2025 IEEE 4th International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things (RAAICON)
27-28 November 2025, Military Institute of Science and Technology, Dhaka, Bangladesh

Survey-Based Analysis of Social Media Behavior and Clinical Scales for Mental Health Prediction Using Machine Learning

1st Md Mahbubur Rahman
Dept. of Computer Science and Information Technology, Computer Science and Engineering, PSTU.
Patuakhali, Bangladesh
mahbub.cse@pstu.ac.bd

2nd Dipraj Howlader
Dept. of Computer Science and Information Technology, Computer Science and Engineering, PSTU.
Patuakhali, Bangladesh
diprajhowladerofficial@gmail.com

3rd Chinmay Bepery
Dept. of Computer Science and Information Technology, Computer Science and Engineering, PSTU.
Patuakhali, Bangladesh
chinmay.cse@pstu.ac.bd

4th Md. Atikqur Rahaman
Dept. of Computer Science and Information Technology, Computer Science and Engineering, PSTU.
Patuakhali, Bangladesh
atik.csit@pstu.ac.bd

Abstract— **The impact of social media on mental health has become a significant global issue. The study uses ethically collected, survey-based behavioral data reflecting individuals' social media usage patterns and their association with mental health outcomes. This research aims to predict mental health outcomes, such as depression, anxiety, and stress, using social media data and machine learning (ML) and deep learning (DL) algorithms. A key challenge in this research is to work with imbalance dataset. Individuals with mental health issues are underrepresented, which makes it hard for predictive models to accurately identify at-risk users. The research applies machine learning algorithms like Logistic Regression, XGBoost, Random Forest, Support Vector Machines, Naive Bayes and so on. It also uses deep learning techniques like Multilayer Perceptron (MLP) to analyze social media usage patterns. The models are evaluated using performance metrics such as accuracy, precision, recall, and F1-score. Also shows correlation between different parameters from user's data. Stratified K-Fold Cross-Validation helps manage the data imbalance and Grid Search Cross-Validation for parameter tunning. The results show that models like Random Forest and Logistic Regression can predict mental health outcomes using behavioral data. The findings emphasize the value of ethically collected, self-reported behavioral data for non-invasive mental health monitoring, especially in resource-limited settings for Bangladeshi people.**

Keywords— Biomedical Informatics, Mental Health Prediction, Behavioral Data Analysis, Social Media Behavior, Artificial Intelligence in Healthcare.

I. Introduction

The rapid growth of social media has changed how people communicate and express their emotions. Platforms like X.Com, Facebook, and Reddit are now central to daily life, offering users spaces to share personal experiences, thoughts, and feelings. This widespread use of social media has raised serious concerns about its impact on mental health. Increasing evidence suggests that excessive use of these platforms is linked to mental health issues such as depression, anxiety, and stress. The data generated on these platforms, including posts, comments, and user interactions, can provide valuable insights into users' emotional states and help in detecting mental health conditions early.

In 2024, about 970 million people around the world live with mental disorders, which is one in eight individuals. This number has risen from 792 million in 2018 and 900 million in 2022. Depression and anxiety are still the most prevalent mental health issues with 5% of adults worldwide affected by depression. Women are twice as likely as men to be diagnosed with depression. This impacts 6% of adult women, compared to 4% of men. Suicide rates are higher among men, who make up 75% of global suicide deaths. Adolescents are also facing growing mental health challenges, as 14% of those aged 10-19 experienced a mental disorder in 2019 [1] [2] [3] [4]. In the U.S., over 5.2 million teens had major depressive cases in 2024 [5]. The COVID-19 pandemic has worsened this crisis, leading to more cases of anxiety and depression. As the demand for mental health support rises, we urgently need to invest in and improve mental health care systems. The economic cost of mental health issues is expected to hit $6 trillion a year by 2030 [6].

Recent research emphasizes the urgent need for systems that can predict mental health issues using social media data. Studies have demonstrated that machine learning models can effectively detect early signs of depression and anxiety by analyzing users' posts and interactions on these platforms [7]. However, challenges remain in achieving high prediction accuracy, finding the correlation, individuals survey based on mental health questionaries (PHQ-9, GAD-7), especially with imbalanced datasets, noisy data, and the varied nature of social media content.

To address these challenges, this study examines the use of various machine learning and deep learning algorithms, including Random Forest, Logistic Regression, Support Vector Machines (SVM), Naive Bayes, and Multilayer Perceptron (MLP), to predict mental health conditions from social media data and finding correlation between different parameters. The focus is on tackling the class imbalance problem using Stratified K-Fold Cross-Validation and GridSearchCV for fine-tuning hyperparameters. The goal is to assess the effectiveness of these algorithms in predicting mental health issues and to aid in developing scalable, non-invasive systems for early detection and intervention.

This paper is organized as follows: Section II reviews the relevant literature on machine learning and deep learning techniques for predicting mental health from social media. Section III presents the proposed methodology, including data preprocessing, model training, and evaluation. Section IV discusses the results of the experiments, and Section V concludes with limitations and recommendations for future research.

II. LITERATURE REVIEW

The rise of social media platforms has changed how people express emotions and interact with others. This shift has led researchers to look at digital footprints as signs of mental well-being. Machine learning (ML) and natural language processing (NLP) are often used to analyze this data for early mental health screening.

For example, sentiment analysis of Reddit posts using models like K-Nearest Neighbors (KNN), Random Forest, and Neural Networks shows that Random Forest achieves an F1-score of 80.6% in detecting depression [7]. While this indicates technical feasibility, the dataset comes from a single platform (Reddit), raising concerns about user diversity, content bias, and generalizability. Additionally, these models usually rely on highly structured or cleaned data that does not reflect real-world usage. Similarly, workplace-based ML studies have used traditional classifiers, including Logistic Regression, Decision Trees, and Random Forests, to predict mental disorders in employees. Decision Trees have shown 81% accuracy in these studies [8]. However, these research efforts tend to focus narrowly on occupational stress and often overlook broader behavioral patterns like sleep, addiction, or platform engagement. This limits the applicability of mental health predictions outside controlled settings. In analyses of doctor-patient consultations, Gradient Boosting (GB) models combined with text embeddings (Word2Vec, BoW, GloVe) achieved 84.2% accuracy in classifying stress and depression [9]. Despite these impressive results, the method relies heavily on clinical dialogue datasets. These datasets are hard to access, not standardized, and unsuitable for predicting outcomes across larger populations. Another NLP framework achieved 77% accuracy for detecting depression based on psychological signals found in social media posts [10]. However, this approach relies only on linguistic signals and misses non-verbal behavioral features like sleep disruption or screen time.

In lower-resource regions, models trained on structured mental health surveys have shown positive results. One study in Kenya used ensemble classifiers to predict depression from PHQ-9 data, achieving 85% accuracy [11]. However, most datasets in this area are either synthetic or collected within narrow age groups, which limits their applicability to larger populations. A thorough review of machine learning models for detecting mental illness found that multimodal data, which combines textual, behavioral, and physiological signals, significantly improves prediction [12]. Still, few studies use these multimodal methods in real-life settings, and even fewer integrate them with platform-level usage data, such as social media activity or device use. Several studies have noted very high accuracy rates. For instance, Logistic Regression reached 99.3% accuracy in predicting poor mental health from survey data involving 481 participants [13], while SVM achieved 93% accuracy on multiple mental disorder screening questionnaires [14]. These results often arise from small, imbalanced, or homogenous datasets, raising concerns about overfitting and external validity. Hybrid approaches like DA-TabSVM, which combines TabNet and SVM, reached up to 99% accuracy using datasets like PHQ-9, GAD-7, and SCL-90 [15]. Likewise, XGBoost achieved over 98% accuracy using DASS42 responses [16]. However, both studies focus on optimization rather than interpretability or behavioral insight, making them unsuitable for clear, real-time interventions. Lastly, a sentiment-based study that used Logistic Regression achieved 87% accuracy in classifying anxiety and bipolar conditions from social media text [17]. While this is promising, it did not consider behavioral context such as sleep, screen time, or addiction. These omissions might weaken the model's effectiveness in real-world applications.

Despite progress in ML-based mental health prediction, important limitations still exist. Many studies rely heavily on pre-labeled or synthetic datasets. These often come from Reddit or clinical sources and lack real behavioral context, which limits their ecological validity. Most models do not include meaningful behavioral features like sleep duration, screen time, or social media addiction. This oversight reduces their interpretability and relevance. Correlation analysis is often underused, with little focus on the strength or direction of behavioral relationships. Furthermore, there is limited research on underrepresented groups in low-resource areas like Bangladesh, as most studies focus on high-income, Western samples. Finally, reproducibility is often weak due to poor reporting of preprocessing, feature engineering, and model validation steps. To tackle these issues, our study combines real-world behavioral data with validated psychological scales (PHQ-9, GAD-7, BSMAS). We apply clear and reproducible ML techniques to create a scalable, non-invasive framework for predicting mental health in underserved populations.

III. PROPOSED METHODOLOGY

The methodology presented in this research aims to predict mental health conditions using machine learning (ML) and deep learning (DL) models, with a focus on social media data. Data collection and data preprocessing initiate the whole process. Then, various algorithms are used, including Random Forest, Logistic Regression, Support Vector Machines, Naive Bayes, and MLP. A simple Logistic Regression model served as a baseline for comparing ensemble and deep learning models. This approach helped ensure a fair evaluation of performance. To address the imbalance in the dataset, Stratified K-Fold Cross-Validation is used, ensuring an equal distribution of mental health and non-mental health cases. Hyperparameter tuning is done with GridSearchCV to improve each model's performance. The evaluation includes standard metrics such as accuracy, precision, recall, and F1-score, which are essential for measuring the model's effectiveness in both binary and multi-class tasks. By combining these methods, the research aims to create an effective and scalable framework for early detection of mental health issues, possibly applicable in real-world settings in the future.

A. Data Collection and Participants

The dataset was created using a Google Form survey shared online with various demographic groups in Bangladesh. Most of the respondents were university and college students. Participants understood the study's purpose and provided consent before submitting their data. The form included three clinically validated tools: the PHQ-9 for depression, the GAD-7 for anxiety, and the BSMAS for social media addiction. Additionally, it contained questions about daily screen time, sleep habits, and emotional reactions to social media.

Inclusion criteria included: (1) Bangladeshi residency, (2) full completion of mental health sections, and (3) consent to participate. Incomplete or non-consenting responses were excluded. Demographic details like age, gender, and occupation are part of the Zenodo dataset repository [18]. Total 750 participants shared their experiences. For external validation, I used an additional dataset titled "Students' Social Media Addiction " from Kaggle Dataset [19].

This dataset includes 705 anonymized survey responses from participants in several countries, including Bangladesh, India, the USA, the UK, and Canada. It covers a variety of age groups, mostly young adults. Respondents consist of high school, undergraduate, and graduate students, representing different academic levels.

B. Data Preprocessing and Feature Engineering

The dataset contained 47 attributes, grouped into four domains: demographics, behavioral indicators, mental health (PHQ-9 and GAD-7), and addiction (BSMAS and binary markers). Categorical variables (e.g., gender, occupation) were encoded numerically using label encoding. Missing values in continuous fields (e.g., sleep duration) were imputed using the mean, and categorical gaps were filled with the mode. Categorical features were numerically encoded to support algorithm compatibility, maintaining ordinal relationships where appropriate (e.g., Occupation: 1 = Student, 2 = Employed, etc.). The dataset was then standardized using a min-max scaling approach to ensure uniformity in feature ranges.

To normalize Likert-based responses, the following transformation was applied:

$$S_m = \left(\frac{\sum_{i=1}^{n} x_i}{n \times 5}\right) \times 10$$

Here, S_m denotes he scaled multi-score, x_i is the response to the i-th item, and n is the number of questions included.

Binary variables (e.g., "Yes"/ "No" indicators for stress or addiction) were encoded and scaled as follows:

$$S_b = \left(\frac{\sum_{i=1}^{n} b_i}{n}\right) \times 10$$

Where S_b is the binary scaled score and $b_i \in \{0,1\}$ for each response. This approach ensures that binary information carries proportionate weight in the dataset. From these, two composite indices were constructed:

Mental Health Score: Reflects inverse distress severity:

$$Mental\ Health\ Score = 10 - \left(\frac{S_m^{(MH)} + S_b^{(MH)}}{2}\right)$$

In this equation, $S_m^{(MH)}$ and $S_b^{(MH)}$ represent the scaled multi-valued and binary-valued mental health responses, respectively.

Addiction Score: Represents overall social media dependency:

$$Addiction\ Score = \left(\frac{S_m^{(AD)} + S_b^{(AD)}}{2}\right)$$

In this equation, $S_m^{(AD)}$ and $S_b^{(AD)}$ refer to the scaled multi-item and binary-item addiction responses.

C. Classification Labels

For binary classification, participants with a Mental Health Score ≤ 5 were labeled as having poor mental health (label = 0), and those above 5 as good mental health (label = 1).

D. Model Development

Several machine learning and deep learning algorithms were explored, including Random Forest, Logistic Regression, Support Vector Machine (SVM), and Multilayer Perceptron (MLP). Each model was trained using Stratified K-Fold Cross-Validation (K=5). This method helped maintain class distribution across folds and avoid bias from data imbalance. Hyperparameters were optimized with GridSearchCV, which tested combinations of key parameters to find the best-performing setup for each algorithm. This methodology aims to provide a scalable and ethically responsible approach to early mental health detection using behavioral data. It focuses particularly on underrepresented populations with limited clinical access. Figure 1 describes the overall procedure.

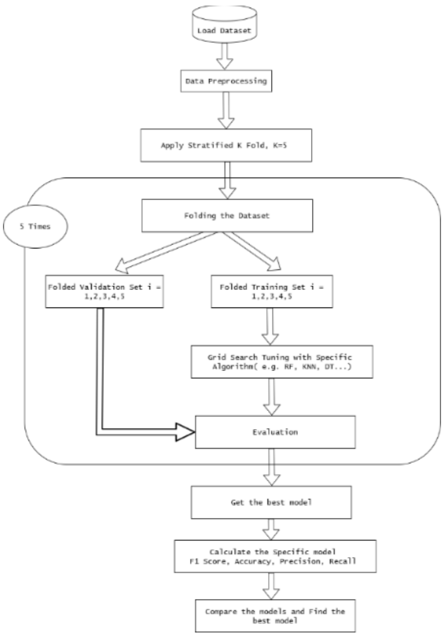

Figure 1: Flowchart of Overall Procedure

IV. RESULT AND DISCUSSION

This section presents the performance of machine learning (ML) models for both binary and multiclass mental health classification with external validation. Evaluation was based on standard metrics: accuracy, F1-score, precision, and recall. A correlation heatmap was also generated to explore behavioral predictors of mental health.

A. Binary Classification Result

Binary classification was performed to differentiate individuals with poor mental health from those with good mental health based on a defined threshold. Random Forest achieved the highest overall accuracy (87.20%) and F1-score (0.9175), making it the most balanced performer. Logistic

Regression showed the highest recall (0.9258), indicating better sensitivity to identifying individuals at risk, while Naive Bayes achieved the highest precision (0.9521), suggesting cautious prediction with fewer false positives. These results are detailed in Table 1 and visualized in Figure 2.

Table 1: Performance Metrics for Binary Mental Health Classification

Model	F1-Score	Accuracy	Recall	Precision
Random Forest	0.9175	0.8720	0.9206	0.9153
SVM	0.9161	0.8693	0.9224	0.9103
Logistic Regression	0.9148	0.8666	0.9258	0.9045
MLP	0.9108	0.8613	0.9172	0.9047
Naive Bayes	0.8803	0.8280	0.8189	0.9521

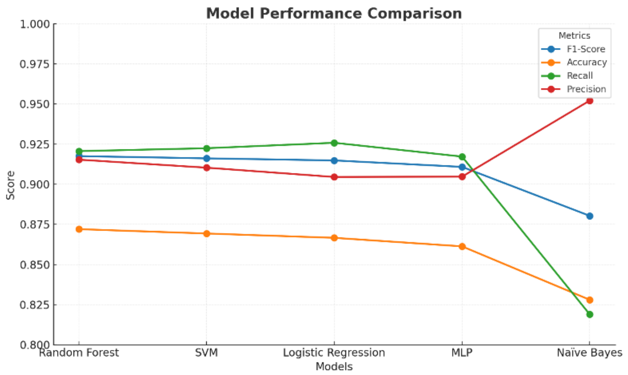

Figure 2: Performance Metrics graph for Binary Mental Health Classification

B. *Multi-Class Classification Result*

For multiclass classification, mental health was categorized into Low, Medium, and High based on score thresholds. XGBoost yielded the best F1-score (0.8346) and shared the highest accuracy (84.67%) with LightGBM, indicating the robustness of gradient boosting algorithms in complex prediction tasks. MLP and Decision Tree performed slightly lower but remained competitive. Results are summarized in Table 2 visualized in Figure 3.

Table 2: Performance Metrics for Multi-Class Mental Health Classification

Model	F1-Score	Accuracy	Recall	Precision
XGBoost	0.8346	0.8467	0.8467	0.8271
LightGBM	0.8333	0.8467	0.8467	0.8226
MLP	0.8309	0.8400	0.8400	0.8267
KNN	0.7976	0.8067	0.8067	0.8021
Decision Tree	0.7893	0.7893	0.7893	0.7918

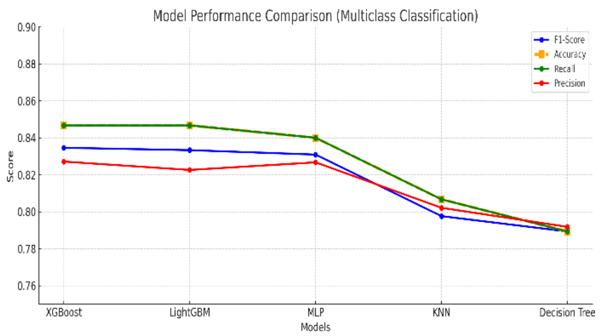

Figure 3: Performance Metrics for Multi-Class Mental Health Classification

C. *External Validation Result*

Table 3 summarizes the results of external validation for four machine learning models. Naive Bayes achieved the highest accuracy at 0.7206 and perfect recall of 1.0000. This shows its strong ability to identify all individuals classified as at-risk. However, its F1-score of 0.5632 was slightly lower, suggesting a moderate balance between precision and recall. Random Forest had an accuracy of 0.6950 and an F1-score of 0.5326. This indicates it was stable across datasets, though it often made conservative predictions.

Table 3: Performance Metrics for external validation dataset

Model	F1-Score	Accuracy	Recall	Precision
Naive Bayes	0.5632	0.7206	1.0000	0.7483
Random Forest	0.5326	0.6950	1.0000	0.7152
MLP	0.5333	0.6906	0.9796	0.7152
Logistic Regression	0.5158	0.6806	1.0000	0.6954

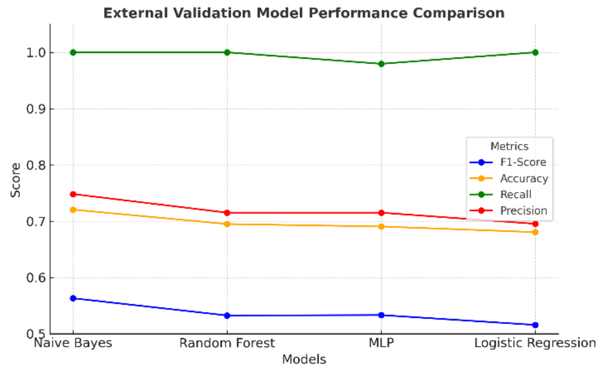

Figure 4: Performance Metrics graph for external validation dataset

D. *Correlation Analysis*

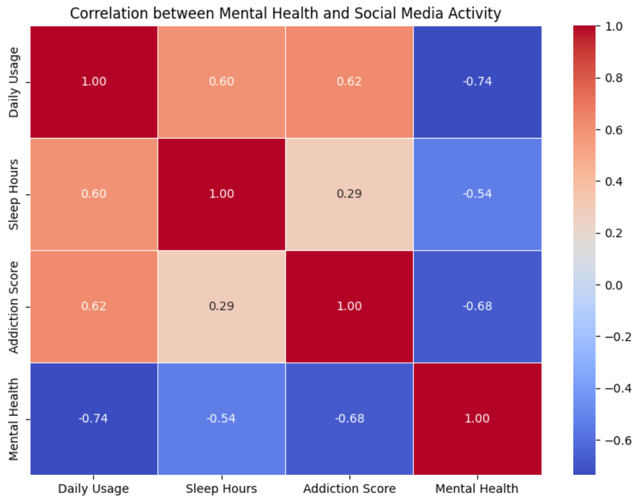

Figure 5: The correlation between mental and other activies

A correlation heatmap (Figure 5) showed significant negative relationships between behavioral features and mental health. Daily social media use had a strong negative correlation with the mental health score (r = -0.74). This was followed by the addiction score (r = -0.68) and sleep duration (r = -0.54). These findings suggest that too much social media use and less sleep are linked to worse psychological well-being. Positive correlations were also noted. Social media use was positively linked to the addiction score (r = 0.62) and sleep duration (r = 0.60). This suggests a possible connection between high usage and changed or longer sleep patterns. A weaker positive correlation between sleep and addiction score (r = 0.29) indicated some dependence among these behaviors.

E. Interpretation and Significance of the Findings

The results show that ML models are effective in classifying mental health status by using behavioral indicators from self-reported data and social media. The strong performance of models XGBoost, Random Forest and Naïve Bayes indicate the possibility of using scalable, non-invasive mental health screening tools.

Additionally, the strong negative link between social media behavior and mental health scores raises concerns about the psychological effects of excessive digital engagement. These findings suggest ways to intervene early through behavioral monitoring and highlight the possibility of integrating these models into digital health systems. When we compare the proposed machine learning models with traditional screening scales like PHQ-9 and GAD-7, we see some interesting results. The traditional scales usually have around 80 to 85% sensitivity and 75 to 80% specificity for detecting depression and anxiety. In contrast, the machine learning models show an accuracy of 86 to 87%. This indicates that the model can work as an additional automated screening tool alongside the current clinical methods. Practical uses might include real-time mental health alerts, wellness tracking apps, or support tools for mental health professionals, especially in low-resource settings.

V. Conclusion

This study looked at how survey-based behavior data on social media use can predict mental health issues among people in Bangladesh, focusing mainly on young adults. The machine learning framework was built using validated tools like PHQ-9, GAD-7, and BSMAS. It showed strong ability to identify individuals at risk for poor mental health. Behavioral signs, such as excessive social media use, digital addiction, and lack of sleep, were strongly linked to anxiety and depression. These results highlight the increasing psychological risks for students and young adults. They also confirm that self-reported, ethically collected behavioral data can be a reliable and non-invasive way to assess mental health early.

To improve reliability, this study included external validation with an independent set of clinically assessed data. This validation confirmed that the model is robust and supports its potential for practical mental health screening. However, there are still limitations. Future research should include validation against data from clinically assessed patients. This will help verify diagnostic reliability and improve real-world applicability in healthcare settings. Moreover, the current dataset mainly consists of students and young adults. To improve model generalization, future studies should include a more diverse group of participants, such as working professionals and people from different socio-economic and cultural backgrounds. Also working with clinical professionals in future studies will help validate medical standards and provide better insight into the accuracy of mental health diagnoses. For real-world use, future studies should compare this model's performance with established screening tools like PHQ-9 and GAD-7 to check its interpretability and fit with clinical standards. The framework could be integrated into digital health platforms or wellness systems in institutions to provide early alerts for counselors, psychologists, and clinicians. When implemented with transparency, fairness, and strong data privacy measures, these tools can enhance traditional mental health assessments and support scalable, technology-assisted well-being solutions.

In summary, this research shows that survey-based behavior data, when paired with clinical scales and externally validated machine learning models, can help create ethical, data-driven, and scalable solutions for early mental health detection. With ongoing improvement and real-world use, these methods can boost mental health awareness, accessibility, and preventive care worldwide.

References

[1] WHO, World Mental Health Report.

https://www.who.int/teams/mental-health-and-substance-use/world-mental-health-report, 2022 (accessed 12 June, 2025)

[2] WHO, Mental Disorders.

https://www.who.int/news-room/fact-sheets/detail/mental-disorders, 2022 (accessed 13 June, 2025)

[3] WHO, Depressive disorder (depression).

https://www.who.int/news-room/fact-sheets/detail/depression, 2023 (accessed 13 June, 2025)

[4] WHO, Mental health of adolescents.

https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health , 2024 (accessed 14 June, 2025)

[5] National Institute of Mental Health, Mental Illness.

https://www.nimh.nih.gov/health/statistics/mental-illness, 2024 (accessed 14 June, 2025)

[6] World Economic Forum, Mental Health

https://intelligence.weforum.org/topics/a1Gb0000000pTDbEAM , 2024 (accessed 14 June, 2025)

[7] K.D. Odja, J. Widiarta, E.S. Purwanto, M.K. Ario, Mental illness detection using sentiment analysis in social media, Procedia Computer Science, 245 (2024), pp. 971-978. https://doi.org/10.1016/j.procs.2024.10.325

[8] S. S. Rautaray, S. Nayak and M. Pandey, "A Machine Learning Based Employee Mental Health Analysis Model," 2023 International Conference on Sustainable Communication Networks and Application (ICSCNA), Theni, India, 2023, pp. 1055-1059. https://doi.org/10.1109/ICSCNA58489.2023.10370378

[9] S. Juanita, A. H. Daeli, M. Syafrullah, W. Anggraeni, M. H. Purnomo, A machine learning approach for text pattern diagnosis in mental health consultations, Decision Analytics Journal, Volume 15, 2025, 100572, ISSN 2772-6622, https://doi.org/10.1016/j.dajour.2025.100572

[10] Yang X, Li G, Psychological and Behavioral Insights From Social Media Users: Natural Language Processing–Based Quantitative Study on Mental Well-Being JMIR Form Res 2025;9:e60286 URL: https://formative.jmir.org/2025/1/e60286 DOI: 10.2196/60286

[11] Y. E. Alharahsheh and M. A. Abdullah, Predicting Individuals Mental Health Status in Kenya using Machine Learning Methods, 2021 12th International Conference on Information and Communication Systems (ICICS), Valencia, Spain, 2021, pp. 94-98, https://doi.org/10.1109/ICICS52457.2021.9464608

[12] Md. Monirul Islam, Shahriar Hassan, Sharmin Akter, Ferdaus Anam Jibon, Md. Sahidullah, A comprehensive review of predictive analytics models for mental illness using machine learning algorithms, Healthcare Analytics, Volume 6, 2024, 100350, ISSN 2772-4425, https://doi.org/10.1016/j.health.2024.100350

[13] S. Purohit, R. Mudgal, S. Vats, P. Rana, O. A. Shah and A. Verma, Analyzing the Impact of Social Media Usage on Mental Health: A Machine Learning Approach, 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1-6, https://doi.org/10.1109/ICCCNT61001.2024.10724267

[14] M. Karunakaran, J. Balusamy and K. Selvaraj, Machine Learning Models based Mental Health Detection, 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT), Kannur, India, 2022, pp. 835-842, https://doi.org/10.1109/ICICICT54557.2022.9917622

[15] Priyanka Arora, Sonika Dahiya, Enhancing Mental Health Diagnosis with DA-TabSVM: A Multi-Class Hybrid Approach for Detecting Depression and Anxiety, Procedia Computer Science, Volume 258, 2025, Pages 1348-1364, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2025.04.368

[16] S. K S, K. K, G. A G, J. D K, P. D. Shenoy and V. K R, Enhancing Mental Illness Prediction using Tree based Machine Learning Approach, 2022 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2022, pp. 1-5, https://doi.org/10.1109/CONECCT55679.2022.9865689

[17] E. Jain and A. Rathour, Understanding Mental Health through Sentiment Analysis: using Machine Learning to Analyze Social Media for Emotional Insights, 2024 5th International Conference on Data Intelligence and Cognitive Informatics (ICDICI), Tirunelveli, India, 2024, pp. 782-786, https://doi.org/10.1109/ICDICI62993.2024.10810855

[18] D. Howlader and M. M. Rahman, Mental Health and Social Media Use Dataset from Bangladesh (Based on PHQ-9, GAD-7, BSMAS) [Dataset], Zenodo, v1, Jul. 31, 2025. https://doi.org/10.5281/zenodo.16621442

[19] Students' Social Media Addiction (Social-Media-Addiction vs Relationships) - *Kaggle Datasets* (A. Shamim, 2023. [Online]. Available: https://www.kaggle.com/datasets/adilshamim8/social-media-addiction-vs-relationships. Accessed: 9 Nov 2025).