

LLM (Large language models) intro

- LLM are a Gen AI tool and fall under deep learning part of Machine learning, they are next word predictors

Give it a piece of text and it predicts one word,

the text in that piece of text

the dog ^{text} jumped

• it is not explicitly programmed

over the Ledge

it learns from training data which

next word

is millions of pieces of text

predicted by LLM

where it has known next words it

then learns from this and is able to predict next words given an unknown piece of text

• a LLM is defined as a complex mathematic function

that can predict what word comes next for any piece of text

* in reality instead of predicting one word with certainty it predicts

Ex: Paris is a city in

many possible next words and gives

ans: France 17%
and 15%
the 9%
Loren 7%
⋮

each one a probability. the highest is the answer.

- in a LLM you add context like in a chatbot for ex

then you take the users prompt and add it onto the

context and input that to the Ex: context: You are a helpful chatbot

LLM then the model repeatedly prompts what to do in Paris?

predicts the next word after this LLM: Sure, there are plenty of ... etc

you append that word to what you

1) context + prompt = Sure

have so far so into the start that context

2) context + prompt + Sure = there

+ prompt and repeat this until

• each output is probability of possible next word

you reach a stop word and stop. 44

etc etc