International Journal of Engineering Trends and Applications (IJETA) – Volume 13 Issue 3, May-Jun 2026

RESEARCH ARTICLE OPEN ACCESS

AI & MENTAL HEALTH: Smart Systems for Early Detection of Mental Disorders

Sarves Kumar Meena*, Surendar Rathor**, Ayush Ray**, Priyanka Choudhary**, Adarsh Kumar Singh**

*Assistant Professor, Department of CSE, Global Institute of Technology, Jaipur, Rajasthan, India
**B.Tech Student, Department of Information Technology, Global Institute of Technology, Jaipur, Rajasthan, India

ABSTRACT

Mental health disorders represent one of the most pressing yet underdiagnosed public health challenges of the twenty-first century. Traditional approaches to diagnosis are often constrained by limited clinical resources, social stigma, delayed help-seeking, and the subjective nature of psychiatric assessment. This paper examines how Artificial Intelligence (AI) encompassing machine learning, natural language processing, and deep learning can function as an intelligent early-warning system for a range of mental health conditions including depression, anxiety, bipolar disorder, and schizophrenia. Drawing on recent published Research, the study surveys data modalities used for AI-based detection (textual, physiological, behavioural, and neuroimaging data), evaluates commonly employed algorithms, highlights the most effective real-world applications, and critically discusses the ethical, privacy, and bias-related challenges that arise when deploying AI in mental healthcare. The paper concludes that AI, when designed responsibly and applied as a complement rather than a replacement to clinical judgement, holds transformative potential for expanding early mental health support globally..

***Keywords* —** Artificial Intelligence, Mental Health, Early Detection, Machine Learning, Natural.

1. Introduction

According to the World Health Organization (WHO), more than 970 million people worldwide live with a mental or substance use disorder, making it the leading cause of disability globally. Despite this staggering burden, over 75% of individuals in low- and middle income countries receive no treatment whatsoever a gap driven by an acute shortage of trained mental health professionals, deeply ingrained cultural stigma, geographical barriers, and the high cost of care. In India alone, the National Mental Health Survey estimates a treatment gap of nearly 83%, meaning the majority of people who need help never access it.

Artificial Intelligence particularly its subfields of machine learning (ML), deep learning (DL), and natural language processing (NLP) has emerged as a potentially transformative force in addressing this gap. AI systems can analyse patterns in language, behaviour, physiological signals, and brain activity at a speed and scale no human clinician can match. More importantly, they can do so passively, unobtrusively, and continuously through devices people already carry: smartphones, wearables, and computers.

This paper investigates how AI can serve as an early detection and monitoring tool for mental health disorders. The objectives of this study are: (1) to review the landscape

of AI techniques used in mental health screening; (2) to examine the types of data these systems use; (3) to survey real-world applications already deployed; and (4) to critically evaluate the ethical considerations that must guide responsible development in this sensitive domain.

2. Background and Literature Review

2.1 The Mental Health Crisis and Technology's Role

The intersection of technology and mental health is not new. Digital mental health interventions ranging from telemedicine to smartphone apps have been studied for over two decades. However, the majority of earlier tools were rule-based, offering scripted Cognitive Behavioural Therapy (CBT) modules or self-assessment questionnaires. What distinguishes the modern AI-driven approach is the capacity for personalisation, continuous monitoring, and predictive analytics.

Torous et al. (2017) were among the first to systematically map the use of smartphones for mental health monitoring, noting that passive sensing data from devices can reveal behavioural signatures of mood disorders. Reece and Danforth (2017) demonstrated that Instagram photographs shared by users later diagnosed with depression differed measurably in colour, brightness, and social engagement from those of healthy controls and that a machine learning model trained on these features outperformed general practitioners in retrospective depression screening.

2.2 Key AI Techniques in Mental Health Research

The following AI methodologies dominate current mental health research:

Support Vector Machines (SVM): Widely used in classification tasks, SVMs have been applied to distinguish between depressed and non-depressed individuals based on speech features and social media text.

Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM): Sequential data such as daily mood logs, Twitter post histories, or heart rate time series is well-suited to RNN/LSTM architectures that can model temporal patterns of deteriorating mental state.

Transformer Models (BERT, RoBERTa): Pre-trained language models have achieved state-of-the-art performance on depression and suicidality detection tasks by capturing nuanced linguistic features such as hopelessness, social withdrawal, and cognitive distortions.

Convolutional Neural Networks (CNN): Applied to neuroimaging (fMRI and EEG) data, CNNs can identify structural and functional brain abnormalities associated with schizophrenia, PTSD, and major depressive disorder.

3. Data Modalities for Mental Health Detection

3.1 Textual and Linguistic Data

Language is perhaps the richest noninvasive indicator of psychological state. AI systems can analyse written text from social media posts, clinical interview transcriptions, diary entries, and chatbot conversations. Studies using Reddit corpora (particularly communities such as r/depression and r/SuicideWatch) have trained NLP classifiers to detect depressive language patterns, identifying features such as first person pronoun use, negative affect vocabulary, absolutist thinking, and diminished future orientation. Coppersmith et al. (2014) demonstrated that linguistic patterns on Twitter could distinguish users with depression, PTSD, and bipolar disorder with statistically significant accuracy.

3.2 Physiological and Wearable Sensor Data

Wearable devices smartwatches, fitness bands, and biosensor patches continuously collect heart rate variability (HRV), electrodermal activity (EDA), skin temperature, sleep patterns, and physical activity levels. These physiological signals are closely linked to the autonomic nervous system, which is dysregulated in most mood and anxiety disorders. Research by Sarker et al. (2021) showed that passively collected actigraphy (movement) data from wrist-worn sensors could predict depressive episode onset up to one week in advance in patients with bipolar disorder.

3.3 Behavioural and Smartphone Usage Data

Modern smartphones generate a continuous stream of behavioural data: call frequency and duration, app usage patterns, screen time, typing speed and error rates, GPS mobility, and even microphone-based acoustic features. Studies from the StudentLife project at Dartmouth College demonstrated that smartphone-derived behavioural data could predict exam stress, depression, and loneliness in college students with moderate-to-high accuracy, providing a proof-of concept for ambient mental health monitoring in real-world environments.

3.4 Neuroimaging and EEG Data

Functional Magnetic Resonance Imaging (fMRI) and Electroencephalography (EEG) provide direct windows into brain activity. Deep learning models applied to fMRI connectivity matrices have achieved classification accuracies exceeding 90% for distinguishing schizophrenia patients from healthy controls. EEG-based approaches are particularly promising for real-time, low-cost deployment given their portability relative to fMRI infrastructure.

However, the clinical translation of these methods remains limited by dataset size, hardware requirements, and the expertise needed for signal acquisition.

4. Real-World Applications

4.1 AI-Powered Chatbots and Virtual Therapists

Conversational AI agents such as Woebot, Wysa, and Youper have been deployed at scale to deliver CBT-based mental health support. These systems use NLP to detect emotional distress in real time, triage severity, and escalate to human clinicians when necessary. A randomised controlled trial published in JMIR Mental Health (Fitzpatrick et al., 2017) found that Woebot significantly reduced depression and anxiety symptoms in college students over a two-week period at a fraction of the cost of in-person therapy.

4.2 Clinical Decision Support Systems

Electronic Health Record (EHR) systems integrated with AI can flag patients at elevated risk of psychiatric deterioration or suicide based on patterns in clinical notes, prescription histories, and follow-up gaps. IBM Watson Health and Google Deep Mind have explored clinical NLP pipelines that scan free-text medical notes for risk indicators. Such systems, if deployed responsibly, could prompt clinicians to schedule timely follow-ups and prevent crisis events.

4.3 Social Media Monitoring

Organisations such as Crisis Text Line and the non-profit Bark use AI to monitor patterns of distress in text messages and social media posts respectively. Bark's AI system reviews communications for signs of cyberbullying, depression, and suicidal ideation, alerting parents and school counsellors when risk indicators are detected. Facebook (now Meta) has also deployed AI based suicide prevention tools that identify at-risk posts and connect users with crisis resources.

5. Ethical Challenges and Limitations

5.1 Privacy and Data Security

Mental health data is among the most sensitive categories of personal information. The collection of social media posts, voice recordings, and physiological signals for AI training raises profound questions about informed consent, data ownership, and the risk of re-identification. In countries governed by frameworks such as India's Digital Personal Data Protection Act (DPDPA) 2023 or the EU's General Data Protection Regulation (GDPR), developers must ensure that data is collected with explicit consent, stored securely, and not repurposed without authorisation.

5.2 Algorithmic Bias and Fairness

Most mental health AI datasets are drawn from English-speaking, Western, predominantly white populations making models poorly generalizable to linguistically and culturally diverse communities. A model trained on English Twitter may fail entirely when applied to Hindi, Urdu, or Tamil text. Moreover, AI systems trained on historical clinical data may replicate and amplify existing healthcare inequities, systematically under detecting disorders in marginalised groups who were already underdiagnosed in the training data.

5.3 Clinical Validity and the Risk of False Positives

No AI system is 100% accurate. In a clinical context, a false positive telling a healthy person they may have a mental disorder can cause significant distress, unnecessary medical intervention, and stigma. A false negative, on the other hand, may provide dangerous false reassurance. It is therefore essential that AI tools are validated in prospective clinical trials, positioned as screening aids rather than diagnostic tools, and paired with trained human oversight.

5.4 Explain ability and Transparency

Deep learning models are frequently described as "black boxes" their internal decision logic is opaque even to their creators. When an AI flags a person as being at risk of suicide, clinicians and patients have a right to understand why. The emerging field of Explainable AI (XAI), encompassing techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), aims to make AI predictions more interpretable. Regulatory frameworks in several jurisdictions now mandate explain ability for high-stakes AI decisions, including those in healthcare.

6. Future Directions

Several directions hold particular promise for the next decade of AI driven mental health research and practice:

Federated Learning: This decentralised approach trains models across multiple institutions without transferring raw patient data, enabling large-scale collaboration while preserving privacy.

Multilingual and Cross-Cultural Models: Investment in diverse, multilingual datasets will be critical to building tools that work equitably across different populations, including India's vast linguistic diversity.

Multimodal Fusion: Combining text, voice, physiological, and visual data in a single unified model is likely to significantly improve detection accuracy and robustness over single-modality approaches.

Human-AI Collaborative Care: The most effective model is not AI replacing clinicians, but AI augmenting them — handling screening and monitoring at scale so that human professionals can focus on nuanced therapeutic work.

7. Conclusion

Artificial Intelligence is not a panacea for the global mental health crisis, but it is an extraordinarily powerful set of tools that, if developed and deployed with care, can substantially narrow the treatment gap. From NLP-powered chatbots providing accessible CBT to passively-sensing wearables that predict mood episodes before they occur, the breadth of AI's application in mental healthcare is remarkable and still expanding.

At the same time, the research community and technology industry carry a significant responsibility. These systems handle some of the most intimate dimensions of human experience. Failures of privacy, bias, or transparency are not merely technical errors they are ethical violations that could cause genuine harm to vulnerable individuals. The path forward requires interdisciplinary collaboration between computer scientists, psychiatrists, ethicists, patients, and policymakers to build systems that are not only accurate, but also fair, transparent, and humane. For undergraduate students in computer science, this domain represents one of the most socially meaningful frontiers available. Building mental health technology is an opportunity to combine technical skill with genuine human impact and to contribute, through engineering and research, to a healthier and more equitable world.

References

[1] G. Coppersmith, M. Dredze, and C. Harman, “Quantifying mental health signals in Twitter,” in Proc. Workshop Computational Linguistics and Clinical Psychology, pp. 51–60, 2014.

[2] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, “Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial,” JMIR Mental Health, vol. 4, no. 2, pp. 1–12, 2017.

[3] A. G. Reece and C. M. Danforth, “Instagram photos reveal predictive markers of depression,” EPJ Data Science, vol. 6, no. 15, pp. 1–12, 2017.

[4] H. Sarker, M. Masud, R. Bhatt, et al., “Passive monitoring of actigraphy data predicts onset of bipolar depressive episodes,” Journal of Affective Disorders, vol. 295, pp. 1318–1327, 2021.

[5] J. Torous, M. V. Kiang, J. Lorme, and J. P. Onnela, “New tools for new research in psychiatry: A scalable and customizable platform to empower data-driven smartphone research,” JMIR Mental Health, vol. 3, no. 2, pp. 1–10, 2017.

[6] World Mental Health Report: Transforming Mental Health for All. Geneva, Switzerland: World Health Organization (WHO) Press, 2022.

[7] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, “StudentLife: Assessing mental health, academic performance, and behavioral trends of college students using smartphones,” in Proc. ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp), pp. 3–14, 2014.

[8] A. B. R. Shatte, D. M. Hutchinson, and S. J. Teague, “Machine learning in mental health: A scoping review of methods and applications,” Psychological Medicine, vol. 49, no. 9, pp. 1426–1448, 2019.

[9] S. Graham, C. Depp, E. E. Lee, C. Nebeker, X. Tu, H. C. Kim, and D. V. Jeste, “Artificial intelligence for mental health and mental illnesses: An overview,” Current Psychiatry Reports, vol. 21, no. 11, pp. 1–18, 2019.

[10] D. D. Luxton, “Artificial intelligence in psychological practice: Current and future applications and implications,” Professional Psychology: Research and Practice, vol. 45, no. 5, pp. 332–339, 2014.

[11] A. Agarwal, R. Joshi, H. Arora, and R. Kaushik, “Privacy and security of healthcare data in cloud based on the blockchain technology,” in Proc. 7th International Conference on Computing Methodologies and Communication (ICCMC), pp. 87–92, 2023.

[12] M. Dahiya, N. Hemrajani, A. Kumar, S. Rani, and S. Rathee, Artificial Intelligence in Medicine and Healthcare. Boca Raton, FL, USA: Taylor & Francis, 2025.

[13] R. Misra, N. Sharma, G. K. Soni, H. Arora, S. Chauhan, and A. Biswas, “A hybrid and ensemble machine learning framework for enhanced brain tumor detection,” in Proc. International Conference on Electronics and Renewable Systems (ICEARS), pp. 1583–1588, 2026.

[14] G. K. Soni, D. Yadav, A. Kumar, P. Jain, and A. Rathi, “Design and SAR analysis of DGS based deformed microstrip antenna for

ON/OFF body smart wearable IoT applications," Physica Scripta, vol. 100, no. 1, pp. 1–28, 2025.

[15] G. K. Soni, D. Yadav, and A. Kumar, "Enhancing healthcare: Flexible and wearable antenna design for tumor detection," in Proc. IEEE International Conference on Distributed Computing and Optimization Techniques (ICDCOT), pp. 1–5, 2024.

[16] G. K. Soni, D. Yadav, A. Kumar, and L. Sharma, "Flexible antenna design for wearable IoT devices," in Proc. IEEE 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 863–867, 2023.

[17] R. Ajmera et al., "Prediction analysis for diabetic patients using clustered based classification," Journal of Emerging and Innovative Research, vol. 5, no. 7, pp. 770–775, 2018.

[18] A. Gautam, R. Ajmera, D. K. Dharamdasani, S. Srivastava, and A. Johari, "Improving climate change predictions using time series analysis and deep learning," Global and Stochastic Analysis, vol. 12, no. 4, 2025.

[19] K. Ahuja, H. Sekhawat, S. Mishra, and P. Jha, "Machine learning in artificial intelligence: Towards a common understanding," Turkish Online Journal of Qualitative Inquiry, vol. 12, no. 8, pp. 1–10, 2021.

[20] M. Mehra, P. Jha, H. Arora, K. Verma, and H. Singh, "Salesforce vaccine for real-time service in cloud," in Sentimental Analysis and Deep Learning: Proc. International Conference on Sentimental Analysis and Deep Learning (ICSADL), Advances in Intelligent Systems and Computing, vol. 1408, pp. 1003–1008, 2022.

[21] P. Jha, G. K. Soni, H. Dogra, D. Goswami, K. Choudhary, and H. Vaishnav, "Plant disease detection and classification using convolutional neural network," in Proc. 4th International Conference on Automation, Computing and Renewable Systems (ICACRS), pp. 1442–1446, 2025.

[22] K. Paliwal, P. Jha, S. Kumari, V. Vaish, N. Vishwakarma, and A. Bansal, "Machine learning in electric vehicle consumption modelling," in Proc. 9th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 727–730, 2026.

[23] P. Jha, P. Jain, A. Kumar, S. Soni, Y. Sharma, and P. Agarwal, "The application of Markov chains to linguistic predictions by utilising its inherent information entropy," in Proc. 9th International Conference on Inventive Systems and Control (ICISC), pp. 1110–1114, 2025.

[24] P. Upadhyay, K. K. Sharma, R. Dwivedi, and P. Jha, "A statistical machine learning approach to optimize workload in cloud data centre," in Proc. 7th International Conference on Computing Methodologies and Communication (ICCMC), pp. 276–280, 2023.

[25] P. Jha, T. Biswas, U. Sagar, and K. Ahuja, "Prediction with ML paradigm in healthcare system," in Proc. Second International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 1334–1342, 2021.