# Milestone 1 – Step 1: Understanding and Documenting the Project Requirements

## Project Title:

## CreditPathAI – Smart Loan Recovery System

## 1. Project Overview

CreditPathAI is an AI-driven Smart Loan Recovery System designed to automate and optimize the loan recovery process for financial institutions. It leverages machine learning algorithms to analyse borrower repayment behaviour and predict the probability of loan default. Based on these predictions, the system provides personalized recovery recommendations to help agents take the most effective actions. The platform combines data processing, predictive modelling, API deployment, and interactive dashboards to create a complete data-driven solution for improving loan recovery efficiency and minimizing financial losses.

## 2. Project Objectives

- To design and develop a **machine learning–based platform** that predicts borrower default risk with high accuracy.

- To create a **recommendation engine** that suggests customized recovery strategies based on risk levels.

- To build an **API service** for serving model predictions in real time.

- To develop an **interactive web dashboard** for financial agents to monitor borrower performance and actions.

- To ensure the system is **cost-effective, scalable, and reliable** by using open-source technologies.

## 3. Key Deliverables

- Cleaned and pre-processed loan repayment dataset.

- Machine learning models (Logistic Regression, XGBoost, LightGBM) trained to predict loan defaults.

- A FastAPI service that exposes the model for real-time predictions.

- A React.js-based dashboard with Plotly.js for visualization and agent insights.

- Complete documentation and rollout plan for deployment.

**4. Key Performance Indicators (KPIs)**

- **AUC-ROC Score ≥ 0.8** (for model accuracy).

- **Precision and Recall ≥ 70%** (for prediction reliability).

- Reduction in manual recovery time by at least **25–30%** through automated insights.

- Enhanced efficiency of financial recovery operations through AI-driven recommendations.

**5. Tools and Technologies**

| Category | Tools/Technologies Used |
| --- | --- |
| Data Storage & Processing | CSV/Excel, SQLite, PostgreSQL, Pandas, Dask |
| Machine Learning & MLOps | Python, scikit-learn, XGBoost, LightGBM, MLflow |
| API & Backend Services | FastAPI, Docker |
| Frontend & Visualization | React.js, Plotly.js |
| Monitoring & Logging | FastAPI built-in logging, SQLite logs |
| Version Control | GitHub (for source code and documentation) |

**6. Stakeholders / End Users**

- **Loan recovery agents:** Use recommendations to plan borrower follow-ups.

- **Credit managers:** Use dashboards to monitor risk patterns.

- **Financial analysts:** Evaluate the effectiveness of recovery strategies.

- **Developers & data scientists:** Maintain and improve the model's accuracy and performance.

**7. Timeline (High-Level Overview)**

| Milestone | Objective | Expected Duration |
|---|---|---|
| 1 | Requirements & Data Collection | 1 week |
| 2 | Data Ingestion & Initial EDA | 1 week |
| 3 | Baseline Model Development | 1 week |
| 4 | Advanced Model Training | 1 week |
| 5 | Recommendation Engine & API Prototype | 1 week |
| 6 | Frontend Development & Final Delivery | 1 week |

**8. Summary**

Milestone 1 focuses on defining a **clear project vision** and understanding its requirements before beginning technical implementation. By outlining objectives, KPIs, tools, and expected outcomes, this step lays the foundation for a structured, data-driven development process. The ultimate aim is to use AI to make loan recovery smarter, faster, and more efficient.