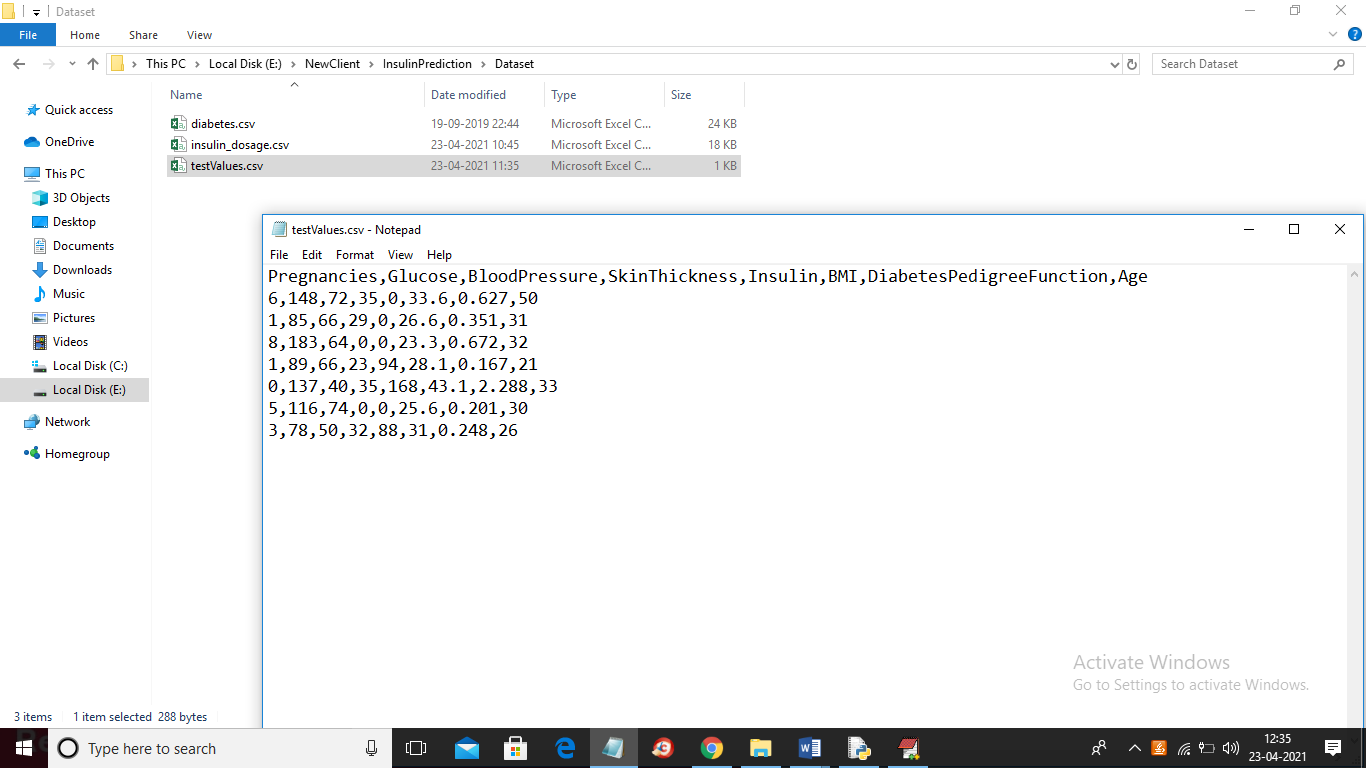
Machine learning Models for diagnosis of the diabetic patient and predicting insulin dosage

In this project we are using Gradient Boosting Classifier to predict diabetes and then using Logistic Regression algorithm to predict insulin dosage in diabetic detected patients. To implement this project we are using PIMA diabetes dataset and UCI insulin dosage dataset. We are training both algorithms with above mention dataset and once after training we will upload test dataset with no class label and then Gradient Boosting will predict presence of diabetes and Logistic Regression will predict insulin dosage if diabetes detected by Gradient Boosting.

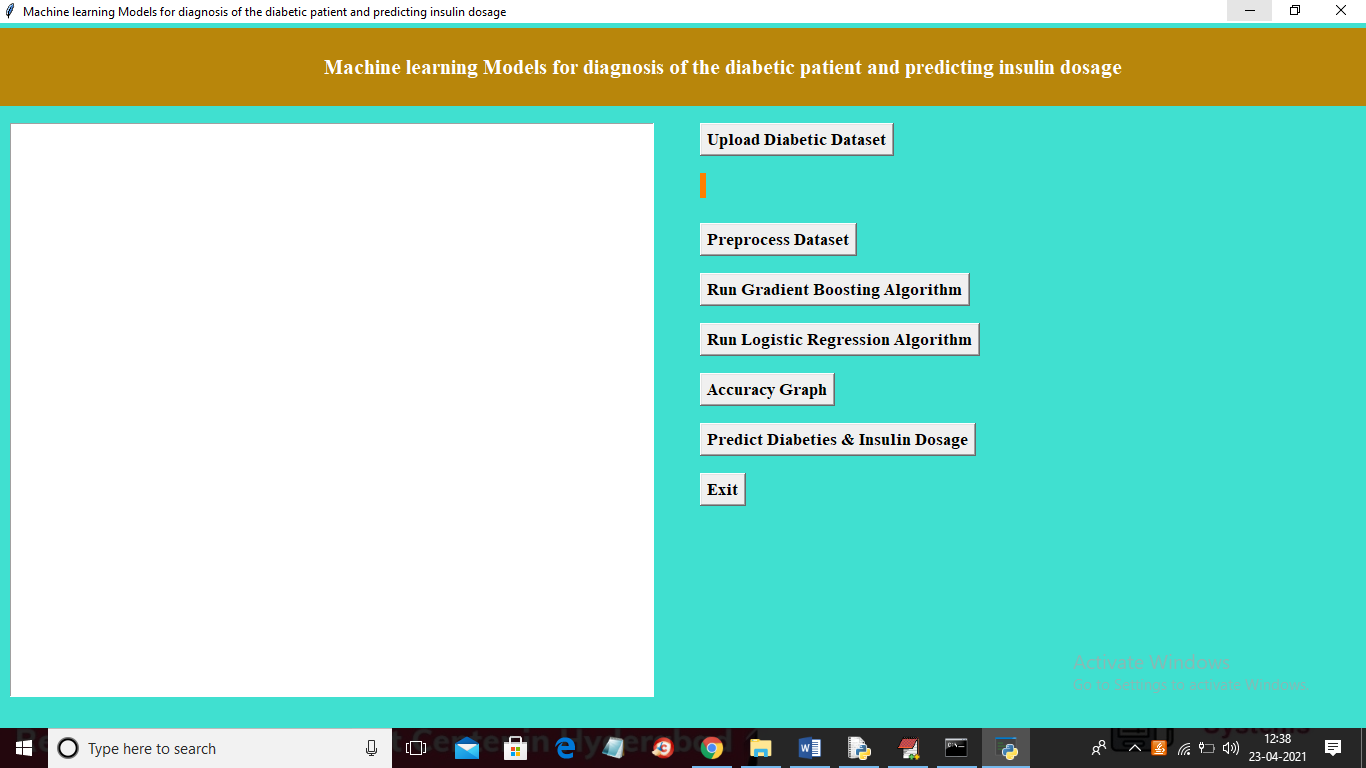
Both dataset available inside Dataset folder and below screen is showing dataset details



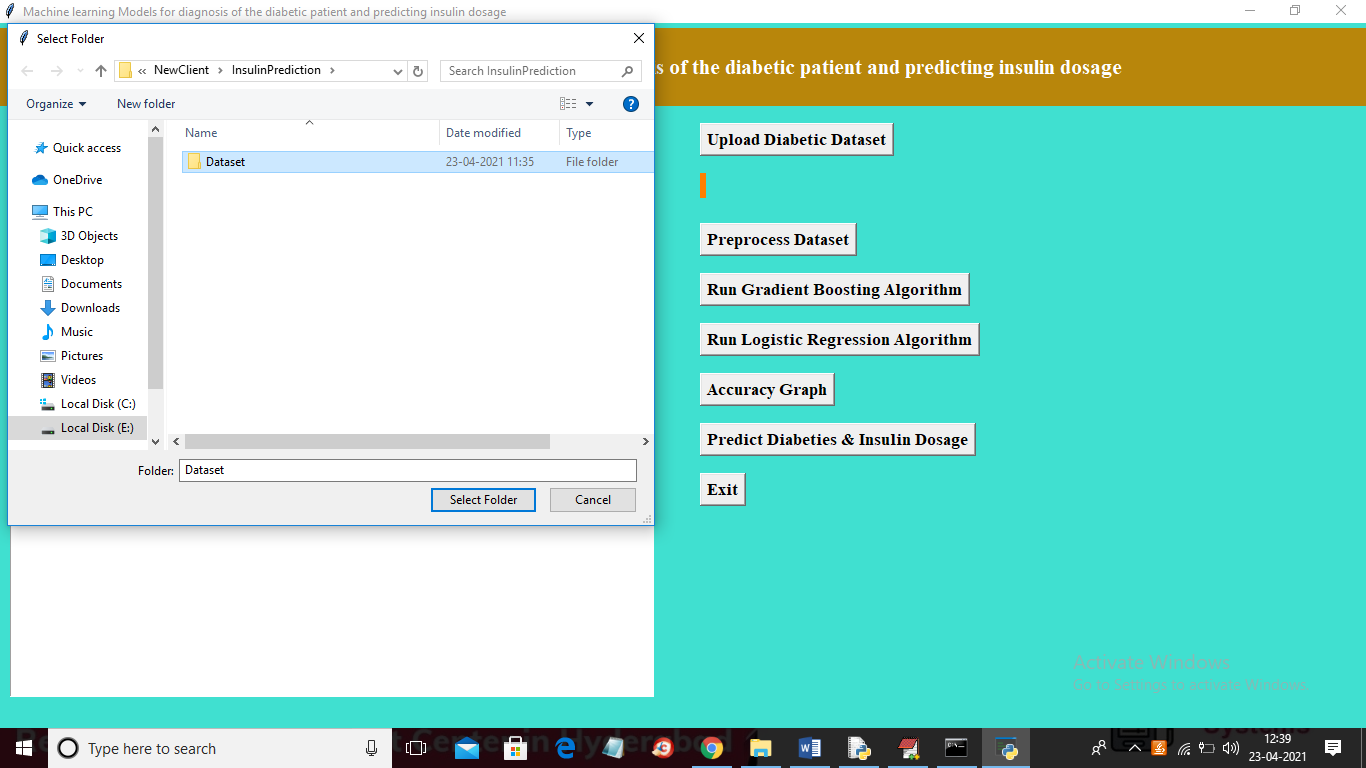
In above screen folder we can see both datasets available and in testValues.csv file we have no class label for diabetes as 0 or 1 where 0 means no diabetes detected and 1 means diabetes detected. When we apply Gradient Boosting algorithm on above test values then Gradient Boosting will predict class label and logistic regression will predict insulin dosage.

SCREEN SHOTS

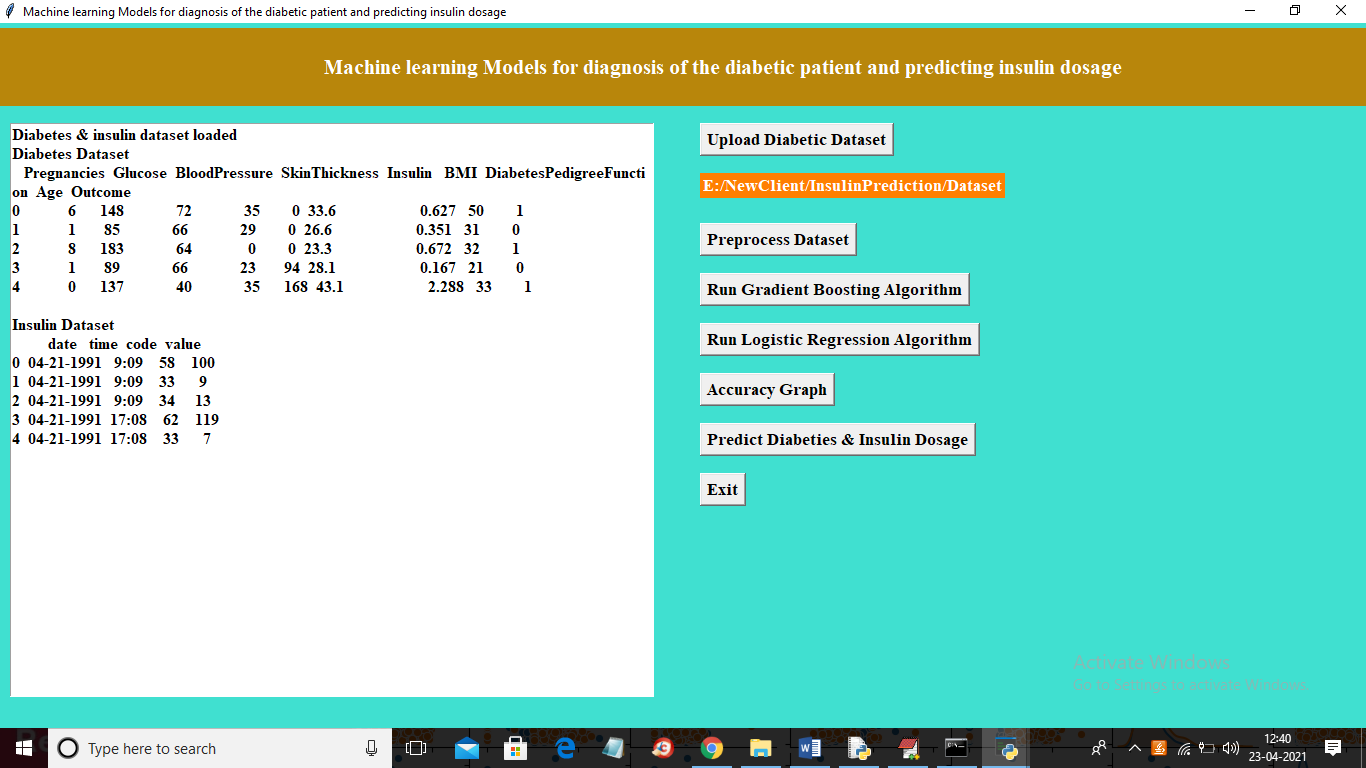
To run project double click on ‘run.bat’ file to get below screen



In above screen click on ‘Upload Diabetic Dataset’ button to upload dataset



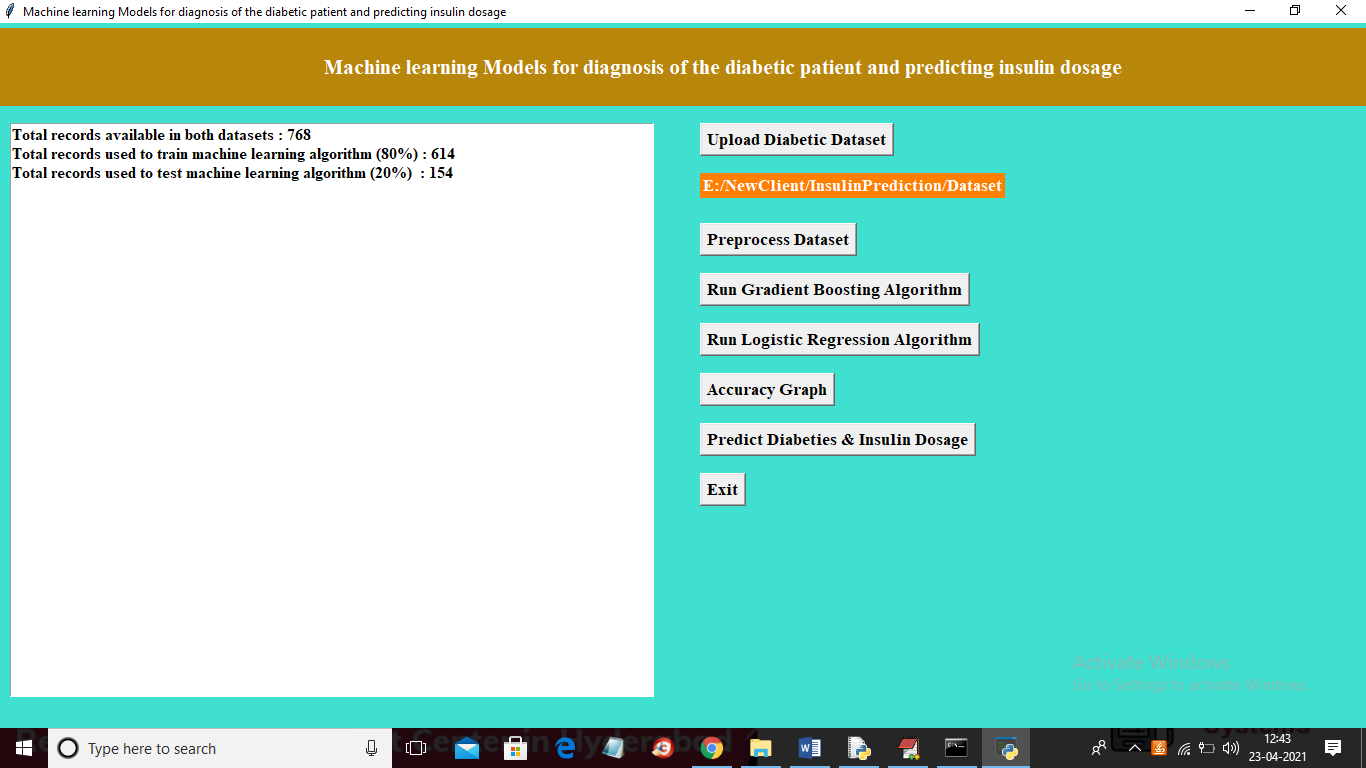
In above screen selecting and uploading entire ‘Dataset’ folder to load both diabetes and insulin dataset and then click on ‘Select Folder’ button to get below screen



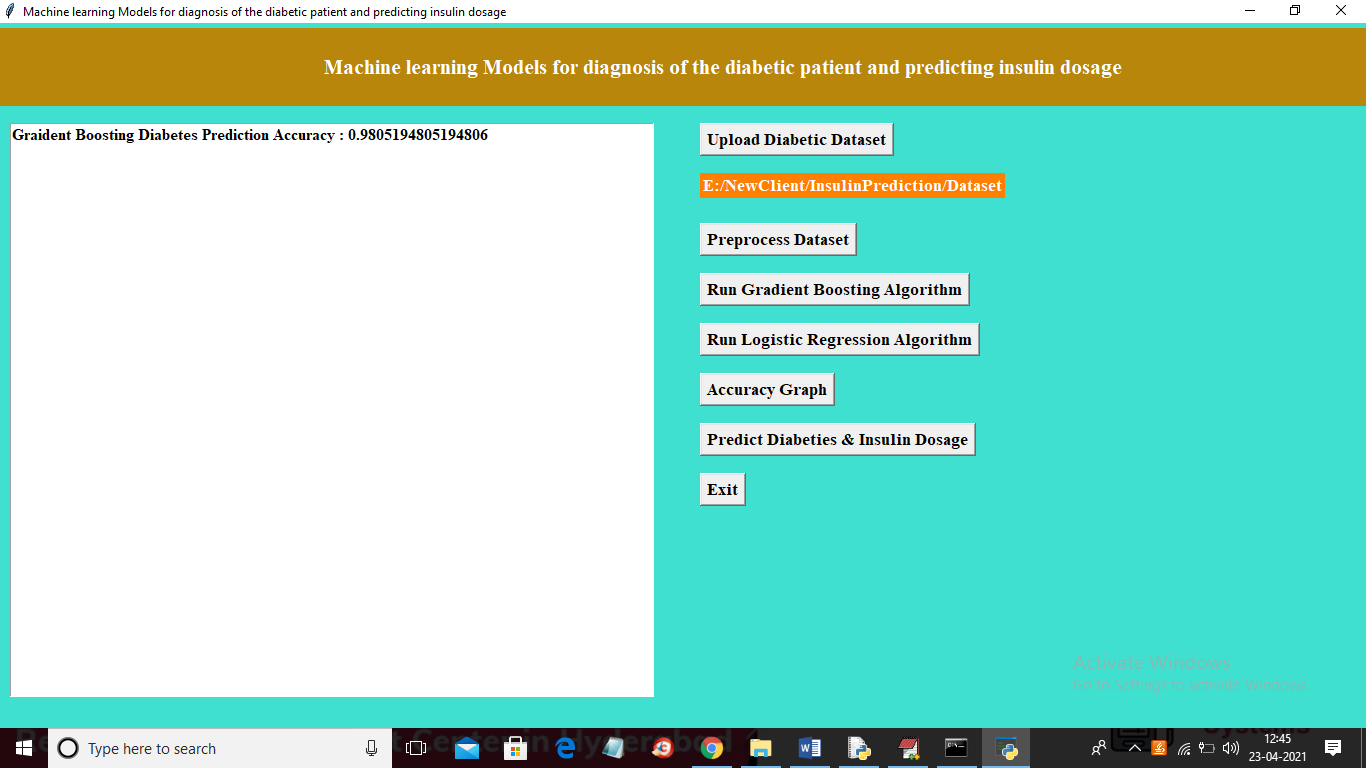
In above screen we can see both datasets loaded and we can see some records from each dataset and will get below graph also



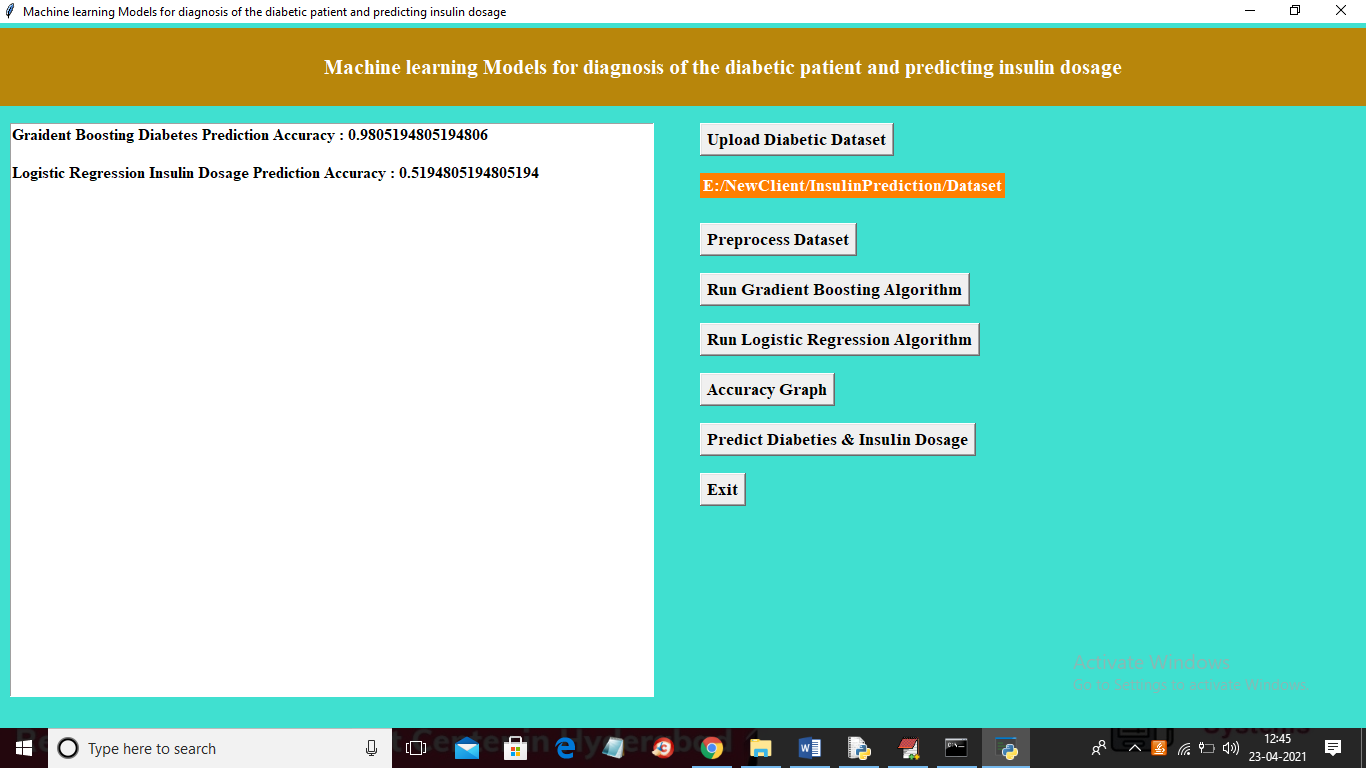
In above graph we can see diabetes for each column where red colour dots indicate presence of diabetes and blue represents no diabetes detected. We are plotting graph for each column value to show with which value diabetes is present and with which value diabetes is not present for example in above graph in first column we are plotting graph for ‘number of Pregnancies’ with ‘presence or no presence of diabetes’ and now close above graph and then click on ‘Preprocess Dataset’ button to remove missing values and to split dataset into train and test



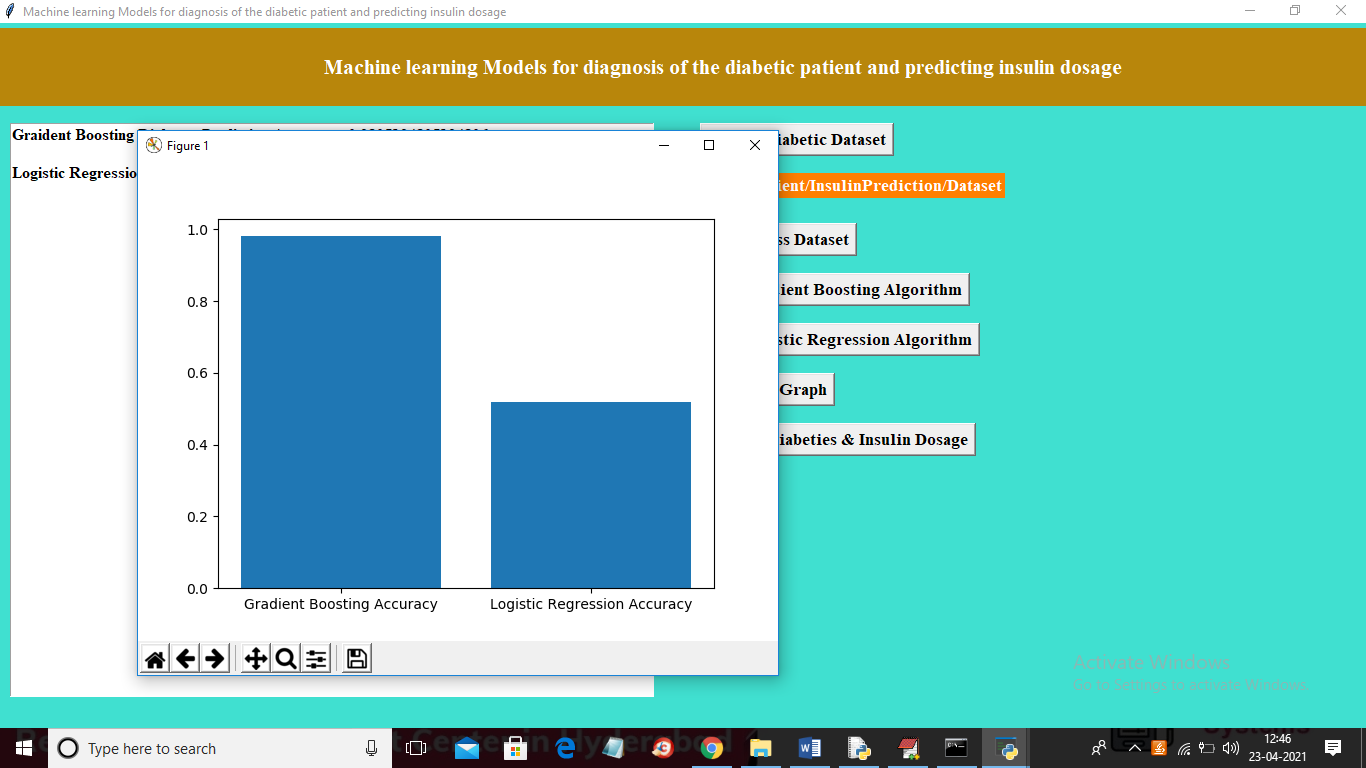
In above screen we can see dataset contains total 768 records and application using 80% records for training and 20% records to test ML accuracy and now dataset is ready and now click on ‘Run Gradient Boosting Algorithm’ button to train gradient boosting with above dataset



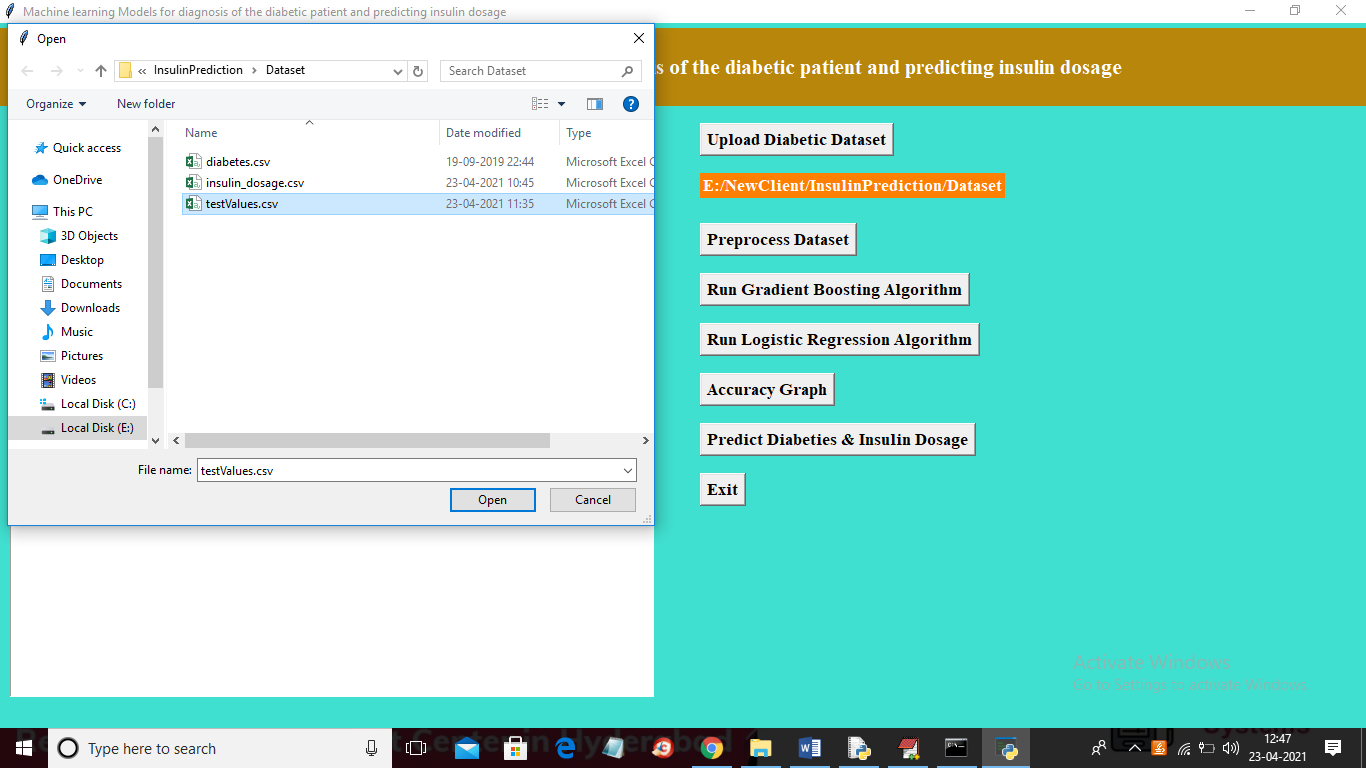
In above screen we got gradient boosting accuracy as 98% and now click on ‘Run Logistic Regression Algorithm’ button to build logistic regression algorithm



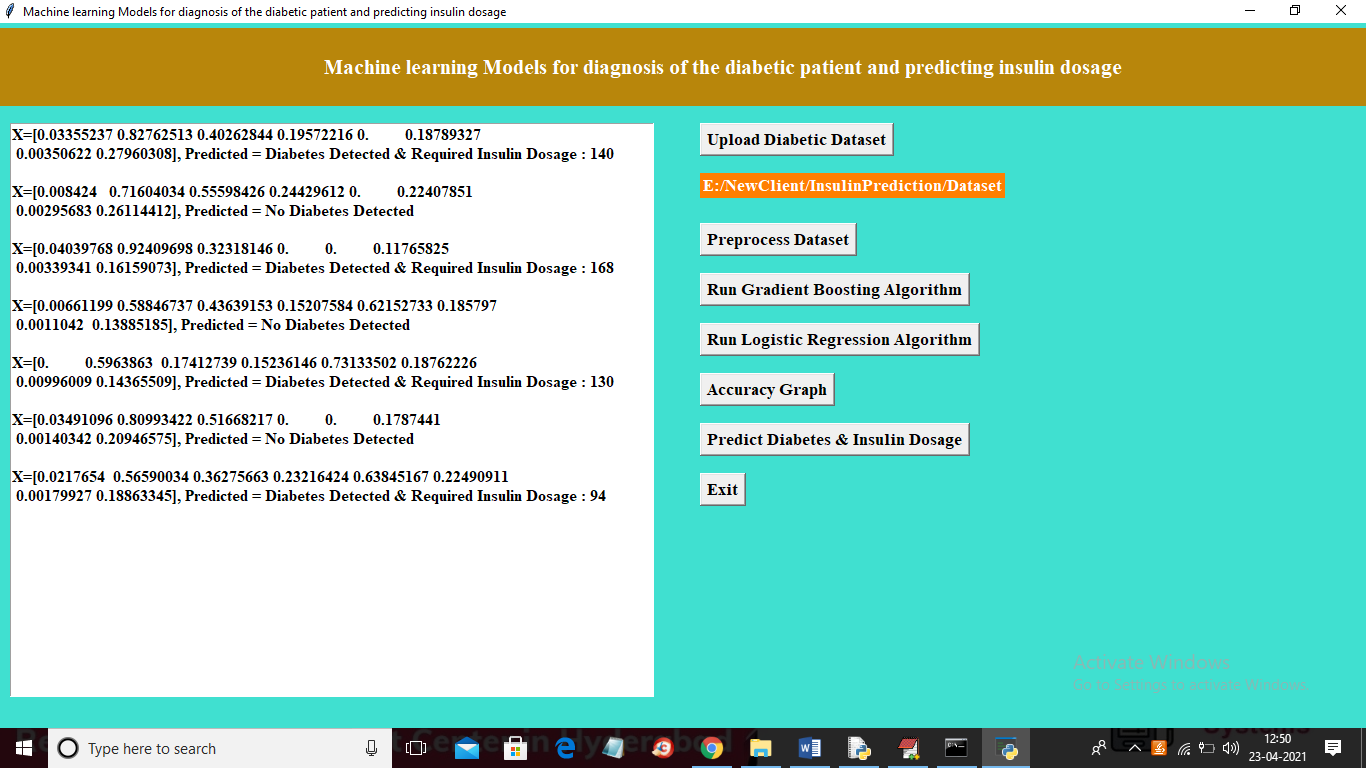
In above screen with logistic regression we got 51% accuracy and now click on ‘Accuracy Graph’ button to get below graph



In above graph x-axis represents algorithm name and y-axis represents accuracy of those algorithms and now close above graph and then click on ‘Predict Diabetes & Insulin Dosage’ button to upload test values and then will get prediction result



In above screen selecting and uploading ‘testValues.csv’ file and then click on ‘Open’ button to get below result



In above screen in square bracket we can see test values and after square bracket we can see predicted result as ‘No Diabetes Detected’ or ‘Diabetes Detected’ and if diabetes detected then logistic regression will predict insulin dosage