ChatDB: Interactive Database Query Assistant

Team Details

Group Category: Course Project ChatDB
Group Name: ChatDB 17

Team Members Background and Skills

Dhyey Desai:

I am a Master's student in Applied Data Science with a solid foundation in machine learning, data analytics, and deep learning. I have hands-on experience with Python, TensorFlow, and scalable infrastructure, along with strong proficiency in data management systems like SQL and Hadoop. My skills in natural language processing and advanced deep learning models have been essential in developing projects such as hate speech detection and medical image segmentation, and I am eager to apply these abilities to tackle complex challenges in dynamic environments.

Prem Doshi:

I am a graduate student specializing in Applied Data Science at the University of Southern California. I have a solid foundation in machine learning and data analysis, with extensive experience using Python, TensorFlow, and various data visualization tools. My expertise in developing advanced predictive models, integrating natural language processing techniques, and utilizing large language models (LLMs) will be crucial for enhancing data-driven insights and user experiences in innovative projects.

Project Requirements

ChatDB aims to create an interactive command-line application that allows users to query both SQL and NoSQL databases using natural language input. The system should:

1. Support connections to both MySQL (SQL) and MongoDB (NoSQL) databases.
2. Allow users to explore database structures and view sample data.
3. Process natural language queries and convert them into appropriate database queries.
4. Execute queries and display results in a user-friendly format.
5. Handle multiple query types, including aggregations, filtering, and joins.
6. Provide a simple and intuitive interface for database interactions.

The application should be flexible enough to work with user-provided databases and handle various query patterns. It should also be easily extendable to support new query types in the future.

Planned Implementation

Our implementation strategy for ChatDB includes:

1. Database Connectivity:
    - Implement connection methods for MySQL and MongoDB.
    - Allow users to input their database credentials.
2. Query Pattern Matching:
    - Define a set of query patterns using regular expressions.
    - Create templates for SQL and NoSQL queries corresponding to each pattern.
3. Natural Language Processing:
    - Develop a system to match user input against predefined patterns.
    - Convert matched patterns into executable database queries.
4. Query Execution:
    - Implement methods to execute queries on both SQL and NoSQL databases.
    - Handle differences in query execution between database types.
5. Result Presentation:
    - Create a function to display query results in a formatted table.
6. User Interface:
    - Develop a menu-driven command-line interface for user interactions.
    - Implement error handling and user guidance features.
7. Extensibility:
    - Design the system to allow easy addition of new query patterns and database types.

Team Responsibilities

Dhyey Desai:

- Database connection and interaction (MySQL and MongoDB)
- Query pattern implementation and natural language processing
- Query execution logic for both SQL and NoSQL databases

Prem Doshi:

- Command-line interface development
- User interaction flow and error handling
- Result formatting and presentation
- Documentation and testing

Both team members will collaborate on system architecture, query pattern design, and overall integration of components.

Timeline

Week	Milestone	Responsible
1	Project setup and architecture design	Both
2	Database connection implementation	Dhyey
3	Basic CLI structure and user input handling	Prem
4	Query pattern definition and matching system	Dhyey
5	Natural language processing implementation	Dhyey
6	Query execution and result presentation	Prem
7	Integration of all components	Prem
8	Testing and bug fixing	Both
9	Documentation and final touches	Both
10	Project presentation preparation	Prem

This timeline is subject to adjustment as the project progresses.