

LLM - [Large Language Model]

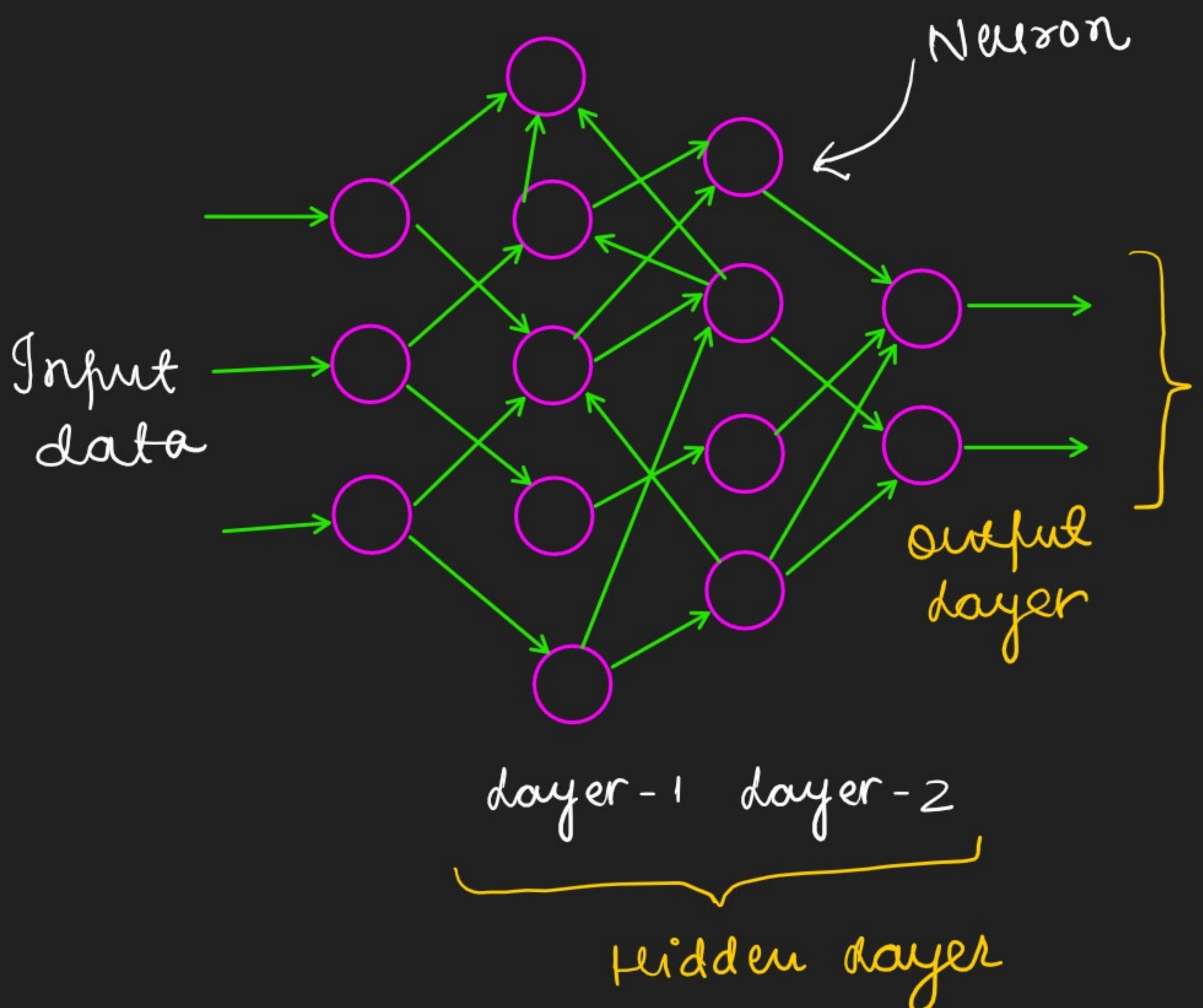
Lecture - 2: Intro to LLM

1. what is large language model?

Neural network designed to understand, generate and respond to human like text.



Deep neural network trained on massive amounts of text data

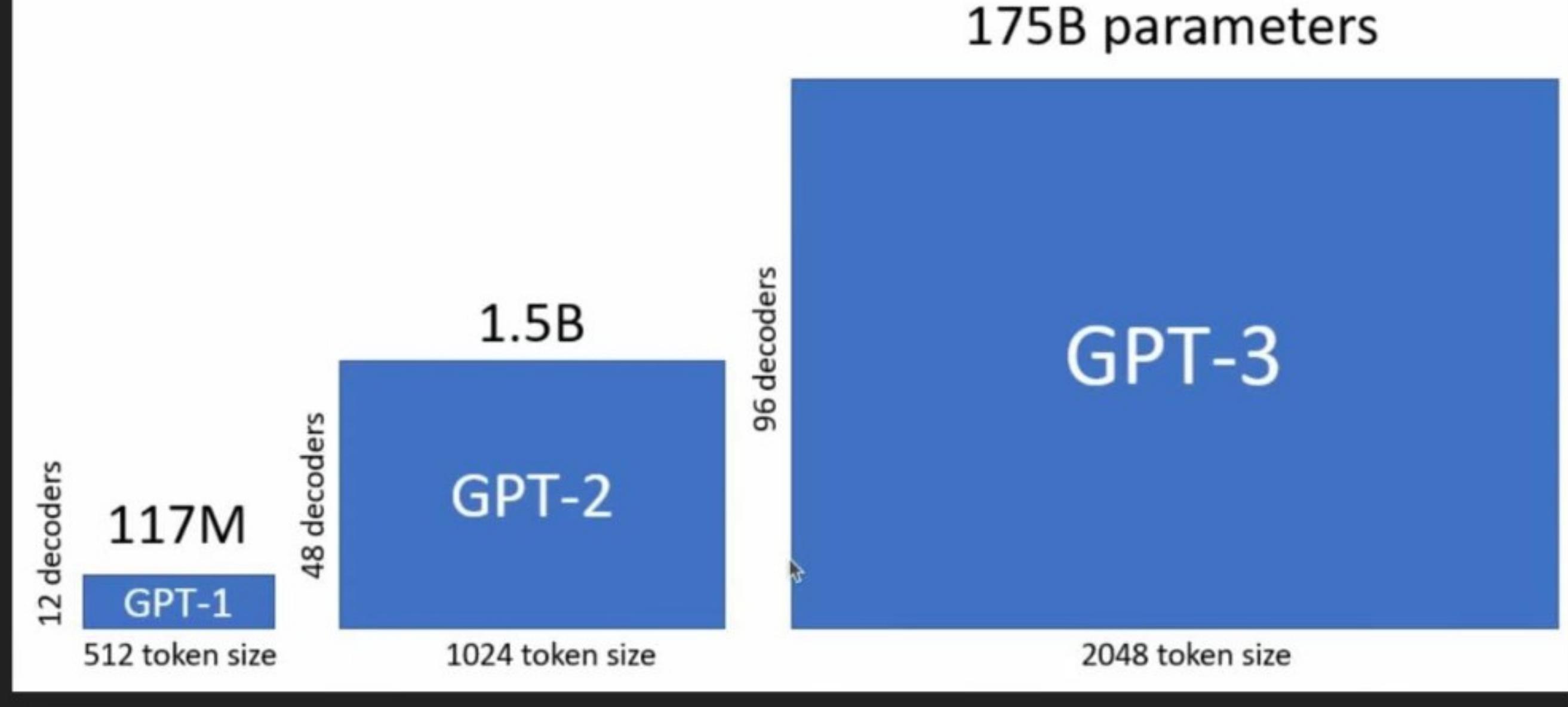


2. Large Language Models (LLMs)

models have billions of parameters

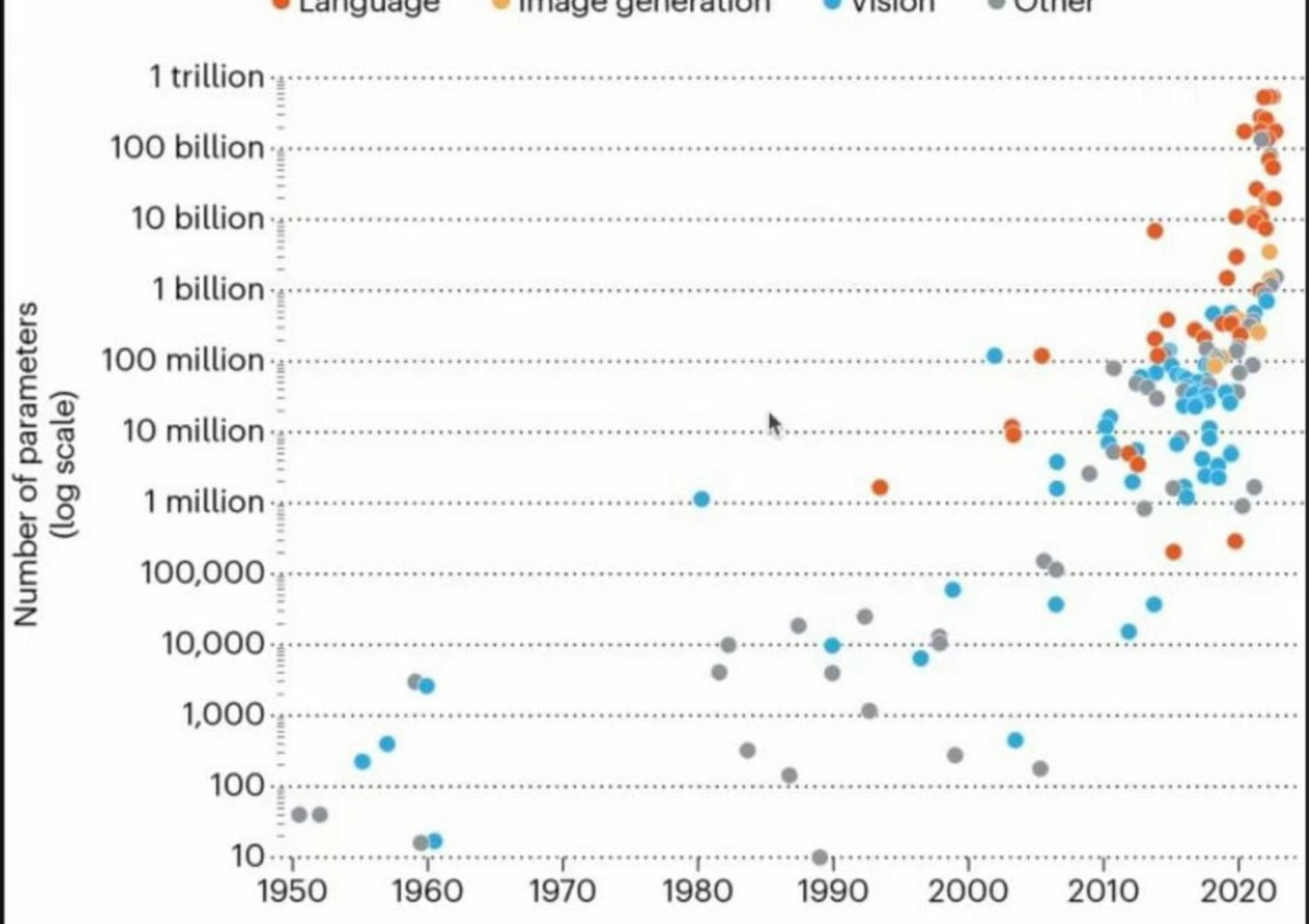
These models do a wide range of NLP tasks: Question answering, translating, sentiment analysis and much more!

Model Name	n _{params}	n _{layers}	d _{model}	n _{heads}	d _{head}
GPT-3 Small	125M	12	768	12	64
GPT-3 Medium	350M	24	1024	16	64
GPT-3 Large	760M	24	1536	16	96
GPT-3 XL	1.3B	24	2048	24	128
GPT-3 2.7B	2.7B	32	2560	32	80
GPT-3 6.7B	6.7B	32	4096	32	128
GPT-3 13B	13.0B	40	5140	40	128
GPT-3 175B or "GPT-3"	175.0B	96	12288	96	128



THE DRIVE TO BIGGER AI MODELS

The scale of artificial-intelligence neural networks is growing exponentially, as measured by the models' parameters (roughly, the number of connections between their neurons)*.



*"Sparse" models, which have more than one trillion parameters but use only a fraction of them in each computation, are not shown.

3. [LLMs] vs Earlier [NLP] models

can do wide range of NLP tasks

Designed for specific tasks like language translation etc.

"Earlier language models could not write an email from custom instruction a task that is trivial for modern LLMs."

4. what makes LLMs so good ??

Transformer Architecture

This is my transformer

It actually look something like this

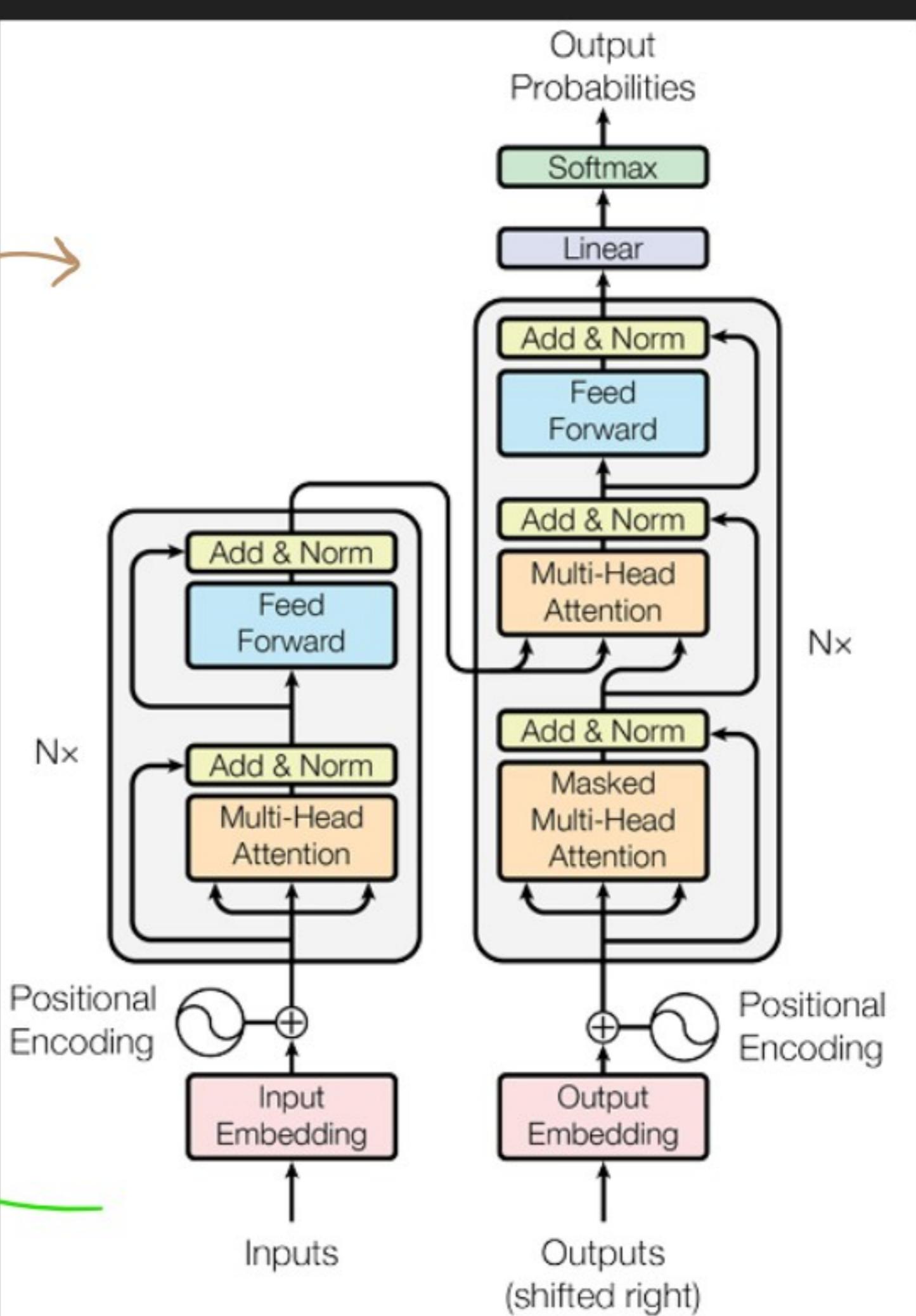
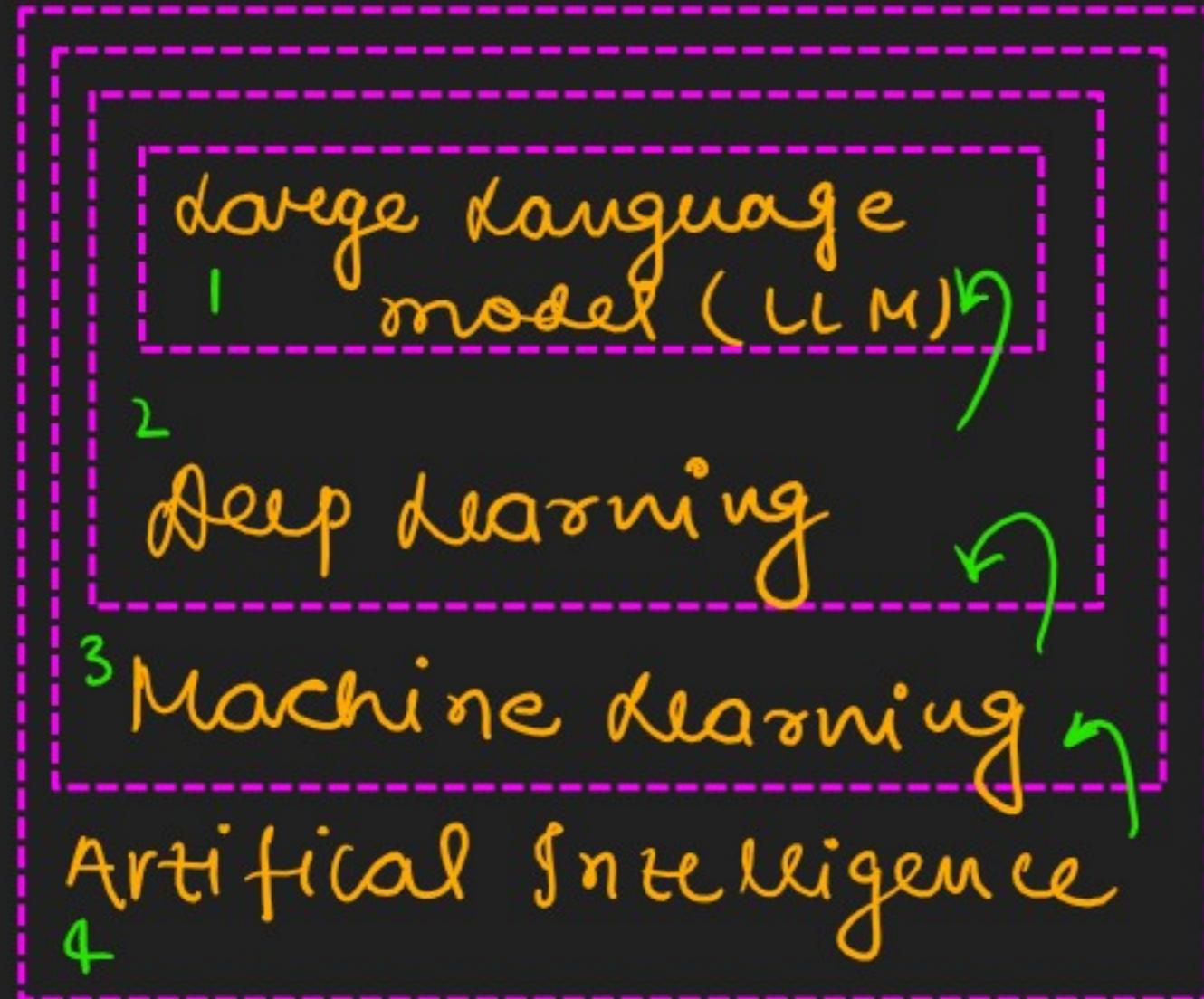


Figure 1: The Transformer - model architecture.

5. LLM vs GenAI vs Deep learning vs Machine learning



→ Big umbrella

1 + 2 → LLM + DL → Generative AI

4 > 3 > 2 > 1

AI > ML > DL > LLM

LLM represent a specific application of deep learning technique, leveraging their ability to process and generate human like text.

using deep neural network to create new content such as text, images, various forms of media.

6. Applications Of LLMs

- Chat-bot virtual assistant.
- Machine translation.
- Novel text generation.
- Sentiment analysis
- Content creation