International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Conference Proceedings: Volume: 05 ACME – 2026 DOI: 10.55041/ISJEM.ACME133
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Human Mental Health Prediction using Machine Learning Models

Susmitha Karanam, Masters of Research Methods in Psychology, University of Strathclyde,Glasgow,UK
Shaik Jumlesha, Professor, Annamacharya Institute of Technology,Tirupathi,India
susmitha.karanam.2025@uni.strath.ac.uk[1] ahmedsadhiq@gmail.com[2]

ABSTRACT

Early detection of mental health issues enables specialists to provide more effective treatments, significantly enhancing patients' quality of life. Mental health encompasses psychological, emotional, and social well-being, influencing how individuals think, feel, and behave. It is vital throughout all life stages, from childhood and adolescence to adulthood. This study focused on identifying and evaluating the accuracy of five machine learning techniques in detecting mental health issues. The techniques analyzed were Logistic Regression, K-Nearest Neighbours (K-NN) Classifier, Decision Tree Classifier, Random Forest, and Stacking. We assessed their performance based on several accuracy criteria to determine the most effective method. After comparing these techniques, we found that the Stacking technique, which combines multiple models to improve prediction accuracy, achieved the highest accuracy rate at 82.75%. This indicates that Stacking is a promising approach for accurately identifying mental health issues, potentially leading to better early intervention and treatment outcomes.

Keywords: Mental health, Early Detection, Machine learning Techniques, Accuracy Evaluation, Health Awareness, Mental Wellbeing.

1 Introduction

Mental wellness refers to a state that enables individuals to manage life's challenges, reach their full potential, study and work effectively, and contribute to their community. It is a vital aspect of health and well-being, supporting our capacity as individuals and as a society to make choices, form bonds with one another, and influence the world we live in. Mental health is essential for socioeconomic, communal, and personal growth. It is experienced differently by each individual and exists on a complex continuum with varying degrees of difficulty and distress, leading to diverse social and therapeutic outcomes.

Individual psychological and biological characteristics can increase a person's susceptibility to mental health issues. These factors include emotional intelligence, substance abuse, and heredity. People are more likely to suffer from mental health disorders when exposed to unfavorable social, economic, geopolitical, and environmental conditions such as poverty, violence, inequality, and environmental deprivation. Early detection of mental illness is essential for improving patient outcomes through timely interventions. Predictive analytics, using machine learning, can analyze healthcare data to foresee potential mental health crises. This enables healthcare providers to take preventative measures based on efficient disease prediction.[10]

A Deloitte study found that over 80% of 3,995 respondents across 12 industries experienced adverse mental health symptoms, with depression (59%) being the most common. Major stress sources included workplace-related issues (47%), financial stress (46%), COVID-19 factors (42%), family relationships (39%), and social connections (37%).[12]
While 26% of Indians are stressed due to problems at the workplace, financial instability affects 17%, according to the report. Long working hours, job security, low wages and fierce competition are the dominant contributors. Poor mental health and well-being are one of the major reasons for employees' poor performance at the workplace dropping out of work. The WHO has identified risks to mental health at work can include:

- under-use of skills or being under-skilled for work;
- excessive workloads or work pace, understaffing;

- long, unsocial or inflexible hours;
- lack of control over job design or workload;
- unsafe or poor physical working conditions;
- organizational culture that enables negative behaviours;
- limited support from colleagues or authoritarian supervision;
- violence, harassment or bullying;
- discrimination and exclusion;
- unclear job role;
- under- or over-promotion;
- job insecurity, inadequate pay, or poor investment in career development; and
- conflicting home/work demands.[13]

Mental health is a critical aspect of overall well-being, yet its detection and diagnosis remain challenging. Traditional methods often rely heavily on subjective assessments and lengthy diagnostic processes. However, with advancements in machine learning (ML) algorithms and the availability of vast datasets, there is a growing opportunity to develop more efficient and accurate methods for mental health detection. Advancing AI techniques can help redefine mental illnesses more objectively than the DSM-5, enable earlier identification for more effective interventions, and personalize treatments. However, caution is needed to avoid over-interpreting preliminary results. More work is required to integrate AI into clinical care effectively[6]. This project aims to leverage ML techniques to create a robust system capable of identifying patterns and markers associated with various mental health conditions, thereby facilitating earlier intervention and personalized treatment plans.

2 Recent Works

In recent years, several studies have explored the application of ML in mental health detection. For instance, researchers have used natural language processing (NLP) techniques to analyse text data from social media platforms and electronic health records, identifying linguistic markers indicative of depression, anxiety, and other disorders. Social Network Sites (SNS) to assess mental health levels through user-generated content (UGC).

Additionally, imaging modalities such as functional magnetic resonance imaging (fMRI) have been combined with ML algorithms to detect brain activity patterns associated with specific mental health conditions. These approaches have shown promising results in terms of both accuracy and scalability, laying the groundwork for further advancements in the field.

S.No	Title of the Paper	Problem Statement	Method/ Algorithm Used	Advantages	Disadvantages
1.	Predicting Depression from Social Media Posts[1]	Many individuals express signs of depression on social media, and detecting these signals is crucial.	Natural Language Processing (NLP) techniques, sentiment analysis.	It offers a non-intrusive way to monitor individuals without the need for direct interaction, allowing for continuous observation.	This method is limited to individuals who express their feelings online, excluding those who do not use or share their emotions on social media.
2.	Detecting Anxiety from Physiological Data[2]	Monitoring physiological data (heart rate, skin conductance) to identify anxiety symptoms.	Wearable devices, physiological sensors, machine learning algorithms.	Facial expression analysis allows for real-time detection of emotions, enabling prompt responses and interventions.	The accuracy of emotion detection relies heavily on the availability and accuracy of facial expression datasets, which may not always capture the full spectrum of human emotions.
3.	Machine Learning for ADHD Detection from Behavioral Data[3]	Detecting ADHD based on behavioral patterns observed through machine learning.	Behavioral analysis, machine learning algorithms.	Potential for early identification and intervention.	Interpretation challenges; behavioral variations among individuals
4.	Suicide Risk Prediction from Online Behavior[4]	Predicting the risk of suicide based on patterns in online behavior.	Machine Learning models, analysis of online activity and language use.	Early intervention and support.	Ethical concerns related to privacy and consent.
5.	Emotion Recognition in Text for Mental Health[5]	Identifying emotional states through the analysis of text data.	Natural Language Processing (NLP), sentiment analysis, machine learning.	Applicable to online forums and text-based communication	Limited context understanding; may miss nuanced expressions.

3 Proposed Work Explanation

The proposed project will build upon existing research by integrating multiple data modalities and advanced machine learning (ML) algorithms to enhance the accuracy and versatility of mental health detection systems. Firstly, we will gather diverse datasets that encompass textual, imaging, and physiological data from sources such as social media, electronic health records, and wearable devices. Next, we will preprocess and feature engineer the data to extract relevant information and mitigate noise and biases.

Subsequently, we will develop and train ML models, including deep learning architectures such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), to learn complex patterns and relationships within the multi-modal data. These models will be optimized using techniques such as cross-validation and hyperparameter tuning to maximize performance metrics like accuracy, sensitivity, and specificity.

Overall, this project aims to advance the field of mental health detection by integrating multi-modal data and advanced algorithms. By doing so, we aim to develop a reliable and scalable system for early detection and intervention, thereby improving mental health outcomes.

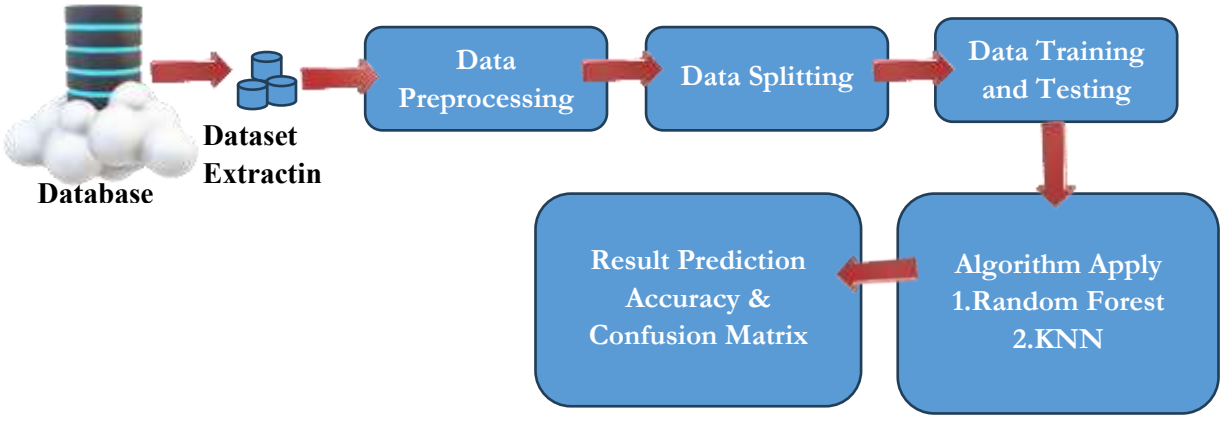
4. Methodology

The methodology for the "Predicting Mental wellness among workspace using Machine Learning models " using Jupyter Notebook and two algorithms, K-Nearest Neighbour (KNN) and Random Forest, involves the following steps:

1. **Data Collection**: The first step involves collecting data on various factors that affect the mental stability of employees working in IT companies. This data can be gathered through surveys, questionnaires, or interviews and may include factors such as job stress, work-life balance, job satisfaction, and personal life factors.
2. **Data Pre-processing**: The collected data needs to be pre-processed to remove any missing values, outliers, and redundant features. This step ensures that the data is clean and ready for model development.
3. **Feature Engineering**: The pre-processed data is transformed into a format suitable for machine learning model development. This involves creating new features that capture the relationship between the input variables and the target variable.
4. **Splitting Data**: The pre-processed data is split into training and testing sets. The training set is used to train the machine learning models, while the testing set is used to evaluate their performance.
5. **Model Development**: Two machine learning models, K-Nearest Neighbour and Random Forest, are developed using the training set. These models are trained on the input features with the target variable being the mental stability of employees.
6. **Model Evaluation**: The models are evaluated using various performance metrics such as accuracy, precision, recall, and F1 score. This step ensures that the models are accurate and reliable, with their performance assessed using the testing set.
7. **Deployment**: The final model, which performs the best on the testing set, is deployed in IT companies to identify employees at risk of developing mental health issues. This allows for timely interventions to prevent the onset of mental health problems.

This methodology uses Jupyter Notebook and the K-Nearest Neighbour and Random Forest algorithms to predict the mental stability of employees in IT companies. The goal is to provide a tool that helps IT companies identify employees who need support and offer timely interventions to improve their overall well-being. The methodology can be customized based on the specific requirements of IT companies.

4.1 ARCHITECTURE FOR PROPOSED SYSTEM



The Architecture of our system is as follows –
1. Imported a dataset from a database.
2. The data was pre processed and splitting will be done for training
3. On the training set of data, the random forest algorithm was used.
4. the testing data was used to apply the evaluation of metrics.
5. Based on the data, that is trained and tested the output was predicted

5. Results and Discussion

The "Human mental health prediction Using Machine Learning models" project, implemented using Jupyter Notebook and two algorithms, K-Nearest Neighbour (KNN) and Random Forest, produced the following results:

1. **Random Forest**: Test accuracy of 0.947273
2. **K-Nearest Neighbour**: Test accuracy of 0.763636

The Random Forest algorithm achieved a higher accuracy score compared to the K-Nearest Neighbour algorithm, indicating that Random Forest is better suited for predicting the mental stability of employees in IT companies based on the given features. The high accuracy score of the Random Forest algorithm suggests it can effectively identify employees at risk of developing mental health issues, allowing for timely intervention to prevent the onset of such problems. Not only employees we want to find mental health issue, now days students also feeling stress , depressed because of various aspects running in their life . Some students will think extreme thoughts and end their life. To prevent those consequences we wish to address their mental health issue on before to become worse.

These results demonstrate the potential of using machine learning algorithms to predict the mental stability of employees in IT companies, students providing valuable insights into their mental well-being. By leveraging the Random Forest algorithm, IT companies can proactively support employees who need it, thereby improving mental health outcomes in the workplace. For students, educational organizations can take counter measures for them to prevent taking wrong decisions. Overall, the project's results highlight the effectiveness of machine learning techniques in enhancing the mental well-being of employees in the IT industry.

Questionnaire used in the survey form :

1. How are you feeling today?
2. How is your eating and sleeping cycle ?
3. (If sad)have you been in the same mental state for the past few days?
4. Is your sadness momentarily or has it been constant for a long time?
5. At what time of the day are you extremely low?
6. Has there been a sudden and huge change in your life?
7. Your stress is related to which of the following areas?
8. How frequently have you had little pleasure or interest in the activities you usually enjoy?
9. How confident you have been feeling in your capabilities recently ?
10. Describe how supported you feel by others around you your friends, family, or otherwise.
11. If you have a mental health condition, do you feel that it interferes with your work?
12. How easy is it for you to take medical leave for a mental health condition?
13. How often do you make use of substance abuse(e.g. smoking, alcohol)?
14. Have you taken any therapy or medication in the near past for mental health?
15. Having trouble concentrating on things, such as reading the newspaper or watching television, or studying?
16. Do you feel bad about yourself or that you are a failure or have let yourself or your family down?

17. If sad, how likely are you to take an appointment with a psychologist or a counsellor for your current mental state?
18. Has the COVID-19 pandemic affected your mental well being?
19. How often do you get offended or angry or start crying ?
20. How likely do you feel yourself vulnerable or lonely?
21. How comfortable are you in talking about your mental health?

The purpose of the questionnaire described above is to collect data on the mental health status of students. It provides a structured method for students to report their experiences by choosing from dropdown options that measure the severity of their symptoms or conditions. The options available typically range from low, indicating minimal impact, to moderate, and high, indicating a significant impact. Once the data collection is done, the user's response is converted using numeric values of 0 to 3, and in some cases 0 to 4.

The process begins with the creation and utilization of a comprehensive dataset to train predictive models tailored for various mental health conditions. These models are developed using machine learning techniques, where the dataset includes previously collected responses from a diverse group of students, annotated with their corresponding mental health conditions.

Students participating in the study are required to complete a survey with 21 carefully crafted questions. Nominal responses like 'Age', 'Country', and 'State' were excluded, and the remaining 18 questions were used as features. Random forest and K-means algorithms were trained using Q15 and Q16 responses as targets to characterize mental illness, followed by a statistical analysis to identify predictors[8] .These questions are designed to capture key indicators of mental health, including their sleeping patterns, the extent of their feelings of loneliness, and their ability to concentrate on tasks. Each response contributes valuable data points that the predictive model uses to assess the student's mental health status.

The primary goal of this survey and predictive model is to provide early identification of potential mental health issues. By analyzing the responses, the model can predict whether a student might be at risk of conditions such as depression, anxiety, or other mental health disorders[9]. This early detection is crucial as it allows for timely intervention, potentially preventing the development or escalation of mental health problems.

Overall, this approach aims to leverage data and technology to enhance mental health support for students, fostering a healthier and more supportive educational environment.

At the same time we also get the survey from the 1091 IT employees and some of them are graduates, based on their working environment, their family history of mental health, how often they get the facility of counselling and therapy from the company , how much important they think mental health as much as physical health , are they work remotely. Same as previous for student data, employees data set also trained by machine learning algorithms random forest and k-nearest algorithm to predict the mental wellness of employee. This will help the companies to improve the productivity from the employees by providing wellness programs by organizational psychologist or counsellors.

Both the Random Forest (RF) and K-nearest neighbour (KNN) algorithms were applied to the dataset, and their results were compared. The Random forest is a technique that solves classification as well as regression problems and it is based on the supervised machine learning approach. However, it is frequently used for classification. Because it combines many decision trees to create a "forest" and feeds random features from the input dataset to them, it is called a random forest and K-Nearest Neighbour is a basic machine learning algorithm that is based on the Supervised earning technique. In the K-NN method, the existing cases and new case/data will be similar. KNN is a non-parametric algorithm that doesn't make any assumption of its underlined data or its distribution. And also it works with multiple classes. The system then posed questions to the users and accurately predicted the stress and composure . Among 1091 employees , 1020 employees and 26 are graduates

were comes under true positive of stress factor and 45 are true negative of composure factor based on the survey questionnaire. The 1067 were in composure state and answered that they suppressed their emotions as a habit or because of the pressure of higher authority. By using K-Nearest neighbouring algorithm the answers were among 1091 , 900 were comes under true positive , 191 are true negative of stress factor. They feel stress but they suppress their emotion being it as a habit or work culture. Some might feel it as stress based on their false negative data from the survey. I took true positive as 'S' ,true negative as 'K' and false positive as ' M' , false negative as ' C' ,Accuracy Rate can be taken as 'R'

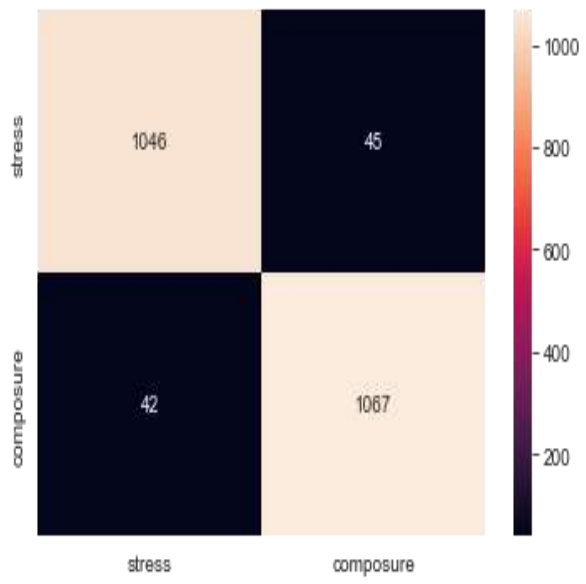
R = sum of diagonals (True Positive) / total number of instances
Error Rate = 1-R

$$\text{Precision (N)} = \frac{S}{S+M}$$

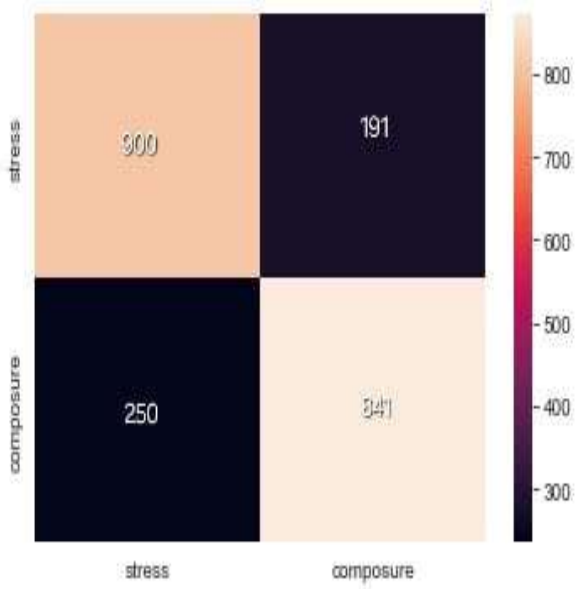
$$\text{Recall (L)} = \text{Sensitivity} = \frac{S}{S+C}$$

$$\text{F1 score} = \frac{2*(N*L)}{N+L}$$

Whereas,
S(True positive) = Diagonals of matrix
C(False Negative) = Sum of the consistent row for class (excluding true positive of that class)
M(False Positive) = Sum of the corresponding column for class (excluding true positive of that class)
K(True Negative) = Sum of the all row and column (excluding row and column of that class)

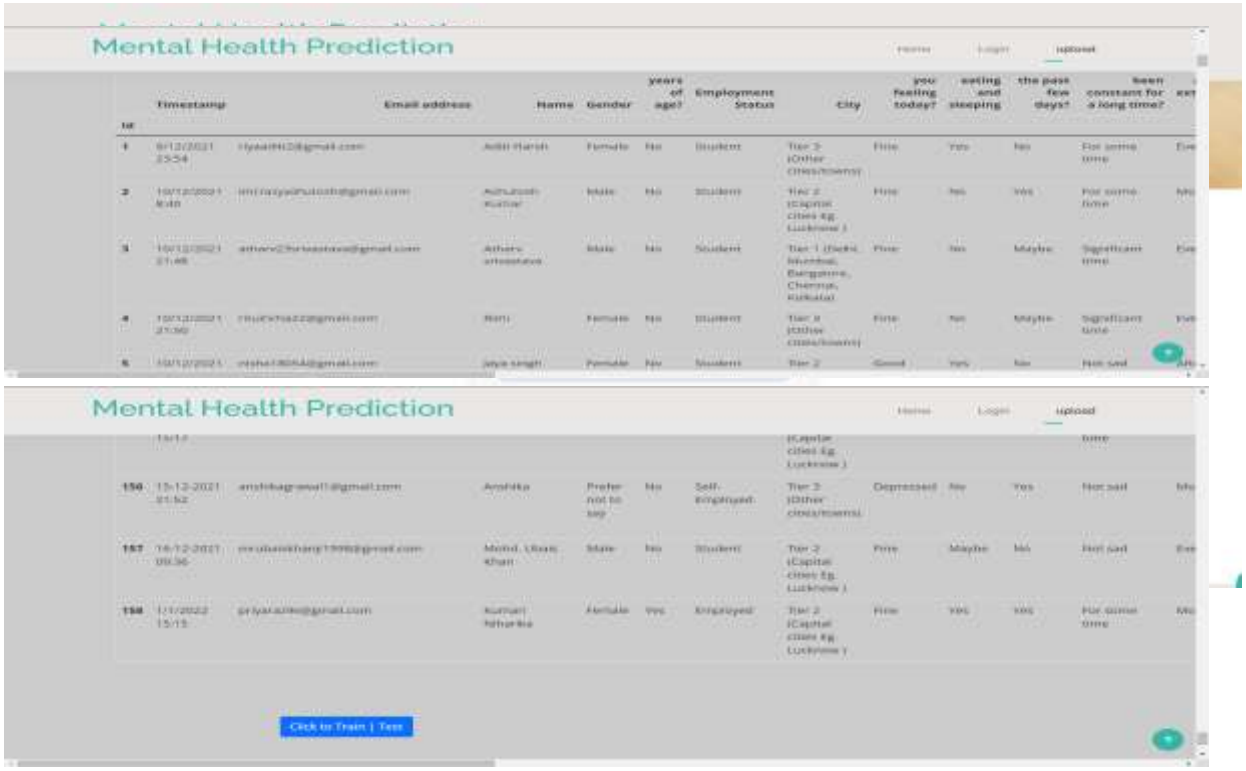


Random Forest



K- Nearest Neighbour

Results :

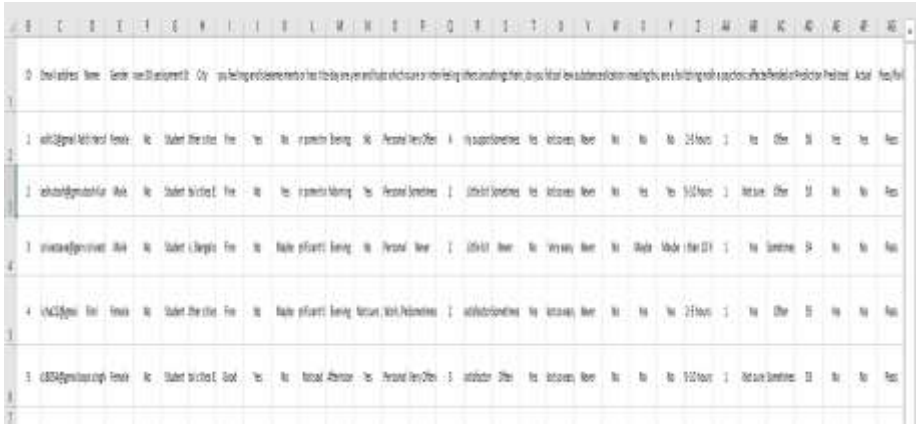


Questions Page :

Output with stress prediction:

5.1 Result Analysis

Dataset testing:



Accuracy plots:

6. Conclusions

In conclusion, the innovative approach introduced for mental health prediction addresses the critical shortcomings of existing methods. By effectively managing the non-linear relationships between feature variables and prediction outcomes, this new method simplifies and enhances the prediction process. Additionally, it tackles the issue of irrelevant and redundant features in prediction datasets, which often hinder the accuracy of conventional models.

A key component of this approach is the calculation of the influence weight of each feature variable, providing valuable insights into how each factor contributes to mental health predictions. This detailed analysis enables a more nuanced understanding of the various determinants impacting mental health outcomes.Overall, the implementation of these advancements promises to significantly enhance the precision and reliability of mental health prediction models. By offering a more accurate and insightful analysis, this method paves the way for more targeted and personalized interventions, ultimately fostering better mental health support and outcomes for individuals.

References

1. Maryam Mohammed Aldarwish, Hafiz Farooq Ahmad," predicting depression from social media posts" ,2017 IEEE 13th International Symposium on Autonomous Decentralized System.

2. Adheena M.A. , Sindhu N., jerritta S," Detecting Anxiety from Physiological Data", IEEE Xplore ,ICSCSS ,2019.

3. Christian Nash (c.nash@newcastle.ac.uk) and Syed Mohsen Naqvi (mohsen.naqvi@newcastle.ac.uk), "Machine Learning for ADHD Detection from Behavioral Data" ,IEEE Xplore,2023.

4. Trehani M Fonseka, Venkat Bhat, Sidney H Kennedy, "Suicide Risk Prediction from Online Behavior" ,volume 53, issue 10, SageJournals, 2019

5. Shiv Naresh Shivhare and Prof. Saritha Khethawat," Emotion Recognition in Text for Mental Health",ResearchGate,2012.

6. Sarah Graham, Colin Depp, Ellen E Lee, Camille Nebeker, Xin Tu, Ho-Cheol Kim, and Dilip V Jeste. Artificial intelligence for mental health and mental illnesses: an overview. Current psychiatry reports, 21(11):1–18,2019.

7. Theodoros Iliou, Georgia Konstantopoulou, Mandani Ntekouli, Christina Lymperopoulou, Konstantinos Assimakopoulos, Dimitrios Galiatsatos, and George Anastassopoulos. Iliou machine learning preprocessing method for depression type prediction. Evolving Systems, 10(1):29–39, 2019.

8. Predicting mental illness at workplace using machine learning December 2022
Mehran University Research Journal of Engineering and Technology 42(1):95
December 2022 42(1):95 DOI:10.22581/muet1982.2301.10

9. M Srividya, S Mohanavalli, and N Bhalaji. Behavioral modeling for mental health using machine learning algorithms. Journal of medical systems, 42(5):1–12, 2018

10. https://therapyhelpers.com/blog/machine-learning-mental-health-prediction/ by Ramesh kaur

11. Mental health and well-being at the workplace Sarkar, Siddharth; Menon, Vikas; Padhy, Susanta; Kathiresan, Preethy : Indian Journal of Psychiatry 66(Suppl 2):p S353-S364, January 2024. | DOI: 10.4103/indianjpsychiatry.indianjpsychiatry_608_23

12. Deloitee. Mental Health and Well-Being in the workplace. 2022. Available from: https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Life-Sciences-Health-Care/gx-mental-health-2022-report-noexp.pdf.

13. World Health Organization. WHO guidelines on mental health at work. 2022. Available from: https://www.who.int/news-room/fact-sheets/detail/mental-health-at-work .