You Choose, We Do It

St. JOSEPH'S COLLEGE OF ENGINEERING

(An Autonomous Institution)

St. Joseph's Group of Institutions

OMR, Chennai - 119.

CREDIT CARD FRAUD DETECTION SYSTEM USING MACHINE LEARNING ALGORITHMS

K . ANA JESSICA
312319104014

V . FEBI
312319104031

ABSTRACT

Financial fraud is a white-collar crime that affects the general public and has a negative impact on the whole economy. Our project focuses on credit card fraud, where criminals can steal your credit card information and misuse it. They could steal your physical card, trick you into entering information on a phishing website or email, buy your details on the Dark Web, or use any number of other credit card scams.In this project, we designed a machine-learning (ML)-based risk scoring system, which introduces early warning against fraud in the banking environment of any e-commerce site during card-not-present transactions. The proposed system analyzes the client's transaction details and can find legit (white), low risk fraud (gray) and high risk fraud (black) in the early stage just after the client authentication and authorization process. This proposed method decreases the number of successful fraud occurrences.

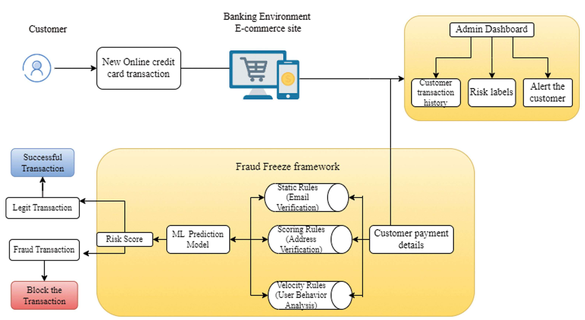
INTRODUCTION

Theft and fraud rates are increasing each year in online banking. Rules based systems rely on hard coded rules that are set to flag transactions if they meet certain criteria.In recent years, machine learning systems have grown in popularity to complement rules-based systems. For big complex systems with large volumes of data, machine learning algorithms can be a powerful tool to find non-liner, non-intuitive patterns.

Our project provides the most robust solutions that require a mixture of both types of systems. One of the most common ways to use the strengths of both machine learning models and rules-based systems is to feed the output of the machine learning model into one of the inputs of the rules-based alerts, thus combining the best of both worlds.

Before the payment gateway process begins, our project, "Fraud Freeze" framework will collect the customer payment details such as IP address of the customer, Address (City, State), E-mail Id and Transaction date and time.

SYSTEM ARCHITECTURE



Our framework can be commonly used across businesses, as they try to detect fraud in their transactions to avoid major profit losses. Our solution applied to many industries like banking, insurance, and e-commerce. With so much at stake and so many variables changing, it's vital to have a robust real-time monitoring system for fraud like this.

SYSTEM ARCHITECTURE

In our project, we built a custom machine learning model, to predict the outcome of each transaction in real-time and we developed a unique fraud score for each transaction. A Fraud Score is an informational tool that helps you gauge the risk involved with each transaction before processing.
The fraud or risk score ranges from 3-9 for every event or transaction. In short, this score indicates the relative risk of fraud. Based on the score, each event is segmented into one of 3 risk levels:

Low Risk (4-6): Low possibility of fraud, but may include false negatives (risk).
High Risk (7-9): High possibility of fraud, but may include false positives.
Legit (3): No possibility of fraud, but may include false negatives (risk).

Using this method, admins are able to prioritize reviews of transactions based on risk.

METHODOLOGY

For privacy and security purposes, the customer's credit card number, CCV (Card Verification Value) or Card Expiry Date is not stored or processed by our framework.
With the collected necessary details, our framework processes the data into required format before it is analyzed by three modules. The three modules are:

1.tatic Rules System
2.Scoring Rules System
3.Velocity Rules System

Each fraud detection rule helps us decide if an activity is fraudulent or not.. It is based on correlation, statistics, and logical comparison. Each system is responsible for three different verifications, whose results are fed into the custom trained machine learning model.

MODULE 1 - STATIC RULES

A static rule is the most basic form of fraud rule and tends to follow a simple if/then logic. It is considered static when the outcome of the rule is strict and inflexible – for instance, blocking a user action. Historically, the earliest static rules were like if an IP was found on a blacklist, the user would be blocked.

However, we have used static rules creatively based on the data we had at hand. We used it to verify if the provided customer's email-id is properly formatted and that its mailbox exists and is able to receive mail, without even sending an email. We also make sure that the domain name is valid and also check whether it's a disposable email address or not.

MODULE 2 - SCORING

Fraud rules don't always have to block actions. Sometimes, you want them to inform your risk strategy, and this is done by risk scores. For instance, we have access to IP information and credit card details. So our address verification rule could be made up of two parameters.

The first parameter will hold the credit card details of a customer that point to the US.
The second parameter will look at IP addresses of the same customer and it may point to Russia.

Thus based on the difference in the country code and the distance between ip address location and the billing address, we assign a risk score.

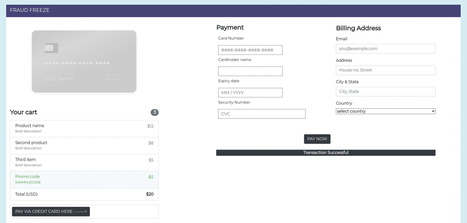
MODULE 3 - VELOCITY RULES SYSTEM

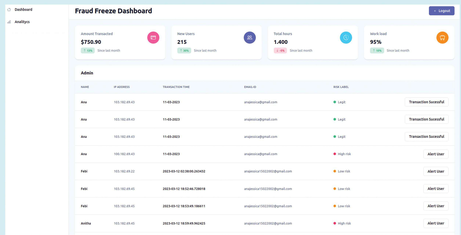
Velocity rules attempt to understand the user behavior by looking at the customer transaction history over a time period.

A good example of a velocity rule would be when a customer whose transaction was previously blocked tries multiple times. If you have no barriers there, they can use credential stuffing or brute force to crack credit card details.

RESULT

All the three modules impact our model's fraud prediction score, and they are fed into the ensemble machine learning model to generate the fraud prediction. Each of them adds a value and ultimately shows how risky the user's action is.





CONCLUSION

By utilizing both machine learning and rules, we capture fraudulent user behavior that is both human intuitive and complex. Rules are powerful when we know the logic and we need to be precise. Machine learning is powerful when we want to be proactive and predictive when approaching a large, complex data set. When it comes to online credit card fraud detection, we use both of these methodologies to provide the best solution.