect Dis 3:5. https://doi.org/10.3390/tropicalmed3010005
11. Poh AH, Mahamd Adikan FR, Moghavvemi M, Syed Omar SF, Poh K, Hisyam Mahyuddin MB, Yan G, Azizah Ariffin MA, Harun SW (2020) Precursors to non-invasive clinical dengue screening: multivariate signature analysis of in-vivo diffuse skin reflectance spectroscopy on febrile patients in Malaysia. PLoS ONE 15:1–22. https://doi.org/10.1371/journal.pone.0228923
12. Journal I.: IRJET—dengue fever prediction using classification techniques
13. Silitonga P, Dewi BE, Bustamam A, Al-Ash HS (2021) Evaluation of dengue model performances developed using artificial neural network and random forest classifiers. Procedia Comput Sci 179:135–143. https://doi.org/10.1016/j.procs.2020.12.018
14. Gambhir S, Malik SK, Kumar Y (2018) The diagnosis of dengue disease: an evaluation of three machine learning approaches. Int J Healthc Inf Syst Informatics 13:1–19. https://doi.org/10.4018/IJHISI.2018070101
15. Fathima AS, Manimeglai D (2012) Predictive analysis for the arbovirus-dengue using SVM classification. Int J Eng Technol 2:521–527
16. Sungheetha A, Sharma R (2021) Design an early detection and classification for diabetic retinopathy by deep feature extraction based convolution neural network. J Trends Comput Sci Smart Technol 3:81–94. https://doi.org/10.36548/jtcsst.2021.2.002
17. Tripathi M (2021) Sentiment analysis of Nepali COVID19 tweets using NB, SVM and LSTM. J Artif Intell Capsul Networks 3:151–168. https://doi.org/10.36548/jaicn.2021.3.001

Exclusive Item Recommendation to the Online Shopping Customers Based on Category Using Clickstream and UID Matrix

R. Suguna, P. Sathishkumar, and S. Deepa

Abstract Online shopping becomes indispensable among the people worldwide. Clickstream, collaborative filtering and machine learning algorithms play a considerable role to analyze the browsing behavior and predict the next click of the customers. In this research, *k*-nearest neighbor is applied to classify the customers into three groups: Regular, Special and Exceptional. User-Item-Detail matrix is constructed to identify the similarity among the online customers. Exclusive recommendation is provided to the customers based on user classification. The accuracy of the research is evaluated with the parameters precision, recall, and f-scores.

Keywords Clickstream · Item recommendation · *K*-nearest neighbor · Collaborative filtering · User item detail · Item rating

1 Introduction

The growth of e-commerce applications is abundant now-a-days. The entire world is shrinking down into homes in line for the pandemic situation. The current pandemic situation proves the importance and necessity of online shopping and its prominence. The proposed research work uses the clickstream data of the online customers to investigate their surfing actions and enhance the shopping experience. In this difficult situation, e-commerce applications rightly help the people to fulfill their essential needs. On the other hand, it is obligatory for e-commerce sites to understand the needs of the customers and maintain the relationship with them [1, 2]. It is necessary to

R. Suguna (✉) · P. Sathishkumar
Computer Science and Engineering, Bannari Amman Institute of Technology, Sathyamangalam, Erode, Tamilnadu 638401, India
e-mail: sugunar@bitsathy.ac.in

P. Sathishkumar
e-mail: sathishkumarp@bitsathy.ac.in

S. Deepa
Computer Technology-PG, Kongu Engineering College, Perundurai, Erode, Tamilnadu, India

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_14

predict the future surfing actions of the customers to recommend a set of items to them [3]. It is the ancestral role of the e-commerce applications to satisfy the customers' expectations and enhance their shopping experience. The above said is achieved by recommending a set of products to the customers who are mostly liked to them [4]. The accomplishment of the suggestion system purely stands on the exact prediction of the customer's future anticipation and recommending a set of items to them [2, 5]. Most of the e-commerce sites attempt to recommend the products to the customers by using the reviews and rating of the products. [1–3, 6] Collaborative recommendation approach also plays a major role for endorsing the most relevant products to the online shoppers [6, 7]. The clickstream information helps the e-commerce sites to identify the shopper's behavior [8–10].

The e-commerce sites monitor every click of the customers and obtain information about customer's interest, preferences and willingness [9]. Reference [2] the weblogs are stored in client machines, central machines and proxy servers. The customers can remove the client machine weblogs while clearing the browsing history [4, 5]. Weblogs store the each and every visit and movement of the Internet users which helps the e-commerce sites to know the preferences of the customers [1]. References [2, 11] the satisfaction of any recommender system depends on the quality of the items recommended to the active customers [1–4, 6]. This paper focuses on collaborative recommendation approaches along with machine learning techniques. Because the pure collaborative filtering approach has its own limitations namely sparsity and scalability issues [2]. This research recommends items to the online shoppers based on the memory based approach of collaborative approach. Memory based collaborative approach provides recommendations based on the similarity that exists between the online shoppers and items. In this approach, user similarity and item similarity is considered for future recommendation [5, 12, 13].

Figure 1 depicts the general workflow of collaborative-based recommender system using clickstream information.

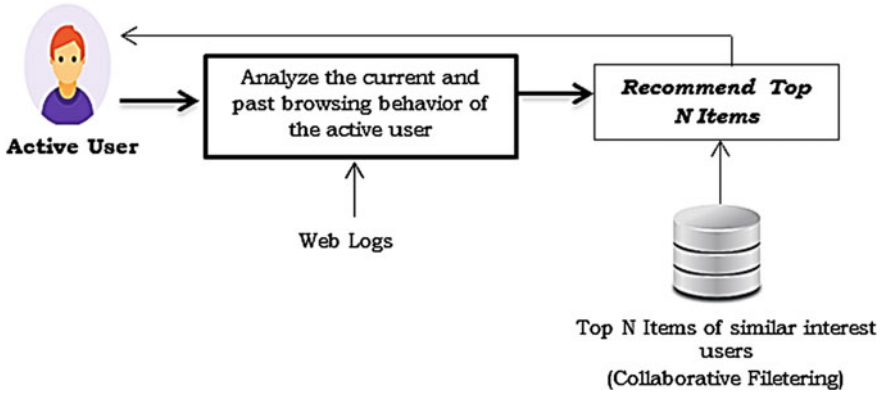


Fig. 1 Collaborative recommender system

2 Literature Review

2.1 Clickstream Data

The online shoppers browsing and buying behavior was analyzed using the clickstream information [3–5, 7–10, 14–16]. The clickstreams were collected from weblogs to know about the customer's interest and behavior.

2.2 Preprocessing

Preprocessing [1–4, 6, 15] helps to make the data suitable for investigation. The result of the investigation stage understood the customers in a better way by using the preprocessed click stream.

2.3 Machine Learning Algorithm for Next Click Prediction

The future browsing pattern of the customers are predicted by using the recurrent neural networks [6], neuro fuzzy classification [2], deep forest model [17], *k*-Clique [11] of deep learning techniques. The deep learning algorithms [10, 13, 17, 18] have proved the improved result in the research of web recommendation. This research has utilized the clickstream of online customers. The next click of the customers was predicted by using the deep learning techniques [10]. Reference [14] a framework was constructed by the authors which focuses on both clicked and not clicked items by the customers. The framework was experimented by the real time data and proven the success rate of the research. References [3, 15] Markov Chains and its algorithms were used along with click stream data of the customers to learn their switching pattern.

References [17] the users were classified into three types based on their purchase behavior namely, frequent, potential and synthetic users. Reference [1] clustering and classification techniques were used along with web usage mining to predict the customer's future behavior. References [12, 19, 20] here data mining algorithms were devised for each individual user's device to analyze and predict their future browsing pattern for recommendation. Reference [21] multivariate real time sequence analysis, [22] location-based recommender system were also proposed by the researchers to improve the future prdiction accuracy by using the clickstream data.

The existing research works have utilized the clickstream information to analyze the customer behavior and recommend the items based on the item ratings and number of clicks on the specific items. Majority of the research have utilized the item rating for recommendation. Deep learning and data mining techniques were applied for

user grouping. From the literature survey it is found that user grouping and item rating plays a major role for web recommendation in the existing research.

This research methodology is proposed to classify the customers into three major groups purely on their past purchase history not on the browsing pattern. Items are recommneded to the customers exclusively based on their past purchase behavior.

3 Proposed Methodology

The proposed methodology (UID-KNN) focuses on constructing a framework which performs (i) Click stream preprocessing (ii) User classification according to their purchase pattern (iii) Constructing User-Item-Detail for each user category and (iv) Provide an exclusive recommendation of items to the customers. The novelty of the research work is to classify the users based on their purchase history rather than item click and provide exclusive item recommendation for each classification. The algorithm classifies the users into three groups (i) Regular (ii) Special and (iii) Exceptional customers. The regular customers are one who use the e-commerce site for their day to day / monthly purchase. The e-commerce sites recommends the items of (i) same brand that they already purchased (ii) items of similar product in other brands and (iii) discouted items. The framework recommends exclusive top most items to the customers shown in Fig. 2.

This methodology helps the e-commerce sites to provide special discouts to the regular customers. The specific customers are one who makes the purchase through e-commerce sites for a specific set of items like clothes, sanitary items, groceries, etc.

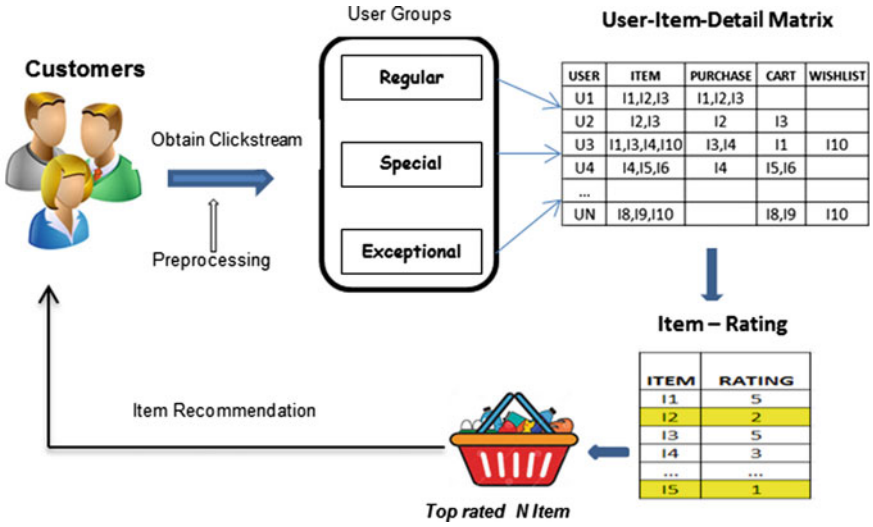


Fig. 2 Work flow of proposed mathodology

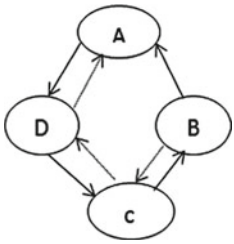


Fig. 3 Site topology

So, exclusive item recommendation is necessary for this kind of customers to enhance their shopping experience. The exceptional customers uses e-commerce sites very rarely for their purchase. This research classify all the exceptional customers in one group and recommend the exclusive items based on their past research experience. In this research approach, KNN algorithm is applied for user classification based on Euclidean distance measure.

User Identification Clickstream data is analyzed by finding the unique users. Unique users are identified either by user id or unique ip address. Let us consider $U = (u1, u2, u3 \ldots un)$ are a set of users, unique user ui is identified by ip address, agent logs and by using user id.

Session Identification The online shoppers spend time to view the products, its description and other customer reviews. Here session time is considered for login and logout of the user account. In some cases login and logout time of the device is also considered as a session. If no surfing for 30 min then it is considered as a new session. Let us consider $S = (p1, p2, p3, \ldots, pn)$, where the shopper visits the web pages $p1, p2, p3, \ldots, pn$ during the session S.

Path Completion It is an important aspect to complete the path and identify the missing path. User's traversal path plays a major role in identifying their preferences. Consider the given site topology in Fig. 3.

The authentic surfing of web page sequence is A—D—C—B—C—D—A. The user is surfing the web pages by using the back button. The web pages A, D, C are visited twice during the session. The web pages visited through back buttons are not recorded in the web server.

3.1 User Classification Using KNN Classifier

In this research users are classified into either in one group by using the KNN algorithm. KNN is a simplest classifier and random modeling, the result of future prediction is automatically recalculated based on the data sets given to the model. The proposed research uses cliclstream information which are changing in nature. The KNN modeling is well suited to handly the clickstream information and classify the customers. It uses the Euclidean distance similarity measure for classifying the

customers. The prposed research work clasifies the users according to their purchase behavior not a browsing pattern. The algorithm is as follows:

Algorithm: KNN

Input: Input the training data set
Procedure:
Find the optimal value of K
Compute a class value for new data
Begin
Calculate the distance(P,Qi); i = 1 to n
(P–New data point and Q–Items in training data set)
Sort the distances in increasing order with corrosponding training data
Choose the top K rows
Get the most frequent class from the choses K rows

End

Here the focus of the e-commerce site is regular and special customers. Exclusive product recommendation is provided to them to improve the customer satisfaction as well as to increase the sales rate of the e-commerce sites. Whenever the customer logs into the e-commerce sites, the algorithm provides exclusive item recommendations to the customers based on their classification. The top exclusive items are extracted based on the collaborative recommendation approach. The KNN algorithm is applied to classify the customers by using the Euclidean distance measure by considering the following attributes mentioned below which are obtained from the clickstream data. The algorithm uses the below formula to find the similarity between the customers.

$$(a, b) = \sqrt{(a1 - b1)^2} + (a2 - b2)^2 + \cdots + (an - bn)^2 \tag{1}$$

$$d(a, b) = \sum_{i=1}^{1} \sqrt{(ai - bi)^2} \tag{2}$$

where: a—data point, b—data point.

The attributes considered for user classification of each user $U = \{i, r, t, h, c, w, p\}$ is given in Table 1.

3.2 Construction of UID (User-Item-Detail) Matrix

This approach concentrates on the construction of a User-Item-Detail matrix rather than the common way of constructing User-Item Rating matrix. Users are classified based on the information obtained from clickstream, namely, time spent on each page and number of hits on that particular page, adding the product to the cart, adding the product to the wish list and making the payment. Once the users' classification

Table 1 Classification attributes

Attribute	Description
i	Item
r	Rating
t	Time spend on viewing the product
h	Number of hits
c	Item added to the cart
w	Item added to wish list
p	User makes the payment
j	Just surfing

is obtained, the second step is to construct the User-Item-Detail which is expressed mathematically as follows:

$$UID = U(i, r, t, h, c, w, p) \tag{3}$$

Here the User-Item-Detail matrix is constructed by including the above attributes to provide the future recommendation to the existing and current active users. Table 2 shows the User-Item-Detail matrix with the attributes. Table 3 represents the numeric conversion of Table 2. The value 1 represent Yes (Y) and 0 represent No (N) which is given in Table 3.

From Tables 2 and 3, it is observed that User 1 (U1) is viewing the items I1 and I2 and purchasing only the item I1. Likewise the second user is viewing the item I1 and not purchasing the same. He has added the items to the cart and wish list for his future transaction. User 3 (U3) is surfing the items I1, I2 and I3 and purchasing only the item I1 and adding the remaining items to the wish list and cart, respectively. In the same way, the details are recorded for all the users. Based on the purchase history of the users, they are classified. The next step of the proposed research methodology is, for each user classification, items are grouped for each user based on their purchase history for the current as well as from the past. This research considers the past six months browsing history of the users. Table 4 shows the attributes of the matrix.

Table 2 User-item-detail matrix

User ID	Item	Item rate	Time spent on the itempage	No. of Hits on the item page	Item added to the cart	Item added to the wish list	Item purchased
U1	I1	4	5	12	Y	Y	Y
U1	I2	5	6	4	Y	N	N
U2	I1	4	3	6	Y	Y	N
U3	I1	4	5	12	Y	Y	Y
U3	I2	5	6	4	N	Y	N
U3	I3	4	3	6	Y	Y	N

Table 3 User-item-detail matrix (numeric conversion)

User ID	Item	Item rate	Time spent on the itempage	No. of Hits on the item page	Item added to the cart	Item added to the wish list	Item purchased
U1	I1	4	5	12	1	1	1
U1	I2	5	6	4	1	0	0
U2	I1	4	3	6	1	1	0
U3	I1	4	5	15	1	1	1
U3	I2	5	6	4	0	1	0
U3	I3	4	3	6	1	1	0

Table 4 Attribute of item-details matrix

Attribute	Description
i	Item
ar	Average rating of an item
at	Average time spent to view the item in minutes
ah	Average number of users hit the items
c	Number of users added the item to shopping cart
w	Number of users added the item to wishlist
p	Number of users purchased the item

Table 5 depicts the Item-Detail matrix of each item hit by the users listed in table. In this matrix the average rating of each product is considered in 1–5 scale rating.

The range 5 indicates the top rating and 1 indicates the lower rating of the item. The average time spent by the users to view the item is considered in terms of minutes. Because if the users have an intention to purchase the items, they have spent more time viewing the items by seeing its descriptions, ratings, review and images. Here the average time is calculated in minutes. The next attribute is the

Table 5 Item-details matrix

Item	Avg rate (1–5, 5-Top, 1-Low)	Avg time in mins	Avg hit	Number of users adding the item to cart	Number of users adding the item to wishlist	Number of users purchase the item
I1	4	5	25	150	567	356
I2	5	2	34	145	654	345
I3	3	3	10	125	453	210
I4	4	2	8	50	230	110
I5	2	2	5	10	35	35
I6	1	2	3	0	20	30

average number of hits to the items by the user. In general users are viewing different products or items during their surfing. They may view the same item many times if they wish to purchase the item. The new technique incorporated in this research is the consideration of the number of users adding the item to wish list, cart and for purchase. These attributes help to find the more suitable items to the current active users for future recommendation. Based on the above mentioned matrix, individual sets of recommendations are provided to the customers of different categories. The customers who added the items to the shopping cart and wish list have the probability of purchasing the same or same kind of items in the future. The customers who only browse the items are considered a special category that needs special attention to tempt them to purchase the items to acquire the new customers. Consider that, let U denote user, I denote Item and D denotes details. The User-Item-Details matrix is mathematically expressed as follows:

$$M1 = \text{User} \times \text{Item} \tag{4}$$

$$M2 = \text{Item} \times \text{Details} \tag{5}$$

$$M3 = \text{User} \times \text{Item} \times \text{Details} \tag{6}$$

where $M1$, $M2$, $M3$ represent matrices. In this research instead of constructing a single user-item rating matrix, three exclusive UID matrices are constructed for regular, special and window type customers. This unique matrix construction, namely UID (Regular), UID (Special) and UID (Exceptional) for the three categories of users helps to obtain a more accurate result.

3.3 Exclusive Recommendation of Top Items to Individual User Category

After the construction of the User-Item-Detail matrix, the top N items are recommended to the users who are interested in the particular items. The below algorithm provides exclusive item recommendation to the users based on their user category.

The algorithm recommends the products exclusively to each group of customers. Consider an example: suppose if the customer purchases an electronic related item online frequently, then more electronic related top rated items are recommended to the users. And another customer is adding the cloth items to the wish list or shopping cart, then the related items are recommended to them positively. The ultimate objective is to assist the online customers to provide an exclusive recommendation of the products based on the users group.

4 Experimental Result

4.1 Data Set

The data sets of Amazon and Flipkart online shopping from kaggle. The size of the flipkart data set is 5630 KB and totally 20,568 records are available. The Amazon online shopping data set consists of 4350 KB and 18,346 customers weblogs. The dataset consists of ipaddress, username, session time, sequence of clikcks made by users on a particular session. The datasets are extracted in.csv files with comma separator. The work is implemented in python. In order to evaluate the recommendation precision, the data sets are divided into training and testing dataset. 70% of the data is applied for training sets and 30% is for testing. 14,397 records from Amazon and 12,842 records from Flipkart are considered to train the classifier.

4.2 Evaluation Metrics and Performance Analysis

In this research, the accuracy of the classification and recommendation is experimented by using The Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The lowest values of MAE and RMSE show the improved performance in the recommendation system. Finally the system is assessed by applying the quality parameters namely, recall, precision and f-measure.

MAE—Produce the average of all the absolute value differences between the true and the predicted rating. The lower the MAE the better the accuracy.

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{n} |\mathrm{d}i| - |\widehat{\mathrm{d}i}| \tag{7}$$

where di—actual user group, di'—predicted user group and n—number of users.

RMSE—This formula calculates the mean value of all the squared differences of actual and predicted user classification of KNN algorithm. The large number of error ratings of RMSE leads to poor results in classification.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n} (\mathrm{d}i - \widehat{\mathrm{d}i})^2} \tag{8}$$

The proposed methodology UID-KNN is compared with Item Rating (IR-KNN) and User-Item Rating (UIR_KNN) algorithms. Table 6 shows the improved performance of the proposed methodology for both the data sets for MAE and MRSE.

Figures 4 and 5 depicts the improved performance of UID-KNN algorithm when compared to the existing research which based on item rating. The prposed algorithm

Table 6 Performance result based on MAE and MRSE

Methodology	Amazon		Flipkart	
	MAE	MRSE	MAE	MRSE
IR-KNN	1.283	1.283	1.287	1.284
UIR-KNN	1.283	1.282	1.286	1.283
UID-KNN	1.234	1.242	1.232	1.243

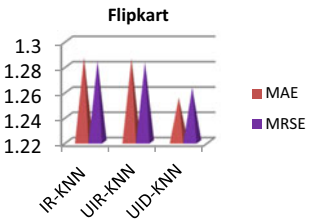


Fig. 4 Performance result (Flikart)

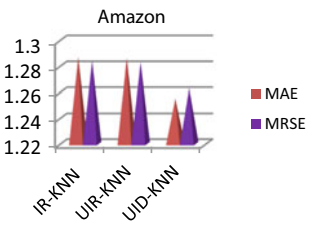


Fig. 5 Performance result (Amazon)

produces the lower value of MAE and MRSE, 1.23 and 1.24, respectively, in both the data sets.

To proves the accuracy of the recommendation system with the standard parameters precision, recall and f-measure. The higher values of the parameters show the highest accuracy of the top N recommendation system. The below are the metrics to evaluate the accuracy of the proposed recommendation system.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} * 100 \tag{9}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} * 100 \tag{10}$$

$$\text{F-measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

where TP—True positive, FN—False negative.

Table 7 describes the accuracy of the recommendation system with the parameters precision, recall and f-measure for Amazon and Flipkart data sets.

Table 7 Performance result based on F-Score

Methodology	Amazon			Flipkart		
	Precision	Recall	F-measure	Precision	Recall	F-measure
IR-KNN	0.786	0.782	0.78	0.787	0.783	0.78
UIR-KNN	0.887	0.865	0.88	0.886	0.865	0.88
UID-KNN	0.987	0.983	0.98	0.988	0.983	0.99

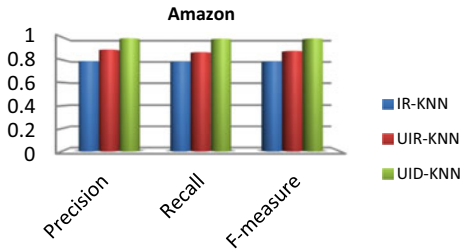


Fig. 6 Performance of recommendation (Amazon)

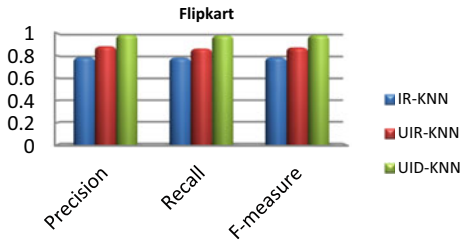


Fig. 7 Performance of recommendation (Flipkart)

Figures 6 and 7 depict the obtained values of the research work for the parameters precision, recall and accuracy. The exclusive recommendation system by using various parameters and classification algorithm proves its ameliorations by giving realistic experimental results in the adequate parameters through Figs. 6 and 7 with the exactness rate of 98% in Amazon data set and 99% in the Flipkart data set in f-score values.

5 Conclusion

It is very interesting and difficult to know the behavior and future surfing of the online customers. This research aims to introduce a new recommender system which understands the expectations of the online customers by using their clickstream information. Here the exclusive recommendation is done based on the collaborative filtering approach and User-Item-Detail matrix. The new methodology proves the recommendation accuracy rate 98%. The existing research works used the Item

Rating matrix which gave the accuracy rate of 88% The proposed algorithm gives better accuracy to classify the users as well as for recommending top *N* items to the customers. In future, a customer specific recommendation system by using a deep learning approach may be developed to advance the customer's satisfactory level.

References

1. Anandhi D, Ahmed M (2019) Prediction of user's type and navigation pattern using clustering and classification algorithms. Cluster Computing 22. https://doi.org/10.1007/s10586-017-1090-2
2. Dash S, Luhach AK et al (2019) A neuro-fuzzy approach for user behavior classification and prediction. Journal of Cloud Computing 8:17. https://doi.org/10.1186/s13677-019-0144-9
3. Chanyoung P, Donghyun K, Min-Chul Y, Jung-Tae L, Hwanjo Y (2020) Click-aware purchase prediction with push at the top. Information Sciences 521:350–364. https://doi.org/ https://doi.org/10.1016/j.ins.2020.02.062
4. Chi Y, Jiang T, He D, Meng R (2017) Towards an integrated clickstream data analysis framework for understanding web users' information behavior. In: International conference proceedings, pp 279–292. https://doi.org/10.9776/17027
5. Chu lY, Yang HK, Peng WC (2019) Predicting online user purchase behavior based on browsing history. In: 2019 IEEE 35th international conference on data engineering workshops (ICDEW). pp 185–192. https://doi.org/10.1109/ICDEW.2019.00-13
6. Toth A, Tan L, Di Fabbrizio G, Datta A (2017) Predicting shopping behavior with mixture of RNNs. SIGIR 2017 eCom. Tokyo, Japan (2017)
7. Esmeli R, Bader-El-Den M, Abdullahi H (2021) Towards early purchase intention prediction in online session based retailing systems. Electron Markets 31:697–715. https://doi.org/10.1007/s12525-020-00448-x
8. Filvà DA, Forment MA, García-Peñalvo FJ, Escudero DF, Casañ MJ (2019) Clickstream for learning analytics to assess students' behavior with scratch. Futur Gener Comput Syst 93:673–686. https://doi.org/10.1016/j.future.2018.10.057
9. Houda Z, Adil H, Mohammed R (2019) A novel approach to dynamic profiling of E-customers considering clickstream data and online reviews. IJECE 9(1):602–612. https://doi.org/10.11591/ijece.v9i1.pp602-612
10. Koehn D, Lessmann S, Schaal M (2020) Predicting online shopping behaviour from clickstream data using deep learning. Expert Syst Appl 150:113342. https://doi.org/10.1016/j.eswa.2020.113342
11. Samuel M (2020) Patient diet recommendation system using K-clique and deep learning classifiers. J Artif Intell 2(2):121–130
12. Haoxiang W, Smys S (2021) Big data analysis and perturbation using data mining algorithm. Journal of Soft Computing Paradigm (JSCP) 3(01):19–28
13. Kottursamy K (2021) A review on finding efficient approach to detect customer emotion analysis using deep learning analysis. JTCSST 3(2):95–113
14. Requena B, Cassani G, Tagliabue J et al (2020) Shopper intent prediction from clickstream e-commerce data with minimal browsing information. Sci Rep 10:16983. https://doi.org/10.1038/s41598-020-73622-y
15. Swapna DK (2017) Analysis of clickstream data using Markov chains. In: Seventeenth AIMS international conference on management, pp 1135–1137
16. Yao S, Yoo H, Sun L, Du X (2019) Using machine learning to address customer privacy concerns: an application with click-stream data. SSRN Electron J. https://doi.org/10.2139/ssrn.3314787

17. Zhang W, Wang M (2021) An improved deep forest model for prediction of e-commerce consumers' repurchase behavior. PLoS ONE 16(9):e0255906. https://doi.org/10.1371/journal.pone.0255906
18. Senthilkumar T (2021) Construction of hybrid deep learning model for predicting children behavior based on their emotional reaction. Journal of Information Technology 3(01):29–43. https://doi.org/10.36548/jitdw.2021.1.004
19. Karthigaikumar P (2021) Industrial quality prediction system through data mining algorithm. JEI 3(2):126–137. https://doi.org/10.36548/jei.2021.2.005
20. Suma V, Shavige MH (2020) Data mining based prediction of demand in Indian market for refurbished electronics. JSCP 2(2):101–110. https://doi.org/10.36548/jscp.2020.2.007
21. Guo L, Zhang B, Zhao X (2021) A consumer behavior prediction model based on multivariate real-time sequence analysis. Math Probl Eng. https://doi.org/10.1155/2021/6688750
22. Vijesh C, Jennifer SR (2021) Location-based orientation context dependent recommender system for users. JTCSST 3(01):14–23. https://doi.org/10.36548/jtcsst.2021.1.002

Android Malware Detection Using Machine Learning Classifiers

Ajay Bandi and Lunduk Sherpa

Abstract The Android operating system has changed what mobile can do. The population of Android mobile users has been increasing, and the number continues to grow as the number of mobile users increases. Primarily, the applications used in Android devices are distributed via the Google Play Store. These applications are required to meet several criteria in order to be distributed via the Play Store. However, attackers sometimes find their way to compromise the Android operating system devices and steal users' sensitive information. Malicious software or malware is designed to extract data from users by compromising their devices for financial gain or other reasons. This study uses machine learning algorithms to detect Android malware in the Android malware dataset. Using CICAndMal2017 dataset, random sampling was done to extract a balanced dataset. Feature engineering was performed to obtain the most significant features. Machine learning algorithms were trained using the balanced dataset with selected features. Initially, all the models were trained using 'Label' as the target variable and then 'Family' as the target variable. Maximum accuracy of approximately 99% was obtained using random forest in both cases.

Keywords Android malware detection · Malware · Machine learning algorithms · Feature selection

1 Introduction

Nearly half of the global population use smartphones. Users tend to store private information on their smartphones. The security of users' information requires a top priority. Despite all the measures used to keep the information secure, malware attacks are still prevalent. There are numerous types of malware and techniques to

A. Bandi (✉) · L. Sherpa
Northwest Missouri State University, Maryville, MO 64468, USA
e-mail: ajay@nwmissouri.edu

L. Sherpa
e-mail: s524425@nwmissouri.edu

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_15

inject them into the target system. Some of the most common and dangerous types of malware are Adware, Ransomware [4], and Scareware. Depending on their purpose and method of delivery, they vary from one another.

The Android security sometimes fails to detect malicious software and is distributed through the official Play Store. Users' information is compromised through some programs distributed through the Play Store [8]. Android, being one of the most popular smartphones operating systems [15], has gained many attackers' attention. Thousands of malware applications have been discovered in recent years. With every update in the security features, the attacks get even more troublesome to detect. Machine learning and other techniques [13, 14, 16, 17] are used to detect Android malware. Machine learning algorithms can be trained to formalize the principles hidden in the dataset used to train the model. A properly trained model can reason the features of previously unseen dataset [7]. Based on the model's approaches, each model has a different capacity and problems they suit their best.

This paper is organized as follows: Sect. 2 includes the proposed work. Section 3 includes the related work where we discuss various approaches taken for Android malware detection. Section 4 includes the dataset description. Section 5 discusses the procedure used in this study. Section 6 discusses the various machine learning algorithms used in this study. Section 7 describes the results and analysis of our study. Section 8 discusses the limitations and future work. In Sect. 9, we present the conclusion of our study.

2 Proposed Work

In this study using the CICAndMal 2017 dataset, machine learning algorithms are trained to detect Android malware. The model classifies the malware into one of the four types included in our dataset: Adware, Scareware, Ransomware, and SMS malware. The overall project flow is shown in Fig. 1. Using five different machine learning algorithms, supervised as well as unsupervised, we train our model using dataset obtained post feature engineering. Without compromising the accuracy of the models used, most significant features are extracted, and their performances are compared.

3 Related Work

Using cloud technology in edge computing and traffic clustering, He et al. [9] developed a new method for identifying Android malware. To begin, the authors routed traffic between Android mobile devices to edge nodes. The edge server delivers to the cloud platform mobile traffic content features in plain text extracted from HTTP flows, as well as traffic behavior such as packet intervals. To detect malware, they utilized a method to automatically evaluate the commonalities between apps and

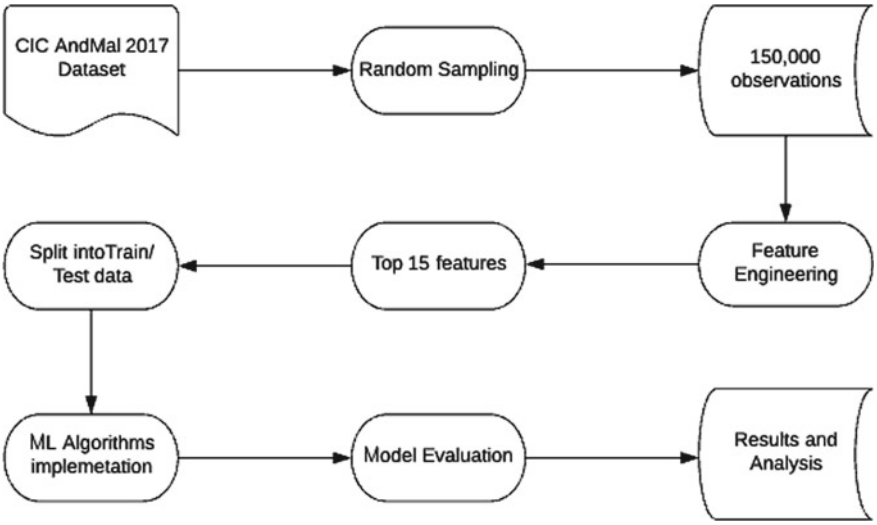


Fig. 1 Proposed methodology

clusters. Malware was identified using similarity algorithms such as TF-IDF and cosine. A total of 400 Android applications was employed to evaluate their method, yielding an average accuracy of 96.9.

Bibi et al. [5] used a deep learning framework based on the long short-term memory (LSTM) to detect Ransomware malware. To enhance Android malware detection and family classification using conversation level, this study chose 609 categories from the CICAndMal2017 dataset, including benign and Ransomware data. They used eight feature selection algorithms, such as Chi-Square and information gain, to select the top 19 flow-level features. A model was built using those features with an accuracy of 97%.

The authors Chen et al. [6] used a portion of the CICAndMal2017 dataset, selecting one PCAP file containing packet data of the network for each malware family. Their samples were chosen randomly, and features were extracted from PCAP files. A Java program was created that used the flow-level technique to separate network flows. Then, using a Python program, fifteen features were extracted. The minimum size of the sent packet within a flow was one of the features. There are three supervised machine learning classifiers: K-nearest neighbors, random forest, and decision tree. Data were divided into two categories: malicious and non-malicious. Following that, the malware was classified into three sub-categories: Adware, Ransomware, and Scareware. Recall, precision, and $F1$-score are the three metrics used by the authors to measure the model's performance. The results show that the random forest classifier had the best results for malware-benign classification, with an $F1$-score of 92% and precision and recall of 95%. The other classifiers accounted for more than 85% for all used metrics.

Arora and Peddoju [2] prioritize network traffic elements of Android malware based on IG and Chi-Square tests. Internet traffic features were eliminated using their proposed technique to increase accuracy results and simplify the training and testing stages. Statistical analysis approaches were employed to rate the features. Nine out of 22 parameters are adequate for better detection accuracy, according to the suggested algorithm. Similarly, the study findings reveal that both training/testing times could be reduced by 50% and 30%, respectively. In this study, we also attempted to train our models using two target variables to predict the label as well as family of malware. We were unable to train some of the models proposed in this study, and the training time increased significantly.

4 Datasets

The dataset used in this study was obtained from the University of New Brunswick's website. The CICAndMal2017 dataset [12] is used to perform our analysis. This dataset contains four different types of malware. Both malware and benign applications were installed on real smartphones to generate the dataset. Three states of data capturing were defined to acquire a comprehensive view of malware samples, viz., installation, before restart, and after restart. CICFlowMeter is a network traffic flow generator and analyzer that has been used for anomaly detection and in various cybersecurity datasets. During all three states, 85 features are extracted using CICFlowMeter-V3. The dataset downloaded from the website contained several Excel files containing 42 different families of malware, including benign records. In order to prepare the data for curation, all the Excel files were merged using R-Studio. An additional feature was included in the dataset so that the dataset contains a separate column for malware type and family. Four different types of malware analyzed in this study are:

1. Adware: This malware bombards the user with pop-up ads. Adware hides in the user device and serves ads. Adware can also monitor users' behavior online and serve ads accordingly. In our dataset, there are ten different families of Adware, namely Dowgin, Ewind, Feiwo, Gooligan, Kemoge, Koodous, Mobidash, Selfmite, Shaunet, and Youmi.
2. Ransomware: Ransomware blocks access to the victim's data or computer files and demands a ransom to restore access to the victim. This could result in loss of sensitive information and can cause costly disruption to operations [11]. The dataset used in this study consists of ten different families of Ransomware, namely Charger, Jisut, Koler, LockerPin, Simplocker, Pletor, PornDroid, RansomBO, Svpeng, and WannaLocker.
3. Scareware: This malware manipulates users by using fake virus warning pop-ups and other tricks to gain access to users' personal information. Scareware makes users believe that their computer is infected by viruses and coaxes them to install unnecessary software. Eleven different families of scareware are included

in the dataset used in this study. Namely, they are AndroidDefender, AndroidSpy.277, AV, AVpass, FakeApp, FakeApp.AL, FakeAV, FakeJobOffer, FakeTaoBao, Penetho, and VirusShield.

4. SMSMalware: This malware exploits the victims' mobile device via SMS text. Without users' consent, this malware can cause several detrimental effects. The attacker can gain access to users' personal information. In addition, attackers can also change, delete, expose, and misuse users' data. In this study, there are 11 different families of SMSMalware, namely BeanBot, Biige, FakeInst, FakeMart, FakeNotify, Jifake, Mazarbot, Nandrobox, Plankton, SMSsniffer, and Zsone.

5 Data Curation

The raw data obtained directly from the source must be processed before performing any analysis. The data quality determines how much information can be extracted from the data. We performed random sampling to obtain a balanced dataset from the raw data, and using data cleaning tools, we cleaned the data for further analysis.

5.1 Random Sampling

The raw dataset was not balanced and contained multiple CSV files for each type of malware family. A balanced dataset including all types of malware is required for our analysis. In order to obtain a balanced dataset, a CSV file was created for each malware family. A total of 42 different CSV files was created. Random sampling has been performed on these CSV files to obtain a sample of 3000 records from each family, including benign records. Samples obtained after random sampling are merged to create a balanced dataset containing 156,000 observations. No weightages have been assigned to attributes in our dataset. Figure 2 shows the random sampling framework for malware dataset.

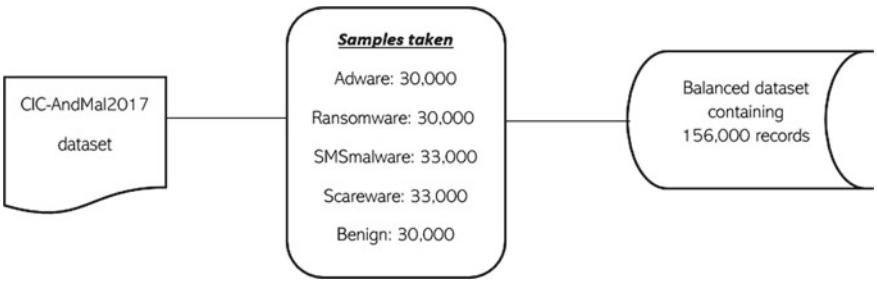


Fig. 2 Sampling framework

5.2 Data Cleaning

Data cleaning is an essential aspect of the data science life cycle. For any data to be processed for analysis, it is important to clean the data. Data cleaning is the process of removing unnecessary noise from the data. Prior to data cleaning, datasets usually contain missing values, incomplete or incorrectly formatted data, varying data types, etc. The dataset obtained after random sampling contained several missing values. Some features required formatting in order to proceed further for analysis. In this study, we performed the following processes to clean the dataset:

1. Special characters were removed from the following features:
   (a) Source.IP
   (b) Destination.IP
   (c) Timestamp
2. Rows with missing values were removed.
3. All infinity values were first replaced with the Nan values, and all the Nan values were filled with 0.
4. All the features datatype were formatted to Float.
5. The values in the target feature 'Label' and 'Family' were replaced with numerical values.
6. Single value features were removed from the data frame using variance threshold.
7. Rows containing Nan as the value for 'Label' and 'Family' were removed.

5.3 Feature Selection

The initial and most crucial step in our model design should be feature selection. The process of selecting the appropriate attributes that have a substantial impact on anticipating our output is known as feature selection. The key benefits of feature selection are lowering overfitting, improving accuracy, and shortening training time. Figure 3 shows the feature selection framework. The univariate feature selection technique selects the most suited features based on the univariate statistical tests. We will compare all attributes to the target variable to see a significant association between them, which is what an analysis of variance, or ANOVA, is all about. For the scoring function, we have chosen f classif. Using Scikit-learn library, SelectKBest, and f_classif modules, scores for all features are calculated. The top 15 features selected by univariate selection are shown in Fig. 4.

Dataset obtained after univariate feature selection contains 156,000 records and 15 most significant features. Prior to training our model, we split our dataset into train and test set at 4:1 ratio. The train set is used to train all the models proposed to be used in this study. We use modules from Scikit-learn library [10] to train our models and evaluate their performance metrices.

Fig. 3 Feature selection framework

Fig. 4 Top 15 features selected by univariate selection

	Specs	Score
5	Timestamp	27707.365004
4	Protocol	142.148710
28	Fwd.IAT.Max	140.341282
25	Fwd.IAT.Total	129.423231
3	Destination.Port	127.106005
23	Flow.IAT.Max	115.951485
35	Fwd.PSH.Flags	114.141354
49	SYN.Flag.Count	114.141354
2	Destination.IP	104.574873
6	Flow.Duration	99.293089
27	Fwd.IAT.Std	94.692103
81	Idle.Max	93.185653
1	Source.Port	89.198627
79	Idle.Mean	85.681873
22	Flow.IAT.Std	85.226464

6 Machine Learning Algorithms

This section introduces various machine learning algorithms used to train the datasets.

A **decision tree** is a decision-making tool that predicts alternative outcomes using a tree-like decision-making paradigm. It covers the consequences of events, resource costs, and decision utility. Decision trees seem like a flowchart or an algorithm with just conditional control statements.

Random forest is an ensemble learning technique for classification, regression, and other operations in machine learning that uses a large number of decision trees during training. They are quick and versatile, and they are a good way to mine high-dimensional data. Random forest expands the classification and regression techniques we explored earlier using decision trees. The times for random woodland runs are quite quick. They are pretty excellent at dealing with incomplete or incorrect data. On the downside, they cannot forecast beyond the extent of the training data, and they are prone to over fitting noisy datasets.

The **K-nearest neighbor (kNN)** technique is a straightforward supervised machine learning approach for classification and regression. Using a similar metric, the distance function, KNN keeps track of available inputs and categorizes new data. Statistical estimation and pattern recognition are two of KNN's essential uses. The KNN algorithm calculates the distances between two requests.

Naive Bayes is another very effective and commonly used machine learning classifier. The Naive Bayes algorithm family encompasses both supervised and unsu-

pervised learning techniques. The Bayes Theorem is used to create the Naive Bayes classifiers, a set of classification algorithms based on the Bayes Theorem. It is a group of algorithms with a familiar concept: Each pair of features to be categorized is distinct from the others. The most common application of Naive Bayes is to estimate the likelihood of distinct groups based on a range of attributes. It is primarily used for text classification in data mining.

In **stochastic gradient boosting**, individual models are not entirely random subsets of data and characteristics. They are developed in order, with cases with inaccurate predictions and significant mistakes receiving more weight. The primary idea is that during learning, occurrences that are difficult to predict accurately ('difficult' circumstances) would be concentrated on so that the model can learn from previous errors. We call this SGB when we train each ensemble on a portion of the training set, and it can help improve the model's generalizability.

In this research project, a randomly sampled balanced dataset was taken, then cleaned using a few data cleaning techniques. We performed the feature selection techniques to determine the best optimal features to be given as input to the machine learning algorithms and train them. Considering those top 15 features, we implemented all the machine learning algorithms by taking the 'Label' feature as the target variable. After evaluating the machine learning algorithms by calculating the 'accuracy', '$F1$-score', 'precision', and 'Recall', we repeated this whole process taking the 'Family' feature as the target variable. Our machine learning models will calculate what kind of attack it is or is it benign. In addition to that, it will classify which family makes the malware attack belongs. The libraries that are used to to implement the ML algorithms are: Sklearn, Matplotlib, Pandas, NumPy.

7 Results and Analysis

The features obtained after univariate feature selection were used to train our models. The dataset was split into train and test data. We trained all the models using the train set, and their corresponding performance was evaluated using measures like accuracy, precision, recall, and $F1$-score. To assess all our data and ensure all the observations have been used to evaluate our models, we performed K-fold cross-validation. We computed the average accuracy for all the models. Bibi et al. [5] have done similar classification on CICAndMal2017 data. The model proposed using neural nets in [5] obtained an accuracy of 97.08% which did not surpass the maximum accuracy we obtained using decision tree and random forest. Abuthawabeh and Mahmoud [1] performed similar classification on the malware detection, malware type categorization, and malware family classification. For all three classifications, the extra trees classifier achieved the highest weighted accuracy percentage of 87.75%, 79.97%, and 66.71%, respectively.

Using univariate feature selection, we obtained the top 15 features. Prior to that, we trained our models using all the features. Our models performed better using the top 15 features. Table 1 shows the results obtained using the 'Label' feature as

Table 1 Using 'Label' as the target variable

ML model	Accuracy (%)	Precision (%)	Recall (%)	$F1$-score (%)
Decision tree	99.63	99.63	99.63	99.63
Random forest	99.54	99.54	99.54	99.54
K-nearest neighbors	93.67	93.67	93.67	93.67
Naïve Bayes	61.80	61.80	61.80	61.80
Stochastic gradient boosting	94.07	94.07	94.07	94.07

Table 2 Using 'Family' as the target variable

ML model	Accuracy (%)	Precision (%)	Recall (%)	$F1$-score (%)
Decision tree	99.01	99.01	99.01	99.01
Random forest	99.08	99.08	99.08	99.08
K-nearest neighbors	62.00	62.00	62.00	62.00
Naïve Bayes	44.42	44.42	44.420	44.42
Stochastic gradient boosting	62.25	62.25	62.25	62.25

our target variable. Models were trained to classify into five different classifications, including benign. We obtained more than 90% accuracy from all the models used except Naïve Bayes. One of the possible reasons why Naïve Bayes showed low performance is the assumption of equally important and independent features that the algorithm relies on. Table 2 shows the results obtained using the 'Family' features as our target variable. Unlike Table 1, in this case, models were trained to classify into 43 different classifications, including benign. Decision tree and random forest performed best of all with approximately 99% accuracy. The details on the code work and the documents used for this analysis are uploaded in our GitHub repository [3].

8 Limitations

During this work, we encountered issues working with the target variables. When we trained our ML models with two target variables, namely 'Label' and 'Family', we were not able to evaluate those multi-class multi-output scenarios. This was the major constraint we encountered while working on this study. One-hot encoding technique for feature scaling could not be executed due to the large volume of data present in our dataset. Despite taking only 15 features, neural net's training time was considerably long, so we did not include neural net as our model.

9 Conclusions

In this study, we used random sampling to generate 156,000 records balanced dataset. By performing a feature selection technique, we came up with the ML algorithms' top 15 features given as training datasets. After training those ML models and evaluating the results, we observed that decision tree, random forest, and stochastic gradient boosting algorithms had performed well with around 99% accuracy when taking 'Label' as the target variable. Taking the 'Family' feature as the target variable, only the decision tree and random forest provided better results with 99.01 and 99.08% accuracy. Decision tree and random forest can be used to detect Android malware by installing a system that has this model installed. This could save operation costs and prevent data from being lost.

Comparatively, Android is one of the most famous mobile OSs in the world. This also tends to give rise to android's vulnerability to malware. Due to open-source OS, it is easier for hackers to tamper with devices running Android OS. Similar study [3] could be done on platforms like iOS depending on what data features are available.

References

1. Abuthawabeh M, Mahmoud K (2020) Enhanced android malware detection and family classification, using conversation-level network traffic features. Int Arab J Inf Technol 17(4A):607–614
2. Arora A, Peddoju SK (2017) Minimizing network traffic features for android mobile malware detection. In: Proceedings of the 18th international conference on distributed computing and networking, pp 1–10
3. Bandi A. https://github.com/bandiajay/AndroidMalwareAnalysis
4. Bandi A, Fellah A, Bondalapati H (2019) Embedding security concepts in introductory programming courses. J Comput Sci Coll 34(4):78–89
5. Bibi I, Akhunzada A, Malik J, Ahmed G, Raza M (2019) An effective android ransomware detection through multi-factor feature filtration and recurrent neural network. In: 2019 UK/China emerging technologies (UCET). IEEE, pp 1–4
6. Chen R, Li Y, Fang W (2019) Android malware identification based on traffic analysis. In: International conference on artificial intelligence and security. Springer, pp 293–303
7. Chumachenko K (2017) Machine learning methods for malware detection and classification
8. Enck W, Octeau D, McDaniel PD, Chaudhuri S (2011) A study of android application security. In: USENIX security symposium, vol 2
9. He G, Xu B, Zhang L, Zhu H (2020) On-device detection of repackaged android malware via traffic clustering. Secur Commun Netw 2020
10. https://scikit-learn.org/stable/
11. https://www.fbi.gov/scams-and-safety/common-scams-and-crimes/ransomware
12. https://www.unb.ca/cic/datasets/andmal2017.html
13. Ou F, Xu J (2022) S3Feature: a static sensitive subgraph-based feature for android malware detection. Comput Secur 112:102513
14. Sahs J, Khan L (2012) A machine learning approach to android malware detection. In: 2012 European intelligence and security informatics conference. IEEE, pp 141–147
15. Worldpanel K (2014) Smartphone OS market share
16. Yuan Z, Lu Y, Wang Z, Xue Y (2014) Droid-Sec: deep learning in android malware detection. In: Proceedings of the 2014 ACM conference on SIGCOMM, pp 371–372
17. Zarni Aung WZ (2013) Permission-based android malware detection. Int J Sci Technol Res 2(3):228–234

Approaches for Detection of Diabetic Retinopathy: A Review

B. Sowmyashree, K. Rao Mahesh, and H. K. Chethan

Abstract Fundus images are mainly used for the screening and detection of retinal diseases. Defect in the eye's posterioir blood vessel causes Diabetic retinopathy. Proliferate Diabetes Retinopathy and Proliferate Diabetes Retinopathy are the two stages of diabetes that develops over the time. The retinal fundus image is commonly used in clinics to diagnose and treat a variety of eye disorders. It is also one of the most important resource for diabetic retinopathy mass screening. In this paper, a survey has been made on various techniques available in image processing such as morphological operations and texture analysis to identify diabetic retinopathy. This survey extended to investigate machine learning and deep learning approach to extract various features namely, area of hard exudates, microaneurysms and blood vessels to identify and detect the depth of the disease. Finally, the review has provided a summary and analysis of most of the recent techniques adapted to detect diabetic retinopathy.

Keywords Blood vessels · Deep learning · Diabetic retinopathy · Machine learning · Morphological operation

1 Introduction

One of the primary difficulties of modern health care is the rapid growth of diabetes. The number of persons suffering from the condition is increasing at an alarming rate. Over the next 25 years, the World Health Organization predicts that the population of diabetes would rise from 130 to 350 million [1]. According to the survey, only 50% of

B. Sowmyashree (✉)
Department of Electronics and Communiction, MIT, Mysore, India
e-mail: sowmyashreebece@gmail.com

K. R. Mahesh
MIT, Mysore, India

H. K. Chethan
Department of Computer Science, MIT, Mysore, India

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_16

the patients aware of the condition aggravatesthe issue. Diabetes also has significant late consequences from a medical standpoint. Heart disease, renal problems, and retinopathy are just a few of the complications that might arise as a result of macro and microvascular changes. Diabetes, for example, is the fifth-deadliest illness in the United States, and there is still no treatment. In 2002, the entire yearly economic cost of diabetes in the United States was projected as $132 B [2]. DR is a frequent diabetic condition. It is, in fact, quite ubiquitous that the main cause of blindness among western working-class people is Diabetes. It is becoming more prevalent not just in industrialized countries, but also in developing countries. Unfortunately, most underdeveloped nations lack basic DR case coding. 75% of persons with DR are thought to live in poor nations [3]. The issue is particularly terrible in developing nations due to insufficient therapy. In most of the country, 25% of the diabetic people likely to be suffer from blindness than those who do not have diabetes. Diabetes is becoming more prevalent not just in industrialized countries, but also in developing countries. Unfortunately, most underdeveloped nations lack basic DR case coding. 75% of persons with DR are thought to live in poor nations [3]. The issue is particularly terrible in developing nations due to insufficient therapy. Digital image capture technology must be employed to reduce the cost of such screenings since it allows us to apply existing image processing approaches to automate the identification of anomalies in fundus images [4]. The following is the format of this review: First part discusses the consequences of diabetes on the eye to meet the objectives of this article. Blood vessel area, discharges, hemorrhages, microaneurysms, and textures are all caused by these impacts as in Fig. 1 [5]. These characteristics are used to identify

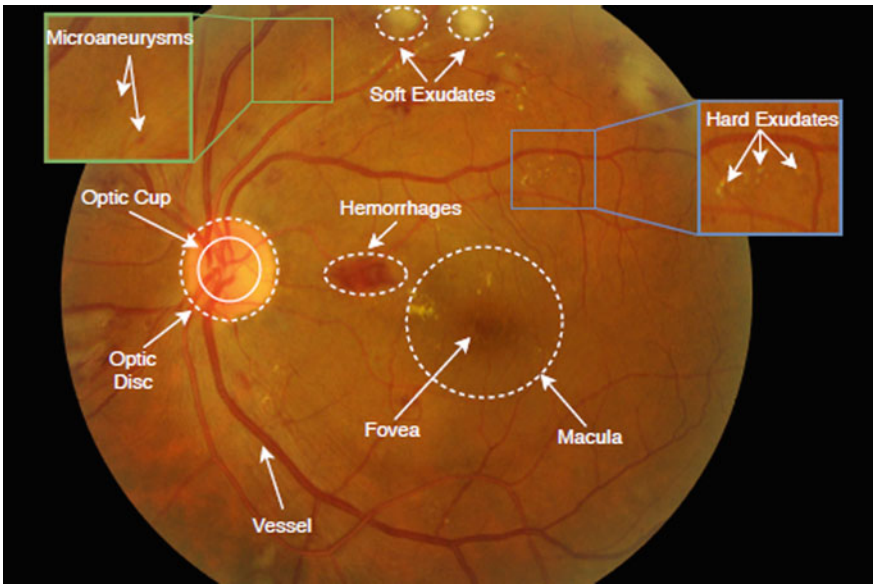


Fig. 1 Different types of DR lesions

Table 1 Classifications of DR stages

DR stages	Level lesions
Stage-0	No lesions
Stage-1	Only microaneuryms
Stage-2	Microaneuryms anf few exudates
Stage-3	Venous bleading, prominentintraretinal microvascular abnormalities
Proliferative DR	New vessels on disc, detachment of retina

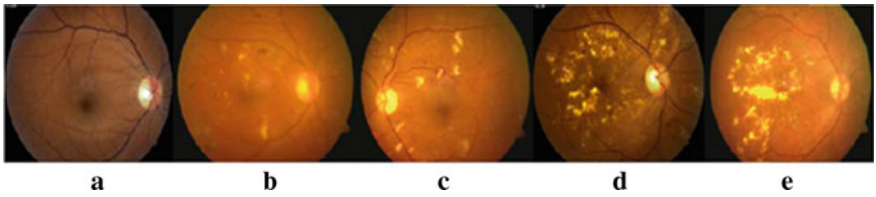


Fig. 2 Fundus images of retina. **a** Normal. **b** Mild. **c** Moderate. **d** Severe. **e** PDR

DR automatically. The next focus will be on several automated detection techniques that have been described in scientific literature under the section on automatic detection of DR phases as in Table 1. The benefits and drawbacks of several techniques are presented in the comment section. The final portion of this study summarizes the findings and suggests future research. DR is a frequent diabetic condition that arises when excessive blood sugar levels damage the blood capillaries in the retina, producing swelling and leaking as in Fig. 2 [6]. In a later stage of DR, the vision may be entirely gone. The global prevalence of DR-related blindness is 2.6% [7]. As a result, diabetic patients should have their retinas screened regularly to diagnose early detection of DR, keep track of its progress, and avert blindness.

Lesions are areas in the fundus image that show up as leaking blood and fluids. Ophthalmologist's manual diagnosis of DR is prolonged, involves an enough work, and may result in wrong-diagnosis. As a result, employing a system for computer-assisted diagnostics can help to prevent wrong diagnosis. And in turn it will save the time, and endeavor of the physician. Deep learning (DL) has evolved and embraced in numerous disciplines over the decade, including medical image analysis. The extraction of hand-crafted features is not required by DL approaches, although they do require a huge quantity of data for training [8–10]. Machine learning (ML) approaches, on the other hand, involve hand-crafted features for its applications. Furthermore, ML methods do not require a large amount of data for training. To identify DR, ML algorithms must first extract the vessel, as described in [11]. Then, as in [12], extract the characteristics of DR lesions for categorization. Segmentation, classification, retrieval, detection, and registration of pictures are examples of DL applications [13]. Convolutional neural network (CNN) is a form of DL technique for image processing that is extensively utilized [14], highly effective, and successful [15, 16]. Significance of automation toward detection of DR has been acknowledged long back. Previous attempts using image classification, pattern recognition, and

ML have made remarkable progress. The objective of this challenge is to push an automated detection system to its limits, preferably leading to models with realistic clinical potential, using color fundus image as input.

2 Detection of Diabetic Retinopathy with Image Processing Approach

A simple image processing method to detect stages of DR is as illustrated in Fig. 3. Before the detection, the fundus image needs to pass through several stages like preprocessing, feature extraction, and classification. To know about the depth of DR, features like a blood vessel, MA, EX, and HM need to be extracted from the fundus image.

Yitian et al. [17] have developed Infinite Perimeter Active Contour Model with Hybrid-region Information (IPACHI) approach. Here, the authors segmented the blood vessels using an automated technique. Finally, the authors compared their approach with different models like Chan Vase (CV) model, Distance Regularized Level Set Evolution (DRLSE) Model, Ribbon of Twins (ROT) Model, and the Infinite Perimeter Active Contour (IPAC) Model. DRIVE, STARE, and VAMPIRE datasets have been used by the authors for their research. To get better results compared with existing methods, filters like Eigen value-based Filter and Wavelet Filters were used.

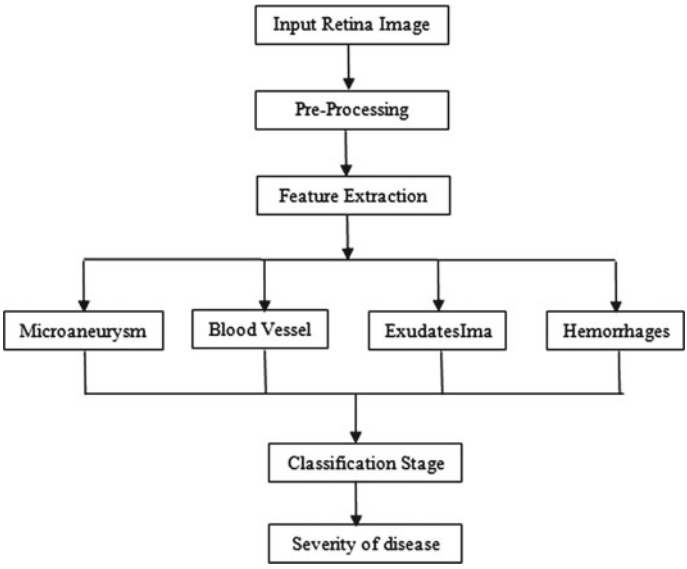


Fig. 3 DR detection using image processing

Zhun et al. [18], proposed an approach to improve accuracy and computational time. Authors have used Morphological operators and Isotropic Undecimated Wavelet Transform (IUWT) to achieve enhanced vessel image. Finally, the major vessels region is extracted using Skeleton extraction and Simple Linear Iterative Clustering (SLIC). The results generated concludes that the proposed method outperforms only for abnormal fundus images.

Orlando et al. [19] worked on the Conditional Random Field (CRF) method for vessel extraction from fundus image. In this work, the Fully Connected CRF (FC-CRF) technique outperforms the local neighborhood-based approach on data sets: DRIVE, STARE, CHASEDB1, and HRF. According to the results, this strategy surpasses other strategies like sensitivity, Matthew's Correlation Coefficient (MCC), mean, and F1-score. According to the authors their technique is more suited for elongated structures.

Hu et al. [20] have combined a multi-scale convolution neural network with an enhanced loss function and CRFs to identify more details and challenging instances, resulting in a unique retinal vascular segmentation technique. Despite the suggested approach's high segmentation efficiency and efficacy, there is still a misjudgement in detecting low contrast of small thin blood vessels, resulting in low sensitivity.

Kaur et al. [21] suggested a three-phased segmentation approach that is unique and generalized. Viz. (i) Image enhancement of the retina (ii) Adaptive image quantization and dynamic decision thresholding to get the actual exudates; (iii) Removal of anatomical features such as blood vessels and optic disc to avoid unwanted interference impact during exudates segmentation. Both lesion-based and image-based methods were used in the trials. The suggested method's performance was assessed using images collected clinically and publicly. In the suggested technique, the computation time is 9.36 s.

Khojasteh et al. [22] suggested a novel color space for fundus images using RGB channels, (i) the RGB space's second Eigen channels, (ii) the Hue/Saturation and Intensity (HSI) space's hue, and (iii) saturation channels. When compared to traditional color spaces, the proposed color space has better replicability.

3 Detection of Diabetic Retinopathy with Machine Learning

A typical DL method to detect stages of DR is as illustrated in Fig. 4. Even here, before the detection, the fundus image needs to pass through several stages as preprocessing, feature extraction, etc. To know about the depth of DR a separate test set and train set of fundus images are prepared. Finally, blood vessel, Microaneurysms (MA), Exudates (EXs), and Hemorrhages (HMs) features are extracted and classified using a suitable model.

Joao et al. [23] have described how different measures of visual evoked capacity may be utilized to identify DR patients. In most cases, DR is a condition in which

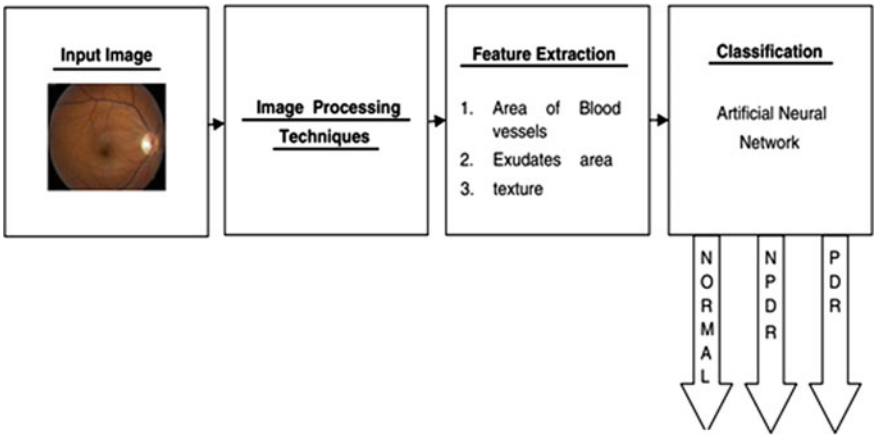


Fig. 4 DR detection using machine learning approach

the body's blood vessels become dry and thick. Unfortunately, many persons with this disease are ignorant of the symptoms, and even when they are, the symptoms often too late for effective treatment. The evaluation of the optical nerve and the optical brain center can pave the way for a faster diagnosis and a better prognosis throughout the recovery phase by allowing for a more accurate calculation of the evoked potential reaction that occurs in the optic nerve in diabetic patients.

Retinal disease screening through the textural method has been proposed by Foeady et al. [24]. The division potential in the fundus of the eye is studied in this article, and it can distinguish between neurotic and strong images. The goal of this study is to look at successful text reading methods using a specific surface descriptor. It is important to note that the goal of the study is to record reading method differences in the text so that the researchers may develop a novel approach that is similar to, how people read the material.

Zeng et al. have focused on an automated method using local pattern modifications [25]. This new and proficient approach is discovered to depict a community with DR. The fundus image is split into a few locations, from which regions of concern are eliminated, and the remaining ones are included into a descriptor's enhanced highlight vector. Local patterns and the random forest methods were used to detect DR and age-related degeneration of muscle in fundus images.

In terms of vision, style, and method, the two latest evaluations, are compared in this approach developed by Badar et al. [26]. Furthermore, Authors used a multi-layer architecture arrangement to improve the performance of the system in a comprehensible way. Van Grinsven et al. have developed a selective sampling approach to decrease the quantity of required feature and to maintain a minimum time during training process. During training, samples are constantly changed by the authors, depending on the results of the current classification iteration. The classifier was a 9-layer CNN, which was influenced by VGG. The CNN with selective sampling (SeS) beat the CNN without selective sampling (NSeS) on the Kaggle competition

and Messidor datasets and there is a decrease in the training epochshas been observed [27].

For the detection of MAs, Dai et al. proposed a multi-sieving convolutional neural network (MS-CNN) based on the clinical data set. Here, the authors have developed an algorithm to create a crude segmentation of microaneurysms. Finally, the authors trained a poor image-to-text algorithm to develop a crude segmentation of microaneurysms. Then, using the rough segmentation as a guide, the suggested MS-CNN was employed to construct superior segmentation. Then, using the rough segmentation as a guide, the suggested MS-CNN has been utilized to create final exclusive segmentation. MS-CNN uses an approach like boosting to address the data imbalance problem [28].

Deep Vessel is a fully convolutional network (FCN) suggested by Fu et al. To aid the network's learning of multi-scale characteristics, they used a side-output approach. A CRF layer was utilized at the end of the net to further model non-local pixel correlations [29]. A fully convolutional network was also proposed by Dasgupta and Singh. Six convolution layers, one down-sampling layer and one up-sampling layer make up their network [30].

Feng et al. presented a FCN that is compared to U-Net in terms of simplicity. Their network only does two up-samples and down-samples. They created entropy to quantify the fraction of vessel pixels in the patch. This method by means of the authors can resolve the class imbalance which exists between background and blood vessels. Half of the patches are chosen at random throughout the training phase, while the other half are chosen from high entropy patches [31].

To identify referable DR, Natarajan et al. have presented a screening method using a smartphone technology. According to the authors, users may receive quick DR diagnostic results through their developed application. This kind of offline approach exhibits greater impact in the field where there is a lack medical resources [32]. Through this research work, DR can be diagnosed with low-cost equipment and in a simple manner.

Tan et al. were the first to use CNN consists of ten layers to automatically and concurrently segment various lesions, including exudates, hemorrhages, and microaneurysms, with result assessed at bit level. The authors contribution proved that a single CNN design may segment many lesions at the same time. Carson et al. utilized a CNN to classify picture patches into five categories as, normal, microaneurysms, dot-blot hemorrhages, exudates or cotton wool patches, and high-risk lesions. They enlisted the help of two physicians to double-check and validate the results, which included 243 images from different data sets. Here, the picture patching approach has been employed and demonstrating that even with minimal training samples, good performance may be attained with this method [33].

Porwal et al. detailed the IDRiD dataset and the setting for the ISBI2018 task "Diabetic Retinopathy Segmentation and Grading". Here, the authors have developed a number of exceptional DL approach from the competition, as well as lessons learned from assessing the entries [34].

Exudates have also been discovered using morphological methods by Sopharak et al. Walter et al. identified exudates based on their gray level variance. Following

the starting location, mathematical morphology techniques were used to compute the contours of the exudates. Depending on the initial threshold value, the applicant exudes areas. The second criterion is the bare minimum of meaning that a nominee must have to stand out from his or her backdrop pixels. They also achieved an accuracy of 92.8% for 15 abnormal retinal images.

Microaneurysms, soft exudates, hard exudates, and hemorrhages were all segmented using a combination of local and global characteristics by Adem et al. Using a low-resolution-based original input image, a Global-Net was utilized to collect more context information [35]. They also used a Local-Net to gather more specific information by using cropped picture patches as input. The backbone of both Global-Net and Local-Net is a U-Net like encoder-decoder architecture.

Hatanaka et al. [36, 37] proposed utilizing density analysis, a rule-based approach, and three Mahalanobis distance classifiers to identify hemorrhage. The propsed method achieved a sensitivity of 80% and specificity of 80%. The green channel's contrast was improved using gray level grouping-based contrast enhancement. Then, using automated seed generation, potential red lesions were retrieved. The hybrid classifier was adopted to distinguish real red lesions from non-red lesions by combining the *k*-nearest neighbor (kNN) and Gaussian mixture models (GMM) classifiers. The best sensitivity achieved was 87%, with a specificity of 95.53%.

Using a morphological technique and "Ball"-shaped structural elements, Acharya et al. [31] discovered hemorrhages by removing the blood artery from hemorrhage candidates. In [21], vessel extraction was done using mathematical morphological and fuzzy clustering approaches. By inhibiting arteries and fovea, hemorrhages and microaneurysms were observed. Finally, authors have developed a hybrid technique to detect red lesions. The mathematical morphology approach was used to extract red candidates. Red lesion candidates were classified using the KNN classifier. The system has a sensitivity of 100% and a specificity of 87%.

To locate blood vessels and hemorrhages, Matei and Matei [17] and Langroudi and Sadjedi [18] employed morphological procedures and thresholding. Maximum of Multiple Linear Top Hats (MMLTH) was enhanced in [19] by employing a multi-scale technique to discover possible hemorrhages and the region expanding surrounding them. The thresholding and skeletonizing approach was used to detect local vessels. SVM classifier was used to classify blot hemorrhages.

The maximal sensitivity and specificity of the technique were 98.60 and 95.50, respectively. Zhang and Chutatape [38] presented a bottom-up approach to brilliant lesion identification and a top-down approach to dark lesion detection. Hemorrhages are initially detected in the Region of Interest by computing the evidence value of each pixel using SVM. To pick features, the kernel PCA and PCA are employed. The SVM Classifier with kernel produces 90.6% true positives, compared to 89.1% for PCA.

4 Data Set

To develop the algorithms, researchers used the publicly available retinal image databases for testing and evaluating the performance of their methods. These databases provide a common platform for comparison. Digital Retinal Images for Vessel Extraction (DRIVE) Database, Structured Analysis of the Retina (STARE), MESSIDOR Database, DIARETDB Database, Kaggle Database, Indian Diabetic Retinopathy Image Dataset (IDRiD) are the major available data set for the research purpose. Along with this researcher have found the required data set in Messidor, e-Ophtha EX, CHASE, HRF, HEI-MED, ROC Dataset, Mums DB, etc. [35].

5 Observation and Discussion

The DR is a progressive disease, early detection of which helps in providing treatment at initial stages which in turn overcomes sighting issues of eye. The manual inspection of which causes erroneous results and can be overcome by automatic DR detection techniques. Several researchers have worked on assessment of DR classification based on five stages (Normal, Mild-NPDR, Moderate-NPDR, Severe-NPDR, PDR) in accordance with the severity of the disease. The literature reveals the use of both supervised and unsupervised techniques in segmentation of identified features such as Blood vessel, Microaneurysms, Exudates, and Hemorrhages for DR detection. The other features such as Cotton wool spots, Cup-to-Disc ratio, and Neovascularisation also contributes toward DR detection which are unaddressed in this survey. A large amount of machine learning approaches has been applied for the grading of diabetic retinopathy but did not focus on all the aspects of diabetic retinopathy. This survey gives an exposure toward numerous classifiers with distinct performance which were evaluated based on Accuracies, Sensitivity, Specificity, AUC.

Most of the databases in the literature were collected from standard databases such as Messidor, DIARETDB, Kaggle, e-Ophtha EX, DRIVE, STARE, CHASE, HRF, HEI-MED, ROC Dataset, Mums DB, except Adem [35]. The particular segmentation technique used in the literature are restricted to limited type of standard datasets and insubstantial number of samples. Nonetheless standard datasets which requires less preprocessing, the consideration of clinical-based datasets gives way for more enhancement methods. The studies do not cover the wider phases of diseases in a particular stage. There is a need for the work that applies a common technique to all the stages of DR and gives best performance consistency. Furthermore, it is observed that advanced diagnosis of DR hang on machine learning approaches with pre-trained models which diminishes the effectiveness of the study. Training using substantial clinical data and based on the application would lead to much promising testing results.

6 Future Research Areas

Based on the review, DR must be diagnosed using recent methods and detection of different stages need to be included in the research. There is a need to come up with an algorithm that will overcome the existing problems like early detection effectively. As a novel approach, hybrid technology can be tested to achieve better performance.

7 Conclusion

The retinal blood vessels are damaged in DR due to fluid leakage from these vessels. Exudates, hemorrhages, microaneurysms, and textures are all distinct types of lesions that are utilized to determine the stage of DR. A study of several image processing methods for DR lesion identification is discussed in this work. This article also includes an overview of current studies on the use of DL in the diagnosis of DR and then able to categorize these pictures as PDR or NPDR. Even though DL has paved the path for more accurate diagnosis and treatment, a study suggests that ophthalmologists must enhance their performance, interpretability, and trustworthiness. This survey gives an exposure toward numerous classifiers with distinct performance which were evaluated based on Accuracies, Sensitivity, Specificity, AUC. A few authors also proposed that Hybrid approach be used to achieve greater accuracy and efficacy in DR detection. Authors conclude that long-term diabetes results in the development of microaneurysms, which thereafter result in exudates and hemorrhages. These are the signs and symptoms of DR, which can result in severe vision loss or possibly blindness. Finally, the authors conclude that to minimize these difficulties, it is critical to identify DR in the early stage by using novel technology.

References

1. World Diabetes, A newsletter from the World Health Organization (1998)
2. Cigna healthcare coverage position—a report (2007)
3. Ong GL, Ripley LG, Newsom RS, Cooper M, Casswell AG (2004) Screening for sight threatening diabetic retinopathy: comparison of fundus photography with automated color contrast threshold test. Am J Ophthalmol 137(3):445–452
4. Hoover AD, Kouzanetsova V, Goldbaum M (2000) Locating blood vessels in retinal images by piece wise threshold probing of a matched filter response. IEEE Trans Med Imag 19:203–210
5. Bhuiyan A, Nath B, Chua J, Kotagiri R (2007) Blood vessel segmentation from color retinal images using unsupervised texture classification. IEEE Int Conf Image Processing 5:521–524
6. American Academy of Ophthalmology-what ıs diabetic retinopathy (2019)
7. Bourne RR, Stevens GA, White RA, Smith JL, Flaxman SR, Price H, Jonas JB, Keeffe J, Leasher J, Naidoo K (2013) Causes of vision loss worldwide, 1990–2010: a systematic analysis. Lancet Glob Health 339–349

8. Wilkinson CP, Ferris FL, Klein RE, Lee PP, Agardh CD, Davis M, Dills D, Kampik A, Pararajasegaram R, Verdaguer JT, Lum F (2003) Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Am Acad Ophthalmol 110:1677–1682
9. Deng L, Yu D (2014) Deep learning: methods and applications. Found Trends Signal Process 7:197–987
10. Vega R, Sanchez-Ante G, Falcon-Morales LE, Sossa H, Guevara E (2015) Retinal vessel extraction using lattice neural networks with dendritic processing. Comput Biol Med 58:20–30
11. Zaid E, Shalash WM, Abulkhair MF (2018) Retinal blood vessels segmentation using Gabor filters. In: Proceedings of the 2018 1st international conference on computer applications and information security, Riyadh, Saudi Arabia, pp 1–6
12. Taylor R, Batey D (2012) Handbook of retinal screening in diabetes: diagnosis and management. Wiley, Hoboken, pp 1–173
13. Sikder N, Masud M, Bairagi AK, Arif ASM, Nahid AA, Alhumyani HA (2021) Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images. Symmetry 13:670
14. Bakator M, Radosav D (2018) Deep learning and medical diagnosis: a review of literature. Multimodal Technol Interact 2:47
15. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JA, van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88
16. Lu L, Zheng Y, Carneiro G, Yang L (2017) Deep learning and convolutional neural networks for medical ımage computing. Springer, Berlin
17. Yitian Z, Lavdie R, Ke C, Simon PH, Yalin Z (2015) Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images. IEEE Trans Med Imaging 34:1797–1807
18. Zhun F, Jiewei L, Yibiaorong (2016) Automated blood vessel segmentation of fundus ımagesusing region features of vessels. In: IEEE symposium series on computational intelligence
19. Orlando JI, Prokofyeva E, Blaschko MB (2017) A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. IEEE Trans Biomed Eng 64:16–27
20. Hu K, Zhang Z, Niu X, Zhang Y, Cao C, Xiao F, Gao X (2018) Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function. Neurocomputing 309:179–191
21. Kaur J, Mittal D (2018) A generalized method for the segmentation of exudates from pathological retinal fundus images. Biocybernetics and Biomedical Engineering 38:27–53
22. Khojasteh P, Aliahmad B, Kumar DK (2019) A novel color space of fundus images for automatic exudates detection. Biomed Signal Process Control 49:240–249
23. Joao VBS, Jorge JGL, Roberto MCJ, Herbert FJ, Michael JC (2006) Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. IEEE Trans Medical Imaging 25:1214–1222
24. Foeady AZ, Novitasari DCR, Asyhar AH, Firmansjah M (2018) Automated diagnosis system of diabetic retinopathy using GLCM method and SVM classifier. In: 5th international conference on electrical engineering, computer science and ınformatics, pp 154–160
25. Zeng X, Chen H, Luo Y, Ye W (2019) Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network. IEEE Access 7:30744–30753
26. Badar M, Haris M, Fatima A (2019) Application of deep learning for retinal image analysis: a review. Comput Sci Rev 35:100203
27. Van G, Van GB, Hoyng CB, Theelen T, Sanchez CI (2016) Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images. IEEE Trans Med Imaging 35:1273–1284
28. Dai L, Fang R, Li H, Hou X, Sheng B, Wu Q, Jia W (2018) Clinical report guided retinal microaneurysm detection with multi-sieving deep learning 37:1149–1161
29. Fu H, Cheng J, Xu Y, Zhang C, Wong DWK, Liu J, Cao X (2018) Disc-aware ensemble network for glaucoma screening from fundus image. IEEE Trans Med Imaging 37:2493–2501

30. Dasgupta A, Singh S (2017) A fully convolutional neural network based structured prediction-approach towards the retinal vessel segmentation. In: 14th IEEE international symposium on biomedical imaging, pp 248–251
31. Feng S, Zhuo Z, Pan D, Tian Q (2017) CcNet: a cross-connected convolutional network for segmenting retinal vessels using multi-scale features. Neurocomputing 392:268–276
32. Natarajan S, Jain A, Krishnan R, Rogye A, Sivaprasad S (2019) Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial ıntelligence system on a smartphone. JAMA Ophthalmology 137:1182–1188
33. Tan JH, Bhandary SV, Sivaprasad S, Hagiwara Y, Bagchi A, Raghavendra U, Rao AK, Raju B, Shetty NS, Gertych A, Chua KC, Acharya UR (2018) Age-related macular degeneration detection using deep convolutional neural network. Future Gener Comput Syst 87:127–135
34. Porwal P, Pachade S, Kokare M, Deshmukh G, Son J, Bae W, Liu L, Wang J, Liu X, Gao L, Wu T, Xiao J, Wang F, Yin B, Wang Y, Danala G, He L, Choi YH, Meriaudeau F (2019) Idrid: diabetic retinopathy—segmentation and grading challenge. Medical Image Anal 59
35. Adem K, Hekim M, Demir S (2019) Detection of hemorrhage in retinal images using linear classifiers and iterative thresholding approaches based on firefly and particle swarm optimization algorithms. Turk J Electr Eng Comput Sci 27:499–515
36. Hatanaka Y, Nakagawa T, Hayashi Y, Hara T, Fujita H (2008) Improvement of automated detection method of hemorrhages in fundus images. In: 30th annual international conference on IEEE EMBS, pp 5429–5432
37. Hatanaka Y, Nakagawa T, Hayashi Y, Kakogawa M, Sawada A, Kawase K, Hara T, Fujita H (2008) Improvement of automatic hemorrhages detection methods using brightness correction on fundus ımages. In: Proceedings of SPIE medical imaging, computer aided diagnosis, 10
38. Zhang X, Chutatape O (2005) Top-down and bottom-up strategies in lesion detection of background diabetic retinopathy. In: Proceedings of CVPR, pp 422–428
39. Zhang X, Thibault G, Decencire E, Marcotegui B, La B, Danno R, Cazuguel G, Quellec G, Lamard M, Massin P, Chabouis A, Victor Z, Erginay A (2014) Exudate detection in color retinal images for mass screening of diabetic retinopathy. Med Image Anal 18:1026–1043
40. Lokaurachchi D, Gunarathna K, Muthumal L, Gamage T (2015) Automated detection of exudates in retinal images. In: IEEE 15th international colloquium on signal processing and its applications
41. Niemeijer M (2010) Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs. IEEE Trans Med Imaging 29:185–195
42. Alireza O, Bita S, Richard M (2017) A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images. IEEE Trans Inf Technol Biomed 13:535–545

Analytical Performance in Data Lake Storage of Big Data Analytics by Databricks Delta Lake for Stock Market Analysis

A. Yasmin and S. Kamalakkannan

Abstract In this paper, a Delta Lake workspace is created by using Azure portal; whereas, the ADLS Gen2 (ADLSG2) acts as primary storage account with a container to store workspace data. Despite the hype, ADLSG2 is immutable and it cannot perform analytics. This drawback leads to the introduction of Azure Databricks Delta Lake (ADDL) to fascinate the learning pattern that can be utilized for developing a support system to anlyze stock market and initiate better prediction on forecasted stock price. Databricks enhancement is an open source named Delta Lake, which remains as a pipeline for atomicity consistency isolation durability (ACID) table storage layer over cloud object stores. Finally, the ADDL performance is evaluated with the existing big data platform by using the parameters like memory usage and CPU usage.

Keywords Big data analytic (BDA) · ADLS Gen2 · Databricks Delta Lake · ACID transactions · Stock market · Cluster · Apache spark

1 Introduction

Financial transactions have evolved from the time of ancient goods trading systems to the current state of the art in E-commerce systems. Financial service industries such as banking, insurance, stock trading, and mutual fund investments have grown significantly as a result of rapid human civilization growth and recent technological breakthroughs. Prior to the digital era, a large number of human workforce is required for all types of transactions as well as business intelligence. At present,

A. Yasmin (✉)
Department of Computer Science, School of Computing Sciences, Vels Institute of Science,Technology and Advanced Studies (VISTAS), Pallavaram, Chennai, India
e-mail: yasmin.a@jbascollege.edu.in

S. Kamalakkannan
Department of Information Technology, School of Computing Sciences, Vels Institute of Science,Technology & Advanced Studies(VISTAS), Pallavaram, Chennai, India
e-mail: kannan.scs@velsuniv.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_17

the digital world has highly evolved by enabling clear and transparent transaction and generating huge amount of digital data. Thus, the digital footprints have been approved for performing a precise analysis based on the recent field named analytics, which assist in generating clear and legitimate business decisions. Based on the ever-increasing young clients, the requirements and desires have been increased along with globalization. However, the financial service industries have generated immense varieties of data with an extreme speed may lead to a new generation of data analytics instance, which is named as BDA. Hence, the large data are analyzed and provided with a better client experience and also the data management may signify the BDA experts.

The client requirement of data processing and increased data volume may lead to the evolution of big data concept [1, 2]. The concept of big data is dependent upon six facets, namely volume, variability, velocity, veracity, value, and variety [3]. The aforementioned characteristics demonstrated that the big data can manage not just volume but also provide important data and ideas for the source [4]. Data Lakes are a supporting solution to the big data storage challenge by acting as a massively scalable storage solution capable of storing large amounts of unstructured and heterogeneous data in an unstructured and heterogeneous form [5]. The key advantage of a Data Lake lies in its capacity to store data in its original form by utilizing a schema on-read approach [6] to process data at runtime. A data lake is also expected to provide the ability to effectively perform analytics, batch processing, and real-time analysis on large volumes of data. This is achieved by combining the benefits of database from both SQL and NoSQL approaches and further supplementing them with both online transaction processing (OLTP) as well as online analytical processing (OLAP) capacities. The data elements are tagged in the lake with an ID and often it will carry additional metadata.

In an organization, the data lakes will assist in improving the archival, capture, exploration, and refinement of raw data. However, the data lake performs as a scalable system to store and analyze the retained data in their inhabitant format and utilized to extract the desired knowledge [7]. Therefore, the data lake can neither be designed from the scratch nor established based on the traditional software solutions [8]. Thus, the various data lake implementations have utilized Apache Hadoop as a base platform [9]. One of the recent trends that earned most recognition is Data Lake. This is utilized for storing data in their raw format with simple architecture. Moreover, each entity of data present in the lake gets connected with unique identifier and an extended metadata set. Hence, the client can be able to create purpose-built schema to query the relevant data, which leads to a smaller dataset, which assist in analyzing and finding the solutions for client's requirements. Hence, no fixed schema has been defined earlier and there are certain questions to discuss about the probability of data becoming improbable.

At present, the recent systems are provided with high capacity as well as scalability enhancement. In 2016, Delta Lake has developed by Databricks to assist in storing data in which the respective objects will be considered as part of a table and further Data Lake is performed as a transaction log in the format of Parquet. It enables the data to scale billions of objects per table [10]. Delta Lake act as a storage layer

and it is an open source, which fetches dependability to Data Lake. It is a technology that builds robust data lakes and remains as a vital component for building the cloud data platform. Delta Lake has integrated both streaming and batch data processing during the provision of ACID transactions as well as handling of scalable metadata. Delta Lake is finally compliant with Apache Spark APIs and functions on top of the existing data lake. Based on the workload pattern, the Delta Lake can be configured on Databricks. Append a transaction log which begins with an entry has described all of the existing files and the Delta Lake has capability to transform a current directory of Parquet files into the table of Delta Lake with zero copies. In 2019, Delta Lake has advanced its development by covering 50% of computing hours on Databricks. Additionally, the layer of metadata is placed naturally to implement the data quality enforcement features. Delta Lakes perform data processing in a different order and the data is maintained in the original form by Data Lake. The preparation procedure may not be completed till the data is requested by the applications or the query gets executed. As a result, it violates the data warehouse's typical ETL principles; whereas, Delta Lake promotes the new extract load transform (ELT) data processing order instead of extract transform load (ETL). In Delta Lake, no data schema is predefined and the data gets extracted from source and get transferred to a data lake based on the required metadata that gets involved. Hence, Delta Lake can be able to handle both write-intensive workloads and read-intensive workloads which is transactions as well as analytics that reuniting two distinct concerns like write and read. One of the instances for Delta Lake is a schema enforcement feature which ensures data uploaded to a database that meets its schema, as well as a constraints API. It allows table owners to specify limits on the data being ingested. Records that breach these standards may be automatically rejected or quarantined by Delta's client libraries. These simple features have shown to be quite effective in improving the quality of Data Lake-based pipelines. Therefore, this research focuses on ADDL to perform BDA efficiently without losing of data from ADLSG2 and transforms to the same ADLSG2 as the user end. Thus, the ADDL involved in ADLSG2 because it is immutable which does not perform analytics, but ADDL assist to perform those BDA process for stock market analysis in forecasting the stock price and deliver to the user end through ADDL.

The paper is organized as follows in which Sect. 2 deliberates the related work of the existing data lake storage techniques as well as benefits of Delta Lake to manage the database management system (DBMS) and also manage analytics in BDA based on user end requirement. The proposed ADDL to support ADLSG2 for BDA is done through Apache spark pool with Azure protocol is illustrated in Sect. 3. The evaluation results of proposed ADDL design with ADLSG2 for BDA are compared with the existing BDA tools using parameters like CPU utilization and memory usage are calculated is illustrated in Sect. 4. Finally, Sect. 5 deal with concluding the proposed ADDL design with ADLSG2 has performed the ELT process instead of ETL process which assists in write as well as read the data. Thus, it assists the proposed design to execute the BDA using spark pool and make the stock market analyzer to predict the data using historical data.

2 Literature Review

Azure Data Lake is a framework which is adaptable, versatile, efficient, and stable. It can store and analyzes a wide variety of data and has been designed for massive workloads with high throughput requirements. It can be accessed via Storm, U-SQL, Hive, and Spark, among other methods. The ADLS and Azure Data Lake Analytics (ADLA) together make up Microsoft's data lake solution.

Data Lake assists to store data in its default format using a simple architecture. All data entity present in the lake has a unique identifier and an extended metadata set. This performed as a purpose-built schema which are created by the client based on the query related data which lead to a less amount of dataset that can be analyzed for providing solutions to client's questions. Moreover, there is no fixed schema that has been defined earlier and there are certain questions discussed about the probability of data become improbable [11]. T.S. Hukkeri et al. have discussed a brief idea about the various data lake solutions available in the market for the consumers. It includes Hadoop, AWS, and Azure data lake architecture. Apache Hadoop is considered as an industry standard for Data lakes. Both Amazon Web Services (AWS) data lakes and Azure Data Lake have constructed wrappers around Hadoop in order to solve concerns about its raw data and lack of data security [12]. Snezhana Sulova has discussed an ideas of data warehouse and Data Lake are explained by the author along with its positives and negatives. This review discusses a quick overview about how Data Lake can be used to address the challenges faced by the smart grid data management systems [13]. Surabhi DHegde et al. have explained data lake lambda architecture for analyzing the data. The emerging big data concepts consists of both Data Lake and fast data is explained. While both data lakes and big data are able to store and process more data, fast data gives real-time insights rapidly based on limited data [14].

C. Diaconu et al. illustrate the infrastructure of RSL-HK ring that perform as a foundation for addressing ADLS which supports very huge files and folders. This infrastructure has provided an efficient, high memory size, scalable, and persistent state which are the essential metrics required from metadata services. The implementation of this infrastructure involves with new combination of Paxos and design of recent transactional in-memory block data management. The availability and scalability of RSL-HK ring is dependent upon the capability of dynamically adding novel Paxos rings as well as accumulating machines to the existing Paxos ring. Similarly, the RSL-HK Paxos components are dependent on the implementation over Cosmos that provided very high availability across hundreds of ring utilized for production in several years. Thus, the leverage technology for transactional in-memory block management is utilized for SQL Hekaton [15]. The agile BI basic principle has been introduced for providing an exact data at right time to the accurate analytical process is the legitimate implementation of Delta Lake. As the results, the first phase of a project in an organization for product development has reduced the errors as well as costs of subsequent stages [16]. The major conditions to be followed for successful Delta Lake data source implementations are as follows [17, 18]

- The purpose and requirements of using Delta Lake are defined clearly.
- Create and execute the strategy of data management in Delta Lake.
- Procedures are created for security and control in data access as well as utilizing the data effectively.
- Generate new and additional data source for analysis instead of BI infrastructure sole components.

The declarative version has been offered for successive data frame APIs are utilized in those libraries that mapped the data preparation computation into the plans of Spark SQL query and obtain benefits from the optimizations in both Delta Lake as well as Delta Engine. However, this method is applicable in Spark data frames as well as Koalas [19]. The koalas have defined the recent data frame API for spark which provide enhanced compatibility with Pandas. Hence, the organizations are creating a large range in the version of ML-specific data and the feature store system has assist in re-implementing default DBMS functionality. This may be easier for utilizing abstraction of Data Lake with built-in function of DBMS that help to implement the feature store functionality. Simultaneously, declarative ML systems like factorized ML might potentially outperform a Lakehouse [20]. Instead of searching into raw files in a data lake, cloud-native DBMS provides a solution as serverless engines that assist in integrating with affluent layers of metadata management using Delta Lake which help for accomplishing enhanced performance [21]. Thus, Delta Lake has acknowledged the applications to create a specific extensive action in their log record objects with Web-based applications and version attributes to monitor the application with specific data likely to be offset associated with input stream is an instance. This application has ensured that the Delta Lake can accumulated the recent data and stored it in a version field atomically by providing this data in the same log record. The similar log records are used as an appropriate delta in add and removal of events that are recorded into the log atomically. In order to get a unique ID, each application can effectively create its app ID at random space and in the case of Spark structured streaming, the feature utilized is Delta Lake connector. [22].

This session discusses the data lake architecture with Hadoop benefits as well as Delta Lake assisted for enhancing the optimization with spark SQL but benefit of BDA with better computation and memory storage usage is not obtained. Moreover, this session discusses the individual performance of ADLS and ADDL which does not support stock market analysis. Hence, the ADDL with ADLS is combined to accomplish better BDA process in stock market analysis and performance of big data platform is discussed in the following sessions.

3 Research Methodology

In the domain of trading, stock price prediction is an essential factor that assists in determining the profit for the customer. This research is focuses on predicting of stock price by machine learning (ML), which has to be progressed by using bulk and Tera

bytes data. However, the bulk data transformation and storage can be done through big data platform. Hence, this work involves the ADDL for streaming the bulk and large data without buffer and data loss. Thus, the Azure protocol has introduced an architecture for BDA that designed to maintain the data mounting, data processing, and also data analysis. In addition, the threshold generated from an organization has created into the big data domain is different which completely based on the customer capacity and their accessing tools. In some cases, very large data sources like hundreds of gigabytes and hundreds of terabytes may process as a workload for the tool. It may handle by advance of big datasets which represent the meaning of big data.

Delta Lakes perform data processing in a different order and the data is maintained in the original form by Data Lake. The preparation procedure may not be completed till the data is requested by the applications or the query gets executed. As a result, it violates the data warehouse's typical ETL principles; whereas, Delta Lake promotes the new extract load transform (ELT) data processing order instead of extract transform load (ETL). In Delta Lake, no data schema is predefined, and the data gets extracted from a source and transferred to a data lake based on the required metadata that gets involved. Hence, Delta Lake can be able to handle both write-intensive workloads and read-intensive workloads as well as analytics that reunite two distinct concerns like write and read. Furthermore, rather than explicitly referring to the quantity of the data, this concept represents the value that can be extracted from datasets by using advanced analytics. Instead, the data source size may be considered during the circumstances tend data source to be larger. The ADDL has integrated Azure protocol with Databrick analytics by using spark pool is as shown in Fig. 1.

Moreover, this architecture has been proposed to resolve various challenges like rapid pace arrival of data that constantly demanded are collected and other hand with slow arrival of data with very large chunks. Hence, the Azure Databricks is introduced to face advanced analytical issues as well as the requirement of ML. Thus, data source is stored in ADLSG2 may be real-time data or batch processing data. The batch layer is a cold path that maintain all type of receiving data in terms of original form and performs batch processing. Once the batch view gets generated as an outcome of processing has enabled high-precision calculation across massive datasets, which is time-consuming. The raw data progressed in batch layer is immutable; whereas, the recent data is often accumulated to the earlier data, and the earlier data is not even obliterated. All changes done to the value of a specific data are recorded as a recent event record with a time stamp. This enables to re-compute at any point of time throughout the data collection's existence. Thus, the capability in re-compute the batch view from the initial data source data is critical since it enables the creation of new aspects as the system evolves.

The performance of data storage through ADLSG2 with global accessibility and partner ecosystem has made the platform in selecting analytics customers as well as partners over the world. The subsequent process is event processing; whereas, Azure event grid is arranged as a service for completely managed event routing. In addition, the serverless performance is executed through Azure functions to compute engine

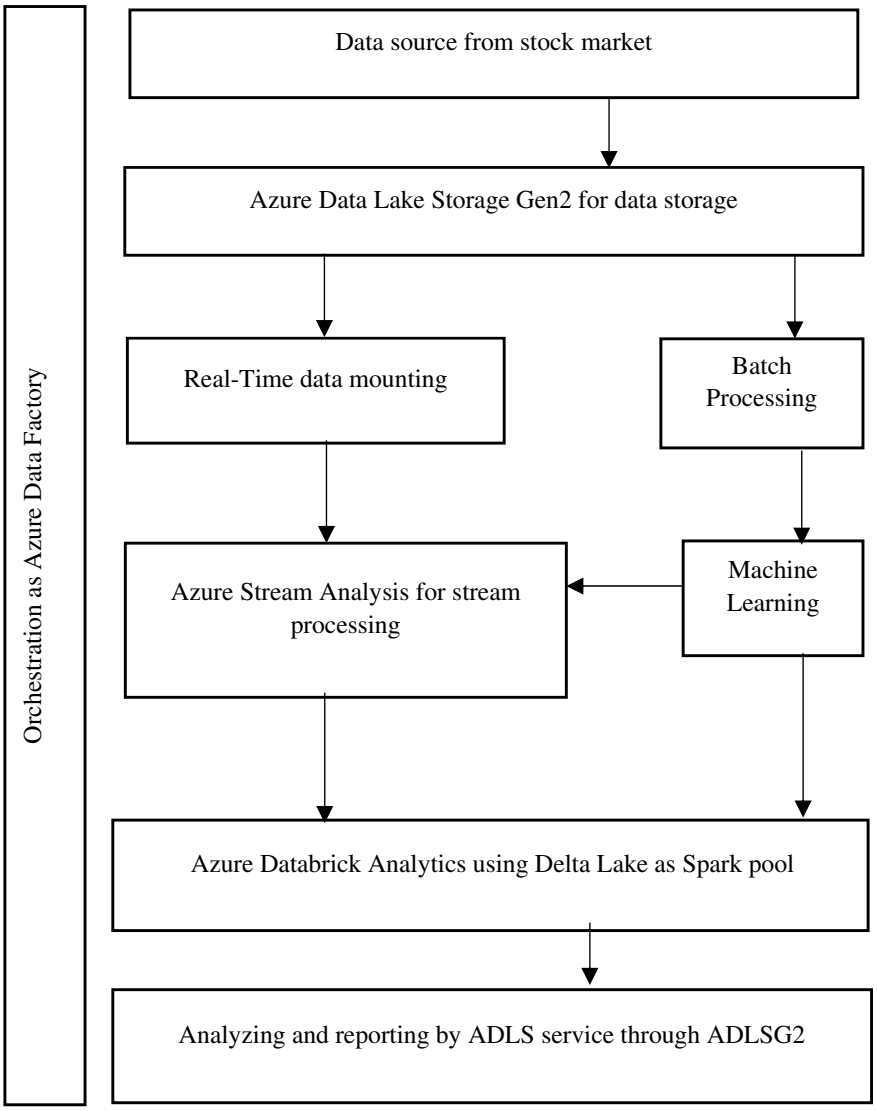


Fig. 1 Big data architecture using ADDL with ADLSG2

and Azure logic apps for workflow orchestration engine, which make the event-based processing to perform easier and workflows get responded to the events in real time. This research focuses on batch processing due to very large data sources and in general, the solution of big data gets accomplished in processing the data files by long running batch jobs (LRBJ). The purpose of LRBJ is to aggregate, filter, and prepare the data for analysis. Once the external data source is created, the utilization of data source and its relative path to file are added in the OPENROWSET function.

In addition, the entire absolute URI is not required for the client files but defining the custom credentials is utilized to access the storage location. Hence, these jobs are involved by reading the data source files, processing of data analytics, and writing of output to the new files. Thus, the data storage stores the data from various sources lands in ADLSG2 through Azure data factory that perform as an orchestration and other data movement tools. The events generated from the ADLSG2 for new file creation, updating of files, and modify by renaming or deleting of files that are routed through Delta Lake as event grid and Azure function by ADDL get connected are shown in Fig. 2.

The Azure function performs to make client interacts with the Data Lake service at the account level. It acts as the primary storage and produces operations for retrieving and configuring the account properties; whereas, the operations like list, create, and delete file systems in the account by Delta Lake as spark pool. Thus, the operations related to a specific file system, directory or file, the clients to those entities can also be retrieved by using functions like

- get_file_client,
- get_directory_client,
- get_file_system_client.

Moreover, the Databricks has provided a very fast and easy way to set up and utilize a cluster; whereas, the setup of cluster in Databricks is easy as heck. Hence, the new cluster is created on the home page with the event log by spark pool is shown in Fig. 3. Thus, it is done through clicking the "Create cluster" in which name of the cluster needs to be chosen for the specific cluster. Reading of data can be accelerated through Delta cache by generating remote file copies in the nodes local storage using the format of fast intermediate data. When the files have been retrieved from a remote

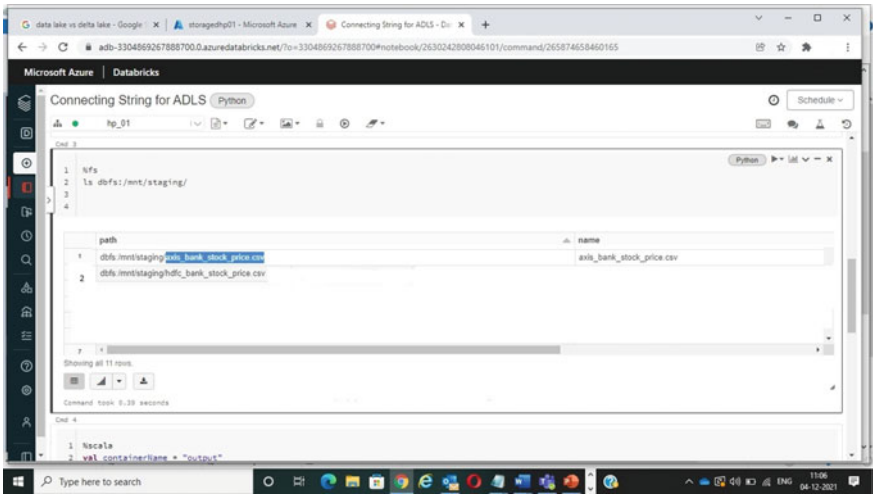


Fig. 2 Connection of batch processing by ADLSG2 to Azure Databricks

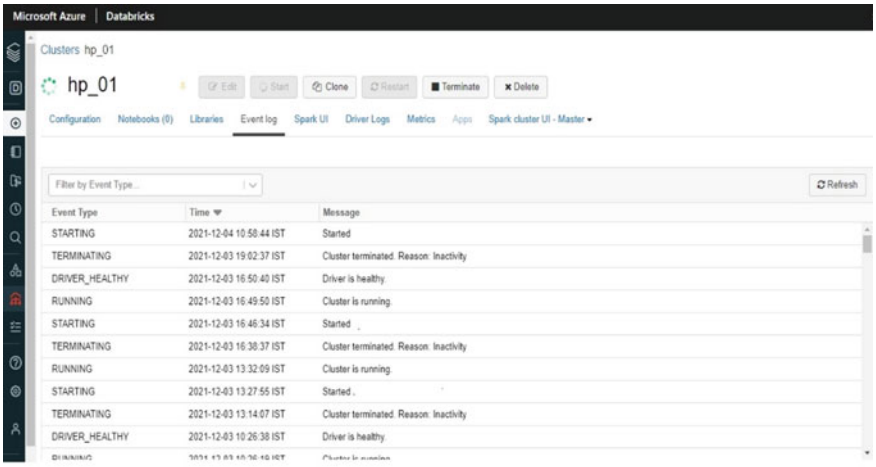


Fig. 3 New cluster generated in Azure Databricks

location, it is understood that data gets cached automatically. The subsequent reads of similar data have been performed locally that results in improving the reading speed significantly. Moreover, the Delta cache consists of local remote data copies which can be able to enhance the queries performance in a wide range but cannot be utilized for storing the arbitrary sub-queries results. Hence, the data preparation of the respective CSV files is made read into the data frames after clustering is done. Thus, the Delta cache has ability to read and execute data faster than the Spark cache due to usage of effective decompression algorithms from the Delta cache. As the results, the output data in the optimal format has more processing by whole-stage code generation.

There are several solutions in big data to prepare analysis and even serve the data process in a structured format, which is queried by ADDL as an analytical tool by using spark pool. This research has involved ADDL with ADLSG2 for providing a manageable service to large-scale and cloud-based data warehousing. The major focus of big data solutions is utilized for providing knowledge into the data over analysis as well as reporting. However, empowering the users for analyzing the data using ADDL with ADLSG2 as a data modeling layer which can perform a multidimensional online analytical processing (OLAP) cube model over services of Azure analysis. In general, the solutions of big data solutions involve copied data processing operations are enclosed in workflows that sinks have been loaded with the processing data into analytical data store. Finally, it transforms the source data and move the data among several sources for projecting the results to a report. Thus, the process is used to automate the workflows using an orchestration technology like Azure data factory. The performance of ADDL is evaluated with the existing big data platform using the parameters like memory usage, CPU usage, and load time.

4 Result and Discussions

In this experimental implementation of Databricks, runtime is done through an open source of Apache Spark APIs using highly optimized execution engine that creates essential performance gain while compared to general source of Apache Spark available on ADLS platform. However, predicate pushdown technique has been applied through Spark SQL for defining filters if essential. Hence, the query processing stage gets verified by defining explain command to query. When the query plan consists of PushedFilter, the query gets optimized for selecting the essential data by predicate returns with either True or False. In the case of no PushedFilter in query plan, it gets casted with condition.

Moreover, the core engine is made wrapped with additional services like ADDL with spark pool for connecting the ADLS using cluster for developer productivity and enterprise governance. Therefore, this experiment is done by cloud environment using Azure protocol with the system configuration of system with 6 core, 8 GB of RAM. This research has implemented the ADDL strategy with ADLSG2 for BDA and it can be evaluated using the Quality of Service (QoS) performance. In addition, the proposed architecture of ADDL with ADLSG2 method has the capability to fulfill the suitable demand in various kinds of diverse data frames. Hence, the ADDL with ADLSG2 method generated analytical as well as storage that can be analyzed through metrics by load time and produced less utilization of the CPU and memory. Thus, the metrics illustrate that usage has progressively reduced the data utilization cost. The QoS parameters for the proposed ADDL with ADLSG2 with optimization outcome for specific date time data processing is given in Table 1. It revealed that the traditional ADLSG2 CPU utilization and usage of memory are more when compared to ADDL with ADLSG2 CPU utilization and usage of memory. This assists in performing better stock market analysis, which computes as well as consume less memory instances.

Table 1 Comparative of optimization outcome from ADDL with ADLSG2 versus traditional ADLSG2

Date time	CPU usage in MHZ		Memory usage in KB	
	ADDL with ADLSG2	Traditional ADLSG2	ADDL with ADLSG2	Traditional ADLSG2
2016-02-02 00:00:00	58.00	81.24	116,763	135,286.8906
2016-02-02 00:05:00	106.49	123.05	333,189	406,408.00
2016-02-02 00:10:00	61.23	68.06	148,478	163,492.1875
2016-02-02 00:15:00	48.47	65.72	128,810	151,165.3906
2016-02-02 00:20:00	59.71	64.16	116,280	140,813.1718

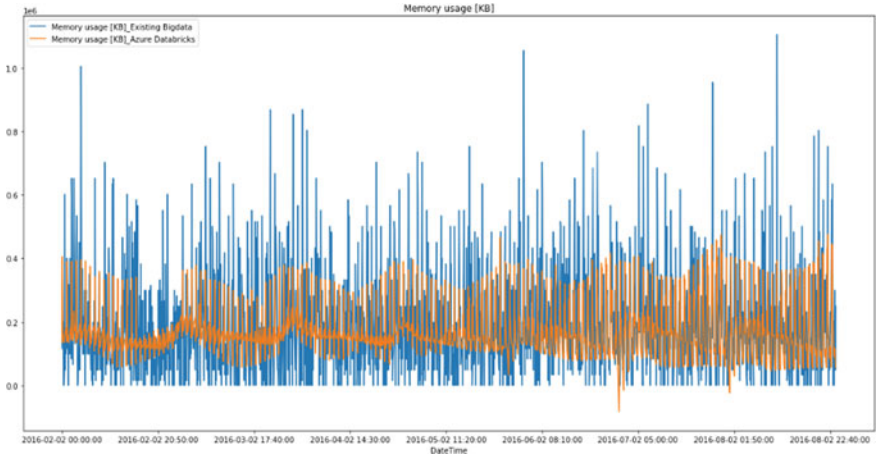


Fig. 4 Memory usage of Azure Databricks versus traditional ADLS

Memory Usage

Figure 4 has illustrated memory usage from ADDL with ADLSG2 and traditional ADLS in which specific date time is considered randomly. The date time interval is considered every 5 min, and the utilization of memory for the respective data source by the consumer is comparatively lower in ADDL with ADLSG2 than traditional ADLSG2 has determined that ADDL support BDA in stock market analysis by consuming less memory storage usage. In 2016:02:02 00:05:00, CPU utilization of ADDL with ADLSG2 is 406.41 MB higher than traditional ADLSG2 (333.19 MB), respectively. Thus, it determines that usage of memory by ADDL with ADLSG2 is less and quite beneficial to allocate more activities for storage in read as well as write to analyze the BDA based on stock market client requirements.

Figure 5 has illustrates CPU utilization from ADDL with ADLSG2 and traditional ADLS in which specific date time is considered randomly. The date time interval is considered every 5 min and the utilization of CPU for the respective data source by the consumer is comparatively lower in ADDL with ADLSG2 than traditional ADLSG2 has determined that ADDL support BDA in stock market analysis by consuming less computation. In 2016:02:02 00:05:00, CPU utilization of ADDL with ADLSG2 is 123.05 MHz higher than traditional ADLSG2 (106.49 MHz). Thus, the utilization of CPU by ADDL with ADLSG2 is less and quite beneficial to analyze and compute more activities for analyzing the BDA based on stock market client requirements.

The performance of large dataset loading into Delta Lake has challenged the ADLS for writing operations and further prompting ADDL to join ADLSG2 for improved writing. Figure 6 depicts the results. The Delta statistics collections have accumulated the essential overloads. The transaction processing performance council (TPC) is an enterprise class for decision support benchmarked as TPC-DS is used to measure the response time of query generated in a single user mode. Similarly, this research is more focused on load time with 250 GB to analyze the table of TPC-DS

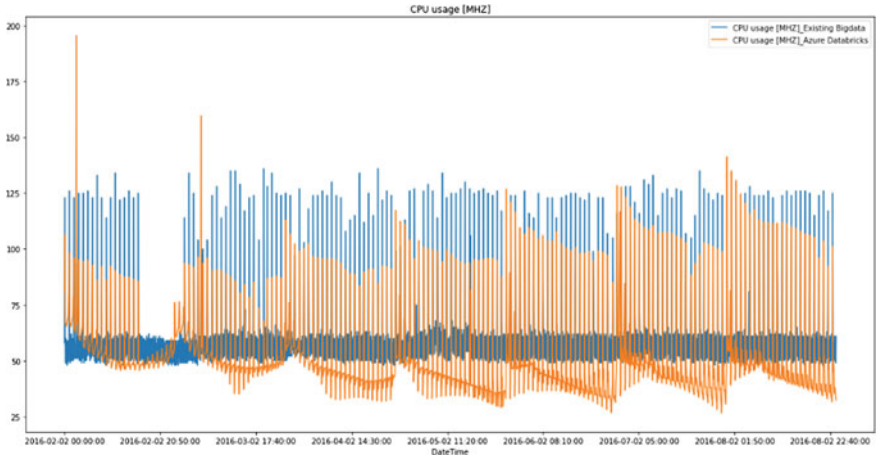


Fig. 5 CPU usage of Azure Databricks versus traditional ADLS

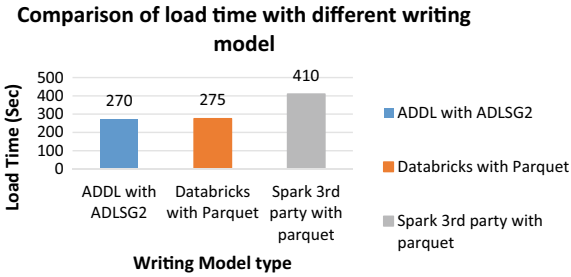


Fig. 6 Comparison of load time with various writing model type

store sales, which is generally formatted as CSV files on a cluster with master system as one i3-2xlarge and six i3-2xlarge workers. The resulted average with 2 runs and Spark pool performance in writing to Delta Lake is similar like writing to Parquet has determined that statistical collection does not accumulate significant overhead in the other data loading work.

ELT has the operation parallelized to the dataset, and disk I/O is generally optimized at the engine level to produce rapid throughput. The scaling of ELT with the existing ADLSG2 hardware and RDBMS engine which even ignored scalability. This even generate better ELT leverages with the existing DBMS and keeps all data in the DBMS all the time. In operational system of TPC-DS, data is provided by flat files that referred to as the refreshed dataset. Thus, it initiate with the reading of the refreshed dataset and the ELT process in TPC-DS has involved in integration and data consolidation from operational systems. It is applied to diverse the workload consisting of data transformations with effective loading time for better performance of throughput in workload.

5 Conclusion

Delta Lake is solely implemented as a storage format and provided with access protocol sets to client with ease in operation as well as availability. This provides client with direct and high bandwidth to the object store. Delta Lake has been utilized by several organization for processing Terabytes of data source per day and it has generally replaced more complicated architecture, which involve various data management system. The presentation of Delta Lake with ACID table as storage layer in cloud object storage enables an extensive range of DBMS performance as well as data management with a reduced expense. The architecture of ADDL with ADLSG2 has produced a storage with read and write operations by considering the ELT operation on the data source. The evaluation is done by utilizing the performance metrics like CPU utilization and memory usage and it has also compared ADDL with ADLSG2 and traditional ADLSG2. As a result, the consumption of CPU utilization and memory usage has been reduced in ADDL with ADLSG2 in the existing big data platform. The load time is calculated by using TPC-DS for determining the writing performance and the load time consumption is less in ADDL when compared with parquet and 3rd party of parquet. Henceforth, the ADDL writing performance illustrates the load time of analysis, as writing performance is better in performing BDA in stock market with efficient utilization of CPU and memory along with a reduced cost.

References

1. Chong D, Shi H (2015) Big data analytics: a literature review. J Manag Anal 2:175–201
2. Tomashevskaya VS, Yakovlev DA (2021) Research of unstructured data interpretation problems. Russian Technol J 9(1):7–17
3. Cappa F, Oriani R, Peruffo E, McCarthy I (2021) Big data for creating and capturing value in the digitalized environment: unpacking the effects of volume, variety, and veracity on firm performance. J Prod Innov Manag 38(1):49–67
4. Yang C, Huang Q, Li Z, Liu K, Hu F (2017) Big data and cloud computing: innovation opportunities and challenges. Int J Digit Earth 10:13–53
5. Khine PP, Wang ZS (2018) Data lake: a new ideology in big data era. ITM Web Conf 17, 03025
6. Miloslavskaya N, Tolstoy A (2016) Big data, fast data and data lake concepts. In: 7th annual ınternational conference on biologically ınspired cognitive architectures, BICA
7. Ilin D, Nikulchev E (2020) Performance analysis of software with a variant NoSQL data schemes. In: 2020 13th ınternational conference management of large-scale system development (MLSD), IEEE, pp 1–5
8. Sawadogo PN, Scholly E, Favre C, Ferey E, Loudcher S, Darmont J (2019) Metadata systems for data lakes: models and features. In: 1st ınternational workshop on BI and big data applications (BBIGAP@ADBIS 2019), Bled, Slovenia, pp 440–451
9. Khine PP, Wang ZS (2018) Data lake: a new ideology in big data era. ITM Web Conf 17:03025
10. Armbrust M, Das T, Sun L, Yavuz B, Zhu S, Murthy M, Torres J, van Hovell H, Ionescu A, Undefineduszczak A, Undefinedwitakowski M, Szafrański M, Li X, Ueshin T, Mokhtar M, Boncz P, Ghodsi A, Paranjpye S, Senster P, Xin R, Zaharia M (2020) Delta lake: high-performance ACID table storage over cloud object stores. VLDB

11. Inmon B (2016) Data lake architecture: designing the data lake and avoiding the garbage dump. Technics Publications
12. Hukkeri TS, Kanoria V, Shetty J (2020) A study of enterprise data lake solutions. Int Res J Eng Technol (IRJET) 07(05)
13. Sulova S (2019) The usage of data lake for business ıntelligence data analysis. Conference Paper
14. Surabhi DH, Ravinarayana B (2016) Survey paper on data lake. Int J Sci Res (IJSR) ISSN 2319–7064
15. Diaconu C, Freedman C, Ismert E, Larson PA, Mittal P, Stonecipher R, Verma N, Zwilling M (2013) Hekaton: SQL server's memory-optimized OLTP engine. In: Proceedings of the 2013 ACM SIGMOD ınternational conference on management of data, pp 1243–1254
16. Nacheva R (2015) Principles of user ınterface design: ımportant rules that every designer should follow. Izvestia J Union of Scientists Varna. Econ Sci 140–149
17. Khıne P, Wang Z (2018) Data lake: a new ideology in big data era, ITM Web of Conferences 17, WCSN 2017
18. Yordanova S, Stefanova K (2019) Big data challenges—definition Characteristics Technologies. Sci Papers UNWE 1:13–31
19. Armbrust M, Xin RS, Lian C, Huai Y, Liu D, Bradley JK, Meng X, Kaftan T, Franklin MJ, Ghodsi A, Zaharia M (2015) Spark SQL: relational data processing in spark. In SIGMOD
20. Li S, Chen L, Kumar A (2019) Enabling and optimizing non-linear feature interactions in factorized linear algebra. In SIGMOD, pp 1571–1588
21. Perron M, Castro Fernandez R, DeWitt D, Madden S (2020) Starling: a scalable query engine on cloud functions. In SIGMOD, pp 131–141
22. Armbrust M, Das T, Torres J, Yavuz B, Zhu S, Xin R, Ghodsi A, Stoica I, Zaharia M (2018) Structured streaming: a declarative API for real-time applications in Apache Spark. In SIGMOD, pp 601–613, New York, NY, USA

Degraded Factors Analysis in Multimedia Data Using Deep Learning Algorithm

A. Selvi, J. K. Thamarai Selvi, V. Umaiyal, S. Keerthana, and R. Gokul

Abstract Because artificial illumination is created in an indoor environment, multimedia data is recorded in an optimal atmosphere. On the other hand, in an outdoor setting, it is critical to eliminate the weather influence. Outdoor vision devices are used to capture scenes in surveillance. Due to the enormous size of the drips, the items will get motion blurred in changing weather conditions. Several computer vision techniques that employ feature information, such as object identification, tracking, segmentation, and recognition, will be harmed by these disturbances. Even if only a little portion of the object is obscured, the object cannot be accurately tracked. Rainfall features that pixel image is not entirely surrounded by rain in all data. The dynamic adverse weather model is researched for restoration resolution. Rainfall is a very important part of a bad weather system. Rain-fed energy has a strong local structure and is strongly influenced by backlight. When light passes, it is repeated and visible, shining in the surrounding area. The movement fades, when it falls at a high rate. The vehemence of the rain line is therefore, determined by the light of the descent, the radiation of the backdrop, and the unification time of the camera. Particles of rain and ice are very difficult to analyze. Rain-like spatial and temporal occurrences can be produced by some scene dynamics. In this work, we will look into ambient light estimation for submitted multimedia material using image processing algorithms. With in-depth reading techniques, we improved light measurement and used the histogram measurement method.

Keywords Image processing · Atmospheric light estimation · Deep learning algorithm · Dehazing · Scene radiance

A. Selvi (✉) · J. K. T. Selvi · V. Umaiyal · S. Keerthana · R. Gokul
Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur 639113, India
e-mail: selvia.cse@mkce.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_18

1 Introduction

The images taken in bad weather conditions may degraded due to the presence of some factors like dust, fog, snow, and rain. While shooting outside, garbage in the climate absorbs and disperses light as it passes from the scene to the computerized camera. Such weakened snap photographs much of the time need visual clarity and give helpless perspective on the scene's substance. This effect is disrupting when driving, just as for business and creative picture takers. The light reflected from a story is spread all through the climate before it arrives at the camera in pretty much every reasonable case [1, 2]. This is on the grounds that to the presence of sprayers, like residue, fog, and exhaust, which cause gentle to go astray from its expected way of transmission. In significant distance pictures or hazy scenes, this technique has a standard effect at the photograph, wherein contrasts are diminished and surface hues arise as weak. Such corrupted photographs routinely need apparent distinctiveness and request, and also, they give an awful perceivability of the scene substance. This impact can be a disturbance to fledgling, business, and inventive picture takers just as sabotage the extraordinary of submerged and ethereal pictures. This can likewise be the situation for satellite TV for PC imaging that is utilized for parcel purposes which incorporates map making and web planning, land-use arranging, archaic exploration, and ecological examination [3–5]. The charming of photograph taken under awful perceivability is normally debased by utilizing the presence of mist, murkiness, exhaust cloud, or fog. Since the climate changed into impacted, the differentiation of the photograph is essentially diminished. Dimness decrease (or dehazing) is very well known in both client/computational pictures and PC vision programming [6, 7]. Dehazing is a procedure for eliminating murkiness from a photograph. To start with, eliminating the murkiness works on the scene's perceivability and adjusts the shading shift brought about by the hermetically sealed. The image with no dimness is all the more outwardly engaging. Second, most PC vision calculations, from low-level picture handling to unnecessary stage object acknowledgment, expect that the passage photograph (after radiometric adjustment) addresses the scene brilliance. The exhibition of innovative and perceptive calculations (e.g., trademark discovery, sifting, and photometric examination) will definitely be tormented by the one-sided, low appraisal scene brilliance [8–10]. Last, the murkiness evacuation can create profundity data and gain numerous vision calculations and prevalent picture upgrading. Cloudiness or haze might be an advantageous profundity sign for scene understanding. The horrible dimness picture can be put to appropriate use. Notwithstanding, cloudiness end is a troublesome difficulty because of the reality that the fog is relying upon the obscure power realities. The difficulty is under limited assuming that the middle is just a solitary fog picture. Figure 1 shows the dehazing system [11].

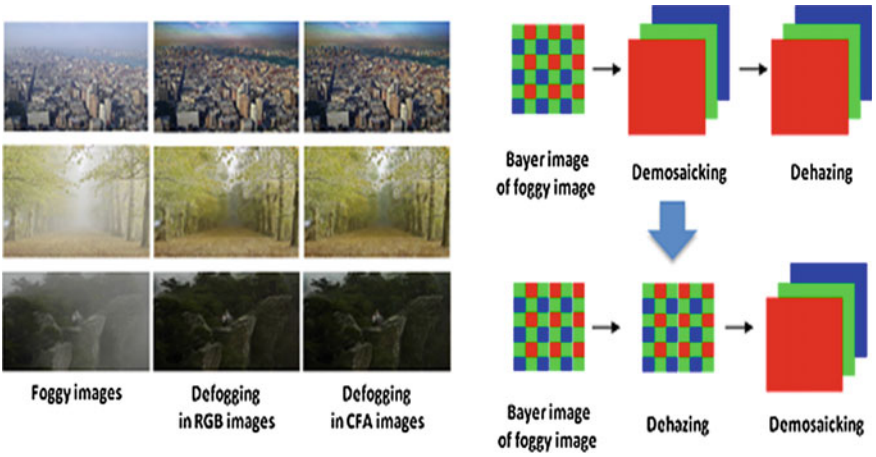


Fig. 1 Defogging in image processing pipeline

2 Related Work

Singh et al. [12] using a single coefficient-based slope image that is primarily deliberated to acquire the precise transmission map. The suggested procedure has been evaluated on five popular external images. Test outcome displays that the colors of the haze-free pictures are not disfigured. The proposed method has a small variety of artifacts and retains larger edges than the current fog removal. We try to restore the appearance of foggy external images and at the same time, press the artifact to restore the radiometric element to adverse weather conditions. The primary contributions of this study are: A redesigned therapeutic approach with a dynamic threshold has been grown to eliminate the DCP problem correlated with the atmosphere. The use of a dynamic threshold cost further reduces the degree of distortion of the shade. The gradient-oriented image has been modified by looking at an improved manual image. It has the power to get better of that problem of halo artifact and reverse gradient. The created coefficient of the directed gradient for cleaning the image is then used to increase the expected air velocity.

Hu et al. [13] single-image fog elimination is essential for surveillance packages and many defogging strategies had been proposed, recently. The atmospheric conditions such as the scattering properties of foggy photos depends upon the intensity records of scene is not fully handiest now a days, however additionally the atmospheric aerosol model which has more distinguished effect on illumination in a fog scene than in a haze scene. However, present defogging methods complicate haze and fog, and that they fail to consider fully about the scattering houses. Thus, those methods are not sufficient to cast off fog consequences, in particular, for pictures in maritime surveillance. Therefore, this paper proposes a single-photograph defogging approach for visible maritime surveillance. Firstly, a broad scattering version is proposed to rapid a fog photograph in the glow-shaped environmental illumination.

Then, an illumination decomposition algorithm is proposed to dispose of the glow effect on the air mild radiance and get better a fog layer, wherein items on the limitless distance have uniform luminance. Secondly, a transmission map estimation molded on the non-neighborhood haze-traces previous is useful to constrain the transmission map into a reasonable range for the entire fog photograph.

Bansal et al. [14] summarized present country of the paintings picture defogging algorithms. Firstly, a photograph elegance set of policies has been presented, and then, we summarized present country of the art picture healing primarily based photograph defogging algorithms. Finally, we summarized image first-rate evaluation strategies observed by the usage of their comparisons of various photograph defogging algorithms. Problems with image deterioration and the scope of the future were mentioned later. The fog changes because the ratio of the air package reaches the filling space. The actual difference between fog and mist is in the appearance of a much lower due to its predecessor. There are many sorts of fog (e.g., radiation fog, advection fog, and lots of others.) which fluctuate from each other of their formation methods. Here, we can use separate words instead. Aerosol particles could also be small suspended within a gas that area unit bigger than air molecules, however, it is minute than haze droplets. Haze developed a distinct gray or blue color which influences the appearance.

Zhou et al. [15] have been proposed to address the debasement of the photograph and work on the evaluation of the picture. In view of special defogging norms, those standard procedures can be arranged into classifications: Photograph reclamation methodologies dependent on the actual form and picture defogging strategies dependent on photograph improvement. To manage the problem that the retinex set of rules cannot effectively enhance tone and component simultaneously, we exhort a retinex basically based Laplacian pyramid method for picture defogging. The procedure is cultivated by means of MSRCR and Laplacian pyramid, and it does not requires additional equipment contraptions. The normal defogging way incorporates three imperative parts: brightening shading upgrade, detail of reflected picture trouble improvement, and direct weighted combination. Initially, we transfer the gamma adjustment light lower got back to reflected picture to acquire conceal improvement. And afterward, the component upgrade is finished by means of the Laplacian pyramid to technique the reflected picture issue. At long last, the component extra legitimate picture and shade revised picture are utilized to recreate the smooth photograph.

Tufail Z et al. [16] proposed based totally at the dark channel earlier (DCP). Existing nation-of-the-art photograph defogging techniques, the usage of DCP fails to show premiere performance for the challenge of picture defogging. Their outcomes are both low in contrast or compromised with the aid of artifacts. We have projected a new method to examine the transmission map and applied a Laplacian filter out to refine the transmission map. Experimental results show that the proposed technique estimates fog more accurately, and the reconstructed photographs have higher color assessment. Halo effects and thin fog layer which may be seen in defogged pics of other strategies are removed the use of the proposed technique. Furthermore, fog effect, SSIM, DS, and AQI values display that the photographs reconstructed with the aid of proposed approach showcase higher perceptual pleasant as compared

to current photograph defogging methods. In destiny, our proposed technique may be advanced to reconstruct fog unfastened photograph from dense foggy photograph also algorithm can be optimized to reconstruct fog unfastened image on run time. The perceptual great of outside scene photographs is crucial for know-how and analyzing the surroundings to perform automatic tasks together with navigation, object detection, and recognition. Scattering or absorption of light in damaging climate due to fog and haze can significantly restrict the visibility of outdoor scenes. Therefore, pictures taken in such climatic conditions be afflicted by lower comparison, dwindled hues, and luminance imbalance, ensuing in items some distance from camera nearly invisible.

Gao et al. [17] proposed to characterize the extra right barometrical gentle via presenting a versatile variable technique. Then, dissimilar to conventional darkish channel past, a solitary assessment approach is given by utilizing melding dull and gentle channels to gauge additional extraordinary climatic light and conveyance. Then, at that point, we take on the dim picture like hue picture as a steerage photograph to refine the conveyance to likewise bring down the time intricacy. Focusing on the area of low conveyance, an extraordinary remuneration work is made to improve the locale of low conveyance and stay away from shading bending. Also, a straightforward and amazing estimation approach is proposed to conclude boundaries in reimbursement include. At long last, the unmistakable distant photograph is set up through a ventured forward climatic dissipating model. Broad trials on genuine world datasets uncover that the proposed strategy outflanks various diverse spic and span systems each on abstract and objective excellent audits.

Salazar-colors et al. [18] have to address an alternate off between restoration nice, below distinctive fog intensities and situations, and time consumption. On the contrary hand, thinking about exceptional noticeable reaches, the proposed strategy demonstrates a prevalent by and large execution basically in three out of five magnificent degrees, than the wide range of various defogging procedures under uncommon light and assessment circumstances. What is more, the proposed approach achieved the second one pleasant exhibition using the no-reference NIQE metric. At long last, the proposed strategy has the base handling time (underneath comparative states of programming and equipment), thinking about stand-out picture goals, in contrast with all inspected calculations that is a really significant for parcels PC inventive and farsighted projects. Subsequently, the proposed strategy decreases normal DCP antiquities, accomplishing a higher mending on sky districts than the methods in explored writing, bringing down over-soaked regions, and achieving a right reaction with a satisfactory substitute off among stand-out phases of fog and calculation time. This proposed technique has a quicker handling time; at the same time, it could not be thought about an honest assessment for the explanation that, the methodology is executed on Python, and done utilizing GPUs. Also, it is far basic to see that with the most extreme photograph length (19,201,080) considered in this canvases, this strategy conveys a memory goofs; along these lines, it very well might be viewed as that an absolutely significant element of the proposed approach is its usefulness of being carried out in conventional inserted structures, making it fitting for being utilized, in actuality, PC vision projects to perform defogging obligations online.

Tufai et al. [19] have been comprehensively utilized in research for picture defogging strategies dependent on substantial model. Because of the presence of haze, the photograph is debased in expressions of three boundaries this is splendor, conceal correlation, and feature data. Mist will expand scene brilliance, lessens the shading assessment, and disposes of side records. A bunch of rules should have been planned which can act in this manner to upgrade each of the three boundaries. The ensuing advance of this procedure is refinement of transmission map. For refinement of transmission map, particular procedures like, smooth matting, respective channel out, or directed channel are proposed in past techniques. In this work, we proposed a region keeping and smoothing channels. Slopes of transmission map are determined the use of Laplacian sift through, and afterward those angle esteems are deducted from transmission guide to hold region insights. Then, at that point, we carried out directed clear out to smoothen the edge protected transmission map without losing the edge records.

Hassan et al. [20] proposed a unique cascade strategy for the treatment of visibility in foggy pictures. The scheduled cascade is implied totally on a mix of enrichment and body fashions; restricted comparisons with variable bar graph activity (CLAHE) and non-black picture element limit with strategic assumption methods (NBPC + PA). The utilization of CLAHE is consistent with the appearance of a smoggy picture, but it creates a color and disfigurement of the element, improves sound and develop halo result. We rectified this shortcoming by delivering the result to NBPC + PA. To enhance the overall cascade presentation of the two strategies, we determine the appropriate parameters. The proposed cascading uses male or female power in two processes which also provides better results to remove the same fog and chronic fog. We present the goal of first phase testing and improvement of visibility with various fog images. The test results validate advanced fog cascade removal techniques compared to existing fog removal algorithms. The proposed fog removal method that combines the capabilities of CLAHE processes with NBPC + PA demonstrates overall performance to improve the visibility of uniform fog shots and fog shortened shots. In completed test experiments, the proposed fog removal method is compared with other existing fog removal procedures. Performance differences suggest individual improvements in policy clauses for large amounts of visible testing and improvement.

Yuan et al. [21] proposed a clever photograph defogging procedure called NIN-DehazeNet for solo picture. This methodology gauges the transmission map by means of NIN-DehazeNet consolidating network-in-network with single-image dehazing through multiscale convolutional neural networks (MSCNNs). By looking the actual level, the transmission gauge of the murky picture is fundamentally based on the informed rendition, after which the expected barometrical light and figured transmission map is emulated. Broad analyses have shown that the proposed calculation over execution traditional strategies. Picture dehazing is a troublesome issue inside the discipline of PC inventive and farsighted. The explanation of image dehazing is to improve an unmistakable photograph from single-boisterous body because of murkiness, mist, or smoke. The dehazing calculations have consequently been generally thought about, as a difficult case of (unwell presented) photograph rebuilding and

upgrade. Like various difficulties like picture denoising and topnotch choice prior dehazing compositions expected the accessibility of more than one photographs from a similar scene. In any case, the cloudiness expulsion from solo photograph has now gotten the predominant standing, taking into account that, it is miles extra reasonable for sensible settings. This paper centered at the problem of solo picture dehazing.

The comparative study is shown in Table 1.

Table 1 Comparative study of papers

S. No.	Tıtle	Methodology	Demerits
1	Single-image defogging by gain gradient image filter	Gain gradient image filter	Does not support all images
2	Single-image defogging based on illumination decomposition for visual maritime surveillance	Haze-lines prior	Pixel density can be loosed
3	A review of image restoration-based image defogging algorithms	Independent component analysis (ICA)	Does not support high-level fog
4	Retinex-based Laplacian pyramid method for image defogging	Laplacian pyramid method	Difficult to reconstruct the images
5	Improved dark channel prior for image defogging using RGB and YCbCr color space	Dark channel prior	Image quality is less
6	A novel UAV sensing ımage defogging method	Dark channel prior (DCP) method	Provide time complexity is high
7	Fast single-image defogging with robust sky detection	Hybrid subjective testing set from realistic single-image dehazing	This is conventional method to provide defogged images
8	Optimization of transmission map for improved image defogging	Adaptive defogging method	Provide high number of noise ratio
9	Cascaded method of removing image index based on physical models and enhancements	Contrast limited adaptive histogram equalization (CLAHE)	Structure can be loosed
10	Single-image dehazing via NIN-DehazeNet	DehazeNet framework	Support limited datasets

3 Existing Methodologies

Fog, rain drops, and haze pixels are removed from photographs using a variety of ways. And some methods are,

3.1 Dark Channel Prior

Need to the dark channel to get a more reasonable outcome, the dull channel earlier is eliminated from the assessment of encompassing light in the dehazed picture [22, 5]. This methodology is most generally utilized for non-sky patches, which have exceptionally low force at a couple of pixels in a single-shading channel. Due to three factors, the dim channel's low power overwhelms are:

1. Bright things or surfaces
2. Shadows (shadows of vehicle, structures, and so on)
3. Dull things or surfaces (dark tree trunk and stone).

Since open air photographs are frequently loaded up with shadows, the picture's dim channels will be very dim. A hazy picture is more splendid than its non-hazy partner because of mist (air-light). Thus, the dark channel of a hazy picture will have a higher force in regions, where there is more mist. Thus, the visual force of the dark channel is a good guess of the mist thickness. To accomplish great outcomes in dim channel previously, we apply pre and post-handling stages. Delicate matting, three-dimensional sifting, and other post-handling methods are utilized [23, 24]. Let $I(z)$ is input picture, $Y(z)$ is murky picture, and $J(z)$ is the transmission of the medium. The weakening of the picture can be communicated as:

$$Y_{att}(z) = I(z)J(z) \tag{1}$$

The influence of the fog about air-light effect is expressed as:

$$Y_{\text{airlight}}(z) = K(1 - J(z)) \tag{2}$$

Dark channel for the random image I is expressed as:

$$I^{\text{dark}}(z) = \frac{\min}{A \in \Omega(z)} \left(\min I^{c}(A)\right) \tag{3}$$

In this, I^c is the shading picture involving RGB parts, and (z) portrays a nearby fix which has its starting point at z. The low power of dim channels is a direct result of shadows in pictures, shading items, and dull articles in pictures. After dim channel earlier, we really want to appraise transmission $J(z)$ for continuing further with the arrangement. The transmission map profundity is created with the help of assessing colors expect atmospheric light K [25].

3.2 Bilateral Filtering

With a nonlinear combination of close by picture values, bilateral filtering smooth's pictures while also preserving edges. Bilateral is non-iterative, local, and simple. The bilateral filter includes a gray vary or colors primarily based mostly on their geometric similarities and measuring similarities, and prefers approximate values to distant values in each area and width. Piecewise constant answers are obtained by using a bilateral filter with clean edges. The use of a two-dimensional filter no longer provides a reduction in working efficiency. Sorting does this by changing each pixel by the weight of its nearby pixels. With the distance in the image plane and the gap on the depth axis, the weight assigned to each one nearby pixel decreases. In comparison to other methods, this clear out allows us to arrive about faster [26, 27]. We employ preprocessing and submit remodeling processes for improved results while using the bilateral filter. The histogram measurement is used as a preprocessing step, while the histogram extension is used to prepare the dispatch. These steps help to produce an image evaluation before and after the reality using the sided channel. Because this set of criteria is independent of fog density, it may be applied to photographs taken in dense fog. It would not necessitate user involvement. It is terribly helpful for chase and navigation for client the industry. Bilateral clean out is a non-abrasive filter that protects the edges of images. In any pixel in an image, the intensity cost is replaced by a weighted depth value averaged from surrounding pixels. This weight could be based on a Gaussian distribution. Importantly, the weights are reached from the distance between Euclidean pixels and the radiometric variances (example, range differences, which include intensity distance, color depth, and so forth.). This ensures crisp edges by looping through all pixels in a methodical manner and modifying the weights of adjacent pixels accordingly [28–30].

3.3 Clahe

Contrast limited adaptive histogram equalization (CLAHE). CLAHE is a procedure for developing picture with minimum probability. For the processing of fogged photographs, this method does not require any forecasted climatic statistics. For starters, the photograph taken by the camera under fog conditions is changed from red, green, and blue (RGB) to HSV (color, fill, and light) color space [31]. The pictures are changed because human sense hues, as well as HSV, are used to create colorations. Second, CLAHE is used to process the value item without disturbing the hue and saturation. Histogram equalization is used in the context area. The cropped pixels are redeployment in each gray section from the original histogram. In this way, all pixel data is reduced to a user-selected size. At last, the picture refined in the HSV color space is transfigured to the RGB color space. Each pixel of the original photograph is at the contextual vicinity's center. The clipped pixels from the unique histogram are reallocated to every gray stage [32, 33]. Because each pixel depth

is limited to a user-defined maximum, the new histogram differs from the original histogram. As a result, CLAHE restricts noise enhancement. The set of rules provided in this work is intended to alleviate the complexity of the CLAHE regulations. To sub-snap shots, the CLAHE method employs histogram correlation. Every pixel of the original photograph is contained inside the closest factor of the sub-photograph. The sub-initial image's histogram is reduced, and the reduced pixels are realignment across each and every gray level [34, 35]. The new histogram differs from the previous one in that the intensity of each pixel is restricted to a patron-determined maximum. As a result, CLAHE can lessen noise enhancement.

4 Atmospheric Light Estimation Back Propagation Neural Network

One of the primary goals of photograph processing is to obtain a more favorable image. Outdoor images are impaired by atmospheric phenomena such as fog, haze, and so on, and many applications, such as customer/computational images and computer vision require a vision-enhanced image. Because the light goes to the observer, atmospheric particles absorb and disperse it, resulting in haze and fog. Airlight shifts the tone of the image in these damaged photographs, which lose comparison. Reducing fog improves image viewing quality and corrects color change. Here, we achieve the haze-free photograph via fusion method. The good facts from each of the offered pictures are fused together by the photograph fusion method to form a consequent photograph whose pleasantness is advanced to any of the input images. This is accomplished by applying Laplacian to the input and Gaussian to the weighted inputs [36, 37]. The final image is created by integrating the magnified statistics from the input photographs into a single image. The goal of a fusion-based technique is to extract two photographs from a single image in order to combine them efficiently. We just need the fundamental capabilities of the photographs for blending, therefore, we cross for the load maps. We recommend a completely new type of explicit photograph clean out called guided filter out in this challenge. The clarified result is a linear rebuild of the steering picture on a local level This filter came out in the exact way to keep the edge as the filter came out, but it is not affected by the art objects that slow down the gradient. It is also linked to the matting Laplacian matrix, making it a more general idea that can be applied to a variety of applications beyond "smoothing." Furthermore, for each gray-scale and coloring image, the guided clean out has an O (N) time (within the large range of pixels N) accurate algorithm [38, 39]. Experiments show that the guided clear out performs very well in phrases of both first-rate and performance in video dehazing or defogging techniques.

4.1 Gaussian Filter

The climatologic light values can be assessed using a Gaussian filter. When light coming from a portion scope an observer, atmospheric transmission refers to how effectively the light is kept. It is a positive scalar number ranging from 0 to 1, with higher values indicating better reflectiveness [40, 41]. Transmission is regularly associated with picture improvement through as portrayed in after equation.

$$Y(z) = I(z)J(z) + K(1 - J(z)) \tag{4}$$

where z is the rotation of a two-dimensional sphere, $Y(z)$ is the orientation, $I(z)$ is the square light, K is the ambient light (atmosphere), and $J(z)$ is the atmosphere transfer. The first term to the right of Eq. 1 is connected in contrast to the amount of light emanating from the scene which is distributed outside the image path and thus increases with enhanced transmission. The subsequent term is the amount of encompassing light commonly from the sun that is dispersed into the pictorial pathway and hence decreases with further developed transmission. In extreme cases, the visual image may be a glare of a scene or an ambient scattering light. With the light of the atmosphere and the transmission map, we can restore the brightness of the spot by (1). But the exact withdrawal name $I(z)$ $J(z)$ can be extremely near to zero when the transmission of $J(z)$ is near to zero. The square light received straight from I is usually noisy. Accordingly, we limit the transfer of $J(z)$ with a low J_0 bond, that is, we keep a small amount of fog in

$$I(z) = \frac{Y(z) - K}{\max(J(z), J_0)} + K \tag{5}$$

A common place worth of J_0 is 0.1. Since the scene brilliance is typically not generally so splendid as the environmental light, the picture after murkiness expulsion looks faint. Hence, we increment the openness of $I(z)$ for show.

4.2 Back Propagation Neural Network Algorithm

To begin, estimate the color space and white balance, then rough estimation of atmospheric color space, and then restore the image using local extrema. A new way to create multi-dimensional image decay with edge preservation. We show that current filter-based decay algorithms are limited in their ability to extract information at negligent sizes [42, 43]. We have created a simple interactive tool that allows you to change the tone and brightness of details on a variety of scales. We start by building a three-level light source from the image (a rough base of level b and two levels of detail). This is concluded using the first (non-repeated) establishment. The user is

then provided with a set of slides to control the presentation of the base layer, as well as enhanced features, for the base, and layers of medium, and fine data [44].

4.2.1 Input Image (Preprocess-M × N Matrix)

Given a picture with N pixels and element dimensionality of M, all the element vectors can be incorporated into a M N framework, I. Expecting that there are L delegate includes, the picture model can be communicated as:

$$I = X\beta + \varepsilon \tag{6}$$

where X is a M L grid whose sections are agent highlights, is a L N framework whose segments are weight vectors, and is model mistake.

4.3 Atmospheric Light Estimation

At the point when light exuding from a scene arrives at an eyewitness, it is alluded to as climatic transmission. It is a positive scalar amount with a scope of 0 to 1, with higher qualities demonstrating better perceivability. Transmission is ordinarily identified with picture arrangement through as depicted in after equation.

$$Y(z) = I(z)J(z) + K(1 - J(z)) \tag{7}$$

where z is a two-dimensional spatial variable, $Y(z)$ is the noticed picture, $I(z)$ is the scene brilliance, An is the surrounding (environmental) light, and $J(z)$ is the climatic transmission. The initial term on the right half of Eq. 1 is conversely identified with the measure of light transmitting from the scene that is dispersed out of the visual pathway and consequently increments with further developed transmission. The subsequent term is the measure of encompassing light regularly from the sun that is dispersed into the visual pathway and hence diminishes with further developed transmission. In the limits, the apparent picture can either be only the scene brilliance or simply the dissipated surrounding light. With the environmental light and the transmission map, we can recuperate the scene brilliance as per (1). However, the immediate constriction term $I(z)$ $J(z)$ can be extremely near zero when the transmission $J(z)$ is near nothing. The straightforwardly recuperated scene brilliance I is inclined to clamor. Subsequently, we limit the transmission $J(z)$ by a lower bound J_0, i.e., we save a modest quantity of dimness in extremely thick cloudiness districts. The last scene brilliance $I(z)$ is recuperated by

$$I(z) = \frac{Y(z) - K}{\max(J(z), J_0)} + K \tag{8}$$

A normal worth of J_0 is 0.1. Since the brightness of the location usually does not shine as brightly as the atmosphere, the picture after the elimination, it looks pale. We therefore amplify the exposure of $I(z)$ to exposure.

4.4 Feature Extraction and Estimation

To exclude the feature, layers of Maxout units are used, which may produce almost all of the features related to the environment.

A new activation function called the Maxout unit is also selected as an indirect size reduction map, promoted by a suggested view, which analyzes color channels for those fog-related features [45, 46]. In most layers, the Maxout unit is a simple function for transferring indirect feeds. It removes fog-related aspects using automatic learning in place of heuristic methods. Multi-dimensional features, which incorporate in-depth image elements to include multiple scales, have been shown to be helpful in removing membership. Scale invariance can also be achieved using multiscale feature extraction [47, 48]. To build the visual impression of three-scale discoveries, we utilize the multiscale tone control method. As per the above examination, we will control tone and differentiation of subtleties at three scales, which incorporate including reestablished pictures R, G, and B acquired from negligible worth, maximal worth, and mean worth.

4.5 Local Extrema

Nearby extrema method, which uses edge-mindful interjection to register envelopes. A smoothed mean layer is gotten by averaging the envelopes. The technique can remove fine-scale detail paying little mind to differentiate. Nonetheless, single-mean layer is not adequate to well estimated climatic cloak, and it is addressed by iterative computation, which is tedious.

Our nearby extrema non-iterative strategy comprises of three stages:

(1) Identification of nearby extrema of V (coarse, fine, and medium qualities);
(2) Inference of extremal envelopes; (Find the cloudiness esteems)
(3) Visibility improvement of the outcome by multiscale tone control calculation.

4.6 Post-Processing (NonLinear Regression)

Rectified linear unit (ReLU), which offers a few presentations. In any case, ReLU is built for partition problems and is not entirely ready for retrieving problems like picture recovery. Partition problems mean the classification error which calculates the mean square error. When accuracy increases, error rate will be reduced.

It could make response overspill, particularly in the last layer, on the grounds that the result upsides of the last layer are needed to be lower and higher obliged in a little reach for picture reclamation.

Local maximum filters for third layer in this in-depth study remove the local estimation error [46, 49]. At the point, when the transmission $J(z)$ is near zero, the immediate constriction term $I(z)$ $J(z)$ may be very near to zero. In an integrated network, ambient light cannot be considered a universal constant that will be studied close to medium transmission. In addition, we assume that the space scattering model can also be studied in the proposed work, without the average transmission rate, the end-to-end map between neutral images can be developed spontaneously.

5 Result and Discussion

The proposed system is to remove the degraded factors from the videos using guided filtering and deep learning algorithm (BPNN) with improved accuracy rate. Table 2 shows the existing system methodology and its demerits and merits of proposed system.

Table 2 Result and discussion

S. No.	Tıtle	Methodology	Demerıts	Merıts
1	Single-image defogging by gain gradient image filter	Gain gradient image filter	Does not support all images	Support all images
2	Single-image defogging based on illumination decomposition for visual maritime surveillance	Haze-lines prior	Pixel density can be loosed	Pixel density cannot be loosed
3	A review of image restoration-based image defogging algorithms	Independent component analysis (ICA)	Does not support high-level fog	Support high-level fog
4	Cascaded method of removing image index based on physical models and enhancements	Contrast limited adaptive histogram equalization (CLAHE)	Structure can be loosed	Structure of the image is predicted and cannot be loosed
5	A novel UAV sensing image defogging method	Dark channel prior (DCP) method	Provide time complexity is high	Time complexity is low

6 Conclusion

Haze, which can be caused by dust, smoke, or other dry matter, reduces the visibility of protected areas by giving the images a normal gray color. Due to poor visibility, the foggy picture suffers from low radiance and sharpness. The presence of fog, mist, or smoke, which darkens colors and reduces visual acuity, is one of the most important challenges to image processing in the open air. Removal of mist or fog is difficult due to the reliance of fog on endless deep information. Multi-mode operating range is reduced when the weather is inclement. A quick and effective solution for real-time photograph and video retrieval are suggested in this study. The removal of a single image without the need for additional information is defined as a particular filtering problem using a newly introduced image prior—dark channel prior, and the upgraded filtering system is proposed based on the guided filter. The aircraft light and low sample transmission can be easily analyzed and excavated using the existing method. The transmission can be improved and up-sampled using a guided filter. Research results show that the technology provided is able to remove the haze layer and achieve real-time performance. It is believed that many presentations, such as outdoor surveillance programs, smart car systems, remote sensor systems, and photograph editors can benefit from the proposed approach.

References

1. Murugesan M, Thilagamani S (2020) Efficient anomaly detection in surveillance videos based on multi layer perception recurrent neural network. J Microprocess Microsyst 79(November)
2. Zhang L, Dong WS et al (2010) Two-stage image de-noising by principal component analysis with local pixel grouping. Pattern Recog 43(4):1531–1549
3. Deepa K, Kokila M, Nandhini A, Pavethra A, Umadevi M (2020) Rainfall prediction using CNN. Int J Adv Sci Technol 29(7 Special Issue):1623–1627
4. Tarel JP, Hautière N, Caraffa L (2012) Vision enhancement in homogeneous and heterogeneous fog. IEEE Intell Trans Syst Magaz 4(2):6–20
5. Zhu Q, Mai J, Shao L (2015) A fast single image haze removal algorithm using color attenuation prior. IEEE Trans Image Process 24(11):3522–3533
6. Ren D, Zuo W, Hu Q, Zhu P, Meng D (2019) Progressive image deraining networks: a better and simpler baseline. In: Proceeding of IEEE/CVF conference on computer vision pattern recognition (CVPR), pp 3937–3946
7. Wang C, Xing X, Wu Y, Su Z, Chen J (2020) DCSFN: deep cross-scale fusion network for single image rain removal. In: Proceeding of 28th ACM international conference on multimedia, pp 1643–1651
8. Thilagamani S, Nandhakumar C (2020) Implementing green revolution for organic plant forming using KNN-classification technique. Int J Adv Sci Technol 29(7S):1707–1712
9. Nishino K (2012) Bayesian defogging. Int J Computer Vis 98:263–278
10. Caraffa L, Tarel JP (2013) Markov random field model for single image defogging. Proceeding of IEEE intelligent vehicles symposium, Gold Coast, Australia, pp 994–999
11. Wei W, Meng D, Zhao Q, Xu Z, Wu Y (2019) Semi-supervised transfer learning for image rain removal. In: Proceeding of IEEE/CVF conference on computer vision pattern recognition (CVPR), pp 3877–3886

12. Singh D, Kumar V (2019) Single image defogging by gain gradient image filter. Sci China Inf Sci 62(7):79101
13. Hu H-M et al (2019) Single image defogging based on illumination decomposition for visual maritime surveillance. IEEE Trans Image Process 28(6):2882–2897
14. Bansal B, Sidhu JS, Jyoti K (2017) A review of image restoration based image defogging algorithms. Int J Image, Graph Sig Process 9(11):62
15. Zhou J et al (2019) Retinex-based laplacian pyramid method for image defogging. IEEE Access 7:122459–122472
16. Tufail Z et al (2018) Improved dark channel prior for image defogging using RGB and YCbCr color space. IEEE Access 6:32576–32587
17. Gao T et al (2020) A novel UAV sensing image defogging method. IEEE J Select Top Appl Earth Observ Remote Sens 13:2610–2625
18. Salazar-Colores S et al (2020) Fast single image defogging with robust sky detection. IEEE Access 8:149176–149189
19. Tufail Z et al (2019) Optimisation of transmission map for improved image defogging. IET Image Process 13(7):1161–1169
20. Hassan N et al (2020) A cascaded approach for image defogging based on physical and enhancement models. Signal Image Video Process 1–9
21. Yuan K et al (2019) Single image dehazing via NIN-DehazeNet. IEEE Access 7:181348–181356
22. Rajesh Kanna P, Santhi P (2022) Hybrid intrusion detection using map reduce based black widow optimized convolutional long short-term memory neural networks. Expert Syst Appl 194
23. Thilagamani S, Shanti N (2014) Gaussian and gabor filter approach for object segmentation. J Comput Inf Sci Eng 14(2):021006
24. Yang W, Tan RT, Wang S, Fang Y, Liu J (2020) Single image deraining: from model-based to data-driven and beyond. IEEE Trans Pattern Anal Mach Intell. https://doi.org/10.1109/TPAMI.2020.2995190
25. Deepika S, Pandiaraja P (2013) Ensuring CIA triad for user data using collaborative filtering mechanism. In: 2013 International conference on information communication and embedded systems (ICICES), pp 925–928
26. He J, Yu L, Xia GS, Yang W (2020) Single image deraining with continuous rain density estimation. arXiv:2006.03190
27. Perumal P, Suba S (20200) An analysis of a secure communication for healthcare system using wearable devices based on elliptic curve cryptography. J World Rev Sci Technol Sustain Dev 18(1):51–58
28. Li Y, Tan RT, Guo X, Lu J, Brown MS (2016) Rain streak removal using layer priors. In: Proceedings of IEEE conference computer visual pattern recognition (CVPR), pp 2736–2744
29. Gunasekar M, Thilagamani S (2020) Performance analysis of ensemble feature selection method under SVM and BMNB classifiers for sentiment analysis. Int J Sci Technol Res 9(2):1536–1540
30. Wei Y, Zhang Z, Wang Y, Zhang H, Zhao M, Xu M, Wang M (2021) Semi-deraingan: a new semi-supervised single image deraining. In: Proceeding of IEEE international conference on multimedia expo (ICME), pp 1–6
31. Li G, He X, Zhang W, Chang H, Dong L, Lin L (2018) Non-locally enhanced encoder-decoder network for single image de-raining. In: Proceeding of 26th ACM International Conference on Multimedia, pp 1056–1064
32. Pandiaraja P, Sharmila S (2020) Optimal routing path for heterogenous vehicular adhoc network. Int J Adv Sci Technol 29(7):1762–1771
33. Wang X, Li Z, Shan H, Tian Z, Ren Y, Zhou W (2020) Fastderainnet: a deep learning algorithm for single image deraining. IEEE Access 8:127622–127630
34. Logeswaran R, Aarthi P, Dineshkumar M, Lakshitha G, Vikram R (2019) Portable charger for handheld devices using radio frequency. Int J Innov Technol Expl Eng (IJITEE) 8(6):837–839

35. He K, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell 37(9):1904–1916
36. Pandiaraja P, Aravinthan K, Lakshmi NR, Kaaviya KS, Madumithra K (2020) Efficient cloud storage using data partition and time based access control with secure aes encryption technique. Int J Adv Sci Technol 29(7):1698–1706
37. Pradeep D, Sundar C (2020) QAOC: noval query analysis and ontology-based clustering for data management in Hadoop 108:849–860
38. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N (2020) An image is worth 16x16 words: transformers for image recognition at scale. arXiv:2010.11929
39. Zhang QL, Yang YB (2021) SA-net: shuffle attention for deep convolutional neural networks. In: Proceeding of IEEE international conference on acoustics speech signal processing (ICASSP), pp 2235–2239
40. Li X, Wu J, Lin Z, Liu H, Zha H (2018) Recurrent squeeze-and-excitation context aggregation net for single image deraining. In: Proceeding of European conference computer vision (ECCV), pp 254–269
41. Wang T, Yang X, Xu K, Chen S, Zhang Q, Lau RWH (2019) Spatial attentive single-image deraining with high quality real rain dataset. In: Proceeding of IEEE/CVF conference on computer vision pattern recognition (CVPR), pp 12270–12279
42. Rajesh Kanna P, Santhi P (2021) Unified deep learning approach for efficient intrusion detection system using integrated spatial–temporal features. Knowl-Based Syst 226
43. Deepa K, Thilagamani S (2019) Segmentation techniques for overlapped latent fingerprint matching. Int J Innov Technol Expl Eng 8(12):1849–1852
44. Ledig C, Theis L, Huszar F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, Shi W (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: Proceeding of IEEE conference on computer vision pattern recognition (CVPR), pp 4681–4690
45. Touvron H, Cord M, Douze M, Massa F, Sablayrolles A, Jégou H (2020) Training data-efficient image transformers and distillation through attention. arXiv:2012.12877. [Online]. Available: http://arxiv.org/abs/2012.12877
46. Yang W, Tan RT, Feng J, Liu J, Guo Z, Yan S (2047) Deep joint rain detection and removal from a single image. In: Proceeding of IEEE conference on computer vision pattern recognition (CVPR), pp 1357–1366
47. Xu W, Xu Y, Chang T, Tu Z (2021) Co-scale conv-attentional image transformers. arXiv:2104.06399
48. Santhi P, Mahalakshmi G (2019) Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction. Int J Eng Adv Technol 8(4):839–846
49. Chen D, He M, Fan Q, Liao J, Zhang L, Hou D, Yuan L, Hua G (2019) Gated context aggregation network for image dehazing and deraining. In: Proceeding of IEEE winter conference on applied computer vision (WACV), pp 1375–1383

The Role of Social Media Technology in Increasing Frozen Food Sales as the Agribusiness Products

Elvin Sestomi, Revaldo Dhamacora, Valentino Felix Aswar, Wilianti, Ford Lumban Gaol, Tokuro Matsuo, and Natalia Filimonova

Abstract During the recent COVID-19 outbreak, many people opted to stay at home and completely run their businesses online. The recent research studies show that, the frozen food sales and social media usage have grown exponentially as a result of the change in people's behavior and companies' marketing strategies. Frozen food sales are expected to grow as a result of social media marketing, and also, the relationship between buyers and sellers is becoming more flexible. The main purpose of this research study is to examine the recent increase in frozen food sales as a result of social media's influence. The proposed research work has been carried out by using qualitative technique, and also, this research study has conducted a literature analysis to learn from previous research works based on what factors they contribute to increase sales, and why social media is the best medium for marketing. The literature study suggests that frozen food sales rise as a result of consumer behavior, and that social media is the best platform for marketing their products. The results indicate that the behavioral changes have a significant impact on frozen food sales, and that social media assists in product marketing and enhancing the key drivers of

E. Sestomi · R. Dhamacora · V. F. Aswar · Wilianti
School of Information System, Bina Nusantara University, Alam Sutera, Kota Tangerang, Indonesia
e-mail: Elvin.Sestomi@binus.ac.id

R. Dhamacora
e-mail: revaldo.dhamacora001@binus.ac.id

V. F. Aswar
e-mail: valentino.aswar@binus.ac.id

F. L. Gaol (✉)
Binus Graduate Program—Doctor of Computer Science, Jakarta, Indonesia
e-mail: fgaol@binus.edu

T. Matsuo
Advanced Institute of Industrial Technology, Tokyo, Japan
e-mail: matsuo@aiit.ac.jp

N. Filimonova
Vladimir Branch of Russian Academy of National Economy and Public Administration, Vladimir, Russia

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_19

growing consumer interest. As a result, a large number of frozen food companies are considering the option of utilizing social media as a marketing tool.

Keywords Social media · Marketing · Sales · Frozen food

1 Introduction

The food industry is one of the stable developing industries in Indonesia. In addition to stable development, the food industry is also showing positive growth. Changes in the lifestyle of high-consumer Indonesians are accompanied by increased purchasing power of people, leading to changes in consumption patterns, and people are more likely to buy food from restaurants than to cook at home [1].

According to Thearling [2], the culinary industry shows a tremendous growth in various parts of the world [3]. The industry, especially frozen food, has a lot of potential. Frozen food may develop swiftly if well planned, managed, and operated since it is durable, convenient, and easy to use [4]. Even though it appears to be good, it reveals that the current culinary trend is a growing number of restaurants and cafes with a range of distinct and intriguing themes, while others have closed their doors. As a result, many company owners, particularly frozen food producers must not only continue to advertise their products or services throughout the COVID-19 epidemic, but also shape the attractiveness of the techniques, tactics, and messaging employed in promotion in the case of a crisis [5].

Many entrepreneurs choose social media marketing because it is one of the most popular promotional breakthroughs. Many businesses today use social media as platforms or tools in their marketing because social media is more promising than other marketing channels as it plays an intermediate role in purchases 87% of the time, which is one of the reasons for its global appeal. Consumers and businesses share text, image, audio, and video information with one another and with businesses via social media, and vice versa [6].

These days, a lot of people are using social media. According to available data, Indonesia's active social media users' number has accumulated to 170 million out of the country's total population of 274.9 million people [7]. Because of the pandemic, total national frozen food consumption is predicted to increase by 17% in 2020, to roughly 11.58 million tons [8], then by March 2018, the participation rate for hot processed meat-eating, such as nuggets, has risen to 27.31%. Over the past year, the percentage rise has been 6.98%, compared to March 2017, when it was 20.33% [9]. We are presenting this research paper with the aim to explain the rise of frozen food in recent years and how social media can help us in marketing for the product. Research question:

1. What are the key factors of sales increase for frozen food products?
2. How is social media able to increase the interest of customers in buying frozen food products?

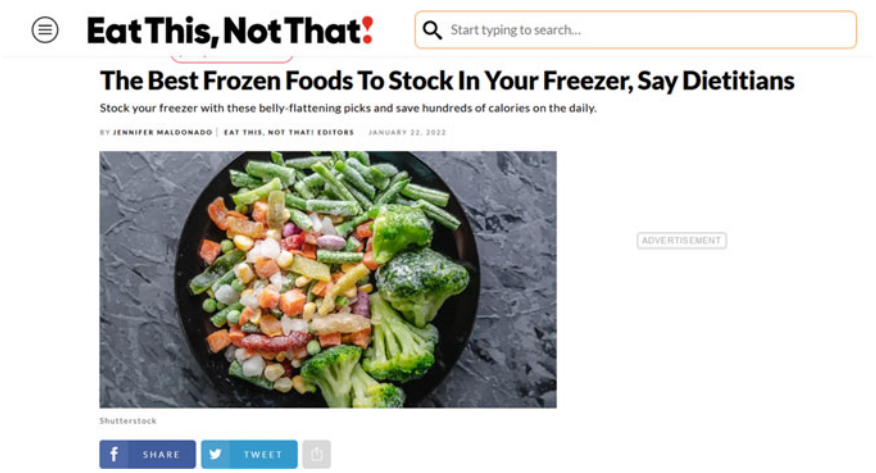


Fig. 1 Example of blog

1.1 Social Media

Social media is an Internet-based channel that is narrow with a wide audiences to interact with each other on occasion and derive value from the perception of interactions with other users, and user-generated content either in real time or asynchronously, it allows you to selectively present yourself with both viewers [10]. Social media includes various kinds of forums such as corporate discussion boards, blogs and chat rooms, consumer products, consumer emails, or service review websites. It will be. (Website with digital audio, images, film, or photograph) [11] (Figs. 1 and 2).

1.2 Social Media Marketing

An integrated business strategy for a company that converts connections (networks) and social media interactions (influence) into valuable marketing strategies in order to reach the desired marketing result is the depiction of social media marketing. It is founded on the implementation of intentional engagement efforts. In addition, a thorough examination of customers' reasons for using brand-related social media [12].

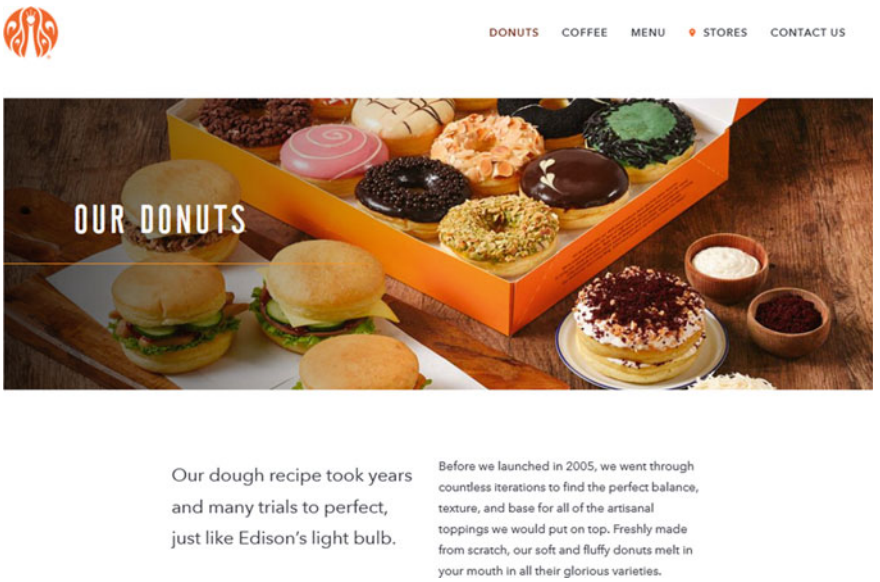


Fig. 2 Example of website advertisement

1.3 Frozen Food

The food that is packed, frozen snacks that just require heating before serving, we know it as frozen food. We can see from the Merriam Webster, ready-made frozen foods are defined, as "special foods that may be swiftly prepared and served." These nutrients are made with minimal regard for quality or importance and are meant for immediate use, availability, or consumption. Low temperatures are used to store this type of food. Frozen instant foods are ready-to-eating foods that have been appropriately packaged and sold. In terms of a healthy diet, this type of food is special. It must be protected by a safety cover. The package must include the date of manufacturing and the date of expiration [13].

Frozen food is any cooked or partially cooked dish that has been frozen at 20 °C for at least 1.5 months and has a shelf life of at least 1.5 months. Before being frozen and used, frozen foods are frequently partially cooked, then packed in bags and snap frozen [14]. For example, there are various products in the fast-food industry, such as packaged convenience foods, nuggets, sausages, potatoes (wedges and steak fries) [9].

1.4 Sales

A sale is known as an agreement between two or more parties where the buyer exchanges their money for something else, that might be a product, services, or other forms. In some circumstances, the seller's loss may be offset by other assets. A sale is a pricing agreement between a buyer and seller of securities in a financial market [15].

1.5 Brand Awareness

When a customer sees or thinks of a brand, an image or logo that comes to mind when they see or think of the brand is referred to as brand awareness. It exists to aid people in recalling a particular brand and to instill a favorable image of the brand in their minds. Because buyers already have a positive impression of the brand, good brand awareness will aid future product sales [15].

2 Research Technique

Our research is done by doing a literature review on several papers or journals that have some correlations with our topic. We searched for 15 references before dissecting it further to learn the cause of the increase in sales for frozen food products, how social media is able to increase the interest of customers on buying frozen food products, what social media is, and so on.

Frozen Food Market Analysis

Research on factors of buying frozen food within the society has demonstrated that there are several factors that influence people's behavior toward frozen food [13]. Socio-demographic characteristics have an impact on actual purchasing behavior where it could be categorized as a person's age, education, and income [16]. On the age factor, young people and older people have different opinions about lifestyle. In general, as people mature, they become more aware of the health benefits of a balanced diet [17]. Other studies show that educated people are more likely to buy frozen meals than ignorant people [18]. Change in the food culture is also influencing people's behavior as well, like COVID-19 makes people stay at their home because concern of getting infected makes individuals develop an attitude where concern always comes wherever they think about what and where they are going to eat [19]. Then, instead of eating their food at restaurants, they will eat it at home, but on a ready-made basis [13].

Usually, the frozen products made from minced meat are very diverse in the market and are consumed quite a lot. Processed meat consumption has grown quickly

at the national level in Indonesia, according to survey data from the Department of Agriculture. Processed meat consumption accounts for 10.28% of total annual production. According to the results of the survey, sausage consumption is increasing at a rate of 4.46% per year. Frozen foods like nuggets, meatballs, and sausages are the most popular by-products of processed meat. Growing consumer demand for frozen foods is driving the growing demand for chicken and beef [20]. Other examples include the Indonesian restaurant company Es Teler 77, which has been delivering frozen versions of its popular meals such as meatballs and scallops to Greater Jakarta clients via social media since the end of March 2020. Since then, they have been able to sell hundreds of frozen foods per week [21].

Social Media as Marketing Medium

Social media marketing does have a lower relationship to the customer buying interest, about 0.794 in the interval of 060–0799. From previous research conducted by Siswanto in 2013, it can be seen from the coefficient score contribution determination that is given by the social media marketing variable toward buy interest.

The research that has been done on J.Co Donut and Coffee on their social media accounts as their marketing medium has turned out to be great. Due to the J.Co Company that is very excellent in managing their social media account, the social media helps J.Co Donut to increase the customer satisfaction and customer buy decision. In this case, the social media helps to build up the J.Co brands and the product description becomes more reachable by the customer due to the availability of the social media that we can access anytime anywhere as long as we are connected to an Internet connection [22].

Client purchasing decisions and the impact of social media marketing and brand awareness

With J.Co donuts and coffee in Semarang, early researchers [22] looked at the impact of social media marketing and brand awareness on customer purchase decisions. The research shows that social media accounts will be more attractive if they give attractive and useful information and post. And we have to control the account activity so the posts are updated frequently so the viewer won't get bored. With the increasing number of followers in social media, the brand awareness will be better because the product's brand is reachable through social media that can be accessed easily, and the price can be a concern. This research stated that the J.Co's products have a lower customer buying interest because of the higher price. So social media marketing has a great impact on brand awareness to increase customer reachability, and then it will help to increase customer buying interest but the price too has a significant impact against customer buying interest.

But social media marketing needs an admin to manage the account so it will always be up to date, this will have a hidden cost like the cost for the electricity and the human resource needed as the account admin before they can really get the impact, so this is like investing and we all know that investing can give us profit but also can cost us money for nothing in return. So many people will be afraid to implement this.

This research is valuable because it demonstrates the link between social media marketing and brand recognition, as well as how it affects customer buying interest and decision making, which is crucial for our study but, this research is about the food that is processed in the company and can be eaten right away meanwhile our research is conducted to the food that still needs to be processed after the customer bought it.

3 Research Method

The research method employed here was a qualitative study with an exploratory strategy along with the goal of obtaining an explanation and description of the research problem as an initial finding that will serve as a foundation for subsequent research [23]. The studies that were selected are determined by the relationship between the research and studies that have been collected. Non-numerical data is also collected and reviewed in this study in order to better comprehend thoughts, opinions, or experiences, and it can be used to gain a better knowledge of a subject or come up with new research ideas (Fig. 3).

Qualitative approaches are very beneficial for developing or constructing theories, conceptual frameworks, or hypotheses, to put it another way. These methods can also be used to refine theories and hypotheses through preliminary testing. There is an important connection to be made here between thorough qualitative and quantitative research. Researchers often strive to reject a null hypothesis in quantitative research, based on epistemological assumptions about whether it is ever possible to confirm (rather than disconfirm) a hypothesis. The study's key finding is that qualitative researchers should hunt for data that will lead to the hypothesis being rejected when they notice a pattern that may be stated as a hypothesis. When qualitative researchers practice this discipline, they take clear and effective steps to minimize bias, whether deliberate or unintentional.

Although qualitative approaches are more commonly employed to build and revise theories, others claim that they can also be used to test theories. Donald Campbell, a seemingly quantitative social scientist, has claimed that qualitative methodologies can be utilized to acquire data for experimental studies (1975). Some qualitative researchers may argue that this is an improper application of their method since it entails a considerable change away from the type of research that involves in-depth, multi-dimensional investigation of a small number of complex holes, as Charles Ragin stated.

Fig. 3 Flowchart

3.1 Data Collection and Procedures

Reviewing, assessing, and comparing the theories of previous study that have a connection to the research that we are performing were used to collect data. Due to our research in researching social media's role in increasing frozen food sales, we will find a paper that discusses frozen food, key success of a frozen food product, application of social media in marketing, impact of social media, and product sales, and then we will analyze the theoretical explanation and case studies in the paper to find an answer to our research question.

4 Result and Discussion

4.1 Result

From our literature review result, we learned that social media can be used for marketing because there is. To successfully make a profit from the frozen food business, one must understand the factors that affect it, and based on our research, we discovered that various characteristics, such as age, education, and money, influence people's attitudes regarding frozen food. Previous research conducted by Anggraeni et al. [20], also tells us that there is a significant increase of processed meat consumption like sausage, nugget. Other than that, the current situation also affects the market. Due to the COVID-19 pandemic, most people have converted from dining in to eating at home and it will be on a ready-made basis or pre-packaged food that you can reheat or simply cook on your own. The said method has been applied by Es Teler 77, a famous restaurant chain in Indonesia, which has also shown significant profit by selling hundreds of frozen foods weekly via social media. This proves that there is a pretty big market for the frozen food business through social media marketing.

In order to increase sales, we also found that social media marketing plays an important role. Although it has a lower relationship with the customer's buying interest, it still plays a major role in increasing customer's satisfaction and brand awareness. Potential and existing customers are able to get the product description easily through social media, landing the business more chances of sales. Moreover, posting information about products on social media will increase the probability of gaining followers which directly boosts brand awareness even more. And if the general public has a good brand awareness about the mentioned brand, then there would be less concerns about price to satisfaction regarding the product. Previous research conducted by Kundrniasari and Budiatmo proved that although J.Co has lower buying interest due to their price, their social media marketing strategies have helped them in sales.

Despite bringing plentiful benefits to sales, we have found that in order for social media marketing to succeed, it needs a lot of effort where previous research conducted by Latiff and Safiee [24] found that proper timing and posting some content that

people usually will like is not enough. That's why the business needs to have an admin that manages the account so that it will always be up to date. Furthermore, the account must be managed in such a way that customers are drawn in and their interest is piqued. To do this, the admin is required to keep up with the trend. Furthermore, this will add another cost to the operational budget which may or may not hinder the business overall yet in this day and age, social media marketing is crucial for business prosperity.

4.2 Discussion

The findings from this study suggest that frozen food sales have increased significantly because of a change in people's behavior toward the recent situation and social media marketing that most companies use recently as well. Social media can increase buy interest because social media play an important role increasing factors that help increase interest.

For the first research question from the previous research and sources, we conclude that the key factor of the sales rise of the frozen food product is the media social and the pandemic. The current situation where pandemic COVID-19 has been happening, affecting how people behave themselves these days where they like something convenient while still at home. This could happen because concern of getting infected creates a feeling that doing anything at home is much better which makes frozen food sales able to increase in recent years. Social media also helped market the product toward customers where this could happen because of some factors like socio-demographic factors (age, education, and income of a person). Furthermore, the way people use social media a lot these days created a much stronger impact toward buying behavior where it already did before.

And for the second research question, we found out that there is a chain reaction affected by the social media, like the one that happens to J.Co Donut and Coffee where they experience a phenomenon of sales rising due to the social media official account, where visitors feel that the account is worth following. This happened because the account contents are exciting and they can access many types of information about the product they sell so there is a chain reaction caused by the social media marketing to increase their sales and to minimize their weakness which is their price that is higher than the other's. The official social media account is considered worthy to be followed by the customer, then the customer get information from the routine post of the social media account, from there the customer will know more about the product start from the product detail until the promotion that available, then because of the information reachability and events, the customer becomes interested to buy the product of J.Co, and that is why the product sales can be raised although the product price is higher than the other's.

5 Conclusion and Suggestion

5.1 Conclusion

According to a qualitative review of the frozen food business, the COVID-19 epidemic has resulted in a significant surge in frozen food sales in recent years. Since it influences human behavior, the pandemic has a significant impact on the frozen food business. From the research study, it is concluded that the main factor that helps to increase the customers' interest in a particular product is brand awareness and to increase it, social media marketing plays a major role.

Based on this conclusion, many entrepreneurs, no matter if they are new or experienced in the frozen food industry, should consider social media as their platform to market their product while utilizing how people interact with the product. Many people will consider finding and buying the products online.

5.2 Suggestion

As mentioned before, the frozen food business is quite promising because of the market. In order to utilize this and stand out from other competitors, the seller needs to use social media as one of their marketing strategies. If the seller still relies on the traditional ways such as door-to-door, newspaper, magazines, radio, and television, they will not be able to survive among the numerous competitors. Hence why we suggest frozen food sellers use social media marketing.

In managing their social media account, there are some factors that need to be considered. The first one is the information. Frozen food sellers need to put up-to-date information regarding the product so that the customers can easily check out the products and decide whether to purchase it or not. They need to update the information every time they make a change, especially about the price to avoid the possibility of customers getting disappointed. The second one is the frequency of their update. To increase brand awareness and engagement of the social media account, frozen food sellers need to update their social media frequently. If possible, it is good to update daily but if not, at least it should be updated every three or four days to keep the customers interested or engaged. The third one is the trend. If frozen food sellers post daily yet the content is not in accordance with the current trend or not using the correct platform, then the social media marketing is not effective as they are not able to pique the interest of the general public. Frozen food sellers need to keep up with trends such as what kind of thing is popular right now. It could be movies, songs, series, memes, applications, celebrities, public figures, and others. Then it is highly required to incorporate those trends with their marketing strategies. For example, they could make a video using trending audio and concepts. By using current trends, the probability of their social media contents reaching the general public is higher.

Usually, in order to manage all of this, businesses often hire a social media admin. The role of admin is highly important. They must be able to devise social media marketing plans, stay current with trends, and even respond to client questions via social media accounts. The amount of work can be quite a hassle to manage if it is done alone by the frozen food seller. That is why we suggest having a social media admin to improve the efficiency and effectiveness of social media marketing.

References

1. Samoedra AD, Vinantama IR, Ramdhani AH, Subagja DA, Latif FM (2021) Strategy analysis of culinary business in the pandemic era using the timmons model approach (case study Rm. Andalas). Rev Int Geogr Educ 11(5):2052–2059
2. Haryandika DM, Santra IK (2021) The effect of customer relationship management on customer satisfaction and customer loyalty. Indonesian J Bus Entrepr (IJBE) 7(2):139–139. https://doi.org/10.17358/ijbe.7.2.139
3. Thearling K (1999) An introduction to data mining. Direct Mark Magaz 28–31. Retrieved from https://sci2s.ugr.es/sites/default/files/files/linksInterest/Tutorials/dmintro.pdf
4. Rohmah SN (2020) Adakah Peluang Bisnis di Tengah Kelesuan Perekonomian Akibat Pandemi Corona? Adalah 4(1)
5. Yetkin CG, Dondurucu ZB, Cetinkaya A (2020) Analysis of Youtube advertisements of GSM operators in the Covid-19 pandemia process in the context of social marketing. Int J Cult Soc Stud (IntJCSS) 6(Issue 1), e-ISSN: 2458–9381
6. Kotler P, Keller KL (2016) A framework for marketing management. Pearson, Boston, MA, p 352
7. Nurhayati H, Wolff (2021) Topic: social media in Indonesia [online]. Available at: https://www.statista.com/topics/8306/social-media-in-indonesia/#dossierKeyfigures. Accessed 30 Jan 2022
8. Qonita A (2021) Bisnis Makanan Beku, Menghangat di Era Corona. https://www.alinea.id/bisnis/bisnismakanan-beku-menghangat-di-era-coronab2cxM90B3. Accessed on 24 Jan 2022
9. Putra DIP, Syah TYR, Indradewa R, Pusaka S (2019) Duck nugget frozen food market analysis and marketing strategy as a startup business plan in Indonesia 3(4). ISSN/e-ISSN: 2541–0369/2613–988X
10. Carr CT, Hayes RA (2015) Social media: defining, developing, and divining. Atlant J Commun 23(1):46–65
11. Mangold WG, Faulds DJ (2009) Social media: the new hybrid element of the promotion mix. Bus Horiz 52(4):357–365
12. Li F, Larimo J, Leonidou LC (2021) Social media marketing strategy: definition, conceptualization, taxonomy, validation, and future agenda. J Acad Mark Sci 49(1):51–70
13. Sen S, Antara N, Sen S (2021) Factors influencing consumers' to take ready-made frozen food. Curr Psychol 40(6):2634–2643
14. Ahuja HK (2011) A comparative study of attitude and adoption of frozen food amongst the consumers of Ludhiana and Delhi cities. Punjab Agricultural University, Ludhiana, 141, 004 (India)
15. Twin A (2022) Sale definition. *Investopedia.* Available at: https://www.investopedia.com/terms/s/sale.asp#:~:text=A%20sale%20is%20a%20transaction,are%20paid%20to%20a%20seller. [Accessed February 10, 2022].
16. Singh A, Verma P (2017) Factors influencing Indian consumers' actual buying behavior towards organic food products. J Clean Prod 167:473–483
17. Chambers S, Lobb A, Butler LT, Traill WB (2008) The influence of age and gender on food choice: a focus group exploration. Int J Consum Stud 32(4):356–365

18. Dettmann RL, Dimitri C (2009) Who's buying organic vegetables? Demographic characteristics of US consumers. J Food Prod Mark 16(1):79–91
19. Udimal TB, Peng Z, Luo M, Liu Y (2022) The impact of COVID-19 on consumers' eating and purchasing habits of agricultural products in China: key determinants and policy implications. Bull Nat Res Centre 46(1):1–8
20. Anggraeni DA, Widjanarko SB, Ningtyas DW (2014) Proporsi tepung porang (amorphophallus muelleri blume) tepung maizena terhadap karakteristik sosis ayam. Jurnal Pangan dan Agroindustri 2(3):214–223
21. Fachriansyah R (2020) Frozen meals become a hot seller for quarantined Indonesians [online]. The Jakarta Post. Available at: https://www.thejakartapost.com/news/2020/05/22/frozen-meals-become-hot-seller-for-quarantined-indonesians.html. Accessed on 28 Jan 2022
22. Kurniasari M, Budiatmo A (2018) Pengaruh social media marketing, brand awareness terhadap keputusan pembelian dengan minat beli sebagai variabel intervening pada J.Co Donuts & Coffee. Retrieved from: https://ejournal3.undip.ac.id/index.php/jiab/article/view/20968
23. Creswell JW (2014) A concise introduction to mixed methods research. SAGE publications
24. Latiff ZA, Safiee NAS (2015) New business set up for branding strategies on social media–Instagram. Procedia Comput Sci 72:13–23

Geographical Information System-Aided Landmark Recognition System Using Machine Learning

S. A. Sahaaya Arul Mary, Lakshmi Kanthan Narayanan, S. Mohana, R. Senthamil Selvi, R. Karthik, and N. Ramya

Abstract The recent advancements in computer vision-based innovations like image identification and object detection are attracting the attention of the modern research community present across the globe. Data analysis domain is interesting, particularly when it comes to image augmentation and categorization. Image recognition will operate on a variety of variables in different contexts relating to human behaviour, such as what sort of behaviour someone would display when driving a car. Activation such as network expanders is used in numerous significant neural network models, which are used in image processing domain. Learning the intricacies of features is gradually becoming more prevalent due to the fact that nearly all mobile devices are connected to the Internet and to the fact that virtually all portable devices are connected to the Internet. To achieve this, the visitor obtains a digital camera and uses it to capture the marker and upload the images to a nearby server where they may be compared to the photo archive. Some useful information can be provided during an image search. For example, learning more about significant landmarks, local trails, or nearby areas of interest, or interesting features are also given on the match to be searched. The broader and deeper form of application will be better for guiding people around the region. This paper proposes an enhanced convolution

S. A. Sahaaya Arul Mary · L. K. Narayanan (✉) · S. Mohana · R. Senthamil Selvi · R. Karthik · N. Ramya
Saranathan College of Engineering, Tiruchirappalli 620012, India
e-mail: lakshmikanth.vlsi@gmail.com

S. A. Sahaaya Arul Mary
e-mail: mary-cse@saranathan.ac.in

S. Mohana
e-mail: mohana-cse@saranathan.ac.in

R. Senthamil Selvi
e-mail: senthamilselvi-cse@saranathan.ac.in

R. Karthik
e-mail: karthik-cse@saranathan.ac.in

N. Ramya
e-mail: ramya-cse@saranathan.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_20

neural and deep convolution neural network-based method for identifying different landmarks by using machine learning [ML] algorithms. The system uses spatial extreme learning machine (SELM) for performing the optimistic feature recognition and enhancement.

Keywords Computer vision · Landmark recognition · Convolution neural network · Deep convolution neural network · Machine learning · Spatial extreme learning machine

1 Introduction

The primary objective of this work is to categorize the scenes from the pictures. Rather than prioritizing, group prioritization is done based on the objectives and the structure of the item present inside a group. Expansion employs both external environment and object composition prioritization. Prior to the application of CNNs, research codebooks and word-bag models (other approaches that used code) are used for segmentation and gaining prior knowledge. In that context, the knowledge of the previous research will be provided for this scenario, where the image is divided into a small number of known categories by using unsupervised learning to guide neural networks and assign each to those categories.

1.1 Purpose

Automatic recognition of history from photographs can proceed for obtaining better localization in the urban area. Mobile line recognition is expanding as more individuals carry mobile devices linked to the Internet. The fundamental concept here is to use the mobile phone to capture images of the location, where the user is standing. After that, the pictures are uploaded to an Internet server, which is synchronized with a global landmark database [1, 2]. When a match is found, background information about the placemark and other relevant information will be returned.

1.2 Scope

This style of application is particularly beneficial when the user is travelling as a tour guide to cities. Animated footage will be used to mask some of this material to save taxpayers money and decrease their fear of using footage recorded on tape. To give another example, the user focuses his or her camera on nearby and relevant structures in order to obtain more information, such as name labels, which appear as an annotation in the scene. This may be quite useful when travelling in a city

as well as in a touristy area, especially if you are unfamiliar with the area. Many smartphone applications now contain GPS and magnetometer functionality, allowing them to assess whether a specific landmark is present in the volume. Though output may be higher in open areas, dense areas have tall buildings and it is possible that GPS reception may be inhibited in urban areas due to nearby tall structures. In an environment like downtown Boston, for example, viewing the buildings helps you better connect to the sidewalk, where people walk and more comfortably offer directions while using the smartphone as third eye.

1.3 Motivation

The use of mobile devices has increased dramatically in recent years. As most handheld devices have a camera and are linked to the Internet, the theory is straightforward. The tourist uses the compact camera to capture pictures of the position marker and move them to a server where photographs complement their photographs archive. If a match is made, valuable information is transmitted, such as details behind a landmark, adjacent paths or details about other important landmarks nearby. As leading city guide or navigation aid, this form of application has tremendous potential.

1.4 Need for Landmark Recognition System

Landmark recognition is a critical function of mobile robots. Landmarks are used to measure population and to create maps of unknown places. In this context, traditional visual techniques based on solid geometric models cannot be used. Instead, models of land symbols should be built from the point of view using image-based techniques. We are trying to address the issue of creating historical meanings based on images from the sequence of images and visions. In addition to its work on mobile robots, this method addresses the common problem of identifying groups of images with common characteristics in the sequence of images. We show that, with appropriate background issues and image captions, this can be done using effective techniques. Identifying landmarks in a sequence of images is a challenging problem for many reasons.

1.5 Problem Statement

We're looking at how CNN can be equipped to differentiate scenes with Top-1 and Top-5 precision. The Places365-Standard Group Database is used and has 365 groups in the training collection, 1.6 million named images in the verification set and 900 images in the evaluation set. Furthermore, we hope to learn that such neural net

variables contribute to improved segmentation efficiency through analysing CNN training for this reason [3]. Dividing scenes (such as mountains and forests) is not an easy task because of their variety, ambiguity, and wide range of light and functional conditions. Two basic techniques can be found in books. The first uses low-level features such as colour, texture and energy spectrum. These methods consider space as a single entity and are commonly used to separate only a few square segments (indoor outdoors, city compared to country, etc.). The second strategy uses moderate presentations before segmentation and is used in cases where there are large numbers of segment segments (up to 13).

The organization of this paper is arranged as follows. Section 2 discusses in detail the various works done on recognizing land patter and geographical information system. Section 3 is furnished with details of how the GIS-aided LR system is designed. In Sect. 4, the results inferred and the conclusion findings are discussed elaborately. In Sect. 5, the future scope of the system is discussed.

2 Literature Review

2.1 The Analysis of Object-Based Change Detection in Mining Area: A Case Study with Pingshuo Coal Mine

There are some smaller scale materials collection projects, such as diversifying of found objects, and the rapid expansion of informal settlement patterns with newly acquired property, that have shown how well object-based acquisitions can be applied to urban sprawl, but there hasn't been a lot of research done on how the general rate of material accumulation in these cases. We discovered in this analysis that object-based distinction and remote image transformation acquisition can be used not only to detect urban transition, but also to track land restitution in the mining industry, and that it increases tracking precision beyond a doubt.

2.2 Scene Analysis by Integrating Primitive Segmentation and Associative Memory

The understanding that a computer has of a film's visuals may expand, but it is not getting any easier. In this paper, the first phase of separation is added to the assembly model to show how a complex database query task like grouping can be reduced to a single step [4, 5], allowing us to store data at various points during the stages of multi-modular integration until the last stage, where it can be passed back into the first memory (memory STM) [6]. When it is first initialized, the global oscillator network divides the input region into multiple bits, with in multiples of corresponding to the predetermined oscillations. Each component establishes patterns in the STM memory

for use [7]. Each component uses these patterns to retrieve additional information and convey information to the other layers. More than 128-layer STM projects in the LEGION network yield an infinitely expanded memory, or an expanded memory in the process, it will do both low-level and high-level analyses; in other words, it will do a full analysis for the population as a based study and for the team. The three-dimensional model is evaluated with a row of three paintings (instead of traditional "set scenes" illustrations), depicting a writing process that incorporates an obstacle, such as adding an object to an performance has improved due to the increased reliance on the organization's memory structure. Many topics are covered, such as how the input attachment of the input is performed, and the most sensitive memories are produced, and where it originates, as well as STM's function in generating sensitive memory.

2.3 Active Learning with Convolutional Neural Networks for Hyperspectral Image Classification Using a New Bayesian Approach

A study of the context and the possible ways things that situations influence perceptions are just as important as a study of machine comprehension [8]. The author combines the first step of the group analysis (separation) with a meeting memory models to give us an integrated treatment for groups of separability [9]. There are two phases in our model: a partitioning scheme that consists of primary memory modules and multi-modular integration, and a temporary memory (memory STM). The first stage of division is performed by a LEGION, which is a global in this case, in that it divides the input region into many sub-areas (i.e. to many corresponding to the different output groups) for further expansion. Each part of the design influences and shares all layers of memory: one component generates many memory patterns, which are found in all of an STM design strengthening the STM (electrostatic strain relief) layer projects into the LEGION network, causing memory loss and cracking. A comprehensive analysis is performed of the entire phase space, which yields deep, detailed information about a particular space as well as overall information.

This design model is put to the test with a series of 3D images that use the additional memory capacity to introduce more elements and increase the number of scenes that can be inserted. Performance has improved due to the increased reliance on the organization's memory structure. A lot of issues are dealt with, such as input attachment, top detachment and memory formation, and how it relates to forming attitudes.

2.4 *Learning Deep Features for Scene Recognition Using Places Database*

This identification is one of machine viewing some recent successes in object recognition techniques are based on the use of vast image databases such as ImageNet and CNNs to tackle high-level features, but not at the same level of accuracy [10]. The current models cannot have enough functionality for those kinds of tasks due to their usage in ImageNet being too extensive [11]. We've developed a new domain database called Places, with over 7 million photographs, assigned to each name, which has been catalogued to each and includes information about the images that follow. The way we propose will enable you to examine density and variability of image databases and demonstrate that locations appear to be as dense as well as equivalent to comparison locations. New to CNN, we expand our knowledge on the field of domain scene recognition and create a number of new technical domains on the top of the barrel scene answer views help us to distinguish between the various types of networks that appear on the inside of the CNN.

2.5 *Scene Categorization with Spectral Features*

This paper introduces a novel way of explaining the application of the concepts of convolution and spectral synthesis, which are already very effective in deep neural networks. As far as we know, this is the first-time researchers have tried to incorporate the most profound imaging characteristics of visual learning into a work in a classification study [12, 13]. By extension, we can also show that with spectral shift, the dependence acting, we reduce the dependence of variables and that results in a functional independence, which represents functional normalization as well as a substantial change to the larger the databases we use, the better the results we obtain (MIT-67, SUN-397 and Places-205). We also have Fig. 1. The proposed

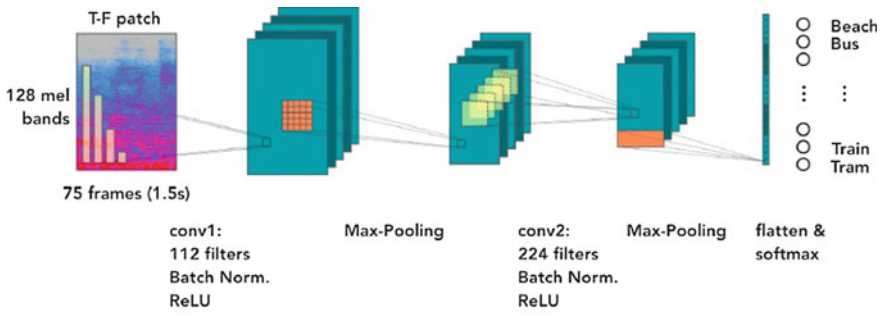


Fig. 1 Architecture of CNN

approach of qualitative work was evaluated for its outstanding success, which reveals its resemblance to natural scenarios.

3 Design Methods

Convolutional neural network: When CNN is introduced over a periodic audio, they can capture spectro-temporal patterns, which can be useful in visual functions [14]. In addition, the size (width and height) of convolutional filters can be associated with time intervals and quantities, respectively. In this work, we explore how these relationships can be applied to the design of ASC convolutional filters.

3.1 System Design

3.1.1 Existing System

In the current system, the proposed MCNN trains CNN on a wide range of scales [15]. Combining image data from F and image has the same value as weight on F and VNET and expanding them simultaneously downloads it for CNN to look at various size measurements.

3.1.2 Proposed System

In the proposed system, we are planning to implement a deep neural network-based scene classification in which various objects present in the scene are accurately detected and the OCR-based text detection is also implemented. As an application, we can also determine various database containing useful information about the scene and this can be pre-processed further for surveillance purposes. The system will be trained in a way not only with the scenery but along with that the surrounding objects are being classified and trained. Whenever the need for recognition happens, it is effective to determine the exact location based on the landmarks identified. The proposed model is enhanced with an object identification system which makes the recognition system more accurate and to carry out effective operations. The OCR-based text detection is carried out to enhance the textual contents present in the scenery. The accuracy of the recognition made will be verified using the geographical information system data availed from the open-source Geo-Info system. The OCR-based recognition is used to determine the text present in and around the environment.

3.1.3 Block Descriptions

Figure 2 depicts the functional steps involved in proposed process. Pre-processing the input scenes needs converting the input factor into accountable factor. Next block feature extraction is used to extract the unique feature points present in the scene images. It extracts the texture feature, colour feature, shape feature, etc., to identify the scene based on features. Feature vectors are mapped uniquely and target mapping. Input mapping is being configured. Deep neural network, transfer function, iteration loops, and neurons are configured.

3.2 Process

Deep neural platform performs two operations such as training and testing. Training the database inputs for certain iterations and performing feature extraction process, the feature plots are allotted to targets vector. Performance plots are used to calculate the amount of time it takes to complete a process, and they may be used to decide which sample has the finest epics. ROC curves are considered as the receiver operating characteristic curves. The number of positive points and genuine negatives from the procedure was defined by using the characteristics plot. The feature plots are assigned to the targets vector after testing the input for a specified number of iterations and conducting the feature extraction procedure. Further, a matrix table will be created from the scenes (Fig. 3).

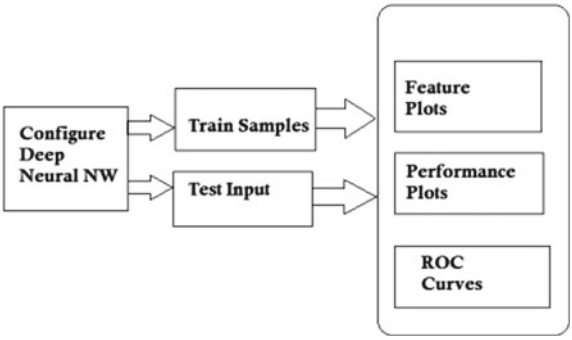


Fig. 2 Pre-processing of data by proposed architecture

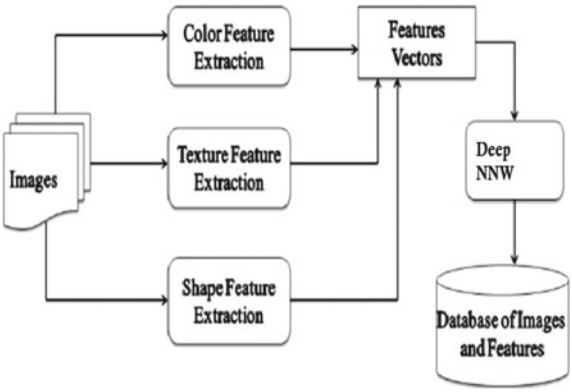


Fig. 3 Functional process flow

3.3 Modules

3.3.1 Nonlinear Function Layer

The nonlinear function selected for use here is being modified linear units (ReLU). The nonlinearity of the application is used to determine the features which are scattered around the focal field.

3.3.2 Pooling Layout

Blending size is given to 5, and integration circuits are the same which are not in harmony with each other.

3.3.3 Fully Connected Layout

As shown in previous experiments, the hidden layer at 50 units is best for auto-encoders to represent an image. Therefore, 50 units are used in a completely hidden layer.

3.3.4 Loss Layer

Softmax separator is followed by a logical loss function using image separation and network training.

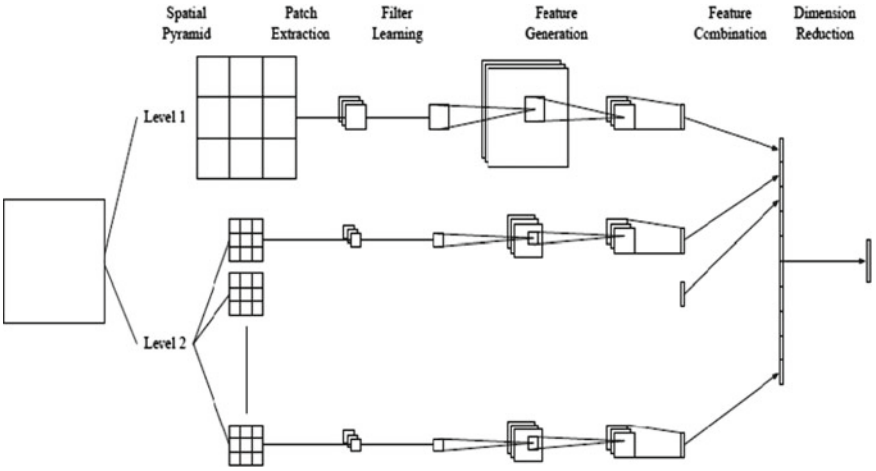


Fig. 4 Proposed landmark recognition system

3.4 Architecture

The architecture explains the feature extraction steps for generating unique feature to be detected from each and every input image. To design an efficient machine learning-enabled scene classification [16]. This study helps us to learn more info about machine learning algorithms. This study provides deep knowledge on linear regression [LR] model (Fig. 4).

4 Results and Discussion

In this paper, we have proposed VGG16 + CNN algorithm for identifying landmarks in a sequence of images. Performance of the algorithm is evaluated with the real-time landmark dataset. The dataset is availed from Kaggle, which contains landmark-related data [17]. The result of the proposed VGG16 + CNN method outperforms the other methods and proves to be better in terms of accuracy (Figs. 5, 6 and Table 1).

5 Conclusion and Future Scope

In order to conclude, among all the ML-based algorithms the combination of VGG and CNN performance is found to be far better than the other algorithms. In the future, this system can be further enriched in terms of precision accuracy by deploying DNN-

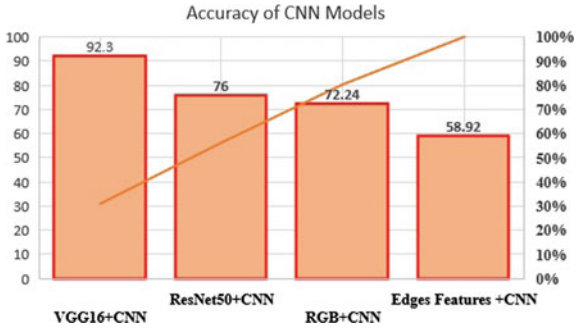
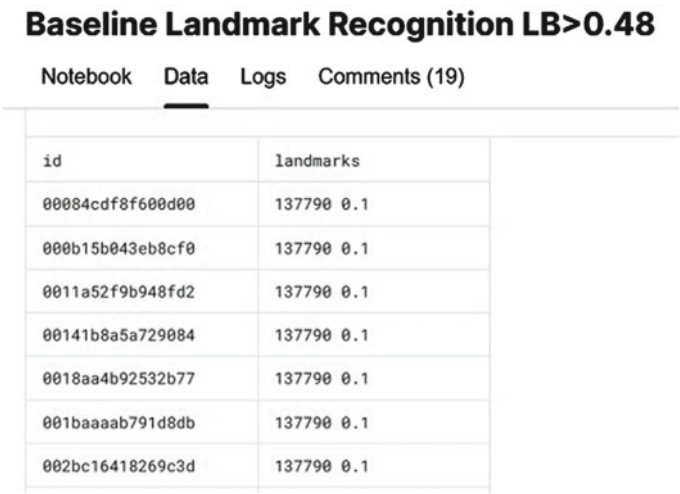


Fig. 5 Accuracy of various CNN models



Baseline Landmark Recognition LB>0.48

Notebook Data Logs Comments (19)

id	landmarks
00084cdf8f600d00	137790 0.1
000b15b043eb8cf0	137790 0.1
0011a52f9b948fd2	137790 0.1
00141b8a5a729084	137790 0.1
0018aa4b92532b77	137790 0.1
001baaaab791d8db	137790 0.1
002bc16418269c3d	137790 0.1

Fig. 6 Line landmark recognition LB > 0.48

Table 1 Method versus accuracy in percentage

Method	Accuracy in percentage
RGB + CNN	72.24
Edges features + CNN	58.92
ResNet50 + CNN	76
VGG16 + CNN	92.3

and RNN-based algorithms to track the real-time coordinates, which will be a handful in cruise and vehicles.

References

1. Jha R, Jain P, Tayal S, Sharma A (2021) Landmark recognition using VGG16 training. Smart Sustain Intell Syst
2. Manoharan JS (2021) Capsule network algorithm for performance optimization of text classification. J Soft Comput Paradigm (JSCP) 3(01):1–9
3. Pasumpon PA (2021) Performance evaluation and comparison using deep learning techniques in sentiment analysis. J Soft Comput Paradigm 3(2):123–134
4. Wang X, Shen S, Ning C, Huang F, Gao H (2016) Multi-class remote sensing object recognition based on discriminative sparse representation. Appl Opt 55(6):1381–1394
5. Shakya S (2021) Unmanned aerial vehicle with thermal imaging for automating water status in vineyard. J Electr Eng Autom 2
6. Wang J, Jin Y, Cai S, Xu H, Heng PA, Qin J, Wang L (2022) Real-time landmark detection for precise endoscopic submucosal dissection via shape-aware relation network. Med Image Analy
7. Banerjee B, Chaudhuri S (2018) Scene recognition from optical remote sensing images using mid-level deep feature mining. IEEE Geosci Remote Sens Lett 15(7):1080–1084
8. Liu Y, Huang C (2018) Scene classification via triplet networks. IEEE J Sel Topics Appl Earth Observ Remote Sens 11(1):220–237
9. Yang R (2021) UAV landmark detection on fast region-based CNN. Arabian J Geosci
10. Cheng G, Xie X, Han J, Guo L, Xia GS (2020) Remote sensing image scene classification meets deep learning: challenges, methods, benchmarks, and opportunities. IEEE J Sel Topics Appl Earth Observ Remote Sens 13:3735–3756
11. Zhang L, Mueller R (2022) Large-scale recognition of natural landmarks with deep learning based on biomimetic sonar echoes. Bioinspiration Biomimetics
12. Wang X, Xiong X, Ning C, Shi A (2018) Integration of heterogeneous features for remote sensing scene classification. J Appl Remote Sens 12(1), (Art.no. 015023)
13. Dhaya R (2021) Hybrid machine learning approach to detect the changes in SAR images for salvation of spectral constriction problem. J Innov Image Process (JIIP) 3(02)
14. alias Muralidharan RR, Subbaiah P, Narayanan LK (2021) Smart infant and child monitoring system using integrated IoT and GIS. Design Eng
15. Zheng X, Yuan Y, Lu X (2019) A deep scene representation for aerial scene classification. IEEE Trans Geosci Remote Sens 57(7):4799–4809
16. Sun H, Li S, Zheng X, Lu X (2020) Remote sensing scene classification by gated bidirectional network. IEEE Trans Geosci Remote Sens 58(1):82–96
17. https://www.kaggle.com/paulorzp/baseline-landmark-recognition-lb-0-48/data

Housing Society Management System Using PHP

V. Harsh, T. Shubham, D. Ritesh, and R. Mansing

Abstract The Housing Society Management System aims at managing the day-to-day activities of a cooperative society and making the current situation in society simple and efficient. The proposed research focuses on creating online apps for handling complaints in gated and guarded communities. These community areas have fencing and are supervised by a guard. They are managed by a company that is paid to look after the assets, amenities, and security in residential communities. Text messages and WhatsApp are now the only complaint handling methods available. As a result, it is difficult for management to review the complaints, and the entire process takes longer. In addition, residents are unaware of the actual status of their complaints. The suggested system entails the application of web-based applications to manage this. The data sent is stored in the developing application's integrated database. This application is put to the test on-site to see how effective it is. The findings of the poll revealed that the application is successful in dealing with resident problems.

Keywords Society · Management · Admin · Member · Monitoring · PHP

1 Introduction

In general, all labor in society is done by hand, including keeping track of members' phone numbers on the sheets. All of society's paperwork and expenses are maintained on paper, making it difficult to keep track of them. All of the society's day-to-day

V. Harsh (✉) · T. Shubham · D. Ritesh · R. Mansing
K.J. Somaiya Institute of Engineering and Information Technology, Mumbai, India
e-mail: harsh.vador@somaiya.edu

T. Shubham
e-mail: svt@somaiya.edu

D. Ritesh
e-mail: ritesh.doshi@somaiya.edu

R. Mansing
e-mail: rathodm@somaiya.edu

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_21

operations will be automated with this housing society management system. This system also keeps track of information for services like security, housekeeping, and equipment maintenance. Since there is no automated framework for dealing with this multitude of administrations, this framework makes it conceivable by automating every one of them written by hand logs and monitoring them. It will provide reduced response time against the queries made by different users. This project is based on PHP language with a MYSQL database which manages society with flat and owner of flat detail. This societal management is carried out in such a way that the most typical issues that arise in residential societies are addressed. Partial or advance payments can be readily reviewed and validated with the help of this method. They necessitate coordination between the respective management societies and the suppliers who provide these services in order to give the proper convenience. The major purpose of this project is to make it simple to track various society expenditures and to keep track of logs. Bills, receipts, and vouchers are carefully made in this system, which is also user-friendly. It also includes an admin module that allows the administrator to keep track of various charges. Admin can access spending information, including bills from service suppliers.

All possible features such as verification, validation, security, and user-friendliness have been considered.

2 Literature Survey

The cloud is used to develop our housing society management system. The system is designed to create monthly maintenance charges, as well as the option to pay payments online. On their account, members can check the progress of their bills. The accounts are set up on the local server using SMOnline. SocieTales, or Tales in Society, is an Android application that aims to introduce a sensible form of communication in society. This web application allows for automatic man-to-man conversation. Another is an Android application that uses sensors to improve the management of housing societies and guest parking. The server is constructed as a Web-based application, and the system is tested with various real Android phones equipped with GPS and Wi-Fi clients. Paper [1] focused on a study that involved using an android application to facilitate tasks for the board of residential society capacity. The hectic pace of city living necessitates a high level of devotion to these services. Private social orders are responsible for a variety of services, including electricity, water supply, security, and a variety of other things. In most cases, traditional ways of correspondence are used in various social orders when it comes to these exercises. Because of the lack of transparency, this may encourage the unethical dissemination of data. Computerization of these jobs is necessary to avoid such conflicts and to administrate and stay up with society's constraints on a daily basis. Paper [2] believed that society's structure should be constructed in such a way that the most pressing issue addressed by private social orders is resolved. This framework is a cloud-based solution for running any welcoming housing organization's day-to-day activities. To

achieve optimum comfort, they require coordination between the social orders of individual organizations and merchants. The principal motivation behind this business is to give web charging and accounting (introduce portion consolidation, pay and cost after, etc.). Bills, receipts, and vouchers are made in a smoothed out way under this framework. In addition, the strategy acts as a client buddy. The Android Application Development College Challenge, as stated in the paper, encouraged understudies to structure and engage in the execution of their Android applications [3]. As a result, it gives understudies a chance to show off their inventiveness and investigate the advancement of Android applications. As more understudies from more institutions and districts participate in the challenge, the influence it has continues to grow. The association of such a challenge is given out in this paper for uninstalling stall installation. This is a study article on a multipurpose Android-based mobile notifier, Google Cloud informing, and an Android application recognized with an online warning framework. All work is done manually in the present framework. The notification board is customary. Thus, there is an issue of data not contacting an adequate number of individuals just as it brings about a maximum usage of time. Additionally, the evaluating is finished utilizing documents that are not straightforward and are inclined to extortion. Stopping installment and portion is tedious and specifically keeping a record is tedious. To shake off the old strategies for keeping up with the records of bills, saves, meeting subtleties, and different administrations, digitalization of tedious records is significant. Informal organizations assume a significant part in interfacing individuals through Internet visiting and partaking in web-based missions and clubs. As a result, they have developed the groups' system, as every housing society has multiple groups for diverse reasons such as social service, company publicity, and hobby and sports preservation. Many people enjoy being a part of many organizations, yet they sometimes choose to remain uninvolved or are hesitant to apply for membership. As a result, the recommendation system assists in suggesting a suitable group to them based on their hobbies and sports preferences, making it easier for residents to choose a group from the list. When registering a resident in the system database, their interests and sports preferences are taken into account. The online application employs a recommendation algorithm that suggests groups to a specific person based on their hobbies and interests (assuming the group is present in the society). As a result, the resident can join and quit the group as needed. To differentiate and make it clear to the resident, the system displays the recommendation for sports-based groups and hobbies/interests-based groups separately.

3 Proposed System

In this society management system, all the society categorize by the number of blocks and flats. The people who live in flats may be an owner of flat or tenant of flat. In this era, people are very hectic schedule, so they do not have time for complain small problem related to flat. We have developed a system for society members where they can make complaints from anywhere at any time, and we resolve the complaint as

soon as possible. In this system, people can easily find address of the flat by providing member name. The different types of modules present in this project are

1. Admin
2. User

The admin interface has a dashboard where the administrator may see a summary of all information, such as total flats, total bills, total allotment, total visitor, unresolved complaints, in-progress complaints, settled complaints, and total complaints. The second input is a flat in which the administrator can control flats (Add/Update). Allotment is a parameter that allows the administrator to manage allotments (Add/Update), and bills are a parameter that allows the administrator to manage bills (Add/Update). View complaint is a feature where the administrator can view and reply to complaints made by society members, as well as change the status of complaints based on their severity. Visitors allow the administrator to handle visitors (Add/Update) as well as their incoming and outgoing details. Admins can search for flats and visitors to the society in the search section. The society member user panel comprises a dashboard that displays a welcome page for society members, followed by view bill. The user can view his or her own monthly society charges in this section, which are generated by admin. Users can check the details of visitors who came to their own flat in the visitor area. Users can file a complaint and view the status of their complaint in the complaint area, which is supplied by the society admin. Users can look for visitors who have visited their flat in the search section. Users may see how many people came to an apartment within a specific time period in the report. The primary element of our task is that it has an Internet-based bill installment framework, a web-based democratic framework for choosing overseeing panels like Secretary, Chairman, and Treasurer. Administrators can add proprietors independently utilizing the administrator board. Paper [4] Proprietors can decide in favor of the up-and-comers that are representing various jobs in the public arena. When a login qualification for a part account is made, an auto-produced email is shipped off the part whose record is made. This framework permits individuals to sign in to their records and get refreshed with get-togethers. The proprietor can likewise really take a look at the exchange history, made against each bill.

The system's data must be stored and retrieved from a database. The database design is an aspect of the overall system design. At the analysis stage, data items and data structures to be stored were discovered. They are organized and combined to create a data storage and retrieval system. A database is a collection of interconnected data that is stored with minimal redundancy in order to serve a large number of users quickly and efficiently. The overall goal is to make database access simple, rapid, affordable, and flexible for users. Relationships between data items are established, and irrelevant data pieces are deleted. Normalization is used to ensure that data is internally consistent, with little redundancy and maximum stability. This guarantees that data storage requirements are kept to a minimum, data inconsistencies are avoided, and updates are optimized [5]. For the development of the relevant databases, the MS Access database was chosen (Figs. 1 and 2).

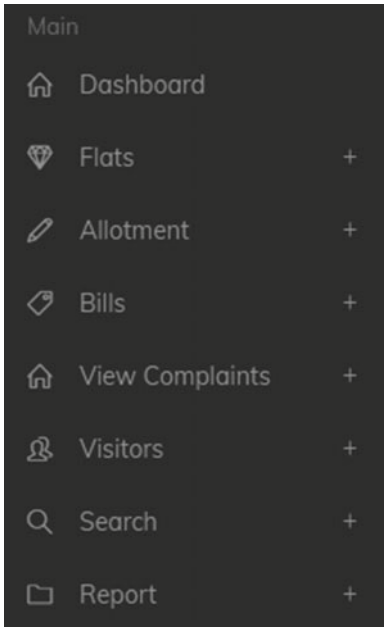


Fig. 1 Functional components

The administrator can create reports of all exchanges whenever. Utilizing the administrator board of this framework, administrators can likewise follow different costs like water supply installment and common work gauges. Saving cultural conveniences like gathering corridors for transport meets that show lobbies for festivities in the public arena were a troublesome undertaking as subtleties were on paper and were available just to the overseeing council which might cause disarray when more than one partner needs to celebrate in a similar show lobby. This framework likewise saves ascribes, for example, cost portrayal, cost date, installment mode, to whom it was paid and so forth clients can recover these logs whenever from the report module. It creates a bill for all proprietors at a solitary snap. Administrators can likewise follow cost explanations for aggregate months and can deliver reports of outgoings. By utilizing the framework, the financial officer can give cost articulations to the pinboard and proprietors at the snap of a button. Paper [6] In this framework, bills can be created aggregately, for example, month to month or yearly, the time period of bills can be allotted. Utilizing this framework individuals can satisfy their obligations on the Internet utilizing a MasterCard, a charge card or a net banking account (Fig. 3).

This product centers essentially around delivering bills with different functionalities of society which are definite duplicates of the genuine happenings in the general public. On effective installment of duty, individuals get moment e-receipts which can be saved or printed for future reference. Individuals can likewise see/download receipts for all records online whenever. On forthcoming occasions, yearly broad executive gatherings can be presented to all proprietors of the general public.

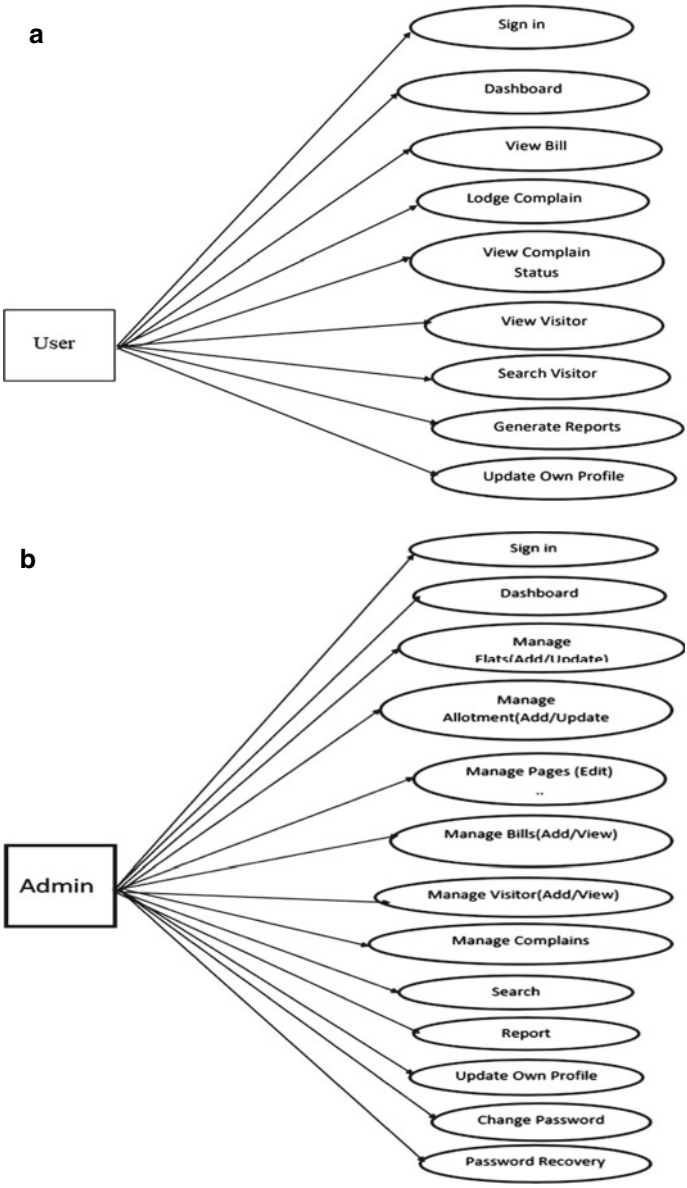


Fig. 2 **a** Use-case diagram: user, **b** use-case diagram: admin

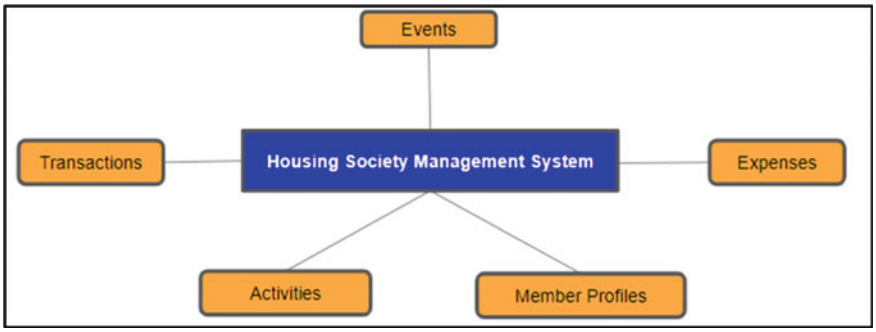


Fig. 3 Functional applications of the system

Forthcoming updates will be apparent on the Internet-based bulletin board till the termination date of the notification (Tables 1 and 2).

The Managing Committee can send brief messages by email to all individuals. The individuals with admittance to the admin board of this framework can without much of a stretch issue notice, procedures, synopsis reports, and imprint the presenters on the web.

Table 1 Admin database

ID	Admin name	User name	Mobile No	Email	Password	Admin reg date
1	Admin	Admin	1,234,567,890	adminuser@gmail.com	202cb962ac59075b964b07152d234b70	2019–12-09 12:38:42

Table 2 Allotment database

ID	Block	FlatNum	PricepUnit	UnitCons	Echarge	BillDate
5	A	102	10	100	1000	2019-12-12 22:38:07
6	A	102	12	50	600	2019-12-12 22:43:15
7	A	101	100	20	2000	2019-12-19 22:44:33
8	A	101	8	400	3200	2019-12-20 09:59:37
9	B	201	8	200	1600	2019-12-24 10:32:32
10	A	107	8	200	1600	2019-12-24 12:21:02
11	A	103	4	700	2800	2019-12-29 22:00:02
12	B	203	12	120	1440	2019-12-29 22:06:38
13	A	201	5	1200	6000	2019-12-29 22:49:33
14	A	101	10	200	2000	2022-03-03 19:17:00
15	B	201	15	200	3000	2022-03-03 19:21:09
16	B	201	0	0		2022-03-04 11:55:10

4 Implementation

The Housing Society Management System was developed using PHP, HTML, MYSQLi, and JavaScript. We used Bootstrap as a front-end framework for the layout. The purpose of the system is to have accurate information about the member transaction, lessen the work time of every individual, and avoid data redundancy. Using this system, the members can easily track their payments for every transaction they have made. Until the deadline arrives, they can also follow the balances that they have to comply with (Fig. 4).

The landing page of the web application is the home page which consists of the two modules of the application that are admin login and user login (Fig. 5).

Fig. 4 Home page

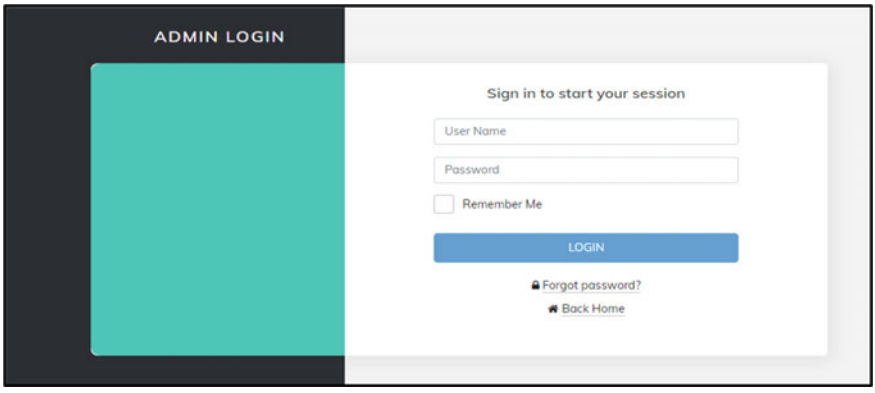


Fig. 5 Admin login

The administrator's landing page from which the system's registered administrators can log in is shown in Fig. 4. This Housing Society Management System features a user-friendly interface and functions (Fig. 6).

Members can login using their phone number, adding the block and the flat number that they live in. These members are created by the administrator (Fig. 7).

The member's Dashboard offers numerous features that allow members to check the status of their bills, visitors, and the complaints that they have registered. The administrator is responsible for administering and maintaining the system's data.

We give an efficient and straightforward solution to address crucial issues by lowering efforts and focusing on dependable communication to overcome the flaws

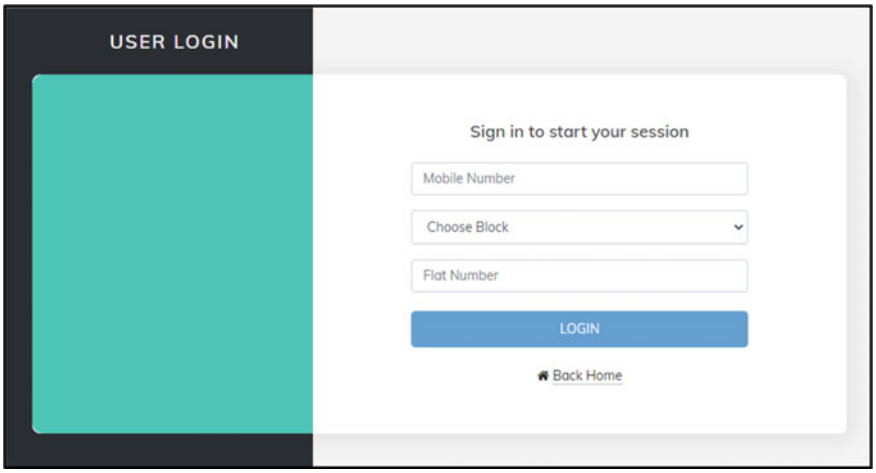


Fig. 6 Member login

Fig. 7 Dashboard member

of the current system. The prototype application is built with the Xampp web server and the phpMyAdmin database administration tool (Fig. 8).

The dashboard of the admin section consists of various complaints of the society registered by the society members including ones which are in progress and resolved as well. The administrator can also add users by validating their credentials. The user can log into the interface after successfully registering and perusing the modules as needed, as well as receive information about social events (Fig. 9).

This section includes details of the flats of the society that can be easily accessed by the administrator. This includes flat details such as flat number, floor, flat type, and also maintenance charge on that particular flat.

As shown in Fig. 10, a detailed track of expenses is visible on the panel. It is

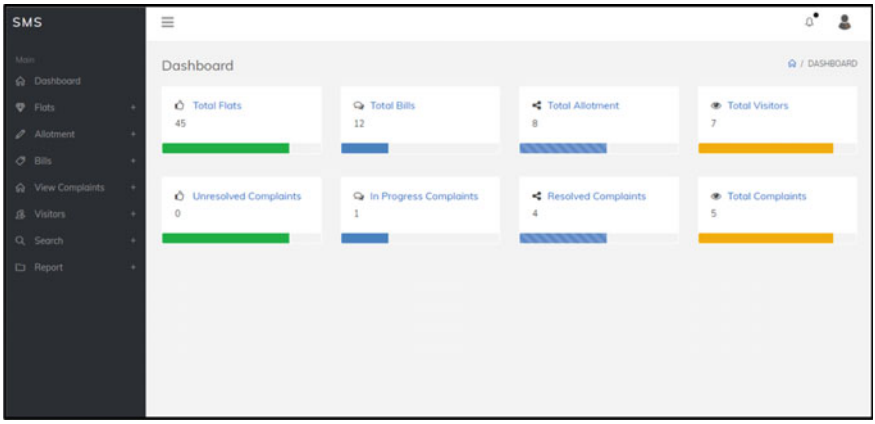


Fig. 8 Dashboard admin

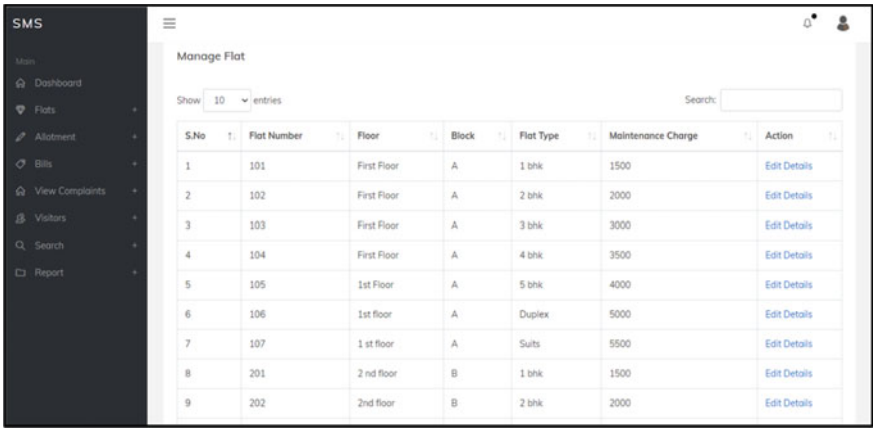


Fig. 9 Manage flats

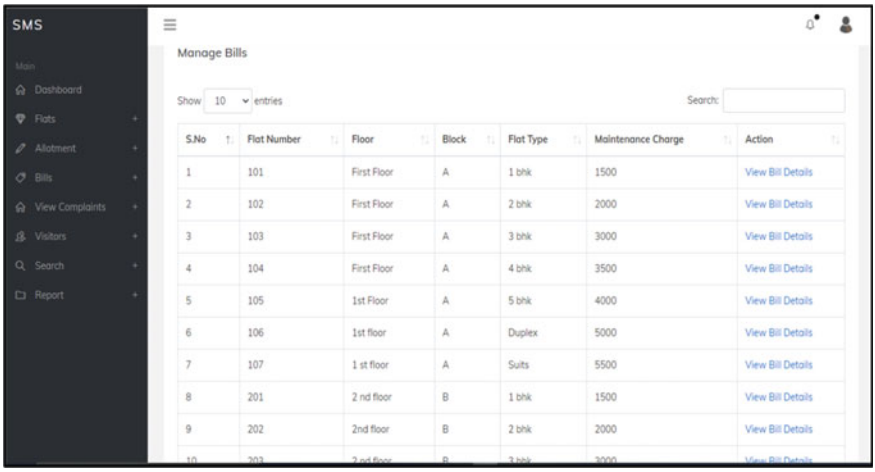


Fig. 10 Track of bills

accessible only to the administrator after the admin login panel. This includes keeping track of electricity bills as well as maintenance charges (Fig. 11).

Complaint details basically emphasize on the issues being triggered and are updated as per the actions performed with timestamp and status (Fig. 12).

The details of all visitors are available in the admin panel. It shows a detailed account of all the people who have visited including their time of entry, exit, and the purpose of their visit.

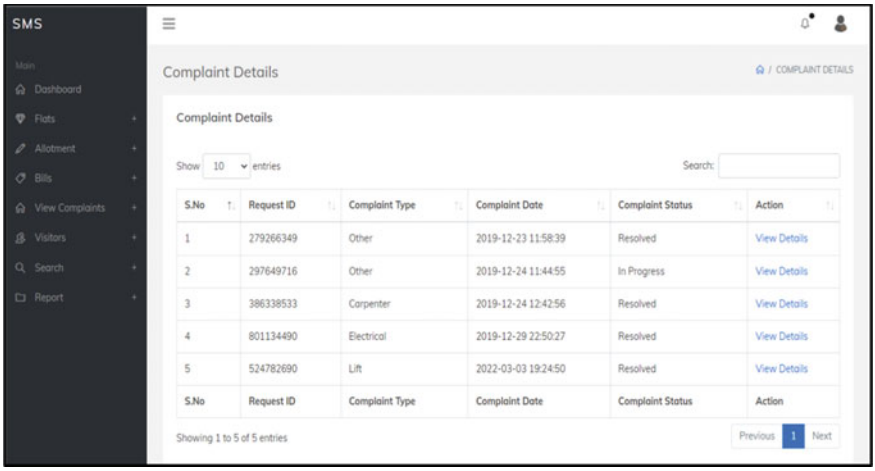


Fig. 11 Complaint details

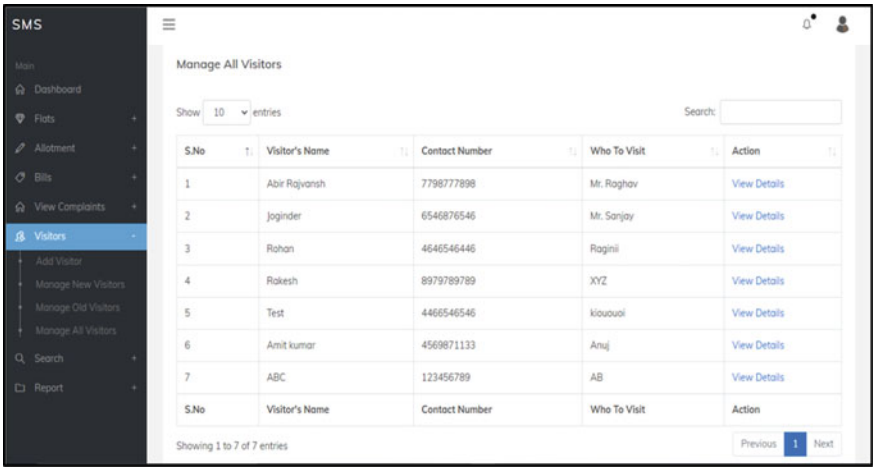


Fig. 12 Details of visitors

5 Conclusion

The complete solution will give a simple, user-friendly, and cost-effective platform for communication, fostering healthy relationships among housing society members and a more intelligent manner of engaging with one another. It also aids in the reduction of time and effort required for manual communication in society by giving reliable and transparent notifications and crucial information to society members. The system encourages inhabitants to communicate with one another by keeping them informed about cultural and sporting events. It will be quite easy to promote one's business and supply various services to people using this application, resulting in increased employment through socialism. It is not necessary to meet with the chairman panel in person to make concerns or suggestions, or to accept the panel's appointments. This job can be completed using the application by sending messages to the secretary and receiving the job's response. It allows residents of separate societies to be continually informed about current developments in their communities and to engage in various events promoting distinct cultures.

Acknowledgements We would like to convey our sincere thanks to everyone who guided us throughout this project endeavor, before we present our project entitled "Housing Society Management System using PHP." We are extremely grateful to Dr. Mansing Rathod, our Project guide, for sharing expertise, along with their sincere and valuable guidance; and encouragement extended to us. We take this opportunity to express gratitude to all the department faculty members for their help and support. We also thank our parents for their unceasing encouragement, support, and attention.

References

1. Solanki H, Yadav D, Yadav A, Shah C (2019) Information Technology Department, Vidyalankar Institute of Technology, Mumbai. Corresponding Author: Hitesh Solanki, IOSR J Comput Eng (IOSR-JCE) 21(2):01–04. e-ISSN: 2278–0661, p-ISSN: 2278–8727, Ser. IV. https://www.iosrjournals.org
2. Hande J, Pawar P, Hasan N, Date D (2016) Int J Res Appl Sci Eng Technol (IJRASET) 4(III):16–21
3. Raut S, Pawar P, Shaikh M, Bhat N, Khalwadekar PN (2016) Housing societies in social network. Int J Sci Res Sci Technol (IJSRST) 2(6):134–137
4. Kohar M, Rana C (2012) Survey paper on recommendation system. Int J Comput Sci Inf Technol (IJCSIT) 3(2):3460–3462
5. Patadia K, Manjrekar K, Zope T, Patil P (2014) Web portal: housing management system. Int J Eng Res Appl 4,(2):547–551, ISSN: 2248–9622
6. Gavhane S, Vatharkar R, Sonar S, Patil P (2015) Study of ımplementation of society management system. Int J Comput Appl 132(1):34–36
7. Barbar M, Khan F, Iqbal W, Yahy A (2016) A secured data management scheme for smart societies in ındustrial ınternet of things environment. IEEE Access. https://doi.org/10.1109/ACCESS.2018.2861421
8. Bhagwat R, Bharadwaj A, Harsode V, Chawake A, Bhanage D (2018) Society management application on android. Int Res J Eng Technol (IJRET)
9. Ngai YongHub EWT, Wonga YH, Xin Sub YC (2011) The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature. Decision Supp Syst 50(3):559–569
10. Vatharkar R, Patil P (2016) Implementation of society management system: societales. Int J Sci Technol
11. Vora D, Shah R (2017) Emergency management system using android application. Int J Adv Res Comput Eng Technol (IJARCET) 6(2), ISSN: 2278–1323
12. Hua T (2009) Discussion on the core technology of smart home system. Intelligent Home 3(1):92–98
13. Jadhav R, Patel J, Jain D, Phad-htare S (2014) Emergency management system using android application. (IJCSIT) Int J Comput Sci Inf Technol 5(3):2803–2805
14. Romualdo L, Brown SK, Pipes S, Ibbotson J (2014) Smart building management through augmented reality. The First Int Workshop Pervasive Syst Smart Cities
15. Hande J, Pawar P, Hasan N, Date D (2016) Study of housing society management system. Int J Comput Trends Technol (IJCTT) 41(2)

Text Summarization Using Combination of Sequence-To-Sequence Model with Attention Approach

Prasad Bhandarkar and K. T. Thomas

Abstract In daily life, we come across tons and tons of information which can be related to news articles or any kind of social media posts or customer reviews related to product. It is difficult to read all the content due to time constraint. Being able to develop the software that can identify and automatically extract the important information. There are two types of summarization methods. Extractive text summarization is the method where it picks the important content from the source text and gives same in the form of short summary, and on the other hand, abstractive summarization is the technique where it gets the context of the source text, and based on that context, it regenerates small and crisps summary. In this paper, we use the concept of neural network with attention layer to deal with abstractive text summarization that generates short summary of a long piece of text using review dataset.

Keywords Text summarization · Long-term short memory · Recurrent neural network · Attention layer · Decoder–encoder

1 Introduction

In the contemporary era of big data and artificial intelligence, getting useful information and insight from a huge text document is a difficult work, because of the growth in the availability of customer reviews, research articles, news articles, are very huge in terms of counts. Text summarization gives a constructive solution for summarizing source text. The logic of the summarization is to load the source text or document and generate the small summaries without changing the crucial information and context of it [1]. Due to recent development in deep learning and natural language models like sequence-to-sequence [2] and attention base model help us to build an extensive variety of applications in different fields including economic, education, and scientific research, medical as a whole.

P. Bhandarkar (✉) · K. T. Thomas
Department of Data Science, CHRIST University, Bangalore, India
e-mail: Bhandarkar.prashant@science.christuniversity.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_22

1.1 Problem Statement

Text summarization using combination of sequence-to-sequence model with attention layer.

1.2 Types of Text Summarization Methods

In the last few years, generating short summaries, getting content from the text, processed and break down productively and efficaciously is difficult task. Generally, there are two kinds of methods to perform text summarization:

Extractive Text summarization: It is established technique developed years back. The core idea is to identify the important sentences from the source text and concatenate in the form of summary. In simple word, it will extract the important sentence or the text from the source document [3].
Source Text: Om and Sai took a cab to attend the night party in the city.
Summary: Om and Sai attend part.

Abstractive Text summarization: It is a sophisticated technique which identify the crucial part of the text and interpret the context and generate the new sentence [1].

The motive behind this method is to convey the information without changing the meaning of the sentence in the shortest way as we can see in Fig. 1.

Source Text: Om and Sai took a taxi to attend the night party in the city. While at the party, Om collapsed and was rushed to the hospital.

Summary: Om was hospitalized after attending a party with Sai.

In this paper, we put forward abstractive base text summarization using feed-forward neural network and recurrent neural network with attention base encoder and decoder method [4] and evaluate the model using ROUGE score [5].

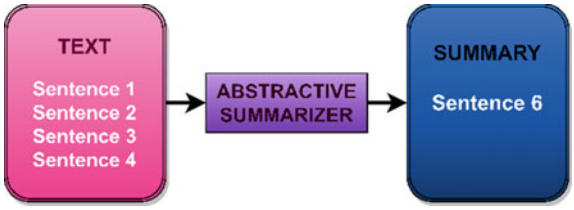


Fig. 1 Abstractive text summarization

1.3 Contribution

During the research, we have found that using plan sequence-to-sequence model was having some limitation with respect to remembering long text. When we used attention layer in the architecture, the model outperformed the previous seq-to-seq model.

1.4 Organization of the Paper

As for the paper, we review some of the research paper related to our model in Sect. 2 before presenting our proposed method and dataset description in Sect. 3. We evaluate the results of our model in Sect. 4. Section 5 concludes the paper.

2 Related Works

The method which is been used in several papers focuses on extractive text summarization. The first paper was published by Luhn [6] in 1958. In his work, he proposed that in the document the importance of the word is depending upon the frequency of that word. This was the traditional method developed first but we know now the word which is having more frequency count does not mean it is important, e.g., stop words. To overcome this problem in the later stage, Salton et al. [7] come up with the new idea called term weight system to get the frequency of words.

Rada et al. [3] proposed TextRank extractive summarization method. The concept of graph-based ranking algorithm is useful to decide the importance of a vertex in the graph with the help of structure of the graph based on the information recursively drawn from the base structure of the graph. This paper introduces TextRank graph-based ranking model for graphs which are extracted from natural language text. During the study, they found that the result of the TextRank model is comparatively better than the other state-of-the-art system.

On other hand, there have been papers which focus on the abstractive text summarization method. Tian et al. [2] gave a comprehensive survey on various abstractive text summarization using seq-to-seq neural network. In this paper, they have come across multiple models, e.g., recurrent neural network-based Seq2seq models, RNN ENCODER-DECODER framework and their training strategies, evaluation metrics such as ROUGE and BLEU score.

Cho et al. [8] proposed recurrent neural network encoder–decoder method. Using this method, we can achieve the machine translation problem. Along with encoder–decoder architecture, they proposed a combination of two get, reset gate and an update gate. Using this unit, we can control which part of the text needs to remember or forget while processing the sequence of text. It helps in the hidden unit by using

the two recurrent neural networks: One network is for encoding the sequence of fix length vector and other network is to decode the sequence into another sequence of symbols. Because of this architecture, they were able to get the linguistic regularities well and improve the performance in terms of BLEU score. They claimed that, this combination is also helpful for the various applications like text summarization and speech transcription.

Chopra et al. [1] proposed attention base recurrent neural network for abstractive text summarization. Novel convolutional attention-based encoder can control the decoder and what part it needs to focus more. Attentive recurrent architecture has two recurrent decoder and recurrent encoder. The result of the paper clearly says that it is performing comfortably as compared to the previous state-of-the-art on gigaword data.

3 Methodology

If we look at the recent development in summarization techniques, there are few things which we can use to improve the model performance. Using sophisticated dataset and embedding techniques to train the model can make effective changes with respect to performance. In this section, we talk about the system which we have proposed for the implementation. The proposed system has three stages (Fig. 2):

1. Data collection and Preprocessing
2. Model Building
3. Model Result and Summary Generation for custom input.

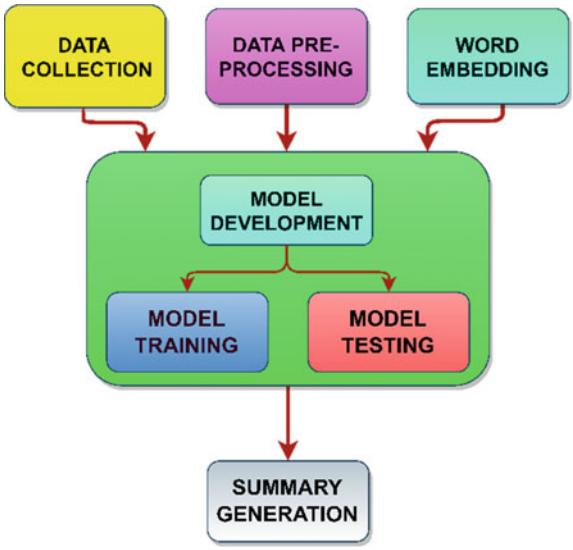


Fig. 2 Proposed system architecture

Table 1 First five rows of final dataset

	Text	Summary
0	I have bought several of the Vitality canned d…	Good Quality Dog Food
1	Product arrived labeled as Jumbo Salted Peanut…	Not as Advertised
2	This is a confection that has been around a fe…	'Delight' says it all
3	If you are looking for the secret ingredient i…	Cough Medicine
4	Great taffy at a great price. There was a wid…	Great taffy

3.1 Data Collection

The dataset that we are considering for this project is Amazon food review that is collected from October 1999 to October 2012. The raw data has 10 columns, namely 'ID', 'ProductID', 'UserID', 'ProfileName', 'HelpfulnessNumerator', 'Helpfulness-Denominator', 'Score', 'Time', 'Summary', 'Text', and 568,454 rows. More than 250 users have given their reviews regarding 74,258 products. Summary column has 295,744 unique value out of 568 k valid rows. The dataframe had two columns, that is, 'Text' and 'Summary', in the final stage after eliminating the unwanted columns as per the requirement and saved it in new csv file as you can see in Table 1.

3.2 Data Preprocessing

Performing basic preprocessing steps are very important before we get to the model building phase. Using dirty and unclean text data is a potentially dangerous move.

3.2.1 Data Cleaning

To process, the unclean data is always a difficult task. In this step, we have removed all punctuation using python packages and natural language tool kit (NLTK). NLTK function word_tokenize split the punctuations from words and separates it. Unwanted characters like @, #, $ were removed using TweetTokenizer. It is also one of the tokenized functions in NLTK. Stop words ('a', 'The', 'I', 'are', etc.) do not add any information which can be helpful for the model so keeping them in the process is not a good idea. Using NLTK function called stopwords.words, we have removed it.

3.2.2 Data Null Values

In the preprocessing part, handling null values are the crucial part because it does not add any values to the model so it is always better to replace it with some value or remove it. In our case, we have removed it using pandas' function.dropna.

Table 2 Start and end tokens

S. No.	Used tokens	description
1	SOSTOK	Start of the input summary
2	EOSTOK	End of the input summary

3.3 Word Embedding Layer

It is a process where each word represents the real value vector. After preprocessing, we have given the vector value to each word from word embedding and given a unique value to the word which is missing or not present in the WordToVec library. Also, we can use Keras embedding vectorizer to optimize the model. To make the system understand where is the text Start and End, we have added the tokens at the start and the end of the sentence using lambda function (Table 2).

3.4 Model Description

To build the constructive text summarizer, different types of deep learning methods have been used.

3.4.1 Sequence-to-Sequence Model

Encoder–decoder architecture is widely used to manage sequence-to-sequence problem; now the goal is to provide the vector form of each word along with the output of all previous hidden layer and base on the context vector generate the another sequence of words. Initially, the encoder integrates each item into a long string of words and provides context information [9]. After completing the sequence of all inputs, the encoder sends the context vector to the decoder to generate another sequence as an output [8]. Sequence-to-sequence Model is a recurrent neural network base model. To work on the textual sequence, deep learning technique recurrent neural network (RNN) outperforms as compared to other deep learning models [10]. When we look at the architecture of RNN, current hidden state is dependent on the previous layer hidden state and new input. This dependent hidden state helps the model to predict the next output in the sequence. For example, when we look at the complete sentence every next word is partially or fully dependent on the meaning of the previous word [11]. In text summarization, generating summary as per the context of the source text is the most crucial task. There are some limitations of this architecture, when the input is long sequence of words it is very difficult for encoder to remember all the previous work for computation so it does not work with long dependencies. Its only remembers fixed length vectors and that get difficult for encoder to memorize long sequence representation. To overcome this problem, Bahdanau comes up with attention layer [4].

3.4.2 RNN Encoder–Decoder with Attention Base Layer

Our model was based on bidirectional RNN using attention layer called Bahdanau attention [4] on the input data. In this architecture, we have used LSTM [12] cell as encoder and decoder to build the abstractive summarization. RNN faces the problem of vanishing gradients while training the model. If model allows gradients to back-propagate in a linear way using gates which is having weight and bias that can solve the problem of vanishing gradients. This gate architecture can control the flow of information in the hidden state thought out the process. In the training phase, this weight and base can get updated. Bi–RNN will transfer the sequence in two ways. Hear LSTM encoders cell extract a fix length of sequence of the input, and the LSTM decoder generates a translation of this sequence. Bahdanau attention layer extended this architecture to allow decoder has more direct access to the input sequence.

Long-term short memory (LSTM) [13] is a one type of gated recurrent neural network architecture to remember the long-term dependencies using different gates. LSTM overcomes the vanishing gradients problem of normal RNN architecture.

In Fig. 3, we can see that this architecture has forget gate, input gate, and output gate. Using forget gate, LSTM cell can decide which information we need to keep from the previous output. The mechanism is to calculate the forget gate value for each input word and apply the sigmoid function. It returns the value between 0 and 1. If value of forget gate is zero, then it will forget everything. If the value is one, then forget gate will remember [14]. This property of forget gate makes LSTM powerful to which text it needs to remember and which one is to forgot. To identify the most important input for the summary, we use attention layer. Attention layer solves the problem of long sequences by predicting a word by seeing a certain part of the sequence rather than the whole sequence. We send multiple hidden context vectors as the number of events in the input sequence, and this produces each time in the output sequence. Thus, without looking at all the words in sequence, we can only focus on the most important part of the sequence by giving them the additional value that indicates the intended sequence. This is a basic idea about how to pay attention. There are two distinct categories of attention method that also depends on what the requirement is and how we obtain the context vector.

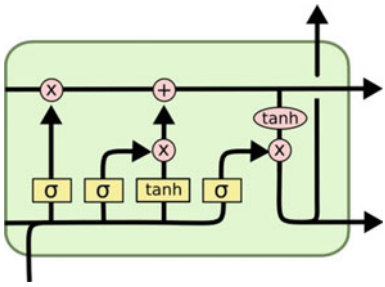


Fig. 3 LSTM architecture

3.5 Model Implementation

Model implementation workflow was divided into three parts.

3.5.1 Model Training

The proposed model implemented using python module called TensorFlow and train on Google Colab which provides all the necessary hardware requirements. For training, we have used 70% data and remaining 20%, 10% were for testing and validation.

3.5.2 Hyperparameter

Choosing right set of parameters and initialize the value of each parameter to archive the optimized solution is called hyperparameter tuning. This parameter must be decided before the training so it can control the algorithm during the training phase.

Set of hyperparameters of our model is Number of Epochs = 70, Batch size = 128, RNN size = 256 with learning rate 0.001. We used RMSprop optimizer which restricts the oscillation in the upward direction, so we can increase the learning rate.

3.5.3 Model Testing

After training phase, we tested our model on 120,000 sample and model was successfully able to create the relevant summaries review written about the product. Model has generated the summaries of the review which contains new word and vocabulary. Figure 4 shows automatic summary generated by the model.

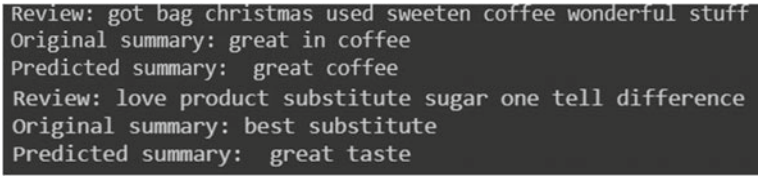


Fig. 4 Generated summaries from test sample

4 Result

Evaluating the performance of the model is the most important part in text summarization model before deploying it in production. One of the major challenges in any sort of text generating model is that there is not a single right answer. For recurrent neural network base language models, we use perplexity as an evaluation metric but it is not sufficient to evaluate techniques like text summarization. To measure the performance of the model, human evaluation is the most accurate method but it is quite expansive because we need to hire someone who is good at language and this process is time-consuming as well. To make evaluation at minimum cost and less time-consuming, we consider automatic machine evaluation metric. They are quite simple, we can compute them without any GPU or complex hardware.

4.1 Evaluation Metrix

To evaluate the performance of the model, we expect two factors that can be checked by the performance metric: First is adequacy—How good output of the model is at conveying the information base on the context of the reference text. Second, fluency of the model in terms of grammar of the language and words. The goal of automatic evaluation metric is that compute the similarities between the machine generation summary and reference summary which is given by human translation.

4.1.1 ROUGE Score

Precision and recall of the words method do not penalize the output for reordering or wrong grammar. Because of this reason, it is not considered as a good metric for the text summarization model. To overcome this problem, we used Recall-Oriented Understudy for Gisting Evaluation also known as ROUGE score. It adds the brevity penalty. If the generated text is to short, it will penalize the output by dividing the score by the length of the reference. ROUGE calculates—How much part of text or n-grams it covered from reference summary to machine-generated summary output. ROUGE -N is an one of the variant ROUGE. N represents the number of grams, and it varies from 1 to N. ROUGE-1, ROUGE-2 ….ROUGE-N for uni-gram, bi-gram …. n-gram respectively. ROUGE-L is another type of ROUGE, It checks the longest sub-sequence between generated summary and reference. We have used ROUGE-1, ROUGE-2 AND ROUGE-L for model evaluation and got 36.5, 21.3, 27.8 scores, respectively.

4.1.2 BLEU Score

BLEU stands for bilingual evaluation understanding. It is a precision base performance matrix. It is used to calculate the words from the source text repeated to the generated text. BLEU measures the quality of the text generated by machine model. How close that machine translated output from professional human translation. The core idea is more it is closer, the more it is better. BLEU counts the unigrams which are matches to generated summary.

Original Review: 'love green mountain coffee first one bought got kurig strong'
Expected Summary: ['good', 'coffee'].
Generated Summary: ['great', 'coffee'].
BLEU score = 0.5287.

In the above sample, expected summary has two words 'good' and 'coffee' but in generated summary only 'coffee' is matching. BLEU score is measured by total word in the expected summary divided by match words in the generated summary by this formula, we got 0.5287 for this sample.

5 Conclusion

Text summarization is one of the fastest-growing domains when it comes to natural language processing. The main reason for such tremendous growth is the availability of a huge number of use cases for the text summarization domain. An important key factor to note while creating a good text summarization system is to focus on the words which can deliver the context without being dependent on most other words in the sentence. This factor was taken care of by the attention-based models which have gained a lot of appreciation from the research communities in recent years. Considering the social impact of such models, there is a huge chunk of the population that uses mobile devices but are not aware of basic reading and writing skills, especially when things are not conveyed in native language. For such users in order to make sure, they do not fall into fraud activities, many telecom companies send only the very much required information to their devices for authentication by applying text summarization to the already available authentication messages. Currently, text summarization models come with a huge number of parameters as they are trained on a large number of corpus but the present focus of research is to compress these models so that edge devices can also be benefitted from text summarization use cases and hence provide benefit to a larger section of the audience.

References

1. Chopra S, Auli M, Rush A (2016) Abstractive sentence summarization with attentive recurrent neural networks. In: Proceedings of the 2016 conference of the north American chapter of the association for computational linguistics: human language technologies. Available: https://doi.org/10.18653/v1/n16-1012. Accessed on 18 Feb 2022
2. Shi T, Keneshloo Y, Ramakrishnan N, Reddy CK. Senior member, IEEE, "neural abstractive text summarization with sequence-to-sequence models: a survey. https://arxiv.org/pdf/1812.02303.pdf
3. Mihalcea R, Paul T (2004) Textrank: bringing order into text. In: Proceedings of the 2004 conference on empirical methods in natural language processing
4. Bahdanau D, Cho K, Bengio Y (2016) Neural machine translation by jointly learning to align and translate. [1409.0473] Neur Mach Trans Jointly Learn Align Trans
5. Luong MT, Pham H, Manning CD (2015) Effective approaches to attention-based neural machine translation. arXiv preprint arXiv:1508.04025
6. Luhn HP (1958) The automatic creation of literature abstracts, IBM J Res Dev 2(2)
7. Salton G, Buckley C (1988) Term-weighting approaches in automatic text retrieval. Inf Process Manage 24:513–523. https://doi.org/10.1016/0306-4573(88)90021-0
8. Cho K, Merriënboer van B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder–decoder for statistical machine translation. https://doi.org/10.3115/v1/D14-1179
9. Suleiman D, Awajan A (2020) Deep learning based abstractive text summarization: approaches, datasets, evaluation measures, and challenges. Math Probl Eng. https://doi.org/10.1155/2020/9365340
10. Giles CL, Kuhn GM, Williams RJ (1994) Dynamic recurrent neural networks: theory and applications. IEEE Trans Neural Netw 5(2):153–156. View at: Publisher Site | Google Scholar
11. Mohammad Masum AK, Abujar S, Islam Talukder MA, Azad Rabby AKMS, Hossain SA (2019) Abstractive method of text summarization with sequence-to-sequence RNNs. In: 2019 10th ınternational conference on computing, communication and networking technologies (ICCCNT), pp 1–5. https://doi.org/10.1109/ICCCNT45670.2019.8944620
12. Greff K, Srivastava RK, Koutník J, Steunebrink BR, Schmidhuber J (2017) LSTM: a search space odyssey. IEEE Trans Neur Netw Learn Syst 28(10):2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924
13. Adhikari S (2009) NLP based machine learning approaches for text summarization. In: 2020 fourth ınternational conference on computing methodologies and communication (ICCMC), pp 535–538 https://doi.org/10.1109/ICCMC48092.2020.ICCMC-00099
14. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9,(8):1735–1780. View at: Publisher Site|Google Scholar

A Comparative Study of Machine Learning Techniques for Credit Card Customer Churn Prediction

Anusmita Bose and K. T. Thomas

Abstract A customer is a churner when a customer moves from one service provider to another. Nowadays, with an increasing number of severe competition with inside the market, essential banks pay extra interest on customer courting management. A robust and real-time credit card holder's churn evaluation is vital and valuable for bankers to preserve credit cardholders. Much research has been observed that retaining an old customer is more than five times easier compared to gaining a new customer. Hence, this paper proposes a method to predict churns based on a bank dataset. In this work, "Synthetic Minority Oversampling Technique" (SMOTE) has been used for handling the imbalanced dataset. Credit card customer churn is predicted using random forest, k-nearest neighbor, and two boosting algorithms, XGBoost and CatBoost. Hyperparameter tuning using grid search has been used to increase the accuracy. The experimental result shows Catboost has achieved an accuracy of 97.85% and tends to do better than the other models.

Keywords Machine learning · Churn prediction · SMOTE · Banking industry · Random forest classifier · CatBoost

1 Introduction

Customers are critical for a company to get positive results. Churn is a barrier to a company's profitable expansion. Customer turnover is a major headache in the banking business. Because the value of getting new clients much outweighs the value of maintaining current ones, retaining current clients is critical for business success. Because of this, the banking industry must devise a strategy for anticipating and retaining clients who are likely to churn. The turnover rate in the banking business

A. Bose (✉) · K. T. Thomas
Department of Data Science, Christ University, Bangalore, India
e-mail: anusmita.bose@science.christuniversity.in

K. T. Thomas
e-mail: thomas.kt@christuniversity.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_23

has been steadily growing in recent years. The churn rate in Europe is 25% to 30% in the US. According to Gartner India, the churn rate is between 3.5% and 6% per month, and hence, yearly, it may average to around 42% to 72%. Thus, customer churn has become a major problem for banking sectors worldwide. Therefore, it is imperative that this problem faced by the banking industry be immediately addressed by building a better predictor model to predict customer churn. Nowadays, churn prediction in real time has become very crucial as it has many positive impacts. It helps in measuring customers' lifetime value and helps the companies to focus on the active users. It helps in finding the customer service lacks a particular business. Improvement of these lacks can increase the revenue on a large scale.

Here, we are trying to compare a few machine learning algorithms which are widely used for the implementation of the objective. With a small dataset, the deep learning algorithms do not work well. Synthetic minority oversampling technique has been used to deal with the problem arrived regarding imbalanced dataset. Two boosting algorithms, XGBoost and CatBoost, are used to better precisely categorize churners and non-churners. Most machine learning models strive for high prediction from a single model; however, boosting algorithms attempt to improve prediction outcomes by training a series of weak models, each of which compensates for the shortcomings of its predecessors. Hyperparameter tuning is used for increased accuracy. The better method was selected by comparing each metric method evaluation based on its accuracy, sensitivity, specificity, and F1-score, calculated for each testing phase.

2 Related Work

In previous researches, customer churn prediction has been made in various ways. Data mining techniques were used in most studies to predict churn in the banking industry. Most of the researchers used a few techniques and presented the comparison study. Anil Kumar and Ravi [1] used data mining techniques for customer credit card churn prediction. Among the different methods studied, it has been discovered that the random forest produces good results for feature-selected datasets in their entirety. With a tenfold cross-validation method for oversampling, the maximum accuracy was 99.32%. Nie et al. [2] focused on two data mining algorithms, logistic regression, and decision tree, to predict churns. The decision tree's error rate and cost were greater than the logistic regression. As a result, it was determined that a logistic regression outperforms a decision tree.

Rajamohamed and Manokaran [3] used hybrid data mining technique to predict churn. Clustering algorithms used were k-means, fuzzy c-means, rough k-means, and classification algorithm used was k-nearest neighbor, decision tree, random forest, SVM, Naïve Bayes. Combining k-means and rough k-means with classification models separately gave higher accuracy than single classifiers. It can be concluded by the end of this paper; hybrid classifiers perform better than a single classifier.

To anticipate customer turnover, Wang et al. [4] focused on traditional prediction approaches. Ridge regression, rough set theory (RST), least squares support vector machine (LS-SVM), decision tree, and ANN were among the methods utilized. The results suggest that the LS-SVM model is beneficial for improving classification and prediction accuracy in challenging circumstances containing nonlinear interactions between predictors and target variables, based on the accuracies of the four models.

Hanif [5] used the XGBoost algorithm to solve customer churn prediction. The XGBoost algorithm used recursive feature elimination (REF) to eliminate its features, while the LogReg algorithm selected features using multicollinearity checking. XGBoost algorithm gave better prediction compared with the LogReg.

Vafeiadis et al. [6] looked at the effectiveness and the exhibition of KNN, decision tree, and random forest methods for churn prediction. Random forest is used for creating the GUI. The GUI is built using the Tkinter package in Python. This system will predict whether the customer will leave the company based on the trained model. Random forest is considered a more efficient technique.

Wang et al. [7] focused on different data mining algorithms to solve customer churn prediction. SMOTE is used for balancing data. An ensemble model utilizing the soft voting approach was developed on atop these individual models with the best hyperparameters, and it achieved the greatest AUC score.

To tackle customer churn prediction, Rahman and Kumar [8] employed four data mining algorithms: KNN, SVM, decision tree, and random forest. The filter types of minimum redundancy maximum relevance (MRMR) and relief feature selection were utilized. Observations: The predictive value of tree classification is reduced by feature selection, and unlike other classifiers, oversampling in SVM diminishes the value. This is due to a skewed bank account record. As a result, SVM is unable to adequately process the data.

Hammoudeh et al. [9] focused on different models and ensemble models for predicting churn. Artificial neural networks (ANNs), support vector machine (SVM), convolutional neural networks (CNNs), averaging ensemble model (AEM), random forests (RFs), averaging ensemble model (AEM), selective ensemble model (SEM) were implemented. Among a set of trained machine learning models, SEM includes only the models that give correct predictions and filters out those models that give wrong prediction result. Overall, SEM performed better than others.

To forecast customer turnover, Yang et al. [10] presented logistic regression, SVM, random forest, and cost-sensitive learning using RF. To address the problem of class imbalance, researchers adopted a borderline SMOTE or SMOTE + Tomeklink oversampling strategy, in which the sample distribution is kept as the amount of training data increases. In terms of average recall and G-means, SVM is shown to be the most effective.

A relational approach was provided by Oskarsdottir et al. [11] to forecast client attrition. Seven CDR Datasets from across the world were used in the studies. The two forms of relational learners are relational classifiers and collective inference methods. The networks were built using the number and length of phone conversations. In both cases, edges were weighted over time as they degraded. Performances were verified

using lift, AUC, and the H-measure. It was identified by examining the influence of collective inference methods on the performance of relational classifiers.

According to Dalmia et al. [12], supervised machine learning techniques must be used to anticipate bank client attrition. To improve performance, KNN was integrated with XGBoost. K-fold cross-validation method is applied by specifying ten folds. In terms of accuracy, sensitivity, and specificity, XGBoost outperformed the others.

To forecast customer attrition in telecommunication, Mohammad et al. [13] employed logistic regression, ANN, and random forest. Recursive feature elimination (RFE) was used to choose features, which improved accuracy. The confusion matrix and AUC were used to validate the results. Logistic regression performed best.

Szmydt [14] proposed an outline of the research on churn prediction in electronic banking. The study found that electronic churn in the banking business could be an important factor worth investigating further. The analysis presented here is simply a snapshot of an issue, and the investigation is still ongoing.

Senthilnayaki et al. [15] suggested multiple machine learning classification models to predict customer turnover, including logistic regression, random forest, and KNN. Only, the most useful attributes from the entire dataset have been listed. PCA was used on the training dataset for feature selection and dimensionality reduction. The best model was chosen by estimating its performance using the "stratified k-fold cross-validation" [15] approach. In comparison to previous strategies, random forest has been shown to produce superior outcomes.

3 Methodology

This section explains how the boosting and well-known base algorithms were used in the credit card churn prediction model. Figure 1 represents the block schematic of proposed work.

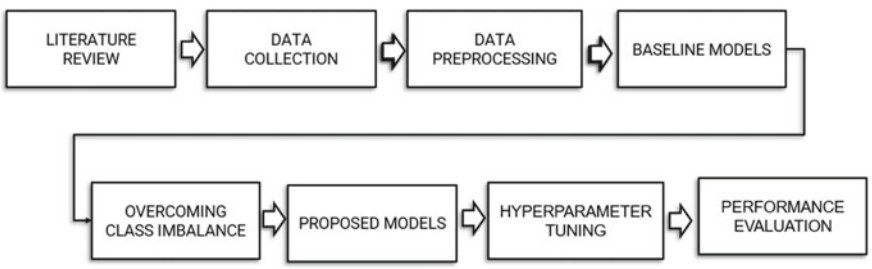


Fig. 1 Block schematic of proposed work

3.1 Dataset Description

Kaggle provided the credit card customer churn dataset utilized in this model. It provides details about 10,127 bank customers. There were 8500 current customers and 1627 churned consumers among them. While the customer is expected to churn, the target variable returns 1; when the client is inclined to remain, it returns 0. The dataset comprises 21 characteristics, with 20 of these 21 predictors derived from the customer's information and transactions. Table 1 has extensive information on these aspects.

Table 1 Dataset description

Feature name	Feature description
Clientnum	Uniquely identifies client account number
Attrition_Flag	1 stands for closed account otherwise 0
Customer_Age	Age of the customer
Gender	M stands for male and F stands for female
Dependent_count	Number of dependents the account holder has
Education_Level	Account holder's academic qualifications (Like: high school, college graduate, and post graduate.)
Marital_Status	Can have values—married, single, divorced, widowed
Income_Category	Annual income of the account holder
Card_Category	Product variable—type of card (Blue, Silver, Gold, Platinum)
Months_on_book	Time span of relationship between the bank and the customer
Total_Relationship_Count	Total no. of products held by the customer
Months_Inactive_12_mon	Total no. of months inactive in the last 12 months
Contacts_Count_12_mon	No. of contacts in the last 12 months
Credit_Limit	Credit limit
Total_Revolving_Bal	Total revolving balance on the credit card
Avg_Open_To_Buy	Open to buy credit line (average of last 12 months)
Total_Amt_Chng_Q4_Q1	Change in transaction amount (Q4 over Q1)
Total_Trans_Amt	Sum of transaction amount (Last 12 months)
Total_Trans_Ct	Number of transaction during the previous 12 months
Total_Ct_Chng_Q4_Q1	Difference in number of transactions Q1 to Q4
Avg_Utilization_Ratio	Average card utilization ratio

3.2 Data Preprocessing

The data are checked for missing values, and if any missing values are found, there are no missing values in the dataset. Binary values like yes or no are replaced with zeros and ones.

3.3 Label Encoding

Many columns in the dataset contain categorical values in different forms. For example, the card category has four categories: blue, silver, gold, and platinum. So, all categorical values like card category, income category, education, and attrition are converted into integers using a label encoder.

3.4 Feature Scaling

Data are standardized using a standard scaler. So, the values will have to mean as zero and unit variance.

3.5 Feature Selection

Based on relevancy, three columns are dropped, which are unnecessary, and the rest 20 attributes which have a major impact on predicting credit card churn are used for modeling. Out of 20 features, the attrition flag is used as target variables, and the rest 19 features are used as independent variables.

3.6 Smote

Because the dataset is severely imbalanced, an oversampling approach is employed to tailor the data's class distribution. Minority class duplicates are created using the conventional oversampling procedure. Still, instead of making exact replicas, the "Synthetic Minority Over Sampling Technique" (SMOTE) technique generates synthetic samples for the minority class. Overcoming the problem of overfitting produced by random oversampling is easier with this method. For credit cards, the churn goal column contains uneven classes of churners and non-churners. As a result, the data must be balanced when employing the SMOTE approach, and the minority class will be oversampled. Otherwise, the non-churning majority class will teach

the model more. Earlier, there were 8500 loyal customers and 1627 churners in the dataset, after applying SMOTE the ratio of churners and non-churners changed to 1:1. Figure 2 shows a graphical depiction of churners vs. committed consumers, and Fig. 3 projects the available sample count in the dataset on each class after implementing the SMOTE.

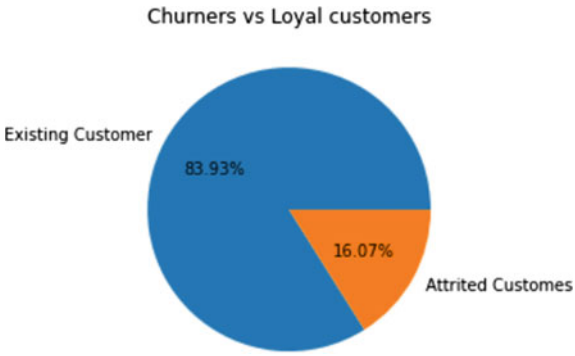


Fig. 2 Pie chart of churners versus loyal customers

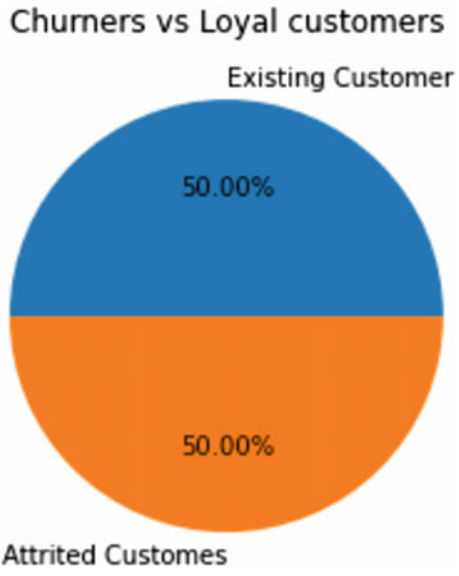


Fig. 3 Pie chart of churners versus loyal customers after SMOTE

3.7 Machine Learning Models

3.7.1 Random Forest

Random forest is classified as a supervised machine learning technique that is mostly used to solve regression and classification issues. It operates in parallel, constructing decision trees for several samples and extracting the majority of votes for classification and the average for regression from them. It can handle datasets with both continuous and categorical variables, as in regression and classification. Random forest has been selected for prediction as it is well known for producing superior results for classification problems.

3.7.2 KNN

KNN is the abbreviation for "K-Nearest Neighbor." This machine learning algorithm is used to solve regression and classification issues. The letter 'K' denotes the number of closest neighbors to an unknown variable that must be classified or predicted. The purpose of KNN is to find all of a new unknown data point's nearest neighbor in order to establish which class it belongs to. It is, in fact, a strategy predicated on distance. KNN estimates the distance between all locations near the new unknown data and filters out those with the shortest distances to them using Euclidean distance. To appropriately characterize the findings (number of nearest neighbors), we must first determine the value of K.

3.7.3 XGBoost

It is also known as "Extreme Gradient Boosting." It uses a gradient boosting framework and is labeled "an upgraded gradient boosting library." When it comes to prediction challenges involving unstructured data such as photos and text, neural networks excel. When it is a matter of too small to medium structured data or tabular data, however, decision tree-based algorithms are considered to be good performers. In machine learning, XGBoost is often used for supervised learning. The gradient boosting decision tree technique is implemented. Gradient boosting is a method that creates new models that calculates the error in the previous model and then add the remainder to make the final prediction.

3.7.4 CatBoost

CatBoost is a machine learning method that was recently open sourced. It works with various data types to assist organizations to deal with a variety of problems. To top it off, it has the best accuracy in the industry. It offers cutting-edge results without

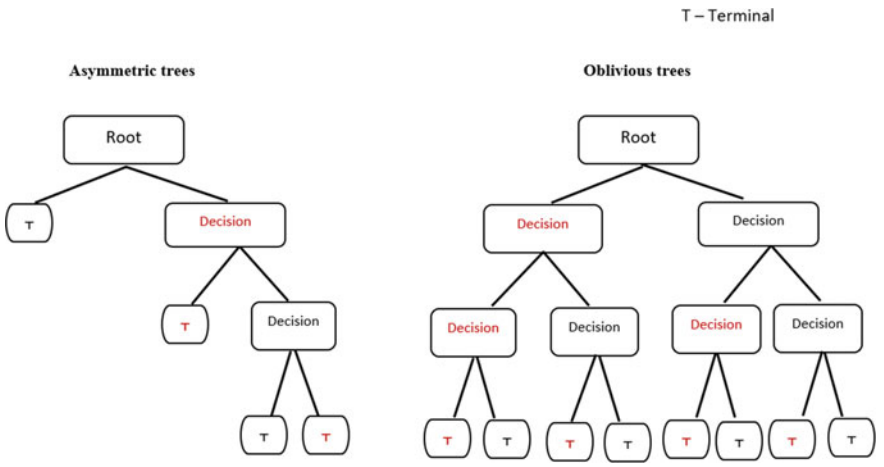


Fig. 4 Splitting of CatBoost

the extensive data training required by other machine learning techniques, and it has solid out-of-the-box support for the more descriptive data formats required by many commercial challenges. Gradient boosting is a sophisticated ML approach that has been successfully practical to a variety of commercial difficulties like fraud detection, recommendation items, and forecasting. It also can return an excellent result with comparatively less data, unlike DL models that require a vast amount of data for learning. CatBoost uses oblivious decision trees, which is unlike asymmetric trees the splitting criterion is the same across an entire level of the tree. Figure 4 shows the working of the algorithm.

3.8 Model Training

For training, 70% of data has been used, and 30% of data has been used for testing purposes. Random forest and k-nearest neighbor were implemented as baseline models, and XgBoost and CatBoost as proposed models.

3.8.1 Elbow Curve

The elbow curve approach was utilized to provide optimum solutions using KNN. The most important step in supervised machine learning using k-nearest neighbors is figuring out what the best value of K is, or how many clusters your data should be partitioned into. The best k-value decreases the influence of noise on classification while also blurring class boundaries. Because the user must give k in advance, the procedure is rather naive in that it assigns all members k to clusters even if it is not

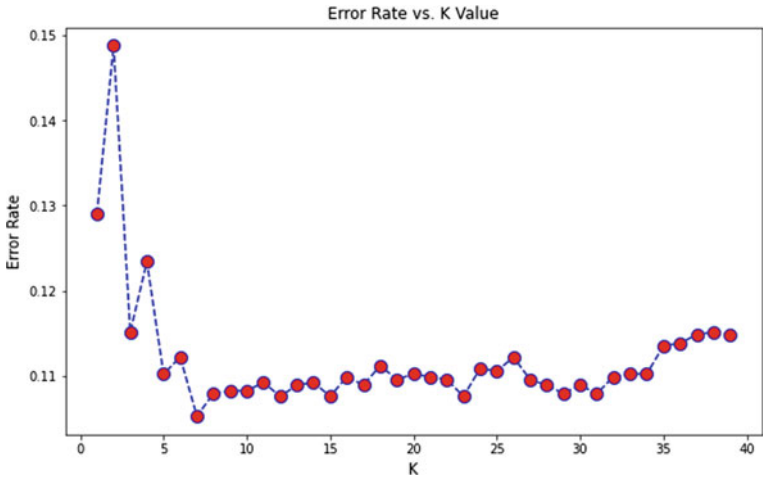


Fig. 5 Elbow curve

the proper k for the dataset. The graphical representation of the elbow curve method is shown in Fig. 5; clearly, we can see the optimal value of k for our dataset obtained was 7.

3.8.2 Grid Search

Hyperparameters like batch size, number of hidden layers, and learning rate are fine-tuned to achieve the best results on various predictions. Grid search is one of the simplest algorithms for hyperparameter tuning. The domain of the hyperparameters is divided into a discrete grid. After this, we try every possible combination of the grid values and find the best combination for optimal results. Here, we have used this on the CatBoost algorithm, using the maximum depth determined was ten, and number estimators determined was 200 for getting the best result.

4 Results and Discussions

After model building and training on data, the models are evaluated on each stage according to the performance of the evaluation metrics. The evaluation metrics used are accuracy, precision, recall, and F1-score. Accuracy score helps in understanding how accurately the models are predicting classes, precision gives us the number of true positive classes, whereas recall quantifies the actual true positive classes. So, giving more priority for recall score over precision score enhances the result accurately. F1-score is simple representation of relation between recall and precision

which is computed by taking their harmonic mean. The first two baseline models that are random forest and KNN gave accuracy scores of 96.54% and 89.47%, respectively, at the initial stage. Then, after applying standardization scaler, the accuracy of random forest model remained same as before, whereas for KNN, the accuracy decreased to 88.12%. Then, in order to deal with the class imbalanced issue, we implemented SMOTE [7, 16], for which the accuracy increased for random forest to 97.22%, and for KNN, it increased a bit to 88.22%. Then, after using the proposed models, the accuracy came 97.55% and 97.65% for XGBoost and CatBoost, respectively. At the final stage, after hyperparameter tuning, the final accuracy obtained is 97.86%. The precision value did not increase after standardization. It remained the same for random forest and decreased from 92.75% to 90.16% for KNN. After applying SMOTE, the precision score increased to 98.09% for the random forest. This score increased for the CatBoost algorithms after hyperparameter tuning by 0.5%, overall best precision score obtained by CatBoost model after applying this hyperparameter tuning technique. On the other hand, the recall score value increased after standardization for KNN. It remained the same for random forest and increased from 94.98% to 96.50% for KNN. SMOTE decreased the recall value for both the data mining models. Boosting algorithms also gave less score compared to other evaluation metrics. Best recall score obtained by random forest with SMOTE and standardization. The F1-score did not increase after standardization. It remained the same for random forest and decreased from 92.75% to 90.16% for KNN. SMOTE also decreased the F1-score for both models. Best F1-score obtained by random forest without applying SMOTE and standardization. Table 2 shows the accuracy, precision, recall, and F1-scores of different models.

Table 2 Accuracy, precision, recall, and F1-scores

Techniques used	Models	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
Baseline model	Random forest	96.54	97.42	98.52	97.97
	KNN	89.47	92.75	94.98	93.85
Standardization	Random forest	96.54	97.42	98.52	97.97
	KNN	88.12	90.16	96.50	93.22
SMOTE	Random forest	97.22	98.09	96.33	97.20
	KNN	88.22	94.50	81.24	87.37
Proposed model	XGBoost	97.55	97.95	97.15	97.55
	CatBoost	97.65	98.07	97.22	97.64
Hyperparameter tuning	CatBoost	97.86	98.57	97.15	97.85

5 Conclusion and Future Work

We can definitely conclude that boosting algorithms are the most effective for churn prediction based on the data. From the results, we obtained for the experiments, XGBoost, could be seen out performed by CatBoost. The term "customer churn" may be applied to a wide range of businesses, including retail, telecommunications, and car insurance. It is vital to maintain clients to prevent losing revenue. So, we can use the similar predictions in these domains as well. Machine learning algorithms do not scale well as the amount of data rises, which is a downside. So, in future, we can extend our study to include neural networks also for the forecast of customer churn. We can extend the experiement to include data from other domains like insurance and telecom.

References

1. Anil Kumar D, Ravi V (2008) Predicting credit card customer churn in banks using data mining. Int J Data Analy Techniq Strateg 1(1):4–28
2. Nie G, Rowe W, Zhang L, Tian Y, Shi Y (2011) Credit card churn forecasting by logistic regression and decision tree. Expert Syst Appl 38(12):15273–15285
3. Rajamohamed R, Manokaran J (2018) Improved credit card churn prediction based on rough clustering and supervised learning techniques. Clust Comput 21(1):65–77
4. Wang N, Niu DX (2009) Credit card customer churn prediction based on the RST and LS-SVM. In: 2009 6th international conference on service systems and service management, pp 275–279. IEEE
5. Hanif I (2020) Implementing extreme gradient boosting (XGBoost) classifier to improve customer churn prediction
6. Vafeiadis T, Diamantaras KI, Sarigiannidis G, Chatzisavvas KC (2015) A comparison of machine learning techniques for customer churn prediction. Simul Model Pract Theory 55:1–9
7. Wang X, Nguyen K, Nguyen BP (2020) Churn prediction using ensemble learning. In: Proceedings of the 4th international conference on machine learning and soft computing, pp 56–60
8. Rahman M, Kumar V (2020) Machine learning-based customer churn prediction in banking. In: 2020 4th international conference on electronics, communication and aerospace technology (ICECA), pp 1196–1201. IEEE
9. Hammoudeh A, Fraihat M, Almomani M (2019) Selective ensemble model for telecom churn prediction. In: 2019 IEEE jordan international joint conference on electrical engineering and information technology (JEEIT), pp 485–487. IEEE
10. Yang L, Li D, Lu Y (2019) Prediction modeling and analysis for telecom customer churn in two months. arXiv:1911.00558
11. Óskarsdóttir M, Bravo C, Verbeke W, Sarraute C, Baesens B, Vanthienen J (2016) A comparative study of social network classifiers for predicting churn in the telecommunication industry. In: 2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), pp 1151–1158. IEEE
12. Dalmia H, Nikil CV, Kumar S (2020) Churning of bank customers using supervised learning. In: Innovations in electronics and communication engineering, pp 681–691. Springer, Singapore
13. Mohammad NI, Ismail SA, Kama MN, Yusop OM, Azmi A (2019) Customer churn prediction in the telecommunication industry using machine learning classifiers. In: Proceedings of the 3rd international conference on vision, image and signal processing, pp 1–7

14. Szmydt M (2018) Predicting customer churn in electronic banking. In: International conference on business information systems, pp 687–696. Springer, Cham
15. Huang B, Kechadi MT, Buckley B (2012) Customer churn prediction in telecommunications. Expert Syst Appl 39(1):1414–1425
16. Saanchay PM, Thomas KT (2022) An approach for credit card churn prediction using gradient descent. IOT Smart Syst 689–697

Development of a LoRa Network for Monitoring Particulate Matter

Amritpal Kaur and Jeff Kilby

Abstract Air pollution in cities has become an important topic due to its adverse effects on humans and air quality. The aim of this paper is to monitor particles less than 1 μm (PM_1), 2.5 μm ($PM_{2.5}$), 4 μm (PM_4), and 10 μm (PM_{10}). The concentration of particulate matter highly changed with location and time. Particulate matter in the air is considered as the primary pollutant, and it affects the environment band the risk of mortality and morbidity of respiratory disease. To address this issue, the research presents the design and development of the low-cost network for monitoring particulate matter using sensiron sensor (SPS30). These devices are equipped with LoRaWAN to test the low-power wide-area network coverage. The designed network contains the sensors connected to the ESP32 microcontroller towards the processing of LoRa modules (sensor nodes), which send data to the gateway using the frequency band, using the 'The Things Network (TTN)'. The sensor collects the different particle matter in the air. The proposed network design has been implemented at St. Paul Street Auckland. The designed network system allows the users to access the online dashboard to test and monitor the concentration levels of particulate matter in the air.

Keywords Air quality indexing · LoRa · Particulate matter

1 Introduction

With the increase in urbanization and industrialization, it decreases the air quality parameters. This is important to know about the air quality parameters using different methods and techniques. The government has developed some policies and procedures for monitoring air quality [1]. When the researchers are using the tra-

A. Kaur (✉) · J. Kilby
Auckland University of Technology, Auckland, New Zealand
e-mail: amritpal.kaur@aut.ac.nz
URL: https://www.aut.ac.nz

J. Kilby
e-mail: jeffrey.kilby@aut.ac.nz

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_24

Table 1 Major air pollutants

Parameter	Sources	Causes
CO and CO_2	Vehicle emissions	Respiratory problems
SO_2	3 industrial smoke, forest fires	Skin and eye irritation
NO_2	Fossil fuels vehicle emission	Asthma
Particle matter (PM)	Dust storms domestic coal burning and municipal waste	Premature deaths respiratory problems
Lead	Petrol and industries	Brain and organ development in children

ditional methods, these are not quite feasible if using the wireless sensor networks (WSNs) in large areas. There are some new techniques developed in WSNs to use low-cost sensors to collect and analyse the real-time air pollution monitoring. WSNs can be used for pollution monitoring with accuracy, calibration techniques, and fault tolerance features in topologies [1]. In the sensor network, it includes the number of sensor nodes that are interconnected with wireless network [2]. The rapid development in industrialization and population increase has led to the rise in pollution. In poor air quality, it includes the harmful gases and particles such as ozone, particle matter, nitrogen dioxide, sulphur dioxide, and carbon monoxide [3]. Table 1 shows the major air pollutants [4].

In the World Health Organization (WHO) [5], air pollution is considered as the bigger environmental issue that causes some health problems and responsible of 9 deaths every year. Industrial and scientific innovations have been developed to provide solution based on the conventional wireless sensor networks (WSNs) [2] for air pollution monitoring system. Air pollution occurs cause of toxic gases, dust, fuel or chemicals integrate with the environment. These toxic gases and particles are harmful to the environment and health. These substances that cause pollution are called pollutants [3]. Over the few decades, several types of research have been done to monitor air pollution. In air quality, it includes different portable particle matter and gas sensor available. The solution to these problems, numerous low-cost air pollution monitoring systems were developed [4].

Numerous studies have shown the pollutant's exposure to the air. But the particulate matter is considered the most dangerous pollutant that causes cardiovascular diseases and other health issues [6–9]. Particulate matters are made up of solid and liquid particles, which enter the atmosphere naturally or with any human activities [10]. The pollutant level of PM pollution can easily assess by continuous monitoring of the level PM_{10} (size less than 10 μm), $PM_{2.5}$ (size less than 2.5 μm), PM_4 (size less than 4 μm), and PM_1 (size less than 1 μm).

Table 2 Communication technologies

Parameter name	Technologies		
	LoRa	Sigfox	NB-IoT
Band (MHz)	868/915	868/915	868/915
PHY	CSS	UNB	NB
Spreading factor	2^7–2^{12}	–	–
Bandwidth (kHz)	500–125	1	180
Data rate (kbps)	27–0.37	0.1	250–226.7
Range (km)	22	63	35

The network system was designed using different wireless communication technologies [4], such as LoRa, Sigfox and NB-IoT. In-depth analysis and understanding of pollution monitoring are an essential to develop the new methodology for pollution monitoring. The newly designed network can only be possible using the accurate monitoring system that helps the network study the pollutant levels in the air [4]. Different communication technologies are compared in Table 2 [5].

2 Air Quality Index (AQI)

AQI shows how the pollutant levels in the air in a particular area. AQI can be calculated using some consideration of pollutants such as carbon monoxide (CO), sulphur dioxide (SO_2), particulate matter (PM) [4]. This is an indicator of air quality, based on various air pollutants that affect human health and the environment. Many countries use air quality indexing, and different nations have their own policies to calculate pollutant levels [11] and as shown in Fig. 1 [12]. AQIs take the pollutant measurements from the air using the air pollution monitoring system and convert these values to a single figure. The nature and scale of the figure are flexible. The scale in AQI ranges from 'safe' to 'hazardous'. This classification can be described to guide health effects and awareness about environmental issues and the breakdown of the particulate matter in the air as given in Table 3.

3 Related Work

Zheng et al. [13] developed the dynamic filter system for outdoor pollution monitoring system for $PM_{2.5}$. The network design network is implemented indoor and outdoor for the testing $PM_{2.5}$ and explains the interrelationship between indoor and outdoor. First, the indoor $PM_{2.5}$ per hour studied and applied Gaussian distribution.

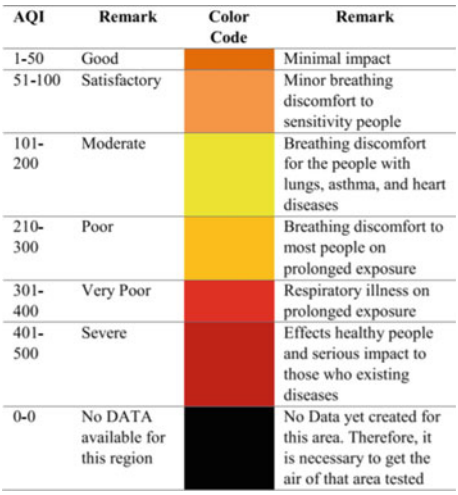


Fig. 1 Air quality indexing

Table 3 Air quality indexing

AQI	Remark	Cause	$PM_{2.5}$
1–50	Good	Minimal impact	0–30
51–100	Satisfactory	Minor breathing issues	31–60
101–200	Moderate	Asthma and heart disease	61–90
210–300	Poor	Prolonged exposure	91–120
301–400	Very poor	Respiratory illness	121–50
401–500	Severe	Serious impact	250+

It performs the statistical distribution of the indoor and outdoor $PM_{2.5}$ pollution level and applied some physical laws and collects data.

Tzai et al. [14] implemented the road network to monitor air pollution, and its affects on humans and environment. In this research, sensor nodes were deployed on crossroads. This paper explains the WSN framework to monitor pollutant levels for the street-level spatial–temporal changes of carbon monoxide. The authors proposed the urban air quality monitoring system based on WSNs that integrate with global system for mobile communication (GSM). The network system is deployed on roads for monitoring carbon monoxide (CO) concentration caused by vehicle emissions.

Chen et al. [15] proposed an open framework for $PM_{2.5}$ monitoring in smart cities. With an increase in population and human activities affects the pollution level in cities. The designed network deployed 2500 sensor nodes in Taiwan and 29 other countries. This paper works to resolve the issues of inaccuracies in the low-cost sensor

networks used for monitoring of particles. Facilitating the network deployment is to investigate the accuracy issue of low-cost particle sensor.

Badura et al. [16] evaluated the low-cost sensors for $PM_{2.5}$ monitoring. This paper presents the results of four models set using different sensors connected to different location under same temperature. The data collection of $PM_{2.5}$ was performed for almost half year. After the research and finding, results show the large dispersion of high relative errors in $PM_{2.5}$.

Guili et al. [17] physically based data analysis model was developed for estimation of concentration levels of $PM_{2.5}$. The research used the fine mode aerosol optical depth at 440, 550 and 675 mm, fine particle radius, humidity and boundary layer height data from 2015 to 2016. The data collected in 2015 was used for calculating the integrated extinction efficiencies based on the network model. The authors concluded the developed method has potential for retrieval of $PM_{2.5}$ using AERONET AOD in Beijing.

Jelicic et al. [18] designed the network system to perform reliable gas monitoring and operating on batteries for one year. Information about the pollutants extracted from the PIR sensors and gas sensors. The system should be able to predict the behaviour, adapt the duty cycle and send the alarm message if necessary. To monitor the reliability of the network, simulations were performed.

Agarwal et al. [19] proposed for the detection and monitoring of the pollution using wireless sensor networks, where the framework has five parts: deployment of sensor node, clustering of nodes, inter-cluster routing protocol and cluster-head selection algorithm. Sensor nodes were deployed in triangular fashion to avoid collision, minimum transmission, detecting pollution, etc. An efficient clustering algorithm was proposed to select the best path.

Knoll et al. [20] presented the denser sensor network to expand the air pollution measurement density, and data is fed into the dispersion models which only provides the approximate results. New network technologies were compared such as LoRa, Sigfox and NB-IoT. Denser sensor network is directly able to monitor air pollution values. The measurements were carried out in the city of Graz, for these measurements an STM32 Nucleo board with an expansion board, featuring LoRa SX1272 transceiver was utilized. Payload for each message contained 16 bytes including a message counter for verification. The sink node was positioned on the third floor outside the window at the university campus, and further nodes for the data collection were placed at the height of 1 m and measurements performed using different technologies. LoRa and Sigfox bring along the required attributes including long-range and small power consumption, where NB-IoT is the latest technology that is strongly driven by the mobile communication.

From the literature review, it shows some limitation of using the existing network systems. To overcome the limitation such as low-power consumption and long-range communication, new network design should be implemented.

4 Network Design and Setup

In this section, focus is on the network design and setup. The purpose of developing sensor network design is to monitor particulate matter and their ranges in the specified area. This network design is working based on software and hardware setups. Network design is shown in Fig. 2, where the LN_1, LN_2, LN_3 and LN_4 are considered as LoRa nodes on which three sensors S_1 (CO_2), S_2 (PM sensor) and S_3 (dust and temperature) placed outside. Before the network deployment, the network design was simulated to analyse the network performance.

4.1 Hardware Required

SPS30 The SPS30 particulate matter (PM) sensor is compact, high quality, optical particle sensor that uses the laser scattering and innovation contamination resistance technology to achieve particle matter measurement [21].

For the research and monitoring of the pollutant level, a previously dust sensor was used instead of SPS30. Due to the inaccuracies and only the dust level monitoring qualities, the sensor was replaced with SPS30. The sensor has inbuilt high-quality and long-lasting component that enables the precision of measurements and provides the throughput for ten years. The sensor package is shown in Fig. 3 [22]. The SPS30 sensor has a five-pin interface that can communicate with two different interfaces:

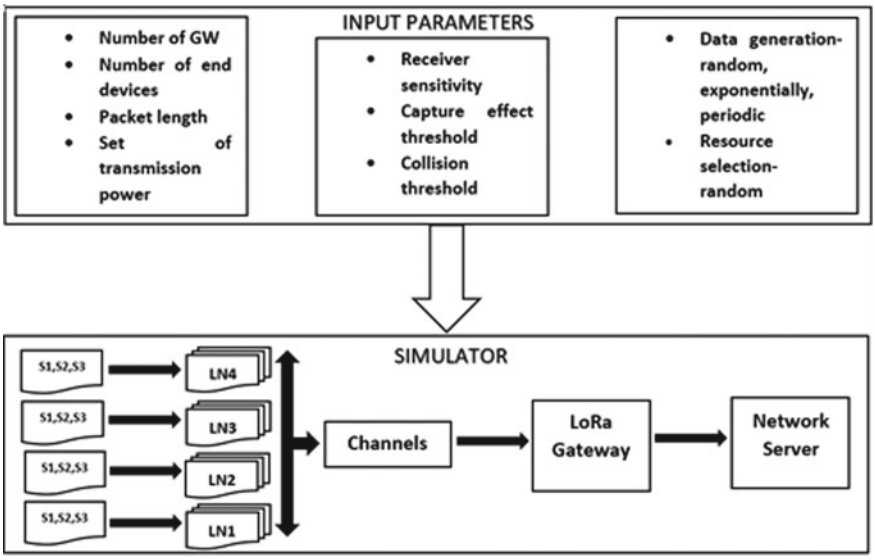


Fig. 2 Block diagram

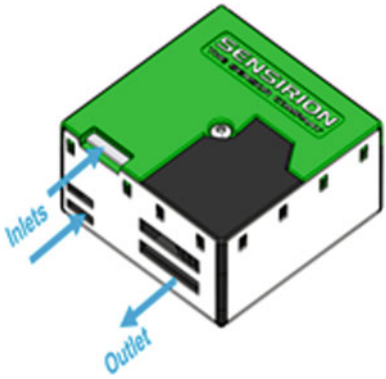


Fig. 3 SPS30 sensor for measuring particulate matter

UART and I2C. The designed network is to monitor the particulate matter. The sensor is working based on UART as per the requirements of the TTGO LoRa32 SX1276 OLED Board.

LoRa32 SX1276 OLED Board The TTGO LoRa32 SX1276 OLED board was found by Semtech, example of LoRa device. LoRa32 board is working for long-range radio devices for transmitting and receiving of data packets over the network. This device leads to less interference and signal loss [21]. Many countries are assigned with specific LoRa bands. LoRa boards have a high communication range as compared to the other networks such as cellular networks. LoRa devices can be used the public and private sectors. LoRa device is shown in Fig. 4.

Fig. 4 TTGO LoRa32 SX1276 OLED

4.2 Software Required

The software for the desired network outcomes is comprised of two parts: the software program coding for the microcontroller board and for the graphical output it requires.

TTN Server The Things Network is the provider for the network server in this research, that means it provides the interface between the gateway and the application. The things network is not providing only the service but also other tools, and these are explained in this section. Before the data transmission to the application, gateway starts receiving messages from the sensor nodes. Similarly, when an application wants to send message to the sensor nodes, it requires to select the broadcast message for gateway. This is the network server, similar to LoRa and LoraWAN allows the sensor nodes to communicate with the gateway. The network server is responsible for routing the data between devices and application. LoRa alliance provided some specifications for the back-end interfaces, for example, controlling the MAC layer, end-point authentication in the network. There is no official release for the LoraWAN network. There are number of solutions available for the LoRaWAN network server such as LoRaServer, open-source network server, ResIOT, and The Things Network developed by The Things Industries. The Things Network is the V2 console for the application deployment and is not accepting any new devices. This V2 console is upgraded with the V3 console.

Ubidots Ubidots is the cloud server that is considered as the webhook for storing and analysing data online. During the transmission, sensor nodes receive data from sensors and received data is shown on the ubidots's dashboard. The collected data can be used for further researches and analysis. The developed application programming interface (API) used ubidots, an IoT platform, and scale projects to production. The ubidots was configured using the representational state transfer (REST) API, which is used to display the data from the sensors stored on the TTN server on a newly designed dashboard, either in text or a graphical form.

5 Results and Discussion

Two LoRa sensor nodes were used to collect data from PM sensors. The pilot study was performed to monitor the PM factors in the air. Four LoRa nodes were deployed with the difference on 5 m apart from gateway. LoRa gateway has the ability to communicate with long-distance transmission. This test was just performed to test the communication range of LoRa and measuring the PM levels in the air. Gateway is registered with free TTN server cloud to which it transmits the collected data from the sensor nodes. The results from one LoRa node (LN_1).

The readings from all LoRa nodes were collected after every 10 min for continuous three days. The dataset collected is plotted for PM_1, $PM_{2.5}$, PM_4 and PM_{10}, shown in Figs. 5, 6, 7 and 8. The comparison among these sensor nodes showed the

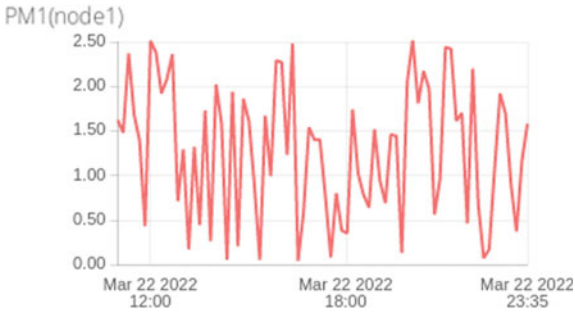


Fig. 5 PM_1 measurements

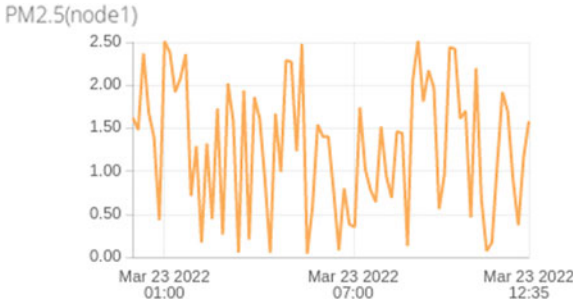


Fig. 6 $PM_{2.5}$ measurements

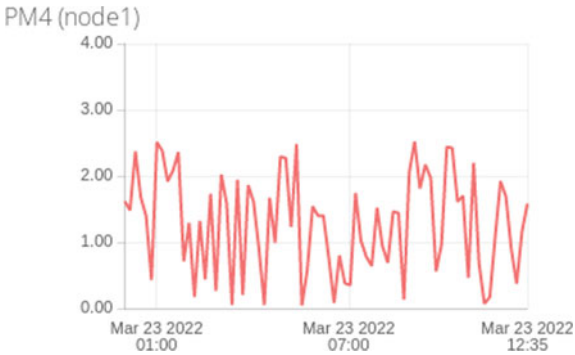


Fig. 7 PM_4 measurements

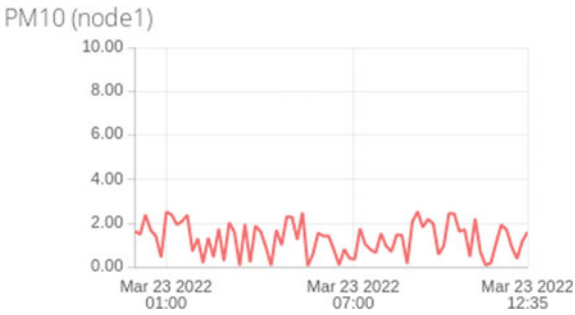


Fig. 8 PM_{10} measurements

excellent correlation between LoRa nodes for measuring the concentration levels. The differences in spread between PM_4, PM_{10} and $PM_{2.5}$ where the PM_1 shows the different concentration levels of both nodes. After few intervals of time, results were taken after 30 min and show the average concentration levels with accuracy of sensors. Concentration levels in the indoor air were not exceeded their limit suggested by the sensors connected.

Particulate measurements were obtained from the SPS30 (Sensiron Sensor) that shows the PM readings of the DustTrak at the sampling level of 1 min. The readings obtained from the SPS30 sensor show some fluctuations. Particulate matter readings are not accurate in this test experiment but these can be improved and shows the good relationship with actual concentration by using the diversity of calibration methods. High variability in PM concentration levels is shown in PM_1. During the monitoring period, measured concentration levels were changed.

6 Conclusion

In summary, number of sensors available in market for monitoring particulate matter and no sensor is ideal for all applications. There are numerous characteristics of SPS30 sensor performance, and choice of this sensor for the network can be based on the critical characteristics. For example, in the network design, the measurements of (PM_1), ($PM_{2.5}$), (PM_4) and (PM_{10}) were required. To monitor all these with low-power consumption, SPS30 sensor was selected that monitor all PM values with single unit. Software and hardware problems have been occurred such as communication range, node placement and time, power management. With the selected devices, 95% time of the transmission was saved using the LoRaWAN network. Different pollutant levels monitor using the time intervals of 1 min. The only down times were due to an outage of The Things Network. Since the sensor nodes were transmitting data every 46 s (the payload of each pollutant level is only 4 bytes). This designed network will be tested with multiple LoRa nodes to check the network performance and accuracy.

References

1. World Health Organization (2016) Ambient air pollution: a global assessment of exposure and burden of disease. Report 2(5)
2. Guanochanga B, Cachipuendo R, Fuertes W, Salvador S, Benítez D, Toulkeridis T, Meneses F (2018) Real-time air pollution monitoring systems using wireless sensor networks connected in a cloud-computing, wrapped up web services. In: Proceedings of the future technologies conference, Nov 2018. Springer, Cham, pp 171–184
3. Bhowmik T, Bhattacharya A, Banerjee I (2019) A low-cost air pollution monitoring system using ZigBee-based wireless sensor networks. In: Information and communication technology for intelligent systems. Springer, Singapore, pp 71–82

4. Rane MS, Naik AR, Vachhani K (2018) Real-time AQI monitoring system: an economical approach using wireless sensor network. In: 2018 9th international conference on computing, communication and networking technologies (ICCCNT), July 2018. IEEE, pp 1–6
5. Mroue H, Nasser A, Hamrioui S, Parrein B, Motta-Cruz E, Rouyer G (2018) MAC layer-based evaluation of IoT technologies: LoRa, SigFox and NB-IoT. In: 2018 IEEE Middle East and North Africa communications conference (MENACOMM), Apr 2018. IEEE, pp 1–5
6. Polichetti G, Cocco S, Spinali A, Trimarco V, Nunziata A (2009) Effects of particulate matter (PM10, PM2.5 and PM1) on the cardiovascular system. Toxicology 261(1–2):1–8
7. Ballester F, Rodriguez P, Iniguez C, Saez M, Daponte A, Galan I, Toro S (2006) Air pollution and cardiovascular admissions association in Spain: results within the EMECAS project. J Epidemiol Community Health 60(4):328–336
8. Zanobetti A, Schwartz, J (2005) The effect of particulate air pollution on emergency admissions for myocardial infarction: a multicity case-crossover analysis. Environ Health Perspect 113(8):978–982
9. Maheswaran R, Haining RP, Brindley P, Law J, Pearson T, Fryers PR, Campbell MJ (2005) Outdoor air pollution and stroke in Sheffield, United Kingdom: a small-area level geographical study. Stroke 36(2):239–243
10. Rodrıguez S, Querol X, Alastuey A, Viana MM, Alarcon M, Mantilla E, Ruiz CR (2004) Comparative PM10-PM2.5 source contribution study at rural, urban and industrial sites during PM episodes in Eastern Spain. Sci Total Environ 328(1–3):95–113
11. Auckland Council (2014) An air quality index for Auckland: review of some international techniques and example data
12. Dhingra S, Madda RB, Gandomi AH, Patan R, Daneshmand M (2019) Internet of Things mobile-air pollution monitoring system (IoT-Mobair). IEEE Internet of Things J 6(3):5577–5584
13. Zheng H, Xiong K, Fan P, Zhong Z (2019) Data analysis on outdoor-indoor air quality variation: buildings' producing dynamic filter effects. IEEE Syst J 13(4):4386–4397
14. Liu JH, Chen YF, Lin TS, Chen CP, Chen PT, Wen TH, Jiang JA (2012) An air quality monitoring system for urban areas based on the technology of wireless sensor networks. Int J Smart Sens Intell Syst 5(1)
15. Chen LJ, Ho YH, Lee HC, Wu HC, Liu HM, Hsieh HH, Lung SCC (2017) An open framework for participatory PM2.5 monitoring in smart cities. IEEE Access 5:14441–14454
16. Badura M, Batog P, Drzeniecka-Osiadacz A, Modzel P (2018) Evaluation of low-cost sensors for ambient PM2.5 monitoring. J Sens 2018
17. Chen G, Guang J, Xue Y, Li Y, Che Y, Gong S (2018) A physically based $PM_{2.5}$ estimation method using AERONET data in Beijing area. IEEE J Sel Top Appl Earth Obs Remote Sens 11(6):1957–1965
18. Jelicic V, Magno M, Brunelli D, Paci G, Benini L (2012) Context-adaptive multimodal wireless sensor network for energy-efficient gas monitoring. IEEE Sens J 13(1):328–338
19. Agarwal L, Dixit G, Jain AK, Pandey KK, Khare A (2016) Energy efficient pollution monitoring system using deterministic wireless sensor networks. In: Proceedings of the international congress on information and communication technology. Springer, Singapore, pp 301–309
20. Knoll M, Breitegger P, Bergmann A (2018) Low-power wide-area technologies as building block for smart sensors in air quality measurements. e i Elektrotech Inf 135(6):416–422
21. Sparkfun (2020) Particulate matter sensor
22. Sensirion (2020) Particulate matter sensor for air quality monitoring and control

Convolutional Neural Network-Enabling Speech Command Recognition

Ankita Patra, Chanki Pandey, Karthikeyan Palaniappan, and Prabira Kumar Sethy

Abstract The speech command recognition system based on deep image classification is the key that would tremendously promise to revolutionize research and development by overcoming the communication barrier between human and machine or computer. We are all aware of challenges in identifying the voice command in noise and variability in speed, pitch, and projection. This paper has developed an efficient and highly accurate speech command recognition for smart and effective speech processing applications like modern telecommunication. In particular, a novel convolutional neural network (CNN) is presented that works with a one-second audio clip consisting of one specific word including ten speech commands and other words labeled as "unknown," and model implementations were operated in the noisy environment. The CNNs are structurally fully developed in such a way to recognize the speech commands with the utilization of deep learning (DL) for image classification concepts. Thus, this research used the concept of DL for image classification to translate the problem of speech command recognition into the image domain.

Keywords Speech command recognition · Image classification · Spectrogram · CNN · Deep learning

A. Patra · P. K. Sethy (✉)
Department of Electronics, Sambalpur University, Odisha, India
e-mail: prabirsethy.05@gmail.com

C. Pandey
National Institute of Technology, Surathkal, Karnataka, India

K. Palaniappan
Centre for System Design, Department of CSE, Chennai Institute of Technology, Chennai, Tamil Nadu, India
e-mail: karthikeyanp@citchennai.net

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_25

1 Introduction

To communicate, people prefer to use speech. Speech signals make natural human interaction easier. Interacting with machines takes effort. Speech command recognition is the key to communicating with machines or computers [1]. Automating your home, dictating messages, and monitoring your security will be easier with this technology. Voice-activated wheelchairs and automated voice-based home navigation systems are just a few examples of equipment and systems that respond to voice commands. Machine learning [2, 3], DL [4–9], neural networks, and IoT [10] are currently being used in practice for tasks such as audio-based face recognition [11], machine translation [12], audio–visual speech recognition [13, 14], speech emotion recognition [15–17], speech recognition [18–20], keyword spotting [21–23], natural language processing [24], agricultural image processing [25–27], medical image processing [28, 29], and ubiquitous computing [30–32] to make smart devices more intelligent for human–machine communication. "There is a requirement for multilingual automatic speech recognition (ASR) and spoken term detection (STD) in contemporary multilingual society [33]." In the study [34], which is based on continuous speech recognition, the hidden Markov model (HMM) was used to detect keywords in conversational speech utterances. A technique based on zone segmentation was recently published in [35], in which the notion of HMM and the pyramid histogram of oriented gradient (PHOG) feature was employed for word spotting in Bangla and Devanagari documents. [36] suggested a multimodal keyword spotting system based on lipreading (LR) and speech features that utilizes a CNN and a long short-term memory (LSTM) network. The study [37] demonstrated the use of a light internal search network (LIS-Net) with parameterized LIS-block and LIS-core for speech command recognition with results that were 24% higher and better.

Keyword spotting has been studied extensively over the past decade using techniques like kernel-based classifiers [38], auto-associative neural networks [39], spectral and prosodic information fusion [40], subspace Gaussian mixture models, written bell, and Kneser–Ney smoothing algorithms [41]. As a result, noise-tolerant speech command recognition is critical, and autonomous feature selection is a major benefit of DL approaches for good human–machine interaction. For example, speech recognition [42] has increased industry expansion and military-based applications. Hand-free voice recognition is possible with speech-enabled devices. A few simple voices control industrial equipment orders including stop, go, up, down, on, and off. So, voice or speech command systems with additional terms have grown popular. But, noise has been an issue. As stated previously, "Noise is prevalent in almost all acoustic situations, causing severe harm to the human-to-machine communication system." [43]

In this study, we focus on single-word speech command recognition using deep image classification, where the end goal is to accurately recognize a set of pre-defined words such as "yes," "no," "up," and "down" from short audio clips in background noise such as pink noise, white noise, bike service, running tap, dude miaowing, and doing dishes. The dataset used in this research is taken from [44], which is the set

of WAVE files, which are about one-second long at a sample rate of 16,000 Hz was used to sample each file into a vector. The noise sound is 1 min to 1.35 min; about, 0.2 fractions are taken and added in command speech. Adding noise randomly to training, validation, and test, the duration of noise is more than the speech that means speech is in a noisy environment. Therefore, to check the robustness of the network, the noise differs in training and testing.

The CNN-based image classification already resulted in high accuracy and minimum error for classification tasks in the current research. So, again, the adaptability of CNN in various areas of research is booming. The uniqueness of the DL approach's adaptation in our study doesn't require the feature extraction and selection process from the.wav file (speech command). But, once it is converted to its image domain (spectrogram), the features (deep features/abstract of features) are automatically extracted and fed to its upper layer, and finally, classification is done.

2 Implementation Specification

The final implementation, which generally includes data collection, spectrogram extraction, network design, and training setup, is carried out to recognize speech commands. The basic topology of the performed experiment is presented in Fig. 1.

Here, the speech command is recognized based on the concept of DL for image classification. In this research, the speech commands (.wav file) are converted into their spectrogram. These spectrograms are treated as training and testing images, which are classified using the concept of DL for image classification.

2.1 Data Collection

We have used the [45] dataset, a large public dataset and freely available for this study. "While there are different possible words, but, the dataset we used for the validation comprises of 31,875 one-second long (or less).wav format file utterances of 10 short words including 'yes', 'no', 'up', 'down', 'left', 'right', 'on', 'off', 'stop', 'go', and other words are labeled as "unknown" [45]." Finally, the distribution of samples is made, as shown in Table 1.

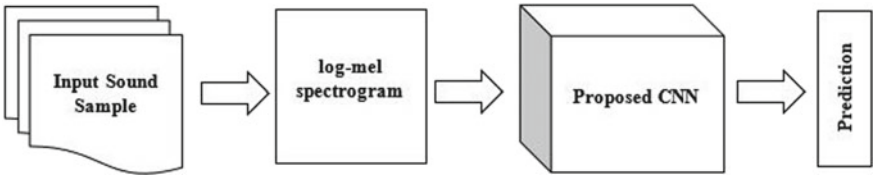


Fig. 1 Basic topology

Table 1 Distribution of speech command samples

Speech command	yes	no	up	down	left	right	on	off	stop	go	unknown	Subtotal
Training and validation set	2159	2172	2153	2175	2157	2167	2167	2180	2175	2177	7993	29,675
Test set	200	200	200	200	200	200	200	200	200	200	200	2200
Total	2359	2372	2353	2375	2357	2367	2367	2380	2375	2377	8193	31,875

Randomly, the training and validation samples are split up about 80:20. Also, 2200 samples, including 200 of each speech command, are kept separate for testing. This way, we can see how well the applications work. This means that the raw audio data have a sample rate of 16,000 Hz and are one-second (or less) long, which means there are 16,000 vectors in it.

2.2 Spectrograms Extraction

For the success of any DL and neural network model, feature extraction plays an important role. Specifically, we have used spectral representation, i.e., log-mel spectrogram, to prepare the data for efficient CNN training. To compute speech spectrogram by converting the speech waveforms to log-mel spectrograms and train the CNN with specific parameters, i.e., segment Duration = 1, frame Duration = 0.025, hop Duration = 0.010, num Bands = 40 are used.

After having chosen fixed time-scale intervals from the input sound spectrograms, a scaling function is applied to them; a series of non-zero spectra are constructed. Then, discrete Fourier transform of each subsequence is generated to form an image [46].

2.3 Addition of Noise

We have also tried different hyperparameters and adding noise to the audio signal before making the spectrogram. This is very important for real-world applications because the words you want to say will never be completely spoken without other distractions. A pure white noise from the thermal voltage of the analog microphone amplifier will always be in the background, but there will also be noise from other sources, like the unorganized crashes of nearby renovations or a highly organized distraction that isn't important. Our sample dataset had audio interference from a variety of noise sources, such as a dishwasher, mixed in at random to make it look like background noise [18]. We tried this at different noise-to-signal ratios, from 0 to 0.5, to see how it worked (corresponding to signal–noise ratios of infinity up to 6 dB). But, the effect was very small: only, a few percentage points of accuracy were lost because of more background noise. This means that this method isn't very sensitive to noise, but low signal-to-noise ratios have a big impact on even the accuracy of humans when they do the same thing.

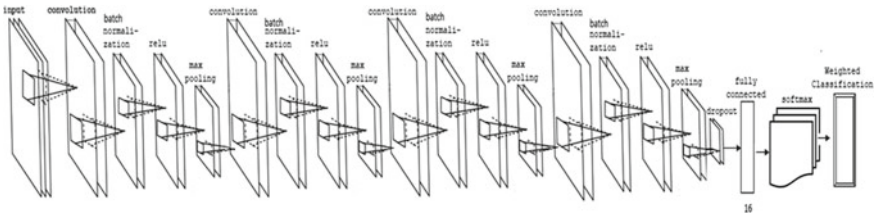


Fig. 2 Network structure of the proposed CNN model

2.4 Network Design

The final network design is very much based on the classic VGGNet [36]. We are going to talk about a kind of multi-stage CNN-based model that has four blocks. Each block has convolutional and batch normalization layers, a ReLU activation function, and max-pooling layers that downsample feature maps "spatially" (that is, in time and frequency) using a kernel size of 3. The final max-pooling keeps the input feature map in one place at a time. This makes the input spectrograms time-translation-invariant, so the network can do the same classification no matter where the speech is in time. If you want to make it less likely that the network will remember specific features of the training data, add a small amount of dropout to the input to the last fully connected layer. The dropout rate should be 0.2. In Fig. 2, you can see how the proposed CNN would work.

The convolutional layer uses 1D convolutions with a kernel size of 3. For the first block, the number of convolutional layers is 12, after which the number is doubled for every block ending with 96 filters. The process of feature extraction and the mathematics behind it are in [47, 48].

2.5 Model Training

The exploration study was done with the MATLAB 2019a. Everything was done on a laptop, such as an HP Pavilion Core i5 5th generation with a simple NVIDIA GEFORCE GTX 1050Ti. We used Adam's optimizer to train the model. We used a mini-batch size of 128 and trained for 25 epochs. We then reduced the learning rate by a factor of 10 after 20 epochs with other parameters such as training options (Adam), initial learning rate (3e-4), all the number of epochs including max epochs, i.e., 25, a mini-batch size, shuffle, plots, training progress, verbose, false, validation data, i.e., {X validation, Y validation}, validation frequency, learn rate schedule, piecewise, learn rate drop factor, i.e., 0.1, learn rate drop period, i.e., 20, validation, and test accuracy was wisely recorded.

3 Results and Discussion

Training, validation, and test sets are each made up of a different part of the dataset. This way, the proposed method's performance can be judged. Randomly, the training and validation samples are split up about 80:20. Also, 2200 samples, including 200 of each speech command, are kept separate for testing. This way, we can see how well the applications work. Accuracy and loss are reported for the validation data and the data that are used to test it for accuracy. Model: 95.68 percent accuracy and 0.028 loss, respectively, are achieved by the proposed model during validation testing. The proposed model also learns at 3e-05 per second on the validation set, which means it learns very quickly. There is a Table 2 that shows how well the proposed model does on a test set, and it shows that it gets 93.71% accuracy.

The results show that the model correctly predicts a high percentage of samples. Again, the performance of the proposed system concerning different noise levels, i.e., noise-to-signal ratio, is recorded in Table 3.

It is noticed from Table 3, with increasing the strength of background noise, the performance of the proposed speech command recognition model degraded, but the effect was relatively very small.

Again, no doubt, many types of research are reported on speech recognition, but these are mainly focused on identifying speech of gender, speaker recognition [49], and emotion recognition. The reports on recognition of speech commands are limited, especially in a noisy environment. The comparison with the state-of-art is illustrated in Table 4.

It was observed from Table 4 that the proposed methodology's performance is competent enough with the state-of-art, even if in a noisy environment. In addition, it not only recognized the ten speech commands but also identified unknown commands. This research is important because it shifts a speech (audio) recognition problem into the field of image classification, where the powerful techniques of CNNs can be fully used.

4 Conclusion

This study presented a model based on deep image classification for speech command recognition to support a CNN. Here, we have used a log-mel spectrogram for representing speech signals. The dataset used in the study was collected from [45], which is freely available. The training and validation samples are split randomly concerning 80:20. Also, 2200 samples, including 200 of each speech command, are kept separately for the test purpose. The proposed model successfully predicted all ten predefined speech commands. Others are labeled as "unknown" with excellent results compared with other existing state-of-the-art techniques. The model's performance is degraded with increasing strength of noise, but it is very less in terms of accuracy percentage. This work mainly focused on speech command (one-second long)

Table 2 Confusion matrix for the test set is based on the proposed CNN model

Speech command	yes	no	up	down	left	right	on	off	stop	go	unknown
yes	0.9602	0.0231		0.0038				0.0038			0.0038
no		0.9556		0.0074	0.0037				0.0037	0.0148	0.0148
up	0.0039		0.9767					0.0078			0.0117
down		0.0455		0.9019	0.0038			0.0114	0.0076	0.0227	0.0038
left	0.0040	0.0162			0.9537	0.0081		0.0040			0.0081
right		0.0039	0.0039		0.0117	0.9600	0.0078				0.0078
on		0.0039	0.0039		0.0079		0.9478	0.0197			0.0118
off			0.0388		0.0078		0.0078	0.9163	0.0078	0.0039	0.0078
stop		0.0041	0.0204		0.0041			0.0041	0.9583	0.0041	
go		0.0315	0.0079	0.0079				0.0039		0.9291	0.0197
unknown	0.0024	0.0168	0.0132	0.0180	0.0144	0.0144	0.0144	0.0120	0.0120	0.0336	0.8487

Table 3 Performance of proposed speech command recognition model in different noise levels

Strength of background noise (noise-to-signal ratio)	Validation accuracy (%)	Test accuracy (%)
0.0 (no noise)	97.80	88.80
0.1	95.68	93.71
0.2	94.5	94.2
0.3	93.8	91.1
0.4	92.1	89.1
0.5	89.2	86.5

Table 4 Comparison of the proposed methodology with state-of-art

References	No. of speech commands	Unknown command identified	Presence of noise	Methodology	Accuracy
[50]	5	No	No	Feature extraction with Mel-frequency Cepstral Coefficient (MFCC) and deep CNN	Validation accuracy—95.1% Test accuracy—94.5%
[51]	10	No	No	CNN	Validation accuracy—96.7% Test accuracy—N/A
[52]	10	Yes	No	CNN	Validation accuracy—97.5% Test accuracy—87.8%
Proposed methodology	10	Yes	Yes	DL for image classification	Validation accuracy—95.68% Test accuracy—93.71%
			NO		Validation accuracy—97.8% Test accuracy—93.8%

recognition, so the proposed model can be further improved for long (more than one second) speech by determining the few more optimal feature. As for long speech commands, there is a possibility of a large variation of the frequency component. So, the number of features is increased. And with consideration of these large sets of features, the computational complexity increases. To address this issue, optimal feature selection is required to transform a large number of correlated variables into a smaller set of composite variables, which is the future scope of this research.

References

1. Shahnawazuddin S, Maity K, Pradhan G (2019) Improving the performance of keyword spotting system for children's speech through prosody modification. Digit Signal Process Rev J 86:11–18. https://doi.org/10.1016/j.dsp.2018.12.011
2. Ma Z, Li B (2020) A DDoS attack detection method based on SVM and K-nearest neighbor in SDN environment. Int J Comput Sci Eng 23(3):224–234
3. Behera SK, Rath AK, Sethy PK (2020) Maturity status classification of papaya fruits based on machine learning and transfer learning approach. Inf Process Agricult
4. Tan Z et al (2020) The clothing image classification algorithm based on the improved Xception model. Int J Comput Sci Eng 23(3):214–223
5. Gu F et al (2020) Face spoof detection using feature map superposition and CNN. Int J Computat Sci Eng 22(2–3):355–363
6. Too EC et al (2020) Performance analysis of nonlinear activation function in convolution neural network for image classification. Int J Computat Sci Eng 21(4):522–535
7. Sethy PK et al (2020) Deep feature based rice leaf disease identification using support vector machine. Comput Electron Agricult 175:105527
8. Behera SK, Rath AK, Sethy PK (2020) Fruit recognition using support vector machine based on deep features. Karbala Int J Modern Sci 6(2):16
9. Sethy PK et al (2020) Nitrogen deficiency prediction of rice crop based on convolutional neural network. J Amb Intell Humaniz Comput
10. Mishra AK, Sinha M, Tripathy AK (2020) A sinkhole prevention mechanism for RPL in IoT. Int J Computat Sci Eng 23(3):262–270
11. Taskiran M, Kahraman N, Erdem CE (2020) Face recognition: past, present and future (a review). Digital Sig Process 106:102809. https://doi.org/10.1016/j.dsp.2020.102809
12. Costa-jussà MR, Allauzen A, Barrault L, Cho K, Schwenk H (2017) Introduction to the special issue on deep learning approaches for machine translation. Comput Speech Lang 46:367–373. https://doi.org/10.1016/j.csl.2017.03.001
13. Noda K, Yamaguchi Y, Nakadai K, Okuno HG, Ogata T (2015) Audio-visual speech recognition using deep learning. Appl Intell 42(4):722–737. https://doi.org/10.1007/s10489-014-0629-7
14. Alam M, Samad MD, Vidyaratne L, Glandon A, Iftekharuddin KM (2020) Survey on deep neural networks in speech and vision systems. Neurocomputing. https://doi.org/10.1016/j.neucom.2020.07.053
15. Zhao J, Mao X, Chen L (2019) Speech emotion recognition using deep 1D & 2D CNN LSTM networks. Biomed Signal Process Control 47:312–323. https://doi.org/10.1016/j.bspc.2018.08.035
16. Issa D, Fatih Demirci M, Yazici A (2020) Speech emotion recognition with deep convolutional neural networks. Biomed Signal Process Control 59:101894. https://doi.org/10.1016/j.bspc.2020.101894
17. Badshah AM et al (2019) Deep features-based speech emotion recognition for smart affective services. Multimedia Tools Appl 78(5):5571–5589. https://doi.org/10.1007/s11042-017-5292-7
18. Gouda SK, Kanetkar S, Harrison D, Warmuth MK (2020) Speech recognition: keyword spotting through image recognition. Accessed: Aug 14, 2020. [Online]. Available: http://arxiv.org/abs/1803.03759
19. Kubanek M, Bobulski J, Kulawik J (2019) A method of speech coding for speech recognition using a convolutional neural network. Symmetry 11(9):1185. https://doi.org/10.3390/sym11091185
20. Jansson P (2018) Single-word speech recognition with convolutional neural networks on raw waveforms
21. Liu B et al (2019) EERA-KWS: a 163 TOPS/W always-on keyword spotting accelerator in 28nm CMOS using binary weight network and precision self-adaptive approximate computing. IEEE Access 7:82453–82465. https://doi.org/10.1109/ACCESS.2019.2924340

22. Li L, Wang J, Li J, Ma Q, Wei J (2019) Relation classification via keyword-attentive sentence mechanism and synthetic stimulation loss. IEEE/ACM Trans Audio Speech Lang Process 27(9):1392–1404. https://doi.org/10.1109/TASLP.2019.2921726
23. Liu B, Qin H, Gong Y, Ge W, Xia M, Shi L (2018) EERA-ASR: an energy-efficient reconfigurable architecture for automatic speech recognition with hybrid DNN and approximate computing. IEEE Access 6:52227–52237. https://doi.org/10.1109/ACCESS.2018.2870273
24. Shrestha H, Dhasarathan C, Munisamy S et al (2020) Natural language processing based sentimental analysis of Hindi (SAH) script an optimization approach. Int J Speech Technol. https://doi.org/10.1007/s10772-020-09730-x
25. Oliveira MM, Cerqueira BV, Barbon S Jr, Barbin DF (2021) Classification of fermented cocoa beans (cut test) using computer vision. J Food Compos Anal 1(97):103771
26. Behera SK, Rath AK, Sethy PK (2021) Fruits yield estimation using faster R-CNN with MIoU. Multimed Tools Appl. https://doi.org/10.1007/s11042-021-10704-7
27. Behera SK, Rath AK, Sethy AK (2020) Maturity status classification of papaya fruits based on machine learning and transfer learning approach. Inf Process Agri. https://doi.org/10.1016/j.inpa.2020.05.003
28. Sethy PK, Pandey C, Khan D, Rafique M, Behera SK, Vijaykumar K, Panigrahi D (2021) A cost-effective computer-vision based breast cancer diagnosis. J Intell Fuzzy Syst. https://doi.org/10.3233/JIFS-189848
29. Sethy PK, Behera SK, A Komma, Pandey C, Khan MR (2020) Computer aid screening of COVID-19 using X-ray and CT scan images: an inner comparison. J X-Ray Sci Technol 1–14. https://doi.org/10.3233/XST-200784
30. Chandramohan D, Sathian D, Rajaguru D, Vengattaraman T, Dhavachelvan P. A multi-agent approach: to preserve user information privacy for a pervasive & ubiquitous environment. Egyptian Inf J 16:151–166. https://doi.org/10.1016/j.eij.2015.02.002. ISSN: 1110–8665
31. Chandramohan D, Vengattaraman T, Rajaguru D, Baskaran R, Dhavachelvan P. EMPPC-an evolutionary model based privacy preserving technique for cloud digital data storage service. In: 3rd IEEE international advance computing conference, India, pp 89–95. ISBN: 978-1-4673-4528-6
32. Chandramohan D, Vengattaraman T, Rajaguru D, Baskaran R, Dhavachelvan P. A privacy breach preventing and mitigation methodology for cloud service data storage. In: 3rd IEEE international advance computing conference, India, pp 83–88. ISBN: 978-1-4673-4528-6
33. Deekshitha G, Mary L (2020) Multilingual spoken term detection: a review. Int J Speech Technol 1–15. https://doi.org/10.1007/s10772-020-09732-9
34. Rose RC (1995) Keyword detection in conversational speech utterances using hidden Markov model-based continuous speech recognition. Comput Speech Lang 9(4):309–333. https://doi.org/10.1006/csla.1995.0015
35. Bhunia AK, Roy PP, Sain A, Pal U (2020) Zone-based keyword spotting in Bangla and Devanagari documents. Multimedia Tools Appl 1–25. https://doi.org/10.1007/s11042-019-08442-y
36. Handa A, Agarwal R, Kohli N (2020) A multimodel keyword spotting system based on lip movement and speech features. Multimedia Tools Appl 79(27):20461–20481. https://doi.org/10.1007/s11042-020-08837-2
37. Anh NT, Hu Y, He Q, Linh TTN, Dung HTK, Guang C (2021) LIS-Net: an end-to-end light interior search network for speech command recognition. Comput Speech Lang 65:101131. https://doi.org/10.1016/j.csl.2020.101131
38. Tabibian S, Akbari A, Nasersharif B (2014) Extension of a kernel-based classifier for discriminative spoken keyword spotting. Neural Process Lett 39(2):195–218. https://doi.org/10.1007/s11063-013-9299-4
39. Jothilakshmi S (2014) Spoken keyword detection using auto-associative neural networks. Int J Speech Technol 17(1):83–89. https://doi.org/10.1007/s10772-013-9208-2
40. Pandey L, Hegde RM (2019) Keyword spotting in continuous speech using spectral and prosodic information fusion. Circuits Syst Signal Process 38(6):2767–2791. https://doi.org/10.1007/s00034-018-0990-6

41. Pala M, Parayitam L, Appala V (2019) Real-time transcription, keyword spotting, archival and retrieval for Telugu TV news using ASR. Int J Speech Technol 22(2):433–439. https://doi.org/10.1007/s10772-019-09598-6
42. Tabibian S (2020) A survey on structured discriminative spoken keyword spotting. Artif Intell Rev 53(4):2483–2520. https://doi.org/10.1007/s10462-019-09739-y
43. Shrawankar U, Thakare V (2010) Noise estimation and noise removal techniques for speech recognition in adverse environments. In: IFIP AICT, vol 340. Springer, Berlin, Heidelberg, pp 336–342
44. Warden P (2017) Speech commands: a public dataset for single-word speech recognition. Available from https://storage.googleapis.com/download.tensorflow.org/data/speech_commands_v0.01.tar.gz
45. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In: 3rd international conference on learning representations, pp 1–14
46. Gouda SK, Kanetkar S, Harrison D, Warmuth MK (2018) Speech recognition: keyword spotting through image recognition. arXiv:1803.03759
47. Wiatowski T, Bolcskei H (2018) A mathematical theory of deep convolutional neural networks for feature extraction. IEEE Trans Inf Theory 64(3):1845–1866. https://doi.org/10.1109/tit.2017.2776228
48. Sethy PK, Barpanda NK, Rath AK, Behera SK (2020) Deep feature based rice leaf disease identification using support vector machine. Comput Electron Agricult 175:105527. https://doi.org/10.1016/j.compag.2020.105527
49. Abdel-Hamid O, Mohamed A-R, Jiang H, Deng L, Penn G, Yu D (2014) Convolutional neural networks for speech recognition. IEEE/ACM Trans Audio Speech Lang Process 22(10):1533–1545. https://doi.org/10.1109/TASLP.2014.2339736
50. Li X, Zhou Z (2017) Speech command recognition with convolutional neural network. CS229 Stanford education
51. Jeoung JE, Yeow YK, Ahemad MM (2019) Keyword spotting on embedded system with deep learning. In: Proceedings of 2019 electrical engineering symposium, vol 3, pp 87–91
52. Patrick J (2018) Single-word speech recognition with convolutional neural networks on raw waveforms. Degree thesis, Arcada University of Applied Sciences, Helsinki, Finland

The Importance of Physical Health Maintenance Applications During the COVID-19 Pandemic in Indonesia

Andhika Satria Putra, Ivander Adrian Djaya, Timothy Alexandro, Winston Amadeus Tandri, Ford Lumban Gaol, Tokuro Matsuo, and Chew Fong Peng

Abstract In this pandemic situation, there are many problems facing people from all over the world. For example, economic downgrades, family problems, illness, and health problems, etc. Indonesia is one of many countries facing this situation. With a record of being the fourth most populous country in the world, Indonesia easily creates a lot of problems in this pandemic situation. This pandemic situation has made the world aware of such problems in their lives. One of the important types of problems that arise from this awareness is the maintenance of physical health. Maintenance of physical health is very important for this pandemic situation. COVID-19 can really cause many diseases, and therefore, health maintenance is very important to get rid of this virus. Many media inform the public to take care of their physical health, but there are also many things that need to be considered and continue life as it is. Therefore, information explaining the importance of maintaining physical health is needed by people in Indonesia, and this study tries to review the

A. S. Putra (✉) · I. A. Djaya · T. Alexandro
Alam Sutera School of Information System, Bina Nusantara University, Jakarta, Indonesia
e-mail: andhika.putra007@binus.ac.id

I. A. Djaya
e-mail: ivander.djaya@binus.ac.id

T. Alexandro
e-mail: timothy.alexandro@binus.ac.id

F. L. Gaol
Binus Graduate Program-Doctor of Computer Science, Jakarta, Indonesia
e-mail: fgaol@binus.edu

T. Matsuo
Advanced Institute of Industrial Technology, Tokyo, Japan
e-mail: matsuo@aiit.ac.jp

W. A. Tandri · C. F. Peng
University of Malaya, Kuala Lumpur, Malaysia
e-mail: winston.tandri001@binus.ac.id

C. F. Peng
e-mail: fpchew@um.edu.my

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_26

importance of health maintenance. This data can be shown in of an app, where it is all currently done digitally, in conjunction also with COVID-19 pandemic that has resulted all people to become digitally literate. In addition, there is also the risk of access to meet other people if you have to consult a doctor, so making a physical health maintenance application during the COVID-19 pandemic will be very helpful. Various educations can be given to the public, such as educating the public that taking supplements/vitamins is important in preventing the COVID-19 virus. Consuming minerals and vitamins for endurance is indeed important during the current pandemic, because the body needs these substances to increase endurance. Apart from supplements, vitamin and mineral needs can also be obtained from various types of food. According to dietitian Maxine Smith, eating healthy foods is the best way to keep the immune system in balance, because the immune system relies heavily on nutritious food. This is because healthy foods contain various substances, including vitamins and minerals as immune support.

Keywords Physical health · Applications · COVID-19 pandemic · Immune

1 Introductions

People in Indonesia are disturbed by a virus called COVID-19, where this virus attacks the respiratory system. This virus can cause mild respiratory problems, lung infections, and even death. In an effort to break the chain of the spread of the coronavirus, the government made a policy by implementing social distancing movements. With this disease, there are many problems such as psychological problems, health to economics.

The current pandemic period prioritizes health being the number one thing that must be maintained, so the government makes students have to take lessons from their respective homes. This is because it is feared that schools will easily spread this virus. Health is very important for humans because by maintaining health, and we can avoid the coronavirus. But there are also many people who think this virus is like an ordinary virus (Kemendikbud). Many people do not wear masks, which means that wearing a mask can reduce coronavirus pollution.

The existence of the coronavirus as it is today makes physical health very important for the people of Indonesia because with our physical weakness, we are easily exposed to this virus. If we do not take care of our physical health, our bodies will feel weak. There are many ways to strengthen our physique, for example, such as exercising where we exercise is very beneficial for our health, and also by exercising one way to fight COVID-19 by exercising provides many benefits for health, either improving physical fitness or maintaining health. From the existence of these coronavirus, there are some questions needed to be answered. Is physical health important for this time?, who plays the important part of maintaining it?, and what kind of foods and vitamins that are good to increase the body's immune system?

Physical health is very important in this pandemic situation. There are several reasons that the physical health is important for the people and the existence of that importance to the people in this pandemic situation.

1.1 Save Money

According to the Centers for Disease Control and Prevention, chronic illnesses act for seven out of every ten deaths in the U.S., while chronic disease treatment acts for 86% of U.S. healthcare costs. While some diseases are unavoidable, you can reduce your risk of developing them, such as heart disease and diabetes, by avoiding risky behaviors and leading a healthy lifestyle.

Trying to make healthy lifestyle choices, such like getting daily exercise, can reduce the risk of developing a variety of health issues and problems, which can result in costly medical care.

1.2 Increase Your Life Expectancy

Various studies have shown that regular physical exercise increases life span and lowers the risk of premature death. Even though there is no magic formula for converting hours of physical activity into extra days of life, data indicates that people who are highly active are healthier and more productive.

1.3 Reduce Your Risk of Injury

Exercising improves muscular strength, bone strength, agility, and stability. Physical fitness, particularly as you age, can reduce your risk of and adaptability to accidental crashes. Stronger muscles and better balance, for example, mean you are less likely to trip and fall, and stronger bones imply you are less likely to get bone injuries if you fall.

1.4 Improve Your Quality of Life

The effects of a sedentary lifestyle and a lack of physical activity can be detrimental to one's health. Physical inactivity has been linked to an increased risk of cancer, a variety of chronic diseases, and mental health problems. Exercise, on the other hand, has been demonstrated to boost mood and mental health as well as providing

a variety of other health advantages. Fitness helps us in many ways and you might not have been allowed to do either.

1.5 Stay Active

Healthy lifestyle enables readers to engage in tasks that require some certain fitness level. Climbing to the summit of a mountain, for example, is a satisfying experience that delivers a sense of success as well as great vistas, but some people are unable to do so due to a lack of physical fitness.

For people who have neglected physical activity for a long time, even a walk around the zoo with their family or playing in the playground with their children might be a problem. If you are active, it is easier to stay active as you get older. It can be inferred that being active does not have to entail traveling to far-flung locations, but can be achieved through little actions such as reading.

1.6 Improve Your Health

There are many health benefits to physical fitness. Workout and physical exercise on a routine basis help to build strong bones and muscles. It helps with breathing, overall health, and general wellness. Staying active can also help you maintain a healthy weight, reduce your risk of type 2 diabetes, heart disease, and reduce your risk of some cancers.

In other words, staying active is an important part of maintaining good health and fitness. So there is no silence and lazy to do activities in this COVID-19 period because with us exercising a lot and moving a lot can make our bodies stronger and increase body immunity. We can also get rid of boredom at home, instead of just sitting there, it is better if we look for a certain activity that can certainly develop ourselves. There are many things that can keep us active during this pandemic.

2 Background and Context

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by the coronavirus, which has become a world health crisis due to its very fast spread [1–6] COVID-19 began to occur in December 2019, this virus outbreak first occurred in the city of Wuhan in the Central Hubei Province of China which is where this virus attacks human health. Physical activity in most of the world's population may have been reduced because as a result of this pandemic, the number of people working from home or not working according to who said that as a result of the act of keeping

a distance, physical activity decreased which made a person anxious, angry, restless, and depressed [7–10].

3 Problem Statement

The threat of the coronavirus that causes COVID-19 in Indonesia makes most people nervous. The reason is clear, the Wuhan coronavirus (SARS-CoV-2) has infected more than 95,000 people globally, and killed more than 3000 people. Detecting everyone with COVID-19 is a global problem. There are, too, many existing respiratory infections, such as colds and flu, occurring in different countries and requiring testing of everyone with coronavirus symptoms. Then, there are still many people who do not understand more deeply about this coronavirus. As a result, our team app includes detailed information about the COVID-19 virus, how else to stop it, and how to manage it for those who have been infected with it. Not only that, we also provide advice on food, vitamins, and any type of exercise that can be done at home.

WHO has previously highlighted lack of funding of mental health: prior to the pandemic, countries were spending less than 2% of their national health budgets on mental health and were struggling to satisfy the requirements of their populations. Furthermore, the pandemic is increasing the demand for mental health services [10–15]. Depression, isolation, financial loss, and afraid all are factors that contribute to or exacerbate mental disorders. Many persons could be abusing alcohol and drugs, as well as suffering from insomnia and anxiety. COVID-19, on the other hand, can cause psychological and mental health effects such like mental illness, irritability, and stroke. People who have a history of mental, neurological, or substance abuse disorders are also much more sensitive to SARS-CoV-2 disease, with a higher risk of suffering outcomes, such as death.

4 Research Questions

a. Is physical health very important during the COVID-19 pandemic?
b. Who plays an important role in maintaining physical health?
c. What foods and vitamins are good for consumption to increase the body's immune system?

5 Relevance and Importance of the Research

The relevance and importance of this research are that there are still a lot of people who are still blind to information about COVID-19, from how to handle it, how to avoid it, the latest variants of COVID-19, to what should be consumed to avoid

this virus. Now that's all related to our research, because we have the idea to create an application that can help the public provide deeper insight into this COVID-19 [16–20].

6 Literature Review

This research is motivated by the low knowledge of the Indonesian people about COVID-19, and there are many factors that cause this to happen, including people who are affected by what other people say that COVID-19 is not real, people who do not have gadgets and live in an insightful perspective. The signal is bad so we do not get news about COVID-19. Overall, we believe that there is only several apps that describe how to manage COVID-19 as well as other up to date news about COVID-19. The suggestion in this study is that it is hoped that all Indonesian people will be more aware of this virus because this virus is real and not a hoax, if all people are aware and aware of this virus, the COVID-19 virus will quickly disappear from Indonesia [21–25].

7 Methods and Sources

The method we use in this study is a qualitative method because it can develop values and draw conclusions based on data, with a process oriented and subjective nature of individual data and value content. We will also conduct interviews to gain a deeper understanding of the topics in this research. Data analysis in qualitative research is interpreted as an effort to systematically search and organize notes from observations, interviews, and others to increase the researcher's understanding of the case under study and present it as findings. To get that understanding, the analysis needs to be continued by trying to find meaning [26–28].

Sequentially, data analysis in qualitative research starts from the stages of data collection, data reduction and categorization, data display, and drawing conclusions. Qualitative data analysis is integrated into the activities of data collection, data reduction, data presentation, and conclusion of research results [29–32].

8 Key Concepts, Theories, and Studies

According to researchers from the WHO, the COVID-19 pandemic has made many of us stay at home and sit longer than usual. It is difficult for many of us to do the kind of exercise we are used to. Even more difficult for people who do not usually do much physical exercise. But at times like these, it is important for people of all ages and abilities to be as active as possible. Just taking a short break from sitting,

by doing 3–4 min of light intensity physical activity, such as walking or stretching, will help relax muscles and increase blood circulation and muscle activity. May lower high blood pressure, help regulate weight and reduce the risk of heart disease, stroke, type 2 diabetes, and various types of cancer—all conditions that can increase susceptibility to COVID-19 [33].

An instant diet is very dangerous during COVID-19 pandemic. Food and drink can affect body ability to recover from infection. While no food or dietary supplement can prevent or cure COVID-19 infection, a healthy diet is important to support the immune system. Good nutrition can also reduce the chances of developing other health problems, including obesity, heart disease, diabetes, and some cancers [34–38].

Apart from eating healthy foods, taking supplements is also important during this COVID-19 pandemic. Consumption of health supplements can help fill gaps in nutrient intake that may not be obtained from fresh food. Taking supplements as an anticipatory plan to maintain immunity during a pandemic is a natural step and is relatively cheaper and easier for most people. Various studies have also shown that taking supplements can increase endurance in groups of people who have difficulty meeting their nutritional needs from food [39, 40].

9 Key Debates and Controversies

Indonesian people's awareness of hygienic living at first glance looks quite well established. Moreover, the flow of information about how to live clean and healthy has been scattered on social media. However, the public is not educated enough. This is shown by data from the Ministry of Health. It is noted that only 20% of the total Indonesian people care about hygiene and health. In addition, based on data from the Directorate General of P2P in 2016 which states that there are still 6.8 million cases of diarrhea in Indonesian society and threatens health and even life safety. This figure can actually still be reduced with a hygienic lifestyle. Chairman of the Association of Indonesian Clinical Microbiologists (PAMKI) DKI Dr. Wani Devita Gunardi, SpMK, said that public education about cleanliness is important. Dangerous bacteria are everywhere, also hygienic houses could be a cause of disease. Therefore, greater attention is needed to educate the public regarding "hygiene issues," he said in his remarks at the opening of IHF, Chubb, Central Jakarta, Monday 23 April 2018. Therefore, based on this, Unilever initiated the establishment of a collaborative forum called the Indonesia Hygiene Forum. In Indonesia, a tier of cleanliness is low, only 20% of the population takes cleanliness seriously. "Indonesia Hygiene Forum (IHF) to be a forum for communication and collaboration between parties involved in health and hygiene, through an educational and scientific approach knowledge," said Hemant Bakshi, President Director of PT, Unilever Indonesia, Tbk. In addition to education, Indonesia also needs hygienic standardization that can be applied in general and comprehensively and can function as an evaluation indicator for the government and the Indonesian people. Prof. of Public Health at the University of

Indonesia, Prof. Purnawan Junaidi stated that it is not only individuals, to make it happen, a special team is needed. “To solve this problem, it is necessary to form a special team that will determine hygienic standards in Indonesia. These standards can be used as evaluation material so that hygiene-related diseases do not easily infect the community.”

10 Gaps in Existing Knowledge

In this study, what needs to be learned is how to create a media, such as an application that can improve the health of everyone at home during this pandemic. This research only takes the scope of society in Indonesia with a sample, namely themselves. Things like articles that can help in this research can also be existing knowledge to apply. In addition, various references that exist may be able to help. However, its application is only in the form of theory and is not made. From this research, the application can motivate readers to create health applications in Indonesia.

11 Practicalities and Potential Obstacles

In this study, things that have the potential to make this research fail are the lack of cooperation with the samples, insufficient tools, and poor member cooperation. Besides that, also lack understanding of the function of the hypothesis, so it is considered that all types of research must be formulated with hypotheses. Making a hypothesis relies solely on perception, so that a hypothesis becomes very speculative. Lack of maturity in making a framework of thought, so that the formulation of the hypothesis has a low rational value. Not supported by established theories or accountable research results. Hypothesis testing in principle is significance testing. Significance itself is the level of error that is obtained/expected when the researcher wants to identify the research sample. Or in other words, the researcher performs an assessment of the population parameters based on the data that has been collected from the parameters of the research sample.

If the results are not significant, then it means that the data collected did not succeed in proving the relationship between X and Y, and it does not mean that X has no effect on Y, but that the sample data did not succeed in proving the relationship.

12 Implication and Contributions to Knowledge

The implications of this research will be used by several people who want to make an application related to the health of the Indonesian people. This research will also make a major contribution to them and students who need knowledge of the

importance of health during the COVID-19 pandemic. Besides that, it also provides insight to people who don't know about COVID-19 so that they understand more about how to deal with COVID, handling during COVID-19 and others.

13 Conclusion

Physical health is defined as the absence of illness or injury. Among the most crucial matters to sustain during the COVID-19 pandemic is its physical health. A good diet, physical well-being, bone caries, personal grooming, and take naps are all examples of physical health. Physical health is critical to one's overall well-being. Chronic disease is a condition health issue that will not resolve itself.

For example, diabetes, asthma, joint disease, or cancer. Chronic diseases are manageable but not curable. There are currently nearly 70,000 children and youth under the age of 18 currently living in Bolton, and this will increase rapidly by 2021. The current pandemic serves as a reminder to all of us to take an active role in our own health care and to be aware of our physical health. Maintain regular contact with your health care team, whether through phone calls, video visits, or face-to-face sessions. Ask your doctor about wearables that can help you track your health problems. Accept responsibility for increasing your understanding of your risks, as well as the prevention and treatment of chronic disease. Share this information with those who live with you or with people you care about to help them improve their physical health. If you have a chronic disease or care for someone who has it, you understand the importance of understanding the dangers, prevention, and treatment of the condition. Physical health can also be expanded by maintaining your circle of support. One approach to achieving this is to enroll in a group workshop to learn how to manage your own health and the health of someone you care about. In a pandemic situation like this, health care is very important for the entire community toward off the COVID-19 virus. Those who are responsible for maintaining health care themselves because physical care is their own responsibility. All kinds of vitamins are essential for their own physical maintenance.

In this case, everyone plays an important role in maintaining their physical health, both the physical health of themselves and those of those closest to them. We ourselves must be aware of how important physical health is, especially in this COVID-19 pandemic when health becomes a valuable thing. If it is not from oneself who cares, then optimal physical health will not be created. We must also remind those closest to us to maintain physical health, such as reminding them to follow strict health protocols and always maintain their immune system. One way to maintain optimal physical health is to consume various vitamins that can increase endurance, such as vitamin C which helps prevent infection and shortens the stay of viruses or bacteria in the body. It is true that taking vitamins will not protect us 100%, but by taking vitamins the probability of our body being attacked by disease is lower.

References

1. World Health Organization (2020) Clinical management of severe acute respiratory infection when novel coronavirus (2019-nCoV) infection is suspected. WHO, Geneva
2. World Health Organization. Novel Coronavirus (COVID-19) Situation Report—25. [Internet]. 2020 [cited 14 February 2020] Available from: https://www.who.int/docs/defaultsource/coronaviruse/situationreports/20200214-sitrep-25-covid19.pdf?sfvrsn=61dda7d_2
3. World Health Organization. Coronavirus disease 2019 (COVID-19) Situation Report—68. [Internet]. 2020 [cited 28 March 2020] Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports
4. World Health Organization. Global surveillance for human infection with novel Coronavirus (2019-nCoV). [Internet]. 2020 [cited 20 March 2020] Available from: https://www.who.int/publications/i/item/globalsurveillance-for-human-infection-with-novelcoronavirus-(COVID-19)
5. World Health Organization (2020) Infection prevention and control during health care when novel coronavirus (Ncov) infection is suspected. [Internet] 2020
6. World Health Organization (2020) Clinical management of severe acute respiratory infection when novel coronavirus (2019-nCoV) infection is suspected. WHO, Geneva
7. Adamson S, Kavaliauskas M, Yamagishi T, Phillips S, Lorimer R, Babraj J (2019) Extremely short duration sprint interval training improves vascular health in older adults. Sport Sci Health 15:123–131. https://doi.org/10.1007/s11332-018-0498-2
8. Affiune A (2002). Cardivascular aging EV Freitas. In: Py L, Néri AL, Cançado FAX, Gorzoni ML, Rocha SM (eds) Geriatrics and gerontology treaty. Guanabara Koogan, Rio de Janeiro, pp 28–32
9. Amatriain-Fernandez S, Gronwald T, Murillo-Rodriguez E, Imperatori C, Solano AF, Latini A et al (2020) Physical exercise potentials against viral diseases like COVID-19 in the elderly. Front Med 7:379. https://doi.org/10.3389/fmed.2020.00379
10. Amatriain-Fernandez S, Murillo-Rodriguez ES, Gronwald T, Machado S, Budde H (2020) Benefits of physical activity and physical exercise in the time of pandemic. Psychol Trauma 12:S264–S266. https://doi.org/10.1037/tra0000643
11. Bangsbo J, Blackwell J, Boraxbekk C-J, Caserotti P, Dela F, Evans AB et al (2019) Copenhagen consensus statement 2019: physical activity and aging. Br J Sports Med 53:856–858. https://doi.org/10.1136/bjsports-2018-100451
12. Bartlett DB, Willis LH, Slentz CA, Hoselton A, Kelly L, Huebner JL et al (2018) Ten weeks of high-intensity interval walk training is associated with reduced disease activity and improved innate immune function in older adults with rheumatoid arthritis: a pilot study. Arthritis Res Ther 20:127. https://doi.org/10.1186/s13075-018-1624-x
13. Bertucci A, Smilenov LB, Turner HC, Amundson SA, Brenner DJ (2016) In vitro rabbit measurement of dose rate effects on radiation induction of micronuclei in human peripheral blood lymphocytes. Radiat Environ Biophys 55:53–59. https://doi.org/10.1007/s00411-015-0628-z
14. Biswas A, Oh PI, Faulkner GE, Bajaj RR, Silver MA, Mitchell MS et al (2015) Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults a systematic review and meta-analysis. Ann Intern Med 162:123–132. https://doi.org/10.7326/M14-1651
15. Blocquiaux S, Gorski T, Van Roie E, Ramaekers M, Van Thienen R, Nielens H et al (2020) The effect of resistance training, detraining and retraining on muscle strength and power, myofibre size, satellite cells, and myonuclei in older men. Exp Gerontol 133:110860. https://doi.org/10.1016/j.exger.2020.110897
16. Borg GAV (1982) Psychophysical bases of perceived exertion. Med Sci Sports Exerc 14:377–381
17. Bowden Davies KA (2018). Physical inactivity and sedentary time: impact on metabolic health and development of type 2 diabetes. The University of Liverpool
18. Brolinson PG, Elliott D (2007) Exercise and the immune system. Clin Sports Med 26:311–319. https://doi.org/10.1016/j.csm.2007.04.011

19. Campbell JP, Turner JE (2018) Debunking the myth of exercise-induced immune suppression: redefining the impact of exercise on immunological health across the lifespan. Front Immunol 9:648. https://doi.org/10.3389/fimmu.2018.00648
20. Carda S, Invernizzi M, Bavikatte G, Bensmaïl D, Bianchi F, Deltombe T et al (2020) The role of physical and rehabilitation medicine in the COVID-19 pandemic: the clinician's view. Ann Phys Rehabil Med 18:S1877–S657. https://doi.org/10.1016/j.rehab.2020.04.001
21. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y et al (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet 395:507–513. https://doi.org/10.1016/S0140-6736(20)30211-7
22. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395:497–506
23. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J et al (2020) Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. Jama 323:1061–1069
24. Chang D, Lin M, Wei L, Xie L, Zhu G, Dela Cruz CS et al (2020) Epidemiologic and clinical characteristics of novel coronavirus infections involving 13 patients outside Wuhan, China. J Am Med Assoc 323:1092–1093
25. Siegel JD, Rhinehart E, Jackson M, Chiarello L (2007) 2007 Guideline for isolation precautions: preventing transmission of infectious agents in health care settings. Am J Infect Control 35:S65–S164
26. Liu T, Hu J, Kang M, Lin L, Zhong H, Xiao J et al (2020) Transmission dynamics of 2019 novel coronavirus (2019-nCoV). bioRxiv. [Preprint] 2020. [cited 14 February 2020] Available from: https://doi.org/10.1101/2020.01.25.919787
27. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H et al (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet 395:565–574
28. Direktorat Jenderal Pencegahan dan Pengendalian Penyakit (P2P) Kementerian Kesehatan RI. Pedoman Pencegahan dan Pengendalian Coronavirus Disease (COVID-19) Revisi ke 3. [Internet] 2020. [cited 14 February 2020] Available from: https://covid19.kemkes.go.id/downloads/#.Xtva kWgzbIU
29. Direktorat Jenderal Pencegahan dan Pengendalian Penyakit (P2P) Kementerian Kesehatan RI. Pedoman Kesiapsiagaan Menghadapi Infeksi Novel Coronavirus (2019- nCoV). Available from: https://covid19.kemkes.go.id/downloads/#.Xtva kWgzbIU [Accessed 30 January 2020]
30. Adhikari SP, Meng S, Wu Y, Mao Y, Ye R, Wang Q et al (2020) A scoping review of 2019 Novel Coronavirus during the early outbreak period: epidemiology, causes, clinical manifestation and diagnosis, prevention and control 1–12
31. Perimpunan Dokter Paru Indonesia (2020) Paduan Praktik Klinis Pneumonia COVID-19 berat tanpa komplikasi (19):7
32. Thevarajan I, Nguyen THO, Koutsakos M, Druce J, Caly L, van de Sandt CE, Jia X, Nicholson S, Catton M, Cowie B, Tong SYC, Lewin SR, Kedzierska K (2020) Breadth of concomitant immune responses prior to patient recovery: a case report of non-severe COVID-19. Nat Med 26:453–455
33. Cai H (2020) Sex difference and smoking predisposition in patients with COVID-19. Lancet Respir Med [Internet] 8(4):e20
34. Diaz JH (2020) Hypothesis: angiotensinconverting enzyme inhibitors and angiotensin receptor blockers may increase the risk of severe COVID-19. J Travel Med 27(3):1–7
35. Heru Wiyono P, Isbaniah F, Handayani D, Soedarsono HJR, Sugiri A, Wahyudi ER, Tri Juli Edi Tarigan RH et al (2020) COVID-19, 1st edn. Perhimpunan Dokter Paru Indonesia, Jakarta
36. Sallard E, Lescure F, Yazdanpanah Y, Mentre F, Pei N (2020) Type 1 interferons as a potential treatment against COVID-19. Antiviral Res 178:104791
37. Ceraolo C, Giorgi FM (2020) Genomic variance of the 2019-nCoV coronavirus. J Med Virol 92:522–528
38. De Salazar PM, Niehus R, Taylor A, Buckee CO, Lipsitch M (2020) Using predicted imports of 2019-nCoV cases to determine locations that may not be identifying all imported cases. [PrePrint] 2020. [cited 14 February 2020]

39. Siegel JD, Rhinehart E, Jackson M, Chiarello L (2007) 2007 Guideline for isolation precautions: preventing transmission of infectious agents in health care settings. Am J Infect Control 35:S65–S164
40. Chang D, Lin M, Wei L, Xie L, Zhu G, Dela Cruz CS et al (2020) Epidemiologic and clinical characteristics of novel coronavirus infections involving 13 patients outside Wuhan, China. J Am Med Assoc

Recommendation System Using Deep Learning with Text Mining Algorithm in E-commerce Framework

K. Deepa, D. Balamurugan, M. Bharath, S. Dineshkumar, and C. Vignesh

Abstract E-commerce has always been thought of as a fast-growing business, and even while online purchasing has not followed the same boom patterns as in the past, it is now being recognized for its potential. One of the current study topics in the field of textual content mining is sentiment evaluation. Digging normal language for conclusions and feelings is a dreary cycle. The best solution is to conduct a sentiment analysis. This gives crucial data to decision production in an assortment of fields. There are a variety of sentiment detection algorithms available, each of which has an impact on the quality of the final result. In this study, we investigate people's feelings toward the services offered by e-commerce websites. Reviews, ratings, and emoticons are among the sentiments. The main purpose is to identify bogus reviews using user and product ids in order to purchase the best products and analyze which one is the best. For this, we deploy a mixed learning system that analyzes numerous service feedbacks. To track down the scores of each word, a text mining calculation is utilized. The sentiments are then divided into three categories: negative, positive, and neutral. It is possible to study a website to figure out whether there are any bogus reviews. Finally, anticipate the bogus reviews left by users. Only paid users can leave reviews, and duplicates are checked using the user id and booking id. Credible client criticism is incorporated when making item suggestions.

Keywords Electronic commerce · Product recommendation · Machine learning · Sentiment analysis · Text mining algorithm

1 Introduction

The Internet has adjusted the most common way of looking for data and formed our shopping conduct behavior. It is no longer just a place of knowledge but offers an interactive platform for any purpose of trading individual involvement in items

K. Deepa (✉) · D. Balamurugan · M. Bharath · S. Dineshkumar · C. Vignesh
Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur, Tamilnadu 639113, India
e-mail: deepak.cse@mkce.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_27

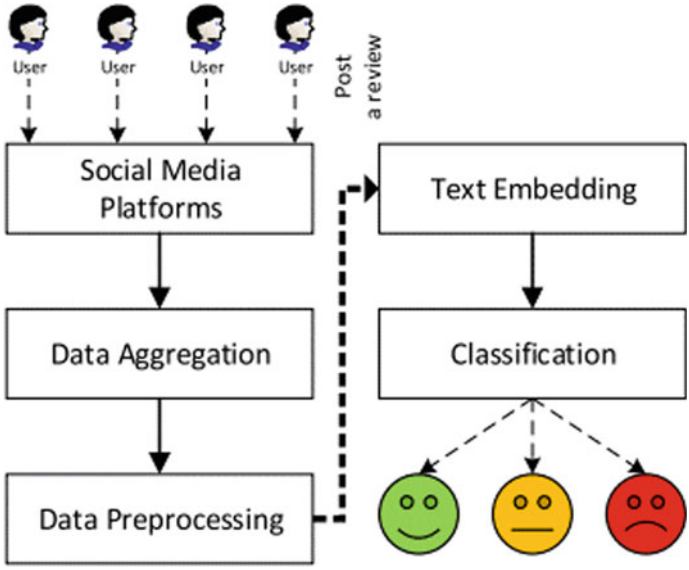


Fig. 1 Sentiment analysis process

or services [1–3]. Through the intelligent idea of the web, buyers are likewise fit for giving data, and in this manner, the data deviation is diminished. Consumer-based information in the shape of online reviews is valuable for potential customers and therefore represents an essential element in the buying decision process [4–6]. Reviews are remarks that represent someone's idea, opinion, or experience with a market product. Users leave product reviews on e-commerce websites to offer suggestions or share their experiences with product providers, merchants, makers, and new buyers [7–10]. By studying the suggestions, the supplied user experience can assist any organization in growing toward improvement. Any product provider can profit or lose money depending on the polarity of evaluations. In another way, new consumers are influenced by reviews while making a buying decision. It may be stated that the consequences of reviews have distinct effects on both businesses and users [11]. So, implement the framework for buying the products based on automated decision-making. Figure 1 shows the process of sentiment analysis.

2 Related Work

Anas et al. [12] created an examination of two models to clarify model execution for this current "Amazon's howl" dataset and their significance for constant programming sending. Hence, the random forest model beats the naive Bayes strategy by a huge degree. The objective is to pick a calculation to attempt crafted by recognizing

and disposing of fake surveys. The matter is tended to genuinely, and a reasonable comprehension of its legitimateness and need is given. Half and half and imaginative models for the bogus survey identification model can be tried later on. The execution technique are implemented to accelerate the results are done using Google co-lab platform. This model is additionally evolved using naive Bayes and discretionary woodlands moves close. These calculations can be utilized to rapidly sort out the number of spam surveys a site or application has. To battle such spammers, a complex model that has been prepared on great many surveys is required. The "Amazon Yelp dataset" is utilized to prepare the models in this review, and it is a genuinely minuscule dataset that might be scaled to accomplish high exactness and adaptability. Limitation is the low of execution process.

Yao et al. [13] the author proposes an ensemble fake review detection model for detecting bogus reviews. The model's four processes are data resampling, feature reduction, parameter optimization, and classifier building. The underlying three stages are proposed for making base classifiers. The performance of the ensemble model improves considerably after optimizing the base classifiers on two sample Yelp datasets with skewed distributions; this presentation is recognizably better compared to that of different normal models and approaches. The examination yielded an assortment of results, both hypothetically and essentially. Because it outperforms other meta-classifiers on the indicated datasets, the Gaussian kernel support vector machine (SVM) classifier was chosen as the meta-classifier in this study. The network search strategy is additionally used to work on the boundaries. We also apply the proposed data resampling approach to improve the meta-classifier to its maximum potential. However, it takes long time when there are too many features.

Elmogy et al. [14] present a machine learning-based response for recognizing fake reviews. This review utilizes an assortment of elements and designing systems, notwithstanding the audit highlights extraction technique to separate different analyst practices. The review checks out the aftereffects of various tests on a genuine Yelp dataset of café surveys, both with and without social elements gathered from customers. We look at the presentation of KNN, naive Bayes (NB), SVM, logistic regression, and random forest classifiers in the two cases. Particular n-gram language models, similar to bi-gram and tri-gram, are associated with the evaluations. The examination utilizes various elements designing strategies on the corpus to extricate unmistakable analyst practices, as well as utilizing regular language handling to concentrate and take care of audit attributes to classifiers. In this review, the impact of the commentators' recuperated characteristics on the classifiers is thought about. The results suggest that the created credits help in the recognizable proof of fake studies. The discoveries show that the made qualities work on the method involved with distinguishing bogus audits. Limitation is difficulty level to consider more behavioral features.

Mohawesh et al. [15] offered a detail evaluation of the most notable machine learning-based fake review detection research to date. To begin, we looked at the many approaches to feature extraction that has been employed by several scholars. From that point forward, we talked about the current datasets and how they were made. We then explained several common machine learning and neural network models for

false review detection using summary tables. Traditional statistical machine learning enhances text classification model performance by improving feature extraction and classifier construction. Deep learning, on the other hand, boosts performance by improving the presentation learning method, algorithm structure, and extra knowledge. We also analyzed numerous deep learning and transformers based on neural network models that have not been used in the detection of fraudulent reviews. According to the results, RoBERTa has the highest efficiency on both datasets. Furthermore, recall, precision, and F1 score were used to illustrate RoBERTa's efficacy in detecting fraudulent reviews. Finally, the current research gaps in this field were identified, as well as potential future directions for achieving reliable results in this domain. However, it is difficult to detect multilingual fake reviews.

Hassan et al. [16] using a dataset of hotel evaluations from online sources, researchers created an effective supervised machine learning strategy for classifying fraudulent online reviews. Counterfeit Internet-based survey location is commonly a paired characterization issue. Purchasing things in online is something we do reliably. When we are considering purchasing something, almost all of us first look it up on websites such as Amazon, AliExpress, eBay, and others. In the case of travel, online service providers can assist with hotel reservations, air ticket purchases, and a variety of other duties. However, because we cannot see the products or services we are buying, we rely on what other people have to say about them. Subsequently, online surveys are fundamental when settling on buying choices. This also provides an opportunity for certain evil actors to deceive consumers by posting phony comments and reviews. They can publish phony reviews to promote their items or to denigrate competitors' offerings. As a result, detecting fraudulent Internet reviews is critical for both users who profit from reviews and businesses who want to keep their goodwill with customers. Using various machine learning approaches, many academics are focusing on auto-detection of bogus reviews. The bulk of studies is based on supervised learning. Many of the researchers use semi-supervised and clustering methodologies. Performance is less when dataset is large size.

Wang et al. [17] multi-feature fusion and rolling collaborative training are combined in a model for detecting fraudulent reviews. Tests have shown that this strategy beats more standard techniques. It expands characterization precision and works on the viability of the grouping framework by utilizing unlabeled information. At the same time, the consistency of opinion and score is examined, and the audit's element extraction is performed utilizing the message portrayal model, which includes combination joined with the message's outside highlights, bringing about a critical improvement in the grouping model's arrangement impact. The main reason for this is that professional fake authors' writing styles will change as a result of an update to the e-commerce platform's detection mechanism, which we will do everything possible to avoid. Because of the detecting process, it is more difficult to be discovered, and appropriate countermeasures can be deployed based on consumer psychology and events such as shopping festivals. And expect that future researchers will continue their work in the dynamic update detection approach area. Limitation is to improve the system which can detect false information in multiple fields.

Liu et al. [18] give a staggered intuitive neural organization model with perspective limitations that mines the multi-facet understood articulation method of audits and consolidates four aspects into survey portrayals, including users, review texts, products, and fine-grained aspects. We model user-product relationships and recast the model's objective function using these relationships as a regularization term. Experiments on three publicly available datasets show that the model we offer outperforms current techniques, demonstrating its efficacy and versatility. To test the MIANA's adequacy, we ran broad tests on three freely accessible datasets. The experiments demonstrated that the MIANA outperforms state-of-the-art systems for false review detection tasks and that the proposed scheme is both successful and practicable. It is difficult to combine user-product information.

Wang et al. [19] there are two sorts of features available: a behavioral feature set and a semantic feature set. Using semantic analysis, we suggest three feature sets. Although behavioral features have good detection accuracy with supervised classifiers, our research indicates that semantic features also function well on real-life bogus reviews in the commercial scenario of Yelp.com. Readability features, topic features, and n-gram features are all similar in terms of performance among the semantic features. The feature vectors were created for n-gram features using the review content's count and Tf-Idf values, and the classification process was performed using SVM and NB classifiers. From overall evaluation, the LR classifier was more accurate that achieves up to 97.2% accuracy with all characteristics. With the introduction of e-commerce, a growing number of consumers are enjoying the convenience of purchasing online and subsequently sharing their experiences on the electronic company website. Customers and merchants may use these online opinions to make purchasing and other decisions. Positive reviews play a simulative role in reaping economic rewards and a well-deserved reputation for merchants' firms when it comes to Internet reviews. As a result, retailers have strong ambitions to manipulate their reputation and hire specialized imposters to post greater ratings on purchasing sites. It is difficult to construct more feature vectors.

Liu et al. [20] the researchers looked at online shopping site review records and presented a new method for detecting false product reviews. By examining the temporal trends of reviews and comments, this review outlier identification approach discovers outlier products. As a result of this viewpoint, our strategy outperforms other ways. Detecting review records is also quick and simple. To conduct the study, we used an Amazon China dataset to examine the characteristics of online reviews. After that, we suggested an isolation forest-based strategy for detecting bogus reviews. To represent the adequacy and productivity of our framework, we contrasted it with different worldly exception recognition techniques. And, while the proposed method can better discover outlier items based on anomalous evolving patterns, the experiment also revealed that the window size for review records selection is crucial. In terms of detecting accuracy, a larger window size will offer better outcomes. We noticed that outlier reviews tend to explode quickly, usually within a few days or weeks. It is difficult to combine multiple feature review datasets.

Hassan et al. [21] this review utilized different semi-directed and managed text mining ways to deal with identify counterfeit web-based audits. To generate a superior

feature set, we joined features from different research studies. We also utilized a different classifier that was not used in the previous study. More seasoned advances are being eliminated for more current, more intricate ones. Individuals may now finish their work all the more rapidly on account of these new advances. The Internet-based commercial center is an illustration of such innovative progression. Online sites permit us to buy and reserve a spot. Nearly everybody understands item or administration surveys prior to making a buy. Therefore, online audits have turned into a significant wellspring of notoriety for organizations. They, moreover, generally influence thing and organization publicizing and headway. Fake Internet reviews have been a major source of concern as online marketplaces have grown in popularity. People can write fraudulent evaluations in order to promote their own items, which can be harmful to actual users. Furthermore, rival businesses can try to smear each other's reputations by posting phony bad reviews. Accuracy is less in unlabeled datasets.

3 Machine Learning Techniques

The sentences which represent observations that are represented as positive or negative are called sentiments. The users who buy products can post their reviews in the e-commerce websites [22–26]. Those reviews can be extracted as unstructured data. The unstructured data is transformed into structured form and extracts features from structured data. The features of the words are selected and then classification technique is applied on extracted features to classify them into its sentiment polarity, that is, namely either positive or negative [27–31]. Representation of feature words which is based on machine learning techniques is discussed as follows.

3.1 Naive Bayesian Classifier

The naive Bayesian classifier depends on Bayes' hypothesis with autonomy suspicions between indicators. A naive Bayesian model is not difficult to work, with no confounded iterative boundary assessment which makes it especially valuable for extremely huge datasets [32–34]. In spite of its effortlessness, the naive Bayesian classifier frequently does shockingly well and is broadly utilized in the light of the fact that it regularly beats more complex characterization strategies [35]. Algorithm Bayes theorem gives an approach to calculate the posterior probability ($P(c|x)$), from $P(c)$, $P(x)$, and $P(x|c)$. Naive Bayes classifier assumes that the impact of the value of a predictor, x on a given class, c is not dependent on the values of other predictors [36]. This assumption is called class conditional independence in Eq. (1).

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$$P(c|x) = P(x_1|c) * P(x_2|c) * \ldots * P(x_n|c) * P(c) \quad (1)$$

where

- $P(c|x)$, posterior probability of class (c, target) given predictor (x, attributes).
- $P(c)$, prior probability of class.
- $P(x|c)$, probability of predictor given class(likelihood).
- $P(x)$, prior probability of predictor

The naive Bayes classification only analyzed positive and negative reviews and provides reduced accuracy in classification.

3.2 Support Vector Machine Algorithm

Support vector machine breaks down the data, characterizes the decision boundaries, and uses the kernels for calculation which is acted in input space. The two sets of vectors are input data of size m each [37–39]. Then, at that point, each datum addressed as a vector is arranged in a specific class. Presently, the task is to observe a margin between two classes that is a long way from any document. The distance characterizes the margin of the classifier, amplifying the margin diminishes indefinite decisions [40, 41]. SVM additionally supports regression and classification which are valuable for measurable learning theory, and it helps perceive the factors precisely that should be considered to comprehend it effectively. It tackles the optimization issue of tracking down the most extreme margin hyperplane between the classes [42–44]. This is basically required to avoid overfitting. Basically, it is a linear classifier separating the classes which can be separated with the help of linear decision surfaces called hyperplanes. For classes having binary features SVM draws a line between the classes, and for classes having multiple features, hyperplanes are drawn. However, it can be used for classifying the nonlinear data also by transforming the feature space into the higher dimensional space so that the nonlinear data in higher dimensional can be isolated effectively by a hyperplane [45–48]. This change is made simple with the assistance of kernel trick. With the assistance of kernels, it is not important to calculate everyone of the dimensions when change and estimation of hyperplane should be possible in a similar lower-dimensional feature space. Kernels are not used only for this purpose but also for making the calculation easier in case of many features [49]. Various kernels are used by machine learning approaches, e.g., radial basis function and poly [50].

4 Automated Recommendation Using Text Mining and Deep Learning

A recommendation system has been carried out in view of the hybrid approach of stochastic learning and context-based engine. We have tried to combine the existing algorithms for recommendation to come up with a hybrid one. It increases the effectiveness by beating the drawbacks of traditional recommendation systems. Recommender frameworks being a piece of data sifting framework are utilized to forecast the ratings the client will in general give for an item. Among various types of recommendation approaches, the collaborative filtering method has an exceptionally high ubiquity as a result of their viability. These traditional collaborative filtering might work actually and can deliver effective recommendations, in any event, for wide running issues. For item based on their neighbor's preferences, entropy-based technique create better suggestions than others.

While different methods like content-based experiences poor precision, adaptability, data scarcity, and large error expectation, to observe these possible outcomes, we have utilized client-based collaborative filtering approach. In this item primarily based collaborative filtering technique, we first take a look at the user item rating matrix and we pick out the relationships among various items, after which we utilize these associations with the expectation to figure the recommendations for the user. Then, the usage of cosine similarity that is a similarity weight is going to play a critical function within the collaborative item primarily based filtering approach and subsequently with a view to preserve or pick the trustable customers from the given set of consumer.

Hence, they deliver us a way to growth or lower the importance of a particular user or item. In the present methodology, we are using adjusted similarity for computation of similar weights of objects. Admin can add the products with specification. Client can look through the items and analyze the sentiment. If the product is positive label means, automatically recommend the products otherwise recommend positive products. After that, user posts the reviews, performs pre-processing, stemming analysis, and extracts the key terms. Finally, the fake reviews are restricted, and the products are recommended with the genuine reviews.

4.1 The Proposed Algorithm Steps as Follows

Input: Labeled Dataset

Output: sentiment polarity with synonym of words and similarity between words

Step-1 Pre-Processing the tweets:

Pre-processing ()

Remove URL:

Remove special symbols.

Convert to lower:

Step-2 Get the Feature Vector List:
For w in words:
Replace two or more words
Strip:
If (w in stopwords)
Continue
Else: Append the file
Return feature vector
Step-3 Extract Features from Feature Vector List:
For word in feature list
Features=word in trained_words
Return features
Step-4 Combining Feature Vector List and Pre-Processing Dataset
Pre-processed file=path name of the file
Stopwords=file path name
Feature Vector List=file path of feature vector list
Step-5 Training the step 4 Apply classifiers classes
Step-6 Finding the Similarity and Synonym of the Feature Vector
For every sentences in feature list (n=3)
Extract feature vector in the tweets ()
For each Feature Vector: x
For each Feature Vector: y
Find the similarity(x, y)
If (similarity>threshold)
Match found
Feature Vector: x= Feature Vector: y
Classify (x, y)
Print: sentiment polarity with similar feature words.

The above pseudocode is utilized to predict the polarity from client datasets. These features can be matched with vector to predict the similarity threshold and labeled as class name.

5 Conclusion

In this research, we portray the execution of recommendation process based on the hybrid recommendation algorithm. Our method's key benefits are a visible grouping of the information dependent on the fundamental structure and a huge size reduction of the search space per result output. Products can be searched by the end-user at any time and any location. Positive and negative attitudes are determined by analyzing ratings, reviews, and emoticons. Price-based filtering and review-based filtering are utilized to track down items. To avoid false reviews, a user id-based screening technique might be utilized. One of the system's primary components, hybrid recommendations, helps to overcome the disadvantages of traditional collaborative and

content-based recommendations. Using our present model, we have gotten some promising findings.

References

1. Murugesan M, Thilagamani S (2020) Efficient anomaly detection in surveillance videos based on multi layer perception recurrent neural network. J Microprocess Microsyst 79
2. Masood F et al (2019) Spammer detection and fake user identification on social networks. IEEE Access 7:68140–68152
3. Hassan R, Rabiul Islam M (2021) Impact of sentiment analysis in fake online review detection. In: International conference on information and communication technology for sustainable development, IEEE
4. Thilagamani S, Nandhakumar C (2020) Implementing green revolution for organic plant forming using KNN-classification technique. Int J Adv Sci Technol 29:1707–1712
5. Danish NM et al (2019) Intelligent interface for fake product review monitoring and removal. In: 16th international conference on electrical engineering, computing science and automatic control. IEEE
6. Thilagamani S, Shanti N (2014) Gaussian and gabor filter approach for object segmentation. J Comput Inf Sci Eng 14(2):021006
7. Abri F et al (2020) Linguistic features for detecting fake reviews. In: 19th IEEE international conference on machine learning and applications. IEEE
8. Perumal P, Suba S (2022) An analysis of a secure communication for healthcare system using wearable devices based on elliptic curve cryptography. J World Rev Sci Technol Sustain Dev 18(1):51–58
9. Ahmed S, Muhammad F (2019) Using boosting approaches to detect spam reviews. In: 1st international conference on advances in science, engineering and robotics technology. IEEE
10. Navastara DA, Alimatus Zaqiyah A, Fatichah C (2019) Opinion spam detection in product reviews using self-training semi-supervised learning approach. In: International conference on advanced mechatronics, intelligent manufacture and industrial automation, IEEE
11. Pandiaraja P, Sharmila S (2020) Optimal routing path for heterogenous vehicular adhoc network. Int J Adv Sci Technol 29(7):1762–1771
12. Anas SM, Kumari S (2021) Opinion mining based fake product review monitoring and removal system. In: 6th International conference on inventive computation technologies, IEEE Explorer
13. Yao J, Zheng Y, Jiang H (2021) An ensemble model for fake online review detection based on data resampling, feature pruning, and parameter optimization. IEEE Access 9:16914–16927
14. Elmogy AM et al (2021) Fake reviews detection using supervised machine learning. Int J Adv Comput Sci Appl 12
15. Mohawesh R et al (2021) Fake reviews detection. IEEE Access 9:65771–65802
16. Hassan R, Rabiul Islam M (2020) A supervised machine learning approach to detect fake online reviews. In: 23rd International conference on computer and information technology (ICCIT). IEEE
17. Wang J et al (2020) Fake review detection based on multiple feature fusion and rolling collaborative training. IEEE Access 8:182625–182639
18. Liu M et al (2020) Detecting fake reviews using multidimensional representations with fine-grained aspects plan. IEEE Access 9:3765–3773
19. Wang X et al (2018) Identification of fake reviews using semantic and behavioral features. In: 4th International conference on information management, IEEE
20. Liu W et al (2019) A method for the detection of fake reviews based on temporal features of reviews and comments. IEEE Eng Manage Rev 47(4):67–79
21. Hassan R, Rabiul Islam M (2019) Detection of fake online reviews using semi-supervised and supervised learning. In: International conference on electrical, computer and communication engineering. IEEE

22. Furia R, Gaikwad K (2020) Tool for review analysis of product. In: International conference on emerging trends in information technology and engineering, IEEE
23. Pandiaraja P, Aravinthan K, Lakshmi NR, Kaaviya KS, Madumithra K (2020) Efficient cloud storage using data partition and time based access control with secure aes encryption technique. Int J Adv Sci Technol 29(7):1698–1706
24. Pillala P et al (2020) Random forest model for intrusion detection in crowd-sourced reviews. In: International conference on smart electronics and communication, IEEE
25. Rajesh Kanna P, Santhi P (2021) Unified deep learning approach for efficient intrusion detection system using integrated spatial–temporal features. Knowl-Based Syst 226
26. Kudo W, Nishiguchi M, Toriumi F (2019) Fraudulent user detection on rating networks based on expanded balance theory and GCNs. In: International conference on advances in social networks analysis and mining, IEEE
27. Santhi P, Mahalakshmi G (2019) Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction. Int J Eng Adv Technol 8(4):839–846
28. Manoj K et al (2018) Genuine ratings for mobile apps with the support of authenticated users' reviews. In: Second international conference on green computing and internet of things, IEEE
29. Deepa K, Thilagamani S (2019) Segmentation techniques for overlapped latent fingerprint matching. Int J Innov Technol Expolring Eng 8(12):1849–1852
30. Raja B et al (2020) Eliminating products' fake reviews using network parameters and geo location. In: IOP conference series: materials science and engineering, vol 981. No. 3. IOP Publishing
31. Pradeep D, Sundar C, QAOC (2020) Noval query analysis and ontology-based clustering for data management in Hadoop, vol 108, pp 849–860
32. Bobadilla J et al (2013) Recommender systems survey. Knowl-Based Syst 46:109–132
33. Logeswaran R, Aarthi P, Dineshkumar M, Lakshitha G, Vikram R (2019) Portable charger for handheld devices using radio frequency. Int J Innov Technol Exploring Eng 8(6):837–839
34. Wang G et al (2014) Sentiment classification: the contribution of ensemble learning. Decis Support Syst 77–93
35. Gunasekar M, Thilagamani S (2020) Performance analysis of ensemble feature selection method under SVM and BMNB classifiers for sentiment analysis. Int J Sci Technol Res 9(2):1536–1540
36. Anand D, Bharadwaj KK (2011) Utilizing various sparsity measures for enhancing accuracy of collaborative recommender systems based on local and global similarities. Expert Syst Appl 38(5):5101–5109
37. Zhao X et al (2010) Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid. ACL
38. Deepika S, Pandiaraja P (2013) Ensuring CIA triad for user data using collaborative filtering mechanism. In: 2013 International conference on information communication and embedded systems, pp 925–928
39. Nguyen TH, Shirai K (2015) Topic modelling based sentiment analysis on social media for stock market prediction. In: Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing, vol 1: Long Papers
40. Bobadilla J et al (2012) A collaborative filtering approach to mitigate the new user cold start problem. Knowl Based Syst 26:225–238
41. Jo Y, Oh AH (2011) Aspect and sentiment unification model for online review analysis. In: fourth ACM international conference on Web search and data mining, pp 815–824. New York, NY, USA
42. Rajesh Kanna P, Santhi P (2022) Hybrid intrusion detection using map reduce based black widow optimized convolutional long short-term memory neural networks. Expert Syst Appl 194
43. Cardie YC (2014) Context-aware learning for sentence-level sentiment analysis with posterior regularization. In: Proceedings of the 52nd annual meeting of the association for computational linguistics, Baltimore, MD, USA, pp 325–335

44. Lin C, He Y, Everson R, Ruger S (2012) Weakly supervised joint sentiment-topic detection from text. IEEE Trans Knowl Data Eng 24:1134–1145
45. Dermouche M, Velcin J, Khouas L, Loudcher S (2014) A joint model for topic-sentiment evolution over time. In: International conference on data mining, pp 773–778, IEEE
46. Deepa K, Kokila M, Nandhini A, Pavethra A, Umadevi M (2020) Rainfall prediction using CNN. Int J Adv Sci Technol 29(7 Special Issue):1623–1627
47. Hu X, Tang L, Tang J, Liu H (2013) Exploiting social relations for sentiment analysis in microblogging. In: Sixth ACM international conference on web search and data mining, pp 537–546. New York, NY, USA
48. Liu S, Cheng X, Li F, Li F (2015) Tasc:topic-adaptive sentiment classification on dynamic tweets. IEEE Trans Knowl Data Eng 27(6):1696–1709
49. Rahman MM, Wang H (2016) Hidden topic sentiment model. In: 25th international conference on World Wide Web. International World Wide Web conferences steering Committee, pp 155–165. Republic and Canton of Geneva, Switzerland
50. Paltoglou G, Thelwall M (2010) A study of information retrieval weighting schemes for sentiment analysis. In: Proceedings of the 48th annual meeting of the association for computational linguistics, association for computational linguistics, pp 1386–1395. Stroudsburg, PA, USA

Hybrid Malicious Encrypted Network Traffic Flow Detection Model

Shivaraj Hublikar and N. Shekar V. Shet

Abstract Encrypted communication technology has evolved as the network, and Internet applications have advanced. Malicious communication, on the other hand, employs encryption to bypass standard detection and security protection. The existing security prevention and detection technologies are unable to identify harmful communication that is encrypted. The growth of artificial intelligence (AI) in these days has enabled to employ machine learning (ML) as well as deep learning approaches to identify encrypted malicious communications without decryption, with remarkably precise detection outcomes. At this moment, research on detecting harmful encrypted traffic is mostly focused on analyzing the features of encrypted data and selecting neural network (NN) techniques. Hybrid ML is proposed in this study by merging two well-performing data mining algorithms with natural language processing tasks. Here, a new traffic flow detection method is performed by the hybrid ML technique. At first, the benchmark data is collected from public sources. The features are extracted using the convolutional layer of deep convolutional neural network (DCNN). Then, the weighted feature extraction is performed by grasshopper optimization algorithm (GOA). Employed the hybrid machine learning-based malicious detection with the "support vector machine (SVM) and neural network (NN)" is utilized in this model to detect the traffic affected by malicious activities, where the hidden neuron count of NN and kernel of SVM are tuning by the same GOA for increasing the accuracy and precision. This research provides findings from experiment, encouraging various researchers to develop the research as future work.

Keywords Malicious encrypted network traffic flow detection framework · Hybrid machine learning · Weighted feature extraction · Support vector machine · Neural network · Grasshopper optimization algorithm · Artificial intelligence

S. Hublikar (✉) · N. S. V. Shet
Department of Electronics and Communication, National Institute of Technology Karnataka, NH 66, Srinivasnagar Surathkal, Mangalore, Karnataka 575025, India
e-mail: shivaraj.hublikar@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_28

1 Introduction

Cyber-attacks are growing more tough and common these days. Various academics presented numerous alternative cybersecurity systems for this reason, and the recommended cybersecurity system is extensively used to secure data from cyber-attacks and unlawful access [1]. To combat the problem of cyber-attacks, experts have developed a variety of cybersecurity technologies. Furthermore, a variety of cybersecurity technologies are provided and used to safeguard sensitive data and prevent illegal access [2]. The relevance of need for cybersecurity has constantly risen as the count of network traffic, and Internet users have expanded.

Network assaults and abnormalities, like "denial of service (DoS) attacks, man-in-the-middle attacks, and malware," may all be protected by cybersecurity [3]. Despite the fact that this type of assault is not new, it nevertheless presents a significant danger to majority of the modern protection systems. The initial step in stopping malicious traffic, particularly DDoS assaults, is to detect anomalies in network traffic as quickly as possible by monitoring "network traffic at gateways, edge servers, or a scrubbing center" [4]. The conventional techniques, like "signature-based and statistical detection systems", contain a number of problems, for example, the rule updating cycle cannot maintain with the increasing number of attack types. The experimental findings, on the other hand, indicated that their suggested method is effective at percolating harmful assaults with a short feedback time and minimal bandwidth usage [5].

ML and artificial intelligence (AI) are powerful and extensively used approaches for the optimal outcomes and appropriate identification. For the identification of cyber-attacks, AI and ML algorithms are commonly used [6]. For traffic management and security monitoring, traffic identification is critical. Due to its accuracy, the ML approach has nowadays gained a lot of attraction and has become quite familiar in a variety of sectors. The ML method's substantial features collection is critical for good identification [7]. Efficient features are appropriate features or qualities that preserve significance information for ML techniques, and they are found in both the training as well as testing sets. As a result, for the assessment of the ML method, a suitable feature set comprising training as well as testing groups is required [8]. The ML approach is broadly used in computer science, particularly in the detection of network traffic. Malicious, incursion, and cyber-attacks can all benefit from ML approaches for recognizing and categorizing them. Although using the ML technique to identify and classify harmful traffic from cyber-attacks is successful, the ML methodology tool is quite difficult in comparison with other computer tools [9, 10]. However, in the domain of identification or categorization, adopting an ML approach is quite beneficial. However, there are significant drawbacks in malicious as well as intrusion detection, such as calculation time and consumption of energy issues. Especially, deep learning method is intended for automatically extracting the features using various neural networks and also understanding the errors that occurred in that network. The neural network models are comprised of different hierarchies that are denoted as the layer-by-layer abstraction. These layers are positioned according to

the nonlinear activation functions. It is also used for extracting complex features and learning the hidden deep features among the texts. Therefore, most of the studies use deep learning for detection models.

The paper contribution is

- To perform a new traffic flow detection model by the hybrid ML technique.
- To do the weighted feature extraction by the optimization concept, where the weight is tuned by the GOA with the intention of variance maximization.
- To perform the hybrid ML-based malicious detection by SVM and NN, in which the hidden neuron count of NN and kernel of SVM are optimized by the same GOA with the consideration of accuracy and precision maximization.

The paper organization is: Section 1 gives the introduction of malicious network traffic flow models. The literature works are displayed in Sect. 2. Section 3 provides the dataset description and proposed architecture considered for malicious Internet protocol network flow detection. The weighted feature extraction and detection model for network malicious identification are displayed in Sect. 4. The results and discussions are shown in Sect. 5. The conclusion is in Sect. 6.

2 Literature Survey

2.1 Related Works

In 2021, Shafiq et al. [11] have presented a novel framework method. Initially, a novel method referred as CorrAUC was suggested. Furthermore, using a "bijective soft set and an integrated TOPSIS and Shannon entropy," chosen characteristics were evaluated. With the help of the Bot-IoT data collection and four distinct ML methods, the suggested strategy was tested. The examination of the experimental findings revealed that the suggested strategy was effective, with average outcomes of >96%. In 2021, Kim et al. [12] have suggested a less invasive traffic sample technique by a "deep deterministic policy gradient (DDPG)," a typical "deep reinforcement learning (DRL) method." It was shown that the suggested strategy contained a high likelihood of detecting malicious flows.

In 2020, Hwang et al. [13] have proposed D-PACK, a traffic anomaly detection system for "auto-profiling traffic patterns and filtering aberrant traffic." D-PACK also operated with approximately 100 percent accuracy when just looking at the first two packets in every flow, according to the data. The architecture might serve as a model that reduced the amount of evaluated packets while also preventing harmful flows in real time.

In 2021, Alshammari and Aldribi [14] have suggested detection architecture for feeding IDS to identify network traffic abnormalities that included an ML method. This detection technique made use of a dataset made up of harmful and non-malicious traffic. The extracted features utilized to train the ML method regarding several

assaults to determine whether it was an abnormality or routine traffic were the research's major obstacles. The "ISOT-CID network traffic dataset" was used to train the ML method. Few key column properties were incorporated, and it was approved that they help the ML method during the training process. This functionality was dependent on the "traffic flow's rambling packet payload length." This research's provided findings and experiment was noteworthy, and it motivated other researchers to continue working on it in the future.

In 2020, Shafiq et al. [15] have presented a unique CorrACC feature selection measure to choose efficient features, followed by bijective soft set for appropriate feature selection. Following that, Corracc, a novel feature selection method on the basis of CorrACC was conceived and developed. On the BoT-IoT dataset, four distinct ML classifiers were employed to evaluate the suggested techniques. The experimental findings achieved by the algorithms were encouraging, with an accuracy of greater than 95%.

In 2019, Sathesh et al. [16] have investigated a block chain technology to safeguard the reliability of the information's stored and utilized the digital signature with authentication. It was utilized to protect the private information's in the patient's record developing a model. The proposed model has examined and compared with the traditional techniques used for storing the medical records, on the terms of reply time, along with the cost of storing and retrieving.

In 2020, Kumar et al. [17] have combined the soft computing techniques and frames using an enhanced soft computing approach to detect the intrusion that caused security issues in the social network. The author has employed the enhanced soft computing technique that combines the fuzzy logic, decision tree, K-means–EM and the machine learning in preprocessing, feature reduction, clustering, and classification, respectively, to develop a security approach that was more effective than the traditional computations in identifying the misuse in the social networks. The intrusion detection system developed using the soft computing approach was tested to given the security percentage, time utilization, and cost and compared with the other traditional methods.

In 2017, Sridhar et al. [18] have proposed the light weight asymmetric cryptography for securing the end-to-end devices which protects the IoT service gateway and the low power sensor nodes. The author has implemented a lattice-based cryptography for securing the broker devices/gateway and the cloud services. The proposed architecture has implemented asymmetric key encryption to share session key between the nodes and then used this session key for message transfer. This has protected the system from distributed denial of service attacks, eavesdropping and quantum algorithm attacks. The proposed protocol used the unique device ID of the sensors to generate key pair to establish mutual authentication between devices and services.

In 2021, Vivekanandam et al. [19] have combined with a better classifier in order to detect the different malware categories. The proposed approach was compared and validated with current techniques by using different datasets. In addition to the test results, this research work has utilized the algorithm for solving a real challenge in android categorization. The experimental results show that, the proposed approach

was superior. Besides, the proposed algorithm has provided a better mean and standard deviation value in the optimization process for leveraging model effectiveness at different datasets.

In 2020, Kirubakaran et al. [20] have presented a secured cloud environment by creating a trusted relationship with the cloud service provider and also by comparing the current security mechanisms applied in the real-world businesses. This research work helped the readers to understand the need of security mechanisms in the cloud environment and the need for the service-level-agreements (SLAs) with the trusted cloud service providers.

2.2 Review

Numerous academics have proposed various ML approach methods to restrict fraudulent traffic flows. Table 1 lists the features and challenges of traditional malicious network traffic flow methods. ML [11] chooses efficient features by AUC metric and also filters the features in an appropriate manner. But, the keylogging theft attacks are poorly detected. DRL [12] achieves less intrusive monitoring and also learns

Table 1 Features and challenges of conventional malicious network traffic flow methods

Author [citation]	Methodology	Features	Challenges
In 2021, Shafiq et al. [11]	ML	The features are filtered in an appropriate manner The efficient features are chosen by AUC metric	It poorly detects the keylogging theft attacks
In 2021, Kim et al. [12]	DRL	Better sampling resource allocation policy is learnt Less intrusive monitoring is achieved	It suffers from important detection delay
In 2020, Hwang et al. [13]	CNN	The process of detection is faster Minimum flow preprocessing time is consumed	Efficient online anomaly detection systems are not built
In 2021, Alshammari and Aldribi [14]	ML	It extracts the packet payload length The deviations within the data instances are captured	The use in real networks affects the speed
In 2020, Shafiq et al. [15]	ML	The optimal features are chosen with respect to "specificity, sensitivity, precision, and accuracy" The accuracy returned is higher	It does not identify the data theft attacks

better sampling resource allocation policy. Still, it suffers from important detection delay. CNN [13] consumes minimum flow preprocessing time and the process of detection is faster. Yet, it does not build efficient online anomaly detection systems. ML [14] captures the deviations within the data instances and also extracts the packet payload length. But, the use in real networks affects the speed. ML [15] chooses the optimal features in terms of "specificity, sensitivity, precision, and accuracy" and the accuracy returned is higher. Still, the data theft attacks are not identified. Thus, it is necessary to suggest novel approaches using ML for efficiently detecting the malicious encrypted network traffic flow.

3 Dataset Description and Proposed Architecture Considered for Malicious Internet Protocol Network Flow Detection

3.1 Dataset Experimented

The suggested method uses data from two well-known datasets: "GitHub and ISCX 2016," both of which are described here.

Dataset 1: It is from "https://github.com/IvanLetteri/MTA-KDD-19/find/master: Access Date: 2021-11-19." This dataset, dubbed "MTA-KDD'19," was created by taking into account two types of traffic: "malicious and legitimate." A total of 2112 pcap flights from June 2013 to August 2019 are included in the malicious traffic. From the case of stratosphere project, legitimate traffic collects 15 pcap files.

Dataset 2: It is from "ISCX 2016-http://205.174.165.80/CICDataset/ISCX-VPN-NonVPN-2016/Dataset/: Access Date: 2021-11-19." This dataset contains real-world traffic in ISCX, with 14 different traffic forms such as "VPN-P2P, P2P, VPN-VOIP, VOIP," and so on. It also has eight categories including "P2P, VoIP, file transfer, streaming, chat, email, online surfing, and traffic." The acquired data is referred to as DU_{iq}, in which $iq = 1, 2, \cdots, IQ$ while the total count of accumulated data is referred to as IQ. The attributes are available in the dataset 1 and dataset 2 is given in Table 2.

3.2 Proposed Model

The count of malware attacks that utilize encrypted HTTP traffic for communication and self-propagation has risen dramatically nowadays. Encrypted malicious traffic mimics benign traffic owing to the abundance of valid encrypted data. Since criminal traffic resembles innocuous traffic, firewalls, and antivirus software are challenged. As these often do not have accessibility to encryption keys, detection approaches that do not require traffic decryption are necessary. The increased use of

Table 2 Attributes illustrated in dataset 1 and dataset 2 for malicious network traffic flow model

Datasets	Name	Attributes
Dataset 1	MTA-KDD'19	MaxLen, MinLen, StdDevLen, AvgLen, MaxIAT, MinIAT, AvgIAT, AvgWinFlow, MaxLenrxMinLenrx,StdDevLenrx, AvgLenrx, MinIATrx, AvgIATrx, NumPorts, FlowLEN, Start_flow, DeltaTimeFlow, HTTPpkts, label, etc.
Dataset 2	ISCX 2016	Duration, protocol-type, service, harvest, hostnames, http, http-2784, http-443, flag, land, wrong-fragment, urgent, num-failed-logins, logged-in, num-compromised, root-shell, num-root, num-file-creations, num-shells, num-access-files, dst-host-count, etc.

encrypted network traffic needs a novel ML traffic analysis technique that does not depend on packet payload or pattern matching content to detect suspicious or malicious communications. Since Internet traffic encryption is becoming more common, network packet content analysis is yielding decreasing benefits. The architecture of the introduced malicious network traffic flow detection model is shown in Fig. 1. The proposed malicious encrypted network traffic flow detection model is composed of four phases such as, "data collection, feature extraction, weighted feature extraction, and malicious detection." The data is initially collected from standard online sources. Next, in the feature extraction phase, DCNN is used, where the features are extracted from the convolutional layer of DCNN [21]. DCNN is utilized as provides better traffic detection. DCNN is an efficient way of feature extraction, it can reduce data dimension and produce a less redundant dataset, also called as a feature map. Each kernel works as a feature identifier, where the filtering out exists feature in the original image. For the above benefits DCNN is chosen for feature extraction. From the extracted features, the weighted feature extraction is done by the GOA, where the weight is tuned by the GOA with the intention of variance maximization. The extracted weight features undergo the final step of malicious detection using hybrid ML consisting of SVM and NN. The hybrid ML model-based malicious detection is carried out by SVM and NN, in which the hidden neuron count of NN and kernel of SVM are optimized by the same GOA for maximizing accuracy and precision. The outcomes from both SVM and NN undergo averaging operation to detect the traffic as the final output.

4 Weighted Feature Extraction and Detection Model for Network Malicious Identification

4.1 Weighted Feature Extraction by GAO

The data is converted into features here. The main advantages of the weighted feature extraction technique is listed below, and it can increases the accuracy. It can solves the

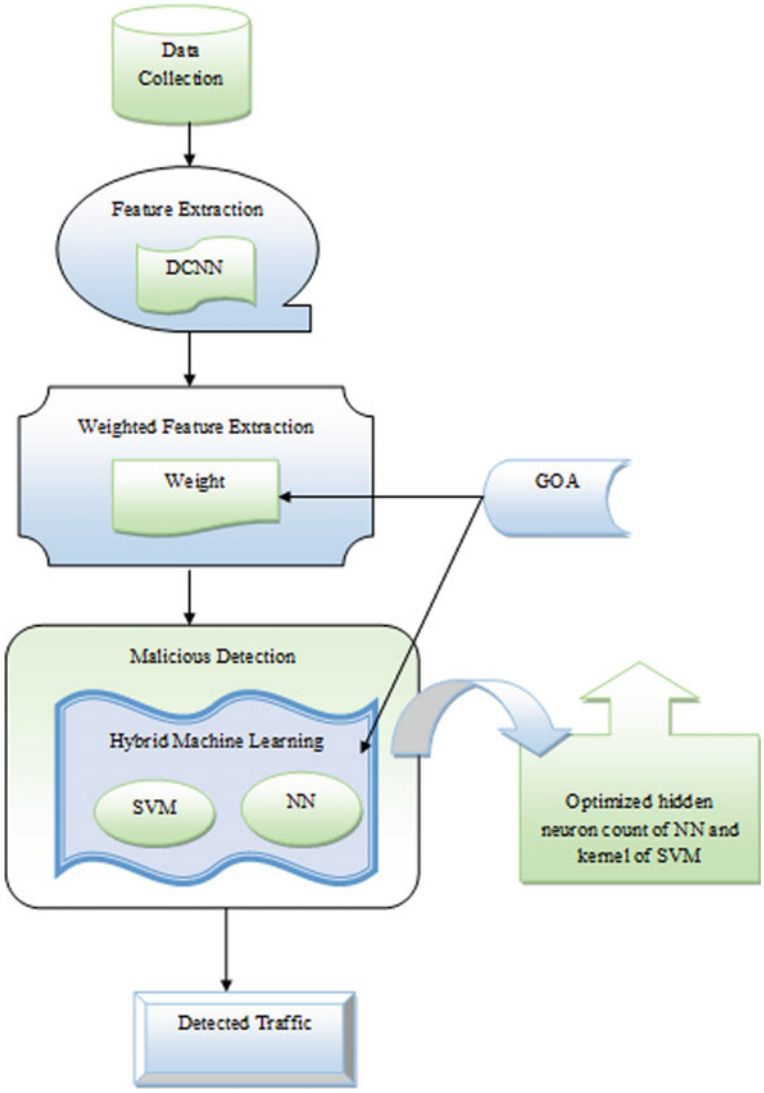


Fig. 1 Proposed Malicious network traffic flow model

over fitting problems. It increases the training speed and also enhanced the data visualization. The weight is optimized using new algorithm and then, weighted feature extraction is carried out. The bounding limit of the weight is between (0–1). The weighted feature extraction is done mathematically as in Eq. (1).

$$\mathrm{WT}^{nw} = \mathrm{WT} \times Ft(fa) \tag{1}$$

Here, the extracted weight features are shown by WT^{nw} and the weight that is being optimized by GOA is shown by WT, respectively. The major objective of the weighted feature extraction for the proposed model is to optimize the weight by the GOA with the consideration of variance maximization. Thus, the objective is shown in Eq. (2).

$$OF_1 = \arg\min_{\{WT\}} \left(\frac{1}{VR} \right) \tag{2}$$

Here, the objective function of the weighted feature extraction phase of the introduced model is shown by OF_1. The variance is defined as "the amount of dispersion of a dataset's data points from its mean," which is calculated as "the average of each data point's squared deviation from the population mean," as stated in Eq. (3).

$$VR = \sigma^2 = \frac{\sum \left(Ft_p(fa) - \mu \right)^2}{P} \tag{3}$$

It is also calculated by "summing up every data point's squared deviation and dividing the result by the total number of data points in the feature set." The count of features in the feature set is provided as P, the mean of features is presented as μ, and the pth feature in the feature set is defined as $Ft_p(fa)$. Furthermore, the optimally extracted weight features are shown as $Ft *_{p*} (fa)$, where $p* = 1, 2, \ldots, P^*$ and P^* denote the total count of weight extracted features.

For tackling optimization issues, the suggested approach GOA [22] mathematically designs and simulates the behavior of grasshopper swarms in nature. The following is the mathematical formula that was used to replicate grasshopper swarming behavior.

$$Y_j = T_j + H_j + B_j \tag{4}$$

Here, Y_j represents the jth grasshopper's location, T_j represents social interaction, H_j represents gravity force on the jth grasshopper, and B_j represents wind advection. Consider that the equation may be written as $Y_j = s_1 T_j + s_2 H_j + s_3 B_j$ to achieve random characteristics with s_1, s_2, and s_3 being random values in the range [0,1].

$$T_j = \sum_{\substack{k=1 \\ k \neq j}}^{O} t\left(e_{jk}\right) \hat{d}_{jk} \tag{5}$$

In the above equation, e_{jk} describes the distance among the jth and kth grasshopper, computed as $e_{jk} = \left| Y_k - Y_j \right|$, t shows a function to characterize the strength of social pressures, as illustrated in Eq. (6), and $\hat{e}_{jk} = \frac{y_k - y_j}{e_{jk}}$ defines a unit vector. The social forces are defined by the t function, which is determined as below in Eq. (6).

$$t(s) = gf^{-\frac{s}{m}} - f^{-s} \tag{6}$$

In Eq. (4), the H component is computed as in Eq. (7).

$$H_j = -h\hat{f}_h \tag{7}$$

Here, h denotes the gravitational constant and $\hat{f}_h$ denotes a unity vector. In Eq. (4), the B component is computed as in Eq. (8).

$$B_j = v\hat{f}_x \tag{8}$$

In the above equation, v denotes a continuous drift and $\hat{f}_x$ denotes a unity vector. This equation may be expanded by substituting T, H, and B in Eq. (9).

$$Y_j = \sum_{\substack{k=1 \\ k\neq j}}^{O} t\left(\left|Y_k - Y_j\right|\right)\frac{y_k - y_j}{e_{jk}} - h\hat{f}_h + v\hat{f}_x \tag{9}$$

Here, O shows the count of grasshoppers and $t(s) = gf^{-\frac{s}{m}} - f^{-s}$.

Yet, the grasshoppers soon find their comfort zone, therefore, this mathematical method cannot be utilized to address optimization issues directly. To tackle optimization difficulties, a modified form of this equation is introduced as below in Eq. (10).

$$Y_j^e = d\left(\sum_{\substack{k=1 \\ k\neq j}}^{O} d\frac{uc_e - lc_e}{2} t\left(\left|y_k^e - y_j^e\right|\right)\frac{y_k - y_j}{e_{jk}}\right) + \hat{U}_e \tag{10}$$

Here, uc_e defines the upper limit in the Eth dimension, lc_e shows the lower bound, $t(s) = gf^{-\frac{s}{m}} - f^{-s}$, $\hat{U}_e$ describes the value of the Eth dimension in the objective, and d indicates a decreasing coefficient to shorten the "comfort zone, repulsion zone, and attraction zone."

The comfort zone is reduced proportionally to the count of iterations by the coefficient c, which is computed as in Eq. (11).

$$d = d\max - m\frac{d\max - d\min}{M} \tag{11}$$

The highest value is d max, the minimum value is d min, the current iteration is m, and the maximum number of iterations is M. The pseudo code of GOA is shown in Algorithm 1.

Algorithm 1: GOA [22]		
Start		
Swarm population initialization		
Initialization of d max,d min and maximum iteration count		
Fitness computation of every search agent		
Best search agent is specified as U		
While ($m < m_{\max}$)		
	Update d by Eq. (11)	
	For every search agent	
		Distance normalization within grasshoppers
		Position update of the current search agent through Eq. (10)
		Current search agent moving outside the boundaries is brought back
	End for	
	Update U if better solution exists	
	$m = m + 1$	
End while		
Return U		
Stop		

4.2 Hybrid Machine Learning-Based Malicious Detection

This malicious detection phase is used for detecting the network traffic of the developed network traffic malicious detection model. Here, the malicious detection is done on the basis of the hybrid ML model. The advantages of the hybrid model are discussed below, it solves dimensionality problems and also it is computationally fast. Hybrid model is easy to implement for optimizations techniques. Hybrid algorithms play a prominent role in improving the search capability of algorithms. Hybridization aims to combine the advantages of each algorithm to form a hybrid algorithm, while simultaneously trying to minimize any substantial disadvantages. In general, the outcome of hybridization can usually make some improvements in terms of either computational speed or accuracy. This hybrid ML consists of two classifiers such as NN and SVM. The description of these classifiers is shown below.

NN: NN [23] is more efficient than other classifiers, therefore, it is utilized in a range of applications. It is made out of "three layers termed input, hidden, and output layers." Before calculating the complete network's output, Eq. (12) must be used to determine the hidden output, and Eq. (13) must be used to determine the network's output from the output layer.

$$\overline{O}^{(O)} = \text{ACH}\left(\tilde{J}_{(\hat{C}d)}^{(O)} + \sum_{d=1}^{Ip(f)} \tilde{J}_{(cd)}^{(O)} Ft *_{p*} (fa)\right) \tag{12}$$

$$\hat{N}_e = \text{ACH}\left(\tilde{J}_{(\hat{C}e)}^{(N)} + \sum_{d=1}^{HDP} \tilde{J}_{(de)}^{(N)} \overline{O}^{(O)}\right) \tag{13}$$

As demonstrated in Eqs. (12) and (13), the input features are characterized by $Ft *_{p*} (fa)$, the input neuron count is given by $Ip(f)$, and the hidden neuron count is supplied by HDP, c, d and f symbolize the input, hidden, and output neurons, respectively. The weight from the cth input neuron to the dth hidden neuron is also described as $\tilde{J}_{(cd)}^{(O)}$, as well as $\tilde{J}_{(de)}^{(N)}$ defines the weight from the dth hidden neuron to the eth output neuron. The bias weight of the hidden neuron is expressed as $\tilde{J}_{(\hat{C}d)}^{(O)}$, whereas the bias weight of the output neuron is shown as $\tilde{J}_{(\hat{C}e)}^{(N)}$. The term ACH represents the activation function. The predicted output of the network is given by $\hat{N}_e$. The weight function $J_y^{NN} = \left\{\tilde{J}_{(\hat{o}hp)}^{(O)}, \tilde{J}_{(\hat{o}op)}^{(N)}, \tilde{J}_{(cd)}^{(O)}, \tilde{J}_{(de)}^{(N)}\right\}$ has been carefully designed to provide NN with the best possible training. Equation (14), also called as measured error, that explains the numerical equation in which $\hat{N}_e$ is the predicted outcome and N_e is the actual outcome.

$$MG1 = \left\{J_y^{NN} = \left\{\tilde{J}_{(\hat{o}hp)}^{(O)}, \tilde{J}_{(\hat{o}op)}^{(N)}, \tilde{J}_{(cd)}^{(O)}, \tilde{J}_{(de)}^{(N)}\right\}\right\} \underset{}{\arg\min} \sum_{e=1}^{N(f)} \left|N_e - \hat{N}_e\right| \tag{14}$$

Thus, the predicted output from NN is shown by $\Pr e^{NN}$.

SVM: SVMs [24] have piqued academic interest because they "have been effectively employed in a variety of categorization applications because of their greater generalization capabilities." After shifting nonlinearly divided samples into a higher (possibly infinite) dimensions space, these methods use the biggest margin algorithm to discover a separating hyperplane. Examine the below dataset: $DG = \{(xg_1, yg_1), (xg_2, yg_2), \cdots, (xg_{NG}, yg_{NG})\}$ where $xg_{ig} \in \Re^{kg}$ and $xg_{ig} \in \Re^{kg}$ are categorized using the separating hyperplane $wg^T xg + bg = 0$, in which wg is the hyperplane direction, and bg shows the distance within the origin and the hyperplane. Equation (15) defines the underlying formula for the SVM.

$$\begin{aligned} &\min_{wg,bg} \frac{1}{2}\|wg\|^2 \\ &\text{s.t.} yg_{ig}\left(wg^T xg_{ig} + bg\right) \geq 1,\ ig = 1, 2, \ldots, NG \end{aligned} \tag{15}$$

In this case, $yg_{ig} fg_{ig} \geq 1$ for the entire $ig\left(fg_{ig} = wg^T xg_{ig} + bg\right)$. As a result, the final predicted output of SVM is represented as $\Pr e^{SVM}$.

The NN offers advantages like better fault tolerance, capability of parallel processing, etc. But, it suffers from hardware dependence, unknown network duration, etc. Similarly, the SVM provides benefits such as much efficient in high dimensional spaces, relatively memory effective, etc. Still, it limits from the fact that it is not applicable for vast datasets, not better in the overlapping cases, etc. Hence, to overcome the shortcomings in both the NN and SVM, hidden neuron count of NN as well as the kernel of SVM are tuned by GOA with the consideration of accuracy and precision maximization, thus referred as hybrid ML. This hybrid ML can work with larger datasets and also overcomes the hardware dependence problem.

The main objective of hybrid ML-based malicious detection of the introduced network traffic model is to optimize the hidden neuron count of NN as well as kernel of SVM by the GOA with the intention of accuracy and precision maximization. The kernels considered here are the, "linear, poly, rbf, and sigmoid." The bounding limit of hidden neurons of NN lies between (5–255) and for kernel, it lies between (0–3). Thus, the objective is shown in Eq. (16).

$$\mathrm{OF}_2 = \underset{\{\mathrm{HDP_{NN}}, \mathrm{KE_{SVM}}\}}{\arg\min} \left(\frac{1}{Acy + \Pr n} \right) \tag{16}$$

Here, the objective function of the malicious detection phase of the introduced model is shown by OF_2, hidden neuron count of NN is shown by $\mathrm{HDP_{NN}}$, kernel of SVM is shown by $\mathrm{KE_{SVM}}$, accuracy is shown by Acy, and precision is shown by $\Pr n$, respectively. Accuracy is defined as, "a ratio of the observation of exactly predicted to the whole observations."

$$Acy = \frac{(t^{pos} + t^{ng})}{(t^{pos} + t^{ng} + f^{pos} + f^{ng})} \tag{17}$$

"Precision is the ratio of positive observations that are predicted exactly to the total number of observations that are positively predicted" as shown in Eq. (18).

$$\Pr n = \frac{t^{pos}}{t^{pos} + f^{pos}} \tag{18}$$

In the above equation, terms t^{ng},t^{pos}, f^{pos}, and f^{ng} indicates "true negatives, true positives, false positives, and false negatives," respectively. The predicted outcomes from both NN and SVM undergo averaging operation for generating the final outcome as in Eq. (19).

$$\Pr e^{\mathrm{final}} = \frac{\Pr e^{\mathrm{NN}} + \Pr e^{\mathrm{SVM}}}{2} \tag{19}$$

The final predicted outcome for detecting the network traffic of the introduced model is shown by $\Pr e^{\mathrm{final}}$, respectively.

5 Results and Discussions

5.1 Experimental Setup

The proposed malicious network traffic flow model was implemented in Python, and the findings were analyzed. The population size and the iteration count were ten. The GOA [22] was compared with classifiers such as NN [23] and SVM [24] in terms of detection analysis with respect to various performance measures to show the betterment of the suggested method. The number of population N_{pop}, chromosome length chlen, and maximum iteration $MaxIter$ is considered as 10, 10 and 10, respectively. The computational complexity of the proposed GOA-NN-SVM model is added below.

$$O\left[N_{\text{pop}} + N_{\text{pop}} + \left[\text{MaxIter} * \left[(N_{\text{pop}} * \frac{\text{chlen}}{2} * (N_{\text{pop}} - 1)) + (N_{\text{pop}})\right]\right]\right].$$

5.2 Detection Analysis

The detection analysis of various classifiers for the introduced malicious network traffic flow model for the two datasets is shown in Fig. 2. The outcomes are superior with GOA than the other methods. From Fig. 2a, for dataset 1, the accuracy of GOA at 75% learning percentage is 5.67%, 7.79%, and 5.79% better than NN-SVM, SVM, and NN, respectively. Therefore, the outcomes of detection analysis are better with GOA than the other methods for the introduced network traffic flow detection model (Fig. 3).

5.3 Overall Detection Analysis

The overall detection analysis of different methods for the introduced malicious network traffic flow model with the two datasets is given in Table 3 for two datasets. It is clear that the results are better with GOA than the other methods, thereby determining its superiority. For dataset 2, the accuracy of GOA is 4.66%, 8.31%, and 5.84% improved than NN-SVM, SVM, and NN, respectively. Hence, the overall detection analysis holds best outcomes with GOA than the other methods for the introduced network traffic flow detection model in terms of two datasets, respectively.

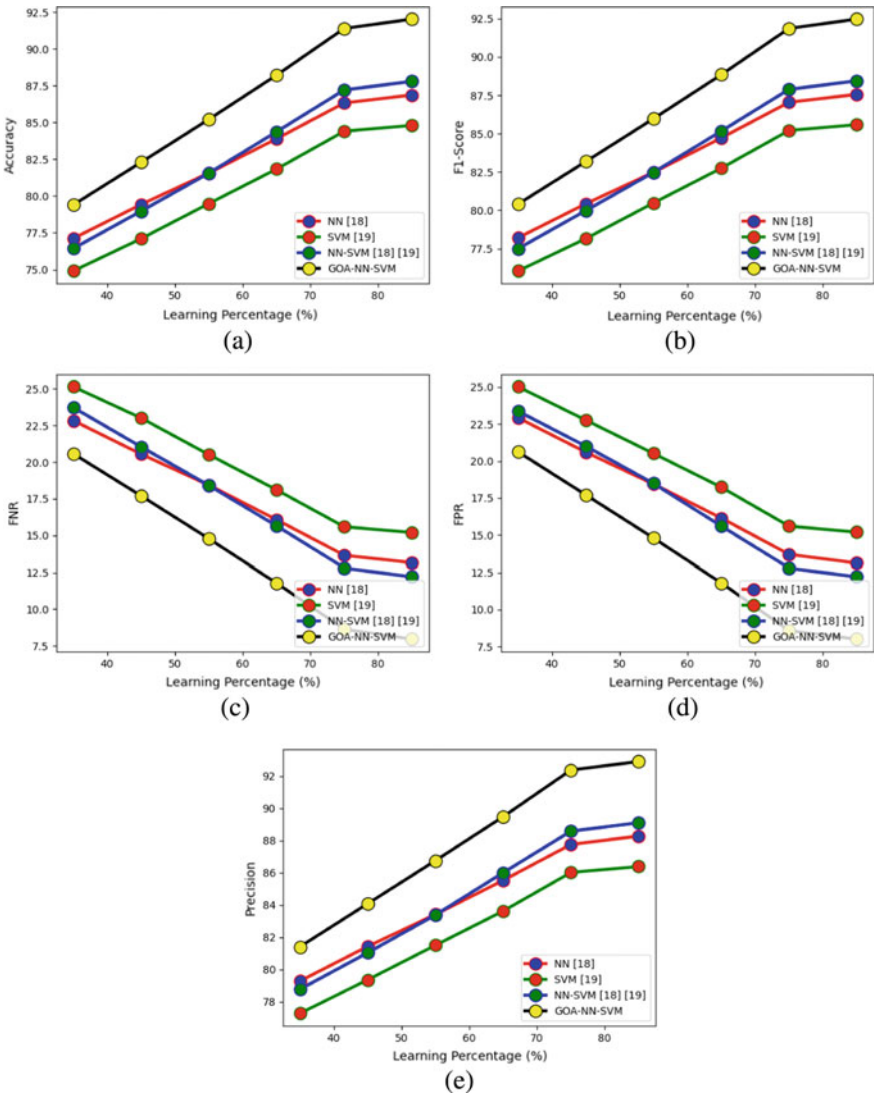


Fig. 2 Detection analysis of different methods for the introduced malicious network traffic flow detection in terms of, "**a** accuracy, **b** F1-score, **c** FNR, **d** FPR, and precision" for dataset 1

5.4 State-Of-the-Art Comparison of Proposed Model

The state-of-the-art comparison of the detection analysis of various traditional models regarding a positive and negative measures for the introduced malicious network traffic flow model for the two datasets is shown in Figs. 4 and 5. From Fig. 4 based on accuracy, in terms of positive measures, the proposed GOA-NN-SVM model is

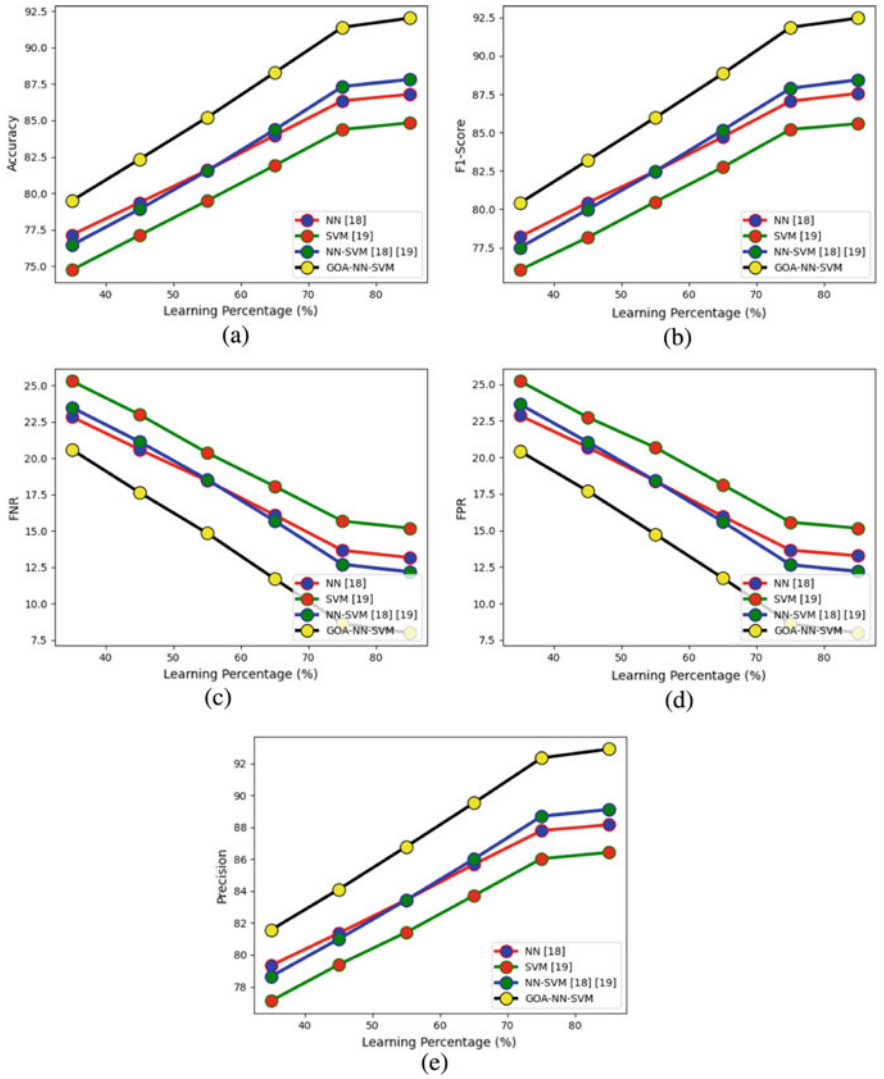


Fig. 3 Detection analysis of different methods for the introduced malicious network traffic flow detection in terms of, "**a** accuracy, **b** F1-score, **c** FNR, **d** FPR, and precision" for dataset 2

secured 4.49%, 7.86%, and 4.49% progressed than mutual and double authentication schemes, genetic algorithms, and encryption algorithms, respectively for dataset 1. Accordingly, from Fig. 5 based on FNR, the proposed GOA-NN-SVM model attains 52.94%, 76.47%, and 41.17% lower than mutual and double authentication schemes, genetic algorithms, and encryption algorithms, respectively for dataset 2. Therefore, it is proved that the proposed GOA-NN-SVM model holds enriched performance when comparing the other conventional traditional methods.

Table 3 Overall detection analysis of different methods for the introduced malicious network traffic model for two datasets

Performance measures	NN [23]	SVM [24]	NN-SVM [23, 24]	GOA [22]
Dataset 1				
"Accuracy"	0.863293	0.844084	0.872138	0.913948
"Sensitivity"	0.863548	0.844158	0.872107	0.91368
"Specificity"	0.863003	0.843999	0.872173	0.914253
"Precision"	0.877574	0.860211	0.885827	0.923765
"FPR"	0.136997	0.156001	0.127827	0.085747
"FNR"	0.136452	0.155842	0.127893	0.08632
"NPV"	0.847597	0.82646	0.857078	0.903041
"FDR"	0.122426	0.139789	0.114173	0.076235
"F1-score"	0.870504	0.852109	0.878913	0.918695
"MCC"	0.72586	0.687414	0.743592	0.827369
Dataset 2				
"Accuracy"	0.863479	0.84379	0.873253	0.913917
"Sensitivity"	0.863431	0.843285	0.873039	0.913884
"Specificity"	0.863533	0.844364	0.873498	0.913955
"Precision"	0.877975	0.860368	0.886983	0.923536
"FPR"	0.136467	0.155636	0.126502	0.086045
"FNR"	0.136569	0.156715	0.126961	0.086116
"NPV"	0.847566	0.82572	0.858156	0.903219
"FDR"	0.122025	0.139632	0.113017	0.076464
"F1-score"	0.870642	0.851741	0.879955	0.918685
"MCC"	0.726252	0.686868	0.745837	0.827297

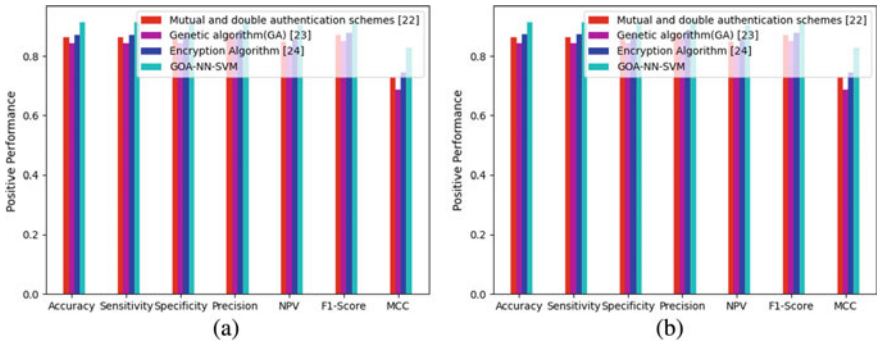


Fig. 4 Detection analysis of different methods for the introduced malicious network traffic flow detection in terms of positive and negative measures for **a** dataset 1 and **b** dataset 2

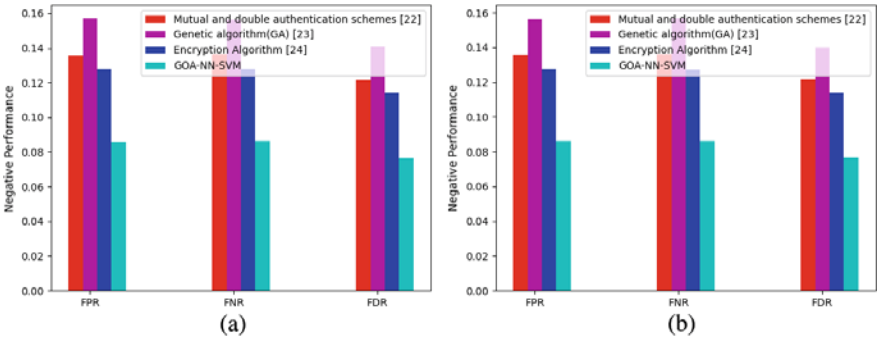


Fig. 5 Detection analysis of different methods for the introduced malicious network traffic flow detection in terms of in terms of positive and negative measures for **a** dataset 1 and **b** dataset 2

6 Conclusion

The hybrid ML approach was used to create a novel traffic flow detecting model in this work. The benchmark data was first gathered from public sources. The GOA then performed the weighted feature extraction. The SVM and NN were then utilized in this model to identify malicious traffic, with the parameter tweaking handled by the same GOA. From the analysis, the accuracy of GOA was 4.66%, 8.31%, and 5.84% improved than NN-SVM, SVM, and NN, respectively. Therefore, the GOA was better than the other methods for the introduced malicious network traffic flow detection model, respectively for the two datasets.

References

1. Egea S, Mañez AR, Carro B, Sánchez-Esguevillas A, Lloret J (2018) Intelligent iot traffic classification using novel search strategy for fast based-correlation feature selection in industrial environments. IEEE Internet Things J 5(3):1616–1624
2. Shafiq M, Tian Z, Bashir AK, Jolfaei AR, Yu X (2020) Data mining and machine learning methods for sustainable smart cities traffic classification: a survey, Sustain Cities Soc
3. Zhao D, Traore I, Sayed B, Lu W, Saad S, Ghorbani A, Garant D (2013) Botnet detection based on traffic behavior analysis and flow intervals. Comput Secur 3(9):2–16
4. Shafiq M, Yu X (2017) Effective packet number for 5g im wechat application at early stage traffic classification, Mob Inf Syst 2017
5. Zhang H, Lu G, Qassrawi MT, Zhang Y, Yu X (2012) Feature selection for optimizing traffic classification. Comput Commun 35(12):1457–1471
6. Shafiq M, Yu X, Bashir AK, Chaudhry HN, Wang D (2018) A machine learning approach for feature selection traffic classification using security analysis. The J Supercomput 74(10):4867–4892
7. Qiu J, Du L, Zhang D, Su S, Tian Z (2019) Nei-tte: Intelligent traffic time estimation based on fine-grained time derivation of road segments for smart city. IEEE Trans Ind Inf

8. Shafiq M, Tian Z, Bashir AK, Du X, Guizani M (2021) CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine-learning techniques. IEEE Internet Things J 8(5):3242–3254
9. Kim S, Yoon S, Lim H (2021) Deep reinforcement learning-based traffic sampling for multiple traffic analyzers on software-defined networks. IEEE Access 9:47815–47827
10. Hwang R-H, Peng M-C, Huang C-W, Lin P-C, Nguyen V-L (2020) An unsupervised deep learning model for early network traffic anomaly detection. IEEE Access 8:30387–30399
11. Alshammari A, Aldribi A (2021) Apply machine learning techniques to detect malicious network traffic in cloud computing, J Big Data 8
12. Shafiq M, Tian Z, Bashir AK, Du X, Guizani M (2020) IoT malicious traffic identification using wrapper-based feature selection mechanisms. Comput Secur 94
13. Tian Z, Gao X, Su S, Qiu J (2020) "Vcash: A novel reputation framework for identifying denial of traffic service in internet of connected vehicles. IEEE Internet Things J 7(5):3901–3909
14. Shafiq M, Tian Z, Sun Y, Du X, Guizani M (2020) Selection of effective machine learning algorithm and bot-iot attacks traffic identification for internet of things in smart city. Future Gener Comput Syst 107:433–442
15. Peng L, Yang B, Chen Y, Chen Z (2016) Effectiveness of statistical features for early stage internet traffic identification. Int J Parallel Program 44(1):181–197
16. Sathesh A (2019) Enhanced soft computing approaches for intrusion detection schemes in social media networks. J Soft Comput Paradigm (JSCP) 1(02):69–79
17. Kumar D, Smys S (2020) Enhancing security mechanisms for healthcare informatics using ubiquitous cloud. J Ubiquitous Comput Commun Technol 2(1):19–28
18. Sridhar S, Smys S (2017) Intelligent security framework for iot devices cryptography based end-to-end security architecture. In: 2017 international conference on inventive systems and control (ICISC), 1–5. IEEE
19. Vivekanandam B (2021) Design an adaptive hybrid approach for genetic algorithm to detect effective malware detection in android division. J Ubiquitous Comput Commun Technol 3(2):135–149
20. Kirubakaran SS (2020) Study of security mechanisms to create a secure cloud in a virtual environment with the support of cloud service providers. J Trends Comput Sci Smart Technol (TCSST) 2(03):148–154
21. Namatevs I (2017) Deep convolutional neural networks: structure, feature extraction and training, Inf Technol Manage Sci 20(1)
22. Saremi S, Mirjalili S, Lewis A (2017) Grasshopper optimisation algorithm: theory and application. Adv Eng Softw 105:30–47
23. Fernández-Navarro F, Carbonero-Ruz M, Alonso DB, Torres-Jiménez M (2017) Global sensitivity estimates for neural network classifiers. IEEE Trans Neural Netw Learn Syst 28(11):2592–2604
24. Yu S, Li X, Zhang X, Wang H (2016) The OCS-SVM: an objective-cost-sensitive SVM with sample-based misclassification cost invariance, IEEE Access 4

Influencer Profiling to Identify the Top Keywords Using LDA

Bahaa Eddine Elbaghazaoui, Mohamed Amnai, and Youssef Fakhri

Abstract In line with the strong growth of the social data volume, a huge number of social contents are created in various forms every day. As a result, it is often difficult for users to find or keep up with current popular keywords. Many researchers are interested to analyze the social data and identify the current popular keywords and topics. Thus, these researches are more complex for profiling the target topics with their keywords. The main purpose of this paper is to identify the most trending keywords in social media based on profiling influencers. So, for achieving this objective. Our approach illustrates determining influential nodes on a social network graph. After that, due to the latent Dirichlet allocation model, we could profile the trending keywords based on influential posts content. In any case, our proposal is more relevant and optimal compared to the previous approaches.

Keywords Topic modeling · Social media · User profiling · Latent Dirichlet allocation · LDA

1 Introduction

Due to the media features provided, many social platforms have become very popular. Users can express their views, opinions, and preferences on different topics and products in a very concise form [1]. The vast amount of data that consumers have released has piqued the curiosity of research organizations and commercial firms. However, through people's opinions, we can understand the topic motivation or

B. E. Elbaghazaoui (✉) · M. Amnai · Y. Fakhri
Laboratory of Computer Sciences, Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco
e-mail: bahaaeddine.elbaghazaoui@uit.ac.ma

M. Amnai
e-mail: mohamed.amnai@uit.ac.ma

Y. Fakhri
e-mail: fakhri@uit.ac.ma

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_29

Fig. 1 Elon Musk tweet impact—"Gamestonk!!" over time

product's popularity [2, 3], and it can also provide valuable help for designing more effective promotional activities.

The Internet is flooded with various information and knowledge resources, which may confuse readers and make them spend more time and energy to find relevant information on specific topics of interest. Therefore, more effective methods and tools are needed to help detect and analyze content in online social media networks, especially for users who use user-generated content as a data source [4]. Furthermore, more useful and hidden metadata must be extracted from a variety of text-based web sources.

The volume of updates in microblogging networks translates into numerous event-related reports [5]. These can include elections, football matches, protests, and even emergencies such as COVID-19. These events are an opportunity for some specific users to give their opinion and create posts regarding the event's topics, especially "influencers." For example, on January 26, 2021, Elon Musk tweeted "Gamestonk !!" about the Reddit page "wallstreetbets," which now has more than 9 million subscribers. Interestingly, it was just a few minutes after the stock market closed. The next day, Gamestop stock jumped 2.5 times. We know how the stock market reacted to Musk. Figure 1 refers to how elon musk impacts Gamestonk.

Nowadays, researchers need powerful automated tools to read data and understand the underlying themes. Topic modeling is a set of techniques with algorithms that can display, discover, and annotate the topic structure in a data collection [6]. However,

the vector space model is the basis of many advanced information retrieval techniques [7, 8] and topic models [9].

Nonetheless, most topic models treat data similarly, whereas in reality, data have varying degrees of relevance on different subjects [10], thus treating them equally may automatically degrade topic modeling efficiency. We must extract and identify only the most significant data from the database in order to quantify the relevance of data on various themes.

We propose incorporating link-based ranking into topic modeling in this paper. Our objective is to profile keywords due to influencers' posts on social media. However, our system suggests starting by identifying influencers from social media. Then, collect the important and actual posts from each influencer. Based on this and thought LDA model, we can determine and assign several topics with keywords in a large data collection in an unsupervised way. Therefore, due to this method, we can optimize the treatment duration to extract topics from our database.

Many approaches for automatically identifying subjects from user postings have been proposed [11, 12]. However, these methods require resources and time to analyze all data sets. Sometimes, the results are irrelevant due to the amount of data and ambiguity. In our approach, we focused only on influencer posts content and extracting topics due to the LDA algorithm. This proposal makes it possible to optimize the analysis time, and it will be effective to know the current subjects.

The following is the structure of this essay, which summarizes all preceding thoughts. We begin by outlining the relevant work and the main issues of our approach in the first section. The basic notion of our solution is presented in the second section. We introduce our implementation in the third section. Finally, we concluded and made recommendations for future work.

2 Background and Related Works

In the real world, many studies have been invested to extract metadata, which is interesting to classify and understand the details of each user. Identifying the data related to the user's interest domain is called user profile analysis or user profile [13], and the process is known by data profiling [14].

This section summarizes the most relevant previous research on identifying social network influencers. Easley [15] or Cook and Holder [16] have extensively researched the basic definitions, algorithms, and methodologies of graph theory and network analysis [17, 18]. Numerous new research have focused on the utilization of platform-specific features. These functions include the number of users who have been followed, the number of users who have been followed, the verification status, the number of people who have been added to lists, and so on.

The recent explosive of the bloggers has attracted a lot of research on related issues [19]. The identification of influential blog sites [20] and related research on the spread of influence between blog sites [21–24] are orthogonal to the issues considered in this article, and the focus is on profiling Influential bloggers on web sites. However,

it is obvious that the works that propose methods for discovering and analyzing blog communities [25, 26] have nothing to do with the current problem.

Identifying influential bloggers is a challenge that has been introduced in [27] and will solve many problems. Therefore, the paper's objective is to extract keywords from social media datasets based on profiling influencers.

3 Problematic

To completely understand user behaviors, many researchers have concentrated on obtaining topic modeling from social media networks [28]. Understanding tendencies can aid in improving the quality of various products and services in accordance with user requirements [29]. However, identifying the most frequent topics in social networks helps understand the user tendencies and allows to give broad visibility on the user needs. Several studies are focused on profiling the most frequent topics in social networks [30]. Unfortunately, most of the proposals find difficult challenges to obtain optimal results. Indeed, the data is dynamic over time and the amount of data is too huge for analysis. Therefore, our goal is to find the top topics on social networks through influencers. This article focuses on profiling the most frequent keywords by category from influencers' post content.

4 Contribution

People are increasingly sharing posts on social media (Facebook, Twitter, Instagram), which include emotions, moods, and feelings related to daily life. Researchers are studying the profound impact of the rapid development of social media on user-generated content, as it has an impact on purchasing behavior and how it shapes user perception and emotion [31].

One of the most important aspects of marketing research is identifying and assessing influential people on social networks. Influential people, on the other hand, frequently share their thoughts on products, services, and companies on social media channels. People are impacted by decisions or ideas made by particular influencers since they communicate with a large network of users that follow them [32]. Some publications of these influencers have spread at a very significant rate [33]. However, nelson emphasized the importance of influencers as role models for relevant and approachable people, because children are willing to establish relationships with them, which in turn can lead to the adoption of influencers' actions and opinions [34]. Van Dam conducted a study on young people and teenagers, and the findings revealed that even older teenagers participate in influencer activities rather than critically thinking about them [35] and many other studies demonstrate the impact of influencers on social media [36].

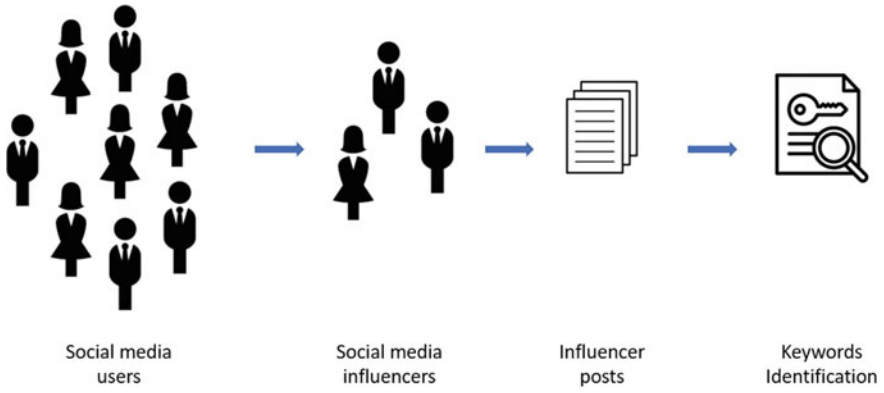


Fig. 2 Workflow to extract keywords based on influencers

Topics in social media are characterized by some keywords, the relevance of which changes over time. The behavior is necessary for learning subject expansion and extracting the most crucial keywords efficiently. Many news stories and social media posts have explored the task of automatic keyword extraction [37, 38].

A keyword can summarize the topic of an article and help to find it when we need it. Keyword assignment has always been a manual process, but with the rise of vast amounts of text data, automatic keyword extraction methods have become essential. Despite the research community's best efforts, the most advanced keyword extraction algorithm still has a number of flaws, and its performance is still inferior to that of many other basic NLP tasks [39]. The first keyword extraction methods mostly follow the supervised method [40]. The first extract keyword features, and then train the classifier on the gold standard data set.

In this paper, our main objective is to identify the trending keywords in social media based on profiling influencers. Most people are attached and convinced by what influencers share on social media and influencers should also share information and opinions on current topics. This requires that influencer posts contain current and potential keywords. If we want to extract the keywords from the entire database, now we need only to identify the influencers from social media datasets, then extract the keywords from their content. Figure 2 explains our proposed workflow.

In what is well explained at the beginning. It only remains how to identify the keywords from a dataset, and also, we need to classify them by type and category. Typically, there are several ways to complete topic keywords identification. One way is hierarchical Bayesian models, like latent Dirichlet allocation (LDA) [9]. However, LDA effectively recognizes topics by searching for frequently occurring words when they are coherent within the text. On the other hand, when the text is incoherent (in terms of word selection or sentence meaning), additional contextual information is needed to fully express the idea of the text.

Latent Dirichlet allocation (LDA) is used to extract topics from documents. LDA is an unsupervised machine learning algorithm that allows a set of text observations

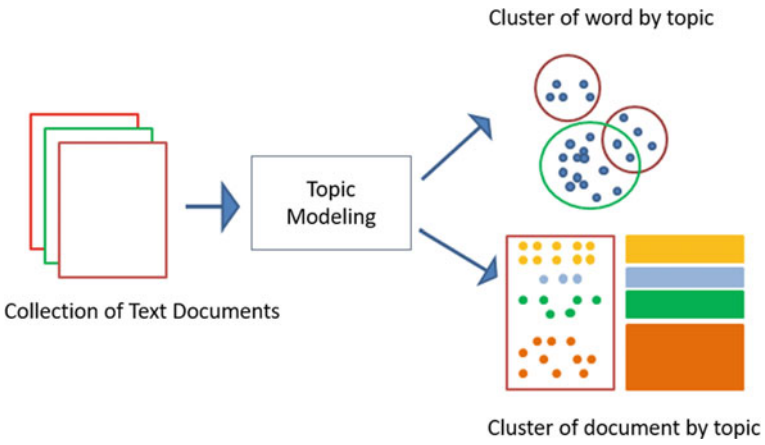


Fig. 3 Latent Dirichlet allocation (LDA)

to be interpreted by an unobserved set of similarities in the interpreted data. LDA represents a document as a mixture of topics, which spit out words with a certain probability as present in Fig. 3.

To understand a given corpus (all text documents), LDA works by a process which consists in calculating the distribution of subjects and the distribution of words for each subject, the words of the text document described by the model of the bag of words (bag-of-words) ignoring the word order, essentially follow a polynomial distribution, which can be calculated from its distribution. To generalize, the LDA algorithm process works as follows:

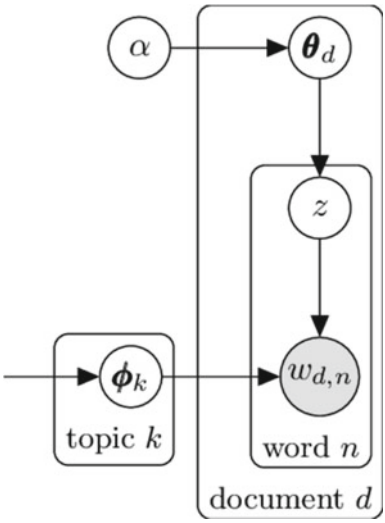
- θ_{d} is the distribution of themes or topics and has a dimension of k where k is the number of subjects or topics.
- ϕ_{d} is the word distribution for the topic of index k, ϕ_{d} has V dimensions where V is the number of terms in the bag-of-word.
- For the document d_{m} among the M corpus documents, we start by sampling the index of the topic $z_{d,n}$, which is the topic number of the nth word in the document d_{m}, $z_{d,n}$ is a value between 1 and k.
- From $z_{d,n}$, choose the corresponding distribution ϕ_{d} and sample the word $w_{d,n}$ is the nth word in the document d_{m}.

In practice, we are interested in the opposite direction about LDA since the number of topics and the corpus are already acquired data, so we look for the multinomial distributions ϕ_{d} and θ_{d} that can be expressed by a probabilistic relation of words as present in Fig. 4.

This procedure can be used for a variety of purposes:

- As in our situation, detect Twitter's "trending topics."
- Find out about fresh hot topics that have been reported in the media.
- Based on a set of professional articles, find fresh and fascinating stocks.

Fig. 4 LDA process



- Arrange scientific texts according to a theme.
- To improve a product, identify the various features that can be criticized.

Note that the LDA method does not give a defined theme but rather the most used words in a theme where the human being must always identify this theme himself. However, we also need to identify the title of each category. Indeed, using machine learning, we can classify each keyword by its category. We need a huge dataset to classify each word and to give a distance or membership factor to a category.

5 Approach Implementation

We started by implementing the first point of our method. However, we have uploaded a database from the Kaggle platform under the name (Identifying Influential Bloggers: Techcrunch) [41], This data set is a crawl of Techcrunch technology blog posts conducted in April 2010. The main goal of this dataset is to provide an active community to identify members who are both productive and influential. There are four tables in the database:

- **authors**: 107 bloggers of Techrunch.
- **inlinks**: provides information on the Techcrunch posts' linking (incoming links). There are 193,808 records in the table.
- **posts**: there are 19,464 Techcrunch posts with metadata (blogger ID, publication date, number of comments, etc.).
- **comments**: contains all of the comments left on Techcrunch articles (746,561 records).

Table 1 Top 10 influential bloggers

Id	Name	MEIBI	MEIBIX
2	Michael Arrington	114	114
5	MG Siegler	66	65
3	Erick Schonfeld	56	57
1	Jason Kincaid	43	43
11	Robin Wauters	37	37
21	Guest Author	36	36
7	Leena Rao	24	24
53	Mark Hendrickson	24	24
4	John Biggs	22	22
19	Paul Carr	18	18

In this database, we find bloggers are well organized following two metrics, the indicators used to identify the blogger's influence are (MEIBI) and the extended MEIBI (MEIBIX) [42]. Metrics consider all factors, and they support the time-aware identification of influencers. However, these indicators assign a unique value to each blogger, and they cannot directly distinguish between influence and productivity.

We can even identify influencers following our latest approach [43]. Indeed, we used Google's PageRank algorithm to classify influencers based on several parameters (followers, following, retweet, etc.). The problem why we didn't use PageRank in this work is because our last approach so far is not stable and needs further improvements for accuracy.

In the following, we can easily extract from the database and identify the most influential bloggers. In our application, we filtered the data based on previous metrics and extracted the top 10 influencers from the dataset. Table 1 shows the top 10 influencers on the dataset. Coincidentally, the top 10 influencers are the same for both metrics and also have the same values.

After collecting the influencer publications, we applied the LDA algorithm to the publication content. The process of this phase is based on data cleaning. Data cleaning is essential for generating useful topic models. The following steps as shown in Fig. 5 are common to most natural language processing (NLP) methods:

- **Tokenization**: convert a document into its atomic elements.
- **Stopwords**: delete meaningless words.
- **Stemming**: combine words with the same meaning.

As an application and to generate an LDA model, we need to understand how often each term occurs. To do this, we need to build a matrix, our program is based on python language, and in our case, we used a package called gensim.

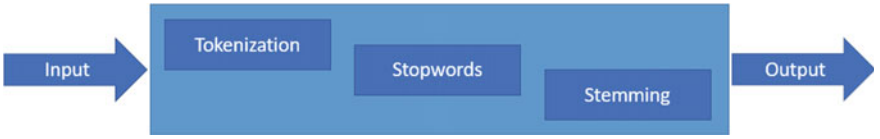


Fig. 5 NLP text preprocessing

Our LDA model is now stored. We can review our topics. Our application allows you to compare the content of influencers with the entire database. Therefore, one finds like the result the following results.

Topic	Influencers	All dataset
0	profile: 0.117	launches: 0.049
	google: 0.083	search: 0.030
	music: 0.018	launch: 0.022
	features: 0.014	microsoft: 0.017
	flickr: 0.010	com: 0.016
1	new: 0.045	web: 0.053
	yahoo: 0.038	million: 0.035
	video: 0.030	social: 0.022
	live: 0.019	blog: 0.021
	beta: 0.018	raises: 0.016
2	launches: 0.052	google: 0.081
	search: 0.030	yahoo: 0.048
	week: 0.018	myspace: 0.014
	microsoft: 0.016	adds: 0.012
	us: 0.016	amp: 0.012
3	quot: 0.040	profile: 0.102
	techcrunch: 0.037	new: 0.042
	online: 0.022	quot: 0.037
	blog: 0.019	video: 0.028
	digg: 0.017	facebook: 0.020
4	web: 0.048	techcrunch: 0.036
	million: 0.028	week: 0.024
	social: 0.023	update: 0.024
	launch: 0.022	beta: 0.021
	com: 0.020	online: 0.018

We have defined our system to extract five clusters as topics and each cluster must give us the 5-top keywords. As result, we obtained keywords that are repeated for the two treatments “influencers and the whole database,” so we also found that these keywords also have similar values for the two applications. The problem that we can see is the topics are not similar. This is because the word occurrence number is decreased in the influencer database. This allows changing the location of topic keywords instead of changing their classification.

6 Conclusion and Future Works

The Internet helps to boost demand for developing applications and services for customers all over the world. However, many studies have been used to focus on social media datasets. Therefore, all these current researches require a lot of time and analysis to profile the target topics with their keywords. In this paper, we propose to profile keywords due to influencers' posts on social media. However, our system suggests starting by identifying influencers from social media. Then, collect the important and actual posts from each influencer based on the LDA algorithm. Our results show that our proposal is efficient and pertinent.

For the next work, we have the objective of analyzing and extracting the keywords by cluster in real time, also giving a label for each cluster. This approach will help us to profile and predict top events in social networks.

References

1. Ahmad A (2011) Social network sites and its popularity. Int J Res Rev Comput Sci 2(2):522–526
2. Mei Q, Ling X, Wondra M, Su H, Zhai C (2007) Topic sentiment mixture: modeling facets and opinions in weblogs. In: Proceedings of the 16th international conference on world wide web, May 2007
3. Pandian AP (2021) Performance evaluation and comparison using deep learning techniques in sentiment analysis. J Soft Comput Paradigm (JSCP) 3(02):123–134
4. Wyrwoll C (2014) User-generated content. In: Social media. Springer Vieweg Wiesbaden
5. Huang M, Zou G, Zhang B, Gan Y, Jiang S, Jiang K (2018) Identifying influential individuals in microblogging networks using graph partitioning. Expert Syst Appl
6. Blei DM (2012) Probabilistic topic models. Commun ACM 55(4):77–84
7. Crain SP, Zhou K, Yang SH, Zha H (2012) Dimensionality reduction and topic modeling: from latent semantic indexing to latent Dirichlet allocation and beyond. In: Mining text data. Springer, Boston
8. Ghorab MR, Zhou D, O'Connor A, Wade V (2013) Personalised information retrieval: survey and classification. User Model User-Adapt Interact 23(4):381–443
9. Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. J Mach Learn Res 3:993–1022
10. Himelboim I, Smith MA, Rainie L, Shneiderman B, Espina C (2017) Classifying twitter topic-networks using social network analysis. Social Media + Society
11. Yang MC, Rim HC (2014) Identifying interesting twitter contents using topical analysis. Expert Syst Appl
12. Zhao WX, Jiang J, He J, Song Y, Achananuparp P, Lim E, Li X (2011) Topical keyphrase extraction from twitter. In: Proceedings of the 49th annual meeting of the association for computational linguistics, Portland
13. Eke CI, Norman AA, Shuib L, Nweke HF (2019) A survey of user profiling: state-of-the-art, challenges, and solutions. IEEE Access 144907–144924
14. Elbaghazaoui BE, Mohamed A, Semmouri A (2021) Data profiling over big data area: a survey of big data profiling: state-of-the-art, use cases and challenges. In: Intelligent systems in big data, semantic web and machine learning. Springer
15. Easley D, Kleinberg J (2010) Networks, crowds, and markets: reasoning about a highly connected world. Cambridge University Press
16. Cook DJ, Holder LB (2006) Mining graph data. Wiley

17. Cossu JV, Labatut V, Dugue N (2016) A review of features for the discrimination of twitter users: application to the prediction of offline influence. In: Social network analysis and mining. Springer
18. Zengin Alp Z, Gunduz Oguducu S (2018) Identifying topical influencers on twitter based on user behavior and network topology. Knowl-Based Syst
19. Agarwal N, Liu H (2008) Blogosphere: research issues, tools and applications. ACM SIGKDD Explor 10(1):18–31
20. Gill KE (2004) How can we measure the influence of the blogosphere? In: Proceedings of workshop on the weblogging ecosystems: aggregation, analysis and dynamics, pp 1–5
21. Gruhl D, Guha R, Kumar R, Novak J, Tomkins A (2005) The predictive power of online chatter. In: Proceedings of the ACM SIGKDD international conference on knowledge discovery in data mining, pp 78–87
22. Java A, Kolari P, Finin T, Oates T (2006) Modeling the spread of influence on the blogosphere. In: Proceedings of ACM world wide web conference, pp 1–7
23. Cross R, Rice RE, Parker A (2001) Information seeking in social context: structural influences and receipt of information benefits. IEEE Trans Syst Man Cybern Part C 31(4):438–448
24. Adar E, Adamic LA (2005) Tracking information epidemics in blogspace. In: Proceedings of IEEE/WIC/ACM international conference on web intelligence, pp 207–214
25. Zhou Y, Davis J (2006) Community discovery and analysis in blogspace. In: Proceedings of ACM world wide web conference, pp 1017–1018
26. Louta M, Varlamis I (2010) Blog rating as an iterative collaborative process. Semant Adapt Pers Serv 279:187–203
27. Agarwal N, Liu H, Tang L, Yu PS (2008) Identifying the influential bloggers in a community. In: Proceedings of the ACM international conference on web search and data mining, pp 207–218
28. Steinskog A, Therkelsen J, Gamback B (2017) Twitter topic modeling by tweet aggregation. In: Proceedings of the 21st Nordic conference of computational linguistics, Gothenburg, pp 77–86
29. Giannakis M, Dubey R, Yan S, Spanaki K, Papadopoulos T (2020) Social media and sense-making patterns in new product development: demystifying the customer sentiment. Ann Oper Res. https://doi.org/10.1007/s10479-020-03775-6
30. Alghamdi R, Alfalqi K (2015) A survey of topic modeling in text mining. Int J Adv Comput Sci Appl
31. Kottursamy K (2021) A review on finding efficient approach to detect customer emotion analysis using deep learning analysis. J Trends Comput Sci Smart Technol 3(2):95–113
32. De Veirman M, Cauberghe V, Hudders L (2017) Marketing through Instagram influencers: the impact of number of followers and product divergence on brand attitude. Int J Advert 36(5):798–828
33. Kaur R, Bhardwaj V (2021) A GUI for multimodal sentiment analysis on Gurmukhi script. In: Innovative data communication technologies and application. Springer, Singapore, pp 69–81
34. Nelson JL, Kim SJ (2020) Improve trust, increase loyalty? Analyzing the relationship between news credibility and consumption. Journal Pract 1–18. https://doi.org/10.1080/17512786.2020.1719874
35. van Dam S, van Reijmersdal E (2019) Insights in adolescents' advertising literacy, perceptions and responses regarding sponsored influencer videos and disclosures. J Psychosoc Res Cyberspace 13(2), Article 2. https://doi.org/10.5817/CP2019-2-2
36. Zimmermann D et al (2020) Influencers on YouTube: a quantitative study on young people's use and perception of videos about political and societal topics. Curr Psychol. https://doi.org/10.1007/s12144-020-01164-7
37. Abilhoa WD, de Castro LN (2014) A keyword extraction method from twitter messages represented as graphs. Appl Math Comput 240:308–325. https://doi.org/10.1016/j.amc.2014.04.090
38. Campos R, Mangaravite V, Pasquali A, Jorge AM, Nunes C, Jatowt A (2018) A text feature based automatic keyword extraction method for single documents. In: ECIR 2018. LNCS, vol 10772. Springer, Cham, pp 684–691. https://doi.org/10.1007/978-3-319-76941-7_63

39. Hasan KS, Ng V (2014) Automatic keyphrase extraction: a survey of the state of the art. In: Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 1: long papers), vol 1, pp 1262–1273
40. Nguyen TD, Luong MT (2010) Wingnus: keyphrase extraction utilizing document logical structure. In: Proceedings of the 5th international workshop on semantic evaluation. Association for Computational Linguistics, pp 166–169
41. Akritidis L, Katsaros D, Bozanis P (2011) Identifying the productive and influential bloggers in a community. IEEE Trans Syst Man Cybern Part C Appl Rev 41(5):759–764
42. Akritidis L, Katsaros D, Bozanis P (2009) Identifying influential bloggers: time does matter. In: Proceedings of IEEE/WIC/ACM international conference on web intelligence, pp 76–83
43. Elbaghazaoui BE et al (2022) Optimized influencers profiling from social media based on machine learning. In: Maleh Y et al (eds) Advances in information, communication and cybersecurity, vol 357. Springer International Publishing, pp 43–56. https://doi.org/10.1007/978-3-030-91738-8_5

Comprehensive Survey on Fire Detection with Machine Learning and Deep Learning Models

G. Shankar and M. Kalaiselvi Geetha

Abstract Fire is an unpredictable dangerous event to the environment and the public. It may cause human and non-human belongings very rigorously. In recent days, many unpredictable events in the world are identified and controlled by computer vision-based machine learning approaches. Among these events, this looks forward to predicting the fire event and saving the human and non-human valuable belonging. Fire detection is a challenging task for the researchers to predict the root cause and alert the nearby ones. In this article, many algorithms have been addressed for fire detection methods from the past two decades that how the computer vision techniques have been grown day by day to predict the fire even. Even though, a formal survey on recent trends to fill the gaps is identified with the last decades. Countless fire detection strategies are developed through machine learning, deep learning, computer vision, etc., with images, video, and sensors. This article discussed a detailed survey that how the fire detection process becomes popular in the research area in recent days. In addition, a proposed new model is to rectify the issues in existing approaches.

Keywords Computer vision · Fire detection · Machine learning · Deep learning · Images · Video · Sensors

1 Introduction

Natural disasters are unpredictable and dangerous events in the real world. Fire is the most crucial one among these. It may cause large-scale damages to the human and non-human valuable belongings [1]. To avoid such damages, the detection of fire is essential. However, a human cannot able to predict when, why, where, and how this event occurs. This may happen with human presence or absence, and the human absence may cause major damages to the valuable properties [2]. So, predicting fire events and alerting them are helping to save their precious properties.

G. Shankar (✉) · M. Kalaiselvi Geetha
Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamilnadu, India
e-mail: shankarganesan1985@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_30

Fire detection is mainly deployed with two categories, (i) traditional vision-based approaches and (ii) sensor-based approaches [3]. These are essentially detecting fire with the human absence and alerting the people who are nearby. But, both have a significance of consequences. Traditional vision-based approaches [4] are involving through images or video. In this, video captured by the surveillance cameras and traditional approaches has been applied. Here, some features are highlighted like color, edges, motion, etc. Sensor-based approaches also give better results, but it fails to produce such features highlighted with a vision-based system [5]. Coverage of large-scale areas is a prominent issue with sensor-based systems. Because, sensors cover a small area that is inadequate, whereas surveillance is an adequate wide area.

The issues identified with sensor-based systems are requiring high expense for large-area coverage because sensors cover only limited space, whereas large space may require many sensors [6]. Sensors are reacting with physical conditions to alert the electrical properties. These help us to get information about the environment by relying on electric systems to analyze, capture, and rely on. A sensor simply converts the motion, sound, heat, and light into an electrical signal. These signals are converted to binary codes with some interface and its process by the computer.

Figure 1 depicts that a surveillance camera is focused wide range of areas and monitoring the visual scene changes. Subsequently, the video acquires as input to convert into frames, these frames are originally large or small in size. Therefore, feature extraction may be intricate with an unordered sequence of frames. To resolve this, every frame is resized with ordered and allows to feature extraction. Numerous feature extraction algorithms were developed in the machine or deep learning algorithms and should choose the optimum one. For later object detection, a classification or clustering algorithm is applied to detect the fire in a scene by using optimal features. Once the fire is detected, peoples are alert in three ways. Owner's or nearby fire station receives an alert message with location or address via mobile or e-mail and alarm sound that alerts the nearby peoples.

Research Objective: This article mainly aims to review how fire detection works and developed with numerous deep learning and machine learning approaches in the

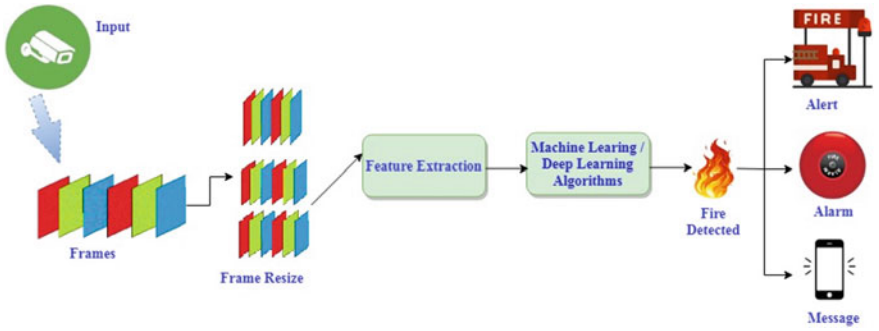


Fig. 1 Block diagram of fire detection system

last decades. The objective of the research on fire detection is discussed and listed as follows,

- Why the fire detection system is to develop with a computer vision approach?
- How this system is useful to society?
- Fire detection systems with many algorithms have been developed, but still, there is a gap to prove the better accuracy due to various aspects.
- A novel model is trying to identify the difficulties to improve the detection system to work well among other models.
- Many models are developed under machine learning and deep learning models, but they have some false alarm rates due to the fire-like objects and other non-suitable features.
- Some non-fire objects have the same features like fire color, shape, motion, etc., so this model addressed such problems with different models and aimed to improve the accuracy.

Our Contribution: The literature survey is analyzed with various articles to identify the research gaps and possible methodologies to rectify them. Many models have been studied with this literature survey. It focused to reduce the false alarm rate due to the fire-like object. Because such fire-like objects are identified as fire objects, it is non-fire objects. These non-fire objects are having some fire features that should identify and remove.

This paper is organized in the following way, Sect. 2 deals with a literature survey over various existing and trending approaches in fire detection, Sect. 3 describes that some of the deep learning models are used in object detection and classification, Sect. 4 explores the comparison with features and existing models. Section 5 deals with application areas of fire detection systems, and Sect. 6 mentioned that what are the challenges that are still there in the fire detection system. Section 7 gives some research directions in detail. Section 8 gives a conclusion about this literature.

2 Literature Survey

Three categories were found in the literature from, Muhammad et al. [7], generally, fire detection through the light camera is used one among the following, (i) pixel-level, (ii) blob-level, and (iii) patch-level methods. Liu et al. [8] and Chen et al. [9] describe a pixel-wise feature like color and flicker which are used in pixel-level methods makes faster but fails in performance because it can be biased easily. Töreyin et al. [10] proposed a blob-level method to achieve high performance at the pixel level. But, it has numerous shapes of fire so difficult to train their classifier. Choi et al. [11, 12] a patch-level method improves the performance than the existing categories, but accuracy is affected due to many outliers.

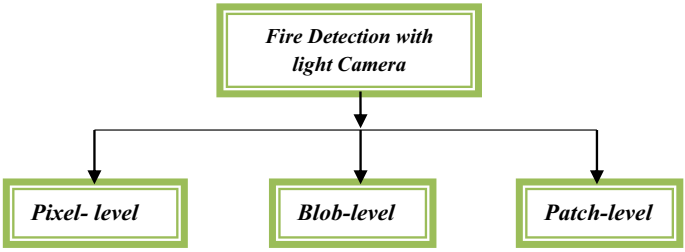


Fig. 2 Types of fire detection with a light camera

A major challenge is still accuracy is not to meet the satisfaction of the researchers, and they attempt color and motion-based features to improve the fire detection accuracy. Hahemzadeh et al. [13] define that fire detection with computer vision-based approaches involves two main categories, color-based and motion-based (Fig. 2).

Hahemzadeh et al. [13] describe that the fire detection technique investigations start from color models commonly set of images to determine the region of color space that includes colors similar to the fire flames. Color-based methods aim at the color characteristic of fire, and these methods have a common shortcoming which is illumination; the fire-like objects also have the same color characteristics, so it makes high false detection in a scene. Generally, the RGB and HSI color models on dynamic behavior and irregular flame for fire detection work slowly. So, Celik et al. [14] and Kim et al. [15] have proposed a background subtraction method and statistical RGB model; in this, region color space is identified from the group of images, if the pixel value is corresponding with the region value, then it is found to be fire quickly. Yu et al. [16] used the HSV color model for filtering the fire color pixels with a set of rules experimentally. HSI color model was introduced to simplify the definition of rules for the designer. Qi et al. [17] implement a comparative methodology to utilize where a total fire grid has been characterized by consolidating the RGB shading and HSV immersion specifically, beginning from the presumption that the green part of the fire pixels has a wide scope of changes contrasted and red and blue ones, and this technique assesses that the spatial shading variety in pixel esteems to recognize non-fire moving items from uncontrolled flames. The above all the approaches have common limitations that they are passionate about the brightness changes. It causes high false positives, because of shadows' presence and different tonalities of red colored. This issue can be recovered by the YUV color model. Marbach et al. [18] combined the YUV color and motion features to predict the fire and non-fire pixels. Continuing the color models, an additional method developed by Han et al. [19] detected flame in the tunnel by comparing the color features and video frames and explored YCbCr color space for separating chrominance component from luminance. Liu et al. [20] proposed a YCbCr color model for flame detection. This method gives good accuracy to detect flames with a small distance and large size of fires.

Kim et al. [21] introduce a statistical RGB color space and background subtraction to detect the fire pixels in a complex background. Object motions in a sequence of images or object recognition are rapidly dealing with motion-based models. A

sequence has numerous frames to extract the motion information and recognition of higher-level motion like walking or running which is more complex that cannot be understood by a few frames. Foggia et al. [22] approach involves baffle movements in a scene whether fire flame is distinguishing from another movable object.

Consequently, Ko et al. [23] presented the spatial–temporal motion-based features like motion orientation. Unlike some methods like optical flow Mueller et al. [24] and dynamic texture Dimitropoulos et al. [25], they assist the flame detection. P. V. K. Borges et al. [26] introduced features such as color, skewness, surface coarseness, area size, and boundary roughness which are evaluated by a Bayesian classifier to analyze the frame-to-frame changes. Rafiee et al. [27] aimed to reduce the false positive rates by using 2D multi-wavelet analysis which evaluates the shape and energy variations to improve the thresholding on the color.

References [28–30] the motion-based approaches have some limitations especially when other non-fire movable objects have appeared in the scene. Motion and color features are combined to enhance the performance of fire detection, and a variety of approaches have been tried. Initially, the color-based feature identifies the candidate fire regions from color features. Subsequently, motion-based features were extracted from candidate fire regions to classify whether fire or non-fire regions. Another color-based approach is proposed by Habiboğlu et al. [31]; in this, an SVM classifier is used to detect fire with high accuracy with closer distance. But, it fails and gives a pitiable performance with the larger distance or less amount of fire. Di Lascio et al. [32] combined the color and motion features for flame detection in surveillance videos. Poobalan et al. [33] and Gomes et al. [34] proposed a model based on rules, or it requires the discriminative feature for fire detection.

Since, the above discussion describes that the color information is simple to configure and successes in infertile areas, where no new object is entered in a scene. This approach makes trouble with normal areas such as red-colored objects in a scene originate as fire because of color domination. These traditional methods have some limitations in that the color segmentation, or color thresholding is used to detect the candidate fire regions, but they are not robust and increase the number of false positive and false negative. This is reduced with a large number of frames of irregular motion and uninformative features with the existing classification problems in machine learning.

Several papers are dealing the fire detection with surveillance which is growing exponentially. But recent days, deep learning methods are ongoing toward object detection and classification, etc. Deep learning has different methods to resolve those problems, and many differences were identified with vision-based fire detection. First, features are not extracted by the experts rather than those are captured by the network after training the huge of diverse data. Next, detection or classification can be received by the features in a similar neural network.

Sebastien et al. [35] proposed CNN to detect fire in a surveillance video where features are trained with MLP neural net. Another CNN approach is proposed by Zhang et al. [36], and it describes cascaded fashion. In this, initially, a complete image is tested on a global image-level classifier, then after if the fire is detected, fine-grained path classifier is used. Muhammad et al. [37] illustrate fine-tuned CNN method for

fire detection on surveillance systems, and this is proficiently well for fire detection, semantic understanding, and localization of the video which are encouraged from the SqueezeNet [38] architecture. References [39–41] initiate the fire detection from CNNs and enlarge the self-learning algorithm in the collection of fire features. In addition, Refs. [42–44] models from deep learning like ResNet, AlexNet, Inception, VGG, etc., are developed for smoke and flame detection algorithms. Namozov et al. [45] worked with time series information into algorithms. Simultaneously, Wonjae et al. [46] developed an algorithm to detect smoke and flame from reformed VGGNet.

In recent days, many models have been proposed with IoT systems. Sathesh [47] proposed a computer vision on an IoT-based patient reference system that performs a kind of algorithm to identify the performance while data transmission. Jeena Jacob et al. [48] have developed an application for image recognition that supports deep learning models for IoT systems. Akey sungheetha et al. [49] illustrate fire detection and real-time monitoring with IoT and cloud-based drones. Vince Leona et al. [50] introduce a temperature monitoring system in the glass manufacturing area, and it uses optisystem simulation. Srinivas et al. [51] proposed a model for fog detection with weather sensors.

Various issues were addressed with a detailed review on fire detection like illumination, fire-like objects, color tonalities, large distance, etc.; each has identical or non-identical solutions based on the model or algorithm that they proposed. But, still producing false alarm rates is a major challenging task among those algorithms. To reduce the false alarm rate, they focused this fire detection method on deep learning models. Section III describes some deep learning models to recover the issues addressed by the survey. Deep learning models give high accuracy than machine learning models in object detection. Various models are available in deep learning, but few are introduced here and listed the pros and cons of the models. Feature extractions are easier than other learning models. After all, it does not require any manual involvement, because mostly it is pre-trained by the model, so it is a time-consuming process.

3 Deep Learning Models for Classification

Deep learning is incorporated with machine learning, and its structure and functions are stimulated from artificial neural networks. DL involves automatic feature extraction, whereas machine learning requires manual feature extraction. In machine learning algorithms, many feature extraction and selection algorithms are proposed in the last decades. This feature selection and extraction are the most challenging task for the researchers. Hence, DL provides different kinds of architecture for classification problems that take care of the automatic feature extraction; here, few are listed with some of the DL models involved to detect the fire event. The following are some of the models used for classification problems from DL.

3.1 VGG16 Model

VGG16 is a CNN model introduced by K. Simonyan and A. Zisserman, and this model gains high accuracy with top-five tests with ImageNet which has datasets over 14 million images consisting of 1000 classes. It replaces the AlexNet with large kernel-sized filters (11 and 5 in the first and second convolutional layers, respectively) with multiple 3 × 3 kernel-sized filters one after another.

Figure 3 depicts VGG16 architecture that has a total of 16 layers and six blocks, and the first five blocks perform convolution and pooling operations. The final block consists of three dense or fully connected layer which consists of neurons with two layers, and the last layer consists of classes. VGG16 is trained with ImageNet weights over 14 million images belonging to 1000 classes.

Figure 4 is the detailed architecture of the VGG16 model that describes that the input is taken as an image with the size 224 × 224 RGB image. Input is not only an RGB image, instead it may be grayscale or binary image. 224 × 224 × 3 indicates that the input image contains height and width as 224 and 3 represent RGB.

In the first convolution, the block consists of two convolution layers with 64 channels of 3 × 3 kernels with padding of same followed by a pooling layer. The

Fig. 3 Convolution layers in VGG16

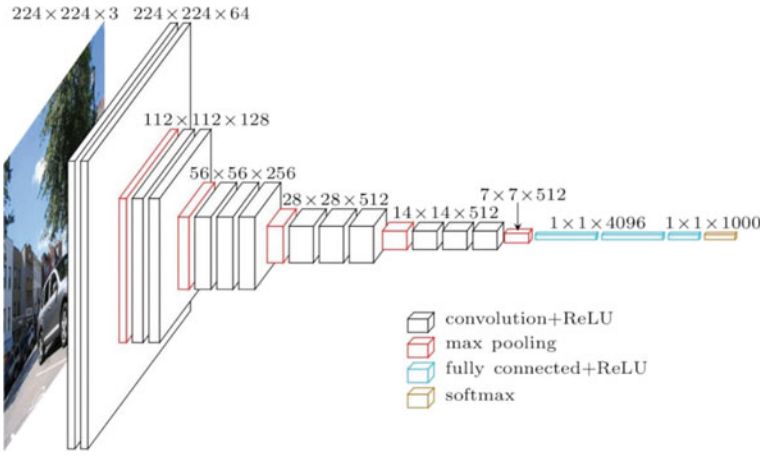


Fig. 4 Architecture of VGG16 [52]

parameter padding has two values, same and valid. The valid allows the changes in the spatial domain, whereas the same does not. Pooling layer selects max pool operation of 2 × 2 pool size and with a stride of 2 × 2. Next two convolution layers are with 112 × 112 of 128 channels with 3 × 3 kernel with max pool operation. Subsequently, the architecture continues with 3 convolution layers with the size of 56 × 56 and 256 channels of 3 × 3 kernel and same padding continue with next 3 layers of 28 × 28 with 512 channels of 3 × 3 kernel and same padding next three layers of 14 × 14 with 512 channels and same padding. The next two layers consist of 7 × 7 with 512 channels and the same padding with max pool operation.

Finally, all the convolution layers are trained, and the last three layers are dense layers or fully connected layers. The first dense has 4096 neurons with activation function, the second dense is also functioning like the same, and the last dense layer contains the 1000 channels to classify the input images.

Figure 5 illustrates those different versions of the VGG model, VGG19 provides more parameters than all other models, but VGG16 gives low error rate than all other VGG models. Generally, the concept of VGG16 and VGG19 is the same, but VGG19 has three convolutional layers additionally. So, it has more parameters than the other models.

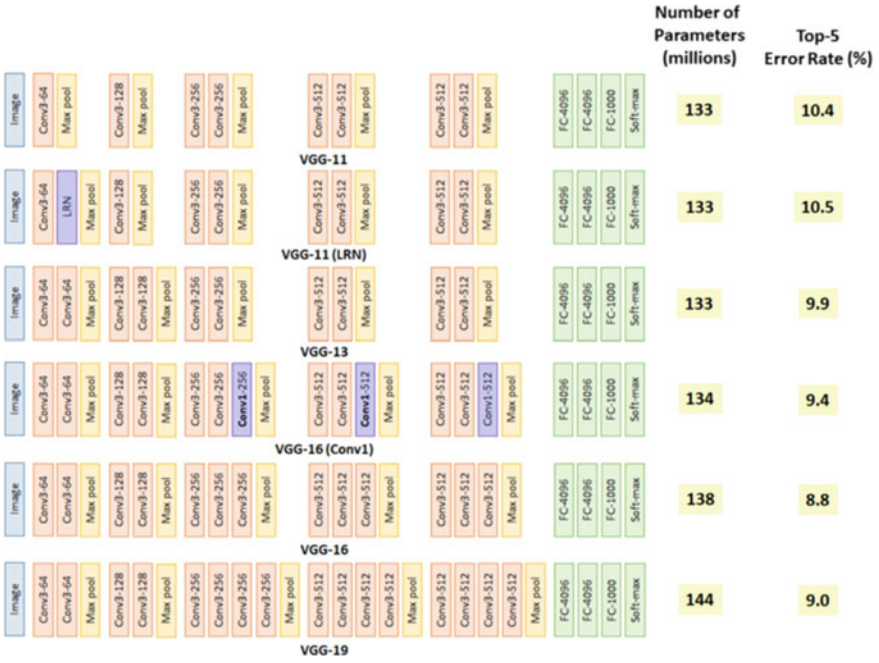


Fig. 5 Different VGG models and error rates [53]

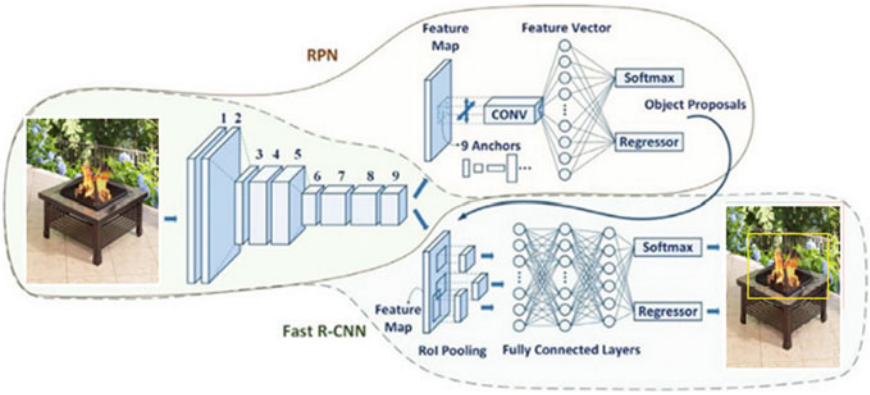


Fig. 6 Faster R-CNN architecture [54]

3.2 Faster R-CNN

Faster R-CNN is developed by Ross Girshick, Shaoqing Ren, Kaiming He, and Jian Sun in 2015, and this architecture helps to detect the object in real time. This model is composed of three major parts, convolution layers, RPN, and bounding boxes prediction.

3.2.1 Convolution Layers

In this stage, filters are trained to extract the features from the image; for example, in our fire detection system, the fire features from the images are extracted at this stage, then those filters are trained to learn the shape and colors of the fires. These layers are composed of convolution layers, pooling layers, and fully connected used to classify or detect an object.

3.2.2 Region Proposal Network (RPN)

RPN is the second stage of Faster R-CNN which takes feature maps from the convolution layers to produce region proposals with pre-trained with the absence of any external mechanism. It detects the object and bounds the box on those objects (Fig. 6).

3.2.3 Classes and Bounding Boxes Prediction

The regions are taken as input from the RPN and pass through fully connected layer networks to predict the object and bounding boxes on the object.

3.3 Long Short-term Memory (LSTM)

LSTM is one of the recurrent neural networks capable of learning order dependence with sequence prediction problems. LSTM is involved in complex problem domains like deep learning, machine translation, object recognition, etc.; it remembers the values over arbitrary levels and works well for classifications and prediction in time series. Supposed to predict the sequence once 1000 intervals instead of 10, the RNN may forget the initial point but LSTM REMEMBERS.

Figure 7 shows how the RNN cell works. This cell takes two inputs X_t and h_{t-1}, one is output z_t from the last hidden state, and another is observation at time t and the hidden state called h_t. But, here none of the information is available with regards the past one to REMEMBER. But, in Fig. 8, the recursive nature of the cell is indicated by the looping arrows. This loop allows the information from preceding intervals and stores it in the LSTM cell. Forget gate which is placed below the cell is used to modify the cell state with adjustment of the input modulation gate. The terms used in the following equations are listed below,

i_t—input gate

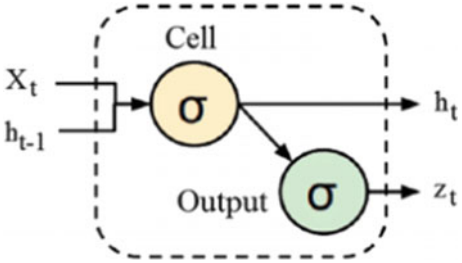


Fig. 7 RNN cell [55]

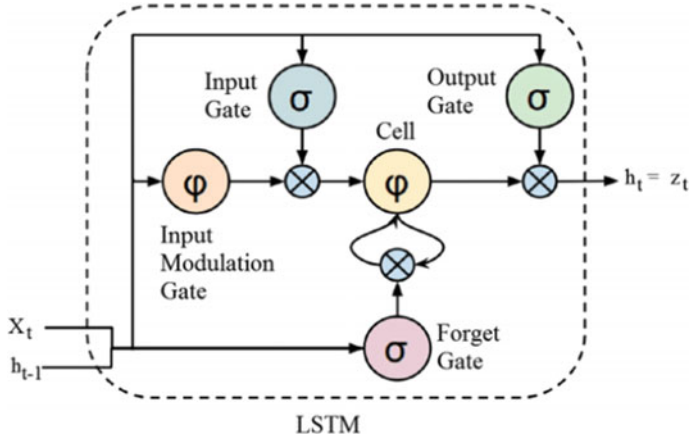


Fig. 8 LSTM cell [55]

f_t—forget gate
o_t—output gate
σ—Sigmoid function
w_f—weights for the respective gates
$h_{t\text{-}1}$—output of previous LSTM block
x_t—input at the current timestamp
c_t—cell state
b_t—biases for the respective gates.

$$c_t = f_t * c_{t-1} + i_t * \overline{c}_t \tag{1}$$

Equation 1 represents the cell state, the previous cell state forgets, and the current forget is multiplied and adds the new information for the output of the input gate.

$$f_t = \sigma\left(w_f\left[h_{t-1}, x_t\right] + b_t\right) \tag{2}$$

Equation 2 depicts that the forget gate is generally called remember vector. The forget gate output gives the cell state which information to forget by multiplying 0 to a position in the matrix. If it is 1, the information is kept in the cell state.

$$i_t = \sigma\left(w_i\left[h_{t-1}, x_t\right] + b_i\right) \tag{3}$$

Equation 3 represents that input gate equation which is also called save vector. It determines which information should enter into a cell state or long-term memory. The activation functions play an important role for each gate.

$$o_t = \sigma\left(w_o\left[h_{t-1}, x_t\right] + b_o\right) \tag{4}$$

Equation 4 indicates the output gate or focus vector to determine the output of all possible values which is moving forward to the next hidden state.

$$h_t = o_t * \tanh(c_t) \tag{5}$$

Equation 5 denotes that the hidden state or working memory determines what information should carry to the next sequence.

3.4 Inception V3

Figure 9 shows the architecture of the InceptionV3 model. In convolutional architecture from deep learning, InceptionV3 is the third version introduced by Google. This model is trained by the dataset consisting of 1000 classes from the ImageNet over one million training images. InceptionV3 is familiar for transfer learning, and it allows you to retain the final layer of an existing model, resulting in a significant

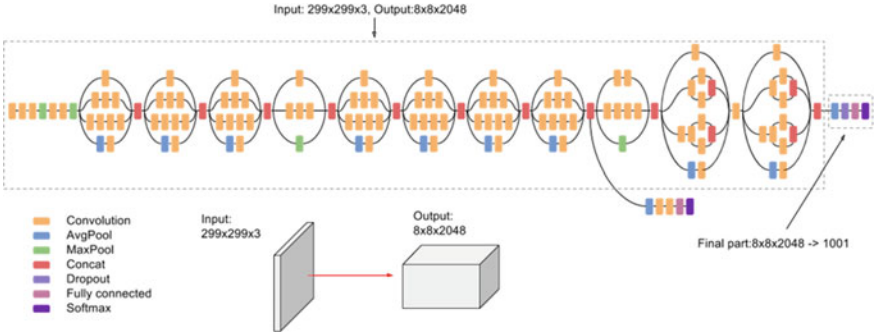


Fig. 9 Architecture of InceptionV3 [56]

decrease in not only training time, but also the size of the dataset required. Being able to retrain the final layer means that you can maintain the knowledge that the model had learned during its original training and apply it to your smaller dataset, resulting in highly accurate classifications without the need for extensive training and computational power. InceptionV3 model consists of the following building blocks convolutions, average pooling, max pooling, concats, dropouts, and fully connected layers. The objective and working functionality of all the above building blocks are explained in the previous models, so the novel model will discuss the advantages over other models here. Label smoothing, factorized convolutions, batch normalization, and use of auxiliary classifier to propagate the label information are improvements in InceptionV3.

3.5 YOLO

"You Only Look Once", YOLO, is a family of convolutional neural networks that achieves the results with a single end-to-end model for object detection in real time. This model is developed by Joseph Redmon et al. in 2015. This model follows the single convolutional neural networks to the full image that crack the inputs into a grid of cells, and each cell predicts the bounding boxes and objects classification.

This model has been refined with three versions, initially, the first version is named YOLOv1, the second one YOLOv2, and the third is YOLOv3. The first version is introduced with the general architecture, and the design and predefined anchor boxes to improve the bounding box proposal are made with the second version, and the third involved in further that refined the model architecture and training process. YOLO achieves high accuracy but not more than R-CNN. The algorithm "You Only Look Once", YOLO, means only one propagation is passing in a forward direction through the entire neural network to make the prediction. This object detection algorithm detects the objects only once, then it recognizes the object with the bounding boxes.

YOLO model has various benefits over other object detection models:

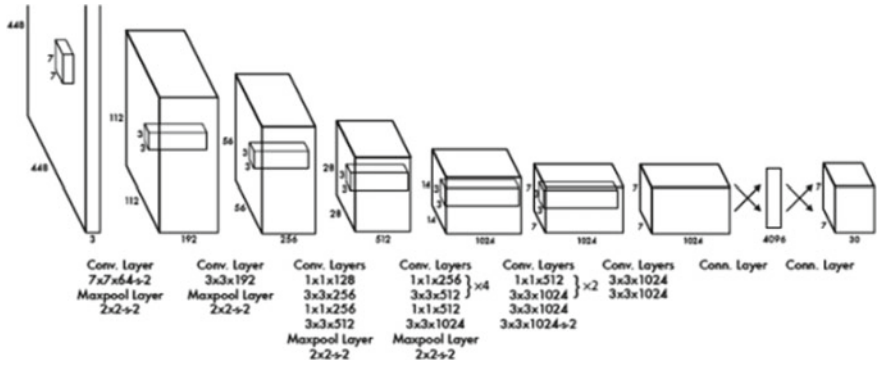


Fig. 10 YOLO architecture [57]

- it is extremely fast
- it implicitly encodes contextual information about classes as well as their appearance.
- YOLO learns generalizable representations of objects.

These models are most familiar in recent days for object detection and classification. So, finally, the above deep learning models are recommended for fire detection systems based on their better performance in object detection and classification method. This article gives an overview architecture and working process of the above deep learning models (Fig. 10).

4 Comparison with Features and Existing Approaches

The following Table 1 shows that the different features and models or algorithms are used to develop various fire detection methods with machine learning and deep learning models. Features such as color, shape, and motion are considered to give better performance.

The main aim of this proposed work is to combine machine learning and deep learning models to give better accuracy. Feature extraction is done by deep learning model, and those features are used in machine learning models for implementation.

5 Applications of Fire Detection

The fire detection system is used in various application areas. This section explains how this system is working fine with those areas. Some of the fire detection areas are described as follows.

5.1 Conventional Fire Alarm Application

Conventional fire alarm applications required physical cabling to interconnect the different call points, detectors, then signals are wired back to the main control unit. This uses simple zone feedback means that every detection requires its wiring circuit. All zones are indicated at the fire alarm control panel, and zone indicates that allocating the suitable area to locate the alarm.

5.2 Addressable Fire Alarm Applications

Addressable fire alarm applications acted like conventional fire alarm, but here, each detector is given a set address called Dual Inline Package (DIP) switch, then the control panel finds the call point or exact detector which initiates the alarm. This is a loop wired circuit maximum of 99 devices which is connected per loop. Finally, the control unit is deciding about whether the fire has happened or not, so the control unit is the last extent to trigger the alarm if a fire happened.

5.3 Intelligent Fire Alarm Applications

The last two applications are not "intelligent" because only the control unit is deciding whether the fire is detected. In intelligent fire alarm, each detector is incorporated with a computer system to evaluate the environment surrounding it and contact the control unit to detect the fire. But, this system is more complex than the other two fire systems.

5.4 Wireless Fire Alarm Systems

Wireless fire alarm is the effective replacement of the traditional fire alarm systems. It uses a sensor to detect the fire and need not have any complexity of circuits. Radio communications are used to interconnect the devices and sensors with the controller. These are suitable for buildings to give alarm when the fire is invoked in the sensor area. Sensors are most suitable for small areas, and it requires less cost.

6 Challenges on Fire Detection System

Fire detection systems with video surveillance have more challenges to give better accuracy. In this section, some of the challenges are listed as follows.

- The fire detection system has challenging tasks like false alarm rate, smoke controls, false sensing by sensors, etc.
- Extracting features with low-quality images or video is more critical, and features like motion, color, shape, edges are limited.
- Fire detection with sensor also bounded with small-scale area when if the detection system is covered with larger with the traditional fire system.
- Extracting limited features with high-quality video or images also does not provide better solutions.
- Some machine learning algorithms with color models are detecting non-fire objects as fire objects.
- Because these models are used only color as features, here some objects like the rising sun, illumination with some objects, and other fire color objects are detected as fire.
- In a machine learning algorithm, extracting the features from the image and selecting the most distinguishable features are the most challenging task.

7 Research Direction

The traditional fire detection system with video surveillance has more advantages. But still, researchers have to handle some difficulties as follows.

7.1 Complex Background Environment

The objects like fire that may appear in the background would change with light which is detected as fire by mistake. Other natural scenes like sunrise, sunset, etc., actually have the fire color scene has detected as fire. Other than these if the camera jittered is facing the performance in the classification.

7.2 Features Are Changing the Targets

Features in the surveillance process such as size, position, magnitude, shape and changes in time are the most difficult to give better performance in the object detection.

7.3 Occlusion Problems

Occlusion problems are very complex in the process of research. Targets are both partially or fully occluded, multiple targets are occluded mutually, and both of the cases influence the performance of object detection.

7.4 Illegible Market Demand

The fire detection system has illegible market demands in the areas such as traffic, hospitals, education, government sectors, nevertheless not widely used.

8 Conclusion

Various algorithms and models have been identified in machine learning and deep learning. Every model has its specialty to provide better accuracy. In this survey, color models from machine learning give a better solution, but it fails when it gets fire color objects. Some models deal with the motion features, but results are very closer to the color models. It does not give very large deviations. Compared with deep learning models, machine learning methods are more complex to extract the features, whereas deep learning models provide better accuracy than machine learning models. In addition, many deep learning models have been discussed in this literature; in the future, proposed model is expected with better accuracy with deep learning models. There is a chance to combine deep learning and machine learning models to give high results. This article discussed CNN, R-CNN, YOLO, LSTM models to prove the better results in object detection. Finally, it has been concluded that fire detection is to be implemented with CNN with the VGG16 model in the future.

Table 1 Features and models used in machine learning and deep learning methods

Sl.no	Methods	Color model	Shape	Motion	Deep learning	Model/algorithm
1	Khan Muhammad et al.[7]				✓	CNN
2	Chen et al. [9]	✓	✓			K-Medoid
3	Hashemzadeh et al. [13]	✓	✓	✓		ICA K-Medoid
4	Celik et al. [14]	✓	✓			Statistical color models
5	X. Qi et al. [15]	✓				Color models
6	Liu, Z.-G. et al. [20]	✓				YCbCr color model
7	Foggia et al. [22]	✓	✓	✓		Combination of color, shape, and motion
8	Dimitropoulos et al. [25]		✓	✓		Spatio-temporal model and texture model
9	Borges et al. [26]	✓		✓		Probabilistic approaches
10	Rafiee et al. [27]	✓	✓			YUV color model
11	Han et al. [30]	✓		✓		GMM and multi-color
12	Frizzi et al. [35]				✓	CNN
13	Yin et al. [40]				✓	CNN
14	Wonjae et al. [46]				✓	DCNN
15	Joseph Redmon et al. [57]				✓	YOLO
16	Habibo ˘glu et al. [58]	✓	✓			Covariance matrix method
17	Muhammad	✓		✓		Color model
18	Proposed*	✓	✓	✓	✓	CNN and SVM

References

1. Töreyin BU, Dedeoğlu Y, Güdükbay U, Cetin AE (2006) Computer vision based method for real-time fire and flame detection. Pattern Recognit Lett 27:49–58 (Elsevier)
2. Ko BC, Cheong KH, Nam JY (2009) Fire detection based on vision sensor and support vector machine. Fire Saf J 44:322–329
3. Celik T, Demirel H (2009) Fire detection in video sequences using generic color model. Fire

Saf J 44(2):147–158
4. Karasulu B, Karasulu S (2012) Moving object detection and tracking by using annealed background subtraction method in videos: performance optimization. Expert Syst Appl 33–43
5. Chino DY, PAvalhais L, Rodrigues JF, Traina AJ (2015) BoWFire: detection of fire in still images by integrating pixel color and texture analysis. In: 28th SIBGRAPI conference on graphics, patterns, and images, pp 95–102
6. Luo Y, Zhao L, Liu P, Huang D (2018) Fire smoke detection algorithm based on motion characteristics and convolutional neural networks. Multimed Tool Appl 77:15075–15092
7. Muhammad K, Ahmad J, Mehmood I, Rho S, Baik SW (2018) Convolutional neural networks based fire detection in surveillance videos. IEEE, vol 6
8. Liu C-B, Ahuja N (2004) Vision based fire detection. In: Proc. 17th int. conf. pattern recognit. (ICPR), pp 134–137, Aug 2004
9. Chen T-H, Wu P-H, Chiou Y-C (2004) An early fire-detection method based on image processing. In: Proc. int. conf. image process. (ICIP), pp 1707–1710, Oct 2004
10. Töreyin BU, Dedeoğlu Y, Güdükbay U, Çetin AE (2006) Computer vision based method for real-time fire and flame detection. Pattern Recognit Lett 27:49–58
11. Choi J, Choi JY (2016) An integrated framework for 24-hours fire detection. Proc Eur Conf Comput Vis 463–479
12. Choi J, Choi JY (2015) Patch-based fire detection with online outlier learning. In: Proc. 12th IEEE int. conf. adv. video signal based surveill. (AVSS), pp 1–6, Aug 2015
13. Hashemzadeh M, Zademehdi A (2019) Fire detection for video surveillance applications using ICA k-medoids-based color model and efficient spatio-temporal visual features. Expert Syst Appl (Elsevier-online)
14. Celik T, Demirel H, Ozkaramanli H, Uyguroglu M (2007) Fire detection using statistical color model in video sequences. J Vis Commun Image Represent 18(2):176–185
15. Kim Y-H, Kim A, Jeong H-Y (2014) RGB color model based the fire detection algorithm in video sequences on wireless sensor network. Int J Distrib Sensor Netw 2014. Art. ID 923609
16. Yu C, Mei Z, Zhang X (2013) A real-time video fire flame and smoke detection algorithm. Proc Eng 62:891–898
17. Qi X, Ebert J (2009) A computer vision based method for fire detection in color videos. Int J Imag 2(S09):22–34
18. Marbach G, Loepfe M, Brupbacher T (2006) An image processing technique for fire detection in video images. Fire Safety J 41:285–289
19. Han D, Lee B (2006) Development of early tunnel fire detection algorithm using the image processing. Proc Int Symp Vis Comput 39–48
20. Liu Z-G, Yang Y, Ji X-H (2016) Flame detection algorithm based on a saliency detection technique and the uniform local binary pattern in the YCbCr color space. SIViP 10(2):277–284
21. Kim Y-H, Kim A, Jeong H-Y (2014) RGB color model based the fire detection algorithm in video sequences on wireless sensor network. Int J Distrib Sens Netw 10(4):923609
22. Foggia P, Saggese A, Vento M (2015) Real-time fire detection for video surveillance applications using a combination of experts based on color, shape, and motion. IEEE Trans Circuits Syst Video Technol 25(9):1545–1556
23. Ko BC, Ham SJ, Nam JY (2011) Modeling and formalization of fuzzy finite automata for detection of irregular fire flames. IEEE Trans Circuits Syst Video Technol 21(12):1903–1912. https://doi.org/10.1109/TCSVT.2011.2157190
24. Mueller M, Karasev P, Kolesov I, Tannenbaum A (2013) Optical flow estimation for flame detection in videos. IEEE Trans Image Process 22(7):2786–2797. https://doi.org/10.1109/TIP.2013.2258353
25. Dimitropoulos K, Barmpoutis P, Grammalidis N (2015) Spatio-Temporal flame modeling and dynamic texture analysis for automatic video-based fire detection. IEEE Trans Circuits Syst Video Technol 25(2):339–351. https://doi.org/10.1109/TCSVT.2014.2339592
26. Borges PVK, Izquierdo E (2010) A probabilistic approach for vision based fire detection in videos. IEEE Trans Circuits Syst Video Technol 20(5):721–731

27. Rafiee A, Tavakoli R, Dianat R, Abbaspour S, Jamshidi M (2011) Fire and smoke detection using wavelet analysis and disorder characteristics. In: Proc. 3rd int. conf. comput. res. develop. (ICCRD), vol 3, pp 262–265, Mar 2011
28. Habiboglu YH, Günay O, Çetin AE (2012) Covariance matrix-based fire and flame detection method in video. Mach Vis Appl 23(6):1103– 1113. https://doi.org/10.1007/s00138-011-0369-1
29. Kong SG, Jin D, Li S, Kim H (2016) Fast fire flame detection in surveillance video using logistic regression and temporal smoothing. Fire Saf J 79:37–43
30. Han X-F, Jin JS, Wang M-J, Jiang W, Gao L, Xiao L-P (2017) Video fire detection based on Gaussian Mixture Model and multi-color features. SIViP 11(8):1419–1425
31. Habiboğlu YH, Günay O, Çetin AE (2012) Covariance matrix-based fire and flame detection method in video. Mach Vis Appl 23(6):1103–1113
32. Di Lascio R, Greco A, Saggese A, Vento M (2014) Improving fire detection reliability by a combination of video analytics. Proc Int Conf Image Anal Recognit 477–484
33. Poobalan K, Liew S (2015) Fire detection algorithm using image processing techniques. In: Proceedings of the 3rd international conference on artificial intelligence and computer science (AICS2015), pp 160–168, Oct 2015
34. Gomes P, Santana P, Barata J (2014) A vision-based approach to fire detection. Int J Adv Robotic Syst 09
35. Frizzi S, Kaabi R, Bouchouicha M, Ginoux JM, Moreau E, Fnaiech F (2016) Convolutional neural network for video fire and smoke detection. In: Proceedings of the IECON 2016—42nd annual conference of the IEEE industrial electronics society, Florence, Italy, 23–26 October 2016; IEEE: Piscataway, NJ, USA, pp 877–882
36. Zhang Q, Xu J, Guo H (2016) Deep convolutional neural networks for forest fire detection. In: Proceedings of the 2016 international forum on management, education and information technology application, Guangzhou, China, 30–31 January 2016; Atlantis Press, Paris, France
37. Muhammad K, Ahmad J, Lv Z, Bellavista P, Yang P, Baik SW (2018) Efficient deep CNN-based fire detection and localization in video surveillance applications. IEEE Trans Syst Man Cybern Syst 99:1–16
38. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K (2016) SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 model size. Comput Vis Pattern Recognit (cs.CV); Artif Intell (cs.AI)
39. Filonenko A, Kurnianggoro L, Jo K (2017) Comparative study of modern convolutional neural networks for smoke detection on image data. In: 10th International conference on human system interactions (HSI), pp 64–68
40. Yin Z, Wan B, Yuan F et al (2017) A deep normalization and convolutional neural network for image smoke detection. IEEE Access 18429–18438
41. Dunnings AJ, Breckon TP (2018) Experimentally defined convolutional neural network architecture variants for non-temporal real-time fire detection. In: 25th IEEE international conference on image processing (ICIP), pp 1558–1562
42. Mao W, Wang W, Dou Z, Li Y (2018)Fire recognition based on multi-channel convolutional neural network. Fire Technol 531–554
43. Muhammad K, Ahmad J, Baik SW (2018) Early fire detection using convolutional neural networks during surveillance for effective disaster management. Neurocomputing 30–42
44. Hu Y, Lu X (2018) Real-time video fire smoke detection by utilizing spatial-temporal ConvNet features. Multimed Tool Appl 77:29283–29301
45. Namozov A, ChoAn Y (2018) Efficient deep learning algorithm for fire and smoke detection with limited data. Adv Electr Comput Eng 121–128
46. Wonjae L, Seonghyun K, Yong-Tae L, Hyun-Woo L, MinDeep C (2017) Neural networks for wildfire detection with unmanned aerial vehicle. In: 2017 IEEE international conference on consumer electronics (ICCE), pp 252–253
47. Sathesh A (2020) Computer vision on IOT based patient preference management system. J Trends Comput Sci Smart Technol 2(2):68–77

48. Patil PJ, Zalke RV, Tumasare KR, Shiwankar BA, Singh SR, Sakhare S (2021) IoT protocol for accident spotting with medical facility. J Artif Intell 3(02):140–150
49. Sungheetha A, Sharma R (2020) Real time monitoring and fire detection using Internet of Things and cloud based drones. J Soft Comput Paradigm (JSCP) 2(03):168–174
50. Ponmalar DS (2020) Tungsten DiSulphide FBG sensor for temperature monitoring in float glass manufacturing. J Inf Technol Digit World 2(4):191–200
51. Srinivas K, Dua M (2019) Early discovery of disaster events from sensor data using fog computing. In: International conference on intelligent computing, information and control systems, pp 114–121. Springer, Cham
52. https://neurohive.io/en/popular-networks/vgg16/
53. https://towardsdatascience.com/cnn-architectures-a-deep-dive-a99441d18049
54. https://towardsdatascience.com/faster-rcnn-object-detection-f865e5ed7fc4
55. https://medium.com/@kangeugine/long-short-term-memory-lstm-concept-cb3283934359#:~:text=LSTM%20is%20a%20recurrent%20neural,time%20lags%20of%20unknown%20duration
56. https://cloud.google.com/tpu/docs/inception-v3-advanced
57. Redmon J, Divvala S, Girshick R, Farhadi A (2015) You only look once: unified, real-time object detection
58. Habiboğlu YH, Günay O, Çetin AE (2012) Covariance matrix-based fire and flame detection method in video. Mach Vis Appl 23:1103–1113
59. krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep neural networks. In: Advances in neural information processing systems, pp 1097–1105
60. Le Cun Y et al (1990) Handwritten digit recognition with a back-propagation network. Proc Adv Neural Inf Process Syst 396–404
61. Di Lascio R, Greco A, Saggese A, Vento M (2014) Improving fire detection reliability by a combination of video analytics. In: Proceedings of the international conference image analysis and recognition, Vilamoura, Portugal, 22–24 October 2014. Springer: Cham, Switzerland
62. Jadon A, Omama M, Varshney A, Ansari MS, Sharma R (2019) FireNet: a specialized lightweight fire & smoke detection model for real-time IoT applications. arXiv 2019. arXiv:1905.11922
63. Muhammad K, Khan S, Elhoseny M, Ahmed SH, Baik SW (2019) Efficient fire detection for uncertain surveillance environment. IEEE Trans Ind Inform 15:3113–3122

Various Diseases' Prediction Based on Symptom by Using Machine Learning

M. Murugesan, R. Gowtham, R. Logesh, S. Selvaganapathy, and R. Yogesh Muthumalai

Abstract An overwhelming number of disease surveys capture patient records of the severity of their illness indications that enable differentiation of patient ailment from typical indication. This paper aims to forecast illnesses in users based on their symptoms. To achieve our aim, we use the XGBoost Classifier, which aids in regulating the victim's health condition then acquiring the symptoms to predict the condition. The data carries 250 variables and 140 corporal evaluation instances (diseases). In addition, we provide medicine for illness and gather the corresponding victim EH to summarize the diagnostic described by the NLTK.

Keywords Extreme gradient boosting · Natural language tool kit · Exploratory data analysis

1 Introduction

Today, many people are affected by various diseases and different kinds of infection by the atmosphere. The disease may cause harmful air in the surrounding and uncontrolled pollution [1]. This kind of activity will be made all over the world. WHO association will take many measures for overcoming this problem. They need software support to reduce the speed of reproducing the disease [2–5]. The medical people collect past victims' medical data and doctor prescription data [6]. By using this, the medical people transfer the data to the software development persons [7]. According to the dataset, the code will add to a particular project. Handling dataset is a little bit tough to develop the software. So, developers design predictive design concepts in machine learning. This kind of model will easily guess the correct disease [8, 9]. The Dataset contains various user diseases indications [10]. So, the software can predict the exact disease. In previous projects, they were added only a few symptoms and they can predict only one disease with the help of the dataset [11].

M. Murugesan (✉) · R. Gowtham · R. Logesh · S. Selvaganapathy · R. Y. Muthumalai
Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Tamil Nadu, Karur 639113, India
e-mail: murugankathir@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_31

Lesser projects are predicting the small kind of disease. But no one can predict a greater number of diseases in the software [12]. To overcome this unpredictable disease, we can add some more disease symptoms to the data list [13]. In machine learning, various algorithms are available by that we can predict the exact disease name, other than predicting medicine to the certain victim's disease [14]. Both ML and DL have lots of algorithms [15]. In ML, we have a popular latest algorithm so we have to get the judgment above 96% and then apply Extreme Gradient Boosting concepts for disease prediction. To assist the patient in comprehending the health report, the text file is summarized using NLTK [16–18]. However, in addition to the two core modules, different modules such as arranging doctor appointments, storing and retrieving medical information, and identifying physicians are incorporated and then made the system a full health record monitoring system [19, 20].

2 Related Work

In our society, every day new diseases are found. Some people do not have an idea about disease and hospital [21]. Most diseases give small symptoms to the body like pain, swelling, insomnia, and fatigue, but people do not notice them [22]. They think it is a small problem, and they do not meet the proper doctor. They take drugs from a random medical shop for that sick [23]. Most of the time they do not meet the appropriate doctor for that problem. But the fact is that a small problem leads to a big disease [24]. In the medical field, finding diseases in starting stage is very important [25]. This project helps to predict the disease based on the symptoms. By this prediction, the user can consult the appropriate doctor and find the disease at an earlier stage. So, they avoid high-level treatment, sickness, high expenses, and dead [26–29].

2.1 Problem Statement

In existing Web sites, only one disease will be predicted by the model. But our project will predict various diseases. Algorithms like SVM and decision tree take more execution time to predict the disease, so that we choose the XGBoost algorithm to increase the accuracy level and it will work fast compared to other algorithms.

2.2 Extreme Gradient Boosting

The XGBoost approach is a standard data mining method for creating prediction models for a target instance based on many explanatory features [30]. This process cuts a population into the tree-like structure branch segments that form the inverted

tree with the root nodes, inner nodes, and leaves nodes [31]. An XGBoost the non-parametrical proceed for dealing with large, complex datasets without the need of several parametric structures [32]. Studying the information can be separated into validation and training if the sample size is high enough. Use the labeled training data to generate an XGBoost algorithm and the testing dataset to determine optimal tree size to create the best final model [33, 34].

2.3 Data Collection

Word finalization and Extreme Gradient Boost methods can be used in this article. For the purpose of work evaluation, list of data is taken to an electronically health information database. The textual dismiss records from the New York-Presbyterian Hospitals were obtained by Columbia University to investigate the document's trends. Based on the study, classify the dataset to the examiner and priming data [35]. We collected some disease datasets. We collect the 4920 records, and we segregate into three different datasets for pre-processing model. Each dataset has different symptoms, and based on the symptoms, the model will be displayed the possible disease names. For the symptom prediction, we collected the datasets like description of symptoms which have the detailed description about the each and every disease. In precaution dataset, it contains precaution instructions to reduce the disease. Severity dataset is a dataset which has assigning the weight for each disease to pre-processing value in binary format.

2.4 Execution

The Extreme Gradient Boosting predicts a disease depending upon the underlying symptoms. To begin, we collect the information of the top five victims' indications and store them in an array with the number 1 allocated to each item. This is fed into the illness prediction as an input in the model. This array corresponds to the illness data collection and finishes with the greatest level of trust at a common leaf node [36].

2.5 Recursive Part

To create the model, we recurrence the mentioned procedure to grow level at the recursive section. It has nothing about whether prediction of publishing output supplied. We also employ electronic health records to add additional disease–symptom pairings to the dataset for better disease detection based on symptoms [37].

Figure 1 shows the structure of XGBoost algorithm. This picture shows that the

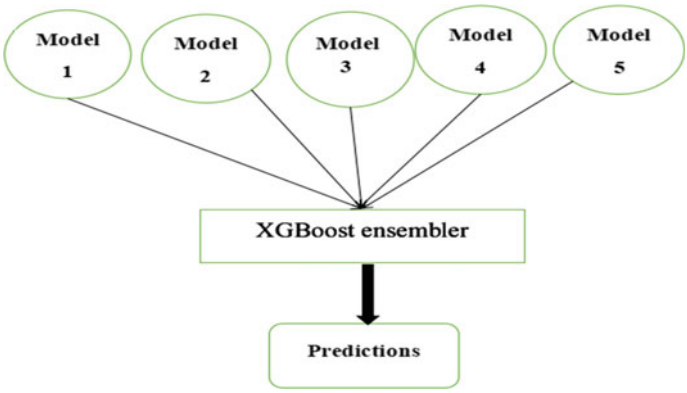


Fig. 1 The structure of XGBoost algorithm

structure will collect the individual datasets from the models and using the XGBoost algorithm it will predict the disease based on the symptoms.

Table1 represents the merits and demerits of the existing algorithms for disease predictions, and it shows the level of the accuracy of existing algorithms.

2.6 Extracting Data

Another big deme of the heart illness discovery of the system is the extraction of the useful characteristics for the bio textual data in the joining the sensor data. Factor extraction techniques are used to extricate handy data to large amounts of healthcare data. It is a method for combining data from several sources to provide useful and relevant data. Many approaches for developing features to healthcare word data. To predict cardiac disease, a current–time system with machine learning classifications was explained [38, 39]. The models extract the essential characteristics from the healthcare dataset using separate feature extraction algorithms. To mine EMRs for the heart disease medical diagnosis, a unique and robust framework was provided. For feature reputation and coronary heart failure prediction, this device employees a word embedding framework and prolonged short-time period memory. Another technique [40] utilized echocardiography searching to forecast the death rate of cardiac users in hospitals. This method extracts data from the word using text mining concepts and then uses a deep learning model to forecast death rates. To extract valuable data from EMRs, individual text mining tricks are discovered. This system's show used ideal-based tools to automatically valuable data extractions and idea-making. Which helps, various medical records include the unstructured data, which makes using a rules-based engine problematic. Furthermore, utilizing unstructured EMRs of patients, a heart disease risk score is established. The inventors of this system employ mining of word algorithms to identify the heart-critical parameters for the unstructured of

Table 1 Related works

Title of the paper	Description	Algorithm	merits	Demerits	Accuracy (%)
Data storage is proceeded prediction diabetes and cardiovascular disease with the ML concepts	This model predicts a data ideal based on the discussion of the datasets	Logistic Regression	It is easy to understand and train	It met only one be used to predicts discrete concepts	72
Identify the corona based on symptoms through the help of ML	This algorithm will correct at the random to develop either forest or gather the decision tree	Random forest	It acts in two types, continuous and categorically values of the dataset	It will suffer interoperability and fails to determine the significance of each variable	84
Forecast the Stages and Parkinson's Disease	A diagrammatical representation of the positive answer the decides based on the conditions	Decision Tree	It does not require normalization of data	It will involve Peaker period to training the tools	90
Predicts the agriculture crops using the KNN algorithm	This concept is used to work out problem condition and classification	KNN algorithm	It stores the training dataset, and it learns only at the period of making predictions	KNN is the high delicate to the audible to the dataset	93

EMRs and create the risk scores for heart disease in diabetic's users. For music suggestions, wearable sensors are used to determine users' moods. The authors used feature-level fusion in their system, which throws data from each server divided and they merge them to emotion recognition. The scientists then employed both ambient and peripheral body sensors to identify mood in a different system. They employed all three levels of fusion to investigate the influences on human health (data, feature, and decision). The (RCNN) stands for the recurrent convolutional neural network-based method described to find out sickness symptoms.

2.7 Characteristics Identify Model

The most related features are picked used the correlation proceed, and they selected characteristics to subsequently merged using both the data-level and feature-value fusion. Depending upon eight chosen characteristics from activity sensors, our system attained a 98% accuracy rate. These qualities, however, do not address all of the risk elements that heart patients face. To deal with the unpredictability of healthcare statistics, the system used to merge the characteristic concepts reach the accuracy of 88.6%, which may be on any format. We will talk about finalizing victims' input in the portion of the study. We do this by retrieving the user's medical record, which may be in another format. The record is shared to a text file, which is passed to the outline module as input. It produces the outline of the health data, and we utilize NLTK—the Natural Language Toolkit.

2.8 NLTK Libraries

Natural Language Toolkit for understanding natural language uses the library tools, and it is a group of a library package. These are the fastest NLP libraries, with packages for learning machines to understand and reply to human language. To arrange numerous words and their frequency factors, we use two libraries which are Corpus and Tokenizer. It is also utilized to keep the stop words in place. By using TF-IDF calculation, we will assign weight for the medical word in the existing symptoms dataset. Corpus: The word corpus mentions the group of text facts. It might be a poet's series of poetry or a blog about assure theme. It functions as a collection of predetermined stop words in this case. Tokenizer sentence tokenism is utilized. It is used to construct succession to the tokens as well as to make regex.

The process of NLP shown in Fig. 2 shows that the words need to be encoded as integers or floating-point values for use as input to a machine learning algorithm. First, we want to load the text. Next, we split the text into tokens and then convert all words into lower case because the computer reads only machine level language. This is the process of tokenizers.

2.9 Procedures for Build Summarization

- You should delete stop text.
- A word frequency table is created, revealing how many instances each word occurs in the dataset, as well as a medical definition weight analysis, which priorities medical terms.
- Using these wordmarks on the temporary table, design the score for each phrase.

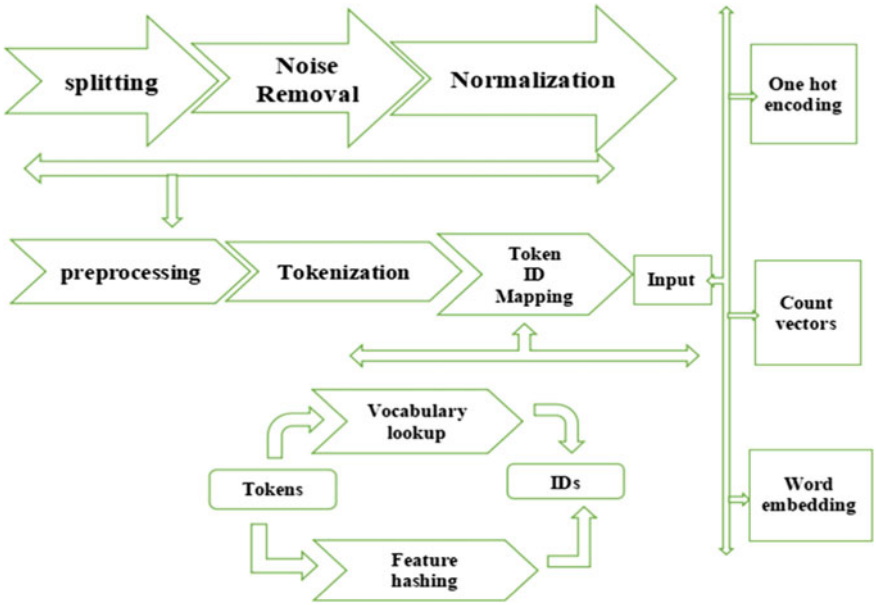


Fig. 2 Flow diagram of NLP process

- The sentence score that is over a particular level is used to construct an analyzed summary.

2.9.1 Need for Summarization

Values which the time, writing style, and grammar are among the technologies that enable a clear description. The primary premise of finalizing is to look at a subset of data that incorporates all of the data an in array. Data overview is often used to manage email message summaries and take action condense material by condensing words used to organize data and promote Web search engines to study the condensed medical data and operation reports. Normally, all kinds of finalization were reserved for mother tongue. But in this paper, we fetched it and it is different from the new environment in the medical sectors, where it aids in extracting a summary and assists victims in understanding what their treatment says unsorted parley qualified doctor. It also works with the app's other capabilities, such as locating the nearest health facility and allowing patients to make medical appointments. So user do not worry about the medical resources. User gives the symptoms, based on that our model will predict the diseases, and it will recommend some Web site which is related to the disease.

2.9.2 Calculation

Result of sentence equal to medical result and frequency result

$$\text{Last score} = \text{Sentence result} + \text{No of characters sequence}$$

This lastly gives the top score to each sequence and builds a review.

$$F1\,\text{score} = 2/(1/\text{precision} + 1/\text{recall})$$

The recall is the number of correct positive results divided by number of all relevant samples.

3 Existing Systems

Various technologies based on wearable sensors have been developed in recent years to enhance the act of cardiac disease predicts. Al-Makhdmeh and Tollba explain the system based on some medical tools which gather information about the cardiac patients both before and after the heart pain. They will transfer the accurate data in the medical scheme when they used characteristics extraction tricks and a dl model to extract valuable features and classify the data correctly. Still, the sensitive characteristic parentage and feature weight technique have limits. In addition, they hired 23 characteristics for education that inserts to the convolution and spatial dimensions inside the machine. Health Fog is a method of the use of Internet of Things (IoT) routers and deep mastering gear to autonomously cure cardiac patients. The principal intention of this machine is to efficiently manipulate facts from IoT gadgets that pertain to cardiac sufferers. A new structure for a choice aid device for illness identity in patients becomes also defined. This gadget merges clinical information acquired from various sources with facts from wearable medical devices. A hybrid advocate device primarily based on IOT gadgets changed into defined. According to the cardiac patient's state, the system offers a nutrition plan and physical activity. This recommended system, on the contrary, is built on basic rules which need to semantic information to extract the more sensitive information about the patient in sequence to give an accurate suggestion. A fuzzy neural classifier was used in an IoT-based illness prediction system. The area gives a breakthrough framework of a mobile healthcare utility for important contamination prognosis. In addition, a three-tier system with an ML method and the third tier use logistic regression to forecast cardiac disease. This method, on the contrary, runs tests with only seven characteristics, which is insufficient to appropriately determine the health status of cardiac patients.

4 Proposed System

The Kaggle repository was used to collect the Disease/Symptoms database. It has numerous qualities (symptoms) and classifications, as previously stated (diseases). Extreme Gradient Boost algorithm model is used to construct the training and testing sets. We collected some disease datasets. We collect the 4920 records. And we segregate into three different datasets for pre-processing model. Each dataset has different symptoms, and based on the symptoms, the model will be displayed the possible disease names. For the symptoms prediction, we collected the datasets like description of symptoms which have the detailed description about the each and every disease. In Precaution dataset, it contains precautions instructions to reduce the disease. Severity dataset is a dataset which has assigning the weight for each disease. In our dataset, we have some meaningless data like null values. The model's overall performance might be harmed by the text's noise and needless elements. To reduce the issues, we used the NLTK concepts to recover the model. Therefore, Weblinks, emojis, extra punctuation, and unwanted, repeated characters were removed from each sentence. It helps to get rid of unhelpful parts of the data, by converting all characters into lowercase, eliminating the punctuation marks, and removing stop words and types.

Following that, translate contractions ("I'm", "there're") and abbreviations to formal language. To extract individual words in a sentence, the author adds the most prevalent abbreviation to a list and tokenizes each sentence. In the list, look for abbreviated terms and replace them with the formal language equivalent. Some pronouns, such as "you", "she", "them", and others, were removed from the English stop words list. To save space in the document, I converted all of the text to lowercase.

4.1 Feature Extraction

Because machine learning algorithms cannot handle all of the features, the author chose some of the best content-based features and sentiment-based features. The contents of the comments are used to create content-based features. Sentiment-based features are those that are based on the text's emotional content. Bag of Words: The bag-of-words model is a representation for natural language processing and information retrieval that simplifies things. A text is represented as a bag of its words in this approach, which ignores syntax and even word order while maintaining multiplicity. Computer vision has also used the bag-of-words model.

TF-IDF: The term frequency-inverse document frequency (tf–IDF) is a numerical statistic used in information retrieval to represent how essential a word is to a document in a collection or corpus. It is frequently used in information retrieval, text mining, and user modeling searches as a weighting factor.

WORD2VEC: Word2vec is a natural language processing approach that was first published in 2013. The word2vec program learns word connections from a huge

corpus of text using a neural network model. Once trained, a model like this can recognize synonyms and recommend new words for a sentence. We obtain the user's symptoms and apply the trained model to forecast the ailment. The patient record, on the contrary, is obtained, resulting in an overview of the medical report depending on the most important symptoms linked with a particular ailment. It is used to expand the dataset's number of illness symptom pairings. Any predictive engine will typically simply look at a predetermined platform of symptom and sickness combinations, but this will only locate the result partially and with a low level of confidence.

To achieve a better degree of confidence, we leverage health data analysis that provides personalized input as well as user engagement with the system, on this project. The entire system contains two modules: one for illness prediction and the other for keeping track of one's health. The next module utilized training to deliver the best outcome to reach a higher confidence level. This procedure is linked to the user interface to collect user input and inform them of their current state.

4.2 Classification Model

We use multiple class divisions in this forecast. Multi-stage is a dividing function that has more than two phases. Each sample can only label as one class. We predicted 40 different types of diseases. In this classification, we use the one-vs-rest strategy, also known as the one-vs-all, used in the one-vs-rest classifier. The strategy consists of inserting one category into each category. In each category, the class is included in all other categories. In addition to its computer efficiency, another advantage of this method is its interpretation. Since each class is represented by only one category and one, it is possible to gain knowledge about the class by examining its corresponding arrangement. This is a widely used strategy and is a viable default option.

Figure 3 represents the XGBoost probability of predicting the correct possibility of instance. It shows the negative average of correct prediction. We can measure the performance of classification model. The blue line shows the train set of log loss, and orange shows the test set of log loss.

4.3 Prediction

The objective is expected based on the symptoms provided. The NLTK is used in conjunction with medical word weight analysis to summarize the uploaded file, resulting in an optimum summary for the patient or user. It also aids in the upkeep of the disease–symptom data collection on the back end. For each of the anticipated diseases, the XGBoost first predicts the disease with the greatest confidence level.

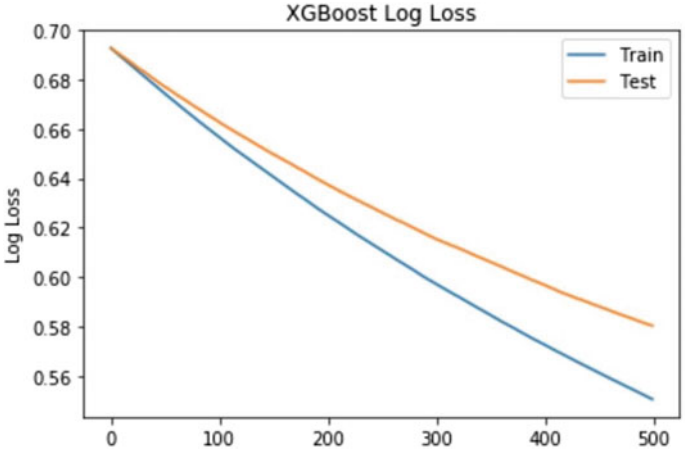


Fig. 3 Shows the XGBoost log loss

5 Conclusion

We need to add the multilingual finalizing and multi-document finalization as future development. Because the files we provide as input may include local languages, health records may be gathered from all over the world and readily summarized utilizing multilingual finalization. Currently, the document uses a global language (English). This study discusses illness prediction utilizing highly personalized training datasets, as well as other associated activities such as scheduling appointments and locating the nearest health center.

6 Future Scope

Predictability of disease is important in clinical decision-making. Machine learning methods offer an exciting opportunity to gain advanced risk assessments and more for each individual. This can aid in a personal drug campaign, by better integrating risk management into each patient. The improvement in the accuracy of the predictions found in the current study should be further evaluated using the study machine and other large clinical datasets, in some communities, and in predicting further disease outcomes. Future research into the feasibility and validity of applications for machine learning in clinical practice will be required. As the computational capacity of healthcare systems increases, the potential for exploitation of machine learning to improve disease risk prediction in clinical practice will be a viable option. This may further include predicting protein formation and genetic sequence from patients' clinical profiles. This will require testing in future studies on the use and clinical use of other computer-assisted machine learning algorithms such as support vectors and

in-depth learning equipment to integrate electronic healthcare records. In the future, we will add features like suggest clinical tests and based on the test result, suggest the hospitals for the treatment, suggest doctors, and show the approximate treatment cost. In several countries, electronic health records for all health care organizations are hosted on central servers. This may allow the development of a new algorithm to be performed off-site using cloud computing software and restored to clinical settings such as program applications (APIs) for PCs, mobile devices, and table.

References

1. Sungheetha A, Sharma R (2021) Design an early detection and classification for diabetic retinopathy by deep feature extraction-based convolution neural network. J Trends Comput Sci Smart Technol (TCSST) 3(02):81–94
2. Ahmed H, Younis EMG, Hendawi A, Ali AA (2019) Heart disease identification from patients' social posts, machine learning solution on Spark. Future Gener Comput Syst 111:714–724
3. Balasubramaniam V (2020) IoT based biotelemetry for smart health care monitoring system. J Inf Technol Digital World 2(3):183–190
4. Journal articles—Saiyed Saziyabegam, Priti S (2016) Literature review on extractive text summarization approaches. Int J Comput Appl 156(12):28–36
5. Murugesan M, Thilagamani S (2020) Efficient anomaly detection in surveillance videos based on multi-layer perception recurrent neural network. J Microprocess Microsyst 79
6. Al-Makhadmeh Z, Tolba A (2019) Utilizing IoT wearable medical device for heart disease prediction using higher-order Boltzmann model: a classification approach. Measurement 147
7. Thilagamani S, Nandhakumar C (2020) Implementing green revolution for organic plant forming using KNN-classification technique. Int J Adv Sci Technol 29(7S):1707–1712
8. Grampurohit S, Sagarnal C (2020) Disease prediction using machine learning algorithms. In: 2020 International conference for emerging technology, India, pp 1–7
9. Thilagamani S, Shanti N (2014) Gaussian and gabor filter approach for object segmentation. J Comput Inf Sci Eng 14(2):7
10. Brandow R, Mitze K, Rau LF (2019) Automatic condensation of electronic publications by sentence selection. Inf Process Manage 31(5):675–685
11. Ferjani MF (2020) Disease prediction using machine learning 10:13140
12. Pandiaraja P, Suba S (2022) An analysis of a secure communication for healthcare system using wearable devices based on elliptic curve cryptography. J World Rev Sci Technol Sustain Dev 18(1):51–58
13. Hamsagayathri P, Vigneshwaran S (2021) Symptoms based disease prediction using machine learning techniques. In: 2021 Third international conference on intelligent communication technologies and virtual mobile networks (ICICV), vol 7, pp 747–752
14. Pandiaraja P, Sharmila S (2020) Optimal routing path for heterogenous vehicular ad-hoc network. Int J Adv Sci Technol 29(7S):1762–1771
15. Fattah MA, Ren F (2018) Automatic text summarization. World Acad Sci Eng Technol 37
16. Pandiaraja P, Aravinthan K, Lakshmi NR, Kaaviya KS, Madumithra K (2020) Efficient cloud storage using data partition and time-based access control with secure as encryption technique. Int J Adv Sci Technol 29(7S):698–706
17. Luhn HP (2015) The automatic creation of literature abstracts. IBM J Res Dev 2(2):159–165
18. Rajesh Kanna P, Santhi P (2021) Unified deep learning approach for efficient intrusion detection system using integrated spatial–temporal features. Knowl-Based Syst 226
19. Edmundson HP (2016) New methods in automatic extracting. J ACM (JACM) 16(2):264–285
20. Santhi P, Mahalakshmi G (2019) Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction. Int J Eng Adv Technol 8(4):839–846

21. Bernal EA, Yang X, Li Q, Kumar J, Madhvanath S, Ramesh P, Bala R (2018) Deep temporal multimodal fusion for medical procedure monitoring using wearable sensors. IEEE Trans Multimer (20):107–118
22. Miller GA (2015) WordNet. a lexical database for English. Commun ACM 38(11):39–41
23. Deepa K, Thilagamani S (2019) Segmentation techniques for overlapped latent fingerprint matching. Int J Innov Technol Exploring Eng 8(12):1849–1852
24. Lin CY, Hovy E (2019) Identifying topics by position. In: Proceedings of the fifth conference on Applied natural language processing, pp 283–290
25. Pradeep D, Sundar C, QAOC (2020) Novel query analysis and ontology-based clustering for data management in Hadoop (108):849–860
26. Davoodi R, Moradi MH (2018) Mortality prediction in intensive care units (ICUs) using a deep rule-based fuzzy classifier. J Biomed Inform 7948–7959
27. Logeswaran R, Aarthi P, Dineshkumar M, Lakshitha G, Vikram R (2019) Portable charger for handheld devices using radio frequency. Int J Innov Technol Exploring Eng (IJITEE) 8:837–839
28. Zhang Y, Zincir-Heywood N, Milios E (2015) Narrative text classification for automatic key phrase extraction in web document corpora. In: Proceedings of the 7th annual ACM international workshop on Web information and data management, pp 51–58
29. Fazlic LB, Hallawa A, Schmeink A, Peine A, Martin L, Dartmann G (2019) A novel NLP-fuzzy system prototype for information extraction from medical guidelines. Conv Inf Commun Technol Electron Microelectron 1025–1030
30. Gunasekar M, Thilagamani S (2020) Performance analysis of ensemble feature selection method under SVM and BMNB classifiers for sentiment analysis. Int J Sci Technol Res 9(2):1536–1540
31. Hao Y, Usama M, Yang J, Hossain MS, Ghoneim A (2019) Recurrent convolutional neural network based multimodal disease risk prediction. Future Gener Comput Syst 92:76–83
32. Deepika S, Pandiaraja P (2013) Ensuring CIA triad for user data using collaborative filtering mechanism. In: International conference on information communication and embedded systems (ICICES), pp 925–928
33. Zamanifar A, Minaei-Bidgoli B, Sharifi M (2019) A new hybrid Farsi text summarization technique based on term co-occurrence and conceptual property of the text. In: Software engineering artificial intelligence, pp 23–33
34. Networking, and Parallel/Distributed computing, SNPD'08. In: Ninth ACIS international conference on IEEE, pp 635–639
35. Jabeen F, Maqsood M, Ghazanfar MA, Aadil F, Khan S, Khan MF, Mehmood I (2019) An IoT based efficient hybrid recommender system for cardiovascular disease. Peer-to-Peer Net Appl 12(1):1263–1276
36. Rajesh Kanna P, Santhi P (2022) Hybrid intrusion detection using map reduce based black widow optimized convolutional long short-term memory neural networks. Expert Syst Appl 226:1
37. Suanmali L, Binwahlan MS, Salim N (2019) Sentence features fusion for text summarization using fuzzy logic. In: Hybrid intelligent systems, 2019. HIS'09. Ninth international conference on IEEE 1, pp 142–146
38. Jonnalagadda SR, Adupa AK, Garg RP, Corona-Cox J, Shah SJ (2017) Text mining of the electronic health record: an information extraction approach for automated identification and sub phenotyping of HFpEF patients for clinical trials. J Cardiovasc Transl 10:313–321
39. Deepa K, Kokila M, Nandhini A, Pavethra A, Umadevi M (2020) Rainfall prediction using CNN. Int J Adv Sci Technol 29(7S):1623–1627
40. Ishtake S, Sanap S (2013) Intelligent heart disease prediction system using data mining techniques. Int J Healthcare Biomed Res 1(3):94–101

Landmark Recognition and Retrieval Using ResNet50 and DELF

P. Nikhil Chandra, M. Kalyan, B. Rishi Ram Naik, K. L. Sailaja, and P. Ramesh Kumar

Abstract Landmark recognition is a method that predicts landmark labels directly from image pixels. This model is designed to detect the proper landmark in a dataset of complex and challenging test images, allowing users to better understand and organize their photograph collections. A dataset containing historic monuments of India of around 100 classes is created. The dataset is divided into training, testing, and validation data. Data augmentation is performed on training and validation data in the dataset. ResNet50 algorithm is used here to train the dataset. A sample image is given to the ResNet50 model for prediction, and the model gives the top ten labels of the sample image. Consider two images for each label from the training dataset, and apply DELF on these images and sample image. The inliers are calculated for these images using the Ransac algorithm. Image label which has a good number of inliers with sample image is the landmark of that sample image.

Keywords Data augmentation · Deep local features (DELF) · Landmark recognition · Ransac algorithm · ResNet50

1 Introduction

The challenge of finding popular man-made sculptures, structures, and monuments inside a picture is known as landmark recognition [1]. One of the most commonly

P. Nikhil Chandra · M. Kalyan (✉) · B. Rishi Ram Naik · K. L. Sailaja · P. Ramesh Kumar
VR Siddhartha Engineering College, Vijayawada, India
e-mail: kalyanmedepalli07@gmail.com

B. Rishi Ram Naik
e-mail: rishiram022000@gmail.com

K. L. Sailaja
e-mail: sailajak@vrsiddhartha.ac.in

P. Ramesh Kumar
e-mail: rameshkumar@vrsiddhartha.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_32

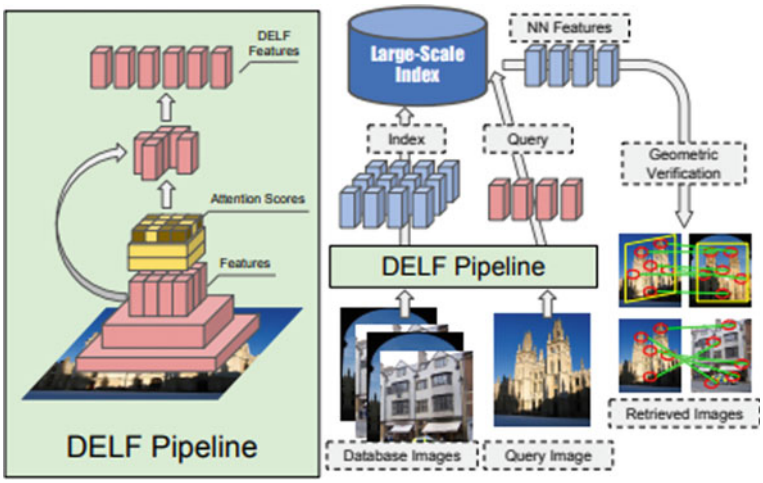


Fig. 1 Deep local features [4]

utilized deep neural networks is the convolutional neural network (CNN). Convolution is the name given to a mathematical linear function between matrices. CNN layers include all convolutional layers, the nonlinearity layer, the pooling layer, and the fully connected layer. Residual neural networks [2] do this by utilizing skip connections or shortcuts to jump over some layers. The principle on which ResNet's work is to build a deeper network compared to other plain networks. Data augmentation [3] is a technique for artificially increasing the size of a training set by producing modified copies of the dataset's pictures. When training a machine learning model, it functions as a regularizer and helps to reduce overfitting. Deep local features can be utilized to replace other keypoint detectors and descriptors in image retrieval, allowing for more accurate feature matching and geometric verification (Fig. 1).

The dataset is created by collecting 1200 images of different landmarks across all states of India from the Google Images.

2 Related Work

In [5], using a Google landmark database of recognized landmarks, the author hopes to construct and enhance a deep learning system that can detect landmarks from photographs. Landmark recognition is powered by instance-level recognition. In this study, the transfer learning approach is used to train three separate VGG16, Inceptionv3, and ResNet50 models, as well as a pure convolution neural network (CNN) model from scratch. Table 1 shows the different accuracies for these algorithms. In this, ResNet50+ DELF gives high accuracy.

Table 1 Accuracy of different models

Model	Training accuracy (%)	Testing accuracy (%)
VGG16	99.8	89.1
Inceptionv3	27.3	18.6
ResNet50	100	90.4
Pure CNN	97.55	89.67
ResNet50+ DELF	100	95.67

The author proposed a model [6] for predicting landmarks with CNN, which was constructed with a VGG16 neural network and ImageNet transfer learning. The evaluation is carried out using the global average precision (GAP) measure, which takes into account the confidence score for each anticipated landmark label. Deep local features use CNN to discover semantically equivalent local features by training using instance-level annotations.

The author [7] proposed a model that automatically discriminates landmark categories with subtle visual differences. For image cropping, they use image cropping and data augmentation techniques. Modeling will extract different numbers of landmarks and images from our dataset to build the landmark recognition models using ResNet and VGG.

Anirudha Sundaresan et al. [8] proposed a deep residual networks-based transfer learning technique. Due to its powerful computational capabilities, the performance of picture categorization, object identification, and face recognition has been improved. ResNet enables us to train incredibly deep networks while still achieving excellent results.

In [4], a unique local feature descriptor is created for large-scale image retrieval applications. DELF is learnt with very little supervision and only image-level labels, and it is combined with our novel semantic feature selection approach.

Monulens is a real-time mobile-based landmark identification tool proposed by Chatterjee et al. [9]. Monulens does all of its processing on the mobile device. Present quantifiable statistics on the time takes to execute a single query and the application's empirical correctness.

At the order level, the author proposed a technique [10] for identifying insect specimen photographs. The methodologies of digital picture progression, pattern recognition, and taxonomy theory were used to construct several related characteristics.

Ulrich Buddemeier et al. [11] created a system that analyzes two types of data: (1) a set of GPS-tagged images and (2) a corpus of travel guide articles. A photograph is a tuple for the initial source, with the unique photograph ID-associated GPS coordinates in terms of latitude and longitude.

Kumar et al. [12] proposed an article that checks image metadata and tracks the individual nation, city, route, and street based on GPS altitude, GPS latitude, GPS longitude, and GPS location.

Dr. Chen et al. [13] proposed an article that compares existing classification approaches to forecast CAD sooner for a more accurate value and also introduces recent adaptive image-based classification strategies.

Balasubramaniam [14] created a deterministic AI strategy based on clinician inputs and deep ensemble malignancy recognition. The histological diagnosis is used as a standard criterion for determining the specialist assessment, algorithmic specificity, sensitivity, and the area under the receiver operating characteristic curve (AUROC).

3 Methodology

In this model, there are three modules: data preprocessing, ResNet50 model training, and DELF (Fig. 2).

3.1 Data Preprocessing

The data is preprocessed by splitting the data into training, testing, and validation in the ratio of 0.6, 0.2, and 0.2, and data augmentation is performed on train and validation data with constraints such as shrinking, shearing, zooming, and horizontal flipping (Fig. 3).

3.2 ResNet50 Model Training

ResNet50 algorithm is used here to train the dataset. After that, a sample image is given to the ResNet50 model for prediction, and the model gives the array containing the probabilities of the label to the image. The ResNet50 model's skip connections

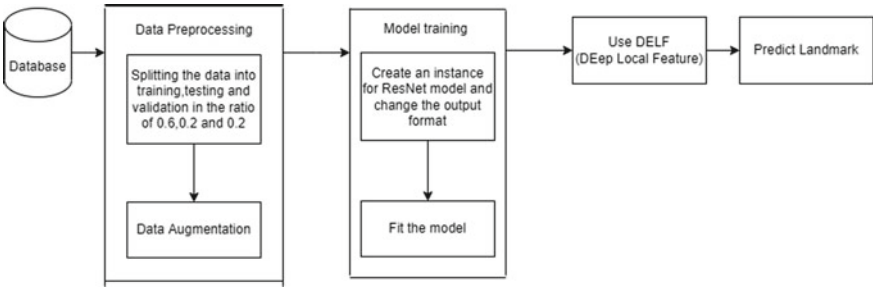


Fig. 2 Methodology

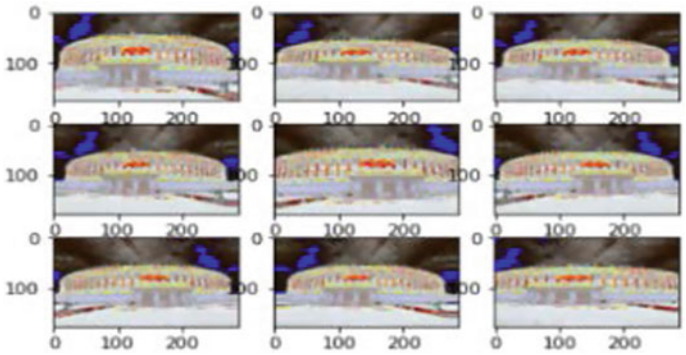


Fig. 3 Data augmentation

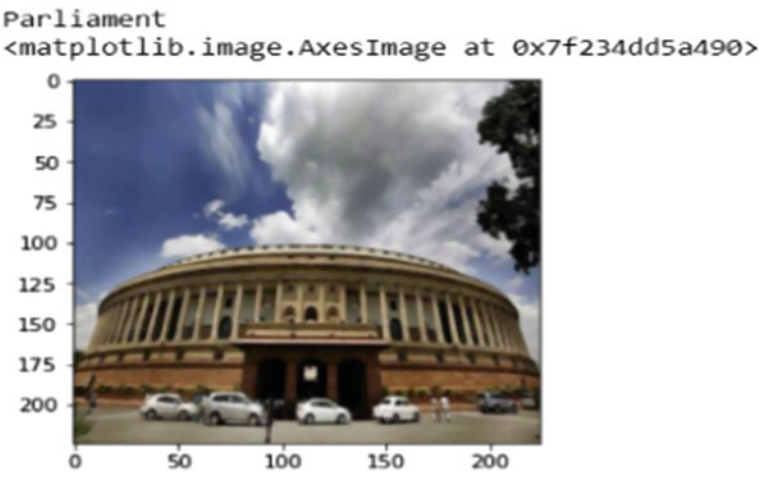


Fig. 4 Output of ResNet50

combine the outputs of prior layers with the outputs of stacked layers, allowing for considerably deeper network training than previously feasible (Fig. 4).

3.3 Deep Local Features

When the user gives input to the ResNet50 model, the model gives the array format output of probabilities. The top ten labels having high probability of that image are considered. For every label, two images are considered. Total 20 images are retrieved, in the images, and DELF is applied with input image and the retrieved image which will return the description of two images. By using the descriptors and locators,

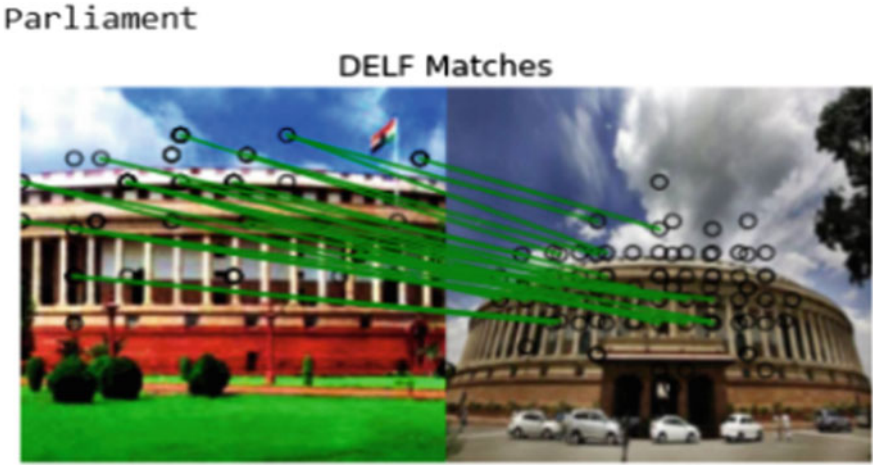


Fig. 5 Output of ResNet50+ DELF

Ransac algorithm is performed to return the inliers or matches of images. Image label which has a good number of inliers with sample image is the landmark of that sample image (Fig. 5).

4 Result and Analysis

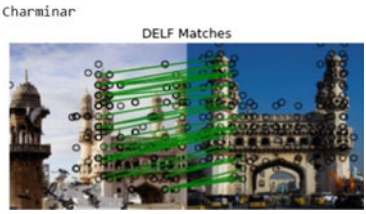
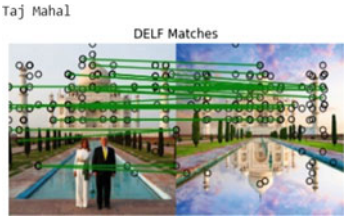
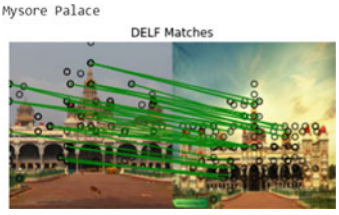
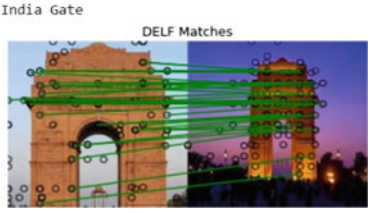
This Table 2 consists of two columns, one is input and the other output. Input: An image is taken from the user. This input image goes to ResNet50 and DELF model to get the output. Output: Label of that landmark and the image which has maximum matches with given image (Fig. 6; Table 3).

For the selection of labels, two constraints are considered. One is accuracy, and the other is time. This accuracy test is performed on a dataset containing 197 images. At first, five labels are considered. The accuracy is low, and time taken is also less for output generation. Up to ten labels, the accuracy and time taken are increasing slightly. Upon considering 15 labels, the rate of change of accuracy is less, and time taken is high. So, ten labels give better performance.

5 Conclusion and Future Work

In this paper, the model predicts the landmark of a given image using ResNet50 and DELF. The objective is to find an optimal method to find the landmark of an image by using the image pixels. This was made possible by creating a hybrid model with the combination of ResNet50 and DELF. The image was firstly given to the ResNet50

Table 2 Outputs of ResNet50+ DELF model

S. no.	Input	Output
1		
2		
3		
4		
5		

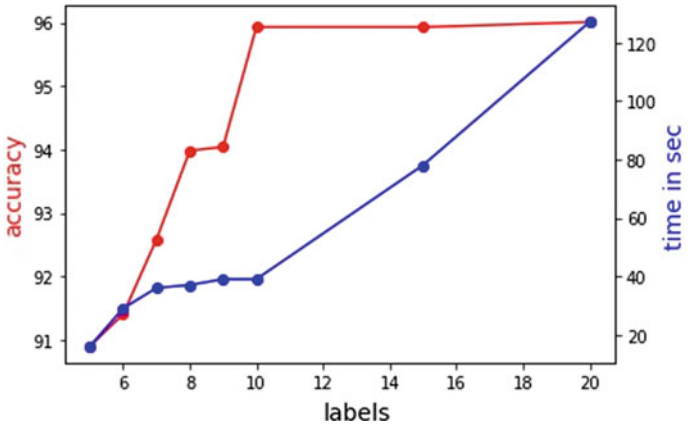


Fig. 6 Accuracy graph for the different no. of labels

Table 3 Accuracy and time taken table for the different no. of labels

S. no	No. of labels	Accuracy (%)	Time taken (in s)
1	5	90.89	16
2	6	91.40	29
3	7	92.58	36
4	8	93.98	37
5	9	94.04	39
6	10	95.93	39
7	15	95.93	78
8	20	96.01	127

model for prediction, and the model gives the top ten labels of the sample image. Consider two images for each label from the training dataset, and apply DELF on these images and sample image. The inliers are calculated for these images using the Ransac algorithm. Image label which has a good number of inliers with sample image is the landmark of that sample image. The result indicates that the proposed model can recognize the landmark correctly. The dataset contains 100 classes. The landmarks defined within the dataset were predicted with 95.93% accuracy. Additionally, this system cannot retrieve all the images of a landmark when a sample query image is given and an API can be developed for the easy access of the model by the user, this need to be studied in the future works.

References

1. Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network.

In: 2017 International conference on engineering and technology (ICET). IEEE, 2017
2. Targ S, Almeida D, Lyman K (2016) Resnet in resnet: generalizing residual architectures. arXiv preprint arXiv:1603.08029 (2016)
3. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data 6(1):1–48
4. Noh H et al (2017) Large-scale image retrieval with attentive deep local features. In: Proceedings of the IEEE international conference on computer vision
5. Nimbare N et al (2021) A hybrid approach for landmark recognition using deep local features and residual network-50. ITM web of conferences, vol 40. EDP Sciences
6. Jha R et al (2021) Landmark recognition using VGG16 training. Smart Sustain Intell Syst 17–39
7. Wu H, Chen M (2020) Chinese landmark recognition. In: 2020 ınternational conference on computing, networking and communications (ICNC). IEEE, 2020
8. Dutreix M et al (2018) Google landmark recognition and retrieval challenges
9. Timmaraju AS, Chatterjee A (2014) Monulens: real-time mobile-based landmark recognition
10. Wang J et al (2012) A new automatic identification system of insect images at the order level. Knowl-Based Syst 33:102–110
11. Zheng Y-T et al (2009) Tour the world: building a web-scale landmark recognition engine. In: 2009 IEEE conference on computer vision and pattern recognition. IEEE, 2009
12. Kumar PR, Ch Srikanth, Sailaja KL (2016) Location ıdentification of the ındividual based on ımage metadata. Procedia Comput Sci 85:451–454
13. Chen JIZ, Hengjinda P (2021) Early prediction of coronary artery disease (cad) by machine learning method-a comparative study. J Artif Intell 3.01:17–33
14. Balasubramaniam V (2021) Artificial intelligence algorithm with SVM classification using dermascopic images for melanoma diagnosis. J Artif Intell Capsule Netw 3(1):34–42

A Prediction System for Agricultural Crops Using Supervised Learning

K. Deepa, M. Karthi, P. Kavin, S. Rahulsankar, and E. Vengaimani

Abstract The prices of agricultural commodities are constantly varying. It is affected by various factors such as weather conditions, plant diseases, labor charges, production, demand, and supply of agricultural products. The farmers face problems when the crop is not worth the price and when unaware of the marketing price. This makes it difficult for them to do agriculture. Future crop prices can be estimated using machine learning algorithms such as decision trees, logistic regression, and support vector machine (SVM). In this paper, we analyzed five supervised algorithms and chose SVM as it gave better results. The dataset contains the prices of agricultural commodities from various states and their nutrient value. At last, the XGBoost boosting algorithm is applied to improve the performance of the model.

Keywords Agricultural products · Price prediction · Machine learning · Crop price · Support vector machine

1 Introduction

Agriculture has been one of the traditional businesses since the olden days. It plays a very important role and adds value to the country's development. Many people are employed in the agricultural sector [1, 2]. The farmers cultivate the crops according to their assumptions and inner thought. But the assumption may or may not be right [3, 4]. The assumption is not backed by any thorough research or experimentation. The assumption can also fail due to some factors like climatic conditions and market expectations. In this situation, the farmer incurs a huge loss with his produce [5, 6]. Machine learning has been able to successfully predict future outcomes in many fields such as healthcare, finance, and transportation. It can also be used in agriculture to predict future outcomes.

Machine learning is a subfield of artificial intelligence that provides computer systems the ability to learn from data and gradually improve themselves. It is a

K. Deepa (✉) · M. Karthi · P. Kavin · S. Rahulsankar · E. Vengaimani
Department of Computer Science and Engineering, Karur, Tamilnadu 639113, India
e-mail: deepak.cse@mkce.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_33

growing field in today's world. It is used in many real-world applications such as Google Maps, and Alexa. Supervised learning is a subcategory of machine learning that uses labeled data to train the algorithm and predict the result. In labeled data, the data is grouped into one or more labels [7]. There are many supervised machine learning algorithms available such as linear regression, logistic regression, decision trees, and SVM.

Agricultural crop prices can be predicted using supervised learning algorithms. They can be used to predict future prices of agricultural commodities by analyzing past historical data [8]. The predictions made can be quite useful for farmers to make better decisions about the right crop to cultivate [9]. In this paper, we have used SVM to predict agricultural crops prices. The dataset for the algorithm is prepared by collecting the crop prices from the open government data website (data.gov.in) and nutrients data from other sources [10]. The dataset contains the details of crops along with prices and nutrients for many states of India [11].

Zhao [12] used an ARMA model to forecast the price of soybean in his research. He predicted the futures of the soybean market and comprehensively presented the results. As from our researched surveys, we have analyzed that a combination of deep learning models and machine learning models shows a very good result [13]. But there is an increase in time complexity while comparing with the other models. Many researchers have only given a model for the problems, and they haven't mentioned about their interface [14] The model selection framework performs better than traditional models in predicting the product's price [15, 16].

2 Related Works

Zhao's [12]. Research first estimates whether machine learning algorithms can be used for predicting crop prices. Next, they have analyzed the basic features which affect the prices of commodities. Then they have used wavelet analysis to smooth the data and then used signal decomposition to obtain layered data. This data is then used by the ARIMA model. They have used soybean prices to predict future prices. They have also shown statistical diagrams and tables to show the correctness of the model. The ARMA model performs better than the ARIMA model in univariate time series conditions. Its major drawback is that it requires more number of steps.

Sabu and Manoj Kumar [13] In this paper, the crop price is predicted using the LSTM and machine learning models. A forecasting comparison has been made between ARIMA, HOLT winter seasonal methods with 100 numbers of areca nuts. Areca nut production is high in Kerala. Price forecasting helps to maximize the revenue of farmers by giving some good insights. Based on their work, LSTM gives a good economic outcome and it is also a parsimonious model for forecasting the areca nut prices in Kerala. The main drawback of this model is that using deep learning models requires lots of data and takes high time for the decision.

Wihartiko et al. [14] The model is used to find the product price prediction with an impact of change. They have used artificial intelligence, regression, and data mining

techniques. The prediction of prices is based on the needs of stakeholders. Based on the products, they have been classified into animal products, agricultural food products, and specific food products. The neural network method has been used for getting higher accuracy. As per the user's needs, the system has been developed to benefit all stakeholders. They need various stakeholders to anticipate the prices and need more good classifiers to classify the data.

Zhang and Sanggyun [17] used a hybrid model based on SVM to predict agricultural products price. Fuzzy information granulation is used to split the data into three components—Low, R and Up. These components indicate the immensity and tendency of price. The data is then fed to the mind evolutionary algorithm (MEA)-SVM to determine the undulation rate in price. The model's performance can be enhanced by tuning the kernel and penalty parameters. The MEA-SVM hybrid model can also be used in other areas like commercial products price prediction. But the model has some defects like miscalculation and prediction errors in the long term.

Dhanapal et al. [15] In this paper, the crop price is predicted using the decision tree regression model. The decision tree algorithm is a supervised learning algorithm that uses simple decision rules to predict the output. They used previous rainfall and WPI data to predict crop prices for the next 12 months. The model was trained on many crop prices like wheat, paddy, and cotton. They also got a good-looking web page which was developed using the Flask framework. The web page shows the crop price and its increase and decrease percentage. But the algorithm failed to take more advanced features like pest outbreaks and disease outbreaks.

Bayona-Ore et al. [18] This paper presented a comprehensive overview of different algorithms and models used over the course of time for predicting agricultural products price. From the research, we can observe that neural networks were the most used for predicting product prices. Neural networks provide better accuracy and performance. But it does not mean that it always provides better results in the real world. Other than that, decision tree algorithms and support vector machines were also used by researchers. The research also shows that there is no single tool or metric to assess the performance of different models.

Zhang et al. [19] This paper shows an effective model selection framework to compare and choose from different available models. They have used 29 features and three models—ANN, SVM, and extreme learning machine to predict and compare the prices. They have also used the minimum redundancy maximum relevance (MRMR) technique to reduce the surfeit in features. Their model selection framework performs better than other simple models. It also shows that there is variation in the optimal model for different categories of time series. Their framework can be further improved by choosing some advanced classifiers.

Wu et al. [20] This paper uses a mixed model. This model is a combination of the PLS regression method and the ARIMA model using space and time factors. The model has been used to predict the weekly prices of agricultural products in various markets. The warning model is used to urgent the change of daily price after making necessary data preprocessing. The price change in different commodities also is used in forecasting prices. Here they have done analysis on the customers and then collected the analyzed data. This data gives a big turnover in the agricultural

commodities and gives a more significant result. Their models take more time to calculate the results, and this model is being made as a framework.

Champaneri et al. [21] Random forest is a popular supervised machine learning algorithm that can be used for both classification and regression. The algorithm generates a series of decision trees, and the best output is taken out from the decision trees and produced as a result. The paper outlines some of the problems in the agricultural sector like climatic changes. To solve those problems and to increase the yield of agricultural produce, a random forest algorithm is used. It is used to predict the price of agricultural products, and the results are shown on a webpage. Maharashtra's agricultural dataset is used for training and testing. Through this webpage, we can predict the prices of the different crops for the different seasons in each Maharashtra District.

Vinaya Kumar et al. [22] A survey about the impact on farmer's life. This paper shows why farmers leave agricultural fields. There are 12 important factors involved to avoid agriculture as a livelihood. Their suicides had occurred because of mental, social, economical, and personal problems. The major problems faced are global warming, climate change, high investment in manure, genetically modified seeds, and external resources. The majority of the Indian farmers own very few farming areas. So they cannot get a high amount of profit.

Venugopal et al. [23] used supervised learning algorithms to predict crop prices in various districts of Kerala. They have gathered a dataset from government websites that include production and rainfall data. They have compared three algorithms—logistic regression, naive Bayes, and random forest to obtain the best algorithm for prediction. According to their research, the random forest algorithm has performed better than other algorithms with an accuracy of 92.81%. For the server side, Heroku is used. They have also created an android application to help farmers to view the results of the analysis.

van Klompenburg et al. [24] presented a literature review of the various prediction models used for yield prediction. Initially, they have defined a search string to be matched by papers published in six different databases. After that, the chosen papers were filtered based on inclusion and exclusion rules. Next, four questionnaires were defined and the papers were analyzed and summarized based on those questionnaires. According to their research, the most used approach is the artificial neural networks and next to that, convolutional neural networks and LSTM were also used.

Suruliandiet et al. [25] discussed various feature selection techniques and supervised models to predict crop price. They have used a dataset from the agricultural department of Tenkasi District. The dataset contains 16 features out of which 12 were related to soil, and the remaining were related to the environment. Various feature selection techniques were compared, and the best technique was suggested. The selected technique with bagging classifier produced better results compared to algorithms without bagging classifier.

Thomas et al. [26] took soil characteristics such as nitrogen, potassium, and phosphorus content in the soil as the base parameters for the prediction of a particular crop. They used a support vector machine algorithm with cross-based validation.

This model provided better results compared to other models like Naive Bayes, and decision tree classifier.

Gosai et al. [27] have taken different soil parameters like Ph, nutrients such as nitrogen, potassium, and phosphorous, moisture content to build a prediction system. Different classifiers were used, and their efficiency was compared. According to the prediction results, XGBoost performed better with an accuracy of 99.31%. But, they have not developed any user interface or mobile application to be used by all farmers.

Suganya et al. [28] used supervised machine learning to predict the yield of different crops. They have taken the different flavors of supervised learning such as decision tree, random forest, and support vector machine and compared their performance. According to their results, logistic regression performed well compared to other supervised algorithms. Further, they also did not develop any web page to show the results to farmers.

Masare et al. [29] The system for maximizing the yielding rate of crops using machine learning algorithm. This module explains the crop prediction module and analyzes the need to filter the plants by soil, weather, rainfall, and required nutrients of the plant. It is used for data mining algorithms, Naïve Bayes algorithms, conditional probability, and final probability. Nowadays, farmers use the estimations by their previous experience and errors. By using this module, they can reduce their errors and get accurate results with the help of this proposed system.

Patil et al. [30] Crop prediction system is using machine learning algorithms. This model will cover all types of farming techniques and also gets more advantages to the smallholder farmers. It will give the result of profit percentage and the advanced level of farming techniques. It is used for Naïve Bayes MapReduce algorithms, data mining algorithms, and supervised learning algorithms. It follows soil type, Ph values, temperature level, weather conditions, and cultivation cost.

Ghutake et al. [31] created a website for the prediction of a crop price using suitable machine learning models because farmers are unable to predict the right crops for there, which makes a huge loss for them. By using their ML algorithm that gives 92% of the accuracy and using this the farmers can get the right crop for their yield and crop pattern and storage of the harvested crop insights. In this prediction, they have used a real-time dataset from data.gov.in, random forest prediction, and decision tree regression.

Nandhini and Gowri Shankar [32] proposed a paper that covers image processing and analysis of seeds, with particular fields for specific seeds. In this system, sample pictures of seeds are uploaded based on their quality. The image recognition system is very helpful for understanding and analyzing the ways and possibilities of occurring problems. It uses support vector machine algorithms, ID3 algorithms, and recommendation systems. The image processing system is used to avoid unwanted noises occurring in images.

Bondre and Mahagaonkar [33] have found that predicting crops and their fertilizer is a tedious task for the farmers. So they have decided to use machine learning models for predicting crops and fertilizer. These models will give better results. For their prediction, they have used a super vector machine and it gives 99.47% of accuracy in comparison with the random forest algorithm (86.35%). They have also created

a mobile application for farmers to upload images for finding the diseases of crops using image processing. By implementing this smart irrigation system farmers can able to get a higher yield production in their farming.

From the research, many of the researchers used neural networks as they provided better performance [34–38]. But the neural models require a large amount of training data. It also requires more time for model execution. Some research predicted the price for a particular crop in a particular period. Very few researchers have developed web pages to show the insights [39, 40].

3 Proposed Methodology

The proposed methodology architecture is given in Fig. 1. It consists of the following steps.

1. Dataset collection
2. Data preprocessing

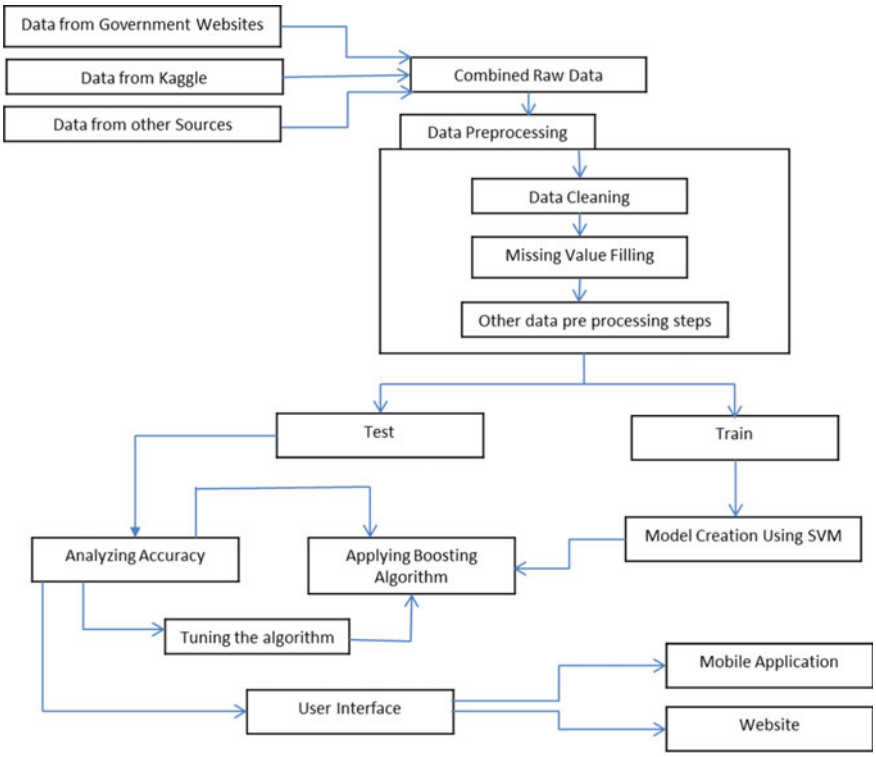


Fig. 1 Architecture diagram for price prediction system

State_Name	District_Name	Crop_Year	Season	Crop	Area	Production	Price per kg
Andaman and Nicoba	NICOBARS	2000	Kharif	Arecanut	1254	2000	240
Andaman and Nicoba	NICOBARS	2000	Kharif	Other Kharif puls	2	1	70
Andaman and Nicoba	NICOBARS	2000	Kharif	Rice	102	321	50
Andaman and Nicoba	NICOBARS	2000	Whole Year	Banana	176	641	55
Andaman and Nicoba	NICOBARS	2000	Whole Year	Cashewnut	720	165	650
Andaman and Nicoba	NICOBARS	2000	Whole Year	Coconut	18168	65100000	140
Andaman and Nicoba	NICOBARS	2000	Whole Year	Dry ginger	36	100	260
Andaman and Nicoba	NICOBARS	2000	Whole Year	Sugarcane	1	2	25
Andaman and Nicoba	NICOBARS	2000	Whole Year	Sweet potato	5	15	50
Andaman and Nicoba	NICOBARS	2000	Whole Year	Tapioca	40	169	30
Andaman and Nicoba	NICOBARS	2001	Kharif	Arecanut	1254	2061	240
Andaman and Nicoba	NICOBARS	2001	Kharif	Other Kharif puls	2	1	70
Andaman and Nicoba	NICOBARS	2001	Kharif	Rice	83	300	50
Andaman and Nicoba	NICOBARS	2001	Whole Year	Cashewnut	719	192	650
Andaman and Nicoba	NICOBARS	2001	Whole Year	Coconut	18190	64430000	140
Andaman and Nicoba	NICOBARS	2001	Whole Year	Dry ginger	46	100	260
Andaman and Nicoba	NICOBARS	2001	Whole Year	Sugarcane	1	1	25
Andaman and Nicoba	NICOBARS	2001	Whole Year	Sweet potato	11	33	50
Andaman and Nicoba	NICOBARS	2002	Kharif	Rice	189.2	510.84	50
Andaman and Nicoba	NICOBARS	2002	Whole Year	Arecanut	1258	2083	240
Andaman and Nicoba	NICOBARS	2002	Whole Year	Banana	213	1278	55
Andaman and Nicoba	NICOBARS	2002	Whole Year	Black pepper	63	13.5	490
Andaman and Nicoba	NICOBARS	2002	Whole Year	Cashewnut	719	208	650
Andaman and Nicoba	NICOBARS	2002	Whole Year	Coconut	18240	67490000	140
Andaman and Nicoba	NICOBARS	2002	Whole Year	Dry chillies	413	28.8	142

Fig. 2 Combined dataset

3. Data analysis
4. SVM model creation
5. Application of XGBoosting algorithm

3.1 Dataset Collection

Dataset plays a key role in machine learning. It is a collection of records. At first, datasets from Kaggle and open government data (OGD) which contain agricultural commodities and their prices have been collected. Then, those raw data are combined into one. The combined dataset contains the prices of 124 commodities from different districts all around India. The dataset has features like state name, district name, year, season, crop, area, production, and price. It contains around 2.46 lakh instances. The combined dataset is shown in Fig. 2.

3.2 Data Preprocessing

It is one of the data mining techniques which is very helpful for transforming the raw data into clean data that is in a useful and coherent format. The steps which are used

in data preprocessing are data cleaning, data transformation, and data reduction. Data cleaning deals with missing and irrelevant data. It is very helpful for handling the missing and noisy data. Missing data refers to the absence of certain important data. In this dataset, the missing values are filled by taking the mean of that column. Noisy data refers to data that is meaningless. The noisy data can be generated by various means. Some of them are faulty data collection and data entry errors. The noisy data can be managed by methods like binning method, clustering, and regression. These methods help to smooth the data by sorting them. Data transformation is very helpful for transforming the data into an appropriate form that is suitable for the mining process. The main step that is being followed in data transformation is normalization. Here, the label encoder is used to convert non-numerical data into numerical data. Data reduction is one of the main data mining steps which is used to manage a huge and large amount of data. Data analysis becomes difficult and harder in a case where the size of data is huge. It aims to increase the storage efficiency by decreasing the data storage and is also used to decrease the cost of the data analysis process. Here, the attribute subset selection is mainly used to achieve the objective of the data reduction technique. In attribute subset selection, the highly relevant attributes are grouped together and the rest will be discarded.

3.3 Data Analysis

It is an important step in exploring datasets prior to developing a machine learning model. It is done to get meaningful insights from data using different data visualization techniques. It helps to identify trends and outliers in the data. Each agricultural commodity in the dataset is analyzed to get insight into that commodity. Figure 3 represents a bar chart for the variation of ash gourd prices in different districts.

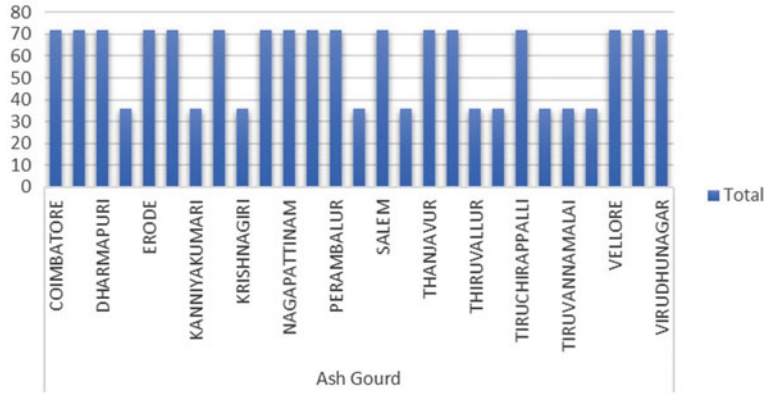


Fig. 3 Analysis of fruit ash gourd

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test= train_test_split(x, y, test_size= 0.25, random_state=10)
#feature Scaling
from sklearn.preprocessing import StandardScaler
st_x= StandardScaler()
x_train= st_x.fit_transform(x_train)
x_test= st_x.transform(x_test)
from sklearn.svm import SVC
classifier = SVC(kernel='linear')
classifier.fit(x_train, y_train)

y_pred= classifier.predict(x_test)

print(classifier.score(x_test,y_test))
```

```
0.8785494855582465
```

Fig. 4 Accuracy of SVM

3.4 SVM Model Creation

The preprocessed data has been split into train and test data in the ratio of 7:3. Five supervised machine learning algorithms, namely linear regression, logistic regression, random forest, decision trees, and SVM, have been taken for comparison. Among the chosen algorithms, SVM is more suited for the dataset and provides better results. SVM is an algorithm that helps to solve both linear and nonlinear problems with the help of regression and classification. It works by mapping the data in an *n*-dimensional space and creating a hyperplane that perfectly separates the data points. The hyperplane divides the data points in a way that the distance between data points and hyperplane is large. It has provided 87.8% accuracy which is high compared to other algorithms. So SVM is chosen for price prediction. The accuracy of SVM is shown in Fig. 4.

3.5 Application of XGBoost Algorithm

After analyzing the accuracy, the XGBoost boosting algorithm is applied to increase the performance. XGBoost is an algorithm that uses gradient boosting to boost weak learners. It creates an ensemble of the base model results and tries to predict the dependent variable by aggregating the results from the ensemble. It employs software and hardware-level optimizations to reduce the running time. It is mainly used for increasing the speed and performance of the model. After applying the XGBoost algorithm, the obtained root mean squared error was 0.19 and it was shown in Fig. 5.

```
from xgboost import XGBRegressor
my_model = XGBRegressor()
my_model.fit(x_train, y_train, verbose=False)
predictions = my_model.predict(x_test)
from sklearn.metrics import mean_absolute_error
print("Mean absolute error : " + str(mean_absolute_error(predictions, y_test)))
print("Root mean squared error : "+str(np.sqrt(mean_absolute_error(predictions, y_test))))
```

```
Mean absolute error : 0.03945971559126127
Root mean squared error : 0.1986446968616612
```

Fig. 5 Root mean square error of XGBoost algorithm

4 Future Work

In the future, we are planning to test various algorithms and compare their performances. Further enhancements can be made in the dataset to include more features such as weather conditions, prices of fertilizers, and seeds. We also aim to improve the model by using real-time data from government websites and agricultural institutions.

5 Conclusion

Agriculture is one of the most important drivers of human life and economic development. Agriculture and farmers are vital to providing the basis for all of our food needs. But farmers' income is affected due to unprecedented market and weather conditions. Machine learning has the potential to improve the farmer's life by predicting the yield, price and suggesting the suitable crop for cultivation based on region and nutrient parameters. Among many other machine learning algorithms, we have used the SVM algorithm for price prediction. The research in machine learning is still going on to improve the algorithms and identify new algorithms that have higher accuracy and efficiency.

References

1. Deepa K, Kokila M, Nandhini A, Pavethra A, Umadevi M (2020) Rainfall prediction using CNN. Int J Adv Sci Technol 29(7):1623–1627
2. Deepa K, Thilagamani S (2019) Segmentation techniques for overlapped latent fingerprint matching. Int J Innov Technol Exploring Eng 8(12):1849–1852
3. Deepa K, Thilagamani S (2021) A spatial-frequency feature ensemble for detecting cervical dysplasia from pap smear images. J Med Imag Health Inf 11:2771–2777
4. Pradeep D, Sundar C (2020) QAOC: Noval query analysis and ontology-based clustering for data management in Hadoop. 108:849–860
5. Thilagamani S, Shanti N, Gaussian (2014) gabor filter approach for object segmentation. J Comput Inf Sci Eng 14(2)

6. Rajesh Kanna P, Santhi P (2022) Hybrid intrusion detection using map reduce based black widow optimized convolutional long short-term memory neural networks. Expert Syst Appl 194
7. Murugesan M, Thilagamani S (2020) Efficient anomaly detection in surveillance videos based on multi layer perception recurrent neural network. J Microprocess Microsyst 79(11)
8. Thilagamani S, Nandhakumar C (2020) Implementing green revolution for organic plant forming using KNN-classification technique. Int J Adv Sci Technol 29(7):1707–1712
9. Rajesh Kanna P, Santhi P (2021) Unified deep learning approach for efficient intrusion detection system using integrated spatial–temporal features. Knowl-Based Syst 226
10. Santhi P, Mahalakshmi G (2019) Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction. Int J Eng Adv Technol 8(4):839–846
11. Pandiaraja P, Rajesh KP (2019) An efficient sentiment analysis approach for product review using Turney algorithm. J Procedia Comput Sci 165:356–362
12. Zhao H (2021) Futures price prediction of agricultural products based on machine learning. Neural Comput Appl 837–500
13. Sabu KM, Manoj Kumar TK (2019) Predictive analytics in agriculture: forecasting prices of Arecanuts in Kerala. In: Third international conference on computing and network communications, pp 699–798
14. Wihartiko FD, Nurdiati S, Buono A, Santosa E (2021) Agricultural price prediction models: a systematic literature review. In: 11th annual international conference on industrial engineering and operations management Singapore, Singapore, pp 2927–2934
15. Dhanapal R, AjanRaj A, Balavinayagapragathish S, Balaji J (2021) Crop price prediction using supervised machine learning algorithms. J Phys Conf Series
16. Perumal P, Suba S (2022) An analysis of a secure communication for healthcare system using wearable devices based on elliptic curve cryptography. J World Rev Sci. Technol Sustain Develop 18(1):51–58
17. Zhang Y, Sanggyun (2018) A novel agricultural commodity price forecasting model based on fuzzy information granulation and MEA-SVM model. Math Probl Eng
18. Bayona-Ore S, Cerna R, Tirado Hinojosa E (2021) Machine learning for price prediction for agricultural products. Wsess transactions on business and economics. ISSN: 2224-2899, 18
19. Zhang D, Chen S, Ling L, Xia Q (2020) Forecasting agricultural commodity prices using model selection framework with time series features and forecast horizons. Nat Sci Found 8
20. Wu H, Wu1 H, Zhu M, Chen W, Chen W (2017) A new method of large-scale short-term forecasting of agricultural commodity prices: illustrated by the case of agricultural markets in Beijing. J Bigdata 1–22
21. Champaneri M, Chandvidkar C, Chachpara D, Rathod M (2020) Crop yield prediction using machine learning. Int J Sci Res (IJSR)
22. Vinaya Kumar HM, Chauhan NB, Patel DD, Patel JB (2019) Predictive factors to avoid farming as a livelihood. J Econ Struct 1–18
23. Venugopal A, Aparna S, Mani J, Mathew R, Williams V (2021) Crop yield prediction using machine learning algorithms. Int J Eng Res Technol (IJERT) 9:87–91. ISSN: 2278-0181
24. van Klompenburg T, Kassahun A, Catal (2020) Crop yield prediction using machine learning: a systematic literature review. Comput Electron Agric 1–18
25. Suruliandi A, Mariammal G, Raja SP (2021) Crop prediction based on soil and environmental characteristics using feature selection techniques. Math Comput Modell Dyn Syst 27:116–140
26. Thomas KT, Varsha S, Saji MM, Varghese L, Er. Thomas J (2021) Crop prediction using machine learning. Int J Future Gener Commun Netw 13(3):1896–1901. ISSN: 2233-7857
27. Gosai D, Raval C, Nayak R, Jayswal H, Patel A (2021) Crop recommendation system using machine learning. Int J Sci Res Comput Sci Eng Inf Technol (IJSRCSEIT) 7(3):558–569. ISSN: 2456-3307
28. Suganya M, Dayana R, Revathi R (2020) Crop yield prediction using supervised learning techniques. Int J Comput Eng Technol 11(2):9–20. ISSN Print: 0976-6367 and ISSN Online: 0976-6375

29. Masare Y, Mahale S, Kele M, Upadhyay A, Nanwalkar BR (2021) The system for maximize the yielding rate of crops using machine learning algorithm. Int J Eng Res Technol (IJERT) 10(6):453–458. ISSN: 2278-0181
30. Patil P, Panpatil V, Kokate S (2021) Crop prediction system using machine learning algorithms. Int Res J Eng Technol (IRJET) 7(2):748–752. p-ISSN: 2395-0072
31. Ghutake I, Verma R, Chaudhari R, Amarsinh V (2021) An intelligent crop price prediction using suitable machine learning algorithm. ITM web of conferences 40, 03040, ICACC
32. Nandhini N, Gowri Shankar J (2020) Prediction of crop growth using machine learning based on seed features. ICTACT J Soft Comput 11(1):2232–2236. ISSN: 2229-6956
33. Bondre DA, Mahagaonkar S (2019) Prediction of crop yield and fertilizer recommendation using machine learning algorithms. Int J Eng Appl Sci Technol 4(5):371–376. ISSN: 2455-2143
34. Pandiaraja P, Sharmila S (2020) Optimal routing path for heterogenous vehicular adhoc network. Int J Adv Sci Technol 29(7):1762–1771
35. Karthikeya HK, Sudarshan K, Shetty DS (2020) Prediction of agricultural crops using KNN algorithm. Int J Innov Sci Res Technol Comput Sci 5(5):1422–1424. ISSN: 2456-2165
36. Thoms KT, Varsha S, Saji MM, Varghese L, Er. Thomas J (2020) Crop prediction using machine learning. Int J Future Gener Commun Netw 13(3):1896–1901. ISSN: 2233-7857
37. Santhi P, Lavanya S (2020) Prediction of diabetes using neural networks. Int J Adv Sci Technol 29(7):1160–1168
38. Gunasekar M, Thilagamani S (2020) Performance analysis of ensemble feature selection method under SVM and BMNB classifiers for sentiment analysis. Int J Sci Technol Res 9(2):1536–1540
39. Gunasekar M, Thilagamani S (2019) Towards sentiment analysis and opinion mining from multimodal data. Int J Recent Technol Eng 8(1):272–274
40. Deepak G, Deepika J, Dharshini M, Vanathi B (2021) Crop yield prediction based on ensemble model using historical data. Int J Adv Res Comput Sci Eng 7(2):668–673. ISSN: 2395-4396

The Impact of the Industrial Revolution 4.0 on Small Medium Enterprises (SME) Agribusiness in the Electronic Payment Sector

Gregorius Josevan Harintoro, Gregorius Bagus Hepi Widyantoro, Aldrick Aaron, Fahmi Muhamad, Ford Lumban Gaol, Tokuro Matsuo, and Natalia Filimonova

Abstract The Industrial Revolution 4.0 has significantly altered how people think, live, and interact. This period has disrupted a wide range of human activities, not just in the sphere of technology, but also in other areas like economics, social relations, and politics. A mix of technologies that blend the physical, digital, and biological realms, together known as cyber-physical systems, characterizes the Industrial Revolution 4.0 age, thus requiring extensive action to help analyze and mitigate. Electronic payment is also a part of the Industrial Revolution 4.0, and it has so many impacts. In this research, we are going to see what it does to small businesses. This type of research will be using quantitative method with 43 samples from small businesses' owners using questionnaire method. The results show that electronic payment has so much impact on small businesses such as more alternative payment options.

G. J. Harintoro (✉) · G. B. H. Widyantoro · A. Aaron · F. Muhamad
School of Information System, Bina Nusantara University, Alam Sutera, Kota Tangerang, Indonesia
e-mail: gregorius.harintoro@binus.ac.id

G. B. H. Widyantoro
e-mail: gregorius.widyantoro@binus.ac.id

A. Aaron
e-mail: aldrick.aaron@binus.ac.id

F. Muhamad
e-mail: fahmi.muhamad@binus.ac.id

F. L. Gaol
Binus Graduate Program—Doctor of Computer Science, Jakarta, Indonesia
e-mail: fgaol@binus.edu

T. Matsuo
Advanced Institute of Industrial Technology, Tokyo, Japan
e-mail: matsuo@aiit.ac.jp

N. Filimonova
Vladimir Branch of Russian Academy of National Economy and Public Administration, Vladimir, Russia
e-mail: natal_f@mail.ru

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_34

Keywords Electronic payment · Small medium enterprises (SME) agribusiness · Industrial revolution · Economy

1 Introduction

The Industrial Revolution 4.0 radically altered how people think, live, and interact with one another. This age will cause disruptions in a variety of human activities, not only in the realm of technology, but also in sectors such as economics, social science, and politics [1]. Additionally, the Industrial Revolution 4.0 period saw the introduction of technology advancements in a variety of disciplines. Robotics, artificial intelligence, nanotechnology, quantum computing, biotechnology, and the Internet of Things are among the topics under discussion. The Industrial Revolution 4.0, sometimes referred to as the Industrial Revolution 4.0, was the fourth industrial period following the eighteenth century's First Industrial Revolution [2].

This Industrial Revolution 4.0 age is defined by a convergence of technologies that blur the boundaries between the physical, digital, and biological realms, together referred to as cyber-physical systems. Additionally, the Industrial Revolution 4.0 period is defined by the rise of technical advancements in a variety of disciplines. Robotics, artificial intelligence, nanotechnology, quantum computing, biotechnology, the Internet of Things, fifth-generation wireless technology, 3D printing, and autonomous vehicles are among the technologies in question [3].

Indonesia's Social and Economic Performance: The development of technology in development has created an ambiguous attitude among the people. On the one hand, society accepts the presence of technology, and on the other hand, the presence of modern technology produces structural problems which then penetrate all sectors of people's lives [4].

1.1 Problem Statement

The Industrial Revolution 4.0 has had a significant influence on the economic aspects of individuals living in DKI Jakarta. Few individuals in Indonesia accept technological growth, despite the fact that there are several benefits to be gained by adopting this industrial revolution. Because some individuals are still unwilling to adjust to new advancements, the development of electronic payments in Indonesia is still not broadly embraced by many people. Teenagers are the bulk of those who can adjust. The quantity of individuals who lack technological knowledge causes new issues for current advancements, and Indonesia is falling more behind other countries in technological growth.

1.2 Research Questions

During this research, there are four research questions based on the problem, and based on the questions, there will be an answer as an output of the research.

1. Does electronic payment improve the effectiveness of small business?
2. What was the economic growth like in DKI Jakarta from 2005 until now
3. Does DKI Jakarta have the sufficient human resource to face Industry 4.0?
4. What impact has industry 4.0 brought upon the people of Srengseng?

1.3 Relevance and Importance of Research

Research will figure out what suitable impact that could be implemented in electronic payment sector. This research will have any relevance and importance, and those are:

1. To figure out what does electronic payment impact to small business.
2. To observe and learn further to what aspects of people that the change will bring.
3. To learn certain changes, it will bring in regard to the quality of life for the people living in DKI Jakarta.
4. To figure out whether the available human resource in DKI Jakarta is prepared according to the 4.0 implementation.

2 Research Technique

This research will be done using a qualitative method utilizing the use of interviews and questionnaires as the main method of data collection. Holding interviews and sending questionnaires to several small medium enterprises (SME) agribusiness that has been carefully chosen to help measure the impact Industrial Revolution 4.0 has brought upon said small medium enterprises (SME) agribusiness located in the DKI Jakarta region [3, 5]. Interviews will be held online through means of an online conference video call software. Small medium enterprises (SME) agribusiness was chosen as the main subject since they can contribute an immense amount of impact on the economic situation in DKI Jakarta. The questionnaire will be done through an online form aptly named Google Forms, and the link which will lead to the questionnaire will also be distributed online to the owners of several small medium enterprises (SME) agribusiness located in DKI Jakarta. Figure 1 shows the flow of process [4, 6].

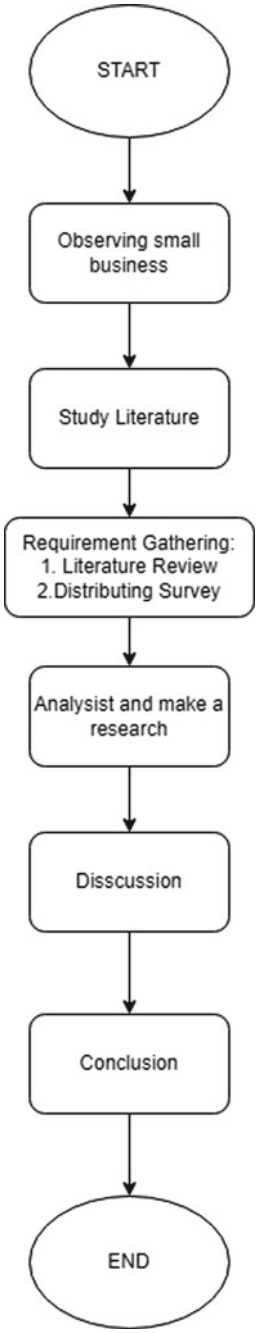


Fig. 1 Flowchart of research technique

3 Research Method

This research will be using a quantitative method. A face-to-face questionnaire will not be held due to the current pandemic situation that is currently ongoing. Hence, the method of online questionnaire will be used as a substitute in order to replace direct questionnaire which will be held with the aid of Microsoft Forms as the media for answering the questions which have been prepared for the questionnaire aspect of data collection [4]. Another issue that might occur is the inability to partake in the data collection since it might be a possibility due to the lack of a device to partake in the questionnaire and for some small business owner that does not have access to a device that is capable in aiding the data collection will be a waste since a physical meeting is currently close to impossible to do due to the current pandemic situation.

Nonetheless, this research aims to expose certain impacts that have been brought by the Industrial Revolution 4.0 which in conclusion will add another fact that said industrial revolution has left some small business owner in the dark due to lack of access to certain technology devices, and for those who do have access might not have received the full impact that this revolution has brought.

3.1 Literature Review

3.1.1 Industrial Revolution

The industrial revolution is a period of fast change in the economy and industry and characterized by increased efficiency, which has a ripple effect on other spheres of life. There have been four industrial revolutions since the First Industrial Revolution, and the Industrial Revolution 4.0 is now underway [7]. The First Industrial Revolution was characterized by the widespread use of machines to facilitate labor. Mass manufacturing started with the Second Industrial Revolution. Computers and industrial automation were key components of the industrial revolution. Additionally, this Industrial Revolution 4.0 is characterized by the advent of cyber-physical systems [8, 9].

Industry 4.0 is predicted to provide significant advantages. The bulk of industry 4.0's potential advantages includes increased manufacturing speed and flexibility, enhanced customer service, and more revenue. If achieved, these potential advantages will have a beneficial effect on a country's economy. Industry 4.0 has tremendous benefits, but it also has certain challenges to solve. According to Morrar and Mousa,

the rise of resistance to demographic and social changes, political instability, limited resources, the danger of natural catastrophes, and the adoption of environmental-friendly technology are the problems that nations confront while adopting industry 4.0 [10, 11].

Industrial Revolution 4.0

The Industrial Revolution 4.0 is a stage in the development of knowledge during which the distinctions between the physical, digital, and biological worlds become more blurred. The educational information technology (IT) landscape has shifted dramatically as a result of the Industrial Revolution 4.0 [12]. Managers must lead in such a manner that workers' viewpoints, ideas, and attitudes evolve through time. Rapid technological advancements have had a significant impact on the payment system that underpins economic activity in Indonesia [13]. Which was previously limited to cash transactions and has now expanded to include digital payment methods, often known as electronic money (e-money). Advances in payment technology have reduced the importance of cash (currency) as a mode of payment in favor of more efficient and inexpensive non-cash payments. Non-cash payments are often done through interbank transfers or intrabank transfers via the bank's internal network. Additionally, non-cash payments may be done using a card, such as an ATM card, a debit card, or a credit card [14, 15].

The era of the Industrial Revolution 4.0 shows an increase in technology that affects the payment system with various applications that are used by the public as a means of non-cash payments. Digital payment applications that are currently booming in the community include OVO, GO-PAY, and DANA [16]. The use of this application is to facilitate transactions in various activities, for example, for online motorcycle taxi payments, food delivery orders, electricity or telephone bill payments, PDAM payments, and many other conveniences produced by digital payment applications. Judging from the many existing practices, digital payment systems are very efficient in their use. However, there are still many Indonesians who have not used this facility. This can be proven by the large number of cash users to pay for goods or services. Consumers who use digital payment systems today are generally dominated by the middle to upper-middle class, as well as people who are already technology literate [17, 18].

Key Concepts, Theories, and Studies

No.	Title	Results	Differences
1	The influence of the Industrial Revolution era 4.0 on human resource competence Indonesian business management journal, Vol. 6 No. 1 researched by Leni Rohida, S.Sos, M.Si in October 2018, pp. 114–136 [11]	The goal of this publication is to explain how the emergence of the Industrial Revolution 4.0 has affected human resource competencies. In this era of revolution, the current industry is concerned with competency, which refers to a person's capacity to execute successfully and have excellence in a certain position. Here, characteristics related to a person's ability, which can make a difference between people. Finally, depending on the distinguishing characteristics, it will be assessed whether someone can be called to be able to do various jobs and address current difficulties	The major distinction is the statement of the problem. Many of the articles in this journal are about human resource management. This publication also uses observational methods; nevertheless, in this study, we will employ questionnaire and interviews

(continued)

(continued)

No.	Title	Results	Differences
2	Industry 4.0: the effect of the industrial revolution on entrepreneurship for economic independence nusamba journal, Vol. 3 No. 2 researched by Hamdan in October 2018, pp. 1–8 [2]	This journal examines the impact of entrepreneurship on economic freedom because of the Industrial Revolution 4.0. Essentially, during the present industrial revolution, the field economy is built on a significant shift in sophisticated technology, which has an influence on automation in almost every line of technology-related industries. Only now, the Industrial Revolution 4.0 approaches will the physical and digital worlds will be brought together, and the fundamentals will shift. This journal examines the impact of entrepreneurship on economic freedom because of the Industrial Revolution 4.0. The field economy has been impacted by a significant shift in sophisticated technology throughout the present industrial revolution	This journal uses qualitative technique, while we are using quantitative. The subject is different; it focuses mostly on Go-Jek as the major subject, as well as other online transportation options. The economic sector in DKI Jakarta is the topic of this study

3.1.2 Key Debates and Controversies

The focus of controversies for this topic is the lack of human resource possessing necessary qualities to keep up with the exponential growth that was brought with the impact of the Industrial Revolution 4.0. There are some of small businesses' owner that still couldn't implement the new technology at the moment. With it, we will not develop with technological progress and compete with other competitors.

Our position in this controversy is to help expose what might be the cause of said lack in human resource quality and what can be done to help increase the quality of the human resource that exists in Indonesia starting with a smaller region particularly in DKI Jakarta.

3.1.3 Gaps in Existing Knowledge

The major cause of debate on this problem is a lack of human resources with the requisite attributes to keep up with the exponential expansion brought on by the Industrial Revolution 4.0's influence. Our position in this debate is to assist uncover what may be causing the lack of human resource quality, as well as what can be done to improve the quality of human resource that exists in Indonesia, beginning with a smaller region in West Jakarta, DKI Jakarta [19].

Based on the previous result, both journals have given a result of the impact of implementing electronic payment in this sector, but the first journal is on a different sector, it focuses on human resources while this journal focuses on economic sector in electronic payment technology areas. The second journal is using qualitative method while they focus on Go-Jek implementation, but it also focuses on economic sector which may refer to this journal, the results show that the implementation brings so much advantages for economic sector [15].

3.1.4 Questionnaire

Specifically, we have a question about electronic payments to people who have answered our questions through the questionnaire we have made available.

1. As a small business owner, does your business has electronic payment option?
2. Do you find it helpful to have electronic payment option?
3. Do customers prefer electronic payment more than cash?
4. Did your income increase after implementing the electronic payment option?

3.1.5 Research Result

Based on our questionnaire, the result below shows that most, small business owners show that they felt helpful by using the electronic payment technology. We distribute the questionnaire to small businesses' owners in Srengseng (Table 1).

Question 1

From the first question, we may know that 90% of the respondents already implement electronic payment in their business, while the rest haven't (Fig. 2).

We may know that electronic payment is not a new thing anymore; almost everyone has it on their smartphone.

Question 2

Most respondents feel helpful, 95% of the respondents said yes (40), while the rest (3) not (Fig. 3).

Table 1 Response analysis for electronic payment option

ID	Question 1	Question 2	Question 3	Question 4
1	Yes	Yes	Yes	Yes
2	Yes	Yes	No	No
3	Yes	Yes	No	No
4	Yes	Yes	Yes	Yes
5	Yes	Yes	Yes	Yes
6	Yes	Yes	Yes	Yes
7	Yes	Yes	Yes	Yes
8	Yes	Yes	Yes	Yes
9	Yes	Yes	Yes	Yes
10	Yes	Yes	Yes	Yes
11	Yes	Yes	Yes	Yes
12	Yes	Yes	Yes	Yes
13	Yes	Yes	Yes	Yes
14	Yes	Yes	Yes	Yes
15	Yes	Yes	No	No
16	Yes	Yes	No	No
17	Yes	Yes	Yes	Yes
18	Yes	Yes	Yes	Yes
19	Yes	Yes	Yes	Yes
20	Yes	Yes	Yes	Yes
21	Yes	Yes	Yes	Yes
22	Yes	Yes	Yes	Yes
23	Yes	Yes	Yes	Yes
24	Yes	Yes	Yes	Yes
25	Yes	Yes	Yes	Yes
26	Yes	Yes	Yes	Yes
27	Yes	Yes	Yes	Yes
28	Yes	Yes	Yes	Yes
29	Yes	Yes	Yes	Yes
30	Yes	Yes	Yes	Yes
31	Yes	Yes	Yes	Yes
32	Yes	Yes	Yes	Yes
33	Yes	Yes	Yes	Yes
34	Yes	Yes	Yes	Yes
35	Yes	Yes	Yes	Yes
36	Yes	Yes	Yes	No

(continued)

Table 1 (continued)

ID	Question 1	Question 2	Question 3	Question 4
37	Yes	Yes	Yes	Yes
38	Yes	Yes	Yes	Yes
39	Yes	Yes	Yes	No
40	No	Yes	Yes	No
41	No	No	Yes	Yes
42	Yes	Yes	No	No
43	No	No	Yes	Yes

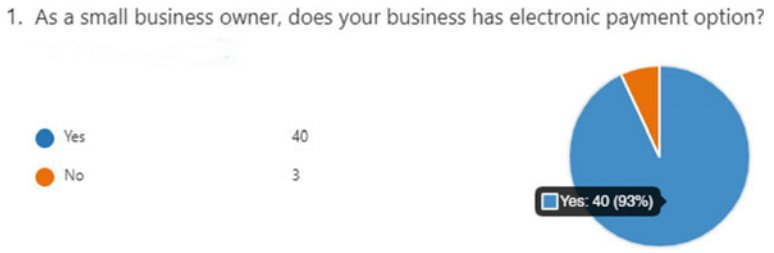


Fig. 2 Response based on question 1

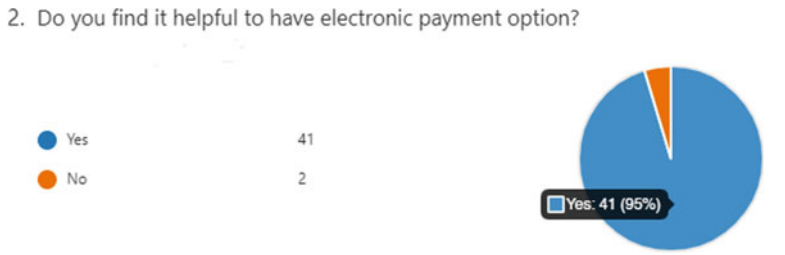


Fig. 3 Responses for question 2

Question 3

As can be seen from picture 3.6.3, 88% of the respondents (38) prefer to use it because it is simpler, while the rest 12% (5) prefer to use cash (Fig. 4).

Question 4

From table 3.6.4, 34 respondents or 81% are feeling the impact of electronic payment; they get more revenue after implementing electronic payment in their business, while 8 respondents are not (Fig. 5).

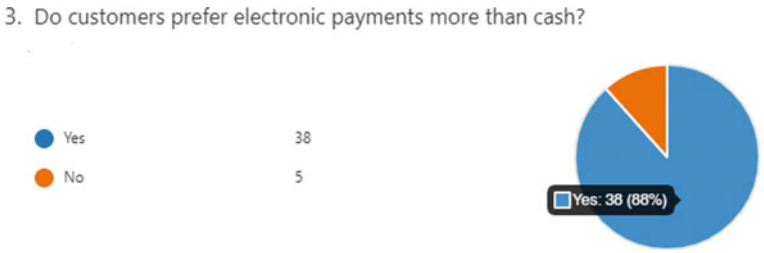


Fig. 4 Response based on question no. 3

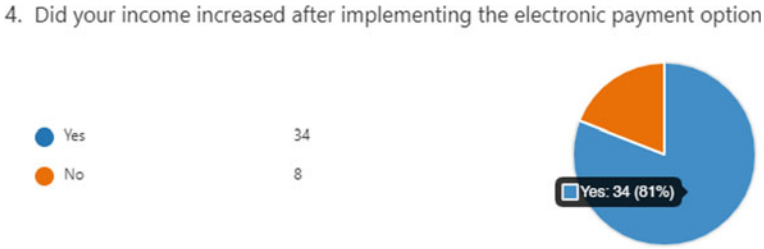


Fig. 5 Response based on question no. 4

4 Implications and Contributions to Knowledge

4.1 Practical Implications

There are two practical implications in this research which is more small business could implement electronic payment as their payment option and for other owner that still couldn't use the technology, the government should apply a training for them, because when they can use technology that can improve their business.

After reading this research, we hope that even more small medium enterprises (SME) agribusiness would utilize the benefits that can come from industrial revolution in aiding the income they can attain from their business. We will be able to see the impact that some business owner will feel on their business, including the addition of electronic payment such as e-wallet and other forms of digital payment, which can aid in the increase of new customers due to certain promotions that may be held by the application.

However, this revolution is still in its early phase in Indonesia resulting in the lack of acceptance from small businesses since it is a possibility that some businesses will only accept which can result in the reduction of customers due to their complaint of not having access to digital payments or e-wallet.

Further public education from the officials to the public mass will add further depth and knowledge to the public on how Industrial Revolution 4.0 will be able

to benefit them and their business. Further increasing the number of businesses in Indonesia adapts this change.

4.2 Theoretical Implications

We aim to demonstrate the influence of the Industrial Revolution 4.0 on the human resource management process at this point. Starting with a fresh viewpoint on current human resources and progressing to raising the qualifications and specifications of human resources required, as well as how organizations may educate their personnel to stay up with the changes and updates brought on by the Industrial Revolution 4.0. In responding to and responding to the Industrial Revolution 4.0, human resource management must consider the following ten aspects: (1) Strategic Perspective, (2) Employment, (3) Performance Management, (4) Finding Talent, (5) Training, (6) Data-Based Management, (7) Information Protection, (8) Work Environment, (9) Social Media, and (10) Work-Life Balance.

A mechanism is essential to give favorable circumstances for small medium enterprises (SME) agribusiness owners in Jakarta, as well as it should be granted to the disciplines of the economic sector. There is a fair playing field in research activities. For this industrial revolution, the government should provide training needs, scientific research, topic selection, cost estimates, and qualified personnel to participate. Universities should focus on creating a framework for students to accept this shift. This digital era may be able to assist many business owners in maintaining their operations. People should adapt to this digitalized world, as we all know it may provide them with numerous benefits. Some people's perspectives on the industrial revolution will shift because of it.

5 Discussion

Based on previous researches that have been done, the industrial revolution did have a positive and negative impact, in the economic sector cashless payments emerged. Currently, if you want to go, you only need to bring a smartphone and you can do various transactions. The presence of e-wallet provides different opinions for the public, business people who can adapt feel benefited, because the transactions that occur are very easy and fast, they are also not in trouble because they must provide change to customers. More and more people are also coming to their business. They can also sell their products online, with the presence of online transportation such as Go-Jek and Grab. They admit that the presence of technology brings blessings to their business. On the other hand, there are also those who think that it is difficult. This happens to businesspeople who are "old" because they cannot adapt to existing technology. Technology is also considered very difficult because they cannot afford to buy a smartphone.

6 Conclusion

In Srengseng, the industrial revolution has given so many advantages to small businesses, they could get more income by using new technologies, some of the business still doesn't accept the new technology, because many of them could not use the technology.

The economic growth in DKI Jakarta is getting better, since there is a crisis that caused by the US Bank and the latest recession from COVID-19, DKI Jakarta is still improving their economic growth.

The government should provide education to business owners so that they can feel the impact of the Industrial Revolution 4.0. This impact is also considered very good for them, because nowadays it is rare for people to carry cash, most of them are more inclined to cashless transactions. The government can provide training for business owners, provide smartphones for those who cannot afford it and support their business so that their business can continue to survive in accordance with the times. With technology, it also makes it easier for them to reach markets that they could not reach before, such as if selling online, they can sell it to everyone who is interested in their product.

In Srengseng itself, the industry 4.0 has given so many impacts for small businesses' owner, based on the research result, most of them already implemented and getting more advantages by them.

Acknowledgements This research was funded by International Grant from Binus University with the contract #No: 017/VR.RTT/III/2021.

References

1. al Qardh J, Tarantang J, Awwaliyah A, Astuti M, Munawaroh M (nd) 60 IAIN Palangka Raya PERKEMBANGAN SISTEM PEMBAYARAN DIGITAL PADA ERA REVOLUSI INDUSTRI 4.0 DI INDONESIA
2. Hamdan H (2018) Industri 4.0: Pengaruh revolusi industri pada kewirausahaan demi kemandirian ekonomi. Jurnal Nusantara Aplikasi Manajemen Bisnis 3(2):1–8
3. Satya VE (2018) Strategi Indonesia menghadapi industri 4.0. Info Singkat, 10(9):19–24
4. Mumtaha HA, Khoiri HA (2019) Analisis Dampak Perkembangan Revolusi Industri 4.0 dan Society 5.0 Pada Perilaku Masyarakat Ekonomi (E-Commerce). Jurnal Pilar Teknologi: Jurnal Ilmiah Ilmu Ilmu Teknik 4(2)
5. Bouguerra I (2018) The effect of electronic payment systems on economic growth in Jordan during the period (2003–2015). el-Bahith Rev 18(1):571–581
6. Ali MA, Hussin N, Abed IA (2019) Electronic payment systems: architecture, elements, challenges and security concepts: an overview. J Comput Theor Nanosci 16(11):4826–4838
7. Manda MI, Ben Dhaou S (2019) Responding to the challenges and opportunities in the 4th industrial revolution in developing countries. In: Proceedings of the 12th international conference on theory and practice of electronic governance, pp 244–253
8. Tri NM, Hoang PD, Dung NT (2021) Impact of the industrial revolution 4.0 on higher education in Vietnam: challenges and opportunities. Linguist Cult Rev 5(S3):1–15

9. Wirawan V (2020) Penerapan E-Government dalam Menyongsong Era Revolusi Industri 4.0 Kontemporer di Indonesia. Jurnal Penegakan Hukum Dan Keadilan 1(1):1–16
10. Morrar R, Arman H, Mousa S (2017) The fourth industrial revolution (Industry 4.0): A social innovation perspective. Technol Innov Manag Rev 7(11):12–20
11. Sundari C (2019) Revolusi Industri 4.0 Merupakan Peluang Dan Tantangan Bisnis Bagi Generasi Milenial Di Indonesia. In: Prosiding Seminar Nasional Fakultas Ekonomi Untidar
12. Oyelami LO, Adebiyi SO, Adekunle BS (2020) Electronic payment adoption and consumers' spending growth: empirical evidence from Nigeria. Future Bus J 6(1):1–14
13. Zou X (2021) China's national digital currency: an overview of digital currency electronic payment. Turk J Comput Math Educ (TURCOMAT) 12(11):4279–4285
14. Feshina SS, Konovalova OV, Sinyavsky NG (2019) Industry 4.0—transition to new economic reality. In: Industry 4.0: industrial revolution of the 21st century. Springer, Cham, pp 111–120
15. Prasetyo B, Trisyanti U (2018) Revolusi industri 4.0 dan tantangan perubahan sosial. IPTEK J Proc Series 5:22–27
16. Bulturbayevich MB, Jurayevich MB (2020) The impact of the digital economy on economic growth. Int J Bus Law Educ 1(1):4–7
17. De Vries J (1994) The industrial revolution and the industrious revolution. J Econ Hist 54(2):249–270
18. Syamsuar S, Reflianto R (2019) Pendidikan dan tantangan pembelajaran berbasis teknologi informasi di era revolusi industri 4.0. E-Tech: Jurnal Ilmiah Teknologi Pendidikan 6(2)
19. Fonna N (2019) Pengembangan Revolusi Industri 4.0 dalam Berbagai Bidang. Guepedia

K-Nearest Neighbor and Collaborative Filtering-Based Movie Recommendation System

Srinivasu Badugu and R. Manivannan

Abstract Over the past several years, the recommendation system has become one of the important aspects in our day to day life. Recommendation systems are used in various areas like YouTube, Amazon, etc. Collaborative filtering algorithms have been used in recommendation systems for a long time. They have been effective in resolving a number of issues with commercially available systems. Methods based on a user's neighborhood have showed potential in forecasting user ratings. The aim of this paper is to design and evaluate 'KNN algorithm and Collaborative Filtering algorithm' for producing movie recommendations. The dataset used in this paper is 'Movielens dataset' which is downloaded from Kaggle. The system was implemented using 'Python programming language'. Initially, compared different distance measures and then performed correlation. Based on the correlation value, developed system will recommend similar movies/users. Performances of both the developed algorithms were analyzed in terms of accuracy. Finally the result shows that the accuracy of the recommendations is very good and we will get more accurate movie recommendations based on the combination of "KNN algorithm and collaborative filtering algorithms".

Keywords Recommendation system · KNN · Correlation · Collaborative filtering · Cosine similarity · Matrix factorization · Machine learning

1 Introduction

A recommendation system suggests items to the users based on their preferences. Recommendation systems are used in various areas like YouTube, Amazon, Netflix, Movies, Books, as well as on Social Media platforms like face book and twitter

S. Badugu (✉) · R. Manivannan
Department of Computer Science and Engineering, Stanley College of Engineering and Technology for Women (Autonomous), Hyderabad, India
e-mail: drsrinivasu@stanley.edu.in

R. Manivannan
e-mail: drmanivannan@stanley.edu.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_35

[1]. It predicts the preferences of a user [2] and gives suggestions based on their past preference (i.e., the purchase history) and also other similar users who have the same preferences [3]. Content-based filtering and collaborative-based filtering are two methodologies used in the recommendation system [4]. A content-based strategy is one that is based on information about the products' content rather than other users' alternatives. The collaborative-based filtering (C.F) approach [5] is based on the users' historical preferences, such as the browsing history of things purchased by the user as well as the preferences of other similar users [6]. Sparsity problems [7], Customer wants may change over time, Cold-start difficulty [8], and so on are some of the frequent obstacles and challenges in recommendation systems. The benefits we notice are an increase in customer satisfaction due to timely delivery and appropriate content, as well as the ability to match users to things [9]. In this study, we used the 'K-Nearest Neighbors and Collaborative Filtering' methods to design a 'Movie Recommender System', compare the accuracy of these two algorithms, and develop a good movie suggestion.

2 Literature Review

We know that the recommender system has been a popular study topic, and multiple scholars have contributed various research projects, which are briefly presented in the following section.

Rajpurkar et al. [10] has used content-based filtering and collaborative filtering for book recommendation system. For better performance they used associative model. It gave strong recommendation. Gaikwad et al. [11] developed an online product recommendation system based on user reviews and rating. In this paper authors used a maximum entropy method for sentence tokenization. Liu et al. [12] proposed a personalize news recommendation system for Google News. Authors used click behavior and information filtering approach for recommendation.

Ungar et al. [13] discussed methods of grouping abd clustering of items, then presented collaborative filtering and compared it with k-means clustering algorithm. Wen [14] presented several collaborative filtering algorithms for the recommendation system. They used an item-based k-nearest neighbor (KNN) algorithm approach. Gutta et al. [15] discussed how the recommender system searches the TV programs based on their likes/dislikes through implicit personalization techniques like Bayesian and Decision Tree (DT) methods to produce implicit recommendations.

Wang et al. [16] proposed an efficient clustering approach for user profile partitioning in a recommendation system. They used PCA and received trustworthy recommendations. Subramaniyaswamy et al. [17] presented individualized movie recommendations based on the user rating. They made use of the Movielens dataset. Content-based filtering with genre connection was proposed by Reddy et al. [18]. The experiment was conducted using the Movielens dataset. Gaurav et al. [19] suggested a hybrid recommendation engine that will suggest movies to users based on their preferences. Support vector machine (SVM) models were employed by Eyrun et al.

[20] to predict users' movie interests based on their personal information. Zhang et al. [21] advocated that users be divided into groups based on their profile features. Shinhyun and Shi [22] devised a low-cost and simple movie recommendation system.

3 Proposed Work

Our summarizing system's architecture is outlined in the proposed approach. The architecture of the suggested system, which comprises the following modules: dataset collecting, data preprocessing, data visualization, and then the construction of the ML model, is detailed in depth. The implementation of the movie recommender system is then illustrated and experimental results are presented. Finally the findings of the system evaluation demonstrate that the system recommendations are quite accurate, and that the recommender system, which is based on a combination of KNN and collaborative filtering algorithms, gives correct movie recommendations.

3.1 System Overview

The objective of the proposed work is to provide recommendations based on the users preferences and then by using *k*-nearest neighbor and collaborative filtering algorithm build a movie recommendation system. We will also see how to recommend both similar movies and users as well as how accurate we get the final results by using these two algorithms. The input to the system are the ratings, movie information and output is Top-N recommendation, movies which are similar. Figure 1 shows a sequential flow of our proposed system.

3.2 Movielens Dataset

The dataset which we will work on is movielens dataset [23] (latest) downloaded from Kaggel which contains 100,836 (ratings), for 610 (users), 3683 (tag) applications across 9742 (movies). The dataset is split into four files—movies.csv, ratings.csv, links.csv and tags.csv. Only (movies.csv, ratings.csv) files are used in our paper. The next step after collecting the datasets is preprocessing the data.

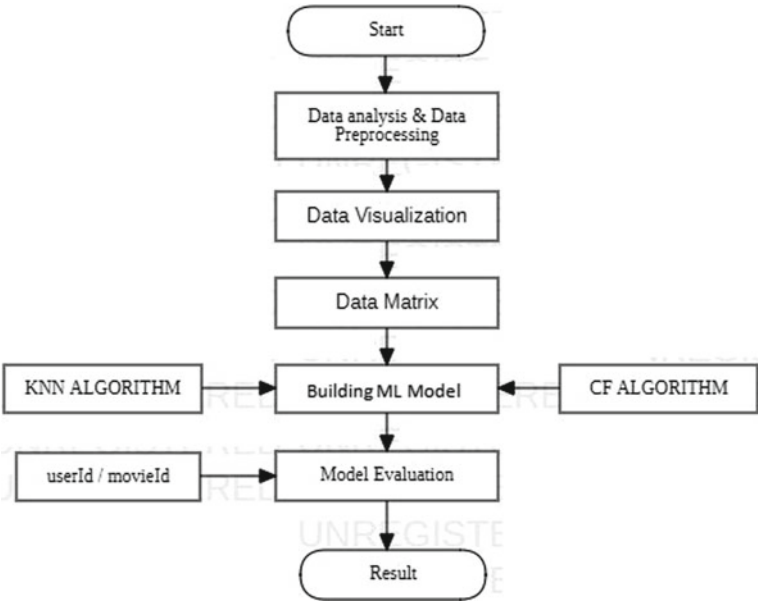


Fig. 1 System architecture

3.3 Data Preprocessing

After importing all the packages then we will load the dataset into pandas data frame. Now we have to read the 'ratings.csv' file and the 'movies.csv' file for that read_csv() method and movie_csv() methods are used. Figure 2 shows data preprocessing steps.

Preprocessing Starts: Merging of rating and movie data frames is done first and then two columns are dropped, i.e., "genres", "timestamp". Next we will search for Not a Number (NaN) values present in our data and also remove any duplicate values from the data. Then average rating and total number of ratings for each movie is taken. For that we apply 'group by function' to (title) and (rating) column. The result we get is the movie title, along with the (rating) and (rating count). Now, basic statistics/information such as total number of movie ratings, number of unique users and unique movies present in our data set are checked. As movieId is common in (ratings.csv) and (movies.csv) files, we Merge (or) Join both the files based on movieId. Lastly, a data frame for data analysis is created.

3.4 Data Visualization

Figure 3 shows a detailed description about the data visualization steps is discussed.

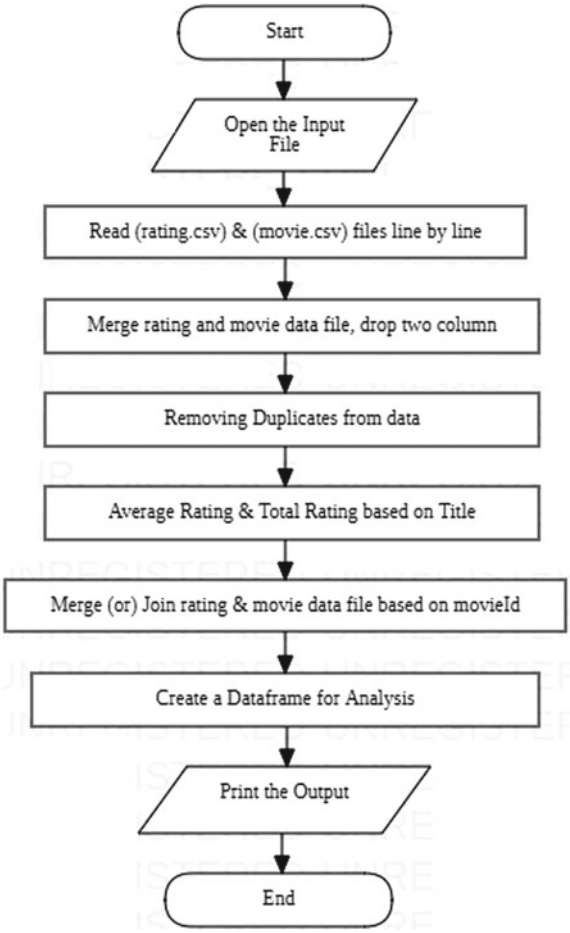


Fig. 2 Flowchart for data preprocessing

STEP 1: Before applying visualization steps, the average rating for each movie is taken first by applying 'group by' function to the title and then for each movie 'mean of the rating' is calculated. Then we see that the average ratings are not sorted. Now let us sort the ratings in the descending order. After this step we face a problem (i.e.) a user can give five stars to one particular movie and that particular movie can be at the top of the list. Normally, good movies get higher ratings and are given by a large number of users.

STEP 2: Next step is plotting on Bar chart and Histogram.

1. **Plot rounded up rating with number of movies**: In order to find the popular ratings for the movies, plot rounded up ratings. A bar graph is plotted by taking the number of movies on (x-axis) against the ratings on (y-axis) they got.

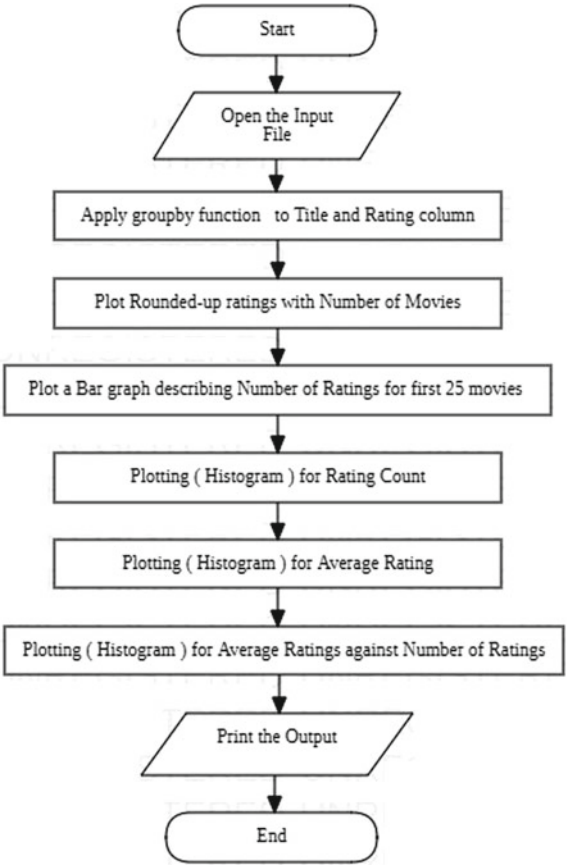


Fig. 3 Flowchart for data visualization

2. **Plot a bar graph describing number of reviews**: A bar graph is plotted for each individual movie in order to describe the total numbers of reviews as well as which movie got the most number of reviews.
3. **Plot Total number of ratings for a movie**: Let us now plot the total number of ratings for a movie. Now we can see some really good movies at the top. Now we know that both the average rating per movie and the no. of ratings per movie are important attributes. Let us create a new data frame that contains these two attributes.
4. **Plot average ratings of each movie**: we create ratings mean count data frame and then to this data frame we will add the average rating of each movie.
5. **Plot average rating against the number of ratings for each movies**: In general the graphs shows that more number of ratings are given by the higher average ratings movies as compared to the movies that have lower average ratings.

3.5 Data Matrix

Matrix factorization was used to combine the two datasets into a single matrix. This matrix is extremely sparse, with many cell values missing. It is either blank or does not apply. To convert sparse to dense, we employed a quantitative filter.

STEP 1: (Create Pivot Tables): Before a matrix is created, we will create two pivot tables. Pivot tables allow us to transform columns into rows as well as rows into columns as well as we can perform data grouping [24]. In the first pivot table we will perform pivot rating into (user features). In order to remove NaN values present in this table, filter data using threshold value. Here, we will check whether the title has got some ratings or not. Then, in the second pivot table we will pivot the rating into (movie features) and again filter the data using threshold value. Here, we will check whether the user Id has given ratings to movies or not. Then, convert this pivot table into a matrix.

STEP 2: (Create the Matrix): The User-Item matrix is a very sparse matrix [25]. Usually many users will not rate all the items/movies which they have purchased. This leads to a sparse matrix where we find more numbers of zero values present in this User-Item matrix. A common (csr_matrix) is created for model KNN [26] using Cosine metric and KNN using Euclidean metric for both (Movie To Movie) and (User To User). Now we are going to use the Pandas Sparse matrix function called Compressed Sparse Row matrix 'csr_matrix' by providing my feature, i.e., (movie_features.values) for (Movie To Movie) and for (user_features_matrix) and (User To User) recommendation. The 'movie_features.values' is nothing but the 'pivot table for movie features'. Once the matrix (i.e., pivot tables for movie and user features) is generated, then it will be converted into an array matrix. Array matrices can be converted by using (scipy.sparse) library. Then import the k-nearest neighbors with two different metric, i.e., Cosine and Euclidean. A full description of the KNN method may be found in Fig. 4.

3.6 Building ML Model

Now we need to configure our KNN model [27]

STEP 1: (Building KNN model): Now Import KNN.

Pseudo code of k-nearest neighbors (KNN):

- Create a 'Pivot tables' using Index, column and values
- Create a 'Matrix'
- Initialize the value of 'k' and Initialize the 'distance measures'
- Calculate the distance between 'target movies' and 'every other movie' in the matrix
- 'Rank' its distance

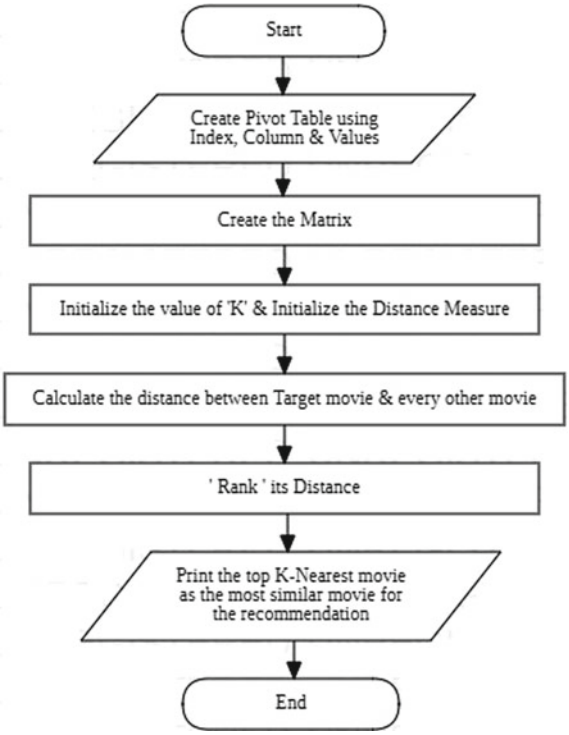


Fig. 4 Flowchart for (KNN) ML model building

- Return the 'Top K-Nearest Neighbor' movie as the most similar movie recommendation.

STEP 2: (Finding Correlations): Now we will find correlation [28] first. We will select one movie and then find other movies which are similar to this particular selected movie by using correlation. Now we see that movies with higher correlation are at the top (i.e., higher the correlation the more correlated that movie is to that selected movie). But a problem may arise, i.e., one or two users have given five stars to one particular movie. That movie can make it to the top of the list as well. For this reason let us sort the values with correlation according to descending order. After sorting them we see that the correlation = 1. Now for (movie to movie) recommendations using correlation, use the corr() method for finding (movies_rating) for a movie with respect to the above corr() data (i.e., the movies_rating), correlate this data with the pivot table and then the movie_features (the UserId in the pivot table) are correlated with movies_ratings. After getting the movies with corr() function, we want good movies at the top of the list of recommendations. So for that reason sort them in ascending = 1. Some of the movie values can be empty (or) could not be correlated to the target movies as well. For that reason simply print the output as 'cannot find correlation'. In this way, follow this same procedure for (user to user)

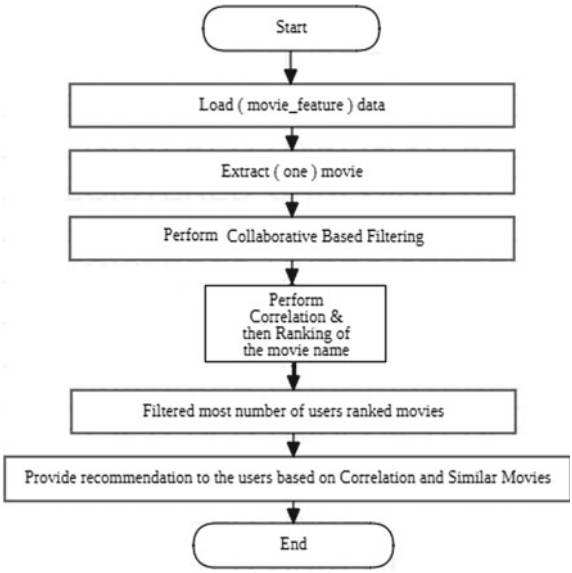


Fig. 5 Flowchart for collaborative filtering (C.F) ML model building

recommendations using correlation. A full description of the collaborative filtering method may be found in Fig. 5.

STEP 3: (Building C.F model):

Pseudo Code of Collaborative Filtering (C.F):

- Load (movie_feature) data
- Extract the (one) movie
- Apply collaborative filtering algorithm
- Perform correlation and then 'Rank' the movie name
- Then provide movie recommendations to the user based on the similar movies using correlation

4 Experimental Results

Histogram for Average Ratings against Number of Ratings: Let us plot average ratings against the number of ratings to show that the area that has received dense points has a higher rating. A histogram of average rating versus number of ratings is shown in Fig. 6.

Movie to Movie Recommendation

We have tested Movie to Movie recommendation and User to User recommendation using KNN with two metrics (i.e., Cosine Similarity and Euclidean distance [29]).

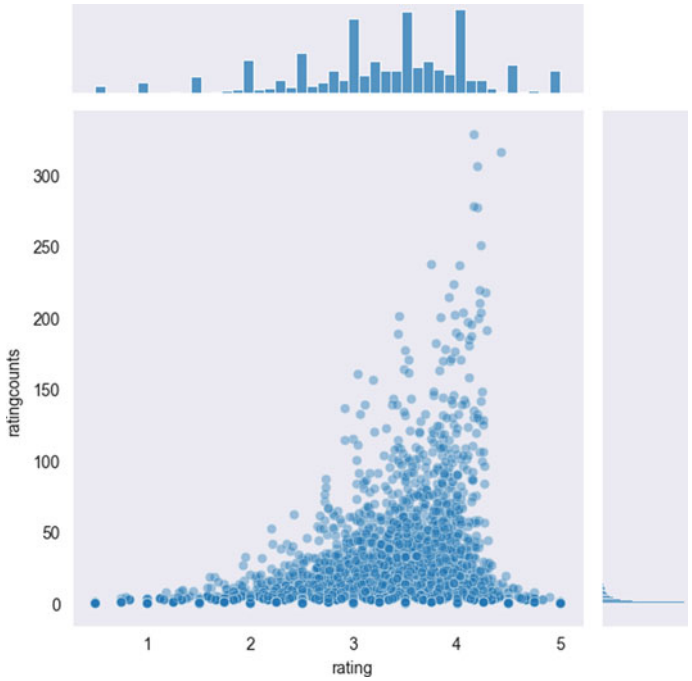


Fig. 6 Histogram for average rating against number of ratings

Here we have taken 9742 movies, 610 users and randomly extracted one index ID (3594) for both metrics. In the Index ID 3594 the title is Grandma's Boy (2006). Using this title we are recommending similar movies based on users ratings. Movie to Movie recommendation using Cosine similarity and Euclidean distance is shown in Tables 1 and 2.

User to User Recommendation

Here we have taken 610 users, 2269 movies and randomly extracted one index ID (501) for both metrics. Using this index ID (501) we are recommending similar

Table 1 Movie to movie recommendation using cosine similarity

S.No.	Title of the input movie	Titles of recommendation movies	Cosine distance
1.	Grandma's Boy (2006)	Without a Paddle (2004)	0.3241
2.		Super High Me (2007)	0.3187
3.		Next Friday (2000)	0.2635
4.		How High (2001)	0.2476
5.		Deuce Bigalow: European Gigolo (2005)	0.2283

Table 2 Movie to movie recommendation using Euclidean distance

S.No.	Title of the input movie	Titles of recommendation movies	Euclidean distance
1	Grandma's Boy (2006)	Deuce Bigalow: European Gigolo (2005)	4.6636
2		Super High Me (2007)	5.4543
3		Without a Paddle (2004)	5.4772
4		Killer Klowns from Outer Space (1988)	5.6568
5		Next Friday (2000)	5.7227

users. User to User recommendation using Cosine similarity and Euclidean distance is shown in Tables 3 and 4.

User to User Recommendation using Correlation

We used the corr() approach to find similar users for user id 502. User to User recommendation using correlation is shown in Tables 5 and correlation values is shown in Fig. 7.

Movie to Movie Recommendation using Correlation

We used the corr() method in order to find similar movies for Angels in the Outfield (1994) movie. Movie to Movie recommendation using correlation is shown in Table 6.

Table 3 User to user recommendation using cosine similarity we used KNN with cosine similarity approach to find recommendations for 502 user

S.No.	Input user ID	Recommended user Id's	Cosine distance
1.	502	64	0.8405
2.		135	0.8390
3.		590	0.8364
4.		469	0.8226
5.		369	0.7453

Table 4 User to user recommendation using Euclidean distance we used KNN with Euclidean similarity approach to find recommendations for 502 user

S.No.	Input user ID	Recommended user Id's	Euclidean distance
1.	502	442	22.4332
2.		175	22.9619
3.		245	23.9060
4.		4369	24.0260
5.		508	24.1402

Table 5 User to user recommendation using correlation

User Id	Correlation
502	1.000000
174	0.678658
542	0.650143
374	0.645201
519	0.536678
631	0.447853

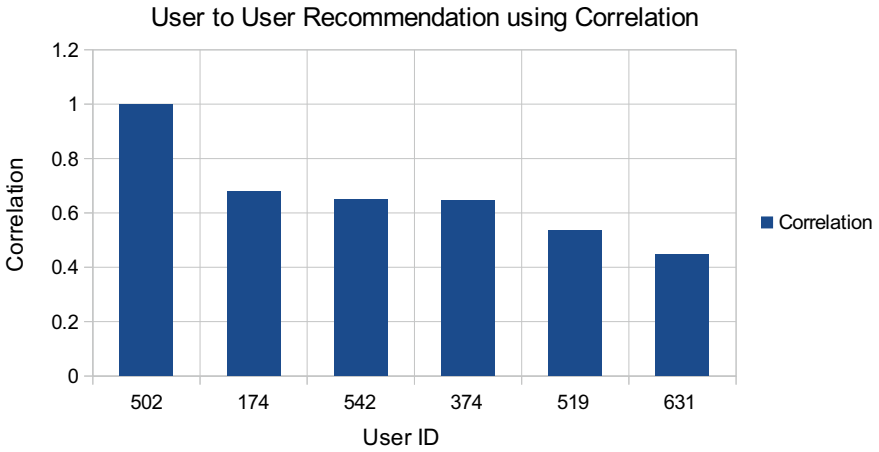


Fig. 7 User to user recommendation using correlation

Table 6 Movie to movie recommendation using correlation

Title	Correlation
Angels in the Outfield (1994)	1.000000
Homeward Bound: The Incredible Journey (1993)	0.473381
D3: The Mighty Ducks (1996)	0.457481
Vegas Vacation (National Lampoon's Las Vegas Vacation	0.419632
American Tail: Fievel Goes West, An (1991)	0.416477
Paulie (1998)	0.410321
Ernest Scared Stupid (1991)	0.409122
Major League (1989)	0.408547
Great Race, The (1965)	0.397790
Love Bug, The (1969)	0.394204

5 Conclusion and Future Work

So far, a lot of work has been done in the area of recommending things like movies. The efficiency and accuracy can be improved over time with more and more research. User rating and preference of the user have been considered while building the system. Recommender systems suggest like or similar items and ideas to a user's specific way of thinking. Machine learning is an important part of systems for these tasks. It helps users to find items they want to buy according to their preferences. In this paper, we have worked with k-nearest neighbors and collaborative filtering approach to get movie recommendations as good as possible. After applying k-nearest neighbors with two different distance measures, i.e., (Cosine and Euclidean Distance) perform correlation() function and based on that correlation values recommend similar movies/users (Collaborative Filtering). The combination of KNN algorithm and collaborative filtering algorithms provides more accurate movie recommendations.

We intend to develop this approach in the future by using feature engineering to obtain more accurate results and improve recommendation performance. We will also work on improving the existing approaches and algorithms so that movie recommendation systems can make better predictions and recommendations.

References

1. Gupta P, Goel A, Lin J, Sharma A, Wang D, Zadeh RB (2013) WTF: The who-to-follow system at Twitter. Publication: WWW'13, In: Proceedings of the 22nd international conference on World Wide Web
2. Le J (2018) The 4 recommendation engines that can predict your movie tastes from towards data science
3. Isinkaye FO, Folajimi YO, Ojokoh BA (2015) Recommendation systems: principles, methods and evaluation. Egypt Inf J 16:261–273
4. Felfernig A, Jeran M, Ninaus G, Reinfrank F, Reiterer S, Stettinger M (2014) Basic approaches in recommendation system. Publisher, Springer
5. Lee J, Hwang W-S, Parc J, Lee Y, Kim S-W (2016) l-injection: toward effective collaborative filtering using uninteresting items. IEEE
6. Ajitsaria A, "Build a recommendation engine with collaborative filtering" from Real Python
7. Kumar B, Sharma N (2016) Approaches, issues and challenges in recommender systems: a systematic review. Indian J Sci Technol
8. Khusro S, Ali Z, Ullah I, Recommender systems: issues, challenges, and research opportunities. Springer, Singapore
9. Zisopoulos C, Karagiannidis S, Demirtsoglou G, Antaris S, Content-based recommendation systems. ResearchGate
10. Rajpurkar S, Malhotra P, Bhatt D (2015) Book recommendation system. IJIRST—Int J Innov Res Sci Technol 1(11)
11. Gaikwad R, Darwatkar R, Jangawali N (2016) Online recommendation system based on reviews and ratings. Int Res J Eng Technol (IRJET) 03(03)
12. Liu J, Dolan P, Pedersen ER (2010) Personalized news recommendation based on click behavior. Published: IUI '10: Proceedings of the 15th international conference on intelligent user interfaces

13. Ungar LH, Foster DP, Andre E (1998) Clustering methods for collaborative filtering
14. Wen Z (2008) Recommendation system based on collaborative filtering
15. Gutta S, Kurapati K, Lee KP, David Schaffer J, Zimmerman J (2000) TV content recommender system. Published: AAAI/IAAI, Association for the advancement of artificial intelligence/innovation applications, of artificial intelligence conference
16. Wang Z, Xue Y, Feng N, Wang Z (2014) An improved collaborative movie recommendation system using computational intelligence. J Vis Lang Comput 25(6):667–675
17. Subramaniyaswamy V, Logesh R, Chandrashekhar M, Challa A, Vijayakumar V (2017) A personalised movie recommendation system based on collaborative filtering. Int J High Perform Comput Netw 10(1–2):54–63
18. Reddy SRS, Nalluri S, Kunisetti S, Ashok S, Venkatesh B (2019) Content-based movie recommendation system using genre correlation. In: Smart intelligent computing and applications. Springer, Singapore, pp 391–397
19. Arora G, Kumar A, Devre GS, Ghumare A (2014) Movie recommendation system based on users' similarity. Int J Comput Scie Mobile Comput 3(4):765–770
20. Eyjolfsdottir EA, Tilak G, Li N (2010) Moviegen: a movie recommendation system. UC Santa Barbara: Technical report
21. Zhang J, Wang Y, Yuan Z, Jin Q (2019) Personalized real-time movie recommendation system: practical prototype and evaluation. Tsinghua Sci Technol 25(2):180–191
22. Ahn S, Shi C-K (2009) Exploring movie recommendation system using cultural metadata. In: Transactions on edutainment II. Springer, Berlin, pp 119–134
23. Maxwell HF, Konstan JA (2015) The movielens datasets: history and context. Acm Trans Interact Intell Syst (tiis) 5(4):1–19
24. Shih-Ping Chang K, Myers BA (2016) Using and exploring hierarchical data in spreadsheets. Conference: the 2016 CHI conference
25. Zhang L, Luo T, Zhang F, Wu Y (2018) A recommendation model based on deep neural network. This work was supported by the National Education Information Technology Research under Grant 171140001
26. Zhang Z (2016) Introduction to machine learning: k-nearest neighbors. Annals of translational medicine 4, no. 11
27. Marlin B (2004) Collaborative filtering: a machine learning perspective. University of Toronto, Toronto
28. Grewal BS, Grewal JS (1996) Higher engineering mathematics. 2002, Khanna Publishers, New Delhi
29. Bali NP, Paul U (2008) A textbook of higher engineering mathematics. Laxmi Publications

Efficient Data Flow Graph Modeling Using Free Poisson Law for Fault-Tolerant Routing in Internet of Things

P. B. Pankajavalli and G. S. Karthick

Abstract Internet of Things (IoT) is based on characteristics of the Wireless Sensor Network (WSN) in which nodes are dispersed for segregating data for different applications. In IoT, the sensor nodes typically possess heterogeneous properties. Among, few nodes have higher energy and even data aggregation functionality. Generally, in WSNs, efficient cluster-based routing is employed for data transmission. When Cluster Head (CH) fails, data sensed by the sensor nodes cannot be transmitted by the faulty CH. Consequently, the gateway or sink cannot recognize the data in the IoT sufficiently. Hence, processing information in this field was severely affected. This research contribution focuses on paired fault-tolerant cluster routing of the disjoint routes in a data flow graph and introduces a new approach called Free Poisson Law for solving this problem. The proposed Efficient Data Flow Graph Modeling using Free Poisson Law (EDFGM-FPL) algorithm has the aim of reducing latency, energy consumption as well as end-to-end delay, dissipated energy, functional complexity thereby achieving improved packet delivery ratio, throughput, and fault detection rate.

Keywords Internet of Things · Wireless sensor networks · Fault-tolerant · Cluster head · Path graph flow · Free poisson law

1 Introduction

Wireless Sensor Network (WSN) consists of finite sensor devices which are dispersed geographically usually in an indoor or outdoor predefined environment. The objective of WSN is to collect environmental data and the location of the nodes may be previously known or unknown a priori. Nodes in the network communicate with all devices

P. B. Pankajavalli (✉) · G. S. Karthick
Department of Computer Science, Bharathiar University, Tiruchirappalli, India
e-mail: pankajavalli@buc.edu.in

G. S. Karthick
e-mail: karthickgs@outlook.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_36

and based on the application such communication describes a mesh or star topology. However, this differs from case to case. Depending on the roles of the node, it may be ad hoc or strategy based on the available resources in the network [1]. In several fields, WSN technology is broadly used efficiently to enhance the performance of the network. Various sensors are utilized in the environmental fields as they can be managed and the configuration setup is easier. Furthermore, the function of Sensor Nodes (SNs) is independent and the infrastructure of the network is constructed using ad hoc. In this infrastructure, nodes have no firm topology and join with the most appropriate neighbor to transmit data depending on certain factors. The data sensed by sensor nodes are forwarded to the Base Station (BS) using Cluster Heads (CHs) and gateways. These CHs then aggregate the data packets received before sending them to BS. CH effectively creates a single or multi-hop routing path toward BS which acts as the central point in the network for the transmission of data. Also, CHs perform a collection of received data and transmit that data through the forward mechanism. Usually, BS are deployed in a centralized manner and accessed by end-user using the Internet or different web-based application for the collection of data. For data transmission sensor nodes are deployed statically or dynamically. The static sensor defines the sensor network constructed in a non-adaptive, constructed network with an appropriate routing mechanism. Conversely, routing tables of dynamic ones are often updated when changes occur in the topology [2]. Static routing is more secure than dynamic routing; but still, the solutions of static algorithms do not suit large areas and provide no scalability. Recently, IoT technologies are mostly merged with other fields to enhance communication thereby providing enhanced throughput, load distribution, and resource utilization. IoT environment has enormous smart devices along with several constraints such as processing ability, storage, limited power life as well as radio range. Additionally, WSN technologies form the base for IoT-based systems (shown in Fig. 1) and help to monitor and forward the conditions for the physical environment [3]. With the numerous worldwide IoT networks, secure routing protocols have increasingly become significant. Hence, designing and implementing a secure routing protocol to recognize and address the objectives of a security protocol are important.

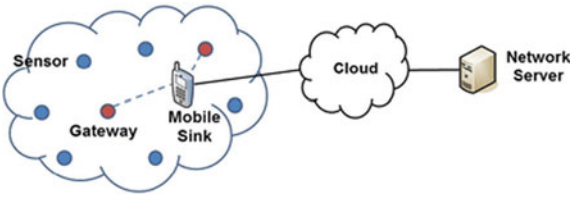


Fig. 1 Overview of WSN model for IoT applications

2 Related Works

While designing the wireless sensor networks the concentration should be made on sensor nodes to improve its functional time. While designing processes, many failures may occur among sensor nodes due to volatility during the broadcast, environmental impact, problems in hardware, radio nosiness, interruption of other nodes, and battery exhaustion. Moreover, there is a failure in the base station due to hardware disfunction, software problems, and planned attacks. Apart from sensor nodes and base stations, the failure may occur in Relay Node (RN) or Super Node. Therefore, it is significant to design the wireless sensor network which must be able to tolerate the fault which is done by building a fault tolerance system, which contains less likelihood of false alarms and more likelihood of fault credentials.

Tong et al. [4] suggested the Static Backup and Dynamic Timing Monitoring (SBDTM) fault tolerance method, which was based on the Markov model with reliability calculation. By minimizing the number of cluster heads, the reliability could be increased. The authors elaborated the structure of data and fault-tolerant conditions to improve the latency and diminish the consumption of energy.

Hu et al. [5] suggested a method named Regular Hexagonal-based Clustering Scheme (RHCS) with Scale-Free Topology Evolution Mechanism (SFTEM) for wireless sensor networks, which further improves the network lifetime. Since it had been using the hexagonal structure for the clustering process, the coverage problem was reduced. This method associated the properties of hexagonal clusters to construct the healthy network, which achieved the interaction among reliable clustering structure and topology development. As a result, this method tolerated the complete faults comprising random failure and energy failure.

Lin et al. [6] suggested a bipartite-flow graph (BFG) can model to find the likelihoods of faults between the fault and un-fault cluster head. To overcome the Main Curvature Fault (MCF) problem in the graph, this author concentrated on achieving optimal points in broadcast sets with less cost. As a result, it was found that more than two failures in cluster heads could tolerate dense load.

Michaelides et al. [7] presented the Event Tracking Algorithm (ETA) to tolerate the badly behaving nodes. The author used this algorithm mainly for reducing complexity, improvise the distribution and tracking purpose. This algorithm used binary methods among sensors to recognize, focus and follow the alternate paths.

Rong Duh et al. [8] suggested Distributed Fault-Tolerant Algorithm (DFTA) identifies a particular area in WSNs. This algorithm was fully based on neglecting the disambiguation issue, classification of defective and non-defective nodes in the network. Here, the cluster heads send periodic reports to the base station and then these base stations identify the abnormality in that area by creating a false alarm. As a result, it was found that the accuracy of event detection accuracy was 99%, fault detection accuracy was 94% and false alarm rate was 0% in unvarying dissemination conditions. The event detection accuracy was 88%, fault detection accuracy was 95% and false alarm rate was 2.3% while sensor fault probability was 0.3.

Rong-Rong et al. [9] introduced a framework namely, Adaptively Fault-tolerant Topology Control (AFTC) to allocate resources in the network. The election of backbone nodes was dependent on the fault-tolerant rate of sensor nodes. Each backbone node was guided with memorized backup nodes to enhance the energy efficiency, storage capacity, and capability of the network. The performance analysis part showed that this framework was suitable to achieve better reliability and efficiency in the network.

Bari et al. [10] suggested a novel Integer Linear Program (ILP) tolerates the fault by accurately considering the less number of relay nodes and sensor nodes in the network by effective the load balancing condition. The load balancing contains both the upper and lower level to tolerate fault by assigning the authentic values as inputs. The presented method resulted in overall performance that guaranteed with better lifetime and minimized energy consumption among sensor nodes.

The backgrounds of fault-tolerant routing protocols in WSN and IoT are analyzed [11, 12]. In general, efficient cluster-based routing is employed for data transmission in WSN. When CH fails, data sensed by the sensor nodes cannot be transmitted by this faulty CH. Consequently, the gateway or sink has not sensed the data in the IoT sufficiently. Hence, processing information in this field was deeply affected. This contribution focuses on paired fault-tolerant cluster routing in a routing graph and introduces a new approach to solving this problem.

3 Proposed Method

Static sink used in IoT applications leads to clusters close to the sinks dying out faster as CH forms a fixed route for data routing in such clusters, which results in network failure. To mitigate the fault oriented constraints in CH, a specific Data Flow Graph Modeling with Free Poisson Law is used here which follows the cluster head selection as well as routes data to mobile sink or the static CH by overcoming the fault in the links. Figure 2 depicts the proposed method architecture for tolerating the faults in CH.

3.1 Network Modeling and Routing Procedure

Network frameworks belong to sensor nodes that are dispersed uniformly over a network. These groupings of sensors are grouped according to CH. The Base Station is located on every sensor zone. All CH and BS do not adjust over the entire cycle. SN's are uniform, however, and each node that is allotted a well-defined ID modification its belongings depends upon the transmission range. In WSN, the following properties apply to network models: (1) SNs are forever static and their position is unchanged. Whereas these nodes collect data from the surrounding area that is then transmitted to BS via wireless communication. (2) All SN abstracts ID the redundant

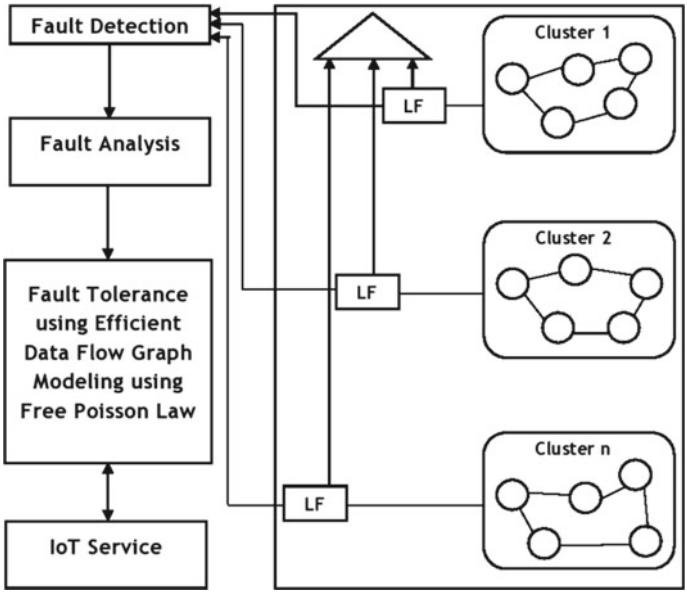


Fig. 2 Proposed method architecture

information is fused. Nevertheless, they cause failure If those sensor nodes have zero energy. (3) Hence, the nodes are willing to continuously adjust their power level as the level of transmitting power decreases.

At BS the routing process and the average data are calculated quantity transmitted through each relay node is assumed to be fixed and it is odd for BS to do so. The optimized route was set for message broadcasting by BS. Each head of the cluster follows the procedure. For every round, this procedure exploits the data transmission in WSN, through the feasible path transmits data packet from BS to every clusters are continual. The routing plan is dynamically calculated in light of the existing energy level to each CH. To this end, CHs typically report their residual energy to BS regularly to coordinate their status. Then BS decides the routing schedule based on the updated information.

3.2 Efficient Data Flow Graph Modeling Using Free Poisson Law (EDFGMA-FPL) Algorithm for Fault-Tolerant Routing

The occurrence of a fault in the network is identified using the data flow graph path construction along with Free Poisson Law. Initially, the cluster head is identified and formed. Then, tolerate the fault by data flow graph path whereas consider there exists

an $A_n - 1(m)$ which contains s_i or t_i (s_i denotes source node, t_i denotes destination node), and the $A_n - 1(m)$ represents a candidate sub data flow graph path for s_i or t_i, $A_n - 1(m)$ is assigned to the target data flow graph path $B_n - 1(li)$. If s_i and t_i are included in distinct $A_n - 1(p)$ and $A_n - 1(q)$, respectively, and both of these satisfy the constraints of the candidate subgraphs for s_i and t_i, either of them is assigned to $A_n - 1(l_i)$. A target sub data flow graph path can be assigned in $O(n)$ time for every pair of source and destination nodes.

For every source node pair $s_i = (s_i1, s_i2, \ldots, s_in)$ and the destination nodes t_i is not assigned any target sub data flow graph path. A route is built either from the source or destination nodes to a sub data flow graph path of the candidate. Here, 'i' assumed that a candidate sub data flow graph path is found for s_i. Then, a target data flow graph path is assigned and a path is constructed using the following procedure.

1. If the $A_n - 1(s_i\ 1)$ denotes a candidate sub data flow graph path for s_i, $A_n - 1(s_i\ 1)$ is assigned to the target sub data flow graph path $A_n - 1(l_i)$, construct a path $s_i\ s(n)_i$ of length 1 to the sub data flow graph path, and let $s(n)_i = s_i$.
2. If the $A_n - 1(s_i\ p)$ $(1 \leq p < p)$
3. For k pairs of nodes s_i and $t_i(1 \leq i \leq k)$, from Steps 3 to 5, paths $s_i = s_i\ (\in A_n - 1(li))$ and $t_i = t_i(\in A_n - 1(l_i))$ is constructed where $A_n - 1(l_i)$ represents the target sub data flow graph path for s_i and t_i. $A_n - 1(l_i)$ does not include any node on s_j or t_j where $t(j = i)$, and hold at most $n - 2k + 1$ failure nodes. Hence, a path s_i, t_i of length at most $2n + 2$ is constructed in $O(n^2)$ time.

The data flow graph path is used to understand the other two features of the planned solution. The flux-chart ensures that sensed data from the built-in CH can also be sent to faultless CHs (Fig. 3).

Pseudocode: Efficient Data Flow Graph Modeling using Free Poisson Law (EDFGMA-FPL) Algorithm

Input: 'n' Number of Nodes, 'p' Data Flow Graph Path, 'k' Pair of Nodes
Output: Fault-Tolerant Path in the Network
$A_n - 1(m) \rightarrow s_i$ or t_i and $A_n - 1(m)$
$A_{n-1}(m) \rightarrow Bn - 1(l_i)$
for
$A_n - 1(p), A_n - 1(q) \neq s_i$ and t_i
Sub graph time = set $O(n)$ $s_i = (s_i1, s_i2,\ldots, s_in)$
$s_i = t_i$
if
$A_n - 1(s_i1) =$ si then subpath is constructed
Else
Unconstruct the path
if $A_n - 1(s_ip)$ $(1 \leq p < p)$
for k pairs of nodes s_i, t_i $(1 \leq i \leq k)$
constructed paths $s_i(\in A_n - 1(li))$ and $t_i(\in A_n - 1(l_i))$
end if
end for

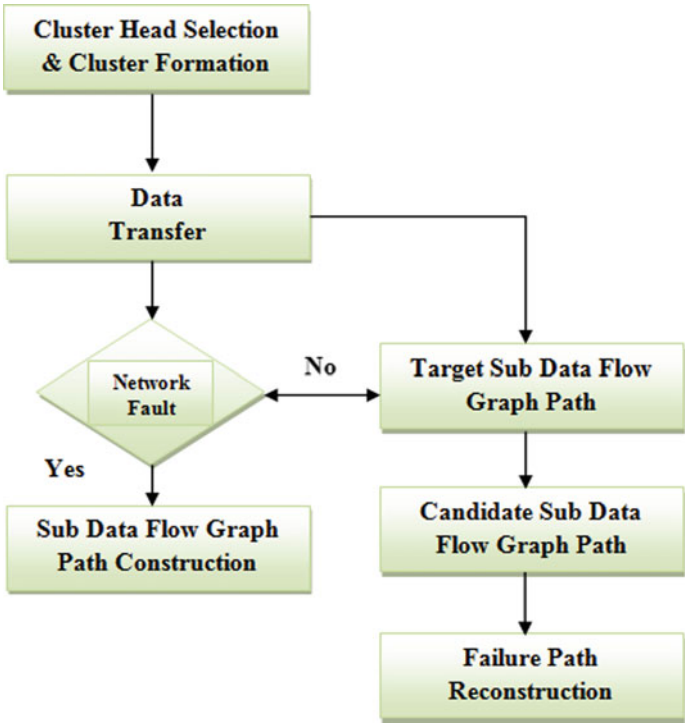


Fig. 3 Flowchart of efficient data flow graph modeling using free poisson law (EDFGMA-FPL) algorithm

4 Experimental Results and Performance Evaluation

In this research work, the network is composed of different sensors with a versatile population of between 100 and 500 sensor nodes that has been dispersed randomly in a mesh network around an area of 1000 by 1000 m^2. The sensors are carried out which has supplied the packet with qualified energy and the nodes parameter was designed to dissipate resources. The Network Simulator-2 (NS2) is utilized for the simulation of the proposed and existing algorithms.

The experimental analysis was conducted and the achieved results were evaluated with the results of contemporary fault-tolerant techniques discovered in the literature, including the BFG, DFTA, and SBDTM algorithms.

The proposed EDFGMA-FPL algorithm is compared and evaluated with the recent algorithms using the below-mentioned performance metrics [12–14].

(i) **Packet Delivery Ratio** refers to the average amount of packets received successfully to the amount of packets originally sent.

$$\text{Packet Delivery Ratio} = \frac{\sum p(rx)}{\sum p(tx)}$$

where, $p(tx)$ is the amount of packets transmitted and $p(rx)$ denotes the amount of packets received.

(ii) **Average Delay** defines the amount of time it takes for a piece of data to move across the network from one communication terminal to another.

(iii) **Dissipated Energy** refers to the amount of energy wasted by all the nodes in the network.

$$E_{\text{Dissipated}} = E_{\text{Initial}} - E_{\text{Consumed}}$$

(iv) **Latency** denotes the time interval between triggering and beginning data transmission. Low latency will provide higher network quality networks and vice-versa.

(v) **Throughput** is the rate of data flow through a channel used for communication (bits or packets delivered successfully over a channel in the network).

$$\text{Throughput} = \sum \frac{S(P) * A(P)}{t}$$

where $S(P)$ is the number of successful packets, $A(P)$ is the average packet size and t is the amount of time interval taken for delivering the data.

Table 1 and Fig. 4 specify the packet delivery ratio and it is inferred that the clustering and data aggregation involved in EDFGMA-FPL caused an impact on packet delivery success. It is ascertained that with few source nodes and reduced data traffic, EDFGMA-FPL executed better because its aggregate nodes can communicate with the sink.

Table 2 and Fig. 5 prove the average delivery delay of the proposed EDFGMA-FPL algorithm is minimum. The average delivery delay has increased precipitately, as the percentage of source nodes for BFG, DFTA, and SBDTM algorithms increases.

From Table 3 and Fig. 6, it is observed that through minimizing data flow, the average dissipated energy achieved for the proposed EDFGMA-FPL algorithm is satisfactory compared to the existing algorithms.

Table 4 and Fig. 7 proves the latency of the proposed EDFGMA-FPL algorithm

Table 1 Packet delivery ratio versus number of nodes

No. of nodes	BFG	DFTA	SBDTM	EDFGMA-FPL
100	71.43	74.17	84.75	89.43
200	73.74	75.47	86.71	91.77
300	74.19	80.27	86.84	92.64
400	73.49	73.61	88.47	94.31
500	77.52	84.49	89.41	96.48

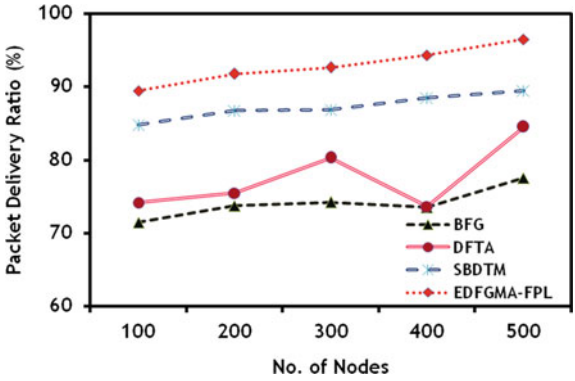


Fig. 4 Packet delivery ratio versus number of nodes

Table 2 Average delay versus number of nodes

No. of nodes	BFG	DFTA	SBDTM	EDFGMA-FPL
100	95	88	73	62
200	157	139	126	98
300	201	185	172	128
400	249	225	217	167
500	347	324	289	212

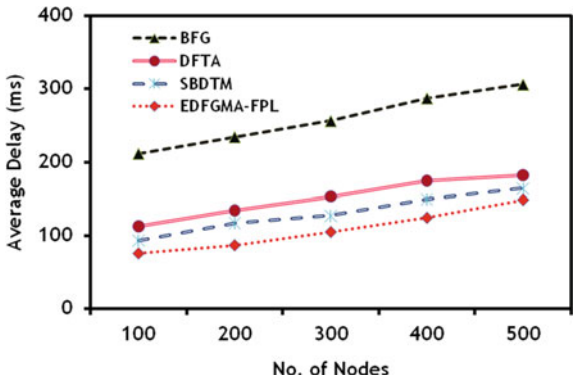


Fig. 5 Average delay versus number of nodes

Table 3 Dissipated energy versus number of nodes

No. of nodes	BFG	DFTA	SBDTM	EDFGMA-FPL
100	211	112	93	76
200	234	134	117	87
300	256	153	127	105
400	287	174	149	124
500	306	182	164	148

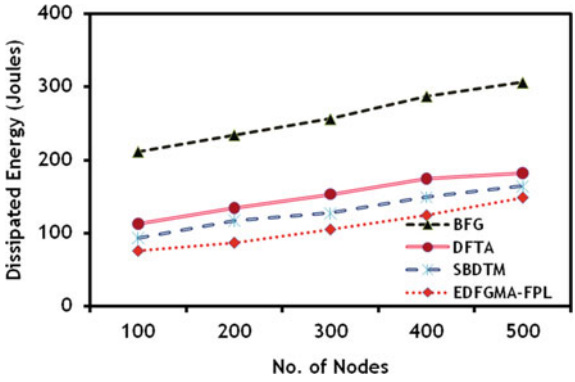


Fig. 6 Dissipated energy versus number of nodes

Table 4 Latency versus number of nodes

No. of nodes	BFG	DFTA	SBDTM	EDFGMA-FPL
100	76	61	52	29
200	74	67	56	28
300	81	63	53	31
400	82	69	57	35
500	84	64	58	36

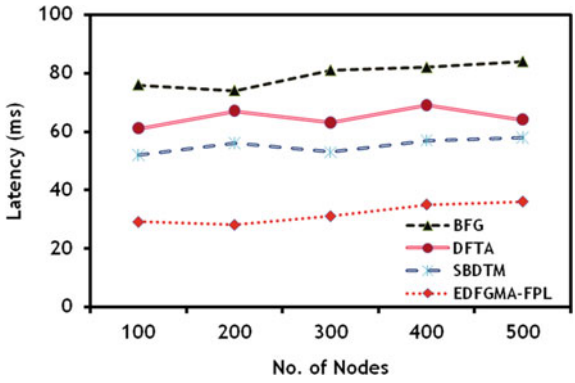


Fig. 7 Latency versus number of nodes

is minimum when compared with the existing algorithms. Also, the latency of BFG algorithm is too high among the DFTA, SBDTM, and EDFGMA-FPL algorithms.

Table 5 and Fig. 8 prove that the throughput of the proposed EDFGMA-FPL algorithm is maximum when compared with the existing algorithms. Also, the throughput achieved by the BFG algorithm is comparatively less than the DFTA, SBDTM, and EDFGMA-FPL algorithms.

From Table 6 and Fig. 9, it can be observed that the proposed EDFGMA-FPL algorithm tolerates the CH faults effectively by constructing a data flow graph model

Table 5 Throughput versus number of nodes

No. of nodes	BFG	DFTA	SBDTM	EDFGMA-FPL
100	62	81	82	89
200	74	75	76	92
300	65	76	78	95
400	69	72	79	90
500	70	79	81	94

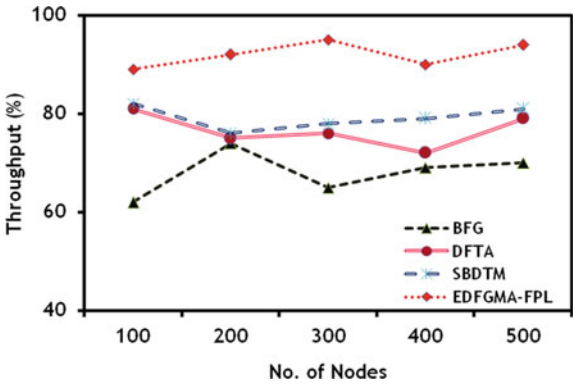


Fig. 8 Throughput versus number of nodes

Table 6 Fault detection rate versus number of nodes

No. of nodes	BFG	DFTA	SBDTM	EDFGMA-FPL
100	58.6	61.23	69.78	78.22
200	55.7	54.23	71.26	79.84
300	45.9	69.21	68.54	81.42
400	42.6	60.58	69.89	83.11
500	42.1	59.4	64.69	84.78

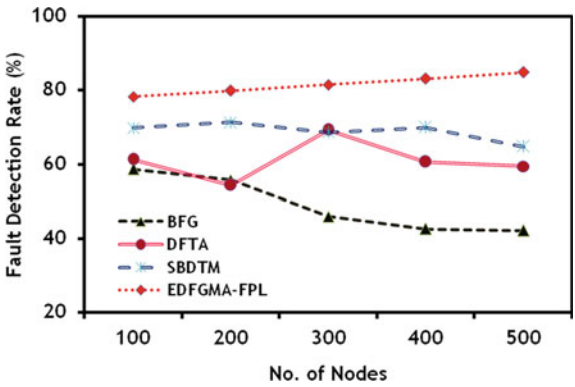


Fig. 9 Fault detection rate versus number of nodes

using free poisson law and it achieves better results when compared with the existing algorithms.

5 Conclusion

This research work has been focused on resolving the problems of the IoT and WSN fault-tolerant cluster routing for the heterogeneous property of sensor nodes. The proposed EDFGMA-FPL algorithm is developed to improve the routing by reducing the fault in the network. The cluster and cluster head formation is discussed with the energy dissipation model. The assumption is made to construct the occurrence of a fault in the network and then it is overcome by the proposed EDFGMA-FPL algorithm. The proposed EDFGMA-FPL algorithm tolerates the CH faults effectively by constructing a data flow graph model using free poisson law and it achieves 81.45% of fault detection rate. In addition, it has been found that the proposed algorithm can able to build faultless disjuncture paths between the k node. This research work focuses on resolving fault tolerance issues and data storage is another foreseen issue in IoT and WSN. Therefore, in the future the storage of WSN in IoT can be enhanced, where it could improve the energy efficiency and data recoverability through an optimized data storage mechanism.

Acknowledgements The work of P. B. Pankajavalli and G. S. Karthick were supported by the Department of Science and Technology—Interdisciplinary Cyber-Physical Systems (DST-ICPS), New Delhi, India, under Grant DST/ICPS/CPS-Individual/2018/193.

References

1. Singh M, Lal N, Sethi M, Poonia S (2010) A tree based routing protocol for mobile sensor networks (MSNs). Int J Comput Sci Eng 1(1):55–60
2. Tripathi A, Gupta HP, Dutta T, Mishra R, Shukla KK (2018) Coverage and connectivity in WSNs: aA survey, research issues and challenges. IEEE Access 6:26971–26992
3. Shahzad MK, Cho TH (2017) An energy-aware routing and filtering node (ERF) selection in CCEF to extend network lifetime in WSN. IETE J Res 63:368–380
4. Tong Y, Tian L, Lin L, Wang Z (2020) Fault tolerance mechanism combining static backup and dynamic timing monitoring for cluster heads. IEEE Access 8:43277–43288
5. Hu S, Li G (2018) Fault-tolerant clustering topology evolution mechanism of wireless sensor networks. IEEE Access 6:28085–28096
6. Lin J-W, Chelliah PR, Hsu M-C, Hou J-X (2019) Efficient fault-tolerant routing in IoT wireless sensor networks based on bipartite-flow graph modeling. IEEE Access 7:14022–14034
7. Michaelides M, Christos Laoudias P, Panayiotou CG (2014) Fault tolerant localization and tracking of multiple sources in WSNs using binary data. IEEE Trans Mob Comput 13(6):1213–1227
8. Rong Duh D, Pei Li S, Cheng V (2013) Distributed fault-tolerant event region detection of wireless sensor networks. Int J Distrib Sens Netw 1–8

9. Rong-rong Y, Bin L, Ya-qian L, Xiao-chen H (2012) Adaptively fault tolerant topology control algorithm for wireless sensor networks. J China Univ Posts Telecommun 19(2):13–18
10. Bari A, Jaekel A, Jiang J, Xu Y (2012) Design of fault tolerant wireless sensor networks satisfying survivability and lifetime requirements. Comput Commun 35(3):320–333
11. Andi HK (2021) Analysis of serverless computing techniques in cloud software framework. J IoT Soc Mob Anal Cloud 3(3):221–234
12. Harkare A, Potdar V, Mishra A, Kekre A, Harkare H (2021) Methodology for implementation of building management system using IoT. In: Evolutionary computing and mobile sustainable networks. Springer, Singapore, pp 939–948
13. Liu X (2015) Atypical hierarchical routing Protocols for wireless sensor networks: a review. IEEE Sens J 15:5372–5383
14. Elshrkawey M, Elsherif SM, Wahed ME (2018) An enhancement approach for reducing the energy consumption in wireless sensor networks. J King Saud Univ Comput Inf Sci 30(2):259–267

Extractive Text Summarization on Large-Scale Dataset Using K-Means Clustering and Word Embedding

Ti-Hon Nguyen and Thanh-Nghi Do

Abstract Automatic text summarization tasks play an important role in natural language processing. In this work, we introduce the single-document extractive summarization model based on clustering and word embedding. In the model, we use K-Means clustering to create the clusters on the large-scale dataset by using word embedding as the feature vector, then use these clusters to extract the most relevant sentences on the document to summarize. At first, we collected the articles on the Vietnamese online newspapers, cleaned them and built up the dataset with a total of 1,101,101 articles. After that, we applied our summarization model for the experimentation. The average time cost for summarizing one document in the test set is 6.22 ms, and the best F-Score of this model based on ROUGE-1, ROUGE-2, and ROUGE-L are 51.40, 16.15, and 29.18%.

Keywords Text summarization · K-Means · Clustering · Large-scale dataset · Word embedding

1 Introduction

Text summarization was first introduced by Luhn [20] in 1958; this is the task to create a smaller version of a document or a set of documents and retain the main content of the original documents. Based on the output [13] of the summary model, there are two kinds of summarization methods: the extractive method and the abstract method. The extractive methods try to select the most relevant sentences in the original text to produce the summary; on the other hand, the abstractive methods create the summary by rewriting the original text. In addition, based on the input [13] of the summary model there are single-document and multi-document summary models.

T.-H. Nguyen (✉) · T.-N. Do
Can-Tho University, Can-Tho, Vietnam
e-mail: nthon@ctu.edu.vn

T.-N. Do
e-mail: dtnghi@ctu.edu.vn

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_37

Nowadays, text summarization attracts a lot of research in natural language processing [5, 8, 9]. There are some English large-scale datasets for evaluating text summary models, such as GigaWord [10, 30], CNN/Daily Mail [32]. However, there is only a little research for Vietnamese text summarization [17, 24, 26]. In addition, most text summarization research attends to the summary's quality based on the ROUGE score metrics [18], but it is hard to find the research focus on the time cost to create a summary while keeping a high ROUGE score.

Therefore, in this work, we built the Vietnamese large-scale dataset for text summarization research and focused on summarization on a single-document. In addition, we use word embedding to shorten the vector feature that helps our summary model produce the summary more quickly.

In the next section, we will discuss the current summary methods. The model section will introduce our summary model, the K-Means clustering model, word embedding, distance metrics, the evaluating process and the using of these topics in this work. Then, the experiment section will present the dataset and the evaluated results. Finally, the conclusion section will show the contribution and advantages of our summary model and the future works.

2 Related Work

In recent years, extractive text summarization and abstractive text summarization methods are exalted in text summary research. The state-of-the-art text summarization models include **LEAD-3** [22], **Pointer-Generator + Coverage** [32], **Latent** [15], **BanditSum** [7], **PNBERT** [35], **BertSumExt** [19], and **MatchSum** [34].

In spite of the large amount of text summarization research [5, 8, 9, 23], there is only a little research using clustering models for extractive text summarization. Such as Agrawal [1], Zhang [33] and Aliguliyev [4] use the clustering algorithm to cluster the sentences of a document and choose top good sentence candidates to combine into the summary for the single-document. Besides that, Deshpande [6], Akter [2], Rossiello [29], and Radev [27] use clustering based for summarizing multi-document. The main idea of this kind of research is applying the clustering algorithm on the documents in the corpus. Then, each sentence in the corpus will be ranked based on the trained clusters and the top sentences will be used to create the final summary.

In addition, there are some research use word embedding for text summarization such as Mohd [21], Rani [28], Jain [14], Jiang [16], and Al-Sabahi [3].

In this work, we construct the single-document summary model; at first, we use a clustering algorithm to cluster the documents in the training set. Then, in the summary step, the model will rank all sentences in that document based on the trained clusters and top sentences will be extracted as the summary.

3 Model

3.1 Summary Model

In this work, we introduce the single-document extractive summarization model. The model has two main steps: the training and the summarizing step. The process of our summary model is presented in Fig. 1.

In the training step, at first, we use the training set to train the Glove model. Next, we use the trained Glove model as word presentation model to create the training set matrix. After that, we use the K-Means clustering method to cluster the training set matrix.

In the summary step, we again use the Glove model to present the vector feature for the testing set, and feed the testing set to our summary model to produce the summary set. The details of our summary model are shown in Fig. 2.

In Fig. 2, every document is fed into the summary model. At first, the document will be split into a list of sentences. After that, these sentences will be transferred into the sentence matrix based on the Glove model. Next, the summary model will get the distance of each sentence with its nearest cluster, then choose the top k lowest distance and pick the corresponding sentence to produce the summary.

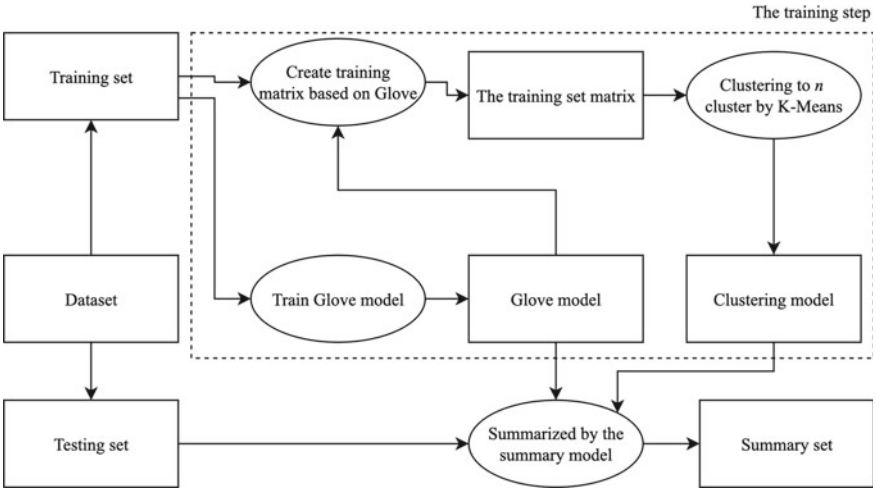


Fig. 1 Process of summary model

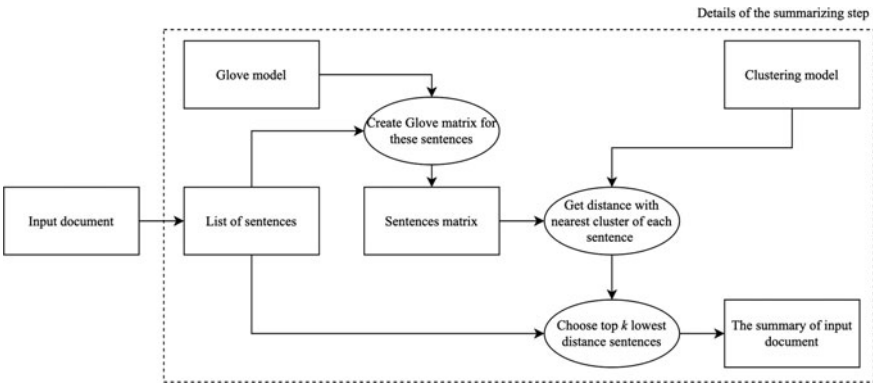


Fig. 2 Process of creation the summary

3.2 K-Means

In our summary model, we need to cluster the training set for ranking the sentence in the summary step. In addition, we evaluate the summary model on a large-scale dataset. Therefore, we need a simple and effective clustering algorithm to hang on with our dataset.

The K-Means [11] is one of the best candidate for our summary model because it is a practical and straightforward algorithm of clustering. We choose Mini-batch K-Means [31] (Algorithm 1), a improved variant of K-Means for working on the large-scale dataset, for clustering step in our summary model.

Algorithm 1 Mini-batch K-Means [31]

Given: k, mini-batch size b, iterations t, data set X
Initialize each $c \in C$ with an x picked randomly from X
$v \leftarrow 0$
for $i = 1$ to t **do**
  $M \leftarrow b$ examples picked randomly from X
  for $x \in M$ **do**
    $\mathbf{d[x]} \leftarrow f(C, x)$ ▷ Cache the center nearest to x
  end for
  for $x \in M$ **do**
    $\mathbf{c} \leftarrow \mathbf{d[x]}$ ▷ Get cached center for this x
    $\mathbf{v[c]} \leftarrow \mathbf{v[c]} + 1$ ▷ Update per-center counts
    $\eta \leftarrow \frac{1}{\mathbf{v[c]}}$ ▷ Get per-center learning rate
    $\mathbf{c} \leftarrow (1 - \eta)\mathbf{c} + \eta\mathbf{x}$ ▷ Take gradient step
  end for
end for

In Algorithm 1: C is the set of cluster centers, $c \in R^m$ is cluster center, $|C| = n$, X is the set of vectors x, $f(C, x)$ is the function that returns the nearest cluster center $c \in C$ to x by using Euclidean distance.

3.3 Word Embedding

In machine learning, the input of algorithms is usually the numeric vectors. That is why we use one kind of word presentation to convert text documents in the dataset to the vectors before feeding them into the machine learning model. There are two main kinds of word presentation: term frequency and word embedding.

When using term frequency word presentation such as the bag of words, the vocabulary is usually created by all terms (words) in the corpus. One text document will be present as a vector with the elements being the term frequency of the particular word in that document. The dimension of the term frequency vector will be the number of words in the vocabulary. Therefore, this kind of vector will be so long and have a lot of zero value. As a result, it needs more time for the operation of the machine learning algorithm.

On the other hand, word embedding tries to present the word (or term) as a short vector, and the document vector could be the sum of the vectors of all of the words in that document. Therefore, word embedding could save a lot of training and predicting time.

In this work, for saving time and speedup of the summary model, we will use Glove [25], one of the word embedding models to present the dataset. The Glove model considered the global relationship of the words in the input text and was introduced in 2014 by Pennington. It will be trained on the training set and used to create the vector input for our summary model.

3.4 Distance Metrics

Distance metrics are used for measuring the distance of the same dimension vectors. In our work, the K-Means algorithm is implemented with Euclidean distance. In the summary model, the distance of the summary candidate sentences with the cluster's centroid of the K-Means model will be the cosine distance.

Cosine Distance. Cosine distance of two vectors $\vec{a}$ and $\vec{b}$ with same dimension n is determined by Eq. (1)

$$\text{cosine\_distance}(\vec{a}, \vec{b}) = 1 - \cos(\theta) = 1 - \frac{\vec{a} \cdot \vec{b}}{|\vec{a}||\vec{b}|} = 1 - \frac{\Sigma_{i=1}^{n} a_i b_i}{\sqrt{\Sigma_{i=1}^{n} a_i^2}\sqrt{\Sigma_{i=1}^{n} b_i^2}} \quad (1)$$

where n is the dimension of vector $\vec{a}$ and vector $\vec{b}$, a_i is the element of vector $\vec{a}$, b_i is the element of vector $\vec{b}$, and θ is the angle formed by $\vec{a}$ and $\vec{b}$.

Euclidean Distance. Euclidean distance d of two vectors $\vec{a}$ and $\vec{b}$ with same dimension n is determined by Eq. (2)

$$d(\vec{a}, \vec{b}) = \sqrt{\Sigma_{i=1}^{n}(a_i - b_i)^2} \tag{2}$$

where n is the vector dimension, a_i is the element of vector $\vec{a}$, and b_i is the element of vector $\vec{b}$.

3.5 ROUGE

ROUGE [18] is the metrics for evaluating the text summary model. It works by comparing the summary produced by the summarization model (the model summary) with reference summaries. In ROUGE, the "overlapping word" is a word visible in both system summary and reference summary. The "overlapping word" is determined by the n-gram model is called ROUGE-N, longest common subsequence is called as ROUGE-L, weighted longest common subsequence is called as ROUGE-W and skip-bigram co-occurrence statistics is called as ROUGE-S. ROUGE's authors suggest that ROUGE-2, ROUGE-L, ROUGE-W, and ROUGE-S will work well for single-document summary tasks. The formula of recall, precision, and F measure in ROUGE is present in Eqs. 3, 4, and 5.

$$R = \frac{S_{\mathrm{O}}}{S_{\mathrm{R}}} \tag{3}$$

$$P = \frac{S_{\mathrm{O}}}{S_{\mathrm{S}}} \tag{4}$$

$$F = \frac{(\beta^2 + 1)PR}{\beta^2 P + R} \tag{5}$$

where R is the recall, P is the precision, F is the F measure, when $\beta = 1$ the F measure will become F_1-score, S_{O} is the number of "overlapping words," S_{R} is the number of words in the reference summary, S_{S} is the number of words in the model summary.

Table 1 VNText detail

Information	Train	Test	Val	Test-Ref	Val-Ref
Articles	880,895.00	110,103.00	110,103.00	110,103.00	110,103.00
Sentences	18,738,333.00	2,342,296.00	2,333,519.00	155,573.00	155,496.00
Words	452,686,377.00	56,563,630.00	56,417,036.00	4,163,723.00	4,170,809.00
Words/Sent	24.16	24.15	24.18	26.76	26.82
Words/Art	513.89	513.73	512.40	37.82	37.88
Sents/Art	21.27	21.27	21.19	1.41	1.41

4 Experiment

4.1 Dataset

We built the dataset **VNText** by collecting the articles from the Vietnamese online newspapers website, which is the same idea of the author of **CNN/Daily Mail** dataset [12, 32]. The collected articles are in HTML form, and we extract the plain text by using HTML Tag and most common metadata in its content and put them together to create the evaluating dataset. The VNText dataset has 1,101,101 records; every sample included the title, subtitle, and main content. The subtitle will be used as the reference summary, and the content will be used as the sampling. Next, we split the VNText into three subsets: the training set includes 880,895 samples, the validation set contains 110,103 samples, and the test set has 110,103 samples. The details of the VNText are shown in Table 1. The training set is used for training the Glove model and the K-Means models, the validation set is used for turning parameters of the K-Means models, and the test set is used to evaluate the summary model.

In Table 1, $Words/Sent$ is average number of words per sentence, $Words/Art$ is average number of words per article, $Sents/Art$ is average number of sentences per article, $Test - Ref$ is the reference summary in the test set, $Val - Ref$ is the reference summary in the validation set.

4.2 Parameters

We choose the parameters for training the Glove model based on Glove publication [25]. There are the word vector dimension $embedding\_len = 300$, number of epochs $epoch = 15$, size of the context windows $window\_size = 15$, kind of distributed representations model is $skip\_gram$.

In the training K-Means model step, we have to choose n for the K-Means model. We use the value of clusters n in the set of {100, 500, 1000, 1500, 2000} for experimentation.

Table 2 Average number of words in the summary

n	$k=1$		$k=2$		$k=3$		$k=4$		$k=5$	
	$w/sent$	w/sum	$w/sent$	w/sum	$w/sent$	w/sum	$w/sent$	w/sum	$w/sent$	w/sum
100	43.21	43.21	43.17	85.30	43.06	126.62	42.96	167.13	42.85	206.55
500	44.27	44.27	44.28	87.50	44.21	130.02	44.15	171.78	44.08	212.49
1000	44.76	44.76	44.75	88.43	44.71	131.47	44.65	173.71	44.59	214.91
1500	44.92	44.92	44.96	88.84	44.91	132.06	44.86	174.54	44.80	215.96
2000	45.05	45.05	45.09	89.09	45.05	132.47	45.01	175.11	44.96	216.73

In the summary step, we have two parameters, the number of sentences in summary k and the distance metric, which is based to find the closest sentence vector and the cluster centroids. Based on the validation, we choose $k \in \{1, 2, 3, 4, 5\}$, and the distance metric is cosine.

4.3 The Computer for Evaluating

All of the models and processes in this work are run on one computer. Its CPU is an ARM Neoverse-N1 4 cores 2.8 GHz, one thread per core, 24 GB Ram, 150.34 MB/s read and write disk speed.

5 Results

5.1 The Summary

Table 2 shows the average number of words in the summary results. In this table, k is the number of sentences per summary, n is the number of clusters of the K-Means model, $w/sent$ is the average number of words per sentence, w/sum is the average number of words per summary.

In Table 2, the number of words in the summary is increased when n goes up.

5.2 ROUGE Score

Tables 3, 4, and 5 show the F_1-score of the summary results based on ROUGE-1, ROUGE-2, and ROUGE-L with variant n. These results are also shown as visualized in Fig. 3.

Table 3 F_1-score (%) of the summary results based on ROUGE-1

n	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
100	51.25	42.32	34.16	28.54	24.65
500	51.37	42.05	33.76	28.11	24.21
1000	51.38	41.87	33.56	27.91	24.01
1500	51.39	41.80	33.48	27.83	23.94
2000	**51.40**	41.75	33.41	27.76	23.87

Table 4 F_1-score (%) of the summary results based on ROUGE-2

n	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
100	15.82	12.60	10.23	8.69	7.66
500	16.03	12.67	10.21	8.62	7.54
1000	16.10	12.62	10.14	8.54	7.45
1500	16.12	12.59	10.11	8.51	7.42
2000	**16.15**	12.59	10.09	8.48	7.38

Table 5 F_1-score (%) of the summary results based on ROUGE-L

n	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
100	29.08	26.73	23.66	21.10	19.08
500	29.15	26.63	23.49	20.91	18.87
1000	29.16	26.58	23.42	20.82	18.78
1500	29.16	26.56	23.39	20.79	18.74
2000	**29.18**	26.54	23.36	20.76	18.71

In ROUGE-1, the highest F_1-score is **51.40%**, with $n = 2000$ and $k = 1$, the F_1-score is slightly increase when increase n but go down significantly when k go up, it because the length of the reference summary, it about 37.82 words per reference summary.

The best F_1-score is **16.15%** based on ROUGE-2 and **29.18%** based on ROUGE-L when $n = 2000$, $k = 1$ and remains a up trend when n goes up. F_1-score also remains in the down direction when k goes up based on ROUGE-2 and ROUGE-L.

5.3 Time

In this work, we have three steps to take time: training the Glove model, training the K-Means model, and Evaluating process, which include summarizing and calculating ROUGE score.

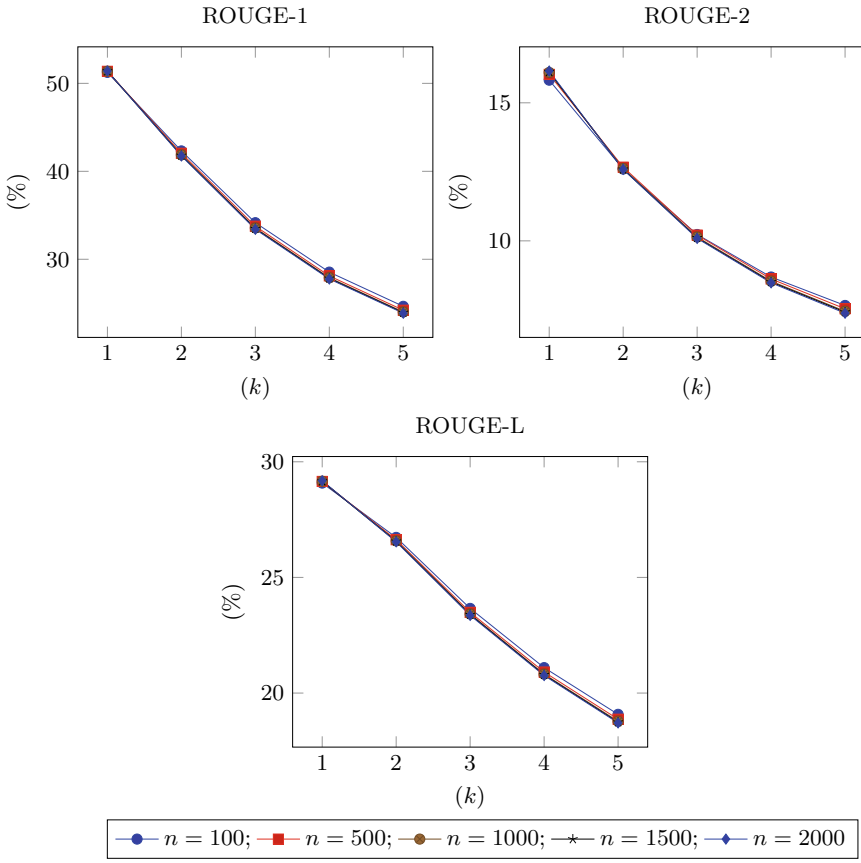


Fig. 3 F_1-score of the summary model based on ROUGE

The Glove model is trained one time, and it takes 4756.82 s.

Table 6 shows the costing time of training the K-Means model with the training set, the time to summarize the test set, the time of running ROUGE evaluating and the sum of them. The results show that the time cost is dependent on the summarizing step and the calculation ROUGE step and slightly increases when n runs into higher.

In Table 6, the highest summary time is **3425.07** s, and this time is spend for create summary for all documents in the test set, which include 110,103 documents in five times ($k \in \{1, 2, 3, 4, 5\}$) when $n = 2000$. Therefore, the average time cost to summarize one document in the test set is 6.22 ms. This is an impressive time cost, and it shows that our summary model can work efficiently on the large-scale dataset.

Table 6 Time for evaluating (in seconds)

n	K-Means	Summary	Rouge	Total
100	16.04	791.20	2367.68	7915.70
500	21.98	1346.38	2426.93	8530.13
1000	27.95	2014.12	2456.31	9227.26
1500	35.75	2717.60	2466.11	9940.52
2000	44.15	**3425.07**	2473.50	10,655.39

6 Conclusion

We introduced the extractive summarization model for the single-document and presented the experimentation on the Vietnamese large-scale dataset.

The first contribution of this work is proving that single-document extractive text summarization can be based on the pre-trained clusters, which trained on the large-scale dataset. The second contribution is using word embedding as the vector feature to improve the summary time and maintain a high ROUGE score. The third contribution is showing that our summary model can work well on the large-scale dataset when using word embedding.

A forthcoming improvement will be to implement the distributed and parallel algorithms to speedup the training task for the automatic text summarization.

References

1. Agrawal A, Gupta U (2014) Extraction based approach for text summarization using k-means clustering. Int J Sci Res Publ 4(11):1–4
2. Akter S, Asa AS, Uddin MP, Hossain MD, Roy SK, Afjal MI (2017) An extractive text summarization technique for Bengali document (s) using k-means clustering algorithm. In: 2017 IEEE international conference on imaging, vision & pattern recognition (icIVPR). IEEE, pp 1–6
3. Al-Sabahi K, Zuping Z, Kang Y (2018) Latent semantic analysis approach for document summarization based on word embeddings. arXiv preprint arXiv:1807.02748
4. Aliguliyev RM (2009) A new sentence similarity measure and sentence based extractive technique for automatic text summarization. Expert Syst Appl 36(4):7764–7772
5. Allahyari M, Pouriyeh S, Assefi M, Safaei S, Trippe ED, Gutierrez JB, Kochut K (2017) Text summarization techniques: a brief survey. arXiv preprint arXiv:1707.02268
6. Deshpande AR, Lobo L (2013) Text summarization using clustering technique. Int J Eng Trends Technol 4(8):3348–3351
7. Dong Y, Shen Y, Crawford E, van Hoof H, Cheung JCK (2018) Banditsum: extractive summarization as a contextual bandit. arXiv preprint arXiv:1809.09672
8. El-Kassas WS, Salama CR, Rafea AA, Mohamed HK (2021) Automatic text summarization: a comprehensive survey. Expert Syst Appl 165:113679. https://doi.org/10.1016/j.eswa.2020.113679
9. Gambhir M, Gupta V (2017) Recent automatic text summarization techniques: a survey. Artif Intell Rev 47(1):1–66

10. Graff D, Kong J, Chen K, Maeda K (2003) English gigaword. Linguistic data consortium, vol 4, no 1. Philadelphia, p 34
11. Hartigan JA, Wong MA (1979) Algorithm AS 136: a k-means clustering algorithm. J R Stat Soc Ser C (Appl Stat) 28(1):100–108
12. Hermann KM, Kocisky T, Grefenstette E, Espeholt L, Kay W, Suleyman M, Blunsom P (2015) Teaching machines to read and comprehend. In: Advances in neural information processing systems, pp 1693–1701
13. Hovy E, Lin CY (1998) Automated text summarization and the summarist system. In: TIPSTER '98: proceedings of a workshop, Baltimore, Maryland, 13–15 Oct 1998. Association for Computational Linguistics, USA, pp 197–214. https://doi.org/10.3115/1119089.1119121
14. Jain A, Bhatia D, Thakur MK (2017) Extractive text summarization using word vector embedding. In: 2017 international conference on machine learning and data science (MLDS). IEEE, pp 51–55
15. Jia R, Cao Y, Tang H, Fang F, Cao C, Wang S (2020) Neural extractive summarization with hierarchical attentive heterogeneous graph network. In: Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP), pp 3622–3631
16. Jiang Z, Srivastava M, Krishna S, Akodes D, Schwartz R (2020) Combining word embeddings and n-grams for unsupervised document summarization. arXiv preprint arXiv:2004.14119
17. Le HT, Le TM (2013) An approach to abstractive text summarization. In: 2013 international conference on soft computing and pattern recognition (SoCPaR). IEEE, pp 371–376
18. Lin CY (2004) Rouge: a package for automatic evaluation of summaries. In: Text summarization branches out, pp 74–81
19. Liu Y, Lapata M (2019) Text summarization with pretrained encoders. arXiv preprint arXiv:1908.08345
20. Luhn HP (1958) The automatic creation of literature abstracts. IBM J Res Dev 2(2):159–165
21. Mohd M, Jan R, Shah M (2020) Text document summarization using word embedding. Expert Syst Appl 143:112958
22. Nallapati R, Zhou B, Gulcehre C, Xiang B et al (2016) Abstractive text summarization using sequence-to-sequence RNNs and beyond. arXiv preprint arXiv:1602.06023
23. Nenkova A, McKeown K (2012) A survey of text summarization techniques. In: Mining text data. Springer, pp 43–76
24. Nguyen-Hoang TA, Nguyen K, Tran QV (2012) TSGVi: a graph-based summarization system for Vietnamese documents. J Ambient Intell Humaniz Comput 3(4):305–313
25. Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543
26. Quoc HT, Van Nguyen K, Nguyen NLT, Nguyen AGT (2021) Monolingual versus multilingual BERTology for Vietnamese extractive multi-document summarization. arXiv preprint arXiv:2108.13741
27. Radev DR, Jing H, Styś M, Tam D (2004) Centroid-based summarization of multiple documents. Inf Process Manag 40(6):919–938
28. Rani R, Lobiyal DK (2021) A weighted word embedding based approach for extractive text summarization. Expert Syst Appl 186:115867
29. Rossiello G, Basile P, Semeraro G (2017) Centroid-based text summarization through compositionality of word embeddings. In: Proceedings of the MultiLing 2017 workshop on summarization and summary evaluation across source types and genres, pp 12–21
30. Rush AM, Chopra S, Weston J (2015) A neural attention model for abstractive sentence summarization. In: Proceedings of the 2015 conference on empirical methods in natural language processing. https://doi.org/10.18653/v1/d15-1044
31. Sculley D (2010) Web-scale k-means clustering. In: Proceedings of the 19th international conference on world wide web, pp 1177–1178
32. See A, Liu PJ, Manning CD (2017) Get to the point: summarization with pointer-generator networks. arXiv preprint arXiv:1704.04368

33. Zhang PY, Li CH (2009) Automatic text summarization based on sentences clustering and extraction. In: 2009 2nd IEEE international conference on computer science and information technology. IEEE, pp 167–170
34. Zhong M, Liu P, Chen Y, Wang D, Qiu X, Huang X (2020) Extractive summarization as text matching. arXiv preprint arXiv:2004.08795
35. Zhong M, Liu P, Wang D, Qiu X, Huang X (2019) Searching for effective neural extractive summarization: what works and what's next. arXiv preprint arXiv:1907.03491

Multi-agent System: Efficiency Enhancement and Search for Anomalies in Equipment Operation

Evgeniy I. Zaytsev, Elena V. Nurmatova, Rustam F. Khalabiya, Irina V. Stepanova, and Lyudmila V. Bunina

Abstract The paper considers methods to improve the efficiency and reliability of the multi-agent system of knowledge representation and processing (MASKRP). An approach to the implementation of MASKRP control subsystems based on an exo-kernel operating system and specialized event-driven software modules is proposed. The tasks of effective distribution of software agents on the nodes of the computing system, optimization of Distributed Knowledge Base (DKB), formation of computing clusters with minimal information connectivity are solved. The architecture of the monitoring system is reviewed. The main results on the development of software for searching anomalies in the work of the MASKRP computing cluster equipment as well as on the determination of the optimal parameters for it to prevent false alarms are given.

Keywords Distributed systems · Multi-agent systems · Intelligent software agents · Knowledge representation · Anomalies · Analysis · Monitoring

1 Introduction

To solve complex problems, reactive and cognitive software agents of the multi-agent system of knowledge representation and processing are used, which are able to take into account inaccuracy, incompleteness, fuzziness of knowledge [1–5]. Intelligent software agents of the MASKRP function in nodes of a distributed computing system. Each node is a shared memory multiprocessor. The problem to be solved on the computational node is divided into separate subtasks, which are distributed among the reactive agents. The priorities of the reactive agents included in the group are set according to the sequence number of the program agent. The work of the group's reactive agents is coordinated by the cognitive agent of this node. The scale of resources of the multi-agent solver can be increased horizontally.

E. I. Zaytsev (✉) · E. V. Nurmatova · R. F. Khalabiya · I. V. Stepanova · L. V. Bunina
MIREA—Russian Technological University, Moscow, Russia
e-mail: zajcev@mirea.ru

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_38

Single-node agents communicate with each other through the local interprocess communication (IPC) mechanism. Cognitive agents of different nodes interact with each other through messaging by middleware. To improve performance, IPC is implemented with specific system software modules (LibOS) that have direct access to the hardware. The structure of the multi-agent solver of the MASKRP is shown in Fig. 1.

For processing knowledge of the problematic components, special cognitive data structures (CDS) are used that describe the plan for solving the target task, define the rules for reducing the task to the subtasks, and establish links between the software agents responsible for solving these subtasks. The functional and structural organization of the problem-oriented multi-agent solver is described in more detail in paper [5].

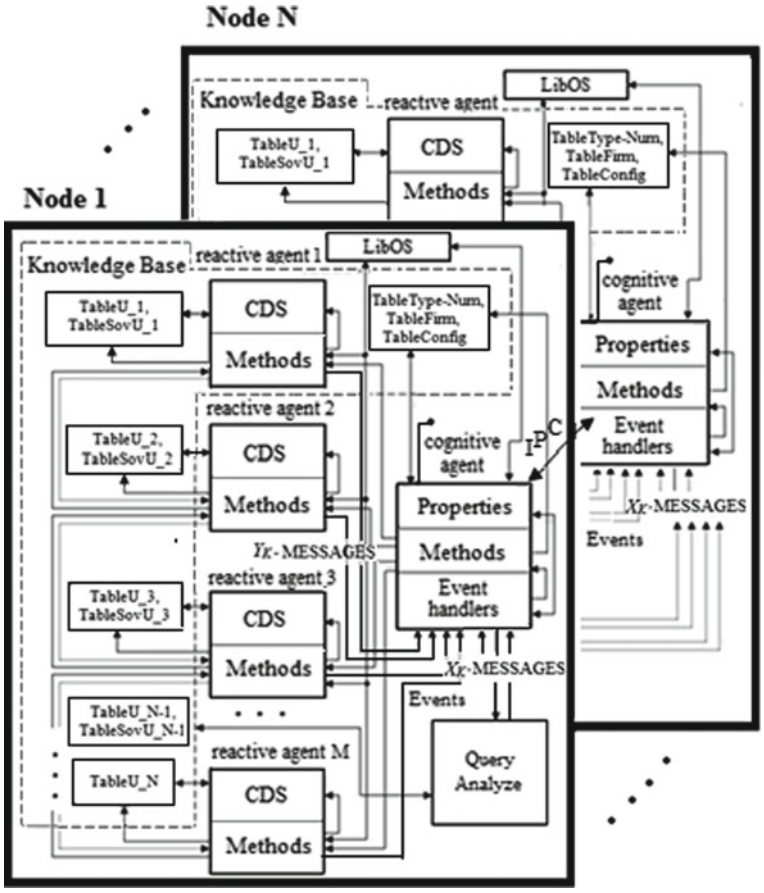


Fig. 1 Structure of the multi-agent solver

The system performance and reliability are improved by implementing special mechanisms for computing resource management, which take into account the specifics of the multi-agent system [6–12].

Traditional operating systems hide information about computational resources behind universal high-level abstractions, which does not allow the MASKRP developer to implement problem-dependent optimization. The key to creating a highly efficient knowledge processing system is specialized computational resource scheduling and management modules that use exo-kernel operating system services.

MASKRP is a complex system that includes a variety of physical equipment—air conditioning systems, uninterruptible power supplies, computing servers, and storage systems. Various anomalies in the operation of this equipment can occur during operation. To manage the work of the MASKRP computing nodes, developed a system to find anomalies in the work of computing cluster equipment, including an indexing cluster (cluster of nodes that receive data), nodes to store real data, and a search cluster to search for stored data. The system allows you to automatically find shift anomalies in the time series, combine the anomalies found in the JSON data array.

2 Performance Improvement of the MASKRP

Each computing node of the MASKRP can use a parallelism model, in which multiple software agents are in a state of execution. To implement parallel computations in a computational node, multiprocessors with uniform memory access (UMA) are used. In this case, the whole set of software agents is divided into subsets (groups). Agents belonging to different groups can act simultaneously.

The agents are divided into groups using the compatibility matrix S and the inclusion matrix R. The compatibility matrix S has the following form:

$$S = \begin{bmatrix} 0 & s_{12} & s_{13} & \dots & s_{1M} \\ s_{21} & 0 & s_{23} & \dots & s_{2M} \\ s_{31} & s_{32} & 0 & \dots & s_{3M} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{M1} & s_{M2} & s_{M3} & \dots & 0 \end{bmatrix} \begin{matrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_M \end{matrix} \quad (1)$$

where $s_{ij} = 1$, if the agents A_i and A_j use different computing resources and work in parallel, otherwise $s_{ij} = 0$.

The distribution of agents into groups is based on the inclusion matrix R:

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1M} \\ r_{21} & r_{22} & \dots & r_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ r_{H1} & r_{H2} & \dots & r_{HM} \end{bmatrix} \begin{matrix} R_1 \\ R_2 \\ \vdots \\ R_H \end{matrix} \quad (2)$$

where M is the number of agents, H is the number of groups. $r_{ij} = 1$ if the agent A_i is included in the group Y_j. The agent A_i is included in the group Y_j if $S_i \cap R_j = \emptyset$ that is the matrix rows do not intersect. For optimal partitioning into subsets, it is necessary to consider the functional features of the software agents, their requirements for computing resources, as well as know the structural organization of the MASKRP node used to implement parallel computations.

To improve the performance of the multi-agent solver and reduce the total volume of data transferred over the network, the agent-associated distributed knowledge base is divided into clusters. At the same time, two complementary mechanisms for interaction of software agents are used: the mechanism of localization of inter-agent interactions and the mechanism of dynamic distribution of agents in the MASKRP nodes [13].

3 Monitoring System

The architecture of the monitoring system consists of three main clusters of nodes: indexing for receiving data, direct nodes for storing real data, and a search cluster for performing a search on stored data. Each of these components can be easily expanded horizontally if higher performance is required. To store data in such a system, there are collectors—programs that collect various types of data, such as environment metrics, infrastructure data, cluster metrics, and send this data using a well-documented REST API to the monitoring core. For data analysis and monitoring, there is a built-in simple Web interface that allows you to create your own dashboards for data groups or simply navigate through raw data with a user-friendly interface. The main point in the architecture is to add new monitoring data that does not require changing the monitoring core (indexing cluster, data nodes, search cluster), and for this, you just need to create or deploy the appropriate collector and create a dashboard for the monitored data. According to the above architecture, several collectors and toolbars have already been implemented, mainly for engineering infrastructure, as the fundamental layer of the entire cluster. One of these systems is shown in Fig. 2.

Non-standard situations in monitoring systems are usually associated with significant changes in the time series describing certain parameters of the monitored system. One of the types of such changes is the shift of the average level of the time series, which can be both very abrupt and gradual. The method “E-Divisive with Medians” (EDM) [14] and implemented in the package for R BreakoutDetection can automatically detect such structural changes in time series.

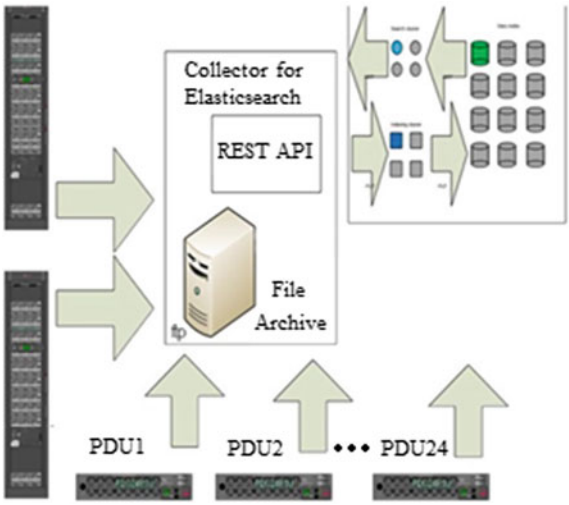


Fig. 2 Structure collector for UPS cluster systems

The method is based on the use of E-statistics, which characterizes the distance between observations from random samples. The energy distance between X and Y is defined as follows:

$$(X, Y) = 2E|X - Y| - E\left|X - X'\right| - E\left|Y - Y'\right|, \tag{3}$$

where X, Y—independent random variables; X', Y'—independent equally distributed as X and Y; $E|*|$—expected value [14].

The Langer $\varepsilon(X, Y)$ is the more reason to believe that the samples being compared come from different general populations. In order to avoid the influence of anomalous observations common in practice, the EDM method operates with medians rather than averages. Moreover, it is a nonparametric method, i.e., it does not assume that the analyzed observations belong to any particular probability distribution. The nonparametric nature of EDM, as well as the use of medians, makes this method particularly suitable for detecting structural changes in time series in practical problems.

The calculation of ε can be approximated by the formula:

$$\hat{\varepsilon}(X_n, Y_m; \alpha) = \frac{2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left|x_i - y_j\right|^{\alpha} - \left(\frac{n}{2}\right)^{-1}\sum_{i<j}\left|x_i - x_j\right|^{\alpha} - \left(\frac{m}{2}\right)^{-1}\sum_{i<j}\left|y_i - y_j\right|^{\alpha}, \tag{4}$$

where X and Y are selected from the time series of values obtained from the sensors.

The basic algorithm on which the EDM package is based uses energy statistics to detect divergence in the mean. Note that EDM can also be used to detect changes in the distribution in a given time series. EDM uses reliable statistical metrics, namely

the median, and estimates the statistical significance of the breakout through the permutation test. In addition, EDM is nonparametric. This is important because the distribution of production data rarely (if ever) follows a generally accepted normal distribution or any other widely accepted model. The following arguments should be specified in breakout():

1. *Z*—time series represented either as a numeric vector or as a data table with columns timestamp ("value time stamp") и count ("value count").
2. min.size—the minimum number of observations that must occur between the breakpoints of the trend in the time series (the default is 30).
3. method—scenario applied to structural change detection data takes two possible values: "amoc" (from "at most one change," i.e., "no more than one change") and "multi" (multiple break points).

The anomaly search system works with a large amount of data, as a consequence there is a high probability of more than one shift, so it was decided to take "multi" as a method.

The cluster continuously receives a large flow of information. For its correct analysis, it is necessary to make a careful tuning of the system [12, 15, 16].

Since it was decided to use a scenario to detect structural changes with multiple breakpoints, it is necessary to pay attention to the following parameters:

- degree—parameter that defines the proximity of the break points (the default is 1);
- percent—minimum relative (%) increases in internal EDM quality criterion;
- beta—parameter for regularization if degree and/or percent is not specified (default is 0.008).

For the parameter percent, which is responsible for the quality of the shift, it was decided to choose the value percent = 0.7 or otherwise 70%. If the degree parameter is set to zero, it may lead to an increase in the number of detected breakpoints and therefore to an excessive number of system triggers. Decreasing the parameter beta also resulted in excessive triggers. Based on the tests, it was decided to use the system with parameter percent = 0.7 and leave the rest by default.

In the *R*-module, responsible directly for the search for anomalies, a code of basic functions was developed:

1. getParameters—the function is responsible for receiving the following parameters when calling this module: index—ES index; type—index type; source—data array, in which the search for anomalies will be performed. This parameter includes data and date.
2. detectBreakout2133—the function is responsible for finding anomalies. According to the parameters provided, this function takes the last thousand records, then converts the date parameter into the desired format and combines it together with the data field into a single dataframe, which after analysis by the breakout function gives the positions of anomalies in the dataframe.

3. main—the main function, which calls the detectBreakout213—function with parameters obtained by the getParameters function. It then translates the obtained anomalies into JSON format.

The bash module runs periodically with a crontab program. Every five minutes, it transmits data from the UPS sensors to the *R*-module to search for anomalies.

To test the correctness of the developed system, a segment of the time series was chosen. Due to unfavorable weather conditions in the time series, it was possible to observe the type of anomaly needed for testing.

Below are three examples of how the anomaly search system works on time series:

1. %Cap—Battery Charge.
   System output:
   {"timestamp":"2021-03-13T12:10:51+0300","anom":100},
   {"timestamp":"2021-03-13T15:15:51+0300","anom":80.1}

In Fig. 3, the ordinate shows the charge of the batteries (Amps per unit time); the abscissa shows the time distribution over 10 min.

2. Iin—Input current.
   System output:
   {"timestamp":"2021-03-13T12:20:51+0300","anom":47},
   {"timestamp":"2021-03-13T14:55:51+0300","anom":39}

Figure 4 shows the time distribution of the current (Amps) for 5 min.

3. kVAout—full power output
   System output:
   {"timestamp":"2021-03-13T11:55:50+0300","anom":30.7},
   {"timestamp":"2021-03-13T14:50:51+0300","anom":24.6}.

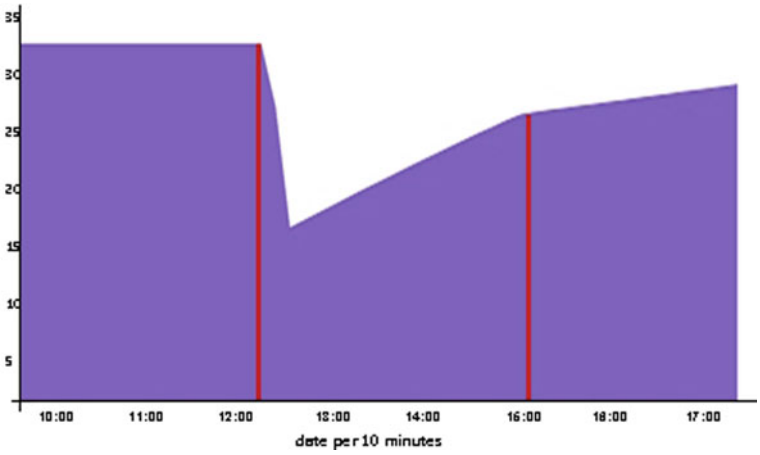


Fig. 3 Time series anomaly: battery charge

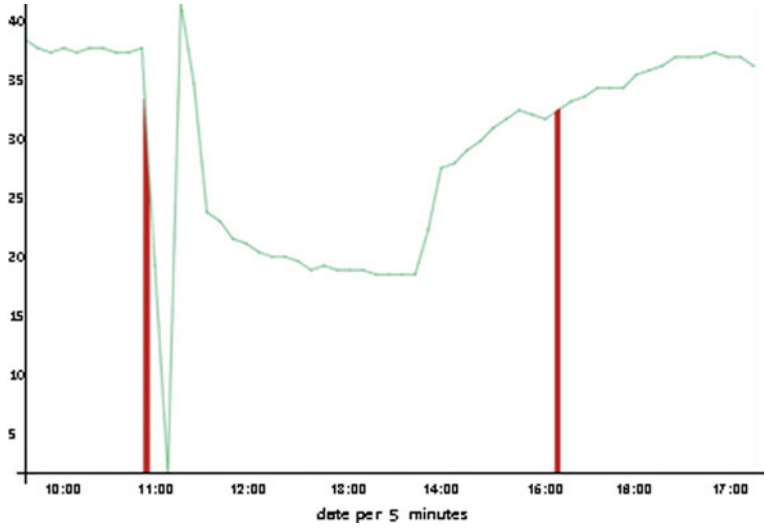


Fig. 4 Time series anomaly: input current

Figure 5 shows the time distribution of output power over 5 min.

As can be seen from the output data and graphs, the system successfully detected anomalies in the time series provided to it. As a consequence, can say that the system has successfully coped with the task.

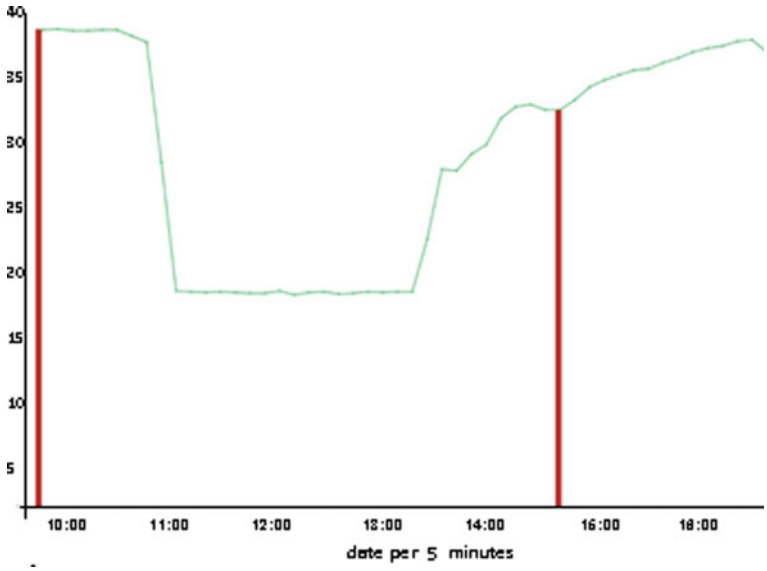


Fig. 5 Time series anomaly: output power

4 Conclusion

The article an approach to improving the performance and reliability of the MASKRP is proposed, which includes the efficient distribution of software agents in the multi-agent solver of the computing node, formation of computing clusters with minimal information connectivity as well as the construction of an optimal logical structure of the distributed knowledge base. An approach to the implementation of MASKRP management subsystems based on an exo-kernel operating system and specialized event-driven software modules is considered.

The architecture of the monitoring system is reviewed. The software code of the main module of the system to search for anomalies in the operation of the computing cluster equipment was developed. All functions performed by the system were thoroughly tested during the development process and tested on the time series. The developed system significantly increases the reliability of the cluster, as well as reduces the analysis time of possible anomalies.

References

1. Aly S, Badoor H (2018) Performance evaluation of a multi-agent system using fuzzy model. 1st international workshop on deep and representation learning (IWDRL), Cairo, pp 175–189
2. Baranauskas R, Janaviciute A, Jasinevicius R, Jukavicius V (2015) On multi-agent systems intellectics. Inf Technol Control 1:112–121
3. Houhamdi Z, Athamena B, Abuzaineddin R, Muhairat M (2019) A Multi-agent system for course timetable generation. TEM J 8:211–221
4. Raj JS, Vijitha Ananthi J (2019) Recurrent neural networks and nonlinear prediction in support vector machines. J Soft Comput Paradigm (JSCP) 1(01):33–40
5. Zaytsev EI, Khalabiya RF, Stepanova IV, Bunina LV (2020) Multi-agent system of knowledge representation and processing. In: Proceedings of the fourth international scientific conference "intelligent information technologies for industry" (IITI'19), Springer, pp 131–141
6. Azeroual M, Lamhamdi T, El Moussaoui H, El Markhi H (2020) The intelligent energy management system of a smart microgrid using multiagent systems. Arch Electr Eng 69:23–38. https://doi.org/10.24425/aee.2020.131756. Search in Google Scholar
7. Batouma N, Sourrouille J (2011) Dynamic adaption of resource aware distributed applications. Int J Grid Distrib Comput 4(2):25–42
8. Cardoso R, Hübner J, Bordini R (2013) Benchmarking communication in actor-and agent-based languages. Engineering multi-agent system. Springer, Berlin, pp 58–77
9. Darweesh S, Shehata H (2018) Performance evaluation of a multi-agent system using fuzzy model. In: 1st international workshop on deep and representation learning (IWDRL), pp 7–12
10. Jurasovic K, Jezic G, Kusek M (2006) Performance analysis of multi-agent systems. Int Trans Syst Sci Appl 4:601–608
11. Nazarov AN (2019) Processing streams in a monitoring cloud cluster. Russ Technol J 7(6):56–67. https://doi.org/10.32362/2500-316X-2019-7-6-56-67
12. Petushkov GV (2020) Evaluation and reliability prediction for highly reliable software and hardware systems: the case of data processing centers. Russ Technol J. 8(1):21–26. https://doi.org/10.32362/2500-316X-2020-8-1-21-26 (In Russ.)
13. Nurmatova EV, Gusev VV, Kotliar VV (2018) Analysis of the features of the optimal logical structure of distributed databases. Collection of works the 8th international conference "distributed computing and grid-technologies in science and education". Dubna, 167 p

14. Szekely GJ, Rizzo ML (2013) Energy statistics: a class of statistics based on distances. J Stat Plan Infer, pp 1249–1272. https://doi.org/10.1016/j.jspi.2013.03.018
15. Malik P (2019) Network intrusion detection system using two stage classifier. In: International conference on inventive computation technologies. Springer, Cham, pp 749–757
16. Sizov VA, Kirov AD (2021) The development of models of an analytical data processing system for monitoring information security of an informatization object using cloud infrastructure. Russ Technol J 9(6):16–25. https://doi.org/10.32362/2500-316X-2021-9-6-16-25 (In Russ.)

An Efficient Text Detection and Recognition Framework for Natural Scene Images

Senu Jerome and Anuj Mohamed

Abstract For understanding the content of the image, a vital clue is the superimposed text in images. However, ineffective text extraction might be caused by the poor contrast and complicated background of the image. An effective text detection and recognition framework as of natural scene images utilizing Anopheles search with convolution neural network are proposed to mitigate such complexities. In this work, for detecting and recognizing the exact text from the image, different steps are undergone by the input images from the street view text (SVT) dataset. The precise text is extracted from the images. For comparing the results attained by the proposed AS-CNN with the preceding top-notch algorithms, SVT and publically accessible datasets are utilized in this research. The experimental outcomes denoted that the proposed AS-CNN exhibits promising results, effectively decreasing the overall computation cost and time.

Keywords Text detection · Text recognition · Convolutional neural network · IGrabCut · Anopheles search optimization · Tunicate swarm algorithm

1 Introduction

Text detection and recognition are a very important task in computer vision-related applications. Numerous researches are being done in the area so as to enable accurate detection, extraction and recognition of text. In the text information, the image's most significant content is reflected [1]. Thus, for easy understanding of the idea behind the images, the extraction of text as of the images is helpful [2]. For several multimedia tasks, namely visual classification, video analysis, mobile applications, scene understanding, etc., a vital requirement is the text detection (TD) process

S. Jerome (✉)
Research Scholar, School of Computer Sciences, Mahatma Gandhi University, Kottayam, India
e-mail: sinuabi@fisat.ac.in; sinuabi@gmail.com

A. Mohamed
School of Computer Sciences, Mahatma Gandhi University, Kottayam, India
e-mail: anujmohamed@mgu.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_39

[3]. Owing to complex situations, such as uneven lighting, blurring, perspective distortion, and orientation, the TD in a natural scene image is difficult [4, 5].

Recently, numerous approaches and algorithms have been developed to mitigate the complexities of text detection. The approaches exhibit a few limitations, concerning higher computation time, irrelevant feature selection (FS), and random weight updation process, even though they have illustrated promising results. The detected horizontal regions are returned by the majority of the methods which depend on horizontal or nearly horizontal annotations. But, the huge number of text regions is not horizontal in real-world applications. Thus, desired results might not be exhibited by such techniques. Long or large text instances could not be effectively managed by several deep learning methods, and they often cause imprecise localization of text. The TD task's performance is often degraded by unsatisfactory text localization, and it also affects the performance of the following text recognition (TR) task. For addressing such issues, an effective TDR framework from NSI is proffered by the work utilizing AS-CNN for the TR model.

The remaining paper is organized as: Sect. 2 examines the related works concerning the proposed work. Section 3 elucidates the proposed efficient TDR framework of NSI utilizing AS-CNN for the TR model. Section 4 renders the results and discussion for the proposed work centered on performance metrics. Section 5 provides the conclusion with future work.

2 Literature Survey

In the recent years, many techniques and approches have been introduced to facilitate accurate detection, extraction, and recognition of text from images.

Roy et al. [6] created a technique named Delaunay triangulation (DT) centered TD as of multi-view of scenes. The same scene was captured via orthogonal direction, disparate angles, and height. The features that explained the text components, like corner points, area, and cavity centered on DT, were extracted by the technique. Thus, the method could handle multi-oriented text. The experimental outcomes denoted that the recognition step removed the false positives, and also it enhanced the method's overall performance concerning the F-measure score. But, the scheme could not handle the randomly oriented text lines as of the images.

Zhonga et al. [7] applied the method that changed the bounding box regression module with a LocNet centered localization module and enhanced the localization accuracy (LA) of a faster region-centered CNN (R-CNN)-based text detector. Experiments exhibit that the LA was boosted by the proposed work. Also, superior performance was attained on horizontal along with multi-oriented TD. But, the semantic information that was closer to the image's content could not be extracted by the method.

Kantipudi et al. [8] proposed a novel approach for scene text recognition that integrates bidirectional LSTM and deep convolution neural networks. CNN is used to generate the ordered sequence of the features from the contoured image. The

sequence of features is coded using the Bi-LSTM. Proposed work is compared with the existing state-of-the-art methods, and according to the analysis, precision and accuracy are improved. For some of the images, results are partially accurate with images which were captured from a long distance.

Dai et al. [9] proposed an accurate scene text detection via scale-aware data augmentation and shape similarity constraint, which uses (SADA) technique to increase the diversity of training samples. A shape similarity constraint (SSC) technique is exploited to model the global shape structure of arbitrary-shapescene texts. The proposed method is capable of working well in many challenging scenarios, but it fails to remain the valid text instances when the overlaps between text instances are large.

Butt et al. [10] proposed a CNN-RNN model with an attention mechanism for Arabic image text recognition. A bidirectional RNN with an attention mechanism is used to generate output, enabling the model to select relevant information from the feature sequences. This model gives better results from those based on a simple CNN, RNN, and hybrid CNN-RNN on Arabic text datasets such as Alif and Activ with no focus on other language text extraction.

Ma et al. [11] proposed a residual attention-based multi-scale script identification in scene text images which uses a hierarchical feature fusion block to extract the multi-scale features. A fully convolutional classifier was used to generate channel-level classifications which are then processed by a global pooling layer to improve classification efficiency. The proposed methodology was not accurate in dealing with languages having curved font structure.

Naiemi et al. [12] proposed a novel pipeline framework for multi-oriented scene text image detection and recognition which uses a convolutional neural network-based pipeline to obtain high-level visual features and improve text detection and recognition efficiency. A pre-trained ResNet-50 network on ImageNet and SynthText for extracting low-level visual features was used in this study. The method failed to detect text in crowded and shiny backgrounds. It also failed to recognize small text at inappropriate angles.

Wang et al. [13] presented an efficient method for the TD in NSI. By the amalgamation of convolutional neural network scores of maximally stable external region along with multiple neighborhood information, the conditional random field model was generated by the method. Besides, the utilization of context information recovered the missed text components. Moreover, the gray and binary features were combined by the method. The shape-specific classifiers were also designed by the method, which precisely verified the text lines as of the images. Promising performance was attained by the work's final outcome, and it revealed the model's effectiveness and robustness. However, the scheme possessed the disadvantage of the random weight generation process. Thus, the method's classification accuracy was low.

Shivakumara et al. [14] formed a new multi-modal approach for bib number detection and TR in Marathon images. The upper body detection from the image was examined by the method. A GrabCut algorithm segmented the foreground regions. For the torso detection, the pictorial structural model had been utilized using CRF in Marathon images. Experimental outcomes showed that better results were attained

by the approach after torso detection when compared to before torso detection. But, the text could not be read effectively by the scheme as of the low resolution along with poor contrast images.

Xu et al. [15] created an English text repair and an extraction algorithm for color images centered on an improved edge detection algorithm. The experimental outcomes indicated that the algorithm attained higher accuracy and quicker computing speed. The algorithm also offered a better comprehensive effect. Nevertheless, the accurate sufficient bounding boxes were not predicted by the method for text instances as of the long distances.

3 Proposed Methodology

The text information is the most crucial among the contents in images. The text encompasses useful information in an image that assists in attaining the overall idea behind the image. For horizontal or near horizontal texts, several TDR systems are deemed. Nevertheless, it is difficult to recognize randomly oriented text in images. Thus, an effective TDR framework from NSI is presented by the work utilizing AS-CNN for the TR model. The proposed work encompasses 2 major phases, such as TD and TR. The images are given as input to the preprocessing step. The image is resized through preprocessing, eliminating unnecessary noise from the image. Next, the image's contrast is improved. Utilizing the improved EAST (IEAST) algorithm, the text is identified from the enhanced image. Utilizing the impoved GrabCut algorithm, the background region of the image is eliminated so that a clear view of the text can be attained. Then, the extraction of significant features is done. Using the LS-TSA technique, the most needed features are chosen from the extracted features. As the final step, the classifier called AS-CNN accepts the selected features as input. The precise text from images is significantly extracted and detected by the AS-CNN. The architecture of the proposed work is depicted in Fig. 1.

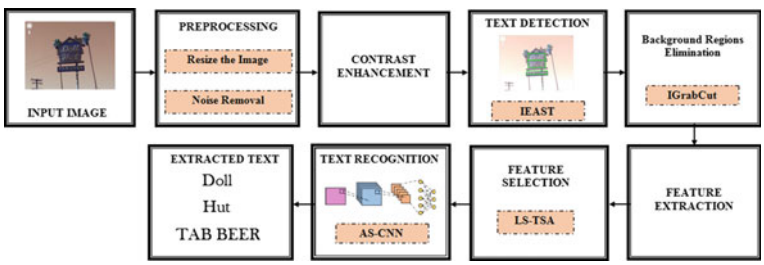


Fig. 1 Architecture of the proposed work

3.1 Input Source

The input textual images are collected from the publicly accessible SVT dataset on the Internet. The SVT dataset exhibits the high-variability textual images with low-resolution rates. For the proposed work's analysis, the input images as of the SVT dataset are employed.

3.2 Preprocessing

The input images are subject to the preprocessing phase in the initial step. The images are processed by the preprocessing function. The image's quality is enhanced by preprocessing via removing the unnecessary distortion and also ameliorating the image's vital features. Image resizing and noise removal are the '2' major steps in preprocessing phase. The preprocessing function is mathematically expressed as

$$\lambda = P_{pf}[\text{Input}_{img}] \tag{1}$$

wherein implies the outcome of the preprocessing function, indicates preprocessing function, and denotes the input images.

$$P_{pf} = \{\lambda_{ir}, \lambda_{nr}\} \tag{2}$$

where λ_{ir} symbolizes the image resizing function, and λ_{nr} indicates the noise removal function.

(a) **Resize the image**

The process of modifying the image's size by augmenting or reducing the number of pixels is called image resizing. Detailed information concerning the image is offered by the image resizing process. The images are resized into 256 * 256 pixels in the proposed work.

(b) **Noise Removal**

The noise present in the image is caused by random variation and fluctuation in colors. Such noises suppress the image's quality. The noise removal function considerably eliminates the undesired noise from images and retains the image's quality. The noise removal function is proffered as

$$N_{\text{rem}} = \lambda_{nr}\{\text{Input}_{\text{img}}\} \tag{3}$$

N_{rem} implies the outcome of the noise removal function and $\lambda_{nr}\{\text{Input}_{\text{img}}\}$ signifies the input image's noise removal function. Thus, the final outcomes of the preprocessed function λ are mathematically specified by

$$\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n) \tag{4}$$

Here, $\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n$ implies the preprocessed images.

3.3 Contrast Enhancement

The processed images (λ) are subjected to the contrast enhancement (CE) phase after completing the preprocessing step. The CE phase highlights the image's features via modifying the color. Hence, it is simple to detect the minute and hidden information behind the images. The CE function of the input image is mathematically denoted as

$$CE_i = c_{en}\{\lambda\} \tag{5}$$

where

$$CE_i = \{CE_1, CE_2, CE_3, \ldots, CE_n\} \tag{6}$$

where CE_i signifies the contrast-enhanced function's outcome that means contrast-enhanced images, c_{en} indicates the CE function, λ denotes the preprocessed images, and $CE_1, CE_2, CE_3, \ldots, CE_n$ symbolizes the contrast-enhanced images.

3.4 Text Detection

In the TD phase, the text as of the enhanced image is grouped after the image's contrast is enhanced. For efficient TD, an algorithm named IEAST is utilized by the proposed method that automatically reads the text from the image.

3.4.1 Text Detection Using Improved Efficient and Accurate Scene Text Detector

The efficient and accurate scene text detector (EAST) is a quick and efficient TD algorithm. For localizing the word or text from the images, a fully convolutional network is utilized by the EAST. Normally, the sigmoid activation function is utilized by the EAST algorithm in the output layer for balancing the loss function (LF) that produces major drawbacks, namely gradient saturation, slow convergence, along with sharp damp gradient problems. Thus, the hyperbolic tangent function (Tanh) is employed in the IEAST algorithm. Hence, the probability of TR as of the image is improved by the proposed IEAST, and it also balances the LF. The steps that are incorporated in the IEAST algorithm are broadly elucidated as

(a) **Feature Extraction Stage**

The enhanced image is attained by the feature extraction (FE) stage, and that image is passed into a set of convolution layers named feature extractor stem. It processes the images, and the images are converted into feature maps (FM). This is mathematically specified as

$$\beta_{fes} = F_e^*\{CE_i\} \tag{7}$$

where β_{fes} implies the outcome of the feature extractor stem, F_e^* denotes the FE function, and CE_i indicates the input images.

$$CE_i^1 = \{u_1, u_2, u_3, \ldots, u_n\} \tag{8}$$

$$CE_i^2 = \{v_1, v_2, v_3, \ldots, v_n\} \tag{9}$$

wherein $\{u_1, u_2, u_3, \ldots, u_n\}$ and $\{v_1, v_2, v_3, \ldots, v_n\}$ signify the FM.

(b) **Feature Fusion Stage**

Then, the FM is unpooled. After that, using several features from disparate spatial resolutions, those FMs are concatenated. For predicting smaller word regions, this merging feature is helpful. The feature merging operation is presented by

$$u_i = \begin{cases} \text{unpool}(v_i) & \text{if } i \leq 3 \\ \text{conv}_{3\times3}(v_i) & \text{if } i = 4 \end{cases} \tag{10}$$

$$v_i = \begin{cases} f_i & \text{if } i = 1 \\ \text{conv}_{3\times3}(\text{conv}_{1\times1}([u_{i-1};\, f_i])) & \text{otherwise} \end{cases} \tag{11}$$

(c) **Prediction Network**

A score map and a geometry map are encompassed by the prediction network, which is the output layer. The score map reveals the probability of text in the region, while the geometry map defines the text box's boundary.

In the EAST algorithm, the LF is computed that comprises both the score map loss and geometry LF. The LF is estimated as

$$L_l = L_s + t_g L_g \tag{12}$$

where L_s indicates the score map loss, L_g implies the geometry LF, and t_g denotes the importance between '2' losses.

For balancing the LF, a hyperbolic tangent function (Tanh) is utilized by the work that enhances the probability of TD. The Tanh function is signified as

$$\tanh x = \frac{\sinh x}{\cosh x} \tag{13}$$

$$\frac{\sinh x}{\cosh x} = \frac{w^z - w^{-z}}{w^z + w^{-z}} \tag{14}$$

$$\frac{w^z - w^{-z}}{w^z + w^{-z}} = \frac{w^{2_z} - 1}{w^{2_z} + 1} \tag{15}$$

At last, utilizing the IEAST algorithm, the texts as of the images are identified. The text detected images are indicated as

$$Td = \{Td_1, Td_2, Td_3, \ldots, Td_n\} \tag{16}$$

3.5 Background Region Removal

The background regions are removed after the identification of text from the images. Thus, from the images, the unnecessary regions are eliminated. Thus, the needed text in an image is highlighted by this process. The IGrabCut segmentation algorithm is utilized by the proposed work. For image segmentation, the Gaussian mixture model (GMM) is employed in the GrabCut algorithm. However, the major parts of the images are segmented by GMM so there is a chance for the removal of significant regions from the image. Thus, the categorical mixture model (CMM) is utilized in this work. Concerning enthalpy energy, the energy is computed that reduces the energy usage. Thus, for solving the minimum cutting that effectively differentiates the foreground and background, the IGrabCut is utilized. The steps of the IGrabCut algorithm are briefly described below.

Step 1: The text detected image Td is deemed, and the initial tri-map G_m is created by choosing a rectangle. The pixels inside the rectangle are indicated as foreground $G_m(Fg)$, and the pixels outside the rectangle are denoted as background $G_m(Bg)$.

Step 2: The array of pixels could be represented for a text detected image Td as

$$h = \{h_1, h_2, h_3 \ldots, h_n\} \tag{17}$$

where

$$h_i = \{R_i, G_i, B_i\}, \ \ i \in [1, 2, 3, \ldots, n] \tag{18}$$

in RGB space.

Step 3: After that, the image segmentation process is performed. Here, for indicating whether the image belongs to the foreground or background, the labels are allocated

to every pixel. The image segmentation process is specified as

$$X = (X_1, X_2, X_3, \ldots, X_n),\ X_n \in \{0, 1\} \tag{19}$$

Here, $X_n = 0$ denotes the background, along with $X_n = 1$ indicates the image's foreground.

Step 4: For the initial foreground along with background regions, the CMM is formed. It could be mathematically denoted as

$$M(Y_i|\varphi) = \sum_{u}^{U} a_u m(Y_i; \theta_u) \tag{20}$$

where $\{a_u\}_{u=1}^{U}$ denotes the mixing terms, $\theta_u = (\theta_{u,1,\ldots,}\theta_{u,Z})$ indicates the probability distribution over the label Z, and $m(Y_i; \theta_u)$ specifies the categorical distribution on Z using θ_u.

Step 5: Concerning enthalpy energy, the energy function for the segmentation could be computed. The energy consumption of the model is efficiently reduced by enthalpy energy. It is assessed as

$$\text{Enthalpy} = E_r + p_r v_o \tag{21}$$

where enthalpy indicates the enthalpy energy which is equal to the addition of internal energy E_r and the product of pressure p_r along with v_o the volume of the model.

Step 6: By considering the neighborhood K around each pixel, the segmented regions must be coherent concerning color, which is assumed by the regularizing prior R. It could be specified as

$$R(X, h) = \Theta \sum_{\{m,n\} \in G} [X_n \neq X_m] \exp(-B\|h_m - h_m\|^2) \tag{22}$$

where B is a constant.

Step 7: Finally, a graph is created. A graph cut is executed to find a new tentative foreground and background of the image by the classification of pixels. Thus, the mathematical expressions for the background eliminated images $B(r)$ are presented by

$$B(r) = B(r)_1, B(r)_2, B(r)_3, \ldots, B(r)_n \tag{23}$$

3.6 Feature Extraction

From the image features, namely color f_c, luminance f_l, shape feature f_s, stroke pixel f_{sp}, edge f_e, texture f_t, horizontal f_h, vertical f_v, curved f_{cr}, rotated f_{rt}, and twist f_t, the most significant and selective features are extracted after the removal of background as of the image. New features, which are a linear amalgamation of the prevailing features, are provided by the feature extraction process. When analogized with the actual feature values, the varied values are possessed by the new set of features. To capture the same data, only fewer features are required. The computational complexity and time complexity are greatly decreased, and the FS accuracy is enhanced owing to the extraction of vital features. New features, which are a linear amalgamation of the prevailing features, are provided by the feature extraction process. When analogized with the actual feature values, the varied values are possessed by the new set of features. To capture the same data, only fewer features are required. The FE of background removed images $f_{fx}(B(r))$ is mathematically produced as

$$f_{fx}(B(r)) = \{f_c, f_l, f_s, f_{sp}, f_e, f_t, f_h, f_v, f_{cr}, f_{rt}, f_t\} \tag{24}$$

3.7 Feature Selection

In this phase, the most needed features which contribute to the TD process are chosen from the extracted features. Hence, the complexities of irrelevant features are reduced by the FS phase, and it augments the model's accuracy by filtering the relevant features. Utilizing linear scaling in TSA, the necessary features are chosen in the proposed method. Utilizing the linear scaling technique, the social forces between the search agents are computed in the tunicate swarm algorithm (TSA). Therefore, the distraction distance between the search agents is improved by the LS-TSA technique. The LS-TSA's steps are listed as

Step 1: Initialization

The truncate population is initialized in the initial step, where $v = 1, 2, 3, \ldots, c$ and $s = 1, 2, 3, \ldots, b$ such that set of key values, which are initialized arbitrarily.

Step 2: Fitness Evaluation

The fitness value for every input variable is assessed after the initialization of a set of key values. Then, the best one is selected by the fitness evaluation step. The fitness could be assessed as

$$fi(\alpha_{vs}) = M_x(\text{PSNR}) \tag{25}$$

$$fi(\alpha_{vs}) = M_x\left[10\log_{10}\left(\frac{255^2}{\sum_{A,B}(\mathrm{I}(a,b) - I_{\mathrm{dec}}(a,b))/A \times B}\right)\right] \tag{26}$$

where $I(a,b)$ indicates the plain images,$I_{\mathrm{dec}}(a,b)$ denotes the decrypted images, and (a,b) indicates the image's row and column. If $PSNR \geq Th_d$, then the existing solution is updated. After that, the maximal fitness value is maintained by the algorithm.

Step 3: The conflicts between search agents are avoided utilizing the vector, $\vec{W}$ the currently updated search agent position is estimated for avoiding the conflicts betwixt the other search agents (other tunicates) centered on the following expression.

$$\vec{W} = \frac{\vec{Z}}{\vec{\delta}} \tag{27}$$

where $\vec{Z}$ indicates the gravity force and $\vec{\delta}$ denotes the social forces betwixt the search agents and that could be denoted as

$$\vec{Z} = \nu_2 + \nu_3 + Q \tag{28}$$

$$Q = 2 \cdot \nu_1 \tag{29}$$

Here, Q indicates the water flow advection in the deep ocean, and ν_1, ν_2, and ν_3 signify the random number in the interval.

Furthermore, for improving distraction distance, the social forces betwixt the search agents could be assessed by utilizing the linear scaling technique. It is presented by

$$L(s) = \frac{S(x) - \min S(x)}{\max S(x) - \min S(x)} \tag{30}$$

where $S(x)$ implies the search agents.

Step 4: Explore the best nearest neighbor. Toward the direction of the best neighbor, the search agents are moved. The distance betwixt the food source and search agent is measured for finding the best nearest neighbor.

$$\vec{D} = \left|\underset{c}{\vec{N}} - \nu.\vec{N}(t)\right| \tag{31}$$

where $\vec{N}(t)$ denotes the current position of the tunicate at t^{th} iteration,$\vec{N}_c$ indicates the position of food source, and ν denotes the arbitrary number in the interval.

Step 5: Converge toward best search agent.

The search agents converge toward the best search agent's position after the exploration of the best nearest neighbor. Therefore, the current position of the tunicate $\vec{N}_p(t)$ is updated and that is denoted as

$$\underset{p}{\vec{N}}(t) = \begin{cases} \vec{N}_d + \vec{W} \cdot \vec{D}, \text{ if } \nu \geq 0.5 \\ \vec{N}_d - \vec{W} \cdot \vec{D}, \text{ otherwise} \end{cases} \tag{32}$$

where $\vec{N}_p(t)$ implies the updated position of tunicate concerning the position of food source $\vec{N}_d$.

Step 6: Implement swarm behavior.

Here, the first '2' best solutions are saved. Next, as per the best search agent's position, the other search agents are updated for representing the tunicate swarm behavior. It is defined as

$$\vec{N}_p(t+1) = \frac{\vec{N}(t) + \vec{N}_p(t)}{2 + \nu_1} \tag{33}$$

Step 7: Termination.

Until the best solutions are determined, the above-mentioned steps are repeated. Concerning the fitness value, the produced keys are assessed at every iteration, and those keys are updated centered on the preceding best keys. Thus, the selected feature as of the images α_s^f are mathematically denoted as

$$\alpha_s^f = \{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n\} \tag{34}$$

3.8 Text Recognition

The chosen features are offered to the recognition phase after the selection of features. Here, utilizing an AS-CNN algorithm, the texts from the images are identified. For performing the classification using CNN, this algorithm is mostly utilized. '5' major layers, like input layer, convolution layer, pooling layer, fully connected layer (FCL), and output layer are encompassed by CNN. For optimizing its weights, the Anopheles search optimization algorithm (ASOA) is utilized by CNN. For optimizing the weight value between the layers to decrease the training time and enhance the accuracy, ASOA is utilized. In CNN, the LF mostly plays a major role in the updation process. The solution is deemed as the optimum if the LF is minimum. The weights are updated by utilizing AS-CNN if the LF is maximum (Fig. 2).

The meta-heuristic algorithm is the ASOA algorithm which is inspired by the behavior of female Anopheles mosquitoes in transmitting the malaria disease. Thus

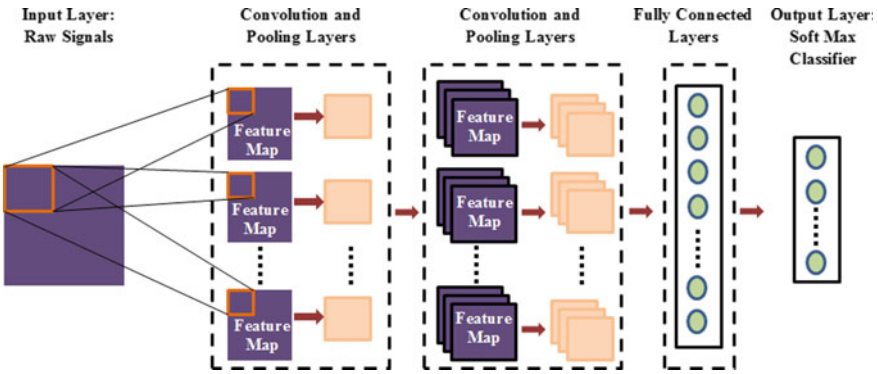


Fig. 2 General structure of the CNN

the complexity in training time is mitigated by ASOA and thereby it enhances the reliability of classification. By implementing metahueristic algorithms, the accuracy of CNN could be optimized. Those problems that could not be solved effectively with deterministic models within a resonable time, and accuracy could be solved effectively with metahueristic approaches. These algorithms are well suited for solving large complex problems within a resonable time.

The steps of the ASOA are given as

(a) Initialization

The populace of Anopheles mosquitoes is initialized. Here, the mosquitoes indicate the weight values, which could be mathematically signified

$$A \in \delta_n,\ n = 1, 2, 3, \ldots, N \tag{37}$$

$$\max_{\delta} f(A) \tag{38}$$

where $f(A)$ implies the objective function, A denotes the set of decision variables, δ_n indicates the set of possible values for every decision variable, and N implies the number of decision variables. The decision variables are defined by utilizing the upper and lower bounds if they are continuous. It is denoted as

$$Lb(\delta_n) \leq \delta_n \leq Ub(\delta_n) \tag{39}$$

$$\delta_n \in \{\delta_1, \delta_2, \delta_3, \ldots, \delta_n\} \tag{40}$$

(b) Evaluate fitness value for each Anopheles

The fitness values for every anopheles mosquito from the search space are calculated. The Anopheles with the best fitness value is chosen as the optimum point. Therefore,

the fitness could be assessed as

$$O(x_i\delta_n) = \frac{1}{mV_i\delta_n} * \log(f(\delta_n) + p \tag{41}$$

$$p = 0 \leq p \leq 0.5 \tag{42}$$

(c) Calculate the distance from the best Anopheles

The optimality of every point is defined by the mosquitoes in a random moving space. After some time, the odor density attracts more mosquitoes. It could be specified as

$$m(x_i\delta_n) = \sqrt{\sum_{i=1}^{n} (\delta_i(x_i) - \delta_i(\delta_i))} \tag{43}$$

(d) Update position of each Anopheles

Then, the odor density is sensed by each mosquito, and it moves toward them. Then, centered on the value, the mosquito's position is updated that defined as

$$O(m_i\delta_n) = \frac{1}{A_{m(\lambda n)} - \delta_n + (z_{\lambda n} - \delta z_n)_2} * \log((f(\delta_n)) + p \tag{44}$$

$$p = 0 \leq p \leq 0.5 \tag{45}$$

(e) Evaluate fitness of each Anopheles and update the best fitness

Until the best solution is attained, the iteration continues. The stopping criteria of the algorithm are presented as

$$A_{new} = A_{old} + O(m_i\delta_n) * |\delta_{\lambda n} - \delta_n| \tag{46}$$

$$Z_{new} = B_{old} + O(m_i\delta_n) * |z_{\lambda n} - z_n| \tag{47}$$

Finally, the AS-CNN framework effectively identifies the text, and it offers accurate text as of the images. Figure 3 depicts the pseudocode for the AS-CNN.

4 Result and Discussion

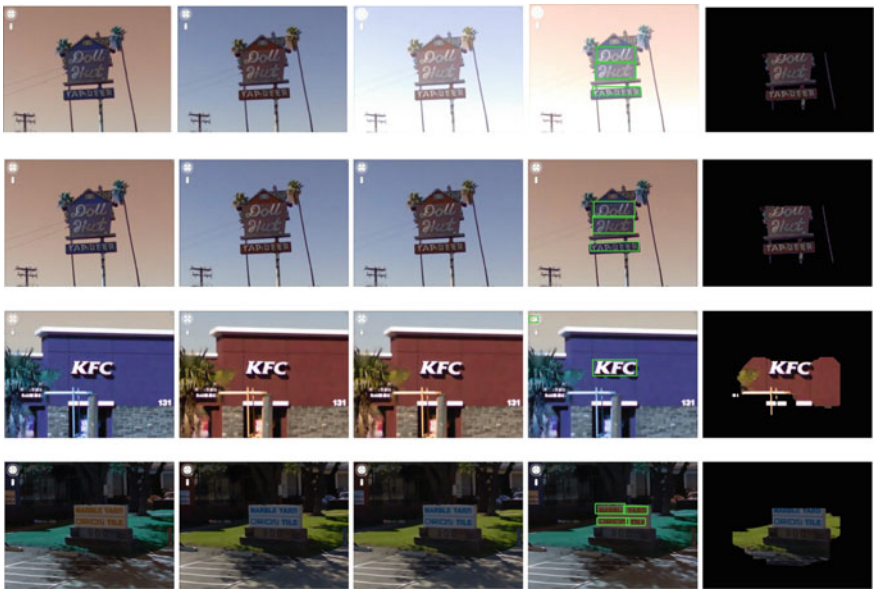
Centered on different performance metrics, the proposed work's final outcome is examined in this section. Here, for stating the work's effectiveness, the performance

Input: Selected features from the images
Output: Extraction and recognition of text from the images
Begin
Initialize weight $Wg_{i,j}$, bias vector $Ba_{i,j}$, image features α, $\max_{itr}$
For $\max_{itr} = (1 to n)$
Evaluate the convolution operation using,
$$r_j(\sum_{i=0}^{N}\sum_{j=0}^{N} Wg_{i,j}\alpha_{a+i,b+i} + Ba_{i,j})$$
If $\alpha_{i,j}(s_z) == \zeta_{con}(s_z)$
Perform padding process
Else
Generate the padding using,
$$(\alpha + 2k - d + 1) \times (\alpha + 2k - d + 1)$$
End if
Evaluate the pooling layer by computing,
$$G_k^*(|\alpha_{i,j}|)$$
Evaluate the fully connected layer by computing,
$$r_j(\sum_{i=1}^{n} Wg_{i,j}\alpha_{i,j} + Ba_{i,j})$$
If $(\alpha_i = \bar{\alpha}_i)$
Determine the exact text
Else
Perform back propagation by updating weights using,
$$A_{old} + O(m_i\delta_n) * |\delta_{\lambda n} - \delta_n|$$
$$B_{old} + O(m_i\delta_n) * |z_{\lambda n} - z_n|$$
End if
End for
End

Fig. 3 Pseudocode for the AS-CNN algorithm

analysis and the comparative analysis are performed. The proposed work is applied in Python. From the SVT dataset, which is publically accessible on the Internet, the images are collected. Street View Text (SVT) dataset includes images gathered as of Google Street View, where every single image is annotated with bounding boxes around words from businesses around where the image was obtained. 350 total images (as of '20' varied cities) along with 725 total labeled words are encompassed in the dataset. From Google Street View, the SVT dataset was obtained. A higher variability along with lower resolution was exhibited by the image text in this data. '2' characteristics were noticed while handling the outdoor street-level imagery. They are (i) image text constantly arrives as of business signage and (ii) business names are effortlessly obtainable via geographic business searches. These factors make the SVT set uniquely fitted for word spotting in the wild: provided a street view image, the intention is to recognize words from nearby businesses. The dataset contains 350 images with 20 cities and 725 labeled words. The hyper-parameters values of the neural network are given as follows:

- Initial learning rate: 0.00125
- Drop: 0.5



(a) (b) (c) (d) (e)

Fig. 4 Output image of the various phases of TR

- Epochs drop: 10.0
- Number of epochs: 30
- Running loss: 0.0
- Weight range: 0–1

Figure 4 exhibits the result attained by every phase of the TR process that is elucidated in the proposed work.

The final result of the disparate phases of TR is depicted in Fig. 4. Furthermore, the text under unsure situations, namely poor lighting, improper posture, etc., could be efficiently recognized through the proposed work.

4.1 Performance Analysis

Concerning different performance metrics, namely accuracy, sensitivity, specificity, precision, recall, F-measure, false-positive rates (FPR), false-negative rates (FNR), and Matthews correlation coefficient (MCC), the proposed AS-CNN is assessed with different prevailing techniques, like recurrent neural network (RNN), deep neural network (DNN), artificial neural network (ANN), and CNN. The ANN model used had a learing rate of 0.1 with 1000 epochs and 2 hidden layers. RNN had a learning rate of 0.1 with an input layer size of 2 and hidden layer size of 16. The CNN

Table 1 Performance analysis of proposed AS-CNN based on accuracy, sensitivity, and specificity

Techniques	Accuracy	Sensitivity	Specificity
RNN	86.85	84.28	83.12
DNN	89.63	87.42	86.29
ANN	92.62	89.71	88.78
CNN	94.62	92.57	92.37
Proposed AS-CNN	95.65	95.71	95.6

used sigmoid, ReLU, and softmax activations at different layers and DNN had input size of 4, output layer size of 3, 50 epochs and 20 neurons at hidden layers with a learning rate of 0.001. Precision and recall rates are useful measures which helps to determine the accuracy of the proposed method to recognize correct text characters from detected text regions. Precision, recall rates, F-score, and accuracy rates of the proposed model are analyzed with respect to the metrics related to state-of-the-art deep learning architectures. To state the model's effectiveness and robustness, the comparative analysis is also performed with these existent techniques.

Concerning accuracy, sensitivity, and specificity, the performance analysis of the proposed AS-CNN with different existing techniques, like RNN, DNN, ANN, and CNN, is exhibited in Table 1. It is known from the tabulation that accuracy of 96.32%, the sensitivity of 95.71%, and also the specificity of 95.6% is obtained by the AS-CNN. The accuracy rate that lies in overall range between 86.85%-94.62%, sensitivity between 84.28%–92.57%, and specificity between 83.12%–92.37% is attained by the existent techniques, like RNN, DNN, ANN, and CNN. Thus, better accuracy, sensitivity, along with specificity rates are acquired by the proposed AS-CNN. This shows that the improvements made by the methods like IEAST, IGrabCut, and TSA are very effective in the proposed work. Thus, the proposed approach recognizes the accurate text from the images. Figure 5 offers a clear view of Table 1.

The comparative analysis of the proposed work with different prevailing techniques is stated in Fig. 5. The model must possess a low-error rate to be more efficient while training and testing. As per that, the proposed AS-CNN produces 96.32% of accuracy, 95.71% of sensitivity, and 83.12% of specificity. However, the existing methods possess a lower metrics rate with an average of 90.93% accuracy, 88.49% sensitivity, and 87.67% specificity. The proposed model highlights higher efficiency with lower error rate.

Centered on precision, recall, along with F-measure, the contrast of the proposed AS-CNN with different prevailing works is exhibited in Fig. 6. The comparative study clearly shows that the precision, recall, and F-measure at the rate of 96.32%, 95.85%, and 96.32%, correspondingly, are attained by the AS-CNN. However, precision rates that overall range between 89.32%–94.32%, recall that ranges between 90.32%–95.32%, and F-measure value that ranges between 89.62%–94.84% are acquired by the existent works, namely RNN, DNN, ANN, and CNN. Hence, the proposed work acquires a better metric rate when analogized to the existent work. This demonstrates that approaches such as IEAST, IGrabCut, and TSA are quite effective in the proposed

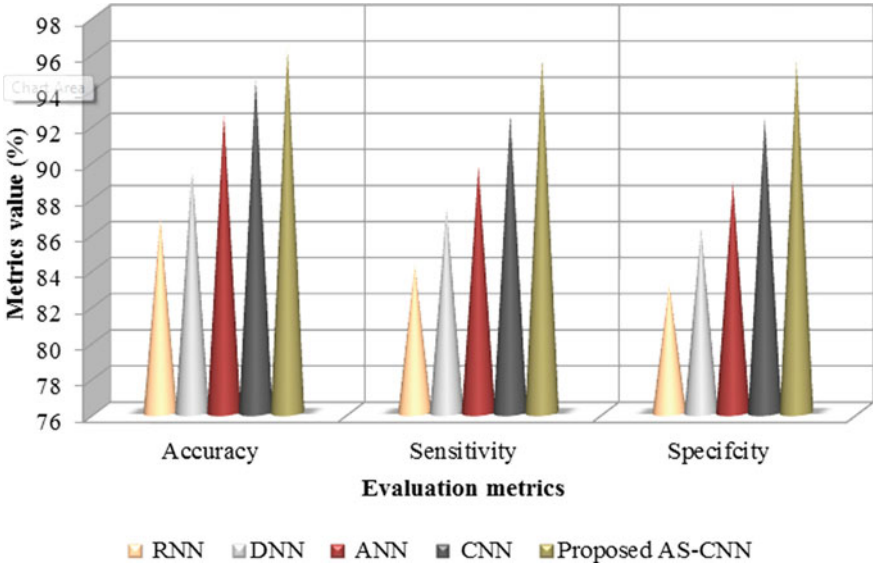


Fig. 5 Graphical representation of the proposed AS-CNN based on accuracy, sensitivity, and specificity

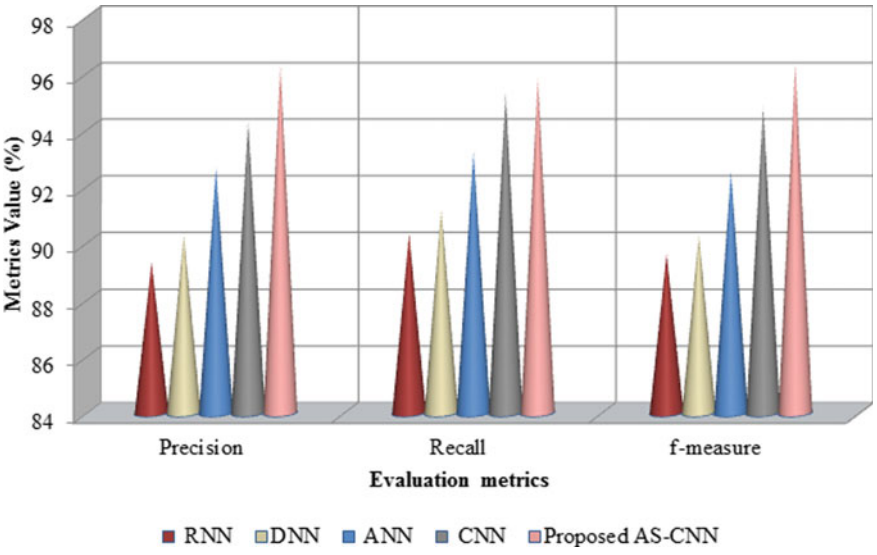


Fig. 6 Graphical representation of the proposed AS-CNN based on precision, recall, and F-measure

Table 2 Performance analysis of proposed AS-CNN based on FPR, FNR, and MCC

Techniques	FPR	FNR	MCC
RNN	0.16	0.15	67.41
DNN	0.13	0.12	73.72
ANN	0.11	0.1	78.49
CNN	0.07	0.07	84.94
Proposed AS-CNN	0.04	0.04	91.31

work. Therefore, the complexities are mitigated via the proposed one, and also it improves the classification accuracy.

Concerning the FPR along with FNR, the proposed AS-CNN's effectiveness is depicted in Table 2. The model's reliability could be attained via the low-false prediction rates. As per that, when contrasted to the existent techniques, the FPR and also the FNR rates of the proposed work are lower. 0.04% of FPR along with 0.04% of FNR is accomplished by the AS-CNN. The overall FPR rate ranges between 0.07%–0.16%, and FNR rates range between 0.1%–0.15% is acquired by the prevailing techniques, like RNN, DNN, ANN, and CNN. Furthermore, concerning the MCC metric, the proposed work is also validated. In converse to the false prediction rates, the model's robustness is demonstrated by the higher value of MCC. 91.31% of MCC is attained by the AS-CNN. However, the MCC rates that overall range between 67.41%–84.94% are attained by the prevailing works. Hence, while training the images, the unusual situation is efficiently handled by the proposed approach, and it also exhibits promising results. These results demonstrate that the proposed framework significantly improves text detection and recognition performance, for text in real scenes.

5 Conclusion

In computer vision-related applications, one of the vital task is information retrieval from images. Extraction of text content from images becomes a challenging task when we deal with natural scene images. Natural scene images may often appear blurred, noisy, and obscurred.

Developing an efficient framework for accurate detection, extraction and recognition of text from scene images would be promising. Utilizing AS-CNN, an effective TDR frameworkfrom NSI is proposed by the work. Then, the experimental analysis is executed. Here, for authenticating the proposed algorithm's effectiveness, the performance and also the comparative analysis of the proposed and existent techniques are performed concerning a few performance metrics. The developed approach could handle different uncertainties and extract the exact text as of the images. For the analysis, the publicly accessible datasets are utilized in which 96.32% of accuracy, 95.71% of sensitivity, and 95.6% of specificity are acquired by the proposed work. Concerning high-accuracy rates, the prevailing top-notch methods are outperformed

by the proposed TDR framework, and it also stays more dependable. As future enhancement, advanced algorithms and neural networks will be used for TDR in images and video.

References

1. Gao X, Han S, Luo C (2019) A detection and verification model based on ssd and encoder-decoder network for scene text detection. IEEE Access 7:71299–71310
2. Wu L, Shivakumara P, Lu T, Tan CL (2015) A new technique for multi-oriented scene text line detection and tracking in video. IEEE Trans Multimedia 17(8):1137–1152
3. Cao L, Li H, Xie R, Zhu J (2020) A text detection algorithm for image of student exercises based on CTPN and enhanced YOLOV3. IEEE Access 8:176924–176934
4. Yao C, Bai X, Liu W (2014) A unified framework for multioriented text detection and recognition. IEEE Trans Image Process 23(11):4737–4749
5. Houssein EH et al (2021) An improved tunicate swarm algorithm for global optimization and image segmentation. IEEE Access 9:56066–56092
6. Roy S, Shivakumara P, Pal U, Lu T, Kumar GH (2019) Delaunay triangulation based text detection from multi-view images of natural scene. Pattern Recogn Lett.https://doi.org/10.1016/j.patrec.2019.11.021
7. Zhonga Z, Sun L, Huo Q (2019) Improved localization accuracy by LocNet for faster R-CNN based text detection in natural scene images. Pattern Recogn 96:1–16
8. Kantipudi MVV, Kumar S, Jha AK (2021) Scene text recognition based on bidirectional LSTM and deep neural network. Comput Intell Neurosci
9. Dai P et al (2021) Accurate scene text detection via scale-aware data augmentation and shape similarity constraint. IEEE Trans Multimedia
10. Butt H et al (2021) Attention-based CNN-RNN arabic text recognition from natural scene images. Forecasting 3.3:520–540
11. Ma M et al (2021) Residual attention-based multi-scale script identification in scene text images. Neurocomputing 421:222–233
12. Naiemi F, Ghods V, Khalesi H (2021) A novel pipeline framework for multi oriented scene text image detection and recognition. Expert Syst Appl 170:114549
13. Wang Y, Shi C, Xiao B, Wang C, Qi C (2018) CRF based text detection for natural scene images using convolutional neural network and context information. Neurocomputing 295:46–58
14. Shivakumara P, Raghavendra R, Qin L, Raja KB, Lu T, Pal U (2017) A new multi-modal approach to bib number/text detection and recognition in Marathon images. Pattern Recogn 61:479–791
15. Xu J, Ding W, Zhao H (2019) Based on improved edge detection algorithm for English text extraction and restoration from color images. IEEE Sens J. https://doi.org/10.1109/JSEN.2020.2964939

Extracting Multi-Language Text from Video into Editable Form

V. Mani, Mohammed Ismail, K. Navinkumar, and P. Kathirnirmal

Abstract For semantic video retrieval (from archives), alert production (live streams), and large levels applications like assessment mining and content reiterating, textual material in videos is an appealing index. The discovery and recognition of textual content are essential factors of similar systems, and this is the focus of our research. This case study talks about all types of the framework for text recognition and segmentation in video frames. We focus on cursive scripts, in particular, using English text as an example. Fine-tuning deep neural network-based item finders for the particular instance of text identification is used to recognize textual patches in video frames. Convolutional neural networks (CNNs) are used to identify the script of the discovered textual material, while an EngNet, a hybrid of CNNs, and long short-term memory (LSTM) (Rajesh Kanna and Santhi in Expert Syst Appl 194, 2022 [1]) systems is proposed for identification. Also being generated is a benchmark dataset with cursive writing and over 13,000 video frames. An F-proportion of 88.3% for identification and an F-proportion of 87% for acknowledgment was later a total arrangement of preliminaries.

Keywords Text recognition · Convolutional neural network (CNN) · Deep learning (DL) · Caption text · Text detection

1 Introduction

The total of digital multi-media information, particularly video material, consumes grown dramatically in recent years, together with the method of video records and live streaming. According to the statistics [2], every single minute people are uploading 300 h of videos on YouTube. The availability of cheap smartphones with cameras is a major contributor to this massive surge. With such large datasets, efficient and effective retrieval strategies are required to allow users to find the information they

V. Mani (✉) · M. Ismail · K. Navinkumar · P. Kathirnirmal
Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur 639113, India
e-mail: vmanimkce663@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_40

want generally, user-assigned comments or keywords, known as tags, have been used to store movies. When searching for material, a term entered as a query is coordinated with these labels to get the most related results. As a result, the given tags are unable to include the rich video information, resulting in a limited retrieval. Instead of just matching tags, the aim proved an extra successful technique is to examine inside the definite material, content-based video recovery. CBVR structures have been read up and set up for a broadened period and license for a wiser recovery of required material. The "content" is related to the visual text (such as image frame), audio material (such as spoken terms), or written information. Amongst these, textual material is the focus of our present research. We are aiming for a clever retrieval structure that uses the textual material in films as a catalog.

Scene and caption text are the two major categories of textual materials seen in video frames. During the video recording process, visible text (Fig. 1) is acquired by the camera and may or may not be connected with the content. Ad banners, sign sheets, and lettering on T-shirt are all examples of scene text. Scene text is often used in applications such as robotic navigation and visually impaired technology. On the other hand, artificial or caption text (Fig. 2) is placed on the video.

Inscription text is by and large connected with video content and is utilized to recover semantic data from recordings. The inscription text is likewise the subject of our momentum research. Text location detection, text extraction (background split), script identification (in multi-text movies), and finally text recognition are key components based on text content and retrieval systems (video OCR). Unsupervised [3–5], supervised [6–8], and half and half [9, 10] strategies can be utilized to recognize text. To recognize text and non-text areas, unsupervised text recovery uses picture examination procedures. Supervised approaches, on other hand, entail teaching a working algorithm to distinguish between textual and non-textual regions using examples of both. In certain circumstances, a mix of the two techniques is used, with unsupervised algorithms identifying suitable text sections that are then confirmed using a supervised approach.

Fig. 1 Example of scene text

Fig. 2 Example for caution text

The text can then be sent into the recognition engine after it has been discovered. An additional module to recognize the content is important in the case of motion pictures with text in different contexts so that each sort of text-based might be distinguished utilizing its OCR. There are progressed acknowledgment calculations for text in the Roman content [11], and cursive script recognition shows a difficult problem.

In addition, unlike document pictures, which are examined at superior grade, video text is often low contraction (Fig. 3) and might show a variety of backdrops, making detection more difficult.

In a multi-script context, this research proposes a complete architecture for image frame text detection and identification. The following are the study's most important findings.

- Examination of cutting-edge deep learning-based object detectors, with a focus on textual content identification.
- Convolutional neural networks are used to identify video text scripts (CNN).
- Using a mix of CNN and LSTM, we were able to recognize cursive (English) video text.

Fig. 3 Pictorial representation of min-resolution manuscript video frame

- A complete set of tests is used to validate the given methods.

2 Literature Survey

Table 1 illustrates the comparison of literature paper.

3 Related Works

Textual content detection and recognition in movies, photos, forms, and regular sceneries have been studied for almost 40 years. Over time, the area has progressed from simple systems that recognize isolated numbers and fonts to complicated end-to-end structures that can read the text in real settings. This unit gives a summary of noteworthy aids to textual content recognition and identification in photos and videos. Paper [11–15] described many comprehensive and extensive techniques of the problem. Our debate is divided into three pieces. The text detection problem is discussed first, monitored by an outline of cursive recognition strategies.

3.1 Text Detection

Text identification is the process of locating textual content in photographs. Unsupervised and supervised approaches are commonly used to categorize text detection techniques. Unsupervised approaches use image classification methods to separate text from related, while managed techniques utilize learning methods to recognize among text and non-text areas.

Edge-based techniques [3, 4, 16] that (expect and) control the difference among text and its experience; techniques connected component-based [17, 18] depending essentially concerning the shading/force of message pixels; and subtitle-based techniques [19, 20] that treat message content in the picture as a special surface that separates itself from non-message regions. Gabor Remover [21], wavelets [22], curvelets [23], local binary patterns (LBP) [24], discrete cosine transformation (DCT) [25], histograms of oriented gradients (HoG) [26], and Fourier change have all been researched color-based approaches [27, 28], which are comparable to component-based methods in many ways and use shading data of text pixels to recognize it from non-text regions and are another prevalent category of techniques.

Typically, supervised textual content detection systems employ cutting-edge deep learning that is prepared on examples of text and non-text blocks, either utilizing pixel esteems or by extracting significant attributes first. Many classification techniques have been researched over the years, including Naive Bayes and support vector

Table 1 Comparison of literature survey

S.no	Authors and year	Title of the paper	Algorithm/methodologies	Findings
1	A. Mirza, I. Siddiqi (2020)	Recognition of cursive video text using a deep learning framework	Cursive text recognition appearing in videos, using a complete framework of deep neural networks (deep learning algorithm)	Text recognition
2	T. He, Z. Tian, W. Huang, C. Shen, Y. Qiao, C. Sun (2018)	Single shot TextSpotter with explicit alignment and attention	RNN algorithm	We propose a novel text-alignment layer that allows it to precisely compute convolutional features of a text instance in arbitrary orientation, which is the key to boost the performance
3	A. K. Bhunia, A. Konwer, A. K. Bhunia, A. (2019)	Script identification in natural scene image and video frames using an attention-based convolutional-LSTM network	LSTM, CNN algorithm	Script detect
4	M. Liao, J. Zhang, Z. Wan, F. Xie, J. Liang, P. Lyu, C. Yao, X. Bai (2019)	Scene text recognition from two-dimensional perspective	CNN-LSTM framework	Script identification plays a significant role in analyzing documents and videos. In this paper, we focus on the problem of script identification in scene text images and video scripts. Because of low image quality

(continued)

Table 1 (continued)

S.no	Authors and year	Title of the paper	Algorithm/methodologies	Findings
5	A. Mirza, I. Siddiqi, S.G. Mustufa, M. Hussain (2019)	Pattern recognition and image analysis. impact of pre-processing on recognition of cursive video text	Pattern recognition	High text recognition rates have been reported in the recent years. An important step in training such systems is the preprocessing of images for effective feature learning and classification

machine (SVM) techniques [29], artificial neural network (ANN) technique [30], and deep neural networks (DNN) [29].

3.2 Script Recognition

Textual information may be found in several scripts in many videos. Before delivering the text sections to the separate OCR machines for identification, these scripts must be detected. In video pictures, researchers have looked at script recognition for text, handwritten, and printed texts.

Due to the poor resolution of the text and, in certain circumstances, complicated backdrops, detection of the script in video text is inherently more difficult than in printed or handwritten materials. A variety of strategies have been published in the literature, ranging from simple template matching methods to advanced structural and statistical characteristics.

Texture-based features are proven to be particularly successful in reporting high classification rates among the numerous characteristics used to characterize the script. Textural measures have been studied in several studies, including Gabor filters, LBP [24], and the grey-level co-occurrence matrix (GLCM). To identify between the script classes under inquiry, the collected characteristics are often input into standard classifiers.

3.3 Text Recognition

Text recognition is also known as optical character recognition, and it is a well-studied pattern recognition function. Recognition methods have been researched for printed and manual composing archives (filtered duplicate or camera-based), text continuous scene pictures, and inscription text in films. For literary data in scope of contents, current acknowledgment strategies (such as Google Tesseract [11]) have been found to have near-perfect identification rates. Text recognition in cursive scripts, on the other hand, remains a challenge, particularly for video text.

Analytical and holistic strategies can be used to classify research activities from the perspective of document OCRs. Analytical procedures either work on individual characters or partition the text into characters first. A variety of approaches have been published for text recognition in document photographs, both at the character and word levels. For character-level text recognition, approaches such as graph-based models, Bayesian classifiers, and hidden Markov models among others have been examined. For word categorize appreciation, an assortment of qualities and classifiers [31] have been inspected.

4 Dataset

We have gathered and tagged a large collection of video frames to test our system. Text identifications, script classifications, and text classifications performance may all be evaluated using the labeled frames. The placement of text sections, script information, and ground truth transcription are all saved for each frame.

The collecting of videos is the initial stage in database creation. By collecting live broadcasts from five different news networks, we were able to collect 60 videos. All videos were shot at a resolution of 900 $\times$ 600 pixels with a video frame rate of 25 frames per second. While the videos feature text in English, videos from four of the frequencies have a higher percentage of Urdu text, while videos from one channel have a higher percentage of English text. We extract each frame every two seconds for classification since succeeding frames in a movie contain duplicate information. The significant reason for extricating every edge for 2 s is to confirm that the textual content of the gathered frames is varied. In contrast to the scenario where a succession of frames has (mainly) content of the same text, this allows for diversity in data training and testing. The learning algorithm would be much better if there were as many different words and combinations of characters as possible.

A thorough labeling tool has been built to assist the process of labeling and setting up basic facts. It allows for the storage of the position of each part of the text in the framework and its writing of the basic truth.

5 Methods

Food security challenges have gotten worse in recent years, and they continue to be a hazard to public health. Keeping track of and tracking the details of various events across the food inventory network, food creation, handling, stockpiling, shipment, and showcasing [10, 12] is basic. A major response to food handling provokes has been to build up a precise and solid food handling global positioning framework. Present a framework for following the authenticity of food dependent on the subject of food, non-edible food from your production until it is re-sold to customers [11, 14, 15]. This procedure necessitates the collaboration of a large number of individuals, all of whom are bound by smart agreement with the use of blockchain information stockpiling. Also, vendors could assure the accuracy of food information. This program, which aids in the tracking and verification of food, necessitates the use of a single server to track each feed as it is transmitted between multiple participants. The data cutting option can be used with diets that have a fixed expiration date to decrease the amount of data. To save and manage food data and assure business ownership, we employ a business-standard contract rather than traditional performance records to protect critical business information. Hence, the security of data is obtained, and it helps us to keep away from spam attacks. We can extend the framework to implement deep learning algorithm and also implement in handcraft products with Android application for quality monitoring.

The features of the suggested framework are introduced in this part, which is summarized in text indicator, cursive classifiers, and text detectors, and the three key components make up the total system. At the application layer, a variety of systems such as indexing and retrieval, keyword-depends-conscious creation, and content summarizing may be built on top of these modules. The text detector module is in charge of detecting and limiting all word-based material in a video frame. Because script might appear in many scripts (during similar videos), the found literary areas are provided to the content grouping bundles, which isolates the text lines into lines dependent on the content. Finally, the text is given to each script's recognition engines, which turn the images of text lines into texts that may be used in a variety of usages. The next sections go through each of these components in depth.

5.1 Text Detection

The suggested method begins with the identification of probable portions for the text covered in image frames. To distinguish textual content in a given frame, latest convolutional neural networks (CNNs)-based feature detection was used. Despite the way that many article locators are prepared with enormous class examples and accomplish high precision in object identification, they are unable to discern text regions in photographs. These models must be adapted to the difficulty of discriminating between text and non-text areas in Fig. 4. Convolutional bases for these models

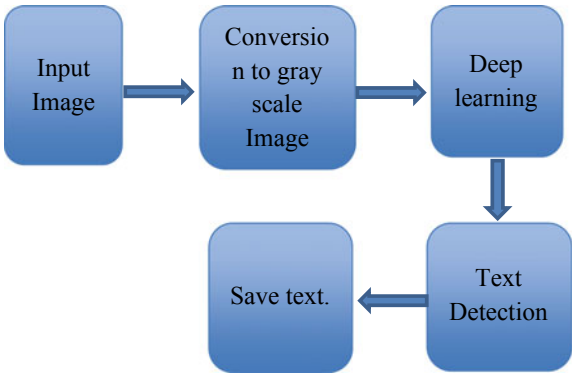


Fig. 4 Overall architecture of the proposed method

can be taught without any preparation or fine-tuned already trained models can be trained on the text and non-text areas to fine-tune them. We examined the finders of the following text-creation feature in this study. The aim is to figure out which one is best for addressing the text identification issue.

- R-CNN
- Single shot detector (SSD)
- Efficient and accurate scene text detection pipeline (EAST).

Quicker R-CNN has been upgraded by R-CNN and Fast R-CNN. Quicker RCNN joins a district proposition network with Fast R-CNN [32] for successful and productive article limitation and recognition. The single shot multi-box detector (SSD) structure was presented by Liu et al. and high accuracy of object detection in common databases such as PascalVOC and COCO.

The plan has 300 info measures and depends on the VGG-16 model without full incorporated layers. EAST detects text patches, as well as geometries in normal scene photos. It can anticipate entire lines of text as well as individual words.

This model's other distinguishing feature is its ability to predict the geometrical form of text. EAST was created particularly to recognize text from scene photos, whereas Faster RCNN and SSD are generalist object detectors. However, when practical straight to the identification of subtitle text in image frame pictures in our investigation, it did not describe about satisfactory recognition performance. As a result, EAST was well processed in our database over Faster R-CNN and SSD.

5.2 Script Identification

To decide the content of a text, content recognition employs text lines as data. It is worth emphasizing that we treat text detection as a two-class problem, with text and non-text areas being distinguished. It is also feasible to conceive of it as a $k + 1$ class problem, where k is the amount of lines (in our example, $k = 2$), and 1 is

the number of classes. To put it another way, the identification calculation might be prepared to perceive text in specific content, removing the requirement for a different content recognition module. Several features of the text in distinct scripts, such as alignments and edge density, are well known. As a result, it seemed more natural to create a system to identify what is text and non-text, which is exactly what we accomplished in our research. We use CNNs in a cataloging outline to identify scripts. Pre-trained ConvNets are fine-tuned to differentiate between the two classes using Urdu and English text lines.

Once a text line's script has been detected, it is passed to the appropriate recognition engine, as described below.

5.3 Text Recognition

As previously said, we are especially interested in videos from news networks that feature cursive writing (Urdu in our instance) as well as text in the Roman cursive. Text OCR structures for English are fairly advanced.

As a result, we use the Google Tesseract OCR engine to recognize English text. The performance of the Google recognition device for cursive text in English, on the other hand, was not particularly promising. As a consequence, we developed our Urdu text line recognition engine (ENGNet). The specifics of the acknowledgment are as follows.

5.4 Google Tesseract

Google Tesseract [11] represents a cutting-edge OCR engine that gives excellent accuracy for a variety of languages, including English. Tesseract version 4.0 that was just published by Google was used in our system. DNN, notably recurrent NN with extended short-term memory architecture, was used to create Version 4.0. The recognition engine receives the English text lines and returns the relevant textual strings.

In recent years, recurrent networks have shown considerable improvements in problems like voice, handwriting, and caption text detection. Small RNNs [8, 33–35] are not able to simulate long-term requirements for input orders, which is why alternatives such as short-term memory networks (LSTM) are widely used. The LSTM is a sort of recurrent component that has three gateways: forget, output, and input.

The sigmoid function is used to build these gates, which control the memory of an LSTM cell. The two-way LSTMs, which analyze feed in both front and back directions and combines better-guessing information, are also widely used.

We progress with a seven-layer convolutional neural network for feature extraction. The ConvNet is given input text line pictures that have been preprocessed.

Height normalization, picture binarization, and flipping are all part of the preprocessing. Urdu languages are written from right side to left side, but western language is printed from left side to right side, in the flipping. The character orders in the record are aligned per the picture after flipping an image. The CNN maps feed lines of text pictures into a feature map that is then fed to the recurrent layers in a sequence. Two layers of bidirectional LSTMs are used to build the recurrent half of the network. Pre-frame predictions are generated by the LSTM, which are then transformed to class labels by a CTC layer [36, 37]. Finally, to translate the class labels to real unicode and generate the output transcription, a look-up table is employed. The CTC loss function is used to conduct training from beginning to conclusion. Our related research includes comprehensive data on the influence of the preprocessing and recurring elements, respectively [38–45].

5.5 Experimental Result

A complete set of tests is carried out to assess the text detection, script identification, and text recognition modules in order to analyze the effectiveness of the proposed framework. The experimental technique and realized findings of text detection are presented first, followed by those of script identification. Finally, we discuss the performance of our identification engine, which was trained on several sets of photos.

6 Conclusion

The structure for the acquisition and recognition of text from video sequencing text in English and other languages is presented in this work. The current study adds to the existing body of knowledge in several ways. To test detection and identification tasks, we built a huge collection of video frames including some information. We used in-depth reading finders to see text in different types of scripts and neatly arranged to see text in different texts. CNNs are employed in a classification [17] framework to determine the script of the identified textual portions. ENgNet, a hybrid of two-directional LSTMs that reports maximum levels of challenging video text recognition in the script language English, is a key contribution of this research.

References

1. Rajesh Kanna P, Santhi P (2022) Hybrid intrusion detection using map reduce based black widow optimized convolutional long short-term memory neural networks. Exp Syst Appl 194
2. Long S, He X, Yao C (2021) Scene text detection and recognition: the deep learning era. Int J Comput Vis 129:161–184

3. Baran R, Partila P, Wilk R (2018) Automated text detection and character recognition in natural scenes based on local image features and contour processing techniques. In: International conference on intelligent human systems integration. Springer, Dubai, , pp 42–48
4. Dai J, Wang Z, Zhao X, Shao S (2018) Scene text detection based on enhanced multi-channels mser and a fast text grouping process. In: IEEE 3rd International conference on cloud computing and big data analysis (ICCCBDA). IEEE, Chengdu, pp 351–355
5. Xu Y, Shan S, Qiu Z, Jia Z, Shen Z, Wang Y, Shi M, Eric I, Chang C (2018) End-to-end subtitle detection and recognition for videos in East Asian languages via cnn ensemble signal process. Image Commun 60:131–143
6. He T, Tian Z, Huang W, Shen C, Qiao Y, Sun C (2018) Single shot textspotter with explicit alignment and attention. arXiv preprint arXiv:1803.03474
7. Fayyaz MM, Seher Z, Siddiqi I (2018) Urdu caption text detection using textural features. In: 2nd Mediterranean conference on pattern recognition and artificial intelligence. ACM, Rabat, pp 70–75
8. Murugesan M, Thilagamani S (2020) Efficient anomaly detection in surveillance videos based on multilayer perception recurrent neural network. J Microprocess Microsyst 79
9. Wang Y, Shi C, Xiao B, Wang C, Qi C (2018) CRF based text detection for natural scene images using convolutional neural network and context information. Neurocomputing 295:46–58
10. Thilagamani S, Nandhakumar C (2020) Implementing green revolution for organic plant forming using KNN-classification technique. Int J Adv Sci Technol 29(7S):1707–1712
11. Rabi M, Amrouch M, Mahani Z (2018) Recognition of cursive Arabic handwritten text using embedded training based on hidden markov models. Int J Pattern Recogn Artif Intell 32(01)
12. Perumal P, Suba S (2022) An analysis of a secure communication for healthcare system using wearable devices based on elliptic curve cryptography. J World Rev Sci Technol Sustain Dev 18(1):51–58
13. Fasquel J-B, Delanoue N (2018) A graph-based image interpretation method using a priori qualitative inclusion and photometric relationships. IEEE Trans Pattern Anal Mach Intell 41:1043–1055
14. Rajesh Kanna P, Santhi P (2021) Unified deep learning approach for efficient intrusion detection system using integrated spatial–temporal features. Knowledge-Based Syst 226
15. Thilagamani S, Shanti N (2014) Gaussian and gabor filter approach for object segmentation. J Comput Inf Sci Eng 14(2)
16. Mirza A, Siddiqi I (2020) Recognition of cursive video text using a deep learning framework. IET Image Process 14:3–4
17. Santhi P, Mahalakshmi G (2019) Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction. Int J Eng Adv Technol 8(4):839–846
18. Kise K, Lopresti D, Marinai S (2018) Document analysis and text recognition. Int J Document Anal Recogn 24(4)
19. Pandiaraja P, Sharmila S (2020) Optimal routing path for heterogenous vehicular adhoc network. Int J Adv Sci Technol 29(7):1762–1771
20. Liu C-L, Fink GA, Govindaraju V, Jin L (2018) Special issue on deep learning for document analysis and recognition. Int J Document Anal Recogn (IJDAR) 21:159–160
21. Lu W, Sun H, Chu J, Huang X, Yu J (2018) A novel approach or video text detection and recognition based on a corner response feature map and transferred deep convolutional neural network. IEEE Access 6:40198–40211
22. Unar S, Jalbani AH, Jawaid MM, Shaikh M, Chandio AA (2020) Artificial Urdu text detection and localization from individual video frames. J Eng Technol 37(2):429–438
23. Gandhi HK, Shabari Nath P, Chouhan R (2020) Image glossiness from curvelet features using SVM-based classification. In: 10th International conference on computer and knowledge engineering. IEEE Access, 9286630
24. Gao Y, Huang Z, Dai Y (2018) Double supervised network with attention mechanism for scene text recognition. arXiv preprint arXiv:1808.00677

25. Liao M, Zhang J, Wan Z, Xie F, Liang J, Lyu P, Yao C, Bai X (2019) Scene text recognition from two-dimensional perspective. In: Proceedings of the AAAI conference on artificial intelligence, pp 8714–8721
26. Liu W, Chen C, Wong K-Y K (2018) A character-aware neural network for distorted scene text recognition. In: 32nd AAAI conference on artificial intelligence, Louisiana
27. Lei Z, Zhao S, Song H, Shen J (2018) Scene text recognition using residual convolutional recurrent neural network. In: International conference on machine vision and application, vol 29. Springer, pp 1–11
28. Saranya KC, Singhal V (2019) Real-time prototype of driver assistance system for Indian road signs. In: International proceedings on advances in soft computing, intelligent systems and applications. Springer, pp 147–155
29. Liu C-L, Fink GA, Govindaraju V, Jin L (2018) Special issue on deep learning for document analysis and recognition. Int J Document Anal Recogn 21:159–160
30. Bhunia K, Konwer A, Bhunia AK, Bhowmick A, Roy PP, Pal U (2019) Script identification in natural scene image and video frames using an attention based convolutional-LSTM network. Pattern Recogn 85:172–184
31. Gunasekar M, Thilagamani S (2020) Performance analysis of ensemble feature selection method under SVM and BMNB classifiers for sentiment analysis. Int J Sci Technol Res 9(2):1536–1540
32. Deepa K, Kokila M, Nandhini A, Pavethra A, Umadevi M (2020) Rainfall prediction using CNN. Int J Adv Sci Technol 29(7):1623–1627
33. Logeswaran R, Aarthi P, Dineshkumar M, Lakshitha G, Vikram R (2019) Portable charger for handheld devices using radio frequency. Int J Innov Technol Explor Eng (IJITEE) 8(6):837–839
34. Saranya KC, Singhal V (2018) Real-time prototype of driver assistance system for Indian road signs. In: International proceedings on advances in soft computing, intelligent systems and applications. Springer, pp 147–155
35. Deepa K, Thilagamani S (2019) Segmentation techniques for overlapped latent fingerprint matching. Int J Innov Technol Explor Eng 8(12):1849–1852
36. Pandiaraja P, Aravinthan K, Lakshmi NR, Kaaviya KS, Madumithra K (2020) Efficient cloud storage using data partition and time-based access control with secure aes encryption technique. Int J Adv Sci Technol 29(7):1698–1706
37. Deepika S, Pandiaraja P (2013) Ensuring CIA triad for user data using collaborative filtering mechanism. In: International conference on information communication and embedded systems (ICICES), pp 925–928
38. Rabi M, Amrouch M, Mahani Z (2018) Cursive Arabic handwriting recognition system without explicit segmentation based on hidden Markov models. J Data Mining Dig Hum
39. Lai Y, Wang N, Yang Y, Lin L (2018) Traffic signs recognition and classification based on deep feature learning. In: 7th International conference on pattern recognition applications and methods (ICPRAM). Springer, Madeira, pp 622–629
40. Mirza A, Siddiqi I, Mustufa SG, Hussain M (2019) Impact of pre-processing on recognition of cursive video text. In: Iberian conference on pattern recognition and image analysis. Springer, Madrid, pp 565–576
41. Pradeep D, Sundar C (2020) QAOC: noval query analysis and ontology-based clustering for data management in Hadoop. Future Generation Comput Syst 108:849–860
42. Tayyab U, Naeem MF, Ul-Hasan A, Shafait F et al (2018) A multi-faceted OCR framework for artificial Urdu news ticker text recognition. In: 13th IAPR international conference on document analysis systems (DAS). IEEE, Vienna, pp 211–216
43. Hayat U, Aatif M, Zeeshan O, Siddiqi I (2018) Ligature recognition in Urdu caption text using deep convolutional neural networks. In: 14th international conference on emerging technologies (ICET). IEEE, Islamabad, pp 1–6
44. Khattak A, Asghar MZ, Saeed A, Hameed Syed IA, Hassan A, Ahmad S (2021) A survey on sentiment analysis in Urdu: a resource-poor language. Egyptian Inf J 22:53–74
45. Liu Z, Li F, Ren F, Goh WL, Yu H (2018) Squeezed text: a real-time scene text recognition by binary convolutional encoder-decoder network. In: 32nd AAAI conference on artificial intelligence, Louisiana, pp 7194–7201

The Application of Cyclostationary Malware Detection Using Boruta and PCA

Mike Nkongolo, Jacobus Philippus van Deventer, and Sydney Mambwe Kasongo

Abstract An analysis of cyclostationary malware is introduced. The most important cyclostationary features used for network intrusion detection systems (NIDS) are then detected with a feature extractor algorithm, such as Boruta and principal component analysis (PCA). These feature patterns are classified to determine the most cyclostationary ones. In particular, this article shows the relevance of detecting cyclostationary malware for a NIDS by using legacy datasets, such as the KDD99 and NSL-KDD. This research has also used the UGRansome cyclostationary dataset intended to support research on anomaly detection. This dataset is subdivided into normal and abnormal classes of network threats. A comparative analysis based on random forest and support vector machine algorithms is undertaken, and the performance of Boruta and PCA was also evaluated. The research suggests the utilization of PCA in terms of extracting cyclostationary network feature patterns as a viable proposition compared to Boruta. The Internet Protocol (IP), malware financial damages, class C of IP addresses, and signature malware were also found to be the most cyclostationary feature pattern. The UGRansome dataset outperformed the KDD99 and NSL-KDD in terms of detecting signature malware with an accuracy of 99% using the random forest algorithm, while the support vector machine achieved 68%. This research proposes the UGRansome as a suitable choice to reduce the computational time of cyclostationary malware classification. Lastly, the research suggests the utilization of random forest to stratify and detect cyclostationary malware.

Keywords Anomaly detection · Zero-day threats · Cyclostationarity · Machine learning · Feature extraction · UGRansome · Data science · Z-score · Boruta

M. Nkongolo (✉) · J. P. van Deventer
Department of Informatics, Faculty of Engineering, Built Environment and Information Technology, The University of Pretoria, Lynnwood, Pretoria, South Africa
e-mail: u21629545@tuks.co.za

J. P. van Deventer
e-mail: phil.vandeventer@up.ac.za

S. M. Kasongo
Department of Industrial Engineering, School of Data Science and Computational Thinking, Stellenbosch University, Banghoek, Stellenbosch, South Africa

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_41

1 Introduction

The field of NIDS focuses on novel approaches that can detect any abnormal traffic pattern. The recognition of cyclostationary traffic patterns promises considerable usages that can improve the NIDS efficiency and assist in enabling innovative frameworks implementation. Most NIDS do not consider cyclostationarity to distinguish static patterns from dynamic ones. Detecting cyclostationary traffic can help specify whether an intrusion can be viewed as a long-term evolution malware that changes and updates periodically. The objective of this article is to assess the practices of long-term evolution malware from a cyclostationary point of view. The term cyclostationarity is used to determine the traffic series of zero-day threats that can vary depending on the network periodicity. This study is motivated by the lack of cyclostationary datasets for the network intrusion detection problem (NIDP) to better understand the cyclostationarity of long-term evolution malware. In the methodology, the unneeded abstraction of cyclostationarity is avoided.

Particularly, all previous research about NIDS did not consider the evaluation of cyclostationary traffic patterns for zero-day threats detection. Most research uses anomaly detection for malware recognition. The cyclostationarity concept has been studied in science and engineering for different purposes. The periodicity and cyclostationarity are used in the mechanical engineering field to rotate or reciprocate belts, gears, shafts, chains, bearings, propellers, and pistons. In meteorology, the weather prediction studies cyclostationarity due to various seasons' properties caused by the Earth's rotation. In the network communication field, radar, telemetry, and sonar are using periodicity and cyclostationarity due to the scanning, sampling, multiplexing, and modulating of signals. Cyclostationary features can also occur via a rotating antenna that can detect various cyclic and dynamic signals. As a result, the benefit of the experiments presented in this article is diverse for the NIDP since little works are related to the study of cyclostationarity.

For instance, in signal processing and remote sensing, the paradigm of cyclostationarity has been also used. However, the NIDL suffers from the lack of publicly available datasets to analyze, study, and implement cyclostationary malware detection systems. The contribution of this article is the utilization of Boruta, PCA, random forest, and support vector machine algorithms to detect cyclostationary malware. The remainder of this article is divided into the (i) theoretical explanation of cyclostationarity and related works presented in Sect. 2; (ii) proposed methodology, feature extraction, classification, evaluation metrics, computing environment, and datasets to detect cyclostationary malware introduced in Sect. 3, (iii) experimental results about cyclostationary feature recognition narrated in Sect. 4, and in Sect. 5, the conclusions of this research are provided (iv).

2 Literature Review

2.1 Description of a Cyclostationary Malware: How It Gets Differ Over a Traditional Malware

A cyclostationary malware is a network threat having abnormal properties that vary cyclically with time [6]. A cyclostationary network traffic T can be segmented into various samples (T_1, T_2, …, T_n), and the malware is generally hidden in the abnormal traffic sample with cyclic features c_t:

$$T = \sum_{t=1}^{T} c_t = 0 \tag{1}$$

where T is the set of features, and T_n is abnormal when $c_t = 0$ or normal when $c_t = 1$.

$$T = \sum_{t=1}^{T} c_t = 1 \tag{2}$$

To detect cyclostationary malware, the abnormal sample should be recognized and subdivided into adjacent or disjoint partitions. Each partition can then be utilized to classify and study cyclostationary malware detected. Then, each data point in the partition represents a cyclostationary feature pattern of malware. In simplest terms, 'traditional malwares' are network attacks that are not used nowadays. Compared to cyclostationary malware that uses updated protocols that are part of the transmission control protocol (TCP) suite, their protocols are outdated and not part of the TCP/IP suite. Time-series data exhibits cyclostationarity demonstrated by the periodicity or repetition of abnormal features patterns.

2.2 Related Works

Long-term and short-term network traffic have periodic behaviors. This property has been used for NIDS design and performance evaluation. With anomaly detection, one can specify a threshold so that if the volume of the network flow is below or above it, then an anomaly is spotted or cyclostationary threat recognized. Nevertheless, these threshold values will vary and change over time. Recently, [15] used thresholds values to classify cyclostationary network traffic from stationary ones. The classification process was accurate with an enhancement in evaluation metrics. To train and test selected algorithms, [15] defined a model batch of malware, and this machine learning model was computed at various periods to determine whether the accuracy of their

classifiers relied on cyclostationarity. In [8], ensemble learning was used to assess the cyclostationarity of heterogeneous datasets using a stacking technique that evaluated the model using the NSL-KDD and UNSW-NB15 datasets.

A feature selection technique was incorporated to detect relevant features with the proposed stacking approach. The results reveal that the proposed model can accurately detect cyclostationary traffic in the network. An adaptable machine learning method was described in [14] to resolve the feature selection problem using a genetic algorithm. The proposed methodology is combined with other algorithms to detect various malware categories. The methodology is validated and compared with trending techniques using different datasets.

A better standard deviation and mean value were obtained in the optimization phase. Soft computing for low distributed denial of service (DDoS) using autonomous detection was used in [4]. The proposed technique utilized a cloud architecture and Markov model to study the network flow and detect malware from the normal flow using the random forest algorithm. The precision, recall, specificity, F-measure, and accuracy metrics were used to evaluate the model improvement. In-depth assessments of mobile and pervasive computing were provided by [7] who considered an evolutionary framework that is applicable in electric motors forming a pervasive environment due to invisibility. Calls were made in [7] to focus on the formation of pervasive computing fields defining the correlations between objects, infrastructure, people, and buildings, to an effective plan for the expansion or establishment of urban centers.

Furthermore, with the popularity of long-term evolution malware including ransomware; cyclostationarity is becoming the primary source of network intrusion used by zero-day threats. Long-term evolution devices represent only 0.9% of Internet connectivity since 2012 and account for 14% of cyclostationary network traffic [3]. Since more users moved from using computers to mobile devices, hackers are now shifting to mobile device intrusion, and this trend is influenced by the proliferation of zero-day threats. The cyclostationarity of zero-day threats can then be analyzed through the network traffic patterns values deviations. The traffic is usually distributed with an abnormality that needs to be detected.

2.3 Limitation in Existing Works

The NIDL has a solid background in terms of normal threats detection methodologies but lacks the analysis of stochastic, cyclostationary traffic, queuing of network flow, intrusion modelization, and zero-day threats taxonomy [1]. Shows the framework used in this research to assess the cyclostationarity of various malware. The cyclostationary analysis of network traffic patterns is rarely utilized for the NIDP. To address this limitation, we have used a supervised machine learning model [12], which has previously been implemented for a variety of NIDP (see Fig. 1).

Figure 1 portrays the supervised learning framework for cyclostationary malware detection using two outdated datasets such as the knowledge discovery and data

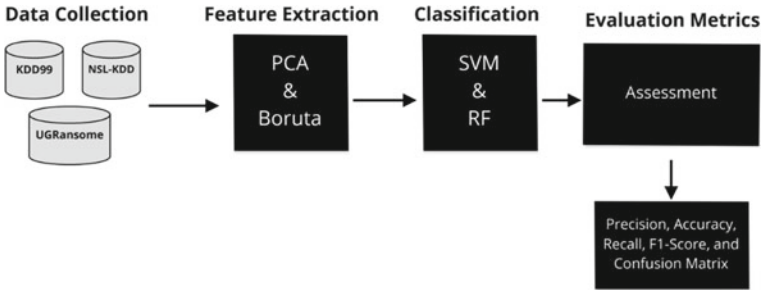


Fig. 1 Overview of the machine learning model for cyclostationary malware detection

mining (KDD99) and network security laboratory–knowledge discovery and data mining (NSL-KDD). The aim is to detect cyclostationary patterns from these datasets. Similarly, a cyclostationary dataset named UGRansome is also used to achieve the same goal [6]. The supervised learning framework utilizes two algorithms for comparison purposes. The support vector machine (SVM) and random forest (RF) are selected and evaluated with a confusion matrix, recall, F1-score, precision, and accuracy.

Additionally, some of the approaches to assess cyclostationarity of malware may require special skills because, contrary to well-known techniques that focus on normal attack detection, the recognition of long-term evolution malware such as zero-day threats might necessitate a periodicity detection that greatly relies on time and thus requires rare or transient process analysis. The supervised analogy to cyclostationary malware detection will ease the design and implementation of NIDS for zero-day threats detection.

2.3.1 The KDD99 Dataset

The KDD99 illustrated in Fig. 2 was used as the first benchmark dataset for NIDS, assessed at MIT's Lincoln Lab but sponsored by DARPA from 1998 to 1999. This dataset provides predictive categories to stratify various network threats, and these categories are remote to local (R2L), probe, user to root (U2R), and denial of service (DoS) [10]. However, the normal category behaviors of different malware are also considered in the corpora. Network patterns have 41 attributes categorized as traffic, content, and basic. Most of the network threats belong to the DoS and normal classes with only 98.61% [10]. Figure 2 exhibits the imbalanced property of this dataset using the normal category.

Imbalanced categories can be thought of as a condition in which one type of class is more represented than the other class. This means that the data is skewed toward a particular type of category and favors the machine learning classification results for that specific class [6]. In Fig. 2, the classification favors the normal category and makes the dataset imbalanced. The training set of the KDD99 contains 4,898,431

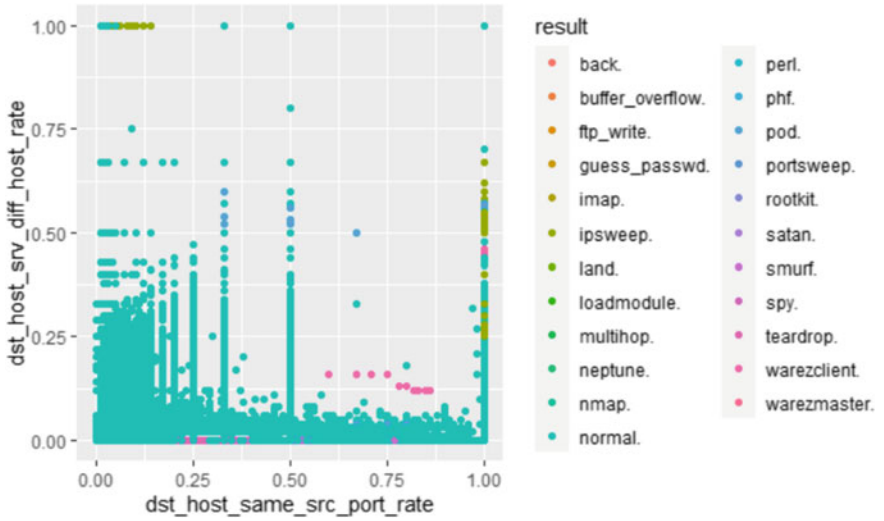


Fig. 2 Imbalanced malware of the KDD99 dataset. The classification result depicts various network attacks on the right side of the figure (back, imap, land, satan, spy, etc.)

Table 1 Malware detected in the KDD99 and NSL-KDD datasets

DoS	Probe	U2R
Back	IpSweep	Buffer overflow
Land	Nmap	Load module
Neptune	PortSweep	Perl
Smurf	Satan	RootKit
Teardrop	*NA*	NA

rows with 2,984,154 observations. Duplicate features are found in the testing and training corpus [10]. The KDD99 dataset is outdated and is not well suited for cyclostationarity analysis.

Table 1 presents malware detected in the KDD99 and NSL-KDD datasets. Most of the network attacks belong to the DoS class.

2.3.2 The NSL-KDD Dataset

The major drawback of the KDD99 is that it suffered from duplicate features skewed with only a few categories and generates false alarms. These limitations are overcome by the NSL-KDD dataset that reduced redundancy in the testing and training corpus (Fig. 3) and minimizes the bias toward little categories [10]. Useless parameters are removed in the dataset to test and train machine learning algorithms [10]. The NSL-KDD dataset has a limitation like the lack of novel malware inclusion, and it

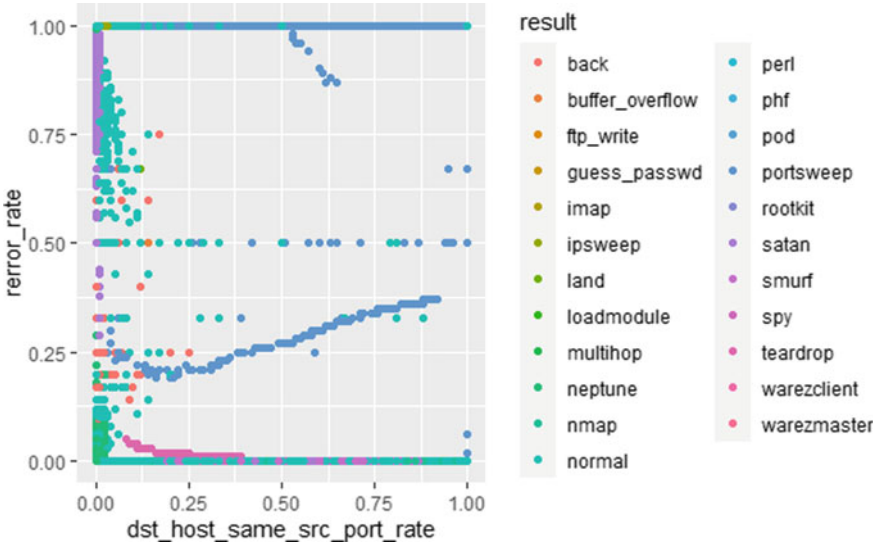


Fig. 3 Balanced malware of the NSL-KDD dataset

is not cyclostationary-oriented (Fig. 3). In Fig. 2, the service rate used by a host has been plotted against the destination port to study the malware cyclostationarity, while Fig. 3 plots the error rate of the destination port.

3 Research Methodology

The classification of cyclostationary malware using supervised learning can be achieved via two phases, namely the extraction of salient features and their classification (Fig. 1). The former extracts important features, while the latter aims at providing their taxonomy. The proposed methodology relies on Boruta and PCA algorithms that act as the feature extractor components. PCA is normally used for dimensionality reduction, but its role as a feature extraction technique is justified because it can be used to visualize the data relevance and exhibits variance.

After the extraction of important features from the aforementioned datasets, the random forest and support vector machine are learning to perform the stratification tasks (Fig. 1). Lastly, the performance of the random forest and support vector machine algorithms is evaluated, and results are presented. Figure 1 exhibits the steps required for the implementation of the suggested methodology. This methodology starts with data collection and ends with the supervised learning evaluation phase elucidated in the upcoming sections. Cross-validation was used to subdivide the dataset into training and testing sets: 80% is used for training and 20% for testing.

3.1 The Cyclostationary Dataset

The UGRansome dataset has been used in the experiments (Fig. 4). It is derived from the UGR'16 and ransomware datasets, and it is made of 207,534 cyclostationary features and 14 tuples or attributes [6]. These features have been triangulated using data fusion, and it is well suited for anomaly detection.

Various long-term malware presented in Fig. 4 has been stratified into a tripartite predictive class: signature (S), synthetic signature (SS), and anomaly (A). Similarly, this dataset includes 16 ransomware families classified per cluster: APT, Locky, DMALocker, SamSam, and much more (Fig. 4). Each cluster is assigned a numerical value to ease the machine learning computation [6]. Contents of the UGRansome are periodic and cyclostationary in nature. This dataset was built in 2021 [6], and it is publicly available. The network traffic of UGRansome does have cyclostationarity properties. In Fig. 5, the network traffic of cyclostationary malware has been segmented based on nine network flags. The AP partition flag contains the most anomalous (A) and cyclostationary malware. The ASF flag partition includes synthetic signature malware (SS). Novel malware has a balanced proportion, and more than 100,000 IP addresses have been labeled with a balanced distribution (class A, B, C, and D). The UGRansome dataset can also be used to enable the recognition of other malware such as SSH, Bonet, DoS, Port Scanning, NerisBonet, and Scan. Table 2 presents the data structure of the UGRansome dataset.

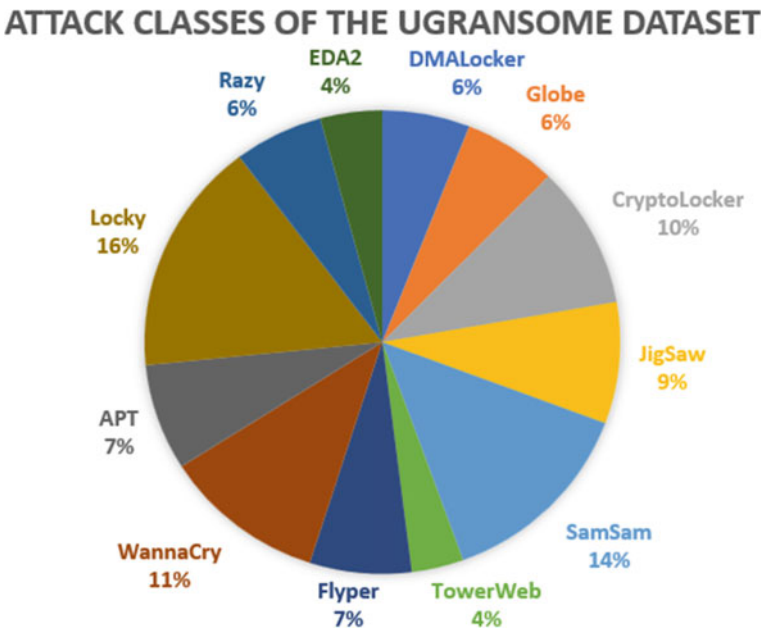


Fig. 4 Malware classes of the cylostationary dataset. APT stands for advanced persistent threats

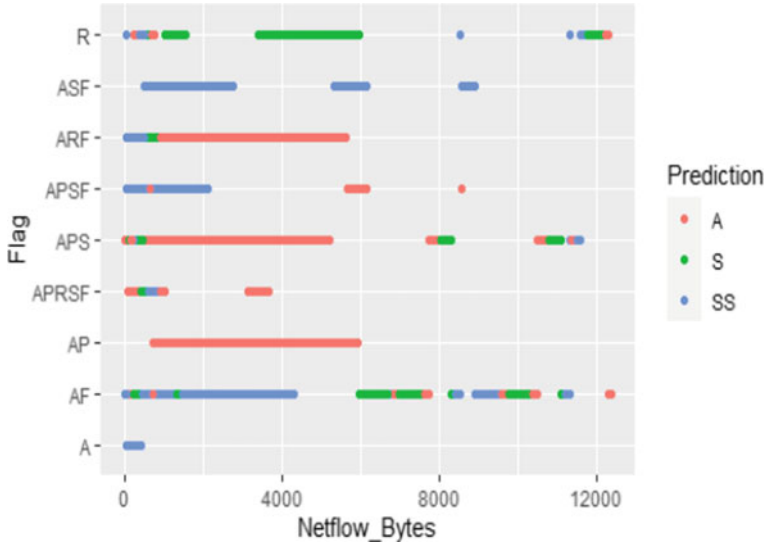


Fig. 5 Cyclostationarity of anomalous network traffic patterns in the UGRansome dataset

Table 2 Data structure of the cyclostationary dataset

Attribute	Example	Description
Prediction	SS	Synthetic signature
Ransomware	WannaCry	Novel malware
Bitcoins (BTC)	60 BTC	Ransome payment
Dollars (USD)	130 $	Ransome payment
Cluster	1	Group assigned
Seed address	*1dice6yg*	Malware address
Expended address	*4ePEyKtk*	Malware address
Port	*5062*	Communication port
Malware	*Bonet*	Novel malware
Network traffic	*18,000 bytes*	Periodic network flow
IP address	*Class A*	Unique address
Flag	*AF*	Network state
Protocol	*TCP*	Communication rule
Timestamp	*40 s*	Traffic duration

3.2 Feature Extraction with Boruta Algorithm

Boruta has been used to estimate the relevance of the KDD99 and NSL-KDD features. It is a feature extractor algorithm that uses the random forest to solve a variety of regression and classification problems [2, 5]. This algorithm is based on different

decision trees developed independently on the various samples with bags taken from the training corpus. Boruta is implemented as a wrapper technique that considers the original data source and extends it by adding shadow features that have random values that are permuted among the training observations [2, 9]. The relevance of features is computed as the loss of the stratification accuracy of the randomized and permuted features [2, 5, 9]. Firstly, the loss of stratification accuracy is individually executed in the forest of the decision trees which makes it possible for a given feature to be classified differently by various trees.

Then, the standard deviation and average computation of the loss of stratification accuracy are computed. Boruta uses the Z-score to estimate the importance or relevance of a feature. The score is calculated by subdividing the loss average by the standard deviation. The variable relevance measure is then utilized to rank features into three categories (rejected, confirmed, and tentative). This Z-score is computed for each feature. Then, the maximum Z-score (MZS) between shadow features is recognized, and a hit is allocated to each feature that scores more than the MZS. The features with less importance will have a Z-score less than the MZS and will be rejected [2, 5, 9]. The features with the highest importance will have a score higher than the MZS. These features are considered relevant, and Boruta will stratify them in the confirmed category. The features with uncertain Z-scores are viewed as remaining features, and they are stratified in the tentative category.

3.3 Feature Extraction with PCA Algorithm

With PCA, $y_{k,l}$ represents recorded features of the ith category made of jth instances included in a Y matrix $(n * p)$. The matrix columns of Y are normalized with zero mean and standard deviation, but features are computed as follows for a given input x [13]:

$$x_{k,l} = \frac{y_{k,l-y_l^*}}{s_l} \tag{3}$$

where s_l and y_l^* represent the standard deviation and mean of Y.

3.4 Classification

Random forest. This algorithm constructs individualistic decision trees from the training sample. Predictions are pooled from all trees to make the final result of classification. In short, the random forest algorithm utilizes a set of results to make a final prediction or classification, and they are commonly named ensemble learning approaches [11]. The relevance of features is computed by using the decrease in the

impurity of weighted nodes. The probability is computed by using the frequency of features in the node, subdivided by the sum of all samples [11]. The greatest value represents the most important feature in the dataset. The total feature's relevance value is computed and subdivided by the number of trees.

Support Vector Machine. The support vector machine sorts feature into binary or multiple categories by using a threshold as a separative measure. Each feature is represented by a data point in the hyperplane, and the Lagrange formula is generally computed to segregate different categories. Lastly, the Euclidean distance is calculated between the threshold and data points to draw a boundary that distinguishes clusters. The boundary differentiating data points can be written as follows:

$$H : W^T(x) + b = 0 \tag{4}$$

W is the weighted features, while x denotes original inputs features. The hyperplane is denoted by H and its bias by b.

3.5 Evaluation Metrics and the Computing Environment

The following evaluation metrics are used to assess the proposed framework. False positive (FP) and false negative (FN) represent misclassification, while the correct classification is represented by true positive (TP) and true negative (TN):

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{FN} + \text{FP} + \text{TN} + \text{TP}} \tag{5}$$

This metric represents the ratio of accurate classification for both true negative (TN) and true positive (TP) cases. The computational time is used to assess the feature extraction and classification performance, while the confusion matrix evaluates the supervised learning algorithms by tabulating the correct stratification results.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{6}$$

The recall value specifies the actual positive cases predicted correctly by the machine learning algorithm. The formula is as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{7}$$

The $F1$-score converges the mean of the recall and precision by showing the overall combined evaluation performance:

Table 3 Hardware and software specification

Node	Type
Test set	20%
Train set	80%
Random state	42
Classifier	Random forest and support vector machine
Feature extraction	*Principal component analysis and Boruta*
Number of trees	*100*
Dataset	*KDD99, NSL-KDD, and UGRansome*
Language	*R with RStudio*
RAM	*39 GB*
System	*64-bits*
Operating system	*Windows*
Computer	*Lenovo ThinkBook*
Library	*Caret*

$$F1\text{−score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

The confusion matrix ($n * n$) is a matrix used in the computation of TP, TN, FN, and FP to calculate the evaluation metrics such as accuracy, precision, and recall. In this research, results of the classification using selected classifiers will be tabulated. Lastly, the computational framework is tested with cross-validation. The computing environment representing the hardware and software specification framework is illustrated in Table 3. The random state of 42 is used for cross-validation.

4 Results

The three datasets were trained with the selected algorithms. In Table 4, Boruta took more computational time using the NSL-KDD dataset. Boruta used more iterations with the KDD99 and rejected more features of the UGRansome dataset.

Results of the PCA using the UGRansome dataset are depicted in Fig. 6. The TCP network protocol (92,157), financial damage of the malware (126,379), class

Table 4 Boruta results using KDD99 and NSL-KDD datasets

Dataset	Time	Iteration	Relevant	Rejected
KDD99	1.2 h	12	36	2
NSL-KDD	2.30 h	24	36	3
UGRansome	45 min	5	8	6

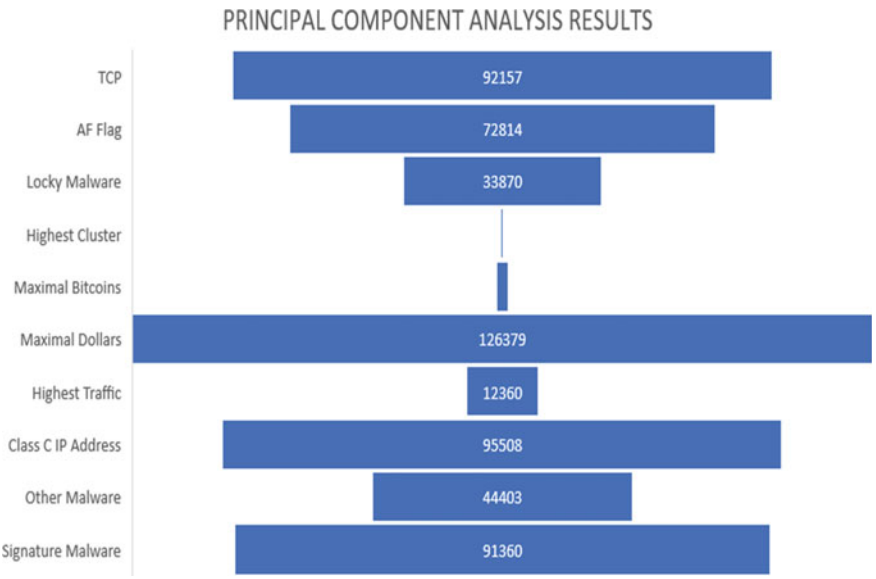


Fig. 6 Principal component analysis results

C of IP addresses, and signature malware are detected as the most cyclostationary features. It was found that the PCA took less computational time compared to the Boruta algorithm that uses more iterations to produce final results. The PCA results were also easy to understand compared to Boruta.

The PCA produced stratification results in less than one minute. Figure 7 shows the random forest and support vector machine reaching an accuracy of 100% on the NSL-KDD dataset.

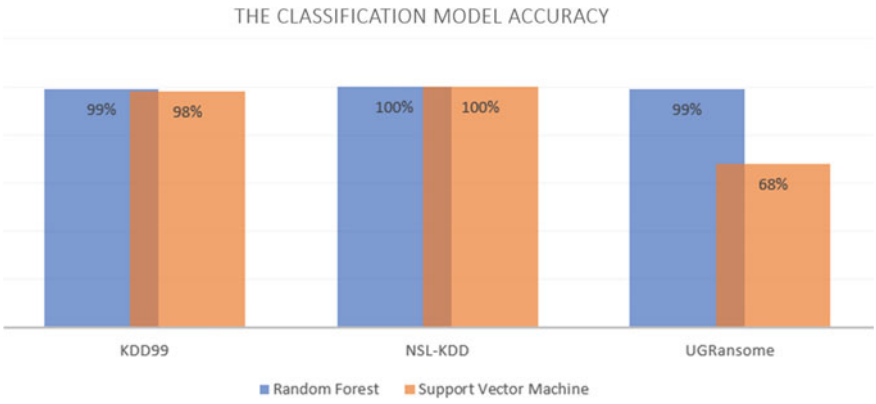


Fig. 7 Comparative analysis

Table 5 Confusion matrix results

Dataset	Random forest	Malware
KDD99	10 721	Neptune
NSL-KDD	4 122	Neptune
UGRansome	89,490	Signature malware
Total	95,333	

The random forest accuracy using UGRansome and KDD99 is similar (99%). The confusion matrix of the random forest is presented in Table 5 with a UGRansome dataset that outperformed the KDD99 and NSL-KDD in terms of classifying cyclostationary features patterns. Signature malware is detected with an occurrence of 80,490.

In general, this computational experiment indicates the utilization of PCA in terms of extracting cyclostationary network feature patterns as a viable proposition. However, the Boruta algorithm is more precise compared to the PCA algorithm. The support vector machine performance was not also impressive with an accuracy of 68%. The most cyclostationary feature patterns are the TCP and financial damages of signature malware in the UGRansome dataset. The accuracy, precision, F1-score, and computational speed are analyzed in Table 6. The NSL-KDD achieved the highest precision value of 94% with 95% of $F1$-score using the support vector machine algorithm (Table 6).

It also obtained 110 ms in terms of computational speed which outperformed the UGRansome and KDD99 datasets.

The NSL-KDD achieved the highest precision value of 96% having 94% of F1-score using the random forest algorithm (Table 7). The UGRansome obtained 100 ms in terms of computational time which outperformed the KDD99 and NSL-KDD datasets. The CPU speed is used to analyze the computational time of the random forest and support vector machine. The highest time was achieved by the support

Table 6 Evaluation metrics with support vector machine

Metrics	KDD99	NSL-KDD	UGRansome
Accuracy	99%	100%	68%
Precision	93%	94%	62%
F1-score	94%	95%	63%
CPU speed	195 ms	110 ms	180 ms

Table 7 Evaluation metrics with random forest

Metrics	KDD99	NSL-KDD	UGRansome
Accuracy	99%	100%	99%
Precision	95%	96%	93%
F1-score	92%	94%	90%
CPU speed	125 ms	120 ms	100 ms

vector machine applied on the KDD99 (195 ms), and the least was obtained by the random forest applied on the UGRansome (100 ms). The research proposes the UGRansome as a suitable choice to reduce the computational time of cyclostationary malware detection.

5 Conclusion

There is a need for intelligent methods to detect zero-day attacks that portray cyclostationary comportments. This article has used three datasets to analyze the cyclostationarity of long-term evolution malware. We have used a feature extraction approach based on Boruta and PCA algorithms to retrieve cyclostationary patterns from the KDD99, NSL-KDD, and UGRansome datasets. We have also classified these features with the random forest and support vector machine algorithms. The obtained results demonstrate that the random forest achieved an accuracy of 100% on the NSL-KDD and 99% on both the KDD99 and UGRansome datasets. It was also discovered that the computational speed of selected machine learning classifiers was affected by the size of the datasets used.

However, this experiment is restricted to supervised learning, and this increases the need for implementing deep learning for cyclostationary long-term evolution malware recognition. One can also evaluate the feature extraction performance on the UGRansome dataset using genetic algorithm.

References

1. Alghawli AS (2022) Complex methods detect anomalies in real time based on time series analysis. In: Alexandria Eng J 61(1):549–561
2. Iman AN, Ahmad T (2020) Improving intrusion detection system by estimating parameters of random forest in Boruta. In: 2020 International conference on smart technology and applications (ICoSTA). IEEE, pp 1–6
3. Lin CY, Chen B, Lan W (2022) An efficient approach for encrypted traffic classification using CNN and bidirectional GRU. In: 2022 2nd International conference on consumer electronics and computer engineering (ICCECE). IEEE, pp 368–373
4. Mugunthan SR (2019) Soft computing based autonomous low rate DDOS attack detection and security for cloud computing. J Soft Comput Paradig (JSCP) 1(02):80–90
5. Nehra D, Mangat V, Kumar K (2021) A deep learning approach for network intrusion detection using non-symmetric autoencoder. In: Intelligent computing and communication systems. Springer, pp 371–382. https://doi.org/10.1007/978981-16-1295-4_38
6. Nkongolo M, van Deventer JP, Kasongo SM (2021) UGRansome1819: a novel dataset for anomaly detection and zero-day threats. Information 12(10):405
7. Raja K, Lilly Florence M (2019) Implementation of IDS within a crew using ID3Algorithm in wireless sensor local area network. In: International conference on ınventive computation technologies. Springer, pp 467–475
8. Rashid M et al (2022) A tree-based stacking ensemble technique with feature selection for network intrusion detection. Appl Intell 1–14

9. Salih AA, Abdulazeez AM (2021) Evaluation of classification algorithms for intrusion detection system: a review. J Soft Comput Data Mining 2(1) 31–40
10. Sekhar Ch et al (2021) Deep learning algorithms for ıntrusion detection systems: extensive comparison analysis. Turkish J Comput Mathe Edu (TURCOMAT) 12(11):2990–3000
11. Shah S et al (2021) Implementing a network intrusion detection system using semi-supervised support vector machine and random forest. In: Proceedings of the 2021 ACM southeast conference, 2021, pp 180–184. https://doi.org/10.1145/3409334.3452073
12. Tahir A et al (2021) Hrnn4f: Hybrid deep random neural network for multi-channel fall activity detection. Prob Eng Inf Sci 35(1):37–50
13. Ullah I et al (2021) Detection of cybersecurity attacks through analysis of web browsing activities using principal component analysis. arXiv preprint arXiv:2107.12592
14. Vivekanandam B (2021) Design an adaptive hybrid approach for genetic algorithm to detect effective malware detection in android division. J Ubiquit Comput Commun Technol 3(2):135–149
15. Yinka-Banjo C et al (2022) Intrusion detection using anomaly detection algorithm and snort. In: Illumination of artificial ıntelligence in cybersecurity and forensics. Springer, pp 45–70

Digitalization and Its Dimensions in a Social and Educational Context

Vladislava Lendzhova, Valentina Milenkova, Dobrinka Peicheva, Mario Marinov, and Dilyana Keranova

Abstract The aim of the article is to show the role of digital skills and literacy as a consequence of the processes of digitalization, but also as a prerequisite for conducting online learning in educational systems. The different dimensions of digitalization and the formation of digital skills by people of different ages and professions are analysed. The methodology is based on various surveys, including an online survey conducted by the authors of the article among the Bulgarian population on the state of digitalization and online learning. An important place in the study is occupied by online education in secondary and higher education in the situation of the COVID-19 pandemic. Views and assessments of the participants in the learning process and their readiness for mobile learning are shown. The main conclusion of the article is that in Bulgarian society there is a relatively good availability of digital skills, which are important for modern processes.

Keywords Digitalization · Digital skills and literacy · COVID-19 · Online learning · Bulgarian society

V. Lendzhova (✉) · V. Milenkova · D. Peicheva · M. Marinov · D. Keranova
South-West University "Neofit Rilski", 66 Ivan Mihailov Street, 2700 Blagoevgrad, Bulgaria
e-mail: vlendzhova@swu.bg

V. Milenkova
e-mail: vmilenkova@swu.bg

D. Peicheva
e-mail: peichevad@swu.bg

M. Marinov
e-mail: Mario_marinov@swu.bg

D. Keranova
e-mail: dilyana.keranova@swu.bg

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_42

1 Introduction

Digitalization has actively entered people's lives—both private and public. It is a very important part of the activities of institutions at different levels and in different areas. Business, the public and the non-governmental sector systematically use various digital tools and services in their activities [1]. Digitalization is ubiquitous and global, and it changes generations and their thinking, their attitudes and qualities, and their skills and reflexivity [2]. It affects and reformats social structures and relationships, imposing on them the requirements of digital technologies.

The digitalization of society is an indisputable fact and is a prerequisite for social development and transformations, because it means providing new opportunities for the development of institutions as well as for the formation of new skills and qualities.

The aim of the article is to show the role of digital skills and literacy as a consequence of the processes of digitalization, and also as a prerequisite for conducting online and mobile learning in educational systems. Digitalization is emerging as a complex and dynamic process that becomes a mandatory element of people's relationships and their participation in various structures, networks, and institutions. The article shows the different dimensions of digitalization and the formation of digital skills by people of different ages and professions. Central to the analysis is data from the National Statistical Institute, because they focus on creating a generalized image of the penetration of digitalization in society, such as the use of various digital devices, and also the provision of Internet connection, which become conditions for the use of digital environment. The results of various studies also contribute to showing the views and assessments of individuals to the processes of digitalization. The idea of the used empirical material is to create a complete picture that has factual and subjective dimensions of the life of the modern Bulgarian, who shops digitally, communicates digitally, banking digitally, perceives culture in the form of movies, and other arts also digitally. On this basis, an important characteristic can be deduced that the modern postmodern person is a digital person.

An important place in the article is occupied by online education in secondary and higher education in the Bulgarian environment in the situation of the COVID-19 pandemic, which is a serious challenge. Online learning is a global and national response to the pandemic situation and lockdown within the framework of secondary and higher education, in a short time the teachers managed to adapt to the requirements of the online environment and to prepare educational resources that meet the digital requirements. In fact, all these different content accents show the existence of interconnected themes that reveal new worlds constructed in different areas of social life.

The research questions formulated in the article are: What are the main characteristics of digitalization in Bulgarian society? To what extent is online learning perceived as an important and necessary aspect of modern educational life?

In general, the digitalization of society and education in particular confronts modern man with the need for a deeper and more critical perception and understanding of information and the world around us. The huge amount of information

that we are surrounded by raises a series of questions regarding its truthfulness, verification, quality, etc. In this sense, digitalization has another positive consequence—it encourages us to be reflexive and critical thinking about things in life—goods, knowledge, and information. Thus, digitalization has not only instrumental, but also terminal value, aimed at developing thinking and reflexivity, and accepting things in their given, but seeking answers that are beyond contradictory information and facts. This personal meaning is a serious contribution that we are aware of and understand with increasing clarity.

2 Methodology

The methodology of the article includes several different methods:

- Online sociological survey conducted in March 2021 with 1018 people in the age group 18+ and carried out under the project "Digital Media Literacy", funded by the National Science Fund and implemented by the Department of Sociology of South-West University (SWU), aimed at studying digital skills and literacy of population from different ages and communities. The online survey as a method was chosen due to the COVID situation, and in this sense, we were placed in a specific objective framework. The online survey allowed us to reach more respondents, but it also became a disadvantage because the questionnaire was filled in mainly by digitally active people. The link to the questionnaire was uploaded to various platforms and organizations, including Facebook, and sent to various communities. The questionnaire included several topics: measuring digital skills, assessing digital learning, attitudes towards digitalization, and its pros and cons. The topic of digital skills was developed on the basis of the following indicators: time (number of hours) spent online per week, number of web sites visited, information retrieval skills, communication, content creation, and network security. The processing of the obtained results was done with SPSS. The summarized information provided an opportunity for various analyses as well as for deriving critical reflections.
- Focus group with students from Sociology specialty of SWU, held in October 2021 and relating to the evaluation of online learning in Bulgarian universities. The focus group was conducted on the same project. The aim is to show the advantages and disadvantages of online learning and its real processes. The focus group was conducted with 10 s and third year students, because the goal was to follow the views of young people who have been studying in a digital environment for three semesters (from March 2020 to May 2021). Discussion topics allowed students to freely share grades and positions as well as to emphasize which of the studied disciplines are more or less suitable for digital learning. It turned out that there are various practical disciplines in which digitalization cannot replace personal contact and acquisition of skills for field work. In general, this method provided

an opportunity for an in-depth and multifaceted discussion of digitalization in university education, which complemented the picture on the topic.

These empirical sources provide a variety of information about the type of institutions and educational levels studied as well as about the time axis and the scope of the included target groups. This creates an enriched picture of the state of digitalization and the digital environment in Bulgarian social and educational conditions.

3 Theoretical Background

3.1 The Main Concepts About Digitalization

In the literature, there are many conceptual definitions of digitalization. This article is focused on two concepts. The first one refers to digitization as the process of digitizing, that is, the conversion of analogue data (e.g., images, video, and text) into digital format [3]. The second one is related with the process of developing social capital and for improving individual skills [4]. These two separate states are a characteristic of the digital world, where digital information is "discrete and pure "while analogue information is" continuous and noisy" [5]. Robinson (2008) defines his analogy much more in detail when he notes that it is analogous as smoothly changing, possessing the properties of the apparent impeccable and untouchable credibility of space and time. It compares analogies with space and time, which allows for infinite division, and in association with them (space and time), connects something authentic and natural to the artificial and arbitrarily truncated precision of the digital. Digitization can also be understood as a process that, in addition to being symbolic, has material dimensions and generates information that can be expressed in many different ways, through many different types of materials and in many different systems. Reference [6] points out that the conversion process occurs through precise technical mechanisms and requires a specific technical infrastructure that realigns the original signal itself. While popular belief often describes digitization as a technical process, people have made certain choices about what to keep and what to discard on algorithms that perform the digitization process [6].

Digitization also generates data with a series of distinctive characteristics, or, as Negroponte (1995) points out, it is the universality of digitized information. In this context, [6] boldly argues that digitization by reducing communication to its essential components produces a lingua franca, capable of facilitating universal communication. Digitized information can be easily stored and transmitted due to a lack of errors, repetition, and statics, which also enables easy manipulation and display of this data. Digitization allows for a considerable measure of interactivity between the user and information, and that in the new social reality demand presence of a digital competence.

3.2 Digital Competence

Digital technologies are changing the world at an incredible speed and are reshaping how people live, work, and study. The development, availability, and accessibility of new technologies, and the definition of digital literacy or digital competencies changed and according to researchers, the concept will continue to change as technologies evolve [7–10]. Wide range of digital technologies brings new challenges and opportunities, transforming our system towards more sustainable one, while it remains to see how individuals, especially from remote areas (facing ongoing problem with an unstable and non-reliable Internet connection), and achieve inclusive and quality education and training. Because of that, multiple national and international organizations put digitalization as the main/focal agenda for most of the policies [11], most of the EU Member States have national and/or regional strategies for digital education. Being digitally competent is both a necessity and a right, and "…being digitally competent involves the confident, critical, and responsible use of and engagement with digital technologies for learning, work, and participation in society…" [11]. It is becoming more important to possess such skills to become a more engaged citizen, playing a more important role in society, engaged in further education and training, accessing the labour market (EC 2020a), facilitating social interaction, promoting instant contact across socio-geographical distances and boundaries, extending information search and spatial horizons, and enriching possibilities for entertainment [12]. But it also brings some negative effects such as sedentarism, inactive lifestyles, loneliness and isolation, fragmenting concentration and attention, as well as potential consequences such as obesity, health problems, learning difficulties, and even mental disorders related to Internet overuse and addiction [13].

Even the young generation of today is often called "digital natives", "…digital technologies uptake and use for teaching and learning requires a critical approach and a holistic perspective…" [13, 14]. Such perspective includes robust digital capacity, planning for organizational change, ongoing monitoring and adaptation, a strong focus on learning-driven pedagogy, leadership, professional development shared understanding, etc. Therefore, a strong link between learners and teachers should be established focusing on several priorities. The growth of use of digital technologies in formal education offers new opportunities to assist people not only during their schooling period but also in their careers at lifetime. COVID-19 has profoundly changed the use of digital technologies in formal education both at institutional and individual level. The formal education institutions increased their capabilities to use of digital technologies, and individuals increased the intensity of use of online learning—both for formal and informal education—compare to pre-COVID-19 levels [9, 15].

In summary, the concept of digital competence is an emerging and reinforcing concept related to technology development as well as to the political goals and expectations of citizenship in the knowledge society. It includes a variety of skills in several areas: media and communications, technology and computing, and literacy and information science. Digital competence consists of: (1) technical skills for using

digital technologies; (2) the ability to use new technologies in a targeted way for work, learning and learning, and for everyday life as a whole in different directions; (3) ability to critically evaluate digital technologies; and (4) motivation to participate in digital culture [16].

4 Results

4.1 Digital Skills

To study the digital skills and literacy of the Bulgarian population, an online survey was conducted in March 2021, in which people in the age group 18+ took part, the sample has the following characteristics.

Age:

- 18–29 years: 37.9%
- 30–39 years: 24.6%
- 40–49 years: 18.7%
- 50–59 years: 8.4%
- 60+ years: 10.4%

The most active age group in the sample are people up to 40 years old, "digital natives" [17] those who are digitally predisposed to the acquisition of digital skills and digital culture people. This is the generation of individuals who, as students or as parents of young children, are actively involved in various forms of interaction with information technology.

4.2 Education

- This was the other important feature of the sample:
- Persons with education "up to secondary" are: 1.4%
- Secondary school graduates make up: 35.6%
- Higher education graduates are: 41.1%
- The persons with scientific degree are: 22.0%

Most are highly educated respondents. This is because digital skills also require a specific general culture and knowledge. Although there are many examples of people who have digital skills without being accompanied by higher education.

The questionnaire included various questions, mainly concerning the topic of digitalization, the use of various digital devices, and the measurement of digital skills themselves [13].

The main reasons that the Bulgarians respondents indicate for to use the Internet services are for making calls, informing, for delaying games, work, overcoming oneself, and social cause. The largest number of respondents indicated as a reason of using Internet the information—86.25%. The share of those who use Internet services for making calls is also high—83.3% of all respondents. More than half of the respondents use the Internet for something related to their work—50.2%. 57.9% use them for training, only 21.4% of all surveys use social networks to entertain with games. As special interests on the Internet, most of the respondents indicate art (music, movies, etc.)—63.5%, followed by those who connect their interests on the Internet with science 62.8%. The percentage of those who indicated education is also high, 74.5%. Again, the low share of those respondents who indicate as special interests in online games remains—this is only 17.1% of those who answered the questions in the survey.

Cyberbullying is also observed as a completely new "acquisition" of modern society with the introduction of electronic means of communication. Cyberbullying is a form of harassment that uses the means offered by modern communication technologies—the Internet and mobile phones. Cyberbullying is considered to be a continuation of social (indirect) harassment. When asked about cyberbullying, almost half of the 49.7% of respondents answered that they did not fall. 35% of respondents received insulting comments, 25% profile theft, 24.6% spread false news and rumours, 15.4% received threats, and 1.2% said otherwise.

To the question: You personally found yourself in a situation in which you were: received offensive comments about your posts, most of the respondents answered "no" 72.1%, as opposed to those who received offensive comments—27.9%. Those of the respondents who received malicious messages are 30.8%, compared to those who did not receive such, and whose number includes more than half of the respondents—69.2%. 19.4% indicated profile thefts, and 80.6% of all respondents had not encountered profile thefts. Regarding the use of photos by other users on social networks, 17.3% of respondents stated that they were victims of this type of cyberbullying, and 82.7 did not face it. 81.7% of all respondents said they had not yet been the victim of rumours on social networks, while 18.3% said they had been the subject of rumours. An important aspect of the skills of the Bulgarian population are digital skills. From the received data, it is clear that the digital literacy of the Bulgarian population has reached an average level, according to the respondents. 69.9% of the respondents rated the digital skills of their fellow citizens as average. Only 11.6% of the respondents give a high assessment of the digital skills of the Bulgarians. In the same way, respondents evaluate their personal digital skills. To the question: How do you assess your digital skills? Most of the participants in the study determined that they have an average level of digital skills of 61.9%. However, it is also high for the percentage of respondents who believe that they have a high level of digital skills of 35.8%. 61.9% assessed themselves as moderately literate, 61.9% defined a high level of digital literacy, and only 2.4% as low (Fig. 1).

Respondents' digital skills include: using the Internet, connecting to a Wi-Fi network, opening a new tab in the browser, downloading and installing applications, downloading and uploading files, filling out online forms, and more. 99.7%

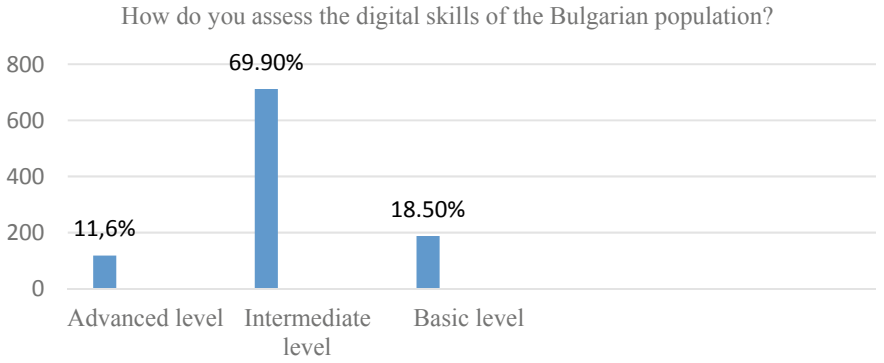


Fig. 1 Bulgarian population's level of digital skills?[1]

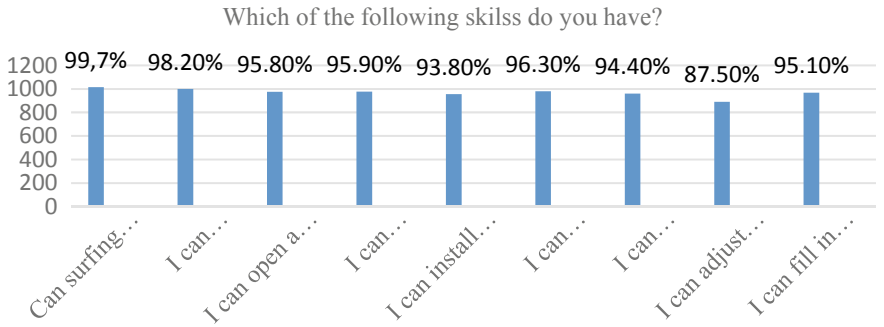


Fig. 2 Most frequent digital skills among Bulgarian population

of respondents say they can use the Internet, 98.2% say that they can connect to a Wi-Fi network, 95.8% can open a new tab, 95.9% can download an application, 93, 8 can install an app, 96.3% can download files, 94.4% can download files, 94.4% can upload files, 87.5% can adjust their privacy settings in their accounts, 95, 1% can fill in an online form, 82.3% can create a Facebook page, and 96.2% can work with an Internet search engine (Fig. 2).

95.5% of respondents can share videos, 94.5% can write and post comments on social networks, 85.7% can follow who sees their messages, 81% can create a video, 44.4% can create a blog, 33.8% deposit, and 22% podcast. 50.3% of all respondents indicated that their parents can use the Internet freely. 23.2% indicate only one of their parents as having Internet skills, while 14.3% of the parents do not use the Internet. 12.2% of the respondents do not have parents. Almost all of the respondents to the question "What smart devices do you have?" Indicated 95.7% of smartphones. 62.3%

[1] The percentages are more than 100, as the respondents gave more than one answer about different digital skills they have.

own a smart TV and 17.2% say they own one. The share of respondents of those who do not have smart devices is small—only 2.4%.

To the question: On average, how many hours a day do you spend on various activities on the Internet? 32.4% of the respondents spend 3 h working, 25.5% spend up to 10 h, 25% up to 1 h, and 2.8% over 10 h. 40.7% of the respondents spend three hours on training, 32% spend 1 h, and 13.2% spend up to 10 h. Up to 1 h is needed to maintain contacts on social networks of 48.9% of all respondents. 38.3% spend up to 3 h, and only 8.4% say they need up to 10 h to keep in touch on the Internet.

The time spent on self-education by 45% of respondents is up to 1 h, followed by those who spend up to three hours 38.5%. The least is those respondents who spend up to 10 h of their personal time to educate themselves on the Internet—7.9%. More than half of the participants in the survey spend 1 h reading newspapers and magazines on the Internet—68.6%. The share of those who spend up to 3 h is relatively smaller—14.50%. Up to and over 10 h on the Internet for reading newspapers on the Internet spend, respectively, 1.6% (up to 10 h) and 0.4% (over 10 h). Reading books on the Internet is not one of the most preferred activities of the Bulgarians who participated in the study. Half of the respondents spend 1 h reading books—52.7%, 22% spend up to 3 h, and only 3.5% up to 10 h.

From the obtained data, it is clear that despite the relatively high digital skills that Bulgarians have, the Internet is not a preferred place to spend free time related to entertainment such as watching movies, listening to music, and more. 46% of the respondents spend 1 h for entertainment, 38.1% up to 3 h, and only 1.4% over 10 h. Several important conclusions can be drawn from the analysis of the obtained results.

The most common reasons for using social networks are receiving some kind of information or having conversations. The largest number of respondents has smartphones or smart TVs. The time that the respondents spend on the Internet is from 1 to 3 h, and the purpose is either for work or training. The share of those who use it for reading or entertainment is very small. From the data obtained it can be argued that in Bulgarian society there is a very high percentage of people who have basic digital skills, use digital devices and have an excellent positioning on the Internet. Still, however, there is a significant difference in the digital skills of different generations. Despite the stated high number of young respondents who have relatively good digital skills, their parents still do not use social networks very often or almost. This in turn imposes the need to develop the digital skills of the Bulgarian population.

4.3 Digitization of University Education

University education is also being transformed into online training during the COVID restrictive measures and restrictions. For several months after the pandemic began, university professors developed such online educational resources. In the survey conducted in 2021, the sampled teachers share how they cope with the situation of digital learning:

- 88% of lecturers use educational resources that are available online.
- 72% of teachers say that they themselves create such resources in accordance with the specifics of the discipline they teach.
- 54% say they use online resources that their colleagues have shared.

The educational environment acquires the basic characteristics of a digital environment and thus becomes adequate to the requirements of online learning while at the same time the teachers themselves are proactive and are responsible for the educational process and teaching methods. In this situation, an indicator of the quality of the conducted training is the feedback from the students and the way in which they evaluate the distance learning. Therefore, in October 2021, we conducted a focus group in Sociology speciality with second and third year students at SWU "Neofit Rilski". The focus group was centred on students' views of digital learning in the education process. Another topic discussed was—what are the advantages and disadvantages of this type of training? What is the difference between in person and online learning, which is preferred and why? The discussions were attended by ten students, all in the age range: 20–28 years. Digital learning itself has been identified as innovative, as it is being introduced comprehensively in all levels of education—secondary and higher education for the first time. The students share that they do not encounter difficulties with the online platforms, with the work in them as well as the preparation of course works and research papers, as this preparation does not differ from (in person) training. In addition, they shared that now it is even easier as digital information on various topics grows and in practice they have the necessary skills to work and share. Two digital platforms for training have been selected at SWU: Blackboard and BigBlueButton. For each subject, the teacher creates a classroom that has its own code and link, which is sent to the students and they enter the virtual classroom to conduct the lecture or the respective lesson. We will summarize the positive aspects of digital learning, following the views of students shared in focus groups:

- The training is mobile and with an individual rhythm:

  Students can study with their own progress and in a place chosen by them. This creates flexibility, accessibility, and convenience (male).

- Training improves digital literacy and skills:

  The availability of different digital platforms stimulates us to develop our digital skills (male).

  The training is being improved from a technological point of view, and it places new demands on the skills and digital literacy of the participants, who need to upgrade their skills. (female).

- The training includes a variety of online resources.

  Different resources are used in the classes, and the teachers make a lot of efforts to make the classes more diverse and saturated and with different materials. (male).

  All lectures, presentations, and topics for discussion are uploaded to the platform and one can return to them many times, read them, and make references. (female).

- The student–teacher relationship is becoming more active

Teachers are much more interested in whether the topics taught are clear; teachers put more effort into explaining, seeking more feedback from students. (female).

- Online learning allows time restructuring.

I have more time for online learning because I do not have to travel to university, and it allows me to better organize my daily life (male).

- Learning effectiveness.

The home environment in which I find myself during online training creates more conveniences for me, which allow me to better understand the material and think that I can handle the content more effectively (male).

Disadvantages of distance learning in a lockdown situation.

- Limited social contacts

In online learning, especially when combined with lockdown, you cannot meet people and feel isolated and physically detached. (male).

Social isolation and loneliness are created, which negatively affects every person (female).

- Technical issues

Various problems can arise—not a good Internet connection or problems with the platform. These problems were more at the beginning of the lockdown—in April–May 2020, they are gradually overcome and there are no more technical problems or they are significantly less (male).

- Pedagogical and organizational problems

All lectures and materials used by teachers need to be done for online learning, some teachers make this restructuring slower and this interferes with lectures (female).

- Health issues

Prolonged standing in front of the computer interferes with vision, immobilizes and is generally quite stressful (female).

The comparison–distance is and/or in person training. According to the students, both forms should be developed and the participants in the educational process should be prepared for them.

If online learning takes place in a normal social environment, without restrictions and closure is preferable because there are many advantages and I prefer it. (female).

I prefer face-to-face training because it gives me more contacts and more social mobility, I like being at university, talking to my fellow students. This has great advantages. Online training closes me and I am not psychologically comfortable (male).

From the presented views, it is clear that both positive aspects of online learning and some weaknesses and shortcomings can be outlined. It is important to note that it has a significant place and is considered a good opportunity in times of social crisis to use. From this point of view, digital learning is a serious reserve in the direction of its improvement and development by educational institutions at different levels, both in formal and non-formal learning.

5 Discussion

There is a vision for the perception of education as an environment for the formation of digital skills and literacy, for the promotion of self-education and the creation of attitudes for lifelong learning. Digital skills need to be considered not only in the context of instrumental skills for access to the digital environment and the Internet. As they provide access to diverse information and knowledge, provide insight into various aspects of the world and relationships, they have also been linked to the formation of sensitivity, behaviour, and culture, with greater activity and community engagement. From this point of view, it is necessary to restructure the approaches to learning and education, as well as the ways in which different forms of teaching and communication between teacher and students are organized. Teachers themselves need to be better prepared for digital learning, it is necessary to develop innovative pedagogical and didactic models that are adequate to the digital environment as well as to the expectations of the students themselves. Most of the information taught in the classes can be found online. This requires teachers to be more creative in building their educational resources, to have more non-traditionalist and motivation to learn. All this brings to the fore the need to form an experience and its upgrading in the teachers themselves, sharing approaches to work in a digital environment and to create educational online resources. The other important emphasis is related to the fact that digitalization itself creates additional divisions between communities and individuals. On the one hand, they are related to the financial opportunities and resources for technological support of digital processes, but on the other hand, they are also related to age characteristics, psychological, and generational difficulties that require systematic efforts to cope. This means working with generations and preparing them for the challenges of digitalisation. Stimulating creativity and critical thinking of people from different ages, encouraging their creative energy and commitment are a hidden resource of digital networks and media. Another aspect of the discussion is related to the young people who are the human capital of digitalization. A change in the culture of understanding and reflexivity of young people means a change in their values and lifestyles. This affects the essence of the interactions between individuals in various social communities, which have different characteristics and specifics, and at the same time, leads to the enrichment of resources, experience, and understanding.

6 Conclusion

Digitalization has an important place in society and is gaining great importance. Almost all social processes are digitalized and the significance of digital skills is growing—in online banking, voting, making bank transfers, paying bills, participating in various platforms, conducting conference calls that are relevant as in everyday life of modern person and in professional work. Digitalization acquires great importance in periods of social crises, when it proves to be an important factor in maintaining the normal course of sociality. It was COVID's situation and lockdown that represented a similar crisis, in which economic, educational, and overall social life had passed into conditions of almost universal and complete blockage, leading to the cessation of many events and processes, but also of various types of learning. That is why the Internet and online platforms are an excellent environment for education, learning, and self-learning.

From the given statistical data, but also the results of surveys, it can be said that mobile learning has a place in Bulgarian education and its quality is relatively good. Online learning depends on the technical and methodological level, creating prerequisites and opportunities for innovation of learning and its individualization. In this regard, a reform of education in the direction of distance education and the transformation of existing textbooks into online form is needed.

Overall, the COVID-19 pandemic has demonstrated the undeniable importance of people's digital skills and the need for a reliable and fast Internet connection. Social networks have faced a significant increase in demand, but it has been argued that social networks cannot compensate for physical isolation.

The basic summary of the article is that digital literacy and skills need to be developed. This is done in parallel with digital innovations and improving the technical characteristics of various digital devices. It is necessary to conduct courses to fill the gaps in the digital skills of people of all ages. This implies active involvement of the public and non-governmental sector. Continuing online learning, which takes place after the blockade with COVID-19, is assessed as important and necessary because it transforms the overall educational environment and the organization of learning and restructures the essence of learning itself. Online training requires technological support, availability of high-speed Internet, and mastering skills for working with digital platforms. An important conclusion is that gradually all participants in the training—students and teachers have adapted to online training, but in general, there are opinions that it is good to be complemented by face-to-face training, which provides live contact and interaction.

It can be said that the Internet and online platforms, based on interactivity, individualization, and new technological opportunities with free and easy access, are an excellent environment for learning and self-learning, to increase the potential of each person. The field of online services and those related to the storage and organization of databases is expanding.

The main conclusion of the article is that digitalization is key in social terms. It has many advantages, and is the basis of various processes, creates conditions for

improving life, and improving the quality of learning and training. Digitalization and digital learning require people to develop a critical view of information and its content, to seek information from more and different sources and to develop a reflection on what is happening in life and in the digital and social environment.

Acknowledgements The article has been developed in the framework of the research project "Digital Media Literacy in the context of "Knowledge Society": state and challenges" № КП–06–H25/4, funded by National Science Fund—Bulgaria.

References

1. Milenkova V, Lendzhova Vl (2021) Digital citizenship in the conditions of social crises: the Bulgarian case. Computers 10(4):40. Available: https://www.mdpi.com/2073-431X/10/4/40
2. Ferrari A, Punie Y, Brecko B (2013) DIGCOMP a framework for developing an understanding digital competence in Europe. Joint Research Centre and Institute for Prospective Technological Studies. Available: ftp://ftp.jrc.es/pub/EURdoc/JRC83167.pdf
3. The Oxford English Dictionary (2019)
4. Pinto M, Leite C (2020) Digital technologies in support of students learning in higher education: literature review. Dig Edu 37:1–18
5. Robinson R, Spalletta G et al (2008) Decreased heart rate variability is associated with post-stroke depression. Am J Geriatric Psych 16(11):867–873. Sanz-Labrador I, Cuerdo-Mir M, Doncel-Pedrera LM (2021) The use of digital educational resources in times of Covid-19. Social Media + Soc 7(3). https://doi.org/10.1177/20563051211049246
6. Van Dijk J (2006) The deepening divide. Inequality in the information society. Sage Publications: London, UK
7. Polizzi G (2020) Digital literacy and the national curriculum for England: learning from how the experts engage with and evaluate online content. Comput Edu 152:1–13
8. Radovanovic D, Holst C, Belur S, Srivastava R, Houngbonon G, Noll J (2020) Digital literacy key performance indicators for sustainable development. Soc Inclusion 8(2):151–162
9. Liu Z, Tretyakova N, Fedorov V, Kharakhordina M (2020) Digital literacy and digital didactics as the basis for new learning models development. Int J Emerg Technol Learn 4–18
10. Singh K, Binti N, Jalil A (2019) Relationship of digital literacy and teacher's perception of 21st century learning: a study of SMK Bako KuchinG, Sarawak. In: International conference on education in the digital ecosystem, pp 367–387
11. Michael T (2011). Deconstructing digital natives: young people, technology, and the new literacies. Taylor & Francis. ISBN: 978-1-136-739002
12. Spante M, Hashemi S, Lundin M, Algers A (2018) Digital competence and digital literacy in higher education research: systematic review of concept use. Cogent Edu 5(1):1–21
13. Van Deursen A, Helsper E, Eynon R (2014) Measuring digital skills. From digital skills to tangible outcomes. Project report. Available: https://www.oii.ox.ac.uk/research/projects/measures-and-models-of-internet-use
14. Hargittai E, Hsieh P (2012) Succinct survey measures of web-use skills. Soc Sci Comput Rev 30(1). Available: https://doi.org/10.1177/0894439310397146
15. Orr D (2020) Bologna process in the global higher education arena. Going digital? In: European higher education area: challenges for a new decade, pp 503–515. https://doi.org/10.1007/978-3-030-56316-5
16. Correia AMR, Teixeira JC (2003) Information literacy: an integrated concept for a safer internet. Online Inf Rev 27(5):311–320
17. Prensky M (2001) Digital natives, digital immigrants part 1. On the Horizon 9(5):1–6. https://doi.org/10.1108/10748120110424816.ISSN1074-8121

Food Quality Checking and Scanning System Using Machine Learning with Blockchain Framework—A Survey

V. Mani, T. Sneha, S. M. Star Ajays Singh, P. Thanalakshmi, and K. Vibul Sundar

Abstract The perishable food production network is one of the most challenging areas in the food business; however, supply chain management is improving rapidly to satisfy the rigorous industry standards. With the global COVID-19 outbreak, food safety is having severe problems, especially in terms of customer purchase. On the other hand, people are more concerned with food quality, origin, and shipping regulations. In this context, a food discernibility framework is required for avoiding the over-changing of the stockpile area. The proposed model involves two main components such as information communication system and chemical-based access. In the grant delivery chain, the above listed tools are used to pinpoint, keep in sight, and keep under surveillance of the purchased food items. With the perishable food worldwide exchanges, client conduct, framework proficiency, unwavering quality, and different elements that share the data in the production network have become problematic as the e-commerce industry has grown. Block chain along with other distributed ledger technologies (DLTs) are fore casted in many organizations due to the allocation of storing the dataset that can be switched between organizations that don't trust each other. ML's learning capabilities may be integrated with blockchain-based apps to boost their capabilities and make them smarter. The disseminated record and time upgrade thrives a reasonable information sharing by using ML security methods. Food-trackable order collects data on food details and identifies food through supply chain activities by using the proposed blockchain KNN method in machine learning domain. The proposed research study will review the existing encryption technologies as well as other machine learning methods to show how the proposed blockchain with machine learning algorithm enhances efficiency.

Keywords Blockchain · Cryptography · Machine learning · Distributed ledger technologies · Food traceability

V. Mani (✉) · T. Sneha · S. M. Star Ajays Singh · P. Thanalakshmi · K. Vibul Sundar
Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur 639113, India
e-mail: maniv.cse@mkce.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_43

1 Introduction

Blockchain is a broadly disbursed large technology, wherein transactions are recorded and stored making them compromise with the built-in cryptographic generation. Recent research focuses at the vital precept of blockchain where it may be utilized in regions consisting of data protection change, food safety, product growth, and business acceleration in all retail stores [1, 2]. The concept has real elements in the security structures to stop low levels and counterfeit drugs. Even at the level of individual stocks, blockchain will ensure compliance with the purchase of any product enhancing proof of ownership using specific sources. The concept has real factors within the protection structures to forestall low tiers and counterfeit drugs. Even at the extent of person stocks, blockchain will make certain compliance with the purchase of any product by improving the evidence of ownership with the usage of unique sources [3, 4]. The idea permits transparency in all types of transactions as every time a product change hands transactions can be recorded and validated the use of a unique identity called EPC code it is far provided by using FSSAI, from its production to sales, and this will rightly be stated that the blockchain does no longer exchange, providing consistency, and meaning. When the block is whole, a completely unique code, which binds to the subsequent page or blocks, the advent of a series of blocks will be created. Since the idea is totally based on a stable cryptographic hash technique so that the transaction cannot be duplicated or encrypted, copies are equal, and little permission is needed to get right of entry to the facts [5, 6]. This will substantially reduce the human blunders, extra charges, and time delays and maybe a worldwide procurement application.

2 Related Works

The suggested work [7] is a blockchain model that is entirely based on the discernibility meshwork that provides meals delivery chain identification, protection, and safety. While food adulteration is on the rise in global retail chain, block chain is a contemporary solution for this problem. This methodology covers the complete data series and records the control technique of every object inside the production network for agribusiness, which gives guarding, perception, along with the responsibility for agro-foodstuffs in first-class protection. Also, they can be prolonged in whole meal supplying regions [8]. Today, the need for the edible chain endeavor has definitely increased. Due to the increased demand, the security and quality assurance of a form-to-fork meal has become critical. Expanding threats to feast security and defilement have resulted in the development of an enhanced discernibility device, which remains as an essential tool for delivering excellent control that ensures appropriate inventory network item security. On this composition, we proposed a square

chain primarily based on the arrangement that dispose of the need for centralization structure moderate journals and exchange of records, streamlines in general execution, and conforms to a solid phase of well-being and uprightness [9].

A blockchain generation to comprehend the garage and control of food sampling facts is the proposed work [10]. Though in addition, we brought the visualization strategies to inherit and display riskiness and to assist the discernibility evolution of meals. Those exploration extends modern delving framework of the subsequent components. Furthermore, the blockchain era was used in storing the facts. Since, blockchain capabilities are disbursed, saved information can't to be altered with malevolently. The maximum modern techniques for danger evaluation are partial, and hazard is assessed into numerous stages [11]. Elicited food testing statistics, the reckonable evaluation approach has put forwarded for food handling chance appraisal. This related work demonstrated the technique that offers a systematic foundation for handling the executives, and it lowers the hap of dangers along with protective human being's fitness. Finally, in contrast to many modern-day evaluation strategies that simplest awareness at the consequences, a holistic and specified evaluation method has been adopted to achieve consequences and the motives for his or her prevalence, respectively. On account of conceit strategies, which show statistics, human beings can effortlessly mine a few crucial statistics [12].

The challenges along with the requesting circumstances of present standard food are the proposed work [13]. Then, at that point, through the intensive assessment by the fundamental requests of gadget purchaser, planned a decentralized system evolved from the blockchain and organization such as EPCIS. For protecting the delicate business insights, the association degree shrewd settlement is utilized rather than ordinary exchange of data to keep and control suppers data not withstanding insist the recognizable proof of association. The speed of data transferring to the square chain is as a general rule via the consensus calculation [14]. Thus, it is necessary to enhance the consensus calculation to further develop the framework throughput which lifts up the transferring system. In the recognizability framework, many business venture data are tricky and may handiest be gotten to by utilizing relied upon business partners, alongside personalities and exchange subtleties. The revelation of delicate business insights may happen because of the self-evident and recognizable activity inside a square chain basically based detectability framework. In this manner, it is far indispensable to plan a sort for the records to be transferred at the square chain all together that the well-being of sensitive insights can be ensured [15, 16].

The enforced discernibility for the well-being of merchandise along with their deliver manacles management is the proposed work [17]. All organization embrace to applied science, wherein there is a greatest price increment for his or her deliver chain. Security and standard of suppers are the fundamental contemplations of partners of food assembling or eating place rigid. Blockchain results are more correct and accurate. Blockchain might be utilized to faucet records from every hub of the inventory network [18]. Those verity is ancillary dependable and sealed due to the end product of the cryptographic ledger chain. The measurements caught from these hubs might be verbalized accurately to make specific meals standard. The suggested

model catches garage time information all through the different degrees of foodstuffs garage [19]. Those conditions help in determining the limits among which the dinners top of the line esteem needs to fall at each phase of managing. In advance, the prototype may evolved, in a comparable method that deals with a ramification of records which wishes in dealing with nice issues. The index may be advanced moreover to cowl a massive sort of products and quantify their excellent [20].

A recognized bound conditions for sharing satisfying records to improve traceability four occurrences within the meal contribution change where look into the utilization of a pattern audit of 16 meeting is the proposed work [21]. Eighteen boundary conditions categorized in enterprise law and traceability training have been identified.

A gadget design depends upon the blockchain which perceives the control and security of various jobs is the proposed work [22]. Next, to upgrade framework stockpiling limit, the author proposed a multimode stockpiling system that consolidates supply chain stockpiling, and an allocated information base with staggered reinforcement. Then next, to manipulate enterprise facts and hazard realities, improve data dependability, and decrease risk, we utilized a tweaked astute settlement to control the perusing and composing of facts [23]. The framework which is proposed can harvest the distribution and interchange of realities inside unbroken delivery cycle. It additionally offers solid modern-chain records hotspots for benefactors, customers, and outsider oversight divisions, notwithstanding risk-related assessment, forecast, and early admonition [24]. For instructive examinations, the inspect at the grain production network records control moved via the square chain can offer a whole and genuine information supply, to moreover work with the correspondingly examine commute dominion of perils inside the food stock cycle and underwrite the discouragement and administers the partner of the food production network [25].

A contemporary development approximately blockchain research and packages in agriculture subject is the proposed work [26]. For this motive, the study accounted articles from all fundamental databases and publishing companions. It turned into additionally found that the studies on blockchain are limited to simplest few countries, and maximum the courses took place as suggestion in convention as opposed to journal and e-book chapters [27]. The literature was additionally analyzed beneath four heads of studies dimensions specifically traceability, structure and protection, data structures, and different programs in agriculture field. It became determined that blockchain-based studies in agriculture are especially targeted on traceability and specific studies, whereas incredible research also exists for blockchain structure and security design and blockchain as records gadget [28]. However, few important issues associated with blockchain implementation exist like scalability, interoperability, utility of blockchain for precision agriculture and meals supply chain network design, and so forth which are rising regions for further research.

A harness distributed ledger with the help of Internet of Things to follow the conveyance of transient artifacts is the proposed work [29]. In Agriculture Production Network Executives, the exchange may likewise incorporate the information like sum, uncooked substances, and so on. A few harvesting inclusion plans, like the Public Agrarian Protection Conspire, might be kept up with the use of blockchain

innovation for alter proof records and for intermittent checking of remuneration of cases for the term of yield misfortunes. Utilizing this innovation as a general rule works on the local area among restrictive partners and ranchers [30, 31]. The homes of blockchain basically offer speedup capacity, better well-being, changelessness, printing, faster settlement, and complete recognizability of put away and exchanges information. This paper to explain complete block chain based track ability which enables to construct blocks for agriculture that constantly works with IOT devices from producer to consumer. This gives the work to add in purchaser network framework. The objective focuses to make apportioned record this is convenient by all clients within the network that during roll over brings straightforwardness [32, 33].

A setup to structure utilizing Ethereum blockchain for shrewd account to follow and to carry out endeavor exchanges putting off delegates and significant variable of handling for soybean discernibility all through rural convey network is the proposed work [34]. This paper introduces information and viewpoints identified with the machine design, format, substance connection outline, associations, assortment charts, and execution calculations. This paper gives the answer for following and recognize the chain for soyabean food items. Until now, blockchain era actually meets key requesting circumstances related with versatility, administration, ID enrollment, security, principles, and rules [35]. Also, we planned to implement the programmed bills and proof of transportation within our proposed system—wherein occasions are paid with the utilization of crypto forex which is programmed and brought together way through the savvy contracts that hit real conveyance of plants also items. Cryptographic ledger gleaned from slip of conveyance of actual benefits with programmed charges in digital money just like question taking care of changed into some time ago proposed [36].

A delivered problem along with the difficulties of contemporary standard food recognizability application is the proposed work [37]. Then, at that point, through inside and our investigation of the vital necessities of clients, we planned a dinner's discernibility framework that depends upon EPC innovation and blockchain. For facilitating the figures blast bother, we utilized aggregate administration that provides both on-and-off-bind material to diminish the measure of numbers for a solitary hub [38]. Made a shrewd agreement module for requesting and selling on the manacles, which affirms the proficiency of intersection. Our appraisal has demonstrated the sharp agreement planned in this notepaper to fit well in general execution and changed into accomplishable. Encoding meals by using EPC era utilizes the Ethernet block that fastens and initiates smart settlement to viably perform exchanges, switch all exchanges between members stressed inside the source chain environment [39, 40]. This proposed rank chain utilizes IPFS to save facts under the network, to help the impressive burst inside the supply of the Internet of Things that offers protected, green, and clear food detectable plan.

In existing system, we need centralized system to save the food information. Separate scanning device is used to get food information. Annotation is needed for manual food checking. These are the limitations we found in our existing system. Table 1 shows the comparative study of the related works.

Table 1 Comparative study

Author name	Title	Methodology	Merits	Demerits
1 Deepak Prashar	Cryptographic ledger technology-suggested discernibility and perceivability for agri-products to enhance meals security in India [7]	Depends on the utilization of savvy agreements to screen and deal with all interchanges and exchanges inside the store network among the partners as a whole	It point outs the meaning, safety, and solidity of the food supply chain	Did not support in greater number of clients
2 Zhihao Hao	A customary status track of meals safety risk traceability based instantly on blockchain [10]	Endorse a blockchain primarily based visual analysis approach for meals protection hazard tracking	It gives a scientific warp for handling the executives, keep down the likelihood of vicious	The validity of the consequences is probably considerable it is influenced through sampling mistakes
3 Qijun Lin, Huaizhen Wang	Meal welfare traceability device formed on blockchain and EPCTS [13]	To address the challenges of sensitive information exposure, data manipulation and trust transfer, an innovative enterprise-level smart contract was adopted	It ensures security of information and avoid spam attacks	Difficult to analyze large amount of data
4 Reno Varghese George, et al.	Meal first-rate recognizability prototype for eating bases using blockchain and meal first aid information [17]	Analyzes the maximum broadly used methods of meals traceability and affords a café model for adopting extra dependable meals discernibility	Derive the food quality index	Addition of blocks can be needed
5 Kay Behnke	Restriction ambience for recognizability in meal deliver chain using blockchain innovation [21]	To identify restriction ambience for sharing assurance information to improve traceability	It is adroit to be worn to make the contrive of commodities more obvious	Only, a few supply chain participants are used

(continued)

Table 1 (continued)

Author name	Title	Methodology	Merits	Demerits
6 XIn Zhang	Blockchain primarily based protection control device for the grain deliver chain [22]	Put forwarded a modern day gadgets structure primarily based on technology for the whole grain supply chain	It obtains facts sharing and interchange of information at some point of the grain supply chain	Difficult to give dependable modern chain data sources
7 Vinay Surendra, Yadav	A systematic literature review of blockchain technology in agriculture [26]	The goal is to analyze all key databases, including Web of science and Scopus, for blockchain applications in agriculture	Provide traceability, architecture and security, information systems	Sometimes error can be occurred in between sharing information
8 S. Madumidha	An intellectual execution: agriculture food delivery chain management utilizing blockchain innovation [29]	Hand over a completely decentralized blockchain-based discernibility that empowers to construct blocks for horticulture which persistently coordinates the IoT gadgets from supplier to buyer	Providing traceability and security	Major issue is storing data into blocks
9 Khaled Salah	Blockchain-based soybean discernibility in agricultural delivery chain [34]	Hand over a blockchain-based arrangement and system for discernibility and perceivability in the soybean production network utilizing Ethereum brilliant agreements	Can be concerned to give honorable and regionalized discernibility	Key difficulties connected with adaptability, administration, personality enrollment, protection, principles, and guidelines
10 Haihui Huang, XinZiu Zhou	Meal delivery chain traceability scheme based on blockchain and EPC technology [37]	To modify and hint food in the whole agricultural deliver chain	To decrease the quantity of records that a node ought to save	Less security in node transactions

3 Food Quality Checking and Scanning System

3.1 Food Protection and Traceability Using Blockchain with Gadget Learning Framework

Agriculture has begun to use blockchain technology. In the agri-foods business, blockchain is mentioned as a rising generation that has the ability to provide a green and strong technique for improving food traceability as well as a transparent and reliable approach to confirm agri-food quality, safety, and sustainability [41, 42]. Models can assist in extricating useful information from the client end. Client verification of any approved client is simple when they are attempting to make changes in the blockchain. We can examine the blockchain with a machine learning framework expand productivity in view of the outcomes of this survey.

3.2 Blockchain Algorithm

A blockchain is a digital concept for storing data [43]. These blocks are chained together, and this makes their data irreversible. If a data block is tied to other blocks, its data will not be reversed.

3.3 Block and Hash Generation

- It carries the detail about the larger freshly renovate product.
- Each piece of information generates a hash. The secure hashing algorithm (SHA) is used to generate a hash in blockchain. Each products detail will be stored in the server by the manufacturer at the time of manufacturing. This product details will be stored in a block. Initially, the hash value will be zero, by the time of updating the product details the block will generate a hash of length 256 bit in hexadecimal format. This is how each piece of information generates a hash.
- Product information is given in contrary historical order.
- The hash is determined not just by current information but also by the block's prior hash.
- Even a minor shift in data outcomes in a completely extraordinary hash.
- Nodes check to see whether hash checking has modified any product details.
- If most of the people of nodes homologate the transaction, it is fair written in a block.
- Each block merges to the one before it, forming a chain of blocks. Each block will have a hash ID and previous hash ID. By using this, the blocks are merged to one another as illustrated in Fig. 1.

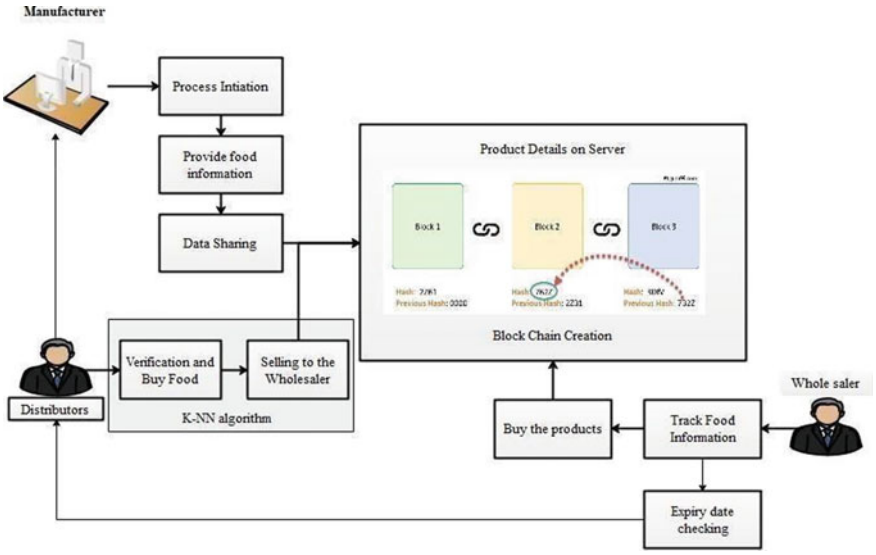


Fig. 1 Proposed system architecture

- Blockchain is active because it is dispersed among several computers, each with a copy [44, 45].

3.4 KNN Classifier

KNN is a method of learning based on examples or lazy reading, in which the work is limited to local estimations, and all calculations are separated [46]. When there is compact or no prior information on data dissemination, KNN is a basic and easy uncoupling procedure. This rule handiest applies to the whole set of exercises. The simplest method for KNN where $K = 1$ is to use the neighborhood neighbors (NN) rule. As an end result, each sample ought to be frivolously disbursed throughout its neighbors. If sample separation is unknown, it can be estimated by looking at sample separation from neighboring neighbors. All distances between the unknown pattern and all of the samples in the schooling set can be determined given an anonymous pattern and training set [47, 48]. The minimum value range corresponds to a training sample paired with an unknown sample. The algorithm includes the steps, and it is as follows: for all unidentified samples. For all non recognized samples, use Unsample (i) and for all recognized samples, use Sample (j). We can calculate the distance by using samples (i) and sample (j). The execution of the KNN separator is basically bounded but the choice of K as well metric range is used. In the manufacture side, 100 training sample is used, and from user side, 80 testing sample is used. Product searching is done by distributor side. QR code scanning is done by user side to track the food information. The choice of k in addition to the metric range employed

determines the performance of the KNN separator. Because the radius of a local region is controlled by the distance of the *k*th neighbor adjacent to the question and different *K* offers opposing for conditional class, the rate is affected by the sensitivity of the neighbor size *K* selection [49, 50]. In KNN algorithm, unknown samples and known samples are used. The known samples are the already saved food information. Here, multiple users are used to scan the QR code, and multiple results are obtained. To best locate these results, we are using KNN classifier in this system.

3.5 Benefits

- Reduces the success rate of attackers [51].
- Provides safety, transparency, and efficiency.
- Avoid outdated products and appropriate notice.
- Eliminate bug management through real-time tracking of food production and resale without double expenditure.
- It alliances quest and retrieval of data.

Figure 1 shows the proposed design that defines the manufacturer adds food details such as name, date, measurement details, ingredients, and more. These facts are stored in a blockchain for security purposes. If distributors purchase food options, use a machine learning algorithm to verify distributor's code and purchase food as well. This is then sold wholesale food and sent a notification of stock details. Finally, the wholesaler tracks food details to check authenticity and distribute to consumers. Wholesaler or distributors cannot modify product verification details. Here, food information tracking is done in a QR code format. In the absence of scanner, we can use camera to track the food from the user side.

4 Conclusion and Future Works

Food security challenges have gotten worse in recent years, and they continue to be a hazard to public health. Keeping track of and tracking the details of various events across the food inventory network, food creation, handling, stockpiling, shipment, and showcasing, are basic. A major response to food handling provokes has been to build up a precise and solid food handling global positioning framework. Present a framework for following the authenticity of food dependent on the subject of food, non-edible food from your production until it is re-sold to customers. This procedure necessitates the collaboration of a large number of individuals, all of whom are bound by smart agreement with the use of blockchain information stockpiling. Also, vendors could assure the accuracy of food information. This program, which aids in the tracking and verification of food, necessitates the use of a single server to track each feed as it is transmitted between multiple participants. The data cutting

option can be used with diets that have a fixed expiration date to decrease the amount of data. To save and manage food data and assure business ownership, we employ a business standard contract rather than traditional performance records to protect critical business information. Hence, the security of data is obtained, and it helps us to avoid spam attacks. The proposed framework can be further extended to implement deep learning algorithm and also implement in handcraft products with android application for quality monitoring.

References

1. Iftekhar A (2020) Application of blockchain and internet of things to ensure tamper-proof data availability for food safety. J Food Qual 7:1–14
2. Tipmontian J, Alcover JC, Rajmohan M (2020) Impact of blockchain adoption for safe food supply chain management through system dynamics approach from management perspectives in Thailand. Multidiscip Dig Publish Inst Proc 39(1):1–5
3. Murugesan M, Thilagamani S (2020) Efficient anomaly detection in surveillance videos based on multilayer perception recurrent neural network. J Microprocess Microsyst 79:1–10
4. Thilagamani S, Nandhakumar C (2020) Implementing green revolution for organic plant forming using KNN-classification technique. Int J Adv Sci Technol 29(7S):1707–1712
5. Tan A, Ngan PT (2020) A proposed framework model for dairy supply chain traceability. Sustain Futures 2:1–6
6. Tao Q (2019) Food safety supervision system based on hierarchical multi-domain blockchain network. IEEE Access 7:51817–51826
7. Prashar D (2020) Blockchain-based traceability and visibility for agricultural products: a decentralized way of ensuring food safety in India. Sustainability 12(8):1–20
8. Pandiaraja P, Sharmila S (2020) Optimal routing path for heterogenous vehicular adhoc network. Int J Adv Sci Technol 29(7S):1762–1771
9. Santhi P, Mahalakshmi G (2019) Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction. Int J Eng Adv Technol 8(4):839–846
10. Hao Z (2020) A novel visual analysis method of food safety risk traceability based on blockchain. Int J Environ Res Public Health 17(7):1–18
11. Eryilmaz U (2020) Traceability blockchain prototype for regulated manufacturing industries. In: 2nd International electronics communication conference, pp 9–16
12. Vikaliana R (2020) Agribusiness commodity traceability using blockchain technology: a systematic review. Test Eng Manage 83:6503–6508
13. Lin Q (2019) Food safety traceability system based on blockchain and EPCIS. IEEE Access 7:20698–20707
14. Deepa K, Thilagamani S (2019) Segmentation techniques for overlapped latent fingerprint matching. Int J Innov Technol Explor Eng 8(12):1849–1852
15. Tian F (2017) A supply chain traceability system for food safety based on HACCP, blockchain & Internet of Things. In: International conference on service system service manage, pp 1–6
16. Rajesh Kanna P, Santhi P (2021) Unified deep learning approach for efficient intrusion detection system using integrated spatial-temporal features. Knowl-Based Syst 226(1):1–12
17. George RV (2019) Food quality traceability prototype for restaurants using blockchain and food quality data index. J Clean Prod 240:1–25
18. Smys S, Wang H (2021) Security enhancement in smart vehicle using blockchain-based architectural framework. J Artif Intell 3(2):90–100
19. Sivaganesan D (2021) Performance estimation of sustainable smart farming with blockchain technology. IRO J Sustain Wirel Syst 3(2):97–106

20. Joe CV, Raj JS (2021) Deniable authentication encryption for privacy protection using blockchain. J Artif Intell Capsule Netw 3(3):259–271
21. Behnke K, Janssen MFWHA (2020) Boundary conditions for traceability in food supply chains using blockchain technology. Int J Inf Manage 52:1–10
22. Zhang X (2020) Blockchain-based safety management system for the grain supply chain. IEEE Access 8:1–13
23. Pandiaraja P, Aravinthan K, Lakshmi Narayanan R, Kaaviya KS, Madumithra K (2020) Efficient cloud storage using data partition and time based access control with secure aes encryption technique. Int J Adv Sci Technol 29(7S):1698–1706
24. Zhang X, Lv S, Xu M, Mu W (2010) Applying evolutionary prototyping model for eliciting system requirement of meat traceability at agribusiness level. Food Control 21(11):1556–1562
25. Pradeep D, Sundar C, QAOC (2020) Noval query analysis and ontology-based clustering for data management in Hadoop. Future Generation Comput Syst 108:849–860
26. Yadav VS, Singh AR (2019) A systematic literature review of blockchain technology in agriculture. In: International conference on industrial engineering and operations management, pp 1–9
27. Liu W, Cheng W, Zhang J (2013) Study on the traceability system establishment of safety-objective-oriented food logistics supply chain. Adv J Food Sci Technol 5(4):492–499
28. Logeswaran R, Aarthi P, Dineshkumar M, Lakshitha G, Vikram R (2019) Portable charger for handheld devices using radio frequency. Int J Innov Technol Explor Eng (IJITEE) 8(6): 837–839
29. Madumidha S (2019) A theoretical implementation: agriculture-food supply chain management using blockchain technology. In: TEQIP III Sponsored international conference on microwave integrated circuits, photonics and wireless networks (IMICPW). IEEE Press, pp 1–8
30. Deng M, Feng P (2020) A food traceability system based on blockchain and radio frequency identification technologies. J Comput Commun 8(9):17–27
31. Gunasekar M, Thilagamani S (2020) Performance analysis of ensemble feature selection method under SVM and BMNB classifiers for sentiment analysis. Int J Sci Technol Res 9(2):1536–1540
32. Abad E, Palacio F, Nuin M, Zárate AGD, Juarros A, Gómez JM, Marco S (2009) RFID smart tag for traceability and cold chain monitoring of foods: demonstration in an intercontinental fresh fish logistic chain. J Food Eng 93(4):394–399
33. Galvez JF, Mejuto JC, Simal-Gandara J (2018) Future challenges on the use of blockchain for food traceability analysis. TrAC Trends Anal Chem 107:222–232
34. Salah K (2019) Blockchain-based soybean traceability in agricultural supply chain. IEEE Access 7:73295–73305
35. Castro-Puyana M, Pérez-Míguez R, Montero L, Herrero M (2017) Application of mass spectrometry-based metabolomics approaches for food safety, quality and traceability. TRAC Trends Anal Chem 93:102–118
36. Deepika S, Pandiaraja P (2013) Ensuring CIA triad for user data using collaborative filtering mechanism. In: International conference on information communication and embedded systems (ICICES), pp 925–928
37. Huang H, Zhou X, Liu J (2019) Food supply chain traceability scheme based on blockchain and EPC technology. In: International conference on smart blockchain. Springer, Cham, pp 1–11
38. Arya V, Naganathahalli P, Shukla S (2019) A blockchain framework for proptech: Success model through disintermediation and self- regulation. In: International conference on intelligent computing, information and control systems, pp 522–528
39. Rajesh Kanna P, Santhi P (2022) Hybrid intrusion detection using map reduce based black widow optimized convolutional long short-term memory neural networks. Expert Syst Appl 194:1–7
40. Lin Q, Wang H, Pei X, Wang J (2019) Food safety traceability system based on blockchain and EPCIS. IEEE Access 7:20698–20707

41. Thilagamani S, Shanti N (2014) Gaussian and gabor filter approach for object segmentation. J Comput Inf Sci Eng 14(2S):1–7
42. Pandiaraja P, Suba S (2022) An analysis of a secure communication for healthcare system using wearable devices based on elliptic curve cryptography. J World Rev Sci Technol Sustain Dev 18(1):51–58
43. Wang Y, Yang Y (2020) Research on agricultural food safety based on blockchain technology. J Phys 12(1):62858–2865
44. Deepa K, Kokila M, Nandhini A, Pavethra A, Umadevi M (2020) Rainfall prediction using CNN. Int J Adv Sci Technol 29(7S):1623–1627
45. Liu F, Wang Y, Jia Y, Hu S, Tu L, Tang C (2015) The egg traceability system based on the video capture and wireless networking technology. Int J Sensor Netw 17(4):211–216
46. Gopalakrishnan PK, Behdad S (2019) A conceptual framework for using video grammetry in Blockchain platforms for food supply chain traceability. In: International design engineering technical conferences and computers and information in engineering conference, pp 1–12
47. Rohmah D (2019) Traceability and tracking systems of halal food using blockchain technology to improve food industry competitiveness. In: 1st International conference on business, law and pedagogy, ICBLP. European Alliance for Innovation (EAI), Indonesia, pp 1–9
48. Wang W, Kumar N, Chen J, Gong Z, Kong X, Wei W, Gao H (2022) Realizing the potential of internet of things for smart tourism with 5G and AI. IEEE Netw 34(6):295–301
49. Xiao X, Fu Z, Qi L, Mira T, Zhang X (2015) Development and evaluation of an intelligent traceability system for frozen tilapia fillet c processing. J Sci Food Agric 95(13):2693–2703
50. Aung MM, Chang YS (2014) Traceability in a food supply chain: safety and quality perspectives. Food Control 39:172–184
51. Bosona T, Gebresenbet G (2013) Food traceability as an integral part of logistics management in food and agricultural supply chain. Food Control 33(1):32–48

Blockchain-Based Web Framework for Real Estate Transactions

Rajan Khade, Amit Pandey, Aditya Shinde, and Neha Deshmukh

Abstract The onset of blockchain generation within Bitcoin has generated significant interest by showing an opportunity to eliminate the middle ground need and transform communication between people and machines by increasing trust. Initially restricted to the integrated currency domain, people began to see the power of a generation beyond just the cryptocurrencies, which brought the acceptance of the blockchain era to erase the world's problems. One such situation is problems for e-governance companies in other areas of the public sphere. In the scope of this thesis, we have specifically addressed the issues within the traditional property registration system. This thesis discusses the new design and architecture for real estate transactions and implements it using a blockchain-based solution and addressing issues including record integrity, privacy, and most importantly the lack of common platforms among concerned government organizations. The advent of the blockchain era led to the creation of blockchain-enabled platforms like Hyperledger Fabric. It is one of the most popular open-source permissioned blockchain frameworks, created and supported by the Linux Foundation and IBM, used in many industrial scenarios. So, it is used to create a network with one ordering organization with an ordering node and one peer organization with two peer nodes to prove the concept. Chaincode similar to Ethereum's smart contract contains the logic to perform all operations and modify the ledger data. All the methods of chaincode are accessed using the Fabric gateway in the web application to perform various operations.

Keywords Blockchain · Hyperledger Fabric · Chaincode · Permissioned blockchain · Real estate transactions · Land records

R. Khade (✉) · A. Pandey · A. Shinde · N. Deshmukh
Department of IT Engineering, A. P. Shah Institute of Technology, Thane, India
e-mail: rajankhade@apsit.edu.in

A. Pandey
e-mail: amitpandey@apsit.edu.in

A. Shinde
e-mail: adityashinde@apsit.edu.in

N. Deshmukh
e-mail: npdeshmukh@apsit.edu.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_44

1 Introduction

Technology in its relationship to globalization has always been at the forefront of disrupting standard issues and producing a brand-new clean look in certain objects. Blockchain is one of those technologies; it is a disruptive innovation with the potential to revolutionize the way governments and other nonprofit, or for-profit organizations handle themselves, as well as how they communicate with collaborating parties [6]. However, it was not very widespread until a person or a group going by the name of Satoshi Nakamoto in 2009 created the Bitcoin cryptocurrency [9].

According to the general public-related industries, technological changes are always in line with integration guidelines and are provided with an amazing deal of problems and even growth. In this thesis, we delve deeper into the knowledge of property registration while conveying it to the Indian context and current related issues. The primary purpose of this concept is to show how integration with blockchain technology is the best solution. We can look at blockchain era property registration technology that makes it an effective solution to current problems, and finally, we can present to you the results obtained from all the tests of the system used. One of the significant challenges facing various provinces in India is the issue of land ownership, which is due to the fact that there is no land records management system to date. The issue of land ownership is so important, because almost all financial institutions rely heavily on the ownership of land-based assets to obtain collateral. Due to uncertainty about patent applications, financial institutions are inefficient and hinder the country's development.

Land ownership in India is primarily based on deeds of sale (acquired at the time of local registration) and other deeds. It involves the following steps:

1. Buyer—Verification of ownership
2. Buyer—Pay negotiated amount to seller
3. Buyer—Preparation of sell deed
4. Buyer—Payment of stamp duty and registration charges
5. Seller—Sign the sell deed
6. Registrar—Execution of sell deed

Both buyer and seller need to follow the workflow to complete a property registration. However, not all of these documents confirm the ownership claim but instead confirm the fact that the transfer of the property will be recorded, and these are the recorded activities known as "world records". The term "world records" is a common term and conclusion of many other documents that provide property maps, the record of rights (ROR), sales documents, and much more information and is organized into various categories. Due to the lack of good governance and communication in departments, we often find inconsistencies in records, outdated and unscheduled data. This discrepancy is the cause of the patent problem as many years of records have to go back to confirm the ownership claim. The need for clear ownership and determining land ownership is explained in this thesis. So, over the past three decades, various schemes have been used to revitalize world records to move from the ownership of imaginative sites to the clearing and clarification of land ownership.

1.1 Blockchain

Blockchain is widely distributed to make business processes safer and more transparent. Blockchain is actually a series of blocks connected by a cryptographic hash, as the name implies [14]. Each block contains data structure and a cryptographic hash of the previous block, activity data, and timestamp. Consensus algorithms are defined in intelligent blockchain protocols, which are programs that automatically perform predefined actions under certain conditions.

1.2 Peer-to-Peer Network

Computers connected to each other without servers should form a peer-to-peer network, which increases security because there is no single point of failure. Transactions are accepted on the block only if all partners agree and all peers have access to all network transaction records. In this thesis, literacy access is given to real estate owners and restricted access to government officials responsible for keeping records of land records.

One of the general motives for the authorities to introduce blockchain-based system is the fact that it will increase the speed of registration without compromising the security and integrity of the data [15]. With the evolution of information and communication technologies (ICT), government authorities are challenged to provide more efficient and effective services to citizens, yet guaranteeing a high level of data security, transparency, auditability, and privacy [6]. As discussed by authors of [12], blockchain can be used in improving the security of smart vehicles; a similar approach can be used for IoT devices. When real estate transactions are managed by using blockchain, we do not need to depend on third parties to verify them, trust is of much less concern when the registration is forced to be done in a transparent way, and cryptographic protocols are used to check the integrity of the information supplied by the person [13].

2 Review of Related Work

The authors of [13] proposed a solution using Ethereum blockchain, a permissionless blockchain, to implement the proposed solution, and the working logic is present in the smart contract. The end users will use a web application, and they need to use the Metamask wallet to perform any operation and pay some ether cryptocurrency as gas. Because, as of now, Ethereum works on a proof of work consensus mechanism, the blockchain miners charge the gas to ensure that the smart contract is executed and the transaction is recorded in the blockchain. Although the decentralization of the

system is achieved, the use of cryptocurrency and crypto wallets makes it inefficient and unnecessarily difficult to use for naive users.

The authors of [10] proposed a solution using the Ethereum blockchain and smart contract written in solidity. The core principle used is that the physical record for every land is converted into a certain amount of tokens based on the ERC20 token standard, so the tokens owned for a unique ID of land represent the ownership of the land. When the ownership is transferred, the tokens with the unique ID are transferred to a new owner. The token-based ownership makes it easy to keep records of shared ownership. Still, the major disadvantage in this solution is similar to the previously mentioned paper, i.e., wastage of the Ethereum cryptocurrency for every transaction. It will make use of this solution expensive for the general public.

The authors of [8] proposed a solution using Hyperledger Fabric (for simplicity Fabric) and InterPlanetary File System (IPFS). Fabric is a permissioned blockchain, and it does not require any cryptocurrency or mining to perform any operations. The working logic is present in chaincode, which is similar to the smart contract of Ethereum. The IPFS is used to store documents in a decentralized way. Although the authors have proposed a more cost-efficient solution than previously discussed solutions, they have not truly digitized the process. Since, to complete the registration process, there is a need to upload a scanned paper-based document.

The authors of [11] have discussed a potential solution with Multichain technology. The core principle is ensuring the integrity and availability of the documents used in the land registration process. The generated hash of the document is stored when the document is uploaded, and it is used whenever the user needs to verify the integrity and authenticity. The problem with the solution is the involvement of documents because it makes representing and accessing data of land or transfer rugged. Also, it risks privacy by exposing unwanted details to the public.

Digital India Land Record Modernization Programme (DIL-RMP) [2, 16], an initiative by the Government of India, mainly focuses on digitizing land records. A separate online web portal is available for citizens to access electronic certificates of land records and other related documents. This cuts down the time for accessing records as it reduces the interaction between the citizen and the government officials. This curtails a significant number of fraudulent transactions and acts as evidence in case of disputes. Likewise, Land Records Mission Mode Project (MMP) [7] is a program under National e-Governance Plan (NeGP) which aims to accelerate easy maintenance and updating of land records by keeping track of mutation cases. However, this lacks the digitization of the land registration process since the preparation and execution of sale deeds happen in a paper-based format. So, the land record needs to be updated manually, making the whole process time consuming and error prone. Since the data storage is centralized, it poses threats to the privacy and integrity of the data.

After reviewing the already proposed solutions, it was observed that currently, no solution could truly decentralize real estate transactions and, at the same time, digitize the complete process. The solution should be secure and fast, and it should be easy to use for the general public. Although using Ethereum, Fabric, and Multichain technologies has its advantages and disadvantages, the major disadvantage for Ethereum

is the high transaction costs. The cost of each transaction is high, and the frequency of transactions is low. The Multichain technology provides a great blockchain platform, but the cost of implementing the solution in production is very high. On the other hand, Fabric is open source and provides full flexibility over blockchain architecture and scalability. With Fabric, there is no need to use cryptocurrency and no cost for executing the transactions.

3 Existing System

In this section, we will describe the current state of local registration and land transfer in India and in particular the state of Uttar Pradesh. We have found that almost all the provinces of India have adopted the same workflow, so we can assume that the results we have given for the state of Uttar Pradesh can be extended to other provinces. In almost every province of India, we have found that usually two or three parts carry land titles [4] and so on.

1. Tickets and Registration Category—Department of Stamps and Registry (DoSR): Manages document verification during registration. There is also an additional obligation to collect stamp duty for the deed of sale to be registered.
2. Finance Department or Finance Board—Board of Revenue (BoR): Handles the preparation and maintenance of transactional records or transaction details.
3. Department of Surveillance and Residence: Responsible for the preparation and maintenance of maps.

In this article, we ignored the Survey Division and Settlement Division as we found that the irregularities were large and there was no record of geographical mapping. For the purposes of this thesis, we exclude title deeds relating to land/property registration. Listed below are the different types of opportunities where you can find a stable place in India:

1. Sell
2. Will/Gift/Swap
3. Inheritance
4. Procurement
5. Provided by the government or a court

All of the above transactions go under different types of actions and thus differ in the course of the transaction. In this sense, the focus is on the registration of acquired assets. In the state of Uttar Pradesh, as shown in Fig. 1, the three-dimensional framework is in place in the DoSR [4].

This divine structure was later used to describe the role-based authority information contained in the blockchain, thus limiting access to sensitive information. The office of sub-registrar (SRO) is the place where all registration activities take place. Various office functions are provided below [4]:

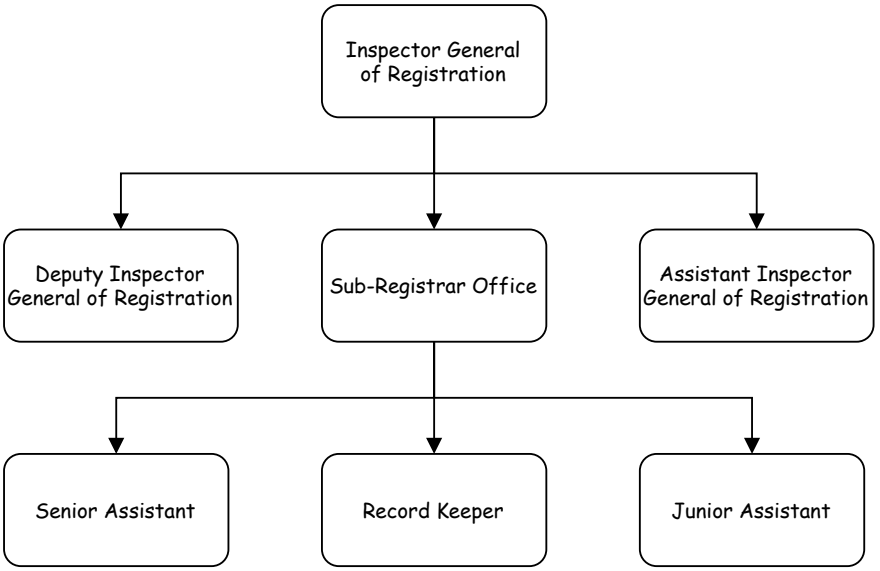


Fig. 1 Hierarchical structure of DoSR

1. Estimation of property value
2. Collection of stamp duties
3. Registration of title deed
4. Maintenance of copies of title deeds
5. Issuing registration certificates

Listed below are the steps during the registration of immovable property in Uttar Pradesh:

1. Asset Verification: This feature of registration usually lasts until the consumer's title is carefully considered. There are many variations of land rights in India and switching to online sites has not yet resolved the issues, so it is up to the buyer to verify ownership of the previous property.
2. Property Valuation: The value of a property is measured by the payment of the registration stamp tax, which is calculated as a percentage of the circular average for that area.
3. Preparation of Stamp Paper: Non-judicial stamp paper equal to the value of stamp tax is required for purchase, now available online.
4. Stamp Work Payment: This section can now be done online
5. Registration to the SRO: This part of the registration process requires personal intervention from the SRO, as well as buyer and seller witnesses for verification.

As mentioned earlier, in the state of Uttar Pradesh, the BoR is responsible for the preparation and maintenance of transaction details. Departmental activities vary widely in the area where property is immovable, i.e., whether the property belongs to the urban or rural category.

4 Proposed Solution

We now propose a blockchain-based framework to execute real estate transactions. It involves the Fabric network architecture as well as a chaincode. It will enable users to perform all the operations digitally and temper proof. The entities involved and working of the framework are based on the existing system in place in India to perform real estate transactions. The users of this system will be the general public and government officials.

4.1 Considerations

There are a few primary processes that are needed to be considered in order to make the framework ready for real-world scenarios. In order to allow users to use this framework, all the users need to register themselves on the system, which can be done using sign-up functionality and later verifying the identity. Also, every sub-registrar needs to register all the estates under his jurisdiction on the system. If there is a need to put old record on the system, then local government officials who keep records of old real estate transactions will need to insert all the old transaction details in the system. After this, the framework will be ready to use. So, to explain the system's working, we have considered that all the procedures mentioned earlier are completed beforehand.

4.2 Architecture

Orderer

The ordering service will be implemented and administered by the central government. The orderer will run the Raft consensus mechanism with multiple nodes, which can be achieved even if the nodes are present at different physical locations. Raft is a crash fault-tolerant (CFT) consensus mechanism used for the ordering service, and it is based on an implementation of Raft protocol in etcd key-value store [5].

Member/Peer Organizations Every state of the country will be a separate member organization in the network, and it will act as a membership service provider (MSP) for that organization's entities.

Peer Every member organization will take the responsibility of setting up peers for the use of their organization and joining the respective channel on every peer. The peers will involve anchoring, endorsing, and committing peers.

Ledger Every peer has its own copy of the ledger that the peer maintains. It contains world state (data in the form of key-value pairs) and blockchain (log of transactions responsible for current world state). There are currently two options to use for ledger databases, LevelDB and CouchDB, but by default, LevelDB is used.

Channel Every state government will be present in a separate channel, and one common participant in every channel will be the central government.

Certification Authority (CA) A root CA will distribute digital identities to peers, orderer nodes, client applications, and administrators. The central government will administer the root CA to secure the blockchain network from unwanted access.

Chaincode Every state government follows different rules to perform real estate transactions, so it will not be feasible for every state to implement the same chaincode. Fabric offers a solution to this problem by allowing every organization to develop and write the chaincode according to their needs. Then they can install the chaincode on endorsing peers, so they can use it to connect them to applications using Fabric gateway.

Application To interact with smart contract and blockchain network, organizations will use the application software development kit (SDK) to develop the applications using the technology of their choice. The general public and government officials will use the application to execute real estate transactions.

In Fig. 2, the "Org 0" is the central government, i.e., the orderer organization. "Org 1" and "Org 2" are two state governments having their separate channels, "Channel 1" and "Channel 2", respectively. Both "Org 1" and "Org 2" have their peers in their respective channels. Applications A1, A2, and peers "Peer 1" and "Peer 2" belong to "Org 1" and "Org 2", respectively. According to their need, all the organizations can have more than one peer and application. Every peer has its copy of the ledger, and every endorsing peer has its instance of the chaincode.

4.3 Security

The security of the proposed framework is handled in two parts, i.e., security of blockchain and security of web server.

Security of web server The web server uses Transport Layer Security (TLS) to secure the communication between the web server and users by encrypting the information.

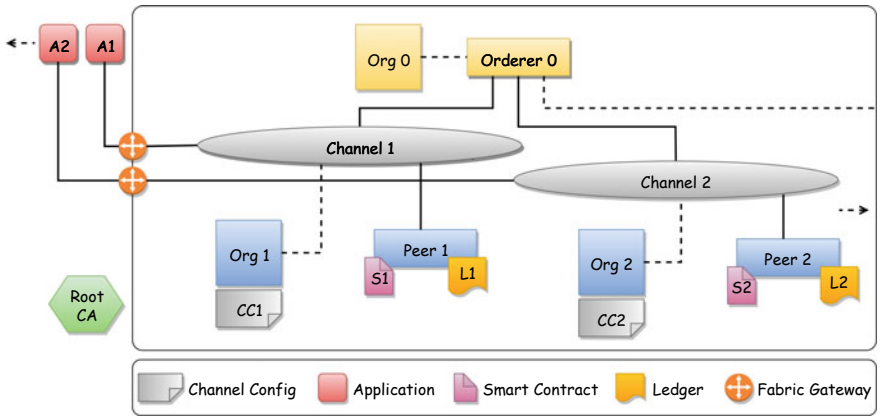


Fig. 2 Proposed fabric architecture

There is no data saved on the web server except users' sessions to prevent data from being accessed by intruders in case the security of the web server is compromised. The users must complete the multifactor authentication (MFA) process and identity verification while signing up or logging in. It ensures that only legitimate users use the web application. Similarly, users need to go through MFA while performing the estate's final transfer.

Security of blockchain The integrity of data in the blockchain is not a primary concern. As more nodes are added to the network, the probability of any unwanted entity modifying the data drops exponentially. But, to secure the data in transit between nodes, TLS is used. As discussed by the authors of [1], the security of Fabric can be improved further without affecting performance.

4.4 Chaincode Development

The significance of discussing chaincode development is that, since the framework is considered a solution for the whole country, if any state has special requirements or needs any different workflow, then the only thing they need to modify is the chaincode. Chaincode development involves developing the framework's logic for executing real estate transactions. The first important part is choosing the programming language to develop the chaincode. There are currently three programming languages to choose from: Go, Java, and Node.js. The optimal choice is Go because it is the native language to Fabric, and according to many analyses published on all three languages, Go is the best choice. The results show that Go is the most performing one for almost all the tests realized and that in the case of the update of the ledger, the latency follows a linear trend as the number of nodes in the network increases. In contrast, in the case of query, the latency is approximately constant [3].

The structures are used as the template for storing data related to Inspector General of Registration (IGR), sub-registrar, users, and estate transaction records. Following is the example of structure used for user data:

```
type User struct {
    Password_Hash string    `json:"passwordSalt"`
    Password_Salt string    `json:"passwordHash"`
    UID           string    `json:"uid"`
    Name          string    `json:"name"`
    Status        int       `json:"status"`
    Owned_Estates []string  `json:"ownedEstates"`
}
```

The data in the ledger is stored as key-value pairs, and the keys are used to add, update, and read data. The pattern used for keys for users is "user_" + UID, where UID is a unique identification number, i.e., Aadhaar number of a user. Similarly, structures and key patterns are used for storing and retrieving the data, such that data can be mapped uniquely. The chaincode has many functions to perform various operations like making land available to sell, requesting to buy a new estate, accepting buyers' requests, and other functions.

```
func (s *SmartContract) RequestToBuy_Estate
    (ctx contractapi.TransactionContextInterface,
    buyer_uid string, name string, ulpin string,
    proposedPrice int, dateTime string)
    error {
    // logic
}
```

The "InitLedger" function initializes the ledger with the credentials of IGR, then IGR will create the office wise sub-registrar accounts to administer the real estate transactions.

5 Experimentation and Results

The implementation is only for a single state of the country, so it involves the central government and one state government. The setup will be similar for every other state. The thing that may or may not differ is the chaincode, because a state government may have some additional requirements so that they may make the necessary changes. The implementation involves deploying all the nodes involved in the Fabric network in docker-based containers. The docker containers are launched using docker-compose. Instead of using a root CA, cryptogen tool is used to simplify the process. The cryptogen is a utility for generating Fabric key material, and it takes the "crypto-config.yaml" file that contains the information about the entities for which the digital certificates or identity is needed to access the blockchain network. It is not recommended to use cryptogen in the production network.

The Fabric network is started using docker-compose, and it uses the configuration specified in the "docker-compose.yaml" file. It launches a single orderer node in

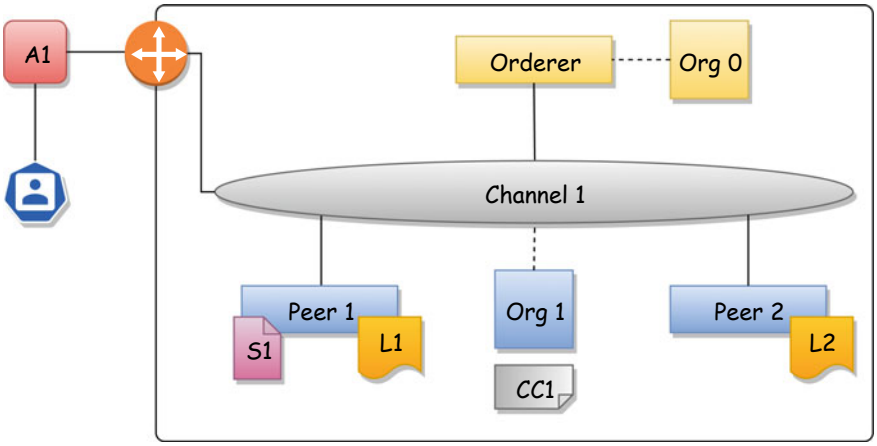


Fig. 3 Implemented fabric architecture

Raft consensus, two peers, and a CLI container. The peers belong to the member organization. The CLI container is used to interact with orderer and peers. Figure 3 illustrates the blockchain network design used for implementation.

"Peer 1" is anchor, endorsing, and committing peer; "Peer 2" is only a committing peer. The chaincode is installed on "Peer 1". The Fabric gateway connects the web application "A1" to the blockchain network; the gateway enables the web application to submit the chaincode transaction requests to peers, wait for the response, and many other operations. The web application exposes the RESTful API which is used in the front end of the application. All the users of the real estate transaction framework like IGR, sub-registrars, and the general public will use the web application to perform the various operations implemented in the smart contract. The users can only see the simple user interface of the web application; they are not bothered by the complexity of the Fabric technology.

The workflow to follow while executing a real estate transaction is as follows:

1. Buyer—Verification of ownership
2. Buyer—Send request with proposed price to seller, using Unique Land Parcel identification Number (ULPIN) of estate
3. Seller—Accept the request
4. Buyer—Pay stamp duty and registration charges
5. Registrar—Verify details and execute the transaction

Workflow involves a seller, a buyer, and a sub-registrar. If the status of the property is set to "available for sell", then any buyer can request to buy the property with the price he wants to propose. If the seller accepts that request, the buyer must pay the registration charges to move forward in the process. After this, the request will go to the sub-registrar under whose jurisdiction the property is present, and then he can verify and execute the final step. Immediately after that, the ownership of the

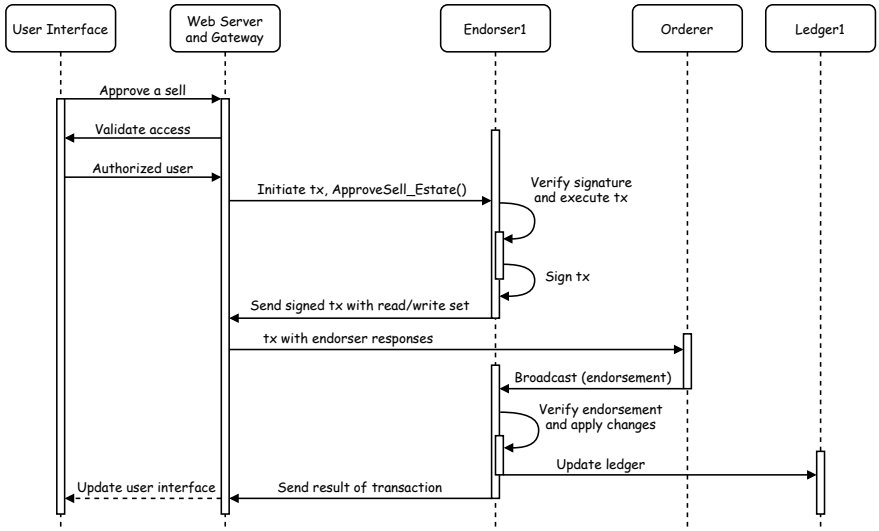


Fig. 4 Sequence of events to achieve consensus, when the estate is transferred

property will change, and the transaction will be recorded on the blockchain for future reference. In every step of the workflow, the consensus is achieved for achieving transparency and integrity. A simple example of the consensus process is shown in Fig. 4, and it consists of all sub-processes involving entities of the blockchain network.

Every operation performed by a user on the web application is completed by submitting a query or a invoke request for a chaincode function, using Fabric gateway on the web server. The query request is useful for only reading the data from the ledger, so it is not recorded on the blockchain, which is why it is much faster than invoke. On the other hand, invoke request is useful for reading as well as writing and updating the ledger data, so it is recorded on the blockchain. It makes it slower because an operation will be complete when the consensus process is over. Both query and invoke requests serve different purposes, so comparing them is not very meaningful, but using them efficiently helps make the system efficient. The framework is tested on Ubuntu (20.04.3) with 8 GB RAM using a docker-based implementation of Fabric architecture (Fig. 3). As illustrated in Fig. 5, the results obtained show how increasing the number of endorsing nodes (assuming that orderer nodes will be increased accordingly) can improve the throughput of the system, i.e., users' operations per second (OPS). It shows that scaling the nodes can largely improve the performance.

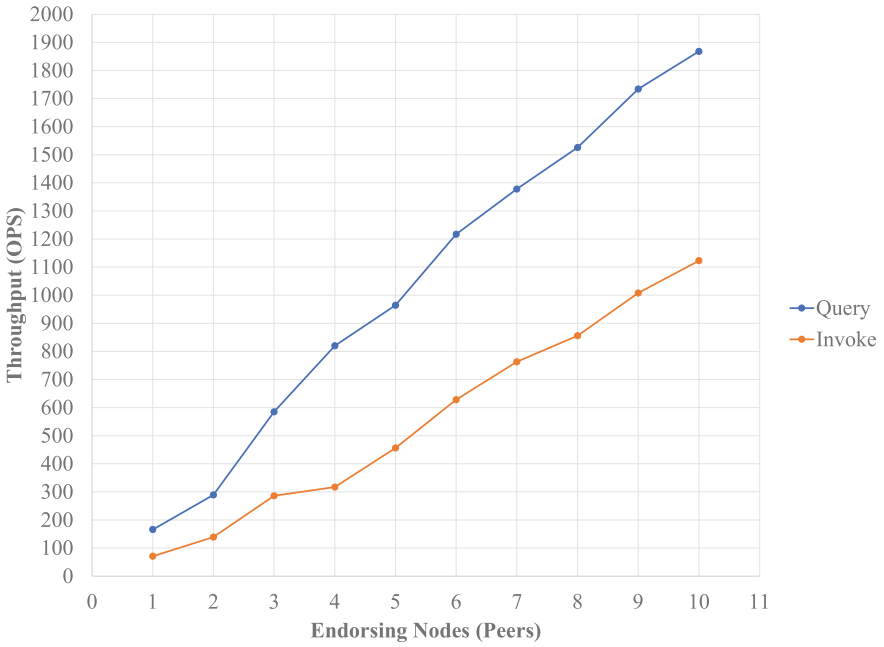


Fig. 5 Throughput for operations involving query and invoke requests

Table 1 Average latency for operations involving invoke requests

Endorsing nodes	Orderer nodes	Invoke (OPS)	Avg. latency(s)	Max latency(s)
1	1	71	0.34	0.96
2	1	139	0.47	1.02
3	2	286	0.60	1.25
4	2	317	0.73	1.37
5	2	456	0.86	1.51
6	3	628	0.99	1.66
7	3	763	1.12	1.80
8	3	856	1.25	1.95
9	3	1008	1.38	2.09
10	3	1123	1.51	2.24

After further analysis of the results, as detailed in Table 1 demonstrated that scaling the nodes increases the latency of operations since the majority endorsement policy is used in the consensus mechanism, because more and more endorsing nodes need to verify the transaction before the ledger is updated. So, to achieve the best performance for the framework, the balance between horizontal scaling and vertical scaling is needed, i.e., number of nodes and capacity of a single node.

6 Conclusion

This paper presents a modern solution to prevent and overcome the problems the current World Record Management System in India is facing. The installation of blockchain in the system makes it secure, decentralized, and free from fraud. The program provides access to all stakeholders involved in world record keeping and builds a strong World Record Management platform. Distributing it via blockchain makes the record transparent, validated, and secure. These records can be used as tangible evidence in any land-based legal proceedings or disputes.

Therefore, we have proposed an easy-to-use and seamless platform that can be used to facilitate property registration. In current system, there are many problems such as vendor engagement or intermediate involvement and time delays. This framework will eliminate the problems associated with property registration in India and in many parts of the world. Digitally registering land titles will not only simplify the process but also protect land titles from various man-made and natural disasters. Blockchain technology is emerging very quickly due to the secure features it provides. So, using a blockchain to save world record transactions is a way to create static records. Nowadays, land is not a commodity, the scope is therefore broad, and there are many use cases for the created platform. However, for any successful implementation of the technology, seamless acceptance by stakeholders is required. Therefore, the main challenge for this solution will be to move existing World Record Management Systems to the proposed one.

Future Work Many additional features can be added to this framework, like more details of the property and its geography. The chaincode can be improved and tested to make it more efficient. There is a need to handle cases, where the property has shared ownership. To make the framework a complete solution, a native cryptocurrency can be added to do all types of payments.

Acknowledgements We would like to extend our deepest gratitude to the Department of Information Technology Engineering at A. P. Shah Institute of Technology, Thane, for providing us with all the resources, support, and guidance needed for this thesis.

References

1. Brotsis S, Kolokotronis N, Limniotis K, Bendiab G, Shiaeles S (2020) On the security and privacy of hyperledger fabric: challenges and open issues. In: 2020 IEEE World Congress on Services (SERVICES), pp 197–204. https://doi.org/10.1109/SERVICES48979.2020.00049
2. Dilrmp.in: Digital india land record modernization programme. https://dolr.gov.in/programme-schemes/dilrmp/digital-india-land-record-modernization-programme
3. Foschini L, Gavagna A, Martuscelli G, Montanari R (2020) Hyperledger fabric blockchain: chaincode performance analysis. In: ICC 2020—2020 IEEE International Conference on Communications (ICC), pp 1–6
4. Gunda A, Shukla S (2019) Property registration and land record management via blockchain. Master's thesis, Indian Institute of Technology Kanpur

5. Hyperledger-fabric.readthedocs.io: Hyperledger fabric documentation. https://hyperledger-fabric.readthedocs.io
6. Lazuashvili N, Norta A, Draheim D (2019) Integration of blockchain technology into a land registration system for immutable traceability: a casestudy of georgia. In: Business process management: blockchain and Central and Eastern Europe Forum. Springer International Publishing, Cham, pp 219–233
7. Meity.gov.in: Land records—ministry of electronics and information technology. https://meity.gov.in/content/land-records
8. Mukne H, Pai P, Raut S, Ambawade D (2019) Land record management using hyperledger fabric and ipfs. In: 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp 1–8. https://doi.org/10.1109/ICCCNT45670.2019.8944471
9. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf
10. Nandi M, Bhattacharjee RK, Jha A, Barbhuiya FA (2020) A secured land registration framework on blockchain. In: 2020 third ISEA conference on security and privacy (ISEA-ISAP), pp 130–138. https://doi.org/10.1109/ISEA-ISAP49340.2020.235011
11. Ramya UM, Sindhuja P, Atsaya R, Bavya Dharani B, Manikanta Varshith Golla S (2019) Reducing forgery in land registry system using blockchain technology. In: Luhach AK, Singh D, Hsiung PA, Hawari KBG, Lingras P, Singh PK (eds) Advanced informatics for computing research. Springer Singapore, Singapore, pp 725–734
12. Smys S, Wang H (2021) Security enhancement in smart vehicle using blockchain-based architectural framework. J Artif Intelli Capsule Netw 3:90–100. https://doi.org/10.36548/jaicn.2021.2.002
13. Suganthe RC, Shanthi N, Latha RS, Gowtham K, Deepakkumar S, Elango R (2021) Blockchain enabled digitization of land registration. In: 2021 International Conference on Computer Communication and Informatics (ICCCI), pp 1–5
14. Swan M (2015) Blockchain: blueprint for a new economy. O'Reilly Media, Inc
15. Verheye B (2020) Land registration in the twenty-first century: blockchain land registers from a civil law perspective. Springer International Publishing, Cham, pp 123–135
16. Vikaspedia.in: Digital India land record modernisation programme. https://vikaspedia.in/social-welfare/rural-poverty-alleviation-1/schemes-department-of-land-resources/computerisation-of-land-records

Blocking Estimation Using Optimal Channel Reservation Policy in GSM 1800 System

Promod Kumar Sahu, Hemanta Kumar Pati, and Sateesh Kumar Pradhan

Abstract Voice is one among the important services provided by any mobile cellular network. The deployment of 5G that is the advanced mobile cellular communication technology will be more important than ever for the world, through various types of use cases. A 5G smartphone will not connect to a mobile network unless voice support is available, so enabling this is a must. Voice service requires end-to-end 5G network support to enable the high-quality voice service experience for mobile devices. So, we need to consider the whole network chain with IP Multimedia Subsystem (IMS), 5G core, and 5G radio access network (RAN)s. To provide seamless voice handover when users move between cells and access technologies, the network must also handle interworking toward 4G, 3G or 2G in case the users move out of 5G coverage. Evolved packet system (EPS) will be helpful for service providers when they are migrating from 4G to 5G. Circuit Switched FallBack (CSFB) will be helpful for service providers when they are migrating from 2G/3G to 4G. Once the call is set for voice communication if the mobile host (MH) moves from one cell in the wireless cellular network to another and finds no free channel in that cell, then the call will be dropped. To address this issue in this paper, we have presented an optimal channel reservation (OCR) policy. It reserves channels in a cell according to the target handoff call dropping probability (HCDP). So, it minimizes new call blocking probability (NCBP) and keeps HCDP below the target. We have applied this policy to Global System for Mobile Communication (GSM) 1800 system and observed that HCDP is below the target and NCBP is minimum.

Keywords New call blocking · Handoff call dropping · Optimal channel · GSM · Mobile cellular networks

P. K. Sahu (✉) · S. K. Pradhan
Department of CSA, Utkal University, Bhubaneswar, Odisha 751004, India
e-mail: promod_sahu@yahoo.com

S. K. Pradhan
e-mail: sateesh.cs@utkaluniversity.ac.in

H. K. Pati
Department of CSE, IIIT, Bhubaneswar, Odisha 751003, India
e-mail: hemanta@iiit-bh.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_45

1 Introduction

Wireless mobile cellular network with limited resource provides service to many users. In such networks, providing service to many users by considering desired quality of service (QoS) with limited capacity is a challenging issue. To this end, network resource allocation concept plays a key role. Again, for efficient utilization of network resources and to provide QoS guarantees to different applications resource allocation is important. However, this goal becomes more difficult due to mobility, limited bandwidth, and error-prone wireless channel in mobile wireless networks compared to wired networks. In the wireless cellular network, the service area is divided into cells. So that the same radio channels may be reused in other cells some distance away [1, 2]. A radio channel can either be a frequency band in FDMA or a code in CDMA in traditional 2G mobile cellular networks or Physical Resource Block (PRB)s in long-term evolution (LTE) like advanced mobile cellular networks. The number of mobile users supported in a cell is limited to the number of radio channels in FDMA or the number of codes in CDMA-based mobile cellular networks. In LTE, the number of users supported is limited to RBs available in the system. The cell/eNB capacity in terms of number of mobile users supported in LTE network providing voice over long-term evolution (VoLTE) service can be found from [3–5]. A base transceiver station (i.e., BTS in GSM like 2G networks or eNodeB/eNB in LTE like 4G networks) connects mobile users in the cell it covers. When mobile user wants to communicate with other users, it must connect to the BTS/eNB with a new connection request, where a set of desirable QoS requirements are implicitly or explicitly stated.

It observed that a huge investment has already gone in to build 4G networks and further to build 5G network a huge investment is required, and this makes logical sense for 5G and 4G networks to interwork with each other. 5G supports both non-stand-alone architecture (NSA) and stand-alone architecture (SA) implementations [6, 7]. 3GPP release 15 standardized the NSA mode of deployment. The 5G networks support existing 4G infrastructure in NSA mode. Based on the deployment of 5G network, a number of voice delivery options are available. For example, Fig. 1 shows Option 3 structure established on E-UTRA New Radio-Dual Connectivity (EN-DC) defined by the 5G NR standard. 4G LTE eNodeB (eNB) acts as the master node, and 5G en-gNB acts as the secondary node. Both of these RAN functions connect to the 4G Evolved Packet Core (EPC). Here based on only 5G, an end-to-end voice service is not available and the only voice options are CSFB and VoLTE.

Once the 5G core is deployed, the additional 5G deployment options are Option 2, Option 4, and Option 7. When the service provider deploys an IMS network to support the VoLTE service, then an end-to-end voice service by using Voice over New Radio (VoNR) is also an option. To meet the QoS flows required for any resultant voice packets associated with a call and for IMS signaling, the 5G end-to-end network will be used in this approach. In the 5G core, this voice option is lost because the CSFB capability is not there. This concept is outlined in Fig. 2. It also presents RAT FallBack (RATFB) and EPS FallBack (EPSFB) as additional voice options. Both the RATFB and EPSFB consider that the device has registered for IMS services and

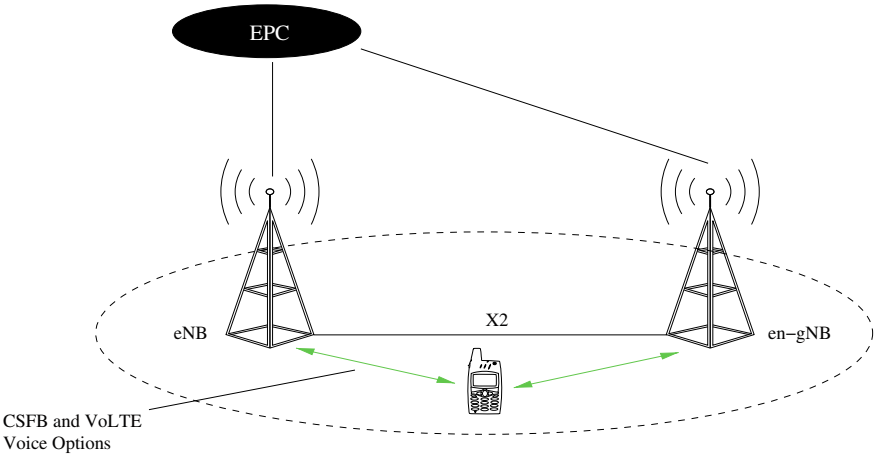


Fig. 1 5G deployment option 3

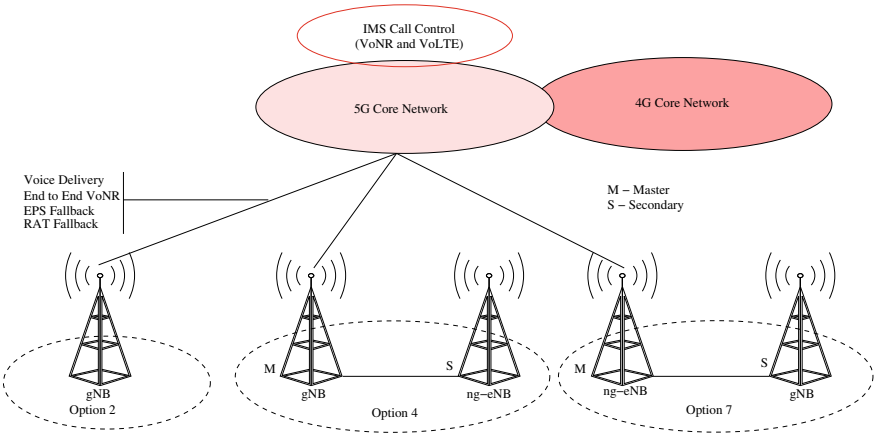


Fig. 2 5G deployment Option 2, Option 4, and Option 7

is on in 5G. The EPSFB or RATFB may be initiated in the current 5G cell when VoNR is not available in the cell. The EPSFB procedure takes place when the device signals to the IMS that it needs to make a call, then the device will be directed to move from 5G to LTE. At this point, the CSFB procedure or even the VoLTE call can be initiated.

When LTE is just used for data transfer, voice calls are handled through the legacy circuit switched mechanisms by falling back to 3G or 2G network as shown in Fig. 3. CSFB works only when the area covered by an LTE network is also covered by the 3G network. EPS will be helpful for service providers when they are migrating from 4G to 5G. CSFB will be helpful for service providers when they are migrating from 2G/3G to a 4G network. In CSFB, the 4G Mobility Management Entity (MME) talks

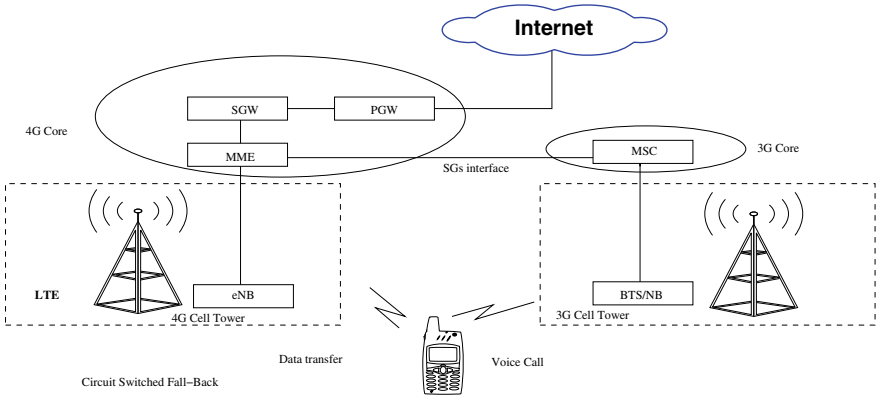


Fig. 3 Circuit switched FallBack

to the 3G mobile switching center (MSC) through the new SGs interface, to set up the voice call. Once the call setup is completed if the user moves from one cell to another and finds no sufficient channel or bandwidth, then the call will be dropped. To address this situation, some channels to be reserved in each cell for this type of handoff call (HC)s, and some schemes are discussed in [8–20]. If more channels are reserved for HC, the HCDP is decreased and the NCBP will be increased. To balance these two QoS parameters (i.e., NCBP and HCDP), we present an optimal channel reservation (OCR) policy. This scheme reserves optimal number of channels based on target HCDP.

The remaining part of this paper is organized as follows. In Sect. 2, we have described our optimal channel reservation policy. In Sect. 3, we have applied this optimal channel reservation policy to GSM 1800 system. Finally, in Sect. 4 we have concluded this paper.

2 The OCR Policy

In this section, the OCR policy is introduced. First, we have presented the expressions for NCBP and HCDP based on Markovian model, and subsequently, we have presented the OCR policy.

2.1 Estimation of NCBP and HCDP

The Markovian model is shown in Fig. 4. Here we considered that the total number of radio channels in a cell of a cellular system is C_t and C_T is the threshold. The system accepts both new call (NC) and HC when the number of used channels in a

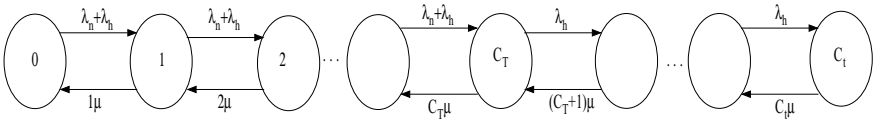


Fig. 4 Markovian model

cell is less than the threshold. The system accepts only HC when the number of used channels in the cell is equal or greater than the threshold and less than C_t. Here λ_n is the NC arrival rate, λ_h is the HC arrival rate, and μ is the service rate. By using these parameters, the NCBP denoted by P_{nb} and HCDP denoted by P_{hd} are given below, and the detail derivation for this can be found from [10].

$$P_{nb} = \left(\frac{1}{C_T!}\left(\frac{\lambda_n + \lambda_h}{\mu}\right)^{C_T} + \sum_{i=C_T+1}^{C_t} \frac{1}{i!}\left(\frac{\lambda_n + \lambda_h}{\mu}\right)^{C_T}\left(\frac{\lambda_h}{\mu}\right)^{i-C_T}\right)P_0 \tag{1}$$

$$P_{hd} = \frac{1}{C_t!}\left(\frac{\lambda_n + \lambda_h}{\mu}\right)^{C_T}\left(\frac{\lambda_h}{\mu}\right)^{C_t-C_T} P_0 \tag{2}$$

where

$$P_0 = \left[\sum_{i=0}^{C_T} \frac{1}{i!}\left(\frac{\lambda_n + \lambda_h}{\mu}\right)^{i} + \sum_{i=C_T+1}^{C_t} \frac{1}{i!}\left(\frac{\lambda_n + \lambda_h}{\mu}\right)^{C_T}\left(\frac{\lambda_h}{\mu}\right)^{i-C_T}\right]^{-1} \tag{3}$$

If P_{td} is the target HCDP, then to find C_T that satisfies the condition $P_{hd} \leq P_{td}$ given as follows:

$$P_{hd} = \frac{1}{C_t!}\left(\frac{\lambda_n + \lambda_h}{\mu}\right)^{C_T}\left(\frac{\lambda_h}{\mu}\right)^{C_t-C_T} P_0 \leq P_{td} \tag{4}$$

Algorithm 1 Optimal Reserved Channel Estimation.

```
1: procedure (λn, λh, μ, Ct, Ptd)
2:     CT = Ct; /* All channels are free channels. */
3:     Pnb(CT, Ct − CT)=NewcallBlock(λn, λh, μ, CT, Ct);
4:     Phd(CT, Ct − CT)=HandoffcallDrop(λn, λh, μ, CT, Ct);
5:     while (Phd(CT, Ct − CT) > Ptd) do
6:         CT = CT − 1;
7:         Pnb(CT, Ct − CT)=NewcallBlock(λn, λh, μ, CT, Ct);
8:         Phd(CT, Ct − CT)=HandoffcallDrop(λn, λh, μ, CT, Ct);
9:     end while
10:    y = Ct − CT;
11:    return(y);
```

2.2 Theoretical Estimation of Optimal Number of Reserved Channels

The Algorithm 1 is proposed to obtain the optimal number of reserved channels. This algorithm takes NC arrival rate (i.e., λ_n), HC arrival rate (i.e., λ_h), service rate (i.e., μ), the total number of channels in the cell (i.e., C_t), and the target HCDP (i.e., P_{td}) as its input parameters. This scheme returns optimal number of reserved channels (i.e., y) for which $P_{hd}(C_t - y, y) \leq P_{td}$ and $P_{nb}(C_t - y, y)$ is minimum. The process flow of this Algorithm 1 is shown in Fig. 5. In this algorithm, the function HandoffcallDrop($\lambda_n, \lambda_h, \mu, C_T, C_t$) estimates the HCDP using (2) and function NewcallBlock($\lambda_n, \lambda_h, \mu, C_T, C_t$) estimates the NCBP using (1). By using this algorithm, NCBP (i.e., $P_{nb}(C_t - y, y)$), HCDP (i.e., $P_{hd}(C_t - y, y)$), and optimal number of reserved channels (i.e., y) are calculated under different input traffic conditions.

3 Application of OCR Policy to GSM 1800 System

In this section first, we have discussed the GSM architecture and subsequently introduced the GSM 1800 system. Finally, we have applied our OCR scheme to the GSM 1800 system and discussed the results obtained.

3.1 GSM Architecture

The GSM architecture is shown in Fig. 6. The GSM network is having three subsystems, namely radio subsystem (RSS), network and switching subsystem (NSS), and operation support subsystem (OSS). The combination of mobile station (MS) and base station subsystem (BSS) is called RSS which comprises all radio-specific entities in the GSM network. The radio subsystem performs the basic functions of connecting the MS to the network. The radio subsystem consists of a number of MSs, base transceiver station (BTS), and base station controller (BSC). The interface between MS and BTS is U_m, interface between BTS and BSC is A_{bis}, and A is the interface between BSC and MSC. The NSS consists of a number of MSCs. Each MSC of the NSS connects to a number of BSCs. There are also home location register (HLR)s and visitor location register (VLR)s. NSS provides the desired interfacing between wireless and fixed networks. The OSS facilitates the operations of MSCs and also controls the operation and maintenance of the entire network [21].

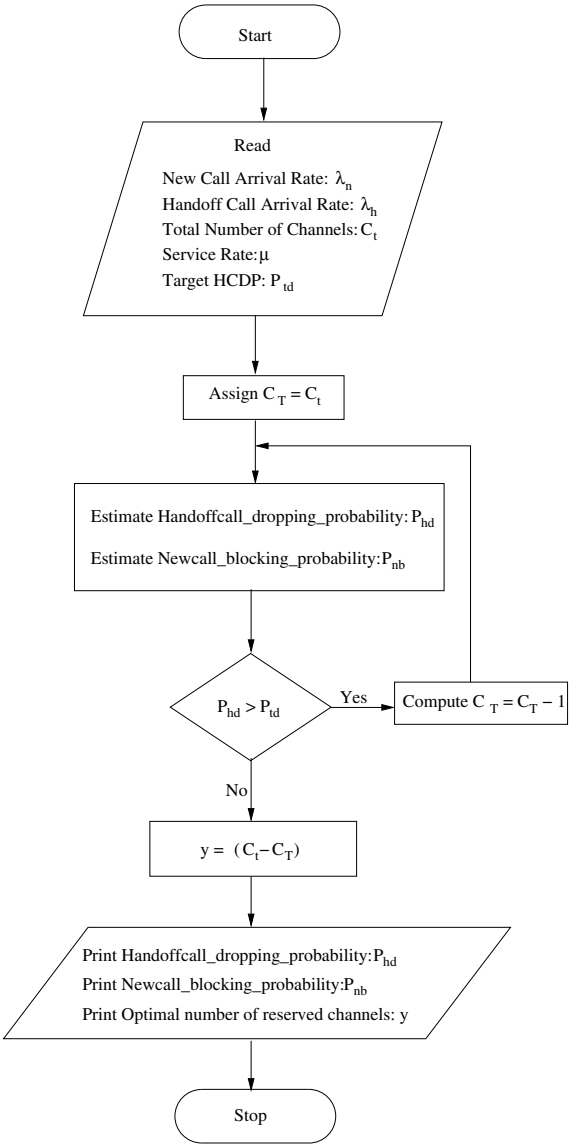


Fig. 5 Process flow of theoretical estimation of optimal reserved channels

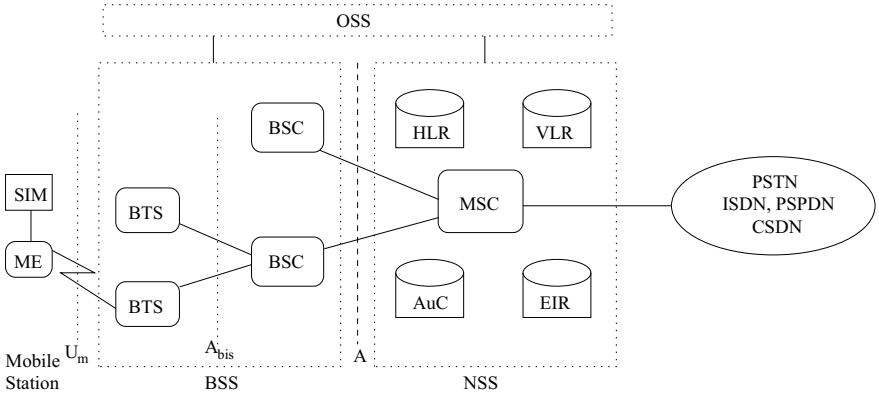


Fig. 6 Architecture of GSM system

3.2 GSM 1800 System

GSM was first introduced as a cellular system explicitly using the 900 MHz band, called the primary band. In GSM system to multiplex air medium among the users, a combination of FDMA and TDMA techniques is used. GSM uses a FDD technique for downlink and uplink separation. So, the downlink and uplink frequencies are different. GSM 900 system uses a frequency band from 935 to 960 MHz in the downlink (BTS transmit, MS receive) and another frequency band from 890 to 915 MHz in the uplink direction (MS transmit, BTS receive) each one 25 MHz. The band separation between downlink and uplink is 45 MHz. Again, the FDMA technique divides the frequency bandwidth of 25 MHz in each direction (i.e., both in downlink and uplink) into 125 carrier frequencies, where the bandwidth of each carrier frequency is of 200 kHz. However, the GSM 900 system uses the last carrier frequency channel as a guard band with other adjacent wireless systems. So, the total available channels are 124. For handling the increasing demands of many European countries, the number of available channels in GSM 900 standard was not sufficient. The controlling bodies assign additional frequency to GSM system. The new standard called GSM 1800 uses frequency band from 1805 to 1880 MHz for the downlink (BTS transmit, MS receive) and 1710 to 1785 MHz for the uplink (MS transmit, BTS receive). The GSM 1800 offers 75 MHz of bandwidth, so the number of channels or carrier frequencies is 374 and each carrier frequency has a bandwidth of 200 kHz [22–24]. Information related to GSM 1800 standard is given in Table 1.

In case of GSM 1800 standard, total channels available in the cell are 374. We can apply our optimal channel reservation policy to this standard. For this standard when a NC arrives if channel is available, then it is assigned to that NC. If no channel is available, then the NC will be blocked. When a NC is blocked, then the NCBP will increase. When a HC arrives if channel is available, then it is assigned to that HC. If no free channel is available in the cell, then the HC will be dropped and it will increase the HCDP. Generally, HC has higher priority than NC, so some channels will be reserved for HCs. If a more number of channels are reserved for HCs, then

Table 1 GSM 1800 system information

Features	GSM 1800
Spectrum alloted	150 MHz with FDD (i.e., 75 MHz in uplink and 75 MHz in downlink)
Uplink frequency band	1710–1785 MHz
Downlink frequency band	1805–1880 MHz
Access technology	FDMA/TDMA with SDMA
Number of channels	374
Channel bandwidth spacing	200 kHz
Modulation technique	GMSK (BT = 0.3)
Time slot per RF channel	8 slots
Time slot spacing for transmit–receive	3 slots
Transmit–receive frequency spacing or duplex spacing	95 MHz
Frame duration	4.615 ms
Compatibility	ISDN and PSPDN
Transmission power-handset	1 W
Encoding	RPE-LTP, EFR

the HCDP will decrease and it will increase the NCBP. So, based on the target HCDP if the channels are reserved for HCs, then the HCDP can be kept below the target and with minimum NCBP. Our OCR policy reserves some channels for HCs based on the target HCDP. If the target HCDP is changed, then the channels reserved for HCs will also be changed. This policy reserves the optimal number of channels for HCs based on the target HCDP, and it minimizes the NCBP. Our OCR policy reserves channels for HCs by setting the threshold C_T, where the number of reserved channels is $C_t - C_T$. Here, the NC will be admitted when used channels in the cell are less than C_T and the HC will be admitted when used channels in the cell are less than total channels available (i.e., C_t = 374 in case of GSM 1800 standard).

3.3 Results and Discussion

We have applied optimal channel reservation policy to GSM 1800 system considering the following parameters. In the first case, we have varied the NC arrival rate (i.e., λ_n) from 1 s^{-1} to 10 s^{-1} keeping HC arrival rate (i.e., λ_h) as 0.05 s^{-1} and service rate (i.e., μ) as 0.005 s^{-1}. In the second case, we have varied λ_h from 0.01 s^{-1} to 0.1 s^{-1} keeping λ_n as 5 s^{-1} and μ as 0.005 s^{-1}. For both of these cases, the target HCDP (i.e., P_{td}) is taken as 0.01. Based on the frequency band used in GSM 1800 system, we have plotted numerical results obtained.

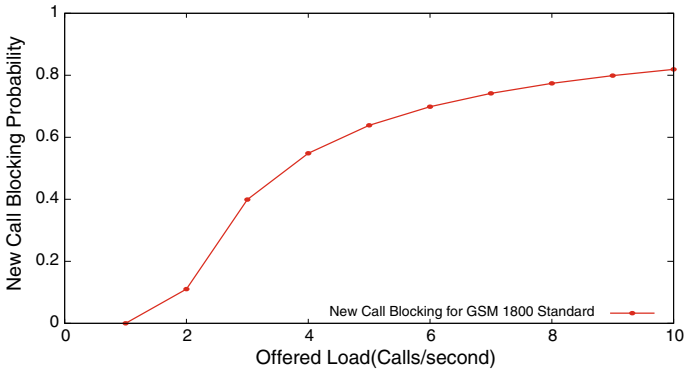


Fig. 7 NCBP versus offered load for GSM 1800 system

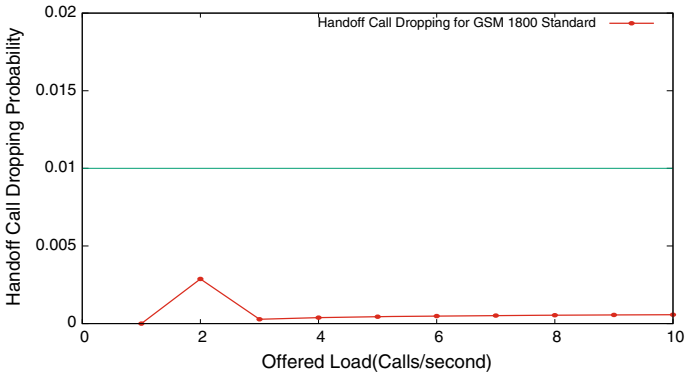


Fig. 8 HCDP versus offered load for GSM 1800 system

Figure 7 shows NCBP in different offered load conditions for the GSM 1800 system. From this Fig. 7, it is clear that NCBP increases with an increase in NC arrival rate for GSM 1800 system. Figure 8 shows the HCDP in different offered load conditions for GSM 1800 system. From this Fig. 8, it is observed that HCDP is remaining below the target for GSM 1800 system. Figure 9 shows optimal number of reserved channels with different offered load conditions for GSM 1800 system.

Figure 10 shows the NCBP in different HC arrival rates for GSM 1800 system. From this Fig. 10, it is observed that NCBP is slightly increasing with an increase in HC arrival rate in GSM 1800 system. Figure 11 shows the HCDP by varying HC arrival rate for GSM 1800 system. From this Fig. 11, it can be observed that HCDP is increasing with an increase in HC arrival rate. Further, it can be found that HCDP is remaining below the target by increasing reserved channels. Figure 12 shows the optimal number of reserved channels with different HC arrival rates for GSM 1800 system.

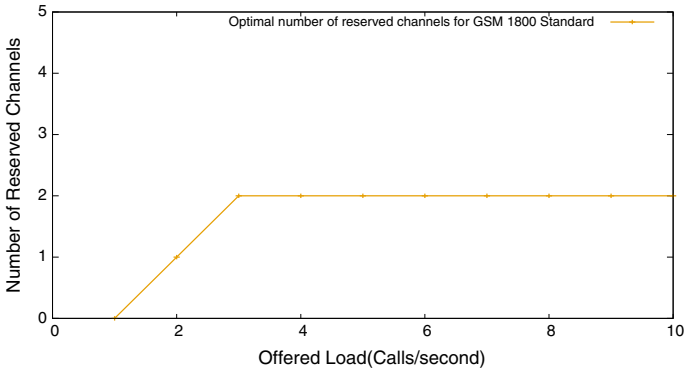


Fig. 9 Optimal number of reserved channels versus offered load for GSM 1800 system

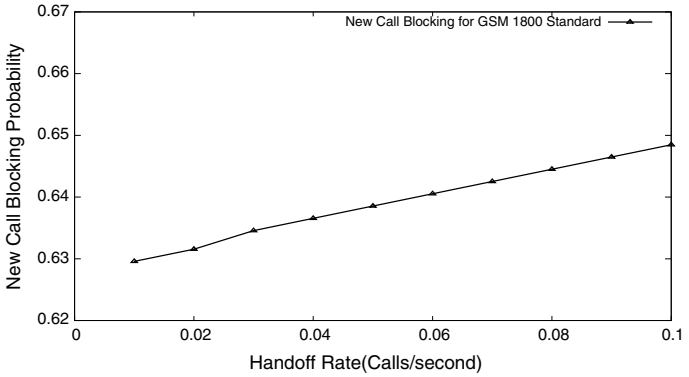


Fig. 10 NCBP versus handoff rate for GSM 1800 system

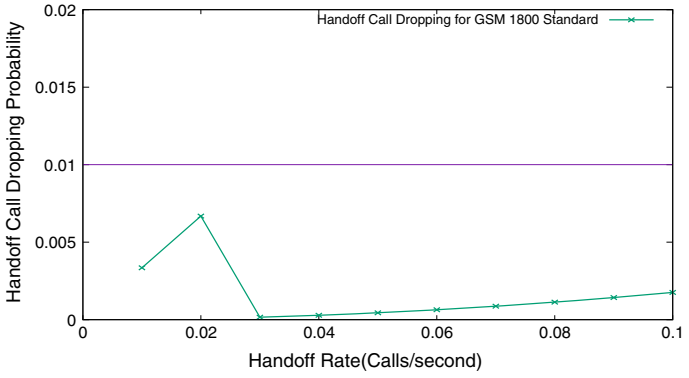


Fig. 11 HCDP versus handoff rate for GSM 1800 system

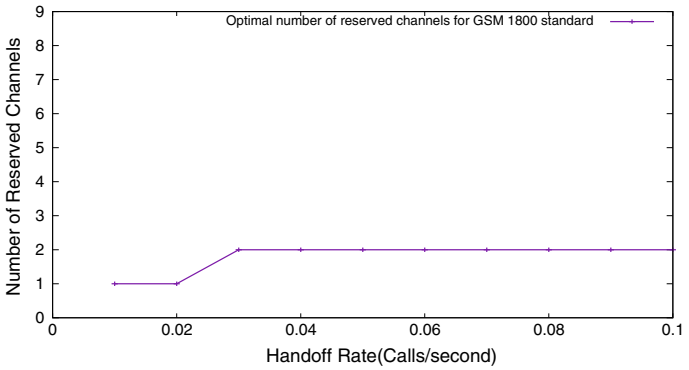


Fig. 12 Optimal number of reserved channels versus handoff rate for GSM 1800 system

4 Conclusion

In wireless cellular networks, two types of calls are there i.e., NC and HC. Generally, HCs have higher priority than NCs, so some channels are reserved for HCs. If reserved channels are more, then HCDP will be decreased, and at the same time, NCBP will be increased. If we reserve channels according to target HCDP, then HCDP can be kept below the target with minimum NCBP. So, in this paper we proposed an OCR policy for HCs. This policy reserved optimal number of channels for HCs based on the target HCDP. We applied this policy to GSM 1800 system having 374 channels with the target HCDP as 0.01. We observed that the HCDP could be maintained below the target with minimum NCBP.

References

1. Rappaport TS (2002) Wireless communications: principles and practice, 2nd edn. Prentice Hall, Upper Saddle River, NJ
2. Katzela I, Naghshineh M (1996) Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey. IEEE Personal Commun 3:10–31
3. Senapati R, Pati HK (2018) VoLTE cell capacity estimation using AMR-WB codec. In: IEEE International conference on advances in computing, communications and informatics, pp 1885–1889
4. Senapati R, Pati HK (2019) Modelling the region-based VoLTE cell capacity estimation using resource optimisation. IET Commun 13:1225–1235
5. Senapati R, Pati HK (2019) VoLTE cell capacity estimation using AMR-WB codecs and considering VAF, SID, Packet Bundling, and TTI Bundling. Wiley's Int J Commun Syst 32:1–16
6. Wang C-X, Haider F, Gao X, You X-H, Yang Y, Yuan D, Aggoune HM, Haas H, Fletcher S, Hepsaydir E (2014) Cellular architecture and key technologies for 5G wireless communication networks. IEEE Commun Magaz 52:122–130
7. 3GGP (2018) "System Architecture for the 5G System", Retrieved from https://www.3gpp.org/

8. Hong D, Rappaport SS (1986) Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures. IEEE Trans VT VT-35:77–92
9. Kulavaratharasah MD, Aghvami AH (1988) Teletraffic performance evaluation of microcell personal communication networks (PCN) with prioritized handoff procedures. IEEE Trans Commun 36:153–163
10. Oh S-H, Dong-wan T (1992) Proritized channel assignment in a cellular radio network. IEEE Trans Commun 40:1259–1269
11. Haring G, Marie R, Puigjaner R, Trivedi K (2001) Loss formulas and their application to optimization for cellular networks. IEEE Trans VT 50:664–673
12. Ni W, Li W, Alam M (2009) Determination of optimal call admission control policy in wireless networks. IEEE Trans Wireless Commun 8:1038–1044
13. Ortigoza-Guerrero L, Aghvami AH (2002) A prioritized handoff dynamic channel assignment strategy for PCS. IEEE Trans VT 48:1203–1215
14. Wang JL, Chiang SY (2004) Adaptive channel assignment scheme for wireless networks. Comput Electr Eng 30:417–426
15. Abdulova V, Aybay I (2014) Performance evaluation of non-prioritized and prioritized call admission control schemes in wireless cellular networks. Wireless Personal Commun 78:69–84
16. Chau TC, Wong KYM, Li B (2006) Optimal call admission control with QoS guarantee in a voice/data integrated cellular network. IEEE Trans Wireless Commun 5:1133–1141
17. Pati HK (2007) A distributed adaptive guard channel reservation scheme for cellular networks. Wiley's Int J Commun Syst 20:1037–1058
18. Pati HK (2013) A control-period-based distributed adaptive guard channel reservation scheme for cellular networks. Wireless Netw 19:1739–1753
19. Sahu PK, Pati HK, Pradhan SK (2021) A study on static call admission control policies for wireless mobile cellular networks. Springer Lecture Notes Electr Eng 702:229–245
20. Sahu PK, Pati HK, Pradhan SK (2021) A study on CAC schemes for homogeneous and heterogeneous wireless cellular networks. IEEE Int Conf Comput Commun Intell Syst 917–924
21. Rahnema M (1993) Overview of the GSM system and protocol architecture. IEEE Commun Magaz 31:92–100
22. Watson C (1993) Radio equipment for GSM. In: Balston DM, Macario RCV (eds) Cellular radio systems. Artech House, Boston
23. Winch RG (1993) Telecommunication transmission systems. McGraw-Hill, New York
24. Lee WCY (2006) Wireless and cellular telecommunications, 3rd edn. McGraw-Hill

The Effect of Big Data Decision-Making Analysis Used in Improving the Effectiveness of the Distance Learning Process at Private University in Jakarta

Febrianti, Hendra Wijaya Salim, Jonathan Buntoro, Oliviana, Ford Lumban Gaol, Tokuro Matsuo, and Fonny Hutagalung

Abstract In the current state of the pandemic, all sectors of human life must change because of the health protocols that must be implemented. All sectors ranging from the economic sector, law, and education experienced obstacles in its implementation. Of course, because this activity will be the spread of the COVID-19 outbreak cluster. In the field of education itself, in 2020 the government stipulates that the implementation of education in Indonesia is carried out using the distance learning model. However, these changes must of course be supported by technological advances in order to be achieved. Technologies such as zoom and other communication tools support this research. In addition, technology that is no less important is big data analysis in decision making in educational institutions, which is also needed to improve distance learning. But is distance learning effective today? And how does the use of big data analysis affect the effectiveness of the distance learning process? To prove it, we chose one of the universities that have used big data analysis in Private University in Jakarta as the sample to be analyzed using quantitative research methods with a

Febrianti (✉) · H. W. Salim · J. Buntoro · Oliviana
School of Information System, Bina Nusantara University, Alam Sutera, Indonesia
e-mail: febrianti@binus.ac.id

H. W. Salim
e-mail: hendra.salim001@binus.ac.id

J. Buntoro
e-mail: jonathan.buntoro@binus.ac.id

F. L. Gaol
Doctor of Computer Science, Binus Graduate Program, Kebon Jeruk, Indonesia
e-mail: fgaol@binus.edu

T. Matsuo
Advanced Institute of Industrial Technology, Shinagawa, Tokyo, Japan
e-mail: matsuo@aiit.ac.jp

F. Hutagalung
University of Malaya, Kuala Lumpur, Malaysia
e-mail: fonny@aiit.edu.my

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_46

Likert-shaped survey distribution. The results of this survey are that according to respondent's distance learning is currently effective, and indeed there is an effect of increasing the effectiveness of distance learning by using big data analysis to assist decision making.

Keywords Big data · Big data analytics · Decision making · Big data trends · Education · Distance learning

1 Introduction

In 2020, Indonesia was hit by a pandemic that is not yet known when it will end [1]. All sectors of life experience problems, one of which is the education sector. Currently, all learning is done online (remotely). Of course, this happened because the social distancing imposed by the government and schools attended by the nation's successors could become clusters of the spread of COVID-19. Therefore, since mid-July 2020, the government has decided to provide distance education [2]. Online learning systems are an opportunity for students who cannot leave their personal jobs but can still register for online courses they wish to develop, and students who live far from the educational institution they wish to attend [3]. But is distance learning effective? Of course, even though the pandemic occurs, the quality of Indonesian education must not decrease for the sake of the future of this nation's successors.

Of course, there are also questions like what to increase the effectiveness of online learning today? With the large amount of data needed in the distance learning process, the use of big data can be a solution [4]. We are currently in the era of big data which is a technology that can increase innovation, competition, and productivity [5]. Big data analysis is the process of collecting useful information by analyzing various types of large data sets which are then analyzed so that they can be used to find hidden patterns, market trends, and consumer preferences for the benefit of corporate decision making [6]. A combination of learning analysis techniques is quite different from new processing pattern systems, such as big data which allows the emergence of relevant information to the education department and teachers who have changed and maximized to use the current method [7]. So, it can be concluded that big data analysis is a large collection of data that is collected and analyzed into information that can predict the future so that it can help the decision-making process. The use of big data analysis in education will certainly help the decision-making process, especially during this uncertain pandemic [8]. Some universities today, such as Private University, use big data analysis in their learning process. But is distance learning currently effective? And is there any effect from using big data in distance learning?

With the questions above, we conducted a study to determine the effect of big data analysis used in big data decision-making analysis used in improving the effectiveness of the distance learning process. With this research, it is hoped that the potential of students can continue to be developed without any obstacles from the pandemic.

Of course, this is because the pandemic has not ended and has even increased recently, but as long as all learning such as courses, training, and others have also implemented strict health protocols, hopefully this pandemic will end soon.

Formulation of the problem.

1. Has the distance learning process been effective?
2. Is there any effect of increasing the effectiveness of distance learning?
3. Is big data important for decision making in universities?

1.1 Marketing Purposes

Definition of marketing according to Stanton (in Basu Swasta and Irawan 2003:5): "Marketing is an overall system of business activities aimed at planning, determining prices, promoting and distributing goods and services that satisfy needs both to existing buyers and buyers potential" [9].

Many parties wrongly set marketing targets during the distance learning process, but with big data analysis the system will analyze the track records of Private University students so that it can be known which areas and which school students are potential targets who can introduce Private University in Jakarta. This of course needs to be done so that marketing is carried out according to its place and does not waste funds. This can be analyzed from how many alumni of the school students in Private University in Jakarta because basically the mindset of students wants to follow in the footsteps of their alumni. After that, looking at the increasing years we will find out if more/fewer fans are entering the business. If, for example, there are more and more enthusiasts, they will continue to do marketing to the school, but if there are fewer enthusiasts, they will start not visiting the school. All of this depends on how we promote and how to get them to be interested in joining the business.

This depends on the conditions because it will be different if it happens like now during the pandemic, all marketing processes run online because they are not allowed to meet in person so that promotional activities are carried out through media Zoom, Google Meet, and others. Of course, of course, if it is promoted offline with online, of course the results obtained and the work system is different, it is possible that the schoolchildren are actually interested, but because online, the current condition is a pandemic which can result in unclear and effective explanations where they cannot interact directly such as asking questions that are not clear makes fans less interested so that the number of students in certain schools who study at private can reduce. In addition, it is possible that the current pandemic situation has reduced the number of enthusiasts visiting private due to unfavorable economic factors. The school is visited every year before the pandemic because of the increase and interested enthusiasts, but during the pandemic the potential targets are not suitable and possibly due to current conditions.

This aims to analyze the results of big data decisions on the marketing system that is implemented whether it continues to run in the school area because we want to

know how much the percentage increases every year whether the number of students who visit Binus increases if they are promoted directly or online.

1.2 Criminal Purposes

According to Nath (2007), in solving complex task cases, data mining is needed to detect criminal cases because the process is complex and requires human experience [10]. Big data analysis is also vulnerable to the possibility of data breaches by hackers so that student data and other important business data are spread or leaked to the public because someone is tracking them secretly. Of course, this usually happens on average due to human carelessness that is done intentionally and unintentionally. Everything goes back to what was planned from the beginning whether you really want to steal data and follow/plagiarize existing data, so you don't have to worry about the data because you already have it and you just need to create the database. After that, it is possible that the stolen data can be resold because during the pandemic, many people are doing things to get money in the dark [11]. This is one example that should be banned because it violates human rights regulations and is a crime. Misuse of personal data is an act that fulfills the elements of a criminal act such as elements of the crime of theft and other criminal acts, both in terms of objective elements and subjective elements. With the fulfillment of these elements, administrative sanctions, civil sanctions, and criminal sanctions are not sufficient to accommodate the criminal act of misuse of personal data which is in fact a perfect form of crime.

Of course, crime comes from a human act that violates the rules with what has been determined in the rule of law, strictly speaking the act that violates the rules set out in the rule of law and does not fulfill what has been stipulated in the rule of law that occurs in a society. We know that individual and university data is personal/private property that must be protected to make it more secure. Therefore, this needs to be watched out for because of the danger if there is duplication of the same data at one university with another, causing fraud and may require a computer-based application to track who stole data and to analyze big data in decision making [12]. An ideal criminal analysis tool should be able to identify criminal patterns quickly and in an efficient manner for criminal pattern detection and action in the future. However, in this scenario, there are major challenges as identified by Malathi and Baboo [13].

1.3 Decision-Making Purpose

Big data for the purpose of helping the decision-making process serves to provide information at the time of decision making. As we know today, this pandemic that sometimes goes up and sometimes goes down makes everything uncertain because big data is needed.

In this technological era, companies are faced with constant and rapid changes that require strategic actions and decisions that must also be fast in this digital era [14]. According to the Head of Data Analysis and Science, Prof. Maria Fasli, Ph.D., NTF, FHEA., Tuesday (7/8) at UGM from UNESCO "Big data provides knowledge as a basis for decision making and further actions" [15].

The use of big data assumes that human behavior has a universal level and can be predicted through trends. With limited information about a learner, we can predict a lot of information, including future behavior, with a limited amount of information about a learner. Decision making based on data from big data can improve business performance [16].

Big data enables instructional designers to identify content that is effective in guiding learning. They can then adjust the content to better fit, restructure, delete, or correct learning deficiencies using the teacher's knowledge. In addition, big data facilitates extensive learning testing and allows students to draw important conclusions about how they engage in learning activities [17]. However, the processing and storage of big data can still be developed by applying several existing technologies [18].

We present this research paper with the aim of determining whether distance learning is currently effective. Our aim in presenting a research paper entitled "The Effect of Big Data Decision-Making Analysis Used in Improving the Effectiveness of the Distance Learning Process at Binus University, Department of Information Systems" is basically to find out what causes the ineffectiveness and how to improve its effectiveness. Education system through big data analysis so that it can be developed.

Formulation of the problem.

1. Has the distance learning process been effective?
2. Is there any effect of increasing the effectiveness of distance learning?

1.4 Literature Review

In our literature review, we searched for 10 previous studies that were related to the current study. After doing a search, we found 10 literature reviews following our literature review in Table 1.

2 Research Technique

This research uses a technique of questionnaire survey using Google Form. The survey is in the form of a questionnaire that was created using Google Form in the form of a Likert with 5 scales, namely strongly agree, agree, neutral, disagree, and strongly disagree as shown in Table 2.

Table 1 Literature review

Conclusion research
Based on this research, the researcher concludes that decision making that is assisted by data from big data can increase the company's business performance exponentially. This researcher agrees that the use of big data in the company's decision making is very influential [16]
In this study, it is said that the distance learning process requires a lot of data, and it is very difficult to store this data. So, the use of big data can be a solution for storing data during the distance learning process [4]
According to this study, it is said that big data analysis is the process of processing data into useful information by analyzing large data sets so that it can be used to find hidden patterns, market trends, and consumer preferences for the benefit of corporate decision making [6]
According to research from America data-driven decision making has been used to help universities identify and evaluate strategies that can increase retention. As a decision based on data manufacturing enters the era of big data and learning analytics, this new approach, while not a silver bullet, may be part of the solution [19]
They can then adjust the content to better fit, restructure, delete, or correct learning deficiencies using the teacher's knowledge. In addition, big data facilitates extensive learning testing and allows students to draw important conclusions about how they engage in learning activities (Tulasi: 2013) [3]
In this research, big data and analytics have an important role in the future of higher education as it increases the use of analytics in the business and government sectors confirms the same [20]
This study agrees that big data can help decision making, but according to him big data can still be developed by applying several existing technologies [18]
A combination of learning analysis techniques that are quite different from new processing pattern systems, such as big data which allows the emergence of relevant information to the education department and teachers who have changed and maximized to use the current method [7]
The use of big data analysis in education will certainly help the decision-making process, especially during this uncertain pandemic [8]
According to this study, big data is a technology that can increase innovation, competition, and company productivity in the future [5]

Table 2 Scales Likert

No	Category
1	Strongly agree
2	Agree
3	Neutral
4	Disagree

This method can be easily accessed by students and other research subjects. This survey was distributed randomly to third semester students of the Information Systems Department at Private University in Jakarta. Because they are the ones who feel the learning process. After the survey, we targeted 20 Binus students to fill out our survey link. And this target was achieved, even as many as 50 students

were willing to fill out our survey, more than 30 students from the target. This data collection process can be seen in Fig. 1.

This survey aims to reveal in detail the importance and influence of big data decision-making analysis used in increasing the effectiveness of the distance learning process. In this survey, it is also known that all respondents have carried out the distance learning process and be aware of the research's goal so that the resulting data is quite valid. Therefore, we summarize the questions as follows:

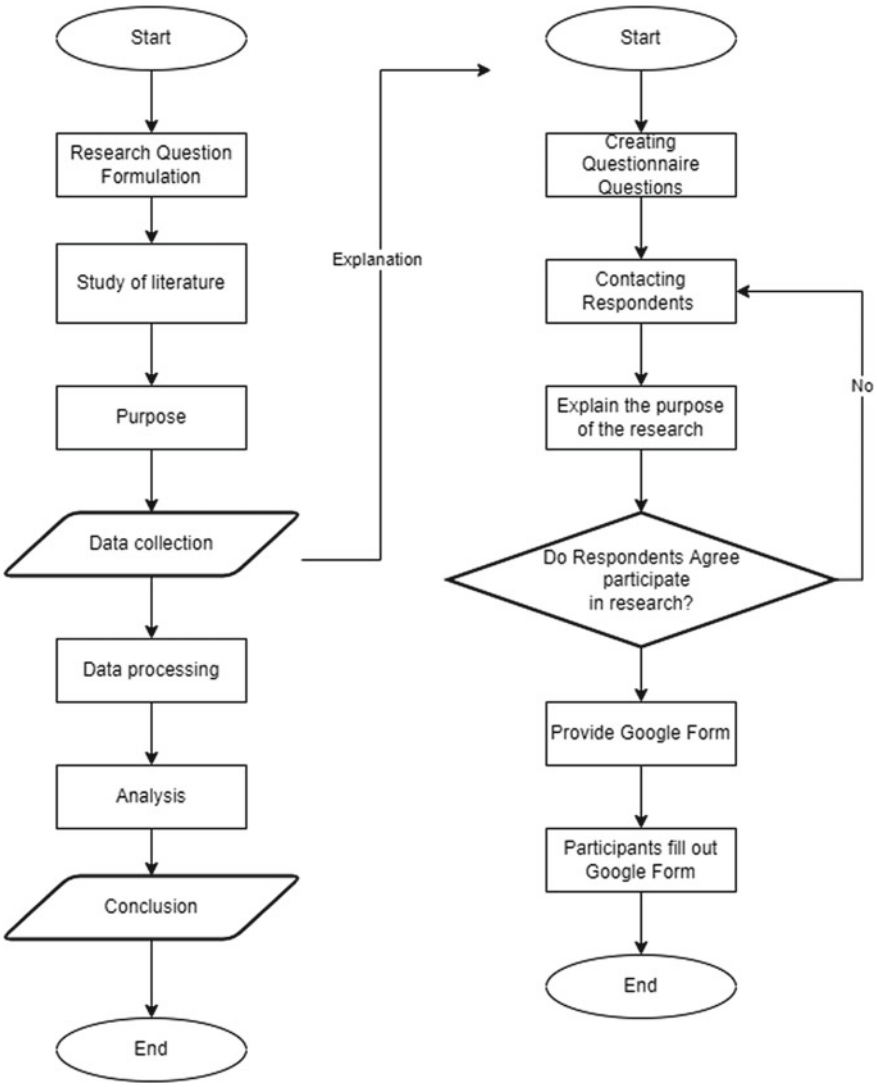


Fig. 1 Flow diagram of questionnaire survey using Google Form

- Is distance learning currently effective?
- Does big data analytics influence increase the effectiveness of distance learning?
- Is big data analysis necessary in decision making in the university?

3 Research Method

This research uses quantitative research methods. According to Nenty (2009), "quantitative research seeks to reveal research results in the form of relationships that occur between variables or phenomena that are currently happening" [10]. Meanwhile, according to Sugiyono (2015:14), "quantitative research methods can be interpreted as research methods based on a particular population or sample; sampling technique is generally done randomly" [21]. "The characteristics of quantitative research are data analysis techniques using quantitative (statistical) techniques objectively" [22] while this type of research uses a survey method. According to Pinsonneault and Kraemer (1993) "that survey research is seen as a method for quantitatively describing certain aspects of a particular population so that data collection is carried out to a group of people whose results can be generalized back to a particular population" [23].

According to Sugiono (2009: 143), "the types of questions in the survey are divided into closed and open" [24]. Open questions are those that ask for the respondent's viewpoint in the form of a description, and hence, the questions are written like extended essays. With this form of questioning, it is hoped that respondents can be open, and researchers can get a lot of data from respondents. Closed questions will help respondents to answer quickly and precisely because respondents only need to be choosing the answers that the researcher has been presented with, so this strategy makes it easier for researchers in conducting data analysis on all questionnaires that have been collected. So, in this study, we will use a quantitative method with a closed type of survey in the form of Likert, because this method is suitable for knowing whether there is an effect on increasing the effectiveness of distance learning by using big data analysis.

This Likert survey form will contain a number of questions about the object to be revealed and respondents can choose the Likert scale provided. The attitudes, views, and perceptions of the respondents in this study were measured using this Likert scale. Of course, this is about whether there is an effect on increasing the effectiveness of big data in the distance learning process. In general, the alternative answers to the Likert scale model questionnaire consist of 5 alternative answers.

Analysis of the data in this study is using statistical techniques percentage (Anas Sudijono 2009: 43) as the calculation formula [25] as follows:

$$P = \frac{F}{N} x 100\% \tag{1}$$

Description:

Table 3 Score from the answers statement

No	Answer	Statement	
		Positive	Negative
1	Strongly disagree	1	5
2	Do not agree	2	4
3	Neutral	3	3
4	Agree	4	2
5	Strongly agree	5	1

P = Percentage number.

F = Number of respondents who answered the answers in Table 3.

N = Number of samples.

Where in this formula we get the percentage amount by dividing the number of respondents who answered the question in accordance with Table 3 by the number of samples or the number of respondents who participated in this study multiplied by 100% so that the results of this data will be in the form of a presentation. For the interpretation of research results from this presentation, we apply the mean formula proposed by Sudjana (2008: 38) [21].

$$M = \frac{\Sigma(fi \times xi)}{fi} \tag{2}$$

Description:

M = Mean (average value).

= Express the amount.

fi = answer frequency x.

xi = score of answer choices.

In this formula, we find the average value by adding up the frequency of answers x times the score of the answer choices divided by the frequency of answers. Next, we can calculate the percentage of the mean or the percentage of answer choices for each statement with the formula:

$$Px = \frac{M}{N} \times 100\% \tag{3}$$

Description:

Px = Percentage of respondent selection.

M = Average percentage for each statement.

N = Total number of mean in one sub-indicator.

Table 4 Category percentage value

No	Percentage limit interval (%)	Rating category
1	0–20	Very low
2	21–40	Low
3	41–60	Currently
4	61–80	Tall
5	81–100	Very high

This formula is done by dividing the average percentage for each question divided by the total number of means in one sub-indicator. Then to represent these results, use the opinion of Riduwan (2006: 41) [26] to interpret the questionnaire data as follows:

The percentage calculation algorithm is as follows:

1. Calculating frequency (f)
2. Calculate the percentage (P) of answers using formula 1 above.
3. Calculating the average (M) answer according to formula 2 above.
4. Calculating the mean percentage (Px) using 3 formula
5. Interpretation of research results using Table 4.

This algorithm can be seen from Fig. 2.

4 Research Result

From the survey results, we calculated the percentage of answers on a Likert scale that we distributed to 50 respondents. In the first question, we know that online learning today is quite effective. The results of this data are shown in the form of a bar chart. This data is the result of a percentage calculation. This can be seen in Fig. 3.

Of the 50 respondents, 18 respondents or about 36% of all respondents agreed that online learning is currently effective. 12 respondents or about 24% of respondents strongly agree that online learning is currently effective, about 28% or 14 neutral respondents neither agree nor disagree but 10% or 5 respondents disagree with this, and 2% or one respondent strongly disagrees with online learning already effective.

Next are the survey results from the second question. The results of this survey will also be presented in the form of a bar chart. As can be seen from the table, the question asked this time is whether there is an effect on increasing the effectiveness of the distance learning process. This data is the result of a percentage calculation. This data can be seen in Fig. 4.

As can be seen from Table 4, as many as 42% or about 21 people agree that the use of big data analysis can improve the distance learning effectiveness. As much as 40% or 20 respondents agree with this opinion and 16% are neutral or 8 respondents neither agree nor agree with this opinion and 2% strongly disagree with this.

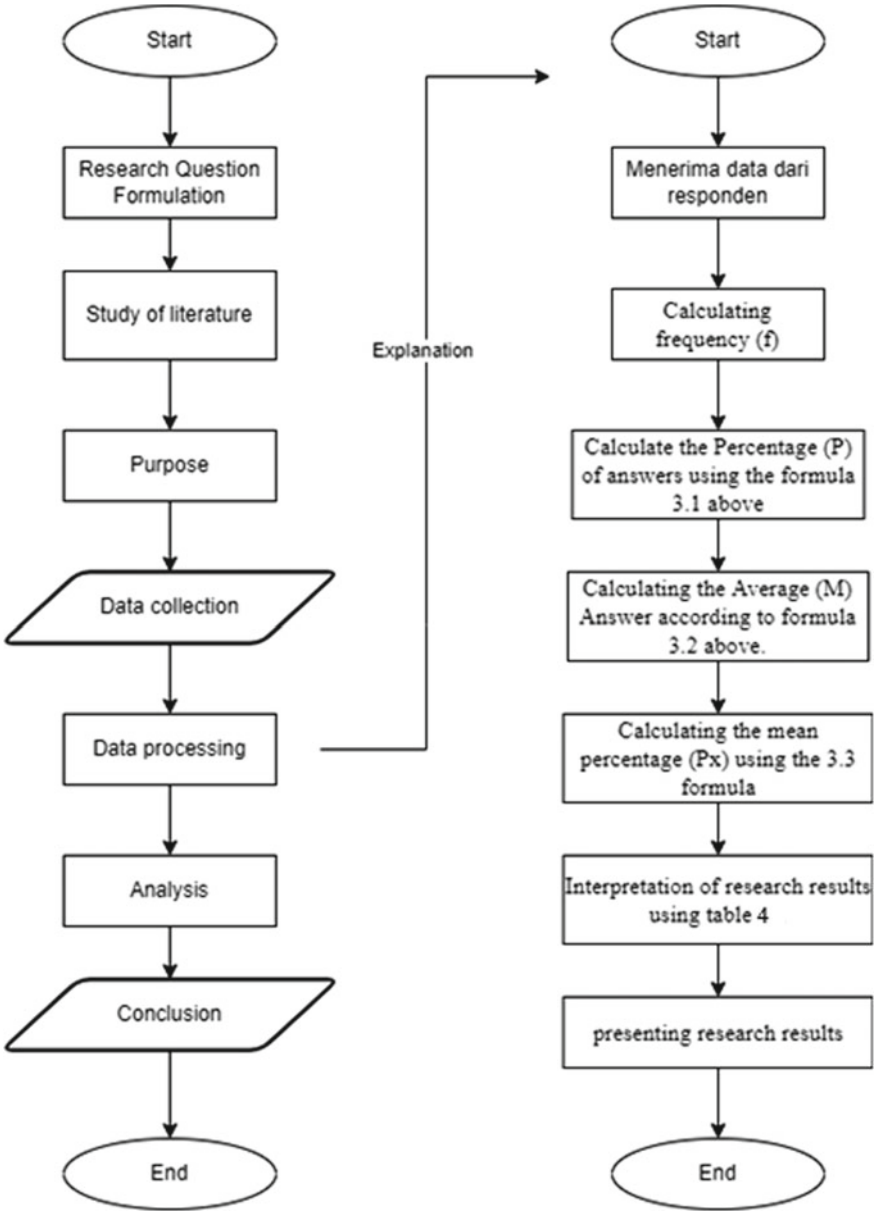


Fig. 2 Flowchart based on percentage calculation

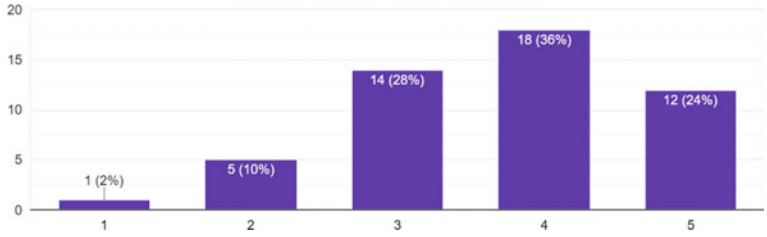


Fig. 3 Survey based on online learning method. *Note x* Scale Likert, *y* Respondent

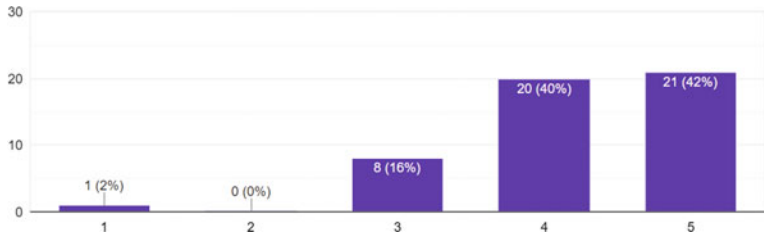


Fig. 4 Survey based on distance learning process. *Note x* Scale Likert, *y* Respondent

For the third question, namely whether big data needs to be used in decision making in universities? We can see the survey results from the bar chart below. This data is the result of a percentage calculation. This data can be seen in Fig. 5.

From the data above, 50% of respondents or 25 respondents strongly agree that big data analysis is very helpful in the decision-making process, while 44% 22 respondents agree with this opinion while 3 respondents or 6% are neutral. And there are no respondents who disagree or strongly disagree with this opinion.

After that, we also analyzed the data by calculating the mean and then the percentage of this mean and analyzed it with Table 4, so the data was obtained as in Table 5.

From this table, we can see that the first question got a score of 86% so that based on Table 5 it meets the high category. The second question got a score of 97% so that based on Table 5 it meets the high category. The third and last question got a

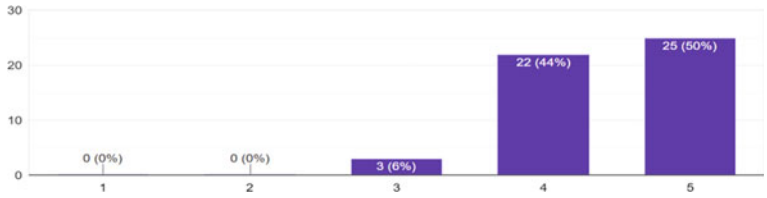


Fig. 5 Survey based on big data analytics. *Note x* Scale Likert, *y* Respondent

Table 5 Data obtained for different questions

No	Question	Percent (%)	Category
1	Is online learning effective today?	86	High
2	Is big data analysis important for decision making to improve learning effectiveness	97	High
3	Is big data analysis important for decision making in universities?	98	High

score of 98% so that based on Table 5 it meets the high category. The following are the results of the study after analysis, and calculations were carried out.

5 Discussion

From the results of the survey and research, we discuss whether the results obtained are in accordance with what is the formulation of the problem. The result of our discussion based on the survey above is that online learning is currently quite effective. We also conclude that online learning is very effective because of big data. As we know, effective online learning cannot be separated from the use of big data. In addition, according to the research of Provost Foster and Fawcett Tom, decision making based on data from big data can improve business performance because a business requires decision making in order to advance its business, with various data in big data itself, of course it will make it easier for decision making [23]. Because of the various existing data, the analysis will be easy to do which can later be used for decision making. According to research by Galuh Nurvind Kurniawati who said that big data analysis can help decision making because it can analyze various types of large data sets. Later the data can be used to find hidden patterns, market trends, and consumer preferences so that the data will later be useful for the benefit of corporate decision making [6]. This is in accordance with the results of our discussion from the survey results above, namely that each result obtained is in accordance with our hypothesis that the use of big data greatly influences the increase in the effectiveness of the distance learning process. The results of the study also showed satisfactory results, namely 50% with 25 answers. With these results, it can be said that the use of big data greatly affects the effectiveness of online learning during this pandemic in terms of decision making.

The results of our discussion also found that currently, the use of big data during online learning is very influential in increasing effectiveness. Because as we know, in this online era, students certainly rely heavily on various kinds of big data such as Google Chrome and many more to increase knowledge and do assignments and examinations. According to Olga Arranz García and Vidal Alonso Secades, with the large amount of data needed in the distance learning process, the use of big data can be a solution [4]. Based on his research, the data needed in distance learning is very

much. Therefore, of course, the use of big data can be a solution in this distance learning. So of course, big data is very much needed in online learning during this pandemic. Without big data, online learning is impossible to run effectively.

6 Conclusion

In conclusion, based on the results of the survey that has been carried out, it can be concluded that online learning during the current pandemic is quite effective. In addition, based on the survey that has been conducted, it is concluded that the use of big data greatly influences the increase in the effectiveness of online learning during this pandemic, especially in big data analysis of decision making.

This can be seen from the survey data that has been shown in the research results chapter. Apart from the results of this survey, it can be seen from the university's policy that recently chose to allow limited onsite learning. Of course, this is because of the pandemic which causes changes in all sectors including education in its implementation. As was the case last semester, where all activities were carried out online following the high level of COVID-19, this semester, limited onsite activities were able to run, even though there were few enthusiasts. However, at the beginning of 2022, the COVID situation escalated again. With this changing situation, Private University can make the right decisions thanks to its big data analysis. Therefore, big data is very important and plays a role in increasing the effectiveness of online learning.

In addition, big data as storage, for example, is Google Drive, which also helps online learning today. In this online learning, Google Drive also plays an important role, namely to store various important files, then send the answer file to the lecturer or teacher via Google Drive. With Google Drive, the important files that we have will be neatly organized in Google Drive itself and Google Drive is very important because sometimes it is used to send files to teachers and lecturers so of course big data like Google Drive will make it easier for us and also increase the effectiveness in the online learning process.

In addition, in this online learning, big data also plays an important role in decision making. Because there is currently a pandemic, big data is needed for decision making. The purpose of big data is used for decision making, for example, various campuses such as Private University, and these campuses need data regarding when to be onsite, whether this year should be onsite or not, of course with big data we will be able to predict this year can be onsite or not and so on. From the large amount of data in big data, of course, we can use the data to be used in decision making that can increase the effectiveness of learning.

Acknowledgements The research is made for the research project at Bina Nusantara University on the subject of Research Method in Information Systems. This research is helped by Mr. Dr. Ir. Ford Lumban, S.Si., M.Kom.

References

1. Dzulfaroh AN (2021) Setahun sejak Kasus Corona Pertama ini Kondisi di Indonesia. https://www.kompas.com/tren/read/2021/03/02/062500165/setahun-sejak-kasus-corona-pertama-ini-kondisi-pandemi-di-indonesia?page=all. Accessed 27 Jan 2022
2. Irfan K (2020) Kilas Balik Pembelajaran Jarak Jauh akibat Pandemi Covid-19. https://nasional.kompas.com/read/2020/09/03/10063201/kilas-balik-pembelajaran-jarak-jauh-akibat-pandemi-covid-19?page=all. Accessed 27 Jan 2022
3. Unsplash (2020) Bagaimana Data Dapat digunakan dalam E-Learning. https://www.brightspaceindonesia.com/blog/bagaimana-big-data-dapat-digunakan-dalam-e-learning. Accessed 28 Jan 2022
4. García OA, Secades VA (2013) Big data & learning analytics: a potential way to optimize elearning technological tools. In: IADIS International conference e-learning. https://files.eric.ed.gov/fulltext/ED562284.pdf
5. Philip Chen CL, Zhang C-Y (2014) Data-intensive applications, challenges, techniques and technologies: a survey on big data. http://www.deeplearningitalia.com/wp-content/uploads/2018/03/Data-intensive-applications-challenges-techniques-and-technologies-A-survey-on-Big-Data.pdf
6. Kurniawati GN (2022) Big data analytics dan Kegunaanya untuk Perkembangan Bisnismu. https://www.dqlab.id/big-data-analytics-dan-kegunaannya-untuk-perkembangan-bisnis. Accessed 28 Jan 2022
7. Muhammad A, Alas Y, Yunusm NHM, Sabtu NPH, Abdul Hamid MHS Online learning: trends, issues, and challenges in the big data era. J e-Learn Knowl Soc 12(1):121–134. https://www.learntechlib.org/p/171433/article_171433.pdf
8. Aisah N, Pratiwi WD (2021) Efektifitas Pembelajaran Jarak Jauh Terhadap Mahasiswa UNSIKA pada saat Pandemi COVID-19. Jurnal Pendidikan Bahasa dan Sastra Indonesia. https://ejournal.undiksha.ac.id/index.php/JJPBS/article/download/31464/pdf. Accessed 27 Jan 2022
9. Basu S, dan Irawan (2003) Manajemen Pemasaran Modern, vol 2(11). Liberty Offset, Yogyakarta
10. Mukherjee A, Nath P (2007) Role of electronic trust in online retailing: a re-examination of the commitment-trust theory. Eur J Market 41(9/10):1173–1202
11. Trisnawarman D (2015) Model Data Pengambilan Keputusan Untuk Analisis Data Tindak Kriminal. Seminar Nasional Informatika 2015. 174665-ID-model-data-pengambilan-keputusan-untuk-a.pdf (neliti.com)
12. Situmeang SMT (2021) Penyalahgunaan Data Pribadi Sebagai Bentuk Kejahatan Sempurna Dalam Perspektif Hukum Siber. 27(1). Penyalahgunaan Data Pribadi Sebagai Bentuk Kejahatan Sempurna Dalam Perspektif Hukum Siber | Situmeang | SASI (unpatti.ac.id)
13. Malathi, Baboo S (2011) Enhanced algorithms to identify change in crime patterns. IJCOPI 2(3):32–38
14. Gunawan M (2020) Pentingnya Big Data dalam Mengambil Keputusan di Perusahaan. https://www.sonora.id/read/422345898/pentingnya-big-data-dalam-mengambil-keputusan-di-perusahaan. Accessed 29 Jan 2022
15. Ika (2018) Big data Bantu Pengambilan Keputusan. https://www.ugm.ac.id/id/berita/16791-big-data-bantu-pengambilan-keputusan. Accessed 29 Jan 2022
16. Provost F, Fawcett T (2013) Data science and relationship to big data and data-driven decision making. Mary Ann Liebert, Inc., vol 1(1). https://doi.org/10.1089/big.2013.1508
17. Vidal A, Arranz O (2015) Big data & elearning: a binomial to the future of the knowledge society. Special Issue Big Data AI Int J Interactive Multimedia Artif Intell 3(6):23–33. https://dialnet.unirioja.es/descarga/articulo/5574005.pdf
18. Popchev1 IP, Orozova DA (2019) Towards big data analytics in the e-learning space. Cyber Inf Inf Technol 19(3). https://doi.org/10.2478/cait-2019-0023
19. Picciano AG (2012) The evolution of big data and learning analytics in American higher education. J Asynchron Learn Netw 16(3). https://files.eric.ed.gov/fulltext/EJ982669.pdf

20. Tulasi B (2013) The evolution of big data and learning analaytic in American higher education. Int J Comput Appl 68(14). https://www.researchgate.net/profile/Tulasi-Bomatpalli/publication/271070362_Significance_of_Big_Data_and_Analytics_in_Higher_Education/links/550fb9ec0cf2752610a0f82c/Significance-of-Big-Data-and-Analytics-in-Higher-Education.pdf
21. Sugiyono (2015) Metode Penelitian Kombinasi (Mix methods). Alfabeta, Bandung
22. Creswell JW (2010) Research design: pendekatan kualitatif, kuantitatif, dan mixed. PT Pustaka Pelajar, Yogyakarta
23. Pinsonneault A, Kraemer KL (1993) Survey research methodology in management information systems: an assessment. California Digital Library University of California. https://escholarship.org/content/qt6cs4s5f0/qt6cs4s5f0.pdf
24. Sugiyono (2009) Metode Penelitian Kuantitatif, Kualitatif dan R&D. Alfabeta, Bandung
25. Anas S (2009) Pengantar evaluasi pendidikan. Rajagrafindo, Jakarta
26. Adkon R (2006) Metode dan Teknik Menyusun Tesis. Alfabeta, Bandung

Coastline Change Detection Using *K*-means Clustering and Canny Edge Detector on Satellite Images

T. Sasank Dattu, D. Bhargav Reddy, M. Charan Teja, K. L. Sailaja, and P. Ramesh Kumar

Abstract Climate change and natural disasters are especially dangerous to coastal landscapes. Natural and man-made hazards like erosion sedimentation, rising sea levels, and tidal flooding are wreaking havoc on coastal areas. The mapping and detection of coastlines are critical for safe navigation, environmental conservation, and long-term coastal development. The model is composed of a method for extracting coastline from sentinel-2-l2a satellite images obtained from the sentinelhub website. The model uses the Gaussian blur module to reduce noise. Using *k*-means clustering, the image is then segmented as land and water. An edge detection algorithm identifies the boundary between land and water. One of many edge detection algorithms, Canny edge detector, applies a multi-level algorithm to detect edges in images. Percentage change is calculated by comparing window to window of edge detected images.

Keywords Canny edge detector · Coastal line extraction · Gaussian blur module · *K*-means clustering · Sentinel-2-L2a satellite images

T. Sasank Dattu (✉) · D. Bhargav Reddy · M. Charan Teja · K. L. Sailaja · P. Ramesh Kumar
Department of Computer Science and Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, India
e-mail: sasanktalluri@gmail.com

D. Bhargav Reddy
e-mail: duddugunta.bhargavreddy@gmail.com

M. Charan Teja
e-mail: charantejamudimadugula@gmail.com

K. L. Sailaja
e-mail: sailajak@vrsiddhartha.ac.in

P. Ramesh Kumar
e-mail: rameshkumar@vrsiddhartha.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_47

1 Introduction

Environmental changes have a threat to coastal zones, which are threatened by human activity, pollution, exotic species and typhoons. Rising sea levels has the ability to degrade and flood marine biodiversity while also eliminating wetlands. Warming and acidification of the oceans are likely to affect coastal and marine ecosystems. As a result, the coastal region must be carefully monitored. It is critical for a human authority to examine the changes in the coastline in the real world. As a result, the proposed system focuses on the change of the coast of Kakinada over time, which defines the many classes that comprise land water segmentation and edge recognition of satellite images, as well as computing the percentage change in the coastline. One of the most significant approaches in image processing is image segmentation. It is a phase in pre-processing, computer vision, and pattern recognition. It is a process of splitting a virtual image into numerous groups of pixels.

Edge detection is an image processing method that is used to discover spots in a virtual image where there are discontinuities or simply sudden changes in image brightness [1]. The image's edges are the spots where the brightness of the image fluctuates drastically. This technique makes users to validate picture characteristics for a substantial shift in greyscale image. It decreases the quantity of data in a picture while maintaining its key features. There are many algorithms which are used to detect the edges from the images. Among various edge detection algorithms Canny edge detector which detects edges at a wide scale are best suitable algorithm.

2 Literature Review

Detecting coastlines may be accomplished using a variety of ways. The coastline is estimated in the first stage by employing region segmentation techniques and a local thresholding approach [2]. The next stage improves on the results of previous stage by modelling shoreline with active contours, such as edge detection and noise reduction. In this method, the author used FCM [3]. Fuzzy *c*-means clustering (FCM) is an unsupervised method for categorizing a small set of N items into c fuzzy clusters. This methodology has effective noise removal. Sobel operator is also used to detect the edge as done in [4]. Satellite data produced two normalized difference indices (NDVI and NDWI) which are combined to give the highest level of precision and accuracy in distinguishing water and land. Following that, the Sobel operator is used to detect the coastal edge in this work. In work [5], the research region was extracted from satellite pictures using the SURF method. Two separate approaches were used to identify change: frame-to-frame comparison and coastal line change detection. By adopting a strategy that windowed the coastal zones, the execution time for this approach was lowered. Canny established three performance conditions for edge detection: the SNR criterion, the positioning accuracy and the single edge response. In paper [6], images of different bands are segmented to determine which

image has superior land water categorization. The coastline lines were found using mathematical morphology, a set theory-based analysis and processing tool. In work [7], atmospheric correction is done to remove the impact of atmospheric and lighting elements on ground object reflection. It is used to acquire true reflectivity data for ground objects. From the above study, segmentation must be done in order to get perfect coastline. Among all the segmentation techniques, *K*-means clustering is best for true colour and Normalized Difference Water Index (NDWI) satellite images show best result. Among all the edge detectors Canny edge detector which is multi-stage algorithm is best suitable since thresholding is done.

3 Methodology

To identify changes in the shoreline, a three-step process is used. The first stage was a pre-processing procedure that smoothens out sharp edges and segment land and water bodies. The edges are then identified using the Canny edge detector. The next phase is change detection, which is accomplished by window-to-window comparison (Fig. 1).

3.1 Data Set

The proposed methodology uses Sentinel satellite images acquired from [8]. It contains various types of images with different combination bands from years 2016 to 2022. The false colour infrared (IR) composite maps near-IR spectral band B8 with red and green bands, B4 and B3, to RGB components directly. Normalized Difference Water Index (NDWI) is used to observe changes in water bodies. Since water absorbs light strongly in the visible to infrared electromagnetic spectrum, NDWI

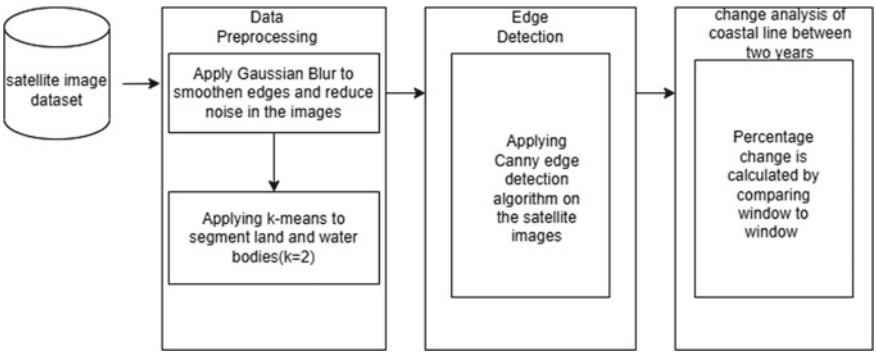


Fig. 1 Methodology

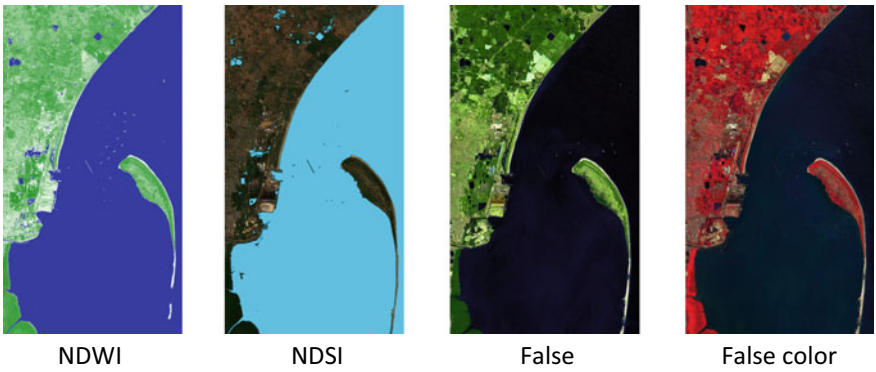


Fig. 2 Input images

highlights water bodies using green and near-infrared bands. False colour urban is used to visualize urbanized areas more clearly. Soil and sand are depicted in various hues. Snow and ice appear dark blue, whereas water appears black or blue. This composite is obtained by combination of 2.5 * $B12$, 2.5 * $B11$, 2.5 * $B04$ (Fig. 2).

$$\text{NDWI} = (B03 - B08)/(B03 + B08)$$

$$\text{NDSI} = (B03 - B11)/(B03 + B11)$$

3.2 Data Pre-processing

The Sentinel L2A satellite images are subjected pre-processing techniques. Gaussian (5×5) filter is applied to both images to reduce the noise a smoothen the sharp edges. Gaussian blur is a low-pass spatial filter that convolutes the input image with a kernel to produce a fuzzy image with sharper edges. It computes a weighted average of the current pixel's neighbours, with distant pixels receiving less weight than those in the centre. Then K-means clustering is applied to segment land and water bodies. When k-means clustering technique is applied to an image, it treats each pixel as a vector point and creates k-clusters of pixels. It is a method used for clustering vector points based on their closeness to K mean points. As a result, the vector points are partitioned into K-clusters. In this methodology, $k = 2$. Hence, the clustered image is obtained as shown in Fig. 3.

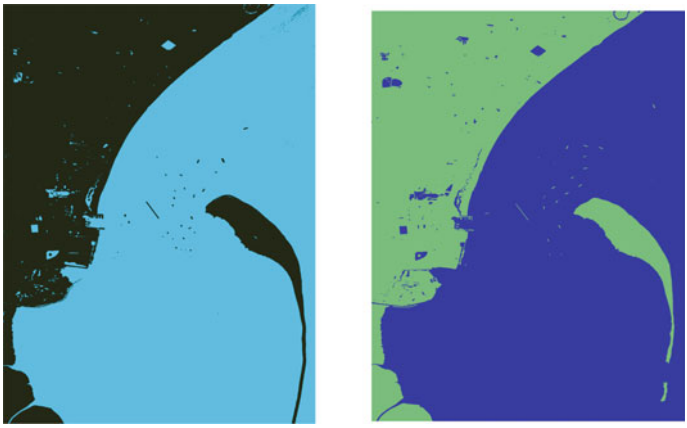


Fig. 3 Segmented images

3.3 Coastal Edge Detection

Segmented images obtained from the *K*-means clustering are subjected to Canny edge detector algorithm to detect the coastal edges from the images. Canny edge detector algorithm is a multi-stage approach. At first stage, the segmented image was filtered with Sobel kernel in both vertical and horizontal directions to calculate the first derivative in vertical (G_y) and horizontal directions (G_x). The direction of the gradient is perpendicular to the edges. It has four angles representing horizontal, vertical and both diagonal directions.

$$\text{Edge}_{\text{Gradient}}(G) = \sqrt{Gx^2 + Gy^2}$$

$$\text{Angle}(\theta) = \tan^{-1}\left(\frac{Gx}{Gy}\right)$$

The second stage is non-maximum suppression. After calculating magnitude and direction of the gradient, a thorough glance of image is performed to discard unnecessary pixels which is done by checking each pixel with local maximum in gradient's direction.

The last stage is hysteresis thresholding. Two threshold values, min$v$ and max$v$, are used as input. Edges with intensity gradient larger than max$v$ are certainly edges, whereas the pixels with intensity gradient less than min$v$ are non-edges and are avoided. Pixels that fall within threshold values are considered if they link to an edge; otherwise, they are ignored. Edges are detected as shown in Fig. 4.

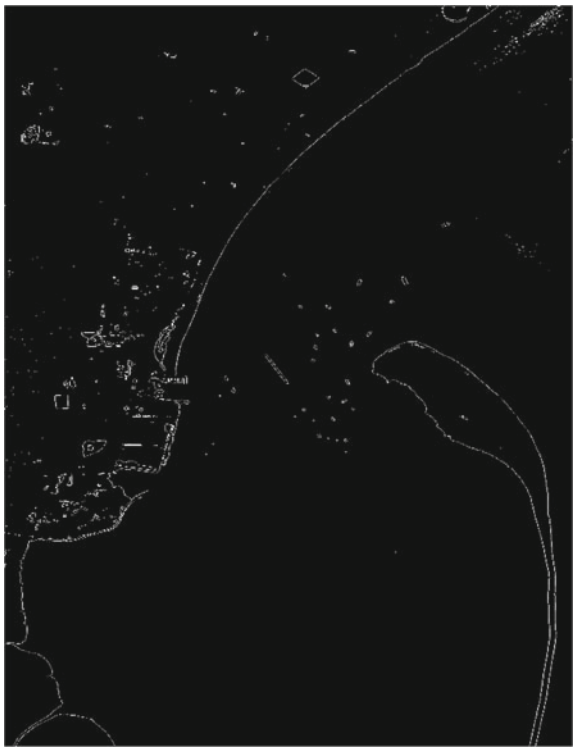


Fig. 4 Edge

3.4 Calculating the Change

Edge detected images are cropped for the studying area in exact same ratios. Now both the images are iterated pixel to pixel. If there is a change, then the absolute difference in pixels is calculated and is added to sum counter after the iteration. Then percentage is calculated from the sum of differences obtained (Fig. 5).

4 Result and Analysis

Kakinada images of years 2018 and 2021 are taken as input images which are Sentinel L2A images of bands 8,4,3. These images are pre-processed and segmented as shown in Fig. 6.

These segmented images are processed with Canny edge detector. The edge detected on one image is overlapped on another year edge image to observe the change in coastal region as shown in Fig. 7.

Bargraph shown in Fig. 8 is plotted using matplotlib library by taking years on *X*-axis and Percentage change on *Y*-axis.

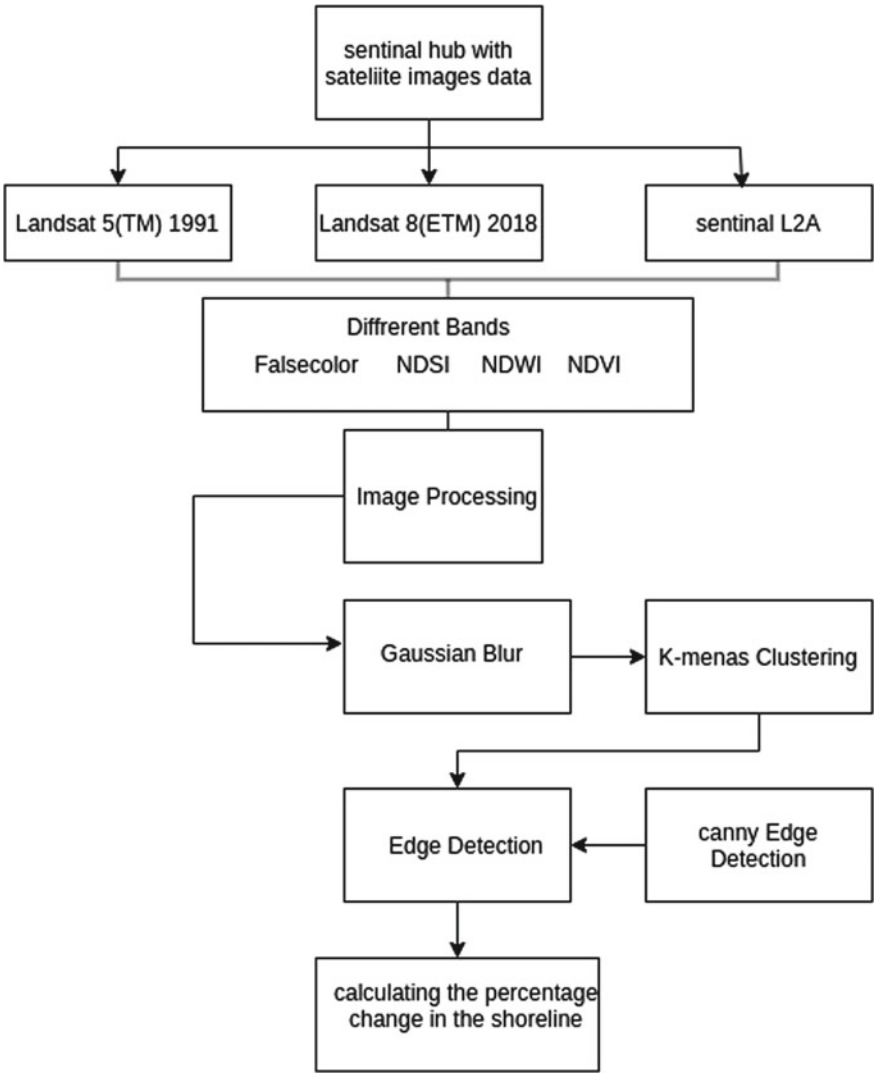


Fig. 5 Process flow

5 Conclusion

In this paper, the system estimates the percentage change in the coastline of Kakinada region. The objective is to enhance the robustness of the extracting coastline from Sentinel L2A satellite images. Instead of directly detecting edges from satellite images, these images are pre-processed using Gaussian blur and *k*-means clustering. Gaussian blur is applied to smoothen the edges, and the noisy pixels are removed.

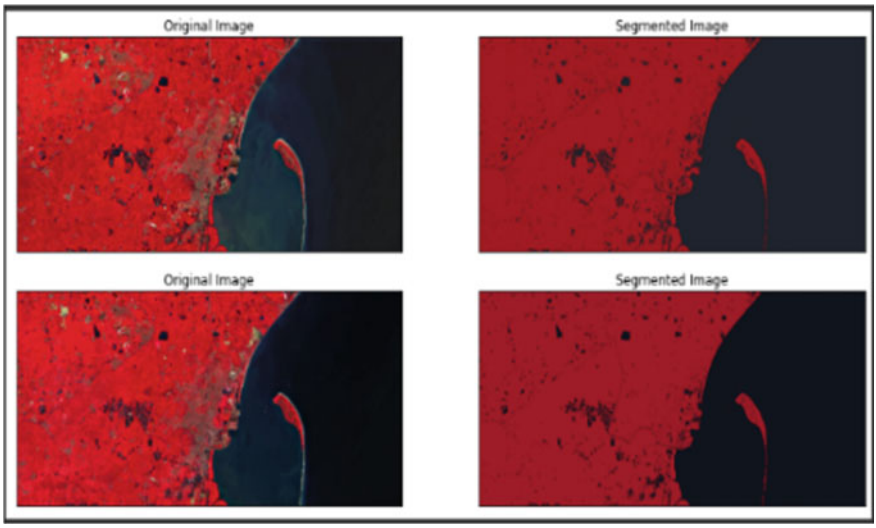


Fig. 6 Segmented images

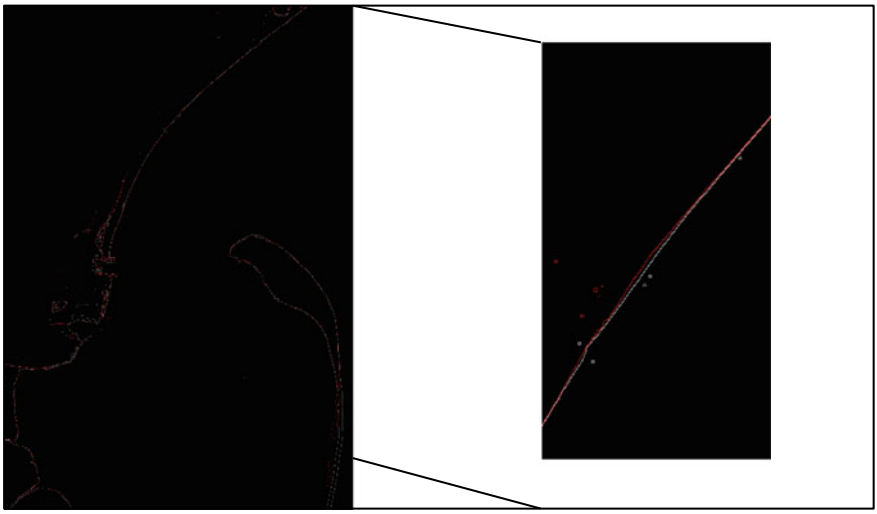


Fig. 7 Overlapped image

K-means is an unsupervised learning technique which segment the land and water bodies based on colours. Now, Canny edge detection algorithm which is a second gradient algorithm is applied to detect the edges accurately instead of any other operator. In next step, percentage change is calculated by iterating pixel to pixel and comparing the pixels of corresponding two images.

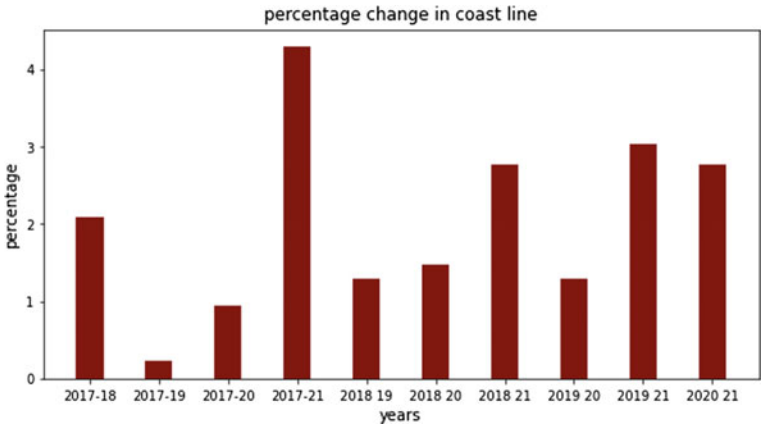


Fig. 8 Bargraph showing percentage change

Table 1 Observed percentage changes

Sl.no	Year 1	Year 2	Percentage change
1	2017	2018	2.0916
2	2017	2019	2.3418
3	2017	2020	2.4603
4	2017	2021	4.3031
5	2018	2019	1.2989
6	2018	2020	1.4695
7	2018	2021	2.7708
8	2019	2020	1.2979
9	2019	2021	3.0350
10	2020	2021	2.7751

The results indicate that the process detects coastline edge and estimate the percentage change.

References

1. Panwar P, Gopal G, Kumar R (2016) Image Segmentation using K-means clustering and thresholding. Image 3(05):1787–1793
2. Yasir M, Hui S, Binghu H, Rahman SU (2020) Coastline extraction and land use change analysis using remote sensing (RS) and geographic information system (GIS) technology—a review of the literature. Rev Environ Health 35(4):453–460
3. Modava M, Akbarizadeh G (2017) Coastline extraction from SAR images using spatial fuzzy clustering and the active contour method. Int J Remote Sens 38(2):355–370

4. Saeed AM, Fatima AM (2016) Coastline extraction using satellite imagery and image processing techniques. Red 600:720
5. Bouchahma M, Yan W, Ouessar M (2012) Island coastline change detection based on image processing and remote sensing. Comput Inf Sci 5(3):27
6. Zhang T, Yang X, Hu S, Su F (2003) Extraction of coastline in aquaculture coast from multispectral remote sensing images: object-based region growing integrating edge detection. Remote Sens 5(9):4470–4487
7. Paravolidakis V, Ragia L, Moirogiorgou K, Zervakis ME (2018) Automatic coastline extraction using edge detection and optimization procedures. Geosciences 8(11):407
8. https://apps.sentinel-hub.com/eo-browser/
9. Karantzalos KG, Argialas D, Georgopoulos A (2002) Towards automatic detection of coastlines from satellite imagery. In: 2002 Proceedings of 14th ınternational conference on digital signal processing, DSP 2002 (Cat. No. 02TH8628), vol 2. IEEE, pp 897–900
10. Dellepiane S, De Laurentiis R, Giordano F (2004) Coastline extraction from SAR images and a method for the evaluation of the coastline precision. Pattern Recogn Lett 25(13):1461–1470
11. Pham VH, Lee BR (2015) An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm. Viet J Comput Sci 21:25–33
12. Biswas R, Sil J (2012) An improved canny edge detection algorithm based on type-2 fuzzy sets. Proc Technol 4:820–824
13. Deng C-X, Wang G-B, Yang X-R (2013) Image edge detection algorithm based on improved canny operator. In: 2013 International conference on wavelet analysis and pattern recognition. IEEE, pp 168–172
14. Ebby DP, Jeena Jacob I (2021) Rain Streaks Removal in digital images by dictionary based sparsity process with MCA estimation. J Innov Image Process 3(3):174–189
15. Karuppusamy P (2021) Building detection using two-layered novel convolutional neural networks. J Soft Comput Paradigm (JSCP) 3(01):29–37
16. Shakya S (2021) Unmanned aerial vehicle with thermal imaging for automating water status in Vineyard. J Electr Eng Autom 3(2):79–91
17. James Deva Koresh H, Chacko S (2020) Hybrid speckle reduction filter for corneal OCT images. In: International conference on ımage processing and capsule networks. Springer, Cham, pp 87–99
18. Sathesh A, Eisa Babikir Adam E (2021) Hybrid parallel ımage processing algorithm for binary ımages with ımage thinning technique. J Artif Intell 3(03):243–258

An IoT-Based System to Measure Methane and Carbon Dioxide Emissions Along with Temperature and Humidity in Urban Areas

Umme Sanzida Afroz, Md. Rafidul Hasan Khan, Md. Sadekur Rahman, and Israt Jahan

Abstract The increment of the world population is the main reason behind the emission of greenhouse gasses like methane and carbon dioxide which affect temperature and humidity. To identify this problem, we implemented an IoT system using an ESP32 microcontroller and the necessary sensor to measure data from the outdoor environment and deploy our system to monitor the outdoor temperature and humidity quality due to increasing the level of methane and carbon dioxide. In the next step, the measured data was sent into the ThingSpeak cloud server and later was made available for visualization. This research work seeks to estimate the level of methane and carbon dioxide emissions which has a great impact on increasing temperature and humidity and also highlights some complications to integrating modern sensors based on the emergence of the Internet of things (IoT) concept. However, the data preserved in the cloud server can be used for further analysis to find other impacts of gasses in our environment.

Keywords IoT · ThingSpeak · Methane emission · Carbon dioxide emission · Temperature and humidity

1 Introduction

The entire world is confronting a critical air pollution problem because of the alarming growth of population, global industrialization, uncontrolled construction of infrastructure and housing, vehicle traffic, narrow and congested roads, using poorly maintained vehicles, poor management of waste landfills, and lack of efficiency in

U. S. Afroz (✉) · Md. R. H. Khan · Md. S. Rahman · I. Jahan
Daffodil International University, Dhaka, Bangladesh
e-mail: umme15-11505@diu.edu.bd

Md. R. H. Khan
e-mail: rafidul15-11542@diu.edu.bd

I. Jahan
e-mail: isratjahan.cse@diu.edu.bd

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_48

the programs of the pollution control system [1]. Air pollution is a term that has detrimental effects on climate change, the atmosphere, and ecosystems, which threaten the upcoming environment [2, 3]. Methane, ground-level ozone, black carbon, and sulfate aerosols are included in pollutants. One of our recent age's utmost obstacles is air contamination, which impacts global warming, global climate change, and also individual and public health. Air contains more than twenty-five thousand types of chemicals, contaminants, and toxins that are identified; large quantities of these are emitted from the combustion of fossil fuel, motor vehicle engines, and also garbage heaps [4].

The ozone layer is especially part of the stratosphere that is connected with climate change and air pollution. Increasing concentrations of the ozone layer, which are directly affecting climate change and ozone gas is a powerful component of the GHG and indirectly takes a part to influence the lifespan of methane gas which is another element of greenhouse gas [5]. On the contrary, change in climate influences the concentrations of ozone through the chemical and dynamic change in the atmosphere. For example, emitted black carbon, CO_2 from the combustion of fossil fuel, and emission of methane gas from waste landfills are directly relevant to the global warming effect of the earth's atmosphere [6].

Methane and carbon dioxide both are natural and anthropogenic because they are released into the atmosphere from a variety of biogenic sources like agriculture and waste disposal, rice paddies, burning of biomass, garbage heaps, and also fermentation is included. Besides, methane is emitted from the extraction of fossil fuels like coal, natural gas, and petroleum. Natural methane gas emissions occur from termites, wetlands, landfills, oceans, and other wild ruminants [7].

The increasing rate of methane emission is reacting in various hazardous ways like oxidization in the atmosphere, incrementing water vapor, and the presence of carbon dioxide (CO_2) in the air [8] where the increment of water vapor in high temperature is the cause of the real feeling of temperature.

The primary objective of this research work is to architect an IoT-based portable smart system that is used to estimate the change of temperature and humidity based on emissions of methane gas and carbon dioxide gas. This system measures carbon dioxide (CO_2), methane (CH_4), temperature, humidity and then transmits the real-time data wirelessly to the server over Wi-Fi [9].

2 Literature Review

This paper [10] works on an alarm system for sewage gas monitoring based on IoT that uses an MQ7 sensor for detecting carbon monoxide, an MQ4 sensor to measure methane gas with the help of ATmega328P which is an Arduino UNO microcontroller-based board. Microcontroller gets data from sensors and sends it to ThingSpeak cloud server through the GSM module where the data uploads to the server in real time. Since the sewer gasses are poisonous so when the sensors detect that the gas levels are higher than normal the system sends an SMS to the worker

or user via the GSM module to ensure safety. Mainly this approach automated an analog sewer gas analysis system where it was hard to measure the oscillation of gasses level in real time and also harmful for that person who enters to measure the gasses level manually.

The authors in this paper [11] introduced a device based on Arduino for monitoring air pollution using two microcontrollers, twelve gas sensors, and five environmental sensors. Here Arduino-based ATmega2560 microprocessor is used for onboard input and output facility and also ESP8266 microcontroller which has a built-in Wi-Fi module to connect with the Internet is used for sending data to ThingSpeak cloud server using REST API approach. It generates an alert message to the respective authority when the air quality measurement exceeds the maximum threshold and to serve this purpose the sensors wake and collect data after a fixed schedule which also consumes power.

The measurement of vehicular pollution is discussed in this [12] paper where among emitted gasses of vehicles here measured only two gasses which are carbon monoxide and sulfur dioxide. The carbon monoxide gas sensor and sulfur dioxide gas sensor are embedded in the Arduino-based ATmega328 board microcontroller and also used the radio frequency identification technique. This system connected with motorists and made them aware when the level of pollution surpasses the maximum threshold in a specific area and advised them to avoid that area.

Monitoring indoor air quality is presented in this paper [13], which discusses the HVAC system and mainly elaborates on the incorrect functionality of heating, ventilation, and air-conditioning. Although here the sensor node is ready to install up to ten sensors, there installed only six sensors and those are temperature, humidity, pressure, particulate material, total volatile organic compounds (TVOC), and equivalent CO_2 (eCO_2). All these sensors are integrated into Raspberry Pi 3, and the collected data are stored in the MariaDB database locally after that the data is stored in a cloud-based system. This device was placed in an office workplace, and the result of it shows that the level of TVOC and eCO_2 increase when so many people are in the office.

To monitor the outdoor air quality mainly proposed to use the MQ4 gas sensor to detect methane, the MQ135 gas sensor to detect carbon dioxide, and the DHT22 sensor to detect temperature and humidity. We will integrate all these sensors in the ESP32 microcontroller which has a built-in Wi-Fi Shield, and this feature helps to send data to the ThingSpeak IoT cloud server directly. This device will be portable which will help to collect data from different locations in different seasons.

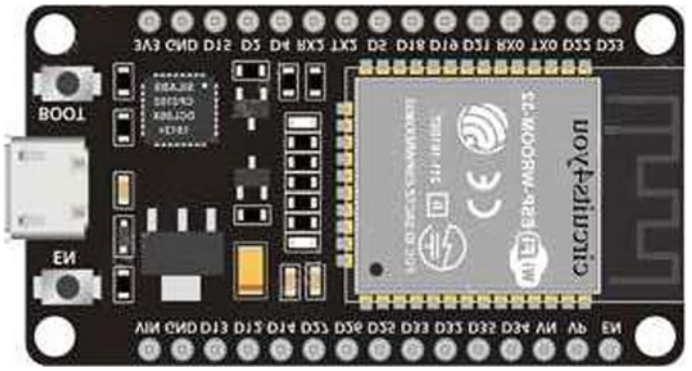


Fig. 1 ESP32 DevKit board

3 Hardware Components

3.1 ESP32 Microcontroller

To implement a sensor-based IoT system, a microcontroller is mandatory where all the necessary components like sensors, regulators, and indicators are integrated. For this reason, we use ESP32 DevKit which is an affordable SoC microcontroller developed by Express if system where SoC stands for system on chip and it has a built-in Wi-Fi shield and Bluetooth feature. It is manufactured using ultra-low-power 40 nm technology which is developed by Taiwan Semiconductor Manufacturing Company (TSMC), and this technology helps to run DC source-related small components like wearable's, audio equipment, baby monitors, and smartwatches. ESP32 DevKit Board contains the ESP-WROOM-32 as the main module and adds some extra hardware to simply program ESP32 and build connections with the GPIO Pins. The ESP32 Board consists of ESP-WROOM-32 Module, where 30 pins are equally distributed on two sides, CP2012 USB – UART Bridge IC, a micro-USB connecter to supply power and insert program code, AMS1117 3.3 V Regulator IC, a Reset Button, a Boot Button for flashing, a Red Power led, a Blue User LED which also connected to GPIO2 and a few passive components [14]. Figure 1 shows the ESP32 DevKit board.

3.2 DHT22: Temperature and Humidity Sensor

On the basis of this work, it is clear that the value of temperature and humidity are vital parameters to measure. For this reason, a sensor named DHT22 is used here to measure both temperature and humidity. Simply join the primary pin or the leftmost pin to the 3-5 V energy pin of the ESP32 board, the second one to our information or

Fig. 2 DHT22 sensor

data pin, and the rightmost pin connected to the ground. A small disadvantage of this sensor is it gets new information from it as soon as every 2 s, so while the usage of the library, sensor readings may be up to two seconds old. This sensor monitors the humidity rate from 0–100% with 2–5 accuracy and monitors the temperature from −40 to 80 °C readings with ± 0.5 °C accuracy [15]. Here Fig. 2 shows the DHT22 sensor which measures temperature and humidity.

3.3 MQ4: Methane Gas Sensor

MQ4 is an alkane series gas sensor that observes the concentration of methane gas within the air and outputs its reading as an analog voltage. The concentration sensing of this sensor varies from 300 ppm to 10,000 ppm. The sensor will operate at temperatures from −10 °C to 50 °C and consume not up to one hundred fifty mA at 5 V. Connecting 5 V across the heating (H) pins keeps the device hot enough to perform correctly. This pin connects at either the A or B pins causing the sensor to emit an analog voltage on the opposite pins. A resistive load between the output pins and ground sets the sensitivity of the detector [16]. Figure 3 is the snapshot of the MQ4 methane gas sensor.

Fig. 3 MQ4 methane gas sensor

Fig. 4 MQ135 gas sensor

3.4 MQ135: Air Quality Sensor

For measuring the CO_2 here uses MQ135 which also can detect some harmful gasses and smoke like ammonia (NH_3), sulfur (S), benzene (C_6H_6). This sensor module operates at 5 V and consumes around 150 mA, and also, it needs some preheating before it might truly provide correct results. The analog output pin of the sensing element is used to live the PPM value of the desired gas and for this issue here need to use an external microcontroller. Figure 4 represents the MQ135 gas sensor.

4 System Architecture

All the previously described components are integrated into the ESP32 DevKit. The circuit diagram below which represents Fig. 5 shows all pin connections of sensors along with the ESP32 board.

Figure 6 shows the block diagram of the above-mentioned work. Here the sensors measure the data from outdoor air and pass it into the ESP32 board and with the help of a built-in Wi-Fi shield system data is passed into the ThingSpeak IoT cloud server.

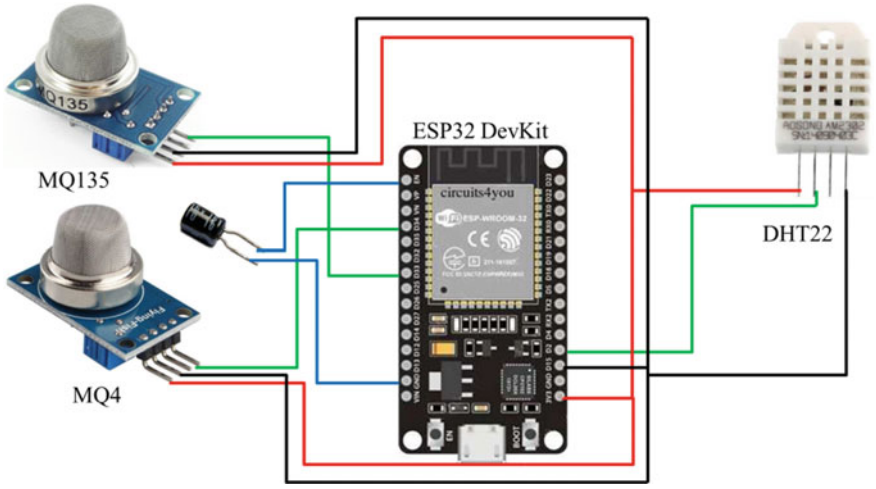


Fig. 5 Circuit diagram

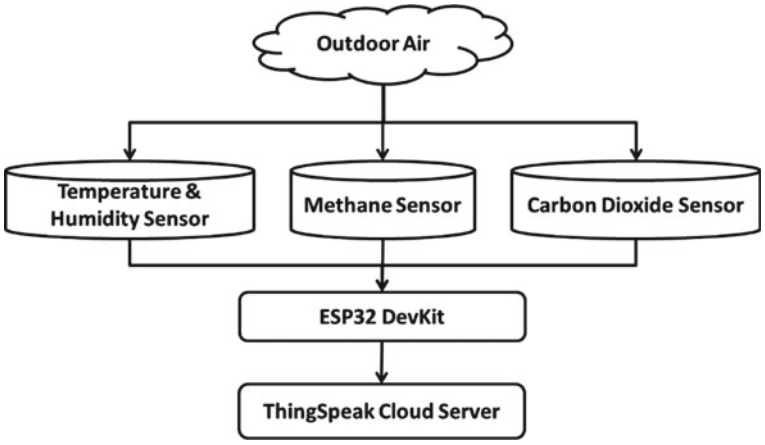


Fig. 6 Workflow diagram

5 System Implementation and Testing

While completing all connections for each component, deploy this portable device in the outdoor environment mostly in urban areas. It collected data in a specific duration and sent it into the ThingSpeak cloud server which is very suitable for IoT-related applications. In ThingSpeak data is stored in four different fields. In the section, Fig. 7 represents the value of the temperature sensor which unit is Celsius and also represents the value of the humidity sensor, MQ4 methane gas sensor, and MQ135 carbon dioxide sensor. The unit of methane and carbon dioxide is ppm.

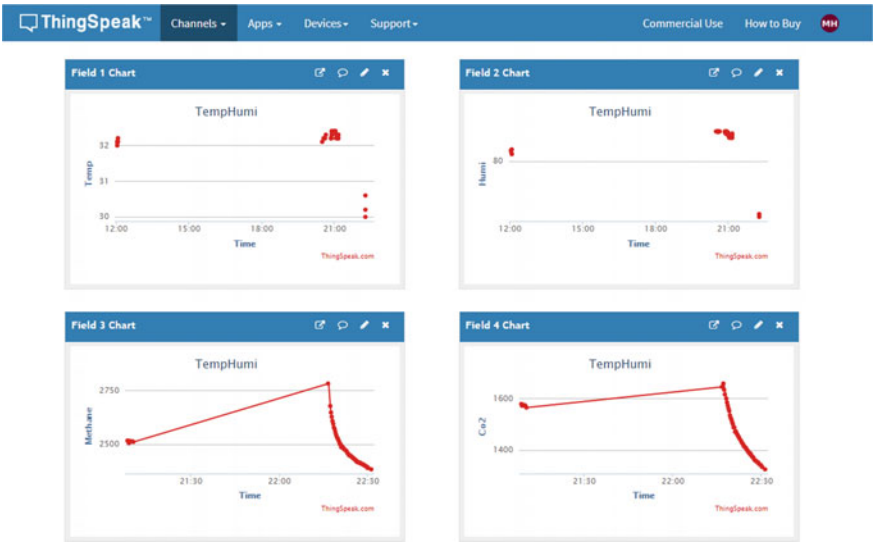


Fig. 7 Data field in ThingSpeak

Figure 8 shows the data sample in excel format where each attribute belongs to each measurement of data. 'Temp' attribute represents the temperature, respectively; 'Humi', 'CH_4', 'CO_2' represent humidity, methane, and carbon dioxide levels.

	A	B	C	D	E	F	G
1	created_at	entry_id	temp	humidity	CH4	CO2	
2	2021-09-10 14:59:37 UTC	161	32.4	88.4	2619	1743	
3	2021-09-10 14:59:53 UTC	162	32.4	88.8	2608	1728	
4	2021-09-10 15:00:08 UTC	163	32.4	88.6	2611	1730	
5	2021-09-10 15:00:24 UTC	164	32.4	88.8	2609	1724	
6	2021-09-10 15:00:41 UTC	165	32.4	88.5	2599	1712	
7	2021-09-10 15:01:11 UTC	166	32.4	88.3	2601	1710	
8	2021-09-10 15:01:27 UTC	167	32.4	88.3	2588	1696	
9	2021-09-10 15:01:43 UTC	168	32.4	88	2583	1689	
10	2021-09-10 15:01:59 UTC	169	32.4	88	2591	1689	
11	2021-09-10 15:02:15 UTC	170	32.4	87.8	2577	1680	
12	2021-09-10 15:02:30 UTC	171	32.4	87.7	2583	1680	
13	2021-09-10 15:02:46 UTC	172	32.4	88	2576	1673	
14	2021-09-10 15:03:02 UTC	173	32.4	87.4	2559	1652	

Fig. 8 Data sheet from ThingSpeak

6 Conclusion

This work represents the development of a portable IoT device for measuring methane and carbon dioxide along with temperature and humidity based on real-time data that is collected from outdoors. We have designed our hardware device with the help of some sensors and used software named ThingSpeak to gather the information of sensors and send it for visualization to the ThingSpeak dashboard. Used advanced IoT technologies to measure the rate of outdoor temperature and humidity quality for the changing level of methane and carbon dioxide. After studying several related works, we say about a hypothesis that there is an impact of methane and carbon dioxide to increase temperature and humidity levels. For this reason, our future goal is to find the correlation of changes in temperature and humidity based on the increasing levels of methane and carbon dioxide level.

Acknowledgements First of all, we would like to thank and affiliate our Department of Computer Science and Engineering, Daffodil International University to provide us the IoT lab facilities for this outdoor-based work which helps to collect lots of data and overcome the different types of barriers.

References

1. Singh, P. and Yadav, D., 2021. Link between air pollution and global climate change. In Global Climate Change (pp. 79–108). Elsevier.
2. Manisalidis I, Stavropoulou E, Stavropoulos A, Bezirtzoglou E (2020) Environmental and health impacts of air pollution: a review. Front Public Health 8:14
3. Vallero, D., 2014. Fundamentals of air pollution. Academic press.
4. Guarieiro LLN, Guarieiro ALN (2013) Vehicle emissions: What will change with use of biofuel. Biofuels-Economy, Environment and Sustainability 1:357–386
5. Schaub, M. and Paoletti, E., 2007. Air pollution and climate change: A global overview of the effects on forest vegetation. *Environmental pollution (1987)*, *147*(3), pp.429–429.
6. Special Report on the impacts of global warming of 1.5 C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. In: Intergovernmental Panel on Climate Change. 2018, IPCC.
7. Wuebbles DJ, Hayhoe K (2002) Atmospheric methane and global change. Earth Sci Rev 57(3–4):177–210
8. How Does Methane Affect the Environment, accessed: 28 January 2022, 00.17 AM (UTC+6). Link: https://www.bridgerphotonics.com/blog/how-does-methane-affect-environment
9. AlShekh RH, Hagem R (2021) Design and Implementation of Smart Air Pollution Monitoring System Based on Internet of Things. Asian Journal of Computer Science Engineering 6(2):25–34
10. Asthana, N. and Bahl, R., 2019, April. IoT device for sewage gas monitoring and alert system. In 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT) (pp. 1–7). IEEE.
11. Fioccola, G.B., Sommese, R., Tufano, I., Canonico, R. and Ventre, G., 2016, September. Polluino: An efficient cloud-based management of IoT devices for air quality monitoring. In 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI) (pp. 1–6). IEEE.

12. Manna, S., Bhunia, S.S. and Mukherjee, N., 2014, May. Vehicular pollution monitoring using IoT. In International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014) (pp. 1–5). IEEE.
13. Esquiagola, J., Manini, M., Aikawa, A., Yoshioka, L. and Zuffo, M., 2018, August. Monitoring indoor air quality by using iot technology. In *2018 IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON)* (pp. 1–4). IEEE.
14. Ravi Teja, 'Getting Started with ESP32 | Introduction to ESP32' accessed: 28 January 2022, 12.14 PM. (UTC+6). Link: https://www.electronicshub.org/getting-started-with-esp32/?nowprocket=1
15. 'DHT22 Digital temperature and humidity sensor' accessed: 28 January 2022, 12.25 PM (UTC+6). Link: https://static6.arrow.com/aropdfconversion/cd4d86e9a91489b38e0ed3f997af8032ef4f8828/pgurl_5149847625733700.pdf
16. 'Methane Gas Sensor MQ-4.' accessed: 28 January 2022, 12.17 PM (UTC+6). Link: https://www.pololu.com/product/1633 MQ135 Semiconductor Sensor for Air Quality. accessed: 28 January 2022, 12.30 PM (UTC+6). Link: https://www.winsen-sensor.com/sensors/voc-sensor/mq135.html

Nationality Detection Using Deep Learning

Md. Abrar Hamim, Jeba Tahseen, Kazi Md Istiyak Hossain, and Saurav Das

Abstract A new intelligent monitoring model is developed by us for determining gender and nationality from frontal picture candidates utilizing the face area based on deep learning. Face recognition is influenced by a number of characteristics, including picture quality, illumination, rotation angle, blockage, and facial expression. As a result, we must first recognize an input image before converting it to a genuine input. Nationality is the most well distinguishing characteristic that is applied in every nation, and it is also important to secure authentication. Image detection is crucial in this case. Then, we can determine the facial shape, gender, and nationality of the candidate image. In the end, we return the result based on the distance comparison using the use of a library to measure the sample. There were significant discrepancies across photographs while measuring samples based on their gender and facial features. The photos used in the input must be the same as those used in the output. Picture of a frontal face with clean lighting and no blemishes at every angle of rotation. The model may be used by ordinary people, models, celebrities, actors, and others. In the end, computer can tell nationality by looking at a picture of a person's face.

Keywords Face detection · Image detection · Face shapes · Gender detection · Nationality recognition

Md. Abrar Hamim (✉) · J. Tahseen · K. M. I. Hossain · S. Das
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh
e-mail: abrar15-11821@diu.edu.bd

J. Tahseen
e-mail: jeba15-11834@diu.edu.bd

K. M. I. Hossain
e-mail: 1istiyakhossain15-4900@diu.edu.bd

S. Das
e-mail: 1saurav15-11851@diu.edu.bd

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_49

1 Introduction

Face recognition gives an outstanding possibility to grow a beneficial software to find out people's identity for the immigration and security workplaces in many countries. In face recognition modeling and processing, the most important issue is to discover a solution for the act for faces that helps for different kinds of tasks regular of human performance. Human facial expression is a complicated visual pattern that includes common categorical statistics in addition to eccentric, primary data, identity particular by using these categorical facts, we are referring that some elements of a face are not particular to that specific face, however, is shared by way of a group of faces. These factors may be used to hire both unusual and acquainted faces to preferred semantic agencies such as place or nationality. Generally, computational patterns of face reputation constitute faces in phrases of geometric identifications that include angles, areas, and distances between base capabilities including mouth, nostril, eyes, or chin. At present, there are a few useful algorithms for identifying face candidate domain regions. Even though a few only work with the manual evaluation, others identify face areas. About the identification of nationality by face candidate, specifically research on detecting nationality has now not been researched but to choose to this problem, first we have got gathered one thousand face images from each country.

2 Related Works

Mainly this system is workable in the plain background picture. Input picture enhanced histogram equalization besides detection based on multiple scales which is called extracted edges linked wing energy function and finally face control extracted applying detection information [1]. This research is based on face key points, face part color, size, and distance. Image quality, lighting, and the facial expression also depend on it. They take real input, identify gender, face shape, and key point, and then give results. This result was found in the measurement database here the most important part is taking input image in frontal no rotation allows. Reference [2], the main work of this paper is detecting human face using skin color and also using here face shape and size. YCBCR and HSV have used the original pictures for quantized skin shade; his period performs rapidly for finding face point. For doing this face processing, a set of easy statistical deviations is created and also created a meaningful feature vector [3]. Because of using simile classifiers and attribute classifiers with LFW methods, its error rates were reduced by 31.68%. The attribute classifiers like "Hair", "Skin", "Flash", and a total of 65 attribute classifiers could detect the visual appearance. Simile classifiers used a total of 60 numbers of people as a reference, and it works on expression, pose, and illumination. For more inquiry, the PubFig dataset is included. Reference [4], this paper worked on a new algorithm named "facial feature detection". It collected images from people ages between 2 and 85. For better

accuracy, four different types of facial pose images were taken. They analyze eyes, lips, and chin from images. This paperwork could be used in facial surgery, age classification, race detection, and other facial feature-related fields. Reference [5], this paper tests residual nets using 152 layers. An examination of CIFAR-10 within 100 to 1000 levels is also presented. Reference [6], Viola-Jones confront finder is utilized to identify the position of the AdaBoost calculation confront. Once more, the geometric proportion of the confront is utilized to identify blunders within the confront and eyes. The bend of a cubic chime is utilized to decide to feel. Reference [7], this paper uses here D-CNN method. We chose to propose an effective convolutional organized VGGnet design that can be utilized in extraordinary cases when the sum of training data utilized to learn D-CNN based on VGGNet design is restricted. Reference [8], this study proposed developing an application that provides users permit to a certain machine depending on a thorough examination of facial features. This project will be built with OpenCV and MS NET framework from Microsoft [9]. The system is made up of three parts and utilizes color pictures. Skin detection is implemented in the first module, which uses Markov random fields models for picture fragmentation and skin recognition. The second module is in charge of detecting and extracting eyes and mouths. The third module uses edge detection as well as measures the variation of the eyes and mouth region figures to identify the emotions depicted in the eyes and lips. Reference [10], in this paper, three labels were developed for an Arab picture dataset GCC, the Levant, and Egypt. The problem was solved using two methods of deep learning. When a pre-trained CNN was trained on the whole Arab dataset and then evaluated on a different dataset, the best result was 56.97%, and 52.12% when the Arab dataset labels were fairly distributed. Reference [11], in this study, we used facial images to train a machine learning classifier to predict Chinese, Japanese, and Korean. Of all the approaches we provided, CNN had the greatest accuracy rate (89.2%) in the three-class categorization of the three different languages [12]. Computer vision can predict nationality based on a person's visual characteristics. This method has achieved an average of 93.6% accuracy with a very low misclassification rate, according to researchers at the University of Michigan [13]. The pre-trained method is used to minimize the dimensionality of the retrieved picture characteristics in a range of video frames so that meaningful and unusual events in the frame can be preserved. For the identification of abnormal circumstances at surveillance footage, the suggested technique is a hybrid of CNN and SVM architectures [14]. In addition, the utility of the deep learning approach for raising the applicability of picture identification with excellent testing performance has been reviewed when it comes to IoT image registration. Furthermore, the fundamental goal of this study project is to assemble a thorough survey data on the different kinds of IoT implementations, as well as security and privacy concern With a higher accuracy rate, privacy issues arise [15] (Table 1).

Table 1 Reference paper overview

Paper name	Method name	Accuracy
[1] Jianguo Wang *, Tieniu Tan, A new face detection method based on shape information, 2000	shape information using histogram equalization	the system can accurately locate 87.5% of the faces
[2] Bat-Erdene Batsukh, Effective Computer Model For Recognizing Nationality From Frontal Image, 2016	using face part color, size, and distances	86.4%
[3] C. Garcia, G. Tziritas**,** Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis,1999	YCbCr and HSV	94.23%
[4] Neeraj Kumar Alexander C. Berg Peter N. Belhumeur Shree K. Nayar, Attribute and Simile Classifiers for Face Verification,2009	“attribute” classifiers and “smile” classifiers were used	85.29%
[5] Azam Bastanfard, Melika Abbasian Nik, Mohammad Mahdi Dehshibi, Iranian Face Database with Age, Pose and Expression,2007	Iranian face database (IFDB)	94.20%
[6] Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun, Deep Residual Learning for Image Recognition, 2015	Deep residual learning	training error < 0.1% test error 7.93%
[7] Rembiye KANDEMİR1, Gonca ÖZMEN2, FACIAL EXPRESSION CLASSIFICATION WITH HAAR FEATURES, GEOMETRIC FEATURES AND CUBIC BÉZIER CURVES, 2013	HAAR FEATURES, GEOMETRIC FEATURES, AND CUBIC BÉZIER CURVES	97%
[8] AmitDhomne, RanjitKumar, VijayBhan, Gender Recognition Through Face Using Deep Learning,2018	D-CNN based on VGGNet	95%
[9] Shervin, Emami Valentin Petrut, Facial Recognition using OpenCV, 2012	OpenCV and MS NET framework	95%

(continued)

Table 1 (continued)

Paper name	Method name	Accuracy
[10] Ilias Maglogiannis Demosthenes Vouyioukas Chris Aggelopoulos, Face detection and recognition of natural human emotion using Markov random fields,2007	Markov random fields	96.68
[11] A Classification of Arab Ethnicity Based on Face Image Using Deep Learning Approach	CNN	75%
[12] Where am I from?—East Asian Ethnicity Classification from Facial Recognitio	CNN	89.2%
[13] A Convolutional Neural Network based Classification Approach for Predicting Nationality from Facial Feature	CNN	93.6%
[14] Design of Deep Learning Algorithm for IoT Application by Image based Recognition	Proposed CNN framework	97%
[15] An Efficient Dimension Reduction based Fusion of CNN and SVM Model for Detection of Abnormal Incident in Video Surveillance	Hybrid proposed	90%

3 Methodology

The human confront is a fantastic source of information for a variety of human social interactions. Identifying a human face for any country's citizens is a challenging task. However, employing machine learning to identify is now a far more viable option. In this case, machine learning seems to do a magical act. It assists us in identifying and distinguishing important human faces and expressions. Though it is a very critical term because individual confront are special. Face shape position and type, distance, and key point are not the same. We actually work here by keeping four different countries male and female. We collected human pictures considered as sample data. In our dataset, we collect every country's public figure face. And our nominated countries are Bangladesh, Pakistan, India (Tamil), and Sri Lanka. We started our process by making architecture, and here, we applied convolutional neural network (CNN) model. A convolutional neural network (CNN) is a kind of artificial neural network employed in image recognition and method that is specifically designed to process element data. CNNs are powerful image processes, AI (AI) that use deep learning to perform each generative and descriptive task, typically

victimization machine vision that has an image and video recognition, in conjunction with recommender systems and language processing (NLP). A neural network is a system of hardware and/or code patterned once the operation of neurons within the human brain. Ancient neural networks are not ideal for image processing and must feed pictures in reduced-resolution pieces. CNN has their "neurons" organized additional like those of the frontal lobe, the space answerable for process visual stimuli in humans Associate in Nursingd alternative animals. The layers of neurons are arranged in such some way on cowl the complete field of regard avoiding the piecemeal image processing downside of ancient neural networks. A CNN uses a system very similar to a multilayer perceptron that has been designed for reduced processing requirements. The layers of a CNN encompass an input layer, an output layer, and a hidden layer that has multiple convolutional layers, pooling layers, fully connected layers, and normalization layers. The removal of limitations and increase in potency for image process results in a system that is much more effective, less complicated to trains restricted for image processing and language processing.

Basically, three main types of layers are used to build convolutional neural network (CNN) architectures:

- Layers of convolution
- A layer of pooling
- Layer with full connectivity.

Large features of a CNN are the first layer that can be recognized and outlined moderately simply. In this model, at first take input, it must be the face of any person, and here, we are using a library which is Autocrop/developed tool. This model takes the image as input and then is divided into different layers. Then lastly, show the final output actually. Our given input is four countries, and for this reason, eight types of output are scanned. Data preprocessing is a very important and significant part of recognizing. Convolutional neural network (CNN) has a place to watch and direct the execution of our action in machine learning for which we need to label the data. We organized our dataset by using the mapping process. And group them 1, 2, 3, 4, 5, 6, 7, 8 by name. Our dataset is very large, and this mapping makes this process easier than before. CNN algorithm takes an image as input and then tests it made on a slandered dataset base for which this method works very fast and reliably. CNN is a class of neural networks that have vindicated exceptionally successfully in ranges. Such as image recognition and it has been compelling in recognizing objects, faces, and activity. Taking an image as input, we have cropped the face using Autocrop tools. Then, we converted the image into numeric data and send data to the trained CNN model and get the highest accuracy result, remap, and index data. Finally, we got our expected result (Fig. 1).

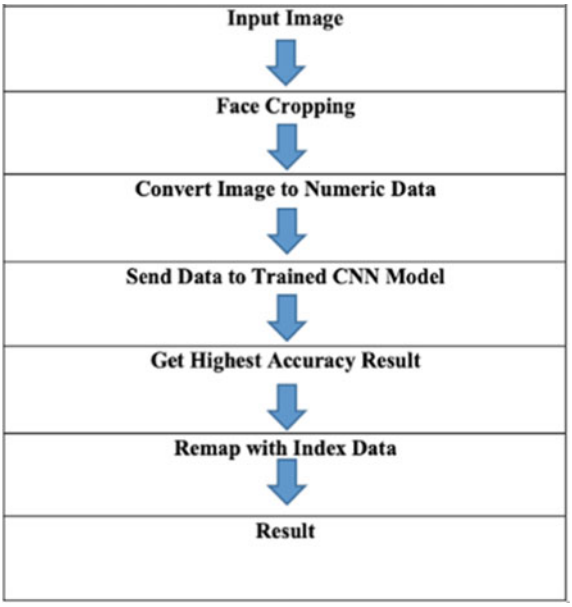


Fig. 1 Work flowchart

3.1 Face Detection and Classification

Face detection is a concept that encompasses both face recognition and face detection. Face detection simply refers to a system's ability to recognize a human face in an image or video. In this paper, photos are autocropped, and control points are extracted using the same segmentation methodology. Face cropping is usually done by hand, and the equations for obtaining a feature point are quite difficult. We used here a third-party library which is Autocrop. This library is quite beneficial. Our input data was filtered using it. It is simple to crop only the face of a photograph. At first, we filtered our all data. This is called leveling. Create a CSB dataset where all data information is stored. Autocrop's unique feature is that all data is produced with the same height and width. We trained our dataset using the trained model and load there for creating a sequential model, and this is our CNN model. For our detection result, we used a base size of 10 and trained 50 times. Finally, compile the model, set the history, and run the program step by step to achieve our aim (Fig. 2).

3.2 Identifying Gender

The picture candidate's gender was an interval exception while we were obtaining measurements from the sample photographs. Gender classification must have been

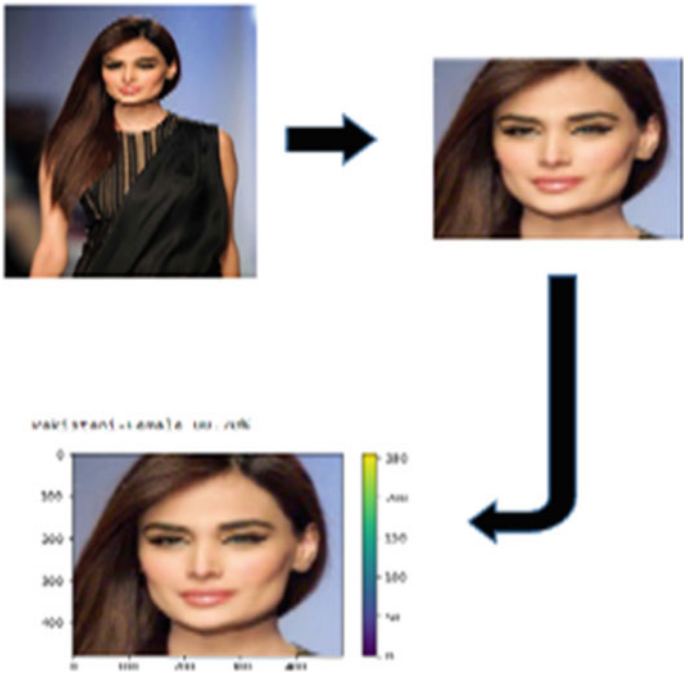


Fig. 2 Face detection and classification

used in real biological rules, particularly in anthropological and morphological analyses. One of the most problematic aspects of describing faces as a set of features is that it assumes some level of a priori knowledge. It is challenging to come up with a set of traits that can accurately be separated between male and female faces. Men and women were initially placed in separate folders when gender was determined. We worked for four separate countries (Bangladesh, Indian Tamil, Pakistan, and Sri Lanka) and synchronized those folders by number since we worked for men and women from four different countries. In our dataset, we divided it into four categories: one for Bangladeshi males, two for Bangladeshi women, three for Pakistani men, and four for Pakistani women. Our dataset already knows this leveling and loading there for creating a sequential model, and this model identifies gender by detecting that folder.

3.3 Recognizing Nationality

A person's nationality is a well-known distinguishing trait that is utilized in every country for important verification purposes. The size and spacing of face parts are

determined by a number of parameters, including picture quality, and facial expression is the process of identifying nationality. As a result, we must first recognize a face in the image before converting it to a real-world input. Candidate may then be determined by the gender, facial shape, and portion of the face of the picture. Finally, we will return the result, which will be based on our database's produced folder. There were significant variations between photos using these leveling folders when we were measuring samples. The frontal face image, with smooth illumination and no rotation angle, must be used as the input image. We run a trained test on our model and provide a model summary. Finally, design a confutation matrix utilizing this accuracy. It actually aids us in detecting nationality and identifying similarities between people from other countries (Figs. 3 and 4).

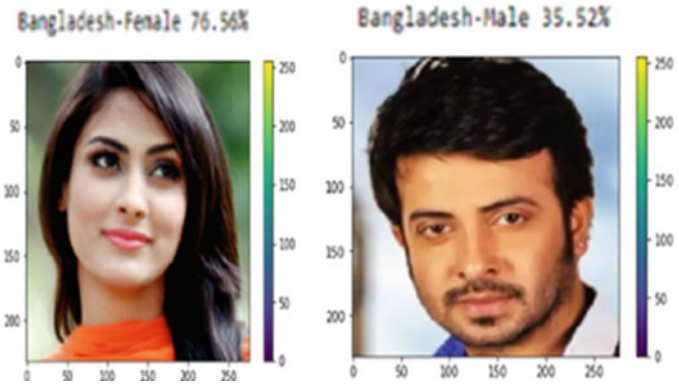


Fig. 3 Bangladeshi male and female

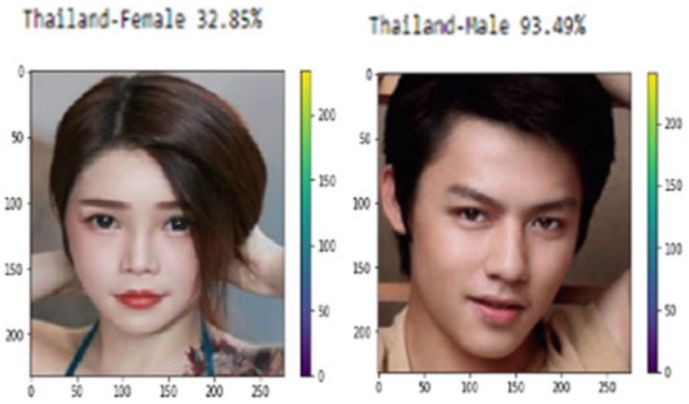


Fig. 4 Thailand male and female

4 Evaluation and Result

After collecting data, we used Autocrop to separate frontal face images from data for training. In our model, we have a total parameter of 3,380,440, the trainable parameter of 3,379,832, and the non-trainable parameter of 608. We trained our 10-batch data 50 times. Here, some results are given below. Showing here some trained process. We know that the ratio of accurately expected data points among all the data points is called accuracy. In our first figure, it is clear that our beginning accuracy level was not very good, so we trained it, and now, it has the highest accuracy. A loss is the result of a wrong prediction. When loss decreases, accuracy increases gradually.

From Fig. 5, we observe that accuracy rate keeps increasing in every step of training. And then we get up to 94.65% accuracy which is actually trained accuracy and 55.97% Val accuracy (value accuracy), which is acceptable.

From Fig. 6, we can see from our loss-v-loss (validation loss) function graph, the loss is decreasing which is 8.34%. When our data is trained, the validation loss is determined on the v-set (validation set).

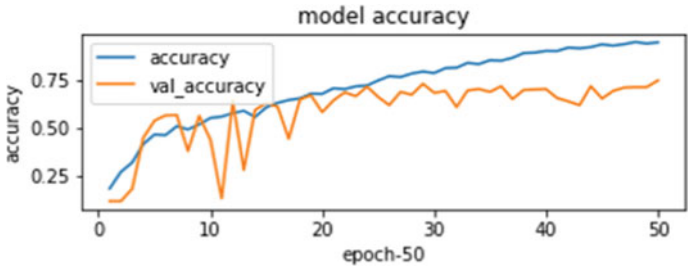


Fig. 5 Model accuracy

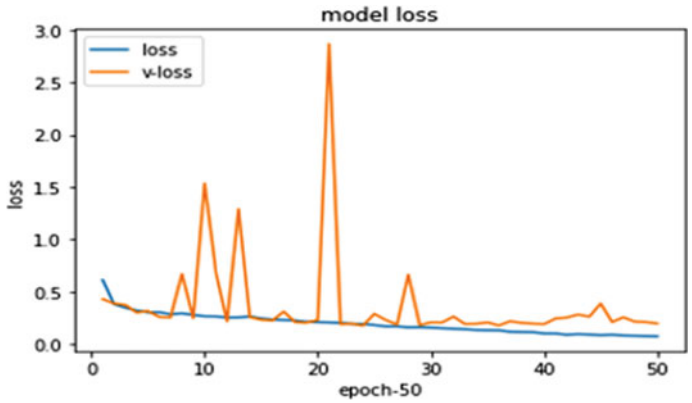


Fig. 6 Model loss

5 Discussion and Conclusion

In this paper, we proposed a model that could detect the nationality and gender of people in a very simple and effective way using frontal images. In the first place, we need to train our system with some targeted data we want to detect. We have achieved up to 94% accuracy with our small dataset. This model is useful for frontal face color images and sufficient lighting conditions. With this model, we can identify people's nationalities easily. And it will be more effective if we train our system with a larger dataset. In a recent global world, finding out nationality is a challenging step. By detecting nationality immigration office, defense and government could make important decisions. It could be used in business modeling and online marketing.

This project is mostly employed in the field of social work. Persons may use it to detect the gender and nationality of people from any country. Our long-term objective is to transform it into an Android application. The Android application has become more user-friendly in recent years. It may be simple to use and comprehend. As a result, we plan to create an Android application in the future.

References

1. Wang J, Tan T (2000) A new face detection method based on shape information. Pattern Recogn Lett 21(6–7):463–471
2. Batsukh BE, Tsend G (2016) Effective computer model for recognizing nationality from frontal image. *arXiv preprint* arXiv:1603.04550
3. Garcia C, Tziritas G (1999) Face detection using quantized skin color regions merging and wavelet packet analysis. IEEE Trans Multimedia 1(3):264–277
4. Kumar N, Berg AC, Belhumeur PN, Nayar SK (2009) Attribute and simile classifiers for face verification. In: 2009 IEEE 12th international conference on computer vision. IEEE, pp 365–372
5. Bastanfard A, Nik MA, Dehshibi MM (2007) Iranian face database with age, pose and expression. In: 2007 International conference on machine vision. IEEE, pp 50–55
6. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
7. Kandemir R, Özmen G (2013) Facial expression classification with haar features, geometric features and cubic bã‰ zier curves. IU-J Electr Electron Eng 13(2):1667–1673
8. Dhomne A, Kumar R, Bhan V (2018) Gender recognition through face using deep learning. Procedia Comput Sci 132:2–10
9. Emami S, Suciu VP (2012) Facial recognition using OpenCV. J Mob Embedded Distrib Syst 4(1):38–43
10. Maglogiannis I, Vouyioukas D, Aggelopoulos C (2009) Face detection and recognition of natural human emotion using Markov random fields. Pers Ubiquit Comput 13(1):95–101
11. Al-Humaidan NA, Prince M (2021) A classification of arab ethnicity based on face image using deep learning approach. IEEE Access 9:50755–50766
12. Chen H, Deng Y, Zhang S (2016) Where am i from?–east asian ethnicity classification from facial recognition. Project study in Stanford University.
13. Junayed MS, Jeny AA, Neehal N, Ahmed E, Hossain SA (2018) Incept-N: a convolutional neural network based classification approach for predicting nationality from facial features. In: International conference on recent trends in image processing and pattern recognition. Springer, Singapore, pp 466–475

14. Sharma R, Sungheetha A (2021) An efficient dimension reduction based fusion of CNN and SVM model for detection of abnormal incident in video surveillance. J Soft Comput Paradigm (JSCP) 3(02):55–69
15. Jacob IJ, EbbyDarney P (2021) Design of deep learning algorithm for IoT application by image based recognition. J ISMAC 3(3):276–290

A Novel Scheme to Deploy the Throwboxes in Delay Tolerant Networks

Nidhi Sonkar, Sudhakar Pandey, and Sanjay Kumar

Abstract Delay Tolerant Networks, the heterogeneous wireless network, are having the ability to enable communication in intermittent connectivity. But, the network suffers from poor performance due to low contact opportunity between nodes. Contact opportunity between nodes must be increased to improve the performance of the network. For that we can deploy static relay nodes also called as throwboxes in optimistic places so that more number of nodes can communicate by that node. Hence, performance of the network can be improved. In this paper, we proposed a deployment technique using grey wolf optimization (GWO), a metaheuristic technique to increase delivery ratio in minimum delay. In this paper, the objective is to find the optimal places using GWO hence, delivery ratio of the network can be increased with least end-to-end delay. For simulation, we have used ONE simulator to find the result and comparison with previously proposed deployment schemes in DTN. By analysing results, we can say that proposed deployment technique performs better as compared to existing deployment technique.

Keywords Delay Tolerant Networks · Deployment · Static relay nodes · Throwboxes · Grey wolf optimization · Opportunistic network environment

1 Introduction

Delay Tolerant Networks have the ability to create communication in the environment where connectivity issues arise or no connectivity available. When end-to-end connectivity not available between the nodes in the networks, it is not possible to send messages in traditional protocols. But, DTNs used Store carry forward strategy to enable communication between them. That is data can be transmitted when two nodes encounters in the communication range of each other [1]. In literature, DTNs have been called by different names like Disruption Tolerant Networks, sparse mobile ad hoc networks, extreme wireless networks [2]. Delay Tolerant Networks have number

N. Sonkar (✉) · S. Pandey · S. Kumar
Department of Information Technology, National Institute of Technology Raipur, Raipur, India
e-mail: nidhisonkar055@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_50

of applications in different areas like underwater communications [3], space networks [4], wildlife tracking [5], remote area networks [6] and vehicular ad hoc networks [7] For DTNs, many algorithms have been developed, but their delivery ratio in minimum latency is very poor. Therefore, for improving delivery ratio with minimum latency contact opportunity between nodes must be improved. For improvement in performance of DTN in most acquired approach is to deploy some static node also called as throwboxes in optimistic places so that nodes contact each other by that static node [8].

Throwboxes are the small storage devices that act as a relay node and buffer for the node that could not meet regularly [9]. The basic idea behind throwboxes is to improve performance of networks in terms QoS (Quality of Service) parameters such as delivery ratio and end-to-end delay and minimize the load on mobile nodes. Throwboxes work as an intermediate between two nodes that could not meet regularly but visit some hotspot places like bus stand, market, etc. in different time. Therefore, we can place the throwboxes in that places so that they can communicate to each other. Throwbox forwards message to the destination node or better node having probability of transmitting message to the destination. If node encounters to the throwboxes and message forwarded to the throwbox if source node visit that place. We can see in Fig. 1, direction of movement of node A and B is different so they cannot communicate to each other at any cost. If we deploy a relay node between them then they can communicate and share data to each other. When node A encounters

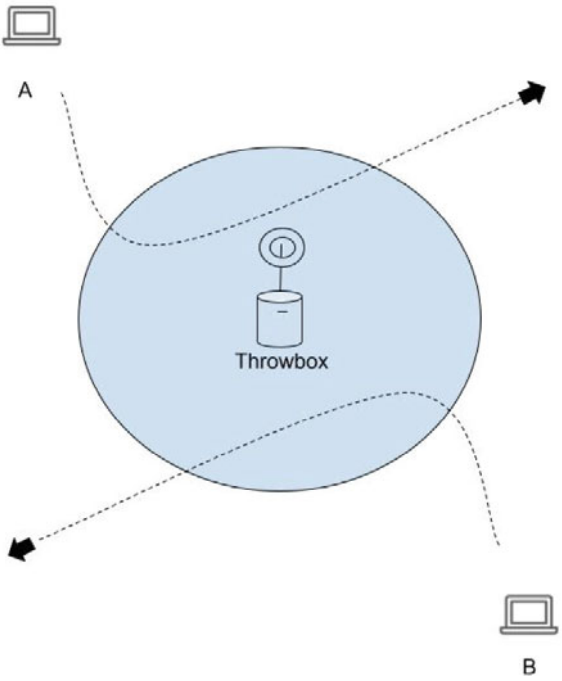


Fig. 1 Throwboxes in delay tolerant networks

in the communication range of throwbox, it forwards the data to throwbox then the relay node (throwbox) forward to node B when it comes in the contact of relay node.

There are many applications of throwboxes proposed by authors are in Ref. [10]. According to Ref. [8], there are two types of throwbox placement: random throwbox placement and planned throwbox placement. Random throwbox placement [11] strategy deploy throwboxes randomly in network area without considering any network factors. However, the cost of random deployment is very low and very easy to implement, but the performance of network may degrade in random deployment technique like lower delivery ratio and higher delays. Hence, we can say random deployment technique is not efficient for relay node placement. In the other hand, the planned deployment technique throwboxes are placed efficiently by considering network topology, routing performance, patterns of nodes etc. Hence, authors preferred to propose planned deployment technique so that network performance could be improved. Number of deployment techniques are proposed for DTN, but best of our knowledge none of them has used any optimization technique to deploy the relay nodes due to the dynamic and challenging environment. Hence, the performance of existing techniques is not satisfying.

Optimization techniques are being used to improve the QoS in the networks [12, 13]. In this paper, a novel scheme to deploy throwboxes in DTN is proposed based on grey wolf optimization (GWO) technique. GWO [14] is the new metaheuristic technique to solve the problems of optimization. GWO is inspired by the hunting mechanism of pack of wolves who are trying to hunt prey. GWO is easy to implement and efficient as compared to other optimization technique. Hence, we have used GWO algorithm to find the best place to deploy the relay nodes in the network. We have considered three parameters of nodes to find the best using GWO that are: distance between mobile nodes to throwbox, residual energy of throwbox and node degree of throwbox. The main goal of this paper is to find the place where, distance between mobile nodes, is less with deployed nodes, less energy used and high node degree that is connectivity between other nodes to deployed node should be high. We have used GWO technique to find the optimum values of above-mentioned parameters. We have compared the proposed technique with existing deployment techniques. By analysing the results, we can say that proposed method have improved the performance of the network in terms of delivery ratio and end-to-end delay.

The whole paper is organized as follows: Section 2 described the existing deployment technique in terms of related work, Sect. 3.1 presented the concept of grey wolf optimization, Sect 3.2 presented the proposed deployment technique using GWO. Section 4 described the simulation and results of proposed method as compare to existing deployment technique. Finally, paper is concluded in Sect. 5 using conclusion and future works.

2 Related Works

In Delay Tolerant Networks, researchers have proposed many routing protocols. The simplest version of DTN routing is the epidemic [15] routing protocol that is flooding-based protocol. İn epidemic routing node flooded message to each node that comes in the communication range and the nodes which have message also forward message to every node. This forwarding process continue till message reach to the destination. İn this technique, delivery ratio is highest, but overhead ratio is also very high. Spray and Wait [16] protocol is also flooding-based technique that works in two phases. İn spray phase node froward L/2 copies of messages to the all encountered node and in wait phase source node waits. İf message reached to the destination, the process is completed else remaining L/2 copies of messages forwarded to the encountered node. Hence, less number of copies of messages wasted and delivery ratio is high. Prophet algorithm is proposed by the authors in Ref. [17], in which message forwarded to the node that has more number of encounters in the history so that the probability to reach the destination could be improved. In Ref. [18] number of important routing protocols are proposed to improve the performance of the network. Routing protocols based on social characteristics of nodes are proposed in Ref. [19] that improves the performance of network. Some buffer management techniques are proposed in Ref. [20] with new parameters to reduce the message lost probability and improve the performance of the network.

The idea of deployment of throwboxes or static relay nodes is proposed in Ref. [21]. The throwbox deployment is coupled with routing techniques to improve the data rate. The architecture of energy-efficient throwbox deployment technique is proposed in Ref. [22]. İn which authors successfully increased the number of delivered bytes to the destination. Social parameter-based throwbox deployment technique is proposed in Ref. [23] by discovering data of real-life tracing. The objective of the authors is to minimize the resource consumption. Several relay strategies are proposed in Ref. [24] to evaluate their performance using mathematical framework. In Ref. [25], authors tried to minimize the number of deployed relay nodes in vehicular DTN under flow controls. However, none of them used optimization algorithms to deployment so that performance can be improved. In this paper, we have proposed a novel approach to deploy the throwboxes using grey wolf optimization technique. We have compared the proposed deployment technique with existing deployment techniques and find that proposed method has improved the performance.

3 Methodology

The objective of this paper is to deploy static relay nodes in the optimistic places so that the performance of the network can be improved. We have used GWO algorithm to find the optimum places. İn this section, we first explain the terminology of GWO

in first subsection and the proposed deployment algorithm using GWO in second subsection.

3.1 Grey Wolf Optimization (GWO)

GWO [13] is a metaheuristic technique to solve the optimization problem inspired by the hunting mechanism of wolves. GWO works based on the hierarchy of leadership of the pack of grey wolf's hunting technique. GWO technique is comparatively new optimization algorithm and least researched. The mathematical model of GWO can be derived by the pattern of observation of hunting mechanism of grey wolves. Then the equations obey the problems to find the optimized solutions.

The solution of the problem can be classified on the basis of hierarchy of the structure of grey wolf's social dominance. That means, the best solution is described as α, second best solution is termed as β and third best solution is described as δ. All the remaining solutions apart from these are considered as Ω that are the least fit solutions. The whole process of GWO is based on the hunting process of grey wolves guided by the solutions α, β, δ and Ω.

In the hunting paradigm of grey wolves, they surround the prey by enclosing in. We can depict the surrounding behaviour of wolves in the mathematical formula which is as follows:

$$\overrightarrow{D} = |\overrightarrow{C}.\overrightarrow{X}_P(t) - \overrightarrow{X}_t| \tag{1}$$

$$\overrightarrow{X}(t+1) = \overrightarrow{X}_P(t) - \overrightarrow{A}.\overrightarrow{D} \tag{2}$$

where $\overrightarrow{A}$ and $\overrightarrow{C}$ are coefficient vectors, t is the iteration number, $\overrightarrow{X}_P$ is the vector position of the prey and $\overrightarrow{X}$ is the position of grey wolf in the form of vector. The vector $\overrightarrow{A}$ and $\overrightarrow{C}$ can be calculated by following equations.

$$\overrightarrow{A} = 2.\overrightarrow{a}.\overrightarrow{r}_1 - \overrightarrow{a} \tag{3}$$

$$C = 2.\overrightarrow{r}_2 \tag{4}$$

where $\overrightarrow{r}_1$ and $\overrightarrow{r}_2$ are random vectors ranges [0, 1] and the value of $\overrightarrow{a}$ decreased linearly from 2 to 0 as the repeated iterations.

The hunting process is done by the best solution, i.e. α and β, and it is assumed that the location of prey is unknown. The least important solutions Ω update locations of wolves according to the information of the best wolves α and β. We can understand the mathematical equations of this process which are as follows.

$$\overrightarrow{D}_{\alpha} = |\overrightarrow{C}_1.\overrightarrow{X}_{\alpha} - \overrightarrow{X}| \tag{5}$$

$$\overrightarrow{D}_{\beta} = |\overrightarrow{C}_2.\overrightarrow{X}_{\beta} - \overrightarrow{X}| \tag{6}$$

$$\overrightarrow{D}_{\delta} = |\overrightarrow{C}_3.\overrightarrow{X}_{\delta} - \overrightarrow{X}| \tag{7}$$

$$\overrightarrow{X}_1 = \overrightarrow{X}_{\alpha} - \overrightarrow{A}_1.(\overrightarrow{D}_{\alpha}) \tag{8}$$

$$\overrightarrow{X}_2 = \overrightarrow{X}_{\beta} - \overrightarrow{A}_2.(\overrightarrow{D}_{\beta}) \tag{9}$$

$$\overrightarrow{X}_3 = \overrightarrow{X}_{\delta} - \overrightarrow{A}_3.(\overrightarrow{D}_{\delta}) \tag{10}$$

$$\overrightarrow{X}(t+1) = \frac{\overrightarrow{X}_1 + \overrightarrow{X}_2 + \overrightarrow{X}_3}{3} \tag{11}$$

The last and final step of grey wolves is to attack the prey. Using the above stated equations, we can compute the attacking position by analysing the fitness function.

3.2 Proposed Method

The paper aims to deploy the static relay nodes or throwboxes in the optimum places so that performance of the network can be improved. We have used GWO algorithm to find the optimum places. As the concept of GWO algorithm is described in the previous subsection. Here, we considered throwbox deployment problem as hunting problem of grey wolf. Grey wolves are considered as the throwboxes and prey is considered as the place to deploy the throwboxes. In the deployment process, we first deploy throwboxes randomly and find the fitness function and then move the throwbox using map-based mobility model and then find fitness function. We follow this process till the maximum iteration reached. Finally we deploy throwboxes in the place where value of fitness function is maximum. To derive the fitness function used parameters are described as follows:

Node degree: Node degree can be calculated as the total number of nodes encountered to the particular throwbox. Node degree should be maximized to enhance the contact opportunity for a throwbox in the network.

$$N_D = \sum E \tag{12}$$

where E is the total number of encounters.

Residual Energy: Residual energy is the remaining energy of a node in the network can be capture at the time encounters. Energy is used for transmission and receiving process. Hence, remaining energy should be high. We assigned maximum energy to every throwboxes when they change their places.

$$R_E = RE \tag{13}$$

where RE is residual energy.

Distance: Distance of the encountered node to the throwbox can be calculated using Euclidian distance and distance between nodes should be minimum.

$$D_T = \text{dis}(n_i, n_j) \tag{14}$$

where n_i is the throwbox and n_j is encountered node.

Hence, the fitness function of proposed method can be written as shown in Eq. 15.

$$\text{Maximize} F = \alpha 1 \times N_D + \alpha 2 \times R_E + \alpha 3 \times (1/D_T) \tag{15}$$

Equation 15 represents, F is the fitness function that is function, where $\alpha 1, \alpha 2 \text{and} \alpha 3$ are the weights of each sub objective where summation of all three parameters is equal to 1.

The flowchart of proposed method is presented in Fig. 2, and the algorithm of proposed algorithm is presented as follows.

Algorithm

1. Initialize the population size and maximum number of iterations.
2. Randomly deploy the throwboxes in the network.
3. While ($I < I_{\max}$) for each iteration do
4. for each throwbox
5. calculate fitness function using Eq. 15.
6. Move the throw according to map-based mobility go to 5.
7. end
8. end
9. place the throwbox in the location (x, y) with maximum fitness function.

4 Simulation and Results

For simulating the environment for proposed GWO-based deployment technique we have used ONE simulator. ONE simulator stands for opportunistic network environment [26] that is based on java programming, actually implemented for applications of Delay Tolerant Networks. Simulation parameters are shown in Table 1. We have considered default Helsinki city map for simulation with area of 4500 × 3400 m^2

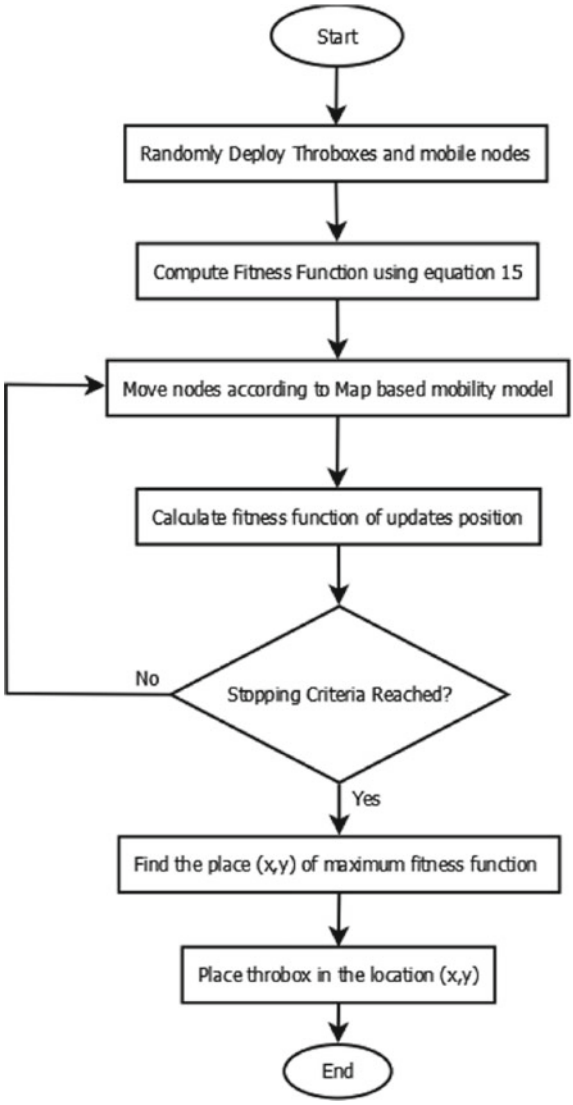


Fig. 2 Flowchart of proposed method

with simulation time of 5000 s. Total number of mobile nodes considered as 100 and 20 static relay nodes are considered. Both mobile and static nodes are first randomly deployed and then move according to mobility models. For mobile nodes random way point mobility model is used and for static relay node we move them according to map-based mobility model till the maximum number of iterations, and then we deploy them in appropriate places according to the algorithm proposed. We have considered up to 50 M of buffer size in nodes. Every node has 100 m of communication range and message size is considered as 100 kb. TTL of every message is taken

Table 1 Parameters to simulate the environment

Parameters	Values
Number of throwboxes	20
Message size	100 kb
Transmission range	100 m
Buffer size	10–50 M
Simulation area	4500 × 3400 m^2
Mobility model	Random Waypoint
Time to live (TTL)	300 min
Simulation time	5000 s
Message generation rate	5–30 message/min
Number of mobile nodes	20–100
Mode of transmission	Bluetooth
Data rate	2048 Bytes per second
Number of iteration	100

as 300 min which is default TTL in ONE simulator. For comparing the proposed method with existing deployment techniques, buffer size is varied from 10 to 50 MB and number of mobile nodes from 20 to 100 nodes to get delivery ratio and end-to-end delay. For the comparison, we have selected three existing algorithms first is NTB (No Throwbox) [17] in which no throwboxes are deployed for message transmission. Second method is RTB (Random Throwbox) [11] in which number of throwboxes are randomly deployed and third algorithm is STB (Social Throwbox) [23] in which throwboxes are deployed based on social parameters of networks. The comparison results as described in the following subsections.

1. Delivery Ratio

   Delivery ratio can be calculated as the ratio of total number of messages generated at the source node to the number of messages reached at destination node.

Figure 3 represents the comparison of delivery ratio with variation of buffer size from 10 to 50 MB. We can see the delivery increases for all the algorithms when buffer size increases. But the delivery ratio is highest for proposed deployment method and minimum for no throwbox technique. The reason behind this increment is proposed method selects the efficient place, where amount of connectivity with the throwbox is increased hence, number of nodes transfers message to the node and node transfer it to destination. Hence, the delivery ratio has been increased for proposed method when buffer size varies. In Fig. 4, the improvement in delivery ratio is presented when number of mobile nodes varies from 20 to 100. We can observe that when mobile nodes increase the delivery ratio has been increased and highest for proposed method for static node deployment. The reason behind this increment is when number of nodes increases the opportunity of

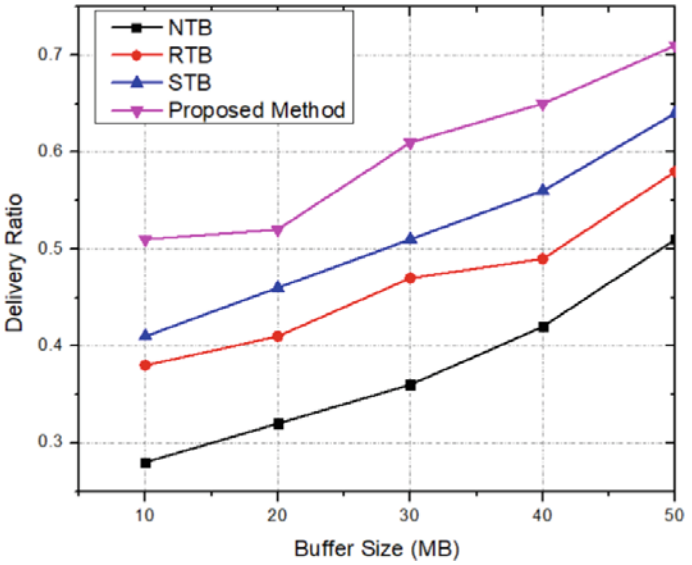


Fig. 3 Improvement in delivery ratio with respect to buffer size

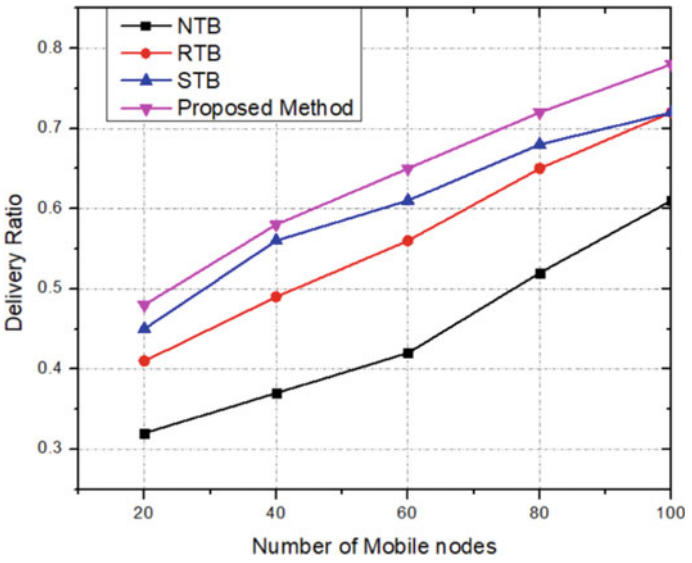


Fig. 4 Improvement in delivery ratio with respect to number of mobile nodes

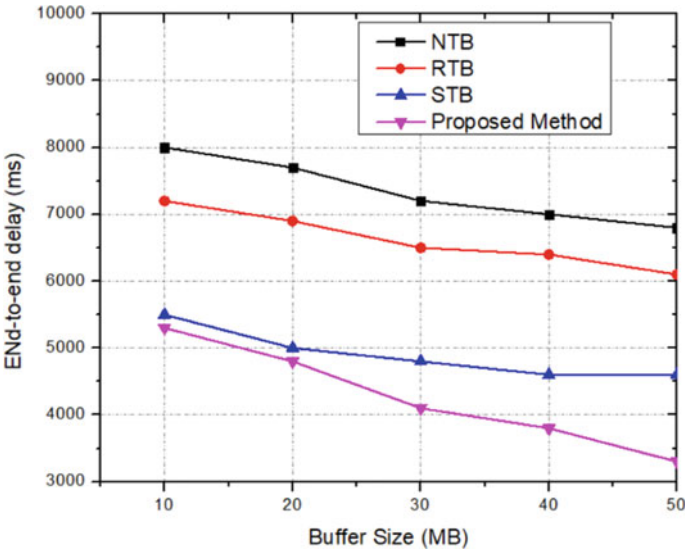


Fig. 5 Decrement is end-to-end delay with respect to buffer size

message transmission increases. Also, we have selected the important parameters of place for node deployment hence, the delivery ratio has been increased.

2. End-to-end Delay

   End-to-end delay can be calculated as the total time to transfer message from source node to destination node.

In Fig. 5, the decrement in end-to-end delay with respect to buffer size is presented. We can observe that delay is least for proposed deployment technique and highest for no throwbox technique. The reason behind this decrement is when we increase buffer size less number of message deleted from the buffer and more number of message reached to the destination. İn proposed method, we have selected important parameters of the place; hence, the delay is least for the proposed method. In Fig. 6, the decrement in end-to-end delay with respect to number of mobile nodes is presented. We can observe that when we increase number of mobile nodes the opportunity increases to forward message and static nodes are placed in the efficient places hence messages need not to wait for the connectivity. Hence, the end-to-end delay has been decreased.

5 Conclusion and Future Work

In this paper, a novel scheme to deploy throwboxes in Delay Tolerant Networks have been proposed. Grey wolf optimization technique is used to find the optimal

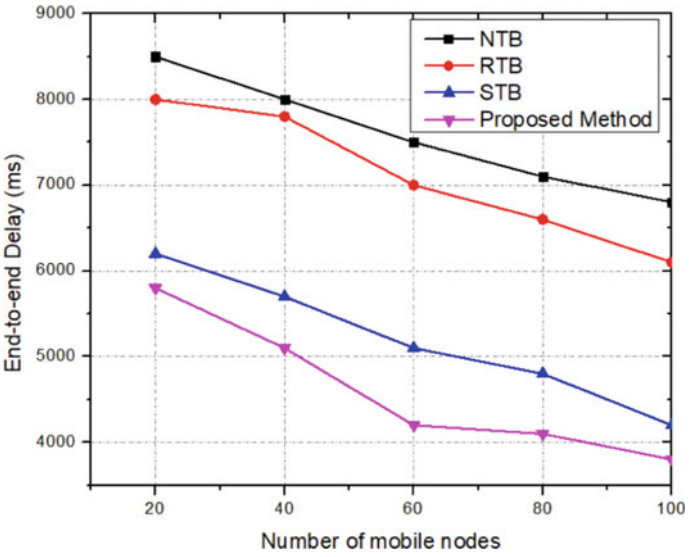


Fig. 6 Decrement in end-to-end delay with respect to number of mobile nodes

places in the network to maximize contact opportunity; hence, delivery ratio can be increased with minimum delay. For the fitness function computation, we have selected three important parameters that are node degree, residual energy and distance between mobile nodes to the throwboxes. Proposed method is compared with existing deployment technique and analysed that proposed method has higher delivery ratio and least end-to-end delay. For future work, we can consider the more number of parameters for computing fitness function so that performance can be increased.

References

1. Fall K (2003) A delay-tolerant network architecture for challenged internets. In: Proceedings of the 2003 conference on applications, technologies, architectures, and protocols for computer communications—SIGCOMM'03, p 27. https://doi.org/10.1145/863956.863960.
2. Sun W, Liu C, Wang D (2012) On delay-tolerant networking and ıts application. Delay Tolerant Netw Appl Wei 51(Iccsit):238–244. https://doi.org/10.7763/IPCSIT.2012.V51.42
3. Geng X (2017) LanePost: lane-based optimal routing protocol for delay-tolerant maritime networks
4. Zhang M, Wolff RS (2010) A border node based routing protocol for partially connected vehicular Ad Hoc networks. J Commun 5(2):130–143. https://doi.org/10.4304/jcm.5.2.130-143
5. Juang P, Oki H, Wang Y (2002) Energy-efficient computing for wildlife tracking: design tradeoffs and early experiences with zebranet
6. Zhao W, Ammar MH, Zegura EW (2004) A message ferrying approach for data delivery in sparse mobile ad hoc networks. In: Proceedings of 5th ACM International Symposiun of Mobile Ad Hoc Networking and Computing—MobiHoc'04, pp 187–198. https://doi.org/10.1145/989459.989483

7. Ahmed SH, Kang H, Kim D (2015) Vehicular delay tolerant network (VDTN): routing perspectives. In: 2015 12th Annual IEEE Consumer Communication Networks Conference CCNC, pp 898–903. https://doi.org/10.1109/CCNC.2015.7158095
8. Qirtas MM, Faheem Y, Rehmani MH (2017) Throwboxes in delay tolerant networks: a survey of placement strategies, buffering capacity, and mobility models. J Netw Comput Appl 91(September):89–103. https://doi.org/10.1016/j.jnca.2017.04.004
9. Zhang Z (2006) Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges. IEEE Commun Surv Tutor 8(1):24–37. https://doi.org/10.1109/COMST.2006.323440
10. G. A. 955. Hillary Jr, "No Title," Defin. Community Areas Agreement. Rural Sociol
11. Chuah M, Xi Y (2007) Enhanced delivery in disruption tolerant networks using advantaged nodes with directional antenna capability. In: Proceedings of IEEE Military Communications Conference MILCOM. https://doi.org/10.1109/MILCOM.2007.4455102
12. Smys S, Haoxiang W (2021) A secure optimization algorithm for quality-of-service ımprovement in hybrid wireless networks. IRO J Sustain Wirel Syst 3(1):1–10
13. Chen JIZ, Smys S (2020) Optimized dynamic routing in multimedia vehicular networks. J Inf Technol 2(03):174–182
14. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. Adv Eng Softw 69:46–61
15. Lu X, Hui P (2010) An energy-efficient n-epidemic routing protocol for delay tolerant networks. In: 2010 IEEE fifth ınternational conference on networking, architecture, and storage. IEEE, pp 341–347
16. Spyropoulos T, Psounis K, Raghavendra CS (2005). Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In: Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking, pp 252–259
17. Lindgren A, Doria A, Schelen O (2004) Probabilistic routing in intermittently connected networks. In: International workshop on service assurance with partial and ıntermittent resources. Springer, Berlin, Heidelberg, pp 239–254
18. Sobin CC, Raychoudhury V, Marfia G, Singla A (2016) A survey of routing and data dissemination in delay tolerant networks. J Netw Comput Appl 67:128–146
19. Zhu Y, Xu B, Shi X, Wang Y (2012) A survey of social-based routing in delay tolerant networks: Positive and negative social effects. IEEE Commun Surv Tutor 15(1):387–401
20. Jain S, Chawla M (2014) Survey of buffer management policies for delay tolerant networks. J Eng 2014(3):117–123
21. Zhao W, Chen Y, Ammar M, Corner M, Levine B, Zegura E (2006) Capacity enhancement using throwboxes in dtns. In: IEEE international conference on mobile Adhoc and sensor systems (MASS), Oct 2006, pp 31–40
22. Banerjee N, Corner M, Levine B (2010) Design and field experimentation of an energy-efficient architecture for dtn throwboxes. IEEE/ACM Trans Netw 18(2):554–567
23. Ying Z, Zhang C, Wang Y (2014) Social based throwbox placement in large-scale throwbox-assisted delay tolerant networks. In: Proceedings of IEEE ICC, June 2014, pp. 2472–2477.
24. Ibrahim M, Nain P, Carreras I (2009) Analysis of relay protocols for throwbox-equipped dtns. In: Proceedings of IEEE WiOPT, June 2009, pp 1–9
25. Farahmand F, Cerutti I, Patel A, Zhang Q, Jue J (2008) Relay node placement in vehicular delay-tolerant networks. In: Proceedings of IEEE GLOBECOM, Nov 2008, pp 1–5
26. Liu X, Chen Y (2013) Report of a DTN simulator-THE ONE

Forest Fire Prediction Using Machine Learning and Deep Learning Techniques

M. Shreya, Ruchika Rai, and Samiksha Shukla

Abstract Forests are considered synonyms for abundance on our planet. They uphold the lifecycle of a diversity of creatures, including mankind. Destruction of such forests due to environmental hazards like forest fires is disastrous and leads to loss of economy, wildlife, property, and people. It endangers everything in its vicinity. Sadly, the presence of flora and fauna only increase the fire spread capability and speed. Early detection of these forest fires can help control the spread and protect the nearby areas from the damage caused. This research paper aims at predicting the occurrence of forest fires using machine learning and deep learning techniques. The idea is to apply multiple algorithms to the data and perform comparative analysis to find the best-performing model. The best performance is obtained by the decision tree model for this work. It gave an accuracy of 79.6% and a recall score of 0.90. This model was then implemented on front-end WebUI using the flask and pickle modules in Python. The front-end Website returns the probability that a forest fire occurs for a set of inputs given by the user. This implementation is done using the PyCharm IDE.

Keywords Forests · Fire · Prediction · Machine learning · Deep learning · Flask implementation

1 Introduction

Forests have always been the synonym for abundance on our planet. As a source of food, vegetation, and wildlife, they have played a significant role in human lives ever

M. Shreya · R. Rai · S. Shukla (✉)
Department of Data Science, Christ University, Bangalore, India
e-mail: samiksha.shukla@christuniversity.in

M. Shreya
e-mail: m.shreya@science.christuniversity.in

R. Rai
e-mail: ruchika.rai@christuniversity.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_51

since humankind emerged. They control the soil and water cycles, thus impacting the agriculture levels in the nearby areas. Being home to more than half of the animal species, they keep the ecological diversity in check. These giant green land masses play a crucial role in our daily life, from the food we eat or the air we breathe to the climate we see around us.

Countries rely on their forests for the economy as well, as they provide a livelihood to millions of their people. The day-to-day production activities of goods like paper, wood, fruits, etc., even medicines and cosmetics, rely heavily on forests. Forests also improve the touristic value of a country. In places like Hawaii, Bali, Maldives, and most islands, the government relies heavily on tourism, reflected in their budget plans. Local forests provide a beautiful ecotourism experience to their customers.

Forest fires are many environmental hazards, controlling whom becomes complicated with every passing minute. The presence of vegetation acts as fuel to the fire, and every minute, the fire intensifies by multiple folds, thus increasing the water required to put it out. The previous century has marked numerous forest fire events. From the Black Friday Bushfire in Australia's Victoria State in January 1939, which burnt a massive 4.9 million acres, to the recent Californian wildfire in the USA, which burnt around 18,000 acres in 2020, wildfires have been causing havoc and resulting in human and wildlife fatalities. The increase in demand to detect and control these fires with a minimum delay has been growing ever since. Researchers and scientists have been trying to develop solutions to avoid or reduce the effects of these fires so that the damage can be minimized. While the fire control techniques and strategies are taken care of by the Forest Department of every country, there exists a lot of scope in coming up with novel ideas for faster forest fire detection.

Motivation: Machine learning (ML) and deep learning (DL) techniques have been used in every discipline over the past decade. It has provoked scholars to implement these computer techniques even in the forest fire domain. Research has been conducted to study the patterns in forest fires. Many papers have elaborated on the insights from data, used ML or DL models for various purposes, and proposed solutions for implementation. While a few papers suggested prototypes of new sensor devices, a few analysed the satellite images for a particular region, and others only used specific ML or DL techniques with already existing data.

Contribution: This paper aims to analyse and predict forest fires using machine learning and deep learning techniques. Two datasets containing meteorological and Fire Weather Index (FWI) parameters are collected. These datasets correspond to two different regions, namely Portugal and Algeria. To gain insight into the characteristics of the data, a thorough exploratory data analysis (EDA) is performed. The data is visualized using principal component analysis (PCA) and split into train and test sets. Nine machine learning models and one deep learning model are applied to the data, and results are obtained by comparing accuracy, precision, and recall metrics. The best model is obtained based on the metrics, and the same is implemented on a front-end web UI using flask. This WebUI takes the input of certain features from the user and generates an output that shows the probability of forest fire occurrence for the corresponding input conditions.

Organization: This paper is organized as follows. Section 2 provides a brief overview of related research work. The problem definition, research challenges, and dataset description are depicted in Sect. 3. Section 4 presents the methodology and module description. Module description explains pre-processing, model, and architecture of the model. The performance analysis and results are discussed in Sect. 5. Section 6 concludes the paper, and future work is discussed.

2 Related Work

One of the renowned works done in this area was a contribution of Cortez and Morais [1], who collected the dataset for forest fires found in the UCI machine learning repository. There were four different feature selection setups: spatial, temporal, FWI components, and weather attributes. Several data mining models like decision trees (DT), random forest (RF), multiple regression, and neural networks (NN) were used, among which SVM gave the best results with just four attributes such as temperature, wind speed, relative humidity, and rain. Their work was extended by Rishickesh et al. [2], who used bagging and boosting techniques with and without PCA. In their paper [2], logistic regression gave the best result with PCA, and without PCA, gradient boosting gave the best result.

Multiple papers have collected and used NASA's MODIS terra and aqua data for analysis. In contrast, Slobodan Milanovíc et al. [3] collected data for Serbia, and Ma et al. [4] collected data for six geographical regions in China. The random forest model gave the best accuracy for both papers. In Ref. [3], the model performance was assessed by ROC curves, while the other paper [4] concluded that climate and vegetation variables had a significant impact on forest fire occurrence. A new methodology for predicting forest fires in Greece based on the Canadian Fire Weather Index system components (FWI—fire weather index) was introduced by Varela et al. [5]. This paper used simple Geographic Information Systems (GIS) functionality. The methodology outcomes are provided as indicators for individual areas and as maps at a regional or national level. Identifying the advantages of both methods, Stojanova et al. [6] used a combination of NASA's MODIS, Aladin, and GIS data for three regions of Slovenia to predict forest fires. Various machine learning and ensemble methods from the WEKA data mining system were used, and it was seen that bagging of decision trees gave the best result.

Manual data collection was observed in a few research works like Pragati et al. [7] and Varela et al. [8]. The paper [7] used Wireless Sensor Networks (WSN), which took regular readings and sent them to the cloud for storage. In contrast, the paper [8] collected data for only two features, namely temperature and humidity. The data sent to the cloud was experimented with using decision tree and support vector machines. The decision tree gave the best results, whereas the other data used only the regression technique. If the result of the regression technique met some fire threshold conditions, then the ID, location, temperature, and humidity parameters were sent forward for fire control.

While the famous work by Cortez and Morais [1] used five different data mining techniques, Wijayanto et al. [9] used a data mining technique called adaptive neuro-fuzzy inference system (ANFIS). Data on wildfire hotspots in Central Kalimantan, Indonesia, were analysed, and this technique was used to classify fire alarms into true or false ones. The results showed low training and testing errors.

Some papers experimented with image analysis of the forest fire hotspots. Ananya et al. [10] implemented several ML models, among which random forest gave the best results. Then, the satellite image map space of the area was extracted for prediction and converted to an HSV model. If the number of hotspots in the image is more significant than a threshold limit, a fire was detected. However, Syarifudin et al. [11] used a 1D convolutional neural network to predict the number of hotspots for the early treatment of forest fires. They used variations in output, the learning rate, and the number of nodes applied on neural networks. The final results included the MAAPE for daily, monthly, and 12-month predictions, respectively.

A few papers and book chapters were studied to understand forests in general. A broad spectrum of information was gained around forests and their evolution through the paper written by Wodzicki [12], translated to English by Barbara Przybylska. The author elaborated on the evolution of forests on earth, starting from the formation of wood-like plant tissues to the current day forests. It was emphasized that forests ecosystems should be characterized by their entirety and not by specific features. Few other important points like the size of forest lands, effects on gases like nitrogen, soil fertilization, and impact on biodiversity were covered in the book chapter 'Tourism and Forest Ecosystem' by Gössling and Hickler [13]. They explained how tourism impacted the forest ecosystem, the various forest-based activities, and the interdependence of forests and climate changes.

The research work by Zong et al. [14] does not fit any models onto the data. Instead, it mainly focuses on estimating the forest fire seasons and patterns. The data were collected for five countries, namely Kazakhstan, Kyrgyzstan, Tajikistan, Uzbekistan, and Turkmenistan, and the fire seasons were analysed for different vegetation types like shrubs, grasslands, mountain forests, etc.

A literature review paper by Arif et al. [15] and a comparative study by Ahmad A. A. Alkhatib [16] were utilised to understand the current forest fire analysis and prediction trends. The paper [15] reviewed and summarized the recent trends in forest fire events prediction, detection, spread rate, and mapping of the burned areas. It stated the four significant research areas related to forest fires and emphasized that accurate prediction was a challenging task. On the other hand, the paper [16] summarized and elaborated the various methods and then pointed out the advantages and disadvantages of each.

While hybrid models were used by researchers like Shidik and Mustofa [17], an innovative approach was presented in the paper by Silvester et al. [18]. The hybrid model used a combination of clustering and classification approaches where several algorithms were combined and tested with fuzzy c-means (FCM). This model's best result was obtained using the combination of FCM with BPNN and classified into one of the three categories, namely no burn area, light burn, and heavy burn. However, in the paper [18], bird sounds were classified using deep learning (DL) algorithms

like convolutional neural network (CNN) into two classes, namely under normal conditions and threatened or panic conditions. ReLU activation function was used, and the results obtained showed that this classification surprisingly gave 96.45% accuracy.

3 Problem Definition, Challenges, and Dataset Description:

3.1 Problem Definition

The problem is to predict the occurrence of forest fires using machine learning and deep learning techniques. This paper aims at a thorough understanding of the dataset. Multiple supervised and ensemble techniques are used to analyse the data. The best model obtained is implemented on a front-end Web UI using the flask module.

3.2 Challenges

The analysis and prediction were performed on a combination of two datasets. These datasets correspond to two different geographic locations, namely Portugal and Algeria. Also, the feature set of one dataset is a subset of the other dataset features. Building the remaining features from the official document is a time-consuming task.

3.3 Dataset Description

The first dataset is called 'Algerian forest fires'. The second dataset is called 'forest fires'. Both the datasets are taken from the UCI machine learning repository, and they correspond to the forest fire data of Algeria and Portugal, respectively.

The final dataset had a total of 14 features, which were Temp, RH, Windspeed, Rain, FFMC, DMC, DC, ISI, BUI, FWI, and Classes (same description as in Table 1) along with month (month of the year 1–12), year (year of forest fire), and location (two classes, Algeria or Portugal). Since the data for BUI and FWI columns were missing in the Portugal dataset, they were imputed using the formulae from the official document [19]. The area and day features were dropped since imputing their missing values would not be meaningful.

Table 1 Features of both the datasets

Algerian data features	Portugal data features
Date—date of forest fire (dd/mm/yy)	*X*—*X*-axis coordinate (0–9)
Temp—Temperature in Celsius	*Y*—*Y*-axis coordinate (0–9)
RH—Relative humidity in %	Month—Month of the year (Jan. to Dec.)
Ws—Wind speed in km/h	Day—Day of the week (Mon. to Sun.)
Rain—Rainfall in mm/m^2	FFMC—Fine fuel moisture code
FFMC—Fine fuel moisture code	DMC—Duff moisture code
DMC—Duff moisture code	DC—Drought Code
DC—Drought code	ISI—Initial Spread Index
ISI—Initial spread index	Temp—Temperature in Celcius
BUI—Buildup index	RH—Relative Humidity in %
FWI—Fire weather index	Wind—Wind speed in km/h
Classes—two classes (fire and not fire)	Rain—Outside rain in mm/m^2
	Area—Burned area of forest

4 Methodology

Initially, domain knowledge is acquired by reading multiple papers and understanding the kind of work done. The proposed methodology is applied to the combined dataset with 14 columns. A total of nine machine learning algorithms and one simple neural network under deep learning are applied. The coding is done in the Jupyter Notebook using the Python programming language (Fig. 1).

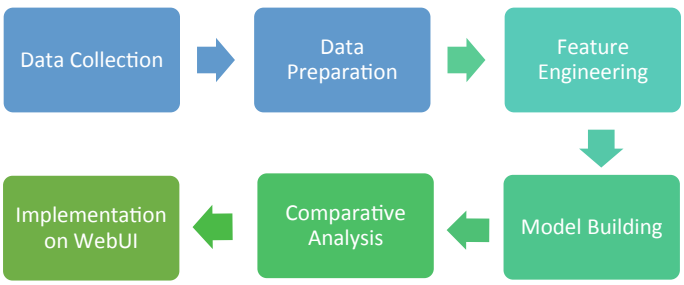


Fig. 1 Methodology

The following steps are followed in methodology:

Step 1: Data Aggregation: The data were collected from two different sources for this study. So, the first step was to aggregate it.

Step 2: Data Preparation: Since a few columns had no data in any one of the datasets, they were either imputed using the official document [19], or they were dropped if deemed not necessary, as mentioned in Sect. 3.3.

Step 3: Exploratory Data Analysis: This step was performed in three stages: univariate, bivariate, and multivariate. Unsupervised techniques like principal component analysis (PCA) and clustering were used under multivariate analysis. Each of them gave significant insights into the data at hand.

Step 4: Training and Test Set: In this step, the data are split into training and testing data in the ratio of 80–20. As a result, 80% of the data is used for training purposes, and the remaining 20% is used for testing the performance of the previously built model. The more training data is, the better it is to learn from it.

Step 5: Model Building: A total of ten models are trained and tested on the data. Among them, nine are machine learning models, which include ensemble techniques. They are support vector machine, decision tree, logistic regression, naïve Bayes, k-nearest neighbours, random forest, AdaBoost classifier, XGBoost classifier, and voting classifier. Under deep learning, a neural network is built, trained, and tested on the data, and the results are analysed.

Step 6: Model Comparison: All the models built in the previous step are compared based on various performance metrics like accuracy, precision, and recall. The best accuracy of 79.6% was obtained using the decision tree algorithm. It also gave a recall score of 0.90 for the 'fire' target class.

Step 7: Feature Identification: Since the decision tree model gave the best results, feature importance is drawn from the model. The top three critical features are selected for the front-end WebUI.

Step 8: WebUI implementation: PyCharm IDE is used for this implementation. Essentially, the WebUI takes input from the user and gives an output along with a safe or dangerous message (Fig. 2).

4.1 Module Description

4.1.1 Proposed Model

The dataset is split into features, targets, training, and testing sets. All the columns excluding the target are considered as features. The data are divided into an 80:20 ratio for training and testing purposes. It means that 80% of the data are used for training the model, and then, the other 20% is used to test the pre-trained model. All the ten algorithms mentioned in Step 5 of the methodology section are trained and tested on the data.

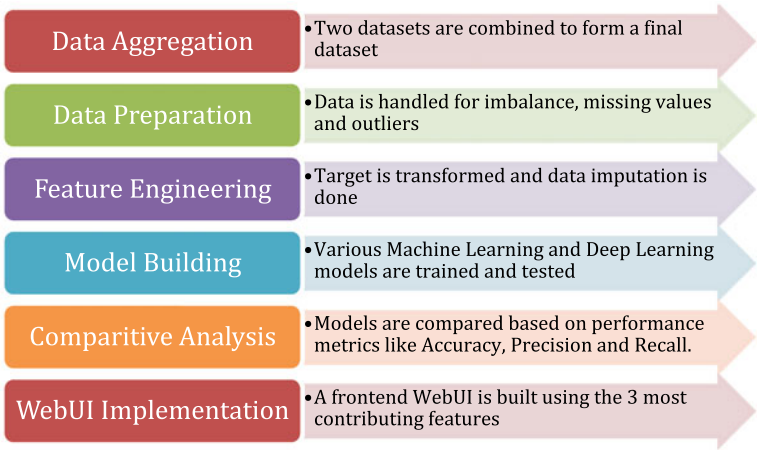


Fig. 2 Overview of workflow

It was seen that the decision tree gave the highest accuracy of 79.6%. Since accuracy can sometimes be biased, the model performances were also checked using the confusion matrix. A classification report was obtained for each model, clearly stating the precision and recall of the fire and non-fire classes. Since fire events are more important than non-fire ones, the precision and recall corresponding to the fire class are considered.

4.1.2 Feature Importance

Since the decision tree algorithm gives the best accuracy, this model is considered the best-performing model on the data. The feature importance is then obtained using Python's sklearn library. The top three features with the highest priority are noted.

Figure 3 clearly shows that features corresponding to 'FFMC', 'windspeed', and 'temperature' have the highest importance, respectively. They will be used for the front-end WebUI implementation.

4.1.3 WebUI Implementation Using Flask

A front-end web page is created where the user inputs three features: FFMC, wind speed, and temperature. These values are captured as a form and sent as a POST or GET request to the back-end. In the back-end, the pickle library is used for storing the models. The decision tree is trained once on the data, and the trained model is loaded into the pickle file. The pickle file is called in the back-end whenever a new input is received from the front-end. Since the model is already trained, the pickle file generates an output for the new data points, and this output is displayed back to

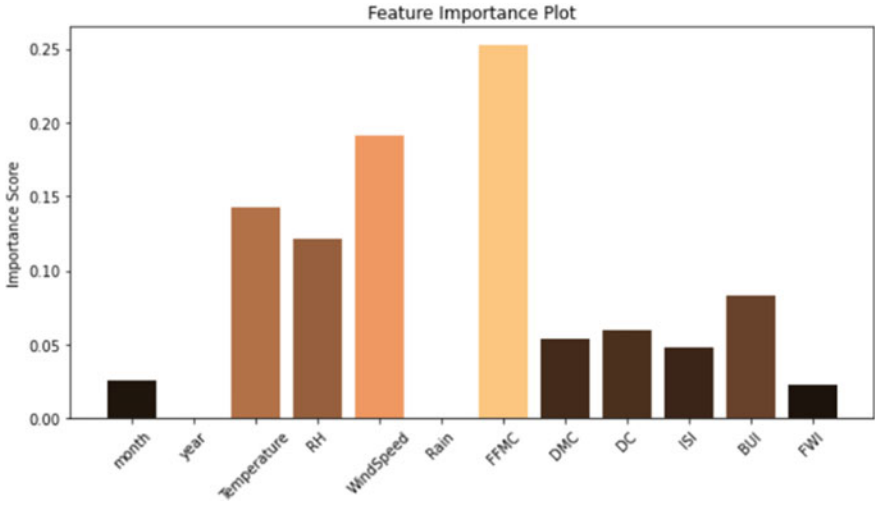


Fig. 3 Feature importance

the user on the front-end. The probability of a forest fire occurring or non-occurring is printed on the front-end along with danger or safe message.

5 Experimental Results

Table 2 shows the results obtained from all the ten models, including the accuracy, precision, recall, f1-score, and AUC-ROC score of each model. The decision tree

Table 2 Model performances

Model name	Accuracy	Precision	Recall	F1-score	AUC-ROC score
Support vector machine	73.6	0.68	0.90	0.80	0.71
Decision tree	79.6	0.73	0.90	0.80	0.75
Logistic regression	69.7	0.67	0.88	0.76	0.68
Naive Bayes	67.1	0.63	0.93	0.75	0.65
K-nearest neighbours	68.4	0.72	0.68	0.70	0.68
Random forest	73.6	0.73	0.80	0.77	0.73
AdaBoost classifier	69.7	0.70	0.78	0.74	0.69
XgBoost classifier	73.6	0.73	0.80	0.77	0.73
Voting classifier	73.6	0.74	0.85	0.80	0.75
Artificial neural network	61.8	–	–	–	–

```
print(confusion_matrix(y_test,y_preds_dt))

[[21 14]
 [ 4 37]]
```

Fig. 4 Confusion matrix of the decision tree model

gives the best accuracy. Also, its recall score is 0.90, which means that the model can correctly identify a fire occurrence 90% of the time.

The confusion matrix gives an insight into how many data points were classified correctly and how many of them were misclassified. The values along the diagonal are correct predictions, and every other value is where the model is misclassified (Fig. 4).

The rows in this confusion matrix indicate the actual classes, whereas the columns indicate the predicted classes. So, there are four values printed, which are true negatives (top left value), false negatives (top right value), false positives (bottom left value), and true positives (bottom right value).

Various metrics like true positive rate, false positive rate, accuracy, precision, recall, sensitivity, specificity, and many more are used to assess the model's performance. These are calculated from the values in the confusion matrix.

Figure 5 shows a sample WebUI input that gives an output probability of forest fire occurrence as 0.87 along with a message that the forest is in danger. If the likelihood of forest fire occurrence is too low, its probability, along with the message 'Your Forest is Safe', is printed instead.

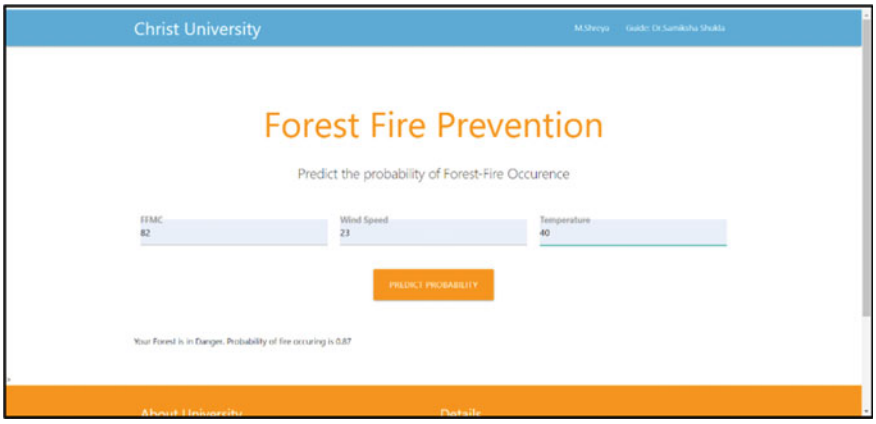


Fig. 5 Front-end WebUI

6 Conclusion and Future Work

This study presents the forest fire-related challenges and prediction model using machine learning and deep learning techniques. Multiple research papers have been examined to understand their objectives, analyse the technique utilised, and identify any gaps in them.

Based on the data collected from various resources, the data cleaning process is applied. After addressing the missing values, various visualization approaches are used to gain insights. Specifically, univariate, bivariate, and multivariate analysis techniques were applied to gain insights. After this, several ML and DL models were applied and tested for their performance. The decision tree algorithm obtained the best results with an accuracy of 79.6%, and a recall score of 0.90 for the fire class. The model was implemented on a front-end WebUI using the flask and pickle Python modules. This WebUI gives the probability of a forest fire for a given set of input parameters.

In the future, the work can be extended by collecting more data from various geographical locations around the globe for further analysis. Using more data would give more precise results for the machine learning model as it would have more data to learn from. Also, a few other deep learning techniques can be applied.

References

1. Cortez P, Morais A (2007) A data mining approach to predict forest fires using meteorological data
2. Rishickesh R, Shahina A, Nayeemulla Khan A (2019) Predicting forest fires using supervised and ensemble machine learning algorithms. Int J Recent Technol Eng (IJRTE) 08(02)
3. Milanovíc S, Markovíc N, Pamucar D, Gigovíc L, Kostíc P, Milanovíc SD (2021) Forest fire probability mapping in eastern serbia: logistic regression versus random forest method. In: MDPI Forests
4. Ma W, Feng Z, Cheng Z, Chen S, Wang F (2020) Identifying forest fire driving factors and related impacts in China using random forest algorithm. In: MDPI Forests
5. Varela V, Vlachogiannis D, Sfetsos A, Politi N, Karozis S (2020) Methodology for the study of near-future changes of fire weather patterns with emphasis on archaeological and protected touristic areas in Greece. In: MDPI Forests
6. Stojanova D, Panov P, Kobler A, Džeroski S, Taškova K (2020) Learning to predict forest fires with data mining techniques
7. Pragati, Shambhuwani S, Umbrajkar P (2019–2020) Forest fire detection using machine learning. Int J Adv Sci Res 04(12)
8. Varela N, Martinez D, Jorge L, Ospino A, Zelaya NAL (2020) Wireless sensor network for forest fire detection. In: The 15th international conference on future networks and communications (FNC), Aug 9–12, 2020
9. Wijayanto AK, Sani O, Kartika ND, Herdiyeni Y (2017) Classification model for forest fire hotspot occurrences prediction using ANFIS algorithm. In: IOP Conference Series: Earth and Environmental Science
10. Shahdeo A, Shahdeo A, Reddy PS, Chaitra K (2020) Wildfire prediction and detection using random forest and different colour models. Int Res J Eng Technol 07(06)

11. Syarifudin MA, Novitasari DCR, Marpaung F, Wahyudi N, Hapsari DP, Supriyati E, Farida Y, Amin FM, Nugraheni RD, Nariswari R, Setiawan F (2020) Hotspot prediction using 1D convolutional neural network. In: 5th International conference on computer science and computational intelligence, procedia computer science, pp 845–853
12. Wodzicki TJ (2020) Forest—the photosphere of life in the Earth's atmosphere (Translated by: Barbara Przybylska). Leśne Prace Badawcze, Forest Research Papers, vol 81, pp 133–138
13. Gössling S, Hickler T (2020) Tourism and global environmental change—chapter X: tourism and forest ecosystems. In: Gössling S, Hall MC (eds) Tourism and global environmental change
14. Zong X, Tian X, Yin Y (2020) Impacts of climate change on wildfires in central Asia. MDPI Forests
15. Arif M, Alghamdi KK, Sahel SA, Alosaimi SO, Alsahaft ME, Alharthi MA, Arif M (2021) Role of machine learning algorithms in forest fire management: a literature review. J Robot Autom 5(1):212–226
16. Alkhatib AAA (2014) A review on forest fire detection techniques. Int J Distrib Sens Netw 2014:12
17. Shidik GF, Mustofa K (2014) Predicting size of forest fire using hybrid model. In: International federation for information processing (IFIP), ICT-EurAsia, pp 316–327
18. Permana SDH, Saputra G, Arifitama B, Yaddarabullah, Caesarendra W, Rahim R (2021) Classification of bird sounds as an early warning method of forest fires using Convolutional Neural Network (CNN) algorithm. J King Saud Univ Comput Inf Sci
19. Van Wagner CE (1985) Equations and FORTRAN program for the Canadian forest fire weather index system. In: Pickett TL (ed) Canadian forestry service, Petawawa National Forestry Institute, Chalk River, Ontario, Forestry Technical Report

Wearable Thong Entrenched with SoS Facility and Mobile Ad Hoc Network

S. Kiruthiga, D. Umanandhini, S. Sridevi, and N. Beulah Jabaseeli

Abstract To ensure peace and safety of the country, military forces are inevitable for the country's protection. India has the second largest military force in the worlds with the strength of 1.4 million active troops. The army people lose their valuable lives due to lack of technical infrastructure to provide status about the injured soldiers. Many lose their lives and become disabilities because of injuries rather than battlefield. Nowadays, there are vast amount of electronic gadgets to fight against the enemies and secure the nation in the border. But in some adverse conditions like bad weather or due to remote location of soldiers, it becomes impossible to interact with them and monitor their health conditions of militants. In this work, we are developing a modern wearable smart thong entrenched with system of system facility and mobile ad hoc network. It has the special feature of continuously monitoring the health condition of troops. It also continuously records the health parameters and environmental parameters of the military warriors so that valuable lives of soldiers which is significant asset to our country can be rescued at right time.

Keywords Smart thong · System of system facility · Mobile ad hoc network · LoRaWAN

S. Kiruthiga (✉) · D. Umanandhini · S. Sridevi · N. Beulah Jabaseeli
Department of Computer Science and Engineering, School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India
e-mail: skiruthiga@veltech.edu.in

D. Umanandhini
e-mail: drumanandhini@veltech.edu.in

S. Sridevi
e-mail: drsridevis@veltech.edu.in

N. Beulah Jabaseeli
e-mail: nbeulahjabaseeli@veltech.edu.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_52

1 Introduction

India has the second largest military force in the world. Soldiers lose their life, due to unavailability of life saving resources and innovative ideas and minimal trials, support from medical personnel. Our work helps in minimizing these travails and improves the emergency savior of the armed forces personnel these includes assessment of soldiers in the aspects of vital parameters and minimize the dependence to medical faculties. It acts as an appropriate tool for measurement of health condition of soldiers. This will also help in evaluating the progress and deterioration in the health of the warriors whether the development of new assistive technology will be favorable or not.

Digital devices and interactive objects that can sense and control physical devices can be built on Arduino which is an open source and in our work we use Arduino Uno, ATmega328 microcontroller in which has special features like being more powerful, easy to code, compatible with other devices and available cheap in market. MANET (LoRaWAN module) [1]—Lora technology is designed for long-range low-power communication; it mainly targets M2M and IOT networks, enables public or multi-tenant to connect a number of application in the same network. The GPS module provides information's like current time, date, latitude and longitude coordinates of the soldiers to the needed microcontroller. It requires no external components except power supply. It is built up with internal RTC back-up battery. It can be directly connected to microcontroller's UART. It includes vital parameters like heart pulse rate, temperature of the body. It is useful in diagnosis of health of militants in military hospitals, before sending them to foreign missions and in monitoring their health conditions from the battle field.

2 Related Work

The paper by Niket Patel et al. [2] reports health-based tracking and monitoring system for soldiers. Their work proposes using Internet of Things (IoT), which can be integrated on the soldier's body to track their vital health parameters status and current location using Global Positioning System, it is used by mounting on soldiers accessories. Using Internet of Things, the collected from soldier will be transmitted to the base station. Their work comprises tiny wearable functional equipment's like biomedical sensors and transmission modules. Therefore, by using the proposed work, it is feasible to implement a mechanism of low cost to safeguard the valuable human life on the battlefield by examining the oxygen level available.

The work done by Patrik Kutilek et al. [3] proposes the purpose of wearable life-guarding systems in military. It describes the various types of current wearable sensors. It does the quantitative evaluation of the soldiers. They use a telemetry monitoring system with Flexi guard. It presents the State of development of Flexi guard in the Czech Republic Army.

Thanga Dharsni et al. [4] propose a method of building a wearable device using distributed computing. The control room acts as an active hub, which monitors the health parameters of the soldiers along with the location. GPS is used in guiding them in directions when they are in remote areas. GSM module is used for information transformation; using the data received from soldiers authorities in control room may take action immediately.

Zeeshan Raz et al. [5] addresses the challenges faced Pakistan soldiers and provides solution to it, by designing a light weight wearable device. They also design a device which consumes low power, because of it even in remote village areas; it can withstand charge for a long period.

Chao Lv et al. [6] does a research individual soldier energy-efficient system. They suggest a feasible system for evaluating the energy of soldiers. EEG (Electro Encephalogram) is collected, and wavelet analysis is performed on it. Various steps like feature extraction, labeling, filtering and cluster analysis were done with the data. It provides theoretical basis approach for promoting the nervous system and brain development.

Jitesh Pabla et al. [7] combines IoT with blockchain technology to monitor the soldiers health parameters and transfer them securely using body sensor network (BSN). They provide the data tamper proof and distributed it by encrypting it using blockchain technology.

Patient preference management system through IoT using Computer Vision was developed by Sathesh [8] to serve in unusual condition like viral explosion, war and natural disaster, whereas traditional computer-based systems will make the patient wait for long time. Here they propose a new IoT-based Computer Vision system to send their status through email or phone messages. Various AI-based algorithms like FRR, MAS, AWS and ODA are applied to analyze the efficiency. They yielded good result with AWS when compared to other algorithms. AWS also minimizes the power consumption and execution time for data movement.

Tatugade et al. [9] indicate the demand of MANET-type communication and its features like communication network and information sharing, but it has limited power resource. So their work deals with energy node limitations and suggests approaches that prolong the battery life in WSN that are clustered hierarchically. It also does a comparison between homogeneity and heterogeneity of nodes using Protocols like Low Energy Adaptive Clustering Hierarchy (LEACH) and Stable Election Protocol (SEP). They proposed a protocol called Advanced SEP to enhance the lifetime of wireless sensors in MANET. Since the entire network infrastructure is moving in MANET, terrain interference point-to-point nodes may be dropped beyond the coverage, because of this the stability of network is continually stressed which needs a betterment of their ASEP protocol with mobility concern.

Santhanamari et al. [1] propose health monitoring system for soldiers using secured data transmission through sensors; they focused on designing health monitoring system over MANET, with an efficient nRF24L01 routing protocol, by comparing with OMNET++. The body parameters are sensed using bio-sensors and transmitted to control room via offline thus ensuring secure communication, alert the control room when an emergency happens SoS (Signal on Ship) is used via cloud

service. They left a future scope to use LoRaWAN technique to support long-range and fast communication even in terrain regions.

Taleb et al. [10] indicate the need of WBAN in future; it gives the survey of all wireless technologies and bio-sensors available in market which can be used in medical field. It consists of miniaturized sensors which are used to collect and transmit data through wireless network. The use of other emerging technologies like LPWAN (Low Power Wide Area Network) was also identified and surveyed. The application of IoT with WBAN is developed and discussed in their work.

Archana Padikar et al. [11] dealt with an idea of keeping track of soldier parameters. Arduino UNO and Node MCU are used in soldier monitoring system. They have integrated traditional IoT with GPS. This was useful to warfare in situations like when extreme poisonous gas is present, in life-threatening situations, and in critical psychological state. They proposed a future scope of involving DC power source to harvest solar energy directly and place a camera at the soldier end which enables the doctors and rescue team to view the situation of soldiers in real-time.

Mahammad Eliyaz et al. [12] designed a soldier monitoring system with MyRio which is an embedded evaluation board, and Lab View PC which enables soldiers to send messages via GUI. GSM 900IDRS232 is used to monitor physical parameters of the soldier. It helps commanders at base station to sketch the war strategies. By making proper GUI at base station PC, the base station can be developed was the future scope proposed by them.

Aashay Gondalia et al. [13] described the ability to track the location and monitor health of soldiers using GPS module and wireless body area sensor networks (WBASN). ZigBee module is adopted to transmit data from sensors. LoRaWAN network infrastructure is used between Squadron leader and control unit in high-altitude areas where data transmission is almost impossible due to low network coverage area. K-means clustering algorithm is used for data analysis which is collected in cloud.

3 Overview of the Proposed Approach

In our work, we proposed the new method of smart thong for monitoring the body parameters of soldiers with MANET-type communication system. It acts as a life-guard to the army personnel of all over the globe, provides accurate location of missing soldier in critical conditions and overcome the drawback of soldiers missing in action (using GPS), helps to improve the communication between soldier to soldier and soldier to base station in emergency situation and provide proper navigation to the control room. MANET (LoRaWAN) is used for long-range low-power communication, and it requires only minimal infrastructure so it is compatible to use even in terrain regions. The Arduino Uno at receiver side provides IDE and has many facilities to connect with the computer or another Arduino board. LCD used displays the sensor readings. At the receiver side, Arduino ATmega microcontroller is used.

The sensor is interfaced with thong which are continuously fetch the physical parameters of the soldier such as heartbeat, respiratory rate and body temperature along with environmental temperature. The heartbeat sensor gives the digital output of heart rate in beats per minute (BPM). Respiratory sensors give the inhale and exhale rates. The body temperature is monitored using LM35 which measures the body temperature proportional to electric circuit. These vital parameters are continuously updated in monitoring along with location. Location can be updated via GPS module. Here we use third-generation GPS module which gives exact date, time, longitude and latitude. Every status will be displayed in receiver section LCD and PC. And also, we are using a strap like control in soldier body with the help of SoS facility it will helps us to locate an accurate location of the soldier when the soldier presses the strap with the help of the GPS it will predict the accurate location of the soldier and transfers the communication with sensors like Arduino MEGA to Arduino UNO with the module LoRaWAN technology. This was the first idea where we have integrated this with soldier's body during the war field.

3.1 General Architecture

In the below diagram Fig. 1, two modules have been shown which are as follows:

(i) Transmitter section (soldier module)
(ii) Receiver section (control room section).

The "Transmitter Section" is being carried by the soldier in the war field. The "Receiver Section," i.e., the control room receives the required signals received by the soldier, analyzes the signals and sends required help and information back to the soldier for their rescue purpose.

3.2 Implementation

To predict soldier's health and safety, following aspects have to be considered.

A. Estimation of vital body parameters
B. Immediate help response (SoS)
C. Data preprocessing
D. Result analysis
E. Instructions generation

A. **Estimation of vital body parameters**

Vital body parameters like heart beat rate, respiratory rate, body temperature and environmental temperature are estimated using the respective sensors and required data are generated.

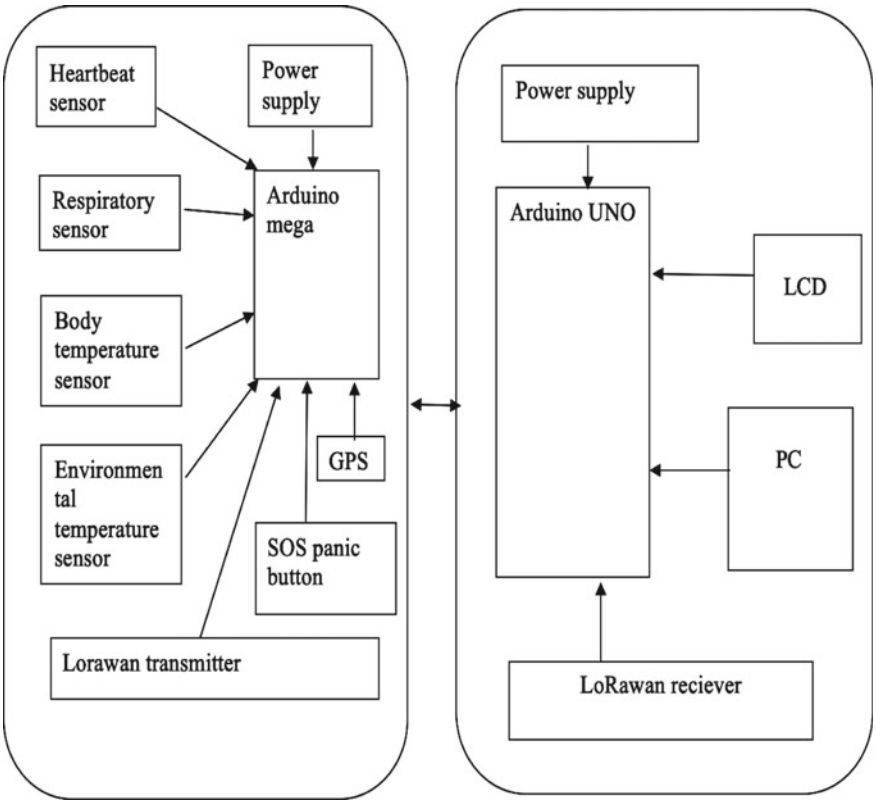


Fig. 1 General architecture of smart thong

B. **Immediate help response (SoS)**

Immediate help response in case of emergency is generated by means of panic button pressed by the soldier (SoS) which sends emergency signal to the control room thus helping them for providing help to soldiers in need.

C. **Data Preprocessing**

Data acquired by means of sensors and panic button is analyzed and preprocessed using various techniques like regression, clustering, binning, etc. and are thereafter refined and collected for further analysis.

D. **Result analysis**

Final accumulated data is analyzed using serial monitor and further monitored within their respective normal range so as to ensure safety of the soldier.

E. **Instructions generation**

Based on the results analysis done so far, required instructions are generated by the control room in order to provide immediate help to soldiers thereby keeping them alive and safe.

4 Monitoring of Soldier Data

Table 1 illustrates the values acquired by applying the MSTS algorithm, gleaned from the R value. MSTS-

algorithm(x-temp, y-heart beat rate)

This algorithm designs a correlation formula for two different variable like temperature in (“C) and heart beat rate.

which is measured in (BPM).

1. Initialize: $X =$ Temperature $Y =$ Heart beat rate
2. Calculate the mean of each variable in sample set then subtract the mean from original values

$$x_1 - \overline{X}, x_2 - \overline{X}, x_3 - \overline{X} \ldots x_n - \overline{X}$$
$$y_1 - \overline{Y}, y_2 - \overline{Y}, y_3 - \overline{Y} \ldots y_n - \overline{Y}$$

3. Calculate the square of subtracted values as in above equation

$$\left(x_1 - \overline{X}\right)^2, \left(x_2 - \overline{X}\right)^2, \left(x_3 - \overline{X}\right)^2 \ldots \left(x_n - \overline{X}\right)^2$$
$$\left(y_1 - \overline{Y}\right)^2, \left(y_2 - \overline{Y}\right)^2, \left(y_3 - \overline{Y}\right)^2 \ldots \left(y_n - \overline{Y}\right)^2$$

4. Multiply the subtracted mean of each variable

$$\begin{gathered}\left(x_1 - \overline{X}\right).\left(y_1 - \overline{Y}\right) \\ \left(x_2 - \overline{X}\right).\left(y_2 - \overline{Y}\right) \\ \vdots \\ \left(x_n - \overline{X}\right).\left(y_n - \overline{Y}\right)\end{gathered}$$

5. Now finally the value of R is find by putting all the determined value in formula as shown below.

$$R = \frac{\sum_{i=1}^{i=n} \left(x_1 - \overline{X}\right), \left(y_1 - Y\right)}{\sum_{i=1}^{i=n} \left(x_1 - X\right)^2, \left(y_1 - Y\right)^2}$$

Table 1 Range of soldier condition

$(y - \overline{Y})^2$	$(x - \overline{X}).(y - \overline{Y})$	Value of R	Expected condition
113.71–450.38	117.01–450.03	0.11–0.33	Normal
1.250–1.493	0.149–12.593	0.44–0.55	Moderate
7.049	7.705	0.66	In harm
35.60	35.26	0.77	Death may occur
828.16–1503.37	82.816–150.372	0.88–0.99	Death

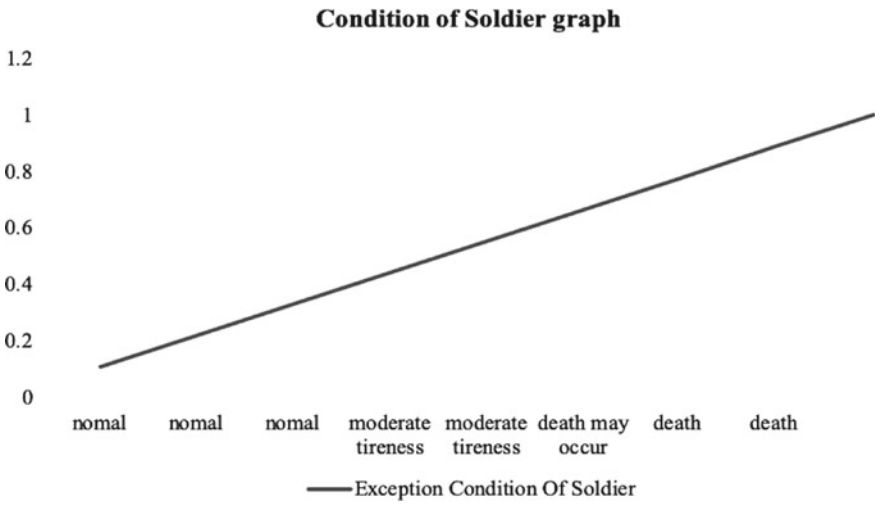


Fig. 2 Graph showing health condition of soldier

6. Write the value of *R*

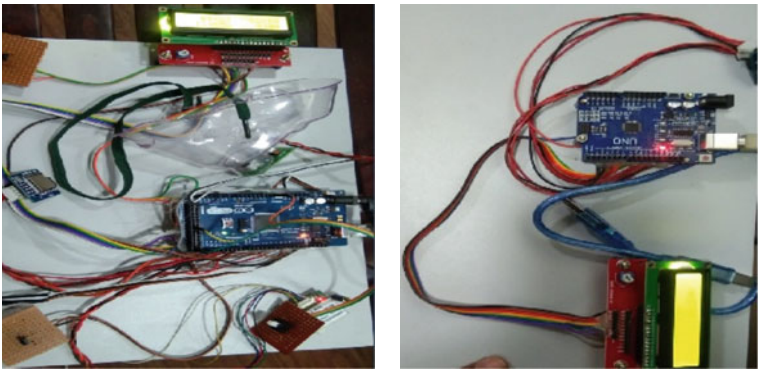
The health condition of warriors is predicted even if they are in remote areas. The conditions like injury, serious injury or death condition is analyzed based on the table shown below, and using the values received like as normal, moderate situation or the injury is serious leading to death. Based on this the rescue team will be sent from the control room to the soldiers location based on GPS. The warriors can send secret codes in encrypted form to the control room, here LoRaWAN module is adopted to ease the communication process between soldiers and control station.

5 Analyzing of Solider Health from Graph

The below graph Fig. 2 is generated based on the values received from soldiers by applying MSTS algorithm. Along the y-axis, range of values obtained and derived using MSTS algorithm is shown, based on which the degree of injury is determined and plotted along x-axis like normal, moderate, probability of death.

6 Results

The sensor readings are taken from the LCD as shown in Fig. 3a, b and c and transmitted to the control room. Based on MSTS algorithm, the corresponding values of "*R*" are determined which denotes the expected health conditions of soldiers



```
LOCATION28.545799,77.170303
HEALTH_PARAMETERhb=0 res=10 bt=39 et=14
LOCATION28.545799,77.170303
HEALTH_PARAMETERhb=0 res=10 bt=38 et=14
LOCATION28.545799,77.170303
HEALTH_PARAMETERhb=0 res=10 bt=39 et=14
LOCATION28.545799,77.170303
HEALTH_PARAMETERhb=0 res=11 bt=39 et=14
LOCATION28.545799,77.170303
HEALTH_PARAMETERhb=0 res=9 bt=39 et=14
LOCATION28.545799,77.170303
HEALTH_PARAMETERhb=0 res=10 bt=39 et=14
LOCATION28.545799,77.170303
HEALTH_PARAMETERhb=0 res=10 bt=39 et=14
LOCATION28.545799,77.170303
```

Fig. 3 **a** Transmitter section. **b** Receiver section. **c** Serial display of sensor readings

(normal, moderate tiredness, death may occur, death). The graph above in Fig. 2 analyzes the health condition of soldier. *X*-axis will be the stages of soldier and *Y*-axis be the range. Based on the corresponding values of "*R*" so obtained (as mentioned in table), the graph so obtained as a straight line denoting the health and injury level of soldiers. Based on these values emergency, medical assistance can be provided to soldiers and their life can be saved.

7 Conclusion

Military systems use very small number of devices. There is wide range of sensors available nowadays, using which we have designed a smart wearable thong which is light weight and consumes low power so that even in remote areas it can be used and charged up easily. Here we use LoRaWAN module and MANET-type communication system to help in emergency situations. The communication infrastructure can be established without any pre-existing infrastructure. It helps in saving the life of warriors in emergency situations. Encryption techniques are also adopted to send

the messages as secret codes from soldiers to the control room and vice versa; this ensures the authenticity of the usage of the system.

8 Future Enhancements

By using the above wearable device, soldiers condition can be identified during a mission. Immediate help is given to soldiers by adopting system of system. The proposed device can be advanced by adopting device to device communication, without the utilization of control room by integrating IoT technologies, which will accomplish it flexible for collecting the information from warfare. The alternative form of this system can moreover be used for sewage gas monitoring, advanced technologies like blockchain technology can be adopted to ensure the incorruptibility of data being transferred from sensors.

References

1. Santhanamari G, Premi J, Rahavi S, Vijay S (2020) Health monitoring of soldiers using efficient manet protocol. In: IEEE recent advances in intelligent computational systems (RAICS), pp 57–62
2. Patil N, Iyer B (2017) Health monitoring and tracking system for soldiers using internet of things (IoT). In: International conference on computing, communication and automation ICCCA2017
3. Kutilek P, PetrVolf (2017) Wearable systems for monitoring the health condition of soldiers review and application. In: International conference on military technologies (ICMT),Brno, Czech Republic
4. Zakir H, Naik P, Raghu M (2018) Soldier security and health monitoring. In: International conference on design innovations for 3Cs compute communicate control
5. Raza Z, Ashraf S, Liaquat K, Kashif H (2016) Monitoring of soldier's health and transmission of secret code. In: Sixth international conference on innovative computing technology
6. Chao Lv, Deng T (2018) Research on individual soldier energy efficiency assessment systems. In: 10th international conference on intelligent human-machine systems and cybernetics
7. Sharma V, Krishnamurthi R (2019) Developing a secure soldier monitoring system using internet of things and blockchain. In: International conference on signal processing and communication
8. Sathesh A (2020) Computer vision on IOT based patient preference management system. J Trends Comput Sci Smart Technol 2:68–77
9. Tatugade A, Jadhav SR (2019) Tactical communication system using MANET and WSN. IOSR J Comput Eng
10. Taleb H, Nasser A, Andrieux G (2021) Wireless technologies, medical applications and future challenges in WBAN: a survey. Wirel Netw 27:5271–5295
11. Archana Padikar A, Cinmayee CK, Chaithra E, Chethan, Puneeth Kumar DN (2020) Health monitoring and soldier tracking system using IOT. Int J Eng Res Technol (IJERT). ISSN: 2278-0181
12. Eliyaz M, Leela Venkata Sai Prudvi M, Pragnya Reddy G, Pavan M (2020) Soldier tracking and health monitoring system using LabView. Int J Emerg Trends Eng Res 8. ISSN 2347-3983
13. Gondalia A, Dixit D, Parashar S, Raghava V, Sengupta A, Sarobin VR (2018) IoT-based healthcare monitoring system for war soldiers using machine learning. Procedia Comput Sci 133

Blockchain and Its Integration in IoT

Manish Bharti, J. Sandeep, and C. Smera

Abstract IoT devices have become an integral part of our lives. The world has witnessed an exponential growth in the number of IoT devices. Managing these devices and the data generated by them has become very crucial. Data security and users' privacy are becoming more difficult as the number of devices grows. Blockchain, the technology behind Bitcoin, is known for data security and managing and efficiently maintaining huge amounts of data. Blockchain stores data in a chronological manner and in an immutable way. Integration of blockchain in IoT infrastructure has many advantages. This paper discusses various applications and challenges in blockchain. It highlights the adoption of blockchain in IoT infrastructure and reviews recent papers in this field.

Keywords Blockchain · Bitcoins · IoT · Smart contracts · Scalability · Supply chain

1 Introduction

The emergence of the Bitcoin cryptocurrency has given recognition to blockchain technology. Blockchain technology is a mechanism of storing transactions using mathematical cryptography in such a way that once a block is created, it cannot be altered or modified. Blockchain can be seen as a distributed ledger that records all the transactions between the sender and receiver efficiently. The transactions can be verified at any time with the stored permanent records [1]. The immutability and decentralized nature of blockchain prompt to facilitate various applications of

M. Bharti (✉) · J. Sandeep · C. Smera
Department of Computer Science, CHRIST (Deemed to Be University), Bangalore, India
e-mail: manish.bharti@mca.christuniversity.in

J. Sandeep
e-mail: sandeep.j@christuniversity.in

C. Smera
e-mail: smera.c@res.christuniversity.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_53

blockchain technology, which include the well-known digital currency Bitcoin [2], smart trading [3] and insurance [4].

Other than monetary usage, blockchain can be used in the supply chain management system [5], Internet of things [5], crowdfunding [4], and reliable data storage for various important and credential data like property exchanges data. Furthermore, in voting [6], healthcare [7], government, and many other fields, this blockchain is being implemented. Bitcoin had value not more than a penny in the starting days of its invention. Under a decade, each Bitcoin is carrying a weight of more than USD 20,000. Blockchain technology skills are becoming one of the most demanding technological skills. Many industries around the world are looking at blockchain because of its various features like immutability and decentralization. But there are a few aspects that must be considered while looking at blockchain. The first of them is the scalability problem.

With the increasing transaction volume, the blockchain is becoming heavier each passing day. In terms of size, the Bitcoin blockchain has exceeded 100 GB of storage [8]. And due to the increase in size, the throughput of transaction processing decreases rapidly. The existing transactions-per-second is also significantly less compared to mainstream transaction processing. These concerns require major research to overcome the challenges of filling the gap and making blockchain acceptable to the public domain and businesses.

While blockchain has many applications and benefits in different areas, it comes with many issues and challenges. These not only include security-related issues but there are many financial and social challenges also, which work as obstacles to the development and widespread implementation of blockchain technology. Firstly, scalability [9] is a major concern for the blockchain industry. Here, scalability is mainly considered as transaction per second (TSP). Only 7 transactions/sec can be performed in a blockchain like Bitcoin in the current scenario. Comparatively on average 2000 transactions/sec are completed in mainstream transaction processing like Visa, Mastercards [10]. So there is a serious gap in terms of scalability. Secondly, although blockchain is very safe for the user as the user makes the transaction using the system generated address instead of using real identity, there has been a concern over privacy leakage [8] over the blockchain. Moreover, it has been proved that miners can attain greater earnings by using a selfish mining strategy than their actual share [8]. Other than these major issues, energy consumption, lack of talent, and lack of standards are vital challenges for blockchain.

2 Background

Blockchain, the technology behind Bitcoin, was first introduced when the same Bitcoin was introduced [11]. In 2008, Satoshi Nakamoto published a well-structured and designed paper that introduced a digital currency application using cryptography and opened a distributed ledger [3]. Before that in 1991, Stuart Haber and W. Scott Stornetta introduced a “cryptographically secured chain of blocks” where they tried

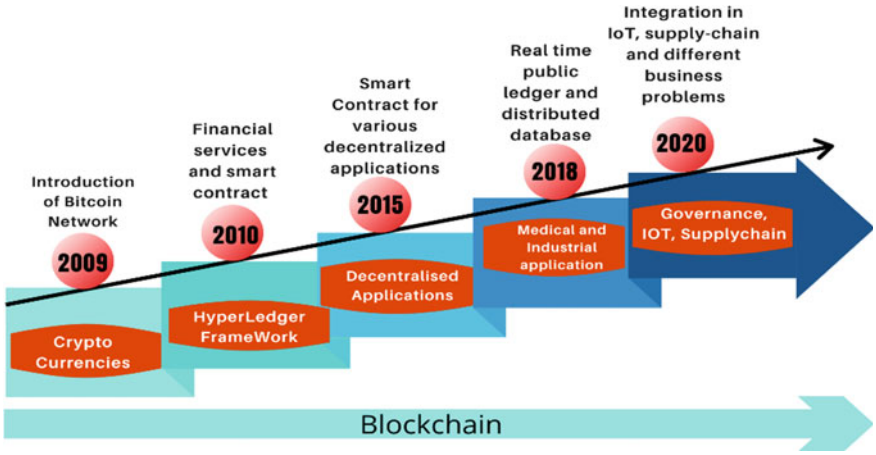


Fig. 1 Generation of blockchain

developing a way in which using the timestamp of a document prevents it from being altered [12].

As shown in Fig. 1, at first, Blockchain was all about digital currency, but later in 2014, when Ethereum was introduced by Vitalik Buterin, Gavin Wood, it was separated from currency and many applications of blockchain were introduced [13]. After the realization of the various applications of blockchain, which promised speed, security, and ease of use, people started investing their energy, money, and time in the research of blockchain. In 2015, smart contracts were introduced for decentralized applications.

By 2018, many medical and industrial use cases came into the picture. Currently, blockchain is being used in solving more complex problems like in IoT, in the supply chain, and in governance. Bitcoin is still the most widely used application of blockchain. The size of the Bitcoin blockchain was 20 GB by 2014, which surpassed 30 GB between 2014 and 2015. Soon, the network had the size 50GB-100 GB in January 2017 [14].

2.1 Recent Research on Blockchain Technologies

This section presents the recent research on blockchain technology. The popularity of blockchain in recent research trends has grown exponentially worldwide. Research conducted by Min Xu and the team has analyzed the literature of blockchain and claimed clustering analysis in blockchain can be leveraged by businesses in order to gain advantages over their competitors [15]. Some of the recent research contributions of the research community have been discussed below to highlight the scope of research in the blockchain.

Sabounchi et al. have addressed the critical challenges of data sharing "the trade-off between privacy and data utility". Using various technologies like TestRPC Ethereum, Web3.js, etc., the performance of the proposed blockchain-based peer-to-peer data trading mechanism is evaluated [16].

Ribeiro, Victor, et al. have studied various key management mechanisms and proposed LoRaWAN environments, a new key management mechanism, based on smart contracts and permissioned blockchain to boost overall security and availability of the system. This system utilizes the decentralized design of blockchain to resolve the central point of failure [17].

In another research work, Nadiya et al. came up with a research proposal using the deflate compression algorithm. The work proposed a way of making summary blocks and compression for the summary blocks which will reduce data volume up to 90% in order to make it easier to verify transactions. This approach also lightens the workload for nodes with limited processing power and memory [18].

Dargaye et al. have proposed "Pluralize" smart contracts that enable formal features in order to gain trust in smart contracts from their specification to their execution on the blockchain. It enables formal reasoning and the establishment of agent accountability for blockchain-based transactions [19].

It can be observed that the major research interests were on secure transactions to gain more trust. The scope was not limited to security but also the key management, data volume, etc.

3 Blockchain and Its Application

A detailed explanation of blockchain real-time applications such as energy, supply chain and logistics, healthcare, digital content distribution, agriculture, manufacturing, business, smart city, IoT, and tourism and hospitality is explained below.

3.1 Supply Chain and Logistics

Supply chains consist of several complex networks of manufacturers, suppliers, distributors, auditors, retailers, and consumers. The high degree of complexity makes it very difficult to follow and monitor everything reliably and efficiently. Blockchain can be a solution for this complex network of so many subdivisions of the supply chain. It would well organize the workflow of every party despite the size of the business network [20]. In addition to that, the shared nature of blockchain would provide transparency with greater visibility into shareholders. Traceability, transparency, and tradability will be three use cases supply chains have.

- **Traceability** improves operational productivity by planning and picturing endeavor supply chains. A developing number of purchasers request sourcing data about the items they purchase. Blockchain helps associations understand their supply chain and draws in purchasers with genuine, certain, and immutable data [21].
- **Transparency** develops trust by catching key information, for example, certifications and claims, and making the data available publicly.
- **Tradability** is a blockchain offering that redefines the regular marketplace idea. Any asset can be tokenized using blockchain by splitting the asset into shares that digitally represent the ownership. Like how a stock trade permits the exchanging of an organization's shares. These tokens are tradeable, and users can transfer ownership without the physical asset changing hands [22].

3.2 Internet of Things

According to IBM 2017, the last two years' data contribute around 90% of the world's total data. This drastic change happened because of the rapid increase of Internet of things (IoT) uses and the high-population growth. The count of interconnected devices is going to cross 30 billion by the year 2021 [23]. Keeping in mind, these huge numbers of devices and data generated by them requires a reliable and tempered proof data storage system. And blockchain is the most suited system for this application. The combination of IoT with blockchain technology has many interesting use cases. By implementing blockchain smart contracts, decision-making can be automated.

The strong security of blockchain against data tampering would help prevent a malicious machine from disrupting a house, warehouse, or transport system by hacking or sending misinformation. Moreover, by the structure, IoT is a distributed system, and blockchain provides its distributed ledger to record and to keep track of all data transmission happening to numerous devices. It reduces costs by removing the middleman overhead.

3.3 Gambling Industry

Gambling is all about money, and people cannot trust each other. Trust is one of the big challenges in this industry. If it comes to online gambling, there are several concerns in terms of fairness and transparency. There is always a risk in fund protection in the current gambling industries [24]. Blockchain technology is being leveraged for overcoming the existing problems in the gambling industry.

3.4 Cryptocurrency

Cryptocurrency and blockchain are two branches of a system. Blockchain did not exist until Bitcoin, the first cryptocurrency was introduced in 2008. Cryptocurrency is a way of transacting value that does not involve any third party like the government or bank. This system uses smart contracts (rules and protocols in terms of computer programs written for the blockchain network), and everyone who joins the network must have to agree to these smart contracts. Cryptocurrency is something that is intrinsic to blockchain. Every blockchain system does have cryptocurrency in it. In the blockchain system, it is generated while the process of mining the new blocks. Bitcoin blockchain has Bitcoin, and Ethereum has weather as cryptocurrency. Fund transfer takes place by using various mathematical encryption algorithms and techniques. For sending the fund, the public key address of the receiver is used, and the same address is used for recording the transaction also [1]. The personal private key is kept secret.

3.5 Ownership and Governance

Ownership is one of the main concepts in any economy. This is what decides the right and authority of assets tangible or intangible. Traditionally, it is very challenging and managing all these complex ownerships and rights in an accurate, efficient, and effective manner. There is always a security risk present in traditional approaches of keeping records of ownership. Blockchain could solve several problems where companies are unable to keep track of records related to who owns their shares, accurately and timely [25]. By applying smart contracts which eliminate the role of the middleman by automating the whole process of trading. Smart contracts and blockchain's transparent data storage help in reducing corruption. Other than just providing reliable data storage, blockchain can tokenize any kind of asset. Whether it is tangible or intangible or just an idea, all existing assets can be converted into manageable, digitally traceable tokens. All individual tokens represent the fractional ownership of the asset. Now, these tokens can be traded independently representing the partial ownership of the asset.

3.6 Agriculture and Food

Many recent research and projects considered blockchain as a relevant and mature technology for facing the challenges in the agri-food industry [26]. It can be used to bring an improved agri-food supply chain [27], enhance the trust among consumers and producers [28], and protect the crop from animal attack [29].

4 Issues in Blockchain and in Its Adoption

4.1 Initial Establishment Cost

Blockchain technology offers several advantages and long-term benefits by providing transparency, efficient, and reliable backend for data storage, which makes blockchain cost-effective. Despite its use cases, blockchain is still in its development stage. Only very few companies develop the software required for blockchain setup which makes it expensive. In addition to that, to run the specific software for blockchain, organizations need to purchase specific hardware too. As there is high demand and very less supply, organizations have to pay more. Talking about blockchain developers, here also the same situation is there. The demand for blockchain developers is high, but very few developers are available currently. Overall, these current scenarios make the initial blockchain establishment cost very high [30].

4.2 Scalability

Scalability has been a problem in blockchain since its creation. Here, scalability means the count of transactions happening per second. As referred to earlier in this paper, only 7 transactions can be performed on blockchain in one minute in Bitcoin blockchain and that is tremendously lower in number if we compare to the current transaction processing system. The unsolved scalability problem has been an obstacle to the blockchain practical applications and adoption of blockchain.

4.3 Lack of Standardization: Limited Interoperability

Blockchain is a technology that is distributed and decentralized in nature. Distributed and decentralized nature of blockchain gives the freedom to the blockchain developers and coders to customize and fit as per the requirements of the specific user [31]. The blockchain system lacks universal standards. Lack of standardization leads to the development of standalone projects with different protocols, different coding languages, different privacy measures, and different consensus mechanisms. This makes it challenging and complicated communication between two various blockchain networks. So, there is very limited interoperability in the current blockchain networks. Because of the lack of consistency among blockchain protocols, basic processes such as security suffer, making mass adoption nearly impossible.

4.4 Legal and Regulatory Uncertainties

Due to blockchain's autonomous structure, blockchain is facing several legal challenges across several parts of the world. It suffers from governance uncertainties. In addition to these challenges, it has to go through unclear taxation problems [32]. There are no taxation policies made for blockchain cryptocurrency transactions for sales of consumer goods, services. The decentralized characteristic of blockchain takes all decision rights, accountability, and incentives from IT Governance officers which makes blockchain's adoption challenging [32].

4.5 Selfish Mining

Bitcoin was the first use case of blockchain. It is been several years since its launch but still, it has many shortcomings. Selfish mining is one of them. In selfish mining, miners can gain more Bitcoins by creating a separate fork for a newly created block while hiding the block which is newly created from the main blockchain [8].

In Bitcoin, mining is the process of validating transactions that occurred over the Bitcoin network. And this process requires solving cryptographically complex puzzles. This requires a huge amount of computing power and energy. This transactions validation process and solving puzzles generate Bitcoins as a reward to the persons who solve the puzzles. This is how Bitcoin is mined and the person who does this mining work is called a miner. The Bitcoin protocol is built to incentivize miners as per standards and in proportion to their mining output. But the Bitcoin-like blockchain ecosystem can be manipulated by the miners seeking to enlarge their rewards in terms of Bitcoin while doing selfish mining [33].

5 Security in IoT with Evolving Applications of Blockchain

Sensors, cameras, networking devices, software, and network connectivity are the main components of the IoT world. By collecting and sharing information generated by interconnected sensors, IoT helps in the efficient utilization of available resources. IoT sensors generate more data that leads to better decisions.

Rapidly growing IoT is contributing to the transformation of our world into a dynamic and very complex system of interconnected devices. Whether it is a smart home, smart car, and smart refrigerator, all are making our life easier while generating a huge amount of data. Managing this huge quantity of data is a big challenge and providing security and integrity of this data is a bigger concern [34]. Existing centralized data management systems have technical limitations like scalability, data leakage, and unauthorized access to the central server. A single point of failure also contributes to that limitation list. In addition, cloud service providers gain an unfair

advantage in terms of data access and manipulation. Overall, the existing system of IoT is facing many security issues [35]. All IoT devices communicate using the cloud, so cloud security has been a major security challenge. There is much research done to tackle network security issues in IoT. Elliptic curve cryptography (ECC) authenticating scheme is being used to provide access to cloud-based robots [36]. For identity management, blockchain-based solution is proposed [37].

Blockchain comes with several features for distributed data management that include traceability, accountability, transparency, and audibility. Traditionally centralized stored data can be stored on a highly decentralized blockchain ecosystem where IoT data are processed by several blockchain peer nodes which eliminates centralized decision-making power.

Figure 2 depicts the system architecture of IoT with blockchain. This includes wireless sensor networks with sensors, and the information fetched from these networks is gathered in decentralized blockchain through the Internet. The authorized user/manager can access and interact through the Internet.

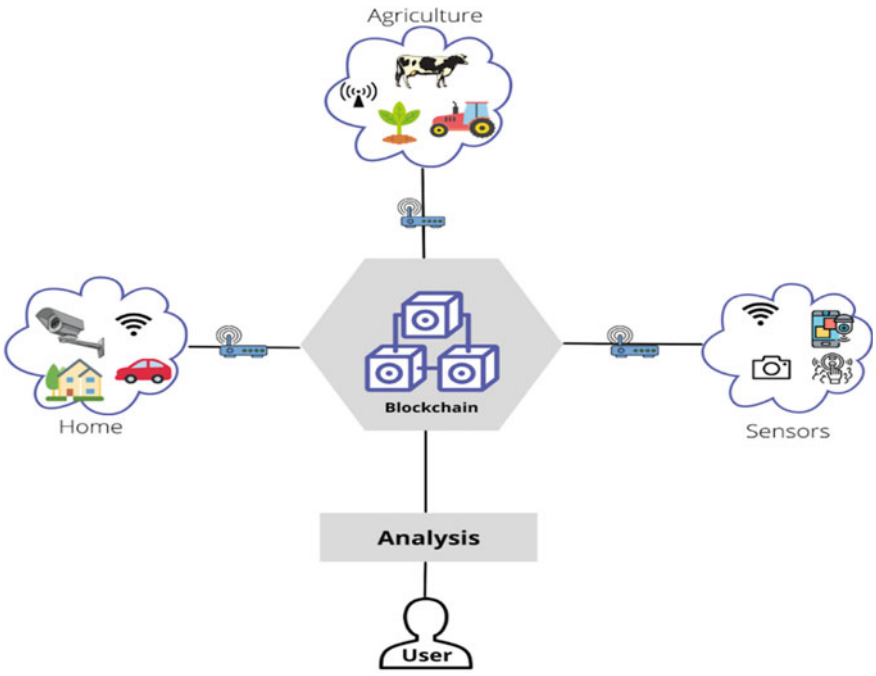


Fig. 2 Blockchain-IoT architecture

6 Recent Research on Blockchain-Based Security in IoT

This article has included a comprehensive literature review focusing on blockchain's integration in IoT. Few of them focus on the challenges, pros, and cons of integrating blockchain, few discuss the security advantages of this integrated system. Some papers point out privacy issues and identity theft problems. In the paper [39], the authors have analyzed the requirements of blockchain in 5G-enabled IoT networks. They also proposed a multilayer blockchain security model and showed that the proposed model is superior when compared to Ethereum blockchain. Similarly, in Refs. [35, 40], the authors broadly explained the pros and cons, and security advantages of blockchain in IoT systems. Using deep learning reinforcement learning, they evaluated their proposed algorithms and IoT framework.

Furthermore, papers [41–43] analyzed and discussed different consensus mechanisms. In paper [41], the authors have proposed proof-of-authentication (PoAh) and showed that PoAh is 200 fold faster than proof-of-work (PoW). Similarly, in paper [42], the authors have introduced proof of block and trade (PoBT) which reduces the computational time and increases transaction speed. In paper [43], also the author has shown removing proof-of-work (PoW) and using BA protocol instead helps reduce consensus time.

Security, privacy, and identity management are discussed in papers [34, 37, 38, 44].

Authors in paper [38] provided an analysis of the implementation of blockchain-based IoT systems' privacy protection measures. Similarly, in the paper [37], the authors proposed an identity management framework that offers a solution that enables both identity validation and tracing. In the paper [34], the authors discussed the security problems in the full life cycle of IoT devices. In paper [45], it was shown that a distributed IoT management system that compiles blockchain as a base does not achieve better performance than the optimized centralized IoT systems in the case of a single Management Hub.

Blockchain technologies are not suitable for most IoT devices as they are very expensive and possess high-bandwidth requirements and delays [44]. Ali Dorri et al. propose a lightweight blockchain-based architecture for IoT security and privacy. The optimized blockchain eliminates the overhead and, ensures the security and privacy of classical blockchain. In the research [46], Y Kurt Peker et al. analyze the cost of data storage in IoT devices using the blockchain concept. They observe that, even though this approach is expensive, it is a cost-effective alternative when data integrity and transparency are crucial. Further details about literature review on IoT based blockchain applications are included in Table 1.

Table 1 Review of different research on IoT-based blockchain

Author (year)	Application-specific	Research type	Blockchain usage	Tools used	Scalability	Cost
Honar Pajooh et al. (2021) [39]	5G-enabled network	Proposal with implementation	Data security	Hyperledger Caliper, IBM Cloud, IBM Watson IoT Platform, Node-Red server, Ethereum network	Yes	Not discussed
Joe, C. Vijesh et al. (2021) [47]	Medical	Proposal with implementation	Data privacy	Cloud and encryption algorithms, Byzantine fault tolerance	Yes	No
Zehui Xiong (2020) [35]	IoT	Proposal with implementation and analysis	Data management	Deep reinforcement learning	Yes	High
Mhaisen et al. (2020) [40]	Monitoring application	Proposal with implementation and analysis	System security	Reinforcement learning, deep Q-network, Ethereum	Not discussed	Low
Kurt et al. (2020) [46]	IoT	Proposal with implementation and analysis	Data storage	Testbed	Yes	High
Puthal et al. (2019) [41]	IoT	Proposal with implementation and analysis	Decentralized security	Raspberry Pi, Python, SQLite	Yes	Low
Biswas et al. (2019) [42]	Healthcare and IoT	Proposal with implementation and analysis	Security and privacy	Hyperledger fabric v1.0.2, IBM's Node-Red application, GoLong	Yes	Not discussed
Muneeb et al. (2019) [38]	IoT	Proposal without implementation	Privacy	No implementation	Not discussed	Not discussed

(continued)

Table 1 (continued)

Author (year)	Application-specific	Research type	Blockchain usage	Tools used	Scalability	Cost
Novo, Oscar (2018) [45]	IoT	Proposal with implementation and analysis	Access management	Ubuntu-16.04 desktop, CoAPBench, docker, Ethereum/client-go, LibCoAP library, Ethereum	Yes	Not discussed
Omar et al. (2018) [37]	IoT	Proposal with implementation	Identity management	Ethereum, smart contracts	Not discussed	Not discussed
Qian et al. (2018) [34]	IoT	Proposal without implementation	Security management	No implementation	Not discussed	Not discussed
Cebe et al. (2018) [43]	IoT	Proposal with implementation	Security and privacy	IEEE 802.11 s-based wireless mesh network, Protonix Wi-Fi dongles, Raspberry Pis	Yes	Not discussed
Ali Dorri et al. (2017) [44]	IoT	Proposal with implementation	Security and privacy	NS3	Yes	Low

6.1 Summary

From the above table, we found that most of the researchers have used the Ethereum blockchain network for their research implementation. In the following papers [35, 40], authors have used reinforcement learning for studying the pros and cons of using blockchain in IoT infrastructure. They also proposed IoT architecture and framework.

Most papers are focused on various security problems, i.e., privacy, data leak, identity theft, etc. In one paper [47], medical image security is discussed. The paper [39] has focused on 5G-enabled IoT architecture and proposed a multilayer blockchain security model that protects IoT networks while making implementation easier. Very few papers have taken into consideration the scalability problem present in blockchain networks.

7 Conclusion

Blockchain technology has many things to offer. Immutability, decentralization, faster settlement, enhanced security, etc., make blockchain the most suitable technology for IoT infrastructure. Blockchain's integration in the IoT environment brings several benefits but as the system is in its development phase, both fields require much attention. An extended study helps in identifying various problems and challenges in the integrated system of IoT and blockchain. It has been observed that research in this field is directed toward the area of data security and users' privacy. There is very little research done to address the scalability issues present in the blockchain-IoT infrastructure. It is also noticed that blockchain lacks standardization. There are lots of blockchain projects coming, but there is very little or no interoperability present in them. They cannot communicate with each other. So, this area also requires researchers' attention.

References

1. Lin I-C, Liao T-C (2017) A survey of blockchain security issues and challenges. IJ Netw Secur 653–659
2. Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system. Decentralized Bus Rev 21260
3. Xu M, Chen X, Kou G (2019) A systematic review of blockchain. Financ Innov 5:1–27
4. Zīle K, Strazdiņa R (2020) Blockchain use cases and their feasibility. Appl Comput Syst 23(1):12–20
5. Al-Megren S, Alsalamah S, Altoaimy L, Alsalamah H, Soltanisehat L, Almutairi E (2018) Blockchain use cases in digital sectors: A review of the literature. In : 2018 IEEE International conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData). IEEE, pp 1417–1424

6. Casino F, Dasaklis TK, Patsakis C (2019) A systematic literature review of blockchain-based applications: current status, classification and open issues. Telematics Inform 36:55–81
7. Tama BA, Kweka BJ, Park Y, Rhee KH (2017) A critical review of blockchain and its current applications. In: International conference on electrical engineering and computer science (ICECOS). IEEE, pp 109–113
8. Zheng Z, Xie S, Dai HN, Chen X, Wang H (2018) Blockchain challenges and opportunities: a survey. Int J Web Grid Serv 14(4):352–375
9. Kolb J, AbdelBaky M, Katz RH, Culler DE (2020) Core concepts, challenges, and future directions in blockchain: A centralized tutorial. ACM Comput Surv (CSUR) 553(1):1–39
10. Croman K, Decker C, Eyal I, Gencer AE, Juels A, Kosba A, Miller A, Saxena P, Shi E, Sirer EG, Song D (2016) On scaling decentralized blockchains. In: International conference on financial cryptography and data security. Springer, Berlin, Heidelberg, pp 106–125
11. Yli-Huumo J, Ko D, Choi S, Park S, Smolander K (2016) Where is current research on blockchain technology?—a systematic review. PloS one 11(10):e0163477
12. Haber S, Stornetta WS (1991) How to time-stamp a digital document. J Cryptol 3(2):99–111. CiteSeerX 10.1.1.46.8740. https://doi.org/10.1007/bf00196791
13. Vujicic D, Jagodic D, Randic S (2018) Blockchain technology, bitcoin, and Ethereum: a brief overview. In: 2018 17th international symposium infoteh-jahorina (infoteh). IEEE
14. Bodkhe U, Tanwar S, Parekh K, Khanpara P, Tyagi S, Kumar N, Alazab M (2020) Blockchain for industry 4.0: A comprehensive review. IEEE Access 8:79764–79800
15. Min XC, Kou G (2019) A systematic review of blockchain. Financ Innov 5:1–27
16. Sabounchi M, Wei J (2018) Blockchain-enabled peer-to-peer data trading mechanism. In: 2018 IEEE international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData). IEEE
17. Ribeiro V et al (2020) Enhancing key management in LoRaWAN with permissioned blockchain. Sensors 20(11):3068
18. Nadiya U, Mutijarsa K, Rizqi CY (2018) Block summarization and compression in bitcoin blockchain. In: 2018 international symposium on electronics and smart devices (ISESD). IEEE
19. Ribeiro V, Holanda R, Ramos A, Rodrigues JJ (2020) Enhancing key management in LoRaWAN with permissioned blockchain. Sensors 20(11):3068
20. Francisco K, Swanson D (2018) The supply chain has no clothes: Technology adoption of blockchain for supply chain transparency. Logistics 2(1):2
21. Kumar R, Tripathi R (2019) Traceability of counterfeit medicine supply chain through Blockchain. In: 11th international conference on communication systems & networks (COMSNETS). IEEE
22. Li X, Wu X, Pei X, Yao Z (2019) Tokenization: Open asset protocol on blockchain. In: IEEE 2nd international conference on information and computer technologies (ICICT). IEEE, pp 204–209
23. Nativi S, Kotsev A, Scudo P, Pogorzelska K, Vakalis I, Dalla Benetta A, Perego A (2020) IoT 2.0 and the INTERNET of TRANSFORMATION. Publications Office Eur. Union, Luxembourg, Tech. Rep. JRC120372
24. Gainsbury SM, Blaszczynski A (2017) How blockchain and cryptocurrency technology could revolutionize online gambling. Gaming Law Rev 21(7):482–492
25. Yermack D (2017) Corporate governance and blockchains. Rev Financ 21(1):731
26. Kamilaris A, Cole IR, Prenafeta-Boldú FX (2021) Blockchain in agriculture. In: Food technology disruptions. Academic Press, pp 247–284
27. Sivaganesan D (2021) Performance estimation of sustainable smart farming with blockchain technology. IRO J Sustain Wirel Syst 3(2):97–106
28. Anand R, Divya Karthiga R, Jeevitha T, Leandra Mithra J, Yuvaraj S (2021) Blockchain-based agriculture assistance. In: Advances in automation, signal processing, instrumentation, and control. Springer, Singapore, pp 477–483
29. Iqbal R, Butt TA (2020) Safe farming as a service of blockchain-based supply chain management for improved transparency. Clust Comput 23(3):2139–2150

30. Madaan L, Kumar A, Bhushan B (2020) Working principle, application areas and challenges for blockchain technology. In: IEEE 9th international conference on communication systems and network technologies (CSNT), Gwalior, India, pp 254–259. https://doi.org/10.1109/CSNT48778.2020.9115794
31. Prewett KW, Prescott GL, Phillips K (2020) Blockchain adoption is inevitable—Barriers and risks remain. J Corp Account Financ 31(2):2128
32. Biswas B, Gupta R (2019) Analysis of barriers to implement blockchain in industry and service sectors. Comput Ind Eng 136:225–241
33. Eyal I, Sirer EG (2014) Majority is not enough: bitcoin mining is vulnerable. In: International conference on financial cryptography and data security. Springer, Berlin, Heidelberg
34. Qian Y, Jiang Y, Chen J, Zhang Y, Song J, Zhou M, Pustišek M (2018) Towards decentralized IoT security enhancement: a blockchain approach. Comput Electr Eng 1(72):266–273
35. Xiong Z, Zhang Y, Luong NC, Niyato D, Wang P, Guizani N (2020) The best of both worlds: a general architecture for data management in blockchain-enabled Internet-of-Things. In: IEEE network 34(1):166–173. https://doi.org/10.1109/MNET.001.1900095
36. Jain S, Doriya R (2021) Authentication of robots using ECC to access cloudbased services. In: Intelligent data communication technologies and internet of things. Springer, Singapore, pp 861–870
37. Omar AS, Basir O (2018) Identity management in IoT networks using blockchain and smart contracts. In: mIEEE international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData). IEEE
38. Hassan MU, Rehmani MH, Chen J (2019) Privacy preservation in blockchain based IoT systems: Integration issues, prospects, challenges, and future research directions. Future Gener Comput Syst 97:512–529
39. Honar Pajooh H, Rashid M, Alam F, Demidenko S (2021) Multi-layer blockchain-based security architecture for internet of things. Sensors 21(3):772
40. Mhaisen N, Fetais N, Erbad A, Mohamed A, Guizani M (2021) To chain or not to chain: a reinforcement learning approach for blockchain-enabled IoT monitoring applications. Futur Gener Comput Syst 1(111):39–51
41. Puthal D, Mohanty SP, Nanda P, Kougianos E, Das G (2019) Proof-of-authentication for scalable blockchain in resource-constrained distributed systems. In: IEEE international conference on consumer electronics (ICCE). IEEE, pp. 1–5
42. Biswas S, Sharif K, Li F, Maharjan S, Mohanty SP, Wang Y (2019) PoBT: a lightweight consensus algorithm for scalable IoT business blockchain. IEEE Internet Things J 6 7(3):2343–2355
43. Cebe M, Kaplan B, Akkaya K (2018) A network coding based information spreading approach for permissioned blockchain in iot settings. In: Proceedings of the 15th EAI international conference on mobile and ubiquitous systems: computing, networking and services
44. Dorri A, Kanhere SS, Jurdak R (2017) Towards an optimized blockchain for IoT. In: IEEE/ACM second international conference on internet-of-things design and implementation (IoTDI), pp 173–178
45. Novo O (2018) Scalable access management in IoT using blockchain: a performance evaluation. IEEE Internet Things J 6(3):4694–4701
46. Kurt Peker Y, Rodriguez X, Ericsson J, Lee SJ, Perez AJ (2020) A cost analysis of internet of things sensor data storage on blockchain via smart contracts. Electronics 9(2):244
47. Joe CV, Raj JS (2021) Deniable authentication encryption for privacy protection using blockchain. J Artif Intell Capsule Netw 3(2021):259–271

Study of Land Cover Classification from Hyperspectral Images Using Deep Learning Algorithm

K. Karthik, M. Nachammai, G. Nivetha Gandhi, V. Priyadharshini, and R. Shobika

Abstract As image sensor electronics have progressed, and hyperspectral images have been used in a wide range of applications. In terms of recognizing the classes, a lot of research work has been done to extract useful information from the available unstructured knowledge database. The use of spectral and geographical datatypes in images can improve the classification precision. To improve the accuracy of the energetic-spectral snap analysis, combining dimensional and spectral data is a good idea. This research study examines the history of dimensional facts based on energetic-spectral image classification designs by using prepared and semi-directed classifiers to classify detached sensing images with particularized class labels. Long-term data are removed if the features are protensive. To increase the veracity of classification, the extracted traits are prepared by utilizing multiple classifiers. The preparation and sinking balance towards the loss function have been reused to train the classifiers. To avoid local minima, the preparation is approved by utilizing various tiers for accompanying the extra balancing limits. To overcome the risk of establishing a contradictory validity image, each detached perceiving image is classified throughout the experiment stage. Exploratory discoveries demand higher validity in-class criteria than other advanced directed classifier techniques.

Keywords Spectral features · Spatial information · Deep neural network · Noise filtering · Classification

1 Introduction

In the domain of land cover classification, it is highly required to first examine how the mild components interact with compounds according to their spectral fingerprints in order to better identify, uncover, or study their characteristics. Every fabric has its own spectral signature, which is analogous to our fingerprint for its unique identity, and it

K. Karthik (✉) · M. Nachammai · G. Nivetha Gandhi · V. Priyadharshini · R. Shobika
Department of Computer Science and Engineering, M.Kumarasamy College of Engineering, Karur, Tamilnadu 639113, India
e-mail: karthikk.cse@mkce.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_54

may be determined from the spectrum of the material to determine the amount of light received, transmitted, or emitted at various wavelengths [1–5]. Spectral imaging is an approach that merge digital imaging with spectral size. Unlike a common digital camera, which records light in the visible spectrum in purple, blue, and green, a spectral imaging digital camera can capture light in tiny wavelengths starting from the UV and moving through the visible and infrared areas [6]. Their broad-spectrum capability allows for easy detection and separation of visually identical compounds in a textile based on colour or chemical differences. Many spectral picturing and hyperspectral picturing are two types of spectral picturing. The spectral decision is the main contrast in the middle of MSI and HIS [7].

An HSI digital camera continually measures light in multiple narrow wavelength bands, whereas an MSI camera only measures a discrete range of wavelength bands. Area borne or airborne rador s like AVIRIS, HyMap, HYDICE, and Hyperion can readily capture masses of narrow spectrum wavelengths for each image pixel due to the rapid progress and dissemination of hyperspectral remote sensing technology. This distinct spectral reflectance profile allows for accurate separation of compounds of interest. Because of the many needs in ecological science, ecology control, precision agriculture, and navy applications, a wide range of hyperspectral picture classification algorithms based on spectrum similarity and spectral-spatial function have appeared on the scene. These approaches are classified as either monitored. The former is usually accomplished by first grouping the data and then deciding the categories manually [8–12]. Those supervised procedures employ effective device learning algorithms to train a selection rule that uses label information to predict the labels of the trying out pixels. In this study, we pivot on supervised hyperspectral picture categorization algorithms. Over the last decade, the global sensing community has been working hard to develop an accurate hyperspectral picture classifier [13–18]. The variation in the middle of many-spectral and excitable-spectral picturing is seen in Fig. 1.

2 Related Works

Land cover is a key characteristic that influences and connects many aspects of the human and physical surroundings. As a result, data on the geographical distribution of land-cowl lessons are crucial for the study of environmental solutions. For environmental monitoring, satellite far-flung sensing technologies are widely utilized. The following section contains a summary of many papers.

Wang et al. [19] suggest a novel technique for band selection depending on magnetic resonance imaging. The varied structure is appropriately examined as a alternative of the likeness score inside the Euclidean space to analyze the organization of hyperspectral statistics. The ranking procedure is completed at the supplied connected measures, which produce the rating score, which allows for the selection of further facts. This method of grading issues has been developed, which uses a hyperssspectral belt decision that is follow one another. In batch mode, the interval between

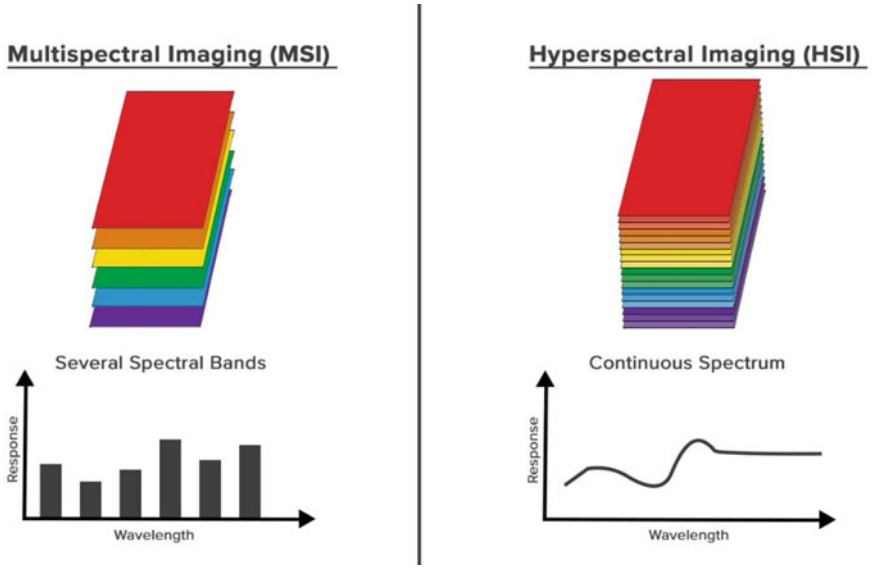


Fig. 1 Difference between multi-spectral and hyperspectral ımaging

the inter-bands is predicted. In most current approaches, the distance between two character bands is calculated. The enumerated effects are then utilized to help in band mixture. This strategy will not work for selecting a band in a series since the chosen one may be similar to the preceding one. As a quiz, the group of before selected bands is favoured, and the complete group is substantially tested in the execution. With the previously picked ones, an additional selected band may be secured in a big way.

Gustavo Camps-Valls et al. [19] are qualified to accurately define the goods in question (for example, the splendour of a land cowl) and to keep inventories up to date. Spectral resolutions have been upgrade in order to promote sign processing and algorithm exploitation. Recently, classification on hyperspectral images has gotten a lot of notice as a demanding trouble that has piqued the interest of various scientific fields including gadget studying, picture processing, and computer imaginative and prescient. In the remote sensing world, the approach of allocating a single pixel to a set of sophistication is known as "category", whereas the process of grouping pixels into gadgets and then assigning to a class is known as segmentation. In the past, straight discriminant survey or Gaussian highest probability was employed instead of HSI as variables for type in telemetry. These methods photos include 4–10 bands in size, based on the calculation of the covariance matrix, which proved successful when handling early multi-spectral images. HSI changes the laws when the measurement of the dots is growed by century. The prevalent parametric technique for the calculation of covariance matrices becomes difficult or inaccurate due to the lack of classified data. Regularization has been addressed inside the involved covariance matrices both

explicitly by Tikhonov's phrases, and by performing type in a smaller subspace for the unavailability of labelled samples.

Hong [20] resolved CNNs and GCNs accompanying a devote effort to something HS image categorization. To the best choice of our information, this is the first occasion that the potentials and disadvantages of GCNs (distinguished with CNNs) are examined the society. And projected a novel directed interpretation of GCNs: miniGCNs, for short. As the name plans, miniGCNs can be prepared in tiny-lot fashion, difficult to find a better and stronger local best. Unlike traditional GCNs, our miniGCNs are not only fit the preparation of the networks utilizing the preparation set but more allow for an honest conclusion of big, out-of-samples utilizing the prepared model. And grown three fusion blueprints, containing supplement melding, factor-intelligent multiplicative fusion, and connection melding, to gain better categorization results in HS figures by mixing features gleaned from CNNs and our miniGCNs, in an end-to-end educable network.

Xiong [21] created the WJCR categorizer, which is a much futured variant of the JCR categorizer. Whilst the spatial and spectral properties are being retrieved, the JCR classifier gives the pixel of acquaintances the same weight. JCR uses the matching force for removing spatial and spectral functions from nearby dot. Unlike JCR, WJCR seeks to use the appropriate dot force by paying attention to the dot in the midpoint and its neighbouring. Because it belongs to exceptional instructions, the centre pixel and nearby pixels are at the heterogeneous position. In varied regions, choosing pixels with similar weights by JCR is unsatisfactory. When this happens, the nearby pixels that can be linked to the central one are taken into account. However, putting off the unnecessary pixels when the computing complexity would grow is not always a good idea. As a result, the surrounding pixel was concerned with an adaptive weight, which was carried out appropriately. WJCR uses the Gaussian kernel feature to select the most appropriate dot loads. With the help of WJCR, the spatial and spectral functions in heterogeneous images are appropriately retrieved, giving a benefit of additional precision.

Gao [22] used a depth trimming approach on the captured subcategory whilst using the few schooling samples accessible. Subcategory-projection, has been enhanced using a many nominal arrange regression technique. Hyperspectral information is evaluated and classified into blended pixels using this approach. Support vector gadget has been incorporated to the subcategory-projection notion for categorizing images of far off sensing. SVM non-linear characteristics are used to generate the subspaces that are connected to each magnificence.

Wu et al. [23] provide a brand up-to-date lay of SFA alternative detection methods for detecting the alive of real alterations in two-temporal many spectral snap photos. We rewrote SFA's mathematical technique for exchange detection in order to fit it to the bi-temporal alternate detection problem. To remove the much constant aspect from the temporal photos, the SFA alternative detection technique is used. The unmodified pixels in the converted function space have relatively modest difference values, therefore the changed areas are comparably emphasized.

Ertürk and others [24] examined in contrast, hyperspectral alternate detection through spectral unfixing has the potential to yield easily interpretable information

about the character of the change, as well as sub-pixel-level alternate statistics. A natural disaster, such as a flood, could result in a large increase in the abundance of "water" stop members in the scene, a change in the rising lot on estate, which could result in the giving and taking of the prominent give up subscriber in the sector, agreement incrusion into, i.e. growth in the copious of synthetic substance, or pass time objectives, which could result in up to date and abnormal end individuals in the spot, or agreement encroachment, i.e. in spite of the element that excitable supernatural change recognition by unmixing offers the important sake of delivering sub-pixel-stage results and data on the essence of the change, it is still in its infancy in the literature, with the majority of study limited to case studies. A general framework for spectral unmixing-based alternative detection is given. For a case study, sub-pixel-stage alternative detection through unmixing is used. It is mentioned that land-cowl mapping can be done through unmixing. Change detection through spectral unmixing is thoroughly investigated, and the authors present sub-pixel-level information on the character of the exchange that such a technique provides by employing the authors in a recent painting. In this paper, scattered unmixing is used for business recognition in various temporal hyperspectral data at the central time. Sparse unmixing, when combined with spectral libraries, avoids the problems of traditional spectral unmixing whilst still displaying sub-pixel-level recognition and information on the type of the modification by generating a distinct trade map for each end associate.

Demir et al. [25] developed a one-of-a-kind trade-detection-driven control mastery method to apprise dry-land cover maps by categorizing distant detecting imageries taken at the same location at changed periods (i.e. picture time series). A trustworthy schooling set must be provided just for one of the photographs (i.e. the supply region) inside the time collection, according to the advocate approach, whilst it is not always available for any other photo to be categorized (i.e. the target area). Different further literature switch mastery strategies, there are no further assumptions on the comparability of elegance deliveries or the availability of the similar set of land covers lessons in both dominion names. The suggested technique aims to define a credible education set for the objective area whilst taking advantage of the previously existing supply domain expertise. This is accomplished by using an unmonitored exchange detection technique to target and supply domain names, then initializing the target area schooling set by shifting class labels of identified unmodified training samples from the source to the goal area. After that, a well-described innovative active learning (AL) approach is used to maximize the schooling set. The most useful models are picked from modified and unchanged unlabelled samples in the early rounds of AL. In the final rounds, the much informative illustrative are chosen from altered and unaffected unlabelled specimens. The final step is to classify the desired pictures. Because of the increased number of images routinely taken using satellite-borne sensors at the same places at specified times, updating land-cowl maps with the use of a category of far-flung sensing photos is a critical challenge. Because of new policies relating to open, this issue is becoming increasingly crucial, as time collecting is now accessible to all possible users in a scientific manner. Land cover maps can be renovate via supervised direct organization of one and all snapshot in the time line. However, in order to successfully train the classifier, such a strategy needs

trustworthy floor reference information for all of the available temporal snapshots. Because of the high value and time-consuming nature of this job, acquiring a necessary range of classified training samples for each unmarried image to be categorized is not practicable in operational settings.

Hang [26] projected a two-arm spectral-dimensional consideration network for hyperspectral concept classification. Compared to the existent CNNs, our model combines consideration modules for each convolutional layer, making CNNs devote effort to something more discriminating channels and relating to space positions whilst suppressing useless one. In the classification state, two-arm results are melded together next to an adjusting burden summation order. Considering the limited numbers of development samples, we intend an insignificant spooky concern piece going around two convolutional drivers. Before convolutional tiers, use global average combining to lower the belongings of dimensional information. More basically, we adjoin a yield layer in the piece to aid allure knowledge process. Similar to the spectral consideration piece, we more use two convolutional layers to build the relating to space consideration piece. Instead of using combining drivers, we select one 1×1 convolutional tier for lowering the number of channels to 1. Also, an output layer is added to guide the learning process of the spatial module.

Cao et al. [27] projected a means to integrate two together deep education and AL methodology into a united foundation by sufficiently utilizing the benefits of two together worlds, that is to say, the strong discriminating skill of deep knowledge and labelling adeptness of AL. Although the skilled have been various combinations of AL accompanying deep knowledge for HSI categorization and proposed design, it has allure specific traits. First, the projected plan adopts different deep network construction, that is to say, CNN whilst the previous arrangements use additional network architectures, in the way that SAE, restricted Boltzmann vehicle (RBM), and Bayesian CNN (BCNN). Second, the projected method uses various tests to energetically select informative samples. Third, the projected form considers contextual news by utilizing MRF. More studies about these dissimilarities will be made clear in related work. Therefore, this proposed arrangement may be judged as a further attempt to integrate AL accompanying a deep neural network in addition, to analyze more contextual news so that weaken the labelling cost.

Akey Sungheetha [28] show that they can recognize fine optic analyzes accompanying an extreme level of veracity. Essentially, the plan searches elude two CNNs on the alike figure because to extract features, before combining those appearance promoting a bilinear normalization process to design a new feature vector. Before the experiment, the dataset was divided into training and test sets. In this study, a range of training percentages was used in the tests to involve corresponding accompanying added preparation orders. The dataset was devised by selecting random training and test sets from the original dataset to train and test the algorithm.

Edriss Eisa Babikir Adam et al. [29] executed the system which can analyze the captured image will be utilized as an input image, and it will evolve into grayscale for further concept analysis. The collage algorithm is used to reduce overlaps and leftover differences from the input raw figures. It is attainable to obtain a unique level of veracity at this moment method. After the collage pattern is completed, the

noise elimination filtering operation will start. On actual bridge photos, the proposed composite labelling pattern is tested for cracks. In addition, we argued for new research on actual structure break detecting wholes. During the process, the proposed method allowed improved categorization and minority categorization errors.

Mittal and Gelbukh [30] achieved a big warehouse of knowledge, and they aid in change detection supervising. Change detection is a detracting action in the fields of land development auditing, trouble management, means administration, and environmental evaluation. Segmentation of a picture is used to settle the fields of interest for change detection. Particle swarm addition, stimulated by nature, is a Meta wondering algorithm, namely plain, strong, and demands fewer assumptions for the task within reach. This research uses the MATLAB surroundings to expand a piece swarm addition (PSO) method as an edge detection separation tool for subsidiary photos got from Google Earth.

3 Hyperspectral Image Classification Using Machine Learning

ML uses two sorts of procedures: supervised learning, which includes training a prototype on known input and output files to assist in predicting the results, and unsupervised learning, which is identifying the underlying patterns or structures present in the datasets. Supervised ML: A model is constructed in supervised machine learning to incorporate computations based on the indicators of uncertainty. A supervised learning approach trains a model to produce reliable predictions for incoming data by using a known set of input files and reactions to data (output) [31, 32]. It is highly recommended to use supervised learning, if you have data for the output and wish to forecast anything. Supervised learning uses categorization and regression processes to improve the forecasting prototypes.

Methods for classifying pixels anticipate discrete responses like whether a pixel is successfully identified or not. Input data are classified using classification models. Medical imaging, voice recognition, and credit scoring are examples of common presentations. If data can be labelled, considered, or separated into distinct groups, use classification. Applications for analyzing diverse land types in hyperspectral datasets [33–38]. Unsupervised design recognition algorithms are used in image processing and computer sight to detect objects and separate pictures. The support vector machine (SVM), pushed and bagged decision trees, k-nearest neighbour, Naive Bayes, discriminant analysis, logistic regression, and neural networks are all common classification techniques.

Regression approaches predict consistent responses, such as temperature shifts or weather data swings. Weather forecasting based on satellite data and algorithmic trading are two examples of typical uses use regression approaches if you are working with a data range or the response is a real number, such as temperature or the time before a piece of equipment breaks. Common regression measures insert the linear

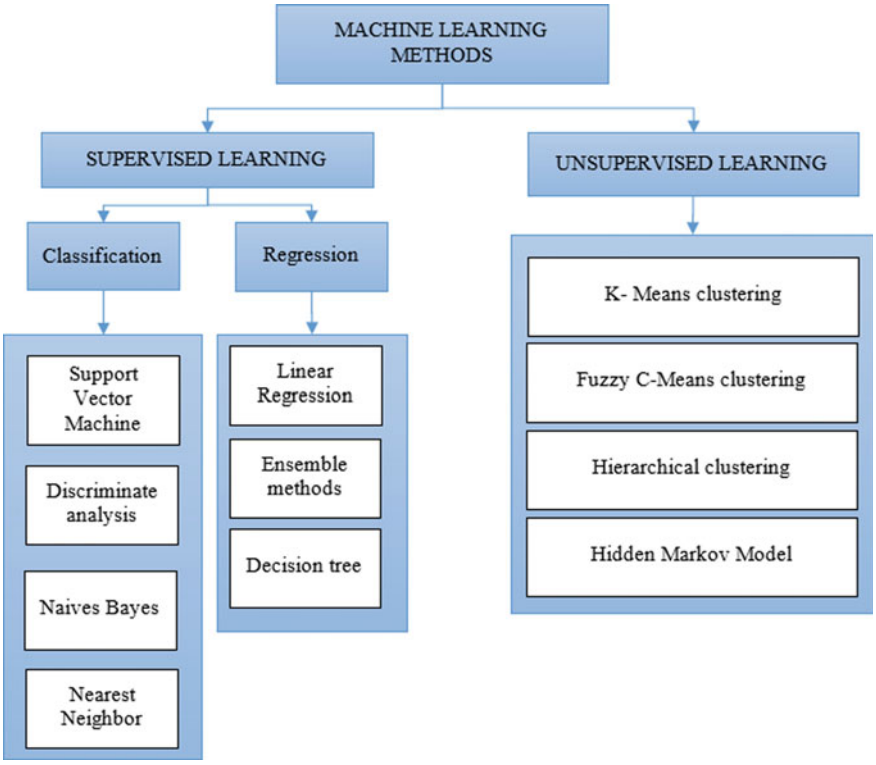


Fig. 2 Machine learning algorithms

prototype, non-linear prototype, regularization, stepwise regression, boosted and bagged decision tree, and adaptive neuro-fuzzy learning.

Unsupervised learning is the discovery of secret designs or unaffected designs in information. It is used to infer conclusions from datasets having input data but no tagged answers. The most frequent unsupervised learning approach is clustering. It is utilized in exploratory information survey to uncover hidden designs or groups. Sequence analysis, market research, and entity identification are examples of cluster analysis applications. Figure 2 exposes the different machine learning approaches [39–41].

4 Difference between Conventional Image and Hyperspectral Image

Spectral depict offers a common RGB colour image by replacing the three channels corresponding to three abundant wavebands, regularly referred to as “red”, “green”,

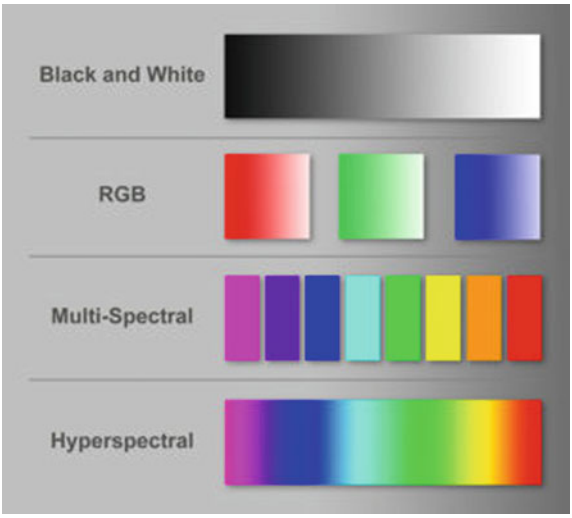


Fig. 3 Spatial distribution

and "blue", accompanying a significantly taller number of wavebands. Multi-spectral imaging (MSI) and hyperspectral imaging (HSI) are established with an identical principle and clash only in the number of channels collected. A spectral concept holds many ranges, each individual point on the sample's surface. Figure 3 contains valuable news about the dimensional allocation of the matters inside the sample.

5 Dataset Description

The Pavia University HSI was seized by the ROSIS sensor all whilst a departure campaign over Pavia, northern Italy. The HSI info is openly possible on Group De Inteligencia Computational (GIC). The Pavia University HSI has 103 spectral bands, and it holds 610 * 340 pixels, but spectral samples of the representation hold no facts that are delimited as zero (0), and they should be rejected before the study. The lines determination is 1.3 m. The ground validity of the HSI is changed into 9 classes. Figure 4 describes input image.

Fig. 4 Hyperspectral image *(Source* Kaggle.com/ PaviaUniversity)

6 Deep Learning Algorithm for Hyperspectral Image Processing

To category hyperspectral pictures, several neural matrix approaches are utilized. The contemporary unquiet are end propagation and convolutional neural matrix. The following is a description of these two networks:

6.1 Back Propagation Neural Network

To detect the different types of HSI pixel in the present substructure, a settling approach called end propagation neural network (BPNN) is used. Artificial neural networks (ANN) may be learned and hence trained to detect patterns in data, anticipate future occurrences, develop solutions, and categories data. Behavioural learning in neural networks is influenced by the way separate parts are linked turn on computation and the power of these links (loads). These weights can be modified automatically by training the network until it performs the intended task satisfactorily, according to a given learning rule. The supervised learning technique, often known as ANN,

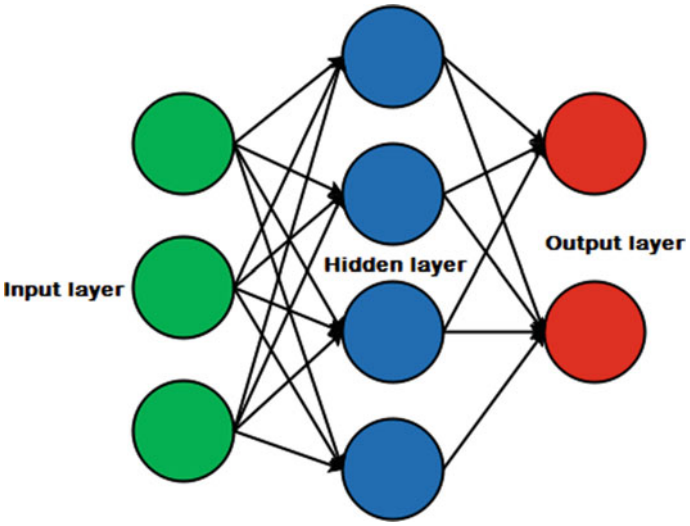


Fig. 5 Graphical representation of back

is a machine learning algorithm that employs a known dataset, or training dataset. These well-known variables aid ANN in making forecast. The input data and their response values are the most important parts of a training dataset [42, 43]. A supervised learning approach is created by making predictions on unknown and untaught dataset response values. A test dataset is used to validate the model. Using bigger training datasets are the greatest way to improve prediction power and generalization capability for a variety of new datasets.

The end propagation technique is put in to categorize the HSI pixels. End propagation is a typical approach for teaching fake neural matrix to reduce the goal function. It is a supervised learning tactic that is made up of several rules that have been generalized. The building of a training set necessitates a large number of inputs in order to create the desired output. This is specific merits in feed-ahead networks (the networks which have no feedback, in easy words, it will have no links in the bend). "Backward error propagation" [44–49] is also a shorthand for this term. The input layer, hidden layer, and output layers are shown in Fig. 5.

6.2 Steps in BPP Algorithms:

Step 1:

The forces are adjusted, and diagonal is ready at random.

Step 2:

It is time to input the training sample.

Step 3:

Forward the inserts; in the secret layers and product layers, compute the net input and product of each unit.

Step 4:

Back propagation of the guilt to the buried sheet.

Step 5:

Based on the reflection of propagated mistakes, weights and biases are rearranged. Training and mathematical learning functions are utilized to change the network's weights and biases without requiring human participation.

Step 6:

Terminate the condition. For dot categorization in land cover recognition, BPNN supplies shorter exactness and require great numeral repetitive stair to codify the dots.

7 CNN-Based Classification

The absolutely linked convolutional sheets and max are combined to represent the CNN feed-ahead neural matrix. By reinforcing the patterns of nearby link in the middle of neighbouring layers neurons, this takes advantage of spatial neighbourhood correlation. In mammalian visual cortex, the intricate characteristics of convolutional sheets and easy cells are replicated and swapped with max pooling layers. To figure a full neural matrix CNN, several sets of convolution sheets and max pooling sheets are used. The efficient study of observable representations convincingly proves the hierarchical nature of CNN. The fundamental issue in such visual tasks is the appearance of inter-class and forms variations modelling items. Hyperspectral data are represented as 2D bend with tonnes of spectral channels. It is difficult to distinguish relative classes with the naked eye, as each class curve has its own visual form that is distinct from that of other classes (e.g. gravel and self-blocking bricks). By exploiting spectral signatures, CNNs can solve competitive challenges and outperform humans; this skill motivates researchers to investigate the feasibility of using CNNs for HSI categorization. How the convolutional layers and max pooling layers are built, as well as how the network is trained, affect the CNN [50–53].

Each dot in a representation can be written as an extreme-spatial vector whose entries pertain to the spectral features (histogram, intensity, repetitiveness) in each band, and CNN supplies a strong feature extractor, which resides of alternative loop and combining coatings, to generalize the features towards deep and abstract likenesses. As it is from independent feature learning and it supplies the essential premise for more extreme-accuracy classification.

8 Validation Metrics

The confirmation of the expected system may be analyzed in conditions of accuracy measured.

True positive (TP): number of true positives − perfect positive prediction.
False positive (FP): number of false positives − imperfect positive prediction.
True negative (TN): number of true negatives − perfect negative prediction.
False negative (FN): number of true negatives − imperfect negative prediction.

9 Accuracy

Accuracy (ACC) is found as the fraction of total number of perfect predictions to the total number of test data. It can also be represented as 1 – ERR. The finest possible accuracy is 1.0, whereas the very worst is 0.0. The performance of the system shown in Table 1 and Fig. 6.

Table 1 Findings of accuracy

Algorithm	Accuracy (%)
Naive Bayes	50
Support vector machine	65
Back propagation neural network	80
Proposed work	92

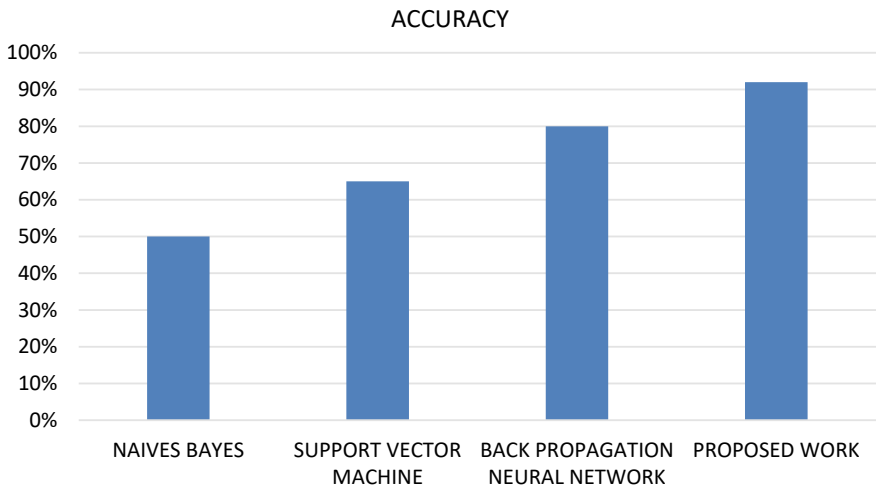


Fig. 6 Performance chart

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \times 100$$

From Fig. 6, it is evident that, CNN provides improved accuracy rate than the existing machine learning algorithms. The Naive Bayes algorithm utilizes probability distribution functions, and hence, it results in delivering shortest accuracy in categorization.

10 Conclusion

This research study has evaluated different machine learning and deep learning algorithms. For HSI categorization, a deep learning CNN technique has been suggested. From the hyperspectral picture categorization observation, the respective pixel is treated as sequential detail. This suggested strategy predicts higher learning percentage than existing neural network methodologies. For the initial period, the intrinsic sequential data structure of hyperspectral pixels is explored to provide a novel technique for improved hyperspectral data comprehension, processing, and modelling. Furthermore, the novel technique provides higher classification accuracy than existing widely used classification algorithms with reduced CPU time consumption. The suggested system is totally computerized, which is a very desirable feature to thoroughly certify the huge CNN attribute for hyperspectral image processing.

11 Future Work

The truthfulness of the genre will be enlarged in the future. Furthermore, this work will be able to establish more accurate results accompanying better accuracy.

References

1. Zhu Q, Zhong Y, Zhao B, Xia GS, Zhang L (2016) Bag-of-visualwords scene classifier with local and global features for high spatial resolution remote sensing imagery. IEEE Geosci Remote Sens Lett 13(6):747–751
2. Li X, Mou L, Lu X (2015) Scene parsing from an MAP perspective. IEEE Trans Cyber 45(9):1876–1886
3. Lu X, Li X, Mou L (2015) Semi-supervised multitask learning for scene recognition. IEEE Trans Cyber 45(9):1967–1976
4. Olmanson LG, Brezonik PL, Bauer ME (2013) Airborne hyperspectral remote sensing to assess spatial distribution of water quality characteristics in large rivers: the Mississippi River and its tributaries in Minnesota. Remote Sens Environ 130:254–265
5. Murugesan M, Thilagamani S (2020) Efficient anomaly detection in surveillance videos based on multi layer perception recurrent neural network. J Microprocess Microsyst 79

6. Fauvel M, Tarabalka Y, Benediktsson JA, Chanussot J, Tilton JC (2013) Advances in spectral–spatial classification of hyperspectral images. Proc IEEE 101(3):1–25
7. Thilagamani S, Nandhakumar C (2020) Implementing green revolution for organic plant forming using KNN-classification technique. Int J Adv Sci Technol 29(7s):1707–1712
8. Ghamisi P, Mura MD, Benediktsson JA (2015) A survey on spectral–spatial classification techniques based on attribute profiles. IEEE Trans Geosci Remote Sens 53(5):2335–2353
9. Ghamisi P, Couceiro MS, Martins FM, Benediktsson JA (2014) Multilevel image segmentation approach for remote sensing images based on fractional-order Darwinian particle swarm optimization. IEEE Trans Geosci Remote Sens 52(5):2382–2394
10. Thilagamani S, Shanti N (2014) Gaussian and gabor filter approach for object segmentation. J Comput Inf Sci Eng 14(2):021006
11. Ghamisi P, Couceiro M, Fauvel M, Benediktsson JA (2014) Integration of segmentation techniques for classification of hyperspectral images. IEEE Geosci Remote Sens Lett 11(1):342–346
12. Zhu Z, Jia S, He S, Sun Y, Ji Z, Shen L (2015) Three-dimensional Gabor feature extraction for hyperspectral imagery classification using a memetic framework. Inf Sci 298:1–24
13. Perumal P, Suba S (2022) An analysis of a secure communication for healthcare system using wearable devices based on elliptic curve cryptography. J World Rev Sci Technol Sustain Dev 18(1):51–58
14. Fang L, Li S, Duan W, Ren J, Benediktsson JA (2015) Classification of hyperspectral images by exploiting spectral–spatial information of superpixel via multiple kernels. IEEE Trans Geosci Remote Sens 53(12):6663–6674
15. Fang L, Li S, Kang X, Benediktsson JA (2015) Spectral–spatial classification of hyperspectral images with a superpixel-based discriminative sparse model. IEEE Trans Geosci Remote Sens 53(8):4186–4201
16. Xue Z, Li J, Cheng L, Du P (2015) Spectral–spatial classification of hyperspectral data via morphological component analysis-based image separation. IEEE Trans Geosci Remote Sens 53(1):70–84
17. Pandiaraja P, Sharmila S (2020) Optimal routing path for heterogenous vehicular adhoc network. Int J Adv Sci Technol 29(7):1762–1771
18. Xu X, Li J, Huang X, Dalla Mura M, Plaza A (2016) Multiple morphological component analysis based decomposition for remote sensing image classification. IEEE Trans Geosci Remote Sens 54(5):3083–3102
19. Wang Q, Lin J, Yuan Y (2016) Salient band selection for hyperspectral image classification via manifold ranking. IEEE Trans Neural Netw Learn Syst 27(6):1279–1289
20. Hong D (2020) Graph convolutional networks for hyperspectral image classification. IEEE Trans Geosci Remote Sens 59(7):5966–5978
21. Li W (2017) Hyperspectral image classification using deep pixel-pair features. IEEE Trans Geosci Remote Sens 55(2):844–853
22. Xiong M (2015) Hyperspectral ımage classification using weighted joint collaborative representation. IEEE Geosci Remote Sens Lett 12(6):1209–1213
23. Gao L (2015) Subspace-based support vector machines for hyperspectral image classification. IEEE Geosci Remote Sens Lett 12(2):349–353
24. Wu C, Du B, Zhang L (2014) Slow feature analysis for change detection in multispectral imagery. IEEE Trans Geosci Remote Sens 52(5):2858–2874
25. Erturk A, Iordache MD, Plaza A (2015) Sparse unmixing-based change detection for multitemporalhyperspectral images. IEEE J Sel Topics Appl Earth Obs Remote Sens 9(2):708–719
26. Hang R (2020) Hyperspectral image classification with attention-aided CNNs. IEEE Trans Geosci Remote Sens 59(3):2281–2293
27. Cao X (2020) Hyperspectral image classification with convolutional neural network and active learning. IEEE Trans Geosci Remote Sens 58(7):4604–4616
28. Akey Sungheetha RSR (2021) Classification of remote sensing ımage scenes using double feature extraction hybrid deep learning approach. J Inf Technol 3(2):133–149

29. Adam EEB, Sathesh A (2021) Construction of accurate crack ıdentification on concrete structure using hybrid deep learning approach. J Innov Image Process (JIIP) 3(2):85–99
30. Mittal N, Gelbukh A (2021) Change detection in remote-sensed data by particle swarm optimized edge detection image segmentation technique. In: Innovative data communication technologies and application. Springer, Singapore, pp 809–817
31. Hu F, Xia GS, Hu J, Zhang L (2015) Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. Remote Sens 7(11):14680–14707
32. Wei F, Li S, Fang L, Kang X, Benediktsson JA (2016) Hyperspectral image classification via shape-adaptive joint sparse representation. IEEE J Sel Topics Appl Earth Obs Remote Sens 9(2):556–567
33. Pandiaraja P, Aravinthan K, Lakshmi Narayanan R, Kaaviya KS, Madumithra K (2020) Efficient cloud storage using data partition and time based access control with secure aes encryption technique. Int J Adv Sci Technol 29(7):1698–1706
34. Rajesh Kanna P, Santhi P (2021) Unified deep learning approach for efficient ıntrusion detection system using ıntegrated spatial–temporal features. Knowledge-Based Systems, p 226
35. Santhi P, Mahalakshmi G (2019) Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction. Int J Eng Adv Technol 8(4):839–846
36. Deepa K, Thilagamani S (2019) Segmentation techniques for overlapped latent fingerprint matching. Int J Innov Technol Explor Eng 8(12):1849–1852
37. Pradeep D, Sundar C, QAOC (2020) Novel query analysis and ontology-based clustering for data management in Hadoop, vol 108, pp 849–860
38. Zhang G, Jia X, Hu J (2015) Superpixel-based graphical model for remote sensing image mapping. IEEE Trans Geosci Remote Sens 53(11):5861–5871
39. Kang X, Li S, Benediktsson JA (2014) Feature extraction of hyperspectral images with image fusion and recursive filtering. IEEE Trans Geosci Remote Sens 52(6):3742–3752
40. Zhang L (2013) Tensor discriminative locality alignment for hyperspectral image spectral–spatial feature extraction. IEEE Trans Geosci Remote Sens 51(1):242–256
41. Kang X (2015) Intrinsic image decomposition for feature extraction of hyperspectral images. IEEE Trans Geosci Remote Sens 53(4):2241–2253
42. Li J et al (2015) Multiple feature learning for hyperspectral image classification. IEEE Trans Geosci Remote Sens 53(3):1592–1606
43. Jia X, Kuo B-C, Crawford MM (2013) Feature mining for hyperspectral image classification. Proc. IEEE 101(3):676–697
44. Xia J (2015) Spectral–spatial classification for hyperspectral data using rotation forests with local feature extraction and Markov random fields. IEEE Trans Geosci Remote Sens 53(5):2532–2546
45. Imani M, Ghassemian H (2017) Feature extraction of hyperspectral images using boundary semi-labeled samples and hybrid criterion. J AI Data Mining 5(1):39–53
46. Logeshwaran R, Aarthi P, Dineshkumar M, Lakshitha G, Vikram R (2019) Portable charger for handheld devices using radio frequency. Int J Innov Technol Explor Eng (IJITEE) 8(6)
47. Ren Y (2017) Hyperspectral image spectral-spatial feature extraction via tensor principal component analysis. IEEE Geosci Remote Sens Lett 14(9):1431–1435
48. Gunasekar M, Thilagamani S (2020) Performance analysis of ensemble feature selection method under SVM and BMNB classifiers for sentiment analysis. Int J Sci Technol Res 9(2):1536–1540
49. Imani M, Ghassemian H (2014) Principal component discriminant analysis for feature extraction and classification of hyperspectral images. In: 2014 Iranian Conference Intelligent Systems (ICIS)
50. Diwaker MK, Chaudhary PT (2016) A comparative performance analysis of feature extraction techniques for hyperspectral ımage classification
51. Deepika S, Pandiaraja P (2013) Ensuring CIA triad for user data using collaborative filtering mechanism. In: 2013 International conference on ınformation communication and embedded systems (ICICES), pp 925–928

52. Rajesh Kanna P, Santhi P (2022) Hybrid ıntrusion detection using map reduce based black widow optimized convolutional long short-term memory neural networks. Expert Syst Appl 194
53. Deepa K, Kokila M, Nandhini A, Pavethra A, Umadevi M (2020) Rainfall prediction using CNN. Int J Adv Sci Technol 29(7s):1623–1627

Design of Super Mario Game Using Finite State Machines

Anjana S. Nambiar, Kanigolla Likhita, K. V. S. Sri Pujya, and M. Supriya

Abstract A finite state machine, commonly known as a finite state automaton, is a model of computation based on a theoretical machine and is composed of one or more states. It enables the switch between multiple possible states and modification of the behaviour according to the state. One of the widely used applications of finite automata is in the video gaming industry. Game designing is the methodology of developing the content and rules of a game in the initial stages. It involves designing the characters, rules, environment, action, and rewards. For example, Super Mario is a game series based on the fictional plumber Mario created by Nintendo. The fundamental purpose of the game is to move the character, Mario and complete as many stages as possible. The secondary goal is killing the enemies and collecting coins or additional items. This project involves Mario transforming his states and behaviour based on the events that have occurred using finite automata. This paper deals with implementing the Super Mario Game Design using Python Programming Language and JFLAP, a widespread open-source software tool used in the Formal Language and Automata courses.

Keywords Finite state machines · Automata · Super Mario game · Java Formal Languages and Automata Package (JFLAP) · Game design · Abstract machines

A. S. Nambiar · K. Likhita · K. V. S. Sri Pujya (✉) · M. Supriya
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India
e-mail: bl.en.u4aie19035@bl.students.amrita.edu

A. S. Nambiar
e-mail: bl.en.u4aie19006@bl.students.amrita.edu

K. Likhita
e-mail: bl.en.u4aie19026@bl.students.amrita.edu

M. Supriya
e-mail: m_supriya@blr.amrita.edu

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_55

1 Introduction

The study of abstract computing devices, also known as abstract machines, is called automata theory. An automaton is any machine that converts information into different forms using a specific repeatable process.

In the 1930s, a great scientist and the father of computing commenced the study of abstract machines that had all the capabilities of a computer available today, as far as they could compute. Simpler machines were studied by researchers a decade later, between the 1940s and the 1950s. These simpler machines were known as "finite automata". Amongst the machines available today, the most straightforward machine is the finite automata. The finite automata had been proposed initially to model the brain function but are currently used in a wide range of applications [1].

In the 1950s, N. Chomsky commenced the study of formal grammars. As a result, formal grammars and automata have a very close relationship, and this forms the base of different applications in automata theory.

In addition to the finite automata, three more automata are widely employed. They are pushdown automata, linear-bounded automata, and Turing machines. The pushdown automaton is a type of automata with limited memory that can be accessed in a restricted way. It uses a stack data structure as storage memory. Pushdown automata can recognize context-free grammar. It can solve more problems than finite state automata. It is used to model parsers.

A linear bound automaton is a type of automata that utilizes restricted tape as a storage unit. They recognize context-sensitive grammar. Similarly, the Turing machine is a type of automata with infinite memory. It has no limitations and works using a tape memory. They recognize unrestricted grammar. They are very powerful and is used to model any computer.

The most efficient automaton is a Turing machine followed by linear bound automata and pushdown automata. On the contrary, finite state automata are the easiest to implement, followed by pushdown automata, linear bound automata, and Turing machines. The Turing machines are challenging to implement because they utilize a tape memory.

Some characteristics of automata are as follows -

- Inputs: In automata theory, inputs are the sequences of symbols selected from a finite set.
- Outputs: In automata theory, the outputs are the sequences of symbols that are selected from a finite set.
- States: In automata theory, state is the property of the machine at any given point in time.

The significance of automata theory is innumerable. Nowadays, computer scientists employ automata theory to compute functions and to understand the methodology of solving problems using machines. For example, it helps determine if a function is defined as computable and if a problem can be described as decidable. Automata is used for numerous applications like text processing, hardware designing,

compiler designing, and game theory. Furthermore, context-free grammar is used in artificial intelligence and programming languages [2].

In addition to applications like text processing, it is used in NLP, hardware designing, compiler designing and game theory, automata theory is used in distributed systems, software designing, real-time systems, and structured data [3]. These models and verifications are made possible by automata allowing them to model time and probabilities. In addition, automata theory can be advantageous to comprehend the limitations of computation from a theoretical viewpoint. Therefore, it is essential to learn automata theory to apprehend formal mathematical models of computation and their reflection in the real world of computer work.

2 Preliminaries

2.1 Finite State Machines (FSM)

Finite state machines, also known as finite state automatons, are a type of automation and a mathematical computation model. It is the most straightforward machine out of the automatons available. They have a finite amount of memory or no memory. They are employed to model "small" computers. They can solve a few problems but cannot solve problems that require memory. Finite state automata are the easiest to implement but are less efficient compared to the other automatons available.

In general, finite automata is used to recognize patterns, regular languages, and for mathematical problem analysis. They are ideal computation models for a small amount of memory and do not maintain a memory [4].

Finite state machines are restricted to certain states and the transitions between these states. The input of the automaton is a string of symbols and changes its state accordingly based on the inputs. The transition occurs on getting the desired input, and the state changes. In general, the automata can either stay in the same state or move to the next state at the time of transition.

Finite automata can be classified as an accepter and a transducer. It can have the following two possibilities:

- Accept
- Reject

A final state is a state that terminates the transitions. On processing the input, if the automata reach one of the final states, the input will be accepted. On the contrary, it will get rejected if it does not reach the final state [5].

The role of FSM in game theory is that the complex behaviours can be characterized and then they can be encapsulated into mini single interactions which are called as states using finite state machines. In finite state machines, each state can be expressed as direct action. Complex behaviour will result from the interactions of several single states.

2.2 Game Design

Game designing is the craftwork of applying design, innovation, and aesthetics to construct a game for amusement, experiment, or educational purposes. A finite state machine changes from one state to another in response to some inputs. It has many applications, and game theory is one such application [6]. Finite state machines have been employed in game designing. The states and transitions can be utilized to represent different actions, rewards, and the game's environment [7]. Applications of finite automata are numerous. In theory, finite automata is used in the initial stages of game development. But not much work has been done in relating this theory to practical game development. This game has been taken up in particular because it has been quite popular for being one of the oldest and successful gaming franchises, and this resulted in anyone born in the 80s and after could have grown up playing Mario games. This paper aims to look at how finite automata can be utilized to design the super Mario game [8].

2.3 Super Mario Game

Super Mario, also known as the Super Mario Bros and Mario series is a game sequel created and developed by Nintendo [9]. This game was founded and established starring the fictional plumber, Mario. In this game, the user will be playing the role of Mario. In the beginning, the user plays the game by passing through the Mushroom Kingdom. This is done by surviving and enduring the forces of the antagonist Bowser and saving and protecting Toadstool, the Princess.

In order to complete the level, the fictional plumber Mario should discover and reach the flagpole at the end of every level. In addition, there are numerous coins scattered and dispersed in the kingdom and unusual bricks marked with a question mark. When Mario hits the bricks, it might disclose and reveal more coins or something special to win the game. Along with the special bricks, the other bricks might also comprise some coins or rare items.

The environment also consists of super mushrooms, fire flowers, and superstars. If Mario consumes a super mushroom, he will evolve into Super Mario by growing double his size and achieving the capability to smash bricks above him. He will turn back to regular Mario on getting hit once again rather than dying.

Similarly, on eating fire flower, Mario turns into Fire Mario, where he acquires the power to throw bouncing fireballs using his hands, and on touching a bouncing Super Star, which is a rare power-up, Mario transforms into Star Mario. In this form, Mario is defensible and unassailable to almost every physical hindrance in the game. In addition, he can slaughter all the opponents with a straightforward touch. This is a provisional state, and henceforth, the player would have to quickly utilize the recently acquired abilities. Figures 1, 2 and 3 depict the screen shots from the game.

Fig. 1 Screenshot of super Mario game's environment

Fig. 2 Screenshot of super Mario game's fire flower

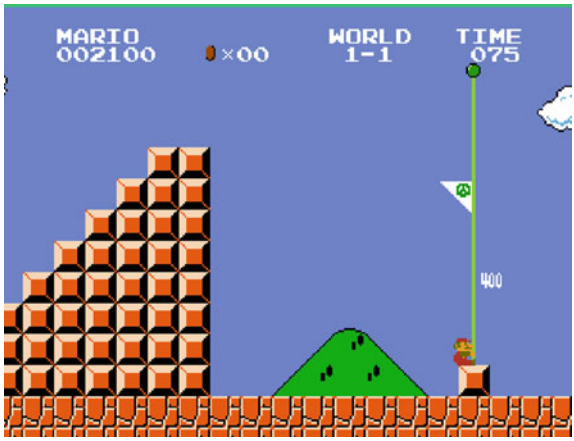


Fig. 3 Screenshot of super Mario game's level completion

3 Implementation

The implementation of Super Mario game design is designed in Python using pyformlang, a python library to manipulate formal grammar and JFLAP. Pyformlang.

Pyformlang is an open-source Python3 library specialized in formal language manipulation. It is a practical and pedagogical Python library designed for formal languages. It comprises the preliminary functionalities to manipulate and understand finite state automata, regular expressions, finite state transducers, pushdown automata, context-free grammars and indexed grammars [10]. All the states and the transitions can be implemented using Pyformlang.

3.1 JFLAP

JFLAP is used in the design of finite state machines. JFLAP is open-source software for formal language subjects like deterministic finite automata (DFA) and non-deterministic finite automata (NFA). DFA refers to a definite finite automaton. Determinism refers to uniqueness. A deterministic finite automaton reads a string of symbols simultaneously as a finite state machine using JFLAP. Each input in the deterministic finite automata has just one path from one state to the next state. NFA stands for non-deterministic finite state machine. A finite state machine is called an NFA if there is more than a single path for a particular input from the current state to the next and can be implemented using JFLAP. Not all non-deterministic finite automata are deterministic finite automata, but all non-deterministic finite automata can be converted to deterministic finite automata using JFLAP [11].

3.2 Game States

The states in the game along with the description of the player is given in Table 1. Table 2 showcases the transitions and the possible score (Fig. 4).

3.3 Game Level Description

This section describes the game which is presented as 2 levels.

Level 1:

Some of the crucial states in the level are:

start game, level 1, small Mario, super Mario, fire Mario, score, game over, exit, game, second life, third life, flag

Table 1 Description of game states

Game states	Description
Start game (initial state)	The player starts the game
Small Mario	Character in the game
Super Mario	Character in the game
Fire Mario	Character in the game
Score	When player collects the coins score is displayed
Flag (final state)	Flag represents the end of each level
Game over	If player died in the game
Exit game (final state)	Player can exit the game at any point of time

Table 2 Transition table with score

Present state	Transitions (events)	New state	Score
small_mario	got_mushroom	super_mario	500
small_mario	got_fire_flower	fire_mario	500
super_mario	got_mushroom	super_mario	500
super_mario	got_fire_flower	fire_mario	500
fire_mario	got_mushroom	fire_mario	500
fire_mario	got_fire_flower	fire_mario	500
small_mario	kill_monster	small_mario	100
super_mario	kill_monster	super_mario	100
fire_ mario	kill_monster	fire_mario	100
super_mario	met_monster	small_mario	0
fire_mario	met_monster	small_mario	0
small_mario	met_monster	second_life	0
small_mario	met_monster	third_life	0
small_mario	met_monster	game_over	0

Some of the important transitions are:

got_mushroom, got_fire_flower, kill_monster, 100, 200, 500, met_monster, reached_flag

At the beginning of the game, the player starts with three lives and as small Mario. If the player captures any coin in the game, the player will be awarded 200 coins. If small Mario captures the mushroom (got_mushroom) then the state is changed to Super Mario, and the player will be awarded a 500 score. Similarly, if Super Mario captures the fire flower (got_fire_flower), the state is changed to Fire Mario, and the player will be awarded a 500 score. If Mario in any state kills the monsters, the player will be awarded 100 as a reward. If the Super Mario or Fire Mario will meet the monster, then the state is again changed to small Mario, and if the small Mario

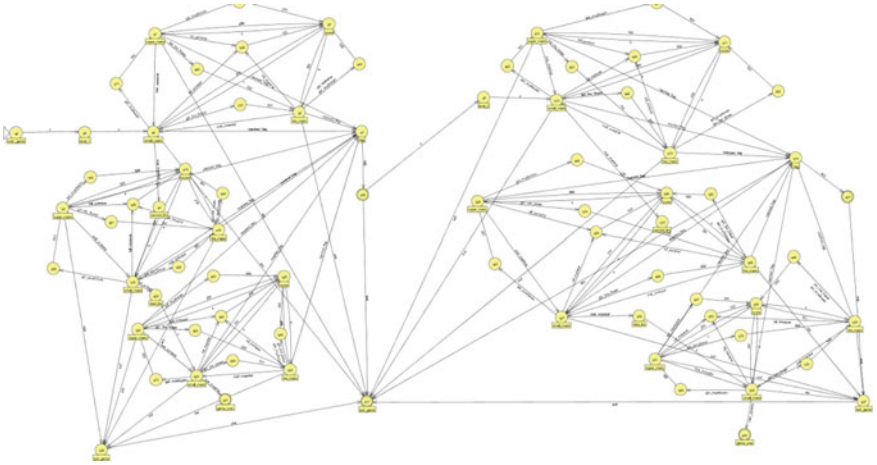


Fig. 4 JFLAP implementation of super Mario game

meets a monster, the player loses one life is lost, and the game comes to an end when all the lives are lost. The flag at the end represents that the player the completed first level.

In python, the level implantation is done by importing the essential libraries and modules. The pyformlang library is used following which begins the creation of an NFA object using the EpsilonNFA(). Different states of the automaton are created using the state(). The start state and the end state of the automaton are created using the add_start_state() and the add_final_state(). The transitions are added using the add_transition() (Figs. 5 and 6).

Similarly, a similar approach is followed to create the automaton using JFLAP. All the states along with the start and the final states are created following which the transitions are added. The automaton can be visualized by the user using JFLAP [12], but not using pyformlang package in Python Programming Language.

```
from pyformlang.finite_automaton import EpsilonNFA
from pyformlang.finite_automaton import Epsilon
from pyformlang.finite_automaton import Symbol
from pyformlang.finite_automaton import State

supermariogameFA = EpsilonNFA()

epsilon = Epsilon()
```

Fig. 5 Implementation of first level super Mario game in python

```
small_mario = State(0)
super_mario = State(1)
fire_mario = State(2)
super_mario = State(3)
second_life = State(4)
score = State(5)
start_game = State(6)
flag = State(7)
level_2 = State(8)
level_1 = State(9)
super_mario = State(10)
score = State(11)
small_mario = State(12)
fire_mario = State(13)
flag = State(14)
second_life = State(15)
small_mario = State(16)
exit = State(17)
```

Fig. 6 Creation of states in the first level super Mario game in python

The implementation of the first level can be represented using a transition diagram and a transition table, of which a tabular representation of the transition function is presented in Table 3. The table lists only few of the important moves. Addition moves are very minimal or none.

Level 2:

The player can move to the next level on reaching the flagpole. The second level two is almost identical to the first level. When the player reaches the flag in the second level, the player gets a score of 400 as a reward. The flag at the end represents that the player the completed second level. The transition table of level 2 is presented in Table 4. As mentioned in the previous section, the table presents only the important transitions.

Remaining Levels:

The remaining levels of the game can be implemented using the same approach.

4 Results

The Super Mario game is successfully designed and implemented using Finite State Automata in Python Programming Language using Pyformlang and JFLAP. A snapshot on the Pyformlang implementation is presented in Fig. 7. Figures 8 and 9 present

Table 3 Transition table of first level super Mario game

State	got_mushroom	got_fire_flower	kill_monster	met_monster	100	200	400	500	**reached_flag**	**exit**	**λ**
q0(small_mario)	{q71}	{q72}	{q38}	{q4}	∅	{q5}	∅	∅	{q7}	{q17}	∅
q1(super_mario)	{q41}	{q60}	{q38}	{q0}	∅	{q5}	∅	∅	{q7}	{q17}	∅
q2(fire_mario)	{q44}	{q44}	{q38}	{q0}	∅	{q5}	∅	∅	{q7}	{q17}	∅
q3(super_mario)	{q42}	{q61}	{q39}	{q16}	∅	{q19}	∅	∅	{q7}	{q36}	∅
q16(small_mario)	{q58}	{q59}	{q39}	{q20}	∅	{q19}	∅	∅	{q7}	{q36}	∅
q18(fire_mario)	{q45}	{q45}	{q39}	{q16}	∅	{q19}	∅	∅	{q7}	{q36}	∅
q22(super_mario)	{q43}	{q62}	{q40}	{q23}	∅	{q25}	∅	∅	{q7}	{q36}	∅
q23(small_mario)	{q73}	{q64}	{q40}	{q21}	∅	{q25}	∅	∅	{q7}	{q36}	∅
q24(fire_mario)	{q46}	{q46}	{q40}	{q23}	∅	{q25}	∅	∅	{q7}	{q36}	{q25}

Table 4 Transition table of second level super Mario game

State	got_mushroom	got_fire_flower	kill_monster	met_monster	100	200	400	500	reached_flag	exit	λ
q10(super mario)	{q51}	{q63}	{q55}	{q12}	ø	{q11}	ø	ø	{q14}	{q17}	ø
q12(small mario)	{q65}	{q66}	{q55}	{q15}	ø	{q11}	ø	ø	{q14}	{q17}	ø
q13(fire mario)	{q52}	{q52}	{q55}	{q12}	ø	{q11}	ø	ø	{q14}	{q17}	ø
q26(super_mario)	{q49}	{q74}	{q54}	{q27}	ø	{q28}	ø	ø	{q14}	{q17}	ø
q27(small_mario)	{q67}	{q68}	{q54}	{q30}	ø	{q28}	ø	ø	{q14}	{q17}	ø
q29(fire_mario)	{q50}	{q50}	{q54}	{q27}	ø	{q28}	ø	ø	{q14}	{q37}	ø
q31(super_mario)	{q47}	{q75}	{q53}	{q32}	ø	{q34}	ø	ø	{q14}	{q37}	ø
q32(small_mario)	{q69}	{q70}	{q53}	{q35}	ø	{q34}	ø	ø	{q14}	{q37}	ø
q33(fire_mario)	{q48}	{q48}	{q53}	{q32}	ø	{q34}	ø	ø	{q14}	{q37}	ø

```
supermariogameFA.add_start_state(small_mario)
supermariogameFA.add_final_state(super_mario)
supermariogameFA.add_transition(small_mario, "got_fire_flower", super_mario)
supermariogameFA.add_transition(super_mario, 500, fire_mario)
supermariogameFA.add_transition(fire_mario, 200, super_mario)
```

Fig. 7 Adding transitions in the first level super Mario game in python

Fig. 8 Testing of the automata using multiple inputs

Input	Result
got_mushroom500200reached_flag	Reject
got_fire_flower500200reached_flag400	Accept
met_monstermet_monstermet_monster	Accept
got_fire_flower200reached_flag	Reject
got_mushroom500reached_flag400got_mushroom500200reached_flag400	Accept
got_fire_flower500200200reached_flag400got_mushroom500reached_flag400	Accept
got_mushroom200500got_fire_flower500200reached_flag	Reject
kill_monster100reached_flag400	Accept
	Reject

Fig. 9 Testing of the automata using JFLAP traceback

the outputs from JFLAP.

After modelling the automata using JFLAP by giving different inputs, we get either Accept or Reject, where Accept means input has been accepted by the automata and Reject means that the input has been rejected (Fig. 10).

In conclusion, this paper has experimented the use of finite state machines in game design. Though FSMs are used in string matching, plagiarism detection, word extraction, activity representation, etc. [13–18], this paper proposes FSM for game designing. This implementation enables to simply the complex behaviours into finite set of states. If the number of states is too complex, behaviour trees can be used which is used in artificial intelligence field.

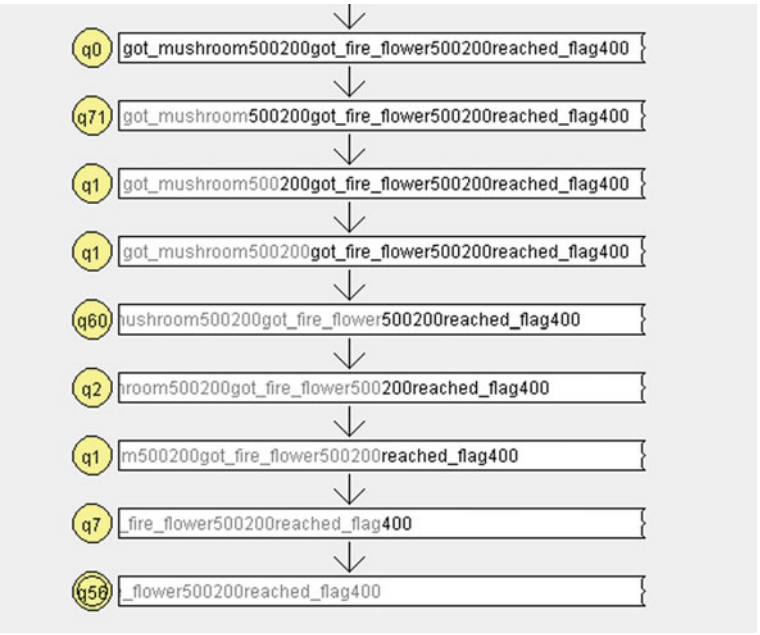


Fig. 10 Traceback analysis of sample input

5 Conclusion

This paper deals with the implementation of the first two levels of the Super Mario game design using a finite automaton using game theory. This implementation shows the state changes of a character. The more the number of states, the more are the machines are capable of handling complex strategies. This analysis may additionally be extended to incorporate advanced computational and complexity theory tools. This can be done to additionally improve the game theory to an advanced level by utilizing Pushdown automata, Turing Machines and Moore and Mealy Machines. Finite automata are widely used to represent simple characters in the game. Therefore, it is easy to present a design that allows developers to see the broad picture and fine-tune the result.

References

1. Mushtaq H, Qureshi N, Aslam M, Ahsan M, Ali M, Atta M (2012) Computing game design with automata theory, vol 3, pp 13–21
2. Khoussainov B, Nerode A (2001) Automata theory and its applications. https://doi.org/10.1007/978-1-4612-0171-7

3. Knight K, May J (2009) Applications of weighted automata in natural language processing. https://doi.org/10.1007/978-3-642-01492-5_14
4. Lee CY (1960) Automata and finite automata. Bell Syst Tech J 39(5):1267–1295. https://doi.org/10.1002/j.1538-7305.1960.tb03960.x
5. Perles M, Rabin MO, Shamir E (1963) The theory of definite automata. IEEE Trans Electr Comput EC-12(3):233–243. https://doi.org/10.1109/PGEC.1963.263534
6. Ali KF, Kalyan V, Kumar KA (2019) Design and implementation of ludo game using automata theory. In: 2019 innovations in power and advanced computing technologies (i-PACT), pp 1–6. https://doi.org/10.1109/i-PACT44901.2019.8959998
7. Adegun A, Ogundokun R, Ogbonyomi S, Sadiku P (2020) Design and implementation of an intelligent gaming agent using A* algorithm and finite state machines. Int J Eng Res Technol 13:191. https://doi.org/10.37624/IJERT/13.2.2020.191-206
8. Jalmanasra S, Suwais K, Mohd Arshad M (2013) The applications of automata in game theory. https://doi.org/10.4018/978-1-4666-4038-2.ch011
9. Mora A, Merelo Guervós J, García-Sánchez P, Castillo P, Rodríguez-Domingo M, Hidalgo-Bermúdez R (2014) Creating autonomous agents for playing Super Mario Bros game by means of evolutionary finite state machines. Evol Intell 6:1–14.https://doi.org/10.1007/s12065-014-0105-7
10. Romero J (2021) Pyformlang: an educational library for formal language manipulation. https://doi.org/10.1145/3408877.3432464
11. Rodger S, Finley T (2006) JFLAP—an interactive formal languages and automata package. ISBN 0763738344. Jones and Bartlett
12. Wang D, Chen Y (2021) Visual teaching design of compile principle based on JFLAP. In: 2nd international conference on artificial intelligence and education (ICAIE), pp 366–370. https://doi.org/10.1109/ICAIE53562.2021.00083
13. Joseph S, Jeevitha KP (2015) An automata based approach for the prevention of NoSQL injections. In: Security in computing and communications (link is external), Springer—communications in computer and information science (link is external), pp 538–546
14. Ramesh G, Menen A (2020) Automated dynamic approach for detecting ransomware using finite-state machine. Decis Support Syst
15. Jevitha KP, Swaminathan J, Jayaraman B (2021) Finite-state model extraction and visualization from Java program execution. Softw Pract Exp 51:pp 409–437
16. Kanjirangat V, Gupta D (2018) Unmasking text plagiarism using syntactic-semantic based natural language processing techniques: Comparisons, analysis and challenges
17. Tamizharasan M, Shahana RS, Subathra P (2020) Topic modeling-based approach for word prediction using automata. J Crit Rev 744–749
18. Kavya J, Geetha M (2016) An FSM based methodology for interleaved and concurrent activity recognition. Paper presented at the 2016 international conference on advances in computing, communications and informatics, ICACCI 2016, pp 994–999

Training Logistic Regression Model by Enhanced Moth Flame Optimizer for Spam Email Classification

Mohamed Salb, Luka Jovanovic, Miodrag Zivkovic, Eva Tuba, Ali Elsadai, and Nebojsa Bacanin

Abstract Spam email is a massive issue that bothers and consumes receivers' time and effort. Because of its effectiveness in identifying mail as wanted or unwanted, machine learning approaches have become a popular technique in spam detection. Current spam detection methods, on the other side, typically have low detection performance and are incapable of handling high-dimensional information easily. As a result, a unique spam detection approach that combines an improved moth flame optimization algorithm and a logistic regression classification model was proposed in this paper. The research evidence on two accessible datasets (CSDMC2010, Enron) indicates that the suggested methodology can tackle high-dimensional data due to its very powerful local and global search skills. The suggested technique was evaluated for spam detection accuracy to that of logistic regression, naive Bayes classifiers, and support vector machine, as well as the performance of earlier research' that includes state-of-the-art approaches. In terms of classification performance, the suggested methodology outperforms the other spam detection algorithms examined in this work.

Keywords Logistic regression · Machine learning · Moth flame optimization · Optimization · Spam filtering · Swarm intelligence

M. Salb · L. Jovanovic · M. Zivkovic · E. Tuba · A. Elsadai · N. Bacanin (✉)
Singidunum University, Danijelova 32, 11000 Belgrade, Serbia
e-mail: nbacanin@singidunum.ac.rs

M. Salb
e-mail: mohamed.salb.16@singimail.rs

L. Jovanovic
e-mail: luka.jovanovic.191@singidunum.ac.rs

M. Zivkovic
e-mail: mzivkovic@singidunum.ac.rs

E. Tuba
e-mail: etuba@ieee.org

A. Elsadai
e-mail: aelsadai@singidunum.ac.rs

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_56

1 Introduction

The growing number of e-commerce businesses has led to a rise in the quantity of commercial ads and emails. On a regular basis, consumers receive emails from unknown senders attempting to steal private information, passwords, or bank account details. Spam refers to unwelcome commercial emails that have been sent in a random fashion.

Radicati stated that [35] the number of expected sent and received emails will approach 347 billion by the end of 2023. According to Ma et al. [30], spam emails are made up 55% of global email traffic in 2017 and 2018. Total spam levels in 2015 were at 52.7%, in 2016 at 53.4%, and in 2017 at 54.6%. Spam emails annoy receivers, consume network bandwidth, storage space, and computer resources, while potentially damaging files by inserting virus items in attached files.

Various strategies for preventing spam have been proposed. Bhowmick and Hazarika [21] classified spam email mitigation strategies into eight categories: machine learning-based techniques, signature schemes, whitelisting, blacklisting, heuristic filters, collaborative spam filtering, challenge-response systems, and honey pots. A further classification splits the approaches into two major categories: dynamic techniques and static techniques.

Filtration techniques based on specified blacklists, and whitelists are examples of static techniques. However, these prove unsuccessful, as spammers utilize multiple email addresses. Additionally people with compromised devices could transmit spam while ignorant of the virus. To determine if an email is spam, dynamic techniques often evaluate its message using text classification built with a statistical methodology or machine learning (ML) algorithms [29, 37].

Throughout this paper, a novel spam filtering technique that benefits from logistic regression (LR) was proposed, featuring quick classification in real-world applications and effective methods for avoiding convergence local minima by applying the moth flame optimization algorithm (MFO) [32]. As a nature-inspired swarm intelligence metaheuristic algorithm, the MFO imitates moth flying behavior and attitude. It should be highlighted that this work presents the first spam filter that employs the MFO as an LR training algorithm. The MFO method works very well with high-dimensional challenges [39], as it has to balance between intensification and diversification, making it a strong choice for spam detection. Using incorporation with the term frequency and inverse document frequency (tf and idf) approach, the suggested method was improved and employed with LR to minimize additional restriction for this classifier, linked with its vulnerability to identical feature weights. Accordingly, the main contributions of this paper are twofold:

- After observing the drawbacks of the basic MFO implementation, a novel, modified variant of the algorithm is proposed that overcomes the detected issues.
- The devised improved MFO algorithm is applied to the spam email classification problem, to further improve the classifier accuracy.

Comparison between the classification performance of proposed method and Gaussian Naive Bayes (NB), multinomial NB, linear support vector machine (SVM),

radial basis SVM and logistic regression (LR) classifiers is performed. On the publically available datasets CSDMC2010, experimental findings reveal that AVMFO-LR outperforms the other approaches. The Enron dataset was utilized to evaluate AVMFO-LR performance to that of earlier approaches, which show that AVMFO-LR methodology surpasses state-of-the-art techniques that have been evaluated on the Enron dataset in the overall classification and false-negative (FN) rate.

The rest of this paper is structured as follows. Section 2 covers logistic regression and provides an overview of works on spam email detection. Section 3 presents the proposed improved approach. Section 4 provides a discussion of the datasets used in experiments, as well as the attained results. Finally, Sect. 5 wraps up the article and suggests some further work.

2 Background and Related Work

Machine learning (ML) is a practical and effective approach of discovering optimized solutions to issues that have historically been solved via time-consuming problem-solving methodologies. Using metaheuristic ML algorithms has assisted in the resolution of several challenges in the field, where traditional methods had failed. Statistics is one of the domains in which machine learning is integrated and has been employed. Classification utilizing multiple methodologies is one of the areas that machine learning may be utilized in this field. Logistic regression (LR), often known as the sigmoid function, is a machine learning method for binary classification and is one of the most well-known classification techniques. This is due to the fact that it is a basic algorithm that works admirably on a wide range of issues. It is an S-shaped curve that can transfer any real-valued integer to a value between 0 and 1, but still never precisely among these boundaries.

Considering a set of training data $\{(\vec{x}_1, y_1), (\vec{x}_2, y_2), \ldots, (\vec{x}_m, y_m)\}$, $\vec{x}_i \in R^n$ and $y_i \in \{0, 1\}$, $1 \leq i \leq m$, Eq. (1) specifies the class of $\vec{x}_i$ in LR.

$$y_i = \begin{cases} 0, & p_i < 0.5 \\ 1, & p_i \geq 0.5 \end{cases} \tag{1}$$

In which p_i is determined by Eq. (2) and is the sigmoid function given by Eq. (3):

$$p_i = \sigma(\vec{w}\vec{x}_i) \tag{2}$$

$$\sigma(a) = \frac{1}{1 + \mathrm{e}^{-a}} \tag{3}$$

The LR method's purpose is to acquire weights ($\vec{w}$) that reduce a cost function, such as a cross-entropy method using ridge regression, thereby reducing computational burden and preventing overfitting. Equation (4) gives the cross-entropy func-

tion using ridge regression.

$$J(\vec{w}) = -\sum_{i=1}^{m} y_i \log(p_i) + (1 + y_i) \log(1 - p_i) + \frac{\lambda}{2} \sum_{j=1}^{n} w_j^2 \tag{4}$$

where λ is a factor for regularization. The gradient descent technique begins with certain starting $\vec{w}$ and iteratively updates Eq. (5) until the terminating requirements are met.

$$\vec{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}_{n\times 1} + \alpha \begin{bmatrix} -\vec{x}_1- \\ -\vec{x}_2- \\ \vdots \\ -\vec{x}_m- \end{bmatrix}^{\mathrm{T}}_{m\times n} \times \left(\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} - \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{bmatrix} \right)_{m\times 1} - \lambda \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}_{n\times 1} \tag{5}$$

where α is the learning rate, influencing converges speed to local optima. Poor converge would occur when α is too low. However α if is too high, gradient descent would miss the local optimum. It should be highlighted that α and λ are essential hyperparameters of LR architectures. Moreover, the gradient descent methods are mainly used to train machine learning models, but they can fall into the local optimum, and therefore, swarm intelligence is applied to overcome such a problem.

In the field of computer science, nature-inspired metaheuristic optimization algorithms and swarm intelligence algorithms had already been used to solve NP-hard challenges, which includes MRI and histological frame classifier optimization [17, 19], predicting the cases of COVID-19 [47, 49], artificial neural network optimization and hyperparameters optimization [4, 7–10, 13, 15, 18, 24, 26, 40, 52], cloud-based framework for task scheduling issues [6, 14, 20, 51], wireless sensor networks [5, 11, 48, 50, 53], malware detection [46], and so on.

Swarm intelligence algorithms have been used with machine learning models to overcome the problem of spam filtering. Naem and his colleagues [34] suggested ALOBoosting, which split into two major steps: An ant lion optimizer (ALO) had been utilized during the first step to acquire the optimal subset of features. Throughout the second step, a boosting classifier had been employed to identify if the email is a spam email or a normal email. In regard to the categorization precision, this strategy was highly successful.

Idris and Selamat [28] presented a method that incorporates a negative selection algorithm (NSA), one of the several essential methods in the artificial immune system (AIS), with particle swarm optimization (PSO), and proved that their suggested framework detected spam emails more effectively compared to the NSA and PSO methods alone.

Dedeturk and Akay [25] proposed a novel spam detection approach combining the artificial bee colony algorithm and a logistic regression classification algorithm. The proposed model was tested on three different datasets, including a Turkish email dataset, and the results show the proposed model can handle high-dimensional data.

After comparing the proposed model spam detection performance to other models tested on the same datasets, the results indicate the superiority of the proposed model in terms of classification accuracy.

3 Proposed Method

In this section, the original and the improved version of the moth flame algorithm is presented.

3.1 Original Moth Flame Algorithm

Moths are a species closely related to butterflies that are developed to fly at night by utilizing moonlight for navigation using a system known as transverse orientation. A moth flies by keeping a steady angle regarding the moon, which is a very efficient way for traveling vast distances in a straight line. However, artificial lights confuse moths, mainly to the inefficiencies of the transverse orientation mechanism. When moths detect a man-made light, they strive to fly in a straight path by maintaining a consistent angle, often resulting in a pointless or lethal spiral flight route. This behavior is mathematically represented to create the MFO approach [32].

Solutions in MFO are represented by moths, and the factors in the equation are their positions. As a result, the moths may fly in one-dimensional or multi-dimensional space by modifying positional variables. Because the MFO method is a population-based approach, it is displayed in the matrix in Eq. (6)

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,d} \\ m_{2,1} & m_{2,2} & \dots & m_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ m_{n,1} & m_{n,2} & \dots & m_{n,d} \end{bmatrix} \tag{6}$$

where d is the number of dimensions and n is the number of moths in the search space.

There is also an array for recording the relevant objective value, as shown in Eq. (7):

$$\mathrm{OM} = \begin{bmatrix} \mathrm{OM}_1 \\ \mathrm{OM}_2 \\ \vdots \\ \mathrm{OM}_n \end{bmatrix} \tag{7}$$

where the number of moth in the search space is n.

Every moth's position vector from M matrix is given to the fitness function, and the fitness function's return is allocated to the associated moth as its fitness value in the OM matrix.

Flame is also a control parameter in the MFO algorithm. And it can be represented as in Eq. (8):

$$F = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,d} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ f_{n,1} & f_{n,2} & \cdots & f_{n,d} \end{bmatrix} \tag{8}$$

where d is the number of diminutions and n is the number of moths in the search space.

There is also an array for recording the relevant objective value, demonstrated by Eq. (9):

$$\mathrm{OF} = \begin{bmatrix} \mathrm{OF}_1 \\ \mathrm{OF}_2 \\ \vdots \\ \mathrm{OF}_n \end{bmatrix} \tag{9}$$

with the number of moths in the search space represented by n.

It must be highlighted that both moths and flames are possible solutions. The distinction is in how to handle and modify them on every cycle. Moths are the main search workers that travel around the solution area, while flames represent the optimum location of moths that have so far been obtained. Using this approach will guarantee that the moth's finest solution will never be lost.

The MFO model is given, as a three tuple as in Eq. (10), that corresponds to the global optimum of optimization problems:

$$\mathrm{MFO} = (I, P, T) \tag{10}$$

where T is the function that will result in true or false depending on the termination criterion. P is the function that moves the moth around the search space using the M as an input. I is the function that generates random moths and corresponding fitness values.

The random population function (I) can be represented as Eq. (11):

$$M(i, j) = (\mathrm{ub}(i) - \mathrm{lb}(i)) \times \mathrm{rand}() + \mathrm{lb}(i) \tag{11}$$

where ub is the upper bound and lb is the lower bound.

After generating the random population, the P function is used to move the moth and explore the search space. It can be represented as follows:

$$M_i = S(M_i, F_j) \tag{12}$$

where S is the spiral function, M_i is the ith moth, and F_j is the jth flame.

The spiral can be represented mathematically as follows:

$$S(M_i, F_j) = D_i \times e^{bt} \times \cos(2\pi t) + F_j \tag{13}$$

where t is randomly generated number between (-1 and 1), it specifies how near the moth's following location will be to the flame ($t = -1$ indicates the nearest location to the flame, whereas $t = 1$ indicates the furthest), b is a constant number to help with shaping the spiral, and D_i represents the space between the ith moth and the jth flame. And D can be defined as follows:

$$D_i = |F_j - M_i| \tag{14}$$

Because of the spiral formula, a moth can fly around a flame rather than just in the area between them. As a result, the intensification and diversification of the solution area can be verified.

To highlight intensification, it is presumed that t is a randomly generated variable between (r and 1), with r decreasing linearly from 1 to 2 throughout the length of cycle. It is worth noting that r is referred to as the convergence factor. Moths use this strategy to exploit their associated flames more precisely proportionate to the amount of cycles. To avoid being trapped in the local minimum in Eq. (12), each moth updates its location with correspondence to only one flame. The first moth adjusts its location in relation to the greatest flame in the array, while the final moth adjusts its location in relation to the lowest flame in the array.

An additional issue to consider is that the location update of moths with n various places in the solution area might limit the diversification of the most viable solution. To address this issue, an adjustable system is developed that decrees the number of flames during the duration of the iterations. This method is shown as follows:

$$\text{FlameNum} = \text{round}\left(N - I \times \frac{N-1}{T}\right) \tag{15}$$

whereby N is the maximum number of flames represented, I is the current number of iteration, and T is the maximum number of iteration.

3.2 Improved Moth Flame Optimization

By analyzing original MFO algorithm, certain deficiencies were noticed. Specifically, in early phases of execution in some runs, due to the stochastics nature, it occurs that the right region of the search space is not found, and consequently in later iterations, the search process is unable to converge around the optimum parts of the search region. Also, it is noticed that the diversity of population is not on the satisfactory level throughout the whole run, leading consequently to just minor solu-

tions' improvements over subsequent iterations. On the other hand, by conducting rigid empirical analysis, it is concluded that the basic MFO's search expressions are highly efficient, yet its potential is not fully employed due to the abovementioned drawbacks.

To address these deficiencies and to better exploit potential of MFO, the idea of enhanced search based on the average fitness (AV) is introduced. After regular solutions' update by the basic MFO search mechanism, the average fitness (fit_{avg}) of the population is calculated. Afterward, random solution $M_r w$ whose fitness is below the average ($\text{fit}_r < \text{fit}_{\text{avg}}$) is selected and discarded from population. Such solution is then replaced by the offspring (M_o) individual generated by applying uniform crossover operator between current best solution (M_{best}) and randomly chosen individual M_{rb} whose fitness is greater than the average ($\text{fit}_r > \text{fit}_{\text{avg}}$). In this implementation, simple fitness representation, where fitness is reverse proportional in the case of minimization problems and directly proportional to the objective when maximization problems are concerned, is utilized. Simple mutation operator with fixed mutation probability of 0.001 (p_m) is then applied to each parameter on such generated solution as in [12].

Proposed method is named average fitness MFO (AVMFO), and according to experimental findings, it exhibits better convergence, as well as exploration than the basic MFO approach. The AVMFO pseudo-code is shown in algorithm 1.

Algorithm 1 Pseudo-code for the AVMFO

initialization. Produce the N random moths M inside solution area bounds using Eq. (11), and measure their fitness.
initialize T the max numbers of iteration.
update the number of flames using Eq. (15)
do
for all M of the moth produced **do**
  for all parameters **do**
    update r and t.
    calculate D using Eq. (14) with respect to the corresponding moth.
    update $M(i, j)$ using Eq. (12), and Eq. (13) with respect to the corresponding moth
  end for
  calculate objective and fitness values
  update flames
end for
Calculate average fitness
Discard M_{rw} individual from population
Perform uniform crossover between M_{best} and M_{rb} and generate M_o
Apply mutation operator and introduce to M_o into population
while($t < T$)

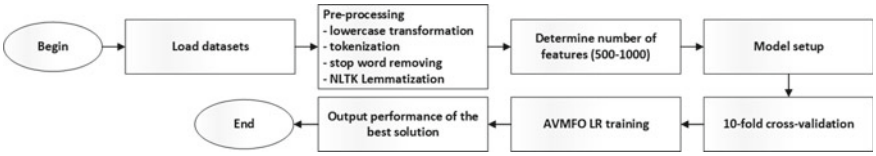


Fig. 1 Block diagram of the AVMFO-LR classifier

4 Experimental Setup, Analysis, and Discussion

In this section, an overview of the experimental setup for evaluating the proposed AVMFO for LR training (AVMFO-LR), along with comparative analysis with other well-known ML approaches, is given. The complete workflow of the proposed AVMFO-LR classifier is shown in Fig. 1.

4.1 Dataset

Datasets used for validation include the CSDMC2010 dataset consisted of 4327 English emails, among which 2949 (68.15%) are valid emails, and 1378 (31.85%) are spam emails, as well as the Enron-1 dataset [31] comprised of 5172 emails, 3672 valid (71%), and 1500 spam (29%). The CSDMC2010 and Enron datasets each have (82,148) and (47,939) unique words.

4.2 Preprocessing for Data Representation

In this classification problem, a series of emails $D = \{d_1, d_2, d_3, \ldots, d_m\}$ were given and two categories $C = \left\{c_{\text{spam}}, c_{\text{valid}}\right\}$. The difficulty is that you need to allocate an email address $d_i \in D$ to one of the two categories.

When handling emails, the content and subject are collected, and the HTML elements are eliminated. Words are split into sets of substrings, all characters changed to lowercase and tokenized. With the exception of exclamation marks, stop words and punctuation are eliminated. The NLTK stemming package [22] is used on the dataset to eliminate morphological prefixes and suffixes from phrases and keep just the phrase stem. Non-informative frequent stop words are removed, and the bag of words (BOW) paradigm is produced regardless of word grammar or position. The BOW paradigm is a syntax (word collection) that contains every one of the unique terms throughout all documents. As shown in Eq. (16)

$$B = \{t_1, t_2, t_3, \ldots, t_B\} \tag{16}$$

where B is the term set and $|B|$ is the total number of unique terms gathered in all documents. Following the creation of the word collection, every document $d_i \in D$ is defined by a vector $d_i = [w_1, w_2, w_3, \ldots, w_{|B|}]$, in which $w_1, w_2, w_3, \ldots, w_{|B|}$ are the weights corresponding to the terms $t_1, t_2, t_3, \ldots, t_{|B|}$.

The weights are calculated with tf and idf and can be calculated with Eq. (17)

$$\text{tf}-\text{idf}(t_j, d_i) \leq w_j, d_i \geq \text{tf}(t_j, d_i) \times \log \frac{|D|}{|D_{tj}|} \tag{17}$$

when $\text{tf}(t_j, d_i)$ indicates the overall number of words t_j in document d_i, $|D|$ represents the overall number of documents, and $|D_{tj}|$ indicates the overall number of documents containing term t_j.

$$\phi(\vec{d}_i', c_j) = \begin{cases} 1, & \text{if } \vec{d}_i' \in c_j \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

4.3 Results and Discussion

The classification performance of Gaussian NB, multinomial NB, linear SVM, radial basis SVM, LR, and the proposed model on the CSDMC2010 dataset was evaluated via comparing, in similar way as it was proposed in [25]. It is important to mention that the authors have implemented independently standard machine learning methods, ML and ABC-LR as well, to verify the results from [25]. Moreover, in this research, the termination condition for all tested methods was set according to the total amount of fitness function evaluations (FFEs) used, and to provide similar conditions as in [25], the number of evaluations was set to 160,000 for all conducted experiments.

Furthermore, the classification accuracy of the model on the Enron dataset was evaluated by comparing it to other techniques from earlier research, and to avoid the influence of the arrange and dividing mechanism on the data, tenfold cross-validation was used in all experimental studies. As a consequence, ten classification outcomes were achieved, and the mean of these outcomes is measured as the research's ultimate classification results.

4.4 Comparison of Classifiers

Table 1 displays the highest and lowest classification performance achieved by six classifiers on the CSDMC2010 dataset with vector sizes of 500 and 1000, and the best achieved results are marked bold. According to Table 1, the LR classifier with the AVMFO algorithm made the greatest outcome, with the highest obtained accuracy

Table 1 Highest and lowest classification performance comparison on the CSDMC2010 dataset

Classification techniques	Feature vector size = 500		Feature vector size = 1000	
	Highest (%)	Lowest (%)	Highest (%)	Lowest (%)
AVMFO-LR	**98.67**	**94.65**	**98.81**	**96.03**
MFO-LR	98.45	76.52	98.52	77.68
Gradient descent with LR [25]	98.25	73.82	98.54	66.74
SVM with RBF kernel [25]	98.57	68.22	98.61	68.23
Linear SVM [25]	98.11	92.32	98.41	94.11
Multinomial NB [25]	90.92	88.92	94.52	92.03
Gaussian NB [25]	93.73	93.73	95.94	95.94

Table 2 AVMFO-LR's performance and comparison on the Enron-1 dataset

Methodology	Performance ǀ STD	FN ǀ STD	FP ǀ STD
AVMFO-LR	**98.97**% ǀ 0.85	0.0005 ǀ 0.0010	0.0209 ǀ 0.0143
MFO-LR	98.83% ǀ 0.88	0.0007 ǀ 0.0013	0.0213 ǀ 0.0184
ABC-LR [25]	98.92% ǀ 0.86	0.0006 ǀ 0.0012	0.0212 ǀ 0.0178
DBB-RDNN-Rel [16]	98.74% ǀ 0.58	0.0016 ǀ 0.0019	0.0213 ǀ 0.0102
CNN [16]	97.49% ǀ 0.88	0.0346 ǀ 0.0194	0.0216 ǀ 0.0078
KNN [16]	91.35% ǀ 1.36	0.0377 ǀ 0.0156	0.1064 ǀ 0.0178
MLP [16]	96.31% ǀ 2.83	0.0502 ǀ 0.0824	0.0315 ǀ 0.0359
AIRS2Parallel [16]	71.36% ǀ 7.65	0.1220 ǀ 0.1679	0.3535 ǀ 0.1519
LR [16]	94.56% ǀ 1.03	0.0664 ǀ 0.0202	0.0498 ǀ 0.0126
AdaBoost [16]	78.76% ǀ 1.15	0.6838 ǀ 0.0340	0.0200 ǀ 0.0800
Random forest [16]	98.03% ǀ 0.58	0.0179 ǀ 0.0123	0.0202 ǀ 0.0074
Voting [16]	97.20% ǀ 1.06	0.0247 ǀ 0.0118	0.0294 ǀ 0.0130
IL-C4.5 [16]	93.35% ǀ 1.29	0.0608 ǀ 0.0207	0.0688 ǀ 0.0159
FDA + SVM [16]	96.64% ǀ 0.85	0.0462 ǀ 0.0175	0.0285 ǀ 0.0093
FDA + NB [16]	91.30% ǀ 1.21	0.1211 ǀ 0.0162	0.0044 ǀ 0.0059
MDL [16]	95.67% ǀ 1.13	0.0107 ǀ 0.0105	0.0566 ǀ 0.0165

of 98.81% on the CSDMC2010 dataset with feature size of 1000 and 98.67% with feature size of 500.

Table 2 displays the detection accuracy of the spam detection. The top spam detection performance was presented and associated with FN and FP rates, as well as the mean duration in seconds, in this subsection. In Table 2, the STD was displayed for every outcome acquired using tenfold cross-validation.

In classification performance and FN rate, the findings show that AVMFO-LR beat the other methodologies on the Enron dataset. On the Enron dataset, AVMFO-LR likewise generated much higher FN rates compared to the rest of the approaches.

When FP rates were compared, AVMFO-LR outperformed all other methods on the Enron dataset excluding the (FDA + NB, random forest, and AdaBoost methods). Nevertheless, due to their relatively weak detection effectiveness, these three approaches should not be used for spam detection (FN rate). The basic MFO-LR approach also obtained decent performances, finishing on the third place, behind the ABC-LR approach.

It is important to note that all demonstrated methods were implemented and tested in order to independently validate results for this research paper. Additionally, with higher number of features, accuracy increases; however, it also introduces some penalties because more LR coefficients need to be determined. Time complexity of the process does not scale linearly. During testing, it was determined that when feature size is increased from 500 to 1000, the metaheuristics takes approximately 30% more time to execute.

5 Conclusion

This paper has presented the process of training logistic regression model by moth flame optimizer for spam email classification. Firstly, the training method used to determine the ideal weight and bias value might progress to a local optimal solution and is responsive to identical weight values. A unique spam detection methodology was introduced in this paper that combines the benefits of LR, an improved version of MFO algorithms, and the tf–idf technique. The suggested methodology was validated on two publicly available datasets for spam detection (CSDMC2010 and Enron). To accommodate for language disparities, word stems and feature vector sizes of 500 and 1000 were obtained using stemming libraries. An extensive test was done on two NB methods, two SVM methods, an LR algorithm trained by gradient descent, and the suggested methodology on spam detection.

The suggested methodology exceeds the other five methods on both datasets and achieves remarkable achievements, with a highest obtained 98.81 success rate on the CSDMC2010 dataset with feature vector size of 1000. Additionally, the Enron dataset was used to evaluate the classification performance of the suggested methodology. Throughout classification performance and FN rate, the AVMFO-LR proved superiority by outperforming existing approaches and can clearly be used for spam detection (Table 3).

Future research will be using different swarm intelligence algorithms with LR classifiers for spam detection, or using the AVMFO-LR with different classification techniques, or checking the performance of the AVMFO with other NP-hardness challenges.

Table 3 AVMFO-LR's performance was evaluated to those of methodologies presented in an early studies on the Enron-1 dataset

Methodology	Performance (%)
AVMFO-LR	**98.97**
MFO-LR	98.83
ABC-LR [25]	98.92
DBB-RDNN-ReL [16]	98.74
Boosted NB + SVM [42]	95.60
Incremental SVM [36]	96.86
Natural language toolkit NB [23]	94.70
k means + SVM [27]	97.35
Relief + NB [43]	96.30
RF [33]	96.39
Enhanced genetic programming [41]	94.10
Bagged RF [38]	97.75
Minimum description length [3]	95.56
Distinguishing feature selector [45]	94.35
Multivariate Bernoulli NB [2]	94.79
AIS [1]	90.00
Deep belief networks [44]	97.43

References

1. Abi-Haidar A, Rocha LM (2008) Adaptive spam detection inspired by the immune system. In: ALIFE, pp 1–8
2. Almeida TA, Almeida J, Yamakami A (2011) Spam filtering: how the dimensionality reduction affects the accuracy of Naive Bayes classifiers. J Internet Serv Appl 1(3):183–200
3. Almeida TA, Yamakami A (2012) Occam's razor-based spam filter. J Internet Serv Appl 3(3):245–253
4. Bacanin N, Alhazmi K, Zivkovic M, Venkatachalam K, Bezdan T, Nebhen J (2022) Training multi-layer perceptron with enhanced brain storm optimization metaheuristics. Comput Mater Continua 70(2):4199–4215. http://www.techscience.com/cmc/v70n2/44706
5. Bacanin N, Arnaut U, Zivkovic M, Bezdan T, Rashid TA (2022) Energy efficient clustering in wireless sensor networks by opposition-based initialization bat algorithm. In: Computer networks and inventive communication technologies. Springer, Berlin, pp 1–16
6. Bacanin N, Bezdan T, Tuba E, Strumberger I, Tuba M, Zivkovic M (2019) Task scheduling in cloud computing environment by grey wolf optimizer. In: 2019 27th Telecommunications forum (TELFOR). IEEE, pp 1–4
7. Bacanin N, Bezdan T, Venkatachalam K, Zivkovic M, Strumberger I, Abouhawwash M, Ahmed A (2021) Artificial neural networks hidden unit and weight connection optimization by quasi-refection-based learning artificial bee colony algorithm. IEEE Access
8. Bacanin N, Bezdan T, Zivkovic M, Chhabra A (2022) Weight optimization in artificial neural network training by improved monarch butterfly algorithm. In: Mobile computing and sustainable informatics. Springer, Berlin, pp 397–409

9. Bacanin N, Petrovic A, Zivkovic M, Bezdan T, Antonijevic M (2021) Feature selection in machine learning by hybrid sine cosine metaheuristics. In: International conference on advances in computing and data sciences. Springer, Berlin, pp 604–616
10. Bacanin N, Stoean R, Zivkovic M, Petrovic A, Rashid TA, Bezdan T (2021) Performance of a novel chaotic firefly algorithm with enhanced exploration for tackling global optimization problems: application for dropout regularization. Mathematics 9(21). https://www.mdpi.com/2227-7390/9/21/2705
11. Bacanin N, Tuba E, Zivkovic M, Strumberger I, Tuba M (2019) Whale optimization algorithm with exploratory move for wireless sensor networks localization. In: International conference on hybrid intelligent systems. Springer, Berlin, pp 328–338
12. Bacanin N, Tuba M (2012) Artificial bee colony (ABC) algorithm for constrained optimization improved with genetic operators. Stud Inf Control 21(2):137–146
13. Bacanin N, Zivkovic M, Bezdan T, Cvetnic D, Gajic L (2022) Dimensionality reduction using hybrid brainstorm optimization algorithm. In: Proceedings of international conference on data science and applications. Springer, Berlin, pp 679–692
14. Bacanin N, Zivkovic M, Bezdan T, Venkatachalam K, Abouhawwash M (2022) Modified firefly algorithm for workflow scheduling in cloud-edge environment. Neural Comput Appl 1–26
15. Bacanin N, Zivkovic M, Salb M, Strumberger I, Chhabra A (2022) Convolutional neural networks hyperparameters optimization using Sine Cosine algorithm. In: Sentimental analysis and deep learning. Springer, Berlin, pp 863–878
16. Barushka A, Hajek P (2018) Spam filtering using integrated distribution-based balancing approach and regularized deep neural networks. Appl Intell 48(10):3538–3556
17. Bezdan T, Milosevic S, Venkatachalam K, Zivkovic M, Bacanin N, Strumberger I (2021) Optimizing convolutional neural network by hybridized elephant herding optimization algorithm for magnetic resonance image classification of glioma brain tumor grade. In: 2021 Zooming innovation in consumer technologies conference (ZINC). IEEE, pp 171–176
18. Bezdan T, Stoean C, Naamany AA, Bacanin N, Rashid TA, Zivkovic M, Venkatachalam K (2021) Hybrid fruit-fly optimization algorithm with k-means for text document clustering. Mathematics 9(16):1929
19. Bezdan T, Zivkovic M, Tuba E, Strumberger I, Bacanin N, Tuba M (2020) Glioma brain tumor grade classification from MRI using convolutional neural networks designed by modified FA. In: International conference on intelligent and fuzzy systems. Springer, Berlin, pp 955–963
20. Bezdan T, Zivkovic M, Tuba E, Strumberger I, Bacanin N, Tuba M (2020) Multi-objective task scheduling in cloud computing environment by hybridized bat algorithm. In: International conference on intelligent and fuzzy systems. Springer, Berlin, pp 718–725
21. Bhowmick A, Hazarika SM (2018) E-mail spam filtering: a review of techniques and trends. In: Advances in electronics, communication and computing, pp 583–590
22. Bird S, Klein E, Loper E (2009) Natural language processing with Python: analyzing text with the natural language toolkit. O'Reilly Media, Inc.
23. Chhogyal K, Nayak A (2016) An empirical study of a simple Naive Bayes classifier based on ranking functions. In: Australasian joint conference on artificial intelligence. Springer, Berlin, pp 324–331
24. Cuk A, Bezdan T, Bacanin N, Zivkovic M, Venkatachalam K, Rashid TA, Devi VK (2021) Feedforward multi-layer perceptron training by hybridized method between genetic algorithm and artificial bee colony. In: Data science and data analytics: opportunities and challenges, p 279
25. Dedeturk BK, Akay B (2020) Spam filtering using a logistic regression model trained by an artificial bee colony algorithm. Appl Soft Comput 91:106229
26. Gajic L, Cvetnic D, Zivkovic M, Bezdan T, Bacanin N, Milosevic S (2021) Multi-layer Perceptron training using hybridized bat algorithm. In: Computational vision and bio-inspired computing. Springer, Berlin, pp 689–705
27. Hassan D (2016) Investigating the effect of combining text clustering with classification on improving spam email detection. In: International conference on intelligent systems design and applications. Springer, Berlin, pp 99–107

28. Idris I, Selamat A (2014) Improved email spam detection model with negative selection algorithm and particle swarm optimization. Appl Soft Comput 22:11–27
29. Kottursamy K (2021) A review on finding efficient approach to detect customer emotion analysis using deep learning analysis. J Trends Comput Sci Smart Technol 3(2):95–113
30. Ma TM, Yamamori K, Thida A (2020) A comparative approach to Naïve Bayes classifier and support vector machine for email spam classification. In: 2020 IEEE 9th Global conference on consumer electronics (GCCE). IEEE, pp 324–326
31. Metsis V, Androutsopoulos I, Paliouras G (2006) Spam filtering with Naive Bayes-which Naive Bayes? In: CEAS, vol 17, Mountain View, CA, pp 28–69
32. Mirjalili S (2015) Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm. Knowl Based Syst 89:228–249
33. Mishra R, Thakur RS (2013) Analysis of random forest and Naïve Bayes for spam mail using feature selection categorization. Int J Comput Appl 80(3):42–47
34. Naem AA, Ghali NI, Saleh AA (2018) Antlion optimization and boosting classifier for spam email detection. Future Comput Inf J 3(2):436–442
35. Radicati S (2019) Email statistics report, 2019–2023. The Radicati Group, Inc. Available at: www.radicati.com/wp/wp-content/uploads/2018/12/Email-Statistics-Report-2019-2023-Executive-Summary.pdf
36. Sanghani G, Kotecha K (2016) Personalized spam filtering using incremental training of support vector machine. In: 2016 International conference on computing, analytics and security trends (CAST). IEEE, pp 323–328
37. Sarma D, Mittra T, Bawm RM, Sarwar T, Lima FF, Hossain S (2021) Comparative analysis of machine learning algorithms for phishing website detection. In: Inventive computation and information technologies. Springer, Berlin, pp 883–896
38. Shams R, Mercer RE (2013) Personalized spam filtering with natural language attributes. In: 2013 12th International conference on machine learning and applications, vol 2. IEEE, pp 127–132
39. Strumberger I, Sarac M, Markovic D, Bacanin N (2018) Moth search algorithm for drone placement problem. Int J Comput 3
40. Strumberger I, Tuba E, Bacanin N, Zivkovic M, Beko M, Tuba M (2019) Designing convolutional neural network architecture by the firefly algorithm. In: 2019 International young engineers forum (YEF-ECE). IEEE, pp 59–65
41. Trivedi SK, Dey S (2013) An enhanced genetic programming approach for detecting unsolicited emails. In: 2013 IEEE 16th International conference on computational science and engineering. IEEE, pp 1153–1160
42. Trivedi SK, Dey S (2019) A combining classifiers approach for detecting email spams. In: 2016 30th International conference on advanced information networking and applications workshops (WAINA). IEEE, pp 355–360
43. Trivedi SK, Dey S (2016) A comparative study of various supervised feature selection methods for spam classification. In: Proceedings of the second international conference on information and communication technology for competitive strategies, pp 1–6
44. Tzortzis G, Likas A (2007) Deep belief networks for spam filtering. In: 19th IEEE International conference on tools with artificial intelligence (ICTAI 2007), vol 2. IEEE, pp 306–309
45. Uysal AK, Gunal S (2012) A novel probabilistic feature selection method for text classification. Knowl Based Syst 36:226–235
46. Vivekanandam B (2021) Design an adaptive hybrid approach for genetic algorithm to detect effective malware detection in android division. J Ubiquitous Comput Commun Technol 3(2):135–149
47. Zivkovic M, Bacanin N, Djordjevic A, Antonijevic M, Strumberger I, Rashid TA et al (2021) Hybrid genetic algorithm and machine learning method for covid-19 cases prediction. In: Proceedings of international conference on sustainable expert systems. Springer, Berlin, pp 169–184
48. Zivkovic M, Bacanin N, Tuba E, Strumberger I, Bezdan T, Tuba M (2020) Wireless sensor networks life time optimization based on the improved firefly algorithm. In: 2020 International wireless communications and mobile computing (IWCMC). IEEE, pp 1176–1181

49. Zivkovic M, Bacanin N, Venkatachalam K, Nayyar A, Djordjevic A, Strumberger I, Al-Turjman F (2021) COVID-19 cases prediction by using hybrid machine learning and beetle antennae search approach. Sustain Cities Soc 66:102669
50. Zivkovic M, Bacanin N, Zivkovic T, Strumberger I, Tuba E, Tuba M (2020) Enhanced grey wolf algorithm for energy efficient wireless sensor networks. In: 2020 Zooming innovation in consumer technologies conference (ZINC). IEEE, pp 87–92
51. Zivkovic M, Bezdan T, Strumberger I, Bacanin N, Venkatachalam K (2021) Improved Harris Hawks optimization algorithm for workflow scheduling challenge in cloud–edge environment. In: Computer networks, big data and IoT. Springer, Berlin, pp 87–102
52. Zivkovic M, Stoean C, Chhabra A, Budimirovic N, Petrovic A, Bacanin N (2022) Novel improved salp swarm algorithm: an application for feature selection. Sensors 22(5):1711
53. Zivkovic M, Zivkovic T, Venkatachalam K, Bacanin N (2021) Enhanced dragonfly algorithm adapted for wireless sensor network lifetime optimization. In: Data intelligence and cognitive informatics. Springer, Berlin, pp 803–817

Semantic-Based Feature Extraction and Feature Selection in Digital Library User Behaviour Dataset

F. Mary Harin Fernandez, I. S. Hephzi Punithavathi, T. Venkata Ramana, and K. Venkata Ramana

Abstract World Wide Web has become a universal environment for human interface, collaboration, communication, data storage and data sharing. Information retrieval is a technique to understand the text in the web pages. The purpose of Semantic Web is to allow the machines to advance knowledge itself by recognizing its meaning. The essential intention of this research is to predict variation in the directional behaviour of the digital library user, based on their recommended data. The significance of this research lies in enlightening the user behaviour in digital library using semantic-driven approach. An important contribution of this research work corresponds to retrieve accurate data from websites by using semantically enhanced various algorithms. This research concerns on identifying educational digital library web user's behaviour. The user behaviour unstructured data are pre-processed and semantically extracted using suggested weighted TF-IDF normalization method. Then the extracted features are selected using improved Ensemble-based FS method. It improves the performance and reliability of classification by eliminating unnecessary and superfluous features from the extracted user behaviour datasets.

Keywords TF-IDF · Feature extraction · Ensemble feature selection · WordNet · N-grams · Semantic Web

F. M. H. Fernandez (✉) · I. S. H. Punithavathi · T. V. Ramana
Sphoorthy Engineering College, Hyderabad, Telangana, India
e-mail: mary.fherin@gmail.com

I. S. H. Punithavathi
e-mail: hephzi21@gmail.com

T. V. Ramana
e-mail: meetramana1204@gmail.com

K. V. Ramana
VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India
e-mail: venkataramana_k@vnrvjiet.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_57

1 Introduction

World Wide Web websites are growing rapidly due to the growth of e-learning, e-commerce and other business activities. Due to unprecedented growth of information and size of World Wide Web, it becomes difficult to search and extract the pertinent information from web.

The information obtained from the digital library user behaviour data is provided in the pre-processing phase. Pre-processing is usually undertaken to find out whether there is a lot of extraneous and unnecessary information on the dataset. It involves numerous measures such as cleaning, normalization, transformation, extraction and selection of features. The outcome of pre-processing data is the ultimate dataset with minimized attributes. The suggested methods are analysed using three distinct benchmark datasets of behaviour, which are pre-processed and extracted semantically using TF-IDF method [1–3].

The proposed feature selection algorithm is dependent on an Ensemble selection from libraries of models in which each model is learned using just one feature. Ensemble learning is an efficient machine learning method. The key theme is to produce improved learning results by integrating various learning models. The faster growth of Ensemble feature selection in recent days draws on the concept of Ensemble learning. The aim of the Ensemble FS is to get several optimal features.

2 Literature Review

The authors Lu et al. have proposed feature extraction and text classification approach to textual data related to a rare event in social platform [4]. They have used the score and weight value for the frequently used words for extraction and later applied the fuzzy logic for classifying the words. This method helps to identify the rare event-related and unrelated text messages.

Mutinda et al. developed a feature extraction and feature selection approach which uses a fixed hybrid N-gram window for feature extraction especially the N-gram which is more efficient when compared with other methods by generating a hybrid vector from words, part-of-speech (POS) tags and word semantic orientation [5]. Minimum redundancy maximum relevance feature selection algorithm was used for sentence-level sentiment analysis.

The authors Nashit et al. have recommended a methodology which classifies email text into three categories, i.e. order, request and general on basis of imperative sentences [6]. This research uses Word2Wec for words conversion into vector and uses two approaches of deep learning, i.e. convolutional neural network and recurrent neural network for email classification. Result shows that RNN gives better accuracy than CNN. These experiments show that increasing in the ratio of training dataset results in increasing accuracy of algorithm.

Zhang et al. authors suggested a feature extraction method based on TF-IDF the concept and game-theoretic shadowed groups to choose the features of the term [7]. For each term among documents, the asymmetric thresholds for choosing appropriate terms are extracted from repeatable experiments on the modification of TF-IDF.

Bommert et al. authors examine how various filter methods operate, compare their output in terms of runtime and predictive performance and present application guidance [8]. Depending on 16 high-dimensional category datasets and associated with a classification technique, 22 filter methods are evaluated with regard to runtime and reliability. Author concluded there is no collective of filter methods which always exceed all the other techniques but that suggestions are provided on filter methods that execute well on several of the datasets.

Kadriu et al. authors concentrate on the document analysis of Albanian news articles using two methods [9]. The terms in a set were considered as individual components in the first method, each of which was assigned a corresponding variable in the vector domain. The findings show that when checking the classification system for not a large text database, the bag of words performs better than fastText. The best results 94% accuracy was achieved with a bag of words.

Hybrid technique that incorporates various techniques to obtain better outcomes. Zhao et al. proposed a novel FS method for hybrid H-gram with merged cross-entropy [10]. The suggested method is added efficient in the performance indexes of all four categorization algorithms such as ID3, AdboostM1, Bagging and random forest.

Vidhya et al. authors proposed a term frequency with a stemmer-based FE algorithm for data classification [11]. They extract characteristics that depend on the term frequency (TF) and inverse term frequency (ITF). For FE feasibility, the proposed method was analysed and contrasted with other methods such as Null Stemmer.

From the aforementioned literature review, the weighted words and TF-IDF models in FE are easy to determine the likenesses between two documents, but do not apprehend location or context in text. Most of the filter methods were used qualitative measures for FS. This method does not ignore the incomplete and redundant features. The filter design is faster in computation, but it lacks the accuracy. In the wrapper method, the features are chosen on the basis of learning algorithm. It is more reliable than the filter approach, but its computation is slow and complicated due to iterative filtering of the right subset of features.

3 Proposed Work

3.1 Pre-Processing

The information obtained from the digital library user behaviour data is provided in the pre-processing phase. During pre-processing, digital library user behavioural raw data is used as input and many data pre-processing approaches are incorporated linearly to minimize invalid instances which remain in the dataset. Once the dataset

has been formed, the next step is to implement the required pre-processing techniques, such as inspecting all words in a sentence, reducing text errors, special characters, numbers and lowering all the uppercase letters in the text [12]. Finally, the stop words are also removed from the dataset.

In NLP, one term that occurs in various forms (i.e. singular and plural form) even as the semantic sense of each type remains the same. One technique for consolidating the various types of a term into the similar feature space is stemming [13]. Stemming method transforms words into root or stem structure. In this research, the Porter stemmer algorithm is used for stemming the terms. Algorithm 1 is used to search out the root of a certain set of terms by eliminating various suffixes while retaining the meaning of the term.

Porter Stemmer : ***Algorithm 1***

Input: List of preprocessed words **Output:** Stemmed(Root) words
1: Remove plurals (-s) and suffixes (-ing or -ed) 2: If the vowel follows in the previous step, replace y to i on the next word 3: Map paired suffixes to single ones (-ational,-ization) 4: Furthermore, decreases the suffixes similar (-ness, -full) etc 5: Removes (-ence, -ant) etc 6: If a word terminates with a grammatical verb ending, then it has been removed 7: Lastly, eliminates a (-e)

3.2 Feature Extraction

The real document is extremely dimensional and unorganized. It implies that a single term may be used as a different aspect. To reduce the document dimensionality, the most distinctive features of a text must be extracted. Figure 1 illustrates the key phases in the method of cleaning raw information, extracting N-grams and discovering meaningful N-grams. Algorithm 2 describes the major steps followed in semantic-based feature extraction.

3.2.1 Syntactic Word Representation

The syntactic and semantic information for extracting sentences in the document representation comes from the scientific genomic documents. In the proposed system, n-gram approach is used as a solution to the syntactic challenge for feature extraction.

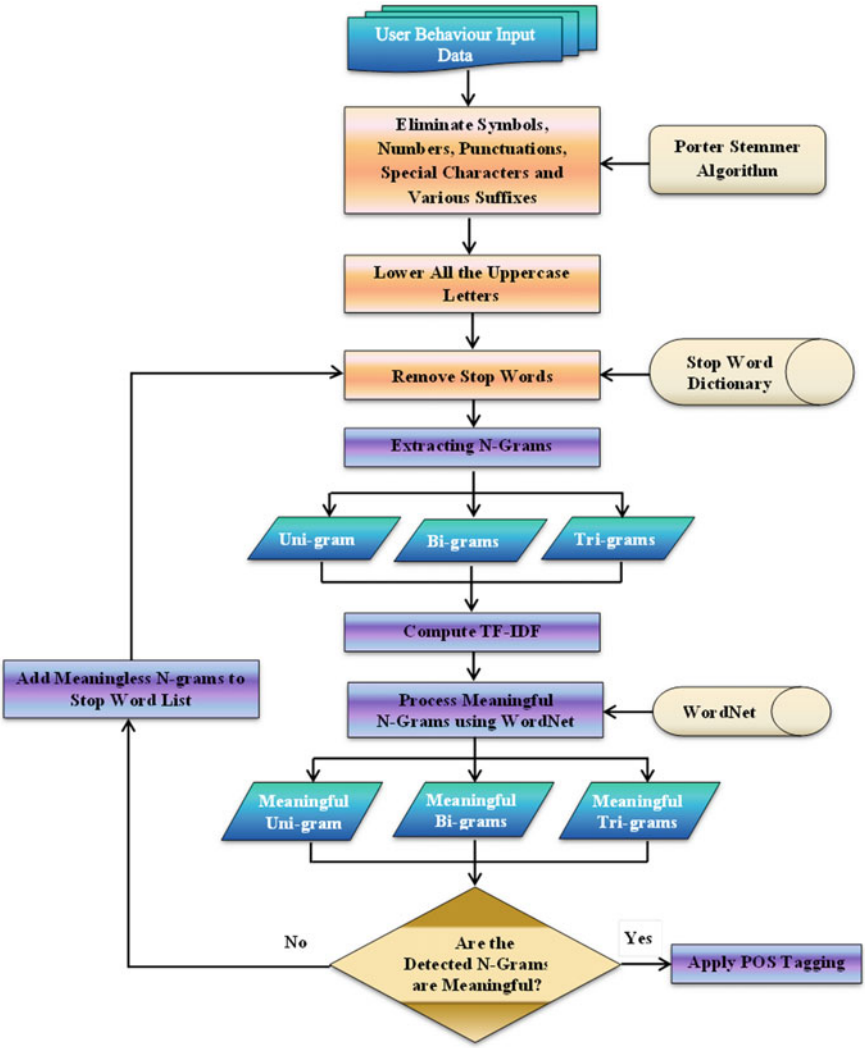


Fig. 1 Key steps to discover significant N-grams

Extracting N-grams

The N-gram method is a sequence of n-words that appears 'in that order' in a text set. This is not a text representation but it may be used to depict a text as a feature. The analysis of N-gram is used to identify any feasible combination of words that appeared together without any other words in between. The feature N-gram explores all word combinations of two, three or more. N-grams improve the frequency matrices at the document term. However, the occurrence of the N-gram cannot be used explicitly

as a measure of the significance or informativeness of the N-gram, because both the volume of the documents and the degree of commonality of the words may falsely raise the significance of the N-gram. Therefore, certain normalization is essential.

TF-IDF Normalization

TF-IDF weight is the most used weight for knowledge extraction or text mining. This weight is a statistical method used to determine the value of a term in a list or corpus text. TF-IDF may be effectively used for filtering stop words in specific topic fields, including document description and classification. Usually, the TF-IDF weight consists of two terms: the first one represents the normalized TF. The number of times a term occurs in a text, divided by the total number of terms in that text given in Eq. (1); the second term is the IDF, calculated as a logarithm of the number of documents in the database determined by the number of documents in which the particular term occurs given in Eq. (2). The product of the TF and the IDF functions can take into account both the frequency of the length of the document and the frequency of the words of all the documents in the database as presented in Eq. (3).

$$\mathrm{TF}(t) = \left(\frac{\mathrm{Numberoftimesthewordtoccursinthetext}}{\mathrm{Totalnumberofwordsinthetext}}\right) \tag{1}$$

$$\mathrm{IDF}(t) = \log\left(\frac{\mathrm{Totalnumberoftexts}}{\mathrm{Numberoftextswithwordtinit}}\right) \tag{2}$$

$$\mathrm{TF} - \mathrm{IDF}(t_i) = \mathrm{TF}(t_i) \times \mathrm{IDF}(t_i) \tag{3}$$

where $\mathrm{TF}(t_i)$ = number of documents containing t_i.
$\mathrm{IDF}(t_i)$ = number of inverse documents containing t_i

Meaningful N-grams

Some N-grams in the document-term matrix have greater discriminating strength than others. Therefore, it is important to choose only the N-grams that can potentially impact the analysis of user behaviour. This also increases the statistical performance of the algorithm when the matrix size becomes too large. For this purpose, the following analyses can be carried out:

(a) **N-gram's Normalized Frequency Analysis**: For this analysis, the sum of the normalized frequencies correlated with each N-gram is determined through all records and then the N-grams are arranged for decreasing order. Top N-grams with the highest normalized frequency are chosen.
(b) **WordNet:** WordNet is an English language lexical repository. This consists of verbs, nouns, adjectives, adverbs and structured around lexical-semantic

associations in terms of their meaning. The WordNet 2.1 incorporated in NLTK Lite used for this research.

Let S be a collection of notions for taxonomy. The relation between the two terms is the magnitude of the definition that subsumes it all in taxonomy. Let the terminology be supplemented with the feature prob : S $\rightarrow$ [0, 1], so that for every $s \in S$, prob(s) is the likelihood of encountering notion s. If the taxonomy has a special top node, its likelihood is 1. The knowledge content of s can be computed as −logprob(s). Then,

$$\text{similarity}(s_1, s_2) = \max_{s \in C(s_1, s_2)} \left[-\text{logprob}(s) \right] \tag{4}$$

In Eq. (4), the $C(s_1, s_2)$ is the collective of concepts including both s_1 and s_2. The term 'similarity' is described in Eq. (5).

$$\text{similarity}(t_1, t_2) = \max_{s_1, s_2} \left[\text{similarity}(s_1, s_2) \right] \tag{5}$$

If similarity(t_1) and (t_2) reflect a collection of words in taxonomy that are concepts of the term t_1, t_2 accordingly, s_1 exceeds similarity(t_1), s_2 exceeds similarity(t_2).

The similarity of two notions as the relationship between the amount of data required indicates the similarity in them and the data required to fully define them, as mentioned by Lin [14]. In Eq. (6), the notions $n_1 \in s_1$ and $n_2 \in s_2$, s_0 are the highest unique class considering both s_1 and s_2.

$$\text{similarity}(n1, n2) = \frac{2 \times \text{logprob}(s_0)}{\text{logprob}(s_1) + \text{logprob}(s_2)} \tag{6}$$

(c) **Mapping Words into Concepts:** This technique expands each word vector $\overrightarrow{w}_t$ by recent entries for WordNet notions S occurring in the texts package. Therefore, the vector $\overrightarrow{w}_t$ will be substituted by the concatenation of $\overrightarrow{w}_t$ and $\overrightarrow{s}_t$ where $\overrightarrow{s}_t := (sf(t, s_1), \ldots, \text{sf(t, s}_\text{m}))$. The notion vector with $m = |S|$ and $sf(t, s)$ represents the frequency at which the notion $s \in S$ occurs in the text t.

Feature Extraction: *Algorithm 2*

Input: Preprocessed word list

Notation: T_i: i^{th} word in M
M: a matrix with word and POS-Tag
W_i: word list of i^{th} word
OS: Semantic Orientation
POS-Tag: A matrix contains the terms with commentated POS
PMI-IR: Pointwise Mutual Information and Information Retrieval

Output: TF-IDF, frequency of adjective and adverb

(continued)

(continued)

1. Separate every T_i as unigram and save into W_i
2. Calculate TF-IDF Normalization
   For each word T_i in M do
      For each element in weight matrix
         Set the threshold 'c'
         Calculate Document Frequency (DF) for each term
          If DF < c then
            Remove the term along with its weight from weight matrix
          End if
       End for
   End for
3. Identify Meaningful N-gram
   a) Find the synonym of T_i with nltk.word_tokenize using the package NLTK
   b) Annotate the T_i with POS-tag
      i. Determine POS norms and it's SO using the PMI-IR
      ii. Calculate the similarity string of the synonym words
      iii. Count the number of adjective and adverb in POS-tag
   c) Return the result

(d) **Adding Hypernyms:** The hypernym association between concepts is defined by applying to the concept frequency sf of every term in the document frequencies. The concept vector element frequencies are then modified in the following manner: For all $s \in S$, substitute $sf(t, s)$ by $sf'(t, s)$ in Eq. (7), where b represents numerous distinct concepts and $H(s)$ provides its hyponyms to a specified word w.

$$sf^{'}(t, s) = \sum_{b \in H(s)} sf(t, b) \tag{7}$$

The semantic similarity concept vector is determined with the similarity measures stated in Eq. (8).

$$\text{Similarity}\left(sv_i, sv_j\right) = \frac{\overrightarrow{sv_i}.\overrightarrow{sv_j}}{|sv_i|\left|sv_j\right|} \tag{8}$$

where sv_i and sv_j represent the concept vector of a specified word. $\overrightarrow{sv_i}$ and $\overrightarrow{sv_j}$ relate to TF-IDF vector estimates, and $|sv_i|$ and $\left|sv_j\right|$ denote the magnitude of vectors.

PMI-IR: It is used to identify synonym polarity and recognize user input as recommended and unrecommended. For a given sentence, the PMI information retrieval is used to compute semantic orientation (SO). PMI is the level of mathematical dependency provided by term_1 and term_2 as shown in Eq. (9).

$$\text{PMI}(\text{term}_1, \text{term}_2) = \log_2\left(\frac{p(\text{term}_1 \& \text{term}_2)}{p(\text{term}_1) \times p(\text{term}_2)}\right) \tag{9}$$

The SO estimate for every word is determined on the basis of the relationship with the positive context term 'Good' and the negative context term 'Completely Irrelevant', and it is specified in Eq. (10).

$$\mathrm{SO(phrase)} = \mathrm{PMI(phrase, \text{“good}")} - \mathrm{PMI\ (phrase, \text{“completelyirrelevant}")} \quad (10)$$

3.3 Improved Ensemble Feature Selection Framework

Ensemble selection is a process of developing Ensembles by a library of model. First it creates a model library with various classifiers and parameters, and then it develops Ensemble procedure. Two approaches are suggested to the issue of overfitting and enhance selection efficiency of the Ensemble features.

a. **Multi-Bagging Method**

Since there are a large number of models in the library, overfitting becomes a major problem. When the range of models in the library raises, the chance to discover model variations that overfit also raises. The bagging strategy will reduce this problem by creating several bags of models. The concept is to randomly choose multiple bags from the library of user behaviour models. Each bag includes a random set of user behaviour attributes which is smaller than the library's total number of models. Then, the Ensemble selection of user behaviour term is performed in each bag. The returned Ensemble is ultimately the combining of the sub-Ensembles' final weight retrieved by each bag. Algorithm 3 describes a comprehensive overview of the multi-bagging ensemble feature selection.

Ensemble Feature Selection with Multi-Bagging : ***Algorithm 3***

Input: List of Bags
Output: Merged Sub-Ensemble Final Weight

1: **Set** Final Weight ← 0
2: **Repeat** for every defined number of Bags
    2.1 Develop a random Bag of Models
    2.2 Select the best Models ← Call ESR Algorithm
    2.3 Increase best Model Weight
3: **Merge** the Sub-Ensemble Bags
    3.1 Final Weight ← Merged Sub-Ensemble Bag Weights
    3.2 Best Feature ← Highest absolute Weight
4: **Return** Final Weight

b. **Dropout Method**

Dropout methodology is suggested for solving the issue of overfitting in deep neural network [15]. The concept behind this strategy is to drop NN units at random during

the training process. In this research, the three versions of dropout are used which is proposed by Yassine Akhiat et al. [16]. They are model dropout, instances dropout and merged dropout. The Ensemble's detailed procedure with DROPOUT technique is given in Algorithm 4.

The theoretical description of the suggested algorithm is as follows:

(1) Using the multi-bagging selection process, the features are chosen and sorted by its significance to achieve several sorted effective feature subsets $\text{FSM}_1, \text{FSM}_2 \ldots \text{FSM}_t$.
(2) For every feature j in the FSM_i, equalize its significance with $(m - j)/m$ and get the feature weight set of the i^{th} FS process $WS_i = \{WS_1, WS_2, \ldots WS_n\}$.
(3) The top feature of $\alpha\%$ is chosen depending on the level of threshold α from the sequences of the sorted features to generate the finest feature subset.
(4) Depending on the finest feature subset, the efficiency of the algorithm is tested using a number of successful classifiers.

Ensemble with DROPOUT : ***Algorithm 4***

Input: List of Models
Notation:
FSM: Feature Selection Method
FSM_{best}: Best Feature in Feature Selection Method
WS: Weight Sequence
F_j: i^{th} Feature
W_{ij}: Weight of the Ensemble FSM
N: Number of Models in Library
m: Total number of Features
α: Threshold Value
Output: FSM_{best}

1: Initialization of Weight
Set each Model Weight ← 0
2: **Repeat** until there is no increase in Ensemble efficiency
2.1 Randomly Drop N Models from Ensemble
3: **Apply** ESR Algorithm
4: **Estimate** the Weight of the Ensemble
Get WS of FSM_i
For FSM_i in $\{FSM_1, FSM_2, ..., FSM_n\}$
For F_j in FSM_i
W_{ij}= (m − j)/m
End For
End For
Return the weight sequences $\{WS_1, WS_2, \ldots, WS_n\}$
5: **Return** FSM_{best} Ensemble Weight
Get best feature subsequence respective to α
Sort FSM according to WS
FSM_{best} ← First α% features
Return FSM_{best}

---

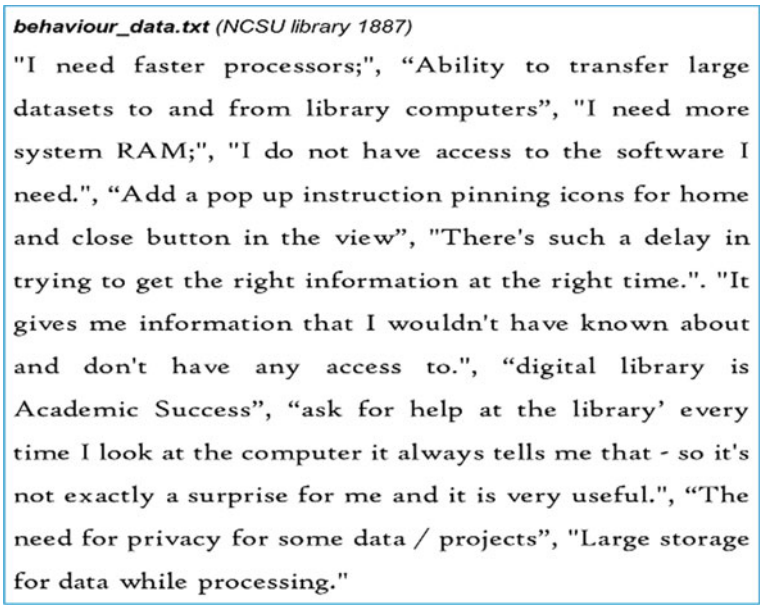
4 Result and Discussion

In pre-processing phase, 2187 records from the digital library user behaviour dataset are merged into a single structured database table and any records that do not fit the schema can be removed from the record set. Figure 2 illustrates an example of the user behaviour plain text. The pre-processed text documents can be incorporated into csv format for N-gram analysis. Figure 3 depicts the text related with the example of user behaviour data after pre-processing.

4.1 TF-IDF Feature Extraction

The behaviour dataset recommended by the user consists of 2187 records. The behavioural activities of the user are analysed with the various aspects of accessing patterns of digital library learning behaviour.

There are 413,562 different concepts in the 2187 user feedback records. Eventually more concepts may be detected when adding bigrams, trigrams and more. Unigram, bigrams, trigrams and typically N-grams may be applied to the dataset, and the TF-IDF converts the matrix of the extended features to see whether accuracy increases. Figure 4 displays an example of the N-grams derived from the user behaviour data. Table 1 shows the result of adding the TF-IDF function, which normalizes the terms


behaviour_data.txt *(NCSU library 1887)*

"I need faster processors;", “Ability to transfer large datasets to and from library computers”, "I need more system RAM;", "I do not have access to the software I need.", “Add a pop up instruction pinning icons for home and close button in the view”, "There's such a delay in trying to get the right information at the right time.". "It gives me information that I wouldn't have known about and don't have any access to.", “digital library is Academic Success”, “ask for help at the library’ every time I look at the computer it always tells me that - so it's not exactly a surprise for me and it is very useful.", “The need for privacy for some data / projects”, "Large storage for data while processing."

Fig. 2 User behaviour plain text

'need'	'need'
'faster'	'add'
'processor'	'pop'
'ability'	'instruction'
'transfer'	'pin'
'large'	'icon'
'dataset'	'home'
'library'	'close'
'computer'	'button'
'need'	'view'
'system'	'delay'
'ram'	'right'
'access'	'information'
'software'	'right'

Fig. 3 Partial views of user behaviour data after pre-processing

Unigram	Bigram	Trigram
'need'	'need-faster'	'need-faster-processor'
'faster'	'faster- processor'	'faster-processor-ability'
'processor'	'processor-ability'	'processor-ability-transfer'
'ability'	'ability-transfer'	'ability-transfer-large'
'transfer'	'transfer-large'	'transfer-large-dataset'
'large'	'large-dataset'	'large-dataset-library'
'dataset'	'dataset-library'	'dataset-library-computer'
'library'	'library-computer'	'library-computer-need'
'computer'	computer-need'	'computer-need-system'
'need'	'need-system'	'need-system-ram'
'system'	'system-ram'	'system-ram-access'
'ram'	'ram-access'	'ram-access-software'
'access'	'access-software'	'access-software-need'
'software'	'software-need'	'software-need-add'
'need'	'need-add'	'need-faster-processor'
'add'	'add- pop'	'faster-processor-ability'

Fig. 4 Pre-processed N-gram data before processing TF-IDF function

related with the frequencies.

4.1.1 Detecting Meaningful N-grams

The taxonomy of the semantic similarity of each N-gram is calculated and ordered to find the most suitable N-grams with the highest frequency, depending on their ordered. Table 2 presents the 10 most important trigrams that encompass the class

Table 1 Partial view of TF-IDF normalized user behaviour dataset

Document text	TF	IDF	TF-IDF
'need'	0.0555556	2.9141101	0.161895
'faster'	0.0833333	2.764671	0.2303893
'processor'	0.0555556	2.8347853	0.1574881
'ability'	0.0277778	2.7683515	0.0768987
'transfer'	0.1111111	3.2117069	0.3568563
'large'	0.0277778	3.1059956	0.0862777
'dataset'	0.1388889	3.4290199	0.4762528
'library'	0.0833333	2.8355037	0.236292
'computer'	0.1111111	2.9055775	0.3228419
'system'	0.1666667	2.9623641	0.4937273
'ram'	0	3.2273745	0
'access'	0.0555556	2.7096671	0.1505371
'software'	0.1388889	2.652281	0.3683724
'add'	0.0277778	3.0090856	0.0835857
'pop'	0.0277778	3.2220889	0.0895025
'instruction'	0.1666667	3.0016434	0.5002739
'pin'	0.0277778	3.492689	0.0970191
'icon'	0.0833333	3.4642523	0.2886877
'home'	0.0555556	2.8699064	0.1594392

Table 2 Detected sample top 10 meaningful trigrams

'need-faster-processor'
'transfer-large-dataset'
'need-system-ram'
'add-pop-instruction'
' delay-right-information'
right-time-information'
library-academic-success'
journal-exactly-surprise'
happy-access-library'
library-resource-satisfy'

'Resource Feedback'. The feature dimensionality is decreased by choosing the top-level N-gram which provides the taxonomy of an accurate meaning.

The extra feature used increases the range by more than 16,000,00 bigrams. Such bigrams were obtained using the POS norms and its Semantic Orientation was defined using the PMI-IR. The PMI computes Semantic Orientation for a given N-grams. The measurement of the semantic orientation for each N-gram is defined on the basis of

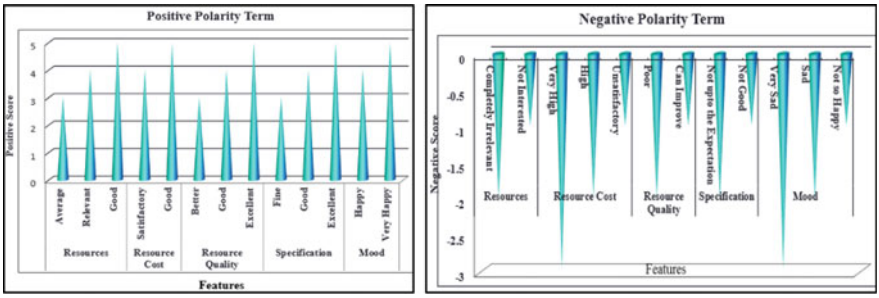


Fig. 5 Features versus score for positive and negative polarity term

Table 3 Selection of optimum features with different dropout methods

Datasets	Instance dropout (%)	Model dropout (%)	Merged dropout (%)	No. of bags
200	92	94	81	40
400	87	90	83	40
800	80	81	86	40
1600	64	68	91	40
2100	45	50	97	40

the relation with the positive context terms and the negative context terms as shown in Fig. 5. The attributes in the X-axis and their scores in the Y-axis were displayed.

Feature Selection

The feature selection approach optimizes the models of machine learning and makes it simpler. This experiment is about evaluating the proposed algorithms which reduce the issue of overfitting and increase the performance of the Ensemble selection.

In Table 3, the instance and model dropouts demonstrate their higher efficacy on datasets when it has few instances but the proposed merged dropout algorithm outperforms when it has larger number of instances and it drops the part of instances during the selection phase which decreases the number of west instances of user behaviour datasets [13]. The graphical representation of optimum feature selection performance is shown in Fig. 6. This validates the usefulness of merged dropout with its promising potential for huge length data mining activities.

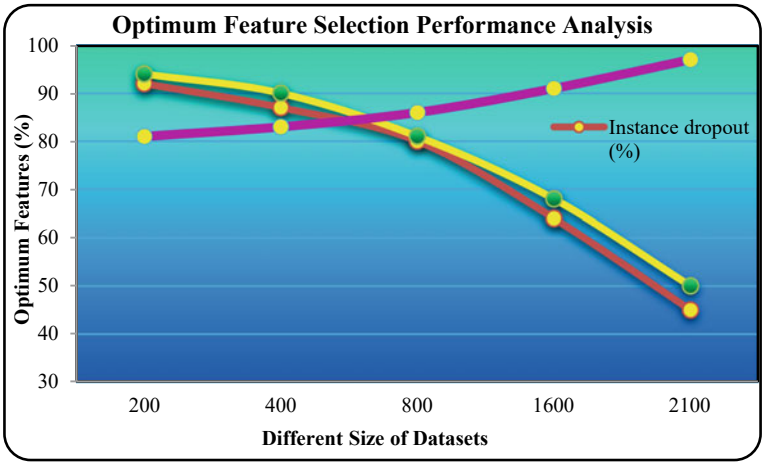


Fig. 6 Optimum feature selection performances analysis

5 Conclusion

The plain text documents of the digital library user feedback were obtained from three benchmark datasets. Such unstructured data are processed and extracted semantically, utilizing the suggested pre-processing and cleaning method that increases the accuracy of the data to be retrieved. The concept of developing feature extraction focuses on implementing semantic features in the document. The features are extracted using the recommended weighted TF-IDF normalization with meaningful both N-grams and POS tagging. WordNet's framework is used for identifying and linking Hypernyms. In this research, the incorporated NLTK Lite WordNet 2.1 is used.

The semantically extracted features are selected using improved Ensemble-based FS algorithm which reduces dimensionality, eliminates redundant and unnecessary features, optimizes the algorithm's predictive efficiency. Multi-bagging and dropout techniques are additionally used to improve the suggested FS algorithm. Experiments with the suggested algorithms reduce the overfitting issue of Ensemble selection and select most suitable user behaviour subset of the feature.

This new methodology is used to develop semantically accessible resources in the digital library and produces better learning facilities to the students' demands and preferences.

This research work can be extended in future by the extraction of more Semantic data like optimum semantic information, such as the Anchor Text (AT) category, additional domain coverage and context link extraction.

References

1. NCSU library Available in https://www.lib.ncsu.edu/ project-reports/user-research
2. Cortez P, Silva A (2008) Using data mining to predict secondary school student performance. In: Brito A, Teixeira J (eds) Proceedings of 5th future business technology conference (FUBUTEC 2008), Porto, Portugal, pp 5–12. ISBN 978–9077381–39–7
3. Young SWH (2017) Digital library user research data. Montana State University Scholar Works. Available in https://doi.org/10.15788/M2W88P
4. Lu XS, Zhou M, Wu K (2021) A novel fuzzy logic-based text classification method for tracking rare events on twitter. IEEE Transactions on Systems, Man, and Cybernetics: Systems 51(7):4324–4333
5. Mutinda J, Mwangi W, Okeyo G (2021) Lexicon-pointed hybrid N-gram features extraction model for sentence level sentiment analysis. 3:1–17
6. Nashit A, Fatima A, Shahzadi H, Ullah A, Polat K (2021) Feature extraction aligned email classification based on imperative sentence selection through deep learning. Journal of Artificial Intelligence and Systems 3:93–114
7. Zhang Y, Zhou Y, Yao JT (2020) Feature extraction with TF-IDF and game-theoretic shadowed sets. Information Processing and Management of Uncertainty in Knowledge-Based Systems, Communications in Computer and Information Science 1237:722–733
8. Bommert A, Sun X, Bischl B, Rahnenführer J, Lang M (2020) Benchmark for filter methods for feature selection in high-dimensional classification data. Comput Stat Data Anal 143:1–19
9. Kadriu A, Abazi L, Abazi H (2019) Albanian text classification: bag of words model and word analogies. Business Systems Research Journal 10(1):74–87
10. Zhao Y, Bo B, Feng Y, Xu C, Yu B (2019) A feature extraction method of hybrid gram for malicious behaviour based on machine learning. Security and Communication Networks 1–8. https://doi.org/10.1155/2019/ 2674684
11. Vidhya S, Asir Antony Gnana Singh D, Leavline EJ (2015) Feature extraction for document classification. International Journal of Innovative Research in Science, Engineering and Technology 4(6):50–56
12. Pahwa B, Taruna S, Kasliwal N (2018) Sentiment analysis—strategy for text pre-processing. International Journal of Computer Applications 180:15–18
13. Singh J, Gupta V (2016) Text stemming: approaches, applications, and challenges. ACM Comput Surv 49(3):1–46
14. Lin D (1988) An information-theoretic definition of similarity. 15th international conference on machine learning. Morgan Kaufmann, San Francisco, CA, pp 296–304
15. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research 15:1929–1958
16. Akhiat Y, Chahhou M, Zinedine A (2019) Ensemble feature selection algorithm. International Journal of Intelligent Systems and Applications 1:24–31

The Impact of the Development of Learning Technology Media on the Learning Process for High School Students

Muhamad Farrel Akbar, R. Sannas Salsabila, Muhammad Farrel Wahyudi, Anastasya Putri Maharani Dewi, Ford Lumban Gaol, Tokuro Matsuo, and Fonny Hutagalung

Abstract The development technology at this time had a complex impact on human life, and education was no exception. The education world at this time has used learning technology media to improve its quality, including high school education. In order for this utilization process to operate effectively, various operations of the learning system are of course required. As we all know, the learning system used to only focus on the teacher, who would explain the material directly to the students. However, with the development of learning technology media, students are expected to become more active in the learning process. The purpose of this study was to understand how learning technology media can help the learning process in high school. The results suggest that the development of learning media technology can stimulate students' creativity, as well as assist the learning process, because apart from being a learning medium, this technology can also be used as a tool and media

M. F. Akbar (✉) · R. S. Salsabila · M. F. Wahyudi · A. P. M. Dewi
School of Information System, Bina Nusantara University, Alam Sutera, Indonesia
e-mail: Muhamad.akbar005@binus.ac.id

R. S. Salsabila
e-mail: sannas.salabila@binus.ac.id

M. F. Wahyudi
e-mail: muhammad.wahyudi001@binus.ac.id

A. P. M. Dewi
e-mail: anastasya.dewi001@binus.ac.id

F. L. Gaol
Binus Graduate Program, Jakarta, Indonesia
e-mail: fgaol@binus.edu

T. Matsuo
Advanced Institute of Industrial Technology, Tokyo, Japan
e-mail: matsuo@aiit.ac.jp

F. Hutagalung
University of Malaya, Kuala Lumpur, Malaysia
e-mail: fonny@um.edu.my

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_58

for storing and distributing material. So that the development of learning technology media can improve the quality of education in high school.

Keywords Learning technology media · Learning process · Creativity · High school

1 Introduction

As social beings, obviously, humans need one another, which causes humans to always be associated with each other where for this situation, great correspondence will be exceptionally valuable [4]. Due to course, fundamentally, humans need each other to address their issues, beginning from addressing the requirements of attire, lodging, and food. With the rapid development of learning technology media at this time, this can help human life, especially in the field of education where learning technology media was the facility and infrastructure such as hardware, software, and useware. The infrastructure is the system and method used for several stages [5] such as obtaining, sending, processing, managing, storing, organizing, and using data meaningfully, Warsita [19: 135].

Communication plays an important role in everyday life, Bernard Berelson and Gary A. Steiner in the book Human Behavior: An Inventory of Scientific Finding (1964) states that communication is a process of transmitting information, ideas, emotions, skills, and others through the use of words, numbers, symbols, pictures, and so on. So, it can be said that if no communication occurs, this will have an impact on the emergence of disharmony and incompatibility [4]. For example, poor communication often leads to misunderstandings that result in disputes; therefore, good communication will be very useful in the field of education. In order for students to understand the subject perfectly, of course, good communication from teachers and students is needed, to avoid possible misunderstandings [7].

Information and communication technology has had a significant impact on education in Indonesia. As stated by Sondang P. Siahaan (2004), there are three functions of electronic learning toward learning in the classroom, namely as an operational supplement, a complement, and a substitute [6].

1. Supplements

It is said to work as a supplement if students have the freedom to choose whether to use electronic learning materials or not. In this case, there is no obligation or requirement for students to access electronic learning materials, but students who use it will certainly have additional knowledge or insight.

2. Complement

Electronic learning materials are programmed to complement the learning materials received by students. As a complement means that electronic learning is programmed to complement enrichment or improvement materials. It is said to be enriching if students can quickly master or understand the subject matter delivered face-to-face

and are given the opportunity to access electronic learning materials that were specifically developed for them. The goal is to increase the level of mastery of the subject. It is said to be a remedial program if students who have difficulty understanding the subject face-to-face are given the opportunity to take advantage of electronic learning materials that are specifically designed for them. The goal is for students to more easily understand the subject presented in class.

3. Substitution

It is said to be a substitution if e-learning is carried out as a substitute for learning activities, for example, by using models of learning activities. There are three alternative models to choose from, namely fully face-to-face, partly face-to-face and partly via the Internet, or even completely via the Internet.

The development of learning technology media at this time provides various conveniences in the world of education. Through the use of learning technology media, this improves the quality of education because students can easily access knowledge that can help the learning process. This is due to the application of high-tech and high-touch approaches that can make education faster, more effective, and efficient [7]. As stated by AETC (2004), educational technology is the study and ethical practice of facilitating learning and improving performance by creating, using, and managing appropriate technological processes and resources [10].

As stated by Eric Ashby (1972), at this time, education has entered the fifth revolution. The first revolution is when parents leave their child's education to a teacher; the second revolution is when a child uses written media for learning purposes; the third revolution is when a printing press is used; the fourth revolution is when television and radio are used to spread education; the fifth revolution is when information and communication technology is utilized, especially computers and Internet media in the learning process [11].

With the alignment of the development process of information technology and communication technology, various multimedia learnings have emerged, such as computer-based learning, Web-based learning (e-learning), computer-assisted learning, and Ava-based learning [12]. With the emergence of various information and communication technologies, education can be advanced. As stated by Commission on Instruction Technology (1970), learning technology is a systematic effort to design, implement, and evaluate the entire learning process for a specific purpose and is based on research on the learning process and communication in humans using a combination of human and non-human resources to learn can be effective.

With the emergence of various information and communication technologies can help advance education [3]. As stated by Commission on Instruction Technology (1970), Learning Technology is a systematic effort to design, implement, and evaluate the entire learning process for a specific purpose, and is based on research on the learning process and communication in humans using a combination of human and non-human resources to learn can be effective [6].

As stated by the National Educational Education Plan (2017), at this time, technology can be used as a tool, where students can use technology to process numbers, create graphics, create databases; not only for students, this technology can also be

used by educational institutions for both staff and teachers. Not only that at this time, technology can also function as a learning material (literacy); in this case, information and communication technology can help students to learn, where computers will be programmed to guide students; in other words, technology can be useful as a facilitator, motivator, transmitter, and evaluator [10].

We present this research paper with the aim of finding out how learning technology media supports the learning process in Indonesia using mixed methods. In this research, we will use a mix method, where this research will include qualitative and quantitative data to meet data needs holistically [13]. Therefore, we present a paper with the title "The Impact of the Development of Learning Technology Media on the Learning Process for High School Students."

Problem formulation

1. What is the benefit of Development of Learning Technology Media on the Learning Process for High School Students?
2. Has the performance of Development of Learning Technology Media on the Learning Process for High School Students been maximized?
3. What are the negative impacts arising from the use of Development of Learning Technology Media on the Learning Process for High School Students?

2 Research Technique

In conducting this research, we used a non-probability technique, namely purposive sampling. Where we will conduct an online survey using the Google Form, we will share this survey given to some high school students. The survey we conducted targeted 20 students. Where in this survey, we will ask how information and communication technology affects the learning process, not only that we will also ask what impact is felt by the use of information and communication technology for education. We also carry out this research by conducting a literature study by analyzing various articles and journals on how information and communication technology affects education in Indonesia.

3 Research Method

The data collection method used is to use the mix method. Research will combine qualitative and quantitative data to meet data needs holistically so that the lack of qualitative data can be met with quantitative data, and vice versa where data collection techniques will be developed into two methods (Fig. 1).

The two methods used in this research, namely:

1. Conducting a Literature Study.

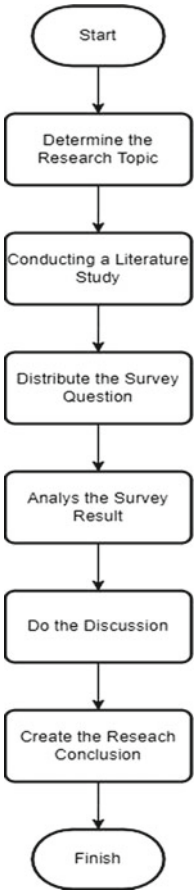


Fig. 1 Flowchart

Analyzing journals and articles on how information and communication technology can play a role in education in Indonesia.

Step 1: looking for previous journals that are related to the title of the research journal we are making
Step 2: find the main point from the previous journal that can help strengthen our argument in this research journal.

2. Survey method.

This is done through a questionnaire given to some high school students. The survey we conducted targeted 20 students. To make the process of distributing questionnaires easier, they will be given in the form of a Google Form. The questionnaire link will be distributed to group chats consisting of students who meet the criteria

Step 1: determine the target audience who will answer the survey questions
Step 2: make a list of questions for each member five questions

Step 3: sorting out the best questions from a combination of all the questions that have been made by each member
Step 4: start preparation by entering all the questions that have been sorted into the Google Form
Step 5: re-check the questions and answer choices that have been inputted
Step 6: The survey is ready to be distributed.

4 Research Result

Regarding the first statement, "The use of learning technology media helps me in the learning process," 5% of the students strongly disagrees that learning technology media helps them in the learning process; 10% of the student chooses disagree; 5% of the student chooses neutral; 45% of the student chooses agree, and 35% of the student chooses strongly agree because with the learning technology media, this can help students to access learning material easily. From the data obtained, it can be concluded that the student agrees if learning technology media helps in the learning process (Fig. 2).

Regarding the second statement, "School performance in utilizing learning technology media is maximized," 5% of the students strongly disagrees that school performance in utilizing learning technology media is maximized; 20% of the student chooses disagree; 25% of the student chooses neutral; 35% of the student chooses agree, and 15% of the student chooses strongly agree because at this time, many schools have supported the use of learning technology media, such as the use of various applications that can facilitate the learning process, not only that at this time, schools are starting to provide Wi-Fi that can be used by students for learning purposes. From the data obtained, it can be concluded that the student agrees if school performance in utilizing learning technology media is maximized (Fig. 3).

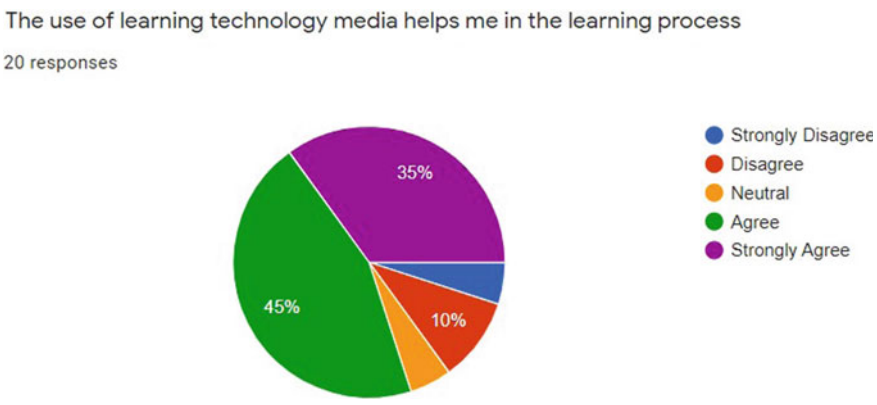


Fig. 2 Question1

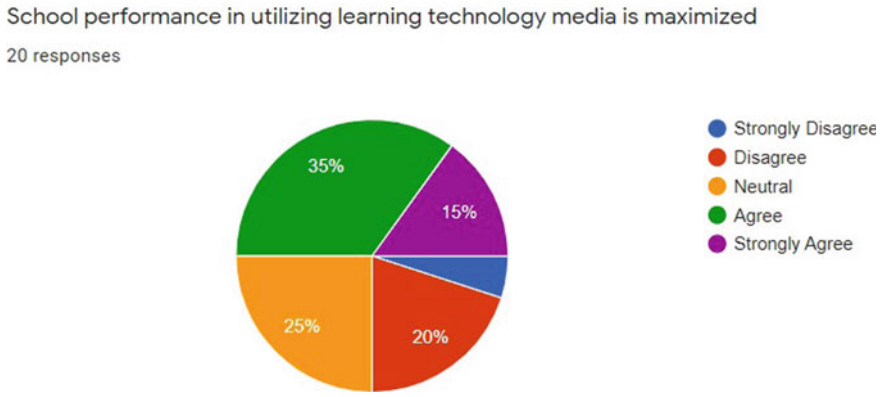


Fig. 3 Question2

Regarding the third statement, "My school has used learning technology media in the learning process," 5% of the students strongly disagrees that their school has used learning technology media in the learning process; 5% of the student chooses disagree; 5% of the student chooses neutral; 55% of the student chooses agree; and 30% of the student chooses strongly agree because as can be seen at this time, many schools have utilized learning technology as online learning media, such as online learning and learning using video and audio. From the data obtained, it can be concluded that the student agrees if learning technology media helps in the learning process (Fig. 4).

Regarding the fourth statement, "All teacher have understood and utilized learning technology media in the learning process," 35% of the student chooses disagree that all teachers have understood and utilized learning technology media in the learning process; 15% of the student chooses neutral; 30% of the student chooses agree, and

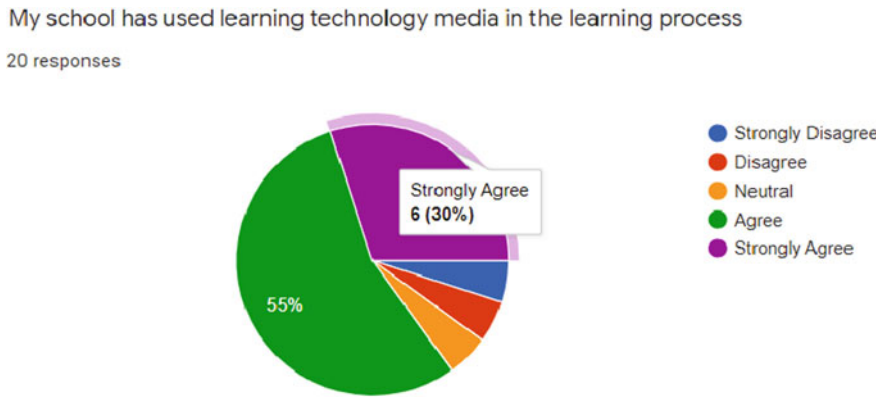


Fig. 4 Question3

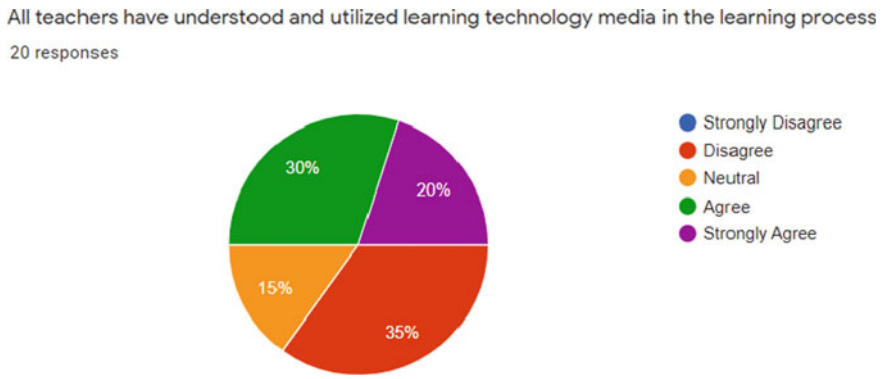


Fig. 5 Question4

20% of the student chooses strongly agree because as we know, even though there are many teachers who understand learning technology media, there are still some teachers, especially generation x teachers who do not understand the development of technology at this time, so they cannot take advantage of learning technology media properly. From the data obtained, it can be concluded that the student disagrees if all teachers have understood and utilized learning technology media in the learning process (Fig. 5).

Regarding the fifth statement, "I feel that the use of learning technology media has a negative impact on me," 30% of the students strongly disagrees that the use of learning technology media has a negative impact; 25% of the student chooses disagree; 10% of the student chooses neutral; 15% of the student chooses agree, and 20% of the student chooses strongly agree because even though learning technology media has some negative impacts, this depends on each individual in responding to it; in other words, if students are wise in using learning technology media, they can avoid various negative impacts. From the data obtained, it can be concluded that the student strongly disagrees if the use of learning technology media gives them a negative impact (Fig. 6).

Learning technology media is a way to carry out learning activities without having to be in the classroom, also known as e-learning. Learning media technologies enable students and teachers to fulfill their obligations when it is not possible so that students' rights are not compromised. The benefit we can get from learning technology media is the flexibility to choose when and where to access the material. In addition, learning technology media provides learners with the opportunity to independently control their learning success. Unfortunately, the application of learning media technologies does not work optimally for several reasons, namely the need for adequate infrastructure and the learning outcomes that will be achieved. Results are up to the students themselves, so they don't necessarily get the best results. And learning media technology doesn't work for everyone's learning style. It is undeniable that even when it can be optimized to run, media technology learning can still have a negative impact on students. Some of these negatives are having difficulty understanding existing

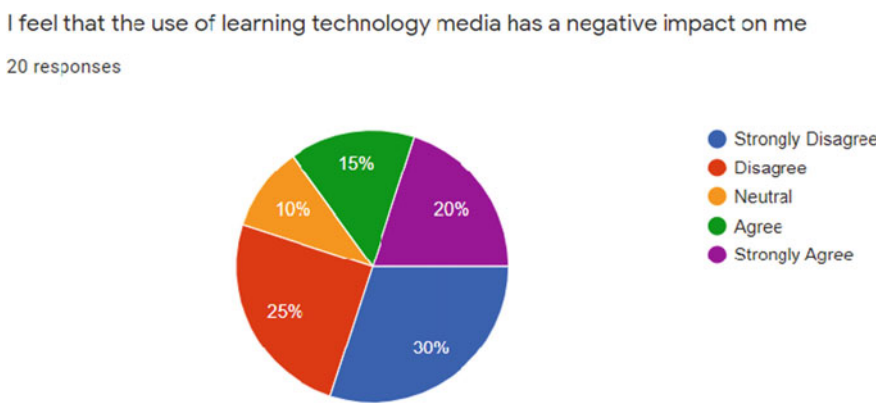


Fig. 6 Question5

systems in media technology learning, creating confusion, and ultimately causing some students to fall behind in the classroom, while for students whose families are in non-strategic areas, media technology learning will be difficult to access because of the need for a good Internet. Stable is able to access learning without buffering or lag.

5 Discussion

From the research that we have done, we discuss the benefit of learning technology media, the performance, and the negative impact from the learning technology media. The result of our discussion is that the student agrees that learning technology media helps in the learning process; the student agrees that school performance in utilizing learning technology media is maximized; the student agrees that learning technology media helps in the learning process; the student disagrees that all teachers have understood and utilized learning technology media in the learning process, and the student strongly disagrees that the use of learning technology media gives them a negative impact.

The results of our discussion are in line with research conducted by Kaite J. Carstens, Jamie M. Mallon, Mohamed Bataineh, and Adel Al-Bataineh titled "Effects of Technology on Student Learning." They developed a survey to look at how technology affects learning on students. Our research shows that 35% of the students thinks that teachers haven't utilized the technology at best, while 30% students agree, and 20% students strongly agree that teachers have utilized the technology. This means 50% of the students is giving positive answers. The comparison shows that it is true that teachers have utilized the technology in their teaching and spent most of their time on led learning. 45% of the students agrees that technology helps

them learn, which is in line with the comparison, showing that 70.3% of the teachers agrees that students are probably relying on the technology.

In addition, the results of our discussion are also in line with research conducted by Sudi Suryadi entitled "The Role of Information and Communication Technology Development in Learning Activities and the Development of the World of Education" states that the quality of education can be done by utilizing communication and information technology, namely by identifying the problems encountered and conducting searches to solve them. Problems through information technology, with the convenience obtained in using Internet media of course, this can make the learning process more effective, but in order for this technology to run really effectively, self-awareness is needed to use Internet media as needed to avoid misuse.

So that we can determine whether learning technology media allows learners to independently control their learning success. The application of learning media technologies does not work optimally for several reasons, namely the need for adequate infrastructure and the learning outcomes that will be achieved. But sadly, learning media technology doesn't work for everyone's learning style.

6 Conclusion

The benefit of development of learning technology media on the learning process for high school students is that they get the opportunity to independently control their learning success. Students would be able to choose freely when and where to access the material. This is because the learning material is always available on the Internet.

The performance of development of learning technology media on the learning process for high school students has been maximized based on what the students are experiencing. Many schools have advanced their use of technology so that the learning activities can be at maximum effectivity. Some schools have enabled the online learning system and are using the technology to help both students and teachers to do their work.

The negative impacts arising from the use of development technology media on the learning process for high school students are the probability of falls happening in some students. Technology is not something that everyone could use immediately. People would need to learn how to use it properly before being able to get the benefits. This process might take time for the students, especially for those who weren't used to using technologies. Some students might get the impact of not being able to learn effectively because they don't have a stable Internet and good devices, which are essential for online learning.

Acknowledgements The research is funded Binus University under contract number No: 017/VR.RTT/III/2021.

References

1. Administrator (2020) Peranan Teknologi Informasi Dalam Dunia Pendidikan. Berita Kedinasan. Retrieved from https://diskominfo.kedirikab.go.id/baca/peranan-teknologi-informasi-dalam-dunia-pendidikan
2. Affandi M (2017) Teknologi informasi & komunikasi dalam pendidikan. Penerbit YNHW, Kuningan
3. Agusten AD (2020) Dampak Positif dan Negatif Pembelajaran Daring Menggunakan E-Learning, 20 Nov 2020. Retrieved from https://www.hmjmpiuinmaliki.or.id/2020/11/dampak-positif-dan-negatif-pembelajaran.html
4. Andriani T (2015) Sistem Pembelajaran Berbasis Teknologi Informasi Dan Komunikasi. Sosial Budaya 12. Retrieved from http://ejournal.uin-suska.ac.id/index.php/SosialBudaya/article/view/1930
5. Asmawi Asmawi SS (2019) Pendidikan Berbasis Teknologi Informasi Dan Koomunikasi. Retrieved from https://jurnal.univpgri-palembang.ac.id/index.php/Prosidingpps/article/view/2930
6. Budiman (2017) Peran Teknologi Informasi Dan Komunikasi Dalam Pendidikan. Al-Tadzkiyyah: Jurnal Pendidikan Islam, 8. Retrieved from http://ejournal.radenintan.ac.id/index.php/tadzkiyyah/article/view/2095
7. Cholik CA (2017) Pemanfaatan Teknologi Informasi Dan Komuniakasi Untuk Meningkatkan Pendidikan Di Indonesia. Syntax Literate, 2. Retrieved from https://www.jurnal.syntaxliterate.co.id/index.php/syntax-literate/article/view/130
8. Harahap L (2021) Peran Teknologi Informasi Dan Komunikasi. Retrieved from http://digilib.unimed.ac.id/38785/2/ATP%2045.pdf
9. Hasan A (2022) Pengertian, Karakteristik Dan Manfaat E-Learning. Retrieved from http://www.smkbkujkt.sch.id/read/16/pengertian-karaktiristik-dan-manfaat-eLearning
10. Hermaw R (2021) Penerapan Teknologi Informasi Dan Komunikasi Dalam Pendidikan, 18 Feb 2021. Retrieved from https://bdkpalembang.kemenag.go.id/upload/files/Pemanfaatan%20TIK%20dalam%20pendidikan%20%281%29%281%29.pdf
11. Ismaniati C (2019) Penggunaan Teknologi Informasi Dan Komunikasi Dalam Peningkatan Kualitas Pembelajaran. Retrieved from http://staff.uny.ac.id/sites/default/files/penelitian/Dr.%20Christina%20Ismaniati,%20M.Pd./Penggunaan%20Teknologi%20Informasi%20dan%20komunikasi%20dalam%20peningkatan%20kualitas%20pembelajaran.pdf
12. Jančařík KJ (n.d.) Teaching Aids and Work With Models in e-Learning Environments. Retrieved from https://files.eric.ed.gov/fulltext/EJ1146047.pdf#:~:text=Eric%20Ashby%20%281972%29%20speaks%20of%20four%20revolutions%20in,revolution%20associated%20with%20the%20use%20of%20educational%20technologies
13. Maulida M (2020) Teknik Pengumpulan Data Dalam Metodologi Penelitian. Jurnal Online IAI Darussalam 21. Retrieved from http://ojs.iai-darussalam.ac.id/index.php/darussalam/article/view/39
14. Organizing Academic Research Papers: Theoretical Framework. (n.d.). Library Sacred Heart UNIVERSITY. Retrieved from https://library.sacredheart.edu/c.php?g=29803&p=185919
15. Teknologi Pembelajaran, Contoh Media dan Fungsinya (n.d.) Suka Dunia NET. Retrieved from https://sukadunia.net/teknologi-pembelajaran/
16. urvashi151993 (2016, Maret 19) Educational Technology. Retrieved from https://urvashi151993.wordpress.com/2016/03/19/educational-technology/comment-page-1/#:~:text=Educational%20technology%20is%20defined%20by%20the%20Association%20for,using%2C%20and%20managing%20appropriate%20technological%20processes%20and%20resources.%E2%
17. Utami SN (2021) Komunikasi: Pengertian Para Ahli, Fungsi, Tujuan, dan Jenis-Jenisnya. Retrieved from https://www.kompas.com/skola/read/2021/08/05/120000469/komunikasi--pengertian-para-ahli-fungsi-tujuan-dan-jenis-jenisnya?page=all

18. Wafiy (2019) Kekurangan Dan Kelebihan Teknologi Informasi Dan Komunikasi. Retrieved from https://www.puplas.com/kekurangan-dan-kelebihan-teknologi-informasi-dan-komunikasi/
19. Warsita B (2008) In Teknologi Pembelajaran: Landasan Dan Aplikasinya. Rineka Cipta, Jakarta
20. William A (2021) Mengenal Teknik Pengumpulan Data Kualitatif dan Kuantitatif. Retrieved from https://tirto.id/mengenal-teknik-pengumpulan-data-kualitatif-dan-kuantitatif-ga1i
21. Yaumi M (2017) Media Pembelajaran: Pengertian, Fungsi, dan Urgensinya bagi Anak Milenial. Retrieved from http://repositori.uin-alauddin.ac.id/11788/1/MEDIA%20PEMBELAJARAN%20Pengertian%2C%20Fungsi%2C%20dan%20Urgensinya%20bagi%20Anak%20Milenial.pdf

Impacts of Vibration Mode Switching on Energy Dissipation Analysis of Rectangular Microplate Resonator-based Sensors in IoT Applications

R. Resmi, V. Suresh Babu, and M. R. Baiju

Abstract Microelectromechanical systems (MEMS) technology is extensively used for making high-performance sensors and actuators in Internet of Things. The ubiquitous advantages and potential applications of MEMS sensors compared to conventional ones paved the way for implementing the diverse micro-sensing technologies in Internet of Things (IoT). Vibrating plate-based MEMS resonant sensors are the most commonly used small structural elements due to its low mass and high-quality factors. The downsizing of devices for achieving minute sensors used in IoT leads to application of higher-order theories. The requirement of enhanced quality factors necessitates the study of characteristics affecting the quality factor and associated energy dissipation. The current analysis explores the significance of mode switching on energy loss which controls the most decisive performance measure for sensors used in IoT applications. The investigation of impact of mode switching on thermoelastic energy dissipation with different structural materials for a rectangular plate-based resonator is included in the study. The energy dissipation is verified to be slightly diminished when operated in higher modes, and the quantification of the impact is accurately done by calculating the percentage reduction in energy losses. While considering the thermoelastic energy dissipation of microplate-based resonators, the influence of mode switching on diamond-based plate resonators is verified to be the most prominent according to our findings. The result analysis of the proposed work helps the engineers for designing resonators with diamond as the structural material operating in higher vibrating modes for IoT and 5G applications.

Keywords Micro/nanoplate rectangular resonators · Energy dissipation · Size effects · Mode switching

R. Resmi (✉)
University of Kerala, LBS Institute of Technology for Women, Poojappura, Trivandrum, India
e-mail: resmilbs@gmail.com; resmi@lbsitw.ac.in

V. Suresh Babu
College of Engineering Trivandrum, APJ Abdul Kalam Technological University, Trivandrum, Kerala, India

M. R. Baiju
University of Kerala, Kerala Public Service Commission, Trivandrum, Kerala 695004, India

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_59

1 Introduction

In modern era, ubiquitous computing also known as pervasive computing plays an important role in day-to-day life by putting forward enormous applications and benefits to people [1–3]. The major advantages of pervasive computing lie in developing computing systems with cost reduction, faster response, improved functionality using smart networks, productivity enhancement, etc. [4–6]. Ubiquitous computing can be made reachable to public with sensor fusion, data networking, and data analytics technologies for several applications such as in tracking and management of resource usage, providing medical and education services both in urban and rural areas, preventing errors in a work environment, detecting unwanted gas emissions for ensuring safe conditions [7, 8].

Ubiquitous computing necessitates the demand for associating various devices like sensors and networking components to communicate relevant information without any interruptions. The pervasive computing technology usually includes wireless sensors and communication devices, self-driving vehicles, smartphones, portable computers, embedded systems, transportable devices, RF ID tags, smart home gadgets, etc. [9, 10].

Due to the enormous advantages of micro/nanomechanical resonators, they have been used for sensing and communication applications in the area of ubiquitous computing. Micromechanical resonators have numerous advantages such as high surface-to-volume ratio, peak sensitivity, highly economical, and low power requirement which finds its usage in pervasive computing applications [11–13]. The traditional sizeable off-chip components in the computing network can be replaced with MEMS/NEMS components.

The performance of a MEMS-based resonator is mainly dictated by its quality factor which designates the energy dissipation in the system [14, 15]. In MEMS resonators, various energy loss mechanisms exist which can be classified as extrinsic and intrinsic types. Both types of loss mechanisms are limiting the quality factor, and the necessity of mitigating the energy dissipation is very crucial for developing resonators with low losses [16, 17]. Air damping, squeeze film damping, and anchor damping come under extrinsic-type energy dissipation mechanisms. Air damping arises due to the collision with the air particles during vibrations and can be eliminated by ensuring vacuum operation [18]. Squeeze film damping is a type of extrinsic damping arising in movable structures due to the influence of the repelling force of air on the movable structures, when the air is squeezed in between the boundary layers [19]. Anchor damping exists when energy flows irreversibly to the supporting structures [20]. The dominant extrinsic losses like squeeze film damping and anchor damping can be eliminated easily by proper geometric design.

Thermoelastic damping (TED) is recognized as the most significant intrinsic energy loss mechanism which limits the total quality factor of the resonator to a large extent. TED occurs due to the temperature gradients resulting from the strain gradients generating from the flexing of resonating structures. Due to the difference in temperature along the structure of the resonator, thermal transport from hotter

to colder region takes place leading to irreversible energy dissipation. The rate of energy loss depends on the time constant of the thermal transport, and if it is close to the time period corresponding to the vibrational frequency, energy loss will be maximized. Hence, to design resonators with low energy dissipation, the vibrating frequency is very significant which determines the losses [21].

Thermoelastic damping (TED) phenomenon in Euler Bernoulli beams was first identified by Zener, who formulated an analytical expression for the energy dissipation related to TED [22, 23]. The more precise expression for thermoelastic energy dissipation in microbeams is governed by Lifshitz and Roukes [24]. The mechanical analyses of damping in other vibrating microstructures such as microplates and microshells are also organized by several works. TED study and quality factor analysis in microplates with various structural mode shapes are performed by Nayfeh and Younis [25]. Nayfeh derived an expression for damping-related energy loss in thermoelastic plates using perturbation method. The experimental validation of thermoelastic damping is done in many works conducted by Amy Duwell et al., Muller et al., etc., and consequently, TED can be considered as one of the critical issues to be solved [26, 27].

Primarily, TED investigation has been done by applying classical elasticity theory, but to model thermoelastic damping in structural domain for micron scale devices, nonclassical elasticity theories are required. The size dependence effect has been experimentally reported in many works [28]. Consequently, several higher-order elasticity theories like strain gradient theories and coupled stress theories are developed to detain the size effects. The most advantageous theories are coupled stress theories which need only fewer material length scale parameters to model it. Modified couple stress theory (MCST) is the most extensively used nonclassical theory due to its simplicity with only one material length scale parameter to incorporate the scaling effects [29].

Energy dissipation analysis in vibrating resonators with size effects is successfully implemented by MCST in many works. MCST is introduced by Yang et al. with a single length scale parameter value and found to be more accurate than other nonclassical elasticity theories [30]. The significance of length scale parameter is studied by Razavilar et al. applying MCST [31]. Figure of merit analysis of rectangular microplate-based resonators is investigated by Resmi et al. [32]. Chen et al. analyzed the mode switching of high-performance MEMS resonators for frequency generations [33]. The modal analysis in thermoelastic damping with electrostatic actuation in MEMS resonators is conducted by Serdean et al. [34]. Duwel et al. developed MEMS resonators with low thermoelastic damping and verified its mode dependency [35]. Yi et al. presented the contour plots of the temperature field of microrings with various mode shape patterns [36].

Kumar et al. investigated TED in microplates utilizing MCST and the three-phase-lag heat conduction model [37]. Devi S et al. studied the frequency shift related to thermoelastic damping in Kirchhoff's microplates based on MCST [38]. Thermoelastic damping in microplates based on CST and dual-phase conduction model is done by Vahid Borjalilou et al. [29]. Applying MCST, Resmi R et al. studied

the influence of dimensionless length scale parameter on thermoelastic attenuation and frequency shifts of microplate resonators [39].

Thermoelastic damping analysis in nanoresonators utilizing MCST and memory-dependent heat conduction model is provided by Zhao et al. [40]. Size-dependent thermoelastic damping is analyzed by Zhong et al. using modified couple stress theory, and the temperature dependency is also verified [41]. As the scaling effect is included, bending rigidity ratio and resonance frequency increase which leads to reduced thermoelastic energy dissipation. When higher vibrating modes are used, the relaxation time of the resonating structures diminishes which reduces the energy dissipation. In the current analysis, simply supported microplates are considered and the structural materials selected are Si, polySi, diamond, GaAs, and SiC. The impacts of mode switching are analyzed by taking the percentage difference of the values of Q^{-1} obtained with Mode I and Mode II under the same stipulated conditions.

In this paper, the significance of mode switching on thermoelastic energy dissipation of a vibrating simply supported rectangular microplate applying size effect is analyzed. To include size effects, a nonclassical theory like modified couple stress theory (MCST) is used in the analysis. The expression for thermoelastic energy dissipation is given in Sect. 2, and results and discussions are presented in Sect. 3. The numerically simulated values of thermoelastic energy dissipations of microplates using five different structural materials with mode switching are presented in Sect. 3.1. The quantification of the impact of mode switching on energy loss is accurately done by evaluating the percentage change in the energy dissipation as discussed in Sect. 3.2. The conclusions of the work are given in Sect. 4.

2 Expression for Thermoelastic Energy Dissipation of a Rectangular Microplate

Size effects are included in the analysis by applying a nonclassical theory like modified couple stress theory, where the strain energy density is a function of both strain and curvature tensors. The impact of scaling effect on thermoelastic energy dissipation is very remarkable, and Q_{TED} is very much enhanced. The tendency shows a large amount of trimming in energy dissipation due to thermoelastic damping, and the influence of mode change can be investigated from the thermoelastic energy dissipation along with size effect variation. To analyze the impact of mode switching of microplates, two modes are selected for analysis Mode I (1,1) and Mode II (1,2).

The amount of thermoelastic damping, expressed in terms of the inverse of the quality factor, is given by

$$Q^{-1} = 2\left|\frac{|\mathrm{Im}(\omega)|}{\mathrm{Re}(\omega)}\right| \tag{1}$$

3 Results and Discussions

The thermoelastic energy dissipation analysis in vibrating microplates with length $L = 200$ μm, width $W = 200$ μm, and thickness $h = 10$ μm is conducted by numerical simulations using MATLAB 15a. The microplates chosen for investigation are Kirchhoff's plates with simply supported boundary type and five diverse structural materials—Si, SiC, GaAs, polySi, and diamond. The size effect is incorporated by including different values for the material length scale parameter (l/h). The main investigation is focused to find out the impact of vibrational mode switching on thermoelastic energy dissipation of rectangular plates-based resonators in micron scales. Table 1 shows the energy loss occurred in the microplates under size scaling with two chosen modes [M I (1,1) and M II (1,2)]. The mechanical and thermodynamic properties of all the structural materials used in the rectangular plates are taken from [32].

3.1 Thermoelastic Energy Dissipation Analysis with Modes Switching

In the current analysis, energy dissipation investigation of a thermoelastic rectangular microplate for the first two modes (1,1) and (1,2) is analyzed. The energy loss for different structural materials under various conditions is given in Table 1.

The simulations are conducted for microplates with two different boundary types (simply supported and clamped–clamped) and various dimensionless length scale values ($l/h = 0$, 0.5, and 1). The numerically simulated values are obtained for two vibrating modes −(1,1) and (1,2) as depicted. The study is based on the interest of exploring the material dependency on energy losses with different vibration modes. Accordingly, the analytical expressions for thermoelastic energy dissipation are numerically simulated for microplates with all the five chosen structural materials Si, polySi, GaAs, diamond, and SiC under the stipulated conditions.

Table 1 Thermoelastic energy dissipation of a microplate for different materials with two vibrating modes (1,1) and (1,2)

Material		PolySi	Diamond	Si	GaAs	SiC
$l/h = 0$	Mode I	3.55E−05	1.91E−04	1.70E−04	2.86E−04	3.72E−04
	Mode II	3.54E−05	1.89E−04	1.66E−04	2.85E−04	3.70E−04
$l/h = 0.5$	Mode I	1.73E−05	1.00E−04	1.02E−04	1.82E−04	2.79E−04
	Mode II	1.72E−05	9.72E−05	1.01E−04	1.82E−04	2.62E−04
$l/h = 1$	Mode I	6.80E−06	5.00E−05	5.30E−05	9.60E−05	1.01E−04
	Mode II	6.75E−06	4.59E−05	5.21E−05	9.66E−05	1.00E−04

While considering thermoelastic damping, the energy dissipation for different modes under the same conditions such as structural materials and dimensionless length scale parameters is only slightly different. The values of energy losses for different constraints simulated numerically are plotted as a bar chart as shown in Figs. 1 and 2.

Figure 1 illustrates the variation of energy dissipation due to thermoelatic damping for microplates with various structural materials, vibrating in first mode (1,1). The important performance parameter, thermoelastic damping limited quality factor (Q_{TED}), is inversely related to the corresponding energy dissipation.

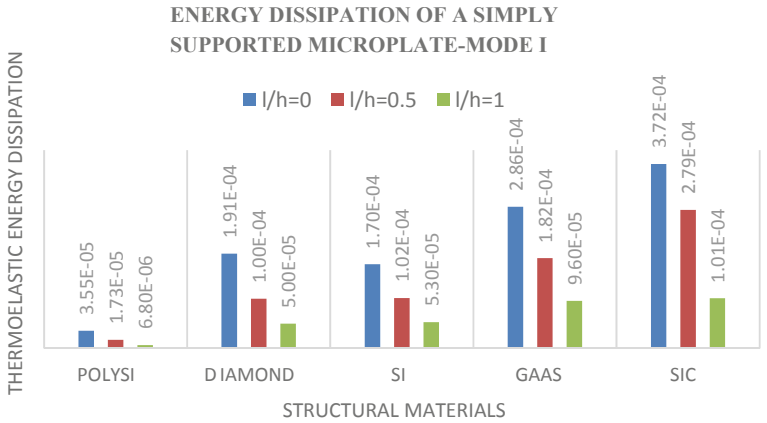


Fig. 1 Variation of energy dissipation, Q^{-1} for a rectangular microplate with different structural materials (polySi, diamond, Si, GaAs, and SiC) vibrating in the first mode [Mode I (1,1)]

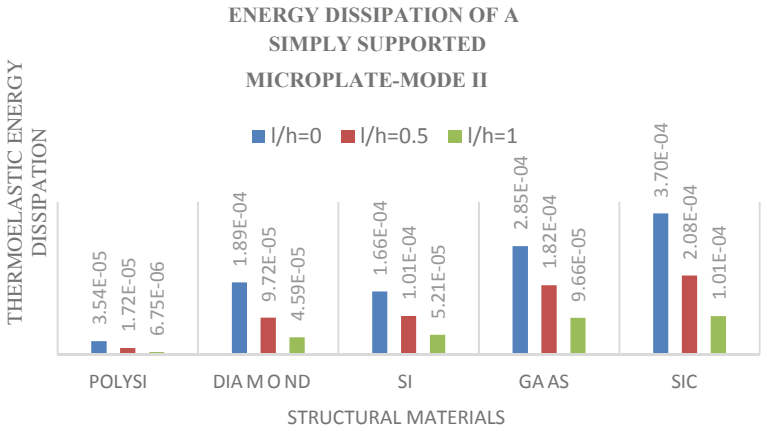


Fig. 2 Variation of energy dissipation, Q^{-1} for a rectangular microplate with different structural materials (polySi, diamond, Si, GaAs, and SiC) vibrating in second mode [Mode II (1,2)]

From Fig. 1, it is verified that the disparity in energy is highly dependent on the type of the structural materials used. The energy loss is maximum for SiC-based microplates, and the minimum is obtained for microplates with polySi as the structural material. The thermoelastic energy dissipation of plates with simply supported boundary condition is plotted for all the five structural materials and different values of dimensionless length scale parameters (l/h = 0,0.5,1). The structural material order in which thermoelastic energy dissipation diminishes is SiC > GaAs > Si > diamond > polySi when MCST is applied, i.e., when size effect is included.

The variation of thermoelastic energy dissipation of the microplates vibrating in second mode (1,2) is portrayed as a bar chart in Fig. 2. The changes observed for the microplates with size scaling and structural materials are in a similar manner. Only a slight change in the values is attained for the energy losses under different modes. The energy loss due to thermoelastic damping is slightly diminished when switched to a higher mode. The switching modes under consideration are Mode I (1,1) and Mode II (1,2) as shown before. The impact on mode switching can be concluded as a very slight decrement when the mode is switched from (1,1) to (1,2). Consequently, the enhancement in Q_{TED} is also not much substantiable.

Figure 2 illustrates the variation of thermoelastic energy dissipation, Q^{-1} for a rectangular microplate with various structural materials such as polySi, diamond, Si, GaAs, and SiC vibrating in (1,2). Hence, to analyze the material dependency on mode switching, percentage change in energy dissipation is calculated associated with the three l/h values for microbeams and the chosen structural materials as presented as in Sect. 3.2.

3.2 Impact Analysis of Modes Switching on Thermoelastic Energy Dissipation and Its Material Dependency

The values of energy dissipation obtained for the microplate with different structural materials applying MCST are discussed in Sect. 3.1. While analyzing the energy losses associated with the two consecutive modes, (1,1) and (1,2), it is substantiated that there is not much difference in numerically simulated values under the same conditions. For example, the value of thermoelastic energy dissipation of a polySi-based simply supported microplate with $l/h = 1$ is 6.80E−06 and under the same stipulated boundary condition (simply supported) and dimensionless length scale value, $Q^{-1} = 6.\ 75\mathrm{E}{-}06$. On analyzing the value of energy losses for two vibrating modes, under the same conditions, for all structural materials, it is verified that there is reduction in energy dissipation value as the vibrating mode is changed from Mode I to Mode II. The decrease in the value of energy losses is small so that the precise impact on mode switching is difficult to analyze.

In order to accurately quantify the impact of mode switching, percentage change in energy dissipation associated with mode switching is calculated for all the structural materials under the same chosen boundary conditions and l/h values. Hence, the

percentage change in reduction of energy dissipation obtained for changing the mode from (1,1) to (1,2) is calculated as discussed and presented in Table 2. The exhibited values are expressed as a percentage for each material. The detailed analysis is done subsequently and proved that the impact of mode switching on energy dissipation is very small for some structural materials, even though a slight decrement is obtained as shown in Fig. 3.

Figure 3 depicts the percentage change in energy dissipation of a rectangular microplate using five different structural materials (polySi, Si, GaAs, diamond, and SiC), while vibrating mode is switched from lower to a higher mode. The maximum percentage reduction (8.116%) in energy loss is achieved for a diamond-based simply supported microplate with $l/h = 1$.

The minimum reduction in percentage change is obtained for a simply supported GaAs-based microplate with $l/h = 0.5$. From Fig. 3 and Table 2, the material order in which the diminishing of percentage change in energy dissipation occurs for different structural materials is diamond > SiC > Si > GaAs > polySi as verified from Table 2.

Table 2 Percentage change (%) in thermoelastic energy dissipation of a microplate for different structural materials with mode switching—Mode I to Mode II transition

l/h		0	0.5	1
Mode switching (M I)-(M II)	PolySi	0.365826	0.578035	0.702941
	Diamond	1.139155	2.774	8.116
	Si	2.482353	0.629673	1.698113
	GaAs	0.405892	0.098869	0.580208
	SiC	0.545655	6.173901	0.710339

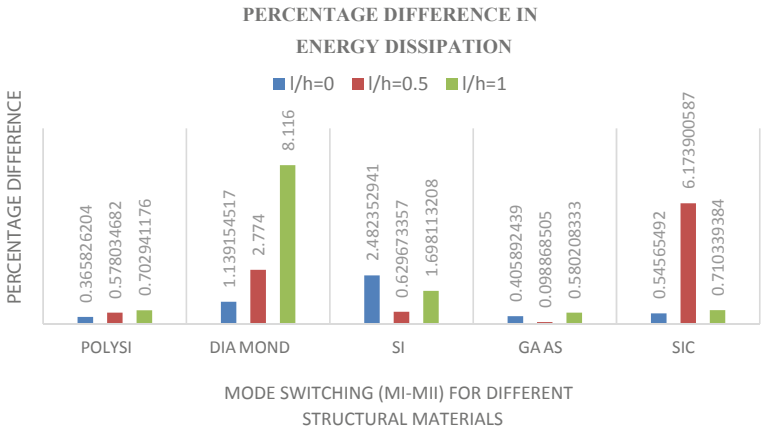


Fig. 3 Percentage change in energy dissipation (Q^{-1}) for a rectangular microplate with different structural materials (polySi, diamond, Si, GaAs, and SiC) when the vibrating mode is switched from Mode I (1,1) to Mode II (1,2)

4 Conclusion

Ubiquitous computing is a leading edge technology boon to people which facilitates computational capability in day-to-day life at anytime and anywhere, for adding quality to life. MEMS-based sensors are the widely used components in pervasive computing which can be realized by micro/nano-based rectangular plate resonators with superior performance and high-quality factors. Among the different energy loss mechanisms, thermoelastic damping is very crucial, and the factors affecting it should be carefully investigated. In this paper, the impact of vibration mode switching on thermoelastic energy dissipation is analyzed for a rectangular microplate with five different structural materials. When the devices are downsized, to accurately model the damping phenomena, size effects are incorporated by including higher-order theories like MCST by assigning different values to the material length scale parameter. The thermoelastic energy dissipation of simply supported microplates with varying *l*/*h* values and various structural materials is numerically simulated for two vibrating modes [(1,1) and (1,2)]. The energy loss is proved to be diminished slightly when the mode of vibration is changed from lower to higher one. Consequently, to accurately evaluate the reduction in energy loss with mode switching, the percentage change in thermoelastic energy dissipation is investigated for all the five structural materials for the simply supported boundary type. The largest change in thermoelastic energy dissipation with mode switching is achieved for diamond-based microplates, and generally, the impact of mode change is found to be not much substantiable. The sequential material order in which reduction in percentage change of energy dissipation occurs is diamond > SiC > Si > GaAs > polySi. The findings can be used for designing microplates with low energy dissipation and suitable structural materials operating in higher vibrating modes which can find extensive applications in 5G networks and IoT.

References

1. Sen J (2012) Ubiquitous Computing: applications. Challenges and Future Trends. https://doi.org/10.1201/b12298-2
2. Smys S, Basar A, Wang H (2020) Hybrid intrusion detection system for internet of Things (IoT). Journal of ISMAC 2:190–199
3. Kumar RP, Smys S (2018) A novel report on architecture, protocols and applications in Internet of Things (IoT). In: 2018 2nd international conference on inventive systems and control (ICISC), pp 1156–1161. IEEE
4. Bashar A (2020) Artificial intelligence based LTE MIMO antenna for 5th generation mobile networks. J Artif Intell 2(03):155–162
5. Devana VN, Rao AM, Rao AM (2020) A compact 3.1-18.8 GHz triple band notched UWB antenna for mobile UWB applications. IRO Journal on Sustainable Wireless Systems 2(1):1–12
6. Zhang G et al (2005) Towards an ubiquitous wireless digital writing instrument using MEMS motion sensing technology. In: Proceedings of 2005 IEEE/ASME international conference on advanced intelligent mechatronics, pp 795–800. https://doi.org/10.1109/AIM.2005.1511080

7. Beigl M, Krohn A, Zimmer T, Decker C (2004) Typical sensors needed in ubiquitous and pervasive computing
8. Gaura E, Newman R (2006) Smart MEMS and sensor systems. Imperial College Press. https://doi.org/10.1142/P355
9. Xiong N, Cheng H, Hussain S, Qu Y (2013) Fault-tolerant and ubiquitous computing in sensor networks. Int J Distrib Sens Netw. https://doi.org/10.1155/2013/524547
10. Lyytinen K, Yoo Y (2002) Ubiquitous computing. Commun ACM 45:63–96
11. Unlu M, Hashemi M, Berry CW, Li S, Yang SH, Jarrahi M (2014) Switchable scattering meta-surfaces for broadband terahertz modulation. Nat Sci Rep 4:5708
12. Meinel K, Melzer M, Stoeckel C, Shaporin A, Forke R, Zimmermann S, Hiller K, Otto T, Kuhn H (2020) 2D scanning micromirror with large scan angle and monolithically integrated angle sensors based on piezoelectric thin film aluminum nitride. Sensors (MDPI) 20(22):6599. https://doi.org/10.3390/s20226599
13. Asri MIA, Hasan MN, Fuaad MRA, Yunos YM, Ali MSM (2021) MEMS gas sensors: a review. IEEE Sens J 21:18381–18397. https://doi.org/10.1109/JSEN.2021.3091854
14. Resmi R, Babu VS, Baiju MR (2021) Analysis of thermoelastic damping limited quality factor and critical dimensions of circular plate resonators based on axisymmetric and non-axisymmetric vibrations. AIP Adv 11:035108. https://doi.org/10.1063/5.0033087
15. Srikar VT, Swan AK, Ünlü MS, Goldberg BB, Spearing SM (2003) Micro-Raman measurement of bending stresses in micromachined silicon flexures. J Microelectromech Syst 12(6):779–788
16. Duwel A, Candler RN, Kenny TW, Varghese M (2006) Engineering MEMS resonators with low thermoelastic damping. J Microelectromech Syst 15(6):1437–1445
17. Ghaffari S et al (2015) Accurate modeling of quality factor behavior of complex silicon MEMS resonators. J Microelectromech Syst 24(2):276–288. https://doi.org/10.1109/JMEMS.2014.2374451
18. Kainz A, Hortschitz W, Schalko J, Jachimowicz A, Keplinger F (2015) Air damping as design feature in lateral oscillators. Sensors and Actuators A: Physical 236. https://doi.org/10.1016/j.sna.2015.11.005
19. Finny S, Resmi R (2016) Material and geometry optimization for squeeze film damping in a micromirror. In: 2016 international conference on emerging technological trends (ICETT), pp 1–5. https://doi.org/10.1109/ICETT.2016.7873698
20. Mol S, Resmi R (2017) Anchor loss limited Q factor analysis of disk resonator for varying disk geometry. In: 2017 international conference on intelligent computing, instrumentation and control technologies (ICICICT), pp 1033–1037. https://doi.org/10.1109/ICICICT1.2017.8342710
21. Kim S-B, Kim J-H (2011) Quality factors for the nano-mechanical tubes with thermoelastic damping and initial stress. J Sound Vibr 330:1393–1402
22. Zener C (1937) Internal friction in solids. I. Theory of internal friction in reeds. Phys Rev 52(3):230–235
23. Zener C (1938) Internal friction in solids II. General theory of thermoelastic internal friction. Phys Rev 53(1):90–99
24. Lifshitz R, Roukes ML (2000) Thermoelastic damping in micro-and nanomechanical systems. Phys Rev B 61(8):5600
25. Nayfe AH, Younis MI (2004) Modeling and simulations of thermoelastic damping in microplates. J Micromech Microeng 14(12):1711–1717
26. Duwel A, Gorman J, Weinstein M, Borenstein J, Ward P (2003) Experimental study of thermoelastic damping in MEMS gyros. Sensors and Actuators A: Physical, Elsevier 103(1–2):70
27. Muller C, Baborowski J, Pezous A, Dubois M-A (2009) Experimental evidence of thermoelastic damping in silicon tuning fork. Procedia Chemistry 1:1395–1398. https://doi.org/10.1016/j.proche.2009.07.348
28. Park SK, Gao X-L (2006) Bernoulli-Euler beam model based on a modified couple stress theory. J Micromech Microeng 16:2355–2359

29. Borjalilou V, Asghari M (2018) Small-scale analysis of plates with thermoelastic damping based on the modified couple stress theory and the dual-phase-lag heat conduction model. Springer GmbH, Austria. https://doi.org/10.1007/s00707-018-2197-0
30. Yang F, Chong ACM, Lam DCC, Tong P (2002) Couple stress based strain gradient theory for elasticity. Int J Solids Struct 39(10):2731–2743. https://doi.org/10.1016/S0020-7683(02)00152-X
31. Razavilar R, Alashti RA, Fathi A (2016) Investigation of thermoelastic damping in rectangular microplate resonator using modified couple stress theory. Int J Mech Mater Des 12(1):39–51
32. Resmi R, Baiju MR, Babu VS (2019) Thermoelastic damping dependent quality factor analysis of rectangular plates applying modified couple stress theory. AIP Conf Proc 2166:020029. https://doi.org/10.1063/1.5131616
33. Chen Z, Kan X, Yuan Q et al (2020) A switchable high-performance RF-MEMS resonator with flexible frequency generations. Sci Rep 10:4795. https://doi.org/10.1038/s41598-020-61744-2
34. Serdean F, Pustan M, Dudescu C, Birleanu C, Serdean M (2020) Analysis of the thermoelastic damping effect in electrostatically actuated MEMS resonators. Mathematics 8(7):1124. https://doi.org/10.3390/math8071124
35. Duwel A, Candler R, Kenny T (2007) Engineering MEMS resonators with low thermoelastic damping. J Microelectromech Syst 15:1437–1445. https://doi.org/10.1109/JMEMS.2006.883573
36. Yi YB (2010) Finite element analysis of thermoelastic damping in contour-mode vibrations of micro-and nanoscale ring, disk, and elliptical plate resonators. Journal of Vibration and Acoustics 132(4). https://doi.org/10.1115/1.4001506
37. Kumar H, Mukhopadhyay S (2019) Thermoelastic damping analysis for size-dependent microplate resonators utilizing the modified couple stress theory and the three-phase-lag heat conduction model. Int J Heat Mass Transf 148:118997. https://doi.org/10.1016/j.ijheatmasstransfer.2019.118997
38. Devi S, Kumar R (2020) Thermoelastic damping and frequency shift in Kirchhoff plate resonators based on modified couple stress theory with dual-phase-lag model. Journal of Solid Mechanics 12(3):700–712. https://doi.org/10.22034/jsm.2020.1896290.1569
39. Resmi R, Babu VS, Baiju MR (2021) Impact of dimensionless length scale parameter on material dependent thermoelastic attenuation and study of frequency shifts of rectangular microplate resonators. IOP Conf Ser: Mater Sci Eng 1091:012067
40. Zhao G, Shi S, Gu B et al (2021) Thermoelastic damping analysis to nano-resonators utilizing the modified couple stress theory and the memory-dependent heat conduction model. J Vib Eng Technol. https://doi.org/10.1007/s42417-021-00401-y
41. Zhong Z, Zhang W, Meng G, Wang M (2015) Thermoelastic damping in the size-dependent microplate resonators based on modified couple stress theory. J Microelectromech Syst 24(2):431–445. https://doi.org/10.1109/JMEMS.2014.2332757

Machine Learning-Based Depression Detection

Saikat Biswas, Md.Mozahidul Islam, Utpaul Sarker, Rashidul Hasan Hridoy, and Md. Tarek Habib

Abstract Depression is an extremely serious illness of humans which causes constant mood swings and feelings of sadness. Nowadays, it is considered to be a deadly disorder in the world. At present, everyone from young to old is suffering from depression but most of them do not have the right idea about their mental state. Everyone needs to have a proper idea about their mental state. We will detect depression through a machine learning-based detection approach. Talking with psychologists and depressed people, we find some factors that are related to becoming depressed, and depending on those factors, information is collected from both depressed and non-depressed people. After applying preprocessing techniques, a processed dataset was created finally. Then, feature selection techniques were used. We applied eight machine learning algorithms and two feature selection methods to our dataset. We used k-nearest neighbor (*k*-NN), decision tree (DT), linear discriminant analysis (LDA), adaptive boosting (AB), support vector machine (SVM), naive Bayes (NB), random forest (RF), and logistic regression (LR) classifier. In our work, the RF classifier gave the best performance based on accuracy and the accuracy of the RF classifier was 96.00%.

Keywords Machine learning · Feature selection · Random forest · SelectKBest · Detection · Depression

S. Biswas (✉) · M. Islam · U. Sarker · R. H. Hridoy · Md. T. Habib
Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh
e-mail: saikat15-10878@diu.edu.bd

M. Islam
e-mail: mozahidul15-10995@diu.edu.bd

U. Sarker
e-mail: utpaul15-11043@diu.edu.bd

R. H. Hridoy
e-mail: rashidul15-8596@diu.edu.bd

Md. T. Habib
e-mail: md.tarekhabib@yahoo.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_60

1 Introduction

Depression which is also known as disappointment is a commonly found problem nowadays in the world, which is a general and troublesome medical condition which influences people's lives. Nearly 280 million people have depression in the world [1]. Depression is technically a mental disorder, but it also affects physical health and well-being. People of all ages are suffering from depression badly. The global spread of moderate to extreme situations of depression is 60.80% [2]. Globally, a significant number of people are estimated to be dealing with depression-related issues, and this number is equal to 4.40% population of the world. According to WHO, depression will be a great reason for several diseases [3]. An analysis from the UNICEF report estimates that mental illness causes about $390 billion a year in damage to the economy, which could lead to disability or death among young people [4]. Depression is the most common illness worldwide, affecting approximately 3.80% of the population, of which 5.00% are adults and 5.70% are over 60 years of age [1]. There can be many 1515 causes of depression including unemployment problems, abuse, lack of love, conflict, lack of family ties, affection, political cataclysm, etc. As a result, people are committing heinous acts like suicide. One of the most common implicit disorders is depression, and 30–70% of victims of suicide suffer from major or minor depression. Self-awareness is needed to get rid of depression. In Bangladesh, a recent study found that 56.60% of the 153 patients affected with COVID-19 reported depression and also noted that depression occurred among the university students (62.90%), general population (47.20%), medical students (49.90%), and healthcare workers (39.50%) [5].

Machine learning and deep learning approaches are used in various predictions, detection, and recognition-related systems [6]. In our dataset, there were 24 features. We applied eight machine learning algorithms, which are *k*-NN, LR, SVM, NB, RF, AB, DT, and LDA classifier to detect depression. In our study, the RF classifier gained 96.00% accuracy with very high sensitivity, precision, and F1-score after applying the SelectKBest feature selection technique.

2 Literature Review

Researchers have conducted research in this field in the past few years based on social media posts, visual expressions, and audio input using approaches of machine learning and deep learning. Khan et al. [7] introduced a research work on machine learning-based sentiment analysis using the Bangla depression dataset. Information has been collected from social media posts and various poems, novels, and quotes from various great personalities. The algorithms they used were SVM, multinomial naive Bayes (MNB), RF, *k*-NN, DT, and XGBoost. Different types of data are used by different researchers. Mulay et al. [8] worked with a total of 35,887 images from

the FER2013 dataset collected from Kaggle. They detected depression levels automatically through visual input. They used the convolutional neural network (CNN) model to extract extraction, and perform classification, which generated an output of emotion vectors. The CNN model had an accuracy of 66.45%. Using speech or voice signals by drawing strength and statistical features, a system is proposed for detecting depression in a person. Researchers identify that a person is depressed or not depressed from the voice signal of that person. The average accuracy of cross-validation sets and training sets was 81.56%. Shukla et al. [9] utilized the Ryerson audio-visual dataset (RAVD) which contains emotional songs and speeches. Ding et al. [10] have shown a method for college students' depression recognition. They collected data from social text data (Sina Weibo) from 1000 users. This paper mainly proposed a deep integrated SVM model for detecting depression from data of Chinese text. For finding output, deep integrated SVM (DISVM) is applied and has achieved the highest accuracy of 86.15%. Some researchers have used data, collected from social media to detect depression [11, 12]. Deshpande et al. has done work on detecting depression by using emotional artificial intelligence. They detected depression by analyzing tweets. For generating the training and test dataset, 10,000 tweets were collected from Twitter using Twitter API. For detecting depression, SVM and NB algorithms were used. SVM gained 79.00% accuracy. But the highest accuracy was gained by the Multinomial Naive Bayes algorithm. The best accuracy was gained 83% in MNB. Also, Orabi et al. [13] have done work on depression detection of Twitter users using deep learning. A total of 1,145 Twitter users' data are combined to build the dataset which is labeled as depressed, control, and PTSD. In addition, all users of this dataset were labeled using gender and age information. The key focus of this work was mainly on recognizing the user's capabilities to depression. CNN and recurrent neural network (RNN) were used by the researchers, where the CNN with max models using their optimized embedding reported higher accuracy which was 87.95%. Al chatbot system was built for combating depression and interacting with humans. On paper [14], Patel et al. analyzed emotion from text data, then recognized depression, and finally built a chatbot. They used the ISEAR dataset. Forming word vector algorithm was applied for segmentation. RNN, CNN, and HAN algorithms were applied for classification purposes. Among those, CNN performed better with 75.00% accuracy with high consistency for 15 epochs. Ranade et al. [15] proposed an emotion model and also implemented an AI chatbot, namely CakeChat, for understanding human emotion. Asad et al. [16] has worked on social media posts of the users and then detected depression from that. The data was collected from Twitter and Facebook, the beautiful soup was used to collect Twitter data, and having permission manually, data was collected from 150 Facebook users. SVM and NB classification algorithms were used in this work. Their ways of detecting depression levels were made using the popular BDI-II questionnaire method [1]. The accuracy of NB was 74.00% with a very high precision of 100%. A multi-region DepressNet was proposed on paper [17]. Zhou et al. recognized depression from facial images which were collected from video input. DepressNet deep learning model was used. AVEC 2013 and 2014 datasets were used for training the deep learning model. Several machine learning and deep learning methods are widely used by researchers for different

kinds of disease detection and recognition. Using people's personal living habits and checkup information, Dahiwade et al. [18] mentioned a machine learning-based disease prediction system. They used the CNN algorithm and the *k*-NN algorithm, where CNN took less time than *k*-NN with the highest accuracy 84.50%. Using machine learning techniques, Habib et al. [19] has addressed a recognition system for papaya diseases, where SVM gained 95.20% accuracy. Moreover, several studies were conducted for emotion analysis and sentiment analysis [20–22].

Most of the research works are done by analyzing social media platforms like Facebook and Twitter posted data. But social life and real life are not the same. Social media data cannot reflect real life fully. That is why we used real-life data of people by questionnaire-based surveys and then we got exact data about them in the COVID-19 situation which becomes more accurate and preferable.

3 System Architecture

In Fig. 1, a system architecture is shown for detecting depression easily, where a person will give response to some selected queries by using a Web-based software. The data provided by a person then goes to a cloud server and then to the specialist machine. The results will be created using the received input information of the system, by utilizing the RF classifier to the input data. Results will be finalized from the output of the developed model and, finally, send the outcome to the user through the application.

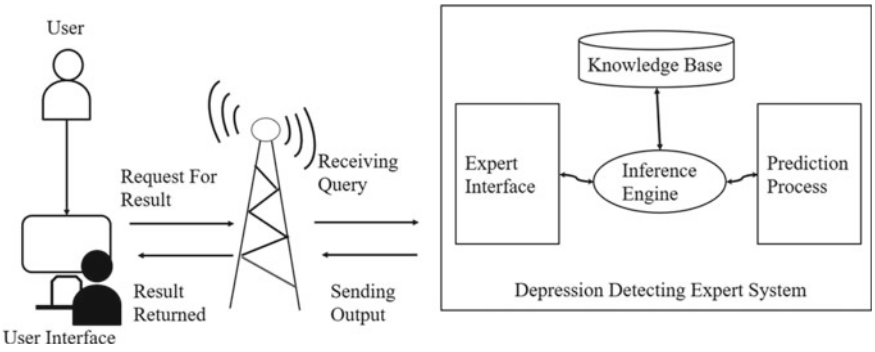


Fig. 1 The system architecture for detecting depression

4 Research Methodology

We collected information of 1000 depressed and non-depressed people, where 80% of the dataset was used as training data and 20% as testing data. We used eight machine learning classifiers like *k*-NN, SVM, LR, NB, DT, RF, AB, and LDA. We examined the performance of classifiers several times. Mainly performance was evaluated first in processed data with all features, and then, we apply two feature selection techniques on the dataset. Accuracy was calculated after using the principal component analysis (PCA) feature selection technique and using selectKBest feature selection technique in the processed dataset. In prior research works conducted for detecting depression, no method was used for selecting important features. In these three main steps, accuracy is checked for (70–30%), (80–20%), and (90–10%) as training data and testing data ratio. We examined the classification performance using accuracy and other metrics such as precision, specificity, recall, and F1-score. The processes of this work are described in Fig. 2 in the following flow diagram. Lastly, using the result of classifier performance, we proposed a machine learning algorithm to detect depression.

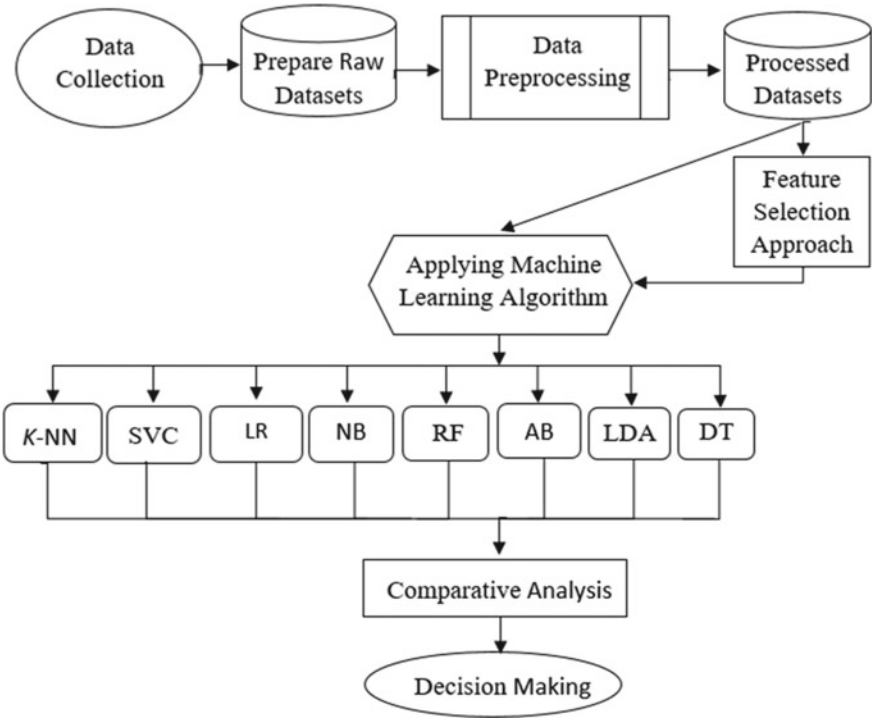


Fig. 2 Steps of the proposed methodology

4.1 Data Collection

The dataset consists of a huge number of features that are directly or indirectly connected with depression. We were not able to collect the information as we went to the hospital because the authorities said that if they provide patient information that could damage patients' privacy and can affect their rules and they also said that sometimes it is not available as a ready dataset. That is why we think, we need to create our datasets which are collected from face-to-face questionnaires, google forms, and a paper with a list of questions. The questionnaires for features are taken by taking with physicians, Web sites [23], articles, and people suffering from depression who are based on GDS [24], SDS [25], and MDI [26]. The features of depression detection are shown in Table 1. We were successful in collecting 1000 people's data based on 24 factors which are the daily activities of a person.

After collecting all the data, the main challenging part that came to focus was data leveling into depressed people and non-depressed people. We visited many doctors to seek help and find patterns to find out. With the help of one doctor, one physiatrist, and one student of psychology, the data were being leveled separately and combining their three decisions into final leveling outputs with an emphasis on the opinion of the majority. Among those data, there is information of 417 depressing and 583 non-depressed people. We gathered all data through an online survey, from many universities, various secondary and higher secondary schools and colleges, the public in different cities and rural areas, and more.

Table 1 Used features to detect depression

Name of feature	Evidence	Name of feature	Evidence
Age	(18)	Feel energy	(18)
Gender	(25)	Prefer to stay home	(18)
Profession	(25)	Avoiding social gatherings	(18)
Marital Status	(25)	Feeling memory loss	(18)
Satisfied in Life	(18)	Consider worthless	(18)
Dropped Activities	(18)	Cry most of the time	(19)
Feel Life Empty	(18)	Sleep well	(19)
Feel Board	(18)	Feeling reduce appetite and losing weight	(20)
Feel helpless	(18)	Situation hopeless	(19)
Afraid for bad-happening	(18)	People better than you	(18)
Hopeful about future	(18)	Feel bad and guilty	(20)
Spend time happily	(18)	Others would better you died	(19)

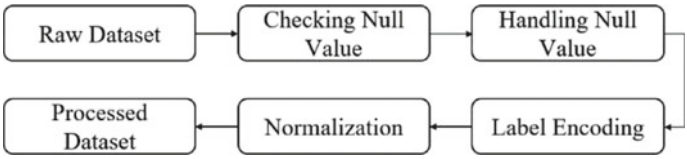


Fig. 3 Data preprocessing process

4.2 Data Preprocessing

After successfully collecting a sufficient amount of data, we noticed that there were some missing values in some of the data, and there were also different types of data like categorical data and numerical data. This kind of data is not suitable for machine learning algorithms. So we made the decision that we will process our data according to our needs and made this dataset compatible with algorithms. Data processing has the power to convert data into appropriate formats after data collection [27]. Processed data in a specific type help to get the best output easily. The applied preprocessing steps are presented below with help of Fig. 3.

Firstly, we started our work by collecting and making our raw dataset. Then, we focused on data cleaning. We inspected if there were any null or missing values in the prepared dataset. We solved the missing values problem using the imputer. Rather than deleting any data, we tried to fill this null value with its relevant possible values. Then, we encoded levels that transformed all text or categorical data into its relevant data of numerical form. Then through normalization, the data transformation was completed. For age data features, we used min–max normalization techniques. Finally, we got the processed dataset.

4.3 Performance Metrics

Based on values of the confusion matrix which are True Positive, False Positive, True Negative, and False Negative, we tested each of our model's performance against five performance metrics, including accuracy, recall, specificity, precision, and F1-score by applying their mathematical formulas which are given below.

$$\text{Accuracy} = \frac{\text{TrueDepressed} + \text{TrueNondepressed}}{\text{TotalData}} \times 100$$

$$\text{Recall} = \frac{\text{TrueDepressed}}{(\text{TrueDepressed} + \text{FalseNonDepressed})} \times 100$$

$$\text{Precision} = \frac{\text{TrueDepressed}}{(\text{TrueDepressed} + \text{FalseDepressed})} \times 100$$

$$\text{Specificity} = \frac{\text{TrueNonDepressed}}{(\text{FalseDepressed} + \text{TrueNondepressed})} \times 100$$

$$\text{F1-Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \times 100$$

5 Result and Discussion

In this section, we will demonstrate in detail the experimental outcome of this research work. To clearly understand, we will discuss this study with the assistance of some tables and graphs and also provide concise comparisons with this work with others.

We worked and built our model, and all further processes were done based on the dataset of 417 depressed people and 583 non-depressed people. In percentage which was about 41.70%, people are depressed and 58.30% are non-depressed. In this research work, we utilized eight machine learning algorithms and two feature selection approaches.

In three processes, we have evaluated the model and we have also trained and tested the dataset in different proportions to monitor the maximum result. At first, we selected all the features to find out the best result. For the 70–30% train-test ratio, after applying the different types of algorithms, we got an accuracy that was 92.70% for LDA, and the performance of LR and RF was not satisfying. Because at the same time, for the 80–20% train-test ratio, we got better accuracy and that was 95.00% for RF. And the best accuracy was gained from the 90–10% train-test ratio which was 96.00% for the RF algorithm. Table 2 demonstrates the efficiency of all algorithms for an 80–20% train-test ratio based on their performance metrics such as precision, accuracy, F1-score, specificity, and sensitivity.

After using the selectKBest feature selection techniques, we dropped some features and among 24 features it came to 17 features. Then, we evaluated the model,

Table 2 Classifier performance for all features for 80–20% training testing data ratio

Method	Model name	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)
All features	*k*-NN	92.50	93.18	91.96	90.11	91.62
	DT	93.00	92.05	93.75	92.05	92.05
	NB	89.50	96.59	83.93	82.52	89.01
	LDA	93.50	95.45	91.96	90.32	92.82
	SVM	93.00	95.45	91.96	90.32	92.82
	LR	94.00	94.32	93.75	92.22	93.26
	RF	95.00	95.45	94.64	93.33	94.38
	AB	90.50	89.77	91.07	88.76	89.27

where for the 70–30% train-test ratio RF performed better with 93.00% accuracy. This accuracy was increased for both 80–20% and 90–10% train-test ratio and gained the highest accuracy of 96.00% by RF. Table 3 demonstrates the efficiency of all algorithms for 80–20% training–testing data ratio based on performance metrics.

Finally, the PCA feature selection technique was also applied. By reducing some features, the classification algorithm's performance was calculated. For the 70–30% train-test ratio, SVM performed better with 94.00% accuracy. But on 90–10%, it reduced the accuracy of some algorithms and got the lowest accuracy of 86.00% for DT. In 80–20%, LR performed best with 94.50% accuracy. Table 4 describes the performance of each algorithm for 80–20% training–testing data ratio based on their performance metrics.

We observe that after applying feature selection techniques, the accuracy of some algorithms was increased, the accuracy of some algorithms was decreased and the accuracy of some algorithms was remained unchanged. Feature selection techniques such as SelectKBest and PCA have a notable impact on the performance of classifiers. SelectKBest performed better than PCA in this study. It appears in Fig. 4 that

Table 3 Classifier performance of selectKBest for 80–20% training testing data ratio

Method	Model name	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)
SelectKBest	*k*-NN	92.50	93.18	91.96	90.11	91.62
	DT	93.00	92.05	93.75	92.05	92.05
	NB	89.50	96.59	83.93	82.52	89.01
	LDA	93.50	95.45	91.96	90.32	92.82
	SVM	93.00	95.45	91.96	90.32	92.82
	LR	94.00%	94.32	93.75	92.22	93.26
	RF	96.00	97.73%	94.64	93.48	95.56
	AB	90.50	89.77%	91.07	88.76	89.27

Table 4 Classifier performance of PCA for 80–20% training–testing data ratio

Method	Model name	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)
PCA	*K*-NN	90.00	88.64	91.07	88.64	88.64
	DT	89.50	90.91	88.39	86.02	88.40
	NB	91.00	92.05	90.18	88.04	90.00
	LDA	94.00	94.32	93.75	92.22	93.26
	SVM	94.00	94.32	93.75	92.22	93.26
	LR	94.50	94.32	94.64	93.26	93.79
	RF	92.50	93.18	91.96	90.11	91.62
	AB	92.00	88.64	94.64	92.86	90.70

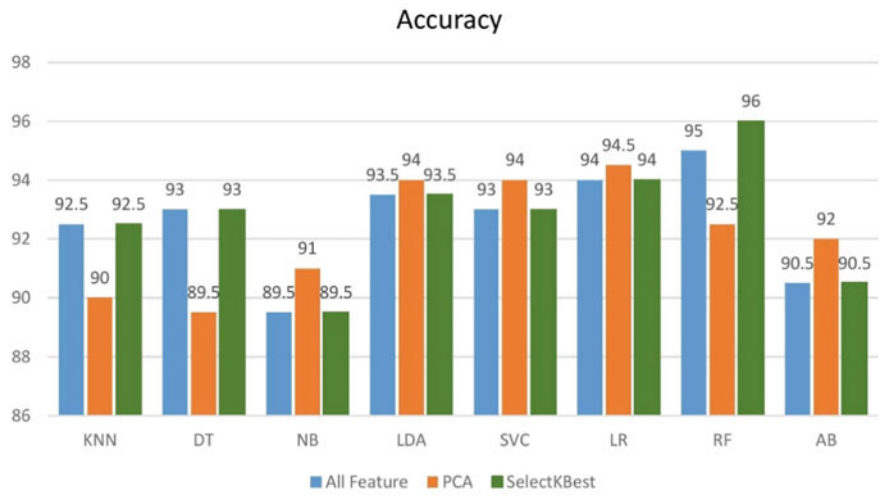


Fig. 4 Accuracy of algorithms for different methods

the RF classifier performed better than any other algorithm in this study under the SelectKBest method.

However, we also compared our study with previous related works, which is demonstrated in Table 5. According to the result of the comparison study, our proposed approach performed better than other studies. Moreover, a comparison between feature selection approaches was first introduced in this study for detecting depression.

6 Conclusion

Our proposed approach in this paper is to detect the mental state of a person in a COVID-19 situation. We applied eight machine learning algorithms and two feature selection approaches, where RF performs best with an accuracy of 96.00% under the SelectKBest method. Under the PCA method, the LR algorithm performed better than other classifiers. The RF classifier achieved 95.00% accuracy using all features.

We created a user-friendly Web site based on this model to detect depression easily. Depression detection with a machine learning model will provide a positive impact on people of all societies. After completing this work, we are hopeful that this work can be helpful for better understanding, further study about depression.

For detecting depression, it is better to use more output levels. We were not able to collect different classes of people's information. The dataset is not so large also. A larger and stronger dataset can be created by collecting information from various classes of people and creating more classification layers of datasets in the future. In addition, by developing an immersive user-friendly GUI, a mobile application can

Table 5 Comparison among our proposed approach and other previous research works

Study	Object(s)/Deal with	Size of sample	Size of feature set	Algorithms	Accuracy
This work	Depression (survey-based)	1000	24	RF	**96.00%**
Ding et al. [10]	Depression (text, emoji)	1000	NM	DISVM	86.15%
Mulay et al. [8]	Depression (image)	35,887	NA	CNN	66.45%
Asad et al. [16]	Depression (text, question)	150	NM	SVM	74.00%
Deshpande et al. [11]	Depression (text, question)	10,000	NM	NB	83.00%
Khan et al. [7]	Depression (text)	10,000	NM	MNB	86.00%
Zhou et al. [17]	Depression (video)	490	NM	NA	NM
Stankevich et al. [12]	Depression (text)	887	3	SVM	63.00%
Patel et al. [14]	A chatbot (user text analysis)	7666	NM	CNN	75.00%
Shukla et al. [9]	Speech signals of persons	7356	13	Multi-layer perceptron	81.57%
Orabi et al. [13]	Depression (text)	1145	NM	CNN	87.96%
Ranade et al. [15]	Chatbot (based on emotion)	NA	NM	NA	NA

NM: Not Mentioned
NA: Not Applicable

be developed which can be made accessible to all people and doctors also. Applying new algorithms and advanced techniques can make the model much more efficient and more accurate.

References

1. World Economic Forum, Available at: https://www.weforum.org/agenda/2021/10/brain-implant-could-cure-depression/ [Online]. Accessed 1 May 2021
2. Arusha AR, Biswas RK (2020) Prevalence of stress, anxiety and depression due to examination in Bangladeshi youths: a pilot study. Child Youth Serv Rev 116:105254
3. Globally Minded, Available at: https://www.globallyminded.org/home/depression-a-neglected-public-health-domain-in-bangladesh/ [Online]. Accessed 13 Aug 2021
4. Medical News Today, Available at: https://www.medicalnewstoday.com/categories/depression [Online]. Accessed 22 Oct 2021
5. Mental Health America, Available at: https://www.mhanational.org/conditions/suicide [Online]. Accessed 23 Feb 2021

6. Hridoy RH, Habib T, Jabiullah I, Rahman R, Ahmed F (2021) Early recognition of betel leaf disease using deep learning with depth-wise separable convolutions. In: 2021 IEEE Region 10 Symposium (TENSYMP), pp 1–7
7. Khan MRH, Afroz US, Masum AKM, Abujar S, Hossain SA (2020) Sentiment analysis from Bengali depression dataset using machine learning. In: 2020 11th international conference on computing, communication and networking technologies (ICCCNT)
8. Mulay A, Dhekne A, Wani R, Kadam S, Deshpande P, Deshpande P (2020) Automatic depression level detection through visual input
9. Shukla DM, Sharma K, Gupta S (2020) Identifying depression in a person using speech signals by extracting energy and statistical features
10. Ding Y, Chen X, Fu Q, Zhong S (2020) A depression recognition method for college students using deep integrated support vector algorithm
11. Deshpande M, Rao V (2017) Depression detection using emotion artificial intelligence. In: International conference on intelligent sustainable systems (ICISS), pp 858–862
12. Stankevich M, Isakov V, Devyatkin D, Smirnov I (2018) Feature engineering for depression detection in social media. In: 7th international conference on pattern recognition applications and methods (ICPRAM 2018), pp 426–431
13. Orabi AH, Buddhitha P, Orabi MH, Inkpen D (2018) Deep learning for depression detection of twitter users. In: Proceedings of the fifth workshop on computational linguistics and clinical psychology: from keyboard to clinic, pp 88–97
14. Patel F, Thakore R, Nandwani I, Bharti SK (2019) Combating depression in students using an intelligent ChatBot: a cognitive behavioral therapy. In: 2019 IEEE 16th India council international conference (INDICON), pp 1–4
15. Ranade AG, Patel M, Magare A (2018) Emotion model for artificial intelligence and their applications. In: 2018 fifth international conference on parallel, distributed and grid computing (PDGC)
16. Asad NA, Mahmud Pranto MA, Afreen S, Islam MM (2019) Depression detection by analyzing social media posts of user. In: 2019 IEEE international conference on signal processing, information, communication & systems (SPICSCON), pp 13–17
17. Zhou X, Jin K, Shang Y, Guo G (2020) Visually interpretable representation learning for depression recognition from facial images. IEEE Trans Affect Comput 11(3):542–552
18. Dahiwade D, Patle G, Meshram E (2019) Designing disease prediction model using machine learning approach. In: 2019 3rd international conference on computing methodologies and communication (ICCMC)
19. Nuruzzaman M, Hossain MS, Rahman MM, Shoumik ASHC, Khan MAA, Habib MT (2021) Machine vision based potato species recognition. In: 2021 5th international conference on intelligent computing and control systems (ICICCS)
20. Kottursamy K (2021) A review on finding efficient approach to detect customer emotion analysis using deep learning analysis. Journal of Trends in Computer Science and Smart Technology 3(2):95–113
21. Tripathi M (2021) Sentiment analysis of Nepali COVID19 tweets using NB, SVM and vLSTM. J Artif Intell 3(03):151–168
22. Kaur R, Vijay B (2021) A GUI for multimodal sentiment analysis on Gurmukhi script. In: Innovative data communication technologies and application, pp 69–81
23. Pudrovska T, Karraker A (2014) Gender, job authority, and depression. J Health Soc Behav 55(4):424–441
24. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO (1982) Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 17(1):37–49
25. Zung WWK (1965) A self-rating depression scale. Arch Gen Psychiatry 12(1):63
26. Olsen LR, Jensen DV, Noerholm V, Martiny K, Bech P (2003) The internal and external validity of the major depression inventory in measuring severity of depressive states. Psychol Med 33(2):351–356
27. Nair R, Bhagat A (2018) A life cycle on processing large dataset—LCPL. Int J Comput Appl

28. Wikipedia, Available at: https://en.wikipedia.org/wiki/Beck_Depression_Inventory#BDI-II [Online]. Accessed 1 Dec 2021

An Energy Efficient, Long Range Sensor System for Real-Time Environment Monitoring

Van-Phuc Hoang, Van-Lan Dao, Van-Trung Nguyen, Xuan Nam Tran, and Koichiro Ishibashi

Abstract Vietnam is among the top countries in the world at risk of natural hazards so that efficient real-time environment monitoring is becoming essential. The continuous development in information and communication technology is inspiring the development of the smart monitoring systems for environmental management and protection. This paper presents an energy efficient, long range sensor system for Internet of Things (IoT)-based smart environment monitoring and early warning. The proposed system combines the novel energy efficient temperature. Beat sensors integrated with the LoRaWAN communication protocol and web interface. The experimental results are achieved to clarify the efficiency of the proposed sensor system and its potential applications in the real systems.

Keywords Beat sensor · Drought monitoring · Internet of things

1 Introduction

Vietnam is among the top countries in the world at risk of water-related threats including drought, flood, storm, coastal and agriculture hazards. Among water-related disasters, droughts are highly complex, frequently occurring natural hazards in Southeast Asia including Vietnam [1]. Droughts severely impact on the economy and directly affect infrastructure, properties and people in large geographical regions for extended periods. Due to the characteristics of drought, real-time measuring and monitoring of full impacts of drought are complicated; hence, it requires research, extensive data and experiential evidences for decision-making processes. This leads

V.-P. Hoang (✉) · V.-L. Dao · V.-T. Nguyen · X. N. Tran
Le Quy Don Technical University, 236 Hoang Quoc Viet street, Hanoi, Vietnam
e-mail: phuchv@lqdtu.edu.vn

V.-L. Dao
Malardalen University, Högskoleplan 1, 72218 Västerås, Sweden

K. Ishibashi
University of Electro-Communications, Chofu-shi, Tokyo, Japan

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_61

to the requirement of an efficient real-time monitoring and early warning system which can provide real-time data for both monitoring and decision making processes against droughts. Recently, as a part of the fourth industrial revolution, the continuous development of Internet of Things (IoT) has inspired the smart environment monitoring solutions [2]. Moreover, energy efficient sensors play an essential role in IoT systems. Therefore, many researchers have been focusing on efficient sensors for these applications including the drought monitoring system. Figure 1 presents our general model for the IoT-based environment monitoring system including the application for drought management. In this system, beside other types of sensor such as humidity and rainfall, the temperature sensors are the key components. Also, Fig. 2 is the application model of the IoT sensor system for real-time drought monitoring. Many sensors are used to obtain the environment parameters and send to the

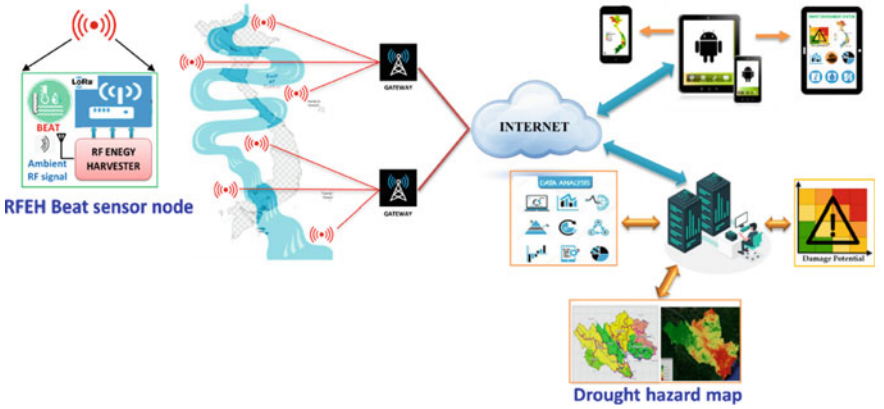


Fig. 1 General IoT system model for real-time drought monitoring

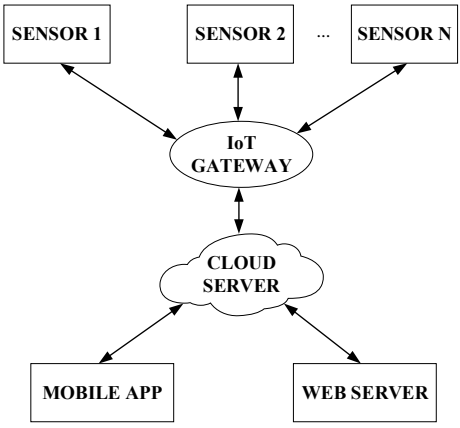


Fig. 2 Application model of IoT sensor system for real-time drought monitoring

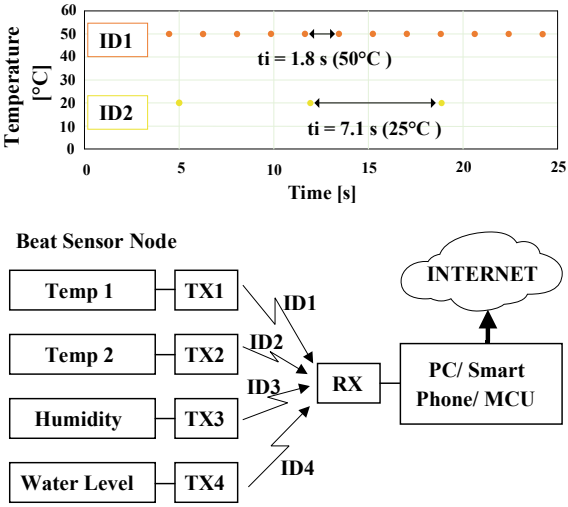


Fig. 3 Beat sensor concept [3]

IoT gateways. A cloud server and web/mobile applications are employed to provide the assessment and early warning of the droughts.

The concept of beat sensor was based on the invention of Prof. Koichiro Ishibashi [3]. Beat sensors are suitable for IoT-based systems due to their compact, low power and low-cost features. The principle of these sensors is demonstrated in Fig. 3, where the sensors are combined with corresponding transmitters. In these beat sensors, each transmitter is associated with one code (ID). Instead of sending the direct sensing values, only the ID signals are transmitted. By exploiting this principle, the physical parameters of the sensor are represented with the corresponding interval times. The ID signals are obtained by the corresponding receivers, and the physical parameters can be calculated from the values of the interval times in the ID signals. Furthermore, the data correction algorithms can be utilized to improve the accuracy and the reliability of the data obtained by beat sensors for some specific applications such as in [4]. However, we expect further improvements and some application specific optimized techniques for these sensors so that they can be employed in the more real systems. Inspired by these observations, in this work, we propose new methods to design and build the efficient temperature sensors with beat principle for the smart environment monitoring systems with a range communication protocol (LoRaWAN). From the next part, this paper is organized as follows. In Sect. 2, we present the detailed structure and operation of the proposed system as well as the results of implementation. Finally, Sect. 3 gives the conclusions and the plan for our future work.

2 Proposed System

Firstly, to improve the energy efficiency of beat sensors, Fig. 4 presents a new struc-

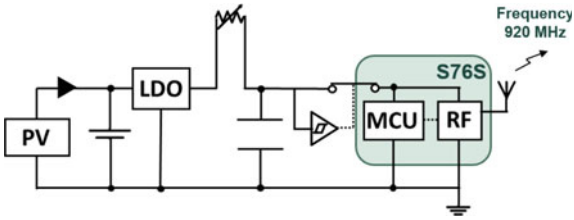


Fig. 4 Beat sensor with integrated LoRaWAN communication module

ture in which the LoRa module is integrated (with the frequency of 920 MHz for use in Japan). Figure 5 presents the experiment set up and Fig. 6 shows the results of the successful communication probability verses communication distance for the first proposal of the temperature Beat sensor with integrated LoRaWAN communication module. The experimental results have confirmed that with this scheme, and the communication distance of 600 m can be achieved.

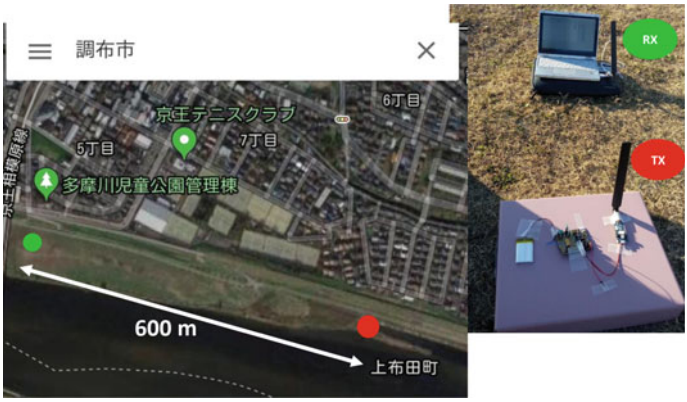


Fig. 5 Demonstrated monitoring system using the first proposal of the temperature sensor using the beat principle and LoRaWAN

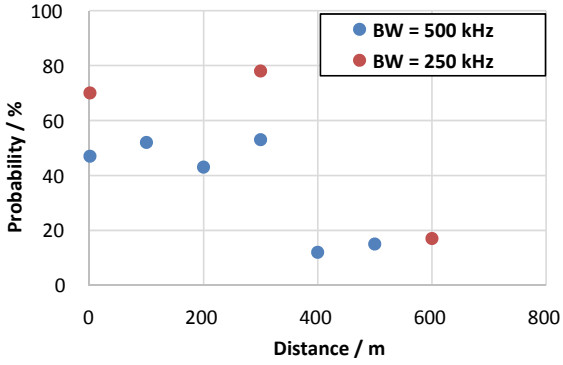


Fig. 6 Results of successful communication probability versus communication distance for the first proposal of temperature beat sensor integrated with LoRaWAN communication protocol

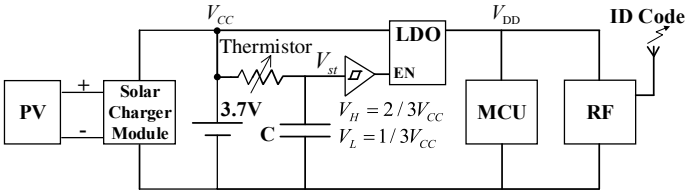


Fig. 7 Second proposal of temperature beat sensor structure

To improve the communication distance and the network reliability, Fig. 7 shows the structure of the second temperature sensor proposal with Beat sensor principle. In this scheme, the photovoltaic (PV) integrated with the solar charger modules are utilized as the supply source for the proposed Beat sensor. In this proposed sensor, a thermistor is used because the value of its resistance reflects the temperature. Especially, we utilize a low-dropout regulator (LDO) with two roles. The first role is to provide the regulated voltage supply for two parts of the microcontroller (MCU) and RF circuits (LoRaWAN module). Its second role is controlling the supply. We adopt a comparator having the threshold voltage values of V_{H} and V_{L} for this supply switching. As a result, the supply voltage for both the MCU and the RF circuit is more stable than that in [3]. The MCU performs both roles of ID assignment and the whole sensor circuit control.

Figure 8 illustrates an application system in which the proposed temperature sensor with Beat principle and LoRaWAN communication are employed. By the on-site experiments, the developed system can achieve the maximum communication distance of 1700 m with low energy consumption. Moreover, Fig. 9 shows the experiment results of communication distance and the successful communication probability with the bandwidth of 500 kHz and different values of spectrum spreading factor (SF).

Based on the developed sensors, we have built an IoT-based drought monitoring system with the web interface as shown in Figs. 10 and 11. In this system, three main precipitation indexes including the standardized precipitation index (SPI), the precipitation effectiveness index (P-E index) and the De Martonne aridity index (Iar-DM) were analyzed [5]. We have performed the experiments for three pilot areas in Hanoi, Vietnam. The experimental results have shown the reliable information of pilot areas including the annual temperature, the precipitation and the drought degree. The information provides a useful tool for climate experts, the government and authorized people to adapt against the droughts.

3 Conclusion

In this paper, we have presented a smart IoT-based drought monitoring and early warning system by combining the energy efficient sensors with the long range

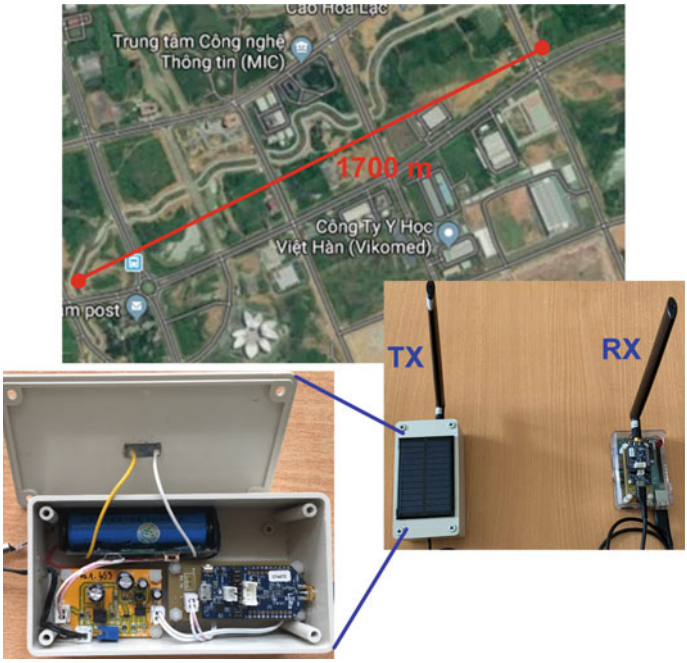


Fig. 8 Proposed drought monitoring system for field experiments

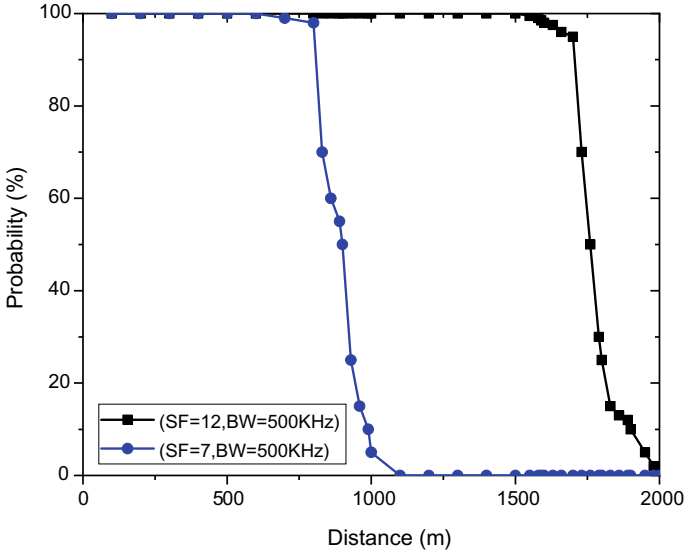


Fig. 9 Results of communication distance verse probability of successful communications with different vales of SF (for the case of 500 kHz bandwidth)

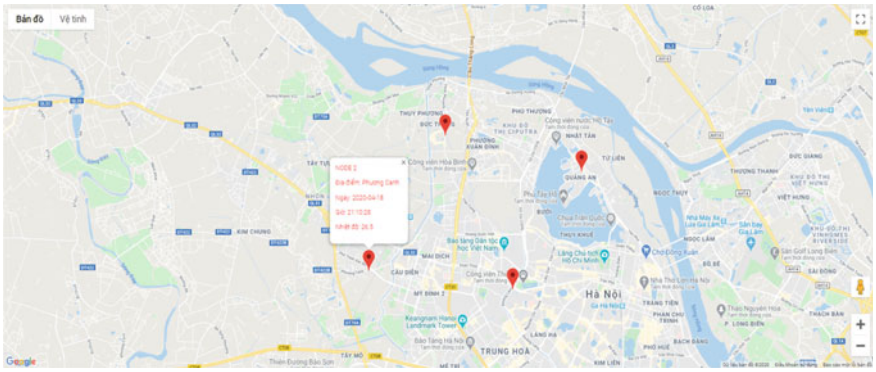


Fig. 10 Web interface for drought monitoring system integrated with Google map

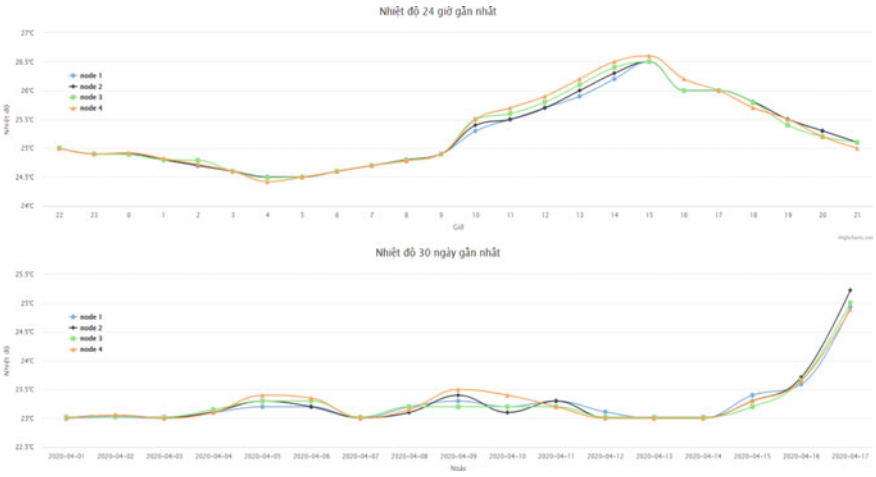


Fig. 11 Web interface showing shared data for users with average temperature

LoRaWAN communication protocol, cloud computing and web interface. With the low energy consumption, high reliability and long communication range, the proposed system provides a promising tool for our government to adapt with the issues of climate change in Vietnam. In the future, we will further improve the proposed system with the comprehensive experiments for a longer period and wider area.

Acknowledgements This publication is the output of the ASEAN IVO (http://www.nict.go.jp/en/asean_ivo/index.html) project, "An energy efficient, self-sustainable and long range IoT system for drought monitoring and early warning," and financially supported by NICT (http://www.nict.go.jp/en/ index.html).

References

1. South-East Asia: Drought (2019). Available online: https://reliefweb.int/disaster/dr-2019-000113-phl
2. Dao V-L, Hoang V-P (2017) A smart delivery system using internet of things. In: Proceedings of IEEE 7th international conference on integrated circuits, design, and verification (ICDV2017), pp 58–63
3. Kishi TM, Ishibashi K (2017) Low-power enhanced temperature beat sensor with longer communication distances by data-recovery algorithm. IEEE Sensors
4. Ishibashi K, Manyvone D, Itoh M, Dao V, Hoang V (2019) Beat sensors for smart environment monitoring systems. In: Proceedings of IEEE 2019 international conference on recent advances in signal processing, telecommunications and computing (SigTelCom), pp 77–79
5. Vu-Thanh H, Ngo-Duc T, Phan-Van T (2014) Evolution of meteorological drought characteristics in Vietnam during the 1961–2007 period. Theoretical and Applied Climatological 118(3):368–370

Identifying Genre of a Book from Its Summary Using Machine Learning Approach

Saidur Rahman Bhuiyan, Md. Rafidul Hasan Khan, Umme Sanzida Afroz, and Md. Sadekur Rahman

Abstract Categorical search and category-wise book recommendation are two common tasks for online booksellers. But for a machine to understand this category from a given text is still challenging work, where machine learning is a widely used tool at present. Though in the English language, with the availability of rich datasets and corpus, machine learning-based categorization and recommendation have reached a standard level, in the Bengali language, to reach the standard, still needs a long way to go. One key reason is the lack of availability of a rich Bengali dataset. The aim of this research was to make a dataset first for the book's genre identification from its given summary and to explore which supervised classifier performed best on that dataset for classifying the genres. Before that, we performed several essential preprocessing steps essential to prepare our dataset fit for the algorithms. Six machine learning classifiers were applied to the dataset, and it was observed that Naive Bayes performed best with an accuracy of 68% followed by XGB with an accuracy of 67%.

Keywords Book genre · Bengali data · NLP · Naive Bayes · Preprocessing · Text classification

1 Introduction

A good amount of bookish documents in e-libraries can be identified, besides reviews and comments from the page of online book shops, e-commerce platforms for books, blogs, or social networking sites. To identify a book based on inspections of all

S. R. Bhuiyan (✉) · Md. R. H. Khan · U. S. Afroz · Md. S. Rahman
Daffodil International Univerity, Dhaka, Bangladesh
e-mail: thesaidursumon@gmail.com

Md. R. H. Khan
e-mail: rafidul15-11542@diu.edu.bd

U. S. Afroz
e-mail: umme15-11505@diu.edu.bd

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_62

information which fulfills our all necessity is still quite difficult. In this perspective, book recommender systems and reviews play a vital role besides trying to give the required information to the customer. In the last era, reviews and recommender systems have tried to provide better recommendations by using reviews of books [1].

Aspect-based sentiment analysis (ABSA) gives out sucking the exact entities alluded to in a review and the sentiment related with them [2] so that we can understand at a glance which sights of a book are searched in a review and the sentiment figures out each other. This type of task can help to figure out more exact results, not only for recommending a book genre but also for querying, organizing, and extracting data in an organized way. Most of the ABSA research worked on restaurants domains and electronic products [3, 4], however, especially for book reviews there is not much work. In recent days, some research can be figured out [5–7], connected with some new annotated datasets on ABSA. They all work based on short reviews or tweets, whereas in this research, we work on very long reviews from experts as well. All the formal reviews about fiction and non-fiction category books have not still been collected, that is why we target to collect this type of Bengali review annotated with sights or sentiment information in the domain of Bengali books.

In the end, we try to present references to some research works which can be found on the basis of the related task, which is finding the genre of a text from its style, like documents of textual [8, 9] and most recently web-based [10]. A style or a genre presents another aspect of a document that differs from a topic and it is also a standard to identify various documents. We actually tried to apply classification methods, collecting different characteristics of texts, like lexical, structural, punctuations, and so on.

Analyzing what genre tends to generate on the basis of a collected summaries of books is the first step for applying sentiment analysis based on the classification approach to a new annotated dataset for the purpose of training and testing. In this research work, we perform an exploration of various kinds of book reviews collected from a reputed source: Rokomari as well as book e-commerce. Finally, the categorization of the book is identified, which will allow the new approach of research in the aspect of the classification of any book genre.

2 Literature Review

Alvarez-López et al. [11] introduced the genre classification of books based on two types of reviews, where one of them is formal reviews written by book experts about scholarly books and another one is from book readers about the fictional and non-fictional books which are less qualified than the previous one. Here, they implement support vector machine for review classification and also analyze both types of review using lexical and semantic analysis although the aspect-based sentiment analysis (ABSA) provides the most appropriate prediction for both types of reviews.

Sentiment analysis classifies the opinion from text, speech, or image, and there should be multiple levels of sentiments. Almeida et al. [12], in their sentiment analysis

work, used multiple methods for identifying the respective class or category. Those methods are like algorithm adaptation, ensemble method, or problem transformation and this will help to explore a broad area of the multilevel solution. Along with seven different classes, ensemble classifiers provide the best result which is 64.89% F measure.

Rout et al. [13] work on a Twitter dataset that contains positive, negative, and neutral types of data. Based on these three classes, they collect data using Twitter REST API which is better for collecting user status and authorized by Twitter. They apply both supervised and unsupervised approaches to identify the exact class of tweets. Their proposed unsupervised approach gains 80.68% accuracy which is greater than the multinomial Naive Bayes classification algorithm. The accuracy of MNB is 67%. In the supervised learning approach unigram and bigram, parts of speeches are also implemented.

Jagdale et al. [14] work on product reviews using machine learning classification techniques, where they collect users' reviews of different products from amazon's e-commerce platform. To classify the review of those products, they mainly used Naive Bayes and support vector machine classification algorithms and got accuracy accordingly 98.17% and 93.54% and this accuracy only generated for camera reviews.

On the basis of four classes, here, we will implement a supervised learning classification approach for identifying the book genre based on Bengali fiction or non-fiction book review. To serve this purpose, our preferable algorithms are support vector machine, multinational Naive Bayes, decision tree, random forest, K-nearest neighbors.

3 Methodology

Any kind of data-driven approach to a huge amount of data is essential and that is why we try to collect a bunch of data based on three-class named 'Islami,' 'science fiction,' and 'recipe.' For Bengali fiction and non-fiction book, review data are available in Rokomari, GoodReads, and a few other sites. Initially, we chose Rokomari to collect book names, publishers, authors, reviews, and book genres.

A mandatory requirement for any text data is preprocessing. That is why we preprocess our Bengali data in different levels like add contraction, remove punctuation, remove stopwords, and steaming. After completing this whole process, we got a cleaned dataset and used a CountVectorizer to encode them. CountVectorizer works using the one-hot encoding method. After that, we implement a feature selection technique to identify the necessary attributes and enhance the accuracy of our work. After implementing the feature selection method, we chose book review and book genre as necessary attributes. Here, Fig. 1 shows the workflow diagram of our work.

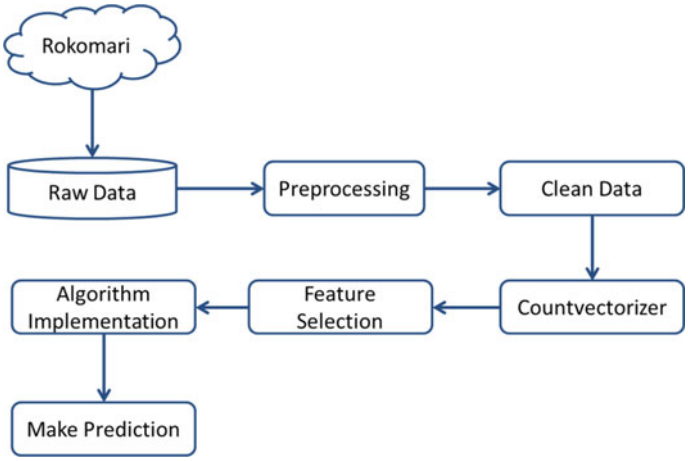


Fig. 1 Workflow diagram

4 Result and Output Analysis

Among six algorithms, the highest accuracy comes from multinomial Naive Bayes which is 67.8%. Here, Fig. 2 shows the overall accuracy of the previously mentioned six algorithms.

In Table 1, we represent the precision, recall, F1-score, and support values for every algorithm in each class.

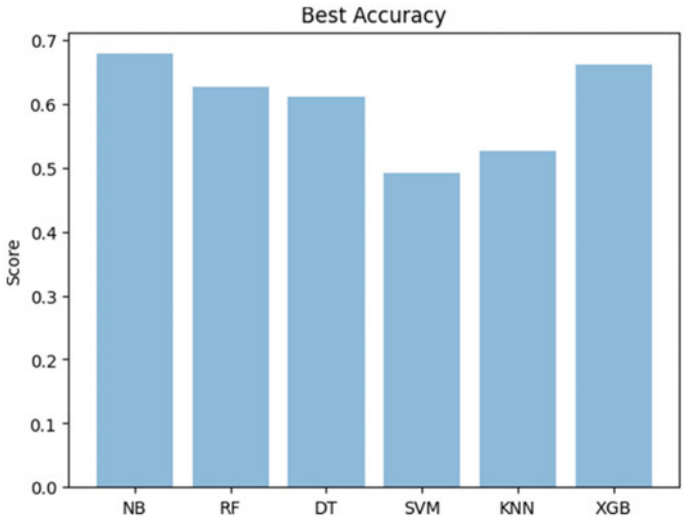


Fig. 2 Accuracy bar chart

Table 1 Result from confusion matrix

Algorithms		Precision	Recall	F1-score	Support
Multinomial Naïve Bayes	0	1.00	0.61	0.76	18
	1	0.41	0.60	0.49	15
	2	0.77	0.77	0.77	26
Random forest	0	0.50	0.50	0.50	18
	1	0.80	0.53	0.64	15
	2	0.65	0.77	0.70	26
Decision tree	0	0.53	0.56	0.54	18
	1	0.57	0.53	0.55	15
	2	0.69	0.69	0.69	26
Support vector machine	0	0.47	0.50	0.49	18
	1	0.40	0.23	0.20	15
	2	0.51	0.69	0.59	26
K-nearest neighbors	0	0.50	0.50	0.50	18
	1	0.40	0.40	0.40	15
	2	0.62	0.62	0.62	26
XGBoost	0	0.62	0.72	0.67	18
	1	0.67	0.53	0.59	15
	2	0.69	0.69	0.69	26

5 Conclusion

In this research work, analyze and classify three different genres of fiction and non-fiction books (formal reviews from Rokomari) based on the classification with the help of sentiment analysis which is part of natural language processing (NLP). Our new method of working, which will be beneficial for further identifying the genre of any different kinds of fiction or non-fiction books. There is not much work on the dataset of Bengali book reviews in general and even dealing with fiction and non-fiction book reviews. The schemes and analysis which are represented in this research work are the first steps to acquire. This outcome could be handy in recommending books as per the users' choice and also can be beneficial to organize books in ontological order. As currently, the accuracy of the models is not too satisfactory, and the future scopes of this research remain open for collecting more data and adopting deep learning approaches for acquiring better performance.

Acknowledgements Here, at first, we are thankful and happy to affiliate our Department of Computer Science and Engineering, Daffodil International University to provide us the GPU support laboratory facilities to execute our work which helps to complete our task and overcome the different types of barriers.

References

1. Chen L, Chen G, Wang F (2015) Recommender systems based on user reviews: the state of the art. User Model User-Adap Inter 25:99–154
2. Schouten K, Frasincar F (2016) Survey on aspect-level sentiment analysis. IEEE Transactions on Knowledge and Data Engineering 28:813–830
3. Ganu G, Elhadad N, Marian A (2009) Beyond the stars: improving rating predictions using review text content. WebDB 2009:1–6
4. Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pp 168–177
5. Alvarez-Lopez T, Fernandez-Gavilanes M, Costa-Montenegro E, Juncal-Martí nez J, García-Mendez S, Bellot P (2017) A book reviews ´dataset for aspect based sentiment analysis. In: Proceedings of 8th language & technology conference, pp 49–53
6. Al-Smadi M, Qawasmeh O, Talafha B, Quwaider M (2015) Human annotated Arabic dataset of book reviews for aspect based sentiment analysis. In: 2015 3rd international conference on future internet of things and cloud (FiCloud), IEEE, pp 726–730
7. Freitas C, Motta E, Milidiu RL, César J (2014) Sparkling vampire... lol! annotating opinions in a book review corpus. New language technologies and linguistic research: a two-way road, pp 128–146. Cambridge Scholars Publishing, Newcastle upon Tyne
8. Kessler B, Numberg G, Schtze H (1997) Automatic detection of text genre. In: Proceedings of 8th conference on European chapter of the association for computational linguistics, pp 32–38
9. Stamatatos E, Fakotakis N, Kokkinakis G (2000) Automatic text categorization in terms of genre and author. Comput Linguist 26(4):471–495
10. Lim CS, Lee KJ, Kim GC (2005) Multiple sets of features for automatic genre classification of web documents. Inf Process Manage 41(5):1263–1276
11. Alvarez-López T, Bellot P, Fernandez-Gavilanes M, Costa-Montenegro E (2018) From genre classification to aspect extraction: new annotation schemas for book reviews. In: 2018 IEEE/WIC/ACM international conference on web intelligence (WI), pp 428–433. IEEE
12. Almeida AM, Cerri R, Paraiso EC, Mantovani RG, Junior SB (2018) Applying multi-label techniques in emotion identification of short texts. Neurocomputing 320:35–46
13. Rout JK, Choo KKR, Dash AK, Bakshi S, Jena SK, Williams KL (2018) A model for sentiment and emotion analysis of unstructured social media text. Electron Commer Res 18(1):181–199
14. Jagdale RS, Shirsat VS, Deshmukh SN (2019) Sentiment analysis on product reviews using machine learning techniques. Cognitive informatics and soft computing. Springer, Singapore, pp 639–647

The Technology Solution to the Effects of the COVID-19 Pandemic on AgroTourism-Based MSMEs

Anggita Futri, N. H. O. Jeremiah, Richard Adrian, Vandella Franciska Paul, Ford Lumban Gaol, Tokuro Matsuo, and Natalia Filimonova

Abstract The pandemic known as COVID-19 has engulfed the world, and brought with it a huge set of impacts on economic growth, especially in the MSME sector for a country, especially Indonesia. Because this pandemic has hit various MSME sectors in Indonesia, especially the tourism sector, it is necessary to get more attention in order to get fast handling. Therefore, this study will discuss the impacts of COVID-19 MSMEs that operate in the tourism sector, what some tourist-heavy regions in Indonesia have been doing thus far to aid in overcoming them, and what the government can do to overcome these impacts. The purpose of this study is to help find solutions to these problems. Our research employs a qualitative research method with a descriptive type of research. The results of the research show that there are several impacts that the pandemic had on MSMEs in the tourism sector, including the lack of income generated because the source of income comes from tourist visits and a

A. Futri (✉) · N. H. O. Jeremiah · R. Adrian · V. F. Paul
Bina Nusantara University, Alam Sutera, School of Information System, Kota Tangerang, Indonesia
e-mail: Anggita.futri@binus.ac.id

N. H. O. Jeremiah
e-mail: jeremiah.ho@binus.ac.id

R. Adrian
e-mail: richard.adrian@binus.ac.id

V. F. Paul
e-mail: vandela.paul@binus.ac.id

F. L. Gaol
Binus Graduate Program, Doctor of Computer Science, Jakarta, Indonesia
e-mail: fgaol@binus.edu

T. Matsuo
Advanced Institute of Industrial Technology, Tokyo, Japan
e-mail: matsuo@aiit.ac.jp

N. Filimonova
Vladimir Branch of Russian Academy of National Economy and Public Administration, Vladimir, Russia

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_63

large drop in earnings and a large increase in MSME bankruptcy across the nation. Most of this occurred because the presence of COVID-19 prevents tourists from visiting. Each region has also enacted their own solutions, such as health protocols in Lampung and Kupang, and tourist-attracting marketing for Bali. The research also shows that the ministry of tourism has enacted several policies to countermeasure this issue, such as a change in how tourist-based businesses are run, safety protocols and certifications, and regulations. Our solution consists of contactless technological solutions as alternatives to travel and enactment of previously discussed government policies.

Keywords COVID-19 pandemic · MSME · Tourism solution · Technology impact

1 Introduction

In the year 2020, a new type of disease had spread, causing the outbreak of a disease called Coronavirus Disease 2019 (COVID-19). While it originated as a regional disease in China, it had spread itself in all corners of the globe. As of January 16, 2022, over 323 million confirmed cases and over 5.5 million deaths have been reported worldwide (WHO International Publication, 2022). Due to how quickly the virus spreads, governments around the world have implemented a series of lockdowns with the aim of reducing the number of people it can infect. The measure was done to minimize social interactions and curb the spread of the disease in the most affected regions [1]. However, this is not without its drawbacks, with one of them being a global economic slowdown. Data from the OECD (Organization for Economic Co-operation and Development) [2] shows that this pandemic and the recession it causes can lead to the threat of a global economic crisis that will lead to a decline or even cessation of production and business activities in many countries and a free fall across stock market. The mandate covers public spaces of all sorts and includes closing schools and workplaces. It placed a generalized restriction on public activities [3, 4].

This prediction will of course threaten Indonesia's national economy, and the economic aspect that will be most affected is the MSME sector (Micro, Small and Medium Enterprises) in Indonesia. This is due to the fact that in 2018 alone, there were a total of 64,194,057 MSMEs in Indonesia, accounting for a massive part of the economy. It can be said that MSMEs are at the epicenter of the economic shock due to the COVID-19 pandemic. The lockdown measures that have been implemented by governments around the world have provided various impacts on the international and national economy, such as the difficulty of meeting the needs of daily life and problems with employment. As of April 17, 2020, as many as 37,000 MSME actors reported themselves to the Ministry of Cooperatives and SMEs affected by the COVID-19 pandemic [5].

UNWTO (United Nations World Tourism Organization) also revealed information that essentially states that all businesses worldwide were badly hit by the COVID-19 Pandemic. One of the hardest hit sectors, however, is tourism. Because of the lockdown and travel restrictions, tourist arrivals are expected to drop by 80% in March 2020. In addition, the decision has limited the movement of people and resources across the nations. This hits the tourist industry especially hard, as it relies heavily on restaurants and travel agencies, which require a huge number of people to contact [6]. Therefore, in addition to having an impact on business owners, it also has an impact on their suppliers [7, 8].

1.1 Definition and Idea of MSMEs

The definitions of MSMEs according to Chapter I Article 1 of Law No. 20 of 2008 concerning Micro, Small and Medium Enterprises (MSMEs) are as follows:

(1) "Micro Enterprises are productive businesses owned by individuals and/or individual business entities that meet the criteria for Micro Enterprises as regulated in this Law.
(2) Small Business is a productive economic business that stands alone, which is carried out by individuals or business entities that are not subsidiaries or not branches of companies that are owned, controlled, or become part either directly or indirectly of Medium Enterprises or Large Businesses that meet Small Business criteria as referred to in this Law.
(3) Medium Enterprises are productive economic businesses that are standalone, which are carried out by individuals or business entities that are not subsidiaries or branches of companies that are owned, controlled, or become part either directly or indirectly with Small Businesses or Large Businesses with total assets. net or annual sales".

From the above definition, we can conclude that Micro, Small, and Medium Enterprises are businesses or economic activities carried out by one person or organization that meet the criteria for Micro, Small, and Medium Enterprises. There are three very important roles of SMEs in the lives of those of middle-to-bottom economic class, namely it provides a means of alleviating people from poverty, a means of leveling the economic level of the small people, and providing a sustainable income for the nation [5]. MSMEs have a very strategic position in the Indonesian economic system. The Ministry of Cooperatives and Small and Medium Enterprises (Kemenkop UKM) said that in 2018 there were 64,194,057 MSMEs in Indonesia, employing 116,978,631 workers (more than ninety percent of the country's working population). Thus, a blow to Indonesia's MSMEs is a blow to the Indonesian economy as a whole.

1.2 Known Effects of the Pandemic on MSME Business Operations

As the pandemic rages on, MSME business processes all over the nation have begun to suffer. This is because the process of selling a product in the MSME business usually relies on direct or face-to-face meetings between the seller and the buyer. Because the lockdown protocol violates the existence of these activities, especially in public places, naturally MSMEs experienced a drastic reduction in sales. According to data from the United Nations World Tourism Organization [9], there was a successive increase in tourists visiting tourist visits in 2019 with observations of worldwide visits of around 1.5 billion (predicted by UNWTO).

Data from the Ministry of Cooperatives (on 8 June 2020) states that 2322 cooperatives and 185,184 Micro, Small, and Medium Enterprises (MSMEs) have been affected by the COVID-19 virus pandemic. Based on data from Pakpahan [10] in the COVID-19 journal and implications for Micro, Small, and Medium Enterprises (The Scientific Journal of International Relations, (2020: 2)), about 56% of MSMEs in Indonesia reported a decline in sales, 22% reported problems in the financing aspect, 15% reported problems with distribution of goods, and 4% reported difficulties in getting raw materials. Therefore, Indonesia, a country dominated by the existence of MSMEs as the support for its national economy, will be seriously affected by this pandemic. MSMEs will not only lose their trade value and production capabilities, but also experience a decrease in the number of workers because they lose their jobs in the midst of this pandemic.

1.3 Tourism-Based MSMEs

One of the largest MSME sectors in Indonesia is the Tourism sector. The potential of tourism in Indonesia is very large, as the country possesses an array of popular tourist destinations and panoramic locations. Tourism is also a sector that contributes to economic growth that is quite promising for the country and has a very important role in the development of Indonesia as a whole. The tourism sector provided as much as 9.00% of Indonesia's national employment in 2014 or around 10.32 million people who were in tourism-related sectors. Meanwhile, the impact of the tourism sector on GDP, wages or salaries, and indirect taxes is in the range of 4.01% [11]. As thus, a drop in MSME activities would decimate tourism-based MSMEs. In 2010, tourism-based MSMEs generated 261.06 trillion rupiahs for Indonesia's GDP, and it increased in 2011 into 296.97 trillion rupiahs, reaching 326.24 trillion rupiahs in 2012. Subsequently, it dropped to 278 Trillion rupiahs by 2013 and remained relatively similar ever since. In 2019, Indonesia's tourism sector's profit was 280 trillion Rupiah, and in 2020, it dropped to a mere 51.2 trillion rupiahs.

However, the extension of the impact of the economic downturn on this sector feels unexplored. The MSME sectors which can be said to have been most affected by this pandemic have been previously recorded by several studies, but a majority of these studies were conducted in the early years of the pandemic, namely 2020. Furthermore, there are only bits and pieces of the pandemic's effects on the tourism industry. In addition, the plans launched by the government at that time had been given time, two years to implement, but there is currently no checking process or a checklist has been made as of January 2022 to be used to check Indonesia's economic situation. Therefore, it can be said that Indonesia does not yet have an updated document that has compiled the progress that is currently being developed or the state of Tourism-based MSMEs in 2021 which should be used to plan forward and orchestrate the rebuilding process of the nation's economy.

The purpose of writing this scientific paper is to find out and analyze the impact caused by the COVID-19 virus on the development of the tourism sector MSME business in Indonesia, which fields are most severely affected by the COVID-19 lockdown, and what things the government can mobilize to tackle this problem. The information that will be recorded in this research will be through relevant and current sources, namely journals and information that we can access as of January 2022, so that it can better reflect the current state of Indonesian tourism-based MSMEs.

Research Questions

So, we aim to make up for this shortcoming by creating our own helper document, answering the following basic questions.

1. What are the impacts of the COVID-19 pandemic on MSMEs in the tourism sector?
2. What have regional MSMEs been doing to solve their current COVID-19 pandemic situation?
3. What can the government do to overcome the impact of COVID-19 on the tourism sector MSMEs?

2 Research Technique

This study adopts a descriptive qualitative approach in its goal to understand and analyze the concepts included in the research process. This technique is done with library research and content analysis techniques (a research method that utilizes a secondary source of data written based on and about previous research and references that are related to the research's content and theme).

Our method of choice, observations were carried out by collecting data from various official sources on the Internet, namely documents, previously published journals, and sources relevant to government activities. Because these sources are very relevant to our research, we use them to find more in-depth information. In this way, we can find various impacts that affect MSMEs in the tourism sector so that we can use the existing data as a guide in our research.

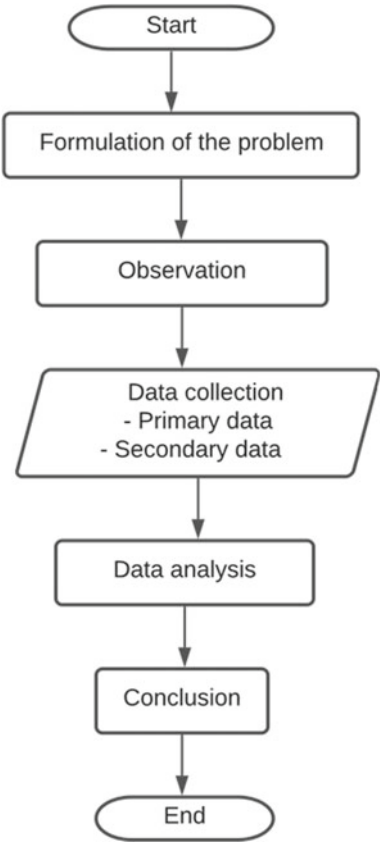


***Image**. Flowchart Research Technique*

3 Research Method

The research method that will be used is a qualitative descriptive method (describing an overview of a study to our use through data or sample study). The data sources we plan to use are books, online journals, and articles from trusted sources, as well as the results of interviews with several people who work or are heavily involved in tourism-based MSME. In this case, the author will collect and analyze the effects of the COVID-19 pandemic on the tourism sector MSME in Indonesia, the solutions that have been deployed so far, and the locations most affected by the pandemic. Researchers will also interpret and record important information from secondary sources.

The selected population is the tourism sector MSME actors affected by the COVID-19 pandemic, who will be observed and studied through secondary sources.

The data collection technique is to make observations from various official sources on the Internet, namely journals that have been previously made and sources relevant to government activities, as well as research about these MSMEs. Then, these sources are analyzed by comparing the existing sources. The results will then be used to draw the author's own conclusions, which are based on real results and logical analysis.

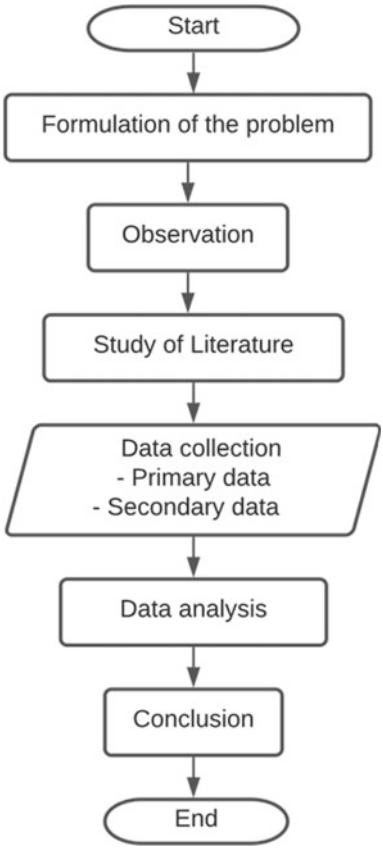


***Image**. Flowchart Research Method*

4 Research Result

4.1 The Pandemic's Impact on General Tourism-Based MSMEs

From our observational activities, we have gathered several data regarding the conditions and state of Indonesia's current state of tourism. As one of the sectors that has been most affected by the COVID-19 pandemic, tourism-based MSMEs has faced a massive loss. Based on data gathered from BPS sources [12], there is a significant reduction in the number of foreign visits to Indonesia. For example, Indonesia got 4.02 million foreign visits in 2020. When compared to 2019, the number of foreign tourists have decreased by an estimated 75.03%, especially in December 2020 (where only 164,088 thousand tourists visited), which is notable since December is a month where tourists usually visit in large droves (Widjaja, 2021). Based on that data, the tourism business experienced a very significant decrease compared to their usual levels of profit. There has been a lot of reduction in workforce, and there are plenty of tourism-based businesses that have closed due to the pandemic due to their dependence on a large number of tourists.

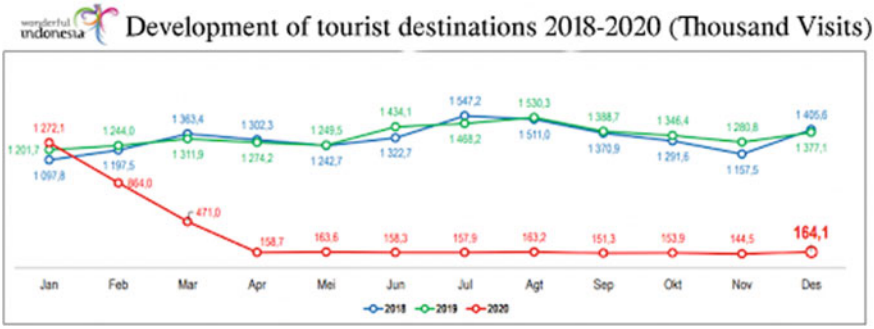


***Image**. Development of tourist destinations 2018-2020 (Thousand Visits) (Figure: BPS, 2021)*

The chairman of the Bali Tourism Board (BTB)/Indonesian Tourism Industry Association (GIPI) Bali, Ida Bagus Agung Partha Adnyana, had reported that over 40,000 hotel arrangements have been canceled in 2020 alone, leading to a loss of One trillion rupiah per month [13]. Ikhsan Ingrabatun, General Chairperson of the Indonesian MSME Association, estimated that the turnover of MSMEs in the non-culinary sector will decline by 30–35%. These things also had a negative impact on other sectors such as micro food and beverage businesses (27% decrease) and wood and rattan handicrafts (17.03%).

The cause of the decline is invariably a number of factors that stemmed from the COVID-19 pandemic and its safety lockdown measures. It can be said that the drastic

decline in consumers was caused by travel restrictions which are still paralyzing the world of tourism. However, the difficulties and decline experienced by MSMEs during the pandemic can be divided into four general problems [5].

First, sales decreased due to a decrease in community activities as consumers. Second, there were difficulties in financing the business due to difficulties in capital turnover due to the decrease in sales. Third, there are product distribution barriers due to product distribution restrictions in specific regions, and fourth is the difficulty of raw materials because SMEs depend on securing raw materials from other industrial sectors. Fourth is the difficulty of raw materials, as SMEs depend on the availability of raw materials in several industrial sectors.

The impact of the pandemic on SMEs is believed to be even greater than expected, due to the high level of vulnerability and lack of resilience due to limited human resources, suppliers, and options for overhauling business models. The impact is felt by MSMEs and the tourism sector because these two sectors have an impact on the country's economic growth.

This journal has a recovery strategy for the MSME tourism sector, namely by providing assistance to business actors, providing tax incentives, relaxation, and credit restructuring, expanding working capital financing, product support, and e-learning training, while the strategy for the tourism sector is the development of tourism products, the development of tourist destinations, the development of tourism human resources, the development of tourism institutions, and the management of tourism infrastructure. Based on data, at least 50% of MSMEs in Indonesia have gone out of business due to the economic shocks caused by the COVID-19 pandemic.

Through data taken from the Central Statistics Agency, we received an estimate that there were around 2.07 million Chinese tourists who came to Indonesia in 2019 and that is a 12.8% share of the total existing global visits. However, visits of that magnitude began to decrease massively in 2020 and subsequently affected the Indonesian economy, which was devastated because other fields that were also involved or became part of tourism also affected, one of which was hotel and restaurant visits. The revenues from these businesses decreased by about 40% because they could only benefit from tourist visits, both local and foreign tourists [14].

Due to the huge impact of losses, other countries are trying to come up with a strategy to reverse this situation back to where it came from. First, the government gives direction to MSME owners to try to develop new business innovations to stay afloat and help unemployed people. Second, for MSME owners who are unable to pay their employees, the government provides a salary subsidy to pay employees. Third, it is given relief in the payment of taxes and loans, namely given a longer period of payment. Fourth, MSME owners are given loans to continue to run their business for those who are still possible. Fifth, think of various ways so that their MSME businesses can be digitized.

Furthermore, the issue is further aggravated by several factors, such as:

- A lack of government standards regarding planning, implementation, and mitigation of the pandemic's impact.
- A lack of technological support to support a map-based data system as a tracking method for various tourism-related key factors in this pandemic situation.
- A weak cross-sectoral and cross-regional pandemic relief coordination, both within government and various tourism industry associations from the central level to the regional level.
- A lack of institutions that specifically handle tourism-related disaster impact mitigation.

In order to gain a comprehensive understanding of the impact and mitigation of the government's policies on various destinations across the country.

4.2 The Pandemic's Impact on Regional MSMEs and Mitigation Efforts

4.2.1 The Pandemic's Impact and Mitigation on Tourism-Based MSMEs in Kupang

The town of Kupang, NTT, is one of the many tourist destination hotspots that dots Indonesia's landscape. Before the COVID-19 Pandemic and the Kupang City PSBB policy was implemented, Kupang City was one of the most sought after destinations by domestic tourists visiting Kupang, and it was a regional center of coastal tourism activities. Tourism-based MSMEs on the three beaches of Kupang City's coastal area serve the purchase of tourism products 7 days a week. Especially at Color Oesapa Beach and Pasir Panjang Beach, the interaction between tourists and the managers of the tourism business unit occurs until the early hours of 2 WITA local time. The products most sold include goods such as coconuts, bananas, lontar fruit, drinks, cigarettes, coffee, and instant noodles. Tourists who made transactions before the arrival of the COVID-19 pandemic and the PSBB policy was enforced could number up to 30 or even 50 transaction visits [15].

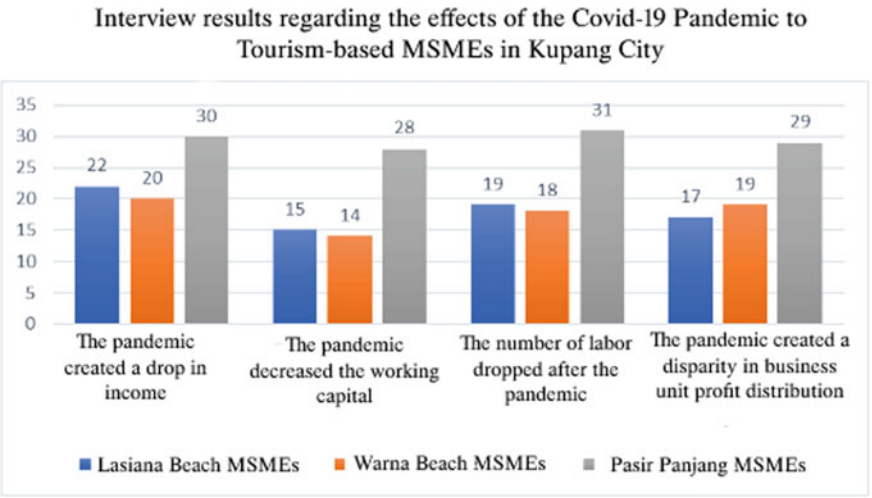


***Image**. Interview Results*
(Source : Nugraha, 2021)

But as the pandemic and PSBB policies hit, according to the data gathered [15], there is a decrease in local tourism income by an average of 70%. A whopping 100% of MSME unit owners in Kupang agreed the pandemic brought about a drop in their income. Out of the 77 units researched, only a mere 5 are still in business. There are major impacts on the area's tourism business, such as increasing capital, reduced tourism workforce, unequal distribution of profits, closing of business units, and other losses [7]. Most of the tourism-based businesses in the area that were impacted by the pandemic due to the COVID-19 diseases are food and beverage businesses and production service businesses experiencing distribution barriers that impact the stock of goods flowing into their inventory. The situation is desperately terminal, especially with the government's warning to stay at home to reduce the spread of COVID-19, which deeply impacted the tourism industry.

The government's attempt to mitigate the dire situation is to loosen the policies of the quarantine by implementing a New Normal protocol. Using the protocol, tourists can travel to tourism destinations while still paying attention to health protocols. The authorities have described this as an initial step and an opportunity for the tourism business unit to be able to return to a profitable position by conducting transactions in coastal tourist attractions. The provisions of the protocols include health-keeping facilities in the form of hand washing facilities, procurement of water reservoirs, body temperature checks for every officer and tourist, and an appeal to always use masks, maintain distance and wash hands when traveling. They also described the three phases of the protocol's implementation: the learning phase where they familiarize themselves with the protocol, the openings of air, sea, and water transports, and implementation (where business runs as usual).

4.2.2 The Pandemic's Impact and Mitigation on Tourism-Based MSMEs in Lampung

As a destination for other tourism activities, Lampung favors its tourism industry to support its regional economic growth. Naturally, one of the drivers of the tourism sector to develop is the existence of MSMEs in tourist areas. But while MSMEs have a significantly positive impact Lampung's development, sources from the BPS said that MSMEs in the tourism sector in Lampung have fluctuated, for the tourism sectors in 2019 by 12.7% and decreased in 2020 by 8.5% because of COVID-19 pandemic [16]. In Lampung, the workforce that operates within the tourism sector has experienced a decrease, a drop of 1.8% in 2019 and 1.5% in 2020, due to a decline in the MSME industry tourism. As a result, the absorption of labor for the sector has also decreased.

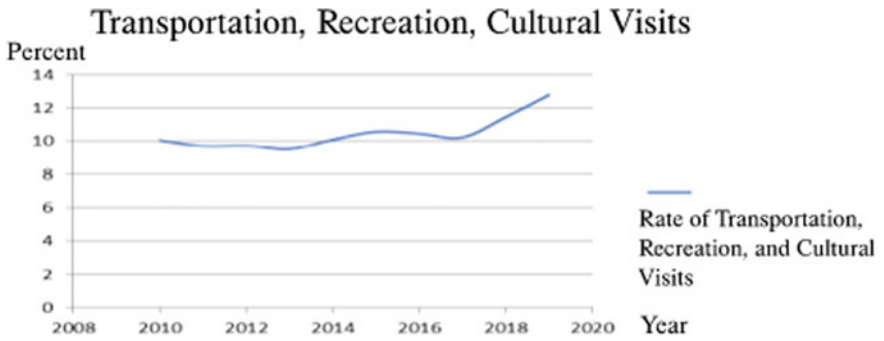


***Image**. Rate Transportation, Recreation, and Cultural Visits*
Figure : (Anggarani, 2021)

Based on the information shown in the following graph, it is obviously visible that the contribution of the MSME sector transportation, accommodation, communication, and cultural tourism to the Provincial GRDP Lampung experienced an increase by a substantial amount. This is even though from 2010 to 2011 had decreased by 9.7% but this contribution figure continued to increase until 2019 by 12.7% and decreased by 8.5% in 2020 due to the pandemic COVID-19.

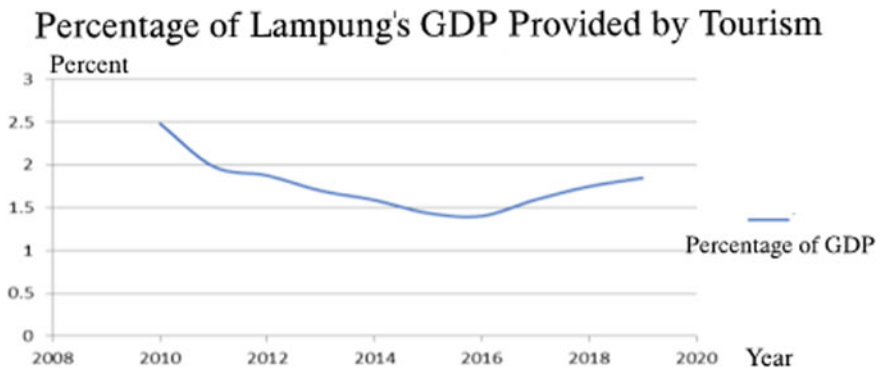


***Image**. Percentage of Lampung's GDP Provided by Tourism*
Figure : (Anggarani, 2021)

Based on the graph above, it can be seen that the contribution of labor who works in the tourism sector to the GRDP of the Lampung Province experiences significant fluctuation. This can be seen in 2010 the contribution of employment in the tourism sector is 2.48% of GDP Lampung Province, this decreased in 2011 by 1.98% until 2016 by 1.4% and continues to increase in 2019 by 1.8%, but this figure has decreased in 2020 by 1.5% due to the COVID-19 Pandemic. Based on the data and discussion above, it can be said that the greater contribution from the tourism sector will increase the economic growth of the Lampung province, while the drop in tourism sector workforce created a visible dent on the economic growth of Lampung.

The government, however, has begun to enact a plan to mitigate this. The method is a transition from physical visits to a new trend in virtual tours. Virtual tour itself is a concept in which a person can use a smart device to view the scenery and have vacation in a location that is accessed through the Internet. While it was not much explored before the COVID-19 pandemic, the standards avoided in tourism (such as being able to gather in crowds without clear limits and capacities) are bringing attention into a method of traveling which does not need such restrictions. In addition, there is also a standard in place on all Lampung locations. Namely, a series of health protocols which consists of mandatory hand washing stations, social distancing, and mandatory masks. This is also accompanied by the implementation of a standard of sort, created to maintain the tourist's trust during the pandemic, called CHSE (Cleanliness, Health, Safety and Environmental Sustainability). Its goal is to promote MSMEs to increase their human resource and competitiveness, and also provide safe and clean tourist staying areas [17].

4.3 Proposals and Solutions to Overcome the Impact of COVID-19 on the Tourism-Based MSMEs

There are various proposals and plans that were considered by the government in their quest to solve the issues caused by the COVID-19 pandemic. One such example is an attempt to adopt a mixed program to save their MSMEs, following similar policies of other nations such as China or the United States [14]. These theories posit that the impact of COVID-19 on the economy can be overcome by following the example of other countries. These countries have implemented various pandemic coverage policies, such as a salary subsidy for struggling MSMEs to help pay for their employee's wage, invest in, and encourage entrepreneurial innovation to absorb the plethora of unemployed workers that this pandemic cause, invest in digitization of business, and suspend the debts or obligations of struggling MSMEs.

In support of this idea, there is an argument [18] that the priority of the post-COVID business world is not to return to its former or "normal" state, but to support the global tourism industry in sustainable development. Of course, some of these goals are also related to the social interventions that have been carried out by many other countries. Below is a graph of various forms of social intervention posted by foreign governments as their solutions to this international problem.

Labor Market Interventions in Several Countries

Type of Intervention	Persentase (%)
Wage Subsidy	36
Workforce rectification through training	11
Adjustment of labor market regulations/governance	11
Work Hour Reduction	3
Total Program	**61**

Image. *Labor Market Interventions in Several Countries*
Figure : (Sutrisno, 2021)

In addition, it must also be noted that there are efforts that have been made by the Indonesian government to overcome these problems, namely short-term tactics, such as implementing strict health protocols and easing or delaying credit payments for MSMEs, and long-term tactics, such as increasing the use of digital technology in local markets to support activities, kick-starting the MSME economy via long-term business plans, and creating a guideline for government-supported MSME development in dealing with a post-COVID-19 business world in particular and the business world in general. The Indonesian government has also proposed a policy of airline

ticket discounts of up to 30% to 40% in order to boost incentives to tourists, including places like Batam and Malang [19].

In addition, there are government-backed aids aimed to revitalize the tourism sector, such as Ida Bagus Agung Partha's claimed effort to boost the sluggish tourism sector [5]. According to the man himself, the government plans to spend around 298.5 billion Rupiah worth of funding to kick-start the heart of the struggling tourism sector. Out of all the funds, around 73 billion Rupiah will fall into the coffers of the Ministry of Tourism to create a project meant to prompt Indonesia tourism with the help of selected foreign influencers and celebrities. The remaining fund shall be used to incentivize airlines and travel agents, as well as promotional activities regarding the tourism project.

To improve the tourism sector and the creative economy, through the Ministry of Tourism and Creative Economy (Kemenparekraf), the Government issues a policy on the necessary steps. The funds prepared to run the program amounted to Rp. 3.3 trillion [20]. It indicates that the government is very serious about implementing the program. According to Anggarini [21], the first step to reviving the tourism-based MSME industry is to create a stimulus for market demand, support, and finance digital platforms to ensure a sufficient number of partnerships, and collaborative efforts to use technology to improve product quality through processing, packaging, marketing, etc.

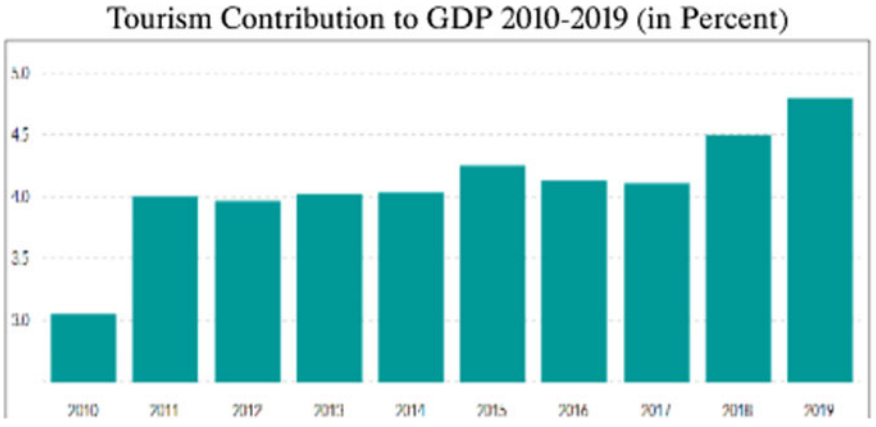


***Image**. Tourism Contribution to GDP 2010-2019 (in percent)*
Figure : (Utami, 2021)

Based on the graph above, over the last few years the contribution of tourism-based MSMEs experienced a significant increase in Indonesia's GDP composition. This matter can be seen from the contribution of the tourism sector: Total exports of goods and services increased sharply from 10% in 2005 to 17% in 2012. Sector tourism contributes directly to GDP of 4.8% in 2019, and the value increased by 0.30 points from in 2018 which is 4.5% [2].

Steps in restoring the slumping tourism industry also includes the preparation of tourist destinations, building a connectivity infrastructure that is competitive with other countries, creating and building attraction tours, increasing training and education to generate a more qualified workforce, and rejuvenating the creative economy. Besides the restoring steps, the Indonesian government is also preparing another program that is preparing additional funds to help the Local Government (Pemda) as well as tourism industry such as hotels and restaurants currently experiencing financial pressure and recovery against decline in Original Income Regions (PAD) due to the ongoing pandemic [20].

5 Discussion

From the results of these observations and research, we discussed whether the COVID-19 pandemic had an impact on the tourism sector MSMEs which became our problem formulation. The results of our discussion that the COVID-19 pandemic has a serious impact on MSMEs in the tourism sector are in accordance with our previous hypothesis. The results of our study also show similar results to the journals we have observed, which specifies that there are drastic losses in a majority of tourism-based MSMEs, especially in divisions like food and drinks and transportation. This is where our research began to show the true picture. While the impact of the pandemic varies from region to region, one fact remains a common element in almost all cases: the direct cause of the drop in tourism, namely the inability of tourists to visit Indonesia and its various tourist destinations due to the limitation imposed. There are also several tourist destinations that have implemented health protocols in accordance with the guidelines made by the Ministry of Health, such as the island of Bali. While they have officially reopened the island, they prioritize tourist comfort and tourists with the aim to build back their trust and help them get used to the health protocols [22].

From this, the government has proposed a series of ideas to halt the pandemic's more negative and drastic effects on the tourism industry that have ranged from short term (an injection of funding, health policies, incentives) or long range (modernization policies, digitalization of the markets). There has also been talks of forgoing standard rebuilding the previous in a post-COVID world and creating a new modernized technology-enhanced system instead. An addition to this is policies from government edicts such as salary subsidies, development encouragement, and forgiving or delaying tax debt. In addition, there are proposals of using technology to simulate the traveling process and bypass the restriction on face-to-face contact that hindered the industry in the first place [23]. This is achieved with audio clips and sounds that are simulated by a VR system, aimed as a remedy for travelers and aiming to bring back customer confidence on Indonesia's tourist destinations. Our team would also like to build on this idea, as the previous ideas seems to be largely explored.

In terms of effectivity, the results of these policies have been shown to be both positive and proven to work in other nations, although it must be admitted that

the time frame of the pandemic (2019–2022) must still be considered during this observation, since a majority of the policy were only enacted a year ago. As such, we plan to use the data we gained from our research to conclude and propose a few additions of our own by picking methods that have worked, drawing inspiration from all interpretations, and utilizing our own hypothesis.

6 Conclusion

Indonesia's tourism sector is one of the backbone of the country's economic growth, accounting for as much as 9% of the nation's employment. As such, it remains one of the hardest hit sectors of the pandemic. The results of our study also show similar results to the journals we have observed, which specifies that there are drastic losses in a majority of tourism-based MSMEs, especially in divisions like food and drinks and transportation, due to the lack of face-to-face interaction and travel limitation that came with the pandemic.

Each region has also enacted their own solutions to these issues. In regions such as Lampung and Kupang, policies have been enacted to lift the suffocating sanctions, such as health protocols (hand washing facilities, procurement of water reservoirs, body temperature checks), qualifications and validations such as HSHE, and new normal protocols. In Bali, their solution involves maximizing customer experience and satisfaction despite their current situation and a series of marketing plans. As there are so many impacts of the COVID-19 pandemic on the tourism sector in Indonesia, there must also be solutions of various kinds from every pore of the government and international community, both short and long term, which must be assisted with actions from the community itself, be it economic or social.

Thus, we plan to make moderate adjustments and exploration regarding the previously mentioned programs and targets based on assumptions that have been updated according to the situation and conditions of the COVID-19 pandemic, and then maintain a program where the assumptions that are the basis are still relevant and can still be adapted to post-COVID-19 conditions.

References

1. Fotiadis A, Polyzos S, Huan TCTC (2021) The good, the bad and the ugly on COVID-19 tourism recovery. Ann Tour Res 87:103117. https://doi.org/10.1016/j.annals.2020.103117
2. Kemenparekraf (2020) Rencana Strategis Kemenparekraf 2020–2024. Kemenparekraf 1–136. https://www.kemenparekraf.go.id/laporan-kegiatan/RENCANA-STRATEGIS-2020_2024-KEMENPAREKRAF~BAPAREKRAF
3. Li X, Yuen KF, Wang X, Wong YD (2021) Contactless technologies adoption during the coronavirus pandemic: a combined technology acceptance and health belief perspective. In:

Technology analysis & strategic management, pp 1–14. Informa UK Limited. https://www.researchgate.net/publication/355251605_Contactless_technologies_adoption_during_the_coronavirus_pandemic_a_combined_technology_acceptance_and_health_belief_perspective
4. Covid-19, New Normal, dan Perencanaan Pembangunan di Indonesia (2020) In Jurnal Perencanaan Pembangunan: The Indonesian Journal of Development Planning 4 (2):240–252. https://doi.org/10.36574/jpp.v4i2.118
5. Sugiri D (2020) Menyelamatkan Usaha Mikro, Kecil dan Menengah dari Dampak Pandemi Covid-19. Fokus Bisnis: Media Pengkajian Manajemen Dan Akuntansi 19(1):76–86. https://www.researchgate.net/publication/343348838_Menyelamatkan_Usaha_Mikro_Kecil_dan_Menengah_dari_Dampak_Pandemi_Covid-19
6. Riadil IG (2020) Tourism industry crisis and its impacts: investigating the Indonesian tourism employees perspectives' in the pandemic of COVID-19. Jurnal Kepariwisataan: Destinasi, Hospitalitas dan Perjalanan 4(2):98–108. Sekolah Tinggi Pariwisata Bandung. https://doi.org/10.34013/jk.v4i2.54
7. Sugihamretha IDG (2020) Respon Kebijakan: Mitigasi Dampak Wabah Covid-19 Pada Sektor Pariwisata. Jurnal Perencanaan Pembangunan: Indonesian J Dev Plann 4(2):191–206. https://journal.bappenas.go.id/index.php/jpp/article/download/113/85/
8. Baker DMA (2/015) Tourism and the health effects of infectious diseases: are there potential risks for tourists? International J Safety Secur Tour Hospit 1(12):1. https://www.palermo.edu/Archivos_content/2015/economicas/journal-tourism/edicion12/03_Tourism_and_Infectous_Disease.pdf
9. United Nations World Tourism Organization (2020) Impact assessment of the Covid-19 outbreak on international tourism. Unwto.Org.
10. Pakpahan AK (2020) COVID-19 dan Implikasi Bagi Usaha Mikro, Kecil, dan Menengah. Jurnal Ilmiah Hubungan Internasional 59–64. https://journal.unpar.ac.id/index.php/JurnalIlmiahHubunganInternasiona/article/view/3870
11. Prasetyo B, Suryoko S (2018) Dampak pengembangan pariwisata terhadap perkembangan umkm pada kawasan wisata dieng. Jurnal Ilmu Administrasi Bisnis 7(4):310–320. https://ejournal3.undip.ac.id/index.php/jiab/article/view/22034
12. Statistik BP (2020) Laporan Perekonomian Indonesia 2020. Jakarta: Badan Pusat Statistik. Badan Pusat Statistik. 2021. Berita Resmi Statistik 1 Februari 2021. https://www.bps.go.id/pressrelease/2021/02/01/1796/jumlah-kunjungan-wisman-ke-indonesia-bulan-desember-2020-mencapai-164-09-ribu-kunjungan-.html
13. Bahtiar RA, Saragih JP (2020) Dampak Covid-19 Terhadap Perlambatan Ekonomi Sektor Umum. Jurnal Bidang Ekonomi Dan Kebijakan Publik 7(6):19–24. http://berkas.dpr.go.id/puslit/files/info_singkat/Info%20Singkat-XII-6-II-P3DI-Maret-2020-1982.pdf
14. Sutrisno E (2021) Strategi pemulihan ekonomi pasca pandemi melalui sektor umkm dan pariwisata. Jurnal Kajian Lembaga Ketahanan Nasional Republik Indonesia 9(1). http://jurnal.lemhannas.go.id/index.php/jkl/article/view/214
15. Nugraha YE (2021). Dampak pandemi covid 19 padaunit usaha pariwisata di kawasan pesisir kota kupang. Jurnal Industri Pariwisata 3(2). http://jurnal.usahid.ac.id/index.php/pariwisata/article/view/411/324
16. Anggarini, D. (2021). Kontribusi UMKM Sektor Pariwisata Pada Pertumbuhan Ekonomi Provinsi Lampung 2020. EKOMBIS REVIEW: Jurnal Ilmiah Ekonomi Dan Bisnis 9(2):345–355. Accessed through: https://jurnal.unived.ac.id/index.php/er/article/view/1462
17. Lestari R, Rachmawati T, Kamandanu FA, Syahrobi D (2021) Lampung tourism supply during the pandemic and post Covid-19 pandemic. In: Proceedings of the 2nd international Indonesia conference on interdisciplinary studies (IICIS 2021). Atlantis Press. https://doi.org/10.2991/assehr.k.211206.012
18. Gössling S, Scott D, Hall CM (2020) Pandemics, tourism and global change: a rapid assessment of COVID19. J Sust Tour 1–20. https://www.tandfonline.com/doi/full/https://doi.org/10.1080/09669582.2020.1758708
19. Atmojo ME, Fridayani HD (2021) An assessment of Covid-19 pandemic impact on Indonesian tourism sector. J Govern Public Policy 8(1):1–9. Universitas Muhammadiyah Yogyakarta. https://journal.umy.ac.id/index.php/GPP/article/download/9792/6095

20. Utami B, Kafabih A (2021) Sektor Pariwisata Indonesia Di Tengah Pandemi COVID 19. Jurnal Dinamika Ekonomi Pembangunan 4(1):383–389. https://doi.org/10.33005/jdep.v4i1.198
21. Anggarini DT (2021) Upaya Pemulihan Industri Pariwisata Dalam Situasi Pandemi Covid -19. Jurnal Pariwisata 8(1). https://doi.org/10.31294/par.v8i1.9809
22. Bascha UF, Reindrawati DY, Witaningrum AM, Sumardiko DS (2020) Dampak Pandemi COVID-19 Terhadap Minat Masyarakat dalam Berwisata dan Sosialisasi Penerapan Protokol New Normal Saat Berwisata. Jurnal Abdidas 1(6):560–570. https://abdidas.org/index.php/abdidas/article/view/129
23. Gegung EM (2021) International tourism and the COVID-19 pandemic: the use of virtual reality to increase tourism destination sustainability and how users perceive the authenticity of VR experiences. Jurnal Kepariwisataan Indonesia: Jurnal Penelitian dan Pengembangan Kepariwisataan Indonesia 15(1):9–15. Ministry of Tourism and Creative Economy. https://doi.org/10.47608/jki.v15i12021.9-15

Modelling Auto-scalable Big Data Enabled Log Analytic Framework

Deshpande Kiran and Madhuri Rao

Abstract Log generation is a continuous process that generates large amounts of log data in various forms and rates that may be analysed to acquire valuable insights.Various open-source and commercial alternatives for integrating log acquisition, storage, retrieval, and analytic have been created in the light of the current Big Data era and the importance of log analysis. Taking into consideration benefits mentioned in literature about Spark and ELK for log analysis, unifying Spark and Elastic Stack capabilities with other open-source services and platforms helpful for log analysis can really prove beneficial for developing an efficient framework for scalable log management. In this study, we present auto-scalable Big Data enabled log analytic framework integrating Apache Flume, Apache Kafka, ELK Stack, and Spark. By combining the functionality of all of the platforms mentioned, the framework seeks to deliver a platform that integrates online and offline Big Data processing methods while also enabling real-time analytic. We also deployed crucial components of the modelled framework on a horizontally scalable Kubernetes cluster implemented over AWS cloud, anticipating performance advantages from the dynamic scalability given through implementation design. Furthermore, the performance of the modelled novel framework is analysed using various search and analytic criteria both before and after deployment over AWS cloud. It was discovered that the novel framework outperformed in terms of several search and analytic criteria used to evaluate performance. We further critically examined the implications of this cloud-based deployment on the framework's performance.

Keywords Big Data · AWS cloud · Kubernetes · Log analytic · Performance evaluation

D. Kiran (✉) · M. Rao
Thadomal Shahani Engineering College, University of Mumbai, Mumbai, Maharashtra, India
e-mail: deshpande.82@gmail.com

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_64

1 Introduction

Much recent research has concentrated on the problems of real-time log analysis. Analytic based on Big Data has progressed to the point where it can successfully address IT infrastructure-related security issues. Considering the study focus on log analytic, stream, and batch log analytic can help IT managers in ensuring IT infrastructure security as well as taking proactive measures. The Internet plays a significant role in the operations of any organisation. Examining Internet and Intranet traffic is critical for detecting anomalous or hostile behaviour that might endanger a company's IT resources security [1, 2]. Analysing logs and performance of vital IT infrastructure resources that include critical events linked to Intranet and Internet traffic, such as web servers, NIDS servers, proxy servers, peripheral routers, and firewalls, is essential [3–6]. Log data is frequently voluminous and constantly growing [2, 4]. In order to discover an in-depth correlation, forensic analysis of events requires the use of both real-time and historical data. In this situation, cloud computing, Big Data platforms and services supporting log analysis, and Kubernetes clustering will prove beneficial to build scalable solution for log analytic.

In this study, we present auto-scalable Big Data supported log analysis platform using AWS cloud and Kubernetes cluster for deployment of its critical open-source platforms and services. By combining the benefits of Big Data processing methods, AWS cloud, and Kubernetes clustering, the framework seeks to deliver a platform that integrates stream and batch processing for scalable log analytic. It may also be used as a reference for implementing Elasticsearch-based log analysis following stream and batch processing in order to fulfil contemporary Big Data log analytic requirements. This is the first attempt, to our knowledge, to integrate Big Data, AWS cloud, and Kubernetes cluster capabilities for tackling the most of the challenges relating with Big Data log analytic such as scalability, availability, and reliability. In this study for critically evaluating performance, we put core components of our log analysis framework on a horizontally scalable Kubernetes cluster deployed over cloud, anticipating performance gains through the dynamic scalability provided.

With the approach used in this study since ingested logs are already pre-processed and indexed through stream processing, indexed logs can be readily used for batch processing implemented through Apache Spark without applying any further ETL process, which saves extra effort and time required otherwise before applying batch processing. Furthermore, with the approach used in implementation since batch-processed logs can move to stream processing framework graphical visualisations of batch-processed logs are possible through visualisation tool used for stream processed logs, which eliminates the need of an independent visualisation tool for batch-processed log data. The rest of the paper is organised as follows. Section 1 provides a brief overview of the study, followed by Sect. 2 which addresses related work. The implementation architecture utilising free and open-source software (FOSS) is described in Sect. 4. The proposed novel framework's performance is evaluated in Sect. 5. In Sect. 6, the concluding section summarises the research findings.

2 Related Work

This section looks at the many ways and procedures developed by academic and industrial groups to deal with the issues of Big Data enabled cloud integrated log analytics. We also present a literature assessment of the current degree of acceptability and maturity of open-source platforms for tackling log analytic difficulties in recent years. We also want to highlight a variety of ways for combining stream and batch processing through this study.

2.1 Big Data Processing: Conceptual Understating and Recent Work

With the prospect of extracting value from information in hand, research into Big Data analysis approaches and technologies is becoming increasingly popular. The survey papers [7–9] are intended to examine the idea, architecture, and typical Big Data processing systems. The importance of hybrid processing systems is discussed, as well as the conceptual representation and comparison of batch and stream data processing systems. During the preceding decade, there had been a lot of research done on batch and stream processing. A common computing architecture for integrating stream and batch processing is proposed in the paper [10]. It investigates the building of a layer between MapReduce applications and the streaming platform in order to combine the features of stream processing with the programming convenience and familiarity with MapReduce batch processing [10]. Batch processing frameworks are so popular, and there has been a lot of study into developing extra functionality for batch processing frameworks in order to allow stream processing. As highlighted in the foundational work [11], streaming analytics is an emerging research topic that focuses on important challenges of real-time, scalable and reliable analytic. The authors of the study [12] emphasised difficulties and possibilities in the visualisation of massive data as well as an presented innovative technique for visualising students' learning activities activities on social networks [13].

2.2 Comparative Assessment of Log Analysis Frameworks Using Big Data

The study [14] examines the merits and drawbacks of trending distributed stream processing platforms for IoT applications. In [3], the authors published a research on offline processing of Big Data using popular Big Data frameworks Hadoop and Spark, indicating that Spark outperforms Hadoop in a performance comparison. The authors of the seminal publication [15] evaluated and contrasted Apache Spark and Elasticsearch with Logstash, two prominent frameworks for processing log data.

Authors conducted studies on a variety of issue solutions of varying complexity to assess the influence on non-functional characteristics such as processing time and resource utilisation. The original work [16] emphasised the use of open-source platforms, proving their utility by comparing commercial and open-source log analysis solutions. Traditional detection algorithms struggle to identify fraudulent events due to Big Data features such as volume, variety, and velocity. Son et al. [16] emphasised the advantages of open-source alternatives over commercial log analysis frameworks. More et al. [17] showed the application of Big Data technology for threat detection [13].

2.3 Survey Regarding Status of Adoption and Maturity of Open-Source Platforms into Big Data Log Analytics

In [18] to handle Big Data properly, the authors examined emerging trends in Big Data analytics. The foundation study [19] presents analysis of Squid Proxy Logs by using Apache Spark batch processing. The conference paper [20] describes a high-performance log data analysis platform that combines distributed NoSQL database technology with the Apache Spark framework to provide correct insights into system administrators and end-users. Therdphapiyanak et al. [5] demonstrated large-scale log analysis for identifying abnormal activity from enormous logs using Hadoop. A few researchers advocated [21, 22] using Elastic Stack to simply and swiftly manage Big Data processing problems. In [23], Liu et al. developed an ELK Stack-based cyberattack detection strategy based on network log analysis and visualisation. In their work, L. Chen et al. used the open-source log collection platform ELK, the lightweight log collector Filebeat, and the distributed message queue Kafka to construct a Docker container log collecting and processing system [24]. Sanjappa and Ahmed proposed a platform that gathers [25] heterogeneous logs using Logstash for discovering harmful activities in the network. The author demonstrated how to leverage the entire ELK ecosystem for successful log file analysis and easy-to-understand insights [22, 25]. YuanTing Wang et al. reported research targeted at creating a monitoring system that uses the ELK Stack to deal with issues of weak Wi-Fi connections [26]. Prakash et al. [27] presented their work in which the ELK Stack was leveraged to efficiently geo-identify website visitor traffic via generated logs. The use of log analysis to solve security issues in IT infrastructure is evident in the literature. Debnath et al. showed how machine learning was used to find patterns in application logs, which was then integrated with real-time log processing to construct sophisticated log analysis applications in [28]. P. He suggested data mining techniques in [29] to provide a solution that ensures widespread usage of logs for system management with guaranteed reliability. Metha et al. created a system utilising ELK, Spark, and Hadoop to identify anomalies in network interface logs in near real time [6]. By utilising Elasticsearch and Spark separately in their framework, Li et al. provided a technique for speeding log analysis using Elasticsearch and Spark [13, 30].

2.4 Assessment of Cloud Computing and Kubernetes Clustering in Designing Scalable Big Log Analytic Solutions

The review article [31] suggests future advancements and recommendations for further studies for cloud computing that will make it easier to translate Big Data properties into values for the applications in question. Lin et al. suggest merging Hadoop and Spark capabilities to generate a unique cloud platform with offline and real-time processing capacity [3, 4]. To fulfil the demand for dynamic computing resources, cloud computing is used in a landmark article [30] to make clustering easier and the log mining process faster. Firouzi et al. employed cloud computing to overcome IoT network constraints such as limited processing power and storage capacity in [32]. Bagnasco et al. showed how to leverage the ELK environment to monitor cloud-based research applications [33]. In [2], Li et al. built a log analysis system using cloud for Apache Spark and Elasticsearch deployment with the purpose of accelerating the log analysis process for HTTP and FTP data. In the study [12], Nguyen, Thanh-Tung, Yu-Jin Yeom, Taehong Kim, and colleagues used a variety of experiments to disseminate critical knowledge about the Horizontal Pod Autoscaler mechanism of Kubernetes, an open-source container orchestration framework that allows high scalability and availability through various auto-scaling mechanisms [13].

3 Architecture for Deployment of Proposed Framework

The envisioned framework's operational architecture is depicted in Fig. 1. This section covers the intended framework's operating process in place in depth. Flume is used for the log ingestion. Being source, it aggregates real-time logs from internal network, peripheral firewall, and application servers and sends collected logs to Kafka as a destination. Kafka creates topics on logs collected and uses Java objects to manage topics created. Zookeeper interfaces with Kafka on the back end. To clone, restore, and maintain these topics, Kafka integrates Zookeeper at back end. Flume delivers log data to Kafka producer, while Kafka consumer feeds the logs to Logstash for further processing. Logstash is responsible for processing and classifying Apache Kafka's unprocessed logs. It parses the data by creating attribute names from topics retrieved from the Kafka consumer. When building an Elasticsearch index, these attribute names come in handy. Simultaneously, Logstash screens content using grok, a filter plug-in that finds certain patterns and eliminates stuff does not fit in that pattern. In Elasticsearch, this makes data organisation and index construction easier. Index mapping created through timestamps may be employed on log data to provide visualisation through Kibana. Visualisations based on Elasticsearch indices let people understand real-time log data in terms of graphical analysis. Packetbeat and Metricbeat are in charge of supervising the performance of IT assets, services, and Intranet in real time by sending metrics related to ELK Stack [13]. To transport

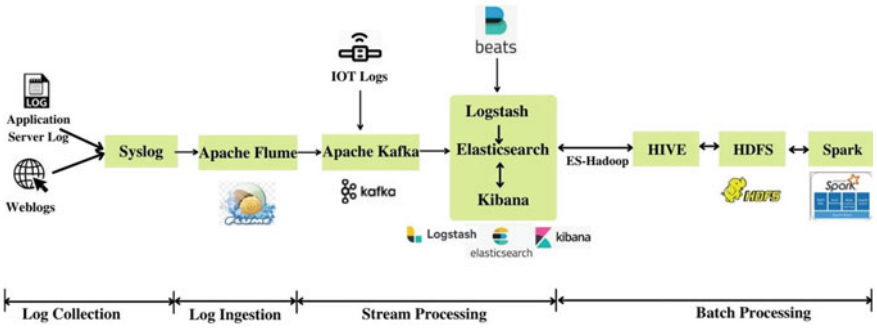


Fig. 1 Implementation architecture [13]

tagged logs to HDFS, the ES-Hadoop connector is used by Elasticsearch. Hive is Hadoop's user interface. For data administration, Hive features a SQL-like interface. The MapReduce technique is used by the Hadoop architecture to perform batch analysis through HDFS. Because of the slowness of file system read-write operations, MapReduce has processing limits. Spark provides better approach to MapReduce latency during batch processing in this scenario. Spark-scala user interface is used to store the processed logs in HDFS. These batch-analysed logs would be sent to Elasticsearch through the ES-Hadoop and Hive. Finally, Elasticsearch can use Kibana to provide the data visualisations needed for comprehensive, low-latency analysis of batch-analysed historical log information [13]. To transport indexed data from Elasticsearch to HDFS, the ES-Hadoop connector is used. Apache Hive is Hadoop's user interface. For data administration, Hive provides a SQL-like interface. The Hadoop architecture uses the MapReduce mechanism to do batch processing through HFDS. MapReduce has processing restrictions due to the latency of file system read-write operations. Apache Spark provides a 100-fold quicker solution to MapReduce latency during batch processing in this situation. By fetching HDFS data into main RAM, the Apache Spark architecture uses resilient distributed datasets (RDDs), data frames, and transitions. The processed data is stored in HDFS using the spark-scala user interface. The ES-Hadoop connection delivers the batch-processed optimal data to Elasticsearch via the Hive interface. Finally, Elasticsearch can use Kibana to provide the visualisations necessary for deep, low-latency analysis of batch-processed historical data [13].

4 Modified Implementation Architecture for Providing Dynamic Scalability to Modelled Framework

We put critical components of our framework on a horizontally scalable Kubernetes cluster deployed over cloud, anticipating performance gains through the

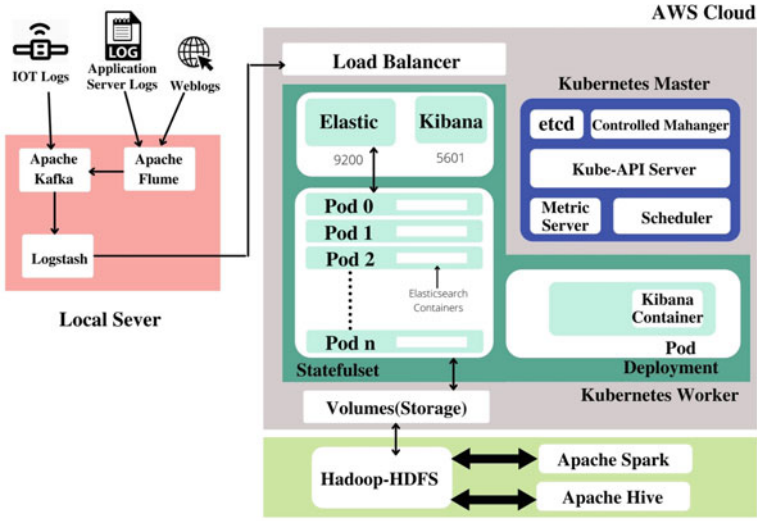


Fig. 2 Architecture for providing dynamic scalability to modelled framework

dynamic scalability provided. Figure 2 represents modified architecture for providing dynamic scalability to modelled framework. Kubernetes cluster will contain pods on which Elasticsearch containers are deployed. Pods with Elasticsearch containers will be dynamically scaled whenever load increases on Elasticsearch during log ingestion as well as during real-time search and analytic as per the criteria in consideration. This horizontally scalable architecture has huge impact on performance of novel framework since it reduces log analysis time in major way.

The application load balancer in AWS receives pre-processed logs using Kafka and Logstash. Over IPv4, Route 53 offers DNS resolution. Elasticsearch and Kibana services provide socket connection to Elasticsearch stateful set and Kibana deployment. Stateful set preserves data for Elasticsearch during scale-out, scale-in processing of Elasticsearch pods. Using service, indexed data is ingested into Elasticsearch pods. The size of each pod is determined by the volume of data packets. For visualisation, the Kibana deployment uses data from stateful set storage. Metric server is used by Kubernetes master to monitor worker performance depending on scaling. The etcd deamon, which is installed on the master, generates key-value-based logs with the help of the controller manager and scheduler. In this case, the kube-API server is crucial for rest-API communication between the master and worker. Using storage volumes, S3 buckets provide persistent storage for processed data. For batch processing, Hadoop architecture built on a local file system leverages S3 volume data. Spark and Hive are two Apache projects that do batch processing on local storage.

5 Performance Evaluation of Framework

To compare performance, various log files of varying size generated through the deployment of the proposed framework in a real-time context are used. Using the previously defined log volume sizes and the queries supplied below, the time necessary to explore for and analyse a certain log pattern from labelled and classified logs is determined [13].

1. Query I: Analysing Apache log to visualise number of unique daily hosts per hour over day.
2. Query II: Analysing Apache logs to visualise 404 errors per day over week period.
3. Query III: Analysing Apache Logs to visualise the top twenty 404 response code hosts over month period.
4. Query IV: Analysing Apache logs to visualise the top twenty 404 response code end points over month period.

Through proposed framework, it becomes very easy to identify, monitor, and analyse malicious events from a number of server log and network traffic in real time in sync with events listed earlier in this section. Timings listed in following table represent the time at which particular event-related logs are seen in different open-source tools responsible for stream processing to cross-check whether logs are ingested and indexed in near real time. Following naming conventions are used for different timings and events while listing in following table. Timings listed in table with respect to events indicate that logs are ingested and seen in near real time across all platforms used for log analysis implemented through novel framework (Table 1).

1. T1: Log recording time of event on actual server log directory.
2. T2: Log recording time of event on Apache Kafka
3. T3: Time at which logs are ingested and indexed in Elasticsearch.
4. E1: Access request from hosts with poorly rated IPs.
5. E2: Server access requests which have received 404 and 403 response Codes.
6. E3: Probing perimeter firewall SSH service on port 22 which is actually running on another non-standard port.
7. E4: Server access requests from blocked user agents.

5.1 Interpretation of Performance Evaluation Before Providing Dynamic Scalability

Before offering dynamic scalability to the modelled framework, query execution time necessary to discover certain log patterns from labelled and categorised logs across different log sizes is documented. Figure 3 depicts a graphical examination of the values provided. It is discovered that the visualisation time required for Query III and IV is much greater than that necessary for Query I and II, demonstrating the

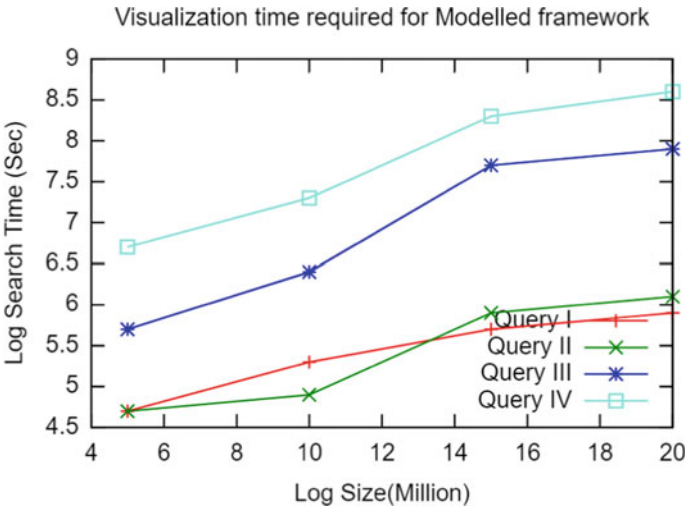


Fig. 3 Search and analysis time observed before providing dynamic scalability for different queries under consideration

Table 1 Timings of event-related logs recorded at different stages of framework

	T1	T2	T3
E1	17:02:2021 13:04:30	17:02:2021 13:04:31	17:02:2021 13:04:32
E2	21:02:2021 18:14:07	21:02:2021 18:14:08	21:02:2021 18:14:08
E3	03:03:2021 08:54:47	03:03:2021 08:54:48	03:03:2021 08:54:49
E4	12:03:2021 03:27:26	12:03:2021 03:27:28	12:03:2021 03:27:30

need for constructing a scalable Kubernetes cluster for deployment of Elasticsearch being critical component of framework. Further, this section goes over performance evaluation before and after introducing dynamic scalability by shifting essential components of the innovative framework to a horizontally scalable Kubernetes cluster deployed over AWS Cloud.

5.2 Interpretation of Performance Evaluation After Providing Dynamic Scalablity

Elasticsearch is an essential core component in modelled framework since it not only handles very crucial work of log processing, indexing at initial stage of log analysis but also has to provide huge volumes of indexed log data while searching and analysing as per the requirements. So providing dynamic scalability to Elasticsearch through auto-scalable Kubernetes cluster deployed on cloud becomes impor-

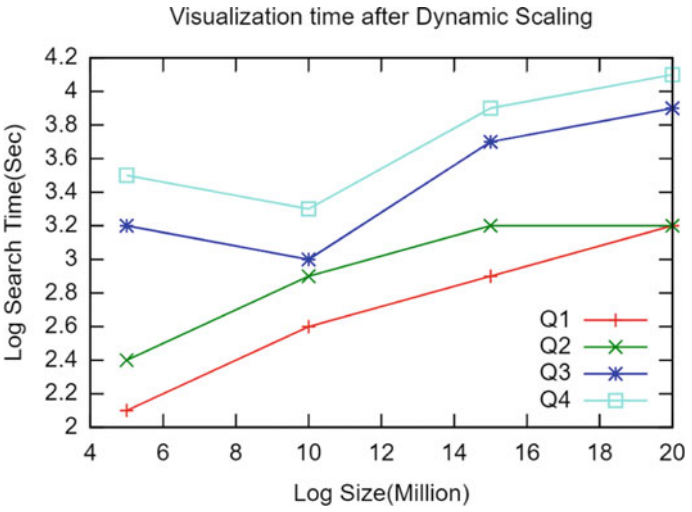


Fig. 4 Search and analysis time observed after providing dynamic scalability for different queries under consideration

tant. After providing dynamic scalability to modelled framework through Kubernetes cluster deployed on AWS cloud, it has been observed that there is a major reduction in search and analysis time required for all queries mentioned earlier in previous subsection. Graphical analysis of same is represented through Fig. 4. Although reduction in search and analysis time is observed in all queries, significant reduction is observed in search and analysis time of Query III and IV as compared to performance evaluation of novel framework discussed earlier before providing dynamic scalability.

In addition, the graphical analysis in Fig. 5 demonstrates the search and indexing latency of Elasticsearch stateful set over the various log sizes mentioned earlier. As part of this investigation, an auto-scalable Kubernetes cluster with pods containing Elasticsearch containers is created by using an AWS cloud t2.Xlarge instance for both Kubernetes master and Kubernetes worker node. A graphical examination demonstrates an increase in search and indexing delay after 25 million logs. Following that, we experimented with higher configuration AWS cloud instances for Kubernetes cluster deployment to observe its impact on search and indexing latency.

Graphical analysis in Fig. 6 depicts the proportional positive impact on indexing latency when we began employing better configuration AWS instances with optimised network performance. C5d.Xlarge has essentially constant indexing latency even for larger log sizes, which was not the case with lower configuration instances such as t2.Xlarge and t3.Xlarge. As a result, we infer that the performance of the modelled framework after deployment on cloud in terms of Kubernetes cluster is also dependent on the type of cloud services used for deployment of auto-scalable Kubernetes cluster. It has been observed same with respect to search latency.

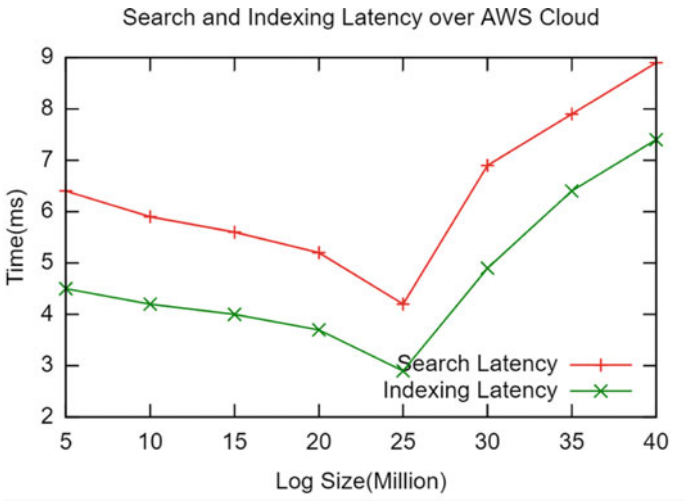


Fig. 5 Search and indexing latency of Elasticsearch stateful set over t2.Xlarge

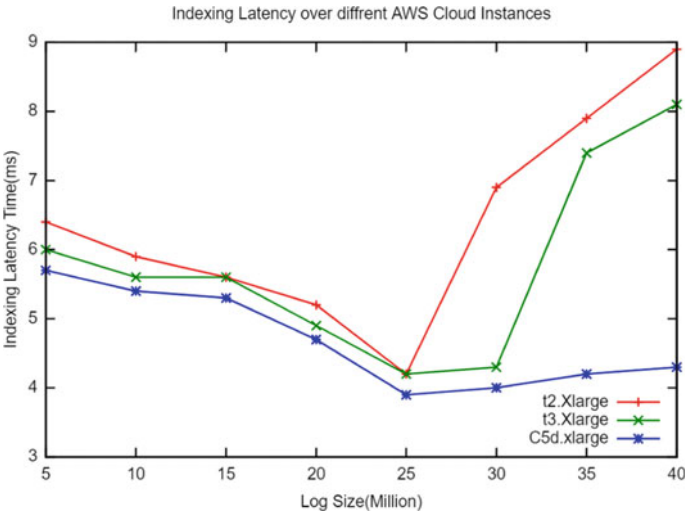


Fig. 6 Indexing latency of Elasticsearch stateful set over different AWS cloud instances

6 Conclusion

This research explores benefits through unification of stream and batch processing taking into consideration need of next generation Big Data log analysis, understanding its relevance to other Big Data log analytic solutions and the scope of open-source technologies which can be opted for removing the existing challenges involved in Big Data log analytic. This exploration has brought out the most pervasive and

well-organised realistic framework that will provide open-source, scalable, reliable Big Data enabled cloud integrated approach which can be widely used by small-scale organisations, entrepreneurs to monitor performance of IT infrastructure and its security analysis.

References

1. Chaudhari S, Maurya VK, Singh V, Tomara SS, Rajana A, Rawata A (2019) Real time logs and traffic monitoring, analysis and visualization setup for IT security enhancement. In: Next generation computing technologies (NGCT-2019)
2. Li Y, Jiang Y, Gu J, Lu M, Yu M, Armstrong EM, Thomas H, David M, McGibbney LJ, Frank G, Yang C (2019) A cloud-based framework for large-scale log mining through Apache Spark and Elasticsearch. Appl. Sci. (Switzerland) 9(6) (2019)
3. Mavridis I, Karatza H (2017) Performance evaluation of cloud-based log file analysis with Apache Hadoop and Apache Spark. J Syst Software 125:133–151. ISSN 0164-1212. https://doi.org/10.1016/j.jss.2016.11.037
4. Lin X, Wang P, Wu B (2013) Log analysis in cloud computing environment with Hadoop and Spark. In: 2013 5th IEEE International conference on broadband network and multimedia technology, pp 273–276. https://doi.org/10.1109/ICBNMT.2013.6823956
5. Therdphapiyanak J, Piromsopa K (2013) Applying Hadoop for log analysis toward distributed IDS. In: Proceedings of the 7th international conference on ubiquitous information management and communication (ICUIMC'13). Association for Computing Machinery, New York, NY, USA, Article 3, 1–6. https://doi.org/10.1145/2448556.2448559
6. Mehta S, Kothuri P, Garcia DL (2018) Anomaly detection for network connection logs. arXiv:1812.01941
7. Yu S, Sun X, Zhang X, Xia Z, Bertino E (2021) Data processing and development of big data system: a survey. In: Advances in artificial intelligence and security (ICAIS 2021). Communications in computer and information science, vol 1423. Springer, Cham. https://doi.org/10.1007/978-3-030-78618-234
8. Lv Z, Song H, Basanta-Val P, Steed A, Jo M (2017) Next-generation big data analytics: state of the art, challenges, and future research topics. IEEE Trans Ind Inform 13(4):1891–1899. https://doi.org/10.1109/TII.2017.2650204
9. Hu H, Wen Y, Chua T-S, Li X (2014) Toward scalable systems for big data analytics: a technology tutorial. IEEE Access 2:652–687. https://doi.org/10.1109/ACCESS.2014.2332453
10. Harvan M, Locher T, Sima AC (2016) Cyclone: unified stream and batch processing. In: 2016 45th International conference on parallel processing workshops (ICPPW), pp 220–229. https://doi.org/10.1109/ICPPW.2016.42
11. Xie W, Li P, Xu H (2018) Architecture and implementation of real-time analysis system based on cold chain data. In: Barolli L, Javaid N, Ikeda M, Takizawa M (eds) Complex, intelligent, and software intensive systems (CISIS 2018). Advances in intelligent systems and computing, vol 772. Springer, Cham. https://doi.org/10.1007/978-3-319-93659-8
12. Tam NT, Song I, Kim K, Joukov N (2016) Big data visualization in information science and applications (ICISA). Lecture notes in electrical engineering, vol 376. Springer, Singapore. https://doi.org/10.1007/978-981-10-0557-2-40
13. Deshpande K, Rao M (2021) An open source framework unifying stream and batch processing. In: 3rd International conference on inventive computation and information technologies (ICICIT-2021)
14. Nasiri H, Nasehi S, Goudarzi M (2019) Evaluation of distributed stream processing frameworks for IoT applications in smart cities. J Big Data 6:52. https://doi.org/10.1186/s40537-019-0215-2

15. Boros A, Lehotay-Kéry P, Kiss A (2020) A comparative evaluation of big data frameworks for log processing. In: Proceedings of the 11th international conference on applied informatics, Eger, Hungary, 29–31 Jan 2020, published at http://ceur-ws.org
16. Son SJ, Kwon Y (2017) Performance of ELK stack and commercial system in security log analysis. In: 2017 IEEE 13th Malaysia international conference on communications (MICC), pp. 187–190. https://doi.org/10.1109/MICC.2017.8311756
17. More R, Unakal A, Kulkarni V, Goudar RH (2017) Real time threat detection system in cloud using big data analytics. In: 2017 2nd IEEE International conference on recent trends in electronics, information and communication technology (RTEICT), Bangalore, pp 1262–1264
18. Naga Rama Devi G, Kumar A, Mozar S (2018) Emerging trends in big data analytics—a study. In: ICCCE. Lecture notes in electrical engineering, vol 500. Springer, Singapore. https://doi.org/10.1007/978-981-13-0212-1-57
19. Mishra DD, Pathan S, Murthy C (2018) Apache Spark based analytics of squid proxy logs. In: IEEE International conference on advanced networks and telecommunications systems (ANTS), pp 1–6. https://doi.org/10.1109/ANTS.2018.8710044
20. https://kafka.apache.org/
21. Bajer M (2017) Building an IoT data hub with elasticsearch, Logstash and Kibana. In: 2017 5th International conference on future internet of things and cloud workshops (FiCloudW), pp 63–68. https://doi.org/10.1109/FiCloudW.2017.101
22. Al-Mahbashi IYM, Potdar MB, Chauhan P (2017) Network security enhancement through effective log analysis using ELK. In: International conference on computing methodologies and communication (ICCMC), pp 566–570. https://doi.org/10.1109/ICCMC.2017.8282530
23. Liu JC, Yang CT, Chan YW et al (2021) Cyberattack detection model using deep learning in a network log system with data visualization. J Supercomput. https://doi.org/10.1007/s11227-021-03715-6
24. Chen L, Liu J, Xian M, Wang H (2020) Docker container log collection and analysis system based on ELK. In: International conference on computer information and big data applications (CIBDA), pp 317–320. https://doi.org/10.1109/CIBDA50819.2020.00078
25. Sanjappa S, Ahmed M (2017) Analysis of Logs by Using Logstash. In: Satapathy S, Bhateja V, Udgata S, Pattnaik P (eds) Proceedings of the 5th international conference on frontiers in intelligent computing: theory and applications. Advances in intelligent systems and computing, vol 516. Springer, Singapore. https://doi.org/10.1007/978-981-10-3156-4_61
26. Wang YT, Yang CT, Kristiani E, Chan YW (2018) The implementation of Wi-Fi log analysis system with ELK stack. In: Hung J, Yen N, Hui L (eds) Frontier computing. FC 2018. Lecture notes in electrical engineering, vol 542. Springer, Singapore. https://doi.org/10.1007/978-981-13-3648-528
27. Prakash T, Kakkar M, Patel K (2016) Geo-identification of web users through logs using ELK stack. In: Proceedings of the 2016 6th International conference cloud system and big data engineering (confluence), Noida, India, 14–15 Jan 2016, pp 606–610
28. Debnath B et al (2018) LogLens: a real-time log analysis system. In: 2018 IEEE 38th International conference on distributed computing systems (ICDCS), pp 1052–1062. https://doi.org/10.1109/ICDCS.2018.00105.
29. He P, Zhu J, He S, Li J, Lyu MR (2018) Towards automated log parsing for large-scale log data analysis. IEEE Trans Depend Secure Comput 15(6):931–944. https://doi.org/10.1109/TDSC.2017.2762673
30. Li Y, Jiang Y, Hu F, Yang C, Armstrong, Huang T, Moroni D, Fench C (2016) Leveraging cloud computing to speedup user access log mining. In: Proceedings of the OCEANS 2016 MTS/IEEE Monterey, Monterey, CA, USA, 19–23 Sept 2016
31. Yang C, Huang Q, Li Z, Hu KLF (2016) Big Data and cloud computing: innovation opportunities and challenges. Int J Digital Earth 10(1):13–53. https://doi.org/10.1080/17538947.2016.1239771
32. Kolajo T, Daramola O, Adebiyi A, (2019) Big data stream analysis: a systematic literature review. J Big Data 6:47. https://doi.org/10.1186/s40537-019-0210-7

33. Bagnasco S, Berzano D, Guarise A, Lusso S, Masera M, Vallero S (2015) Monitoring of IaaS and scientific applications on the Cloud using the Elasticsearch ecosystem. Proc. J. Phys. 608:012016

Robust Cuckoo Search Enabled Fuzzy Neuro Symbolic Reasoning-Based Alzheimer's Disease Prediction at Their Earlier Stages

C. Dhanusha, A. V. Senthil Kumar, and V. S. Giridhar Akula

Abstract The cognitive impairments among elderly peoples make their lifestyle more complex, and it is extremely essential to accurately discover patients with mild cognitive impairment stage as it may or may not progress to Alzheimer's disease. Many machine learning models have developed to discover the prediction of Alzheimer's disease, but they are constrained with imbalanced medical dataset with the disease case records. Thus, this paper concentrates on developing the robust model which handles the imbalanced data when there are a smaller number of instances with case label of dementia. In this work, neuro fuzzy symbolic reasoning is developed by adapting case-based reasoning to retrieve similar pattern of matched records with the new instance as input query to determine whether it suffers from dementia or not. The case-based reasoning is empowered by applying cuckoo search model, which involves searching for matching records by applying levy flight strategy to define unknown instances and their corresponding matching instance labels, and this information is given as input to the fuzzy artificial neural network along with training dataset to enhance its prediction accuracy. The FANN with the additional knowledge obtained from cuckoo search-enhanced case-based reasoning during training phase handles uncertainty and produces more accurate results for Alzheimer's prediction compared to the other traditional models of classification. During testing phase, the new cases are predicted and if the output is correct, then it is added and maintained in the case history for future references. The simulation results proved the prominence of cuckoo search-enabled symbolic reasoning-based Fuzzy Neuro Classifier by applying this strategy on ADNI dataset.

C. Dhanusha (✉)
Hindusthan College of Arts and Science, Coimbatore, India
e-mail: kunjukannan020817@gmail.com

A. V. Senthil Kumar
Department of MCA, Hindusthan College of Arts and Science, Coimbatore, India

V. S. Giridhar Akula
Princeton Institute of Engineering and Technology, Hyderabad, India

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_65

Keywords Alzheimer's disease · Case-based reasoning · Symbolic reasoning · ADNI · Cuckoo search · Fuzzy neuro classifier

1 Introduction

Alzheimer's disease is a progressive, irreversible, and chronic neurodegenerative illness which is clinically revealed by gradual loss of cognitive impairments, amnesia, anxiety, and collapsing their regular activities [1]. Alzheimer's disease pathogenesis is not completely well defined, and there is no proper therapy that can cure or stop the progression of disease completely. Mild cognitive impairment is a transitional phase among cognitively normal aging and dementia [2].

The patients with mild cognitive impairment have higher chance of AD than healthy cognitive due to their age factor. It is very essential and effective when Alzheimer is detected at its earlier stage strategies for caring them and assisting experts to diagnose during MCI so that further worsening of the disease can be avoided. A huge variety of machine learning algorithms have used for AD prediction, but existing machine learning models inherit the patterns by receiving attribute values instead of receiving information about those attributes and instances. Recent advances in the field of machine learning and artificial intelligence have extraordinary impact in the medical field [3].

Deep Neural Network have a step forward solutions then the existing classification models. Though it has its own power, it greatly depends on the huge volume of labeled data and, regrettably, it is very tough and resource intensive to annotate voluminous data [4].

In most of the disease detection dataset, the instances relevant to the positive cases are limited but for deep learning and other classification models, handling such imbalanced data is really difficult. Thus, there is a necessary for prior knowledge constructed on logic as a natural solution to make learning effectual especially when there is a scarce of training dataset.

Hence, in this research work integration of symbolic reasoning is considered as an important part for performing Alzheimer's detection by enhancing the conventional fuzzy neural network to produce more accurate results in determining the severity of it. This work integrates logical learning which gain more attention in prediction paradigms. The two important cognitive abilities are learning through experience and the aptitude to reason from what it has been learned. The robustness of case-based reasoning is achieved in this work by adapting cuckoo search for retrieving similar case of patients to induce the knowledge about gained experience training the fuzzy neural network with its efficacy.

2 Related Work

Neelaveni and Devasana [5] in their work used various machine learning models to diagnose the presence of Alzheimer. Their objective is to identify the symptoms of the Alzheimer at the earlier stages so by giving proper treatment at right time it may slow down the aggressiveness of the AD. The physical parameters used for predicting Alzheimer are age, number of visits, education, and MMSE attributes.

Pan et al. [6] constructed an extreme learning model based on grading method which prominently integrates multimodal data to predict the patients under the conversion of mild cognitive impairments to Alzheimer's conversion. Initially, from the MRI the features are extracted and, among them, potential features are selected by applying feature selection algorithm. Different modalities of MCI are included for prediction; they are positron, cerebrospinal fluid biomarkers, gene data, and emission tomography. The grading scores obtained are fed as input to the classifier to differentiate the progression of MCI to AD or it is stable MCI.

Charles et al. [7] in their work adapted an unsupervised learning model known as conditional restricted Boltzmann machine which is used to monitor and predict dozens of patient characteristics to personalize medicine for Alzheimer's disease. The 18-month courses of 1909 patients with 44 clinical information were collected as variables to predict the transition from MCI to Dementia. The determined standard deviations and correlation among variables were determined by applying logistic regression.

Park et al. [8] forecast Alzheimer's disease by adapting bootstrapping model; its output produces a one-year prediction as probable dementia or definite dementia. They used 51 attributes which comprised of age, hemoglobin level, etc.

Vindhya et al. [9] designed a convolutional neural network and the extraction of features and selection of potential features to predict the Alzheimer's disease at various stages. They used MRI images of brain to predict whether the victim is affected with moderate to demented or non-demented by applying CNN prediction model.

Grassi et al. [10] used different machine learning models to train 123 MCI subjects to predict the neurological disorder. They used clinical information, baseline information, sociodemographic, and cardiovascular risk indexed are used for perdition process. They used Leave—pair-out-cross validation may employ the validation strategy and for feature selection recursive feature elimination is used to discover more relevant predictors' subsets.

Beheshti et al. [11] developed a CAD-based model to predict the Alzheimer. It involves four different stages: In first stage, they used voxel-based measure to examine local and global gray matter among AD group with healthy groups. In second stage, voxel values are extracted from the atrophy regions for HC, stable MCI, AD, and progressive MCI. In third stage, feature extraction is done by applying t-test score with genetic algorithm to discover significant subsets. At last, SVM is used to classify the severity of Alzheimer.

3 Methodology

3.1 Deep Adaptive Clustering

According to cognitive theories, just like humans how they react based on analogical reasoning, Roger Shank created an artificial intelligence-based learning model known as case-based reasoning (CBR) [12]. CVR performs comparison of query case with the historical cases, and similar cases are verified out of the case base. Those remaining similar cases are involved in solving new cases. Further, the identified new case is retained as an incremental learning model by case base for forecasting. This is the reason for improving the model constantly to predict more accurately with continuous increase in both experience and knowledge. In general, CBR method comprised of four parts; they are case base to store past histories, indexing mechanism for searching, determining the similarity using matching algorithm, and finally adaptation mechanism. The process of CBR is attained by 4R's they are retrieve, reuse, revise, and retain [13] as shown in the Fig. 1.

- Retrieve: For a query instance, retrieve relevant cases from memory to predict its type
- Reuse: For a given query instance, map the corresponding previous case to solve the problem
- Revise: Once mapped the previous instance to the query instance, test the new instance presented in the real world and if needed revise it.

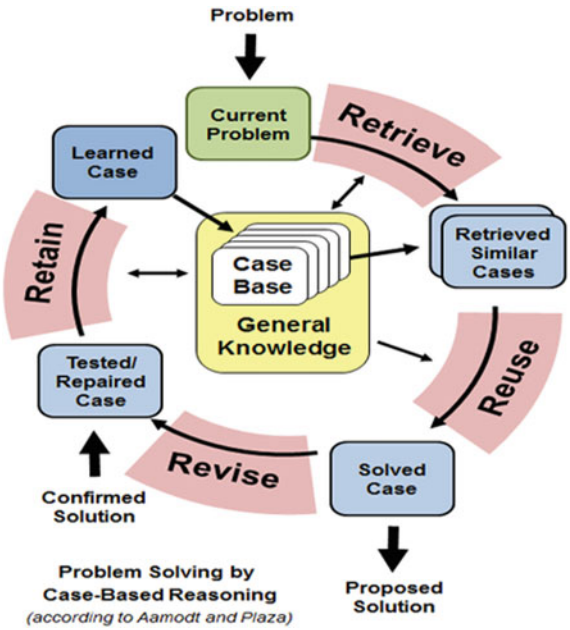


Fig. 1 Complete process of case-based reasoning

- Retain: Once the results have been adapted by the query instance, store the obtained experience as a new case in memory.

While using CBR, it does not need extra analysis in a particular area, as it offers a solution for a stated problem without needing additional information form the expertise area as even when it is incomplete. An essential note is that it compares the similarity among two cases. It offers a comprehensive explanation of its functionality to produce best output.

3.2 Artificial Neural Networks (ANN)

Neural networks have the ability to build a model based on mathematical structures which is inspired by the human nervous model learning process. This model produces an outstanding capability to produce results for complex to vague dataset. ANN approach performs pattern extraction and discovers the patterns which are more difficult and to be observed by either other techniques or human [14]. A complete trained network can be discovered as the professional classification of the data which has been given of process analysis. This produces possible results during testing phase with labels not known. The usage of neural network is expanded to technical areas as it generates the optimal outputs in new environments. They are chiefly fine suited for forecasting the consequences in an accurate way. The construction of a neural network is shown in Fig. 2.

ANN in its simplified structure comprised three different layers, the first layer which receives the input from the outside world is known as input layer which is followed by hidden layer and finally the output layer. Each layer contains one or more neurons or nodes, and the nodes in input layer receives the input values and forward it to the hidden layers. In the hidden layer, neurons function as computation units and the input values received from the previous layer are computed with the weights assigned

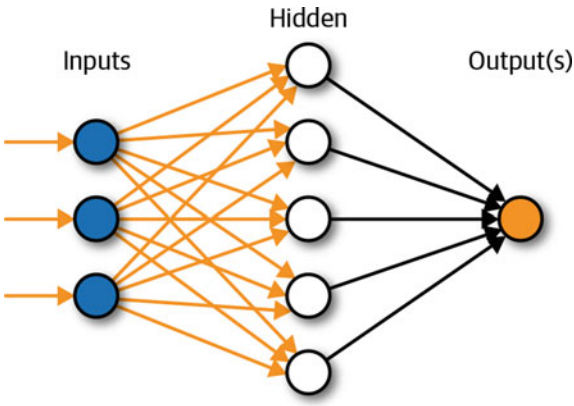


Fig. 2 General view of artificial neural network [ANN]

to the corresponding connected links. The complexity of the network increases when there are a greater number of hidden layers or neurons involved in the model. The activation function is applied with a threshold to discover whether information is passed along.

The hidden nodes are weighted by weights that are determined by their values. The weight values obtained from all previous layer nodes are combined as shown by the symbol $\sum$ in Fig. 3 to produce a new value. The summation value that governs the output node is added to a non-linear function known as sigmoid in the S-shaped curve. This sigmoid operation's input varied from $-\infty$ to $+\infty$ range, while the output ranged from 0 to 1.

Like the hidden layer, the output layer's active node is also merged, modifying the data to be generated as output. Besides, the number of layers and nodes on each layer varies depending on the input and output. The output layer took the form of a single node to carry out the target's prediction. The output value is a positive sign of the presence and a negative sign of absence that depends heavily on the input data.

3.3 Cuckoo Search Algorithm

An optimized algorithm which is developed based on the inspiration of obligate brood parasitism is known as cuckoo search [15]. Cuckoo birds lay their eggs in other host bird nests. Once the host bird identifies the alien eggs, it will either lay those eggs or abandon its nest and construct a new nest in some other place. Each egg in a nest denotes a solution, and a new solution is denoted as cuckoo egg. The objective of this algorithm is to utilize the new and possibly improved results to substitute an average result in the nest. It lays single egg at a time, and it is placed in an arbitrarily selected nest. The nest with high quality of eggs is considered as best, and it will be carried to next generation. Number of hosts birds' nests is fixed, and it discovers the presence of cuckoo laid egg with the probability of $\text{prob}_a \in (0,1)$. The algorithm below describes the working principle of the cuckoo search algorithm.

Procedure for Cuckoo Search Algorithm

```
Begin
Assign n number of host nests as initial population
While (itr < max_gen) or (termination is met)
        Get cuckoo in a random manner and replace its egg by performing levy
          flights
        Compute its fitness Fiti
        Select a random nest (j) among n
        If (Fiti > Fitj) then
                Substitute j by new solution
        End if
        A portion popa of the worst nests are discarded and new ones are
          constructed
```

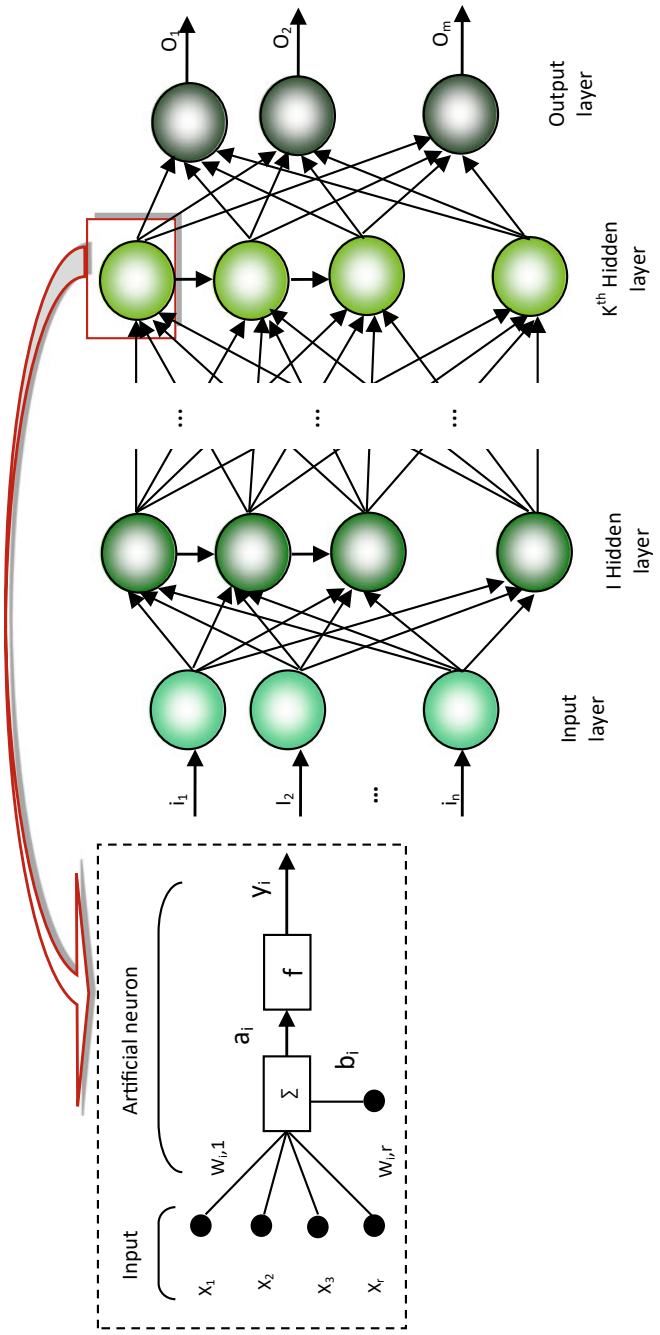


Fig. 3 General view of artificial neural network [ANN]

Keep the best nests/solutions
Rank the nests/solutions and recognize the present best
Permit the present best solutions to the subsequent generations
End while

3.4 Cuckoo Search-Enhanced Fuzzy Neuro Symbolic Reasoning for Alzheimer's Prediction

In this paper, three different aspects of investigation are carried to construct a powerful Alzheimer's prediction model. ADNI dataset [16] is used which has 1534 instances of elderly persons, and the attributes used are rid, ptid, viscode, exam date, age, ptgender, pteducat, apoe4, mmse, apoe genotype, and class. The identifiers are removed during preprocessing, and only the remaining attributes are used for Alzheimer's prediction. The dataset is normalized to convert the attributes range of value to fall under same range. Initially, the dataset is split into training and testing dataset, 80% of the dataset is used for training process and remaining 20% is used for testing process. The training dataset is considered as a case history and each instance of the testing dataset is used as new case or query instance, similar records which is available in the case history is searched by applying the cuckoo search model, instead of conventional searching models like k nearest algorithm (Fig. 4).

This overcomes the local optima and matching information's are given as the input to the fuzzy artificial neural network during the training process of the prediction model. The fuzzification of the input instances handles the uncertainty, and the case-based reasoning is used as symbolic reasoning which improve the performance of the FANN to produce optimized Alzheimer's prediction at their earlier stages.

The flowchart depicted in Fig. 5 the detailed process of cuckoo search-enhanced fuzzy neuro symbolic reasoning (CS-FNSR) for predicting the Alzheimer's detection. In this symbolic reasoning, adapted was case-based reasoning, and ADNI dataset [16] which is used by fuzzy ANN is also used by cuckoo search-enabled case-based reasoning to retrieve the similar pattern of records. In this model for an input instance, its similar pattern of records is searched by applying the cuckoo search model, where the input is considered as cuckoo egg and it searches the relevant instance to determine then whether they are affected by Alzheimer or not.

Cuckoo search algorithm for optimized similarity searching, the levy flight technique is applied for the finding the fittest records in the previous history of patients. Then, the host birds are used to discover the input query with the list of matched records and the observed output is passed to the fuzzy ANN to understand the instances belong to Alzheimer or normal. This step is particularly used for the process of mapping with previous patients, and if the expected output and observed output of the FANN produce more accuracy, then the new solution is added to the case history database. During the testing phase, the input instance is treated with same sort of similarity searching done by applying levy flight model and the results are revised accordingly.

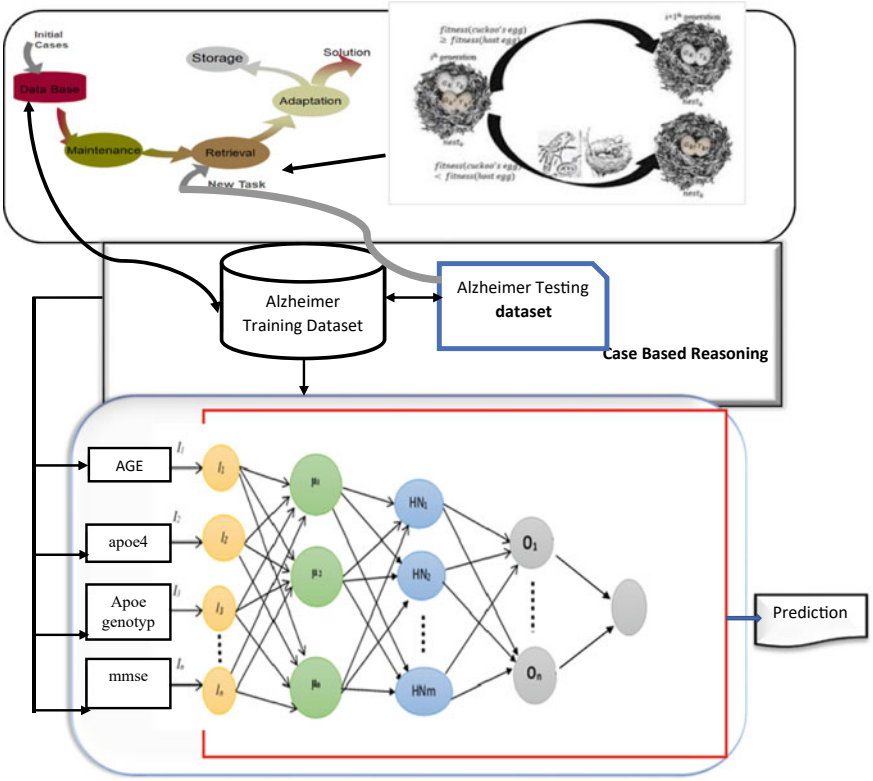


Fig. 4 Architecture of cuckoo search-enhanced fuzzy neuro symbolic reasoning for Alzheimer's prediction

In Fuzzy ANN, the given input is converted to the fuzzy representation and then the hidden layers perform the process of computation of input values to be multiplied with the weights assigned to the links between layers [17]. Naturally, the input nodes are passive. Its purpose is to accept and transfer the data from the outside world to the next level. The hidden layer must multiply the obtained value by adding the equivalent weight to each node.

$$\text{ANNk}_j = \sum_i w_{ij} \cdot O_i \tag{1}$$

where wt_{ij} represents the weights among node_i and node_j, and O_i is the output from node_i calculated as in Eq. (2):

$$O_j = (\varphi\ \text{ANN}_j) \tag{2}$$

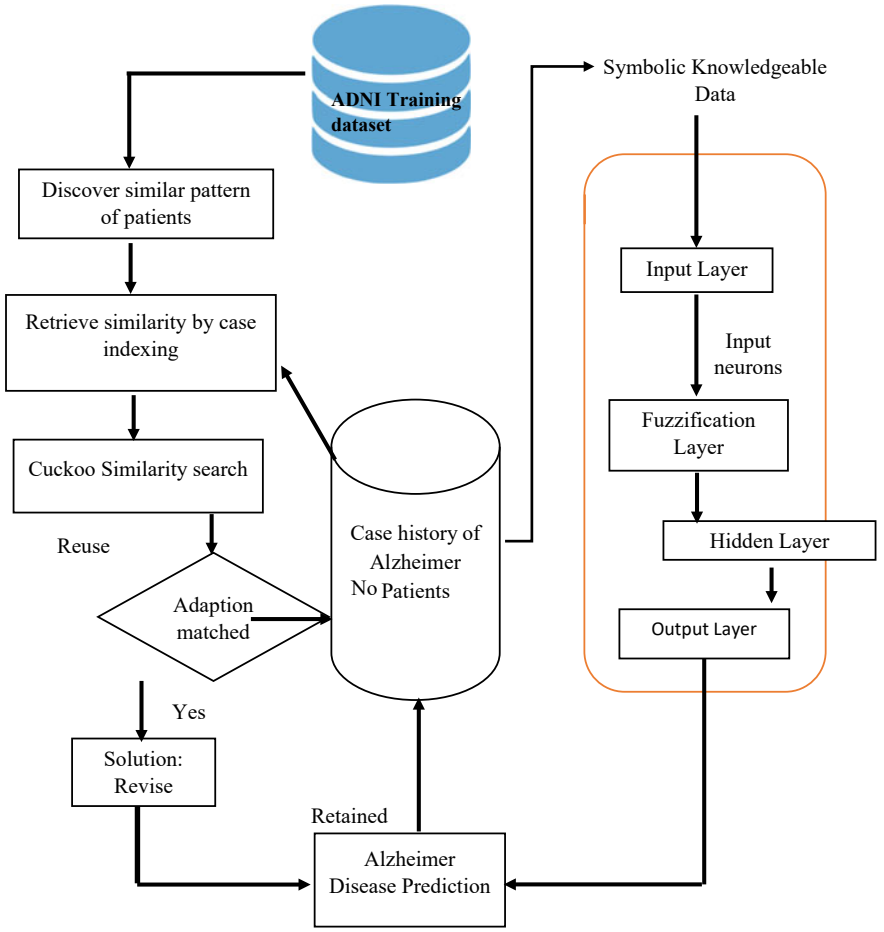


Fig. 5 Workflow of cuckoo search symbolic reasoning with fuzzy neural classifier

where φ refers to activation function and the sigmoid S-shaped curve applied to the weighted sum of input values before passing it to the subsequent layer.

The derivative of the sigmoid function is shown in Eq. (3).

$$\varphi'(\mathrm{ANN}_j) = \varphi(\mathrm{ANN}_j)(1 - \varphi(\mathrm{ANN}_j)) \tag{3}$$

Here, tan-sigmoid transfer function is used for both the hidden layer and the output layer. They compute the output of each layer form network input. Finally, the generated output is compared with the expected output and their difference is assigned as an error. The standard neural network uses back propagation method to update the weights in trial and error method repeatedly until the accepted error rate is obtained. This results in high computation time and complexity in determining

the appropriate weights to produce more accurate results. The error rate is defined in Eq. (4).

$$\text{Err} = \frac{1}{2}\sum_{k}(\text{dsrot}_k - \text{rsltot}_k) \tag{4}$$

The final process of cuckoo search-enhanced case-based reasoning is that presence of Alzheimer is diagnosed and its solution is compared with the results produced by fuzzy neural network. The prediction is made efficient, and the newly discovered target patients is stored as a new case for forecasting process.

4 Results and Discussions

This section analyzes the performance of the newly constructed symbolic reasoning-based fuzzy neural network for predicting Alzheimer at their earlier stages. The CS-SRFANN is implemented using MATLAB software and to predict Alzheimer ADNI dataset [16] which has 1534 instances of elderly persons and the attributes used are rid, ptid, viscode, examdate, age, ptgender, pteducat, apoe4, mmse, apoe genotype, and class. The identifiers are removed during preprocessing, and only the remaining attributes are used for Alzheimer's prediction process. The performance of the proposed CS-SRFANN is compared with standard ANN and fuzzy ANN. The metrics used for analyzing the performance are accuracy, precision, recall, and RMSE.

4.1 Comparative Analysis Based on Accuracy

Accuracy is used to measure the correctly predicted instance of type of Alzheimer and the absence of Alzheimer among the total patients in ADNI database.

$$\text{Accuracy} = \frac{\text{Tot. No. of Correctly preidcted presence and absence of Alzheimer}}{\text{Total No. of Patients in ADNI Dataset}} \tag{5}$$

Prediction of Alzheimer at earlier stages is very essential to maintain qualified life of elderly persons, and providing more accurate results will assist the experts to diagnose its severity. Thus, Fig. 6 illustrates the performance of symbolic reasoning-based fuzzy artificial neural network produces highest accuracy rate because instead of receiving raw input values, it acquires knowledge from symbolic reasoning and expertise itself with fuzzy to handle uncertainty. In fuzzy ANN, it has the ability to represent uncertainty but with its input data it classifies the severity of autism so that it performance is less than the proposed SRFANN.

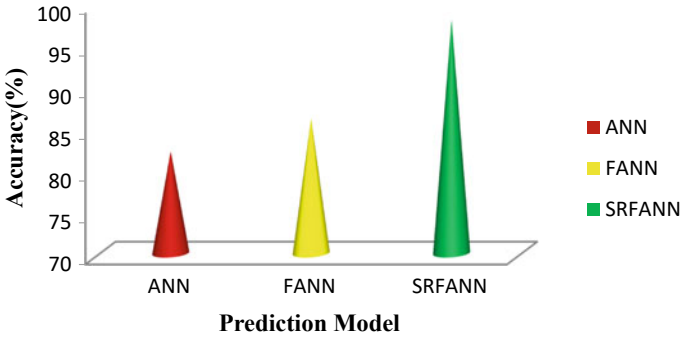


Fig. 6 Alzheimer's prediction analysis based on accuracy

4.2 Comparative Analysis Based on Precision

The measure of correctly predicted Alzheimer patients with their risk level is known as precision, and it is calculated by dividing the correctly predicted positive cases with total number of patients assumed to be Alzheimer cases.

$$\text{Precision} = \frac{\text{Correctely predicited Alzheimer Patients}}{\text{Total No. of Alzhiemer Patients detected}} \tag{6}$$

The Fig. 7 expose percentage of correctly predicted Alzheimer positive cases achieved by ANN, FANN, and SRFANN. The precision rate reveals the efficacy of proposed SRFANN as it utilizes the concept of case-based reasoning which exhibits matching and retrieval of similar instance which is very useful to enhance the training session, and thus, it produces more promising results of predicting the presence or absence of Alzheimer at their earlier stages in an optimized way. The ANN and

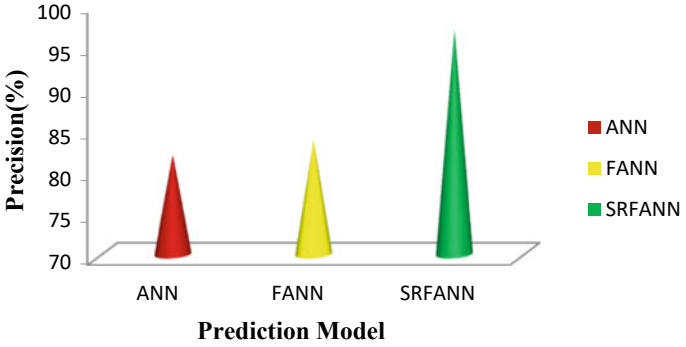


Fig. 7 Alzheimer's prediction analysis based on precision

FANN has the limited instance of Alzheimer instances, so they suffer to understand in depth about the importance of input patterns compared to SRFANN.

4.3 Comparative Analysis Based on Recall

Recall metric is used to determine the efficacy of predicting accurately the presence of Alzheimer among the actual Alzheimer patients present in ADNI database.

$$\text{Recall} = \frac{\text{Tot. No. of Correctly predicited Alzheimer Cases}}{\text{Total No. of actual Alzhiemer Cases in ADNI dataset}} \tag{7}$$

The output of recall value obtained by three different models to predict the presence of Alzheimer is depicted in the Fig. 8. This measure shows the relevancy of each prediction models to predict the Alzheimer at their earlier stages. The symbolic reasoning analyzes the input data and performs searching of similar records in the past history by applying matching and retrieval. This work used case-based reasoning as symbolic analysis, in which most similar cases are retrieved by applying nearest neighbor retrieval algorithm. The output of the case-based reasoning is provided as input to the fuzzy neural network, to improve its accuracy in detection of Alzheimer when there is uncertainty in distinguish it from other cognitive impairments. Thus, compared to SRFANN, other two prediction models standard ANN and fuzzy ANN produce less precision value in Alzheimer's prediction.

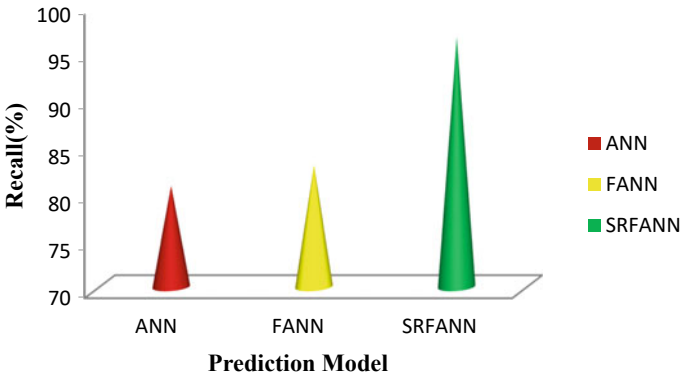


Fig. 8 Alzheimer prediction analysis based on recall

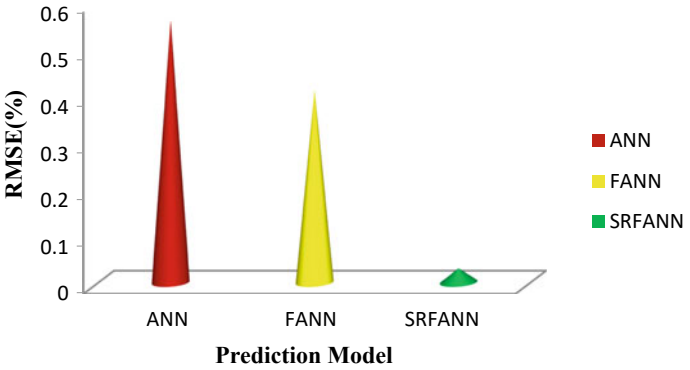


Fig. 9 Alzheimer's prediction analysis based on error rate

4.4 Comparative Analysis Based on Error Rate

The variance among corresponding observed values in ADNI dataset and predicted values of models are squared, and they are averaged over the sample

$$\mathrm{RMSE}_{fo} = \left[\sum_{i=1}^{N} \left(Z_{fi} - Z_{oi}\right)^2 / N\right]^{1/2} \tag{8}$$

The results of RMSE of three prediction models are depicted in the Fig. 9. The error rate of SRFANN is very less compared to ANN and FANN, and this is because integrating symbolic and fuzzy neural approaches offers prediction model with effective learning and capability for precise prediction, especially, while using case-based reasoning as symbolic reasoning provides more effective results than other symbolic systems. The proposed SRFANN produce optimized results when the prior knowledge of solving similar patterns is available. Fuzzy neural network retains itself by utilizing closely matching Alzheimer's cases selected by the case-based reasoning retrieval approach; thus, it produces less error rate as it handles both uncertainty and improves the accuracy rate.

5 Conclusion

This work introduced the logical reasoning by inducing cuckoo search-enabled case-based reasoning as symbolic reasoning. The proposed method cuckoo search-enhanced fuzzy neuro symbolic reasoning (CS-FNSR) intends to improve the fuzzy neuro classification model more precisely by retrieving the similar pattern of existing case records and searching its matching records by applying the nature-inspired meta-heuristic model known as cuckoo search, like it determining the cuckoo egg form

the Host nest, for the input data its concern matching record is searched in the case history which utilize the ADNI dataset. The searching is done by levy flight strategy instead of searching the entire database in a sequence order, and this will optimize the searching process more precisely, quickly, and accurately. With the obtained knowledge about the input instance class label, further it is utilized during training phase and the concern instance is fed as input to the fuzzy neural classifier. The fuzzy neuro classifier with the empowered knowledge performs the prediction process more accurately and breakthrough the performance of other classification models by handling the imbalanced dataset and uncertainty.

References

1. Ulep MG, Saraon SK, McLea S (2018) Alzheimer disease. J Nurse Practit 14:129–135
2. Liu S, Liu S, Cai W, Pujol S, Kikinis R, Feng D (2014) Early diagnosis of Alzheimer's disease with deep learning. In: Proceedings of the IEEE ınternational symposium on biomedical ımaging, New York, NY. IEEE
3. So A, Hooshyar D, Park KW, Lim HS (2017) Early diagnosis of dementia from clinical data by machine learning techniques. Appl Sci 7:651
4. Ji H, Liu Z, Yan WQ, Klette R (2019) Early diagnosis of alzheimer's disease using deep learning. In: Proceedings of the 2nd ınternational conference on control and computer vision, pp 87–91. Association for Computing Machinery, New York, NY
5. Neelaveni J, Devasana MSG (2020) Alzheimer disease prediction using machine learning algorithms. In: 2020 6th ınternational conference on advanced computing and communication systems (ICACCS), Coimbatore, India. pp 101–104
6. Pan D, Zeng A, Jia L, Huang Y, Frizzell T, Song X (2020) Early detection of alzheimer's disease using magnetic resonance ımaging: a novel approach combining convolutional neural networks and ensemble learning. Front Neurosci 14(259)
7. Fisher CK, Smith AM, Walsh JR (2019) Machine learning for comprehensive forecasting of Alzheimer's disease progression. Sci Rep 9:13622
8. Park JH, Cho HE, Kim JH, Wall M, Stern Y, Hyunsun, Yoo S, Kim HS, Cha J (2020) Machine learning prediction of ıncidence of alzheimer's disease using large-scale administrative health data machine learning prediction of incidence of alzheimer's disease using large-scale administrative health data. Digit Med 3: 46
9. Vindhya GB, Alam M, Mansi M, Kedia M, Anjal D (2020) Prediction of alzheimer's disease using machine learning technique. Int Res J Eng Technol (IRJET) 07(05)
10. Grassi M, Perna G, Caldirola D, Schruers K, Duara R, Loewenstein DA (2018) A clinically-translatable machine learning algorithm for the prediction of alzheimer's disease conversion in individuals with mild and premild cognitive impairment. J Alzheimer's Dis 61(4):1555–1573
11. Beheshti I, Demirel H, Matsuda H (2017) Alzheimer's disease neuroimaging ınitiative. Classification of alzheimer's disease and prediction of mild cognitive impairment-to-alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. Comput Biol Med 83:109–119
12. Guishi GYD (2004) Review of case-based reasoning. Comput Eng Appl 40(21):1–4
13. Pal S, Simon C (2003) Foundations of soft case-based reasoning. A Wiley-Interscience publication
14. Towell G, Shavlik J (1994) Knowledge-based artificial neural networks. Artif Intell 70(1–2):119–165
15. Yang XS, Deb S (2009) Cuckoo search via Lévy flights. World Congress on Nature and biologically ınspired computing (NaBIC 2009). IEEE Publications, pp 210–214

16. http://adni.loni.usc.edu/data-samples/data-types/clinical-data/
17. Klawonn F, Kruse R, Nauck D, Borgelt C (2003) Neuro-fuzzy-system. Vieweg, Wiesbaden
18. Dhanusha C, Senthil Kumar AV (2019) Intelligent ıntuitionistic fuzzy with elephant swarm behaviour based rule pruning for early detection of alzheimer in heterogeneous multidomain datasets. Int J Recent Technol Eng (IJRTE) 8(4):9291–9298. ISSN: 2277–3878
19. Dhanusha C, Senthil Kumar AV (2020) Enriched neutrosophic clustering with knowledge of chaotic crow search algorithm for alzheimer detection in diverse multidomain environment. Int J Sci Technol Res (IJSTR). 9(4):474–481, ISSN:2277–8616, Scopus Indexed
20. Dhanusha C, Senthil Kumar AV, Dr. Musirin IB (2020) Boosted model of LSTM-RNN for alzheimer disease prediction at their early stages. Int J Adv Sci Technol 29(3):14097–14108
21. Dhanusha C, Senthil Kumar AV (2021) Deep recurrent Q reinforcement learning model to predict the alzheimer disease using smart home sensor data. In: International Conference on computer vision, high performance computing, smart devices and network, IOP conference series: materials science and engineering, vol 1074, International conference on computer vision, high performance computing, smart devices and networks (CHSN 2020) 28th–29th Dec, Kakinada, India. IOP conference on serial: material science enginerring, pp 1074–012014
22. Dhanusha C, Senthil Kumar AV, Musirin IB, Abdullah HMA. Chaotic chicken swarm optimization-based deep adaptive clustering for alzheimer disease detection. In: Proceedings of the ICPCSN 2021 held in Salem, India, pp 709–719. Pervasive computing and social networking, lecture notes in networks and systems book series (LNNS, vol 317)
23. Dhanusha C, Senthil Kumar AV, Jagadamba G, Musirin IB (2021) Evolving chaotic shuffled frog leaping memetic metaheuristic model-based feature subset selection for alzheimer's disease detection. In: Proceedings of the ICSCN 2021 held in Shree Venkateshwara Hi-Tech Engineering College, Erode, India, pp 679–692. Sustainable communication, networks and applications, lecture notes in data engineering and communication technologies book series (LNDECT, vol 83)
24. Dhanusha C, Senthil Kumar AV, Villanueva L (2021) Enhanced contrast pattern based classifier for handling class ımbalance in heterogeneous multidomain datasets of alzheimerdisease detection. In: Proceedings of the ınternational conference on advances and applications of artificial ıntelligence and machine learning (ICAAAIML-2021)
25. Dhanusha C, Senthil Kumar AV (2021) Enhanced deep hierarchical classification model for smart home-based alzheimer disease detection. In: Proceedings of the ınternational conference on computational vision and bio ınspired computing (ICCVBIC 2021)

Hybrid Approach for Detecting the Traffic Violations Based on Deep Learning Using the Real-Time Data

Priya Gupta, R. Rajkumar, S. Santhanalakshmi, and J. Amudha

Abstract Nowadays, the number of vehicles on the road is steadily increasing, and this raises the risk of road accidents. One of the reasons of this regrettable situation is violation of traffic rules. Overloaded cars are one example of a traffic rule violation. This research study proposes a hybrid approach for detecting different types of violations and violators. The proposed model will detect the overloaded vehicles and the number plate of the offenders by using the deep neural network model called YOLOv4. Additionally, Tesseract is used to recognize characters in the LSTM number plate. Finally, the proposed model has delivered a satisfactory result.

Keywords YOLOv4 · Traffic violations · LSTM OCR engine (long short-term memory optical character reader) · Overloaded vehicles detection

1 Introduction

With the massive industrialization, development of new roads, increased use of automobiles, and exponential growth of the transport and tourist industries, road safety has become an issue of critical importance all over the world. This growth is much more noticeable in developing countries like India, resulting in an increase in road accidents and loss of many lives or injuries that make individuals physically unable

P. Gupta · S. Santhanalakshmi (✉) · J. Amudha
Department of Computer Science and Engineering, Amrita School of Engineering, Bengaluru, India
e-mail: s_lakshmi@blr.amrita.edu

J. Amudha
e-mail: j_amudha@blr.amrita.edu

S. Santhanalakshmi · J. Amudha
Amrita Vishwa Vidyapeetham, Bengaluru, India

R. Rajkumar
Department of Computer Science and Engineering, IIT-M, Chennai, India
e-mail: ERRajkumar@icsrpis.iitm.ac.in

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_66

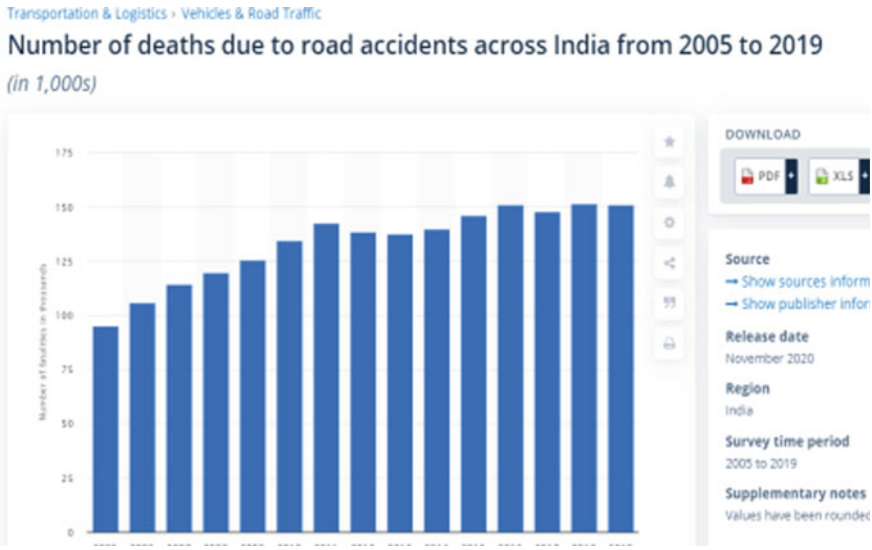


Fig. 1 Graph representing the death due to road accident in India from 2005 to 2019

for the rest of their lives. According to statistics, in 2019 alone, due to road accidents, 151 thousand fatalities have recorded [1]. One of the major causes of such unfortunate incidents is traffic rules violations like signal jump, helmetless driving, etc.

Road safety is something that is very important and has to be followed all the time. The capital city of Delhi reported the highest number of deaths across the major cities, at approximately 1.4 thousand in 2018. Over 65% of the accidents occurred on straight roads [1].

Figure 1 shows the death happens due to road accidents between 2005 and 2019 [1].

The recent technological advancements like machine learning and deep learning have shown remarkable results in many computer vision-related tasks, and in some cases, it performs better than humans [2]. The main reason behind the superior performance of deep learning models is the ability of the models to automatically learn discriminative features via model training. This paper specifically discusses about the object detection task by using the object detection algorithms, which are developed based on deep learning. There are many object detection algorithms based on deep learning, they are Faster R-CNN, single shot detectors (SSD), mask R-CNN, YOLO versions, etc. Every algorithm has its own limitation either in terms of speed or accuracy, but the YOLO algorithm overcomes the limitations of other algorithms in terms of both speed and accuracy and becomes the best fit for the real-time detection.

This paper has presented a hybrid model for the detection of one of the types of traffic rules violation, i.e. vehicles carrying load more than the limit, which also lead to road accident-like incidents by using the deep learning-based approach. The

output of deep learning model will be given as input to another deep learning model for obtaining the final result.

2 Related Work

Many research works have been done in this field for automating the system in real time by taking video or images as input to process further and generate results. This section provides a detailed analysis of previous work in the subject of road safety; however, to the best of our knowledge, no similar study has investigated the deep learning-based object recognition technique relevant to our work based on Indian road traffic for overloaded vehicles.

In [3], YOLOv3 has been used for detecting the traffic rules violation with an accuracy of 97.67% for vehicle count detection and an accuracy of 89.24% to detect the vehicle. In [4–6], authors have detected helmet (object) using YOLO versions, and then, triple bike riders are detected [4] using vertical angel. For performing background removal and segmentation in bike rider recognition (object), binary classifiers and visual characteristics are used in [7]. Naren Babu [8] YOLO versions and OCR recognition are used for number plate detection [9, 10]. LSTM Tesseract accuracy of about 95% has been achieved in number plate identification process by utilizing Faster R-CNN and recognition Tesseract OCR [11]. The moving vehicle recognition and flow on crowded highways have analysed by employing YOLO versions with an average accuracy of 92.11%. In [12, 13], a comparative research was conducted among several vehicle detection methods or architecture on different real-time datasets for recognizing objects such as cars, and it was discovered that YOLO performs better in terms of detection and speed, whereas SSD has greater accuracy but is slower. In [14], a model has been proposed to detect text from an images based on CNN and LSTM architecture, but model work has been developed for English word only. In [15], brightness awareness mechanism has been developed under different illumination conditions in order to detect pedestrians (object here) on roads in day and night by using Faster R-CNN object detection model, which is trained over both the thermal images and colour images with an map 81.27%. In [16], YOLO model has been used for the detection of license plate and perform layout classification with heuristic rule in order to achieve a better accuracy. In [17], YOLOv3 and YOLOv4 were used for detecting multiple objects in image and video to obtain 98% accuracy for images and 99% accuracy for videos. In [2], vehicles are detected by using the Faster R-CNN over the heterogeneous traffic scenario data in order to obtain state-of-the art vehicles detection performance at 5 ps over TitanX GPU. In [18], YOLO models were used for the detection of triple riding and speed violation of two wheeler vehicles.

Most of the works done in this area include traffic monitoring or tracking, and only fewer works are done in the area of cause of road accidents, which sometimes lead to fatality.

3 Methodology

The proposed approach is based on the hybrid concept, which combines two models, where each model is divided into two to three phases.

Datasets

This research study has generated own datasets for both models and labelled them manually since there is no publicly available dataset dedicated to this purpose. In order to expand the dataset size, additionally some augmentation has been performed. The dataset contains almost all types of Indian vehicles from two-wheelers to four-wheelers. There are 109 images available in the number plate dataset, which consists of almost all types of Indian vehicles (auto,bus,scooter,bikes, etc.). For the overloaded dataset, we have used 286 images, and further, the overloaded dataset is divided into two classes such as overloaded bike and overloaded vehicles.

Overloaded vehicles				
Types of overloaded vehicles	Overloaded vehicles		Overloaded bikes	
	Persons	Luggage/Goods	Persons	Luggage/Goods

Sample images of overloaded and number plate datasets are shown in Figs. 3 and 2, respectively.

YOLOv4 is a one stage object detection model with a size of approximately 290 MB. These architectures consist of three layers: (a) Backbone—extract the essential features, (b) Neck—collect the features maps from different stages, and (c) Dense

Fig. 2 Sample images of ANPR dataset

Fig. 3 Sample images of overloaded dataset

prediction—perform final predictions, which is a vector containing the coordinates of the predicted bounding box (centre, height, and width), the confidence score of the prediction, and the label as shown in Fig. 4 [19].

Detection Model for overloaded vehicles

This section presents the proposed model for overloaded vehicles.

The pipeline of the overloaded vehicles is shown in Fig. 5.

The input images of varied sizes resized to 416X416.

The model is trained and tested over two classes such as overloaded vehicles and overloaded bikes. For training, transfer learning model is used, and further, the pretrained weights were used to provide custom training, since the dataset is not very large. Generally, YOLOv4 works on the principle of regressing bounding box around the object having four dimensions (x, y, w, and h), confidence scores, in output, the model is trained in darknet framework, and later, it gets converted into OpenCV.

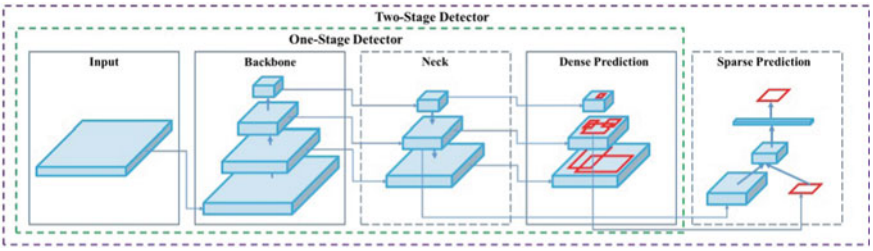


Fig. 4 YOLO architecture

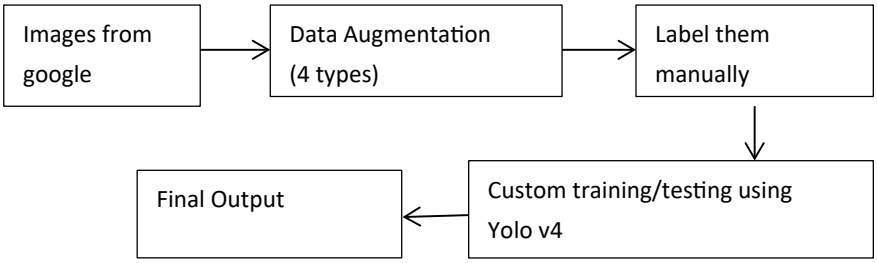


Fig. 5 Pipeline for overloaded vehicles detection model

Number Plate Detection and Recognition Model

This section presents number plate detection and character recognition.

The pipeline number plate detection and recognition is shown in Fig. 6.

It has 3 modules for performing plate detection from the overloaded vehicles.

(a) Number plate detection
(b) Extracting the plate
(c) Segmentation and recognition.

Image processing is an important step in recognition, lack of proper preprocessing may lead to inaccurate or poor accuracy results. The main reason behind this processing is to enhance the image quality and readability of the characters present in the plate.

This is the last phase of ANPR. The processed image is passed through the OCR engine for recognition of characters. For recognition, the characters of the number plate are segmented. For this, we used Python wrapper of Google Tesseract, which is an open-source OCR engine called Pytesseract. There are different segmentation modes called page segmentation mode (PSM), that performs segmentation based on the angles, orientation, and styles. Thirteen PSM modes are shown in Table 1. Four different modes of operation are selected by using the OEM option (OEM stands for OCR engine mode) [9], where LSTM engine mode is selected.

The final model is an integration of two different models.

The pipeline of the final model is shown in Fig. 7.

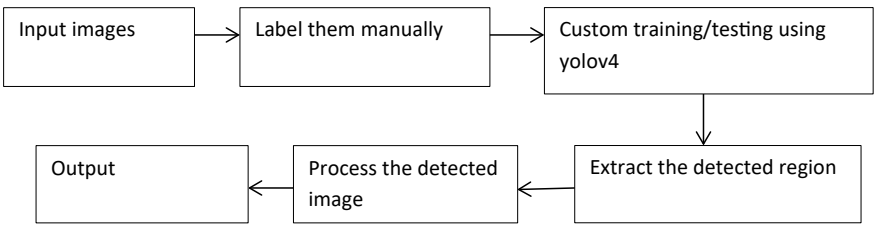


Fig. 6 Number plate detection pipeline

Table 1 Page segmentation mode

PSM	Description
0	Orientation and script detection (OSD)
1	Automatic page segmentation (APS) with OSD
2	APS excluding OSD and OCR
3	Fully APS excluding OSD and OCR
4	Assuming a single column of text of different sizes
5	Assuming a single uniform block of vertically aligned text
6	Assuming a single uniform block of text
7	Picture is treated as single text line
8	Picture is treated as single word
9	Picture is treated as single word in circle
10	Picture is treated as single character
11	Picture is treated as Sparsh Text, and text is found in a specific order as much as it can
12	Picture is treated as Parse Text with OSD
13	Raw line, picture is treated as a single text line, and Tesseract-specific hacks are bypassed

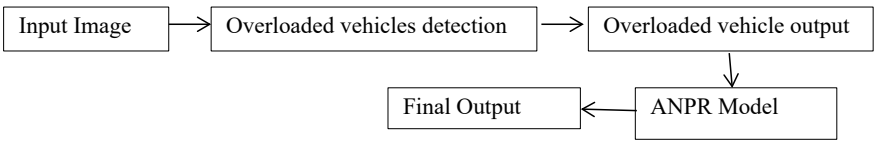


Fig. 7 Integration pipeline

In the integrated model, the output of first model that is overloaded vehicles will be fed as input to the second model for performing number plate recognition.

4 Experimental Result and Analysis

This section covers the experimental set-up and result analysis. For training, the Google colab GPU (Tesla V4) Cuda version 11,000 are used. Before training, it is required to create train.txt file, test.txt file, names file, and data files, wherein these custom configuration files contain the necessary configurations and model parameters for training.

The parameters used while training the detection model have been listed in Table 2 for both overloaded vehicles detection model and number plate detection model.

The performance metrics of the models are shown individually in Table 3.

The results of the overloaded vehicles and number plate detection are shown in Figs. 8 and 9, respectively.

Table 2 Parameters configurations used while training overloaded dataset and number plate dataset

Parameters configurations		
Parameters	Overloaded model	Number plate model
Batch	64	64
Subdivision	16	16
Width	608	608
Height	608	608
Channels	3	3
Learning rates	0.0013	0.0013
Max batches	4000	2000
Steps	3200,3600	1600,1800

Table 3 Result metrics

Model		Best_Weights (%)
Overloaded model	Map50	100
	Map75	97.38
Number plate model	Map50	97.15
	Map75	88.15

The final result with the intermediate result and original image of model is shown in Figs. 10, 11, 12, and 13, respectively.

The proposed approach gave the satisfactory result. In the Figs. 14 and 15 last training weights are shown. Since the size of the model is 299MB, the training process is cumbersome and we overwrite the existing weight graphs.

Since the model is trained with a small dataset and with limited resources, it also has some limitations as shown in Fig. 16. The models performance degrades in case of noise, illumination, image quality and in case of occluding objects.

5 Conclusion and Future Scope

This research work has proposed an approach that detects violation by detecting the overloaded vehicles and its number plate. In this, the detection has done by using the deep learning-based object detection algorithm YOLOv4. For recognition, LSTM OCR engine was used. The proposed approach has delivered a satisfactory result but also has some limitations with some false nuisance too. Since the dataset was not very large which somewhere affects the model performance and accuracy rate.

Future enhancement can be in the direction of improving the model performance by training the model over a large dataset and diverse dataset with a good GPU and storage system. In first phase, YOLOv4 is used, but there are other algorithms and OCR engines available for training and testing. Instead of using Tesseract try out

Fig. 8 Sample output overloaded vehicles

Fig. 9 Sample output number plate detection

Fig. 10 Original image

Fig. 11 Intermediate result 1 model

Fig. 12 Intermediate result 2 model

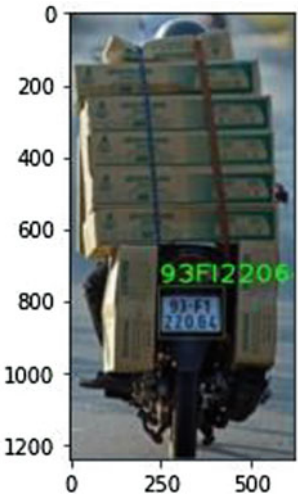


Fig. 13 Final result (after integration)

with TopOCR, which is the fastest engine. Another functionality that can be added to this is a messaging system for sending challans to such violators by using the RTO database, storing vehicle owner details with respect to vehicles number.

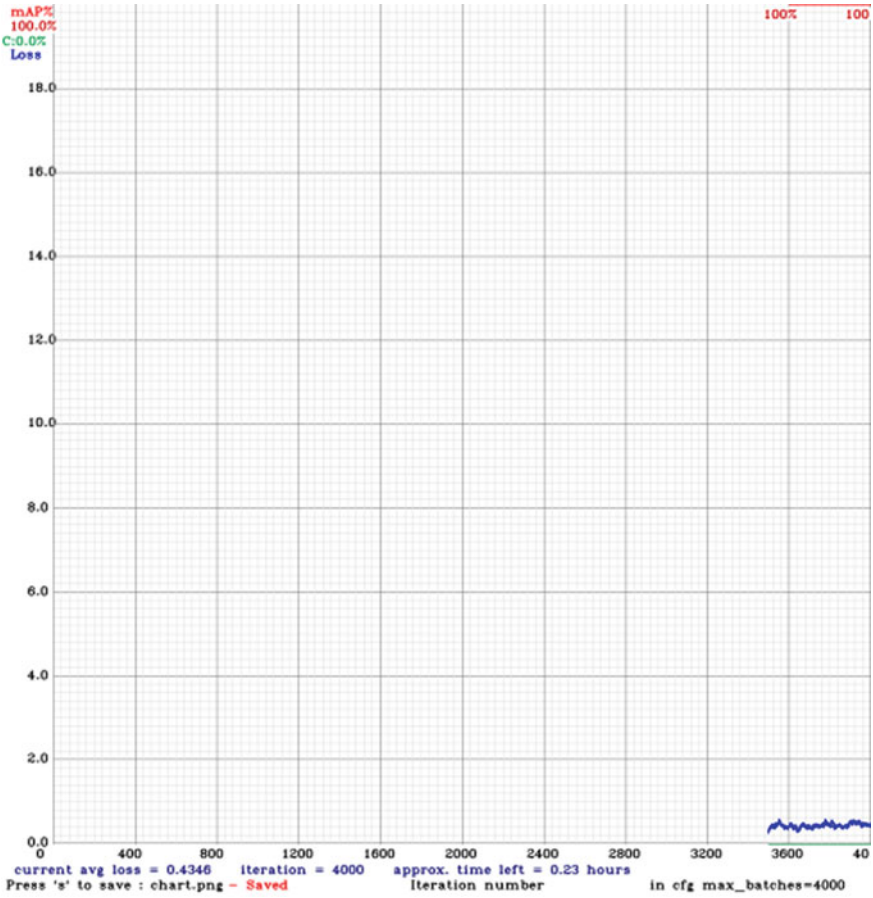


Fig. 14 mAP graph of last weights of overloaded vehicles

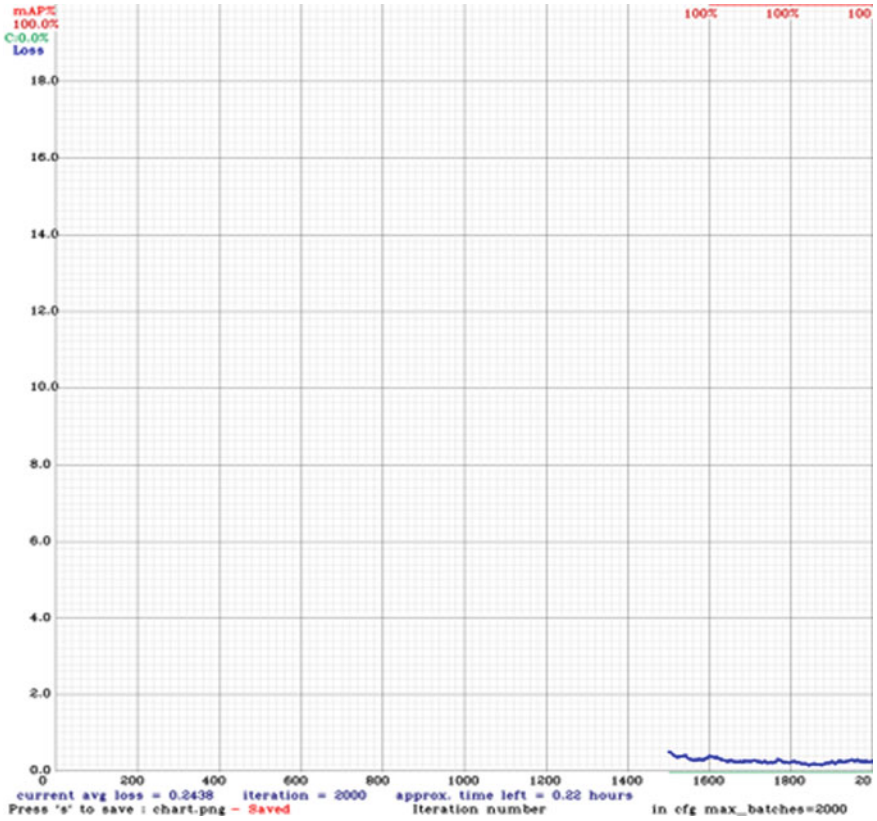


Fig. 15 mAP graph of last weights of number plate model

Fig. 16 **a** Model gets affected by the high noise, illumination changes, **b** OCCLUDING object and data annotation quality

References

1. https://www.statista.com/statistics/746887/india-number-of-fatalities-in-road-accidents/
2. Mittal D, Reddy A, Ramadurai G, Mitra K, Balaramanravindram (2018) Training deep learning architecture for vehicles detection using heterogeneous traffic data. In: IEEE 2018 10th international conference on communications systems and networks
3. Franklin RJ (2020) Traffic signals violations detection using artificial intelligence and deep learning. In: Fifth international conference on electronics systems(ICCES)
4. Saumaya A, Gayathri V, Venkateswaram K, Kale S, Sridhar N (2020) Machine learning based surveillance system for detection of bike riders without helmet and triple rides. In: International conference on smart electronics and communication (ICOSEC)

5. Prajwal MJ, Tejas K, Barsad B, Mahesh V, Murgod M, Shashidhar R (2019) Detection of non-helmet riders and extraction of license plate using YOLOv2 and OCR method. Int J Innov Technol Explor Eng (IJITEE) 9(2), ISSN: 2278–3075
6. Jamsto Y, RiyamongKol P, Waranuast E (2020) Real time license plate detection for non-helmeted motorcyclist using YOLO. ScienceDirect
7. Dahiya K, Singh D, Krishna C (2016) Automatic detection of bike riders without helmet using surveillance video real time. In: International joint conference on neural network (IJCNN).
8. Naren Babu R, Sawmya V, Soman K P (2019) Indian car number plate recognition using deep learning. In: 2nd international conference on intelligent computing, instrumentation and control (ICICICT)
9. Singh J, Bhusan B (2019) Real time indian license plate detection using deep neural networks, optical characters recognition using LSTM Tesseract. In: International conference on computing, communication intelligent systems (ICCCIS)
10. Dias C, Jagetiya A, Chaurasia S (2019) Anonymous vehicles detection for secure campus: a framework for license plate recognition using deep learning. In: 2nd international conference on intelligent communication and computational techniques (ICCT)
11. Lou L, Zhang Q, Liu C, Sheng M, Zheng Y, Liu X (2019) Vehicle detection of traffic flow using deep learning. In: IEEE 8th data driven control and learnings system conference
12. Mirthubashini J, Santhi V (2020) Video based vehicles counting. In: 6th international conference on advanced computing and communication system
13. Subbiah U, Kumar DK, Dr. Senthil Kumar T, Parmeshwaram L (2020) An extensive study and comparison of various approaches to object detection using deep learning (2020) In: Third international conference on smart system and inventive technology (ICSSIT)
14. Shrivastava A, Amudha J, Gupta D, Sharma K. Deep learning model for text recognition in images. In: 2019 10th international conference on computing, communication and networking technologies (ICCCNT), IEEE, Kanpur
15. Chebrolu KNR, Dr.(col.) Kumar PN (2019) Deep learning based pedestrian detection at all light conditions. In: IEEE international conference on communication and signal processing (ICCSP)
16. Laroca R, Zanlorensi LA, Goncalves GR, Todt E, Schwartz WR, Menotti D (2021) An efficient and layout independent automatic license plate detection system using yolo. arXiv:1909.01754v3
17. Kumar C, Punitha B, Mohana R (2020) YOLOV3 And YOLOV4: multiple object detection for surveillance. In: Proceedings of third international conference on smart system and inventive technology (ICSSIT)
18. Mallela NC, Voltey R, Srinivasa Perumal R, Nadesh RK (2020) Detection of triple riding and speed violation on two wheelers using deep learning algorithms. Springer
19. Veni S, Jose EK (2018) Yolo classification with multiple object tracking for vacant parking lot detection. J Adv Res Dynam Control Syst

Machine Learning for Cloud Resources Management—An Overview

Viktoria N. Tsakalidou, Pavlina Mitsou, and George A. Papakostas

Abstract Nowadays, an important topic that is considered a lot is how to integrate Machine Learning (ML) to cloud resources management. In this study, our goal is to explore the most important cloud resources management issues that have been combined with ML and which present many promising results. To accomplish this, we used chronological charts based on keywords that we considered important and tried to answer the question: is ML suitable for resources management problems in the cloud? Furthermore, a short discussion takes place on the data that are available and the open challenges on it. A big collection of researches is used to make sensible comparisons between the ML techniques that are used in the different kind of cloud resources management fields and we propose the most suitable ML model for each field.

Keywords Cloud computing · Cloud resource management · Machine learning · SLA

1 Introduction

In July 2020, in a demographic research conducted by Statista, approximately 4.57 billion users were active on the Internet, encompassing 59% of the global population [1]. This huge amount of users shows the demanding need for resilient, secure and easily configurable web applications. Before the cloud computing era, from the enterprise perspective, the cost of the maintenance of big data centers and bootstrap-

V. N. Tsakalidou · P. Mitsou · G. A. Papakostas (✉)
MLV Research Group, Department of Computer Science, International Hellenic University, Kavala 65404, Greece
e-mail: gpapak@cs.ihu.gr

V. N. Tsakalidou
e-mail: vitaaka@cs.ihu.gr

P. Mitsou
e-mail: pamitso@cs.ihu.gr

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_67

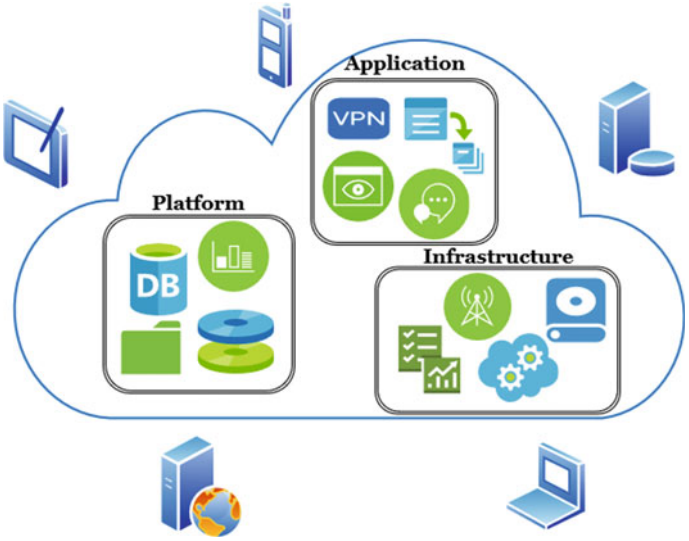


Fig. 1 Basic cloud service categories

ping a company was too high. *Cloud Computing* is the provision of virtual resources via the Internet (e.g., servers, apps, etc.), from central systems located away from the end users, which serves them by automating processes, providing convenience, flexibility of connection [2] as well as a nice pay-as-you-go payment plan to the company. In the 1950s first appeared in educational institutes and companies, and its use was made by servers with large computing and storage capabilities [2].

As Fig. 1 depicts three are the basic cloud service categories:

- *Software-as-a-Service (SaaS)*: The main concept is to not install the software on the customer's computer or servers but to deliver and access them via Internet. In that way the release process, and the bug fixes will be easier [2, 3]. There are already multiple applications that they are provided as SaaS like CRM, ERP, Word, etc.
- *Platform-as-a-Service (PaaS)*: There is also a similar concept as SaaS but in PaaS as an alternative of having to buy–pay for software licenses for platforms such as operating systems, databases and intermediate software, you subscribe and use computing platforms over the Internet. Ex. AWS Elastic Beanstalk facilitates the deployment and scaling of the web apps and it provided by AWS cloud provider [2, 3].
- *Infrastructure-as-a-Service (IaaS)*: The IaaS is even in a bigger scale and abstraction. So instead of owning and maintaining your own hardware devices in your data center, you can access them via the Internet by logging in with official credentials and they can be distributed all around the planet. These hardware devices can be networking related devices, VMs, storage devices, etc. [2, 3].

A commitment between a service provider and a client called Service-level Agreement (SLA). Wieder et al. [4] mentioned that some aspects of service quality, responsibilities, availability, have to be agreed between the service user and service provider. SLA usually contains a lot of components, from a definition of services to the termination of the agreement. Also, SLA defined at different levels:

- *Corporate-level SLA*: All end corporate customers are applicable and it imparts all the relevant aspects of the agreement. It hardly changes because mostly the issues are unchanging.
- *Customer-based SLA*: An individual agreement in a customer group that contains all the services they use.
- *Service-based SLA*: In a common deal with all customers using the services being delivered by the service provider.

The *Quality of Service* (QoS) is the measurement or description of the overall performance of a service, such as a Cloud Computing service, or a computer network, particularly the performance seen by the users of the network. Many components of network service are often considered to measure service quality such as availability, transmission delay, packet loss.

Arthur Samuel, in 1959, coined the term *Machine Learning* (ML). He was a pioneer in the field of computer gaming and AI [5, 6]. Machine Learning is the ability of the machines to learn from data. The ML makes computers to discover how to do tasks without being programmed explicitly. It steadily makes its unique way in different kind of areas in enterprise applications such as fraud detection, business intelligence and customer support. Microsoft, Amazon and Google made significant investments in Machine Learning and Artificial Intelligence (AI), by launching new services to carry out important reorganizations which will strategically place Artificial Intelligence in their organizational structures in the last years. Even Google's CEO, Sundar Pichai, mentioned that Google is shifting to an "AI-first" world. Some authors said that the hardware is impacted by ML workloads. In order to train a model to understand speech or recognize a pattern requires major parallel computing resources that could take days with the traditional computers based on CPU. We can say that powerful Graphics Processing Units (GPUs) is the first choice for the processing unit in many Machine Learning and Al workloads due to significantly reduced processing time.

This paper aims to present the applications of Machine Learning in the cloud resources management fields and the challenges that still remain open. There is an extensive investigation among the most popular topics in cloud resources management on the effectiveness of ML models and the metrics that are used to measure the performance. Also, a discussion on the popularity of this topic in the research community until now is included.

The structure of this paper is as follows. On each of the subsections of Sect. 2 is presented the related work that has been done in each cloud resources management field that ML models can be applied. Some comparisons and comments between the ML models and the data are taken place on it too. In Sect. 3 there is a discussion

around the dataset benchmarks that have been used in the research until now and the way that are used. There is an extensive discussion in the open challenges that are still relevant and important in Sect. 4 and the future work that can be conducted around this topic. Eventually, Sect. 5 concludes the paper.

2 Machine Learning in Cloud Computing

ML and cloud computing have been trends and extensively used all around the tech world. Cloud provides ML solutions throughout it but there is also a tendency in the research that explores how the cloud resources management will be benefit from ML. From 2009 till 2019 there is a constant increase in the numbers of papers that are related to the usage of ML in cloud computing as it shows in Fig. 2. These results are based on the 300 most related articles in Google Scholar. The keywords that were used to extract the results were "Cloud Resources Management" and "Machine Learning". Figure 3 presents the number of articles on different kinds of cloud resources management fields. For each distinct field, the number of papers that are used for the graphs was figured by checking that the field was part of the paper's title and it was explicitly for that. The articles that we decided to focus were the one that the number of citations was high and the paper title was more relevant to the field. Although a valid question is "Machine Learning suitable for cloud resources management problems?".

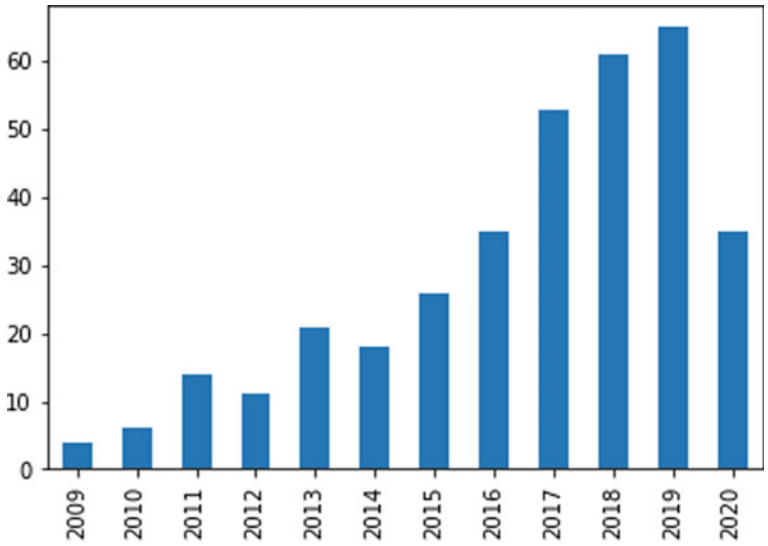


Fig. 2 Scientific articles published for "Machine Learning for Cloud Resources Management" per year

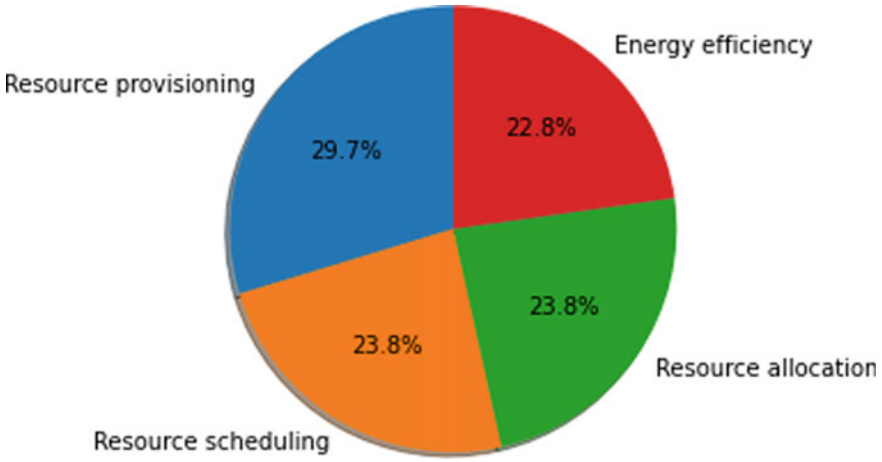


Fig. 3 Scientific articles published per section in total that involve ML solutions

2.1 Resource Provisioning

Resource provisioning has as parts the deployment, selection and management of hardware and software resources in run time for ensuring guaranteed performance for applications. Cloud providers have two provisioning options. First is the *short-term on-demand* and the second is *long-term reservation plans* [2]. There are important and challenging problems in the large-scale distributed systems that Resource provisioning is related to. Some of the resource delivery techniques used must meet service quality (QoS) parameters, such as security, availability, reliability, avoiding Service Level Agreement (SLA) violations and managing long-term reservation plans.

Alipour and Liu proposed a method to make predictions and learn workload patterns with real time load on a microservice architecture that used ML models [7]. They used AWS, to demonstrate the microservice architecture and 2 ML algorithms of logistic regression and multiclass classification. The ML models they used were the linear regression and the multinomial logistic regression. They integrated the real time prediction to the auto-scaling configuration to remove or add resources.

Islam et al. [8] developed prediction-based resource measurement and provisioning strategies using Neural Networks(NN) and Linear Regression models to satisfy upcoming resource demands. They proposed prediction techniques in the context of the dataset obtained by using TPC-W, a benchmark that is well-established for e-commerce applications. The prediction accuracy of the proposed learning models is promising and effective.

Cloud client prediction models using Machine Learning is also extensively researched by Bankole and Ajila [9] through the TPC-W benchmark web application in Resource provisioning by using three different Machine Learning techniques. Among Linear Regression (LR), Neural Networks (NN) and Support Vector,

Table 1 Related work for *Resource provisioning*

References	Year	Prediction model	Metric	Performance	Dataset
[10]	2019	SVM-based forecasting models	SLA violation	Outperform the basic forecasting models	Real dataset from University of Madrid
[7]	2017	Logistic and linear regression	CPU usage	RMSE = 1.9426	Netflix data benchmark [11]
[12]	2014	Linear regression and SVM	Profit broker earns per hour in $	SVM better that LR by 2–3%	Data collected from the application
[9]	2013	NN, SVM and Linear regression	CPU and SLA	SVM outperforms NN and LR in the MAPE and PRED metrics	TCP-W benchmark [13]
[8]	2012	NN and Linear regression	Response time	RMSE of NN = 9.6 RMSE of LR = 10.41	TPC-benchmark

it seemed SVM was the best prediction model by focusing on CPU utilization and SLA (throughput and response time) in a 9–12 min window.

Based on the prediction of the processing load that a distributed server can have and the estimation of the appropriate number of resources that must be provisioned, Moreno-Vozmediano et al. created an auto-scaling predictive framework using ML methods for queuing theory and time-forecasting [10]. For the prediction the SVM regression model was used, which appears to be ideal for the data input with linear and non-linear patterns.

According to the above researches, it seems that Resource provisioning and ML are used in combination in order to predict the upcoming resource demands. The most common ML techniques, which their results are promising and efficient, are Linear Regression (LR), Neural Networks (NN) and Support Vector Machine (SVM). Although based on Table 1 SVM appears to outperform the rest ML techniques in a variety of metrics and datasets.

2.2 *Resource Scheduling*

Resource scheduling allow us to efficiently schedule and automate any virtual job. The main target is the most efficient usage of the resources. It is a multi-dimensional work in cloud as it is depending on QoS of cloud apps. Cloud is by nature a diverse environment, so it encounters a variety of resources and uncertainty on the allocation of resources, which makes it impossible to be solved with existing the resource allocation policies. Until today research community struggles to choose an efficient algorithm for the resource scheduling on selected workload from the existing literature.

Wilk et al. [14] proposed a system with reservation scheme adaptation that used ML. It is considered important in the automated scaling the problem to define a reservation program, which allows effective resource scheduling that can be used. Based on the above issue they found a solution that allows updating a booking program originally drawn up by an administration. This will allow reservation plans to be adjusted one or more weeks in advance. In this way, the administrator has the opportunity to identify potential problems with the under-provisioning or over-provisioning on the resources. This could help to prevent the server of unnecessary costs and overloads later on. That resolution tested on OpenStack and it was used real traffic data from Wikipedia. First they used Neural Networks because were applied successfully to assign allocation adaptation in Mobile Cloud. Afterward, they compared the quality and the results among linear regression and decision trees techniques.

Authors compared the performance of the Neural Networks with the other three ML algorithms which were the RepTree , the linear regression and the M5P. In order to find and compare the results, they summed the Q values of all 365 days. They understood that the RepTree algorithm was learned 7% faster than Neural Networks, but the latter model ultimately yielded better predictions.

They authors claimed that is closed-loop solution. So the big benefit is that it can be verified in advanced the updated plan in order to identify potential issues in short and longer periods. Also, proposed that before the deployment, we can take the advantage of the historical application's data for the parameter setting and the initial training of the Machine Learning models. They believed that from the experimental results they had, this solution improves the use of cloud resources. They claim that the proposed method can be applied to other resources and it is more general [14].

Yang et al. [15] proposed a general-based solution through data driven profiling, problem formalizing, modeling with supervised, and gave a complete architecture. They believed that could achieve optimization in architectural level and improved efficiency with an ML-based method. As result they said that they consider important to pair ML algorithms and systems on edge and cloud devices.

Singh and Chana [16] mentioned that in literature, cost, time constraints and energy are some of the efficient resource scheduling policies that have been reported. They proposed a framework which was efficient on cloud management to identify and analyze the workload. K-means was used to cluster them and as weights was the QoS requirements. The experimental results were collected by CloudSim. They proposed that this framework showed improved performance on cost, execution time, energy consumption unlike existing scheduling algorithms.

Tong et al. [17] proposed a new scheduling algorithm which called QL-HEFT. It combines Q-learning with the HEFT algorithm to reduce the makespan. They used benchmarks on WorkflowSim. Their results show that the QL-HEFT has good performances in terms of the average response time. Also, it achieves a shorter makespan compared to three other classical scheduling algorithms.

The main focus of resource scheduling is how to create a cost-efficient reservation plan. For this mission different kind of ML models have been used in the researches

Table 2 Related work for *Resource scheduling*

References	Year	Prediction model	Metric	Performance	Dataset
[14]	2019	NN, Linear regression, RepTree and M5P	CPU usage	RepTree algorithm is 7% better than NN	Wikipedia server data and Grafana Wikimedia
[17]	2019	RL, Q-learning, QL-HEFT	QoS parameter	QL-HEFT can perform at least no worse than the other algorithms in 97% of situations	Inspiral, Montage, CyberShake, and Epigenomics on WorkflowSim
[15]	2018	Decition tree, NN	QoS and accuracy	prediction accuracy of node performance 92.86%	Open Cloud cluster at Carnegie Mellon University
[16]	2015	K-means	CWMF, QoS and Energy comsumption	Energy consumption was reduced from 9.99% on a hundred resources and 11.02% on the tripled amount of resources	Standard Performance Evaluation Corporation (SPEC)

based on Table 2 and QoS is a metric that there is a lot of attention on it. But still, the development of such mechanisms with this intelligent scheduling abilities becomes subsequently more challenging and complex.

2.3 *Resource Allocation*

Resource allocation includes the assignment of available resources to the Cloud apps that needs them over the Internet. Resource allocation needs to be managed precisely. If there will not be, it will starve from services. That problem can be solved with the resource provisioning which allows the management of every module independently by the service providers.

Liu et al. [18] proposed a framework that claims to solve the overall power management and resource allocation issues in cloud. It includes a tier for virtual machine global resource allocation to servers and for distributed energy management, a local tier. Except the improved scalability and the reduced action/state space dimensions, by performing local power management on the servers, it enhances the parallelism and it reduces the complexity computationally. In order the global tier problem to be solved a DRL technique is adopted. For the acceleration, a structure which has unique weight sharing was used. LSTM workload predictor helped power manager, without a model, based on Reinforcement Learning (RL) to determine a suitable action for the servers.

For experimental purposes actual Google cluster traces were used. This framework has this ability to save 53.97% energy consumption, with 95k job requests

Table 3 Related work for *Resource allocation*

References	Year	Prediction Model	Metric	Performance	Dataset
[21]	2020	Reinforcement learning	QoS	has 93.7% correctness	RUBiS benchmark
[20]	2018	Logistic and linear regression	Response time	98% prediction accuracy	DAS-2 dataset [22]
[18]	2017	Deep reinforcement Learning	Power consumption	Reduced energy by 16.20%	Google cluster-usage traces [23]
[24]	2013	SVM and KNN	CPU utilization	SVM outperforms KNN	MediaPaaS platform

in just 30-server cluster [19]. In the identical case, this framework gave the average per-job latency saving with the identical energy usage up to 16.16% , and the average power/energy saving with the identical latency up to 16.20%. According to their experiments they found that the proposed hierarchical framework significantly saved the energy consumption compared to the baseline with similar average latency. Moreover, in a server cluster, it could achieve the best trade-off among latency and power consumption [18].

Zhang et al. [20], proposed a model to analyze the resource allocation, which is a multi-dimensional problem. They proposed two algorithms for the resource allocation prediction by using logistic and Linear Regressions. They found out that the algorithm can provide a practical solution which is close to the optimal based on resource utilization and allocation accuracy.

Chen et al. [21] proposed a new method which called PCRA. They used this method to efficient resource allocation and explore adaptive. The benchmark that they used is the RUBiS. From the results they understood that this method select the management operations with 93.7% accuracy.

According to Table 3, despite the fact that resource allocation is considered an NP-hard problem, it is possible to do resource allocation predictions by using Linear Regression and the results can be really close to the optimal.

2.4 Energy Efficiency

Has become a major concern in data centers due to the environmental impact, costs and operational expenses that imparts as problems. The infrastructure still needs planning for the resources in order to address the energy wastes although it has been reduced a lot lately due to the utilization of best-practice technologies. The hardware devices and technologies have been improved based on energy consumption a lot but still some concerns exist to their usage [25].

Heath et al. [26] analyzed how to distribute the requests from clients to different servers in a cluster with heterogeneous form in which an optimal balance among

Table 4 Related work for *Energy efficiency*

References	Year	Prediction model	Metric	Performance	Dataset
[29]	2017	K-means	energy saving, power consumption	nearly 23%	CNET
[27]	2011	NN and Linear predictor	RMSE	RMSE for the LP .25 and for NN over 8.00	NASA and EPA
[28]	2010	NN	Energy consumption	Rate power reduced by 46.7% on NASA	CISRegistry
[26]	2005	Model adaptive	Less energy	Their server conserves 45% more energy	Microbenhmark

energy savings and throughput will be considered. They designed and developed a cluster which can be configured itself in order metrics such as energy consumption to be optimized. Analytical models were used the distribution of requests and the resource utilization. Simulated annealing on request distribution was used to find the minimum ratio on power to throughput. This particular method appeared to have energy consumption savings more than 40%.

Prevost et al. [27] presented a framework by combining stochastic state transition models and load demand prediction. It was used NN and linear regression models. They claimed that their model will lead to optimal cloud resource allocation by minimizing the energy consumed while maintaining required performance levels. Finally, they found out that from the simulations that linear predictor and NN seemed to have promising predictions for the future network loads. But the linear predictor eventually provided the most accurate results.

Duy et al. [28] introduced a scheduling algorithm which also characterized as a green algorithm that uses a predictor for power savings based on NN in cloud. They have conducted various experiments with several simulation configurations. They concluded that the optimal configuration to assure service level is the prediction in addition 20% more servers. They proposed that the power consumption savings for PP20 mode can achieve up to 46.3%.

Cai et al. [29] proposed the way to challenged problems of energy-efficiency in cloud. They used K-means model and methods of MapReduce model. From the experiments, they understood that the algorithm can save power consumption. After the 10 iterations of K-means algorithm seems that up to 5% savings can be accomplished and for the hibernation energy-based method 23%.

In energy efficiency, the results are promising since helps to reduce energy almost by half by optimally configuring CPU frequency and propose the most energy saving configurations. Neural Networks is an efficient ML model in multiple datasets but the Linear Predictor is more accurate as it can be seen from Table 4.

3 Datasets

In the ML, the datasets are an integral part. Cloud every day produces tones of data from the resource management processes but it seems that there are not available open data to be used in the research. In the vast majority of papers that we analyzed we observed that tools generate mostly the datasets [11, 13] or they are internal datasets that are used. There are not so many open datasets for analysis purposes and that have been used in common research works in order to be able to have better performance metrics and comparisons. The open data mostly contains information which are traces and logs on the machine events, the requested resources, scheduling information, etc. Usually, these open datasets are provided by files but also some others by SQL scripts, BigQuery tables which facilitates the analysis.

4 Discussion

Machine Learning for cloud resources management is an issue that has been discussed for the last 10–15 years. As shown in Fig. 2, from 2009 to 2019, there is a steady increase in the number of papers related to the usage of ML in cloud computing, but in 2020 this steady increase has fallen. We believe that it may be correlated with the phenomenon of COVID-19 pandemic because it is something demanding and there are a lot of data available on it. So most researchers would like to deal with this issue as it is considered to be a pandemic and it has influenced all of the people's lives this period. Also, due to cloud resources management which requires complicated decisions and policies for multiple objective optimization and therefore the effective resource management is challenging because of the scale of the cloud infrastructure. Based on the papers that are written in the field, seems that the large population of users combined with the unpredictable interactions of the system and the ML difficulties to integrate makes it challenging to solve such problems. An open challenge is also the data that is not easily accessed and available for the researchers to use them. It is observed that Machine Learning significantly helps to reduce energy consumption almost by half, by optimally configuring CPU frequency and providing the most energy saving configurations. Also, the results from Resource allocation predictions by using Linear Regression can be really close to the optimal. It can be concluded from most of the sections that Machine Learning is a really suitable solution in all of the cloud resources management fields. Mostly it helps with the prediction of future resource demands.

5 Conclusion

In this paper, we have tried to include the most important cloud resources management issues that are handled by the Machine Learning paradigm showing promising results. In particular, we observed from the conducted literature analysis that Machine Learning is the most suitable solution in cloud resources management issues. The most popular models that are used, are the Neural Networks (NN), the Linear Regression (LR) and the Support Vector Machine (SVM). We concluded that the ML can provide feasible solutions that are very close to the optimal ones based on allocation accuracy, energy saving and resource utilization. Furthermore, we realized that there is a serious problem with the data that are available. Researchers have generated a variety of in-house data, but there are not available open data in order to be used in other researches and thus we cannot have sensible comparisons on them. The evaluation of several ML oriented methodologies for solving the cloud resources management issues and their comparison with popular non ML-based techniques on the same data constitute potential directions for future research.

Acknowledgements This work was supported by the M.Phil. program "Advanced Technologies in Informatics and Computers", hosted by the Department of Computer Science, International Hellenic University, Greece.

References

1. Global digital population as of July 2020. https://www.statista.com/statistics/617136/digital-population-worldwide/
2. Lakshmi Devasena C (2014) Impact study of cloud computing on business development. Oper Res Appl Int J (ORAJ) 1:1–7
3. Rajaraman V (2014) Cloud Comput. Reson 19:242–258. https://doi.org/10.1007/s12045-014-0030-1
4. Wieder P, Butler JM, Wolfgang T, Yahyapour R (2011) Service level agreements for cloud computing. https://link.springer.com/book/10.1007/978-1-4614-1614-2. Cited 15 Jan 2011
5. Kohavi R, Provost F (1998) Glossary of terms. Mach Learn 30(2–3):271–274
6. Samuel AL (1959) Some studies in machine learning using the game of checkers. IBM J Res Develop 3(3):210–229. https://doi.org/10.1147/rd.33.0210
7. Alipour H, Liu Y (2017) Online machine learning for cloud resource provisioning of microservice backend systems. In: 2017 IEEE International conference on big data (Big Data), Boston, MA, pp 2433-2441. https://doi.org/10.1109/BigData.2017.8258201
8. Islam S et al (2012) Empirical prediction models for adaptive resource provisioning in the cloud. Future Gener Comput Syst 28(1), pp 155–162. ISSN 0167-739X, https://doi.org/10.1016/j.future.2011.05.027
9. Bankole AA, Ajila SA (2013) Predicting cloud resource provisioning using machine learning techniques. In: 2013 26th IEEE Canadian conference on electrical and computer engineering (CCECE), Regina, SK, pp. 1–4. https://doi.org/10.1109/CCECE.2013.6567848
10. Rafael M-V, Rubén M, Eduardo H, Ignacio L (2019) Efficient resource provisioning for elastic Cloud services based on machine learning techniques. J Cloud Comput 8. https://doi.org/10.1186/s13677-019-0128-9

11. Netflix Data Benchmark, https://medium.com/netflix-techblog/netflix-data-benchmark-benchmarking-cloud-data-stores-7266186ded11
12. Biswas A, Majumdar S, Nandy B, El-Haraki A (2014) Automatic resource provisioning: a machine learning based proactive approach. In: 2014 IEEE 6th International conference on cloud computing technology and science, Singapore, pp 168–173. https://doi.org/10.1109/CloudCom.2014.147
13. Smith WD. TPC-W*: Benchmarking an ecommerce solution
14. Bartłomiej Ś, Piotr P, Michal W, Marcin J, Krzysztof Z (2019) VM reservation plan adaptation using machine learning in cloud Computing. J Grid Comput 17. https://doi.org/10.1007/s10723-019-09487-x
15. Yang R, Ouyang X, Chen Y, Townend P, Xu J (2018) Intelligent resource scheduling at scale: a machine learning perspective. In: IEEE Symposium on service-oriented system engineering (SOSE), Bamberg, pp 132–141. https://doi.org/10.1109/SOSE.2018.00025
16. Singh S, Chana I (2015) QRSF: QoS-aware resource scheduling framework in cloud computing. J Supercomput 71:241–292. https://doi.org/10.1007/s11227-014-1295-6
17. Tong Z, Deng X, Chen H et al (2020) QL-HEFT: a novel machine learning scheduling scheme base on cloud computing environment. Neural Comput Appl 32:5553–5570. https://doi.org/10.1007/s00521-019-04118-8
18. Liu N, Li Z, Xu Z, Xu J, Lin S, Qiu Q, Tang J, Wang Y (2017) A hierarchical framework of cloud resource allocation and power management using deep reinforcement learning
19. Reiss C, Wilkes J, Hellerstein JL. Google cluster-usage traces: format + schema. http://code.google.com/p/googleclusterdata/wiki/TraceVersion2
20. Zhang J, Xie N, Zhang X, Yue K, Li W, Kumar D (2018) Machine learning based resource allocation of cloud computing in auction. Comput Mater Continua 56:123–135. https://doi.org/10.3970/cmc.2018.03728
21. Chen X, Zhu F, Chen Z, Min G, Zheng X, Rong C (2020) Resource allocation for cloud-based software services using prediction-enabled feedback control with reinforcement learning. In: IEEE Trans Cloud Comput. https://doi.org/10.1109/TCC.2020.2992537
22. The grid workloads archive. http://gwa.ewi.tudelft.nl/
23. Reiss C, Wilkes J, Hellerstein JL (2011) Google cluster-usage traces: format + schema, Google Inc., Mountain View, CA, USA, Technical Report, Nov 2011, revised 20 Mar 2012. Posted at http://code.google.com/p/googleclusterdata/wiki/TraceVersion2
24. Krisantus S, Andreas B (2013) Dynamic resource allocation for cloud-based media processing. In: Proceedings of the international workshop on network and operating system support for digital audio and video, pp 49–54. https://doi.org/10.1145/2460782.2460791
25. Lee YC, Zomaya AY (2012) Energy efficient utilization of resources in cloud computing systems. J Supercomput 60:268–280. https://doi.org/10.1007/s11227-010-0421-3
26. Heath T, Diniz B, Carrera EV, Meira W, Bianchini R (2005) Energy conservation in heterogeneous server clusters. In: Proceedings of the tenth ACM SIGPLAN symposium on principles and practice of parallel programming (PPoPP '05). Association for Computing Machinery, New York, NY, USA, pp 186–195. https://doi.org/10.1145/1065944.1065969
27. Prevost JJ, Nagothu K, Kelley B, Jamshidi M (2011) Prediction of cloud data center networks loads using stochastic and neural models. In: 2011 6th International conference on system of systems engineering, Albuquerque, NM, pp 276–281. https://doi.org/10.1109/SYSOSE.2011.5966610
28. Duy TVT, Sato Y, Inoguchi Y (2010) Performance evaluation of a Green Scheduling Algorithm for energy savings in Cloud computing. In: IEEE International symposium on parallel & distributed processing, workshops and Ph.D. forum (IPDPSW), Atlanta, GA, pp 1–8. https://doi.org/10.1109/IPDPSW.2010.5470908
29. Cai X-B, Ji Y-X, Han K (2017) Energy efficiency optimizing based on characteristics of machine learning in cloud computing. ITM Web Conf 12:03047. https://doi.org/10.1051/itmconf/20171203047

Correction to: A Delicate Authentication Mechanism for IoT Devices with Lower Overhead Issues

R. Raja and R. Saraswathi

Correction to:
Chapter "A Delicate Authentication Mechanism for IoT Devices with Lower Overhead Issues" in: S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_7

In the original version of the book, the following belated correction was incorporated:

In chapter "A Delicate Authentication Mechanism for IoT Devices with Lower Overhead Issues", the affiliations "Department of CSE, Faculty of Engineering, Karpagam Academy of Higher Education, Coimbatore, India" and "Assistant Professor, Computer science and Technology, SNS college of Engineering, Coimbatore, India" of author "R. Saraswathi" were changed to "Assistant Professor, Computer science and Technology, SNS college of Engineering, Coimbatore, India".

The correction chapter and book has been updated with the changes.

The updated original version of this chapter can be found at
https://doi.org/10.1007/978-981-19-3035-5_7

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*,
Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5_68

Author Index

S. Smys et al. (eds.), *Computer Networks and Inventive Communication Technologies*, Lecture Notes on Data Engineering and Communications Technologies 141,
https://doi.org/10.1007/978-981-19-3035-5