Optimisation Model for Investigating Fraudulent Transactions in the Banking Sector

Methodology, Modelling and Consulting Skills

Project By:

Ting-Hung Tsai s2442379
Takae Hayashi s2447325
Yukta Bhange s2554720
Rasika Dalvi s2511090

THE UNIVERSITY *of* EDINBURGH

sopra steria

Executive Summary

Within the UK banking systems, frauds and scams in transactions are a significant concern, with a major portion of these scams being committed online. According to the Crime Survey 3.7 million cases of fraud occurred in the year 2022. Fraudsters exploit vulnerabilities in online systems, with a possibility that a significant portion of fraud perpetrated against the UK is partially or fully committed from overseas. To address this issue, the UK government and regulatory bodies are implementing measures such as requiring banks to reimburse fraud victims who have been tricked into sending money to scammers. This makes it more important to identify fraud and improve fraud prevention. Existing automated fraud detection systems rely on algorithms to identify potential problems, such as unusual transactions or spending behavior. Statistical analysis is also a cornerstone of banking fraud detection, involving the gathering of data to establish patterns related to risk using algorithms, which excel at catching new types of fraud.

The importance for this optimization comes from the need to focus on fraudulent transactions and curtail fraud while improving the system using past patterns and maintaining the balance between available resources. To address growing concerns about fraudulent activity in the banking industry, this report presents an optimization model designed to investigate fraudulent transactions within the resource bound. The study uses a dataset provided by Sopra Steria of over 315,000 transactions among five participating banks from October 2023 to July 2024. The primary aim is to formulate a systematic approach to identify potential fraud cases by optimizing investigation strategies.

The proposed optimization method maximizes the amount recovered from fraudulent transactions while minimizing the cost of investigation associated to external investigators. The model seeks to determine the optimal allocation of internal and external investigators to cases, considering workload, investigator availability, and priorities. In the given objective function, the strategy aims to strike a balance between maximizing expected recovery and minimizing costs, ensuring efficient utilization of investigative resources in the banking sector. The optimization prioritizes cases with higher expected amounts so after identifying actual frauds from the previous day, probabilities are updated using "odds" of being fraud. Further to control the number of cases selected for investigation, a linear model is built that predicts the budget required for hiring investigators, considering transaction details and fraud probabilities. This budget is then directed towards cases with the highest fraud rates and expected amounts while excluding others. The strategy works on the idea that transactions with high fraud rates likely already carry substantial expected amounts after adjusting for odds.

The results indicate that the models have good accuracy of 87% in the second run with updated odds but this is achieved at the cost of investigating a high number of cases. Decreasing the budget limit might help with the number of cases investigated by external investigators. The model has high recall (88-99%) but low precision (1-3%). This detects almost all fraud while increasing unnecessary investigations. The model opts for an extremely high recall, low precision approach to minimize fraud losses.

Contents

1 Introduction **1**

2 Problem Statement **1**

3 Solution Approach **2**

3.1 Updating System 2

3.2 Assumptions 2

4 Linear Programming Model **3**

4.1 Variable Definition 3

4.2 Objective function 4

4.3 Constraints 4

5 Results **5**

5.1 Optimal Solutions and Validation Analysis 5

5.2 Further Analysis for Additional Questions 7

5.2.1 Q4: Seasonal Trends 7

5.2.2 Q5: International vs Domestic 7

5.2.3 Q7: Vulnerable Bank 8

5.2.4 Q8: Transactional Outliers - Fraud vs Non-Fraud 9

6 Conclusion and Suggestions **9**

1 Introduction

The world is experiencing an explosive increase in financial fraud. According to the UK Finance's report[1], In UK, over £1.2 billion was stolen through fraud in 2022. That is equivalent to £2,300 every minute of last year. Consumers in the US also lost around $8.8 billion (approximately £7 billion) last year, despite their record investment in detection and prevention. In many countries, the amount of damage has increased rapidly compared to the previous year. Fraud is now the most common crime in the UK, affecting one in 15 people. In total, around 3 million frauds took place across the UK, with payment card fraud being the most common. Other common types of fraud were purchase fraud, romance fraud, and investment fraud. This trend of financial fraud is now one of the biggest threats facing the industry. Not only does the financial industry incur costs in fighting fraud, but it also risks losing the trust of victimized customers.

Additionally, as fraud techniques become more sophisticated, the question of who is responsible for the damage becomes increasingly difficult. In the UK, for instance, victims of unknown transactions (for example, if someone copies your credit card and uses it) are legally protected from loss. However, if someone tricks you into paying, who is responsible is less clear. In July, the UK Supreme Court ruled that banks cannot be held liable for simply following the instructions of customers who were tricked into sending money overseas[3]. In other words, banks are legally obligated to refund money for so-called unauthorised fraud, however, they currently do not have to cover the cost of an authorized push payment scam (AAP), where victims are tricked into agreeing to send money to fraudsters. Some payment service providers have signed on to the Voluntary Contingent Reimbursement Model (CRM) Code, which began operational in 2019, and have also made independent voluntary commitments to victim reimbursement. While the UK Government welcomes and encourages this, they also recognise that reimbursement to victims of APP fraud remains inconsistent and many victims continue to suffer losses without reimbursement[2].

It is therefore apparent that financial industry organizations, including banks and technology companies, urgently need a wide range of proactive measures to prevent fraud in order to protect their customers and maintain their trust. As an example of these measures, we aim to provide meaningful insights to the consultancy firm Sopra Steria and its client banks by applying mathematical approaches to explore optimal solutions to minimize the occurrence of fraud across multiple banks in this project. We believe that this could contribute to banks find ways to protect their customers who may be targets of fraud.

2 Problem Statement

The problem at hand revolves around transaction dataset originating from a bank. There are 5 domestic banks involved in the transactions along with other international banks. There are mainly 3 types of transactions taking place in the dataset. The first type of the transaction is within one of the domestic banks itself where the sender's and receiver's bank is the same. The second type of the transaction is between 2 domestic banks out of the aforementioned 5 banks where sender's and receiver's banks are different. Lastly the third type of transaction is where transaction takes place between one of the domestic banks and an international bank. The dataset mainly consists of debit transactions along with some credit transactions such as regular income, transfers between accounts, and interest earned.

In addition, the data includes customer probabilities as well as transaction and description probabilities of fraud from banks' internal system. The probabilities for customer, transaction and description ranges between 0.2 - 0.9, 0.1 - 0.8 and 0.1 - 0.9 respectively. Each bank has it's individual investigator team for the rectification of the fraudulent cases. Bank A has 8 internal investigators, Bank B has 12 internal investigators whereas Banks C, D and E has 10 internal investigators. Priorities can be assigned to transactions based on the time taken for the investigation which ranges between 0.25 - 2 days. Banks can hire external investigators for the investigations at a certain hiring cost.

At the end of each daily iteration of the optimization model, the transactions will be categorized as true positives, false positives, false negatives, and true negatives. In this report we are required to create a stochastic model to adjust the probability scores and priority criteria in order to improve the performance of the system through learning from the past results.

Table 1: Transaction Priority and Hiring Costs

Transaction Priority	Investigation Time (days)	Cost of Hiring Externally (£)
1	0.25	40
2	0.5	60
3	1	100
4	2	150

3 Solution Approach

3.1 Updating System

The objective is to build a model to detect potential fraud cases for the investigation that maximizes the expected fraudulent amount investigated and minimizes investigation costs. Namely, the linear programming model would be sensitive to a case with a higher expected fraudulent amount. As a result, we aim to add parameters that increase the weight of certain situations, such as a case belonging to an electronics category, and it would contribute to the expected fraudulent amount. After the optimization model is completed on the current day, actual frauds are identified. Then the probabilities of different events being fraudulent and their odds could be computed. The formula of the odds is given as follows

$$\frac{P(A)}{1 - P(A)}$$

where $P(A)$ is the probability of an event A is a fraud. The odds are used to weight cases based on relative fraud likelihood in different categories.

To better differentiate between cases, the transaction probabilities, customer probabilities and description probabilities are separated into levels. A probability between 0 (excluded) and 0.1 (included) would be level 1, between 0.1 and 0.2 would be level 2, etc. This categorization allows an odds ratio to be calculated for each level based on historical fraud rates, rather than treating all probabilities equally.

The transactions are categorized into 23 category types such as bank fees, electronics, groceries etc. The odds is calculated for each category based on its historical fraud rate compared to the overall rate. For instance, there were 2 frauds out of 9 electronics purchases in day 1, giving a 22.2% fraud rate, much higher than the overall 0.8% rate. This results in an odds ratio of 0.2857 for electronics. In contrast, the bank fees category had 0 frauds in 2 cases, hence odds ratio of 0. If an electronics and bank fee case both had £1 transactions, the expected "payout" from investigating them would be £1 and £1.2857 due to the category odds adjustment. This differentiation in category odds allows the model to prioritize investigation of categories that have shown higher historical fraud rates, maximizing the likelihood of uncovering issues on new similar cases.

Besides, we compound the odds in the current day calculated after seeing the results with previous odds by multiplying them to be the odds using for the next day. We initialize the all the odds as 0 in the beginning.

$$\text{odds}_{\text{next}} = (\text{odds}_{\text{previous}} + 1)(\text{odds}_{\text{current}} + 1) - 1$$

To prevent selecting too many unnecessary cases, we have developed a linear model that utilizes a fixed true positive rate to estimate the budget by calculating the cost of hiring additional investigators. The linear model predicts the budgets by using the number of cases with different categories, transactions, customers, and probability levels of the description. The budget is set to select cases with a top fraud rate and expected benefits while truncating the others. This approach is based on our expectation that the high fraud rate cases are already of high expected amount after the odd adjustment.

3.2 Assumptions

In order to solve this problem as a linear programming model, below stated assumptions[4] are considered:

1. As the above stated problem is being solved as linear optimization problem which assumes linearity, the objective function and constraints are formulated as linear functions of decision variables.

2. Similarly, the decision variables and the objective function or constraints is assumed to exude proportionality.

3. The decision variables are assumed to exhibit continuity within their feasible range.

4. Rationality in the decision-making process is assumed.

4 Linear Programming Model

This optimization problem is applied to each from day 1 to 305 iteratively and the availability of employees from the previous n^{th} day is updated to $n+1^{th}$ as data input.

4.1 Variable Definition

Let

Index

N be the case number of the day	$n \in N$
$Q = \{1, 2, \cdots, 50\}$ be the internal investigator label	$q \in Q$
$B = \{1, 2, 3, 4, 5\}$ be the bank A, B, C, D and E	$b \in B$
$P = \{1, 2, 3, 4\}$ be the priority 1, 2, 3 and 4	$p \in P$

Decision Variable

x_n be the binary indicator of whether to investigate the case n or not
y_n be the binary indicator of whether the case n is investigated by an internal investigator or not
z_n be the binary indicator of whether the case n is investigated by an external investigator or not
$w_{n,q}$ be the time of the investigator q spent on the case
$u_{n,q}$ be the binary indicator of whether the home bank investigator q is involved in the investigation of the case n or not
$v_{n,q}$ be the binary indicator of whether the paid-out bank investigator q is involved in the investigation of the case n or not
l_q be the availability of the internal investigator q for the next day
$e\_\text{one}_{n,q}$ be the binary indicator that supports the second constraint
$e\_\text{two}_{n,q}$ be the binary indicator that supports the third constraint
e_three_q be the binary indicator that supports the fourth constraint
e_four_q be the binary indicator that supports the sixth constraint
e_five_q be the binary indicator that supports the seventh constraint
e_six_q be the binary indicator that supports the eighth constraint

Data

pc_n be the customer probability of the case n
pt_n be the transaction probability of the case n
pd_n be the description probability of the case n
c_p be the cost of hiring an external investigator to investigate the case associated priority p
k_p be the investigation duration of the case with priority p
m_n be the amount of transaction n
α_n be the category odd of the case n
β_n be the transaction odd of the case n
γ_n be the description odd of the case n
δ_n be the customer odd of the case n

$d_{n,p}$ be the binary indicator of whether the case n is with priority p or not
$f_{n,b}$ be the binary indicator of whether the case n is belonging to the home bank b or not
$g_{n,b}$ be the binary indicator of whether the case n is belonging to the paid-out bank b or not
$h_{q,b}$ be the binary indicator of whether the investigator q is belonging to the bank b or not
r_n be the investigation responsibility ratio of the home bank for the case n
s_q be the binary indicator of whether the internal investigator q is availability for the current day
budget be budget for the current day
M be sufficiently large number that does not limit how many investigators can work on a case

4.2 Objective function

$$\max \sum_{n\in N}\Big(\frac{1}{3}(\alpha_n+1)((\beta_n+1)\,\text{pt}_n m_n + (\gamma_n+1)\,\text{pd}_n m_n + \text{pc}_n m_n)\,(\delta_n+1)\,x_n - z_n\sum_{p\in P} d_{n,p}c_p\Big) - \sum_{q\in Q}\frac{c_1}{k_1}(1-l_q)$$

The linear programming maximizes the expected fraudulent amount and minimizes the cost of hiring extra investigators and the loss of assigning a priority 4 tasks to internal investigators.

4.3 Constraints

1. If the transaction is belonging to the home bank b then an investigator in the home bank b will investigate this case.

$$y_n f_{n,b} = \sum_{q\in Q} u_{n,q}h_{q,b} \qquad \forall b\in B,\ \forall n\in N$$

$$y_n = \sum_{q\in Q} u_{n,q} \qquad \forall n\in N$$

2. If the transaction is paid-out to the bank b then an investigator in the paid-out bank b will investigate this case.

$$1-\sum_{q\in Q} v_{n,q}h_{q,b} \le e\_\text{one}_{n,q} \qquad \forall b\in B,\ \forall n\in N$$

$$\sum_{q\in Q} v_{n,q}h_{q,b} - 1 \le e\_\text{one}_{n,q} \qquad \forall b\in B,\ \forall n\in N$$

$$y_n f_{n,b} - y_n g_{n,b} \ge e\_\text{one}_{n,q} - 1 \qquad \forall b\in B,\ \forall n\in N$$

3. If the paid-out bank is the same as the home bank then there is no extra investigator in the the home bank taking this case. Additionally, if the paid-out place is the international bank or withdraw then there is no extra investigator in the home bank taking this case.

$$\sum_{q\in Q} v_{n,q}h_{q,b} \leq e\_\text{two}_{n,q} \qquad \forall b\in B,\ \forall n\in N$$

$$\sum_{q\in Q} v_{n,q}h_{q,b} \leq e\_\text{two}_{n,q} \qquad \forall b\in B,\ \forall n\in N$$

$$y_n f_{n,b} - y_n g_{n,b} \leq 1 - 2e\_\text{two}_{n,q} \qquad \forall b\in B,\ \forall n\in N$$

4. If the internal investigator is not available for the current day then the investigator will not be assigned to a case.

$$s_q \geq e\_\text{three}_q \qquad \forall q\in Q$$

$$\sum_{n\in N}(u_{n,q}+v_{n,q}) \leq Me\_\text{three}_q \qquad \forall q\in Q$$

5. The investigator in the home bank b will take a certain proportion of the workload of the case.

$$w_{n,q} = \sum_{p\in P} d_{n,p}k_p\left(u_{n,q}r_n + v_{n,q}\left(1-r_n\right)\right) \qquad \forall q\in Q,\ \forall n\in N$$

6. If the investigator spent more than one day on a case then the investigator can only take one case.

$$2 - \sum_{n\in N} w_{n,q} \geq e\_\text{four}_q \qquad \forall q\in Q$$

$$\sum_{n\in N}(u_{n,q}+v_{n,q}) - 1 \leq Me\_\text{four}_q \qquad \forall q\in Q$$

7. If the investigator does not involve in the investigation of a case with priority 4 then the workload should not exceed one day.

$$\sum_{n\in N} w_{n,q} \leq e\_\text{five}_q + 1 \qquad \forall q\in Q$$

$$\sum_{n\in N} d_{n,4}\left(u_{n,q}+v_{n,p}\right) \geq e\_\text{five}_q \qquad \forall q\in Q$$

8. If the investigator spent more than one day on a case then the investigator is not available for the next day.

$$\sum_{n\in N} w_{n,q} \leq 2 - e\_\text{six}_q \qquad \forall q\in Q$$

$$l_q \leq e\_\text{six}_q \qquad \forall q\in Q$$

9. The case will be assigned either to the internal investigator or the external investigator.

$$x_n = y_n + z_n \qquad \forall n\in N$$

10. The budget constraint for external hiring.

$$\sum_{n\in N}\sum_{p\in P} d_{n,p}c_p z_n \leq \text{budget}$$

Note that M here is set to be 100 because $\max\sum_{n\in N} u_{n,q}+v_{n,q}$ is 100. A single can be assigned to all 50 internal investigators in both home banks and paid-out banks.

5 Results

5.1 Optimal Solutions and Validation Analysis

In this project, the model was built with the goal of maximizing both the accuracy of determining which transactions should be investigated for fraud and the amount of money saved from fraud,

Table 2: Validation of Optimization Model with and without Budget Constraint

	First Run		Updated Run	
	Model w budget	Model w/o budget	Model w budget	Model w/o budget
Accuracy	87.8%	97.7%	87.6%	99.2%
Precision	0.034	0.013	0.033	0.010
Recall	0.886	0.974	0.883	0.992

while minimizing the total expenditure across banks on fraud. According to our solution approach, we first run the model without the budget and keep updating the odds after that using the budget predicted by the regression we have a second updated run with and without the budget constraint. The results of the model for the first run and the updated run in terms of different metrics are given in Table 1.

The results of the optimal solution obtained in the updated run with budget constraint can be seen in Table 3. This includes the amount of money correctly determined and protected from fraud (True Positive: TP), the amount of loss due to fraud that was overlooked (False Negative: FN), and the amount of additional cost for hiring external investigators for cases incorrectly determined to require investigation (false positives: FP).

Table 3: Optimal Solutions

Condition	Transaction Count	Amount Description	Amount(£)
TP	2190	Saved from fraud	236993.20
FN	291	Loss - Fraud taken	16147.20
FP	63967	Loss - Cost waste	1110810.00
		Total Loss	1126957.20

This result explains that the cost of false positive cases (cost of external investigation to identify fraud) is over 68 times greater than the cost of false negative cases (money lost due to fraud). In short, our model is a very conservative model that emphasizes minimizing the amount of damage to customers rather than the additional costs of investigation.

This validation result indicates that the precision rate (also known as sensitivity) is low and the recall rate is noticeably high. In most cases, there is an inverse relationship between precision and recall, where it is possible to increase one at the cost of reducing the other. In general, when the cost of false positives is high, precision would be a better choice, and when the cost of false negatives is high, a recall rate would be emphasized. The costs mentioned here mean, for example, the former is the impact of economic loss due to the time and cost of additional investigation into cases that do not require investigation, and the latter is the impact of loss of trust from customers who have actually been harmed by fraud.

That is to say, since our current model has a high recall rate, which increases the chance of detecting all fraud cases (positive outcome), but also increases the chance of investigating unnecessary cases that are not fraud (negative outcome). In this way, it is a complete trade-off as to which influence to emphasize and which consequences to obtain.

This may be an extreme statement, however, if *Sopra Steria* and its client banks have a business decision that they place greater importance on reducing the cost of additional investigations than on the amount of damage suffered by bank customers and the resulting loss of trust, the model could also be considered adjusting to make the precision higher. On the other hand, also in the inevitable case such as not being able to hire a sufficient number of additional investigators, precision will increase and recall will decrease as a result. Therefore, the result will have the same trend as in the hypothetical case mentioned above, which leads to a reduction in the bank's total expenditures.

5.2 Further Analysis for Additional Questions

We would like to answer for several additional questions here by reviewing plots of the given raw data and comparing between different perspectives, such as months, banks, international/domestic, categories, fraud types, for the fraud transactions and non-fraud ones.

5.2.1 Q4: Seasonal Trends

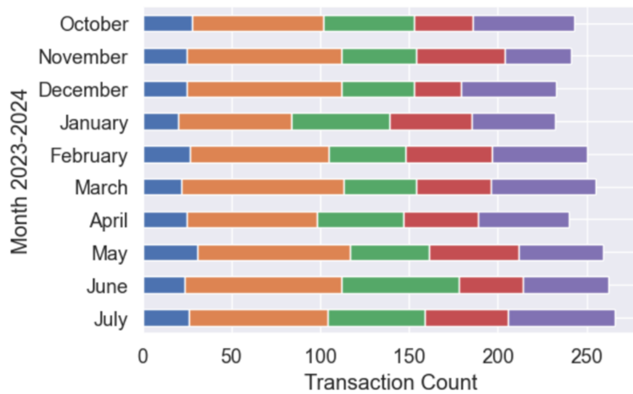

Figure 1: Q4 - Seasonal Trends for Fraud Transactions

By reviewing the number of fraud transaction, the data reveals noticeable seasonal patterns over the span of 10 months. Fraud transaction numbers exhibit a decline during the winter months, hitting lows during the December to February timeframe. This downward trend reverses in spring, as fraud transactions incrementally increase from March through May. The summer months see further escalation, with transaction figures peaking in July before declining slightly in the fall months. The seasonal effects observed in damage amount follow a broadly similar pattern, with winter months also typically associated with lower levels of damages from fraudulent activity. Although the damage amounts undergo an unusual and remarkable rise in March touching but gradually drop towards spring. The damage amounts also show a slight increase as in July. The analysis points to winter as the period of reduced fraud transaction activity, while late spring and summer are characterised by increased fraud count.

5.2.2 Q5: International vs Domestic

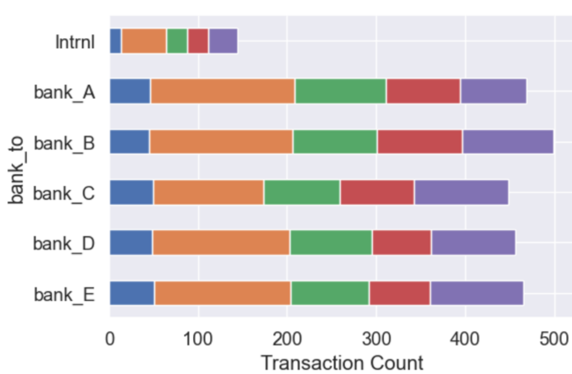

Figure 2: Q5 - International vs Domestic for Fraud Transactions

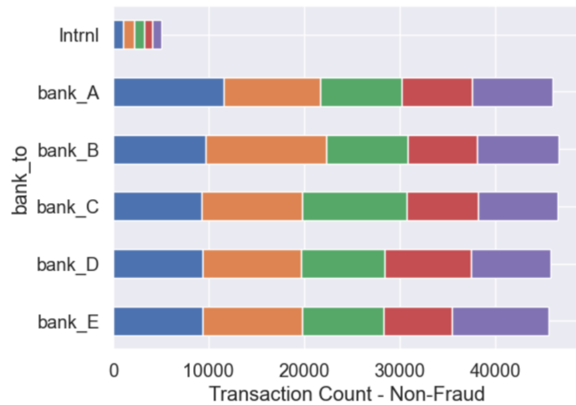

Figure 3: Q5 - International vs Domestic for Non-Fraud Transactions

The total number of international transactions is far fewer than that of domestic ones and this trend can be seen in both fraud and non-fraud cases in common. When it comes to the amount lost by fraud, the international transactions surpass £1200. This is a remarkably higher amount compared to the total damage amount for transactions between two domestic banks which reach to only around £400 for each home bank. On the other hand, the total amount of money for non-fraud transactions are similar each other of transactions international as well as domestic ones in range between £600 and £700.

5.2.3 Q7: Vulnerable Bank

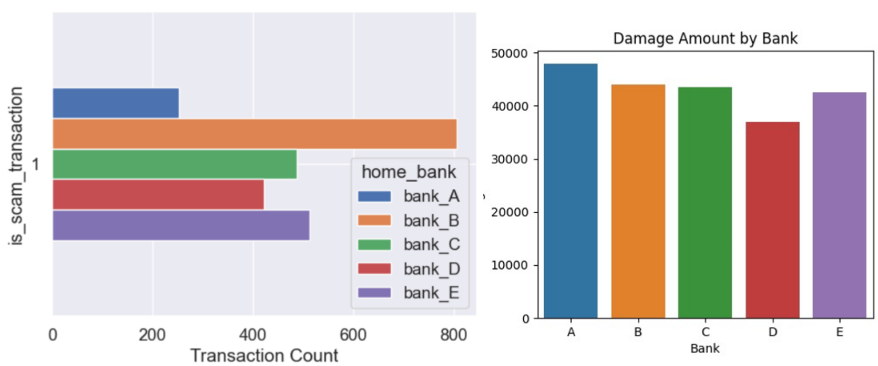

Figure 4: Q7 - Bank Fraud Transaction Count & Damage Amount

The data showed that Bank B experienced an unusually high number of fraudulent transactions (approximately 800) compared to the other banks. Banks A, C, D and E saw fraudulent transaction figures that were significantly lower and within a closer range (400-500). This indicates that Bank B stood out as an outlier with high fraud activity. However, a comparative look at overall fraud amount told a more different story. Here, the total damages linked to fraudulent schemes were remarkably similar across all five banks approximately near, despite Bank B's outsized incident volume. This suggests that while Bank B saw more fraudulent transactions, the value per transaction was lower than bank A which had less number of fraud transactions. The analysis indicates that Bank B's exposure profile is characterized by high frequency, low value fraud. As for banks A, C, D and E, their risks lean towards lower frequency but higher value transactions. In essence the data shows that bank B is more vulnerable to scams in terms of number of transactions but all the banks are at same stage of vulnerability when it comes to total amount.

5.2.4 Q8: Transactional Outliers - Fraud vs Non-Fraud

Losses from fraud in all categories except Investment range from approximately £200 to £600 in each category, however, we can see that losses in the Investment category exceed £1100, which is significantly higher than in other categories. On the other hand, for non-fraud transactions, the total amount for Housing was approximately £6,800, which is remarkably greater than the amount for transactions in all other categories (less than £1,500 each). This suggests that Housing transactions are not a major target for fraud, at least for now. Therefore, we recommend that banks consider paying closer attention to Investment-related transactions and increasing the priority of fraud investigations. In terms of the number of transactions, we can see that there are many more fraudulent transactions in Utilities, Shopping, and Holiday than other categories. Utilities and Shopping each have around 20,000 non-fraud transactions, so it might be more difficult to correctly identify whether or not they are fraudulent than other categories. Therefore, we recommend that banks also focus on monitoring transactions in these categories.

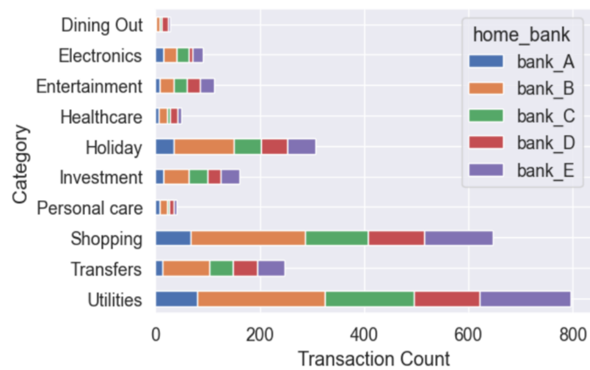

Figure 5: Q8 - Comparison in Categories for Fraud Transactions

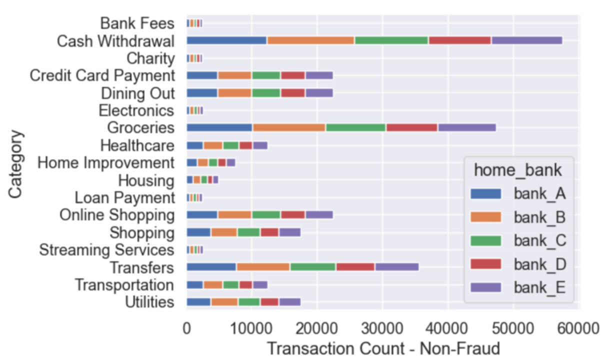

Figure 6: Q8 - Comparison in Categories for Non-Fraud Transactions

6 Conclusion and Suggestions

The results indicate that the model has a good accuracy of 87% in the second run with updated odds but this is achieved at the cost of investigating a high number of cases. High recall and low precision detect almost all fraud while increasing unnecessary investigations to minimize fraud losses. Higher precision would reduce investigation costs but allow more undetected fraud. The model could be adjusted to increase precision if banks prefer to minimize investigation costs over minimizing fraud losses for customers. But this may hurt customer trust. Expected values are only increased by the odds there should be a system to handle these, the sensitivity towards the bank being more prone to fraud is not considered. The model should be modified such that additional biases like international transactions should always be investigated as they are generally more vulnerable are accounted for.

References

[1] UK Finance. Annual fraud report 2023. https://www.ukfinance.org.uk/policy-and-guidance/reports-and-publications/annual-fraud-report-2023, 2023.

[2] Government approach to authorized push payment scam reimbursement. https://www.gov.uk/government/publications/government-approach-to-authorised-push-payment-scam-reimbursement/government-approach-to-authorised-push-payment-scam-reimbursement, 2022.

[3] Ainsley Thomson Nabila Ahmed, Adam Haigh and Ellie Harmsworth. Bloomberg article: Deepfake imposter scams are driving a new wave of fraud. https://www.bloomberg.com/news/articles/2023-08-21/money-scams-deepfakes-ai-will-drive-10-trillion-in-financial-fraud-and-crime, 2023.

[4] H. Paul Williams. *Model Building in Mathematical Programming*. John Wiley Sons, 2013.