CrediSphere - Breach 2025 FinTech Hackathon

Hackathon: Breach 2025 FinTech Hackathon

Project: CrediSphere

Team: The Impressionists

Team Members:

- Tanay Shah – Backend
- Anirudhha Dalal – AI/ML
- Dhup Thumbadiya – Frontend Development
- Jeet Patel – Frontend Development

Dates: 21st - 23rd March 2025

Track: Web/API Development, Data Engineering, AI/ML

Problem Statement & Real-World Scenario

The Credit Risk Assessment Challenge

Financial institutions depend on **credit bureau data** for borrower risk assessment. However, **inconsistent scoring, bureau downtime, and delayed decision-making** lead to inefficiencies. Banks require a solution to **aggregate, normalize, and analyze credit data** from various bureaus for accurate, timely risk evaluations.

Real-World Scenario

ABC Bank is processing a **home loan** application for Ravi, a salaried professional.

- The bank receives credit scores from three bureaus:

Bureau A: 750 (**Good Credit Score**)

Bureau B: 600 (**Borderline Risk**)

Bureau C: Unavailable (API Downtime)

- **Conflicting scores** hinder accurate risk assessment, delaying loan approval.
- **Manual verification takes days**, inconveniencing Ravi and the bank.
- **Ravi opts for a faster lender**, resulting in **ABC Bank losing a potential customer**.

Proposed Solution

Our **Multi-Bureau Credit Scoring & Risk Aggregation System**:

Leverages **LLMs** to select **APIs** based on user context.

Manages API downtime for uninterrupted credit assessment.

Cleans, normalizes & aggregates data from multiple sources.

Utilizes **XGBoost ML** to predict **risk scores**.

Delivers **AI-driven financial insights** for enhanced lending decisions.

Provides an **AI-powered chatbot** for **contextual financial analysis**.

System Architecture & Implementation

Frontend

React.js, TailwindCSS – Ensures a fast, responsive, and modern UI.

Backend

Node.js, Express.js – Manages API calls, data processing, and authentication , flask

Database

Supabase – Stores user authentication, report data, and chatbot history.

AI/ML

- XGBoost – Predicts borrower risk score based on cleaned credit data.
- LLM (Large Language Model) – Generates financial insights from risk analysis.
- RAG (Retrieval-Augmented Generation) – Powers the chatbot to fetch relevant insights.

Infrastructure

- REST APIs, API Gateway
- and real-time data processing.

System Data Workflow

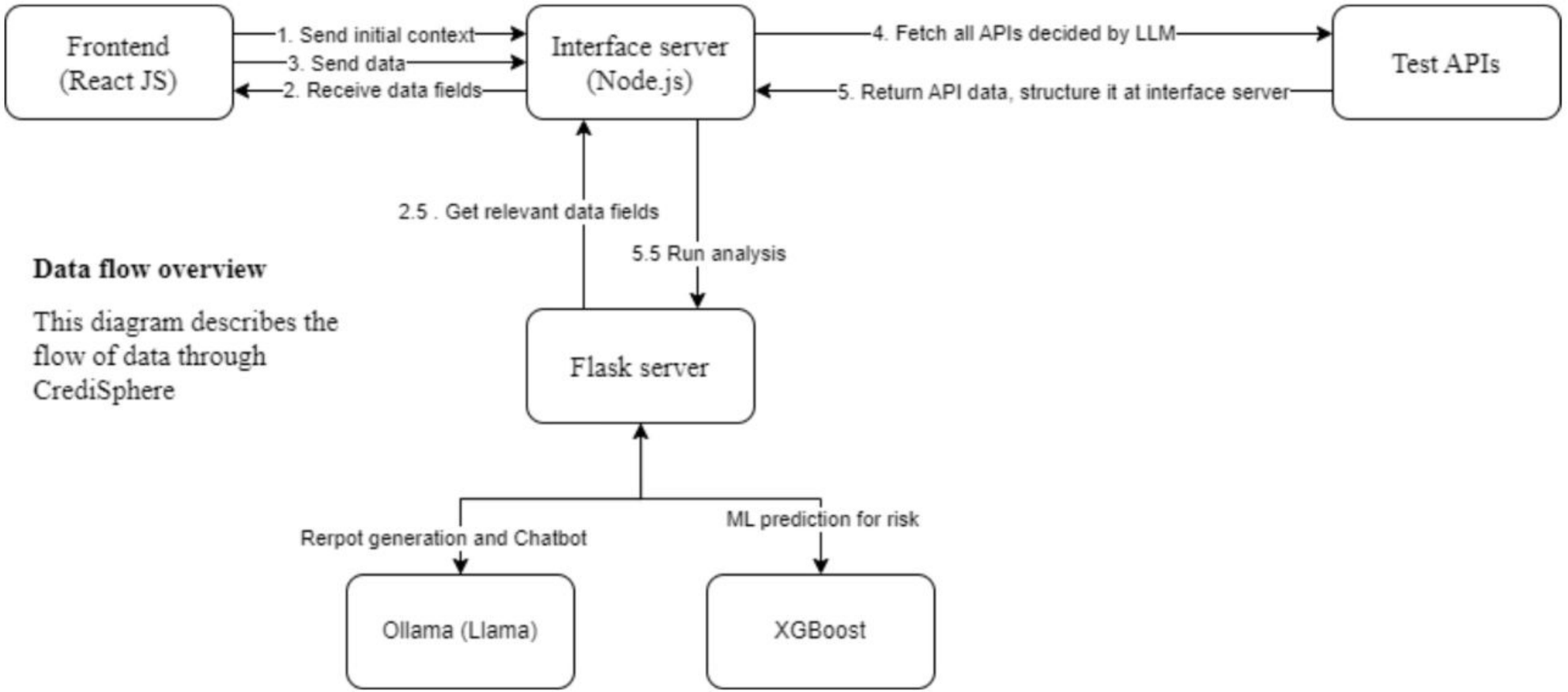

Figure 1: System data flow diagram

Prototype Demonstration & Key Features

1

Loan Context Input

Users provide loan specifics, which the system analyzes to dynamically choose relevant credit bureau APIs.

2

Real-Time API Calls & Data Aggregation

The backend retrieves and normalizes credit scores from multiple bureaus.

Fallback protocols guarantee continuous risk assessment, even during API downtime.

3

Risk Score Prediction

An XGBoost model predicts the borrower's risk score based on aggregated credit data.

The result is stored for subsequent AI analysis.

4

AI-Generated Financial Insights

A context-aware LLM processes the risk score and generates detailed credit risk insights.

These insights explain creditworthiness, repayment risk, and financial trends.

5

PDF Upload & AI-Assisted Processing

PDF upload is enabled post insight generation.

Users can upload financial statements/reports, enabling the chatbot to deliver enhanced risk assessments.

6

AI Chatbot for Credit Analysis

A RAG-powered chatbot retrieves insights from AI reports and user-uploaded PDFs.

This chatbot provides contextual, memory-enabled responses to aid lenders in their decisions.

Challenges, Solutions & Future Enhancements

Key Challenges

We faced several challenges in building CrediSphere:

Inconsistent Credit Scoring: Standardizing data from different bureaus.
Bureau Downtime: Ensuring uninterrupted risk assessment.
Loan Approval Delays: Accelerating the evaluation process.
Large Financial Documents: Efficiently processing PDFs.

Our Solutions & Future Enhancements

Here's how we addressed these challenges and plan to improve CrediSphere:

Unified Credit Profile: Data normalization and aggregation.

Fallback API Selection: Seamless risk evaluation.

Real-Time Insights: ML-based risk prediction and LLM analysis.

RAG-Powered Chatbot: Targeted data retrieval from PDFs.

- **Future:**
  - More data sources
  - Enhanced ML model
  - Scalable cloud infrastructure

Team Contributions

Tanay Shah (Backend)

API Integration, Database, Authentication

Anirudhha Dalal (AI/ML)

XGBoost Risk Model, LLM Insights

Dhup Thumbadiya (Frontend)

UI, Forms & Reports

Jeet Patel (Frontend)

UX, visualization

Use Cases

- **Banks & NBFCs** – Automate **credit scoring** and **loan approvals** with real-time risk assessment.
- **Loan Marketplaces** – Standardize **multi-bureau credit data** for accurate borrower profiling.
- **FinTech Lenders** – Improve **borrower evaluation** and reduce **default risks** using AI-driven insights.

References

APIs & Data:

Credit Bureau APIs - cibil , Experian , Equifax

XGBoost ML Model

Normalization Methods - Pattern Recognition and Machine Learning

Links:

[GitHub Repo], [LinkedIn Post]