



### →SOME IMPORTANT POINTS←

- Only the first query is used for webscrapping and next queries use the already webscrapped content to generate answers. So serper and BeautifulSoup is used only once for the first query.
- The whole webscrapped content is then sent to llama model. The model will return a list of many sentences which are important and collected from all the websites.
- This list of sentences are stored in FAISS vector database. This data can be used for faster retrieval of related data based on future queries in the conversation.
- Memory is used to store all the conversations done between user and the AI model.
- Based on the query entered on streamlit, first the top related content from webscrapped data is retrieved. This data along with query is passed to LLAMA model.also conversation history is also passed. Based on the query and related data, the model gives response.
- REMEMBER- based on first query only webscrapping is done and it is used for further responses.
- To start a fresh conversation, i.e to delete the already webscrapped content and want to webscrap about new topic enter "exit".
- After entering "exit" the database will be cleared. Now the new entered query is used to webscrap new content and further conversation will be based on the new query only.

## →STEPS TO FOLLOW←

→Step 1: create conda environment. open command prompt and enter:

```
conda create -p venv python==3.10
```

```
conda activate venv
```

→Step 2: install all packages present in requirements.txt

```
pip install -r requirements.txt
```

→Step 3: copy paste the api keys of serper and groqcloud

Go to <https://serper.dev/api-key> and copy the key and paste it in SERPER\_API\_KEY in .env file

Go to <https://console.groq.com/keys> and copy the key and paste in GROQ\_API\_KEY in .env file

→Step 4: change the file location of faiss\_index.bin according to your file location.

Go to app.py in streamlit\_app file, in line 5 give the location of faiss\_index.bin in faiss\_index\_file variable. Even though faiss\_index.bin is not present, please enter the location to your flask\_app file.

i.e- faiss\_index\_file=r"path to flask\_app\faiss\_index.bin"

→Step 5: open a new command prompt and run app.py in flask\_app file

```
cd flask_app → goes to your flask_app file loc
```

```
python app.py → runs the app.py in flask_app
```

→Step 6: open another command prompt and run app.py in streamlit\_app file

```
cd streamlit_app → goes to your streamlit_app file loc
```

```
streamlit run app.py → runs the app.py in streamlit_app
```

→Step 7: enter your first query to webscrap the content. This content will be used again and again for further conversations. So the first query should be your main topic for doing conversation. This step will take some time to process :)

→Step 8: now enter your query about main topic. Ask your questions related to the topic.

→Step 9: if you want a new conversation and stop the present topic. Enter “exit” to delete the previous webscrapped data and to again webscrap about the new query.