Natural Language SQL Querying with Llama 3, LangChain, and PostgreSQL

Tech Con 2025 Abstract 288

HPE Edge;

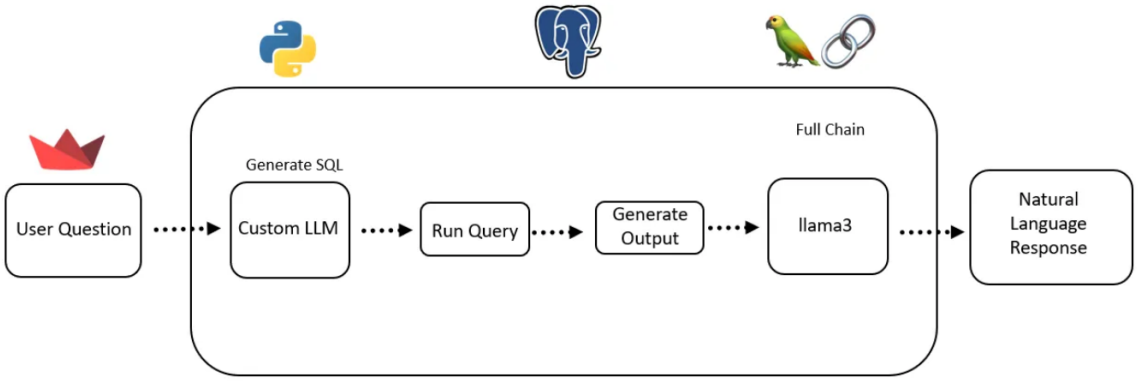
Abstract: The "Natural Language SQL Querying" project leverages the capabilities of Llama 3, LangChain, and PostgreSQL to develop a chatbot that allows users to perform SQL queries through natural language interactions. This innovative application transforms user queries from plain English into SQL commands to fetch data from a database, enhancing accessibility and ease of use for those without technical SQL knowledge.

Problem Addressed: Traditionally, querying SQL databases requires specific technical knowledge of SQL syntax, which can be a significant barrier for non-technical users. This limitation restricts access to data analysis and can delay decision-making processes in environments where quick data retrieval is crucial.

Unique Solution Overview: Our solution integrates a custom large language model (LLM) built with Llama 3 and hosted locally via Ollama, ensuring data privacy and security within the user's network. LangChain, an open-source framework, connects the LLM with PostgreSQL databases, automating the conversion of natural language queries into SQL commands.

Innovative and Unique Features:

- **Custom LLM Integration:** The use of a bespoke LLM for generating SQL queries from natural language inputs is central to our solution, allowing for precise and context-aware translations that enhance query accuracy.
- **Local Hosting of LLM:** By hosting the LLM in user/enterprise data center, we ensure all data processing remains within the user's controlled environment, significantly enhancing data security and privacy.
- **Streamlit Chatbot Interface:** Users interact with the database through a user-friendly chatbot built with Streamlit, which simplifies the query process and makes it as intuitive as having a conversation.
- **LangChain Automation:** LangChain automates the workflow from query generation to execution, streamlining the entire process and reducing the potential for errors.



Evidence the Solution Works: Preliminary testing and user feedback indicate that the system effectively translates natural language queries into SQL with high accuracy. Users without prior SQL knowledge were able to retrieve correct information from the database, validating the model's utility and effectiveness.

Competitive Approaches: Unlike many traditional SQL querying tools that require manual coding, our solution stands out by automating the translation process and simplifying user interactions. While some competitors offer natural language processing capabilities, they often require cloud-based data handling, which can pose security concerns. Our approach maintains all data processing locally, offering a significant competitive edge in terms of data security.

Current Status: The project is currently in the final stages of the development phase, with core functionalities implemented and being refined based on trial use. User testing has been overwhelmingly positive, with high satisfaction in the tool’s ability to simplify complex querying tasks.

Next Steps: The next phase involves extensive scalability testing to ensure the system can handle larger datasets and a higher volume of queries efficiently. Additionally, we plan to integrate more advanced natural language understanding features to cover a broader range of query types and complexities. Following successful scalability testing, the solution will be prepared for a wider rollout, targeting various industries that rely heavily on data-driven decision-making.