A

SYNOPSIS

ON

BANK FRAUD DETECTION

SUBMITTED BY :-
~GOURAV RANA
~KEVIN PETER
~CH POOJA SATHVIKA

SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FOR MAJOR PROJECT SUBMISSION

Skill VERTEX

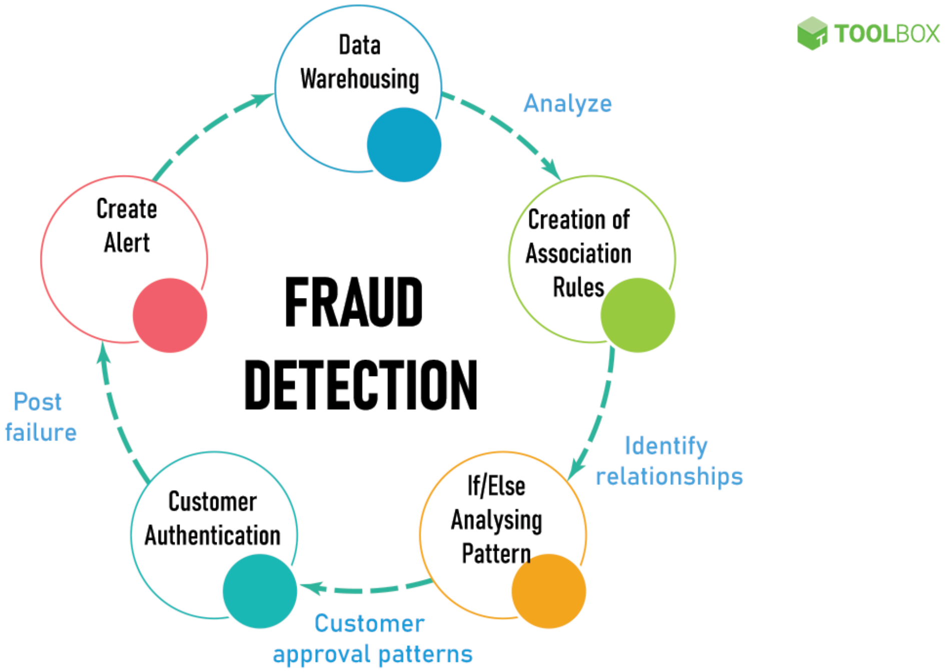
EDUCATION DOESN'T ASSURE EMPLOYMENT BUT SKILL DOES

TABLE OF CONTENTS~

INTRODUCTION

Fraud detection is a set of activities undertaken to prevent money or property from being obtained through false pretenses. Fraud detection is applied to many industries such as banking or insurance. In banking, fraud may include forging checks or using stolen credit cards. Other forms of fraud may involve exaggerating losses or causing an accident with the sole intent for the payout.

With an unlimited and rising number of ways someone can commit fraud, detection can be difficult to accomplish. Activities such as reorganization, downsizing, moving to new information systems or encountering a cybersecurity breach could weaken an organization's ability to detect fraud. This means techniques such as real-time monitoring for frauds is recommended. Organizations should look for fraud in financial transactions, location, devices used, initiated sessions, and authentication systems.



OBJECTIVE OF THE PROJECT

The objectives of credit card fraud detection are to reduce losses due to payment fraud for both merchants and issuing banks and increase revenue opportunities for merchants.

FRAUD DETECTION TECHNIQUES

Fraud is typically an act that involves many repeated methods; making searching for patterns a general focus for fraud detection. For example, data analysts can prevent insurance fraud by making algorithms to detect patterns and anomalies.

Fraud detection can be separated by the use of statistical data analysis techniques or artificial intelligence (AI).

Statistical data analysis techniques include the use of:

- Calculating statistical parameters
- Regression analysis
- Probability distributions and models.
- Data matching

AI techniques used to detect fraud include the use of:

- Data mining- This can classify, group and segment data to search through up to millions of transactions to find patterns and detect fraud.
- Neural networks- Which can learn suspicious-looking patterns, and use those patterns to detect them further.
- Machine learning- Which can automatically identify characteristics found in fraud.
- Pattern recognition- Which can detect classes, clusters and patterns of suspicious behavior.

TYPES OF FRAUD

Fraud can be committed in a number of different ways and in a number of different settings. For example, fraud can be committed in banking, insurance, government and in healthcare sectors.

One common type of fraud in banking is customer account takeover, where someone illegally gains access to a victim's bank account using bots. Other examples of fraud in banking include the use of malicious applications, the use of false identities, money laundering, credit card fraud and mobile fraud.

Fraud in insurance can include premium diversion fraud, which is the embezzlement of insurance premiums; or frees churning, which is excessive trading by a stockbroker to maximize commissions. Other forms of insurance fraud include asset diversion, workers compensation, car accident, stolen or damaged car, and house fire fraud. The motive behind all insurance fraud is financial profits.

Government fraud is committing fraud against federal agencies such as the departments of Health and Human Services, Transportation, Education, or Energy. Types of government fraud include billing for unnecessary procedures, overcharging for items that cost much less, providing old equipment when billing for new or reporting hours worked for a worker that does not exist.

Healthcare fraud includes drug fraud and medical fraud, as well as encompassing some insurance fraud. Healthcare fraud is committed when someone defrauds an insurer or government health care program.

SYSTEM SPECIFICATIONS

Software Engineers have been trying various tools, methods and procedures to control the process of software development in order to build high quality software with high productivity. This method provides how it is for building the software while the tools provide automated or semi automated support for the methods. They are used in all stages of software development process, namely, planning, analysis, design, development and maintenance. The software

development procedure integrates the methods and tools together and enables rational and timely development of the software system.

OLD SYSTEM

By doing manually on paper~
By doing manually fraud has been a major issue in sectors like banking, medical, insurance, and many others. Due to the increase in online transactions through different payment options, such as credit/debit cards, PhonePe, Gpay, Paytm, etc., fraudulent activities have also increased. Moreover, fraudsters or criminals have become very skilled in finding escapes so that they can loot more. Since no system is perfect and there is always a loophole them, it has become a challenging task to make a secure system for authentication and preventing customers from fraud

NEW SYSTEM

ML-based Fraud Detection Algorithms

In the rule-based approach, the algorithms cannot recognize the hidden patterns. Since they are based on strict rules, they cannot predict fraud by going beyond these rules. But in real world, fraudsters are very skilled and can adopt new techniques every time to commit a crime. Therefore, there is a need for a system that can analyze patterns in data and predict and respond to new situations for which it is not trained or explicitly programmed.

Hence, we use Machine Learning for detecting fraud. Here, a machine tries to learn by itself and becomes better by experience. Also, it is an efficient way of detecting fraud because of its fast computing. It does not even require the guidance of a fraud analyst. It helps in reducing false positives for transactions as the patterns are detected by an automated system for streaming transactions that are in huge volume.

Now, we will look at the two most commonly used Machine Learning models for detecting fraud in transactions.

Supervised Learning Used in Fraud Detection Algorithms

Supervised Learning models are trained on tagged outputs. If a transaction occurs, it is tagged as either ‘fraud’ or ‘non-fraud.’ Large amounts of such tagged data are fed into the supervised learning model in order to train it in such a way that it gives a valid output. Also, the accuracy of the model’s output depends on how well-organized your data is.

Kick-start your career in Artificial Intelligence with the perfect Artificial Intelligence Course now!

Unsupervised Learning Used in Fraud Detection Algorithm

Unsupervised learning models are built to detect unusual behavior in transactions which is not detected previously. Unsupervised learning models involve self-learning that helps in finding hidden patterns in transactions. In this type, the model tries to learn by itself, analyzes the available data, and tries to find the similarities and dissimilarities between the occurrences of transactions. This helps in detecting fraudulent activities.

So, both these models, supervised and unsupervised, can be used independently or in combination for detecting anomalies in transactions.

Need for the Fraud Detection Machine Learning Algorithms

Human beings always search for methods, tools, or techniques that reduce the human effort for performing a certain task efficiently. In Machine Learning, algorithms are designed in such a way that they try to learn by themselves using past experience. After learning from the past experience, the algorithms become quite capable of reacting and responding to conditions for which they are not explicitly programmed. So, Machine Learning helps a lot when it comes to fraud detection. It tries to identify hidden patterns that help in detecting fraud which has not been previously recognized. Also, its computation is fast as compared to the traditional rule-based approaches.

Why do we use Machine Learning in Fraud Detection?

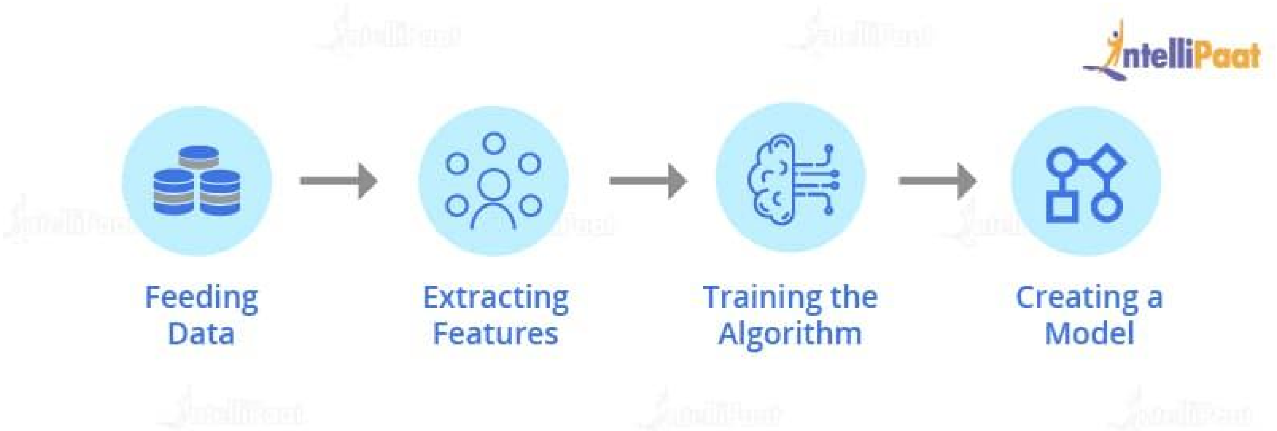
Here are some factors for why Machine Learning techniques are so popular and widely used in industries for detecting frauds:

- Speed: Machine Learning is widely used because of its fast computation. It analyzes and processes data and extracts new patterns from it within no time. For human beings to evaluate the data, it will take a lot of time and evaluation time will increase with the amount of data. Rule-based fraud prevention systems are based on written rules for permitting which type of actions are deemed safe and which one’s must raise a flag of suspicion. Now, this Rule-based system is inefficient because it takes much time to write these rules for different scenarios. And that’s exactly where Machine Learning based Fraud Detection algorithms succeed in not only learning from these patterns it is capable of detecting new patterns automatically. And it does all of this in a fraction of the time that these rule-based systems could achieve.

- Scalability: As more and more data is fed into the Machine Learning-based model, the model becomes more accurate and effective in prediction. Rule-based systems don't evolve by themselves as professionals who developed these systems must write these rules meeting various circumstances. But for Machine Learning based algorithms, a dedicated team of Data Science professionals must be involved in making sure these algorithms are performing as intended.
- Efficiency: Machine Learning algorithms perform the redundant task of data analysis and try to find hidden patterns repetitively. Their efficiency is better in giving results in comparison with manual efforts. It avoids the occurrence of false positives which counts for its efficiency. Due to their efficiency in detecting these patterns, the specialists in Fraud detection could now focus on more advanced and complex patterns, leaving the low or moderate level problems to these Machine Learning based algorithms.

How does a Machine Learning system work for Fraud Detection?

The below picture shows the basic structure of the working of fraud detection algorithms using Machine Learning:



Feeding Data: First, the data is fed into the model. The accuracy of the model depends on the amount of data on which it is trained, more data better the model performs.

For detecting frauds specific to a particular business, you need to input more and more amounts of data into your model. This will train your model in such a way that it detects fraud activities specific to your business perfectly.

Extracting Features: Feature extraction basically works on extracting the information of each and every thread associated with a transaction process. These can be the location from where the transaction is made, the identity of the customer, the mode of payments, and the network used for transaction.

- Identity: This parameter is used to check a customer’s email address, mobile number, etc. and it can check the credit score of the bank account if the customer applies for a loan.
- Location: It checks the IP address of the customer and the fraud rates at the customer’s IP address and shipping address.
- Mode of Payment: It checks the cards used for the transaction, the name of the cardholder, cards from different countries, and the rates of fraud of the bank account used.
- Network: It checks for the number of mobile numbers and emails used within a network for the transaction.

Training the Algorithm: Once you have created a fraud detection algorithm, you need to train it by providing customers data so that the fraud detection algorithm learns how to distinguish between ‘fraud’ and ‘genuine’ transactions.

Creating a Model: Once you have trained your fraud detection algorithm on a specific dataset, you are ready with a model that works for detecting ‘fraudulent’ and ‘non-fraudulent’ transactions in your business.

The advantage of Machine Learning in fraud detection algorithms is that it keeps on improving as it is exposed to more data.

There are many techniques in Machine Learning used for fraud detection. Here, with the help of some use cases, we will understand how Machine Learning is used in fraud detection.

Techniques of Machine Learning for Fraud Detection Algorithms

1. Fraud Detection Machine Learning Algorithms Using Logistic Regression: Logistic Regression is a supervised learning technique that is used when the decision is categorical. It means that the result will be either ‘fraud’ or ‘non-fraud’ if a transaction occurs.
2. Fraud Detection Machine Learning Algorithms Using Decision Tree: Decision Tree algorithms in fraud detection are used where there is a need for the classification of unusual activities in a transaction from an authorized user. These algorithms consist of constraints that are trained on the dataset for classifying fraud transactions.
3. Fraud Detection Machine Learning Algorithms Using Random Forest: Random Forest uses a combination of decision trees to improve the results. Each decision tree checks for different conditions. They are trained on random datasets and, based on the training of the decision trees, each tree gives the probability of the transaction being ‘fraud’ and ‘non-fraud.’ Then, the model predicts the result accordingly.
4. Fraud Detection Machine Learning Algorithms Using Neural Networks: Neural Networks is a concept inspired by the working of a human brain. Neural networks in Deep Learning uses different layers for computation. It uses cognitive computing that helps in building machines capable of using self-learning algorithms that involve the use of data mining, pattern recognition, and natural language processing. It is trained on a dataset passing it through different layers several times.

TECHNOLOGY USED:-

Language ~ Python
Libraries ~ Numpy, Pandas,

SYSTEM REQUIREMENT:-

Minimum ram	- 256 MB
Hard disk	- 40
Processor	- Intel Pentium 4

CONCLUSION:-

Machine learning-based fraud prevention is an exciting new development in the prevention of illicit payments.By replacing outdated rule-based systems with modern machine learning solutions, banks and payment processors can reduce the losses they incur due to fraud, lower their security system-related expenses, and reduce payment friction for their clients.

As for the companies that are hesitant to switch, the costs associated with maintaining their legacy payment fraud systems will eventually outweigh the investment necessary to introduce the more modern system. It is predicted that all major financial industry players will eventually transition to machine learning-based payment fraud prevention systems.

GROUP MEMBER DETAILS:~

Member 1;

Name- Gourav Rana

Ph no- +917015208168

E-mail- gourav.rana70000@gmail.com

Member 2;

Name-Kevin Peter

Ph no- +919595855577

E-mail- peter.kevin809@gmail.com

Member 3;

Name- Ch Pooja Sathvika

Ph no- +919900877436

E-mail- poojasathvika999@gmail.com