**Introduction:**

This project is focused on being guide for tourist who are visiting Egypt by developing flask based chat bot the can answer with brief about a famous tourist site and the timing of the events that are likely to happen there then gives a recommendation on what to do after the event.

**Key features:**

**1.Document embedding storage:**

This project utilizes ChromaDB to store the dataset that was embedded into vectors, ChromaDB is a database that is designed to handle embeddings so that the storage be scalable and efficient making it possible to retrieve a large chunk of embedded data.

**2.Query based text retrieval:**

The user can enter a query which will retrieve the relative data from ChromaDB, ChormaDB has the capabilities to find the highest ranked relative text chunks based on the users query , providing users with precise and contextually relevant search results.

**3. Generative Response Generation:**

To enhance the user experience, the application integrates Google's Gemini API for generating natural language responses based on the retrieved text passages. This feature allows for more interactive and informative search results.

Manual:

The user will go to the website and write about the place he wants to go in Egypt, the response will be a brief about the place and the time the likely events will take place and a recommendation on what to do after the events.

**Description of the data:**

The data was scrapped from a website and put in a pdf file, The data contains 55 places with a brief about each place and the activities that can be done there.

Preprocessing of the data:

1. pdf text extraction:

The text is extracted from each page of the pdf file which is then compiled to a single string of texts.

2.text splitting:

The extracted text is split into smaller, manageable chunks using sentence tokenization. This ensures that each chunk is of the optimal size for embedding and processing.

3. Generating Embeddings:

Each text chunk is passed through the Google Gemini API to generate embeddings. These embeddings capture the semantic meaning of the text, making them suitable for similarity searches and information retrieval.

4. Storing Embeddings:

aThe generated embeddings are stored in ChromaDB, a specialized database designed for handling embeddings. This allows for efficient storage and retrieval of large volumes of embedded text data.

5. Querying the Database:

Users can query the database to retrieve relevant passages based on their input. ChromaDB’s query capabilities are used to find and rank the most relevant text chunks.

6.Generating Responses:

To enhance the user experience, the application generates natural language responses based on the retrieved text passages using the Google Gemini API.

Refrence :

* https://www.thrillophilia.com/best-places-to-visit-in-egypt