CREDIT CARD FRAUD DETECTION USING DATASCIENCE

INTRODUCTION:

Credit card fraud is the act of using another person's credit card to make purchases or request cash advances without the cardholder's knowledge or consent. These criminals may obtain the card itself through physical theft, though increasingly fraudsters are leveraging digital means to steal the credit card number and accompanying personal information to make illicit transactions.

PROBLEM STATEMENT:

The problem is to develop a machine learning-based system for real-time credit card fraud detection. The goal is to create a solution that can accurately identify fraudulent transactions while minimizing false positives. This project involves data preprocessing, feature engineering, model selection, training, and evaluation to create a robust fraud detection system.

Card present fraud

Card present fraud is when the criminal uses a physical card, which is either stolen or duplicated, to make fraudulent purchases. Card present fraud can be the result of the theft of a card through robbery, pickpocketing, or mail theft.

Card-not-present fraud

Card- fraud not-present is when the criminal uses the details associated with the card, such as the card number, accountholder name, and CVV code, without having the card in their possession.

Creditcard Fraud:

According to the FBI, credit card fraud is *"the unauthorized use of a credit or debit card, or similar payment tool to fraudulently obtain money or property."* All players involved in the card-based payment process can potentially fall victim to scammers, including:

- cardholders,
- online merchants,
- payment processing companies,
- credit card payment systems,
- card issuers (issuing banks), and
- acquirers (acquiring banks).

DESIGN THINKING STEPS:

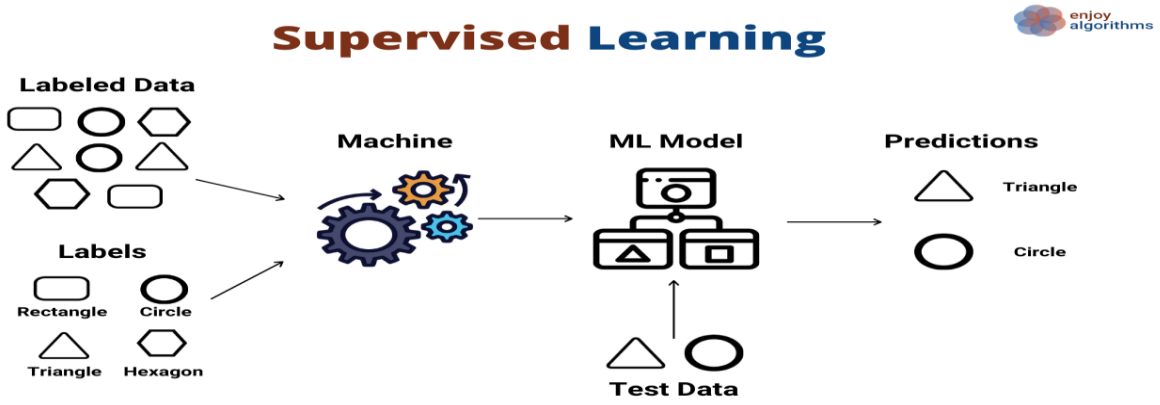
1. Data Source:
2. Data Preprocessing
3. Feature Engineering
4. Model Selection
5. Model Training
6. Evaluation

How machine learning helps with fraud detection:

1. Higher accuracy of fraud detection.
2. Less manual work needed for additional verification
3. Fewer false declines
4. Ability to identify new patterns and adapt to changes.

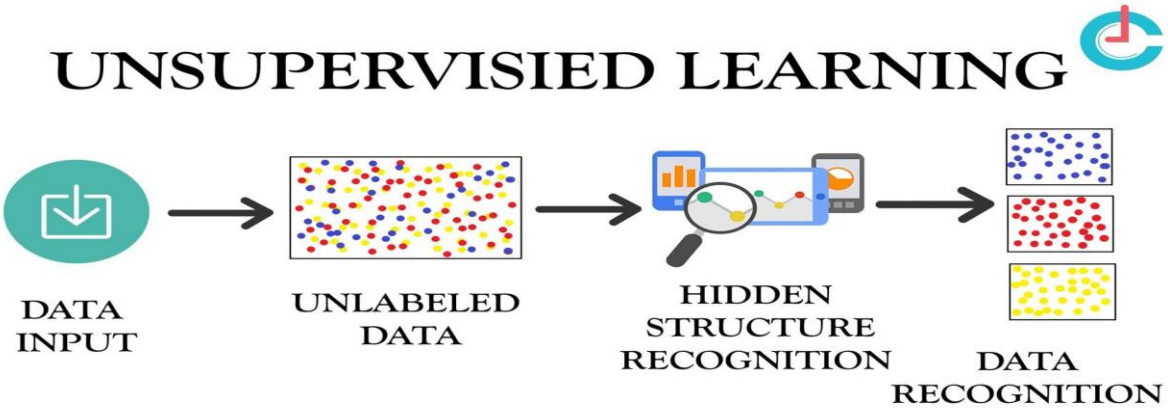
Supervised learning :

This means that a model learns from previous examples and is trained on labeled data. In other words, the dataset has tags that tell the model which patterns are related to fraud and which represent normal behavior.



Unsupervised learning :

This is also called anomaly detection as it automatically captures unusual patterns. In this case, training datasets come without any labels or instructions.



PROBLEM DEFINITION:

Online payment does not require physical card.

Anyone who know the details of card can make fraud transactions.

Currently,card holder comes to know only after the fraud transaction is carried out.

No mechanism to track the fraud transaction.

SOLUTION:

To determine the given transaction is fraud or not

Hidden Markov Model works on the basis of a spending habit of a user.

Classifies the user.

IMPLEMENTATION:

HTML & CSS – Interface designing

Javascript – Client side validation

PHP – Server side script

MYSQL – Database

ADVANTAGES:

Reduction in number of fraud transcation

User can use safely the credit or debit card

Added layer of the security