LLM

Description:

In this task, you will use a large language model (LLM) to generate answers to questions regarding data stored in a database. The objective is to develop a chatbot that enables the users to inquire about jobs, the parameters of the jobs, and the comparisons between jobs.

You will be provided with an OpenAI key and can use the openai library, langchain library, or any other library you wish.

substrata_interview.sqlite3

Familiarise yourself with the database above The main tables are:

- `substrata_api_fw3djob`

- `substrata_api_project`

- `substrata_api_jobparameters`

Tasks

1. Create a simple utility that allows the user to enter their questions, and returns a brief answer after querying the SQLite database.

2. The chatbot should be able to answer as many as the prompts below as possible.

3. Add memory to this chatbot so it infers the job and project that is being talked from the message history.

Prompts:

1. What are the names of all the projects that were run that begin with F?

2. How many jobs were run for the "Sensitivity-Kernels-2024" project?

3. Summarize the parameters that were used for the ET015_psN_topY_atypeN_rtypeN job.

   - We expect to see a summary of the parameters that were used, response should not include parameters that were not set (ie the ones that have value set to NULL).

4. Summarize the AWI parameters and their values for the job named "ET015_psN_topY_atypeN_rtypeN".

   - We expect to only see a summary of the parameters that begin with 'awi_', such as: awi_normfilt, awi_max_freq etc

5. What are the parameter differences between jobs with id "ebb4c99d7b944b1eaf7ac62f9181fe85" and "457207b0f11445b9889e5b04279e9f09"?

6. What are the parameter differences between jobs with names "E007_awi_NFN_amin_Y_r" and "ET007_awi_NFN_amin_Y"?