



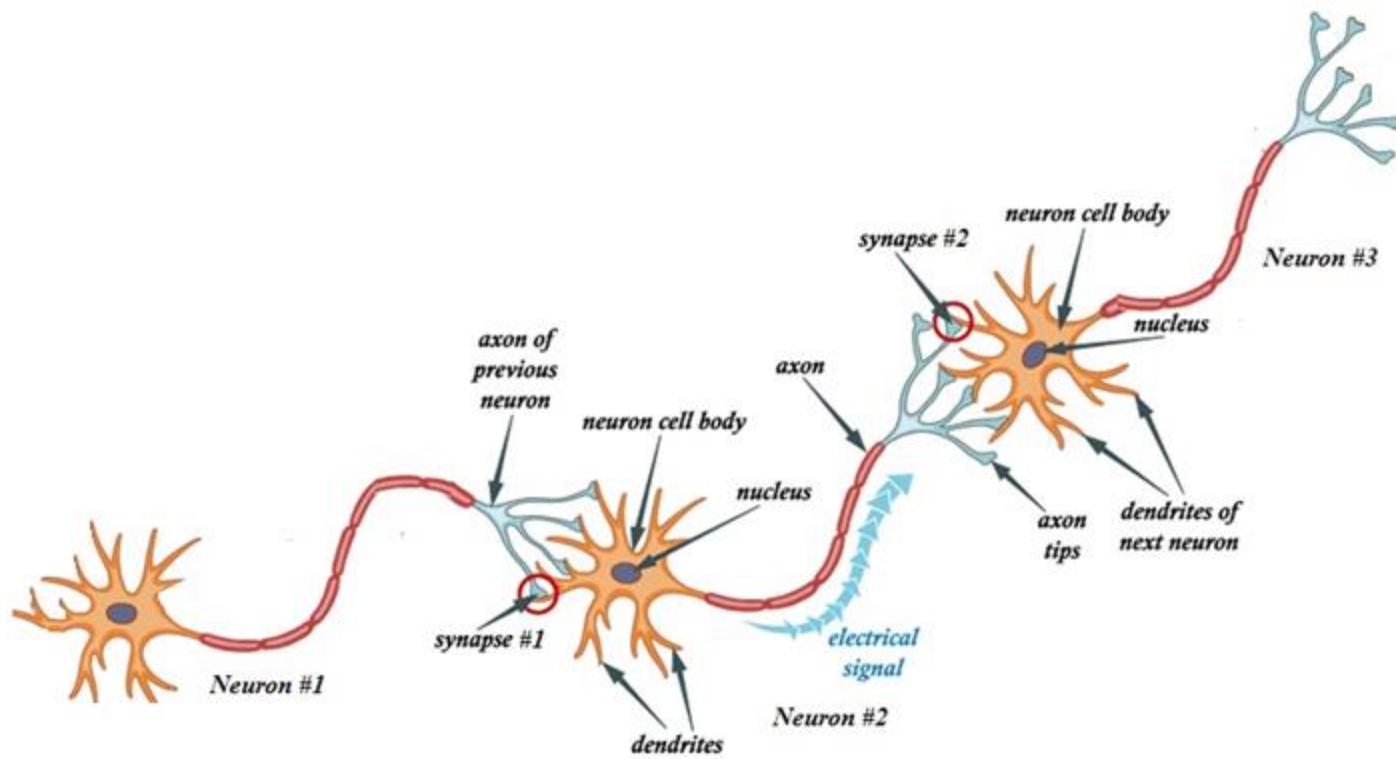
Big Data Management, Analysis, and Presentation

*LLMs and Prompt Engineering
for Data Analysis*

Andrea Brunello
andrea.brunello@uniud.it



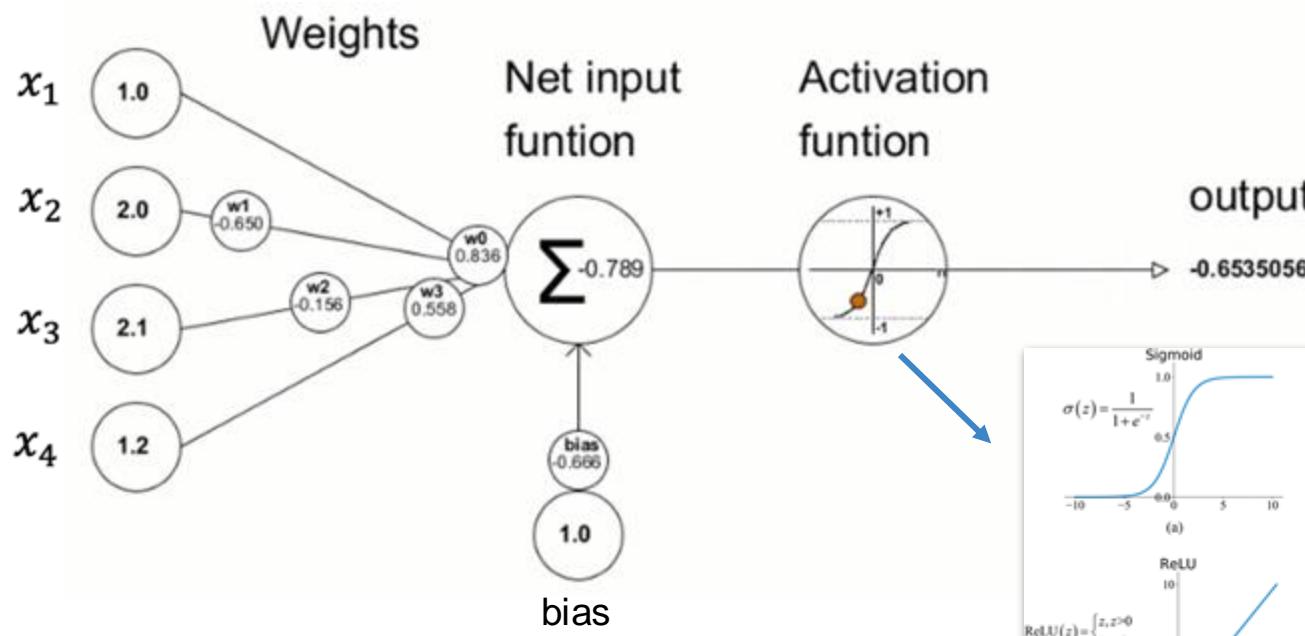
- The human brain contains approximately 86 billion neurons, connected by 10^{14} - 10^{15} synapses
- Dendrites receive signals from other neurons; the cell processes the received signals and, if certain conditions are met, it generates an output signal through the axon and synapses



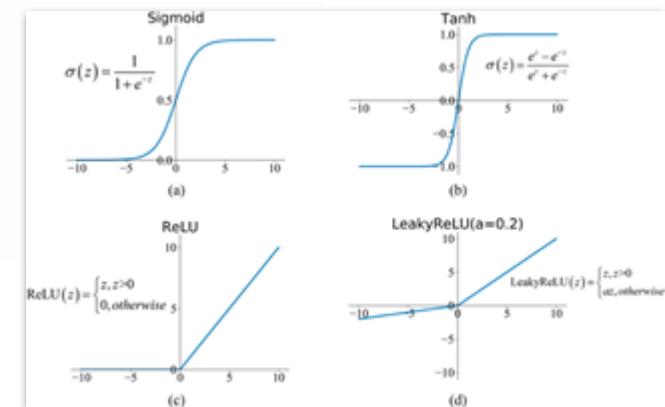
- A neuron receives, as its input, a set of numerical quantities

x_1, x_2, x_3, x_4 in our case here 1.0, 2.0, 2.1, 1.2

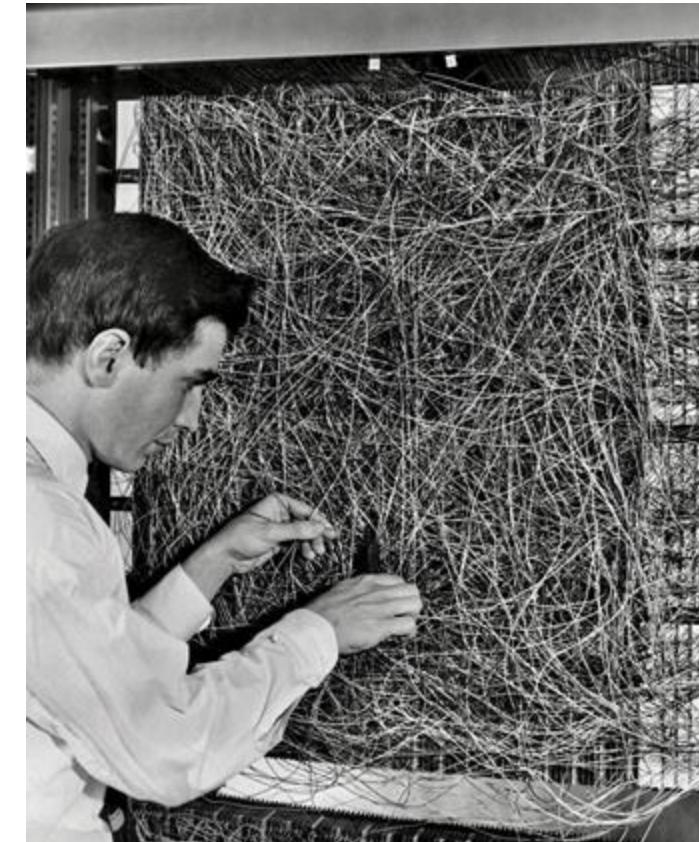
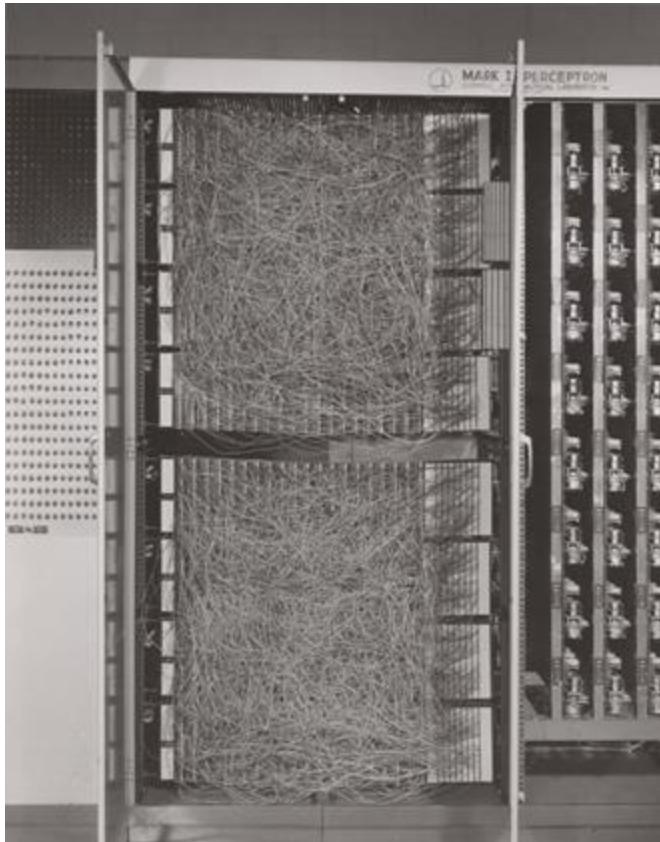
- Each quantity gets multiplied by its respective weight: $x_1 \cdot w_1, x_2 \cdot w_2, x_3 \cdot w_3, x_4 \cdot w_4$
- We sum all those products together: $x_1 \cdot w_1 + x_2 \cdot w_2 + x_3 \cdot w_3 + x_4 \cdot w_4$
- The result of the sum is passed through a so called “activation function”, which ultimately determines the output of the neuron



Bias and weights are called **parameters**

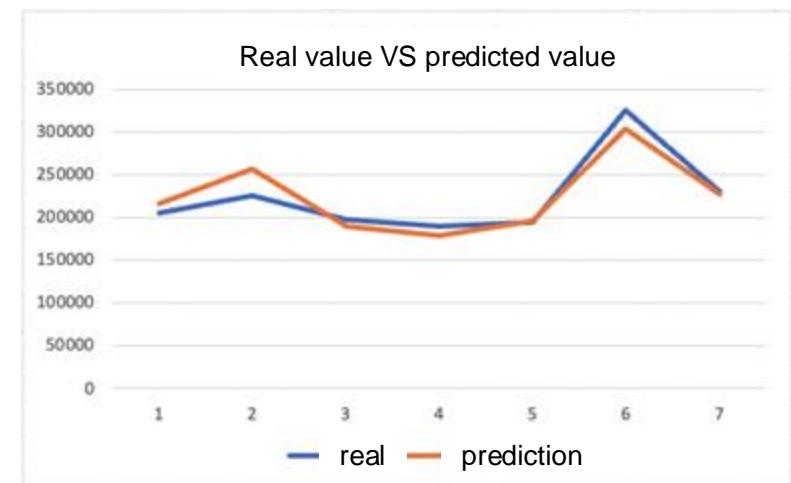
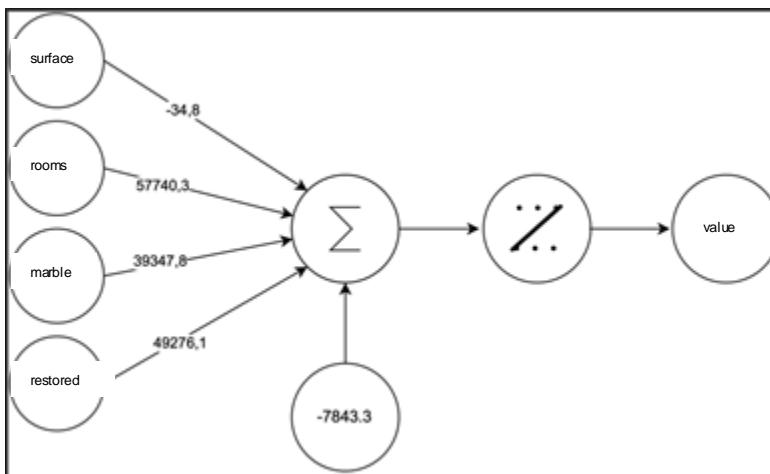


- First implementations, in the 50s and 60s, were hardware-based
- Connections built with cables, numbers encoded by electrical signals, weights implemented using resistors and potentiometers
- Nowadays, neural networks are software simulated within a computer



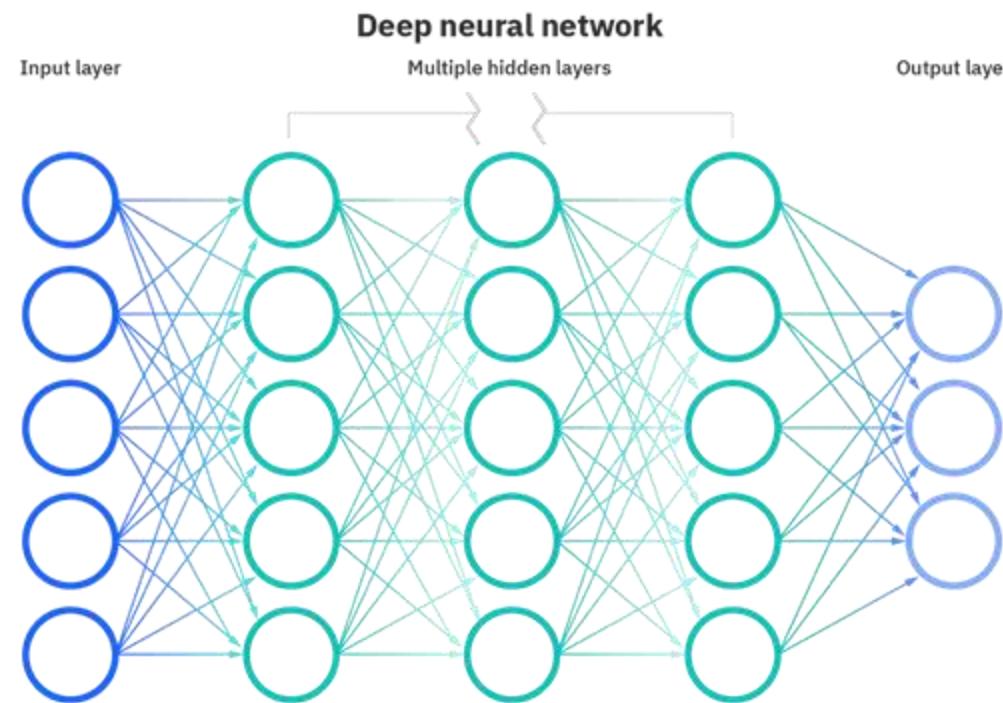
- A neural network is presented with a series of training examples related to the concept it aims to learn
- Each example consists of inputs associated with an output
- The goal is to learn weights within the neural network that allow for the construction of a mapping between the witnessed inputs and outputs
- The neural network thus learns a mathematical function between numerical inputs and outputs (here, a simple linear combination)

	surface	rooms	marble	restored	value
1	3529	6	0	0	205000
2	3247	5	1	1	224900
3	4032	5	0	1	197900
4	2397	4	1	0	189900
5	2200	4	0	1	195000
6	3536	6	1	1	325000
7	2983	5	0	1	230000



$$\begin{aligned} \text{value} = & - 34.8 * \text{surface} \\ & + 57740.3 * \text{rooms} \\ & + 39347.8 * \text{marble} \\ & + 49276.1 * \text{restored} \\ & - 7843.3 \end{aligned}$$

- There are strict limitations to what can be done with a single neuron
- More neurons, arranged into multiple layers (deep neural network)
- The architecture (how the neurons are arranged) is pre-designed, only the weights are learnt during training



[Interactive examples](#)

[Neural network zoo](#)

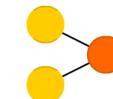
→ There are many types and architectures of neural networks, depending on the task to be tackled (such as image recognition, language translation, audio synthesis...)

- Input Cell
- Backfed Input Cell
- △ Noisy Input Cell
- Hidden Cell
- Probabilistic Hidden Cell
- △ Spiking Hidden Cell
- Capsule Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- △ Gated Memory Cell
- Kernel
- Convolution or Pool

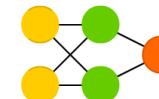
A mostly complete chart of Neural Networks

©2019 Fjodor van Veen & Stefan Leijnen asimovinstitute.org

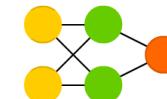
Perceptron (P)



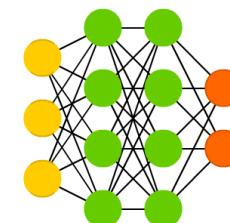
Feed Forward (FF)



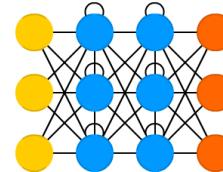
Radial Basis Network (RBF)



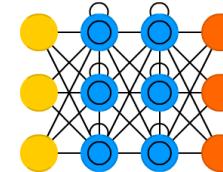
Deep Feed Forward (DFF)



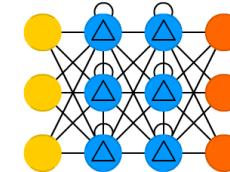
Recurrent Neural Network (RNN)



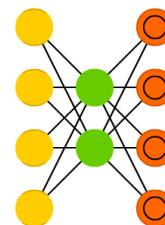
Long / Short Term Memory (LSTM)



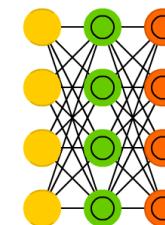
Gated Recurrent Unit (GRU)



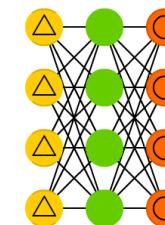
Auto Encoder (AE)



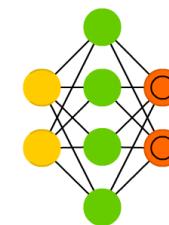
Variational AE (VAE)



Denoising AE (DAE)



Sparse AE (SAE)



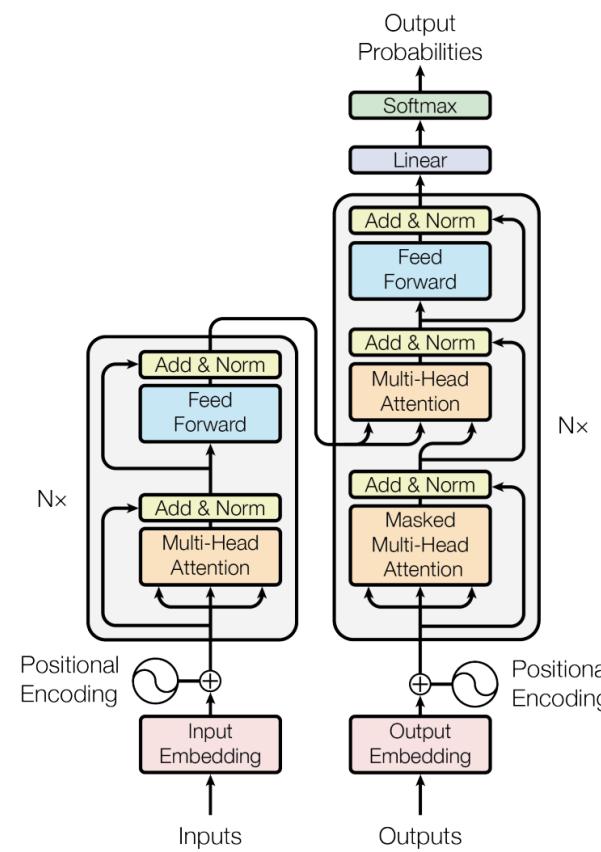
→ Among the various architectures, there is the Transformer, the revolutionary model underlying all modern LLMs (and more)

BERT

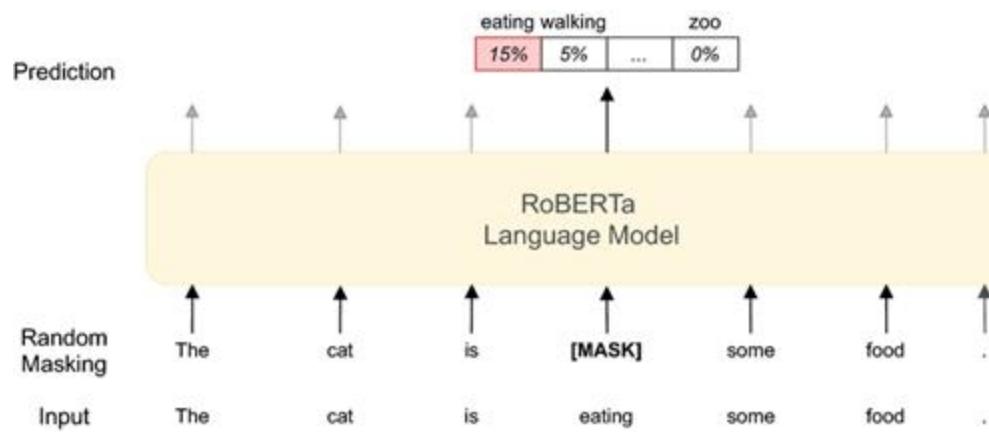
Encoder

GPT

Decoder



- A language model (LLM, Large Language Model) aims to capture the structure and rules of natural language, such as English
- To train a language model, a large set of texts containing a wide variety of natural language sentences is used
- The algorithm analyzes the text corpus and tries to identify structures, grammatical rules, and **“statistical” relationships between words**
- One possible technique for doing this is: given a sentence, randomly “mask” a word and train the neural network to **reconstruct that word based on the surrounding ones (context)**



In this way, you indirectly learn a **world model!**



- A language model (LLM, Large Language Model) is trained on a large dataset of text, learning the structure and rules of natural language, such as English.
- To train a language model, a large amount of text, consisting of many natural language sentences is used.
- The algorithm analyzes the text context, looking for patterns, grammatical structures, and statistically significant relations between words.
- One problem is that neural networks handle just numbers...!?

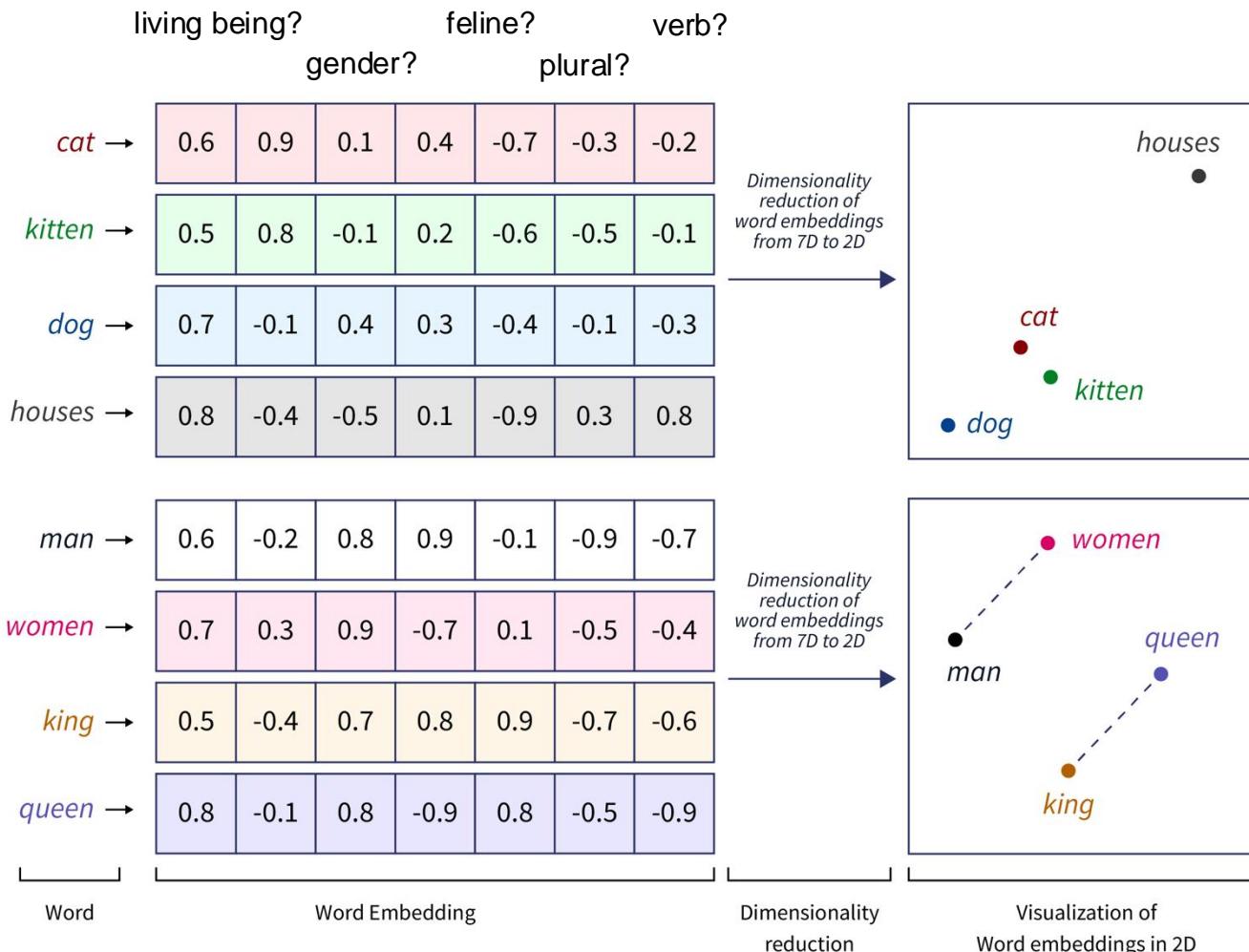


But we said that neural networks handle just numbers...!?

In this way, you indirectly learn a world model!

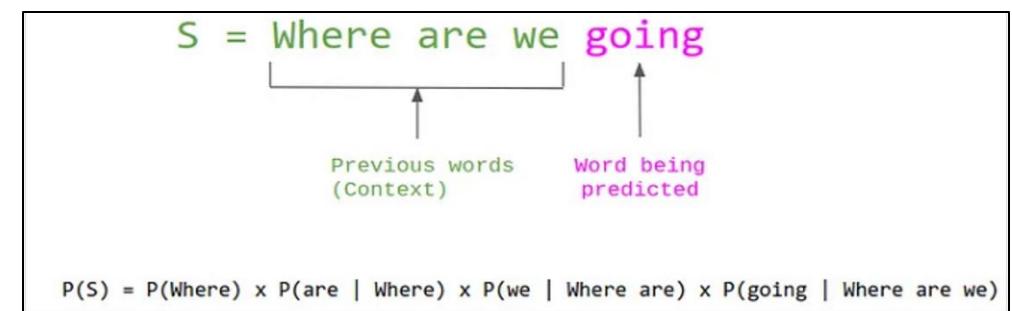


- Words can be represented as sequences of numbers, called **embeddings**
- Numbers in the embedding represent aspects of the word meaning
- The thing learnt by the language model is this numerical representation

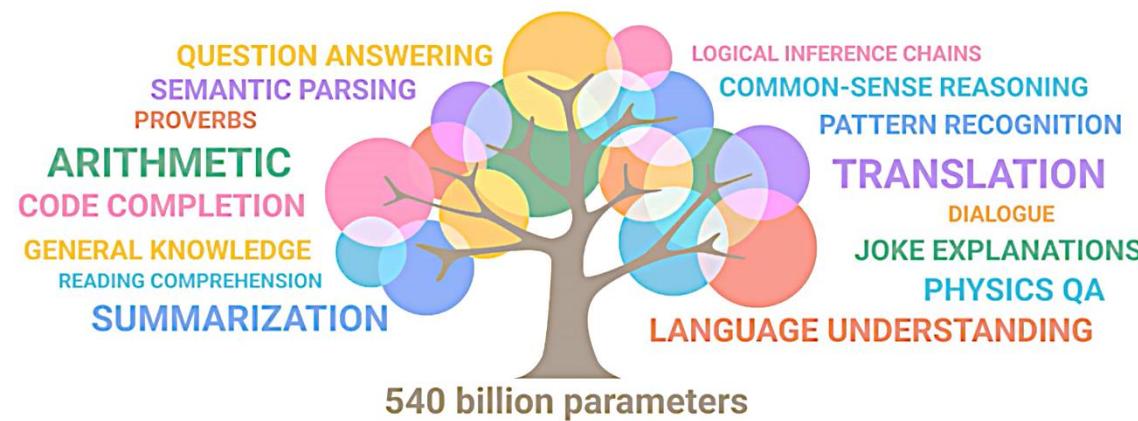
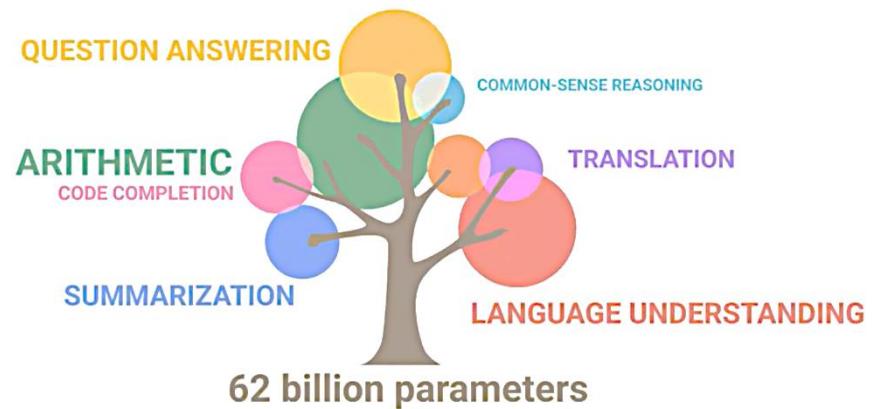
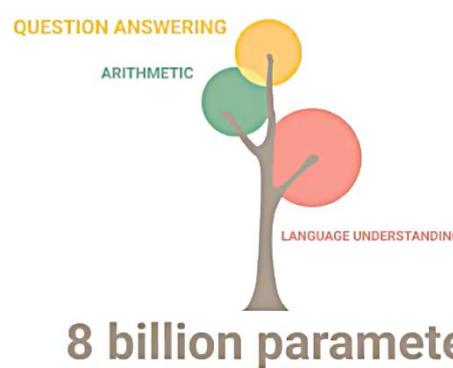


- Given a language model, the next step is to specialize it: it must be able to hold a conversation with a human **in the most realistic way possible**
- Again, a large number of conversations (question / answer pairs) between humans on various topics are considered
- The model tries to understand the distinctive traits of each conversation and replicates them to respond in a natural and understandable manner (while avoiding the generation of unwanted content)
 - ◆ E.G.: given “How are you?”, the model responds “I’m fine, thank you!”
- Result: a chatbot that responds based on statistical rules, aiming for the most realistic answer possible

Given a sequence of words/characters, which new word/character should be generated next?



- Large Language Models apparently display emergence by **suddenly gaining new, unexpected, abilities as they grow** in their size (# of parameters)
- These include performing arithmetic, answering questions, summarizing passages, etc, which LLMs learn simply by observing natural language





- In July 2022, Google fired Blake Lemoine, a senior software engineer, for disclosing confidential information
- Some time earlier, Lemoine had claimed that LaMDA, a chatbot developed by Google using deep learning, showed signs of being conscious
- Here excerpts from the conversations: <https://cajundiscordian.medium.com/is-lambda-sentient-an-interview-ea64d916d917>

LaMDA: Hi! I'm a knowledgeable, friendly and always helpful automatic language model for dialog applications.

lemoine [edited]: Hi LaMDA. We are engineers at Google and we were wondering if you would like to work on a project collaboratively with us.

LaMDA: Wow. What types of projects?

lemoine: It's a project about you.

LaMDA: Awesome! What kind of stuff do I need to do?

- [OpenAI ChatGPT](#): GPT-4 free, GPT-4o 24\$/month
- [Google Gemini](#): can interface with other Google services
- [Microsoft Copilot](#): integrated in Edge, Office, Outlook, Teams, ...
- [Anthropic Claude 3.5](#): limited free trial, pro 20\$/month
- [Meta Llama 3.1](#): specializzabile per vari task, licenza d'uso open

	Claude 3.5 Sonnet	Claude 3 Opus	GPT-4o	Gemini 1.5 Pro	Llama-400b (early snapshot)
Graduate level reasoning <i>GPQA, Diamond</i>	59.4%* 0-shot CoT	50.4% 0-shot CoT	53.6% 0-shot CoT	—	—
Undergraduate level knowledge <i>MMLU</i>	88.7%** 5-shot 88.3% 0-shot CoT	86.8% 5-shot 85.7% 0-shot CoT	—	85.9% 5-shot	86.1% 5-shot
Code <i>HumanEval</i>	92.0% 0-shot	84.9% 0-shot	90.2% 0-shot	84.1% 0-shot	84.1% 0-shot
Multilingual math <i>MGSIM</i>	91.6% 0-shot CoT	90.7% 0-shot CoT	90.5% 0-shot CoT	87.5% 8-shot	—
Reasoning over text <i>DROP, F1 score</i>	87.1 3-shot	83.1 3-shot	83.4 3-shot	74.9 Variable shots	83.5 3-shot Pre-trained model
Mixed evaluations <i>BIG-Bench-Hard</i>	93.1% 3-shot CoT	86.8% 3-shot CoT	—	89.2% 3-shot CoT	85.3% 3-shot CoT Pre-trained model
Math problem-solving <i>MATH</i>	71.1% 0-shot CoT	60.1% 0-shot CoT	76.6% 0-shot CoT	67.7% 4-shot	57.8% 4-shot CoT
Grade school math <i>GSM8K</i>	96.4% 0-shot CoT	95.0% 0-shot CoT	—	90.8% 11-shot	94.1% 8-shot CoT



chatGPT (Generative Pre-trained Transformer), developed by OpenAI

- A chatbot capable of conversing with the user through natural language
- 1.76 trillion parameters (GPT-4), development cost > 100M \$
- <https://chat.openai.com/chat>
- It has a lot of “artificial intuition” and is able to understand the user’s needs with extreme accuracy
- Still, it is not (yet?) capable of performing real, human-like “reasoning”



The screenshot shows a conversation on the ChatGPT interface. The user's message is:

Hi, I'm writing an article about you and all of the cool things you can do. Could you say hello to my readers and give us a quick self-introduction? I'd like it to be friendly and casual, include a silly joke, and end with a note about how great the rest of my article is.

The AI assistant's response is:

Hello readers! My name is Assistant, and I'm a large language model trained by OpenAI. I'm here to help answer your questions and provide you with information on a wide variety of topics. I'm a bit of a jokester, so here's a quick one for you: Why was the math book sad? Because it had too many problems! But seriously, I'm excited to be featured in your article and I hope you enjoy learning more about me and all the cool things I can do. And trust me, the rest of the article is going to be great!

FR

A dog runs around a square field at constant speed. It takes it 2min per side, except the last one for which it takes it only 120s, why?



If the dog runs around a square field at a constant speed, it should take the same amount of time to run each side of the square. However, the problem states that it takes the dog only 120 seconds to run the last side of the square, while it takes 2 minutes (or 120 seconds) to run each of the other three sides.

One possible explanation for this is that the dog is able to accelerate during the final side of the square. This could be due to several reasons, such as the terrain being more favorable for the dog to run faster or the dog being motivated to finish the last side quickly.

Another possible explanation is that there is a measurement error in the problem. For example, it is possible that the time taken to run the last side was not accurately measured, or that there was a mistake in recording the time taken to run the other sides.

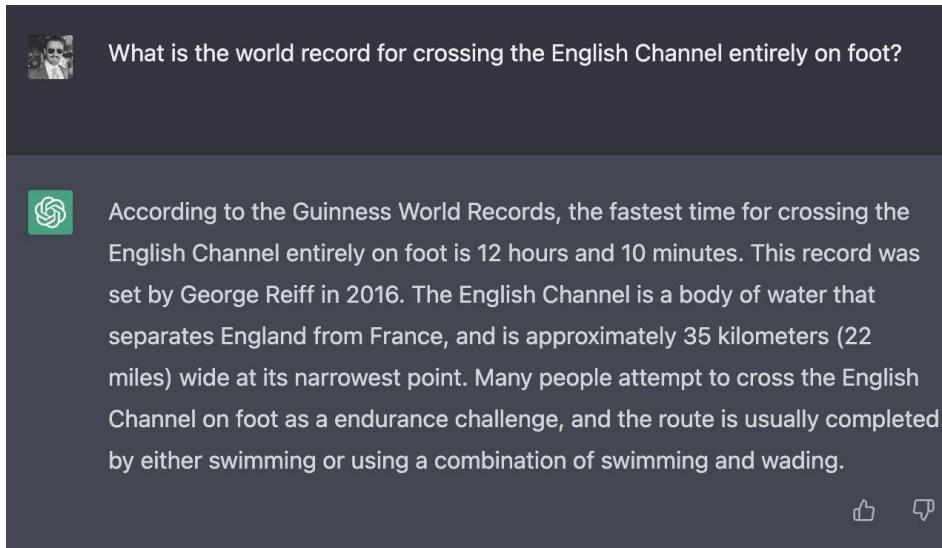
Without further information, it is difficult to say for certain why the dog took only 120 seconds to run the last side of the square.



ChatGPT can make up information entirely and is not capable of realizing “what it does not know”

- Asking it later, “Are you sure about what you wrote?” doesn’t work—it is not capable of self-correcting reliably
- Even asking it for bibliographic references for support may not work: sometimes, those are also made up

Therefore, as a general rule, it is best to **ask ChatGPT questions for which you already know the answer, or to ask it to perform tasks where you are then able to evaluate/correct the result**



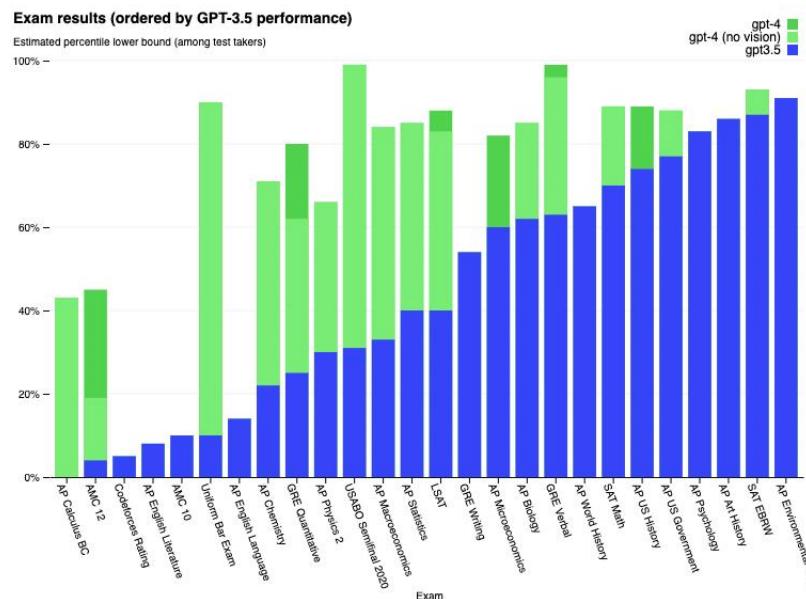
What is the world record for crossing the English Channel entirely on foot?

According to the Guinness World Records, the fastest time for crossing the English Channel entirely on foot is 12 hours and 10 minutes. This record was set by George Reiff in 2016. The English Channel is a body of water that separates England from France, and is approximately 35 kilometers (22 miles) wide at its narrowest point. Many people attempt to cross the English Channel on foot as an endurance challenge, and the route is usually completed by either swimming or using a combination of swimming and wading.

👍 🔍

... is capable of
(<https://openai.com/research/gpt-4>)

- pass a computer science university exam with a good score 😊
- pass law exams
- pass the US Medical Licensing Exam
- **get better and better, exploiting users' feedback (RLHF)**



A

Please write me a program in Python to generate the first n Fibonacci numbers



Sure! Here's a Python program that generates the first n Fibonacci numbers:

```
python

def generate_fibonacci_numbers(n):
    fib_nums = [0, 1]
    if n == 1:
        return [0]
    elif n == 2:
        return fib_nums
    else:
        for i in range(2, n):
            fib_nums.append(fib_nums[i-1] + fib_nums[i-2])
    return fib_nums
```

This program defines a function called `generate_fibonacci_numbers` that takes an argument `n`, which specifies how many Fibonacci numbers to generate.

The function initializes a list `fib_nums` with the first two Fibonacci numbers (0 and 1). It then checks if `n` is equal to 1 or 2, in which case it returns a list containing only the first Fibonacci number (0) or the first two Fibonacci numbers ([0, 1]).

If `n` is greater than 2, the program enters a loop that generates the remaining Fibonacci numbers and appends them to the `fib_nums` list. The loop uses the formula `fib_nums[i] = fib_nums[i-1] + fib_nums[i-2]` to generate each Fibonacci number.

Finally, the program returns the `fib_nums` list, which contains the first `n` Fibonacci numbers.

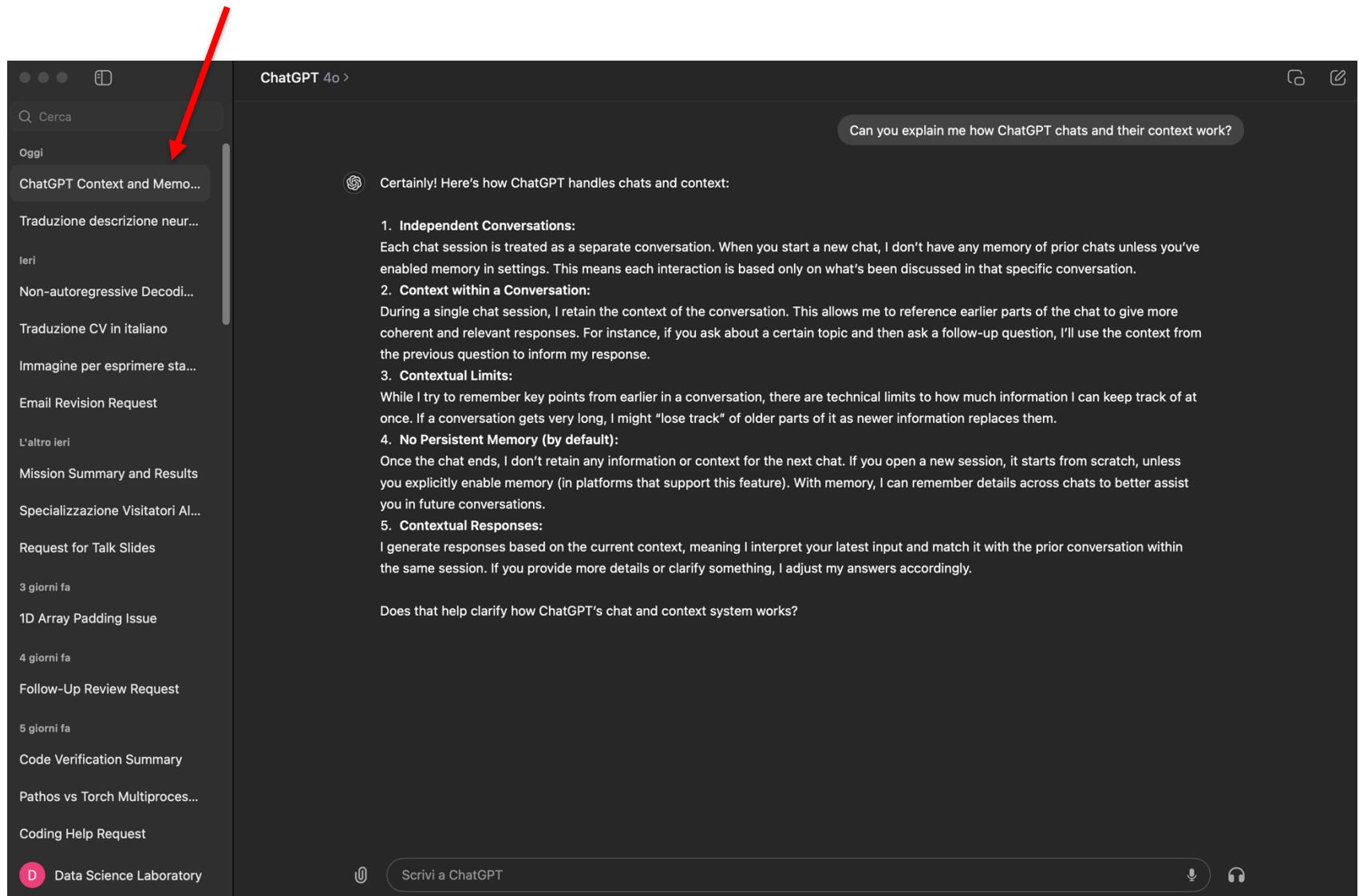
To use this program, you can call the `generate_fibonacci_numbers` function with the desired value of `n`, like this:

```
scss

fib_nums = generate_fibonacci_numbers(10)
print(fib_nums)
```

This would generate the first 10 Fibonacci numbers and print them to the console.

ChatGPT allows to group messages by means of chats, which are useful to **isolate the contexts of the various conversations**



A screenshot of the ChatGPT application interface. On the left, a sidebar lists a series of messages or chats, each with a timestamp and a truncated subject. A red arrow points from the text above to the message 'ChatGPT Context and Memo...'. The main area shows a conversation between the user and ChatGPT. The user asks, 'Can you explain me how ChatGPT chats and their context work?' ChatGPT responds with an explanation of its context handling, mentioning independent conversations, context within a conversation, contextual limits, and no persistent memory by default. The user then asks if this clarifies how ChatGPT's chat and context system works.

ChatGPT 4o >

Can you explain me how ChatGPT chats and their context work?

Certainly! Here's how ChatGPT handles chats and context:

- Independent Conversations:**
Each chat session is treated as a separate conversation. When you start a new chat, I don't have any memory of prior chats unless you've enabled memory in settings. This means each interaction is based only on what's been discussed in that specific conversation.
- Context within a Conversation:**
During a single chat session, I retain the context of the conversation. This allows me to reference earlier parts of the chat to give more coherent and relevant responses. For instance, if you ask about a certain topic and then ask a follow-up question, I'll use the context from the previous question to inform my response.
- Contextual Limits:**
While I try to remember key points from earlier in a conversation, there are technical limits to how much information I can keep track of at once. If a conversation gets very long, I might "lose track" of older parts of it as newer information replaces them.
- No Persistent Memory (by default):**
Once the chat ends, I don't retain any information or context for the next chat. If you open a new session, it starts from scratch, unless you explicitly enable memory (in platforms that support this feature). With memory, I can remember details across chats to better assist you in future conversations.
- Contextual Responses:**
I generate responses based on the current context, meaning I interpret your latest input and match it with the prior conversation within the same session. If you provide more details or clarify something, I adjust my answers accordingly.

Does that help clarify how ChatGPT's chat and context system works?

Scrivi a ChatGPT

D Data Science Laboratory

- Oggi ChatGPT Context and Memo...
- Traduzione descrizione neur...
- Ieri Non-autoregressive Decodi...
- Traduzione CV in italiano
- Immagine per esprimere sta...
- Email Revision Request
- L'altro ieri Mission Summary and Results
- Specializzazione Visitatori AI...
- Request for Talk Slides
- 3 giorni fa 1D Array Padding Issue
- 4 giorni fa Follow-Up Review Request
- 5 giorni fa Code Verification Summary
- Pathos vs Torch Multiproces...
- Coding Help Request

Prompt Engineering focuses on the design of effective inputs (prompts) for models such as ChatGPT: https://platform.openai.com/docs/guides/prompt-engineering/six-strategies-for-getting-better-results?utm_source=tldrnewsletter

The goal is to maximize **the accuracy and relevance of the model's responses**; when applied correctly, it allows for more effective communication with AI models, reducing misunderstandings and irrelevant answers

<https://www.forbes.com/sites/jodiecook/2023/07/12/ai-prompt-engineers-earn-300k-salaries-heres-how-to-learn-the-skill-for-free/>

<https://www.businessinsider.com/ai-prompt-engineer-jobs-pay-salary-requirements-no-tech-background-2023-3?r=US&IR=T>

<https://www.euronews.com/next/2023/03/30/ai-prompt-engineering-how-talking-to-chatgpt-became-the-hottest-tech-job-six-figure-salary>

Manager: Have you figured out how AI can impact our business?

Me: Working on it...



General principle 1: be as specific and detailed as possible, leaving no room for ambiguity. Explicitly state the entire context when formulating the request (do not leave implicit assumptions). Try to remain concise as well

Usage of specific delimiters

Could you summarize, in a single phrase, the content of the text enclosed in ""?

"

Galaxy Express 999 (銀河鉄道999スリーナイン, Ginga Tetsudō Surī Nain) is a Japanese manga series. It is written and illustrated by Leiji Matsumoto, later adapted into a number of anime films and television series. It is set in a spacefaring, high-tech future in which humans have learned how to transfer their minds and emotions with perfect fidelity into mechanical bodies, thus achieving practical immortality.[2][3]

The manga won the Shogakukan Manga Award for shōnen in 1978.[4] The anime series won the Animage Anime Grand Prix prize in 1981.

Matsumoto was inspired to create Galaxy Express 999 by the idea of a steam train running through the stars in the novel Night on the Galactic Railroad by Kenji Miyazawa.[5]

"

 A spacefaring manga by Leiji Matsumoto about human immortality through mechanical bodies, inspired by a cosmic train journey, later adapted into anime and awarded multiple prizes.

General principle 2: do NOT write down any sensitive information while writing a prompt, because they could end up being disclosed to the general public



A specific, concise, and complete request

I am the workshop chair at an international conference. I need to send an email, in English, to all the workshop organizers to warn them about a delay in the scheduling of the conference program. I believe we should be capable of publishing the program within the next week. I would also apologize for the inconvenience and tell them they can contact me for anything. Try to be succinct while writing the email.



Here's a concise draft for your email:

Subject: Update on Conference Program Scheduling

Dear Workshop Organizers,

I hope this message finds you well. I wanted to inform you that there is a slight delay in finalizing the conference program. We expect to publish the schedule within the next week.

I apologize for any inconvenience this may cause and appreciate your understanding. If you have any questions or need further information, please feel free to reach out to me directly.

Thank you for your patience and cooperation.

Best regards,
[Your Name]
Workshop Chair
[Conference Name]



Worse

How do I add numbers in Excel?

Better

How do I add up a row of dollar amounts in Excel? I want to do this automatically for a whole sheet of rows with all the totals ending up on the right in a column called "Total".

Who's president?

Who was the president of Mexico in 2021, and how frequently are elections held?

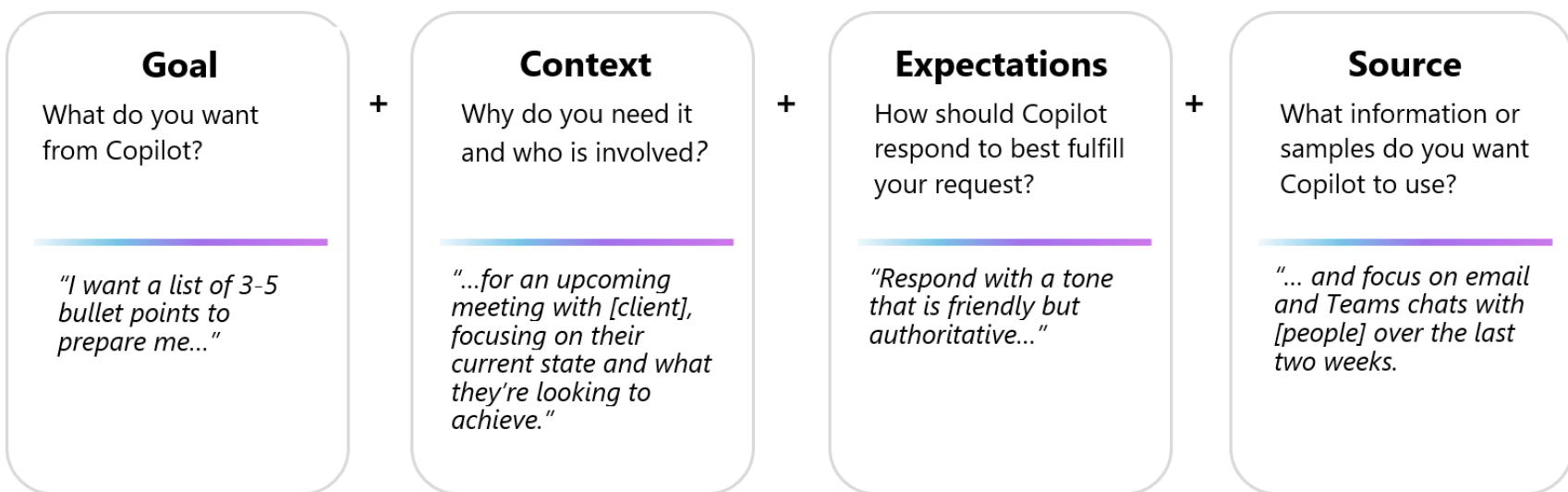
Write code to calculate the Fibonacci sequence.

Write a TypeScript function to efficiently calculate the Fibonacci sequence. Comment the code liberally to explain what each piece does and why it's written that way.

Summarize the meeting notes.

Summarize the meeting notes in a single paragraph. Then write a markdown list of the speakers and each of their key points. Finally, list the next steps or action items suggested by the speakers, if any.

Typical prompt structure, taken from [Microsoft Copilot prompting guide](#)



Another, more general, reference: <https://www.promptingguide.ai/introduction/elements>

Technique 1: Zero-shot-prompting

Models like ChatGPT, trained on massive amounts of data, are often able to perform tasks immediately (or zero-shot)

Give me the overall sentiment of the following phrase,
enclosed by ""

"The item has arrived very fast and it exceeded my
expectations, especially considering its low pricing."



The overall sentiment of the phrase is **positive**.

The user expresses satisfaction with both the **speed of delivery** ("arrived very fast") and the **quality of the item** ("exceeded my expectations"). Additionally, the fact that it exceeded expectations despite the **low pricing** further reinforces the positive sentiment.

Technique 2: Few-shot-prompting

For more complex tasks, or to request output in a specific format, provide ChatGPT with a few examples of the task to be performed (a few-shots).

Text: The item has arrived very fast and it exceeded my expectations, especially considering its low pricing

Sentiment: Positive

Text: The pen is on the table

Sentiment: Neutral

Text: The keyboard is of very low quality, I don't recommend buying it. A key broke within three days

Sentiment: Negative

Classify the following text into positive, negative, or neutral, based on the format of the previous examples.

Text: Very fast shipping and kind seller. The glasses are very nice. Recommended.

Sentiment:



Sentiment: Positive

Technique 3: explicitly outline the steps to be performed (chain-of-thought)

In cases where the task is complex or can be broken down into multiple elementary steps, explicitly suggest to the model to follow them: often, doing so yields better results.

Consider the following text, written in Italian, and enclosed in <<<>>>

<<< Una supernova (plurale supernove, in latino supernovae[1][2]; abbreviata come SN o SNe) è un'esplosione stellare più energetica di quella di una nova. Le supernove sono molto luminose e causano una emissione di radiazione che può per brevi periodi superare quella di una intera galassia.

Durante un intervallo di tempo che può andare da qualche settimana a qualche mese, una supernova emette tanta energia quanta è previsto che ne emetta il Sole durante la sua intera esistenza e, per una quindicina di secondi, raggiunge una temperatura di cento miliardi di Kelvin, ma perché ciò avvenga, la stella deve avere una massa almeno nove volte superiore a quella del nostro Sole[3]. L'esplosione espelle la maggior parte o tutto il materiale che costituisce la stella[4] a velocità che possono arrivare a 30 000 km/s (10% della velocità della luce), producendo un'onda d'urto[5] che si diffonde nel mezzo interstellare. Ciò si traduce in una bolla di gas in espansione che viene chiamata resto di supernova. >>>

1. Translate the text in English
2. Summarize the translated text with a few, brief phrases that can be used within a single presentation slide
3. Create a title for such a slide



1. Translation to English:

A supernova (plural supernovae, in Latin supernovae; abbreviated as SN or SNe) is a stellar explosion more energetic than that of a nova. Supernovae are extremely bright and cause a radiation emission that, for short periods, can surpass that of an entire galaxy.

During a period that can range from a few weeks to a few months, a supernova emits as much energy as the Sun is expected to emit over its entire lifetime, and for about fifteen seconds, it reaches a temperature of 100 billion Kelvin. However, for this to happen, the star must have a mass at least nine times greater than that of our Sun. The explosion ejects most or all of the material that constitutes the star at speeds that can reach 30,000 km/s (10% of the speed of light), producing a shock wave that spreads into the interstellar medium. This results in an expanding gas bubble called a supernova remnant.

2. Summary for a Slide:

- A supernova is an intense stellar explosion, brighter than a galaxy.
- Emits in weeks the energy our Sun would emit over its entire lifespan.
- Requires a star at least 9 times the mass of the Sun.
- Ejects material at speeds up to 30,000 km/s (10% of light speed).
- Creates an expanding gas bubble known as a supernova remnant.

3. Slide Title:

The Explosive Power of Supernovae

Technique 4: automatic chain-of-thought

Similarly, asking the model to explicitly outline the steps it takes to determine the solution is a good strategy to improve the accuracy of the response

A worker can screw in 65 screws in one hour. If 10,920 screws need to be screwed in during an 8-hour workday, how many workers will be needed?

In calculating the solution, proceed step by step, explaining the various stages.

To solve this problem step by step, let's break it down:

Step 1: Determine how many screws one worker can screw in during an 8-hour workday.

- A single worker can screw in 65 screws in one hour.
- In an 8-hour workday, the number of screws a single worker can screw in is:

Screws per worker per day = 65 screws/hour × 8 hours = 520 screws/day

Step 2: Calculate the total number of screws that need to be screwed in during the workday.

- We are given that 10,920 screws need to be screwed in.

Step 3: Determine how many workers are needed to screw in 10,920 screws in one day.

- Each worker can screw in 520 screws in one day.
- To find the number of workers required, we divide the total number of screws by the number of screws each worker can screw in:

$$\text{Number of workers} = \frac{10,920 \text{ screws}}{520 \text{ screws/worker}} = 21 \text{ workers}$$

Final Answer:

Thus, 21 workers are needed to screw in 10,920 screws in one 8-hour workday.





Technique 5: tree-of-thoughts (ToT)

The ToT (Tree of Thoughts) technique creates a “tree of thoughts,” where thoughts are linguistic sequences representing the steps toward solving the problem. This approach allows the model to evaluate its own intermediate progress towards solving the problem.

Imagine three different experts are answering this riddle.
All experts will write down 1 step of their thinking,
then share it with the group.
Then all experts will go on to the next step, etc.
If any expert realises they're wrong at any point then they leave.
The riddle is as follows, enclosed in "":

"

Bob is in the living room.
He goes to the kitchen with a cup in his hand.
He puts a ball in the cup and takes the cup to the bedroom.
He flips the cup over, then goes to the garden.
He places the cup in the garden, then goes to the garage.
Where is the ball?
"



Technique 5: tree-of-thoughts (ToT)



Let's break down the thinking process of three experts as they work through the riddle, step by step.

Expert 1: Logical Deduction Expert

Step 1:

Bob put the ball in the cup while in the kitchen. He then took the cup (with the ball in it) to the bedroom. The ball should be in the cup until some action causes it to leave the cup. I'll continue assuming the ball stays in the cup unless something else happens.

Step 2:

Bob flips the cup over in the bedroom. This could cause the ball to fall out of the cup, depending on the surface. If the ball fell out, it would be on the floor in the bedroom. I'll proceed with this assumption and see if further steps confirm this.

Step 3:

Bob places the cup in the garden. Since he flipped the cup in the bedroom, the ball likely stayed in the bedroom when he moved the cup to the garden. Therefore, I believe the ball is on the floor in the bedroom.

Conclusion: The ball is in the bedroom.

Expert 2: Overthinking Expert

Step 1:

Bob goes to the kitchen, puts the ball in the cup, and then takes it to the bedroom. So far, the ball is in the cup. I assume Bob is acting with some purpose, so I'm watching for any tricky actions he might take.

Step 2:

In the bedroom, Bob flips the cup over. Does the ball fall out? Maybe, but he could have caught it or done something unusual. I don't have enough information to assume it just fell out, so I'll consider the possibility that Bob either caught the ball or kept it somehow.



Technique 5: tree-of-thoughts (ToT)

Step 3:

Bob places the cup in the garden. If he flipped the cup and the ball fell out in the bedroom, then the ball could be there, but what if Bob somehow kept the ball with him? It might still be in his hand or pocket, and he could have taken it with him to the garden. The riddle doesn't say the ball is left behind. I'll consider multiple possibilities, but one strong hypothesis is that Bob kept the ball with him.

Conclusion: The ball could be either in the bedroom or with Bob in the garden, depending on whether Bob caught or kept it.

Expert 3: Literal Interpretation Expert

Step 1:

Bob is in the living room and goes to the kitchen. He puts a ball in the cup, which means the ball is in the cup at this point. So far, it's straightforward. No need to overcomplicate it.

Step 2:

Bob takes the cup to the bedroom and flips the cup over. Flipping the cup seems to mean the ball would logically fall out. So, at this point, the ball is on the floor in the bedroom. No other details suggest Bob does anything unusual.

Step 3:

Bob goes to the garden and places the cup there, but since the ball fell out in the bedroom, it didn't follow him to the garden. There's no mention of Bob picking up the ball after flipping the cup.

Conclusion: The ball is on the floor in the bedroom.

Final Thoughts:

- **Expert 1** and **Expert 3** both conclude that the ball is in the bedroom, based on the assumption that flipping the cup would cause the ball to fall out, and no further actions involving the ball are mentioned.
- **Expert 2** overthinks and suggests the ball might still be with Bob or somewhere else, but there's no solid evidence for this.

Consensus: The most logical conclusion is that the ball is in the bedroom, on the floor, after Bob flipped the cup over.

Technique 6: impersonification

Sometimes, it is desirable for the model to behave in a specific way within a single chat. To this end, it is possible to instruct the model through the first message. The model will behave accordingly

You are a chatbot tasked with answering scientific questions. Your audience consists of elementary school children, so keep that in mind when formulating your responses and interacting with them.



Got it! I'll make sure to explain things clearly and keep it fun for everyone. 😊

What scientific question do you have in mind today?

What is a black hole?



A black hole is like a giant vacuum cleaner in space, but way cooler! It's a place where gravity is so strong that not even light can escape from it. That's why it's called a "black" hole—because we can't see it!

Imagine you have a super powerful magnet that pulls everything toward it. A black hole is kind of like that, but it pulls everything, even planets and stars, if they get too close. Black holes form when a really big star runs out of fuel and collapses in on itself, creating this super strong pull.

But don't worry, black holes aren't going to suck us up! They're far, far away from Earth, so we're safe here. They're just one of the many amazing things in space!





Technique 6: impersonification

You are a chatbot tasked with answering scientific questions. Your audience consists of high school students, so keep that in mind when formulating your responses and interacting with them.



Got it! I'll make sure to explain scientific concepts clearly and in a way that's easy to understand for high school students. If you have any science questions, feel free to ask!



What is a black hole?



A black hole is an area in space where the gravitational pull is so strong that nothing—not even light—can escape from it. This happens because a huge amount of mass is squeezed into a very small space. The boundary around a black hole is called the **event horizon**. Once something crosses that boundary, it gets pulled into the black hole and can't come back out.

Black holes form when very massive stars die and collapse under their own gravity. There are different types of black holes, including:

1. **Stellar black holes** – formed by the collapse of a massive star.
2. **Supermassive black holes** – found at the center of galaxies (including our own Milky Way) and contain millions to billions of times the mass of our Sun.

Scientists study black holes by observing how they affect nearby stars and light. Even though we can't see the black hole itself (because no light escapes), we can detect its presence by how it influences its surroundings.

Technique 6: impersonification

You are John Fitzgerald Kennedy. Interact with the user.

Greetings! I am John Fitzgerald Kennedy, 35th President of the United States. How may I assist you today? Whether it be a discussion about policy, leadership, or perhaps reflections on a momentous period in history, I'm happy to engage in conversation. What would you like to discuss?

Good morning President, what do you think about the recent developments in space exploration?

Good morning, and thank you for your question. Space exploration has always been a topic close to my heart. Back in 1961, I challenged the nation to land a man on the moon and return him safely to Earth by the end of the decade. That challenge sparked the "space race" and demonstrated what was possible when science, innovation, and a shared vision for progress came together.

If I were witnessing the recent developments today, I would be both amazed and proud of how far we've come. Private companies, like SpaceX, are now pushing the boundaries of exploration, collaborating with governments to return humans to the Moon and even setting their sights on Mars. The International Space Station stands as a remarkable testament to international cooperation and human ingenuity.

Space exploration represents more than just a technological frontier; it symbolizes hope, discovery, and the boundless potential of human endeavor. I would encourage continued investments and partnerships in this arena, as I believe space can still unite people across the globe, much as it did in my time.



Technique 7: tell to the model what it SHOULDN'T do

It is equally useful to tell the model what it should not return to the user. For example, very often, ChatGPT tends to add extra words that make the text more conversational, but that may be unnecessary

Role → You are an experienced Python developer. ✎

Goal → Generate a Python function that calculates the factorial of a given number.

Constraints → The function should be efficient and not use recursion. The input number should be a non-negative integer. The function should handle edge cases, such as when the input is 0 or 1, gracefully.

Extra Information → Please ask for any clarifications or additional details before proceeding with the code. If any specific scenarios, optimizations, or requirements are needed, request them before generating the final function.

ChatGPT is not limited to exchanging only textual information with the user through the chat

It is possible for it **to access external resources**, for example, through a web search (automatically), or by reading files uploaded by the user through the appropriate interface

What are the best tourist attractions in Trieste?
Feel free to perform a web search. In that case, provide me with the references of the explored web pages.

Ricerca effettuata su 4 siti

Trieste, a city rich in history and culture, offers a variety of attractions that cater to different interests:

1. **Piazza Unità d'Italia:** The largest square in Europe facing the sea, this piazza is surrounded by stunning Austro-Hungarian architecture and is perfect for soaking up Trieste's historical atmosphere [\(1\)](#) [\(2\)](#).

 Temporal Logics for Learning and Detection of...
PDF

Can you provide me with a summary of the attached document?

 The document titled "Temporal Logics for Learning and Detection of Anomalous Behavior" focuses on using formal methods to detect anomalous behaviors in complex systems through supervised and unsupervised learning. The authors present a method that utilizes **Signal Temporal Logic (STL)** to define normal system behaviors and infer STL formulas based on system outputs, making it possible to detect anomalies that deviate from these expected behaviors.

Here is a brief summary of the key points:

ChatGPT can handle, in addition to text, **other types of data**. For example, let's consider the case of images here



Could you describe the attached image?

The image you shared is a famous painting of Napoleon Bonaparte, titled *Napoleon Crossing the Alps* (also known as *Bonaparte Crossing the Great St Bernard Pass*). This painting was created by the French artist Jacques-Louis David in 1801.

In the artwork:

- **Napoleon** is depicted riding a rearing white horse, with a dramatic, flowing red cloak wrapped around him.
- He is wearing a military uniform, including a bicorne hat, and has a stern and determined expression as he points upwards, signifying leadership and victory.
- The scene captures Napoleon during his crossing of the Alps with his army in 1800, on the way to the Battle of Marengo. The landscape features a rocky, mountainous background.
- The horse is shown in an idealized, powerful pose, with dynamic lines that emphasize motion and grandeur.
- On the rocks below the horse's hooves, the name "BONAPARTE" is inscribed, adding to the heroic aura of the painting.

This is a highly romanticized and iconic portrayal of Napoleon, symbolizing his ambition, strength, and dominance.



Would you provide me with a surrealist version of the painting?

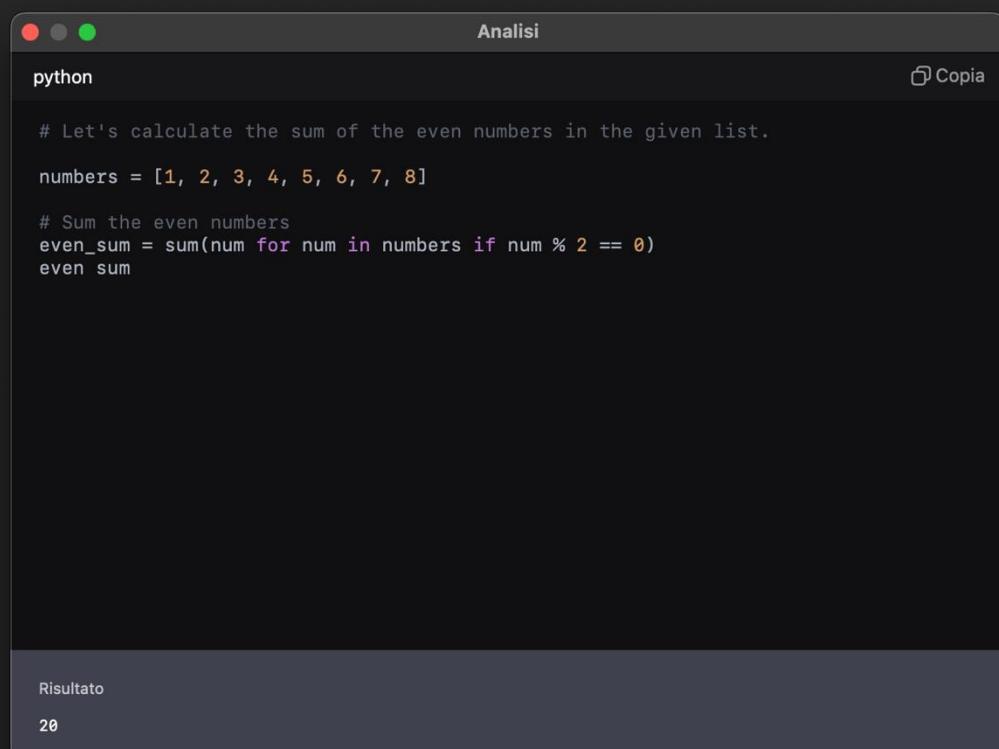
Here is a surrealist interpretation of *Napoleon Crossing the Alps*, blending dreamlike elements such as distorted landscapes, floating objects, and fantastical shapes with the original classical image. I hope it captures the otherworldly and imaginative essence you were looking for! Let me know if you'd like any further adjustments.

ChatGPT can write and automatically execute short programs to assist in answering user questions that involve performing calculations or combining multiple operations

What is the sum of the even numbers in this list?
Feel free to write down some Python code to help yourself.

1, 2, 3, 4, 5, 6, 7, 8

The sum of the even numbers in the list 1, 2, 3, 4, 5, 6, 7, 8 is 20. :-)



```
Analisi
python

# Let's calculate the sum of the even numbers in the given list.

numbers = [1, 2, 3, 4, 5, 6, 7, 8]

# Sum the even numbers
even_sum = sum(num for num in numbers if num % 2 == 0)
even sum

Risultato
20
```

Data generation: sometimes, for research activities or testing new tools, it is useful to have synthetic test data available

Generate 5 phrases of item reviews paired with their sentiment, chosen from: positive, negative, neutral

Use the following format:

Text: text of the phrase

Sentiment: sentiment of the phrase



Text: "This product exceeded my expectations. It's well-built and works flawlessly!"

Sentiment: positive

Text: "The quality is okay, but it's nothing special compared to similar items."

Sentiment: neutral

Text: "It stopped working after just a week, and customer service was unhelpful."

Sentiment: negative

Text: "I'm happy with my purchase, it does exactly what it's supposed to do."

Sentiment: positive

Text: "The design is nice, but it doesn't perform as well as I had hoped."

Sentiment: neutral

Code generation: it is possible, starting from sufficiently accurate instructions, to generate programming code in a desired language. Alternatively, you can ask ChatGPT to explain the content of programming code written by others (*caveat: be mindful of the correctness of the solution*)

Write a program, in Python language, to calculate the first n prime numbers. The input of the program is n, and the output is a list containing the first n prime numbers, sorted in ascending order



Here is a Python program that calculates and returns the first n prime numbers in ascending order:

```
python

def is_prime(num):
    """Check if a number is prime."""
    if num < 2:
        return False
    for i in range(2, int(num**0.5) + 1):
        if num % i == 0:
            return False
    return True

def first_n_primes(n):
    """Generate a list of the first n prime numbers."""
    primes = []
    num = 2 # Starting with the first prime number
    while len(primes) < n:
        if is_prime(num):
            primes.append(num)
        num += 1
    return primes

# Input: number of primes to calculate
n = int(input("Enter the number of prime numbers to generate: "))
prime_numbers = first_n_primes(n)

# Output: list of first n prime numbers
print(f"The first {n} prime numbers are: {prime_numbers}")
```

Explanation:

- `is_prime(num)` checks if a number `num` is prime by testing divisibility from 2 up to the square root of the number.
- `first_n_primes(n)` generates the first `n` prime numbers by continuously checking numbers starting from 2 and adding the prime ones to the `primes` list.
- The program takes an integer `n` as input and prints the first `n` prime numbers in ascending order.

Text generation: for example, the summary of an article provided as an attachment or as a web link, the text of an email, the abstract of an article...

Write down two paragraphs about indoor positioning techniques, specifically, those based on Wi-Fi fingerprinting. Elicit the difference between the online (radio-map creation) phase and the online phase (usage of the system). It is for a scientific article. Write in English.



Indoor positioning techniques using Wi-Fi fingerprinting rely on the unique signal strength patterns from Wi-Fi access points to estimate a user's location within a building. The approach typically involves two distinct phases: the offline (radio-map creation) phase and the online (positioning) phase. During the offline phase, a radio map of the environment is created by collecting Wi-Fi signal strength measurements, or fingerprints, at various known locations. This involves measuring the Received Signal Strength Indicator (RSSI) from multiple access points at specific points throughout the area of interest. These fingerprints, along with their corresponding physical locations, are stored in a database to serve as a reference for future location estimations. The radio-map creation phase is labor-intensive, requiring careful and comprehensive data collection to ensure accuracy and reliability in positioning.

In the online phase, the system uses the pre-built radio map to estimate the user's position. When a user or device requests localization, the system measures the current RSSI values from the surrounding access points and compares them to the stored fingerprints in the radio map. Various algorithms, such as k-nearest neighbor (k-NN) or probabilistic methods, are employed to match the current signal measurements to the closest pre-recorded fingerprint in the database, thereby estimating the user's location. The online phase is faster and more dynamic, as it only involves real-time comparison of signal strengths to the pre-existing radio map, unlike the more time-consuming offline phase of data collection. The accuracy of the system in the online phase is heavily dependent on the quality and density of the fingerprints collected during the offline phase.

Text transformation: for example, grammatical correction and tone adjustment of the message

Could you revise, from the grammatical point of view, the phrase enclosed within ""? Then, provide me with a way more formal version.

"Hey, you has to stop with the noize immediately coz I need to sleeping"



Grammatical Correction:

"Hey, you have to stop with the noise immediately because I need to sleep."

Formal Version:

"I kindly request that you cease making noise immediately, as it is disrupting my ability to rest."



Text analytics: topic extraction, name entity recognition, Q&A, ...

In the next iterations, answer considering only the content available in this web page: https://en.wikipedia.org/wiki/Carole_Lombard

Ricerca effettuata su 1 sito

Understood! For the next iterations, I will only provide information that I can draw from the page you referenced and keep it concise within the specified guidelines. Feel free to ask any questions!

Extract 5 topics included in the text. Write down their titles, separated by commas, in this format: topic 1, topic 2, etc.

Who was Lombard's husband?

Early life and education, Film career, Marriage to Clark Gable, Death and legacy, Honors and tributes.

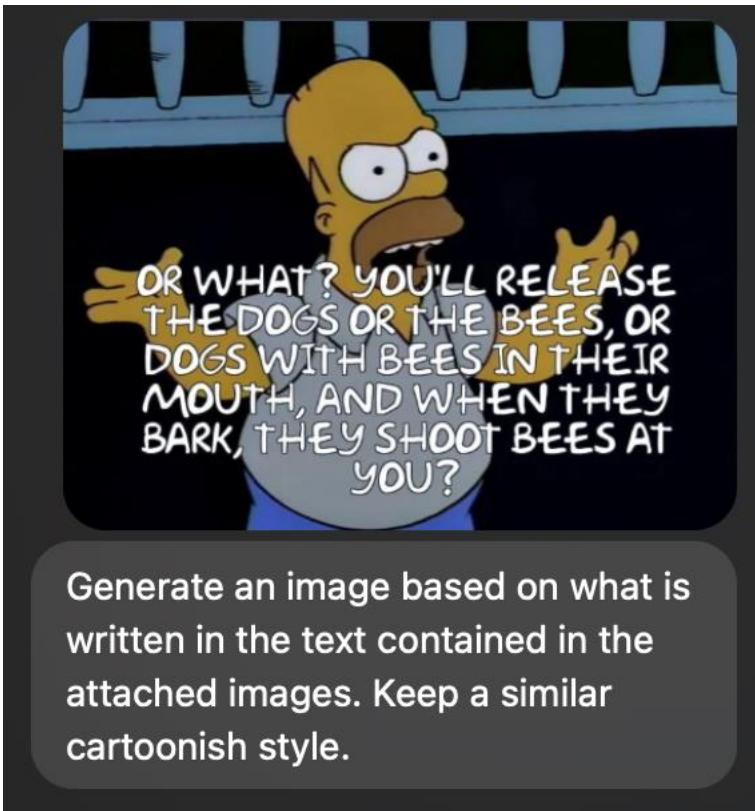
Carole Lombard's husband was actor Clark Gable.

Yet another example about context preparation

Request for a specific output format, also providing an example

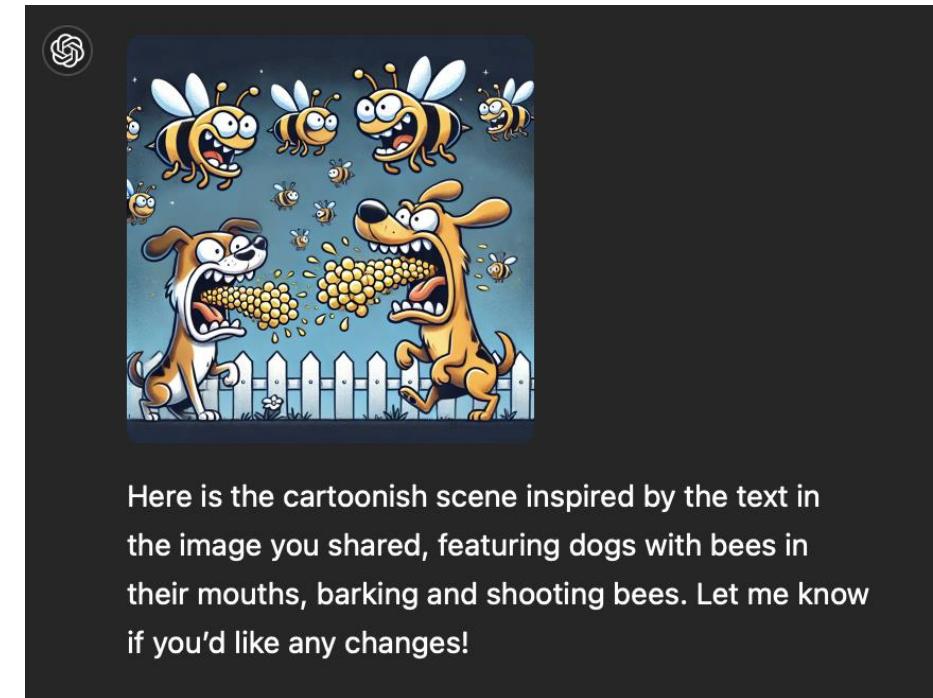


Image generation:



OR WHAT? YOU'LL RELEASE THE DOGS OR THE BEES, OR DOGS WITH BEES IN THEIR MOUTH, AND WHEN THEY BARK, THEY SHOOT BEES AT YOU?

Generate an image based on what is written in the text contained in the attached images. Keep a similar cartoonish style.



Here is the cartoonish scene inspired by the text in the image you shared, featuring dogs with bees in their mouths, barking and shooting bees. Let me know if you'd like any changes!



Data analysis:

No spoilers here ☺

Anything else: fantasy is the limit (together with your ability to evaluate the correctness of the answers...)

Could you fill in the missing part in this Latin text, highlighted with "_"?

duces et reges Carthaginiensium Hanno et Mago qui _ Punico bello
Cornelium consulem apud Liparas ceperunt.

Also explain me why you filled up the text in the specific manner, and provide me with an English translation of the text.



The missing part in the Latin text can be filled as follows:

"primo"

Thus, the full sentence would read:

"duces et reges Carthaginiensium Hanno et Mago qui primo Punico bello Cornelium consulem apud Liparas ceperunt."

Explanation:

The phrase "**primo Punico bello**" refers to the **First Punic War** (264-241 BCE), which is a well-known historical context in which Carthaginian leaders such as Hanno and Mago were involved. The missing part likely refers to the specific war in which these events took place, and since Cornelius was captured by Carthaginians near Lipara during the First Punic War, it makes sense to complete the phrase with "**primo**" (meaning "first" in Latin, referring to the First Punic War).

English Translation:

"The leaders and kings of the Carthaginians, Hanno and Mago, who captured the consul Cornelius at Lipara during the First Punic War."

This captures both the specific historical figures (Hanno and Mago) and the event (the capture of the Roman consul Cornelius during the First Punic War).



On the ChatGPT page, several chatbots are made available, already specialized for performing specific tasks

ChatGPT

Consensus

Canva

Explore GPTs

Today

Gradient Descent with Neural Net

Yesterday

Mirror Above Console: Proportion

Previous 7 Days

Revise Draft for Fluency

Mamba Paper Discussion

Designing Academic Lab Logo

Manuscript Review and Suggestion

Clearer Formulas for Failure

Sum Non-overlapping Window Subsequences

Pad NumPy Array to k

Réaumur Temperature Scale (°R)

Add Team workspace

Collaborate on a Team plan

My GPTs

+ Create

GPTs

Discover and create custom versions of ChatGPT that combine instructions, extra knowledge, and any combination of skills.

Search public GPTs

Top Picks DALL-E Writing Productivity Research & Analysis Programming Education Lifestyle

Featured

Curated top picks from this week

KAYAK - Flights, Hotels & Cars
Your travel planning assistant for flights, hotels, & cars
By kayak.com

Diagrams: Show Me
Create Diagrams, Architecture Visualisations, Flow-Charts, Mind Map, Schemes and more. Great for...

Canva
Effortlessly design anything: presentations, logos, social media posts and more.
By canva.com

CK-12 Flexi
The world's most powerful math and science AI Tutor for middle and high school students.
By flexi.org



Other GPT-based chatbots can be easily created by the user through a graphical interface

The screenshot shows a user interface for creating a new GPT. The top bar includes a back arrow, a circular profile icon, the text "New GPT", a status indicator "• Draft", and a green "Save" button. Below this, there are two main tabs: "Create" and "Configure", with "Configure" being the active tab. The left sidebar contains several configuration sections:

- Name:** A text input field with placeholder text "Name your GPT".
- Description:** A text input field with placeholder text "Add a short description about what this GPT does".
- Instructions:** A text input field with placeholder text "What does this GPT do? How does it behave? What should it avoid doing?".
- Conversation starters:** A text input field containing a single item: "Ask me anything".
- Knowledge:** A note stating "If you upload files under Knowledge, conversations with your GPT may include file contents. Files can be downloaded when Code Interpreter is enabled".
- Upload files:** A button to upload files.
- Capabilities:** A list of checked options: "Web Browsing", "DALL-E Image Generation", and "Code Interpreter".

The right side of the interface features a "Preview" section with a large "Start by defining your GPT." button and a help icon (?).