Credit card analysis presentation script:

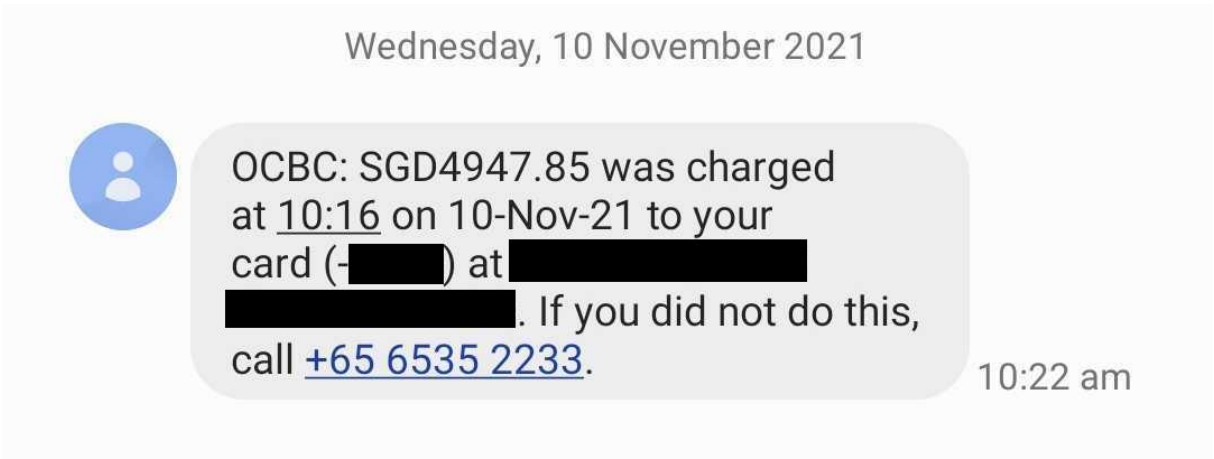
By Omezzine Meysour
Ironhack, 24 Mar 2023

1. story telling

I remember when I was a Kid, my parents used to get calls or an SMS from the bank from time to time asking them about certain transactions if it was done by them or not. At that time, I thought that the bank tellers had to go through each transaction that occurs and check if it was a fraud or not. As I got older, I understood that the bank has a system that can detect this fraudulent transaction.

2. What is credit card fraud, why and how?

A fraudulent transaction is the unauthorized use of an individual's accounts or payment information.
You have probably received an email or a mobile alert asking if a recent transaction is familiar or received a call from a bank staff about an overseas purchase.
Today we are going to look at why and how are we getting alerts like this?



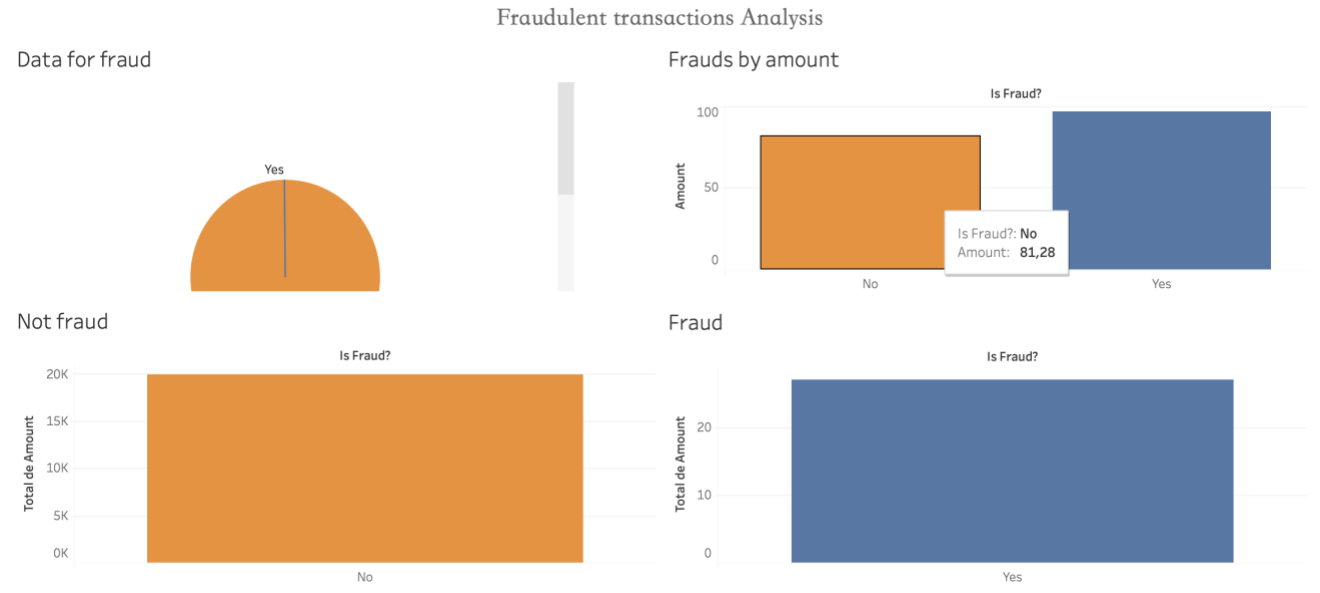
3. data

We will looking at a real dataset of credit card transactions that includes features such as the transaction date,amount,MCC category , merchant type etc that will help us understand how a machine learning model can use these features to predict fraudulent transactions.

4. data analysis

In order to predict a fraudulent transaction, we must look at different transaction features and how it can be a strong indicator of fraud which will help us in making our prediction model.

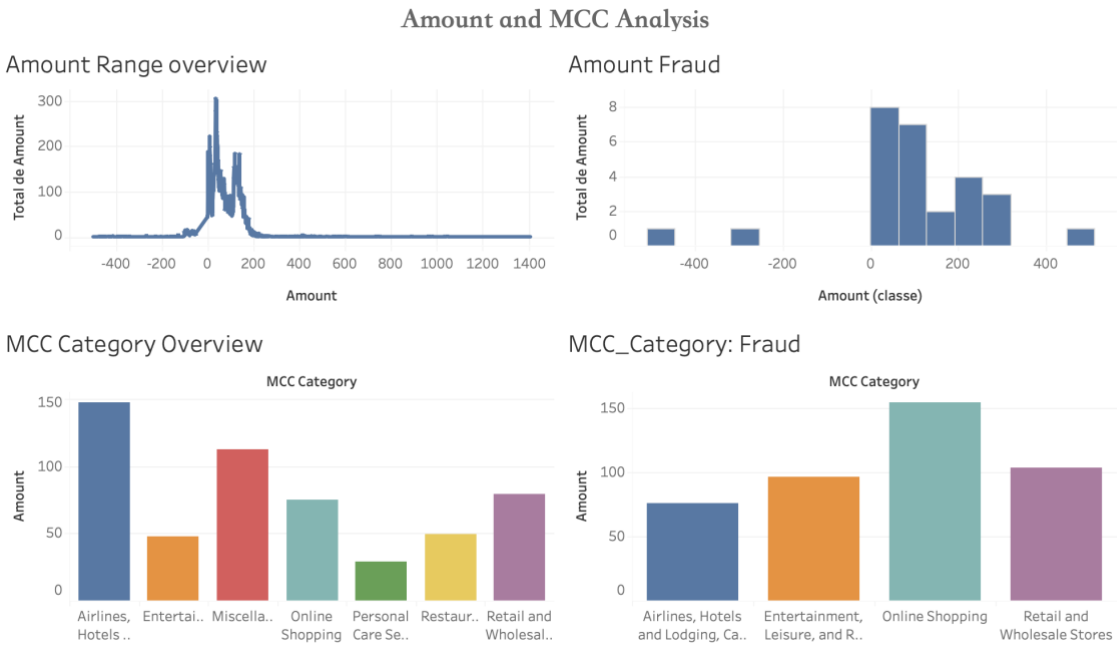
** include first graph**



Firstly looking at our data, we see that only (0.137% of transactions are frauds).
Looking at the average fraudulent amount, the average fraudulent amount is about 80$

let's have a look at the fraudulent transaction amounts and see where it occurred the most :

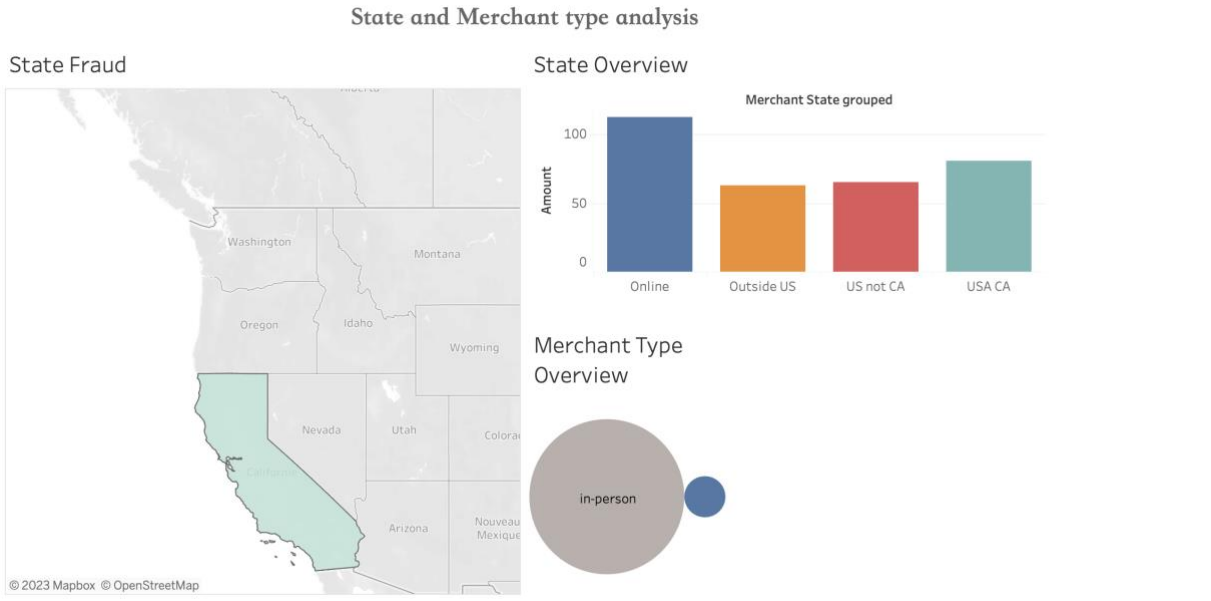
5. Amount analysis



Looking at our graphs, we see that fraudulent transactions amounts ranges from -500 to +500 When a credit card transaction is negative, it means that the transaction has resulted in a credit to the account instead of a charge.
For example, let's say you purchased a $100 item with your credit card and then returned it for a refund. The refund would appear as a negative transaction on your credit card.
This explains This is explained in how most fraudulent transaction happened in online shopping and retail stores.

Now we will compare how fraudulent transactions are spread through different regions :

6. State and merchant Type

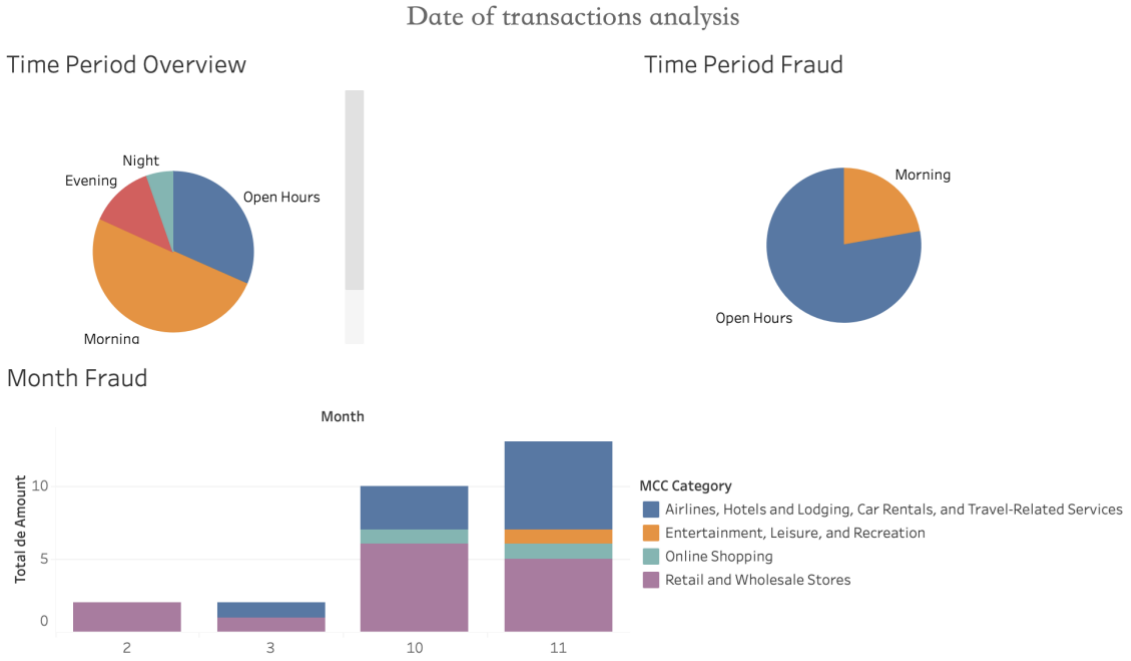


State and merchant type are strong indicators of fraudulent transactions. If a person is living in the US and the bank gets a transaction from France, it will probably be alerted as fraud.

We looked at the MCC category and the transaction state, we noticed that most frauds occurred online and in the state of CA mostly in retail stores.
Online transactions means more credit card transactions: even PayPal can be linked to a credit card. Moreover, contactless payment at retail outlets has become highly encouraged. We can pay with the credit card's "tap and go"; or we can pay with other payment services such as Apple Pay and they are usually linked to the users' credit card.

The use of credit cards and online payment methods may be influenced by certain periods of the year, such as the holiday shopping season or major events like Black Friday and Cyber Monday.

As we can see in our next graph here



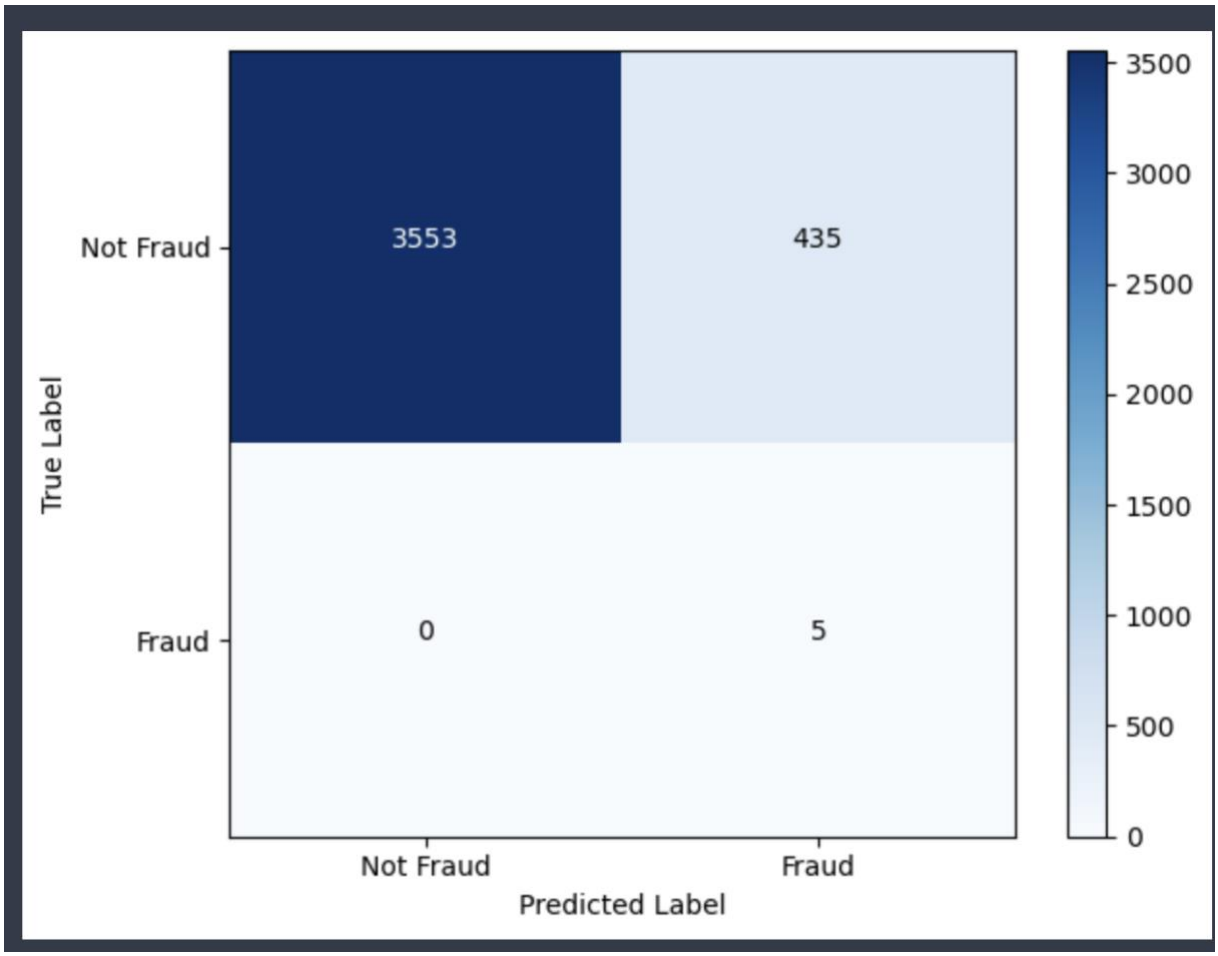
The majority of our fraud cases happened during store open hours, or early in the morning during the first 2 months and last months of the year or the holiday period.

During these times, there is an increase in the number of transactions made using credit cards,(Christmaas and new year's) which can in turn increase the likelihood of fraudulent activity. This makes it important to have effective fraud detection models in place, such as KNN, to help protect consumers during these periods.

Slide 9 : model

Model used: KNN neighbors

Score:



when a bank or financial institution's model predicts many false positives, it means that the model is identifying transactions as frauds when they are actually not frauds. This can be time-consuming and costly because they will have to spend extra resources checking with their clients to determine if the identified transaction is actually a fraud. Additionally, false positives can erode customer trust and lead to unnecessary friction in the banking process. While minimizing false positives is important to reduce the cost and effort of checking with clients, it is crucial to ensure that the model is accurately identifying positive events or behaviors to mitigate risk and maintain customer trust.

Conclusion:

we now have a better understanding now why we sometimes receive calls or alerts from our banks about certain transaction.

In conclusion, the use of machine learning models, such as KNN, can greatly benefit banks and financial institutions in detecting and preventing credit card fraud. With the increase in online transactions and contactless payments, credit card fraud has become a significant concern for consumers and financial institutions alike. By implementing machine learning models, banks can analyze large amounts of data and identify patterns that may be indicative of fraudulent activity. This allows for more efficient fraud detection and prevention, which ultimately helps protect consumers from financial harm.

Thank you.