UPI Fraud Detection & Investigation Prioritization System

Improving investigation efficiency through risk-based alert prioritization

---

1. Executive Summary

Fraud monitoring teams receive a large number of alerts every day, but investigators cannot manually review every transaction.
The real operational challenge is not only detecting fraud — it is deciding **which alerts should be reviewed first**.

This project builds a fraud monitoring workflow that combines rule-based detection and machine learning to improve investigation efficiency.

The system identifies suspicious transactions, ranks them by risk, and produces readable investigation notes for each case.

Key Findings

- Rule-based signals effectively capture obvious fraud patterns
- Machine learning improves prioritization rather than overall detection
- Reviewing only top alerts significantly increases fraud concentration
- A structured investigation queue reduces manual effort

The project demonstrates how analytics supports operational decision-making rather than replacing human investigators.

---

2. Business Problem

Fraud teams face two conflicting constraints:

1. Reviewing all transactions is impossible
2. Missing fraud leads to financial loss

Most systems generate alerts but do not help investigators decide **what to review first**.

This leads to:

- wasted investigation time
- delayed fraud response
- inconsistent decision quality

The business needed a solution that answers:

1. Which transactions are suspicious?
2. Which suspicious transactions are most important?
3. How can investigators understand alerts quickly?

---

3. Data & Approach

Data Structure

The project simulated a production-style environment using structured datasets:

Table	Purpose
transactions	payment activity
users	customer behavior profile
behavior_signals	risk indicators

Data extraction was performed using SQL and analysis using Python.

Methodology

The project followed an operational workflow instead of a pure modeling approach.

Step 1 — Risk Signal Identification

Behavioral indicators were created:

- unusually high transaction amount
- deviation from customer's normal activity
- late-night transaction behavior

These represent real fraud monitoring signals.

Step 2 — Rule-Based Risk Scoring

Transactions were scored using weighted rules:

- behavioral anomaly (strongest signal)
- unusual amount
- suspicious timing

This concentrates obvious fraud into a manageable set of alerts.

Step 3 — Machine Learning Detection

A logistic regression model was trained to detect subtle patterns not captured by rules.

Because fraud is rare, class imbalance handling was applied to ensure meaningful detection.

Step 4 — Hybrid Scoring

Rules identify suspicious transactions
Machine learning ranks which to review first

The combined score produces a prioritized investigation queue.

Step 5 — Explainable Output

Readable investigation notes were generated describing why a transaction was flagged.

This reduces analyst interpretation time.

---

4. Key Insights

Rules vs Machine Learning

Rules effectively detect clear fraud behavior.
Machine learning improves prioritization rather than replacing rules.

Investigation Efficiency

Fraud concentration increases significantly when reviewing top-ranked alerts.
This means investigators can focus effort where risk is highest.

Operational Value

The system supports decision-making rather than fully automating detection.

---

5. Recommendations

Priority-Based Review

Investigators should review alerts in risk order rather than sequentially.

Hybrid Detection Strategy

Maintain rule detection while using machine learning for ranking.

Explainable Alerts

Provide readable explanations for faster investigation decisions.

---

6. Business Impact

The proposed workflow improves:

- investigation efficiency
- response speed
- decision consistency
- operational productivity

Instead of trying to detect every fraud automatically, the system optimizes human effort.

7. Conclusion

Fraud detection is not only a prediction problem — it is a prioritization problem.

This project shows that combining rules and machine learning enables better operational outcomes than relying on either alone.

By ranking alerts and explaining risk factors, organizations can move from reactive alert handling to **structured investigation management**.