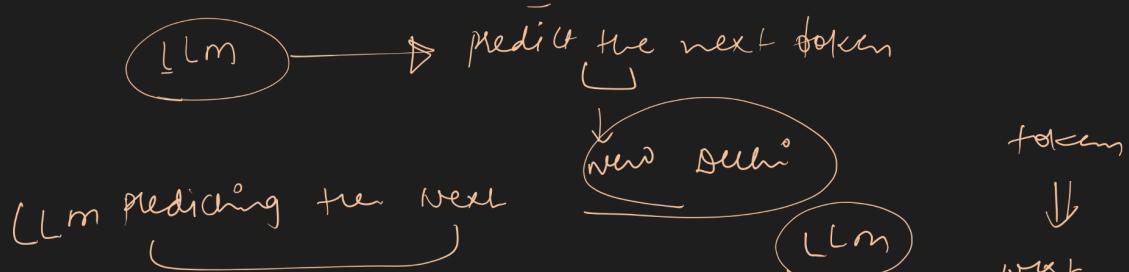


* LLM Foundations -

↳ If you don't understand how LLM works,
you'll struggle with

- {
 - ① Prompting
 - ② RAG
 - ③ Evaluation
 - ④ Production Reliability

* What problem does an LLM solve :-



$$\textcircled{1} \text{ member} \rightarrow x$$

$$\textcircled{2} \text{ New Delhi} \rightarrow y$$

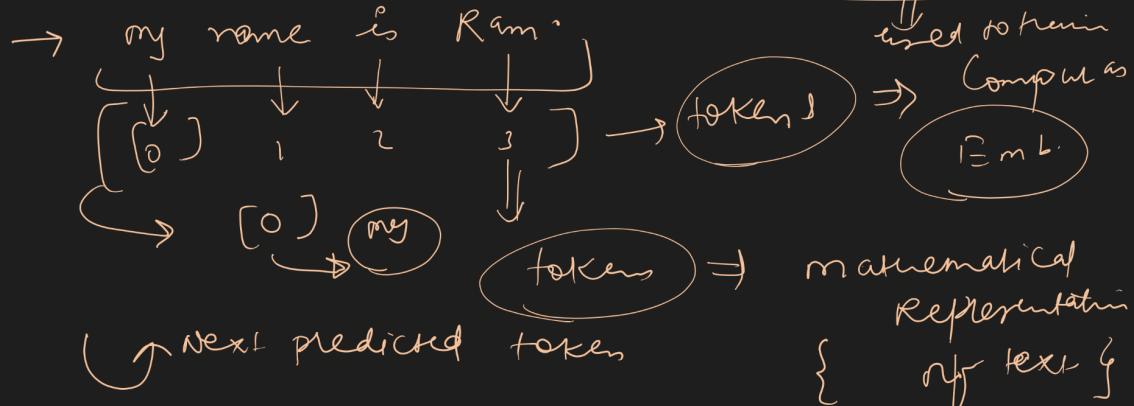
$$\textcircled{3} \text{ Jaipur} \rightarrow z$$

$$\left\{ \begin{array}{l} y > x \\ y > z \end{array} \right\}$$

* Tokens - How text becomes number

what are these tokens

⇒ Random value to
these text



* Transformers & attention -

↓ attention based Arch -
DL based principle

⇒ Cat

RNN → GRU

↓ sequential learning

This is a cat, it's color is white.

↑
token was → the other token
are related to it

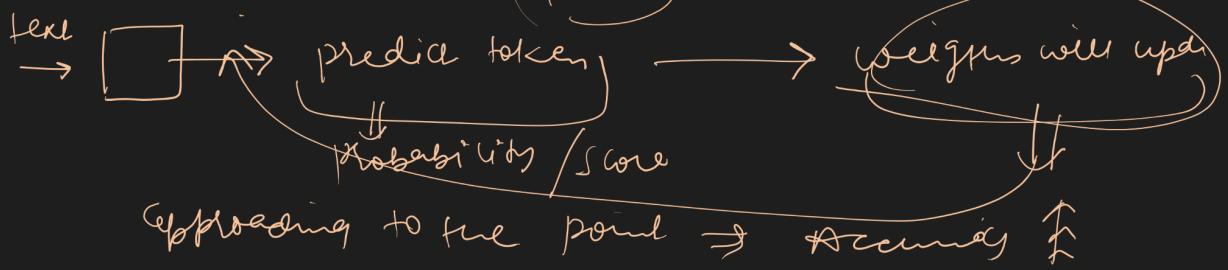
* How LLMs are trained:-

Training source can be:-

{ ↳ Book pdf / Any other file
↳ blog

↳ Conversations [Audio files → Speech
podcasting
E = $y_p - y_t$]
↳ contacts]

training loops.



* Hallucination in LLMs:-

Context is missing / limiting context

No training happens

