

FRAUD DETECTION

In Financial Transaction

Team Memembrs

- MAI KAMEL
- BASMALA AHMED
- TASNEEM SALAMA
- MAI MOHAMMED

Group Code

CAI1_AIS5_S8E

Technical Instructor

ENG : ESLAM ELREEDY

Training company

AST

PROBLEM STATEMENT

The exponential growth of digital financial transactions has led to a parallel rise in fraudulent activities, making fraud detection a critical issue for the financial industry.

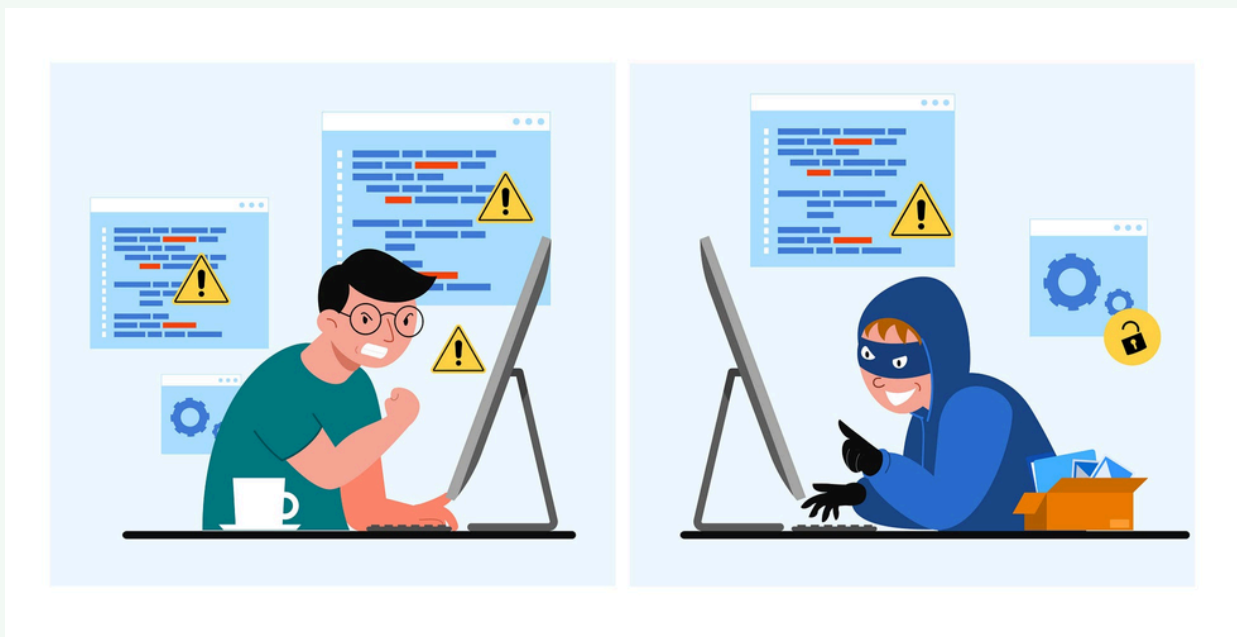
As online payment systems expand and evolve, so do the methods employed by fraudsters, posing significant challenges to the security and integrity of financial operations.

Detecting and preventing fraud is essential not only to safeguard customer assets and privacy but also to maintain the credibility and trust in financial institutions.

The growing sophistication of fraud tactics has made advanced fraud detection systems essential.

Machine learning-powered models offer significant improvements by analyzing vast transactional data in real-time, adapting to new patterns, and providing more accurate, proactive fraud prevention.

This results in faster detection and fewer false positives compared to traditional rule-based approaches.



KEY SUCCESSES



Fraud detection using credit card



Fraud detection using text mail

This project integrates various frameworks like NumPy, Pandas, TensorFlow, and to data analysis, model building, and deployment.

It begins with data cleaning to address missing values, duplicates, and outliers, followed by preprocessing steps such as feature scaling and encoding.

Exploratory Data Analysis (EDA) with Matplotlib and Seaborn guides model selection,

With machine learning models like Logistic Regression, Decision Trees, and Random Forest evaluated for optimal performance.

For text data, NLP techniques using NLTK or spaCy are applied, while TensorFlow's GANs generate synthetic data for testing.

MLflow manages the entire machine learning lifecycle, ensuring efficient model deployment and scalability.

TIMELINE

Gather datasets from platforms like Kaggle, Hugging Face, and other sources.

Clean and preprocess the data, handling missing values, scaling, and encoding.

Choose the best model (Logistic Regression, Decision Tree, Random Forest) based on data analysis.

Train the model on Colab GPU, test and optimize its performance.

Implement a Generative Adversarial Network (GAN) to generate synthetic data that enhances model testing and validation, ultimately improving its robustness and performance.

Deploy the model via Streamlit for user interaction and monitor its real-world performance.

01

Data Collection

02

Data Preparation

03

Model Selection

04

Model Training & Testing

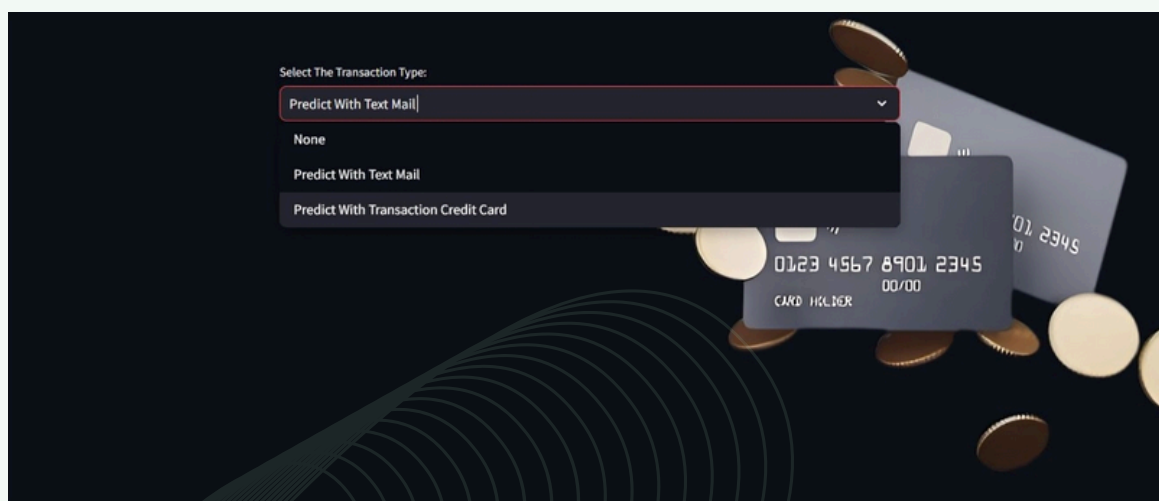
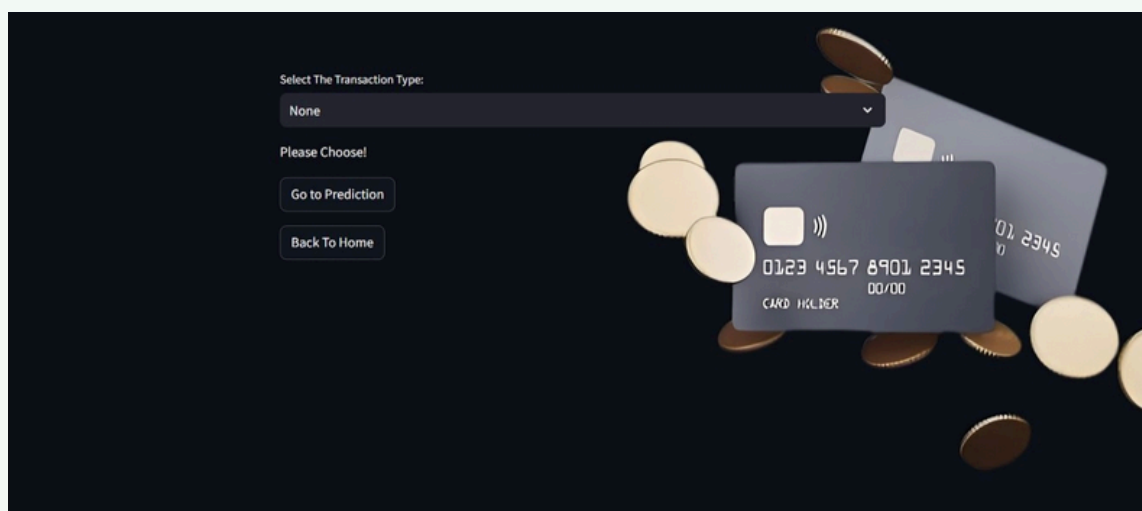
05

GANs for Data Generation

06

Model Deployment

OUTPUTS SCREEN



OUTPUTS SCREEN

Predict With Text Mail

Enter the email content for fraud prediction:

Hi, I hope you are well. I just wanted to follow up after the viewing we had on Friday at the Royal Winchester House. Have you made a decision on whether to proceed or not?

Predict

This transaction seems legitimate.

Back

Predict With Text Mail

Enter the email content for fraud prediction:

Congratulations! You've won a prize! Claim your reward now by clicking this link.

Predict

This transaction is likely to be fraudulent.

Back

Predict With Transaction Credit Card

Year

2023

Month

12

Day

17

Hours

2.00

Amount

0.00

Zip

10000

OUTPUTS SCREEN

MCC
4000

Select the transaction type:
Swipe Transaction

Select the Merchant City:
La Verne

Select the Merchant State:
CA

Select the Notes Error:
Technical Glitch

Predict

This transaction seems legitimate.

Back

Predict With Transaction Credit Card

Year
2023

Month
6

Day
25

Hours
6.00

Amount
150000.00

Zip
89109

MCC
7995

Select the transaction type:
Swipe Transaction

Select the Merchant City:
Las Vegas

Select the Merchant State:
NV

Select the Notes Error:
Bad CVV

Predict

This transaction is likely to be fraudulent.

Back