The Journal of Supercomputing (2020) 76:5923–5947
https://doi.org/10.1007/s11227-019-03081-4

# Predicting the customer's opinion on amazon products using selective memory architecture-based convolutional neural network

Trupthi Mandhula[1] · Suresh Pabboju[2] · Narsimha Gugulotu[3]

Published online: 19 November 2019


**Abstract**

Opinion mining and sentiment analysis are useful to extract subjective information out of bulk text documents. Predicting the customer's opinion on amazon products has several benefits like reducing customer churn, agent monitoring, handling multiple customers, tracking overall customer satisfaction, quick escalations, and upselling opportunities. Though performing sentiment analysis is a challenging task for the researchers to identify the user's sentiments from the large datasets, it is unstructured in nature, and also includes slangs, misspells, and abbreviations. To address this problem, a new proposed system is developed in this research study. Here, the proposed system comprises of four major phases; they are data collection, pre-processing, keyword extraction, and classification. Initially, the input data were collected from the dataset: amazon customer review. After collecting the data, pre-processing was carried out for enhancing the quality of collected data. The pre-processing phase comprises of three systems: lemmatization, review spam detection, and removal of stop words and URLs. Then, an effective topic modelling approach latent Dirichlet allocation along with modified possibilistic fuzzy C-Means was applied to extract the keywords and also helps in identifying the concerned topics. The extracted keywords were classified into three forms (positive, negative, and neutral) by applying an effective machine learning classifier: Selective memory architecture-based convolutional neural network. The experimental outcome showed that the proposed system enhanced the accuracy in sentiment analysis up to 6–20% related to the existing systems.

**Keywords** Convolutional neural network · Latent Dirichlet allocation · Lemmatization · Modified possibilistic fuzzy c-means · Adam optimization algorithm · Sentiment analysis

✉ Trupthi Mandhula
trupthjan@gmail.com

Extended author information available on the last page of the article

## 1 Introduction

Presently, online shopping is growing effectively in the fields of e-commerce in which the customer experiences are achieved after the purchase over interest. The non-rational factors play an important role in the development and activity of social movements through online media that have the main impact on viral spreading [1, 2]. The multi-analysis approaches mainly concentrated on the concerns in online social websites that depend on the emotion passed with the messages [3, 4]. The cognitive effort lessens the heuristics and also the information is easily accessible; this plays a crucial role in online purchase decisions in order to have a good sale [5, 6]. In social websites, the sentiment analysis is very useful to shed light on the role of emotion in both offline and online [7–10]. Investigating the effect of online user reviews increases product awareness [11–14]. The social web text identifies the false sentiment patterns that are irrelevant to the topic, which is extensively applied in a variety of dissimilar social web contexts [15–19]. Nowadays, sharing the information becomes a trend among business partnerships that is a mutually beneficial way to increase productivity. The main purpose of this research study is to propose a proper keyword extraction methodology and classification approach for classifying the opinions of customers as neutral, negative, and positive forms utilizing amazon customer review dataset.

In this research, sentiment analysis was performed on a reputed dataset, i.e. amazon customer review dataset. After the collection of input data, pre-processing was carried out by applying lemmatization, review spam detection, and removal of stop words and URLs from the input data. The lemmatization converts the words of a sentence into a dictionary form in order to extract the proper lemma. In addition, review spam detection identifies the customers' untruthful opinions like positive spam reviews and negative spam reviews. Then, the pre-processed data were utilized to extract the keywords by applying an effective topic modelling approach LDA. In addition, modified PFCM was used to cluster the extracted keywords on the basis of amazon products. In conventional PFCM, a quantum-inspired methodology was included to obtain the correct cluster number (three). The quantum-inspired methodology operates on the smallest information representation named as a quantum bit (qubits). The output of modified PFCM was given as an input for SMA-CNN classifier to classify the opinions of the customers for amazon products: neutral, positive, and negative. Here, a new optimization algorithm (Adam optimization algorithm) was applied to optimize the moments of feature values in CNN–long short-term memory networks (LSTM) classifier that helps to identify the condition of best accuracy, and also to minimize the estimated error. Finally, the proposed system performance was compared with other existing systems in light of recall, f-measure, precision, classification accuracy, and area under curve (AUC).

This research paper is arranged as follows. Several recent papers on sentiment analysis are reviewed in Sect. 2. Problem statement about the existing methods is shown in Sect. 3. Detailed explanation of the proposed system is given in Sect. 4. Section 5 illustrates the quantitative analysis and comparative analysis of the proposed system. The conclusion is given in Sect. 6.

## 2 Literature review

Researchers developed numerous methodologies on dissimilar stages of sentiment analysis. In this literature section, a brief review of some important contributions to the existing literatures is given in Table 1.

## 3 Problem definition and solution

This section describes a problem statement in sentiment analysis and also detailed about how the proposed system gives the solution to the problem statement.

- Expert knowledge is required to select a suitable keyword modelling approach

After pre-processing the collected data, keyword extraction is carried out to find the optimal keywords from the huge data. In sentiment analysis, finding the keywords from the huge database is one of the emerging concerns for researchers. The high-dimensional data increase the system complexity because if the volume of the data space increases then the collected data becomes sparse which leads to "curse of dimensionality" issue. In sentiment analysis, LDA is a well-known topic modelling approach that is designed to find the keywords from huge dataset. Though, LDA is a well-defined generative model that can be easily extended in other complicated approaches. To further improve the performance of keyword extraction, a new clustering algorithm (modified PFCM) is combined with LDA for finding the optimal keywords. In this research study, after implementing the LDA algorithm, the extracted keywords are optimized by using modified PFCM. In PFCM, quantum-inspired method is included for identifying the similarity between the objects, which effectively reduces the computational complexity of the system.

- Expert knowledge is required to select an appropriate classifier

After extracting the keywords from the pre-processed data, classification is carried out to classify the opinions of the customers for amazon products. In sentiment analysis, binary classifiers like support vector machine is a well-known classifier that is designed for the two-class problem. The success of binary classifiers depends on the decision boundary that delivers good generalization performance [28–30]. The two major problems in binary classifiers are ineffective in high-dimensional data and applicable only for two-class classification. To address these issues, multiclass classification approaches are developed. Solution: In this research work, a new classifier SMA-CNN is implemented for multiclass classification. The SMA-CNN classifier effectively diminishes the size of the resulting dual issue by developing a relaxed classification error bound. In addition, the undertaken classification approach speeds up the training process by maintaining a competitive classification accuracy.

**Table 1** Review of existing literatures

| References | Methodology | Advantage | Limitation |
|---|---|---|---|
| Giatsoglou et al. [20] | Hybrid machine learning approaches | The developed system adopts a machine learning approach with textual documents for training a polarity classification model. Here, many document vector representation approaches were applied: lexicon-based, word embedding-based and hybrid vectorizations for improving the sentiment classification task | A number of features were considered for data feature extraction, which may decrease the accuracy of data classification |
| Alsinet et al. [21] | Argumentative approach | In this work, a natural extension of the system was introduced and investigated, in which relationships between tweets were related to a probability value, which indicates the uncertainty that the relationships hold | The major drawback of this research work was the developed approach did not concentrate on the pre-processing of twitter data |
| Balahur and Perea-Ortega [22] | Multilingual sentiment analysis system | In this research paper, hybrid features, multilingual, machine-translated data attained better relevant features for sentiment classification and also it increases the precision of sentiment analysis systems | In some cases, the developed system inefficient for large-scale databases like amazon customer review, twitter-sanders-apple 2 datasets that leads to two major problems: computational complexity and poor data classification |
| Bouazizi and Ohtsuki [23] | A pattern-based approach for multi-class sentiment analysis in twitter data | The developed method is scalable and classifies the texts into more classes. Here, a new tool SENTA was utilized for selecting the extensive variety of features for better classification | Mostly, the binary classification supports only structural data. It does not support unstructured data, which was considered as one of the major drawbacks in this research work |
| Yu et al. [24] | Hierarchical topic modelling | This system automatically mines the hierarchical dimension of tweets' topics that help to improve the classification accuracy. The experimental results demonstrate that the developed approach outperformed other current topic models in mining and constructing the hierarchical dimension of tweeter topics | The LDA (topic modelling approach) cannot support latent sub-topics within a determined label or any global topics |

**Table 1** (continued)

| References | Methodology | Advantage | Limitation |
|---|---|---|---|
| Bharathi et al. [25] | Hybrid features with sentence level sentiment score algorithm | The developed approach delivered high accuracy in stock market detection by merging the Sensex points with simple tweets and syndication news feeds | Performing classification was very difficult in this research paper, when the dimensions of the collected data values were very high |
| Saif et al. [26] | SentiCircles, (lexicon-based approach) | In this research paper, the developed approach delivers static and fixed prior sentiment polarities. The SentiCircles consider dissimilar word patterns in the tweets for capturing the semantics and to update the pre-assigned polarity and strength in sentiment lexicons | The lexicon-based approaches need more repeated words for updating the pre-assigned strength and polarity of the words |
| Ren et al. [27] | LDA and support vector machine | Here, LDA automatically extracts the keywords from the collected data and then the support vector machine classified the twitter sentiments. The experimental outcome showed that the topic-enhanced word embedding was very effective in twitter sentiment classification | SVM is a binary classifier, which supports only two-class classification |

# 4 Proposed system

In recent periods, sentiment analysis gained much attention among the researchers, and it plays a major role in several applications such as healthcare, marketing, retail industry, and education. This research study tackles the issue of sentiment polarity categorization that is one of the major issues of sentiment analysis. The input data utilized in this research are online product reviews collected from amazon.com. The proposed system majorly consists of four phases: data collection, pre-processing, keyword extraction, and classification. Figure 1 shows the workflow of the proposed system. The detailed explanation about the proposed system is described below.

## 4.1 Data collection

At first, the input data are collected from the dataset: amazon customer review dataset. It is comprised of customer reviews from the amazon website. The time span of the amazon customer review dataset is 18 years that include approximately 35 million reviews up to March 2013. The reviews comprise product ratings, user information, product information, and a plain text review. Table 2 describes the data characteristics of the amazon customer review dataset.

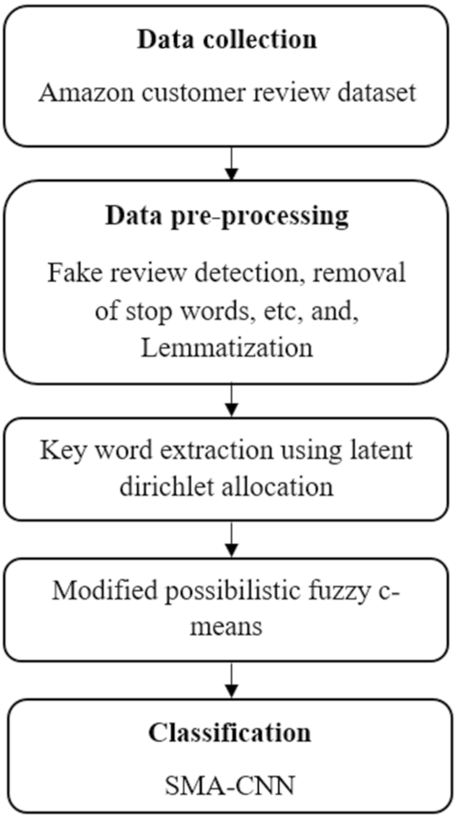


**Fig. 1** Work flow of proposed system

**Table 2** Data statistics

| Dataset statistics | |
|---|---|
| Number of users | 6,643,669 |
| Number of reviews | 34,686,770 |
| Number of products | 2,441,053 |
| Time span | June 1995–March 2013 |
| Median no. of words per review | 82 |
| Users with > 50 reviews | 56,772 |

### 4.2 Data pre-processing

After the collection of data from amazon customer review dataset, data pre-processing is carried out to enhance the quality of collected data. Generally, the raw data contains more noise in terms of stop words, and URLs, which are all removed effectively from the collected data. In addition, lemmatization technique and review spam detection are applied to further enhance the quality of data.

- *Review spam detection* The main task of review spam detection is to identify the customers' untruthful opinions like positive spam reviews and negative spam reviews.
- *Lemmatization* It transforms the words of a sentence into dictionary form. In order to extract the proper lemma, it is essential to analyse each morphological word [31, 32]. An example of lemmatization is denoted in Table 3.

### 4.3 Keyword extraction

After pre-processing the collected data, keyword extraction is performed using LDA [33, 34]. It is a probabilistic topic methodology, where each document is indicated as a random combination of latent topics. In LDA, each latent topic is considered as a fixed set of words that is utilized for identifying the primary latent topic structure on the basis of observed data. Generally, the words are generated in a two-phase mechanism for every document. In the first phase, the random distribution over the topic is selected for every document. A word is a distinct data from a vocabulary index $\{1, \ldots V\}$, a series of $N$ words are indicated as $w = (w_1, w_2, \ldots, w_n)$ and a collection of $M$ documents are signified as $D = w = (w_1, w_2, \ldots, w_M)$. LDA is a three layer representation, where the parameters $\mu$ and $\pi$ are analysed during the corpus generation.

**Table 3** A sample example of lemmatization

| Form | Morphological information | Lemma |
|---|---|---|
| Studying | Gerund of the verb study | Study |
| Studies | Singular number, third person, present tense of the verb study | Study |

For each document, the document-level topic values are analysed. Besides, the word-level values are also analysed for every word in the document. The joint distribution of random values is denoted as the generative mechanism of LDA. The probability density function of $k$-dimensional Dirichlet random value is evaluated using Eq. (1). Besides, the probability of a corpus and the joint distribution of a topic mixture are estimated by utilizing Eqs. (2) and (3).

$$p(\aleph|\pi) = \frac{\Gamma\left(\sum_{i=1}^{k} \pi_i\right)}{\prod_{i=1}^{k} \Gamma\left(\pi_i\right)} \aleph_1^{\pi_1 - 1} \ldots\ldots \aleph_k^{\pi_k - 1} \tag{1}$$

$$p(\aleph, x, y|\pi, \mu) = p(\aleph|\pi) \prod_{n=1}^{N} p\left(x_n|\aleph\right) p\left(y_n|x_n, \beta\right) \tag{2}$$

$$p(D|\pi, \mu) = \prod_{d=1}^{M} \int p\left(\aleph_d|\pi\right) \times \left( \prod_{n=1}^{N_d} \sum_{x_{dn}} p\left(x_{dn}|\aleph_d\right) p\left(y_{dn}|x_{dn}, \mu\right) \right) d\aleph_d \tag{3}$$

where $N$ is indicated as a number of words, $\aleph$ is signified as document-level topic values,$\mu$ is stated as topics, $x$ is specified as a per word topic assignment, $y$ is stated as observed word, $M$ is denoted as a document, and $\pi$ is specified as Dirichlet parameter.

In the LDA method, the estimation of the posterior distribution of the hidden value in a document is an essential task. The exact interpretation of the posterior distribution of the hidden value is a crucial issue. The grouping of LDA with approximation algorithms like Markov chain, variational approximation, Laplace, and Gibbs sampling is widely used for keyword extraction. The positive and negative keywords are extracted with individual weight values, and the extracted keywords are stored in the dictionary. The testing data are coordinated with the dictionary in the testing phase for obtaining the negative and positive weight values. After obtaining the negative and positive weight values, the clustering process is carried out by using the modified PFCM algorithm.

### 4.4 Modified possibilistic fuzzy c-means

Clustering is a task that identifies the hidden groups in a set of objects accurately. Also, it is an unsupervised approach, so clustering does not require previous knowledge of both outputs and inputs. In this research study, the PFCM approach is utilized for clustering the membership grade. Generally, the PFCM clustering considers every object as a member of each cluster with a variable degree of “membership function”. In modified PFCM clustering, quantum-inspired method is included to find the similarity between the objects that effectively lessens the computational complexity of the system. In this research study, quantum-inspired method plays an essential role in obtaining correct clusters; here, the optimal cluster size is three. The quantum-inspired method operates on the smallest information representation named as a quantum bit (qubits) [35, 36]. The classical bits are represented as "1" and "0" that store the information at a time, where a single qubit has the capability to store number of information at a time with the help of a probability feature. Qubit states the linear superposition of "1" and "0" bits probabilistically, which is stated in Eq. (4).

$$Q = \alpha|0> + \beta|1> \tag{4}$$

where $\alpha$ and $\beta$ are represented as complex numbers, which appears in two states, state "0" and state "1". Thus, $\alpha^2$ and $\beta^2$ are denoted as probabilities of a qubit in state "0", and state "1", which is described in Eq. (5).

$$\alpha^2 + \beta^2 = 1; \quad 0 \le \alpha \le 1, \quad 0 \le \beta \le 1 \tag{5}$$

As mentioned in Eq. (5), the qubit is denoted as linear superposition of two states: state "0", and state "1". For instance, one and two qubit systems perform the operation on two and four values. Thus, $n$-qubit performs the operation on $2n$ values. So, quantum bit individual contains a string of $q$ quantum bits. Let us consider the example of two quantum bits that are represented in Eqs. (6) and (7).

$$Q = \left\langle \begin{matrix} 1/\sqrt{2}|1/\sqrt{2} \\ 1/\sqrt{2}|1/\sqrt{2} \end{matrix} \right\rangle \tag{6}$$

$$Q = \left( \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} \langle 00 \rangle + \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} \langle 01 \rangle + \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} \langle 10 \rangle + \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} \langle 11 \rangle \right) \tag{7}$$

After identifying the exact clusters, the concept of quantum bits' representation is used for achieving the global optimization in PFCM. Here, PFCM depends on the reduction in the objective function that is mathematically represented in Eq. (8), (9), and (10).

$$J_{\text{PFCM}}(U, T, V) = \sum_{i=1}^{C} \sum_{j=1}^{n} (u_{ij}^{m} + t^{n}) d^2(x_j, v_i) \tag{8}$$

where

$$\sum_{i=1}^{c} \mu_{ij} = 1, \quad \forall j \in \{1, \dots, n\} \tag{9}$$

$$\sum\nolimits_{j=1}^{n} t_{ij} = 1, \quad \forall i \in \{1, \dots, c\} \tag{10}$$

where $T$ is denoted as typicality matrix, $J_{\text{PFCM}}$ is indicated as objective function, $U$ is denoted as partition matrix, and $V$ is represented as vector of cluster centres. In this work, the objective function is evaluated by utilizing the number of cluster centres and the degree of membership, as mathematically denoted in Eqs. (11)–(13).

$$\mu_{ij} = \left[ \sum_{k=1}^{c} \left( \frac{dx_j, v_i}{dx_j, v_k} \right)^{\frac{2}{m-1}} \right]^{-1}, 1 \le i \le c, \quad 1 \le j \le n \tag{11}$$

$$t_{ij} = \left[ \sum_{k=1}^{n} \left( \frac{dx_j, v_i}{dx_j, v_k} \right)^{\frac{2}{n-1}} \right]^{-1}, 1 \le i \le c, \quad 1 \le j \le n \tag{12}$$

$$v_i = \frac{\sum_{k=1}^{n} \left( u_{ik}^{m} + t_{ik}^{n} \right) x_k}{\sum_{k=1}^{n} \left( u_{ik}^{m} + t_{ik}^{n} \right)}, \quad 1 \le i \le c \tag{13}$$

where $n$ is indicated as number of data points that are described by the coordinates $(x_j, v_i)$, which are utilized for calculating the distance between cluster centres and datasets, and $c$ is stated as number of cluster centres. Modified PFCM clustering algorithm creates memberships and possibilities with normal cluster centres and prototypes of each cluster [37]. Here, selecting the objective function is the crucial aspect for enhancing the performance of clustering method. In this paper, the clustering performance is evaluated based on the objective function. In order to develop an effective objective function, the following requirements are considered:

- The distance between the clusters needs to be diminished.
- The distance between the data points (assigned in the clusters) needs to be decreased.

In addition, the objective function of modified PFCM clustering algorithm is enhanced by using driven prototype learning of parameter $\alpha$. The learning mechanism of $\alpha$ is dependent between the clusters, which is updated at each iteration. The parameter $\alpha$ is mathematically stated in Eq. (14).

$$\alpha = \exp\left( -\min_{i \ne k} \frac{\| v_i - v_k \|^2}{\beta} \right) \tag{14}$$

where $\beta$ is signified as the sample variance that is denoted in Eq. (15).

$$\beta = \frac{\sum_{j=1}^{n} \| x_j - \bar{x} \|^2}{n} \tag{15}$$

where $\bar{x} = \frac{\sum_{j=n}^{n} x_j}{n}$

Then, a weighting parameter is developed for calculating the value of $\alpha$. Every point of the dataset comprises of a weight in relationship with every cluster. A better classification result is attained by using the weighting parameter, especially in the case of noisy data. Formula of weighting parameter is given in Eq. (16).

$$w_{ji} = \exp\left(-\frac{\| x_j - v_i \|}{\left(\sum_{j=1}^{n} \| x_j - \bar{v} \|^2\right) \times c/n}\right) \tag{16}$$

where $w_{ji}$ is stated as weight function of the point $j$ with the class $i$. The working procedure of modified PFCM is shown in Fig. 2 and also effectively explained below.

- *Initialization* Initially, the number of clusters is assumed by the user on the basis of quantum-inspired evolutionary algorithm.
- *Estimation of similarity distance* After assuming the number of clusters, evaluate the distance between data points and centroids for each and every segment.
- *Estimation of typicality matrix* After calculating distance matrix, typicality matrices are evaluated, which are attained from the modified PFCM clustering algorithm.
- *Estimation of membership matrix* Estimate membership matrix $M_{ik}$ by evaluating the membership value of data point that are collected from the modified PFCM.
- *Update centroid* After generating the clusters, centroid modernization is updated. After extracting the keywords, data classification is carried out by using SMA-CNN classifier.

### 4.5 Classification of data using SMA-CNN classifier

Generally, SMA-CNN is a multi-layer feed-forward network, which is designed to recognize the features in the sentiment data. The proposed SMA-CNN classifier contains

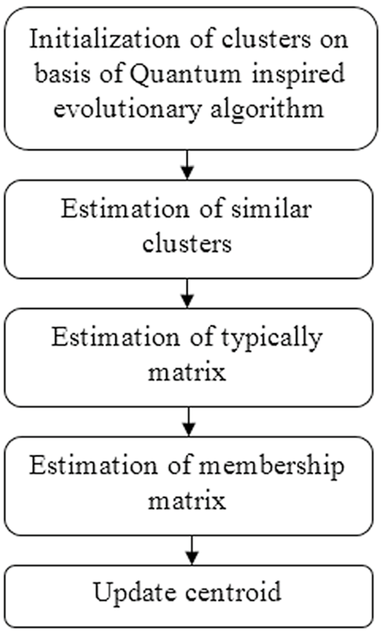


**Fig. 2** Working procedure of modified PFCM

seven layers: three convolutional layers, two LSTM layers, and two dense layers. The neurons in SMA-CNN classifier consider a small portion of the data that are named as sub-data [38]. Then, the respective sub-data are used for feature extraction, for instance, a feature may be a vertical line, arch, or circle. The features are captured by the respective feature maps of the network. A combination of features is utilized to classify the data. In addition, multiple different feature maps are used to make the network more robust. In this research, an optimization algorithm Adam optimization algorithm is used to optimize the moment of features values from one layer to another layer. Due to this action, the unwanted convolutions happening in the convolutional layer are avoided. A few major advantages of Adam optimization algorithm are: works effectively even with a little tuning of hyper-parameters and relatively low memory requirements [43]. Figure 3 represents the general architecture of the SMA-CNN classifier.

The convolutional layer is a primary layer in CNN classifier, which extracts the local information of the data. Moreover, convolutional operation improves the input features and reduces noise interference. The mapping operation in the convolution process is mathematically expressed in Eq. (17).

$$x_j^l f_c\left(\sum_{i\in M_j} x_i^{l-1}\times k_{i,j}^l + \theta_j^l\right) \tag{17}$$

where $x_j^l$ is specified as the $j$th mapping set of convolutional layer $l$, $x_i^{l-1}$ is denoted as the $i$th feature set indicating in the ($l$-1) convolutional layer, and $k_{i,j}^l$ is indicated as the convolutional kernel between the $i$th feature set and $j$th mapping set in the convolutional layer $l$. The variable $\theta_j^l$ is represented as bias and $f_c$ is denoted as activation function. The next step is the pooling process, which reduces the possibility of over-fitting during training process. The pooling process is mathematically denoted in Eq. (18).

$$x_j^l = f_p\left(\beta_j^l \text{down}\left(x_i^{l-1}\right) + \theta_j^l\right) \tag{18}$$

where down(·) is represented as the downsampling approach from layer ($l$-1) to layer $l$th, $\theta_j^l$ and $\beta_j^l$ are indicated as the additive bias and multiplicative bias, and $f_p(\cdot)$ is represented as the activation function. Generally, the pooling process is sub-divided into two types such as, average and maximum pooling. The final pooling layer (matrix features) are arranged to form a rasterization layer, which is further connected with the fully connected layer. The output of node $j$ is mathematically stated in Eq. (19).

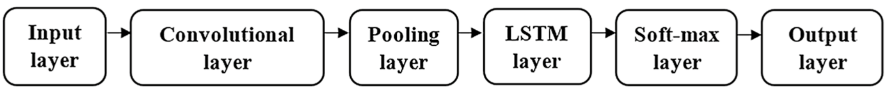


**Fig. 3** General architecture of SMA-CNN classifier

$$h_j = f_h\left(\sum_{i=0}^{n-1} w_{i,j}x_i - \theta_j\right) \tag{19}$$

where $w_{i,j}$ is denoted as the connection weight of input vector $x_i$, $\theta_j$ is stated as the node threshold, and $f_h(\cdot)$ is represented as the activation function. The next layer is the LSTM layer that helps to capture the sequential data by considering the prior data. This layer considers the output vectors from the pooling layer as inputs. The LSTM layer has a number of cells or units and the input of every cell is the output from the pooling layer. The final output of this layer has a similar number of units in the network.

If the LSTM layer deals with the multiclass issue, the softmax classifier is utilized in the fully connected layer. The loss function of softmax classifier is denoted in Eq. (20).

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k} l\{y^{(i)} = j\} \log \frac{e^{\theta_j^l}}{\sum_k e^{\theta_k^l}}\right] \tag{20}$$

where $e^{\theta_j^l}$ is represented as the input of $j$th neuron in the $l$ layer, $\sum_k e^{\theta_k^l}$ is denoted as the input of all the neurons, $\frac{e^{\theta_j^l}}{\sum_k e^{\theta_k^l}}$ is indicated as the output of $j$th neuron, $e$ is stated as the constant, and $l(.)$ is represented as the indictor function. If the value in the brace is true, the result of the indicator function is one. If the value in the brace is false, the result of the indicator function is zero. Then, add the rule items in $J(\theta)$ to prevent from falling into local optimum. The loss function of softmax classifier $J(\theta)$ after adding the rule items is mathematically stated in Eq. (21).

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k} l\{y^{(i)} = j\} log \frac{e^{\theta_j^l}}{\sum_k e^{\theta_k^l}}\right] + \frac{\rho}{2}\sum_{i=1}^{k}\sum_{j=0}^{n}\theta_{ij}^2 \tag{21}$$

where $\frac{\rho}{2}\sum_{i=1}^{k}\sum_{j=0}^{n}\theta_{ij}^2$ is represented as the weighted term that helps to stabilize the excessive parameters in the training set. In proposed classifier, each layer contains a reluctant linear unit (ReLU) activation function for activating the neurons in each layer. The ReLU activation function makes the proposed classifier (SMA-CNN) more redundant, because it effectively solves the exploding and vanishing gradient problems completely. ReLU activation function initially calculates the prediction error and then estimates the gradients utilized to update each weight in the network, so that less error is accomplished next time. At last, the obtained results in three forms are positive, negative, and neutral. Pseudo-code of proposed SMA-CNN is determined below.

### 4.5.1 Pseudo-code of SMA-CNN classifier

```
Read dataset
Remove stop words
Perform stemming and lemmatization
Get words count
Check if word count < threshold
          Remove word.
          Check if threshold < helpfulness ratio
                    Fake Review
          Else:
                    Real Review
Else:
          Get usage ratio
          Convert word to int
          Initialize layers of SMA-CNN
          Train SMA-CNN
          Get summarized text
          Calculate sentiment polarity of summarized text
Classify positive/negative/ neutral review
```

## 5 Experimental result and discussion

This section details the experimental result and discussion of the proposed system and also explains the performance metric, experimental setup, quantitative, and comparative analyses. The proposed system was implemented using Python with 4 GB RAM, 1 TB hard disc, 3.0 GHz Intel i5 processor. The performance of the proposed system was compared with other classification methods and existing research papers based on the amazon customer review dataset for assessing the efficiency of the proposed system. The performance of the proposed system was evaluated in light of recall, classification accuracy, precision, f-measure, and AUC.

### 5.1 Performance measure

Performance measure is the procedure of collecting, reporting, and analysing information about the performance of a group or individual. Mathematical equation of accuracy, f-measure, precision, and recall is denoted in Eqs. (22)–(25).

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \times 100 \tag{22}$$

$$F - \text{measure} = \frac{2\text{TP}}{(2\text{TP} + \text{FP} + \text{FN})} \times 100 \quad (23)$$

$$\text{Precision} = \frac{\text{TP}}{(\text{FP} + \text{TP})} \times 100 \quad (24)$$

$$\text{Recall} = \frac{\text{TP}}{(\text{FN} + \text{TP})} \times 100 \quad (25)$$

where TP is signified as true positive, TN is indicated as true negative, FP is specified as false positive, and FN is indicated as false negative.

### 5.2 Quantitative analysis

Amazon customer review dataset is used for evaluating the performance of the proposed system and other existing classification approaches like random forest, decision tree, and Naive Bayes. In this research study, the collected data are classified into three forms: positive, negative, and neutral classes. In Tables 4, 5, 6, and 7, the performance evaluation of the proposed system and existing classification approaches are evaluated in terms of accuracy, recall, precision, f-measure, and AUC. Here, the performance evaluation is validated with 80% training of data and 20% testing of data. Among 2,441,053 amazon products, eight products are considered for experimental investigation such as, amazon instant video, books, electronics, home and kitchen, movie review, media, kindle, and camera.

Here, Tables 4 and 5 show the performance investigation of proposed and existing classification methods for four amazon products: amazon instant video, books, electronics, and home and kitchen. The average classification accuracy of the proposed classifier (SMA-CNN with FCM) is 90.32%, and the existing classification approaches (random forest, decision, and Naive Bayes) achieved 72.30%, 71.61%, and 82.78% of classification accuracy. Similarly, the average classification accuracy of the proposed classifier (SMA-CNN with modified KFCM) is 92.85%, and the existing classification approaches (random forest, decision, and Naive Bayes) achieved 75.44%, 73.85, and 85.76% of classification accuracy. Correspondingly, the average recall, precision, f-measure, and area under curve of the proposed classifier are better than the existing classifiers with FCM, because the proposed system effectively calculates the linear and nonlinear properties of collected data and also significantly preserves the quantitative relationship between the high- and low-level features. The graphical comparison of the proposed approach with dissimilar classifiers for the amazon products, amazon instant video, books, electronics, and home and kitchen is represented in Figs. 4 and 5.

In addition, the comparative study of proposed and existing classification methods with FCM and modified KFCM is carried out for another four amazon products like a movie review, media, kindle, and camera. Here, the performance evaluation is validated with 80% training and 20% testing of data. Inspecting Tables 6 and 7, the proposed classifier (SMA-CNN) outperformed with the average classification accuracy

**Table 4** Performance analysis of proposed approach with dissimilar classifiers in light of recall, precision, and f-measure

| Classifiers | Keyword extraction | Class | Precision (%) | Recall (%) | F-measure (%) |
|---|---|---|---|---|---|
| Random forest | LDA–FCM | Amazon instant video | 68.15 | 65.90 | 65.89 |
| | | Books | 72.61 | 73.58 | 64.42 |
| | | Electronics | 68.96 | 71.46 | 68.33 |
| | | Home and kitchen | 66.96 | 65.25 | 67.30 |
| | LDA-modified PFCM | Amazon instant video | 74 | 67.43 | 69.76 |
| | | Books | 75.09 | 76.91 | 67.31 |
| | | Electronics | 73.34 | 73.42 | 71.11 |
| | | Home and kitchen | 70.98 | 68.34 | 70.61 |
| Decision tree | LDA–FCM | Amazon instant video | 68.01 | 58.10 | 82.37 |
| | | Books | 73.44 | 82.93 | 82.21 |
| | | Electronics | 68.99 | 68.74 | 77.83 |
| | | Home and kitchen | 65.71 | 65.88 | 78.57 |
| | LDA-modified PFCM | Amazon instant video | 79.74 | 62.10 | 83.88 |
| | | Books | 81.56 | 86.32 | 84.74 |
| | | Electronics | 78.72 | 70.02 | 79 |
| | | Home and kitchen | 70 | 67 | 79.94 |
| Naive Bayes | LDA–FCM | Amazon instant video | 66.14 | 84.42 | 82.90 |
| | | Books | 66.19 | 78.68 | 78.96 |
| | | Electronics | 65.06 | 76.56 | 84.91 |
| | | Home and kitchen | 66.41 | 74.87 | 86.90 |
| | LDA-modified PFCM | Amazon instant video | 81.23 | 88 | 86.45 |
| | | Books | 76.45 | 81.24 | 80.80 |
| | | Electronics | 77 | 77.76 | 86.87 |
| | | Home and kitchen | 89.98 | 77.02 | 89 |
| SMA-CNN | LDA–FCM | Amazon instant video | 85.99 | 90.54 | 94.09 |
| | | Books | 88.85 | 89.98 | 93.69 |
| | | Electronics | 88.53 | 90.40 | 93.66 |
| | | Home and kitchen | 91.41 | 93.87 | 86.82 |
| | LDA-modified PFCM | Amazon instant video | 90 | 93.44 | 96.78 |
| | | Books | 91.11 | 92.49 | 94.80 |
| | | Electronics | 92.67 | 94 | 96.67 |
| | | Home and kitchen | 93.33 | 95.55 | 90.62 |

**Table 5** Performance analysis of proposed approach with dissimilar classifiers in light of AUC and accuracy

| Classifiers | Keyword extraction | Class | AUC (%) | Accuracy (%) |
|---|---|---|---|---|
| Random forest | LDA–FCM | Amazon instant video | 52.58 | 75.19 |
| | | Books | 57.79 | 73.86 |
| | | Electronics | 55.79 | 68.53 |
| | | Home and kitchen | 68.87 | 71.63 |
| | LDA-modified PFCM | Amazon instant video | 56.02 | 78.09 |
| | | Books | 60.98 | 77 |
| | | Electronics | 56.80 | 72.24 |
| | | Home and kitchen | 70.07 | 74.44 |
| Decision tree | LDA–FCM | Amazon instant video | 61.11 | 67.08 |
| | | Books | 72.87 | 67.97 |
| | | Electronics | 64.91 | 71.76 |
| | | Home and kitchen | 67.96 | 79.64 |
| | LDA-modified PFCM | Amazon instant video | 62.34 | 69.78 |
| | | Books | 74.61 | 70.48 |
| | | Electronics | 68 | 73 |
| | | Home and kitchen | 69.023 | 82.17 |
| Naive Bayes | LDA–FCM | Amazon instant video | 77.28 | 86.35 |
| | | Books | 73.13 | 80.51 |
| | | Electronics | 70.39 | 80.76 |
| | | Home and kitchen | 81.48 | 83.52 |
| | LDA-modified PFCM | Amazon instant video | 79.87 | 88 |
| | | Books | 76.95 | 84.31 |
| | | Electronics | 73.60 | 83.90 |
| | | Home and kitchen | 83.65 | 86.86 |
| SMA-CNN | LDA–FCM | Amazon instant video | 76.73 | 88.32 |
| | | Books | 68.56 | 88.98 |
| | | Electronics | 78.63 | 95.11 |
| | | Home and kitchen | 77.29 | 88.88 |
| | LDA-modified PFCM | Amazon instant video | 78.75 | 90.65 |
| | | Books | 71.11 | 92 |
| | | Electronics | 80.64 | 96.87 |
| | | Home and kitchen | 79 | 91.9 |

of 92.8% as compared to the traditional classification methods: random forest, decision tree, and Naïve Bayes and existing clustering algorithm (FCM). In addition, the existing classifiers achieved minimum recall, precision, f-measure, accuracy, and AUC as related to the proposed classifier (SMA-CNN). In this research study, the computational time differs based on the data and the number of features in each and every review created by the user and product. So, the user cannot able to justify the stranded time scale for any text analysis mechanism. The graphical comparison of

**Table 6** Performance analysis of proposed approach with dissimilar classifiers by means of recall, precision, and f-measure

| Classifiers | Keyword extraction | Class | Precision (%) | Recall (%) | F-measure (%) |
|---|---|---|---|---|---|
| Random forest | LDA–FCM | Movie | 64.37 | 56.34 | 61.27 |
| | | Media | 40.87 | 59.27 | 59.23 |
| | | Kindle | 52.08 | 51.35 | 62.30 |
| | | Camera | 47.04 | 58.97 | 64.51 |
| | LDA-modified PFCM | Movie | 67.11 | 58.34 | 62.43 |
| | | Media | 44 | 60.78 | 61.34 |
| | | Kindle | 53.29 | 52.80 | 66 |
| | | Camera | 49.967 | 60 | 67.554 |
| Decision tree | LDA–FCM | Movie | 64.99 | 56.42 | 75.09 |
| | | Media | 64.67 | 61.89 | 69.47 |
| | | Kindle | 60.78 | 66.88 | 76.71 |
| | | Camera | 69.61 | 75.77 | 76.35 |
| | LDA-modified PFCM | Movie | 68.80 | 60.36 | 77.72 |
| | | Media | 66.32 | 65.22 | 70.87 |
| | | Kindle | 64.71 | 70 | 79 |
| | | Camera | 70.94 | 77 | 79 |
| Naive Bayes | LDA–FCM | Movie | 75.85 | 79.72 | 82.52 |
| | | Media | 71.84 | 82.15 | 76.95 |
| | | Kindle | 72.18 | 80.05 | 77.35 |
| | | Camera | 75.97 | 82.43 | 81.57 |
| | LDA-modified PFCM | Movie | 78.53 | 80.98 | 86.05 |
| | | Media | 73.90 | 84.23 | 79.78 |
| | | Kindle | 74 | 83 | 80.34 |
| | | Camera | 77.01 | 85.54 | 83.32 |
| SMA-CNN | LDA–FCM | Movie | 87.89 | 90.91 | 90.15 |
| | | Media | 88.41 | 92.70 | 91.80 |
| | | Kindle | 88.84 | 85.93 | 92.13 |
| | | Camera | 85.99 | 85.72 | 87.92 |
| | LDA–modified PFCM | Movie | 89.77 | 92 | 91.84 |
| | | Media | 90.41 | 94.27 | 94.50 |
| | | Kindle | 90 | 87.23 | 94.19 |
| | | Camera | 88 | 89.67 | 90.98 |

proposed and existing classifiers with modified KFCM for amazon products like a movie review, media, kindle, and the camera is shown in Figs. 6 and 7.

### 5.3 Comparative analysis

Comparative study on performance of existing works and the proposed work is given in Table 8. Han et al. [39] developed a new sentiment classification approach

**Table 7** Performance analysis of proposed approach with dissimilar classifiers by means of AUC and accuracy

| Classifiers | Keyword extraction | Class | AUC (%) | Accuracy (%) |
|---|---|---|---|---|
| Random forest | LDA–FCM | Movie | 53.68 | 54.90 |
| | | Media | 45.03 | 57.14 |
| | | Kindle | 50.35 | 60.34 |
| | | Camera | 57.34 | 62.56 |
| | LDA-modified PFCM | Movie | 54.90 | 58.12 |
| | | Media | 46.89 | 60.67 |
| | | Kindle | 51.67 | 62.98 |
| | | Camera | 58.92 | 65.57 |
| Decision tree | LDA–FCM | Movie | 65.78 | 72.97 |
| | | Media | 63.66 | 75.00 |
| | | Kindle | 65.14 | 73.73 |
| | | Camera | 69.95 | 74.18 |
| | LDA-modified PFCM | Movie | 67.52 | 76.90 |
| | | Media | 67 | 77 |
| | | Kindle | 69.12 | 75.119 |
| | | Camera | 73.57 | 78 |
| Naive Bayes | LDA–FCM | Movie | 77.34 | 84.05 |
| | | Media | 78.00 | 78.33 |
| | | Kindle | 85.79 | 79.79 |
| | | Camera | 83.33 | 86.73 |
| | LDA-modified PFCM | Movie | 80.25 | 87.6 |
| | | Media | 81.58 | 81.59 |
| | | Kindle | 86.91 | 83.33 |
| | | Camera | 84.47 | 89.907 |
| SMA-CNN | LDA–FCM | Movie | 87.13 | 89.79 |
| | | Media | 87.17 | 89.76 |
| | | Kindle | 88.25 | 92.18 |
| | | Camera | 89.68 | 91.46 |
| | LDA-modified PFCM | Movie | 89.38 | 92.68 |
| | | Media | 88.30 | 91.50 |
| | | Kindle | 92.20 | 93.55 |
| | | Camera | 90.81 | 94 |

(SentiWordNet (SWN)), which was utilized as the experimental sentiment lexicon, and then reviewed the data of four amazon products, which were collected from amazon customer review dataset. The experimental results showed that the bias processing strategy reduced the polarity bias rate and improved the performance of lexicon-based sentiment analysis. The developed algorithm achieved 69.79% of accuracy for DVD product, 68.72% of accuracy for electronics product, 68.17% of accuracy for books, and 71.41% of accuracy for kitchen products. Additionally, Liu

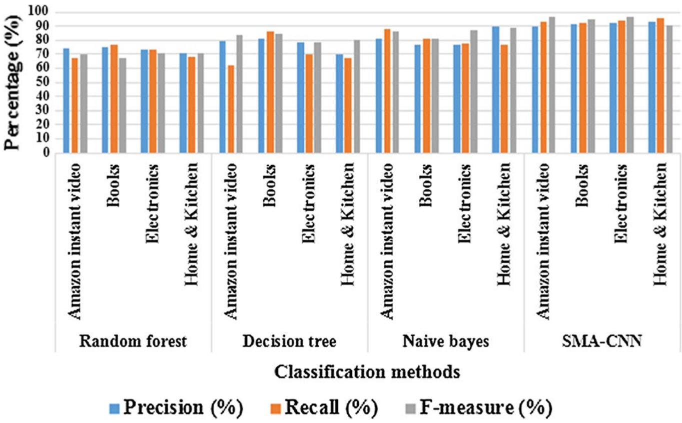


**Fig. 4** Graphical evaluation of proposed approach with dissimilar classifiers by means of recall, precision, and f-measure

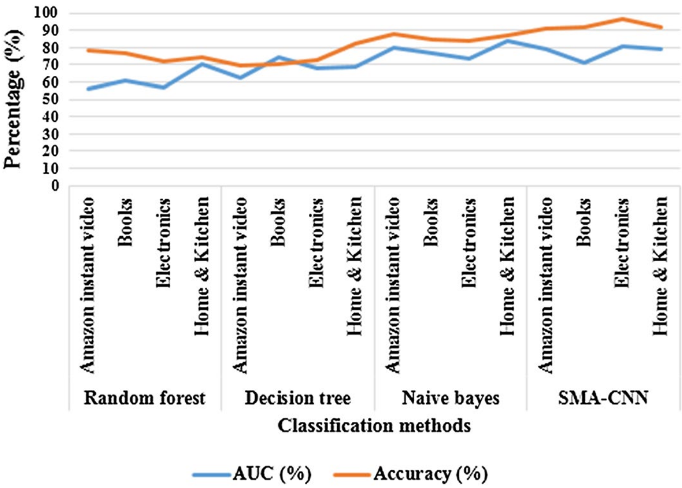


**Fig. 5** Graphical evaluation of proposed approach with dissimilar classifiers by means of AUC and accuracy

et al. [40] evaluated the scalability of Naive Bayes classifier in amazon customer review dataset. In this research paper, Naive Bayes classifier was used for achieving fine-grain control of the analysis process. This classifier achieved 82% of classification accuracy on movie reviews.

Rain [41] developed an effective algorithm for sentiment analysis. Initially, the features from the collected data were extracted by applying sentence length, bag of words, part of speech tags, spell checking, handling negation, and collocations. The extracted features were classified using Naive Bayes and decision list classifiers for tagging given reviews positive or negative. Their algorithm achieved 84% of the accuracy for book, 84% of accuracy of kindle product, and 79.93% of accuracy

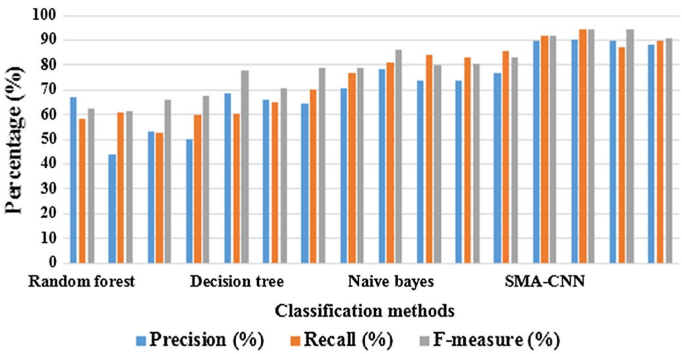


**Fig. 6** Graphical evaluation of proposed approach with dissimilar classifiers by means of recall, precision, and f-measure

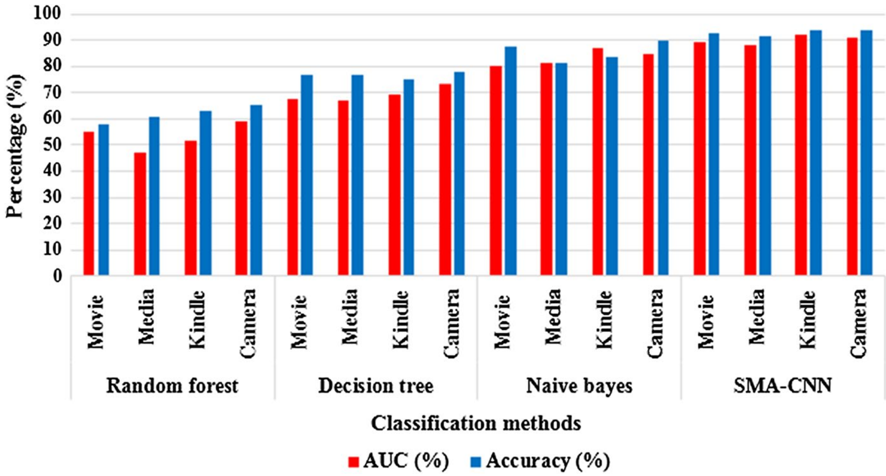


**Fig. 7** Graphical evaluation of proposed approach with dissimilar classifiers by means of AUC and accuracy

for media. In addition, A. Ghose, and Ipeirotis [42] examined the relative position of the three broad feature categories such as, "review readability", "review subjectivity", and "review-related" features, and also found the three feature sets results in a statistically equivalent performance by using all available features. Initially, the econometric, text mining, and predictive modelling methods were integrated towards a more complete analysis of the information collected by user-generated online reviews for valuing the economic impact. Finally, the random forest was used to classify the reviews of customers in three forms: neutral, negative, and positive. Their system achieved 78.79% of accuracy for DVD product, 87.57% of audio and video product, and 87.68% of accuracy for digital cameras. Compared to these existing papers, the proposed system achieved better performance that was almost 6–20% higher than the existing papers.

**Table 8** Comparative analysis of proposed and existing papers

| Method | Amazon products | Accuracy (%) |
|---|---|---|
| SentiWordNeT [39] | DVD | 69.79 |
| | Electronics | 68.72 |
| | Books | 68.17 |
| | Kitchen | 71.41 |
| Naive Bayes classifier [40] | Movie review | 82 |
| Naive Bayes and decision list classifiers [41] | Books | 84 |
| | Kindle | 84 |
| | Media | 79.93 |
| Random forest [42] | DVD | 78.79 |
| | Audio and video | 87.57 |
| | Digital camera | 87.68 |
| Proposed system (SMA-CNN) | DVD | 92.34 |
| | Electronics | 96.87 |
| | Books | 92 |
| | Kitchen | 91.90 |
| | Movie review | 92.68 |
| | Kindle | 93.55 |
| | Media | 91.50 |
| | Audio and video | 90.65 |
| | Digital camera | 94 |

## 6 Conclusion

In this research study, a new supervised system developed to classify the opinions of the customers for amazon selling products. The main aim of this experiment is to develop a proper keyword extraction method and classification approach for classifying the opinions of customers as neutral, negative, and positive forms using the amazon customer review dataset. In this scenario, a keyword extraction method (LDA) along with modified PFCM is used for selecting the appropriate keywords. The obtained keywords are classified using the classifier SMA-CNN. The development of an automated system for analysing the customer's opinion on amazon products has numerous advantages like able to handle multiple customers, effective in agent monitoring, track overall customer satisfaction, etc. Compared to the existing papers, the proposed system delivered an effective performance by means of quantitative analysis and comparative analysis. From the experimental analysis, the proposed system averagely achieved around 92.83% of classification accuracy, but the existing methodologies attained limited accuracy in the amazon customer review dataset. In future work, an effective unsupervised system will be developed in order to further improve the classification accuracy of sentiment analysis.

**Funding** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Compliance with ethical standards

**Conflict of interest** None.

## References

1. Hassan MK, Shakthi SP, Sasikala R (2017) Sentimental analysis of Amazon reviews using naïve bayes on laptop products with MongoDB and R. IOP Conf Ser Mater Sci Eng IOP Publ 263(4):042090
2. Glorot X, Bordes A, Bengio Y (2011) Domain adaptation for large-scale sentiment classification: A deep learning approach. In: Proceedings of the 28th International Conference on Machine Learning (ICML-11), pp 513–520
3. Deng ZH, Luo KH, Yu HL (2014) A study of supervised term weighting scheme for sentiment analysis. Expert Syst Appl 41(7):3506–3513
4. Fang X, Zhan J (2015) Sentiment analysis using product review data. J Big Data 2(1):5
5. Pandey AC, Rajpoot DS, Saraswat M (2017) Twitter sentiment analysis using hybrid cuckoo search method. Inf Process Manag 53(4):764–779
6. Ghiassi M, Lee S (2018) A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach. Expert Syst Appl 106:197–216
7. Alharbi ASM, de Doncker E (2019) Twitter sentiment analysis with a deep neural network: an enhanced approach using user behavioral information. Cogn Syst Res 54:50–61
8. Daniel M, Neves RF, Horta N (2017) Company event popularity for financial markets using Twitter and sentiment analysis. Expert Syst Appl 71:111–124
9. Abid F, Alam M, Yasir M, Li C (2019) Sentiment analysis through recurrent variants latterly on convolutional neural network of Twitter. Future Gener Comput Syst 95:292–308
10. Öztürk N, Ayvaz S (2018) Sentiment analysis on Twitter: a text mining approach to the Syrian refugee crisis. Telemat Inform 35(1):136–147
11. Singh T, Kumari M (2016) Role of text pre-processing in twitter sentiment analysis. Proc Comput Sci 89:549–554
12. Philander K, Zhong Y (2016) Twitter sentiment analysis: capturing sentiment from integrated resort tweets. Int J Hosp Manag 55:16–24
13. Schumaker RP, Jarmoszko AT, Labedz CS Jr (2016) Predicting wins and spread in the Premier League using a sentiment analysis of twitter. Decis Support Syst 88:76–84
14. Da Silva NF, Hruschka ER, Hruschka ER Jr (2014) Tweet sentiment analysis with classifier ensembles. Decis Support Syst 66:170–179
15. Wang Y, Sun L, Wang J, Zheng Y, Youn HY (2017) A novel feature-based text classification improving the accuracy of twitter sentiment analysis. Advances in computer science and ubiquitous computing. Springer, Singapore, pp 440–445
16. Le B, Nguyen H (2015) Twitter sentiment analysis using machine learning techniques. Advanced computational methods for knowledge engineering. Springer, Cham, pp 279–289
17. Jalaja G, Kavitha C (2019) Sentiment analysis for text extracted from Twitter. Integrated intelligent computing, communication and security. Springer, Singapore, pp 693–700
18. Yang M, Qu Q, Chen X, Guo C, Shen Y, Lei K (2018) Feature-enhanced attention network for target-dependent sentiment classification. Neurocomputing 307:91–97
19. Araque O, Corcuera-Platas I, Sanchez-Rada JF, Iglesias CA (2017) Enhancing deep learning sentiment analysis with ensemble techniques in social applications. Expert Syst Appl 77:236–246
20. Giatsoglou M, Vozalis MG, Diamantaras K, Vakali A, Sarigiannidis G, Chatzisavvas KC (2017) Sentiment analysis leveraging emotions and word embeddings. Expert Syst Appl 69:214–224
21. Alsinet T, Argelich J, Béjar R, Fernández C, Mateu C, Planes J (2018) An argumentative approach for discovering relevant opinions in Twitter with probabilistic valued relationships. Pattern Recogn Lett 105:191–199
22. Balahur A, Perea-Ortega JM (2015) Sentiment analysis system adaptation for multilingual processing: the case of tweets. Inf Process Manag 51(4):547–556
23. Bouazizi M, Ohtsuki T (2017) A pattern-based approach for multi-class sentiment analysis in twitter. IEEE Access 5:20617–20639

24. Yu D, Xu D, Wang D, Ni Z (2019) Hierarchical topic modeling of Twitter data for online analytical processing. IEEE Access 7:12373–12385
25. Bharathi S, Geetha A, Sathiynarayanan R (2017) Sentiment analysis of Twitter and RSS news feeds and its impact on stock market prediction. Int J Intell Eng Syst 10(6):68–77
26. Saif H, He Y, Fernandez M, Alani H (2016) Contextual semantics for sentiment analysis of Twitter. Inf Process Manage 52(1):5–19
27. Ren Y, Wang R, Ji D (2016) A topic-enhanced word embedding for Twitter sentiment classification. Inf Sci 369:188–198
28. Preethi PG, Uma V (2015) Temporal sentiment analysis and causal rules extraction from tweets for event prediction. Proc Comput Sci 48:84–89
29. Kumar KA, Rajasimha N, Reddy M, Rajanarayana A, Nadgir K (2015) Analysis of users' sentiments from kannada web documents. Proc Comput Sci 54:247–256
30. Amolik A, Jivane N, Bhandari M, Venkatesan M (2016) Twitter sentiment analysis of movie reviews using machine learning techniques. Int J Eng Technol 7(6):1–7
31. Cambria E, Poria S, Gelbukh A, Thelwall M (2017) Sentiment analysis is a big suitcase. IEEE Intell Syst 32(6):74–80
32. Ghorbel H, Jacot D (2011) Sentiment analysis of French movie reviews. Advances in distributed agent-based retrieval tools. Springer, Berlin, pp 97–108
33. Boyd-Graber J, Resnik P (2010) Holistic sentiment analysis across languages: Multilingual supervised latent Dirichlet allocation. In: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, pp 45–55
34. Colace F, De Santo M, Greco L, Moscato V, Picariello A (2016) Probabilistic approaches for sentiment analysis: latent dirichlet allocation for ontology building and sentiment extraction. Sentiment analysis and ontology engineering. Springer, Cham, pp 75–91
35. Patel OP, Bharill N, Tiwari A (2015) A quantum-inspired fuzzy based evolutionary algorithm for data clustering. In: IEEE International Conference on FUZZY SYSTEMS (FUZZ-IEEE), pp 1–8
36. Chakhmakhchyan L, Cerf NJ, Garcia-Patron R (2017) Quantum-inspired algorithm for estimating the permanent of positive semi definite matrices. Phys Rev A 96(2):022329
37. Trupthi M, Pabboju S, Narsimha G (2018) Possibilistic fuzzy c-means topic modelling for twitter sentiment analysis. Int J Intell Eng Syst 11(3):100–108
38. Alayba AM, Palade V, England M, Iqbal R (2018) A combined CNN and LSTM model for arabic sentiment analysis. In: International Cross-Domain Conference for Machine Learning and Knowledge Extraction, Springer, Cham, pp 179–191
39. Han H, Zhang Y, Zhang J, Yang J, Zou X (2018) Improving the performance of lexicon-based review sentiment analysis method by reducing additional introduced sentiment bias. PLOS One 13(8):e0202523
40. Liu B, Blasch E, Chen Y, Shen D, Chen G (2013) Scalable sentiment classification for big data analysis using naive bayes classifier. In 2013 IEEE International Conference on Big Data, pp 99–104
41. Rain C (2013) Sentiment analysis in amazon reviews using probabilistic machine learning, Swarthmore College
42. Ghose A, Ipeirotis PG (2011) Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. IEEE Trans Knowl Data Eng 23(10):1498–1512
43. Balles L, Hennig P (2017) Dissecting adam: the sign, magnitude and variance of stochastic gradients. *arXiv preprint* arXiv:1705.07774

## Affiliations


### Trupthi Mandhula[1] · Suresh Pabboju[2] · Narsimha Gugulotu[3]

Suresh Pabboju
Plpsuresh@gmail.com

Narsimha Gugulotu
trupthi@cbit.ac.in

[1] Information Technology, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana 75, India

[2] Deptartment of Information Technology, CBIT, Gandipet, Hyderabad 75, India

[3] Computer Science Engineering, College of Engineering, JNTUH, Hyderabad 85, India