ISSN: 2581-3404 (Online) IF: 5.86 (SJIF)
IJIRTM, Volume-10, Issue-5, October-2026

Depression Prediction using Machine Learning Algorithms : Survey & Discussion

Krishna Kant Magariya[1]
Research Scholar, Computer Science & Engineering Department[1]
Maharana Pratap College of Technology, Gwalior, India[1,2]
Priya S Jain[2]
Assistant Professor, Computer Science & Engineering Department[2]

Abstract. *Depression is a major mental health concern that requires effective methods for early detection and prediction. This review examines the application of machine learning (ML) and deep learning (DL) techniques for depression prediction using 16 research papers. Various algorithms, including Logistic Regression, Support Vector Machine, Decision Tree, Random Forest, XGBoost, CNN, LSTM, and ensemble models, have been investigated using psychological, demographic, clinical, EEG, facial expression, and social media data. The reviewed studies demonstrate the potential of these techniques to identify depression-related patterns and improve prediction performance. Explainable AI methods, such as SHAP, also support the interpretation of model predictions. However, challenges related to data quality, class imbalance, model generalization, privacy, and clinical validation remain. This review highlights the need for reliable, interpretable, and generalizable ML-based frameworks to support early depression screening and assist healthcare professionals in decision-making.*

Keywords: Depression Prediction, Machine Learning, Deep Learning, Artificial Intelligence.

Introduction

Depression is a complex mental health disorder that affects the emotional, psychological, and physical well-being of individuals. It is generally characterized by persistent sadness, reduced interest in daily activities, changes in sleeping and eating patterns, fatigue, difficulty concentrating, feelings of worthlessness, and reduced motivation. In severe cases, depression may lead to self-harm, suicidal thoughts, social isolation, and difficulties in maintaining personal and professional relationships. The consequences of depression extend beyond individuals and can affect families, educational institutions, workplaces, and healthcare systems. Therefore, early identification and effective prediction of depression have become important research areas in healthcare and computational intelligence. Traditional depression diagnosis primarily relies on clinical interviews, psychological assessments, standardized questionnaires, and observations made by mental health professionals. Commonly used assessment instruments include the Patient Health Questionnaire-9 (PHQ-9), Depression Anxiety Stress Scale (DASS), Center for Epidemiologic Studies Depression Scale (CES-D), and Burns Depression Checklist (BDC). These instruments provide structured information about an individual's psychological condition and help healthcare professionals evaluate the severity of depressive symptoms. However, conventional assessment approaches may be influenced by subjective reporting, differences in individual experiences, social stigma, cultural factors, and the willingness of individuals to disclose their mental health

conditions. Furthermore, limited access to trained professionals and healthcare facilities can delay diagnosis and treatment, particularly in resource-constrained communities. These limitations have encouraged researchers to investigate automated and data-driven approaches for identifying depression.

In recent years, Artificial Intelligence (AI) has emerged as an important technology in healthcare because of its ability to analyze large volumes of complex information and identify meaningful patterns. Machine Learning, a major branch of AI, enables computational systems to learn relationships from historical data and generate predictions without requiring explicitly programmed rules for every decision. In depression prediction, ML algorithms analyze demographic, psychological, behavioral, socioeconomic, and clinical characteristics to identify patterns associated with depressive symptoms. These techniques can support the classification of individuals into different depression-related categories, such as normal, mild, moderate, and severe, depending on the dataset and assessment methodology.

Several researchers have investigated traditional ML algorithms for depression prediction. Sharif et al. examined Support Vector Machine, Decision Tree, and Light Gradient Boosting methods using 604 surveyed participants and the Burns Depression Checklist. Their findings demonstrated that Decision Tree achieved the highest accuracy of 95.66%, followed by Light Gradient Boosting at 94.58% and SVM at 91.48%. These findings illustrate the potential of conventional classification algorithms for questionnaire-based depression assessment. Similarly, Zhai et al. (2024) developed predictive models using XGBoost, Random Forest, Decision Tree, and Logistic Regression to identify anxiety and depressive disorders among 61,619 college students from 133 higher education institutions. Their research demonstrated the usefulness of ML models for identifying students at increased risk and highlighted financial circumstances, campus belonging, disability status, and age as important predictive variables.

Literature Review

Several researchers have explored machine learning and deep learning approaches for depression prediction using different datasets and assessment methods. Sharif et al. investigated SVM, Decision Tree, and LightGBM using the Burns Depression Checklist and reported the highest accuracy of 95.66% with Decision Tree [1]. Zhai et al. developed predictive models using XGBoost, Random Forest, Decision Tree, and Logistic Regression to identify anxiety and depressive disorders among college students [2]. Luo et al. applied Random Forest to identify depression risk factors among Chinese college students and achieved 87.5% accuracy with an AUC of 0.927 [3]. Li et al. presented a comprehensive survey of depression prediction using different modalities, including EEG, brain imaging, audiovisual signals, and text [4]. Hasan et al. compared seven machine learning algorithms for predicting depression, anxiety, and stress among Bangladeshi university students [5]. Cao et al. reviewed deep learning-based depression recognition using facial expressions, focusing on spatial and spatial-temporal features [6]. Vu et al. evaluated six machine learning algorithms using NHANES data and identified XGBoost as the best-performing model [7]. Lin and Zhou proposed an ensemble neural network combining fully connected and dropout-based networks for postpartum depression prediction [8]. Abir et al. investigated machine learning and deep learning techniques for EEG-based psychiatric disorder prediction and highlighted the effectiveness of CNN-based approaches [9]. Tahir et al. reviewed machine learning and deep learning techniques for detecting depression through social media data [10]. Mumenin et al. introduced DDNet, a three-stage stacked ensemble model, achieving 98.98% and 99.16% accuracy on two university-student

datasets [11]. Ndikumana et al. applied machine learning models to predict mental health vulnerability among Rwandan youth [12]. Ballı et al. developed machine learning models for identifying suicidal ideation using non-suicidal predictors and external validation [13]. Zhou and Mohd proposed a BERT-BiLSTM-based framework for depression detection from social media text [14]. Zuo and Yang developed an interpretable XGBoost-based model for assessing depression risk among stroke patients [15]. Finally, ShamsEldin et al. evaluated five machine learning algorithms using DASS-42 psychometric data and reported the highest depression prediction accuracy of 99.3% with SVM [16]. Overall, these studies demonstrate the effectiveness of machine learning and deep learning for depression prediction while highlighting the need for improved generalization, explainability, and clinical validation.

Table 1: This table show comparative study for depression prediction.

Ref.	Author(s) and Year	Techniques Used	Performance Parameters / Reported Results	Limitations and Future Scope
[1]	Sharif et al. (2025)	Decision Tree (DT), Support Vector Machine (SVM), LightGBM	DT: 95.66% accuracy; LightGBM: 94.58%; SVM: 91.48%	The study relies on questionnaire-based data, which may introduce self-reporting bias. Future work can use larger and more diverse clinical datasets and validate the models with real-world patient data.
[2]	Zhai et al. (2025)	Machine learning predictive models	AUC: 0.74 for anxiety and 0.77 for depression	The study focuses on college students, which may limit generalization to other age groups. Future research can examine longitudinal data and validate models across different populations.
[3]	Luo et al. (2025)	Random Forest and other ML classifiers	Random Forest: 87.5% accuracy, AUC: 0.927	The study is based on a specific student population. Future work can include participants from different geographical regions and investigate additional psychological and environmental factors.
[4]	Li et al. (2025)	Traditional ML and DL methods; multimodal approaches	Reviews performance using accuracy, F1-score, MAE, RMSE, and other modality-specific metrics.	Limited availability of clinical datasets, data heterogeneity, privacy concerns, and lack of model interpretability. Future work should focus on multimodal fusion, longitudinal studies, explainable AI, and clinical validation.

[5]	Hasan et al. (2025)	KNN, RF, GBM, XGBoost, CatBoost, LR, SVM	For depression, SVM: 56.93% accuracy, 75.60% precision. For anxiety, LR: 69.48% accuracy.	The study is limited to students from two universities and uses self-reported questionnaire data. Future work can include larger, more representative populations and additional clinical or behavioral features.
[6]	Cao et al. (2025)	Deep learning methods for facial-expression recognition, including spatial and spatiotemporal approaches	Reviews recognition performance across facial-expression-based methods.	Facial expressions may not consistently represent an individual's psychological condition. Future work can integrate facial expressions with speech, physiological signals, and contextual information.
[7]	Vu et al. (2025)	LR, RF, Naïve Bayes, SVM, XGBoost, LightGBM; SHAP	XGBoost achieved the best accuracy, sensitivity, specificity, precision, AUC, and F1-score.	The analysis uses a specific NHANES survey period. Future work can use more recent and longitudinal data and validate the model on independent clinical populations.
[8]	Lin and Zhou (2025)	Ensemble neural network combining Fully Connected Neural Network and Dropout Neural Network	FCNN achieved the best accuracy, sensitivity, specificity, precision, AUC, and F1-score.	The study focuses specifically on postpartum depression. Future work can collect more disease datasets and develop an online prediction platform for practical use.

Problem Statement

Depression is a complex psychological condition influenced by multiple interacting factors, including demographic characteristics, socioeconomic circumstances, family history, academic pressure, behavioral patterns, physical health, and psychological well-being. Accurate identification of depression is challenging because individuals may experience different symptoms and may not openly disclose their psychological difficulties. Conventional assessment techniques depend significantly on clinical interviews and self-reported questionnaires, which can introduce subjective variations and delay the identification of individuals requiring psychological support. The reviewed research studies demonstrate that machine learning algorithms can identify depression-related patterns from structured and unstructured datasets.

However, developing a reliable prediction model involves several technical and practical challenges. To develop and evaluate a reliable machine learning-based framework for depression prediction that can effectively identify depression-related patterns from relevant psychological, demographic, behavioral, and clinical attributes while addressing data-quality issues, class imbalance, feature relevance, classifier performance variations, and model generalization. The research focuses on identifying suitable classification algorithms and evaluating their effectiveness through a systematic experimental process. The ultimate objective is to develop a data-driven prediction framework that can assist mental healthcare professionals in preliminary depression risk assessment and support timely clinical intervention.

Conclusion

The reviewed research demonstrates that machine learning and deep learning techniques have significant potential for the early prediction and detection of depression. Traditional classifiers such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Logistic Regression (LR), and K-Nearest Neighbors (KNN), along with advanced approaches such as XGBoost, CatBoost, CNN, LSTM, and ensemble neural networks, have been applied to different types of depression-related data. These studies indicate that prediction performance depends on the selected algorithm, dataset characteristics, feature representation, and preprocessing techniques. The literature review works, also highlights the importance of using diverse data sources, including psychological questionnaires, demographic information, clinical records, EEG signals, facial expressions, speech, and social media text. However, challenges remain regarding dataset diversity, class imbalance, model generalization, privacy, interpretability, and clinical validation. Therefore, future research should focus on developing a reliable, explainable, and generalizable prediction framework used with machine learning techniques with feature selection, optimization supported.

References

[1] M. S. Sharif, A. Zorto, A. T. Kareem, and R. Hafidh, “Effective machine learning based techniques for predicting depression,” University of East London, London, U.K. 2024, pp. 1-6.

[2] Y. Zhai, Y. Zhang, Z. Chu, B. Geng, M. Almaawali, R. Fulmer, Y.-W. D. Lin, Z. Xu, A. D. Daniels, Y. Liu, Q. Chen, and X. Du, “Machine learning predictive models to guide prevention and intervention allocation for anxiety and depressive disorders among college students,” Journal of Counseling & Development, vol. 103, no. 1, pp. 110–125, 2025, doi: 10.1002/jcad.12543.

[3] L. Luo, J. Yuan, C. Wu, Y. Wang, R. Zhu, H. Xu, L. Zhang, and Z. Zhang, “Predictors of depression among Chinese college students: A machine learning approach,” BMC Public Health, vol. 25, Art. no. 470, 2025, doi: 10.1186/s12889-025-21632-8.

[4] Q. Li, X. Liu, X. Hu, M. A. R. Ahad, M. Ren, L. Yao, and Y. Huang, “Machine learning-based prediction of depressive disorders via various data modalities: A survey,” IEEE/CAA Journal of Automatica Sinica, vol. 12, no. 7, pp. 1320–1349, Jul. 2025.

[5] M. E. Hasan et al., “Prevalence, associated factors, and machine learning-based prediction of depression, anxiety, and stress among university students: A cross-sectional study from Bangladesh,”

Journal of Health, Population and Nutrition, vol. 44, Art. no. 361, 2025, doi: 10.1186/s41043-025-01095-8.

[6] X. Cao, L. Zhai, P. Zhai, F. Li, L. He, and T. He, "Deep learning-based depression recognition through facial expression: A systematic review," Neurocomputing, vol. 627, Art. no. 129605, Apr. 2025.

[7] T. Vu, R. Dawadi, M. Yamamoto, J. T. Tay, N. Watanabe, Y. Kuriya, A. Oya, P. N. H. Tran, and M. Araki, "Prediction of depressive disorder using machine learning approaches: Findings from the NHANES," BMC Medical Informatics and Decision Making, vol. 25, Art. no. 83, 2025, doi: 10.1186/s12911-025-02903-1.

[8] Y. Lin and D. Zhou, "A method for predicting postpartum depression via an ensemble neural network model," Frontiers in Public Health, vol. 13, Art. no. 1571522, 2025, doi: 10.3389/fpubh.2025.1571522.

[9] S. I. Abir et al., "Machine learning and deep learning techniques for EEG-based prediction of psychiatric disorders," Journal of Computer Science and Technology Studies, vol. 7, no. 1, pp. 46–63, 2025, doi: 10.32996/jcsts.2025.7.1.4.

[10] W. B. Tahir, S. Khalid, S. Almutairi, M. Abohashrh, S. A. Memon, and J. Khan, "Depression detection in social media: A comprehensive review of machine learning and deep learning techniques," IEEE Access, vol. 13, pp. 12789–12816, 2025, doi: 10.1109/ACCESS.2025.3530862.

[11] N. Mumenin, M. A. Yousuf, M. O. Alassafi, M. M. Monowar, and M. A. Hamid, "DDNet: A robust, and reliable hybrid machine learning model for effective detection of depression among university students," IEEE Access, vol. 13, 2025, doi: 10.1109/ACCESS.2025.3552041.

[12] F. Ndikumana et al., "Machine learning-based predictive modelling of mental health in Rwandan youth," Scientific Reports, vol. 15, Art. no. 16032, 2025, doi: 10.1038/s41598-025-00519-z.

[13] Deepak Kumar Rathore, Dr. Praveen Kumar Mannepalli, "Recent Trends in Machine Learning for Health Care Sector", International Journal of Innovative Research in Technology and Management, Vol-5, Issue-2, 2021.

[14] S. Zhou and M. 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, 2025, doi: 10.1109/ACCESS.2025.3559170.

[15] W. Zuo and X. Yang, "Network-based predictive models for artificial intelligence: An interpretable application of machine learning techniques in the assessment of depression in stroke patients," BMC Geriatrics, vol. 25, Art. no. 193, 2025, doi: 10.1186/s12877-025-05837-5.

[16] T. ShamsEldin, S. Gaber, S. Ansari, R. Elgohary, M. A. Shawky, M. Elbahnasawy, and M. Abdrabou, "Artificial intelligence for predicting depression anxiety and stress using psychometric data," Scientific Reports, vol. 15, Art. no. 37282, 2025, doi: 10.1038/s41598-025-21301-1.

[17] Deepak Kumar Rathore, Dr. Praveen Kumar Mannepalli, "A Review of Machine Learning Techniques and Applications for Health Care ", International Conference on Advances in Technology, Management & Education, 2021, IEEE proceeding, 978-1-7281-8586-6/21.

[18] M. Ballı, A. E. Dogan, S. H. Senol, and H. Yapici Eser, "Machine learning based identification of suicidal ideation using non-suicidal predictors in a university mental health clinic," Scientific Reports, vol. 15, Art. no. 13843, 2025, doi: 10.1038/s41598-025-97387-4.