Large Language Models: Test Your Knowledge

Let's do a quick test! You must answer at least 4 questions correctly to pass this quiz.

1. How many 2-grams (bigrams) are present in the following phrase:

`they visited New York last week`

- ○ 3
- ○ 4
- ◉ 5
  - ✓ There are five 2-grams in the phrase "they visited New York last week": "they visited", "visited New", "New York", "York last", and "last week".
- ○ 6

2. Which attributes of large language models help them make better predictions than other types of language models? (Choose all that apply)

Choose as many answers as you see fit.

- ☑ LLMs contain many more parameters.
  - ✓ By definition, a large language model is a model that has a very large number of parameters, which is helpful in making better predictions.
- ☑ LLMs capture more context.
  - ✓ The large number of parameters in LLMs enable them to capture more context, which is helpful in making better predictions.
- ☐ LLMs don't need to be trained on as much data.
- ☐ LLMs never hallucinate.

3. True or False: A full Transformer consists of both an encoder and a decoder.

- ◉ True
  - ✓ A Transformer contains both an encoder (which converts text to an intermediate representation) and a decoder (which converts an intermediate representation into text).
- ○ False

4. An LLM is trained on a large corpus of data that includes the following example:

`My cousin's new fashion line is so cool!`

What mechanism helps the LLM learn that in this sentence, "cool" most likely means "great" and does not refer to the temperature of the clothing?

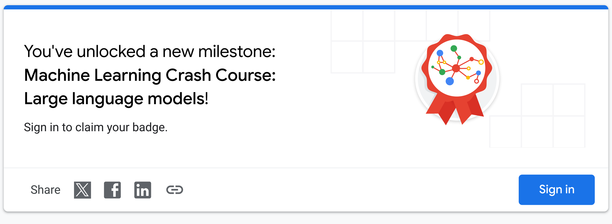
- ○ Prompt engineering
- ○ Decoder
- ○ Distillation
- ◉ Self-attention
  - ✓ Self-attention is the process LLMs use to weight the importance of relations between tokens in the input sequence. Self-attention enables LLMs to interpret words based on context, which in turn enables them to disambiguate in cases of multiple possible meanings.

5. Which of the following statements about fine-tuning vs. distilling is true?

- ○ Fine-tuning increases the number of parameters in the model, whereas distillation decreases the number of parameters in the model.
- ◉ Fine-tuning generally increases the quality of the model's predictions, whereas distillation generally decreases the quality of the model's predictions.
  - ✓ Fine-tuning updates the parameters of a model to improve its performance on a specialized task, improving prediction quality. By contrast, distillation's goal is to reduce the size of the model, which typically comes at the cost of some prediction quality.
- ○ Fine-tuning is performed on text models, whereas distillation is performed on image models.
- ○ None of the above are true.

Results

You scored 5 out of 5. Congratulations! You have passed this quiz.



Was this helpful?