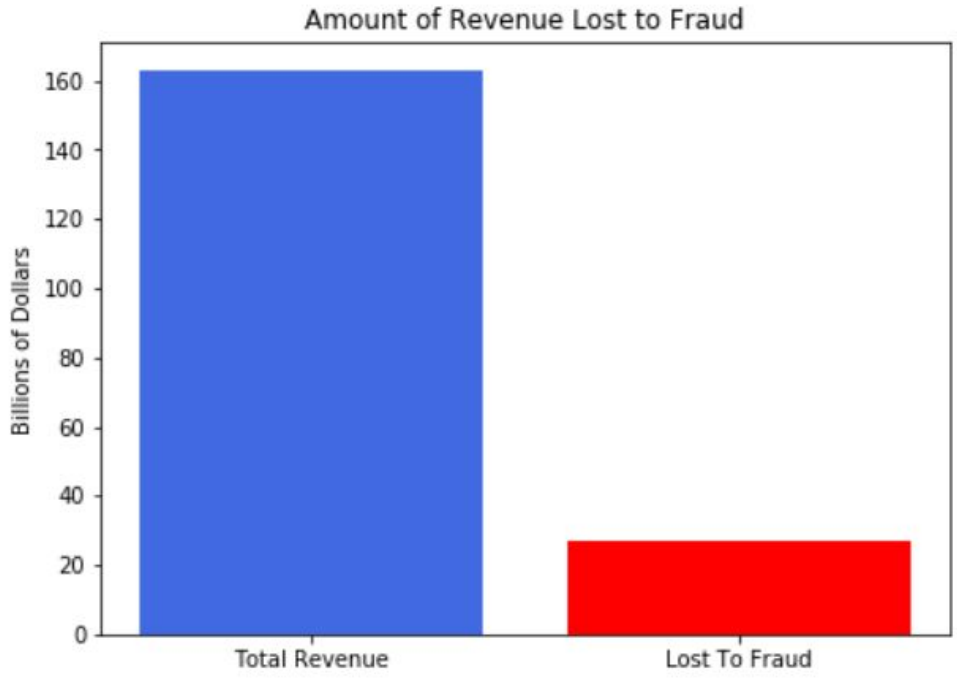
Business Understanding

Credit cards have become extremely common and are used for just about everything. In 2016 credit card lenders earned $163 Billion in revenue.[1] However it is common for credit cards to be used fraudulently. When a cardholder reports misuse of their card, they are legally only liable for at most a small portion of it.[2] The credit card lender ends up recompensing the card holder. In 2018 over $27 Billion was lost by credit card lenders because of fraud.[3] Together that means that about one sixth of the revenue is being lost and it can be prevented.



There are several ways for credit card information to be stolen. Some obvious ways include stealing the physical card or scamming someone out of their information (phishing). A somewhat less thought about method is when criminals implant devices in card readers to steal the information and then sell it anonymously online. Some ways to prevent loss are to educate card holders to the dangers of identity theft, make the physical cards and card readers more resistant to hacking, and to use machine learning models to flag cards for further inspection. The last of these three is what I will be working on in this project.

Given characteristics of the card holder's account, we should be able to gain some sense to the likelihood that the card information is being misused. Then, once our models flag the card, we can decline all transactions and set up a call with the client to work from there to set up a new card and gather more information on the cause of the problem, if it is indeed a case of theft. Even though there are most likely models already in place working on this problem, we can add/update models at the end of the pipeline to lower this large amount of preventable loss.

As for the performance metrics on my model I will use recall and precision. Recall is essentially a percentage of the known frauds that the model is able to flag. Precision is a measure of the inverse of false flags. Meaning that the higher the precision the less often the model is flagging legitimate cards and saying there is a problem when there isn't. These two metrics have an inverse relationship and the model can be tuned to prefer one over the other. Recall is more important because paying off fraudulent transactions will cost us more money than irritating a fraction of a percent of our customers with a false flagging of their account. Not to mention that our customers would be extremely displeased if they had to deal with a fraud case that could have been prevented.

Sources

1
This Is How Credit Card Companies Hauled in $163 Billion in 2016
The Ascent Staff
https://www.fool.com/the-ascent/credit-cards/articles/this-is-how-credit-card-companies-hauled-in-163-bi/

2
Credit Card Fraud
Wikipedia
https://en.wikipedia.org/wiki/Credit_card_fraud#United_States

3
Payment Card Fraud Losses Reach $27.85 Billion
The Nilson Report
https://www.prnewswire.com/news-releases/payment-card-fraud-losses-reach-27-85-billion-300963232.html#:~:text=The%20United%20States%20accounted%20for,There%20was%20some%20good%20news.