

what is a language model?

what is a neural network language model?

what is GPT-3?

what is few-shot learning?

how to work with GPT-3 and few-shot learning?

brainstorm the topics

what is a language model?

1. Probability of a sentence $P(\text{"oops i did it again"})$
2. Probability of an upcoming word $P(\text{"again"} | \text{"oops i did it"})$

what is a language model?

n-gram LMs

$$P(i | \text{oops}) = ?$$

how many times “oops” was followed by “i” out of all times we've used “oops”?

$$P(\text{did} | \text{oops } i) = ?$$

how many times “oops i” was followed by “did” out of all times we've used “oops i”?

what is a language model?

n-gram LMs

$$P(i | \text{oops}) = 3/5$$

$$P(\text{did} | \text{oops } i) = 2/3$$

oops i

oops you

oops we did it

oops i did it

oops i did it again

what is a language model?

n-gram LMs

Markov assumption:

$P(\text{blah} \mid \text{my next word in this long sentence will be})$

=

$P(\text{blah} \mid \text{sentence will be})$

=

$P(\text{blah} \mid \text{will be})$

what is a language model?

n-gram LMs

$P(\text{my next word in this long sentence will be blah})$

=

$P(\text{my}|\text{START}) * P(\text{next}|\text{my}) * P(\text{word}|\text{next}) * P(\text{in}|\text{word}) * P(\text{this}|\text{in}) * P(\text{long}|\text{this}) * P(\text{sentence}|\text{long}) * P(\text{will}|\text{sentence}) * P(\text{be}|\text{will}) * P(\text{blah}|\text{be}) * P(\text{END}|\text{blah})$

what is a neural network language model?

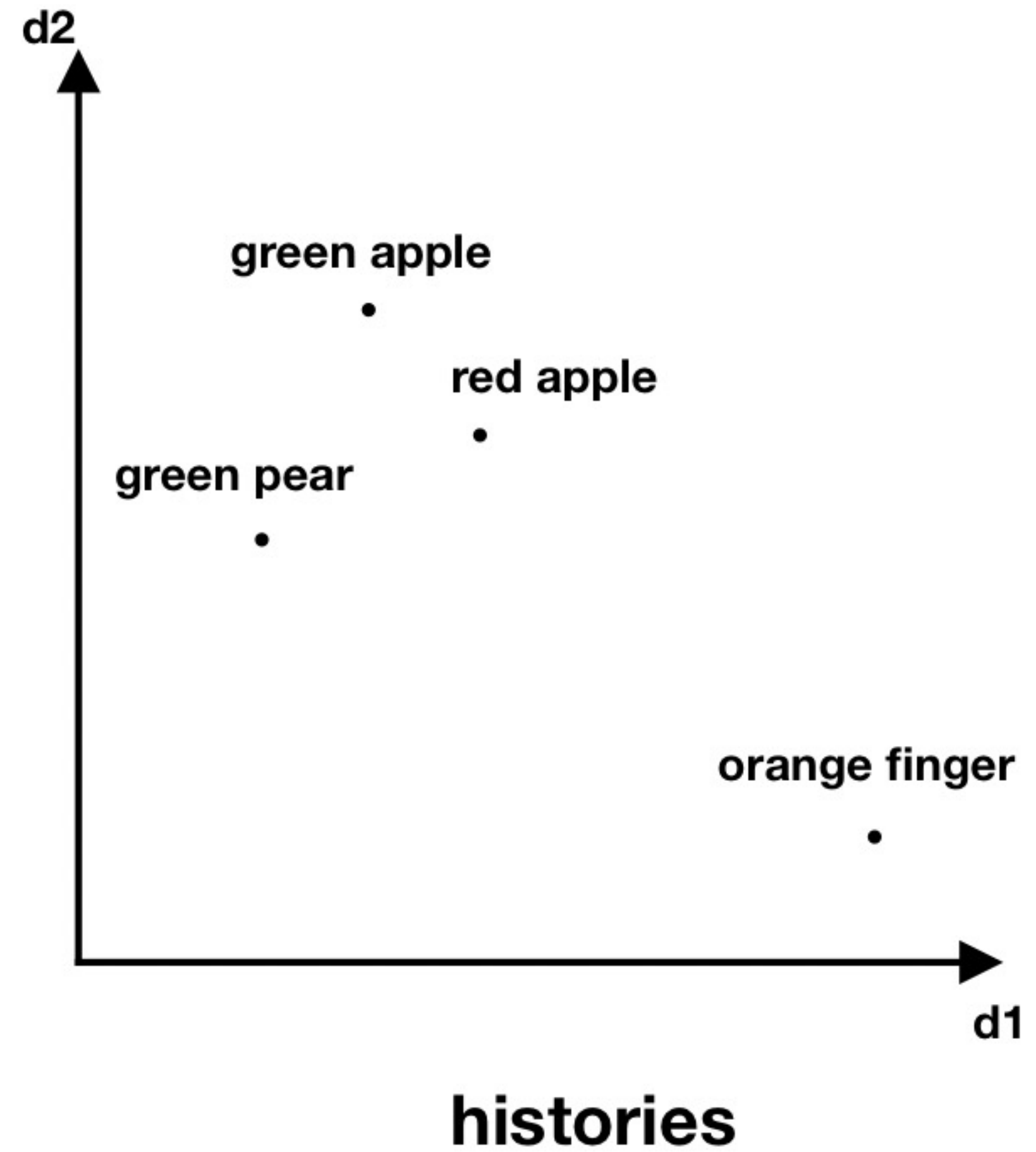
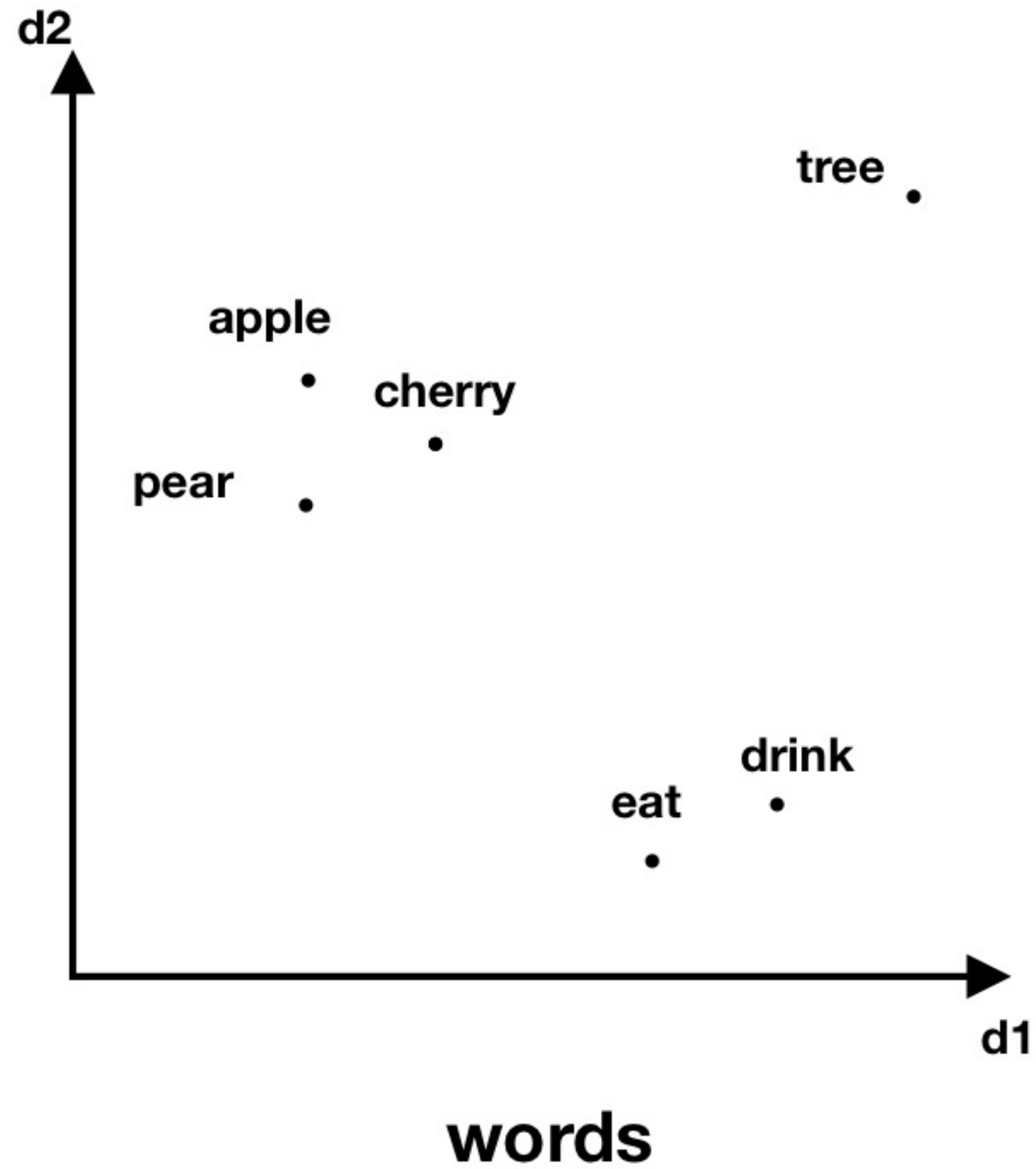


what is a neural network language model?

$V = [\text{oops, i, did, it, again, yesterday}]$



what is a neural network language model?



what is GPT-3?

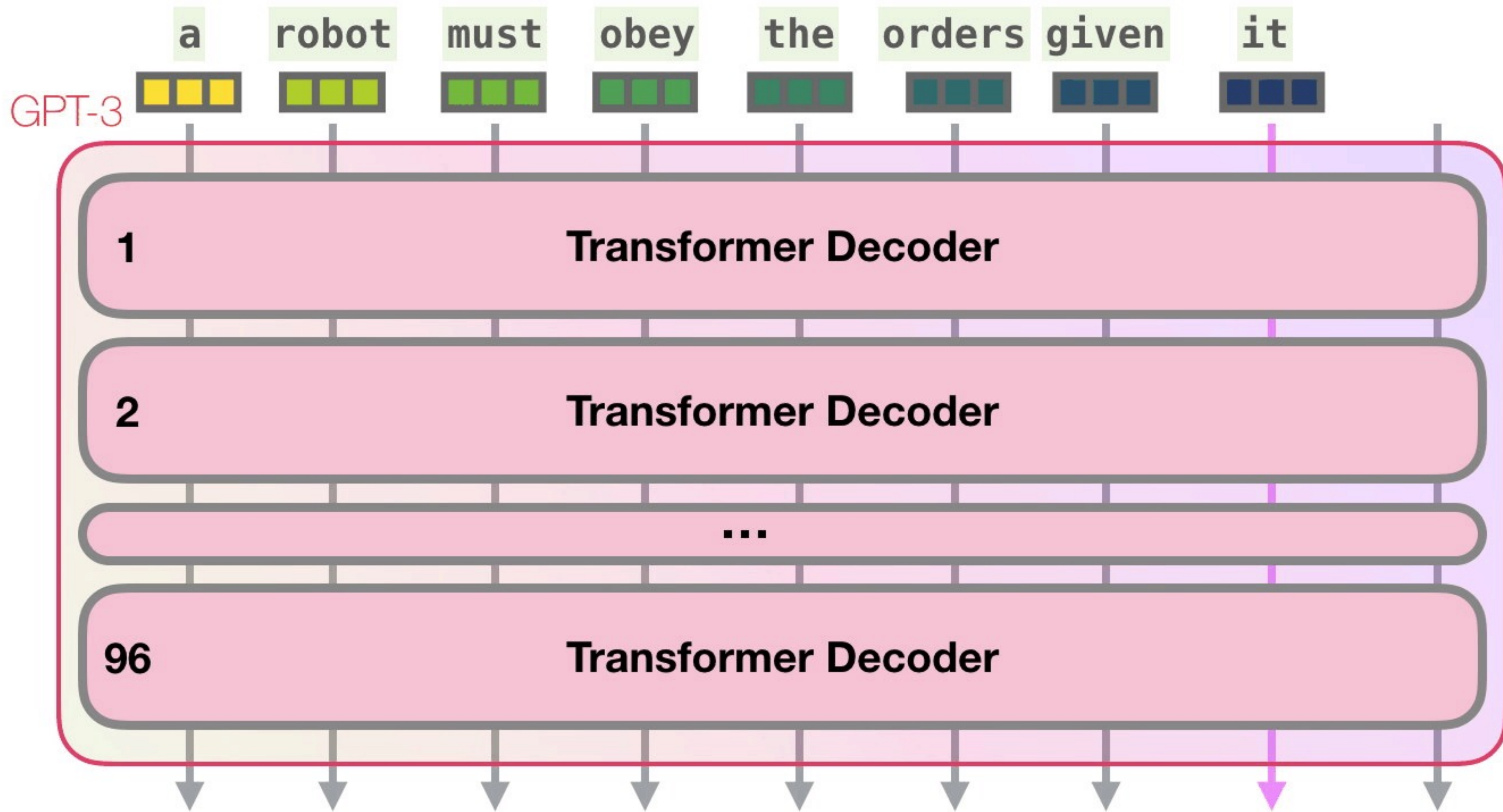
Input Prompt:

Recite the first law of robotics

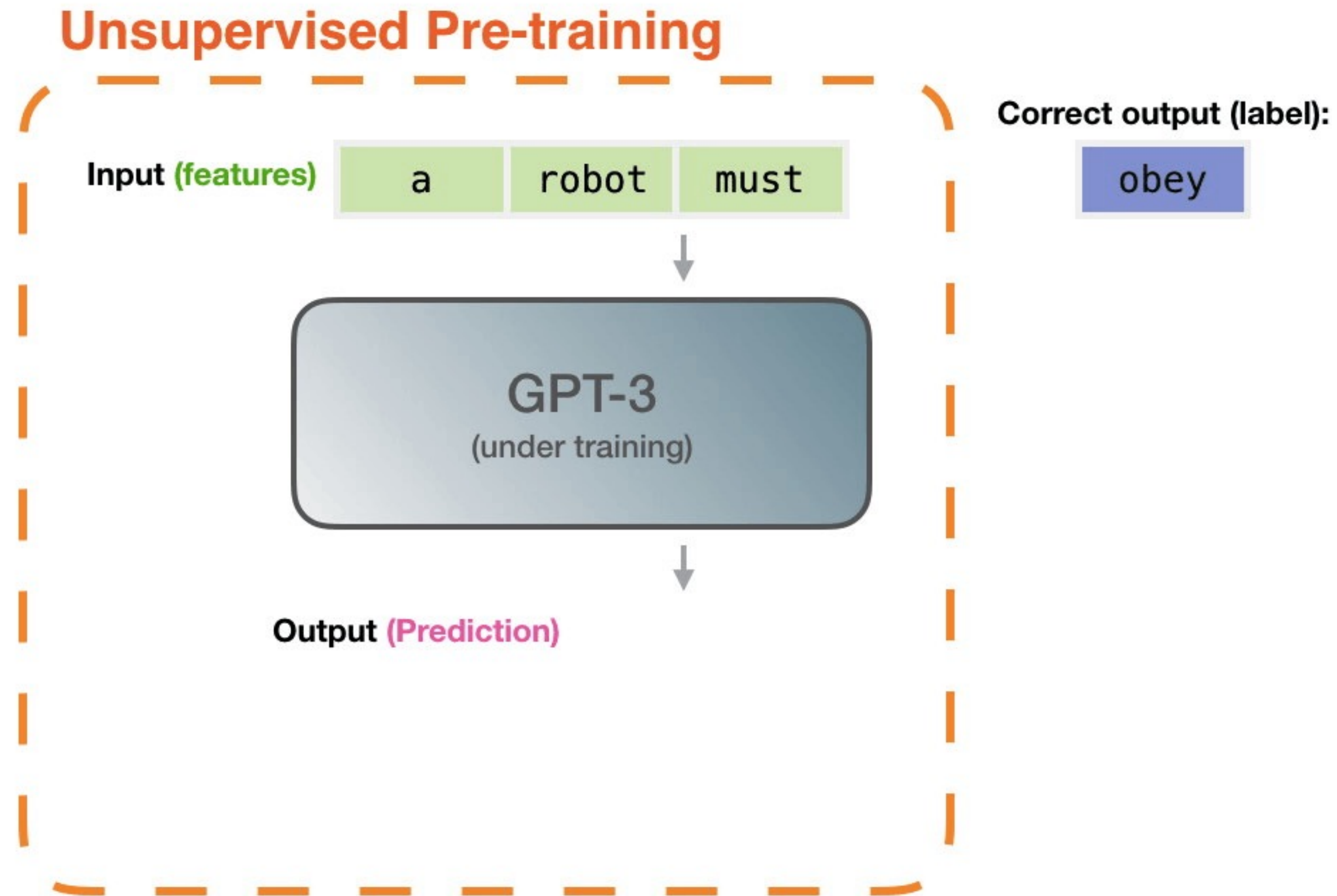


Output:

what is GPT-3?



what is few-shot learning?



what is few-shot learning?

Fine-tuning

The model is trained via repeated gradient updates using a large corpus of example tasks.



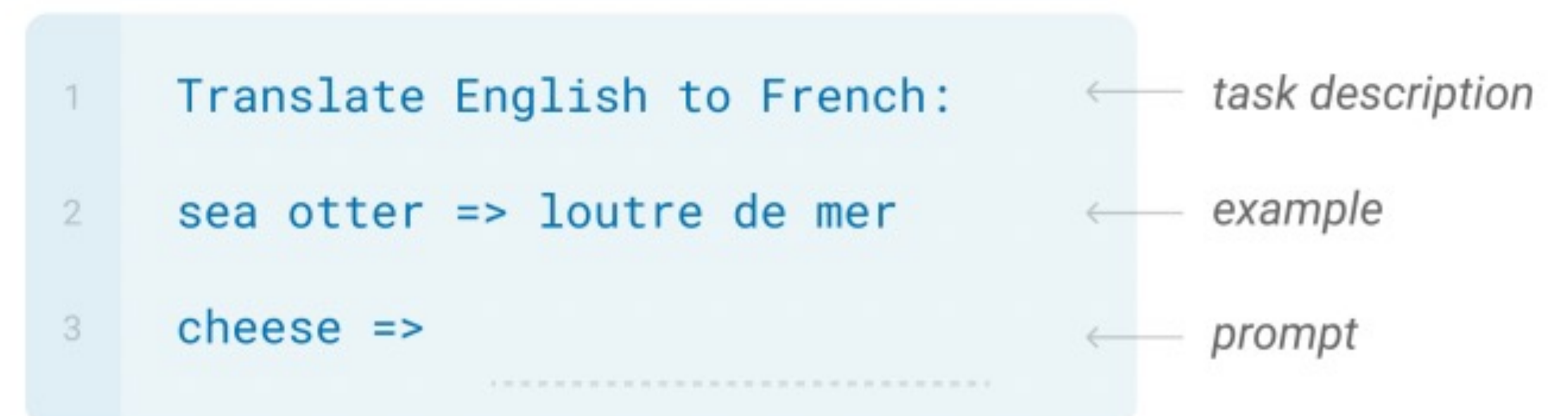
Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.



One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.



what is few-shot learning?

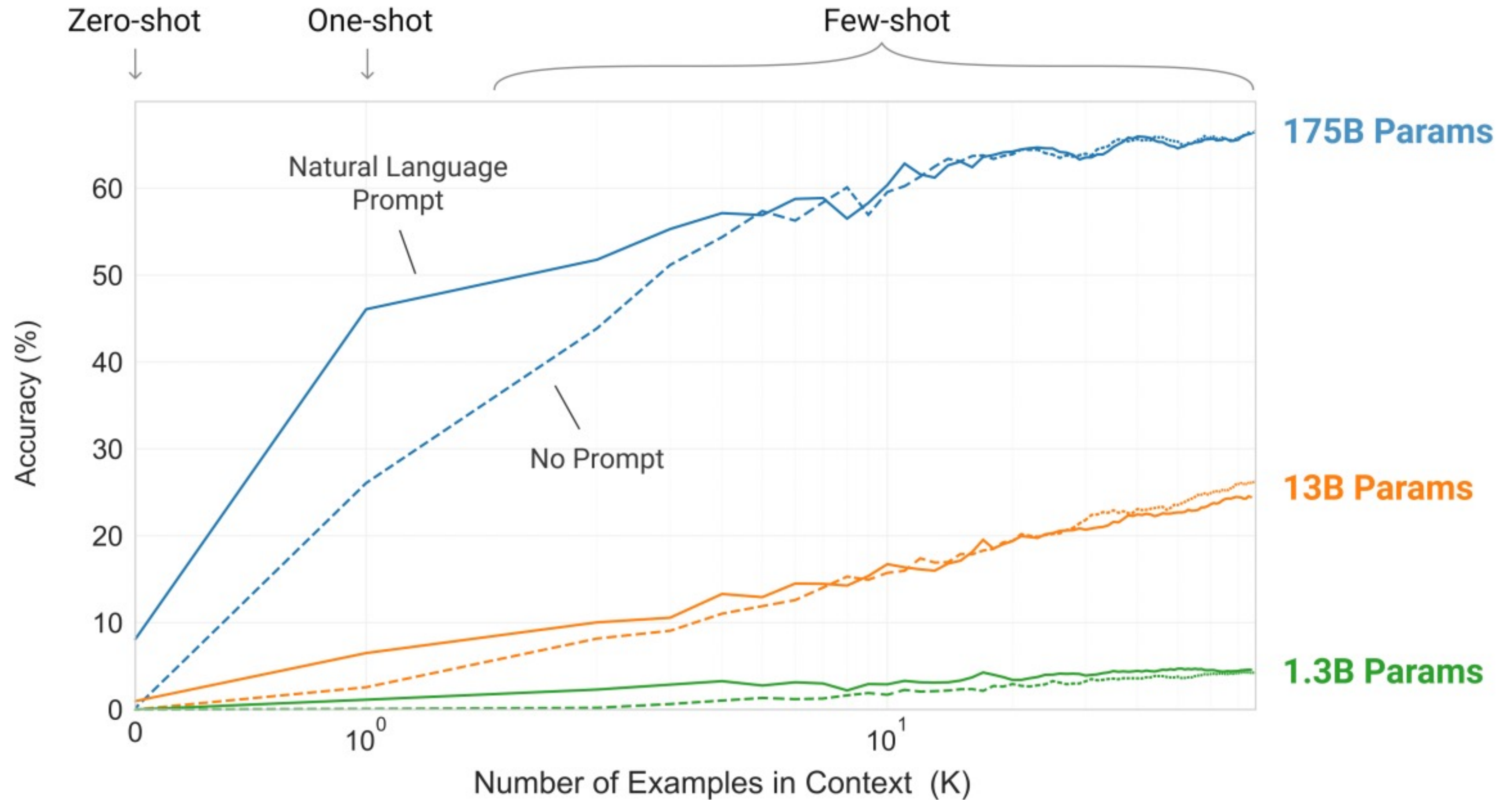
Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

The diagram shows a prompt structure for a few-shot learning task. It consists of five lines of text, each with a line number on the left. The first line is the task description. The next three lines are examples of the task. The fifth line is the prompt to be completed. Arrows on the right point to each line, with labels: 'task description' for line 1, 'examples' for lines 2-4, and 'prompt' for line 5.

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← examples
3 peppermint => menthe poivrée ←
4 plush girafe => girafe peluche ←
5 cheese => ..... ← prompt
```

what is few-shot learning?



how to work with GPT-3 and few-shot learning?

**PRE-SOFTWARE:
SPECIAL-PURPOSE
COMPUTER**



**SOFTWARE 1.0:
DESIGN
THE ALGORITHM**



**SOFTWARE 2.0:
DESIGN
THE DATASET**



**SOFTWARE 3.0:
DESIGN
THE PROMPT**



topics

generate names for emo bands

finish limericks

generate movie titles based on movie descriptions