CREDIT CARD FRAUD DETECTION

AKALYA SK	UG STUDENT, COMPUTER SCIENCE AND ENGINEERING, SNS COLLEGE OF TECHNOLOGY, COIMBATORE, INDIA- 641035
ANUVARSHINI SS	UG STUDENT, COMPUTER SCIENCE AND ENGINEERING, SNS COLLEGE OF TECHNOLOGY, COIMBATORE, INDIA- 641035
AKSHAYA BS	UG STUDENT, COMPUTER SCIENCE AND ENGINEERING, SNS COLLEGE OF TECHNOLOGY, COIMBATORE, INDIA- 641035
ABHINAV S	UG STUDENT, COMPUTER SCIENCE AND ENGINEERING, SNS COLLEGE OF TECHNOLOGY, COIMBATORE, INDIA- 641035
SATHISH KUMAR S	ASSOCIATE PROFESSOR, COMPUTER SCIENCE AND ENGINEERING, SNS COLLEGE OF TECHNOLOGY, COIMBATORE, INDIA- 641035

ABSTRACT:

Fraud detection with Machine Learning becomes possible since it is easy to learn from historical fraud patterns and recognize them in further transactions. ML algorithms are able to find sophisticated fraud traits which humans can't detect. Fraud models can be tackled with both supervised and unsupervised ML algorithms. In the first case, traditional classifications are used. Whereas in second case, anomaly detection techniques are used.

KEYWORDS:

FRAUD DETECTION, MACHINE LEARNING, TRANSACTIONS, CLASSIFICATION ALGORITHM, FRAUD PATTERNS.

INTRODUCTION

The things we used to buy at stores are now purchased online. There are over 12 – 24 million e- commerce websites available. This triggers members of criminal world to find ways to take over the victim's wallet. This ML algorithm helps in recognizing fraud patterns in credit card transactions

CREDIT CARD FRAUD DETECTION:

A. PROBLEM STATEMENT:

Credit card fraud events take place frequently which results in huge financial losses. Number of online transactions has grown in large numbers and online credit card transactions hold a huge share of these transactions.

B. OBJECTIVE:

This project mainly focuses on the objective of detecting fraudulent credit card transactions. As years go by, the number of online transactions have raised so much. This may lead to many fraud practices by criminals. Thus, this algorithm goes through the historical fraud patterns and detects fraud transactions. Organizations, banks and merchants are put into risk when a data breach leads to monetary theft. This might be helpful for large organizations, banks and companies where a large number of transactions take place.

C .EXISTING SYSTEM:

Large online merchants and payment service providers are no strangers to credit card fraud and its consequences. They have been building their risk management strategies for years, being among early adopters of machine learning. Some of these pioneers share experience with the general public, even giving open access to their anti - fraud solutions.

D. PROPOSED SYSTEM:

This project aims to develop a secure environment for online credit card transactions. The main objective of this credit card detection system is to identify suspicious events and report them to an analyst. This project actually represent that any e-commerce site need not be as big as Amazon or PayPal to adopt machine learning. Machine learning can be more effective as it dramatically reduces the number of false positive transactions that require manual verification. The approach with ML algorithm can bring significant improvements to the process

The ML algorithms used in the detection of credit card fraud are:

- ○ PCA (Principal Component Analysis)
- ○ LOF (Local Outlier Factor)
- ○ One-class SVM (Support Vector Machine)
- ○ Isolation Forest (IF)

E. ALGORITHM DESCRIPTION:

1) *PCA (Principal Component Analysis):* PCA enables the execution of an exploratory data analysis to reveal the inner structure of the data and explain its variations. PCA is one of the most popular techniques for Anomaly Detection. PCA searches for correlations among features

— which in the case of credit card transactions, could be time, location, and amount of money spent — and determines which combination of values contributes to the variability in the outcomes. Such combined feature values allow the creation of a tighter feature space named *principal components*

2) *LOF (Local Outlier Factor):* LOF is the score factor that helps understand how high the chance is for a certain data sample to be an outlier (anomaly). This is another of the most popular Anomaly Detection methods. To calculate LOF, the number of neighboring data points is considered to figure out its density and compare it to the density of other data points. If a certain data point has a substantially low density compared to its close neighbors, it is an outlier.
3) *One-class SVM (Support Vector Machine):* SVM is a classification algorithm that helps to identify outliers in data. This algorithm allows one to deal with imbalanced data-related issues such as Fraud Detection. The idea behind One-class SVM is to train only on a solid number of legitimate transactions and then identify anomalies or novelties by comparing each new data point to them.
4) *Isolation Forest (IF):* IF is an Anomaly Detection method from the Decision Trees family. The main idea of IF, which differentiates it from other popular outlier detection algorithms, is that it precisely detects anomalies instead of profiling the positive data points. Isolation Forest is built of Decision Trees where the separation of data points happens first because of randomly selecting a split value amidst the minimum and maximum value of the chosen feature. Subsequently, if we have a set of legitimate transactions, the Isolation Forest algorithm will define fraudulent credit card transactions because of their values — which are often very different from the values positive transactions have.

RESULTS:

The algorithm will intimate the Data analyst / Data Scientist. Then it will go by authentication factors. This pattern will be changing while the users continues a new trend. Then according to the pattern the algorithm will work.

When a user requests a transaction the algorithm will check the probability of the fraud rate. If the fraud rate is below 10% the transaction is allowed, when it has 10% - 80% then it will have security checkups like sending an OTP, a SMS and some security questions. If the probability is more than 80% then the transaction will be frozen then it will require a manual verification.

DISCUSSION:

Unlike robbery, identity theft can go completely unnoticed before the victim encounters a dramatic loss. The “red flags” for understanding that you have become a victim are unknown transactions or increasing debt on a credit card, the source of which is unknown to you. Mail about the spent money can go to another address of the criminal’s choice. Thus, you will not know about the situation right away.

CONCLUSIONS:

As long as the modern world is overwhelmed with card-not-present transactions online, the Banking and Retail sectors are under threat and face many fraud cases. Email phishing, payment fraud, identity theft, document forgery, and fake accounts contribute to the high level of criminal attacks on vulnerable users’ data and lead to data breaches. As old rule-based algorithms for fraud detection fade into the past, new top-notch methods based on Machine Learning algorithms for fraud detection and prevention are bringing greater value to businesses with their real-time work, speed, and efficiency.

In future, we decided to integrate this algorithm in a website. Because by keeping as a program will not help us to solve anything so we planned to embed in a live website and we will see the results. If this works better, we can actually provide those algorithms to small - start-ups which will help them a lot.

ACKNOWLEDGMENTS:

The authors would like to thank Mr.S.Sathish Kumar (ASP / CSE/SNSCT) and Mr.G.Ramesh Kalyan (AP/CSE/SNSCT) for their useful suggestions and discussions.

FIGURES:

GRAPHICAL REPRESENTATION OF THE PROJECT:

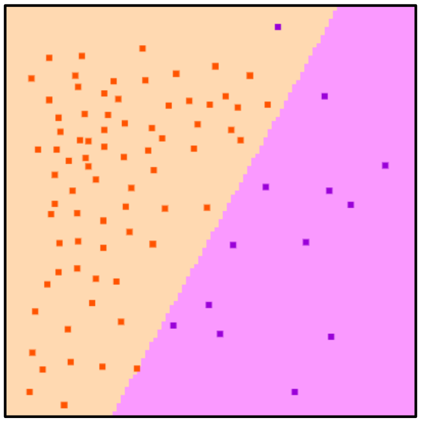

FIG. 1 PICTORIAL REPRESENTATION OF ALGORITHM ACCURACY

This is a pictorial representation of our algorithm. In this the orange region shows the genuine transactions and the purple region shows the fraud / suspicious transactions.

REFERENCES

1. https://www.sciencedirect.com/science/article/pii/S187705092030 065X

2. https://www.researchgate.net/publication/336800562_Credit_Card_Fraud_Detection_using_Machine_Learning_and_Data_Science

3. https://www.kaggle.com/code/marcelotc/credit card-fraud-logistic- regression-example

4. https://www.ijert.org/credit-card-fraud-detection

5. https://towardsdatascience.com/credit-card-fraud-detection-using-machine-learning-python 5b098d4a8edc

6. https://www.sciencedirect.com/science/article/pii/S187705091500 7103

7. https://journalofbigdata.springeropen.com/articles/10.1186/s40537-022-00573-8

8. https://www.ripublication.com/ijaer18/ijaerv13 n24_18.pdf

9. https://asianssr.org/index.php/ajct/article/view/11 51

10. https://digitalcommons.aaru.edu.jo/cgi/view content.cgi?article=11 05&context=fcij

11. https://www.sciencepubco.com/index.php/ijet/article/view/9356

12. https://fraud.net/n/five-methods-of-banking-fraud -prevention/

13. https://seon.io/resources/banking-fraud-detection-and-prevention/

14. https://www.crisil.com/en/home/our-businesses/global-research-and-risk-solutions/our-offerings/non-financial-risk/financial- crime-management/fraud-management/fraud-detection-and-analytics.html