رواد مصر الرقمية

وزارة الاتصالات
وتكنولوجيا المعلومات

FRAUD DETECTION

In Financial Transaction

وزارة الاتصالات
وتكنولوجيا المعلومات

رواد مصر الرقمية

Team Memebrs

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

Group Code

CAI1_AIS5_S8E

Technical Instructor

ENG : ESLAM ELREEDY

Training company

AST

رواد مصر الرقمية

وزارة الاتصالات
وتكنولوجيا المعلومات

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.

رواد مصر الرقمية

وزارة الاتصالات وتكنولوجيا المعلومات

TIMELINE

01 Data Collection

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

02 Data Preparation

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

03 Model Selection

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

04 Model Training & Testing

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

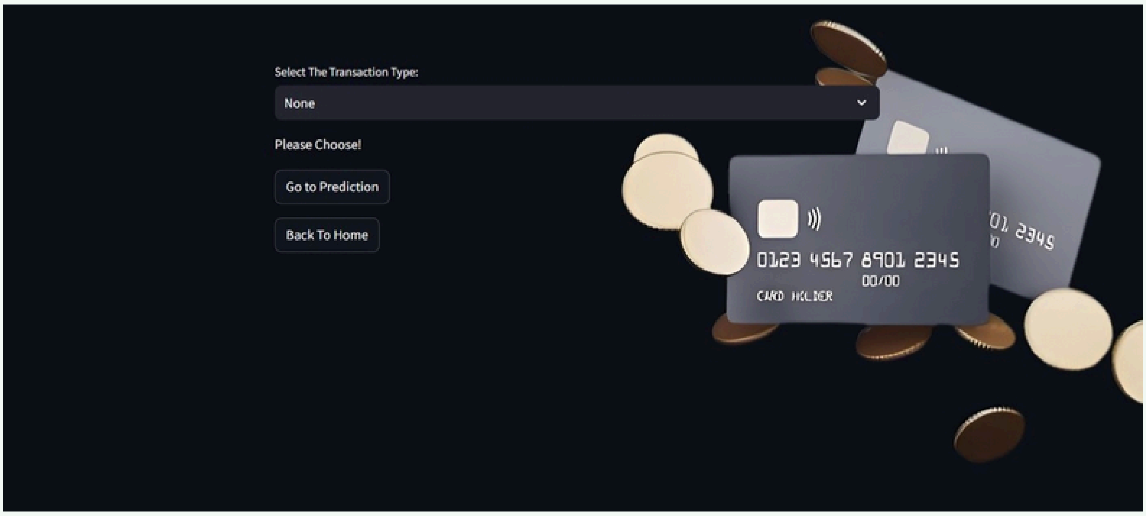
05 GANs for Data Generation

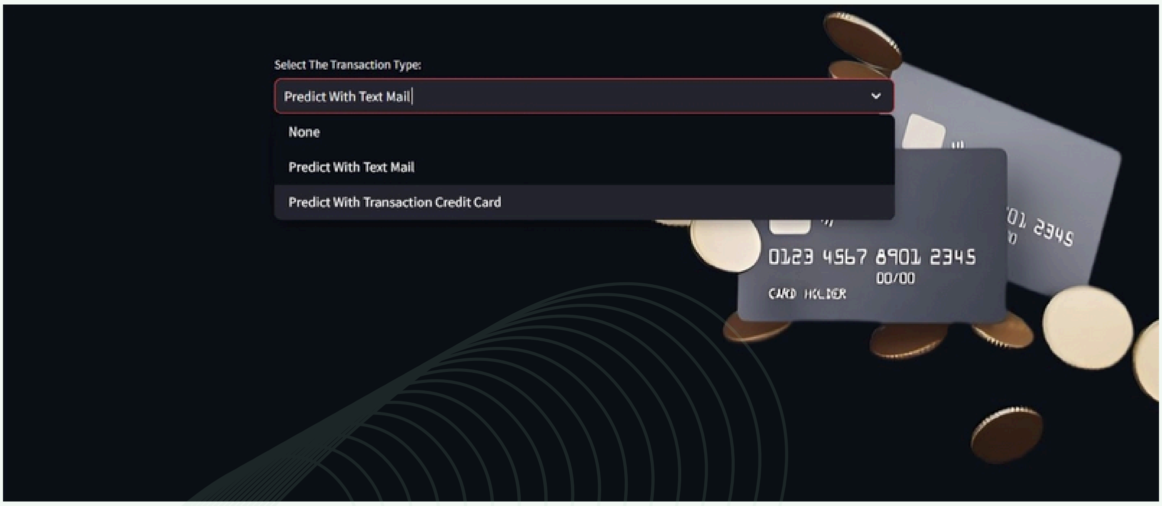
Implement a Generative Adversarial Network (GAN) to generate synthetic data that enhances model testing and validation, ultimately improving its robustness and performance.

06 Model Deployment

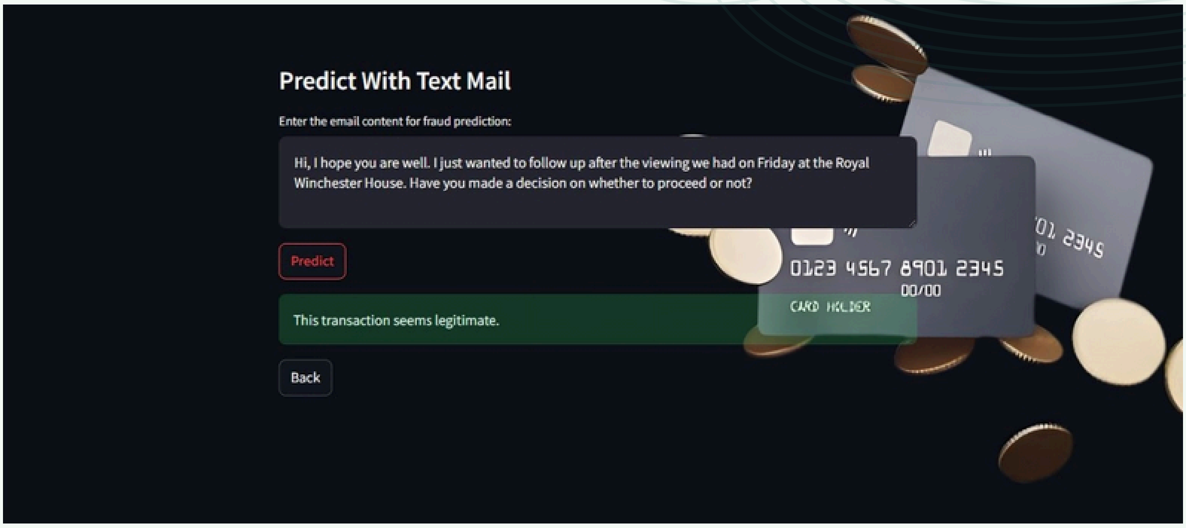
Deploy the model via Streamlit for user interaction and monitor its real-world performance.

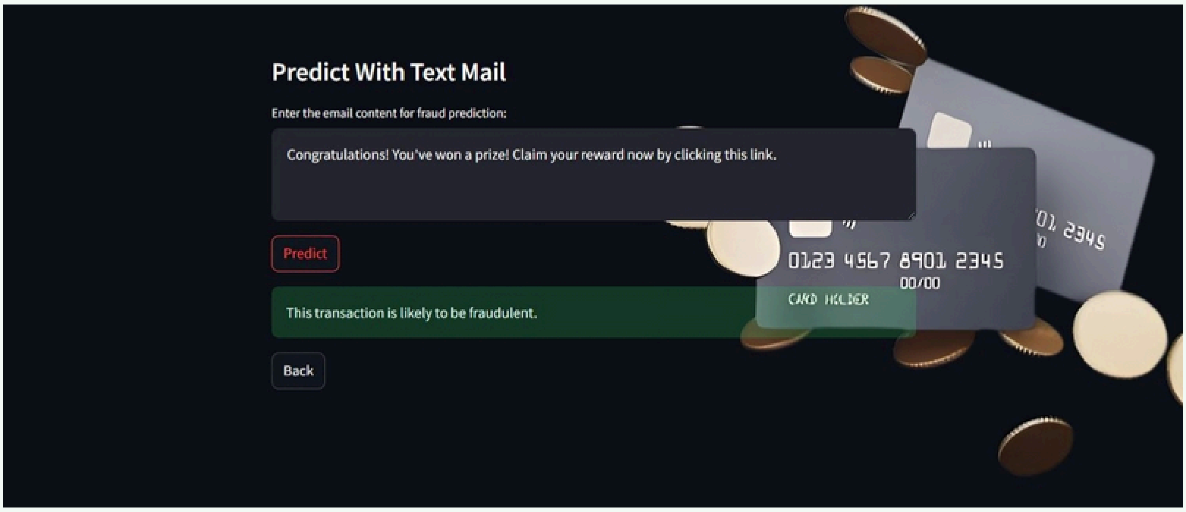
OUTPUTS SCREEN

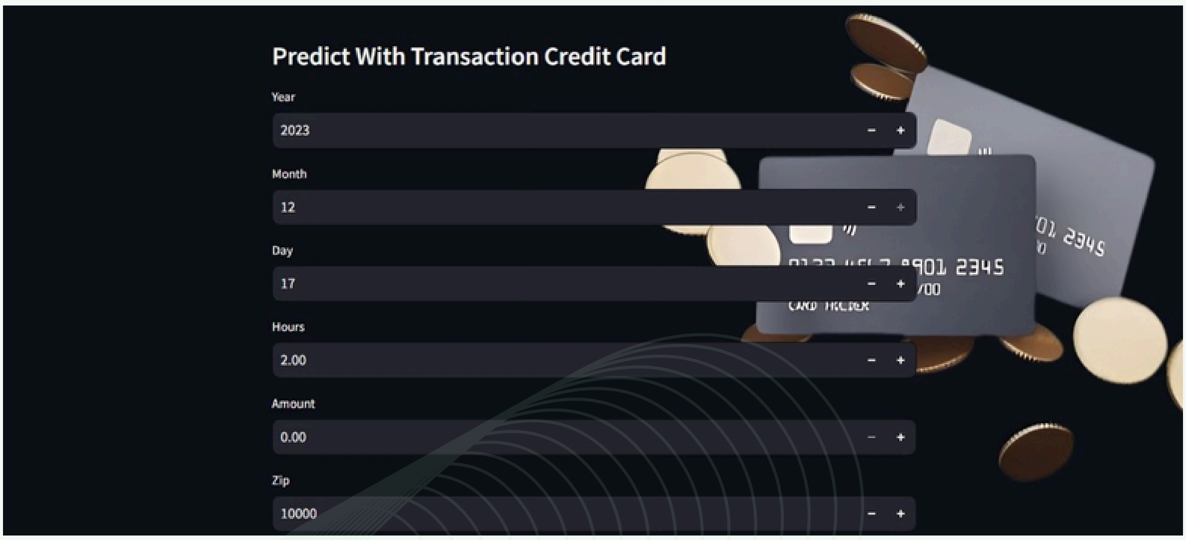




OUTPUTS SCREEN







OUTPUTS SCREEN