

Project Proposal Title: Fraud Detection in Financial Transactions using Anomaly Detection and PCA

Group Members:

Alfred Okoronkwo (alfred.okoronkwo@gmail.com)

Background and Motivation:

Fraudulent transactions pose a significant threat to financial institutions and individuals, leading to monetary losses and security concerns. Traditional rule-based systems often fail to detect sophisticated fraud patterns, especially when fraudsters adapt quickly. To address this challenge, we aim to use anomaly detection techniques that can identify unusual patterns in transactional data. By incorporating dimensionality reduction techniques such as Principal Component Analysis (PCA), we can simplify the dataset while retaining the most informative features. This will enhance the performance of anomaly detection models and allow us to better identify fraudulent behavior.

Data:

We plan to use publicly available Credit Card Fraud Detection dataset from Kaggle, which contains anonymized transaction features. We will also generate synthetic transactional logs to simulate different fraud scenarios. The Kaggle dataset is highly imbalanced, with fraudulent transactions making up only a small fraction of the data. We will apply preprocessing steps such as handling imbalance, normalization, and feature engineering to ensure we have quality data for the project.

Scope:

We will start by performing Exploratory Data Analysis (EDA) to understand feature distributions, correlations, and class imbalance. Next, we will apply PCA for dimensionality reduction to focus on the most significant variables. For fraud detection, we will experiment with Isolation Forest, ensemble methods, and precision-recall based evaluation metrics, which are better suited for imbalanced classification problems. This project will allow us to gain hands-on experience with anomaly detection, dimensionality reduction, and performance evaluation on highly imbalanced datasets, all of which are critical skills in financial fraud analytics.