Detecting Attempted Credit Card Fraud

By George Bennett

Image source (https://en.wikipedia.org/wiki/Credit_card)

Problem Statement

- Classification
- Highly Imbalanced
- Flag Fraud
- Don’t Flag Too Much

- I built classification models
- The dataset was highly imbalanced. %0.17 Fraud
- My goal is to be able to flag as many incoming transactions as possible
- I must avoid having my models flag too many transactions as that would make the model impractical

Business Value

- Prevent As Much Fraud As Possible

Image Source (https://datafloq.com/read/will-analytics-technology-end-credit-card-fraud/2121)

- My model will be able to give lenders added protection against fraud

Methodology

- Transform the data different ways
- NearMiss
- Several machine learning models

- My methodology was to transform the data several ways using different scalers
- I used NearMiss to resample the training data to make it balanced
- I used several machine learning algorithms

Findings

- Support Vector Machine
- Recall : %79
- Precision: %63

- My best model was a support vector machine
- It had a recall score of %79 meaning it would potentially prevent %79 of fraud
- It had a precision score of %63 meaning there would be very few instances of transactions being flagged falsely (less than %1 of transactions)

Future Work

- SMOTE
- machine learning models
- More expansive parameter searches

- SMOTE
- More machine learning models (such as neural networks)
- More expansive parameter searches

Thank you