International Journal of Engineering Trends and Technology
ISSN: 2231-5381/ https://doi.org/10.14445/22315381/IJETT-V74I4P110

Volume 74 Issue 4, 132-143, April 2026

Original Article

Cockroach Optimized Progressive Laplace Extreme Learning Machine for Depression Prediction in Social Media Text

R.Geetha[1], D.Vimal Kumar[2]

[1,2]*Department of Computer Science, Rathinam College of Arts and Science, Coimbatore, Tamil Nadu, India.*

[1]*Corresponding Author : gthram92@gmail.com*

Received: 17 November 2025 Revised: 11 February 2026 Accepted: 19 February 2026 Published: 29 April 2026

Abstract - *Early detection of depression is crucial for timely intervention and effective treatment. Social media platforms play a significant role in sharing individual thoughts and opinions through textual posts. Conventional deep learning models in depression face challenges in improving the accuracy of depression detection with minimal time consumption. To improve the accuracy, a novel model, Cockroach Optimized Progressive Laplace Extreme Learning Machine (COPLELM), is proposed for sentiment analysis of Twitter social media text with minimal time consumption. The data acquisition is to collect text data from the Twitter Dataset. Progressive Laplace Kernelized Extreme Learning Machine is employed for depression prediction with several layers. First, a number of Twitter text data are given to the input layer. Texts are transferred to a hidden layer for performing pre-processing, which involves tasks such as tokenization, stop-word removal, and word stemming. A censored regressive cockroach swarm optimization algorithm is employed in the subsequent hidden layers to extract optimal keywords. Finally, the prediction is performed in the next hidden layer using Gestalt pattern matching. Depression and non-depression Twitter texts are accurately detected in the output layer with minimal time. Experimental evaluations are conducted using various performance metrics, including prediction accuracy, precision, recall, F1-score, specificity, error rate, and depression prediction time, confusion matrix, and ROC. The results demonstrate that the proposed COPLELM model achieves higher accuracy in depression prediction, with reduced time and error rates compared to existing methods.*

Keywords - *Depression detection, Social Media Text, Progressive Extreme Learning Machine, Laplace Kernel, censored regressive cockroach swarm optimization algorithm, Gestalt pattern matching.*

1. Introduction

Depression is the most widespread mental illness that affects many people worldwide. An attention-based CNN-BiLSTM Attention (CBA) model was developed in [1]. A deep temporal model was introduced in [2]. Advanced transformer-based models were developed in [3]. A multimodal fusion cross-modality method was introduced in [4]. A semi-supervised learning model was proposed in [5]. The BERT model was introduced in [6]. Multiple hybrid machine-learning models were developed in [7]. Machine learning techniques and implementation steps were developed in [8]. Several machine learning algorithms were developed in [9]. A hybrid deep learning model was proposed in [10]. The Additive Cross-Modal Attention Network model was developed in [11]. A low-covariance multimodal combined spatio-temporal converter method was developed in [12]. A deep-knowledge-aware depression detection model was designed in [13]. BERT with a Bidirectional Long Short-Term Memory model was developed in [14]. Artificial intelligence and machine learning models were developed in [15]. Depression is a serious mental health disorder that leads to a series of emotional and physical problems that affect daily functioning and the overall quality of people's lives. Several existing machine-learning methods have been developed for depression detection. The major issues identified in most of the existing depression methods include a longer time, huge amounts of data that cannot be concentrated, and failure to select optimal keywords. Early recognition of depression is crucial for providing timely intervention and treatment. Accuracy and recall were insufficient when using conventional classification techniques. To overcome the gaps, the proposed COPLELM is developed in depression for higher accuracy and less time in the early stage. In this paper, the COPLELM model utilizes text pre-processing and optimal keyword extraction. The TextBlobtokenizer, Laplace kernel-based stop word removal, and Lovins Stemmer are performed in pre-processing tasks with a lesser depression prediction time. Followed by, a censored regressive cockroach swarm optimization algorithm is employed to choose an optimal keyword with higher accuracy. The association between the keywords was estimated and verified with Gestalt pattern matching. Depressed and non-depressed samples are determined via the sigmoid activation function with higher accuracy and less time for the early stage.

1.1. A Novel Contribution of the COPLELM Model

- The proposed COPLELM model has been developed to improve depression prediction accuracy from social media text by integrating text pre-processing, keyword extraction, and classification.
- Text pre-processing is performed in COPLELM by using TextBlobtokenizer, Laplace kernel-based stop word removal, and Lovins Stemmer. TextBlobtokenizer is utilized to extract words. The Laplace kernel function is used to eliminate stop words. The root word is achieved with the Lovins Stemmer method. In this way, depression prediction time is minimized.
- Censored regressive cockroach swarm optimization is employed in COPLELM. The fitness is measured via censored regression for selecting the optimal keywords and eradicating the others. With this, accuracy is improved.
- Innovation of Gestalt pattern matching is applied in COPLELM to measure the contextual relationship between the keywords. Also, the sigmoid activation function is employed for predicting the depressed and non-depressed samples. In this way, precision and recall are improved.

1.2. Paper Organization

The paper is organized into different sections. Section 2 reviews the related works. Section 3 provides a description of the COPLELM model with a neat diagram. Section 4 outlines the dataset utilized in this work. Section 5 provides a comparative analysis of the proposed COPLELM model and existing methods using various performance metrics. Section 6 explains the discussion part of the proposed model. Finally, Section 7 concludes the paper.

Table 1 shows the summary of the literature review.

Table 1. Summary of literature review

Reference No and Method	Objectives	Merits	Demerits
CBA [1]	Find user-level clinical depression based on a user's temporal social media posts.	Higher accuracy	Failed to focus on minimizing the time consumption
Deep temporal model [2]	Achieve robust performance of depression detection across different social media contexts.	Time reduced	Efficacy and robustness were not enhanced
Advanced transformer-based models [3]	Classify mild, moderate, and severe depression by using text data sourced from Reddit.	The error was minimized	Failed to focus on integrating a large dataset
Multimodal fusion cross-modality method [4]	Identify depression by analyzing speech features and linguistic textual features.	Recall was enhanced	A deep learning model was not used
Semi-supervised learning model [5]	Predict mental health conditions from textual data.	Improved accuracy	High Macro-F1 score not achieved
BERT [6]	Early prediction of user depression using social media data.	Precision enhanced	A deep learning approach was not implemented
Multiple hybrid machine-learning models [7]	Detect signs of depression	Accuracy increased	Failed to explore the use of other feature extraction methods
Machine learning techniques [8]	Find depression	Higher precision	Advanced hybrid machine learning models were not employed
Machine learning algorithms [9]	Early-stage depression detection using sentiment analysis	F1 score enhanced	Lacked integration of emotion recognition features
Hybrid deep learning model [10]	Detect depression	Lesser time	Failed to achieve a significant improvement in specificity

2. Related Works

A hybrid model was developed in [16]. An SVM-based classification approach was developed in [17]. A hybrid deep learning approach was proposed in [18]. LLMs were developed in [19, 28]. A FastText-based weighted soft voting ensemble model was developed in [20]. Transformer-based models were developed in [21]. A deep learning model was proposed in [22]. Various ML and transformer models were developed in [23]. A hypergraph neural network was developed in [24]. An ensemble of hybrid model-based techniques was introduced in [25]. The feature group partitioning approach was developed in [26]. The DeepFM network was developed in [27]. Automated estimation methods were developed in [29]. The encoder-only Transformer architecture was designed in [30].

3. Proposal Methodology

Depression is a common mental health disorder that affects a large number of people around the world. A novel COPLELM model is developed for accurate depression detection using social media text. Figure 1 presents the architecture of the COPLELM model developed for accurate depression prediction. The number of tweet samples from the input images is fed into the input layer, where weights and bias functions are assigned. The input and hidden layer weights remain fixed throughout the process. The input text tweet samples are then sent to the neurons of the hidden layer, where text pre-processing is carried out to obtain the number of words. The output from the pre-processed results is subsequently transferred to the next hidden layer.

In that layer, cockroach swarm optimization is applied to select the keywords from the tweets. With the selected keywords, the depression detection is performed based on the contextual relationship measure between the words by using Gestalt pattern matching. Finally, the binary classification outcome is observed at the output layer with a sigmoid activation function.

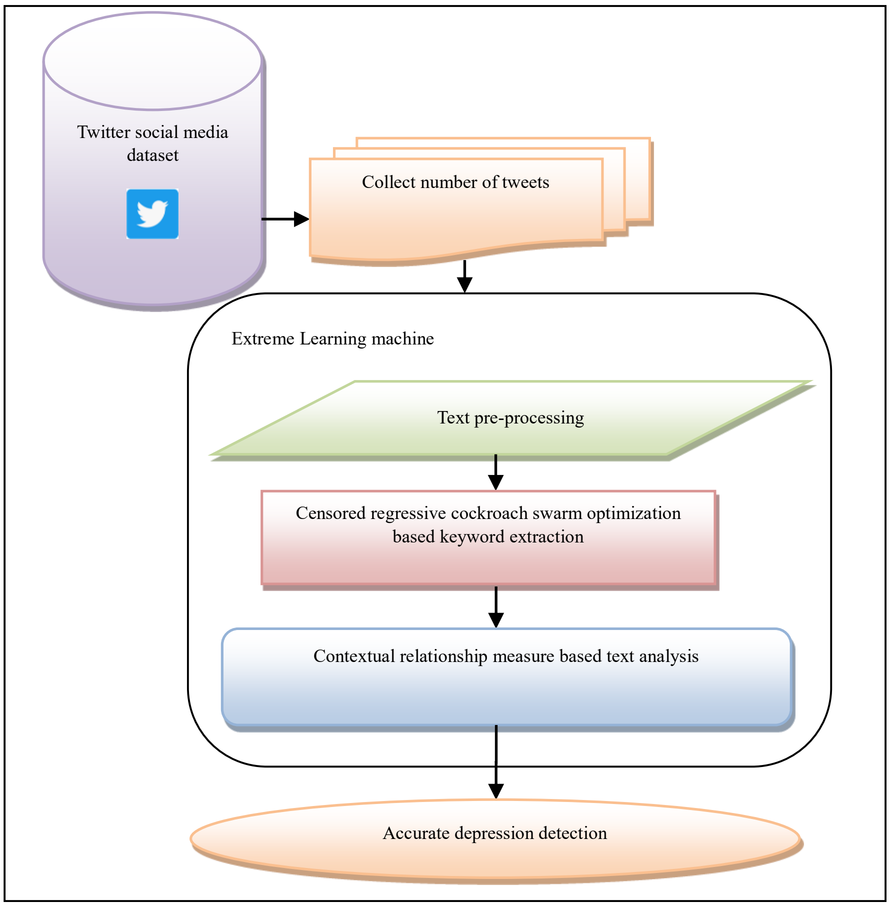

Fig. 1 Architecture diagram of the COPLELM model

3.1. Tweet Data Sample Acquisition

The initial phase of the COPLELM model involves tweet data sample acquisition, where relevant information is collected from Depression: Twitter Dataset + Feature Extraction https://www.kaggle.com/datasets/infamouscoder/mental-health-social-media.

3.2. Progressive Extreme Learning Machine

Progressive ELM is an expansion of standard ELM that is designed to handle sequential learning or incremental data. Sequential learning is a machine learning approach in which the PELM learns from data that arrives one instance or group at a time, updating its parameters incrementally without needing to retrain on the entire dataset.

Compared to other traditional methods, PELM includes many advantages, such as extremely fast learning, low computational burden, especially for large-scale and continuous datasets.

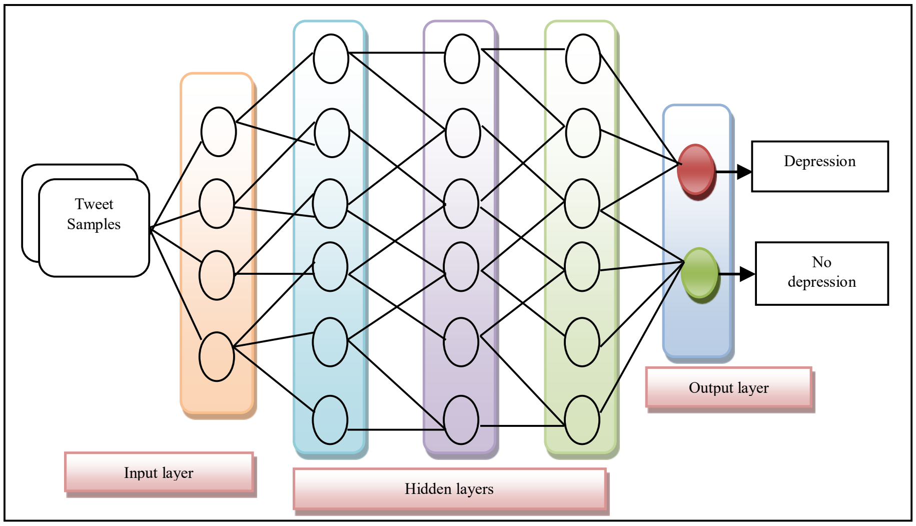

Fig. 2 Schematic structure of PELM

Figure 2: Both input and output layers are always a single layer; the hidden layer includes multiple internal layers. Each layer in the PELM network architecture is made up of basic processing units called neurons or nodes. It helps to strengthen the connection between layers PELM network classifier model considers the training set $\{TS_i, Y_k\}$ where TS_i indicates training sample'tweet samples$TS = \{TS_1, TS_2, TS_3, \dots TS_n\}$' and Y_k indicates a classification of outcomes as depression and no depression.

The number of tweet samples is given as input to the input layer of the PELM network classifier. Each neuron in the input layer receives these tweet samples and forwards them to the neuron of the hidden layer, applying corresponding weights and biases. Each neuron in the hidden layer computes a weighted sum as,

$$Z = \sum_{i=1}^{n}(TS_i * \vartheta_{ih}) + B \quad (1)$$

Where Z indicates a weighted sum function, TS_i represents a tweet sample, ϑ_{ih} indicates a weight related to the connection between a neuron in the input layer and a neuron in the hidden layer, and B denotes a bias for a neuron.

In the first hidden layer, text pre-processing is carried out to ensure that the raw tweet samples are transformed into a suitable format for modeling and further analysis. This phase involves cleaning, organizing, and standardizing the text samples to remove the inconsistencies and enhance the performance of depression detection.

3.2.1. Text Data Samples Pre-Processing

Pre-processing of the Twitter text data sample is a fundamental step for efficient data analysis. It involves capturing the sequence of operations designed to clean and standardize the dat. The following pre-processing steps are applied to the collected text samples, such as tokenization, stop word removal, and word stemming. The TextBlob tokenizer is employed in the pre-processing method. This tokenizer automatically splits the text into a number of words, which includes punctuation, spaces, and commas. The tokenization process of COPLELM is expressed as,

$$TS_i = [\text{'}W_1\text{'}, \text{'}W_2\text{'}, \text{'}W_3\text{'}, \dots, \text{'}W_m\text{'}] \quad (2)$$

Where TS_i denotes a tweets sample, W_1', 'W_2', 'W_3' ... 'W_m denotes words extracted from the tweets using TextBlobtokenizer. Finally, the tokenizer returns a list of words, which are cleanly split from the text for further processing. After the tokenization, the stop word removal step is executed to remove frequently occurring words that carry less meaningful information.

These words are typically removed using the Laplace kernel function. This statistical technique evaluates the relationship between words in the input text samples and a predefined list of stop words. The similarity computation is carried out as described below,

$$K_L = exp\left[-\frac{|W_m - W_l|}{v}\right] \quad (3)$$

Where K_L represents a Laplace kernel function, W_m indicates the number of words in the given tweet samples and W_l represents a list of stop words, v denotes a deviation, $|W_m - W_l|$ indicates the difference between the two words.

$$Q = \begin{cases} K_L > T, & Stopword \\ K_L < T, & Normalword \end{cases} \quad (4)$$

Where Q indicates a stop word removal output, K_L denotes a Laplace kernel function, T indicates a threshold. If the Laplace kernel similarity score exceeds a predefined threshold, the word is classified as a stop word. Otherwise, it is considered a normal word. Word stemming is another important text pre-processing technique used in Twitter text samples to decrease the words to their root or base form, known as the stem. The Lovins Stemmer method is employed to obtain the root word. The example of the Lovins Stemmer process is shown in Table 2.

Table 2. Example of word stemming

Word extracted fromtext	Stemming	Root word
Ending	Ing	End
Employed	Ed	Employ
Totally	ly	Total

In Table 2, words containing suffixes like 'ing', 'ed', and 'ly' are stemmed by removing these endings and obtaining end, employ, and total. The resulting processed words are sent to the next hidden layer for keyword extraction.

3.2.2. Censored Regressive Cockroach Swarm Optimization-Based Keyword Extraction

After the pre-processing, keyword extraction is performed to reduce the time consumption by minimizing the dimensionality of the words in the given text. Keyword extraction is the process of choosing the significant keywords and removing the others from the text. COPLELM utilizes the censored regressive cockroach swarm optimization to select the optimal keywords for accurate detection. The cockroach swarm algorithm is a metaheuristic optimization inspired by the foraging behavior. Each cockroach moves toward the position of another cockroach with better fitness. At first, the populations of cockroaches, i.e., words, are initialized in the search space.

$$W_j = \{W_1, W_2, \dots, W_b\} \tag{5}$$

After that, the fitness 'f' of each word is measured based on word frequency.

$$WF = \left[\frac{NFO}{b}\right] \tag{6}$$

Where WF denotes a word frequency, b indicates the total number of words, NFO represents the number of words that frequently occur in the given text samples.

To compute fitness, the optimization algorithm utilizes censored regression for finding more optimal keywords and removing the others. Censored regression is a machine learning technique where the dependent variables, i.e., keywords, are partially observed based on a threshold or limit.

$$F = \begin{cases} WF > \delta, & Selectkeyword \\ WF < \delta, & Removeword \end{cases} \tag{7}$$

Where F indicates a fitness, WF denotes a word frequency, δ indicates a threshold. If the word frequency score exceeds a predefined threshold 'δ', the word is selected as a keyword. Otherwise, it is considered a normal word. Based on the fitness evaluation, the current best keywords are selected and executing three foraging behaviors of cockroaches, such as chase-swarming, dispersing, and ruthless. At first, the algorithm begins by applying the chase-swarming strategy, during which each cockroach updates its position by moving in the direction of the global best optimal solution. Therefore, the new position of the cockroach is updated as follows,

$$X_{i+1} = X_i + S * r * [0.5 * |X_{best} - X_i|] \tag{8}$$

Where, X_{i+1} indicates a newly updated position of the cockroach, X_i denotes a current best solution, 'S' denotes a step size, r indicates a random number [0, 1], X_{best} denotes the global best solution.

From (8), the function $0.5 * |X_{best} - X_i|$ refers to a Jensen-Shannon divergence for measuring the relationship between the global best solution 'X_{best}' and the current local best solution 'X_i'. Next, the optimization algorithm applies the dispersion behavior of the cockroach to preserve diversity within the population. During this phase, some cockroaches move away from current solutions to explore unexplored areas of the search space. Dispersion behavior is expressed as,

$$X_{i+1} = X_i + q\ (1, d) \tag{9}$$

Where X_{i+1} specifies an updated solution, X_i indicates a current position, $q\ (1, d)$ denotes a d-dimensional random vector. Ruthless behavior is executed by replacing the randomly chosen individual to attain a better solution.

$$X_R = X_{best} \tag{10}$$

X_R designates a randomly chosen individual and X_{best} denotes the global best position. Process continued until the maximum number of iterations was reached. A global optimal keyword is extracted for accurate depression detection.

Figure 3 depicts the flowchart of the optimization technique aimed at selecting the optimal keywords for detecting the depression and non depression tweets samples. The population of words in the tweets is initialized within the defined search space.

For each word within this population, the fitness is evaluated based on a word frequency measure. Subsequently, the current best word is identified, and the positions of words are updated accordingly based on three behaviors of the cockroach.

This iterative process gets repeated until it reaches the maximum iterations. After reaching the maximum iteration, the optimal keyword is selected for further processing and removal of the other words.

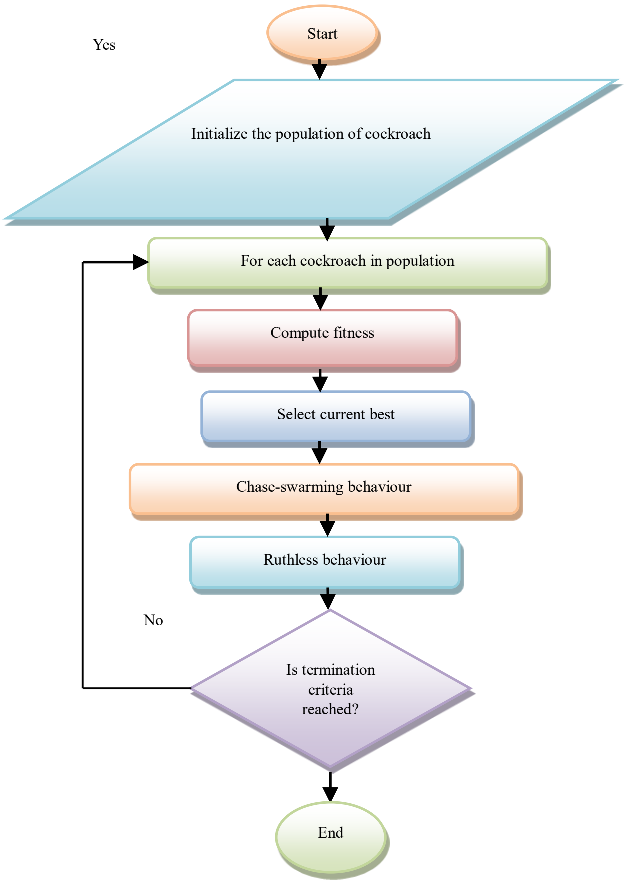

Fig. 3 Flow chart of the censored regressive cockroach swarm optimization-based keyword extraction

3.2.3. Depression Detection

Finally, the depression detection is performed in the next hidden layer with the extracted optimal keywords by means of applying a Gestalt pattern matching. This pattern-matching technique analyzes the contextual relationship between the words. Based on analyses, depression and non-depression Twitter texts are accurately detected. In this layer, a large lexical database is utilized to extract similar related words based on a set of optimal keywords. This contextual or similar word helps to expand the original keyword set, capturing related meanings to recognize conceptually similar expressions. The extracted contextual words are used to generate a vector model.

$$K_{opt} = [RW_1, RW_2, RW_3 \dots . RW_k] \quad (11)$$

$$M = \left[K_{opt}, RW_1, RW_2, RW_3 \dots . RW_k\right] \quad (12)$$

Where, K_{opt} denotes an optimally selected keyword from the previous hidden layer, $[RW_1, RW_2, RW_3 \dots . RW_k]$ denotes semantically related words for the particular keyword that was extracted from the large lexical database, M denotes a word vector constructed by combining the keyword with its semantically related terms. After forming the word vector, the Gestalt pattern matching technique is employed for analyzing the word vector in the training set and words in the depression testing set. The pattern matching is expressed as follows,

$$\varphi_{pm} = 2 * \frac{[W_M \cap W_t]}{|W_M||W_t|} \quad (13)$$

$$\varphi_{pm} = \begin{cases} 1, & depression \\ 0, & nodepression \end{cases} \quad (14)$$

Where φ_{pm} a pattern matching score, W_M denotes words in the vector model, W_t indicates words in the testing set, $|W_M|$ and $|W_t|$ represents the cardinality of the set, i.e., the number of strings in the words W_M and W_t. The pattern matching score produces output values between 0 and 1. If the score is equal to 1, the corresponding tweet is classified as a depression. Otherwise, the tweet is classified as no depression. In an extreme learning network, the sigmoid activation function is typically applied in the output layer for binary classification problems.

$$Y = F_s[\vartheta_{ho} * h_t)] \quad (15)$$

Where Y indicates a predicted classification output of a deep extreme learning model, F_s represents a sigmoid activation function, ϑ_{ho} denotes a weight between the hidden and output layer.

From the output results, the depression and non depression text tweet samples are correctly predicted with minimal time consumption as well as error rate.

4. Experimental Settings

Experimental assessment of the COPLELM model and CBA [1] and Deep temporal model [2] are implemented in Python using Depression: Twitter Dataset + Feature Extraction**.**

5. Performance Analysis

Performance of the COPLELM model and existing [1, 2] models are discussed using various metrics.

5.1. Performance Comparison Analysis of Depression Prediction Accuracy

It represents the ratio of accurately classified tweets as depressed and non-depressed to the total number of tweet samples collected from the dataset. The calculation is performed as follows:

$$DDA = \left(\frac{Tps+Tng}{Tps+Tng+Fps+Fng}\right) * 100 \quad (16)$$

Table 3 illustrates that the COPLELM model improves the performance of accuracy by 5% and 3% than the [1, 2].

Table 3. Comparison of depression prediction accuracy

Number of tweet samples	Depression prediction accuracy (%)		
	COPLELM	CBA	Deep temporal model
2000	98.5	94	96
4000	98.06	93.56	95.66
6000	98.36	93.78	95.05
8000	98.07	94.05	95.88
10000	98.74	93.12	95.32
12000	98.69	93.45	95.74
14000	98.03	93.11	95.33
16000	98.78	93.1	95.74
18000	98.06	93.08	95.87
20000	97.45	93	95.88

5.2. Performance Comparison of Precision

It denotes the proportion of correctly identified depressed and non-depressed tweets among all tweets. Precision is calculated using the following formula,

$$PRC = \left(\frac{Tps}{Tps+Fps}\right) \quad (17)$$

In Figure 4, precision is improved by 5% and 3% than the [1, 2] applying the COPLELM model.

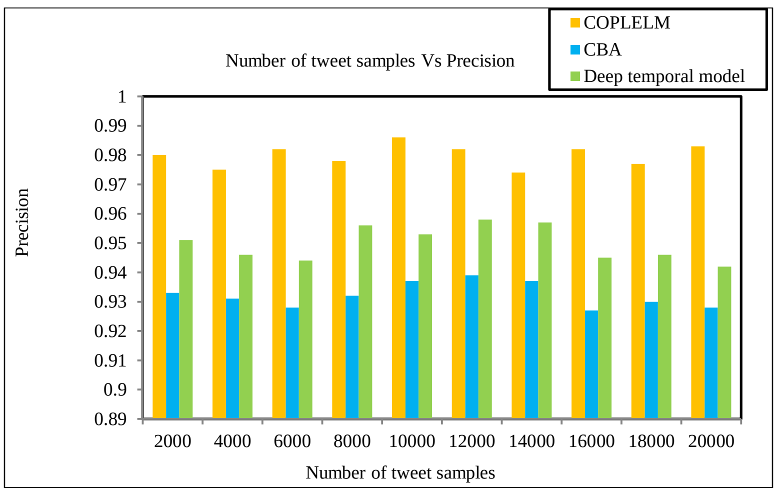

Fig. 4 Graphical illustration of precision

5.3. Performance Comparison of Recall

It indicates the proportion of correctly identified depressed and non-depressed tweets relative to the total number of actual tweets in each category within the dataset. This measure, also known as sensitivity, is calculated as:

$$REC = \left(\frac{Tps}{Tps+Fng}\right) \quad (18)$$

Table 4 shows that recall is improved by 5% and 3% than the existing [1, 2].

Table 4. Comparison of recall

Number of tweet samples	COPLELM	CBA	Deep temporal model
2000	0.990	0.951	0.970
4000	0.985	0.948	0.963
6000	0.992	0.944	0.962
8000	0.986	0.943	0.965
10000	0.993	0.94	0.964

12000	0.994	0.947	0.959
14000	0.991	0.943	0.966
16000	0.995	0.942	0.969
18000	0.991	0.945	0.964
20000	0.987	0.938	0.961

5.4. Performance Comparison of F1 score

It is referred to as the F-measure; this classification performance metric provides a harmonic mean of precision and recall. It is defined using the following formula,

$$F1\_Score = 2 * \left(\frac{PRE*REC}{PRE*REC}\right) \tag{19}$$

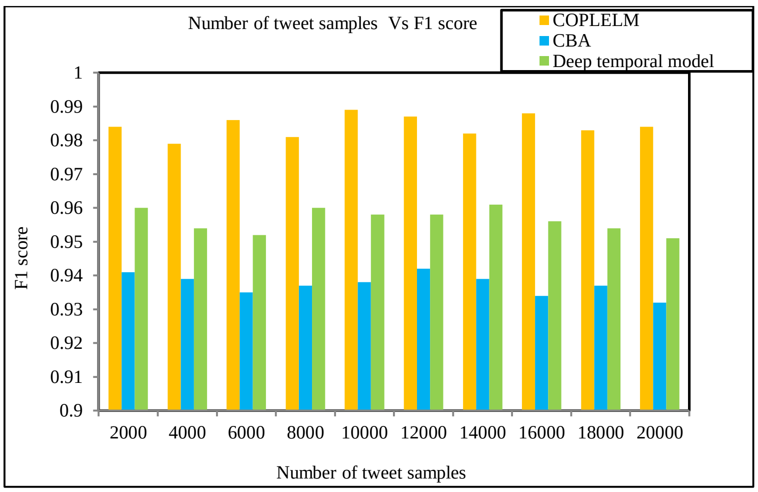

Fig. 5 Graphical illustrations of the F1 score

Figure 5 illustrates that the F1-score of the COPLEL is improved by 5% and by 3% than the [1, 2].

5.5. Performance Comparison of Specificity

In depression prediction, specificity evaluates the model's ability to correctly identify no depressed tweet samples as true negatives. The specificity metric is calculated as follows,

$$SPE = \left(\frac{Tng}{Tng*Fps}\right) \tag{20}$$

Table 5 demonstrates that COPLELM increases the specificity by 6% compared to [1] and 3% compared to [2].

Table 5. Comparison of specificity

Number of tweet samples	COPLELM	CBA	Deep temporal model
2000	0.979	0.927	0.948
4000	0.975	0.925	0.936
6000	0.976	0.922	0.942
8000	0.977	0.92	0.936
10000	0.969	0.911	0.935
12000	0.975	0.918	0.933
14000	0.972	0.917	0.938
16000	0.97	0.92	0.94
18000	0.968	0.927	0.943
20000	0.962	0.925	0.944

5.6. Performance Comparison of Error Rate

Error rate refers to the incorrect predictions made by the model. It is calculated as follows,

$$ER = \left(\frac{Fps+Fng}{Tps+Tng+Fps+Fng}\right) \tag{21}$$

Figure 6 demonstrates that the COPLELM minimizes the error rate by 74% compared to [1] 61% compared to [2].

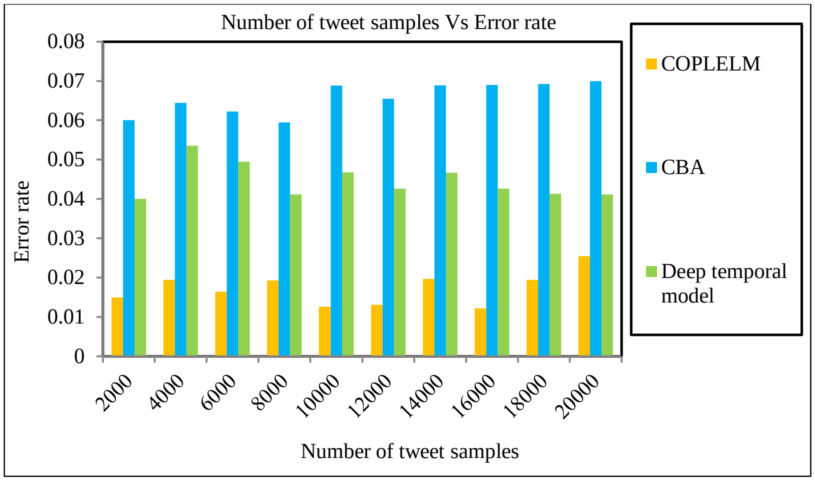

Fig. 6 Graphical illustration of error rate

5.7. Performance Comparison of Depression Prediction Time

This evaluation metric measures the amount of time consumed by the algorithm to identify depressed and non-depressed tweet samples.

The time consumption is determined using the following calculation:

$$DPT = \sum_{i=1}^{n} TS_i * TM(DD) \tag{22}$$

Table 6. Comparison of depression prediction time

Number of tweet samples	Depression prediction time (ms)		
	COPLELM	**CBA**	**Deep temporal model**
2000	34	56	40
4000	41.2	60.3	47.6
6000	46.6	65	55.8
8000	52.3	73.6	62.3
10000	58.7	85.5	68.9
12000	65.8	89	75.6
14000	77.8	95.6	86.8
16000	86.5	105.7	96.5
18000	93.6	113.5	106.7
20000	106.2	126.6	120.6

Table 6 demonstrates that the depression prediction time was reduced by 26% and 13% than the [1, 2].

5.8. Confusion Matrix

A confusion matrix is an important assessment tool in classification tasks, used to analyze the performance of the proposed COPLELM model.

	Total number of tweet samples: 20000	Actual value: Positive	Negative	
Predicted value	Positive	TP=9700	FP=300	10000
	Negative	FN=200	TN=9800	10000
		9900	10100	

Fig. 7 Confusion metrics using the COPLELM model

It summarizes the results of depression prediction by utilizing the counts of true positives, false positives, true negatives, and false negatives.

Figure 7 illustrates confusion matrix outcomes for COPLELM, evaluated using 20,000 tweet samples.

5.9. Performance of the ROC curve

The ROC (Receiver Operating Characteristic) curve is utilized to evaluate the performance of the proposed COPLELM model, with existing CBA [1] and deep temporal model [2] in predicting the positive class. It visually represents the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR).

Figure 8 illustrates that the COPLELM model achieves better prediction accuracy and more reliable classification in depression prediction tasks, outperforming the existing [1, 2] methods.

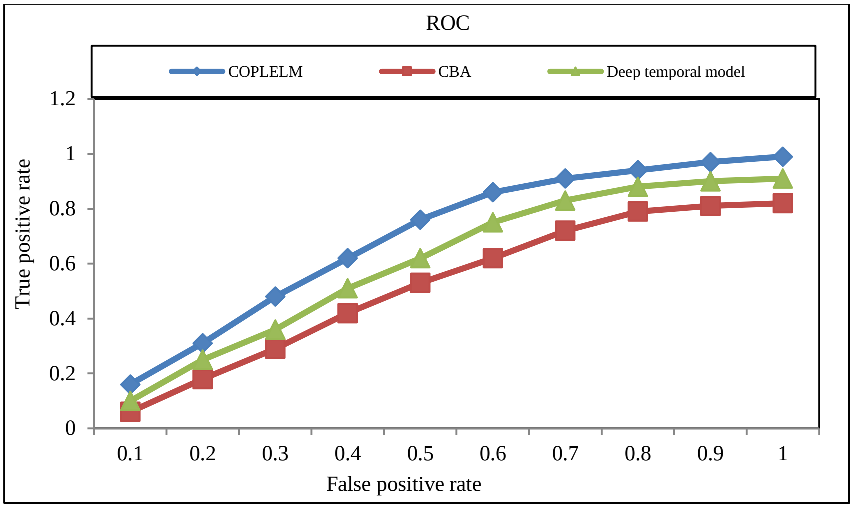

Fig. 8 Graphical results of the ROC curve

6. Discussion

This research work compares the proposed COPLELM model with conventional CBA [1] and Deep temporal model [2] using Depression: Twitter Dataset + Feature Extraction based on different parameters, namely, depression prediction accuracy, precision, recall, F1 score, specificity, error rate, prediction time, confusion matrix, and ROC analysis with respect to the number of tweets samples. In this approach, a novel method, the COPLELM model, has been developed for accurate depression prediction with minimal time consumption. The text pre-processing was performed to minimize the complexity of the depression prediction. Subsequently, the optimal keyword extraction processes further reduced the time consumption of depression prediction. Finally, the progressive extreme learning machine enhanced the accuracy of depression by minimizing the error rate. The results confirm that the proposed COPLELM method achieves improved 4% accuracy, precision, recall, and F1 score, with reduce the 68% error rate and 20% depression prediction time when compared to existing CBA [1] and Deep temporal model [2] methods using Depression: Twitter Dataset + Feature Extraction dataset.

7. Conclusion

Depression is the most widespread and serious mental health disorder, and it leads to tiredness, poor concentration, feelings of sadness, and so on. The COPLELM model has been developed for accurate depression prediction with minimal time consumption. The proposed model is compared with the two existing methods (CBA and the deep temporal model). The research findings of the proposed COPLELM model achieve a higher 4% accuracy, precision, recall, and F1 score. The proposed COPLELM model considerably minimizes the 68% error rate and 20% depression prediction time compared to existing methods. In the future, the proposed model is further suggested to use advanced DL to improve the accuracy of depression detection and reduce time by means of various datasets.

Conflict of Interest

The authors declare no potential conflict of interest.

Consent for Publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Data Availability Statement

Data are available within the article. The data have been gathered from (https://www.kaggle.com/datasets/infamouscoder/mental-health-social-media).

Acknowledgements

None

References

[1] Joel Philip Thekkekara, Sira Yongchareon, and Veronica Liesaputra, "An Attention-based CNN-BiLSTM Model for Depression Detection on Social Media Text," *Expert Systems with Applications*, vol. 249, pp. 1-9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Nawshad Farruque et al., "Deep Temporal Modelling of Clinical Depression through Social Media Text," *Natural Language Processing Journal*, vol. 6, pp. 1-14, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Amna Qasim et al., "Detection of Depression Severity in Social Media Text using Transformer-based Models," *Information*, vol. 16, no. 2, pp. 1-23, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[4] Eunchae Lim et al., "A Lightweight Approach based on Cross-Modality for Depression Detection," *Computers in Biology and Medicine*, vol. 186, pp. 1-14, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[5] Nawshad Farruque et al., "Depression Symptoms Modelling from Social Media Text: An LLM Driven Semi-Supervised Learning Approach," *Language Resources and Evaluation*, vol. 58, no. 3, pp. 1013-1041, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Muhammad Asad Abbas et al., "Novel Transformer based Contextualized Embedding and Probabilistic Features for Depression Detection from Social Media," *IEEE Access*, vol. 12, pp. 54087-54100, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Shakir Khan, and Salihah Alqahtani, "Hybrid Machine Learning Models to Detect Signs of Depression," *Multimedia Tools and Applications*, vol. 83, no. 13, pp. 38819-38837, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Filza Rehmani et al., "Depression Detection with Machine Learning of Structural and Non-Structural Dual Languages," *Healthcare Technology Letters*, vol. 11, no. 4, pp. 218-226, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Bayode Ogunleye, Hemlata Sharma, and Olamilekan Shobayo, "Sentiment Informed Sentence BERT-Ensemble Algorithm for Depression Detection," *Big Data and Cognitive Computing*, vol. 8, no. 9, pp. 1-16, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Umesh Kumar Lilhore et al., "Prevalence and Risk Factors Analysis of Postpartum Depression at Early Stage using Hybrid Deep Learning Model," *Scientific Reports*, vol. 14, no. 1, pp. 1-24, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Ngumimi Karen Iyortsuun et al., "Additive Cross-Modal Attention Network (ACMA) for Depression Detection based on Audio and Textual Features," *IEEE Access*, vol. 12, pp. 20479-20489, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Yongfeng Tao et al., "DepMSTAT: Multimodal Spatio-Temporal Attentional Transformer for Depression Detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 7, pp. 2956-2966, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Wenli Zhang et al., "Depression Detection using Digital Traces on Social Media: A Knowledge-Aware Deep Learning Approach," *Journal of Management Information Systems*, vol. 41, no. 2, pp. 546-580, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Cao Xin, and Lailatul Qadri Zakaria, "Integrating Bert with CNN and BiLSTM for Explainable Detection of Depression in Social Media Contents," *IEEE Access*, vol. 12, pp. 161203-161212, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] M. Kiran Myee et al., "Detection of Depression in Social Media Posts using Emotional Intensity Analysis," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 16207-16211, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Vandana, Nikhil Marriwala, and Deepti Chaudhary, "A Hybrid Model for Depression Detection using Deep Learning," *Measurement: Sensors*, vol. 25, pp. 1-10, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] İsmail Baydili, Burak Tasci, and Gülay Tasci, "Deep Learning-based Detection of Depression and Suicidal Tendencies in Social Media Data with Feature Selection," *Behavioural Sciences*, vol. 15, no. 3, pp. 1-28, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[18] T. Srajan Kumar, M. Narayanan, and Harikrishna Kamatham, "A Deep Learning Framework with a Hybrid Model for Automatic Depression Detection in Social Media Posts," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4, pp. 3217-3231, 2024. [Publisher Link]

[19] Eliseo Bao, Anxo Pérez, and Javier Parapar, "Explainable Depression Symptom Detection in Social Media," *Health Information Science and Systems*, vol. 12, no. 1, pp. 1-18, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Muhammad Rizwan et al., "Depression Intensity Classification from Tweets using Fast Text based Weighted Soft Voting Ensemble," *Computers, Materials and Continua*, vol. 78, no. 2, pp. 2047-2066, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Loukas Ilias, Spiros Mouzakitis, and Dimitris Askounis, "Calibration of Transformer-based Models for Identifying Stress and Depression in Social Media," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 2, pp. 1979-1990, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Shiwen Zhou, and Masnizah Mohd, "Mental Health Safety and Depression Detection in Social Media Text Data: A Classification Approach based on a Deep Learning Model," *IEEE Access*, vol. 13, pp. 63284-63297, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[23] Biodoumoye George Bokolo, and Qingzhong Liu, "Advanced Comparative Analysis of Machine Learning and Transformer Models for Depression and Suicide Detection in Social Media Texts," *Electronics*, vol. 13, no. 20, pp. 1-22, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] Xiaolong Li et al., "Hypergraph Neural Network for Multimodal Depression Recognition," *Electronics*, vol. 13, no. 22, pp. 1-20, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Dip Kumar Saha et al., "Ensemble of Hybrid Model based Technique for Early Detecting of Depression based on SVM and Neural Networks," *Scientific Reports*, vol. 14, no. 1, pp. 1-18, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[26] Tumpa Rani Shaha et al., "Feature Group Partitioning: An Approach for Depression Severity Prediction with Class Balancing using Machine Learning Algorithms," *BMC Medical Research Methodology*, vol. 24, no. 1, pp. 1-18, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[27] Ziling Ruan et al., "Automatic Depression Detection among Higher Education Students based on DeepFM," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1-10, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[28] Shahid Munir Shah et al., "Advancing Depression Detection on Social Media Platforms through Fine-Tuned Large Language Models," *Online Social Networks and Media*, vol. 46, pp. 1-16, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[29] Esteban A. Ríssola et al., "On the Incidence of Depression Symptoms on Social Media," *Journal of Computational Social Science*, vol. 8, no. 2, pp. 1-30, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[30] Robinson Narvaez Burbano, Oscar Mauricio Caicedo Rendon, and Carlos. A. Astudillo, "An Encoder-Only Transformer Model for Depression Detection from Social Network Data: The DEENT Approach," *Applied Sciences*, vol. 15, no. 6, pp. 1-17, 2025. [CrossRef] [Google Scholar] [Publisher Link]