BANK ANTI-FRAUD TRANSACTION DETECTION ANALYSIS : BASED ON THE XGBOOST MODEL

BUSINFO 704
2024 Quarter 3-Group 27:
Nikki Yao - syao713
Tiange Wang - twan545
Yongkang Li - yli688
Zheying Lu - zlu793

*The data used for this project were synthetically generated.

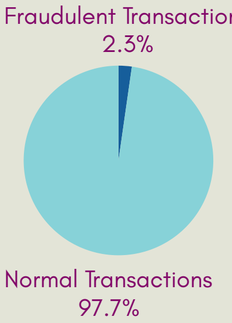
INTRODUCTION

ASB Bank, a major financial institution in New Zealand, is committed to secure banking services. However, the rise of online and mobile banking has led to increasingly sophisticated fraudulent transactions, causing direct monetary losses and eroding customer trust. This project focuses on developing a robust classification model to accurately identify fraudulent transactions and provide recommendations to enhance ASB's fraud detection capabilities, reduce losses, and rebuild customer confidence.

OVERVIEW

- 50,0000 Transactions
- 10,0000 Customers
- 2024.05-2024.07

In the dataset, 2.3% of the transactions are fraudulent, while 97.7% are normal:



METHODOLOGY

- ***Assumptions made***

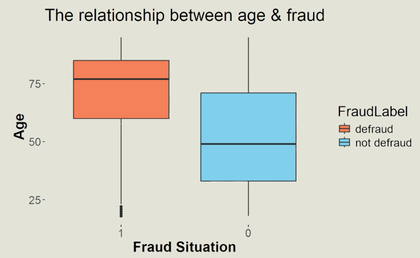
Independence of Data Records

Representative Sample

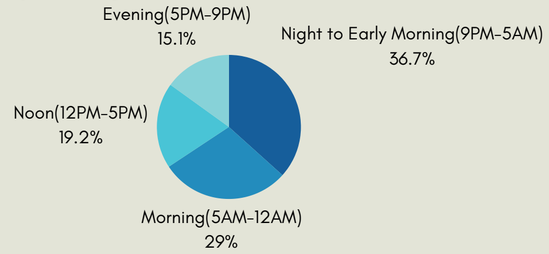
Model Evaluation Consistency

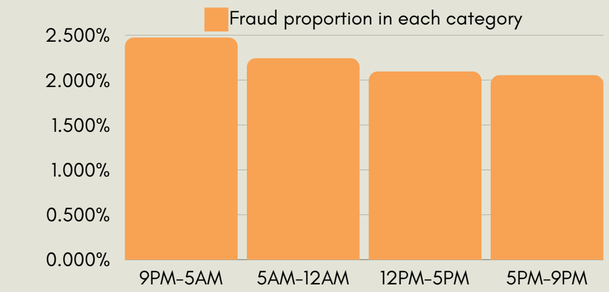
- ***Data Preparation***
- Predictor Variables

1.**Age**: Age patterns differ across different fraud types.

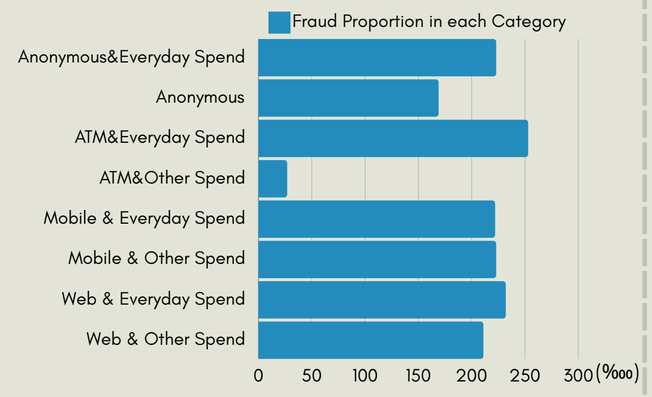


2.**Trading Period(new)**: The Hour extracted from the Transaction Date variable. The peak period is Night to Early Morning.





3.**Payment_equipment_amount(new)**
= Agent type + Transaction Amount type
(amounts ≤ 523 "Everyday Spend", amount> 523 "Other Spend", based on two-steps clusters).



- Data Transformation

Normalising--Converting--Upsampling--Removing

- Data Spliting

- ***Techniques used***
- Cross-validation

K-fold cross-validation was used, folding the data **seven times**, to assess the model's performance across different data splits.

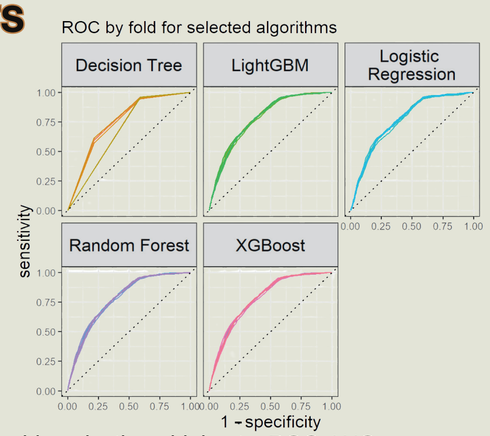
- Model Selection

Testing five models, including:

- Logistic Regression
- Random Forest
- Decision Tree
- LightGBM
- XGBoost

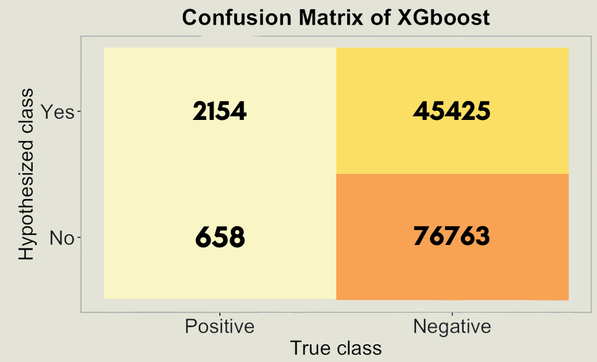
To find the best-performing classification model

RESULTS



XGBoost achieved the highest ROC-AUC score, consistently demonstrating high sensitivity and specificity across cross-validation folds, making it the optimal choice for real-time fraud detection at ASB Bank.

Model	bal_accuracy	roc_auc	sensitivity	specificity
XGBoost	0.697	0.772	0.766	0.628



The confusion matrix shows that XGBoost correctly identified **76.6**% fraudulent and **62.8**% non-fraudulent transactions in the testing data.

CONCLUSIONS AND RECOMMENDATIONS

Age

- **Family Involvement:** The bank could encourage seniors to use joint accounts with trusted family members.
- **Offline Services:** The bank could provide more in-person services tailored to seniors.

Trading Period & Payment_equipment_amount

- **Transaction limits:** Setting higher transaction limits during fraud-peak hours is suggested.
- **Security pattern:** The bank is suggested to use passcodes or biometrics for added security on unknown devices.
- **Install Protective Equipment:** The bank could install anti-peeping shields and encrypted keypads on ATMs, and use devices that can detect and block illegal card skimmers.

Automated Messaging System

38% of 1.2 Million Clients (456K) are messaged
4.53% of them (20.7K) are truly defrauded
One Message costs $0.1
Total cost is $0.0456M
Frauded Amount's Median of
per transaction is $100
Total defrauded amount is $2.07M
We assume that 50% can be recovered
$1.035M
Total recovered amount is to ***$989.4K***

Reference Code