Nat. Volatiles & Essent. Oils, 2021; 8(6): 2843-2857

NVEO
Natural Volatiles &
Essential Oils

A Novel Hybrid Two-Pass Optimized Deep Neural Network Classifier (Cnn-Olstm) For Twitter Sentiment Analysis

Sailaja Kumar K and Evangelin Geetha D

Department of Master of Computer Applications, M S Ramaiah Institute of Technology, India

Abstract

The goal of sentiment analysis is to identify and categorize the emotions expressed in tweets, messages, and user reviews.Social media platforms like Instagram, Facebook, and Twitter generate a lot of emotionally charged data, which may be quite beneficial when trying to improve the quality of both products and services uniformly.Even though various machine learning algorithms have been designed to recognize the emotions of Twitter users, sentiment analysis still faces significant obstacles.A two-pass optimized deep neural network has been designed to address this.The long short-term memory (LSTM) classifier and the convolution neural network (CNN) are used to create the proposed hybrid two-pass classifier (CNN-OLSTM).In addition, the Adaptive sunflower optimization(ASFO)technique is proposed to optimize the parameter existing in a two-pass classifier to improve its performance.Various measures are used to evaluate the effectiveness of the proposed strategy.

Keywords: -Sentiment analysis, Twitter, Convolution neural network, Long-short term memory, Adaptive sunflower optimization technique.

1. Introduction

Online media is now widely regarded as the best platform for distributing facts around the world [1].Twitter is one of the most widely used social media platforms.Also, it is one of the most widely used microblogging platforms[2]. Twitter has garnered increased interest in education and industry due to a growth in the number of users registered online, commercial prospects, management needs, and security concerns [3]. Twitter allows people to express their feelings and opinions about items and services.A tweet is a 140-character message that contains phrases and words that are not commonly used in language processing.Twitter has increased the text limit to 280 characters today.Twitter is an excellent tool for gaining insight into public opinion and sentiment analysis [4]. For learning about public opinion and sentiment analysis, Twitter is an excellent resource. Sentiment analysis is particularly useful for social media monitoring.Sentiment analysis deals with processing and transforming reviews and comments, which is a challenging task [5].New strategies have been introduced to improve the precision of sentiment analysis.

Sentiment analysis is the process of automatically extracting opinions from data and estimating their polarity (positive, negative, or neutral).Machine language or deep learning models, as well as natural language processing (NLP) approaches, are used to classify the data [6]. Deep learning is the most focused of these strategies since it allows researchers to obtain significant outcomes for NLP tasks [7]. The assumption behind sentiment analysis in social networks is that the texts given by users are independent and consistently distributed.With the use of opinion mining and sentiment analysis, policymakers incorporate public input into the formulation of any new policy [8]. Sentiment analysis is an algorithmic procedure that determines whether a text segment contains objective or opinionated content.It can also determine the text's polarity of sentiment.

Artificial neural networks (ANN), support vector machines (SVM), k-nearest neighbor (KNN), deep neural network (DNN), longshort-term memory (LSTM), and convolution neural networks (CNN) are some of the machine learning and deep learning methods for sentiment analysis that have been proposed in the literature[9-11].Recurrent neural networks (RNN) are one of the most extensively used approaches for classification and modeling.This is capable of handling sequences of any length.The disappearing gradient problem, on the other hand, is the most challenging aspect of the RNN structure.Other types of RNN are created to circumvent RNN restrictions [12].

In this paper, the hybridization of the LSTM and CNN was introduced to mitigate the impact. An adaptive sunflower optimization (ASFO) technique is suggested in this research to identify inputs as positive or negative depending on polarity.Contributions of the proposed approach (CNN-OLSTM) are described below:

- To begin, pre-processing is used to lessen the complexity of Twitter reviews. Stop words are removed, stemming is done, and segregation is done as part of the pre-processing.
- The words are turned into vectors for pre-processing. A skip-gram model is employed for vector conversion.
- A CNN classifier is used in the dimension reduction procedure.
- The LSTM classifier receives the reduced dataset.
- The parameters of the classifier are appropriately selected using the ASFO technique to improve its effectiveness.
- The proposed approach’s effectiveness is measured in terms of accuracy and mean squared error.

2. Literature survey

Many academics have developed sentiment analysis for Twitter data.Among these, some of the works are examined in this article;Sentiment analysis using an LSTM-CNN hybrid classifier was first introduced by Ishaani Priyadarshini and Chase Cotton [13].The authors employed a variety of criteria to evaluate their findings, including accuracy, precision, sensitivity, specificity, and the F-1 score.They used several datasets to compare their work against baseline algorithms such as CNN, KNN, and LSTM in an experimental study.The LSTM-CNN strategy had a maximum accuracy of 96 percent, which was greater than the baseline methods.

Enhanced continuous neural network with Firefly-Oriented Multi-Verse Optimizer (FF-MVO) based sentiment analysis on financial data gathered from Twitter was introduced by Samik Datta and Satyaji Chakrabarti [14].They start by extracting the characteristics. These features are converted using the Word2vec conversion method.The FF-MVO algorithm was used to calculate polarity scores, and its performance was compared to that of other algorithms.Muhammad Umer and his associates used the unified CNN-LSTM network model to analyze sentiment in Tweets [15].The model's efficacy in this technique was evaluated using various classifiers.The feature extraction approaches employed in this research were frequency-inverse document frequency and word2vec.In order to assess the performance, three datasets were employed.As a result, the CNN-LSTM classifier had greater accuracy than other classifiers, according to the authors.

Accounting Local Contextual Semantic was introduced by Itisha Gupta and Nisheeth Joshi for improved sentiment analysis of Twitter data [16]. A feature vector based on SentiWordNet (SWN) was employed as input in this method.The authors provided a naive and novel shift strategy in their proposed approach.This approach was used as a modifier and a sentiment-bearing word.Following that, the score of SWN words was adjusted depending on their context semantics and guesses from nearby terms.Zhao Jianqiang et al. used Deep- CNN to do sentiment analysis on Twitter data [17].In this approach, the authors presented a word embedding system derived from unauthorized learning based on the large Twitter corpora. This method used latent contextual semantic relationships between words on Twitter and co-occurrence statistical properties. Further, feature sets for training and predicting sentiment classification labels integrated into the Deep-CNN. The authors in [18] introduced sentiment analysis of Twitter using Deep CNN. In this approach, they used CNN with multiple filters for different window sizes. Using GloVe Twitter word vectors, the authors achieved better results.The proposed approach is the test dataset reported from the SemEval 2015 Task 10. The approach does not depend on any hand-crafted features or polarizing dictionaries.

Mingda Wang and Guangmin Hu had introduced sentiment analysis of Twitter using a new method of Attentional-Graph Neural Network (AGN) [19]. In this approach, a three-layered neural structure for AGN-TSA integration on tweet text and the user link information is proposed. The proposed layout includes the word-embed layer, the user-embed layer, and the graphical network layer of focus. For analysis, the authors used a real-world database of the 2016 US presidential election. From the result, the authors achieve 5% better performance compared to multiple metrics. From the results, the authors proved AGN-TSA was efficient and feasible for TSA tasks.

Willian Becker et alintroduced multilingual Sentiment Analysis in Twitter based on Efficient Deep Neural Architecture [20]. In this approach, the authors presented an efficient method for performing multilingual sentiment analysis over tweets. The proposed system and each architecture are based on a distinct embedding strategy. In this approach, they used character-based models with convolutional neural networks in multilingual contexts. For analysis, the authors used different computational resources for classifying tweets. From the results, the proposed architecture, Conv-Char-R, reaches competitive results when compared with other deep neural models for sentiment classification.

3. Methodology

Using the proposed approach, it is possible to automatically identify whether a tweet is negative or positive based on its sentiment polarity. Figure1 illustrates the overall structure of the proposed approach.The proposed model consists of three stages, namely, pre-processing, converting words into vectors, and making predictions with CNN-OLSTM algorithms. The performance of the CNN-OLSTM classifier is enhanced by using the ASFOtechnique. It is used to determine whether a user's tweets regarding specific product brands are positive or negative after the prediction procedure.The information from this approach can be used to analyze the product details, movie polarity, stock market, best brand products, and political data.

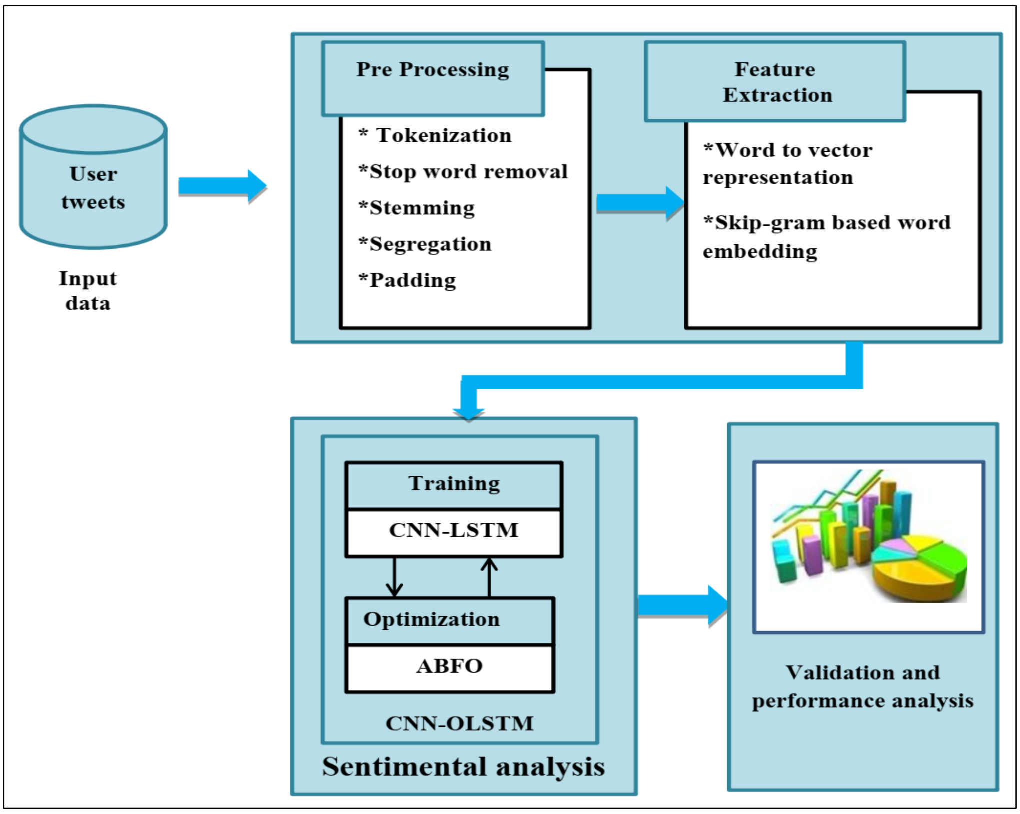

Figure 1. CNN-OLSTM Model for Sentiment Analysis

3.1 Data Pre-processing

The first step in the sentiment analysis process is pre-processing.It will convert the raw data into a format that the user can comprehend and interpret.Experimental analysis of social media data is conducted in this paper. Due to its origin from various types of individuals and sources, social media data are large and highly susceptible to noise, fragmentation, missing data, and conflicting information.Therefore, pre-processing is necessary before processing the data.Pre-processing involves the following steps:

Tokenization: It is the process of splitting large text strings into smaller pieces or tokens. Sentences may be tokenized into words, and paragraphs can be tokenized into sentences.

Stop words removal: In this step, the unwanted words are removed. The stop words are available in the natural language toolkit.

Stemming: Streaming is the process of extracting affixes from a text to obtain a text stem. For example, the terms “working” and “teaching” are simplified to their core word “work” and “teach” respectively.

Segregations: In segregation method, the individual characters such as “ ’ ? ! ; : # $ % & ( ) * + − / <> = [ ] n ˆ _ { } | ~ are separated from the tweets.

Padding: This is the process of truncating the words. This will help for further processing.

POS tagging: The part-of-speech tagging (POS) allows to automatically mark the word of each text based on which part it belongs to noun, pronoun, adverb, adverb, interjection, intensify, etc. In sentiment analysis, POS-tagging is regarded as the most crucial component. It is essential to separate the review sentence's qualities or strengths. Each word in the text is labeled by POS-Tagger.

3.2 Word2vec Conversion

Word2vec was invented by Weik Mikolov and is capable of capturing semantic and syntactic significance from the content. In this paper, the stochastic gradient method (SGM)usesWord2vec conversion as shown in Figure 2. The most relevant to a word is discovered by using the SGM method, which uses non-supervised learning techniques [17]. It presents a target word as input and context words as output.

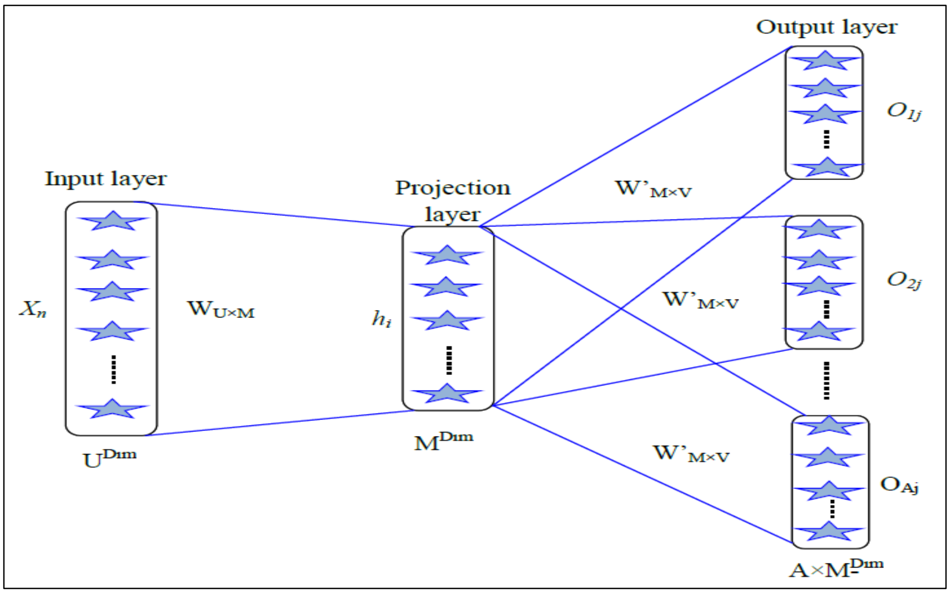

Figure 2. Structure of SGM

In SGM, the input layer is designed based on a one-hot vector. In a one-hot vector, the word present in the corresponding position is considered as 1 and all the other position is placed as zeros. Activation functions are not applied to nodes on the hidden layer, while regression classifiers are applied to nodes on the output layer. To maximize the objective function, the SGM model is trained.The objective function of SGM is given in equation (1).

$$S(W_{c,i} = W_{o,c} | W_I) = O_{c,i} = \frac{\exp(U_{c,i})}{\Sigma_{i'=1}^{V} \exp(u_{i'})} \quad (1)$$

Where

$W_{c,i}$ → Output layer i[th] word on the c[th] panel

$W_{O,c}$ →Actual c[th] word in the output context words

W_I → Input word

$O_{c,i}$ →Output of c[th] panel of i[th] unit of the output layer

$U_{c,i}$ →Net input of c[th] panel of i[th] unit of the output layer.

The weight values of input to hidden and hidden to output matrix is given in below equations;

$T'^{(new)}_{wi} = T'^{(old)}_{wi} - \eta . EI_i .h$ (2)

$T'^{(new)}_{wI} = T'^{(old)}_{wI} - \eta . EI^T$ (3)

Where

$T_{w,i}$ → Column of the hidden to output weight matrix W' for w_i

$T_{w,I}$ → Row of the input to hidden weight matrix W for w_I

EI_i → Sum of prediction errors overall context words

3.3 Feature Selection Using CNN

Once the vectors have been converted to Word2vec, they are fed into the CNN classifier as input.The main aim of CNN is to reduce the output size of the vector as shown in Figure 3. In this, unwanted features are removed. This will increase the prediction accuracy and reduce the computation complexity. In CNN, initially, the extracted input is given to the convolution layer. The first convolution layer gives the 16 feature map output. Here, the ReLU implementation function is used. The output of the last convolution layer acts as the input to the next convolution layer. The next convolution layer contains 32 3x3 sized kernel filters that are used for each functional graph recovered from the last layer. Specific functions such as ReLU and maximum pooling are performed to generate 64x64 pixel sub-sample data. The same functions are performed on the third layer of the last layer, which uses 64 filters of 3x3 size kernels that generate 32x32 pixel data. The third convolution layer contains the output of 64 32x32 pixel feature maps. These features are equated to a length vector of 32x32x64 = 65536, which acts as the input to the fully attached layer. These features are used to evaluate whether the review type is positive or negative.

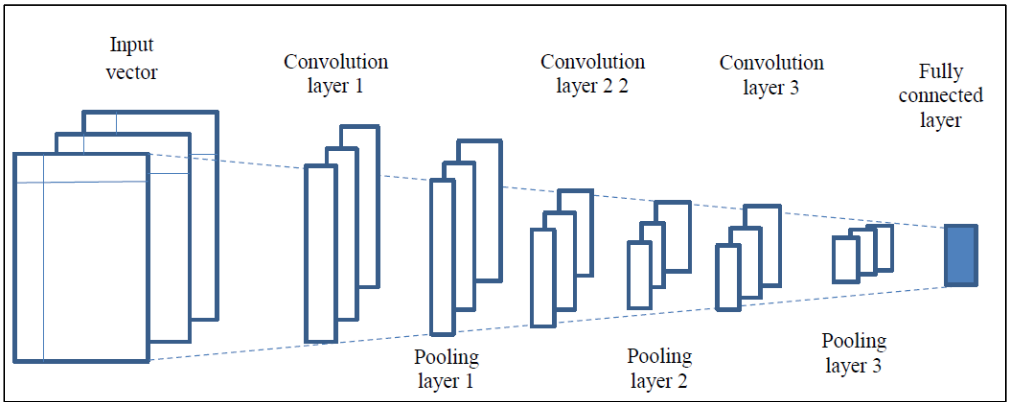

Figure 3: Structure of Convolution layer

3.4 PredictionUsing Optimized LongShort-Term Memory

Following the dimension reduction process, the reduced dimension data is fed into the LSTM's input. The LSTM is the RNN based enhancing structure. For realizing the utilization of information in sentences of the greater distance and to overcome the issue of gradient disappearance in RNN, the LSTM presents a gated mechanism and memory unit. The LSTM unit has three control gates that are input, forget, and output gates. Also, the memory cell state is utilized to store and update information. The mathematical expressions of the LSTM model [13] are defined in equations (4)-(9):

The output of the input gate is defined using (4)

$$I_t = ([h_{t-1}, v_t]W_I + b_I)\,\sigma \quad (4)$$

The output of the forget gate is defined using (5)

$$F_t = ([h_{t-1}, v_t]W_F + b_F)\,\sigma \quad (5)$$

The output of the output gate is defined using (6)

$$O_t = ([h_{t-1}, v_t]W_O + b_O)\,\sigma \quad (6)$$

The current state of the input vector is defined using (7)

$$\tilde{U}_t = \tanh([h_{t-1}, v_t]W_U + b_U) \quad (7)$$

The update state at time t is defined using (8)

$$U_t = \tilde{U}_t * I_t + U_{t-1} * F_t \quad (8)$$

The hidden state output at time t is defined using (9)

$$h_t = \tanh(U_t) * O_t \quad (9)$$

Where tanh and σ denote the hyperbolic tangent function and the sigmoid activation function respectively. v_t denotes the input vector. I_t, F_t, and O_t denote the output of the input, forget, and output gates at time t respectively. b and W denote the bias and weight of the control gates. $\tilde{U}_t$ represents the input's current state. U_t and h_t represent the update state and output at time t. To improve the efficiency of LSTM, the weight values are optimally selected using ROA.

3.5 Parameter Optimization Using ASFOTechnique

The LSTM consist of four weight values $\{W_F, W_I, W_V \text{ and } W_O\}$. These values are degrading the performance of LSTM. In this stage, the weight value is optimally selected by using ASFO.SFO is a population-based heuristic, inspired by nature. Its main idea is to simulate the orientation of sunflowers to receive sunlight [59]. A sunflower row is arranged at regular intervals. They start the day with a movement to observe the sun. In the evening, they go in the opposite direction. It is thought that each sunflower produces only one pollinated comet. To enhance the performance of SFO, the oppositional-based learning (OBL) strategy is adapted with SFO. OBL is used to increase the searching ability and reduce the processing time. The step by step process of ASFO technique based parameter selection is explained below;

Step 1: Solution encoding: In this section, the weight values are considered as the solutions.At first, the solutions are createdarbitrarily. The solutions are represented as flowers. The flower format is given in equation (10).

$$S_i = \{F_1, F_2, \dots, F_n\} \quad (10)$$

$$F_i = \{W_F, W_I, W_V \text{ and } W_O\} \quad (i)(11)$$

Where,

- W_F→weight parameter of forget gate,
- W_I→ weight parameter of the input gate neurons
- W_o→ weight parameter of the output gate neurons
- W_V→Weight parameter oftanH layer

Step 2: Opposite Solution Creation: Then, the opposite solutions are generated based on the initial solutions. The purpose of opposite solution generation is to increase the searching ability.Where $S \in [a,b]$ are a real number and the opposite solution $\overline{S}$ is estimated as follows

$$\overline{S} = a + b - S \qquad (12)$$

Step 3: Fitness Evaluation: Once, we defined the initial flower means, we find out the fitness for each flower. The accuracy is considered for the fitness function. A good classification should have maximum accuracy. The fitness function is given in equation (13).

$$Fitness = Max(Accuracy) \qquad (13)$$

Step 4: Updation using ASFOTechnique: After fitness calculation, the solutions are updated using the ASFO technique.For updation, initially, each flower's (solution) absorbed radiation is calculated using equation (14).

$$R_i = \frac{P}{4\pi d_i^2} \qquad (14)$$

Where; P represent the source power and d_i represent the distance between flower and sun.

The sunflower's orientation to the sun is calculated using equation 15.

$$\vec{O} = \frac{S^* - S_i}{\left\| S^* - S_i \right\|} \qquad (15)$$

Where, S^* represent the present plantationand S_i represent the i^{th} plantation.

The step in the specific direction is calculated using equation (16).

$$D_i = \gamma \times P_i\left(\left\|S_i + S_{i-1}\right\|\right) \times \left\|S_i + S_{i-1}\right\| \qquad (16)$$

Where; γ is a constant value, $P_i(\ )$ represent the pollination probability.

The algorithm controls the maximum given step given in equation 4 to avoid local optimization.

$$X_{\max} = \frac{\left\|S_{\max} - S_{\min}\right\|}{2 \times N_{pop}} \qquad (17)$$

Where, N_{Pop} represent the whole amount of plants $S_{\max}$ and $S_{\min}$ are the minimum and maximum iteration, respectively.

The new plantation is updated using equation (18).

$$\hat{S}_{i+1} = \hat{S}_i + D_i \times O_l \qquad (18)$$

To increase the diversity of the population and increase the local search capability of the SFO, one new operation is added to SFO namely Gauss mutation, which is arbitrarily choosing Si from the solution and creatinga new solution using equation (19).

$$\hat{S}_i = S_i \times e \qquad (19)$$

Where; $e \sim N(0,1)$, in this using equation (15) $\hat{S}_i$ will be mapped to a new location.

$$\hat{S}_i = S_{\max} + \left|\hat{S}_i\right| \% \left(S_{\max} - S_{\min}\right) \qquad (20)$$

To enhance the performance of SFO, Levy Flight (LF) strategy is adapted with SFO. This Levy Flight (LF) strategy is utilized to enhance the initial integration problem, which follows random behavior to deal with local search conditions using Equation (21) - (24). The distribution of taxes is approximate as follows:

$$L(S) \approx |S|^{-1-\varepsilon} \quad (21)$$

$$S = \frac{a}{|b|^{1/\varepsilon}} \quad (22)$$

$$a \sim M\left(0, \sigma_a^2\right),\ b \sim M\left(0, \sigma_b^2\right) \quad (23)$$

$$\sigma_a = \gamma \sqrt{\left\{\frac{\Gamma(1+\varepsilon)\sin(\pi\varepsilon/2)}{\Gamma[(1+\varepsilon)/2]\varepsilon^{2(\varepsilon-1/2)}}\right\}},\ \sigma_b = 1 \quad (24)$$

Based on the LF function, the new updation is calculated using (21).

$$S_{i+1} = (S_i + D_i \times O_l) \times L(S) \quad (25)$$

Step 5: Stopping Criteria: The algorithm is continued until choosing the best flower or best parameter values. Once the best flower is reached, then the algorithm will be stopped. This optimized flower value is used to run the SVM classifier.

3.6CNN-OLSTM Based Sentiment Analysis

The following steps explain how CNN-OLSTM based sentiment analysis works:

Step 1: Pre-processing layer: Initially, the collected data are pre-processed using tokenization, stop word removal, and stemming.

Step 2: Embedding layer: In this layer, the words are converted into the vector using the Word2Vec model. For Word2Vec conversion, SGM is utilized.

Step 3: Convolution layer: The convolution layer gets the input from the embedding layer. With the help of the pooling layer, the convolution layer convolves the information. This will minimize the dimension of information sentences, input boundaries, calculation in the network and will reduce the overfitting issues.

Step 4: Global max-pooling: It gives worldwide good outcomes from the entire network in the wake of applying distinctive convolution layers.

Step 5: Dropout: This layer prevents excess material from the sample. This improves the performance of the model and drops inappropriate data from the network that does not contribute to processing.

Step 6: LSTM layer: After the CNN process, LSTM is used. LSTM uses three types of gates and cells to handle the flow of information in a network.

Step 7: Softmax layer: The final output is carried out in this layer. The obtained score value is 0 indicates the review as negative and 1 indicates, the review as positive

4. Result and Discussion

The results obtained from the proposed sentiment analysisare illustrated in this section. The proposed approach is implemented using Java language on a 2 GHz dual-core PC machine with 8GB of main memory running on the 64-bit version of Windows.The efficiency of the suggested approach is analyzed using accuracy, precision, recall, F-measure, sensitivity, specificity, and training time. For experimental analysis, the data are collected from Twitter using APIs.

4.1 Evaluation Metrics

The effectiveness of the proposed approach is analyzed based on the different metrics such as accuracy, sensitivity, specificity, precision, recall, and F-measure.

Table 1: Evaluation metrics

Metrics	Formula
Accuracy (A)	$A = \frac{TR_P + TR_N}{TR_P + TR_N + FA_P + FA_N} \times 100$
Precision (P)	$P = \frac{TR_P}{TR_P + FA_P} \times 100$
Recall (R)	$R = \frac{TR_P}{TR_P + FA_N} \times 100$
F-Measure (F)	$F = 2 \times \frac{\Pr ecision \times \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call}$
Time	Time is taken for complete process

In table 1, True positive is represented as TR_P, True negative is represented as TR_N, Fault positive is represented as $FA_{P,}$ and fault negative is represented as FA_N.

4.2 Dataset Description

The sentiment analysis in this paper is based on Twitter review data.The Twitter review allows us to analyze the responses to different products, brands, and topics.The dataset is collected from https://www.kaggle.com/kazanova/sentiment140. The dataset consists of 1.6 million tweets. Positive, negative, happy, sad, and other emotions are expressed in the tweets.The use of Twitter reviews allows us to select the best products, brands, etc.

4.3 Performance Analysis

The effectiveness of the suggestedmethod is analyzed in this section. The effectiveness of the suggested approach is discussed based on the different metrics such as accuracy, precision, recall, F-Score, and training time. In this paper, the ASFO technique and LSTM classifier is used for computation. Here, the weight parameter of the LSTM is optimally selected using ASFO Technique. The purpose of the optimal weight selection process is to improve LSTM output and reduce activation time. On the basis of accuracy, Figure 4 analyzes the efficiency of the suggested approach.Analysis of Figure 4 shows that the training process yielded 98.1% and the test process 97.7%.As compared to other approaches, the suggested method provided better results.

In Figure 5, the efficiency of the suggested approach is analyzed based on Precision. Figure 5 shows that the suggested technique had a maximum precision of 97.2 percent for 100 data, which was 82 percent for SVM-based forecasting, 83.5 percent for CNN-based forecasting, and 80 percent for DNN-based forecasting.In comparison to the other ways, the suggested methodology yielded better results, as shown in Figure 5. This is due to the weight optimization procedure using the ASFO approach.

On a recall basis, the efficiency of the suggested strategy is examined in Figure 6.The recall is used to determine how accurate a forecast is.In the case of Figure 6, the suggested approach had a maximum recall of 95.3 percent, with 90 percent for 200 data, 83 percent for 300 data, and 88 percent for 500 data.The value of the recall rapidly diminishes as the number of data or Twitter users grows.In this case, the suggested method outperformed other methods.

In Figure 7, the efficiency of the suggested approach is analyzed based on the F-score. Figure 7 shows that the proposed method has a maximum F-score of 97 percent, which is greater than the existing alternatives.

In Figure 8, the efficiency of the suggested approach is analyzed based on training time. The time it takes to complete the process is measured in time. Figure 8 shows that as the number of data grows, so does the amount of time.

Figure 9 shows statistics of sentiment values from SVM, CNN, DNN, and ASFO-LSTM approaches.For better polarity orientation, the best technique will use a smaller number of words retrieved from the dataset for the sentiment.In comparison to other existing approaches such as SVM, CNN, and DNN, the ASFO-LSTM is recommended since it has the necessary properties to match the impression from a variety of datasets, resulting in better polarity.The results show that the proposed method outperformed alternative approaches.

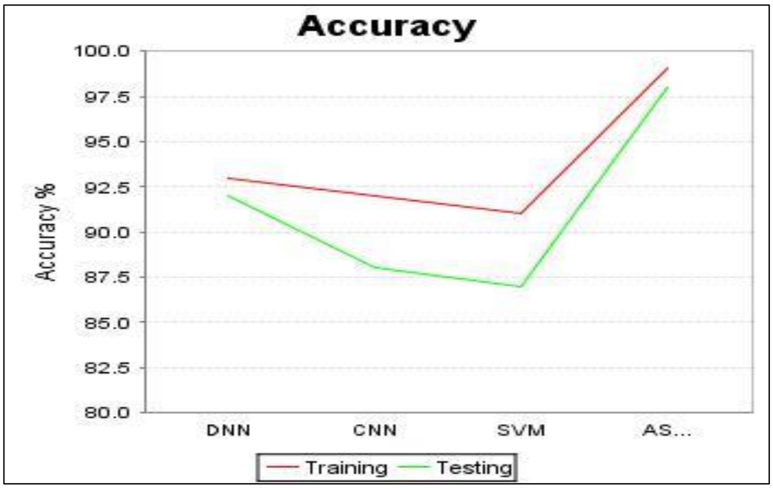

Figure 4: ComparativeAnalysis Based on the Accuracy

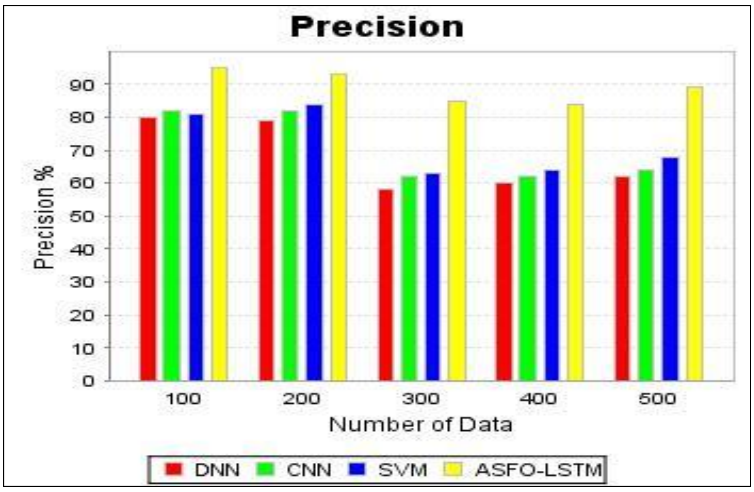

Figure 5.ComparativeAnalysis Based on Precision

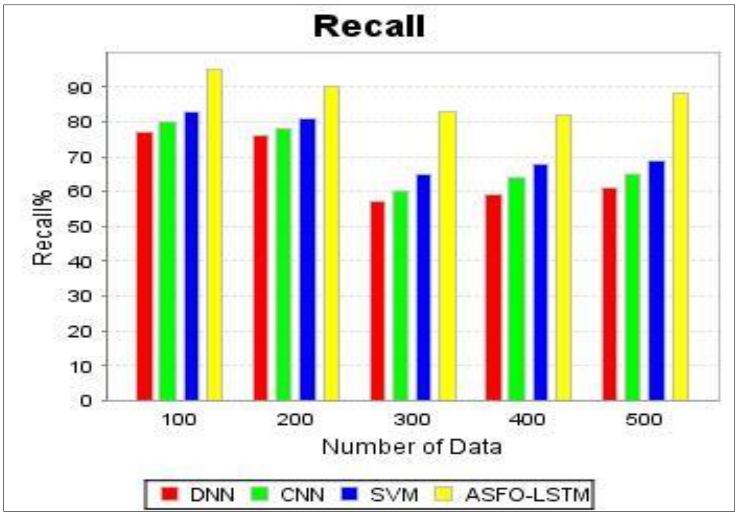

Figure 6.ComparativeAnalysis Based on Recall

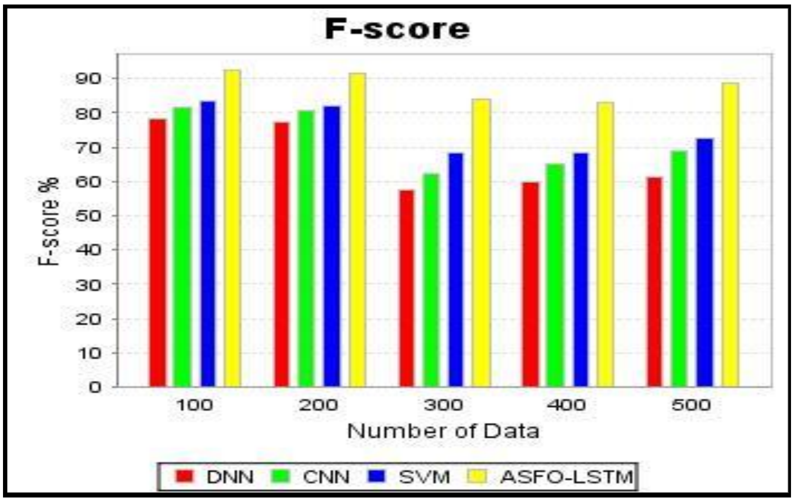

Figure 7: ComparativeAnalysis Based on F-Score

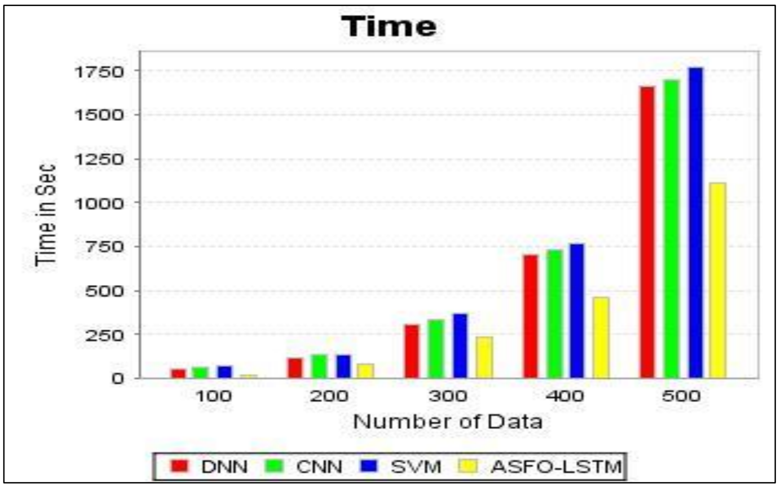

Figure 8.ComparativeAnalysis Based on Training Time

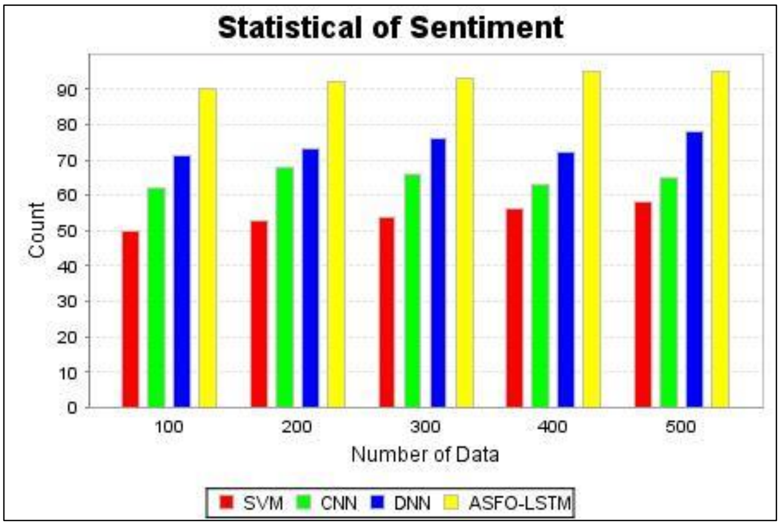

Figure 9.Statisticsof Sentiment Analysis

4.4 ComparativeAnalysis with Published Work

The proposed work was compared with published research [21-25] to illustrate the efficiency of the suggested approach.SVM-based sentiment analysis for Movie Reviews is described in [21].A machine learning algorithm is used to analyze movie reviews and automobile reviews in [22]. As well, in [23], a machine learning algorithm is used to analyze an unsupervised dataset of movie reviews.Hybrid classifiers are used in [24] and [25] to analyze multiple reviews.

Table 2. Comparative analysis results

References	Accuracy (%)
[21]	95.5
[22]	66
[23]	71
[24]	85.4
[25]	66.8

Proposed	98.1

Table 2 summarizes comparative results. Table 2 shows that the recommended approach has proven to be more successful than other approaches.As shown in Table 2, the suggested approach achieved the best accuracy of 98.1%, compared to 95.5% for SVM-based sentiment analysis[21], 66% for unsupervised machine learning algorithm based sentiment analysis[22], and 71% for unsupervised machine learning algorithm based sentiment analysis[23].

5. Conclusion

To deal with text data limitations and high-dimensional difficulties, this paper offers a hybrid model that merges the LSTM mechanism with the CNN network for sentiment analysis.The collected data are first pre-processed, and then the reviews are converted into vector format using the skip-gram model.The extracted feature was then given to the CNN network for input.Following that, the reduced dataset was given to the LSTM.The performance of the proposed CNN-OLSTM classifier is enhanced by using the ASFO technique. Further, different metrics and outcomes were used to compare the performance of the proposed approach with existing methods.The proposed method achieved a maximum accuracy of 93.5% and a maximum precision of 95%.Summarization and entity-based recognition will be the focus in the future.

References:

[1]. B. O'Connor B. R. Routledge, Smith N. A., "From tweets to polls: Linking text sentiment to public opinion time series", InProceedings of the Fourth International AAAI Conference on Weblogs and Social Media,pp: 122-129, 2010.

[2]. José Van Dijck, "Tracing Twitter – The Rise of a Microblogging Platform", International Journal of Media and Cultural Politics, 7(3):333-348, December 2011.

[3]. Mostefai Abdelkader, "A Probabilistic Deep Learning Approach for Twitter Sentiment Analysis", International Journal of Distributed Artificial Intelligence, 2020

[4]. Pak, A., & Paroubek, P, "Twitter as a corpus for sentiment analysis and opinion mining". In Proceedings of the International Conference on Language Resources and Evaluation,Vol. 10, 2010.

[5]. Vishal A. Kharde, S.S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques", International Journal of Computer Applications, 139(11), April 2016.

[6]. Tao Hauff C. Houben G.J. Abel F.,Wachsmuth G., "Facilitating Twitter data analytics: Platform, language, and functionality", IEEE International Conference on Big Data, pp: 421-430, 2014.

[7]. Yin Kann K. Yu M., Schutze H. "Comparative study of CNN and RNN for natural language processing", arXiv preprint arXiv:1702.01923, 2017.

[8]. Neha Gupta and Rashmi Agrawal. "Hybrid Computational Intelligence Challenges and Applications, A volume in Hybrid Computational Intelligence for Pattern Analysis and Understanding", 2020.

[9]. Kalchbrenner Grefenstette E., Blunsom P, "A convolutional neural network for modeling sentences", arXiv preprint arXiv:1404.2188,2014

[10]. Kim Y. "Comparative study of CNN and RNN for natural language processing", arXiv preprint arXiv:1408.5882,2014.

[11]. Hochreiter S., Schmidhuber J, “Long short-term memory, Neural computation”, 9(8): 1735–1780, 1997.

[12]. Denil, Misha & Demiraj, Alban & Kalchbrenner, Nal & Blunsom, Phil & Freitas, Nando, “Modelling, Visualising and Summarising Documents with a Single Convolutional Neural Network”, 2014.

[13]. Ishaani Priyadarshini and Chase Cotton, “A novel LSTM-CNN-grid search-based deep neural network for sentiment analysis”, The Journal of Supercomputing, 77:13911-3932, 2021.

[14]. Samik Datta and Satyaji Chakrabarti, “Aspect based sentiment analysis for demonetization tweets by optimized recurrent neural network using fire fly-oriented multi-verse optimizer”, Sādhanā vol.46, Article number: 79, 2021.

[15]. Muhammad Umer, Imran Ashraf, Arif Mehmood, Saru Kumari, Saleem Ullah, Gyu Sang Choi, “Sentiment Analysis of Tweets Using a Unified Convolutional Neural NetworkLong Short-Term Memory Network Model”**,** Computational Intelligence,37:409-434, 2021.

[16]. Itisha Guptaand Nisheeth Joshi, “Enhanced Twitter Sentiment Analysis Using Hybrid Approach and by Accounting Local Contextual Semantic”, Journal of Intelligent Systems, 29(1):1611-1625,2020.

[17]. Zhao Jianqiang, Gui Xiaolin and Zhang Xuejun, “Deep Convolution Neural Networks for Twitter Sentiment Analysis”, IEEE Access, Vol.8, 2018.

[18]. Dario Stojanovski, Gjorgji Strezoski, Gjorgji Madjarov, and Ivica Dimitrovski, “Twitter Sentiment Analysis using Deep Convolutional Neural Network”, In Proceedings of the Conference: HAIS, June 2015.

[19]. Mingda Wang and Guangmin Hu, “A Novel Method for Twitter Sentiment Analysis Based on Attentional-Graph Neural Network”, Information, 11(2), 92, 2020.

[20]. Willian Becker, JˆonatasWehrmann, Henry E. L.Cagnini, Rodrigo C. Barros, “An Efficient Deep Neural Architecture for Multilingual Sentiment Analysis in Twitter”, Thirtieth International Florida Artificial Intelligence Research Society Conference, May 2017.

[21]. A. Abbasi, H. Chen and A. Salem, “Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums,” In ACM Transactions on Information Systems, 26(3):1-34, 2008.

[22]. P. Turney, “Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews”, Proceedings of the Association for Computational Linguistics (ACL), pp. 417-424, 2002.

[23]. A. Harb, M. Planti, G. Dray, M. Roche, Fran, O. Trousset and P. Poncelet, “Web opinion mining: how to extract opinions from blogs?”, presented at the Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology, Cergy-Pontoise, France, 2008.

[24]. L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, “Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis”, Technical report, HP Laboratories, 2011.

[25]. Ji Fang and Bi Chen, “Incorporating Lexicon Knowledge into SVM Learning to Improve Sentiment Classification”, In Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), pp. 94-100, 2011.