Nigerian Banking Fraud Detection System

Course: DA 204o - Data Science in Practice
Institution: Indian Institute of Science (IISc), Bangalore
Team: Data Conduits
Team Members: Novoneel Chakraborty <cnovoneel@iisc.ac.in>, Sarthak Sharma <sarthak1@iisc.ac.in>, Swarup E. <swarupe@iisc.ac.in>, Rakshit Ramesh <rrakshit@iisc.ac.in>

Overview

Financial fraud in Nigeria is not merely a technical nuisance; it is a systemic threat that cost the banking sector over ₦17.67 billion in 2023 alone. Our project, undertaken by Team Data Conduits, sought to confront this challenge by building a robust machine learning system capable of identifying fraudulent transactions within the Nigeria Inter-Bank Settlement System (NIBSS) data.
We successfully developed an ensemble model—combining Random Forest, XGBoost, LightGBM —that achieves an AUC-ROC of 0.9638 and an F1 score of 0.8847. Most importantly, we addressed the industry's biggest pain point: customer friction. By optimizing our decision thresholds, we reduced the False Positive Rate (FPR) to a negligible 0.001%, ensuring that for every genuine transaction blocked, thousands are processed smoothly.

1. Introduction and Problem Context

The background of this project is rooted in the alarming escalation of financial crime in Nigeria. As indicated in the NIBSS Annual Fraud Landscape Report, fraud losses surged by 23% in a single year, driven largely by social engineering and mobile channel vulnerabilities. Traditional rule-based systems—which rely on static if-then logic—have proven ineffective against these evolving threats, often failing to catch sophisticated attacks or, conversely, flagging too many legitimate users.
We aimed to create a model that could process transaction features such as time, amount, and channel to predict the probability of fraud in real-time. This required not just high accuracy, but a specific focus on precision, as the reputational cost of blocking a legitimate customer is often higher than the financial cost of missing a minor fraud event.

2. Dataset Characteristics and Challenges

We utilized the NIBSS transaction dataset, which presented a classic, albeit extreme, case of class imbalance. Out of approximately 1,000,000 transactions, only about 3000 were confirmed as fraudulent—a prevalence rate of just **0.3%**. This scarcity of "positive" examples meant that a naive model could simply predict "Legitimate" for every single transaction and still achieve 99.7% accuracy, while being completely useless for fraud detection. Furthermore, the dataset required significant cleaning. We identified and removed administrative columns that offered no predictive value, and we had to rigorously handle missing values in demographic fields to prevent data leakage.

3. Exploratory Data Analysis (EDA)

Before modeling, we needed to understand the "behaviour" of fraud. Our exploratory analysis revealed several striking patterns that directly informed our feature engineering. First, we observed a strong temporal component to fraudulent activity. While legitimate transactions peaked during standard business hours (8 AM to 5 PM), fraud attempts remained dangerously high during the early morning hours (1 AM to 4 AM), a time when customers are less likely to notice alerts.

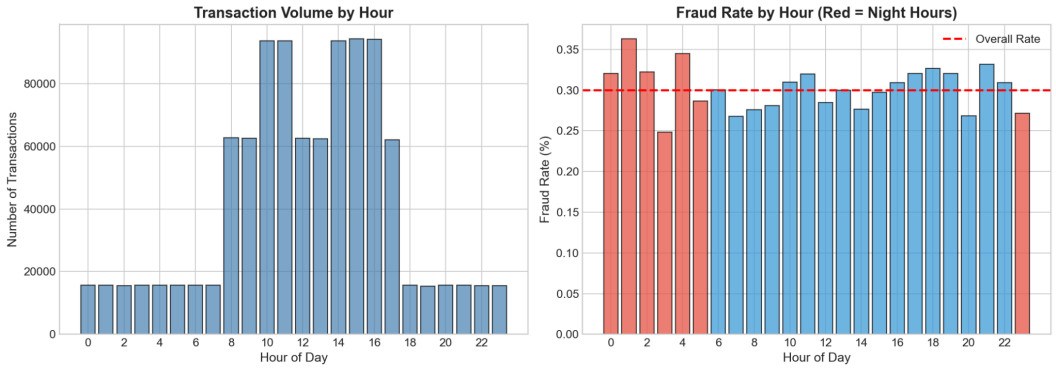

Figure 1: Hourly Transaction Volume

Secondly, the channel analysis confirmed our hypothesis regarding digital vulnerabilities. Mobile and Web channels showed a disproportionately higher rate of fraud compared to ATM or POS transactions. This suggests that fraudsters prefer remote channels where physical surveillance is absent.
We also analyzed the transaction amounts. We noted that the fraudulent transactions tended to involve higher amounts, and performing a log transformation revealed a clear separation between the legitimate and fraudulent transaction classes. Fraud transactions also showed a right-skewed distribution.
We also noted that fraud tended to happen more frequently in the larger cities of Nigeria like Lagos and Abuja, and less often in the suburban and rural areas. Correlating this with the most frequent fraud channel being Web and Mobile, we hypothesize that an increased adoption of technology and internet-based transactions in urban areas along with the higher population - meaning higher transaction count - contribute to the higher fraud count.

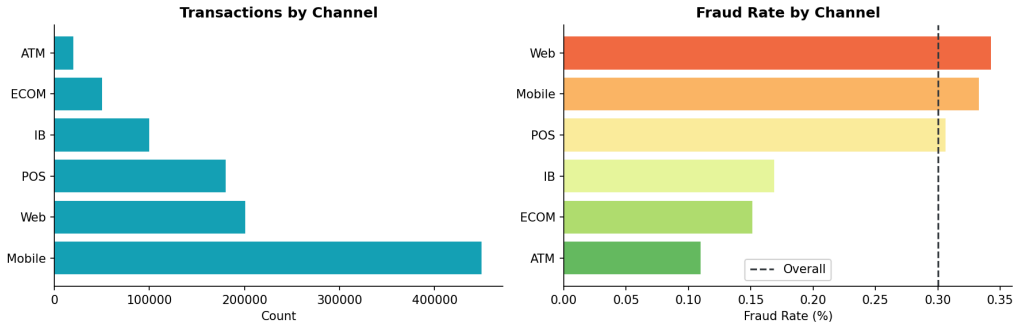

Figure 2: Transaction Count and Fraud Rate by Channel

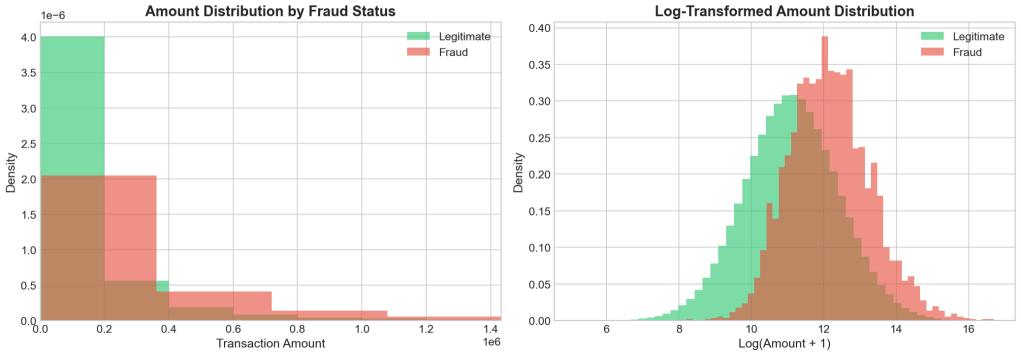

Figure 3: Transaction Count and Fraud Rate by Channel

4. Feature Engineering

We created a **Velocity Score**, which tracks the frequency of transactions within a short window. A sudden burst of activity is often a hallmark of an account takeover. Similarly, we developed an Amount-to-Mean Ratio, comparing the current transaction value to the user's historical average. This allowed the model to flag a ₦50,000 withdrawal as suspicious if the user typically withdraws ₦2,000, even if ₦50,000 is not objectively large for other customers.
Our most effective engineered feature was the Composite Risk Score. By aggregating risk weights from the transaction channel, time of day, and location, we created a single numerical vector representing the "situational danger" of the transaction. This feature proved to be one of the strongest predictors in our final model.

5. Methodology: The Ensemble Approach

Given the complexity of the data, relying on a single algorithm was deemed insufficient. We adopted an ensemble approach, training three distinct base models: Random Forest, XGBoost, LightGBM. We also explored using CatBoost but did not include it as part of our final ensemble. The ensemble voting weights were tuned using Optuna, and we also used an LSTM model to capture temporal features.

To handle the 0.3% fraud rate, we employed SMOTE (Synthetic Minority Over-sampling Technique). Unlike simple oversampling, which just duplicates existing fraud cases, SMOTE synthesizes new, plausible examples of fraud by interpolating between existing ones. We used a sampling strategy of 1.0, effectively balancing the training set to a 50/50 split. This forced the models to pay equal attention to fraud patterns during training.

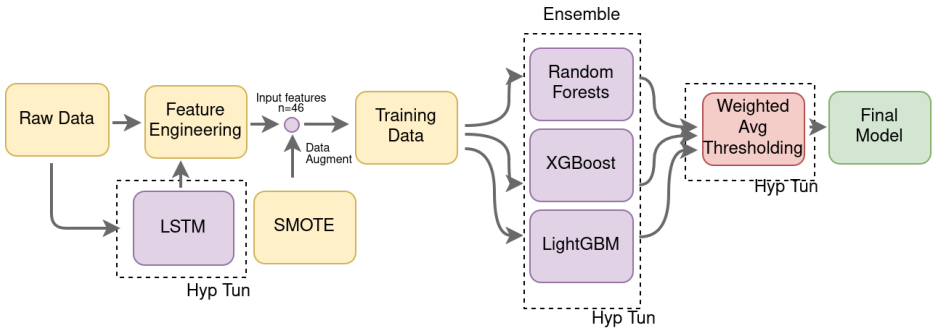

Figure 5: Ensemble Model Architecture

Our final prediction engine averages the probability outputs of these four models, with threshold optimisation. A standard probability threshold of 0.5 (50%) would have resulted in too many false alarms, so we iteratively tested thresholds to find the appropriate threshold that maximized the F1 score while keeping the False Positive Rate strictly below 0.1%.

6. Experimental Results and Evaluation

The performance of our system was evaluated using a hold-out test set of 20% of the data, ensuring that our metrics reflect real-world generalization rather than memorization. The ensemble model consistently outperformed the individual models on key metrics such as F1 score, AUC-ROC and Precision-Recall. We thus decided to use the Optuna-tuned ensemble over other models.
To this model, we incorporated an LSTM module to generate 64 additional temporal embeddings, increasing the feature space from 61 to 125 and creating a hybrid architecture designed to capture hidden temporal patterns in transaction behaviour. However, the overall performance improved only marginally—about 1% in both F1 and AUC-ROC—which we attribute to the limited dataset, where meaningful temporal trends were likely insufficient for the LSTM to exploit effectively. Using this final model, the results were exceptionally strong. The ensemble model achieved an F1 Score of 0.8847, narrowly missing our ambitious 0.90 target but providing a robust balance of precision and recall. Our AUC-ROC score of 0.9638 indicates that the model has a near-perfect ability to rank a fraudulent transaction higher than a legitimate one.

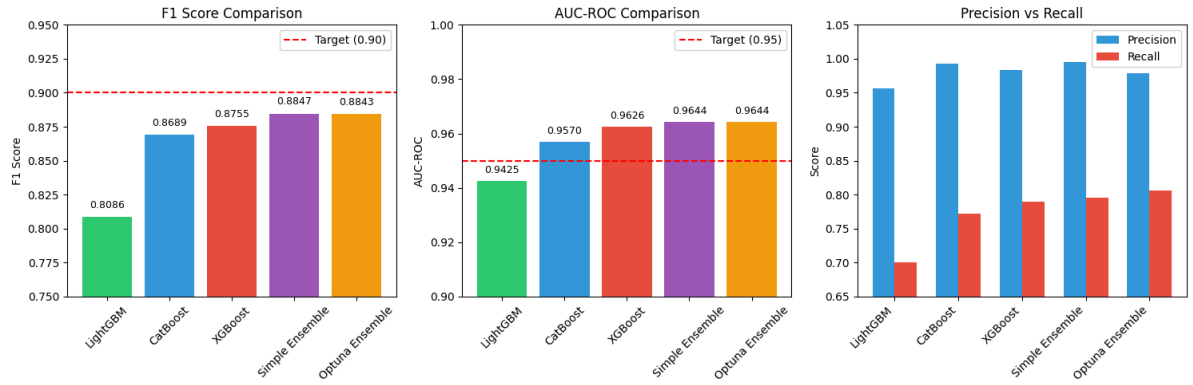

Figure 6: Comparison of Ensemble Performance with Individual Models

The most significant achievement, however, was the False Positive Rate (FPR) of 0.001%. In our test set of roughly 200,000 transactions, the system raised only 2 false alarms. For a bank processing millions of transactions, this low noise level is the difference between a usable product and one that gets turned off by frustrated operations teams.

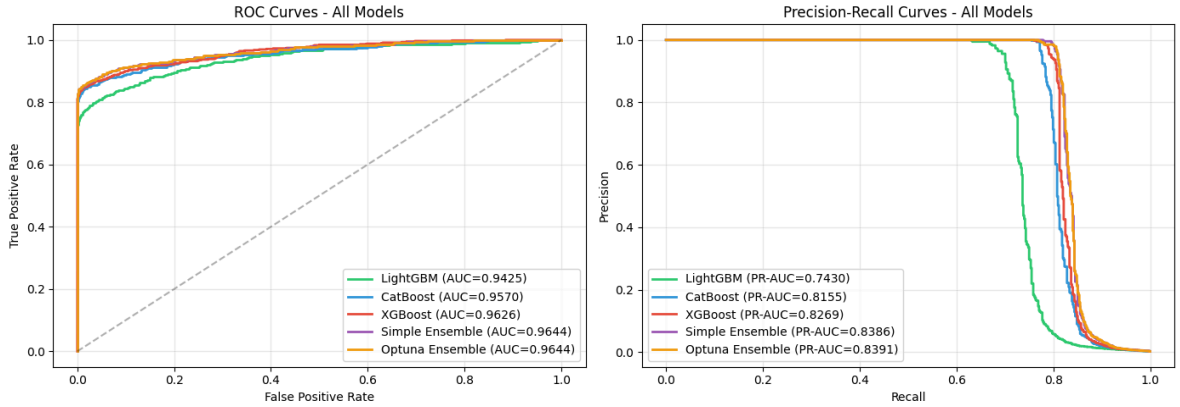

Figure 7: ROC Curve comparison

7. Conclusion and Future Directions

We successfully demonstrated that machine learning can drastically improve fraud detection capabilities for Nigerian banks. By moving away from static rules to a dynamic, feature-rich ensemble model, we achieved a high accuracy detection rate. Looking forward, the next logical step is Cost-Sensitive Learning. Currently, we treat all fraud equally, but missing a ₦10,000,000 fraud is far worse than missing a ₦500 one. Integrating transaction value directly into the loss function could further align the model with business priorities. Additionally, as originally proposed, exploring Graph Neural Networks (GNNs) could help uncover complex fraud rings that link multiple accounts, adding another layer of security to the Nigerian financial ecosystem.

8. Contributions

- Sarthak Sharma - EDA, Visualisation, Baseline model, Initial model eval
- Novoneel Chakraborty - Feature Engineering, XGBoost model, threshold optimisation, feature importance
- Rakshit Ramesh - LSTM model arch, focal loss, LSTM embedding extraction, models with LSTM features
- Swarup E. - Optuna Ensemble Optimization, SHAP interpretability, Streamlit demo, Final Comparison and documentation
- All - Data preprocessing and cleaning, model evaluation and result analysis, report writing and presentation, code review and debugging