Aspect based sentiment analysis for demonetization tweets by optimized recurrent neural network using fire fly-oriented multi-verse optimizer

SAMIK DATTA* and SATYAJIT CHAKRABARTI

Department of Computer Science and Engineering, University of Engineering and Management, Kolkata 700156, India
e-mail: samikdatta123@gmail.com; satyajit.chakrabarti@iemcal.com

Abstract. In this paper, it is proposed to understand the opinion of the public regarding the policy of demonetization that is implemented recently in India through Aspect-based Sentiment Analysis (ABSA) that predicts the sentiment of specific aspects present in the text. The major aim is to identify the relevant contexts for various aspects. Most of the conventional techniques have adopted attention mechanisms and deep learning concepts that decrease the prediction accuracy and generate huge noise. Another major disadvantage with the attention mechanisms is that the sentiment related to few context words alters with various aspects, and hence it cannot be concluded from itself alone. This paper adopts the optimized deep learning concept for performing the ABSA for demonetization tweets. The proposed model involves various phases such as pre-processing, aspect extraction, polarity feature extraction, and sentiment classification. Initially, the different demonetization tweets collected from the Kaggle dataset are taken. Pre-processing is done with the help of four phases like stop words removal, punctuation removal, lower case conversion, and stemming from minimizing the data to its reduced format. This pre-processed data is further performed with aspect extraction to extract the opinion words. These extracted aspect words are converted to the features with the help of polarity score computation and Word2vec. The weight of the polarity scores is optimized using hybridization of two meta-heuristic algorithms like FireFly Algorithm (FF), and Multi-Verse Optimization (MVO), and the new algorithm is termed as Fire Fly-oriented Multi-Verse Optimizer (FF-MVO). Further, combined features are subjected to a deep learning algorithm called Recurrent Neural Network (RNN). As a modification to the existing RNN, the hidden neurons are optimized by the hybrid FF-MVO, FF-MVO-RNN classifies the positive and negative sentiments. Finally, the comparative analysis of different machine learning algorithms proves the competent performance of the proposed model.

Keywords. Aspect-based sentiment analysis; demonetization tweets; recurrent neural network; fire fly-oriented multi-verse optimizer; aspect extraction; sentiment classification; Vader sentiment intensity analyzer; word2vec; weighted polarity score.

1. Introduction

Sentiment analysis or opinion mining is the technique of investigating the behavior of feelings, opinions, and emotions for a particular object or topic that is indicated in textual data [1]. The companies use sentiment analysis to increase their product's sales and services [2, 3]. In the case of private business, sentiment analysis shapes the needs and views of the user that leads to the institution's recognition of the service of the customer. A fine-grained sentiment analysis task is referred to as the ABSA. Its motive is to recognize the sentiment polarity of one sentence in the path of an aspect word that is also known as the target word.

ABSA involves two subtasks, such as sentiment classification and aspect detection. Aspect information is important in the case of a specific product feature or quality, and it is the user-produced content [4, 5]. Unexpected demonetization is not a novel experience for India [6].

Many larger denomination banknotes were available only with banks at that time, the distribution of the larger denomination banknotes was very less. As stated by RBI records in 2016, Indian rupee notes were distributed to the public that rewards around 16,664 billion. Among these 86% banknotes, some of the banknotes that worth 14.180 billion are in Rs 1000 and Rs 500. The dishonest people hold money in 1000 and 500 rupee banknotes [7, 8]. Hence, the government emphasized that the demonetization of Rs 1000 and Rs 500 notes will eradicate the black money. Still,

*For correspondence

time was given to the people from November 2016 to swap their old rupee notes with the banks. Additionally, people could save their old rupee notes in their particular bank accounts [9]. Moreover, the government allowed the public to employ these rupee notes for various favors like obtaining diesel, petrol, rail tickets, and air tickets [10, 11]. As the announcement was given, people responded to various responses to these plans [12, 13]. Recurrent neural networks are useful for deep learning works, such as LSTMs, GRUs, and NTMs [14]. Because of COVID-19 pandemic, social media platforms plays a huge role to express about the health care, feelings about the individuals and governments for communicating about the COVID-19 [15].

ABSA is the technique where the sentiments with distinct aspects are recognized. Aspects are defined as the characteristics, attributes, or features of a service or a product [16, 17]. ABSA helps these industrial business companies to be noticed of the good comments given to their service or product by the public [18]. Therefore, depending on the online reviews of the internet platform, they can enhance their service or product. One more benefits of ABSA in industrial business companies are that it can conserve budget, manpower and time. The major studies in ABSA are unsupervised and supervised. The unsupervised studies are classified into linguistic resources-based and topic model-based strategies. Still, two major disadvantages are there in conventional researches [19, 20]. In the first case, ABSA has concentrated on the English language. So, it has gained low concentration in remaining languages owing to the unavailability of unsupervised effective techniques and the poor linguistic resources. Secondly, most of the present techniques cover only a portion of the three labeled subtasks, but in the ABSA system, all the three must be completely executed [21]. Therefore, ABSA on demonetization policy has to be considered by different views of the people in an efficient manner.

1.1 *Motivation and problem statement*

Sentiment analysis is used for finding the positive and negative sentiments. It also defines the process, which gathers information about how customers reacting to products or services. It monitors and measures sentiments in social media and gains insights using large volumes of text data. All the feedbacks are impossible to read by a single person. By using sentiment analysis we can get how customers feel about different areas without having to read the whole feedback. ABSA goes one step further than sentiment analysis by automatically assigning sentiments to specific features or topics. It involves breaking down text data into smaller fragments and provides more granular and accurate insights from the data. Hence, this paper proposes an ABSA for demonetization tweets.

- To perform the ABSA for demonetization tweets by gathering the various tweets related to the demonetization policy from the Kaggle dataset.
- To execute the pre-processing phase for reducing the data to its minimized format using various steps like stop words removal, remove punctuation, lower case conversion, and stemming.
- To carry out the aspect extraction phase to extract the opinion words and next the polarity scores are measured with Vader sentiment intensity analyzer, and Word2vec converts the features to words.
- To develop the optimal weighted polarity score by optimizing the weight of the polarity score by means of proposed FF-MVO to maximize the classification accuracy by concerning the trained information.
- To develop an optimized RNN by tuning or optimizing the hidden neurons of RNN to classify the final sentiment as positive or negative.
- Since hybrid optimization algorithms have been applied successfully to diverse Engineering applications and can solve real-world and complex optimization challenges, adopting hybrid FF-MVO in both weighted feature extraction and classification is the key contribution.

The paper is arranged in the following manner. Section 1 provides the introduction about the ABSA for the demonetization tweets. Section 2 explains the various literary works that contributed towards the demonetization tweets related to the demonetization tweets in ABSA. The ABSA for demonetization data is explained in section 3. Section 4 describes the hybrid FF-MVO for ABSA. The pre-processing of demonetization data and aspect extraction is described in section 5. Section 6 explains the weighted polarity score computation for ABSA. Section 7 discusses the results and conclusions. In the last section 7, the conclusions are provided.

2. Literature survey

Sentiment analysis is used to check whether the given text is positive, negative, or neutral. There are three types of sentiment analysis: (a) Document-level sentiment analysis, (b) Sentence-level sentiment analysis, and (c) Aspect-level sentiment analysis. The document-level Sentiment Analysis module examines text to check whether it is a positive or negative statement. It works better when the text has 40 characters in length. Sentence-level sentiment analysis aims to check the opinions expressed in sentences are positive, negative, or neutral statements.

The aspect level sentiment analysis is considered as the better approach because it helps the businesses for analyzing a huge amount of data. It also saves money and time. It focuses on the most important task and completes it within a particular time.

2.1 *Document-level sentiment analysis*

In 2008, Denecke [22] developed a methodology in a multilingual framework to determine the polarity of text. By using standard translation software, a document that is not an English language is translated into English. The translated document is classified as positive and negative. Here, it combines the existing technologies, standard translation software, and existing sentiment analysis for classifying text according to positive and negative sentiments. SentiWordNet can be applied to any opinion related tasks for sentiment analysis for better results. The testing for accuracy is impossible; it requires annotation of WordNet.

2.2 *Sentence-level sentiment analysis*

In 2016, Mishra *et al* [23] developed a sentiment analysis of Twitter data, and it conveys the opinion about Mr Modi's Digital India Campaign. The sentiments were collected, and the polarity of sentiments is classified in the opinion with positive, negative, or neutral. In this paper, the Dictionary Based Approach is used to analyze the data based on different users. Polarity classification is done from the data obtained through this approach. The challenges related to the sentiment analysis are sarcasm sentences handling, Negation handling, and sentences with emoticons and expressing the way of their opinions.

2.3 *Aspect-level sentiment analysis*

In 2018, Pong-inwong *et al* [24], have developed a new sentiment analysis method called phrase pattern matching. The comments which are posted in a teaching evaluation system in the form of open-ended questions are extracted. The students are allowed to provide feedback on factors that affects the classroom and studying in the classroom. The goal of this research is to collect feedback through open-ended questions and to determine the best classification of responses to open-ended questions by classifying attitude as positive or negative. Without a proper grammatical structure, it is difficult to find out the sentiment phrase in each sentence. The method which is proposed in this paper is called SPMM. This method is flexible for finding patterns of languages.

In 2019, Meng *et al* [25] have proposed an aspect-level neural network for sentiment analysis known as CNN-BiLSTM (FEA-NN). This technique employed CNN to remove a larger-level phrase representation series from the embedding layer that gave efficient help for successive coding tasks. BiLSTM was employed to safeguard the semantic information and to enhance the value of context encoding, and to detain both temporal and global sentence semantics and local features of phrases. Further, the interaction relationships were designed between sentences and aspect words using an attention mechanism to concentrate on those targeted keywords to grasp most efficient context representation. The developed technique was estimated on Laptop, Restaurant, and Twitter. The efficiency of FEA-NN was revealed by the vast simulations.

In 2019, Fatima *et al* [26] had addressed a co-extraction technique containing purified word embeddings to utilize the dependency structures without the help of syntactic parsers. A deep learning-based multilayer dual-attention technique was developed to utilize the indirect relationship between the opinion and aspect terms. Moreover, rather than the Word2Vec technique, word embeddings were purified by giving different vector representations to various sentiments. To defeat the conflict of identical vector characterizations of conflicting sentiments, a sentiment refinement method was used for the pre-trained word embedding technique. The performance of the developed technique was estimated with the three benchmark datasets of SemEval Challenge 2014 and 2015. The outcomes revealed the efficiency of the ABSA model when it was differentiated from the conventional techniques.

In 2019, Siyam *et al* [27] had developed an ABSA hybrid technique to examine the entity's smart apps reviews, which combined domain rules and lexicons. The developed technique categorized the relative sentiments and removed the significant aspects from the comments. This technique labelled various sentiment analysis disputes by approving language processing rules, methods, and lexicons to generate concise outcomes. As stated by the reported outcomes, when the implicit aspects were taken into account, it shows an important enhancement in the aspect extraction accuracy. Additionally, the combined classification technique exceeded the lexicon-based baseline and the several rules by 5% concerning average accuracy. Moreover, the developed technique exceeded machine learning techniques, which employed SVM when the same dataset was used. Therefore, providing these rules and lexicons as input features to the SVM technique has better accuracy than the remaining SVM models.

In 2020, Khoshavi and Dastjerdi [28] introduced an unsupervised paradigm ABSA that was easy to handle in various languages and universally carried out the subtasks for ABSA. The technique was composed of three coarse-grained phases that were divided into numerous fine-grained functioning. In the first step, aspect word sets and preliminary polarity lexicon were selected to derive the prior domain knowledge from the dataset as symbolic of aspects. This primitive knowledge was given to an expectation-maximization algorithm to detect the probability based on sentiment and aspect. In the final step, the polarity of aspect was described by breaking the document into its comprising aspects and computing the probability of every polarity/aspect based on the document. This technique was estimated using two datasets in Persian and English languages, and the outcomes were differentiated with several baselines. The outcomes have revealed that the developed

technique exceeded the baselines for opinion-word extraction, aspect, and aspect-level polarity classification.

In 2020, Narapareddy *et al* [29] had developed a new IGCN for understanding the mutual relation among the target and the comparative review context with the help of a bidirectional gating mechanism. The target's sentiment was forecasted using the positional information of contextual words in terms of POS tags, provided target, and domain-specific word embeddings. The efficiency of the developed IGCN technique was revealed by the outcomes on SemEval 2014 datasets.

In 2020, Meskele and Frasincar [30] developed a hybrid method known as ALDONAr for sentence-level ABSA. The effect of every word on an aspect's sentiment value was calculated by the bidirectional context attention technique. The complex structure of a sentence was modeled by the classification module. The field-specific knowledge was employed by the manually produced lexicalized domain ontology. When differentiated with the conventional ALDONA technique, ALDONAr made use of regularization, distinct model initialization, BERT word embeddings, and the Adam optimizer. Additionally, the classification module resulted in advanced outcomes with two 1D CNN layers on standard datasets.

In 2020, Liu and Shen [31] developed a new end-to-end memory neural network, known as ReMemNN, to reduce the problems in word embeddings. A special module known as embedding adjustment learning module was modelled to transmit the pre-trained word embeddings into adjustment word embeddings for handling the demerits of pre-trained word embeddings. A multi-element attention mechanism was modelled to handle the fragile communication in the attention mechanism and to produce more specific aspect dependent sentiment representation and powerful attention weights. Moreover, an explicit memory module was modelled to produce representations and hidden states and to stock these distinct representations. The simulation outcomes revealed that ReMemNN gained the existing performance and exceeded classical baselines. These outcomes revealed that ReMemNN was dataset type-independent and language-independent.

In 2020, Feng *et al* [32] had developed an efficient and lightweight sentiment analysis technique known as DNet, for an on-device inference that was dependent on gated CNN. DNet minimized the size of the model by gaining better performance with low inference latency and could refine aspect-aware context information continuously from the unstructured text by joining the attention mechanism with stacked gated convolution. The simulations on ACL14 Twitter and SemEval 2014 Task 4 datasets revealed that this technique gained the existing performance. Moreover, DNet enhanced the responsiveness by 24 times and minimized the model size by more than 50 times when it was differentiated with the BERT-based technique.

In 2021, Damarta *et al* [33] developeda sentiment analysis of PT PLN (Persero) Twitter account service quality using *k*-nearest neighbors classifier. The opinion from the public about the service of the Twitter account is viewed by using the *k*-nearest neighbors algorithm. It also considers the quality of the public opinion. PT PLN is controlled by using the text mining method. Initially, the data is collected and send to the pre-processing stage. This is done by using the *k*-nearest neighbors classifier to classify data into positive, neutral, and negative classes. This model is used to identify the file, which consists of new tweet data updated by the users. Here, the accuracy is based on the data quality, and the prediction is difficult for a large set of data.

In 2021, Lymperopoulos [34] had developed a path-breaking model called the RC Tweet for the popularity of the tweets. It explains analogy within charging dynamics of capacitor and retweet cascades in an RC circuit. In the data analytics, the RC tweet model has a sound impact in marketing fields. Online networking sites play an important role in interacting marketing messages to a huge audience. At the time to post retweets, the RC Tweet truly reflects the popularity of tweets. The RC tweet does not need any training for retweet cascades. RC tweet suits for a real time popular forecasts. The RC Tweet cannot be captured by the macroscopic and the mechanistic description of the retweet rate.

In 2020, Hong Wei *et al* [35] developed a system, which is called as firefly for finding the news of a given geographical area. The online geotagging procedure are followed and amount of tweets are significantly increased. The locality-aware keywords are found by this method and they are grouped together for detecting the news. This system is performed well in the Twitter and it might face the challenges to the generality in the social medias. Detecting and extracting small news for a local place is a challenge in this method. This method is used to overcome the data sparsity and captures many local stories per day.

In 2021, Girolamo *et al* [36], developed a method in the aim of revealing feom a social stream by using the online clustering method based on the Game theory. The first subsection develops a series of informations which supports the online clustering solution. The second subsection defines a pre-processing phase. The pre-processing phase is performed before the clustering approach, then the final approach defines the game formulation of grouping. This method is used for realizing the online clustering of tweets by grasping the replicator dynamics and the evolutionary game theory. The online detection of arbitrary events within OSN is difficult and proposes a rule-based approach for tweet filtering. The features and challenges of state-of-the-art Aspect Based Sentiment Analysis models are given in table 1.

Table 1. Features and challenges of state-of-the-art aspect based sentiment analysis models.

Author [citation]	Methodology	Features	Challenges
Khoshavi and Dastjerdi [26]	Unsupervised paradigm	It is very easy to use in various languages	It does not rely on weighting methods, language-dependent methods, and aspect extraction algorithms
Meng *et al* [23]	FEA-NN	It universally performs the subtasks for aspect-based sentiment analysis	It does not improve the results with semantic rules or linguistic resources such as intensifier determination or negation
		It provides effective support for consequent coding tasks	It consumes more time for the entire process
Fatima *et al* [24]	Deep learning model	It achieves the best sentiment classification accuracy	
		It exploits the indirect relationships among the opinion and aspect terms	The datasets cannot be applied for in-depth analysis
Narapareddy *et al* [27]	IGCN	It overcomes the problem of similar vector representations of opposite sentiments	
		It predicts the sentiment of a target	It does not improve the accuracy
		It effectively allocates appropriate weights to the constituent words of the target	
Meskele and Frasincar [28]	ALDONAr	It handles the complex structure of a sentence	It does not concentrate on the classification of implicit aspects
Siyam *et al* [25]	Language processing technique	It utilizes field-specific knowledge	It does not investigate semi-automatic methods for domain ontology building
		It significantly improves the aspect extraction accuracy	It does not identify the improvement areas and innovative ideas through the analysis of customer's reviews
Liu and Shen [29]	ReMemNN	It handles the key challenges in aspect-based sentiment analysis	
		It is language-independent and dataset type-independent	It does not explore the specific role of each aspect dependent sentiment representation

Table 1 continued

Author [citation]	Methodology	Features	Challenges
Feng et al [30]	DNet	It generates powerful attention weights and more expressive inner aspect dependent sentiment representations	It does not combine the pre-trained embedding like Bert with the existing models
		It explores a trade off between model complexity and model accuracy	It is difficult to judge the correct emotional polarity
Denecke. [20]	SentiWordNet	It enables on-device inference for opinion analysis and opinion mining	It suffers from limited memory, computing power, and resources
		It is applied to any opinion related tasks for sentiment analysis	The accuracy of testing is impossible
Mishra et al [21]	Dictionary Based Approach	It gives better results fastly	
		It is used to analyze the data based on different user	It is difficult to handle sarcasm sentences
Pong-inwong et al [22]	SPMM	This method is flexible for finding the patterns of languages	It is hard to handle sentences with emoticons and expressing the way of their opinions
R Damarta et al [31]	PT PLN	It is used to identify new tweet data uploaded by users	It is difficult to find out the sentiment phrase in each sentence without grammar
Ilias N. Lymperopoulos [32]	RC Tweet	It doesnot need any training for retweet cascades	It is difficult to predict a large set of data
Hong Wei et al [33]	Firefly	It suits for real time popular forecasts	It cannot be captured by using the macroscopic and the mechanistic description
Girolamo et al [34]	Game theory	It is used to overcome the data sparsity	It is difficult for detecting and extracting small news for a local place
		It is used for realizing the online clustering of tweets	The online detection of arbitrary events within OSN is difficult.

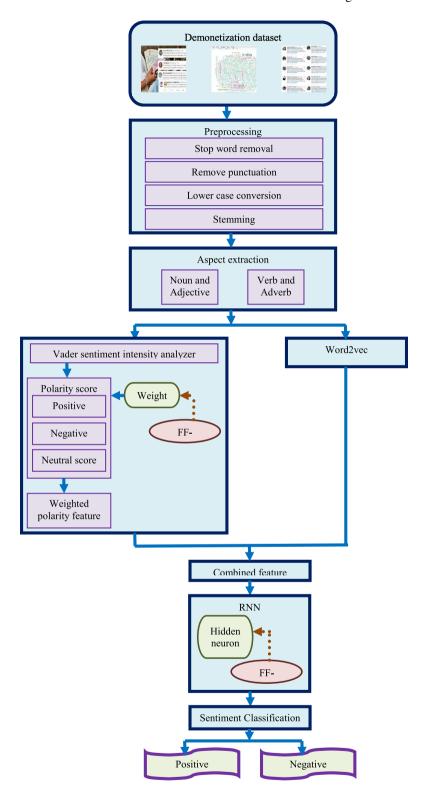**Figure 1.** Proposed architecture of ABSA for demonetization tweets.

3. Hybrid firefly-based multi-verse optimization for aspect-based sentiment analysis

3.1 *Intelligent model for ABSA*

The goal of ABSA is to examine the sentiment polarities having multiple aspect targets or categories with the provided aspect terms (targets) or aspect categories that are spread in sentences. The two various subtasks present in ABSA are the ATSA and ACSA, respectively. In the case of ACSA, an aspect category displays completely in the sentence. While considering ATSA, a particular entity is characterized by an aspect-term (target) that happens clearly in the sentence. The various issues in ABSA are the classification, identification, and aggregation. In most of the conventional techniques, ABSA is considered as a classification problem, in which the information regarding the aspect is integrated. The two major challenges in ABSA related to the extracted features are the classification of the sentiment polarities of aspects and extraction of aspect-specific context features. Since the sentiment orientation is described by a few significant context words, the ABSA needs to extract the aspect specific context features. Existing techniques give the aspect-related information for the classifier by integrating the aspect-dependent features of a sentence, in which the outcomes are based on the features.

RNNs are the familiar models in ABSA to extract the context features via producing sentence representations and representing words as real-value vectors. The NLP tasks involving sentences, question answering, machine translation, learning distributed representations of words, documents, and automatic summarization were dominated by the RNN and its variants like GRU and LSTM. Through semantic correlation modelling, ABSA provided attention mechanisms to identify the sentiment context regarding the provided aspect and to calculate the semantic contribution of every context word in the region of the aspects. The significant features were extracted using the feature extraction and pre-processing phases, and a classifier model can be created using the machine learning and deep learning approaches. The proposed architecture of ABSA for demonetization tweets is shown in figure 1.

The proposed architecture initially gathers different reviews regarding the demonetization tweets from Kaggle. The collected data from Twitter cannot directly undergo analysis or learning. Therefore, the data must be pre-processed to carry out different techniques. Hence, these data are subjected to the pre-processing phase that includes various steps like stop word removal, remove punctuation, lower case conversion, and stemming. Once pre-processing is done, aspect extraction is performed to extract opinion words in the form of the noun, adjectives, verb, adverb, and its pairs. These extracted aspect words are subjected to the Vader sentiment intensity analyzer and word2vec for converting the words to the features. The vader sentiment intensity analyzer helps to determine the polarity score of the corresponding aspect word using positive score, negative score as well as the neutral score. A neutral score involves irrelevant tweets and off-topic comments, and non-English texts. A positive score involves the positive opinion or positive sentiment, and negative scores include negative sentiments. Since the polarity score of the word may not match with the dataset detail (training information), it is essential to add a weight to all three scores of a particular aspect. Hence, the weight added to each score is optimized with the help of hybrid meta-heuristic algorithm known as FF-MVO to maximize the classification accuracy. Thus, the optimal weighted polarity feature is generated. Moreover, the extracted aspect words are converted to vector with the help of Word2vec. The weighted polarity feature is combined with the features obtained from the Word2vec. The combined features are given to the deep learning algorithm called RNN. An improvement is made to the RNN by optimizing the hidden neurons with the same FF-MVO. The training data is created with the positive, negative, and neutral labels of the sample dataset. In the testing phase, this optimized RNN predicts the output by classifying the sentiment as positive or negative.

3.2 *Conventional MVO*

MVO [37] is a familiar theory among physicists. There exist multiple big bangs, and every big bang results in the universe's birth. In the multi-verse theory, multiple universes crash with one another. MVO is motivated by three major concepts like black holes, white holes, and wormholes.

A roulette wheel mechanism is employed to transfer the universe's objects and to mathematically design the black/white hole tunnels. In every round, the universes are sorted in the order of the rate of inflation and one among them is selected by the roulette wheel to contain a white hole. This is represented in Eq. (1).

$$X = \begin{bmatrix} w_1^1 & w_1^2 & \cdots & w_1^c \\ w_2^1 & w_2^2 & \cdots & w_2^c \\ \vdots & \vdots & \vdots & \vdots \\ w_m^1 & w_m^2 & \cdots & w_m^c \end{bmatrix} \qquad (1)$$

Here m represents the count of candidate solutions or universes and c represents the count of variables or parameters. The parameter is initialized as in Eq. (2).

$$w_k^l = \begin{cases} w_j^l & ra1 < NI(Xk) \\ w_k^l & ra1 \geq NI(Xk) \end{cases} \qquad (2)$$

In the above equation, Xk represents the kth universe, $ra1$ represents a random number in the interval range of $[0, 1]$, w_k^l represents the lth parameter of the kth universe, $NI(Xk)$ represents the normalized inflation rate of the kth universe,

and w_j^l represents the lth parameter of the jth universe that is chosen by a roulette wheel mechanism.

The determination and selection of the white holes are performed by the roulette wheel on the basis of the normalized inflation rate. The lower rate of inflation makes the larger probability of passing objects via black/white hole tunnels. The universes are needed to transfer objects for exploring the search space. Without disruption, the universes continue transferring objects. Every universe contains wormholes for carrying out the exploitation and to transfer objects via space in a random manner. Without considering the rate of inflation, the wormholes randomly alter the objects. The formula for this mechanism is represented as in Eq. (3).

$$w_k^l = \begin{cases} \begin{cases} Z_l + TDR \times ((ub_l - lb_l) \times ra4 + lb_l) & ra3 < 0.5 \\ Z_l - TDR \times ((ub_l - lb_l) \times ra4 + lb_l) & ra3 \geq 0.5 \end{cases} & ra2 < WEP \\ w_k^l & ra2 \geq WEP \end{cases} \tag{3}$$

Here, TDR represents a coefficient, lb_l represents the lower bound of lth variable, w_k^l represents the lth parameter of kth universe, Z_l represents the lth parameter of optimal universe, WEP represents a coefficient, ub_l represents the upper bound of lth variable, and $ra2, ra3, ra4$ represents the random numbers in the interval range of $[0, 1]$.

The WEP describes the probability of the presence of a wormhole in the universes. To achieve exploitation as the sequence of the process of optimization, WEP is enhanced linearly over the rounds. TDR describes the distance rate at which an object can be transmitted by a wormhole across the optimal universe. TDR is enhanced over the rounds to hold a more exact local search across the optimally obtained universe. The mathematical formula for both these coefficients is represented in Eq. (4) and Eq. (5).

$$WEP = \min + t \times \left(\frac{\max - \min}{T} \right) \tag{4}$$

Here, max represents the maximum, T represents the maximum iterations, min represents the minimum, and t represents then present iteration.

$$TDR = 1 - \frac{t^{1/r}}{T^{1/r}} \tag{5}$$

In the above equation, r represents the exploitation accuracy over the rounds. The more the value of r, the faster and more accurate exploitation occurs. The process of optimization begins by generating a group of random universes. In each round of the universe, objects with larger inflation rates move towards the lesser inflation rates through black/white holes. This process continues till the stopping criterion is reached.

3.3 Conventional FF

FF [38] criticizes the firefly's social characteristics. With the help of bioluminescence, fireflies search for prey, communicate and find mates using different flashing behaviours. The agent's initial positions are randomly described in the search space as in Eq. (6).

$$w_{t,l}^{(0)} = w_{t,\min} + ra \cdot (w_{t,\max} - w_{t,\min}), \quad t = 1, 2, \ldots, M \tag{6}$$

Here $w_t^{(0)}$ represents the initial value of the tth variable for the lth agent, $w_{t,\min}$, and $w_{t,\max}$ are the minimum and maximum permitted values for the tth variable.

When the distance from the source increases, the attractiveness and light intensity decreases. Thus, the attractiveness and the variations of light intensity should be considered as decreasing functions. Eq. (7) describes the Gaussian form by combining the absorption and the inverse square law.

$$J(s) = J_o e^{-\gamma s^2} \tag{7}$$

In the above equation, J_o represents the original light intensity, J represents the light intensity, and γ represents the light absorption coefficient and it can be considered as a constant. Eq. (8) describes the attractiveness β of a firefly.

$$\beta(s) = \beta_o e^{-\gamma s^2} \tag{8}$$

Here β_o represents a constant, and it produces the attractiveness at $s = 0$. The cartesian distance, represented as $s_{tl} = |w_t - w_l|$ is defined as the distance among two fireflies t and l at w_t and w_l. Equation (9) describes the movement of a firefly t attracting to another firefly l.

$$\Delta w_t = \beta_0 e^{-\gamma s_{tl}^2} (w_l^k - w_t^k) + \alpha \varepsilon_t, \quad w_t^{k+1} = w_t^k + \Delta w_t \tag{9}$$

In the above equation, the first term represents the attraction and the second term represents the randomization, α represents the randomization parameter and ε_t represents the vector of random numbers from a Gaussian distribution. The second term can be replaced with the Levy distribution, in which the step size represents a random number as in Eq. (10).

$$Q(p) = Bp^{-(1+\lambda)}, \quad B = \lambda \Gamma(\lambda) \sin \left(\frac{\lambda \pi}{2} \right) \frac{1}{\pi} \tag{10}$$

Here λ represents the exponent of the distribution and $\Gamma(\lambda)$ represents a Gamma function. The geometrical annealing schedule beginning from the initial α_0 is represented by a function as in Eq. (11).

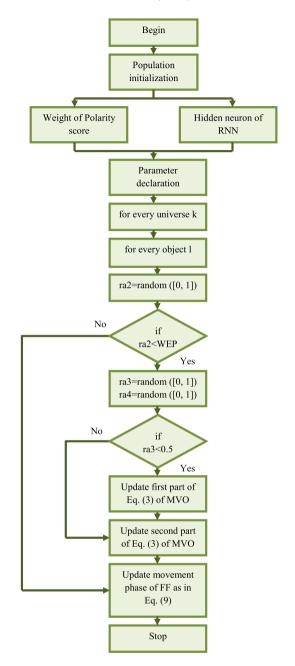$$\alpha = \alpha_0 \theta^k \tag{11}$$

Here, the randomness reduction constant is represented as $0 < \theta < 1$. When γ reaches zero, the brightness and attractiveness become constant and therefore a firefly can

be viewed by all the remaining fireflies. When γ is large, the brightness and attractiveness decrease, and therefore all the fireflies move randomly, corresponding to a random search method. FF finds global optima and local optima very efficiently. Distinct fireflies work in an independent manner providing a parallel implementation. The constraints can be handled by the penalty function technique.

3.4 *Proposed FF-MVO*

Optimization algorithms achieve much attention between the researchers. Complex problems are solved by the optimization algorithms through several modifications and improvements. Through optimization procedures, decision-making systems and expert systems have been produced [39]. Recently, classifier and prediction performance are based on the optimization algorithms. The current ABSA for demonetization dataset, optimal polarity score and optimized RNN-based classification depends on a hybrid meta-heuristic algorithm. The conventional MVO [37] is inspired by three concepts such as the black hole, white hole, and wormhole. The mathematical models of these concepts perform the exploitation, exploration, and local search respectively. MVO finds its strong application in searching the optimal solution. It solves the global optimization problems in a very effective manner. Though it has several advantages, it suffers from various disadvantages like it does not increase the exploration rate for the search space and it also does not enhance the exploitation capability. Therefore, to overcome the drawbacks of the existing MVO, FF is integrated into it, and the new algorithm is termed as FF-MVO. Optimization mechanisms or procedures are joined to produce a hybrid optimization algorithm [40]. For a specific set of search-related problems, hybrid optimization algorithms find their frequent usage. To attain fast convergence, the applications of various optimization algorithms are made by the hybrid optimization algorithms. Compared to the existing algorithms, it attains better convergence behaviour. FF [38] can handle large non-linear multi-modal optimization problems effectively and naturally. The convergence speed is very high in FF in the case of probability of searching the global optimized solution. It has the capability of combining various optimization methods to produce hybrid tools. Moreover, it does not need a good initial solution to begin its iteration process.

Generally, in the conventional MVO, if $ra2 \geq WEP$, the algorithm is updated using Eq. (3) of MVO. But, in the proposed FF-MVO, if $ra2 \geq WEP$, the algorithm is updated

Figure 2. Flowchart for FF-MVO oriented aspect based sentiment analysis.

using the movement phase of FF as in Eq. (9). The flowchart for FF-MVO oriented aspect based sentiment analysis is shown in the Figure 2. The pseudo code of the proposed FF-MVO is shown in Algorithm 1.

Algorithm 1: Proposed FF-MVO

```
Begin
Population initialization
Declare the parameters
for every universe  k
for every object  l
    ra2 = random([0,1])
    if  ra2 < WEP
        ra3 = random([0,1])
        ra4 = random([0,1])
        if ra3 < 0.5
            Update the process with the first part of Eq. (3)
        Else
            Update the process with the second part of Eq. (3)
        end if
    Else
        Update the process with the movement phase of FF as in Eq. (9)
    end if
    end for
end for
Stop
```

3.5 *Objective function*

The major objective function of the proposed ABSA for demonetization data is to maximize the accuracy. Here the weight to be added to the polarity score, as well as the hidden neurons of RNN, is optimized with the help of the proposed FF-MVO. The polarity score is measured by the Vader sentiment intensity analyzer in terms of positive score, negative score and neutral score. The polarity score may differ based on the aspect, and hence, it is correlated with the trained data by adding weight. If the weight is added to the polarity score, then the scores of positive, negative, and neutral get altered. On the basis of the dataset, the weights are optimized with the help of the proposed FF-MVO to obtain the weighted polarity score that could generate the maximum accurate result. The combined features with weighted polarity score and Word2vec is applied to the RNN. In the RNN, the hidden neurons are optimized with the same FF-MVO to classify the final sentiment as positive or negative. The proposed objective function can be described as in Eq. (12).

$$Obj\ fun = \underset{\{W_{po}, W_{ne}, W_{nt}, HN_{RNN}\}}{\arg\max} (Accu) \quad (12)$$

In the above equation, W_{po}, W_{ne}, and W_{nt} represents the weight to be added to positive, negative, and neutral polarity scores, HN_{RNN} represents the hidden neurons of the RNN that are to be optimized, and $Accu$ represents the accuracy respectively. Accuracy is described as, "the closeness of a measured value to a standard or known value". Equation (13) describes the mathematical representation of the accuracy.

$$Acc = \frac{TrPo + TrNe}{TrPo + TrNe + FaPo + FaNe} \quad (13)$$

Here, True positive is represented by $TrPo$, True negative is represented by $TrNe$, False positive is represented by $FaPo$, and False-negative is represented by $FaNe$.

4. Pre-processing of demonetization data and aspect extraction

4.1 *Pre-processing*

The pre-processing techniques, such as stop words removal, remove punctuation, lower case conversion, and stemming are performed to eradicate unnecessary information from the tweets. The gaps were reduced by the stemming. The data that we are got from Twitter has a lot of HTML entities. These entities are removed in the processing steps. The HTML parser is used to convert entities to standard HTML tags.The description of these techniques is explained below.

Stop words removal: Stop words are the repeatedly occurring words in a natural language. Examples of stop words include conjunctions, adverbs, prepositions, and articles of the English language.

Remove punctuation: Punctuation marks are used to read the comprehension of a text properly. Punctuation marks have various uses and shapes. Punctuation marks help to understand the correct meaning of the sentence. Removing these punctuation marks helps to minimize the processing time of the system.

Lower case conversion: The lower case letters are the most frequently used ones, and the consumer's power of sight lies mostly with the lowercase letters. Within the same lines, the lowercase letters are mostly viewed than the upper case letters because of their familiar usage.

Stemming: It reduces the meaningless words in the sentence. The initial parts of the word that are mostly found in a sentence are found and then the suffixes are obtained by their difference.

4.2 *Aspect extraction*

Aspect extraction is a basic task of opinion mining. It is defined as the process of identification and extraction. In this process, the aspects are found from the opinionated text. The aspects are of both explicit and implicit. This step extracts the instances of modifiers and product aspects, which describe the opinion regarding a specific aspect. Based on particular syntactic dependency paths, the pairs of words are extracted with the dependency parser tree available in Python's spaCy package [41]. This step pro-

duces a group of nouns, adjectives, verbs, adverbs, and pairs as output for the next step. The usage of a noun is to identify groups of places, people, or things. Adjective represents a word that names an attribute of a noun. Verb reports an occurrence, action, or a state of being. An adverb alters an adjective, verb, preposition, clause, or sentence. The extraction of noun aspects can differentiate the subject and the competitor, and thus the proper sentiment classification is possible. In the review sentence, the syntactic grammatical dependency relation present between words is returned by the dependency parsing with the help of Stanford Parser. The identification of sentiment word, potential noun phrase aspect, and the aspect-sentiment word pairs is exploited with proper dependency relation.

5. Weighted polarity score computation for aspect-based sentiment analysis

5.1 *Measuring polarity score*

The polarity score of the aspect word in terms of noun, adjectives, verb, adverbs, and its pairs is measured with the help of VADER. VADER is a sentiment analysis tool that is based on rules and lexicons. It is particularly familiar to the sentiments that are displayed in social media. It provides many applications in several areas like NY Times editorials, product reviews, social media texts, and movie reviews. It not only provides the negativity and positivity score but also provides how much sentiment is negative or positive. Under the MIT License, VADER is completely open-sourced. Most of the ratings of the VADER are provided by Amazon's Mechanical Turk. When compared with the conventional sentiment analysis, VADER provides a lot of advantages: It derives multiple domains, and it also performs better on social media type text; There is no need for training data, yet it is built from a valence-based, generalizable, and a sentiment lexicon of the human-curated gold standard; For every extracted feature, the VADER returns the positive score, negative score as well as the neutral score. VADER is specifically designed for the given body of the text created in an online media platform. It works efficiently in online media texts. It can find the polarity (detecting the positive and negative of the statement) of a given body of the text. VADER has the highest precision in grey regions of emotion. The grey regions explain the neutral sentiment score.

5.2 *Optimal weighted polarity score-based feature*

Once the polarity scores of each aspect are determined, it is required to provide the weight to each polarity score. It is because of the changes in polarity weight when compared to the trained information. Hence, weighted polarity score is accomplished, in which the weight of each polarity score

Figure 3. Optimal weighted polarity score.

is optimized by the proposed FF-MVO in order to maximize the classification accuracy of sentiment. If the weight is added, then the polarity scores of positive, negative, and neutral changes. On the basis of the dataset, the scores get optimized. This optimal weighted polarity score is developed to attain an accurate score to each aspect based on exact meaning of the review. Therefore, owing to optimization, the polarity is optimized on the basis of class in the corresponding dataset. The solution encoding of the weighted polarity score is shown in figure 3.

The bounding limit of the weight of the polarity score lies in the range of (0–1). From figure 3, the term W_{po}^*, W_{ne}^*, *and* W_{nt}^* denotes the weights to be added to the positive, negative, and neutral polarity scores. Hence, the optimal weighted polarity feature is represented as *FP*.

5.3 *Word2vec-based feature extraction*

To convert the extracted aspect words to the vector, Word2vec is used. Word2vec represents a low dimensional dense vector y_v with the help of every word v in a vocabulary Y that is present in an embedding space $\Re^E$. These word vectors y_v, $\forall v \in Y$ are learned from a training corpus, and so the spatial distance among the words represents the syntactic or semantic similarity among the words. On the basis of the distributional hypothesis, these word representations describe words with similar context containing similar meaning. The target word is predicted from the surrounding context with the help of SGNS. Figure 4 shows the architecture of Skip-gram model.

In the training phase, the Skip-Gram predicts the surrounding words present in a sentence. Let $\{v_1, v_2, \ldots, v_k\}$ represents the sequence of words then the average log
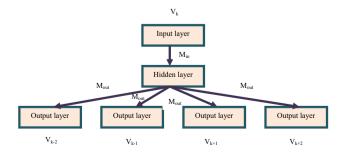

Figure 4. Architecture of skip-gram model.

probability is maximized by the Skip-Gram model as in Eq. (14).

$$I(\Omega) = \frac{1}{K}\sum_{k=1}^{K}\sum_{-d \le i \le d, i \ne 0} \log p\left(v_{k+i}/v_k\right) \tag{14}$$

Here, the training context's size is represented by d, the model parameters to be optimized are represented by Ω, and the probability of seeing word v_{k+i} when the centre word v_k is given is represented by

$$p\left(v_{k+i}/v_k\right).$$

As in figure 4, this probability function is designed as a one hidden layer neural network model. The network is composed of a hidden layer, input layer, and some softmax output layers that correspond to an output word. The input is represented as $v_k \in \Re^Y$, in which the size of the vocabulary is represented by Y that generates a hidden state $g \in \Re^E$, in which the embedding space's dimension or the hidden layer's size is denoted by E, which returns $v_{k+i} \in \Re^Y$ as output. The term N_{out} represents the matrix weight at the output layers, in which various layers are fully connected and shared within all output words. The model parameter is thus represented as $\Omega = \left\{N_{in}^{Y \times E}, N_{out}^{Y \times E}\right\}$. A sparse vector of a $1 - of - Y$ encoding represents an input v_k. The element related to the input word v_k is set to 1 and the remaining components are set to 0. Hence, the softmax function is described by the basic Skip-Gram as in Eq. (15).

$$p(v_O/v_I) = \frac{\exp(\langle y_{in}^{v_I}, y_{out}^{v_O}\rangle)}{\sum_{v=1}^{Y}\exp(\langle y_{in}^{v_I}, y_{out}^{v_O}\rangle)} \tag{15}$$

In the above equation, $y_{out}^{v_O}$ and $y_{in}^{v_I}$ denotes the output and input representations of v, related to the corresponding rows of model parameter matrices N_{out} and N_{in}, and the inner product among two vectors is represented by $\langle .,. \rangle$. The performance of word2vec is enhanced by introducing negative sampling to approximate the softmax log as in Eq. (16).

$$\log p(v_O/v_I) \approx \log \sigma\left(\langle y_{in}^{v_I}, y_{out}^{v_O}\rangle\right) + \sum_{l=1}^{L} \mathrm{E}_{v_l} \approx P_q(v)\left[\log \sigma\left(-\langle y_{in}^{v_I}, y_{out}^{v_I}\rangle\right)\right] \tag{16}$$

The expectations are calculated from a sampling distribution $P_q(v)$, $\forall v \in Y$ and the sigmoid function is represented by $\sigma(y) = \frac{1}{1+\exp(-y)}$. Thus, the features obtained from word2vecor are represented as FW.

5.4 *Combined feature*

The weight of the positive, negative and neutral polarity score is optimized with the proposed FF-MVO to obtain the weighted polarity feature for increasing accuracy. In the other case, using Word2vec, the extracted aspect words are converted to the vector. These two features are joined to generate the combined feature. Equation (17) represents the combined feature.

$$FC_N = FP + FW \tag{17}$$

Here, FC_N represents the combined feature, $N = 1, 2, \dots n$, and n represents the total number of combined features, FP represents the polarity feature, and FW represents the Word2vec feature.

5.5 *Optimized RNN*

RNN [42, 43] takes the combined features as input in order to classify the sentiments of demonization reviews based on aspects. RNN is a division of ANN. In a sequence of data, the directed graph is formed by the connections between the nodes. The temporal characteristics are efficiently modelled for a sequence of time. The output is based on the earlier computations. The sequential progression is modelled by the relationship among the present sequence's output and the earlier one. LSTM is a general form of RNN structure in which the explosion and gradient mass are unsolved problems. LSTM is composed of an input, forget, and output gates that are the three gate units and a memory cell unit. LSTM captures the necessary information and rejects the irrelevant information via the three gates with the help of updating the condition of memory cell. A more general kind of LSTM is the GRU. RNN is made simpler by the GRU by excluding the memory cell from LSTM. Hence, the RNN model can be constructed by the GRU. In RNN, words are given in a high dimensional vector space, and features are extracted and applied to the given neural network. RNN can understand the structure of sentences. These characteristics made RNN efficiently suitable for sentiment analysis.

GRU has joined the output and forget gates into a single update gate a. Through linear interpolation, the single update gate receives the present state of the output. It performs easier training and has fewer parameters. The input features of the earlier hidden state and the jth image slice are represented as (y_j, g_{j-1}). Equations (18) and (19) calculates the reset gate q and the update gate a as given below.

$$a_j = \sigma\left(V^{ya}y_j + V^{ga}g_{j-1}\right) \tag{18}$$

$$q_j = \sigma\left(V^{yq}y_j + V^{gq}g_{j-1}\right) \tag{19}$$

In the above equations, a logistic sigmoid function is represented by σ and the corresponding weight matrices are represented by V^{ya}, V^{ga}, V^{yq}, and V^{gq} respectively. Equation (20) describes the candidate state of the hidden unit.

$$\tilde{g}_j = \tan\left(V^{yg}y_j + V^{gg}\left(g_{j-1}\Theta q_j\right)\right) \tag{20}$$

Figure 5. Optimized RNN.

Here, an element-wise multiplication is represented by Θ. When q_j is near to 0, the reset gate q forgets the earlier computed state and reads the initial symbol of an input sequence. Eq. (21) describes the linear interpolation of jth hidden activation state g_j of GRU among the candidate state $\tilde{g}_j$ and the earlier state g_{j-1}.

$$g_j = (1 - a_j)\Theta\tilde{g}_j + a_j\Theta g_{j-1} g_j = (1 - a_j)\Theta\tilde{g}_j + a_j\Theta g_{j-1}$$
(21)

As an improvement to the RNN, The hidden neurons of the RNN are optimized with the proposed FF-MVO to classify the final sentiments as positive or negative with high accuracy. Figure 5 shows the solution encoding of optimized RNN.

The bounding limits of the hidden neurons of RNN lie in between the value of (5-35). Here, HN_{RNN} represents the hidden neurons of the RNN that are being optimized with the proposed FF-MVO to classify the final sentiments as positive or negative.

6. Results and discussions

6.1 *Experimental procedure*

The proposed ABSA for demonetization data was implemented in Python and the analysis was carried out. The population size of the proposed FF-MVO was taken as 10 and the maximum number of iterations performed was 25. The performance of the proposed FF-MVO-RNN was compared with polarity feature over-weighted feature and several existing optimization algorithms like PSO-RNN [44], GWO-RNN [45], MVO-RNN [37] and FF-RNN [38] and the results were computed. Additionally, the performance of the proposed FF-MVO-RNN was compared with various traditional machine learning algorithms like DT [46], NB [47], KNN [48], SVM [49], NN [50], and RNN [51] and the results were computed. Moreover, the performance of the proposed FF-MVO-RNN was analyzed in terms of Accuracy, Sensitivity, Specificity, Precision, FPR, FNR, FDR, NPV, F1 Score and MCCto confirm the performance of the proposed FF-MVO.
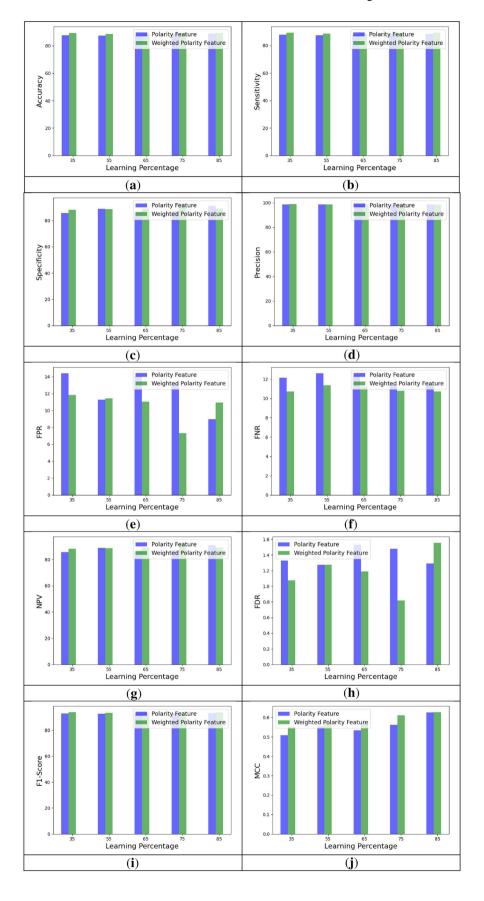
6.2 *Tweet collection*

The demonetization tweets dataset was collected from Kaggle [52]. It is the highest community of data science containing powerful resources and tools. It is a predictive modelling platform and is a repository containing a large amount of data. The collected data includes a tweet id, the tweet text, and the sentiment label. The information regarding the tweets, such as creator name, creation date, etc. is collected for Twitter using the python libraries. All Tweets are available on a public Twitter profile. There are no barriers to stop collecting the data from Twitter.This dataset involves 14940 data tweets related to demonization. The parameters involved in the dataset are serial number, text or review, favorited (true or false), favourite count (0 or 1), reply to SN, created on which date, truncated (true or false), reply to SID, id, reply to UID, status source, screen name, retweet count, is retweet (true of false), retweeted (true or false). There is no particular base work for this research done on the same dataset, and the dataset is available publically in Kaggle source. The tweets are monolingual (English) language.

6.3 *Effect of weighted polarity feature*

The comparative analysis of the polarity feature over-weighted feature for ABSA using demonetization dataset with respect to various performance measures is displayed in figure 6. It can be seen that when the weight is added to the polarity score, the polarity score of the positive, negative, and neutral gets changed. Depending upon the dataset, the scores get optimized. But, in few cases, the dataset contains some positive opinion score, but on the basis of this weight, it will be negative. So, the polarity can be optimized on the basis of the class in the corresponding dataset. From figure 6a, at learning percentage = 75%, the accuracy of the weighted polarity feature is 2.38% advanced than the polarity feature without weight. On considering figure 6b, at learning percentage =85%, the sensitivity of the weighted polarity feature is 1.19% better than the polarity feature without weight. In figure 6c at learning percentage = 75%, the specificity of the weighted polarity feature is 7.32% improved than the polarity feature without weight. While considering figure 6d, at learning percentage =75%, the precision of the weighted polarity feature is 2.04% progressed than the polarity feature without weight. From figure 6e, at learning percentage = 85%, the FPR of the weighted polarity feature is 22.22% superior to polarity feature without weight. In figure 6f, at learning percentage = 85%, the FNR of the weighted polarity feature is 8.70% upgraded than the polarity feature without weight. On considering figure 6g, at learning percentage = 85%, the NPV of the weighted polarity feature is 2.27% advanced than the polarity feature without weight. From figure 6h, at learning percentage = 85%, the FDR of the weighted polarity feature is 21.54% better than the polarity feature without weight. While considering figure 6i, at learning percentage = 85%, the F1-score of the weighted polarity feature is 1.16% improved than the polarity feature without weight. In figure 6j, at learning

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

◄**Figure 6.** Comparative analysis of polarity feature over weighted feature for aspect-based sentiment analysis using demonetization dataset in terms of (**a**) Accuracy, (**b**) Sensitivity, (**c**) Specificity, (**d**) Precision, (**e**) FPR, (**f**) FNR, (**g**) NPV, (**h**) FDR, (**i**) F1-score and (**j**) MCC.
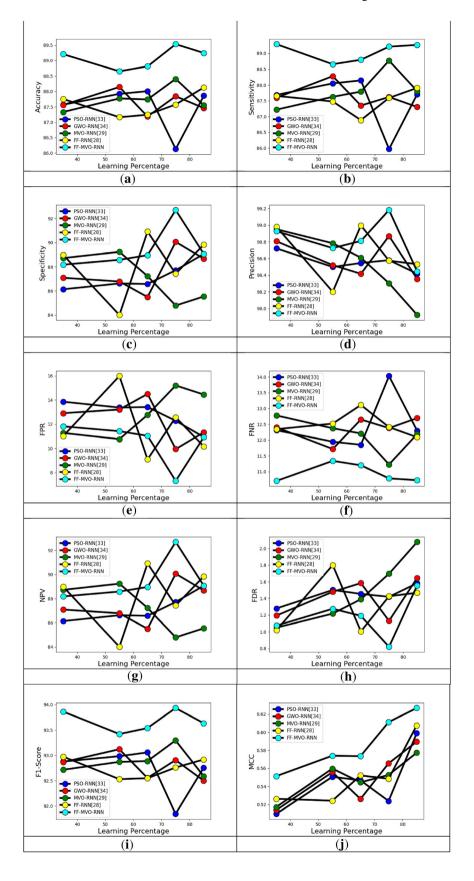
percentage = 75%, the MCC of the weighted polarity feature is 9.09% progressed than the polarity feature without weight. Therefore, it can be concluded that the weighted polarity feature performs better than the polarity feature without weight for ABSA using the demonetization dataset based on the proposed FF-MVO-RNN.

6.4 *Analysis on optimized RNN*

The performance analysis of the proposed and existing heuristic-based RNN for ABSA for demonetization dataset is displayed in figure 7. In the case of proposed FF-MVO-RNN, the positive measures such as accuracy, sensitivity, specificity, precision, NPV, F1 Score and MCC show an increment and negative measures such as FPR, FNR and FDR shows a decrement with respect to the existing algorithms which shows the betterness of the proposed FF-MVO-RNN. From figure 7a, at learning percentage = 85%, the accuracy of the proposed FF-MVO-RNN is 1.71% better than PSO-RNN, 2.17% better than GWO-RNN, 2.06% better than MVO-RNN and 1.36% better than FF-RNN. In figure 7b, at learning percentage = 75%, the sensitivity of the proposed FF-MVO-RNN is 3.96% improved than PSO-RNN, 1.94% improved than GWO-RNN, 0.68% improved than MVO-RNN and 2.06% improved than FF- RNN. On considering figure 7c, at learning percentage =75%, the specificity of the proposed FF-MVO-RNN is 6.18% progressed than PSO-RNN, 3.34% progressed than GWO-RNN, 9.69% progressed than MVO-RNN and 6.30% progressed than FF- RNN. In figure 7d, at learning percentage = 75%, the precision of the proposed FF-MVO-RNN is 1.64% superior to PSO-RNN, 0.38% V GWO-RNN, 0.92% superior to MVO-RNN and 0.62% superior to FF-RNN. While considering figure 7e, at learning percentage = 75%, the FPR of the proposed FF-MVO-RNN is 41.46% upgraded than PSO-RNN, 26.53% upgraded than GWO-RNN, 52% upgraded than MVO-RNN and 41.94% upgraded than FF-RNN. From figure 7f, at learning percentage = 85%, the FNR of the proposed FF-MVO-RNN is 13% advanced than PSO-RNN, 15.75% advanced than GWO-RNN, 12.30% advanced than MVO-RNN and 11.57% advanced than FF-RNN. While considering figure 7g, at learning percentage = 75%, the NPV of the proposed FF-MVO-RNN is 6.42% better than PSO-RNN, 3.11% better than GWO-RNN, 9.69% better than MVO-RNN and 6.67% better than FF- RNN. From figure 7h, at learning percentage = 75%, the FDR of the

proposed FF-MVO-RNN is 63.64% improved than PSO-RNN, 27.27% v GWO-RNN, 51.22% improved than MVO-RNN and 42.85% improved than FF- RNN. In figure 7i, at learning percentage =85%, the F1 score of the proposed FF-MVO-RNN is 1.08% progressed than PSO-RNN, 1.30% progressed than GWO-RNN, 1.19% progressed than MVO-RNN and 0.86% progressed than FF-RNN. From figure 7j, at learning percentage = 85%, the MCC of the proposed FF-MVO-RNN is 5.70% superior to PSO-RNN, 8.24% superior to GWO-RNN, 10.53% superior to MVO-RNN and 4.30% superior to FF-RNN. Hence, it can be concluded that the proposed FF-MVO performs better performance analysis in terms of various performance measures when it is compared with several optimization-based RNN for ABSA using demonetization dataset.

6.5 *Machine learning models for ABSA*

The comparative analysis of the proposed and existing machine learning models for ABSA using demonetization dataset is shown in figure 8. While considering figure 8a, at learning percentage = 75%, the accuracy of the proposed FF-MVO-RNN is 3.45% improved than DT, 7.14% improved than NB, 4.65% improved than KNN, 3.45% improved than SVM, 4.65% improved than NN, and 2.27% improved than RNN. From figure 8b, at learning percentage = 75%, the sensitivity of the proposed FF-MVO-RNN is 1.19% progressed than DT, 3.66% progressed than NB, 2.41% progressed than KNN, 2.41% progressed than SVM, 3.66% progressed than NN and 1.19% progressed than RNN. On considering figURE 8c, at learning percentage = 75%, the specificity of the proposed FF-MVO-RNN is 3.53% superior to DT, 6.02% superior to NB, 3.53% superior to KNN, 4.76% superior to SVM, 2.33% superior to NN, and 2.33% superior to RNN. In figure 8d, at learning percentage = 75%, the precision of the proposed FF-MVO-RNN is 1.03% upgraded than DT, 2.08% upgraded than NB, 1.03% upgraded than KNN, 1.03% upgraded than SVM, 1.03% upgraded than NN, and 1.03% upgraded than RNN. While considering figure 8e, at learning percentage = 75%, the FPR of the proposed FF-MVO-RNN is 48.15% advanced than DT, 57.58% advanced than NB, 44% advanced than KNN, 48.15% advanced than SVM, 30% advanced than NN, and 30% advanced than RNN. From figure 8f, at learning percentage = 85%, the FNR of the proposed FF-MVO-RNN is 20.29% better than DT, 21.43% better than NB, 26.67% better than KNN, 21.43% better than SVM, 15.38% better than NN and 8.33% better than RNN. In figure 8g, at learning percentage = 75%, the NPV of the proposed FF-MVO-RNN is 3.45% improved than DT, 5.88% improved than NB, 3.45% improved than KNN, 4.65% improved than SVM, 2.27% improved than NN and 2.27% improved than RNN. On considering figure 8h, at learning percentage = 85%, the FDR of the proposed FF-MVO-RNN is 34.78% progressed than DT, 6.25%

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

◄**Figure 7.** Performance analysis of proposed and existing heuristic-based RNN for aspect-based sentiment analysis for demonetization dataset in terms of (**a**) Accuracy, (**b**) Sensitivity, (**c**) Specificity, (**d**) Precision, (**e**) FPR, (**f**) FNR, (**g**) NPV, (**h**) FDR, (**i**) F1-score and (**j**) MCC.
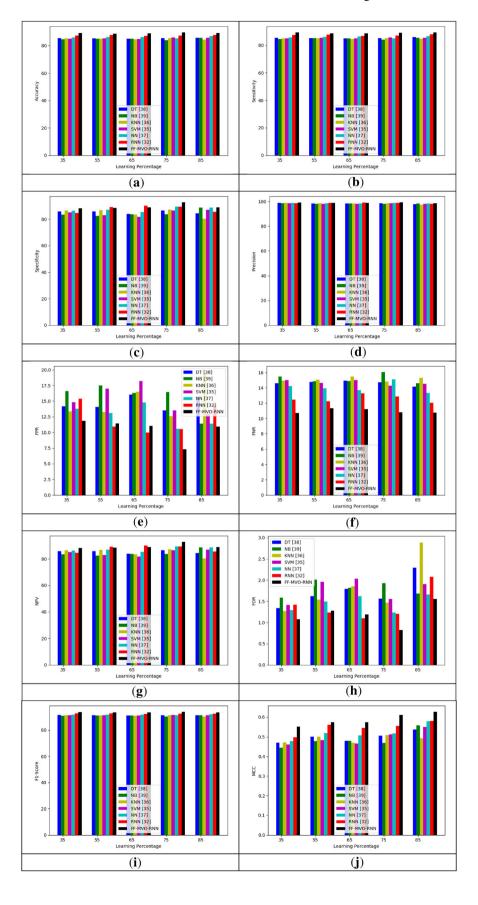
progressed than NB, 46.43% progressed than KNN, 16.67% progressed than SVM, 6.25% progressed than NN and 25% progressed than RNN. From figure 8i, at learning percentage = 75%, the F1 score of the proposed FF-MVO-RNN is 2.27% superior to DT, 3.45% superior to NB, 2.27% superior to KNN, 2.27% superior to SVM, 2.27% superior to NN and 1.12% superior to RNN. While considering figure 8j, at learning percentage = 85%, the MCC of the proposed FF-MVO-RNN is 20% surpassed than DT, 15.38% surpassed than NB, 33.33% surpassed than KNN, 20% surpassed than SVM, 9.09% surpassed than NN and 9.09% surpassed than RNN. FF-MVO-RNN has high performance when compared to the other classifiers. The reasons for the high performance of the proposed classifier is that it uses the RNN, which posses several merits. RNN remembers every piece of information through time. RNN is used to find optimal solutions.Also, it has the advantage of objective function value, image quality measurement, convergence performance, and robustness.RNN can remember the past data for the sentiment analysis. Also, the RNN is trained by the proposed FF-MVO, which improves the classification performance of the RNN. Hence, It is concluded that the proposed FF-MVO-RNN performs better comparative analysis with respect to various performance measures when it is compared with traditional machine learning algorithms.

6.6 *Overall performance analysis*

The overall performance analysis of ABSA for demonetization dataset in terms of various performance measures are listed in tables 2, 3, and 4, respectively. From table 2, the accuracy of the proposed FF-MVO-RNN is 1.57% advanced than PSO-RNN, 2.04% advanced than GWO-RNN, 1.94% advanced than MVO-RNN, 1.27% advanced than FF-RNN, 4.17% advanced than DT, 4.07% advanced than NB, 6% advanced than KNN, 4.17% advanced than SVM, 2.68% advanced than NN and 1.78% advanced than RNN. The sensitivity of the proposed FF-MVO-RNN is 1.78% better than PSO-RNN, 2.25% better than GWO-RNN, 1.66% better than MVO-RNN, 1.55% better than FF-RNN, 3.99% better than DT, 4.54% better than NB, 5.41% better than KNN, 4.42% better than SVM, 2.96% better than NN and 1.49% better than RNN. Moreover, the specificity of the proposed FF-MVO-RNN is 0.01% improved than PSO-RNN, 0.44% improved than GWO-RNN, 4.11% improved than MVO-RNN, 0.87% improved than FF-RNN, 5.63% improved than DT, 0.49% improved

than NB, 10.79% improved than KNN, 2.30% improved than SVM, 0.49% improved than NN and 4.11% improved than RNN. On considering table 2, the FPR of the proposed FF-MVO-RNN is 0.01% progressed than PSO-RNN, 3.45% progressed than FF-RNN, 24.32% progressed than MVO-RNN, 7.69% progressed than FF-RNN, 30.27% progressed than DT, 3.83% progressed than NB, 44.22% progressed than KNN, 15.48% progressed than SVM, 3.83% progressed than NN and 24.32% progressed than RNN. The FNR of the proposed FF-MVO-RNN is 12.70% superior to PSO-RNN, 15.48% superior to GWO-RNN, 11.98% superior to MVO-RNN, 11.25% superior to FF-RNN, 24.20% superior to DT, 26.55% superior to NB, 29.94% superior to KNN, 26.04% superior to SVM, 19.32% superior to NN, and 10.88% superior to RNN. Additionally, the FDR of the proposed FF-MVO-RNN is 1.72% upgraded than PSO-RNN, 5.49% upgraded than GWO-RNN, 25.16% upgraded than MVO-RNN, 5.96% upgraded than FF-RNN, 32.18% upgraded than DT, 7.52% upgraded than NB, 46.15% upgraded than KNN, 18.45% upgraded than SVM, 6.13% upgraded than NN and 25.04% upgraded than RNN. In table 4, the precision of the proposed FF-MVO-RNN is 0.03% surpassed than PSO-RNN, 0.09% surpassed than GWO-RNN, 0.53% surpassed than MVO-RNN, 0.09% surpassed than FF-RNN, 0.76% surpassed than DT, 0.13% surpassed than NB, 1.37% surpassed than KNN, 0.36% surpassed than SVM, 0.10% surpassed than NN and 0.53% surpassed than RNN. The NPV of the proposed FF-MVO-RNN is 0.01% advanced than PSO-RNN, 0.44% advanced than GWO-RNN, 4.11% advanced than MVO-RNN, 0.87% advanced than FF-RNN, 5.63% advanced than DT, 0.49% advanced than NB, 10.79% advanced than KNN, 2.30% advanced than SVM, 0.49% advanced than NN, and 4.11% advanced than RNN. The F1-Score of the proposed FF-MVO-RNN is 0.95% better than PSO-RNN, 1.22% better than GWO-RNN, 1.13% better than MVO-RNN, 0.77% better than FF-RNN, 2.45% better than DT, 2.44% better than NB, 3.49% better than KNN, 2.49% better than SVM, 1.60% better than NN and 1.03% better than RNN. Moreover, the MCC of the proposed FF-MVO-RNN is 4.65% improved than PSO-RNN, 6.31% improved than GWO-RNN, 8.57% improved than MVO-RNN, 3.16% improved than FF-RNN, 16.78% improved than DT, 12.25% improved than NB, 27.06% improved than KNN, 14.05% improved than SVM, 8.29% improved than NN and 8.09% improved than RNN. Hence, it can be concluded that the proposed FF-MVO-RNN performs better performance analysis for various performance measures when it is compared with the traditional machine learning algorithms.

6.7 *Analysis based on cross-fold validation*

Cross-validation is a technique that makes near-optimal use of existing data by frequent pre-processing classifiers on

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)

(j)

◄**Figure 8.** Comparative analysis of proposed and existing machine learning models for aspect-based sentiment analysis for demonetization dataset in terms of (**a**) Accuracy, (**b**) Sensitivity, (**c**) Specificity, (**d**) Precision, (**e**) FPR, (**f**) FNR, (**g**) NPV, (**h**) FDR, (**i**) F1-score and (**j**) MCC.

various subsets of data, generally with huge training and a little validation in each iterative process. For example, 5 fold-cross-partitions uses 50% of the data for training. This process happens 5 times till all the data give validation data once. The accuracy of comparative methods based on

Table 2. Overall performance analysis of aspect-based sentiment analysis for demonetization dataset in terms of accuracy, sensitivity, and specificity over conventional models.

Methods	Accuracy	Sensitivity	Specificity
PSO-RNN [42]	0.8786256135653726	0.8770780856423174	0.890625
GWO-RNN [43]	0.8746095493083444	0.8730478589420655	0.88671875
MVO-RNN [35]	0.8755020080321285	0.8780856423173804	0.85546875
FF-RNN [36]	0.8813029897367247	0.8790931989924433	0.8984375
DT [44]	0.8566964285714286	0.8584382871536524	0.8431372549019608
NB [45]	0.8575892857142857	0.853904282115869	0.8862745098039215
KNN [46]	0.8419642857142857	0.8468513853904283	0.803921568627451
SVM [47]	0.8566964285714286	0.854911838790932	0.8705882352941177
NN [48]	0.8691964285714285	0.8670025188916877	0.8862745098039215
RNN [49]	0.8768406961178046	0.8795969773299748	0.85546875
FF-MVO-RNN	0.892458723784025	0.8926952141057934	0.890625

Table 3. Overall performance analysis of aspect-based sentiment analysis for demonetization dataset in terms of FPR, FNR, and FDR over conventional methods.

Methods	FPR	FNR	FDR
PSO-RNN [42]	0.109375	0.12292191435768261	0.01582815149802148
GWO-RNN [43]	0.11328125	0.1269521410579345	0.016458569807037457
MVO-RNN [35]	0.14453125	0.12191435768261964	0.020786516853932586
FF-RNN [36]	0.1015625	0.12090680100755667	0.014680971202710334
DT [44]	0.1568627450980392	0.1415617128463476	0.0229357779816513763
NB [45]	0.113372549019607843	0.14609571788413098	0.016821345707656612
KNN [46]	0.19607843137254902	0.1531486146095718	0.028885037550548817
SVM [47]	0.12941176470588237	0.145088161209068	0.01907514450867052
NN [48]	0.113372549019607843	0.13299748110831233	0.01657142857142857
RNN [49]	0.14453125	0.12040302267002519	0.02075154234436343
FF-MVO-RNN	0.109375	0.10730478589420656	0.015555555555555555555

Table 4. Overall performance analysis of aspect-based sentiment analysis for demonetization dataset in terms of precision, NPV, F1-Score, and MCC over conventional methods.

Methods	Precision	NPV	F1-Score	MCC
PSO-RNN [42]	0.9841718485019785	0.890625	0.9275439531166756	0.5989059236393571
GWO-RNN [43]	0.9835414301929626	0.88671875	0.9250066720042701	0.5895351880043516
MVO-RNN [35]	0.9792134831460674	0.85546875	0.9258964143426295	0.5772608389788214
FF-RNN [36]	0.9853190287972897	0.8984375	0.9291799787007454	0.6075185927313828
DT [44]	0.9770642201834863	0.8431372549019608	0.913917940466613	0.5366741878037092
NB [45]	0.9831786542923434	0.8862745098039215	0.9139929900242653	0.558333262376804
KNN [46]	0.9711149624494512	0.803921568627451	0.9047362755651238	0.49325604471723844
SVM [47]	0.9809248554913295	0.8705882352941177	0.9135935397039031	0.549515146149316
NN [48]	0.9834285714285714	0.8862745098039215	0.9215528781793842	0.5787460036661942
RNN [49]	0.9792484576556366	0.85546875	0.9267515923566879	0.5798531245663597
FF-MVO-RNN	0.9844444444444445	0.890625	0.936327608982827	0.626735110148875

Table 5. Accuracy of comparative methods based on cross-fold validation.

Methods	PSO	GWO	WOA	SLNO	Proposed
k = 1	0.8725678653726	0.8770780850023174	0.872525	0.880302889736	0.890625
k =2	0.8736095493083444	0.8754078589420655	0.88671875	0.871322889736	0.880625
k = 3	0.8755021080321285	0.8780856423173804	0.87546875	0.8892134831460674	0.8915852296
k = 4	0.8713029897367247	0.8790931989924433	0.884375	0.883778786190287	0.8955226788296
k = 5	0.8723789890	0.87267872141057934	0.880625	0.8844444444444445	0.897989

Table 6. Error analysis.

Methods	Error analysis
PSO-RNN [42]	0.1213743864346274
GWO-RNN [43]	0.1253904506916556
MVO-RNN [35]	0.1244979919678715
FF-RNN [36]	0.1186970102632753
DT [44]	0.1433035714285714
NB [45]	0.1424107142857143
KNN [46]	0.1580357142857143
SVM [47]	0.1433035714285714
NN [48]	0.1308035714285715
RNN [49]	0.1231593038821954
FF-MVO-RNN	0.107541276215975

cross-fold validation is shown in table 5. For k = 1, the accuracy of the proposed method is 2.02% better than the PSO, 1.52% better than the GWO, 2.03% better than the WOA, and 1.15% better than the SLNO. Similarly, for k = 5, the accuracy of the proposed method is 2.85% better than the PSO, 2.82% better than the GWO, 1.93% better than the WOA, and 1.51% better than the SLNO.

6.8 *Error analysis*

The error analysis for the proposed method is shown in table 6. The error value of the classifiers, such as PSO-RNN, GWO-RNN, MVO-RNN, FF-RNN, DT, NB, KNN, SVM, NN, and RNN is 0.1214, 0.1254, 0.1245, 0.1187, 0.1433, 0.1424, 0.1580, 0.1433, 0.1308, 0.1236, respectively. On the otherhand, the proposed FF-MVO-RNN has the minimum error value of 0.1075, which shows the effectiveness of the proposed classifier in sentiment analysis.

7. Conclusion

This paper has performed the ABSA for demonetization tweets by adopting the optimized deep learning concept. The various demonetization tweets were gathered from the Kaggle dataset. Pre-processing was done by means of four phases such as stop words removal, remove punctuation, lower case conversion, and stemming for minimizing the data to its reduced format. The pre-processed data has undergone aspect extraction and the extracted aspect words

were converted to the features using Word2vec and polarity score computation. The weight of the polarity scores was optimized with the proposed FF-MVO for generating the weighted polarity score. The combined features were applied to a deep learning algorithm known as RNN. An enhancement was made by optimizing the hidden neurons with the same FF-MVO and therefore the positive and negative sentiments were classified. At last, the comparative analysis over various machine learning algorithms proved the competent performance of the proposed model. From the analysis, the accuracy of the proposed FF-MVO-RNN was 1.57% better than PSO-RNN, 2.04% better than GWO-RNN, 1.94% better than MVO-RNN, 1.27% better than FF-RNN, 4.17% better than DT, 4.07% better than NB, 6% better than KNN, 4.17% better than SVM, 2.68% better than NN and 1.78% better than RNN, which proved the better performance of the proposed method. The proposed FF-MVO has some limitations, such as (a) When addressing the real-world optimization problem, the MVO may need modification and changes. (b) It is quite difficult to manage the difficulties in the multi-modal search procedure. These limitations will be considered in future work, and a more effective algorithm will be developed. Also, the experimentation will be carried out on large data to analyze the performance of the proposed work.

Abbreviations

ABSA	Aspect-based Sentiment Analysis
KNN	K-Nearest Neighbours
RBI	Reserve Bank of India
FPR	False Positive Rate
FF	Fire Fly Algorithm
NB	Naïve Bayes
FEA-NN	Feature Enhanced Attention-Neural Network
MVO	Multi-Verse Optimization
CNN	Convolutional Neural Network
GWO	Grey Wolf Optimization
RNN	Recurrent Neural Network
FDR	False Detection Rate
IGCN	Interactive Gated Convolutional Network
FF-MVO	Fire Fly-oriented Multi-Verse Optimizer
MCC	Matthews Correlation Coefficient
ALDONAr	A Lexicalized Domain Ontology and a Regularized Neural Attention model

PSO	Particle Swarm Optimization
FNR	False Negative Rate
SVM	Support Vector Machine
ReMemNN	Recurrent Memory Neural Network
NPV	Net Present Value
DNet	Distillation Network
BERT	Bidirectional Encoder Representations from Transformers
NIFTY	National Stock Exchange Fifty
DT	Decision Tree
BSE	Bombay Stock Exchange Stock Exchange
SENSEX	Sensitive Index
GDP	Gross Domestic Product
ATSA	Aspect-Term Sentiment Analysis
NLP	Natural Language Processing
NN	Neural Network
GRU	Gate Recurrent Unit
LSTM	Long Short-Term Memory
ACSA	Aspect-Category Sentiment Analysis
TDR	Travelling Distance Rate
VADER	Valence Aware Dictionary and sEntiment Reasoner
WEP	Wormhole Existence Probability
SGNS	Skip-Gram with Negative Sampling
ANN	Artificial Neural Network

References

[1] Liu B 2012 Sentiment analysis and opinion mining; *Synthesis Lect. Human Lang. Technol.* 5: 1–167

[2] Shams M and Baraani-Dastjerdi A 2017 Enriched LDA (ELDA): Combination of latent Dirichlet allocation with word co-occurrence analysis for aspect extraction; *Expert Syst. Appl.* 80: 136–146

[3] Yang C, Zhang H, Jiang B and Li K 2019 Aspect-based sentiment analysis with alternating coattention networks; *Commun. Comput. Inf. Sci.* 56: 463–478

[4] Schouten K and Frasincar F 2016 Survey on aspect-level sentiment analysis; *IEEE Trans. Knowl. Data Eng.* 28: 813–830

[5] Subrahmanian V S and Reforgiato D 2008 AVA: Adjective-verb-adverb combinations for sentiment analysis; *IEEE Intell. Syst.* 23: 43–50

[6] Che W, Zhao Y, Guo H, Su Z and Liu T 2015 Sentence compression for aspect-based sentiment analysis; *IEEE/ACM Trans. Audio Speech Lang. Process.* 23: 2111–2124

[7] Wang W, Pan SJ, Dahlmeier D and Xiao X 2016 Recursive neural conditional random fields for aspect-based sentiment analysis. In: *Proceedings of EMNLP*, pp. 616–626

[8] Tang D, Qin B and Liu T 2016 Aspect level sentiment classification with deep memory network. In: *Proceedings of EMNLP*, pp. 214–224

[9] Alqaryouti O, Siyam N and Shaalan K 2019 A sentiment analysis lexical resource and dataset for government smart apps domain. In: *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics on Advanced Intelligent Systems and Informatics 2018 (AISI2018)*, Springer, Cham

[10] Chen P, Sun Z, Bing L and Yang W 2017 Recurrent attention network on memory for aspect sentiment analysis. In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*

[11] Poria S, Cambria E, Winterstein G and Huang G-B 2014 Sentic patterns: Dependency-based rules for concept-level sentiment analysis; *Knowl. Based Syst.* 69: 45–63

[12] Singh P, Sawhney R S and Kahlon K S 2018 Sentiment analysis of demonetization of 500 and 1000 rupee banknotes by Indian government; *ICT Express* 4: 124–129

[13] Dhanya N M and Harish U C 2018 Sentiment analysis of twitter data on demonetization using machine learning techniques. In: *Computational Vision and Bio Inspired Computing*, pp. 227–237

[14] Srinivas T A S and Manivannan S S 2020 Prevention of hello flood attack in IoT using combination of deep learning with improved rider optimization algorithm; *Comput. Commun.* 163: 162–175

[15] Sumit H, Sailee B and Abhishek B 2021 Rapid digitization of healthcare: a review of COVID-19 impact on our health systems; *Int. J. Res. Educ. Sci. Methods* 9(2): 1457–1459

[16] Asghar M Z, Khan A, Ahmad S and Kund I F M 2014 A review of feature extraction in sentiment analysis; *J. Basic Appl. Sci. Res.* 4: 181–186

[17] Fu X, Liu W, Xu Y and Cui L 2017 Combine HowNet lexicon to train phrase recursive autoencoder for sentence-level sentiment analysis; *Neurocomputing* 241: 18–27

[18] Akhtar M S, Gupta D, Ekbal A and Bhattacharyya P 2017 Feature selection and ensemble construction: a two-step method for aspect based sentiment analysis; *Knowl. Based Syst.* 125: 116–135

[19] Cui Z, Shi X and Chen Y 2016 Sentiment analysis via integrating distributed representations of variable-length word sequence; *Neurocomputing* 187: 126–132

[20] Cambria E 2016 Affective computing and sentiment analysis; *IEEE Intell. Syst.* 31: 102–107

[21] Chen P, Sun Z, Bing L and Yang W 2017 Recurrent attention network on memory for aspect sentiment analysis. In: *Proceedings of the 2017 conference on empirical methods in natural language processing*, pp. 452–461

[22] Denecke K 2008 Using SentiWordNet for multilingual sentiment analysis. In: *Proceedings of the IEEE 24th International Conference on Data Engineering Workshop*, pp. 507–512

[23] Mishra P, Rajnish R and Kumar P 2016 Sentiment analysis of Twitter data: case study on digital India. In: *Proceedings of the 2016 International Conference on Information Technology (InCITe)—The Next Generation IT Summit*, pp. 148–153

[24] Pong-Inwong C and Songpan W 2018 Sentiment analysis in teaching evaluations using sentiment phrase pattern matching (SPPM) based on association mining; *Int J Mach Learn Cybern* 10: 2177–2186

[25] Meng W, Wei Y, Liu P, Zhu Z and Yin H 2019 Aspect based sentiment analysis with feature enhanced attention CNN-BiLSTM; *IEEE Access* 7: 167240–167249

[26] Rida-E-Fatima S, Javed A, Banjar A, Irtaza A, Dawood H, Dawood H and Alamri A 2019 A multi-layer dual attention

deep learning model with refined word embeddings for aspect-based sentiment analysis; *IEEE Access* 7: 114795–114807

[27] Alqaryouti O, Siyam N, Monem, A A and Shaalan K 2019 Aspect-based sentiment analysis using smart government review data. *Appl. Comput. Inform.*

[28] Shams M, Khoshavi N and Baraani-Dastjerdi A 2020 LISA: language-independent method for aspect-based sentiment analysis; *IEEE Access* 8: 31034–31044

[29] Kumar A, Narapareddy V T, SrikanthV A, Neti L B M and Malapati A 2020 Aspect-based sentiment classification using interactive gated convolutional network; *IEEE Access* 8: 22445–22453

[30] Meskele D and Frasincar F 2020 ALDONAr: a hybrid solution for sentence-level aspect-based sentiment analysis using a lexicalized domain ontology and a regularized neural attention model; *Commun. Comput. Inf. Sci.* 57: 102211

[31] Liu N and Shen B 2020 ReMemNN: A novel memory neural network for powerful interaction in aspect-based sentiment analysis. *Neurocomputing*

[32] Ren F, Feng L, Xiao D, Cai M and Cheng S 2020 DNet: a lightweight and efficient model for aspect based sentiment analysis. *Expert Syst. Appl.*

[33] Damarta R, Hidayat A and Abdullah A S 2021 The application of k-nearest neighbors classifier for sentiment analysis of PT PLN (Persero) twitter account service quality; *J. Phys. Conf. Ser.* 1722: 1–7

[34] Lymperopoulos I N 2021 RC-Tweet: modeling and predicting the popularity of tweets through the dynamics of a capacitor; *Expert Syst. Appl.* 163: 1–23

[35] Wei H, Sankaranarayanan J and Same T H 2020 Enhancing local live tweet stream to detect news. *GeoInformatica*, pp. 1–31

[36] Girolamo R D, Esposito C, Moscat V and Sperli G 2021 Evolutionary game theoretical on-line event detection over tweet streams; *Knowl. Based Syst.* 211: 106563

[37] Mirjalili S, Mirjalili S M and Hatamlou A 2016 Multi-verse optimizer: a nature-inspired algorithm for global optimization; *Neural Comput. Appl.* 27: 495–513

[38] Gandomi A H, Yang X-S, Talatahari S and Alavi A H 2013 Firefly algorithm with chaos; *Commun. Nonlinear Sci. Numer. Simul.* 18: 89–98

[39] George A and Rajakumar B R 2013 On hybridizing fuzzy min max neural network and firefly algorithm for automated heart disease diagnosis. In: *Fourth International Conference on Computing, Communications and Networking Technologies*, Tiruchengode, India

[40] Beno M M, Valarmathi I R, Swamy S M and Rajakumar B R 2014 Threshold prediction for segmenting tumour from brain MRI scans; *Int. J. Imaging Syst. Technol.* 24: 129–137

[41] Spacy, Industrial-Strength NLP in Python, spacy.io/

[42] Nguyen T H and Shirai K 2015 Phrase RNN: phrase recursive neural network for aspect-based sentiment analysis, pp. 2509–2514

[43] Basiri M E, Nemati S, Abdar M, Cambria E and Acharya U R 2021 ABCDM: an attention-based bidirectional CNN-RNN deep model for sentiment analysis; *Future Gener. Comput. Syst.* 115: 279–294

[44] Pedersen M E H and Chipperfield A J 2010 Simplifying particle swarm optimization; *Appl. Soft Comput.* 10: 618–628

[45] Mirjalili S, Mirjalili S M and Lewis A 2014 Grey wolf optimizer; *Adv. Eng. Softw.* 69: 46–61

[46] Tsang S, Kao B, Yip K Y, Ho W-S and Lee S D 2011 Decision trees for uncertain data; *IEEE Trans. Knowl. Data Eng.* 23: 64–78

[47] Chen S, Webb G I, Liu L and Ma X 2019 A novel selective naïve Bayes algorithm. *Knowl. Based Syst.*

[48] Huang M, Lin R, Huang S and Xing T 2017 A novel approach for precipitation forecast via improved K-nearest neighbor algorithm; *Adv. Eng. Inform.* 33: 89–95

[49] Borkar G M, Patil L H, Dalgade D and Hutke A 2019 A novel clustering approach and adaptive SVM classifier for intrusion detection in WSN: a data mining concept; *Sustain. Comput. Inform. Syst.* 23: 120–135

[50] Beck G, Duong T, Lebbah M, Azzag H and Cerin C 2019 A distributed approximate nearest neighbors algorithm for efficient large scale mean shift clustering; *J. Parallel Distrib. Comput.* 134: 128–139

[51] Li F and Liu M 2019 A hybrid convolutional and recurrent neural network for hippocampus analysis in Alzheimer's disease; *J. Neurosci. Methods* 323: 108–118

[52] Kaggle dataset is available at https://www.kaggle.com/abhimicro3/demonetization-twitter-sentiment-analysis