

Team 4

Project Proposal: Loan Default Prediction

Project Overview

Banks and financial institutions face major financial risks when customers fail to repay their loans. This project aims to develop a **machine learning model** that predicts whether a customer is likely to default on a loan based on their financial and demographic information.

By identifying **high-risk borrowers**, banks can make better lending decisions, reduce losses, and design fairer credit policies.

Problem Statement

When customers cannot pay back their loans, it affects the bank's revenue and increases operational risk.

Traditional credit scoring methods often fail to capture complex relationships between features such as **income**, **loan amount**, **credit score**, and **repayment history**.

Our goal is to use **AI and data-driven modeling** to:

- Identify patterns that indicate loan default risk
 - Help financial institutions manage credit more effectively
 - Demonstrate the real-world value of predictive analytics in banking
-

Objectives

1. Perform **data cleaning** and handle missing or inconsistent values.
 2. Conduct **exploratory data analysis (EDA)** to find trends and correlations.
 3. Apply **classification algorithms** (Logistic Regression, Random Forest, XGBoost).
 4. Evaluate models using **accuracy, precision, recall, and ROC-AUC**.
 5. Visualize and explain **key risk factors** behind defaults.
 6. Summarize findings with clear **business recommendations**.
-

Dataset

Source: Kaggle – Loan Default Prediction Dataset

The dataset contains anonymized financial data of loan applicants, including:

- `loan_amount`, `term`, `interest_rate`, `income`, `credit_score`,

- employment_status, payment_history, and loan_status (target variable)

Expected Outcome

- A trained ML model that predicts whether a borrower is **low-risk** or **high-risk**.
- Visualizations showing which features have the most impact on loan default.
- A clean, reproducible notebook and a clear business summary for presentation.
- Practical insights that demonstrate how **AI can enhance credit risk management**.