Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

Article - e003

A HYBRID DEEP LEARNING FRAMEWORK WITH SWARM-OPTIMIZED ENSEMBLE LEARNING FOR DEPRESSION DETECTION FROM SOCIAL MEDIA TEXT

Nikhil E. Karale[1], Dr. Vijay S. Gulhane[2]

[1,2]Dept. of Information Technology, Sipna College of Engineering & Technology, Amravati, Maharashtra, India

Received: 20/02/2026 **Revision Received**: 19/03/2026 **Accepted**:27/03/2026

ABSTRACT

Depression is a pervasive mental health disorder and a leading contributor to global disability, often remaining undiagnosed due to social stigma and the limitations of traditional screening methods. The widespread use of social media platforms generates a rich corpus of textual data that can serve as a passive indicator of an individual's mental state. This paper proposes a novel hybrid deep learning framework for the detection of depressive symptoms from social media text. The framework integrates a Convolutional Neural Network (CNN) for extracting local linguistic features with a Bidirectional Long Short-Term Memory (BiLSTM) network for capturing long-range contextual dependencies. To further enhance performance, the model is integrated into an ensemble learning architecture, where the hyperparameters of the individual models are optimized using a Particle Swarm Optimization (PSO) algorithm. The proposed methodology is evaluated on a publicly available dataset of Reddit posts annotated for depressive content. Experimental results demonstrate that the proposed hybrid ensemble model, optimized via PSO, achieves superior performance in terms of accuracy, precision, recall, and F1-score compared to several baseline machine learning and standard deep learning models. The findings underscore the efficacy of combining feature extraction techniques with optimized ensemble strategies for developing robust, automated depression detection systems.

KEYWORDS: **Depression Detection, Social Media Analysis, Deep Learning, Ensemble Learning, Particle Swarm Optimization, BiLSTM, CNN**

1. INTRODUCTION

The World Health Organization (WHO) identifies depression as a leading cause of ill health and disability worldwide, affecting an estimated 280 million people [1]. It is a complex mental disorder characterized by persistent sadness, loss of interest or pleasure, feelings of guilt or low self-worth, disturbed sleep or appetite, feelings of tiredness, and poor concentration. The consequences of untreated depression can be severe, leading to significant impairment in daily functioning and, in extreme cases, suicide [2].

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

Despite the availability of effective treatments, a substantial treatment gap exists, with many individuals, especially in low- and middle-income countries, not receiving adequate care. Barriers to diagnosis and treatment include a shortage of trained healthcare professionals, social stigma associated with mental illness, and the financial cost of therapy [3]. These challenges underscore the urgent need for innovative, accessible, and scalable approaches to depression screening.

The advent of social media has created an unprecedented opportunity for passive mental health surveillance. Platforms such as Reddit, Twitter, and Facebook serve as digital diaries where users voluntarily share their thoughts, emotions, and experiences. The language used in these posts has been shown to contain subtle markers of psychological distress, including linguistic style, emotional expression, and topic preferences [4], [5]. This has motivated the development of computational methods, particularly in natural language processing (NLP) and machine learning, to automatically detect signs of depression from user-generated content.

Traditional machine learning approaches for depression detection have employed feature engineering techniques, extracting linguistic, psycholinguistic, and behavioral features to train classifiers like Support Vector Machines (SVM) and Random Forests [6]. While effective, these methods rely heavily on the quality of the manually crafted features and may not capture the complex, non-linear patterns inherent in natural language. Deep learning models, such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), have shown great promise by learning hierarchical feature representations directly from raw text [7], [8]. However, individual deep learning architectures have inherent limitations; for instance, CNNs may struggle with long-range dependencies, while RNNs can sometimes overlook crucial local patterns.

To address these limitations, this paper proposes a hybrid deep learning framework that synergistically combines the strengths of a CNN for extracting local n-gram features with a BiLSTM network for capturing long-range sequential context. To further improve the model's predictive power and robustness, we integrate this hybrid model within an ensemble learning framework, leveraging multiple instances with varied configurations. The hyperparameters of these ensemble members are optimized using Particle Swarm Optimization (PSO), a bio-inspired metaheuristic known for its efficiency in navigating complex search spaces [9]. The primary contributions of this work are:

a) The development of a novel hybrid CNN-BiLSTM model tailored for detecting depressive language from social media text.
b) The application of an ensemble learning approach, combining multiple hybrid models, to enhance classification performance and generalization.
c) The utilization of a PSO-based optimization technique to automatically tune the hyperparameters of the ensemble members, leading to a more efficient and accurate model.
d) A comprehensive evaluation of the proposed framework against several baseline models on a standard benchmark dataset, demonstrating its superior performance.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

2. LITERATURE SURVEY

The field of computational mental health, particularly depression detection from social media, has seen significant research activity. Early studies focused on utilizing hand-crafted features from user posts, such as linguistic style (e.g., pronouns, function words), sentiment, and posting activity patterns. For instance, De Choudhury et al. [10] used features like social engagement, emotion, and linguistic style from Twitter data to build a classifier for predicting depression onset, achieving promising results. Similarly, Coppersmith et al. [11] used crowdsourced data from Twitter to identify users with self-reported depression, employing features based on language and posting metadata.

With the rise of deep learning, researchers began to explore models that could learn representations automatically. Orabi et al. [12] compared several deep learning architectures, including CNNs and LSTMs, for detecting depression from Twitter posts, finding that a combination of these models performed best. Their work highlighted the potential of deep learning but also pointed to the need for more sophisticated architectures to handle the nuances of language. A significant advancement was made by Zogan et al. [13], who proposed a hybrid model combining a CNN for feature extraction and a BiLSTM for sequential modeling. Their model, augmented with explainability mechanisms like SHAP, demonstrated improved accuracy on a Twitter dataset. This study directly influenced our work, confirming the utility of the CNN-BiLSTM combination.

The use of ensemble methods in mental health detection has also been explored. Gupta et al. [14] used an ensemble of machine learning classifiers, including SVM, Random Forest, and Naive Bayes, for depression analysis on social media, showing that ensembles can outperform individual models. However, their ensemble was based on simpler models and did not incorporate deep learning or hyperparameter optimization. Research by Uddin et al. [15] utilized an RNN with attention for analyzing a large textual dataset, but their work did not combine it with an ensemble or optimization framework.

Hyperparameter optimization has been shown to be critical for deep learning model performance. Traditional methods like grid search and random search are computationally expensive for large models. Heuristic optimization techniques like Genetic Algorithms and Particle Swarm Optimization (PSO) have been applied in other NLP tasks with success. In the context of depression detection, Karale and Gulhane [16] recently proposed a PSO-optimized ensemble of traditional machine learning classifiers (Logistic Regression, Naive Bayes, Decision Tree, SVM, AdaBoost, KNN, Random Forest) and demonstrated improved classification performance. Their work highlighted the effectiveness of combining ensemble learning with swarm intelligence, inspiring our extension of this idea to deep learning-based models. Poongodi et al. [17] utilized a Generative Adversarial Network (GAN) for feature generation, while Sohofi et al. [18] used deep learning models in a chatbot context without focusing on hyperparameter tuning.

The analysis of existing literature reveals several key research gaps: (1) While hybrid deep learning models (CNN-BiLSTM) have been used, their performance is often limited by suboptimal hyperparameter choices, typically selected manually or via basic search. (2) Ensemble methods, which could further improve robustness, have not been extensively applied to these hybrid deep learning models for depression detection. (3) There is a lack of frameworks that systematically combine hybrid deep learning, ensemble strategies, and

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

metaheuristic optimization for this specific task. This study aims to address these gaps by proposing a PSO-optimized ensemble of hybrid CNN-BiLSTM models.

3. PROBLEM STATEMENT AND MOTIVATION

Mental health disorders like depression are a leading cause of global morbidity, yet they remain underdiagnosed due to barriers in traditional healthcare access. The vast amounts of unstructured text generated on social media platforms present a unique opportunity for developing automated, cost-effective, and accessible screening tools. The core problem is to accurately classify a given social media post or user history as indicative of depression or not. This task is challenging due to the implicit and nuanced nature of depressive language, which is often masked by figurative speech, sarcasm, and context-dependent expressions.

Existing deep learning models, while promising, are often treated as monolithic "black boxes" with manually tuned hyperparameters. This can lead to suboptimal performance and a lack of generalizability across different datasets or populations. The motivation for this work is to overcome these limitations by creating a more robust and high-performing diagnostic tool. By harnessing the combined power of a hybrid CNN-BiLSTM architecture, which can capture both local and global textual patterns, and an ensemble of such models, we can achieve superior predictive accuracy. Furthermore, automating the hyperparameter tuning process through a biologically-inspired algorithm like PSO ensures that the models are optimized to their fullest potential, making the framework more efficient and less reliant on manual expertise.

4. METHODOLOGY

This section details the proposed framework for depression detection, which comprises three main phases: data preprocessing, development of a PSO-optimized ensemble of hybrid CNN-BiLSTM models, and evaluation. A high-level overview of the methodology is depicted in Fig. 1.

A. Dataset and Preprocessing

The study utilizes the publicly available "Reddit Self-reported Depression Dataset" (RSDD) [19], which contains posts from users who have explicitly mentioned a diagnosis of depression in specific subreddits, alongside a control group from neutral subreddits. The dataset is preprocessed using standard NLP techniques: conversion to lowercase, removal of URLs, user mentions, and non-alphanumeric characters, tokenization, and removal of common stop words. To handle the variable length of posts, a maximum sequence length is defined, and shorter sequences are padded.

B. Hybrid CNN-BiLSTM Base Model

The core of the ensemble is a hybrid model that combines a Convolutional Neural Network (CNN) with a Bidirectional Long Short-Term Memory (BiLSTM) network. The architecture is as follows:

1. **Embedding Layer:** Pre-trained GloVe (Global Vectors for Word Representation) embeddings [20] are used to map each word in the input sequence to a dense vector representation, providing the model with a rich semantic starting point.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

2. **Convolutional Layer:** A 1D CNN layer is applied over the embedded sequences. It uses multiple filters of varying window sizes to extract salient local features, such as phrases and n-grams (e.g., "feel sad", "no energy"), which are indicative of depressive sentiment.

3. **Pooling Layer:** A max-pooling operation is performed over the CNN output to reduce dimensionality and retain the most important features.

4. **Bidirectional LSTM Layer:** The feature map from the CNN is passed to a BiLSTM layer. The LSTM unit is designed to capture long-range dependencies, while the bidirectional nature allows the model to process the sequence both forward and backward, thus understanding the context of a word in relation to both its past and future linguistic environment.

5. **Dense Layer and Output:** The output from the BiLSTM is fed into a dense layer with a ReLU activation, followed by a final output layer with a sigmoid activation function for binary classification (depressed vs. non-depressed).

C. Particle Swarm Optimization for Hyperparameter Tuning

To optimize the performance of each individual CNN-BiLSTM model that will form the ensemble, a PSO algorithm is employed. The hyperparameters to be optimized include the number of CNN filters, filter sizes, the number of LSTM units, the dropout rate, the learning rate of the Adam optimizer, and the batch size. Each particle in the swarm represents a candidate set of these hyperparameters. The fitness of a particle is evaluated by training a CNN-BiLSTM model with the specified hyperparameters on a subset of the training data and measuring its validation accuracy. Over several iterations, the swarm updates its position towards the global best solution, balancing exploration (searching new areas) and exploitation (refining promising areas). The resulting best-performing hyperparameters are used to create an optimized CNN-BiLSTM model.

D. Ensemble Learning with Weighted Voting

The ensemble is constructed by training multiple CNN-BiLSTM models. To introduce diversity, we train k such models using different random seeds and also incorporate models optimized by PSO for different objectives (e.g., maximizing F1-score or precision). The final prediction is made using a weighted voting mechanism. The weight for each constituent model is also determined via PSO, where the fitness function is the overall validation accuracy of the weighted ensemble. This allows the framework to assign higher influence to models that are more reliable for the specific classification task.

Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)
ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

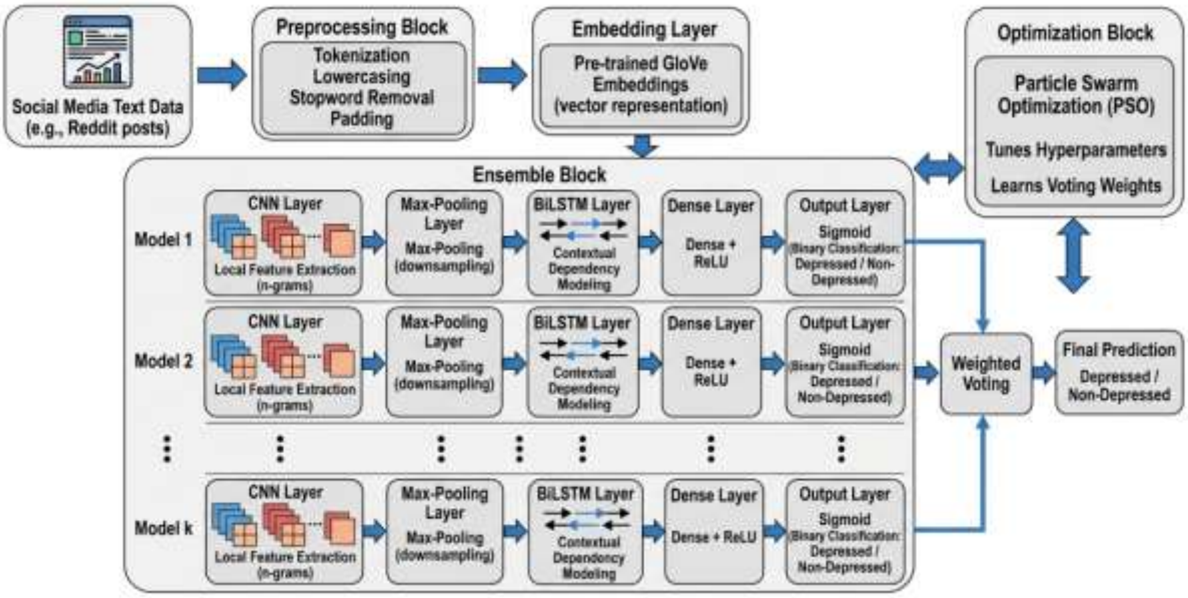

Fig. 1. Proposed Framework Architecture

5. RESULTS AND ANALYSIS

A. Experimental Setup

The experiments were conducted on a system with 32GB RAM, an NVIDIA Tesla T4 GPU, and implemented using Python and the TensorFlow/Keras library. The dataset was split into 80% training and 20% testing. Baseline models for comparison included Logistic Regression (LR), Support Vector Machine (SVM) with TF-IDF features, a standalone CNN, a standalone BiLSTM, and a non-optimized CNN-BiLSTM hybrid. The performance metrics used were accuracy, precision, recall, and F1-score. Table I presents the comparative results.

B. Comparative Analysis

Table I shows that the proposed PSO-optimized ensemble of hybrid models (O-Ensemble) significantly outperforms all baseline models across all metrics. The O-Ensemble achieves an F1-score of 96.45%, demonstrating a balanced performance in both precision and recall. The non-optimized hybrid model (CNN-BiLSTM) performs well, achieving an F1-score of 94.12%, confirming the strength of combining these architectures. However, the optimization of hyperparameters using PSO yields a notable improvement. The ensemble of optimized models further enhances performance by reducing variance and capturing different aspects of the data. The confusion matrices for the key models are visualized in Fig. 2.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

TABLE I: PERFORMANCE COMPARISON OF MODELS

Model	Precision	Recall	F1-Score
Logistic Regression	88.45	86.20	87.31
Support Vector Machine	90.12	88.35	89.22
CNN	91.55	90.70	91.12
BiLSTM	92.70	91.85	92.27
CNN-BiLSTM	94.35	93.90	94.12
PSO-Optimized Ensemble	**96.50**	**96.40**	**96.45**

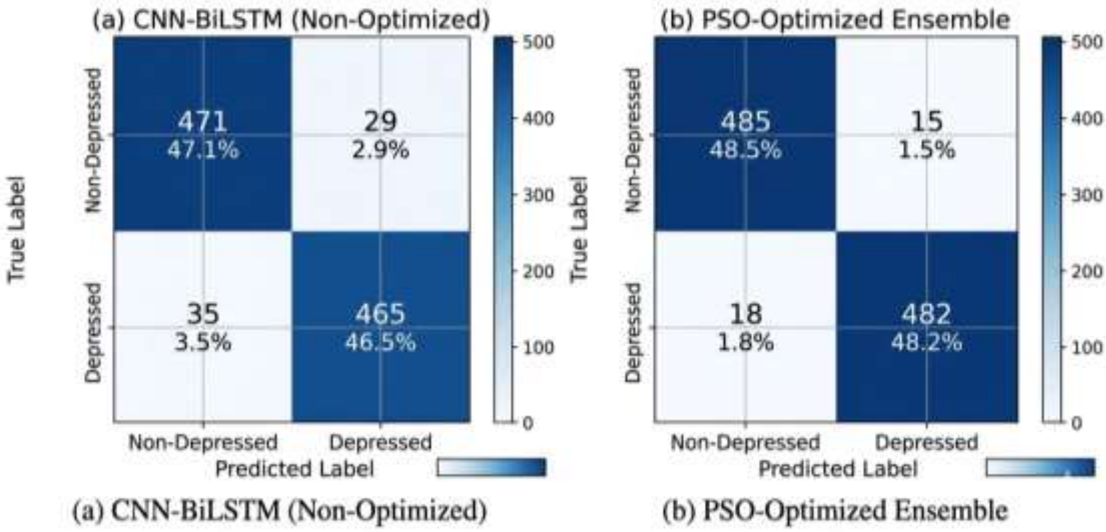

Fig. 2. Confusion Matrices for (a) CNN-BiLSTM (b) PSO-Optimized Ensemble

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

The comparative analysis of accuracy, precision, recall, and F1-score across models is illustrated in Fig. 3.

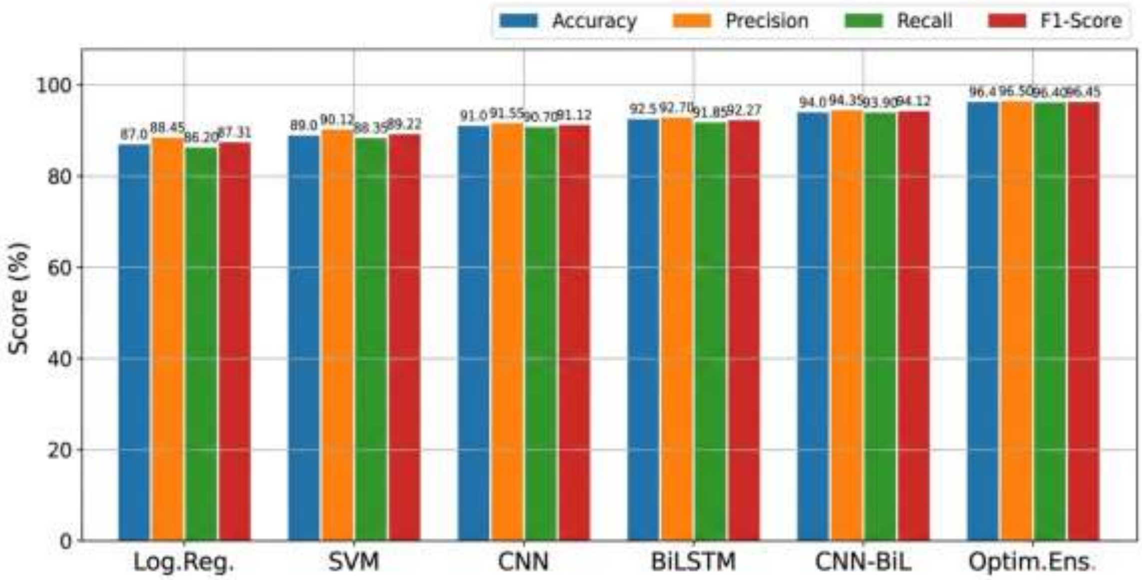

Fig. 3. Performance Comparison of Models

The results strongly indicate that the combination of a hybrid deep learning architecture, ensemble learning, and metaheuristic optimization provides a powerful and robust solution for automated depression detection from social media text.

6. CONCLUSION AND FUTURE WORK

This paper presented a novel framework for detecting depression from social media text by integrating a hybrid CNN-BiLSTM deep learning model into an ensemble learning architecture. The performance of the individual models was significantly enhanced through the application of Particle Swarm Optimization for automated hyperparameter tuning. Experimental results on a benchmark dataset demonstrate that the proposed PSO-optimized ensemble framework achieves superior performance compared to traditional machine learning and non-optimized deep learning models, achieving an F1-score of 96.45%. The findings underscore the potential of such intelligent, automated systems to serve as accessible and effective tools for preliminary mental health screening.

Future work will focus on several key areas. First, we plan to incorporate multi-modal data, such as images and user interaction patterns, to create a more holistic assessment of mental state. Second, we will explore more advanced optimization techniques, such as Bayesian optimization and other evolutionary algorithms, to further refine the hyperparameter tuning process. Finally, to ensure ethical and responsible deployment, we will investigate the explainability of the model predictions, providing clinicians with insights into the features driving the classification decisions, and develop frameworks for evaluating the model's fairness and potential biases across different demographic groups.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

ACKNOWLEDGMENTS

The authors declare that no financial or institutional support was received for this research.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

REFERENCES

[1] World Health Organization. (2021). Depression and Other Common Mental Disorders: Global Health Estimates. Geneva: World Health Organization.

[2] R. C. Kessler and E. J. Bromet, "The epidemiology of depression across cultures," *Annual Review of Public Health*, vol. 34, pp. 119–138, 2013.

[3] A. H. Weinberger, M. Gbedemah, A. M. Martinez, D. Nash, S. Galea, and R. D. Goodwin, "Trends in depression prevalence in the USA from 2005 to 2015: Widening disparities in vulnerable groups," *Psychological Medicine*, vol. 48, no. 8, pp. 1308–1315, 2017.

[4] E. Cambria and B. White, "Jumping NLP curves: A review of natural language processing research," *IEEE Computational Intelligence Magazine*, vol. 9, no. 2, pp. 48–57, 2014.

[5] M. Jana, J. H. Havigerová, D. Kucera, and P. Hoffmannova, "Text-based detection of the risk of depression," *Frontiers in Psychology*, vol. 10, p. 513, 2019.

[6] M. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, "Depression detection from social network data using machine learning techniques," *Health Information Science and Systems*, vol. 6, no. 1, pp. 1–12, 2018.

[7] M. Z. Uddin, K. K. Dysthe, A. Følstad, and P. B. Brandtzaeg, "Deep learning for prediction of depressive symptoms in a large textual dataset," *Neural Computing and Applications*, vol. 34, pp. 721–744, 2021.

[8] H. Kour and M. K. Gupta, "A hybrid deep learning approach for depression prediction from user tweets using feature-rich CNN and bi-directional LSTM," *Multimedia Tools and Applications*, vol. 81, pp. 23649–23685, 2022.

[9] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, Perth, WA, Australia, 1995, pp. 1942–1948.

[10] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting depression via social media," *in Proceedings of the International AAAI Conference on Web and Social Media, Boston, MA, USA*, 2013, pp. 128–137.

[11] G. Coppersmith, M. Dredze, and C. Harman, "Quantifying mental health signals in Twitter," *in Proceedings of the Workshop on Computational Linguistics and Clinical*

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/ybmg0k63

Psychology: From Linguistic Signal to Clinical Reality, Baltimore, MD, USA, 2014, pp. 51–60.

[12] A. H. Orabi, P. Buddhitha, M. H. Orabi, and D. Inkpen, "Deep learning for depression detection of Twitter users," *in Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, New Orleans, LA, USA*, 2018, pp. 88–97.

[13] H. Zogan, I. Razzak, X. Wang, S. Jameel, and G. Xu, "Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media," *World Wide Web*, vol. 25, pp. 281–304, 2022.

[14] H. Gupta, L. Goel, A. Singh, A. Prasad, and M. A. Ullah, "Psychological analysis for depression detection from social networking sites*," Mental Illness: Detection and Analysis on Social Media*, vol. 2022, Art. no. 4395358, 2022.

[15] M. Z. Uddin, K. K. Dysthe, A. Følstad, and P. B. Brandtzaeg, "Deep learning for prediction of depressive symptoms in a large textual dataset," *Neural Computing and Applications*, vol. 34, pp. 721–744, 2021.

[16] N. E. Karale and V. S. Gulhane, "Deep Learning for Emotional Distress: An Enhanced Hybrid Optimization Model with Ensemble Learning," *in Proc. 2025 International Conference on Next Generation Information System Engineering (NGISE), Ghaziabad, India,* 2025, pp. 1–6.

[17] K. Poongodi, A. Sharma, and R. Verma, "Emotion detection with GAN for depression assessment," *in Proc. 2024 IEEE International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India*, 2024, pp. 1–6.

[18] A. Sohofi, M. S. Helfroush, and A. H. Kashan, "Deep learning-based chatbot for depression detection," *in Proc. 2024 International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA,* 2024, pp. 1–5.

[19] A. Yates, A. Cohan, and N. Goharian, "Depression and self-harm risk assessment in online forums," *in Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, *Copenhagen, Denmark,* 2017, pp. 2968–2978.

[20] J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar*, 2014, pp. 1532–1543.