Grenze International Journal of Engineering and Technology, June Issue

An Ensemble Learning Model for Accurate Mental Stress State Prediction

Prashant G. Ghulaxe[1], Jatin Umesh Mahajan[2], Ayush Rajendra Titirmare[3], Piyush Gathode[4] and Mandar Milind Lakhe[5]
[1-5]Department of Computer Science & Engineering, Yashwantrao Chavhan College of Engineering, Nagpur, Maharashtra, India
Email: pghulaxe@gmail.com, jatinmahajan9890@gmail.com, titirmareayush21@gmail.com, ghatodepiyush7@gmail.com, mandarlakhe150@gmail.com

***Abstract*— The tremendous use of technology and social media and also tremendous work from working place so that people suffer from several mental disorders like anxiety, and stress become critical global concerns, affecting individuals' quality of life and overall productivity. This study proposes an intelligent system that integrates NLP and ML to predict the mental state issues by analyzing linguistic and behavioral patterns from textual data on social media. This paper proposed the ML based ensemble model that involves selecting the feature such as emotional tone, sentiment polarity, and word frequency, followed by text preprocessing techniques like tokenization, stop-word removal, and lemmatization. Various baseline ML models were proposed and trained and evaluated to identify individuals at potential risk of low, medium and high. XGBoost obtained accuracy of 0.9636, F1-score of 0.9594, and MCC of 0.9409. The Light GBM obtained an accuracy of 0.9573 and F1-score of 0.9526. RF obtained accuracy of 0.9281 and an MCC of 0.8820. The ensemble model shows the best overall results to obtained accuracy of 0.9842, MCC of 0.9744 and Cohen's kappa of 0.9743. The findings show the strong capability of NLP-based ML approaches in detection of mental health conditions.**

***Index Terms*— Mental State Prediction, Social Media, Behavioral Patterns, Ensemble Model, Sentiment Polarity.**

I. Introduction

Mental health has become one of the most critical aspects of human well-being in the 21st century. With the growing pace of modern life, increased social media engagement, and workplace pressures so that several mental disorders are becoming alarmingly common across all age groups [1]-[2]. According to the WHO, approximately one in four people worldwide were suffered from neurological disorder at some point in their lives [3]. Despite this, a significant proportion of individuals suffering from mental health issues remain undiagnosis, lack of awareness, and limited healthcare facilities [4].

Early identification and intervention help to prevent the progression of mental disorders and improving patient outcomes. However, traditional approaches to diagnosis of mental disorders using clinical interviews, psychological assessments, and questionnaires were time-consuming, subjective, and sometimes inaccessible [5]. With advancements AI and NLP, there is growing potential to used data-driven insights for early prediction and monitor the mental health through textual or speech-based communication data [6].

Grenze ID: 01.GIJET.12.2.186_22

This paper focuses on applying computational techniques to identify linguistic cues, emotional patterns, and sentiment markers that may indicate mental health concerns. By analyzing textual data from online platforms, chat transcripts, or surveys, ML models can predict the mental illnesses with high accuracy that enables early intervention and timely support.

Contribution of the paper is

- To analyze the key linguistic and behavioral factors influencing mental health based on textual expressions
- To design the ensemble models based on the several ML models for predicting mental health state.
- To measure the effectiveness of proposed models based on several evaluation parameters.

The structure of the paper is as follows

The section 2 reviews the previous work in mental state prediction. The section 3 presents the complete methodology of mental state prediction consists of step-by-step process discussed. The section 4 discussed the experimental findings of the proposed models. Finally conclude the study and discussed the future direction.

II. Related Work

Current advancement in ML and NLP significantly enhanced the automatic prediction of mental states, particularly through the analysis of behavioral, psychological, and social media data. Mandal K. P. (2025) investigated mental health treatment-seeking behavior using multiple ML models and reported that a stacking-based ensemble achieved the highest accuracy of 81.7% shows the importance of workplace interference and perceived anonymity as influential factors [7]. Narang et al. (2025) suggested NLP-driven SVC model on large-scale social media over 27,000 posts and obtained 92.00% accuracy to identify the mental health [8].

Another study focused on mental-state classification over the Twitter data. Nabil et al. (2023) also suggested the traditional and ML models such as DT, SVM, and RNN obtained the best performance with 76.00% accuracy in binary mental-state classification [9]. Kabir et al. (2023) proposed and LR model to predict the stress from textual data and obtained the accuracy score of 83.00% precision, 98.00% recall, and an F1-score of 91.00% [10]. Rawat et al. (2024) applied DL models for depressive state detection, where an LSTM model obtained 99.27% accuracy on the full dataset, while a BERT–MLP model showed improved robustness on balanced data with 98.49% accuracy [11].

More recent work has integrated ensemble learning and transformer-based language models for improved generalization. Abdullah et al. (2024) evaluated traditional ML gradient boosting, and LLM such as RoBERTa to predict future mental disorders from social media content; however, performance on non-clinical datasets remained limited, with an F1-score of 0.43, indicating challenges in real-world deployment [12]. Gaonkar et al. (2025) shows that transformer-based model consistently outperform conventional methods, obtained the 98.00% accuracy in depression detection tasks [13]. Earlier reviews, such as that by Mohammadi et al. (2020), highlighted wide performance variability (42%–87% accuracy) across studies due to dataset heterogeneity and reliance on publicly available social media posts [14]. Vijayalakshmi et al. (2024) further confirmed the effectiveness of tweet-based mental-state prediction, reporting 98.25% accuracy using NLP pipelines and 96.71% accuracy with ML models such as RF and XGBoost [15]

TABLE I. SUMMARY OF PREVIOUS RELATED WORK IN MENTAL STATE PREDICTION

Author	Datasets	Methods	Findings
Mandal K. P. (2025)	Survey-based mental health data	ML + Stacking Ensemble	Accuracy: 81.7%
Narang et al. (2025)	27,000+ social media posts	NLP + SVC	Accuracy: 92%
Nabil et al. (2023)	Twitter data	DT, SVM, RNN	Accuracy: 76% (RNN)
Kabir et al. (2023)	Textual stress data	Logistic Regression	Precision: 83%, Recall: 100%, F1: 91%
Rawat et al. (2024)	Twitter data	LSTM, BERT-MLP	Accuracy: 98.27% (LSTM)
Abdullah et al. (2024)	Social media posts	LR, XGBoost, RoBERTa	F1-score: 0.43
Gaonkar et al. (2025)	Social media datasets	NLP + Transformers	Accuracy: >98%
Mohammadi et al. (2020)	Public social media posts	Review of ML methods	Accuracy: 42%–87%
Vijayalakshmi et al. (2024)	Twitter data	NLP, RF, XGBoost	Accuracy: 98.25%

Table 1 shows the existing studies confirm the effectiveness of ML and DL models for mental-state prediction, particularly when combined with large-scale textual data and advanced ensemble or transformer-based models. However, challenges related to generalization, class imbalance, and interpretability remain open research directions.

II. METHODOLOGY

This section presents the complete workflow of proposed model of the mental health prediction. The model consists of several steps data collection. After collection, the text preprocessing includes cleaning, removing noise, and transform into standardize format. The ML based ensemble model trained on preprocess training data and predict the mental health into low, medium, and high risk.

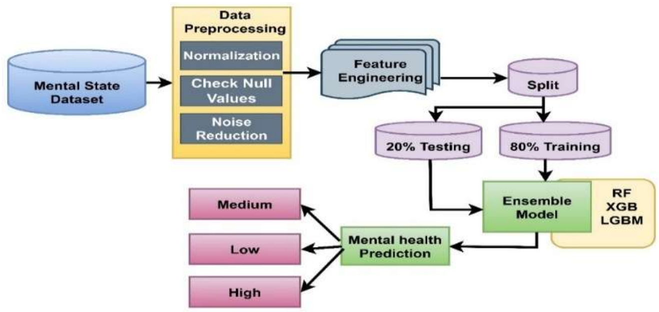

Figure 1. Complete architecture of mental state prediction

Figure 1 shows the complete architectural workflow of proposed model. The proposed model used data-driven and iterative techniques and, allows to refine and optimization at each step based on intermediate results.

B. Dataset Description

In this study the dataset is obtained from publicly available on Reddit posts, Twitter datasets, or online forums dedicated to mental health discussions. The datasets contain 6850 samples that express various emotional and psychological conditions rang from normal mental states that indicate the distress, anxiety, and depression. Each psychological risk categories into medium, low and high with 3000, 2850 and 1000 text entry is associated with a label.

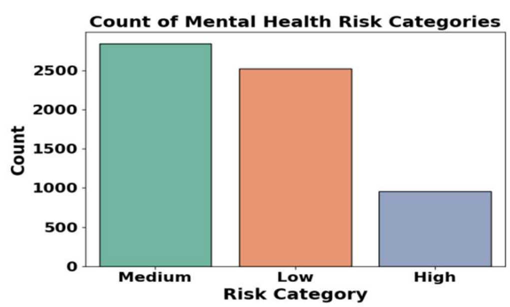

Figure 2. Distribution of the Dataset

Figure 2 shows the class wise distribution of the dataset that shows the risk factor of medium, low and high.

C. Data Preprocessing

The raw data contain noise, redundant, and irrelevant data that can negatively affect model performance. In this step prepares the data for analysis and feature extraction. This process involves a sequence of cleaning and transformation operations to improve dataset quality and uniformity. The preprocessing steps include:
Lowercasing: Tranformed all the words to lowercase to avoid case sensitivity issues.
Stop-word Removal: Eliminate the common but uninformative words such as “the”, “is”, “and” that do not contribute meaningful to sentiment analysis.
Lemmatization/Stemming: Reducing words to their root form such as running → run to minimize redundancy to capture linguistic meaning.
Noise Removal: Deleting special characters, URLs, numbers, and punctuation that may distort analysis.

D. Feature Engineering

In this step meaningful representations of the text are created in a numerical form for baseline ML models. In NLP-based prediction, features capture semantic, syntactic, and sentiment-related aspects of language.

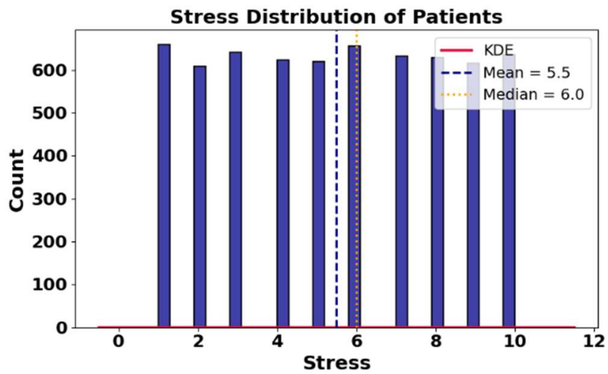

Figure 3. Stress Distribution of the Patient in dataset

Figure 3 shows the stress distribution of the Patient in dataset. This indicates the directly influence the model to detect subtle linguistic cues associated with mental state. For instance, the presence of negative emotional words (sad, hopeless, anxious) or cognitive distortion patterns (always, never, worthless) can be strong indicators of psychological distress.

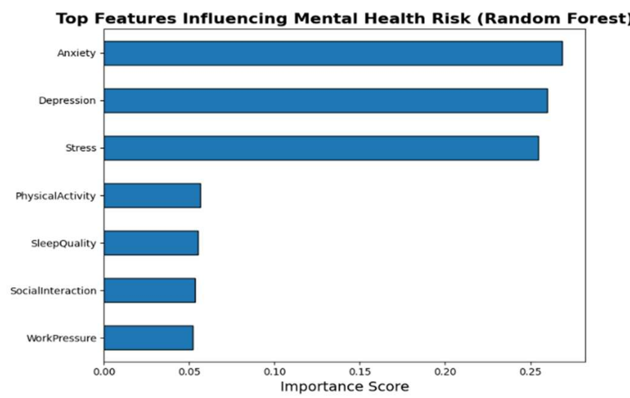

Figure 4. Top features for Mental health risk

Figure 4 shows the feature importance analysis that indicate the most crucial features of mental health risk and contribute the substantially higher importance scores than other variables. Behavioral and lifestyle factors such as physical activity, sleep quality, social interaction, and work pressure show lower influence that shows the secondary but supportive role in risk analysis.

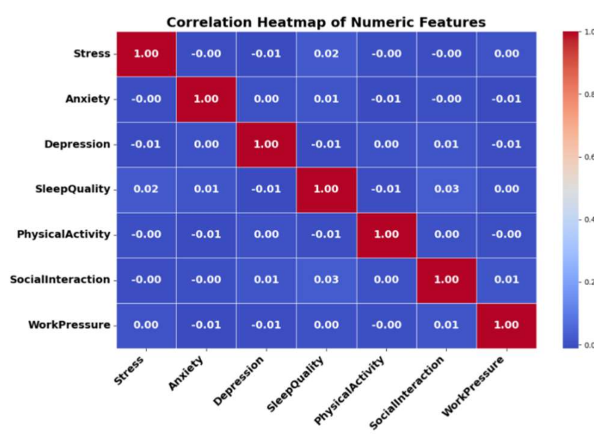

Figure 5. Heatmap of each Feature Correlation

Figure 5 shows the correlation heatmap of the mental-health features like Stress, Anxiety, Depression, Sleep Quality, Physical Activity, Social Interaction, and Work Pressure show that the dataset exhibits extremely weak linear relationships among all variables, with correlation coefficients range to -0.01 to 0.03. As expected, each feature shows a perfect self-correlation of 1.00 along the diagonal. Minor positive associations are observed, such as SleepQuality–SocialInteraction (r = 0.03) and Stress–SleepQuality (r = 0.02), while negligible negative correlations appear between Stress–Depression (r = –0.01), Anxiety–PhysicalActivity (r = –0.01), and WorkPressure–Anxiety (r = –0.01). The uniformly low values indicate minimal multicollinearity, suggesting that each feature contributes independent information about mental-state variation.

E. Building Ensemble Model

The proposed ensemble model uses a stacking model integrates multiple heterogeneous base classifiers to enhance predictive performance and model generalization. In this proposed ensemble model used three ML classifiers such as RF, XGBoost, and LightGBM are first trained independently on the same training dataset. Each model captures different structural patterns: Random Forest used bootstrap aggregation and feature randomness, XGBoost models complex non-linear interactions through gradient boosting, and LightGBM introduces leaf-wise tree growth for efficient high-dimensional learning. Their predictions are then passed, along with the original feature set (enabled by passthrough=True), to a meta-learner, implemented as Logistic Regression. A 5-fold cross-validation (cv=5) strategy that the stacked model is trained on out-of-fold predictions, preventing overfitting issue.

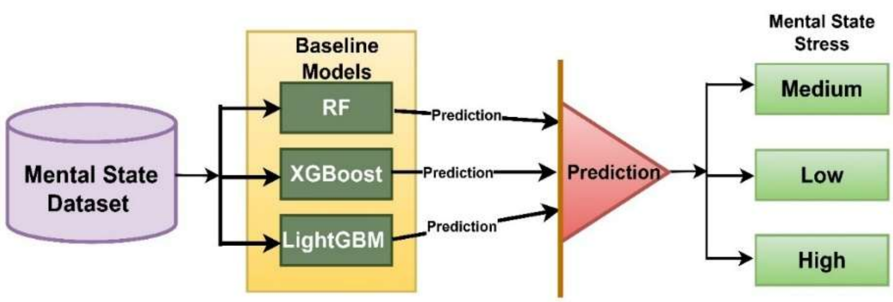

Figure 6. Stacking Ensemble Model

Figure 6 shows the ensemble model for mental-state prediction, where the input mental-state dataset is simultaneously processed by three baseline models such as RF, XGBoost, and Light GBM. Each model independently learns feature representations and generates its own prediction output.

IV. Result and Discussion

This section presents the performance of the baseline ML models and trained using the training dataset for fair comparison.

The simulation of model was carried out on Google Colab Pro using python programming. The dataset is divided into 80% of training and 20% of testing. The Cross-validation $(k = 5)$ was used to address the overfitting and improve the proposed model performance.

TABLE II. Performance Analysis of proposed Models

Model	Accuracy	Precision	Recall	F1 Score	MCC	Kappa
RF	0.9281	0.9421	0.9120	0.9253	0.8820	0.8828
XGBoost	0.9636	0.9603	0.9586	0.9594	0.9409	0.9410
LightGBM	0.9573	0.9565	0.9489	0.9526	0.9305	0.9306
Ensemble	0.9842	0.9842	0.9842	0.9842	0.9744	0.9743

Table 5.1 shows the performance of the proposed baseline and ensemble models that indicate the ensemble model outperforms as compared to baseline models across all evaluation metrics. Among the single models, XGBoost obtained the highest accuracy of 0.9636, F1-score of 0.9594, and MCC of 0.9409. The Light GBM model obtained the accuracy of 0.9573 and F1-score of 0.9526. The RF shows comparatively lower performance that obtained accuracy of 0.9281 and an MCC of 0.8820. The ensemble model shows the best overall results to obtained an accuracy of 0.9842, with the highest MCC of 0.9744 and Cohen's kappa of 0.9743.

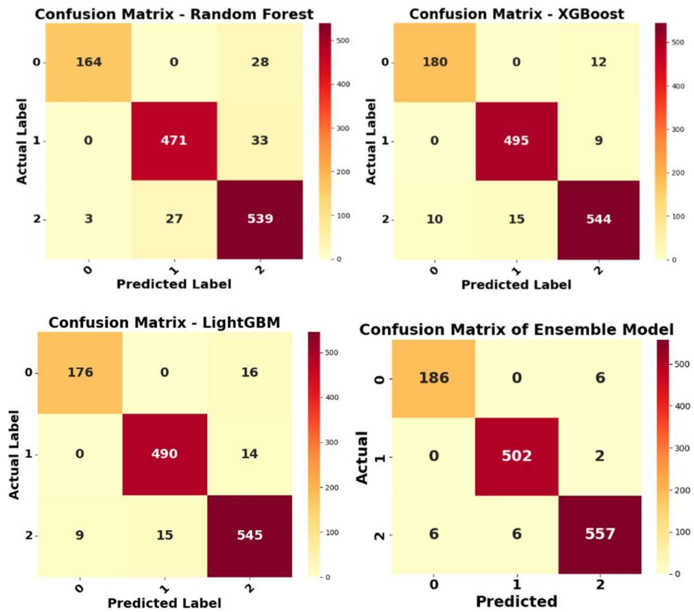

Figure 7. Confusion Matrix of Baseline and Ensemble Models

Figure 7 shows the confusion matrix of the baseline and ensemble models shows highly accurate classification of mental stress states across the three categories. For Class 0, the ensemble model correctly predicted 186 out of 192 instances, misclassifying only 6 samples as Class 2. Class 1 shows near-perfect performance, with 502 correct predictions and only 2 instances misclassified as Class 2. Class 2 also exhibits strong accuracy, with 557 out of 569 samples correctly identified, while 6 samples were misclassified as Class 0 and another 6 as Class 1.

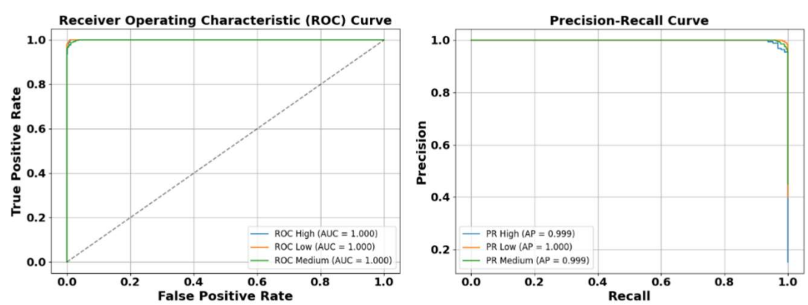

Figure 7. ROC-AUC and PR Curve Ensemble Models

Figure 8 shows the ROC-AUC and PR curve of the ensemble model shows the high discriminative performance for mental stress state prediction. All three classes such as High, Low, and medium stress that exhibit an AUC value of 1.000 that indicate perfect separation between positive and negative instances across all thresholds. The curves rapidly rise to a TPR of 1.0 with near-zero FPR, and precision and recall in PR shows that the model consistently identifies each stress class with no misclassification trade-offs.

TABLE III. COMPARATIVE ANALYSIS OF PROPOSED MODELS WITH EXISTING MODELS

Authors	Models	Accuracy
Mandal K. P. (2025)	ML + Stacking Ensemble	81.7%
Narang et al. (2025)	NLP + SVC	92.00%
Nabil et al. (2023)	DT, SVM, RNN	76.00%
Rawat et al. (2024)	LSTM, BERT-MLP	98.27%
Gaonkar et al. (2025)	NLP + Transformers	98.00%
Mohammadi et al. (2020)	RF	87.00%
Vijayalakshmi et al. (2024)	NLP, RF, XGBoost	98.25%
Proposed	Ensemble	98.42%

Table 3 shows the comparison of ensemble model that obtained the superior performance against existing models in mental-state prediction. Earlier ML–based models such as the stacking ensemble by Mandal K. P. (2025) and traditional classifiers reported by Nabil et al. (2023) obtained lower perfromance of 81.7% and 76.0%, respectively. Some Models used NLP and SVC-based approach by Narang et al. (2025) and the RF model suggested by Mohammadi et al. (2020) shows moderate improvements with accuracy of 92.0% and 87.0%. More recent DL-based and transformer-based models such as LSTM and BERT-MLP proposed by Rawat et al. (2024) and NLP-transformer model by Gaonkar et al. (2025) obtained accuracy score of 98.0% to 98.27%. The proposed ensemble model obtained accuracy of 98.42% outperform with existing model that indicate the effectiveness of combine the baseline models for robust and accurate mental stress state prediction.

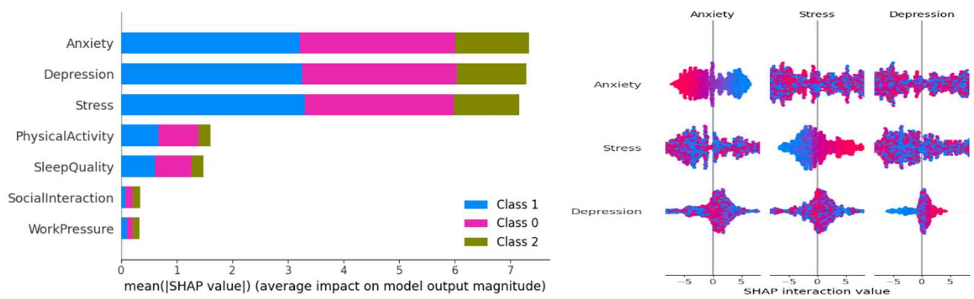

Figure 9. SHAP interaction analysis showing pairwise feature interactions among key psychological factors

Figure 9 shows the SHAP global importance and interaction plot that indicate the Anxiety, Depression, and Stress are the dominant features across all risk factor that contribute the highest average impact in mental health risk stratification.

V. Conclusion and Future Scope

Mental stress is a critical global issue that impact on emotional well-being, cognitive performance, and physical health. Early prediction of stress patterns helps timely medication and prevent from severe psychological disorders. In this paper, three baseline ML models such as RF, XGBoost, and LightGBM were trained and evaluated the mental state. XGBoost shows the strongest individual performer accuracy = 96.36%, followed by LightGBM accuracy = 95.73% and RF accuracy = 92.81%. However, to capture complementary strengths of all models and reduce individual biases, a stacking ensemble model was designed using LR as the meta-learner. This ensemble model improved performance that obtained the 98.42% accuracy, 0.984 macro precision, and the highest MCC of 0.9744 and kappa of 0.9743 scores among all evaluated models.

Future work can expand on the use of modern neural models such as LSTM for understanding emotional sequences, BERT for capturing contextual meaning, and CNNs for extracting sentence-level emotional cues. The DL-based models enhance the system’s ability to interpret subtle linguistic signals, and improving overall prediction accuracy.

References

[1] Magomedova, A., & Fatima, G. (2025). Mental Health and Well-Being in the Modern Era: A Comprehensive Review of Challenges and Interventions. Cureus, 17(1), e77683. https://doi.org/10.7759/cureus.77683

[2] Naslund, J. A., Bondre, A., Torous, J., & Aschbrenner, K. A. (2020). Social Media and Mental Health: Benefits, Risks, and Opportunities for Research and Practice. Journal of technology in behavioral science, 5(3), 245–257. https://doi.org/10.1007/s41347-020-00134-x

[3] Feigin, V. L., Vos, T., Nichols, E., Owolabi, M. O., Carroll, W. M., Dichgans, M., Deuschl, G., Parmar, P., Brainin, M., & Murray, C. (2020). The global burden of neurological disorders: translating evidence into policy. The Lancet. Neurology, 19(3), 255–265. https://doi.org/10.1016/S1474-4422(19)30411-9

[4] Meghrajani, V. R., Marathe, M., Sharma, R., Potdukhe, A., Wanjari, M. B., & Taksande, A. B. (2023). A Comprehensive Analysis of Mental Health Problems in India and the Role of Mental Asylums. Cureus, 15(7), e42559. https://doi.org/10.7759/cureus.42559

[5] Goyal, A. K., Bakshi, J., Panda, N. K., Kapoor, R., Vir, D., Kumar, K., & Aneja, P. (2023). Accuracy of a self-reported Measure in Psychological Assessment when the Instrument is self-administered by the Patient or when Administrated by the Clinician. Indian journal of otolaryngology and head and neck surgery : official publication of the Association of Otolaryngologists of India, 75(2), 1266–1270. https://doi.org/10.1007/s12070-022-03379-5

[6] Zafar, F., Fakhare Alam, L., Vivas, R. R., Wang, J., Whei, S. J., Mehmood, S., Sadeghzadegan, A., Lakkimsetti, M., & Nazir, Z. (2024). The Role of Artificial Intelligence in Identifying Depression and Anxiety: A Comprehensive Literature Review. Cureus, 16(3), e56472. https://doi.org/10.7759/cureus.56472
[7] Mandal, K. P. (2025). A Comprehensive Study: Mental Health Status Prediction Using Machine Learning. Indian Scientific Journal Of Research In Engineering And Management, 09(06), 1–9. https://doi.org/10.55041/ijsrem50947
[8] Narang, G., & Marken, G. (2025). Predicting Mental Health Crises Using Social Media Sentiment Analysis: A Machine Learning Approach. 1–6. https://doi.org/10.1109/icctdc64446.2025.11158851
[9] Nabil, N. F., Galib, A., & Sase, T. (2023). Mental State Detection From Tweets By Machine Learning. International Journal on Perceptive and Cognitive Computing, 9(2), 1–7. https://doi.org/10.31436/ijpcc.v9i2.396
[10] Kabir, Md. H., Samrat, N., Mahmud, A. A., Akter, R., & Raihan, M. (2023). Mental Stress Prediction From the Text of Social Media Using Machine Learning Techniques. 1–7. https://doi.org/10.1109/icccnt56998.2023.10308343
[11] Rawat, N., Chauhan, S., & Awasthi, L. K. (2024). Deep Learning Approaches for Predicting Mental States through Tweet Analysis. https://doi.org/10.1109/icitiit61487.2024.10580551
[12] Abdullah, M., & Negied, N. K. (2024). Detection and prediction of Future Mental disorder from Social Media Data using Machine Learning, Ensemble Learning, and Large Language Models. IEEE Access, 1. https://doi.org/10.1109/access.2024.3406469
[13] Gaonkar, S. L., Naik, V. A., Kumar, A., & Nagamani, S. (2025). Effectiveness of Data Mining Techniques in Identifying Early Signs of Mental Health Issues from Social Media Usage. International Journal of Innovative Research in Engineering & Multidisciplinary Physical Sciences, 13(4). https://doi.org/10.37082/ijirmps.v13.i4.232640
[14] Mohammadi, M. A., Heydarabadi, M. F. M., Zardary, S., & Heysieattalab, S. (2020). Identifying Psychological Disorders Based on Data in Virtual Environments Using Machine Learning. 7(4), 1–12. https://jcp.khu.ac.ir/article-1-3218-en.html
[15] Vijayalakshmi, A., Aruna, T., Divya, N., Chakravarthy, V. J., & Seenivasan, M. (2024). Text Classification on Tweets predicting the Mental Illness using Natural Language Processing and Machine Learning Classifiers. 1–9. https://doi.org/10.1109/iceeict61591.2024.10718368.