

## **Critical discussion regarding the chatbot's model**

Alireza Rafiei

The utilization of large language models (LLMs) has been growing for various applications. In this assignment, I aimed to evaluate the performance of GPT-3.5 in addressing a model-based machine learning problem. To achieve this, the chatbot should first establish a pipeline to generate synthetic data, which would augment the available training data for developing a machine learning model. Subsequently, it should generate the necessary code for implementation.

However, since the chatbot cannot directly access or analyze the data, it remains unable to suggest a relatively optimal pipeline for a given problem. As a result, its proposed workflows are subjective and heavily reliant on user input. When users define the problem, the chatbot can typically offer a broad array of potential solutions, some of which the user may be unfamiliar with. Nonetheless, these solutions are often relatively basic, and the chatbot has yet to demonstrate the ability to provide more complex, sophisticated solutions.

I found the chatbot to be more effective in providing code snippets for different aspects of the problem when given specific input questions and asked to debug incrementally, rather than expecting a complete solution at once. Although the chatbot can usually generate relatively optimal code that addresses a given problem, its inability to access the data directly may lead to issues with data format and structure, causing potential errors in the provided scripts.

A prime example of this occurred during my trial. The chatbot accurately suggested machine learning models for my task but failed to identify the appropriate data structure required for inputting data into the model. Since we were not permitted to manually correct any part of the code provided by the chatbot, the issue remained unresolved, ultimately preventing the generation of any results.

Compared to our previous approach, the chatbot offered a more simplified workflow. In terms of generating synthetic data, we incorporated actual sounds from the hospital environment into the records, while the chatbot recommended basic signal processing techniques to produce synthetic data. With regard to the machine learning aspect, we had developed more sophisticated models, but the chatbot suggested more elementary methods.

To sum up, I believe that chatbots may not be highly effective in proposing pipelines or developing various models for complex problems. However, they can be incredibly useful in providing code snippets for specific issues defined by the user. This assistance can significantly increase coding speed, enabling users to explore multiple approaches to problem-solving more rapidly, without getting bogged down by memorizing various libraries and syntaxes. Therefore, I anticipate that in the coming years, the ability to identify appropriate approaches and workflows for a problem will become far more important than raw coding ability. As a result, students like myself should prioritize familiarizing themselves with diverse methods for solving problems.