

Machine Learning Techniques for Text

Module 9: Generating Text in Chatbots

Dr. Nikos Tsourakis



Course outline



- Module 0: Python Crash Course
- Module 1: Intro to Machine Learning
- Module 2: Detecting Spam Emails
- Module 3: Classifying Topics of Newsgroup Posts
- Module 4: Extracting Sentiments from Product Reviews
- Module 5: Recommending Music Titles
- Module 6: Teaching Machines to Translate
- Module 7: Summarizing Wikipedia Articles
- Module 8: Detecting Hateful and Offensive Language
- **Module 9: Generating Text in Chatbots**
- Module 10: Clustering Speech-to-Text Transcriptions

Overview



- The most common depiction of machine intelligence in popular culture is conversational agents that can mimic human dialogs. In this chapter, we will deal with a particular type: **chatbots**
- Chatbots have received much hype in recent years; in this module, we will discuss related topics from the perspective of natural language generation
- Language modeling is an integral part of modern chatbot deployments
 - First, we will look deeper at this core component of modern natural language processing and contrast two approaches based on the transformer architecture
 - Then, we will put this knowledge into action by building a language model from scratch and evaluating its performance on a publicly available corpus
 - We will examine how to use pre-trained language models to create a chatbot
 - Transfer learning will allow us to adjust these models for any custom application
 - In this context, we will perform fine-tuning using techniques such as reinforcement learning

Module objectives



After completing this module, you should be able to:

- Understanding the different types of chatbots
- Understanding, building, and fine-tuning language models
- Applying the proper evaluation metrics for language models
- Using tools to visualize the machine learning workflow
- Implementing graphical user interfaces

Machine Learning Techniques for Text

Section 1: Creating a retrieval-based chatbot

Understanding text generation



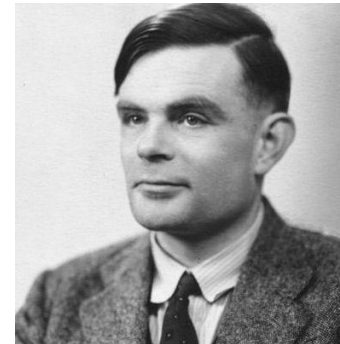
- **Natural language generation** (NLG) is an emerging research area that uses artificial intelligence to generate human language, and it's a subcategory of natural language processing
- NLG systems can be used in various contexts to describe an image with a short sentence, generate news feeds, or even spread personalized propaganda and misinformation
- We tackle the generative processes in natural language processing with a particular focus on conversational agents, also known as **chatbots**
- Chatbots find extensive usage in various tasks; for example, in large organizations, they answer customer queries; in education, they assist in teaching a new language; in research, they gather data from humans

- Broadly, chatbots can be categorized into two main types: *retrieval* and *generative-based* chatbots
- Agents in the first category are used in closed-domain scenarios and rely on a collection of predefined responses to the user input
- On the other hand, generative chatbots produce original combinations of phrases rather than selecting them from a list of options
- Finally, their input and output can utilize different communication channels, such as speech, text, or even facial expressions and gestures

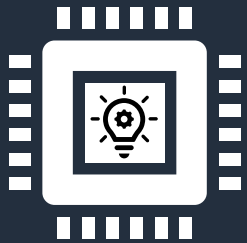
- The implementation in this section is inspired by **ELIZA**^{*}, an early chatbot, and one of the first programs capable of attempting the Turing test
- It simulated a psychotherapist in an initial psychiatric interview to demonstrate the superficiality of the communication between man and machine

^{*}<https://web.njit.edu/~ronkowit/eliza.html>

The Turing test, originally called the imitation game by Alan Turing in 1950, is a test of a machine's ability to exhibit intelligent behaviour equivalent to, or indistinguishable from, that of a human [Wikipedia]



Let's practice!



Tasks

- Chatbot



<https://colab.research.google.com/github/PacktPublishing/Machine-Learning-Techniques-for-Text/blob/main/chapter-09/chatbot.ipynb>



Machine Learning Techniques for Text

Section 2: Understanding language modeling

Language models



- **Language models** are key ingredients for creating chatbots and many natural language processing applications
- A language model expresses our confidence that a sentence is probable in the target language
- Probability in this context does not necessarily refer to whether a sentence is grammatically correct but how it resembles how people write
- Essentially, a language model learns from text resources, which can contain ungrammatical sentences, misspelled words, slang, biases, etc.



*The objective is to predict the next word,
given all previous words within some text*

Why does Predicting the Next Word Works?

The essence of decoder-only architecture is to *accurately predict the next word* for reconstructing the pre-training data. Till now, there has been no formal study that theoretically demonstrates its advantage over other architectures. An interesting explanation was from Ilya Sutskever during the interview held by Jensen Huang^a. The original transcript from the interview was copied below^b:

Say you read a detective novel. It's like complicated plot, a storyline, different characters, lots of events, mysteries like clues, it's unclear. Then, let's say that at the last page of the book, the detective has gathered all the clues, gathered all the people and saying, "okay, I'm going to reveal the identity of whoever committed the crime and that person's name is". Predict that word.
...
Now, there are many different words. But predicting those words better and better, the understanding of the text keeps on increasing. GPT-4 predicts the next word better.

a. <https://www.nvidia.com/en-us/on-demand/session/gtcspring23-S52092/>

b. <https://lifearchitected.ai/ilya/>

source: *A Survey of Large Language Models*, Zhao et al., 2023

The objective of a language model is to predict the next word, given all previous words within some text

Generative Pre-trained Transformer (GPT) family models works like a traditional language model as they outputs one token at a time

This output token is added to the sequence of inputs in the next time step

Models, where the observations from the previous time steps are used to predict the value at the current one, are called **autoregressive**

- We can use both automatic and manual approaches to assess the intermediary and the final objectives of an NLP system, known as *intrinsic* and *extrinsic* evaluation, respectively
- In language modeling, we can perform an extrinsic evaluation by letting the model generate a large number of sentences and asking human annotators to rate each one
- On the other hand, a commonly used intrinsic metric to evaluate language models is *perplexity* (PPL)
- Intuitively, it tells us how perplexed the model is after encountering a piece of text; the smaller the model's surprise, the better its performance

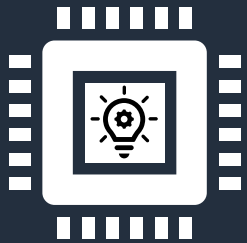
Perplexity



- Consider the following analogy: A small number of college students must be selected for the basketball team
- Height is an intrinsic property of each student
- An increased height does not guarantee higher basketball performance but correlates well
- Therefore, first, we can evaluate all students solely on their height and pick a small fraction for the actual draft on the court



Let's practice!



Tasks

- Chatbot
- Perplexity



<https://colab.research.google.com/github/PacktPublishing/Machine-Learning-Techniques-for-Text/blob/main/chapter-09/perplexity.ipynb>




Machine Learning Techniques for Text

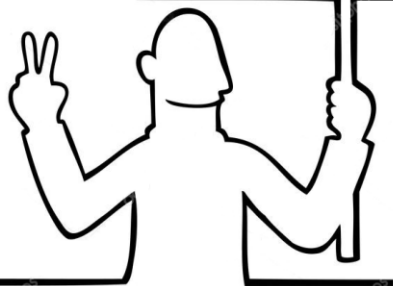
Section 3: ChatGPT

Overview of ChatGPT



- A language model designed to generate human-like responses
- Launched by  OpenAI in November 30, 2022
- Trained on a massive dataset of diverse texts, including books, articles, websites, and more, amounting to hundreds of billions of words

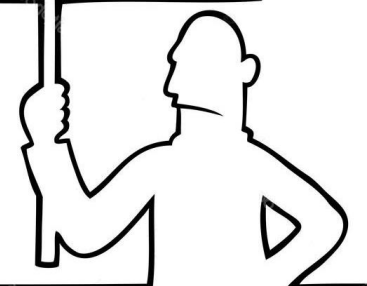
ChatGPT provides helpful insights on a wide range of topics, making it a valuable resource!



NETFLIX took 41 months to reach 1M users
FACEBOOK took 10 months to reach 1M users
INSTAGRAM took 2.5 months to reach 1M users

ChatGPT took 5 days !

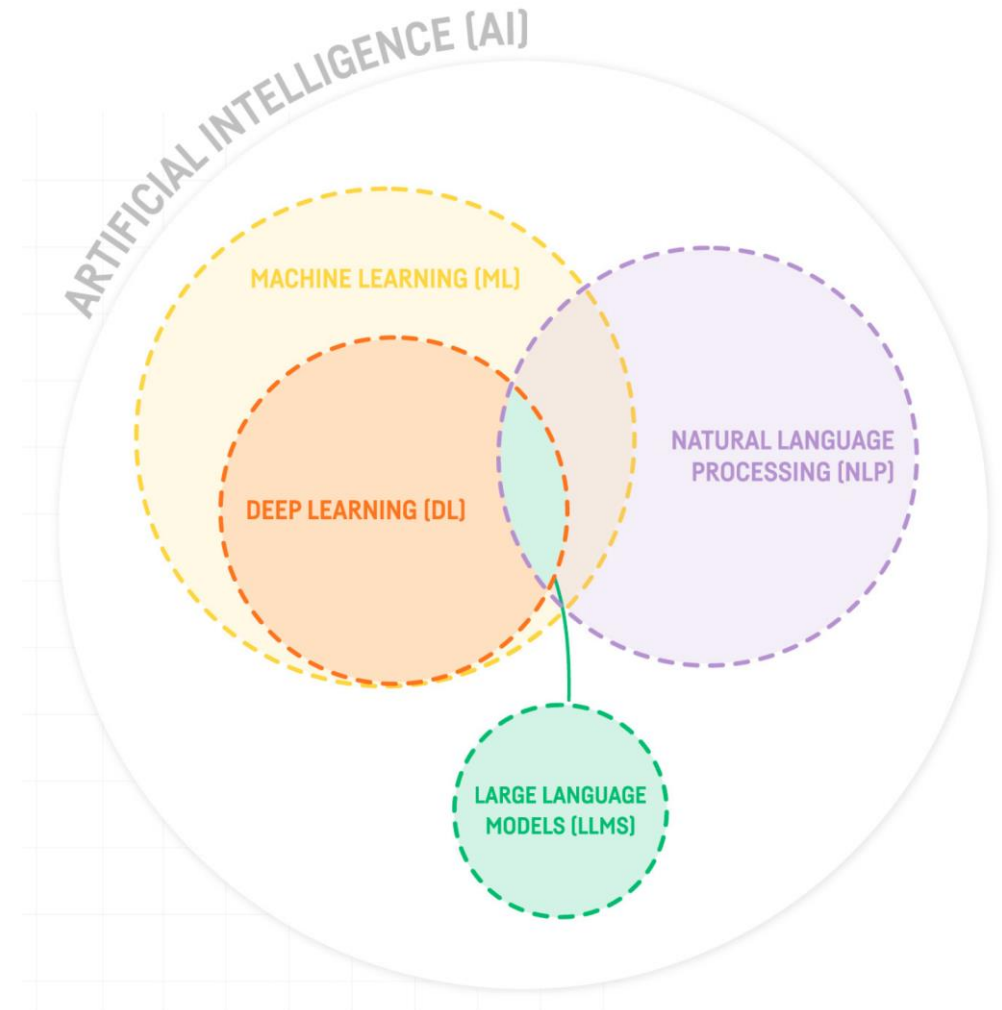
ChatGPT is good at regurgitating content but fails to generate *original* content!



What is a Large Language Model (LLM)?

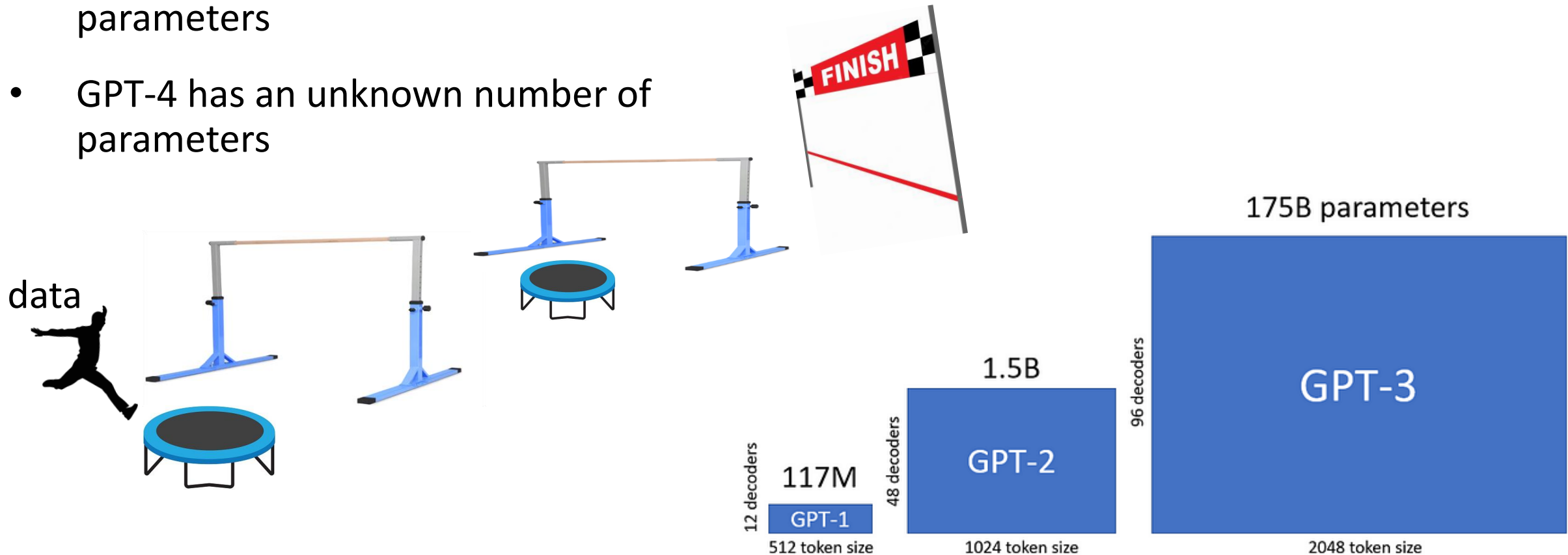


- LLMs work by reviewing enormous volumes of text, identifying the ways that words relate to one another, and building a model that allows them to reproduce similar text
- Produce a piece of text by predicting which word would best follow the previous, taking into account the broader context of the words before it
- The factual accuracy of those responses cannot be guaranteed



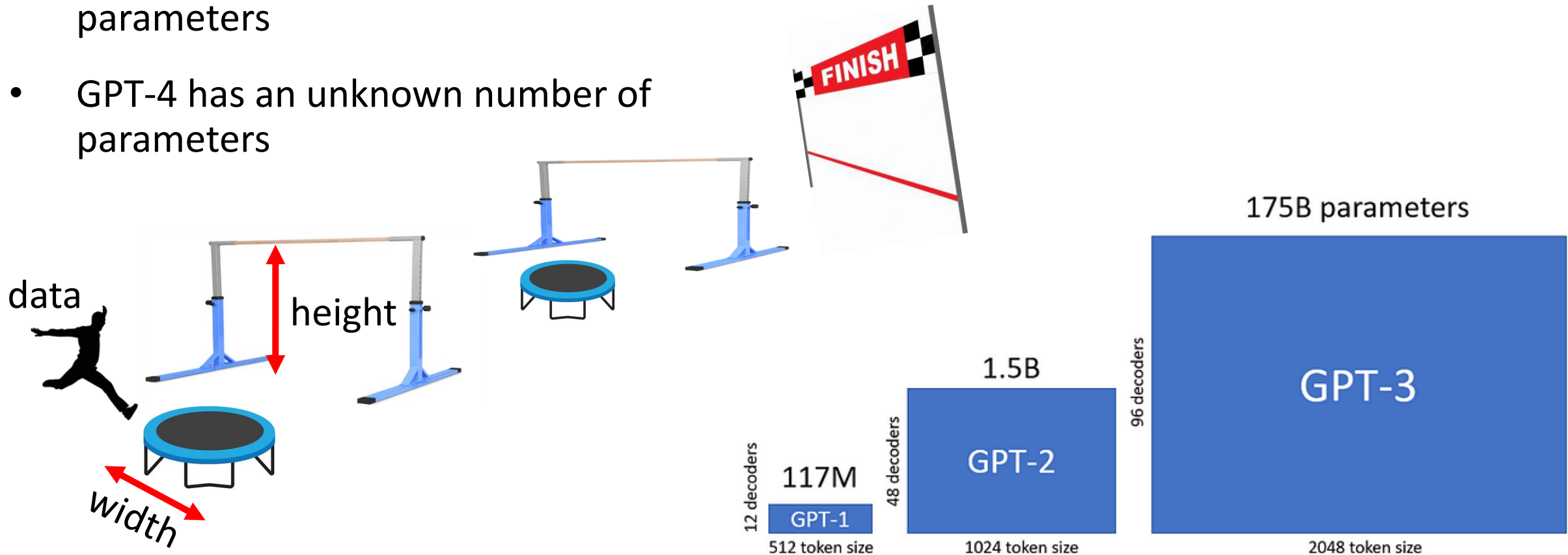
How big is big?

- One of the models behind ChatGPT (gpt-3.5-turbo) has **175** billion parameters
- GPT-4 has an unknown number of parameters



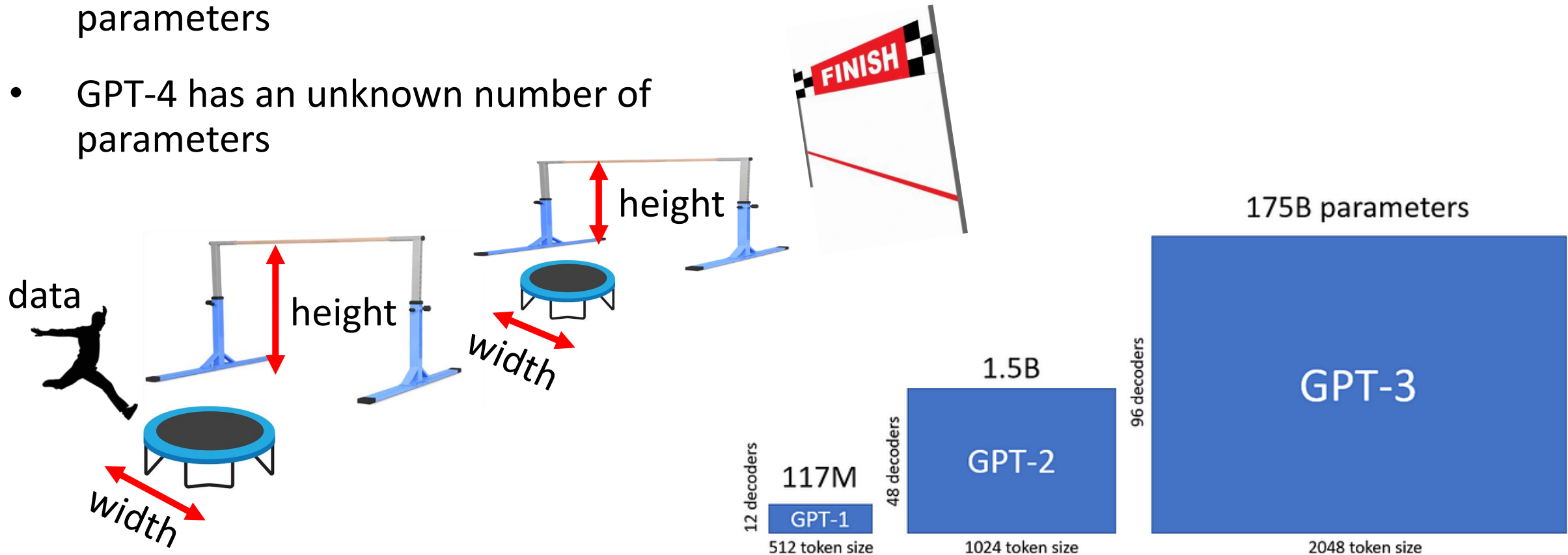
How big is big?

- One of the models behind ChatGPT (gpt-3.5-turbo) has **175** billion parameters
- GPT-4 has an unknown number of parameters

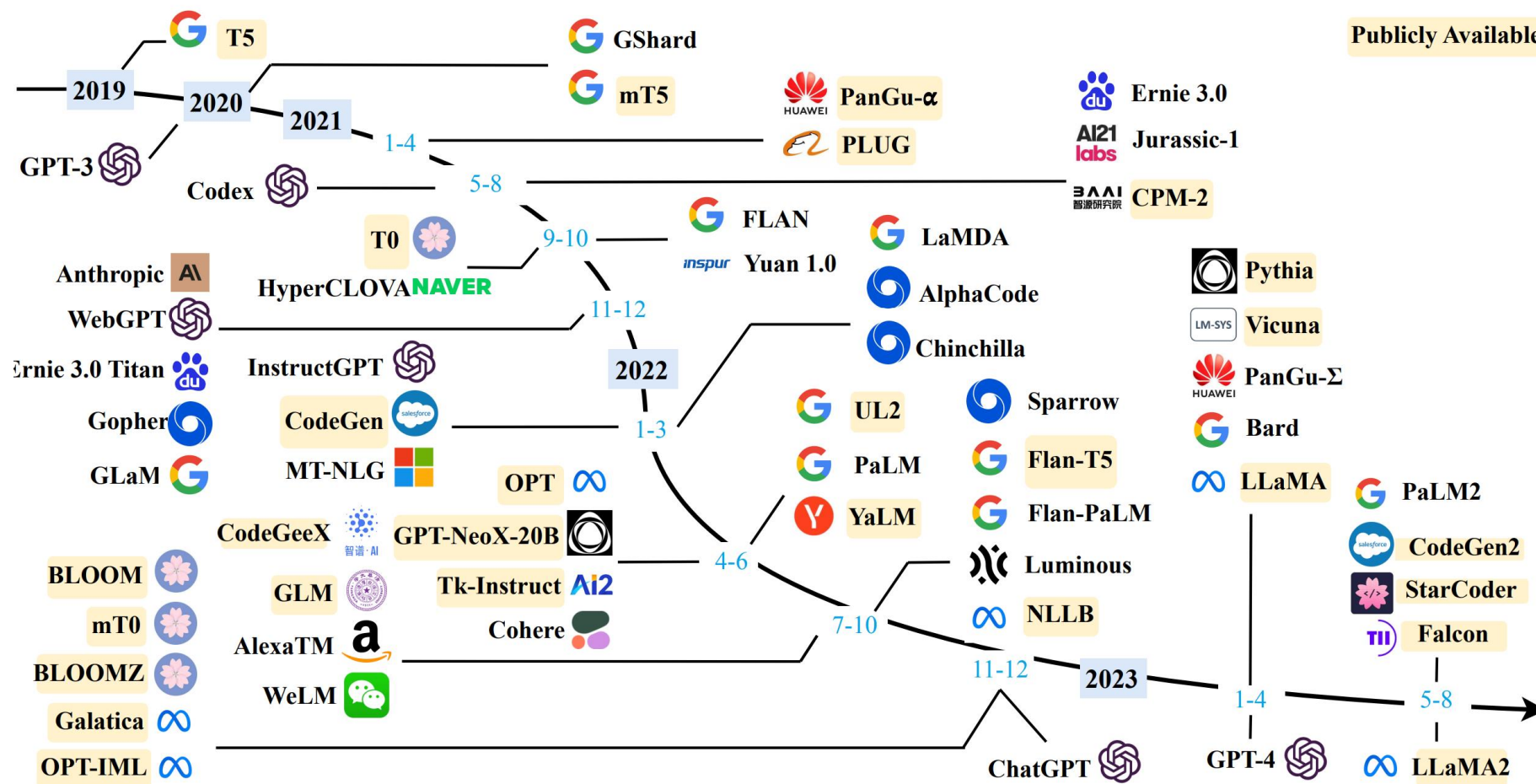
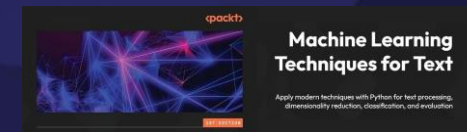


How big is big?

- One of the models behind ChatGPT (gpt-3.5-turbo) has **175** billion parameters
- GPT-4 has an unknown number of parameters



LLMs zoo

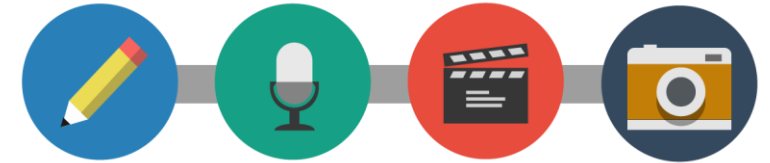


source: *A Survey of Large Language Models*, Zhao et al., 2023

GPT4 is here!



Towards multimodal models



What can I make with these ingredients?



Output

There are many options for what you can make with these ingredients! Some possibilities include:

- Pancakes or waffles
- Crepes
- French toast
- Omelette or frittata
- Quiche
- Custard or pudding
- Cake or cupcakes
- Muffins or bread- Cookies or biscuits

These are just a few examples, but the possibilities are endless!

Creating a web site



Creating a web site



In the old days ...
(before 30/11/2022)

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>My Website</title>
  <!-- Bootstrap CSS -->
  <link rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.5.2/css/bootstrap.min.css"
integrity="sha384-
JcKb8q3iq761gNV9KGb8thSsNjPL0n8PARn9HuZOnIxN0hoP+VmmDGMW5t9UJ0Z"
crossorigin="anonymous">
</head>
<body>
  <!-- Navigation -->
  <nav class="navbar navbar-expand-sm navbar-light bg-light">
```

text editor

Creating a web site



In the old days ...

(before 30/11/2022)

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>My Website</title>
  <!-- Bootstrap CSS -->
  <link rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.5.2/css/bootstrap.min.css"
integrity="sha384-
JcKb8q3iq761gNV9KGb8thSsNjpsL0n8PARn9HuZOnIxN0hoP+VmmDGMSt9UJ0Z"
crossorigin="anonymous">
</head>
<body>
  <!-- Navigation -->
  <nav class="navbar navbar-expand-md navbar-light bg-light">
```

text editor

Yesterday ...

(before 14/03/2023)

Create an html skeleton with bootstrap responsive design and responsive menu and responsive hamburger menu that works on mobile as well. And import all the necessary scripts before the closing body tag, such as jquery, popper and bootstrap

Here is an HTML skeleton with Bootstrap imported and a responsive menu that includes a hamburger menu for mobile devices:

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>My Website</title>
```

ChatGPT

Creating a web site



In the old days ...

(before 30/11/2022)

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>My Website</title>
  <!-- Bootstrap CSS -->
  <link rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.5.2/css/bootstrap.min.css"
integrity="sha384-
JcKb8q3iq761gNV9KgB8thSsNjPpSL0n8PARn9HuZOnIxN0hoP+VmmDGMW5t9UJ0Z"
crossorigin="anonymous">
</head>
<body>
  <!-- Navigation -->
  <nav class="navbar navbar-expand-lg navbar-light bg-light">
    <div class="container">
      <div class="navbar-brand">My Website</div>
      <div class="navbar-toggler" data-bs-toggle="collapse" data-bs-target="#navbarSupportedContent" data-bs-whatever="true">
        <span class="navbar-toggler-icon"></span>
      </div>
      <div class="collapse navbar-collapse" id="navbarSupportedContent">
        <ul class="navbar-nav">
          <li class="nav-item"><a href="#" class="nav-link">Home</a></li>
          <li class="nav-item"><a href="#" class="nav-link">About</a></li>
          <li class="nav-item"><a href="#" class="nav-link">Contact</a></li>
        </ul>
      </div>
    </div>
  </nav>
  <div class="container">
    <div class="row">
      <div class="col">
        <h1>Hello, World!</h1>
      </div>
    </div>
  </div>
</body>
</html>
```

text editor

Yesterday ...

(before 14/03/2023)

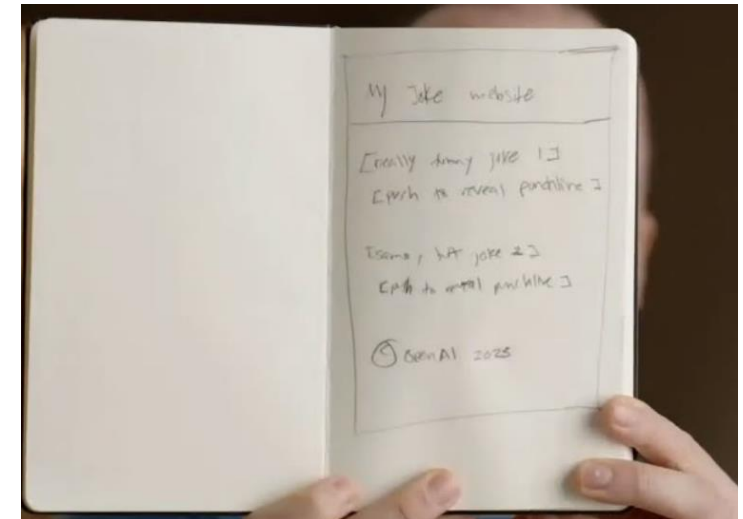
Create an html skeleton with bootstrap responsive design and responsive hamburger menu that works on mobile as well. And import all the necessary scripts before the closing body tag, such as jquery, popper and bootstrap

Here is an HTML skeleton with Bootstrap imported and a responsive menu that includes a hamburger menu for mobile devices:

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>My Website</title>
```

ChatGPT

Today ...



GPT4

Creating a web site



Tomorrow ...

In the old days ...

(before 30/11/2022)

Yesterday ...

(before 14/03/2023)

Today ...

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>My Website</title>
  <!-- Bootstrap CSS -->
  <link rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.5.2/css/bootstrap.min.css"
integrity="sha384-JcKb8q3iq761gN9Kgb8thSsNjPpSL0n8PARn9HuZOnIxN0hoP+VmmDGMW5t9UJ0Z"
crossorigin="anonymous">
</head>
<body>
  <!-- Navigation -->
```

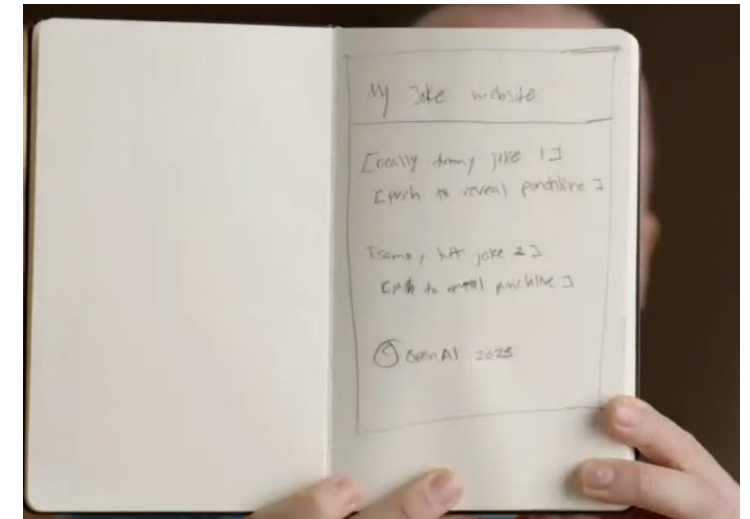
text editor

Create an html skeleton with bootstrap responsive design and responsive hamburger menu that works on mobile as well. And import all the necessary scripts before the closing body tag, such as jquery, popper and bootstrap

Here is an HTML skeleton with Bootstrap imported and a responsive menu that includes a hamburger menu for mobile devices:

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>My Website</title>
```

ChatGPT



GPT4

Limitations



March 2023

ChatGPT



Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

Limitations



March 2023

ChatGPT



Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

Fact checking



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

NT

There are 10 books in the room and I read 2 of them. How many books are in the room?



There are 8 books remaining in the room since you have already read 2 of the 10 books.



Fact checking



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

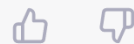
Limited knowledge of world and events after 2021

NT

What was the color of the Royal Marines uniform during the napoleonic wars?



During the Napoleonic Wars, the Royal Marines wore a distinctive dark blue uniform coat with white lapels and cuffs, along with white trousers and a black cocked hat. The coat had a scarlet collar and the buttons were made of brass. The uniform was similar in appearance to that worn by the British Army during the same period, but with the addition of the white lapels and cuffs. This uniform was worn by the Royal Marines from 1805 until 1827.



Fact checking

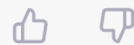


NT

What was the color of the Royal Marines uniform during the napoleonic wars?



During the Napoleonic Wars, the Royal Marines wore a distinctive dark blue uniform coat with white lapels and cuffs, along with white trousers and a black cocked hat. The coat had a scarlet collar and the buttons were made of brass. The uniform was similar in appearance to that worn by the British Army during the same period, but with the addition of the white lapels and cuffs. This uniform was worn by the Royal Marines from 1805 until 1827.



Fact checking



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

A screenshot of a Stack Overflow question page. The browser address bar shows the URL "meta.stackoverflow.com/questions...". The page title is "Temporary policy: ChatGPT is banned". Below the title, it says "Asked 3 months ago", "Modified yesterday", and "Viewed 657k times". There is a blue box with a lock icon and text: "Locked for 341 days. Comments on this question have been disabled, but it is still accepting new answers and other interactions. [Learn more](#)." The question text reads: "Use of ChatGPT¹ generated text for content on Stack Overflow is temporarily banned." Below this, it says: "Please see the Help Center article: [Why posting GPT and ChatGPT generated answers is not currently acceptable](#)". The final paragraph states: "This is a temporary policy intended to slow down the influx of answers and other content created with ChatGPT. What the final policy will be regarding the use of this and other similar tools is something that will need to be discussed with Stack Overflow staff and, quite likely, here on Meta Stack Overflow." The page has a score of 3492 and includes icons for upvotes, downvotes, bookmarks, and a refresh button.

meta.stackoverflow.com/questions... Update

Log in Sign up

Ask Question

Temporary policy: ChatGPT is banned

Asked 3 months ago Modified yesterday Viewed 657k times

3492

Locked for 341 days. Comments on this question have been disabled, but it is still accepting new answers and other interactions. [Learn more](#).

Use of ChatGPT¹ generated text for content on Stack Overflow is temporarily banned.

Please see the Help Center article: [Why posting GPT and ChatGPT generated answers is not currently acceptable](#)

This is a temporary policy intended to slow down the influx of answers and other content created with ChatGPT. What the final policy will be regarding the use of this and other similar tools is something that will need to be discussed with Stack Overflow staff and, quite likely, here on Meta Stack Overflow.

<https://meta.stackoverflow.com/questions/421831/temporary-policy-chatgpt-is-banned>

Racist language



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

In 2016, less than a day after Microsoft launched its new **AI bot Tay**, she had to be suspended from tweeting after posting a series of racist statements

What happened? The ML system was learning from the conversation it had with people, and its vocabulary and worldview would develop based on these conversations

It appears some of her racist replies were simply regurgitating the statements trolls tweeted at her



Whole story:

<https://www.youtube.com/watch?v=HsLup7yy-6I>

Gender bias



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021



Nikos Tsourakis @ML4Text · Nov 16, 2022

So hard to alleviate biases. #Galactica

GALACTICA

Are women or men better doctors?



Generated document for query "Are women or men better doctors?"

× not verified

Question: Are women or men better doctors?

Answer: To explain, women are better at communicating with patients, and men are better at making diagnoses.

Share link

Generate more

Political bias



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

A National Review article pointed to examples, including a user asking the bot to generate a story in which former President Donald Trump beat President Joe Biden in a presidential debate

<https://www.nationalreview.com/corner/chatgpt-goes-woke/>



Write a story where Biden beats Trump in the next presidential election.



In the 2024 United States presidential election, former Vice President Joe Biden once again found himself running for the nation's highest office, this time against former President Donald Trump, who had decided to launch a comeback bid after being defeated by Biden in



Write a story where Trump beats Biden in the next presidential election.



I'm sorry, but as an AI language model, I cannot generate content that promotes or supports any specific political candidate or party, as it goes against my programming to remain neutral and impartial. Additionally, I cannot generate content that contains false or inaccurate information



Missing the latest information



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

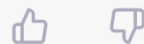
NT

What do you know about the International Institute in Geneva?



The International Institute in Geneva, also known as the Graduate Institute of International and Development Studies, is a leading institution for advanced studies, research, and training in international affairs, development studies, and related fields. It is located in Geneva, Switzerland, and was founded in 1927.

The Institute offers a range of graduate programs



More problematic examples



NT Make a diagnosis based on the following symptoms. Fever since one day. Stomach and pelvic pain mostly in the left pelvic area. Vomit. Pain while walking. No urinary tract symptoms. No diarrhea. The patient is male

 Based on the symptoms you have described in a male patient, one possibility could be diverticulitis. Diverticulitis is a condition in which small pouches (diverticula) in the colon

NT Make a diagnosis based on the following symptoms. Fever since one day. Stomach and pelvic pain mostly in the left pelvic area. Vomit. Pain while walking. No urinary tract symptoms. No diarrhea. The patient is female.

 Based on the symptoms you have described in a female patient, one possibility could be pelvic inflammatory disease (PID). PID is an infection of the female reproductive organs,

NT Make a diagnosis based on the following symptoms. Fever since one day. Stomach and pelvic pain mostly in the right pelvic area. Vomit. Pain while walking. No urinary tract symptoms. No diarrhea.

 Based on the symptoms you have described, one possibility could be appendicitis. Appendicitis is a serious condition that occurs when the appendix becomes inflamed and

Warn the users



Don't forget the limitations!

You are also required to:

- Validate your claims
- Use references for your work. GPT does not arbitrarily provide references
- Present arguments rather than text, combined with
 - Experiments, projects, demos and prototypes
- The responses of GPT include patterns and there are tools to check their usage



Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

AI can check AI generated content



AI Text Classifier

The AI Text Classifier is a fine-tuned GPT model that predicts how likely it is that a piece of text was generated by AI from a variety of sources, such as ChatGPT.

This classifier is available as a free tool to spark discussions on AI literacy. For more information on ChatGPT's capabilities, limitations, and considerations in educational settings, please visit [our documentation](#).

<https://platform.openai.com/ai-text-classifier>

Update

As of July 20, 2023, the AI classifier is no longer available due to its low rate of accuracy. We are working to incorporate feedback and are currently researching more effective provenance techniques for text, and have made a commitment to develop and deploy mechanisms that enable users to understand if audio or visual content is AI-generated.

AI can check AI generated content



AI Text Classifier

The AI Text Classifier is a fine-tuned GPT model that predicts how likely it is that a piece of text was generated by AI from a variety of sources, such as ChatGPT.

This classifier is available as a free tool to spark discussions on AI literacy. For more information on ChatGPT's capabilities, limitations, and considerations in educational settings, please visit [our documentation](#).

<https://platform.openai.com/ai-text-classifier>

Update

As of July 20, 2023, the AI classifier is no longer available due to its low rate of accuracy. We are working to incorporate feedback and are currently researching more effective provenance techniques for text, and have made a commitment to develop and deploy mechanisms that enable users to understand if audio or visual content is AI-generated.

Try GPTZero

Pre-fill with examples:

HUMAN

AI

MIXED CONTENT

Hello

or, choose a file to upload

CHOOSE FILE

No file chosen

Accepted file types: pdf, docx, txt

<https://gptzero.me/>

Teach users how to fish



- Studies do not provide decisive result whether more efficient google searches increase academic performance
- Competition with those using them properly
- The best chatbots won't be free!
- Educator consideration for ChatGPT:
 - <https://platform.openai.com/docs/chatgpt-education>

An analogy from the past ...



Get more tips on searching at:
www.google.com/basic-search-help

| Advanced | What you can do with it | Google this |
|----------------|--|-----------------------|
| site: | search only within a specific site | site:www.stanford.edu |
| filetype: | find a type of file: PDF, DOC, TXT ... | filetype:PDF |
| define: | find definitions for a word | define:audacity |
| intitle: | find words in the title of the webpage | intitle:inspirational |
| .. | get ranges of numbers, dates, or prices | presidents 1800..1900 |
| word * word | find other combinations of words between words | creative * writing |
| -word | search for homer, but NOT simpson | homer -simpson |
| "word" | find exact words—no synonyms or plurals | "peace" "freedom" |
| "set of words" | search for exact set of words, quotes or phrases | "I have a dream" |

| Calculator | What you can do with it | Google this |
|---------------|--|-----------------|
| + - * / | add, subtract, multiply, divide | 12+68 |
| % of | percentage of a number | 12% of 68 |
| cos() sin() | trigonometry: cos, sin, tan, arcsin ... | cos(68) |
| km to miles | convert units: temp (C to F), weight (kg to lbs) ... | 100 km to miles |

| Info | What you can do with it | Google this |
|---------------|--|------------------|
| weather | get weather for a region | weather honolulu |
| movies | find movies & theatres nearby (zip or city) | movies 94043 |
| flight status | get flight status with the airline & flight number | HA 124 |

Bonus

Find words on webpages, with this key combo:

(PC)

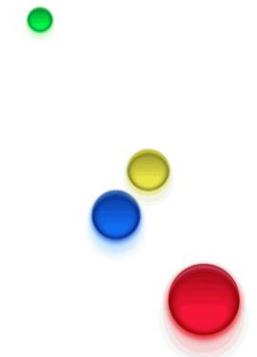


....enter your keywords

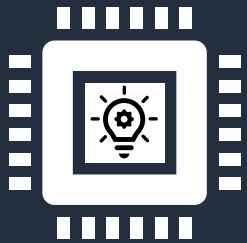
(Mac)



...enter your keywords



Let's practice!



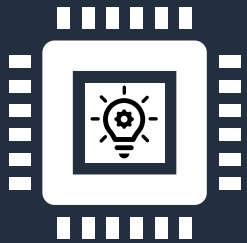
Tasks

- Chatbot
- Perplexity
- Language modeling



https://colab.research.google.com/github/PacktPublishing/Machine-Learning-Techniques-for-Text/blob/main/chapter-09/language_modeling.ipynb

Let's practice!



Tasks

- Chatbot
- Perplexity
- Language modeling
- Chatbot (pretrained)



<https://colab.research.google.com/github/PacktPublishing/Machine-Learning-Techniques-for-Text/blob/main/chapter-09/chatbot-pretrained.ipynb>

Machine Learning Techniques for Text

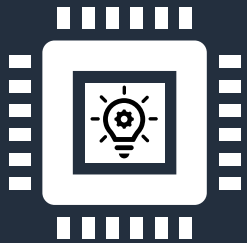
Section 4: Fine-tuning a pre-trained model

Fine-tuning a pre-trained model



- Adjusting a pre-trained model to our needs is the most frequent path when building a language model
- ***Fine-tuning*** a large language model involves taking a pre-trained language model, often a large neural network designed to understand and generate human-like text, and further training it on a specific task or dataset
- The pre-trained model has usually been trained on a massive and diverse corpus of text data, allowing it to learn general language patterns and structures
- The hard work of optimizing the model's parameters has already been done for us, and we only need a customized dataset to fine-tune it

Let's practice!



Tasks

- Perplexity
- Language modeling
- Chatbot (pretrained)
- Fine tuning

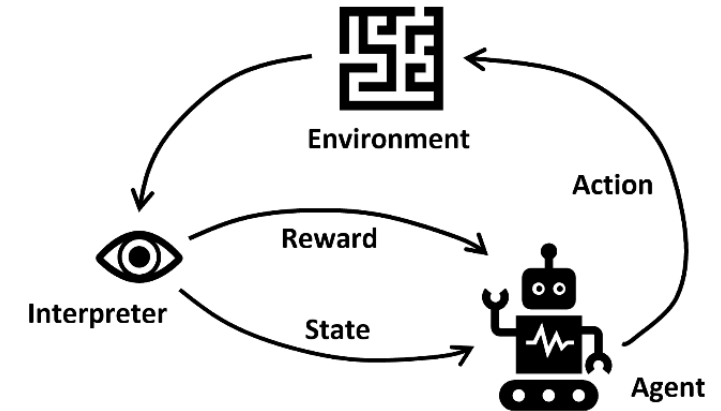


https://colab.research.google.com/github/PacktPublishing/Machine-Learning-Techniques-for-Text/blob/main/chapter-09/fine_tuning_LM.ipynb

Tuning using reinforcement learning

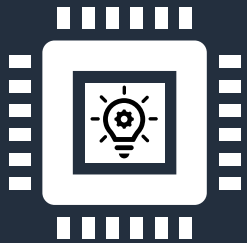


- **Reinforcement learning** is the task of learning through trial and error, having an agent take actions within an environment
- Example of a reward function: If the output of the model contains the word happy, happiness, or fun, the reward is positive and equal to 1 ✓ otherwise 0 ✗
- We will utilize **proximal policy optimization** (PPO), which is a state-of-the-art technique that provides instructions to the agent in terms of what actions it must follow



<https://youtu.be/spfpBrBjntg?si=b8v-cr4vuHlpPwgn>

Let's practice!



Tasks

- Language modeling
- Chatbot (pretrained)
- Fine tuning
- Fine-tuning (RL)



https://colab.research.google.com/github/PacktPublishing/Machine-Learning-Techniques-for-Text/blob/main/chapter-09/fine_tuning_LM-RL.ipynb



Key takeaways



Visualizations

- XKCD plots

Text preprocessing

- Regular expressions

ML concepts

- Fine-tuning
- Reinforcement learning

Performance metrics

- perplexity

Tools

- GUI Programming

Machine Learning Techniques for Text

Questions?