

Project Initialization and Planning Phase

Date	5 July 2025
Team ID	SWTID1749835773
Project Title	Applicant Credibility Prediction for Loan Approval
Maximum Marks	3 Marks

Project Proposal (Proposed Solution) template

This project proposal outlines a solution to address a specific problem. With a clear objective, defined scope, and a concise problem statement, the proposed solution details the approach, key features, and resource requirements, including hardware, software, and personnel.

Project Overview	
Objective	To develop a machine learning–based system that predicts the credibility of loan applicants using historical loan data, thereby supporting banks in making fair, data-driven, and efficient loan approval decisions.
Scope	The project includes data analysis, model development, and optional model deployment. It will involve data cleaning, exploratory analysis, model training, evaluation, and fairness checks using the uploaded loan prediction dataset. The system will aim to enhance decision-making efficiency and reduce human bias.
Problem Statement	
Description	Financial institutions manually process large volumes of loan applications, often relying on static criteria and human judgment. This approach is prone to inconsistency, delays, and bias.
Impact	A credible automated solution will: <ul style="list-style-type: none"> • Improve efficiency in the loan approval process. • Reduce human errors and subjectivity. • Enable responsible lending using consistent and scalable AI methods.
Proposed Solution	

Approach	<p>We will follow a data-driven development methodology:</p> <ol style="list-style-type: none"> Data Understanding: Analyze the dataset structure and features. Data Preprocessing: Handle missing values, categorical encoding, and outliers. EDA: Visualize relationships between applicant attributes and loan status. Modeling: Train models (e.g., logistic regression, decision trees, random forest). Evaluation: Use metrics like accuracy, precision, recall, and AUC. Fairness Analysis: Investigate demographic bias. (Optional) Deployment via a simple web interface using Flask.
Key Features	<ul style="list-style-type: none"> AI-powered loan credibility scoring. Transparent model with interpretable features. Bias detection and mitigation strategies. Optional UI for model usage by loan officers.

Resource Requirements

Resource Type	Description	Specification/Allocation
Hardware		
Computing Resources	CPU/GPU specifications, number of cores	8-core CPU or GPU-enabled environment (e.g., Google Colab, Kaggle Kernel)
Memory	RAM specifications	8–16 GB RAM
Storage	Disk space for data, models, and logs	1–2 GB for dataset and model artifacts
Software		
Frameworks	Python frameworks	Python, Flask (for optional deployment)
Libraries	Additional libraries	pandas, numpy, sklearn, matplotlib, seaborn, xgboost
Development Environment	IDE, version control	Jupyter Notebook, Git for version control

Data		
Data	Source, size, format	Provided CSV dataset (loan_prediction.csv) Structured CSV with ~600+ records and attributes such as Gender, Education, ApplicantIncome, LoanAmount, Loan_Status, etc.