

Technical Assessment and Final Interview Preparation

Congratulations on passing the first round of interview with Synapxe Data Science and AI team and you have been selected to proceed to the next stage of our hiring process:

1. Technical assessment
2. Final interview

Deliverables

Technical assessment

The objective of the technical assessment is to gauge your technical proficiency and critical thinking skills as an AI engineer. You would be given a problem statement below related to Generative AI. You are required to submit the following within 7 days from receiving this email:

- Develop the **working codes** for the back-end pipeline and provide the link to your **Github repository** with the following:
 - Document the activities, instructions on running the code and references in a clear and concise manner in a README.md file
 - Requirements.txt file for environment setup
 - Clean code base with comments at appropriate sections for easy review
 - Results folder to include sample output generated from the GenAI pipeline. Refer to the section on '*Sample Output*' below for the suggested format.
- If Postman is used to send and receive input/output, then provide a screenshot and sample of the fields required in the README.md file.
- Ideally, the codes should be able to run on any laptop/operating environment.

You are not required to develop the front-end application of the chatbot.

If you require deadline extension for this assessment, you can provide your reasons and write in to Synapxe HR point-of-contact.

Final Interview

The objective of the final interview is to provide the opportunity for you to present your implementation and analysis in a concise and convincing manner. This is to assess your ability to analyse complex problems, break them down into smaller components and develop innovative solutions. We would also be assessing your ability to explain complex concepts and translate technical value to business value. You are required to prepare for:

- Technical slides to be presented within 20 minutes, some suggestions (not limited to the following) include:
 - Outline your technical implementation and architecture
 - An exploration and analysis of various methods for each technical component, and its suitability to the specific use case on hand
 - Highlight any assumptions made
 - Results and evaluation
 - Areas for improvement

You do not need to submit this deck of slides before the interview. You can present the technical slides during the final interview, if you are selected after the technical assessment.

Problem Statement for Generative AI

A local healthcare company published multiple articles containing healthcare facts, information, and tips. It wishes to create a conversational chatbot that can address readers' concerns in natural language using information from the trusted articles and in the healthcare context.

The conversational chatbot should answer readers' queries using only the information from the published articles. Where appropriate, it should adopt an empathetic and understanding tone.

Requirements

Develop the back-end AI framework for a working conversational chatbot that could answer questions from an external data source, i.e. the set of websites provided in the 'Dataset' section below, using large language models. Here are some considerations:

1. Ability to answer readers' queries using only the information from the published articles
2. Devise a strategy to reduce LLM hallucinations
3. [**Optional Challenge**]: Assess the performance of the answers generated from the chatbot, given that there are no ground truth Q&A pairs provided to you.

Please refer to the 'deliverables' section above on the submission requirements.

Dataset

A list of 13 public Singapore-based URLs related to diabetes are listed below. These source website articles and/or PDFs containing information on diabetes that you should use to augment the conversational chatbot's responses.

URLs
https://www.ace-hta.gov.sg/docs/default-source/default-library/managing-pre-diabetes-(updated-on-27-jul-2021)c2bfc77474154c2abf623156a4b93002.pdf
https://www.ace-hta.gov.sg/docs/default-source/acgs/acg-t2dm-personalising-medications.pdf
https://www.ace-hta.gov.sg/docs/default-source/acgs/gdm---an-update-on-screening-diagnosis-and-follow-up-(may-2018).pdf
https://ch-api.healthhub.sg/api/public/content/8fcd2bf22c2f472c8efbe5c650e3969e?v=a8fc8264
https://www.healthhub.sg/a-z/diseases-and-conditions/diabetes-treatment-capsules--tablets
https://www.healthhub.sg/a-z/diseases-and-conditions/diabetic_foot_tts
https://www.healthhub.sg/a-z/diseases-and-conditions/diabetic-ulcer
https://www.healthhub.sg/a-z/diseases-and-conditions/diabetes-treatment-insulin
https://www.healthxchange.sg/diabetes/living-well-diabetes/diabetes-recommended-vaccinations-children-adults
https://ch-api.healthhub.sg/api/public/content/57349b96ccfe47319fb49d902a064022?v=6a601253&_gl=1*ok09oc*_ga*MTk3Mjg0Mjc5Ni4xNzA3Nzg5OTUw*_ga_VQW1KL2RMR*MTcxMDc1MzAxMS40LjEuMTcxMDc1NDQxNC4xMi4wLjA_
https://ch-api.healthhub.sg/api/public/content/70e56617c5b140139bb5cb37db3cdef0?v=a7dd48f7&

_gl=1*11zfakr*_ga*MTk3Mjc0Mjc5Ni4xNzA3Nzg5OTUw*_ga_VQW1KL2RMR*MTcxMDc1MzAxMS40LjEuMTcxMDc1NDQzMC42MC4wLjA.
https://ch-api.healthhub.sg/api/public/content/0aa3838b8eeb40e6b34a2340a65fdd03?v=eca50be0&_gl=1*1tq2sup*_ga*MTk3Mjc0Mjc5Ni4xNzA3Nzg5OTUw*_ga_VQW1KL2RMR*MTcxMDc1MzAxMS40LjEuMTcxMDc1NDQzMC42MC4wLjA.
https://ch-api.healthhub.sg/api/public/content/57c95e67895840888240d7322f00117c?v=f00ca9b4&_gl=1*1tq2sup*_ga*MTk3Mjc0Mjc5Ni4xNzA3Nzg5OTUw*_ga_VQW1KL2RMR*MTcxMDc1MzAxMS40LjEuMTcxMDc1NDQzMC42MC4wLjA.

Here are some sample reader queries:

1. What is gestational diabetes and how is it diagnosed?
2. What are some healthy eating tips for people with diabetes?
3. How can my outpatient bill for diabetes be covered?

Sample Outputs

A csv/excel file with some samples that you tried. The csv could look like this:

Reader Query	Information Retrieved from source	LLM Response
What is gestational diabetes and how is it diagnosed?
What are some healthy eating tips for people with diabetes?
How can my outpatient bill for diabetes be covered?
...

If there are any additional information that you deem necessary (such as evaluation metrics), you may also include them as additional columns.