Alipay Fraud Detection Model - Monitoring Framework

# 1. Background and Objective

The Alipay Fraud Detection Model is a rule-based system deployed by the bank to assess and mitigate fraud risks associated with global remittance transactions processed through World Asia Pvt. Ltd. on the Alipay platform. The model evaluates each transaction in real time using a predefined set of fraud detection rules. If a transaction satisfies all rules, it is approved for processing; otherwise, it is flagged as suspicious and routed for manual review. This real-time decisioning framework enables the bank to maintain secure, compliant, and efficient processing of global remittances.  
  
The primary objective of ongoing monitoring is to ensure that the Alipay Fraud Detection Model operates effectively, remains stable over time, and complies with the bank’s risk governance standards. Monitoring will cover key areas including the model’s ability to detect suspicious transactions accurately, the operational efficiency of real-time decisioning, the integrity and consistency of input data, and the performance and relevance of individual fraud detection rules in adapting to evolving transaction behaviors.

## Governance Details

| Parameter | Value |
| --- | --- |
| Model Owner | Fraud Risk Analytics Team, Bank XYZ |
| Validation Owner | Model Risk Management Team, Bank XYZ |
| Model Tier | Tier 2 (Medium Risk, Real-time System) |
| Monitoring Frequency | Monthly (Automated Reports); Quarterly (In-depth Review) |

# 2. Monitoring Metrics

These metrics are derived solely from the model’s output in the Alipay Payload, which provides a binary decision: Approved or Flagged.

| Metric Name | Formula / Computation | Business Explanation |
| --- | --- | --- |
| Flag Rate | (Flagged Transactions / Total Transactions) × 100 | Indicates the proportion of transactions being flagged. Spikes or drops may signal issues in rule sensitivity. |
| Approval Rate | (Approved Transactions / Total Transactions) × 100 | Complements flag rate; helps assess throughput and model pass rate. |
| Daily/Hourly Volume | Count of Transactions per Day or Hour | Monitors traffic trends and potential anomalies or outages in transaction flow. |
| Flag Volume Trend | Count of Flagged Transactions over Time | Helps track how suspicious activity or model flagging behavior changes over time. |
| Flag Distribution by Channel/Source | (Flagged in Channel / Total in Channel) × 100 | Identifies anomalies or risk concentrations across specific transaction sources or geographic regions. |
| Escalation Rate\* | (Sent to Manual Review / Flagged Transactions) × 100 | If CRM-linked: Assesses reliance on manual review, indicating how confident the model is in its rule-based decisions. |

# 3. Data Sources

| Source Name | Description | Usage in Monitoring |
| --- | --- | --- |
| Alipay Payload | Real-time transaction payload containing model decision (Approved/Flagged) | Primary source for computing operational and output metrics. |
| CRM for Alipay | Internal system containing customer profile and manual review outcomes | If integrated, supports escalation metrics and future performance validation with feedback. |

# 4. Governance Framework

Defined thresholds are used to proactively identify anomalies and trigger escalation protocols based on severity.

| Metric | Thresholds | Color Band | Action / Escalation Path |
| --- | --- | --- | --- |
| Flag Rate | < 5% or > 15% of total transactions | Red | Immediate review of rule settings and model input changes |
| Flag Rate | 5% – 8% or 12% – 15% | Amber | Monitor closely, validate with recent transaction patterns |
| Flag Rate | 8% – 12% | Green | No action required |
| Approval Rate | < 85% or > 95% | Red | Check for data anomalies or rule misconfigurations |
| Transaction Volume | ±25% deviation from trailing 7-day avg | Red | Investigate for system outages or external volume shifts |
| Flag Distribution | > 30% flag rate in any single channel | Red | Investigate for channel-specific risk or rule misalignment |
| Escalation Rate\* | > 60% of flagged cases escalated | Red | If available, suggests over-flagging or low rule precision |

# 5. Sample Dashboards and Visualizations

These dashboards support real-time and historical monitoring of model activity and behavior:

## 5.1 Output Distribution Dashboard

- Flag vs. Approved Transactions Over Time – Line chart comparing monthly flag/approval trends.

- Flag Rate by Channel/Region – Bar chart segmented by transaction source or geography.

## 5.2 Volume and Load Monitoring

- Daily Transaction Volume – Line or bar chart showing total transaction counts per day/hour.

- Flag Volume Trend – Line chart showing flagged transactions over recent periods.

## 5.3 Operational Alerts

- Threshold Breach Alerts – Conditional color-coded indicators for flag rate, approval rate, and transaction volume anomalies.