Lidia Puspa N

Developing a Fraud Detection Model at GrowBank

Executive Summary

Fraud remained a persistent and growing issue in GrowBank's B2C transactions throughout 2023. That may reflect either actual fraud incidents or misclassification of legitimate transactions.

The Facts

USD 1,6 M
Fraudulent transactions out of 4.9 M total in 2023

32%
Average Monthly Fraud Rate

3x
Avg. failed attempts in 7d for fraudulent transactions

0.8
Median risk score of fraud transactions

Goal

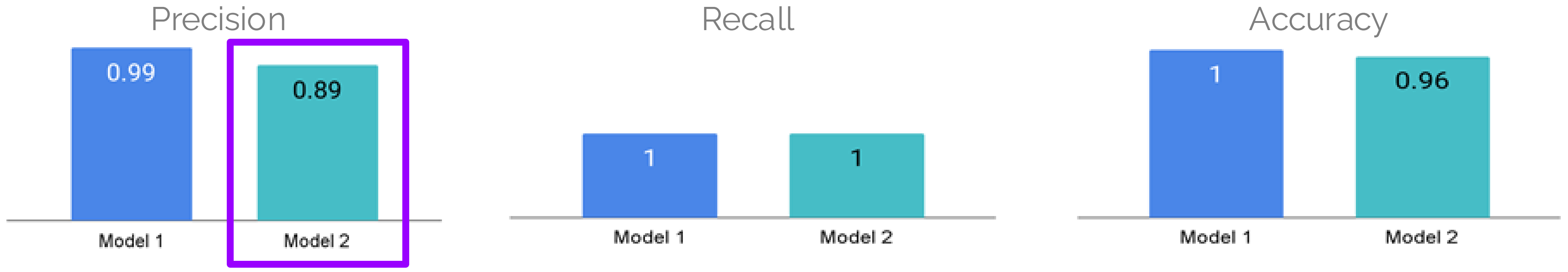
Develop fraud detection model with precision > 80%

Key Metrics

Precision (%)

False Positive Rate (%)

Key Findings



Model 1 had perfect metrics but was overly overfit and unrealistic for real-world conditions.
Model 2 was more realistic with 89% precision — there were still some false positives, but no fraud cases were missed. This balance made Model 2 the preferred choice for production.

Recommendation

- Adaptive thresholds monitoring for risk scores
- Add new features
- Develop anomaly detection

Tools Used: python Google Sheets tableau

Grow Bank

Outline

01

Background

02

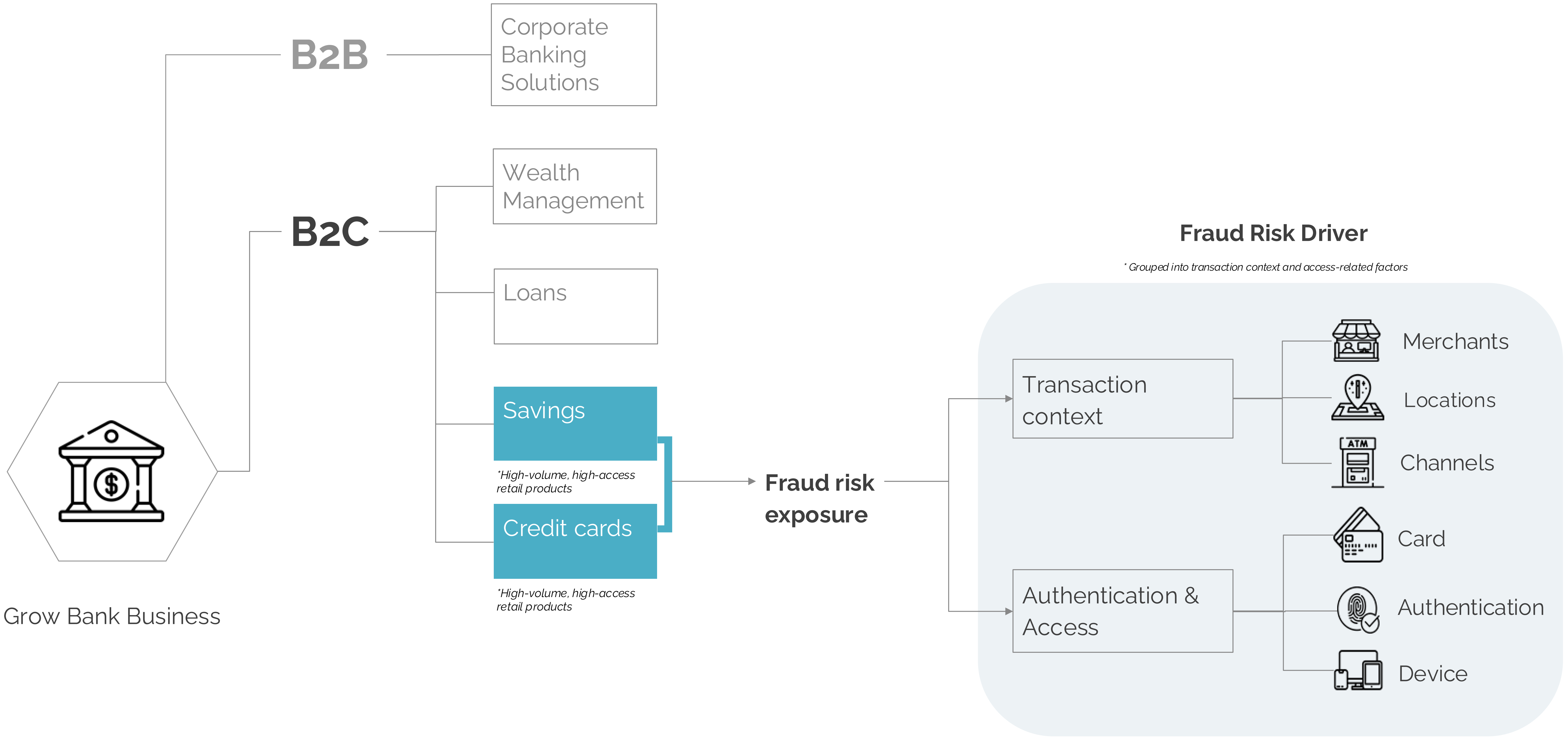
Analysis & Model Development

03

Insight & Recommendation

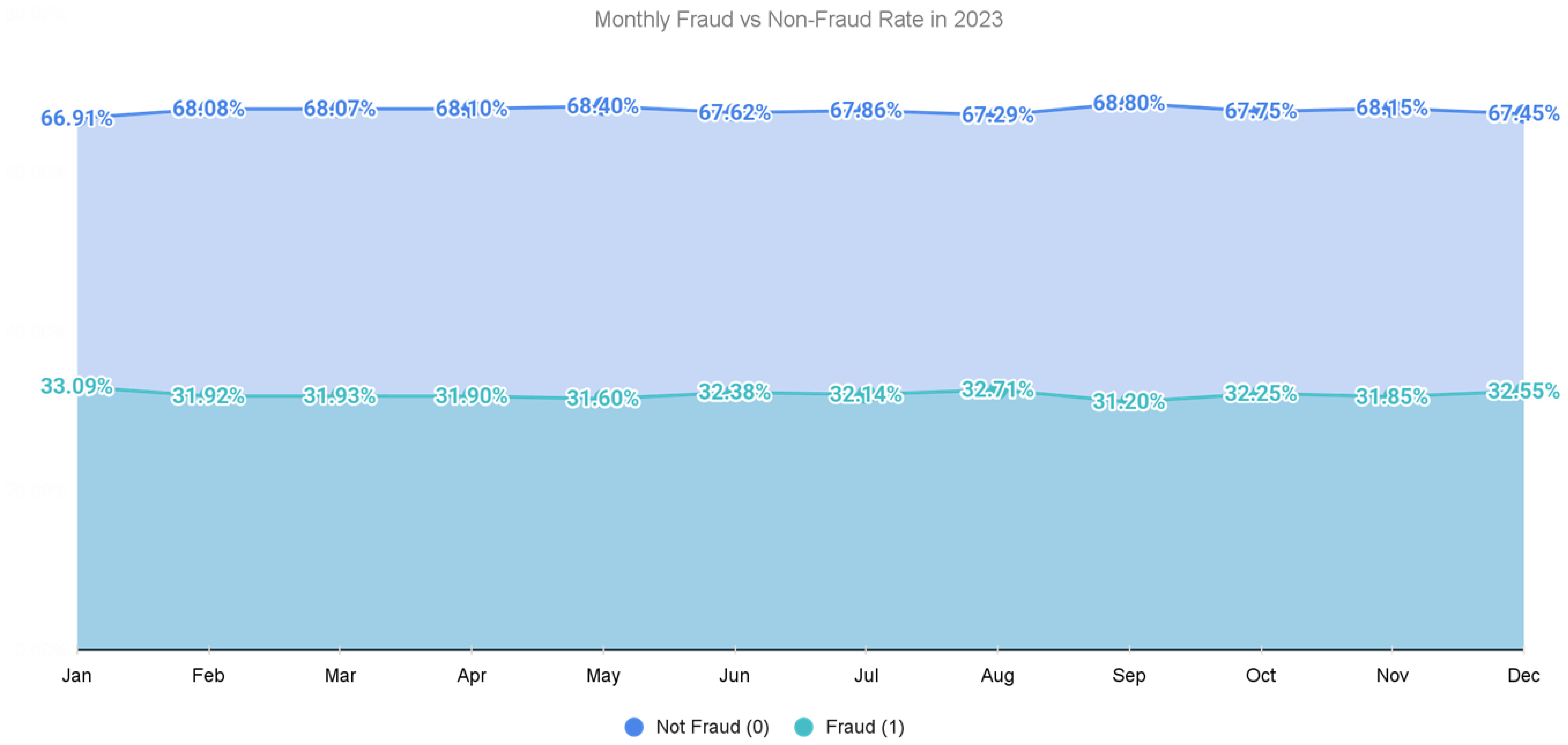
Background

High Transaction Volume and Broad Access in Savings and Credit Cards Expose Fraud Risk



Consistently High Fraud Flags Signal Possible Over-Detection

- 32% of B2C transactions were flagged as fraud every month in 2023, totalling USD 1.6 million for the year.
- This indicates a persistently aggressive detection pattern likely capturing not only real fraud but also false positives.



Over-Detection Drives Down Trust and Missed Revenue Opportunities

Up to 2/3 of declined transactions are false positives

37% Of customers with bad fraud experience closed or abandoned their accounts.

USD 1 Million

in estimated transaction value lost due to wrongly declined legitimate payments in 2023

≈ USD 20,5 K

in estimated lost revenue from false positive declines

(assuming 70% credit card @ 2.5% fee and 30% debit @ USD 1 flat fee)

Source: A new approach to fighting fraud while enhancing customer experience (McKinsey, 2022)

Grow Bank

Problem Statement

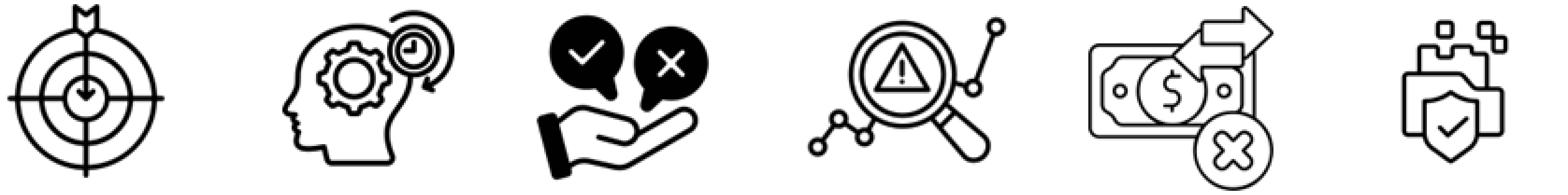
Chief Risk Officer (CRO) → Data Analyst Team

How can we analyze the Grow Bank transaction data to **identify patterns** of fraud and develop a predictive **fraud detection model** that achieves at least **80% precision** within the **next 3 months**?

Objectives

- **Analyze** trends and characteristics of the fraudulent transactions.
- **Identify** Key Fraud Indicators
- **Develop** a Fraud Detection Model to improve detection precision

Metrics



- Precision (%)
- Recall (%)
- False Positive Rate (%)
- Fraud Rate (%)
- Failed Transaction Attempts (%)
- Loss Prevented (USD)

DARCI

ROLE	PIC
Decider	Chief Risk Officer
Accountable	Head of Data Protection & Fraud Risk
Responsible	Data Analyst
Consulted	Fraud Prevention Executive Cyber Security Team Legal Team
Informed	Operations Team IT & Infrastructure Team Finance Team

About The Data

50,000 transaction records in 2023

Risk & Security Indicators:
IP_Address_Flag, Previous_Fraudulent_Activity, Risk_Score, Authentication_Method

Location-Based Features:
Location, Transaction_Distance

Target Variable:
Fraud_Label (0 = Not Fraud, 1 = Fraud)

Transaction Details:
Transaction_ID, Transaction_Amount, Transaction_Type, Merchant_Category, Timestamp, Is_Weekend

User Information:
User_ID, Account_Balance, Card_Type, Card_Age, Device_Type

Historical Behaviour::
Daily_Transaction_Count, Avg_Transaction_Amount_7d, Failed_Transaction_Count_7d

This synthetic transaction data is from *[Fraud Detection Transactions Dataset (Kaggle.com)](https://www.kaggle.com)* *and* intended solely for educational and research purposes.

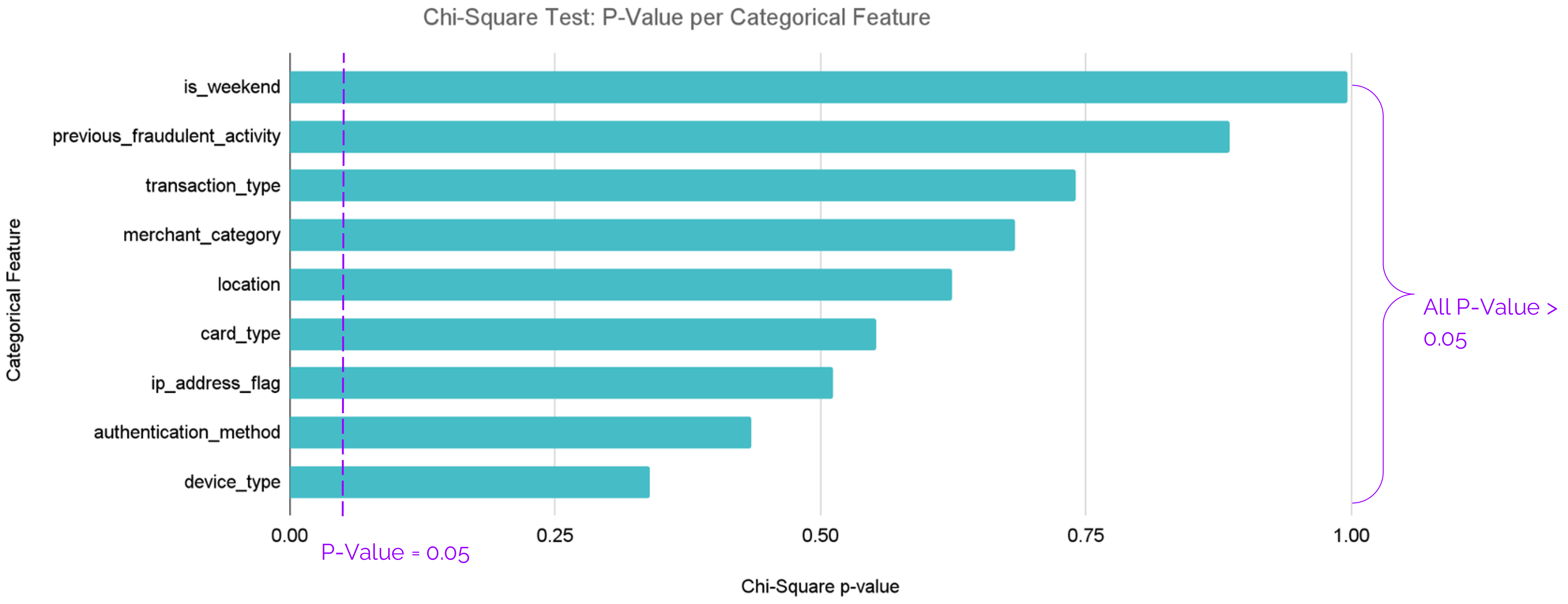
Analysis Results

How to Develop Fraud Detection Model?

- **How can we develop a model with at least 80% precision in 3 month?**
  - transaction related factor
    - Unusual Spending Behavior
      - Fraudulent transactions tend to have unusually high amounts compared to a user's normal spending.
      - Fraud often happens on weekends when monitoring is weaker.
    - Suspicious Locations & Devices
      - Fraudsters often make purchases far from the user's usual location.
      - Using flagged IP addresses increases the fraud risk.
  - User & Authentication Risks
    - Risky Users with Past Fraud History
      - Users who committed fraud before are more likely to do it again.
      - Users who fail multiple transaction attempts may be testing stolen data.
      - A high risk score may indicate a user's involvement in fraud.
    - Weak Authentication & Payment Methods
      - Fraudsters may prefer easy-to-bypass security methods like passwords.
      - New cards are often used for fraud before detection systems catch them.

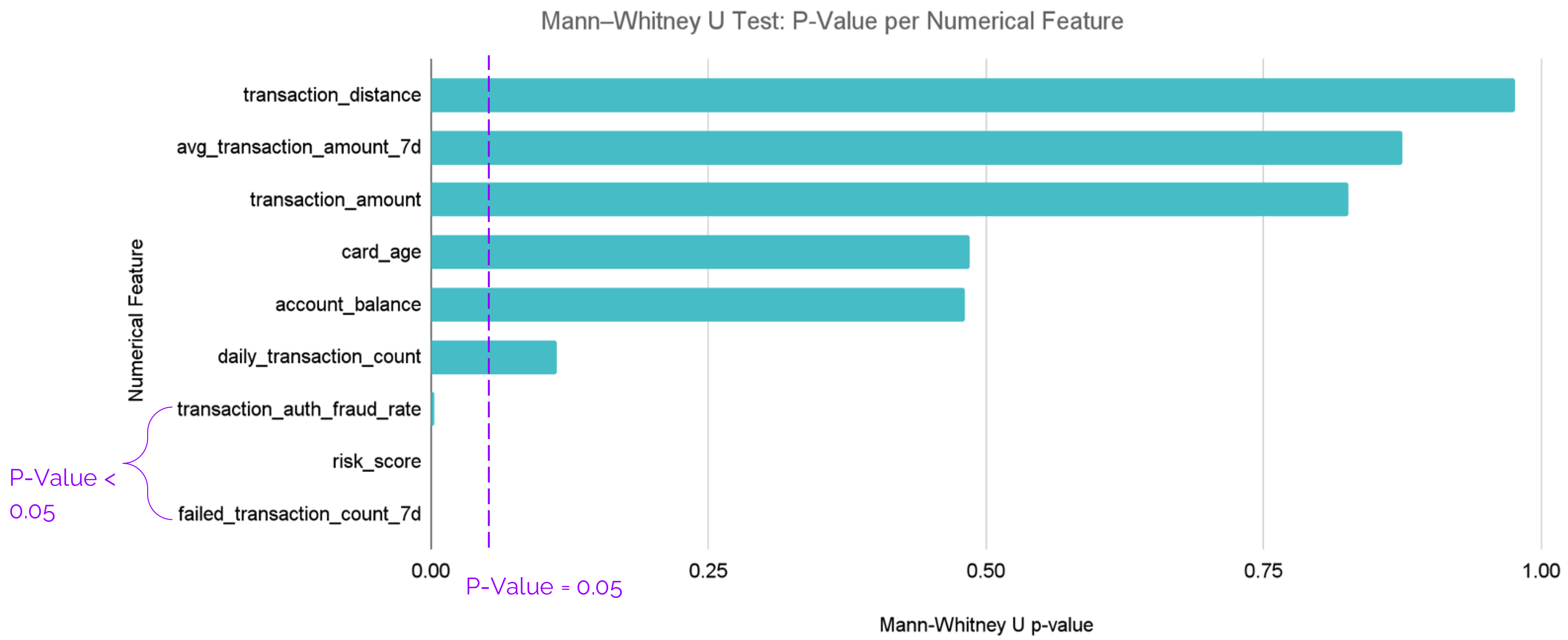
Categorical Features Aren't Statistically Differentiate Fraud

Chi-Square tests show no significant difference between fraud and non-fraud for features like transaction type, device, and location—indicating low predictive value.



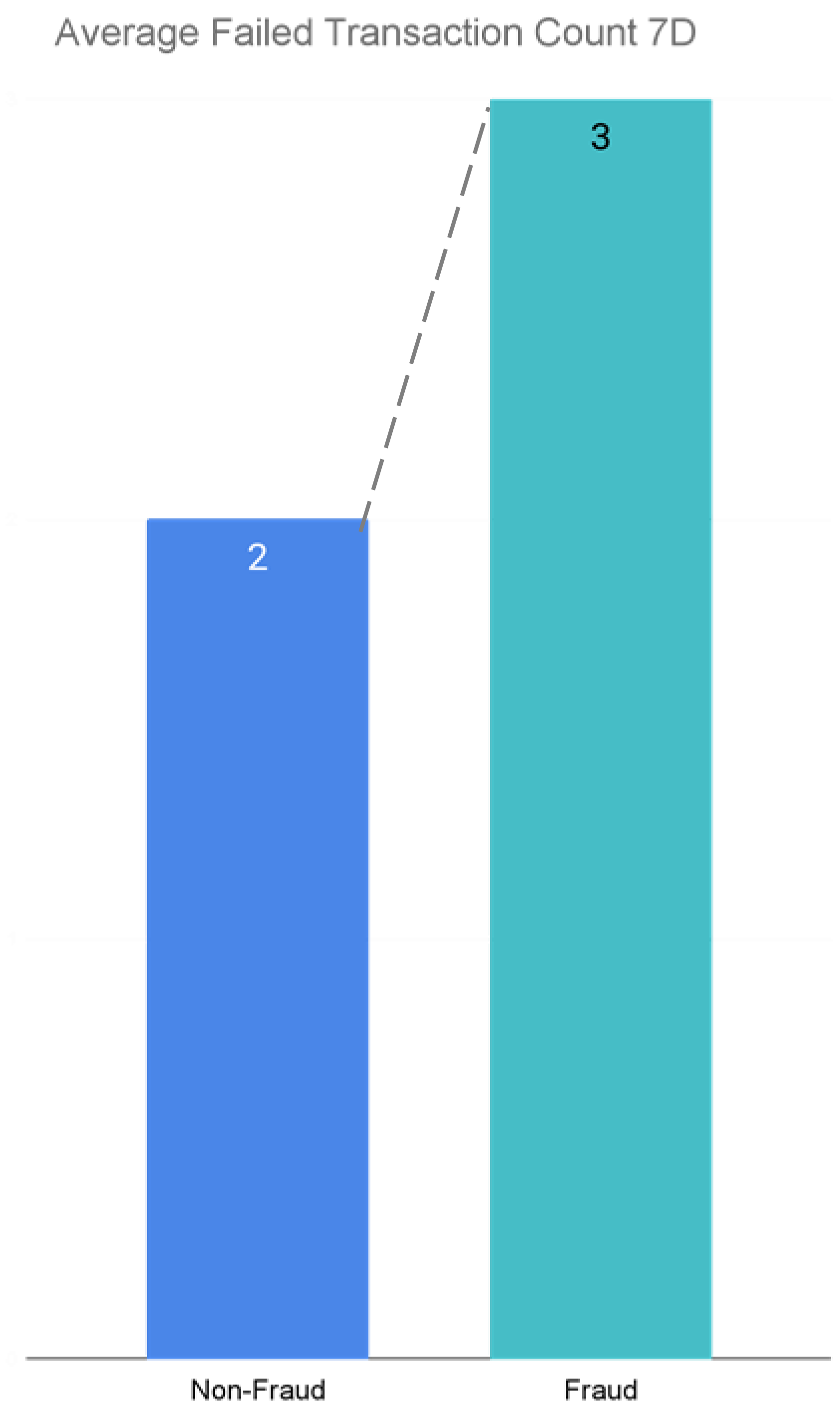
3 Features Strongly Differentiate Fraud from Legitimate Transactions

Failed transaction attempts, risk score, and authentication fraud rate—show statistically significant differences in distribution between fraud and non-fraud transactions *($p < 0.05$)*.



Failed Attempts Are a Strong Signal of Fraud

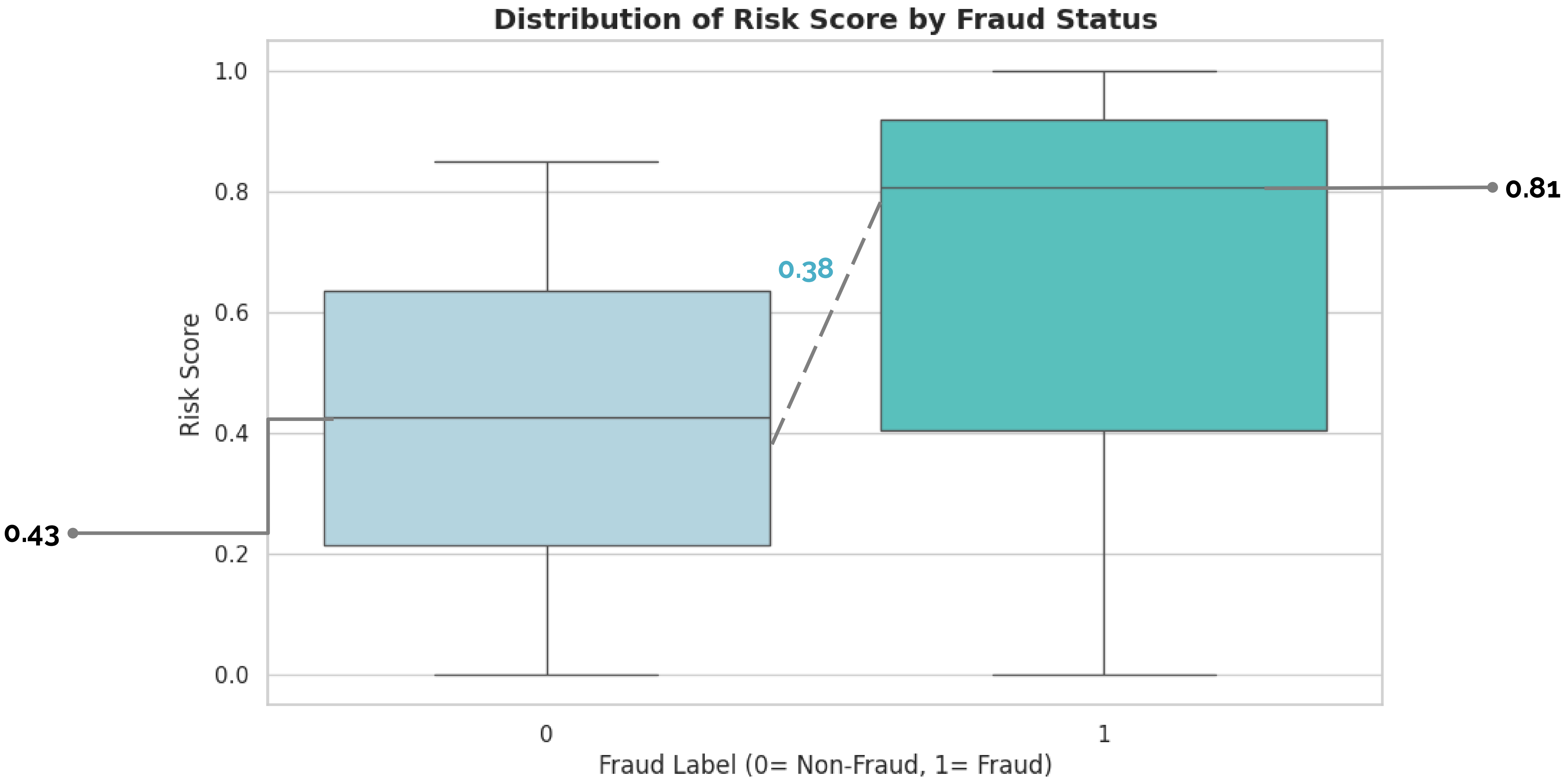
- Fraudulent transactions tend to have a slightly higher average number of failed transactions in the past 7 days compared to non-fraudulent ones
- This suggesting potential behavioral differences worth further exploration.



Grow Bank

Fraudulent Transaction Tend To Have A Higher Risk Score

- Fraud transactions show a 0.38-point advantage in median risk score compared to non-fraud.
- Fraud transactions are more concentrated in high risk score ranges, indicating strong signal.
- Non-fraud transactions tend to cluster in lower risk score ranges, showing less threat.

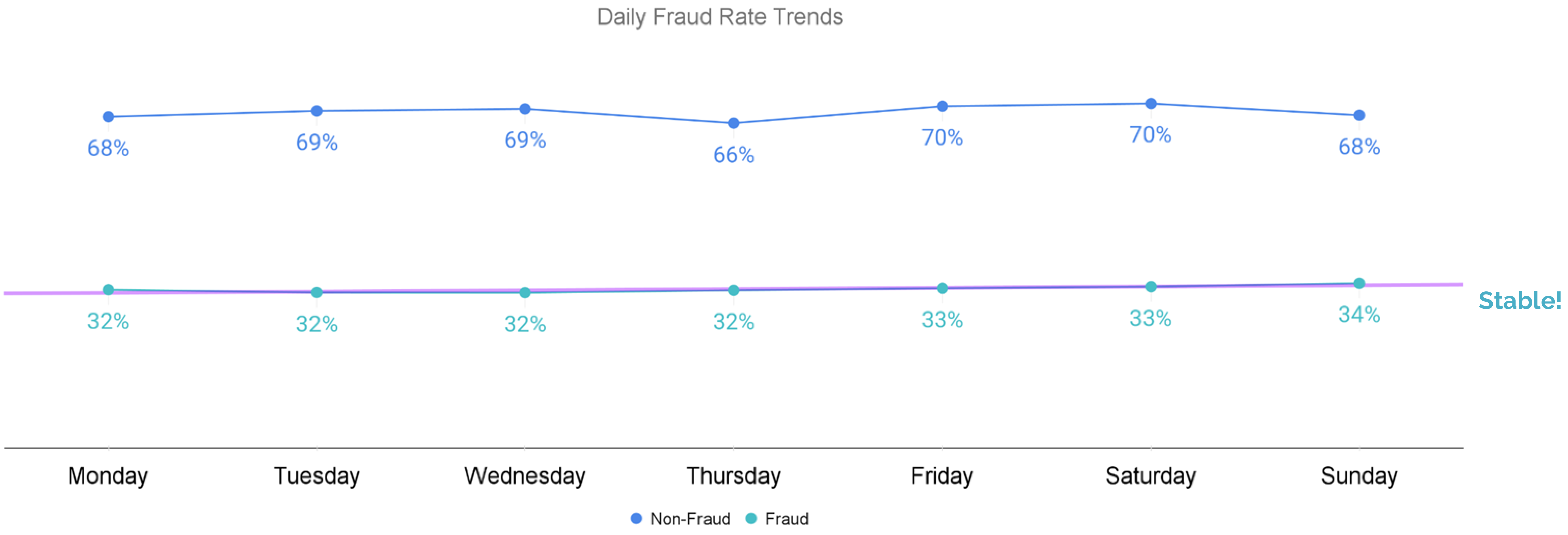


Transaction Authentication Fraud Rate Slightly Differentiates Fraudulent from Legitimate Transactions

While the median difference is minimal (0.322 vs 0.321), this feature still offers a weak but valuable signal of fraud risk linked to specific transaction-authentication combinations.

Fraud Occurs Steadily Throughout the Week, With No Clear Daily Pattern

- The distribution of fraudulent transactions remains fairly stable throughout the week, no sharp spikes indicating specific time-based fraud patterns.
- This suggests that fraud attempts occur regularly, regardless of day
- Reinforcing the need for continuous, 24/7 fraud monitoring rather than relying on time-based rules.



EDA Findings & Model Development Plan

Findings

- 32% of transactions are fraudulent
- Fraud transactions fail more often (avg. 3X)
- Transaction Auth. Fraud Rate can slightly differentiates fraud
- No Categorical Feature can differentiate fraud
- High risk scores = strong fraud indication

Model Experiments using Random Forest Algorithm

Model	Key Feature	Remarks
Model 1	risk_score	Baseline model using all feature
Model 2	risk_score_bin	Prevent overfitting

Metrics: Precision,Recall, F1-Score, Accuracy and ROC-AUC

Model Result and Evaluation

Grow Bank

Risk Score Binning Improves Robustness with Minimal Trade-offs

Model	Treatment	Precision (fraud)	Recall (fraud)	F1-Score (fraud)	ROC-AUC	Accuracy	Strengths	Weaknesses
1	risk_score	0.99	1.00	1.00	1.00	1.00	Very high precision and recall, perfect accuracy.	High risk of overfitting, may not generalize well
2	risk_score_bin	0.89	1.00	0.94	0.988	0.96	Strong recall, better generalization, reduced overfitting risk	Slightly more false positives, but still within acceptable range

Recommendation?

Model 2

A reasonable trade-off, maintaining 100% recall while minimizing the overfitting risks observed in Model 1.

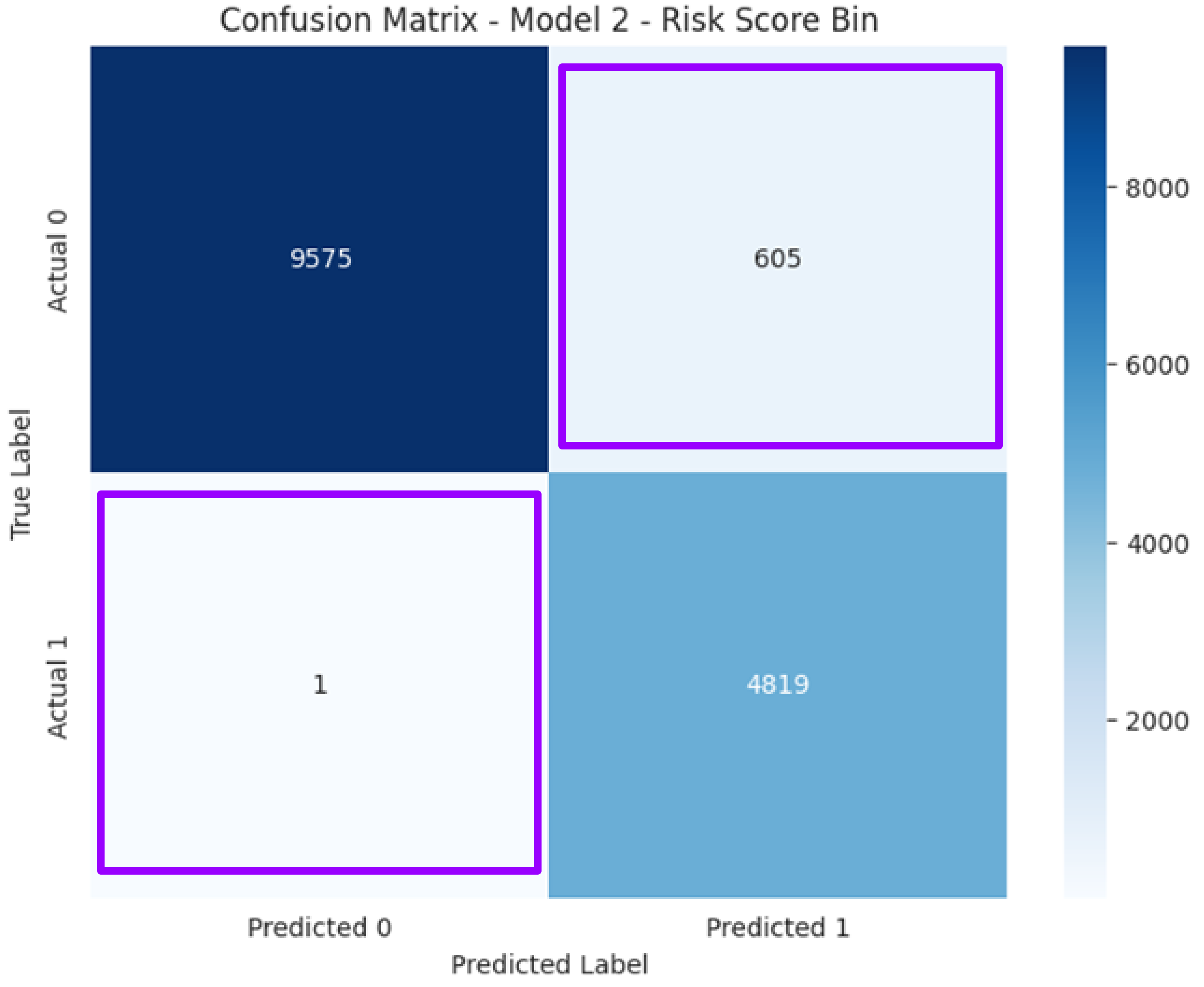
Model 2 Evaluation Result

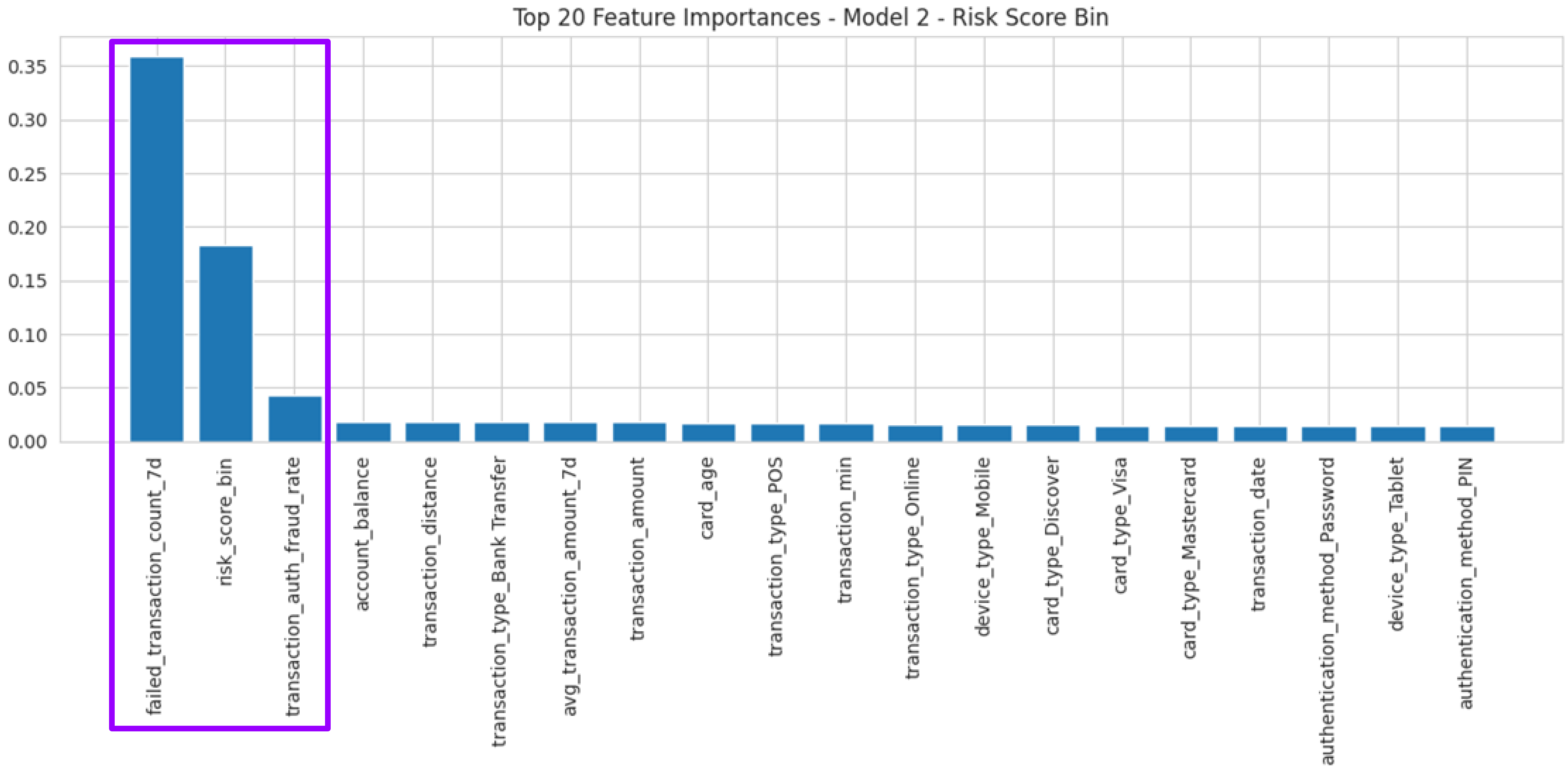
- 1 false negatives, 605 false positives, The model maintains strong recall performance but experiences a slight drop in precision.
- Feature contribution is more balanced compared to Model 1, although failed transaction count (7 days) and risk score bin remain the top contributors.
- This suggests reduced overfitting risk, as the model now learns from a broader set of features beyond the two dominant ones.

Classification Report:

	precision	recall	f1-score	support
0	1.00	0.94	0.97	10180
1	0.89	1.00	0.94	4820
accuracy			0.96	15000
macro avg	0.94	0.97	0.96	15000
weighted avg	0.96	0.96	0.96	15000

ROC AUC Score: 0.9886





Key Insights

Fraud occurs consistently across all times and channels.

Rule-based detection alone is not enough. A real-time, behavior-based approach is required.

Risk Score and Failed Transaction Count 7d are strong fraud indicators, but raw risk score leads to overfitting.

Using Risk Score Bin provides a safer and more balanced model for real-world application.

The current model is stable but still requires behavioral features to improve.

To better capture subtle fraud patterns and reduce false positives.

Proposed Solutions

Grow Bank

Adaptive Thresholds Monitoring For Risk Scores

Ensuring stable fraud detection thresholds through data-driven monitoring

1. Set Threshold Based on Real Data

- **Low (< 0.4)** → based on the median of non-fraud risk scores
- **High (> 0.8)** → based on the median of fraud risk scores
- **Mid-range (0.4–0.8)** is treated as a gray zone

2. Define Tolerance for Re-Evaluation

- Example:
  tolerance = ±0.05 → threshold is reviewed if the fraud median shifts more than +0.05 or lower than -0.05
- Can be determined based on **historical behavior**, business impact, or risk appetite

Optional dual-layer system:

- **0.05** → *Soft alert* (flag for monitoring)
- **0.1** → *Hard review* (requires immediate threshold evaluation)

This helps align system responsiveness with operational needs.

3. Monitor and Track Deviation (Quarterly)

Every quarter:

- Calculate the **median fraud score based on data in current quarter**
- Measure **deviation from the active threshold**

Deviation from Threshold	Status
≤ ±tolerance	KEEP
> ±tolerance	EVALUATE❗

4. Take Action ,Log Results and Track Through KPIs

- If status is **EVALUATE** → review threshold **before next quarter**
- Document all evaluations:
  - Period, median, deviation, status, action
- Keep a clear audit trail
- To ensure effectiveness, we tracked our success using KPIs such as Average Alert Evalute Time

This ensures auditability and supports future policy decisions.

Threshold Monitoring Simulation

Period	Threshold (Active)	Median Fraud	Median Non-Fraud	Deviation Median Fraud vs. Non-Fraud	Threshold - Median Fraud	Alert	Action
Q4-2023	0.80	0.80	0.42	0.38	0.00	KEEP	No Action Needed
Q1-2024	0.80	0.86	0.41	0.45	0.06	**EVALUATE❗**	**Update Threshold❗**
Q2-2024	0.86	0.83	0.42	0.41	0.03	KEEP	No Action Needed
Q3-2024	0.86	0.80	0.42	0.38	0.06	**EVALUATE❗**	**Keep Threshold**
Q4-2024	0.80	0.78	0.42	0.36	0.02	KEEP	No Action Needed

**Threshold = 0.80, Tolerance = 0.05*

Grow Bank

New Features Recommendation

Transaction Time Gap

Measures the time interval between a user's most recent transactions.

Example:
If the previous transaction happened at 13:00, and the current one at 13:05 → **Time Gap = 5 minutes**

Repeat Transaction Amount

Detects users who repeat transactions with the same or similar amount.

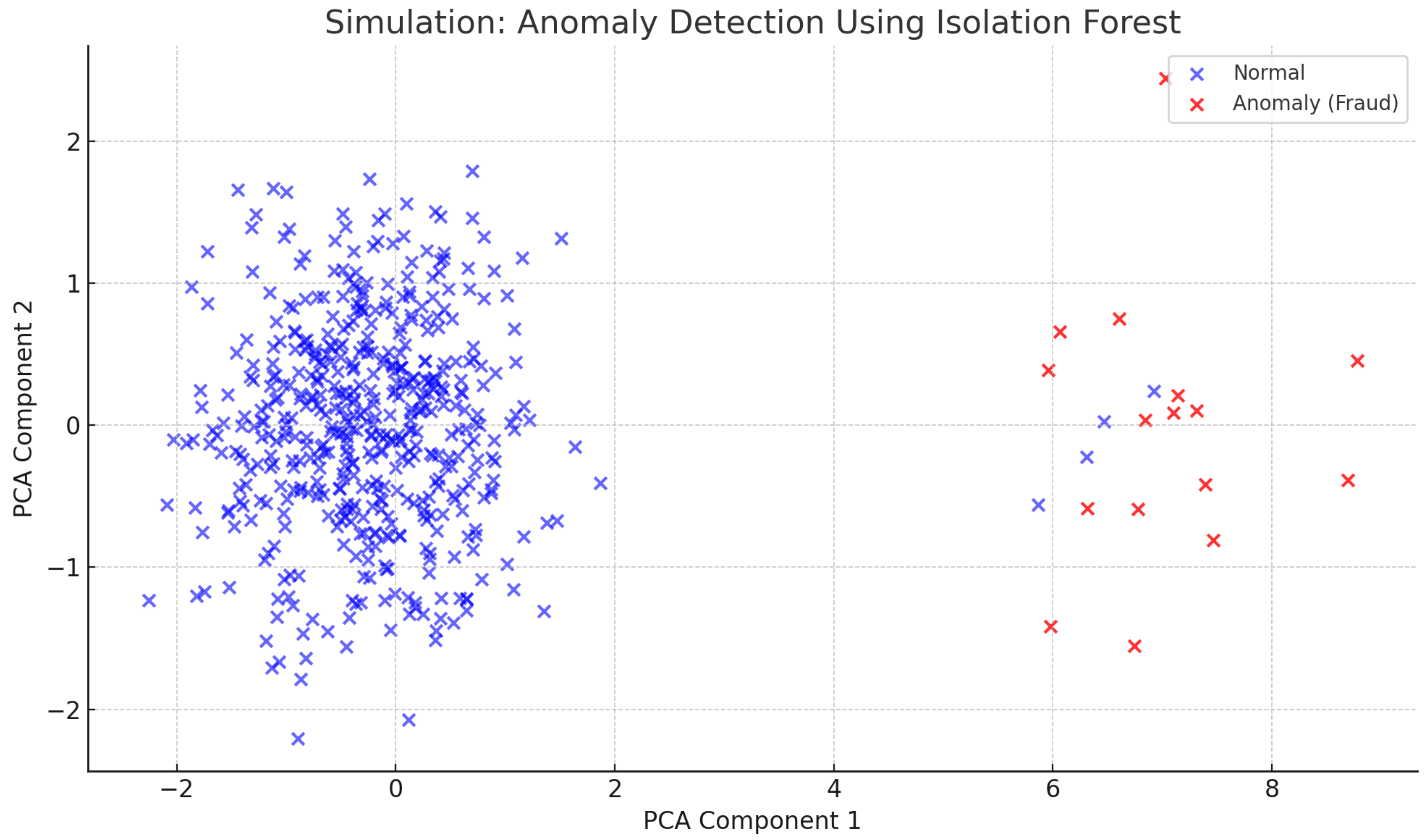
Example:
If a user transfers Rp 100,000 three times in one day → this is considered a **repeat**

Rare Device Flag (Device Familiarity)

Identifies whether the device used in the current transaction is **rarely used** by the user.

Example:
If the user usually logs in from "Mobile Android", but now logs in from "Windows Desktop" for the first time → this is considered a **rare device**

Anomaly Detection with Isolation Forest



This visualization illustrates how Isolation Forest effectively isolates suspicious transactions that deviate from the normal cluster.

Purpose

Implement an unsupervised learning model as an additional feature to identify "anomalous" transactions.

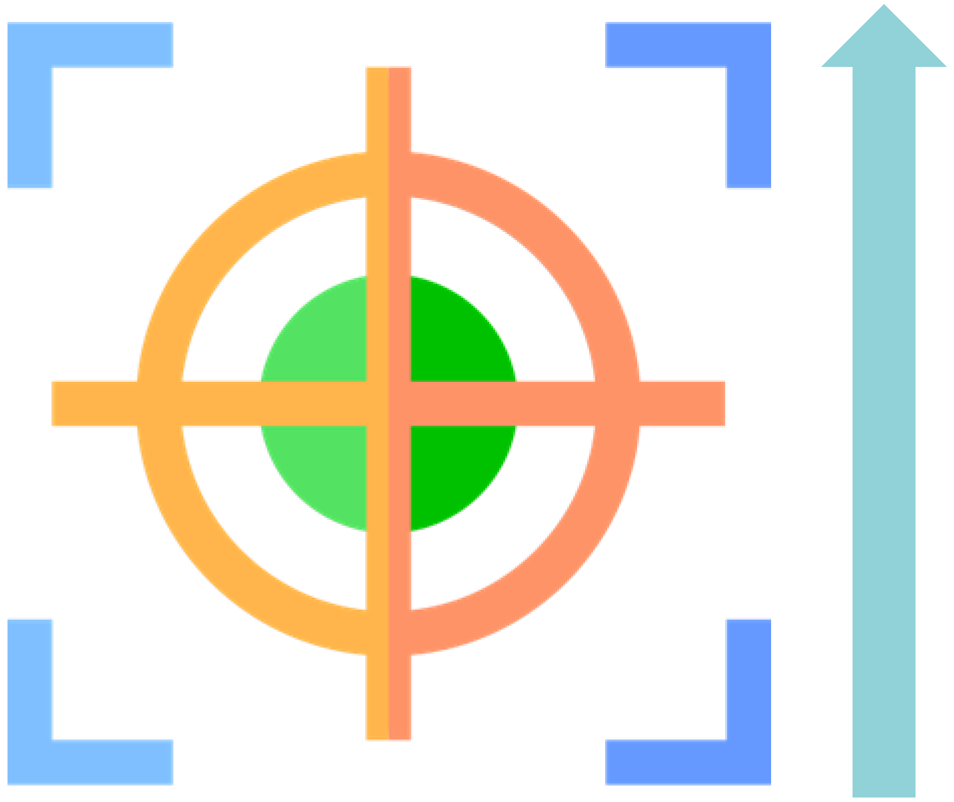
How it Works

Each transaction is assigned an anomaly score. This score is added as an input to the main fraud model, helping it better distinguish real fraud from false positives.

Key Benefits

- Improves precision by reducing false positives.
- Helps the model catch suspicious behaviour early.

Projected Impact on Precision Goal: Estimated 93% by End of 2024



Current Precision (Baseline)
➔ 89% (Q4–2023)

Adaptive Threshold Monitoring

Dynamically aligns the fraud score cut-off with shifting fraud patterns
→ Potential Precision Gain: +1%

Behavioral Feature Enrichment

Adds context-aware signals like transaction gap, repetition, and rare device use
→ Potential Precision Gain: +1–2%

Anomaly Detection

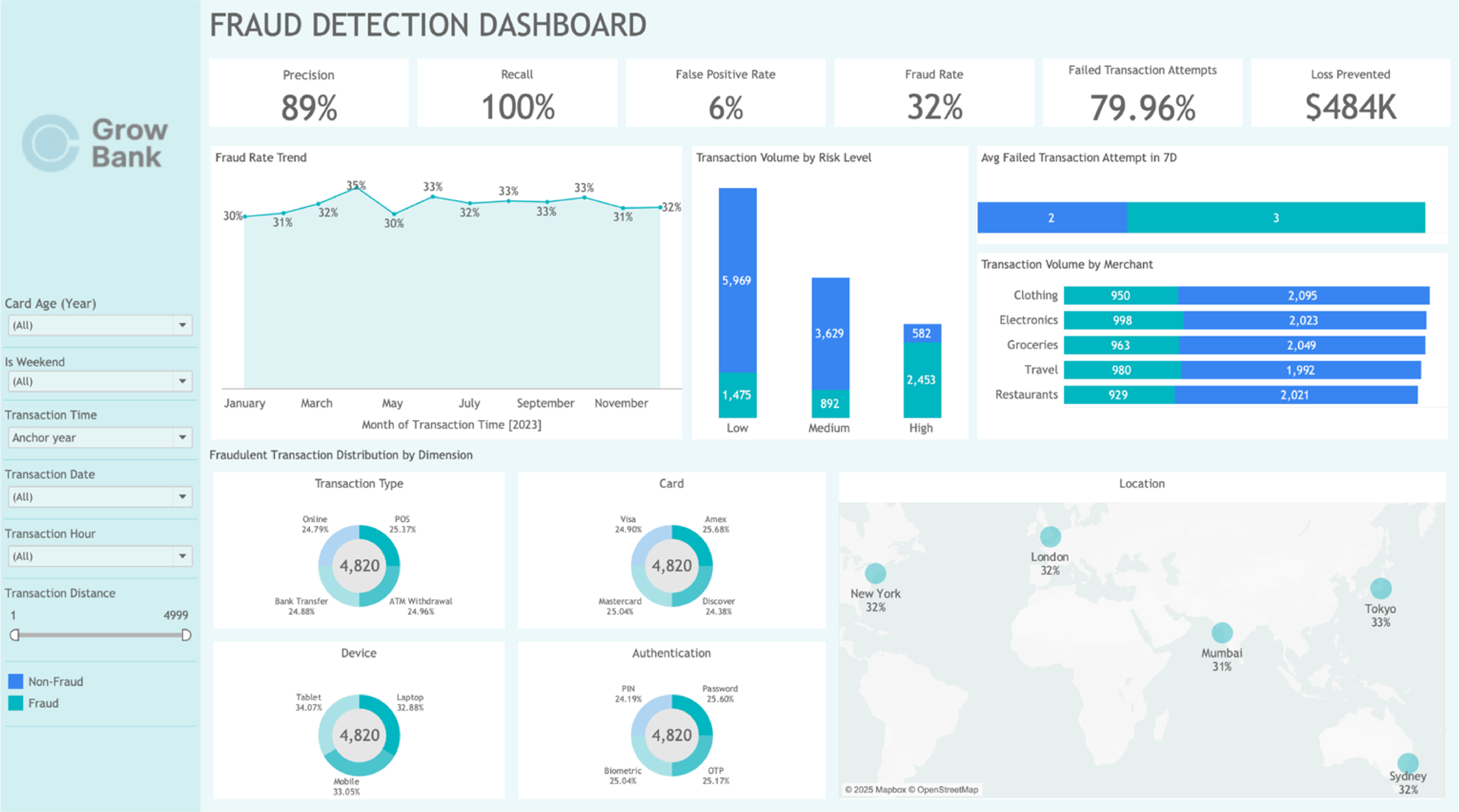
Add as an additional feature
→ Potential Precision Gain: +1-2%

With these three strategies, Grow Bank is projected to reach 92–93% precision by the end of 2024. If false positives are reduced from two-thirds to one-third, the estimated financial loss from false declines could drop from USD 1 million to USD 500K.

"By focusing on smarter thresholds, stronger features and anomaly detection, we not only prevent fraud—but prevent misjudging customers."

Fraud Detection Dashboard

Grow Bank

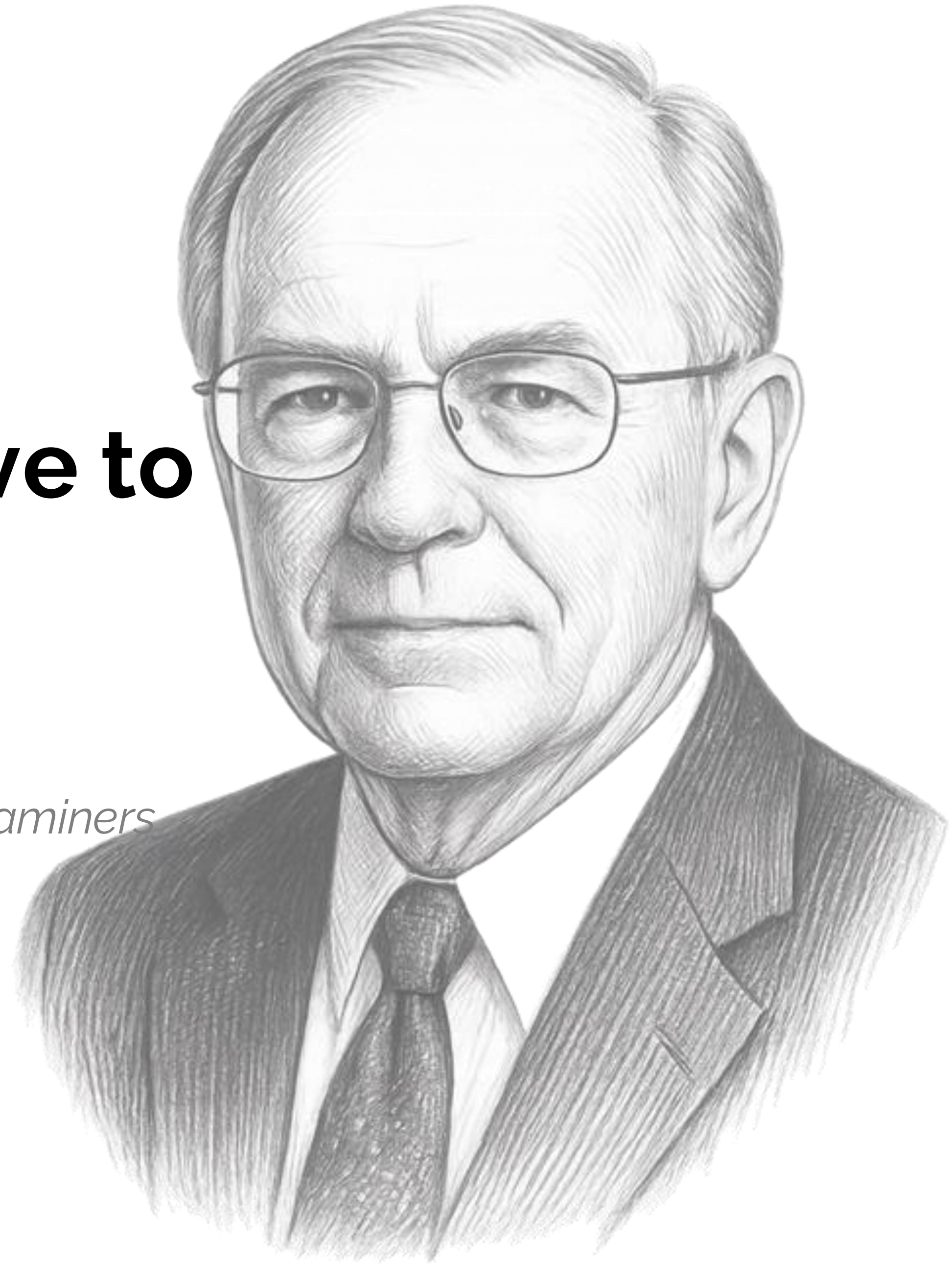


[LINK]

"Every fraud has a trail. You just have to know where to look."

— Joseph T. Wells

founder and Chairman of the Board of the Association of Certified Fraud Examiners (ACFE), the world's largest anti-fraud organization.



Grow Bank

Thank You

Appendix

Documentation

Project Scope

Dataset Description

The dataset consists of 50,000 transaction records from 2023 across 8,963 unique users in five countries.

Focus Areas

Identify trends and characteristics associated with fraudulent transactions and developing a predictive fraud detection model.

Exclusions

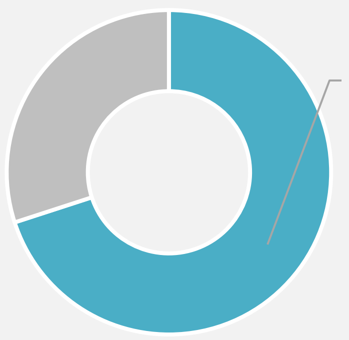
- The analysis will not include non-transaction-related data such as customer service interactions or account management processes.
- This project will not compare or analyze existing models due to limitations in available information.

Precision isn't Just a Technical Metric. It Protects Revenue, Reputation, and Long-term Customer Loyalty

$5.9B

Fraud losses in the US (2021)
(+436% vs 2017)

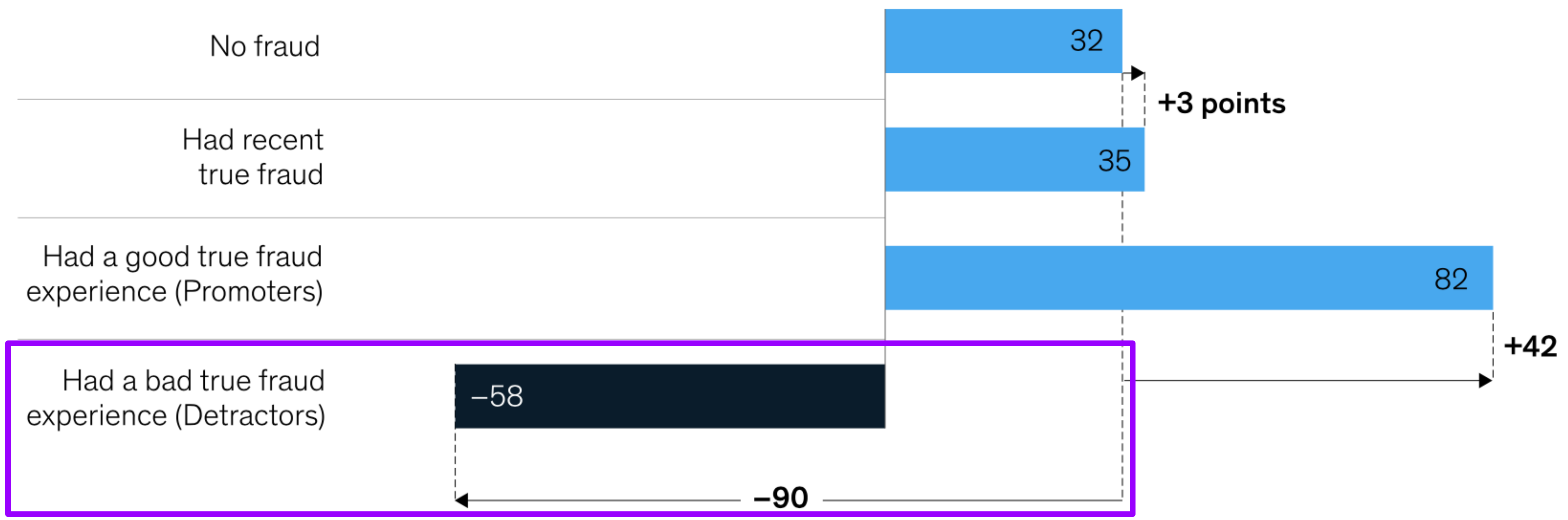
Up to 2/3 of declined transactions are false positives



70% Fraud victims report stress & dissatisfaction

When companies respond well to fraud events, customers report higher levels of satisfaction.

Average customer satisfaction score for different customer groups, illustrative

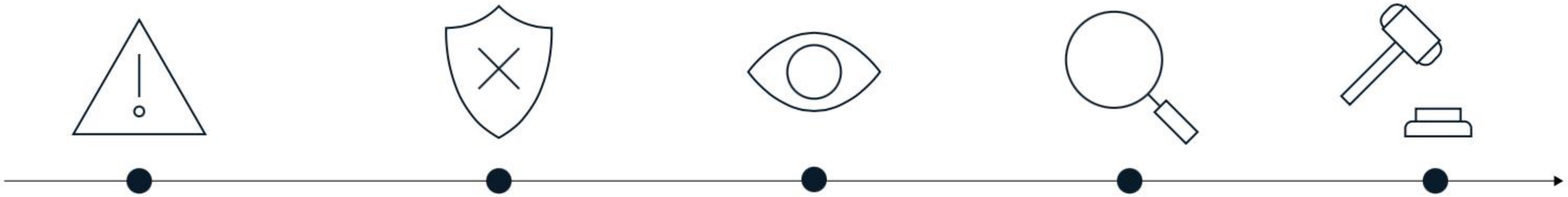


37% of customers with bad fraud experience closed or abandoned their accounts.

Source: A new approach to fighting fraud while enhancing customer experience (McKinsey, 2022)

A Best-practice Fraud Prevention Strategy Should Be Geared Toward Preserving Positive Client Experience

Leading organizations use machine-learning algorithms and strive to utilize all available data to achieve a step change in the accuracy of fraud detection. They seek to reduce noise (false positives) and the risk that fraudulent transactions are missed (false negatives).



Deterrence

- Cross-industry collaboration
- Intelligence
- Public communication

Prevention

- Risk assessment
- Controls and usage strategy
- Awareness and education
- Authentication

Detection

- Data analytics
- Operationalization

Investigation

Dispute handling

McKinsey & Company

Source: A new approach to fighting fraud while enhancing customer experience (McKinsey, 2022)

Dataset Overview

Dataset Overview

The dataset contains **50,000 transaction records in 2023** from **8,963 unique users**, with transactions sample originating from **five countries**: **London, Mumbai, New York, Sydney, and Tokyo**. It includes **21 columns** representing various aspects of fraud monitoring in financial transactions, categorized as follows:

- **Transaction Details:** Transaction_ID, Transaction_Amount, Transaction_Type, Merchant_Category, Timestamp, Is_Weekend
- **User Information:** User_ID, Account_Balance, Card_Type, Card_Age, Device_Type
- **Behavioral Features:** Daily_Transaction_Count, Avg_Transaction_Amount_7d, Failed_Transaction_Count_7d
- **Risk & Security Indicators:** IP_Address_Flag, Previous_Fraudulent_Activity, Risk_Score, Authentication_Method
- **Location-Based Features:** Location, Transaction_Distance
- **Target Variable:** Fraud_Label (0 = Not Fraud, 1 = Fraud)

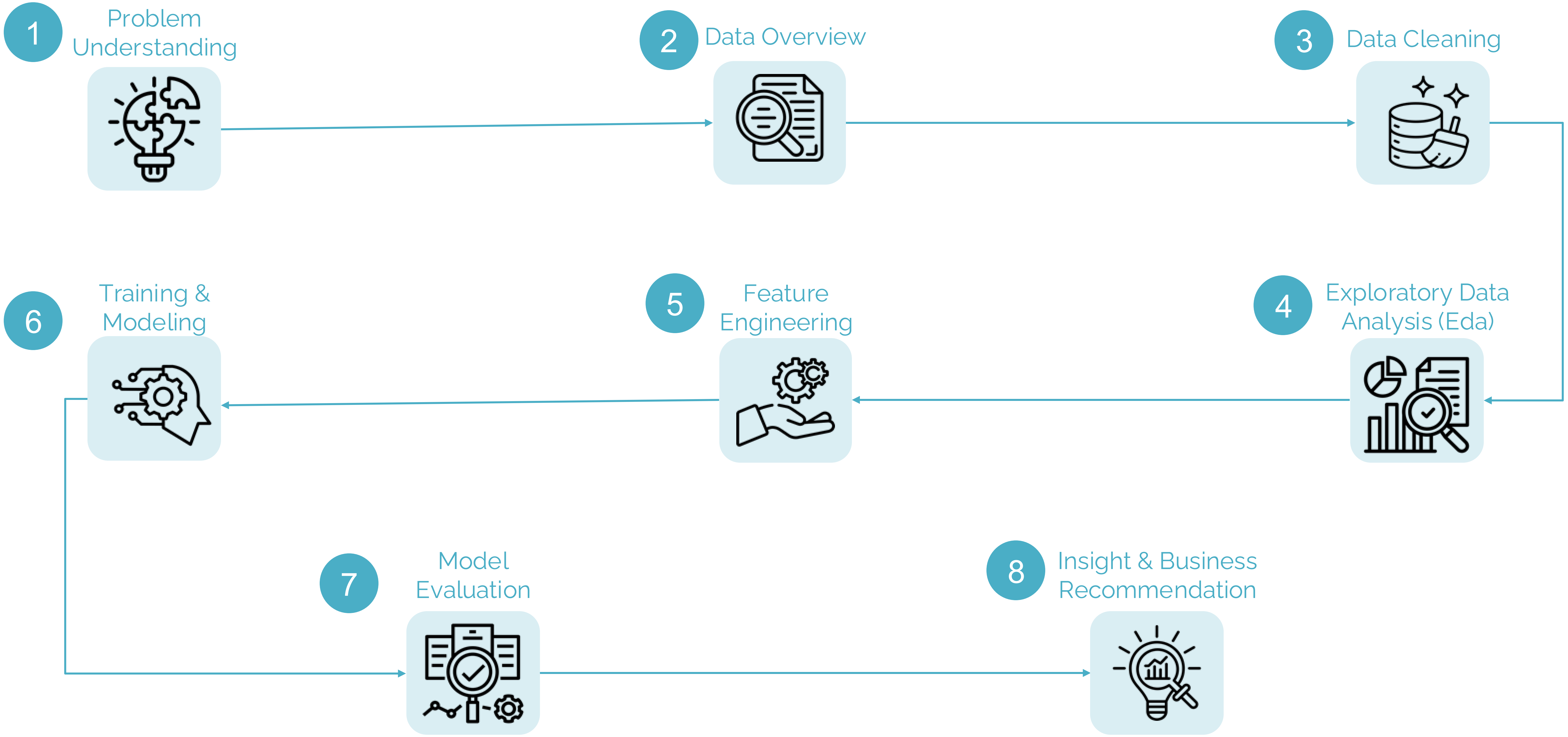
Dataset Source

This synthetic transaction data is from **[Fraud Detection Transactions Dataset (Kaggle.com)]**.

Disclaimer

- This dataset is intended **solely for educational and research purposes**.
- This analysis in based on the public data from Kaggle and the author created this dataset using **simulated financial transaction data to reflect common patterns in fraud detection.**
- This project focuses on exploratory data analysis (EDA) and the development of a fraud detection model.
- Grow Bank is a fictional company created for analytical purposes.
- The transactions in the dataset are assumed to be made using either a savings account or a credit card.

Methodology



Grow Bank

Data Preprocessing

1. Raw Dataset

- 50,000 Transactions Data
- 9 Categorical features
- 12 Numerical features

2. Data Cleaning

- No Missing Values
- No Duplicates
- Data Type Correction on timestamp feature: converted from obj to datetime
- Lowercase all Feature name (Timestamp -> timestamp, etc)
- Transform transaction_id to index for

3. Feature Engineering

add new feature:

1. transaction_min
2. transaction_hour
3. transaction_date
4. trasaction_month
5. transaction_auth_fraud_rate
6. risk_score_bin

Drop non-informative column (user_id, timestamp)

4. Encoding Categorical Features

Encoding categorical features (transaction_type, device_type, location, merchant_category, card_type, authentication_method)

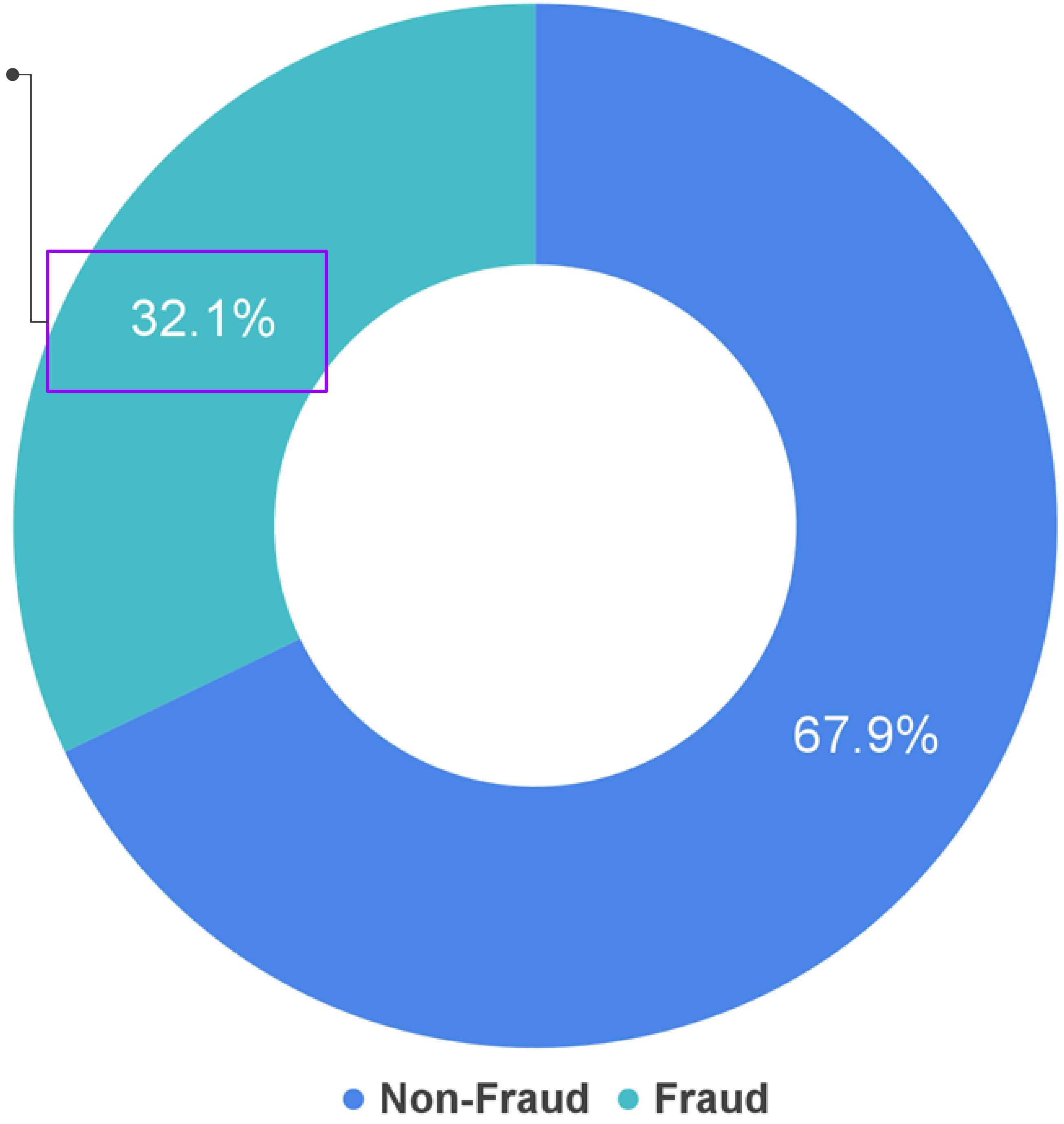
5. Data for Train & Test

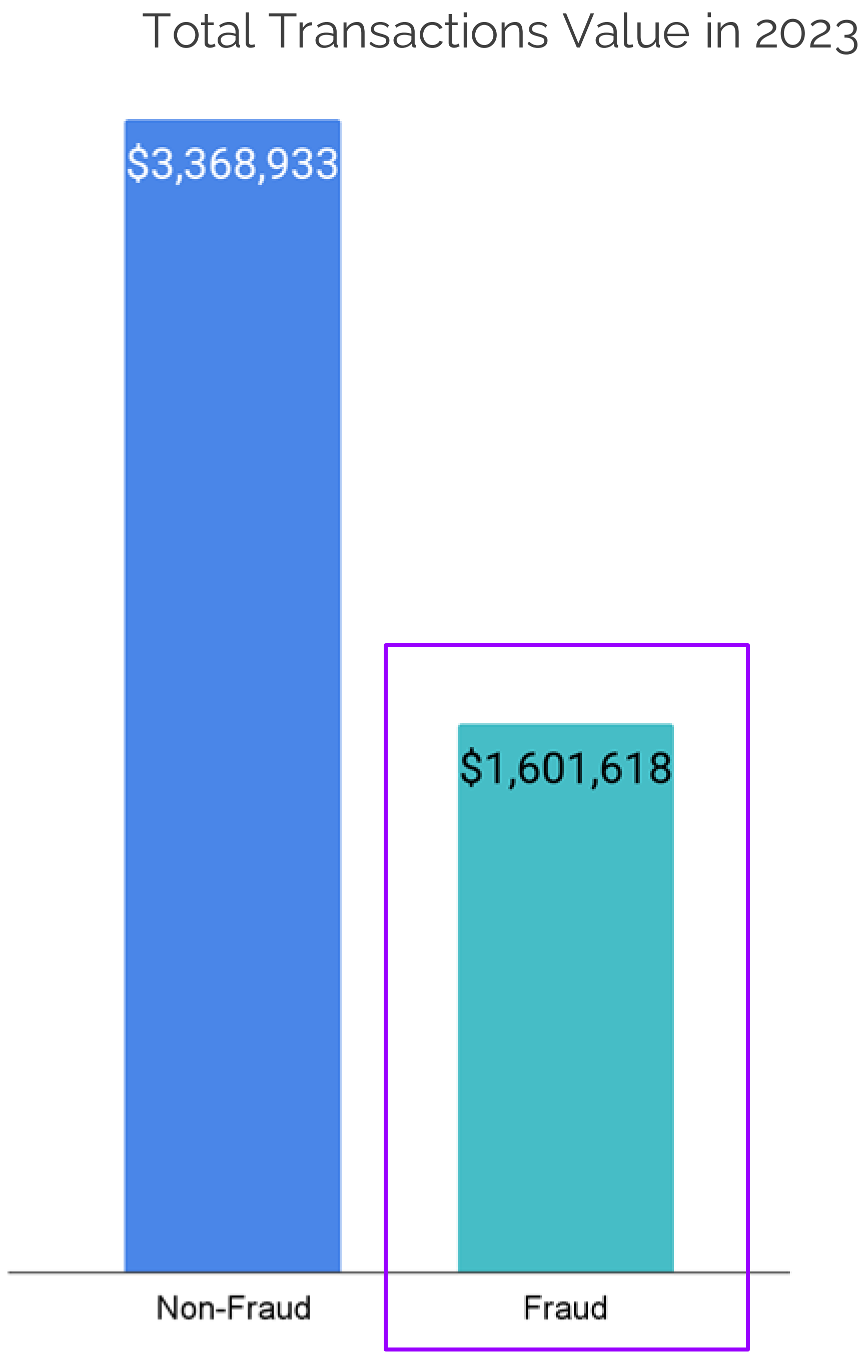
- 50,000 Transactions Data
- 36 Numerical features
- 1 target feature

fraud_label

High Proportion Of Fraud In 2023

1/3 transactions in 2023 is a fraud

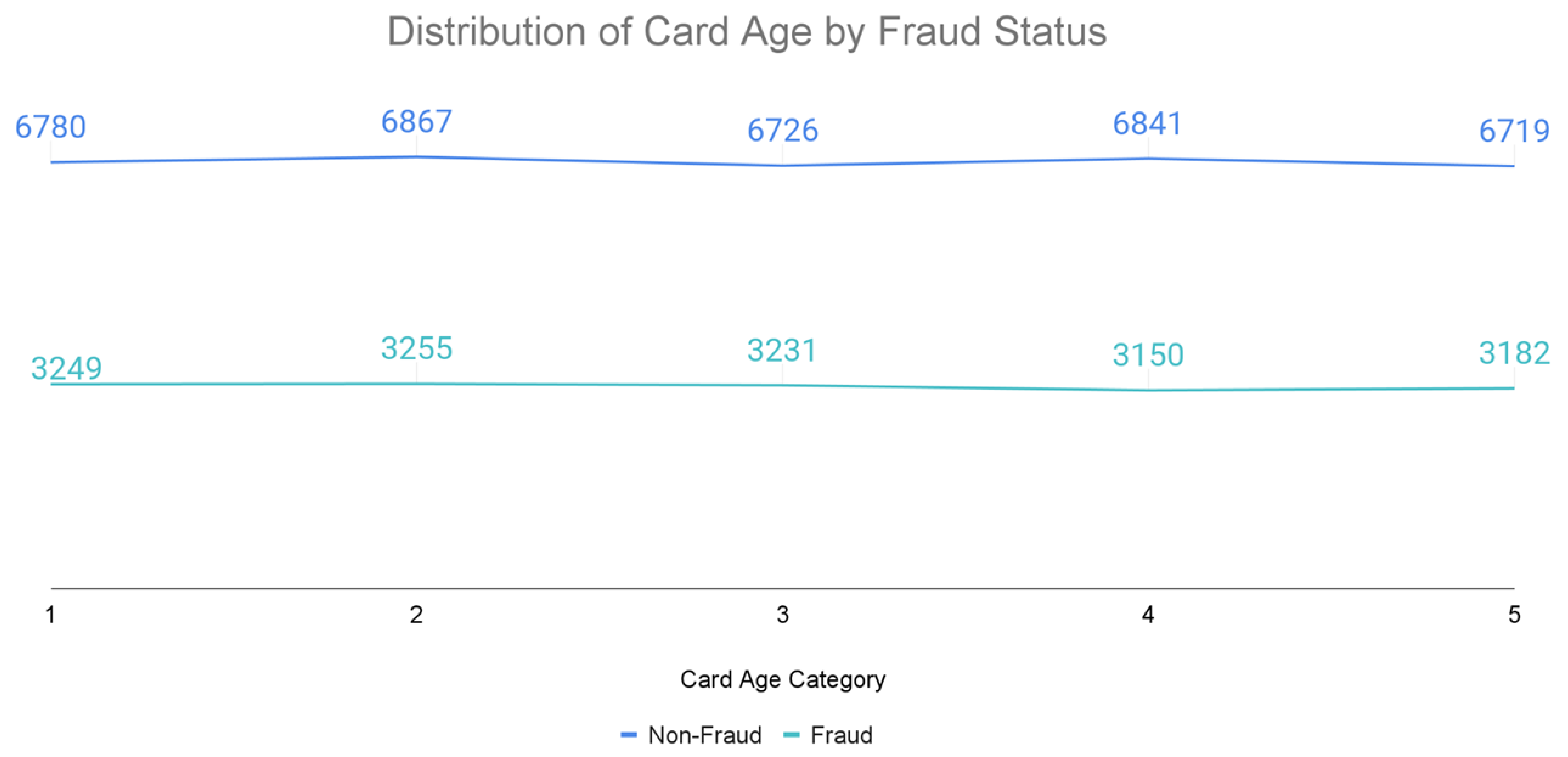




Grow Bank

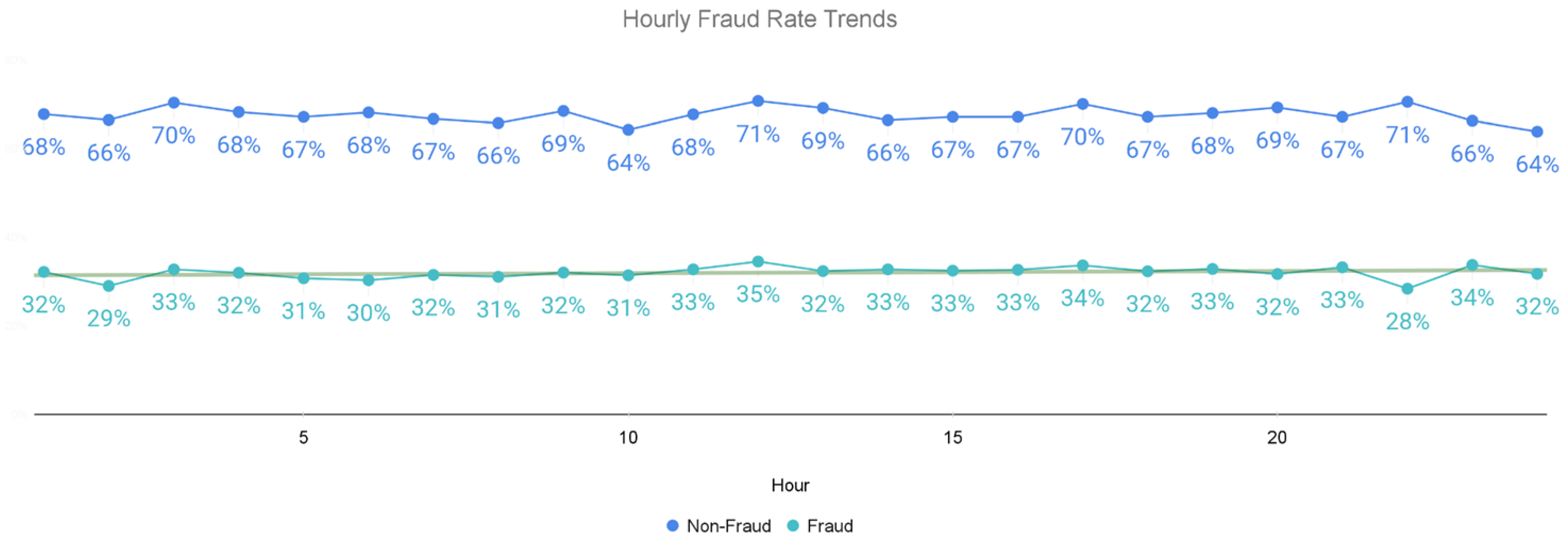
There Is No Significant Difference In The Number Of Fraudulent Transactions Across Card Age Categories

The age of a card doesn't appear to have a strong impact on whether a transaction associated with it is fraudulent or not.



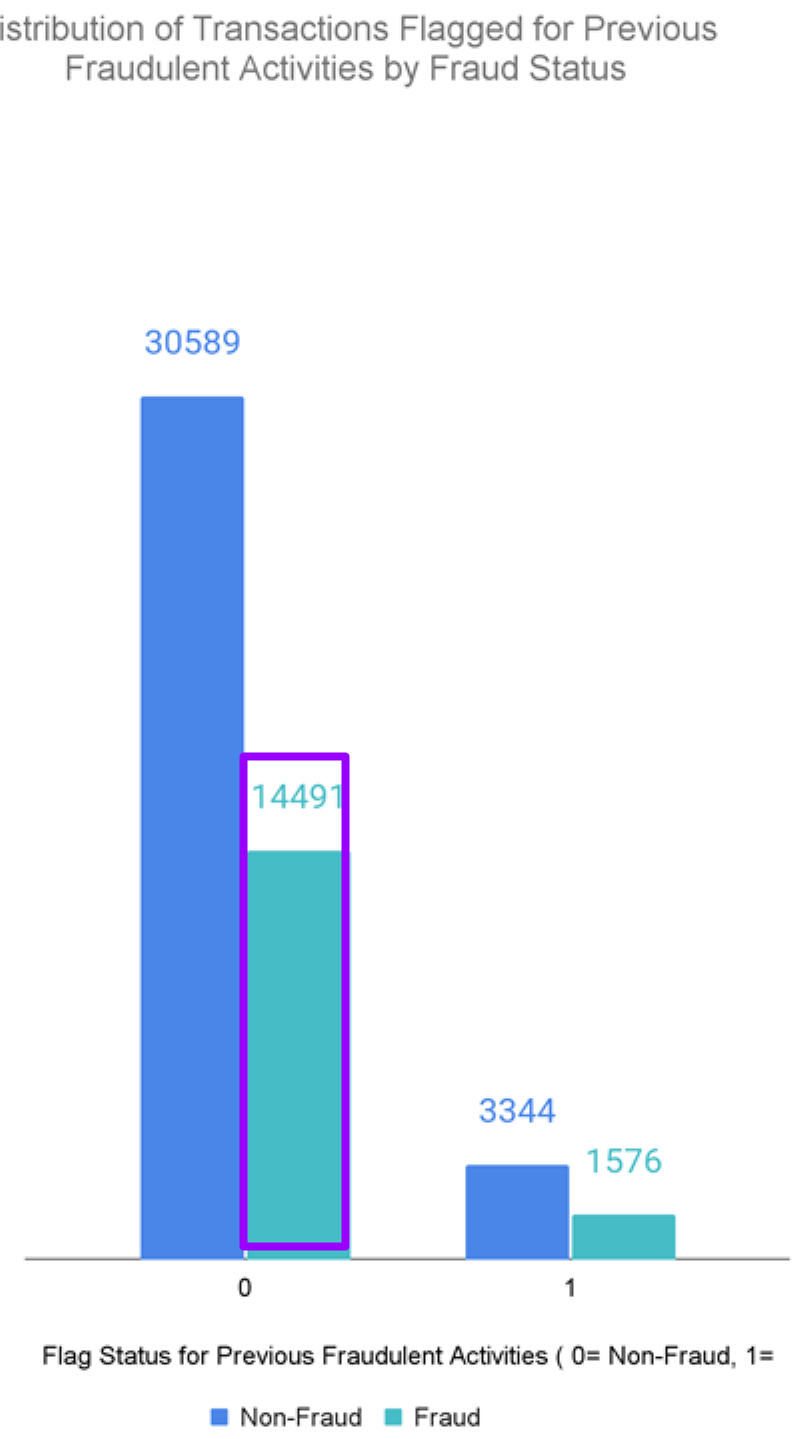
Fraud Occurs Consistently Across All Hours - No Specific Pattern Detected

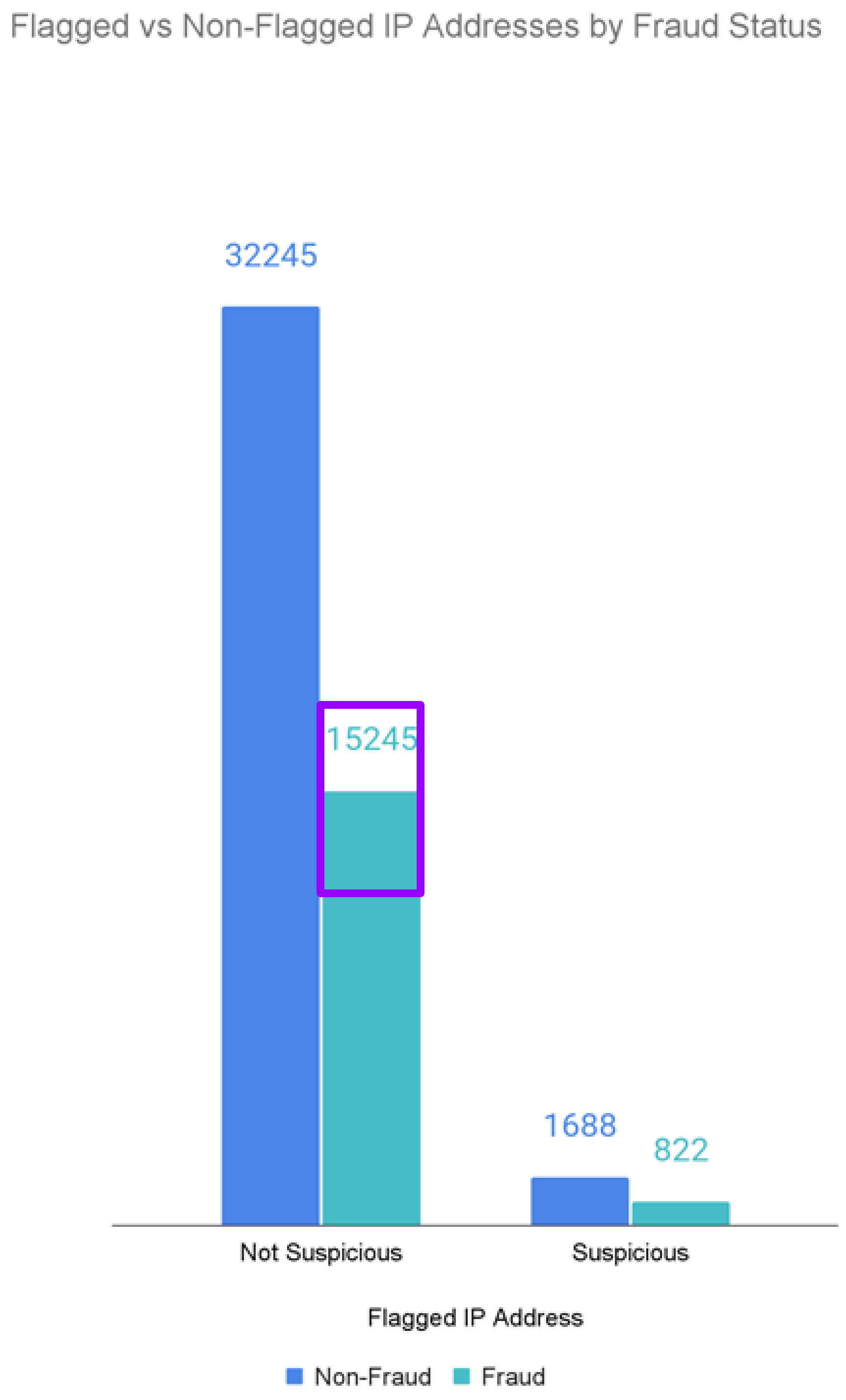
- The distribution of fraudulent transactions remains fairly stable throughout the hours, no sharp spikes indicating specific time-based fraud patterns.
- This suggests that fraud attempts occur regularly, regardless of the hours
- Reinforcing the need for continuous, 24/7 fraud monitoring rather than relying on time-based rules.



Known Risk Markers Are Helpful, but Most Fraud Still Happens Outside Them

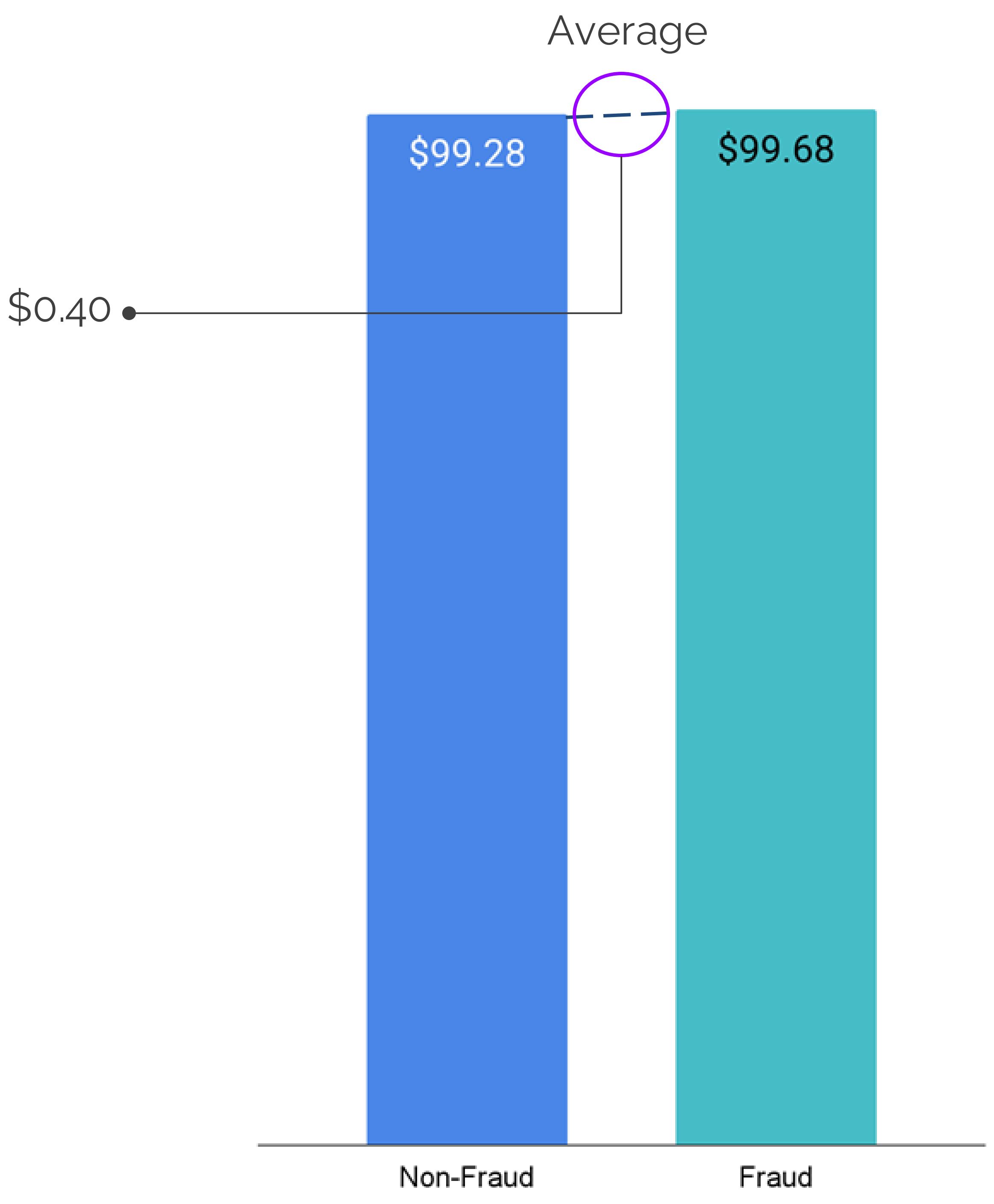
- While suspicious IP addresses and users with prior fraud flags are important indicators, a large number of fraudulent transactions still originate from IPs not previously marked as suspicious, and from users without prior fraud history.
- This shows that fraud is not always concentrated in known risk profiles — highlighting the need for adaptive, behavior-based detection strategies.

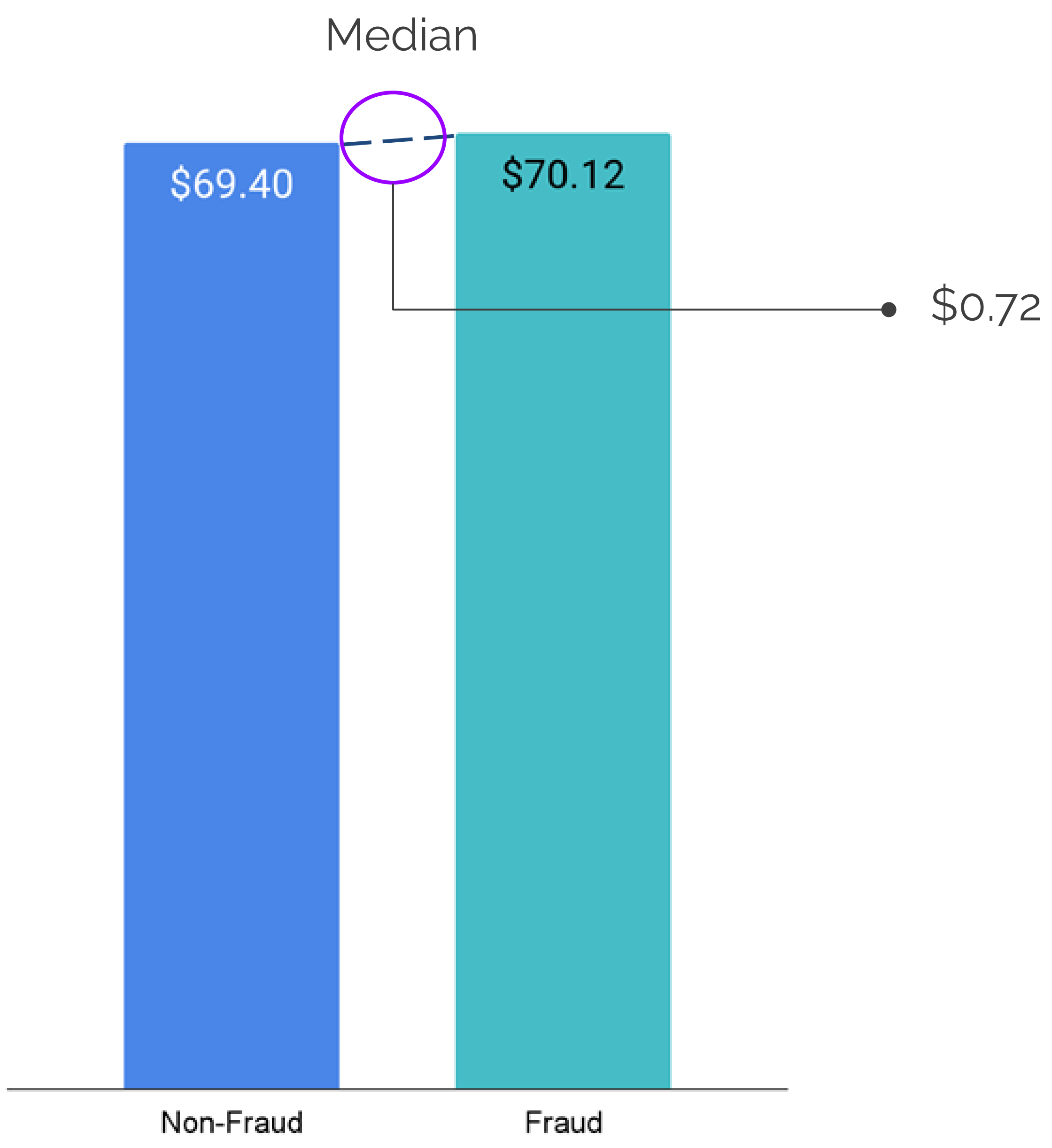




Fraud and Non-Fraud Have Similar Transaction Values

Fraud and non-fraud transactions show nearly identical values, both in average and median, suggesting that transaction amount alone is not a strong fraud indicator.



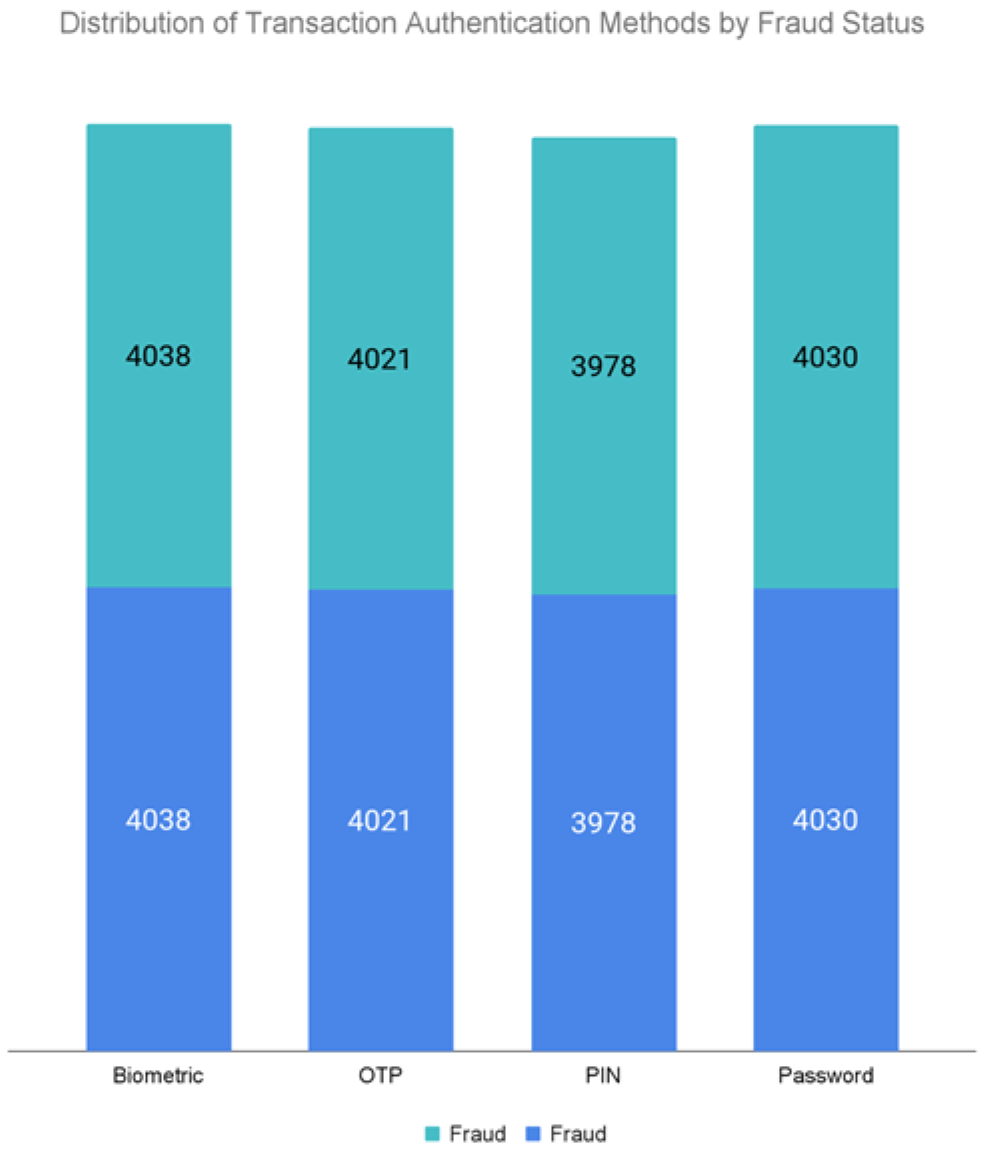


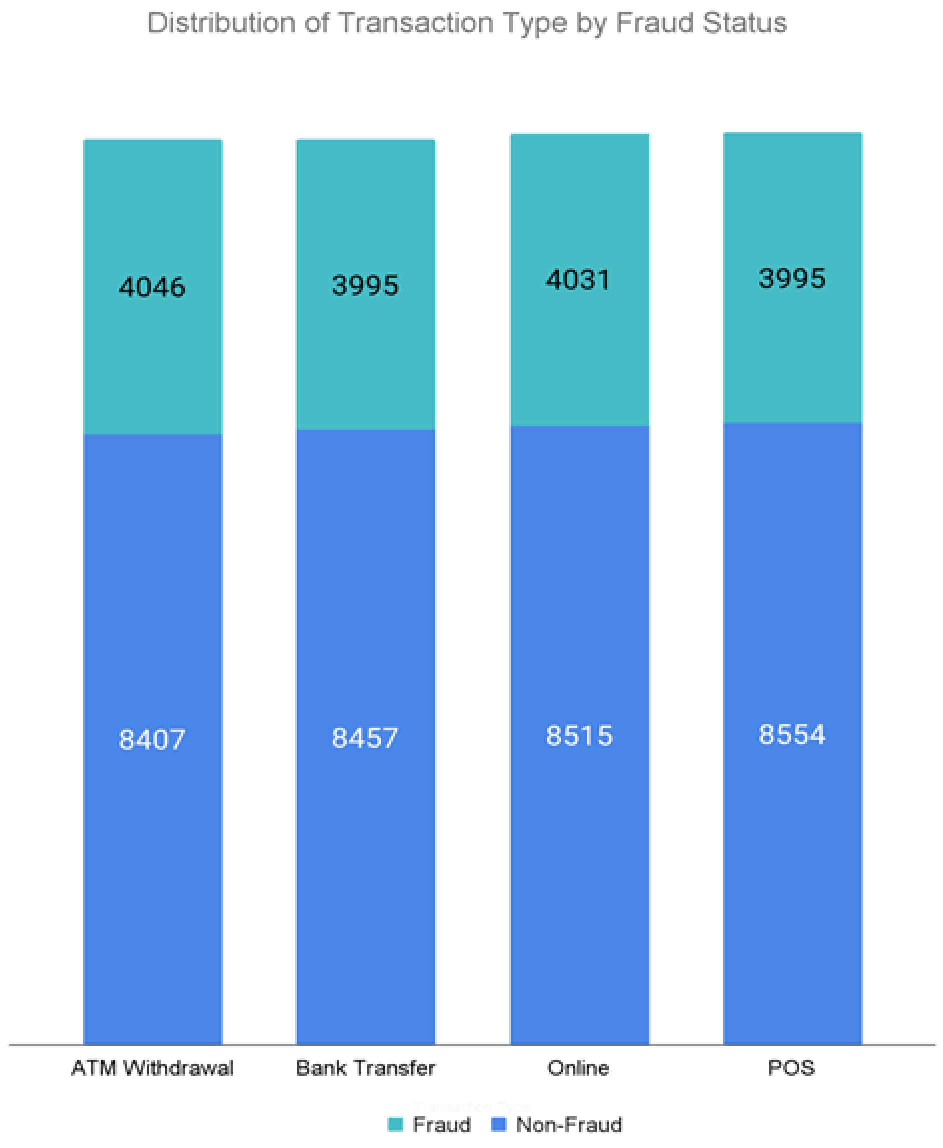
Distance Doesn’t Differentiate Fraud

- No meaningful difference in average distance between fraud (2,499) and non-fraud (2,499).
- Median values are nearly identical as well (2,488 vs 2,497).
- This suggests that distance is not a discriminative feature for fraud detection in this context.

Fraud Happens Across Channels And No Authentication Method is Immune

- All transaction authentication methods and transaction types have relatively similar distributions of fraud and non-fraud cases.
- No specific method or type stands out as significantly riskier than the others.
- This suggests that fraud is not concentrated in one particular channel or authentication method,
- Highlighting the need for more advanced behavioral features to improve fraud detection.

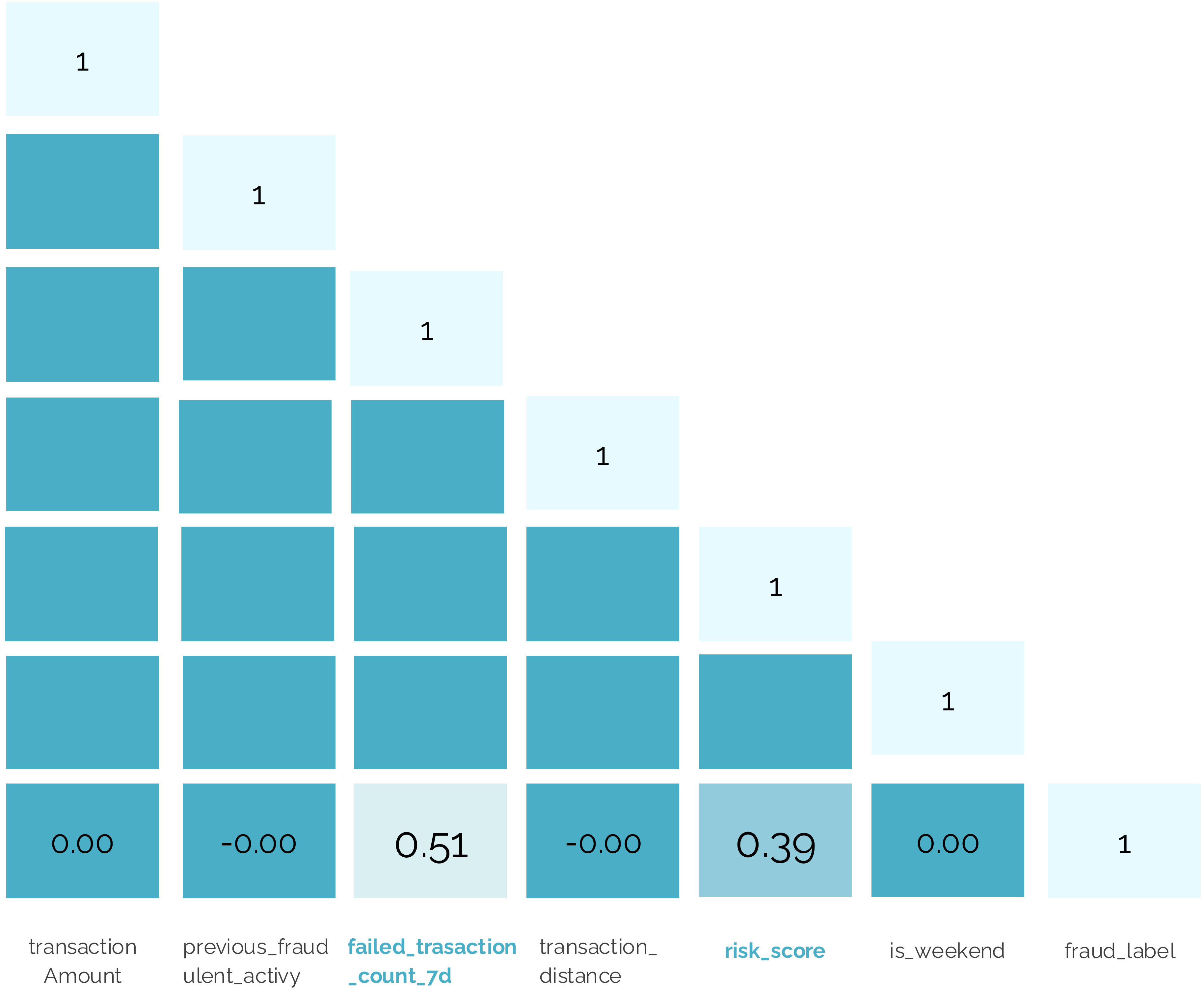




Failed Attempts & Risk Score Show Early Fraud Signals

Grow Bank

Before Features Engineering And Encoding



- failed_transaction_count_7d has a strong positive correlation (0.51) with fraud label.
  →Users with more failed attempts in the past 7 days are more likely to be involved in fraud.
- risk_score shows a moderate positive correlation (0.39) with fraud.
  → If the risk score increases, the likelihood of fraud also rises—making it a key signal for early detection.
- transaction_distance, amount, and is_weekend show near-zero correlation, indicating they are less useful in raw form.
- This supports early hypothesis: behavioral signals (like failures and risk score) are stronger fraud indicators than static attributes.

Adaptive Threshold Implementation For Risk Score Bin

Risk_Score	*fraud*
Mean	0.66
Standard Error	0.00
Median	0.81
Mode	0.86
Standard Deviation	0.31
Sample Variance	0.09
Kurtosis	-0.95
Skewness	-0.69
Range	1.00
Minimum	0.00
Maximum	1.00
Sum	10650.88
Count	16067.00
Largest(1)	1.00
Smallest(1)	0.00
Confidence Level(95.0%)	0.00

Risk_Score	*non fraud*
Mean	0.43
Standard Error	0.00
Median	0.43
Mode	0.63
Standard Deviation	0.24
Sample Variance	0.06
Kurtosis	-1.19
Skewness	0.00
Range	0.85
Minimum	0.00
Maximum	0.85
Sum	14426.90
Count	33933.00
Largest(1)	0.85
Smallest(1)	0.00
Confidence Level(95.0%)	0.00

Risk bins are defined based on actual distribution:

- Low threshold (< 0.4) is derived from the non-fraud median (0.43), indicating dominance of legitimate transactions.
- High threshold (> 0.8) comes from the fraud median (0.81), highlighting where fraud is most concentrated.

This strategy ensures that the thresholds are data-driven, not assumptions, and reflect real behavior patterns in historical data.

Risk Threshold	Score Range	Rationale
Low (0–0.4)	Score < 0.4	Dominated by non-fraud (low median & mean)
Medium (0.4–0.8)	0.4 – 0.8	Gray zone – overlapping fraud and non-fraud
High (0.8–1)	Score > 0.8	Majority of fraud transactions (high median & mode)

Feature Simulation: Transaction_auth_fraud_rate Calculation And Application

Grow Bank

This engineered feature allows the model to understand historical fraud likelihood based on specific transaction type and authentication patterns — a strong behavioral fraud signal.

1. **Dataset Overview**

transaction_id	user_id	transaction_type	authentication_method	fraud_label
T001	U001	Online	Biometric	0
T002	U001	Online	Biometric	1
T003	U002	POS	PIN	0
T004	U002	POS	Biometric	1
T005	U003	Bank Transfer	Password	0
T006	U004	POS	Biometric	0
T007	U005	ATM Withdrawal	OTP	0

We start with a sample dataset containing:

- Transaction ID
- User ID
- Transaction type (e.g., Online, POS, ATM)
- Authentication method (e.g., Biometric, PIN, OTP)
- Fraud label (0 = not fraud, 1 = fraud)

2. **Calculate Fraud Rate by Combining Transaction Type and Authentication Method**

transaction_type	authentication_method	Fraud Count	Total Count fraud per combination	transaction_auth_fraud_rate
Online	Biometric	1	2	0.5 (50%)
POS	PIN	0	1	0%
POS	Biometric	1	2	0.5 (50%)
Bank Transfer	Password	0	1	0%
ATM Withdrawal	OTP	0	1	0%

We group the data by transaction_type and authentication_method, then calculate:

- Total number of transactions for each combination
- Total fraud cases in each group
- Resulting fraud rate per combination
  (e.g., for Online + Biometric → 1 fraud out of 2 = 50% fraud rate)

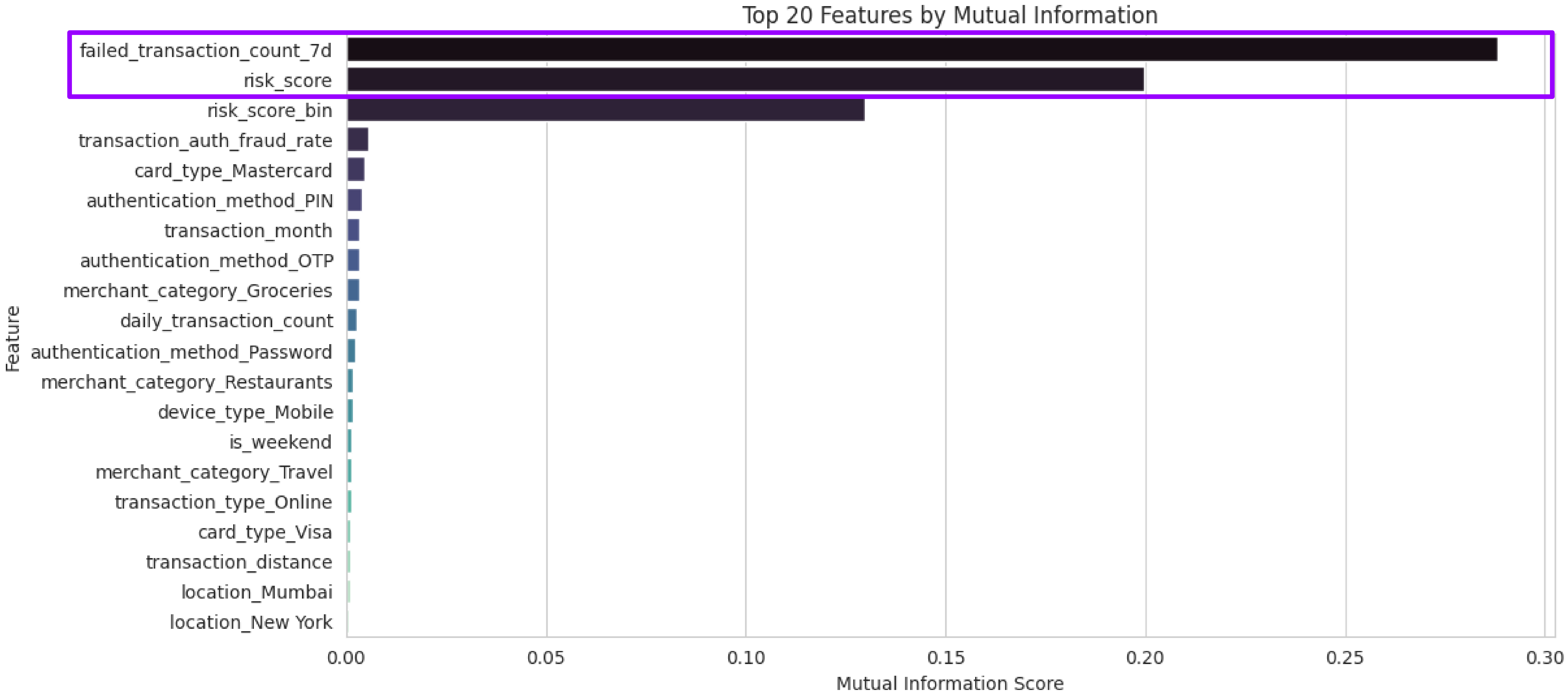
3. **Merge our transaction_auth_fraud_rate combination table with our data set**

transaction_id	user_id	transaction_type	authentication_method	fraud_label	transaction_auth_fraud_rate
T001	U001	Online	Biometric	0	0.5
T002	U001	Online	Biometric	1	0.5
T003	U002	POS	PIN	0	0.0
T004	U002	POS	Biometric	1	0.5
T005	U003	Bank Transfer	Password	0	0.0
T006	U004	POS	Biometric	0	0.5
T007	U005	ATM Withdrawal	OTP	0	0.0

We merge the calculated fraud rate (transaction_auth_fraud_rate) back into the original dataset.
Each transaction now carries an inherited fraud risk score based on its transaction_type and authentication_method combination.

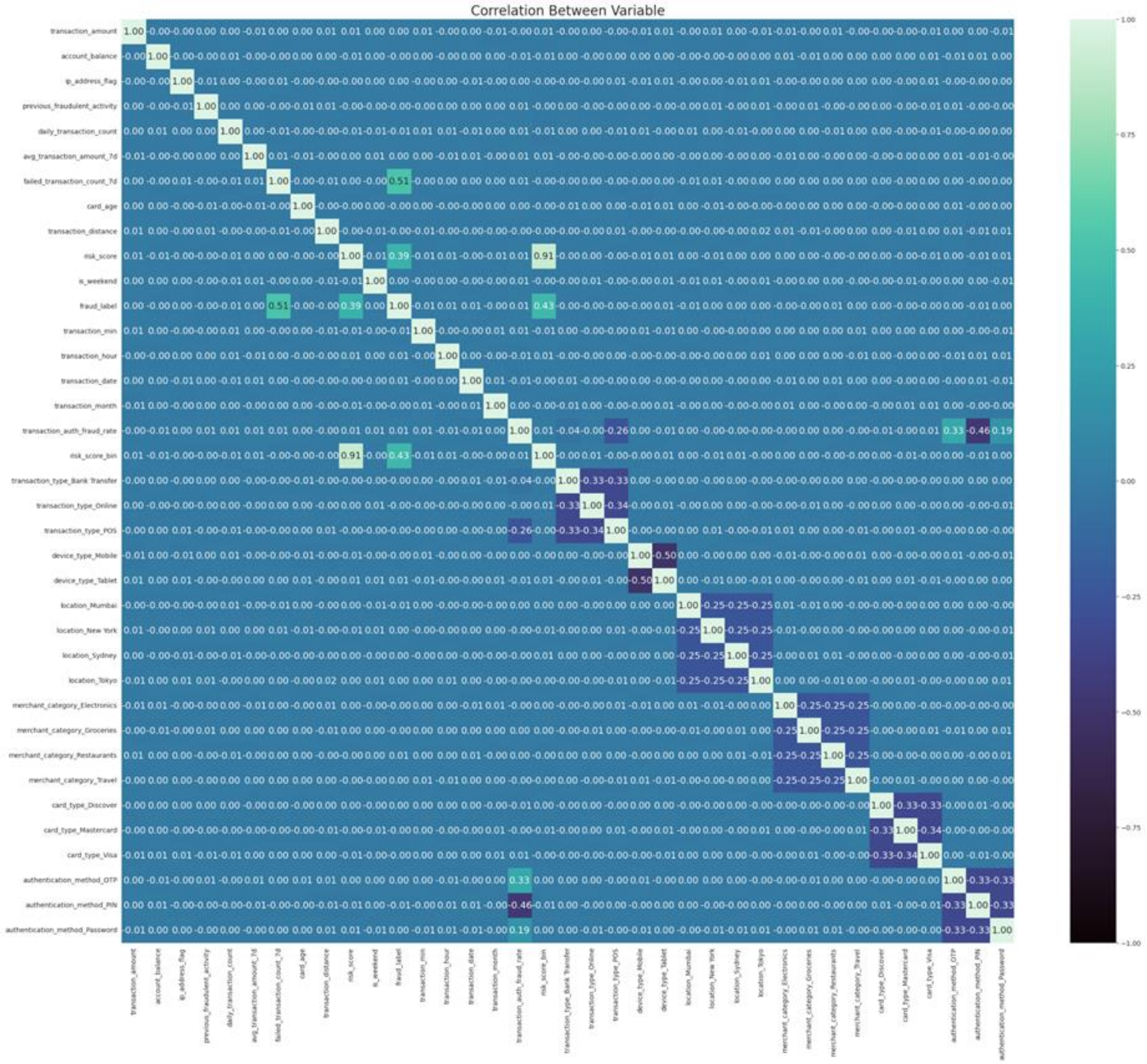
Top 20 Key Features Influencing Fraud Detection

Failed Transaction Count (7 Days) and Risk Score are the most dominant feature, indicating strong information contribution that may enhance model accuracy but raising a concern about potential overfitting due to heavy reliance on these features.



Correlation Across Variables On Transaction Data

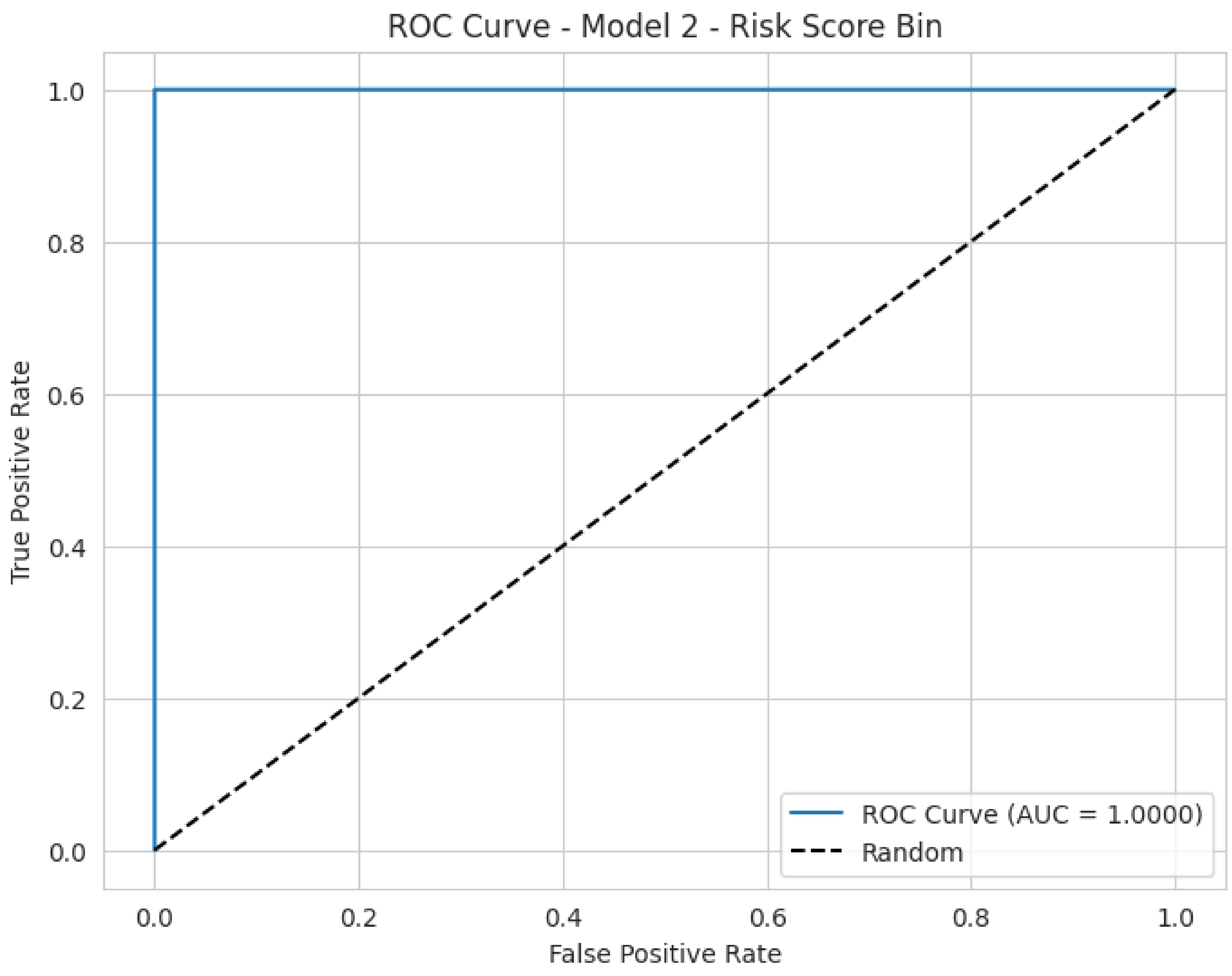
After Features Engineering And Encoding



- Initial correlation analysis reveals that failed_transaction_count_7d (+0.51) and risk_score (+0.39) show strong - moderate positive correlations with fraud_label.
- risk_score_bin, a binned version of risk_score, also shows a moderate correlation (+0.43) with fraud_label but is highly correlated with risk_score itself (+0.91).
- To avoid multicollinearity, two separate models will be developed — one using risk_score, and one using risk_score_bin.

ROC Curve for Model 2

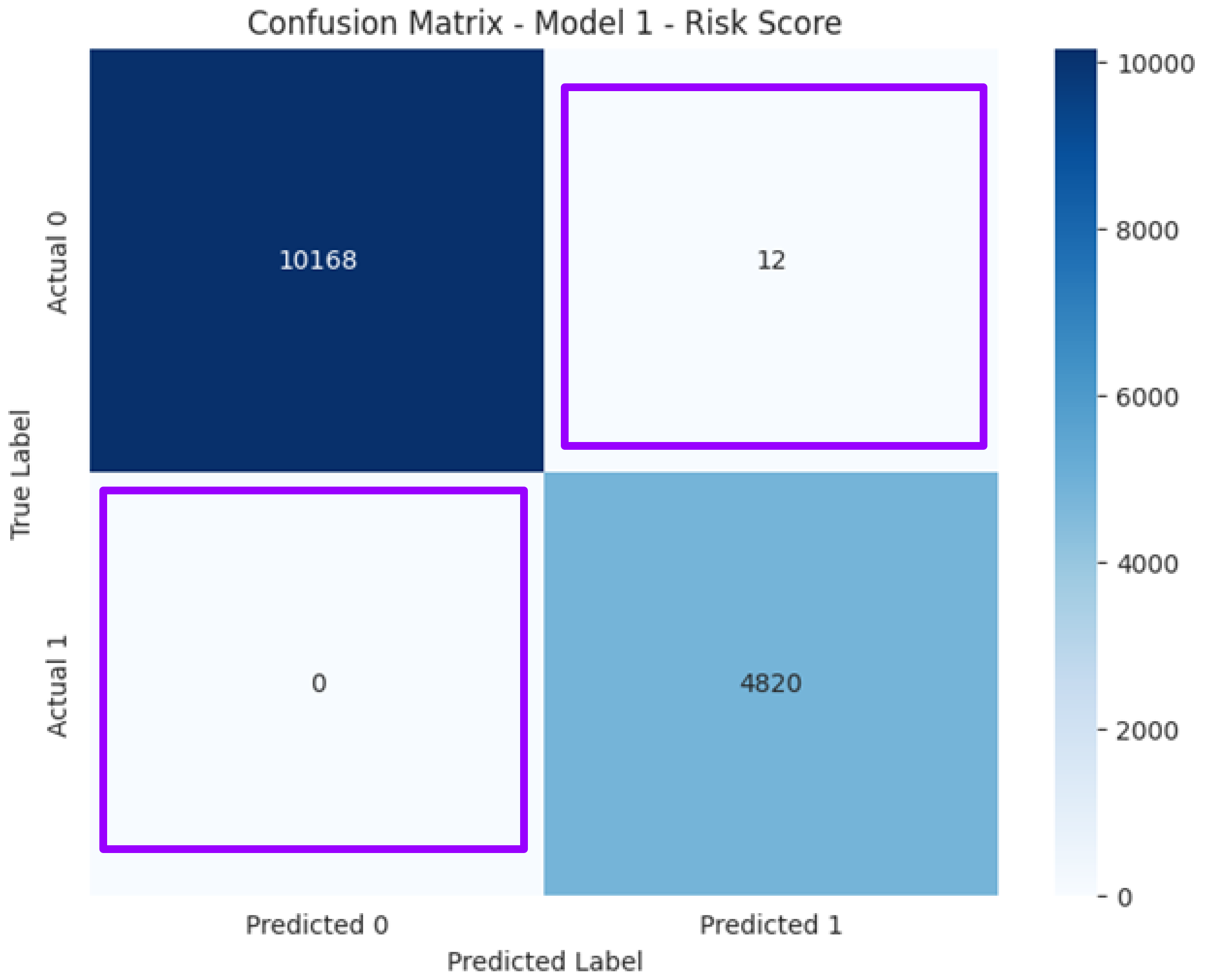
Model 2 shows perfect fraud classification (AUC = 1.00) — indicating ideal performance on test data, but also a risk of overfitting. Threshold monitoring is recommended to ensure long-term reliability.

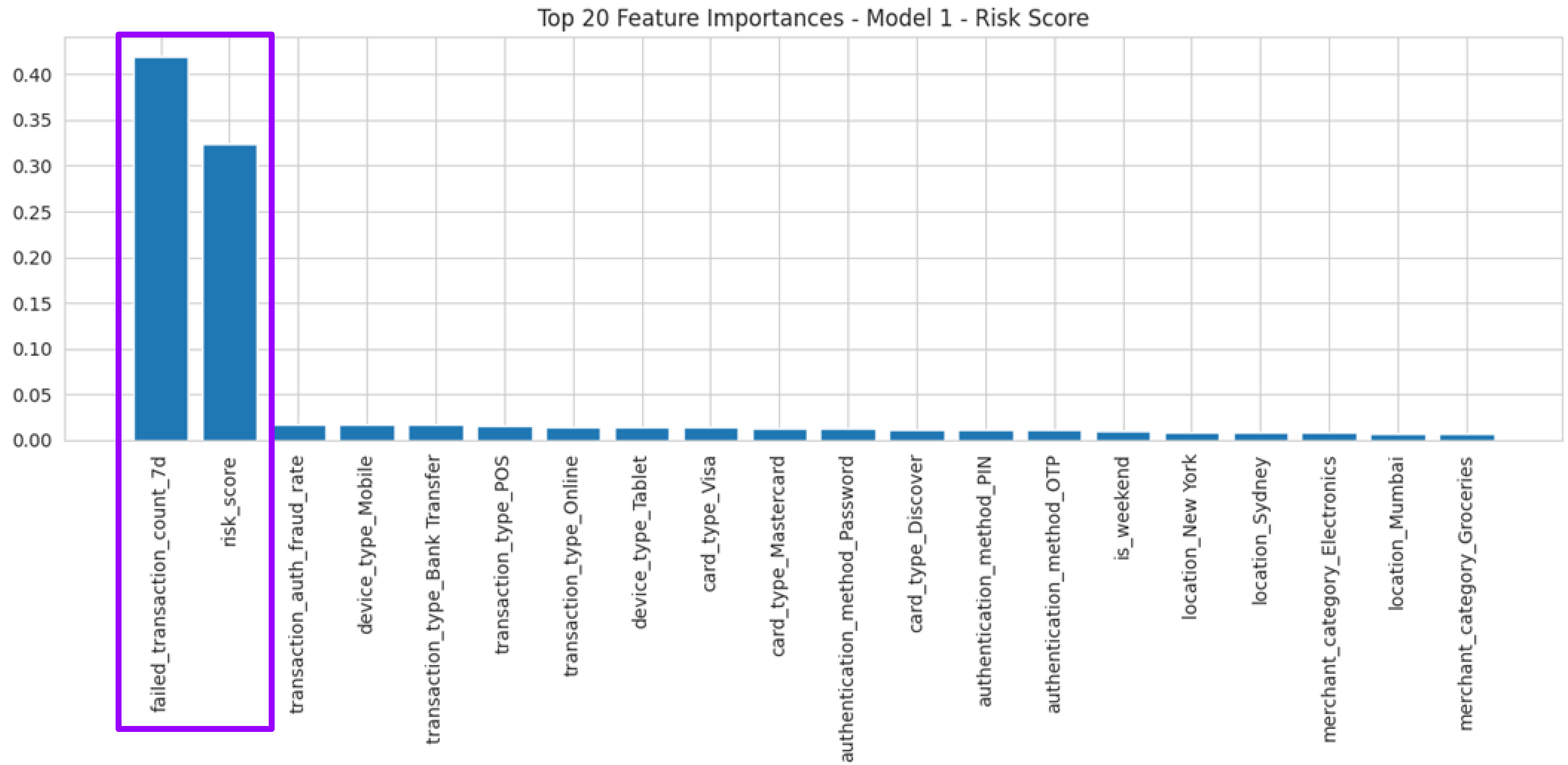


Model 1 Evaluation Result

Zero false negatives, 12 false positives. As we suspected, there are two highly dominant features: failed transaction count (7 days) and risk score, The perfect scores across all evaluation metrics may indicate that the model is overfitting.

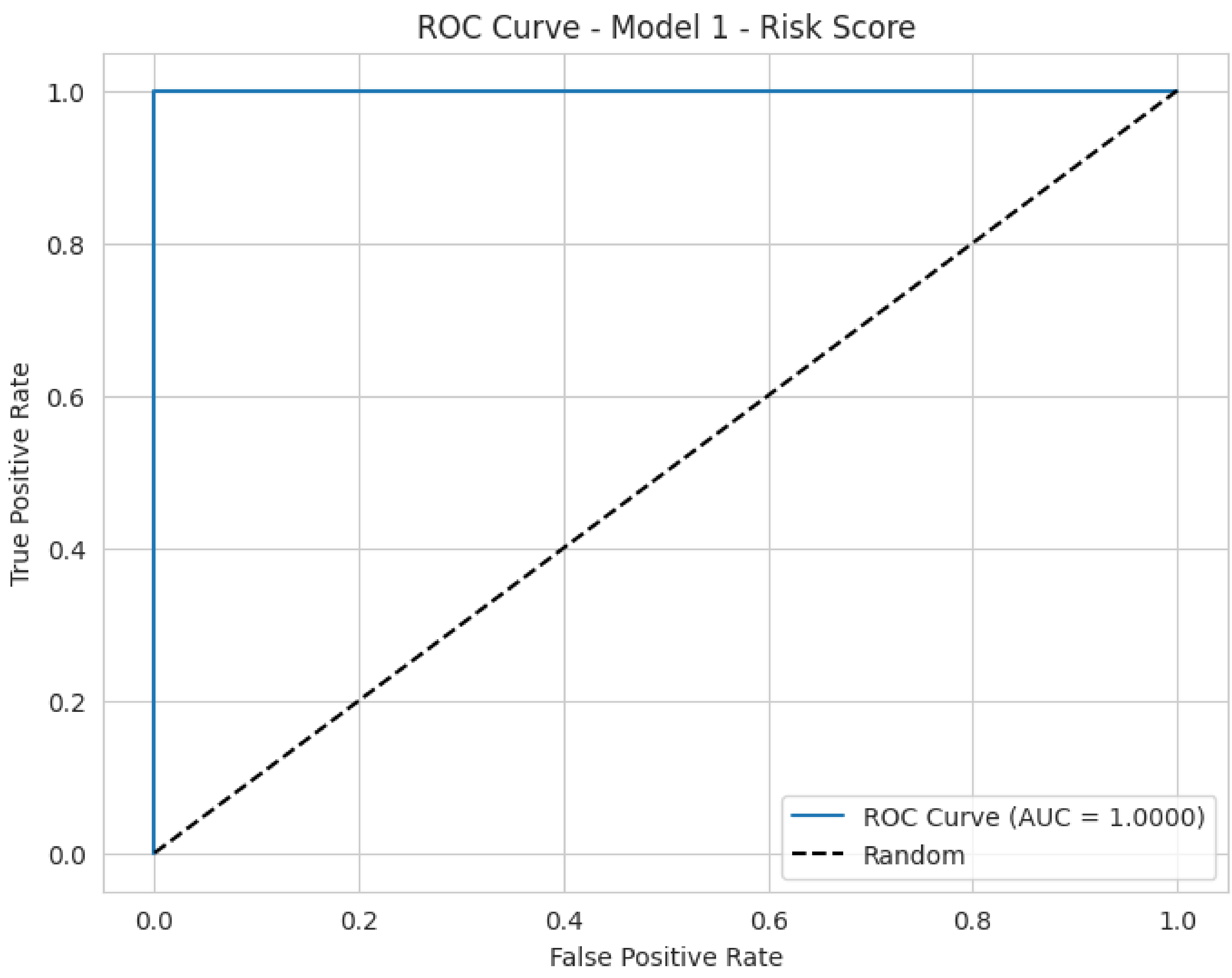
```
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     10180
           1       1.00      1.00      1.00      4820

    accuracy                           1.00     15000
   macro avg       1.00      1.00      1.00     15000
weighted avg       1.00      1.00      1.00     15000

ROC AUC Score: 1.0000
```



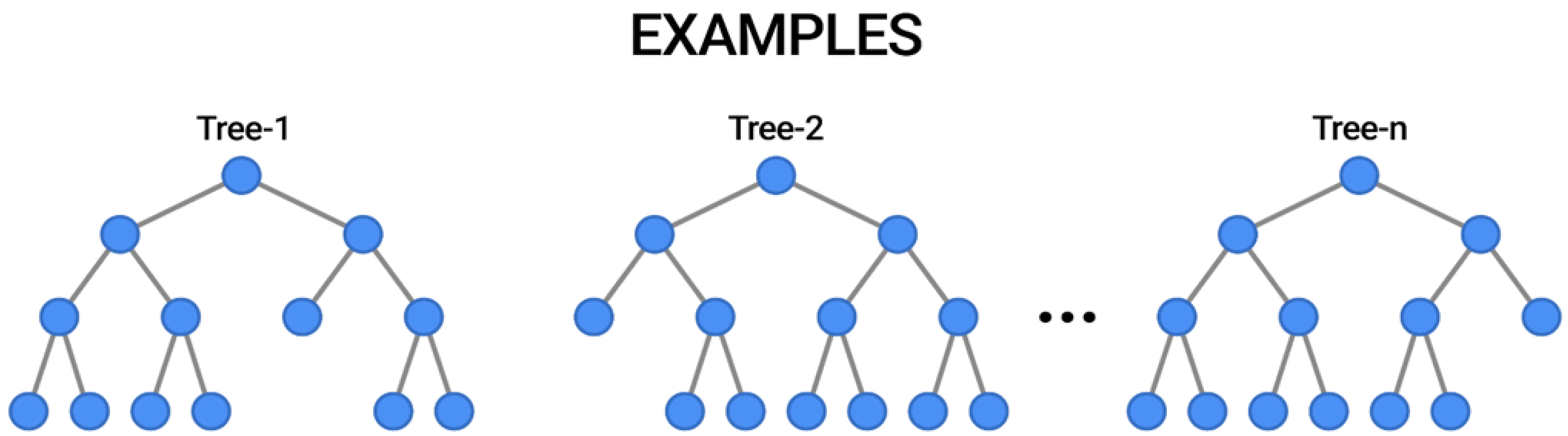


ROC Curve for Model 1

Model 1 shows perfect fraud classification (AUC = 1.00) — indicating ideal performance on test data, but also a risk of overfitting.



Why Use Random Forest Algorithm?



1. Handles Overfitting

Uses majority voting across trees to lower the risk of overfitting.

2. Feature Importance

Provides insights into feature significance, helping identify which predictors (like failed_transaction_count_7d and risk_score) are most influential.

3. High Accuracy

Generally offers strong performance for classification tasks, including fraud detection, due to its ensemble nature.

4. Resilience to Noise

Performs well even with missing values and noisy data, which is common in real-world datasets.