

Reductive Bias with Gated Recurrent Sentiment Analysis for Social Media Texts Using the Twitter Dataset

D. Srilatha^{1*}, Harikrishna Bommala²

¹Research Scholar, Computer Science and Engineering, Bharatiya Engineering Science and Technology Innovation University, Gorantla, Anantapur, Andhra Pradesh, India

²Professor, Computer Science and Engineering, Reddy College of Engineering and Technology, Moinabad, Hyderabad, 501504, India

*E-mail: 2022pcse017@bestiu.edu.in; haribommala@kgr.ac.in

Received for publication
July 09, 2025.

Abstract

Recently, people are progressively expressing their feelings and opinions on social media sites such as Facebook, twitter, and Instagram, and interacting with other people worldwide. However, the classification of customer opinions is challenging due to the lack of relevant information about categorizing them into specific classes. To solve this problem, a reductive bias-based gated recurrent unit (RD-GRU) approach is proposed to enhance the classification of sentiments in the Twitter dataset effectively. Initially, the data from Sentiment140 and Twitter API are preprocessed by four different techniques, then the features are extracted by Log term frequency (LTF) and modified inverse class frequency (MICF) technique. The features are selected based on principal component object (PCO) and the proposed IBAO algorithm efficiently. Finally, classification of user review is performed by the SA-Bi-LSTM approach that categorizes the reviews into Positive, Negative, and Neutral classes. The results of the proposed model for sentiment analysis (SA) is evaluated by the performance metrics which attained accuracy of 98.92% and 97.84% for Sentiment140 dataset and Twitter API datasets is superior than existing models such as convolutional neural network (CNN) and long short-term memory (LSTM) approaches.

Keywords

gated recurrent unit, log term frequency-modified inverse class frequency, principal component analysis, reductive bias, sentiment analysis

1. Introduction

Sentiment analysis (SA) is an application of natural language processing (NLP), where individual data is certainly extracted from the text data [1, 2]. Nowadays, people share their opinions on social media with each other, worldwide. The main objective of SA is to analyze the sentiments of users based on text data on social media, which is obtained from tweets that are in the form of unstructured data [3–5]. The text-based SA for social media is performed according to three categories: sentence-level, aspect-level, and

document-level categorization. Sentence-level classification determines, for a given sentence with several words, whether the specific text falls into a positive, negative, or neutral category [6–8]. Aspect-level analysis is determined by analyzing the texts based on various aspects and classifying them into positive, neutral, and negative sentiments, respectively [9]. Lastly, document-level SA is that which analyzes massive textual data and sentences in the document of a user rather than delivering the whole basic aspect of the entity.

Nowadays, for analyzing the sentiments expressed by people in tweets, artificial intelligence

techniques are widely used to categorize the sentiments by learning the context information about the texts. Initially, machine learning (ML) methods were utilized for analysis of sentiments in tweets, but these methods failed to provide better classification and accurate results. Due to a lack of relevant features, the model has difficulty in differentiating the sentiments efficiently [10–11]. Thus, deep learning (DL), which is an ML-based method, is used to classify the sentiments in the texts accurately by extracting the relevant features. In DL methods, the neural network with several layers can learn complex representations of sentiments, which helps to analyze the sentiments effectively [12–13]. However, existing DL methods struggle to analyze and categorize the positive, negative, and neutral classes of emotions precisely from tweets, due to a lack of relevant information about these emotion classes; moreover, similar characteristics between the positive, negative, and neutral classes lead to misclassification. To overcome this limitation, a reductive bias-based gated recurrent unit (RD-GRU) is proposed to analyze- the sentiments in

Twitter data by differentiating the positive, negative, and neutral classes efficiently [14–15]. The main contributions of this research are given below.

Figure 1 illustrates SA as a NLP methodology employed to discern emotions and opinions from textual data, particularly from social media platforms such as Twitter. It functions at three tiers: sentence-level, aspect-level, and document-level, each concentrating on distinct granularities of sentiment classification. Initially, ML techniques were employed but encountered challenges with accuracy owing to inadequate feature extraction. DL methodologies, especially neural networks such as gated recurrent units (GRU), enhance performance by discerning intricate patterns in textual data. Nonetheless, DL models encounter difficulties in differentiating across analogous sentiment categories. A unique RD-GRU model is developed, including a reduction bias mechanism into GRU to prioritize sentiment-laden words (e.g., adjectives, adverbs, negations). This improves the model's capacity to precisely categorize attitudes as good, negative, or neutral in tweets.

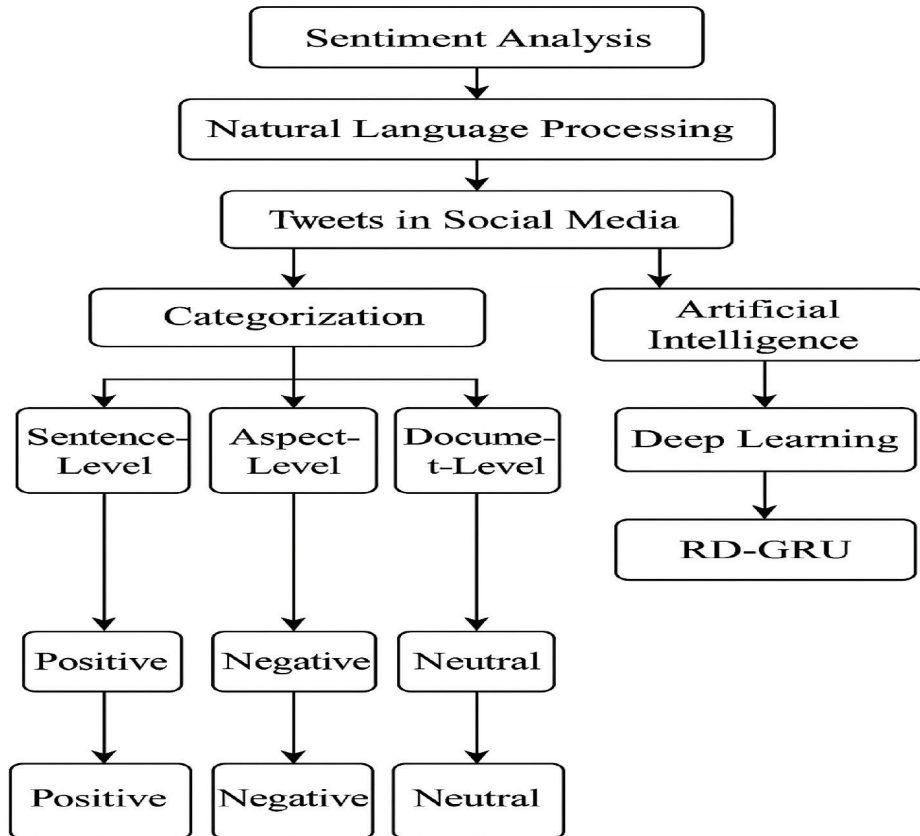


Figure 1: Methodology of reductive bias in gated recurrent SA of twitter data. RD-GRU, reductive bias-based gated recurrent unit; SA, sentiment analysis.

- GRU captures long-term dependencies and temporal patterns in tweet sequences effectively, enabling the model to understand sentiment flow over time.
- The reductive bias technique is incorporated with GRU's update gate to emphasize sentiment-rich components such as adjectives, adverbs, or negations.
- The remaining part of this research is organized as follows: Section II explains the literature review. Section III describes the methodologies implemented for this research. Section IV illustrates the experimental results. Section V concludes the paper.

II. Literature review

Halawani et al. [16] developed an NLP model along with dimensionality reduction based on support vector machine, k-nearest neighbor, and Naive Bayes methods for SA for social media tweets. The developed ML models were employed individually to analyze the sentiments in the tweets from social media. However, the feature selection technique utilized in the developed SA model has poor search ability, which impacts the selection of relevant features and reduces the model's performance.

Aslan et al. [17] presented a model that utilizes DL techniques for SA of the Twitter dataset. The presented model utilized a convolutional neural network (CNN) and long short-term memory (LSTM) to analyze the sentiments from tweets. The advantage of the presented DL model is that a self-attention mechanism helps the model focus more on specific information, which enhances the classification performance of sentiments. However, the presented DL model was unable to analyze certain words due to multiple meanings for the same word, thus leading to poor analysis.

Parveen et al. [18] designed a sentiment classification model for a social media text dataset based on a hybrid gated attention network. The hybrid model was designed with the White Shark optimizer to analyze and categorize the various sentiments from tweets. The main advantage of the designed hybrid gated attention recurrent network is its use of term-weighting-based feature extraction and the White Shark optimizer, which selects features based on polarity to enhance the performance of the hybrid model. However, the designed hybrid gated attention recurrent network fails to differentiate the sentiments of positive, negative, and neutral, which leads to inaccurate results.

Vidyashree and Rajendra [19] presented an SA model based on stochastic gradient descent optimization with a stochastic gate neural network for the Twitter dataset. The represented gate neural network model, which was incorporated with the gradient descent optimization method, was utilized to analyze the sentiments effectively. The advantage of the representational gate neural network model was that stochastic gradient optimization enhanced its effectiveness. However, the represented gated neural network model failed to analyze sentiment for certain words, because those words shared similar characteristics with other sentiments.

Serpil Aslan et al. [20] designed CNN-based sentiment for the social media Twitter dataset. The SA CNN model, which incorporates an arithmetic optimization model, is designed for text classification of tweets to address existing limitations by detecting and eliminating stop words. The arithmetic optimization model fine-tuned the developed model, enhancing its detection performance. However, the designed CNN model with the arithmetic optimization model struggles to accurately classify emotions in SA due to irrelevant features.

III. Methodology

The main objective of the proposed DL method is to analyze the sentiment/emotions present in the tweets from social media. This proposed framework involves five phases: dataset, preprocessing, feature extraction, feature selection, and classification. Figure 2 represents the block diagram for DL-based SA for social media text. At first, the tweets from social media are acquired from two Twitter datasets, Sentiment140 and SemEval, and then preprocessed by various techniques to make a useful format of the text by removing unnecessary words in tweets. After preprocessing, these transformed texts are passed to log term frequency-modified inverse class frequency (LTF-MICF) for feature extraction based on the term weighting process. Next, the extracted features are passed to the proposed feature selection process, which selects optimal features by using an optimization algorithm. Finally, with the selected features as input, the proposed DL-based classifier model analyzes and categorizes the sentiments present in the tweets, which are social media texts.

a. Datasets

In this research, the social media datasets such as Sentiment140 and SemEval based on Twitter are used as input for DL-based SA.

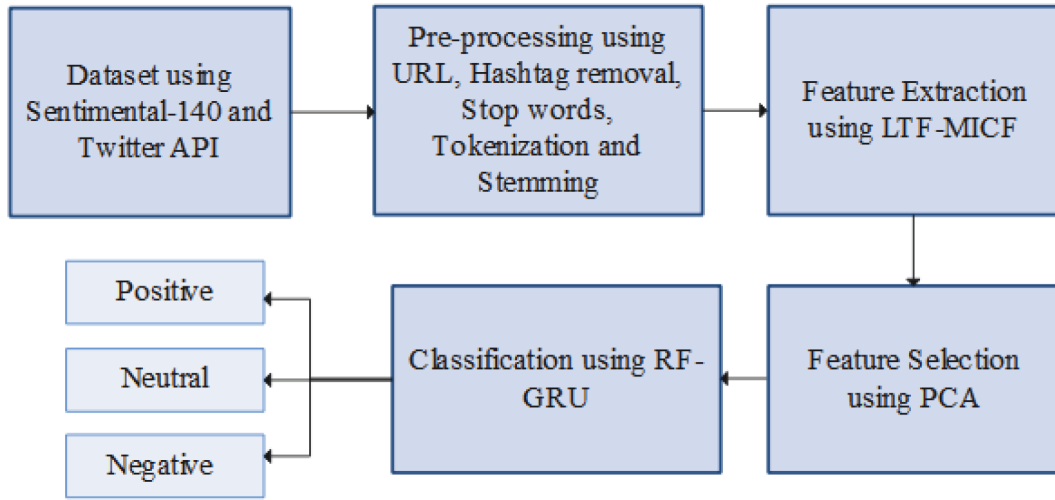


Figure 2: Block diagram for SA for social media texts. LTF-MICF, log term frequency-modified inverse class frequency; SA, sentiment analysis.

a.i Sentiment140 dataset

This is a publicly available Twitter dataset that is extracted from the Twitter API and comprises 1,600,000 tweets. These tweets are categorized into three groups, such as positive, neutral, and negative, based on the score values, where negative tweets are represented by the score value “0,” neutral tweets are denoted by the score value “2,” and positive tweets are indicated by the score value “4.” According to the score values, there are 22,542 negative tweets, 18,318 neutral tweets, and 20,832 positive tweets in the dataset with six attributes. Table 1 illustrates the attributes of the Sentiment140 dataset.

a.ii Twitter API dataset

The Twitter API dataset, a publicly available dataset, focuses on real-time applications that gather text data from Twitter. This dataset has the data, which

Table 1: The six attributes of Sentiment140 dataset

S. no.	Attributes
1.	Text of the tweet
2.	User who tweeted
3.	Whether there is a query on the tweet or not?
4.	Date of tweet
5.	Tweet ID
6.	Polarity of the tweet

is categorized into three categories: positive tweets, negative tweets, and neutral tweets from the user. These raw data of tweets are fed to the initial processing phase of the preprocessing stage to enhance the data by removing unwanted texts.

b. Preprocessing

The collected raw text data is initially fed into the preprocessing phase to transform the texts into a useful format for enhancing the SA. This research employs preprocessing techniques such as stemming, tokenization, and stop word removal to transform the raw text sentences obtained from the dataset. The process of these techniques is explained as follows:

- **URL and Hashtags removal:** The tweets include several hashtags and URLs, which are helpful to the users; however, they contain noise that is not used to further the process. Thus, these hashtags and URLs are removed from texts.
- **Stop words removal:** After eliminating the hashtags, the tweets are passed to the stop words removal technique, which identifies and eliminates irrelevant texts in the sentence. In stop word removal, the frequently utilized stop words like “and,” “is,” “the,” and “so,” which have no sentiment, are removed from the tweets. The techniques include Z-methods, mutual information, the term-based random sampling method (TBRS), and others.
- **Tokenization:** The tokenization technique divides a sentence into small texts, phrases, and other

meaningful forms, which are referred to as “tokens.” Moreover, this tokenization helps identify the most meaningful words that enhance accurate SA. These tokens are further passed to the stemming process.

- i. **Stemming:** After tokenization, the tokens are transformed into their root words by this stemming process. In this process, the prefixes and suffixes like “un,” “dis,” “ed,” “im,” and “ing” are removed to make the text into a root word or dictionary form. During the stemming process, the two processes mentioned below were considered:
- ii. The words with different meaning should keep separately.
- iii. The words of morphological forms that have similar meaning should be mapped with the same stem words. These preprocessed texts (stem words) are forwarded to the next phase of feature extraction.

c. Feature extraction

After preprocessing, the preprocessed data is fed to the feature extraction phase based on a new term weighting technique, namely, LTF-MICF. This feature extraction method combines two distinct term weighting strategies, LTF and MICF, which extract the features based on frequently occurring terms in the sentence or tokens, referred to as term frequency. The frequently occurring terms in the sentence/tokens provide insufficient term frequency. Therefore, this hybrid feature extraction technique is employed to address this issue by integrating term frequency with the MICF technique. The inverse class frequency is determined as the inverse ratio of the entire class of terms that occurs during training the tweets to the total classes. The LTF-MICF is mathematically formulated as given in Eq. (1):

$$LTF - MICF(tp) = T_i(tp) * \sum_{r=1}^m w_{sp} - [C_i(tp)] \quad (1)$$

Where, w_{sp} notes the weighting factor (WF) for every class; tp presents the preprocessed terms; and the specific WF is expressed mathematically as in Eq. (2):

$$w_{sp} = \log \left(1 + \frac{s_i \bar{t}}{\max(1, s_i \bar{t})} \cdot \frac{s_i \bar{t}}{\max(1, s_i \hat{t})} \right) \quad (2)$$

Where, s_i notes the number of tweets in a class; $s_i \bar{t}$ dictates the preprocessed terms; $s_i \hat{t}$ presents

that the class does not have preprocessed terms; and the hybrid feature extraction method is utilized to allocate weights for a given Twitter dataset, which is known as the WF. The constant “1” is used to avoid the negative weights in the texts. These extracted features are forwarded to the feature selection phase for further process of SA.

d. Feature selection

The extracted features from the hybrid method are further passed to the principal component analysis (PCA) technique-based feature selection phase. This PCA technique is extensively used for unsupervised learning algorithms in ML techniques for dimensional reduction. Similarly, in this research, PCA is employed to minimize the dimensions of the extracted feature, the high-dimensional feature vector set, to a new low-dimensional vector set and also selects the most relevant features. The PCA method normalizes the data first, then constructs a covariance matrix by acquiring eigenvectors; thereafter, representative eigenvectors are identified. This PCA technique assumes that most of the feature information in the three classes includes significant variations. The process of reducing dimensional reduction of the feature vector using PCA is mentioned below:

- 1) Estimate the mean value.
- 2) Identify the covariance matrix of extracted features.
- 3) Determine the covariance matrix to acquire eigenvalues and eigenvector, then arrange it based on descending order. Then, obtain p dimensional feature subspace ($k \leq M$) by calculating p largest eigenvalues and equivalent eigenvector.

e. Classification

After dimensionality reduction and feature selection by PCA, the selected features are forwarded as input to the proposed classification method for SA of social media texts into positive, negative, and neutral. For efficient classification of sentiments, a DL-based method is proposed in this research, namely, the GRU, which is a version of the recurrent neural network (RNN) model. The GRU model allows the prediction results of past contextual information of the words and also makes a sequence operation of features, which is more stable compared with individually processing each selected feature. The proposed GRU model includes two gates: an update gate that

combines the input, and the forget gates found in LSTM. A function of the update gate is to choose the memory information that makes the model continue to be retained until the present stage. The reset gate controls how much past information is fed to the network. The mathematical formulations of these two gating states, based on the hidden layer state and the input of the present node, are presented in Eqs (3) and (4):

$$z_t = \sigma(W_{xz}x^t + W_{hz}h^{t-1} + b_z) \quad (3)$$

$$r_t = \sigma(W_{xr}x^t + W_{hr}h^{t-1} + b_r) \quad (4)$$

Where, σ notes sigmoid function; x^t dicates input text; z_t and r_t represent update and reset gates; h^{t-1} ands for hidden layer output; W_{xz} and W_{xr} note weights of update and rest gates; and b_r and b_z represent the reset and update gates.

The information flow extracted from the selected features in the network is leveraged by the two gates in GRU. The above-mentioned update and reset gates are used to determine information which are to be retained and those that are to be forgotten, also alleviating the gradient descending problem. However, the GRU has certain limitations like overfitting, long-term dependencies, and give less importance to frequently occurring patterns in the dataset that affects to differentiate sentiments efficiently that lead to misclassification.

e.i Reductive bias

In DL-based algorithms, residual is known as the “Residual term or error”, which is a difference among the observed and analyzed or predicted values that are created by the model. In a linear regression, the residual term at every data point is estimated by subtracting the observed value of the dependent variable from the analyzed value, which is determined as corresponding to the independent variables and the GRU model parameters. The mathematical representation of reductive bias r_t time stamp t expressed in Eq. (5):

$$r_t = (\hat{Y}_t^{\tau} - x_t) \quad (5)$$

Where, $\hat{Y}_t^{\tau}$ notes the prediction or classification or analysis results of the GRU model. Then the average of residual r_t elds bias (b) of model which is written as in given Eq. (6):

$$b = \left(\frac{1}{m}\right) \sum_{t=1}^m (r_t) \quad (6)$$

In terms of bias ($\hat{b}$) corresponding to overall evaluation, is represented in Eq. (7):

$$\begin{cases} b^+, & \text{if } b > 0 \\ b^-, & \text{otherwise} \end{cases} \quad (7)$$

The incorporation of this reductive bias is in the hidden state of the GRU model that enhances the rest gate and update gate of the GRU with this bias term. This enhancement utilizes only the information that is most relevant from the selected features to that sentiment. Utilizing the proposed reductive bias technique addresses the issue of the GRU model and improves its ability to differentiate between the positive, negative, and neutral classes effectively.

IV. Results and Discussion

The performance evaluation of the proposed DL-based classifier for SA of social media text data, which includes the Sentiment140 and Twitter API datasets, is presented here. The performance of the proposed classifier and feature extraction methods is analyzed with approaches used in SA, as illustrated in Tables 2–5 and Figures 2 and 3. The proposed SA framework employs a variety of performance measures to assess the effectiveness of its methods. We use performance metrics like accuracy, precision, recall, and F1-score. The mathematical representation of these performance measures is expressed in Eqs (8)–(11):

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100 \quad (8)$$

$$Precision = \frac{(TP)}{(TP + FP)} \times 100 \quad (9)$$

$$Recall = \frac{(TP)}{(TP + FN)} \times 100 \quad (10)$$

$$F1 - score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \times 100 \quad (11)$$

Where, *TP* True Positive, *FP* False Positive, *TN* True Negative, and *FN* False Negative.

a. Performance analysis

The quantitative and qualitative analysis of the proposed DL-based classification of emotions utilized for SA of social media texts using two benchmark datasets are depicted in this section. Existing DL approaches include CNN, LSTM, RNN, and GRU. Tables 2 and 3 and Figures 1 and 2 represent the performance analysis of feature extraction and the proposed classification models.

b. Comparative analysis

Comparative analysis evaluates the proposed method used to analyze sentiments in texts obtained from two benchmark Twitter datasets. Existing SA approaches are utilized to evaluate the proposed approach for comparison. Performance metrics such as accuracy, precision, recall, and F1-score are utilized for comparative analysis of the proposed method. Tables 4 and 5 represent the comparative analysis of the proposed method for Sentiment140 dataset and Twitter API datasets, respectively.

V. Discussion

The proposed approach achieved better results with precise classification of sentiments or emotions

of users from the Amazon review dataset. SA of user reviews by various ML and DL models has a problem in classifying emotions accurately due to a lack of informative features from the review dataset, which is accomplished by the presented RD-GRU method. IBAO Improves Feature Quality by selecting only the most relevant features; it reduces noise and enhances the model's ability to learn meaningful patterns and boosts classification accuracy. Studies indicate that models using IBAO outperform traditional GRU and LSTM models in sentiment classification tasks. SVM, the feature selection technique utilized in the developed SA model has poor search ability, which impacts selection of the relevant features and reduces the model's performance. CNN-LSTM was unable to analyze certain words due to multiple meanings for the same word, thus leading to poor analysis. The CNN model with arithmetic optimization model SA based on the developed model struggles to classify emotions accurately because of irrelevant features. To solve this problem, an IBAO approach is proposed to effectively enhance sentiment classification in the Twitter dataset [21]. However, RD-GRU captures long-term dependencies and temporal patterns in tweet sequences effectively, enabling the model to understand sentiment flow over time [22]. The incorporation of this reductive bias in the hidden state of the GRU model enhances the reset gate and update gate of the GRU with this bias term [23]. This enhancement utilizes only the

Table 2. Performance analysis of the proposed method for Sentiment140 dataset

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CNN	91.11	92.93	89.96	91.42
LSTM	93.79	94.41	91.40	92.88
RNN	95.36	96.95	93.15	95.01
GRU	96.74	97.42	95.63	96.52
Proposed RD-GRU Method	98.92	98.49	97.54	98.01

CNN, convolutional neural network; GRU, gated recurrent unit; LSTM, long short-term memory; RD-GRU, reductive bias-based gated recurrent unit; RNN, recurrent neural network.

Table 3. Performance analysis of feature extraction method for Sentiment140 dataset

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Word2Vec	91.09	92.48	93.04	92.76
Glove	94.13	95.69	95.80	95.74
Skip gram	96.33	96.29	96.36	96.32
LTF-MICF	98.92	98.49	97.54	98.01

LTF-MICF, log term frequency-based modified inverse class frequency.

information that is most relevant from the selected features to that sentiment.

Figure 3 and Figure 4 illustrate the performance of the RD-GRU model an RD-GRU model for the SA of Twitter data. In contrast to conventional ML

and DL models, which frequently encounter difficulties in accurately classifying attitudes due to feature overlap and insufficient contextual focus, RD-GRU incorporates a reductive bias mechanism inside the GRU framework. This approach improves the

Table 4. Comparative analysis of the proposed method for Sentiment140 dataset

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
SVM [17]	84.25	85.83	86.37	86.13
CNN-LSTM [18]	97.86	96.65	96.76	96.70
IBAO [21]	98.73	97.56	96.46	95.89
Proposed RD-GRU	98.92	98.49	97.54	98.01

CNN-LSTM, convolutional neural network-long short-term memory; RD-GRU, reductive bias-based gated recurrent unit.

Table 5. Comparative analysis of the proposed method for Twitter API dataset

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
TSA-CNN-AOA (KNN)	95.09	94.23	93.23	93.71
Proposed RD-GRU	97.84	97.37	96.95	97.15

RD-GRU, reductive bias-based gated recurrent unit.

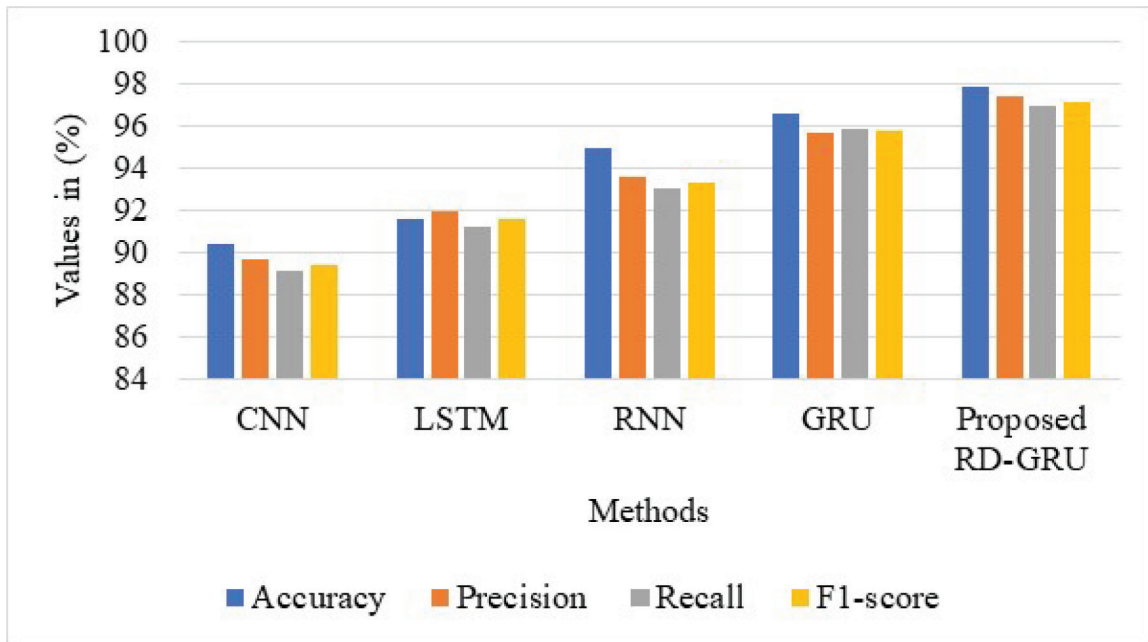


Figure 3: Performance analysis of the proposed method for Twitter API dataset. API, CNN, convolutional neural network; GRU, gated recurrent unit; RD-GRU, reductive bias-based gated recurrent unit; RNN, recurrent neural network.

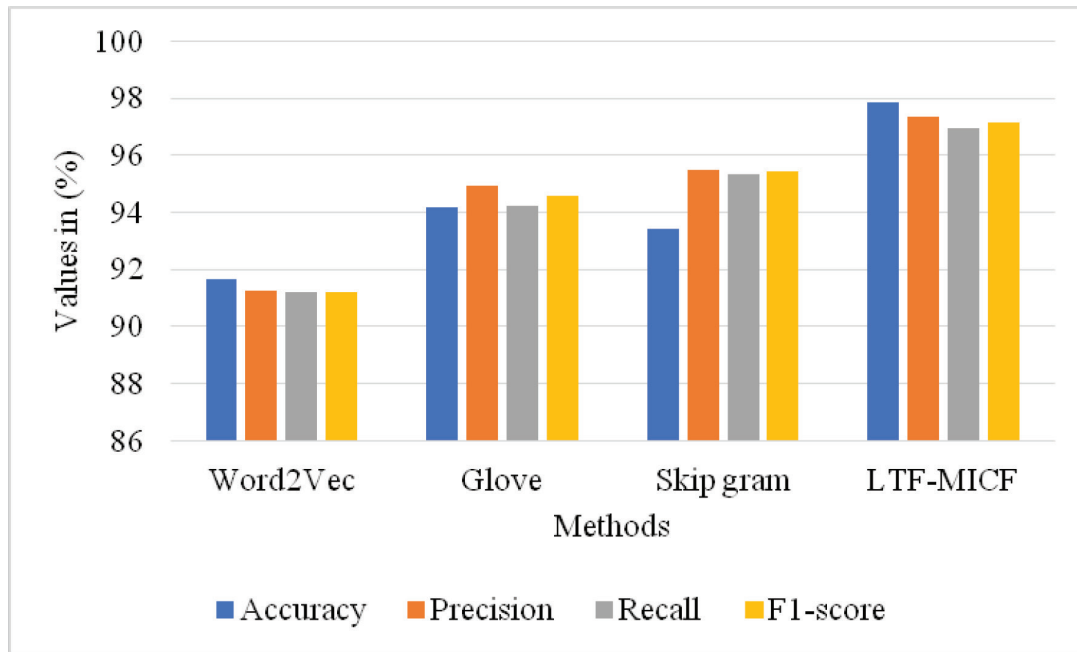


Figure 4: Performance analysis of feature extraction method for Twitter API dataset. API, LTF-MICF, log term frequency-based modified inverse class frequency.

model's capacity to concentrate on sentiment-laden elements such as adjectives, adverbs, and negations by altering the update gate. Consequently, RD-GRU enhances the distinction among positive, negative, and neutral sentiment categories, resulting in more accurate classification. The model adeptly captures long-term dependencies and temporal trends in tweet sequences, providing a more profound comprehension of sentiment progression across time. This focused improvement renders RD-GRU a unique and more efficient method for SA in chaotic, unstructured social media data.

VI. Conclusion

The RD-GRU approach is proposed to enhance classification of sentiments in the Twitter dataset effectively. Initially, the data from Sentiment140 and the Twitter API are preprocessed by four different techniques, and then features are extracted by the LTF and MICF techniques. The features are selected based on PCA and the proposed IBAO algorithm efficiently. Finally, classification of user reviews is performed by the SA-Bi-LSTM approach that categorizes the reviews as positive, negative, and neutral classes. The results of the proposed model for SA are evaluated by the performance metrics, which attained an accuracy of 98.92% and 97.84% for the Sentiment140

dataset and Twitter API datasets, and are superior to existing models such as CNN and LSTM approaches. In the future, transformer-based models will be used to enhance the classification of sentiments in the social media text for various Twitter datasets.

References

- [1] V. Mahalakshmi, P. Shenbagavalli, S. Raguvaran, V. Rajakumareswaran and E. Sivaraman, "Twitter sentiment analysis using conditional generative adversarial network." *International Journal of Cognitive Computing in Engineering*, vol. 5, pp. 161-169, 2024.
- [2] A. Alsayat, "Improving sentiment analysis for social media applications using an ensemble deep learning language model." *Arabian Journal for Science and Engineering*, vol.47, no. 2, pp. 2499-2511, 2022.
- [3] S.T. Kokab, S. Asghar and S. Naz, "Transformer-based deep learning models for the sentiment analysis of social media data." *Array*, vol.14, pp. 100157, 2022.
- [4] L. Khan, A. Amjad, K.M. Afaq and H.T. Chang, "Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu text shared in social media." *Applied Sciences*, vol. 12, no. 5, pp. 2694, 2022.

- [5] R. Obiedat, R. Qaddoura, A.Z. Ala'M, L. Al-Qaisi, O. Harfoushi, M.A. Alrefai and H. Faris, "Sentiment analysis of customers' reviews using a hybrid evolutionary SVM-based approach in an imbalanced data distribution." *IEEE Access*, vol. 10, pp. 22260-22273, 2022.
- [6] N. Aslam, F. Rustam, E. Lee, P.B. Washington and Ashraf I, "Sentiment analysis and emotion detection on cryptocurrency related tweets using ensemble LSTM-GRU model." *IEEE Access*, vol. 10, pp. 39313-39324, 2022.
- [7] M. Loukili, F. Messaoudi and M. El Ghazi, "Sentiment Analysis of Product Reviews for E-Commerce Recommendation based on Machine Learning." *International Journal of Advances in Soft Computing & Its Applications*, vol. 15, no. 1.
- [8] C. Ahmed, A. ElKorany and E. ElSayed, "Prediction of customer's perception in social networks by integrating sentiment analysis and machine learning." *Journal of Intelligent Information Systems*, vol. 60, no. 3, pp. 829-851, 2023.
- [9] G. Revathy, S.A. Alghamdi, S.M. Alahmari, S.R. Yonbawi, A. Kumar and M.A. Haq, "Sentiment analysis using machine learning: Progress in the machine intelligence for data science." *Sustainable Energy Technologies and Assessments*, vol. 53, pp. 102557, 2022.
- [10] W. Aljedaani, F. Rustam, M.W. Mkaouer, A. Ghallab, V. Rupapara, P.B. Washington, E. Lee and I. Ashraf, "Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry." *Knowledge-Based Systems*, vol. 255, pp. 109780, 2022.
- [11] N. Chintalapudi, G. Battineni and F. Amenta, "Sentimental analysis of COVID-19 tweets using deep learning models." *Infectious disease reports*, vol. 13, no. 2, pp. 329-339, 2021.
- [12] C.S.R. Priya and P. Deepalakshmi, "Sentiment analysis from unstructured hotel reviews data in social network using deep learning techniques." *International Journal of Information Technology*, vol. 15, no. 7, pp. 3563-3574, 2023.
- [13] H. Kaur, S.U. Ahsaan, B. Alankar and V. Chang, "A proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets." *Information Systems Frontiers*, vol. 23, no. 6, pp. 1417-1429, 2021.
- [14] U. Sehar, S. Kanwal, K. Dashtipur, U. Mir, U. Abbasi and F. Khan, "Urdu sentiment analysis via multimodal data mining based on deep learning algorithms." *IEEE Access*, vol. 9, pp. 153072-153082, 2021.
- [15] A. Baqach and A. Battou, "CLAS: A new deep learning approach for sentiment analysis from Twitter data." *Multimedia Tools and Applications*, vol. 82, no. 30, pp. 47457-47475, 2023.
- [16] H.T. Halawani, A.M. Mashraqi, S.K. Badr and S. Alkhalaf, "Automated sentiment analysis in social media using Harris Hawks optimisation and deep learning techniques." *Alexandria Engineering Journal*, vol. 80, pp. 433-443, 2023.
- [17] S. Aslan, S. Kiziloluk and E. Sert, "TSA-CNN-AOA: Twitter sentiment analysis using CNN optimized via arithmetic optimization algorithm." *Neural Computing and Applications*, vol. 35, no. 14, pp. 10311-10328, 2023.
- [18] N. Parveen, P. Chakrabarti, B.T. Hung and A. Shaik, "Twitter sentiment analysis using hybrid gated attention recurrent network." *Journal of Big Data*, vol. 10, no. 1, pp. 50, 2023.
- [19] K.P. Vidyashree and A.B. Rajendra, "An improved sentiment analysis model on twitter data using stochastic gradient descent (SGD) optimization algorithm in stochastic gate neural network (SGNN)." *SN computer science*, vol. 4, no. 2, pp. 190, 2023.
- [20] Serpil Aslan, A deep learning-based sentiment analysis approach (MF-CNN-BILSTM) and topic modeling of tweets related to the Ukraine–Russia conflict, *Applied Soft Computing*, Vol. 143, 2023, 110404, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2023.110404>.
- [21] Goismi Mohamed, Hamou Reda Mohamed, Tomouh Adil, Maaskri Moustafa Enhancing Twitter Sentiment Classification with a Hybrid Bio-Inspired Feature Selection Approach" Vol. 10 No. 53s (2025) e-ISSN:2468-4376, JISEM.
- [22] S. Hundekari, S.A.V. Nair, P. Shanthi, C.M. Rao, M. Rafeeq, T.B. Sivakumar, "Turbocharged AI: Harnessing Federated Learning and Model Parallelism for Efficient Deep Learning on Distributed System," *Communications on Applied Nonlinear Analysis*, vol.31, 2024.
- [23] M.D. Rafeeq, V.H. Shastri, A. Ba, V. Shanmugasundaram, P.R. Kumar, D. Jyothsna, "An Intelligent Road Traffic Data Management for Upgrading Real Time Performance," *AIP Conference Proceedings*, vol. 2426, 2022.