"Banks lose money when customers don't repay loans, and they need a way to predict this risk earlier."

Problem Statement

Banks face financial losses when customers default on loans. To mitigate these losses, banks need a way to assess the credit risk of loan applicants before granting a loan. The goal is to develop a predictive system that helps banks identify high-risk applicants early, enabling better lending decisions.

Problem Objective

Build a model to predict the probability of loan default using customer and financial information before the loan is issued.

Stakeholders

Stakeholder	Role in Problem	Interest/Impact
Bank Management	Decision makers	Minimize loan defaults
Loan Officers	Using model to assess risk	Make smarter lending choices
Data Science Team	Model developers	Build accurate predictions
Customers	Loan applicants	Fair and timely evaluation

Features

- Demographic information (age, gender, education)
- Income & employment details
- Credit history (credit score, past defaults)
- Existing debts or liabilities
- Loan-related info (amount, purpose, duration)

Target Variable

Binary Classification:

- 1 = Default (customer fails to repay loan)
- 0 = No Default (customer repays loan successfully)

Outcome (Success)

- Business Metric: Reduce loan default rate, increase profitability
- Model Metric: High accuracy, precision, recall, and especially AUC-ROC for imbalanced data

Limitations and Constraints

- Data availability and quality
- Fairness and bias concerns
- Regulatory compliance
- Real-time decision-making constraints

Risks

- Misclassification: Approving high-risk applicants or rejecting low-risk ones
- Customer dissatisfaction due to opaque decision process
- Model drift over time as economic conditions change

Solution

Supervised Model

Risk Scoring System

API or Web Interface

Dashboard for risk scoring