IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

BANKING FRAUD DETECTION USING MACHINE LEARNING

Mr.MANISH BHADANE, Mr.SWAPNIL KATKADE, Mr.SANDIP GADADE, MR.ARVIND DATE.

Abstract- Financial services are used everywhere and operate in a very complex manner. With the increase in online transactions comes an alarming increase in fraud. An automatic fraud detection system is therefore required. With millions of transactions in progress, it is virtually impossible to detect fraud manually with good speed and accuracy. We design a system that provides a robust, cost effective, efficient, yet accurate solution for fraud detection in both online payment transactions and credit card payments. The proposed solution is a machine learning model that will be used to detect "fraudulent" and "genuine" transactions in real time. This is beneficial for all sectors that are even slightly attuned to finance. The solution will help them analyse based on various factors whether an ongoingtransaction may be malicious and prevent manyunfortunate incidents.

Keywords- fraud detection, Retail Frauds, credit card,online payments, fraudulent.

INTRODUCTION

Frauds in online transactions are seen in abundance with the increase in the use of online methods within the government, private organizations and throughout various sectors in order to have faster payment infrastructure. High reliance on technology has resulted in increased banking transactions in today's world. However, frauds in the banking sector have accelerated jointly as transactions online and offline. As transactions became a commonly used mode of payment, there is a lot of research being done in this area to combat against the ever increasing financial frauds online. The current systems in place are not able to efficiently handle the pace at which the number of transactions are happening. Studies suggest 93% of the merchants perform manual review anywhere between 1% - 10% of orders for fraud detection and this is costly, time consuming and it also leads to higher falsenegatives and human errors.

LITERATURE REVIEW

In a world driven by technological advances and complex socio- economic dynamics towards full digitization, products disappearing from the shelves are not the only or worst threat to retailers and their customers.

As the 2020 Nelson report points out, electronic payment fraud is agrowing phenomenon that affects several industries and generated $28.65 billion in global losses in 2019 alone.

Machine learning systems are also a powerful weapon in the fight against identity theft by monitoring users' shopping habits and transaction data.

TYPES OF FRAUDS:

The types of retail fraud:

1. Transaction fraud

It is also called card-not-present (CNP) fraud where the fraudster uses a stolen credit card for online purchases. The company loses money when the original owner of the card demands a chargeback.

2. Return fraud:

As the e-commerce industry eased its return policies for the convenience of its customers, these became the favoured target for fraudsters to exploit and abuse. Some of the most common instances are wardrobing, receipt fraud, price switching or open box fraud, price arbitrage, and bricking

3. Chargeback guarantee fraud:

Many online retail fraud prevention solutions guarantee that they will block all transactions and friendly frauds and even pay the admin fee out of their pocket.

4. Triangulation Fraud:

Triangulation fraud is when a customer makes a genuine purchase on a third-party marketplace (for example Amazon or Sears.com), but the product they receive was fraudulently purchased from a different retailer's website. This practice harms businesses of all kinds Customers usually aren't aware.

Techniques for the avoid Fraud in retail:

1. Predictive analytics:

It leverages analytics tools and platforms for large-scale customers and transactional data to detect fraudulent activity linked to previous incidents of fraud.

AI fraud detection solutions backed by predictiveanalytics can synchronise with retail payment processing infrastructure at the point of sale. Then ML algorithms that fuel fraud detection systems learn to identify Trends and characteristics links.

2. Anomaly Detection:

AI fraud detection systems for retail transactions function by analysing massive amounts of previous and contemporary transaction data to discover underlying motives and detect anomalies.

When an anomaly is spotted, AI-driven anomaly analytics solutions can restrict a user, a transaction, or inform retailers, depending on the documented principles.

Benefits:

Quick and accurate detection of potential frauds. Reduces the cost incurred due to fraudulent activities. Impactful real-time data processing.

The Types of Credit card frauds:

POS scam

In this type of scam, small skimming devices are attached to normal Point-of-Sale (POS) devices to hack your data. These devices scan and store card information while the customer completes the download transaction. This is usually a merchant or store employee sharing these details with malicious intent. Similar attachments can also be attached to ATM card slots to clone card information, while a camera is secretly placed on the keypad to capture yourPIN.

Phishing and vishing

These include pretending to be official communications from the bank, which in turn act as bait to get you to click on fake links. This will usually get you to websites that look authentic. Once you enter your card details on these fake links, fraudsters can access them and use them to their advantage. Another version is when scammers make phone calls impersonating bank officials and ask you to share a one-time password in order to 'verify your card' or 'earnrewards points' or 'extend your rewards points'.

Keystroke recording

Nowadays, with most financial transactions taking place online, hackers have come to rely on keystroke logging using malicious software to obtain credit card information. This usually starts after you click on asuspicious link and unknowingly install malware on your system. The software records every keystroke on the system, potentially stealing card details, PINs and more.

Application fraud

This is a type of identity theft where fraudsters impersonate a real customer using their stolen or forged documents to obtain a credit card. While this can be detected after thorough background checks, if done, it allows criminals to use a valid credit card with a fake paper trail. A similar type of fraud involves taking over a valid credit card account by impersonating a customer using a similar fake paper trail.

The Types of Online payment Frauds Identity Theft Identity theft, where cybercriminals steal sensitive personal and payment information and use it to make fraudulent online payments.

Friendly Fraud

Friendly deception is not friendly in the real world. This is a trick used by the real cardholder to deceive the seller by falsely claiming that they did not receive the purchased item or report a damaged product and request a refund.

Triangulation fraud

Triangulation fraud, also known as phishing, is a type of fraud where a seller creates fake and malicious pages to lure customers into fake products and collect their credit card details in order to make other fraudulent transactions in the future.

Algorithm:

XG Boost is the algorithm that dominates recently applied machine learning and Kaggle competitions for structured or tabular data. XG Boost is an implementation gradient boosted decision trees designed for rate a performance. The XG boost algorithm is based on a gradient boosted decision trees. With the help of these decision tree he classified data for fraud or not. Because it is based on a decision tree, it can they give us pretty good accuracy as well as efficiency. This The algorithm has some key properties that are optimal results and at high speed. Some of the advantages of the XG Boost algorithm:

A. Regularization: XG Boost has a built-in L1 (Lasso regression) and L2 regularization (ridge regression). which prevents the model from being displaced. That's why XG Boost is also called a regularized form of GBM (Gradient Fitness machine).When using the Scikit Learn library, we pass two hyper parameters (alpha and lambda) to XG Boost related to regulation. alpha is used for L1 regularization and lambda is used for L2 regularization.

b. Parallel processing: XG Boost harnesses the power parallel processing and therefore much faster than GBM. It uses multiple CPU cores to run the model. When using the Scikit Learn library, the n thread hyperparameter is used for parallel processing. a thread represents the number of CPU cores to be used. If you want to use all available cores, do not specify any value for n thread and the algorithm automaticallydetects.

C. Handling missing values: XG Boost has a built-in ability to handle missing values. When XG Boost encounters a missing value in a node, tries both left and the right-hand split and learns the path leading to the higher loss for each node. He then does the same when working on test data.

d. Cross-validation: XG Boost allows users to run crossvalidation at each iteration of the boosting process and so it is easy to get the exact optimum number boosting iterations in a single run. This is unlike GBM where we need to run a grid search and only limited values can be tested.

E. Effective tree pruning: GBM would stop splitting a node when it encounters a negative split loss. Thus, it's more of a greedy algorithm. XG Boost on the other hand does the split up to the specified Max depth and then he begins to prune the tree back and remove the splits behind which there is no positive gain.

FUTURE IMPROVEMENTS

Although we couldn't reach the goal of 100% fraud accuracy detection, we finally created a system that can enough time and data to get very close to that goal. As with everyone such a project, there is some room for improvement. The very nature of this project allows for multiple algorithms to be integrated together as modules and their results can be combined to increase the accuracy of the final result. this model can be further improved by adding more algorithms into it. However, the output of these algorithms must be in the same format as the others. Once that condition is met, modules can easily be added as they were done code. This provides a high degree of modularity and versatility to the project. further room for improvement can be found in the dataset. As demonstrated earlier, the accuracy of the algorithms increases when the size of the data set increases so there will be more data will certainly refine the model in fraud detection and reduce the number of false positives. However, it requires official support from the banks themselves.

CONCLUSION

Fraud detection in current trends in payment methods hours are needed. The solution proposed in this article is a robust, fast and accurate way to detect fraud that occur in both online and credit payments cards. Because the solutions are based on a high-performance machine Learning algorithms provide some cost- effective and quick predictions and thereby reducing the risk of fraud. This approach may be suitable for real-world applications where companies providing financial services, banks, financial institutions, etc. can deploy it and achieve maximum customer satisfaction by preventing fraud that occurs in transactions carried out by them users So Fraud Detect will fulfil the demand in real time detecting fraud before it ever happens.

REFERENCES

https://www.infosysbpm.com/blogs/fraud-retail/howto-prevent- fraud-in- retail.html#:~:text=Use%20fraud%20detection%20technology%3A&text=An%20AI%2Dbased%20retail%20 fraud,it%20before%20it%20causes%20damage.

https://recosenselabs.com/blog/ai-for-fraud-detectionin-retail- powerful-use-cases

https://www.itransition.com/machine-learning/frauddetection

1. K.Chaudhary, J.Yadav, “A review of fraud: Acomparative study.”decis . Support syst, vol 50, no3, pp.602-613,2011

2. Katherine J. Barker , Jackie D’Amato ,Paul Sheridon,2008 “Credit card fraud :awareness and prevention”, Journal+- of financial Crime ,Vol. 15issue:4,pp.398-410

3. Dipti Thakur ,salamis Bhatia “distribution data Mining approach to credit card fraud detection” SPIT IEEE colloquium and international conference , volume4,48,issue2002.

4. “CreditCard Fraud Detection Based on Transaction Be haviour - by John Richard D. Kho, Larry A. Vea” published by Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017.

5. Customer Transaction Fraud Detection Using Xgboost Model -by Yixuan Zhang, Ziyi Wang, Jialiang Tong, Fengqiang Gao June, 2020

6. Jerome H. Friedman. Greedy function approximation: a Gradient Boosting machine. The Annals of Statistics, 29(5):1189 – 1232, 2001.

7. Wang, M., Yu, J., & Ji, Z. (2018). Credit Fraud Risk Detection Based on XGBoost-LR Hybrid Model.

8. A. Mishra, C. Ghorpade, “Credit Card Fraud Detection on the Skewed Data Using Various Classification and Ensemble Techniques” 2018 IEEE International Students' Conference onElectrical, Electronics and Computer Science (SCEECS) pp. 1-5. IEEE.