**CreditPathAI – Smart Loan Recovery System**

**Project Overview**
CreditPathAI is an end-to-end AI-driven loan recovery decision support system designed for banks and NBFCs. It predicts borrower default risk, classifies borrowers into risk categories, and recommends recovery actions through an interactive dashboard.

**ML Lifecycle Followed**
Data Ingestion → EDA → Feature Engineering → Model Training → API → Frontend Dashboard

**Folder Structure**
**data/**: Stores raw and processed datasets. final_features.csv is used for model training.
**notebooks/**: Contains step-by-step ML pipeline including EDA, feature engineering, and modeling.
**src/**: Backend layer using FastAPI serving predictions via API.
**frontend/**: React.js dashboard for business users and recovery agents.
**docs/**: Documentation for data sources and KPIs.
**logs/**: Stores experiment and EDA summaries.

**Model Summary**
Logistic Regression (Baseline): AUC ~0.74
Random Forest: AUC ~0.73
XGBoost (Final Model): AUC ~0.73 with better defaulter recall

**Why XGBoost?**
Industry-grade, handles imbalance well, strong performance on tabular data, preferred in BFSI projects.

**What is creditpath.db?**
creditpath.db is a SQLite database used during EDA and feature engineering to manage large multi-table credit datasets efficiently. It enables fast joins, low memory usage, and SQL-based aggregation. Final training uses final_features.csv for simplicity and deployment efficiency.

**Business Impact**
• Reduces manual recovery cost
• Helps prioritize risky borrowers
• Improves recovery success rate

**Why This Fits Infosys**
End-to-end ML lifecycle, real BFSI use case, production-ready APIs, frontend integration, and strong engineering practices.