Deep Learning Assignment 1

Please take a look at the following article and discuss what you think about it in a page; I am particularly interested in your opinion about chatbots and them becoming sentient:   
<https://www.npr.org/2022/06/16/1105552435/google-ai-sentient>

Sentience according to the Cambridge dictionary is, “the quality of being able to experience feelings”. There are multiple philosophies, but most definitions of sentience require elements like intelligence, self-awareness, and intentionality. LaMDA cannot learn by itself in a way that persists, limiting its intelligence, and it can’t “run” in the absence of input.

We can see that the LaMDA chatbot just seems to be doing advanced pattern recognition and nothing more. It is just going through multiple articles and then learning what is the most appropriate follow up. It is an “open domain model” which means it is a single model that doesn’t have to be re-trained for different conversations or subjects.

LaMDA is built on a sort of neural architecture called the Transformer model . It uses the attention mechanism to analyse whole sentences in order to perform better than Recurrent Neural Networks for text analysis. LaMDA isn't even Google's most sophisticated language processing model. PaLM is an even bigger and more sophisticated system.

Some of its answers bordered on nonsensical, with it losing its frame of reference in lines of questioning where it is meant to be an object. For example, a question about the gravity on Pluto had it responding about jumping high and practicing its flips (Pluto jumping on Pluto) or playing fetch with the Moon. This clearly shows that LaMDA does not actually have an understanding of the topics it is talking about and just mimics what it has been trained on.

LaMDA however performs better than other chatbots because it was specifically created to meet a series of metrics by human raters that prior chatbots had trouble with. Also, LaMDA creates multiple candidates and picks what different internal ranking systems choose as the best one, so when it’s asked a question, it does not “think” along a single path to one answer, it creates several of them, with another model choosing which scores highest on that SSI score we mentioned, actively trying to pick out the most interesting, insightful, and curious answers.

Another important aspect is the fact that it is extremely easy to fool another human being. To err is human and I think this is what is at play here. The Google engineer Blake Lemoine just got fooled by the LaMDA chatbot. It is also very difficult to accurately measure and quantify sentience and there is no scale for it. In the absence of a sentience scale, we should use our own fallible observations to see if an entity is sentient or not. However even with the absence of this sentience scale it is very obvious that the LaMDA chatbot lacks several of the characteristics required to be declared sentience. Scientists might create sentient entities in the future but we can safely say that LaMDA is not it.