Anish Almeida

Introduction to Machine Learning

Project Proposal: Credit Card Fraud Detection

A common occurrence in everyday life is the use of one’s credit card. Most individuals make small purchases with their credit cards ranging from small groceries to lunch and dinner receipts. A vast majority of Americans have up to 2 to 3 credit cards. With multiple cards per person, keeping a check on one’s day to day expenses is challenging enough- not to mention time consuming. Most individuals aren’t even aware if there is fraudulent activity on their cards unless abnormally large purchases are made within a relatively close time frame to each other. Only when large transactions are made do the bank or credit card companies alert the individual. While these services are improving, and most credit card companies have dedicated fraud and identity theft prevention teams- it is certainly a growing topic of importance.

The scope of my project is aimed towards understanding how machine learning algorithms are being used to detect fraud. I will be basing my program from existing algorithms such as the Local Outlier Factor. Similar to how a credit card company tracks each purchase- the Local Outlier Factor in my program will make use of each purchase as a data point. A few things to consider additionally would be including not only daily purchases but perhaps a monthly average as well. (Such as if usual monthly expenditures are $1,000 and one month is $1,500). Will still need to work out the finer points and am strongly considering the addition of another machine learning algorithm to perhaps compare and contrast the two and see which yielded better results or if the two were able to come to the same result.

My dataset: <https://www.kaggle.com/mlg-ulb/creditcardfraud>

While it is important to acknowledge that Credit card companies may have a faster response time and the benefit of having a personal representative for fraud detection- it is certainly important and beneficial to explore different programs and algorithms as this can improve existing code and perfect existing programs.