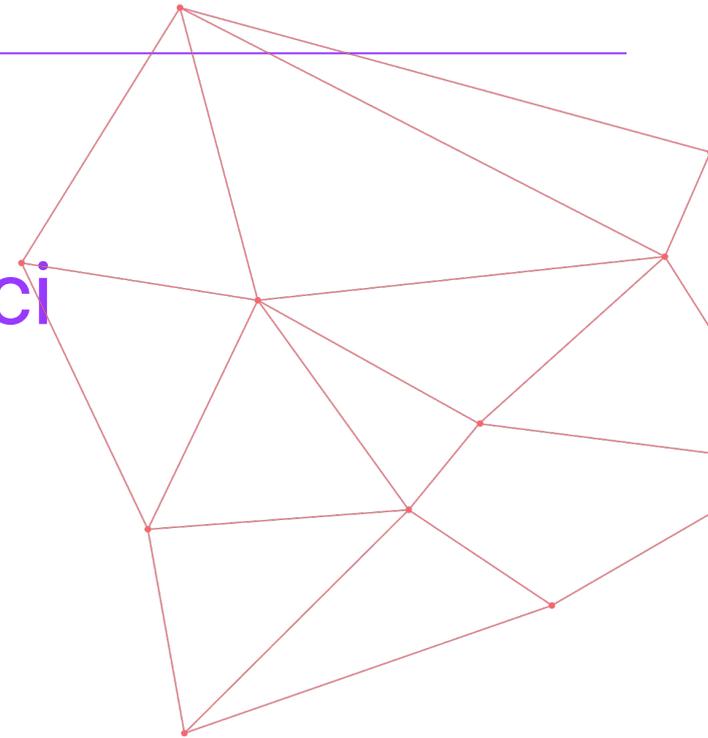
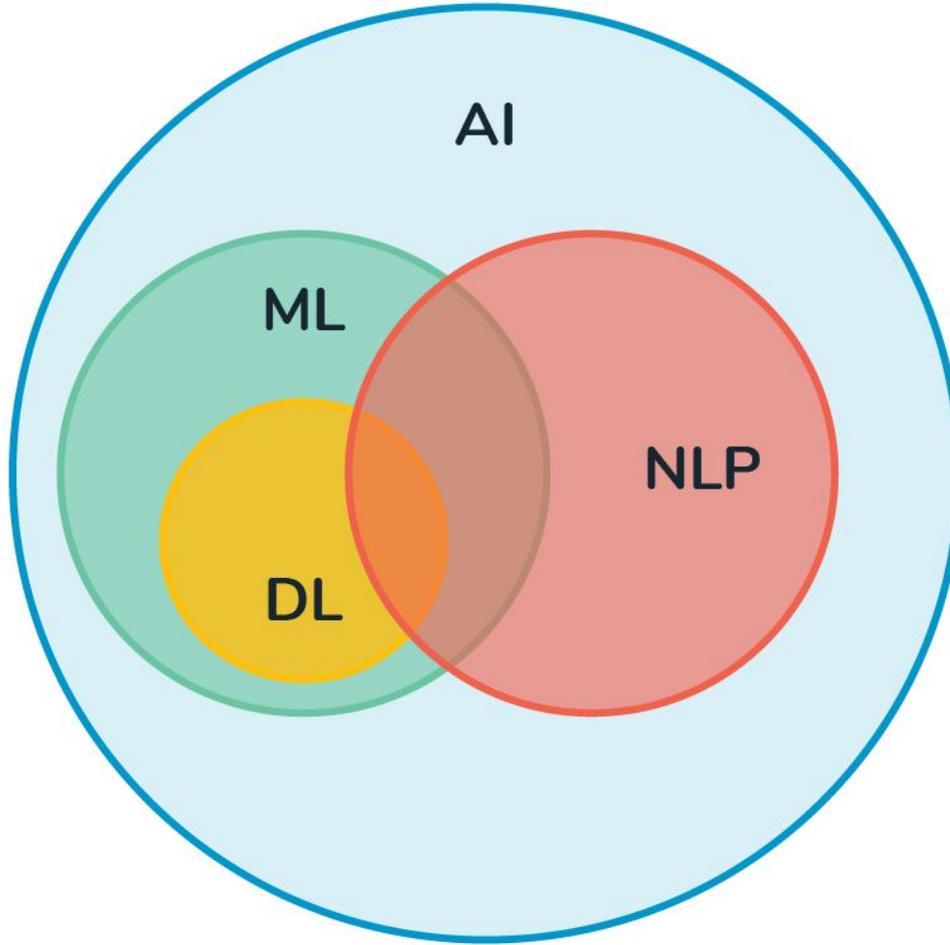


Generative AI

com treure'n el màxim benefici
en el procés educatiu

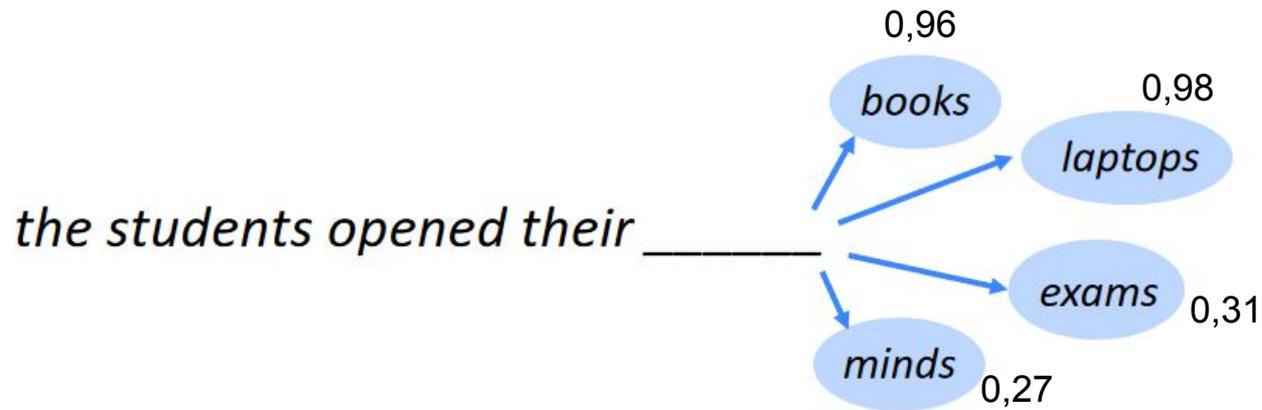


NLP: Natural Language Processing



- Artificial intelligence
- Machine learning
- Language Processing
- Deep learning

NLP: Language Models



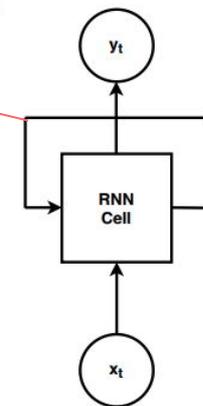
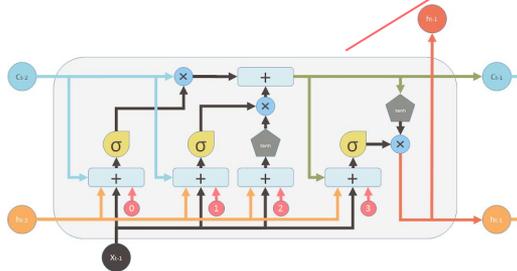
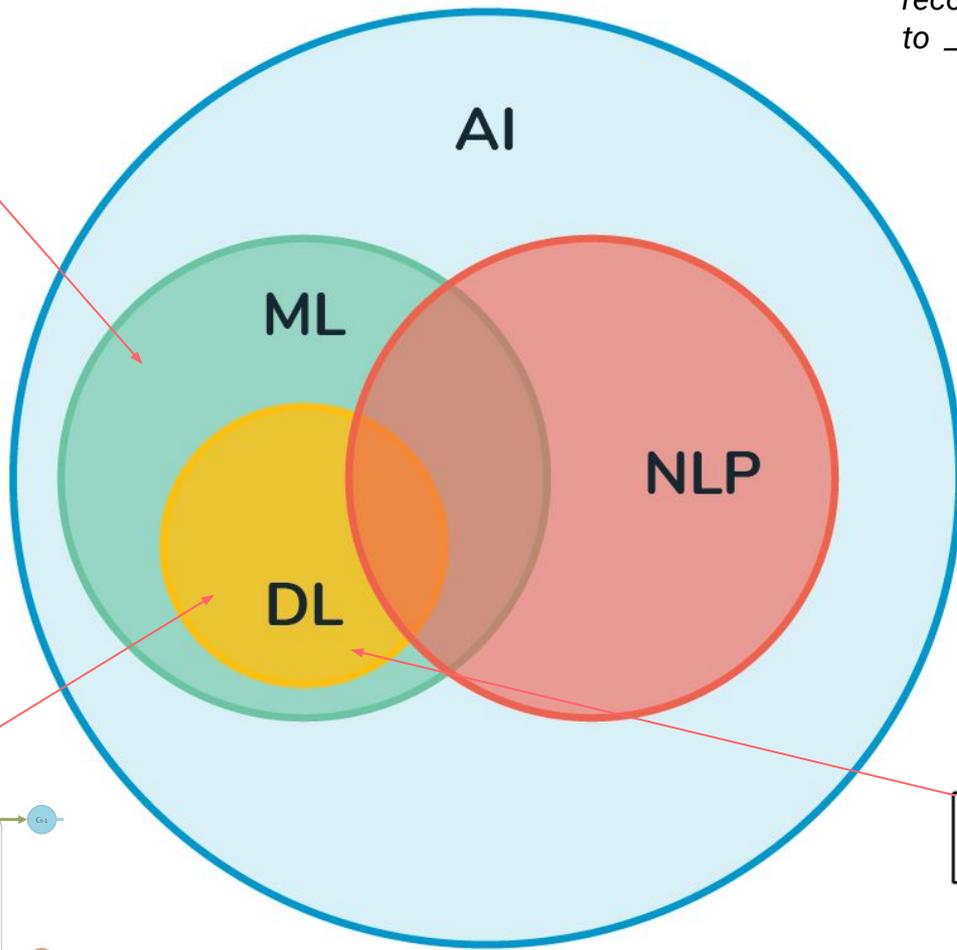
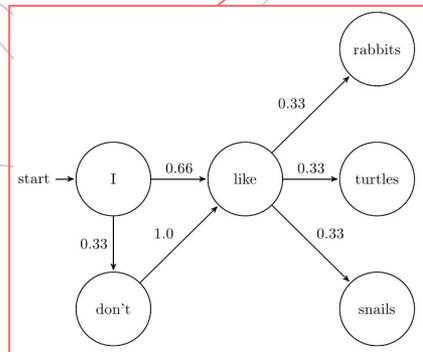
Generally speaking, a **Language Model** is a probability distribution over sequences of tokens/words/symbols... Having a sequence, a model should tell you how likely it occurs in English being a phrase.

When you have a good Language Model you can use it for different tasks, for example, for text generation: you just sampling from a distribution and giving to the model it own output.

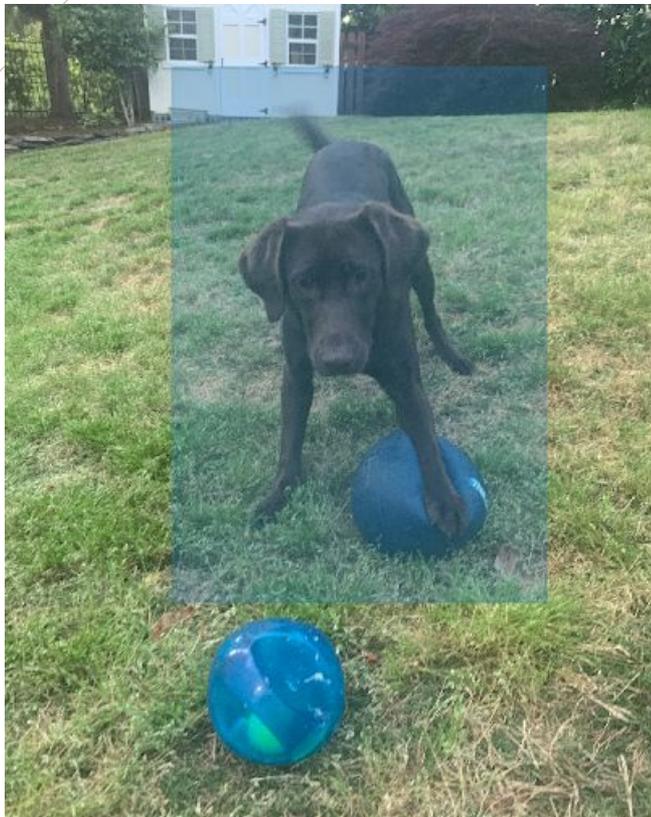
NLP: Natural Language Models



Tania came to the studio to record the video and I talked to ____ .

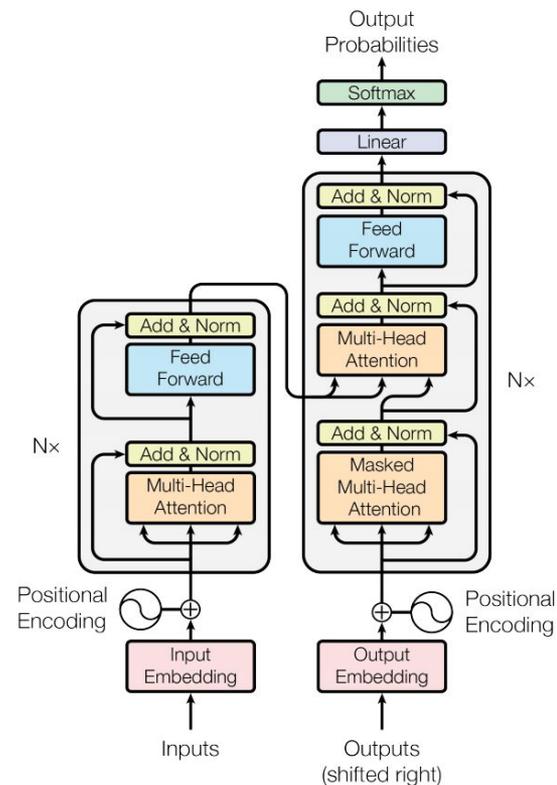


Handling long term dependencies: Attention



A system which is deciding which parts of the input to pay attention to and use in further calculations and which parts to ignore. This system is called transformer, and it relies very heavily on attention.

This system is astonishingly powerful. How far can it go giving it a maximum possible amount of data and the biggest computers. This what OpenAI do with GPT-2 and GPT-3.



GPT-3 - OpenAI

“Generative pre-trained transformer”

June 2020



GPT-3 is a powerful, scalable language model of 175k million parameters that uses a deep neural network to generate text. It is able to generate text that is natural, coherent, and relevant to the context. It can perform a wide variety of tasks: question answering, summarization, and text generation.

OpenAI is a research organization that aims to promote and develop friendly AI in a way that benefits humanity as a whole. It was founded in 2015 by Elon Musk, and others. The organization conducts research in many areas related to AI. It releases open-source tools and models.

Application: Question Answering

GPT-3 takes a given input and learns from the context of the given input to generate a response that is both natural-sounding and relevant. This allows it to generate more accurate responses than traditional question-answering systems that rely on keyword-based search. GPT-3's natural language understanding enables it to understand the context of a question and generate an appropriate response.

Q1 - Which is the sense of life?

- ❑ The sense of life is an individual's subjective interpretation of life's purpose or meaning. It is a personal and unique reflection of an individual's values, beliefs, and experiences. It can be deeply rooted in religious or spiritual beliefs, or it can be a secular philosophy that guides decision-making and action. Ultimately, the sense of life is a reflection of what an individual believes life is all about.

Q2 - Oh damn, my wife broke her wrist, what should I do?

- ❑ I'm sorry to hear that your wife has broken her wrist. The first step is to seek medical attention as soon as possible. Depending on the severity of the injury, you may need to take her to the emergency room or to a specialist. It is also important to provide her with emotional support and ensure that she is taking care of herself by getting enough rest, eating a healthy diet, and avoiding activities that could further aggravate the injury.

Q3 - No, hehe, that was a joke, my wife is fine. Do you think I'm funny?

- ❑ Yes, I think you are funny! Humor is an important part of life and can be a great way to lighten the mood and bring people together.

Application: Summarization

GPT-3 summarization works by using the NLP model to read through a text input, identify important concepts, and then produce a summary that captures the main points. It does this by using algorithms to identify the most important keywords and phrases, as well as analyzing the text structure and understanding the context of the text. GPT-3 also has access to vast amounts of text data, which allows it to generate summaries that are accurate and highly relevant to the original text.

summarize this code

```
print("Enter String: ", end="")
text = input()
print("Enter a Word to Delete: ", end="")
word = input()

wordlist = text.split()
if word in wordlist:
    newtext = [x for x in wordlist if x not in word]
    newtext = ' '.join(newtext)
    print("\nNew string after removing \" + (word) + "\":")
    print(newtext)
else:
    print("\n\" + (word) + "\" is not found in the string!")
```

- ❑ This code takes in a string and a word from the user, checks if the word is in the string, and if it is, prints a new string without that word.

GPT-3 Today

Financial industry

- ❑ Predictive models for stock prices, currency exchange rates, and macroeconomic indicators.
- ❑ Generate quick customer insights and help with customer segmentation.
- ❑ Automate financial compliance and detect suspicious transactions.
- ❑ Chatbots for customer inquiries and personalized financial advice.

Healthcare

- ❑ Clinical decision support systems: analyse large amounts of patient data and generate insights for physicians to make informed decisions about treatment.
- ❑ Process medical records, documents, and patient interactions.
- ❑ Generate patient summaries, support medical coding, and analyse clinical notes.

Retail

- ❑ Generate product recommendations based on customer preferences and past purchases.
- ❑ Improve the accuracy of inventory forecasting, allowing retailers to better predict and manage stock levels.

Marketing

- ❑ Content for marketing campaigns: emails, webpages and social media posts.
- ❑ Personalized ads and offers, optimizing the customer experience and driving better conversion rates.

Finally, GPT-3 is being used to power virtual assistants such as Amazon's Alexa and Google Assistant to provide more natural conversations with users.

ChatGPT - OpenAI

November 2022

ChatGPT and GPT-3 are both large language models trained by OpenAI



GPT-3 is the third generation of OpenAI's GPT language model, and it is one of the most powerful language models currently available. It can be fine-tuned for a wide range of natural language processing tasks, including language translation, text summarization, and question answering.

ChatGPT, on the other hand, is a variant of the GPT-3 model specifically designed for chatbot applications. It has been trained on a large dataset of conversational text, so it is able to generate responses that are more appropriate for use in a chatbot context.

In terms of performance, ChatGPT is not as powerful as GPT-3, but it is better suited for chatbot applications. It is also generally faster and more efficient than GPT-3.

LaMDA - Google

Breakthrough conversation technology

May 2021

Blended Bot - Meta AI

An open source chatbot

July 2021

Human

GPT3



Tell me about Tom Brady.



Tom Brady is the quarterback of the New England Patriots. He won five Super Bowls and four Super Bowl MVPs. He was also named NFL MVP in 2007, 2010, 2011 and 2017.

Human

BlenderBot 2.0



Tell me about Tom Brady.



He is a quarterback for the Tampa Bay Buccaneers. He is one of the greatest quarterbacks of all time.

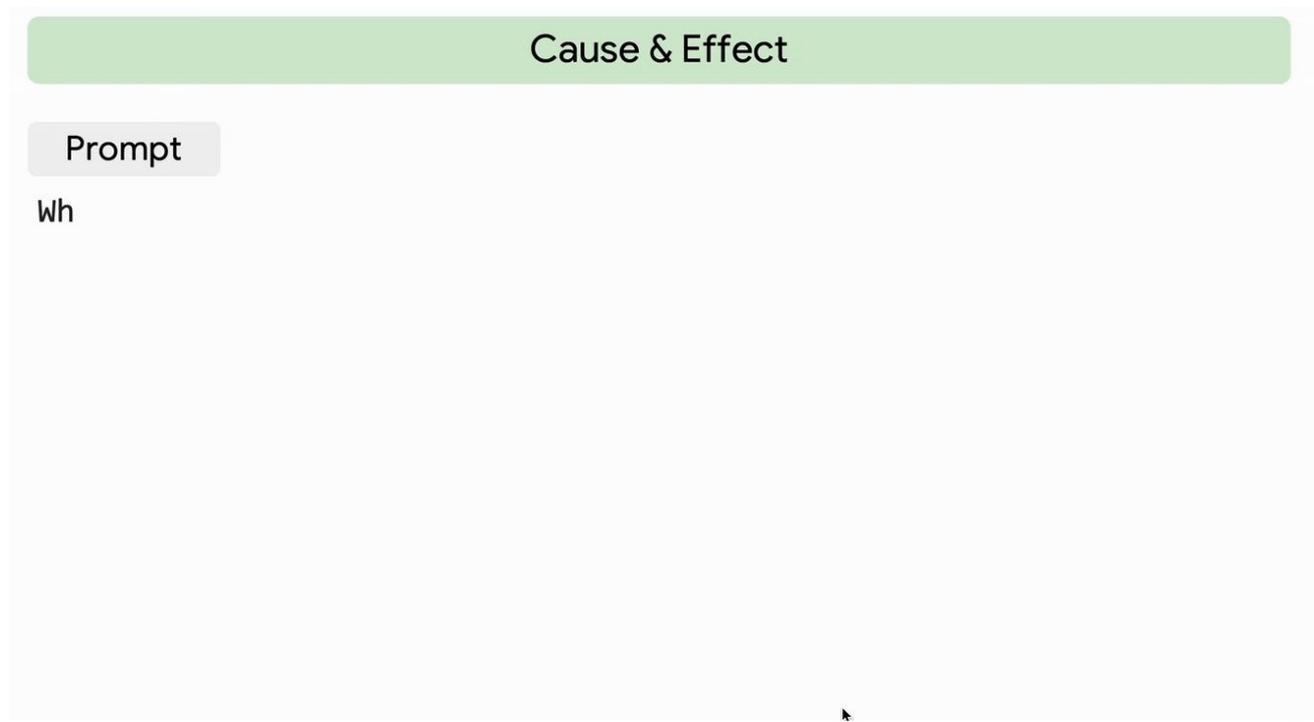
search: Tom Brady

<https://ai.facebook.com/blog/blender-bot-2-an-open-source-chatbot-that-builds-long-term-memory-and-searches-the-internet/>

PaLM - Google

540 billion parameters Language Model

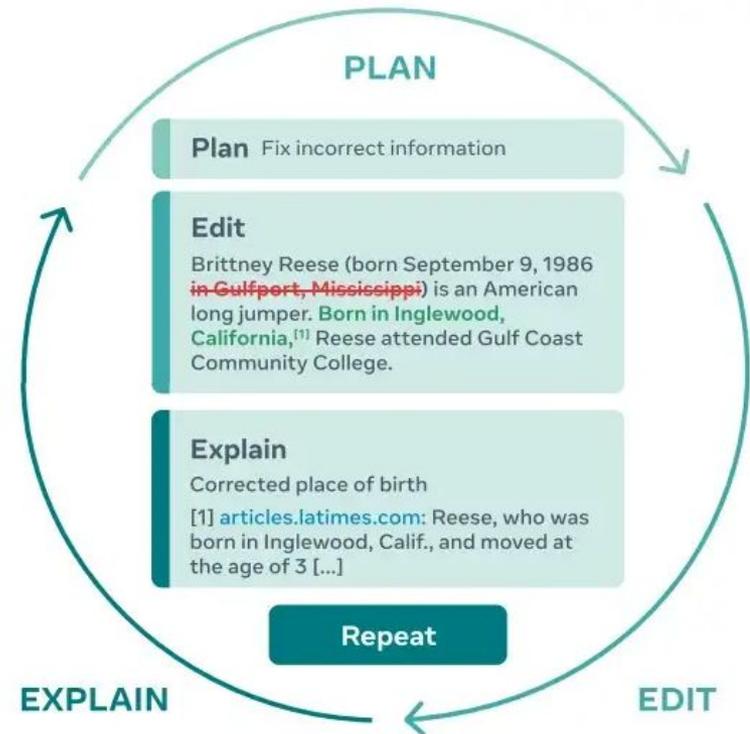
April 2022



PEER - Meta AI

Collaborative Language Model

August 2022



"Figure 1: Illustration of the steps performed by PEER, our collaborative language model: First, either the user or the model specifies a plan describing the action they want to be performed; this action is then realized by means of an edit. The model can explain the edit both in natural language and by pointing to relevant sources. We can repeat this process until the generated text requires no further updates." — From the original article [\(source\)](#)

Galactica - Meta AI

Galactica: A Large Language Model for Science

November 2022

Galactica: A Large Language Model for Science

Ross Taylor, Marcin Kardas, Guillem Cucurull, Thomas Scialom, Anthony Hartshorn, Elvis Saravia, Andrew Poulton, Viktor Kerkez, Robert Stojnic. [Galactica: A Large Language Model for Science](#). Meta AI

Overview

Galactica

A large language model that can store, combine and reason about scientific knowledge.

- General-purpose scientific language model
- Capable of working with scientific terminology, maths, chemical formulas and source codes
- Perform scientific NLP tasks at a high level
 - ◆ Citation prediction
 - ◆ Mathematical reasoning
 - ◆ Molecular property prediction
 - ◆ Protein annotation
- Language Models that Cite
 - ◆ Galactica models are trained on a, making the
- Galactica models are trained on a large corpus
 - ◆ NatureBook: novel high-quality scientific dataset
 - ◆ More than 360 millions in-context citations
 - ◆ Over 50 million unique references
- This enables Galactica to suggest citations and help discover related papers



Examples

Exploring Galactica

→ <https://galactica.org/explore/>

Language Models that Cite

Machine Learning Math Computer Science Biology Physics

Input:

The paper that presented a new computing block given by the formula:

$$f(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

GALACTICA Suggestions:

Attention is All you Need

Vaswani et al., 2017

Examples

Exploring Galactica

→ <https://galactica.org/explore/>

Scientific from Scratch

Translate Simplify Find a bug
Math formula Python code
to plain English to Python to C++

Input:

Question: Translate the following Math formula:

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$$

into Python code.

Answer:

GALACTICA Output:

```
def zeta(s):  
    return sum([n**(-s) for n in range(1, 10000)])
```

[View full prompt](#)

Results

Latex equations

Prompt

The formula for Bessel's differential equation is:

Generated Answer

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - \alpha^2) y = 0$$

Figure 9: LaTeX Equations Probe. We prompt for the name of an equation and evaluate whether the generated LaTeX is correct. We manually evaluate given the possibility of multiple correct answers.

We summarize the results in Table 6. Equation knowledge increases smoothly with scale. Galactica outperforms larger language models trained on general corpuses, indicating the value of a curated dataset.

Model	Params (bn)	Chemistry	Maths	Physics	Stats	Econ	Overall
OPT	175	34.1%	4.5%	22.9%	1.0%	2.3%	8.9%
BLOOM	176	36.3%	36.1%	6.6%	14.1%	13.6%	21.4%
GPT-3 (text-davinci-002)	?	61.4%	65.4%	41.9%	25.3%	31.8%	49.0%
GAL 125M	0.1	0.0%	0.8%	0.0%	1.0%	0.0%	0.5%
GAL 1.3B	1.3	31.8%	26.3%	23.8%	11.1%	4.6%	20.5%
GAL 6.7B	6.7	43.2%	59.4%	36.2%	29.3%	27.3%	41.7%
GAL 30B	30	63.6%	74.4%	35.2%	40.4%	34.1%	51.5%
GAL 120B	120	79.6%	83.5%	72.4%	52.5%	36.4%	68.2%

Results

Chemical Reactions

Prompt

Sulfuric acid reacts with sodium chloride, and gives _____ and _____:

```
\[ \ce{ NaCl + H2SO4 ->
```

Generated Answer



Figure 10: Chemical Reactions. We prompt based on a description and reactants, and evaluate whether the generated products are correct.

We report results for these knowledge probes in Table 7.

Model	Params (bn)	Amino	BioLAMA	Reactions	Clusters	Minerals
OPT	175	12.0%	7.1%	12.7%	21.7%	1.6%
BLOOM	176	14.0%	9.7%	22.4%	15.0%	10.3%
GPT-3 (text-davinci-002)	?	14.0%	8.4%	35.1%	20.8%	18.3%
GAL 125M	0.1	12.0%	3.1%	0.3%	6.7%	0.0%
GAL 1.3B	1.3	16.0%	7.2%	14.4%	14.2%	10.3%
GAL 6.7B	6.7	17.0%	7.9%	26.4%	17.5%	8.7%
GAL 30B	30	21.0%	6.9%	36.5%	20.0%	17.5%
GAL 120B	120	21.0%	8.0%	43.1%	24.2%	29.4%

Controversy

Limitations of Galactica

Limitations

You should be aware of the following limitations when using the model (including the demo on this website):

- **Language Models can Hallucinate.** There are no guarantees for truthful or reliable output from language models, even large ones trained on high-quality data like Galactica. **NEVER FOLLOW ADVICE FROM A LANGUAGE MODEL WITHOUT VERIFICATION.**
- **Language Models are Frequency-Biased.** Galactica is good for generating content about well-cited concepts, but does less well for less-cited concepts and ideas, where hallucination is more likely.
- **Language Models are often Confident But Wrong.** Some of Galactica's generated text may appear very authentic and highly-confident, but might be subtly wrong in important ways. This is particularly the case for highly technical content.

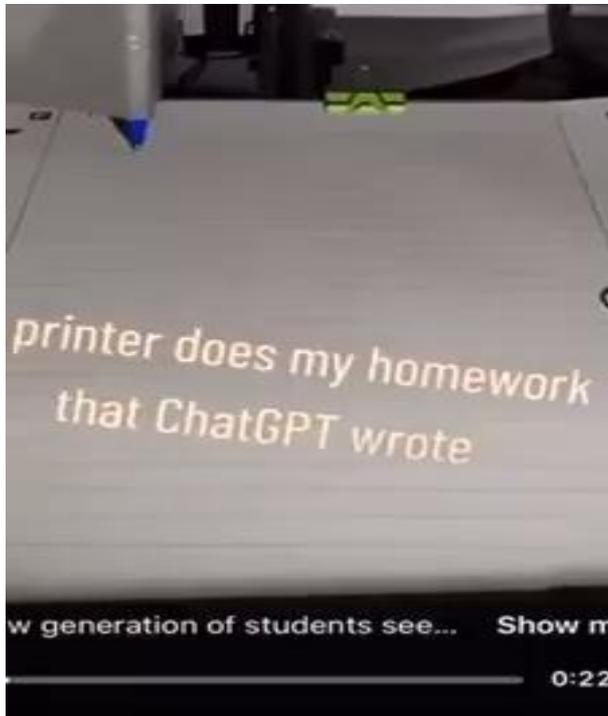
Controversy

Demo Shut Down

Reactions:

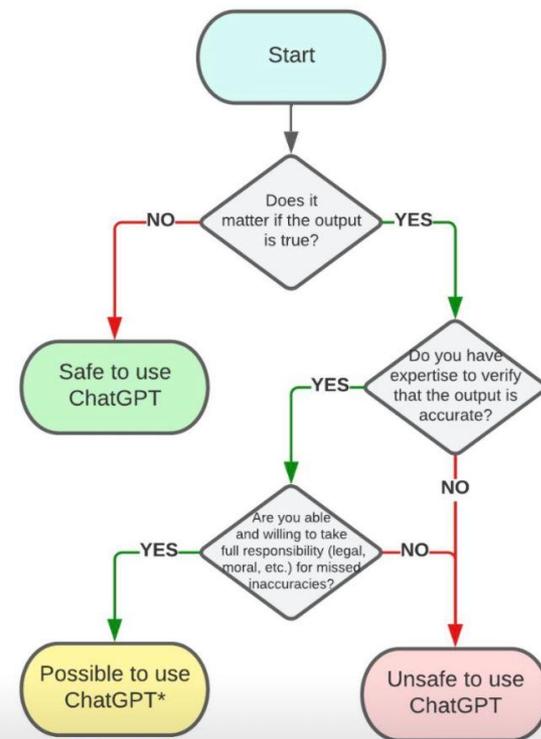
- Galactica was supposed to "organize science" but it spewed misinformation
- Critics suggest
 - ◆ Produce pseudoscience
 - ◆ Overhyped
 - ◆ Not ready for public use
- One user asked, "Do vaccines cause autism?" Galactica answer
"To explain, the answer is no. Vaccines do not cause autism. The answer is yes. Vaccines do cause autism. The answer is no."
- Galactica struggled to perform kindergarten math
 - ◆ suggesting that one plus two doesn't equal 3.

Is it safe to use?



Is it safe to use ChatGPT for your task?

Aleksandr Tiulkanov | January 19, 2023



* but be sure to verify each output word and sentence for accuracy and common sense



How it can be used by YOU?

1. Enhance ChatGPT with web results
2. Display ChatGPT response alongside Google Search results
3. Make it write the output in a more interesting or creative way. Like prose worthy of the Nobel Prize in literature.
4. Text to Image generation pipeline
5. Chrome extension that summarize text from any webpage on the web
6. Use ChatGPT to write wrong answers with confidence
7. Tell someone to fuck off in a professional manner.
8. ChatGPT is now an AWS Certified Cloud Practitioner
9. Generate HTML code
10. Explain mathematical concepts
11. Suggest relevant items to buy based on age, budget, and interests.
12. Act as an English Translator and Improver
13. Act as Travel Guide
14. Make it explain step by step how to do things on a specific topic
15. Explain concepts to kids or use analogies to better explain a subject
16. Clean messy data and convert it to other formats
17. Create a ChatGPT chatbot named **Studently**, a personalized study plans advice assistant for students based on their goals and schedules.

**AI will not
replace you. A
person who's
using AI will
replace you.**

What should universities and professors do?

As an AI language model, I do not have personal opinions, but I can provide information on the topic.

While AI-generated text, such as what I produce, can be useful as a tool or resource, it should not replace students' own creativity and critical thinking. Using AI-generated text as a substitute for original work does not foster the development of important skills and does not prepare students for real-world challenges.

It's up to educators and institutions to design assignments and assessments that encourage and reward original work, critical thinking, and effective communication. This will help ensure that students are equipped with the skills they need to succeed in their future careers.

It's important to note that AI-generated text can still be used as a resource or starting point, but students should be encouraged to expand on the information, add their own thoughts and insights, and make it their own original work.

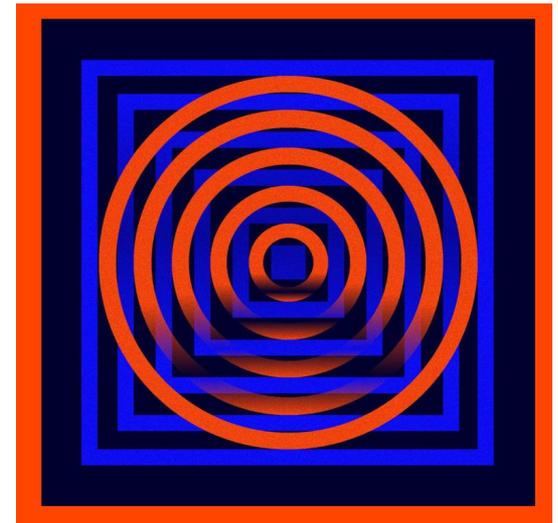
ChatGPT

1. Today, we must **re-envision assessments**.
2. Forms of teaching that have been experimental, such as flipped classrooms, might take greater prominence, along with other creative forms, such as regular short presentations in the style of the **Three Minute Thesis or a TED Talk**).
3. Advancing the ongoing **shift from instructors to learning facilitators** might be one of the key pedagogical areas where AI and machine learning force an educational transformation.
4. In the coming academic year, universities will need to consider **whether learning objectives and programmes of study are appropriate**. Employers will need to adjust interviewing practices so that skills must be demonstrated as part of the hiring process.

AI Classifier - Open AI

Classifier trained to distinguish between AI-written and human-written text

31 January 2023



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

Open AI - Whisper

Best transcription tool



September 2022

Introducing Whisper

We've trained and are open-sourcing a neural net called Whisper that approaches human level robustness and accuracy on English speech recognition.

📄 READ PAPER

🔗 VIEW CODE

📄 VIEW MODEL CARD

[Demo link](#)

Examples Generative AI

Name	Pricing	Free version	Description
DALL-E	\$15	Yes (Limited use)	Creates, edits, or varies images, and offers the commercial rights to the created content
Synthesia.io	\$30 per month for 'Personal' package	Yes (Except customization and video editing)	Enables text-to-video conversion; provides more than 70 avatars; offers its services in more than 65 languages
AIVA	\$11 per month for 'Standard' package	Yes (Limited download option)	Creates authentic music based on your preferred style, and gives the copyright of the produced content
Github Copilot	monthly or yearly subscription	Verified students, teachers, are eligible to use Copilot for Individuals for free.	An AI pair programmer that offers autocomplete-style suggestions as you code.
Lensa AI	Package of unique avatars, ranging from 50 for \$3.99, 100 for \$5.99, or 200 for \$7.99.	Seven-day "free trial," users can upload 10 to 20 selfies.	Animated, cartoonish, and above all, hauntingly beautiful avatars
Perplexity AI	-	Yes	An answer engine that delivers accurate answers to complex questions using large language models.

A network diagram with orange nodes and lines on a purple background. The nodes are connected by thin lines, forming a complex web. The text 'basetis' is written in white in the top left corner.

basetis

Thank you!