Machine Learning in Mental Health

R.Soundharya[*1], Nabeeha Banu[#2], Niveditha Nanda Kumar[#3], Shravya Kamath U[#4], Moazzami Mukthar Ahmed[#5]

[*]Assistant Professor, Dept. of Computer Science and Engineering, Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal, Udupi, India.

[#]UG Student, Dept. of Computer Science and Engineering, Shri Madhwa Vadiraja Institute of Technology and Management, Bantakal, Udupi, India.

***Abstract*— Modern people suffer from a variety of behavioral and mental diseases as a result of the added stress and pressure in their daily lives. Just a few examples of mental health problems include schizophrenia, bipolar disorder, anxiety, depression, and stress. Physical and emotional symptoms can both be present in mental illness. This examination will determine whether a person has a mental disorder based on their behaviors and thoughts. Symptoms of mental illness include delusions and hallucinations, as well as panic attacks, sweating, palpitations, melancholy, worry, and excessive thinking. Each symptom denotes a certain type of mental disorder. For this study, only a few machine-learning techniques were used, including decision trees, logistic regression, SVM, and KNN. In this work, we used a feature selection strategy that included a second tree classifier with other pre-processing methods. Based on a patient's symptoms and the feature extraction technique, mental health stability has been identified using a machine-learning algorithm. The effectiveness of machine learning models was evaluated using Recall, Accuracy, Precision, and F1-score.**

***Keywords*— Mental Health, Machine Learning, Machine Learning Algorithms, Accuracy, Supervised Learning,**

I. INTRODUCTION

From birth through adulthood, mental health affects every phase of a person's life. A person's circumstances, views, psychological state, and physical state all change with time. Because these modifications directly affect their ability to think, feel, behave, and be stable, they can all be referred to as having a mental health disease. Psychiatric problems may affect people's sense of stability and security as well as the environment. Due of the rise in mental health difficulties and the demand for high-quality medical care, researchers have applied machine learning algorithms to solve challenges related to mental health. Workers and pupils are among those who battle with mental health concerns. To spot any health-related irregularities, it is essential to monitor the mental health profiles of different populations. It is fairly clear that experts are regretfully more concerned with the current problems that people with conditions face than those who are mentally ill. Most of the time, they are ignorant of the concepts and techniques used to promote mental wellness.

As a result, mental health issues are ignored or underappreciated. Some of the most common mental health conditions include bipolar disorder, depression, and anxiety disorders. They typically have early-stage symptoms which require time to develop. The pervasive and sometimes unquantifiable symptom of economic hardship is what fuels the vicious cycle of sickness and poverty. Technology developments have enabled all classes to be reached. The ability to analyse these data using machine learning has grown. Mobile phones, social networks, neurology, and technological advances have allowed medical professionals and psychology researchers to swiftly gather an enormous quantity of data. We are now able to obtain certain forecasts in ways that were previously not conceivable thanks to machine learning. Machine learning algorithms may be useful in recognising crucial behaviours of the user in order to measure his mental health. The article presents an application that utilizes machine learning that addresses to track a user's mental well-being.

II. LITERATURE REVIEW

Norah Saleh Alghamdi et al. [1] want to investigate the use of a text-analysis tool in an intelligent application to enhance mental wellbeing. It detects depression and anxiety via multiple techniques integrated into sophisticated electronics and image sensors. They evaluated several machine learning classification algorithms and found that the SVM they chose worked most effectively, scoring 79.81% on the text evaluation tool.

Yash Jain et al. [2] offer an extensive healthcare management system that takes into consideration both physical as well as mental health. The Arduino UNO chip and the IOT sensor network are utilised, and an emotion acceptance gadget is included to keep track of the person's feelings. The emotion recognition system obtained a 66% accuracy after retraining CNN.

Piyush Kumar et al [3] used a previously recorded dataset to predict mental illness in humans and characterise the patient's perspective. They used Logistic Regression, KNN, and SVM in Ensemble model 1 to get an accuracy rate of 89.603%, and Decision Tree, Nave Bayes, and SVM in Ensemble model 2 to achieve an accuracy rate of 87.539%.

Ahmad Rauf Subhani, et al [4]. The main contribution of this research is the development of an experimental paradigm for successfully inducing stress at various levels as well as the provision of a framework combining EEG data analysis for stress detection at various levels. The recommended framework effectively identified stress with a maximum accuracy of 94.6% between

two degrees of stress and the control and 83.43% between stress and the other levels of stress. The results suggest that stress levels may be precisely determined by EEG signals.

The article by Ruyi Wang et al[5] advocates employing a chatbot to track and assess the psychological well-being of pregnant women. It uses controlled machine learning to develop a model to measure the levels of anxiety, depressive disorders, and low mood index of perinatal women using a Support Vector Machine (SVM) method.

The authors [6] propose a system design based on the Android app "Mental Health Tracker," which detects a user's mental health. For the sake of privacy of information, the application incorporates a verification method. For sentiment evaluation, the system combines Neural Processing Language (Text Blob), Machine Learning (algorithm random forest), and MySQL.

Apoorva Bagul and colleagues [7] The group plans on using a mental health tracker to analyze the user's state of mind, figure out if they are distress, and provide gauges to help them out of their current state of mind. A dashboard reflects a user's emotional state based on their replies on particular questions and advises steps based on their responses.

Braden Tabisula, et al [8] proposes a smartphone application that users can download. The user shouldn't squander time pondering their current strategy while under stress, worry, or sadness. Instead, the app can start the process for a speedier and more efficient resolution to their symptom after asking a few questions to assess the individual.

Philip Moore, et al [9]. In order to effectively monitor patients with mental problems in Smart-Psychiatric Intensive Care Units and the community, this study examines the practical difficulties that must be overcome. It was concluded that all parties involved in the treatment of mental diseases will profit from efficient patient monitoring.

Jetli Chung, et al[10] summarizes the findings of a recent in depth analysis of machine learning algorithms for identifying mental health issues in this paper. The systematic review was carried out utilizing the PRISMA methodology. It takes into account the challenges and limitations that machine learning researchers in the field of mental health face. They also provided detailed suggestions for further research.

III. METHODOLOGY

Our intention is to forecast the development of mental disease, and supervised learning is the ideal approach for this. Supervised learning is the most extensively used technique in various sorts of research, investigations, and experiments, particularly in predicting sickness in the medical area. In supervised learning, all data instances should represent the words, characteristics, and values. In more technical terms, supervised learning is a classification approach that takes use of structured training data. The collection of datasets from the data source, analysis of the data, and data processing. Choose the finest attributes as inputs for improved prediction. Using matplot lib or pyplots, generate graphs for visualisation of data and research. The dataset has to be developed and verified. Fit the machine learning model, predict using inputs, evaluate the model, and determine the accuracy. Combine a machine learning model with the Streamlit web framework to improve user interface and experience.

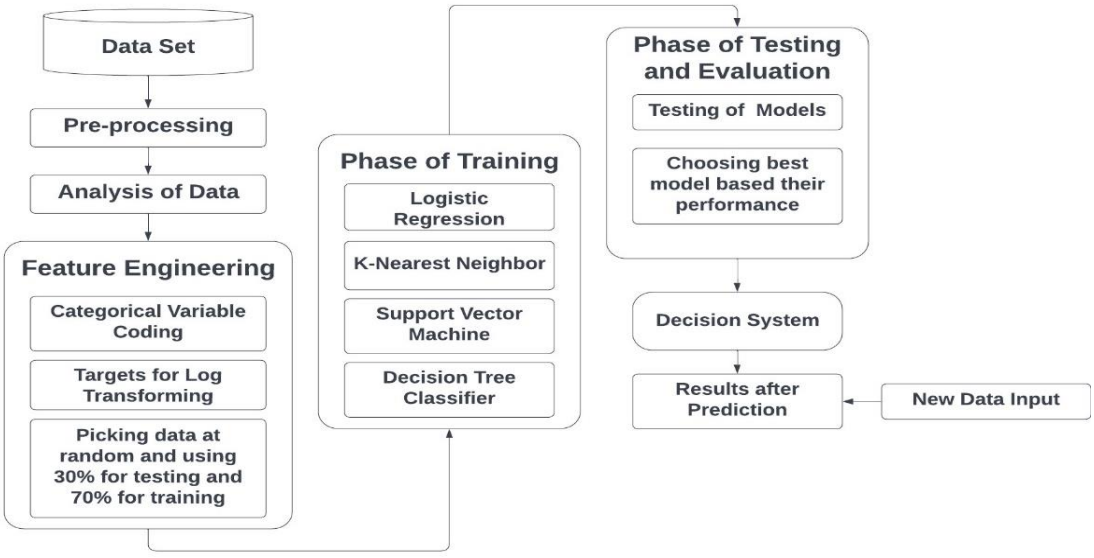

Fig.1 System Architecture

In order to forecast success, we use a total of four different machine learning algorithms: SVM, KNN, Decision Tree and Logistic Regression. We assessed different metrics to evaluate the efficiency of a machine learning algorithm.
The project is divided into six major modules:

- ❖ Data Collection
- ❖ Data Preparation
- ❖ Model Selection
- ❖ Model Training
- ❖ Model Evaluation
- ❖ Model Prediction

A. Data Collection

For your machine-learning model to be able to uncover the right trends, you must first gather data that can be trusted. The accuracy of the data that you feed the computer determines how accurate your model will be when you create it. Using data that is faulty or obsolete will result in inaccurate results or predictions which are irrelevant on a relational, hierarchical basis if the data is inaccurate or outdated.

B. Data Preparation

To determine the outcome, after the data has been collected, the next step is to combine all of the information and randomize it. By doing so, you ensure that the data will be dispersed evenly and that the order of the data will not affect the learning process in any way. There are many steps that need to be taken to clean the data for removing any unwanted information, such as missing values, rows, columns, duplicate values, and converting the data type. If you have to rearrange the dataset, then you may have to change the row and column arrangements or even the indexes of those rows and columns. Visualize the data to see how it's organized and the relationships between the various variables and classes that are there. The production of two sets of cleaned data: a training set and a testing set. The training set is the one from which your model learns. After training, a testing set is utilized to assess the accuracy of your model.

C. Model Selection

Choosing the model After performing a machine learning algorithm on the obtained data, a machine learning model selects the output. Choosing a model that is appropriate for the task at hand is crucial. Scientists and engineers have developed many models for diverse tasks such as speech recognition, image recognition, prediction, and so on. We must also determine whether your model is suitable for numerical or categorical data and make the appropriate choice.

1. Logistic Regression: A primary use for the supervised learning technique known as logistic regression is classification tasks, where the objective is to estimate the likelihood that a given instance belongs to a particular class.

2. Support Vector Machine (SVM): A supervised learning system called a support vector machine is employed for classification and regression issues. Many people highly favor the support vector machine because it generates correct results with a notable degree of efficiency. It is primarily applied to classification issues.

3. KNN Classifier (KNN): It is a supervised machine learning approach that is non-parametric and utilized for regression and dataset classification. To select the K value that minimizes error and enables the algorithm to accurately predict results, we must continually undo the code.

4. Decision Tree Classifier (DT): In data mining, it is the most often used supervised machine-learning technique. A decision tree is used to visually illustrate a statistical likelihood or a sequence of events, actions, or results.

D. Model Training

Training is the step of machine learning that is most important. To detect patterns and generate predictions, you give the prepared data to your machine learning model during training. As a result, the model gains knowledge from the data and succeeds in achieving the given goal. With practice, the model improves its ability to forecast.

E. Model Evaluation

You must evaluate your model's performance after it has been trained. This is accomplished by putting the model to the test on previously unseen data. The testing set that you split our data into before is the unseen data used. If you test on the same data that

was used for training, you will not get an accurate measure because the model is already familiar with the data and recognizes the same patterns. This will result in unreasonably high precision. We receive an accurate measure of how your model will perform and its speed when you utilize it on testing data.

Tuning: Examine whether your model's accuracy can be improved in any manner once you've constructed and tested it. By fine-tuning the model's parameters, you can achieve this. The variables in the model that the programmer chooses are known as parameters. The accuracy will be the highest at a specific value of your parameter. Parameter tweaking is the process of locating these settings.

F. Model Prediction

"Predictions" refer to the results produced by an algorithm once it has been trained on a historical dataset and applied to new data to forecast the likelihood of a certain outcome from new data through an algorithm. In the end, we can use our model to create accurate predictions in the absence of unknown data by applying it to unknown data.

Data Flow Diagram: A data flow diagram (DFD) depicts the information flow in this system. It displays data inputs, outputs, storage locations, and routes between each destination using predefined symbols such as rectangles, circles, and arrows as well as brief text labels. The entity, process, data store, and data flow are the four essential components of all data flow diagrams. In this instance, the DFD is split into two levels to make it simpler to understand the data flow.

1. Data Flow Diagram Level 0

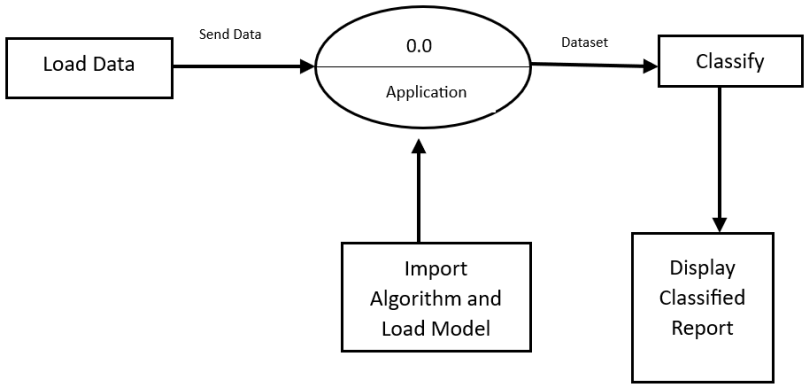

Fig. 2 DFD 0

The diagram up top is a representation of DFD0, which gives you an overview or context diagram of the entire system. It is intended to provide a quick view that presents the system as a single high-level process. Here, data from the file is loaded into the application, where it is sent to the classification unit, which uses the trained model file to predict the outcome, and the output is known.

2. Data Flow Diagram Level 1

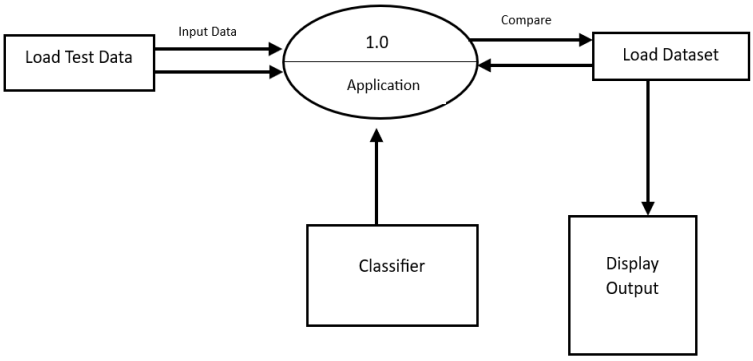

Fig. 3 DFD 1

The diagram above is a representation of DFD1. The Level 0 DFD is further subdivided into the Level 1 DFD. Level 1 DFD illustrates the system's basic modules and the flow of data between them. The data from the file is loaded into the system, and the loaded data is passed to the classification unit to forecast the outcome, and classes are classed and labeled.

IV. RESULTS

All the discussed models were implemented in Python using Scikit-learn to test the prediction if a person needs treatment or not. This section will evaluate and discuss the data about the machine learning methodologies that we have employed to conduct a prediction or diagnosis for mental health issues. The logistic regression, support vector machine, k nearest neighbor, and decision tree are the four machine learning methods that this study found. The performance of the machine learning algorithms will also be evaluated and assessed. The outcomes are shown graphically and tabulated as follows:

TABLE I
PERFORMANCE OF TRAINED MODELS

Method	Accuracy	Precision	Recall	F1-Score
Logistic Regression	83.17%	43.75%	13.72%	20.89%
SVM	84.12%	100.0%	1.96%	3.84%
KNN	85.07%	83.33%	9.80%	17.54%
Decision Tree	83.80%	50.0%	5.88%	10.52%

Table 1 contains the evaluation and tabulation of the performance of several trained models. Through the use of machine learning methods, the model was trained. KNN is the most accurate algorithm when compared to all the models.

REFERENCES

[1]Alghamdi, Norah Saleh. "Monitoring mental health using smart devices with text analytical tool." In 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT), pp. 2046-2051. IEEE, 2019.
[2]Atharva Burte, Jain Yash, Aditya Vora, and Hermish Gandhi. "Mental and physical health management system using ML, computer vision and IoT sensor network." In 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 786-791. IEEE, 2020.
[3]Kumar, Piyush, Rishi Chauhan, Thompson Stephan, Achyut Shankar, and Sanjeev Thakur. "A Machine Learning Implementation for Mental Health Care. Application: Smart Watch for Depression Detection." In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 568-574. IEEE, 2021.
[4]Subhani, Aamir Saeed Malik, Wajid Mumtaz, Mohamed Naufal Bin Mohamed Saad, Ahmad Rauf, and Nidal Kamel. "Machine learning framework for the detection of mental stress at multiple levels." IEEE Access 5 (2017): 13545-13556.
[5]Yuan Liao, Wang, Jiankun Wang, Ruyi, and Jinyu Wang. "Supervised machine learning chatbots for perinatal mental healthcare." In 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), pp. 378-383. IEEE, 2020.
[6]Jadhav, Vaishnavi N., Manasi S. Bhamare, Manasvi U. Vegurlekar, and Vidya S. Kubde. "MENTAL HEALTH TRACKER."
[7]Bagul, Apoorva, Pooja Sinkar, Priyanka Jadhav, Deepali Ahire, and D. D. Sharma. "A MENTAL HEALTH TRACKER BUILT USING FLUTTER AND FIREBASE."
[8]Tabisula, Braden, and Chinazunwa Uwaoma. "The Need for an Adaptive Sociotechnical Model for Managing Mental Health in a Pandemic." In 2022 IEEE International Conference on Digital Health (ICDH), pp. 66-68. IEEE, 2022.
[9]Tarik Qassem, Moore, Bin Hu, Philip, Andrew Thomas, and Nik Bessis. "Monitoring patients with mental disorders." In 2015 9th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, pp. 65-70. IEEE, 2015
[10]Chung, Jetli, and Jason Teo. "Mental Health Prediction Using Machine
Learning: Taxonomy, Applications, and Challenges." Applied Computational Intelligence and Soft Computing (2022)