**ONLINE PAYMENT FRAUD DETECTION SYSTEM BASED ON**

**MACHINE LEARNING**

**Abstract**

Online transactions are rapidly increasing, and it will mostly be done by credit cards. Loss of physical credit cards or loss of credit card information will make a person pay much more amount. The hackers are there to do fraud with people. So, there was a need to detect fraudulent transactions and secure online credit card transactions. For analysis of this problem and to overcome this fraud related to the credit cards, we have developed a new system namely “Online credit card fraud detection and prevention system” using machine learning. We have experimented with multiple algorithms to check the accuracy of the system. This system chooses a random forest algorithm to check whether the transaction is normally unfraudulent. Apart from random forest algorithms, there are also other algorithms like unbalanced classification, logistic regression, multilayer perception, and Naive Bayes. This system uses multiple decision trees at the backend of the random forest algorithm

**Introduction**

For a long time, online transaction fraudsters and detectors have played a complex role. Transaction fraud is more widespread than it has ever been, especially in the Internet era, and it causes enormous financial losses. The Nilsson research delved into the global landscape of online transaction fraud in considerable depth. Fraudulent online transactions cost the economy $21 billion in 2015, $24 billion in 2016, and more than $27 billion in 2017. The global rate of online transaction fraud is expected to rise year after year, reaching $31.67 billion in 2020.

Banks and financial service providers may be forced to implement an automated online fraud detection system to detect and monitor online transactions as a result. By isolating aberrant activity patterns from a huge number of transactional records, fraud detection systems are able to detect and monitor incoming transactions.

Machine learning has been shown to be quite good at detecting these patterns. A vast quantity of transaction records, on the other hand, may be used to train a high-performing fraud classifier. Despite the fact that supervised learning has been shown to be quite successful at detecting fraudulent transactions, transactional fraud analysis technology will continue to advance. Small adjustments might help a corporation save a lot of money. There are certain problems with the unique approach of unsupervised and regulated online fraud detection.

**METHODOLOGY**

**Decision Tree Algorithm**

A decision tree is a decision support tool. It is a nonparametric supervised learning method used for classification and regression. it works on training data and produces a decision tree. Decision trees are a very excellent tool for helping to choose between a number of courses of action. The main aim of the decision tree is to create a training model for predicting class.

**Decision Tree Algorithm Steps**

1. set the best attribute of the dataset at the root of the tree

2. divide the training set into subsets. Subsets should be made in the degree of a way that each subset contains data with the same value for an attribute.

3. Repeat steps 1 and 2 on each subset until you got leaf nodes in all the branches of the tree.

**• Disadvantage of Decision Tree Algorithm**

1. In the decision tree Algorithm high chances of overfitting

2. Training time is relatively expensive

3. There may be difficult to handle non-numeric data

**PROJECT**

**Problem Statement**

Create a model to estimate online payment fraud detection.

**Data**

The data is the most important aspect of a machine learning assignment, to which special attention should be paid. Indeed, the data will heavily affect the findings depending on where we found them, how they are presented if they are consistent if there is an outlier, and so on. Many questions must be addressed at this stage to ensure that the learning algorithm is efficient and correct.

To obtain, clean, and convert the data, many sub-steps are required. We will go through these steps to understand how they've been used in my project and why they're helpful for the machine learning section.

**EXPERIMENTAL SETUP**

**Steps to Create Model**

1. Import Libraries

2. Load Dataset

3. Exploratory Data Analysis

4. Data Cleaning

5. Feature Engineering

6. Data Visualization

7. Building a Model

**Tools used**

1. Anaconda

2. Jupiter Notebook

3. Google Collaboratory

**Technologies used**

Python

**CONCLUSIONS**

Online payment method fraud signifies a very serious commercial problem. These frauds can lead to huge losses. online paying money by the credit card is increased as also credit card fraud so there is a need to detect this happened fraud transactions and provide security to the users of credit cards. So, the main purpose of this paper is for detecting as well as preventing fraud during transactions. There are still some problems in the previous system such as accuracy. This paper has surveyed the performance of the Decision Tree Algorithm. Decision Tree obtains good results on small data sets as well as large data set.