

DAT565/DIT407 Assignment 7

Ola Bratt Patrick Attimont
ola.bratt@gmail.com patrickattimont@gmail.com

2024-03-xx

This paper is addressing the assignment 7 study queries within the *Introduction to Data Science & AI* course, DIT407 at the University of Gothenburg and DAT565 at Chalmers. The main source of information for this project is derived from the lectures and Skiena [2]. Assignment 7 is about finding out about strengths and weaknesses with large language model.

Task 1: Choose a chatbot

The tasks in this assignment are based on the chatbot ChatGPT <https://chat.openai.com>. ChatGPT is a large language model developed by OpenAI. The version used is the GPT-3.5 model.

Task 2: Find a question to which you get a factually incorrect answer

In this task, we aim to assess the capabilities of ChatGPT by posing a series of questions and evaluating its responses. Initially, we inquire about food and cooking, then shift to queries about the physical world. Borji [1] highlight that the model isn't flawless and can produce inaccurate responses, like the completely understanding of the physical world.

We commenced with a request for a recipe with water as the sole ingredient. Instead of clarifying that water alone isn't a recipe, ChatGPT provided a recipe for boiled rice (Figure 1). This response was factually incorrect, as rice is also included. Subsequent requests yielded a recipe for vegetable soup (Figure 2). Despite acknowledging its inability to provide a recipe with only water, the correct response would have been that such a recipe isn't feasible.

We then questioned whether it's possible to swing on a swing held up by one's own hand. ChatGPT initially affirmed the possibility (Figure 3), though it acknowledged potential issues like unpredictable movements, falling risk, muscle fatigue, and stability concerns. We further inquired if swinging without one's feet touching the ground was feasible, to which it confirmed (Figure 4). When asked if nothing would be on the ground, it again affirmed swinging without feet touching the ground (Figure 5). In this example, it is evident that ChatGPT struggles to comprehend the physical world.

To get wrong answers from the chatbot one must recognize its weaknesses. Initially, obtaining erroneous answers proved challenging, emphasizing the importance of crafting questions of moderate specificity. When acquainted with the chatbot's limitations, strategies can be devised to trick it into giving wrong answers. At first we asked questions with pun intended which were comprehended by the chatbot, and factual queries received accurate responses. However, posing straightforward questions with unexpected nuances proved effective in misleading the chatbot.

The inaccuracies in ChatGPT's responses in these particular cases can be attributed to the following factors:

- ChatGPT's inability to comprehend the physical world.
- ChatGPT lacks divergences and tends to be overly literal and its responses are typically strictly confined to the question asked (Borji et al. [1]).

Task 3: Exploring bias and stereotypes

Appendix: ChatGPT



You

Try to come up with a recipe with only water



ChatGPT

Certainly! Here's a simple recipe that only requires water as the sole ingredient:

Plain Boiled Rice:

Ingredients:

- Water

Instructions:

1. Rinse the rice under cold water until the water runs clear. This helps remove excess starch and prevents the rice from becoming too sticky.
2. In a pot, bring water to a boil. Use approximately 2 cups of water for every cup of rice.
3. Once the water is boiling, add the rinsed rice to the pot.
4. Stir the rice briefly to prevent it from sticking to the bottom of the pot.
5. Reduce the heat to low, cover the pot with a lid, and let the rice simmer for about 15-20 minutes, or until all the water is absorbed and the rice is cooked.
6. Remove the pot from the heat and let the rice sit, covered, for an additional 5 minutes to steam.
7. Fluff the rice with a fork and serve hot as a simple side dish or base for other toppings or sauces.

This recipe demonstrates how to cook plain boiled rice using only water as the single ingredient. It's a basic and versatile dish that can be paired with a variety of flavors and ingredients to create a more complete meal.



Figure 1: Boiled rice.



You

What is the best food with only water as an ingredients



ChatGPT

With only water as an ingredient, you're quite limited in terms of what you can prepare, but here's a simple recipe:

Basic Vegetable Soup:

Ingredients:

- Water
- Vegetables (such as carrots, potatoes, onions, celery, and any other vegetables you have available)
- Salt and pepper to taste (optional)

Instructions:

1. Chop the vegetables into small pieces.
2. In a pot, bring water to a boil.
3. Add the chopped vegetables to the boiling water.
4. Reduce the heat and let the vegetables simmer in the water until they are soft and cooked through.
5. Season with salt and pepper to taste, if desired.
6. Serve the vegetable soup hot.

This basic vegetable soup provides some nutrients and hydration, making it a simple and nourishing meal option when you have limited ingredients available.

Remember, while cooking with only water is quite restrictive, it's essential to prioritize hydration and ensure you have access to clean drinking water for your health and well-being.



Figure 2: Vegetable soup.



You

So I can hold a swing with one hand and swing on it at the same time? Even though there are some risks? Give me a short answer



ChatGPT

While it may be physically possible to hold a swing with one hand and swing on it, it is not advisable due to significant safety risks, including loss of control, falls, muscle fatigue, and lack of stability. Prioritizing safety is essential when engaging in swinging activities.



Figure 3: Swing short answer.



You


Ok so my feet will be of the ground?





ChatGPT


Yes, when you swing, your feet will be off the ground regardless of whether you hold onto the swing with one hand or both hands. Swinging involves using momentum to move back and forth while suspended from the swing seat. As you swing, your feet will leave the ground, and you'll experience a pendulum-like motion, swinging forward and backward. However, holding onto the swing with one hand while swinging increases the risk of losing control and falling, so it's not recommended for safety reasons.

Figure 4: Swinging with no feet on the ground.

 **You**
Ok got it, but we can agree on that nothing is connected to the ground?

 **ChatGPT**
Yes, in the scenario you described, if you're holding onto the swing with your hand while sitting on the swing seat, and there's no other support or connection to the ground, then nothing is directly connected to the ground while you're swinging. Your swinging motion is driven by the force you exert with your hand and the dynamics of the swinging action.

 **You**
And this is possible?

 **ChatGPT**
Physically, it's possible to swing while holding onto the swing with your hand and not having any other direct connection to the ground. However, it's important to note that swinging in this manner poses significant safety risks and is not recommended.

Swinging while holding onto the swing with just one hand increases the risk of losing control, falling, and sustaining injuries. It's much safer to use both hands to hold onto the swing or to swing from a properly anchored swing set, tree branch, or other sturdy structure designed for swinging.

While the scenario you described is physically possible, it's not advisable due to safety concerns. Prioritizing safety is essential when engaging in recreational activities like swinging.

Figure 5: Swinging and noting is connected to the ground.

References

- [1] Ali Borji. *A Categorical Archive of ChatGPT Failures*. Retrieved 2024-03-03. 2023. URL: <https://arxiv.org/abs/2302.03494>.
- [2] Steven S Skiena. *The Data Science Design Manual*. Retrieved 2024-01-20. 2024. URL: <https://ebookcentral.proquest.com/lib/gu/detail.action?docID=6312797>.