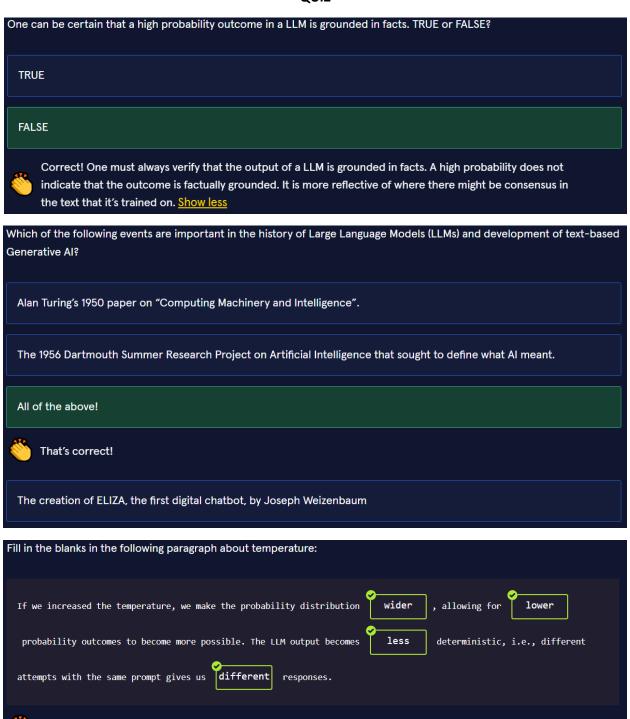
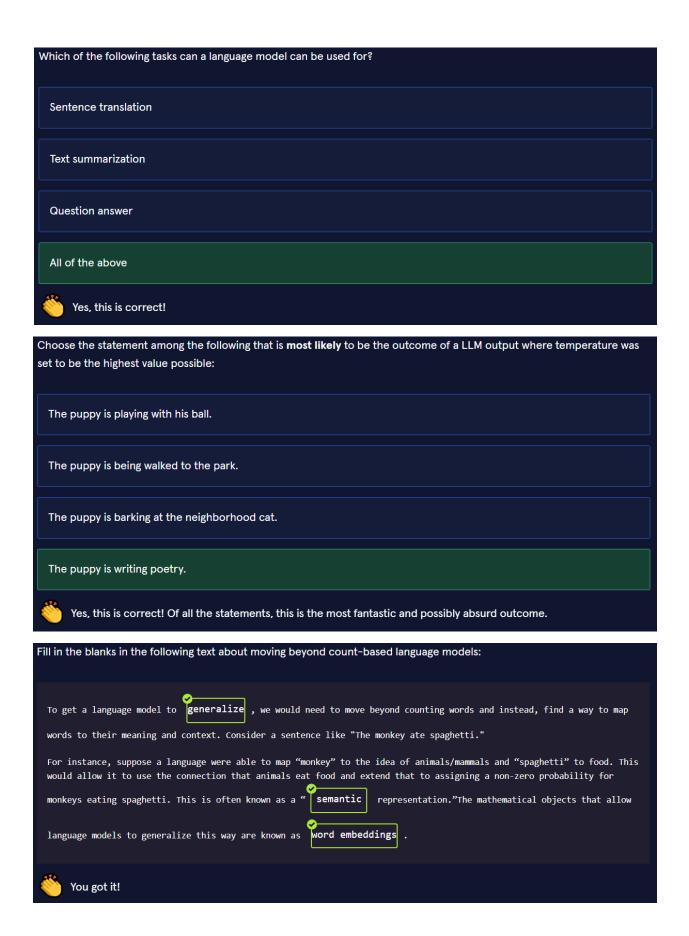
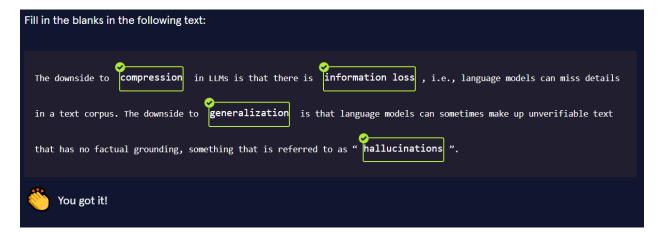
QUIZ



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.