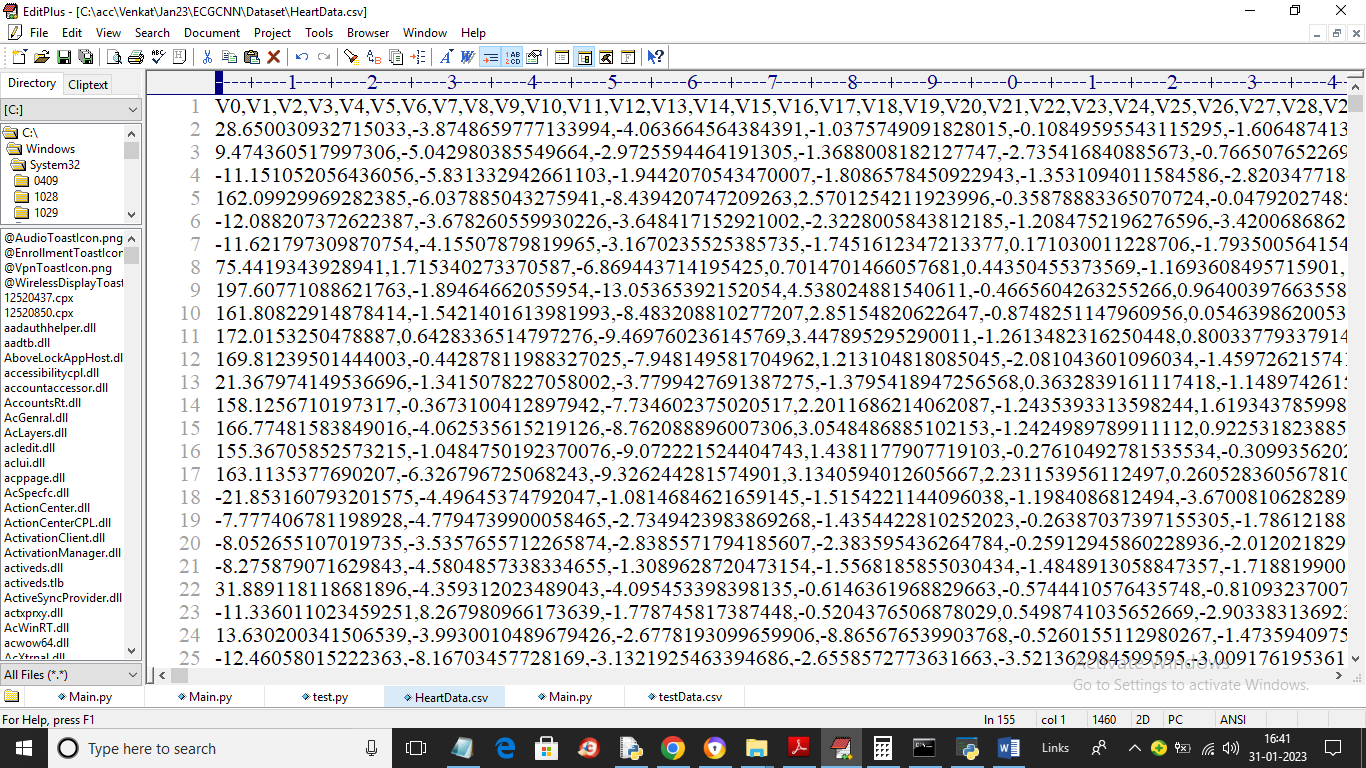
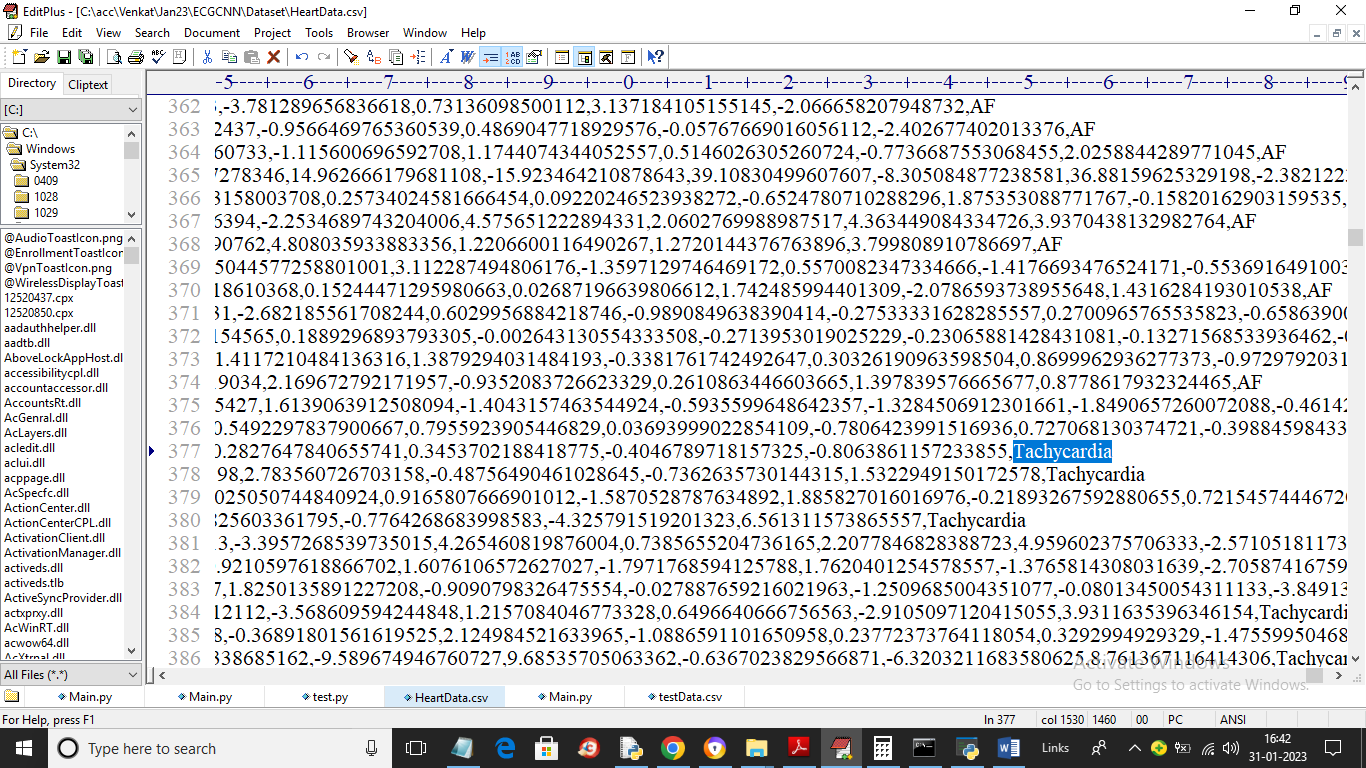
Machine Learning for Real-Time Heart Disease Prediction

In this paper author employing XGBoost algorithm on Heart Disease dataset to predict 7 various heart diseases such as Normal, AF, Tachycardia, Bradycardia, Arrhythmia, Other or Noisy. Propose algorithm giving accuracy more than 99%. To train XGBOOST we are using Heart Disease ECG dataset from KAGGLE and below screen showing heart disease dataset details



In above dataset screen first row contains dataset column names and remaining rows contains ECG signal values and last column contains class labels showing in below screen

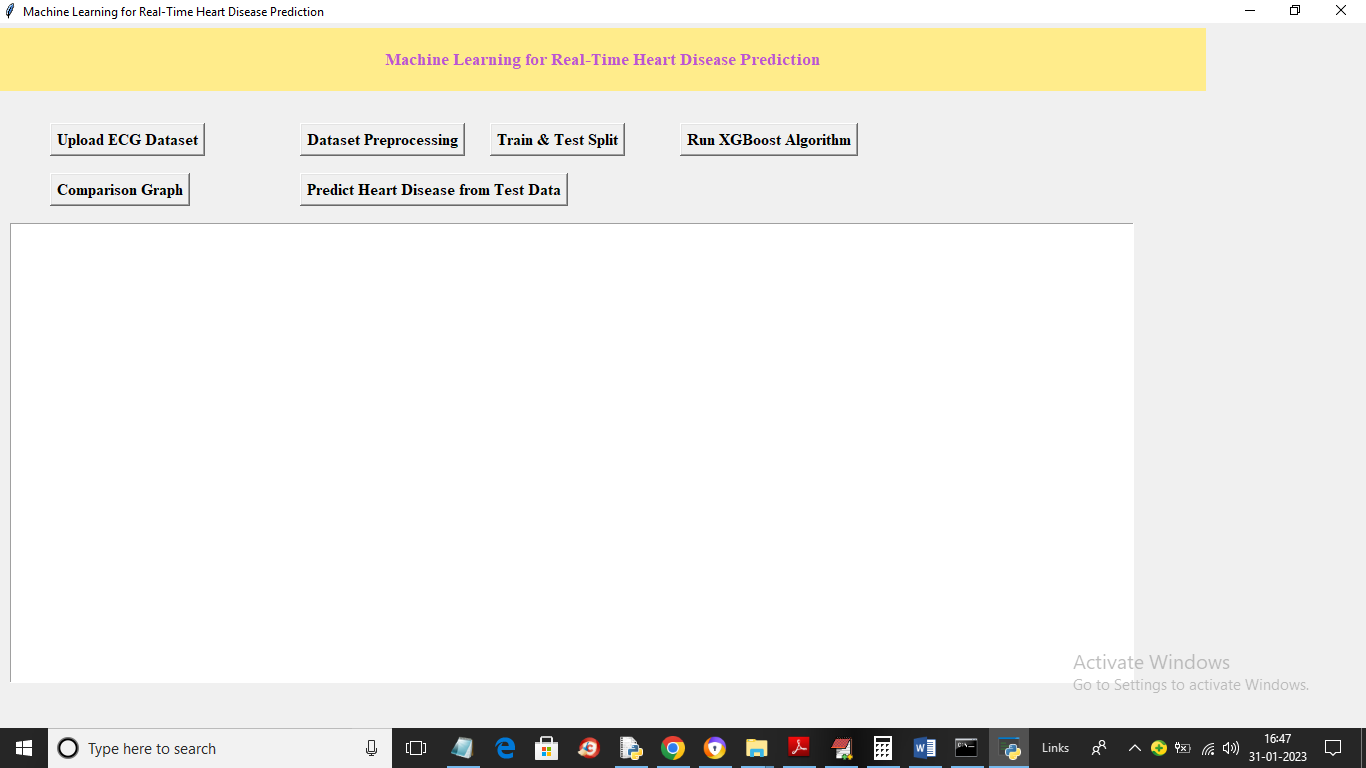


In above screen in last column we can see class label

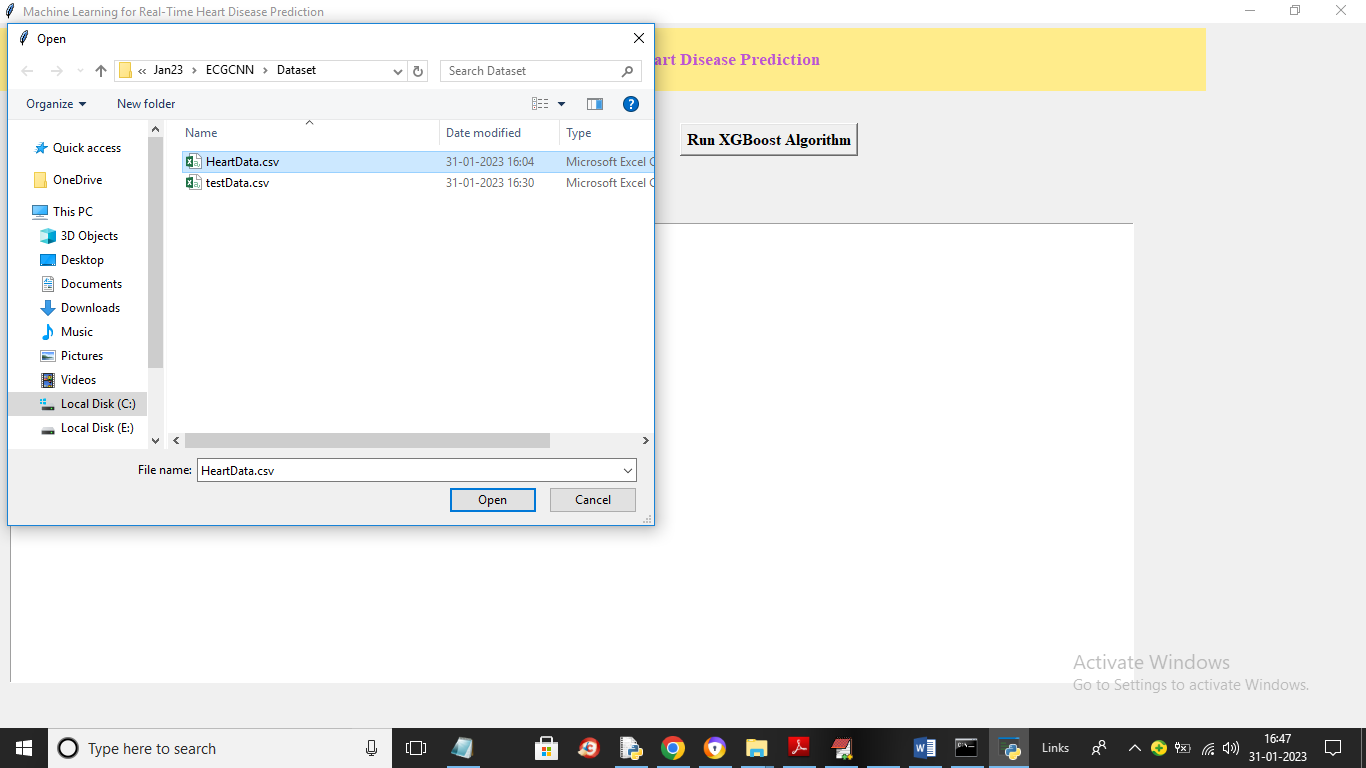
To implement this project we have designed following modules

1. Upload ECG Dataset: using this module we will upload dataset to application
2. Dataset Preprocessing: using this module we will read all dataset values and then remove missing values and then normalize and shuffle the values
3. Train & Test Split: processed values will be split into train and test where application using 80% dataset for training and 20% for testing
4. Run XGBoost Algorithm: using this module we will trained XGBOOST algorithm on 80% training data and then apply on 20% data to calculate prediction accuracy
5. Comparison Graph: using this module we will plot accuracy and precision graph
6. Predict Heart Disease from Test Data: using this module we will upload test data and then application will analyse test data and give predicted labels as normal or heart disease.

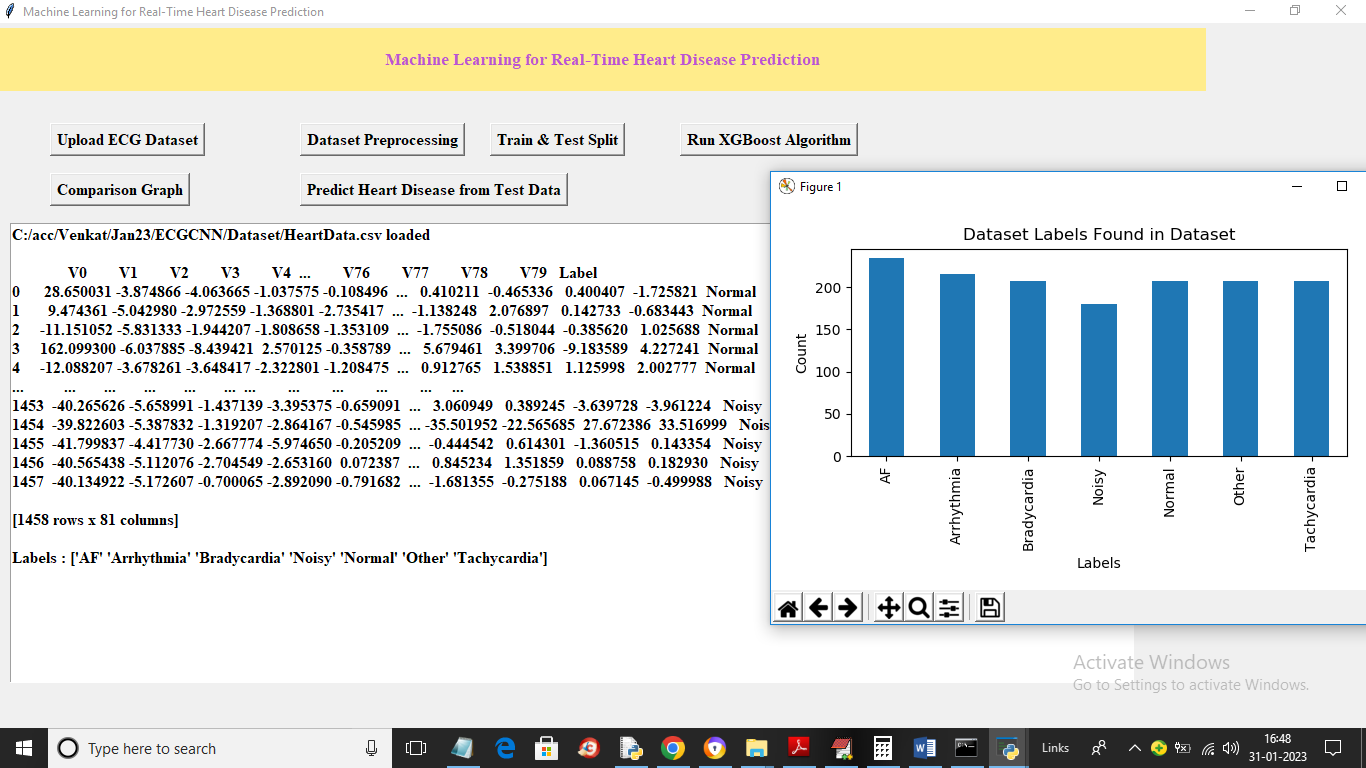
To run project double click on run.bat file to get below screen



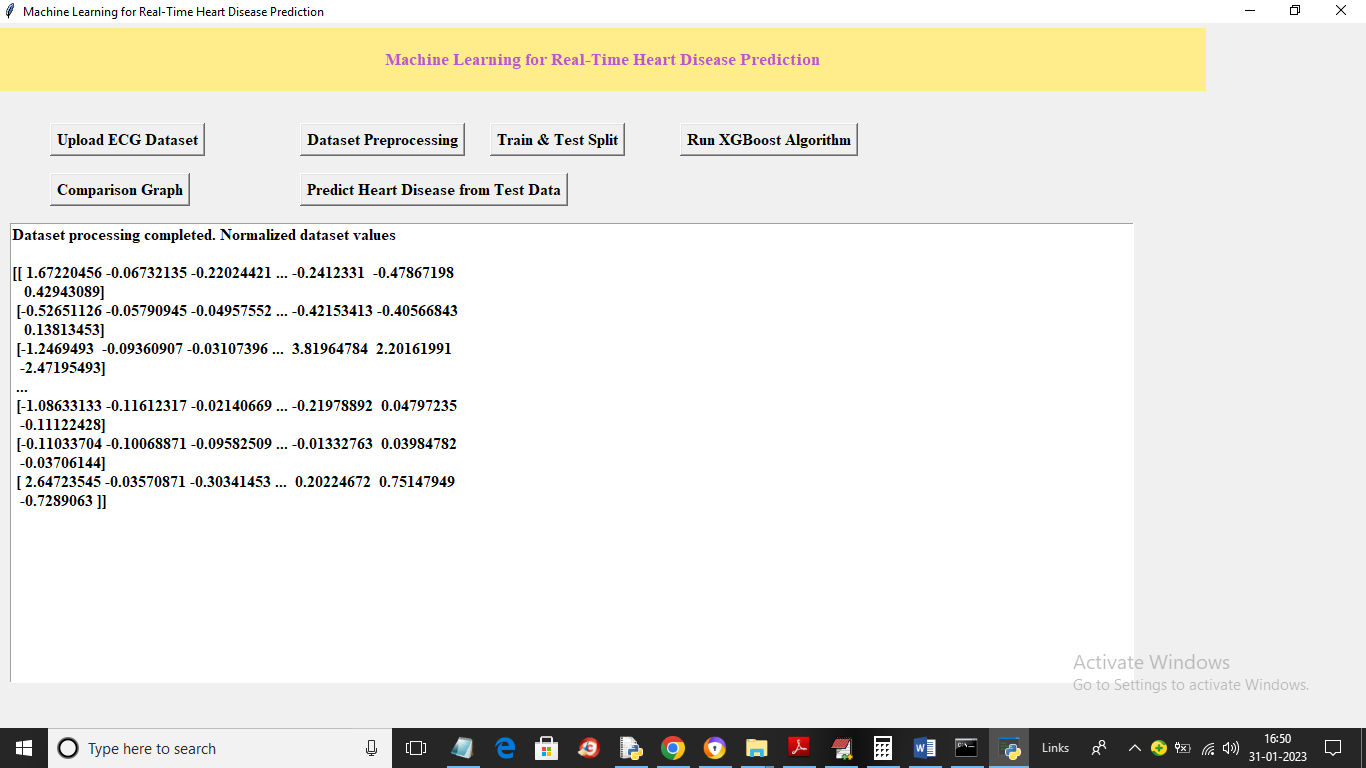
In above screen click on ‘Upload ECG Dataset’ button to upload dataset and get below output



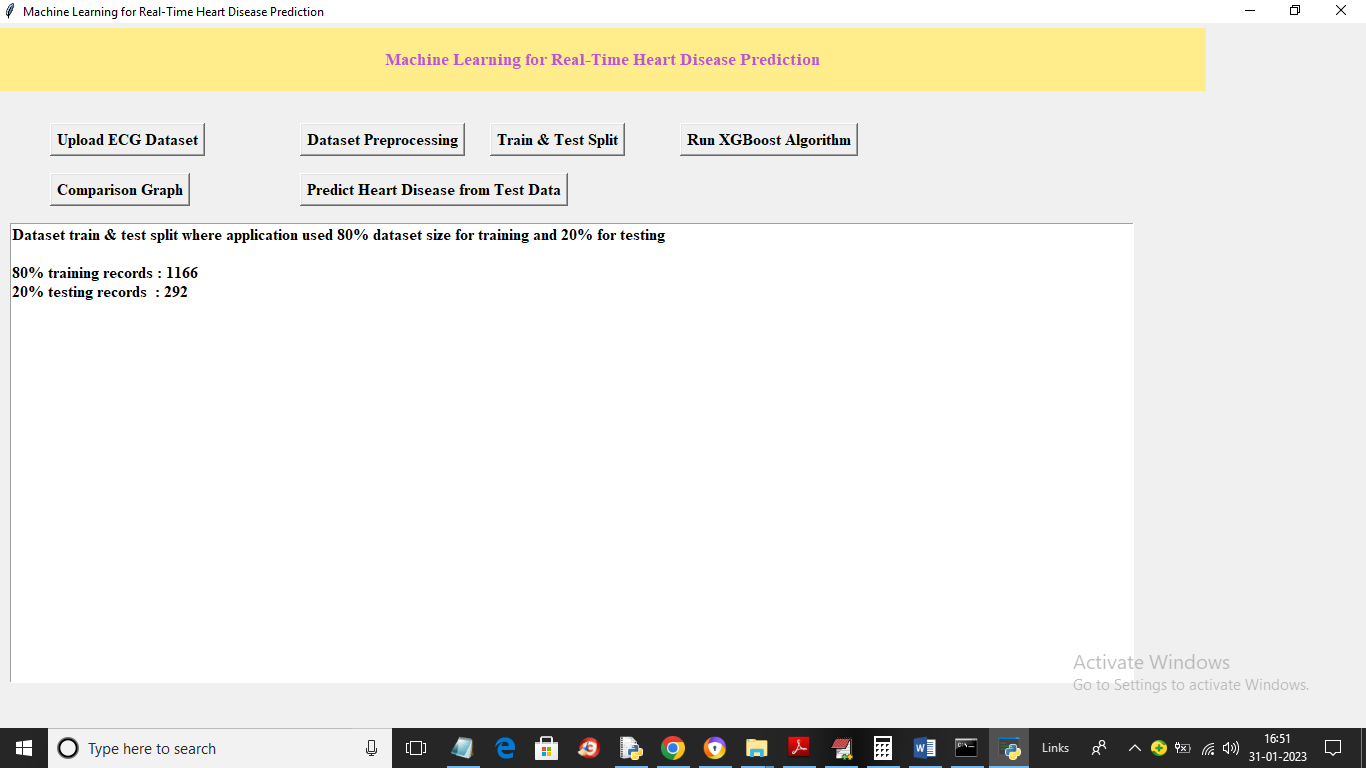
In above screen selecting and uploading ‘HeartData.csv’ file and then click on ‘Open’ button to load dataset and get below output



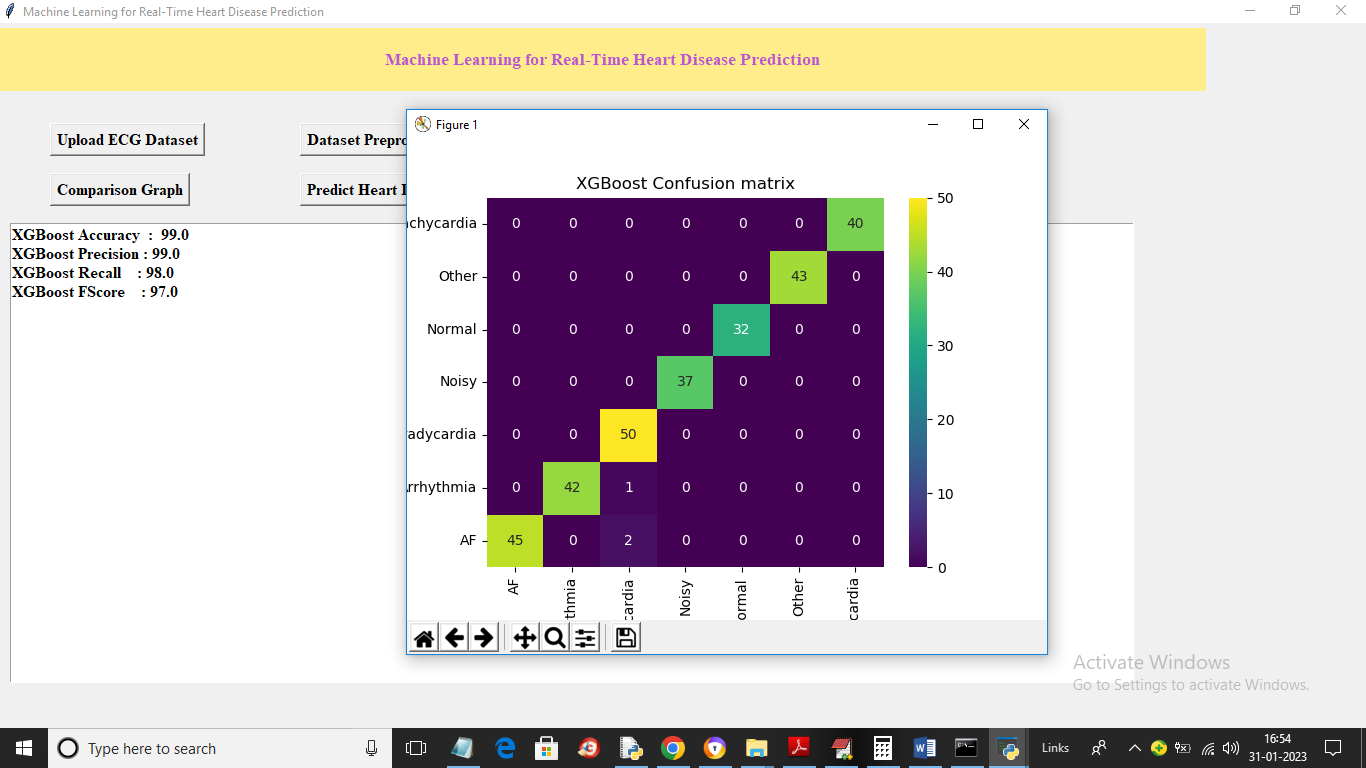
In above screen dataset loaded and we can see dataset contains numeric and non-numeric data but machine learning algorithm will take only numeric data so we need to preprocess data to convert to numeric and in above graph x-axis represents heart disease name and y-axis represents count of that disease records found in dataset. Now close above graph and then click on ‘Dataset Preprocessing’ button to process dataset and get below output



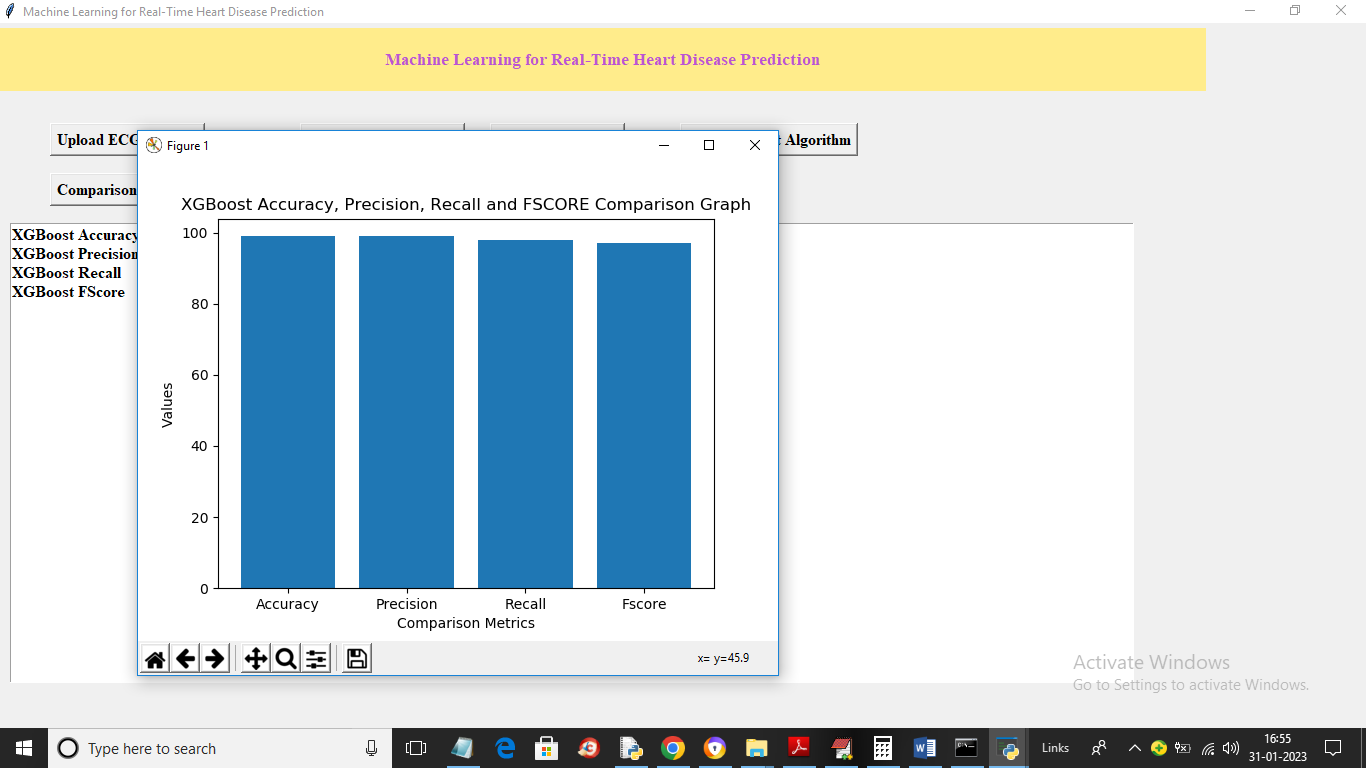
In above screen all values converted to numeric format and now click on ‘Train & Test Split’ button to split dataset into train and test and get below output



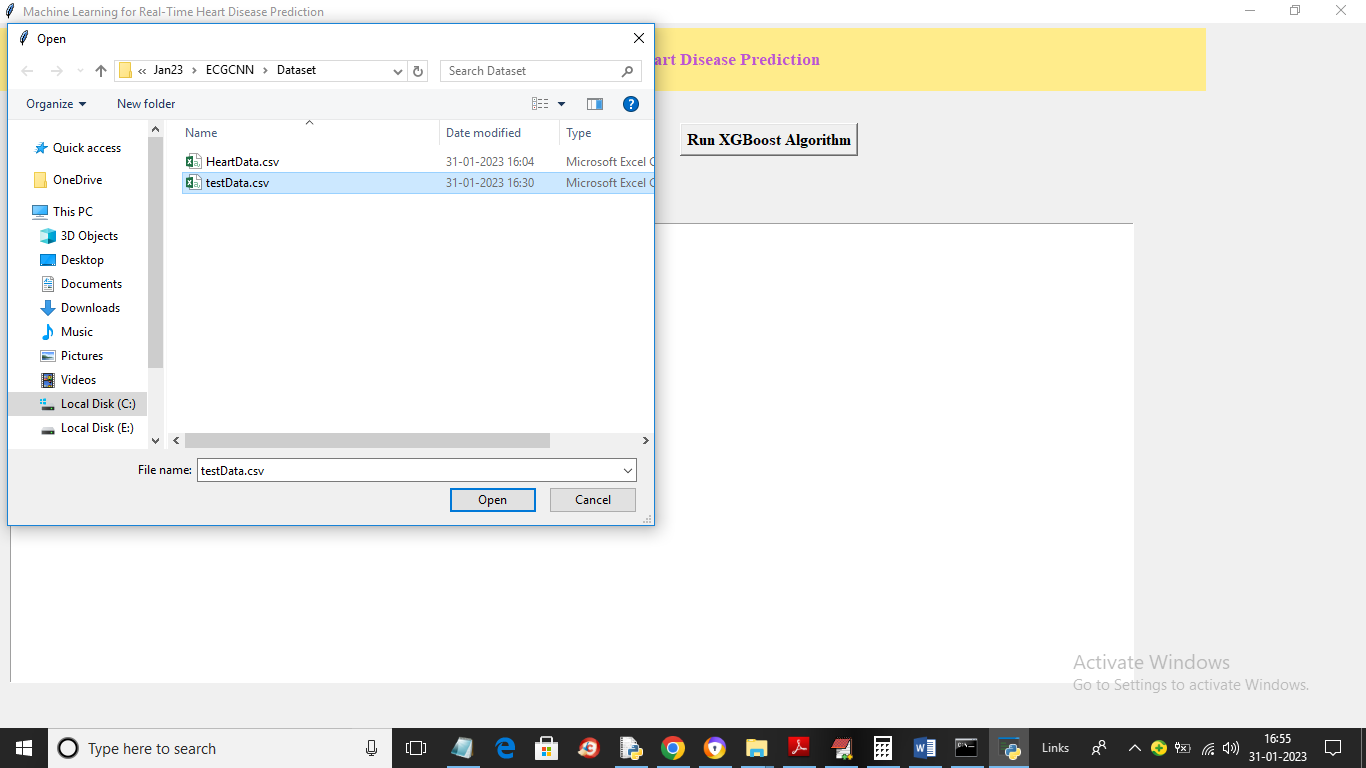
In above screen dataset using 1166 records for training and 292 for testing and now click on ‘Run XGBoost Algorithm’ button to train XGBOOST and get below output



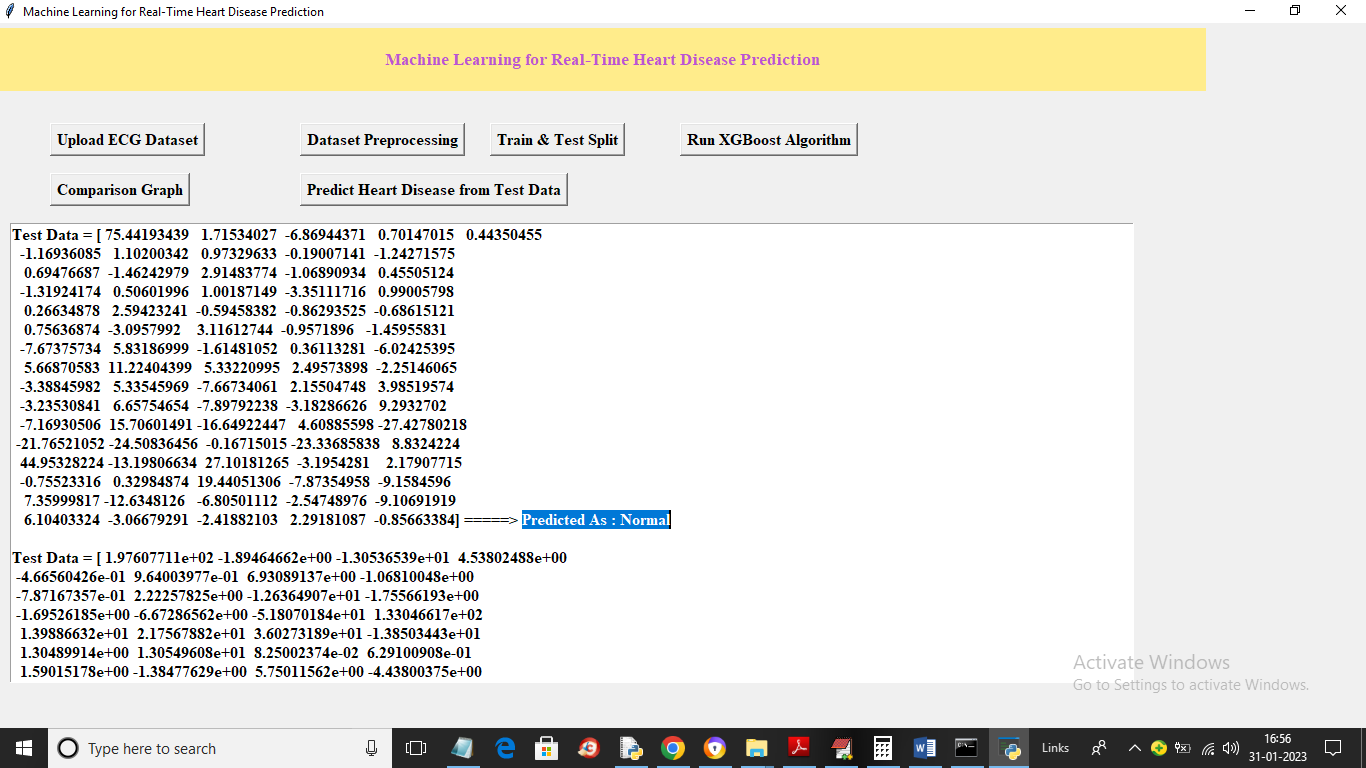
In above screen with XGBOOST we got 99% accuracy and in confusion matrix graph x-axis represents Predicted Labels and y-axis represents True Labels and all different colour boxes represents correct prediction count and blue boxes contains incorrect prediction count which are very few. Now close above graph and then click on ‘Comparison Graph’ button to get below graph



In above graph x-axis represents XGBOOST metrics and y-axis represents values which are closer to 100%. Now click on ‘Predict Heart Disease from Test Data’ button to upload test data and get below prediction output



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



In above screen in square bracket we can see ECH signal test data and after arrow symbol =🡺 we can see predicted output and scroll down above output screen to view other prediction output

