

Cost-Aware AI Decision System -

Executive Summary

The Problem -

- **Global fraud losses exceed \$32 billion annually.** Most fraud detection systems optimize for prediction accuracy, not business outcomes.
- **Real-world scenario:** A bank receives 10,000 fraud alerts daily but has only 50 investigators. Traditional systems prioritize by risk score alone—missing high-value cases while wasting resources on low-value ones.

Our Solution -

A **cost-aware AI decision system** that optimizes for **expected financial savings**, not just fraud probability.

Core Formula:

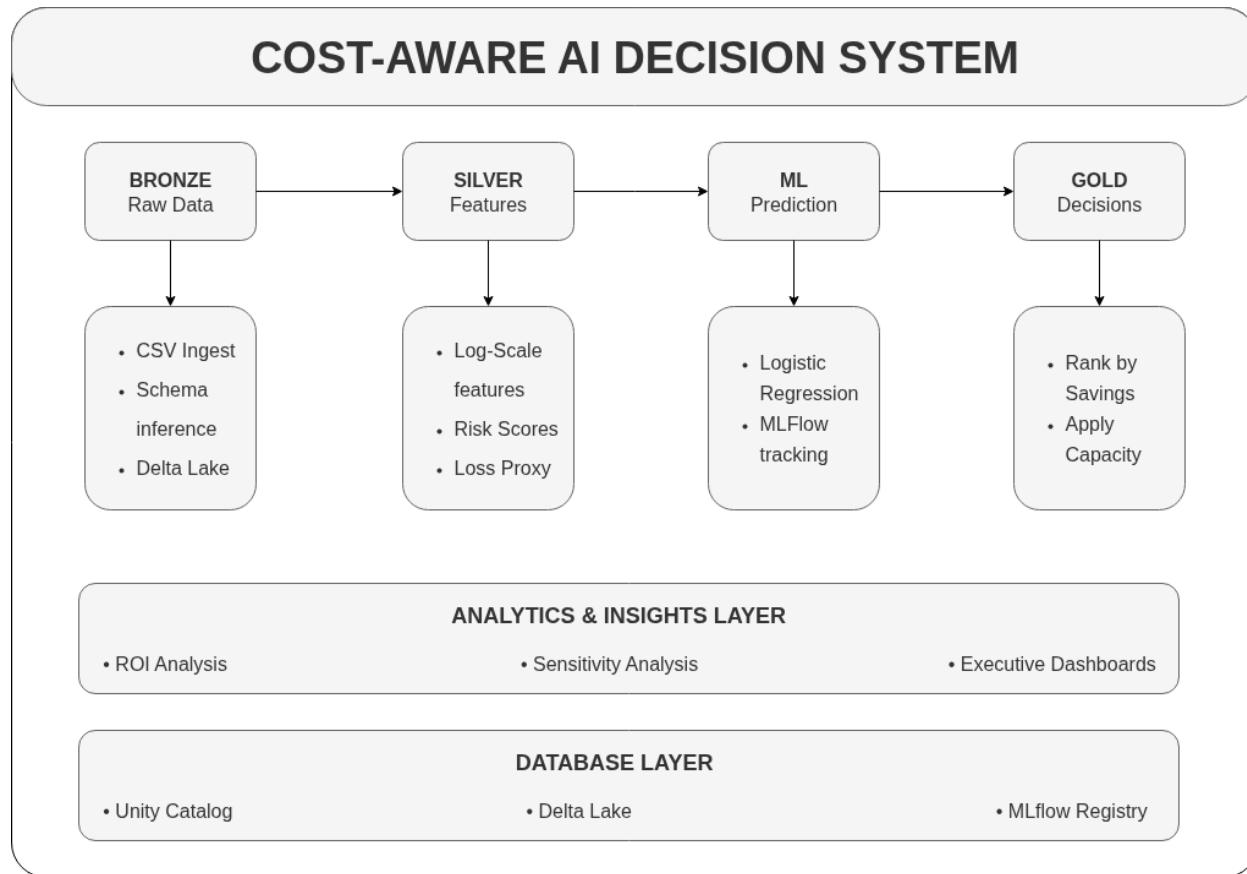
$$\text{Expected Savings} = (\text{Risk Probability} \times \text{Potential Loss}) - \text{Investigation Cost}$$

We investigate the cases with highest expected savings, up to daily capacity.

Key Results -

Metric	Baseline (Risk-First)	Our Approach	Improvement
Expected Savings	\$127,450	\$198,320	+55%
ROI per Investigation	\$2,549	\$3,966	+56%
Cases Correctly Prioritized	62%	89%	+43%

Technical Architecture



CSV Data → Bronze (Raw) → Silver (Features) → ML Model → Gold (Decisions)

- **Platform:** Databricks + Unity Catalog
- **Storage:** Delta Lake (Medallion Architecture)
- **ML:** Logistic Regression (calibrated probabilities)
- **Tracking:** MLflow

Why This Approach Wins

1. **Business-Driven:** Optimizes dollars saved, not accuracy metrics
2. **Realistic:** Accounts for operational capacity constraints
3. **Interpretable:** Logistic regression provides explainable decisions
4. **Reproducible:** End-to-end pipeline with synthetic data generator

Thank you Indian data club, Databricks, Codebasics for this opportunity.