**Title: LLAMA3 Streamlit App Documentation**

**Project Overview:**

This project showcases the use of advanced NLP models (LLAMA3) to generate context-aware responses. The application uses LangChain for integrating LLAMA3 with a knowledge base and Ollama to generate both grounded and ungrounded responses. It uses FAISS and Sentence Transformers for efficient knowledge retrieval.

**Features:**

* **Ungrounded Response**: Generate a response using LLAMA3 without additional context.
* **Grounded Response**: Generate a response using LLAMA3 with context retrieved from a knowledge base.
* **Interactive UI**: A Streamlit-based interface for inputting queries and displaying responses.

**Loom Video Walkthrough:**

You can watch a video walkthrough of the LLAMA3 Streamlit App using this Loom video link:  
<https://www.loom.com/share/116917808943414896af219925769ce2?sid=25eb0291-f2ee-4084-9712-c9926cfdc027>

**Setup and Installation:**

* **Prerequisites**: Python 3.9 or higher, Docker (optional, for containerization)

1. **Clone the Repository**:  
   git clone https://github.com/yourusername/llama3-streamlit-app.git  
   cd llama3-streamlit-app
2. **Install Dependencies**:  
   pip install -r requirements.txt
3. **Run the Streamlit App**:  
   streamlit run llama\_streamlit\_app.py
4. **Access the App**:  
   Open a web browser and go to http://localhost:8501.

**Usage:**

1. Enter a query in the text input box.
2. Click the "Generate Response" button.
3. View both the grounded and ungrounded responses generated by the LLAMA3 model.

**Dockerization:**

1. **Build the Docker Image**:  
   docker build -t llama\_streamlit\_app .
2. **Run the Docker Container**:  
   docker run -p 8501:8501 llama\_streamlit\_app
3. **Access the App**:  
   Open a web browser and go to http://localhost:8501.

**Project Structure:**

bash

/your\_project\_directory

├── llama\_streamlit\_app.py # The Streamlit application code

├── Dockerfile # Dockerfile for building the Docker image

├── requirements.txt # Python dependencies

├── knowledge\_base.txt # The knowledge base file

└── README.md # Project README file

**Technologies Used:**

* **Streamlit**: Interactive UI for generating responses.
* **Ollama**: Python library for interacting with LLAMA3.
* **LangChain**: Framework for building language model applications.
* **FAISS**: Fast similarity search for knowledge retrieval.
* **Sentence Transformers**: Embedding models for semantic search.

**Contributing:**

Contributions are welcome! Please open an issue or create a pull request if you have any suggestions or improvements.

**Contact:**

For any questions or issues, please contact anuragsrivastava56789@gmail.com.

**Screenshots:  
  
A screenshot of a computer

Description automatically generated  
  
A screenshot of a computer

Description automatically generated  
  
A screenshot of a computer screen

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

A screenshot of a computer program

Description automatically generated