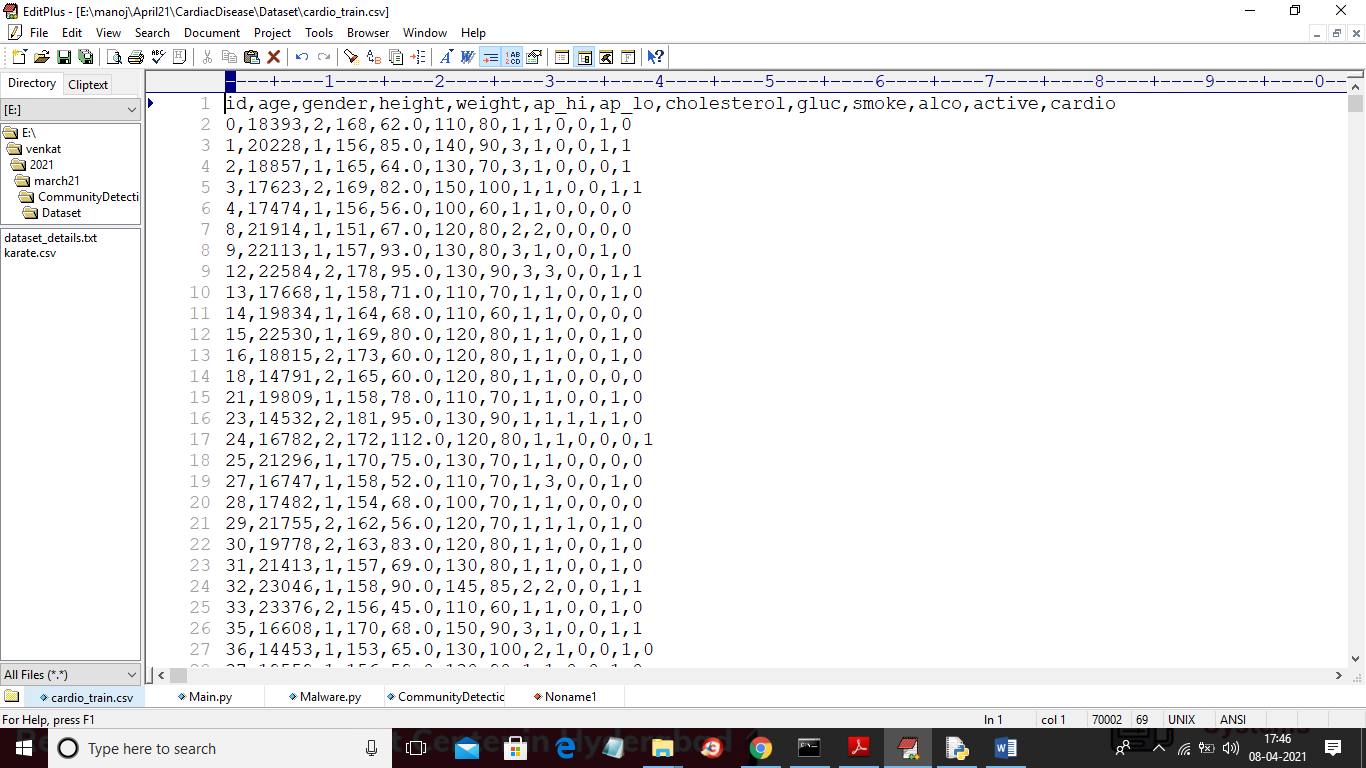
Prediction of Cardiac Disease using Supervised Machine Learning Algorithms

In this paper author is using many supervised machine learning algorithms to predict cardiac disease and this algorithms will be trained on Cardiac dataset obtained from KAGGLE website. Author using algorithms such as SVM, Decision Tree, Random Forest, Naïve Bayes, KNN and Logistic Regression. Among all algorithms decision tree is giving better prediction accuracy. In this paper author is evaluating performance of each algorithms in terms of Precision, Recall, Accuracy and FSCORE.

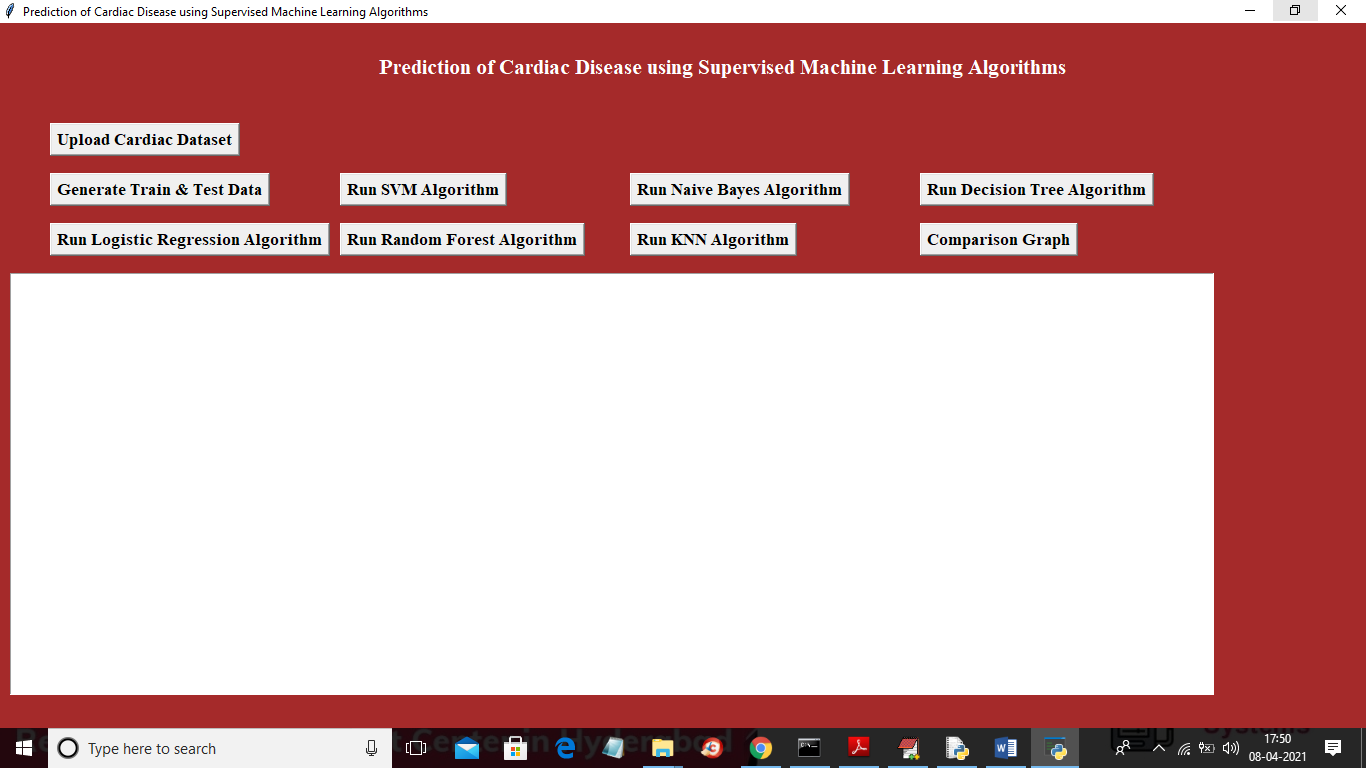
Dataset used in this project is available inside ‘Dataset’ folder and below screen shots showing somes records from dataset



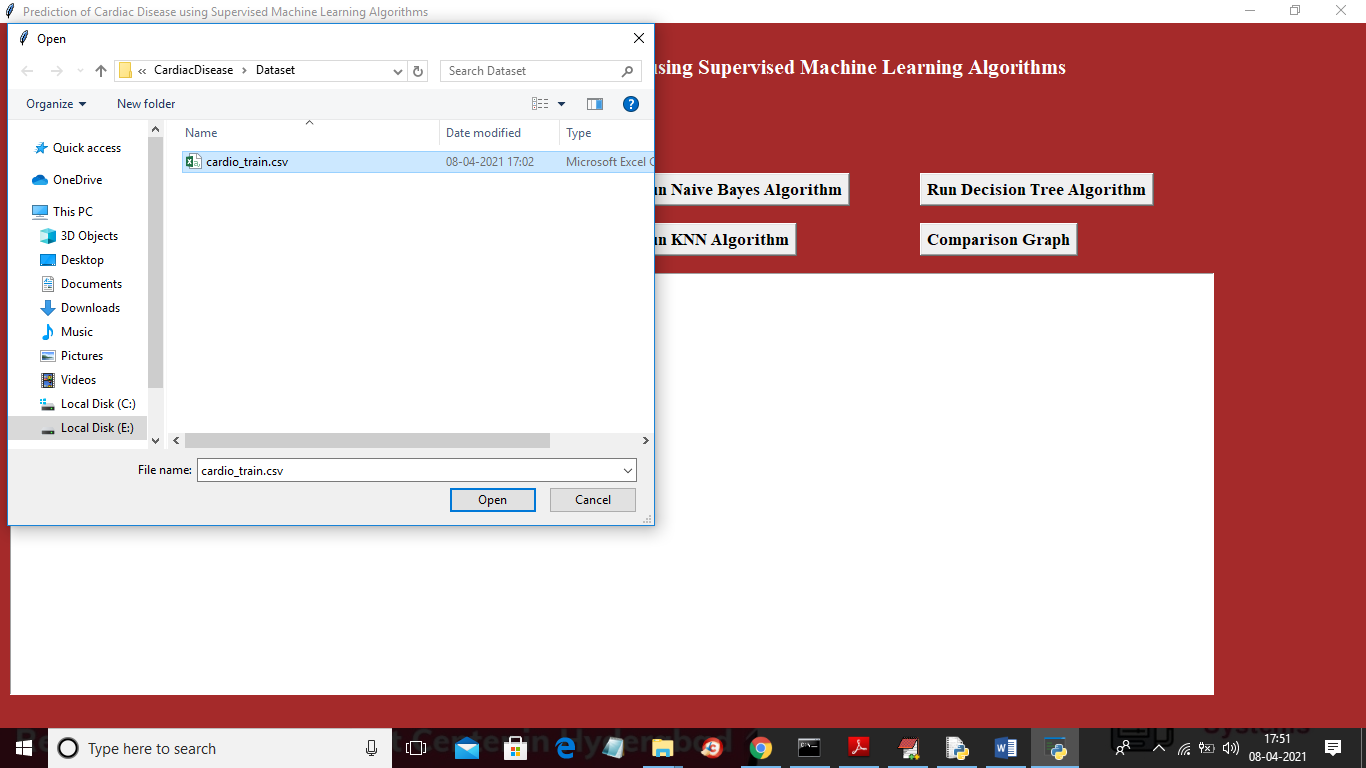
In above screen first row contains dataset column names and other rows contains dataset values and last is the class label as 0 or 1 where 0 means no disease and 1 means cardiac disease detected.

SCREEN SHOTS

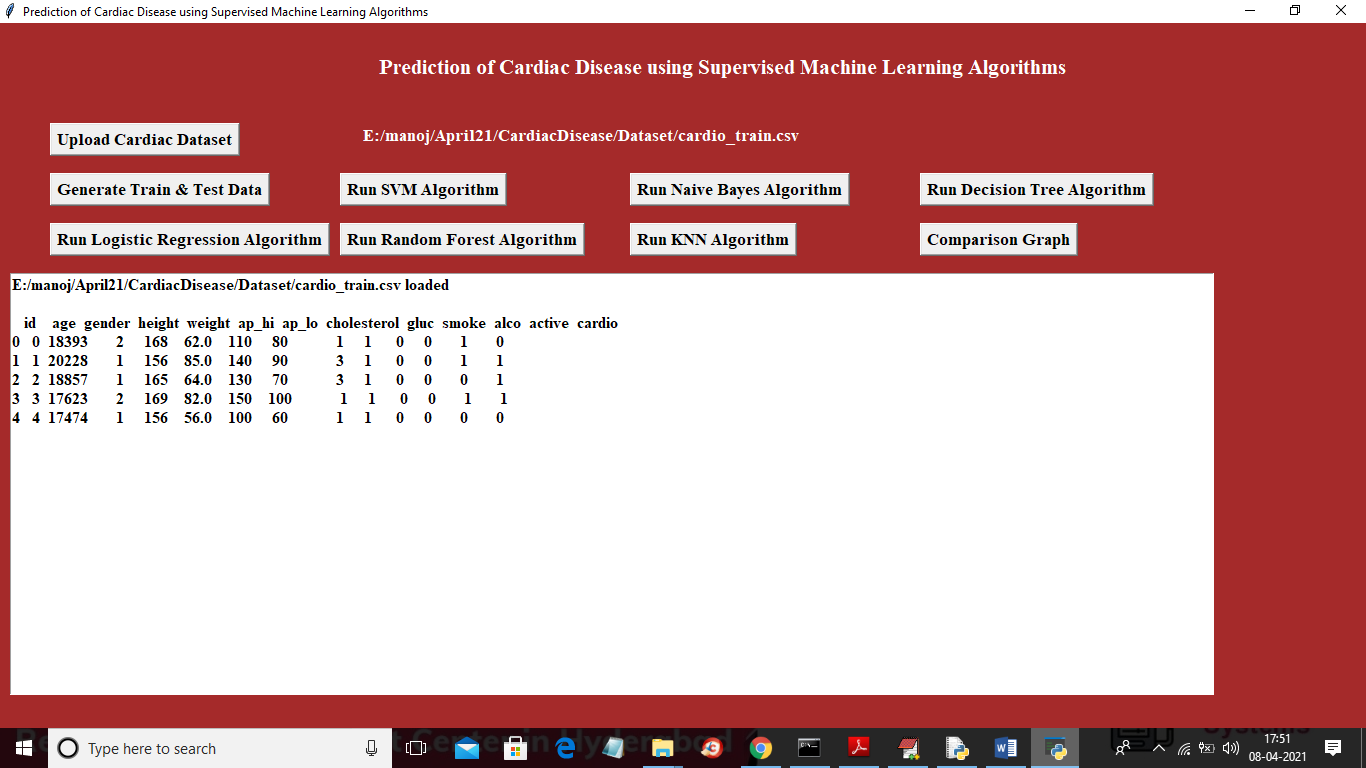
To run project double click on ‘run.bat’ file to get below screen



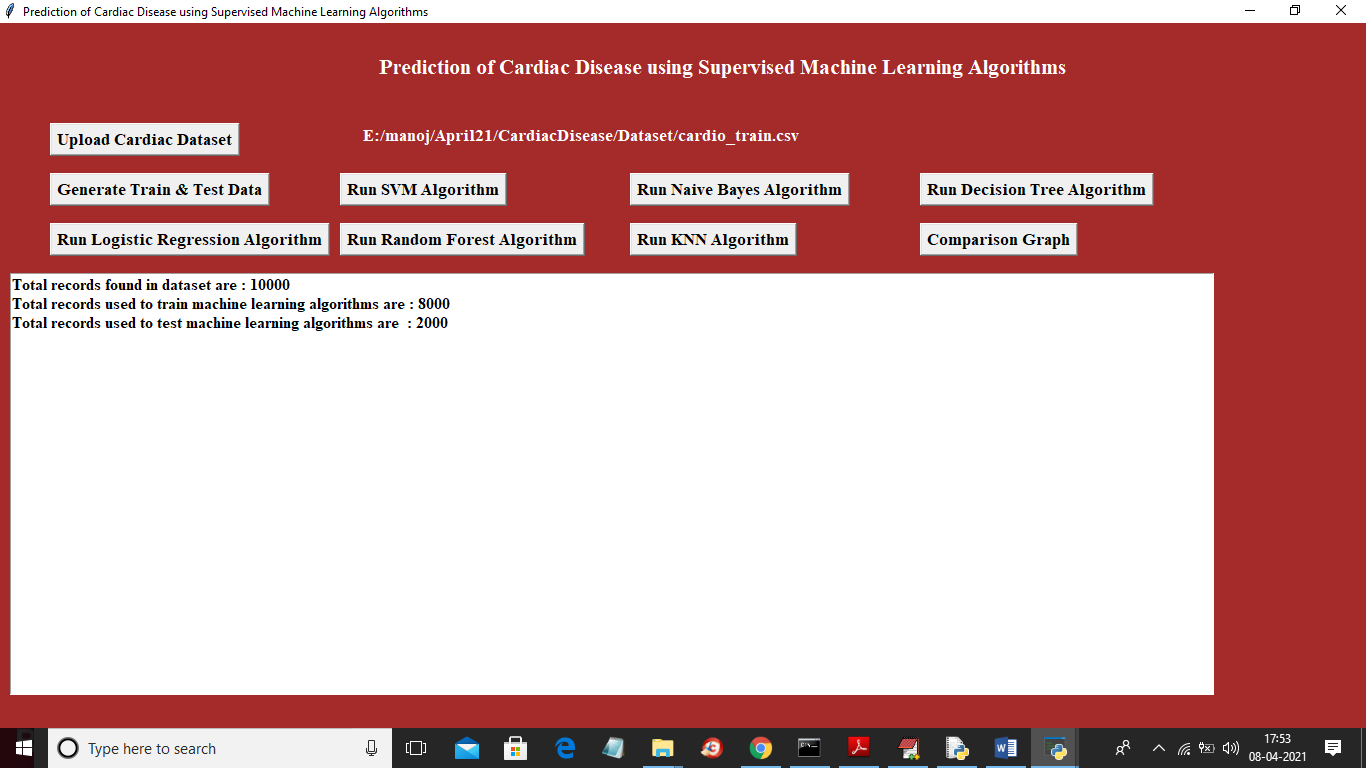
In above screen click on ‘Upload Cardiac Dataset’ button and upload dataset



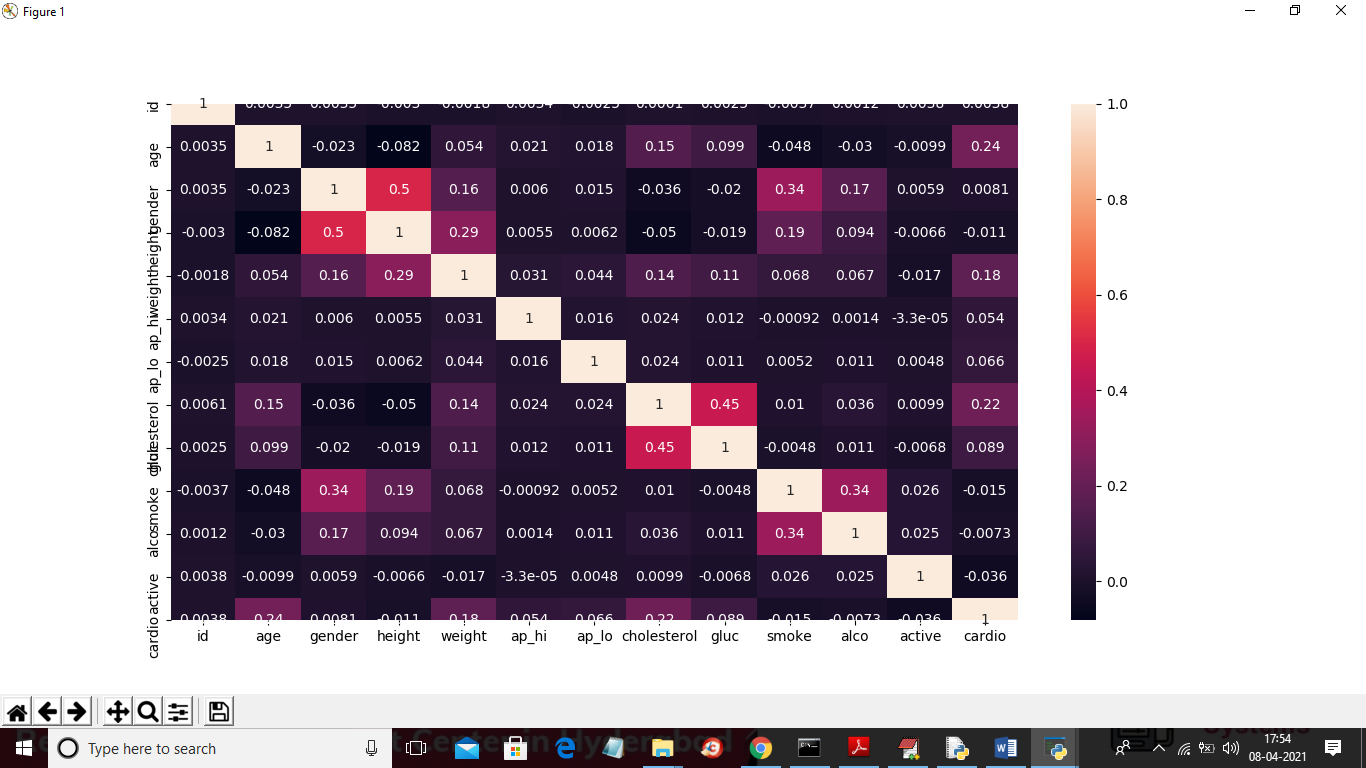
In above screen selecting and uploading ‘cardiac\_train.csv’ file and then click on ‘Open’ button to load dataset and to get below screen



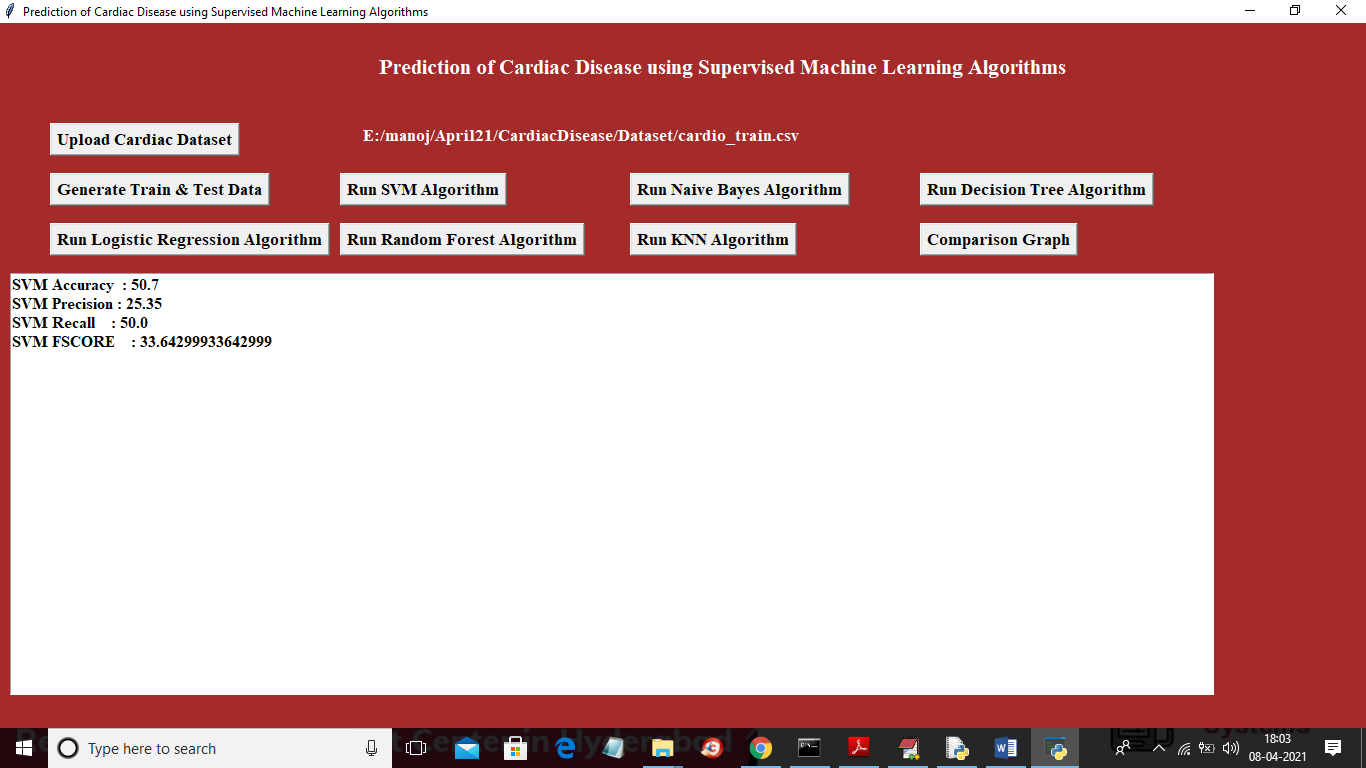
In above screen dataset loaded and we can see some records from dataset and now click on ‘Generate Train & Test Data’ button to divide dataset into train and test part where application used 80% dataset to train machine learning algorithms and used 20% to test the accuracy of those algorithms.



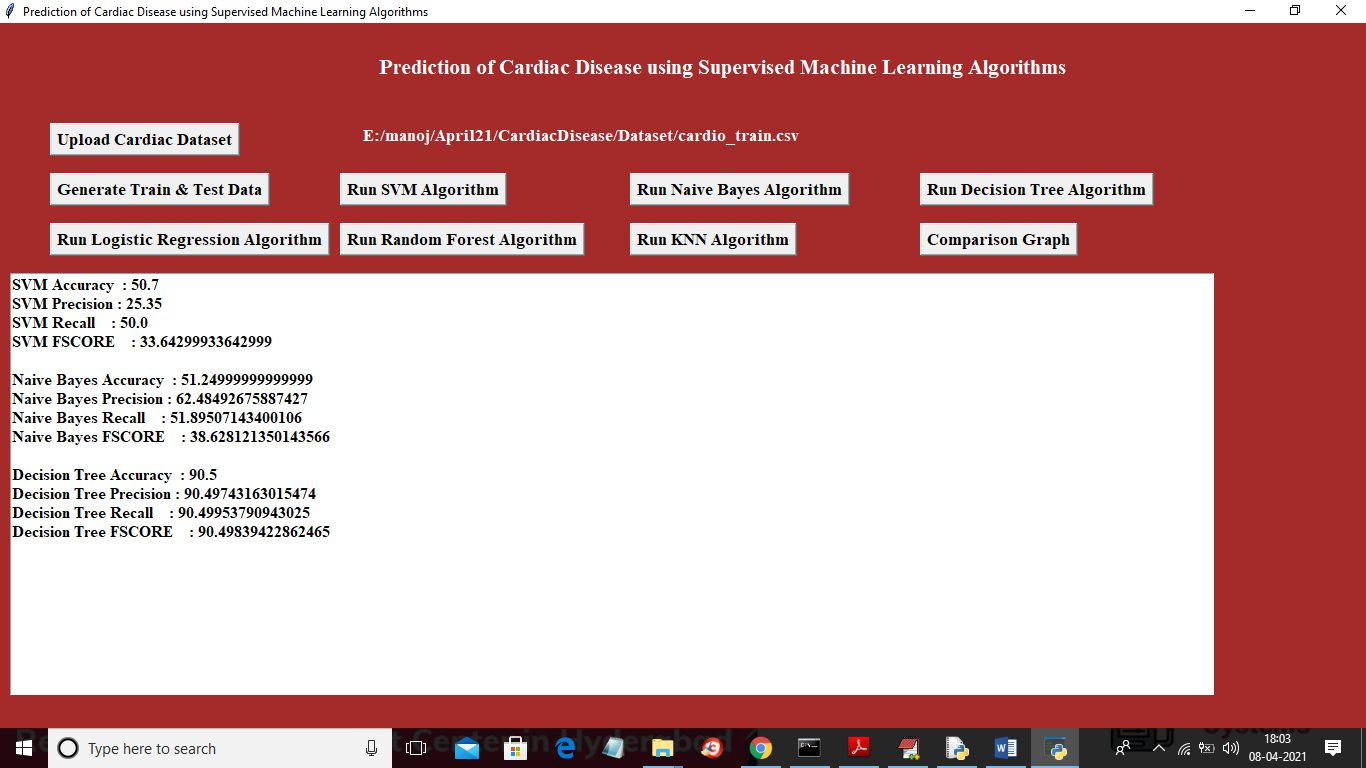
In above screen we can see dataset contains 10000 records and application using 8000 records for training and 2000 records for testing and below graph is the features importance graph where any feature whose correlation value close to 1 will be consider as important feature



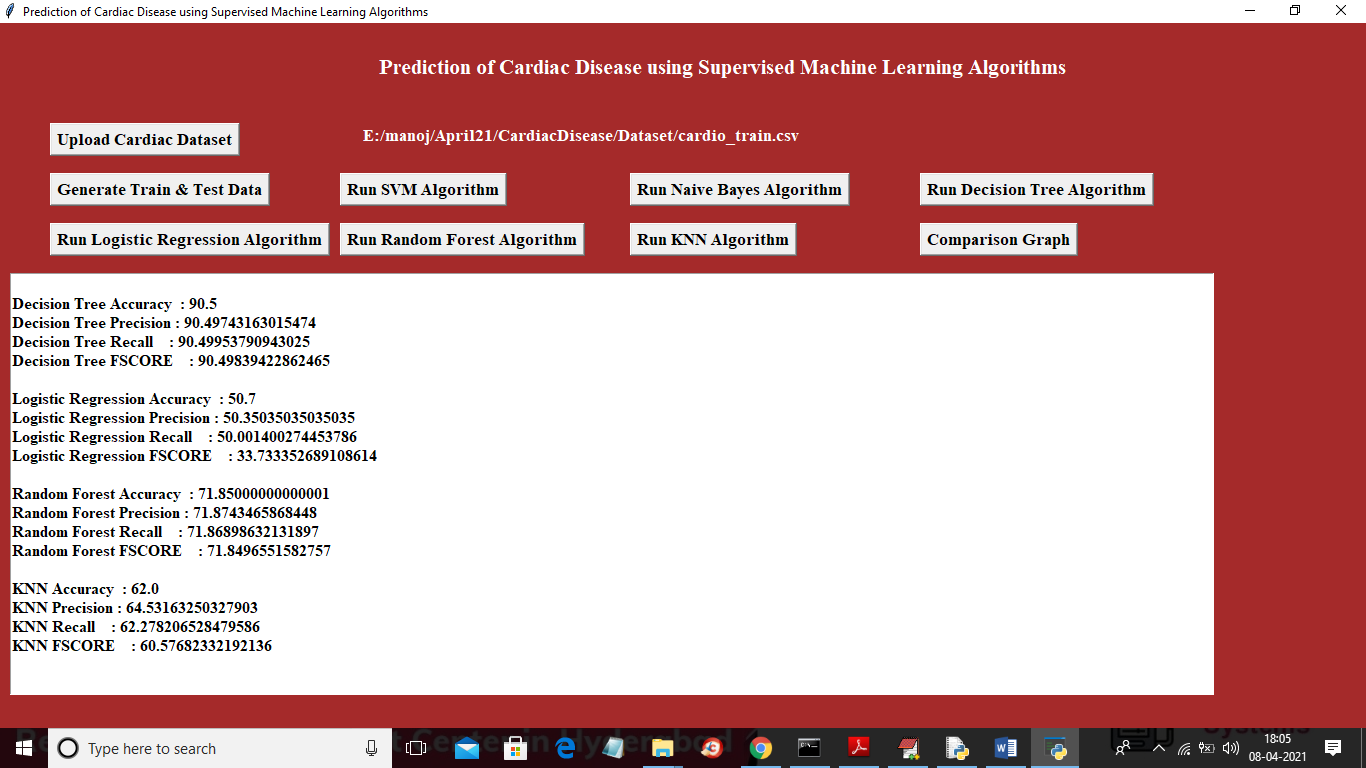
In above screen x and y-axis contains features/column names and y-axis contains correlation values for that features and the feature which has value closer to 1 will be consider as important attributes. In above screen we can see many columns has values 1 and names in that column will be consider as important attribute. Now train and test data is ready and now click on ‘Run SVM Algorithm’ button to train SVM on above dataset



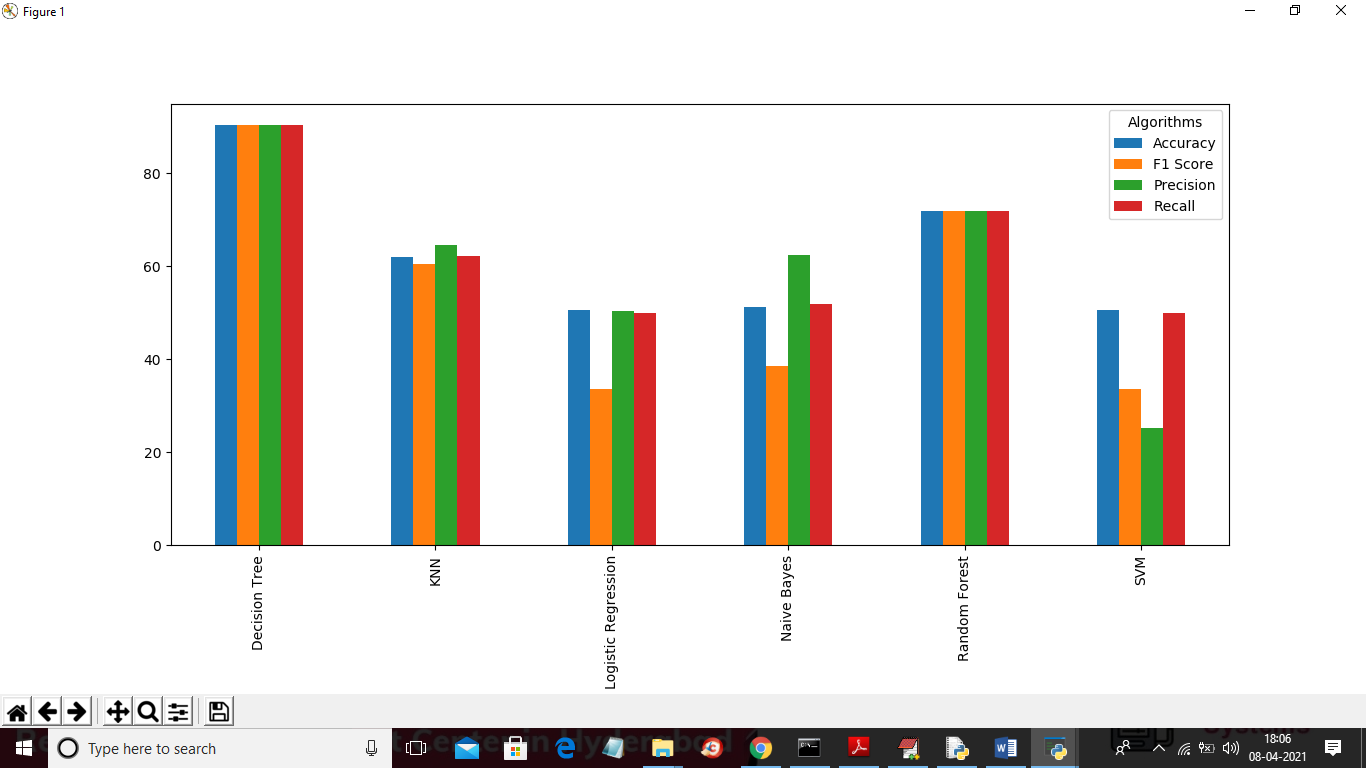
In above screen SVM got 50% accuracy and now click on Naïve Bayes and Decision Tree button to get its accuracy



In above screen with Naïve Bayes we got 51% accuracy and Decision tree we got 90% accuracy and now click on Logistic Regression, Random Forest and KNN Algorithms button to get their prediction accuracy



In above screen with logistic regression we got 50% and random forest got 71% and KNN got 62% accuracy. In above screen we trained all algorithms on cardiac dataset and we got highest accuracy for decision tree algorithm and now click on ‘Comparison Graph’ button to get below comparison graph



In above graph x-axis represents algorithm name and y-axis represents accuracy, precision, recall and FSCORE. From above graph we can conclude that decision tree gave better prediction accuracy.