



SenDemonNet: sentiment analysis for demonetization tweets using heuristic deep neural network

Şafak Kayıkçı¹

Received: 17 July 2021 / Revised: 26 November 2021 / Accepted: 3 January 2022 /
Published online: 17 February 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Sentiment analysis is one of the efficient models for extracting opinion mining with identification and classification from unstructured text data such as product reviews or microblogs. It is used to gain feedback from political campaigns, brand reviews, marketing analysis, and customers. The sentiment analysis on Twitter data is a recent research field in the natural processing. The dataset is gathered from the “Twitter” package in R along with Twitter API. The main intent of this paper is to understand the public opinion on the recently implemented demonetization policy using the proposed SenDemonNet. Initially, the tweet preprocessing was done, which is intended for cleaning the text data. Then, the feature extraction is performed by Bag of n-grams, TF-IDF, and the word2vec algorithm. The main objective of this work is a weighted feature selection that is developed by the hybrid Forest–Whale Optimization Algorithm (F-WOA) to get the best classification outcome. With these features, the Heuristic Deep Neural Network (HDNN) is adopted for classification, where the proposed FOA and WOA tune the parameter of DNN for reaching the maximum accuracy rate. From the statistical analysis, the performance of the designed F-WOA-DNN is 1.8%, 1.9%, 1.86%, and 2% enhanced than PSO-DNN, GWO-DNN, WOA-DNN, FOA-DNN, SVM, CNN, LSTM, and DNN respectively. Extensive experimental results show that SenDemonNet outperforms its competitors, producing an impressive increase in the classification accuracy on the benchmark dataset.

Keywords Sentiment analysis · SenDemonNet · Demonetization policy · Twitter data · Forest-whale optimization algorithm · Weighted feature selection · Heuristic deep neural network

✉ Şafak Kayıkçı
safak.kayikci@ibu.edu.tr

¹ Department of Computer Engineering, Bolu Abant İzzet Baysal University, BAİBÜ Gölüköy Yerleşkesi, 14030 Merkez/Bolu, Turkey

1 Introduction

Demonetization is “the process of removing currency from general usage or circulation.” In India, the Demonetization was announced on November 8, 2016, to stop the 500 and 1000 rupee banknotes [34]. It is the act of changing a currency unit of their status as an authorized one, which is occurred when there is a need of changing the national currency based on the enforcement of Demonetization [6] like improvement of trade, eradication of crime and corruption, and management of inflation [39]. Among them, corruption is the major cause for enforcing Demonetization [43], which assists in tax evasion and control frauds that also discourages a cash-based economy. Moreover, inflation must be handled carefully to increase the value of money and to minimize the huge soaring costs. Around the world, different countries have been announced Demonetization. As an example, in the United States, silver was demonetized for addressing the economic depression in 1873. Similarly, European nations have broadcasted Demonetization by changing the Euro as their official currency. Similarly, Zimbabwe demonetized their Zimbabwean dollar into US dollars in 2015 like Botswana pula and South Africa. The most significant objective of the Indian government is to solve the problems in the Indian economy. It is required for solving economical-related shortcomings like tax evasion and fraud currency [33]. It also aims to eradicate illegal activities like money laundering and offering money to terrorist groups. The Indian government also focuses on promoting a cashless economy for future growth [41]. However, the sudden Demonetization leads to serious problems among general people, which is not a new experience to India, in which India was demonetized in 1946 and 1978. Though, the exchange of huge denomination banknotes was also more restricted and taken by the bank itself. While considering the recent Demonetization in India, Reserve Bank of India (RBI) records confirm that “Indian rupee banknotes worth 16,664 billion are being circulated among the public.” This demonetization opinion should be analyzed using the sentiment analysis [47] method through text-based analysis from the reviews of the general public [42].

Sentiment analysis [3] is the procedure of “analyzing the meaning of the text and the sentiment hidden in the text by using natural language processing technology” for disassembling, modeling, extracting, reasoning, and classifying the text [18]. The major aim of sentiment analysis [4] is to discover the attitude of bipolar prospects on specific targets in a sentence [41]. In recent years, diverse approaches have been proposed for sentiment analysis by considering the complex text polarity [44]. Several researchers focus on knowing about the sentiments of the texts in various constraints [32]. Sentiment analysis is defined as computational learning for assessing the opinions about individuals, appraisals, events, people’s issues and attitudes, topics, entities, and status of mind along with their attributes [42]. It also focuses on the automatic uncovering of the considered attitudes [10]. Understanding the user’s perspective is significant for different applications like marketing analysis, product feedback, political campaigns, product reviews, and public relations [36]. The sentiment analysis has also been applied in complex problems like security threats such as monitoring conversations regarding terrorism [45]. Thus, sentiment analysis is helpful for different applications like demonetization [40].

To present a host-based intrusion detection system using a C4.5-based detector on top of the popular Consolidated Tree Construction (CTC) algorithm, it works efficiently in the presence of class-imbalanced data. This research work aims to recognize a robust classifier, which is suitable for consideration as the base learner while designing host-based or network-based intrusion detection [23] systems. This study aimed to show how methodologies

borrowed from different areas including computer science, econometrics, statistics, data mining, and sociology may be used to analyze Facebook data to investigate the patients' perspectives on a given medical prescription. It analyzes the increasing instances of cyber racism during the COVID-19 [7] pandemic, by assessing emotions and sentiments associated with tweets on Twitter. An enhanced Information-Centric Networking-Internet of Things(ICN-IoT) [12] content caching strategy by enabling Artificial Intelligence (AI)-based collaborative filtering within the edge cloud to support heterogeneous IoT architecture. A new version of the standard Optimized Link State Routing (OLSR) protocol for SGs to improve the management of control intervals that enhance the efficiency of the standard OLSR protocol [27] without affecting its reliability. Machine learning approaches are the most significant for sentiment analysis, which are logistic regression, decision tree, naive Bayes, Support Vector Machine (SVM), etc., have attained better results [19, 29]. However, different hand-crafted features need manual design and adjustment that are cost-intensive and time-consuming. The deep learning approaches are most efficient while using large-scale corpus, which has become a major research field in sentiment analysis [8, 46]. However, sentence-level sentiment analysis is complex due to the different challenges. On the other hand, neural network-based approaches are not adequately encoded and understand the semantic information of the sentences, which may result in poor classification accuracy and results [31].

The major contribution of this developed SenDemonNet model is given here.

- The main objective of this paper is to attempt to understand the public opinion on the recently implemented demonetization policy using the proposed SenDemonNet. This dataset is gathered from the “Twitter” package in R along with Twitter API. To present a new SenDemonNet model for analyzing the tweets on demonetization policy with intelligent approaches like weighted feature selection and classification using the new meta-heuristic-based algorithm to get the opinions like positive and negative.
- To propose a novel weighted feature selection method for choosing the most significant features using the F-WOA algorithm. This weighted feature selection improves the classification efficiency to get suitable opinions.
- To implement a new classification model named HDNN using the F-WOA algorithm for attaining the final sentiments like positive and negative opinions to maximize the multi-objective function concerning accuracy and precision.
- The novelty of this work is to suggest a new algorithm termed F-WOA to optimize weight function to get the weighted optimal features and the optimization of hidden neurons of DNN for improving the classification rate. This proposed algorithm aims to improve the convergence behavior of the designed SenDemonNet model.
- To validate the efficiency of the suggested SenDemonNet model with the different standard performance measures along with k-fold validation by comparing with multiple optimization algorithms and machine learning-based algorithms.

The remaining sections are discussed as follows. The existing sentiment analysis models are reviewed in Section 2. The enhanced model of proposed demonetization-based sentiment analysis using deep learning is explained in Section 3. The optimal weighted feature selection for accurate sentiment analysis for the demonetization dataset is reviewed in Section 4. The HDNN for developing the SenDemonNet is given in Section 5. The designed SenDemonNet is analyzed in Section 6. The conclusion of the proposed SenDemonNet model is given in Section 7.

2 Literature survey

2.1 Related works

In [34] have investigated the policy of government based on the general public perspective through sentiment analysis theory by considering the Twitter data. This model has investigated the state-wide analysis by geolocation for future elucidating the causes of unhappiness between the people of each state. The major part of the proposed model was to investigate the consequence of the demonetization policy carried out with the Indian government. This model has attained the result as the huge range of Indian people was happy regarding this policy. At the initial stage, most of the people were unsatisfied and shown negative result due to the complexities of getting new banknotes. As the day's passes, the people were satisfied with this policy once the notes were properly regulated. This model has considered the 30 states for analysis, in which the 21 states have accepted this policy and the remaining were not satisfied. Diverse social-economic constraints like an agriculture-based economy and a huge percentage of the rural population have comprised the difficulties faced in collecting new banknotes.

In [6] have intended to get the general public perspective regarding the recently suggested demonetization policy in India. It was analyzed by using the sentiment analysis concept, which was done based on the Twitter dataset through different classification methods like Naïve Bayes, Support Vector Machine(SVM), and decision tree. This model has collected the Twitter data from the November 9th to December 3rd with the complete set of 5000 tweets, which were given to the preprocessing stage for removing the unwanted tweets using the approaches like stemming, twitter specific terms, emoticons, and removing punctuations. Here, each tweet was categorized into three classes like “positive, negative and neutral”. Different machine learning classifiers were employed for finding the best sentiment analysis based on “positive, negative and neutral.” The experimental analysis has shown a positive review and also increases the accuracy of prediction while using the SVM.

In [43] have proposed a new sentiment analysis system termed “Sentiment Lexicon on Chinese Based and Deep Learning (SLCABG)” through the integration of Convolution Neural Network (CNN) and “attention-based Bidirectional Gated Recurrent Unit (BiGRU).” The developed model has combined the features of both deep learning and sentiment lexicon approach and has solved the challenges of conventional sentiment analysis models through the product reviews. Initially, the sentiment features present in the reviews were enhanced using the sentiment lexicon. Furthermore, the BiGRU and CNN were employed for extracting the major features like context features and sentiment features in the reviews and employ the attention strategy for evaluating the process with adding weight. Finally, the weighted sentiment features were classified. The numerical results have shown that the developed model can successfully improve the efficiency of text sentiment analysis.

In [26] have implemented a new method using a new feature ensemble model for the classification of fuzzy sentiments in the Twitter commands. It has included the elements like sentiment polarity, position, semantic, word-type, and lexical of words. The real time dataset was collected and experimented with here, and the performance analysis was carried out in terms of F1-score. The feature ensemble model was constructed for translating each tweet into a tweet embeddings through the extraction of features concerning tweets with a fuzzy sentiment like a) embeddings of words by the GloVe model; b) the distance among the words; c) sentiment score of words such as fuzzy semantic words, primary sentiment words, and

negation words; d) N-grams of words; and e) Part-of-Speech (POS) tags. The developed model has created the tweet embeddings, and CNN was applied for improving the performance of sentiment analysis. The suggested model was offered by considering the integration of the deep learning algorithm, feature ensemble model, and the divide-and-conquer approach. The suggested approach was focused on enhancing the efficiency of the sentiment analysis method.

In [47] have implemented a new model termed ADeCNN from the “aspect-level sentiment analysis based on deformable CNN”. The developed model has also suggested the Bi-directional Long Short-Term Memory Network (Bi-LSTM) with deformable CNN along with sentence-level attention for extracting the sentiment features for solving the challenges of existing models. Then, a new “Gated end-to-end memory network (GMemN2N)” was used for integrating the target into the sentiment feature extraction procedure for obtaining the sentiment features. Additionally, the correlation among the target and the words in the sentence was enhanced using the Aspect-level Sentiment Analysis (ADeCNN) model, which was experimented on “SemEval 2014 Task4 and SemEval 2017 Task4 datasets”. The experimental results have demonstrated the performance and functional efficiency of the suggested model in terms of prediction accuracy.

In [3] had provided an approach named Contextual Analysis (CA) to build a correlation among the sources and words, which was done in a tree structure known as “Hierarchical Knowledge Tree (HKT).” Furthermore, Tree Differences Index (TDI) and Tree Similarity Index (TSI) were generated from the tree structure to get the similarities and differentiation among the actual and trained datasets. The proposed model has used regression analysis to show the performance and correlation among the accuracies of supervised machine learning (SML) and TSI. Therefore, the developed model has reduced the estimation error and also improved the performance of the SML approach. The experimental results have captured and understood the variations among the negative and positive words utilized in the sentiment analysis process by considering the influential nodes in the tree construction.

In [4] had gathered 2000 k Tweeter commands from Goods and Services Tax (GST) from June 2017 to December 2017. This model has collected a topic-sentiment system using LSTM by considering the tweets regarding GST. The sentiment words were identified using different existing lexicons and allotted polarity rating for each tweet. However, a polarity-popularity structure was developed for extracting the right words with implicit GST, in which the popular words were ranked based on sentiments. Then, an Long short-term memory(LSTM) model was trained by considering the rated words to predict the sentiment on GST tweets, which has attained an improved accuracy rate. The developed model has attained superior performance during training.

In [45] have suggested a new model for solving the low classification accuracy and insufficient semantic understanding problems in the text sentiment classification approaches. The proposed model was named as “a sentiment classification model based on capsule network (SC-BiCapsNet)” for improving the capability of text information. The semantic coding structure was optimized using the COL-Att model for enhancing the consequences of text semantic expression in the text semantic framework. Further, the classification phase was increased to get more accuracy rate, which was evaluated on the “Internet Movie Database (IMDB)and Natural Language Processing and Chinese Computing (NLPCC) 2014 dataset”. Finally, the performance of the developed model was improved in terms of different metrics like F1-score. However, this model has suffered from a lack of generalization ability and takes more time.

In [1] have performed data cleaning by removing the stop words, followed by classifying the tweets as positive and negative by polarity of the words. It has generated the word cloud. Finally, it has generated positive and negative word clouds, comparison of positive and negative scores to get the current public pulse and opinion.

In [13] have introduced the concepts, backgrounds, and pros and cons of edge computing. It has to explain how it operated and its structure hierarchically with artificial intelligence concepts. They intended to clarify various analyses and opinions regarding edge computing and artificial intelligence.

In [34] have analyzed the government policy from the common person's perspective by using the concept of sentiment analysis and taking Twitter as a tool. In addition to performing a nationwide analysis, and also performed state-wide analysis using geolocation to further elucidate the reasons of displeasure among people of respective states.

In [17] have implemented a user-friendly, stand-alone access control system based on human face recognition at a distance. The local binary pattern (LBP)-AdaBoost framework was employed for face and eyes detection, which was fast and invariant to illumination changes. For fast face recognition with high accuracy, the Gabor-LBP histogram framework was modified by substituting the Gabor wavelet with Gaussian derivative filters.

In [5] have adopted the optimized deep learning concept for performing the Aspect-based Sentiment Analysis (ABSA) for demonetization tweets. The weight of the polarity scores was optimized using hybridization of two meta-heuristic algorithms like FireFly Algorithm (FF), and Multi-Verse Optimization (MVO), and the new algorithm was termed as Fire Fly-oriented Multi-Verse Optimizer (FF-MVO). Further, combined features were subjected to a deep learning algorithm called Recurrent Neural Network (RNN).

In [25] have introduced Boundary Equilibrium Generative Adversarial Network with Constrained Space (BEGAN-CS), which was improved in terms of the loss function. The discriminator structure of BEGAN-CS was AutoEncoder (AE), which cannot create a particularly useful or structured latent space. AE was considered to be related to the occurrence of mode collapse.

In [14] have reviewed sentiment analysis. Sentiment analysis was an upcoming field of the text mining area. Sentiment analysis was the process of tracing opinions views or suggestions of a particular Twitter dataset. Many algorithms have been used to find the opinion in sentiment analysis. Retrieving documents by the subject was a goal of information retrieval. There were some aspects of textual content, which from equally valid selection criteria.

In [6] have to understand the public opinion on the recently implemented demonetization policy in India. Sentiment analysis was carried out on Twitter data set using machine learning approaches. The data set was pre-processed for cleaning the data and making it possible for analysis. A final set of 5000 tweets were analyzed using machine learning techniques like SVM, Naïve Bayes classifier, and Decision tree, and the results were compared.

2.2 Problem statement

Enhancing the accuracy of text sentiment analysis assists the government regarding different decision-making and guides public opinion. Though, short text includes different characteristics like poor sentence integrity, irregular grammatical framework, and high colloquialism, which creates more complexities in sentiment analysis. Numerous sentiment analysis approaches are proposed in the literature, the contribution of sentiment analysis using demonetization dataset seems to be less. Few of the recently published machine learning and deep

learning-based sentiment analysis [28] models are reviewed in Table 1. The challenges for existing algorithms are given by, the computational complexity is very low, and it attains a low convergence rate. It does not use for real-time applications. It suffers from class imbalance, exploding gradient, and over-fitting issues, and also it takes more time for processing. For these disadvantages, the new algorithm termed F-WOA is used to improve the convergence rate. To implement a new classification model named HDNN using the F-WOA algorithm for maximizing accuracy and precision. These challenges motivate the researchers to focus on developing a new sentiment analysis by considering the demonetization tweets.

3 Enhanced model of proposed demonetization sentiment analysis using deep learning

3.1 Proposed model for SenDemonNet

The sentiment analysis on different social networks like Facebook or Twitter has become the major learning tool for learning users' opinions, which has broadened the range of real time applications. This research work aims to extract emotion recognition from text. Moreover, the machine learning methods are not as efficient compared with the deep learning approaches due to the lack of training performance and take more training time. Though, the accuracy and efficiency

Table 1 Features and challenges of existing sentiment analysis approaches

Author [citation]	Methodology	Features	Challenges
[34]	scoring method	<ul style="list-style-type: none"> • It helps to find the literacy rate in each state. • It also evaluated the happiness of the states using a scoring approach. 	<ul style="list-style-type: none"> • Though, it attains the negative sentiments for the policy of Demonetization.
[6]	Ensemble learning	<ul style="list-style-type: none"> • It attains a high accuracy rate. 	<ul style="list-style-type: none"> • The less accurate rate is attained for Tamil and Hindi tweets. • It also has more neutral counts.
[43]	CNN and BiGRU	<ul style="list-style-type: none"> • It increases the performance of text sentiment analysis. • It also improves the service quality of users and classification accuracy. 	<ul style="list-style-type: none"> • This model is not suitable for sentiment refinement application with more constraints.
[26]	CNN	<ul style="list-style-type: none"> • It increases the performance of sentiment analysis. 	<ul style="list-style-type: none"> • This model is only applicable for fuzzy sentiments.
[47]	ADeCNN	<ul style="list-style-type: none"> • It improves the classification accuracy with text sentiment analysis. 	<ul style="list-style-type: none"> • Conversely, this model does not perform on ironic text sentiment analysis and also its sentiment polarities.
[3]	SML	<ul style="list-style-type: none"> • It increases the prediction performance. 	<ul style="list-style-type: none"> • This model is not suitable for real time applications.
[4]	LSTM	<ul style="list-style-type: none"> • It attains an accuracy rate of 84.51% and also increases the testing accuracy. 	<ul style="list-style-type: none"> • This model does not consider the bi-lingual font-mixed tweets.
[45]	BiCapsNet	<ul style="list-style-type: none"> • It increases the performance and classification accuracy. • It also improves the expression capability of lexical. 	<ul style="list-style-type: none"> • It suffers from the lack of generalization ability and time consumption.

of sentiment analysis are affected through the limitations present in the natural language processing, where the lack of suitable labeled data negatively influences the performance. Therefore, deep learning is adopted for sentiment analysis owing to the automatic learning ability. Moreover, the existing research works face some technical and theoretical problems on sentiment analysis that results in attaining less accuracy. The better performance is achieved only by using the shorter text datasets and more readability in terms of sentiment classification. Hence, the reliability of the model is depended on the domain or size of the data. Consequently, the recent research works focus on developing a deep learning-based sentiment analysis model that consists of different benefits and efficiency. These algorithms solve the sentiment analysis problems like aspect-based sentiment classification and sentiment polarity due to the automatic learning and extraction of features, which helps in improving the performance and accuracy. However, the major challenging issue in sentiment analysis is to summarize and detect the overall sentiments.

Similarly, the sentiment analysis has the major complication as the gathered tweets are from the local language. Likewise, the sentiment analysis using Twitter text face diverse limitations like complications in designing the application-specific techniques and algorithms for analyzing the human language linguistics precisely, usage of acronyms, emoji's, abbreviations, URL and hashtags on Twitter, restricted sign about sentiment, shorter messages, and tweets in an unceremonious language. Manual analysis is almost impracticable for sentiment analysis. These challenges are considered while implementing a new sentiment analysis model. Further, this paper attempts to understand the sentiments and public opinion on the current demonetization policy announced in India, which is carried out by gathering data from well-known social media named Twitter. The proposed sentiment analysis model on tweets regarding demonetization policy in India is represented in Fig. 1.

This paper has developed a new sentiment analysis model on Indian demonetization policy for recognizing human emotions, which has classified the sentiments like negative and positive sentences. The sentiment analysis of tweets is considered as the major task because of the informal language of tweets leads to giving more creative and new opinions. The suggested SenDemonNet model consists of different stages like “(a) pre-processing, (b) feature extraction, (c) weighted feature selection and (d) classification”. The gathered tweets on the Indian demonetization policy are preprocessed using the processes like stop word removal, blank space removal, and punctuation and tags removal, which processes the raw input data for improving the quality. The preprocessed data are further given to the feature extraction process for getting the significant and most representative features for minimizing the complexities, which is carried out using the approaches like Word2vector, Term Frequency — Inverse Document Frequency (TF-IDF), and Bag of n-grams. As the total number of extracted features is more in number, the dimensionality of the features is reduced using the PCA to increase the interpretability without information loss. Furthermore, the dimensionality reduced features are given to the weighted feature selection process, where the more suitable features is selected using the F-WOA, and it is transformed to a weighted feature using a weight function. The weight function is also optimized using the F-WOA for choosing the most suitable and required features for better sentimental analysis. Finally, the weighted features are given to the HDNN-based sentimental analysis along with the F-WOA for the optimization of hidden neurons, which gets the final sentiments as the positive and negative opinions. The major objective of the developed model on tweets regarding demonetization policy in India is the maximization of accuracy and precision as a multi-objective function. Thus, this research work discovers the way for the extraction of emotions automatically from the social media text based on sentimental analysis on demonetization policy in India.

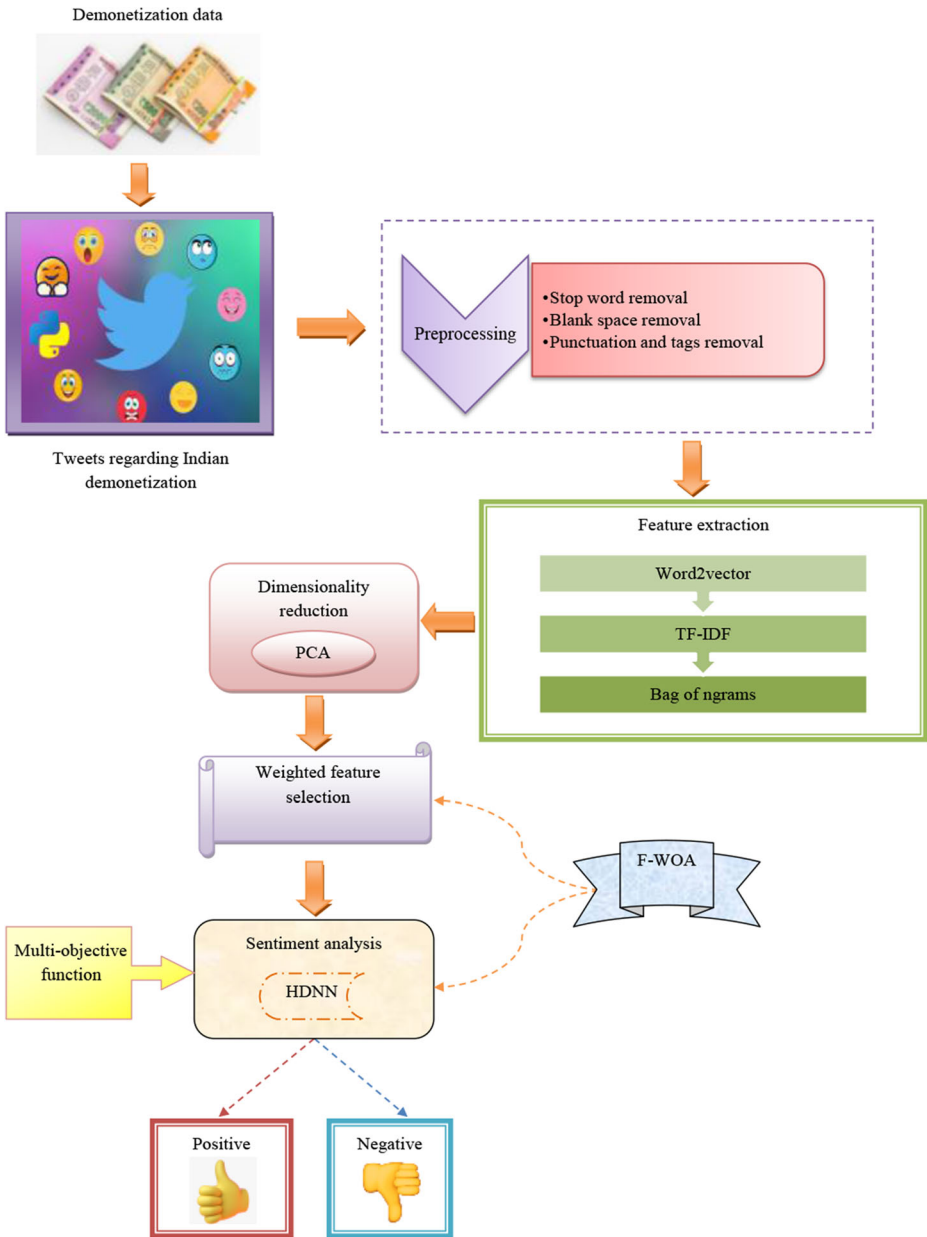


Fig. 1 Sentiment analysis model on demonetization policy in India

3.2 Description of dataset

The proposed sentiment analysis model has collected the input web page data from the “<https://www.kaggle.com/arath2/demonetization-in-india-twitter-data>: access date: 12-05-2021”. This dataset is gathered from the “Twitter” package in R along with Twitter API. The Government of India has announced the Demonetization of 1000 and 500 banknotes on

8 November 2016. The number of tweets as 6000 based on the Demonetization is considered as the data, which includes 14 columns and 6000 rows for 6000 tweets, respectively. The 14 columns are considered to be created as replyToSID, screenName, is Retweet, favorite, replyToSN, replyToUID, favorite Count, Text (Tweets), retweeted, retweet count, truncated, id, and status Source. The datasets for the proposed SenDemonNet model are described in below Tables 2 and 3.

The collected data is termed as T_r , where $r = 1, 2, \dots, R$ and the total number of tweets present in the dataset is termed as R , which is given to the further processing.

Summary: Twitter API The data set for the analysis is collected from Twitter API. This data is pre-processed and unwanted tweets are removed and a collection of 6000 tweets with demonetization. The summary collected the dataset contains the following information. Each entry contains a tweet id, the tweet text, and the sentiment label. SenDemonNet model using the demonetization dataset was implemented in MATLAB 2020a. MathWorks matlab r2020a is a multi normal form numerical computing environment and special programming language developed by MathWorks. It is mainly focused on the high-tech computing environment of visualization, interactive programming, and scientific computing and also it is the most powerful, practical, and powerful business mathematics software in the world. MATLAB libraries for Twitter can be used to collect information about the tweets like creation date, creator name, etc. Each tweet is classified into two categories such as positive, and negative. Positive is used for showing positive sentiment or positive opinion towards demonetization, Negative specifies for showing negative sentiment towards the movement.

3.3 Data pre-processing

Data preprocessing is essential for sentiment analysis, which is intended for cleaning the text data T_r . As the proposed model use tweets on demonetization policy from the Twitter dataset, the comments have lots of unknown and unnecessary words. It consists of casual language, frequent words, adverbs, conjunction, prepositions, and article, blank space in sentences, and tags and punctuation in comments. The data generated from Twitter is usually not suitable for learning or analysis directly. Hence the data should be normalized to make it in a better format before applying any techniques. In our proposed method, pre-processing techniques are used to eliminate unwanted text from the tweets and hence reduce the features. Thus the data is made suitable for all the learning algorithms. Stemming is also used for pre-processing the data. One of the significant data mining methods is called as data preprocessing, which transforms the raw data into a comprehensible format. Due to the nature of real-world data like inconsistent, incomplete, lacking in specific trends or behaviors leads to more number of errors. Thus, data

Table 2 Dataset description for the proposed SenDemonNet model

Description	Values
Number of rows	6000
Number of columns	14
Number of tweets	6000
Number of classes	2

Table 3 Training and testing samples for the proposed SenDemonNet model

Sl.no	Samples	Positive	Negative
1	Did you vote on #Demonetization on the Modi survey app?	✓	
2	Was PayTM informed about the # Demonetization edict by PM?	✓	
3	Section 144 issued by District Magistrate, #Indore prohibiting #socialmedia posts on #demonetization, Why??	✓	
4	Do you support the Opposition's agitation against #demonetization?	✓	
5	How does it impact the common man?	✓	
6	Demonetization 21K Crore Deposited In #JanDhanAccount Most From Bengal, Karnataka. The reason for protests?		✓
7	And you want to fire @narendramodi on # demonetization?		✓
8	Why is it so difficult for BJP to have coherent, cogent communication on a critical situation like #demonetization?		✓
9	So how does CM stone-pelting stop after #Demonetization If they don't care about 4 bullets or pellets why do they care so much 4 money?		✓
10	#demonitization didn't stop terrorism?		✓

preprocessing is the most required method for this sentimental analysis model to solve these problems. This suggested SenDemonNet use stop word removal, blank space removal and punctuation and tags removal for improving the quality of raw data [15].

- (a) Stop word removal: This approach removes the frequently occurred words like adverbs, conjunction, prepositions, and article. The removal of these words reduces the dimensionality of the datasets. It also includes the most common words like 'they', 'she', 'but', 'if', 'he' and 'we', etc. Hence, the stop words should be eliminated to improve the quality of sentiment analysis. The stop word removed text data is termed as T_r^{SW} , which is further given to the blank space removal process.
- (b) Blank space removal: The blank space in the sentences is removed for better sentiment analysis, which is represented as T_r^{BS} .
- (c) Punctuation and tags removal: The Twitter comments include more symbols like "." & ";" HyperText Markup Language (HTML) tags, which have to be eliminated for enhancing the efficiency of the SenDemonNet model. The preprocessed data after removing the punctuation and tags from the sentences are specified as T_r^{PT} that is further given to the feature extraction process.

4 Optimal weighted feature selection for accurate sentiment analysis for demonetization dataset

4.1 Feature extraction

The developed SenDemonNet model gathers the most significant features from the preprocessed data T_r^{PT} . It is the new emerging area in the sentiment analysis field that helps in knowing about the nature of comments like feature has positive or negative comments. This feature extraction stage is intended to minimize the computational complexity of the sentiment analysis model. This model extracts the features using three effective approaches like Bag of n-grams, TF-IDF, and Word2vector. Initially, the preprocessed data T_r^{PT} is given to the Bag of n-grams for dividing the text or comments from data.

- (a) Bag of n-grams [37]: It is used for validating the continuous words present in the given sequence of text, which is employed to analyze the sentiment of the document or text. This model consists of three types like unigram, bigram, and trigram. Here, the unigram represents the single word; bigram specifies the pair of words, whereas the trigram expresses the set of words equivalent to the count of three. The sentences extracted using a bag of n-grams are termed as T_r^{SE} .
- (b) Word2vector [9]: The extracted sentences T_r^{SE} are given to the word2vector algorithm for building the word vector representation. It determines and learns the vector representation of every word. In word2vec, two training techniques are there, which are Continuous Bag-Of-Words (CBOW) and skip-gram model. This technique discovers the contextual similarity among the words and phrases in a specific document. Generally, each word identifies the uniqueness based on the Bag of words to represent the no contextual correlation among two words.

CBOW: It predicts the context of a particular current word in the separate window, in which the input layer, projection layer, and output layers are formulated in the CBOW that takes the input as the current word. The hidden layer represents the word to be projected at the output layer by fixing the number of dimensions. CBOW model is more efficient when compared with the skin gram model.

Skip-gram model: This model has also found the context words at a specific window size based on the current word. The output layer generates the resultant as the context words. Therefore, this model gets the contextual words, which must be appeared closer to the specific input word at a similar time. This technique is used to represent rare phrases and words, which is also more suitable for small-scale data. Finally, the extracted features using the word2vec technique are attained as 3600 and is represented as fs_b^{W2V} .

- (iii) TF-IDF [2]: The extracted sentences T_r^{SE} are given to TF-IDF. It is one of the most effective approaches for retrieving the major features from the tweets. It is also called as the weight measure for determining the significance of a word for a specific document, in which the TF is used for measuring the “number of times a particular term s occurred in a document.” When the term has occurred different times in the document, the frequency is increased. Moreover, the TF is measured in Eq. (1).

$$TF(s, DC) = \frac{Ns}{TN} \quad (1)$$

Here, the total number of termssin a document DC is denoted as Ns and the number of times term s appeared in a document DC is represented as TN . IDF measures the importance of terms, which gives most significance to the rarely occurred terms in the document as derived in Eq. (2).

$$IDF(s) = \log_e \left(\frac{ND}{TD} \right) \quad (2)$$

In Eq. (2), the total number of documents with terms presents in it that is denoted as TD , and the total number of documents in the dataset T_i is indicated as ND . Finally, the weight for the term

sis formulated in Eq. (3).

$$TF-IDF(s, DC) = TF(s, DC) \times IDF(s) \quad (3)$$

Finally, the features extracted using TF-IDF are specified as f_s^{TF-IDF} and obtained as 29,020.

Finally, the extracted features are represented as $f_s^{FE} = \{f_s^{W2V}, f_s^{TF-IDF}\}$, where $b = 1, 2, \dots, B$ and B denote the total number of features attained as 32,620 by concatenating both f_s^{W2V} and f_s^{TF-IDF} . Here, the length of features is represented as $1 \times 32,620$, which is similar to the number of features.

4.2 PCA-based dimension reduction

As the total number of features extracted is more in this developed SenDemonNet model, there is a need to reduce the features by applying the Principle Component Analysis (PCA) [38]. It reduces the dimensionality of the feature sets, which assists in speeding up and simplifying the computations. It is a broadly utilized statistical approach for minimizing the dimension of the feature set, which can transform the original set of correlated variables into a smaller set of uncorrelated variables. Assume the word vector features as f_s^{FE} , where the matrix of word data vector is considered as $A \times C$. Here, the term A denotes the number of comments in the features text and the attributes of the features are considered as C . Initially, the covariance matrix Mf_s^{FE} is computed. Then, the Eigenvalues ξ_c and the Eigenvectors v_c of M are formulated, where $c = 1, 2, \dots, C$. Furthermore, the dimensionality of the data is reduced based on the variance of the Eigenvalues. The standard transformation matrix TM is computed, in which each column is formulated based on Eq. (4).

$$g_c = \frac{e_c}{\sqrt{\xi_c}}, c = 1, 2, \dots, l \quad (4)$$

The principal component variables are derived for each comment, as shown in Eq. (5).

$$f_s^{PCA} = f_s^{FE} \times g_c \quad (5)$$

In Eq. (5), the final dimensionality reduced text features are specified as f_s^{PCA} with the $a \times l$ matrix of principal component variables, which includes the variance of one and the mean of zero.

PCA is used for highlighting their differences and similarities among extracted features and thus, it is more essential step for developed SenDemonNet model. It mainly focuses on reducing the dimension and obtaining the maximum variance of data. It reduces dimensionality by eliminating the later principal components. The attained Eigenvalues are arranged from highest to lowest order for significance, where the less significant components are eradicated. From the list of Eigenvectors, the matrix is formed. Thus, the higher significant components are selected as 1×50 in the model. Finally, the final dimensionality reduced text features are specified as f_s^{PCA} , where $c = 1, 2, \dots, C$ and C denotes the total number of dimensionality reduced features, which is equivalent to the higher significant components, and thus, 50 features are extracted using PCA.

4.3 Weighted feature selection by heuristic F-WOA

The primary aim of this proposed SenDemonNet model is to select the weighted features using the F-WOA. The PCA-reduced features f_s^{PCA} are also more in number for training, and thus the feature selection is necessary with the optimization algorithm. From the 50 features f_s^{PCA} , the most suitable features are extracted using F-WOA that is termed as f_s^{WF*} and obtained as 5. These features are further multiplied with the weight function Wg_c , which is optimized using F-WOA. To map the high scale between the features, the weighted features are extracted using the F-WOA. Here, the length of the weight function is equivalent to the number of selected features. This weighted feature selection is formulated in Eq. (6).

$$f_s^{WF*(new)} = f_s^{WF*} \times Wg_c \quad (6)$$

Here, the new weighted features are termed as $f_s^{WF*(new)}$, which is attained by multiplying the weighted function Wg_c with the optimal features f_s^{WF*} , where the weight Wg_c is optimized using the F-WOA and $c = 1, 2, \dots, C^{WF*}$ indicates the total number of weighted features that is extracted as 5. From the 50 features f_s^{PCA} , the extracted features are attained as 5, which is simpler for processing that minimizes the computational complexity of the sentiment analysis. This optimal feature selection minimizes the feature sizes and enhances the classification accuracy of sentiment analysis. It reduces the non-informative features and gets the most significant features for increasing the classification efficiency. The unnecessary features are removed using the optimization of weighted function Wg_c through F-WOA. The designed weighted feature selection is depicted in Fig. 2.

5 Heuristic deep neural network for developing SenDemonNet

5.1 Developed F-WOA

A new F-WOA is developed based on features of both WOA [20] and FOA [11] techniques. It is developed for performing the weighted feature selection to get the suitable features for reducing the complexities. F-WOA algorithm is also used for the optimization of hidden neurons in the DNN. This fine-tuning offers better performance to get reasonable opinions on tweets with a demonetization policy regarding positive and negative comments with efficient sentiment analysis. The hybrid optimization F-WOA algorithm is inspired by the integration of two or more features to increase the convergence behavior of the designed sentiment analysis model. In the proposed F-WOA, the positions are updated based on the random parameter, where if $k < 0.5$ then the positions are updated based on the encircling the prey, and global search mechanisms of WOA or else the solutions are updated using FOA based on the local seeding and global seeding.

WOA has different benefits like has better exploration and exploitation abilities in solving different real-time constrained and unconstrained approaches. Though, this method suffers from the local optima problems. Therefore, a new algorithm called FOA is adopted here, which provides superior performance for solving the real non-linear optimization problems. Hence, the proposed F-WOA algorithm poses superior search abilities with better convergence behavior.

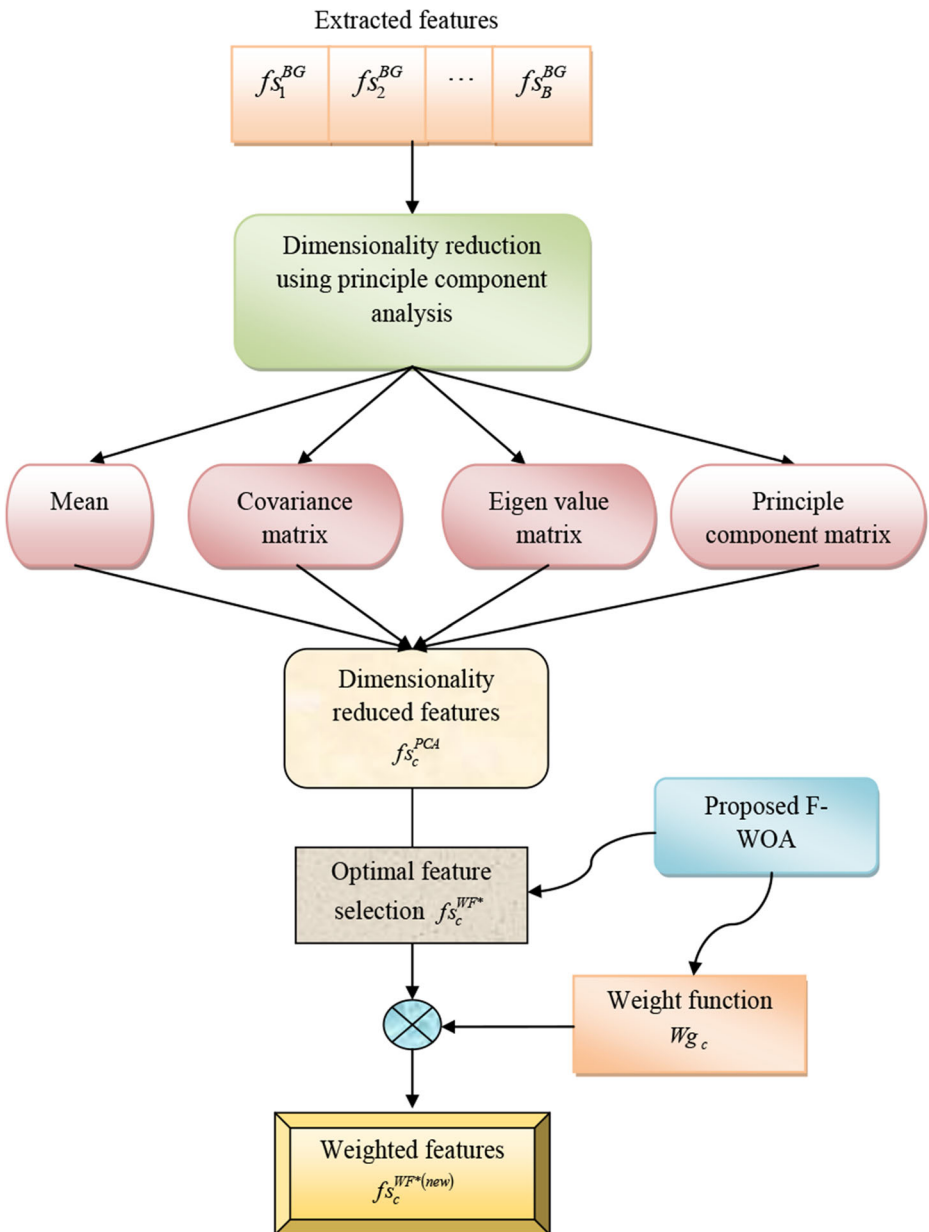


Fig. 2 Designed weighted feature selection using F-WOA

WOA is inspired by the hunting nature of humpback whales owing to their ability to recognize the prey and encircling nature. It also follows the bubble-net hunting scheme that offers a competitive performance than other existing approaches. WOA consists of diverse stages like “encircling prey, spiral bubble-net feeding maneuver,

and search for prey,” where the spiral updating is replaced with the FOA technique. The encircling phase is formulated in Eq. (7).

$$\vec{Y} = \left| \vec{M} \cdot \vec{U}^*(j) - \vec{U}(j) \right| \quad (7)$$

$$\vec{U}(j+1) = \vec{U}^*(j) - \vec{N} \cdot \vec{E} \quad (8)$$

In the abovementioned equations, the best position vectors among the attained vectors is represented as \vec{U}^* , and the element-by-element product is referred to as “.”, the absolute function is noted as $||$ and the coefficient vectors are denoted as \vec{N} and \vec{M} that is formulated here.

$$\vec{M} = 2 \cdot \vec{r}nd \quad (9)$$

$$\vec{N} = 2\vec{n} \cdot \vec{r}nd - \vec{n} \quad (10)$$

Here, the random vector is specified as $\vec{r}nd$ that is ranging among $[0, 1]$. The bubble-net technique of WOA has two mechanisms like shrinking encircling and spiral updating mechanisms. The shrinking encircling method works by reducing the value of \vec{n} in Eq. (10).

Spiral bubble-net feeding maneuver based on FOA: It is motivated by the nature of trees in the forest. Few trees in the forests survive for multiple years, whereas some trees only live for a specific time. Hence, the FOA is formulated based on the seeding procedure of trees, which can be done by animals that feed on the fruits or seeds, seed distribution based on natural procedures, and some seeds fall under the trees. It is more efficient for solving non-linear optimization problems.

FOA consists of three different phases like “local seeding of the trees, population limiting and global seeding of the trees”. Initially, tree population is formulated, where a tree consists of variables that denote the age of the specified tree. At first, the value of ‘0’ is assigned as the age of a tree. Once the initialization of trees is done, the local seeding operator generates the new young tree, and these trees are added to the forest. Further, the age of older trees is increased by ‘1’ except for newly produced trees. Further, a control on the population of the trees in the forest is limited by some control, where some of the trees are avoided from the forest, which will create a new candidate population for generating the global seeding phase. Here, in the forest, some percentage of the trees in the population is selected for moving them. This stage has added few possible solutions to the forest for solving the local optimal solutions. Then, the fitness value is used for ranking the trees, where the largest fitness value is considered as the best tree and its age is assigned to 0 for avoiding the aging. Further, the best tree in the forest is removed due to the age increment of all the trees in the local seeding stage. These phases are continued until the termination condition is met. These stages are performed with the following parameters like “Local Seeding Changes” or “LSC”, “transfer rate” and “Global Seeding Changes” or “GSC”.

The search for prey or exploration phase is formulated here.

$$\vec{Y} = \left| \vec{M} \cdot \vec{U}_{rand} - \vec{U} \right| \quad (11)$$

$$\vec{U}(j+1) = \vec{U}_{rand} - \vec{N} \cdot \vec{E} \quad (12)$$

Here, a random position vector is termed as \vec{U}_{rand} , \vec{N} is used for searching the prey, the WOA works by using \vec{N} and random parameter k . The proposed F-WOA algorithm, initially 50 extracted features using PCA has been given as an input. The extracted features are termed as f_s^{PCA} . From the 50 features, the fitted features are extracted by the F-WOA algorithm. Then, the weight is given as an input to the F-WOA algorithm. The unwanted features are removed using the optimization of the weighted function Wg_c . Here, the new weighted features are termed as $f_s^{WF^*(new)}$. The extracted features are attained as 5, which is simpler for processing that minimizes the computational complexity of the sentiment analysis. The output of the F-WOA algorithm is a weighted feature. The DNN model is used for feature selection. For the DNN model, the weighted features are given as input. The total number of nodes in the hidden layer hn is formulated using the F-WOA algorithm. Finally, the output has been got as a number of hidden neurons getting optimal results. Here, the length of the weight function is equivalent to the number of selected features. Hence, the pseudo-code of the designed F-WOA is given in Algorithm 1.

Algorithm 1: Proposed F-WOA
Initialize population \vec{U}_i where $(i = 1, 2, \dots, ik)$
Determination of the fitness for every search agent
Compute the best search agent as U^*
Assign random parameters
While $(j < j_{max})$
for each search agent
if $(k < 0.5)$
Solution update based on WOA
if $(N < 1)$
Update the position of search agent based on encircling of prey using Eq. (7)
else if $(N \geq 1)$
Choose a random search agent U_{rand}
Update the position of recent search agent based on exploration phase by Eq. (12)
end if
else if $(k \geq 0.5)$
Solution update based on FOA
Update the solutions using local seeding phase of FOA
Update the solutions using global seeding phase of FOA
end if
end for
end for
Compute fitness of each search agent
Update the best search agent U^*
$j = j + 1$
end while
return U^*

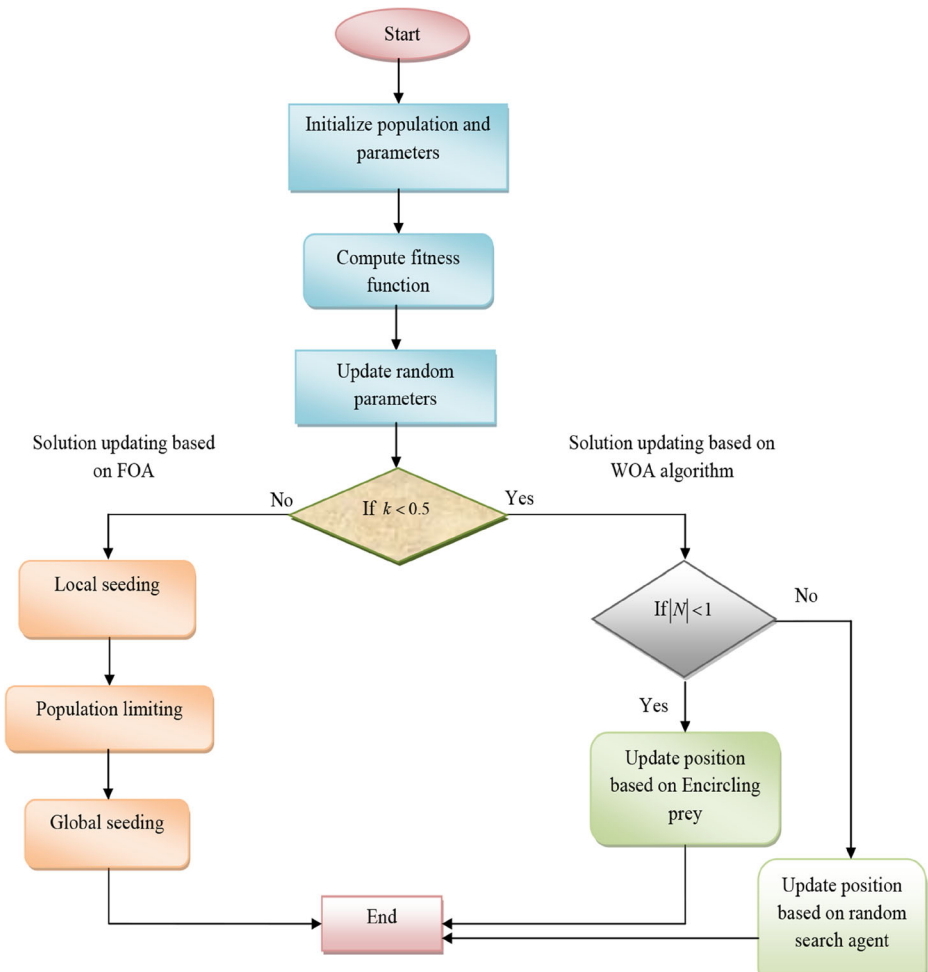


Fig. 3 Flowchart of the designed F-WOA

The flowchart of the developed algorithm is given in Fig. 3.

5.2 HDNN-based sentimental analysis

The ablation studies of the proposed DNN are described given below, 6000 tweets were taken from the Twitter dataset. The DNN model is used for feature selection. For the DNN model, the weighted features are given as input. The hidden layers of DNN are tuned using the F-WOA algorithm to attain the maximum accuracy rate. The HDNN approach is used for the proposed model to improve the performance with efficient sentiment classification. In the comparison between the NN and machine learning algorithms, SVM is not suitable as the computational complexity is very high because the required training time is higher. For KNN, the prediction time is high. Required high memory for storing the training data. DNN has been improved to HDNN based on the abovementioned disadvantages. The main advantage of HDNN is given by, features that are automatically deduced and optimally tuned for the desired

outcome. The same neural network-based approach can be applied to many different applications and data types. The deep learning architecture is flexible to be adapted to new problems in the future. The proposed SenDemonNet model uses the for the efficient classification of opinions in terms of positive and negative comments. Although the DNN has the ability of automatic feature extraction, it suffers from extracting the features due to the complexities in data like inadequate learning rate, an inadequate number of hidden neurons, unsuitable structure, large scale dataset, noisy data, and bad input selection. Therefore, there is a need to select the suitable features from the input data, which offers good performance for processing the data, removes the noisy features, and saves time. Although it has several features, DNN affects the generalization performance, complications in training, and needs more data for processing, and thus it is not suitable for small-scale datasets. Therefore, this paper intends to optimize the hidden neurons of DNN using the F-WOA, which results in improving the performance improvement with efficient sentiment classification.

The DNN [29] includes three layers such as input layer, hidden layer, and output layer. It takes the input as the optimal selected weighted features $f_s^{WF*(new)}$. The number of hidden neurons plays a major role in the classification in terms of accuracy and training speed, which is present in the hidden layers of DNN. Thus, the optimization is performed using F-WOA. The total number of nodes in the hidden layer hn formulated by Eq. (13).

$$hn = \sqrt{in} + z + p \quad (13)$$

In Eq. (13), the number of nodes in the input layers is termed as in , the number of nodes in the output layers is considered as z and the constant value is termed as p in the range of [1, 10]. An activation function is derived in the hidden layer of DNN for improving the capability of the non-linear fitness values. Here, the activation function is considered based on the sigmoid function, as shown in Eq. (14).

$$\sigma = \frac{1}{1 + e^{-f_s^{WF*(new)}}} \quad (14)$$

Here, the network input data is mentioned as $f_s^{WF*(new)}$, which is activated using the mapping functions ψ as formulated in Eq. (15).

$$\psi = \mathbf{sigm} \left(w_{is} f_s^{WF*(new)} + \zeta_{is} \right) \quad (15)$$

In Eq. (15), the bias and the weight matrix are termed as ζ_{is} and w_{is} among the output and hidden layer. Finally, the proposed SenDemonNet model gets the classified outcomes as positive and negative opinions with the aid of the F-WOA technique, which is demonstrated in Fig. 4.

5.3 Objective model for SenDemonNet

The reason behind deploying a new SenDemonNet method for tweets is described given below, sentiment analysis is research and developing a stream of Natural Language Processing (NLA) methods. Classifying the sentiment of Twitter messages is most similar to sentence-level sentiment analysis for the limited-sized tweets and paragraph-level sentiment analysis for more than one sentence. Twitter follows a micro-blogging nature, small tweets, and also supports different natural languages, so Twitter is the best source for sentiment

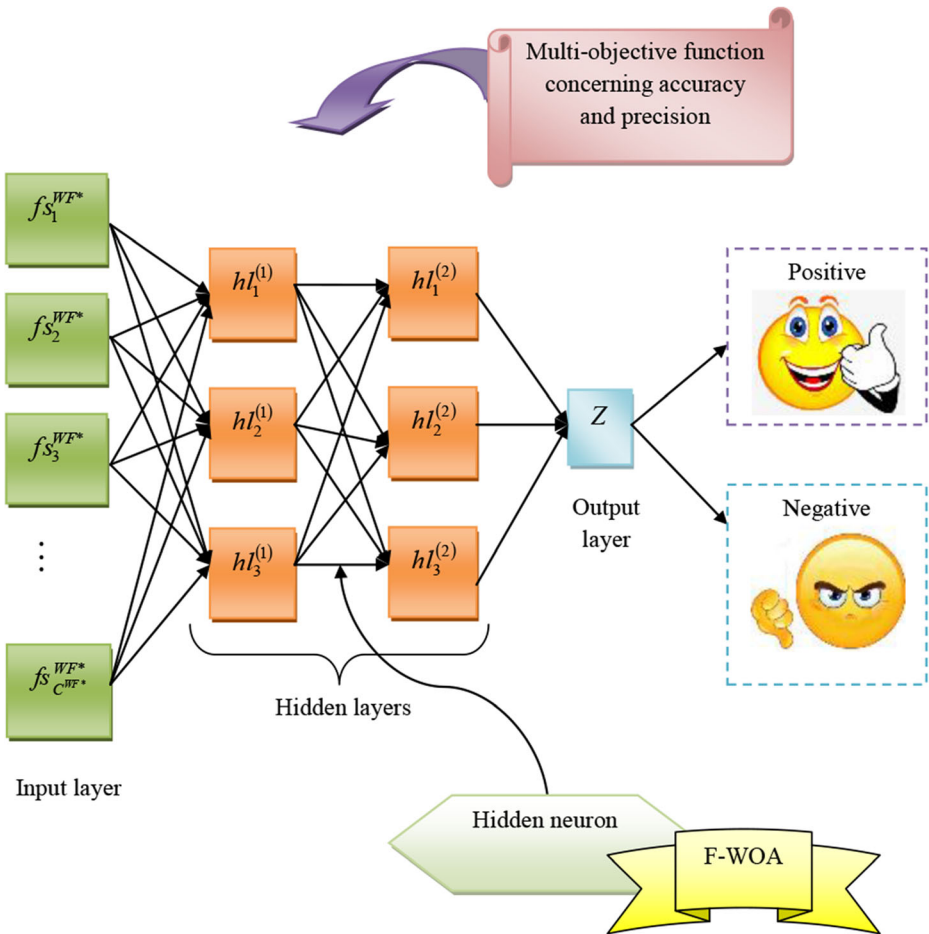


Fig. 4 HDNN-based Sentimental Analysis using F-WOA

analysis. The designed SenDemonNet model intends to improve the performance of sentimental analysis based on the weighted feature selection and improved classification using HDNN, which is performed using the F-WOA. This optimization increases the convergence rate along with better efficiency. The optimization performed in the SenDemonNet model is diagrammatically represented in Fig. 5.

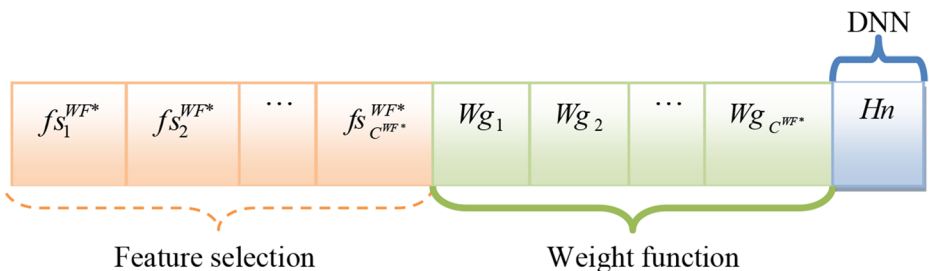


Fig. 5 Solution encoding of the proposed sentimental Analysis using F-WOA

The designed SenDemonNet model improves the efficiency by optimal feature selection f_s^{WF*} , weighted feature selection $f_s^{WF*(new)}$ using weight function Wg_c , and hidden neuron optimization Hn through F-WOA. The weighted features f_s^{WF*} are attained as 1 to the total number of PCA features. The range of weight function is considered among 0 to 1 and the hidden neurons of DNN Hn lie in the range of 1 to 50.

The proposed SenDemonNet model considers the multi-objective function concerning accuracy and precision using the F-WOA with the hidden neuron optimization of DNN for efficient classification on tweets of demonetization policy. The multi-objective function Obj for the designed HDNN using F-WOA is formulated in Eq. (16).

$$Obj = \underset{\{Hn, f_s^{WF*(new)}, Wg_c\}}{\operatorname{argmin}} \left(\frac{1}{Ac + Pr} \right), \quad (16)$$

Here, precision Pr is “the ratio of positive observations that are predicted exactly to the total number of observations that are positively predicted,” as formulated in Eq. (17).

$$Pr = \frac{Tu^{ps}}{Tu^{ps} + Fe^{ps}} \quad (17)$$

Accuracy Ac is a “ratio of the observation of exactly predicted to the whole observations” as derived in Eq. (18).

$$Ac = \frac{(Tu^{ps} + Tu^{ng})}{(Tu^{ps} + Tu^{ng} + Fe^{ps} + Fe^{ng})} \quad (18)$$

Here, $Tu^{ps}, Tu^{ng}, Fe^{ps}, Fe^{ng}$ refer to the “true positives, true negatives, false positives, and false negatives,” respectively. Therefore, the designed SenDemonNet model enhances the performance of sentiment analysis in terms of precision and accuracy with more accurate classified outcomes.

5.4 Methods used for comparison

As a new hybrid optimization algorithm has been suggested in this model, the well-performing optimization algorithms and classifiers have been used only for experimentation to show the effectiveness of the designed F-WOA technique with DNN for SenDemonNet model. The experimental analysis is performed with different algorithms like PSO [35], GWO [30], WOA [20], and FOA [11]. The performance of the proposed model is also compared over the conventional models such as SVM [21, 24], CNN [16], LSTM [22], and DNN [29] in terms of “Type I and Type II measures.

PSO: It is more eminent due to its simple concept, easier implementation, robustness to control parameters, and computational efficiency while comparing with other heuristic optimization and mathematical algorithms. Moreover, it requires a small number of parameters and a correspondingly lower number of iterations.

GWO: It is chosen here because of its easier implementation owing to its less computational requirements, less storage, simple structure, faster convergence regarding the ability to avoid local minima, continuous reduction of search space, and fewer decision variables. It provides very competitive results compared to these well-known meta-heuristics.

SVM: It can scale the high dimensional data, usage of the kernel function, performs well in semi-structured and unstructured data, and gets better results. It can be applied to both classification and regression tasks. It also discovers the optimal boundary among the probable outputs.

CNN: It can automatically detect significant features without human intervention. It is also computationally efficient, which solves image classification and computer vision problems. It gets higher accuracy than SVM.

LSTM: LSTM has better abilities due to the utilization of input and output biases, learning rates, and huge range of parameters. It does not need any finer adjustments. It gives less computational complexity. It has the capability of learning long-term dependencies.

Due to these features, all these techniques have been selected for comparing our F-WOA-DNN model to validate the superiority.

6 Results and discussions

6.1 Experimental setup

The proposed SenDemonNet model using the demonetization dataset was implemented in MATLAB 2020a. Here, Type I measures are positive measures like accuracy, sensitivity, specificity, precision, Negative Predictive Value (NPV), F1Score and Mathews correlation coefficient (MCC), and Type II measures are negative measures like False positive rate (FPR), False negative rate (FNR), and False Discovery Rate (FDR)". The experimentation was carried out by considering the number of population as 10 and the maximum number of iterations as 25. PSO algorithm the acceleration constants are given by $c_1 = 1$, $c_2 = 2$ and the maximum weight and minimum length are considered as 0.90,0.1. FoA, the transfer rate, and area limit are considered as 15, 20. In WOA, the acceleration constants are given by, $b = 1$. For our proposed model the hidden neuron count is considered as 12. SVM classifier, the k-fold, and verbose values are given by 10, 0. The fit posterior is taken false rate and the learners are SVM. The kernel function is linear. In CNN, the convolution 2d layer size is [3,16]. The optimizer and verbose have been taken as sgd and false. LSTM, the number of hidden units, and the max epochs are given by, 2 and 6. The mini-batch size and the initial learning rate are given by, 20 and 0.1. The gradient threshold and the verbose values are taken as 1 and 0. And the DNN classifier, the hidden neuron count is given by 10.

6.2 Performance measures

Different performance metrics are considered for evaluating the performance of the proposed SenDemonNet model and it's described below.

- (a) F1 score: "harmonic mean between precision and recall. It is used as a statistical measure to rate performance".

$$F1score = \frac{2Tu^{ps}}{2Tu^{ps} + Fe^{ps} + Fe^{ng}} \quad (19)$$

(b) MCC: “correlation coefficient computed by four values”.

$$MCC = \frac{Tu^{ps} \times Tu^{ng} - Fe^{ps} \times Fe^{ng}}{\sqrt{(Tu^{ps} + Fe^{ps})(Tu^{ps} + Fe^{ng})(Tu^{ng} + Fe^{ps})(Tu^{ng} + Fe^{ng})}} \quad (20)$$

(iii) NPV: “probability that subjects with a negative screening test truly don’t have the disease”.

$$NPV = \frac{Tu^{ng}}{Fe^{ng} + Tu^{ng}} \quad (21)$$

(iv) FDR: “the number of false positives in all of the rejected hypotheses”.

$$FDR = \frac{Fe^{ps}}{Fe^{ps} + Tu^{ps}} \quad (22)$$

(e) FPR: “the ratio of count of false positive predictions to the entire count of negative predictions”.

$$FPR = \frac{Fe^{ps}}{Fe^{ps} + Tu^{ng}} \quad (23)$$

(f) FNR: “the proportion of positives which yield negative test outcomes with the test”.

$$FNR = \frac{Fe^{ng}}{Tu^{ng} + Tu^{ps}} \quad (24)$$

(g) Sensitivity: “the number of true positives, which are recognized exactly”.

$$Se = \frac{Tu^{ps}}{Tu^{ps} + Fe^{ng}} \quad (25)$$

(h) Specificity: “the number of true negatives, which are determined precisely”.

$$Sp = \frac{Tu^{ng}}{Tu^{ng} + Fe^{ps}} \quad (26)$$

6.3 Performance analysis on different heuristic-based algorithms

The proposed SenDemonNet model using tweets on Indian Demonetization is evaluated in terms of different algorithms with diverse heuristic-based techniques as given in Fig. 6 by varying the learning percentages. The accuracy of the designed HDNN with F-WOA is 2.3%, 2.3%, 3% and 2.2% superior to PSO-DNN, GWO-DNN, WOA-DNN and FOA-DNN, respectively at 35%. The F1-score of the developed F-WOA-HDNN is superior to 15.5% PSO-DNN, 6.3% superior to GWO-DNN, and 24% superior to WOA-DNN and 34% superior to FOA-DNN at 55%. Similarly, the superior performance is observed by suggested HDNN with F-WOA for all the performance measures. Finally, the suggested SenDemonNet model using HDNN along with F-WOA and weighted feature selection is outperformed the existing heuristic-based algorithms.

6.4 Performance analysis on different classifiers

The performance of the designed SenDemonNet model using HDNN with F-WOA is analyzed with existing classifiers as given in Fig. 7. The accuracy of the proposed HDNN with F-WOA is 1%, 1%, 0.5% and 2.1% advanced than SVM, CNN, LSTM and DNN, respectively at 35%. The FNR of the developed F-WOA-HDNN is 3.8%, 40.4%, 37.5% and 26% progressed than SVM, CNN, LSTM and DNN, respectively at 65%. The FPR of the proposed F-WOA-HDNN is 5.5%, 26%, 10.5% and 27.6% enhanced than SVM, CNN, LSTM and DNN, respectively at 75%. The specificity of the developed F-WOA-HDNN is 1.3%, 2.8%, 1.5% and 2.6% superior to SVM, CNN, LSTM and DNN, respectively at 95%. Likewise, the better performance is observed for all the performance measures. Further, the SenDemonNet model using F-WOA-HDNN and weighted feature selection establishes the superior performance when compared to the conventional classifiers.

6.5 Analysis on K-fold validation

The designed SenDemonNet model using HDNN along with F-WOA and weighted feature selection is validated in terms of k-fold validation with different meta-heuristic-based

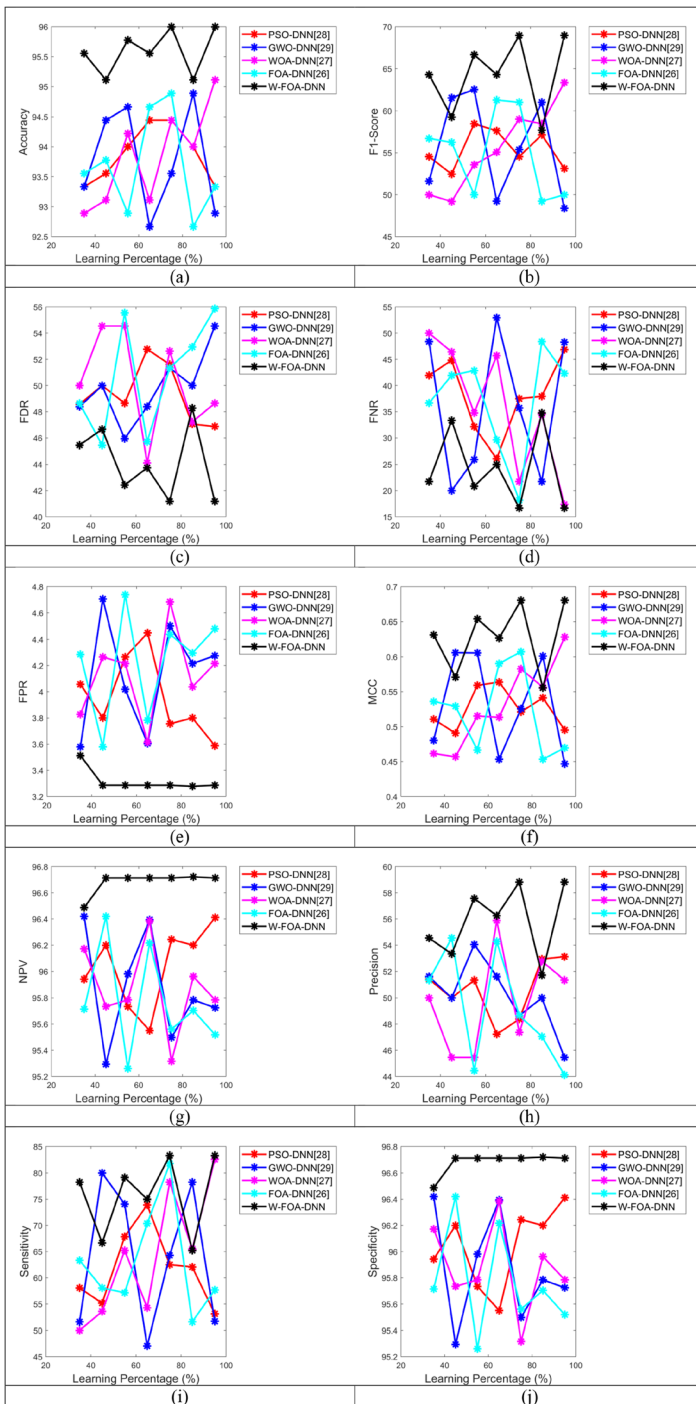


Fig. 6 Performance analysis of the proposed SenDemonNet model for demonetization tweets with traditional meta-heuristic-based algorithms concerned with “(a) Accuracy, (b) F1-score, (c) FDR, (d) FNR, (e) FPR, (f) MCC, (g) NPV, (h) precision, (i) Sensitivity and (j) Specificity”

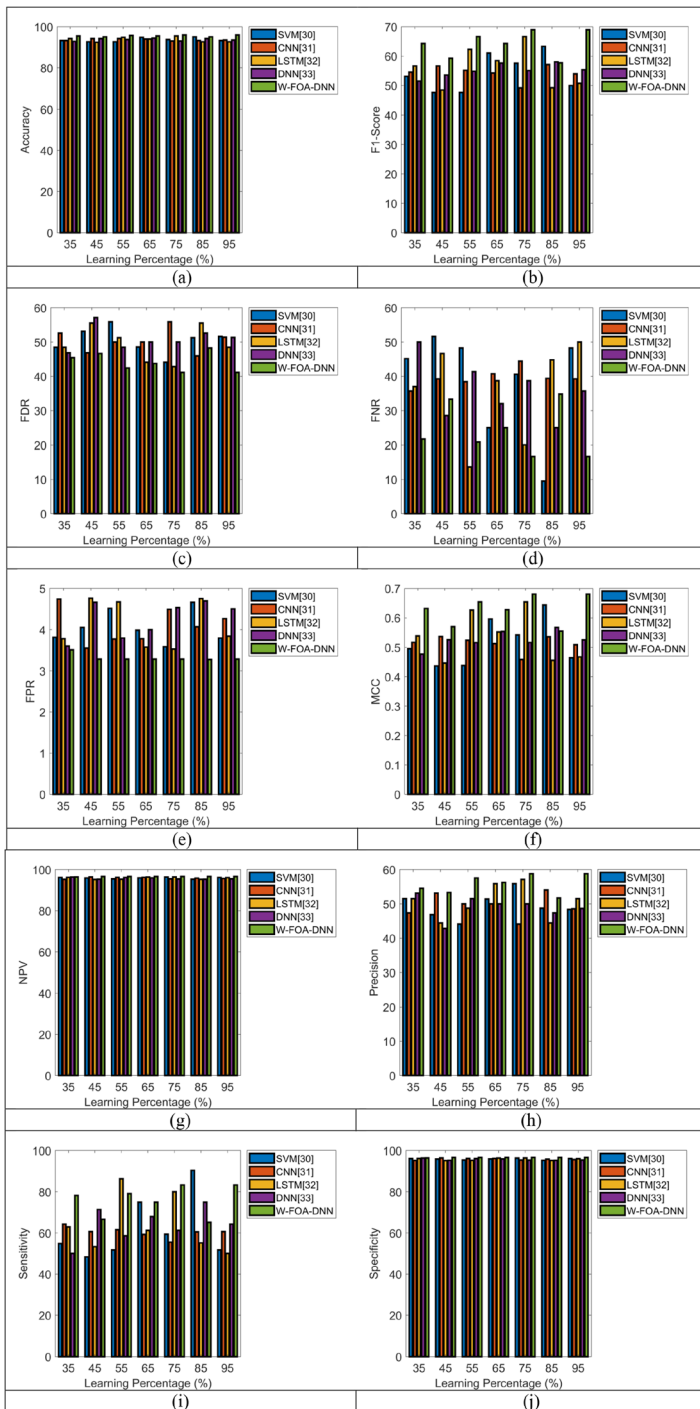


Fig. 7 Performance analysis of the proposed SenDemonNet model for demonetization tweets with traditional classifiers concerned with “(a) Accuracy, (b) F1-score, (c) FDR, (d) FNR, (e) FPR, (f) MCC, (g) NPV, (h) precision, (i) Sensitivity and (j) Specificity”

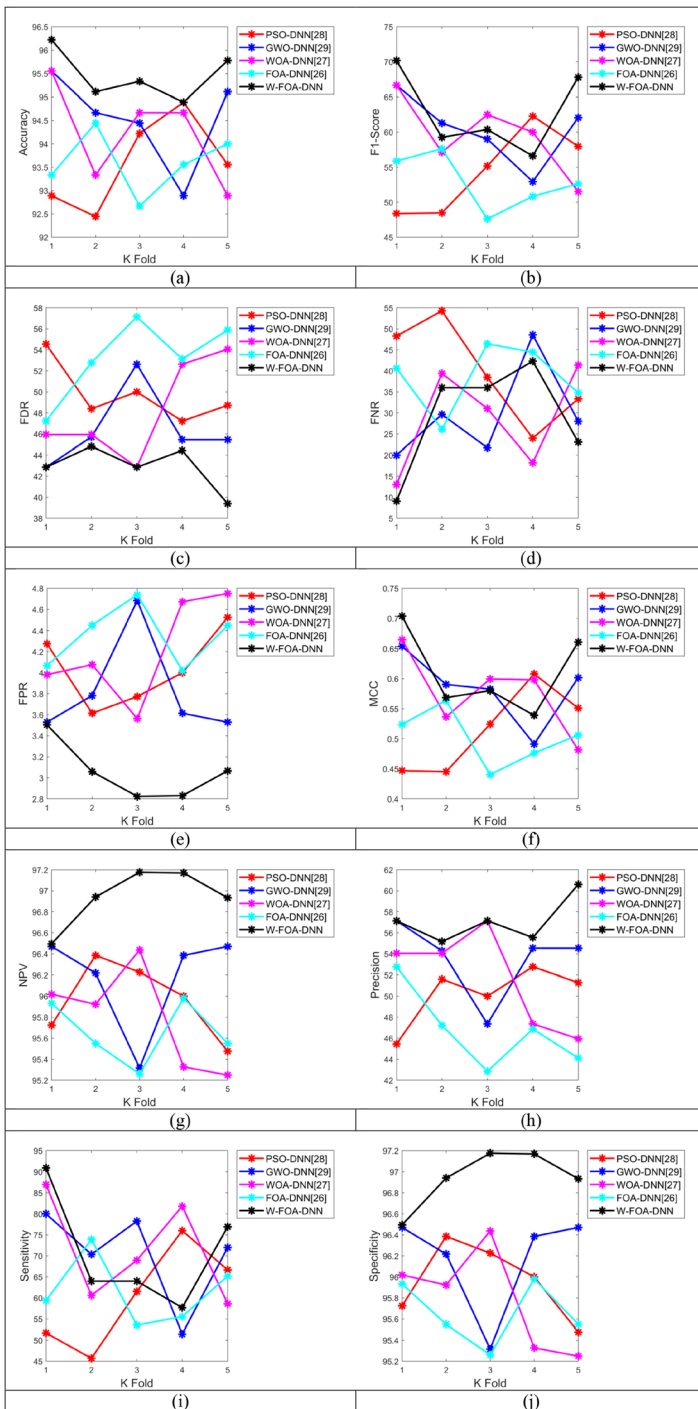


Fig. 8 K-fold validation on the proposed SenDemonNet model for demonetization tweets with traditional classifiers concerned with “(a) Accuracy, (b) F1-score, (c) FDR, (d) FNR, (e) FPR, (f) MCC, (g) NPV, (h) precision, (i) Sensitivity and (j) Specificity”

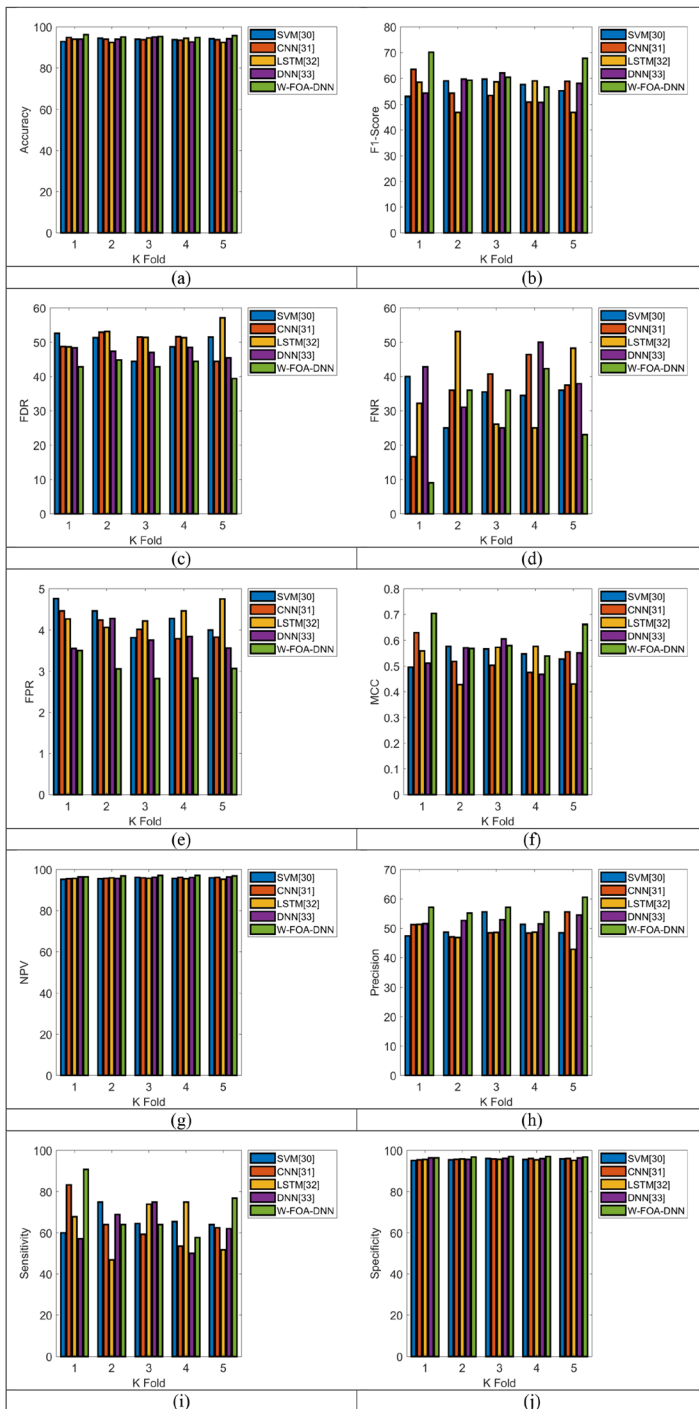


Fig. 9 K-fold validation on the proposed SenDemonNet model for demonetization tweets with traditional classifiers concerned with “(a) Accuracy, (b) F1-score, (c) FDR, (d) FNR, (e) FPR, (f) MCC, (g) NPV, (h) precision, (i) Sensitivity and (j) Specificity”

Table 4 Overall performance analysis for the proposed SenDemonNet model over diverse meta-heuristic-based approaches

Measures	PSO-DNN [35]	GWO-DNN [30]	WOA-DNN [20]	FOA-DNN [11]	F-WOA-DNN
“Accuracy”	0.94444	0.92667	0.93111	0.94667	0.95556
“Sensitivity”	0.73913	0.47059	0.54286	0.7037	0.75
“Specificity”	0.9555	0.96394	0.96386	0.96217	0.96714
“Precision”	0.47222	0.51613	0.55882	0.54286	0.5625
“FPR”	0.044496	0.036058	0.036145	0.037825	0.032864
“FNR”	0.26087	0.52941	0.45714	0.2963	0.25
“NPV”	0.9555	0.96394	0.96386	0.96217	0.96714
“FDR”	0.52778	0.48387	0.44118	0.45714	0.4375
“F1-score”	0.57627	0.49231	0.55072	0.6129	0.64286
“MCC”	0.56387	0.45344	0.51349	0.59046	0.62697

algorithms and classifiers as given in Figs. 8 and 9, respectively. The designed HDNN with F-WOA attains 3.7%, 0.8%, 0.8% and 3.1% higher accuracy than PSO-DNN, GWO-DNN, WOA-DNN and FOA-DNN, respectively by considering the k-fold as 1. The specificity of the developed F-WOA-HDNN is 2%, 1.8%, 1% and 0.5% progressed than SVM, CNN, LSTM and DNN, respectively while mentioning the k-fold as 4. Therefore, the deep learning-based SenDemonNet model using weighted feature selection and HDNN along with F-WOA show the better performance than existing methods.

6.6 Overall performance analysis

The comparative analysis of the designed SenDemonNet model using F-WOA-DNN is evaluated with existing algorithms and classifiers as shown in Tables 4 and 5, respectively. The accuracy of the designed F-WOA-DNN is 1.1%, 3%, 2.6%, 0.9%, 0.7%, 1.6%, 1.6% and 1.1% enhanced than PSO-DNN, GWO-DNN, WOA-DNN, FOA-DNN, SVM, CNN, LSTM and DNN respectively. Similarly, the suggested F-WOA-DNN model gives better performance than other methods. Hence, the proposed SenDemonNet model using F-WOA-DNN shows the superior performance than other conventional approaches.

Table 5 Overall performance analysis for the proposed SenDemonNet model over diverse classifiers

Measures	SVM [21]	CNN [16]	LSTM [22]	DNN [29]	F-WOA-DNN
“Accuracy”	0.94889	0.94	0.94	0.94444	0.95556
“Sensitivity”	0.75	0.59259	0.6129	0.68	0.75
“Specificity”	0.96009	0.96217	0.9642	0.96	0.96714
“Precision”	0.51429	0.5	0.55882	0.5	0.5625
“FPR”	0.039906	0.037825	0.0358	0.04	0.032864
“FNR”	0.25	0.40741	0.3871	0.32	0.25
“NPV”	0.96009	0.96217	0.9642	0.96	0.96714
“FDR”	0.48571	0.5	0.44118	0.5	0.4375
“F1-score”	0.61017	0.54237	0.58462	0.57627	0.64286
“MCC”	0.59576	0.51263	0.55304	0.5547	0.62697

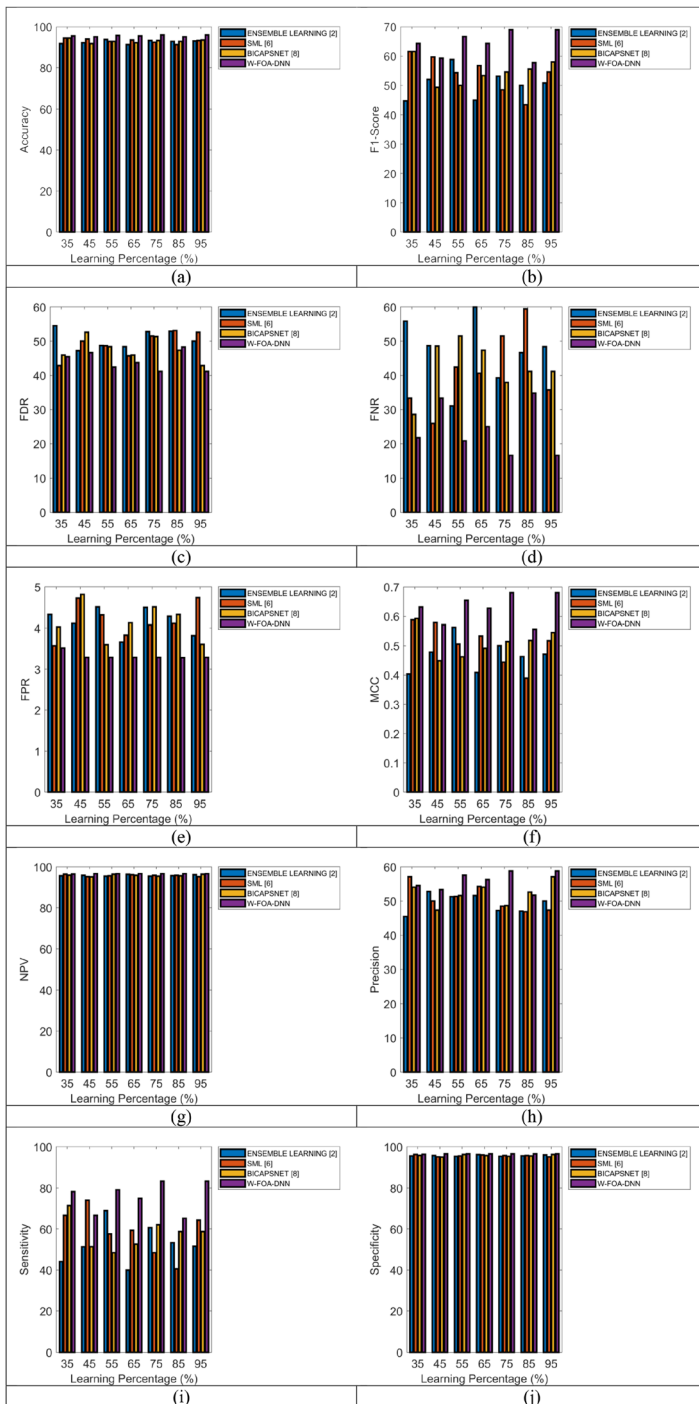


Fig. 10 Performance analysis of the proposed SenDemonNet model for demonetization tweets with existing sentiment analysis models concerned with “(a) Accuracy, (b) F1-score, (c) FDR, (d) FNR, (e) FPR, (f) MCC, (g) NPV, (h) precision, (i) Sensitivity and (j) Specificity”

Table 6 Statistical analysis for the proposed SenDemonNet model over diverse meta-heuristic-based approaches

Measures	PSO-DNN [35]	GWO-DNN [30]	WOA-DNN [20]	FOA-DNN [11]	F-WOA-DNN
Best	0.94444	0.94889	0.95111	0.94889	0.96
Worst	0.93333	0.92667	0.92889	0.92667	0.95111
Mean	0.93873	0.93778	0.93841	0.93683	0.95587
Median	0.94	0.93556	0.94	0.93556	0.95556
Standard deviation	0.004776	0.008889	0.008287	0.008399	0.003725

6.7 Performance analysis on existing methods

The performance of the designed SenDemonNet model using HDNN with F-WOA is analyzed with existing sentiment analysis models as given in Fig. 10. The superior performance is observed by suggested strategy compared with traditional methods due to the challenges listed in Section 2.2. Thus, the performance is high by designed model SenDemonNet model using F-WOA-HDNN and weighted feature selection.

6.8 Statistical analysis

The statistical analysis of the designed SenDemonNet model model for Demonetization Tweets with diverse meta-heuristic-based approaches and diverse classifiers is given in Tables 6 and 7, respectively. The considered techniques are in stochastic nature and the experiment is executed five times. This analysis is carried out by considering the measures like best, worst, mean and standard deviation. “The mean is the average value of the best and worst values and the median is referred to as the center point of the best and worst values whereas the standard deviation is represented as the degree of deviation between each execution”. The performance of the designed F-WOA-DNN is 1.8%, 1.9%, 1.86% and 2% enhanced than PSO-DNN, GWO-DNN, WOA-DNN, FOA-DNN, SVM, CNN, LSTM and DNN respectively. Similarly, the suggested F-WOA-DNN model gives better performance while comparing with the other methods for all the values.

6.9 Discussion

The parameter values of the proposed model is given to ensure reproducibility of the results in Table 8.

To show the effectiveness of our hybrid Heuristic F-WOA technique with DNN, the well-performing optimization algorithms have been adopted with DNN for validating the high accuracy rate of suggested model in terms of sentiment extraction. When comparing with the

Table 7 Statistical analysis for the proposed SenDemonNet model over diverse classifiers

Measures	SVM [21]	CNN [16]	LSTM [22]	DNN [29]	F-WOA-DNN
Best	0.95111	0.94222	0.95556	0.94444	0.96
Worst	0.92667	0.93111	0.92444	0.92889	0.95111
Mean	0.93683	0.93683	0.93778	0.93746	0.95587
Median	0.93333	0.93556	0.94	0.93778	0.95556
Standard deviation	0.009843	0.0046	0.012172	0.005939	0.003725

Table 8 Parameters used for the proposed SenDemonNet model

Description	Values
Weighted features	1 to the total number of PCA features
Range of hidden neurons of DNN	1 to 50
The number of population	10
The maximum number of iterations	25

suggested HDNN using F-WOA algorithm, CNN or LSTM gives lesser performance due to its challenges. CNN suffers from class imbalance, exploding gradient and over-fitting issues. However, LSTM takes more time for processing. Moreover, F-WOA provides superior performance for solving the real non-linear optimization problems. Hence, the proposed F-WOA algorithm poses superior search abilities with better convergence behavior. Due to these features, the designed model has attained promising results when compared with other classifiers. Moreover, the proposed model gets high accuracy, but lower F-score because the F1-score is independent from the true-negatives while accuracy is not. It also does not have imbalanced classes and thus, F1-score is low than accuracy rate. But, in this model, single classifier with heuristic-improvement is used, in future; the hybrid classifier can be adopted for showing the efficiency. The computational complexity of the designed F-WOA algorithm is given based on big oh notation, where the computational complexity of computing fitness function in WOA is given as $O(YDJ)$, where the whale population is given as Y , the dimension of search agents is denoted as D and the number of maximum generations or iterations is termed as J . The computational complexity of FOA and DNN is given as $O(d)$, respectively where the input data is given as d . Thus, the computational complexity of F-WOA algorithm with DNN is given as $O(YDJ) + O(d) + O(d)$. The time complexity of the suggested model with existing approaches are given in Table 9.

6.10 Error analysis

The Error analysis of the proposed SenDemonNet model for demonetization tweets with existing sentiment analysis models is given in Fig. 11. From Fig. 11 the training loss of the designed F-WOA-DNN for iteration 5 is secured as, 3.48%, 2.32%, 4.65%, and 1.162% lower than PSO-DNN, GWO-DNN, WOA-DNN, and FOA-DN respectively. Hence, the proposed SenDemonNet model using F-WOA-DNN shows that the proposed model obtained low error rate than the other conventional approaches for all iterations.

Table 9 Computational time for the proposed SenDemonNet model over diverse approaches

Measures	PSO-DNN [35]	GWO-DNN [30]	WOA-DNN [20]	FOA-DNN [11]	F-WOA-DNN
Computational time (sec)	8.5	6.1	7.3	8	5.5
Measures	SVM [21]	CNN [16]	LSTM [22]	DNN [29]	F-WOA-DNN
Computational time (sec)	6.5	9.8	8.6	7.2	5.5

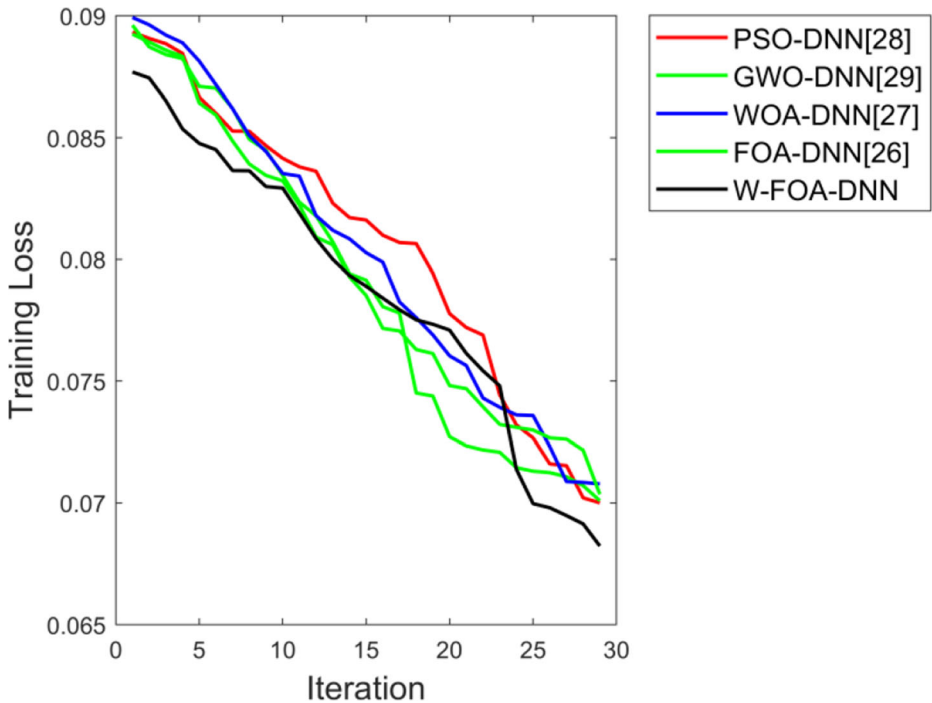


Fig. 11 Error analysis of the proposed SenDemonNet model for Demonetization Tweets with existing sentiment analysis models

6.11 Training scale with batch size

The training scale with a batch size of the proposed SenDemonNet model for demonetization tweets with existing sentiment analysis models is given in Fig. 12. From Fig. 12, the accuracy of the designed F-WOA-DNN at a learning rate 35 obtained, 4.4%, 5.5%, 3.3%, and 6.6% superior to PSO-DNN, GWO-DNN, WOA-DNN, and FOA-DN respectively. Hence, the proposed SenDemonNet model using F-WOA-DNN shows that the proposed model is better than the other conventional approaches.

6.12 Computation time analysis

The computation time of the proposed SenDemonNet model for demonetization tweets with the sentiment analysis model is given in Table 10. The computation time of the designed F-WOA-DNN is obtained, 47.763%, 54.987%, 56.286%, and 49.398 higher to SVM, CNN, LSTM, and DNN respectively. Hence, the proposed SenDemonNet model using F-WOA-DNN shows that the proposed model is better than the other conventional approaches.

6.13 Performance analysis over existing approaches

The performance analysis of the different algorithms and classifiers of the proposed SenDemonNet model for demonetization tweets with the sentiment analysis model is given

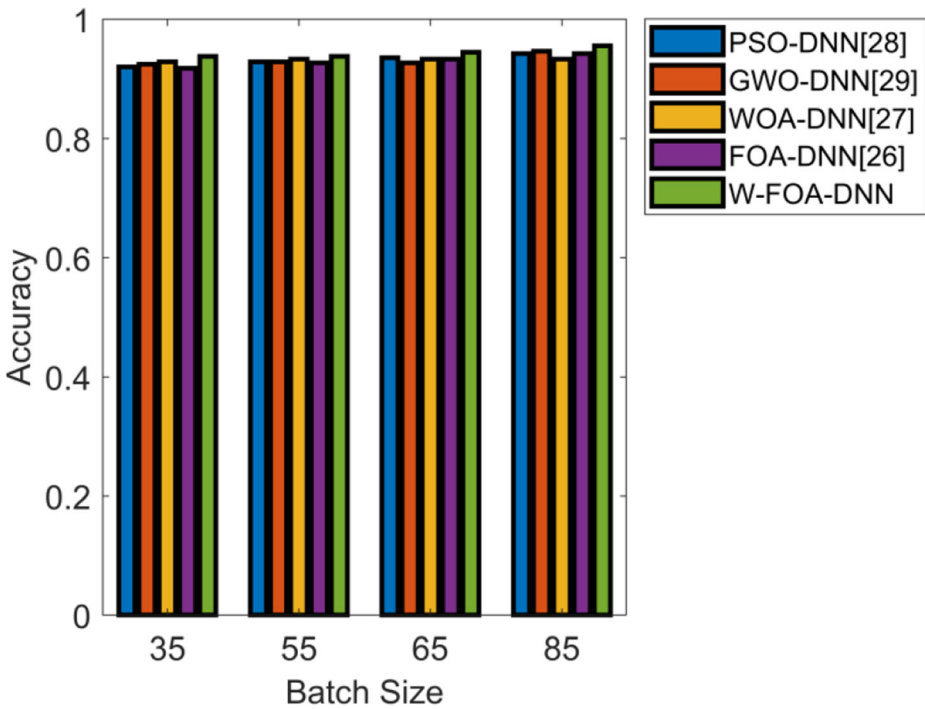


Fig. 12 Training scale with a batch size of the proposed SenDemonNet model for demonetization tweets with existing sentiment analysis models

in Tables 11 and 12 at learning percentages as 95. From Table 11, the accuracy of the designed F-WOA-DNN is acquired, 0.933%, 0.928%, 0.951 and 0.933% superior to PSO, GWO, WOA, and FOA respectively. From Table 12, the accuracy of the designed F-WOA-DNN is achieved 0.933%, 0.935%, 0.926 and 0.933% superior to PSO, GWO, WOA, and FOA respectively. Hence, the proposed SenDemonNet model using F-WOA-DNN shows that the proposed model is better than the other conventional approaches.

6.14 Comparison of feature extraction with deep learning techniques

The comparison of feature extraction techniques with a deep learning algorithm is shown in Fig. 13. From Fig. 13, the accuracy of the designed F-WOA-DNN is acquired, 5.154%, and 04.123% superior to Text-CNN and RESNET respectively. Hence, the proposed SenDemonNet model using F-WOA-DNN shows that the proposed model is better than the other conventional approaches.

Table 10 The computational time for the proposed SenDemonNet model over diverse approaches

Classifiers	SVM [21]	CNN [16]	LSTM [22]	DNN [29]	F-WOA-DNN
Computation Time(sec)	47.763	54.987	56.286	49.398	46.465

Table 11 The performance analysis of a different heuristic algorithm for the proposed SenDemonNet model over diverse approaches

Terms	PSO-DNN [35]	GWO-DNN [30]	WOA-DNN [20]	FOA-DNN [11]	F-WOA-DNN
Accuracy	0.93333	0.92889	0.95111	0.93333	0.96
Sensitivity	0.53125	0.51724	0.82609	0.57692	0.83333
Specificity	0.96411	0.95724	0.95785	0.95519	0.96714
Precision	0.53125	0.45455	0.51351	0.44118	0.58824
FPR	0.035885	0.042755	0.042155	0.044811	0.032864
FNR	0.46875	0.48276	0.17391	0.42308	0.16667
NPV	0.96411	0.95724	0.95785	0.95519	0.96714
FDR	0.46875	0.54545	0.48649	0.55882	0.41176
F1-score	0.53125	0.48387	0.63333	0.5	0.68966
MCC	0.49536	0.44693	0.62846	0.46977	0.68056

7 Conclusion

This paper has developed a new sentiment analysis model named SenDemonNet on demonetization data from Twitter. The dataset is gathered from the “Twitter” package in R along with Twitter API. The main intent of this paper was attempted to understand the public opinion on the recently implemented demonetization policy using the proposed SenDemonNet. Initially, the gathered data was given to the preprocessing stage for preparing the text data for further process. It was then fed to the feature extraction stage to get the most significant features using different approaches like Word2vector, TF-IDF, and Bag of n-grams. Further, these features were subjected to the PCA for reducing the dimensionality of the features to reduce the complexity of processing. As a major contribution, this model has proposed a weighted feature selection method using a newly developed algorithm named F-WOA for selecting the most representative features. The optimal weighted features were further processed by the HDNN through F-WOA for getting better classification results in terms of negative and positive opinions. The experimental analysis has shown the efficiency of the suggested SenDemonNet through F-WOA, which was analyzed in terms of different performance measures. The accuracy of the designed HDNN-based F-WOA was 0.7%, 1.6%, 1.6%, and 1.1% enhanced than SVM, CNN, LSTM, and DNN, respectively. Thus, the performance of the designed SenDemonNet model has outperformed the existing approaches. The computation time of the designed F-WOA-DNN was secured, 47.763%, 54.987%, 56.286%, and 49.398 higher to

Table 12 The performance analysis of different classifiers for the proposed SenDemonNet model over diverse approaches

Terms	SVM [21]	CNN [16]	LSTM [22]	DNN [29]	F-WOA-DNN
“Accuracy”	0.93333	0.93556	0.92667	0.93556	0.96
“Sensitivity”	0.51724	0.60714	0.5	0.64286	0.83333
“Specificity”	0.962	0.95735	0.96154	0.95498	0.96714
“Precision”	0.48387	0.48571	0.51515	0.48649	0.58824
“FPR”	0.038005	0.042654	0.038462	0.045024	0.032864
“FNR”	0.48276	0.39286	0.5	0.35714	0.16667
“NPV”	0.962	0.95735	0.96154	0.95498	0.96714
“FDR”	0.51613	0.51429	0.48485	0.51351	0.41176
“F1-score”	0.5	0.53968	0.50746	0.55385	0.68966
“MCC”	0.46462	0.50913	0.46792	0.5257	0.68056

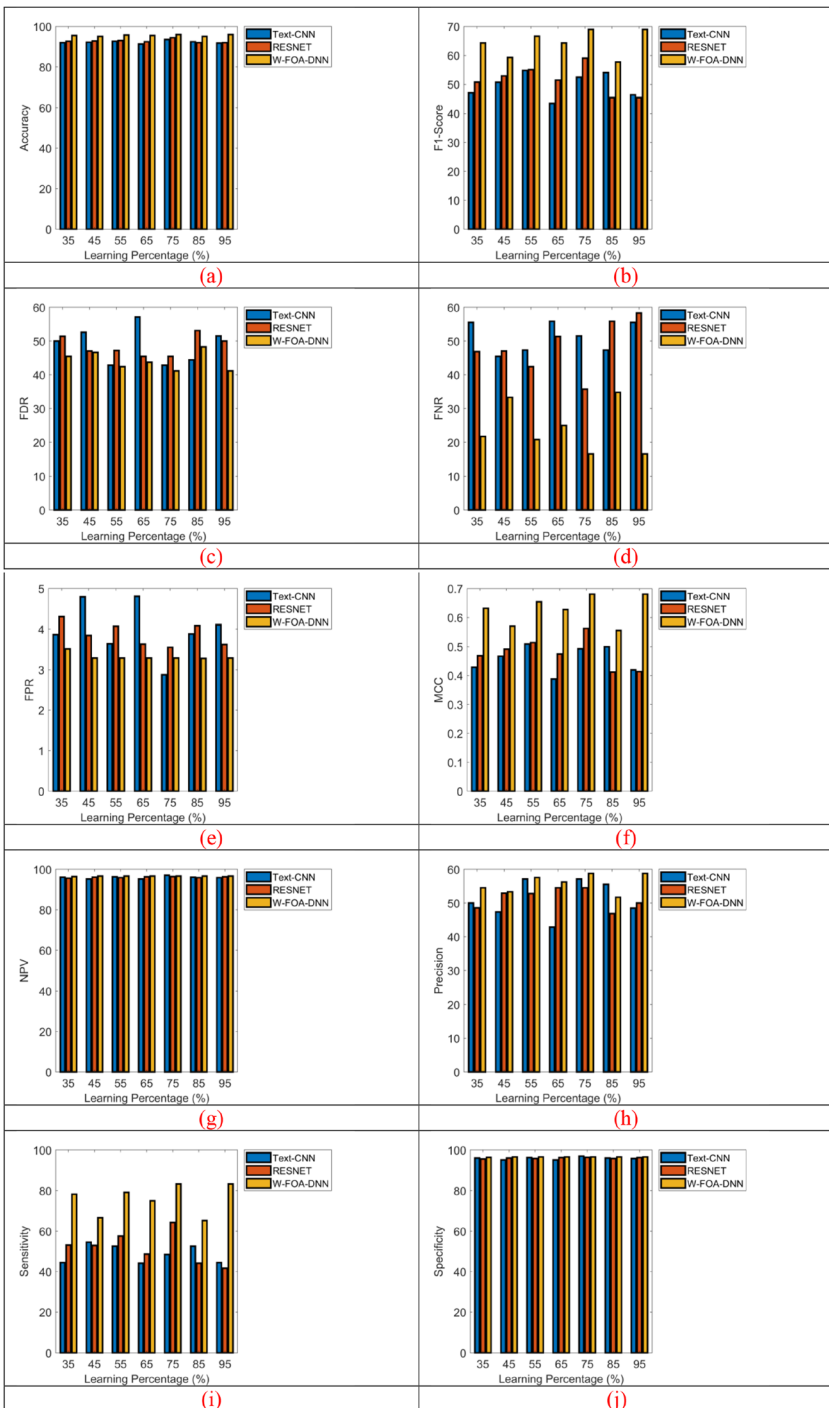


Fig. 13 Comparison of feature extraction techniques with a deep learning algorithm of proposed SenDemonNet model for demonetization tweets with existing sentiment analysis models

SVM, CNN, LSTM, and DNN respectively. Hence, the proposed SenDemonNet model using F-WOA-DNN shows that the proposed model was better than the other conventional approaches. Some popular sentiment analysis applications include social media monitoring, customer support management, and analyzing customer feedback. This research work does not consider sarcasm tweets. In the future, this model can be extended by using more large datasets which can have a huge amount of tweets with intelligent approaches like hybrid classifiers or ensemble learning with heuristic algorithms to get a good result.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. Arun K, Srinagesh A, Ramesh M (2017) Twitter sentiment analysis on demonetization tweets in India using R language. *Int J Comput Eng Res Trends* 4(6):252–258
2. Avinash M, Sivasankar E (2018) A study of feature extraction techniques for sentiment analysis. *Emerg Technol Data Mining Inf Secur* 475–486
3. Aziz AA, Starkey A (2020) Predicting supervise machine learning performances for sentiment analysis using contextual-based approaches. *IEEE Access* 8:17722–17733
4. Das S, Das D, Kolya AK (2020) Sentiment classification with GST tweet data on LSTM based on polarity-popularity model. *Sādhanā* 45:140
5. Datta S, Chakrabarti S (2021) Aspect based sentiment analysis for demonetization tweets by optimized recurrent neural network using fire fly-oriented multi-verse optimizer. *Sādhanā* 46(2):1–23
6. Dhanya NM, Harish UC (2018a) Sentiment analysis of Twitter data on demonetization using machine learning techniques. In: *Computational vision and bio inspired computing*. Springer, Cham, pp 227–237
7. Dubey AD (2020) The resurgence of cyber racism during the COVID-19 pandemic and its aftereffects: analysis of sentiments and emotions in tweets. *JMIR Public Health Surveill* 6(4):e19833
8. El-Affendi MA, Alrajhi K, Hussain A (2021) A novel deep learning-based multilevel parallel attention neural (MPAN) model for multidomain Arabic sentiment analysis. *IEEE Access* 9:7508–7518
9. Fauzi MA (2018) Word2Vec model for sentiment analysis of product reviews in Indonesian language. *Int J Electr Comput Eng* 9(1):525
10. Feng Y, Cheng Y (2021) Short text sentiment analysis based on multi-channel CNN with multi-head attention mechanism. *IEEE Access* 9:19854–19863
11. Ghaemi M, Feizi-Derakhshi M-R (2014) Forest optimization algorithm. *Expert Syst Appl* 41(15):6676–6687
12. Gupta D, Divya, et al. (2021) Edge caching based on collaborative filtering for heterogeneous ICN-IoT applications. *Sensors* 21(16):5491
13. Huh J-H, Seo Y (2019) Understanding edge computing: engineering evolution with artificial intelligence. *IEEE Access* 7:164229–164245
14. Kanimozhi P, Elavarasi D (2018) Survey on sentiment analysis using Twitter dataset. *Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE
15. Kharde VA, Sonawane SS (2016) Sentiment analysis of Twitter data: a survey of techniques. *Int J Comput Appl* 139(11):5–15
16. Kim H, Jeong Y-S (2019) Sentiment classification using convolutional neural networks. *Appl Sci* 9(11):2347
17. Lee H, Park S-H, Yoo J-H, Jung S-H, Huh J-H (2020) Face recognition at a distance for a stand-alone access control system. *Sensors* 20(3):785
18. Li Z, Li R, Jin G (2020) Sentiment analysis of Danmaku videos based on Naïve Bayes and sentiment dictionary. *IEEE Access* 8:75073–75084
19. Liu H, Chatterjee I, Zhou M, Lu XS, Abusorrah A (2020) Aspect-based sentiment analysis: a survey of deep learning methods. *IEEE Trans Comput Soc Syst* 7(6):1358–1375
20. Mirjalili S, Lewis A (2016) The whale optimization algorithm. *Adv Eng Softw* 95:51–67

21. Nafis NSM, Awang S (2021) An enhanced hybrid feature selection technique using term frequency-inverse document frequency and support vector machine-recursive feature elimination for sentiment classification. *IEEE Access* 9:52177–52192
22. Pal S, Ghosh S, Nag A (2018) Sentiment analysis in the light of LSTM recurrent neural networks. *Int J Synth Emot* 9(1):33–39
23. Panigrahi R, Borah S, Bhoi AK, Ijaz MF, Pramanik M, Kumar Y, Jhaveri RH (2021a) A consolidated decision tree-based intrusion detection system for binary and multiclass imbalanced datasets. *Mathematics* 9(7):751
24. Panigrahi R, Borah S, Bhoi AK, Ijaz MF, Pramanik M, Jhaveri RH, Chowdhary CL (2021b) Performance assessment of supervised classifiers for designing intrusion detection systems: a comprehensive review and recommendations for future research. *Mathematics* 9(6):690
25. Park S-W, Huh J-H, Kim J-C (2020) BEGAN v3: avoiding mode collapse in GANs using variational inference. *Electronics* 9(4):688
26. Phan HT, Tran VC, Nguyen NT, Hwang D (2020) Improving the performance of sentiment analysis of tweets containing fuzzy sentiment using the feature ensemble model. *IEEE Access* 8:14630–14641
27. Rani S et al (2021) An optimized framework for WSN routing in the context of industry 4.0. *Sensors* 21(19):6474
28. Rocchetti M et al (2017) Attitudes of Crohn's disease patients: infodemiology case study and sentiment analysis of Facebook and Twitter posts. *JMIR Public Health Surveill* 3(3):e51
29. Sadr H, Pedram MM, Teshnehlab M (2020b) Multi-view deep network: a deep model based on learning features from heterogeneous neural networks for sentiment analysis. *IEEE Access* 8:86984–86997
30. Salam MA (2020) Optimizing extreme learning machine using GWO algorithm for sentiment analysis. *Int J Comput Appl* 176(38):22–28
31. Salur MU, Aydin I (2020) A novel hybrid deep learning model for sentiment classification. *IEEE Access* 8: 58080–58093
32. Schouten K, Frasincar F (2016) Survey on aspect-level sentiment analysis. *IEEE Trans Knowl Data Eng* 28(3):813–830
33. Singh P, Dave A, Dar K (2017) Demonetization: sentiment and retweet analysis. In: *Proceedings of the International Conference on Inventive Computing and Informatics*
34. Singh P, Sawhney RS, Kahlon KS (2018a) Sentiment analysis of demonetization of 500 & 1000 rupee banknotes by Indian government. *ICT Express* 4(3):124–129
35. Sousa T, Silva A, Neves A (2004) Particle swarm based data mining algorithms for classification tasks. *Parallel Comput* 30(5–6):767–783
36. Tang T, Tang X, Yuan T (2020) Fine-tuning BERT for multi-label sentiment analysis in unbalanced code-switching text. *IEEE Access* 8:193248–193256
37. Tripathy A, Agrawal A, Rath SK (2016) Classification of sentiment reviews using N-gram machine learning approach. *Expert Syst Appl* 57:117–126
38. Vinodhini G, Chandrasekaran RM (2013) Effect of feature reduction in sentiment analysis of online reviews. *Int J Adv Res Comput Eng Technol (IJARCET)* 2(6):2165–2172
39. Wang C, Xiao Z, Liu Y, Xu Y, Zhou A, Zhang K (2013) SentiView: sentiment analysis and visualization for internet popular topics. *IEEE Trans Hum Mach Syst* 43(6):620–630
40. Wang Y, Huang G, Li J, Li H, Zhou Y, Jiang H (2021) Refined global word embeddings based on sentiment concept for sentiment analysis. *IEEE Access* 9:37075–37085
41. Wu J, Lu K, Su S, Wang S (2019a) Chinese micro-blog sentiment analysis based on multiple sentiment dictionaries and semantic rule sets. *IEEE Access* 7:183924–183939
42. Xu G, Yu Z, Yao H, Li F, Meng Y, Wu X (2019a) Chinese text sentiment analysis based on extended sentiment dictionary. *IEEE Access* 7:43749–43762
43. Yang L, Li Y, Wang J, Sherratt RS (2020) Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning. *IEEE Access* 8:23522–23530
44. Zhai G, Yang Y, Wang H, Du S (2020) Multi-attention fusion modeling for sentiment analysis of educational big data. *Big Data Min Analytics* 3(4):311–319
45. Zhang K, Jiao M, Chen X, Wang Z, Liu B, Liu L (2019) SC-BiCapsNet: a sentiment classification model based on bi-channel capsule network. *IEEE Access* 7:171801–171813
46. Zhou J, Huang JX, Chen Q, Hu QV, Wang T, He L (2019) Deep learning for aspect-level sentiment classification: survey, vision, and challenges. *IEEE Access* 7:78454–78483
47. Zhou J, Jin S, Huang X (2020) ADeCNN: an improved model for aspect-level sentiment analysis based on deformable CNN and attention. *IEEE Access* 8:132970–132979