**Credit Card Fraud Detection**

**TEAM MEMBERS**

* Asmita Sood RA2011031010084
* Laraib Hanif RA2011031010088
* Shashank Tripathi RA2011031010096
* Harshit Anand RA2011031010110

**ABSTRACT**

With the growing usage of credit card transactions, financial fraud crimes have also been drastically increased leading to the loss of huge amounts in the finance industry. Having an efficient fraud detection method has become a necessity for all banks in order to minimize such losses. In fact, credit card fraud detection system involves a major challenge: the credit card fraud data sets are highly imbalanced since the number of fraudulent transactions is much smaller than the legitimate ones. Thus, many of traditional classifiers often fail to detect minority class objects for these skewed data sets.

This Project deals with **CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING**. Machine learning models can recognise unusual credit card transactions and fraud involving collecting and sorting raw data, which is then used to train the model to predict the probability of fraud.

**INTRODUCTION**

A credit card is a thin handy plastic card that contains identification information such as a signature or picture, and authorizes the person named on it to charge purchases or services to his account - charges for which he will be billed periodically. Today, the information on the card is read by automated teller machines (ATMs), store readers, and banks and is also used in online internet banking systems. They have a unique card number which is of utmost importance. Its security relies on the physical security of the plastic card as well as the privacy of the credit card number.

There is a rapid growth in the number of credit card transactions which has led to a substantial rise in fraudulent activities. Credit card fraud is a wide-ranging term for theft and fraud committed using a credit card as a fraudulent source of funds in a given transaction. Generally, statistical methods and many data mining algorithms are used to solve this fraud detection problem. Most of the credit card fraud detection systems are based on artificial intelligence, Meta learning and pattern matching. A proper and thorough literature survey concludes that there are various methods that can be used to detect credit card fraud detection.

**MOTO**

The challenge is to recognize fraudulent credit card transactions so that the customers of credit card companies are not charged for items that they did not purchase.

**METHODOLOGY**

Credit Card Fraud Detection Using Machine Learning

* **Unsupervised Learning** - Machine Learning Algorithms such as Isolation Forest, One-class SVM, LOF, etc., do not require labelled data for training the model. They identify patterns in the data and try to group the data points based on observed similarities in patterns.
* **Supervised Learning** - Machine Learning Algorithms such as Ensemble Models (RandomForest, XGBoost, LightGBM, etc.), KNN, Neural Networks, Autoencoders, etc. These algorithms are trained on labelled data, and the model learns to predict the labels for the unseen data. Labelled data can be expensive to gather.

Like humans, machine learning algorithms learn from past transaction data and use that information to analyse future transactions with the same lens. While machines might not be as intelligent as humans are and might need some supervision on top of it, the advantage lies in the speed of data processing and computation. Also, machines can identify and remember more patterns in vast volumes of data compared to humans. Generally, these algorithms are known as **Anomaly Detection**.

The benefits of using machine learning for credit card fraud detection:

* Faster Detection
* High Accuracy
* Improved efficiency with larger data

**Main challenges involved in credit card fraud detection are:**

* Enormous Data is processed every day and the model build must be fast enough to respond to the scam in time.
* Imbalanced Data i.e., most of the transactions *(99.8%)* are not fraudulent which makes it really hard for detecting the fraudulent ones
* Data availability as the data is mostly private.
* Misclassified Data can be another major issue, as not every fraudulent transaction is caught and reported.
* Adaptive techniques used against the model by the scammers.

**How to tackle these challenges?**

* The model used must be simple and fast enough to detect the anomaly and classify it as a fraudulent transaction as quickly as possible.
* Imbalance can be dealt with by properly using some methods which we will talk about in the next paragraph
* For protecting the privacy of the user, the dimensionality of the data can be reduced.
* A more trustworthy source must be taken which double-check the data, at least for training the model.
* We can make the model simple and interpretable so that when the scammer adapts to it with just some tweaks, we can have a new model up and running to deploy.