How LLMs work

Andrea López Incera, Mirte van der Eyden
Universität Innsbruck

Teaching in AI (TAI)
Erasmus+ Project
Funded by the European Union

Goal of the LLM

Predict the next word in an incomplete sentence

I crossed the street through the **zebra** ?

Goal of the LLM

Predict the next word in an incomplete sentence

Predicted word

I crossed the street through the **zebra** crossing

Context

Focus word

In practice…

Context

"Write a story in first person about a woman in a city"

It was a cold morning, the skyscrapers rose in the middle of the fog. I needed some air, so I decided to go for a walk. I left my apartment, I crossed the street through the zebra crossing

In practice…

Context

"Write a story in first person about a woman in a city"

It was a cold morning, the skyscrapers rose in the middle of the fog. I needed some air, so I decided to go for a walk. I left my apartment, I crossed the street through the zebra crossing ,

In practice…

Context

"Write a story in first person about a woman in a city"

It was a cold morning, the skyscrapers rose in the middle of the fog. I needed some air, so I decided to go for a walk. I left my apartment, I crossed the street through the zebra crossing, and

In practice…

Context

"Write a story in first person about a woman in a city"

It was a cold morning, the skyscrapers rose in the middle of the fog. I needed some air, so I decided to go for a walk. I left my apartment, I crossed the street through the zebra crossing, and I

In practice…

Context

"Write a story in first person about a woman in a city"

It was a cold morning, the skyscrapers rose in the middle of the fog. I needed some air, so I decided to go for a walk. I left my apartment, I crossed the street through the zebra crossing, and I started

In practice…

Context

"Write a story in first person about a woman in a city"

It was a cold morning, the skyscrapers rose in the middle of the fog. I needed some air, so I decided to go for a walk. I left my apartment, I crossed the street through the zebra crossing, and I started

to

In practice…

Context

"Write a story in first person about a woman in a city"

It was a cold morning, the skyscrapers rose in the middle of the fog. I needed some air, so I decided to go for a walk. I left my apartment, I crossed the street through the zebra crossing, and I started to feel

In practice…

Context

"Write a story in first person about a woman in a city"

It was a cold morning, the skyscrapers rose in the middle of the fog. I needed some air, so I decided to go for a walk. I left my apartment, I crossed the street through the zebra crossing, and I started to feel better

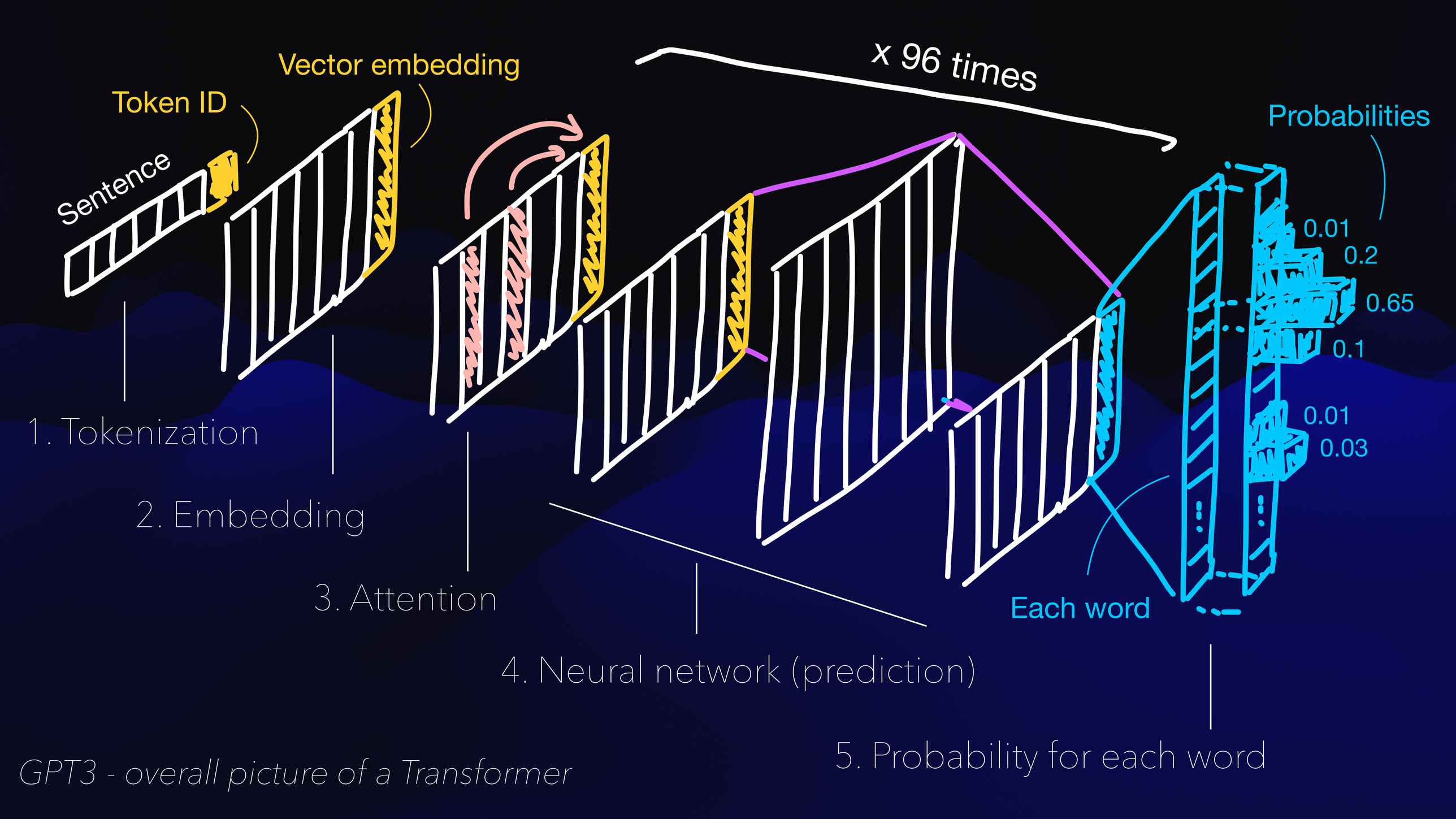
In practice…

Context

"Write a story in first person about a woman in a city"

It was a cold morning, the skyscrapers rose in the middle of the fog. I needed some air, so I decided to go for a walk. I left my apartment, I crossed the street through the zebra crossing, and I started to feel better .

How does it work?



GPT3 - overall picture of a Transformer

1. Tokenization



I crossed the street through the zebra

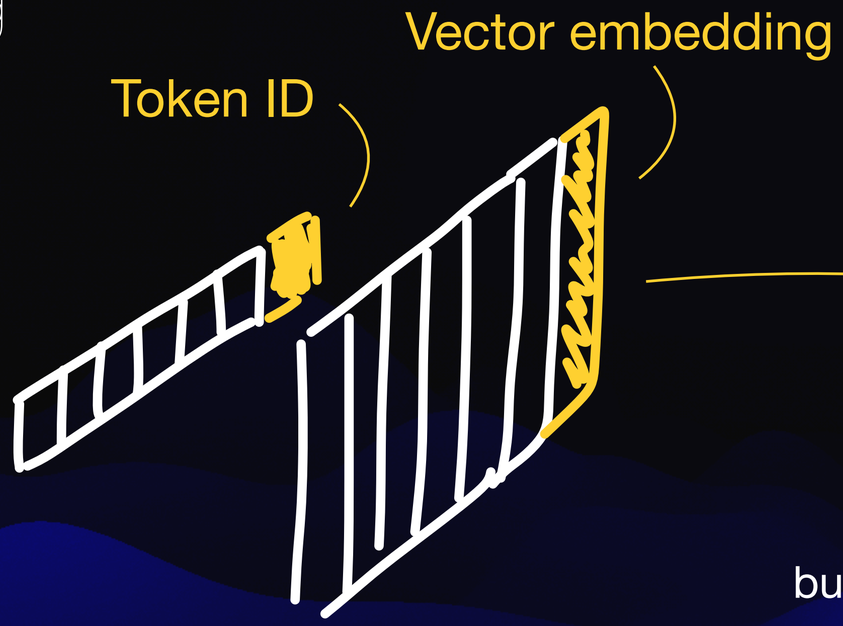
Token IDs: 14 105 30 3 875 567 4 95

Total number of tokens: 50k

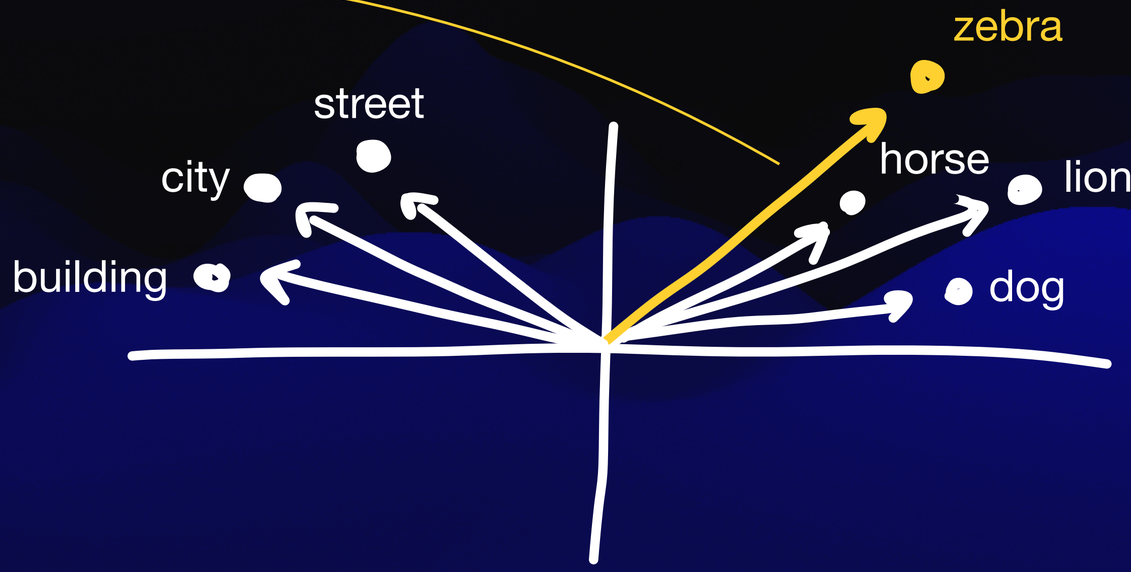
*This is an example, the actual ID numbers may be different

2. Embedding

(Learned)



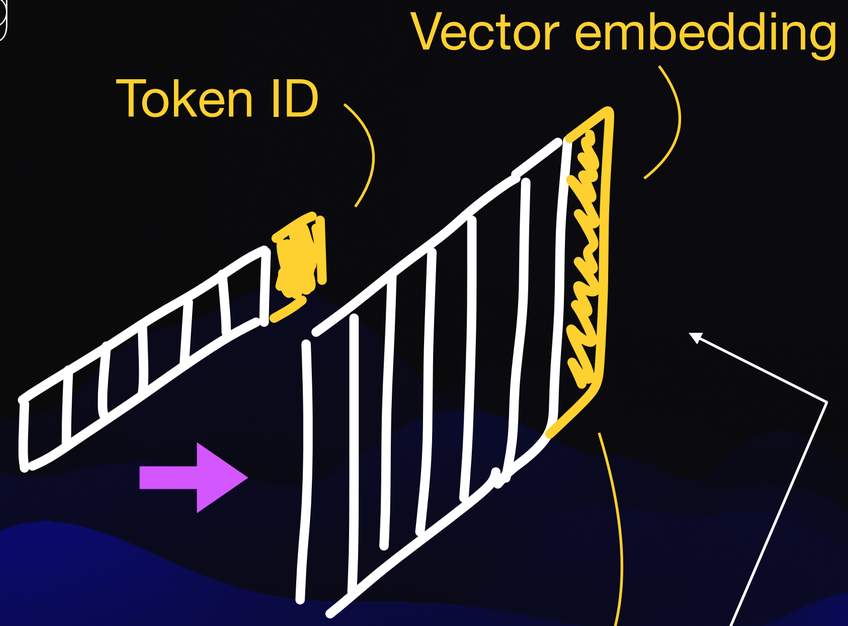
Words are clustered
by meaning



**Visualization in 2D: actual space has 12.288 dimensions!*

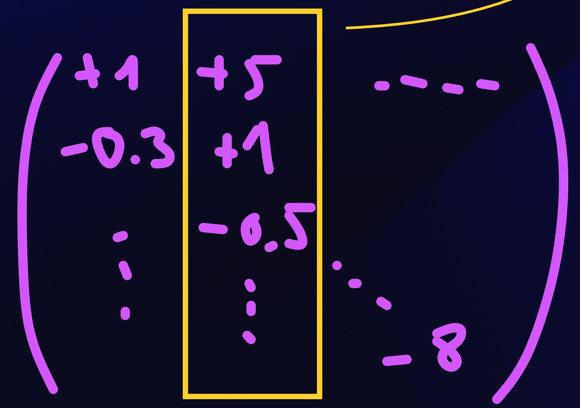
2. Embedding

(Learned)



50k (all tokens)

12.288 numbers (Parameters)

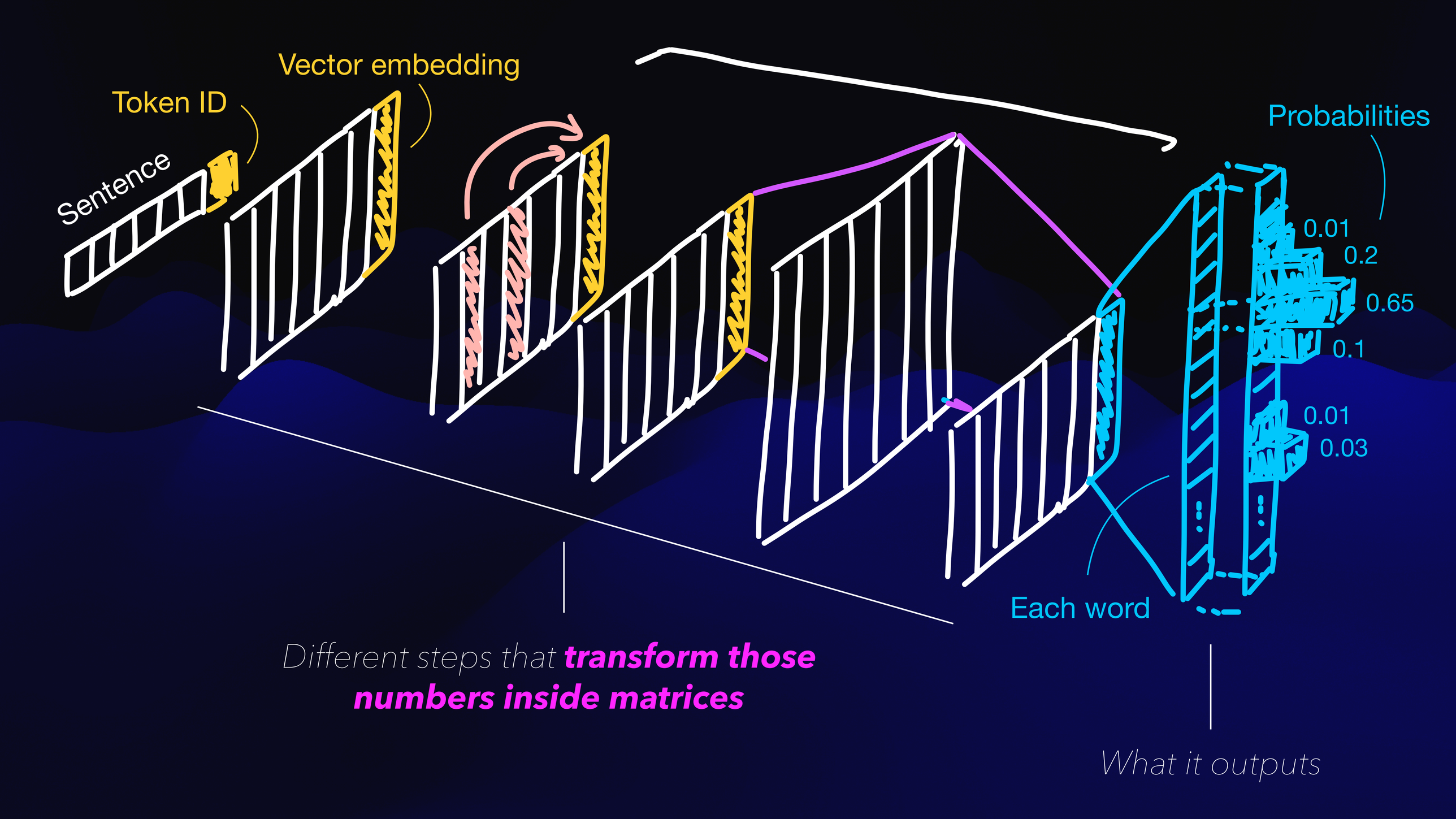


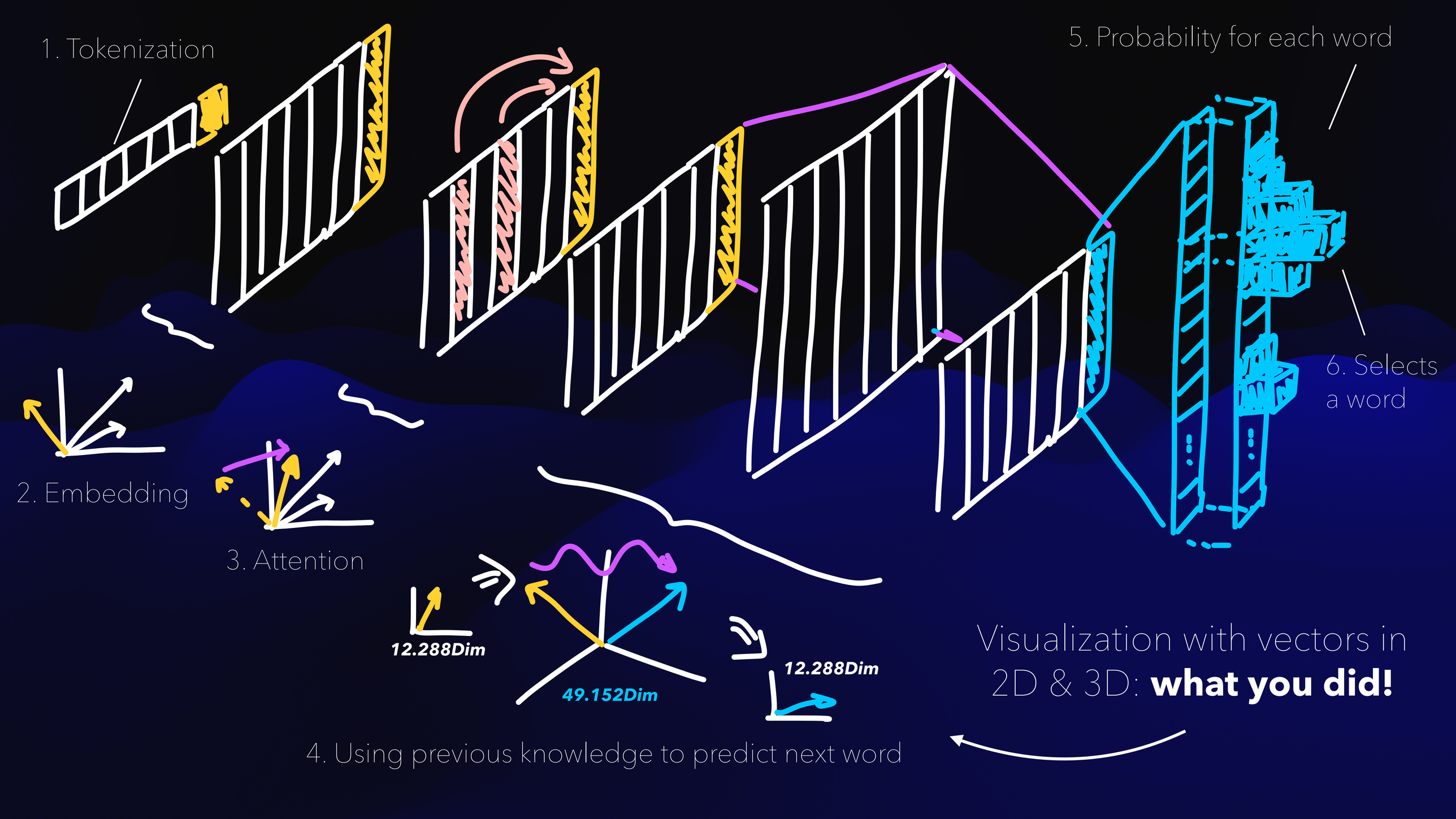
Token ID ⟶ Vector

WHAT THE MACHINE SEES

(and changes with learning):

- Numbers arranged in **matrices**
- Each column is a ***vector*** (numbers that indicate where to place the point in the grid)


Vector embedding
Token ID
Sentence
Probabilities
0.01
0.2
0.65
0.1
0.01
0.03
Each word
Different steps that **transform those numbers inside matrices**
What it outputs


1. Tokenization
5. Probability for each word
6. Selects a word
2. Embedding
3. Attention
12.288Dim
49.152Dim
12.288Dim
4. Using previous knowledge to predict next word
Visualization with vectors in 2D & 3D: **what you did!**

Predicted word

I crossed the street through the **zebra** crossing

How does it learn?

With (MANY) examples!
from the entire internet

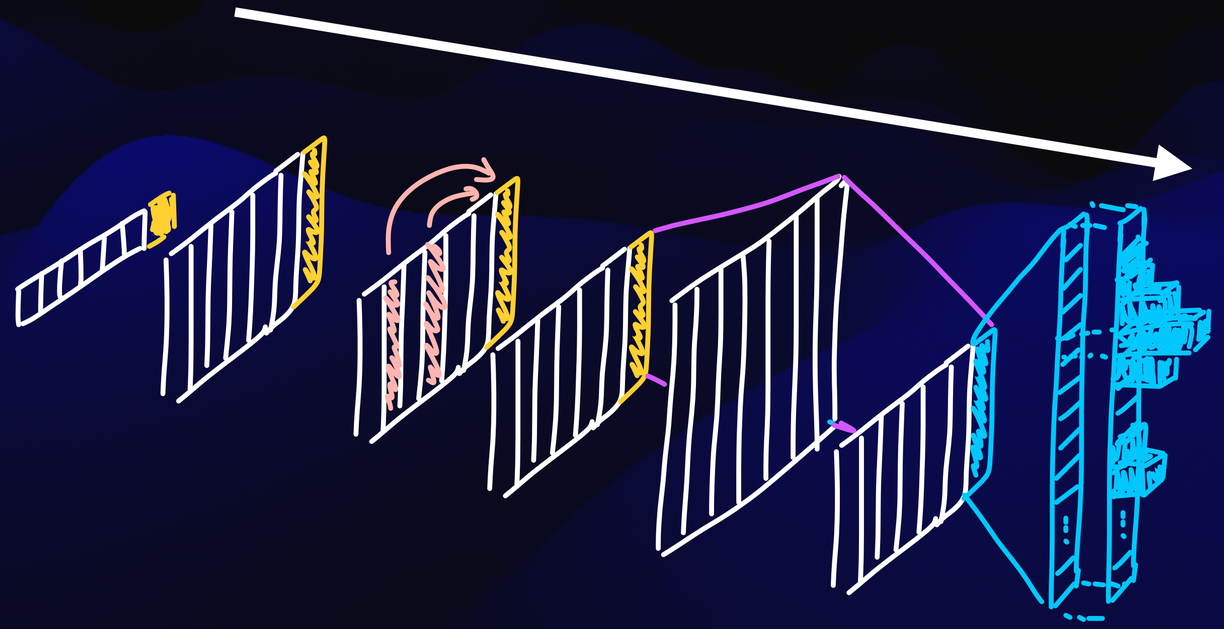
Example text

The candles drifted higher as the hall filled with whispers, their flames trembling like they were listening. A young wizard tightened his grip on his **wand**, feeling magic hum beneath his skin as the doors slowly creaked open. Somewhere above, a staircase shifted, rearranging itself as if amused by the newcomers' nerves. It felt like the castle itself was watching, quietly deciding what kind of night this would be.

With (MANY) examples!
from the entire internet

Example text

The candles drifted higher as the hall filled with whispers, their flames trembling like they were listening. A young wizard tightened his grip on his **wand**.

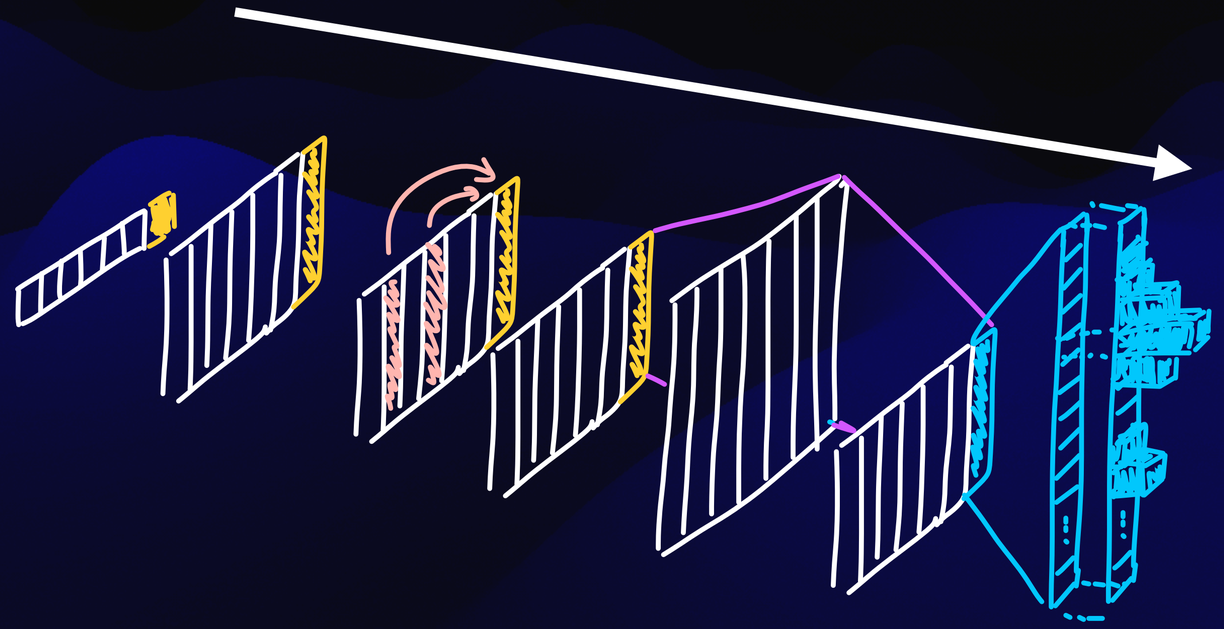


Do one pass: predict a word

With (MANY) examples!
from the entire internet

Example text

The candles drifted higher as the hall filled with whispers, their flames trembling like they were listening. A young wizard tightened his grip on his **wand**, feeling magic hum beneath his skin as the doors slowly creaked open. Somewhere above, a staircase shifted, rearranging itself as if amused by the newcomers' nerves. It felt like the castle itself was watching, quietly deciding what kind of night this would be.



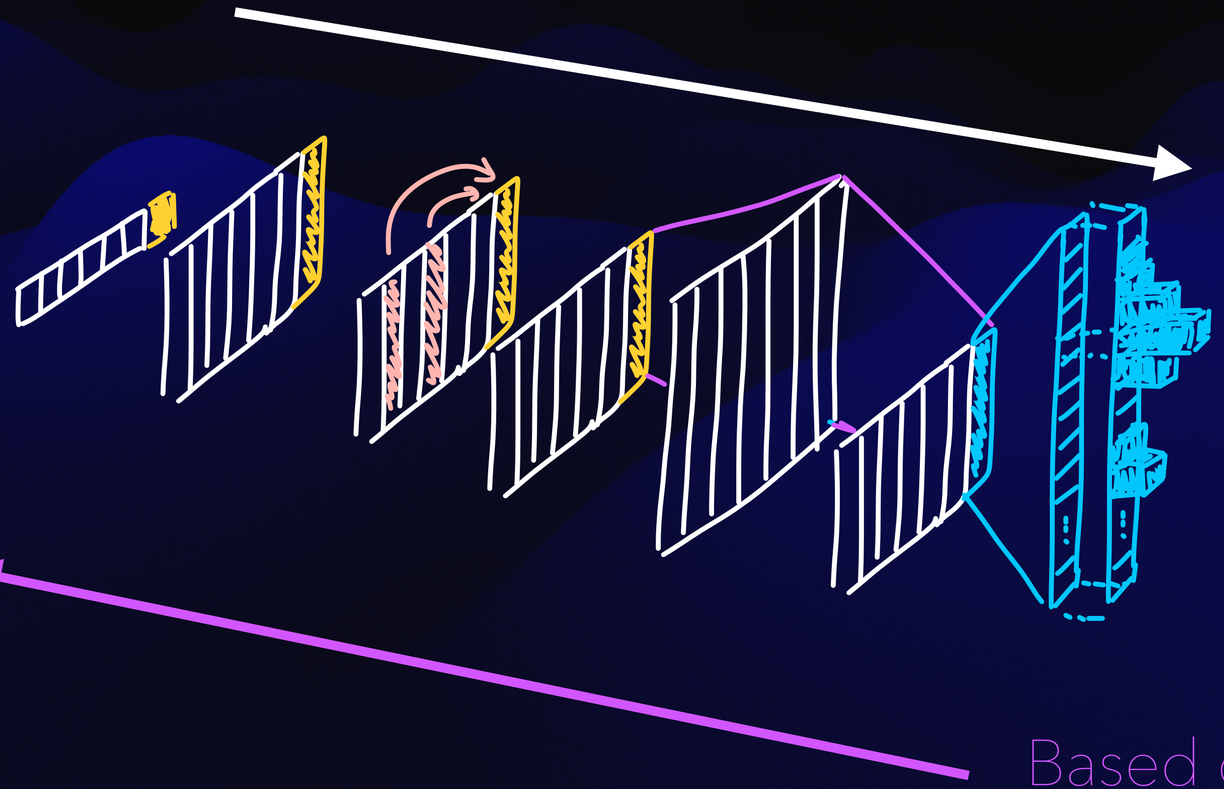
Do one pass: predict a word

Check how well it did

With (MANY) examples!
from the entire internet

Example text

The candles drifted higher as the hall filled with whispers, their flames trembling like they were listening. A young wizard tightened his grip on his **wand**, feeling magic hum beneath his skin as the doors slowly creaked open. Somewhere above, a staircase shifted, rearranging itself as if amused by the newcomers' nerves. It felt like the castle itself was watching, quietly deciding what kind of night this would be.



Do one pass: predict a word

Check how well it did

Based on that, update ALL parameters
(change pink matrices with numbers)

+ Reinforcement Learning with human feedback



People check the answers and give a positive reward
if they feel they "sound human"

GPT 3 *173.181.291.520 parameters (173 billion!)*

GPT 5 ***1,7 - 1,8 trillion (estimate - not published)***

If I stack 1,7 trillion sheets of paper, I get halfway to the Moon!

A lot of energy to train them!
(GPT 3: 1.287 MWh = approx. 120 US homes for a year)*

** Data from here*

Current models also use **external access to the internet, or to other programs** like Python to do maths, or more complex tasks.

References

Material inspired by this series of videos from 3Blue1Brown

Cho et al., "Transformer Explainer: Interactive Learning of Text-Generative Models", arXiv:2408.04619. Simulation and explanation, with video here .

Vaswani et al., "Attention is all you need", arXiv:1706.03762.