1. Problem Statement

A mid-sized NBFC is facing challenges in recovering overdue loans. Traditional collection methods are proving inefficient and often lead to customer dissatisfaction. The company wants to leverage data science and AI to:

1. Predict which customers are most likely to miss next payment.
2. Recommend the most effective recovery strategy for each customer.
3. Automate customer interactions using a chatbot that adapts its tone and messaging based on customer personas (e.g., cooperative, evasive, aggressive, confused).

2. Objectives

a. **Explore and analyse** a loan collection dataset.
b. **Build a predictive model** to identify potential loan defaulters.
c. **Design a strategy recommendation engine** based on customer behaviour and sentiment.
d. **Develop a persona-based chatbot** that customizes its responses to different customer types.

3. Steps

A: Predictive Modelling

- **Objective**: Build a model to predict the likelihood of missing next payment.
- **Techniques**: Use Machine Learning , feature engineering, and interpretability.

B: Strategy Recommendation Engine

- **Objective**: Based on prediction and customer profile/persona, recommend the best collection strategy. This strategy can also be revised based on chat-bot response.

C: Persona-Based Chatbot Development

- **Objective**: Develop a basic chatbot that:
  - Identifies customer persona (e.g., cooperative, evasive, aggressive, confused).
  - Adapts tone (e.g., empathetic, assertive, informative) and content accordingly.
- **Techniques**:
  - Sentiment and intent detection.
  - Prompt engineering or fine-tuning using LLM APIs.
  - Basic UI mock-up or command-line interface.

4. **Dataset Description:**

- **Customer demographics**: age, income, location, employment status.
- **Loan details**: amount, tenure, interest rate, type.
- **Repayment history**: missed payments, delays, partial payments.
- **Interaction logs**: previous communication attempts, sentiment scores, response times.
- **Behavioural data**: app usage, website visits, complaint history.

Column Name	Description
CustomerID	Unique identifier for each customer
Age	Age of the customer
Income	Annual income in INR
Location	Urban/Suburban/Rural
EmploymentStatus	Employment type
LoanAmount	Loan amount in INR
TenureMonths	Loan tenure in months
InterestRate	Annual interest rate (%)
LoanType	Type of loan (Personal, Auto, etc.)
MissedPayments	Number of missed payments
DelaysDays	Total delay in days
PartialPayments	Number of partial payments made
InteractionAttempts	Number of contact attempts made
SentimentScore	Sentiment score from past interactions (-1 to 1)
ResponseTimeHours	Average response time in hours
AppUsageFrequency	App usage frequency score
WebsiteVisits	Number of visits to the loan portal
Complaints	Number of complaints registered
Target	Miss next payment

5. Evaluation Criteria

Category	Weight
Data Exploration & Insights	10%
Predictive Model Accuracy & Interpretability	20%
Innovation in strategy recommendation	20%
Chatbot Design & Adaptability	30%
Presentation and storytelling	20%

6. Submission Format

- Jupyter Notebook (.ipynb)
- PPT and PDF report
- Any additional files (e.g., chatbot screenshots , visualizations)