Credit Card Fraud Detection

Introduction

Credit card fraud is a growing concern in the digital economy. Detecting fraudulent transactions efficiently helps financial institutions and customers avoid significant monetary losses. This project focuses on identifying fraudulent transactions using machine learning techniques.

Abstract

This project uses a dataset containing transactional information to classify whether a given transaction is fraudulent or not. We use preprocessing techniques to encode categorical data, and apply XGBoost, a powerful machine learning classifier, for prediction. Evaluation metrics such as the confusion matrix and ROC curve are used to measure the model's performance.

Tools Used

- Python
- Pandas & NumPy
- Scikit-learn
- XGBoost
- Matplotlib
- Streamlit
- Git & GitHub

Steps Involved in Building the Project

1. Load and explore the dataset.
2. Select relevant features: category, amount, gender, and city population.
3. Encode categorical data using Label Encoding.
4. Split the dataset into training and testing sets.
5. Train an XGBoost classifier.
6. Evaluate the model using confusion matrix and classification report.
7. Plot the ROC curve to assess model performance.
8. Save the trained model and integrate with Streamlit app for user interaction.
9. Deploy the app using a user-friendly web interface.

Credit Card Fraud Detection

Conclusion

The model successfully identifies fraudulent credit card transactions using features derived from the dataset. XGBoost performs well with strong classification metrics and AUC scores. A Streamlit web interface allows users to input transaction details and receive real-time predictions. This project demonstrates the potential of machine learning in combating financial fraud.