A BEGINNER'S GUIDE TO LARGE LANGUAGE MODELS

By: Luis Ramirez
(Mentor) Dr. Daniel Huang

Why AI is Important:

One of the many reasons is it could do our homework for us when we need help.
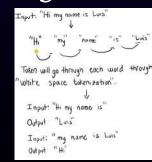
Input:

Output:

Machine Learning and AI:

Machine learning is when artificial intelligence gives the computer the ability to learn without being programmed. It is used in order to focus on the data and algorithms in order to make it have better accuracy.
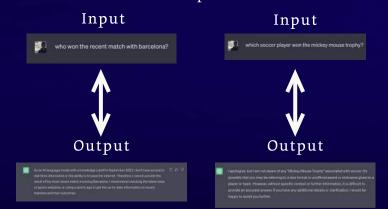
Language Model Background:

Machine learning is when artificial intelligence gives the computer the ability to learn without being programmed. It is used in order to focus on the data and algorithms in order to make it have better accuracy. There are also "tokens", these tokens are able to process the input of a user and predict a response.
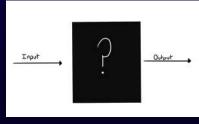
The Hallucination Problem:

Hallucinations can occur when an AI gives us a false response.

Input
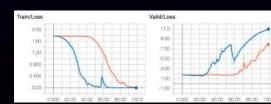
Output

Input

Output

Model's Are Blackbox Processes?:

In a black box we don't know how something works. However, we just know that it works. This is the same thing that happens with Artificial Intelligence, we know that it can give us an input and we'll get something back. One of the few ways we we can try to understand these models is by using prompt engineering. We can give these models a prompt and see the response we can get back with and without the prompt.

GPUs:

GPUs are essential for large language models due to their optimized performance for training AI models. Being able to have access to multiple GPUs will make it easier and faster to train models. The model shows that the orange line (single model) is having a more challenging time than the blue line (multiple GPUs). However, GPUs are very expensive and being able to buy multiple is a crazy amount.

Prompt Engineering (ChatGPT):

Video Game Example:

Simple Prompt: Complex Prompt:

Soccer Example:

Simple Prompt: Complex Prompt:

Prompt Engineering (Google Bard):

Video Game Example:

No Prompt: Prompt:

Soccer Example:

No Prompt: Prompt:

Conclusion:

In conclusion, we are starting to see large language models appear more and more. If we are able to understand how to use them effectively whether it would be with prompt engineering, in order for us to prevent hallucinations. However, there is still a lot more to research into these models in order to get a deeper understanding of what is going on.

References:
- https://arxiv.org/abs/2201.11903
- https://www.alignmentforum.org/s/yivyHaCAmMJ3CqSyj/p/LbrPTJ4fmABEdEnLf
- https://arxiv.org/abs/2201.02177
- https://arxiv.org/pdf/2212.07919v1.pdf