MENTAL HEALTH PREDICTION USING MACHINE LEARNING
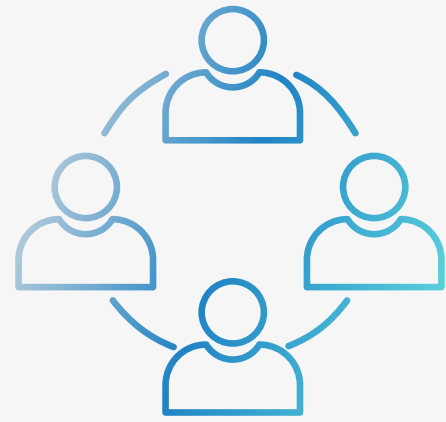
Group Project Presentation
Feb 23 2023

TABLE OF CONTENT

1. GROUP MEMBERS

L JEYA SURIYA
19/11/EC/032

jeya47_soe@jnu.ac.in
+91 99400 44565

ABHISHEK JAIN
19/11/EC/036

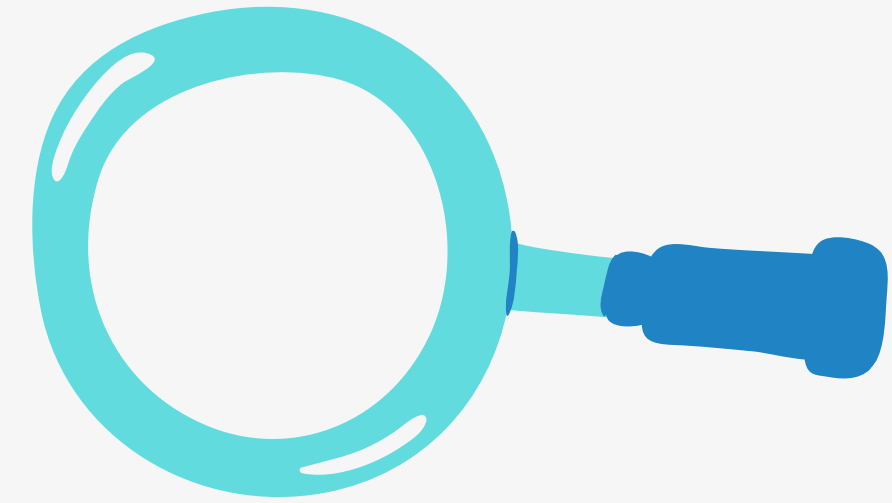abhish34_soe@jnu.ac.in
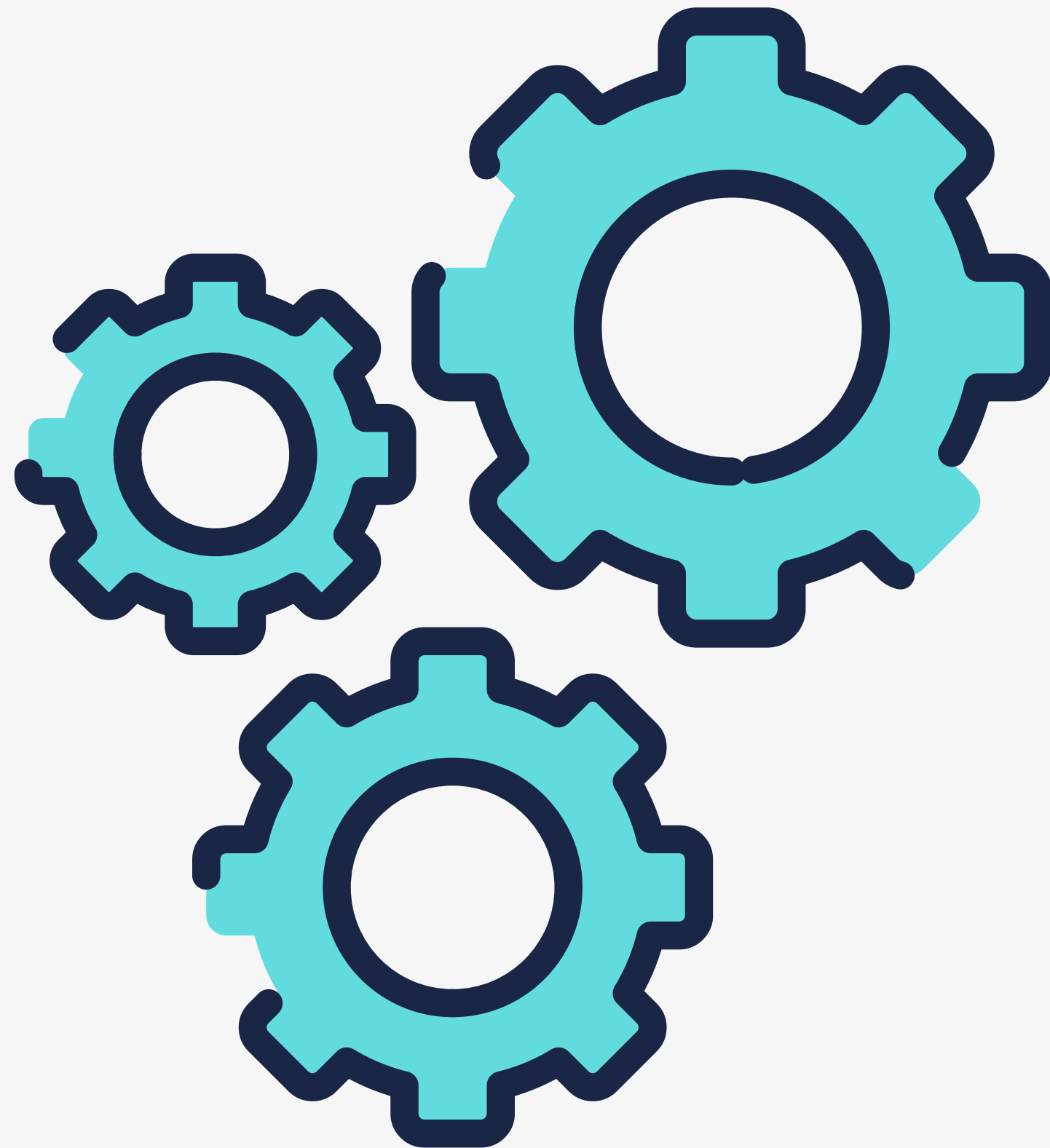91113 64144

JAGESHWAR
19/11/EC/053

jagesh93_soe@jnu.ac.in
62605 28024

2. PROBLEM STATEMENT

To build a machine learning model using classification algorithms to predict an individual's mental health.

3. INTRODUCTION

Mental health relates to a person's emotional, psychological, and social well-being. It is more important because it affects every aspect of our lives, and mental health conditions influence everything we do, think, or say. Mental and physical well-being are equally crucial components of general wellness. Keeping positive mental fitness can boost productivity, enhance our self-image, and improve relationships with our loved ones.
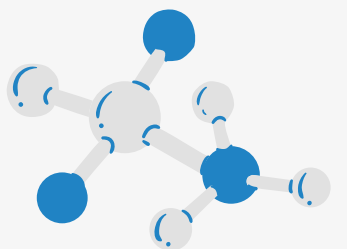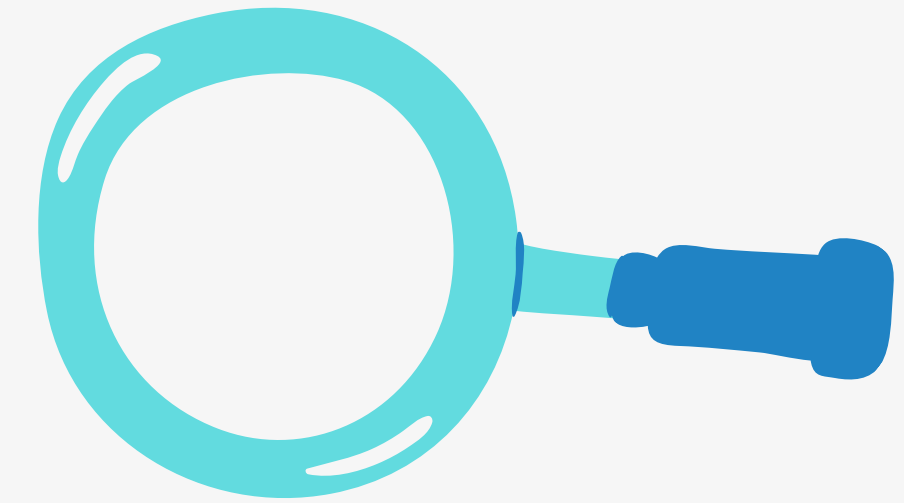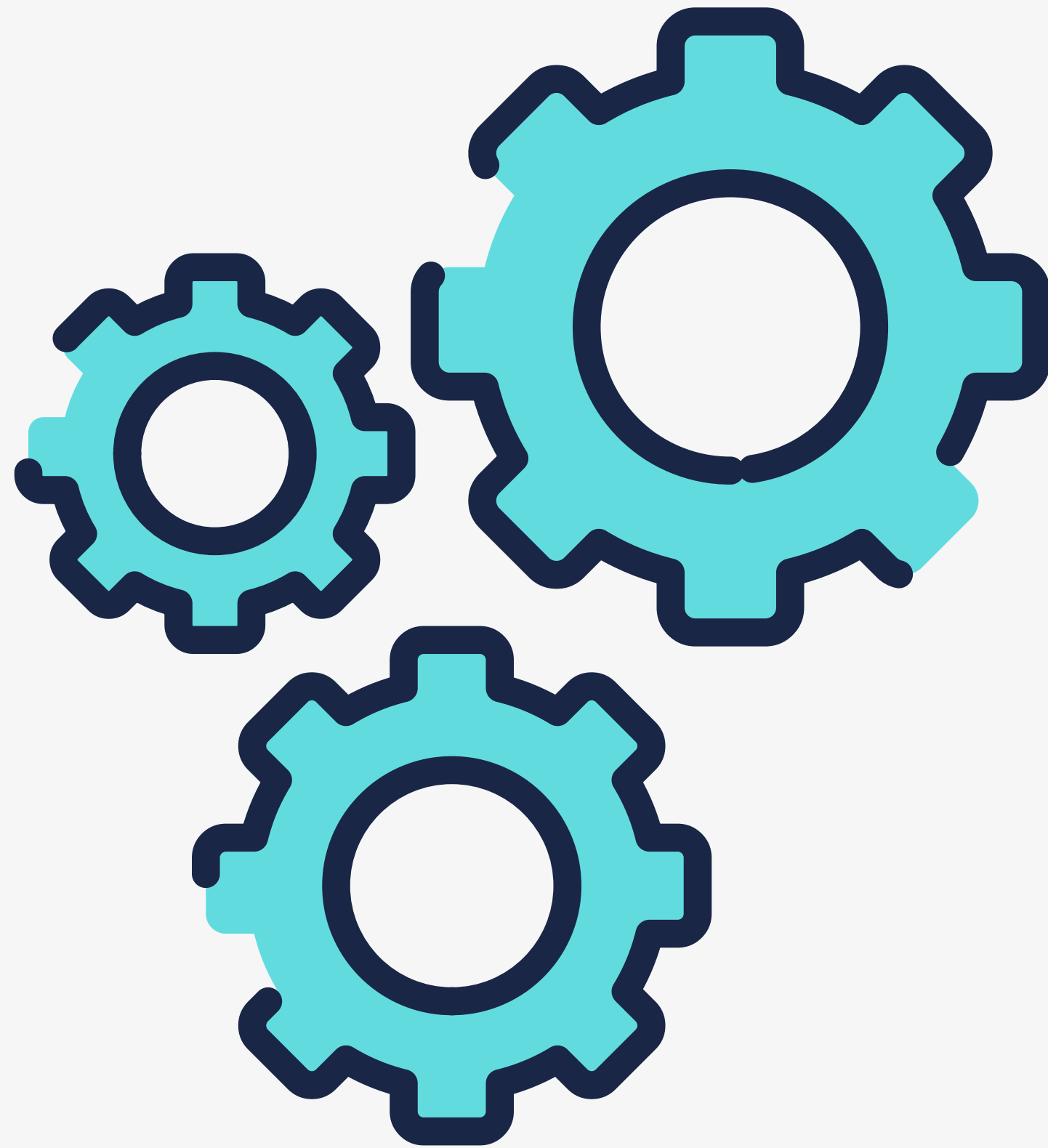
At each point in life, mental health is vital, from childhood to adulthood. Numerous factors contribute to mental health issues that lead to mental illness like stress, social anxiety, depression, etc. The onset of mental illness should be determined without flaws for maintaining an appropriate life balance. Therefore, an ML model that determines the mental health of a person with higher accuracy can be highly beneficial to the medical field.
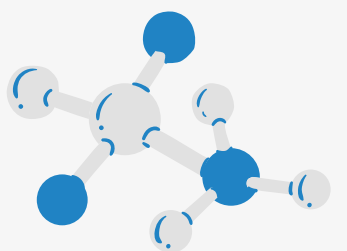
4. METHODOLOGY

Uploading Dataset

We begin by uploading an existing dataset..

Data Cleaning

Here, incomplete, or missing data will be modified, replaced, or eliminated based on necessity.

Data Encoding

Here, categorial data will be converted into integer form for the convenience of our model and more accurate predictions.

Finding Confusion Matrix and ROC curve

The confusion matrix is a matrix used to determine the performance of the classification models for a given set of test data. It can only be determined if the true values for test data are known.

Feature Scaling

This process helps us to handle highly varying units.

4. METHODOLOGY

Tuning

It is done to enhance the model's performance.

Evaluation of Machine Learning ALgorithms

The model is then evaluated using different machine learning algorithms which are Logistic Regression, Adaboost CLassifier, Decision Tree classifier, KNN classifier, Random Forest Classifier, Gradient Boost Classifier, and XGB Classifier.

Calculation of Accuracy

After that, we will calculate the accuracy of different models. The accuracy of any classifying algorithm will depend upon how well the algorithm will classify the data set which is being tested.

Integration of ML model and Web-Page

Then, we'll select the best Machine Learning Model based on the accuracy and integrate it with the already made Web-Page using Flask.
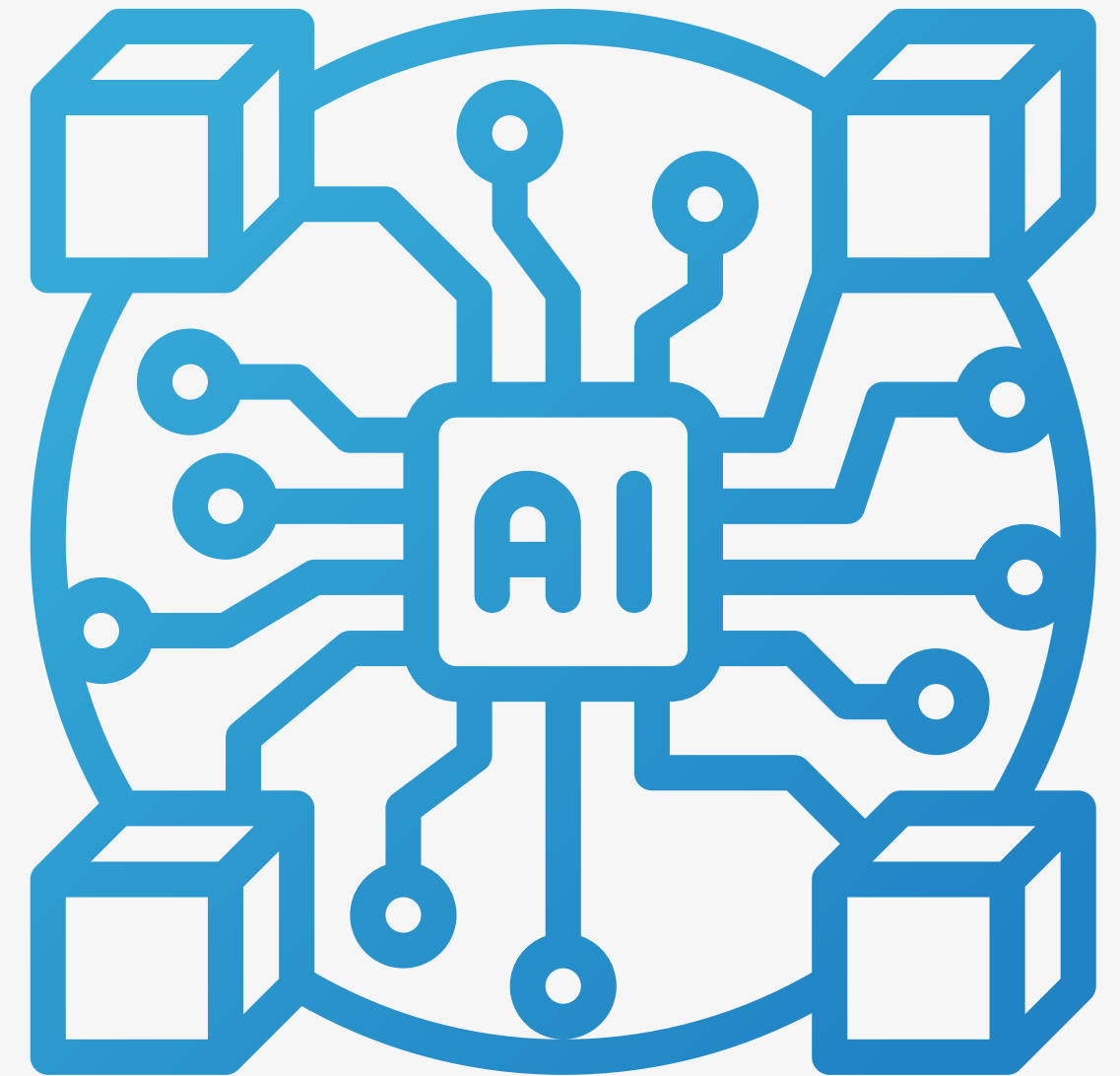
5. ALGORITHMS USED

We'll now look at some algorithms that have been tested during our "evaluation model" phase.

1). Logistic Regression
Logistic regression is a classification algorithm that uses supervised learning to predict the likelihood of a target variable.

2). AdaBoost Classifier
AdaBoost is a method for iterative ensemble construction. The AdaBoost classifier creates a strong classifier by combining multiple low-performing classifiers, resulting in a high-accuracy strong classifier.
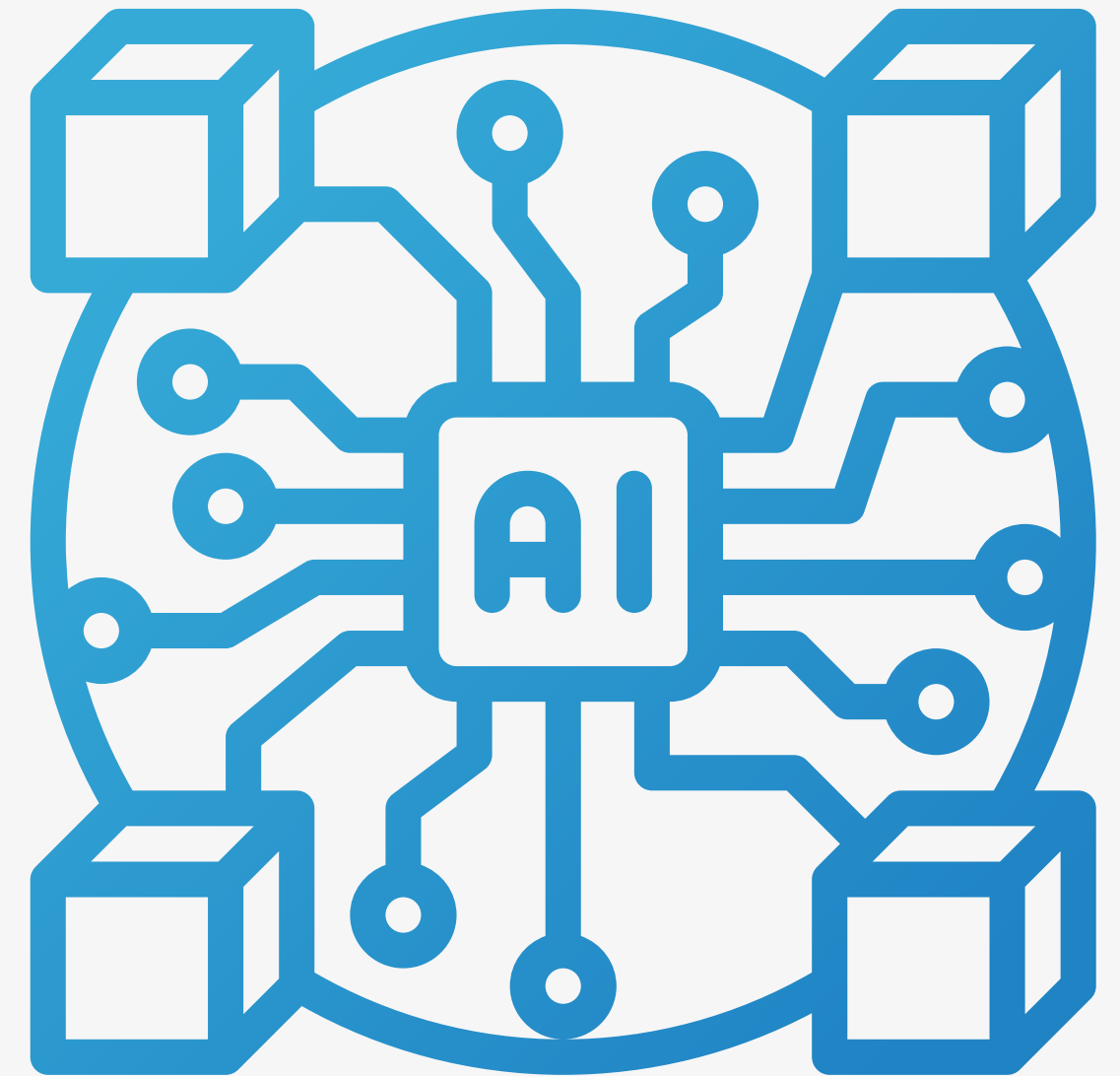
3). Decision Tree Classifier

The Decision Tree algorithm in machine learning is one of the most widely used algorithms today. It is a supervised learning algorithm used for problem classification. It is effective for categorizing both categorical and continuous dependent variables.

4). KNN Classifier

The K-nearest neighbours (KNN) algorithm is a type of supervised machine learning (ML) algorithm that can be used for both classification and regression predictive problems. This algorithm predicts the values of new data points based on 'feature similarity,' which means that the new data point will be assigned a value based on how closely it matches the points in the training set.
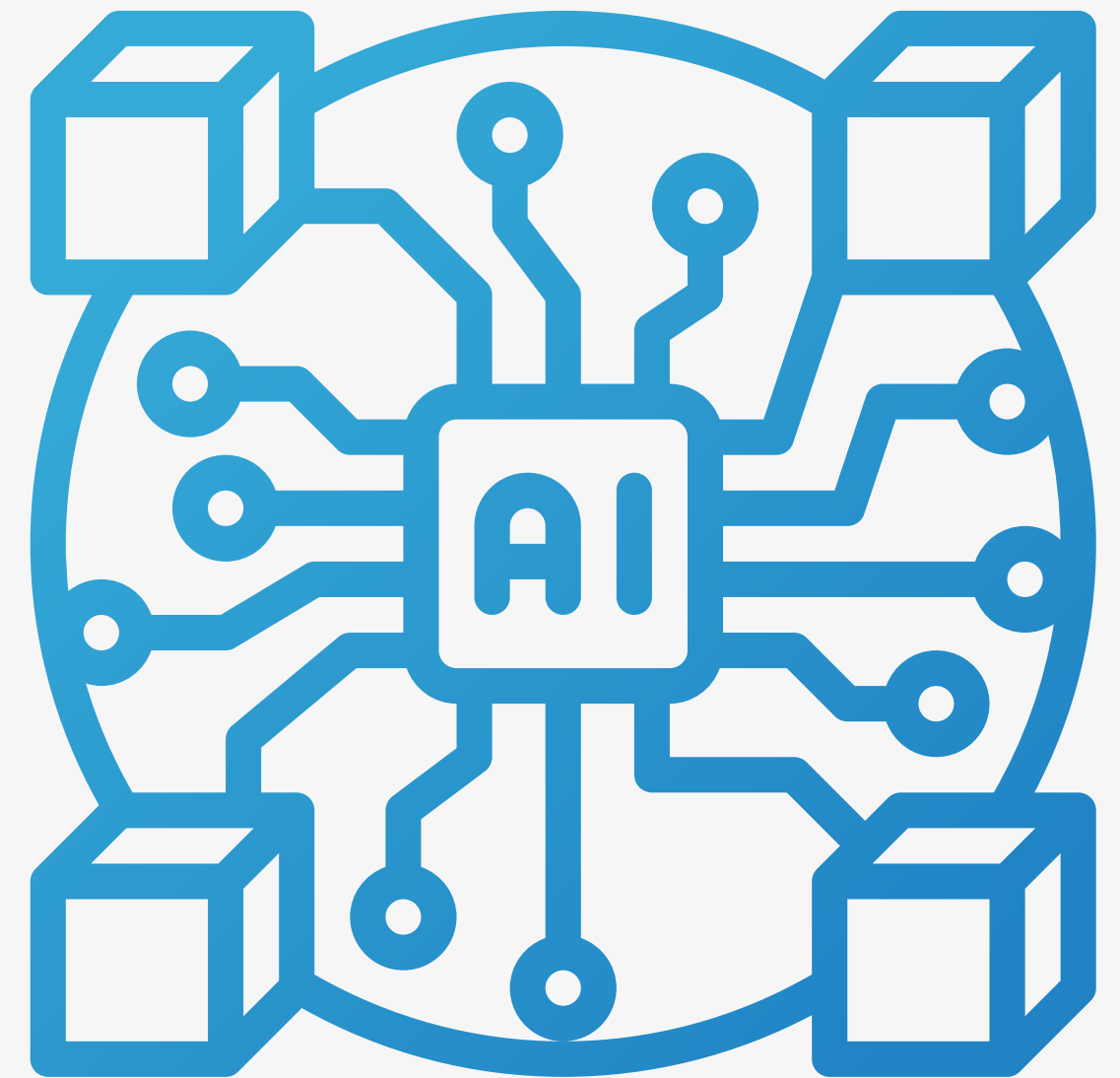
5). Random Forest Classifier

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression.

6). Gradient Boosting Classifier

Gradient Boosting is a functional gradient algorithm that repeatedly selects a function that leads in the direction of a weak hypothesis or negative gradient so that it can minimize a loss function. Gradient boosting classifier combines several weak learning models to produce a powerful predicting model.

7). XGB Classifier

XGBoost is a distributed gradient boosting library that has been optimized for efficient and scalable training of machine learning models.

6. RESULTS

The accuracy of **AdaBoost Classifier** is more accurate compared to other classifiers with an accuracy rate of 86.93.

```
In [55]:  abc_tuned = AdaBoostClassifier(random_state=49,n_estimators=11, learning_rate=1.02)
          abc_tuned.fit(X_train,y_train)
          pred_abc_tuned = abc_tuned.predict(X_test)
          print('Accuracy of Adaboost (tuned)=', accuracy_score (y_test, pred_abc_tuned))

          Accuracy of Adaboost (tuned)= 0.8693333333333333
```

7. CONCLUSION

In this project, we have compared seven different techniques of machine learning which are used to classify the dataset on various mental health problems. It is very clear from the results that all seven machine learning techniques give more accurate results. Among them, the AdaBoost classifier gave the most accurate results. Then, it was integrated with a web-page using flask.

8. FUTURE SCOPE

The data set used in the research is very minimal and in the future, a large data set can be used and the research can be applied to the same for more accuracy

Along with the web-page, a web-app can also be made for the use of the user.