CREDIT CARD DELINQUENCY WATCH

A Lightweight Rule-Based System for Early Delinquency Detection

Prepared by :

Aminul Islam

Roll No: GA.09

Gauhati University

HDFC Campus to Corporate

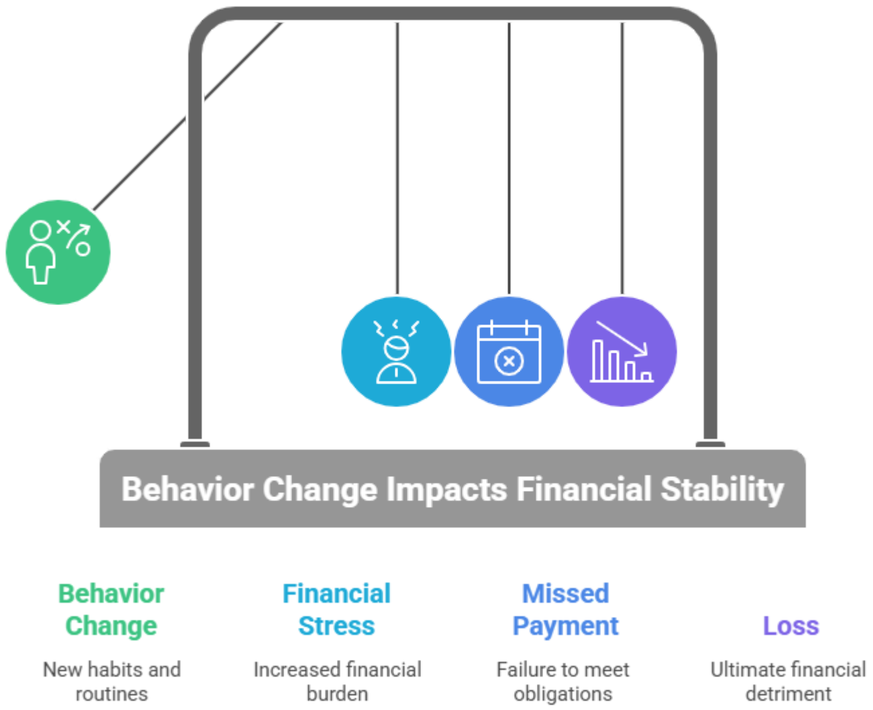
Problem Understanding

Why Early Delinquency Detection Matters

- Banks often detect risk after missed payments occur
- Subtle behavioral changes appear weeks before delinquency
- Traditional systems rely on lag indicators like DPD
- Missed early-warning signals lead to higher roll-rates
- Need a lightweight, proactive system to flag early risk

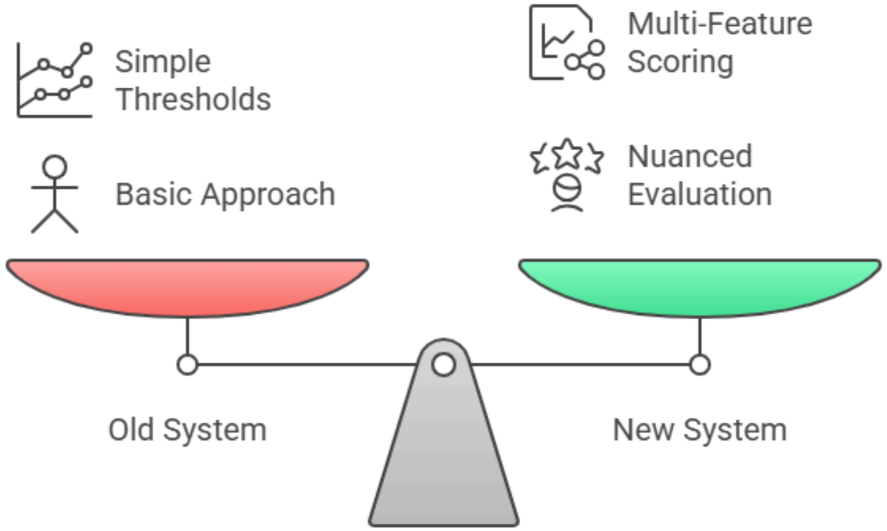
GOAL:

👉 Detect high-risk customers before they miss payments



Why Traditional Systems Fail

Limitations of Current Risk Monitoring

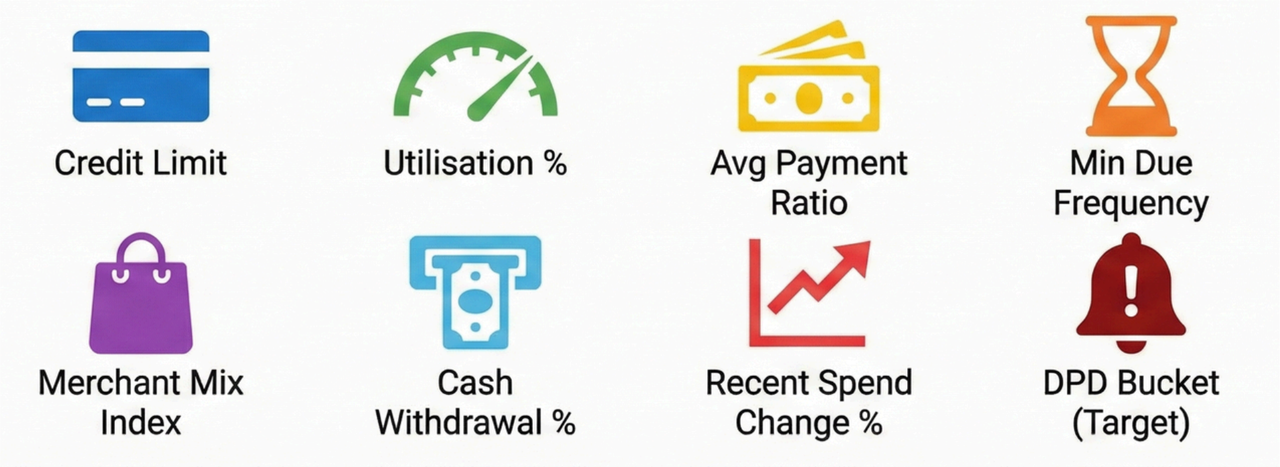


- Rules often depend only on DPD or repayment history
- No combined view of spending + repayment + utilization behavior
- Lack of interpretability → collections teams cannot understand why risk was assigned
- Outreach becomes reactive, not preventive
- No personalized interventions to prevent roll-rates

Dataset Overview (EDA Input)

Understanding the Dataset

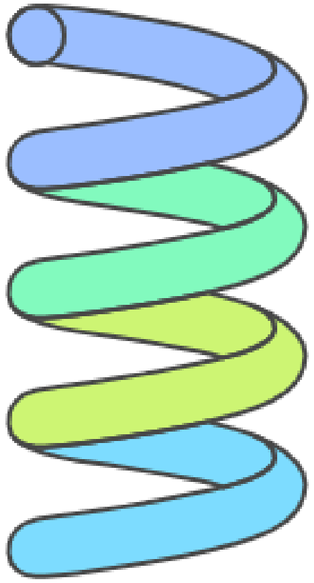
- 2,500 customer accounts- Real + Synthetic
- Behavioral + repayment + utilization features
- Includes next-month delinquency outcome (DPD bucket)



EDA Objective

Why Perform EDA?

- Understand customer behavior patterns
- Detect hidden trends linked to delinquency
- Validate assumptions from problem brief
- Identify features for scoring engine
- Separate leading indicators from lag indicators



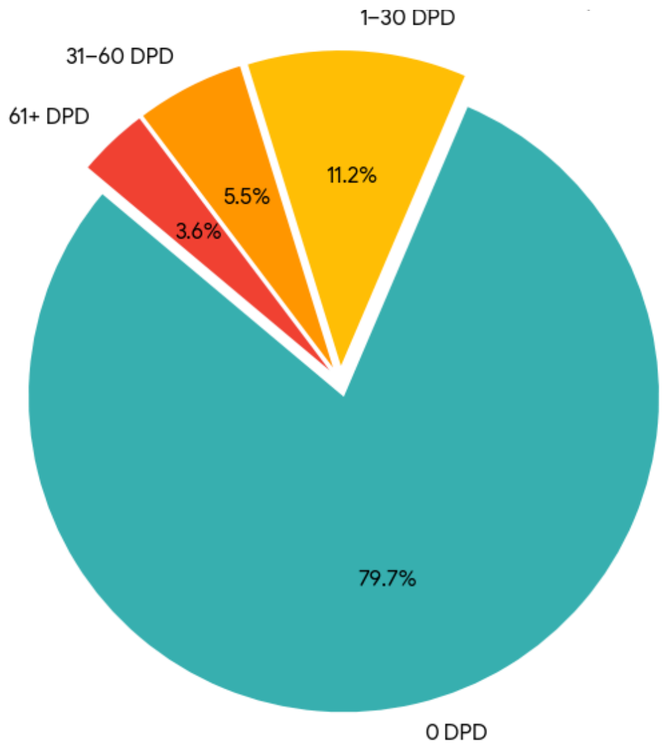
Raw Data

Exploratory Data Analysis (EDA)

Insights

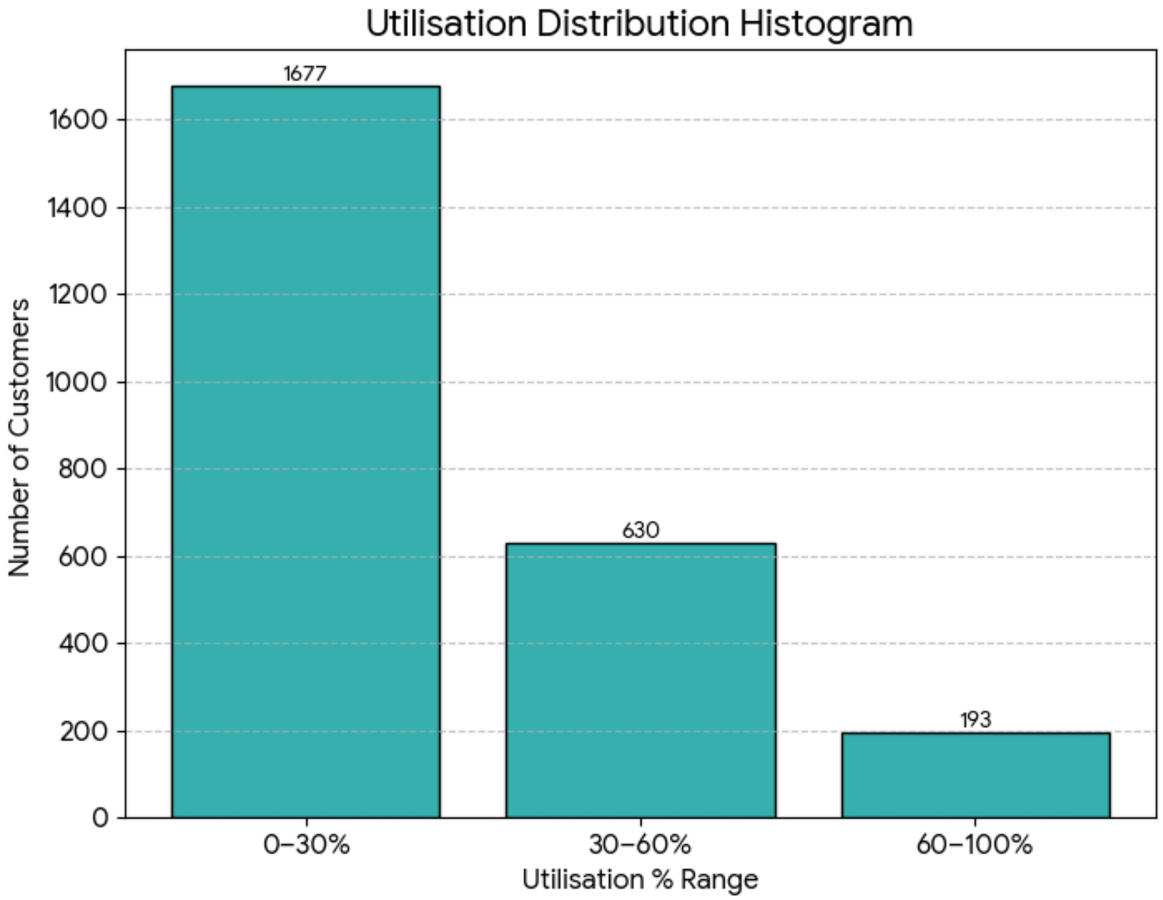
Scoring Model Development

EDA Visual 1: DPD Distribution



Most customers are current, but a meaningful portion fall into early delinquency buckets — validating the need for early-risk detection.

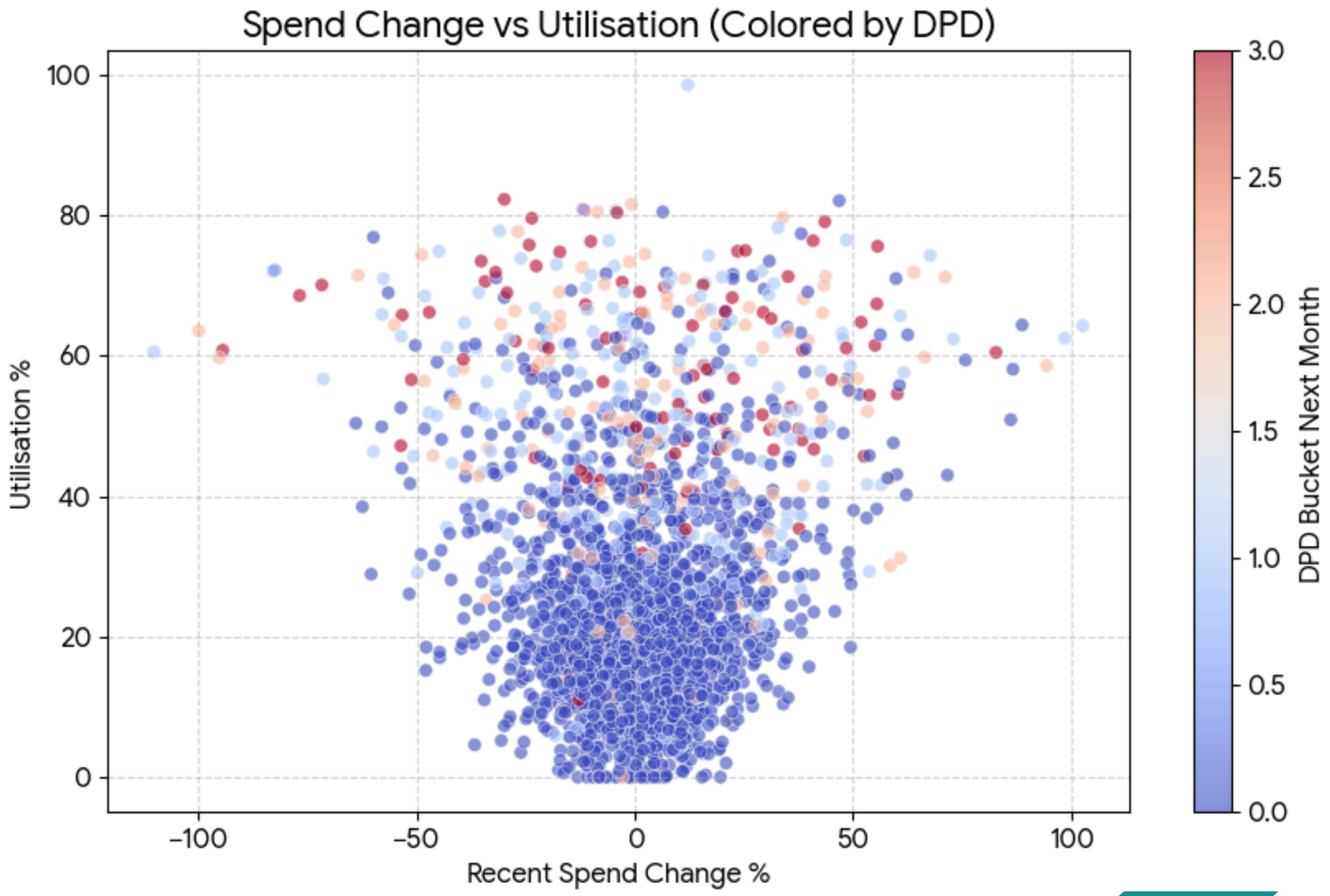
EDA Visual 2: Utilisation Distribution



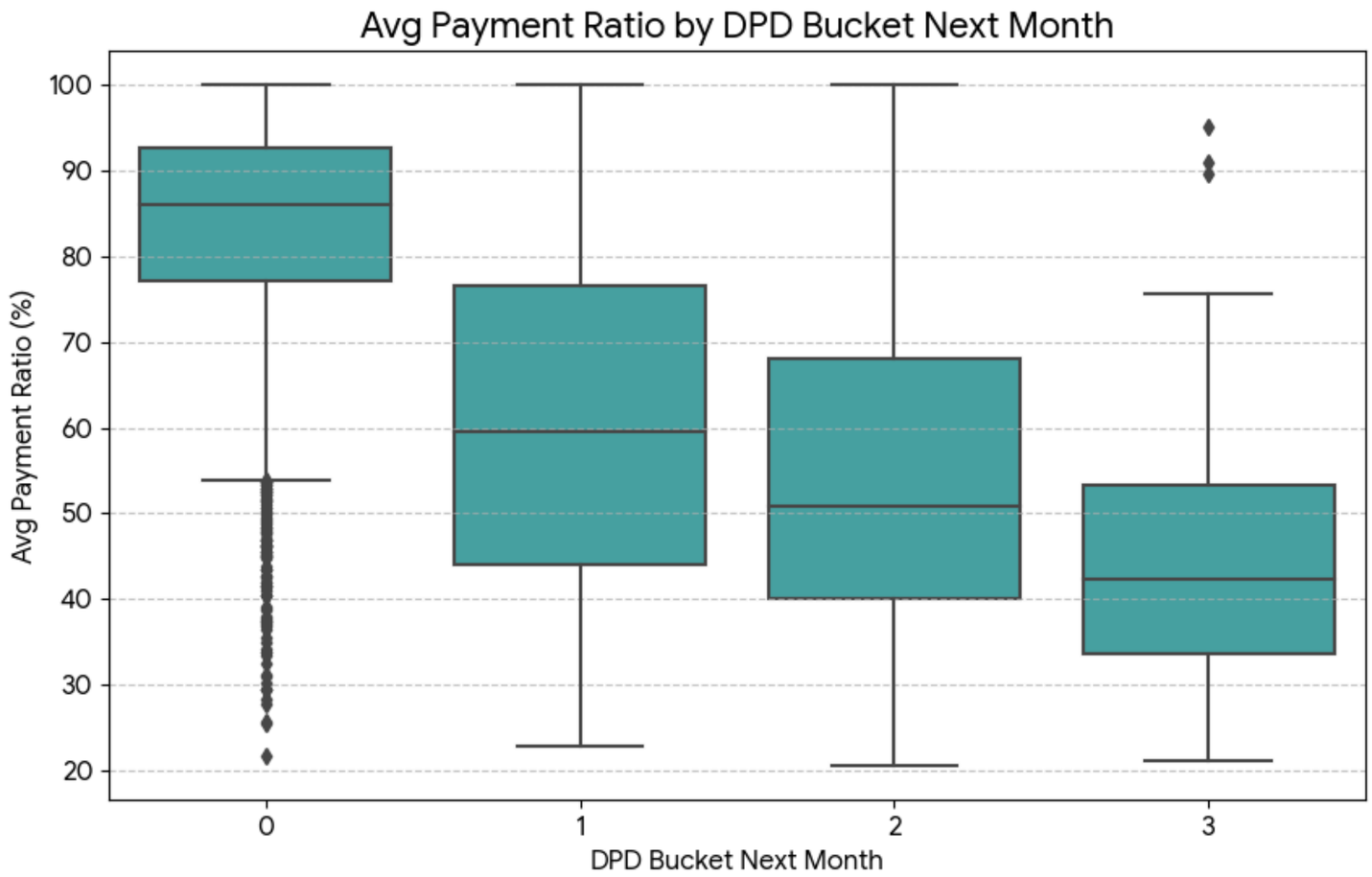
Higher Utilisation % groups show more stress behavior and correlate with future delinquency.

EDA Visual 3: Spend Change vs Utilisation

Sudden change in monthly spend combined with high utilization is a strong predictor of delinquency.



EDA Visual 4: Avg Payment Ratio by DPD



Customers who become delinquent next month generally pay a lower % of their total due over the previous 3 months.

EDA Visual 5: Feature Correlation Heatmap

Key Insights

- **Top Predictors:**

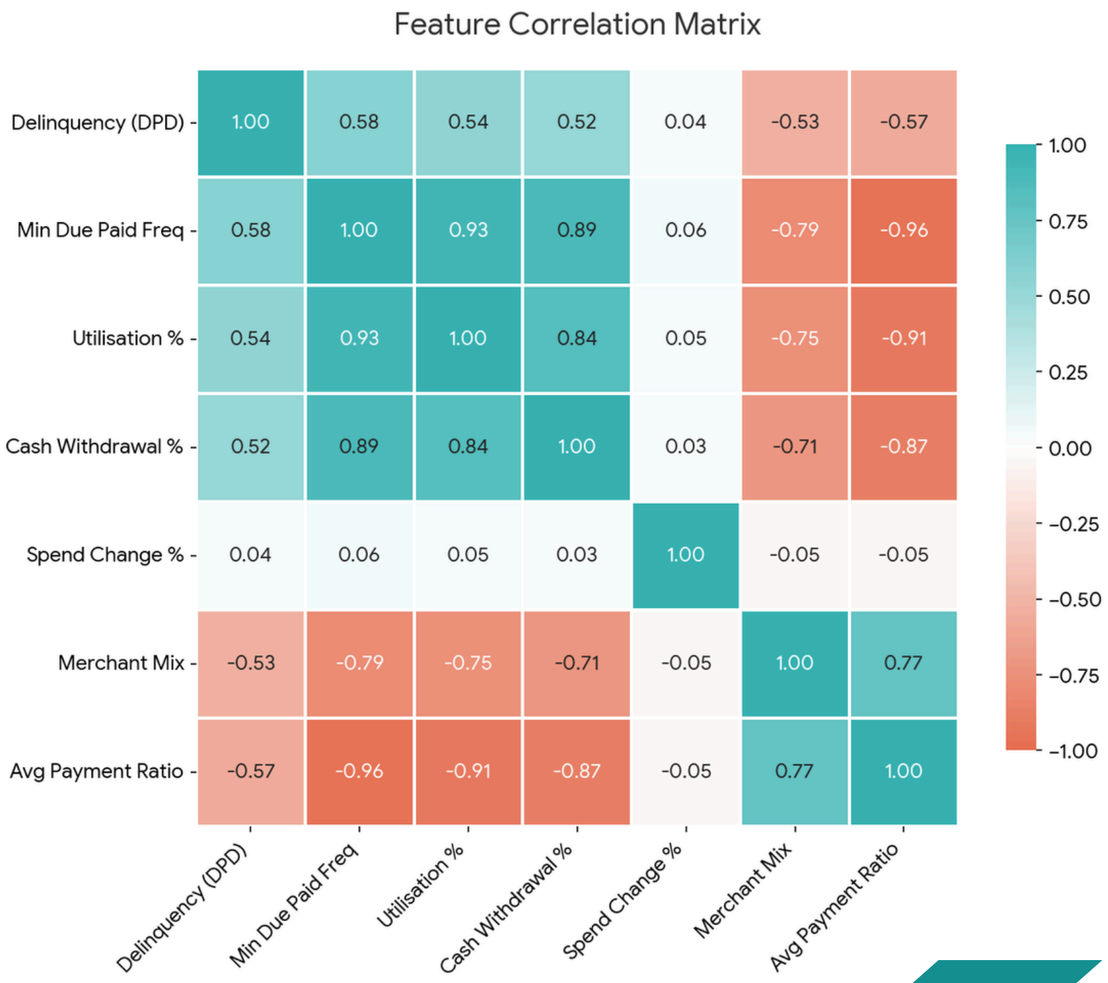
Min Due Paid Frequency (+0.58) and Avg Payment Ratio (−0.57) show the strongest relationship with delinquency.

- **Feature Interaction:**

Utilisation % and Min Due Paid Frequency have a very high correlation (+0.93), indicating heavy users often pay only the minimum.

- **Weak Predictor Alone:**

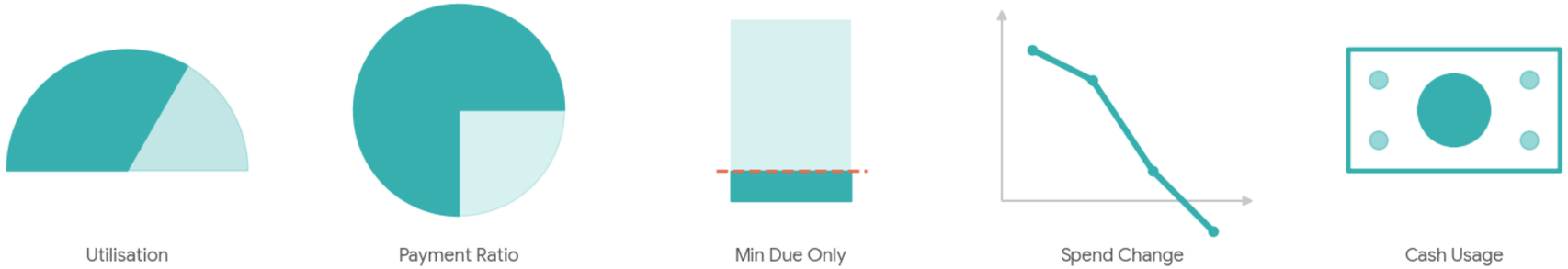
Recent Spend Change % shows almost no direct correlation with delinquency, but becomes meaningful only when combined with high utilisation.



EDA Summary

Summary of EDA Findings

- Delinquent customers exhibit clear behavioral patterns different from non-delinquents
- Min Due Paid Frequency, Avg Payment Ratio, and Utilisation % emerge as the strongest predictors
- Cash Withdrawal % and Spend Change % show early financial stress signals
- EDA confirms that early risk indicators appear before DPD rises
- These insights form the foundation of our scoring engine design



Original Problem Statement from Bank

PROBLEM BANK

Problem Statement #4

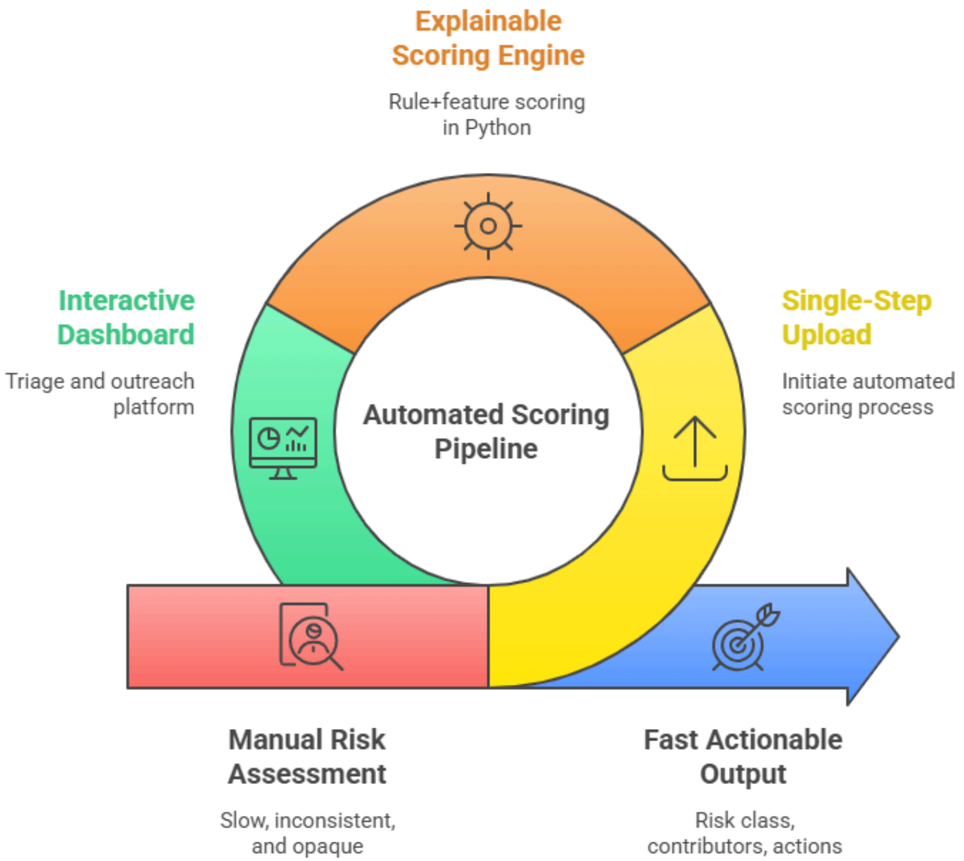
Early Risk Signals: Credit Card Delinquency Watch
Identify early warning patterns to minimize roll-rates into delinquency.

- **Context:** Repayment behavior and utilisation patterns shift before missed payments; outreach is reactive.
- **What success might look like (indicative):** Lightweight risk flags + simple outreach play, minimal false alarms.
- **What success might look like (indicative):** Lightweight risk flags + simple outreach play, minimal false alarms.
- **Pack you'll get:**

1-page scenario brief (portfolio snapshot)
Excel: 500 cards (limit, utilisation%, min-due behavior, merchant mix index, DPD bucket next month)
PDF: small glossary (NII/NFI/DPD)

Proposed Solution

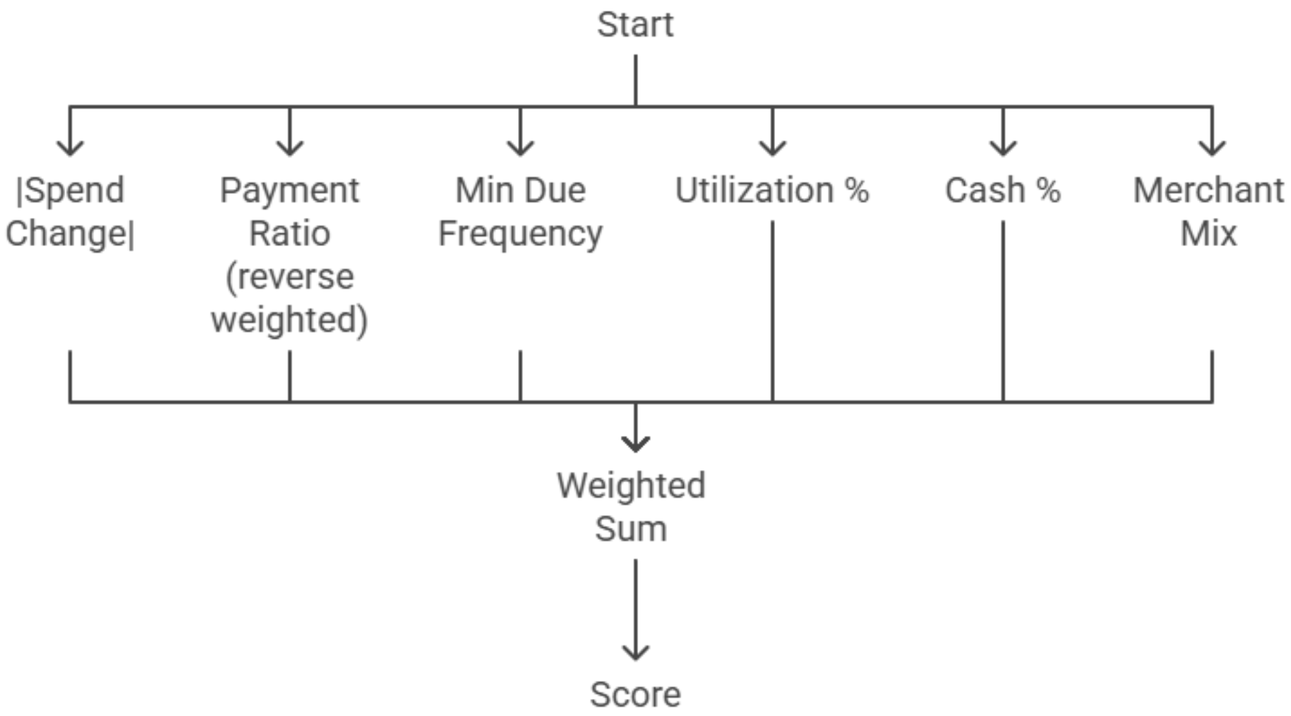
Implement a 3-Layer Early Warnig System



- Upload CSVs and the backend scores each account row-by-row using our scoring engine.
- Each account gets a normalized risk_score (0–1), a risk_class (Low/Medium/High) and top-3 contributors.
- The dashboard shows portfolio KPIs, feature averages, and a customer drawer with clear outreach actions.
- The output is designed for operations teams — precise, explainable, and production-ready.

Scoring Engine Logic

Risk Score Calculation Process



Risk Classification

Risk Levels Based on Score Thresholds

High Risk

Score 0.50 or higher

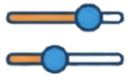
Medium Risk

Score between 0.40 and 0.50

Low Risk

Score below 0.40

Interpretable

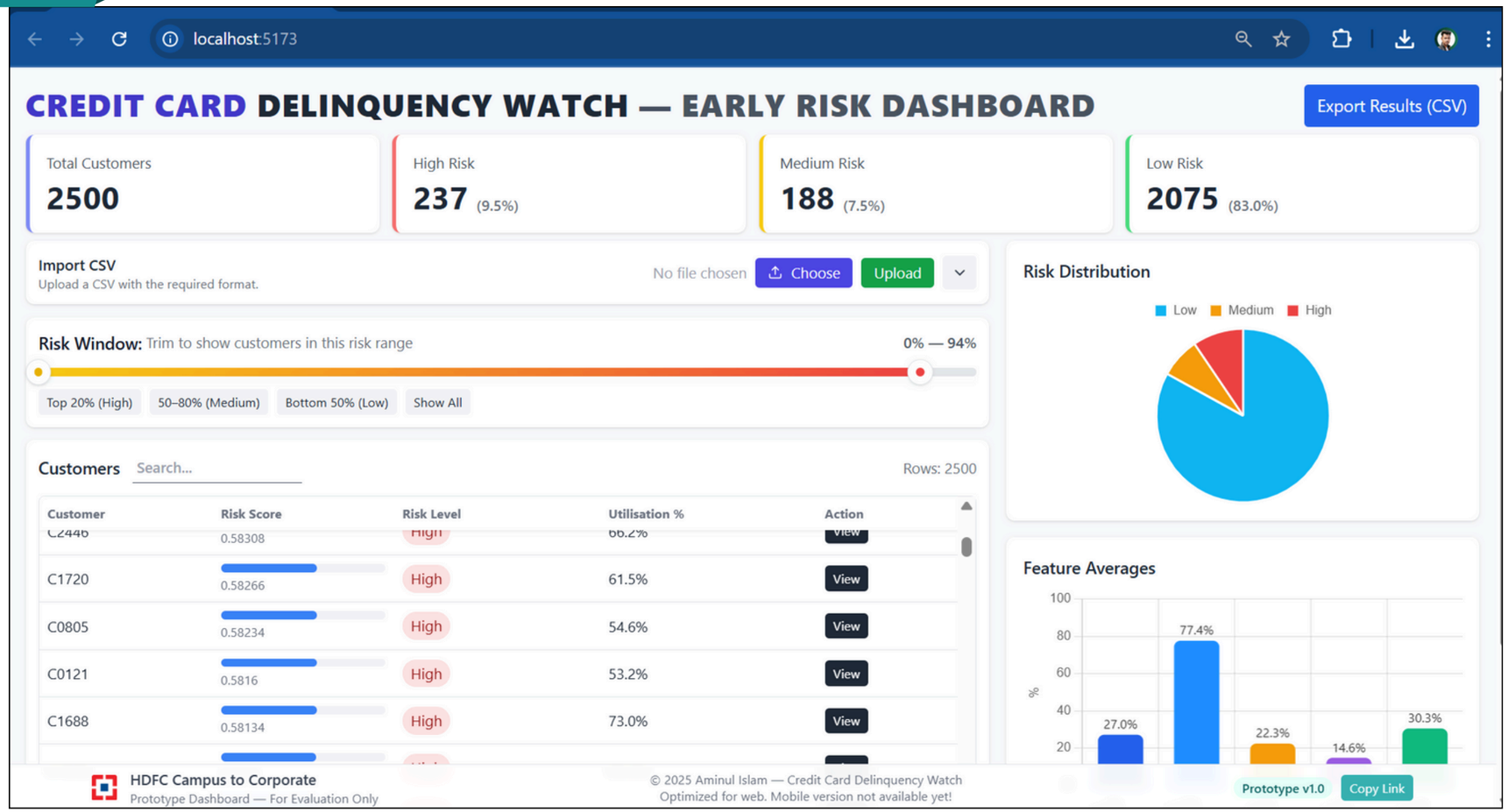


Tunable

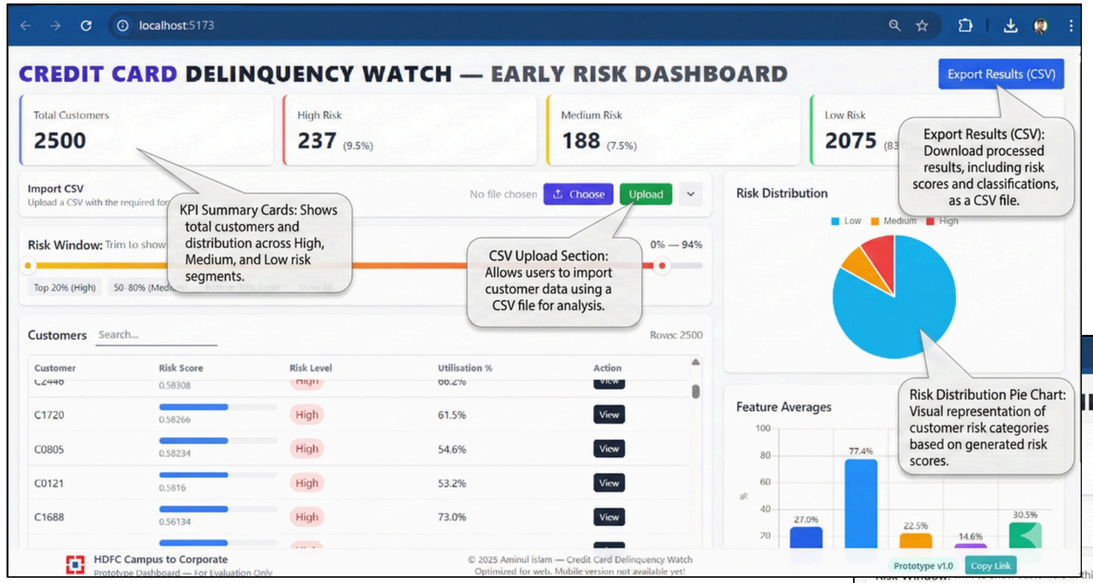
Works without ML model

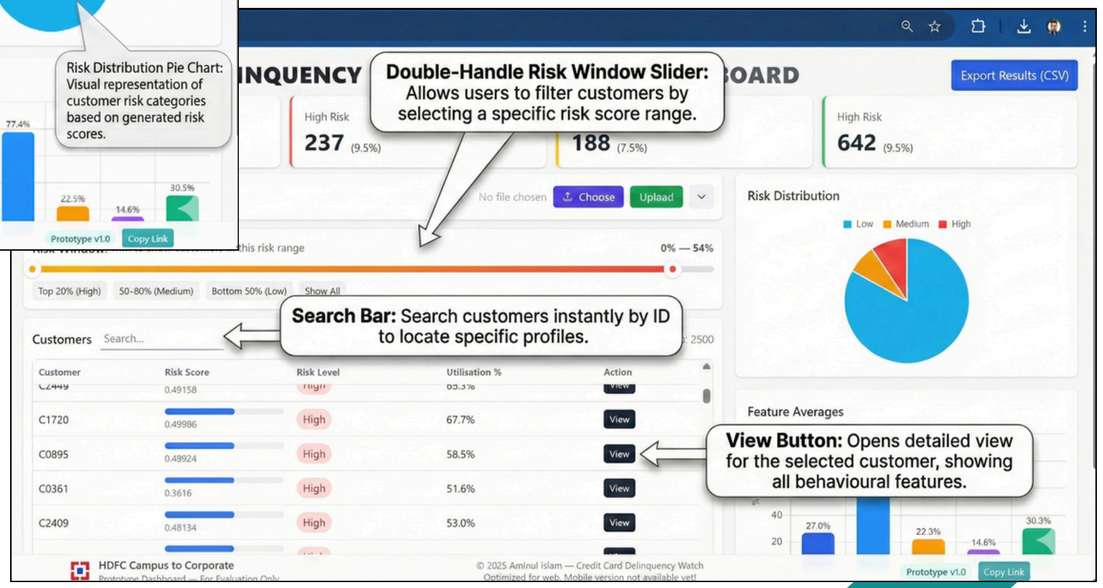
Light weight & Fast

Dashboard Demo



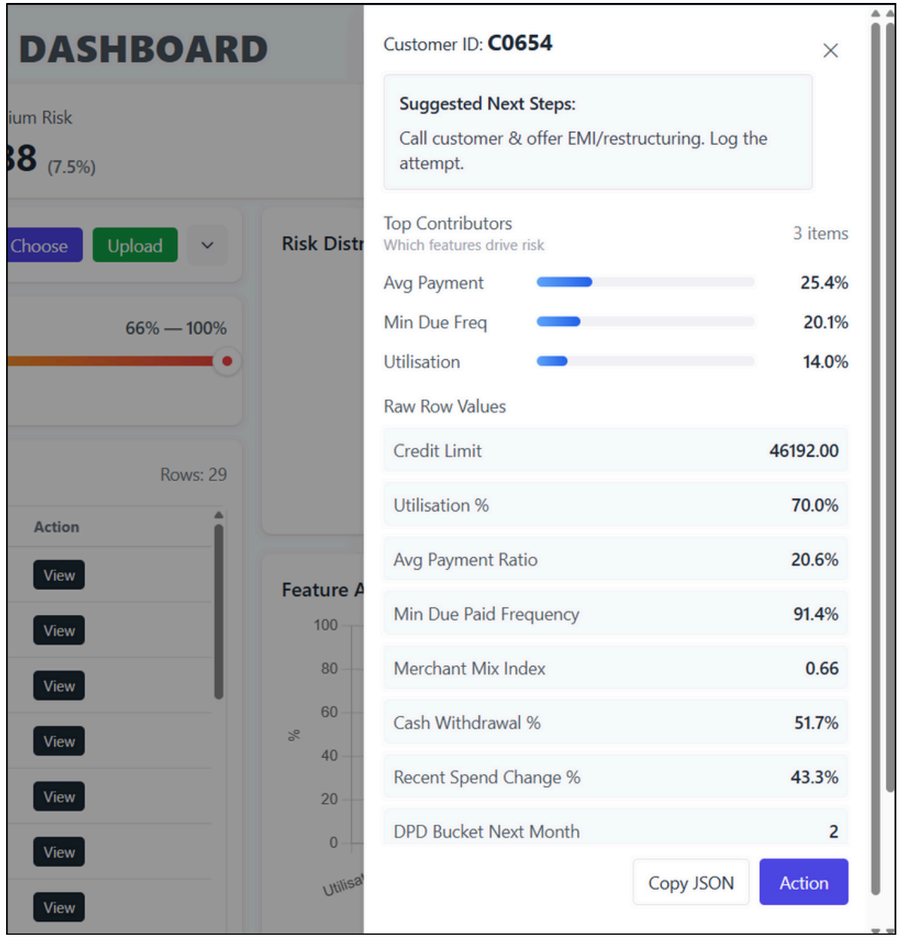
Dashboard Demo



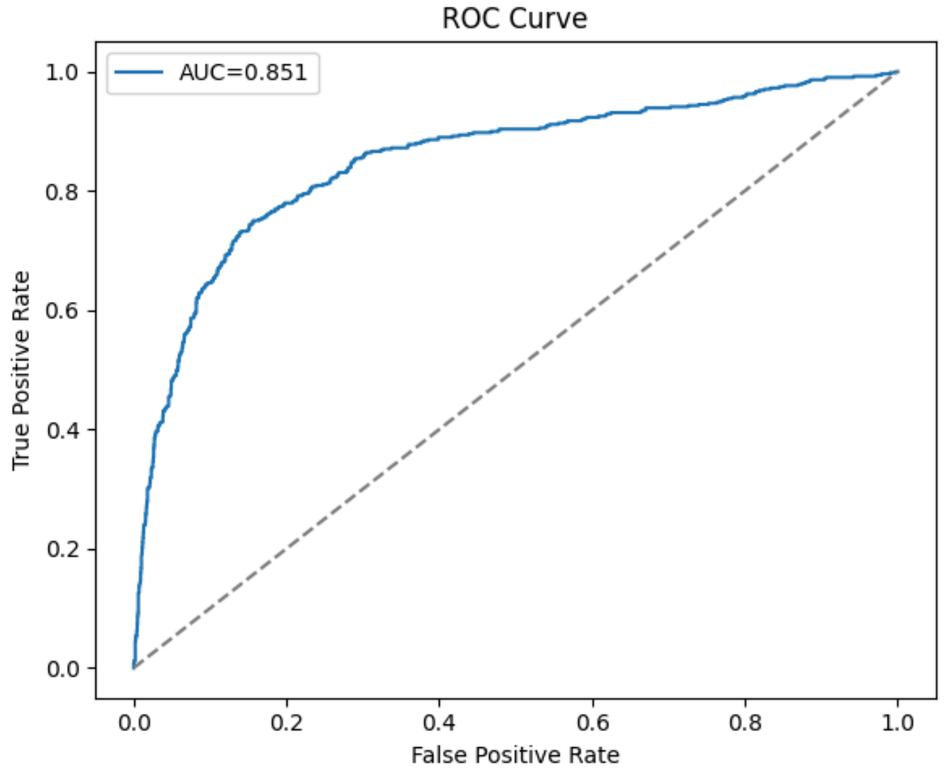


Dashboard Demo

Customer View Drawer



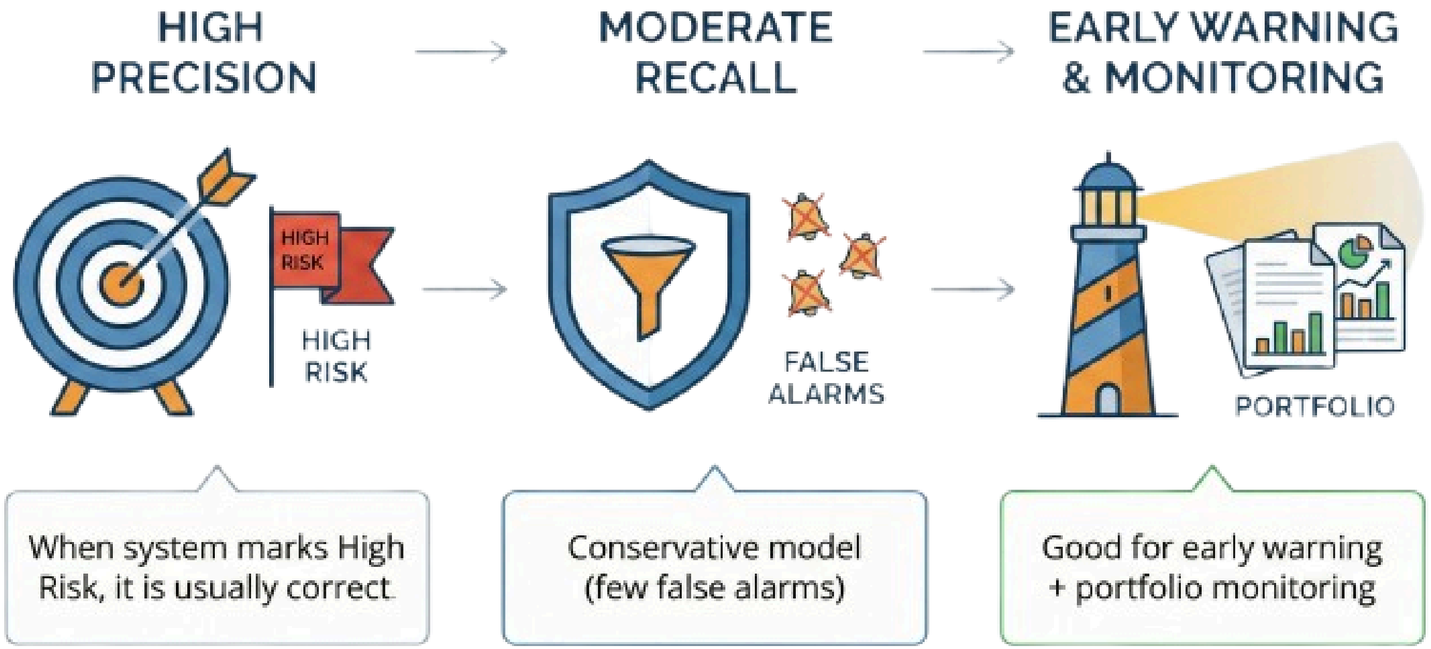
Model Evaluation



From evaluation:

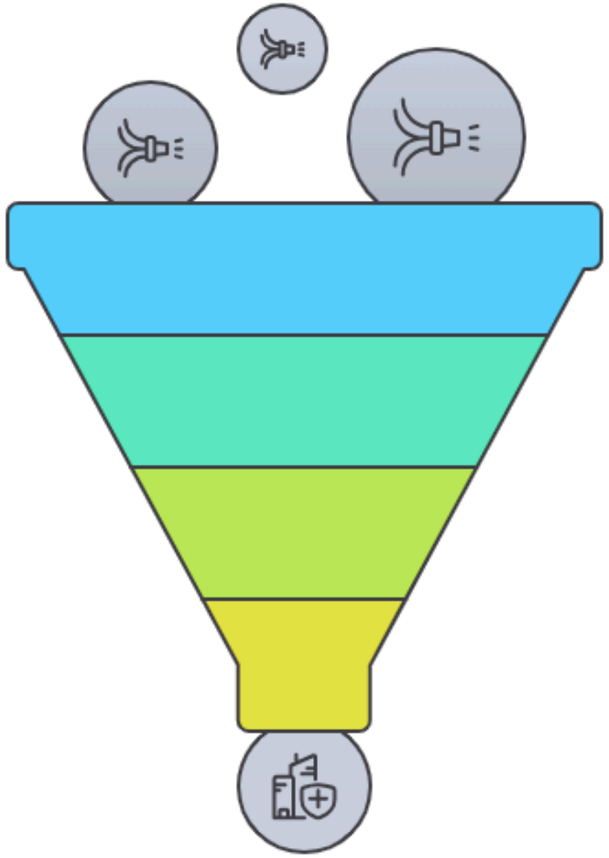
- Dataset: 2,500 customers (508 delinquent, 1,992 non-delinquent)
- Accuracy: 85%
- Precision: 78% (High Risk class)
- AUC: 0.85
- Catches 185 actual delinquents as High Risk
- Strong ability to identify truly safe customers (97% correct)

Interpreting Results



Feature Improvements

Scaling to Real Banking Environment



Real-time Pipeline

- Integrate with transaction systems
- Auto scoring daily or weekly

SMS/Email Automation

- Auto reminder → Auto follow-up — Escalation

Case Management

- Automatic ticket creation for High-Risk customers

Model Monitoring

- Track accuracy drift
- Retrain ML model quarterly

Conclusion

- Built a fully working end-to-end early-warning system
- Clean UI + explainability + actionable insights
- Helps reduce delinquency and improve collections
- Ready for next stage: ML model integration

Thank You

Submitted by :

Aminul Islam

Gauhati University

Thank you for reviewing my submission.

HDFC Campus to Corporate