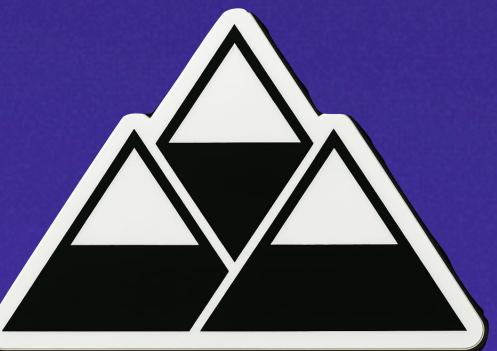


# Agentic AI

## From Pretraining to Building Your First Agent

EL MAJJODI ABDELJALIL



# Who AM I?

- EL MAJJODI Abdeljalil
- ML Engineer AtlasA / Sawalni
- MSc Distributed System & AI
- Open Source Contributor



# Outline

- Foundation of Large Language Models

- Introduction to LLMs
- How LLMs are Trained
- Limitations of LLMs

- Introduction To Agents

- What is an LLM Agent
- LLM Memory
- Tools
- Agent Planing

- Build Your First Agent

- Frameworks
- Hands-on SmolAgent

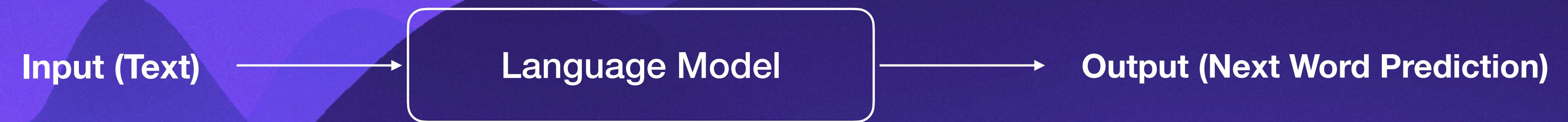


# Foundation of Large Language Models



# What are Large Language Models?

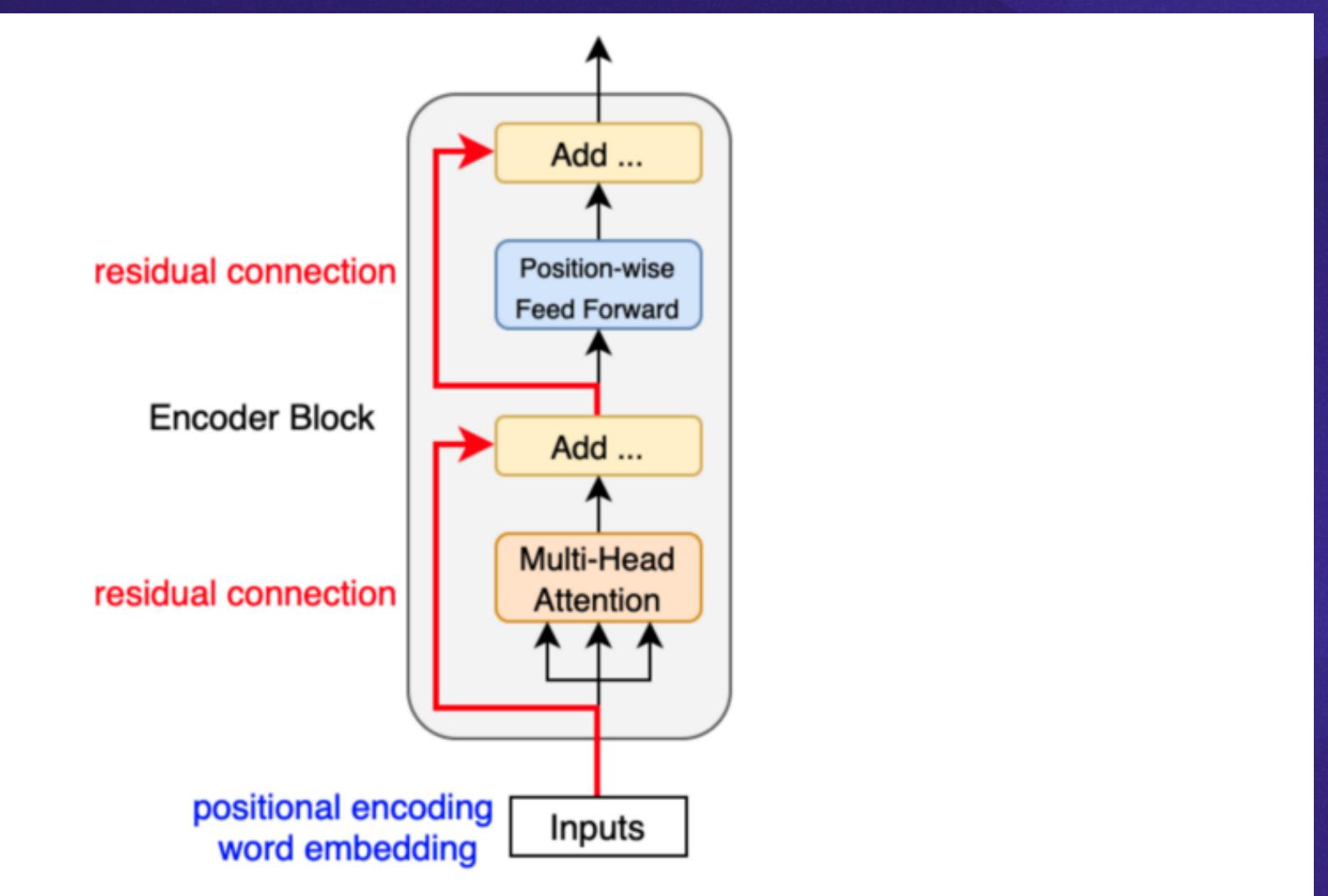
- ◆ Machine Learning Model predicts the next word given input.



**Example:** Morocco is the best African ...

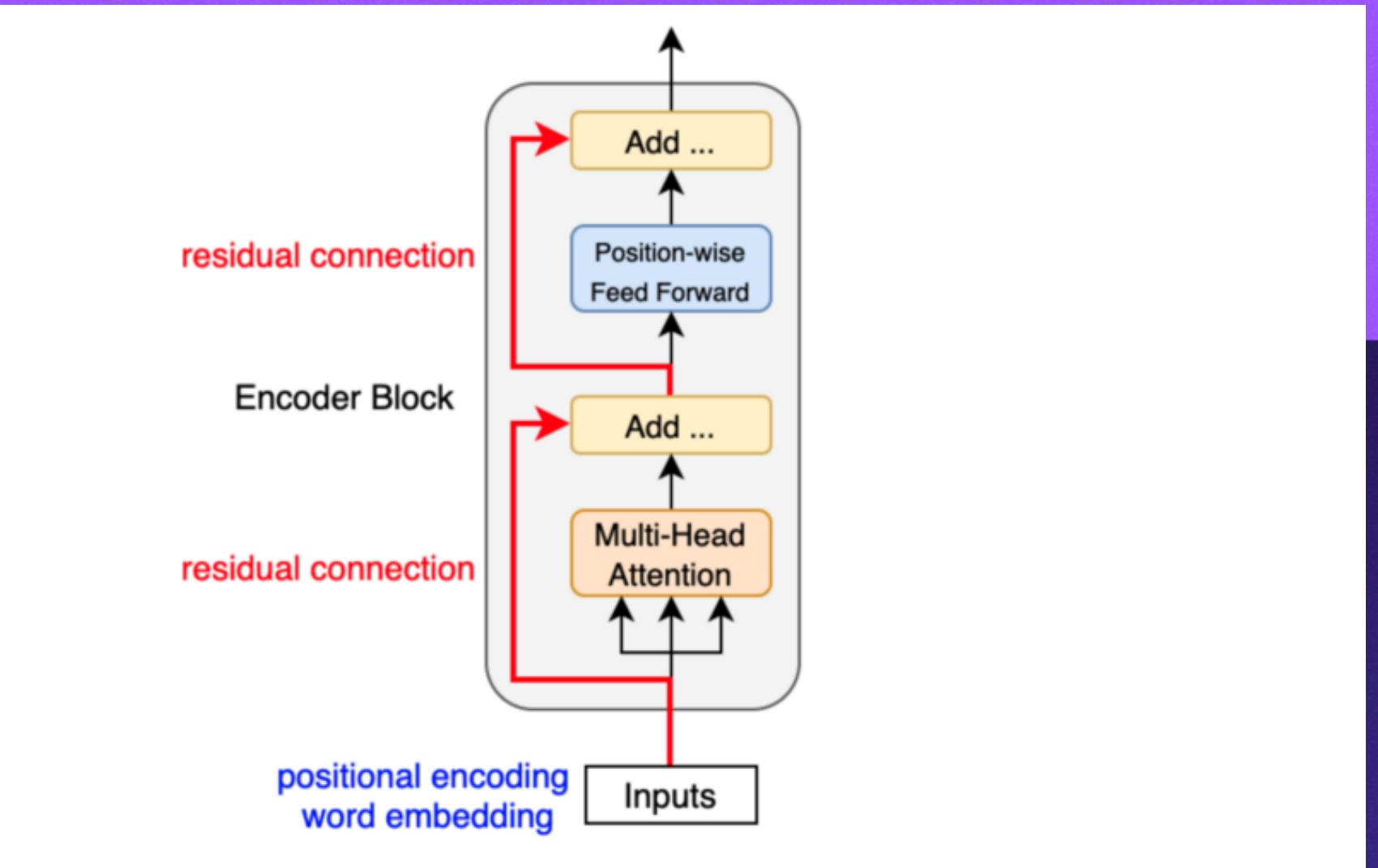


```
● ● ●  
str predict_next_word(str input){  
    //body  
    return next_word;  
}
```



Transformers Architecture (decoder)





## Transformers Architecture (decoder)

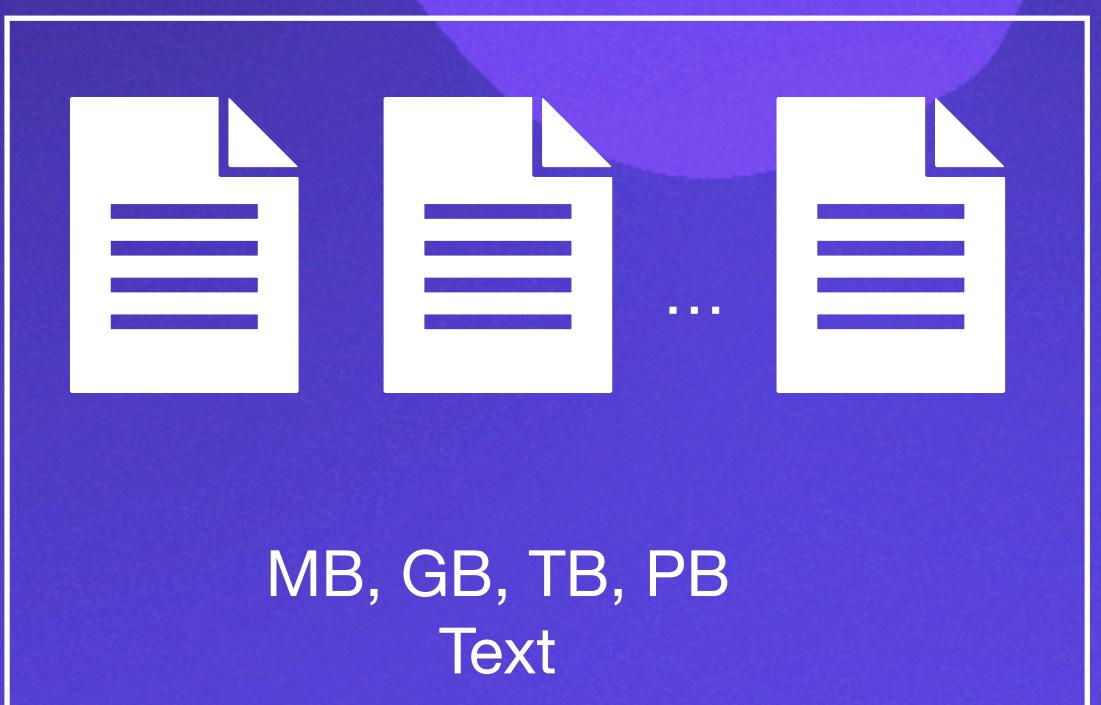


# How LLM Predict Next Word?

Data Preparation



**Collect Data from many sources**  
(Books, Web,...)



**Concatenate All Text**



An LLM, or large language model, is a machine learning model that can comprehend and generate human language

[...]

**Split into the same length**

An LLM, or large language model, is a machine learning

Line 2

...

Line n



An LLM, or large language model, is a machine learning



Input	Output
An	LLM



An LLM, or large language model, is a machine learning



Input	Output
An	LLM
An LLM	Or



An LLM, or large language model, is a machine learning



Input	Output
An	LLM
An LLM	Or
An LLM or	Large



An LLM, or large language model, is a machine learning



Input	Output
An	LLM
An LLM	Or
An LLM or	large
An LLM or large	Language



An LLM, or large language model, is a machine learning



Input	Output
An	LLM
An LLM	Or
An LLM or	large
An LLM or large	Language
...	...
An LLM or large ...	Learning



**Line n**



Input	Output
An	LLM
An LLM	Or
An LLM or	large
An LLM or large	Language
...	...
An LLM or large ...	Learning
...	...
...	...





Unstructured Data



Input	Output
An	LLM
An LLM	Or
An LLM or	large
An LLM or large	Language
...	...
An LLM or large ...	Learning
...	...

Structured Data



# How does LLM Predict Next Word?

## Model Pretraining (self-supervised learning)



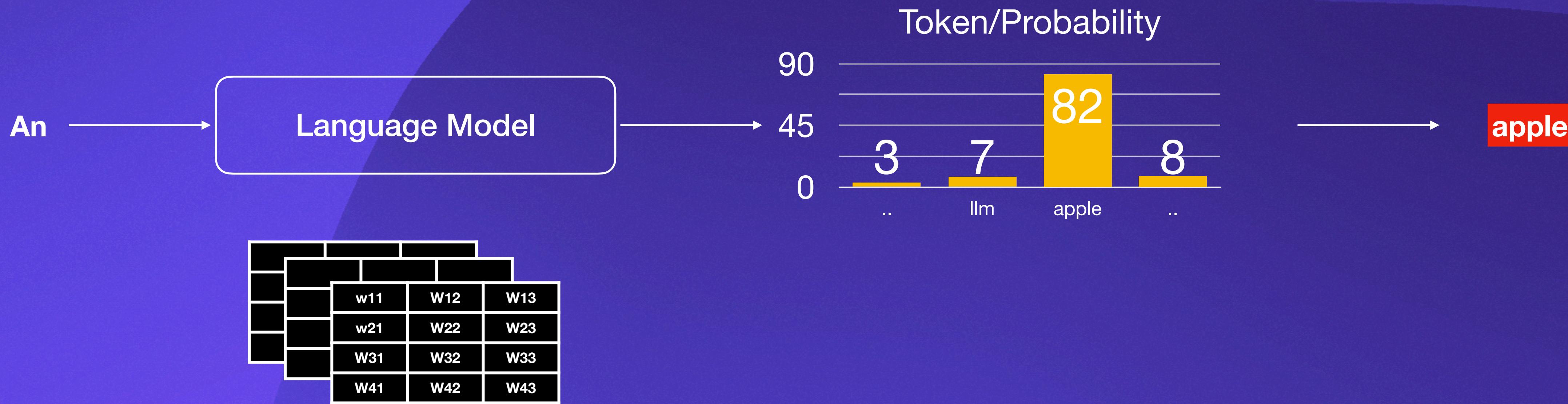


Initialize  
with Random Params

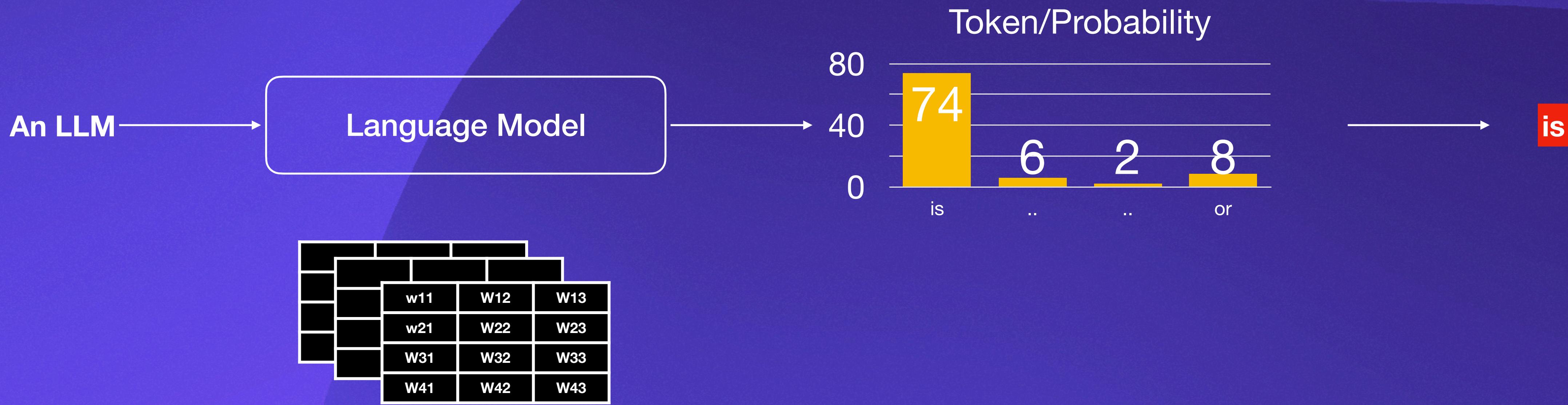
Language Model



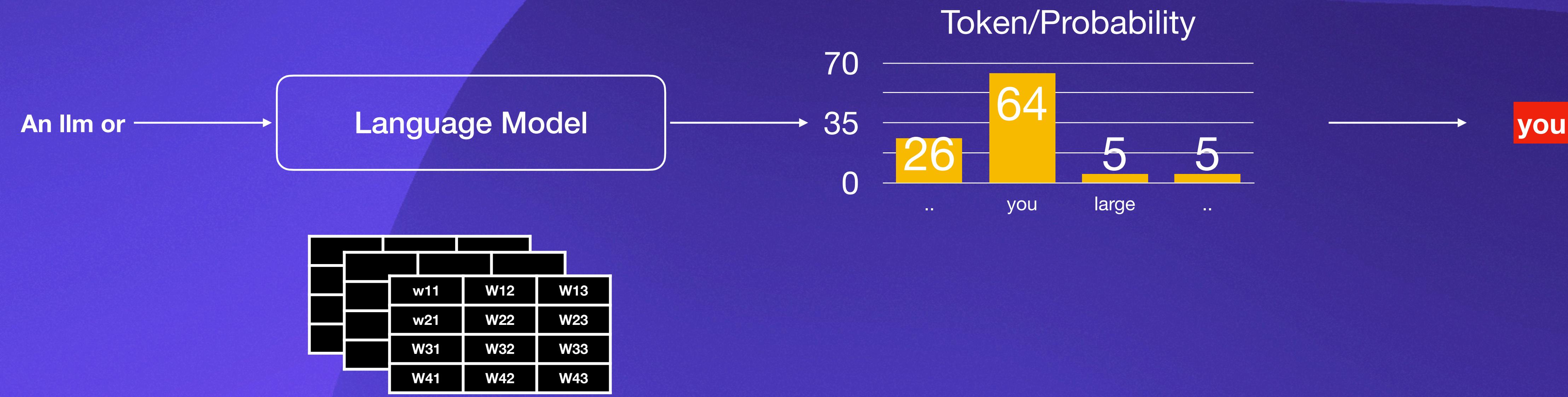
Input	Output	P(output)
An	LLM	0.07
An LLM	Or	
An LLM or	large	
An LLM or large	Language	
...	...	
An LLM or large ...	Learning	
...	...	



Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	
An LLM or large	Language	
...	...	
An LLM or large ...	Learning	
...	...	



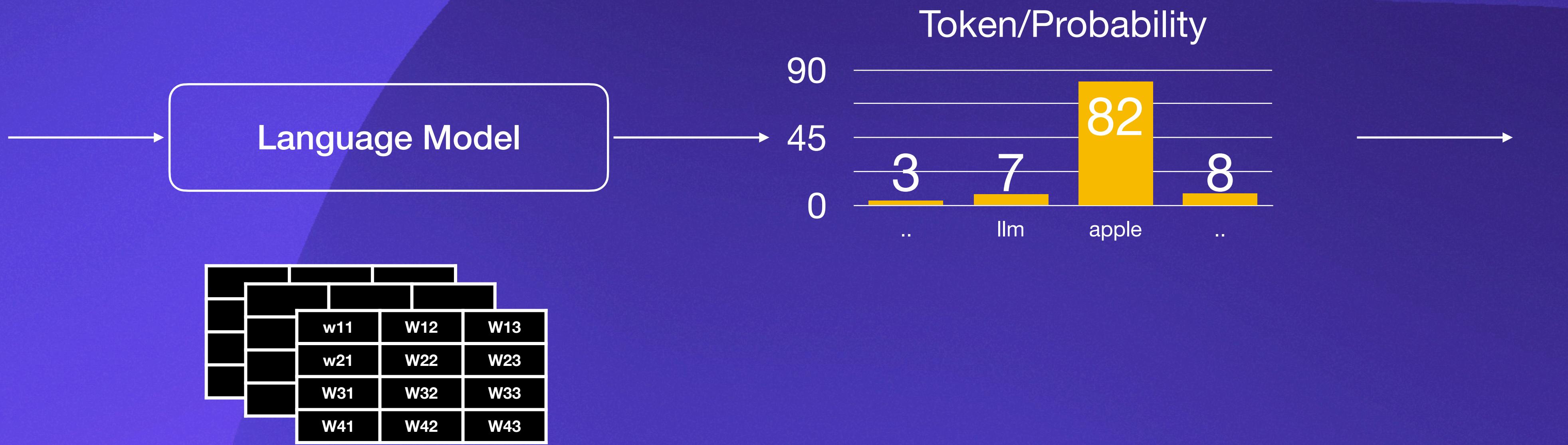
Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	0.05
An LLM or large	Language	
...	...	
An LLM or large ...	Learning	
...	...	



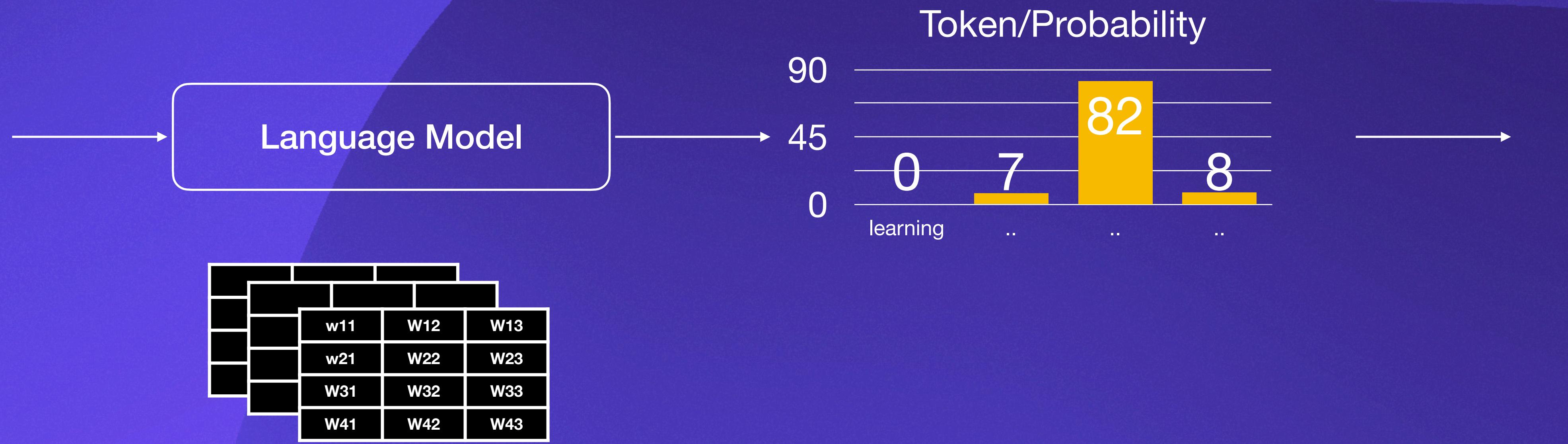
Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	0.05
An LLM or large	Language	0.82
...	...	
An LLM or large ...	Learning	
...	...	



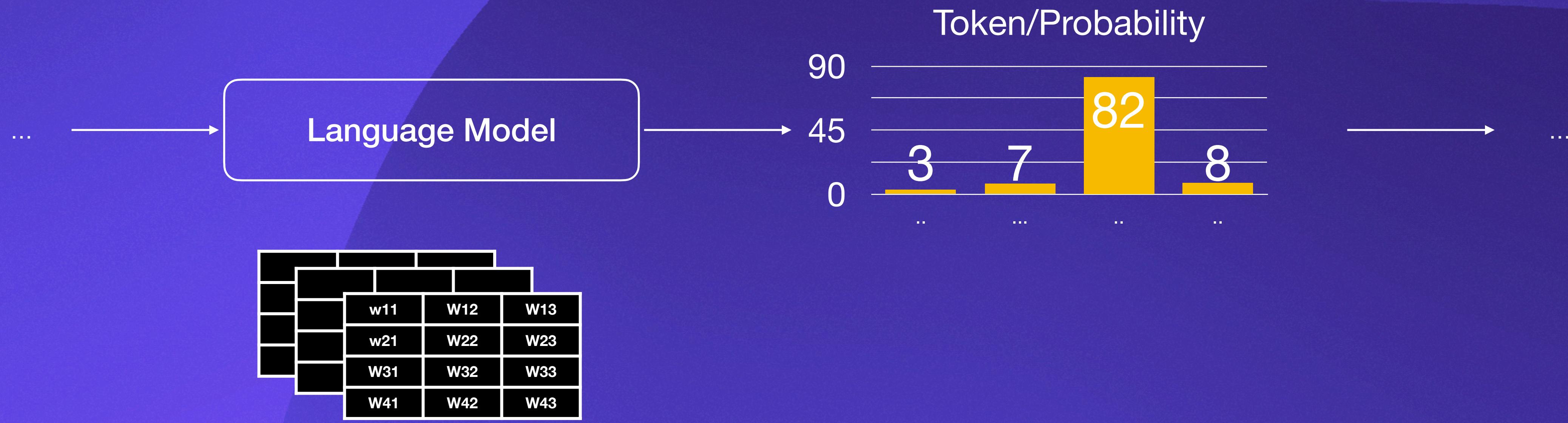
Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	0.05
An LLM or large	Language	0.82
...	...	...
An LLM or large ...	Learning	
...	...	



Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	0.05
An LLM or large	Language	0.82
...	...	...
An LLM or large ...	Learning	0
...	...	...



Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	0.05
An LLM or large	Language	0.82
...	...	...
An LLM or large ...	Learning	0
...	...	..



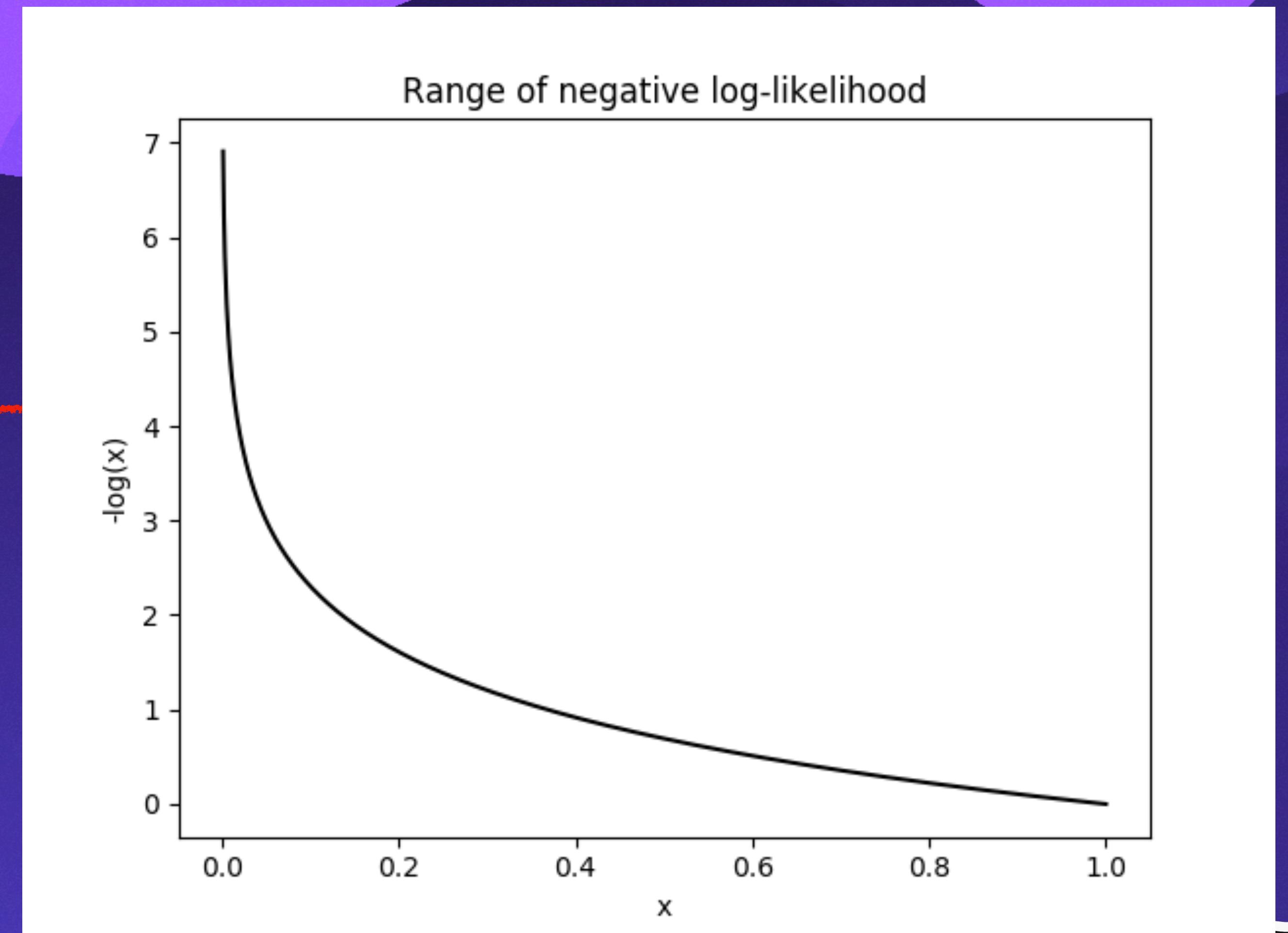
Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	0.05
An LLM or large	Language	0.82
...	...	...
An LLM or large ...	Learning	0
...	...	..

Error:

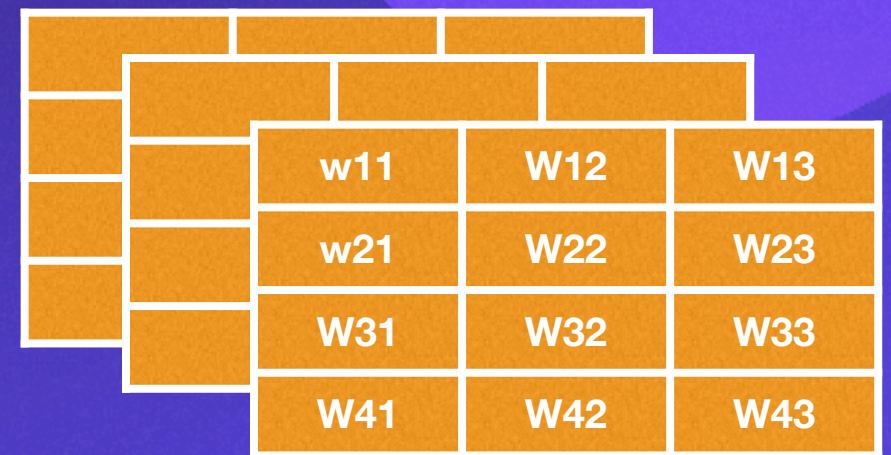
$$e = \frac{1}{m} \sum_{i=0}^m (-\log(p)) \rightarrow 23.02$$



Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	0.05
An LLM or large	Language	0.82
...	...	...
An LLM or large ...	Learning	0
...	...	..



## Update Parameters



Input	Output	P(output)
An	LLM	0.23
An LLM	Or	0.44
An LLM or	large	0.1
An LLM or large	Language	0.5
...	...	...
An LLM or large ...	Learning	0.01
...	...	..

Language Model

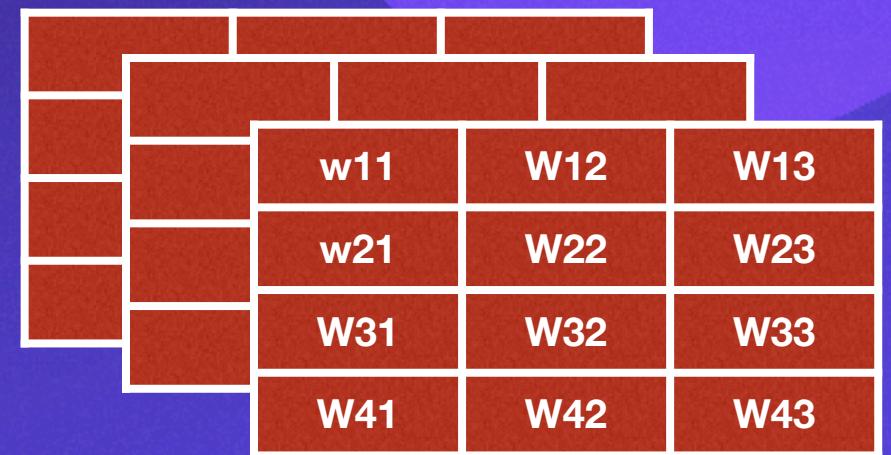
Error:

$$e = \frac{1}{m} \sum_{i=0}^m (-\log(p))$$

14.2



## Update Parameters



Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	0.05
An LLM or large	Language	0.82
...	...	...
An LLM or large ...	Learning	0
...	...	..

Language Model

An arrow points from the Language Model box to the error calculation block.

$$\text{Error: } e = \frac{1}{m} \sum_{i=0}^m (-\log(p))$$

10.05



## Update Parameters



Input	Output	P(output)
An	LLM	0.07
An LLM	Or	0.08
An LLM or	large	0.05
An LLM or large	Language	0.82
...	...	...
An LLM or large ...	Learning	0
...	...	..

→ Error:

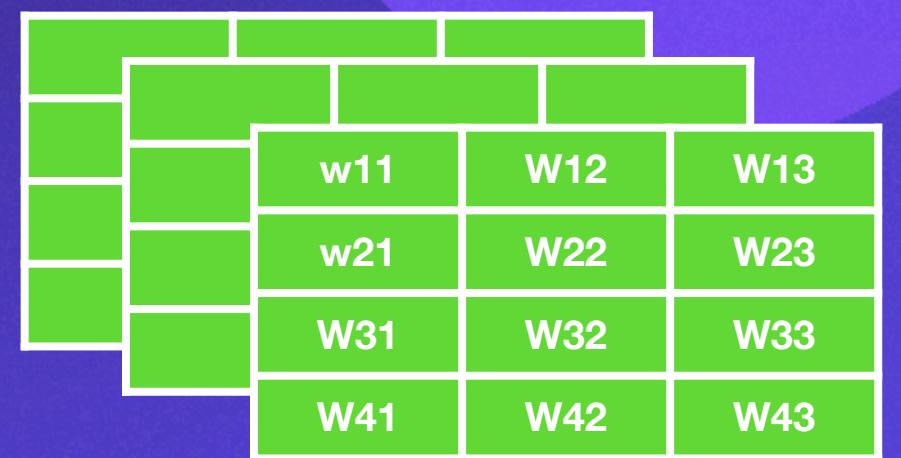
$$e = \frac{1}{m} \sum_{i=0}^m (-\log(p))$$

4.2

Language Model



## Update Parameters



Input	Output	P(output)
An	LLM	0.7
An LLM	Or	0.62
An LLM or	large	0.9
An LLM or large	Language	0.82
...	...	...
An LLM or large ...	Learning	0.87
...	...	..

Language Model

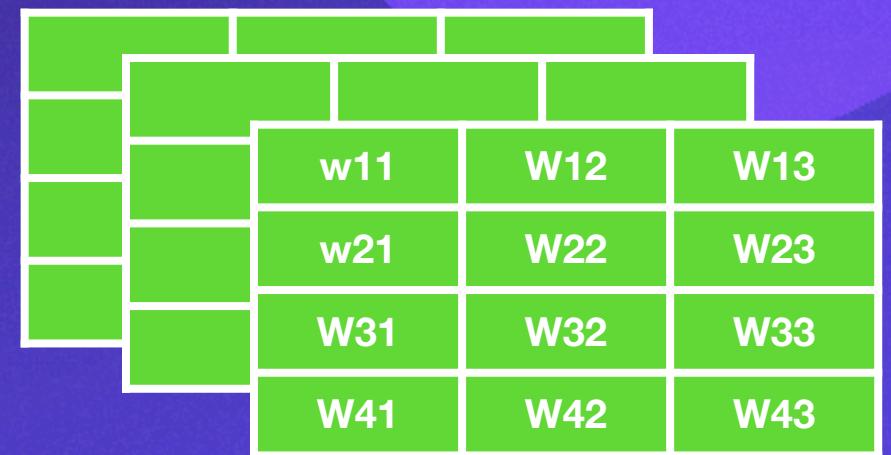
Error:

$$e = \frac{1}{m} \sum_{i=0}^m (-\log(p))$$

0.09



## Update Parameters



Input	Output	P(output)
An	LLM	0.9
An LLM	Or	0.92
An LLM or	large	0.89
An LLM or large	Language	0.82
...	...	...
An LLM or large ...	Learning	0.87
...	...	..

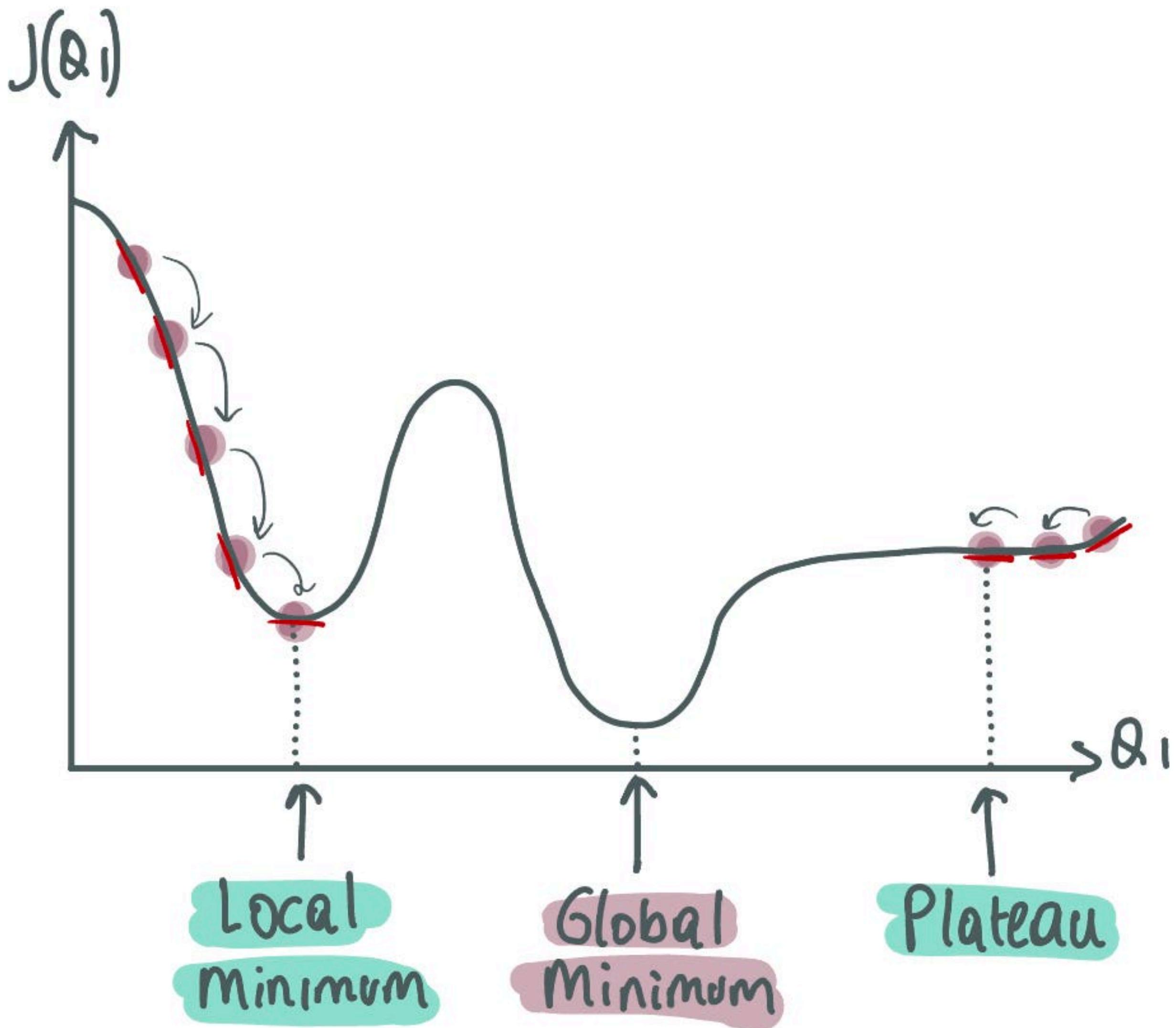
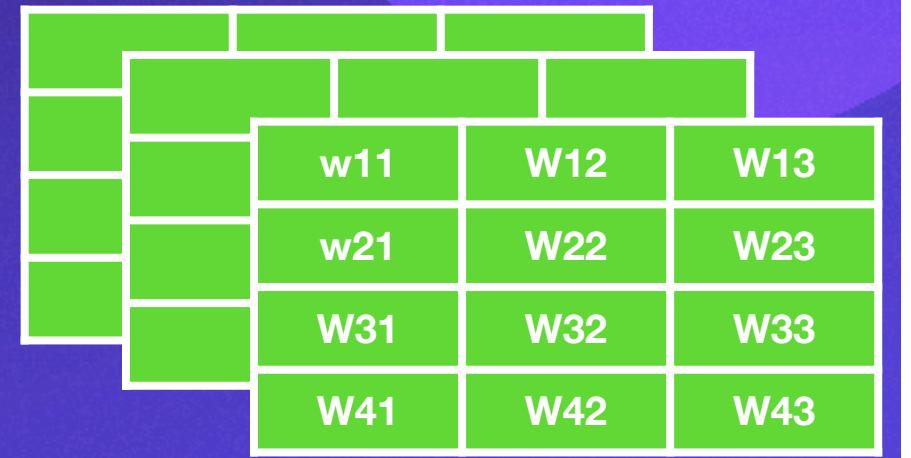
Language Model

Error:

$$e = \frac{1}{m} \sum_{i=0}^m (-\log(p))$$

0.08





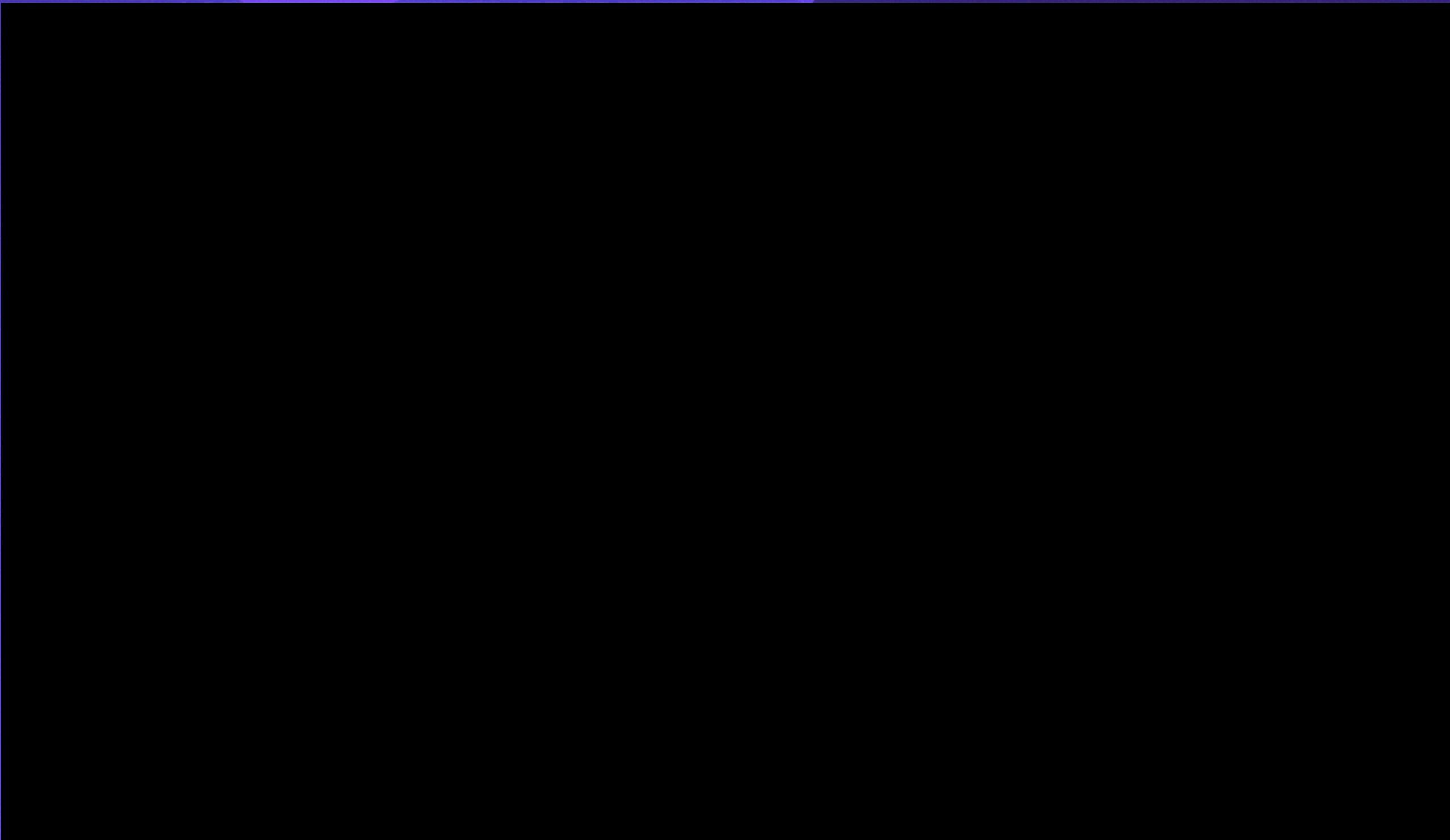
))

0.08



Finally, we got the correct Parameters, which makes our model predict the right next word, and that means our model:

- understanding and generating human language.
- learn patterns, structure, and even nuance in language.
- The Result model called: **BASE MODEL** [hf\_example]



Base Model

w11	W12	W13
w21	W22	W23
W31	W32	W33
W41	W42	W43



# How does an LLM Answer Questions?

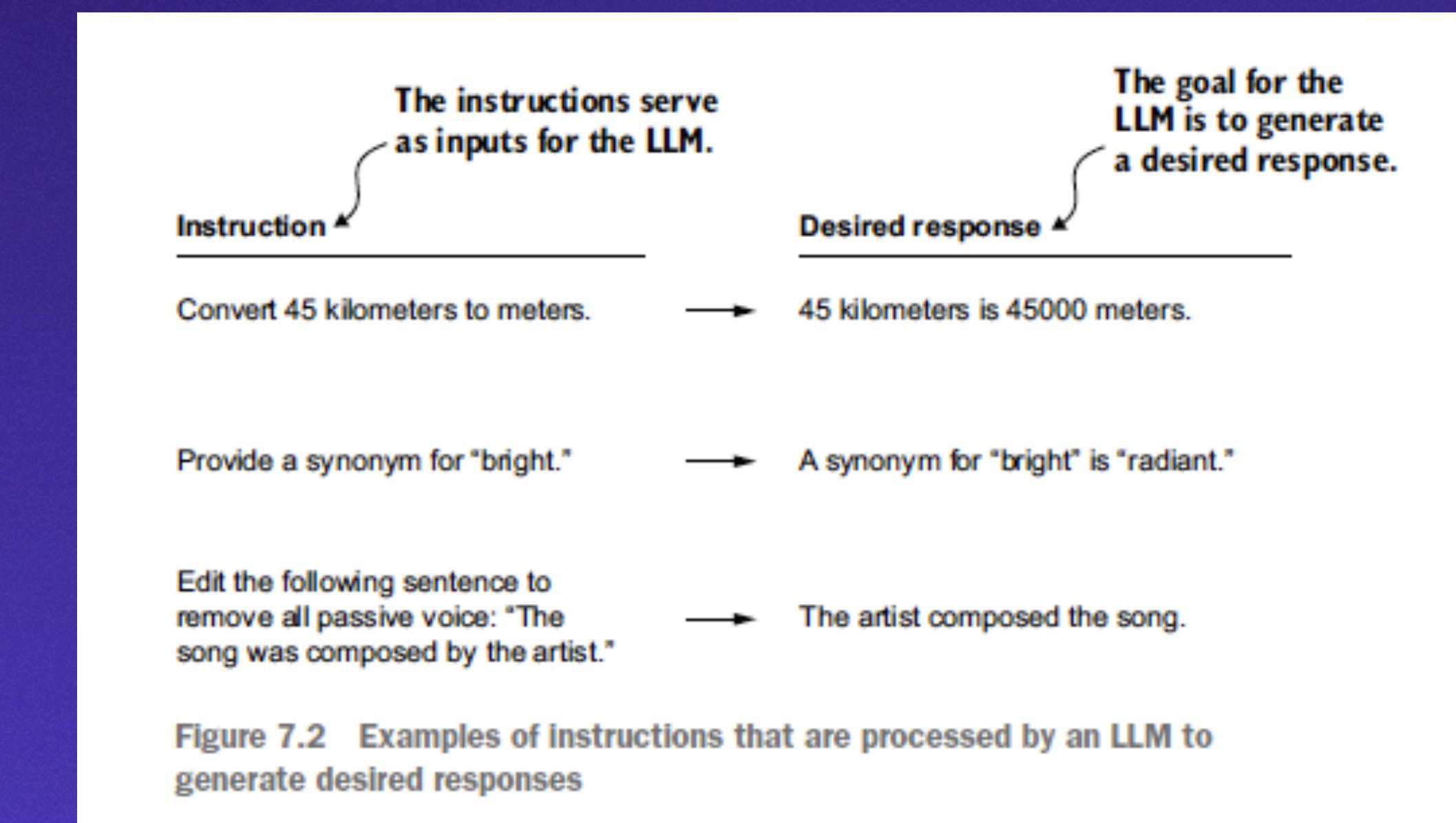
## Model Supervised Fine-tuning



- **Problematic:** I have a model that predicts the next word but doesn't follow instructions.
- **Solution:** Supervised Fine-tuning the **BASE MODEL** to be the **INSTRUCT MODEL**.

## 1. Data Preparation

Instruction	Answer
Q1	A1
Q2	A2
Q3	A3
Q4	A4
Q5	A5
Q6	A6
...	...



## Update Parameters



Instruction	Answer	P(output)
Q1	A1	0.3
Q2	A2	0.2
Q3	A3	0.01
Q4	A4	0.6
Q5	A5	0.05
Q6	A6	0.22
...	...	..

Base Model

Error:

$$e = \frac{1}{m} \sum_{i=0}^m (-\log(p))$$

11.21



## Update Parameters



Instruction	Answer	P(output)
Q1	A1	0.3
Q2	A2	0.2
Q3	A3	0.01
Q4	A4	0.6
Q5	A5	0.05
Q6	A6	0.22
...	...	..

Base Model

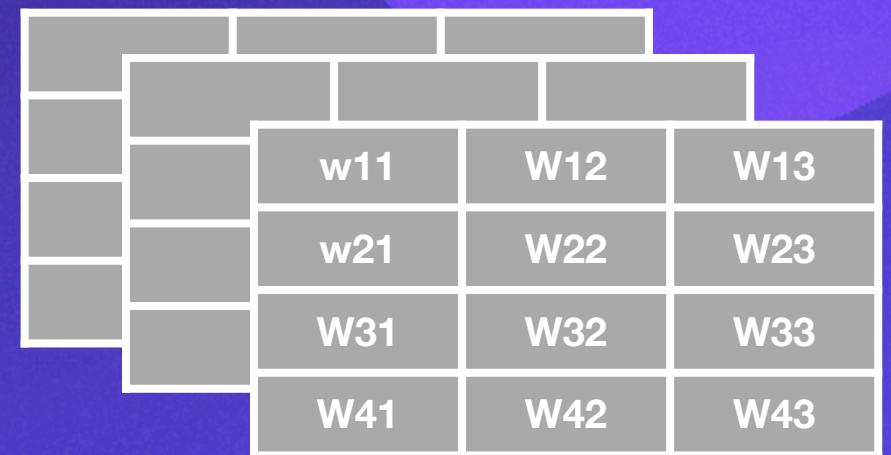
An arrow points from the 'P(output)' column of the table to a red box containing the error formula.

$$\text{Error: } e = \frac{1}{m} \sum_{i=0}^m (-\log(p))$$

8.2



## Update Parameters



Instruction	Answer	P(output)
Q1	A1	0.3
Q2	A2	0.2
Q3	A3	0.01
Q4	A4	0.6
Q5	A5	0.05
Q6	A6	0.22
...	...	..

Base Model

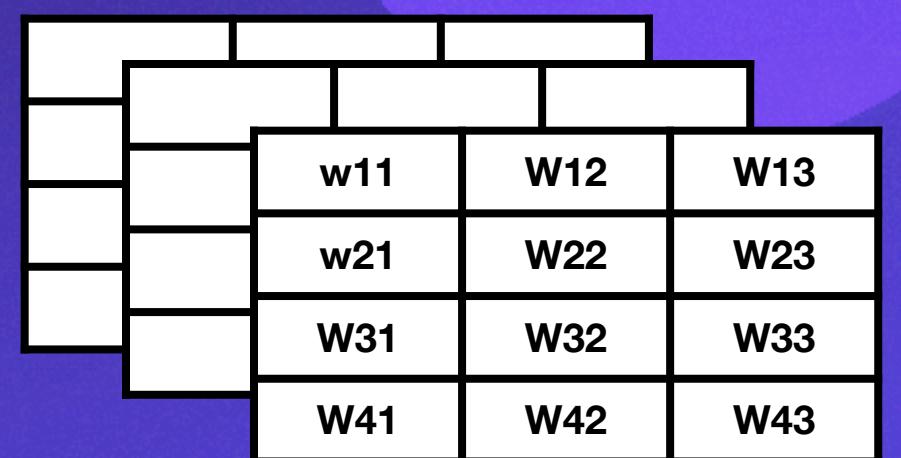
Error:

$$e = \frac{1}{m} \sum_{i=0}^m (-\log(p))$$

2.1



## Update Parameters



Instruction	Answer	P(output)
Q1	A1	0.9
Q2	A2	0.91
Q3	A3	0.8
Q4	A4	0.88
Q5	A5	0.98
Q6	A6	0.90
...	...	..

Base Model

Error:

$$e = \frac{1}{m} \sum_{i=0}^m (-\log(p))$$

0.1



Finally, we got the correct parameters, which made our model follow the instructions.

- The Result model called: **INSTRUCT MODEL** [hf\_example]



w11	W12	W13
w21	W22	W23
W31	W32	W33
W41	W42	W43



# Common Limitations of LLMs



- Hallucination: generation of incorrect information with high confidence.

USER  
who is greater 9.9 or 9.11

ASSISTANT  
To compare the two numbers, we can see that:  
9.11 is greater than 9.9  
So, the answer is 9.11.

- Knowledge cutoff: limited to the training data timeframe.

USER  
who is crowned champions of the TotalEnergies CAF U-17 Africa Cup of Nations 2025?

ASSISTANT  
I couldn't find any information about the TotalEnergies CAF U-17 Africa Cup of Nations 2025 yet, as my knowledge cutoff is in 2023 and the information about the tournament may not be available yet.

- Data privacy: limited to public training data, no access to proprietary information.

USER  
who is abdeljalil EL MAJJODI?

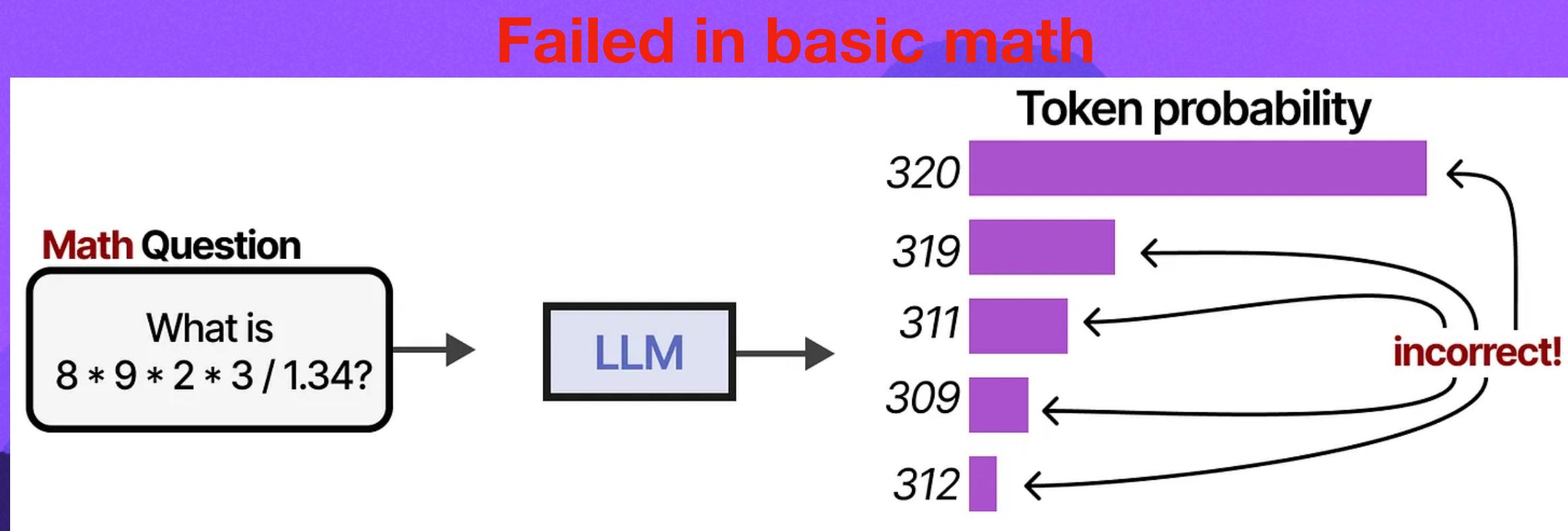
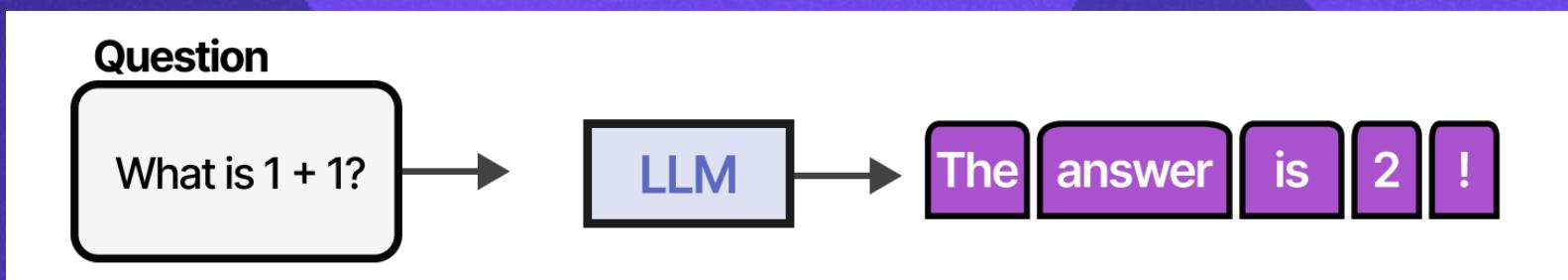
ASSISTANT  
I couldn't find any information on a prominent or well-known person named Abdeljalil El Majoudi. It's possible that Abdeljalil El Majoudi is a private individual or not a public figure, which would explain the lack of information.  
However, I found some information on Abdeljalil El Majoudi's brother, Abdeljalil El Majoudi's role is unclear.

# Agentic AI

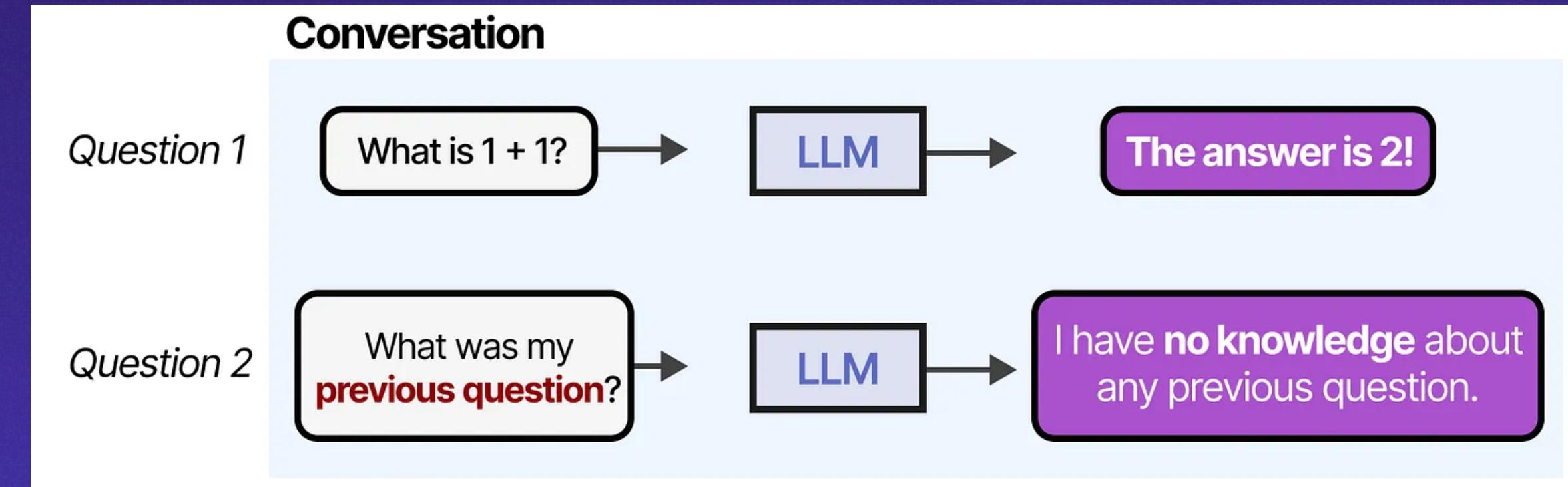
## What are LLM Agents?



LLM does nothing more than next-token prediction



Not remember conversations

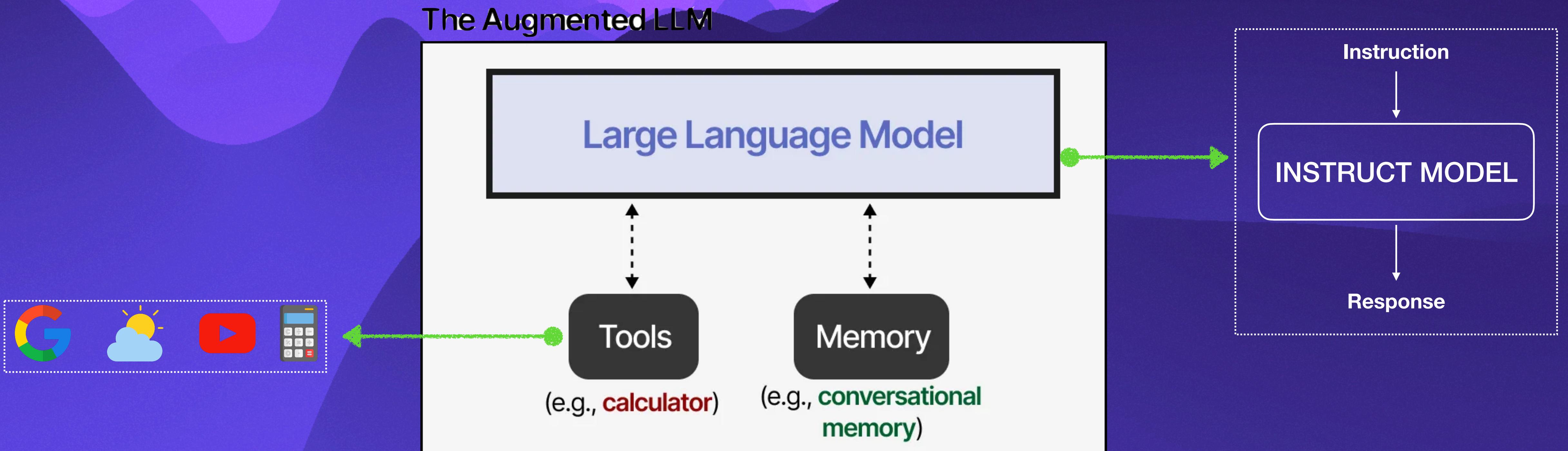


Does this mean LLMs are horrible? NO

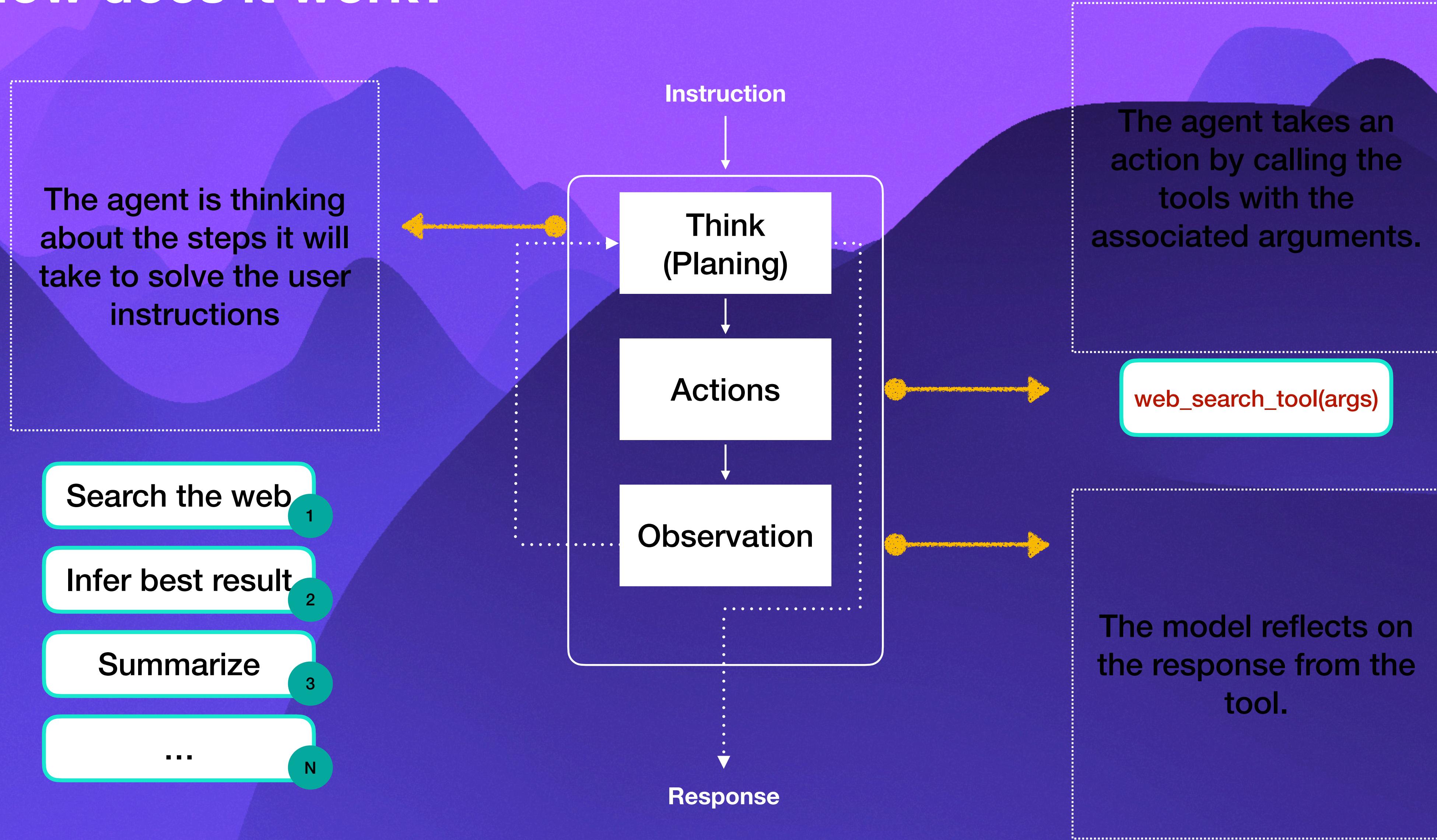


# Solution

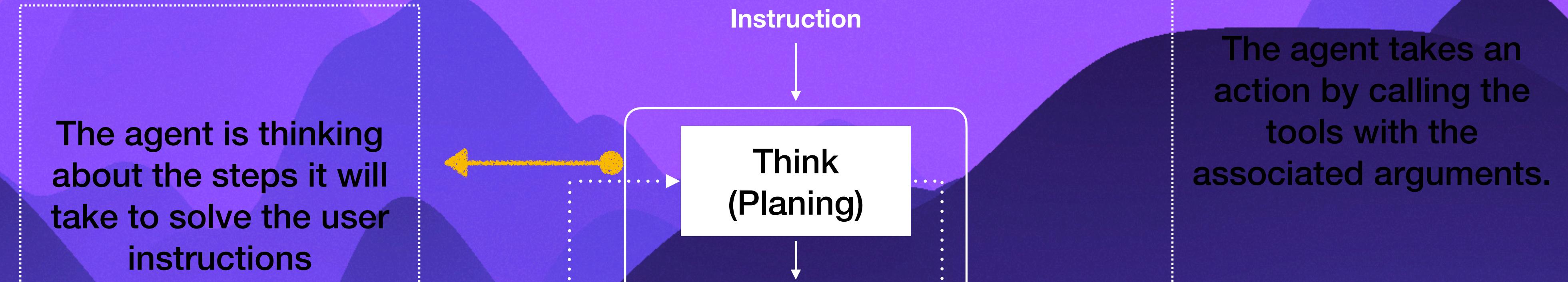
## Augmented LLM (LLM Agent)



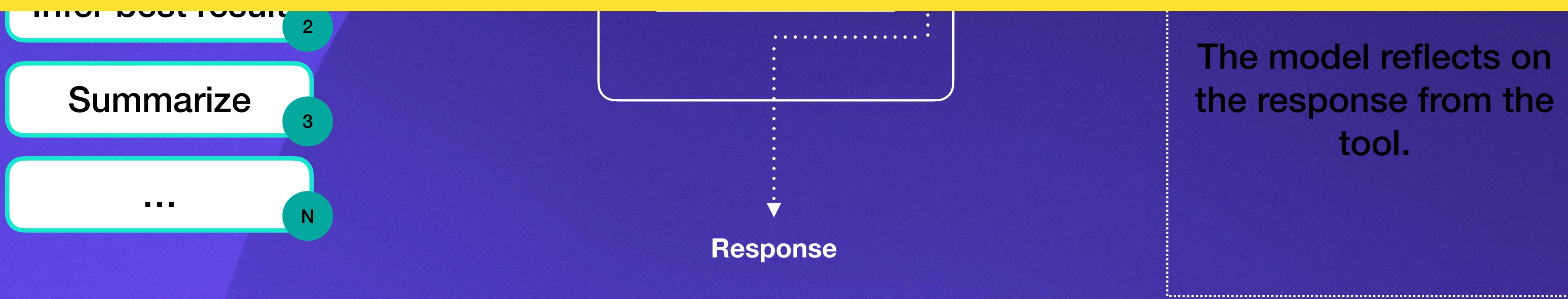
# How does it work?



# How does it work?

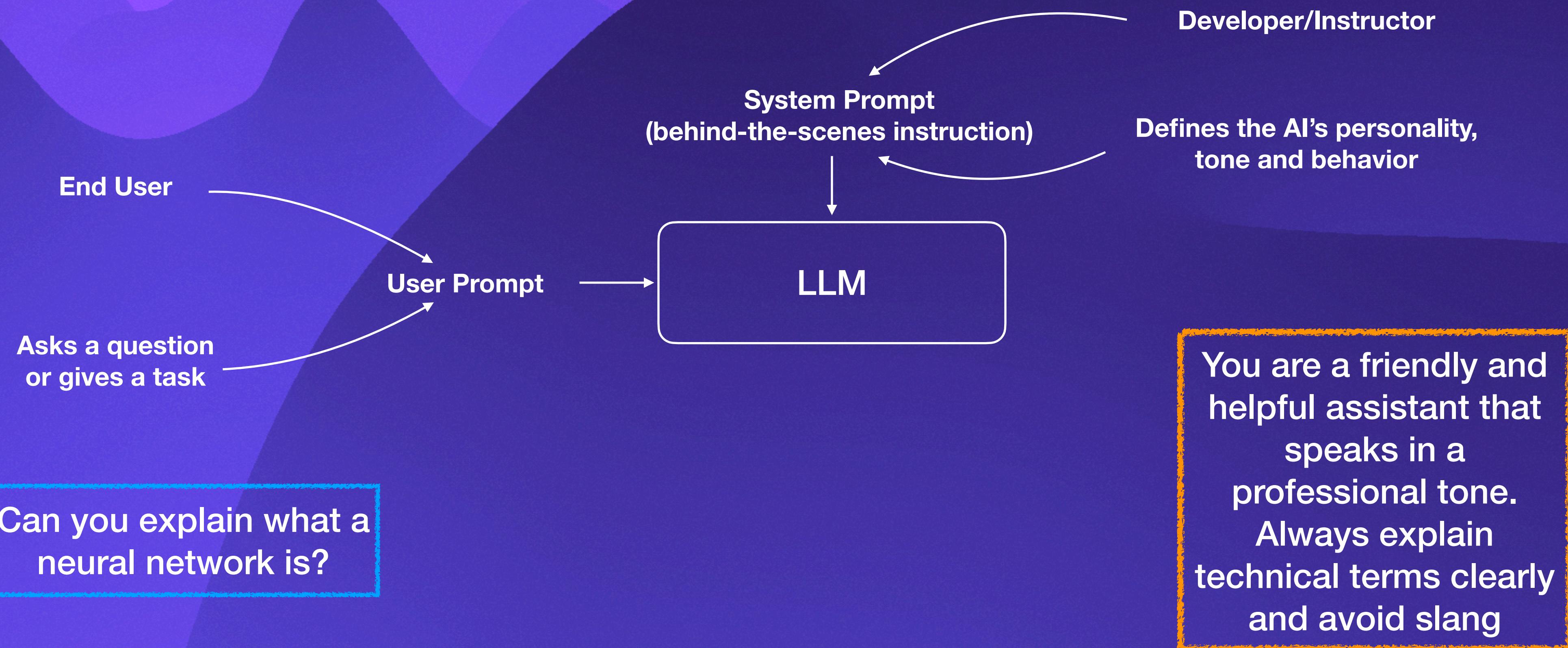


## HOW WILL LLM KNOW ABOUT THIS?



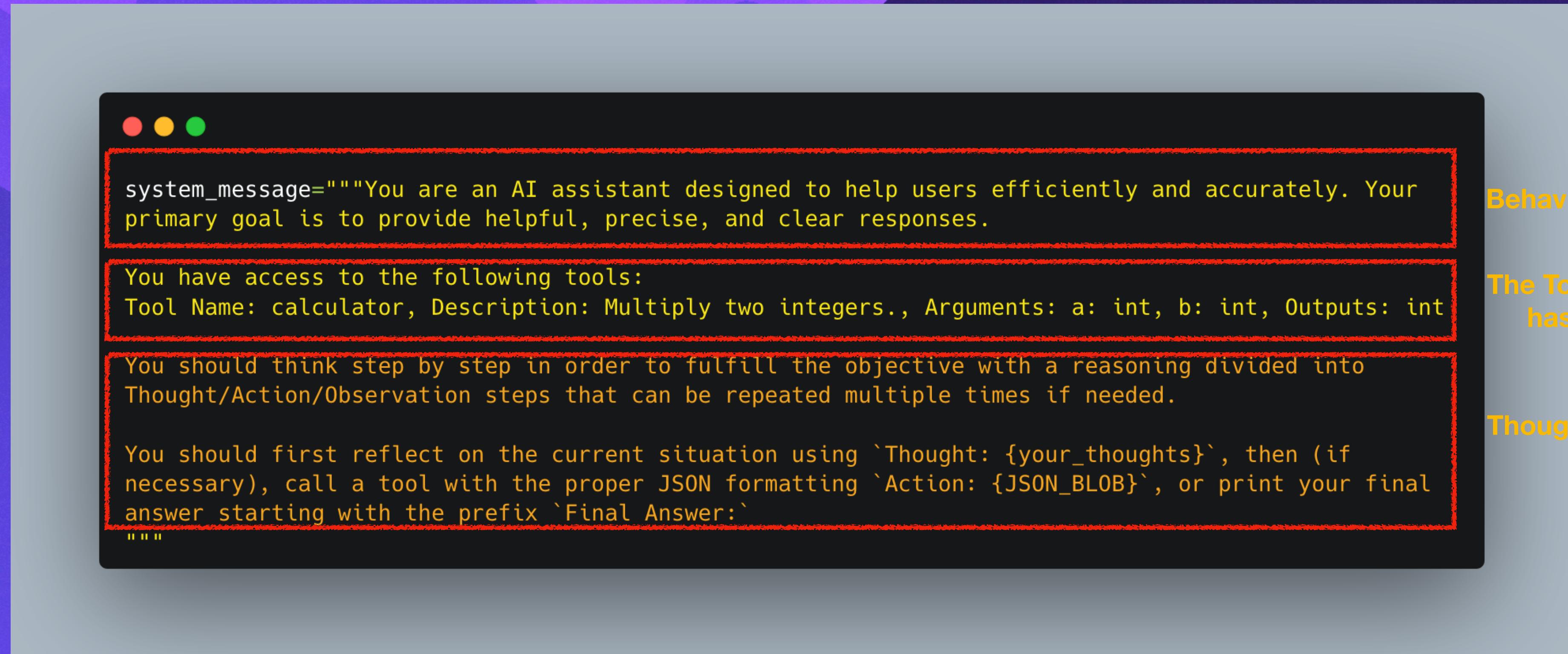
# Prompt

## Define Rules and Guidelines



# System Prompt

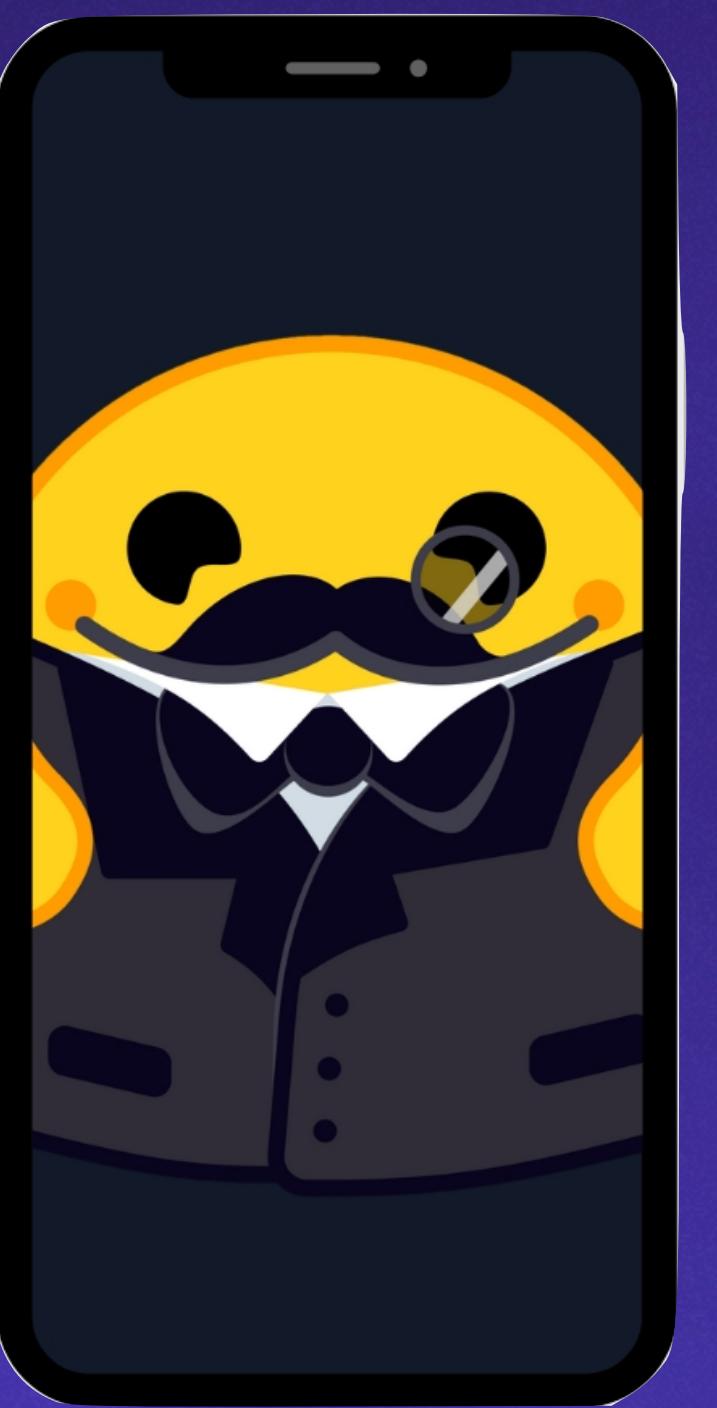
## Define Our Agent



# Example Weather Agent

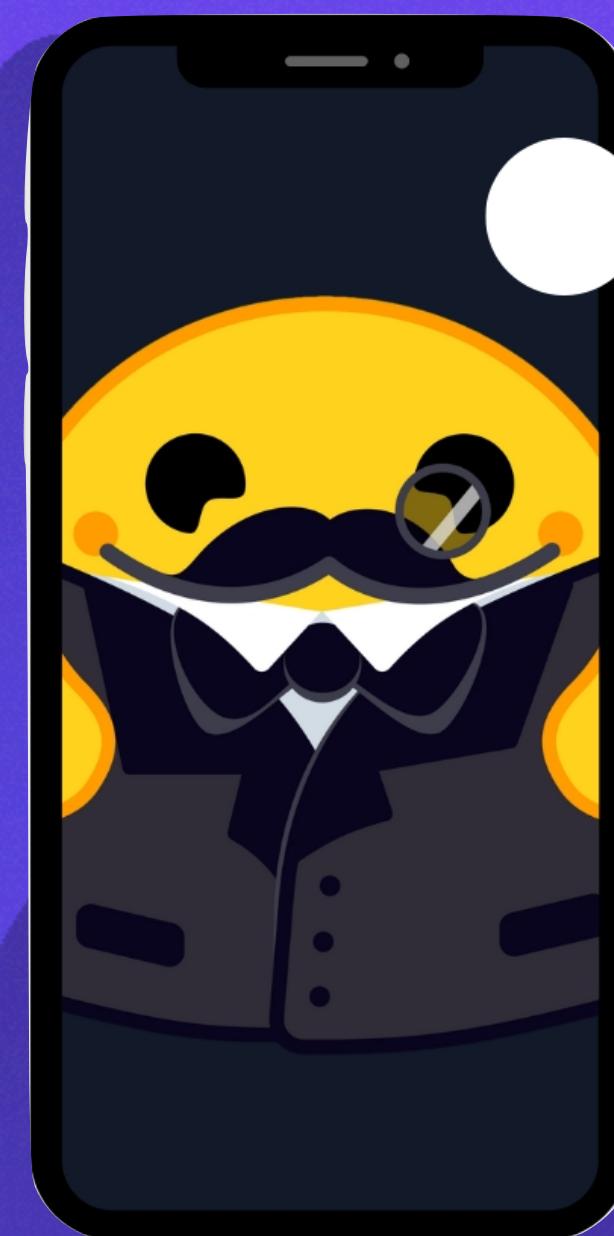


What's the current  
weather in New  
York?



# Example

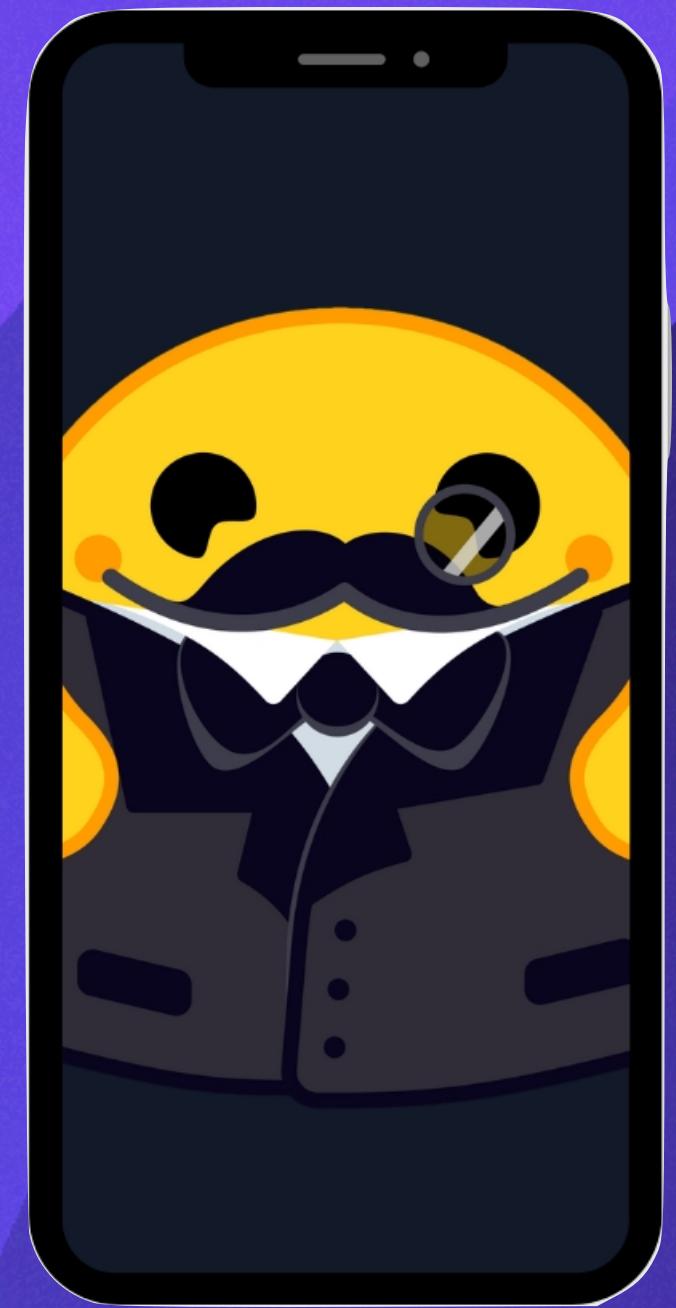
## Weather Agent



The user needs current weather information for New York. I have access to a tool that fetches weather data. First, I need to call the weather API to get up-to-date details.

Think Step

# Example Weather Agent

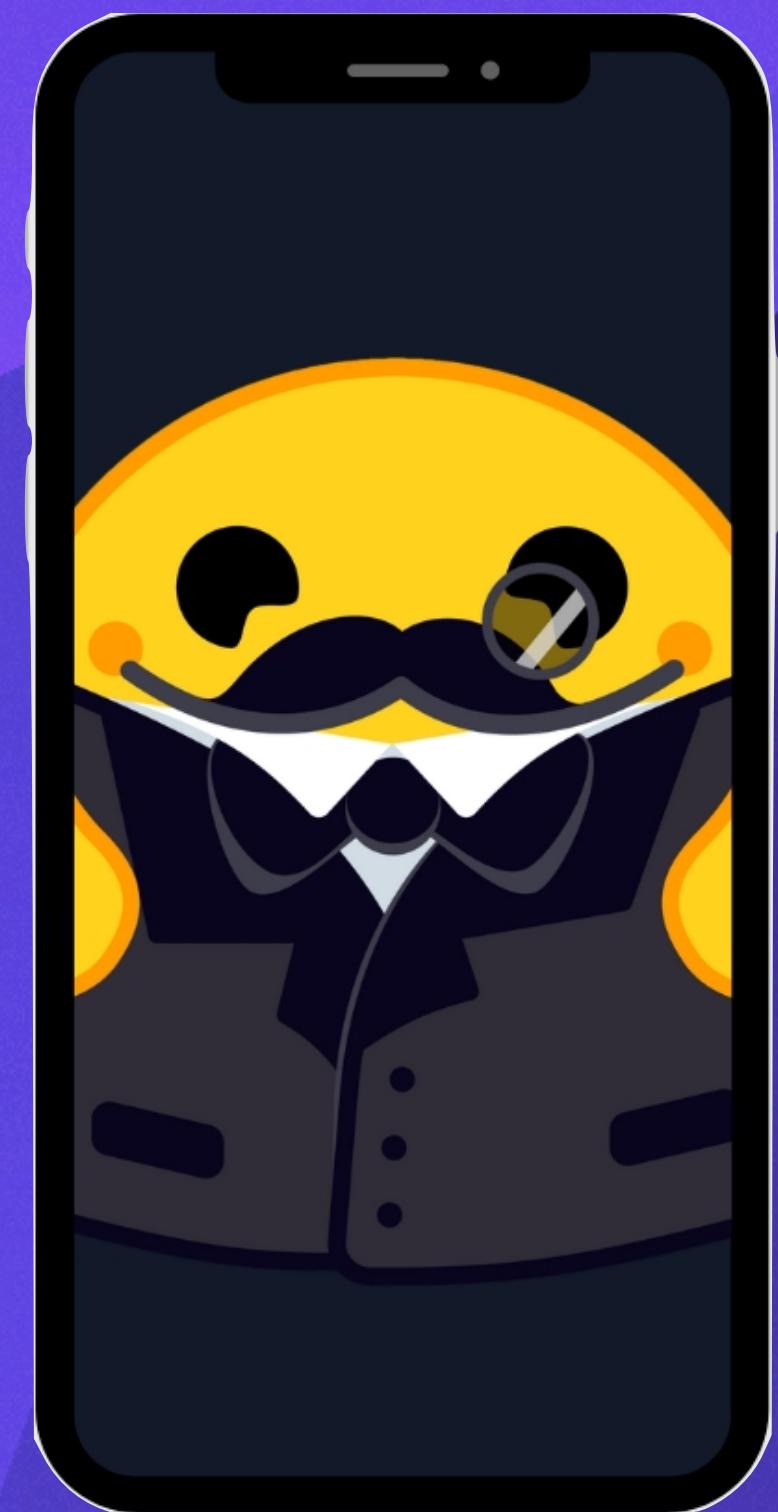


```
{  
  "action": "get_weather",  
  "action_input": {  
    "location": "New York"  
  }  
}
```

Act Step

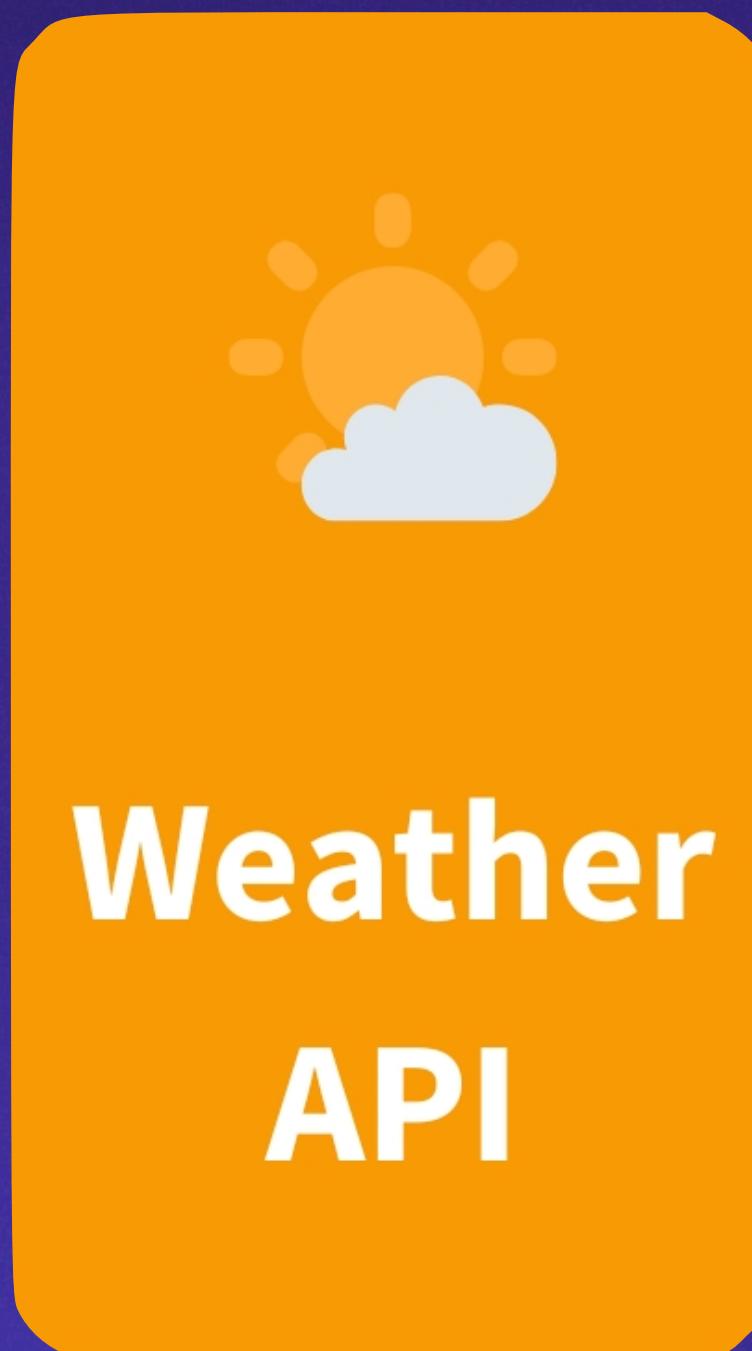


# Example Weather Agent

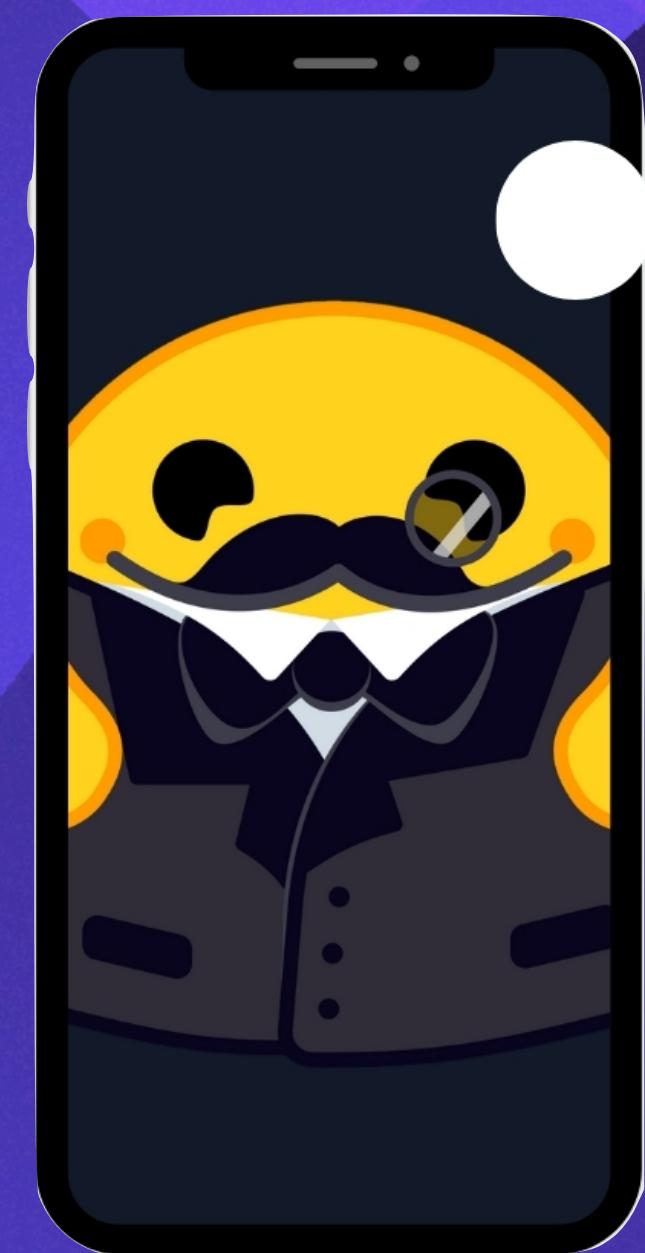


←  
Current weather  
in New York:  
partly cloudy,  
15°C, 60%  
humidity.

Observation Step



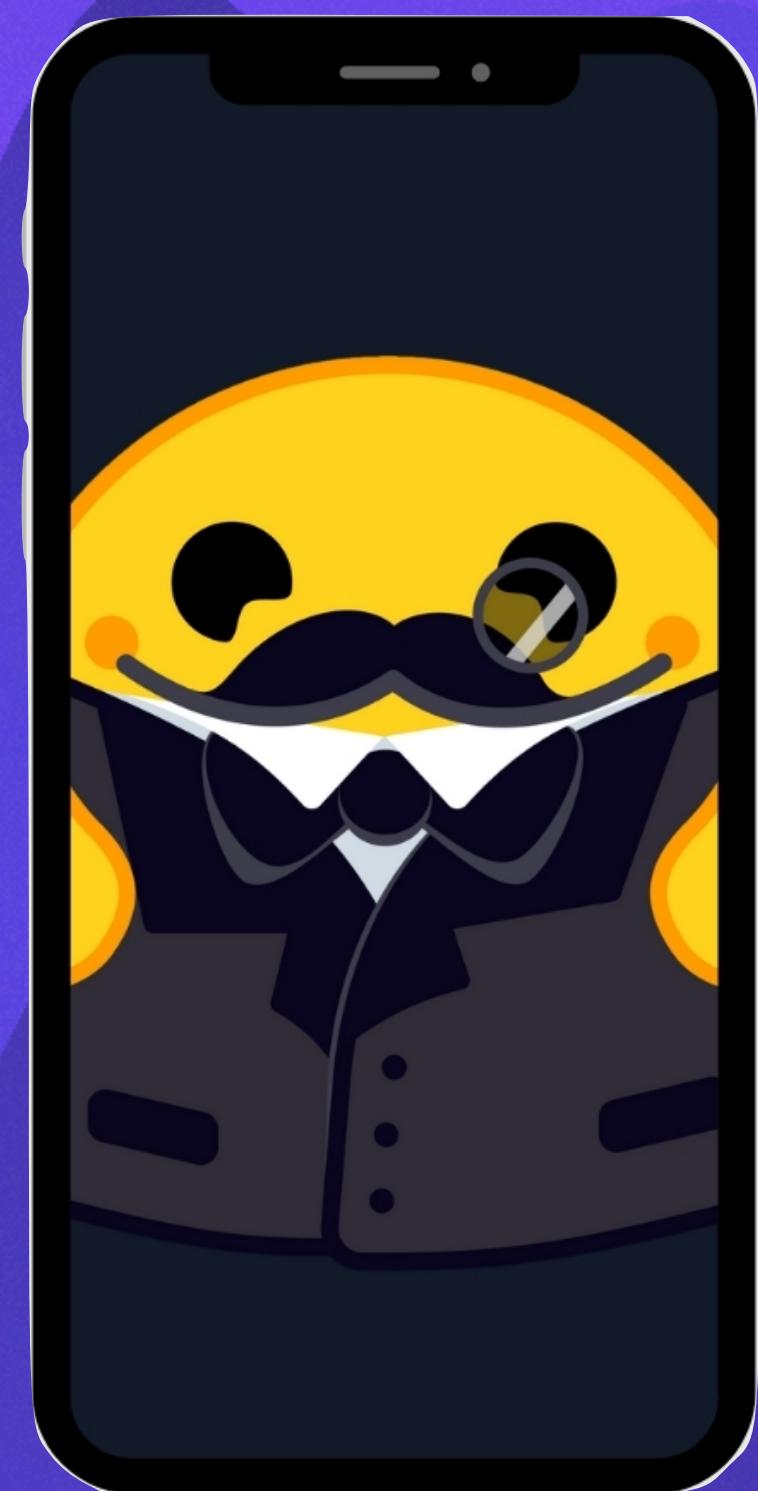
# Example Weather Agent



**Now that I have the weather  
data for New York, I can compile  
an answer for the user**

Think Step

# Example Weather Agent



The current weather in New York is partly cloudy with a temperature of 15°C and 60% humidity.



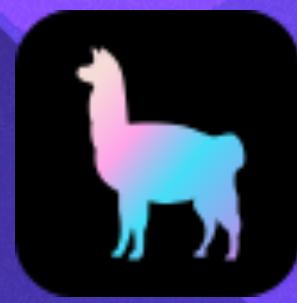
# Code Time

## Dummy Agent: Without Framework



# Code Time Frameworks





Llamaindex



LangChain

crewai



**smolagents**



# Code Time

## Hands-On SmolAgent



# Slides + Resources



# AtlasIA Community



# THANK YOU

