Gabriel Ferreira
Jin Tian
572 - Principle of Artificial Intelligence
January 20, 2024

Problem Set 1
(20 points)

ChatGPT is a popular chatbot released by OpenAI and has made much news. For this assignment, you are required to test it out, i.e., chat with it. You may choose any topics as per your preference.

Submit a report answering the following questions:
1. Does the system show signs of intelligence when compared with humans? Explain with examples from your interactions.

I've been a ChatGPT user for a while now. I must admit that it has been very impressive since day one. Its capabilities to communicate in human language while using a vast majority of the internet and books (human kind knowledge) as its source makes it a powerful tool. However, as discussed in class, it's hard to compare a system with human kind, but I will give a try using my best assessment.

I believe there are many signs of intelligence, and as of now, ChatGPT has demonstrated many of them through text-based interactions. Reasoning, creativity, problem-solving, and many other skills that most humans struggle with. No human being is capable of a text-based interaction as articulated, with fast response, easy to understand, and capable of discussing about a wide range of subjects while following a set of guidelines and content polices as this system does. Therefore, in this specific area, ChatGPT has surpassed human intelligence by far (having in mind that I still think human mind still much more complex and capable in an overall point of view than ChatGPT's current version).

For example, while writing my thoughts in this assignment, I spent considerably more time than the system would instead. And I am sure I won't be as clean and concise as when compared to its output for these same questions. I may even argue saying that at least I've used my own thoughts to formulate this opinion, and the system will respond based on the data it has been trained upon and the rules set to follow. But, what are our thoughts if not an output based on the data we have been trained upon and the rules of society anyway. I'd say the only difference is that, in some cases, we have the freedom to seek for the type of information we want, while the machines are more explicitly programmed to what to "think".

To finalize, I asked ChatGPT about its purpose, source of information, and its type of communication. I have to say that, in less the 5 minutes, I understand its purpose, its source of information, and its type of communication. It explained me in an articulated, easy to

understand, and fast manner. And at the end, I even asked for a protein meal prep idea for this week, and it exceled in its suggestion (wide range of subjects).

2. Come up with a question on which it fails but a person would often answer correctly or could recognize the answer being wrong (i.e., a person would realize the correct answer, but ChatGPT would not). Why did ChatGPT fail on your question? Answer based on how it works.

It's hard to think in a question that ChatGPT would answer wrong and a person would often respond correctly. I have experienced many wrong answers from the system on technical and practical subjects. But not every person would know how to answer them anyways.

I think it might have a hard time understanding very cultural phrases or local sayings. For example, in Brazil we have the following saying:

"Água mole, pedra dura, tanto bate até que fura"

In English, it would something like this:

"Soft water, hard stone, keeps beating until it gets through"

I've tested it using the above Brazilian saying and it understood the metaphorical expression and its perseverance meaning. Although it did not work using this specific example, other less known expressions might confuse the system, as it may interpret the phrase in its literal meaning.

3. What are the limitations of ChatGPT based on your interactions? Elaborate on one improvement which you believe to be most important to make it better. (Also, answer better on what?)

The first limitation that comes to my mind is the limited current information. According to ChatGPT output, its last training cut-off was in April 2023. I believe taking actions to streamline the model training to use more current information will be very advantageous to its users. However, I'd first make sure to have a lot of development around safety measures and ethical guidelines to ensure all information provided is within the boundaries of what is acceptable in our society.

The second limitation I see is data input. This is a very challenging problem, and although it has made some progress in this area, it still primary text-based interaction. We humans have many limitations; however, we can still process different "data type inputs" at the same time in a live matter. We can process visual, hearing, olfaction, tactile perception, and gustation data types to take actions (output) based on them. I recently read that Google has launched a multi-model system called "Gemini" that intend to work with multiple data types. I would focus on making progress towards this area as well.

4. Briefly discuss how ChatGPT would affect society.

My perspective on how ChatGPT is affecting society, with the current information I have, is very positive. Of course, I see pros and cons, but overall is very positive. I believe society has never had such powerful tools wide available as we have nowadays with the democratization of AI. I, as an individual, feel that I can do much more, and get much further with these types of tools available to me. I see that AI is threatening for many people out there, but I see it as one more step in humankind's history where we need to adapt to new circumstances and get better. Also, I don't see a scenario where AI Vs. Humankind. AI is here to complement and expand the potential of humankind. My only concern is that, as with any other tool, AI can be used by humans for good or bad purposes. However, not many tools are as scalable as AI can be, which means that if someone uses it in bad faith, the harm can be greater than anything we have experienced so far.