I am writing this blog as recently I became a victim of credit card fraud. My card was not stolen but suddenly one night I got the alert from the credit card company that fraudulent activity has taken place in my card and my card was on hold. when similar incident happened to me fifteen years ago, I did not get any alert. I found about it accidentally by checking my statement online. This time I was surprised and happy that the bank found immediately and cautioned me. I was curious on the following.

1. How bank identified the fraudulent transaction happened in my card and alerted me?

2. what detection mechanism that bank use to identify fraud?

3. Since the card was not stolen, what was the source of this fraud?

It made me to do some research on the credit card fraud being a data science student. During that week, I made only two transactions. one is 2 hours before in a furniture retail store and another one is 2 days before in a gas station. This is not the regular gas station I go to. research shows that most of the skimmer fraud happens at ATMs and gas pumps.

Gas pump skimming: How big is the risk? According to the National Association for Convenience Stores:

- 37 million Americans refuel every day.
- Of them, 29 million pay for fuel with a credit or debit card.
- When skimming occurs at a gas station, it usually takes place at only one pump.
- A single compromised pump can capture data from 30 to 100 cards per day.

Fraud detection methods: I found some of the following methods used in fraud detection.

Predictive Analytics: Institutions collect vast amounts of data in the course of doing business – data that can be used to detect fraud patterns to determine future probabilities and trends. Although predictive analytics won't reveal what type of fraud will happen, it will point to what might happen, with an acceptable degree of reliability. Worldpay's fraud management, for example, uses predictive models based on billions of payment transaction data and consumer profiles. Using historical data combined with customer insights, Analytics and reporting is able to predict future fraud events so that actions can be taken to avoid and mitigate these events.

Outlier Models: In addition to using historical data, a fraud detection solution must also be able to adjust dynamically to the stream of data because fraud patterns are often not linear. Outlier models can be especially helpful for detecting fraud in emerging markets where sufficient data to make predictions does not yet exist. A flexible, self-calibrating model provides a great advantage because it requires less data and adjusts in real-time based on the transaction system– critical to saving money.

Classification: Classification is a machine learning paradigm that involves deriving a function that will separate data into categories, or classes, characterized by a training set of data containing observations (instances) whose category membership is known. This function is then used in identifying in which of the categories a new observation belongs. ***Binomial Logistic Regression*** algorithm is used to detect fraud.

The most commonly used fraud detection methods are ANNs, rule-induction techniques, decision trees, Support Vector Machines (SVM), LR, and meta-heuristics such as genetic algorithms, k-means clustering and nearest neighbor algorithms. These techniques can be used alone or in collaboration using ensemble or meta-learning techniques to build classifiers.

The sad part is bank does not go after the person who is committing fraud as its expensive for them.