QUIZ

One can be certain that a high probability outcome in a LLM is grounded in facts. TRUE or FALSE?

TRUE

FALSE

👏 Correct! One must always verify that the output of a LLM is grounded in facts. A high probability does not indicate that the outcome is factually grounded. It is more reflective of where there might be consensus in the text that it's trained on. Show less

Which of the following events are important in the history of Large Language Models (LLMs) and development of text-based Generative AI?

Alan Turing's 1950 paper on "Computing Machinery and Intelligence".

The 1956 Dartmouth Summer Research Project on Artificial Intelligence that sought to define what AI meant.

All of the above!

👏 That's correct!

The creation of ELIZA, the first digital chatbot, by Joseph Weizenbaum

Fill in the blanks in the following paragraph about temperature:

If we increased the temperature, we make the probability distribution wider , allowing for lower probability outcomes to become more possible. The LLM output becomes less deterministic, i.e., different attempts with the same prompt gives us different responses.

👏 You got it!

Which of the following tasks can a language model can be used for?

- Sentence translation
- Text summarization
- Question answer
- All of the above

👏 Yes, this is correct!

Choose the statement among the following that is **most likely** to be the outcome of a LLM output where temperature was set to be the highest value possible:

- The puppy is playing with his ball.
- The puppy is being walked to the park.
- The puppy is barking at the neighborhood cat.
- The puppy is writing poetry.

👏 Yes, this is correct! Of all the statements, this is the most fantastic and possibly absurd outcome.

Fill in the blanks in the following text about moving beyond count-based language models:

To get a language model to generalize , we would need to move beyond counting words and instead, find a way to map words to their meaning and context. Consider a sentence like "The monkey ate spaghetti."

For instance, suppose a language were able to map “monkey” to the idea of animals/mammals and “spaghetti” to food. This would allow it to use the connection that animals eat food and extend that to assigning a non-zero probability for monkeys eating spaghetti. This is often known as a “ semantic representation.”The mathematical objects that allow language models to generalize this way are known as word embeddings .

👏 You got it!

Fill in the blanks in the following text:

The downside to compression in LLMs is that there is information loss , i.e., language models can miss details in a text corpus. The downside to generalization is that language models can sometimes make up unverifiable text that has no factual grounding, something that is referred to as “ hallucinations ”.

You got it!

Which of the following statements about detecting patterns in text is FALSE?

To mathematically model language, one needs to find a way to turn text into math. The only way to do this is to consider the smallest unit of text to model to be a word.

Correct, this is not true! The smallest unit of language to model can be letters, words or tokens which can be a collection of words or even subwords.

Word embeddings are a way to map text into a mathematical model where aspects of the meaning and context of text are retained.

Count-based approaches to detecting patterns in text have been popular for over a century. These approaches, while being mathematically simpler get computationally inefficient as the volume of text they’re modeled on increases.

The field of Natural Language Processing (NLP) involves finding mathematical representations of language to capture statistical regularities in text. There are many techniques to do this and some are more computationally efficient than others.

The type of neural network that’s used in the GPT language models is:

CNNs (Convolutional Neural Networks)

Transformer

Yes, this is correct! Transformers are a kind of neural network architecture developed in 2017, perform great at language-related tasks and represent the “T” in GPT.

LSTM (Long Short Term Memory networks)

RNNs (Recurrent Neural Networks)

Which of the following statements comparing count-based and neural language models is FALSE?

The size of count-based models increases with the size of the training data and the same is true of neural language models as well.

They're both autoregressive, i.e., they both predict the "future based on the past".

If a specific sequence of words has never appeared in a corpus of text, both language models will assign a zero probability to the likelihood of this sequence appearing.

Correct, this statement is FALSE! Neural language models can generalize to unseen instances while count-based language models cannot.