ALGORITHM FOR CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING TECHNIQUES

In this project you are asking to use dataset which has low, high or medium risk but on internet we don’t have any such dataset and there is only one European Credit Card dataset is there which has only two types of class labels such as NORMAL or FRAUD and we are using same dataset to implement this project.

You can read complete details from below URL about dataset used in this project

Dataset URL: https://www.kaggle.com/mlg-ulb/creditcardfraud

In this project we have designed following modules

1) Upload dataset: using this module we will upload dataset to application

2) Preprocess Dataset: In this module we will read all records and then preprocess them to remove missing values or to drop TIME column which is not require to build machine learning module. After preprocessing we will split dataset into train and test part where application use 80% dataset for training and 20% dataset for testing accuracy of trained prediction model.

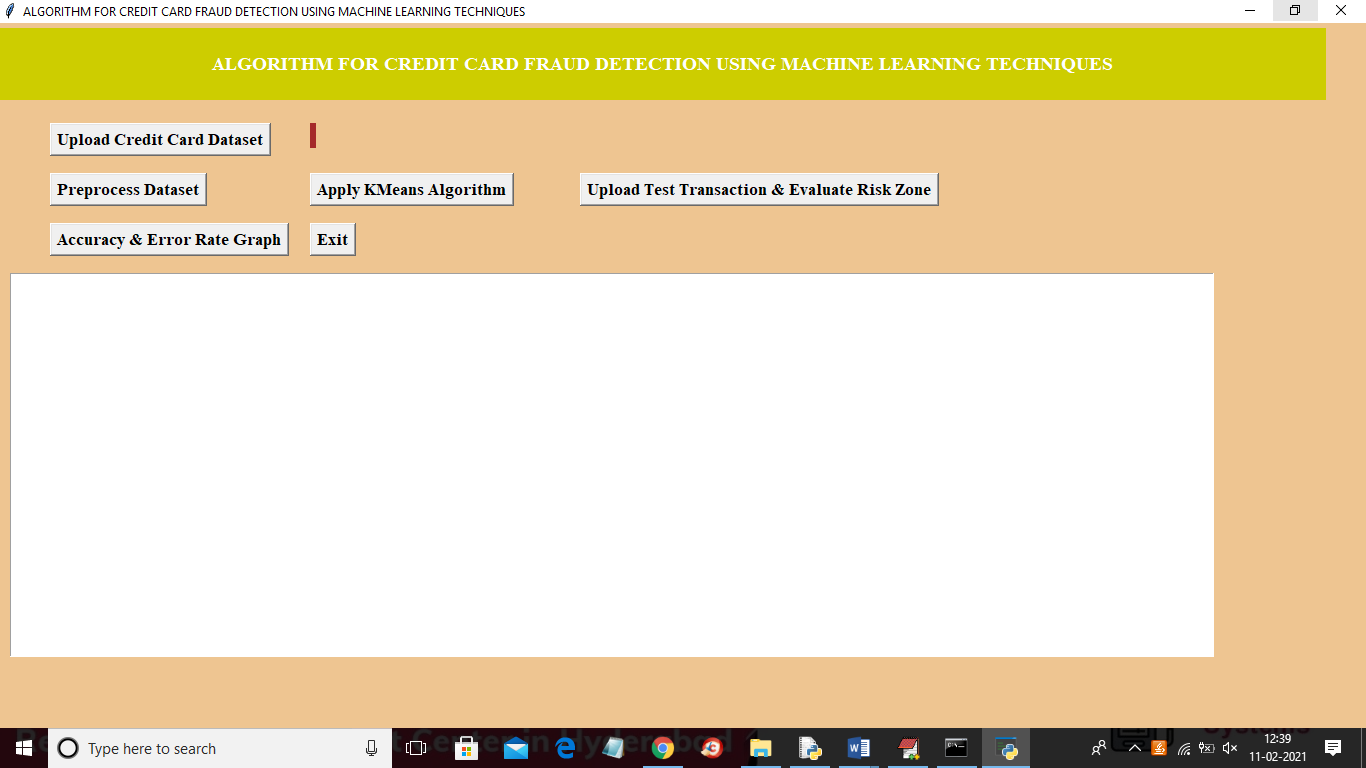
3) runKMEANS: Using this module we will divide train data into 2 clusters such as normal or fraud and then build prediction model.

4) Upload Test Transaction & Evaluate Risk Zone: In this module we will upload new test transactions data and then apply fuzzy logic and KMEANS prediction model to calculate accuracy of transaction as normal or fraud and if accuracy of transaction closer to zero then it will consider as NORMAL transaction and if its accuracy is closer to 1 then it will consider as fraud transaction.

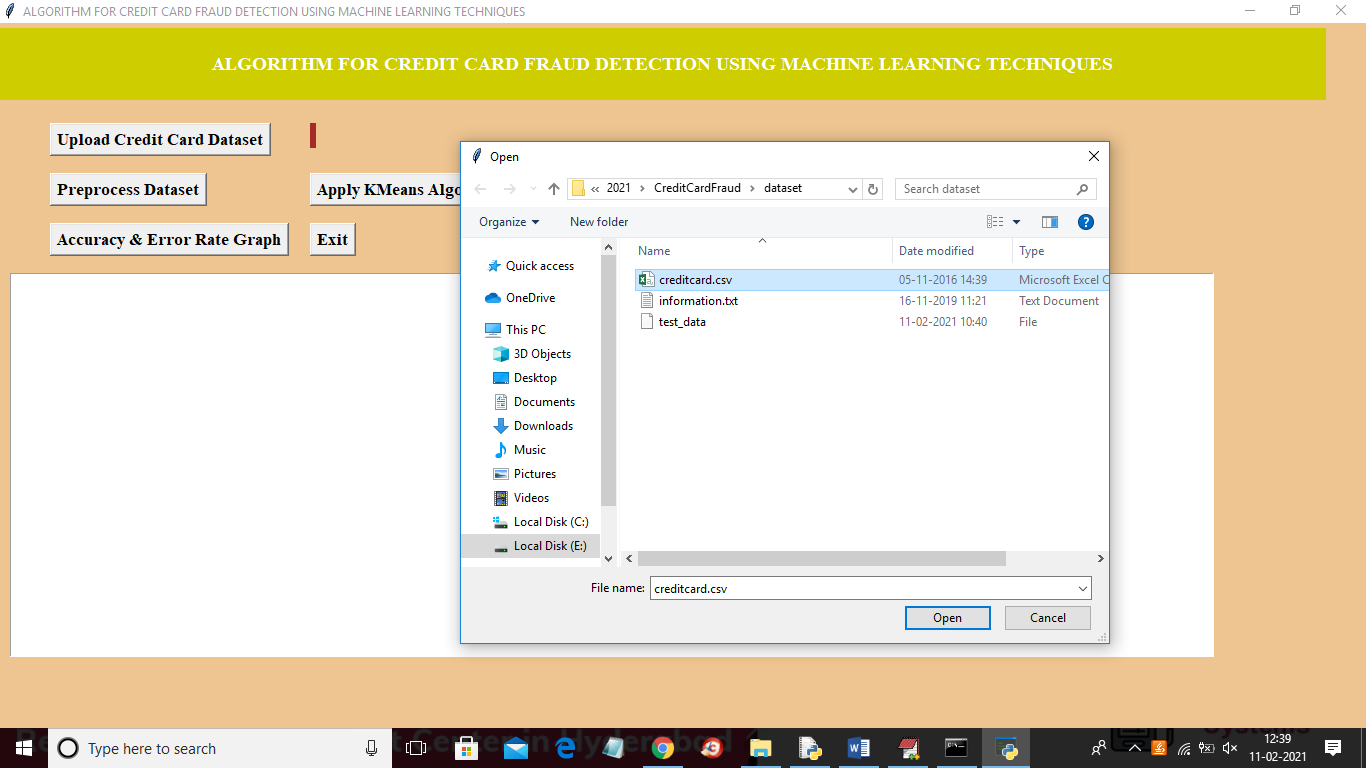
Note: You are asking to get ACK and to perform deny transaction process but we don’t have any real time application or data to perform this step just we build machine learning model using KMEANS and then we upload transaction test data and then predict test transaction signature is closer to NORMAL or FRAUD.

SCREEN SHOTS

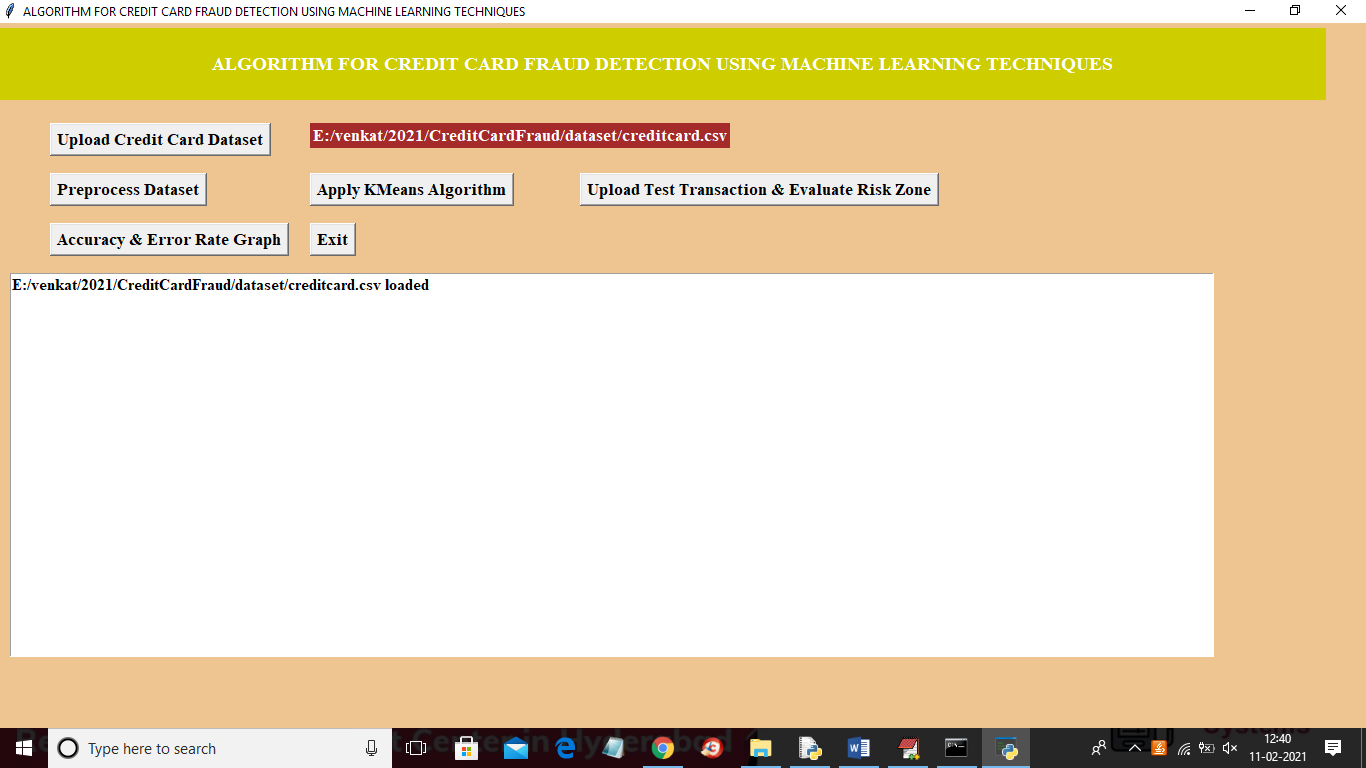
To run project double click on ‘run.bat’ file to get below screen



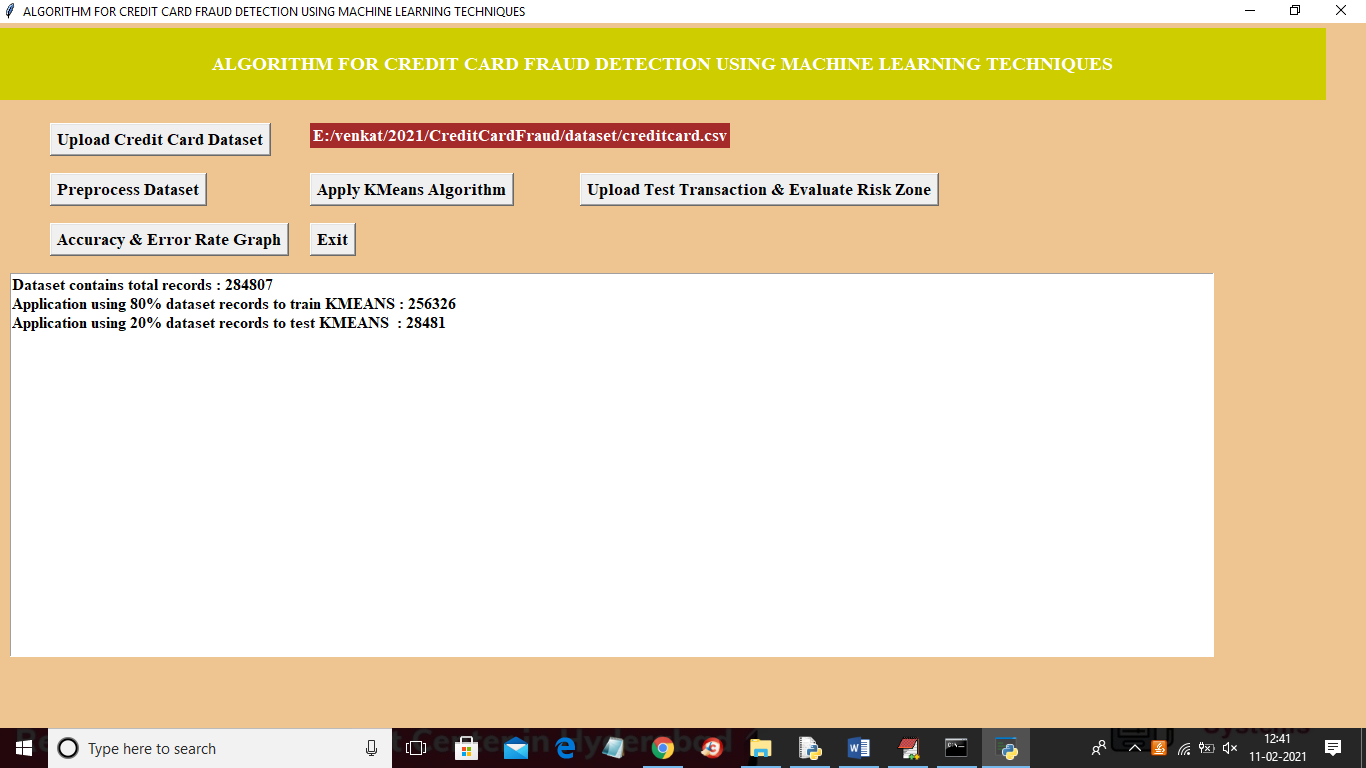
In above screen click on ‘Upload Credit Card Dataset’ button and then load dataset



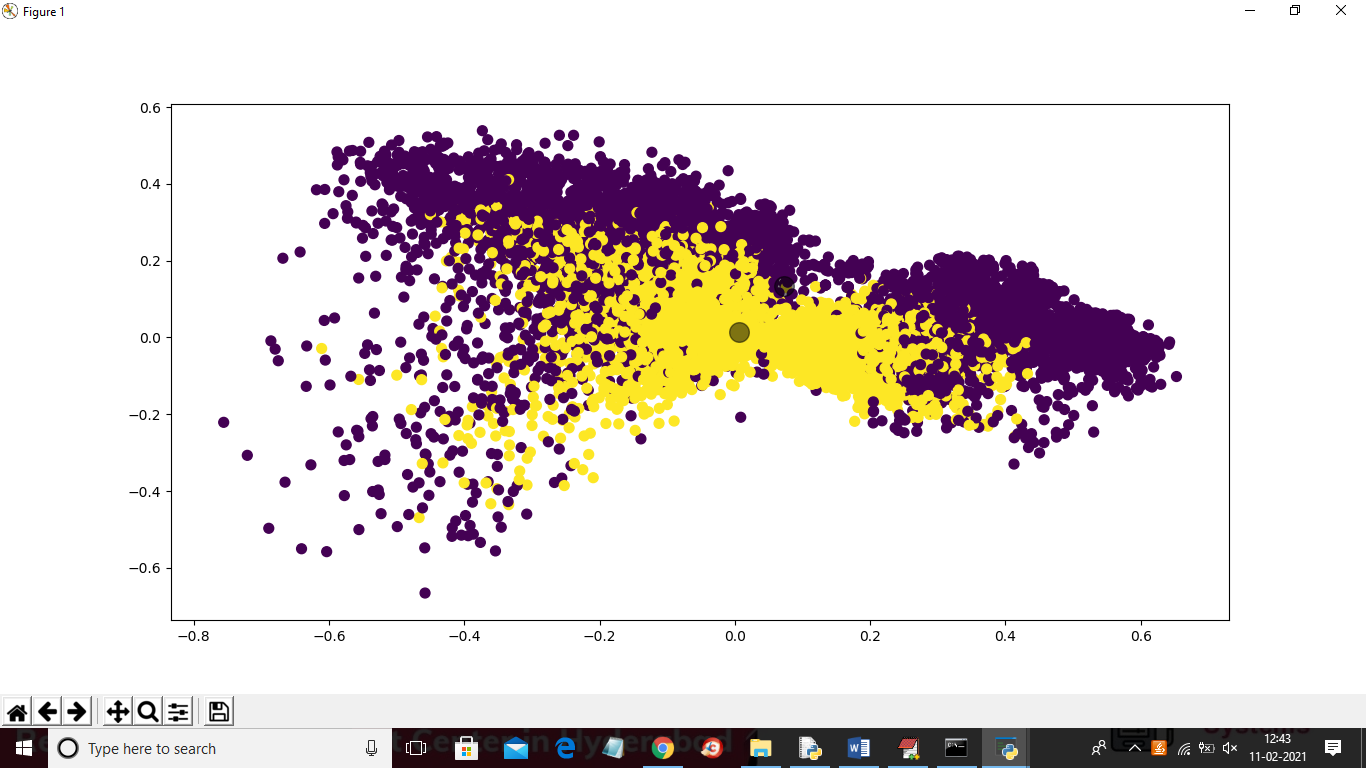
In above screen selecting and uploading ‘creditcard.csv’ file and then click on ‘Open’ button to load dataset and to get below screen



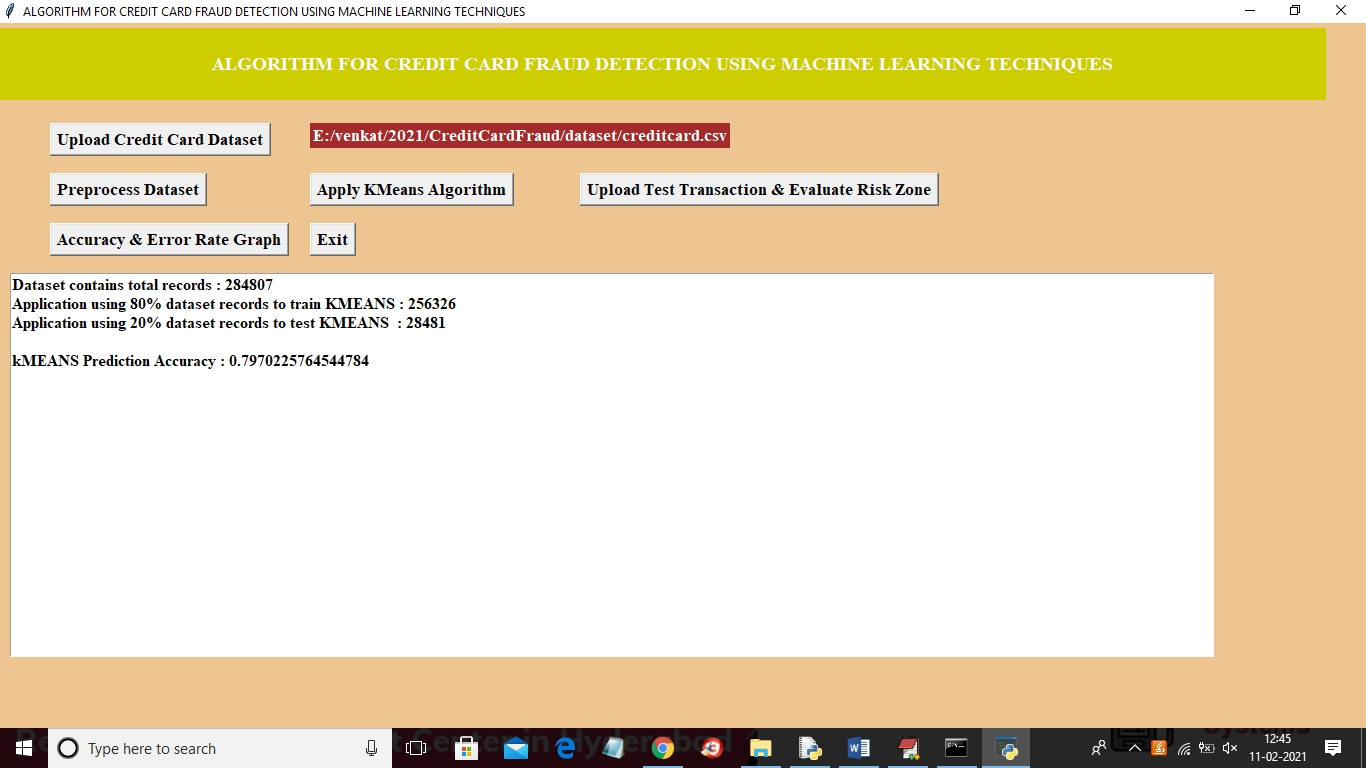
In above screen dataset loaded and now click on ‘Preprocess Dataset’ button to remove missing values and to remove transaction TIME column and the split dataset into train and test part



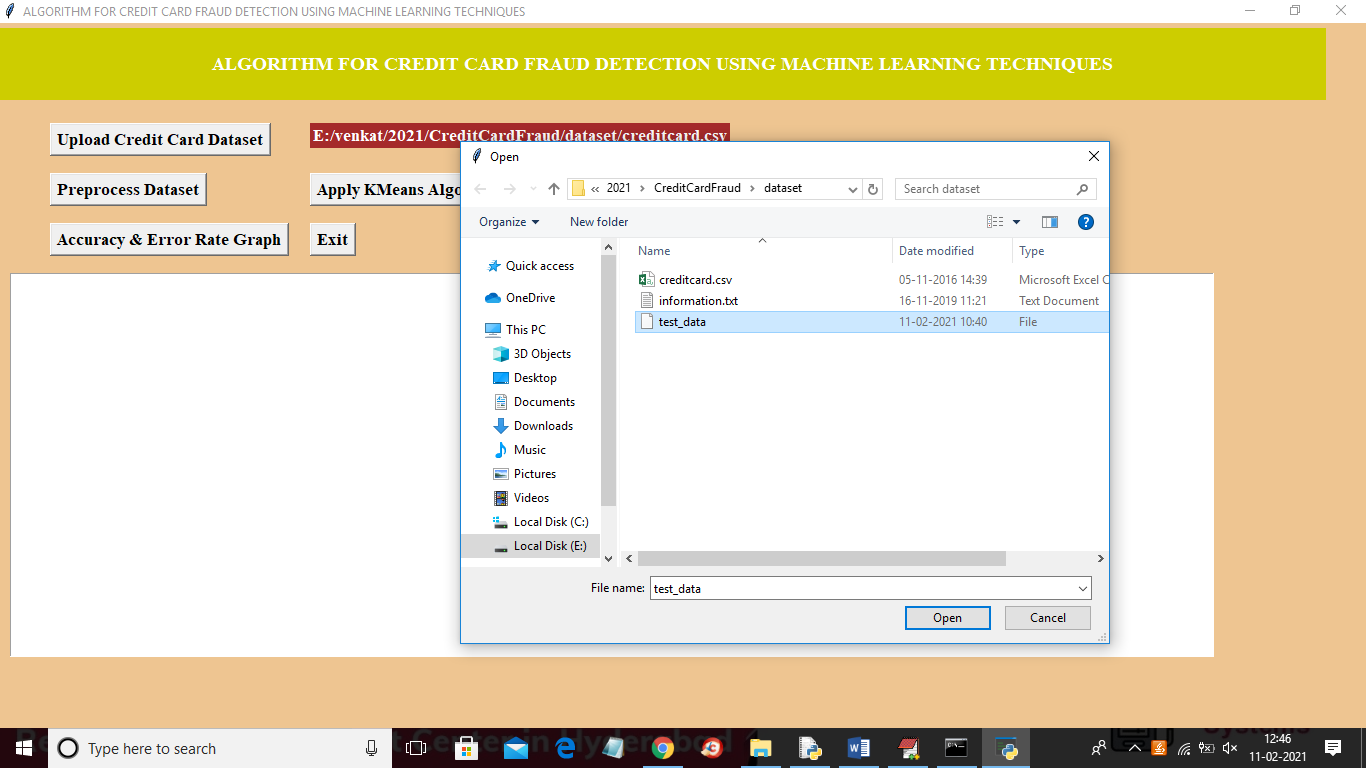
In above screen we can see dataset contains 284807 records and then application using 90% (256326 records) for training and 28481 for testing and now both train and test data is ready and now click on ‘Apply KMeans Algorithm’ button to divide data into NORMAL and FRAUD and then build prediction model and to get below screen



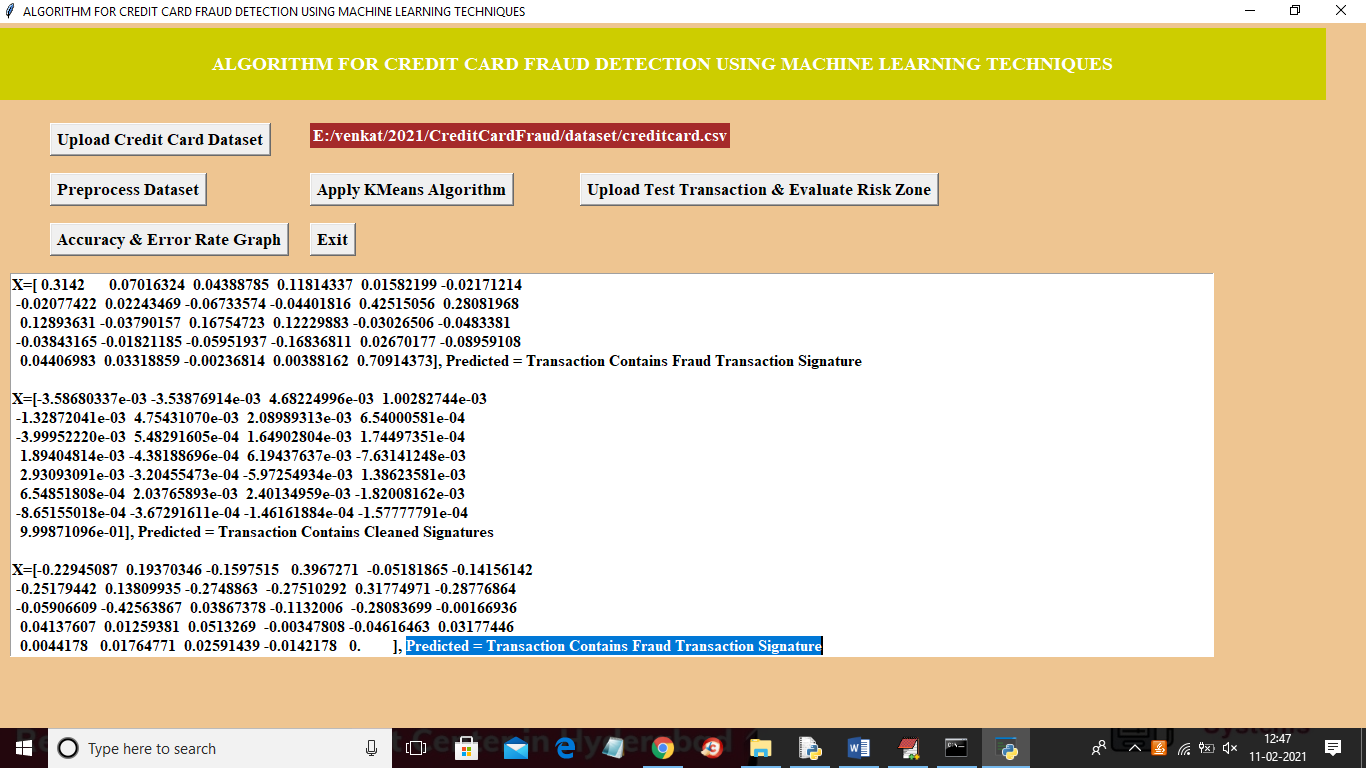
In above graph back colour dots are the normal transaction and yellow colour dots are the fraud transaction which is generated from KMEANS two clusters and in below screen we can see KMEANS accuracy



In above screen KMEANS prediction accuracy is 0.79% and now model is ready and now click on ‘Upload Test Transaction & Evaluate Risk Zone’ button to upload test data and then application apply fuzzy logic to evaluate test data signature into normal or fraud



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



In above screen data inside square brackets are the test data or transaction signature and after square bracket we can see prediction result as transaction contains normal or fraud signature.

Similarly you can upload other transaction signature and perform prediction