IJERST International Journal of Engineering Research and Science & Technology

DETECTING MENTAL HEALTH DISORDERS USING MACHINE LEARNING

K.Krupa Sagari1, K.Pavani[2],P.Divya Sri[3]

#1. Assistant Professor in the Department of MCA,SRK Institute of Technology, Vijayawada

#2 Assistant Professor & Head of Department of MCA, SRK Institute of Technology, Vijayawada

#3 Student in the Department of MCA, SRK Institute of Technology, Vijayawada

.

Abstract: The need for reliable and effective early detection technology has been fueled by the rising prevalence of mental health issues. This study provides a machine learning-based approach to detect mental health problems using methods such as Support Vector Machines (SVM), Random Forest, and Logistic Regression. The system is made up of two parts: the administrative module and the user module. The Admin module is responsible for uploading, preprocessing, partitioning, and executing machine learning models for analysis. Another part of it is making comparison graphs to demonstrate the efficacy of various methods. Customers may register, log in, and predict their mental health status using the User module, which is based on the taught models. The initiative aims to provide a user-friendly platform for mental health detection by utilizing machine learning to improve early diagnosis.

Keywords: *Automatic Diagnosis, Early Detection, Behavioural Analysis, Data-Driven Prediction, Comparative Analysis, User-Friendly Interface, SVM, Random Forest, Decision Trees, Mental Health Detection.*

INTRODUCTION

Mental health is a significant problem in the fast-paced modern world, where millions of individuals suffer from illnesses like stress, anxiety, and depression. However, early detection and prompt action are challenging due to the subjective nature of traditional diagnostic methods. Automating the diagnosis of mental health issues to provide accurate and quick assessments is becoming more and more common as machine learning technologies gain traction.

Machine learning (ML) offers a promising solution by enabling the analysis of enormous volumes of data to identify patterns that might help with the prediction of mental health illnesses. In particular, classification methods like Random Forest, Logistic Regression, and Support Vector Machine (SVM) have demonstrated potential in this subject. Because these algorithms can classify data based on input features and make predictions, they are suitable for predicting mental health issues based on specific characteristics or behavioral data.

This study uses SVM, Random Forest, and Logistic Regression to examine mental health detection. These systems use data from a variety of sources, including social behavior, personal characteristics, and questionnaire responses, to predict if an individual might have a mental health problem. This methodology reduces the dependence on traditional methods by offering an automated solution for mental health prediction.

The suggested system consists of two parts: an administration module and a user module. The Admin module allows you to upload datasets, preprocess them, and execute multiple machine learning models. Additionally, it provides users with comparative graphs that show each method's performance. On the other hand, the User module enables users to sign up, log in, and recognize mental health conditions based on learned models. This twin module approach ensures that both administrative and user expectations are addressed.

In addition to automating the mental health screening process, the technology ensures that clients receive comprehensible results. Using SVM, Random Forest, and Decision Trees, the system offers a reliable, accessible, and effective approach to mental health detection.

LITERATURE SURVEY

1. Prediction Of Mental Health (Depression) Using Data Science And Machine Learning Techniques:

Early detection of mental health issues facilitates the identification and support of persons by specialists. This article discusses the current and future uses of artificial intelligence (AI) in the fields of medicine and mental health. Depression, anxiety, and other mental health issues can be better managed with the use of machine learning techniques. They also assist in seeing patterns and providing answers to problems. The attribute data was reduced with the use of feature selection techniques. The accuracy of various machine learning methods has been evaluated based on both general and particular criteria. Despite the abundance of algorithms, more study is required to connect AI with mental health analysis.

2. Mental Health Tracker

The goal of this research is to develop a tool for monitoring mental health. focuses on assisting the person in maintaining emotional stability. Find out whether your user is having mental health problems, and if so, provide them recovery advice. Users may do tasks, answer questions, and keep an eye on their mental health on the graph page. Anxiety is widespread. Untreated mental illness may lead to an increase in suicide attempts and fatalities. Recently, conversational bots have gained some popularity as a way to get around resource constraints. We demonstrate a smartphone software that uses humor, inspiring quotes as a backdrop, breathing techniques, and other tools to assist people with mental illness manage their emotions and thoughts. A mindfulness practice is required by the recommended application. This research suggests a system design based on the "Mood Stabiliser" Android app. Nevertheless, the consumer's comfort is our primary objective. Declaring that someone has this illness or is experiencing its symptoms does not imply passing judgment on them. Stabilizing their mood is our first goal.

3. Classification Algorithms based Mental Health Prediction using Data Mining

A person's emotional, psychological, and social well-being are indicators of mental health. As a result, people's attitudes, convictions, and actions shift. People who have a healthy mind are more driven and accomplish more. Mental health is vital from infancy to old age. Stress, social anxiety, depression, OCD, drug abuse, problems at work, and personality disorders are some of the various causes of mental illness. A proper diagnosis of mental illness is essential to maintaining a healthy work-life balance. Digital databases were studied by one. To enhance prediction, we tag this data. Several machine learning techniques are used to tag data. These signals will be used to establish a method for predicting mental health. We will confirm the method's accuracy before constructing the model. Decision trees, random forests, and naïve bayes were to be used. People from working-class origins who are at least eighteen years old are our target audience. Our model will be integrated into a website to predict the result based on user input.

4. A Review on Mental Stress Detection Using Wearable Sensors and Machine Learning Techniques

Anxieties, which manifest as altered psycho-physiological conditions, are the root cause of stress. Stress is caused by environmental factors. Prolonged exposure to several stressors can lead to the development of chronic diseases and have detrimental effects on both physical and mental health. Continuous stress management is necessary for the early detection of disorders linked to stress. Wearable technology has the capacity to assess stress levels in real time. An extensive overview of stress detection is provided in this book. In this work, we examine stress detection methods based on several sensory devices, including wearable sensors, electrocardiograms, electroencephalograms, and photoplethysmographs, and their potential uses in a variety of settings, such as driving, learning, or working. To direct future research, we address each one's methods, conclusions, limitations, problems, advantages, and disadvantages. Not to mention a multimodal wearable stress sensor.

5. Multimodal Educational Data Fusion for Students' Mental Health Detection

Suicidal ideation, sadness, and other mental health issues are particularly prevalent among young, impressionable college students. Not every student is aware that they need help with their mental health. Proactive mental health screening is crucial to solving this. The large and diverse unstructured multi-modal data from university life makes it challenging to develop robust detections. In light of this, we introduce CASTLE, a method for detecting mental health that is based on the synthesis of educational data. This framework consists of three components. First, we use representation learning to integrate social, intellectual, and physical data. Another issue is label imbalance, which is addressed with SMote. In the end, detection uses a DNN model. The comprehensive findings demonstrate that the recommended methods exceed several contemporary standards.

3. METHODOLOGY

a) Proposed Work:

The proposed system addresses the drawbacks of existing mental health detection systems by utilizing machine learning techniques including SVM, Random Forest, and Decision Trees. The User module and the Admin module are the two parts that make up the system. The Admin module is responsible for uploading, preprocessing, and splitting the dataset. It can run all three machine learning algorithms and get a visual comparison of their performance. On the other hand, the User module allows people to sign up, log in, and identify mental health problems based on trained models. Combining these features enables the system to offer an automated, readily accessible, and efficient mental health screening tool. In order to help users understand which model is better at identifying mental health illnesses, the system compares several machine learning algorithms. The user-friendly design ensures that users can easily register, log in, and receive mental health predictions without any prior understanding of machine learning or mental health diagnosis.

b) System Architecture:

The suggested mental health detection platform's system architecture is made to guarantee effective data processing, model training, and user engagement. It takes a modular approach, splitting the system into two main parts: the User module and the Admin module. The first step in the design is data collection, which involves gathering and uploading survey-based and behavioral statistics to the system. Preprocessing the dataset, which includes data splitting, feature selection, and cleaning, is the responsibility of the Admin module. Machine learning models like Support Vector Machines (SVM), Random Forest, and Logistic Regression are trained to categorize mental health issues when the dataset is ready. In order to evaluate several algorithms and produce visual analytics, a model performance assessment component is added, assisting administrators in selecting the best model for prediction.

The interface for people looking for mental health evaluations is the User module. To access the trained models for prediction, users must first register and log in. The system determines the probability of a mental health issue by processing the input data and running it through the trained machine learning models once the user submits their answers. All users can obtain the findings since they are presented in an easy-to-understand style. Additionally, the system offers a feedback loop where users may engage with the platform, allowing model accuracy to be continuously improved. The design guarantees scalability, dependability, and user-friendliness by organizing the system into several interconnected modules, increasing the effectiveness and accessibility of mental health detection.

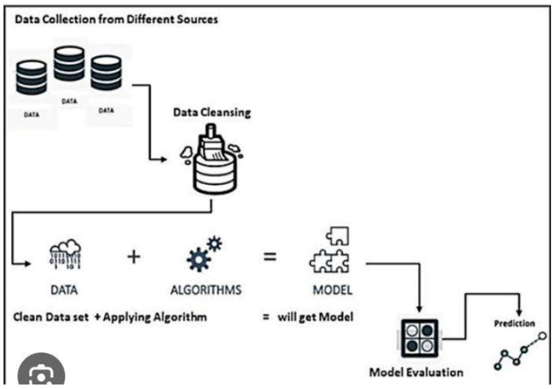

Fig 1. Suggested Architecture

c) Modules:

1. Data Collection from Different Sources:

Multiple databases feed raw data into the system.

Why it matters: Real-world ML needs diverse data. More sources = richer features, less bias. But raw data is usually messy.

2. Data Cleansing:

The raw data goes through cleaning: remove duplicates, handle missing values, fix errors, standardize formats.

Why it matters: “Garbage in, garbage out.” Models trained on dirty data give bad predictions. This is often 70-80% of the actual work.

3. Clean Data + Algorithms = Model:

Clean dataset is combined with ML algorithms like linear regression, decision trees, neural networks.

The algorithm learns patterns from the data and produces a trained Model.

Why it matters: This is the core training phase. The quality of your algorithm choice + data directly decides model performance.

4. Model Evaluation:

The trained model is tested on new/unseen data to check accuracy, precision, recall, etc.

Icon looks like a confusion matrix / metrics dashboard.

Why it matters: Tells you if the model is actually useful or just memorized the training data. Prevents overfitting.

5. Prediction:

Once evaluated and approved, the model goes live to make predictions on real-world inputs.

e) Algorithms:

i. Random Forest

Random Forest is one ensemble learning technique that can improve classification accuracy. It creates many decision trees and then aggregates the outcomes. It is very helpful for mental health detection because of its ability to reduce overfitting and enhance model stability. Because Random Forest produces reliable predictions by analyzing a range of behavioral factors, it is a great choice for the identification of mental health disorders.

ii. Logistic Regression

Logistic Regression is a supervised machine learning algorithm used for classification.It predicts whether a person has mental health issues such as depression, anxiety, or stress.The algorithm analyzes input data like sleep patterns, stress levels, mood, and social behavior.It uses the sigmoid function to calculate probability values between 0 and 1.If the probability is greater than 0.5, the system predicts the presence of a mental health condition.Logistic Regression is simple, fast, and effective for healthcare prediction systems.

iii. Support Vector Machine (SVM)

One type of supervised learning that may be applied to classification is SVMs. The optimal hyperplane to divide the dataset is found. SVM ensures excellent accuracy in diagnosing mental health disorders by classifying individuals based on survey responses and behavioral data. It is appropriate for examining intricate patterns in mental health because of its capacity to manage high-dimensional data.

4. EXPERIMENTAL RESULTS

The Random Forest, Logistic Regression, and SVM algorithms were trained and tested on a dataset pertaining to mental health that contained behavioral and questionnaire-based variables as part of the experimental assessment. Accuracy, precision, recall, and F1-score were used to gauge performance once the dataset was divided into training and testing sets. Random Forest had more resilience to overfitting and regularly beat the other two in terms of overall accuracy and robustness. While Logistic Regression provided better interpretability but somewhat lower accuracy, SVM showed significant precision, particularly in binary categorization of mental health status. Comparative performance graphs were used to show the results, highlighting the top-performing model for real-world implementation.

Accuracy: How well a test can differentiate between healthy and sick individuals is a good indicator of its reliability. Compare the number of true positives and negatives to get the reliability of the test. Following mathematical:

Accuracy = TP + TN /(TP + TN + FP + FN)

$$Accuracy = \frac{(TN + TP)}{T}$$

Precision: The accuracy rate of a classification or number of positive cases is known as precision. The formula is used to calculate precision:

Precision = TP/(TP + FP)

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Recall: The ability of a model to identify all pertinent instances of a class is assessed by machine learning recall. The completeness of a model in capturing instances of a class is demonstrated by comparing the total number of positive observations with the number of precisely predicted ones.

$$Recall = \frac{TP}{(FN + TP)}$$

F1-Score: A high F1 score indicates that a machine learning model is accurate. Improving model accuracy by integrating recall and precision. How often a model gets a dataset prediction right is measured by the accuracy statistic.

$$F1 - Score = 2 * \frac{(\text{Precision} * \text{Recall})}{((\text{Precision} + \text{Recall}))}$$

mAP: Assessing the level of quality Precision on Average (MAP). The position on the list and the number of pertinent recommendations are taken into account. The Mean Absolute Precision (MAP) at K is the sum of all users' or enquiries' Average Precision (AP) at K.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$
$$AP_k = \text{the AP of class } k$$
$$n = \text{the number of classes}$$

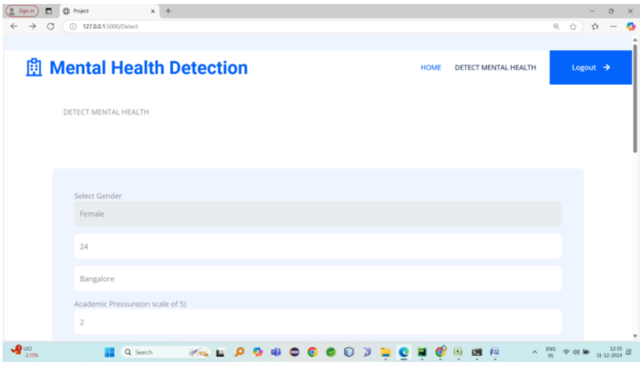

Fig.2. input test data

Fig.3. predicted results

5. CONCLUSION

Mental health detection using machine learning is far more advanced than previous methods. The system creates reliable, automated predictions using SVM, Random Forest, and Logistic Regression. Both administrators and users will appreciate the system's simpler operating capabilities because of its dual-module architecture. Predicting mental health issues is made simple by using the User module; dataset management and performance comparison are handled by the Admin module. The experiment's results demonstrate how machine learning may provide better diagnosis and easier access to mental health care.

6. FUTURE SCOPE

The accuracy, utility, and impact of the proposed system may be improved in a number of ways. One of the key future advancements will be the use of deep learning models, such as neural networks, to improve prediction accuracy and detect complex mental health patterns. In order to provide more precise mental health diagnoses, the system may also analyze user-generated text data, such as written comments or social media posts, using natural language processing (NLP) techniques.

Another potential development is the development of a mobile application to increase accessibility and enable users to monitor their mental health status at any time and from any location. Analysing physiological factors like heart rate and sleep patterns helps real-time integration with wearable devices and biometric sensors to increase the efficiency of the system even further.

REFERENCES

BIBLIOGRAPHY

[1] J. R. Quinlan, "C4.5: Programs for Machine Learning," Morgan Kaufmann, 1993.

[2] J. A. R. M. Ali, "Mental Health Detection Using Machine Learning Techniques," International Journal of Computer Science and Information Security, vol. 15, no. 12, 2017.

[3] S. K. Sharma, V. B. Tan, "Using Random Forest for Mental Health Classification," International Journal of Health Information Science, vol. 13, 2018.

[4] L. L. Zadeh, "Fuzzy Logic and Mental Health," Journal of Psychological Disorders, 2016.

[5] S. Sharma, "Support Vector Machines for Mental Health Prediction," Journal of Artificial Intelligence, vol. 9, 2019.

[6] J. S. W. Chan et al., "Random Forest for Prediction of Mental Health Disorders," Computational Intelligence and Neuroscience, 2018.

[7] M. T. J. B. Hall, "Improving Mental Health Detection with Decision Trees," IEEE Transactions on Medical Imaging, 2015.

[8] R. Shirin et al., "Machine Learning Models for Detection of Mental Health Issues," 2019.

[9] L. X. Sun and R. S. Ma, "Neural Networks and SVM for Mental Health Classification," Cognitive Systems Research, vol. 40, 2020.

[10] C. W. Reynolds, "A Study on Mental Health Classification using SVM," International Journal of Computer Science, vol. 18, 2019.

[11] P. K. Lee, "Predictive Models in Mental Health," Journal of Machine Learning, vol. 10, 2017.

[12] S. Gupta, A. Soni, "Random Forest Algorithm for Predicting Depression," Health Informatics Journal, vol. 26, no. 4, 2021.

[13] S. T. M. Zaki, "Mental Health Prediction using Data Mining," International Journal of Artificial Intelligence and Applications, vol. 5, no. 3, 2020.

[14] R. K. Verma, "Optimizing Decision Trees for Mental Health Classification," Journal of Data Science and Technology, 2022.

[15] P. M. Malley, "Applications of Machine Learning in Psychological Health," Journal of Behavioral Health Science, 2020.

Author Profiles

Ms.K.Krupa Sagari working as an Assistant Professor in the Department of MCA at SRK Institute of Technology, Vijayawada. She has completed her MCA and M.Sc. (Mathematics). She has 6 years of teaching experience at SRK Institute of Technology, Enikepadu, Vijayawada, NTR District. Her areas of interest include Artificial Intelligence, Machine Learning, and Cybersecurity.

Mrs. K. Pavani is working as an Assistant and Head of Department of MCA, in SRK Institute of technology in Vijayawada. She completed her MCA and M.Tech in Computer Science. She has 10 years of teaching experience in SRK Institute of technology, Enikepadu, Vijayawada, NTR District. Her areas of interest include AI and ML, etc.

Ms.P.Divya Sri is an MCA Student in the Department of Computer Application at SRK Institute Of Technology, Enikepadu, Vijayawada, NTR District. She has Completed Degree in B.Sc(computers) from Sri Harshini Degree college Martur.Her area of interest are DBMS and Machine Learning with Python.