Building Human-Centered AI Applications

Homework 04 — An iSchool Chatbot Using RAG

Goal: Create a chatbot that answers useful questions about student organizations.

Instructions:

1. Copy Lab4.py content to a new file HW4.py and add HW4.py to your app's navigation
2. Create a vector database
    a. Use all the HTML files provided as part of your RAG
        i. Copy all the HTML pages from the zipped folder to your project area
        ii. Build a vector DB with those documents
            1. "Chunk" each document (create two mini-documents for each)
            2. Explain (in comments) which chunking method you're using and why you chose that method
    b. Create the vector DB file only if it does not already exist (this way you can run the app multiple times but only create it once)
3. Using Streamlit, accept a query (prompt/question) from the user
    a. Create a memory conversation buffer chatbot, storing up to the last 5 interactions
    b. Use the vector DB to augment the information provided to the LLM
    c. Use any LLM you want
4. Output the results/answers via a chat interface
5. Evaluate the chatbot with 5 questions
    a. Provide the questions
    b. Explain why you chose these questions
    c. Tell me how well the chatbot performed, given your configuration, and whether you would change anything now that you have tested it with your questions

Deploy the application

- Deploy to Streamlit Community Cloud
- Be sure to update requirements.txt as needed
- Make sure your application works as expected before submitting to Blackboard

Submit: The link to your application; HW4.py; your responses to step 5

After you have submitted your assignment to Blackboard:
Sign up for an API key from Open Weather Map

- NOTE: This must be done the night prior to class
- The API key takes a few hours to become active
- We will use this key for Lab 5