

# LLM Fine-Tuning

## Part 1

**Constantine (Kostas) Kopsinis, Qitong Wang, Stratos Idreos**

kkopsinis@g.harvard.edu



What is supervised learning?

What is an LLM?

What is pre-training?

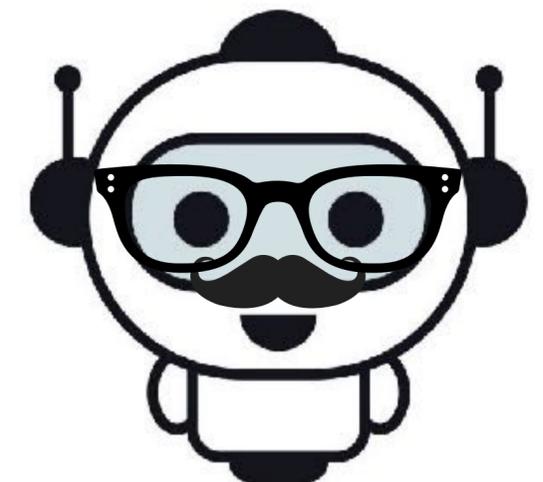
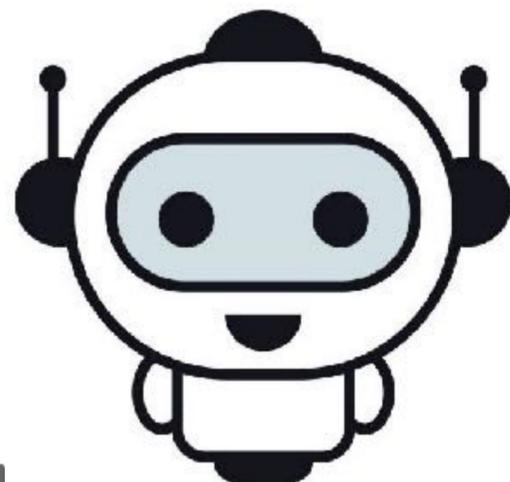
What is an agent?

What is tuning?

# LLM Fine-Tuning

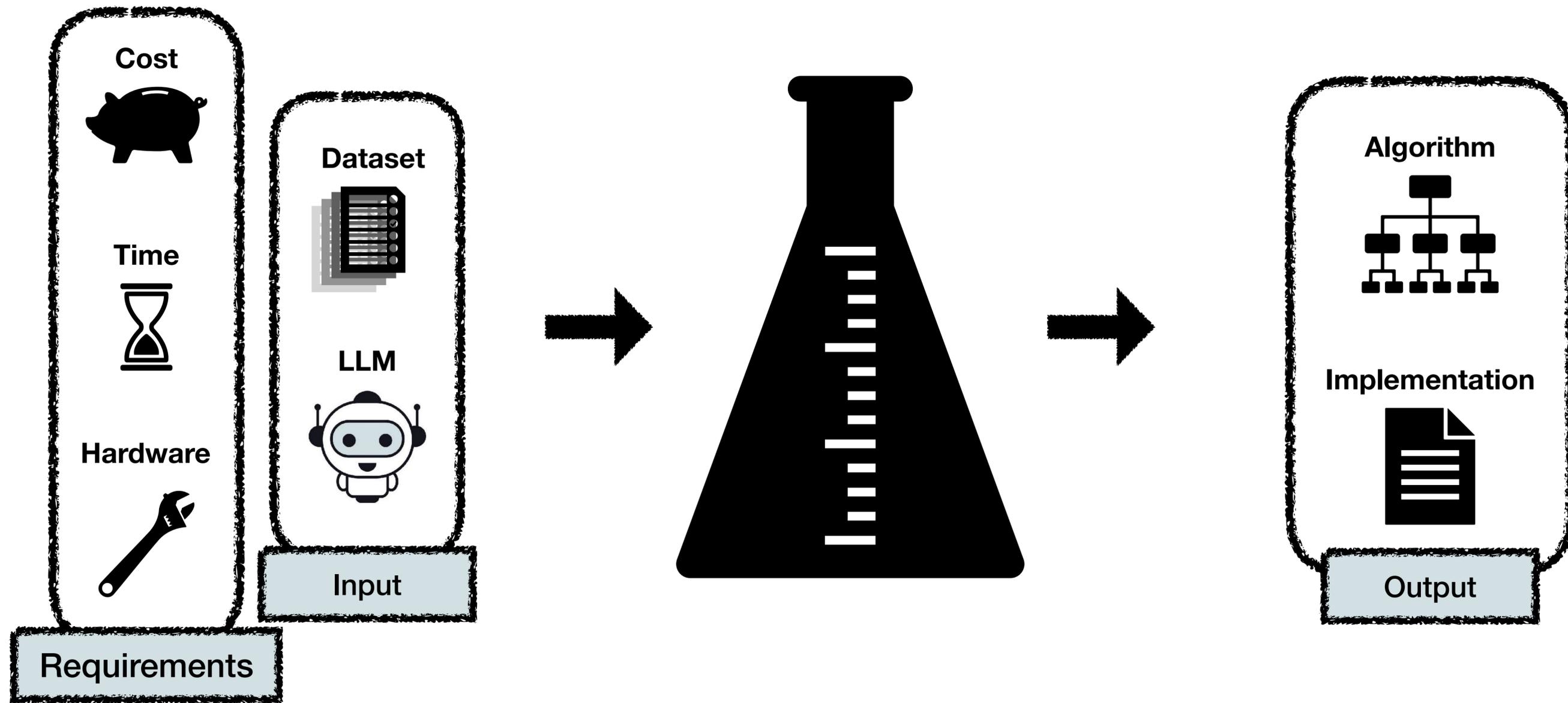
What is post-training?

What is fine-tuning?



# LLM Fine-Tuner

The vision for a self-designing system



**Learning Algorithm**

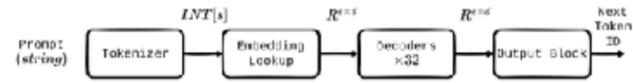
**Task/  
Dataset**

**LLM  
Fine-Tuning**

**Weight  
Update**

**Architecture**

## Overview of an LLM



$s$ : number of tokens in our prompt.  
 $d$ : dimension of vectors we use for embeddings.

**We Inspect All of Them!**



## LLM Inference Lectures Part 2: LLM

(Harvard CS96G - DATA/AT Systems LAB)



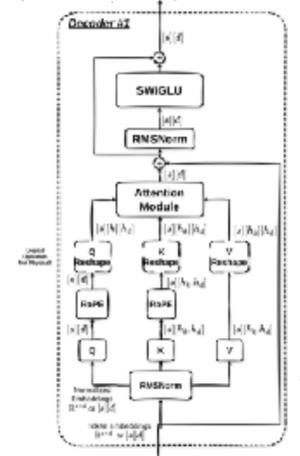
Milad Rezaei Hajidehi  
(milad@seas.harvard.edu)



**Last Missing Piece**



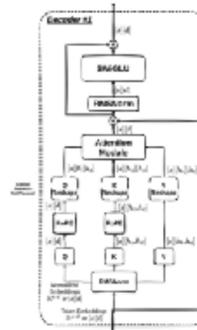
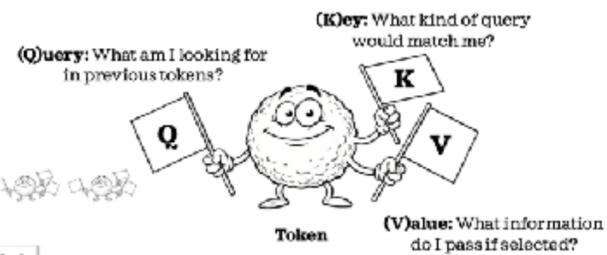
**Goal: Transform Initial Embedding  $\rightarrow$  Contextual Representation for Output Block**



# Architecture

**Input of Decoder:** One embedding per token.

**Output of Decoder/Attention(H):** One embedding per token. But transformed!  
 based on context and previous tokens!



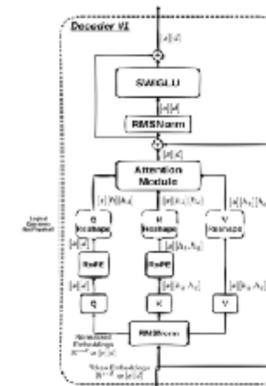
How much  $j$  can impact  $i$ ?  
(normalized)

$$\alpha_{ij} = \frac{\exp\left(\frac{Q_i K_j^T}{\sqrt{128}}\right)}{\sum_k \exp\left(\frac{Q_i K_k^T}{\sqrt{128}}\right)}$$

Embeddings Transformed.

$$\text{Attention}(Q, K, V) = AV$$

! In LLAMA3, We use GQA rather than Vanilla, which is slightly different.



# Task/Dataset

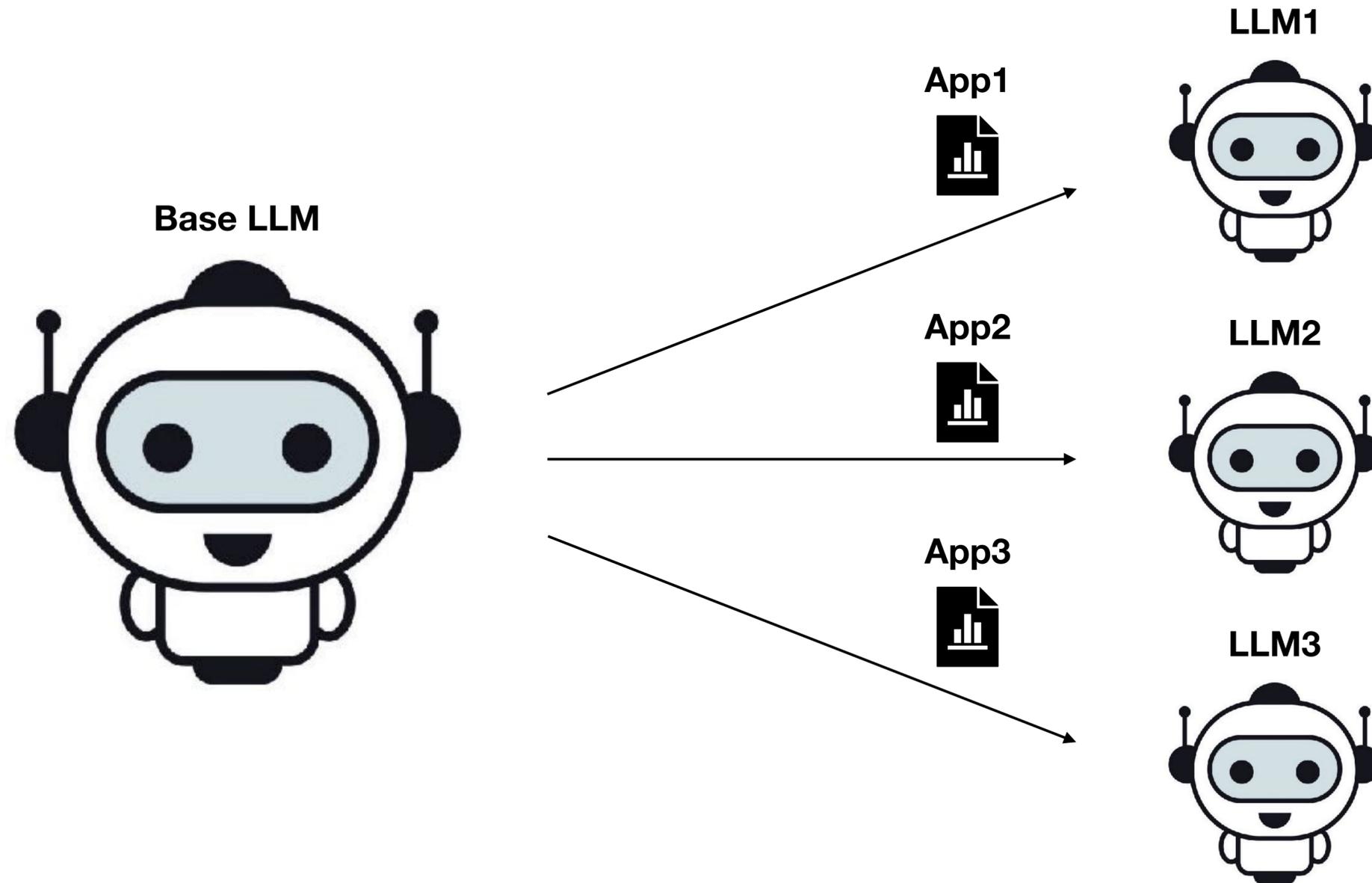
# Task/Dataset



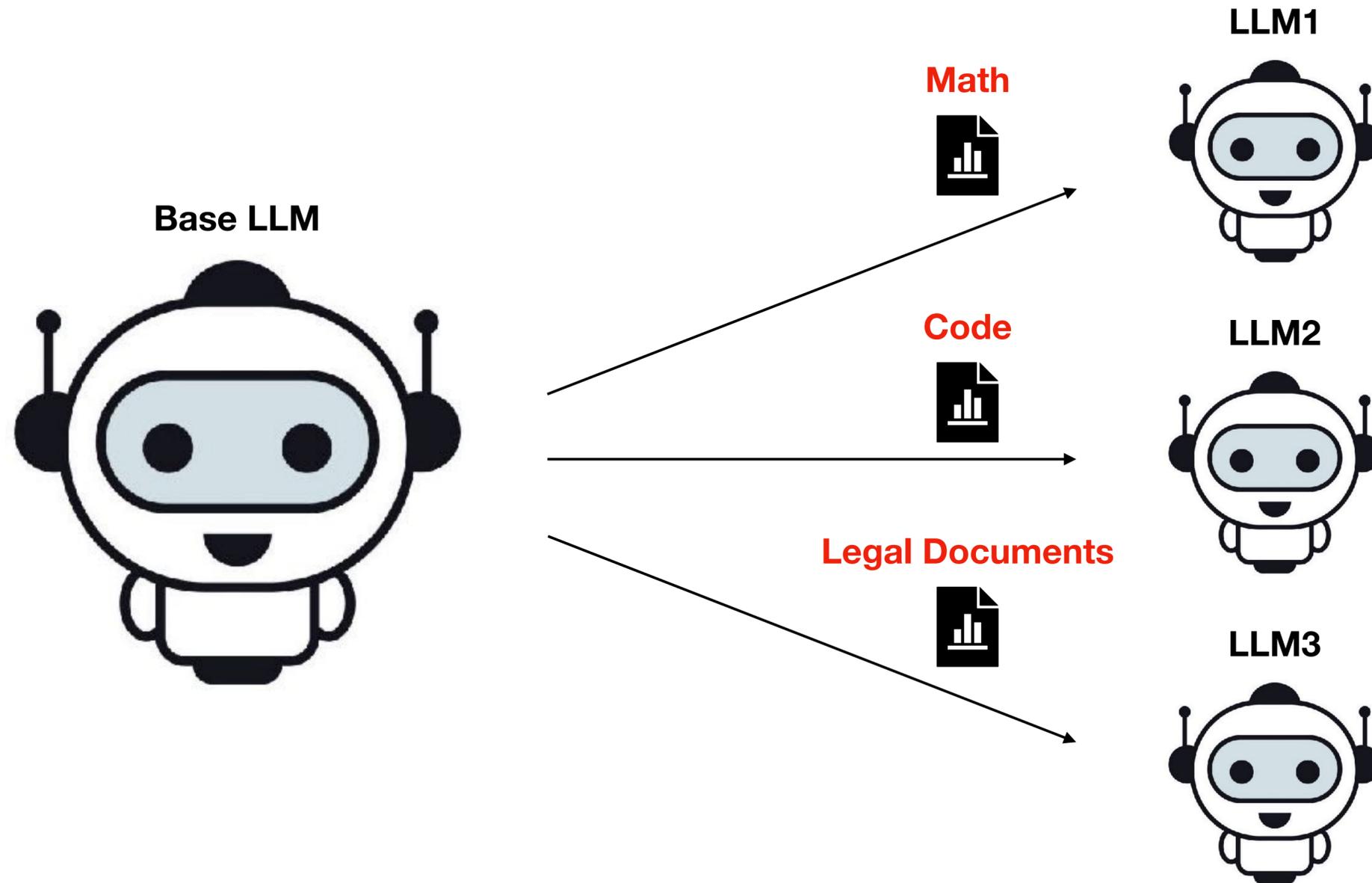
## Domain

## Behavior

# Domain



# Domain



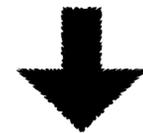
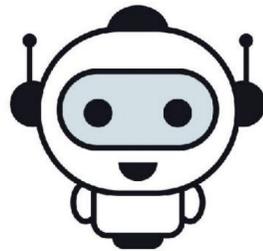
# Behavior

## Instruction Tuning

**prompt:**

Can you explain what a neural network is in simple terms?

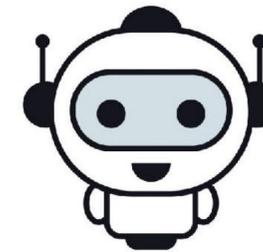
**Base LLM**



**output:**

I have a vague idea but I've yet to completely understand it.

**Instruction Tuned LLM**



**output:**

A neural network is a system of connected units that learn patterns from examples, similar to how the brain processes information.

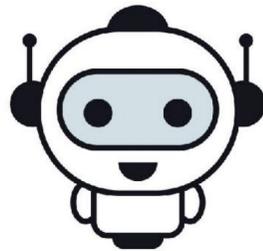
# Behavior

## Instruction Tuning

**prompt:**

```
PROMPT: Rewrite the sentence below as a polite email request. Keep it to one sentence, under 20 words.  
"Send me the assignment deadline."  
ANSWER:
```

**Base LLM**



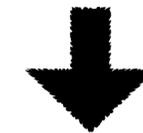
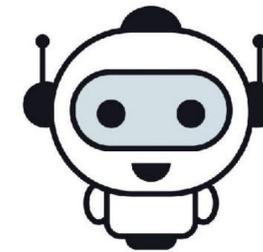
**output:**

```
Hi, I was wondering if you could send me the assignment deadline when you have a chance.
```



Both models understand English.  
Only one of them is trying to do what we asked.

**Instruction Tuned LLM**



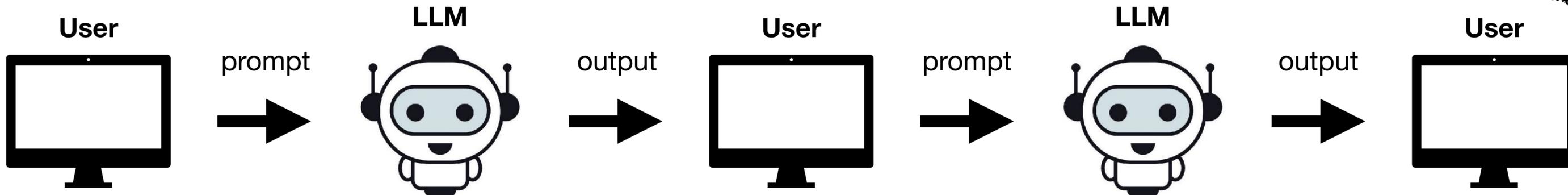
**output:**

```
Hi, could you please let me know the assignment deadline?
```

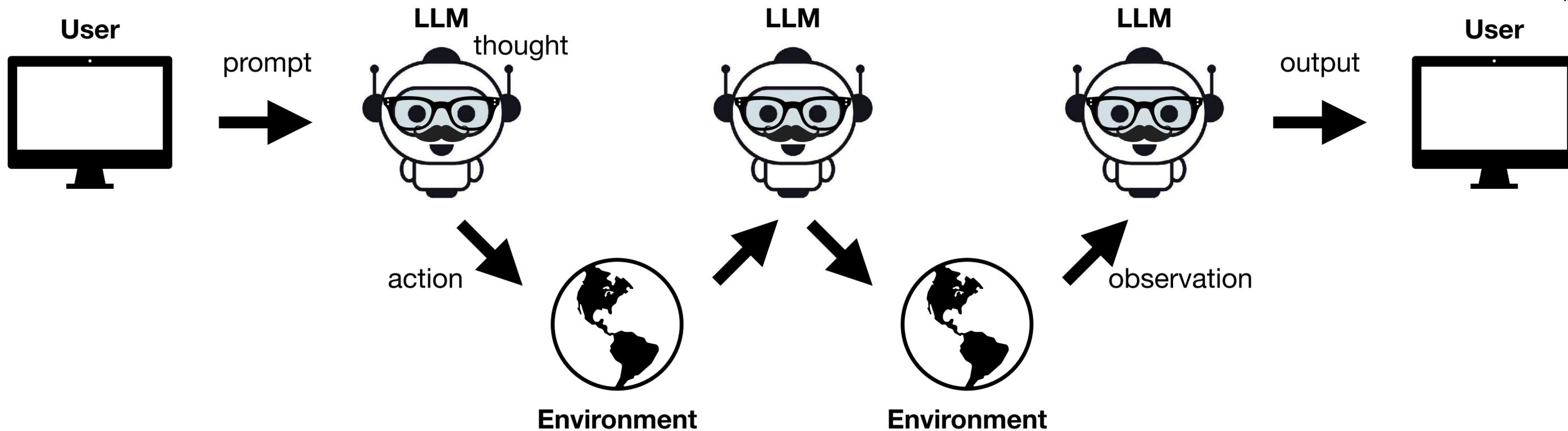
### Single Turn



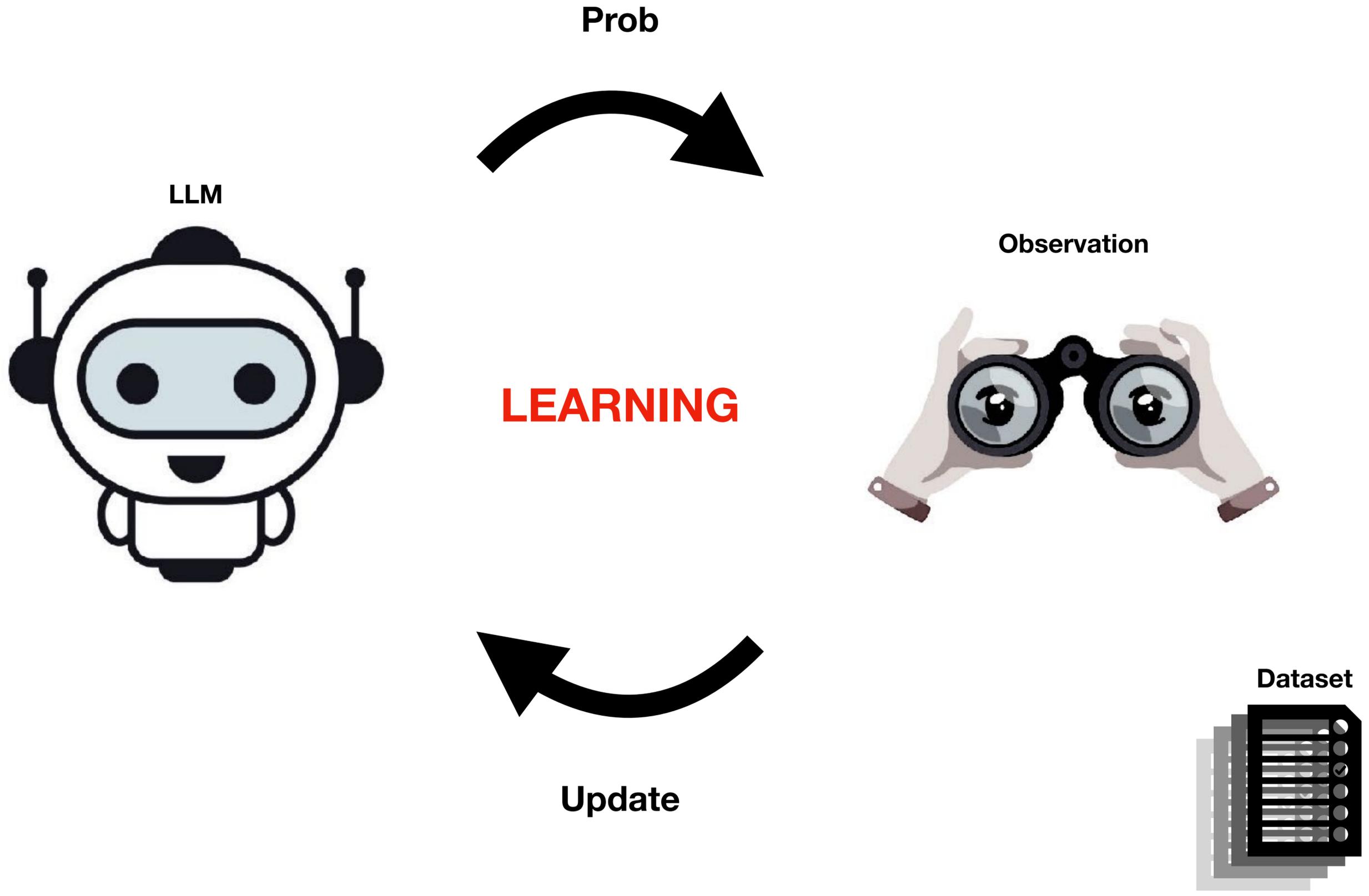
### Multi Turn



### Agent



# Learning Algorithm



## Supervised Learning

## Reinforcement Learning

Objective ?

distance  
minimization

Learning Paradigms

reward  
maximization

Labels from...

Reward from...

internal  
structure of  
data

external  
annotations

environment

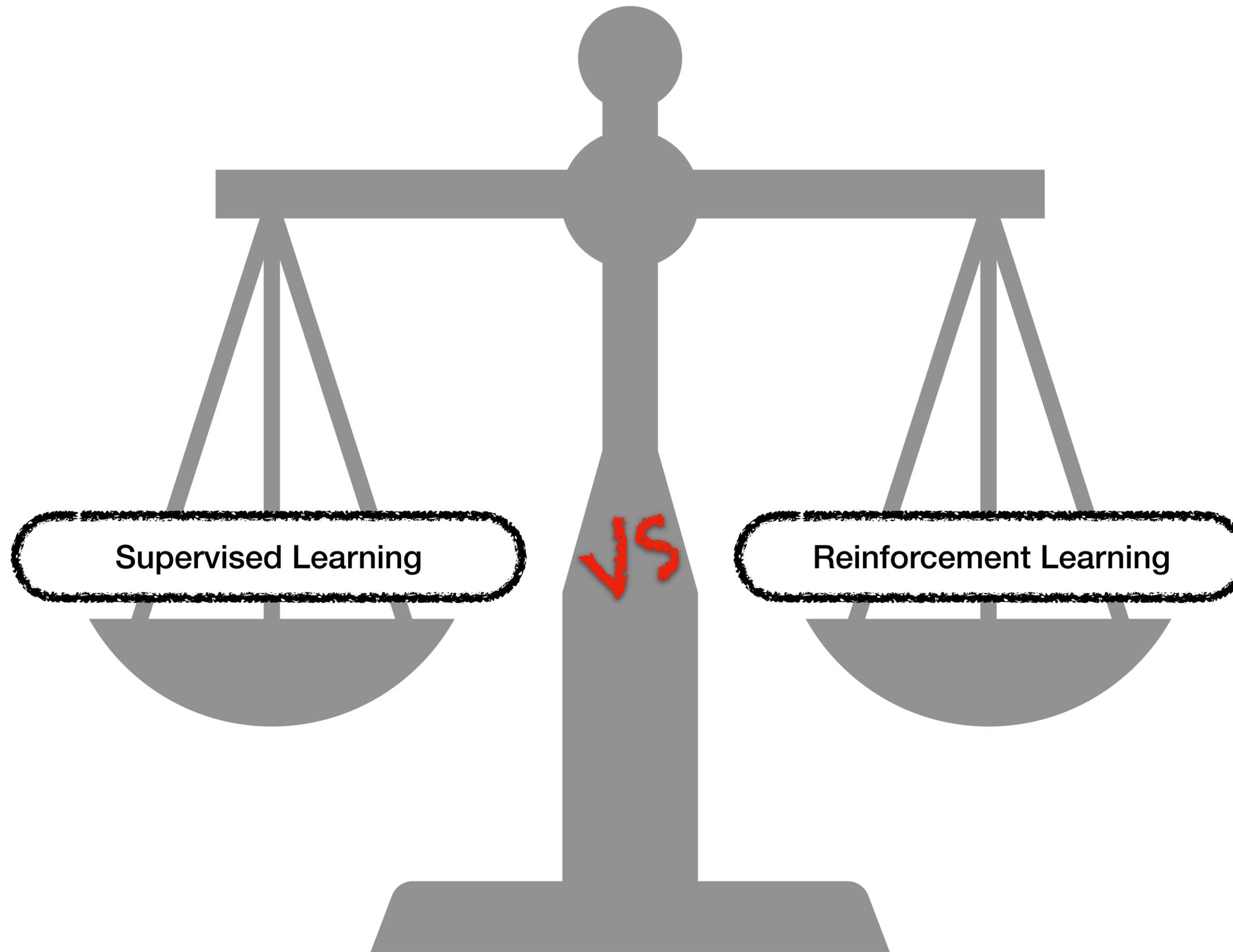
human  
preferences

Self-Supervised  
Fine-Tuning

Supervised  
Fine-Tuning  
(SFT)

Reinforcement  
Fine-Tuning  
(RFT)

Reinforcement Learning  
with Human Feedback  
(RLHF)



Supervised Learning

Reinforcement Learning

Objective ?

distance  
minimization

reward  
maximization

Learning Paradigms

Labels from...

Reward from...

internal  
structure of  
data

external  
annotations

environment

human  
preferences



Self-Supervised  
Fine-Tuning

Supervised  
Fine-Tuning  
(SFT)

Reinforcement  
Fine-Tuning  
(RFT)

Reinforcement Learning  
with Human Feedback  
(RLHF)

pre-training

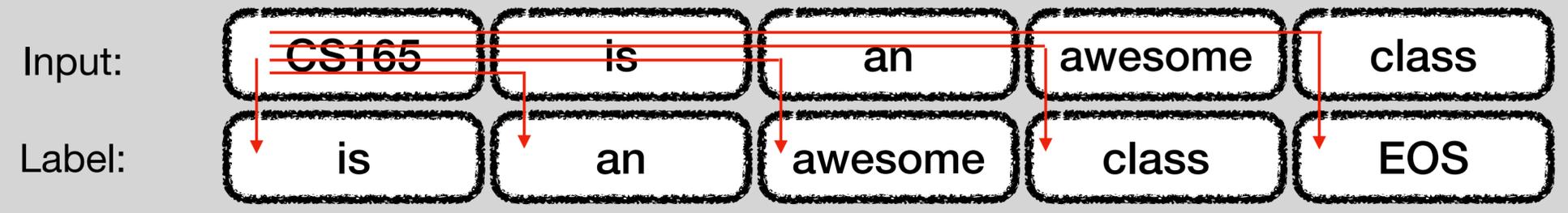


# Self-Supervised Fine-Tuning

Decoder Architecture

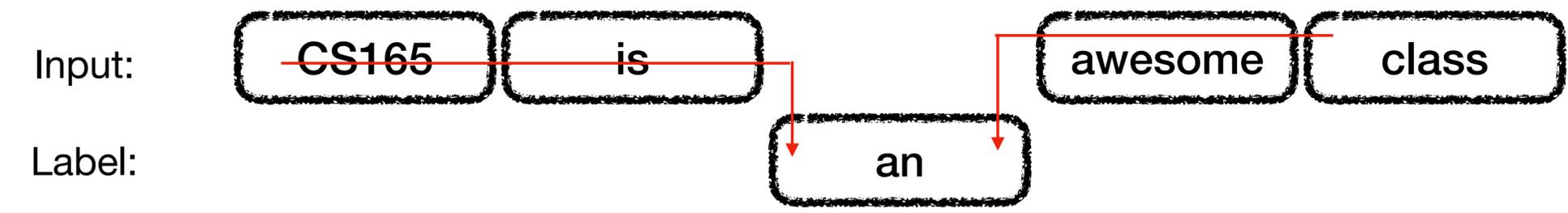
## Causal Language Modeling

(predict next token left → right)



Encoder Architecture

## Masked Language Modeling

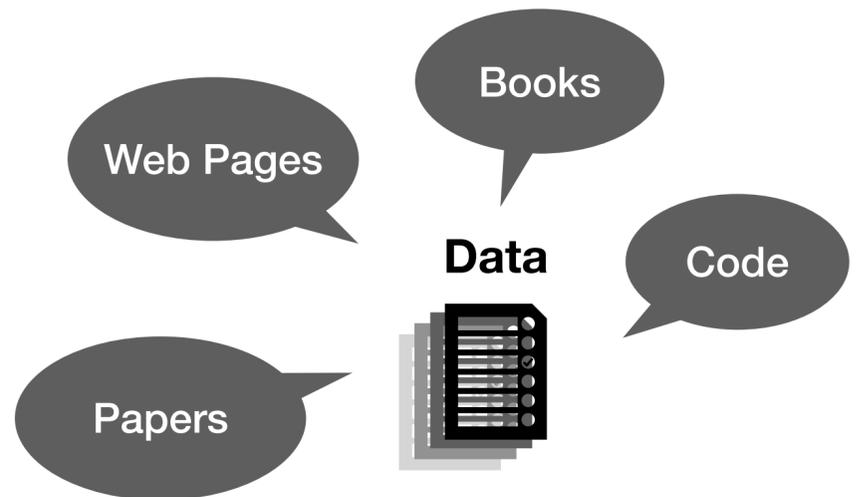


pre-training



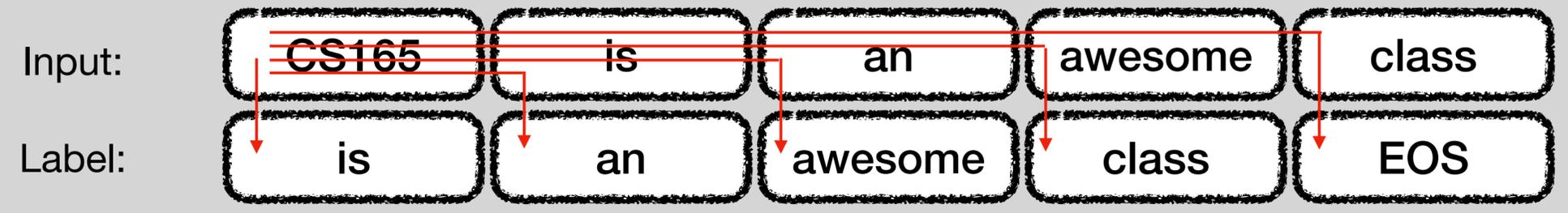
# Self-Supervised Fine-Tuning

## Decoder Architecture



### Causal Language Modeling

(predict next token left → right)



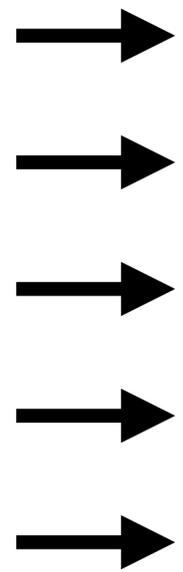
## pre-training



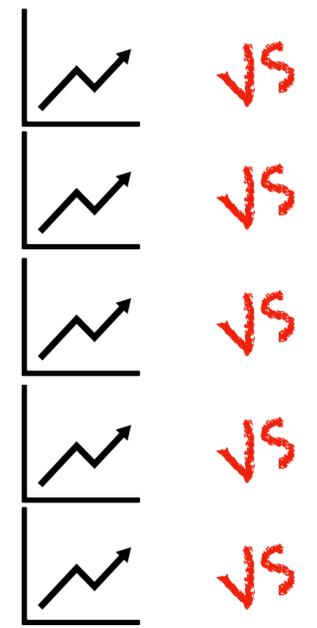
# Self-Supervised Fine-Tuning

Input:

- CS165
- is
- an
- awesome
- class



Decode:



Label:

- is
- an
- awesome
- class
- EOS



loss function

Likelihood

$$\begin{aligned}
 & p(\text{is, an, \dots, EOS}) \\
 &= p(x_1, x_2, \dots, x_N) \\
 &= \prod_{i=1}^N p(x_i | x_1, \dots, x_{i-1})
 \end{aligned}$$

Cross-Entropy

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N \log p(x_i)$$

Perplexity

$$e^{\mathcal{L}} = \exp \left( -\frac{1}{N} \sum_{i=1}^N \log p(x_i) \right)$$

pre-training

Supervised Learning

Reinforcement Learning

Objective ?

distance  
minimization

reward  
maximization

Learning Paradigms

Labels from...

Reward from...

internal  
structure of  
data

external  
annotations

environment

human  
preferences



Self-Supervised  
Fine-Tuning

Supervised  
Fine-Tuning  
(SFT)

Reinforcement  
Fine-Tuning  
(RFT)

Reinforcement Learning  
with Human Feedback  
(RLHF)

pre-training

**Learning Algorithm**

**Task/  
Dataset**

**LLM  
Fine-Tuning**

**Weight  
Update**

**Architecture**