Project : Credit Card Fraud Detection

Problem Statement:

1. A credit card is a small thin plastic or fiber card that incorporates information about the person such as a picture or signature and the person's name on it to charge purchases and services to his linked account. Charges are debited regularly. Nowadays, card data is read by ATMs, swiping machines, store readers, banks and online transactions.
2. Each card has a unique card number which is very important. Its security mainly relies on the physical security of the card and also the privacy of the credit card number. There is a rapid growth in credit card transactions which has led to substantial growth in scam cases.
3. Credit card fraud is expanding heavily because fraud financial loss is increasing drastically. Multiple data mining and statistical techniques are used to catch fraud. Therefore the detection of fraud using efficient and secured methods are very important.

Tasks to be performed:

1. Load the dataset using the pandas module.
2. Perform missing value analysis on the dataset.
3. From the dataset, calculate the number of genuine transactions, number of fraud transactions and the percentage of fraud transactions.
4. Using the visualization module, visualize the genuine and fraudulent transactions using a bar graph.
5. Using the Standard Scaler module, normalize the amount column and store the new values in the NormalizedAmount column.
6. Split the dataset in train and test set and have a 70:30 split ratio for the model.
7. Now use a decision tree and random forest model for training on top of the train set.
8. Compare the predictions of both models using predict().
9. Compare the accuracy of both models using score().
10. Check the performance matrix of both models and compare which model is having the highest performance.

