**Project Name:** Credit Card Fraud Detection

**Github Link:** https://github.com/projectsforstudents2022/Credit\_Card\_Fraud\_Detection.git

**Why was this project created?**

Credit card fraud is the word used to describe the unlawful use of payment cards, such as debit or credit cards, to purchase goods or services. From unsafe websites, hackers or scammers may be able to get their hands on the card's private information. Everybody involved in the process loses when a fraudster steals someone's credit or debit card, from the person whose private information has been exposed to the organizations who issue the credit card and the retailer who is completing the transaction with a purchase. Because of this, it is crucial to spot fraudulent transactions right away.

**What problem is it solving?**

The goal here is to minimize inaccurate fraud categories while detecting 100% of the fraudulent transactions. A classic example of classification is the detection of credit card fraud. We have concentrated on analyzing and preprocessing data sets in this process, as well as deploying several anomaly detection algorithms.

**Entire explanation of project**

* **PROPOSED APPROACH**

We are importing the datasets that include credit card transaction data. We will investigate the information that is there in credit card data. Following that, we'll scale our data using the scale () function and show the credit card information using the head and tail functions. This will be applied to our credit card data amount's amount component. Another name for scaling is feature standardization. The data is organized according to a defined range with the use of scaling. We will divide our dataset into a training set and a test set with a split ratio of 80% & 20% after we have standardized the entire dataset.

The first model will be fitted. With logistic regression, we'll start. For estimating the likelihood of a result in a class, such as pass/fail, positive/negative, and in our instance, fraud/not fraud, logistic regression is used. We will draw the ROC curve in order to evaluate the effectiveness of our model. Receiver Optimistic Characteristics, or ROC, is another name for them. In order to do this, we will first load the ROC package before plotting our ROC curve and evaluating its effectiveness.

Algorithm for creating next word prediction model :

**Step 1:** Import Libraries & Load Dataset

**Step 2:** Data Preprocessing

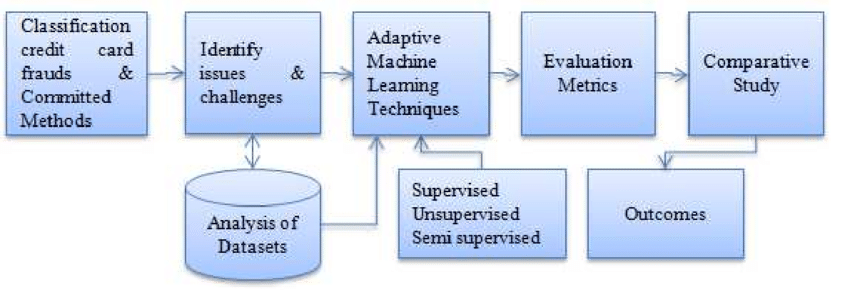
**Step 3:** Label Encoding

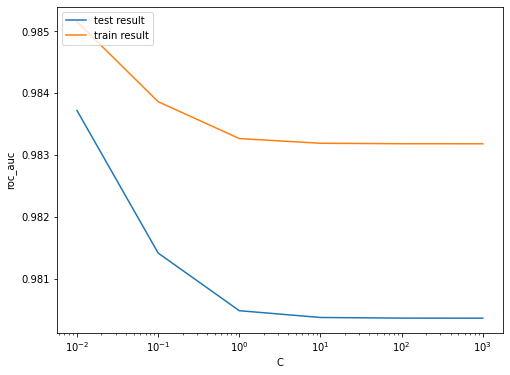
**Step 4:** Build Logistic Regression Classifier

**Step 5:** Train Model

**Step 6:** Testing & Visualization

* **DATA FLOW DIAGRAM**



* **RESULT**
* **CONCLUSION**

Although the algorithm achieves over 99.6% accuracy, when only a tenth of the data set is considered, its precision is only 28%. The precision increases to 33% when the system is fed the whole dataset, though. This high accuracy rate is expected given the vast disparity between the number of transactions that are valid and those that are genuine.