

## Intelligent Machine Learning with Metaheuristics Based Sentiment Analysis and Classification

R. Bhaskaran<sup>1,\*</sup>, S. Saravanan<sup>1</sup>, M. Kavitha<sup>2</sup>, C. Jeyalakshmi<sup>3</sup>, Seifedine Kadry<sup>4</sup>, Hafiz Tayyab Rauf<sup>5</sup> and Reem Alkhamash<sup>6</sup>

<sup>1</sup>Department of Information Technology, PSNA College of Engineering and Technology, Dindigul, 624622, India

<sup>2</sup>Department of Electronics and Communication Engineering, K. Ramakrishnan College of Technology, Trichy, 621112, India

<sup>3</sup>Department of Electronics and Communication Engineering, K. Ramakrishnan College of Technology, Tiruchirappalli, 621112, India

<sup>4</sup>Department of Applied Data Science, Noroff University College, Kristiansand, Norway

<sup>5</sup>Department of Computer Science, Faculty of Engineering & Informatics, University of Bradford, Bradford, United Kingdom

<sup>6</sup>English Department, University College, Taraba, Taif University, Taif, 21944, Saudi Arabia

\*Corresponding Author: R. Bhaskaran. Email: bhaskarpsna@gmail.com

Received: 15 October 2021; Accepted: 29 December 2021

**Abstract:** Sentiment Analysis (SA) is one of the subfields in Natural Language Processing (NLP) which focuses on identification and extraction of opinions that exist in the text provided across reviews, social media, blogs, news, and so on. SA has the ability to handle the drastically-increasing unstructured text by transforming them into structured data with the help of NLP and open source tools. The current research work designs a novel Modified Red Deer Algorithm (MRDA) Extreme Learning Machine Sparse Autoencoder (ELMSAE) model for SA and classification. The proposed MRDA-ELMSAE technique initially performs pre-processing to transform the data into a compatible format. Moreover, TF-IDF vectorizer is employed in the extraction of features while ELMSAE model is applied in the classification of sentiments. Furthermore, optimal parameter tuning is done for ELMSAE model using MRDA technique. A wide range of simulation analyses was carried out and results from comparative analysis establish the enhanced efficiency of MRDA-ELMSAE technique against other recent techniques.

**Keywords:** Sentiment analysis; data classification; machine learning; red deer algorithm; extreme learning machine; natural language processing

### 1 Introduction

The continuous growth of internet has increased the amount of opinions expressed by users in social media and other digital platforms. Many users express their emotions and views through comments in internet platforms [1,2]. E.g., Product-related comments are given in e-commerce websites like Taobao and Jingdong while hospitality-related comments are produced in travel websites like ELong and Ctrip. These comments convey the views and emotions of internet users about hot events, products, and so on. Merchants can reap customer satisfaction with suitable product comments. Prospective buyers and potential users can estimate a product by looking at such product comments or reviews. With a dramatic



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

increase in the number of comments, it is challenging to investigate the comments automatically. Therefore, Information Technology (IT) is exploited to mine the sentiment tendency enclosed in a text which is otherwise known as text sentimental analyses techniques. Text sentimental analyses represent the tendency to mine text sentiments using technology. Based on the granularity of text, text sentimental analyses can be separated under three stages such as chapter level, word level, and sentence level. SA aims at offering fast data by treating the posted reviews through Machine Learning (ML) methods rather than manual reading which is a difficult practice to follow, based on previous experiences posted online on web platforms [3].

Supervised and unsupervised ML approaches are utilized in SA which extracts the basic information from structured and unstructured text to assist the decision maker [4]. The supervised technique proves to be efficient in describing the polarity of sentiments. However, it needs huge amount of labelled information that is not simple for accomplishing. On the other hand, unsupervised techniques are still useful in treating the data without any labels, though the technique cannot be said as a better one. In classification process, the major phase is to elect the appropriate features from information. In addition, Parts of speech, Word2vec, Term Frequency (TF), and TF-IDF are the widely employed features from SA [5]. The application of certain features with different classification models yields a certain result. Therefore, appropriate method would be analyzed to use the special features with distinct classifiers and its performance should be examined. A high performance of uni- and bigram features, in terms of sentiment classification, is observed in literature. Nonetheless, a dedicated classifier is not suitable for twitter whereas a voting/integration ensemble of various classification methods shows remarkable performance for SA [6]. ML technique is generally used in the implementation of Sentiment Analysis (SA). It is broadly categorized into two types such as Supervised- and Unsupervised learning models [7]. Supervised learning approach depends on labelled datasets which are used to train the classification model whereas ML method executes the classification. Generally, partial and hierarchical clustering methods are utilized in unsupervised models [8]. In ML technique, SA precision mainly depends on the trained database based on which the classification is performed. When the size of database increases, the precision decreases upon using comparable trained database. Lexicon method does not depend upon a trained database which makes it appropriate for SA [9,10]. In literature, the researchers developed an effective pattern-based technique for feature extraction and opinion phrase extraction. SA was performed using Lexicon-based method. Various stages have been accomplished while performing feature level SA including feature selection, pre-processing, review extraction, POS tagging to positive or negative, and review classification.

Li et al. [11] presented a Sentiment Information-based Network Model (SINM) in which Transformer encoding and LSTM can be utilized as model mechanisms. Chinese emotional dictionary was used in this research to automatically define the sentiment skill from Chinese comments. In SINM, a hybrid task learning technique was planned for learning valuable emotional terms and forecast sentiment tendency. Initially, SINM needs to learn the sentiment skill from text. During the auxiliary control of emotional data, SINM has to pay further attention towards sentiment data compared to useless data. Wang et al. [12] examined sentiment diffusions by analyzing a phenomenon named sentiment reversal. The authors determined any stimulating characteristic in sentiment reversal. Afterwards, it can be regarded as inter-relationship between textual data of Twitter message and sentiment diffusion pattern. This iterative technique was presented in the name of SentiDiff for predicting sentiment polarity observed in Twitter messages. In order to achieve the best of ability, this case is the initial step to utilize sentiment diffusion designs that can help Twitter SA to enhance.

A comprehensive sentiment dictionary was constructed in the study conducted earlier [13]. This wide sentiment dictionary comprises of fundamental sentiment word, field sentiment word, and polysemic sentiment word that enhance the accuracy of SA. Naive Bayesian (NB) technique was utilized to

determine the polysemic sentiment and the field of text. Based on this model, the sentiment value of polysemic sentiment words from the fields is attained. By employing comprehensive sentiment dictionaries and the planned sentiment score rule, the sentiment of text is scored. In [14], a Multi-Attention Fusion Modeling (Multi-AFM) was presented that combines global as well as local attention with gating unit control so as to generate a reasonable contextual representation and gain enhanced classification outcomes. The experimental outcomes demonstrate that Multi-AFM technique is superior to the presented approaches in the field of education and other such fields.

Li et al. [15] made a danmaku sentiment dictionary and proposed a novel technique with the help of sentiment dictionary and NB to SA of danmaku review. This technique has been significantly useful in the supervision of entire expressive orientation of danmaku videos and forecasting its popularity. With procedures that involve the removal of expressive data in danmaku video, categorizing sentiment, and visualization information, the time distribution of seven sentiment dimensions was attained. Wu et al. [16] presented a technique for Chinese micro-blog sentiment computation with difficult sentences for clauses to clauses-to-words, and with emoji. This technique accurately categorized the comments in Chinese micro-blog as positive, negative, and neutral. Liang et al. [17] introduced the construction of a Gaussian Process Dynamic Bayesian Network for modelling dynamic as well as interactive sentimental issues on social media like Twitter. It utilized Dynamic Bayesian Network for modelling the time series of sentiments expressed for the compared issues and learns connections amongst them. The network technique itself implemented Gaussian Process Regression to model the sentiments at provided time point based on the issues compared at preceding time.

The current research work proposes a new Modified Red Deer Algorithm (MRDA) Extreme Learning Machine Sparse Autoencoder (ELMSAE) model for Sentiment Analysis and classification. The proposed MRDA-ELMSAE technique initially performs pre-processing to transform the data into a compatible format. In addition, TF-IDF vectorizer is employed for extraction of features. Followed by, ELMSAE model is applied for the classification of sentiments. Besides, optimal parameter tuning of ELMSAE model is conducted with the help of MRDA technique. An extensive experimental validation was performed on benchmark dataset and the results were examined under varying aspects. In short, the contributions of the paper are summarized herewith.

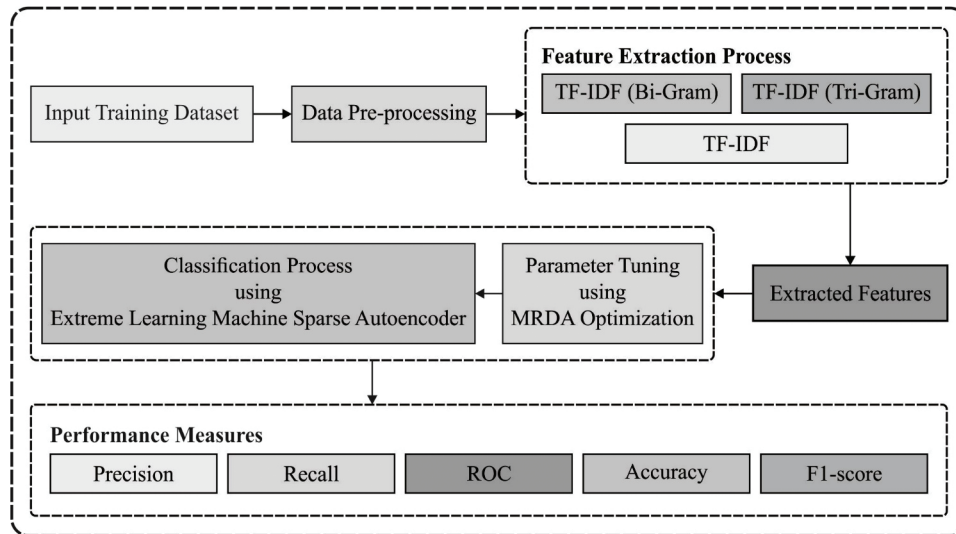
- A novel MRDA-ELMSAE technique is proposed to detect and classify the sentiments under different classes.
- Encompasses different sub-processes namely pre-processing, extraction of features, classification, and parameter optimization.
- An effective TF-IDF-based feature vector transformation and ELMSAE-based classification for SA are designed.
- A novel MRDA-based parameter optimization technique is developed for optimal adjustment of the parameters in ELMSAE model.
- The proposed MRDA-ELMSAE technique is validated for its performance using benchmark dataset.

Rest of the paper is organized as briefed herewith. Section 2 elaborates the proposed MRDA-ELMSAE technique while Section 3 offers a detailed overview about performance validation for the proposed model. Lastly, Section 4 draws the conclusion.

## 2 The Proposed Model

In this study, a new MRDA-ELMSAE technique is designed for the classification of sentiments under positive and negative polarities. The proposed MRDA-ELMSAE technique encompasses four different stages namely, pre-processing, feature extraction using TF-IDF technique, classification using ELMSAE,

and parameter tuning using MRDA technique. Fig. 1 exhibits the overall working process of the proposed MRDA-ELMSAE technique.



**Figure 1:** Overall process of MRDA-ELMSAE model

## 2.1 Pre-Processing

Data pre-processing is the foremost step in this model which is employed for the removal of incomplete and noisy data. The dataset, utilized in this case, has enormous amount of redundant data which tend not to act another role from the forecasted one. Both testing as well as training time enhances, when dataset is the most important entity. The removal of unwanted data accelerates the trained model. Primary phase is an analytics step with missed values that recognizes and removes the missed information since it reduces the classification performance. Afterward, the mathematical value is divided into text as it does not give information about learning of the classifier. It decreases the complexity of the trained classifier. At few instances, the analysis process may have to encounter special symbols like thumb sign, hear sign, and so on. These symbols are also detached to decrease the feature dimensions and improve performance. Then, subsequent punctuations [ ] ( ) / ; . ' are divided during analysis as it does not give meaning to text studies. It cripples the ability of the model in separating punctuation and other character. In the succeeding stage, the words are altered for lowercase. If this phase is not implemented, the ML modules sum up the character. Lastly, stemming is applied which is a vital pre-processing phase. During this phase, the attaches in the word are removed. Stemming modifies this word to its original/root procedures and uses it for increasing the performance of classification.

## 2.2 TF-IDF Based Feature Extraction

At this stage, the preprocessed data is provided as input to TF-IDF model. A python library called Scikit learns is utilized [18] as per the literature. This library is a perfect choice to execute a few tasks with TF-IDF vectorizer method. This model consists of TF-IDF vector which denotes a term 'relative importance' in the record or altogether. The second feature of this model is Term Frequency (TF) which is significant. It implies the frequency of a term that appears from the datasets (defined as 'term frequency' while facing data exploration). The equation to find the TF is given herewith.

$$TF(t, d) = \frac{\text{Number of times } t \text{ takes place from the doc 'd'}}{\text{Total word count of doc 'd'}} \quad (1)$$

IDF, Inverse Document Frequency is the next entity that should be defined to ensure appropriate functioning of the algorithm. It is utilized to measure the importance of a word in whole dataset. The equation for IDF is given herewith.

$$IDF(t) = \log_e \left( \frac{\text{Total number of documents}}{\text{Number of documents with term } t \text{ in it}} \right) \quad (2)$$

Hereafter, it is defined as TF-IDF. TF-IDF is equivalent to Inverse Document Frequency, incorporated with TF, as given herewith.

$$TF-IDF(t, d) = TF(t, d) * IDF(t) \quad (3)$$

The TF-IDF method removes the feature engineering and calculates the appropriate term from fake and real news under this dataset. These reasons help in attaining high efficiency. TF-IDF vectorizer method is the next step in this working mechanism. TF-IDF Vectorizer employs an in-memory jargon (python dictionary) for planning the most consecutive word to emphasize the files and streamline the word event recurrence (scanty) networks. Using TF-IDF vectorizer, recurrent weightings and tokenized records are archived.

### 2.3 ELMSAE-Based Classification

ELMSAE model receives the extracted features as input and performs the classification process. In this stage, the universal approximation capacity of the ELM is used for the purpose of AE. Furthermore, sparse constraints are included in AE optimization due to which it is called ELM sparse autoencoder. Different from AE (*viz.*, *BP related NN*) which is applied in conventional DL methods, the input weight of the presented ELM sparse AE is determined by seeking the path back from arbitrary space. ELM concept has shown that the ELM trained by randomly-mapped input weight is highly effective in approximating the input data. In other words, if AE is trained after ELM concept after AE is initiated, then there is no need to fine-tune the parameters. Additionally, for generating compact and sparse features of the input,  $\ell_1$  optimization is performed for ELM-AE establishment. However, it is dissimilar to ELM-AE and  $\ell_2$ -norm singular value is estimated for feature depiction. Fig. 2 depicts the architecture of ELM. The optimization method for the presented ELM sparse AE is given herewith

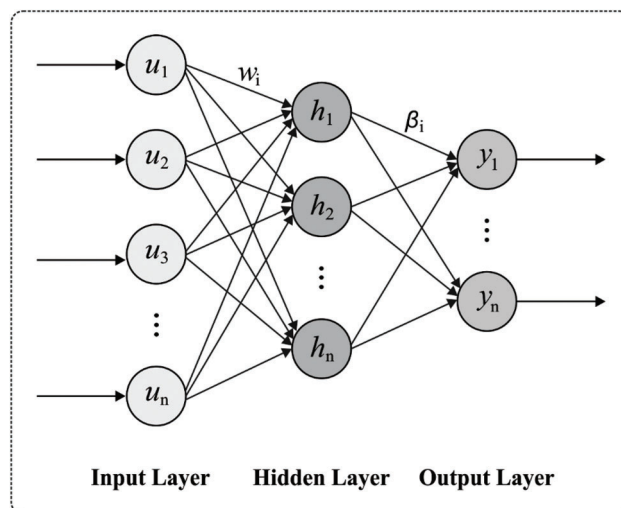


Figure 2: ELM structure

$$O_\beta = \arg \min_{\beta} \{ \|H\beta - X\|^2 + \|\beta\|_{\ell_1} \} \quad (4)$$

Here,  $X$  represents the input data,  $H$  denotes the random mapping output, and  $\beta$  indicates the hidden layer weights to be attained. In present DL method,  $X$  is the encoding output of the base  $\beta$  that should be altered at the time of iteration. But, in the presented AE, random mapping is employed for hidden layer feature depiction.  $X$  represents the actual information and  $H$  indicates the randomly-initiated output which needs no enhancement [19].

Moreover, the following subsection shows that it will not assist in enhancing the trained time and learning precision. Hereafter, the optimization method is explained for  $\ell_1$  optimization issue. For certain depictions, the objective function is altered as given herewith.

$$O_\beta = p(\beta) + q(\beta) \quad (5)$$

where  $(\beta) = \|H\beta - X\|^2$ , and  $q(\beta) = \|\beta\|_{\ell_1}$  is the  $\ell_1$  penalty term of the training model.

FISTA method is adapted to resolve the problems. FISTA decreases a smooth convex function using difficulty of  $(1/j^2)$ , whereas  $j$  indicates the iteration time. The performance of FISTA is given herewith.

- (1) Estimate the Lipschitz constant  $\gamma$  of the gradient of smooth convex function,  $\nabla p$ .
- (2) Begin the iteration by considering  $y_1 = \beta_0 \in R^n$ ,  $t_1 = 1$  as the primary point. Next, for  $j(j \geq 1)$ .
  - (a)  $\beta_j = s_\gamma(y_j)$ , whereas  $s_\gamma$  denotes the following equation.

$$s_\gamma = \operatorname{argmin}_{\beta} \left\{ \frac{\gamma}{2} \|\beta - (\beta_{j-1}) - \frac{1}{\gamma} \nabla p(\beta(j-1))\|^2 + q(\beta) \right\}. \quad (6)$$

$$(b) \ t_{j+1} = \frac{1 + \sqrt{1 + 4t_j^2}}{2}.$$

$$(c) \ y_{j+1} = \beta_j + \left( \frac{t_j - 1}{t_{j+1}} \right) (\beta_j - \beta_{j-1}).$$

After the iteration process is computed, it can handle the retrieval of data point from the corrupted one completely. With resulting base  $\beta$  as the weight of the presented AE, the inner products of both input and learned feature reflect the compact representation of actual information.

## 2.4 MRDA-Based Parameter Tuning

For optimal adjustment of the parameters involved in ELMSAE model, MRDA technique is applied. The mathematical modeling of MRDA technique and distinct phases of the modified method are discussed in this section. Similar to original RDA, this altered form is directed towards the identification of global or local optimum solution for the problems considered. In solution space (P), possible solutions are regarded as Red Deer (RD) which is determined by

$$RD = [P1, P2, P3, \dots, P_D] \quad (7)$$

Here,  $D$  indicates the dimension of solutions. The badness or goodness of a solution is defined by well-determined fitness function as expressed in Eq. (8)

$$f(RD) = f(P1, P2, P3, \dots, P_D) \quad (8)$$

Initially, the population is generated by making  $NP$  number of RD solutions. According to the fitness values, optimal RD is regarded as  $N_{male}$  whereas the remaining solutions are considered as  $N_{hind}$ . Now, the elitism property is monitored to choose the RD in these two sets. According to this property, the number of males is regarded as follows.



$$N_{male} = round\{\alpha_1.NP\} \quad (9)$$

Here,  $\alpha_1$  represents an arbitrary constant value i.e.,  $\alpha_1 \in [0.2, 0.5]$ . The amount of hinds ( $N_{hind}$ ) is represented as  $N_{hind} = NP - N_{male}$ . The position of male RDs, regarded as the best solution, makes an attempt to upgrade the best results in neighborhood too. The natural observations of the new method are preserved whereas the arbitrariness for upgrading the locations of male RDs is also integrated. Instead of utilizing two constant values which are arbitrarily made from the original form of approach, a single arbitrary constant value is employed at this point. In order to provide a gradient in adaptability and assist an increasingly low rate of change in parameter value, the arithmetic equation is designed precisely as given herewith.

$$male_{new} = \begin{cases} male_{old} - (1 - a) * (ub - lb) * \frac{n}{i^2 + n}, & a < 0.5 \\ male_{old} + a * (ub - lb) * \frac{n}{i^2 + n}, & a \geq 0.5 \end{cases} \quad (10)$$

Here,  $a$  represents a haphazardly-made real number ( $a \in [0, 1]$  the present generation and the overall amount of generations are represented as  $i$  and  $n$ , correspondingly. Different kinds of males exist in nature. According to original RD method, both stags and commanders are two kinds of RD which are considered under this altered form. The amount of commander males ( $N_{cm}$ ) is determined as follows.

$$N_{cm} = round\{\alpha_2.N_{male}\} \quad (11)$$

whereas  $\alpha_2$  represents the arbitrary constant values i.e.,  $\alpha_2 \in [0.2, 0.5]$ . So, 20% to 50% of males are chosen as commanders. Rest of the males are regarded as stags and the quantity of stags ( $N_{stag}$ ) is calculated as given herewith.

$$N_{stag} = N_{male} - N_{cm} \quad (12)$$

In view of original RD method, the fighting process between stags and commanders is based on a significant randomization number. The two arbitrary constant parameters perform a significant part. Based on novel RDA, a set of hinds in male commanders is called 'harem'. The size of harem is based on the strength of commanders. During harem formation process in novel RDA, the fundamental concept is to assist the accessibility of Female-Hinds, according to the objective fitness of stags and commanders. This occurs in the ascending order of magnitude based on its natural and particular domain behavior. During MRDA, the allocation of hinds is distinctively conceived based on the number of commanders within the population and randomization parameter,  $R$ . In order to achieve the structure of harem, the following steps are performed to select the hinds in control of commander

- $P$  = Overall amount of hinds/Amount of unallocated commanders
- $R$  = Random integer in (1,  $P$ )
- $N$  = ( $R$ \* Amount of unallocated hinds)  $IP$

At first, the  $N$  number of hinds is allocated to the initial commander. In the following generation, the  $N$  number of hinds is allocated to the following commanders.

In novel RDA, a commander's mate exists with specific kind of hind in his harem. The arithmetical equation is given below.

$$N_{of1} = a * cm + (1 - a) * hn - a * (ub - lb) * \frac{n - \frac{i}{n}}{i^2 * n}, a > 0.5 \quad (13)$$

$$N_{of^2} = (1 - a) * cm + a * hn - (1 - a) * (ub - lb) * \frac{n - \frac{i}{n}}{i^2 * n}, a \leq 0.5 \quad (14)$$

The arithmetical model is engineered distinctively to conserve less recessive features, and more of dominant features with certain goals of availing a gradient alteration from the value. However, stags are permitted to randomly mate with every hind under the population. However, a specific stag does not take part in mating several times in a certain generation. In order to choose the population for upcoming generations, the quality of offspring RDs is estimated based on fitness value. According to the fitness value of present offspring and the RD of preceding generation, the best RD is chosen for upcoming generation. Roulette wheel selection [20] approach is utilized for this election procedure. This step is repeated for a specific number of iterations for time interval or optimal solution is not altered for a longer period.

### 3 Experimental Analysis Results

The current section discusses the results of experimental investigation conducted upon MRDA-ELMSAE technique using a benchmark dataset including mobile application analyses for Google app. The dataset encompasses 64,295 instances with ‘App’, ‘Translated\_Reviews’, and ‘Sentiments’. It includes three kinds of sentiments namely, positive, negative, and neutral.

Tab. 1 demonstrates the results of comparative analysis accomplished by MRDA-ELMSAE model using distinct types of features under different measures. Fig. 3 provides the results of the analysis achieved by MRDA-ELMSAE technique under TF-IDF features. The figure conveys that the proposed MRDA-ELMSAE technique accomplished improved outcome over existing methods. For instance, MRDA-ELMSAE technique classified the negative label with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 98.7%, 92.3%, and 95.4% respectively. Moreover, MRDA-ELMSAE method classified the neutral label with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 93.6%, 95%, and 94.2% correspondingly. Furthermore, MRDA-ELMSAE method categorized the positive label with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 99%, 99.5%, and 99.1% correspondingly. Similarly, RF model accomplished poor performance on the classification of sentiments. RF model classified the negative labels with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 87%, 71%, and 78% respectively. Meanwhile, RF method categorized the neutral labels with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values such as 79%, 82%, and 81% correspondingly. Eventually, RF method classified the positive labels with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 89%, 94%, and 92% correspondingly.

**Table 1:** Comparative study of MRDA-ELMSAE technique under distinct types of features

Models	Negative_Class			Neutral_Class			Positive_Class		
	$Pre_n$	$Rec_l$	$F_{score}$	$Pre_n$	$Rec_l$	$F_{score}$	$Pre_n$	$Rec_l$	$F_{score}$
TF-IDF Feature									
Proposed MRDA-ELMSAE	98.7	92.3	95.4	93.6	95.0	94.2	99.0	99.5	99.1
BBSO-FCM	97.0	89.0	92.0	90.0	94.0	91.0	97.0	99.0	98.0
Gradient Boosted SVM	92.0	84.0	88.0	85.0	90.0	87.0	94.0	96.0	95.0
Gradient Boosting Tree	92.0	82.0	87.0	80.0	94.0	86.0	95.0	95.0	95.0
Support Vector Machine	89.0	83.0	86.0	84.0	86.0	85.0	93.0	95.0	94.0
Logistic Regression	91.0	75.0	82.0	85.0	71.0	77.0	88.0	97.0	92.0
Random Forest	87.0	71.0	78.0	79.0	82.0	81.0	89.0	94.0	92.0

(Continued)



**Table 1 (continued).**

TF-IDF Bigram Feature									
	Negative_Class			Neutral_Class			Positive_Class		
	$Pre_n$	$Rec_l$	$F_{score}$	$Pre_n$	$Pre_n$	$Rec_l$	$F_{score}$	$Rec_l$	$Pre_n$
Proposed MRDA-ELMSAE	88.0	64.8	73.1	69.3	46.4	42.9	85.9	99.7	93.3
BBSO-FCM	86.0	63.0	72.0	68.0	45.0	42.0	84.0	99.0	92.0
Gradient Boosted SVM	75.0	56.0	64.0	60.0	26.0	36.0	76.0	92.0	83.0
Gradient Boosting Tree	76.0	36.0	49.0	46.0	18.0	26.0	71.0	94.0	81.0
Support Vector Machine	83.0	49.0	61.0	90.0	20.0	32.0	74.0	97.0	84.0
Logistic Regression	88.0	21.0	34.0	92.0	3.0	6.0	67.0	99.0	80.0
Random Forest	78.0	49.0	60.0	55.0	26.0	35.0	75.0	93.0	83.0

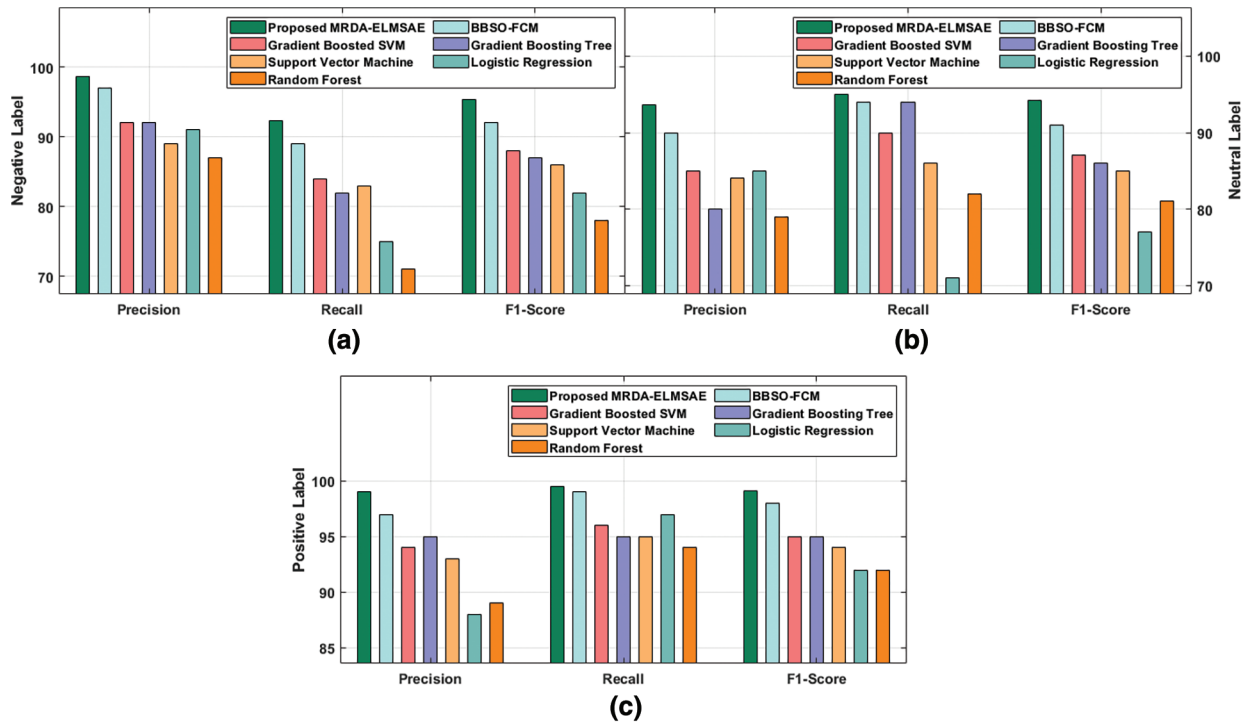
  

TF-IDF Trigram Features									
	Negative_Class			Neutral_Class			Positive_Class		
	$Pre_n$	$Rec_l$	$F_{score}$	$Pre_n$	$Rec_l$	$F_{score}$	$Pre_n$	$Rec_l$	$F_{score}$
Proposed MRDA-ELMSAE	93.7	46.8	60.7	78.3	76.4	74.3	80.8	100.0	88.1
BBSO-FCM	90.0	46.0	60.0	77.0	76.0	74.0	79.0	99.0	87.0
Gradient Boosted SVM	83.0	37.0	51.0	69.0	19.0	30.0	71.0	96.0	82.0
Gradient Boosting Tree	84.0	18.0	30.0	77.0	11.0	19.0	67.0	99.0	80.0
Support Vector Machine	96.0	32.0	48.0	94.0	17.0	29.0	70.0	99.0	82.0
Logistic Regression	97.0	11.0	20.0	100.0	1.0	3.0	65.0	100.0	79.0
Random Forest	41.0	32.0	47.0	84.0	18.0	30.0	70.0	99.0	82.0

Fig. 4 gives the results of the analyses accomplished by MRDA-ELMSAE system of the projected method under TF-IDF Bi-gram features. The figure shows that the proposed MRDA-ELMSAE approach achieved maximal results over existing models. For instance, MRDA-ELMSAE approach categorized the negative label with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values such as 88%, 64.8%, and 73.1% correspondingly. Furthermore, MRDA-ELMSAE system categorized the neutral label with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 69.3%, 46.4%, and 42.9% correspondingly. Moreover, the proposed MRDA-ELMSAE approach classified the positive label with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values such as 85.9%, 99.7%, and 93.3% correspondingly. Likewise, RF approach achieved poor performance on the classification of sentiments. RF method classified the negative labels with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 78%, 49%, and 60% correspondingly. Meanwhile, RF approach classified the neutral labels with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 55%, 26%, and 35% correspondingly. In the end, RF method classified the positive labels with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values such as 75%, 93%, and 83% correspondingly.

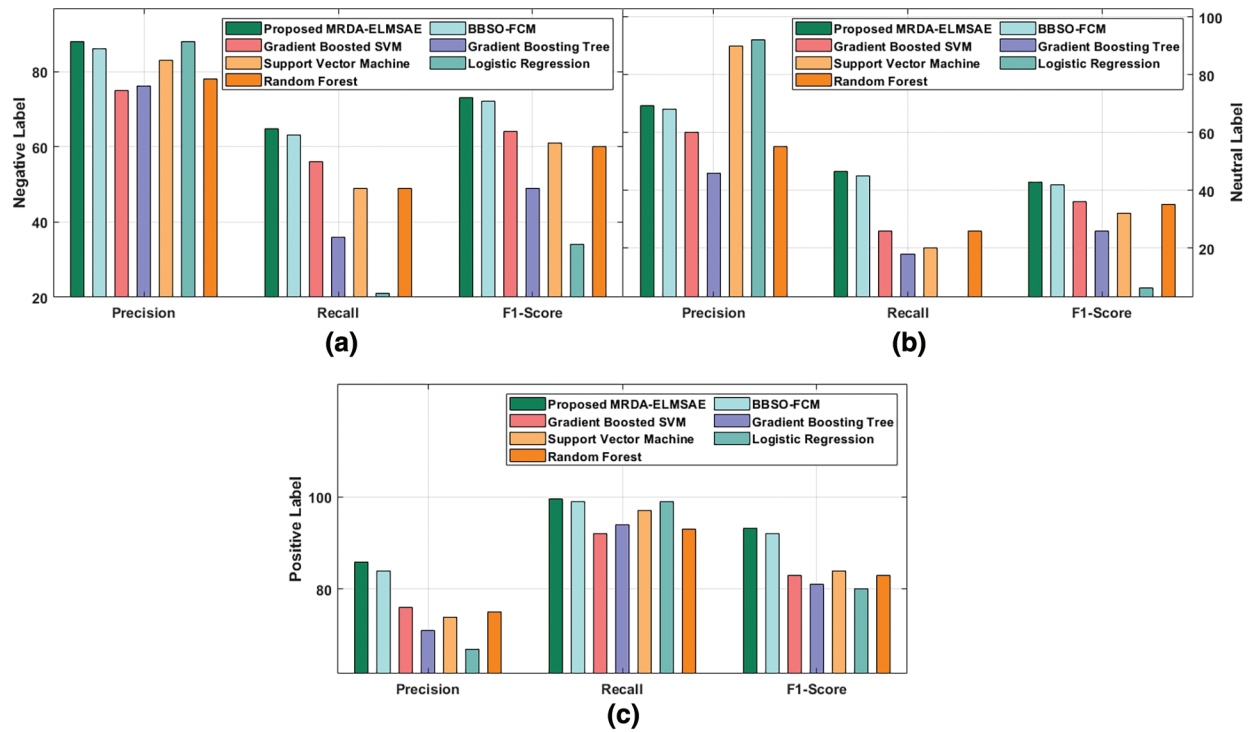
Fig. 5 portrays the results accomplished by MRDA-ELMSAE method of the presented technique under TF-IDF Tri-gram features. The figure showcases that the proposed MRDA-ELMSAE algorithm achieved high results over existing approaches. MRDA-ELMSAE method classified the negative label with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 93.7%, 46.8%, and 60.7% correspondingly. Furthermore, MRDA-ELMSAE approach classified the neutral label with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values such as 78.3%, 76.4%, and 74.3%

correspondingly. Moreover, MRDA-ELMSAE method classified the positive label with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 80.8%, 100%, and 88.1% correspondingly. Likewise, RF system reached poor efficiency on the classification of sentiments. RF method classified negative labels with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values being 41%, 32%, and 47% correspondingly. Meanwhile, RF approach classified the neutral labels with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  values such as 84%, 18%, and 30% correspondingly. At the end, the RF method identified the positive class with  $Pre_n$ ,  $Rec_l$ , and  $F_{score}$  of 70%, 99%, and 82% correspondingly.

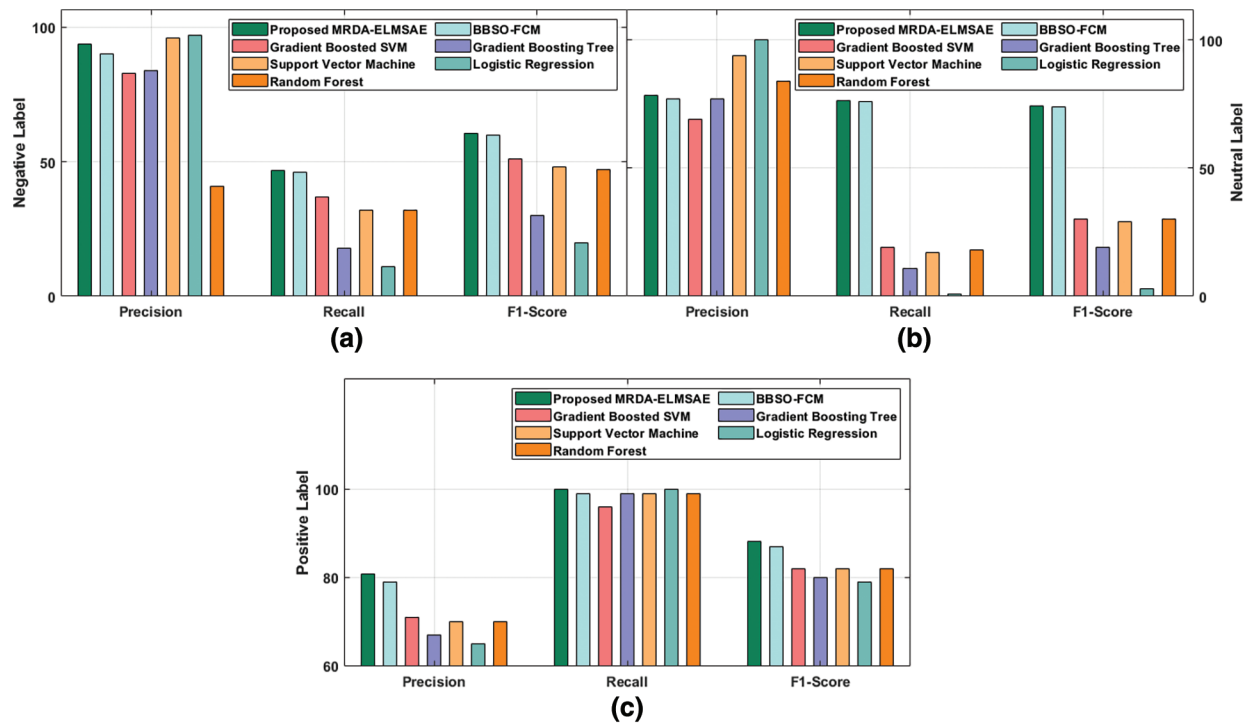


**Figure 3:** Comparative analysis results of MRDA-ELMSAE model with TF-IDF features (a)  $Pre_n$ , (b)  $Rec_l$ , and (c)  $F_{score}$

Finally, a comparative accuracy analysis was conducted between the proposed MRDA-ELMSAE technique against existing approaches and the results are shown in Tab. 2 and Fig. 6. The figure shows that R\_NB\_KNN technique gained poor outcomes with low accuracy (73%). At the same time, Stochastic GDCLR, GCN, SGCN, and NABoE techniques obtained moderate performance with accuracy values being 88%, 88%, 89%, and 86% respectively. Moreover, BBSO-FCM and Gradient Boosted SVM techniques too accomplished near optimal accuracy i.e., 96% and 93% respectively. However, the proposed MRDA-ELMSAE technique resulted in increased accuracy of 98%. Based on the results and discussion so far, it can be confirmed that the proposed MRDA-ELMSAE technique is an effective tool for Sentiment Analysis.



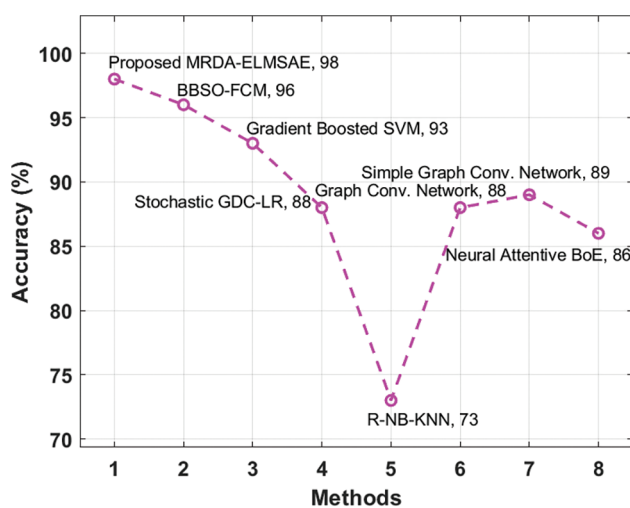
**Figure 4:** Comparative analysis results of MRDA-ELMSAE model with Bigram TF-IDF features (a)  $Pre_n$ , (b)  $Rec_l$ , and (c)  $F_{score}$



**Figure 5:** Comparative analysis results of MRDA-ELMSAE model with Trigram TF-IDF features (a)  $Pre_n$ , (b)  $Rec_l$ , and (c)  $F_{score}$

**Table 2:** Comparison of proposed MRDA-ELMSAE technique against existing techniques

Methods	Accuracy (%)
Proposed MRDA-ELMSAE	98.00
BBSO-FCM	96.00
Gradient-Boosted SVM	93.00
Stochastic GDCLR	88.00
R_NB_KNN	73.00
GCN	88.00
SGCN	89.00
NABoE	86.00

**Figure 6:** Accuracy analysis results of MRDA-ELMSAE model against existing approaches

#### 4 Conclusion

In current research a novel MRDA-ELMSAE technique is derived for the classification of sentiments under positive and negative polarities. The proposed MRDA-ELMSAE technique encompasses four different stages such as pre-processing, feature extraction using TF-IDF technique, classification using ELMSAE, and parameter tuning using MRDA technique. The application of MRDA technique to adjust the weight and bias of ELMSAE model helps in achieving enhanced classification performance. A wide range of simulation analyses was carried out and the results from comparative analysis established the enhanced performance of MRDA-ELMSAE technique than other recent approaches. In future, feature selection approaches can be included to boost the overall classification results.

**Funding Statement:** We acknowledge Taif University for Supporting this study through Taif University Researchers Supporting Project number (TURSP-2020/173), Taif University, Taif, Saudi Arabia.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] H. Peng, E. Cambria and A. Hussain, "A review of sentiment analysis research in Chinese language," *Cognitive Computation*, vol. 9, no. 4, pp. 423–435, 2017.
- [2] M. Dragoni and G. Petrucci, "A neural word embeddings approach for multi-domain sentiment analysis," *IEEE Transactions on Affective Computing*, vol. 8, no. 4, pp. 457–470, 2017.
- [3] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016.
- [4] C. Diamantini, A. Mircoli and D. Potena, "Social information discovery enhanced by sentiment analysis techniques," *Future Generation Computer Systems*, vol. 95, no. 6, pp. 816–828, 2019.
- [5] S. Kashfia and A. Reda, "Emotion and sentiment analysis from Twitter text," *Journal of Computational Science*, vol. 36, no. 9, pp. 1–42, 2019.
- [6] J. Zhao and X. Gui, "Comparison research on text pre-processing methods on Twitter sentiment analysis," *IEEE Access*, vol. 5, pp. 2870–2879, 2017.
- [7] S. Stieglitz and L. D. Xuan, "Emotions and information diffusion in social media-sentiment of microblogs and sharing behavior," *Journal of Management Information Systems*, vol. 29, no. 4, pp. 217–248, 2013.
- [8] D. Zeng, Y. Dai, F. Li, J. Wang and A. K. Sangaiah, "Aspect based sentiment analysis by a linguistically regularized CNN with gated mechanism," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 5, pp. 3971–3980, 2019.
- [9] P. Ji, H. Y. Zhang and J. Q. Wang, "A fuzzy decision support model with sentiment analysis for items comparison in E-commerce: The case study of PConline.Com," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 49, no. 10, pp. 1993–2004, 2019.
- [10] D. Tang, F. Wei, B. Qin, N. Yang, T. Liu *et al.*, "Sentiment embeddings with applications to sentiment analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 2, pp. 496–509, 2016.
- [11] G. Li, Q. Zheng, L. Zhang, S. Guo and L. Niu, "Sentiment information based model for chinese text sentiment analysis," in *2020 IEEE 3rd Int. Conf. on Automation, Electronics and Electrical Engineering (AUTEEE)*, Shenyang, China, pp. 366–371, 2020.
- [12] L. Wang, J. Niu and S. Yu, "SentiDiff: Combining textual information and sentiment diffusion patterns for twitter sentiment analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 10, pp. 2026–2039, 2020.
- [13] G. Xu, Z. Yu, H. Yao, F. Li, Y. Meng *et al.*, "Chinese text sentiment analysis based on extended sentiment dictionary," *IEEE Access*, vol. 7, pp. 43749–43762, 2019.
- [14] G. Zhai, Y. Yang, H. Wang and S. Du, "Multi-attention fusion modeling for sentiment analysis of educational big data," *Big Data Mining and Analytics*, vol. 3, no. 4, pp. 311–319, 2020.
- [15] Z. Li, R. Li and G. Jin, "Sentiment analysis of danmaku videos based on naïve bayes and sentiment dictionary," *IEEE Access*, vol. 8, pp. 75073–75084, 2020.
- [16] J. Wu, K. Lu, S. Su and S. Wang, "Chinese micro-blog sentiment analysis based on multiple sentiment dictionaries and semantic rule sets," *IEEE Access*, vol. 7, pp. 183924–183939, 2019.
- [17] H. Liang, U. Ganeshbabu and T. Thorne, "A dynamic bayesian network approach for analysing topic-sentiment evolution," *IEEE Access*, vol. 8, pp. 54164–54174, 2020.
- [18] G. M. Raza, Z. S. Butt, S. Latif and A. Wahid, "Sentiment analysis on covid tweets: an experimental analysis on the impact of count vectorizer and tf-idf on sentiment predictions using deep learning models," in *2021 Int. Conf. on Digital Futures and Transformative Technologies (ICoDT2)*, Islamabad, Pakistan, pp. 1–6, 2021.
- [19] Y. Zeng, L. Qian and J. Ren, "Evolutionary hierarchical sparse extreme learning autoencoder network for object recognition," *Symmetry*, vol. 10, no. 10, p. 474, 2018.
- [20] S. De, S. Dey, S. Debnath and A. Deb, "A new modified red deer algorithm for multi-level image thresholding," in *2020 Fifth Int. Conf. on Research in Computational Intelligence and Communication Networks (ICRCICN)*, Bangalore, India, pp. 105–111, 2020.