**MINDSPA – CREATE A BETTER YOU**

A report submitted for the course of

**Application Development\_Machine Learning Explore**

**III B. Tech I Semester**

**by**

## 

**A.Ajithkumar-2011CS030004**

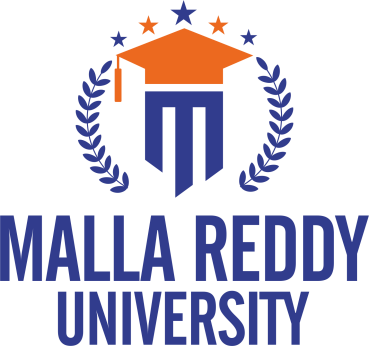
**A.Asmitha -2011CS030006**

**G.Mounika -2011CS030060**

Under the esteemed guidance of

**Dr.D.B.Jaggannadha Rao**

**(Associate Professor)**



**Department of Computer Science & Engineering (Data Science)**

**Malla Reddy Univeristy**

Maisammaguda, Dulapally,

Hyderabad, Telangana 500100  
www.mallareddyuniversity.ac.in

**2022-23**

## CERTIFICATE

This is to certify that this bonafide record of the application development entitled **“MINDSPA-create a better you”** submitted by Mr.Ajith kumar(2011cs030004), Ms.Asmitha(2011cs030006), Ms.G.Mounika(2011cs030060) of III year I sem to the Malla Reddy University, Hyderabad.This bonafide record of work carried out by us under the guidance of our supervision.The contents of this report, in full or in parts, have not been submitted to any other Organization for the award of any Degree.

|  |  |  |
| --- | --- | --- |
| **INTERNAL GUIDE** | **APP DEVELOPMENT**  **CONVENER** | **HEAD OF THE**  **DEPARTMENT** |
| **Dr.D.B.JaggannadhaRao** | **Dr. K Suresh Kumar Dr.GS NaveenKumar** | **Dr.GS NaveenKumar** |
| **Associate Professor** | **Associate Professor**  **CSE (Data Science)** | **CSE (Data Science)** |

**Date:**

**External Examiner**

## 

## 

## 

## DECLARATION

I certify that

1. The work contained in this report is original and has been done by me under the guidance of my supervisor(s).
2. The work has not been submitted to any other Institute for any degree or diploma.
3. I have followed the guidelines provided by the Institute in preparing the report.
4. I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.
5. Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the report and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.
6. **A.Ajithkumar-2011CS030004**
7. **A.Asmitha -2011CS030006**
8. **G.Mounika -2011CS030060**

## ACKNOWLEDGEMENT

We would like take this opportunity to acknowledge everyone who has helped us in every stage of this project.

We ineffably indebted to **Dr.D.B.JaggannadhaRao, Designation, CSE(Data Science)** for conscientious guidance and encouragement to accomplish this Application.

We extremely thankful to **Dr. K Suresh Kumar**, **Associate Professor, App Development Convener, Dept. of CSE(DS), Malla Reddy University** for his continuous support.

We wish to convey our sincere thanks to **Dr.GS Naveen Kumar**, **Head of Department, Dept. of CSE(DS), Malla Reddy University** for his continuous support.

We wish to convey our gratitude **to Prof. Dr. K. Harikrishna , Associate Dean-SoE, Malla Reddy University** for his continuous support.

We wish to convey our extended gratitude **to Prof. K Kailasa Rao, Dean-SoE, Malla Reddy University** for his continuous support.

We would like to express our special gratitude and thanks to **Dr. VSK Reddy, Vice Chancellor, Malla Reddy University** for his continuous support.

We are also grateful to all administrative staff and the technical staff of Computer Science and Engineering (Data Science) Department, people who provided us the useful and necessary information needed for this project.

## 

## ABSTRACT

Mental Health is a very important issue in today’s world . As in these days, a lot of people now working from home and even companies are preferring work from home, the mental health situation of the people has nose- dived to unbelievable levels. The main objective of the proposed project focuses on building an application that will try to collect the mental state of the user and provide close to accurate remedies – assigning tasks, connecting a doctor( or psychiatrist) etc. The application tries to find if an app user is suffering from any mental issue by asking a few psychological questions, and it suggests measures that the app users can believe to recover from their issue.The application will suggest tasks to them. The mental health of a person is predicted by using the machine learning algorithms. Apart from predicting the mental health of a person it also includes the basic information of mental health, and the disorders that a person faces due to the mental health that can be found in the causes, myths& facts and disorders in the website.

## 

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| Title Page |  | 1 |
| Certificate | | 2 |
| Declaration |  | 3 |
| Acknowledgement | | 4 |
| Abstract |  | 5 |
| Contents |  | 6 |
| List of Figures | | 7 |
| List of Tables | | 8 |
| Abbreviations | | 9 |
| Chapter 1 | Introduction | 10 |
|  |  |  |
|  |  |  |
|  |  |  |
| Chapter 2 | Review of Relevant Literature | 11 |
| Chapter 3 | Methodology | 14 |
| Chapter 4 | Results and Discussions | 33 |
| Chapter 5 | Conclusions and Future Scope of Study | 38 |
| References |  | 39 |

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| **Table** | **Title** | **Page** |
| 2.1 | Literature survey | 12 |
| 4.1 | Accuracies of ml algorithms | 34 |
| 4.2 | Classification report 1 | 37 |

4.3 Classification report 2 37

## LIST OF FIGURES

|  |  |  |
| --- | --- | --- |
| **Figure** | **Title** | **Page** |
| 1.1 | Fig Name1 | 3 |
| 1.2 | Fig Name2 | 4 |
| 1.3 | Fig Name3 | 8 |
|  |  |  |

## LIST OF ABBREVIATIONS

ANNs Artificial Neural Networks

BM Bayesian Models

DM Deep Learning

DR Dimensionality Reduction

DT Decision Trees

EL Ensembles Learning

IBM Instance Based Models

SVMs Support vector Machines

**CHAPTER – 1**

**INTRODUCTION**

Mental comfort of a human being is the current state of mind and it shows an overview of general behaviour and mental illness. To provide suggestion and mental treatment are difficult task and its needs to understand the behavior and mental nature of person. To understand the mental health related issues, many communities are formed and it acts as indicator for effectively treating the ailments of the individual and predict mental health related divergence. The section can be broadly summarized as teenage adolescents, college goers and working peoples. The common notion that for categories of the population are subject to stress and depression. It is a necessity to address the mental wellness of different categories at different times in order to prevent any serious illness. As per statistics of 2011 World Health Organization (WHO) predicted that by 2031, misery will be the leading cause of global disease and it affects to all ages.

The report also includes that the mental health profile of a person by healthcare providers will be made compulsory in the coming years, it provide better medication and also assist in faster recovery of a mental health diseases.

Few researcher are also discussed how the predictive analytics will revolutionize the healthcare field globally and provide better support to health sector. The behavior includes emotional, psychological, and social well-being and also affects how we expect, sense, and behave. It also determine how people can handle pressure, relate to others, and make options. As per analysis mental health prediction is important at every phase of life, from pre-teens and adolescence through adulthood. It is measured by a high grade of impairment, such as affective disorder that results in depression and different anxiety disorders. One in four people in the world will be surely a victim of mental or brain disorders at some point in their lives. In developing and developed countries 25 % people suffer mental health disorder. World Health Organization (WHO) report also revealed that 7.5 percent of the Indian population suffers from some form of mental disorder. The system can be widely used in across sectors: Education, Family and children, services, child psychology, pediatrics, counselling and social work.

In education sector, if mental health services are embedded within school systems it can improve both mental health and educational attainment for children. In child psychology sector, our project will be useful by knowing what kind of mental illness problem is affecting children’s the most. In this research the person fills questionnaire consist of questions related to day to day activities of person like how much time he/she is using social media in a day, how much hours they are sleeping etc.

To resolve the mental well-being machine learning technique play important role. It holds great promise to transform mental health care. Its tools also hold the potential to extend the current capabilities of clinicians, to deal with complex problems and ever-expanding information streams that stretch the limits of human ability. In the world of advancement of technologies we are trying to create a system which will predict the mental health status of a person. The outcome of the system is to give prediction based on data (questionnaire) or digital data.

Our proposed different levels of questionnaire and based on the results of that provide free checking of a person’s mental state and help him by suggestions, tasks and the doctor recommendations if required.

**CHAPTER - 2**

**REVIEW OF RELEVENT LITERATURE**

The application of Artificial Intelligence (AI) and big-data technologies helps diagnose mental health problem and it has great potential for personalizing treatment selection, prognosticating, monitoring for relapse used for improving mental healthcare through AI and big-data. Also discussed the various data mining algorithms and techniques to explore Mental Health and presents survey of the existing research works in this domain, referring to the techniques and algorithms. They also discussed how machine learning techniques helps in mental health and it provide the analysis and prediction of machine learning (ML) on different real time applications, discussed the Mental Health Problems among Children Using Machine Learning algorithm and has identified various machine learning techniques and has compared their performances on with different measures of accuracy on five basic mental health problems, discussed how to detect and treat mental Illness on Social Networks datasets and reviews. Also discussed the various machine learning techniques to detect depression. Author used the large social media dataset for identification of common traits among depressed people and identify them using various machine learning algorithms.

**Life satisfaction and mental health of Chinese older adults in different living arrangements**, In order to compare life satisfaction and mental health status of older adults in different living arrangements, we investigated a sample of 1, 915 Chinese older adults using Satisfaction with Life Scale and Mental Health Inventory for the Urban Elderly. Difference in life satisfaction between three living arrangements is significant even after controlling age, education and income. Life satisfaction of older adults living in elderly apartments is better than those living at home. Income and co-residence with children interact to influence life satisfaction. Among older adults with high income, older adults living alone or with spouse have the greater life satisfaction than those living with children; while among older adults with average and low income, living with or separate form children have no difference on life satisfaction. The result indicates that institutionalized older adults have equal mental health with those living at home and even greater life satisfaction

Data Science in Public Mental Health: A New Analytic Framework. Understanding public mental health issues using data science and finding solutions based on the findings from the data science projects can be complex and requires advanced techniques, compared to conventional data analysis projects. It is important to have a comprehensive project management process to ensure that project associates are competent and have enough knowledge to implement the data science process. Therefore, this paper presents a new framework that mental health professionals can use to solve challenges they face using data science. Although a large number of research papers have been published on public mental health, few have addressed the use of data science in public mental health. Recently, Data Science has changed the way we manage, analyze and leverage L.G.N.S.C.O.E, Department of Computer Engineering 2021-2022 6 data in healthcare industry. Data science projects differ from conventional data analysis, primarily because of the scientific approach used during data science projects. One of the motives for introducing a new framework is to motivate healthcare professionals to use ”Data Science” to address the challenges of mental health. Having a good data analysis framework and clear guidelines for a comprehensive analysis is always a plus point. It also helps to predict the time and resources needed in the early in the process to get a clear idea of the problem to be solved.

Jan Bohacik , Ivan Skula , Michal Zabovsky , ”Benefiting from online mental status examination system and mental health diagnostic system ”, Computer Science and Information Systems (FedCSIS) 2020 15th Conference on, pp. 27-30, 2020. In this really hectic world, quite a number of people are exposed to situations where mental stress Is unavoidable. This leads to people having all kinds of mental health problems that eventually may turn to chronic mental disorders. People with mental health problems normally have the tendency of not admitting their health problems because of the stigma attached to these kinds of illnesses. Most of them are in denial state, and this situation may cause very serious social problems since people with mental problems will develop some kind of mental disorders, and as a result, they might be harmful to others around them. People with mental health problems must receive proper treatments and medications. If their mental status can be assessed and examined easily, then most probably their mental problems can be detected at a very early stage, and can be easily controlled and cured. The above scenarios become the motivation for conducting this research. This research paper presents some findings on mental health and disorders on past research study’s results and also proposes an online mental status examination (MSE) system that examines individuals’ mental health status. The result of the MSE system is used in determining whether the respective person needs to undergo a more detailed diagnosis for more specific mental disorders. It is hoped that the outcomes of this research study are able to assist new psychotherapists and psychiatrists in examining and diagnosing those who are affected by some kind of mental disorders in a more efficient manner.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sr.**  **No.** | **Ref No.** | **Years** | **Study Proposal** | **Techniques and**  **Algorithm** | **Advantages and open**  **issues** |
| 1 | [1] | 2018 | A Systematic Review of data mining technique in health sector | Data Mining and algorithms ( KNN, DT ) | Prediction model discover the significant predictors that lead to the disease and discuss review analysis, future direction. **Gaps:** Data Mining technique discussed,  scalability of system is big challenge. |
| 2 | [3] | 2019 | Machine Learning Techniques are used in depression analysis | Machine learning techniques (SVM ) | 1. Improve the performance of the data modeling, 2. Maximize modeling performance,   **Gaps:** Suitable for digital data |
| 3 | [9] | 2017 | With the help of social network detect and Treat Mental states. | Machine learning algorithm ( Predictive model ) | 1. Model detect user’s poor mental health and provide them interventions on social network platforms. 2. Contribute to the early detection and treatment.   **Gaps:** Applicable only for social data |
| 4 | [5] | 2017 | Predicting Depression Levels Using Social Media Posts | Artificial Intelligence, Support  Vector Machine (SVM), and Naïve Bayes. | 1. Screening tool classify the user using artificial intelligence (UGC on SNS) 2.The SNS overcome the  problems regarding self-reporting.  **Gaps:** Difficult to handle MRI, ECG data. |
| 5 | [2] | 2016 | Machine Learning Techniques are helpful to predict the mental health problem of children’s | Machine learning Techniques, Feature Selection algorithms | 1. The tool provides various measures to understand the categorization. 2. Suitable for large data set, high accuracy.   **Gaps:** Trained prior to the implementation of any technique Applicable for small data, |
| 6 | [4] | 2019 | Big Data Analytics and AI in Mental Healthcare | Artificial Intelligence (AI) and big-data technologies | 1. The AI technologies offers potential to gain greater insight on disease progression. 2. System scalability maintained by big data and work on all types of data.   **Gaps:** Difficult to handle ethical issues. |
| 7 | [7] | 2019 | Diagnosis process can be interp of a new and unknown case for available set of medical data. | Machine Learning Techniques | 1. The data patterns provide useful information to prevent mental illness and assist in delivery of efficient mental health services.  **Gaps**: Works for few dataset. |
| 8 | [8] | 2018 | Web based app responsible for | Data Mining technique, KNN, random forest, Naïve base and decision tree. | 1. Offline web application, Chatbot assistant and Mobile access (web app) **Gaps**: 1. Inefficiency in dealing with patients who cannot type. 2. Notification System and Voice command oriented app and mobile device not supported. |

**Table:2.1 Literature survey**

**CHAPTER - 3**

**METHODOLOGY**

**MACHINE LEARNING**

Machine learning is a growing technology which enables computers to learn automatically from past data. Machine learning uses various algorithms for **building mathematical models and making predictions using historical data or information**. Currently, it is being used for various tasks such as **image recognition, speech recognition, email filtering, Facebook auto-tagging, recommender system,** and many more.

But can a machine also learn from experiences or past data like a human does? So here comes the role of **Machine Learning**

With the help of sample historical data, which is known as **training data**, machine learning algorithms build a **mathematical model** that helps in making predictions or decisions without being explicitly programmed. Machine learning brings computer science and statistics together for creating predictive models. Machine learning constructs or uses the algorithms that learn from historical data. The more we will provide the information, the higher will be the performance.



**Fig : 3.1 ML**

**3.1 PROBLEM DEFINITION**

Mental Health is a very important issue in today’s world . The mental health situation of the people has nose- dived to unbelievable levels. The main objective of the proposed project focuses on building an application that will try to predict the mental state of the user .The mental health of a person is predicted by using the machine learning algorithms. . In the world of advancement of technologies we are trying to create a system which will predict the mental health status of a person. The outcome of the system is to give prediction based on data (questionnaire) or digital data providing free checking of a person’s mental state and help him by suggestions, tasks and the doctor recommendations if required.

**3.2 OBJECTIVE**

1. To ensure availability and accessibility of minimum mental health care for all in the foreseeable future, particularly to the most vulnerable and underprivileged sections of the population.
2. To encourage the application of mental health knowledge in general health care and in social development.
3. To promote community participation in the mental health services development and to stimulate efforts towards self- help in the community.
4. To know the major cause of mental illness through mental health analysis.

**3.3 SYSTEM DESIGN**

**ML**

**ALGORITHMS**

**DATA**

**PREPROCESSING**

**DATASET**

**COLLECTION**

**USER**

**INTERFACE**

**DEPLOY USING**

**FLASK**

**QUESTIONNAIRE**

**RECOMMENDATIONS**

**PREDICT**

**Fig:3.2 System design**

**DATASET COLLECTION**

This dataset is from a 2014 survey that measures attitudes towards mental health and frequency of mental health disorders in the tech workplace.

**Content**

This dataset contains the following data:

* **Timestamp**
* **Age**
* **Gender**
* **Country**
* **state**: If you live in the United States, which state or territory do you live in?
* **self\_employed**: Are you self-employed?
* **family\_history**: Do you have a family history of mental illness?
* **treatment**: Have you sought treatment for a mental health condition?
* **work\_interfere**: If you have a mental health condition, do you feel that it interferes with your work?
* **no\_employees**: How many employees does your company or organization have?
* **remote\_work**: Do you work remotely (outside of an office) at least 50% of the time?
* **tech\_company**: Is your employer primarily a tech company/organization?
* **benefits**: Does your employer provide mental health benefits?
* **care\_options**: Do you know the options for mental health care your employer provides?
* **wellness\_program**: Has your employer ever discussed mental health as part of an employee wellness program?
* **seek\_help**: Does your employer provide resources to learn more about mental health issues and how to seek help?
* **anonymity**: Is your anonymity protected if you choose to take advantage of mental health or substance abuse treatment resources?
* **leave**: How easy is it for you to take medical leave for a mental health condition?
* **mental**health**consequence**: Do you think that discussing a mental health issue with your employer would have negative consequences?
* **phys**health**consequence**: Do you think that discussing a physical health issue with your employer would have negative consequences?
* **coworkers**: Would you be willing to discuss a mental health issue with your coworkers?
* **supervisor**: Would you be willing to discuss a mental health issue with your direct supervisor(s)?
* **mental**health**interview**: Would you bring up a mental health issue with a potential employer in an interview?
* **phys**health**interview**: Would you bring up a physical health issue with a potential employer in an interview?
* **mental**vs**physical**: Do you feel that your employer takes mental health as seriously as physical health?
* **obs\_consequence**: Have you heard of or observed negative consequences for coworkers with mental health conditions in your workplace?
* **comments**: Any additional notes or comments

**DATA PREPROCESSING**

Data collected from a 2014 survey that measures attitudes towards mental health and frequency of mental health disorders in the tech workplace, they are not in standard form and has many irrelevant and noisy data. With the help of data pre-processing eliminate irrelevant and redundant attributes using Best First Search technique.

In our dataset in gender column we have noisy data or mistyped data.By using replace function we have converted all those noisy data into standard values.

From the dataset the independent variable will be treatment and remaining all the columns taken as independent variables.

To convert categorical values to numeric values so that machine learning models can understand the data and find hidden patterns we used labelencoder and onehotencoder.

**MACHINE LEARNING ALGORITHMS**

ML involves the use of advanced statistical and probabilistic techniques to construct systems with an ability to automatically learn from data. Firstly, the dataset is trained where we considered the target variable as “treatment” column as we need to predict the mental health of a person. The following algorithms will be helpful in predicting the mental health of a person

1.Logistic Regression

2.K-Nearest Neighbor

3.Decision Tree

4.Random Forest

5.AdaBoost Classifier

6.Gradient Boosting Classifier

**LOGISTIC REGRESSION**

Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables.Logistic regression predicts the output of a categorical dependent variable. Therefore the outcome must be a categorical or discrete value. It can be either Yes or No, 0 or 1, true or False, etc. but instead of giving the exact value as 0 and 1, it gives the probabilistic values which lie between 0 and 1.

Logistic Regression is much similar to the Linear Regression except that how they are used. Linear Regression is used for solving Regression problems, whereas **Logistic regression is used for solving the classification problems**.

In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1).

The curve from the logistic function indicates the likelihood of something such as whether the cells are cancerous or not, a mouse is obese or not based on its weight, etc.

Logistic Regression is a significant machine learning algorithm because it has the ability to provide probabilities and classify new data using continuous and discrete datasets.

Logistic Regression can be used to classify the observations using different types of data and can easily determine the most effective variables used for the classification. The below image is showing the logistic function

## Logistic Function (Sigmoid Function):

The sigmoid function is a mathematical function used to map the predicted values to probabilities.

It maps any real value into another value within a range of 0 and 1.

The value of the logistic regression must be between 0 and 1, which cannot go beyond this limit, so it forms a curve like the "S" form. The S-form curve is called the Sigmoid function or the logistic function.

In logistic regression, we use the concept of the threshold value, which defines the probability of either 0 or 1. Such as values above the threshold value tends to 1, and a value below the threshold values tends to 0.



**Fig:3.3 Sigmoid Function**

## Assumptions for Logistic Regression:

The dependent variable must be categorical in nature.

The independent variable should not have multi-collinearity.

## Logistic Regression Equation:

The Logistic regression equation can be obtained from the Linear Regression equation. The mathematical steps to get Logistic Regression equations are given below:

We know the equation of the straight line can be written as:

Logistic Regression in Machine Learning

In Logistic Regression y can be between 0 and 1 only, so for this let's divide the above equation by (1-y):

Logistic Regression in Machine Learning

But we need range between -[infinity] to +[infinity], then take logarithm of the equation it will become:

Logistic Regression in Machine Learning

The above equation is the final equation for Logistic Regression.

## Type of Logistic Regression:

On the basis of the categories, Logistic Regression can be classified into three types:

**Binomial:** In binomial Logistic regression, there can be only two possible types of the dependent variables, such as 0 or 1, Pass or Fail, etc.

**Multinomial:** In multinomial Logistic regression, there can be 3 or more possible unordered types of the dependent variable, such as "cat", "dogs", or "sheep"

**Ordinal:** In ordinal Logistic regression, there can be 3 or more possible ordered types of dependent variables, such as "low", "Medium", or "High".

**K- NEAREST NEIGHBOR**

K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique.

K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.

K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm.

K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems.

K-NN is a **non-parametric algorithm**, which means it does not make any assumption on underlying data.

It is also called a **lazy learner algorithm** because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset.

KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

**Example:** Suppose, we have an image of a creature that looks similar to cat and dog, but we want to know either it is a cat or dog. So for this identification, we can use the KNN algorithm, as it works on a similarity measure. Our KNN model will find the similar features of the new data set to the cats and dogs images and based on the most similar features it will put it in either cat or dog category.



**Fig:3.4 KNN**

Why do we need a K-NN Algorithm?

Suppose there are two categories, i.e., Category A and Category B, and we have a new data point x1, so this data point will lie in which of these categories. To solve this type of problem, we need a K-NN algorithm. With the help of K-NN, we can easily identify the category or class of a particular dataset. Consider the below diagram:



**Fig:3.5 Before KNN Fig:3.6 After KNN**

## How does K-NN work?

The K-NN working can be explained on the basis of the below algorithm:

**Step-1:** Select the number K of the neighbors

**Step-2:** Calculate the Euclidean distance of **K number of neighbors**

**Step-3:** Take the K nearest neighbors as per the calculated Euclidean distance.

**Step-4:** Among these k neighbors, count the number of the data points in each category.

**Step-5:** Assign the new data points to that category for which the number of the neighbor is maximum.

**Step-6:** Our model is ready.

Suppose we have a new data point and we need to put it in the required category. Consider the below image:



**Fig:3.7 Understanding KNN**

Firstly, we will choose the number of neighbors, so we will choose the k=5.

Next, we will calculate the **Euclidean distance** between the data points. The Euclidean distance is the distance between two points, which we have already studied in geometry. It can be calculated as:



**Fig:3.8 Euclidean Distance**

By calculating the Euclidean distance we got the nearest neighbors, as three nearest neighbors in category A and two nearest neighbors in category B. Consider the below image:



As we can see the 3 nearest neighbors are from category A, hence this new data point must belong to category A.

## How to select the value of K in the K-NN Algorithm?

Below are some points to remember while selecting the value of K in the K-NN algorithm:

There is no particular way to determine the best value for "K", so we need to try some values to find the best out of them. The most preferred value for K is 5.

A very low value for K such as K=1 or K=2, can be noisy and lead to the effects of outliers in the model.

Large values for K are good, but it may find some difficulties.

**DECISION TREE**

Decision Tree is a **Supervised learning technique**that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where**internal nodes represent the features of a dataset, branches represent the decision rules** and**each leaf node represents the outcome.** In a Decision tree, there are two nodes, which are the **Decision Node** and**Leaf Node.** Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches.

The decisions or the test are performed on the basis of features of the given dataset.

It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions**.**

It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure.

In order to build a tree, we use the **CART algorithm,** which stands for **Classification and Regression Tree algorithm.**

A decision tree simply asks a question, and based on the answer (Yes/No), it further split the tree into subtrees.



**Fig:3.9 Decision Tree**

Decision Trees usually mimic human thinking ability while making a decision, so it is easy to understand.

The logic behind the decision tree can be easily understood because it shows a tree-like structure.

**Decision Tree Terminologies**

1. **Root Node:** Root node is from where the decision tree starts. It represents the entire dataset, which further gets divided into two or more homogeneous sets.

2. **Leaf Node:** Leaf nodes are the final output node, and the tree cannot be segregated further after getting a leaf node.

3. **Splitting:** Splitting is the process of dividing the decision node/root node into sub-nodes according to the given conditions.

4. **Branch/Sub Tree:** A tree formed by splitting the tree.

5. **Pruning:** Pruning is the process of removing the unwanted branches from the tree.

6. **Parent/Child node:** The root node of the tree is called the parent node, and other nodes are called the child nodes.

**How does the Decision Tree algorithm Work?**

**Step-1:** Begin the tree with the root node, says S, which contains the complete dataset.

**Step-2:** Find the best attribute in the dataset using **Attribute Selection Measure (ASM).**

**Step-3:** Divide the S into subsets that contains possible values for the best attributes.

**Step-4:** Generate the decision tree node, which contains the best attribute.

**Step-5:** Recursively make new decision trees using the subsets of the dataset created in step -3. Continue this process until a stage is reached where you cannot further classify the nodes and called the final node as a leaf node.

**RANDOM FOREST**

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning,** which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.

As the name suggests, ***"*Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.*"*** Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

**The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.**

The below diagram explains the working of the Random Forest algorithm:



**Fig:3.9 Random Forest**

## Assumptions for Random Forest

Since the random forest combines multiple trees to predict the class of the dataset, it is possible that some decision trees may predict the correct output, while others may not. But together, all the trees predict the correct output. Therefore, below are two assumptions for a better Random forest classifier:

There should be some actual values in the feature variable of the dataset so that the classifier can predict accurate results rather than a guessed result.

The predictions from each tree must have very low correlations.

## Why use Random Forest?

Below are some points that explain why we should use the Random Forest algorithm:

<="" li="">

It takes less training time as compared to other algorithms.

It predicts output with high accuracy, even for the large dataset it runs efficiently.

It can also maintain accuracy when a large proportion of data is missing.

## How does Random Forest algorithm work?

Random Forest works in two-phase first is to create the random forest by combining N decision tree, and second is to make predictions for each tree created in the first phase.

The Working process can be explained in the below steps and diagram:

**Step-1:** Select random K data points from the training set.

**Step-2:** Build the decision trees associated with the selected data points (Subsets).

**Step-3:** Choose the number N for decision trees that you want to build.

**Step-4:** Repeat Step 1 & 2.

**Step-5:** For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority votes.

The working of the algorithm can be better understood by the below example:

**Example:** Suppose there is a dataset that contains multiple fruit images. So, this dataset is given to the Random forest classifier. The dataset is divided into subsets and given to each decision tree. During the training phase, each decision tree produces a prediction result, and when a new data point occurs, then based on the majority of results, the Random Forest classifier predicts the final decision.

**ADABOOST CLASSIFIER**

AdaBoost was the first really successful boosting algorithm developed for the purpose of binary classification. AdaBoost is short for Adaptive Boosting and is a very popular boosting technique that combines multiple “weak classifiers” into a single “strong classifier”. It was formulated by Yoav Freund and Robert Schapire. They also won the 2003 Gödel Prize for their work.

**Algorithm:** 

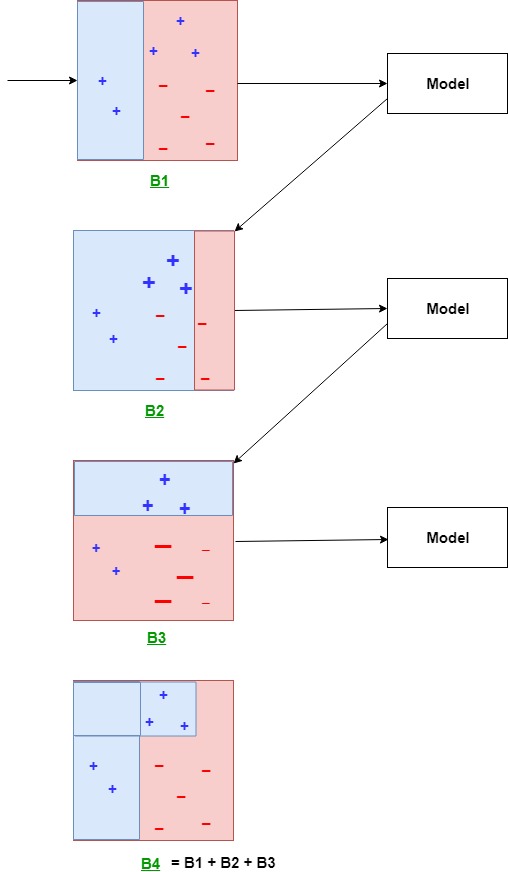
Initialise the dataset and assign equal weight to each of the data point.

Provide this as input to the model and identify the wrongly classified data points.

Increase the weight of the wrongly classified data points.

if (got required results)   
  Goto step 5   
else   
  Goto step 2 

End



**Fig:3.10 Adaboost classifier**

**Explanation:**   
The above diagram explains the AdaBoost algorithm in a very simple way. Let’s try to understand it in a stepwise process:

**B1** consists of 10 data points which consist of two types namely plus(+) and minus(-) and 5 of which are plus(+) and the other 5 are minus(-) and each one has been assigned equal weight initially. The first model tries to classify the data points and generates a vertical separator line but it wrongly classifies 3 plus(+) as minus(-).

**B2** consists of the 10 data points from the previous model in which the 3 wrongly classified plus(+) are weighted more so that the current model tries more to classify these pluses(+) correctly. This model generates a vertical separator line that correctly classifies the previously wrongly classified pluses(+) but in this attempt, it wrongly classifies three minuses(-).

**B3** consists of the 10 data points from the previous model in which the 3 wrongly classified minus(-) are weighted more so that the current model tries more to classify these minuses(-) correctly. This model generates a horizontal separator line that correctly classifies the previously wrongly classified minuses(-).

**B4** combines together B1, B2, and B3 in order to build a strong prediction model which is much better than any individual model used.

**GRADIENT BOOSTING**

Machine learning is one of the most popular technologies to build predictive models for various complex regression and classification tasks. **Gradient Boosting Machine** (GBM) is considered one of the most powerful boosting algorithms.

Although, there are so many algorithms used in machine learning, boosting algorithms has become mainstream in the machine learning community across the world. Boosting technique follows the concept of ensemble learning, and hence it combines multiple simple models (weak learners or base estimators) to generate the final output. GBM is also used as an ensemble method in machine learning which converts the weak learners into strong learners. In this topic, **"GBM in Machine Learning"** we will discuss gradient machine learning algorithms, various boosting algorithms in machine learning, the history of GBM, how it works, various terminologies used in GBM, etc. But before starting, first, understand the boosting concept and various boosting algorithms in machine learning.

## What is Boosting in Machine Learning?

Boosting is one of the popular learning ensemble modeling techniques used to build strong classifiers from various weak classifiers. It starts with building a primary model from available training data sets then it identifies the errors present in the base model. After identifying the error, a secondary model is built, and further, a third model is introduced in this process. In this way, this process of introducing more models is continued until we get a complete training data set by which model predicts correctly.

AdaBoost (Adaptive boosting) was the first boosting algorithm to combine various weak classifiers into a single strong classifier in the history of machine learning. It primarily focuses to solve classification tasks such as binary classification.

## Steps in Boosting Algorithms:

There are a few important steps in boosting the algorithm as follows:

Consider a dataset having different data points and initialize it.

Now, give equal weight to each of the data points.

Assume this weight as an input for the model.

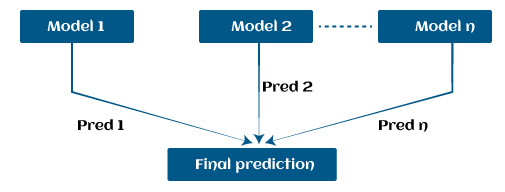
Identify the data points that are incorrectly classified.

Increase the weight for data points in step 4.

If you get appropriate output then terminate this process else follow steps 2 and 3 again.

### **Example:**

Let's suppose, we have three different models with their predictions and they work in completely different ways. For example, the linear regression model shows a linear relationship in data while the decision tree model attempts to capture the non-linearity in the data as shown below image.



**Fig:3.11 Gradient boosting**

Further, instead of using these models separately to predict the outcome if we use them in form of series or combination, then we get a resulting model with correct information than all base models. In other words, instead of using each model's individual prediction, if we use average prediction from these models then we would be able to capture more information from the data. It is referred to as ensemble learning and boosting is also based on ensemble methods in machine learning.

## Boosting Algorithms in Machine Learning

There are primarily 4 boosting algorithms in machine learning. These are as follows:

**Gradient Boosting Machine (GBM)**

**Extreme Gradient Boosting Machine (XGBM)**

**Light GBM**

**CatBoost**

**ACCURACIES**

These are the algorithms used after training the datasets. Each algorithm has its own accuracy values that are as follows:

1.Accuracy of Logistic Regression : 0.85333333333334

2. Accuracy of K-Nearest Neighbor : 0.72

3. Accuracy of Decision Tree : 0.76

4. Accuracy of Random Forest : 0.834666666667

5. Accuracy of AdaBoost Classifier : 0.834666666667

6. Accuracy of Gradient Boosting Classifier : 0.8373333333334

When we observe the accuracy of AdaBoost Classifier it is 0.834666666667 but when we applied a technique called “**RandomizedSearchCV**” the accuracy increased to 0.8613333333333

**RandomizedSearchCV:** RandomizedSearchCV implements a “fit” and a “score” method. It also implements “predict”, “predict\_proba”, “decision\_function”, “transform” and “inverse\_transform” if they are implemented in the estimator used. The parameters of the estimator used to apply these methods are optimized by cross-validated search over parameter settings. Randomized Search is a yet another technique for sampling different hyper parameters combination in order to find the optimal set of parameters which will give the model with most optimal performance / score. As like Grid search, randomized search is the most widely used strategies for hyper-parameter optimization.

**DEPLOY USING FLASK**

Flask is a lightweight [WSGI](https://wsgi.readthedocs.io/) web application framework. It is designed to make getting started quick and easy, with the ability to scale up to complex applications. It began as a simple wrapper around [Werkzeug](https://palletsprojects.com/p/werkzeug) and [Jinja](https://palletsprojects.com/p/jinja) and has become one of the most popular Python web application frameworks.

It is a micro web framework written in Python that speeds up application development by providing essential back end components for programmers to build upon. Flask is simple and lightweight—one of the most manageable frameworks around—and contains only the vital necessities for web development. It is, however, also designed to be highly extensible so developers can customize it however they see fit.

It’s also a good idea to have acquired some basic knowledge of programming and web development fundamentals including HTML and CSS.

**HTML** stands for**HyperText Markup Language**. It is used to design web pages using a markup language. HTML is the combination of Hypertext and Markup language. Hypertext defines the link between the web pages. A markup language is used to define the text document within tag which defines the structure of web pages.

HTML is a markup language that is used by the browser to manipulate text, images, and other content to display it in the required format.

**CSS (Cascading Style Sheets)** is a stylesheet language used to design the webpage to make it attractive. The reason for using CSS is to simplify the process of making web pages presentable. CSS allows you to apply styles to web pages. More importantly, CSS enables you to do this independent of the HTML that makes up each web page.

There are three types of CSS which are given below:

**Inline CSS**: In Inline CSS, we add the style to the tags using the “style” attribute inside the tag which we want to design

**Internal or Embedded CSS:** Internal CSS allows us to style our page by adding the <style> tag inside the <head> tag. Inside the <style> tag, we add the design that we want to give to our page.

**External CSS** : External CSS lets us add style to our HTML page externally. We can add our styles in a different file with extension .css and link this page to our HTML page.

In our application, we used templates to render HTML which will display in the user’s browser. This folder contains our HTML pages such as home.html, index.html,myths.html,disorders.html and output.html.

**Flask script:** we need to download flask and some other libraries. Here, we make use of a virtual environment, where all the libraries are managed which makes both the development and deployment job easier.

Installing Flask: pip install flask

**Flask script:**Before starting with the coding part, we need to download flask and some other libraries. Here, we make use of a virtual environment, where all the libraries are managed which makes both the development and deployment job easier.

Installing Flask: pip install flask

Here we imported the libraries, then using app=Flask(\_\_name\_\_) we create an instance of flask. @app.route(‘/’) is used to tell flask what URL should trigger the function home() and in the function index, we use render\_template(‘home.html’) to display the script home.html in the browser.Like this we have created @app.route(‘/pred’), @app.route(‘/output’), @app.route(‘/myths’). @app.route(‘/disorders’), @app.route(‘/remedies’) to tell flask what URL should trigger.

We used render\_templates index.html,output.html,myths.html,disorders.html and remedies.html.

**Predicting the mental state of the user:**When someone submits the form, the webpage should display the predicted value of mental state. For this, we require the model file (model.pkl) we created before in the same project folder.  
Here, after the form is submitted, the form values are stored in the variable pred in the form of a dictionary. We convert it into a list of the dictionary’s values and pass it as an argument to predict() function. In this function, we load the model.pkl file and predict the new values and return the result.   
This result/prediction describes whether the person requires mental health treatment or not.If the predicted value is 1 then the output will “The person requires mental health treatment” ad if the value is 0 then the output will be “The person does not require mental health treatment”.

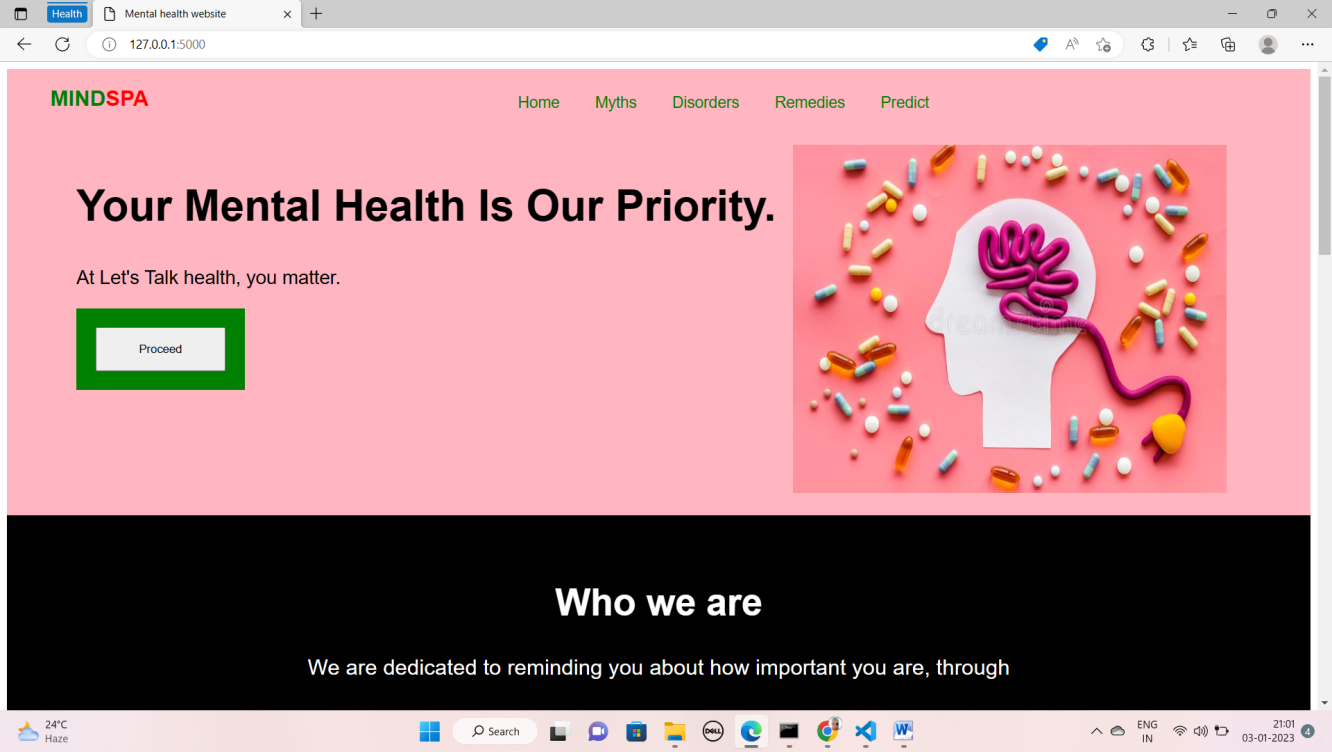
To run application:python app.py

This should run the application and launch a simple server. Open http://127.0.0.1:5000/ to see the web page

**USER INTERFACE**

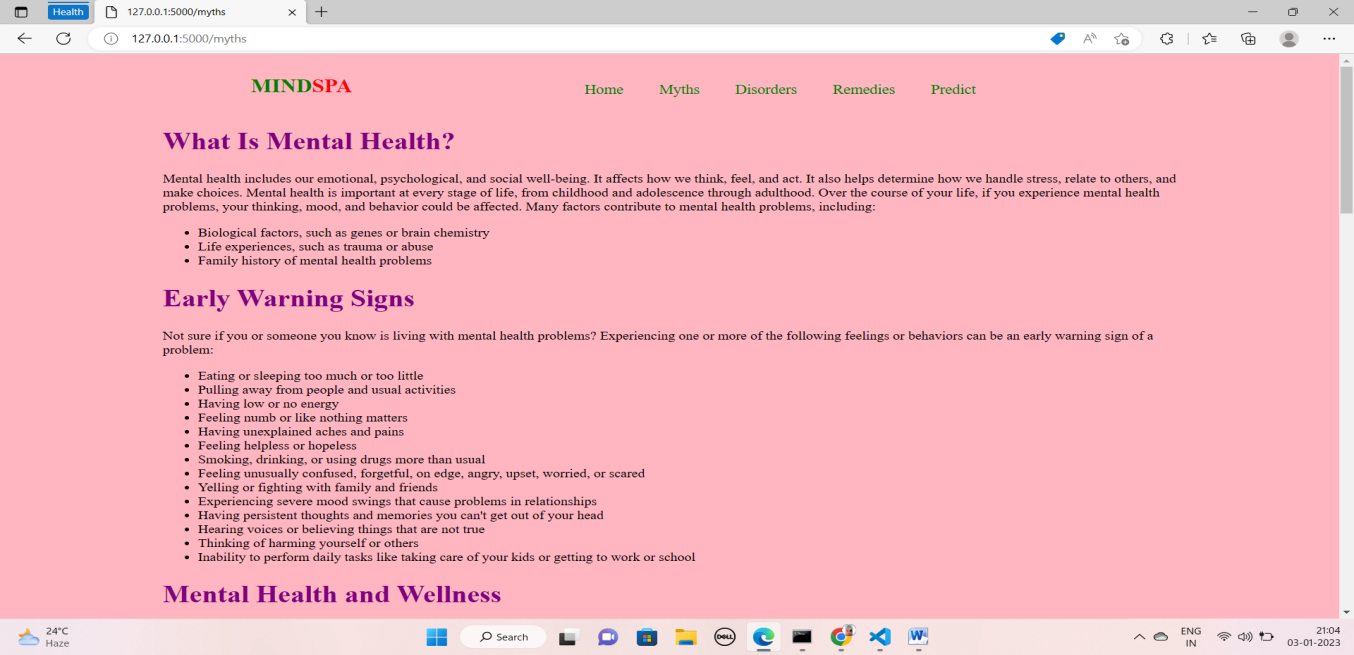
When we see the web page it loads a web page which consists of navigation bar containing home , myths, disorders, remedies and predict tabs.When we click on each tab it will redirect to particular web page.

**Home.html:**



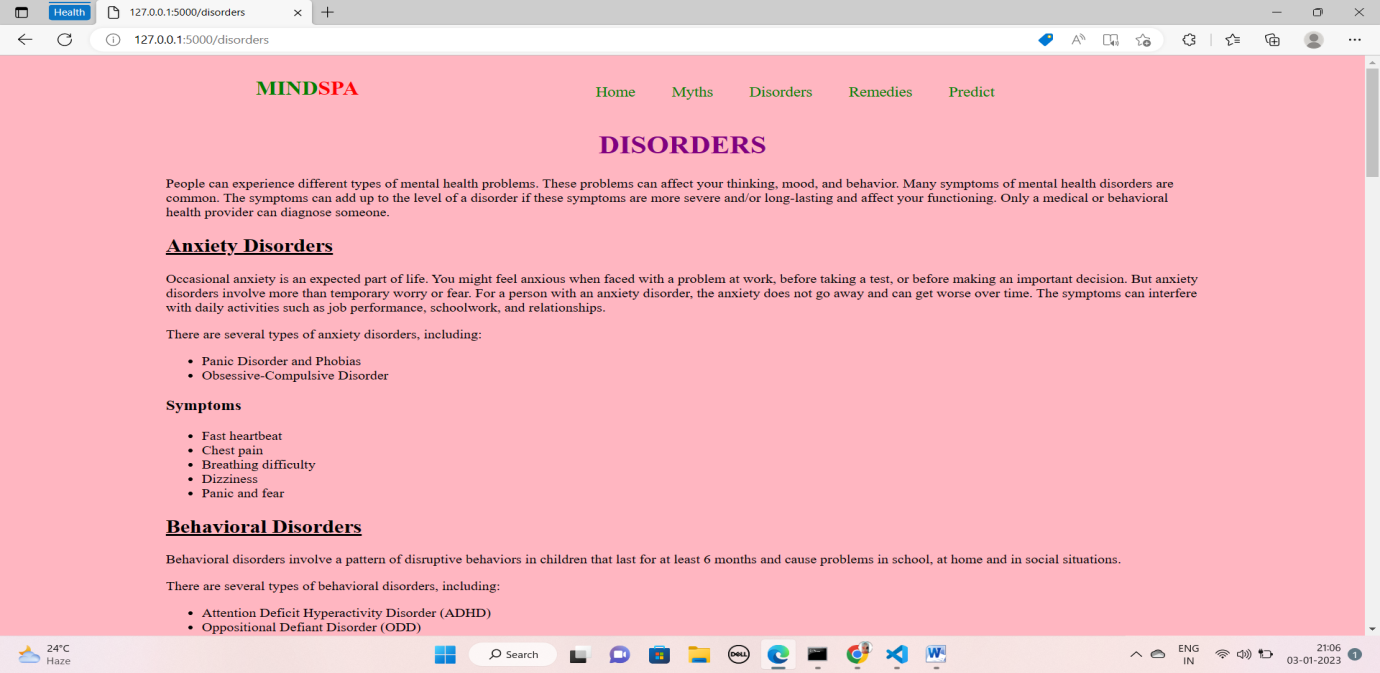
**Fig:3.12 home.html**

**Myths.html:**

****

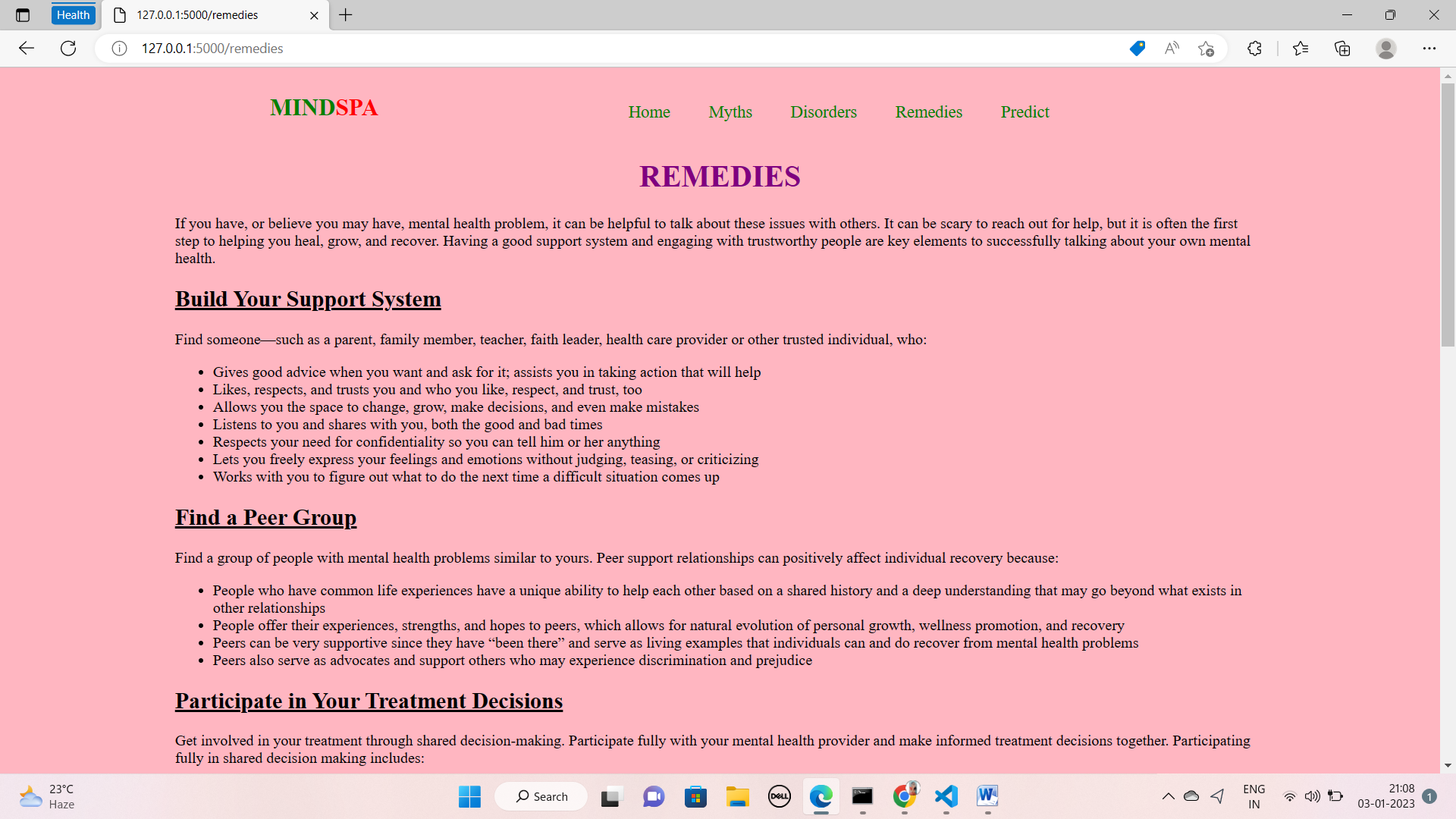
**Fig:3.13 myths.html**

**Disorders.html:**

****

**Fig:3.14 disorders.html**

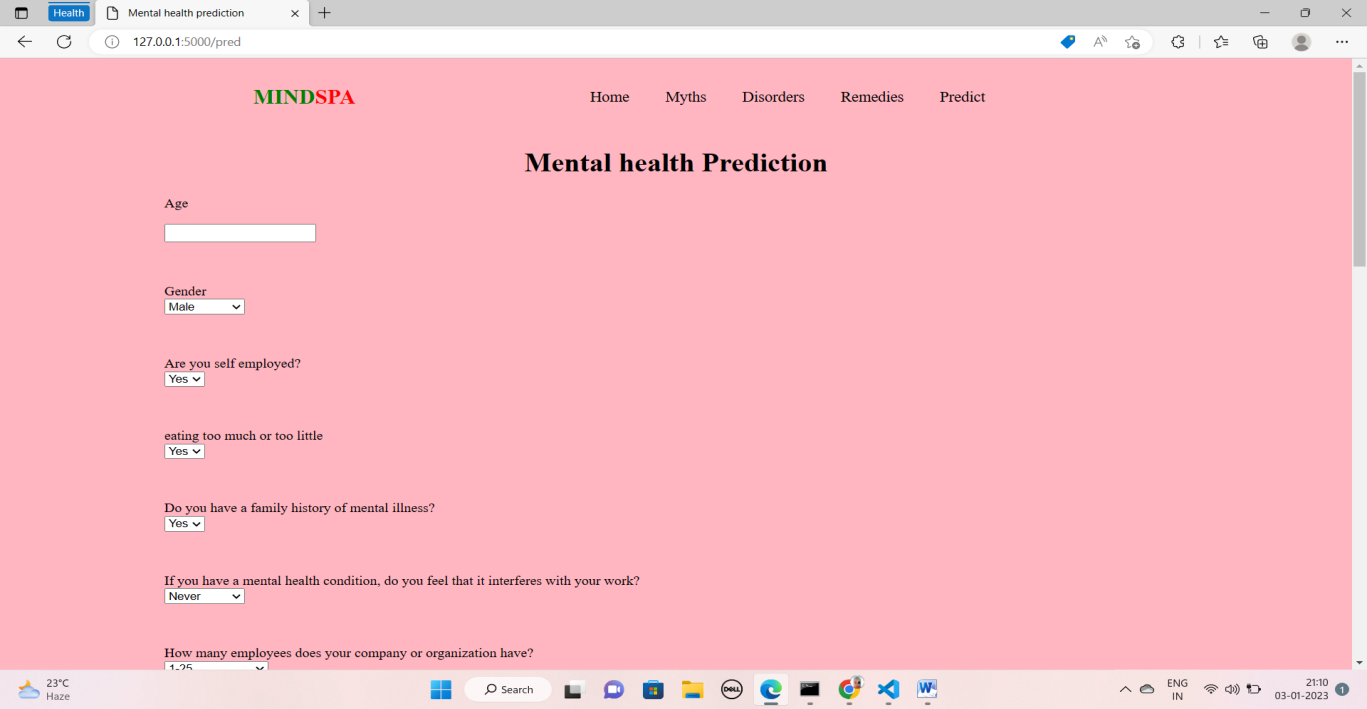
**Remedies.html:**

****

**Fig:3.15 remedies.html**

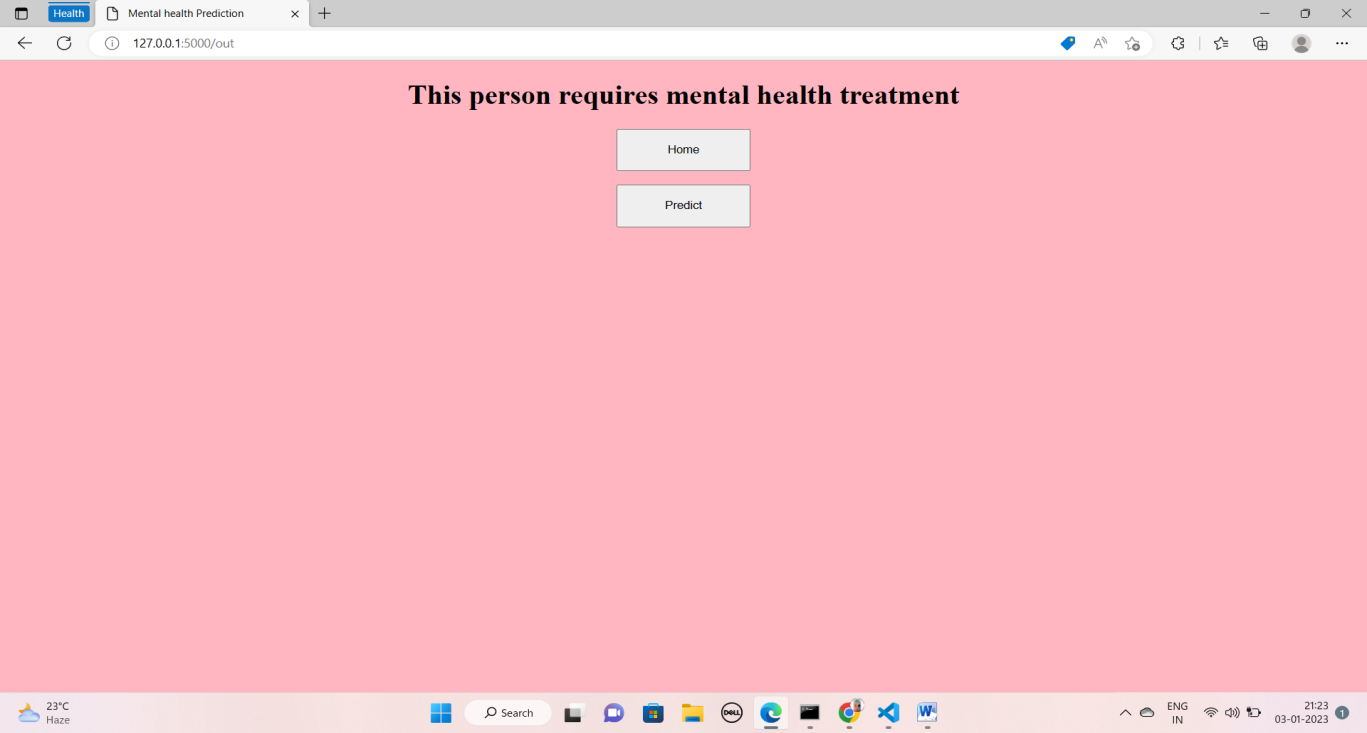
**Prediction:**

Here the user is having questionnaire based on the answers to the questions out application will predict the mental state of that user if he requires treatment or not.

****

**Fig:3.16 index.html**

**Output.html:**

****

**Fig:3.17 output.html**

**CHAPTER-4**

**RESULTS AND DISCUSSIONS**

* 1. **EXPERIMENT SETUP:**

The system requires hardware, Python 3.9, Scikit learn, Numpy,Flask, Visual Code , Pandas.

**Hardware :** The word hardware in computing refers to the physical, tangible parts of a computer system, electrical, electronic, electromechanical, and mechanical component Input devices are keyboard and mouse. The memory contains pieces of information in sections, like books in the library. The programs and data are stored there. It is also known as RAM. The processor works like the human brain, and also it examines and alters data, loads, and carries out program instructions. Also searches for memory programs to run.Cables, cabinets or boxes, peripherals of all kinds, and any other physical element involved, make up the Hardware or material support.

**Python:** Python was designed to be highly [extensible](https://en.wikipedia.org/wiki/Extensibility) via modules. This compact modularity has made it particularly popular as a means of adding programmable interfaces to existing applications. Van Rossum's vision of a small core language with a large standard library and easily extensible interpreter.

**Scikit learn:** Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistence interface in Python. This library, which is largely written in Python, is built upon NumPy, SciPy and Matplotlib.

**Numpy:** NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra, fourier transform, and matrices. NumPy was created in 2005 by Travis Oliphant. It is an open source project and you can use it freely. NumPy stands for Numerical Python.

In Python we have lists that serve the purpose of arrays, but they are slow to process. NumPy aims to provide an array object that is up to 50x faster than traditional Python lists. The array object in NumPy is called ndarray, it provides a lot of supporting functions that make working with ndarray very easy. Arrays are very frequently used in data science, where speed and resources are very important.

**Flask:** Flask is a lightweight [WSGI](https://wsgi.readthedocs.io/) web application framework. It is designed to make getting started quick and easy, with the ability to scale up to complex applications. It began as a simple wrapper around [Werkzeug](https://palletsprojects.com/p/werkzeug) and [Jinja](https://palletsprojects.com/p/jinja) and has become one of the most popular Python web application frameworks.

It is a micro web framework written in Python that speeds up application development by providing essential back end components for programmers to build upon. Flask is simple and lightweight—one of the most manageable frameworks around—and contains only the vital necessities for web development. It is, however, also designed to be highly extensible so developers can customize it however they see fit.

**Visual Code:** Visual Studio is the fastest IDE for productivity. Target any platform, any device. Build any type of application. Work together in real time. Diagnose and stop problems before they happen. It makes the stuff you do every day more fluid and responsive.

**Pandas:** Pandas is an open-source library that is made mainly for working with relational or labeled data both easily and intuitively. It provides various data structures and operations for manipulating numerical data and time series. This library is built on top of the NumPy library. Pandas is fast and it has high performance & productivity for users.

* 1. **ACCURACY**

# **Precision Score Estimation:**

|  |  |
| --- | --- |
| **ALGORITHM** | **PRECISION SCORE** |
| Logistic regression | 0.853 |
| KNN classifier | 0.720 |
| Decision tree classifier | 0.760 |
| Random forest classifier | 0.834 |
| AdaBoost classifier | 0.861 |
| Gradient Boosting classifier | 0.837 |

**Table:4.1 Accuracies of ml algorithms**

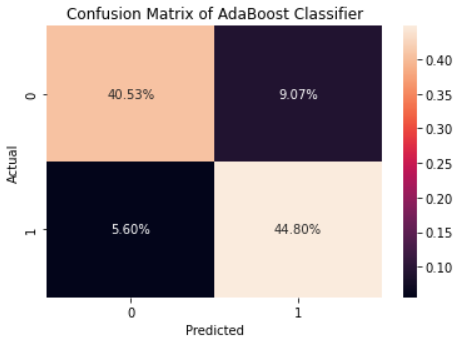
The above mentioned are the accuracies of the machine learning models after they are trained.

Final model we have taken for our predictions is adaboost classification its actual accuracy score is 85.33.

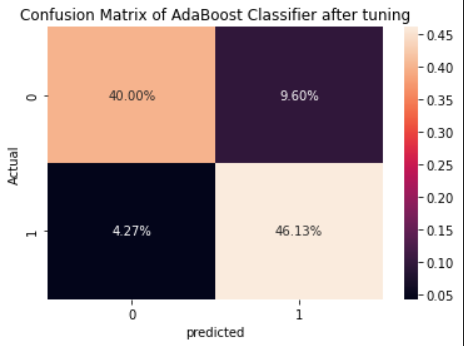
To increase accuracy we have performed a technique called tuning.For tuning we imported RandomizedSearchCv from sklearn.model\_selection.

After tuning the accuracy increased upto 0.86133.

**4.3 CONFUSION MATRIX**



**Fig:4.1 Confusion matrix before tuning**

****

**Fig:4.2 Confusion matrix after tuning**

**4.4 ROC CURVE**

**ROC or Receiver Operating Characteristic curve represents a probability graph to show the performance of a classification model at different threshold levels**. The curve is plotted between two parameters, which are:

* **True Positive Rate or TPR**
* **False Positive Rate or FPR**

In the curve, TPR is plotted on Y-axis, whereas FPR is on the X-axis.

TPR or True Positive rate is a synonym for Recall, which can be calculated as:

**TPR=TP/TP+FN**

FPR or False Positive Rate can be calculated as:

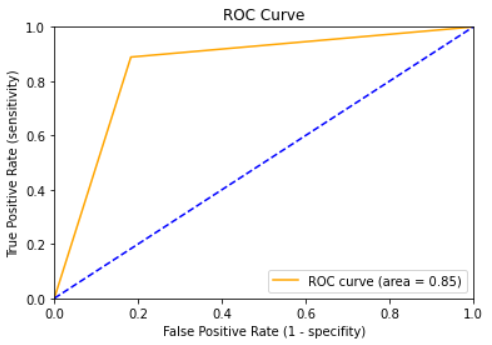
**FPR=FP/FP+TN**

Here, **TP**: True Positive

**FP**: False Positive

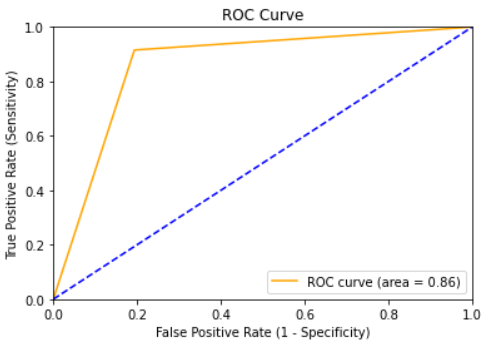
**TN**: True Negative

**FN**: False Negative

****

**Fig:4.3 ROC curve before tuning**

**ROC CURVE AFTER TUNING**

****

**Fig:4.4 ROC curve after tuning**

**4.5 CLASSIFICATION REPORT**

It is one of the performance evaluation metrics of a classification-based machine learning model. It displays your model’s precision, recall, F1 score and support. It provides a better understanding of the overall performance of our trained model.

**Precision:** Precision is defined as the ratio of true positives to the sum of true and false positives.

**Recall:** Recall is defined as the ratio of true positives to the sum of true positives and false negatives.

**F1 Score:** The F1 is the weighted harmonic mean of precision and recall. The closer the value of the F1 score is to 1.0, the better the expected performance of the model is.

**Support:** Support is the number of actual occurrences of the class in the dataset. It doesn’t vary between models, it just diagnoses the performance evaluation process.

**Classification report before tuning:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **F1 score** | **support** |
| **0** | 0.88 | 0.82 | 0.85 | 186 |
| **1** | 0.83 | 0.89 | 0.86 | 189 |
| **accuracy** |  |  | 0.85 | 375 |
| **macro avg** | 0.86 | 0.85 | 0.85 | 375 |
| **weighted avg** | 0.85 | 0.85 | 0.85 | 375 |

**Table:4.2 Classification report 1**

## Classification report after tuning:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **precision** | **recall** | **F1 score** | **support** |
| **0** | 0.90 | 0.81 | 0.85 | 186 |
| **1** | 0.83 | 0.92 | 0.87 | 189 |
| **accuracy** |  |  | 0.86 | 375 |
| **macro avg** | 0.87 | 0.86 | 0.86 | 375 |
| **weighted avg** | 0.87 | 0.86 | 0.86 | 375 |

**Table:4.3 Classification report 2**

**CHAPTER – 5**

**CONCLUSION AND FUTURE SCOPE OF STUDY**

For the effective and efficient prediction of mental health disease at an early stage many trained systems have been used in the medical department which uses various Machine learning technique so it becomes of utmost importance to analyse and approach the best system for the domain as per the requirement. In developing and under-developed countries, where mental health is considered as low priority Al, can be harnessed fully to predict mental illness at a very early stage. Mental health condition can alone create great distress and impairment for an individual; as per the WHO report where depression alone can affect 11% of the world population so it will be a worthy investment for the developing nations to invest in mental health prognosis using AI-based technology for getting their patients treated at an initial stage and conquering greater heights in various fields. The capacity of AI can be completely utilized to provide greater intuition to the disease breakthrough and impact on healthcare.

To resolve the mental well-being machine learning technique play important role. It holds great promise to transform mental health care. Its tools also hold the potential to extend the current capabilities of clinicians, to deal with complex problems and ever-expanding information streams that stretch the limits of human ability. In the world of advancement of technologies we are trying to create a system which will predict the mental health status of a person. The outcome of the system is to give prediction based on data (questionnaire) or digital data.

Our proposed different levels of questionnaire and based on the results of that provide free checking of a person’s mental state and help him by suggestions, tasks and the doctor recommendations if required.

**FUTURE WORK**

Further advancement includes user logins and dashboard updation based on users mental state and also predicting various disorders that will caused by mental stressUp-gradation of a prediction model with advanced algorithm for patients in psychological distress by determining more significant predictors that lead to the disease. Traditional psychotherapy and treatment can be replaced by the digital application using a virtual psychiatrist and personalized patient monitoring will provide clinicians far more profound understanding of the patient's illness. Implementation of such system will reduce the limitation of non-hospitalized patients who are devoid of treatment and visit to the clinicians on a frequent basis.

**REFERENCES**

[1] M. Srividya1 & S. Mohanavalli1 & N. Bhalaji, “Behavioral Modeling for Mental Health using Machine Learning Algorithms”, Springer Science+Business Media, LLC, part of Springer Nature 2018.

[2] Ms. Sumathi M.R., Dr. B. Poorna Prediction of Mental Health Problems Among Children Using Machine Learning Techniques, International Journal of Advanced Computer Science and Applications, Vol. 7, No. 1, 2016.

[3]Shatte ABR, Hutchinson DM, Teague SJ. “Machine learning in mental health: a scoping review of methods and applications”, Psychological Medicine 49, 1426–1448. S0033291719000151, 2019. [4] Ariel Rosenfeld, David Benrimoh, “Big Data Analytics and AI in Mental Healthcare”, arXiv:1903.12071v1 [cs.CY] 12 Mar 2019.

[4] Ariel Rosenfeld, David Benrimoh, “Big Data Analytics and AI in Mental Healthcare”, arXiv:1903.12071v1 [cs.CY] 12 Mar 2019.

[5] Maryam Mohammed Aldarwish , Hafiz Farooq Ahmed, “Predicting Depression Levels Using Social Media Posts”, 2017 IEEE 13th International Symposium on Autonomous Decentralized Systems, 2017.

[6] Sharath Chandra Guntuku, David B Yaden, Margaret L Kern, “Detecting depression and mental illness on social media: an integrative review”, 2352-1546/ã 2017 Elsevier Ltd.

[7] Dr. E. Chandra Blessie , Bindu George, A Novel approach for Psychiatric Patient Detection and Prediction using Data Mining Techniques, International Journal of Engineering Research & Technology, Volume 7, Issue 05-2019.

[8]Kevin Daimi, Shadi Banitaan, Using Data Mining to Predict Possible Future Depression Cases, International Journal of Public Health Science (IJPHS) Vol.3, No.4, December 2014, pp. 231 ~ 240 ISSN: 2252-8806.

[9] Susel Góngora Alonso1 & Isabel de la Torre-Díez1 & Sofiane Hamrioui2 & Miguel López-Coronado1 & Diego Calvo Barreno1 & Lola Morón Nozaleda3 & Manuel Franco4, Data Mining Algorithms and Techniques in Mental Health: A Systematic Review, Journal of Medical Systems (2018) 42: 161

[10] Vivek Patel , Piyush Mishra , J C Patni, “PsyHeal: An Approach to Remote Mental Health Monitoring System”, International Conference on Advances in Computing and Communication Engine

[11] Hui Yang and Peter A. Bath, “The Use of Data Mining Methods for the Prediction of Dementia: Evidence From the English Longitudinal Study of Aging”, IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 24, NO. 2, FEBRUARY 2020.

[12]Vidhi Mody, Vrushti Mody, “Mental Health Monitoring System using Artificial Intelligence: A Review”, 2019 5th International Conference for Convergence