**CSE3013**

**ANALYZING STUDENT’S MARKS AND PREDICTING PERFORMANCE BASED ON ACADEMICS**

**Team Members-**

**Shraddha Mehta 19BCE1722**

**Palak Tandel 19BCE1539**

**Muskan Dhingra 19BCE1148**

**Flow of Code Execution with Steps**

**Software used - Jupyter Notebook**

**Student-mat.csv** file consists of the dataset of students’ Maths subject scores.

**Student-por.csv** file consists of the dataset of students’ Portuguese subject scores.

**Students.csv** file consists of a dataset of all students i.e a merged dataset of Maths and Portugese students.

Step 1: Importing all the required Python libraries

Step 2: Load the Dataset

Step 3: Renaming the variables

Step 4:Creating another column ‘FInal Grade’ based on ‘final\_score’

Step 5: Creating dataframe for classification

Step 6: Label encode fnal\_grade

Step 7: Dataset train\_test\_split

Step 8: Get dummy variables

**DECISION TREE:**

Step 9: Finding the optimal number of minimum samples leaf for decision tree

Step 10: Calculating the Final model Score

**RANDOM FOREST:**

Step 11: Find a good number of estimators

Step 12: Find a good number of minimum samples of leaf for random forest classifier

Step 13: Calculating the Final model Score

**SUPPORT VECTOR CLASSIFICATION:**

Step 14: Calculating the Final Model Score

**LOGISTIC REGRESSION:**

Step 15: Find optimal number of features to use in the model

Step 16: Plotting the feature selection

Step 17: Calculating the Final Model Score

**ADA BOOST CLASSIFICATION:**

Step 18: Calculating the Final Model Score

### **Stochastic Gradient Descent Classification:**

Step 21: Calculating the Final Model Score

**K-Nearest Neighbors:**

Step 22: Define data

Step 23: Find optimal number of features to use in the model

Step 24: Calculating random state

Step 25: Plotting the ROC curve

Step 26: Display the confusion matrix

**GENETIC ALGORITHM:**

In a new Python Notebook, perform the following steps:

Step 27: Perform Steps of loading the dataset (step 1-8) and perform logistic regression(step 15-19)

Step 28: Defining various steps required for the genetic algorithm

Step 29: Define functions for initialization\_of\_population, fitness\_score, selection, crossover, mutation, generations.

Step 29: Accuracy score after genetic algorithm is applied in logistic regression