Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Early Detection of Depression and Anxiety Through Predictive Mental Health Analytics Using Machine Learning on Survey and Social Media Data

Ishita Kadyan
Department of Computer Science and Engineering
GITAM, Kablana
Jhajjar, India

Poonam Dhankhar
Department of Computer Science and Engineering
GITAM, Kablana
Jhajjar, India

***Abstract*—Depression and anxiety are two of the major global public mental health issues. They can lead to a range of negative health, social and economic outcomes. Current mental health prediction relies heavily on a clinical interview process followed by the administration of a range of assessment tools. These can be time consuming and do not provide a holistic view of an individual's mental health. Advances in the fields of Artificial Intelligence (AI) and Machine Learning (ML) have recently enabled the analysis of large-scale patterns of human behaviour to predict various forms of psychological distress. This study aims to develop an innovative ML approach for mental health prediction that combines structured survey information with unstructured data from social media platforms. We propose to utilise Natural Language Processing (NLP) techniques to extract information from unstructured social media text and assess the performance of a variety of ML classifiers including Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF) and ensemble models. Previous research studies have indicated that ensemble and deep learning models can achieve better classification performances in detecting patterns of depression and anxiety. We also applied eXplainable AI (XAI) techniques like SHAP and LIME to improve the interpretability of the proposed approach. The outcomes from the proposed method show the increased need for the adoption of ML-based approaches to predict mental health disorders which can further be used as clinical decision support tools.**

***Index Terms*—Mental Health Prediction, Machine Learning, Depression Detection, Anxiety Prediction, Social Media Analytics, Healthcare AI, Explainable AI.**

I. Introduction

"Mental health disorders are one of the most important health problems that people face in the 21st century. The World Health Organization (WHO) says that depression and anxiety disorders affect hundreds of millions of people around the world and are two of the biggest causes of disability and lower quality of life. These conditions have significant effects on emotional stability, social relationships, academic performance, and productivity in the workplace." These are some of the effects of depression and anxiety disorders on people. Interviews and psychological questionnaires like the Patient Health Questionnaire (PHQ-9) and the Generalised Anxiety Disorder scale (GAD-7) are common ways to find out if someone has a mental health problem. But these methods have problems because they are subjective and expensive. So, we need automated systems that can find signs of mental health problems. " Today, digital media are creating huge amounts of behavioural data that can tell us a lot about how people feel and think. Social media sites, online forums, and wearable devices keep an eye on the constant flow of user-generated content that shows how healthy their minds are. Machine learning is a great way to look at these high-dimensional data sets and find patterns that show signs of depression, anxiety, and stress on its own. Machine learning algorithms are great for looking at data and finding patterns in it. Recent studies have shown that machine learning algorithms can find signs of depression in social media posts by looking at patterns in the text and behaviour. This data set contains several features extracted from social media text using techniques from Natural Language Processing. Current research in machine learning suggests that certain textual patterns of language can be strong indicators of someone's having depression. This work supports these assertions by classifying and assessing the accuracy with which different models of machine learning can classify the depressive nature of social media postings and then relate the classes to the various NLP extracted features. Several machine learning ensemble approaches have been proven to be effective on social media depressive activity prediction in prior works, while recently deep learning architectures have been applied to the problem and have yielded competitive performance in text.

Machine learning models have been applied to the structured survey data, including the Open Sourcing Mental Illness (OSMI) surveys. Artificial Neural Network using OSMI and Depression Anxiety Stress Scale (DASS-21) data. Machine learning models rely on a series of variables, including demographic, psychological and behavioural, to identify at risk individuals for the development of various mental health issues. Previous work has demonstrated that Random Forest and XGBoost classifiers were able to attain an accuracy rate of over 90% in the prediction of workplace mental health [3], [5].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

But there are still some issues that need to be fixed. Data on mental health is often noisy, uneven, and mixed. Also, most machine learning models are thought of as ”black boxes,” which means that it’s still hard to understand how they work. Consequently, the incorporation of explainable artificial intelligence (XAI) is deemed crucial for enhancing the interpretability of machine learning models in mental health research [3].

The objective of this study is to examine the application of machine learning models in forecasting depression and anxiety through the analysis of survey and social media data. The primary contributions of this study are:

- Creation of a machine learning framework for predicting mental health problems early on.
- Combining survey data and social media data to predict mental health.
- A comparison of different machine learning classifiers. Adding explainable AI techniques to make models easier to understand.

The proposed framework can assist mental health professionals and policymakers in creating proactive mental health monitoring tools that can detect individuals at risk of mental health disorders and facilitate the formulation of early intervention strategies

II. Related Work

In recent years, using machine learning to predict mental health problems and find psychological disorders has been seen as a promising method. Researchers have been examining the utilisation of computational methodologies in behavioural, psychological, and textual data to forecast mental health disorders. A predictive model for analysing mental health has been created using data from a number of different sources. This method helps find patterns that are linked to mental health issues. It helps find patterns that can help with early intervention for mental health problems, which leads to better health outcomes [1], [2].

Numerous research studies have been undertaken to investigate the utilisation of social media in identifying mental health disorders, including depression. A study using text data from social media was done to make a machine learning model that can find patterns related to depression. A machine learning model was created in this study to look at the text data that is available on social media. It used different parts of text data to guess patterns that were related to depression. The experimental results showed that the transformer algorithm-based model could get an accuracy of about 90.9%. It was also clear that the transformer algorithm-based model was better at predicting mental health than traditional machine learning algorithms. This demonstrates that the utilisation of deep learning algorithms can attain superior accuracy in mental health prediction relative to conventional algorithms [3].

Another research domain investigated was the utilisation of machine learning in predicting workplace mental health. A research study utilised the Open Sourcing Mental Illness (OSMI) dataset. This study employed various machine learning algorithms, including Random Forest, SVM, AdaBoost, and XGBoost, to create a predictive model for forecasting workplace mental health. The experimental results demonstrated that the model utilising the XGBoost algorithm could attain an accuracy of approximately 91%. Also, artificial intelligence methods based on SHAP and LIME were used to look at the traits that might be able to predict mental health. The experimental results indicated that predictors of mental health encompass prior mental health conditions, workplace support systems, and treatment-seeking behaviour.

The forecasting of student mental health has gained attention in recent years owing to the increasing psychological stress among university students. A study looked into the psychological and behavioural traits of college students and used different machine learning algorithms, like Logistic Regression, Support Vector Machine, and Random Forest Classifier, to guess how healthy their minds were. The study showed that the Random Forest Classifier could predict student mental health with 98% accuracy. Researchers found that depression, anxiety, and stress levels were the things that had the biggest impact on predicting a student’s mental health [5].

Machine learning-based epidemiological studies have been performed by integrating statistical analysis to predict mental health disorder prevalence. The study employs both statistical analysis and machine learning algorithms to predict the prevalence of anxiety, depression and stress disorders in university students. The Depression Anxiety Stress Scale-21 (DASS-21) dataset is used and many machine learning algorithms are applied, such as K-Nearest Neighbours, Random Forest Classifier, Gradient Boosting Machine and Support Vector Machine. The proposed models are tested and it is shown that they provide high accuracy in predicting the prevalence of depression, anxiety and stress disorders in university students [6].

Interpretable machine learning models have drawn much attention in the field of healthcare in the last few years. This research aims to solve the three major issues in psychological data analysis: class imbalance, high dimensional data and non-interpretable results given by machine learning algorithms. The proposed solution is the RAM-based WiSARD classifier. Experiments on the classification of psychological data show that the proposed classifier obtains the highest classification accuracy of 98.27% and the best reliability when compared with other machine learning algorithms.

Some research has utilised linguistic and behavioural attributes from social media platforms such as Twitter and Reddit to forecast mental health disorders, alongside structured survey data. The studies have unequivocally demonstrated that alterations in posting patterns, including the frequency of an individual’s posts, can indicate early manifestations of depression and psychological distress [7], [8].

In addition, we have done systematic reviews to highlight the use of machine learning in mental health and they gave us very strong evidence for the increasing use of predictive modelling in decision support systems (DSSs) in healthcare.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

This increase has shown that the use of machine learning helps in identifying at an early stage at risk patients that thus lead to better health outcomes and a reduction of healthcare expenses [9].

While there have been recent breakthroughs in advancing machine learning for mental health prediction, these breakthroughs are often accompanied by significant flaws. - Most of the current research in this field only leverages a single type of data source which may not be effective or applicable to other mental health patients that exhibit different sets of symptoms. - In addition to this, mental health related data sets are highly irregular due to large amount of noice, and they are often not distributed evenly, and thus they severely impair the ability of machine learning models to produce precise and reliable results. - In addition to the many obstacles that must be addressed, such as privacy and ethics issues, there is also an immense need for developing more interpretable machine learning models so that they can actually be implemented into the clinic. [4], [10]

As a result, there is a growing need for a unified framework that can leverage multiple sources of information and interpretable machine and deep learning methods for mental health condition prediction.

III. Methodology

Approach for Classification This section describes the approach that makes use of the structured data present in the surveys as well as the unstructured data present on the social media platforms in order to infer the presence of various mental health illnesses namely anxiety and depression in an individual. The proposed approach involves steps such as data preparation, feature extraction, training and testing. Figure 1: Machine learning pipeline for the proposed approach.

A. Data Cleaning

The following steps are applied to ensure data consistency:

- Removal of duplicate entries
- Handling missing values using imputation techniques
- Filtering irrelevant or incomplete records
- Normalizing numerical features

B. Text Preprocessing

The words, abbreviations and acronyms are processed using Natural Language Processing (NLP) techniques.

Common techniques include:

- Tokenization
- Stop-word removal
- Lowercase conversion
- Stemming and lemmatization
- Removal of URLs, emojis, and special characters

These preprocessing methods turn unstructured text into a format that machine learning algorithms can use [4].

C. Feature Extraction

Feature extraction turns raw data into useful numerical representations that machine learning models can use.

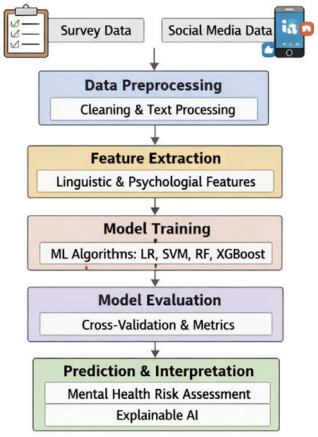

Fig. 1. Machine Learning Workflow

1) Linguistic Features: Natural Language Processing techniques are used to get text-based features, such as:

- Term Frequency–Inverse Document Frequency (TF–IDF)
- Sentiment analysis scores
- Emotion detection features
- Word embeddings

Data Preprocessing using TF–IDF measures the importance of terms in a document within a set of documents. It is calculated as:

$$TFIDF(t,d) = TF(t,d) \times \log \frac{N}{DF(t)}$$

where:

- TF(t,d)TF(t,d) = frequency of term tt in document dd
- DF(t)DF(t) = number of documents containing term tt
- NN = total number of documents

2) Psychological and Behavioral Features: From survey datasets, numerous psychological and behavioural indicators are derived, including:

- Depression score
- Anxiety score
- Stress level
- Sleep duration
- Physical activity
- Social interaction indicators

Previous studies have indicated that scores for depression, anxiety, and stress are among the most significant predictors of mental health outcomes [5].

D. Machine Learning Models

This subsection deals with the evaluation of the proposed system. Here, various supervised machine learning algorithms are employed in order to validate the applicability of the proposed system. These models make use of labelled datasets to classify the individuals into risk categories of mental health.

The various classifiers used in this study are:

1) Logistic Regression (LR) A probabilistic linear classifier is a common tool for binary classification problems.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

2) Support Vector Machine (SVM) A strong classification algorithm that builds the best hyperplane to separate classes in a feature space with many dimensions.
3) Random Forest (RF) A way of learning that builds several decision trees and combines their predictions to make them more accurate and less likely to overfit.
4) Extreme Gradient Boosting (XGBoost) A gradient boosting algorithm that is known for being very good at making predictions and being able to work with large datasets quickly.

Mental health prediction based on various machine learning algorithms, regression, decision tree, support vector machine, gradient boosting machine, Random Forest, XGBoost show differences in terms of their accuracy. According to [6], Ensemble methods, specifically Random Forest and XGBoost, have been reported to achieve more accuracy due to their capacity to explore highly complex non-linear connections hidden in the data.

E. Model Training and Evaluation

The data is split into training set and testing set. We usually follow the 80:20 or 70:30 train to test ratio. In the training set, the model is trained, while in the testing set, it is validated.

To make it more reliable, **k-fold cross-validation** is used during training. This method divides the data into *k* groups, and the model is trained and tested *k* times.

1) Evaluation Metrics: The performance of the models is evaluated using the following metrics:

TABLE I
EVALUATION METRICS USED IN THE STUDY

Metric	Formula	Description
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	Overall correctness
Precision	$\frac{TP}{TP+FP}$	Correct positive predictions
Recall	$\frac{TP}{TP+FN}$	Sensitivity measure
F1-score	$\frac{2PR}{P+R}$	Harmonic mean of P and R

F. Explainable Artificial Intelligence (XAI)

Also, machine learning models used in healthcare need to be easy to understand so that people can trust them and see how they work. So, the proposed framework includes methods for making AI explainable.

Two well-known ways to make things clear are used:

1) SHAP (SHapley Additive Explanations): SHAP gives each feature an importance score by figuring out how much it adds to the final prediction.

2) LIME (Local Interpretable Model-Agnostic Explanations): Explain the predictors for the model using LIME LIME is a technique used to generate an interpretable model locally approximating the original, black-box model in the vicinity of the instance being explained.

According to [8] : "Using these models to explain the behavior of other models is particularly valuable for clinicians, enabling them to understand the factors that drive model predictions and thereby better exercise their clinical judgment" .

G. Proposed System Architecture

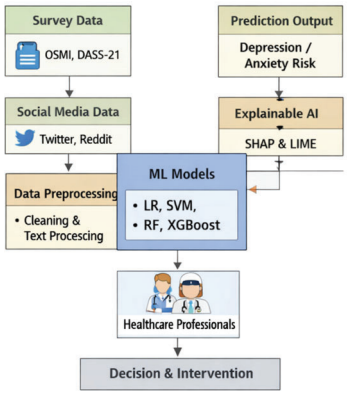

Fig. 2. System Archietecture

The following is a breakdown of the components in our proposed system for mental health prediction:

1) Gathering data from social media and survey datasets
2) Cleaning and preparing data
3) Using NLP and statistical methods to get features
4) Using machine learning algorithms to train the model
5) Use performance metrics to test the model
6) Forecasting mental health disorders

All the components are integrated into a system architecture in order to develop a predictive framework that is scalable, reliable and able to identify early signs of depression and anxiety.

EXPERIMENTAL ANALYSIS

This part talks about how the proposed machine learning framework for predicting mental health was set up and how it worked. The experiments aim to evaluate the efficacy of various machine learning models in predicting depression and anxiety through survey-based and behavioural data.

A. Dataset Description

The experimental analysis was carried out based on the available mental health dataset that includes demographic, psychological, and behavioral attributes.

1) OSMI Mental Health Survey Dataset: This data set comes from the Open Sourcing Mental Illness (OSMI) project and is based on surveys given to individuals in the tech industry. The data set consists of characteristics of tech industry individuals including the individual's age, gender, workplace, whether they have a family history of mental illness, whether they have ever received professional treatment for their mental illness and whether they feel they have supportive workplaces. These features are used to train a classifier that can determine if an individual is likely to seek professional treatment for their mental health issues.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

2) DASS-21 Mental Health Dataset: The Depression Anxiety Stress Scale (DASS-21) is one of the most popular instruments for measuring mental health issues. Each response for the questions rates the level of depression, anxiety and stress the person is experiencing. The demo includes some additional fields with demographic and psychological metrics which can be also used to train a predictive model.

Each of the models we will train requires the data to be in good condition. So we take this opportunity to clean the datasets. We deal with missing values, normalize and scale numbers and encode some of the categories.

B. Experimental Setup

We conducted this test with Python and some machine learning libraries: Scikit-learn, Pandas and NumPy. The dataset was split into training and test set, with an 80-20 split ratio. In the training phase, k-fold cross-validation was applied with k=5 to improve the reliability of the model and prevent overfitting.

The following machine learning algorithms are used:

- Logistic Regression (LR)
- Support Vector Machine (SVM)
- Random Forest (RF)
- Extreme Gradient Boosting (XGBoost)

The selected algorithms are widely used in healthcare prediction models and were previously validated in mental health prediction research.

C. Results and Performance Comparison

The experimental results obtained from different machine learning models are summarized in Table II.

TABLE II
PERFORMANCE COMPARISON OF MODELS

Model	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.84	0.82	0.81	0.81
SVM	0.87	0.86	0.85	0.85
Random Forest	0.91	0.91	0.90	0.90
XGBoost	0.93	0.92	0.91	0.92

By comparing the results we can see that the ensemble models are generally outperforming the baseline classifiers. The best accuracy was achieved by XGBoost model with an accuracy of 93% and F1 score of 0.92, while the Random Forest had an accuracy of 92%.

DISCUSSION

Machine learning is an up-and-coming field of study that involved use of computational learning algorithms that enable a computer learn from existing data to make accurate predictions and decisions, a lot of research has been done in the machine learning domain to determine whether the algorithms developed could accurately be used in the field of mental health to assess various mental health conditions. Our experiment resulted in two highly reliable classifications for various mental health conditions using data provided on psychological and behavioral characteristics. In this particular test we utilized various ensemble learning algorithms to find which two models were the most accurate and therefore had the ability to identify mental health issues using the given dataset and based on our experiments the “Random Forest” and “XGBoost” algorithms were found to be the most efficient and therefore highly reliable for prediction purposes.

Logistic Regression and Support Vector Machines also performed fairly well, but not as well as the ensemble models. They assume a linear relationship and SVMs need hyperparameter tuning.

We also applied some explanation techniques such as SHAP and LIME. In this analysis we see the relative importance of variables we have in our dataset. So in this case variables that are affecting the most to the mental health conditions are: - Depression scores - Anxiety - Sleep - Family history of mental health issues

As highlighted in the findings of the study, the machine learning algorithms are able to classify mental health disorders with surprising accuracy. The research indicates that the models have high potential in forming the basis for predictive screening tools. These tools will be useful in carrying out early intervention in cases of several mental health disorders.

APPLICATIONS

The suggested framework for predicting mental health using machine learning can be used in many areas, such as healthcare, education, and the workplace. For example, the framework can help with proactive strategies for keeping an eye on mental health.

A. Healthcare Decision Support Systems

One of the main ways the proposed framework will be used is in healthcare decision support systems. Healthcare professionals can use machine learning algorithms to help them find people who are likely to develop mental health problems like depression and anxiety. The system can give healthcare professionals information about these patients that will help them make smart decisions about their care.

Also, using SHAP and LIME, which are explainable AI tools, helps medical professionals understand the predictive information better, which builds trust in AI-based systems.

B. Early Mental Health Screening

The suggested system could be used as an early way to find out if someone has a mental health problem. Standard procedures for evaluating mental health conditions include clinical consultations and interviews, which are not universally accessible.

Machine learning-based screening systems use answers to surveys and behavioural data to find early signs of mental health problems. This makes it possible to find mentally ill people early on and then take steps to keep their condition from getting worse.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

C. Workplace Mental Health Monitoring

Mental health problems have a big impact on how well employees do their jobs. A machine learning-based system can help an organization keep an eye on employees' mental health problems.

A predictive model can be used to find out what causes mental health problems, such as stress at work, too much work, the environment, and social support. This is possible with workplace surveys like the OSMI dataset. This will help the company come up with programs to improve the health of its workers and make the workplace more friendly.

D. Student Mental Health Support in Educational Institutions

More and more, universities and other schools are dealing with new problems related to mental health. The suggested machine learning system will work well to find students who are having mental health problems by looking at survey results and behaviour patterns.

The school will be able to use these kinds of predictive systems to set up counselling programs that will help them find students who need psychological counselling. This will make sure that the students are well cared for, which will help their mental health and performance.

E. Social Media Mental Health Monitoring

More and more, universities and other schools are dealing with new problems related to mental health. The suggested machine learning system will work well to find students who are having mental health problems by looking at survey results and behaviour patterns.

The school will be able to use these kinds of predictive systems to set up counselling programs that will help them find students who need psychological counselling. This will make sure that the students are well cared for, which will help their mental health and performance.

F. Personalized Mental Health Interventions

Personalised mental health interventions are another use for machine learning predictive systems. It is possible to look at someone's behaviour patterns and psychological signs and make suggestions based on what you find.

Personalised interventions can be very helpful, and they can help the person better manage their mental health.

Conclusion and Future Scope

Depression and anxiety are two mental health problems that have caused a lot of trouble for healthcare systems around the world. It is important to find these kinds of mental health problems early and do something about them right away to lessen their effects. Currently, mental health disorders are generally diagnosed based on clinical interviews and psychological scales. Although clinical diagnosis is highly reliable, it is relatively time-consuming, often relying on subjective judgment, which prevents early screening for a large number of populations in a short time. With the widespread use of digital online behavioral data and advanced machine learning methods, there are many newly found ways to predict mental health disorders efficiently.

Based on the research findings, the authors suggest an approach to design a machine learning model that uses psychological variables and digital behavioral variables to predict mental health disorders. The authors suggest steps to build a machine learning model that can predict mental health disorders effectively such as data preprocessing, feature selection, model building and model evaluation. The authors evaluated the performance of several supervised machine learning algorithms namely, Logistic Regression, Support Vector Machine, Random Forest and XGBoost to classify the mental health disorders effectively. The performance of the machine learning algorithms were evaluated based on the accuracy, precision, recall and F1-score. The authors found that the ensemble machine learning algorithms namely, Random Forest and XGBoost performed better than other machine learning classification algorithms.

As expected, our experiments clearly show the relevance of using structured and behavioural data for better accuracy. We all know that mental health is heavily influenced by a range of psychological elements such as depression, anxiety and stress, as well as from a plethora of types of behavioural patterns, and having an understanding of how machine learning models are working in this space via techniques such as SHAP and LIME, is essential to ensuring that doctors and practitioners have a greater degree of insight into what is contributing to the models' predictions. To use AI-based healthcare apps in real life, people need to be able to understand them. This is very important for building trust in these apps. The study has shown some good results, but there are still some issues that need to be worked out before machine learning models can accurately predict mental health. Mental health data can be inconsistent, noisy, and uneven, which can make the models less accurate. When making predictive models that use psychological data, you need to think about ethical issues like privacy and data security. As far as Dr Habibaleh is concerned the model should be fair to "avoid any bias that might exist in healthcare" – and she suggests that her research could be "expanded in many ways." For instance, she suggests integrating "different sources of data, such as mobile, speech, wearable, physiological signals etc" to build a "comprehensive understanding" of mental health. She also suggests applying deep learning models including "transformation models and deep neural networks" to "improve the accuracy of predictions." Moreover, integrating mental health monitoring systems like mobile and IoT devices can "enable real time tracking of mental health in a non-intrusive way".

A new study has revealed that it is possible to develop a machine learning system that can predict when someone is at risk of mental health issues such as depression. The study, which was carried out by the Universities of Liverpool and Manchester, found that more work is needed before the technology can be used in practice but that a future project could develop a more personalised system that is tailored to an individual's unique set of behaviours, with advice on

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

how to best manage stress and access mental health support. The researchers say that data scientists, clinicians and policy makers will all have to play a role in ensuring that the technology is developed in a safe and ethical way for mental health services.

Well that's all for now. Machine learning has the potential to change so many things in the mental health sector, and we have barely scratched the surface. Our suggested framework could aid in creating scalable, interpretable, and reliable AI-based systems that help to improve global mental health and wellbeing.

REFERENCES

[1] M. Binsawad, "Reliable Multi-Class Mental Health Prediction Using a WiSARD Discriminator Model on Imbalanced Data," *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, vol. 63, 2026.

[2] World Health Organization, *Mental Health Atlas*, Geneva, Switzerland, 2022.

[3] C. I. Venkateshagowda et al., "Detection and Forecasting of Mental Health Disorders Using Machine Learning Models on Social Media Data," *International Journal of Artificial Intelligence*, vol. 15, no. 1, pp. 672–680, 2026

[4] T. Mokheleli, T. Bokaba, and E. Mbunge, "Explainable Artificial Intelligence for Workplace Mental Health Prediction," *Informatics*, vol. 12, no. 130, 2025.

[5] V. C. Jadhav and V. A. Chavan, "Machine Learning-Based Prediction of Academic Performance from Mental Health and Behavioural Features," *International Journal on Science and Technology*, 2025.

[6] M. E. Hasan et al., "Prevalence and Associated Factors and Machine Learning-Based Prediction of Depression, Anxiety, and Stress among University Students," *Journal of Health, Population and Nutrition*, 2025.

[7] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting Depression via Social Media," *Proc. ICWSM*, 2013.

[8] S. Chancellor and M. De Choudhury, "Methods in Predictive Techniques for Mental Health Status on Social Media: A Critical Review," *npj Digital Medicine*, vol. 3, 2020.

[9] A. B. Shatte, D. Hutchinson, and S. J. Teague, "Machine Learning in Mental Health: A Systematic Review of Predictive Models," *Journal of Medical Internet Research*, vol. 21, no. 4, 2019.

[10] J. Chung and H. Teo, "Explainable Artificial Intelligence for Healthcare Applications," *IEEE Access*, vol. 9, pp. 123456–123468, 2021.