Concourse AI Take Home Exercise

Take-Home Assignment: Context Selection for Improved LLM Performance

This assignment evaluates your ability to design a function that selects the most relevant dataset for an LLM (Large Language Model) to answer a user's question effectively. Don't take more than 2 hours to build a solution, any work left to be done can motivate a discussion about next steps with your technical interviewer. On the call we'll focus on your implementation to solve this problem, the tools you chose (if any?), and why. Be prepared to discuss not only the structure of your approach but how those specific tools are functioning under the hood.

Datasets: You will be provided with two small datasets (only 30 rows each) in a CSV format. The datasets provide background info for very different kinds of questions.

- **shopping_habits.csv:** This dataset provides basic demographic data about consumers, and a score related to how much they shop / spend
- **weekly_searches.csv:** This dataset shows a weekly breakdown of search engine searches for three popular programming languages.

Task:

1. **Function Design:** Develop a Python function that takes a user question as input and analyzes it to determine which of these toy datasets is most relevant for the LLM to access when answering the question.
2. **LLM Integration:** Integrate your function with the LLM. This would involve:
    - Passing the selected dataset to the LLM in a format it can use, along with the user question and any additional information you feel will help it answer.
    - Extracting the LLM's response as a simple string.

Evaluation Criteria:

- **Function Logic:** Clarity and effectiveness of the function's approach for selecting the relevant dataset.
- **Code Quality:** Readability, efficiency, and proper documentation of your code.
- **LLM Integration:** Completeness and success of the integration with the LLM API
- **Example Responses:** Relevance and coherence of the responses generated by the LLM using the selected dataset.

Note: This assignment is designed to be completed within a reasonable timeframe (2 hours or less). The focus is on your approach and problem-solving skills rather than achieving perfect accuracy.