Multimedia Tools and Applications (2023) 82:43273–43296
https://doi.org/10.1007/s11042-023-14767-6

Deep learning-based hybrid sentiment analysis with feature selection using optimization algorithm

D. Anand Joseph Daniel[1] · **M. Janaki Meena**[1]

Received: 16 August 2021 / Revised: 20 January 2023 / Accepted: 31 January 2023 /
Published online: 27 April 2023

Abstract

In the past few years, sentiment analysis (SA) of online content has gained more attention in the research area due to the enormous increase of online content from various sources like websites, social blogs, etc. Many organizations use SA techniques to determine the opinion of users and to ensure their satisfaction. Numerous techniques are suggested by many researchers to identify the sentiments of online content. Among them, hybrid of deep learning and lexicon-based SA techniques are gaining more attention due to their outstanding performance than other approaches. Though the lexicon-based SA approaches integrated with deep learning SA approaches possess more advantages they suffer from lack of accuracy and scalability issues due to the high-dimensional features. To eliminate this issue, a hybrid SA approach is proposed in this paper with a bio-inspired feature selection technique. The Valence Aware Dictionary for Sentiment Reasoning (VADER) approach is integrated with the hybrid deep learning approach of attention-based bidirectional long short-term memory and variable pooling convolutional neural network (VPCNN-ABiLSTM) for SA. The optimal features are selected to minimize the scalability issue by integrating the chimp optimization algorithm with the opposition-based learning technique. The performance of the proposed approach is evaluated for four types of benchmark datasets in terms of precision, accuracy, recall, and F1 score. The proposed approach with OBL-CHOA based feature selection technique achieved higher accuracy of 97.1% with the reduction of 13.6% features. The accuracy of the proposed approach with the feature selection technique is 6.9% higher than the existing BiLSTM-CNN based SA approach.

Keywords Sentiment analysis · VADER · Convolutional neural network · Deep learning · Long short-term memory · Chimp optimization algorithm

✉ D. Anand Joseph Daniel
danny02.20099@gmail.com

M. Janaki Meena
janakimeena.m@vit.ac.in

1 School of Computer Science and Engineering, VIT Chennai, 600127, Tamil Nadu, India

1 Introduction

With the exponential advance of network technology, users express their evaluations, opinions, sentiments, appraisals, emotions, and attitudes on the internet towards entities like organizations. The organizations utilize these online data for users'sentiment identification to improve the satisfaction of users [27]. An opinion extraction process called sentiment analysis (SA) is used by many online communities to extract and identify users' opinions from online reviews. This process can define whether the opinion of the user is neutral, positive, or negative. The SA techniques are mainly classified into lexicon-based and machine learning (ML) based approaches [32]. In lexicon-based methods, the sentiment score in databases is determined through the semantic orientation of words [55]. The main advantage of this technique is labeled data not required to analyze the sentiments. The over fitting problem at any instance is prevented by defining the texts independently [46]. Though this approach shows robust performance on multiple domain databases [31], it requires manual maintenance and the accuracy is less than ML based approaches. In addition, abbreviations in the non-standard form could not be identified due to the limited coverage of this technique on informal texts [11]. At the same time, the ML-based SA approaches utilize various learning algorithms for unstructured content and informal texts [2]. Furthermore, this approach does not require predefined lexicons and so provides more flexibility. However, it needs labeled data to train the classifiers such as Naive Bayes (NB), Support Vector Machine (SVM), Neural Network(NN), Decision Tree (DT), and k-Nearest Neighbor (KNN) [61].

The deep learning-based ML classifier is a powerful classifier implemented in many works due to its effective performance in SA [14]. The Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) models are mostly applied in Natural Language Processing (NLP) tasks to gain effective and remarkable results [10]. The LSTM model captures long-term dependencies between the sequences of word. The presence of max-pooling and convolutional layers in the CNN model can efficiently extract the high-level features. On the other hand, hybrid deep learning models can give improved results than individual models [52]. Different hybrid methods are recommended by researchers to combine the benefits of both ML and lexicon approaches [53]. It can further minimize the limitations and deficiencies faced during individual usage of both these approaches. The combination of two deep learning techniques called CNN-LSTM shows better performance in sentiment classification [5, 35, 56]. To improve the accuracy of the CNN-LSTM approach, Valence Aware Dictionary for Sentiment Reasoning (VADER) [26] strategy is adopted for product data labeling which in turn is directed to train the CNN-LSTM classifier [40, 41]. The main purpose of VADER sentiment is the classification of text sentiments into either positive or negative sentiment. Though this approach shows better performance, the accuracy of these approaches are low and scalability issue may arise when dealing with the high-dimensional dataset. Scalability issue will arise due to higher number of features on training. During classification, considering all possible features on training cannot fit into the memory and causes scalability issues. The application of a feature selection method achieves better scalability and maintains the classifier performance even for larger datasets [13, 30]. The vital role of feature selection is to selectoptimal feature subset without manual involvement. This is helpful to enhance the classifier performance, improve the scalability and reduce the time cost. Nowadays evolutionary approaches are mostly exploited for optimal feature selection due to their effective performance [43]. The two traditional algorithms applied to resolve the complex problems are genetic algorithm (GA) [23] and Particle swarm optimization

(PSO) [12]. A novel optimization technique called Chimp Optimization Algorithm (CHOA) is inspired by the sexual motivation and individual intelligence of chimps in their group hunting process [33]. It is proved to obtain better performance than existing optimization algorithms like PSO and GA [33]. In this paper, LSTM-CNN based classifier with CHOA based feature selection approach is proposed to improve the accuracy while reducing the feature size to eliminate scalability issues. The CHOA is incorporated with opposition-based learning (OBL) to improve its rate of convergence. Different models such as Glove [49], FastText [3], Word2Vec [38], attention mechanism [6], and BERT [17] are commonly used for NLP tasks because of their high-quality word embedding outcomes. The authors of [10] proposed variable pooling convolutional neural network (VPCNN) to extract the semantic features between word vectors. It is a modified version Text CNN network structure which ignores the semantic information between word vectors and only considers the length of the sentence and the maximum feature value without considering other information. The authors of [15] proposed a novel attention mechanism based bidirectional LSTM (BiLSTM) and the convolutional layer text classifier with varying convolution window size and the stride size, which has outperformed the some state-of-the-art DNNs.

The Glove model failed to generate word embedding for those words absent in the dictionary. The attention mechanism adds more weight parameters to the model, which can increase training time especially if the input data for the model are long sequences. Though the BERT model shows better performance, it takes more training time. In the proposed work, the Word2 vec model pre-trained with Fast Text word embedding is used to perform initial feature-based word embedding.

The main contribution of the suggested sentiment analysis technique is summarized as follows.

- To combine the benefits of both the lexicon-based sentiment analysis approach and the deep learning-based sentiment analysis approach, the hybrid of VADER sentiment lexicon and VPCNN-ABiLSTM classifier is used to classify the sentiments of online content.
- To improve the performance of the hybrid VADER and VPCNN-ABiLSTM model by minimizing the scalability issue, a novel feature selection technique is introduced with the integration of Opposition based Learning and chimp optimization algorithm (OBL-CHOA).
- The performance of the proposed approachis evaluated with four datasets such as Amazon, Yelp, IMDB, and Travel. In addition, it is compared with six existing classifiers such as CNN, GRU, CNN-GA, Hybrid GA, BERT, CNN-LSTM, and CNN-BiLSTM in terms of recall, F1 score, precision, and accuracy.
- Moreover, the OBL-CHOA based feature selection is compared with conventional algorithms like PSO, GA, and original CHOA.

This paper is arranged as follows. The related works on SA and optimal feature selection techniques are discussed under Section 2. The background of the proposed approach is detailed in Section 3. The hybrid VPCNN-ABiLSTM classifier with architecture and the mathematical system model is explained in Section 4. Experimental results of the proposed approach, performance comparison and discussion with existing techniques, various classifiers, and datasets are provided in Section 5. In Section 6, this paper is concluded with future work.

2 Related works

In this section, the related works on SA are explained. The main two approaches of SA are lexicon-based and ML-based approaches.

2.1 Lexicon-based approaches

Lexicon-based approaches are the traditional approach for SA [10]. Dictionary-based and corpus-based approaches are the two important techniques to create sentiment lexicons [15]. In dictionary-based approaches, the SA is performed for each word with their polarity scores in the dictionary. It starts with a small sentiment word set and performs well for general purpose applications. The lexicons are then expanded with antonyms and synonyms found in the existing dictionaries. In corpus-based approaches, the SA is performed with sentiment probability for both positive and negative sets of words. This approach can be used for specific domains. VADER, follows certain rules and performs well compared to traditional SA lexicons [52]. Borg and Boldt [22] employed the sentiment lexicon based on the VADER method to predict the sentiments of customer response with SVM fora Swedish telecommunication corporation. It shows better mean values with 0.834 as the F1 score and 0.896 as the AUC. Tanjim et al. [9] offered five different ML models trained with supervised learning for SAof Amazon product reviews. They have achieved 94.02% classification accuracy. Though the lexicon-based SA approaches show better performance, a predefined dictionary is required to label the dataset.

2.2 Machine Learning-based approaches

The ML approaches are classified into three types: unsupervised, semi-supervised, and supervised techniques. Supervised learning uses labeled data for training and gives correct results for a new set of testing data. The unsupervised learning proposed by Li and Liu [4] uses unlabeled data for training. This method group the similar unlabeled data detected through common words into clusters. This is suitable for unstructured content and informal texts. The semi-supervised learning proposed by Da Silva et al. [20] for Twitter data SA utilizes both unlabeled and labeled data for training. In [37], biased Regularized Least squares (bRLS) and biased SVM (bSVM) based semi-supervised models are presented. The performance of the suggested approach outperformed the existing semi-supervised learning methods in terms of computational complexity. Wu et al. [8] proposed semi-supervised Dimensional Sentiment Analysis (DSA) with a variation auto encoder algorithm. This method calculates the sentiment score for several dimensions. Hence, it is obvious that the recent deep learning-based classifiers can give improved performance than existing SVM, NB, KNN, and DT classifiers.

Lee et al. [25] and Nogueira et al. [58] exploited CNN model for SA. This method is computationally less expensive and faster than LSTM, GRU, and Recurrent Neural Network (RNN). Nevertheless, CNN fails to handle long-term dependency. Hassan and Mahmood [36] suggested an RNN model for SA that can handle long-term dependency features. Though RNN performs better than CNN in handling long-term dependency, it fails to handle very long sequences. Moreover, it suffers from exploding and vanishing gradients problems. Huang et al. [45], Peng et al. [21], and Singhal et al. [24] employed LSTM based SA to prevent exploding and vanishing gradients problems in RNN by remembering the previous step input. After calculating the output each time, this technique is computationally expensive for applying the back propagation algorithm. Verma et al. [48] and Rana et al. [54] recommended the GRU model for SA. Though it is less complex than LSTM, its performance is poor on larger texts. The combination of LSTM and CNN has shown better accuracy in [52]. An ensemble model of CNN and LSTM is proposed by Minaee et al. [57]. The performance of this method is better than the existing methods. As a consequence,a similar hybrid VPCNN-ABiLSTM classifier is used in this paper for sentiments classification.

2.3 Hybrid approaches

The text sentiments can be obtained by examining the polarities and frequencies of the positive and negative words using a predefined dictionary in lexicon-based methods [51]. The ML approach predicts a text label by generating features from the text through various learning algorithms. Many hybrid techniques designed from the lexicon-based and ML approaches are employed to improve the accuracy of SA. In these approaches,the ML classifiers are trained with term polarities from lexicons of extra features [39]. Kang et al. [7] combined lexicon-based NB classifier for SA of restaurant review dataset. A similar hybrid approach is developed by Xia et al. [29]. Govindrajan [59] established an arcing hybrid for SA. This approach analyzes the accuracy of an ensemble classifier with GA and NB. A set of hybrid approaches are presented by Gautam et al. [18] for SA of the product reviews. Ortigosa et al. [16] obtained a SA approach using hybrid ML and lexicon methods for Facebook to automatically detect the sentiments with an online product review dataset. An approach for imbalanced threshold classification is introduced by Zou et al. [47]. Agrawal and Nandi [62] found atwo-layered hybrid SA in which one layer is based on ML and another layer is based on the lexicon approach. Rajganesh et al. [44] endorsed SA for a feedback-based recommendation system. The SA for Amazon product reviews is performed by the hybrid method in [50]. The polarity-based lexicon approach is combined with ML to propose a hybrid approach in [42]. Therefore, it is clear that hybrid techniques provide better performance than individual classifiers.

2.4 Optimized feature selection

The accuracy of SA systems can be enhanced via the proper selection of features. Nowadays evolutionary algorithms are proven to obtain better results in feature selection. Besides, bio-inspired algorithms can solve different optimization problems [19]. Kristiyanti et al. [60] applied principal component analysis (PCA), GA, and PSO for feature selection in the SVM classifier. These three algorithms enhanced the classifier accuracy with suitable features selected from Amazon products reviews. Cat Swarm Optimization (CSO) based LSTM model with greedy feature selection is proposed in [34] for SA of big data. Kalarani et al. [1] applied the firefly algorithm (FA) for feature selection in SA of movie reviews with ANN and SVM classifiers. The training time is reduced and the accuracy is improved in this approach. Iqbal et al. [27] proposed GA based feature selection approach to improve the scalability issues of SA in reviews from Yelp, Amazon, and IMDB. The performance of this algorithm is compared with six classifiers in terms of recall, accuracy, F1 score,and precision. In this paper, a similar approach is constructed with VADER sentiment lexicon, OBL-CHOA based feature selection and CNN-LSTM based ML classifier.

3 Background

This section describes the methods used in the proposed SA approach. The details about the CHOA approach, OBL strategy, CNN-LSTM classifier are detailed in this section.

3.1 Chimp optimization algorithm (CHOA)

The CHOA is inspired by the sexual motivation and individual intelligence of chimps in their group hunting. The four types of chimps in a chimp colony are driver, chaser, barrier, and attackers. The prey is followed by the driver. A dam is built by the barriers themselves in a tree across the movement of prey. The movement of chasers is rapidly to catch up with the prey. At last, the attackers hunt the prey. Different abilities of the chimps are considered for a successful hunt. The chimps' hunting process is mainly categorized into two phases: exploration and exploitation. The participation of the chaser, barrier, and driver chimps in the process of hunting is irregular and there is no data about the prey's best location. For mathematical modeling, it is assumed that the optimum location of the prey is informed to the chaser, driver, and barrier by the first attacker. The best four solutions are saved and positions of other chimps are updated based on the location of the best chimps.

3.2 Opposition Based Learning (OBL)

The OBL technique is used by many researchers as an optimization technique to increase the quality of initial populations by diversifying the population. This strategy searches the search space in both directions. The original solution and opposite solution are included in the two directions. The worst solutions are taken from both the solutions. The opposition is applied to the worst solution. The opposite position of original position $x \in [a, b]$ in the jth dimension will be computed using (1).

$$\widetilde{x_j} = a_j + b_j - x_j \tag{1}$$

Where, j= 1, 2, 3,..., i. Here i is the problem dimension. The original position x and opposite position $\widetilde{x}$ will be denoted by the (2) and (3).

$$x = [x_1, x_2, x_3, ..., x_j] \tag{2}$$

$$\widetilde{x} = [\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3, ..., \widetilde{x}_j] \tag{3}$$

If the fitness value of opposite population f $(\widetilde{x})$ is superior than fitness value of original position f(x), then x= $\widetilde{x}$, else x=x. Thus, optimization is performed using the opposite population.

3.3 Variable Pooling Convolution Neural Network and attention-based bidirectional Long Short-Term Memory (VPCNN-ABiLSTM)

The layers of the VPCNN-ABiLSTM classifier are explained below.

Embedding layer: A meaningful and unique words sequence is provided by the Preprocessed dataset. In this layer, random weights are assigned by initializing the words.

Variable Pooling Convolution layer: The VP CNN model has variable convolution layer, variable pooling layer and a fully connected layer. The sentences from the embedding layer are passed to the variable convolutional layers. The variable convolution layers extract features from sentence with length L and word embedding has dimension d respectively. The variable pooling layer convolves the input. The representation of computation in the network, input parameters, input sentences are reduced and the over fitting in the network is controlled with the help of the variable pooling layer.

Variable Pooling layer: It is a layer of CNN used to minimize the computational complexity. The output size of one stack layer to the next is reduced by the pooling techniques to preserve the important information. The most used pooling techniques are max-pooling and average pooling, which is used to pool the features extracted from variable convolution layer and gets detailed semantic information and features between word vectors.

Activation Function: At negative values, this function gives zero, and with positive values, it increases.

ABiLSTM layer: BiLSTM is used to exploits both the preceding and succeeding context representations from the variable convolution layer extracted higher-level phrase representations of the word embedding vectors. Attention mechanism can give important information highlighted contextual word vector by setting varying weights based on its subject importance.

Dense layer: This layer is employed to implement classification on the extracted features from convolutional layers.

SoftMax: This is the final layer. The result of this layer ranges between 0 and 1.

4 Proposed methodology

This section explains the proposed hybrid CNN-LSTM and VADER-based SA with the CHOA-based optimized feature selection approach. The complete architecture of the suggested model is detailed in Fig. 1. The proposed model takes the collected data as the input. Then it performs preprocessing to eliminate unnecessary data. The pre-processed data is labeled using the VADER sentiment lexicon. The labeled dataset is processed into vector form using Word2Vec and the FastText word embedding layer. To minimize the dimensionality of the feature vector, only optimal features are selected using the feature selection

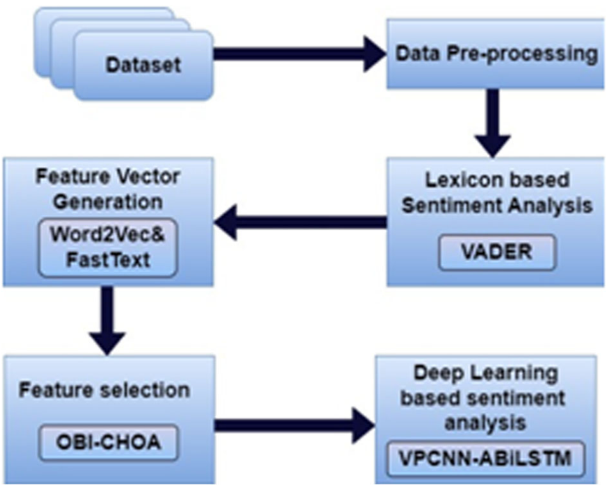

Fig. 1 Proposed model

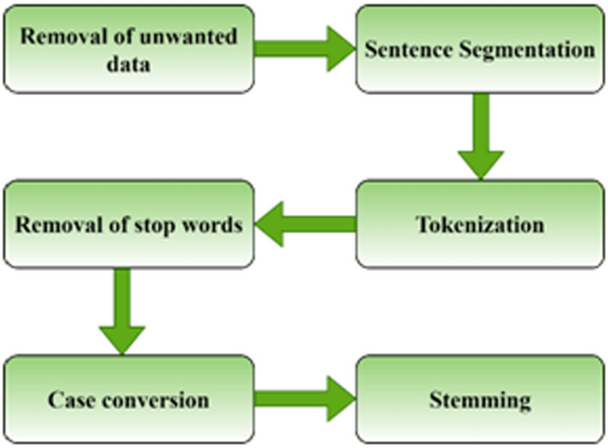

Fig. 2 Data Pre-processing

technique before training the classifier. The OBL-CHOA technique is used in the feature selection approach. The VPCNN-ABiLSTM classifier is trained using the selected features. The high-level features are extracted by the convolution layer and long-term dependencies between words are detected using the ABiLSTM layer.

4.1 Data pre-processing

After manual processing, the dataset is kept in the memory for pre-processing. Segmentation of sentence, Tokenization, removal of stop word, case conversion, and stemming are the steps involved in the preprocessing process as shown in Fig. 2. At first, the unwanted data such as symbols, URLs, digits, web addresses, and online links are cleared from the dataset. All the sentences from the text are segmented in the sentence segmentation step. Then the phrases, symbols, tokens, and keywords are separated from the sentence using the tokenization process. This process is used to remove the punctuation marks. The stop words are cleared from the dataset to improve the accuracy because they have no meaning. All the reviews should be in the same case to process. Hence, all the reviews are converted into lower case. At last, the words are converted into their root for musing the stemming process.

4.2 VADER sentiment labeling

VADER sentiment is a lexicon-based approach that can analyze the sentiment of text into either negative or positive sentiment. The words of pre-processed data and their sentiment intensity are extracted by the VADER sentiment lexicon. The sentiment intensity of each word in the sentence lies in the range of -4 to +4. Here, -4 indicates extremely negative sentiment and +4 indicates extremely positive sentiment. However, the overall scores of words in a sentence are normalized within the range [-1, +1]. This normalized score is termed as a compound score which gives the sentiment of the text. This compound score of the sentence can be calculated as shown in (4).

$$Compoundscore = \frac{x}{\sqrt{x^2 + \alpha}} \tag{4}$$

Table 1 Sentiment labeling using VADER sentiment

Text	VADER Compound score	Sentiment
Good quality benches	0.4404	Positive
Not a good investment	−0.3412	Negative

Here, α is constant and x represents the sum of sentiment intensity of all the words in the sentence. The sentence is labeled as positive if the compound score is greater than 0, others are labeled as negative. Table 1 shows the example for sentiment labeling of furniture data using the VADER sentiment score.

4.3 Feature vector generation

The ML classifiers can understand only numerical values. Hence, it is necessary to convert the features into embedding vectors. It is an efficient and better-performing latent semantic analysis model. The syntactic and semantic information about the words can be extracted via this technique. The semantically same words are arranged to nearby areas and dissimilar words are mapped far away in the vector space. FastText learns each word as characters of n-gram instead of taking the words directly as in the Word2vec model. Thus, this technique breaks down the words into n-grams to vectorize the words that are not in the pre-trained model. The Word2vec model creates word embedding by using a shallow two-layered neural network as shown in Fig. 3. The input layer utilizes one hot encoder and it represents the input word with 0 and 1. It places 1 at the corresponding position of the input word and places 0 at other positions. The hidden layer didn't have any activation function and the softmax classifier is utilized in the output layer. The output is the probability of the input word. In the proposed work, the Word2vec model is combined with pre-trained FastText word embedding to vectorize the features of the dataset. The pre-trained FastText word embedding is utilized to train the Word2Vec model. FastText word embedding is employed

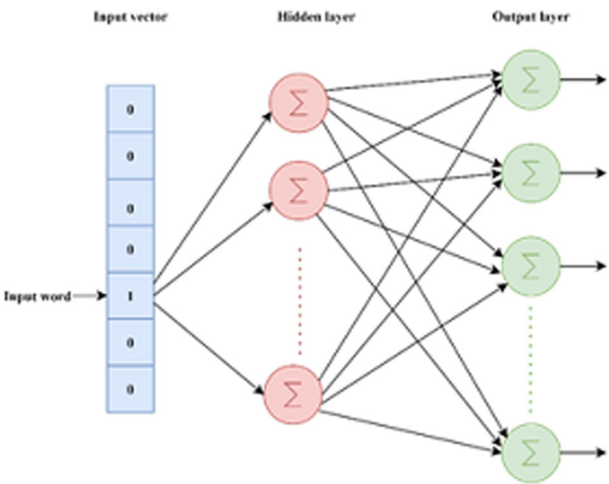

Fig. 3 The architecture of the Word2Vec model

to load the pre-trained word embedding. The text is tokenized into words and the Word2vec returns the embedding of words through pre-trained FastText word embedding.

4.4 Feature selection using OBL-CHOA

The feature vectors with the sentiment label are utilized to train the CNN-LSTM classifier. If the dataset has more features, training the CNN-LSTM classifier may cause the scalability issue because it contains 80% of the input data. This problem may worsen as the dataset size grows larger. To minimize this issue, a novel optimized feature selection technique is introduced in this paper. Proposed opposition-based CHOA generates optimal features from the given feature subset. The flow chart of the suggested OBL-CHOA based feature selection technique is displayed in Fig. 4.

The chimp population $\overrightarrow{P_p}$ is initialized randomly. The parameters m,a & c are initialized using (5–7).

$$c = 2.w_2 \tag{5}$$

$$a = 2.f.w_1 - f \tag{6}$$

$$m = chaoticvalue \tag{7}$$

Here, the iteration process f is non-linearly reduced from 2.5 to 0. The vectors w_1 and w_2 represents the random vectors in the range between 0 and 1. The chaotic vector m is measured from different chaotic maps. Based on the OBL process, the opposite population $\overrightarrow{OP_p}$ is taken from the chimp population $\overrightarrow{P_p}$ using (1). The position of $\overrightarrow{(P_p \cup OP_p)}$ is calculated using (8).

$$x_{chimp}(t+1) = x_{prey}(t) - d.a \tag{8}$$

$$d = |x_{prey}(t).c - m.x_{chimp}(t)| \tag{9}$$

Where the count of the current iteration is indicated by t. The value of d is calculated using (9). x_{prey} represents the position of the solution. The fitness value of x_{chimp} is calculated to determine the best search agents. The fitness function calculates the error rate for the prediction of sentiment values. The search agent whose position possesses a minimum error rate is taken as the best search agent. The fitness value is measured with (10).

$$Fitness = \min\{1 - accuracy\} \tag{10}$$

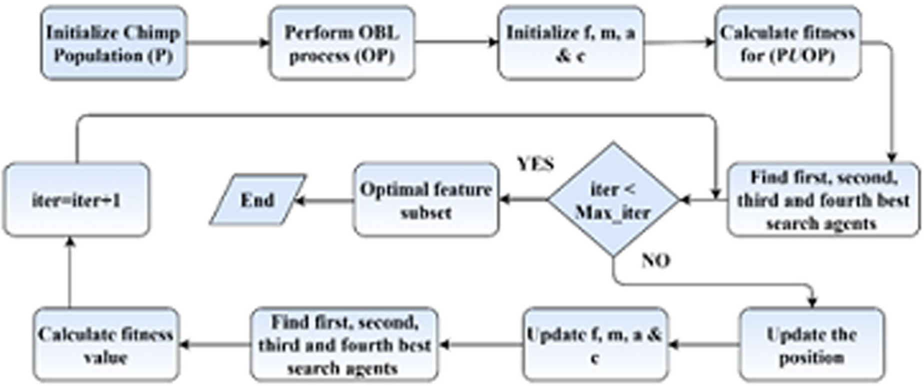

Fig. 4 Feature selection approach flow chart for the proposed OBL-CHOA

The error rate is calculated by using the accuracy of predicted sentiments of x_{chimp}. To calculate the accuracy, the predicted sentiments of search agents' positions are compared with the VADER sentiment labeling. The first, second, third, and fourth best search agents x_a, x_b, x_c, x_d are determined by using (11–14).

$$x_1 = x_a - a_1(d_a) \tag{11}$$

$$x_2 = x_b - a_2(d_b) \tag{12}$$

$$x_3 = x_c - a_3(d_c) \tag{13}$$

$$x_4 = x_d - a_4(d_d) \tag{14}$$

Here, the d_a, d_b, d_c, d_d values are calculated using (15–18). The a_1, a_2, a_3, a_4 values are calculated using (6).

$$d_a = |c_1 x_a - m_1 x| \tag{15}$$

$$d_b = |c_2 x_b - m_2 x| \tag{16}$$

$$d_c = |c_3 x_c - m_3 x| \tag{17}$$

$$d_d = |c_4 x_d - m_4 x| \tag{18}$$

Here, the m_1, m_2, m_3, m_4 are chaotic values that can be calculated using (7). The values of c_1, c_2, c_3, c_4 are calculated using (5). The optimum location of new chimps is updated using (19).

$$x(t+1) = \frac{x_1 + x_2 + x_3 + x_4}{4} \tag{19}$$

This process is repeated until the maximum iteration is reached. The optimal feature set is obtained at the end of this OBL-CHOA-based optimal selection of features. The optimal best solution is taken as the selected feature set. Algorithm 1 explains the pseudocode of OBL-CHOA based feature selection.

4.5 Sentiment classification using VPCNN-ABiLSTM classifier

The architecture of the VPCNN-ABiLSTM classifier is displayed in Fig. 5. The embedding layer is not used in the suggested VPCNN-ABiLSTM model since the vectors are generated by Word2Vec and the FastText model. The VPCNN module includes fully connected layers, varying convolutional layers, and varying pooling layers. The features from sentence, which having sentence length L and word embedding dimension d, are extracted by Variable convolution layer. The selected feature matrix from the feature selection process are represented in the form of matrix of size as $L_i \times d_i$. In the varying pooling layer, adaptive varying max-pooling and average pooling are used to filter the features which has been extracted from variable convolution layer, this method can reduce the computational complexity by minimizing the dimensionality of the feature vector. The activation function employed is Relu. The feature vector output from VPCNN is sent as input to the ABiLSTM through the fully convolution layer. The sigmoid function is used in the classification layer to map the output in the range between 0 and 1.

The four datasets are divided into testing and training data. The model is trained with training data. Then the trained model is used to predict the sentiments of test data. The actual labeling of the dataset is used to compare the results.

5 Experimental results and discussion

This section includes the dataset details, performance metric details, and parameter setting.

Input: Chimp population $\overrightarrow{P_p}$ is Feature set,
Output: Optimal best features
Procedure OBL-CHOA-feature-selection
// Initialization
1. Initialize population $\overrightarrow{P_p}$
2. Calculate opposite population $\overrightarrow{OP_p}$
3. Initialize the parameters f, m, a & c.
//Compute fitness
4. Optimal features $\longleftarrow$ Compute Fitness ($\overrightarrow{P_p \cup OP_p}$)
5. Find the first, second, third, and fourth search agent
while ($Iter < Max_{iter}$) do
6. Update the position of best features
7. Update the parameters f, m, a & c.
8. Find the first, second, third, and fourth search agent
9. Calculate fitness value
10. iter= iter+1
11. end while
12. return Optimal features
end procedure
procedure Compute Fitness ($\overrightarrow{P_p \cup OP_p}$)
13. for $i \longleftarrow$ 1 to N do
14. Fit [i] = Fitness Function ($\overrightarrow{P_p \cup OP_p}(\overrightarrow{l}$, :))
15. end for
16. $Fit_{best} \longleftarrow Best(Fit[])$
17. return Fit_{best}
end procedure

Algorithm 1 OBL-CHOA based feature selection.

5.1 Dataset collection

The reviews of Amazon, IMDB, Yelp, and Travel datasets are utilized in this experiment. The Amazon reviews were collected from Amazon (https://s3.amazonaws.com/amazon-reviews-pds/tsv/index.txt). IMDB reviews were collected from Kaggle (https://www.kaggle.com/lakshmi25npathi/imdb-dataset-of-50k-movie-reviews). Yelp reviews were collected from Kaggle (https://www.kaggle.com/omkarsabnis/yelp-reviews-dataset). Travel reviews were collected from the UCI ML repository (https://archive.ics.uci.edu/ml/datasets/Travel+Reviews). From each category, 1000 negative reviews and 1000 positive reviews have been taken after some manual data cleaning process. Because of the unstructured format of the dataset, it is pre-processed before training the classifier.

5.2 Performance metrics

The performance measures utilized in this paper are accuracy, recall, F1 score, and precision. These performance measures are discussed below.

- **True Positives** The number of correctly identified positive text.

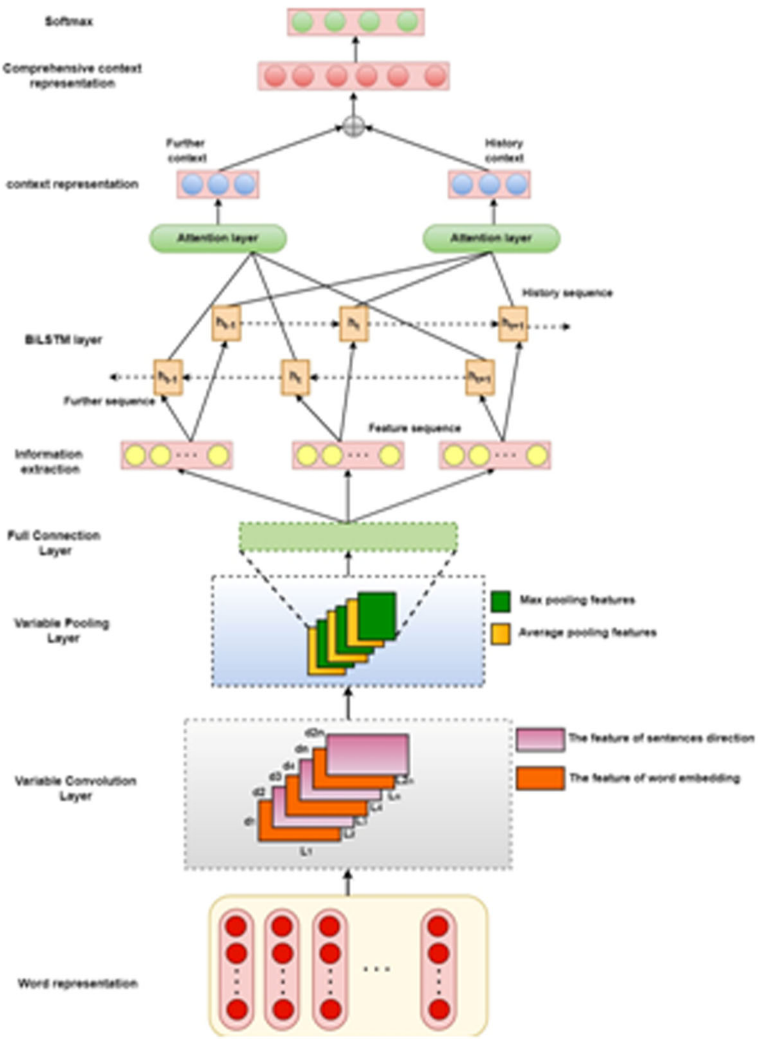

Fig. 5 VPCNN-ABiLSTM classifier

- **False Positives** The number of wrongly identified negative text.
- **True Negatives** The number of correctly identified negative text.
- **False Negatives** The number of wrongly identified positive text.
- **Accuracy** The ratio between the correctly identified text and all texts.It can be given as in (20).

$$Accuracy = TP + TN/TN + TP + FN + FP \tag{20}$$

- **Precision** The ratio of accurately identified positive text over all of positive identified text as mentioned in (21).

$$Precision = TP/FP + TP \tag{21}$$

Table 2 Model parameters for classifier

Parameters	ABiLSTM	VPCNN
Learning rate	0.01	0.01
Dropout	0.2	0.2
Step size	20	20
No of filters	256	256
Embed size	300	300
Batch size	64	64

- **Recall** The ratio of accurately identified positive text overall text actually belonging to that class as mentioned in (22).

$$Recall = TP/FN + TP \tag{22}$$

- **F1-Score** The weighted average of recall and precision is known as F1-score and it can be mentioned as given in (23).

$$F1 - score = 2 * precision * recall/recall + precision \tag{23}$$

5.3 Parameter settings

HP laptop with Intel Core i3 processor having 4GB of RAM and 2.3 GHz frequency, Windows 10 operating system is used to perform the experiments. To evaluate the accuracy of the suggested SA approach, the 10-fold cross-validation method is used. MATLAB R2021a is the software employed to implement the proposed SA approach. The parameters of the proposed classifier are tabulated in Table 2.

The parameters of optimization algorithms set in the proposed model are given in Table 3. In this section, the experimental results of the suggested OBL-CHOA based VPCNN-ABiLSTM approach is evaluated with four different datasets. The results of the proposed model are compared with the conventional ML approaches. Moreover, the feature selection algorithm is compared with existing optimization algorithms.

5.4 AMAZON dataset

5.4.1 Comparing DL classifier under OBL-CHOA optimization

Figure 6 shows the comparison chart for different DL classifiers under OBL-CHOA optimization. The proposed VPCNN-ABiLSTM shows better performance than other classifiers. The accuracy of the VPCNN-ABiLSTM classifier is 96.99% that is 4.31% more

Table 3 Model parameters for feature selection

Algorithm	Parameters	Values
GA	Rate of Mutation	0.01
	Rate of Crossover	0.8
PSO	Weight	0.2
	Constant	2
CHOA	r_1, r_2	random
	m	chaotic

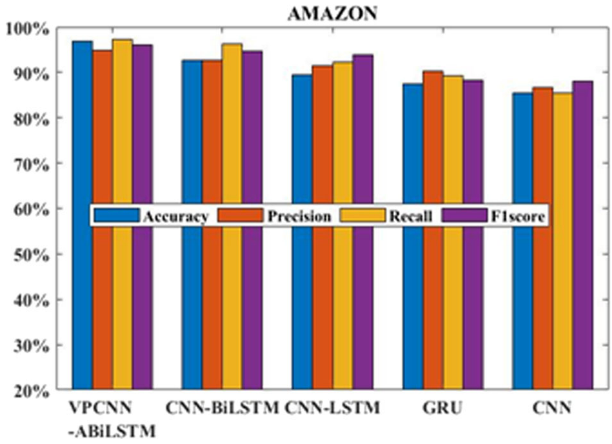

Fig. 6 Performance measures comparison on Amazon dataset

than the CNN-BiLSTM classifier. All the classifiers achieved more than 87.52% accuracy. The precision of CNN-BiLSTM, GRU, CNN-LSTM and proposed VPCNN-ABiLSTM are more than 90% and other classifier achieved more than 85% precision. All the classifiers achieved more than 90% recall. The F1 score of all the five classifiers are more than 90% and the proposed approach achieved a higher F1 score of 96.04%.

5.4.2 Comparison for SA approaches

Three SA approaches like the VADER lexicon-based approach, VPCNN-ABiLSTM based approach without OBL-CHOA, and proposed VPCNN-ABiLSTM based approach with OBL-CHOA are compared. Table 4 displays the performance comparison of three approaches for the Amazon dataset. The performance of the proposed VPCNN-ABiLSTM based approach with OBL-CHOA is better than the other two approaches. The precision of VPCNN-ABiLSTM based approach without OBL-CHOA is 1.4% higher than the suggested approach. But recall, accuracy, and F1 score of the proposed approach are higher than other approaches that are 98.3%, 96.4%, and 97% respectively.

Table 4 Performance measures comparison of three approaches for Amazon dataset

Techniques	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	STD (%)
VADER	93	92.9	96.5	94.7	0.09
VADER+VPCNN-ABiLSTM without chimp	95	95.2	94	94.9	0.07
VADER+VPCNN-ABiLSTM with chimp (Proposed)	96.4	95.1	98.3	97	0.05

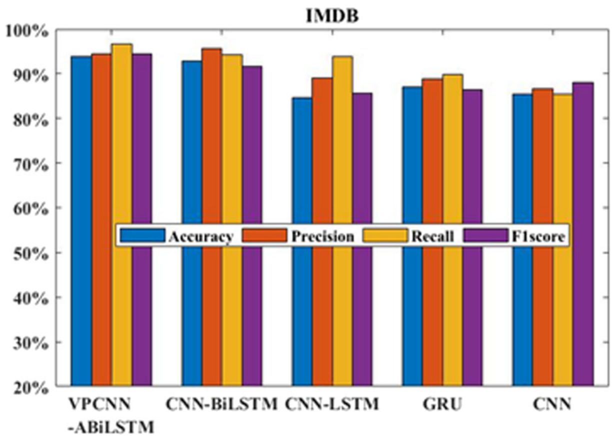

Fig. 7 Performance measures comparison on IMDB dataset

5.5 IMDB dataset

5.5.1 Comparing DL classifier under OBL-CHOA optimization

Figure 7 shows the comparison chart for different DL classifiers under OBL-CHOA optimization using the IMDB dataset. The proposed VPCNN-ABiLSTM shows better performance than other classifiers. The accuracy of the VPCNN-ABiLSTM classifier is 94.87% that is 2% more than the CNN-BiLSTM classifier. All the classifiers achieved more than 85.52% accuracy. The precision of CNN-BiLSTM and proposed VPCNN-ABiLSTM are more than 90% and other classifiers achieved more than 86% precision. All the classifiers achieved more than 91% recall. The F1 score of all the six classifiers except CNN is more than 91% and the proposed approach achieved a higher F1 score of 95.55%.

5.5.2 Comparison for SA approaches

Table 5 displays the performance comparison of three approaches for the IMDB dataset. The performance of the proposed VPCNN-ABiLSTM based approach with OBL-CHOA is better than the other two approaches. The precision, accuracy, recall, and F1 score of the proposed approach is higher than other approaches that are 94.4%, 93.6%, 96.5%, and 95.6% respectively.

Table 5 Performance measures comparison of three approaches for IMDB dataset

Techniques	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	STD (%)
VADER	87	86	88	90	0.16
VADER+VPCNN-ABiLSTM without chimp	92	91.2	89.4	91.6	0.11
VADER+VPCNN-ABiLSTM with chimp (Proposed)	93.6	94.4	96.5	95.3	0.09

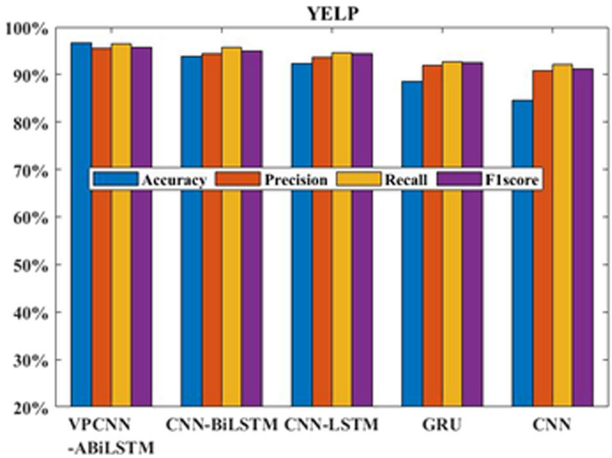

Fig. 8 Performance measures comparison on Yelp dataset

5.6 YELP dataset

5.6.1 Comparing DL classifier under OBL-CHOA optimization

Figure 8 shows the comparison chart for different DL classifiers under OBL-CHOA optimization using the YELP dataset. The proposed VPCNN-ABiLSTM shows better performance than other classifiers. The accuracy of the VPCNN-ABiLSTM classifier is 96.65% that is 4.23% more than the CNN-LSTM classifier. All the classifiers except CNN achieved more than 85% accuracy. The precision of CNN-LSTM, CNN-BiLSTM and proposed VPCNN-ABiLSTM are more than 93% and other classifiers except CNN achieved more than 91% precision. All the classifiers except CNN achieved more than 92.76% recall. The F1 score of all the five classifiers except CNN and GRU is more than 92.53% and the proposed approach achieved a higher F1 score of 95.71%.

5.6.2 Comparison for SA approaches

Table 6 shows the comparison of performance for three SA approaches with the Yelp dataset. The performance of the proposed VPCNN-ABiLSTM based methods with OBL-CHOA is better than the other two approaches. The precision, accuracy, recall, and F1 score of the proposed approach is higher than other approaches that are 94.5%, 96.5%, 96.7%, and 96.4% respectively.

Table 6 Performance measures comparison of three approaches for Yelp dataset

Techniques	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	STD
VADER	88	89.3	95.1	92.2	0.15
VADER+VPCNN-ABiLSTM without chimp	92	93.3	92.6	93.7	0.12
VADER+VPCNN-ABiLSTM with chimp (Proposed)	96.5	94.5	96.7	96.4	0.06

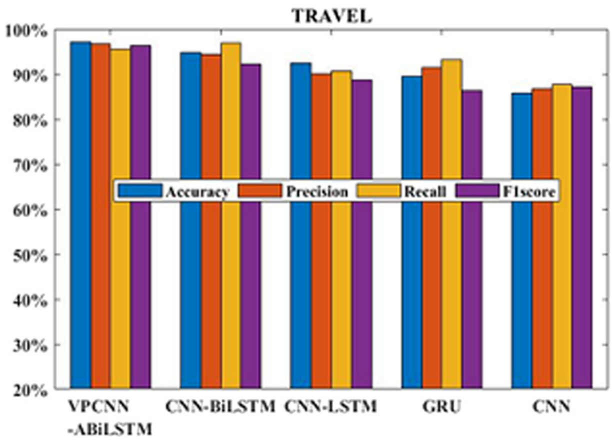

Fig. 9 Performance measures comparison on travel dataset

5.7 Travel dataset

5.7.1 Comparing DL classifier under OBL-CHOA optimization

Figure 9 shows the comparison chart for different DL classifiers under OBL-CHOA optimization using the travel dataset. The proposed VPCNN-ABiLSTM shows better performance than other classifiers. The accuracy of the VPCNN-ABiLSTM classifier is 97.15% that is 3.4% more than the CNN-BiLSTM classifier. All the classifiers achieved more than 90% accuracy. The precision of GRU, CNN-LSTM, CNN-BiGRU, and proposed VPCNN-ABiLSTM are more than 91.5% and other classifiers achieved more than 89% precision. All the five classifiers achieved more than 88% recall. The F1 score of all the four classifiers except GRU and CNN is more than 95% and the proposed approach achieved a higher F1 score of 96.45%

5.7.2 Comparison for SA approaches

Table 7 displays the performance comparison of three approaches for the travel dataset. The performance of the proposed VPCNN-ABiLSTM based approach with OBL-CHOA is better than the other two approaches. The precision, accuracy, recall, and F1 score of the proposed approach is higher than other approaches that are 97.5%, 97.1%, 93%, and 97.6% respectively. It can be known from the results that the hybrid approach can give better results than individual approaches.

Table 7 Performance measures comparison of three approaches for Travel dataset

Techniques	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	STD (%)
VADER	92	91.5	90.2	92.34	0.11
VADER+VPCNN-ABiLSTM without chimp	95	96.1	97.1	96.4	0.08
VADER+VPCNN-ABiLSTM with chimp (Proposed)	97.1	97.5	93	97.6	0.05

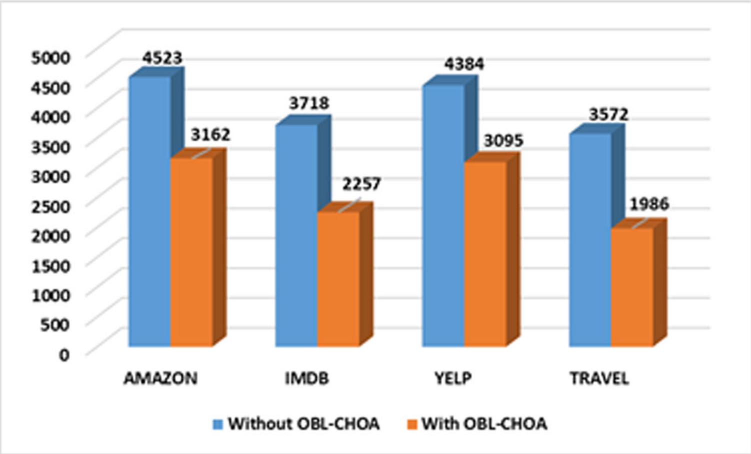

Fig. 10 Feature size comparison before and after feature selection

5.8 Feature size comparison

The feature size comparison of four datasets before and after the feature selection process is calculated to display the effectiveness of the proposed feature selection algorithm. Figure 10 shows the comparison chart for feature size. For the Amazon dataset, 13.6% of features are reduced by the suggested feature selection approach. For the Yelp dataset, the feature selection approach minimized the size of the features from 4384 to 3095. The features size of the IMDB dataset before the feature selection process is 3718. It is reduced to 2257 by the feature selection approach. For the travel dataset, 15.9% of features are reduced by the feature selection approach. Thus the proposed approach can reduce the scalability problem by reducing the dimensionality of the dataset.

5.9 Analysis of the variance (ANOVA) test

The test for important variance among the mean scores for the four datasets is taken by the analysis of variance. The results of the ANOVA test are shown in Table 8. From the results, it can be known that among the sentiment score mean there is no difference. The ANOVA test depicts that with OBL-CHOA select less number of features which are more important, it retain all more meaningful information.

Table 8 ANOVA test r0 for different datasets

Dataset	Feature selection with OBL-CHOA				Feature selection without OBL-CHOA			
	N	Mean	F value	Sig	N	Mean	F value	Sig
AMAZON	92	35.28	21.67	0.232	96	37.32	22.34	0.218
IMDB	63	35.97			65	36.01		
YELP	24	46.98			27	47.34		
TRAVEL	20	28.28			23	29.79		

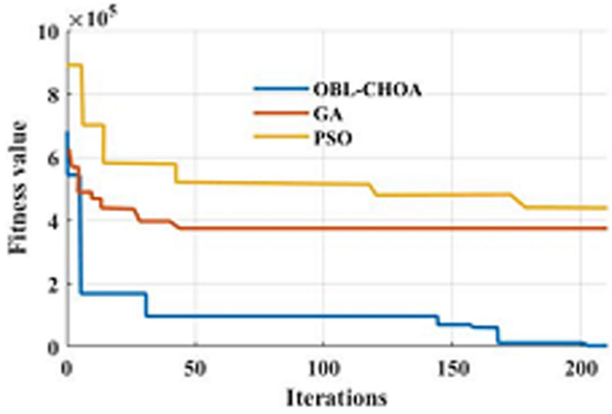

Fig. 11 Convergence curve comparison for PSO, GA, and OBL-CHOA

5.10 Comparing OBL-CHOA with GA and PSO

To show the effectiveness of the convergence curve of the proposed OBL-CHOA is compared with two popular conventional algorithms: PSO and GA. The results are taken only with the proposed classifier.

Figure 11 illustrates that in the feature selection task, the optimum solution is found by OBL-CHOA faster than the existing optimization algorithms with an improved convergence speed. Thus, the balance between the exploitation and exploration is maintained with the proposed OBL-CHOA based feature selection algorithm and it can escape from the local optima issue.

5.11 Comparison of the proposed approach with existing works

Table 9 displays the comparison between the suggested method and related works in terms of accuracy. The hybrid deep learning and lexicon-based SA approach are employed in the proposed approach with OBL-CHOA based feature selection approach. It improves

Table 9 Proposed model with the existing model

Methods	Dataset	Accuracy (%)
CNN-GA [28]	Amazon	92
	IMDB	91.5
Hybrid GA [27]	Amazon	78.3
	IMDB	76.7
	Yelp	75.1
LSTM-CNN [52]	IMDB	86.6
BERT [17]	Amazon	88.48
Proposed model	Amazon	96.4
	Yelp	96.5
	IMDB	93.6

the scalability of the proposed work in terms of execution time. When compared to existing works, the proposed approach shows improved performance with reduced feature size. The maximum accuracy achieved by the proposed SA technique is 95.2% on the Amazon dataset while other datasets achieved lesser accuracy. For the IMDB dataset, the proposed ABiLSTM-VPCNN based model with feature selection technique achieved 5.9% improved accuracy than the ABiLSTM-VPCNN approach [52] without feature selection technique.

5.12 Discussion

At first, the OBL-CHOA based feature selection technique is applied to seven existing DL/ML classifiers such as CNN, CNN-GA, Hybrid GA, BERT, GRU, CNN-LSTM, CNN-BiLSTM, and VPCNN-BiLSTM. From the comparison results, it can be known that the VPCNN-ABiLSTM classifier with OBL-CHOA based feature selection technique can give better performance. Then the results are taken for the VADER sentiment lexicon. Then the results are taken for the hybrid of VADER and VPCNN-ABiLSTM approach without feature selection method. At last, the results are taken with the feature selection technique. From these results, it can be known that the proposed feature selection approach with a hybrid of VADER and VPCNN-ABiLSTM can minimize the scalability and maximize the accuracy. To evaluate the convergence performance of the suggested feature selection algorithm, the convergence curve comparison is taken for the proposed OBL-CHOA, PSO, and GA. The proposed OBL-CHOA shows better convergence speed compared to other algorithms. Though existing approaches for sentiment analysis have better accuracy, the scalability issue is there. So that, the proposed approach considers a novel feature selection technique that can give better performance than other feature selection techniques.

6 Conclusion

In this paper, a hybrid deep learning technique based on the VPCNN-ABiLSTM classifier is proposed with VADER based sentiment lexicon to label the data. Moreover, the scalability issue is solved by the integration of a novel opposition learning-based chimp optimization feature selection algorithm. Four types of datasets such as Amazon, IMDB, YELP, and Travel are used to evaluate the working of the proposed approach. The precision, accuracy, recall, and F1 score of the proposed approach are calculated and compared with seven existing DL/ML approaches such as CNN, CNN-GA, Hybrid GA, BERT, GRU, CNN-LSTM and CNN-BiLSTM. The optimization algorithm is compared with existing algorithms like PSO, GA, and original CHOA.The proposed VPCNN-ABiLSTM based Sentiment Analysis approach with OBL-CHOA based feature selection technique has achieved higher accuracy of 96.5% with the reduction of 13.6% feature. The proposed approach with 15.9% reduced features achieved 5.9% higher accuracy than the ABiLSTM-VPCNN approach without feature selection technique. The future work aims to propose a deep learning-based ensemble classifier for all types of datasets.

Acknowledgments I confirm that all authors listed on the title page have contributed significantly to the work, have read the manuscript, attest to the validity and legitimacy of the data and its interpretation, and agree to its submission.

Author Contributions I confirm that all authors listed on the title page have contributed significantly to the work, have read the manuscript, attest to the validity and legitimacy of the data and its interpretation, and agree to its submission.

Funding Funding information is not applicable because no funding was received

Data Availability Data will be shared on the reasonable request.

Declarations

Ethics approval Funding information is not applicable because no funding was received, so he name of the approving body and the approval numberID not applicable to this manuscript.

Consent for Publication I, give my consent for the publication of above Article.I declare that I shall not submit the paper for publication in any other Journal or Magazine till the decision is made by journal editors.

Conflict of Interests The authors declare that they have no conflict of interest.

References

1. Alarifi A, Tolba A, Al-Makhadmeh Z, Said W (2018) A big data approach to sentiment analysis using greedy feature selection with cat swarm optimization-based long short-term memory neural networks. J Supercomput 76:4414–4429. https://doi.org/10.1007/s11227-018-2398-2
2. Boiy E, Moens M-F (2008) A machine learning approach to sentiment analysis in multilingual Web texts. Inf Retr 12(5):526–558. https://doi.org/10.1007/s10791-008-9070-z
3. Bojanowski P, Grave E, Joulin A, Mikolov T (2017) En-riching word vectors with subword information. Transactions of the Association for Computational Linguistics 5:135–146. https://doi.org/10.1162/tacl-a-00051
4. Borg A, Boldt M (2020) Using VADER sentiment and SVM for predicting customer response sentiment. Expert Syst Appl 162:113746. https://doi.org/10.1016/j.eswa.2020.113746
5. Cahyadi A, Khodra ML (2018) Aspect-based sentiment analysis using convolutional neural network and bidirectional long short-term memory. In: 2018 5th international conference on advanced informatics: concept theoryand applications (ICAICTA), pp 124–129. https://doi.org/10.1109/ICAICTA.2018.8541300
6. Cai Y, Huang Q, Lin Z, Xu J, Chen Z, Li Q (2020) Recurrent neural network with pooling operation and attention mechanism for sentiment analysis: a multi-task learning approach. Knowl-Based Syst 203:105856. https://doi.org/10.1016/j.knosys.2020.105856
7. Cho H, Kim S, Lee J, Lee J-S (2014) Data-driven integration of multiple sentiment dictionaries for lexicon-based sentiment classification of product reviews. Knowl-Based Syst 71:61–71. https://doi.org/10.1016/j.knosys.2014.06.001
8. Da Silva NFF, Coletta LFS, Hruschka ER, Hruschka ERJ (2016) Using unsupervised information to improve semi-supervised tweet sentiment classification. Inform Sci 355–356:348–365. https://doi.org/10.1016/j.ins.2016.02.002
9. Darwich M, Noah SAM, Omar N, Osman N (2019) Corpus-Based Techniques for sentiment lexicon generation: a review. J Digit Inf Manag 17(5):296–305. https://doi.org/10.6025/jdim/2019/17/5/296-305
10. Dong M, Li Y, Tang X, Xu J, Bi S, Cai Y (2020) Variable convolution and pooling convolutional neural network for text sentiment classification. IEEE Access 8:16174–16186. https://doi.org/10.1109/ACCESS.2020.2966726
11. Ducange P, Fazzolari M, Petrocchi M, Vecchio M (2019) An effective Decision Support System for social media listening based on cross-source sentiment analysis models. Eng Appl Artif Intell 78:71–85. https://doi.org/10.1016/j.engappai.2018.10.014
12. Eberhart R, Kennedy J (n.d.) A new optimizer using particle swarm theory. MHS'95. In: Proceedings of the Sixth international symposium on micro machine and human science, pp 39-43. https://doi.org/10.1109/mhs.1995.494215
13. Fathi M, Haghi Kashani M, Jameii SM, Mahdipour E (2021) Big data analytics in weather forecasting: a systematic review. Archives of Computational Methods in Engineering 27:1247–1275. https://doi.org/10.1007/s11831-021-09616-4
14. Fu X, Yang J, Li J, Fang M, Wang H (2018) Lexicon-enhanced LSTM with attention for general sentiment analysis. IEEE Access 6:71884–71891. https://doi.org/10.1109/access.2018.2878425
15. Gang L, Jiabao G (2019) Bidirectional LSTM with attention mechanism and convolutional layer for text classification. Neurocomputing 337:325–338. https://doi.org/10.1016/j.neucom.2019.01.078

16. Gautam G, Yadav D (2014) Sentiment analysis of twitter data using machine learning approaches and semantic analysis. In: 2014 Seventh international conference on contemporary computing (IC3), pp 437–442. https://doi.org/10.1109/ic3.2014.6897213
17. Geetha MP, Renuka D (2021) Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model. Int J Intell Netw 2:64–69. https://doi.org/10.1016/j.ijin.2021.06.005
18. Govindarajan M (2013) Sentiment analysis of movie reviews using hybrid method of naive Bayes and genetic algorithm. Int J Adv Comput Res 3(4):139–145
19. Gupta I, Joshi N (2019) Enhanced twitter sentiment analysis using hybrid approach and by accounting local contextual semantic. J Intell Syst 29(1):1611–1625. https://doi.org/10.1515/jisys-2019-0106
20. Haque TU, Saber NN, Shah FM (2018) Sentiment analysis on large scale Amazon product reviews. In: 2018 IEEE international conference on innovative research and development (ICIRD), pp 1–6. https://doi.org/10.1109/icird.2018.8376299
21. Hassan A, Mahmood A (2018) Convolutional recurrent deep learning model for sentence classification. IEEE Access 6:13949–13957. https://doi.org/10.1109/ACCESS.2018.2814818
22. Hemmatian F, Sohrabi MK (2017) A survey on classification techniques for opinion mining and sentiment analysis. Artif Intell Rev 52(3):1495–1545. https://doi.org/10.1007/s1046 2-017-9599-6.
23. Holland JH (1975) Adaptation in natural and artificial systems: an introductory analysis with applications to biology. Control, and Artificial Intelligence, The MIT Press
24. Huang W, Rao G, Feng Z, Cong Q (2018) LSTM with sentence representations for document-level sentiment classification. Neurocomputing 308:49–57. https://doi.org/10.1016/j.neucom.2018.04.045
25. Hussain A, Cambria E (2018) Semi-supervised learning for big social data analysis. Neurocomputing 275:1662–1673. https://doi.org/10.1016/j.neucom.2017.10.010
26. Hutto CJ, Gilbert E (2015) VADER: A parsimonious rule-based model for sentiment analysis of social media text. Proceedings of the 8th International Conference on Weblogs and Social Media 8(1):216–225. https://doi.org/10.1609/icwsm.v8i1.14550
27. Iqbal F, Maqbool J, Fung BCM, Batool R, Khattak AM, Aleem S, Hung PCK (2019) A hybrid framework for sentiment analysis using genetic algorithm based feature reduction. IEEE Access 7:14637–14652. https://doi.org/10.1109/access.2019.2892852
28. Kalarani P, Selva Brunda S (2018) Sentiment analysis by POS and joint sentiment topic features using SVM and ANN. Soft Comput 23:7067–7079. https://doi.org/10.1007/s00500-018-3349-9
29. Kang H, Yoo SJ, Han D (2012) Senti-lexicon and improved naïve Bayes algorithms for sentiment analysis of restaurant reviews. Expert Syst Appl 39(5):6000–6010. https://doi.org/10.1016/j.eswa.2011.11.107
30. Karimi Y, Haghi Kashani M, Akbari M, Mahdipour E (2021) Leveraging big data in smart cities: a systematic review. Concurrency and Computation: Practice and Experience 33(21):e6379. https://doi.org/10.1002/cpe.6379
31. Katz G, Ofek N, Shapira B (2015) Consent: context-based sentiment analysis. Knowl-Based Syst 84:162–178. https://doi.org/10.1016/j.knosys.2015.04.009
32. Kausar S, Huahu X, Shabir MY, Ahmad W (2020) A sentiment polarity categorization technique for online product reviews. IEEE Access 8:3594–3605. https://doi.org/10.1109/access.2019.2963020
33. Khishe M, Mosavi MR (2020) Chimp optimization algorithm. Expert Syst Appl 149:113338. https://doi.org/10.1016/j.eswa.2020.113338
34. Kristiyanti DA, Wahyudi M (2017) Feature selection based on Genetic algorithm, particle swarm optimization and principal component analysis for opinion mining cosmetic product review. In: 2017 5th international conference on cyber and it service management (CITSM), pp 1–6. https://doi.org/10.1109/citsm.2017.8089278
35. Kumar SS, Kumar MA, Soman KP (2017) Sentiment analysis of tweets in malayalam using long Short-TermMemory units and convolutional neural nets. In: Ghosh A, Pal R, Prasath R (eds) Mining intelligence and knowledgeexploration. Springer, Cham
36. Lee G, Jeong J, Seo S, Kim CY, Kang P (2018) Sentiment classification with word localization based on weakly supervised learning with a convolutional neural network. Knowl-Based Syst 152:70–82. https://doi.org/10.1016/j.knosys.2018.04.006
37. Li G, Liu F (2014) Sentiment analysis based on clustering: a framework in improving accuracy and recognizing neutral opinions. Appl Intell 40(3):441–452. https://doi.org/10.1007/s1048 9-013-0463-3
38. Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. ICLR. arXiv:301.3781
39. Minaee S, Azimi E, Abdolrashidi A (2019) Deep-sentiment: sentiment analysis using ensemble of cnn and bi-lstm models. https://doi.org/10.48550/arXiv.1904.04206
40. Mishra RM, Urolagin S, Angel AJJ, Neogi A, Nawaz N (2021) Deep learning-based sentiment analysis and topic modeling on tourism during Covid-19 pandemic. Frontiers in Computer Science 3. https://doi.org/10.3389/fcomp.2021.775368

41. Mujahid M, Lee E, Rustam F, Washington PB, Ullah S, Reshi AA, Ashraf I (2021) Sentiment analysis and topic modeling on tweets about online education during COVID-19. Appl Sci 11(18):8438. https://doi.org/10.3390/ app11188438
42. Mumtaz D, Ahuja B (2017) A lexical and machine Learning-Based hybrid system for sentiment analysis. Studies in Computational Intelligence, pp 165–175. https://doi.org/10.1007/978-981-10-4555-4-11
43. Nabaei A, Hamian M, Parsaei MR, Safdari R, Samad-Soltani T, Zarrabi H, Ghassemi A (2016) Topologies and performance of intelligent algorithms: a comprehensive review. Artif Intell Rev 49(1):79–103. https://doi.org/10.1007/s10462-016-9517-3
44. Nandi V, Agrawal S (2016) Political sentiment analysis using hybrid approach. Int Res J Eng Technol (IRJET) 3(5):1621–1627
45. Nogueira CSD, Gatti M (2014) Deep convolutional neural networks for sentiment analysis of short texts. In: Proceedings of 25th international conference on computational linguistics, pp 69–78
46. Ortigosa A, Martín JM, Carro RM (2014) Sentiment analysis in Facebook and its application to e-learning. Comput Hum Behav 31:527–541. https://doi.org/10.1016/j.chb.2013.05.024
47. Ortigosa A, Martín JM, Carro RM (2014) Sentiment analysis in Facebook and its application to e-learning. Comput Hum Behav 31:527–541. https://doi.org/10.1016/j.chb.2013.05.024
48. Peng H, Ma Y, Li Y, Cambria E (2018) Learning multi-grained aspect target sequence for Chinese sentiment analysis. Knowl-Based Syst 148:167–176. https://doi.org/10.1016/j.knosys.2018.02.034
49. Pennington J, Socher R, Manning C (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp 1532–1543. https://doi.org/10.3115/v1/D14-1162
50. Rajganesh N, Asha C, Keerthana AT, Suriya K (2018) A hybrid feedback-based book recommendation system using sentiment analysis. IJSRCSEIT 3(3):2456–3307
51. Rana R (2016) Gated recurrent unit (GRU) for emotion classification from noisy speech. arXiv:1612.07778
52. Rehman AU, Malik AK, Raza B, Ali W (2019) A hybrid CNN-LSTM model for improving accuracy of movie reviews sentiment analysis. Multimed Tools Appl 78:26597–26613. https://doi.org/10.1007/s11042-019-07788-7
53. Shang L, Zhou Z, Liu X (2016) Particle swarm optimization-based feature selection in sentiment classification. Soft Comput 20(10):3821–3834. https://doi.org/10.1007/s00500-016-2093-2
54. Singhal P, Bhattacharyya P (2016) Sentiment analysis and deep learning: a survey. In: Center for Indian Language Technology, Indian Institute of Technology, Bombay. https://code.google.com/archive/p/word2vec/
55. Taboada M, Brooke J, Tofiloski M, Voll K, Stede M (2011) Lexicon-Based Methods for sentiment analysis. Comput Linguist 37(2):267–307. https://doi.org/10.1162/coli-a-0049
56. Umer M, Ashraf I, Mehmood A, Kumari S, Ullah S, Choi GS (2021) Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model. Comput Intell 37(1):409–434. https://doi.org/10.1111/coin.12415
57. Verma S, Saini M, Sharan A (2018) Deep sequential model for review rating prediction. In: 10th international conference on contemporary computing (IC3), pp 1–6. https://doi.org/10.1109/IC3.2017.8284318.
58. Wu C, Wu F, Wu S, Yuan Z, Liu J, Huang Y (2019) Semi-supervised dimensional sentiment analysis with variational autoencoder. Knowl-Based Syst 165:30–39. https://doi.org/10.1016/j.knosys.2018.11.018
59. Xia R, Zong C, Li S (2011) Ensemble of feature sets and classification algorithms for sentiment classification. Inf Sci 181(6):1138–1152. https://doi.org/10.1016/j.ins.2010.11.023
60. Yadav A, Vishwakarma DK (2020) A comparative study on bio-inspired algorithms for sentiment analysis. Clust Comput 23:2969–2989. https://doi.org/10.1007/s10586-020-03062-w
61. Yuan Z, Wu S, Wu F, Liu J, Huang Y (2018) Domain attention model for multi-domain sentiment classification. Knowl-Based Syst 155:1–10. https://doi.org/10.1016/j.knosys.2018.05.004
62. Zou Q, Xie S, Lin Z, Wu M, Ju Y (2016) Finding the best classification threshold in imbalanced classification. Big Data Res 5:2–8. https://doi.org/10.1016/j.bdr.2015.12.001

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.