Credit Card Fraud Detection System

Project Report

Introduction

Credit card fraud detection is a critical application of machine learning in the financial sector, where billions of dollars are lost annually due to fraudulent transactions. This project develops an intelligent fraud detection system capable of identifying suspicious credit card transactions in real-time while providing interpretable insights to help understand the factors contributing to fraud predictions.

The primary objective was to create a robust, accurate, and explainable machine learning model that can distinguish between legitimate and fraudulent transactions using features such as transaction amount, location, time, merchant category, and user demographics.

Abstract

This project presents a comprehensive credit card fraud detection system built using machine learning with explainable AI capabilities. The system processes transaction data through a pipeline including data preprocessing, class balancing using SMOTE, feature engineering, and XGBoost classifier training.

The final model achieves exceptional performance with 99% precision, recall, and F1-score on both fraud and legitimate transactions. The system incorporates SHAP for model interpretability, revealing that transaction amount, geographical coordinates, and transaction category are the most influential features. A Streamlit web application provides real-time fraud prediction with visual explanations and trend analysis.

Tools Used

Core Technologies:

- **Python 3.x** with Jupyter Notebook/Google Colab
- **XGBoost**: Primary classification algorithm
- **scikit-learn**: Data preprocessing and model evaluation
- **imbalanced-learn**: SMOTE implementation for class balancing
- **SHAP**: Model explainability and feature importance
- **Streamlit**: Interactive web application framework
- **Pandas/NumPy**: Data manipulation and numerical computations
- **Matplotlib/Seaborn**: Data visualization

- **Joblib**: Model serialization and deployment

Steps Involved in Building the Project

1. Data Preprocessing and Balancing

Challenge: Original dataset had severe class imbalance (866 fraudulent vs 200,922 legitimate transactions - 0.43% fraud rate).

Solution: Applied strategic undersampling by selecting 10x fraud samples from legitimate class, creating a balanced dataset with 9,526 total samples. Removed missing values and engineered temporal features (hour, day) from transaction timestamps.

2. Feature Engineering and Encoding

Extracted key features: amount, gender, city population, category, latitude, longitude, hour, and day. Applied one-hot encoding to categorical variables (gender, transaction category) ensuring consistent feature representation for model input.

3. Advanced Class Balancing with SMOTE

Applied SMOTE after initial preprocessing to create synthetic fraud samples, achieving equal representation of both classes (5,196 samples each). This generated realistic synthetic samples based on feature space neighborhoods.

4. Model Training and Evaluation

Algorithm: XGBoost classifier chosen for superior tabular data performance. **Training:** 70% training, 30% testing with stratified sampling (7,274 training, 5,196 testing samples). **Results:** Achieved exceptional 99% precision, recall, and F1-score for both classes.

```
Performance Metrics:
              precision    recall  f1-score   support
Legitimate       0.99      0.99      0.99      2598
Fraudulent       0.99      0.99      0.99      2598
Accuracy                             0.99      5196
```

5. Model Interpretability with SHAP

Integrated SHAP explainer for model transparency, calculating feature importance using mean absolute SHAP values. Key insights revealed:

- **Transaction Amount**: Most influential feature for fraud detection

- **Geographical Location (lat/long)**: Strong predictors indicating location-based patterns
- **Transaction Category**: Certain merchant categories show higher fraud propensity
- **Temporal Features**: Time-based patterns influence fraud likelihood

6. Web Application Development

Developed comprehensive Streamlit interface featuring:

- **Real-time Prediction**: Interactive input form with 0.4 probability threshold
- **Visual Explanations**: SHAP-based feature importance charts
- **Fraud Analytics**: Historical trend analysis by hour and day
- **User Experience**: Intuitive design with clear indicators (✅ legitimate, 🚨 fraud)

Conclusion

This fraud detection system successfully demonstrates advanced machine learning application to financial security challenges. Key achievements include:

Technical Success:

- Developed highly accurate model (99% across all metrics) for real-time fraud detection
- Successfully addressed class imbalance through strategic sampling and SMOTE
- Implemented comprehensive explainability using SHAP for stakeholder trust
- Created end-to-end pipeline from preprocessing to deployment

Business Impact:

- Provides financial institutions with reliable fraud prevention tool
- Reduces false positives through balanced training approach
- Offers interpretable predictions supporting decision-making processes
- Enables real-time transaction monitoring with immediate risk assessment

The modular system architecture supports easy integration of additional features and model updates while maintaining the existing interface, demonstrating that effective fraud detection requires high-performing algorithms, thoughtful data handling, interpretable results, and user-friendly interfaces.

Project Completion Date: 25th July 2025
Technologies: Python, XGBoost, SHAP, Streamlit, scikit-learn