IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Machine Learning for the Determination of Mental Health -SVM,Random Forest and DecisionTree

Ankem Tarakram[1,a)] and Ayancha Bharath[2,b)] ,Singareddy V.N.Swetha[3,c)] , C.S.R.D.Santosh[4,d)] Javvaji Mounika[5,e)]

[1] Asst. Professor, Department of Computer Science and Engineering, NRI Institute of Technology, Agiripalli-521212, Vijayawada, Andhra Pradesh, India

[2,3,4,5]B. Tech Student, Department of Computer Science and Engineering, NRI Institute of Technology, Agiripalli-521212, Vijayawada, Andhra Pradesh, India

Abstract— Mental health issues are increasingly prevalent, necessitating reliable and automated diagnostic tools. This study leverages survey and behavioral data to predict mental health illnesses using Support Vector Machine (SVM), Random Forest, and Decision Tree algorithms. The system comprises two essential modules: Admin and User. The Admin module facilitates dataset management, preprocessing, model training, and performance comparison through visual analytics. The User module allows individuals to sign up, log in, and receive mental health predictions from trained models. By integrating various machine learning techniques, the system enhances diagnostic accuracy and model efficiency. This approach minimizes reliance on traditional survey-based evaluations while enabling real-time, user-friendly mental health detection. Machine learning-driven mental health prediction holds significant potential for expanding access to early assessment and intervention.

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

I. INTRODUCTION

Mental health has become a critical concern in today's fast-paced world, with conditions such as depression, anxiety, and stress affecting millions globally. Traditional methods for diagnosing mental health disorders rely on surveys, self-assessments, and professional evaluations, which can be time-consuming and subjective. Early detection and timely intervention remain significant challenges, as individuals may not always recognize symptoms or seek professional help due to stigma or accessibility issues. The need for automated systems is therefore increased and efficient mental health detection systems that can provide early warnings and assist in timely intervention.

Machine learning (ML) has emerged as a powerful tool for analyzing large datasets and identifying patterns that may indicate mental health conditions. By leveraging classification algorithms such as SVM, RF, and Decision Trees, it is possible to predict mental health disorders based on behavioral traits, social interactions, and responses to structured surveys. These algorithms enable data-driven decision-making, reducing the reliance on traditional diagnostic methods and offering a more objective and efficient solution. The Admin module is responsible for managing datasets, preprocessing data, training models, and comparing the performance of different algorithms using visual analytics. Meanwhile, the User module allows individuals to register, log in, and predict their mental health status based on trained models. This dual-module approach ensures ease of use for both administrators and end-users, making mental health detection more accessible and reliable.

By integrating multiple machine learning algorithms, the proposed system provides a comparative analysis of model performance, helping users understand which approach works best for mental health prediction. The system aims to enhance early diagnosis, provide real-time assessments, and offer a user-friendly interface for individuals seeking insights into their mental health. This approach has the potential to revolutionize mental health diagnostics by minimizing subjectivity, improving accessibility, and enabling data-driven decision-making.

II. LITERATURE SURVEY

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

https://sist.sathyabama.ac.in/sist_naac/documents/1.3.4/1822-b.e-cse-batchno-145.pdf

Professionals are better able to recognise and help people with mental health issues when they are caught early. Discussed in this essay are the current and future uses of artificial intelligence (AI) in the field of mental health and healthcare. Depression, anxiety, and other mental health issues can be better managed with the use of machine learning techniques. Furthermore, they have the ability to spot patterns and provide answers to problems. By utilising Feature Selection techniques, the attribute data was minimised. The accuracy of various machine learning algorithms has been evaluated on both all and certain attributes. Despite the abundance of algorithms, further study is required to reconnect AI with mental health analysis.[1]

Mental Health Tracker

https://www.academia.edu/102929972/Mental_Health_Tracker

This project is developing a tool to monitor mental health. Deals with stabilising the user's mood as well. Find out if your user is suffering mentally, and if so, give them advice on how to get well. Users may track their mental state, complete tasks, and answer questions all on the graph page. Anxieties are prevalent all throughout the world. The number of suicide attempts and fatalities caused by untreated mental illness may rise. As a solution to resource restrictions, conversational bots have become increasingly popular in recent years. We present a mobile app that includes breathing exercises, encouraging quotes as background, jokes, and more to aid those with mental illness in managing their emotions and thoughts. The application recommends including a mindfulness practice. Using the "Mood Stabiliser" software for Android as a foundation, our essay suggests a system design. However, reassuring the user is our primary objective. Therefore, stating that someone has this illness or is suffering from it does not constitute criticism of them. Stabilising their mood is our first objective.[2]

Classification Algorithms based Mental Health Prediction using Data Mining

https://ieeexplore.ieee.org/document/9137856

A person's emotional, psychological, and social health are indicators of their mental health. The way individuals feel, think, and behave is impacted. Having a healthy mind allows people to put in more effort and accomplish more. From infancy to old age, one's mental health is vital. Stress, social anxiety, depression, OCD, drug abuse, problems at work, and personality disorders are just a few of the many causes of mental illness. Accurate diagnosis of mental illness is crucial for maintaining a healthy work-life balance. Digital databases were consulted. For better prediction, this data is tagged. Data is labelled using a variety of machine learning techniques. These labels will be used to develop a model for predicting mental health. We shall check the algorithm's accuracy before we create the model. Classification methods such as Decision Tree, Random Forest, and Naïve Bayes were to be employed. We're aiming for adults from working-class backgrounds who are 18 and over. To predict the result depending on user input, our model will be integrated into a website.[3]

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

ttps://www.researchgate.net/publication/352099909_A_Review_on_Mental_Stress_Detection_Using_Wearable_Sensors_and_Machine_Learning_Techniques

Anxieties are the root cause of stress, which manifests as an increased psycho-physiological condition. Stress is caused by environmental factors. Negative effects on mental and physical health, as well as the development of chronic diseases, can result from prolonged exposure to many pressures. The early detection of stress-related illnesses requires continuous stress monitoring. Wearable gadgets can monitor stress in real-time. This work offers a complete examination of stress detection. In this research, we take a look at stress detection methods that rely on various sensory devices, such as wearable sensors, electrocardiograms, electroencephalograms, and photoplethysmographs, and how they might be applied in various settings, such as while driving, studying, or working. To help direct future studies, we mention each one's methods, results, limitations, concerns, strengths, and weaknesses. Not to mention a multi-modal stress detecting system that may be worn[4]

Multimodal Educational Data Fusion for Students' Mental Health Detection

https://www.researchgate.net/publication/361652706_Multi-modal_educational_data_fusion_for_students%27_mental_health_detection

Depression, suicidal ideation, and other mental health issues are common, particularly among young, impressionable college students. When it comes to mental health, not all students know they need help. Proactive screening for mental health is necessary to address this issue. Unstructured multi-modal data from campus life is rich and variable, making it challenging to develop solid detections. In light of this, we provide CASTLE, a method that uses educational data fusion to detect mental health issues. This structure is composed of three parts. We begin by using representation learning to integrate physical, academic, and social data. The second issue is that SMOTE deals with label imbalance. In the end, a DNN model is used for detection. According to the detailed findings, the proposed methods are superior to many state-of-the-art benchmarks.[5]

III. METHODOLOGY

A) Proposed System

The proposed system improves mental health detection by utilizing ML algorithms such as SVM, RF, and Decision Trees. These models analyze behavioral data and survey responses to provide automated, real-time predictions, reducing reliance on traditional, time-consuming survey-based assessments. The system ensures accurate and efficient mental health detection, enabling early intervention and minimizing subjectivity in diagnosis. By integrating multiple machine learning techniques, it offers a comparative

analysis of algorithm performance, allowing users to understand which model works best for mental health prediction.

The system is divided into two modules: the Admin module and the User module. The Admin module handles dataset management, including uploading, preprocessing, and splitting the dataset for training, while also running machine learning models and generating performance comparison graphs. The User module allows individuals to register, log in, and receive mental health predictions using the trained models. With a user-friendly interface, the system ensures accessibility for non-technical users, making mental health assessments more convenient and widely available. This structured, data-driven approach enhances the reliability and usability of mental health detection, contributing to better awareness and timely diagnosis.

B) System Architecture

The system architecture of the proposed mental health detection platform is designed to ensure efficient data processing, model training, and user interaction. It follows a modular approach, dividing the system into two key components: the Admin module and the User module. The architecture begins with data collection, where behavioral and survey-based datasets are gathered and uploaded to the system. The Admin module is responsible for preprocessing the dataset, which includes cleaning, feature selection, and data splitting. Once the dataset is prepared, machine learning models such as Support Vector Machines (SVM), Random Forest, and Decision Trees are trained to classify mental health conditions. A model performance evaluation component is integrated to compare different algorithms and generate visual analytics, helping administrators choose the most effective model for prediction.

The User module serves as the interface for individuals seeking mental health assessments. Users can register and log in to access the trained models for prediction. When a user submits their responses, the system processes the input data and runs it through the trained machine learning models to determine the likelihood of a mental health condition. The results are displayed in a clear and comprehensible format, ensuring accessibility for all users. The system also provides a feedback loop where users can interact with the platform, enabling continuous improvement of model accuracy. By structuring the system into these interconnected modules, the architecture ensures scalability, reliability, and ease of use, making mental health detection more efficient and accessible.

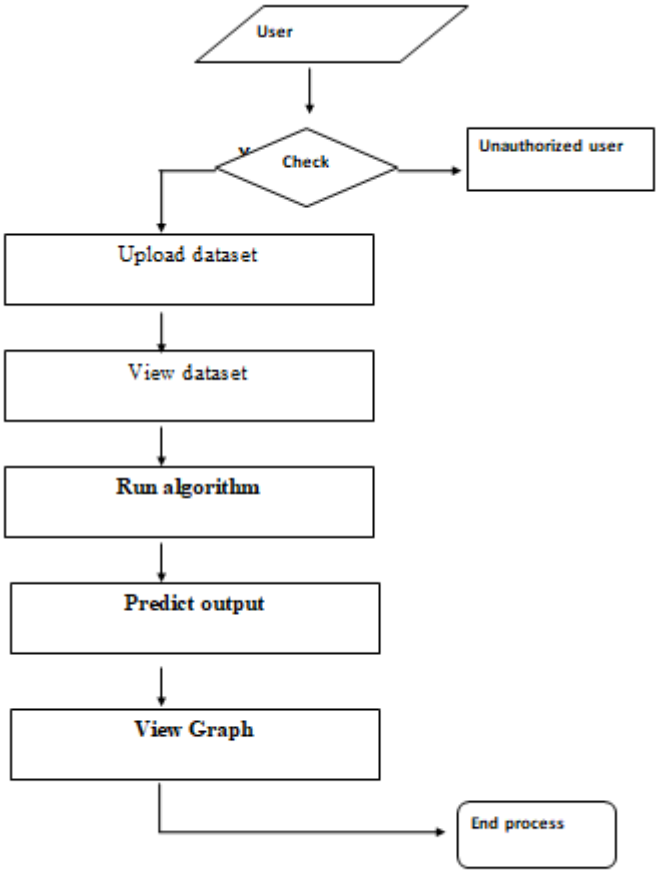

Fig 1. Suggested Architecture

Fig 1 presents system architecture which depicts the mental health detection of an user and they can check their mental health conditions through the application in which the data is extracted from large data set of users and user registrations and login and then the model runs the algorithm after loading the dataset into machine and predicts the desired output from the model.

C) MODULES

a) Admin Module

The Admin module is responsible for managing the dataset and machine learning models. Administrators can upload datasets, preprocess data (cleaning, feature selection, and splitting), and train machine learning algorithms such as Support Vector Machines (SVM), Random Forest, and Decision Trees. This module also includes a comparison feature that generates performance graphs, allowing administrators to analyze which model provides the most accurate predictions. Additionally, the Admin module ensures that the system remains updated with the latest datasets and optimized models for improved detection accuracy.

b) User Module

The User module provides a user-friendly interface for individuals to register, log in, and predict their mental health status. Users can input their behavioral or survey-based responses, which are then processed by the trained ML models to generate predictions. The system presents the results in an easy-to-understand format, helping users gain insights into their mental health condition. This module ensures accessibility for individuals without technical expertise and enables real-time mental health assessment, making early diagnosis more convenient and effective.

D) Algorithms

a) Random Forest

One ensemble learning approach that can increase classification accuracy is Random Forest. It builds numerous decision trees and then combines their results. Due to its capacity to decrease overfitting and improve model stability, it is incredibly useful for mental health detection. Random Forest is an excellent option for the detection of mental health illnesses since it generates trustworthy predictions by examining a variety of behavioural parameters.

b) Decision Tree

A straightforward and effective classification technique, Decision Tree divides data into branches according to feature values. Because it is easy to understand and visualize, it finds extensive usage in mental health screening. This algorithm aids in comprehending the elements that lead to mental health disorders by organizing decisions in a tree-like fashion, rendering forecasts more explicable and visible.

c) Support Vector Machine (SVM)

SVMs are a kind of supervised learning that may be used for categorization. It locates the best hyperplane to partition the dataset. In mental health detection, SVM helps classify individuals based on their behavioral data and survey responses, ensuring high accuracy in identifying mental health conditions. Its ability to handle high-dimensional data makes it suitable for analyzing complex mental health patterns.

IV. EXPERIMENTAL RESULTS

Accuracy: How well a test can differentiate between healthy and sick individuals is a good indicator of its reliability. Find out how reliable a test is by comparing real positives and negatives. Following mathematical:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \text{ -------- (1)}$$

Precision: The accuracy rate of a classification or number of positive cases is known as precision. Accuracy is determined by applying using the one that follows:

$$\text{Precision} = \frac{TP}{TP+FN} \text{ --------- (2)}$$

Recall: The recall of a model is a measure of its capacity to identify all occurrences of a relevant machine learning class. A model's ability to detect class instances is shown by percent of correctly anticipated positive observations relative to total positives.

$$\text{Recall} = \frac{Tp}{TP+FN} \text{ --------- (3)}$$

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.

$$\text{F1 Score} = 2.\frac{Precision.Recall}{Precision+Recall} \text{ -------- (4)}$$

MAP: Information retrieval system performance is measured by MAP, which stands for Mean Average Precision. It finds the mean precision for all classes or queries. While accuracy measures the validity of results, precision determines the mean accuracy for all queries. MAP evaluates the system's performance by averaging the AP scores across all queries or classes.

$$\text{MAP} = \frac{1}{N} \sum_{i=1}^{N} \text{AP}_i$$

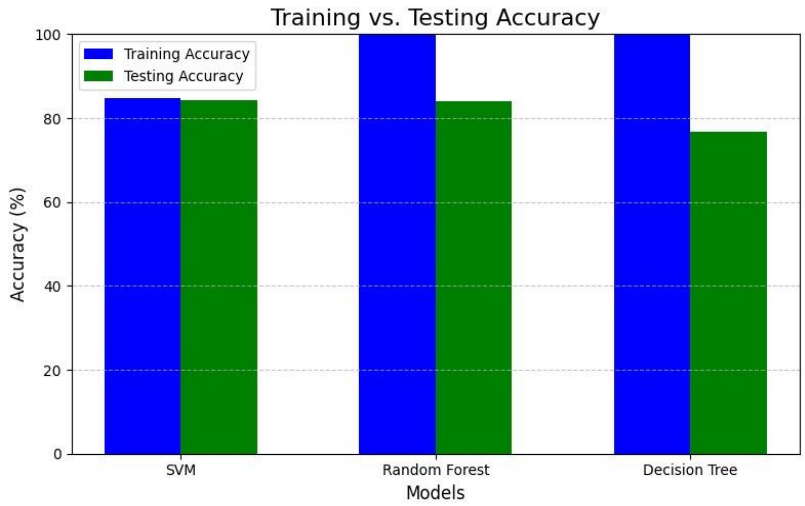

Fig.2. input test data

Fig 2 presents accuracy of **Mental Health Detection** system. It shows that the support vector machine(SVM),Random Forest, Decision Tree model learns strongly without much overfitting, showing no overfitting, with validation accuracies being 84.23, 84.14, 76.86 respectively.

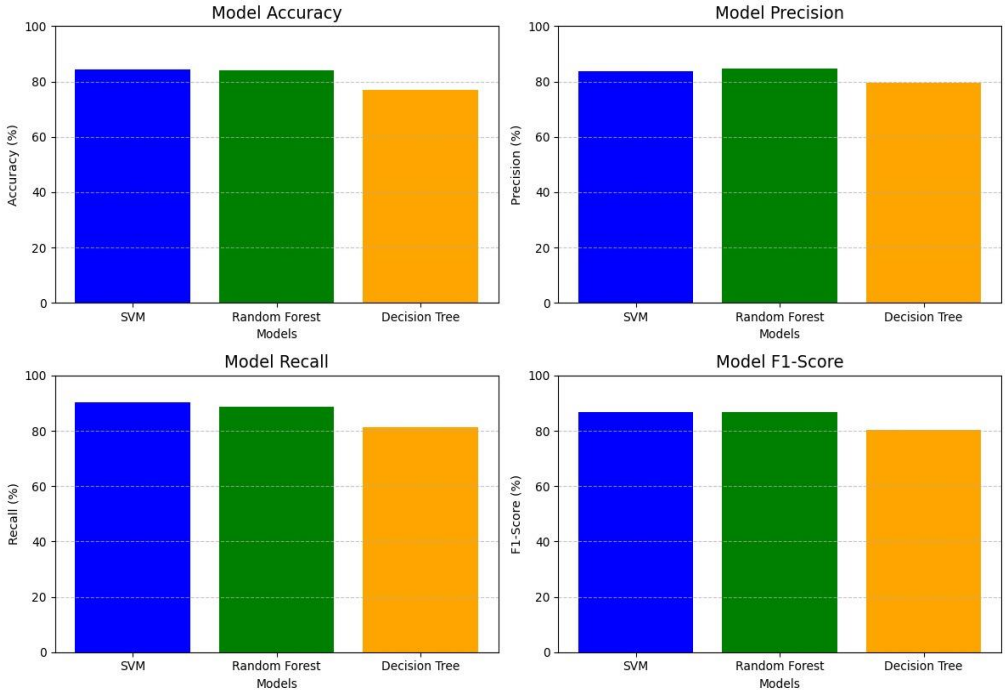

Fig.3. predicted results

Fig 3 presents the overall evaluation metrics (model accuracy, Precision, Recall, F1-score) of the model which is trained on a data set with the well classified machine learning algorithms support vector machine, Random Forest, Decision Tree.

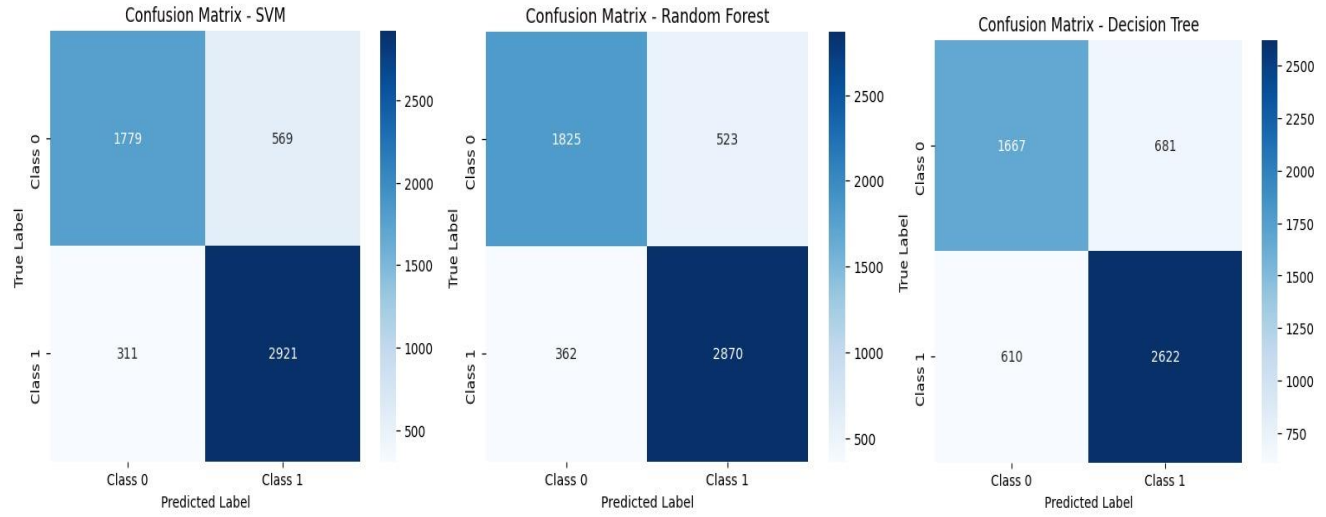

Fig 4:Confusion matrix

Fig 4 presents the confusion matrix of the machine learning model which is trained with SVM identified the model as discriminating 2921 positive tests from 311 negative ones with 569 errors in each category. Hence, it is evident that SVM model works well with high accuracy. The other two matrix shows other two model's accuracies. Machine learning model which is trained with Random Forest identified the model as discriminating 2870 positive tests from 362 negative ones with 523 errors in each category. Hence, it is evident that Random Forest model works well with high accuracy Machine learning model which is trained with SVM identified the model as discriminating 2921 positive tests from 311negative ones with 569 errors in each category. Hence, it is evident that SVM model works well with high accuracy

Results of three machine learning models

SNo	Metric	SVM	Random Forest	Decision Tree
1	Accuracy	84.23	84.14	76.86
2	Precision	83.7	84.59	79.38
3	Recall	90.38	88.8	81.13
4	F1-Score	86.91	86.64	80.24

Table 1:Comparision Table

This table presents a **comparison of different machine learning algorithms based on their accuracy**. It helps in evaluating which model performs the best for a given classification or prediction task.

- **SVM has the highest accuracy (84.23%)**, making it the best-performing model among the three.
- **Random Forest follows closely with 84.05% accuracy**, showing competitive performance.
- **Decision Tree has the lowest accuracy (76.2%)**, indicating it might not be the best choice for this dataset.

V. CONCLUSION

Comparing to previous methods, machine learning-based mental health identification is light years ahead. Automated, efficient, and reliable predictions are made by the system using SVM, Random Forest, and Decision Trees. Admins and users alike will appreciate the system's streamlined operation thanks to its dual-module design. Predicting mental health issues is a breeze using the User module, while the Admin module takes care of dataset management and performance comparison. The results of this experiment demonstrate the potential of machine learning to provide easier access to mental health care and better diagnoses.

VI. FUTURE SCOPE

The proposed system can be enhanced in several ways to improve its accuracy, usability, and impact. One of the key future improvements is integrating deep learning models such as Neural Networks to enhance prediction accuracy and detect complex mental health patterns. Additionally, incorporating Natural Language Processing (NLP) techniques can allow the system to analyze user-generated text data, such as social media posts or written responses, to provide more accurate mental health assessments.

Another potential advancement is the development of a mobile application to increase accessibility, allowing users to check their mental health status anytime and anywhere. Real-time integration with wearable devices and biometric sensors can further improve the system's efficiency by analyzing physiological parameters like heart rate and sleep patterns.

REFERENCES

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.