Project Proposal: Banking Fraud Detection and Real-Time Transaction Monitoring

Introduction

In this capstone I will primarily focus on designing and implementing an intelligent system that is capable of detecting fraudulent transaction for online banking transactions. I will be leveraging a large-scale bank transaction dataset from Deutsche Bank which constitute about 6 million records. The features of this dataset have been masked to preserve anonymity of the actual customer data. The objective will be to accurately predict fraudulent transactions in real time by evaluating several machine learning and deep learning algorithms and selecting the best performing model that optimizes certain classification metrics like recall, precision and f1-score.

Problem Statement

Fraudulent banking transactions presents sever financial risks and reputational damage to financial institutions. Online payments have rapidly expanded in the last couple of years and continues to as it becomes the norm in this modern age. Detecting anomalies in transaction pattern is becoming more complex due to data imbalances and the dynamic nature of fraud. The primary challenge in this capstone will be developing a machine learning model that can accurately identify fraudulent activities with high recall and precision, thereby minimizing false negatives and safeguarding customer assets.

Project Objectives and Methodology

The project prototype and final paper will comprehensively cover all four curricular areas to build an operational model capable of processing transaction data in real time. Using data analytics and data management, I will be conducting an extensive exploratory data analysis (EDA) to understand feature correlations and transaction patterns. The large dataset will be effectively handled and preprocessed by implementing data cleaning, sampling (reducing over 6 million records to a manageable representative subset) and transforming categorical variables to ensure data integrity. Effective data storage and retrieval methods will support the high-volume transaction records. This will be developed using python and jupyter notebooks. This will also include integrating machine learning libraries (e.g., scikit-learn, TensorFlow for deep learning). The system will support future expansion and potential deployment into a production environment. Addressing the secure handling of sensitive financial data is paramount. The final paper will include an in-depth analysis of real-time alert mechanisms, transmission protocols and strategies for defending against adversarial attacks.

Conclusion

This project presents a technically rigorous and interdisciplinary approach to solving a critical problem in financial fraud detection. The prototype will serve as a proof of concept, while the final paper will offer a comprehensive academic analysis across all curricular areas.