# Credit Card Fraud Detection using Machine Learning

1. The aim of the project is to predict fraudulent credit card transactions using machine learning models.
2. Fraud detection, and its prevention is crucial from the bank's as well as customer's perspective.
3. Banks cannot afford to lose the customers' money to fraudsters.
4. Every fraud is a loss to the bank, and the bank is responsible for fraudulent transactions.

Steps in Implementing Credit Card Fraud Detection;

1. Reading, understanding, and visualizing the data.
2. Preparing the model for modelling.
3. Building the model.
4. Evaluating the model.

Algorithms used for the project;

- Logistic Regression
- K-Nearest Neighbor
- Decision Tree
- Random Forests
- XGBoost

Four Techniques are used for Under sampling or Oversampling of the Data;

- Random Oversampling
- Random Under sampling
- TOMEK Links Under sampling
- Cluster Centroid Under sampling

## Solutions Approach

1. Understanding and exploring data.
2. Data Cleaning
    - Handling missing values
    - Treatment of outliers
3. Exploratory data analysis
    - Univariate Analysis
    - Bivariate Analysis
4. Preparing the data for modelling
    - Check the skewness of the data and mitigate it for analysis.
5. Splitting the data into training data, and testing data.
6. Model Building
    - Train the model with various algorithms such as Logistic Regression, SVM, Decision Trees, XGBoost, etc,.
    - Tuning the hyperparameter with Grid Search Cross Validation, and find the optimal values of the hyperparameter.
7. Model Evaluation
    - The data is heavily imbalanced. Accuracy may not be the correct measure for this model.
    - We have to look for balance between Precision and Recall over accuracy.
    - It is required that we obtain a good ROC score with high TPR and low FPR in order to lower number of misclassifications.