

RETAIL INDUSTRY CHATBOT

This is an LLM project based on Google Palm and LangChain. I have built a system that can talk to a MySQL database. The user asks questions in a natural language, and the system generates answers by converting those questions to a SQL query and then executing that query on the MySQL database. AtliQ Tees is a T-shirt store where they maintain their inventory, sales, and discounts in a MySQL database.

Project Structure:

- `main.py`: The main Streamlit application script.
- `langchain_helper.py`: This has all the LangChain code
- `requirements.txt`: A list of required Python packages for the project.
- `few_shots.py`: Contains few shots prompts
- `.env`: Configuration file for storing your Google API key.

Project Highlights:

- AtliQ Tees is a t shirt store that sells Adidas, Nike, Van Heusen and Levi's t shirts
- Their inventory, sales and discounts data are stored in a MySQL database
- We will build an LLM based question and answer system that will use following,
- Google Palm LLM
- Hugging face embeddings
- Streamlit for UI
- LangChain framework
- Chromadb as a vector store

Key Technologies

1. LangChain

- **Description:** LangChain is a Python library designed for chaining together components for building applications with language models.
- **Role in Project:** It serves as the backbone for integrating various NLP functionalities, like question-answering and document retrieval.

2. Google Palm (LLM - Large Language Model):

- **Description:** Google Palm is utilized as the Large Language Model (LLM) for natural language processing tasks.
- **Role in Project:** The primary language model for generating embeddings and providing answers based on the content of news articles.

3. Streamlit

- **Description:** Streamlit is an open-source app framework for Machine Learning and Data Science projects.
- **Role in Project:** It is used to build and deploy the user interface of the news research tool, making it interactive and easy to use.

4. FAISS (Facebook AI Similarity Search)

- **Description:** FAISS is a library for efficient similarity search and clustering of dense vectors.
- **Role in Project:** It is utilized for storing and retrieving document embeddings, enhancing the tool's ability to find relevant news articles quickly.

5. Environment Management (.env)

- **Description:** Environment variables are used for configuration and setting up important parameters like API keys.
- **Role in Project:** Secure handling of the Google Palm API key and other sensitive configurations.

Sample Questions:

- How many total t shirts are left in total in stock?
- How many t-shirts do we have left for Nike in XS size and white color?
- How much is the total price of the inventory for all S-size t-shirts?
- How much sales amount will be generated if we sell all small size adidas shirts today after discounts?