Algorithms for Intelligent Systems
Series Editors: Jagdish Chand Bansal · Kusum Deep · Atulya K. Nagar

Deepak Garg
Neeraj Kumar
Rahat Iqbal
Suneet Kumar Gupta *Editors*

Innovations in Information and Communication Technologies

Proceedings of ICIICT 2022

Springer

Algorithms for Intelligent Systems

This book series publishes research on the analysis and development of algorithms for intelligent systems with their applications to various real world problems. It covers research related to autonomous agents, multi-agent systems, behavioral modeling, reinforcement learning, game theory, mechanism design, machine learning, meta-heuristic search, optimization, planning and scheduling, artificial neural networks, evolutionary computation, swarm intelligence and other algorithms for intelligent systems.

The book series includes recent advancements, modification and applications of the artificial neural networks, evolutionary computation, swarm intelligence, artificial immune systems, fuzzy system, autonomous and multi agent systems, machine learning and other intelligent systems related areas. The material will be beneficial for the graduate students, post-graduate students as well as the researchers who want a broader view of advances in algorithms for intelligent systems. The contents will also be useful to the researchers from other fields who have no knowledge of the power of intelligent systems, e.g. the researchers in the field of bioinformatics, biochemists, mechanical and chemical engineers, economists, musicians and medical practitioners.

The series publishes monographs, edited volumes, advanced textbooks and selected proceedings.

Indexed by zbMATH.

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

Deepak Garg · Neeraj Kumar · Rahat Iqbal ·
Suneet Kumar Gupta
Editors

Innovations in Information and Communication Technologies

Proceedings of ICIICT 2022

Springer

Editors
Deepak Garg
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, India

Neeraj Kumar
Thapar Institute of Engineering and Technology
Patiala, Punjab, India

Rahat Iqbal
Big Data and IoT Research Group
Coventry University
Coventry, UK

Suneet Kumar Gupta
School of Computer Science Engineering and Technology
Bennett University
Greater Noida, India

ISSN 2524-7565 ISSN 2524-7573 (electronic)
Algorithms for Intelligent Systems
ISBN 978-981-19-3795-8 ISBN 978-981-19-3796-5 (eBook)
https://doi.org/10.1007/978-981-19-3796-5

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Organization

Honorary Co-chairs

Vincenzo Piuri, Universita' degli Studi di Milano, Italy
Jaume Anguera, Universitat Ramon LLull, Spain
Sajal K. Das, Missouri University of Science and Technology, USA
Mohammad S. Obaidat, University of Sharjah, UAE

General Co-chairs

Deepak Garg, Bennett University, India
Neeraj Kumar, Thapar University, India
Suneet Kumar Gupta, Bennett University, India
Rahat Iqbal, Coventry University, UK
Gaurav Singal, Netaji Subhas University of Technology, India

International Advisory Board

Valentina E. Balas, Aurel Vlaicu University of Arad, Romania
Ali Kattan, Knowledge University, Iraq
Ajay Shrestha, Hospital and Health Care, New York, USA
Neeraj Mittal, University of Texas Dallas, USA
Neetesh Saxena, Cardiff University, UK
Ali Kashif Bashir, Manchester Metropolitan University, UK

Program Chair

Lalit Garg, University of Malta, Malta

Technical Program Committee

Ahmed A. Elngar, Beni-Suef University, Egypt
José Rolim, University of Geneva, Switzerland
Sheroz Khan, International Islamic University, Malaysia
S. B. Goyal, City University, Malaysia
Mohammad Israr, Sur University College, Sultanate of Oman
Osama Mohammed Elmardi Suleiman Khayal, Nile Valley University, Sudan
Raad Farhood Chisab, Technical Institute, Kut, Middle Technical University, Iraq
Jeremiah Chukwuneke, Nnamdi Azikiwe University, Nigeria
Kailash Kumar Maheshwari, Saudi Electronic University, Saudi Arabia
Chidi E. Akunne, Nnamdi Azikiwe University, Nigeria
António Espírito Santo, Universidade Da Beira Interior, Portugal
P. Ganesh Kumar, K. L. N. College of Engineering, India
Amit Kumar Saraf, Jagannath University, India
Anuj Kumar Gupta, Chandigarh Group of Colleges, India
Renuka Mohanraj, Maharishi International University, India
Arjan Singh, Punjabi University, India
Balasubramani R., NMAM Institute of Technology, India
Balkrishan Y., Punjabi University, India
Ch. Chengaiah, Sri Venkateswara University, India
Chandresh Kumar Chhatlani, Rajasthan Vidyapeeth, India
Mani Jindal, Mangalmay Institute of Management and Technology, India
Charu Saxena, Chandigarh University, India
D. Sivabalaselvamani, Kongu Engineering College, India
Daisy Joseph, Bhabha Atomic Research Centre, India
Dattatraya V. Kodavade, DKTE Society's Textile and Engineering Institute, India
D. N. Sujatha, BMS College of Engineering, India
H. L. Sharma, Government College Jukhala, India
Kamlesh Lakhwani, Lovely Professional University, India
Latika Kharb, JIMS, India
Mahesh Kumar Porwal, Sree Chaitanya College of Engineering, India
Manoj Challa, CMR Institute of Technology, India
R. Sujatha, Vellore Institute of Technology, India
Rajiv Pandey, Amity University, India
Rashmi Soni, Dayananda Sagar Institutions, India
S. N. Panda, Chitkara University, India
Sanjay C. Patil, Thakur College of Engineering and Technology, India

Sanjeev Kumar, NIMS University, India
Sankit Ramkrishna Kassa, SNDT Women's University, India
Santosh Kumar Singh, JECRC Foundation Jaipur, India
Sarvesh Tanwar, Amity University, India
Shankar Chatterjee, NIRDPR, India
Shuchita Upadhyaya, Kurukshetra University, India
Shyamal Dutta, University of Burdwan, India
Soumendra Darbar, Jadavpur University, India
Sunil Kumar Gupta, BCET, Gurdaspur, India
Vaibhav Bhatnagar, S. S. Jain Subodh PG Autonomous College, India
Williamjeet Singh, Punjabi University, India
Yash Paul, Guru Jambheshwar University of Science and Technology, India
Yogesh Kumar Sharma, Shri J. J. T. University, India
M. Nirmala, Kumaraguru College of Technology, India
Charu Saxena, Chandigarh University, India
Baljit Singh Khehra, BBSB Engineering College, India
Akash Saxena, CIITM, India
Shirish Joshi, Symbiosis International Deemed University, India
Arjan Singh, Punjabi University, India
Neeraj Mangla, M. M. D. U. Mullana, India
Dinesh Kumar, Sri Sai College of Engineering and Technology, India
Monika Mathur, SKITM, India
S. Mathivilasini, Ethiraj College for Women, India
Sandhya Save, Thakur College of Engineering and Technology, India
Girish P., TOCH Institute of Science and Technology, India
Arjun Choudhary, Sardar Patel University of Police, Security and Criminal Justice, India
Deepak Kumar Mohapatra, Trident Academy of Technology, India
Harsh Khatter, ABES Engineering College, India
Hemant Sahu, Geetanjali Institute of Technical Studies, India
K. G. Revathi, DMI College of Engineering, India
Lalji Prasad, Sage University, India
Mamoon Rashid, Lovely Professional University, India
Naiyer Mumtaz, Jharkhand University of Technology, India
Prateek Agrawal, Lovely Professional University, India
Puja Das, Techno India University, India
S. Raviraja, Royal Research Foundation, India
Saravanan K., Anna University Regional Campus, India
Shashikant Sharma, B. K. Birla Institute of Engineering and Technology, Pilani, India
Subramanyam Kunisetti, RVR and JC College of Engineering, India
Suresh M., Kongu Engineering College, India
Vijay Kumar, Manav Rachna International Institute of Research and Studies, India
Vishu Madaan, Lovely Professional University, India
G. Jayalakshmi, VR Siddhartha Engineering College, India

Rohit Raina, Model Institute of Engineering and Technology, India
S. N. Tirumala Rao, Narasaraopeta Engineering College, India
V. Gomathi, National Engineering College, India
Rajat Subhra Goswami, National Institute of Technology Arunachal Pradesh, India
S. Rao Chintalapudi, Pragati Engineering College, India
D. G. Harkut, Prof. Ram Meghe College of Engineering and Management, India
D. Bujji Babu, QIS College of Engineering and Technology, India
G. Kishor Kumar, Rajeev Gandhi Memorial College of Engineering and Technology, India
R. Balaji Ganesh, Ramco Institute of Technology, India
B. Raveendra Babu, RVR and JC College of Engineering, India
Vijay Dhir, SBBS University, India
G. Prabhakar, Scient Institute of Technology, India
Shailendra Aswale, Shree Rayeshwar Institute of Engineering and Information Technology, India
Kamal Mehta, Shri Shankracharya College of Engineering, India
S. Ravi Chandra, Shri Vishnu Engineering College for Women, India
P. Kavitha Rani, Siddharth Institute of Engineering and Technology, India
M. Naresh Babu, Sree Vidyanikethan Engineering College, India
S. Jyothi, Sri Padmavati Mahila Visvavidyalayam (Women's University), India
R. Anuradha, Sri Ramakrishna Engineering College, India
K. Shirin Bhanu, Sri Vasavi Engineering College, India
V. Janardhan Babu, Sri Venkatesa Perumal College of Engineering and Technology, India
D. Nagaraju, Sri Venkateswara College of Engineering and Technology, India
Deepthi Jordana, Srinivasa Ramanujan Institute of Technology, India
D. Haritha, SRK Institute of Technology, India
Jatindra Kumar Dash, SRM University, AP, India
A. Swarnalatha, St. Joseph's College of Engineering, India
A. Kiran Mayee, Talla Padmavathi College of Engineering, India
Aditya Maheshwari, Techno India NJR Institute of Technology, India
Shachi Natu, Thadomal Shahani Engineering College, India
Sudipta Roy, Assam University, India
E. Sreenviasa Reddy, University College of Engineering and Technology Acharya Nagarjuna University, India
L. Sumalatha, University College of Engineering, India
A. Rajanikanth, Vardhaman College of Engineering, India
K. Giri Babu, Vasireddy Venkatadri Institute of Technology, India
Manne Suneetha, Velagapudi Ramakrishna Siddhartha Engineering College, India
D. Murali, Vemu Institute of Technology, India
Shalu Chopra, VES Institute of Technology, India
Ramani Bai V., Vidya Academy of Science and Technology, India
D. Aruna Kumari, Vidya Jyothi Institute of Technology, India
E. Lakshmi Lydia, Vignan's Institute of Information Technology, India
D. Srinivasarao, VNR Vignana Jyothi Institute of Engineering and Technology, India

T. Samraj Lawrence, Francis Xavier Engineering College, India
Patil V. P., Smt. Indira Gandhi College of Engineering, India
Manjunath V. Joshi, Dhirubhai Ambani Institute of Information and Communication Technology, India
Balamurgan, Galgotias University, India
P. Kumar, Rajalakshmi Engineering College, India
Mita Parikh, Sarvajanik College of Engineering, India
Vaibhav Gandhi, B. H. Gardi College of Engineering and Technology, India
Krishna K. Warhade, MIT World Peace University, India
Nishu Gupta, Vaagdevi College of Engineering, India
Premchand Bhagwan Ambhore, Government College of Engineering, India
Muhammad Naufal Bin Mansor, Universiti Malaysia Perlis, Malaysia
Anshul Tripathi, RGPV, India
Hadi Erfani, Islamic Azad University, Iran

Preface

The objective of 2nd International Conference on Innovations in Information and Communication Technologies (ICIICT-2022) was to provide the platform to researchers, developers and academicians for sharing the thoughts on recent developments in the computer science domain. In ICIICT-2022, there were a series of keynote lectures and oral presentations. The focused areas of all these activities were recent developments in the domain of artificial intelligence, blockchain, computer network, data science, computer graphics, etc. The motive of the conference was to support the computer society to get benefits from the advances of next-generation communication and computer technology.

The conference was scheduled on April 15 and 16, 2022, in online mode, and many researchers have exchanged their thoughts on recent developments and applications. Acceptance rate in ICIICT-2022 is only 10%, and four IEEE/ACM fellows were at the key positions of the conference to make this conference a grand success.

Greater Noida, India — Deepak Garg
Patiala, India — Neeraj Kumar
Coventry, UK — Rahat Iqbal
Greater Noida, India — Suneet Kumar Gupta
April 2022

Contents

About the Editors

Dr. Deepak Garg is Director of leadingindia.ai; Dean, International Relations & Corporate Outreach and School of Computer Science Engineering & Technology Bennett University. He is also Director of NVIDIA-Bennett Center of Research on Artificial Intelligence. He is leading the largest Development, Skilling and Research initiative in AI in India with more than 1000 institutional collaborators. He is a chief consultant for algorithmguru.in. He has done his Ph.D. in efficient algorithm design in 2006. He served as chair of IEEE Computer Society, India IEEE Education Society (2013–15). He has handled funding of around INR 700 million including RAENG, UK on MOOCs, Machine Learning and AI. He has 100+ publications with 1250+ citations and Google h-index of 18. In his 20 years of experience he has delivered 300+ invited talks and conducted 100+Workshops and 15+ Conferences across the country. He has Supervised 13 Ph.D. and 35 PG students. He is a blogger in Times of India named as breaking shackles.

Prof. Neeraj Kumar (SMIEEE) (2019, 2020 highly-cited researcher from WoS) is working as a Full Professor in the Department of Computer Science and Engineering, Thapar Institute of Engineering and Technology (Deemed to be University), Patiala (PB), India. He is also adjunct professor at Asia University, Taiwan, King Abdul Aziz University, Jeddah, Saudi Arabia and Charles Darwin University, Australia. He has published more than 500 technical research papers in top-21 cited journals and conferences which are cited more than 31269 times from well 22 known researchers across the globe with current h-index of 97. He is highly cited researcher in 2019 in the list released by Web of Science (WoS). He has guided many research scholars leading to Ph.D. and M.E./M.Tech. His research is supported by funding from various competitive agencies across the globe. His broad research areas are Green computing and Network management, IoT, Big Data Analytics, Deep learning and cybersecurity.He has also edited/authored 10 books with International/National Publishers like IET, Springer, Elsevier, CRC. Security and Privacy of Electronic Healthcare Records: Concepts, paradigms and solutions (ISBN-13: 978-30 1-78561-898-7), Machine Learning for cognitive IoT, CRC Press, Blockchain, Big Data and Machine learning, CRC Press, Blockchain Technologies across industrial vertical, Elsevier, Multimedia Big Data Computing for IoT Applications: Concepts, Paradigms and Solutions (ISBN: 978-981-13-8759-3), Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019) (ISBN 978-981-15-3369-3). Probabilistic Data Structures for Blockchain based IoT Applications, CRC Press. One of the edited text-book entitled, *Multimedia Big Data Computing for IoT Applications: Concepts, Paradigms, and Solutions* published in Springer in 2019 is having 3.5 million downloads till 06 June 2020. It attracts attention of the researchers across the globe. He is serving as editors of *ACM Computing Surveys, IEEE Transactions on Sustainable Computing, Elsevier Journal of Networks and Computer Applications Elsevier, Computer Communication, Wiley International Journal of Communication Systems*. Also, he has organized various special issues of journals of repute from IEEE, Elsevier, Springer. He has been a workshop chair at IEEE Globecom 2018, IEEE Infocom 2020 and IEEE ICC 2020 and track chair of Security and privacy of IEEE MSN 2020. He is also TPC Chair and member for various International conferences such as IEEE MASS 2020, IEEE MSN2020. He has won the best papers award from IEEE Systems Journal and IEEE ICC 2018, Kansas-city in 2018. He won the best researcher award from parent organization every year from last eight consecutive years.

Dr. Rahat Iqbal is a senior academic with 20 years' experience in industrial and academic roles. His main research interests lie in Big Data Analytics, Internet of Things (IoT), Information Retrieval, Artificial Intelligence and Energy Efficient Communication Protocols. He is leading the Research Group in Big Data and IoT in the Institute for Future Transport and Cities. Dr. Iqbal has supervised to completion 17 Ph.D. students. He serves on programme committee and advisory committee of many international conferences and journals. He has published more 130 papers in international journals, conferences, book chapters and workshops. He has also co-invented 5 patents relating to biologically inspired deep machine learning approaches for fault detection and isolation, route optimisation and monitoring driving. Dr. Iqbal is a principle investigator on a 3.5-year Ph.D. project (£104,892) titled "Urban flood monitoring and forecasting" funded by EIT Digital and UK Industrial partner (2021–2025). Dr. Iqbal is a principle investigator on a 3.5-year Ph.D. project (£104,892) titled "Passenger profiling for autonomous car using AI/machine learning" funded by EIT Digital and UK Industrial partner (2021–2025). Dr. Iqbal is a principle investigator on a 2-year project (£105,310) titled "Digital Connected and Autonomous Vehicle Proving Ground (DIGICAV)" funded by Innovate UK in partnership with HORIBA MIRA (2019–2020). DIGICAV. Dr. Iqbal is a principle investigator on a two-year project (£750,000) titled "Flood Ultra-Cognitive Dendrite (FLUD)" funded by Innovate UK and PlaTcom Venturs, Malaysia under the Technology Programme "Newton fund: UK-Malaysia urban Innovation Challenge 2017 Competition for funding". Dr. Iqbal is a principle investigator for knowledge transfer partnership and spin-out company, a supplier of high-tech intelligent data-driven solutions. He is seconded to work as CEO at Interactive Coventry Limited. Dr. Iqbal was principle investigator on a one-year project (2012–2013) "Driver's intent prediction: self-learning car initiative" funded (£49,301) by Technology Strategy Board funded in collaboration with Jaguar Land Rover Ltd. From 2013–2015, Dr. Iqbal worked on a number of Knowledge Exchange and Enterprise Network (KEEN) projects in collaboration with Infinity Wireless Ltd, Modus IT Ltd to develop high-tech solutions relating to data recovery apps and big data analytics for businesses. He was principle investigator on a two-year project (2011–2013) "Development of an innovation teaching tool using web/mobile and service oriented technologies" funded by Technology Strategy Board Project in collaboration with Currant limited. From 2009–2013, he was principle investigator for an EPSRC funded (£85,052) project titled "Task Sensitive Search Recommender System Based on Relevance Feedback". The project was carried out in collaboration with Trinity Expert Systems. From 2007–2009, Dr. Iqbal was involved in a number of projects to introduce novel teaching, learning and assessment methods. These projects were funded by Coventry University, CILASS: Centre for Inquiry-based Learning in the Arts and Social Sciences, University of Sheffield. Also over this period, he was involved in a number of industrial consultation activities. Dr. Iqbal was also involved in the project management and development of the EU FP7 project CHIL (Computers in Human Interaction Loop) at the Technical University of Eindhoven, Netherlands.

Suneet Kumar Gupta is a working as Associate Professor at School of Computer Science Engineering & Technology at Bennett University, Greater Noida, India. His current research interests are Wireless Sensor Network, Internet of Things, Natural Language Processing and Brain-Computer Interaction. Dr. Gupta has completed Wireless Sensor Network based project funded by Department of Science and Technology, Uttar Pradesh. Dr. Gupta is the reviewer of many international journals i.e. *Computer Network, Wireless Personal Communication, Information Science, IEEE Transactions* etc. Dr. Gupta is also part of a project funded by the Royal Academy of Science London titled with Leadingindia.ai. Under this project, Dr. Gupta has organized 70 workshops on Deep Learning and Internet of Things (IoT). He has 80+ publications with 1600+ citations and Google h-index of 20.

Chapter 1
Deep Learning Model for Arrhythmia Classification with 2D Convolutional Neural Network

S. Nithya and Mary Shanthi Rani

1 Introduction

Cardiac arrhythmia disease is a medical condition that transpires in the heart chamber due to heartbeat irregularities. An irregular heartbeat depends on age, body mass index (BMI), stress level, family history, and activity. The beats that are more than 100/m beats per minute (bpi) are represented as tachycardia, and beats below 60/m (bpi) are denoted as bradycardia. These irregularities are also defined as palpitations. Generally, arrhythmia is categorized into two major types, ventricular arrhythmias and supraventricular arrhythmias. Ventricular arrhythmias, which occur in the ventricle, are more rapidly dangerous than supraventricular arrhythmia in the upper chamber of the heart. Arrhythmia is classified into several subtypes, atrial flutter, atrial fibrillation, ventricular fibrillation, bradycardia, tachycardia, premature ventricular contractions (PVC), etc. Pellucid classification of arrhythmia could aid clinicians' diagnosis of cardiovascular disease [1].

The most commonly used non-invasive and inexpensive method for diagnosing cardiac disease is electrocardiogram signal recordings. The electrical impulse generated by the sinus node (SA) is recorded with electrodes placed in the chest [2].

Clinical experts take more time to make predictions by identifying abnormalities in the ECG signal. This has motivated the researchers to explore the intervention of artificial intelligence methods in cardiac arrhythmia detection for fast diagnosis and timely treatment to avoid fatalities [3]. Deep learning has witnessed rapid growth in

S. Nithya · M. S. Rani (✉)
The Gandhigram Rural Institute (Deemed to be University), Gandhigram, Tamilnadu 624 302, India
e-mail: drmaryshanthi@gmail.com

S. Nithya
e-mail: nithyaselvaraj2019@gmail.com

D. Garg et al. (eds.), *Innovations in Information and Communication Technologies*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-19-3796-5_1

recent decades with its significant performance in various domains such as robotics, compression, satellite imaging, medical imaging, agriculture, and much more [4–7].

Electrocardiogram signals are recorded as one-dimensional signals represented as time series data which can be analyzed with deep learning algorithms [8].

Generally, one-dimensional ECG (1D-ECG) signals accumulate lots of noise during recording, and hence, preprocessing is essential to eliminate the outliers, which is less versatile. To avoid this tedious process, a 1D-ECG signal is converted to two-dimensional ECG images (2D-ECG) to speed up accurate classification [9, 10].

The spur of this research work is transforming 2D ECG images with 1D-ECG for detecting arrhythmia. These 1D signals are segmented into beats. Each beat is analyzed segmented based on the annotation symbols separated as beats and non-beats.

This proposed work is sequentially systemized as follows: Sect. 2 presents the related works. The methodologies of proposed methods are presented in Sect. 3, and a discussion on experimental results is given in Sect. 4. Finally, the conclusion is given in Sect. 5.

2 Related Works

The challenge in identifying abnormal beats remains in the research fields. Several methods have been developed to distinguish various classes of arrhythmia.

Güler and Übeylı [11] applied discrete wavelet transform (DWT) to decay the time–frequency signal. At the second stage, the four classes of arrhythmia are detected and an accuracy of 96.94% was reached with a combined neural network model. Yu and Chou [12] developed a technique to unify independent component analysis (ICA) with neural networks to classify ECG beats. Probabilistic neural network (PNN) with backpropagation neural network (BPNN) was used as classifiers to separate eight beats. The PNN showed a better result of 98% accuracy. Melgani and Bazi [13] introduced a scheme to generalize the support vector machine (SVM) classifier along with the support of particle swarm optimization (PSO). The scheme classified six classes of arrhythmia, including the normal type. The new PSO-SVM reached overall accuracy of 89.72%. Ubeyli [14] trained with Levenberg–Marquardt algorithm on segmented ECG beats. Four types of arrhythmia are classified with RNN combined with eigenvector methods. Dutta et al. [15] introduced an automated medical diagnostic tool designed along the least square support vector machine (LS-SVM). The developed tool classified three types of regular, PVC, and other beats by reaching an accuracy of 95.82%. Kumar and Kumaraswamy [16] suggested arrhythmia classification by utilizing discrete cosine transform (DCT) in the RR interval and reached a 90% accuracy. Park et al. [17] presented a new method to extract complex QRS waves and P waves using the Pan–Tompkins algorithm to classify beats of the nearest neighbor algorithm to achieve an accuracy of 97%. Sahoo et al. [18] identified the complex QRS complex with less error rate and reached high accuracy of 98.39% with SVM and tried with neural network and attained 96.67%

accuracy for detecting ECG beats. Singh et al. [19] suggested a classification of normal and abnormal arrhythmia by implementing a recurrent neural network (RNN) and attained an 88.1% of accuracy when compared with other RNN methods. Altan et al. [20] developed a novel technique implementing a deep belief neural network in the task of classifying five classes of arrhythmia, which works in multiple stages to attain an accuracy of 94.15% in five classes of arrhythmia. Wavelet packet decomposition, discrete Fourier transform, morphology, and higher-order statistics are applied to extract the features. Izci et al. [21] applied transformation of ECG signals to image and applied 2D CNN and achieved an accuracy of 97.42% without any preprocessing. Rajkumar et al. [22] introduced a deep learning model to classify irregular rhythms using CNN. Seven classes irregular rhythm. The results are compared with different activation functions, and ELU activation provided better performance among other activation functions. Ullah et al. [23] developed a novel technique to identify eight classes of arrhythmia 1D CNN and with 2D CNN. The input image they have used is with and without augmentation techniques. The highest accuracy reached by the 2D CNN is 98.92%.

Earlier research works reported on 1D-ECG signal, and now, the direction of research is focused on noise-free 2D-ECG images for classification. Hence, this research work uses 2D-ECG images for accurate prediction of arrhythmia by incorporating more unexplored arrhythmia classes with two developed models ARBC and ARMC.

3 Materials and Methods

3.1 Experimental Setup

The evaluation of the experiment of the proposed method is done with the publicly available MIT-BIH arrhythmia dataset. It has several arrhythmia signals annotated by two or more cardiologists separately based on the types of arrhythmia. Every record is obtained from two lead channels: modified limb lead II (MLII) and V1, V2, V4, or V5. Since there is deformation in the second channel for the development of the proposed model, MLII lead recording is taken. Every record in the ECG signal is a 30 min recording with 360 Hz sampling [24].

3.2 Methodology

In this section, we have proposed two arrhythmia classification methods, which involve the following key phases.

Beat Segmentation

Each heartbeat signal in the MIT-BIH arrhythmia dataset is annotated by a cardiologist depending on the QRS structure. Based on the annotations, the heartbeats are segmented with the help of Waveform Database Toolbox (WFDB) in python. The QRS complex in the ECG beat is identified based on the annotation with this WFDB toolbox.

Transformation of Signals

Generally, conventional methods for arrhythmia detection mostly experimented with 1D-ECG signals. The ECG signals are recorded with electrode motion artifacts, electromyographic (EMG) noise, powerline interference, and baseline wander. Detecting arrhythmia with these noises is too tricky, and hence, high-pass and low-pass filters are applied to eradicate noise before classification. The proposed method eliminates these hurdles by converting 1D ECG signals to 2D images as this transformation process effectively suppresses noise. Each beat is saved as a 128 $\times$ 128-pixel grayscale image to preserve the resolution. Conversion of RGB to grayscale is performed as RGB color ECG signal has no significant impact in distinguishing arrhythmia.

Furthermore, this grayscale conversion reduces the image dimension, thereby shrinking the computational complexity of the classification procedure. The converted images are saved as two separate datasets. The first dataset holds normal and abnormal beats for binary class classification. The second dataset contains 12 classes of beats, namely normal, APB, FPBN, FVCN, LBBB, PAB, PB, PVC, RBBB, VEB, and VW, for multi-class classification. The sample of segmented beats in the multi-class dataset is visualized in Fig. 1.

3.3 Convolutional Neural Network

A CNN is a robust architecture which is constructed for analyzing 2D data formats like images and videos. Feature extraction is not needed in CNN. Automated feature extraction and classification are combined to pick out the robust features from the input images. CNN has been designed with three layers in the architecture, namely convolution layer, pooling layer, and fully connected layer [25]. Our proposed model has been structured with three convolution layers, three pooling layers, and a fully connected layer depicted in Fig. 2.

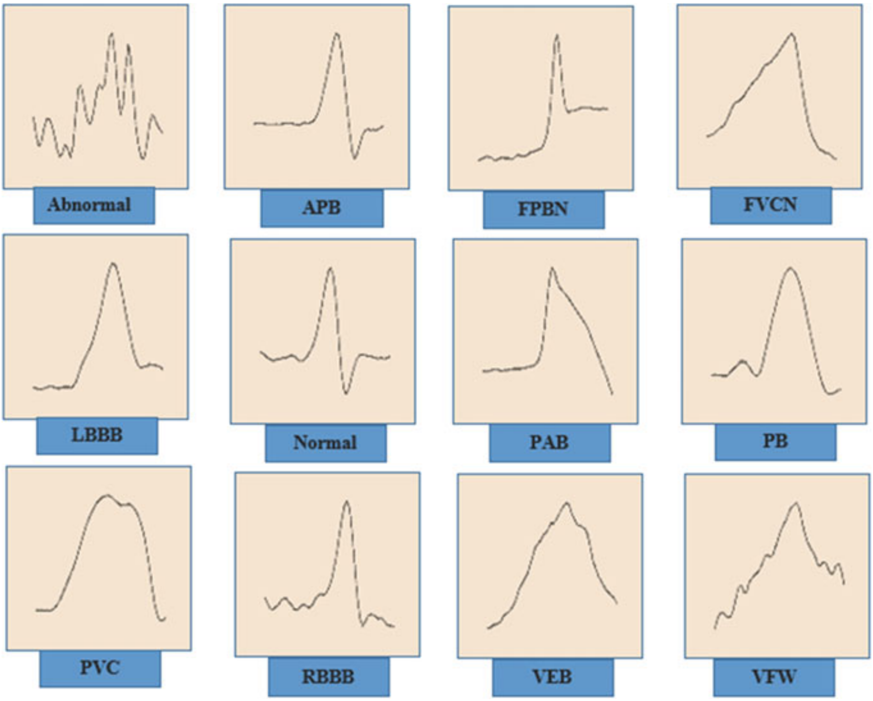


Fig. 1 Segmented sample beats

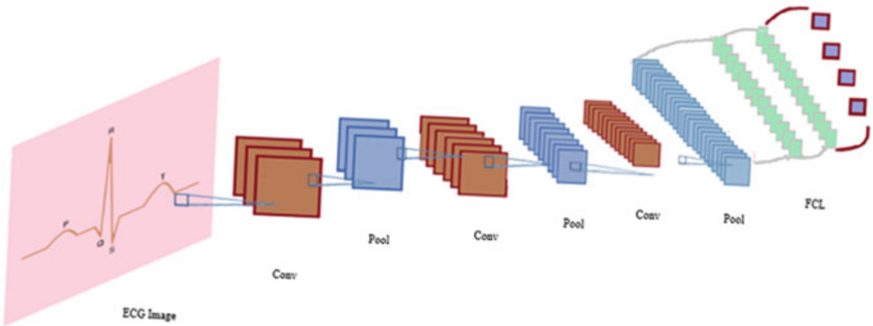


Fig. 2 Proposed architecture of the 2D-CNN model

3.4 Arrhythmia Binary Classifier (ARBC)

The objective of this ARBC is to do the binary classification of abnormal and normal beats in the raw signal. Among several annotated beats, non-beats are wiped out

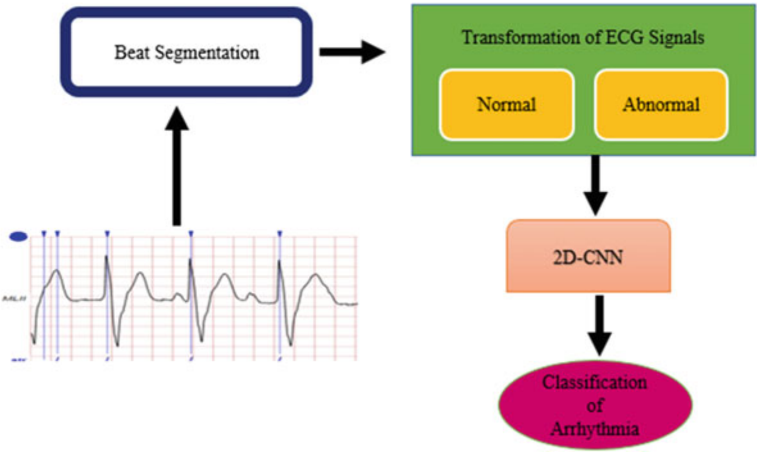


Fig. 3 Diagrammatic depiction of ARBC

from segmentation. The remaining arrhythmia beats are segmented as normal and abnormal beats to form binary class beats dataset. This dataset is fed as an input to a 2D convolutional neural network (2D-CNN) for training. The dataset is subdivided in the ratio of 3:1:1 for training, validation, and testing, respectively. The trained model is validated and tested with test dataset. This model classifies the input ECG signal as abnormal or normal. The various stages in the implementation of the proposed ARBC are summarized below.

Stage 1: Beat segmentation of input signals into 12 classes.
Stage 2: Transformation segmented beats into 2D binary dataset representing normal and abnormal beats.
Stage 3: Implementation of ARBC using 2D-CNN with binary class dataset.
Stage 4: Binary classification of arrhythmia.

The pictorial presentation of the ARBC model is depicted in Fig. 3.

3.5 Arrhythmia Multi-class Classifier (ARMC)

This ARMC method focuses on the multi-class classification of arrhythmia from the raw ECG signal. As discussed in 3.2, the non-beats are eradicated and the abnormal beats are categorized into twelve classes forming a multi-class dataset. This dataset has been utilized for classifying multiple types of arrhythmia beats. The stages in ARMC are listed below.

Stage 1: Beat segmentation of input signals into 12 classes.
Stage 2: Transformation of segmented beats into 12 classes.
Stage 3: Implementation of ARMC using 2D-CNN with multi-class dataset.

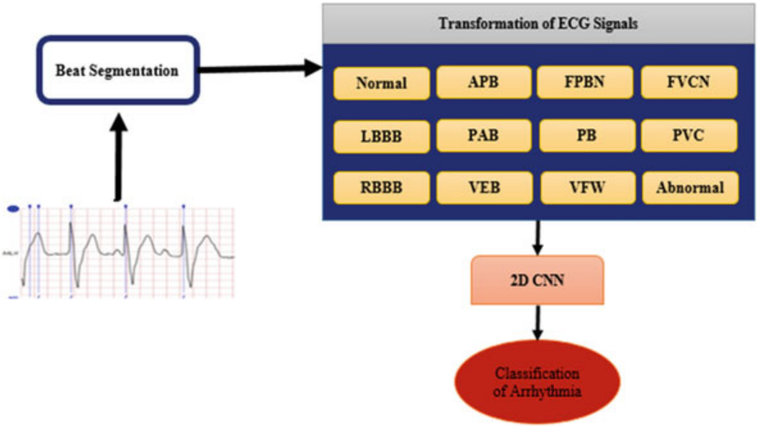


Fig. 4 Schematic representation of ARMC model

Stage 4: Multi-class classification of arrhythmia.

The diagrammatic representation of the ARMC model is portrayed in Fig. 4.

4 Results and Discussion

The proposed model is implemented with intel(R) Core(TM) i5-9300H CPU @ 2.40 GHz, NVIDIA® GeForce® GTX 1050,64-bit operating system, ×64-based processor, 8 GB RAM. Experiments were performed with two datasets, binary dataset representing normal and abnormal beats and multi-class dataset representing 12 classes. The description of the multi-class dataset with the no. of beats in each class and respective annotation symbol is shown in Table 1.

Binary classification and multi-class classification are done separately with 2D CNN. The dataset has been divided into 3:1:1 ratios for training, testing, and validating, respectively, for both datasets. The parameters used in the proposed 2D CNN are exhibited in Table 2.

The performance of the proposed methods is assessed using the standard metrics, accuracy, F1-score precision, and sensitivity, which are evaluated using Eqs. (1–4). Accuracy is calculated by the total no. of exactly classified values divided by the total number of images. True positive (TP) denotes when both actual and predicted values are similar. When it is different, it is said to be true negative (TN). When the predicted value is true and the actual result is false, it is determined as false positive (FP). False negative (FN) indicates when the predicted value is false and the actual value is true.

$$\text{Precision} = \frac{\text{TP} * 100}{\text{TP} + \text{FP}} \tag{1}$$

Table 1 Beat segmentation based on annotation

Beat type	Classes	Annotation symbol	No. of beats segmented
Atrial premature beat	APB	A	2696
Left bundle branch block beat	LBBB	L	8075
Right bundle branch block beat	RBBB	R	7259
Fusion of paced and normal beat	FPBN	f	982
Ventricular flutter wave	VFW	!	472
Premature ventricular contraction	PVC	V	7130
Original paced beat	PAB	/	7032
Nodal premature beat	PB	J	85
Fusion of ventricular and normal beat	FVCN	F	803
Ventricular escape beat	VEB	E	106
Normal beat	Normal	N	75,012

Table 2 Hyperparameter in the proposed model ARBC and ARMC

Parameters	ARBC	ARMC
Total parameters	158,786	1,157,260
Optimizer	Adam	Adam
Learning rate	0.0001	0.0001
Loss function	Binary cross-entropy	Categorical cross-entropy

$$\text{F1-score} = 2 * \frac{(\text{Precision} * \text{recall})}{(\text{Precision} + \text{recall})} * 100 \tag{2}$$

$$\text{Sensitivity} = \frac{\text{TP} * 100}{\text{TP} + \text{FN}} \tag{3}$$

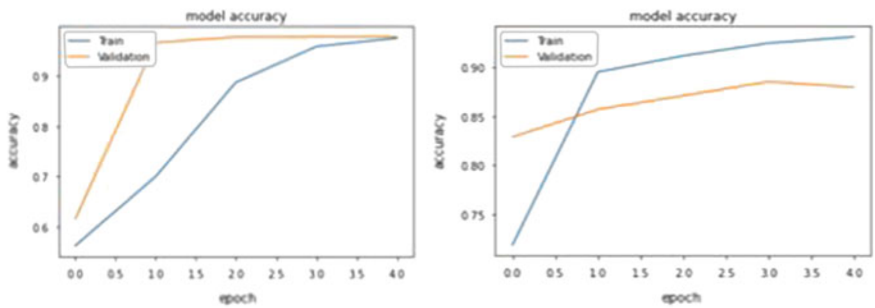
$$\text{Accuracy} = \frac{(\text{TN} + \text{TP}) * 100}{\text{TN} + \text{FP} + \text{TP} + \text{FN}} \tag{4}$$

Experimental results achieved by models ARBC and ARMC are shown in Table 3. Test results demonstrate that the binary classification of normal and abnormal classes reached the highest accuracy of 96.88%, sensitivity 96%, precision 97%, and F1-score 97%. It is also apparent from Table 3 that the ARMC model has reached

Table 3 Demonstrates the performance of the ARBC and ARMC

Methods	No. classes	Accuracy	Precision	Sensitivity	F1-score
ARBC	2	96.88	97	96	97
ARMC	8	98.98	98	94	96

an accuracy of 98.98%, the sensitivity of 94%, precision of 98%, and 96% in F1-score. Though the multi-class dataset contains 12 classes, the ARMC model can classify only eight classes, namely APB, FPBN, FVCN, LBBB, PAB, PVC, RBBB, and normal. The reason is an imbalance in the distribution of samples among the various classes, specifically PB, VEB, VW, and abnormal classes, in the dataset. The balanced distribution in the class is needed for deep learning models, where the classified heartbeats are not equal in number, which causes an imbalance in the dataset, with a significant impact on classification accuracy. The impact of the number of epochs on the accuracy for both the models is pictorially represented in Fig. 5.



Accuracy reached by Binary Class

Accuracy reached by Multi Class

Fig. 5 Accuracy reached by ARBC and ARMC

Table 4 shows the performance comparison of ARBC and ARMC methods with other state-of-the-art methods in terms of performance metrics for binary and multi-class arrhythmia classification.

Table 4 reveals that the proposed models outperform similar standard methods by achieving higher accuracy. It is also notable that ARMC reached the highest accuracy of 98.98% in multi-class classification among all other methods.

5 Conclusion

This work proposes a novel work on classifying two classes and eight classes of ECG arrhythmia classification. The segmented ECG beats are converted into 2D images to suppress noise for accurate arrhythmia classification. Furthermore, the dimensionality reduction of the image reduces the computational complexity as well. Experimental results prove the superior performance of the proposed ARBC and ARMC models with the highest accuracy of 98.98% when analyzed with the other state-of-art methods. In the future work, the imbalance ratio in the dataset is to be addressed for the classification of the remaining classes of arrhythmia.

Table 4 Comparison of the ARBC and ARMC with other state-of-art methods

Methods	No. classes	Accuracy	Precision	Sensitivity	F1-score	Specificity
FFNN [11]	4	96.4	–	96.31	–	97.78
PNN [12]	8	98.71	–	–	–	99.65
LS-SVM [15]	3	95.82	97.01	86.16	0.91	99.17
SVM [13]	6	91.67	–	93.83	–	90.49
RFT [16]	3	92.16	–	–	–	–
RNN [14]	4	98.06	–	98.15	–	97.78
KNN [17]	17	97	–	96.60	–	95.80
SVM [18]	6	98.39	–	–	–	–
1D-CNN [26]	5	96.40	79.20	68.80	0.73	99.50
LSTM [19]	2	88.1	–	92.4	–	83.35
DBN [20]	5	94.15	–	–	–	–
2D-CNN [21]	5	97.42	–	–	–	–
1D CNN [22]	7	93.60	–	–	–	–
2D CNN [23]	8	98.92	98.69	97.26	0.98	99.67
ARBC	2	96.88	97	96	97	–
ARMC	**8**	**98.98**	98	94	96	–

To indicate the proposed model reached the highest accuracy compared with other state-of-methods. To indicate ARBC and ARMC is the proposed model

References

1. Aziz S, Ahmed S, Alouini MS (2021) ECG-based machine-learning algorithms for heartbeat classification. Sci Rep 11(1):1–14
2. Sree V, Mapes J, Dua S, Lih OS, Koh JE, Ciaccio EJ, Acharya UR (2021) A novel machine learning framework for automated detection of arrhythmias in ECG segments. J Ambient Intell Humaniz Comput 12(11):10145–10162
3. Bhagyalakshmi V, Pujeri RV, Devanagavi GD (2021) GB-SVNN: Genetic BAT assisted support vector neural network for arrhythmia classification using ECG signals. J King Saud Univ Comp Inf Sci 33(1):54–67
4. Mary Shanthi Rani M, Chitra P, Lakshmanan S, Kalpana Devi M, Sangeetha R, Nithya S (2022) DeepCompNet: a novel neural net model compression architecture. Computational Intelligence and Neuroscience
5. Devi MK, Rani MMS (2021) A comparative study of machine learning classifiers for diabetic retinopathy detection. In: Advances in Automation, Signal Processing, Instrumentation, and Control: Select Proceedings of i-CASIC 2020. Springer Singapore, pp 735–742
6. Sangeetha R, Rani MMS (2021) A novel method for plant leaf disease classification using deep learning techniques. In: Machine learning, deep learning and computational intelligence for wireless communication. Springer, Singapore, pp 631–643
7. Karthikeyan N, Rani MS (2022) ECG Classification Using Machine Learning Classifiers with Optimal Feature Selection Methods. In Evolutionary Computing and Mobile Sustainable Networks, Springer, Singapore, pp 277–289
8. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK (2018) Medical image analysis using convolutional neural networks: a review. J Med Syst 42(11):1–13

9. Wu Y, Yang F, Liu Y, Zha X, Yuan S (2018) A comparison of 1-D and 2-D deep convolutional neural networks in ECG classification. arXiv preprint arXiv:1810.07088
10. Zhao J, Mao X, Chen L (2019) Speech emotion recognition using deep 1D & 2D CNN LSTM networks. Biomed Sig Process Control 47:312–323
11. Güler İ, Übeylı ED (2005) ECG beat classifier designed by combined neural network model. Pattern Recogn 38(2):199–208
12. Yu SN, Chou KT (2008) Integration of independent component analysis and neural networks for ECG beat classification. Expert Syst Appl 34(4):2841–2846
13. Melgani F, Bazi Y (2008) Classification of electrocardiogram signals with support vector machines and particle swarm optimization. IEEE Trans Inf Technol Biomed 12(5):667–677
14. Ubeyli ED (2009) Combining recurrent neural networks with eigenvector methods for classification of ECG beats. Digit Sig Process 19(2):320–329
15. Dutta S, Chatterjee A, Munshi S (2010) Correlation technique and least square support vector machine combine for frequency domain based ECG beat classification. Med Eng Phys 32(10):1161–1169
16. Kumar RG, Kumaraswamy YS (2012) Investigating cardiac arrhythmia in ECG using random forest classification. Int J Comput Appl 37(4):31–34
17. Park J, Lee K, Kang K (2013) Arrhythmia detection from heartbeat using the k-nearest neighbor classifier. In: 2013 IEEE international conference on bioinformatics and biomedicine. IEEE, pp 15–22
18. Sahoo S, Kanungo B, Behera S, Sabut S (2017) Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities. Measurement 108:55–66
19. Singh S, Pandey SK, Pawar U, Janghel RR (2018) Classification of ECG arrhythmia using recurrent neural networks. Procedia Comp. Sci. 132:1290–1297
20. Altan G, Allahverdi N, Kutlu Y (2018) A multistage deep learning algorithm for detecting arrhythmia. In: 2018 1st international conference on computer applications & information security (ICCAIS). IEEE, pp 1–5
21. Izci E, Ozdemir MA, Degirmenci M, Akan A (2019) Cardiac arrhythmia detection from 2d ECG images by using deep learning technique. In: 2019 Medical technologies congress (TIPTEKNO). IEEE, pp 1–4
22. Rajkumar A, Ganesan M, Lavanya R (2019) Arrhythmia classification on ECG using deep learning. In: 2019 5th International conference on advanced computing & communication systems (ICACCS). IEEE, pp 365–369
23. Ullah A, Anwar SM, Bilal M, Mehmood RM (2020) Classification of arrhythmia by using deep learning with 2-D ECG spectral image representation. Remote Sens 12(10):1685
24. Moody GB, Mark RG (2001) The impact of the MIT-BIH arrhythmia database. IEEE Eng Med and Biol (PMID: 11446209) 20(3):45–50
25. Yıldırım O, Pławiak P, Tan RS, Acharya UR (2018) Arrhythmia detection using deep convolutional neural network with long duration ECG signals. Comput Biol Med 102:411–420
26. Ince T, Kiranyaz S, Eren L, Askar M, Gabbouj M (2016) Real-time motor fault detection by 1-D convolutional neural networks. IEEE Trans Industr Electron 63(11):7067–7075

Chapter 2
Sign Language Recognition System for Deaf People

Adarsh Sharma, Shantanu Pingale, Uddhav Sabale, Nikita Patil, and Shital Dongre

1 Introduction

American Sign Language (ASL), visual gesture language, is used by most of the deaf community all over the world [1]. Nowadays, establishing communication with deaf and mute people is of utmost importance and necessity. These gestures are powerful means of communication within such especially abled people also in normal people. Also, it becomes difficult to find a well-experienced and educated translator for sign language every time and everywhere [1]. The main aim of this paper is to discuss Sign Language Recognition from the application point of view.

This paper will talk starting from data acquisition to recognition and real-time use also about past research made in this field. With many constraints, factors and limitations, Sign Language Recognition (SLR) is a challenging and motivating task [2]. The sign language recognition system is an important task that directly impacts society by bridging the communication gap and encouraging the deaf–mute community.

A Sign Language Recognition system is difficult to implement because of the variations in hand gestures, facial expressions, body movements, some external factors, and many other variations and constraints [3]. Despite having difficulties, the main motivation behind developing the model is a social impact, overcoming the limitations of current implementations, cost constraints, etc.

2 Literature Review

Sign Language to Text Conversion Using Deep Learning by Kartik P.V.S.M.S., Sumanth K.B.V.N.S., Ram V.N.V.S., Prakash P proposed a model for sign language

A. Sharma (✉) · S. Pingale · U. Sabale · N. Patil · S. Dongre
Vishwakarma Institute of Technology, Pune, India
e-mail: adarsh.sharma18@vit.edu

D. Garg et al. (eds.), *Innovations in Information and Communication Technologies*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-19-3796-5_2

conversation using CNN, image processing and deep learning methods. The accuracy achieved is 98.67% [1].

Sign language Recognition Using Machine Learning Algorithm by Prof. Radha S. Shirbhate, Mr. Vedant D. Shinde, Ms. Sanam A. Metkari, Ms. Pooja U. Borkar, Ms. Mayuri A. Khandge introduces a project for sign language recognition using techniques such as artificial neural networks, skin segmentation and SVM. The model had achieved an accuracy of 95% [1].

Hand Gesture Recognition and Voice Conversion System for Dumb People by V. Padmanabhan, M. Sornalatha introduced a tool as dumb communication interpreter using flex sensor, accelerometer sensor, gesture recognition (PIC microcontroller), speaker, etc. It fits to every user and accuracy of the model is inflated [4].

Sign Language Recognition Application Systems for Deaf–Mute People: A Review Based on Input-Process-Output by Suharjito, Ricky Anderson, Fanny Wiryana, Meita Chandra Ariesta, Gede Putra Kusuma proposed a paper to find out best method used by researchers for sign recognition considering [5].

3 System Design and Workflow

Flow Chart

See Fig. 1.

A dataset is created by capturing sample number of images through webcam. When the system starts up, it begins collecting gesture images from the previously generated dataset. The images are then preprocessed and cropped so that only the hand gesture is visible. This is accomplished by drawing contours and determining the contour with the greatest area. The contour region is then photographed and resized to 128 × 128 pixels. These images are then divided into three categories: Training, Validation and Testing. After this, the model starts building by adding different layers to the image for the accuracy purpose to all the input images. Then, different layers are added to esteemed model. After adding of the layers, the training of CNN model begins. With the help of CNN, the model builds a network and creates a prediction model which will be used in the prediction of the result. After this, 'validation' and 'testing' of model is performed to check whether model is over-fitting or underfitting. Further steps are taken to increase the accuracy of model later.

4 Experimentation

1. Dataset Creation

We are using OpenCV library to capture images through webcam. The main objective of using OpenCV library is to apply various filters provided by the library. Filters like grayscale conversion and Gaussian blur were applied on the target images. Grayscale

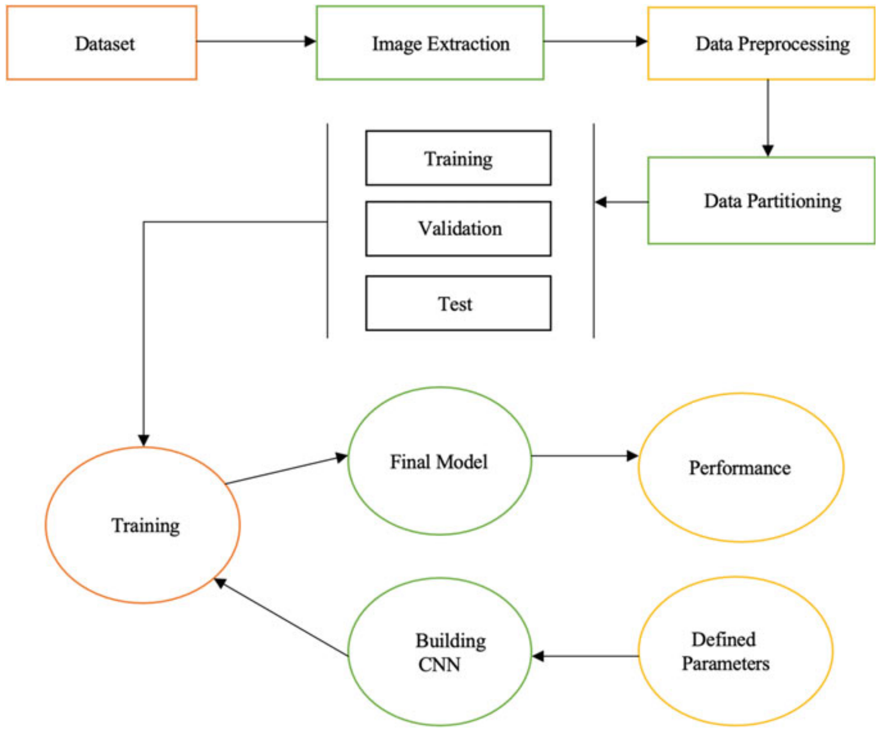


Fig. 1 System design and architecture

conversion simply converts the image into black and white image. The input image is in color which is converted into a grayscale image. Gaussian blur removes some of the noise before further processing the image. We are using the ‘adaptive-threshold’ function to highlight the image borders. The above functions are discussed in detail in further subsections.

We are using Region of interest (ROI) of 300 × 300 pixels for capturing the hand signs. Below is a code snippet of how the gestures in region of interest are captured.

We have resized all the images to 128 × 128 resolution to minimize the training time of the model.

We have created two folders in our dataset, namely Train and Test, both consisting individual folders having 600 images of each symbol.

2. **Image preprocessing**

We are using two types of images to train the dataset namely binary images and canny edge images. Binary image consists of pixels of only two colors, usually black and white. Canny edge image shows all the edges present in the image which increase the amount of data to be processed but in turn helps in getting useful structural information from an image.

For less symbol classes (10 symbols), using binary images to train the dataset yields good results. But for all 26 English alphabets, canny edge data was found to be more useful than binary data images.

For creating a dataset, we implemented a program and captured the live feed through webcam.

These steps are followed while preprocessing dataset:

1. Capture a RGB image (Fig. 2).
2. Convert the RGB image into grayscale and apply Gaussian blur to it (Fig. 3).
3. Applying desired threshold to convert grayscale into binary image. Applying a threshold means selecting some specific value for pixels in image and converting the pixels in image such that (Fig. 4).
4. Cropping and resizing the (Region of Interest) ROI to desired size (Fig. 5).
5. Converting the grayscale image into canny edge using adaptive-threshold filter. Adaptive-threshold method calculates threshold value for smaller regions and helps in finding more features in image (Fig. 6).
6. Cropping and resizing the ROI to desired size (Fig. 7).
7. Saving both images into local disk.

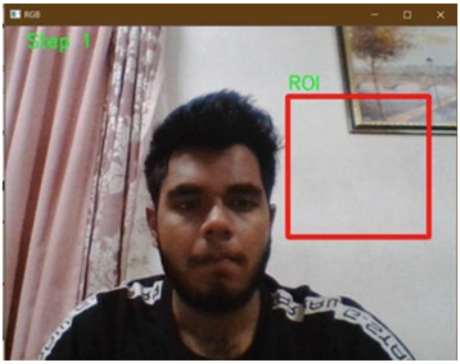


Fig. 2 RCB image

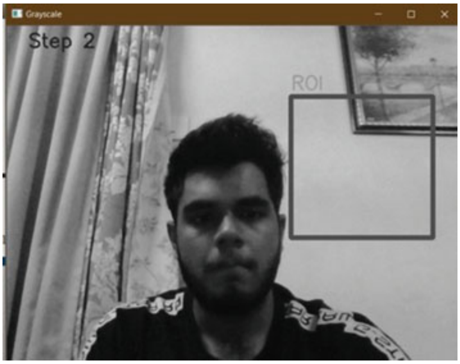


Fig. 3 Convert into grayscale

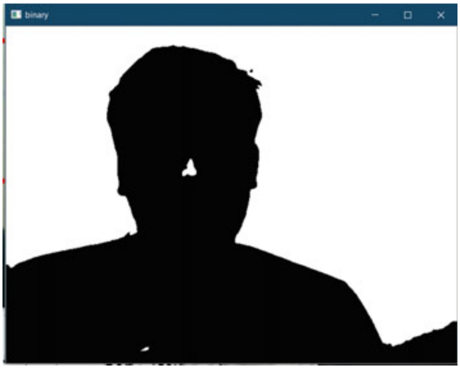


Fig. 4 Convert into binary image

Fig. 5 Cropping and resizing

Fig. 6 Converting grayscale image into canny edge

Fig. 7 Cropped and resized to desired size

3. CNN Model

The model is being trained with CNN.

Why CNN?

1. CNNs are automatic feature extractors.
2. We can use max pooling to reduce size of image without compromising with image quality.
3. CNN provide usage of convolutions, which act as feature emphasizer (Fig. 8).

The OpenGenus IQ [6] team conducted a performance study of various image classifiers in order to determine which was the best classifier for image classification. The findings of an image classification task used to distinguish lymphoblastic leukemia cells from non-lymphoblastic ones have been presented to support their performance review. The features were extracted using a convolutional neural network, and they were fed to different classifiers (Table 1).

Training a CNN

In our model, one convolutional layer which will help in highlighting different features like gesture boundaries in binary images and minute finger features like emerging thumb between ring finger and middle finger. The convolution layer's

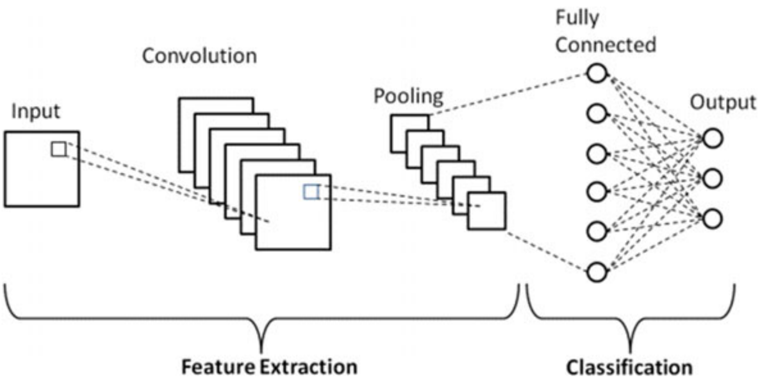


Fig. 8 CNN model flow

Table 1 Comparing of classifiers in terms of accuracy

Classifier	Accuracy (%)	Precision	Recall	ROC
SVM	85.68	0.86	0.87	0.86
Decision Tree	84.61	0.85	0.84	0.82
KNN	86.32	0.86	0.86	0.88
ANN (for 100 epochs)	83.10	0.88	0.87	0.88
CNN (for 300 epochs)	91.11	0.93	0.89	0.97

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 126, 126, 32)      320
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 63, 63, 32)        0
_________________________________________________________________
flatten (Flatten)            (None, 127008)            0
_________________________________________________________________
dense (Dense)                (None, 128)               16257152
_________________________________________________________________
dropout (Dropout)            (None, 128)               0
_________________________________________________________________
dense_1 (Dense)              (None, 96)                12384
_________________________________________________________________
dropout_1 (Dropout)          (None, 96)                0
_________________________________________________________________
dense_2 (Dense)              (None, 64)                6208
_________________________________________________________________
dense_3 (Dense)              (None, 5)                 325
=================================================================
Total params: 16,276,389
Trainable params: 16,276,389
Non-trainable params: 0
```

Fig. 9 Training and overflow of CNN

output is a feature map that highlights the gesture's characteristics. We have added a max pool layer, which helps in reducing the size of input data to learning layers of model while ensuring that the details in image are not lost. This is done in order to reduce the computing time.

Lastly, we have three fully connected layers with 128, 96 and 64 neurons each and an output layer. This layer learns from the input data and produces weights which are then used to classify signs (Fig. 9).

4. **Sign to Speech Conversion**

We have implemented Text-to-Speech Conversion using Pyttsx3, a speech engine which converts text to speech. It is Python library which works offline and converts Text into Speech. The characters which are recognized using CNN model are grouped together to form meaningful words which are then pronounced using Pyttsx3.

5. **Speech to Sign Conversion**

Speech recognition is being used to turn voice commands into text. The text is then passed through a converter function, with the results shown as a series of images on the screen. This is a stand-alone feature with its own graphical user interface.

6. **Word Suggestion Functionality**

Even though we are creating words from recognized characters, it may take some time to spell big words. Also, there is a possibility of incorrect spellings. This is where" Word suggestion and correction" feature comes into play.

We are using US English dictionary to suggest words based on some characters in word. This is done using Hunspell module in python.

7. **Graphical User Interface**

We want a product that allows users to predict sign language in real time. Since the aim is to make it easily accessible without many dependencies, we have created both desktop application and web application.

Desktop Application

This application is created using Tkinter module in Python and has all features of system built in it.

The features include:

- Sign to Text Conversion.
- Sign to Speech Conversion.
- Speech to Sign Conversion.
- Tet to Sign Conversion.
- Word Suggestions.
- Shortcut Mode.
- Emergency Mode.

The images below display 'Sign to Speech/Text Conversion' page and 'Text/Speech to Sign Conversion' page (Figs. 10 and 11).

Web Application

We have used Streamlit to create the Web application. Streamlitis an open-source app framework for creating and deploying data science applications.

The features included in current version of web app are:

- Sign to Text Conversion
- Text to Sign Conversion
- Real time Audio–Video Feed.

5 Results and Discussion

In current CNN model, we are using one convolution layer and one max pooling layer. We are using two datasets which are:

- Binary dataset
- Canny edge dataset.

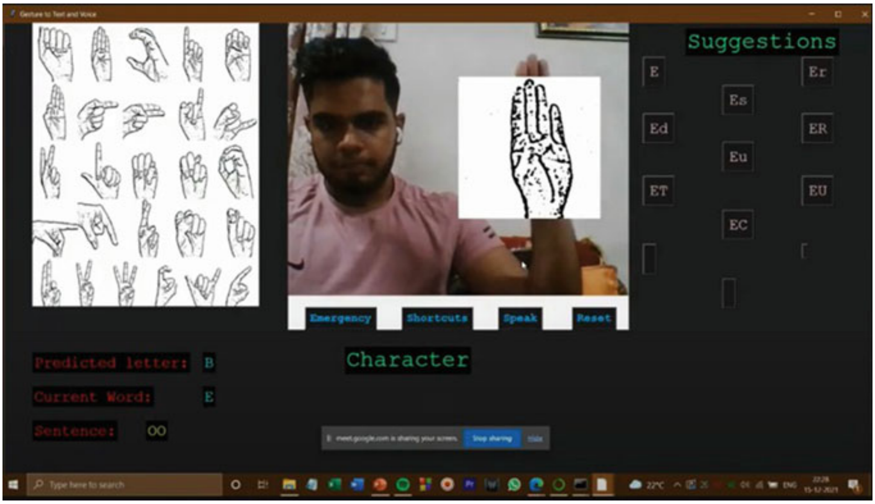


Fig. 10 Sign to speech

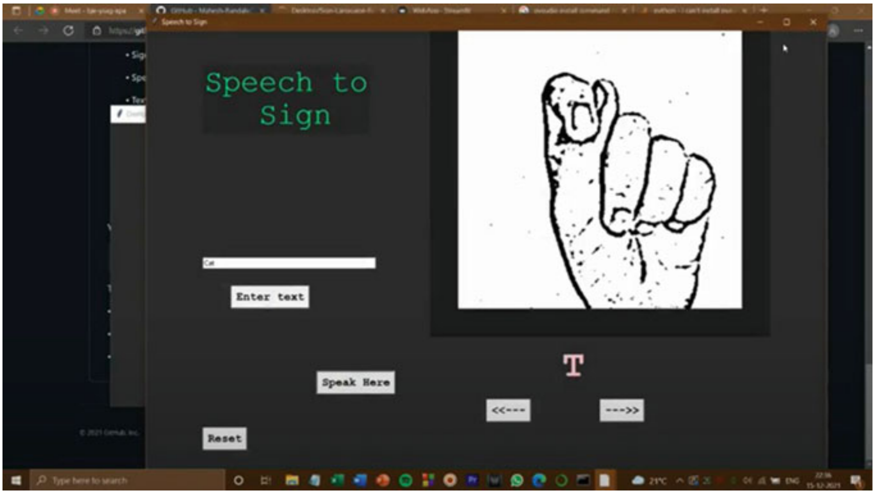


Fig. 11 Speech to sign

We trained the model until it had an accuracy of 99% on training dataset. After the training, during validation of version model, we have achieved:

- 98% accuracy while using canny edge dataset.
- 100% accuracy while using binary dataset.

Improvements in Binary Dataset

While using binary dataset, the model was over-fitting. We stopped the training phase early after a certain number of iterations to prevent over-fitting.

The model was either providing low accuracy of 80% or was over-fitting, even after adding different methods to avoid over-fitting.

Furthermore, the model was unable to differentiate between signs with slight variations, such as (S, M, N) and so on. Binary images were not capable of showing minute differences in different characters.

Hence, we decided to work on canny edge dataset, as it shows a lot of promise for improving the accuracy rate.

Improvements in Canny Edge Dataset

The canny edge dataset had good accuracy on trained data, but 89% accuracy on test data implied that the model needs to be trained with a variety of input images.

We improved the accuracy of the canny edge dataset by using image augmentation, which allows us to access an infinite number of images. By using max pooling, which shows the most common features of existing feature maps, we were able to increase accuracy even further. Following these measures, we were able to achieve a 98% accuracy on the canny edge dataset.

6 Future Scope

1. Introduce hand tracking algorithm in existing system.
2. Extend the model to recognize motion gestures.
3. Deploy the web application with web hosting platforms like Heroku, Streamlit, etc.

7 Conclusion

In this paper, the model project model uses CNN for detection of 26 English ASL alphabets using different image enhancement techniques. This model also uses convolution layer to highlight the dominant features and max pooling layers to reduce computing cost. The current version of model has one convolution, one max pooling and three fully connected layers. These layers may change while in- creasing the accuracy of model. The current version of model has accuracy of 98%. Furthermore, added features provide the means for an effortless two-way communication between the corresponding entities.

References

1. Adewale VA, Olamiti AO Dr (2018) Conversion of sign language to text and speech using machine learning techniques. J Res Rev Sci 5:58–65. https://doi.org/10.36108/jrrslasu/8102/50(0170)
2. Vishwas S, Hemanth Gowda M, Vivek Chandra HN, Tannvi, Ravi Kumar V Dr (2018) Sign language translator using machine learning. Int J Appl Eng Res 13(4) (2018). ISSN 0973-4562
3. Kartik PV, Sumanth KB, Ram VN, Prakash P (2021) Sign language to text conversion using deep learning. In: Ranganathan G, Chen J, Rocha Á (eds) Inventive communication and computational technologies. Lecture notes in networks and systems, vol 145. Springer, Singapore
4. Zhou Z, Chen K, Li X et al (2020) Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays. Nat Electron 3:571–578. https://doi.org/10.1038/s41928-020-0428-6
5. Ojha A, Pandey A, Maurya S, Thakur A, Dayananda P (2020) Sign language to text and speech translation in real time using convolutional neural network. Int J Eng Res Technol (IJERT) NCAIT–2020. 8(15)
6. Paolanti M, Romeo L, Felicetti A, Mancini A, Frontoni E, Loncarski J (2018) Machine learning approach for predictive maintenance in industry 4.0. In: 2018 14th IEEE/ASME international conference on mechatronic and embedded systems and applications (MESA), pp 1–6. https://doi.org/10.1109/MESA.2018.844915

Chapter 3
Efficient Question Answering in Chatbot Using TF-IDF and Cosine Similarity

Rajesh Shrivastava, Abinash Pujahari, Simar Preet Singh, and Tanmay Bhowmik

1 Introduction

An organization uses a chatbot to solve clients' queries. These days, all Web sites and other social media platforms use a chatbot to help their customers. When a user asks a query, a chatbot will look to a file or database for the answer. If the user asks simple queries related to who, what, or predetermine questions, the user gets a correct solution. But if a user asks for a complex query like how to detect symptoms of kidney disease, maybe you get multiple answers or wrong answers. This situation happens because various keywords match in this query. So, the chatbot engine selects the wrong solution from the database file. Finding the correct answer from the available documentation is a real-world problem. Natural language processing (NLP) is used to solve this problem. There are various solutions available for this type of problem, like counting the frequency of keywords and selecting an answer with the highest frequency [1]. Most of the available chatbots are rule based. Once users ask a query, they search in a database as per pre-defined rule [2]. Chatbots follows a pattern to accept user query. There is not a chatbot that can take user queries and search solutions on an available document. The author mainly discusses an algorithm to

R. Shrivastava · S. P. Singh (✉)
Bennett University, Greater Noida, India
e-mail: simarpreet.singh@bennett.edu.in

R. Shrivastava
e-mail: rajesh.shrivastava@bennett.edu.in

A. Pujahari
Department of Computer Science and Engineering, SRM UNIVERSITY-AP, Amaravti, Andhra Pradesh 522502, India
e-mail: abinashpujahari@ieee.org

T. Bhowmik
Pandit Deendayal Energy University, Gandhinagar, Gujarat, India
e-mail: tanmaybhowmik@gmail.com

D. Garg et al. (eds.), *Innovations in Information and Communication Technologies*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-19-3796-5_3

solve this problem. We develop solutions through Google's Collab. To construct a database, we downloaded Wikipedia pages of a kidney disease topic.

Term frequency-inverse document frequency (TF-IDF) mainly helps to calculate the most valuable word from the dataset. Traditional methods calculate the frequency of the word document, but the frequency is always not a good idea. In the question, multiple lines support the keyword, so selecting the correct word is challenging. TF-IDF helps us choose the statistical score based on the usage of expression in the document. The cosine similarity finds the distance vector and selects the most appropriate answer from the datasets. Our contribution to this paper is as follows.

- Develop and design a framework to take user queries and effectively find them using TF-IDF and cosine similarity.
- Compare our result with other available methods in the parameter of accuracy and time.

The experimental results support our hypothesis, and our proposed method achieves 85% accuracy and provides correct answers to a user query.

The rest of the paper is organized as follows: Sect. 2 reviews the recent work in NLP related to search question answers in documents. Section 3 includes the details of the proposed solution and describes different algorithms used therein. Section 4 includes an evaluation of the proposed solution. Finally, Sect. 5 finishes with concluding remarks on the work.

2 Related Work

Chao et al. [1] wrote a review paper to discuss the problem of NLP. In that paper, Chao et al. [3] tried to solve the healthcare patent problem by finding titles from the available patent record. He focuses on the need for an effective method that helps to find duplication of a patent in the database. In that paper, he used TF-IDF for ranking and other distance methods to find the similarity.

Lahitani et al. [4] used cosine similarity to solve essay assessment problems. He applied TF-IDF in the education system to automate essay assessment. This process helps teachers to score the essay. But in that work, if the student writes an article that will not be matched with a model solution, he will not get a good grade. The automation is purely dependent upon checking the content.

Chen et al. [3] presented a YMC model, which is used to capture video features and classify them using TF-IDF and cosine similarity. The author focused on masking the unrelated area from the object detection result in that research.

Rajkumar and Ganapathy [5] developed a chatbot for an E-learning platform. The chatbot helps to solve student queries. But it will give only fixed pattern answers. If the user needs more information or asks a question in a different style, the chatbot cannot answer.

Miklosik et al. [6] proposed a chatbot, but their model has certain limitations. Their model does not search the entire database, and results are not as expected.

Chao et al. [7] proposed a detailed review on chatbots. As per his review, chatbots are used pre-defined data set for answering. To detect intelligent answering, chatbot needs to search the dataset and figure out correct answers. For the same purpose, the most common method is to identify the frequency of the input word into the data set. As per the frequency result, the output is selected. But the false-positive rate is high in this method.

In this paper, we used TF-IDF and cosine similarity. The cosine and TF-IDF methods calculate similarity distance different from earlier approaches. Early researchers widely used frequency-based solutions to select the correct answer.

3 Proposed Method

Figure 1 explains the working model.

The data collection is done through web scrolling using the beautiful soup utility of Python. We build our document by collecting various information related to health care. We collected information about various diseases. In the collected data, a manual

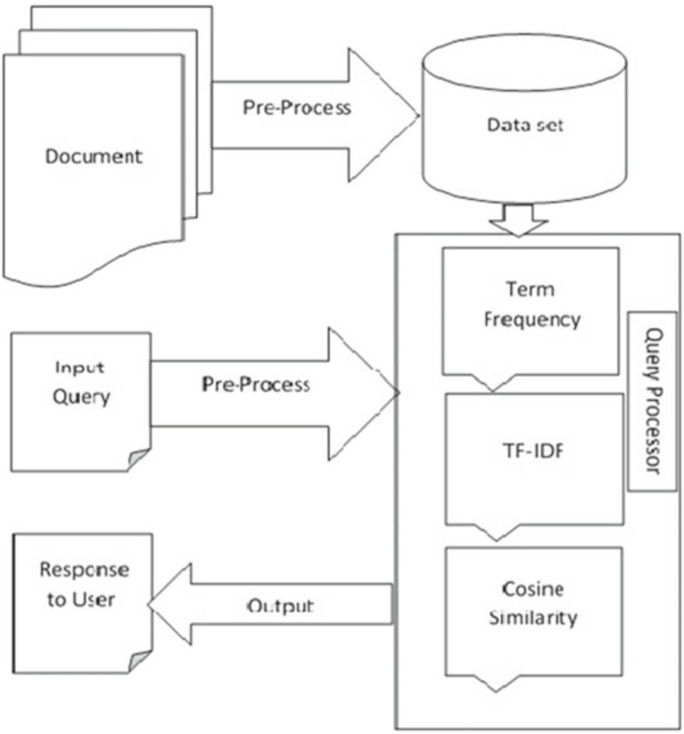


Fig. 1 Chatbot working model

search is not possible. We build an intelligent algorithm that helps us find the correct answer from these datasets. We already perform essential operations on datasets like converting input text into tokens, removing punctuation marks, creating metadata that links all related words together. For example, send, sent, sending are combined. All comments which have similar meanings are clubbed into metadata. We applied the following algorithm to get the most appropriate answer to the input query.

Algorithm 1: Calculate similarity using TF-IDF and cosine

```
Input: Document (D); a dataset, User input query (Q)
Output: Reply to input query
Remove the punctuations and stop words from the input query.
while each word in Q do
    TF(Q, D) = Number of times Q appear in dataset/Total tokens in dataset
    IDF(Q, D) = 1 + log(D/Q)
    Calculate TF-IDf(Q, D) vector;
end
result = Calculate cosine(Q, D);
if result == 0 then
    Print ("Sorry, No result found")
    return false;
end
else
    Print "Hi Solution of your query is "+ result
end
return true;
```

Algorithm 1 explains the method to search the answer of user queries from the datasets. The data set is already stored in the token format with metadata. Once a user asks a question to the chatbot, the query will search in the datasets. TF calculates the number of times a word occurs in a document.

$$\text{TF}(Q, D) \leftarrow \frac{\text{Word from the input query appear in the document}}{\text{total word in the document}}.$$

After calculating term frequency, the algorithm searches for the relevance of a document concerning the user input query. Since TF will not differentiate in all frequencies, IDF adds weights to the more relevant answer. IDF uses a log which helps us to solve this problem.

$$\text{IDF}(Q) \leftarrow \frac{1 + \text{loge(Total Number of Documents)}}{\text{(Number of Documents with term in input user query } (Q))}$$

Then, we calculate TF * IDF. It will give a calculative score which will help identify a correct document.

To find similarity, we derive a vector from each document. Now, we have a vector space containing each element's vector. Using the cosine similarity, we can find the mapping between any documents.

$$\text{Cosine}(Q, D) \leftarrow \frac{\text{Dot\_product}(Q, D)}{||Q|| * ||D||}$$

where Dot_product(Q, D) is $Q * D \leftarrow |Q| * |D| * \cos(\Theta)$

- $|Q|$ = magnitude of vector Q
- $|D|$ = magnitude of vector D
- Θ is the angle between Q and D.

The cosine similarity matches the correct answer by comparing two virtual lines in a vector space. If two points are matched, the angle between the two vectors is closer. Otherwise, the angles are wide. The value of cosine is in the range of −1 to 1. The smaller angle is near to 1, and the wider angle 180 °C is close to −1.

4 Result

We experiment with ten different documents and calculate the TF-IDF and cosine similarity score. Table 1 displays the similarity score. Our result is matched with document number 8.

Figure 2 discusses the impact achieved by TF-IDF and cosine similarity. Document 8 contains the correct answer to the user query. Its score is more similar to query document Q. If we compare our result with the existing frequency-based

Table 1 Similarity score

S. No.	Document	Cosine	Degree
1	d10	0.7071	45
2	d7	0.766	40
3	d6	0.788	38
4	d3	0.848	32
5	d9	0.866	30
6	d1	0.88	28
7	d4	0.906	25
8	d5	0.906	25
9	d2	0.92	23
10	d8	0.939	20
11	Q	0.933	21

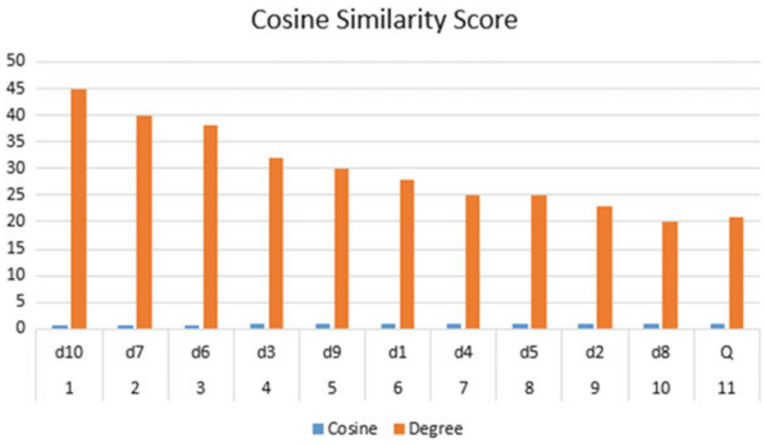


Fig. 2 Cosine similarity

solution, then our method achieves better accuracy. We received almost 85% of precision during the experiment, whereas the other available methods got up to 78% of accuracy.

5 Conclusion

We develop an interactive chatbot that will search all user queries related to health care and respond to the user. In this process, TF-IDF and cosine similarity help to process effectively. In chatbot processing, TF-IDF and cosine similarity help search for a correct answer. This method helps the query processing engine to find the accurate result in a faster time. Researchers will also try to reduce search time with better accuracy in the future. We can also experiment with other methods to improve the accuracy.

References

1. Chao MH, Trappey AJ, Wu CT (2021) Chatbots: a review and trend forecast using intelligent ontology extraction and patent analytics. In: Emerging technologies of natural language-enabled complexity
2. Thorat SA, Jadhav V (2020) A review on implementation issues of rule-based chatbot systems. In: Proceedings of the international conference on innovative computing and communications (ICICC), Apr 2020
3. Chen TY, Chiu YC, Bi N, Tsai RTH (2021) Multi-modal chatbot in intelligent manufacturing. IEEE Access

4. Lahitani AR, Permanasari AE, Setiawan NA (2016) Cosine similarity to determine similarity measure: study case in online essay assessment. In: 2016 4th international conference on cyber and IT service management, Apr 2016. IEEE, pp 1–6
5. Rajkumar R, Ganapathy V (2020) Bio-inspiring learning style chatbot inventory using brain computing interface to increase the efficiency of e-learning. IEEE Access 8:67377–67395
6. Miklosik A, Evans N, Qureshi AMA (2021) The use of chatbots in digital business transformation: a systematic literature review. IEEE Access
7. Chao MH, Trappey AJ, Wu CT (2021) Emerging technologies of natural language-enabled chatbots: a review and trend forecast using intelligent ontology extraction and patent analytics. Complexity 2021

Chapter 4
Power Allocation Scheme Based on DRL for CF Massive MIMO Network with UAV

Prakhar Consul, Ishan Budhiraja, Deepak Garg, and Abhay Bindle

1 Introduction

Unmanned aerial vehicles (UAV) grab a lot of attention these days because of their energy efficient models that can be used to perform various critical tasks with ease and automation [1]. These UAVs can bring excellent results if integrated with advanced cellular networks like the cell-free (CF) massive MIMO system. These networks are highly complex and dense in nature. Moreover, in big cities, high-rise buildings and other obstacles degrade their performance. UAV-assisted on-fly communication can enhance the quality of services by means of expanding ground coverage. There are more users equipments (UEs) to experience line of sight communication (LoS) which improves channel gain [2, 3]. Power optimisation among the various resources in the network is a critical issue, since the inter-user interference and gain of the channel also depend upon the power allocation schemes.

In an asymptotic sense, massive MIMO systems are the best-known solution for interference free communication in the last few years. But due to pilot contamination problems and unpredicted user activity, which reduced the performance of the network. Although massive MIMO system increases the energy efficiency and spectral efficiency [4] but not to the desire extend as required in the future communication networks. So, there needs to be some architectural changes in the current design of communication systems that can work more efficiently compared to massive MIMO

P. Consul (✉)
School of Computer Science Engineering, Bennett University, Greater Noida, Uttar Pradesh, India
e-mail: e21soep0004@bennett.edu.in

I. Budhiraja · D. Garg
Bennett University, Greater Noida, Uttar Pradesh, India
e-mail: ishan.budhiraja@bennett.edu.in

A. Bindle
Maharishi Makendeshwar University, Ambala, India

D. Garg et al. (eds.), *Innovations in Information and Communication Technologies*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-19-3796-5_4

systems [5]. In a cell-free massive MIMO system, a large number of access points (APs) is distributed in the coverage area, which are connected to a CPU through a backhaul network unlike a massive MIMO architecture where ground users served by a particular AP in a particular cell [6].

Reaping all the benefits of a massive MIMO system, APs in a CF massive MIMO system can serve or jointly serve each UE with the same time and frequency resources [7]. As there are a large number of APs distributed in a particular geographical area, there are very few chances of shadow fading. Also, problems like cell edge users and handoffs are completely eliminated. Furthermore, in this paper, a UAV-assisted CF massive MIMO system based on deep deterministic policy gradient (DDPG) has been proposed which minutely optimises the power and also improves the overall quality of service (QoS).

1.1 Related Work

In this section, the preexisting deep learning (DL)-based methods are discussed. However, many existing power allocation schemes show some good results [8–10], but their computational complexity is very high. For example, the weight minimum mean square error (WMMSE) involves complex mathematical operations like matrix inversions and large numbers of iterations. Recently, some deep reinforcement learning (DRL)-based algorithms based on data-driven approaches have been proposed [11]. Reference [12] proposed an ensembling deep neural network-based power control technique that maximises the user sum rate of a multi-user fading channel. Various other works, deep neural network (DNN), deep Q learning (DQL), artificial neural networks (ANNs), were used for power optimisation [13–15]. Also, various aspects of UAV assisted communication network has been explored in the last decade. DRL-based algorithms perform the same task without involving any complex mathematical calculations.

1.2 Motivation

Motivated by the above-mentioned survey, we study UAV-assisted CF massive MIMO communication for power optimisation among various resources in a network. UAV-assisted wireless communication is quite a new concept, and not much work has been done in this direction. The concept of establishing an aerial base station via UAV is quite exciting. Moreover, the practical application of UAVs in such areas as precision agriculture, aerial photography, and search and rescue will make it beneficial for future generations. The performance of the proposed system has been analysed by considering uplink data transmission, propagation model, and uplink channel estimation.

1.3 Contributions

We have first derived an expression for the propagation model, uplink data transmission, and uplink channel estimation for a UAV assisted CF massive MIMO system. From our study, it is found that the ratio of uplink and downlink transmission affects the spectral efficiency (SE). The proposed system model is basically converted into a machine learning model through Markov's decision process (MDP). DDPG reinforcement learning-based actor critic model has been designed to make the whole process automatic. Our results show the UAV-assisted cell-free MIMO system performs better than the normal CF massive system.

1.4 Organisation

The organisation of the paper is as follows: in Sect. 2, we designed a system model in terms of propagation model, uplink channel estimation, uplink data transmission of cell-free massive MIMO. In Sect. 3, we proposed a solution based on the DDPG actor critic model. Also, a model of DDPG with UAV is being evaluated. In Sect. 4, specific simulation results are given. The last section, Sect. 5 summarises the whole paper in terms of conclusion.

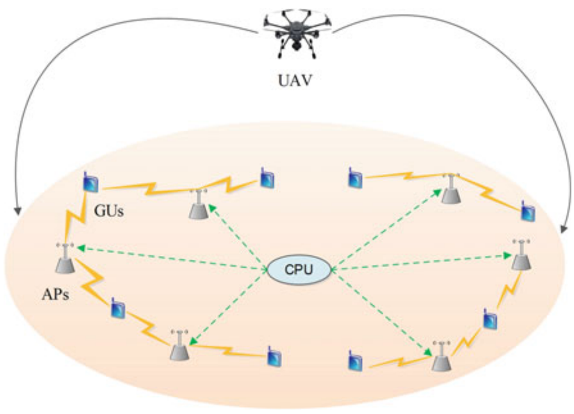
2 System Model

Considered a cell-free massive MIMO system (CFM-MIMO)-assisted via UAV as shown in Fig. 1. The system includes a UAV, various APs, and various ground users (GUs) over a geographical area of $A \times A$. A single antenna is equipped via each AP and GU. Let $\mathcal{I} = \{1, 2, \ldots, i, \ldots, I\}$ and $\mathcal{J} = \{1, 2, \ldots, j, \ldots, J\}$ are the sets of APs and GUs, respectively. The channel gain is denoted via h_{ik} between ith AP and kth GU. It is defined as $h_{ik} = \sqrt{\alpha_{ik}} f_{ik}$. $\sqrt{\alpha_{ik}}$ and $f_{ik} \sim \mathcal{CN}(0, 1)$ denotes the large-scale and small-scale fading for the AP-GU link, respectively. We consider the uplink scenario, estimation of channel and data uplink data transmission. We consider the uplink scenario that comprises the phases of estimation of channel, pilot transmission and data uplink data transmission.

2.1 Estimation of Channel and Pilot Transmission

In the pilot transmission phase, all GUs provide pilot data to the APs to facilitate channel estimation. During this phase, GUs send their pilot sequences of length t

Fig. 1 System architecture



symbols to the APs at the same time. The message signal received at ith AP is given as follows:

$$\mathbb{Y}_{p,i} = \sqrt{t\sigma_p}\sum_{j=1}^{J} h_{ik}\theta_j + W_{p,i}, \tag{1}$$

where $\sqrt{t\sigma_p} \in \mathbb{C}^{t\times 1}$ represents the assigned pilot sequence to the ith GU with $||\theta_j||^2 = 1$, $W_{p,i} \in \mathbb{C}^{t\times 1}$ the additive noise for ith AP having i.i.d $\sim \mathcal{CN}(0, 1)$ elements. The ith AP estimates the channel gain h_{ik} $\forall j$, which is determined by the obtained pilot signal $\mathbb{Y}_{p,i}$. The projection of $\mathbb{Y}_{p,i}$ onto θ_j^H is denoted by $\overline{Y}_{p,ij} = \theta_j^H \mathbb{Y}_{p,i}$. Therefore,

$$\overline{Y}_{p,ij} = \sqrt{t\sigma_p}\left(h_{ik} + \sum_{j\neq j'}^{J} h_{ik'}\theta_j^H\theta_{j'}\right) + \theta_j^H W_{p,i}. \tag{2}$$

Given $\overline{Y}_{p,ij}$, the MMSE estimation of h_{ik} is formulated as follows:

$$\hat{h}_{ij} = \frac{\mathbb{E}\{\overline{Y}^*_{p,ij}h_{ik}\}}{\mathbb{E}\{|\overline{Y}_{p,ij}|^2\}} = c_{ij}\overline{Y}_{p,ij}, \tag{3}$$

where c_{ij} is defined as follows:

$$c_{ij} = \frac{\sqrt{t\sigma_p}\alpha_{i,j}}{t\sigma_p\sum_{j'=1}^{J}\alpha_{i,j'}|\theta_j^H\theta_{j'}|^2} \tag{4}$$

2.2 Uplink Data Transmission

During the uplink, the data of all GUs send to the APs at the same time. Let $x_j = \sigma q_j s_j$ represents the transmitted signal from the jth GU, where s_j be the transmit symbol having $\mathbb{E}\{|s_j|^2\} = 1$. σ, q_j denote the normalised uplink SNR and power control coefficient of jth GU. The range of q_j is 0–1. Now, from all users, the received signal at the ith AP is as follows:

$$\mathbb{Y}_i = \sum_{j=1}^{J} h_{ij} x_j + W_i = \sqrt{\sigma} \sum_{j=1}^{K} \sqrt{q_j} s_j + W_i, \tag{5}$$

where W_m represents the additive noise at the ith AP. The scaled received signals are then transferred to the CPU for joint detection after match filtering is performed at each AP using the locally acquired channel estimate $\hat{h_{ij}}$. To detect s_j, the CPU uses the aggregated received signal r_j. The large-scale fading α_{ij} is assumed to be known. Signal r_j is given as follows:

$$r_j = \sqrt{\sigma} \sum_{j'=1}^{J} \sum_{i=1}^{I} \hat{h}_{ij}^* h_{ij'} \sqrt{q_{j'} s_{j'}} + \sum_{i=1}^{J} \hat{h}_{ij}^* W_i \tag{6}$$

Let $\beta_{ij} = \mathbb{E}\{|\hat{h_{ij}}|^2\} = \sqrt{t\sigma} \alpha_{i,j} c_{ij}$. Also, $\mathbb{DS}$, $\mathbb{BU}$, and $\mathbb{UI}$ denotes the desired signal, the beamforming gain uncertainty, and interference produced by the jth GU, respectively. According to [5], for the jth user, the uplink rate can be derived as:

$$\mathbb{R}_{\beth}^{\mathbb{U}} = \log_2 \left(1 + \frac{|\mathbb{DS}_j|^2}{\mathbb{E}\{|\mathbb{BU}_{\beth}|^2\} + \sum_{j' \neq j}^{J} \mathbb{E}\{|\mathbb{UI}_{\beth}|^2\} + \sum_{i=1}^{I} \beta_{ij}} \right), \tag{7}$$

where

$$|\mathbb{DS}_j|^2 = q_j \left(\sum_{i=1}^{I} \beta_{ij} \right)^2 |\theta_j^H \theta_j|^2, \tag{8}$$

$$\mathbb{E}\{|\mathbb{BU}_{\beth}|^2\} = \sum_{J' \neq j}^{J} q_{j'} \left(\sum_{i=1}^{I} \beta_{ij} \frac{\alpha_{ij'}}{\alpha_{i,j}} \right)^2, \tag{9}$$

$$\mathbb{E}|\mathbb{UI}_{\beth}|^2 = q_{j'} \sum_{i=1}^{I} \beta_{ij} \alpha_{i,j'} \tag{10}$$

3 Proposed Solution

3.1 Markov Decision Process

In the model, at a centralised location the information is processed in a centralised manner (e.g. at access point). The UAV's mobility consumes the most energy. As a result, in order to enhance the operational period, the UAV is assumed to stay in a fixed place at the centre of the cluster. The access point is considered as an agent (AP) for enhance data rate at GUs. With the help of MDP, the optimisation problem can be defined as:

$$\text{MDP} = (\mathtt{S}, \mathtt{A}, \mathtt{P}, \mathtt{R}, \mathbb{\Gamma}) . \tag{11}$$

From the model, the centralised optimisation can be described as:

State space: To obtain good data rate of GUS, the agent should be aware of local information such as beamforming gain uncertainty and interference produced by the jth GU. The state space can be defined as:

$$\mathtt{S} = \left[(\mathbb{BU}_1, \ldots, \mathbb{BU}_j, \ldots \mathbb{BU}_J); (\mathbb{UI}_1, \ldots, \mathbb{UI}_j, \ldots \mathbb{UI}_J)\right] . \tag{12}$$

Action space: To improve the data rate at the base station in the MIMO system is the main aim. So, action space can be represented as:

$$\mathtt{A} = \left[(q_1, q_2, \ldots, q_j, \ldots, q_J)\right]. \tag{13}$$

At the state s^{t-1}, agent performs the action $a^{t-1} = (q_1, q_2, \ldots, q_j, \ldots, q_J)$. The agent moves towards the next state s^t after performing action a^{t-1}.

Reward function: To maximise the data rate fairness of all GUs, the reward function (RF) is expressed as:

$$\mathbb{RF} = \sum_{i=1}^{I} \sum_{j=1}^{J} \mathbb{R}^{\mathbb{U}}{}_J. \tag{14}$$

3.2 Centralised Optimisation

A DRL scheme such as DDPG is employed to achieve the best optimal policy. The DDPG is works as an hybrid model that has two networks: the actor network and the critic network. Actor networks rely on value functions. Critic networks, on the other hand, rely on policy search. To improve convergence speed and reduce calculations, the techniques used by DDPG algorithm are replay buffer and target network. The executed transition $\left(s^{t-1}, a^{t-1}, r^{t-1}, s^t\right)$ is saved in a limited or restricted memory

Algorithm 1 Data rate fairness maximisation for APs using centralised optimisation technique

Input

- Environment: (a) APs (b) GUs (c) UAV
- $q_i \geq q_i^{\min}$: APs Min requirement
- $q_j \geq D_j^{\min}$: GUs Min requirement

Initialisation:

- $\mathbb{Z}(s, a, \zeta_x)$ = With variable ζ_x the critic network.
- $\nu(s; \zeta_\nu)$ = With variable ζ_ν the Actor network.
- $\mathbb{Z}'(s, a, \zeta_x')$ = With variable $\zeta_{x'}$ the target critic network.
- $\nu'(s; \zeta_{\nu'})$ = With variable $\zeta_{\nu'}$ the Target actor network.
- C = Replay Exp.

Output: α, β

```
for episode = 1, …, Θ do
    Action process starts
    the starting state s^0 obtained
    for iteration = 1, …, 𝒯 do
        The achieved action a^{t−1} executed at state s^{t−1}
        In accordance with (14) the modified reward is r^{t−1}
        Noticify the upcoming state s^t
        In the replay buffer (s^{t−1}, a^{t−1}, r^{t−1}, s^t) saved as transition
        from C a mini-batch of E transitions (s^i, a^i, r^i, s^{i+1}) sample randomly
          Using loss function in (15) update critic parameter descent by stochastic gradient
          The actor policy parameter updated
          Target critic network parameters updated
          (ζ_{x'}, ζ_{ν'}) according to (18) and (19)
          State updated s^{t−1} = s^t
      end for
end for
```

of capacity C using an experienced replay buffer. A little batch U of transitions is selected randomly from collected samples. For updating new samples and rejecting old ones, a finite size memory is used. To calculate the target value for both the critic and the actor networks, target networks are used.

Let the critic network represented as $\mathbb{Z}(s, a; \zeta_x)$ with variable ζ_x and the target critic network represented as $\mathbb{Z}'(s, a; \zeta_{x'})$ with variable $\zeta_{x'}$. Similarly, the actor network is represented as $\nu(s, a; \zeta_\nu)$ with the variable ζ_ν and the target actor network is represented $\nu'(s, a; \zeta_{\nu'})$ with variable $\zeta_{\nu'}$. To examine the actor and critic network over a little batch of U samples, stochastic gradient descent (SGD) is used. Now the critic is updated by minimising as:

$$MN = \frac{1}{U} \sum_{i}^{P} \left(y^i - \mathbb{Z}(s^i, a^i, \zeta_x)\right)^2, \tag{15}$$

with the target

$$y^i = r^i(s^i, a^i) + \Gamma \mathbb{Z}(s^{i+1}, a^{i+1}; \zeta_{x'}|_a^{i+1} = \nu'(s^{i+1}; \zeta_{\nu'}). \tag{16}$$

Algorithm 2 Using centralised optimisation technique for D2D users throughput maximzation

Input

- Environment: (a) DMPs (b) CUEs (c) NOMA-integrated BS.
- $D_i^n \geq D_i^{n,\min}$: CUE min requirement
- $D_j^n \geq D_j^{n,\min}$: DPs min requirement

Initialisation:

- $\mathbb{Z}(s, a, \zeta_x)$ = With variable ζ_x the critic network.
- $\nu(s; \zeta_\nu)$ = With variable ζ_ν the Actor network.
- $\mathbb{Z}'(s, a, \zeta_x')$ = With variable $\zeta_{x'}$ the target critic network.
- $\nu'(s; \zeta_{\nu'})$ = With variable $\zeta_{\nu'}$ the target actor network.
- C = Experience replay

Output: α, β

```
for episode = 1, ..., Θ do
    Action process starts
    the starting state s^0 obtained
    for iteration = 1, ..., 𝒯 do
        The achieved action a^{t-1} executed at state s^{t-1}
        In accordance with (14) the modified reward is r^{t-1}
        Notify the upcoming state s^t
        In the replay buffer (s^{t-1}, a^{t-1}, r^{t-1}, s^t) saved as transition
        from C a mini-batch of E transitions (s^i, a^i, r^i, s^{i+1}) sample randomly
            Using loss function in (15), update critic parameter descent by stochastic gradient
            The actor policy parameter updated
            Target critic network parameters updated
            (ζ_{x'}, ζ_{ν'}) according to (18) and (19)
            State updated s^{t-1} = s^t
    end for
end for
```

The updated parameter of actor network as follows:

$$\vec{\nabla}_{\zeta_\nu} K \approx \frac{1}{U} \sum_{i}^{U} \vec{\nabla}_{a^i} \mathbb{Z}(s^i, a^i; \zeta_x)|_{a^i=\nu(s^i)} \vec{\nabla}_{\zeta_\nu} \nu(s^i; \zeta_\nu). \tag{17}$$

To update the ζ_x and the ζ_ν', soft target updates are used as follows:

$$\zeta_{x'} \leftarrowtail \chi \zeta_x + (1-\chi)\zeta_{x'} \tag{18}$$

$$\zeta_{\nu'} \leftarrowtail \chi \zeta_\nu + (1-\chi)\zeta_{\nu'}, \tag{19}$$

where χ is in the range between 0 and 1.
Where Θ represents the number of maximum episodes, and $\mathscr{T}$ denoted time step in the proposed algorithm.

4 Evaluation of the Performance

The evaluation of the performance of the scheme is discussed in this section. It has two sections: (i) numerical approach and (ii) results and discussion.

4.1 Simulation Parameters

Consider a circular area of 5 km^2 comprised of 30 APs and 15 GUs. The rest of the simulated parameters is derived from [12], and listed in Table 1.

Table 1 Simulation parameters

Parameters	Values
Bandwidth	20 MHz
Carrier frequency	1.9 GHz
UAV transmission power	4 W
Number of UAVS	$U = 1$
UAV's coverage	400 m
Path-loss parameter	2.1
Thermal noise	−1.74 dBm
Discounting factor	0.85
Initial batch size	$E = 30$
Activation function	ReLu
Noise figure	9 dB
User association	Cell-free
Traffic model	Full buffer
Optimiser	Adam
Software used	Python 3.8 and Tensorflow 1.13.1

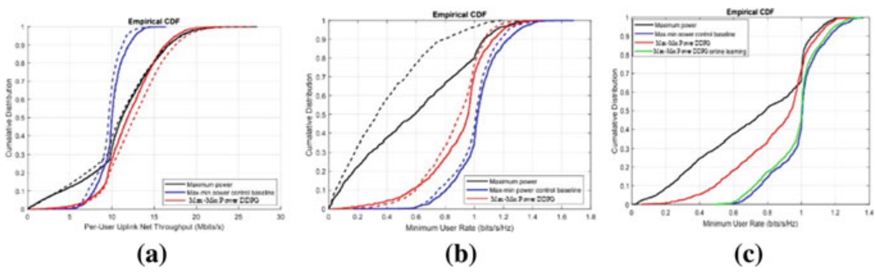


Fig. 2 Comparative analysis **a** CDF versus per-user net throughput. **b** CDF versus minimum user rate under fixed APs. **c** CDF versus minimum user rate under fixed APs and movable users

The relationship between the cumulative distribution function (CDF) and per-user net throughput is shown in Fig. 2a. In the lower net throughput range, the graph suggests that the DDPG performs similarly to the baseline. In both network setups, the baseline and DDPG have nearly identical 97% expected per-user net throughput. The disparity between the baseline and DNN performance, on the other hand, implies that the DDPG may have learned a power allocation strategy that is not ideal.

Figure 2b represents the graph between the CDF and minimum user rate. The graph depicts how effectively the DDPG has met the desired goal of maximum–minimum user fairness. As compared to the ideal baseline solution, the DDPG has a lower minimum user rate and it outperforms the maximum power scenario significantly.

The CDF is impacted by the minimal user rate under the conditions of fixed APs and moveable GUs, as shown in Fig. 2c. According to the graph, DDPG implementations achieve near-optimal minimum user rate performance with online training, outperforming the random AP/user scenario. The claimed performance was obtained with only a training set of 10,000 samples and an offline model training time of less than 10 min. As it has the low complexity offline training and the online processing and comparable performance, the proposed DDPG algorithm has the potential for rapid and easy deployment in a realistic environment.

5 Conclusion

In the paper, the target is to enhance the net throughput while minimising the interference and maintaining the QoS of the ground users. To achieve the target, we used the CFM-MIMO system with an UAV assist for the power optimisation of the network. Here, we optimise the power of the APs along with the UAV to improve the usernet throughput. Numerical results represents that the proposed scheme gives good result in comparison with the baseline schemes.

References

1. Shen C, Chang T-H, Gong J, Zeng Y, Zhang R (2020) Multi-UAV interference coordination via joint trajectory and power control. IEEE Trans Signal Process 68:843–858
2. Björnson E, Sanguinetti L (2019) Making cell-free massive MIMO competitive with MMSE processing and centralized implementation. IEEE Trans Wireless Commun 19(1):77–90
3. Özdogan Ö, Björnson E, Zhang J (2019) Performance of cell-free massive MIMO with Rician fading and phase shifts. IEEE Trans Wireless Commun 18(11):5299–5315
4. Zhang J, Björnson E, Matthaiou M, Ng DWK, Yang H, Love DJ (2020) Prospective multiple antenna technologies for beyond 5G. IEEE J Sel Areas Commun 38(8):1637–1660
5. Ngo HQ, Ashikhmin A, Yang H, Larsson EG, Marzetta TL (2015) Cell-free massive MIMO: uniformly great service for everyone. In: IEEE 16th international workshop on signal processing advances in wireless communications (SPAWC), Aug 2015, pp 201–205
6. Ngo HQ, Ashikhmin A, Yang H, Larsson (2017) Cell-free massive MIMO versus small cells. IEEE Trans Wireless Commun 16(3):1834–1850

7. Zhang J, Chen S, Lin Y, Zheng J, Ai B, Hanzo L (2019) Cell-free massive MIMO: a new next-generation paradigm. IEEE Access 7:99878–99888
8. Zhao Y, Niemegeers IG, De Groot SH (2020) Power allocation in cell-free massive MIMO: a deep learning method. IEEE Access 8:87185–87200
9. Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. Neural Netw 2(5):359–366
10. Sanguinetti L, Zappone A, Debbah M (2018) Deep learning power allocation in massive MIMO. In: 52nd Asilomar conference on signals, systems, and computers. IEEE, pp 1257–1261
11. Budhiraja I, Kumar N, Tyagi S, Pham Q-V, Tanwar S (2020) Energy efficient mode selection scheme for wireless powered D2D communications with NOMA underlaying UAV. In: IEEE conference on computer communications workshops (INFOCOM WKSHPS), pp 877–882
12. Van Chien T, Nguyen Canh T, Björnson E, Larsson EG (2020) Power control in cellular massive MIMO with varying user activity: a deep learning solution. IEEE Trans Wireless Commun 19(9):5732–5748
13. Budhiraja I, Kumar N, Tyagi S (2021) Deep-reinforcement-learning-based proportional fair scheduling control scheme for underlay D2D communication. IEEE Internet Things J 8(5):3143–3156
14. Budhiraja I, Kumar N, Alazab M, Tyagi S, Tanwar S, Srivastava G (2020) Energy management scheme for wireless powered D2D users with NOMA underlaying full duplex UAV. In: Proceedings of the 2nd ACM MobiCom workshop on drone assisted wireless communications for 5G and beyond, pp 7–12
15. Consul P (2015) Triple band gap coupled microstrip U-slotted patch antenna using L-slot DGS for wireless applications. In: 2015 communication, control and intelligent systems (CCIS), pp 31–34

Chapter 5
Machine Learning Techniques for Predicting Dengue Outbreak

M. B. Anuranjan, Chandana Divya Vani, Chanpreet Singh, Samarjeet Barman, Kuldeep Chaurasia, and P. V. Arun

1 Introduction

Dengue is an exponentially emerging commonly prone viral disease in many parts of the world. Dengue flourishes in urban poor areas, suburbs and the countryside but also affects more affluent neighbourhoods in tropical and subtropical countries. Apart from causing a severe flu-like illness, sometimes it is prone to cause a lethal complication called severe dengue. According to World Health Organization (WHO), 50–100 million infections are now estimated to occur annually [1]. The Aedes aegypti mosquito transmits the viruses that cause dengue. Because dengue is carried by mosquitoes, the transmission of dengue is related to meteorological and environmental data such as temperature, precipitation and vegetation. A vaccine to prevent dengue is licenced and available in some countries. Surprisingly, the vaccine manufacturer has announced in 2017 that people who have not been previously infected and received the vaccine may be at risk of developing severe dengue if they get infected following the vaccination [2]. The backbone for the treatment of dengue is supportive

M. B. Anuranjan (✉)
Rajalakshmi Engineering College, Chennai, Tamil Nadu, India
e-mail: anuranjan.mb.2018.cse@rajalakshmi.edu.in

C. Divya Vani
JNTUK, Kakinada, Andhra Pradesh, India

C. Singh
Amrita University, Kollam, Kerala, India

S. Barman
IIIT Manipur, Heingang, Manipur, India

K. Chaurasia
Bennett University, Greater Noida, India

P. V. Arun
Indian Institute of Information Technology, Sri City, Chittoor, India

D. Garg et al. (eds.), *Innovations in Information and Communication Technologies*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-19-3796-5_5

care [3]. It becomes vital to recognize the three phases of dengue [4]: febrile, critical and recovery. Hence, serious efforts are required to control and prevent this disease. This makes predictions on dengue outbreaks very important. With the help of this prediction, health departments across the world can take preventive measures to combat dengue fever before the outbreak begins, saving millions of lives. The objective of the research work is to find the best algorithm to predict the spread of Dengue in the area under study.

2 Related Work

Several approaches have been used to predict dengue outbreaks. Rachata et al. [5] worked on the use of artificial neural networks with an entropy method to build a predictive model for dengue outbreaks in Thailand. A study in Sri Lanka, using past patterns of weather and past dengue cases as inputs, introduced an artificial neural network (ANN) to predict the dengue outbreak in the Kandy district [6]. Cheong et al. [7] showed that land-use factors other than human settlements, like different types of agricultural land, water bodies and forest, can be associated with reported dengue cases in the state of Selangor, Malaysia, and used boosted regression to account for nonlinearities and interactions between these factors. Dharmawardana et al. [8] proposed that the mobility of humans has a significant effect on the outset of dengue to the immunological dengue 'naive' region. The data set was derived using mobile network big data in Sri Lanka. The predictions were made using ANN and XGBoost. Recent work of Muhilthini et al. [9] proposed a dengue possibility forecasting model; the data set contains information about several dengue cases observed every week for several years in any country. It contains details about the conditions of weather like precipitation amount, temperature, humidity and so on using GBR to find the dependencies in the given training data set and predict the amount of dengue cases for the given week and year of a country in the test data set. Tate et al. [10] proposed a model for the prediction of dengue, diabetes and swine flu using random forest. The main aim of this model is to predict the disease by using the symptoms taken from patients, and to recommend a specialized doctor, from this, the risky cases of that particular disease in a week of that particular area was also calculated. Ong et al. [11] used random forest regression to predict the risk rank of dengue transmission in 1 km grids, with dengue, population, entomological and environment data in Singapore. More than 80% of the observed risk ranks fell within the 80% prediction interval.

3 Study Area and Data Resources

The two cities considered for this experiment are San Juan, Puerto Rico, and Iquitos, Peru. San Juan is the capital city of Puerto Rico and also happens to be the most populous municipality. Iquitos is the capital city of the Mayan Province and Loreto region and is claimed to be the ninth most populous city of Peru. (Fig. 1). The data used in this study were collected by various U.S Federal Government Agencies (including the Centers for Disease Control and Prevention, the National Oceanic and Atmospheric Administration and the U.S. Department of Commerce). The consolidated data set was acquired from the "Data Download" section of the openAI competition "DengAI: Predicting Disease Spread" hosted by drivendata.org. The data set was a combination of meteorological data including the features. Some of the important features are city abbreviations: sj for San Juan and iq for Iquitos, week_start_date, maximum temperature, minimum temperature, average temperature, total precipitation, mean air temperature, pixel southeast of city centroid, pixel northeast of city centroid etc.

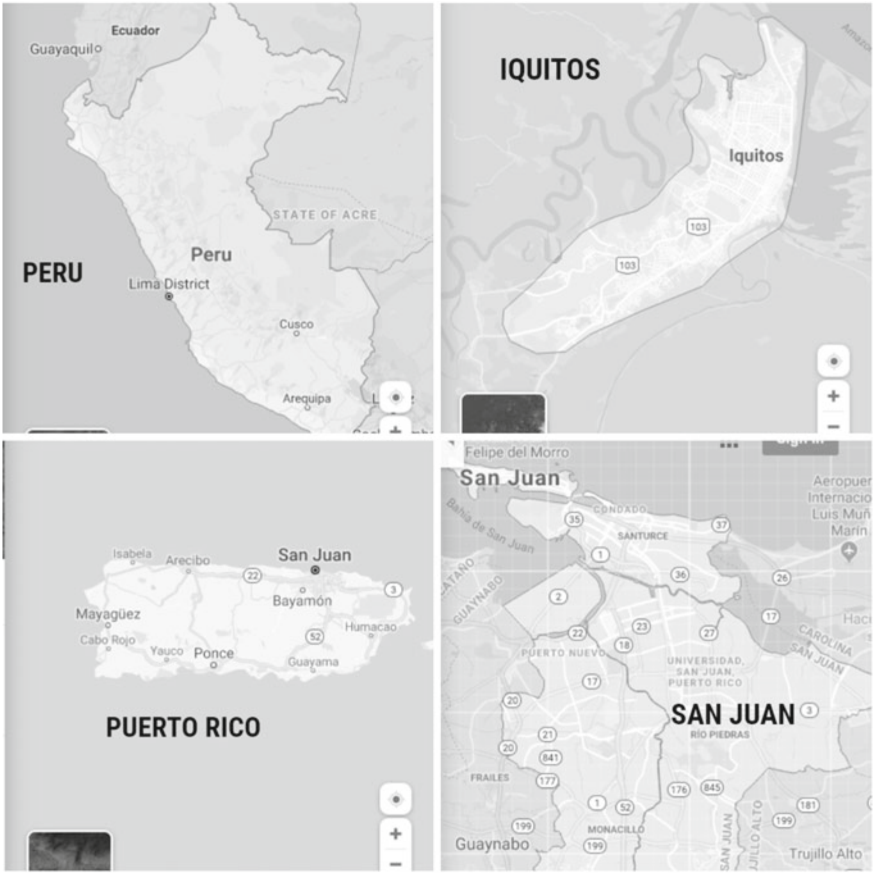


Fig. 1 Study area [13]

4 Methodology

The methodology involved seven steps, namely acquiring the data set, preprocessing the data, selecting a relevant predictive model, building the model, fitting and training the model with the training data, tuning the parameters for optimal performance and predicting the values (Fig. 2).

4.1 Dataset Preprocessing

The acquired data had to be preprocessed for maximum performance of the models. Various data preprocessing techniques were used in this study. To begin with, the data set had to be divided based on the two cities. A new feature called "month"

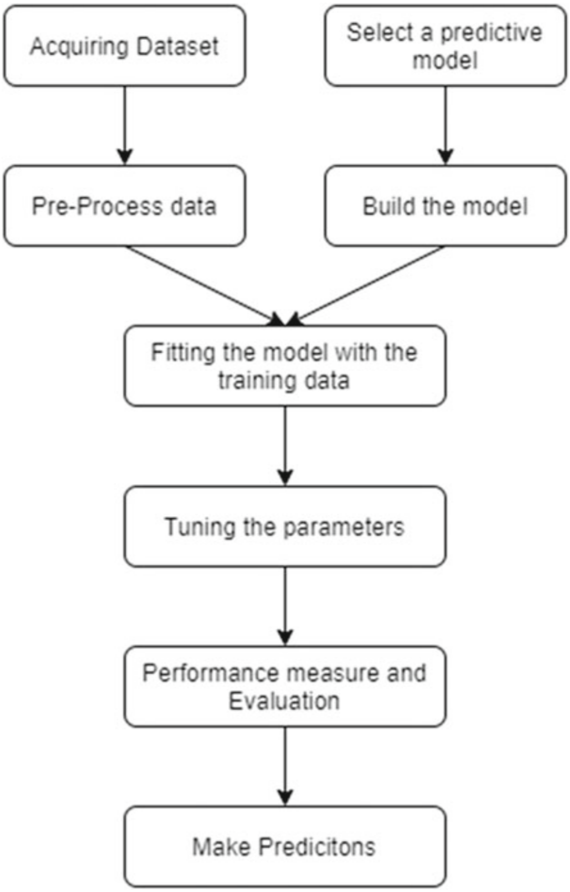


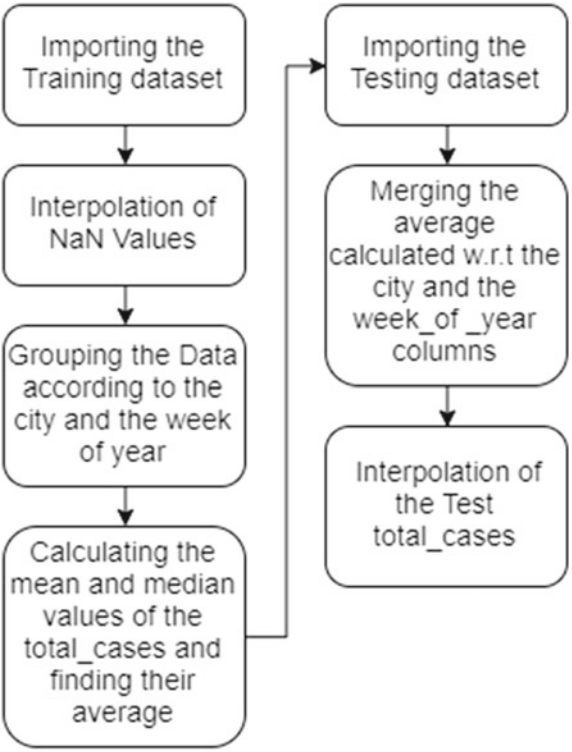
Fig. 2 General methodology applied to make various predictive models

had to be extracted from the "week_start_date" since the latter is a little harder to work with. The Not a Number (NaN) values were replaced with the mean of their respective columns. It can be noticed that a few features associated with temperature were recorded in Kelvin and a few others in Celsius. The features were all converted to Celsius.

4.2 Interpolation

Interpolation is the process of acquiring a function from a set of data points such that the function passes through all the given data points and can be used to appraise data points in-between the given ones. Using interpolation (Fig. 3) and the weekly pattern of the recorded cases, it is possible to make a plausible prediction. To attain a weekly pattern, the data from the training set can be grouped by the week_of_year column. The mean and median of this grouped data are averaged to get an appreciable week pattern. The pattern is then later used to predict the total_cases reported in the testing data. Interpolation of the predicted total_cases is done for accurate results.

Fig. 3 Algorithm used to make the interpolation model



4.3 Time Series Forecasting with Random Forest

Random forest is a subset of the supervised learning algorithms which uses an ensemble learning method. It can be used for both classification and regression. The trees in random forests are made to run in parallel meaning there is not any interaction between these trees. A random forest combines many decision trees. The differences lie in the number of features that can be split at each node and the randomness added by random selection of sample data from the original data by the decision trees to avoid over-fitting. A time series is a vector of values that are indexed by time. Sometimes, it is necessary to perform some preprocessing to make it comprehensible (Fig. 4).

This algorithm involved the prediction of total cases using time series forecasting with random forest. The data set after being imported was preprocessed. The data preprocessing involved replacing the NaN values using the forward and backward fill method.

It was necessary to split the data set based on the cities for accurate predictions. The next step involved introducing time series based lag which served as the base for the time series forecasting of the model. A random forest regressor with the necessary parameters was built. The next step involved the fitting of the data set to the model. Two parameters were tuned for better performance. The two parameters considered are n_estimators and max_depth. After trying out various values, the above-tabulated values (Table 1) were considered for the cities as using these parameters gave the

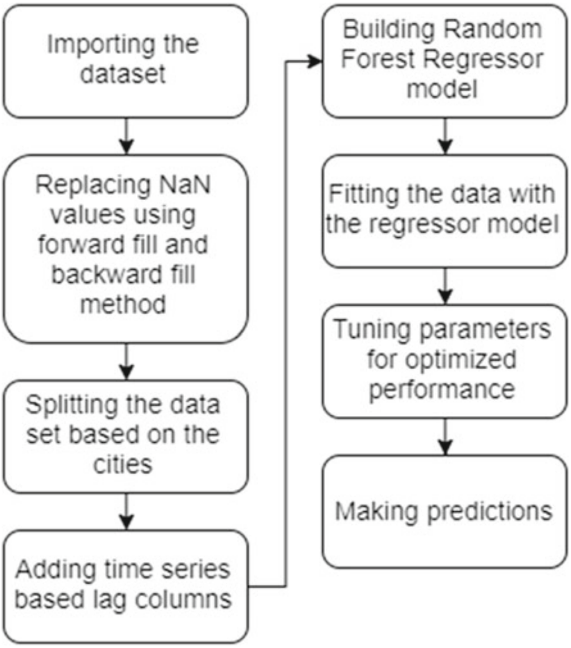


Fig. 4 Algorithm used with time series prediction with random forest regression

Table 1 Details of various parameters tuned in the random forest model

Parameter tuning			
Parameter	Values	San Juan	Iquitos
n estimator	10,20,50,100,200,300,500	20	100
max depth	10,15,20,40,50,100	10	10

best result following which the model has trained again with the tuned parameters and used to predict the total_cases. The mean absolute error computed (using a time series-based algorithm) for both the cities was better than the MAE calculated using the normal random forest algorithm.

4.4 Gradient Boosting for Time Series Prediction

Gradient boosting is one of the most powerful techniques for building predictive models. It is a machine learning technique widely used for regression and classification problems. Gradient boosting aims at producing an ensemble of weak prediction models. The algorithm repetitively leverages the patterns in residuals and bolster a model with weak prediction to improve it. This algorithm involved the prediction of total_cases using time series forecasting with gradient boosting. The data set after being imported needed to be preprocessed. The data preprocessing involved replacing the NaN values using the forward and backward fill method. It was necessary to split the data set based on the cities for more accurate predictions. The next step involved introducing time series-based lag which served as the base for the time series forecasting of the model. A gradient boosting regressor with the necessary parameters was built (Fig. 5).

The next step involved the fitting of the data set to the model. Four parameters were tuned for better performance (Table 2).

4.5 Performance Measures

The performance of both models has been assessed by computing the mean absolute error (MAE). The mean absolute error (MAE) [12] measures the closeness of forecasts predictions to the actual outcomes. It is given by (Eq. 1):

$$\text{MAE} = \frac{1}{n}\sum_{1}^{n}\left|y_i - y_{\hat{i}}\right| = \frac{1}{n}\sum_{1}^{n}|e_i|. \tag{1}$$

$$\text{AE} = |e_i| = \left|y_i - y_{\hat{i}}\right| \tag{2}$$

where actual $= y_i$ and predicted $= y_{\hat{i}}$.

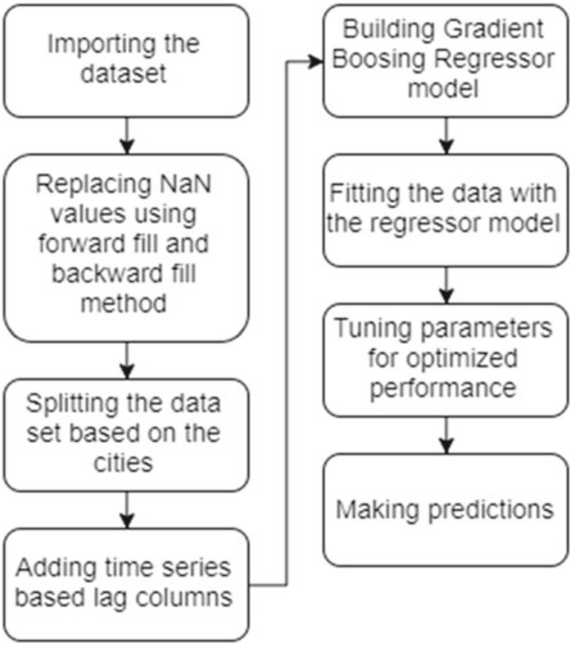


Fig. 5 Algorithm used in implementing the gradient boosting regression model

Table 2 Various parameters tuned in the gradient boosting model

Parameter tuning			
Parameter	Values	San Juan	Iquitos
max iter	10,20,50,100,200,300,500	50	10
max leaf nodes	10,16,32,64,128,256	10	16
max depth	4,8,16,32	8	10
max bins	10,16,32,64,128,256	10	32

5 Results and Discussion

Time series-based prediction is a good alternative for ordinary predictive models. The three algorithms used were manipulation using interpolation, times series-based random forest and gradient boosting. Although these algorithms perform well, the time series-based gradient boosting algorithm seems to have the best of results.

The MAE of this algorithm is better than those of others. The respective MAE values have been tabulated (Table 3).

The random forest regressor seems to work the best in this particular problem. It can be observed from the graph (Fig. 6) that the predicted values were almost the same as the reported cases in most of the instances. Although the values might not

Table 3 Comparison of results of various predictive model

Comparison of MAE using different algorithms		
Algorithms	San Juan	Iquitos
Random forest	26.66	6.7
Gradient boosting	24.11	7.36
Interpolation	25.98	25.98

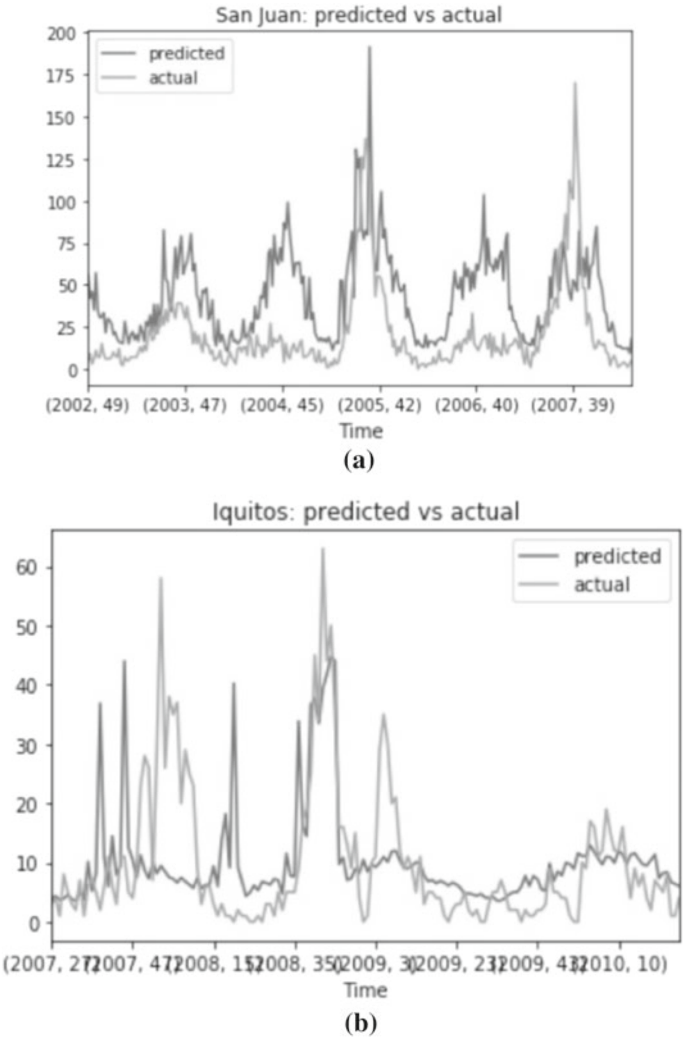


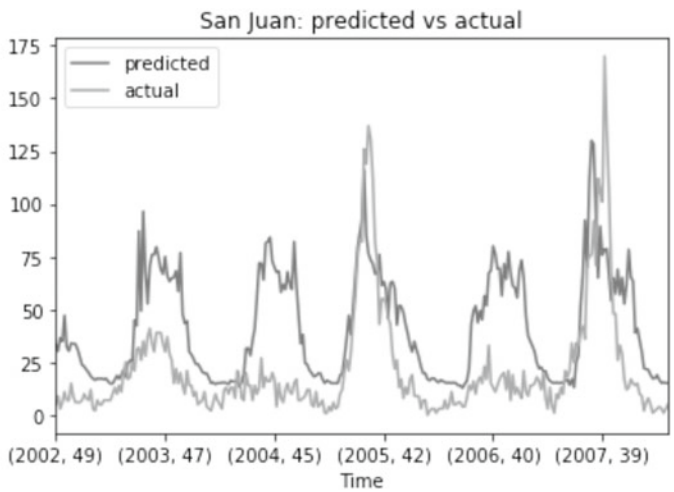
Fig. 6 Plotting result using random forest: **a** San Juan region, **b** Iquitos region

be very close, the regressor had done a good job in predicting and understanding the pattern involved.

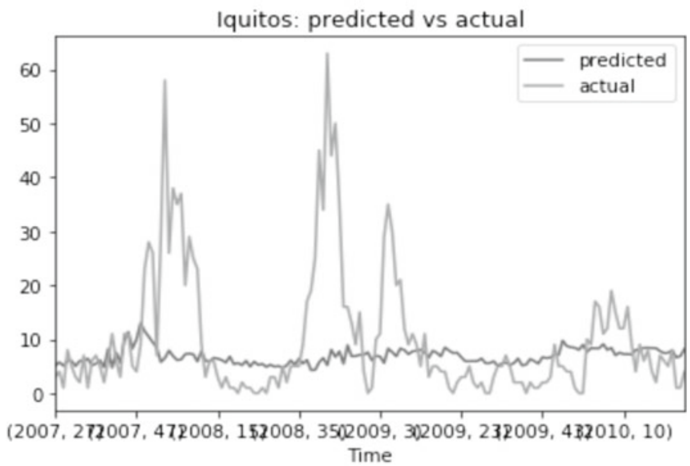
The gradient boosting algorithm does not seem to have worked the best in this situation (as shown in Fig. 7). The algorithm did a good job of predicting the number of reported cases in Iquitos.

The algorithms might have made better results with a deep analysis of the data set and more fine-tuning of the model.

Overall, the gradient boosting algorithm seems to have worked the best among the other algorithms used. The other algorithm used involves a random forest regressor



(a)



(b)

Fig. 7 Plotting result using gradient boosting: **a** San Juan region, **b** Iquitos region and interpolation technique which seems to be providing an acceptable result. The following limitations made the process of achieving better results harder and also undermined the scope of using deep neural networks to make predictions: data set given data is limited or less, two cities are located in different geographical areas, and dengue occurs at particular months of the year. This result can further be developed by using models that are known to work better for time series-based data like the ARIMA model.

6 Conclusion

This study considered three different modelling techniques, namely interpolation, gradient boosting regression and random forest regression, to predict the amount of dengue cases reported in two cities. Parameters were tuned and adjusted for the most optimal performance. Comparisons of results were made based on mean absolute error (MAE). The performance was analysed, and the result points out that the gradient boosting regression performs significantly better than the other models and is therefore considered to be a better approach. Future results can be improved by obtaining large amounts of meaningful data and implementing better models associated with time series predicting.

Acknowledgements The authors would like to thank Bennett University for supporting the project and providing all necessary facilities. The authors would also like to thank the following U.S Federal Government Agencies for providing the data—Centers for Disease Control and Prevention (CDC), National Oceanic and Atmospheric Administration and the U.S. Department of Commerce. The research problem was proposed through a series of competitions held by drivendata.org. The authors would like to acknowledge drivendata.org for this opportunity.

References

1. Safitri MD, Yusniar H (2019) Association between environmental factors and the presence of mosquito larvae to dengue hemorrhagic fever (DHF) in Karimunbesar Island, Indonesia. Int J Health, Educ Soc (IJHES) 2(12):18–25
2. Halstead SB, Deen J (2002) The future of dengue vaccines. The Lancet 360(9341):1243–1245
3. Wallace D, Canouet V, Garbes P, Wartel TA (2013). Challenges in the clinical development of a dengue vaccine. Curr Opin Virol 3(3):352–356
4. Gubler DJ (1998) Dengue and dengue hemorrhagic fever. Clin Microbiol Rev 11(3):480–496
5. Rachata N, Charoenkwan P, Yooyativong T, Chamnongthai K, Lursinsap C, Higuchi K (2008) Automatic prediction system of dengue haemorrhagic-fever outbreak risk by using entropy and artificial neural network. Iscit 210–214
6. Herath PHMN, Perera AAI, Wijekoon HP (2014) Prediction of dengue outbreaks in Sri Lanka using artificial neural networks. Int J Comput Appl 101(15):1–5
7. Cheong YL, Leitão PJ, Tobia L (2014) Assessment of land use factors associated with dengue cases in Malaysia using boosted regression trees. Spatial Spatio-temporal Epidemiol 10(2014):75–84

8. Dharmawardana KGS, Lokuge JN, Dassanayake PSB, Sirisena ML, Fernando ML, Perera AS, Lokanathan S (2017) Predictive model for the dengue incidences in Sri Lanka using mobile network big data. In: 2017 IEEE international conference on industrial and information systems (ICIIS), pp 1–6. IEEE
9. Muhilthini P, Meenakshi BS, Lekha SL, Santhanalakshmi ST (2018) Dengue possibility forecasting model using machine learning algorithms. Int Res J Eng Technol 5(3):1661–1665
10. Tate A, Gavhane V, Pawar J, Rajpurohit B, Deshmwch GB (2017) Prediction of dengue diabetes and swine flu using a random forest classification algorithm. Int RJ Eng Tech 4(2017):685–690
11. Ong J, Xu L, Rajarethinam J, Kok SY, Liang S, Tang CS, Cook AR, Ng LC, Yap G (2018) Mapping dengue risk in Singapore using random forest. PLoS Neglected Trop Diseas 12(6):e0006587
12. Willmott CJ, Matsuura K (2005) Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Res 30(1):79–82
13. Google maps. [Online] Available: https://www.google.com/maps

Chapter 6
Human Activity Recognition Using Wearable and Inertial Sensors

Ajit Singh Rawat, Aranaya Ratra, Shaleen Govil, Neha Gupta, and Shikha Rastogi

1 Introduction

Our daily lives have become more reliant on mobile gadgets. This is due to an increase in the sophistication of mobile device development, as well as the integration of high-quality sensors, high processing power, massive storage capacity, and constant connectivity. People constantly engage with their low-cost, small-sized cell phones in their everyday activities, which have resulted in a rise in research into extracting information from data gathered by ubiquitous sensors in mobile devices [1, 2]. Lifelogs may be used to track simple physical activities like walking, running, sitting, and so on, as well as more complex activities like eating, working, and exercising [3, 4]. This has a wide range of applications in domains like medicine [5], augmented reality [6], computer–human interaction [7], text classification [8], and sound classification [9].

The most challenging part of creating a comprehensive lifelog is gathering activity data from various wearable sensors [10] and properly categorizing human activity based on the data. Sensors such as accelerometers, gyroscopes, GPS, magnetometers, proximity sensors, and barometers [3] are used in mobile devices. Researchers can easily collect sensor data with minimal infrastructure by using smartphones. Based on the data collected, modern machine learning algorithms may be used to detect and recognize human behavior based on the data obtained. A basic smartphone might assist in preserving detailed records of a user's everyday activities. The utilization of a range of sensors, as well as advances in deep learning and feature selection approaches, has the potential to push the frontiers of human activity detection to deeper ontological levels.

A. Singh Rawat · A. Ratra · S. Govil · N. Gupta (✉) · S. Rastogi
Bharati Vidyapeeth's College of Engineering, Paschim Vihar, New Delhi 110063, India
e-mail: coedelhi@bharatividyapeeth.edu; neha.gupta@bharatividyapeeth.edu

D. Garg et al. (eds.), *Innovations in Information and Communication Technologies*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-19-3796-5_6

Human activity recognition (HAR) is a great example of current smartphones' versatility [11]. Aside from smartphones, any digital gadget equipped with specialized sensors, such as tablets, MP3 players, or smartwatches, may recognize human movements and gestures and respond with a set of tasks [12]. The readings from two specific sensors, the accelerometer and the gyroscope, are usually considered in activity signal-based HAR.

We have studied and analyzed various research papers based on human activity recognition. Post completion of the literature survey, we became aware of the various methods that can be applied to recognize human activities. We were also enlightened by the applicability of HAR in various fields.

In [1], Jovanov et al. discussed wearable health monitoring systems that are integrated into a telemedicine system and can help with early diagnosis of abnormal states and prevention of harmful effects. The joint coordinates of the three-dimensional skeletons received from the depth sensor were used by Yalcin et al. [2] to automatically recognize human actions using CNN and LSTM deep learning algorithms. In [3], Wang et al. introduced a deep learning-based approach for healthcare applications that can distinguish between individual actions as well as transitions between two different short-duration, low-frequency activities. In [13], Li et al. used a variety of machine learning techniques to investigate the impact of different algorithms and multi-features on time series. Body activity recognition is formulated as a classification issue using data acquired by wearable sensors by Yani Guan et al. in [10].

The paper is organized as follows: Sect. 2 covers the proposed methodology, Sect. 3 illustrates the datasets used, and Sect. 4 presents performance evaluation of the models and discussion about it. Lastly, Sect. 5 concludes the paper.

2 Proposed Methodology

First, using accelerometer data from a mobile device worn around a person's waist, we trained a deep neural network (DNN) to distinguish the type of movement (walking, running, jogging, etc.). For the same, we used the WISDM data set [14].

1. Data from accelerometer is loaded from the WISDM dataset.
2. Convert accelerometer data to a time-sliced representation and reformat it.
3. Create a visual representation of the accelerometer data.
4. Restructure the multidimensional tabular data such that Keras can accept it.
5. Split up the data set into three components, i.e., training, validation, and testing.
6. For data on human activity recognition data, train the deep neural network.
7. Via the learning curve and confusion matrix, validate the trained neural network's performance against the test data (Fig. 1).

We discovered that employing more complex neural network architectures, such as long short-term memory, has a lot of promise for improving the model. Using the

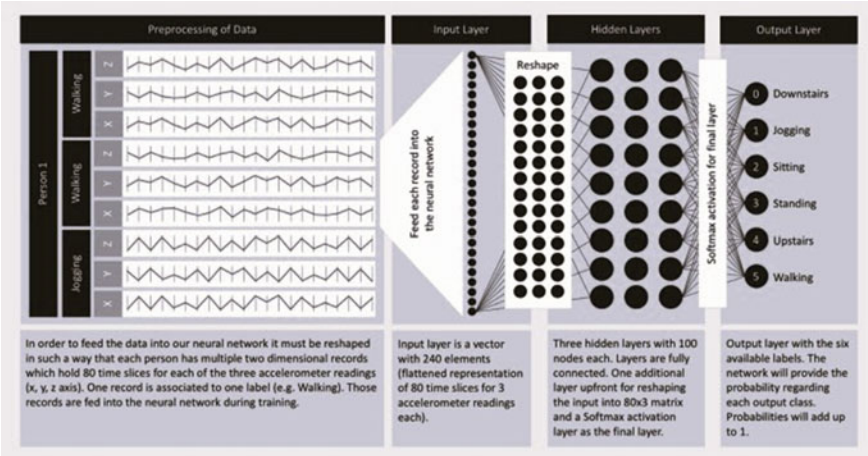


Fig. 1 DNN example

same dataset, we trained an LSTM neural network for human activity recognition [15], which was coded in TensorFlow.

As we progressed, we learned that human activities differ from one individual to the next and are difficult to divide into a finite number of categories. As a result, additional data from various actions was required for improved HAR research. This was the driving force for our decision to use the KU-HAR dataset [16]. We intended to combine more sophisticated samples with the ones we already had.

We have built two models using KU-HAR dataset:

1. Decision trees
2. Random forests (Fig. 2)

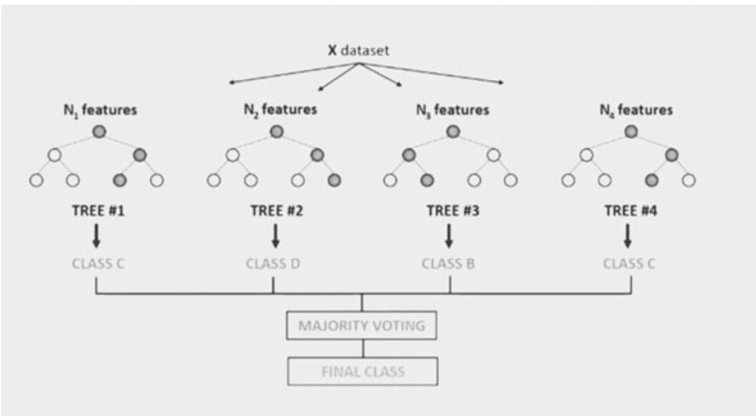


Fig. 2 Random forests

3 Dataset Description

We have used the WISDM dataset from UCI machine learning repository [14]. It is a compilation of accelerometer data collected from roughly 30 people's smartphones while they were doing six different activities (walking, walking upstairs, walking downstairs, sitting, standing, and laying). The x-, y-, and z-axis accelerations for each action were measured and recorded with a timestamp and a person ID (Fig. 3).

To expand the scope of our dataset in terms of number and complexity of activities, we have used the KU-HAR dataset [16]. It is an open dataset for recognizing different types of human activity. This dataset provides data on 18 distinct activities that were obtained using smartphone sensors from 90 people. ML-based HAR models will be able to learn from these examples as a result and become more reliable in detecting actions (Fig. 4).

	user-id	activity	timestamp	x-axis	y-axis	z-axis
0	33	Jogging	49105962326000	-0.7	12.7	0.5
1	33	Jogging	49106062271000	5.0	11.3	1.0
2	33	Jogging	49106112167000	4.9	10.9	-0.1
3	33	Jogging	49106222305000	-0.6	18.5	3.0
4	33	Jogging	49106332290000	-1.2	12.1	7.2
5	33	Jogging	49106442306000	1.4	-2.5	-6.5

Fig. 3 WISDM dataset sample

Fig. 4 KU-HAR dataset sample

Column no.	Information
1 – 300	Accelerometer X-axis readings
301 – 600	Accelerometer Y-axis readings
601 – 900	Accelerometer Z-axis readings
901 – 1200	Gyroscope X-axis readings
1201 – 1500	Gyroscope Y-axis readings
1501– 1800	Gyroscope Z-axis readings
1801	Class ID
1802	Length of each axial data
1803	Serial no. of the subsample

4 Results and Discussion

We have made two models based on the WISDM data set (DNN and LSTM) and two models based on the KU-HAR data set (decision tree and random forest). We have split the data sets for model training and model testing purposes (Table 1).

4.1 Based on WISDM Dataset

In the DNN model to segregate training and validation data for model training, we have utilized an 80:20 split. Figures 5 and 6, respectively, show the confusion matrix and learning curve of the model.

The LSTM model contains 2 fully connected and 64 units of 2 LSTM layers which are stacked on each other. We have trained our model for 50 epochs and measured accuracy and error. Figures 7 and 8, respectively, show the confusion matrix and learning curve of the model.

Table 1 Hyperparameters used

Hyperparameter	DNN model	LSTM model
Batch size	400	1024
No. of epochs	50	50
No. of hidden layers	3	2

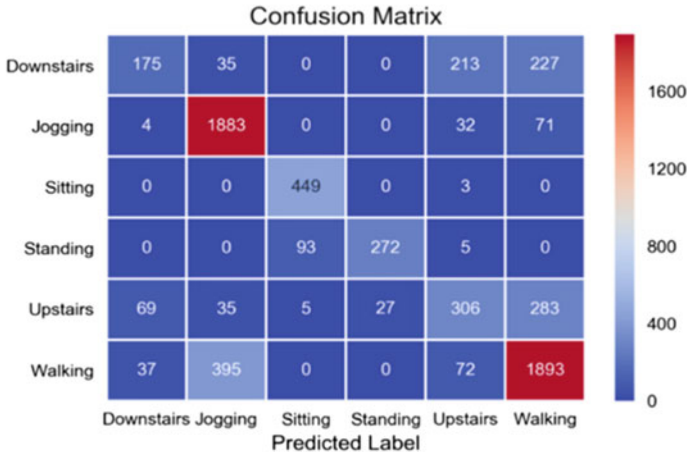


Fig. 5 DNN confusion matrix

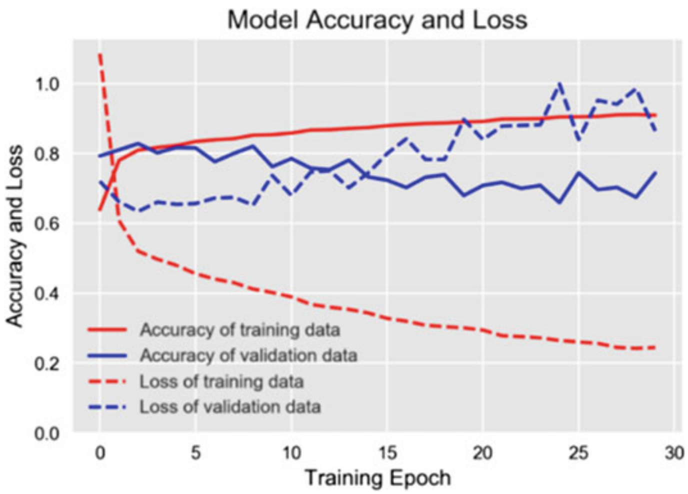


Fig. 6 DNN loss curve

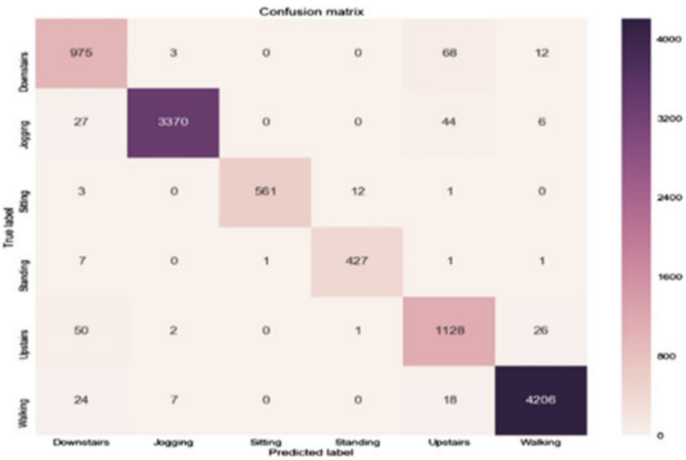
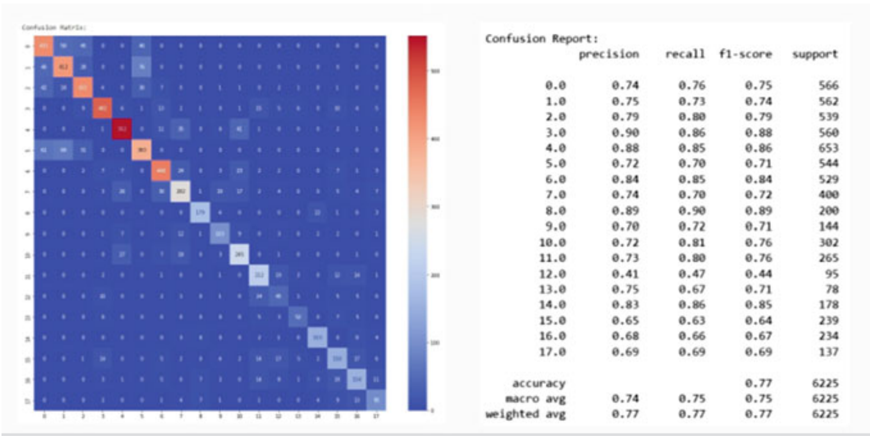


Fig. 7 LSTM confusion matrix

4.2 Based on KU-HAR Dataset

We have used the Scikit learn library of machine learning to build the models. We have utilized an 70:30 train–test split. Decision tree is a type of supervised learning technique where features of the dataset are represented by internal nodes and outcome is represented by leaf node [4]. Figure 9 shows the confusion matrix and confusion report of the model.

Fig. 8 LSTM loss curve



	precision	recall	f1-score	support
0.0	0.74	0.76	0.75	566
1.0	0.75	0.73	0.74	562
2.0	0.79	0.80	0.79	539
3.0	0.90	0.86	0.88	560
4.0	0.88	0.85	0.86	653
5.0	0.72	0.70	0.71	544
6.0	0.84	0.85	0.84	529
7.0	0.74	0.70	0.72	400
8.0	0.89	0.90	0.89	200
9.0	0.70	0.72	0.71	144
10.0	0.72	0.81	0.76	302
11.0	0.73	0.80	0.76	265
12.0	0.41	0.47	0.44	95
13.0	0.75	0.67	0.71	78
14.0	0.83	0.86	0.85	178
15.0	0.65	0.63	0.64	239
16.0	0.68	0.66	0.67	234
17.0	0.69	0.69	0.69	137
accuracy			0.77	6225
macro avg	0.74	0.75	0.75	6225
weighted avg	0.77	0.77	0.77	6225

Fig. 9 Decision tress confusion matrix and confusion report

Random forests belong to the supervised learning technique. It is built on the technique of ensemble learning, which combines multiple classifiers to solve a complex problem and improve the performance of the model. Figure 10 shows the confusion matrix and confusion report of the model.

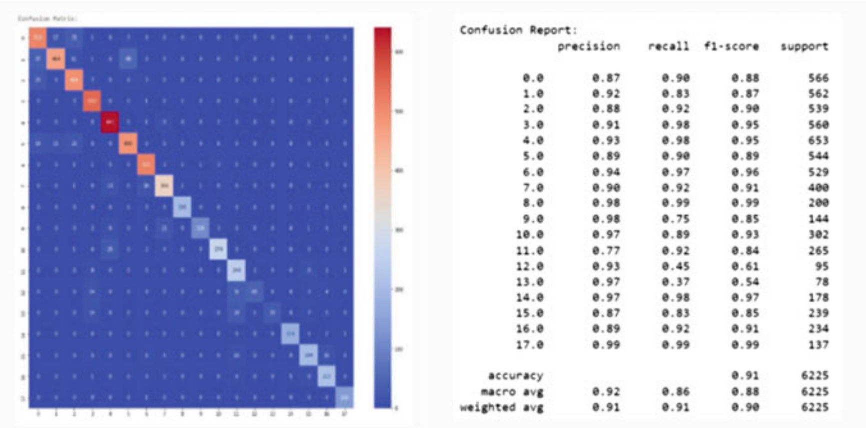


Fig. 10 Random forest confusion matrix and confusion report

Table 2 Performance evaluation

	DNN	LSTM	Decision tree	Random forest
Accuracy on testing data (%)	76	97	77	90
Average accuracy (%)	86.5		83.5	

4.3 Performance Evaluation

We used Google Collab with a GPU hardware accelerator. The accuracy of the DNN model on the test dataset is 76%. It implies that our model performs well for the unseen data. The correctness of the model is well for predicting jogging, standing, sitting, and walking. The model is not performing well for upstairs and downstairs activities. This model has the least accurate confusing matrix owing to the most mispredicted activities. Thus, the DNN model is the least accurate model (Table 2).

The LSTM model can achieve over 97% accuracy on the test dataset to predict human activities. This model has the most accurate confusion matrix as it has the least mispredictions. Hence, the LSTM model has the most accuracy and subsequently becomes the best proposed model.

Further, we have achieved an accuracy of 77% on the test data in the decision trees model. The model is great for predicting the activities like standing, sitting, talking while sitting, talking while standing, stand_sit, lay, and lay_stand activities.

Random forest model is able to predict the activities with an accuracy of 90% on the test set.

5 Conclusion

LSTM has the ability to see long-term dependencies in the dataset. It identifies the directionality of the variations in the true data. DNN can achieve the same effect, but it requires collecting the input vector over time and then feeding it to a huge layer. Hence, we have observed 21% increase in accuracy.

A random forest combines numerous decision trees. It makes use of the strength of multiple decision trees rather than relying on a single decision tree's feature relevance. Therefore, the accuracy of the model has increased by 13%.

We envisage extending our model into the detection of complex and detailed human activity using the inertial sensors, such as preparing a meal and answering a phone call. Thus, resulting in it being applicable to efficiently tackle real-world problems.

References

1. Milenkovic A, Jovanov E, De Groen PC, Otto C (2005) A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. J Neuroeng Rehabil 2:1–10. https://doi.org/10.1186/1743-0003-2-6. (Conference 2016, LNCS, vol 9999, pp 1–13. Springer, Heidelberg (2016))
2. Yalcin H, Tufek N, Yalcin M (2018) Activity recognition of interacting people. 26th IEEE Signal Process Commun Appl Conf SIU 2018:1–4, 2018. https://doi.org/10.1109/SIU.2018.8404173. (9th International Proceedings on Proceedings, pp 1–2. Publisher, Location (2010))
3. Wang H et al (2020) Wearable sensor-based human activity recognition using hybrid DL techniques. Secur Commun Netw 2020. https://doi.org/10.1155/2020/2132138
4. Badshah M (2019) Sensor—based HAR using smartphones 7:1008–1012
5. Gaikwad S (2021) Study on artificial intelligence in healthcare. In: 2021 7th international conference on advanced computing and communication systems, ICACCS 2021, pp 1165–1169. https://doi.org/10.1109/ICACCS51430.2021.9441741
6. Aggarwal R, Singhal A (2019) A R and its effect on our life. Proceedings of the 9th international conference on cloud computing, data science and engineering, confluence 2019, pp 510–515
7. Nambissan G, Mahajan P, Sharma S, Gupta N (2021) The variegated applications of deep learning techniques in human activity recognition. In: 2021 thirteenth international conference on contemporary computing (IC3–2021). Association for computing machinery, NY, USA, pp 223–233. https://doi.org/10.1145/3474124.3474156
8. Lamba D, Goyal S, Chitresh V, Gupta N (2020) An integrated system for occupational category classification based on resume and job matching (May 21, 2020). In: Proceedings of the international conference on innovative computing and communications (ICICC) 2020, Available at SSRN: https://ssrn.com/abstract=3607282
9. Agarwal I, Yadav P, Gupta N, Yadav S (2020) Urban sound classification using machine learning and neural networks. In: Proceedings of 6th international conference on recent trends in computing: ICRTC 2020, p 323. Springer Nature
10. Cheng L, Power K, Corporation T, Guan Y, Zhu K, Li Y (2017) Learning methods with wearable sensors
11. Ahmed N, Rafiq JI, Islam MR (2020) Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model. Sensors (Switzerland) 20(1). https://doi.org/10.3390/s20010317

12. Gupta N, Gupta SK, Pathak RK, Jain V, Rashidi P, Suri JS (2022) Human activity recognition in artificial intelligence framework: a narrative review. In: AI Review (Issue 0123456789). Springer Netherlands
13. Li L, Wu Y, Ou Y, Li Q, Zhou Y, Chen D (2018) Research on ML algorithms and feature extraction for time series. In: IEEE International symposium on personal, indoor and mobile radio communications, PIMRC, 1–5 Oct 2017
14. Walse K, Dharaskar R, Thakare VM (2016) Performance evaluation of classifiers on WISDM dataset for human activity recognition
15. Valai Ganesh S, Agarwal M, Gupta SK, Rajakarunakaran S (2021) Static and dynamic activities prediction of human using machine and deep learning models. In: Innovations in computer science and engineering, pp 1–7. Springer, Singapore
16. Sikder N, Nahid A (2021) KU-HAR: an open dataset for heterogeneous human activity recognition. Pattern Recogn Lett 146:46–54

Chapter 7
Concern Levels During COVID-19: An AI-Based Approach for Social Media Analysis

Divya Gupta, Shampa Chakraverty, Aryaman Sharma, and Bhavya Nagpal

1 Introduction and Related Work

Uncertainty about **CO**rona**VI**rus **D**isease (COVID-19) is continually posing various threats. The dynamic behaviour of the virus forced people to be restricted to their homes. Social media during confinement opened up the platform where people could convey their thoughts and concerns. Micro-blog analysis has been used in studies to investigate the various aspects of COVID-19. Sentiment analysis, information dissemination, and the extraction of popular themes and public opinions from microblogs gained relevance. Ridhwan et al. investigated the prevalence of various emotional intensities during the pandemic [6]. Another study by Kabir et al. looked at the emotional intensities at different locations [1]. Yousefinaghani et al. used microblogs to extract public opinion in the context of immunization through tweets [9]. Mahdikhani conducted a study to determine the emotional intensity relative to popular topics [5]. Researching on the informative role of social media, Kankanamge et al. discovered that social media posts aimed at increasing situation awareness attract more community attention than other posts [2]. Not only for these applications, but also to support governance and urban planning, social media for smart cities has grown in importance. Troisi et al. investigated sentiments in a smart city in relation to user technology [7].

From the literature study, we understand the importance of social media for a community. Despite the fact that several aspects have been studied, we discover a dearth of studies analysing levels of concern and awareness about the crisis. We believe that examining them can assist in determining whether a concern is alarming. Knowing this, administrative authorities can take steps to resolve the situation contributing to the emergence of a resilient community.

D. Gupta (✉) · S. Chakraverty · A. Sharma · B. Nagpal
Netaji Subhas University of Technology, New Delhi, India
e-mail: divyaguptaatwork@gmail.com

D. Garg et al. (eds.), *Innovations in Information and Communication Technologies*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-19-3796-5_7

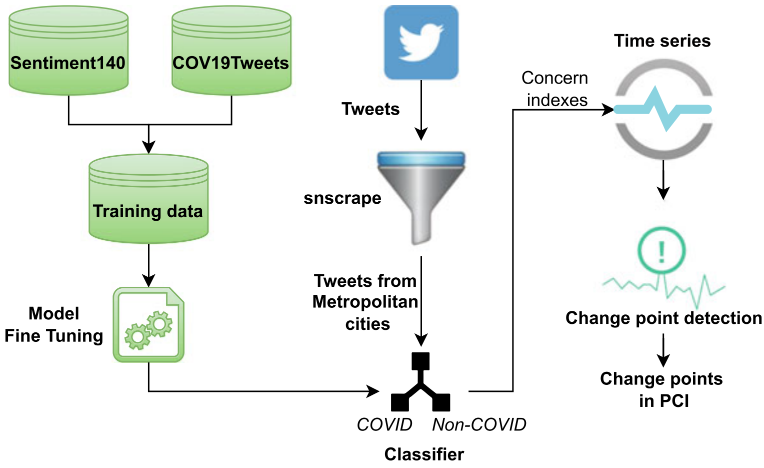


Fig. 1 Workflow for analysing concern levels during pandemic

In our work, we aim to automate the process of analysing the concern levels using a temporal tweet data. Figure 1 illustrates the workflow of our research. We fine-tune a deep learning model to identify whether a tweet is COVID-19-related. The components coloured green are all included in training the classifier. We extract tweets of an area without employing any hashtag or keyword. By analysing these tweets, we generate a time series of social media metric. Further, change point detection identifies the points of transitions in it. The main contributions of our work are:

1. Evaluate the concern levels by calculating a social media metric, Public Concern Index (PCI).
2. Analyse the variation in PCI during the second and third wave of the pandemic.

The remaining paper is laid out as follows. Section 2 describes the data collection, datasets used, and the methodology used to realize our approach. Section 3 presents the results of our analysis. Section 4 discusses the findings from the results. We conclude in Sect. 5.

2 Data and Methods

2.1 Datasets Used

Sentiment140 It is the dataset containing 1.6 million English tweets collected from 6 April 2009 to 25 June 2009.[1] We use this dataset to provide instances of *Non-COVID* tweets, i.e. tweets not related to COVID-19, for training the classifier. The rationale

[1] https://alt.qcri.org/semeval2017/.

for using it is that it was prepared long before the prevalence of COVID-19 was established. Additionally, it contains a massive set of tweets. Therefore, it provides a reliable sample of tweets not related to the pandemic.

COV19Tweets Lamsal proposed a dataset containing English tweets specific to COVID-19 collected across globe [3]. The dataset contains the IDs and sentiment scores of the pandemic-related tweets posted 20 March 2020 onwards. To extract the text associated with the tweet ID, different applications/libraries such as *Hydrator*,[2] *twarc*[3] are used.

As our work does not use the sentiment scores, we omit it from the dataset. Then, we select the tweets posted from October 2020 to December 2020 because it spans for the globally peaking stage of the first wave of pandemic.[4] This results in a dataset containing 1.4 million tweets which we term as *subCOV19Tweets*. It serves as a sample of *COVID* tweets, i.e. tweets which are COVID-19 related, for our classifier.

2.2 Data Collection

Twitter Data Our implementation extracts tweets posted during pandemic on the popular social media platform, Twitter. To comply with our aim, it is necessary to retrieve tweets without employing any hashtag or keyword in the extraction process. Since, this would result in massive amount of tweets, we limit our search to the ones originating from the metropolitan cities in India. The ground for such selection is that metropolitan cities are densely populated which can provide a good sample for our study.[5] Furthermore, the majority of population residing in such cities is digitally connected.[6] So, we work on the three metropolitan cities of India—Delhi, Mumbai, and Bangalore. As our training dataset includes the tweets posted during and before December 2020, we retrieve the geo-tagged tweets from January 2021 to January 2022. This time frame spans the onset, peak, and fading stages of both the second and third waves in India.

For extracting the tweets, we use a scraper for social networking services (SNS), *snscrape*.[7] This scraper can extract tweets using a text search query which includes fields such as the keywords, beginning date *since*, the end date *until*, the search location *near*, and the kilometre radius around a location *within*. For retrieving tweets, we use a text search query without any keyword *since: 2021-01-01 until: 2022-01-31 near: <city> within: 10 km*. The *city* in the given query is *Delhi, Mumbai, and Bangalore*. In total, we obtain a total of 3.9 million tweets.

---

2 https://github.com/DocNow/hydrator.

3 https://github.com/DocNow/twarc.

4 https://ourworldindata.org/explorers/coronavirus-data-explorer.

5 https://www.census2011.co.in/city.php.

6 https://timesofindia.indiatimes.com/business/india-business/for-the-first-time-india-has-more-rural-net-users-than-urban/articleshow/75566025.cms.

7 https://github.com/JustAnotherArchivist/snscrape.

2.3 Data Preprocessing

Data preprocessing is an essential step prior to building any model on a dataset. For preprocessing our datasets, we remove duplicate and null entries, user mentions, punctuations, emoticons, emojis, symbols, pictographs, transport and map symbols, and stopwords. We use Ekphrasis' word segmenter[8] trained on the twitter corpus to retain the hashtag content and then lemmatize.

2.4 RoBERTa

Liu et al. proposed a transformer-based **R**obustly **o**ptimized **BERT a**pproach (RoBERTa) pre-trained on a large English corpus [4]. The pre-trained RoBERTa model can be fine-tuned to perform classification tasks. Prior to fine-tuning, it tokenizes each input sentence adding special token *[CLS]* in the beginning and *[SEP]* at the end. After fine-tuning, the model produces the output embedding of the sentence including the pooled *[CLS]* which provides the overall contextual information of the input. For performing classification, the pooled *[CLS]* is fed to a linear layer which uses a sigmoid activation function to produce the labels. Some of the hyper-parameters of the model include the *optimizer*, number of epochs e, learning rate lr, maximum sequence length len, and batch size b.

Our implementation builds a classifier that can identify COVID-19-related microblogs from the collection of tweets. Therefore, we fine-tune RoBERTa to assign label *1* (*COVID*) or label *0* (*Non-COVID*) to tweets.

2.5 Social Media Metric

Wang et al. used a social media metric **P**ublic **C**oncern **I**ndex (PCI) to measure the awareness and concern about flooding in people [8]. It is calculated as:

$$\text{PCI}_{\text{topic}} = \frac{\text{Count of foreground tweets, i.e. tweets related to topic } (N_t)}{\text{Total number of tweets posted } (N)} \quad (1)$$

The value of PCI ranges from 0 to 1. Our study employs this metric to measure the level of concern and awareness in people about COVID-19. Recording PCI for our study period generates a time series data. Further, change point detection extracts the transition points from this time series. We also analyse the acceleration in the PCI time series. Acceleration in PCI is the difference between the current momentum and the momentum 7 days ago. A negative slope in acceleration indicates a decrease in PCI, and positive slope depicts an increase.

[8] https://libraries.io/pypi/ekphrasis.

Table 1 Sample tweets of the training data

S. No.	Tweet text	Label
1	Okkkay possibly time bed good night	0
2	Approved sale vaccine licenced doe vaccine approval number wondering lab rat	1
3	Congrats new gig think blast working gang jealous	0
4	Ahh come mandatory card hold record vaccination…	1
5	Mom husband tested positive COVID mom high risk terrified	1
6	School closure due affect million sparse Internet connection shortage COVID child Iraq	1
7	Never tweet back keep trying love adore come back Argentina soon incredible	0

2.6 Change Point Detection

Change point segments a time series into sections having uniform statistical characteristics. A change point detection identifies the change points in the underlying model of the time series. In our work, change points would signify the transitions in concern levels. For change point detection in PCI time series, we use Facebook's *Kats*[9] library. *CUSUMDetector* of the library extracts multiple change points from the time series using two windows—historical window of length *hw* and scan window of length *sw*. Additionally, a step value *s* is used to slide the windows over the time series. We appropriately select the values of *hw*, *sw*, and *s* for our time series. The detector outputs different attributes related to change points such as *p-value*, mean before the change point PCI_{bmean}, and mean after the change point PCI_{amean}.

In the next section, we present the results of model fine-tuning and time series analysis.

3 Experimental Results

Our implementation uses Python 3.8 on Kaggle[10] platform with Nvidia Telsa P100 GPU, 16 GB RAM.

[9] https://github.com/facebookresearch/Kats.

[10] https://www.kaggle.com/.

Table 2 Hyper-parameters of the $\text{RoBERTa}_{\text{BASE}}$ classifier

Optimizer	Number of epochs *e*	Maximum sequence length *len*	Learning rate *lr*	Batch size *b*
Adam	2	35	1e−5	32

3.1 Training Data Preparation

For training RoBERTa, we merge both datasets *Sentiment140* and *subCOV19-Tweets*. Prior to merging, the tweets are cleaned following the steps mentioned in Sect. 2.3. After preprocessing, *Sentiment140* contains 1.4 million tweets and *subCOV19Tweets* has 1.2 million tweets. We combine these tweets to form the training set with 2 columns—*tweet text* and *label*. Table 1 presents some of the preprocessed sample tweets of the training data. The *tweet text* column contains the text content of the tweet. The *label* column contains *0* for all the tweets of *Sentiment140* and *1* for all tweets of *subCOV19Tweets*.

3.2 Model Performance

For fine-tuning pre-trained 'roberta-base', we use the training data in 7:3 for training and testing. Table 2 lists all the hyper-parameters of the classifier. The classifier exhibits a precision of 98.9%, recall of 96%, and $F1$ score of 97.5%.

3.3 Generating PCI Time Series

For calculating the PCI score, we first preprocess the tweets acquired in Sect. 3.1 and label them using the classifier. Table 3 presents the results of the labelled preprocessed tweets. Then, we use Eq. 1 to evaluate the PCI score for each day from the labelled results. The count of tweets labelled *1* is assigned to N_t for a day. And, the count of total tweets of that day to N.

3.4 Change Points in PCI Time Series

We use *CUSUMDetector* with $hw = 9$, $sw = 5$, and $s = 4$. Figure 2 presents the result of change point detection in PCI time series. The presence of *CP1*, *CP2*, *CP3*, *CP4*, *CP5*, *CP6*, and *CP7* indicates seven transitions in PCI. Table 4 lists out the descriptive

Table 3 Sample of labelled tweets of the collected twitter data

S. No.	Preprocessed tweet content	label
1	Good idea concentration corona virus high closed area air circulation inside theatre in heal high concentration virus lung	1
2	New year welcome year happy new year	0
3	India heading towards serious disaster due cruelty towards innocent voiceless soul even many disaster corona epidemic one learn lesson	1
4	Brilliant yes special mag cricket whole lot sport	0

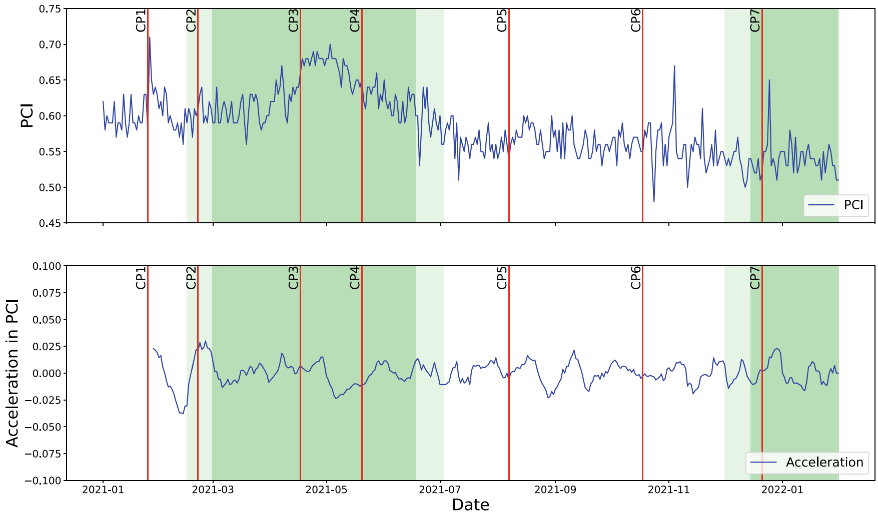


Fig. 2 Results of change point detection in PCI time series

of these change points. We report all the change points which establish statistically significant change in PCI mean with a *p-value* less than 0.05. Along with PCI time series, the plot in Fig. 2 shows the acceleration in PCI. The plot displays two green pockets—*P1* representing the second wave and *P2* marking the third wave. The gradient of green colour along each pocket from left to right reflects the stages of the wave. The dark green colour of the pocket indicates the peaking stage, light green on left depicts the onset stage, and light green on the right represents the fading stage. The white area accounts for the fading stages of the previous waves. *P1* contains the transitions along *CP2*, *CP3*, and *CP4*. And, *P2* packs the transitions associated with *CP7*. *CP1*, *CP5*, and *CP6* are not contained in any pocket. *CP1* overlaps with the fading stage of the first wave. *CP5* and *CP6* occur in the fading stage of the second wave. Further, we observe PCI_{bmean}, PCI_{amean}, and acceleration along each change point.

Table 4 Descriptives of the change points detected in PCI

Change point sequence number	Date	Mean PCI before change point PCI_{bmean}	Mean PCI after change point PCI_{amean}	p-value
CP1	25-01-2021	0.597	0.710	0.023
CP2	21-02-2021	0.592	0.635	0.005
CP3	17-04-2021	0.633	0.677	0.036
CP4	20-05-2021	0.652	0.615	0.003
CP5	08-07-2021	0.585	0.540	0.005
CP6	18-10-2021	0.558	0.580	0.013
CP7	21-12-2021	0.522	0.553	0.001

Change points contained in pockets: Our analysis reveals a positive acceleration in region of *CP2* with $PCI_{bmean} = 0.592$ and $PCI_{amean} = 0.635$. $PCI_{bmean} = 0.633$ and $PCI_{amean} = 0.677$ of *CP3* indicates a further positive acceleration. Constrastingly, *CP4* presents a negative acceleration with $PCI_{bmean} = 0.652$, $PCI_{amean} = 0.615$. Similarly, we observe *CP7* with $PCI_{bmean} = 0.522$, $PCI_{amean} = 0.533$ and a positive acceleration.

Change points contained in the fading stages: We observe *CP1* to be the intermittent peak with $PCI_{bmean} = 0.597$, $PCI_{amean} = 0.710$ and positive acceleration. Another point *CP5* shows a negative acceleration with $PCI_{bmean} = 0.585$ and $PCI_{amean} = 0.540$. Next to the decreasing trend along *CP5*, *CP6* associated with an intermittent peak reveals a positive acceleration with $PCI_{bmean} = 0.558$ and $PCI_{amean} = 0.580$.

We discuss the findings associated with the change points and the future scope in the next section.

4 Discussion

From the results, it is evident that majority of pandemic-related tweets were posted during the onset and peaking stages of the pandemic. We found an accentuation in PCI level during the onset and peaking stages. Interestingly, the level dropped as we enter the fading stage. The intermittent peaks observed could be attributed to the uncertain circumstances that come along. The peaks in the PCI inside *P1* and *P2* indicate higher concern levels during second wave in comparison with the third wave. Surprisingly, the acceleration pattern in these regions seem similar. Therefore, we inferred that the increase in the prevailing concern levels fqueryacross the two waves was similar.

During our research, we uncover potential areas for improvement. Future research could design a generic algorithm to identify crisis-related tweets to explore the public sentiment index.

5 Conclusion

We fine-tuned a RoBERTa model exhibiting an $F1$ score of 97.5% to identify pandemic-related tweets in our research. From the labelled results, we generated a PCI time series which was subjected to change point detection. The change points suggested transitions in the level of concern in the community during COVID-19. Our analysis reasoned from the findings that pandemic has a greater impact on people's levels of concern and awareness. Also, we believe that administrative authorities might benefit from real-time monitoring of PCI in relation to certain events.

References

1. Kabir MY, Madria S (2021) EMOCOV: machine learning for emotion detection, analysis and visualization using COVID-19 tweets. Online Soc Netw Media 23:100135
2. Kankanamge N, Yigitcanlar T, Goonetilleke A (2020) How engaging are disaster management related social media channels? The case of Australian state emergency organisations. Int J Disaster Risk Reduct 48:101571
3. Lamsal R (2021) Design and analysis of a large-scale covid-19 tweets dataset. Appl Intell 51(5):2790–2804
4. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V (2019) RoBERTa: a robustly optimized BERT pretraining approach. arXiv
5. Mahdikhani M (2022) Predicting the popularity of tweets by analyzing public opinion and emotions in different stages of covid-19 pandemic. Int J Inf Manage Data Insights 2(1):100053
6. Mohamed Ridhwan K, Hargreaves CA (2021) Leveraging twitter data to understand public sentiment for the COVID-19 outbreak in Singapore. Int J Inf Manage Data Insights 1(2):100021
7. Troisi O, Fenza G, Grimaldi M, Loia F (2022) Covid-19 sentiments in smart cities: the role of technology anxiety before and during the pandemic. Comput Hum Behav 126:106986
8. Wang B, Loo BPY, Zhen F, Xi G (2020) Urban resilience from the lens of social media data: responses to urban flooding in Nanjing, China. Cities 106:102884
9. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S (2021) An analysis of COVID-19 vaccine sentiments and opinions on twitter. Int J Infect Dis 108:256–262

Chapter 8
An Assessment of Application of Artificial Intelligence in Retail

P. Dhanya Mohan and M. Rakesh Krishnan

1 Introduction

Retailing around the world continuously evolves and innovates over time. Technological advancements in various sectors have helped all the retailing formats to a newer and better way of doing business. The COVID pandemic has now accelerated digitization in retail for the last couple of years. Various sectors adapt to changes due to the social distancing norms and the fear of pandemic. Digitization, artificial intelligence, and other technologies have helped people worldwide cope with the new normal. With the emergence of the Internet and related technologies in the early 1990s, the retail sector had also started adopting newer technologies.

Similarly, a sudden halt in sales during the country-wide shutdown led shopkeepers to look for alternate ways to keep their businesses running. Any organized retailer, whether small, medium, or large, would endeavor to adapt to the changes occurring in the environment and create brand awareness while also being cognizant of the store image [1]. A bibliometric evaluation aims to analyze the technology evolution in retail. The review will aid in locating published research on the topic of future technologies and artificial intelligence in the retail segment.

Artificial intelligence and virtual reality, and other digital strategies have been in the market for some time now. However, these technologies were used by larger organized retail chains as a strategic move to gain consumers' interest in them [2]. Now, in the pandemic, these technologies have become common and affordable for more prominent retailers, which is accelerated by the new social norms that have come to place. The study will explore the emerging technologies used in retail and the theoretical framework. Also, this would help identify the dynamics of the research area concerning various technologies used in the retail segment. On searching the

P. Dhanya Mohan (✉) · M. Rakesh Krishnan
School of Management Studies, Cochin University of Science and Technology, Kochi, Kerala, India
e-mail: dhanyamohanp@cusat.ac.in

D. Garg et al. (eds.), *Innovations in Information and Communication Technologies*, Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-19-3796-5_8

Scopus database, nearly five thousand articles were fetched with the keyword future of retailing from the past two decades.

Technology in retail marked its presence nearly five decades ago with the dawn of standard product codes and home shopping networks in the 1980s. The traditional brick-and-mortar sales have evolved to modern retail today, passing through various stages of technological advancement. This paper aims to look into the present literature available in technology in retailing and propose future areas for research work in retailing.

The advancement in technologies in the past few years and its application in retail is evident. However, the research on the retailer's acceptance and adoption of newer technologies have been limited. The adoption of various novel technologies has accelerated post the pandemic as to the merchants around the world are finding new ways of exploring new opportunities. The purpose of this work is to examine the research work conducted in the area of future retailing and to propose new areas of research for upcoming scholars. Hence, the paper aims to acknowledge the scholarly publication in Web of Science as well as the new arenas of exploring the technological prospects of organized and unorganized retailing.

Research objectives

- The dynamics of the technology adoption research over the three decades.
- To identify and group the most influential studies in the area of future technology also to identify the most cited and the most significant authors in the retailer technology adoption.
- The pattern and collaboration of countries and authors using biblioshiny package.
- To propose the themes for future research.

The article is divided in three parts. First is introduction and background of the study where the basic technologies used in retailing are discussed. Second, the research methodology is discussed and also the tools and bibliometric analysis. And finally, the analysis of the articles selected is incorporated in the final part under Analysis section along with the future scope for research work.

1.1 Background of the Study

There are several articles published on the future trends in retailing. The Web of Science database fetched more than one thousand five hundred results, searching the term future of retailing in the system. The unique combination of keywords fetched the results, which were again refined by journal articles published in business and management. Popular bibliometric packages such as Vos-viewer and biblioshiny in 'R' are used for fetching the network diagrams and author, country statistics. The primary digital technologies a retailer uses are mobility and social media. There are cloud-based technologies and retail analytics that merchants have used for the past decade. All other technologies such as AI—artificial intelligence, VR—virtual reality, various digital payments, AR—augmented reality, and others come under the aforementioned digital technologies. Singh pointed out the technological factors and non-technological factors that affect the business strategies relating to digitization or otherwise called digital business strategies [3].

The retail sector has evolved through two centuries in the world. Technologies and digitization have marked their presence since the 1970s. The big fintech companies are now looking at the small and medium business and B2B segment to penetrate the market, which holds a prospect for them. Also, the speed of Internet penetration to the deeper section of society is putting traditional merchants in a challenging position [4]. They need to cope with the changes around them to keep pace with the modern retailing options that the consumers are opting for their convenience. [5] Varadarajan identified technologies as interactive technologies, communication technologies, and infrastructural technologies in retail. Also, he proposed a process model for retailing strategies impact on interactive technologies.

The search was performed using a set of keywords, and there are exclusion and inclusion criteria. Hence, the first article published in the search query was in 1992. The steady growth of the articles started in 2008, and the most significant number of publications was in the year 2021. In 2022, there are around 38 articles available in the WOS database. The present study reviews the literature of retailers' technology adoption and proposes the research agenda for future research. The keywords are selected based on the previously published literature and close synonyms. The subject categories are business and management and its close subject categories. No specific time limit is fixed for the search; however, the first published article in the WOS database was in 1992. Hence, the timeframe will be between 1992 and 2022.

1.2 Research Methodology

The first step is to conduct the bibliometric analysis in artificial intelligence adoption; hence, the bibliometric analysis methodology is used. Vos-viewer is a computer program developed by Eck and Waltman [6] for bibliometric mapping. This helps analyze the critical keywords, co-citation network, and most popular authors and sources in the selected area. Also, the biblioshiny technique in R software is used for building data matrix with co-citation, keyword analysis, and collaboration analysis. Content analysis will help identify the techniques and theoretical models of the refined articles and help identify future research directions.

Research protocol followed in the review process was to develop a plan for review. The documents were selected based on Table 1 criteria. There was a combination of keywords used for the comprehensive search, and the most suitable synonyms were also included also incorporated with 'OR' Boolean operators. All keywords were identified based on the author's judgment on the topic.

Table 1 Literature search criterion

Literature search rule applied	Results
Document	Research article
Database	Web of Science
Technique	Bibliometric analysis
Tool	Vos-viewer, biblioshiny
Collection	1990–2022
Exclusion criterion	
Language	English only
Documents	Book review, conference proceedings
Subject/area	Business management, commerce

2 Bibliometric Analysis

The bibliometric analysis will help identify the most influential aspects of the future of retailing technologies from published sources. The database authors have relied on is Web of Science, and the bibliometric review is based on the contributions made per the database. Vos-viewer and biblioshiny package in R software is used to visualize bibliometric review. As per the biblioshiny package in Fig. 1, three field plot depicts the significance of keywords in the future of retailing and their flow from journals and their contributing countries. The figure shows that the USA has the most significant contribution on top, followed by China and U.K. respectively. The keywords 'impact,' 'technology,' 'behavior' are significant keywords that are predominantly used in A* rated Journals in ABDC journal quality list. Table 2 represents the main information

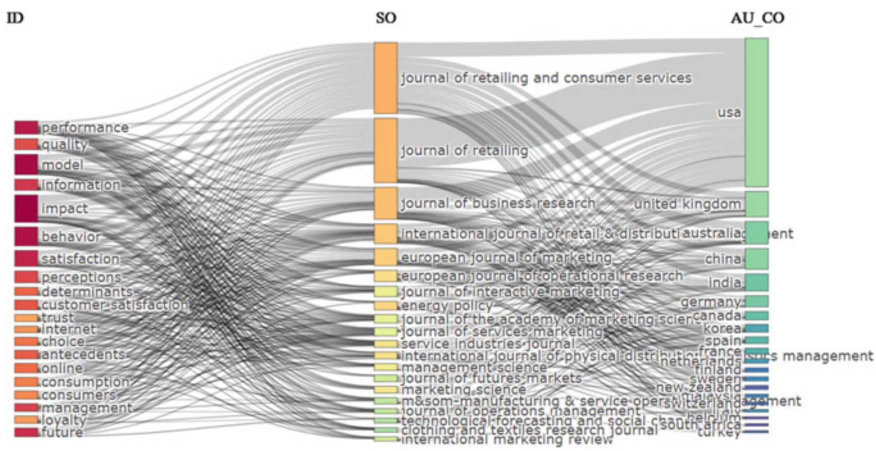


Fig. 1 Three field-plot from biblioshiny bibliometrics

Table 2 Main information about the data

Main information from web of science	Results
Timespan	1989:2022
Sources (journals, books, and other sources)	308
Documents	1405
Average years from publication	8.66
Average citations per documents	41.88
Average citations per year per doc	3.988
References	62,932
Keywords plus (ID)	2868
Author's keywords (de)	4097
Authors	3273
Single-authored documents	174
Documents per author	0.429
Authors per document	2.33
Co-authors per documents	2.74
Collaboration index	2.52

extracted from the Web of Science database.

The three-field plot shows that a particular set of themes influences the future of retailing research. A thematic evolution in biblioshiny package also reveals about future of retailing, digital technologies, and omnichannel retailing. Figure 2 represents the thematic evolution in the future of retailing. The thematic map has specific parameters. Such as the keywords are author keywords derived from the database. Also, the map is based on the Web of Science database. The timeframe is divided into

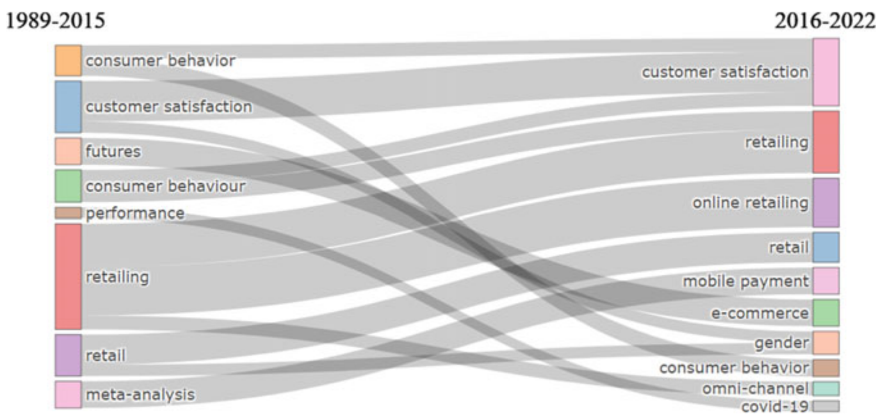


Fig. 2 Thematic analysis of author keywords

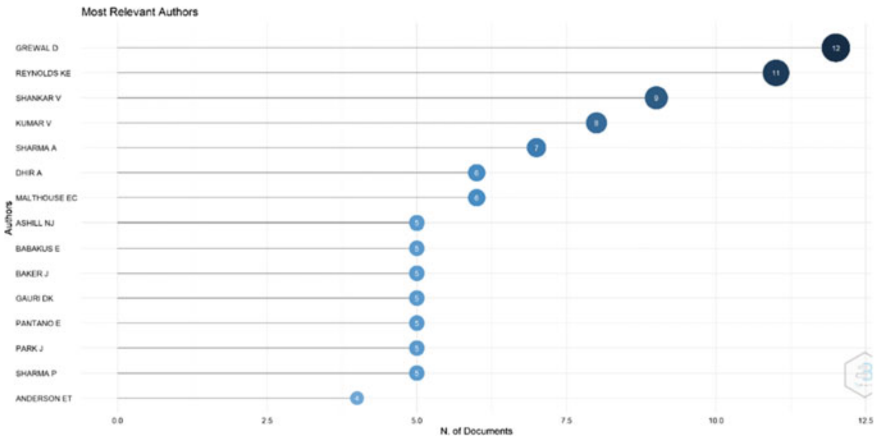


Fig. 3 Authors contribution in the retailing research area

two. Moreover, the analysis shows how the keywords transitioned from consumer behavior to online retailing to e-commerce, omnichannel, and other technologies.

Authors contribution in the research area

The analysis from the bibliometric package in R software shows the most contributing author in the technology adoption and artificial intelligence sector.

The most relevant authors in the retailing research area are depicted in Fig. 3. As per this document, Grewal has published the most significant number of articles. Figure 4 depicts the country's contribution to the future of the retailing area. The network diagram as per Fig. 4 is derived from Vos-viewer software used for bibliometric visualization. As per the network diagram, the USA has published more literature on the future of retailing and technologies. The co-authorship map also shows the countries like Australia, Germany, and England's contribution.

3 Analysis

The digitization and Internet of Things have revolutionized the retail sector and the research work in the area. The contribution of the researchers around the world is represented in Fig. 5. Here, the country collaboration for the contribution in the research is shown where the USA is most significant in the map as the collaboration network is robust and vast compared to other countries. South American countries like Brazil and Argentina also have contributed to the literature as per the figure.

Figure 6 represents the conceptual structure map based on the multiple correspondence analysis, also known as the MCA of the keywords. Here, a conceptual structure is made using the keywords associated with the resilience articles selected for the study. The analysis compresses the extensive data into low-dimensional space,

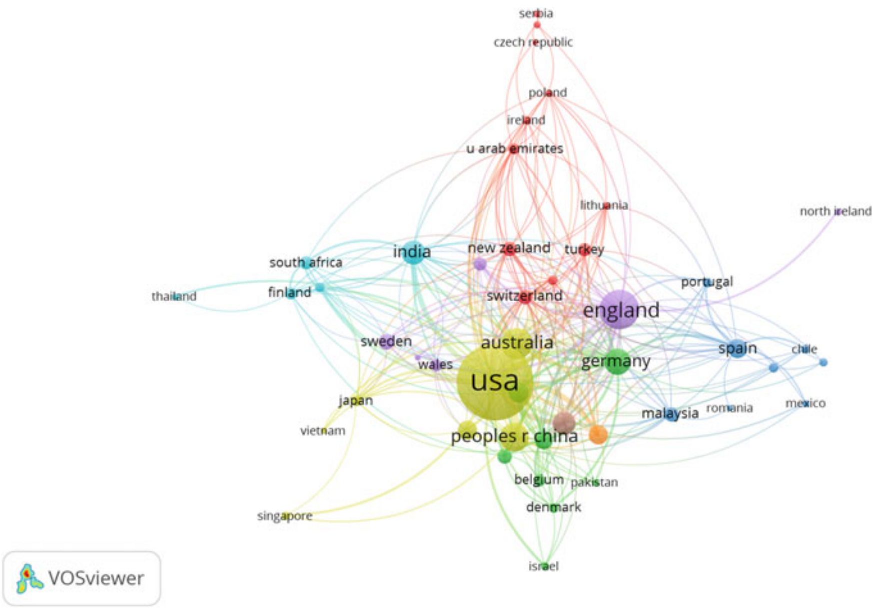


Fig. 4 Co-authorship map of countries

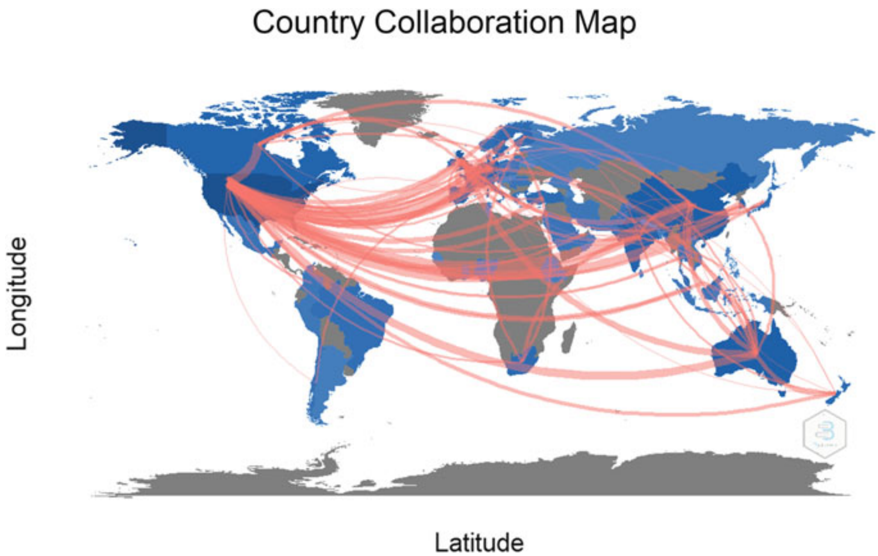


Fig. 5 Country collaboration map from biblioshiny package

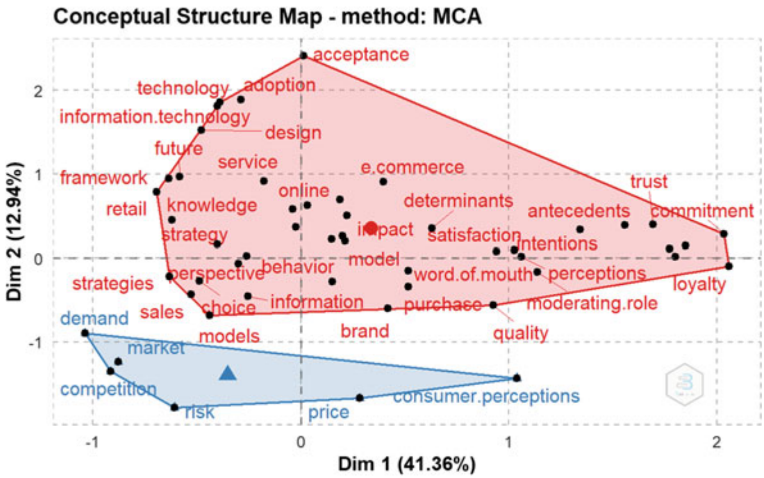


Fig. 6 Multiple correspondence analysis maps from biblioshiny package

forming a two-dimensional graph as in Fig. 6, which shows the distance and similarity of keywords. Cluster one in red represents the more similarly distributed keywords in red color. It also represents the more essential keywords that represent the most studied area.

3.1 Digitization of Retail

Digitization and technological advancement in the retail sector has been present since the 1970s. The developed countries mainly concentrated on researching and developing technologies, whereas the emerging countries concentrated on commercializing innovation. Reinartz et al. [7] examined the challenges and opportunities of retailing innovations in a global context. Further, the authors have divided the market into mature, emerging, and least developed countries for analysis. The significant contributions made in the digitization and the future of the retailing area are from the ABDC journal list A* rated journals in retailing.

3.2 Future of Digital Formats in Retailing

Internet and communication technologies are not confined only to retail sales or inventory management systems. AI-driven demand forecasting is an area of advancement in retail. COVID pandemic resulted in panic buying among consumers followed by random lockdowns and restrictions around many countries in the world. This resulted in machine learning-based demand forecasting by the merchants, which would enable them to examine the current demand and anticipate future demand. Digital technologies are also facilitated by machine learning techniques that accelerate automated forecast processing and identify the hidden pattern of the data collected. Many of the more giant retailers have developed their machine learning packages to help analyze the data and create a robust system with high adaptability to frequent changes.

3.3 Retailing Innovation

The indoor positioning system is a new technology introduced for locating and finding an object or human where the global positioning system lacks precision in results. The IPS technology is now successfully used in retail chains and shopping malls, which helps retailers identify and track visitors' statistics. The GPS also helps in location analytics and optimization and layout of their stores [8]. IPS offers a range of benefits to the consumers, which helps them locate their targeted store and product quickly in a big store.

Societal media and the use of such media are also an innovation concerning traditional retailers. Social media here means the platforms like Facebook and Instagram and market places like Amazon and the virtual stores on the Internet. Verhoef et al. [9] asserted the consumer behavior in multiple channels while shopping. Verhoef also looked at the scope of the social media. Torres de Oliveira et al. [10] Torres exerted that a moderating role of digital capabilities and channels positively impacted the use of social media.

3.4 Data Confidentialities and Apprehensions

Confidentiality is well defined as 'the ability of the individual to control the terms under which personal information is acquired and used' [11]. Data privacy is mainly concerned with the consumers and retailers concerning regulatory issues and data compromises. The risk of access to the newest technologies by retailers poses the threat of knowing their customers too well, which poses a lack of trust among consumers. Bleier and Eisenbeiss [12] insisted on the perceived usefulness of the personalized advertisement without hindering the customers' privacy. Ponte and Bonazzi [13] proposed that there is a willingness to use innovative systems from consumers. Data privacy concerns are concerned with consumers and retailers and other regulators. The data privacy concern can be addressed through social media, traditional brick-and-mortar, and mobile channels [14].

Fig. 7 AI adoption index. *Source* (IBM, 2021)

3.5 AI, AR, and Interactive Technologies in Retail

Artificial intelligence and other emerging technologies research have benefited businesses theoretically and practically. [15] Voice-based artificial intelligence systems such as Siri, Cortana, and Alexa are popular applications that rely on Artificial intelligence. As per the prospects, [16] predicts the deployment and exploring the artificial intelligence is shown in Fig. 7. In the case of virtual reality, the technology aims to create an experience of a place without actually visiting it or experiencing the same. Any well-organized retailer, whether small, medium, or large, would strive to adapt to changes in the environment and promote the brand while keeping the shop's image in mind [1]. Virtual stores on the web, activities on platforms such as Facebook and Instagram, and marketplaces such as Amazon [9] have facilitated buyers to move through these channels and buy goods. Consumer behavior across multiple channels when shopping was asserted [10]. Torres explained how digital proficiencies and channels play a moderating function in social media usage.

Digital technologies in manufacturing are machine learning, artificial intelligence, predictive analytics, where the companies can monitor and direct the facilities on time. The companies that use disruptive and innovative technologies get the extra mile of improved facilities and make real-time decisions. Whiteside [17] explains how demand forecasting error is decreased by 20%, which further reduces lost sales by 30%, reducing the planners' workload by 50%. AI in manufacturing is mainly used in predictive maintenance, which helps to reduce the unprecedented downtime in factories. Also, AI creates an optimal design that helps reduce time and increases efficiency. Managing process quality is as old as a century; Walter A Shewhart introduced a statistical control chart, which significantly increased the quality of the manufacturing process. Senoner et al. [18] explained a data-driven model which improves the process quality in manufacturing. Here, artificial intelligence is adapted and used in quality management. Zeba et al. [19] examined the technologies used in business, such as intelligent manufacturing, deep learning, and real-time forecast algorithms.

conversational AIs popularly known as voice commerce is expected to grow beyond 70 million users in the United States of America in the next three years [20]. AR capable smart mirrors are another technology introduced a few years ago and widely accepted during the pandemic in 2020. The most important advantage is that intelligent mirrors are not intrusive, and hence, they do not breach consumers' privacy.

4 Scope for Future Research

Further studies can be done as to how the penetration of various technologies to the lower strata of the society and the informal sector of the economy help in the inclusion of these populations into the digital transformation. Secondly, there is a scope of exploring the niche segments in retail with the advancement of technologies. Thirdly, innovation and adoption of technologies are critical for the success of retail firms in the future, especially in the case of firms belonging to the informal economy. Specific industry challenges are also a prospective area for future studies when adopting newer technologies. The current generation of technology developments, e.g., quantum computing, could accelerate the progression of AGI, artificial general intelligence. Hence, governments and policymakers are encouraged to look at such trends, and laws need to be installed to limit the growth of the machine's cleverness and help mankind from the threats of robots. Also, outlets need to follow to rules that might dedicate them to being genuinely, socially accountable businesses whose interest into AI-driven revolution and employment are for customer's good along with personnel the society.

5 Conclusion

Artificial intelligence primarily are those packages reshaping how customers select channels, pick out services and products, and make purchases. AI-based innovations help the users make the proper decisions, feeling much less time stress and growing confidence in addition to delight with their choices. However, retailers do not prefer to embody the rising improvements if it shields their competitiveness. AI is not assumed to update humans. Digital-centric retail approaches will continue to disrupt the business in the coming years and the adoption of these technologies in the emerging economies and the lesser developed economies. The study has certain limitations as Web of Science relied upon published sources in artificial intelligence and other digital technologies. Articles from other sources also can be considered for studying the broader span of technologies and adoption. The darker side of artificial intelligence needs to be studied thoroughly as it poses greater challenges for the consumers and there are limited studies on data privacy and related areas.

References

1. Graciola AP, De Toni D, Milan GS, Eberle L (2020) Mediated-moderated effects: high and low store image, brand awareness, perceived value from mini and supermarkets retail stores. J Retail Consum Serv 55:102117. https://doi.org/10.1016/J.JRETCONSER.2020.102117
2. Grewal D, Gauri DK, Das G, Agarwal J, Spence MT (2021) Retailing and emergent technologies. J Bus Res 134:198–202. https://doi.org/10.1016/J.JBUSRES.2021.05.004
3. Singh G, Gaur L, Agarwal M (2017) Social sciences ANS humanities factors influencing the digital business strategy. Pertanika J Soc Sci Hum 25:1489–1500
4. Inman JJ, Nikolova H (2017) Shopper-facing retail technology: a retailer adoption decision framework incorporating shopper attitudes and privacy concerns. J Retail 93:7–28. https://doi.org/10.1016/J.JRETAI.2016.12.006
5. Varadarajan R, Srinivasan R, Vadakkepatt GG, Yadav MS, Pavlou PA, Krishnamurthy S, Krause T (2010) Interactive technologies and retailing strategy: a review, conceptual framework and future research directions. J Interact Mark 24:96–110. https://doi.org/10.1016/J.INTMAR.2010.02.004
6. van Eck NJ, Waltman L (2010) Software survey: VOS viewer, a computer program for bibliometric mapping. Scientometrics 84:523–538. https://doi.org/10.1007/S11192-009-0146-3/FIGURES/7
7. Reinartz W, Dellaert B, Krafft M, Kumar V, Varadarajan R (2011) Retailing innovations in a globalizing retail market environment. J Retail 87:S53–S66. https://doi.org/10.1016/j.jretai.2011.04.009
8. Indoor positioning technologies as a rising force in retail sales—intellias, https://intellias.com/indoor-positioning-technologies-as-a-rising-force-in-retail-sales/
9. Verhoef PC, Kannan PK, Inman JJ (2015) From multi-channel retailing to omni-channel retailing: introduction to the special issue on multi-channel retailing. J Retail 91:174–181. https://doi.org/10.1016/J.JRETAI.2015.02.005
10. Torres de Oliveira R, Indulska M, Steen J, Verreynne ML (2020) Towards a framework for innovation in retailing through social media. J Retail Consum Serv 54:101772. https://doi.org/10.1016/j.jretconser.2019.01.017
11. Westin AF (1968) Washington and lee law review privacy and freedom. Lee L Rev 166
12. Bleier A, Eisenbeiss M (2015) The importance of trust for personalized online advertising. J Retail 91:390–409. https://doi.org/10.1016/J.JRETAI.2015.04.001
13. Ponte D, Bonazzi S (2021) Physical supermarkets and digital integration: acceptance of the cashierless concept. https://doi.org/10.1080/09537325.2021.1994942. https://doi.org/10.1080/09537325.2021.1994942
14. Martin KD, Palmatier RW (2020) Data privacy in retail: navigating tensions and directing future research. J Retail 96:449–457. https://doi.org/10.1016/J.JRETAI.2020.10.002
15. Grewal D, Roggeveen AL, Nordfält J (2017) The future of retailing. J Retail 93:1–6. https://doi.org/10.1016/J.JRETAI.2016.12.008
16. IBM: Global AI Adoption Index 2021, 1–13 (2021)
17. Whiteside J. Three ways AI is improving manufacturing operations | Manufacturing Global. https://manufacturingglobal.com/ai-and-automation/three-ways-ai-improving-manufacturing-operations

18. Senoner J, Netland T, Feuerriegel S (2021) Using explainable artificial intelligence to improve process quality: evidence from semiconductor manufacturing. https://doi.org/10.1287/mnsc.2021.4190. https://doi.org/10.1287/MNSC.2021.4190
19. Zeba G, Dabić M, Čičak M, Daim T, Yalcin H (2021) Technology mining: artificial intelligence in manufacturing. Technol Forecast Soc Change 171:120971. https://doi.org/10.1016/J.TECHFORE.2021.120971
20. U.S.: number of Smart Homes 2017–2025 | Statista, https://www.statista.com/forecasts/887611/number-of-smart-homes-in-the-smart-home-market-in-the-united-states

Author Index

D. Garg et al. (eds.), *Innovations in Information and Communication Technologies*,
Algorithms for Intelligent Systems, https://doi.org/10.1007/978-981-19-3796-5