Please note

While you can use AI tools to help you complete this assignment, please make sure you don't just copy-paste the AI output.

If your aim is to get selected for this opportunity, please spend some time and energy doing this assignment properly. All the best!

You have applied for a part-time, paid internship at BeyondChats. During the internship, you will work closely with BeyondChats' Product Team to design LLMs-led solutions to complex problems that people in the industry face.

This assignment will help us get a better understanding of how well your skills fit our requirements. If you have any queries, please send us a message on Internshala.

All the best!

Your Task

Evaluate LLM responses' reliability.

Design an **LLM evaluation pipeline** (an actual python script) that would help you automatically test AI answers to user queries against the following parameters in real-time:

- Response Relevance & Completeness
- Hallucination / Factual Accuracy
- Latency & Costs

Input: 2 JSONs. One json contains the chat conversation, the other JSON contains context vectors fetched from the vector database against a specific user message - to be used for generating AI response.

Here are some sample JSONs for your convenience:
https://drive.google.com/drive/folders/1Uu0Jr7TW0Gb96uCJJ50DYJpg_tIUnJRa

Submission Guidelines

- A git repo which contains your code as well as a ReadMe file with the following details:
    - Local setup instructions
    - The architecture of the evaluation pipeline you have created.
    - Why did you decide to build the solution this way, and not some other way?
    - If we run your script at scale (millions of daily conversations between humans and AI), how have you ensured that the latency and costs for real-time evaluations will remain at a minimum?

- Make sure your git repo is publicly accessible so our team can review your submission.

Technologies to Use:
Feel free to use any tools or libraries or programming languages needed to accomplish this task.

Note:

- Please make the submission before the due date. **(Sunday) 14 Dec, 11:59 pm IST.**

- It is recommended to follow PEP-8 guidelines

- Your code is your property and we promise to not use any part of your submission unless we select you for this paid internship.

- This internship can also be extended into a longer-term internship or even a full-time position post internship completion.

ALL THE BEST!