

Enhancing Business Analytics with LLM-SQL Integration

PROJECT ABSTRACT

BACHELOR OF ENGINEERING

Computer Engineering

SUBMITTED BY

Neeti Kurulkar
Komal Ambhore
Tejal Dhoot
Soham Sarade
Shrey Doshi

Under the guidance of
Prof. Dr. Mrs. S. A. Itkar



Department of Computer Engineering
P. E. S. Modern College of Engineering,
Pune.

2025-2026

Signature of Guide
A.Y.

1. Introduction

While large language models (LLMs) have demonstrated remarkable success in natural language generation and unstructured reasoning tasks, they continue to struggle with structured data access—particularly when it involves logical operations, relational schemas, and precision. In contrast, most enterprise-critical data such as profit and loss statements, balance sheets, and KPIs is stored in structured formats like spreadsheets or relational databases. This creates a bottleneck where non-technical finance teams must rely on analysts or BI tools for actionable insights

2. Purpose / Hypothesis (Problem Statement)

To develop a conversational interface that leverages large language models (LLMs) to translate natural language questions into executable SQL queries, allowing non-technical users to interact with structured financial databases and retrieve performance insights without requiring prior knowledge of SQL or BI tools.

3. Algorithmic Approach

- Accept user query in natural language via **React** frontend
- Extract database schema using **PostgreSQL** introspection (tables, columns, datatypes)
- Generate **SQL structure** using **LLM** with **keyword-constrained decoding**
- Fill in table/column names using **schema-aware few-shot prompting**
- Validate **SQL syntax** and execute on database using **safe parameterized queries**
- Feed results into **LLM-based Structured Reasoning Module** for explanation
- Output **human-readable summary** through the frontend **UI**

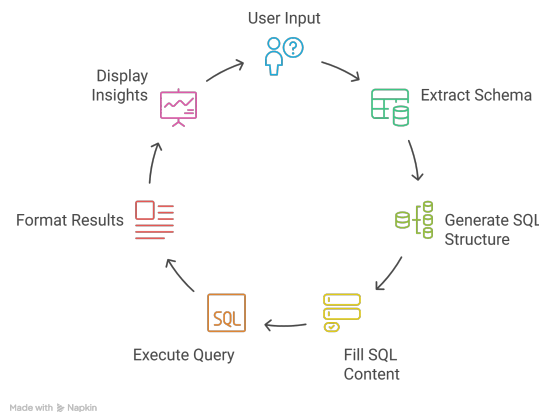


Figure 1: Algorithmic Approach

4. Block Diagram

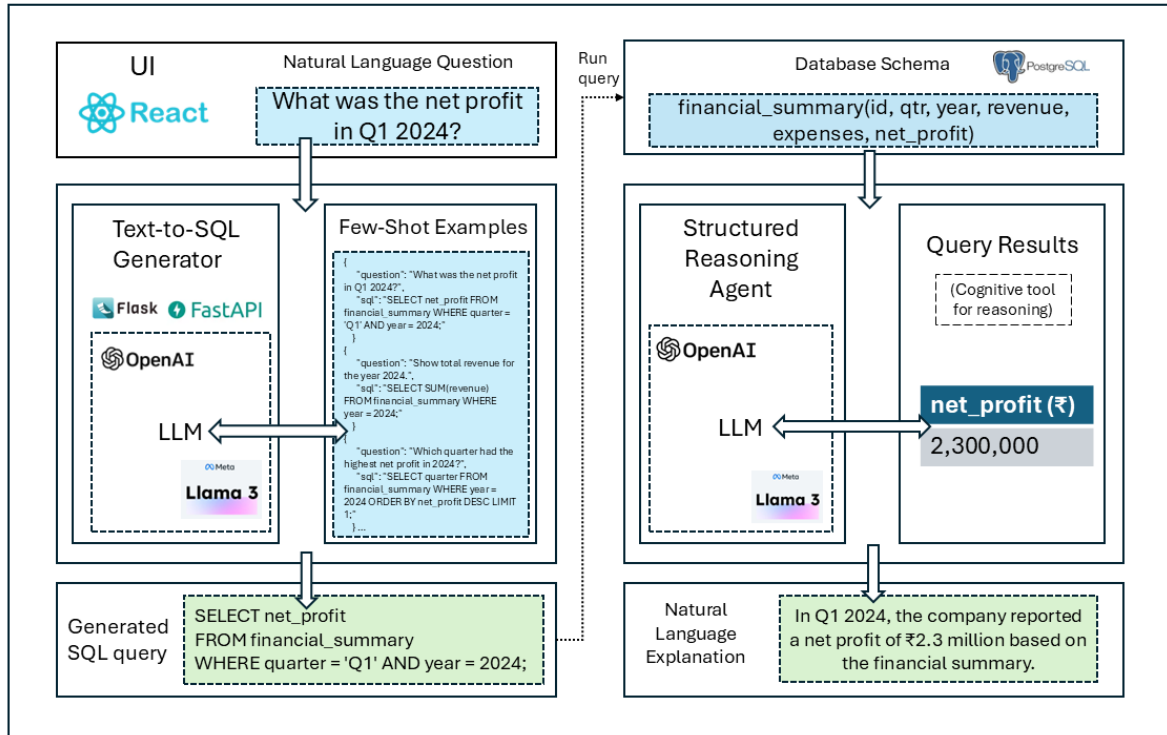


Figure 2: Architecture Diagram

5. Dataset / Input

Structured financial data such as:

Input:

- Profit & Loss (P&L) statements
- Balance Sheets
- All data is stored in a relational database format with a well-defined schema.

Dataset:

- Evaluated using the **BIRD benchmark dataset**, designed for testing Text-to-SQL accuracy on complex, multi-table queries.

6. Conclusions / Applications

- The system empowers non-technical users (e.g., finance, sales, HR) to interact with structured data using natural language.
- Enables real-time access to financial performance metrics without reliance on analysts or SQL knowledge.
- The system is scalable and can be extended to other business domains like HR, Sales, and Operations.