Innovative Fraud Detection in Mobile Money Systems for Low-Resource Settings: A Machine Learning Approach Using Synthetic Data

Authors:

- Yiga Gibert
- Muhindo Mubaraka

Presented by:

Group N

Makerere University, Kampala

Uganda.

Contents of Presentation

Introduction to Fraud Detection in Mobile Money Systems

- Emphasize mobile money as a critical financial service in Africa, where access to traditional banking is limited.
- Describe the unique challenges of fraud detection in low-resource environments, such as limited data availability and high costs of real-time monitoring.
- State that due to the scarcity of real-world data, synthetic datasets like PaySim offer valuable insights into fraud patterns, especially in resource-constrained regions.

Research Problem Definition

- Predicting fraud in mobile money transactions is difficult due to the class imbalance problem and incomplete real-world data.
- Existing fraud detection models often rely on artificially balanced datasets, which may not translate well to real-world scenarios.
- There is a need for a robust model that can effectively identify fraudulent transactions without relying on data balancing techniques.

Research Aims & Objectives

- Objective 1: Develop a machine learning framework tailored to fraud detection in mobile money systems, specifically addressing low-resource environments.
- Objective 2: Implement feature-engineering techniques to enhance fraud detection capabilities, focusing on features indicative of fraudulent patterns in synthetic datasets.
- Objective 3: Validate the approach through precision-recall metrics, aiming to demonstrate applicability and scalability in real-world scenarios with limited data.
- Objective 4: Contribute insights into creating accessible, low-cost fraud detection solutions suitable for deployment in African mobile money platforms.

Research Approach/Methodology

- Exploratory Data Analysis
  - Identified key features differentiating fraud and non-fraud transactions
  - Engineered new features like "errorBalanceOrig" and "errorBalanceDest"
- Machine Learning Model
  - Trained an XGBoost classifier on the cleaned dataset
  - Optimized model hyperparameters for best performance
- Interpretability Techniques
  - Leveraged SHAP and LIME to explain model predictions

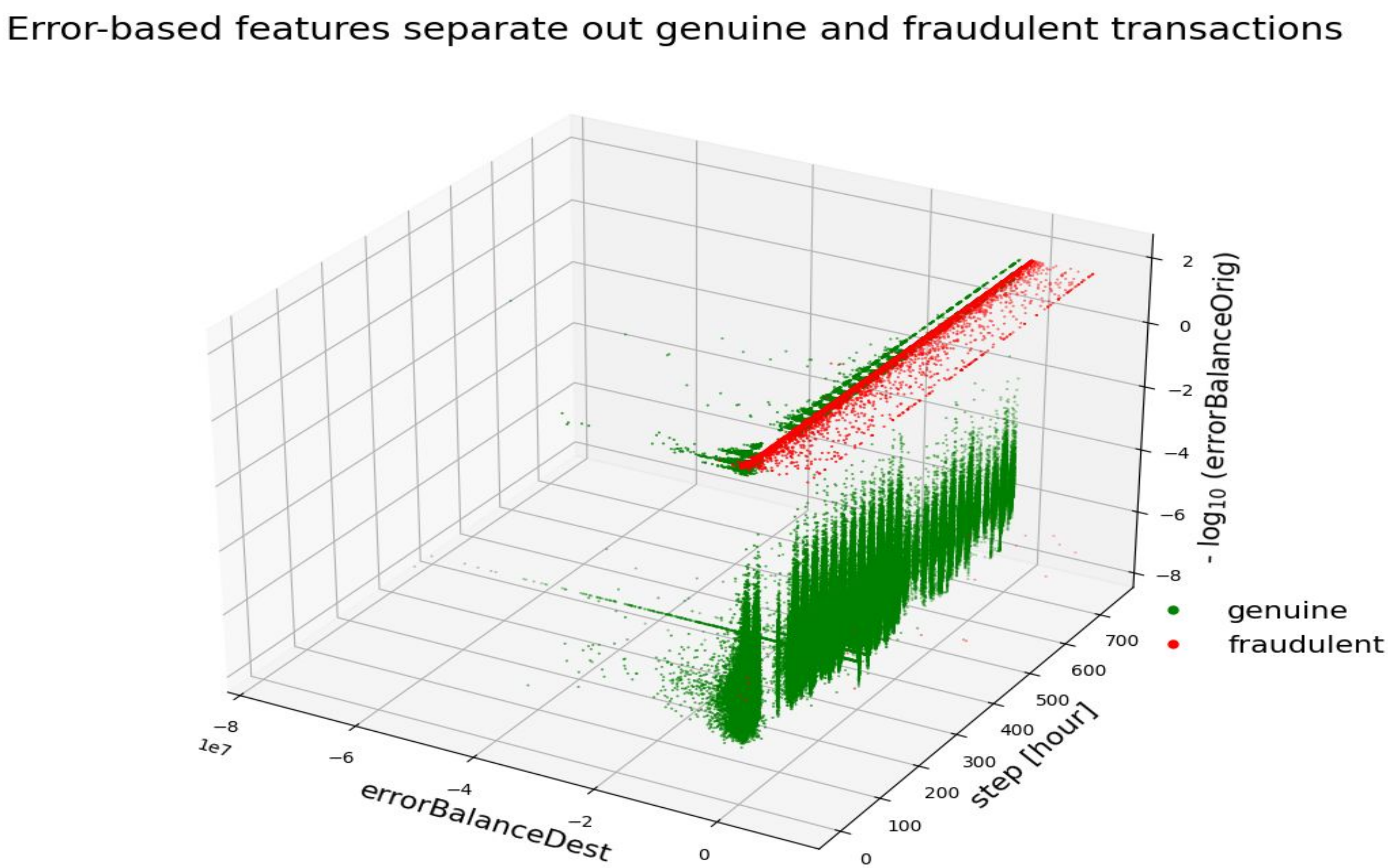
Key Findings and Data Insights.

- Fraud primarily occurs in "TRANSFER" and "CASH_OUT" transaction types
- The "isFlaggedFraud" feature was found to be unreliable
- New engineered features like "errorBalanceOrig" and "errorBalanceDest" were effective at separating genuine from fraudulent transactions

Exploratory Data Analysis

- Fraudulent transactions show more homogenous distribution over time compared to genuine transactions
- Fraudulent transactions have distinct patterns in terms of transaction amount and errors in destination account balances
- 3D visualization using engineered features clearly separates genuine and fraudulent transactions
- Correlation heatmaps reveal different fingerprints for genuine and fraudulent transactions

Key Findings and Data Insights.



Key Findings and Data Insights.

Model Implementation

- Having obtained evidence from the plots above that the data now contains features that make fraudulent transactions clearly detectable, the remaining obstacle for training a robust ML model is the highly imbalanced nature of the data.

skew = 0.002964544224336551

Selection of metric

- Since the data is highly skewed.

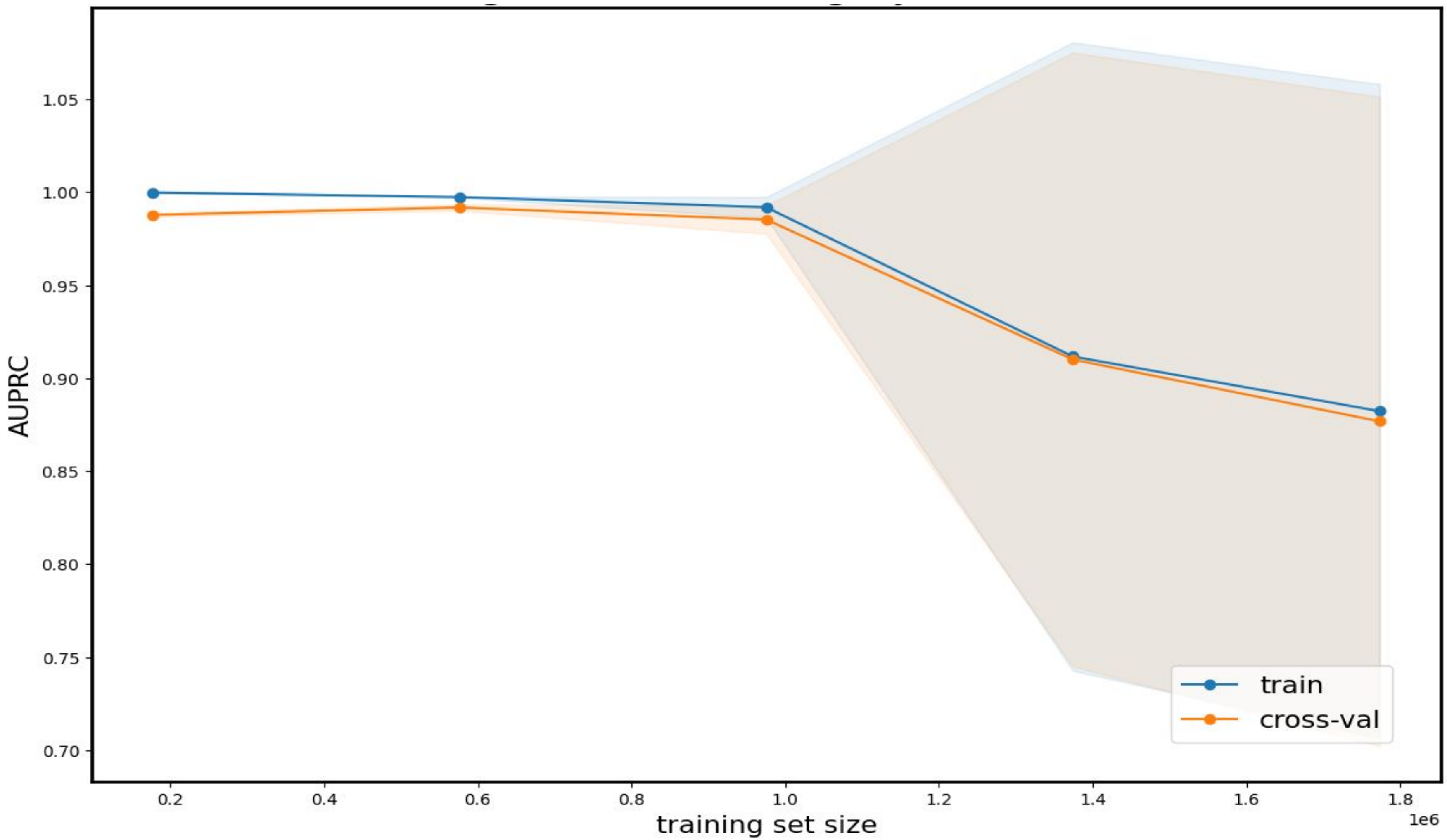
- I use the area under the precision-recall curve (AUPRC) rather than the conventional area under the receiver operating characteristic (AUROC).

- This is because the AUPRC is more sensitive to differences between algorithms and their parameter settings rather than the AUROC.

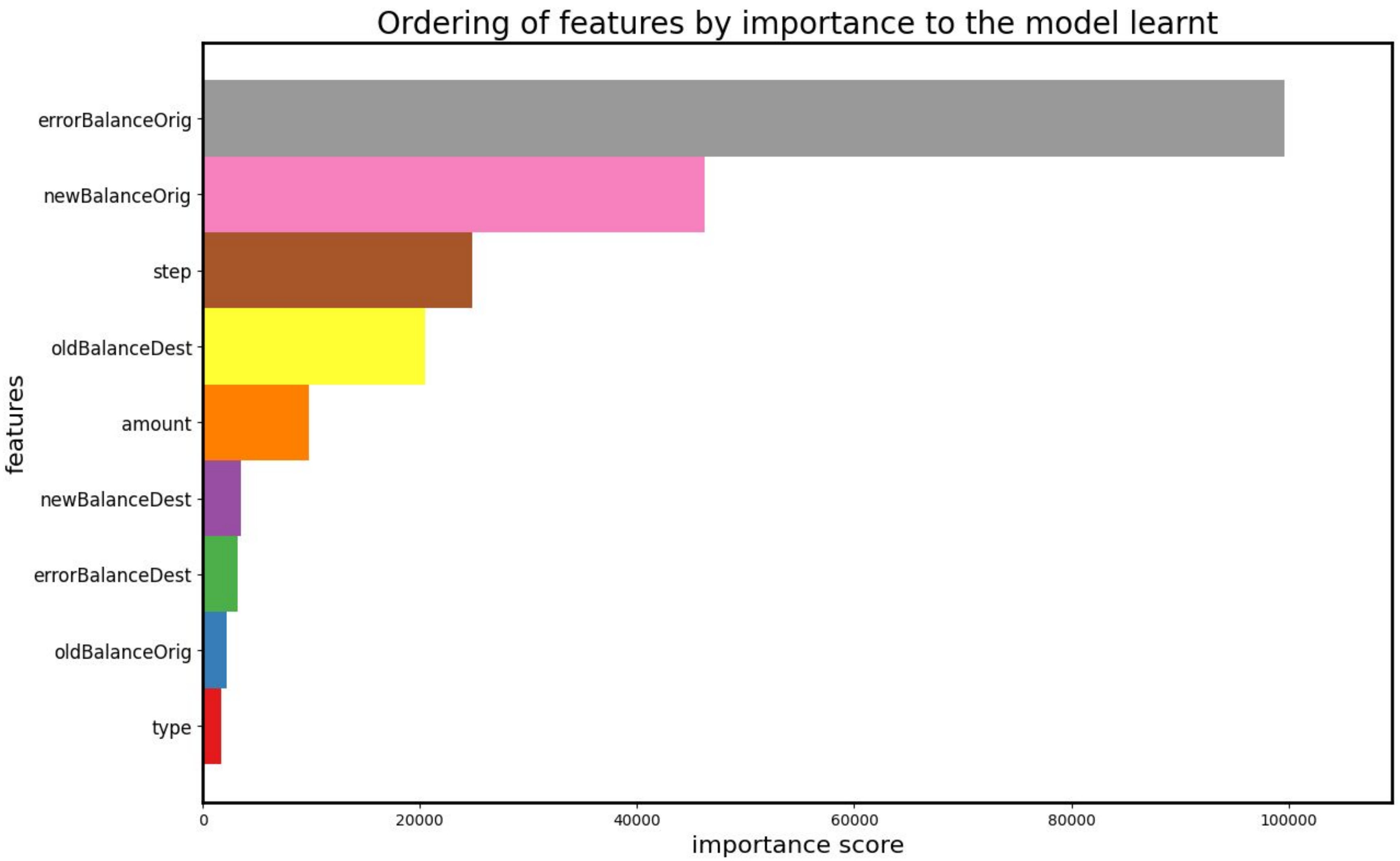
Model Implementation

- The XGBoost model achieved an AUPRC(area under the precision-recall curve) ~0.9 on the test set.
- "errorBalanceOrig" was the most important feature for the XGBoost model.
- Learning curves indicated the model was slightly underfit.

Model Implementation



Key Findings and Data Insights.



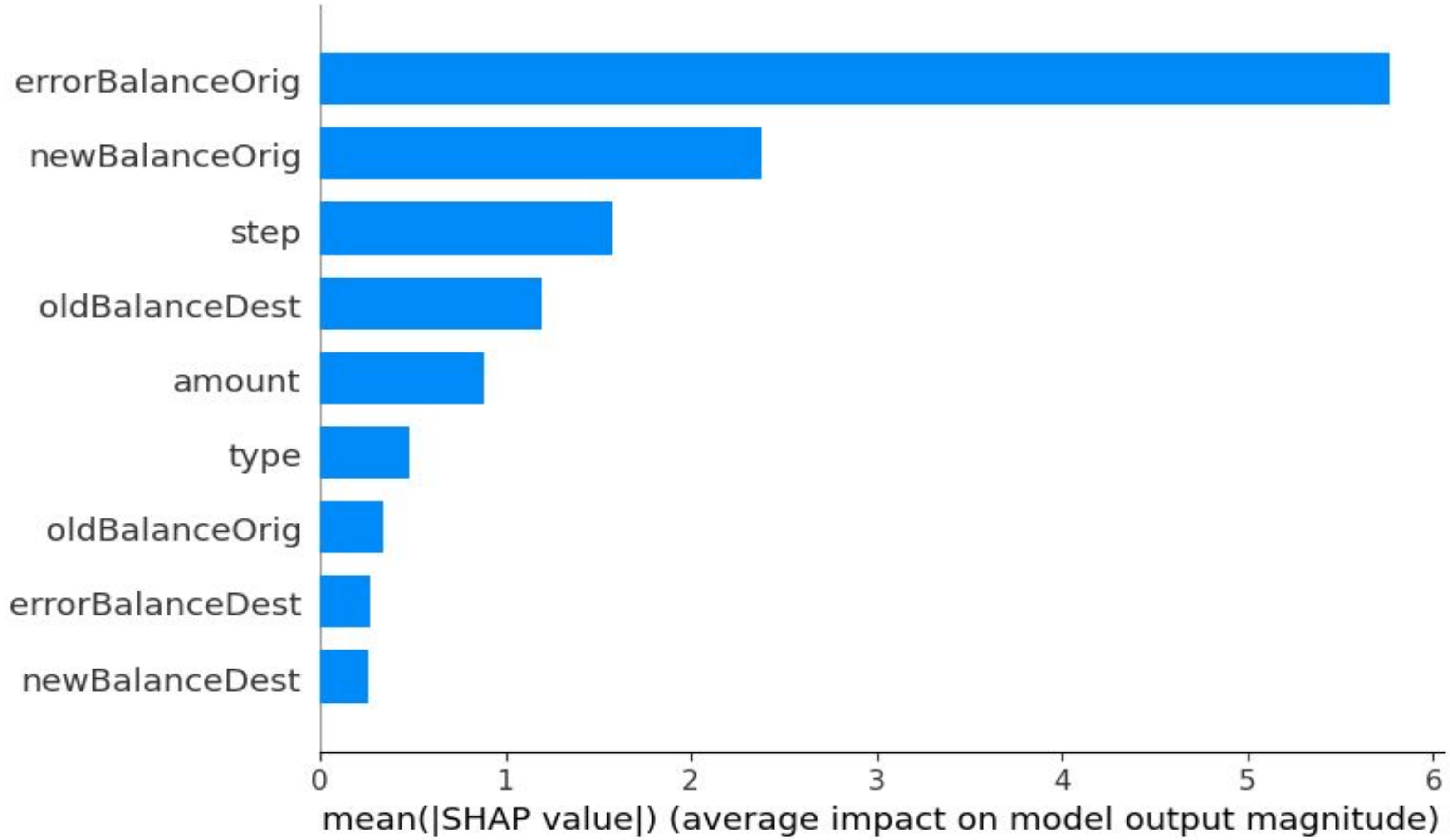
Model Implementation

- The classification report showed strong precision, recall, and F1-score for both classes.
- SHAP and LIME analyses provided insights into feature importance and local explanations.

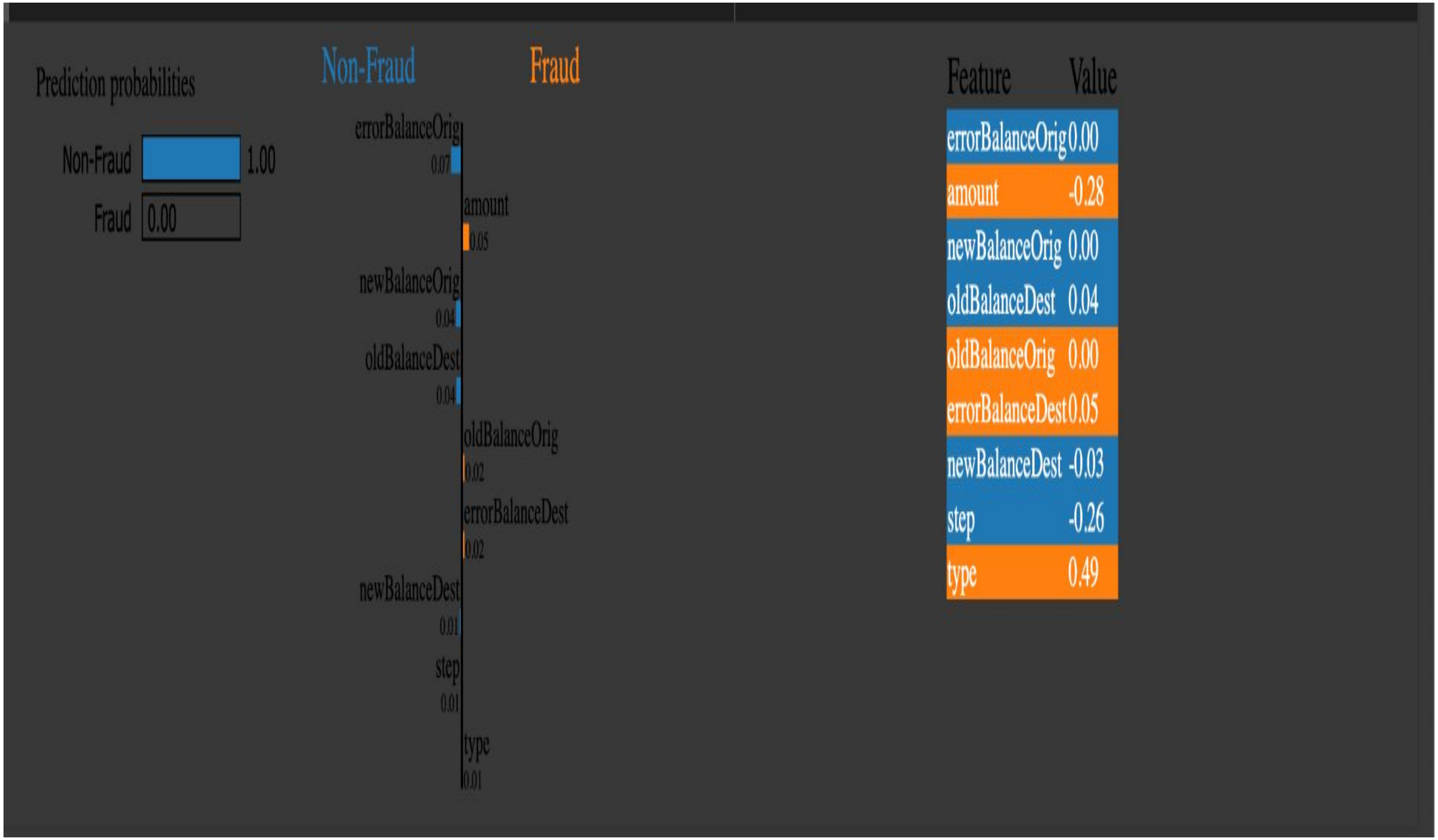
Model Implementation

	precision	recall	f1-score	support
Non-Fraud	1.00	1.00	1.00	552412
Fraud	0.85	1.00	0.92	1670
accuracy			1.00	554082
macro avg	0.92	1.00	0.96	554082
weighted avg	1.00	1.00	1.00	554082

SHAP Feature importance



LIME Feature importance



Limitations and Future Work

- Limited by the reliability of the "isFlaggedFraud" feature in the dataset
- Explore additional feature engineering and model tuning to further improve performance.
- Investigate the root causes behind the identified fraudulent transaction patterns.