CS513 Final Project Proposal
Project Group 2

Team Members:

Udit Patel - 20006998
Krina Shah - 10459200
Riya Kale - 20008491
Varun Reddy - 10474196

Credit Card Fraud Detection through classification algorithms

1. Problem Statement:

- Unauthorized card operations hit an astonishing amount of 16.7 million victims in 2017.
- It's critical for credit card firms to be able to spot fraudulent credit card transactions so that customers aren't charged for things they didn't buy.
- We will be detecting credit card frauds of customers using classification algorithms based on their information.

2. Dataset description and source of data:

Source: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

The Data set has 284,808 Rows and 31 Columns. It contains only numeric input variables resulting from a PCA transformation. Unfortunately, due to confidentiality issues, the data does not include the original features and more background information about the data. The task is to classify the credit card fraud of the customer. Some algorithms we will be implementing will be Logistics Regression, Decision Tree, Random Forest, and XGBoost.

Variables	Description
Time	Number of seconds elapsed between this transaction and the first transaction in the dataset
V1-V28	The result of a PCA Dimensionality reduction to protect user identities and sensitive features(v1-v28)
Amount	Transaction amount
Class	1 for fraudulent transactions, 0 otherwise

3. Implementation strategy:

- Load the data into Dataframe
- Data Preprocessing
- As data is unbalanced, oversample and downsample it
- Use “Class” as Target variable for classification algorithm
- Divide data as training, testing and validation data
- Apply classification algorithm over oversample, un-sample and downsample data, and compare the results.

4. Team Members name and allocate tasks in the project:

- Udit and Varun will work on Data preprocessing
- Krina and Riya will work on downsampling and oversampling the data
- Udit will work on the logistic regression algorithm
- Riya will work on the decision tree algorithm
- Krina will work on Random forest algorithm
- Varun will work on XGBoost algorithm