Received December 23, 2020, accepted January 14, 2021, date of publication January 29, 2021, date of current version February 11, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3055507

Spider Monkey Crow Optimization Algorithm With Deep Learning for Sentiment Classification and Information Retrieval

AARTI CHUGH[1], VIVEK KUMAR SHARMA[2], SANDEEP KUMAR[3], ANAND NAYYAR[4,5], (Senior Member, IEEE), BASIT QURESHI[6], (Member, IEEE), MANJOT KAUR BHATIA[7], AND CHARU JAIN[1]

[1]Department of Computer Science and Engineering, Amity School of Engineering and Technology, Amity University, Gurugram 122413, India
[2]Faculty of Engineering and Technology, Jagannath University, Jaipur 303901, India
[3]Department of Computer Science and Engineering, CHRIST (Deemed to be University), Bangalore 560074, India
[4]Graduate School, Duy Tan University, Da Nang 550000, Vietnam
[5]Faculty of Information Technology, Duy Tan University, Da Nang 550000, Vietnam
[6]Department of Computer Science, Prince Sultan University, Riyadh 11586, Saudi Arabia
[7]Department of Computer Science, Jagan Institute of Management Studies, Delhi 110085, India

Corresponding authors: Anand Nayyar (anandnayyar@duytan.edu.vn) and Basit Qureshi (qureshi@psu.edu.sa)

This work was supported in part by the Robotics and Internet of Things Laboratory, Prince Sultan University.

ABSTRACT The epidemic increase in online reviews' growth made the sentiment classification a fascinating domain in academic and industrial research. The reviews assist several domains, which is complicated to gather annotated training data. Several sentiment classification methodologies are devised for performing the sentiment analysis, but retrieval of information is not accurately performed, less effective, and less convergence speed. In this paper, we propose a sentiment paper proposes a sentiment classification model, namely Spider Monkey Crow Optimization algorithm (SMCA), for training the deep recurrent neural network (DeepRNN). In this method, the telecom review is employed to remove stop words and stemming to eliminate inappropriate data to minimize user's seeking time. Meanwhile, the feature extraction is performed using SentiWordNet to derive the sentiments from the reviews. The extracted SentiWordNet features and other features, like elongated words, punctuation, hashtag, and numerical values, are employed in the DeepRNN for classifying sentiments. To retrieve the required review, the Fuzzy K-Nearest neighbor (Fuzzy-KNN) is employed to retrieve the review based on a distance measure. With rigorous assessments and experimentation, it is observed that the proposed SMCA-based DeepRNN performs better in terms of accuracy of 97.7%, precision of 95.5%, recall of 94.6%, and F1-score 96.7%, respectively.

INDEX TERMS Sentiment classification, deep recurrent neural network, information retrieval, SentiWordNet, fuzzy K-nearest neighbours.

I. INTRODUCTION

Social media has emerged due to the usage trends and innovative services offered by the prevalent social networking sites. Social media's impending role is considered an expediter that enhances the learning and analysis of massive data [1]. The development of web-based techniques is unparalleled, wherein social media technologies have to turn into an essential component in everyday lives. The accessibility of social media is no more ostensible than at the universities, wherein the technologies are transforming the ways students interact and collaborate [2]. The processor in the telecom system generates precise responses within a particular time interval.

Moreover, the trades are not processed in the same order as they occur, but they are processed according to certain priorities. The needs on the quality, usability, reliability, and availability of these models are very rigorous [3]. Generally, the software based on telecom is incessantly upgraded due to the long era of manipulation. To fulfill customer requirements, the software packages are produced, delivered, and specified rapidly in a precise manner. Recently, the ISP containing system downtime is highly important for sales argument. Telecom software development's augmentations should be carefully controlled, predictable, planned, and

The associate editor coordinating the review of this manuscript and approving it for publication was Sabah Mohammed.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

measurable [3]. However, it is essential to briefly study the classification for detecting the sentiments hidden in the post [4].

Sentiment analysis is extensively categorized into ternary sentiment classification, binary sentiment classification. The purpose of classifying Twitter sentiment is to inevitably find the divergence of sentiment in tweets as negative or positive. Twitter sentiment classification monitors the machine learning techniques for building a classifier from tweets using polarity labels of manually annotated sentiments [5]. Munjal *et al.* developed a framework for predicting twitter sentiments [6] and also proposed a model for opinion dynamics [7], [8]. The classification of multi-class sentiment determines all the sentiments contained in the provided tweets and attributes with different scores considering the sentiments portraying the weight or to determine how pertinent they are considered in a tweet. The classification of multi-class sentiment is based on the writing patterns considering the set of features to classify ternary sentiment using the tweets. The scores of the sentiment analysis are employed to filter the sentiments using the tweets [9]. A model, namely GloVe-Deep Convolution Neural Network (DCNN), is devised to classify the sentiment. Recently, the deep learning methods gained interest, which significantly improved the classification accuracy. The sentiment information is represented as word representation and words embedding that concatenated the feature of prior polarity score. The feature sets are integrated and subjected in the DCNN to predict the sentiment labels using the tweets obtained from the social media that helped in information retrieval methods [5].

Information retrieval (IR) is an extensive and interdisciplinary exploration that contained relevance rankings, search engines, and information indexing. Moreover, the evaluation measures used for sentiment classification involved recall and precision. IR is concerned with determining the critical information based on user queries [1]. The IR arena also covered supporting users in filtering or browsing data or further process the set of retrieved data [10]. The domain of information retrieval has emanated a long way and facilitated the more comfortable and fast discovery of information [11]. The purpose of IR is to counterpart the user's requirements for acquiring the required information [1]. To compute information retrieval effectiveness in a standardized way, a test collection with three things should be mainly considered. The first is the collection of documents; the second is the generation of queries using test suites, and the third is the set of relevant judgments wherein a binary assessment of either non-relevant or relevant for each pair of query-document [10]. IR integrates the space and dimensions for dealing with classical retrieval issues. Besides, the queries required a collective learning system for determining the matching between the spatial domain and thematic domain.

Moreover, it is essential to provide a match between the user's query and data. The relative incompetence of IR models is mainly caused by the inaccuracies using a query [1]. The field of attacking may cause serious issues that users face in the information-hidden world [12]. Due to the exponential growth of information, the IR can play a progressively imperative role in the future [11].

In this research, proposed SMCA-based DeepRNN is employed for classifying the sentiment from the telecom reviews. Here, the pre-processing is adapted to eliminate irrelevant reviews and reduce the seeking time using stemming and removing stop words. The mining of features is carried out with SentiWordNet, wherein positive score, negative score, polarity, and subjectivity features are extracted with other features, like elongated words, punctuation, hashtag, and numerical values. The DeepRNN is employed for classifying the sentiments using the proposed SMCA, which is designed by integrating SMO and CSA. The sorting of sentiments manually is a complex, time-consuming, and expensive task. Thus, employing automatic sentiment classification classifiers allows making sense of these issues to yield effective insights for sentiment classification. Moreover, the classifier's training with proposed SMCA assists analysis of trends considering user sentiments from the accumulated reviews.

The sentiments analysis is a trending domain in the real-time applications for classifying the sentiments using telecom reviews. However, the primary issue in analyzing the sentiment is faced while classifying the polarity of specified text in sentences, data, or features to predict the uttered opinion. **Thus, this sentiment classification process must be automated with the adaption of a novel technology that categorizes positive and negative reviews to help users make effective decisions.**

The major contribution of the paper includes:

- A new approach proposed by hybridizing SMO and CSA and named SMCA.
- The SMCA based DeepRNN is used for sentiment classification and tested over two datasets.
- The FuzzyKNN is used for information retrieval from Amazon unlocked mobile reviews dataset and telecom tweets.
- The newly proposed approach outperformed considered existing approaches and achieved 97.7% accuracy, 95.5% precision, 94.6% recall, and 96.7% F1-score.

The paper is assembled subsequently: Section II portrays the assessment of conventional strategies and their challenges available in the literature. Section IV deliberates the proposed sentiment classification framework along with information retrieval. Section V elaborates the proposed method with conventional strategies, and section VI concludes the paper with future scope.

II. LITERATURE REVIEW

The eight classical strategies based on sentiment classification are portrayed below, along with their limitations. Jianqiang and Xiaolin [5] employed DCNN for generating word embedding based on unsupervised learning on massive twitter corpora, which employed relative semantic relations using tweets. These word embedding were combined using the polarity of sentiment score for performing a sentiment feature using the given tweets. The features set were combined with DCNN for training and computing the labels of classified sentiments. However, the inherent relation between the semantics and improved feature information amongst words present in tweets were not extracted. Bouazizi and Ohtsuki [4] developed a Multi-class sentiment approach method using ternary and binary classifications. The method was scalable and can be implemented for classifying the texts into different classes. The technique was utilized to choose features for implementing the sentiment classification.

Ducange *et al.* [13] modeled a decision-making system for efficiently supporting managing the companies for promoting the marketing campaigns. The DSS computes brands' reputation and offers feedback considering the digital marketing campaign with the reviews. The method estimated the sentiments of users based on positive, negative, or neutral polarities. However, the method failed to embed a sentiment analysis engine for computing the sentiments. Ohtsuki [9] devised a Sentiment Quantification approach for analyzing the sentiments and for mining the opinions. This method referred to the automatic detection of opinions by evaluating the posts and publications. Here, the multi-class sentiment analysis was devised to identify the user's exact sentiment using sentiment polarities. The method used a conventional multi-class classification using the Twitter datasets. The method was employed for automatically attributing different scores for each sentiment and selecting the highest score for the analysis. However, the sentiments considering different polarities were not addressed.

Li *et al.* [14] developed a probabilistic model, namely Social Working Memory (SWM), for sentiment classification. Initially, a semantic hierarchy was established using long-term memory for encoding personal information. The method employed self-regulation for performing adaptation using human personality. Two learning techniques were devised for training the probabilistic SWM model. However, the method failed to integrate social cognitive skills into the social functions. Gao *et al.* [15] designed an interval-at-a-time (IAAT) model for classifying the sentiments. Moreover, the method relied on the inverted index that made it well-suited with existing search engines. The method was fast due to fast insert and appends operations, but the method was ineffective when different retrieval constraints were considered.

Mata-Rivera *et al.* [1] designed Geographic Information Retrieval (GIR) method on the basis of three matching queries using ontologies of heterogeneous data sources. Here, the technique to imitate contextualization of queries was devised to control heterogeneous data using the societal network. This method was considered a collaborative learning system devised based on user's experience level with many domains. However, the method failed to offer expertise levels by interacting with users.

Virmani *et al.* [16] devised a query processing model for retrieving the information using a social network that extracted the user's intent considering different social networks. Here, the machine learning techniques, ranking algorithm, and Latent Dirichlet Algorithm (LDA) was adapted to enhance the results generated by each query and obtained the information using the data mining method. The model exclusively contributed to a user-centric query retrieval model based on of natural language with temporal metrics. The method increased users' determination to increase the business to collaboratively explore the promotions to determine a new set of networks. The method significantly improved the recall and precision, respectively, but the method was inapplicable to limited features.

A. CHALLENGES AND RESEARCH GAP

The challenges confronted by the conventional sentiment classification methodologies are illustrated below:

- In [17], a genetic algorithm-based feature reduction technique is devised to address scalability issues.
- It helped to precisely compute the views and public sentiments considering different domains that involved terrorism, global conflicts, and social problems. Even though the method minimizes the feature size, agencies' recommendations depending on user opinions are not produced.
- In [18], the inverse document frequency technique is designed for sharing short text messages like tweets. Here, multilingual tweets are accumulated using characteristics of corpus considering different languages. This method performed prominent information by sentiment analysis, but the clash between the tweets are not analyzed.
- The Convolutional neural networks are employed by incorporating the information of user behavioral information considering a given tweet. The method effectively attained the sentiment classification, but tweet classification was not attained considering other networks [19].
- In [20], the query expansion approach was designed for selecting the terms utilized in query reformulation. The method was employed for choosing terms utilized in query reformulation. This method is employed in the milieu of structured IR. However, the removal of irrelevant nodes in the process of term selection highly affects the classification performance.
- A feature selection method and the voting model are utilized for performing sentence-level sentiment analysis using classifiers. This method acquired improved accuracy with high entropies, but the labeled feature sets are incapable of transforming through the vectors [21].

III. RELATED TERMINOLOGIES

A. SPIDER MONKEY OPTIMIZATION ALGORITHM

The SMO algorithm has six steps based on global and local leaders' learning and decision activities while searching for food sources. After initialization, it performs the local and global leader phase. Now global leader and local leader update their position in the learning phase. Finally, a local leader and a global leader decide local optimal and fusion of subgroups.

The SMO algorithm starts with a swarm of randomly initialized N spider monkeys in the D dimension. Local leaders and global leader phases follow the initialization to update individuals by utilizing local and global leaders' knowledge. In the next phase, SMO follows learning by the global and local leaders to update their position. Finally, a local leader and global leader take a decision. Detail of the SMO algorithm is given in Algorithm 1.

Algorithm 1 Spider Monkey Optimization [22]

Initialization of parameters: Population, *CR*, *GlobalLeaderLimit*, and *LocalLeaderLimit*.
Evaluate fitness of each solution.
Choose Leaders (Global and Local).
while meet the termination criteria **do**
  Update individuals in local groups.
  A greedy selection approach employed for all solutions based on fitness.
  Evaluate probability ($prob_i$) of selection for each member.
  engender new location for group member (identified based on $prob_i$) by utilizing experience of global leader and group members and self-experience.
  Apply greedy selection strategy to update the position of Global and Local Leader and group members.
  cycle = cycle + 1
  A particular group's member redirected for foraging if a local leader not able to update her position and exceeds *LocalLeaderLimit*.
  Global leader create new subgroups if she is not able to update her position.
end while Result the best solution detected so far.

The SMO algorithm is one of the most efficient swarm-based algorithms, and it is an updated number of ties for different applications. Sharma *et al.* [23] introduced the concept of aging in SMO and considered individuals' age. Kumar *et al.* introduced non-linear perturbation rate in SMO based on exponential function [24], chaotic function [25], and hyperopic function [26]. Major applications of SMO are engineering optimization [27], [28], antenna design [29], placement of capacitor [30], image segmentation [31], PIDA controller design [32], clustering [33] and many more. Sharma *et al.* [34] discussed working an example of SMO. A detailed study about the SMO algorithm is available in [35] with an improved variant and a new perturbation strategy in SMO proposed by Sharma *et al.* [36].

B. CROW SEARCH ALGORITHM

The CSA algorithm was developed by Askarzadeh [37] in 2016. It mimics the intelligent behavior of the crow. This algorithm is based on the crow's natural behavior while searching, stealing food, hiding excess food, and reiterating the location of food after a long time. They also warn group members if they detect some fear. Algorithm 2 illustrate pseudo code for CSA.

Algorithm 2 Crow Search Algorithm [37]

Initialization population of crow and memory for each crow.
Evaluate each solution.
while $iter < iter_{max}$ **do**
  for Repeat for each crow **do**
    Randomly select j^{th} one crow.
    Define an awareness probability (AP).
    if $AP^{j,iter} \leq r_j$ **then**
      $x^{i,iter+1} = x^{i,iter} + r_i \times fl^{i,iter} \times (m^{j,iter} - x^{i,iter})$
    else
      randomly selects one position.
    end if
  end for
  Evaluate new solution and select if highly fitted then existing one.
  Update memory.
end while

The CSA deployed for various optimization problems like feature selection [38], parkinson disease diagnosis [39], load dispatch [40], image thresholding [41], energy problems [42], engineering optimization [43], and many more.

IV. SMCA BASED DeepRNN FOR SENTIMENT CLASSIFICATION AND INFORMATION RETRIEVAL

In this section, the proposed SMCA-based DeepRNN is illustrated for performing sentiment classification and information retrieval over extensive telecom reviews using features. In this model, the information retrieval is incorporated to retrieve the required review from the set of reviews.

Figure 1 portrays a schematic view of the proposed SMCA-DeepRNN of sentiment classification with Fuzzy KNN based information retrieval. The model employs three phases that involves pre-processing, extraction of feature, and classification of sentiment to perform sentiment classification.

The pre-processing is carried out for handling reviews using sentiments, which facilitate information retrieval. Here, elimination of stop word and the process of stemming are employed to increase speed of sentiment classification. Initially, the keywords using the input telecom reviews are fed to pre-processing for eliminating the irrelevant words. Then, feature extraction is performed using SentiWordNet [44] for finding keywords of the document. Then, the sentiment classification is carried out using the SMCA-DeepRNN, and the last step is information retrieval, using fuzzyKNN by facilitating the matching between the new query and the classified data. The aim of the proposed SMCA-DeepRNN [45] is to select optimal weights to process boundary degree technique

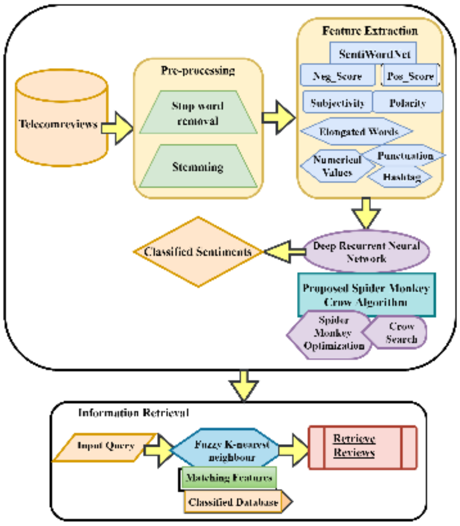


FIGURE 1. Schematic view of proposed SMCA-based DeepRNN for sentiment classification and FuzzyKNN for information retrieval.

and is devised by combining SMO [22] and CSA [37]. Thus, effective sentiment classification is performed using the proposed SMCA-DeepRNN, and information retrieval is carried out using the fuzzy KNN.

Assume an input data represented as D using different attributes (refer Eq. 1).

$$D = \{D_{gh}\}; (1 \leq g \leq K)(1 \leq h \leq L) \tag{1}$$

where, D_{gh} represent documents contained in database D with h^{th} attribute in g^{th} data. Here, K data points and L number of attributes are employed.

A. PRE-PROCESSING

Here, the pre-processing is adopted as an essential phase for effectively arranging massive documents. The pre-processing phase is employed to process the documents for obtaining improved representations. The pre-processing is performed for eliminating noise from the reviews by correcting spellings and eliminating ellipsis. Initially, the reviews are collected and accumulated in the relational database. In this phase, the removal of stop words and irrelevant words from reviews is done to maximize the proposed sentiment classification model's performance.

1) ELIMINATION OF STOP WORD

The stop word represents a word, which contains no data. Deleting stop words is a process in which stop words are eliminated using large-scale documents. The elimination of the stop word is used to save storage space and provide faster processing.

2) STEMMING

The process of stemming is adapted for transforming words to the stem. In massive sized data, many words are adapted which expresses similar notion. The noteworthy technique employed to reduce words to roots.

B. DRAWING SIGNIFICANT FEATURES FOR SENTIMENT CLASSIFICATION

This section portrays noteworthy features obtained from selected documents, and the implication of extracting features is to produce pertinent features for improved sentiment classification and information retrieval. The extraction of features is employed after pre-processing by eliminating keywords with SentiWordNet.

SentiWordNet [44] is employed as a lexical resource for extracting opinions. SentiWordNet allocates each synset of WordNet using three arithmetical scores, like positive Pos(s), negative Neg(s), and objective Obj(s). Different words have different polarities using lemma for identifying different word senses. SentiWordNet is the major tool for computing the score of specific words from the reviews. The SentiWordNet is employed to detect the polarity of the provided review. The utilization of speech tagger using tag words and search has maximized the performance and eliminated the stop words. The method determined the polarity of each attribute in a multi-aspect sentence. SentiWordNet provides the automatic annotation of synsets based on positivity, negativity, and neutrality. Thus, the SentiWordNet is employed to determine positivity, negativity, and neutrality are the words contained in the sysnet.

Other features, like several elongated words, several punctuation, hashtag, and numerical values, are extracted from the sentence for sentiment classification. Once the SentiWordNet is applied, the significant features and the sentence-based features are extracted, which is accumulated in the feature vector F, and the obtained feature vector is fed to the proposed SMCA based DeepRNN for useful sentiment classification.

C. SENTIMENT CLASSIFICATION USING SMCA-BASED DeepRNN

The proposed SMCA based DeepRNN is employed to categorize the sentiments of telecom data. Here, the SMCA based DeepRNN is devised by integrating SMCA in the Deep-RNN model to choose optimum weights present in the DeepRNN. The proposed SMCA is employed for optimizing DeepRNN [45] by selecting optimum weights. The structural design of DeepRNN and the process of training are portrayed below.

1) ARCHITECTURE OF DeepRNN

The features extracted using input telecom reviews are accumulated as a feature vector, S. This vector is passed to the DeepRNN classifier as an input. The Deep RNN [45] represents the network hierarchy of multiple repeated hidden layers. These hidden layers have recurrent connections among them. The classifier has an efficient capability to work with the varying lengths of input features derived from any information. It utilized the information of the preceding time step as a component in the present forecast, and processed the iteration with information of hidden states. Features are handled efficiently due to the recurrent nature of Deep RNN. The Deep RNN is the most favorable deep learning classifier amongst the available conventional methods because of sequential information patterns. The structural design of Deep RNN is portrayed in Fig. 2.

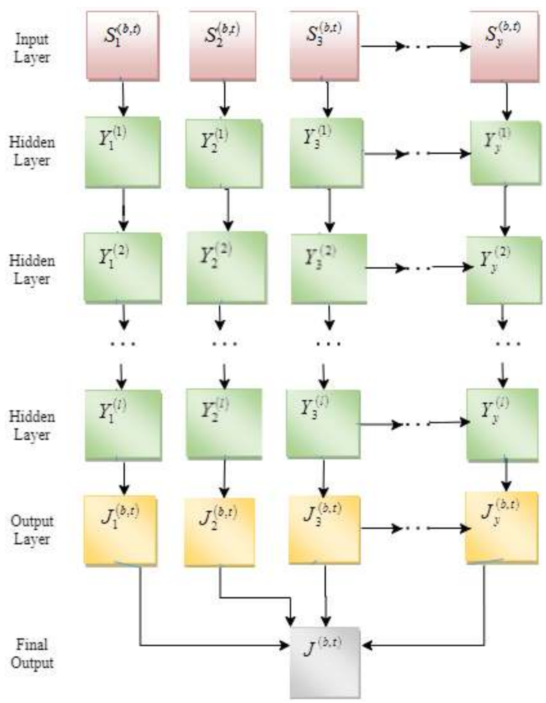


FIGURE 2. **Architecture of Deep RNN classifier.**

The organization of Deep RNN starts here from b^{th} layer at t^{th} time. The input feature vector is $S^{(b,t)} = \{S_1^{(b,t)}, S_2^{(b,t)}, \ldots S_i^{(b,t)}, \ldots S_y^{(b,t)}\}$ which produces output vector expressed as $J^{(b,t)} = \{J_1^{(b,t)}, J_2^{(b,t)}, \ldots J_i^{(b,t)}, \ldots J_y^{(b,t)}\}$. Here, each element of input and output vectors is called as unit. In this classifier, i represent random unit of b^{th} layer, and y signifies total units of b^{th} layer. Moreover, random unit number and total units of $(b-1)^{th}$ layer is expressed as j and E. Thus, weight of input transmission varies from $(b-1)^{th}$ layer to b^{th} layer is expressed as, $W^{(b)} \in A^{y \times E}$, and recurrent weight of b^{th} layer is expressed as $w^{(b)} \in A^{y \times y}$. Here, A symbolize weight set. Thus, the units of input vector is formulated using Eq. 2.

$$S_i^{(b,t)} = \sum_{z=1}^{E} p_{iz}^{(b)} J_z^{(b-1,t)} + \sum_{i'}^{y} x_{ii'}^{(b)} J_{i'}^{(b,t-1)} \tag{2}$$

where, $p_{iz}^{(b)}$ and $x_{ii'}^{(b)}$ indicate attributes of $W^{(b)}$ and $w^{(b)}$. i' represent random unit of b^{th} layer. The attribute of output vector in b^{th} layer is expressed by Eq. 3

$$J_i^{(b,t)} = \beta^{(b)}(S_i^{(b,t)}) \tag{3}$$

where, $\beta^{(b)}$ symbolize activation function. Different activation functions, such as sigmoid function as $\beta(S) = tanh(S)$, ReLU as $\beta(S) = max(S, 0)$, and logistic sigmoid function as $\beta(S) = \frac{1}{1+e^{-S}}$ exists and are used in many applications.

Further, 0^{th} weight as $p_{i0}^{(b)}$ and 0^{th} unit as $J_0^{(b-1,t)}$ are introduced to simplify working of model and finally, bias is expressed through Eq. 4 as below.

$$J^{(b,t)} = \beta^{(b)} \times (W^{(b)} \times J^{(b-1,t)} + W^{(b)} \times J^{(b,t-1)}) \tag{4}$$

Here, $J^{(b,t)}$ signifies classifier final output

2) TRAINING OF DeepRNN USING PROPOSED SMCA ALGORITHM

The training procedure of Deep RNN classifier is performed with proposed SMCA algorithm. However, the weight of the classifier is trained using the proposed SMCA for obtaining optimal solution. SMCA is used for training Deep RNN and is designed by combining SMO and CSA. Sentiment classification assist monitoring of social media as it permits to acquire an overview of public opinion under certain set of domains. The monitoring of social media makes the process quicker and simpler by monitoring their capabilities in real-time. Moreover, the analysis of sentiments helps to tackle the issues of customers and elevate their services. However, the impulsive and automatic detection of sentiments from the telecom review is the major issue faced by multinational organizations. Thus, the optimization methods are employed to improve accuracy of sentiment classification and expedite automatic evaluation of data. To attain effective sentiment classification, the optimization, namely SMCA is devised. The SMO [22] is highly effective as count of training data is very flexible in improving speed of computing. On the other hand, CSA [37] algorithm is motivated from the intelligent behavior of crows, which leads to optimal solutions. In CSA, an awareness probability is utilized, which helps to control the diversity of algorithm and is simpler to implement. By integrating the SMO with CSA, the parametric features from both the optimization are inherited, which boost the performance of classification accuracy. The algorithmic procedure for sentiment classification is given as follow:

1) **Initialize the population**: The foremost step is population initialization which is represented as, B (refer Eq. 5) with total l spider monkeys, such that $1 \leq k \leq l$.

$$B = \{B_1, B_2, \ldots, B_k, \ldots, B_l\} \tag{5}$$

where, l is total solution, and B_k indicate the k^{th} solution.

2) **Evaluation of error**: The mean square error (MSE) is computed by subtracting value of predicted output from actual output. The solution with less error is selected as best solution and is evaluated using Eq. 6.

$$E_c = \frac{1}{K} \times \sum_{g=1}^{k} J_v^{b,t} - k_v \tag{6}$$

L_n signifies fitness of n^{th} spider monkey, $J_v^{b,t}$ indicate output from DeepRNN classifier, and k_v symbolize estimated output.

3) **Determination of update position**: The SMO algorithm facilitates improved accuracy with efficient classification of sentiments. According to the SMO algorithm [22], the update of local leader phase of SMO is given by Eq. 7.

$$B_{m,n}(q+1) = B_{m,n}(q) + X(0,1) \times (R_{s,n} - B_{m,n}(q)) + X(-1,1) \times (B_{fn}(q) - B_{m,n}(q)) \quad (7)$$

where, $B_{m,n}(q)$ indicate the n^{th} dimension of m^{th} spider monkey, $R_{s,n}$ represent the leader position of s^{th} local group leader in n^{th} dimension, $X(0,1)$ specifies the uniform random number between 0 to 1, and $X(-1,1)$ indicate the uniform random number between -1 to 1, $B_{fn}(q)$ signifies the n^{th} dimension of f^{th} spider monkey that is chosen randomly based on s^{th} group. The Eq. 7 transformed into Eq. 8 and subsequently into Eq. 9.

$$B_{m,n}(q+1) = B_{m,n}(q) + X(0,1) \times R_{s,n} - X(0,1) \times B_{m,n}(q) + X(-1,1) \times B_{fn}(q) - X(-1,1) \times B_{m,n}(q) \quad (8)$$

$$B_{m,n}(q+1) = B_{m,n}(q) + [1 - X(0,1) - X(-1,1)] + X(0,1) \times R_{s,n} + X(-1,1) \times B_{fn}(q) \quad (9)$$

In order to obtain global optimal solutions in sentiment classification, the CSA [37] is utilized in this algorithm. Thus, the update position of crows based on CSA algorithm is given by Eq. 10.

$$B_{m,n}(q+1) = B_{m,n}(q) + c_{m,n} \times f_{mn}(q) \times (d_{an}(q) - B_{m,n}(q)) \quad (10)$$

where, $c_{m,n}$ is random number, $f_{mn}(q)$ indicate the flight length, $d_{an}(q)$ represent memory of m^{th} crow at dimension n in iteration q, $B_{m,n}(q+1)$ specifies position of m^{th} crow at dimension n in $q+1$ iteration, and $B_{m,n}(q)$ indicates the position of crow m at n dimension. The Eq. 10 transformed into Eq. 11, Eq. 12 and Eq. 13.

$$B_{m,n}(q+1) = B_{m,n}(q) + c_{m,n} \times f_{mn}(q) \times (d_{an}(q) - c_{m,n} \times f_{mn}(q) \times B_{m,n}(q)) \quad (11)$$

$$B_{m,n}(q+1) = B_{m,n}(q)(1 - c_{m,n} \times f_{mn}(q)) + c_{m,n} \times f_{mn}(q)(d_{an}(q)) \quad (12)$$

$$B_{m,n}(q) = \frac{1}{1 - c_{m,n} \times f_{mn}(q)} \times [B_{m,n}(q+1) - c_{m,n} \times f_{mn}(q)(d_{an}(q))] \quad (13)$$

The proposed SMCA update equation is obtained by substituting the Eq. 13 in Eq. 9 that is given by Eq. 14.

$$B_{m,n}(q+1) = \frac{B_{m,n}(q+1) - c_{m,n} \times f_{mn}(q)(d_{an}(q))}{1 - c_{m,n} \times f_{mn}(q)} \times [1 - X(0,1) - X(-1,1)] + X(0,1) \times R_{s,n} + X(-1,1) \times B_{fn}(q) \quad (14)$$

As CSA is based on the hiding place of m^{th} crow $d_{a,n}(q)$ and the local leader phase of SMO is based on leader position and thus $d_{a,n}(q) = R_{s,n}$. Thus, local leader phase of proposed SMCA is updated as shown in Eq. 15 to Eq. 20.

$$B_{m,n}(q+1) = \frac{B_{m,n}(q+1) - c_{m,n} \times f_{mn}(q) R_{s,n}}{1 - c_{m,n} \times f_{mn}(q)} \times [1 - X(0,1) - X(-1,1)] + X(0,1) \times R_{s,n} + X(-1,1) \times B_{fn}(q) \quad (15)$$

$$B_{m,n}(q+1) = \frac{B_{m,n}(q+1)[1 - X(0,1) - X(-1,1)]}{1 - c_{m,n} \times f_{mn}(q)} - \frac{c_{m,n} \times f_{mn}(q) R_{s,n}[1 - X(0,1) - X(-1,1)]}{1 - c_{m,n} \times f_{mn}(q)} + X(0,1) \times R_{s,n} + X(-1,1) \times B_{fn}(q) \quad (16)$$

$$B_{m,n}(q+1) - \frac{B_{m,n}(q+1)[1 - X(0,1) - X(-1,1)]}{1 - c_{m,n} \times f_{mn}(q)} = X(0,1) \times R_{s,n} + X(-1,1) \times B_{fn}(q) - \frac{c_{m,n} \times f_{mn}(q) R_{s,n}[1 - X(0,1) - X(-1,1)]}{1 - c_{m,n} \times f_{mn}(q)} \quad (17)$$

$$B_{m,n}(q+1)\left(1 - \frac{1 - X(0,1) - X(-1,1)}{1 - c_{m,n} \times f_{mn}(q)}\right) = X(0,1) \times R_{s,n} + X(-1,1) \times B_{fn}(q) - \frac{c_{m,n} \times f_{mn}(q) R_{s,n}[1 - X(0,1) - X(-1,1)]}{1 - c_{m,n} \times f_{mn}(q)} \quad (18)$$

$$B_{m,n}(q+1)\left(\frac{1 - c_{m,n} \times f_{mn}(q) - 1 + X(0,1) + X(-1,1)}{1 - c_{m,n} \times f_{mn}(q)}\right) = X(0,1) \times R_{s,n} + X(-1,1) \times B_{fn}(q) - \frac{c_{m,n} \times f_{mn}(q) R_{s,n}[1 - X(0,1) - X(-1,1)]}{1 - c_{m,n} \times f_{mn}(q)} \quad (19)$$

$$B_{m,n}(q+1)\left(\frac{X(0,1) + X(-1,1) - c_{m,n} \times f_{mn}(q)}{1 - c_{m,n} \times f_{mn}(q)}\right) = X(0,1) \times R_{s,n} + X(-1,1) \times B_{fn}(q) - \frac{c_{m,n} \times f_{mn}(q) R_{s,n}[1 - X(0,1) - X(-1,1)]}{1 - c_{m,n} \times f_{mn}(q)} \quad (20)$$

The final update equation of proposed SMCA is expressed by Eq. 21.

$$B_{m,n}(q+1) = \frac{1 - c_{m,n} \times f_{mn}(q)}{X(0,1) + X(-1,1) - c_{m,n} \times f_{mn}(q)} \times [X(0,1) \times R_{s,n} + X(-1,1) \times B_{fn}(q)$$

$$-\frac{c_{m,n} \times f_{mn}(q) R_{s,n}[1 - X(0,1) - X(-1,1)]}{1 - c_{m,n} \times f_{mn}(q)}] \tag{21}$$

4) **Error evaluation for update solutions**: The error of updated solutions is evaluated, wherein weights linked to the minimum error is employed for training Deep-RNN.
5) **Terminate**: The optimum weights are obtained in iterative manner until maximal iteration is accomplished. The pseudo code of proposed SMCA algorithm is illustrated in Algorithm 3.

Algorithm 3 Proposed SMCA Algorithm

Select input B, Output $B_{m,n}(q+1)$
Initialize population of spider monkeys B and Compute error
for Each solution **do**
  Choose global leader and local leader by greedy selection
  while Stopping Criteria not satisfied **do**
    Generate new position
    Compute error with Eq. 6
    Generate new position using Eq. 21
    Return best position
  end while
end for

D. INFORMATION RETRIEVAL USING FuzzyKNN

On the other hand, whenever a new user query or review arrives, the review is pre-processed and significant features are obtained. The features of test review are matched with the classified database using the FuKNN for retrieving the query information. Thus, the information retrieval is done effectively. The FuKNN algorithm allocates class membership to a sample. Consider Q represent new query wherein the pre-processing and feature extraction is carried out and the extracted features are matched with classified database using FuzzyKNN. Consider the extracted features be given as $F = \{f_1, f_2, \ldots, f_t, \ldots, f_u\}$ with a set of u data samples. Assume $u_o(f)$ represent a membership function of feature f_t and $u_{o,v}$ represent membership function of o^{th} class and t^{th} vector. The membership function of feature f_t is given by Eq. 22.

$$u_o(f) = \frac{\sum_{v=1}^{M} u_{o,v}\left(\frac{1}{\|f - f_v\|^{\frac{2}{P-1}}}\right)}{\sum_{v=1}^{M}\left(\frac{1}{\|f - f_v\|^{\frac{2}{P-1}}}\right)} \tag{22}$$

where, P denote the relative distance. Based on FuzzyKNN, the features are matched and the information is retrieved from the categorized data.

V. RESULTS AND DISCUSSION

The analysis of methods with precision, accuracy, F1-Score and recall is done by varying the training data. Moreover, the effectiveness of proposed SMCA-Deep RNN is illustrated.

A. EXPERIMENTAL SETUP

The execution of proposed SMCA-Deep RNN is performed in MATLAB considering Windows 10 OS, Intel i3 core processor and 2GB RAM.

B. DESCRIPTION OF DATASET

The analysis is performed on dataset that are taken from Amazon unlocked mobile reviews dataset (Link: https://www.kaggle.com/PromptCloudHQ/amazon-reviews-unlocked-mobile-phones) and telecom tweets for performing the sentiment classification.

1) AMAZON UNLOCKED MOBILE REVIEWS DATASET

This dataset [46] is extracted from 400 thousands reviews of mobile phones put on the market in Amazon.com to identify the insight with respects to reviews. The rate of product is ranging between 1-5. The maximum and minimum price of the product is 2598 and 226.86 respectively.

2) TELECOM TWEETS

The telecom tweets are collected from twitter website and the sentiments are identified using the comments of telecom tweets [47]. Further, the positive and negative comments are analyzed and stored in the database. The generated database is adapted for the experimentation.

C. PERFORMANCE MEASURES

The analysis of performance is done with performance measures named, accuracy, precision, False alarm rate, and recall, which are detailed in this section.

1) ACCURACY

This measure defines the nearness to the right sentiment classification using the reviews, and it is expressed by Eq. 23.

$$Accuracy = \frac{T^p + T^n}{T^p + F^p + T^n + F^n} \tag{23}$$

where, T^p indicate rate of true positives, T^n represent rate of true negatives, F^p denote rate of false positives, and F^n is the rate of false negatives.

2) PRECISION

Precision refers to the highest level of exactness, which is given by Eq. 24.

$$Precision = \frac{T^p}{T^p + F^p} \tag{24}$$

3) RECALL

Recall refers the ratio of the true positive with respect to addition of true positive and false negative and is formulated by Eq. 25.

$$Recall = \frac{T^p}{T^p + F^n} \tag{25}$$

4) F-MEASURE

It refers the harmonic mean of the recall and precision for computing the classification performance and is modelled by Eq. 26.

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{26}$$

D. COMPETING METHODS

The comparative methods include DeepCNN [48], Pattern based approach [4], Natural language processing (NLP) [49], and proposed SMCA+DeepRNN.

E. COMPARATIVE ANALYSIS

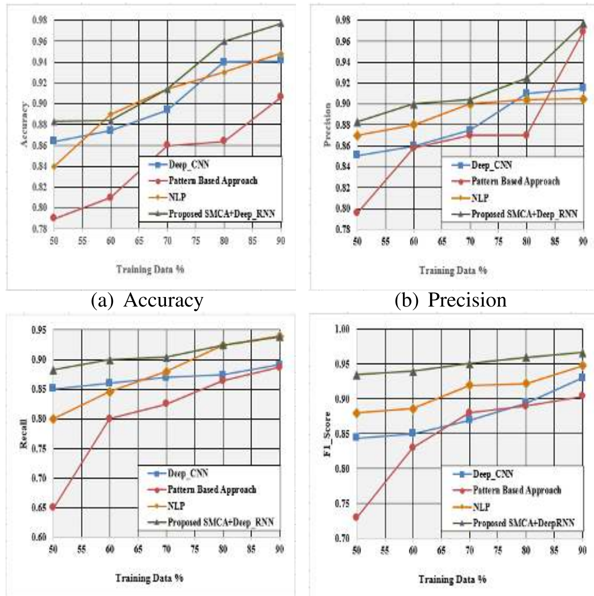
The evaluation of methods using two datasets is devised wherein the analysis is carried out on the methodologies based on certain metrics by altering training data and number of retrievals.

1) ANALYSIS USING AMAZON UNLOCKED MOBILE REVIEWS DATASET

The analysis based on Amazon unlocked mobile reviews dataset considering precision, recall, F1-Score, and accuracy is depicted below. The analysis is done by varying the number of retrievals and training data.

a: ANALYSIS BASED ON CLASSIFICATION

The analysis of methods using 20 Newsgroup dataset considering accuracy, recall, precision, and F1-score is illustrated in Fig. 3. The analysis using accuracy parameter for sentiment classification is illustrated in Fig. 3(a). For 50% training data, the accuracies computed by existing DeepCNN, Pattern based approach, NLP are 0.868%, 0.796%, 0.843%, whereas accuracy of proposed SMCA+DeepRNN is 0.885%, respectively. Similarly, for 90% training data, the accuracies computed by existing DeepCNN, Pattern based approach, NLP are 0.941%, 0.906%, 0.948% whereas accuracy of proposed SMCA+DeepRNN is 0.977%, respectively. The analysis based on precision parameter for sentiment classification is illustrated in Fig. 3(b). For 50% training data, the precision computed by existing DeepCNN, Pattern based approach, NLP are 0.849%, 0.798%, 0.869%, whereas precision of proposed SMCA+DeepRNN is 0.879, respectively. Similarly, for 90% training data, the precision computed by existing DeepCNN, Pattern based approach, NLP are 0.913%, 0.969%, 0.905%, whereas precision of proposed SMCA+DeepRNN is 0.955%, respectively. The analysis based on recall parameter for sentiment classification is illustrated in Fig. 3(c). For 50% training data, the recall computed by existing DeepCNN, Pattern based approach, NLP are 0.850%, 0.643%, 0.797%, whereas recall of proposed SMCA+DeepRNN is 0.868%, respectively. Similarly, for 90% training data, the recall computed by existing DeepCNN, Pattern based approach, NLP are 0.891%, 0.887%, 0.940%, whereas recall of proposed SMCA+DeepRNN is 0.939%,respectively.



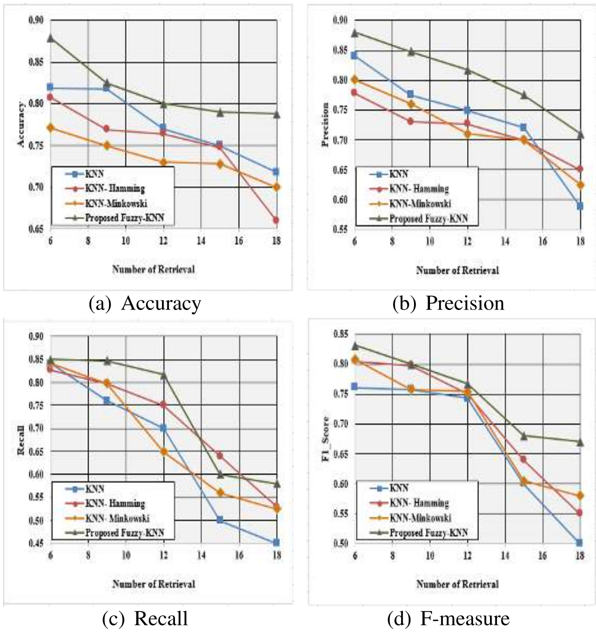
(a) Accuracy (b) Precision (c) Recall (d) F1-score

FIGURE 3. Analysis based on sentiment classification using dataset-1.

The analysis based on F1-score parameter for sentiment classification is illustrated in Fig. 3(d). For 50% training data, the F1-score computed by existing DeepCNN, Pattern based approach, NLP are 0,.833%, 0.728%, 0.852%, whereas F1-score of proposed SMCA+DeepRNN is 0.933%, respectively. Similarly, for 90% training data, the F1-score computed by existing DeepCNN, Pattern based approach, NLP are 0.930%, 0.904%, 0.948%, whereas F1-score of proposed SMCA+DeepRNN is 0.967%, respectively.

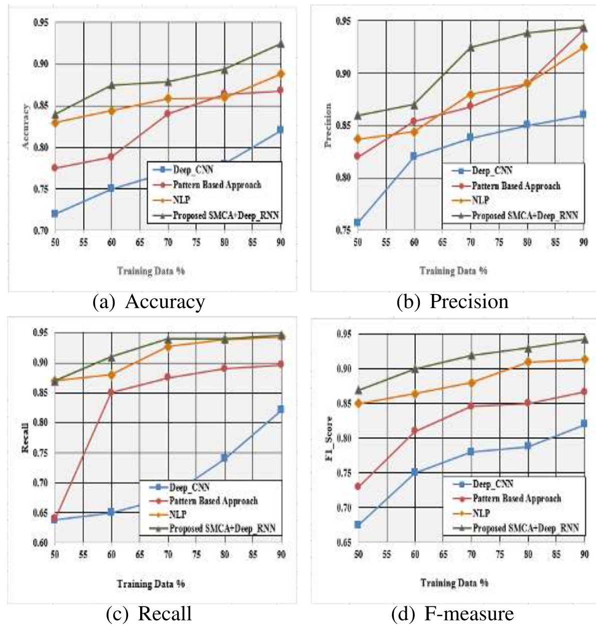
b: ANALYSIS BASED ON INFORMATION RETRIEVAL

The analysis of methods using 20 Newsgroup dataset considering accuracy, recall, precision, and F1-score is illustrated in Fig. 4. The analysis using accuracy parameter for Information retrieval is illustrated in Fig. 4(a). When number of retrieval is 6, the accuracies computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.819%, 0.807%, 0.771%, whereas accuracy of proposed Fuzzy-KNN is 0.879%, respectively. The analysis based on precision parameter for Information retrieval is illustrated in Fig. 4(b). When number of retrieval is 6, the precision computed by existing KNN, KNN-Hamming, KNN-Minkowski, proposed Fuzzy-KNN are 0. 840%, 0.778%, 0.804%, and 0.879%. Likewise, when number of retrieval is 18, the precision computed by existing are KNN, KNN-Hamming, KNN-Minkowski, 0.590%, 0.653%, and 0.619%, whereas the precision computed proposed Fuzzy-KNN is 0.708%, respectively. The analysis based on recall parameter for sentiment classification is illustrated in Fig. 4(c). When number of retrieval is 6, the recall computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.841%, 0.827%, and 0.842% whereas recall of proposed Fuzzy-KNN is 0.857%, respectively. Likewise, when number of retrieval is 18, the recall



(a) Accuracy (b) Precision (c) Recall (d) F-measure

FIGURE 4. Analysis based on information retrieval using dataset-1.



(a) Accuracy (b) Precision (c) Recall (d) F-measure

FIGURE 5. Analysis based on sentiment classification using dataset-2.

computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.460%, 0.527%, 0.522%, whereas recall of proposed Fuzzy-KNN is 0.581%, respectively. The analysis based on F1-score parameter for sentiment classification is illustrated in Fig. 4(d). When number of retrieval is 6, the F1-score computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.761%, 0.805%, 0.808%, whereas F1-Score proposed Fuzzy-KNN is 0.831%, respectively. Likewise, when number of retrieval is 18, the F1-score computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.505%, 0.547%, 0.582% whereas proposed Fuzzy-KNN is 0.671%, respectively.

2) ANALYSIS USING TELECOM TWEETS

The analysis based on telecom tweets dataset considering precision, recall, F1-Score, and accuracy is depicted below. The analysis is done by varying the number of retrievals and training data.

a: ANALYSIS BASED ON CLASSIFICATION

The analysis of methods using telecom tweets dataset considering accuracy, recall, precision, and F1-score is illustrated in Fig. 5. The analysis based on accuracy parameter for sentiment classification is illustrated in Fig. 5(a). For 50% training data, the accuracies computed by existing DeepCNN, Pattern based approach, NLP are, 0.714%, 0.770%, and 0.823% whereas the accuracy of proposed SMCA+DeepRNN is 0.832%, respectively. Similarly, for 90% training data, the accuracies computed by existing DeepCNN, Pattern based approach, NLP are 0.815%, 0.870%, and 0.888%, whereas the accuracy of proposed SMCA+DeepRNN is 0.922%, respectively. The analysis based on precision parameter for sentiment classification is illustrated in Fig. 5(b). For 50% training data, the precision computed by existing DeepCNN, Pattern based approach, NLP are 0.756%, 0.816%, and 0.832%, whereas the precision of proposed SMCA+DeepRNN is 0.856%, respectively. Similarly, for 90% training data, the precision computed by existing DeepCNN, Pattern based approach, NLP are 0.860%, 0.942%, and 0.925% whereas the precision of proposed SMCA+DeepRNN is 0.944%, respectively. The analysis based on recall parameter for sentiment classification is illustrated in Fig. 5(c). For 50% training data, the recall computed by existing DeepCNN, Pattern based approach, NLP are 0.637%, 0.631%, 0.881%, whereas the recall of proposed SMCA+DeepRNN is 0.881%, respectively. Similarly, for 90% training data, the recall computed by existing DeepCNN, Pattern based approach, NLP are 0.821%, 0.896%, 0.943% whereas the recall of proposed SMCA+DeepRNN is 0.946%, respectively. The analysis based on F1-score parameter for sentiment classification is illustrated in Fig. 5(d). For 50% training data, the F1-score computed by existing DeepCNN, Pattern based approach, NLP are 0.673%, 0.730%, 0.846% whereas F1-score of proposed SMCA+DeepRNN is 0.866%, respectively. Similarly, for 90% training data, the F1-score computed by existing DeepCNN, Pattern based approach, NLP are 0.820%, 0.867%, 0.914%, whereas F1-score of proposed SMCA+DeepRNN is 0.943% respectively.

b: ANALYSIS BASED ON RETRIEVAL

The analysis of methods using Reuter dataset considering accuracy, recall, precision, and F1-score is illustrated in Fig. 6. The analysis using accuracy parameter for sentiment classification is illustrated in Fig. 6(a). When

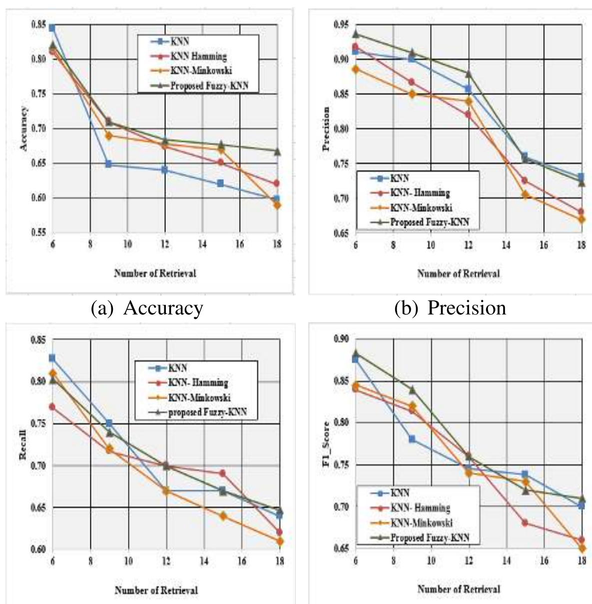
TABLE 1. **Comparative analysis for classification.**

Dataset	Metrics	DeepCNN	Pattern-based approach	NLP	Proposed SMCA+DeepRNN
Dataset-1	Accuracy	0.941	0.906	0.948	0.977
	Precision	0.913	0.969	0.905	0.955
	Recall	0.891	0.887	0.940	0.939
	F1-score	0.930	0.904	0.948	0.967
Dataset-2	Accuracy	0.815	0.870	0.888	0.922
	Precision	0.860	0.942	0.925	0.944
	Recall	0.821	0.896	0.943	0.946
	F1-score	0.820	0.867	0.914	0.943

TABLE 2. **Comparative analysis for information retrieval.**

Dataset	Metrics	KNN	KNN- Hamming	KNN-Minkowski	Proposed Fuzzy-KNN
Dataset-1	Accuracy	0.819	0.807	0.771	0.879
	Precision	0.840	0.778	0.801	0.879
	Recall	0.841	0.827	0.842	0.842
	F1-score	0.761	0.805	0.808	0.831
Dataset-2	Accuracy	0.844	0.810	0.816	0.821
	Precision	0.911	0.918	0.886	0.937
	Recall	0.827	0.769	0.810	0.803
	F1-score	0.875	0.839	0.845	0.883

number of retrieval is 6, the accuracies computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.844, 0.810, 0.816 whereas accuracy of proposed Fuzzy-KNN is 0,821. Likewise, when number of retrieval is 18, the accuracies computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.592, 0.613, 0.588 whereas proposed Fuzzy-KNN is 0.666. The analysis based on precision parameter for Information retrieval is illustrated in Fig. 6(b). For 6 retrieval, the precision computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.911, 0.918, 0.886 whereas proposed Fuzzy-KNN is 0.937,. Likewise, for 18 retrieval, the precision computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.724, 0.692, 0.682 whereas proposed Fuzzy-KNN is 0.720, respectively. The analysis based on recall parameter for Information retrieval is illustrated in Fig. 6(c). For 6 retrieval, the recall computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.827, 0.769, 0.810 whereas proposed Fuzzy-KNN is 0.803. Likewise, for 18 retrieval, the recall computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.633, 0.617, 0.607 whereas proposed Fuzzy-KNN is 0.642, respectively. The analysis based on F1-score parameter for Information retrieval is illustrated in Fig. 6(d). For 6 retrieval, the recall computed by existing KNN, KNN-Hamming, KNN-Minkowski, are 0.875, 0.839, 0.845 whereas proposed information retrieval is 0.883,. Likewise, for 18 retrieval, the recall computed by KNN, KNN-Hamming, KNN-Minkowski, are 0.696, 0.656, 0.651, whereas proposed information retrieval is 0.642, respectively.



(a) Accuracy (b) Precision

(c) Recall (d) F-measure

FIGURE 6. **Analysis based on information retrieval using Reuter dataset.**

F. COMPARATIVE DISCUSSION

Table 1 and Table 2 elaborates analysis of methods considering performance shown by the methods, with maximum training data for sentiment classification and number of retrievals for information retrieval. Considering sentiment classification, the maximal accuracy is shown by proposed SMCA+DeepRNN with accuracy value of 0.977, the precision of 0.955, maximal recall is 0.946, and maximal F1-Score is 0.967. Considering information retrieval, the maximal accuracy is shown by proposed Fuzzy-KNN with accuracy value of 0.879, the precision of 0.879, proposed Fuzzy-KNN with recall value of 0.842, and maximal F1-Score is 0,883 respectively. Through the analysis, it is clear that proposed SMCA+DeepRNN showed improved performance as compared to other existing methods.

VI. CONCLUSION

The analysis of sentiment is carried out with Deep RNN, which aimed to improve sentiment classification performance. The conventional techniques of sentiment classification with neural networks reveal poor performance in the existence of irrelevant data addressed using the proposed sentiment classification method. The proposed SMCA is adapted for training classifier to derive optimal weights, and the proposed SMCA is a combination of SMO and CSA. The training of the classifier is carried out with features obtained using input telecom reviews. The features are extracted using SentiWordNet, which provided features like polarity, subjectivity, positive score, negative score, and other features, like many elongated words, punctuation, and hashtag and numerical values. The classification of sentiments is done using the features mentioned above with the proposed SMCAbased DeepRNN model. The experimentation is carried out using two datasets, namely Amazon unlocked mobile reviews dataset and telecom tweets dataset. The proposed SMCAbased DeepRNN outperformed other methods with a maximum accuracy of 97.7%, precision of 95.5%, recall of 94.6%, and F1-score 96.7%, respectively.

The future work will be done by adopting advanced features like n-grams and feature weighting to attain improved accuracy. Additionally, this work may be extended to detect fake reviews, deceptive reviews, fake comments, fake blogs, fake reviewers, review spammer, opinion spammer, and many more. A new algorithm may be deployed for feature extraction and selection.

FUNDING

This work is partially funded by the Robotics and Internet of Things Lab at Prince Sultan University.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] F. Mata-Rivera, M. Torres-Ruiz, G. Guzmán, M. Moreno-Ibarra, and R. Quintero, "A collaborative learning approach for geographic information retrieval based on social networks," *Comput. Hum. Behav.*, vol. 51, pp. 829–842, Oct. 2015.

[2] P. A. Tess, "The role of social media in higher education classes (real and virtual)—A literature review," *Comput. Hum. Behav.*, vol. 29, no. 5, pp. A60–A68, Sep. 2013.

[3] S. Golubić and D. Marušić, "Reviews and inspections-an approach to the improvement of telecom software development aprocess," in *Proc. 1st Workshop Inspection Softw. Eng. (WISE)*, 2001, pp. 283–290.

[4] M. Bouazizi and T. Ohtsuki, "A pattern-based approach for multi-class sentiment analysis in Twitter," *IEEE Access*, vol. 5, pp. 20617–20639, 2017.

[5] Z. Jianqiang, G. Xiaolin, and Z. Xuejun, "Deep convolution neural networks for Twitter sentiment analysis," *IEEE Access*, vol. 6, pp. 23253–23260, 2018.

[6] P. Munjal, M. Narula, S. Kumar, and H. Banati, "Twitter sentiments based suggestive framework to predict trends," *J. Statist. Manage. Syst.*, vol. 21, no. 4, pp. 685–693, Jul. 2018.

[7] P. Munjal, S. Kumar, L. Kumar, and A. Banati, "Opinion dynamics through natural phenomenon of grain growth and population migration," in *Hybrid Intelligence for Social Networks*. Springer, 2017, pp. 161–175.

[8] P. Munjal, L. Kumar, S. Kumar, and H. Banati, "Evidence of ostwald ripening in opinion driven dynamics of mutually competitive social networks," *Phys. A, Stat. Mech. Appl.*, vol. 522, pp. 182–194, May 2019.

[9] M. Bouazizi and T. Ohtsuki, "Multi-class sentiment analysis in Twitter: What if classification is not the answer," *IEEE Access*, vol. 6, pp. 64486–64502, 2018.

[10] H. Schütze, C. D. Manning, and P. Raghavan, *Introduction to Information Retrieval*, vol. 39. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[11] A. Singhal, "Modern information retrieval: A brief overview," *IEEE Comput. Soc. Tech. Committee Data Eng.*, vol. 24, no. 4, pp. 35–43, Jan. 2001.

[12] I. Shah, M. El Affendi, and B. Qureshi, "SRide: An online system for multi-hop ridesharing," *Sustainability*, vol. 12, no. 22, p. 9633, Nov. 2020.

[13] P. Ducange, M. Fazzolari, M. Petrocchi, and M. Vecchio, "An effective decision support system for social media listening based on cross-source sentiment analysis models," *Eng. Appl. Artif. Intell.*, vol. 78, pp. 71–85, Feb. 2019.

[14] L. Li, Q. Xu, T. Gan, C. Tan, and J.-H. Lim, "A probabilistic model of social working memory for information retrieval in social interactions," *IEEE Trans. Cybern.*, vol. 48, no. 5, pp. 1540–1552, May 2018.

[15] L. Gao, Y. Wang, D. Li, J. Shao, and J. Song, "Real-time social media retrieval with spatial, temporal and social constraints," *Neurocomputing*, vol. 253, pp. 77–88, Aug. 2017.

[16] C. Virmani, D. Juneja, and A. Pillai, "Design of query processing system to retrieve information from social network using NLP," *KSII Trans. Internet Inf. Syst.*, vol. 12, no. 3, pp. 1–21, 2018.

[17] F. Iqbal, J. M. Hashmi, B. C. M. Fung, R. Batool, A. M. Khattak, S. Aleem, and P. C. K. Hung, "A hybrid framework for sentiment analysis using genetic algorithm based feature reduction," *IEEE Access*, vol. 7, pp. 14637–14652, 2019.

[18] S. Mansour, "Social media analysis of User's responses to terrorism using sentiment analysis and text mining," *Procedia Comput. Sci.*, vol. 140, pp. 95–103, Jan. 2018.

[19] A. S. M. Alharbi and E. de Doncker, "Twitter sentiment analysis with a deep neural network: An enhanced approach using user Behavioral information," *Cognit. Syst. Res.*, vol. 54, pp. 50–61, May 2019.

[20] M. Mataoui, F. Sebbak, F. Benhammadi, and K. B. Bey, "Query expansion in XML information retrieval: A new approach for terms selection," in *Proc. 6th Int. Conf. Modeling, Simulation, Appl. Optim. (ICMSAO)*, May 2015, pp. 1–4.

[21] A. U. Hassan, J. Hussain, M. Hussain, M. Sadiq, and S. Lee, "Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2017, pp. 138–140.

[22] J. C. Bansal, H. Sharma, S. S. Jadon, and M. Clerc, "Spider monkey optimization algorithm for numerical optimization," *Memetic Comput.*, vol. 6, no. 1, pp. 31–47, Mar. 2014.

[23] A. Sharma, A. Sharma, B. K. Panigrahi, D. Kiran, and R. Kumar, "Ageist spider monkey optimization algorithm," *Swarm Evol. Comput.*, vol. 28, pp. 58–77, Jun. 2016.

[24] S. Kumar, B. Sharma, V. K. Sharma, H. Sharma, and J. C. Bansal, "Plant leaf disease identification using exponential spider monkey optimization," *Sustain. Comput., Informat. Syst.*, vol. 28, Dec. 2020, Art. no. 100283.

[25] S. Kumar, B. Sharma, V. K. Sharma, and R. C. Poonia, "Automated soil prediction using bag-of-features and chaotic spider monkey optimization algorithm," *Evol. Intell.*, pp. 1–12, Nov. 2018, doi: 10.1007/s12065-018-0186-9.

[26] S. Kumar, A. Nayyar, N. G. Nguyen, and R. Kumari, "Hyperbolic spider monkey optimization algorithm," *Recent Adv. Comput. Sci. Commun.*, vol. 13, no. 1, pp. 35–42, Mar. 2020.

[27] K. Gupta, K. Deep, and J. C. Bansal, "Spider monkey optimization algorithm for constrained optimization problems," *Soft Comput.*, vol. 21, no. 23, pp. 6933–6962, Dec. 2017.

[28] K. Gupta, K. Deep, and J. C. Bansal, "Improving the local search ability of spider monkey optimization algorithm using quadratic approximation for unconstrained optimization," *Comput. Intell.*, vol. 33, no. 2, pp. 210–240, May 2017.

[29] U. Singh, R. Salgotra, and M. Rattan, "A novel binary spider monkey optimization algorithm for thinning of concentric circular antenna arrays," *IETE J. Res.*, vol. 62, no. 6, pp. 736–744, Nov. 2016.

[30] A. Sharma, H. Sharma, A. Bhargava, N. Sharma, and J. C. Bansal, "Optimal placement and sizing of capacitor using Limaçon inspired spider monkey optimization algorithm," *Memetic Comput.*, vol. 9, no. 4, pp. 311–331, Dec. 2017.

[31] S. S. Pal, S. Kumar, M. Kashyap, Y. Choudhary, and M. Bhattacharya, "Multi-level thresholding segmentation approach based on spider monkey optimization algorithm," in *Proc. 2nd Int. Conf. Comput. Commun. Technol.* India: Springer, 2016, pp. 273–287.

[32] A. Sharma, H. Sharma, A. Bhargava, and N. Sharma, "Optimal design of pida controller for induction motor using spider monkey optimization algorithm," *Int. J. Metaheuristics*, vol. 5, nos. 3–4, pp. 278–290, 2016.

[33] N. Mittal, U. Singh, R. Salgotra, and B. S. Sohi, "A Boolean spider monkey optimization based energy efficient clustering approach for WSNs," *Wireless Netw.*, vol. 24, no. 6, pp. 2093–2109, Aug. 2018.

[34] H. Sharma, G. Hazrati, and J. C. Bansal, "Spider monkey optimization algorithm," in *Evolutionary and Swarm Intelligence Algorithms*. Springer, 2019, pp. 43–59.

[35] V. Swami, S. Kumar, and S. Jain, "An improved spider monkey optimization algorithm," in *Soft Computing: Theories and Applications*. Springer, 2018, pp. 73–81.

[36] B. Sharma, V. K. Sharma, and S. Kumar, "Sigmoidal spider monkey optimization algorithm," in *Soft Computing: Theories and Applications*. Springer, 2020, pp. 109–117.

[37] A. Askarzadeh, "A novel Metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm," *Comput. Struct.*, vol. 169, pp. 1–12, Jun. 2016.

[38] G. I. Sayed, A. E. Hassanien, and A. T. Azar, "Feature selection via a novel chaotic crow search algorithm," *Neural Comput. Appl.*, vol. 31, no. 1, pp. 171–188, Jan. 2019.

[39] D. Gupta, S. Sundaram, A. Khanna, A. Ella Hassanien, and V. H. C. de Albuquerque, "Improved diagnosis of Parkinson's disease using optimized crow search algorithm," *Comput. Electr. Eng.*, vol. 68, pp. 412–424, May 2018.

[40] F. Mohammadi and H. Abdi, "A modified crow search algorithm (MCSA) for solving economic load dispatch problem," *Appl. Soft Comput.*, vol. 71, pp. 51–65, Oct. 2018.

[41] D. Oliva, S. Hinojosa, E. Cuevas, G. Pajares, O. Avalos, and J. Gálvez, "Cross entropy based thresholding for magnetic resonance brain images using crow search algorithm," *Expert Syst. Appl.*, vol. 79, pp. 164–180, Aug. 2017.

[42] P. Díaz, M. Pérez-Cisneros, E. Cuevas, O. Avalos, J. Gálvez, S. Hinojosa, and D. Zaldivar, "An improved crow search algorithm applied to energy problems," *Energies*, vol. 11, no. 3, p. 571, Mar. 2018.

[43] H. Zamani, M. H. Nadimi-Shahraki, and A. H. Gandomi, "CCSA: Conscious neighborhood-based crow search algorithm for solving global optimization problems," *Appl. Soft Comput.*, vol. 85, Dec. 2019, Art. no. 105583.

[44] M. Ghosh and A. Kar, "Unsupervised linguistic approach for sentiment classification from online reviews using sentiwordnet 3.0," *Int J Eng Res Technol*, vol. 2, no. 9, pp. 1–6, Sep. 2013.

[45] M. Inoue, S. Inoue, and T. Nishida, "Deep recurrent neural network for mobile human activity recognition with high throughput," *Artif. Life Robot.*, vol. 23, no. 2, pp. 173–185, Jun. 2018.

[46] PromptCloud. (Nov. 2020). *Amazon Unlocked Mobile Reviews Dataset*. Accessed: Dec. 12, 2020. [Online]. Available: https://www.kaggle.com/PromptCloudHQ/amazon-reviews-unlocked-mobile-phones

[47] NLTK Data. (Nov. 2020). *Reuterdataset Reuter Database*. Accessed: Nov. 23, 2020. [Online]. Available: https://www.kaggle.com/nltkdata and https://www.kaggle.com/nltkdata/reuters

[48] J. Feng, S. Cai, and X. Ma, "Enhanced sentiment labeling and implicit aspect identification by integration of deep convolution neural network and sequential algorithm," *Cluster Comput.*, vol. 22, no. S3, pp. 5839–5857, May 2019.

[49] M. Kanakaraj and R. M. R. Guddeti, "NLP based sentiment analysis on Twitter data using ensemble classifiers," in *Proc. 3rd Int. Conf. Signal Process., Commun. Netw. (ICSCN)*, Mar. 2015, pp. 1–5.

AARTI CHUGH is currently an Assistant Professor with Amity University, India. She is also a Research Scholar with Jagannath University, Jaipur, India. She has published 15 research articles in well-known journals. She presented her research in various conferences, and attended several national and international conferences and workshops. Her research interests include big data, database systems, image processing, software testing, and computational intelligence.

VIVEK KUMAR SHARMA is currently a Professor and the Director of the Research with Jagannath University, Jaipur. He has published more than 50 articles in various international journals and conferences. His research interests include nature-inspired optimization techniques and fluid dynamics.

SANDEEP KUMAR received the B.E. degree from Engineering College, Kota, the M.Tech. Degree from RTU, Kota, and the Ph.D. Degree in computer science and engineering. He is currently an Associate Professor of Computer Science and Engineering with CHRIST (Deemed to be University), Bangalore, India. He has published more than 50 articles in well-known SCI/SCOPUS indexed international journals and conferences. He attended several national and international conferences and workshops. He has authored/edited four books in the area of computer science. His research interests include nature inspired algorithms, Swarm intelligence, soft computing, and computational intelligence. He edited Special Issues for many journals, including JDMSC, JIM, JSMS, JIOS, IJGUC, IJIIDS, RPCS, and RACSC.

ANAND NAYYAR (Senior Member, IEEE) received the Ph.D. degree in computer science from Desh Bhagat University, in 2017, in the area of wireless sensor networks and Swarm intelligence. He is currently working with the Graduate School, Duy Tan University, Da Nang, Vietnam. A Certified Professional with 75 Professional certificates from CISCO, Microsoft, Oracle, Google, Beingcert, EXIN, GAQM, Cyberoam, and many more. He published more than 350 Research articles in various National and International Conferences, International Journals (Scopus/SCI/SCIE/SSCI Indexed) with High Impact Factor. He has authored/coauthored cum edited 25 books of *Computer Science*. He is a member of more than 50 Associations as a Senior Member and a Life Member and also acting as an ACM Distinguished Speaker. He was associated with more than 500 International Conferences as a Programme Committee/the Chair/an Advisory Board/a Review Board Member. He was awarded 27 Awards for Teaching and Research. The Young Scientist, the Best Scientist, the Young Researcher Award, the Outstanding Researcher Award, and the Excellence in Teaching. He is acting as the Editor-in-Chief of IGI-Global, USA, Journal titled *International Journal of Smart Vehicles and Smart Transportation* (IJSVST).

BASIT QURESHI (Member, IEEE) received the bachelor's degree in computer science from Ohio University, Athens, OH, USA, in 2000, the master's degree in computer science from Florida Atlantic University, in 2002, and the Ph.D. degree in computer science from the University of Bradford, U.K., in 2011. He is currently an Associate Professor and the Chair of the Department of Computer Science with Prince Sultan University, Saudi Arabia. He has published over 60 research works in the area of robotics, the Internet of Things, machine learning, and computational intelligence. He is a member of the IEEE Computer Society, the IEEE Communications Society, and the ACM.

MANJOT KAUR BHATIA received the Ph.D. degree in information security from the University of Delhi, Delhi, India. She is currently working as a Professor of Computer Science and the Head of the Research Department with the Jagan Institute of Management Studies, Delhi. She has more than 20 years of teaching and research experience in the areas of information security, databases, Linux and operating system. She has published more than 20 research articles in scopus/sci/dblp indexed journals. She has presented number of papers in international conferences. Her research interests include cloud computing, steganography, data hiding, information security, and software testing. She has been the Session Chair in various international conferences. She is a Reviewer of various international journals.

CHARU JAIN is currently an Assistant Professor with Amity University Haryana, India. She has published 15 research articles in well-known journals. She presented her research in various conferences, and attended several national and international conferences and workshops. Her research interests include computational intelligence, image processing, artificial intelligence, and biometric.