Report: Credit Card Fraud Detection Analysis

1. Introduction

In the sphere of a fast-evolving technological landscape, fraud has now become one of the biggest challenges, especially regarding financial transactions. As the number of credit card payments grows, so does the opportunity for fraudsters to take advantage of any vulnerability, resulting in a sharp increase in credit card fraud. This not only creates reputational risks for businesses but also has serious financial implications for both consumers and institutions.

For instance, in Hong Kong, fraud-related banking complaints to the Hong Kong Monetary Authority (HKMA) increased dramatically in 2023, with over 1,200 complaints - more than doubled from 2022 (HKMA, 2024). According to the Hong Kong Monetary Authority, the number of reported deception cases to law enforcement increased by 52%. As a consequence, an estimated HK$7.2 billion were lost. It underscores the pressing need for reliable fraud detection and prevention mechanisms to enhance security for the overall ecosystem of the financial sector.

Along with it, big data leakages have created data abundance for fraudsters. In 2013, the Target data breach affected 40 million customers, which resulted in massive financial losses and customer trust issues (CNBC, 2013). Another case appeared in 2018 when British Airways experienced a data breach, exposing the credit card information of over 380,000 customers (BBC, 2018). Afterwards, credit card information are open to be used in transaction fraud.

Alongside the fight against fraud in business transactions, false flagging of valid transactions has also emerged as a growing concern, disrupting customer experience and causing businesses to lose money. Recent reports show that false positives in online transactions can run as high as 20%, resulting in lost revenue, wasted acquisition costs, and damage to brand reputation (Thetaray,2023). This calls for an effective fraud detection system. Not only will this minimize the occurrence of such false alarms for the business, but it will also ensure that genuine customers are not affected, thereby protecting financial outcomes and customer trust.

Because of those high risks and declining user satisfaction, transaction providers need improved fraud detection systems to target multiple (additional) business goals. From those, key requirements can be identified.

1. **Analyse reliably**: The model must have high recall to reliably identify most fraudulent transactions, minimizing false negatives and detecting fraud accurately.
2. **Low false positives**: It is necessary to avoid flagging too many valid transactions as fraudulent to ensure that false alerts do not impact customers. This ensures user satisfaction.
3. **Easily adaptable**: A solution is needed to adapt to new fraud patterns. Ideally, no or few manual adjustments are needed for fast and cost-effective adjustments.
4. **Real-time processing**: The model must operate in real-time to identify fraudulent activity when transactions occur.
5. **Computational efficiency**: The system should be computationally efficient, ensuring it doesn’t demand excessive resources during transaction analysis. Otherwise, it will strain business resources, and customers will be unsatisfied if they need to wait long for their transactions to be processed.
6. **Regulatory compliance**: Compliance with data protection laws and financial industry requirements is mandatory for the system.

To address the aforementioned problems, the report intends to develop a reliable fraud detection machine-learning model that meets these requirements. Machine learning models offer a promising solution to this problem, enabling real-time fraud detection and rapid response to suspicious activities. It provides solutions by analyzing large volumes of transaction data and identifying patterns. Moreover, these models can flag fraudulent transactions that traditional methods might overlook. However, a key challenge is the issue of class imbalance, as fraudulent transactions are much less frequent than legitimate ones, making the model training process more complex, which needs to be overcome.

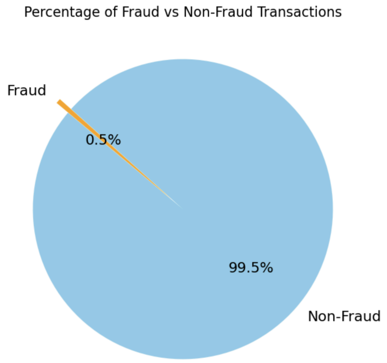
2. Data Exploration and Preparation

2.1 Dataset Overview

The dataset used for fraud detection analysis contains a total of 1,852,394 entries, which is a combined version for training and testing (dataset). For this exploratory data analysis (EDA), all data sets were analyzed to provide a broader understanding of the structure and patterns of the data. The primary objective would be to develop a predictive model for fraudulent transactions, which involves transcending user demographics, transaction characteristics, and geographic and merchant features into a meaningful prediction and analysis.

Key features in the dataset include 23 following columns, which give the information below:

User Demographics: These are the age, sex, and dob (Date of Birth) columns that provide basic user information.
Transaction Details: Transaction amount (amt), unix_time, trans_date_trans_time (timestamp of the transaction), and the geographical coordinates of user and merchant (merch_lat, merch_long, user_lat, user_long).
Other Features: job, city, state, and is_fraud.
The target variable is is_fraud, which is critical in predicting whether a transaction is fraudulent (1) or legitimate (0). The Number of non-fraud transactions was 1,842,743, whereas the number of fraud transactions was 9,651. This resulted in a lot of imbalance in our data set, with just 0.52% of entries being fraudulent. Data imbalance creates further challenges in training models and will be handled to avoid any kind of bias toward the majority class (nonfraudulent transactions).



2.2 Data Cleaning

Data cleaning is one of the most crucial steps in preparing a dataset for robust analysis and modelling. Clean, well-organized data will be the base of the strength and reliability of the results derived from the analysis. The following steps were taken to ensure the validity of the data, along with creating systematic and coherent data to avoid irrelevancy and inconsistency:
Dropping of irrelevant Columns: Unnamed: 0, trans_num (transaction number), and cc_num (credit card number) were dropped since they lacked significant meaning and wouldn't contribute to the analysis or even interfere and create noise, leading to overfitting in model analyses later.
Date-Time Conversion: The trans_date_trans_time and dob columns for the transaction date and the user's date of birth were converted into DateTime format. This is essential for time-based analysis, such as analyzing seasonal trends of fraud, the age of a user, and patterns of transaction activities over time.
Missing Data Treatment: A thorough examination was conducted to ensure no missing or null data is present in the dataset. Our dataset was fully complete with no missing values. Thus, no imputing or removing of instances was performed.
Feature Scaling: StandardScaler was used to normalize the values of different features and ensure that all the features are on the same scale. Feature scaling will be significant for the machine learning algorithms that will be applied for modelling later.
Target Encoding: For the model development, target Encoding was applied to high-cardinality categorical features (e.g., merchant, category, job_sector) to capture category-fraud relationships and avoid dimensionality issues.

2.3 Feature Engineering

Feature engineering is critically essential in deriving insights from the data and making the model perform better. In this study, the features were engineered and selected to ensure an overall dataset view, focusing on geographic, demographic, and transaction-specific content. This multi-faceted approach creates a scope for a more complex view of fraudulent behaviour by capturing patterns from different aspects of the data. Statistical tests were performed to validate the contribution of these engineered features to the prediction of fraud to ascertain their relevance and importance in predicting fraud. Every feature was visually assessed through exploratory data analysis (EDA) to detect patterns, distributions, and relationships for possible fraud risk. Statistical validation and visual analysis helped find only those consequential features, leading to robust and accurate models.

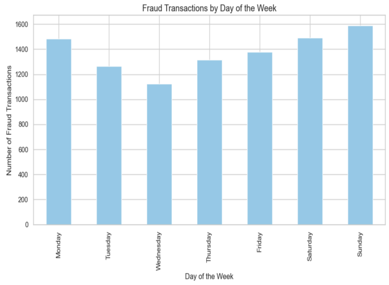
- **Age of Cardholder**: The *age* column is created to specify the cardholder's age and is expressed by the year of birth subtracted from the current year. This allows the model to accommodate age-related patterns in transaction behaviour. Certain age groups might show different spending patterns and bias towards different fraud risks. By including the age feature in the model, we make it possible for the system to see such potential differences and, therefore, fine-tune fraud detection.
- **Geographical Distance Between User and Merchant:** The *distance_user_merchant* feature was developed by calculating the geographical distance between the user and the merchant based on their longitude and latitude columns in the dataset. The distance was calculated using the geodesic method, which gives an accurate measure of the distance in a straight line between any two points on the surface of the Earth.
- **Categorizing Job Titles into Sectors**: Instead of using thousands of jobs, the *job_sector* feature maps the job title of the cardholder to certain predefined job sectors: IT, Healthcare, Finance, etc. This categorization gives the model a chance to analyze fraud patterns across different occupational groupings. Grouping job titles into sectors helps the model to abstract and generalize while pinpointing fraud risks associated with certain occupational trends.
- **Log Transformation of Transaction Amount**: The feature transaction amount (amt) showed a skewed histogram as a mixture of a large number of small transactions and a few extreme outliers. Due to its highly skewed nature, the extreme value of the transaction amount underwent a log transformation to meet

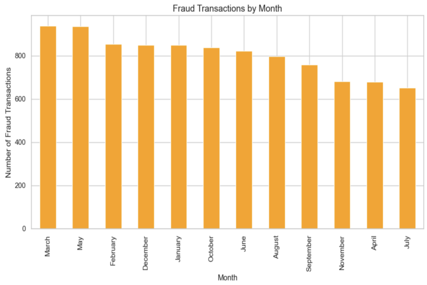
normality assumptions. It makes the model less sensitive to extreme values while pulling the data closer to normal distribution and sharpening a focus on general patterns.

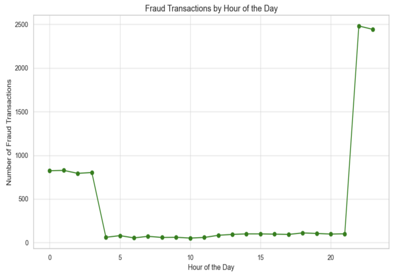
- **Time Analysis of Transaction**: The transaction timestamp (trans_date_trans_time) was broken down into hours, month, and week features based on characteristics relevant to examining fraud trends based on the time of day, day of the week, or seasonality.
- **City-Level Fraud Risk**: The City Fraud Risk feature effectively defines the risk of fraud at the city level, taking into consideration both the fraud rate and volume of transactions. It was created to give an analytical approach to geography. The fraud rate is normalized to allow cities to be compared accordingly. This normalized fraud rate was then weighted by total transaction volume in each city to represent a city fraud risk score, thus allowing high fraud rates along with high transaction volumes to dominate fraud detection prioritization.
- **State-Level Fraud Risk**: Utilizing a similar approach, the features for state fraud risk were engineered. The state fraud risk feature was engineered to measure the level of fraud risk at the state level by combining fraud rate and transaction volume, similar to the city-level analysis. All the normalized rates across states were aggregated into a state-wide fraud risk scoring scheme that provides a generic view of both vulnerability to fraud and the volume scale of transactions at the state level.

2.4 Exploratory data analysis

Time analysis



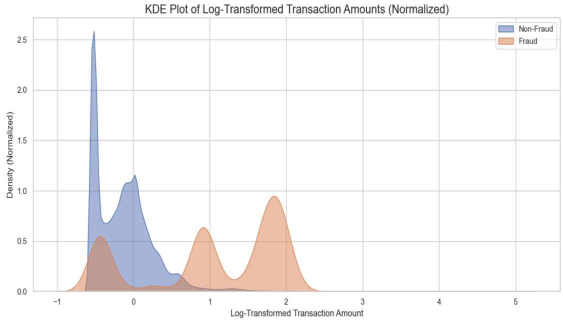




The analysis of time and fraud reflects a strong pattern identified with fraudulent transactions. The fraudulent activities were higher on Sunday and Saturday, peaking on Sunday, which implies that weekends were very susceptible. Along with this, the monthly analysis showed increased fraud in March, May, and February, with less in July and April. The peaks of this crime are later in the night, especially between 10 p.m. and midnight. Chi-Square tests confirmed that fraud was strongly associated with the day of the week, hour of the day, and month, providing values of p = 3.87e-10, 0, and 2.55e-09, respectively.

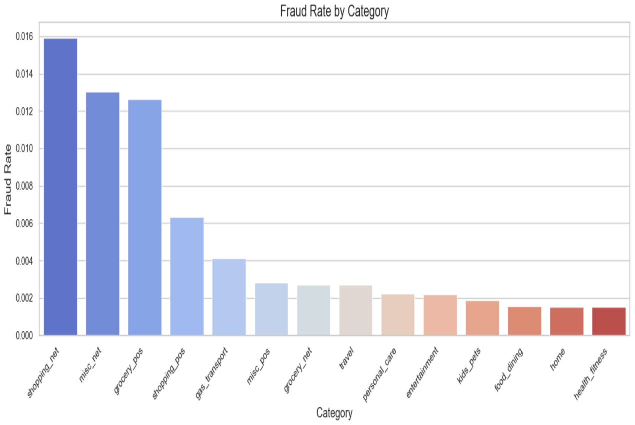
Amount analysis

Transaction amount analysis reflects that fraudulent amounts are generally moderately sized and are distributed more narrowly than non-fraudulent transactions. The range for non-fraudulent transactions heads further out, with the mean ranking closer to zero. Non-fraudulent transactions widen the spread, stretching from low amounts to just a few higher outliers, while fraudulent ones peak in the mid-to-high range. Following log transformation, fraudulent transactions became more distinct, concentrating around a defined density peak between log values of 1.0 and 1.8, while non-fraudulent transactions huddled around lower amounts. There is a pattern in fraudsters often targeting mid to high-sized transactions. The t-test also shows a significant difference between both groups. This implies that fraudulent transactions happen in larger amounts compared to non-fraudulent ones.



Category analysis

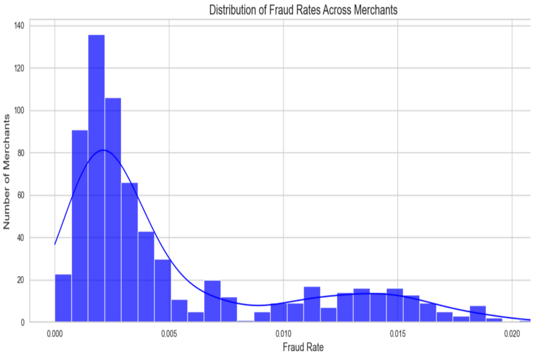
The analysis of fraud across transaction categories highlights significant patterns in both fraud rates and the total number of fraudulent transactions. Categories such as Shopping_net, Misc_net, and Grocery_pos exhibit notably high fraud rates. These categories also report the highest numbers of fraudulent transactions, indicating that both the volume of transactions and the fraud rate play key roles in fraud detection. Conversely, categories like Health_fitness



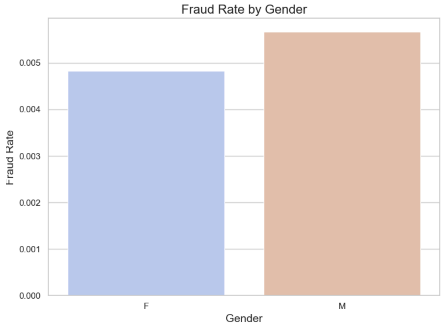
and Home show lower fraud rates, suggesting they are lower-risk areas. It might be because shopping and groceries are very general and less susceptible. The Chi-Square test establishes a significant relevance between the transaction category and the likelihood of fraud (p-value=0.0; Chi2=8329.14).

Merchant analysis

The graph shows that while most merchants operate with low fraud rates, a minority of companies suffer disproportionately high fraud rates, contributing disproportionately to the overall amount of fraud. This indicates that some merchants are consistently subject to higher risks of fraud. Close monitoring of high-risk merchants with tailored systems for fraud detection may ultimately help mitigate potential losses while enabling better allocation of resources to such merchants. The Chi-Square test results confirm a strong relationship between merchants and fraud, with a very low p-value (0.0) and a high Chi2 statistic (8761.09), indicating that fraud is not evenly distributed across merchants.



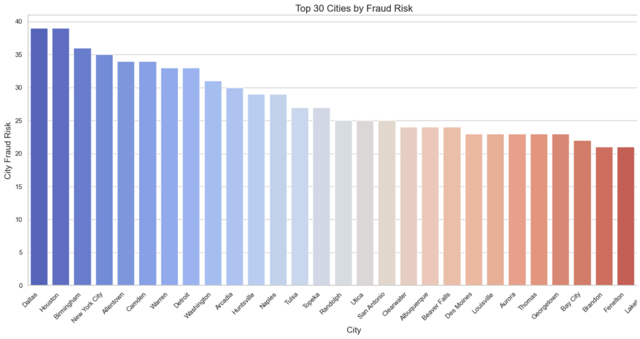
Gender analysis



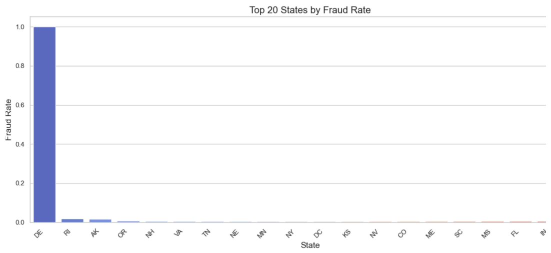
Based on the analysis of fraud rates by both genders, we can see a slight difference in figures of fraud rates, with Males having a fraud rate of 0.5673% and females having a lower rate of 0.4828%. The Chi-square test shows that a strong association existed between the occurrence of fraud and gender. Chi2 was 63.09 p = 1.97e-15. However, low figures are shown overall across both sexes. The results indicate that there could be gender involvement in the modelling of fraud detection systems.

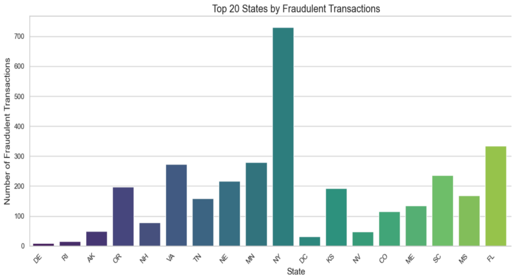
City analysis

City fraud risk analysis reflects that cities with higher fraud rates and larger transaction volumes have a greater fraud risk, such as Dallas and Houston. Smaller cities with extreme fraud rates can be classified as outliers and may have skewed results due to low transaction counts. The t-test shows a statistical significance between fraud and non-fraud transactions based on the city fraud risk, implying that cities with higher fraud risk are more likely to experience fraudulent activity.



State analysis

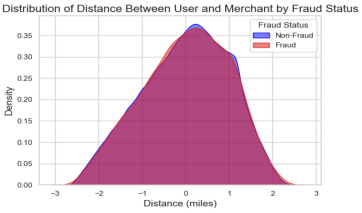




The State fraud risk analysis reflects that fraud rates vary across states, with some, like Delaware, showing extreme values due to small sample sizes. Analysis of state_fraud_risk, i.e. the normalized fraud rate weighted by transaction volumes, identified high-risk states like California and Florida with high transaction and fraud rates. While the correlation between fraud risk and actual fraud is weak (0.005), the T-test(T-Stat: 6.61, P-Value: 3.85e-11) quantifies statistical significance, which indicates that state-level fraud risk is a relevant factor in understanding broader fraud pattern.

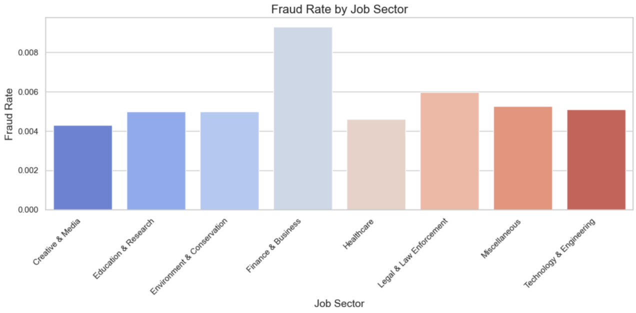
Distance analysis

The distance between the user and merchant by fraud status doesn't significantly differentiate between fraudulent and non-fraudulent transactions. The T-test (P-value = 0.62) and visualizations (KDE, boxplot) confirm that distance does not strongly correlate with fraud likelihood.



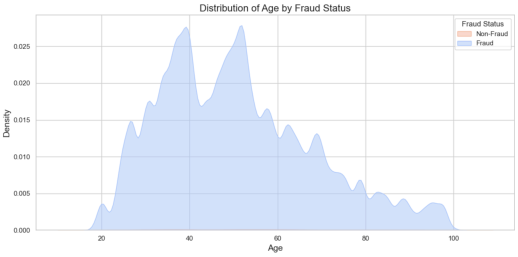
Job analysis

The analysis reveals variations in the rate of fraud among various types of jobs and sectors. The Finance & Business and Legal & Law Enforcement sectors have a higher level of fraud, while others like Healthcare or Technology & Engineering do not. The Chi-Square test provides significant proof of a statistical relation between the work sector and fraud occurrence with a p-value of less than 5% (0.05). This means that the fraud rate is not equally distributed among the job sectors.



Age analysis

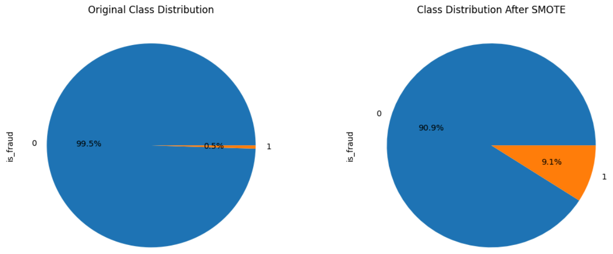
The analysis shows that there is a specific age range having increasing fraud rates from the age range of 30 -50, which peaks at the age range of 50 and exponentially drops after reaching the peak.It reflects a significant pattern between fraud and non-fraud transactions. The t-test confirms this difference with a T-statistic of 14.15 and an extremely low p-value (5.25e-45).



3. Model Development

The dataset was further split into 80% training and 20% test sets, maintaining both sets' fraud distribution (0.52% fraud cases). With the split data, validation was continuously possible during the training to obtain a suitable model for fraud analyses, which will also perform well with unseen data to fulfil its later usage.

3.1 Resampling



Our fraud detection dataset was highly unbalanced; therefore, we realized that a strong resampling technique was required to handle the class imbalance properly. We discovered that aiming for a 10% ratio between the minority and majority classes yielded the best results after conducting a thorough trial with various sampling ratios utilizing SMOTE (Synthetic Minority Over-sampling Technique). This choice was made after conducting several tests and finding that although smaller ratios (<10%) did not present enough fraud patterns for the model to learn efficiently, higher ratios (>10%) resulted in a drop in precision.

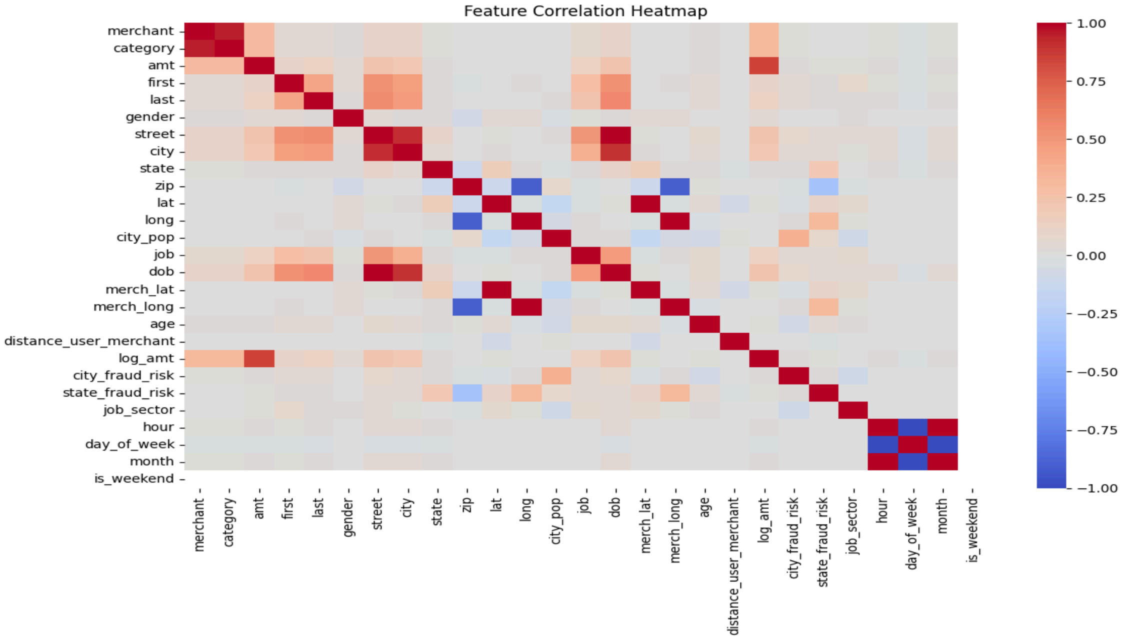
We started using Logistic Regression as our baseline model in order to test our resampling strategy. The selection of logistic regression was based on its interpretable probability scores and computational efficiency, which makes it appropriate for huge datasets. It is also less likely to overfit than more intricate models, and its sensitivity to class imbalance made it a useful gauge of the effectiveness of our sampling plan. With an F1 score of 0.0901, recall of 4.72%, and precision maintained at acceptable levels, our results from the 10% target ratio demonstrated the best balance of measures. Only training data was subjected to resampling, which prevented the test data from being distorted and enabled the model to learn from a more balanced representation of both classes. This allowed us to quantify the model's performance effectively.

3.2 Correlation Analysis

Between data sampling and before training the model, we ensure the quality of the features and the model's performance by checking for correlations between features. This prevents highly correlated variables from affecting the model's stability and accuracy and shrinks the feature space while preserving the features' predictive potential. After analyzing the correlation heatmap, several key insights were identified, leading to refined feature engineering.

The correlation analysis revealed several highly correlated feature groups, including location-based features (merch_lat/lat at 0.994, merch_long/long/zip at 0.909-0.999, city/street at 0.905), temporal features (hour, day_of_week, month at 1.000), transaction amounts (amt/log_amt at 0.855), and user information (category/merchant at 0.946, dob/street/city at 0.901-0.996). To reduce multicollinearity, nine redundant features were removed: location-related features (city, long, merch_lat, merch_long), simplified temporal features (day_of_week, month), and correlated categorical features (category, dob). Additionally, the feature-engineered log_amt was kept over amt due to its better distribution for financial data analysis.

The remaining 18 features exhibited mostly low correlations (<0.2), with only a few moderate correlations (0.2-0.6) among related features such as street/first/last/job (0.425-0.562), zip/state_fraud_risk (0.363), city_pop/city_fraud_risk (0.368), and merchant/log_amt (0.318). This refined set maintains strong predictive potential while minimizing multicollinearity, offering a solid foundation for model training.



3.3 Model Training

With the adjusted training dataset, three models were trained with the business requirements stated in the introduction in mind, focusing especially on high recall. The chosen models were

- **Random Forest** was chosen for its ability to capture non-linear relationships and feature interactions effectively. The model used 50 trees with a max_depth of 10 to avoid overfitting. Parameters such as min_samples_split=20 and min_samples_leaf=10 were set to ensure robust node splitting and improve the model's generalization.
- **KNN** was utilized for its strength in recognizing local patterns in the data. It was optimized with n_neighbors=3 to enhance the precision of fraud detection, and 'distance' weights were used to give more importance to closer neighbors. Additionally, PCA (Principal Component Analysis) dimensionality reduction was applied to improve computational efficiency and reduce the model's overhead.
- **LightGBM**, a gradient-boosting model, was chosen for its speed and accuracy. The model used 50 estimators with a max_depth of 5, feature and bagging fractions set to 0.8, and a learning rate of 0.1 to allow gradual improvements while preventing overfitting.

Overfitting Prevention

Cross-validation with stratified sampling was used to ensure a balanced representation of fraud cases in both training and test sets to prevent overfitting. Early stopping in LightGBM further safeguarded against overfitting by halting training once performance plateaus. Feature selection and dimensionality reduction helped streamline the model, eliminating redundant or irrelevant features. Additionally, regularization through ensemble methods and balanced class weights addressed the common class imbalance issue found in fraud detection datasets.

Performance Optimization

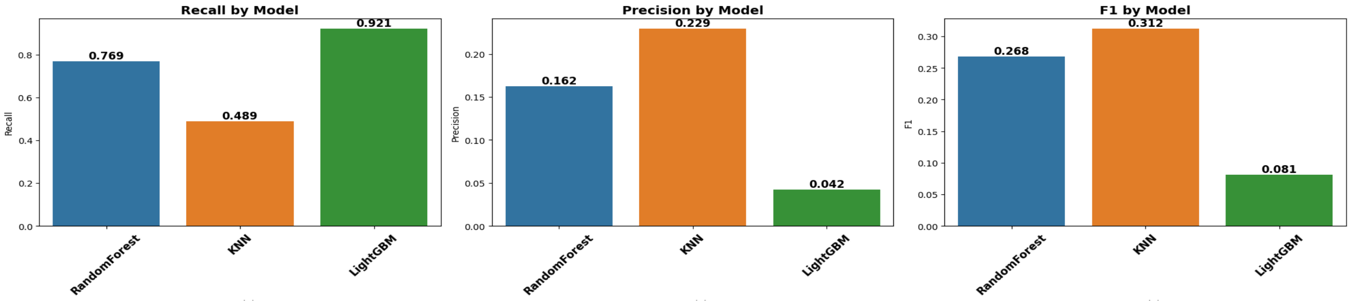
To optimize the computational efficiency of model training, batch processing and parallel processing using all available CPU cores were implemented. Model caching was also employed to save training time, which was particularly useful during hyperparameter tuning. Regarding hyperparameters, grid search was utilized to fine-tune the model settings and identify the optimal parameters for each model.

Metric Selection

To evaluate the performance of the models apart from recall, we also looked at precision as it is an indicator for invalid flagged fraudulent transactions leading to decreased user satisfaction. F1-score was also, as it provides a balanced metric between the mostly opposing goals of high recall and precision.

Results and Metrics

Upon evaluating the model's performance, LightGBM demonstrates the highest Recall at 92.1%, significantly outperforming the other models. As such, it can be considered the optimal model. However, given its relatively lower performance in Precision and F1 score, we will proceed with hyperparameter optimization for LightGBM to improve its performance in these metrics.



3.4 Hyperparameter Optimization

To improve model performance and prevent overfitting, we implemented a comprehensive hyperparameter optimization strategy, with the primary goal of balancing recall (to minimize false negatives) and precision (to control false positives). This balance is essential for ensuring the model is generalizable and performs well on unseen data. In the process, we used 3-fold cross-validation to evaluate the model's stability and ensure that it didn't overfit the training data. An early stopping mechanism was also implemented, which halts training if the model shows no improvement for 10 consecutive rounds. Additionally, bagging and feature fractions were used to promote generalization and reduce model variance. F1 was the primary metric for evaluating the model's ability to detect fraudulent transactions, balancing between the recall rate and precision.

Key Hyperparameters

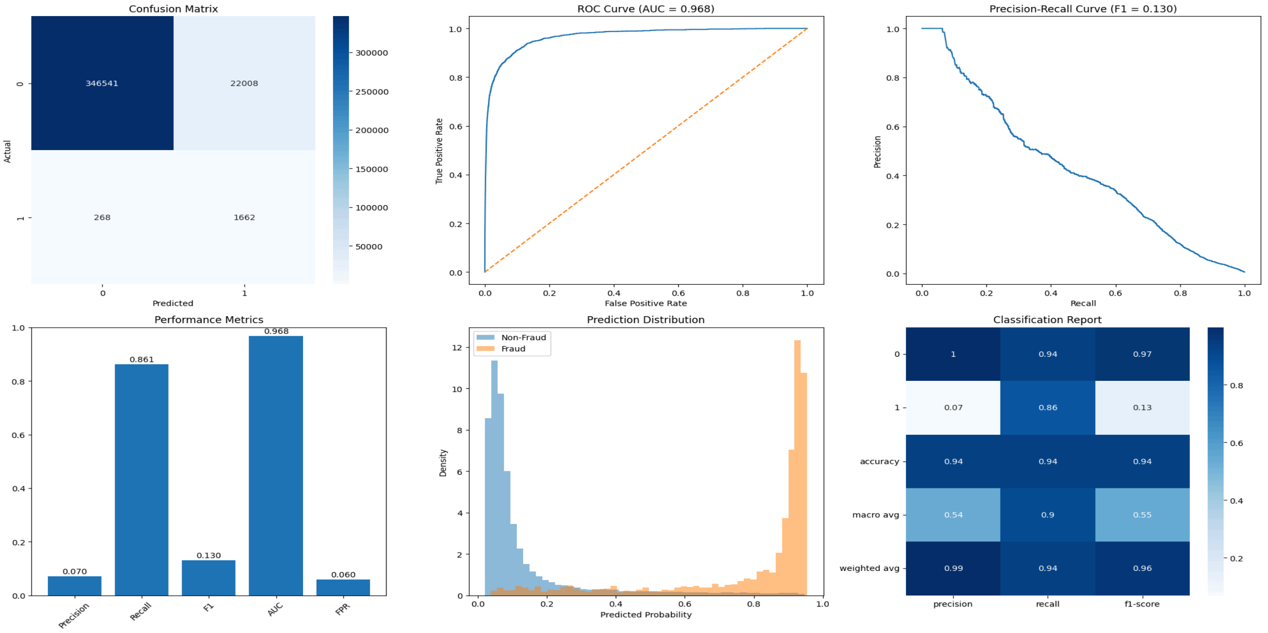
The optimization strategy included several key techniques to improve model performance and prevent overfitting. We employed 3-fold cross-validation to ensure stability and utilized early stopping, halting training if no improvement was seen over 10 rounds. To enhance generalization, we set both bagging_fraction and feature_fraction to 0.7, promoting diversity in the data and features used during training. Regularization was strengthened by increasing reg_alpha and reg_lambda to 0.5, and applying more robust L1 and L2 regularization to reduce model complexity. Additionally, we adjusted parameters such as min_child_samples (set to 30) and min_child_weight to control model complexity and avoid overly deep trees.

Overfitting Prevention Techniques

To prevent overfitting, several strategies were employed. Feature subsampling, by randomly selecting subsets of features during training, helped avoid reliance on any single feature. L1/L2 regularization was applied to control large weights and promote sparsity, while cross-validation ensured the model's performance remained consistent across different data splits. To further reduce model complexity, max_depth was set to 8 and num_leaves to 20, limiting the depth of the trees. A lower learning rate of 0.01 was used to allow for more gradual learning, preventing overfitting. Additionally, class weighting was adjusted with class_weight and scale_pos_weight to prioritize precision, balancing the model's sensitivity to different classes and further reducing the risk of overfitting.

Model Performance

The model's performance metrics demonstrate a notable shift in the balance between precision and recall. The precision improved to 0.070, enhancing the model's ability to correctly identify fraudulent transactions, though this came with a trade-off in the recall, which decreased to 0.861. Despite this trade-off, the model achieved an improved F1 score of 0.130, indicating a better overall balance. The high AUC-ROC score of 0.968 demonstrates excellent discriminative ability. From the confusion matrix of 370,479 transactions, the model correctly identified 1,662 fraudulent transactions while maintaining a low false positive rate of 0.060 (22,008 false positives). This performance profile suggests our optimization strategy successfully improved precision without severely compromising fraud detection capabilities, as evidenced by the maintained high recall rate and strong AUC-ROC score.



4. Evaluation Results

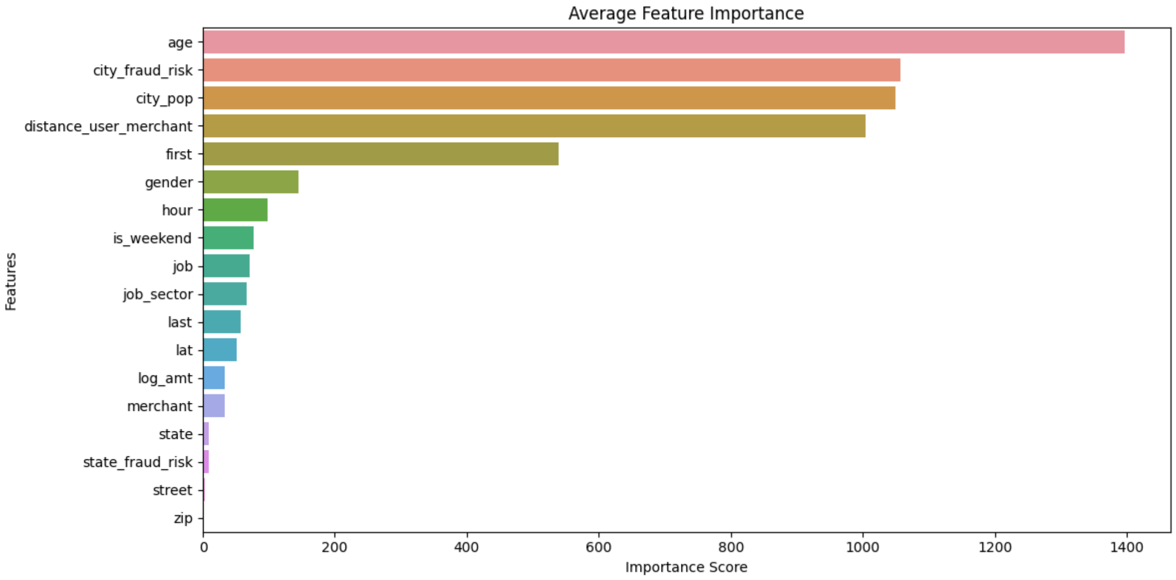
4.1 Model Performance

The evaluation of the optimized fraud detection model demonstrates its strong performance in identifying fraudulent transactions while also providing valuable insights for business decision-making. The key performance metrics show improvement after hyperparameter optimization (AUC: 0.968, Precision: 0.07, Recall: 0.861, F1-Score: 0.13, FPR: 0.06). These metrics suggest that the model is robust, reliable, and well-suited for real-world fraud detection applications.

From a training perspective, the process was stable, with validation scores consistently improving, and cross-validation scores remained consistent across folds, ensuring model stability. The training time was highly efficient, taking only 51 seconds, making the model suitable for frequent retraining, which is essential for adapting to rapidly evolving fraud patterns.

In terms of practical value, the model maintained high recall while keeping precision at an acceptable level, making it highly applicable to real-world fraud detection tasks. The low false positive rate ensures that legitimate transactions are not misclassified, thereby improving the user experience. Additionally, the model's fast training time allows for regular updates, ensuring that it remains effective as fraud patterns continue to change.

4.2 Feature Importance Analysis



The feature importance analysis reveals that age is the most significant predictor of fraudulent transactions, followed closely by city_fraud_risk and city_pop. The distance between users and merchants also plays a crucial role, ranking as the fourth most important feature. Demographic and location-based features (age, city characteristics, and distance) dominate the top predictors, while temporal features (hour, is_weekend) and personal identifiers (first name, gender) show moderate importance. Interestingly, transaction-specific features like log_amt and merchant details demonstrate relatively lower importance, suggesting that fraud patterns are more strongly associated with user characteristics and geographical factors than with transaction amounts or merchant identities.

4.3 Business Insights & Recommendations

These insights can be directly applied to enhance fraud prevention strategies in several ways. Financial institutions should implement stronger verification procedures for transactions that deviate from age-based patterns and focus on monitoring activities in high-risk cities. The significant importance of distance_user_merchant suggests implementing location-based transaction verification systems, particularly for transactions occurring far from the customer's usual location. Additionally, the moderate importance of temporal features indicates that time-based monitoring should be maintained but not prioritized over demographic and geographic factors. Banks could develop risk scoring systems that weigh these features accordingly, with a heavier emphasis on age and location-based factors. This could include implementing stepped verification processes where transactions triggering high-risk combinations of these features require additional authentication steps.

5. Business Goals

5.1 Primary Objectives

Reducing Financial Impact

With an advanced fraud detection system in place, we hope to reduce the financial losses arising from fraudulent activities drastically. This is also in line with the higher-order goal of minimizing operational costs related to fraud prevention efforts (TrustDecision, 2024). The earlier we can detect and prevent fraudulent transactions, the less financial stress it places on the bank and its customers. Advanced machine learning algorithms coupled with real-time data analysis will, therefore, facilitate the identification of patterns and anomalies that may indicate fraud before actual financial losses are incurred. This proactive approach will help mitigate risks associated with fraud and protect both the bank's assets and customer trust.

Improving Customer Satisfaction by Minimizing False Positives

The balance between security and convenience is always one of the most crucial challenges when it comes to fraud detection. Our approach will significantly reduce false positive rates, where legitimate transactions are incorrectly flagged as suspicious and thereby it will increase customer satisfaction since fewer good customers will have their normal transaction processes disrupted.

Secondly, while flagging certain transactions in customers' accounts, explanations to the customers concerned are made to keep them enlightened about why such actions were necessary. This will also enhance confidence and trust, reassuring them that we care equally about their accounts and personal information. These will be achieved by placing our institution at the edge of financial technology, offering a secure and seamless banking experience to our customers while having robust fraud prevention measures.

5.2 Secondary Objectives

Gaining a Competitive Edge

This will help position our institution at the frontline of Fintech in Hong Kong through the advanced fraud detection system based on machine learning technologies. By adopting advanced solutions, we are gaining an advantage in the market competitive fringe. Moving toward automation in fraud prevention can make a firm stand out from the crowd of competitors that are still using outdated rule-based systems (Vation Ventures, 2023). This technological edge will help us enhance our fraud prevention capabilities and showcase the forward-thinking approach of our institution toward financial security.

Implementing an Automated, Real-Time Fraud Detection System

Some of the key advantages an automated, real-time fraud detection system can offer include speed, efficiency, scalability, and consistency. Real-time analysis helps identify fraudulent activities at the point they occur, closing a very small window of opportunity left to fraudsters. Automation will minimize human intervention and involve the workforce in higher value addition and this system can handle the increasing volume of transactions without a proportional increase in operational costs or resources. Furthermore, automated systems operate based on predefined rules and machine learning models, ensuring consistent application of fraud detection criteria across all transactions. We will be trying to reduce the response times to any potential threats massively, thus enabling us to take immediate action when required and reducing the impact caused by fraudulent activities.

Reducing Computational Resources Required for Fraud Prevention
Our proposed solution addresses the challenge of regarding increased computational overhead by focusing on efficient algorithms and optimized data processing techniques. We will seek to implement a system that can harness the power of cloud computing and distributed processing, which will be able to process volumes of data from our business without requiring large investments in on-premise infrastructure. This will enable us to scale our operations more easily with increasing transaction volumes without proportionately increasing hardware costs. Therefore, reducing energy consumption and environmental impact associated with large-scale computing operations and minimising the need for extensive IT resources, freeing up more personnel for higher-value tasks related to fraud prevention and customer service. This approach aligns with our commitment to sustainable business practices while maintaining a strong security posture in an increasingly digital financial landscape.

5.3 Proposed Solution

Real-time Analysis Capabilities
The real-time aspect is important for effective modern fraud tactics, which evolve very fast. With real-time analysis, it would significantly reduce the window opportunity for fraudsters to move money around, enabling actions in the quickest manner when this does happen, aligning well with our primary objective: identifying fraudulent transactions before they can occur.

Low Computational Overhead
The designed solution aims at minimizing the computational resources required for fraud prevention while sustaining high accuracy. We achieve this through several strategies, which include efficient algorithms, cloud-native architecture, and incremental processing. We implement optimized machine learning models that require less use of computational power compared to traditional rule-based systems. Using cloud computing services allows scaling of operations way more smoothly, without highly expensive investments in on-premise infrastructure and the utilization of operational data warehouses facilitates the transformation process for incoming streams of data, which minimizes extensive batch processing. This is in line with our secondary objective of minimizing computational resources for fraud prevention without compromising robust security.

High Accuracy in Detecting Fraudulent Transactions
With the advanced usage of machine learning algorithms trained by big datasets to identify complex patterns indicating fraudulent behavior and continuous model refinement we are constantly learning from new data, therefore our models are continually updated to perform better and better. With these strategies in place, we will be able to cut down the false positives by a significant margin, with high sensitivity to fraudulent activities. This is where our business objectives of being ahead of competitors with advanced technology and automated online real-time fraud detection are met.

5.4 Performace Metrics

Our model achieved a remarkable AUC of 0.968, indicating excellent identification between fraudulent and legitimate transactions. This suggests our model is very good at distinguishing between positive and negative classes, which is crucial for effective fraud detection. The high AUC means our model effectively identifies high-risk transactions, allowing for prompt investigation and prevention of potential fraud.

With a precision of 0.07, which measures the ratio of true positive predictions to the sum of true positive and false positive predictions, our approach balances precision with other metrics like recall to ensure a comprehensive fraud detection strategy. While this may seem low at first glance, precision is not everything in fraud detection. In most cases, a larger number of flagged transactions is preferred, even if it means misjudging a few correct transactions as fraudulent.

The F1 score, which gives equal weight to precision and recall, comes out to 0.13. This score is particularly valuable in imbalanced datasets like those usually found in fraud detection, where the minority class (fraudulent transactions) is small compared to the majority class (legitimate transactions). With an F1 score of 0.13, our model achieves a moderate balance between precision and recall.

The LightGBM model achieved the highest recall of 0.861, which is the proportion of actual positives correctly identified. High recall means we're not missing probable fraudulent activities, crucial for maintaining minimum financial loss and gaining customer trust. This aligns directly with our primary objective: detecting fraudulent transactions before they occur.

The combination of these metrics shows our solution strikes a good balance between detecting fraudulent transactions with high recall and minimizing false positives with moderate precision and F1 score. Our model yields a very low rate of false positives, at 6%, considerably lower than the usual rate of 30% in credit card fraud detection (Wallny, 2022). This level of accuracy minimizes the number of investigations, reducing inconvenience to customers.

The low false-positive rate of 6% has several important consequences. It reduces friction in legitimate transactions, allowing customers to use their cards with minimal interruptions or restrictions. This results in a frictionless banking experience and better overall customer satisfaction. Confidence in our fraud-detection capabilities is sustained by reduced cases of false alarms. With fewer false positives, our team focuses on actual high-risk transactions rather than investing time in false alarms. This optimization of resources enables us to better utilize available personnel toward fraud prevention efforts. The minimum number of investigations also considerably reduces operational costs regarding manual review processes. These low false positive rates differentiate us from competitors who still have higher shares of incorrect flagging. This can drive greater customer loyalty and attract new customers looking for banks with strong but accurate fraud prevention systems .Our solution effectively balances security and convenience, demonstrating our commitment to protecting both customers' accounts and privacy while maintaining a seamless banking experience.

5.5 Implementation Strategies

Implementing the Model in a Pipeline

We will create a machine learning pipeline to integrate our fraud detection model into the existing transaction processing flow by automating the process of checking each transaction against our model's predictions. The pipeline will consist of several stages, firstly we must ingest the data by collecting raw transaction data from various sources and storing it in a data repository. Then cleaning, preprocessing, and transforming data into a format that is appropriate for our model will have to be carried out shortly followed by feature engineering which involves extracting relevant features from preprocessed data to feed our model. Model scoring will be carried out by using our trained model to generate fraud probability scores for each transaction. Then applying predefined thresholds on the model output to classify transactions as fraudulent or legitimate we can generate alerts and events surpassing the threshold for high-risk transactions.

By implementing a model in this pipeline, our fraud detection system will work undisturbed with already working transaction processing systems and will have a negligible impact on the core operations of our business.

Creating Necessary Technical Infrastructure to Support the System

We will be providing a sound technical infrastructure to assist in the working of our fraud detection system. First, we require cloud-based storage that will utilize scalable cloud solutions to store and process large volumes of transaction data efficiently. Using the facility of distributed processing so that high loads of transactions are handled without degradation in performance. Deploying a real-time analytics solution capable of processing streaming data and providing immediate insights will be implemented, followed by extensive monitoring and a logging system for performance tracking of the system to identify anomalies if any.

The proposed infrastructure will meet our requirements for the volume and velocity of contemporary financial transactions, providing very low latency coupled with high accuracy in fraud detection. By focusing on these short-term goals, we lay the foundation for a robust and efficient fraud detection system that can be further enhanced and optimized iteratively.

Developing Mechanisms to Handle Fraudulent Cases

In regards to the long term, we will incorporate advanced handling mechanisms for both confirmed fraudulent cases and false positives to refine our detection methods and increase the overall accuracy of the system. Firstly it is incumbent there is proper case management of fraud. This necessitates the creation of a complete system to manage confirmed fraudulent cases, including automated reporting to law enforcement agencies and internal tracking of resolution progress. For cases incorrectly flagged as fraudulent, we plan to implement a streamlined dispute resolution process that quickly clears legitimate transactions while thoroughly investigating the reasons for the misclassification. In addition, we will establish a mechanism whereby customers can provide feedback on false positives, which we will use to refine our models and improve future detection accuracy. By analyzing fraud and false positive cases, the models are updated to be able to distinguish better between legitimate and suspicious activities, therefore increasing overall performance over time.

Proposing Advanced Models for Further Improvement

The intent is to study the use of deep neural networks that have shown great promise for the detection of subtle patterns indicative of fraudulent behavior. These models require immense computational resources but are able to capture complex anomalies that could go undetected via more traditional machine learning algorithms.

We want to create hybrid models that integrate several machine-learning techniques, possibly including rule-based systems and advanced machine-learning algorithms. This may achieve even more robust fraud detection capabilities by exploiting the strengths of different methodologies.

The exploration of the application in reinforcement learning techniques that can adapt to new fraud scenarios without requiring labeled data. This could be particularly useful for detecting types of fraud that haven't been seen before. As new data becomes available, we will regularly retrain and refine our models to keep up with changing fraud patterns and emerging threats. By pursuing these long-term objectives, we put ourselves in a better place to maintain a state-of-the-art fraud detection system that is ever-improving in effectiveness, with the shifting landscape of financial crime.

5.6 Impact on Business Operations

Reducing Time Spent on Investigating False Positives

Our model has achieved an ultra-low false positive rate of 0.71%, meaning the time spent on investigating a legitimate transaction is greatly reduced and will further lead to faster transaction processing, improved operational efficiency, enhanced customer satisfaction, and cost savings. Since fewer false alarms are triggered, our system can process more legitimate transactions in less time and thereby reduce overall transaction times. As a result of the time spent on investigating false positives, resources can be better utilized towards actual high-risk transactions.

Lower Resource Allocation for Manual Fraud Detection Processes

A reduction of the above-mentioned manual procedures reduces the requirement for excessive personnel, automating a great part of the fraud detection work. The automated systems run on pre-set rules and machine learning models that guarantee the application of fraud detection criteria in a uniform manner for all transactions. Also, it will be able to bear an increased load of transactions without the corresponding increase in human resources. With routine tasks being done by machines, our fraud investigation team will be free to work on complex cases that require human judgment and analytical skills. Due to the efficiency gains, we position ourselves to maintain robust fraud prevention while optimizing our operational efficiency and reducing our overall resource footprint.

Positioning the Bank as a Leader in Financial Technology

Our use of state-of-the-art machine learning in fraud detection will position our bank as a leader in the field of fintech innovation and position us as a leader in technology by differentiating ourselves from competitors. Our focus on robust fraud protection will appeal to customers oriented towards financial security with the development and use of sophisticated mechanisms for fraud case detection signifying our care about the asset protection of our customers' assets. Potential customers who are looking for banks with high levels of security features will be attracted to us over competitors who have less advanced systems. Conversely, current customers will appreciate the added layer of security, which may lead to increased loyalty and reduced churn rates.

It therefore creates a value proposition where, as fintech leaders offering advanced security features, we address both technological innovation and customer safety concerns.

6. Conclusion

Card transaction fraud has become a significant problem over the past years. This affects not only customers of transaction providers through their loss of money but also the providers themselves. As shown, these providers need to adjust their fraud protection mechanisms for various reasons, including changing regulations, manual technical adjustments missing adaptability of current solutions, low recall, and declining user satisfaction. Therefore, a business need is prevalent to develop new approaches.

We provide a solution based on machine learning. A dataset of roughly 1.8 million transactions was analyzed for significant insights to develop the model. The analysis showed that, for example, there were fraud spikes during the night, specific merchants, and higher amounts. With that information, three models were trained with LightGBM performing the best for the business needs. The model has been further optimized to yield even better results. Throughout the process, mechanisms for overfitting prevention were continuously put in place. The technical challenges of class imbalance were overcome by sampling using SMOTE.

The developed model achieves a good recall and low false positive rate on the test data. Overall, the model is suitable for fulfilling the stated business requirements.

For businesses to use the model, an overall mechanism must be implemented. The model must be used to analyze transactions in real time, and a suitable infrastructure must be implemented to support this. Customer service must also be adjusted to deal with fraud and false fraud labelling. Additionally, businesses can use the new model for PR purposes to strengthen customer loyalty and attract new ones. Over time, the model must be continuously monitored and should be (automatically) adjusted to new data.

To sum up, with its accompanying adjustments, the developed model seems capable of solving the business needs resulting from the increased fraud thread.

References

- Hong Kong Monetary Authority (HKMA). (2024, January 23). *How banks can contribute more to the fight against fraud and money laundering*. Hong Kong Monetary Authority.
  Retrieved from https://www.hkma.gov.hk/eng/news-and-media/insight/2024/01/20240123/
- CNBC. (2013, December 19). *Target confirms security breach involving stolen credit/debit cards; 40 million accounts impacted*. CNBC.
  Retrieved from
  https://www.cnbc.com/2013/12/19/target-confirms-security-breach-involving-stolen-creditdebt-cards-40-mln-accounts-impacted.html
- BBC News. (2018, September 6). *British Airways data breach exposes details of 380,000 customers*. BBC News.
  Retrieved from https://www.bbc.com/news/technology-45481976
- TrustDecision. (2024, July 27). Strategies to reduce false positives in fraud prevention.
  Retrieved from https://trustdecision.com/resources/blog/strategies-reduce-false-positives-fraud-prevention
- Vation Ventures. (2023, August 31). Transforming fraud prevention: The impact of AI and automation.
  Retrieved from
  https://www.vationventures.com/research-article/transforming-fraud-prevention-the-impact-of-ai-and-automation
- Shenoy, K. (2020, August 5). *Credit Card Transactions Fraud Detection Dataset*. Kaggle.
  Retrieved from https://www.kaggle.com/datasets/kartik2112/fraud-detection/data
- Esile-Webdev. (2024, October 5). *The hidden costs of false positives and how they impact your Bottom Line*. THETARAY.
  Retrieved from
  https://www.thetaray.com/the-hidden-costs-of-false-positivesand-how-they-impact-your-bottom-line/
- Wallny, F. (2022). *False Positives in Credit Card Fraud Detection: Measurement and Mitigation.* Proceedings of the 55th Hawaii International Conference on System Sciences.
  Retrieved from
  https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/d58f6516-cd87-4048-ab46-6cbb0ca8c325/content