Iteration 2 Report: FinTech Chatbot Project

Pratham Pradeep Mahajan (002527465)
Ishit Arhatia (002010960)

October 27, 2025

1. Project Kickoff

Project Goals and Expected Outcomes

- Develop a domain-specific FinTech Chatbot capable of answering financial queries using LLM and NLP.

- Integrate a structured database for storing user interactions, financial data, and chatbot responses.

- Implement ETL pipelines to process and clean financial datasets for analysis.

- Generate SQL reports and analytics for insight into user behavior and financial trends.

- Deliver a working prototype demonstrating chatbot interactions with real financial data.

Project Scope

- **In Scope:** Database design, data acquisition (public datasets), ETL pipelines, NLP-based chatbot implementation, SQL reporting.

- **Out of Scope:** Paid APIs, proprietary datasets, extensive front-end deployment.

Key Deliverables per Phase

1. Database design: ER diagrams, relational schema, sample data population.

2. ETL & Data Processing: Scripts to clean, transform, and load financial datasets.

3. Chatbot Implementation: NLP pipelines including tokenization, intent classification, and sentiment analysis.

4. SQL Reporting: At least 5 advanced analytical queries.

Milestones and Deadlines (25-Day Timeline)

Phase	Tasks	Deadline
Database Design	ER diagrams, relational schema, sample data population	Day 5
ETL Pipelines	Data cleaning, transformation, loading into DB	Day 10
Chatbot Integration	NLP model implementation, LLM integration, testing	Day 15
SQL Reporting	8 analytical queries, dashboards	Day 20
Documentation & Final Review	Report, tracker, repository update, final testing	Day 25

Team Capability Assessment

- Both members possess strong database, SQL, and NLP skills.

- Gaps identified: fine-tuning LLM for finance-specific queries.

- Mitigation: use open-source resources, tutorials, and libraries.

Data Availability

- Public datasets from Kaggle, Yahoo Finance, stock summaries, and company filings.

2. Team Discussions

Core Skills and Contributions

Team Member	Core Skills / Contributions
Pratham Mahajan	Database design, SQL queries, ETL pipelines, Python scripting
Ishit Arhatia	NLP implementation, chatbot logic, LLM integration, Python/ML libraries

Tools and Technologies

- **Existing Experience:** Python, Pandas, NumPy, SQL, PostgreSQL, basic NLP.

- **To Learn:** Hugging Face Transformers, spaCy advanced modules.

Programming Languages and Platforms

- Python (primary for NLP, ETL, and chatbot logic)

- SQL / PostgreSQL (database management)

- Jupyter Notebook / VS Code (development environment)

Missing Skills & Challenges

- Fine-tuning LLM for domain-specific finance queries.

- Ensuring chatbot responses are accurate for financial terminology.

3. Skills, Tools, and Initial Setup

External Resources

- Open-source finance datasets, NLP libraries, Hugging Face models.

- Community forums, tutorials, and documentation.

Tools, Frameworks, and Libraries

- **Database:** PostgreSQL

- **ETL & Data Processing:** Pandas, NumPy

- **NLP & LLM:** Hugging Face Transformers, spaCy, NLTK

- **Version Control:** Git, GitHub

Development Environment & Version Control

- Python 3.11, PostgreSQL 15, VS Code as IDE

- Jupyter Notebook for prototyping

- GitHub repository with branching: `main` (stable), `dev` (active development)

Task Assignment

Task	Assigned To
Database Design	Pratham / Ishit (collaborative)
ETL Pipelines	Pratham
Chatbot NLP	Ishit
LLM Integration	Ishit / Pratham (reviewed collaboratively)
SQL Reporting	Pratham / Ishit
Documentation	Both