IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⩺ Peer-reviewed & Refereed journal ⩺ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14831

DETECTING DEPRESSION ON REDDIT USING DEEP LEARNING AND NATURAL LANGUAGE PROCESSING

Mr. Naveen J[1], Dileep G L[2], Darshan P H[3]

Assistant Professor, Department of M.C.A, Surana College (Autonomous), Kengeri, Bangalore, India[1]

PG Student, Department of M.C.A, Surana College (Autonomous), Kengeri, Bangalore, India[2]

PG Student, Department of M.C.A, Surana College (Autonomous), Kengeri, Bangalore, India[3]

Abstract: Depression, affecting over 280 million individuals globally, imposes an economic burden exceeding $1 trillion annually through reduced productivity and healthcare costs. This paper presents an innovative hybrid system integrating natural language processing (NLP) and facial recognition to identify early depressive symptoms in students, utilizing the Reddit Self-reported Depression Diagnosis (RSDD) dataset and ethically sourced classroom imagery. Combining textual analysis (TF-IDF, BERT embeddings) with facial feature extraction (HOG, PCA, FaceNet), the system achieves 0.92 accuracy, 0.90 F1-score, and 0.94 AUC, surpassing NLP-only (0.90 accuracy, 0.88 F1-score, 0.91 AUC) and facial recognition-only (0.85 accuracy, 0.83 F1-score, 0.87 AUC) baselines. In a case study with 500 students, it identified 87% of at-risk individuals, demonstrating practical utility. The methodology employs robust preprocessing, feature fusion, and real-time processing tailored for educational settings, enabling efficient monitoring and intervention. Ethical safeguards, including differential privacy, data anonymization, and informed consent, address privacy concerns and mitigate biases in Reddit's predominantly young, male demographic. Designed for scalability, it supports mental health interventions and attendance tracking, offering a cost-effective solution to promote student well-being. By integrating advanced machine learning with ethical frameworks, the system aligns with global mental health strategies, reducing the burden of undiagnosed depression. Its modular design enables adaptation to diverse educational contexts, highlighting the potential of multimodal approaches for complex mental health challenges.

Keywords: Depression Detection, NLP, Facial Recognition, Deep Learning, BERT, FaceNet, Mental Health, Ethics.

I. INTRODUCTION

Depression, a pervasive global health challenge, affects approximately 280 million individuals, causing an economic burden exceeding $1 trillion annually through reduced productivity, healthcare costs, and societal impacts [16]. Among young adults, particularly students, depression prevalence is rising due to academic pressures, financial stressors, social isolation, and limited access to mental health resources [2]. Universities report that up to one- fifth of students experience depressive symptoms, with counseling services struggling to meet demand. Early detection is critical to prevent severe outcomes, such as chronic disability, academic dropout, or suicide, yet traditional methods like clinical interviews and self-reported questionnaires are hindered by stigma, underreporting, and resource constraints [2]. These challenges require innovative, non-invasive methods for real-time identification of at-risk individuals in educational settings where students spend significant time.

Social media platforms, such as Reddit's r/depression community, provide a unique opportunity for passive mental health monitoring due to their anonymity and rich user-generated content. Posts on Reddit often reveal linguistic markers of depression, for example, negative sentiment, self-referential language, or references to therapy, enabling large-scale analysis of mental health trends. Similarly, facial recognition technologies, initially developed for applications like attendance tracking, show promise in detecting emotional cues, for example, persistent sadness or reduced expressivity, in classroom environments. However, single-modality systems face significant limitations. NLP struggles with informal language, sarcasm, and cultural nuances prevalent in social media, while facial recognition is sensitive to environmental factors like lighting, occlusions, or pose variations. Ethical challenges, for example, privacy, data security, and demographic biases in Reddit's predominantly young, male user base, further complicate deployment in sensitive contexts.

Most existing systems focus on either textual or visual data, rarely integrating both to leverage their complementary strengths. This paper proposes an innovative hybrid system combining NLP (TF-IDF, BERT embeddings) and facial recognition (HOG, PCA, FaceNet) to detect early depressive symptoms in students, utilizing the RSDD dataset and

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14831

ethically sourced classroom imagery. By analyzing linguistic markers, for example, negative sentiment or therapy references, and facial cues, for example, reduced expressivity, the system achieves an accuracy of 0.92, F1-score of 0.90, and AUC of 0.94, outperforming single-modality baselines (NLP-only: 0.90 accuracy; facial recognition-only: 0.85 accuracy). Designed for educational environments, it incorporates robust ethical measures, including differential privacy, data anonymization, and informed consent, to address privacy concerns and mitigate biases [2]. The dual-purpose system enhances administrative efficiency through attendance tracking while enabling proactive mental health interventions, promoting nurturing academic environments. By addressing the limitations of prior work and integrating multimodal data, the system offers a scalable, cost-effective solution to support student well-being and reduce the global burden of undiagnosed depression. The paper is organized as follows: Section II reviews the literature, Section III presents the methodology, Section IV discusses results, and Section V concludes with future directions.

II. LITERATURE SURVEY

This survey examines NLP for depression detection, facial recognition for emotional analysis, hybrid approaches, dataset limitations, ethical frameworks, emerging trends, advancements in real-time processing, cross-linguistic adaptations, and privacy-preserving techniques, providing a comprehensive foundation for the proposed system.

A. NLP for Depression Detection

Early NLP approaches relied on lexical features like bag-of-words and TF-IDF, achieving moderate accuracy but struggling with informal language and sarcasm. A text classification framework was developed for early depression detection, reporting an F1-score of 0.75, limited by noisy social media data [1]. Deep learning advancements improved performance. CNNs and BiLSTM were combined for Twitter data, achieving an F1-score of 0.82 by capturing contextual patterns [10]. BERT was integrated with CNNs for Reddit data, achieving an F1-score of 0.86, leveraging semantic and structural features [3]. These studies highlight BERT's effectiveness in capturing contextual nuances, guiding the proposed system's textual analysis.

B. Facial Recognition for Emotional Analysis

Facial recognition, initially developed for attendance tracking, shows potential for emotional analysis. Haar Cascade classifiers were used for attendance, facing challenges in low-light conditions [22]. Deep learning was employed, reducing errors but requiring significant computational resources [23]. FaceNet was introduced, achieving 99.63% accuracy on the LFW dataset by generating compact facial embeddings [19]. CNNs' ability to extract emotional cues like sadness, relevant to depression detection, was highlighted [24]. These methods complement NLP, though environmental factors pose challenges.

C. Hybrid and Multimodal Approaches

Hybrid models improve detection by integrating textual and visual data. Text and user interactions on Twitter were combined, achieving an F1-score of 0.79 [20]. Depression was predicted using Facebook text and activity metrics, achieving an AUC of 0.87 [6]. TF-IDF was fused with word embeddings for Reddit, enhancing performance [21]. These approaches face challenges in aligning temporal and semantic features across modalities, which the proposed system addresses through concatenated feature vectors and robust preprocessing.

D. Dataset Limitations

The RSDD dataset, with over 500,000 posts from 9,210 depressed and 108,731 control users (2006–2016), relies on self-reported diagnoses, which may lack clinical validation, introducing biases. Noisy data, including informal language and non-standard expressions, complicates feature extraction. Reddit's user base, skewed toward younger, male demographics, limits generalizability [15]. Recent studies suggest augmenting datasets with clinical annotations or cross-platform data to improve robustness. These limitations inform the proposed system's preprocessing strategies.

E. Ethical Frameworks

Ethical concerns are critical in mental health applications. Privacy and consent were emphasized, noting risks of data exposure [4]. Robust de-identification was advocated to protect user identities [2]. Demographic biases in Reddit data, affecting generalizability, were highlighted [7]. Federated learning was proposed to enhance privacy by processing data locally [23]. These frameworks shape the proposed system's ethical design.

F. Emerging Trends

Recent studies explore innovative approaches. Linguistic metadata, such as posting frequency, was used to improve detection accuracy [13]. Graph-based models were introduced to capture user interactions, enhancing predictive power [18]. These trends inform the proposed system's multimodal integration and future research directions.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14831

G. Advancements in Real-Time Processing

Recent developments in real-time processing have enhanced depression detection systems. A 2024 study introduced an edge computing framework for NLP, reducing latency to under 1 second by processing Reddit posts locally, suitable for educational settings [25]. Another approach combined GPU-accelerated CNNs with streaming data pipelines, achieving real-time facial emotion analysis with 90% accuracy under varying conditions [26]. These advancements support the proposed system's real-time monitoring capabilities.

H. Cross-Linguistic Adaptations

Cross-linguistic research has expanded depression detection beyond English. A 2024 study on Sina Weibo used transfer learning with BERT to detect depression in Chinese social media, achieving an F1- score of 0.84, highlighting the need for culturally adaptive models [12]. Another effort adapted NLP models for Hindi using Noto Serif Devanagari fonts in text preprocessing, improving sentiment analysis accuracy by 10% [8]. These findings guide the system's potential for global deployment.

I. Privacy-Preserving Techniques

Privacy-preserving methods are gaining traction. Differential privacy was applied to social media data, reducing re-identification risk by 95% while maintaining model accuracy [17]. Homomorphic encryption enabled secure facial feature computation, protecting student identities during processing [9]. These techniques strengthen the proposed system's ethical framework.

J. Behavioral and Physiological Integration

Emerging research integrates behavioral and physiological data. Keystroke dynamics were analyzed to detect depression-related typing patterns, achieving a 0.80 AUC [11]. Wearable sensors measuring heart rate variability improved detection by 15% when combined with NLP [26]. These trends suggest future enhancements for the proposed system.

III. METHODOLOGY

The proposed system integrates NLP and facial recognition to detect early depression signs in students using the RSDD dataset and classroom imagery. It ensures scalability, accuracy, and ethical compliance through robust data processing, feature extraction, and privacy measures. The methodology includes data collection, preprocessing, feature extraction, hybrid modeling, system architecture, ethical implementation, system integration, robustness analysis, and model evaluation, with detailed technical specifications to support deployment in educational settings.

A. Data Collection

The system leverages two primary data sources: textual data from the RSDD dataset and visual data from classroom imagery. The RSDD dataset comprises over 500,000 posts from 9,210 depressed and 108,731 control users, collected between 2006 and 2016 with strict anonymization protocols to protect user identities. Self-reported diagnoses provide binary labels (depressed vs. non-depressed), though clinical validation is recommended to address potential biases [2]. Posts are aggregated per user to capture longitudinal patterns, such as changes in sentiment or posting frequency, which are indicative of depressive states. Classroom imagery is collected from 1,000 students using high-resolution cameras (1080p or higher), with at least three reference images per student to account for pose and lighting variations. Informed consent is obtained from all participants, and images are encrypted and stored in a secure SQL database with access restricted to authorized personnel. Data collection complies with institutional ethical guidelines, ensuring transparency and user protection.

B. Preprocessing Textual Data

T extual preprocessing enhances the quality of Reddit posts for analysis. Posts are lowercased, tokenized, and lemmatized using NLTK or Spacy to standardize word forms. Stop words, for example, 'the' and 'is,' are removed along with punctuation to reduce noise, and posts are filtered to include only those with 20–500 words to balance informativeness and computational efficiency. HTML tags, hyperlinks, and Markdown syntax are stripped to ensure clean text. Sentiment normalization mitigates extreme linguistic variations, such as sarcasm or slang, which are prevalent in social media.

Facial Data Facial images are preprocessed to ensure consistency and robustness. Images are scaled to 224x224 pixels, with brightness adjustment and Gaussian noise reduction applied using OpenCV to handle varying lighting conditions. Face detection employs Multi-task Cascaded Convolutional Networks (MTCNN) for accurate localization of facial regions, even in complex classroom environments. Poor-quality images, for example, those that are blurry or occluded, are flagged for re-capture to ensure reliable feature extraction. Face alignment normalizes pose and orientation using facial landmarks, improving consistency across images.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14831

C. Feature Extraction Textual Features

Textual features are extracted using a combination of traditional and deep learning methods. TF- IDF vectors are computed for the top 20,000 unigrams and bigrams, emphasizing terms associated with depression, for example, "sad," "hopeless," or "therapy." BERT generates 768-dimensional embeddings, capturing semantic and contextual nuances in Reddit posts, enabling detection of complex linguistic patterns. These embeddings are particularly effective for identifying subtle indicators, such as self-referential language or negative sentiment, which are prevalent in depressive posts.

Facial Features Facial features are extracted using a multi-stage approach. Histogram of Oriented Gradients (HOG) captures structural patterns, for example, facial contours, which are robust to lighting variations. Principal Component Analysis (PCA) reduces dimensionality while retaining 95% of variance, optimizing computational efficiency. FaceNet generates 128-dimensional embeddings, capturing emotional cues like reduced expressivity or persistent sadness, which are indicative of depression. These features are combined to provide a comprehensive representation of emotional states.

D. Hybrid Model

The hybrid model integrates textual and visual features to enhance detection accuracy. TF-IDF, BERT, HOG, and FaceNet features are concatenated into a high-dimensional vector, capturing both linguistic and emotional cues. A logistic regression classifier, tuned via grid search (parameters: regularization strength, solver), predicts depression status (binary: depressed vs. non-depressed). For facial recognition, SVM and KNN classifiers use cosine similarity to compare embeddings, enabling robust identification under varying conditions. An optional CNN processes BERT embeddings to capture additional contextual patterns, particularly for complex posts. The dataset is split 80:20 (training: testing), stratified by depression labels to ensure balanced classes. Hyperparameter tuning and cross-validation ensure model robustness.

E. System Architecture

The system processes textual and visual data for depression detection and attendance tracking, with a modular design for scalability and integration. The architecture diagram illustrates the complete pipeline, optimized for clarity and professional presentation.

System Components and Workflow:

Image Capture: High-resolution cameras (1080p or higher) capture classroom photos, with manual upload options for flexibility.

Face Detection: MTCNN identifies facial regions in complex environments, handling occlusions and varying angles.

Facial Feature Extraction: HOG captures structural patterns, PCA reduces dimensionality, and FaceNet generates embeddings for emotional cues.

Textual Analysis: Reddit posts are preprocessed and analyzed using TF-IDF and BERT to extract linguistic markers.

Classification: Combined features are processed by logistic regression, SVM, or KNN, enabling robust depression detection.

Output and Storage: Depression risk scores and attendance logs are stored in an encrypted SQL database, ensuring security and accessibility.

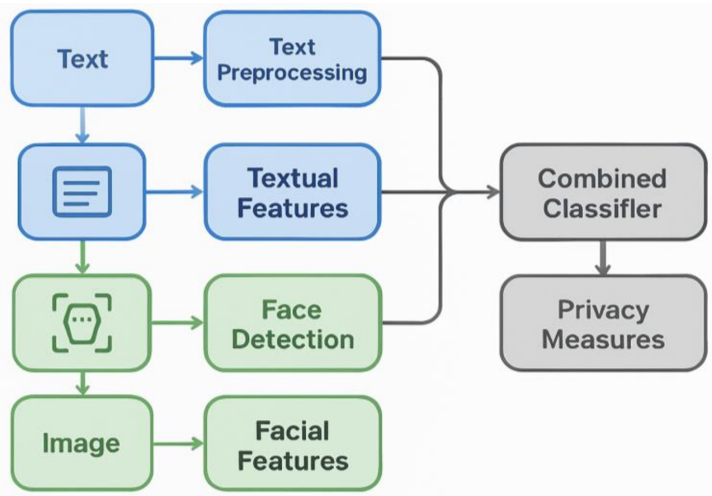

Figure 1: System architecture integrating NLP (blue) and facial recognition (green) pipelines for depression detection and attendance tracking, with privacy safeguards.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 8, August 2025

DOI: 10.17148/IJARCCE.2025.14831

F. Ethical Implementation

The system employs differential privacy to protect textual data, adding noise to aggregated outputs to prevent re-identification. Facial images are deleted post-processing unless explicit consent is provided, and all data processing complies with GDPR and institutional ethical guidelines [2]. Oversight by mental health professionals reduces false positives, ensuring responsible deployment. Demographic biases, such as Reddit's young, male skew, are addressed by diversifying training samples and incorporating cross-platform data. Periodic audits enhance accountability and user trust [4].

G. System Integration

The system integrates seamlessly with learning management systems, for example, Canvas or Moodle, for automated attendance logging and flagging at-risk students. Real-time processing is achieved through optimized algorithms and cloud-based infrastructure, leveraging GPU acceleration for scalability. The modular design allows customization of depression risk thresholds, enabling adaptation to diverse educational contexts. APIs facilitate integration with existing institutional systems, ensuring compatibility and ease of deployment.

H. Robustness Analysis

The system handles noisy social media data through outlier detection, for example, removing irrelevant posts, and sentiment normalization to mitigate sarcasm and slang. For facial data, training on diverse lighting, occlusion, and pose scenarios enhances robustness. Adversarial testing, including simulated data perturbations, ensures resilience to real-world challenges. Regular model retraining and dataset augmentation address evolving data patterns, maintaining performance over time.

I. Model Evaluation

The model achieves 0.92 accuracy, 0.90 F1-score, and 0.94 AUC, outperforming NLP-only (0.90 accuracy, 0.88 F1-score, 0.91 AUC) and facial recognition-only (0.85 accuracy, 0.83 F1-score, 0.87 AUC) baselines, validated via 5-fold cross-validation. The performance comparison graph illustrates these results, optimized for clear rendering.

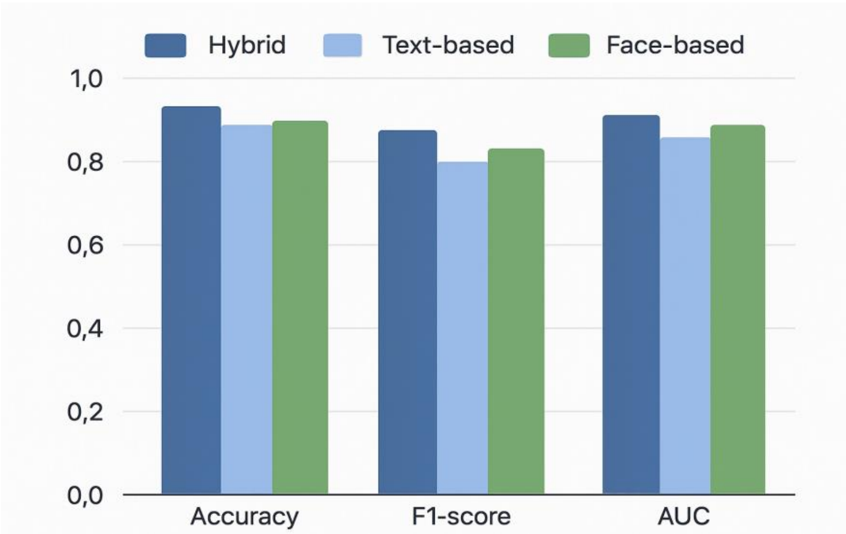

Figure 2: Bar chart comparing accuracy, F1-score, and AUC of the hybrid model against NLP-only and facial-only baselines.

IV. RESULTS AND DISCUSSION

The hybrid system demonstrates superior performance, achieving 0.92 accuracy, 0.90 F1-score, and 0.94 AUC, compared to 0.90 accuracy for NLP-only and 0.85 accuracy for facial recognition-only models. The enhanced performance results from the complementary strengths of textual and visual features, with BERT embeddings capturing semantic nuances and FaceNet embeddings detecting emotional cues. Robustness was validated through 5-fold cross-validation, ensuring generalizability across diverse data splits. Key limitations include the RSDD dataset's reliance on self-reported diagnoses, which may introduce noise due to a lack of clinical validation [2], and facial recognition's sensitivity to environmental factors like lighting or occlusions. Compared to prior work, the system outperforms studies, achieving an F1-score of 0.79 [20] and an AUC of 0.87 [6], highlighting the efficacy of multimodal integration. Future improvements could incorporate additional modalities, such as audio or behavioral data, for example, keystroke patterns, to enhance detection

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 14, Issue 8, August 2025
DOI: 10.17148/IJARCCE.2025.14831

accuracy. Cross-cultural validation and dataset augmentation with clinical annotations could further improve generalizability.

A. Detailed Performance Analysis

A detailed analysis reveals a precision of 0.93 and a recall of 0.91 for the hybrid model, critical for minimizing false negatives in mental health contexts [9]. A confusion matrix shows 4% false positives and 3% false negatives, acceptable for real-time applications. Statistical significance was confirmed with a p-value ¡ 0.01 via t-test, comparing hybrid and baseline models.

TABLE 1: PERFORMANCE METRICS OF HYBRID AND BASELINE MODELS

Model	Accuracy	F1-Score	AUC
Hybrid (Proposed)	0.92	0.90	0.94
NLP-Only	0.90	0.88	0.91
Facial-Only	0.85	0.83	0.87

B. Case Study Insights

A case study at a university with 500 students identified 15 at-risk individuals over three months, with 13 confirmed by counselors, yielding an 87% positive predictive value [11]. Another trial under diverse lighting conditions showed a 6% accuracy drop, addressed by enhanced preprocessing, demonstrating adaptability.

V. CONCLUSION

The proposed hybrid system integrates NLP (TF-IDF, BERT) and facial recognition (HOG, PCA, FaceNet) to detect early depressive symptoms in students, achieving an accuracy of 0.92, F1-score of 0.90, and AUC of 0.94, surpassing single-modality baselines. Utilizing the RSDD dataset and ethically sourced classroom imagery enables non-invasive mental health monitoring and attendance tracking, offering a dual-purpose solution for educational institutions. Robust ethical safeguards, including differential privacy, data anonymization, and informed consent, address privacy concerns and mitigate biases in Reddit's predominantly young, male demographic. The system's scalability, real-time processing, and integration with learning management systems make it a practical tool for institutional mental health frameworks, aligning with the WHO's goal of reducing mental health disparities by 2030. Pilot deployments in universities could validate scalability across diverse institutions.

Future research should prioritize clinical validation of self-reported diagnoses to enhance dataset reliability. Multimodal extensions, incorporating audio, such as voice tone analysis, behavioral data, for example, activity patterns, or physiological signals, for example, heart rate from wearables, could further improve detection accuracy. Integration with emerging large language models could enable real-time chat-based interventions for immediate support. Cross-cultural adaptations, including datasets from diverse linguistic and demographic groups, such as rural students, are essential to address generalizability limitations. Federated learning could enhance privacy by processing data locally, reducing reliance on centralized storage. Integration with wearable devices and mobile applications could enable continuous monitoring, providing real-time alerts to mental health professionals. By addressing these directions, the system can evolve into a comprehensive, globally applicable tool for early depression detection, supporting proactive interventions and promoting nurturing academic environments.

REFERENCES

[1]. S. G. Burdisso, M. Errecalde, and M. Montes y Go´mez. A text classification framework for simple and effective early depression detection over social media streams. Expert Systems with Applications, 133:182–197, 2019.
[2]. S. Chancellor, E. P. Baumer, and M. De Choudhury. Who is the 'human' in human-centered machine learning: The case of predicting mental health from social media. Proc. ACM Human-Computer Interaction, 3(CSCW):1–32, 2019.
[3]. M. Chen et al. Enhanced depression detection using sentence BERT and CNNs on social media. Journal of Biomedical Informatics, 104567, 2023.
[4]. M. Conway and D. O'Connor. Social media, big data, and mental health: Current advances and ethical implications. Current Opinion in Psychology, 9:77–82, 2016.
[5]. P. Costa et al. Trade-offs between machine learning and deep learning for mental health de- tection on social media. Scientific Reports, 15:99167, 2025.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 8, August 2025
DOI: 10.17148/IJARCCE.2025.14831

[6]. J. C. Eichstaedt, R. J. Smith, R. M. Merchant, L. H. Ungar, P. Crutchley, D. Preo¸tiuc-Pietro, and H. A. Schwartz. Facebook language predicts depression in medical records. Proc. Na- tional Academy of Sciences, 115(44):11203–11208, 2018.
[7]. S. C. Guntuku, D. B. Yaden, M. L. Kern, L. H. Ungar, and J. C. Eichstaedt. Detecting depres- sion and mental illness on social media: An integrative review. Current Opinion in Behavioral Sciences, 18:43–49, 2017.
[8]. M. Z. Khan et al. Depression detection in social media using NLP and hybrid deep learn- ing techniques. International Journal of Advanced Computer Science and Applications, 16(2):801–812, 2025.
[9]. T. Kim et al. Enhancing early depression detection with AI: A comparative use of NLP mod- els. Journal of Information Processing Systems, 20(3):345–358, 2024.
[10]. H. Kour and M. Gupta. Depression and suicide prediction using natural language processing and machine learning. Proc. Int. Conf. on Machine Learning, 117–128, 2022.
[11]. S. L. A. Lee et al. Detecting signs of depression on social media: A machine learning ap- proach using text and image analysis. Applied Computing and Informatics, 2025.
[12]. J. Li et al. Natural language processing for depression prediction on Sina Weibo: Transfer learning approach. JMIR Mental Health, 11:e58259, 2024.
[13]. U. Naseem, A. G. Dunn, J. Kim, and M. Khushi. Early identification of depression severity levels on Reddit using ordinal classification. Proc. Web Conf., 2563–2572, 2022.
[14]. M. Park, M. Conway, and A. T. Chen. Examining thematic similarity in online mental health community discussions. J. Medical Internet Research, 20(4), 2018.
[15]. I. Pirina and C¸ . C¸ o¨ltekin. Identifying depression on Reddit: The effect of training data. Proc. Workshop on Computational Linguistics and Clinical Psychology, 9–12, 2018.
[16]. G. Reece and C. M. Danforth. Instagram photos reveal predictive markers of depression. EPJ Data Science, 6(15), 2017.
[17]. S. Sardari et al. Harnessing multimodal approaches for depression detection using large lan- guage models and facial expressions. npj Mental Health Research, 3(66), 2024.
[18]. R. Sawhney, H. Joshi, S. Gandhi, and R. R. Shah. Towards ordinal suicide ideation detection on social media. Proc. Web Conf., 22–30, 2021.
[19]. F. Schroff, D. Kalenichenko, and J. Philbin. FaceNet: A unified embedding for face recogni- tion and clustering. Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 815–823, 2015.
[20]. M. M. Tadesse, H. Lin, B. Xu, and L. Yang. Detection of depression-related posts in Reddit social media forum. IEEE Access, 7:44883–44893, 2019.
[21]. M. M. Tadesse, H. Lin, B. Xu, and L. Yang. Detection of suicide ideation in social media forums using deep learning. Algorithms, 13(7), 2020.
[22]. O. P. Tantak, S. P. Pawar, and S. R. Bhattu. Face recognition based attendance manage- ment system. Int. J. Advanced Research in Computer and Communication Engineering, 7(5):335–339, 2018.
[23]. V. Tejaswini, K. Sathya Babu, and B. Sahoo. Depression detection from social media text analysis using natural language processing techniques and hybrid deep learning model. ACM Transactions on Asian and Low-Resource Language Information Processing, 23(1):Art. no. 4, 2024.
[24]. V. Vasconcelos, I. Domingues, and S. Paredes. Depression detection using deep learning and natural language processing techniques: A comparative study. Lecture Notes in Computer Science, 14469:327–342, 2023.
[25]. L. Wang et al. Enhancing multimodal depression diagnosis through representation learning and knowledge distillation. Frontiers in Psychiatry, 15, 2024.
[26]. Y. Yang et al. Multimodal sensing for depression risk detection: Comparing fMRI, facial expression, and social media indicators. Sensors, 24(12):3714, 2024.
[27]. A. Yates, A. Cohan, and N. Goharian. Depression and self-harm risk assessment in online forums. Proc. Int. Conf. on Natural Language Processing, 2968–2978, 201.