Credit Card Fraud Detection

Slides prepared by Seth Boswell, Oluwaseyi Olayeye, Malli Montano, and Daniel Saulsberry

Overview of Project

- To protect their customers, banks and other financial institutions use rule-based algorithms designed to detect fraudulent transactions.
- Due to the increasing scale and complexity of these fraudulent transactions, companies have turned to more advanced models, such as machine learning models, to help aide in the detection of fraudulent transactions
- In our project, we attempt to identify fraudulent transactions from a large dataset, which contains 492 fraudulent transaction out of 284, 807 total transactions

Imbalanced Dataset

- Fraud cases only account for 0.172% of total transactions in this dataset, making it highly imbalanced.
- To adjust for the imbalance, we employ the following techniques:
  - Oversampling
  - Undersampling
  - A cross-validation framework for our logistic regression, random forest, support vector machine, and k-means models.
  - Recall and accuracy are used to determine the best classifier model