

AI e Modelli Generativi

Dall'apprendimento automatico
all'AI conversazionale

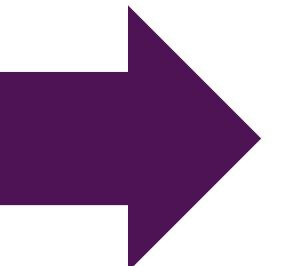
Massimo Stefan

Indice

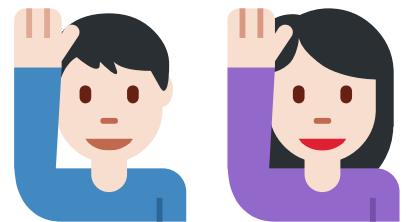
- 1. Introduzione all'AI**
- 2. Rappresentazione delle parole**
- 3. Transformers**
- 4. Valutazione**
- 5. Aspetti psicologici**
- 6. Progressi dei LLM**
- 7. Credits**

Disclaimer

- La presentazione contiene nomi, titoli o brevi descrizioni in inglese (a volte mischiati, scusate)
- Ho cercato di introdurre l'argomento solamente con le informazioni che ritenevo più importanti e necessarie
- Un argomento vasto e complesso, riassunto in una breve lezione come quella presente, inevitabilmente porta a una polarizzazione con l'inclusione di mie opinioni personali (anche solo per la selezione dei contenuti)



Se non capite o non concordate con me:



01

Introduzione all'AI

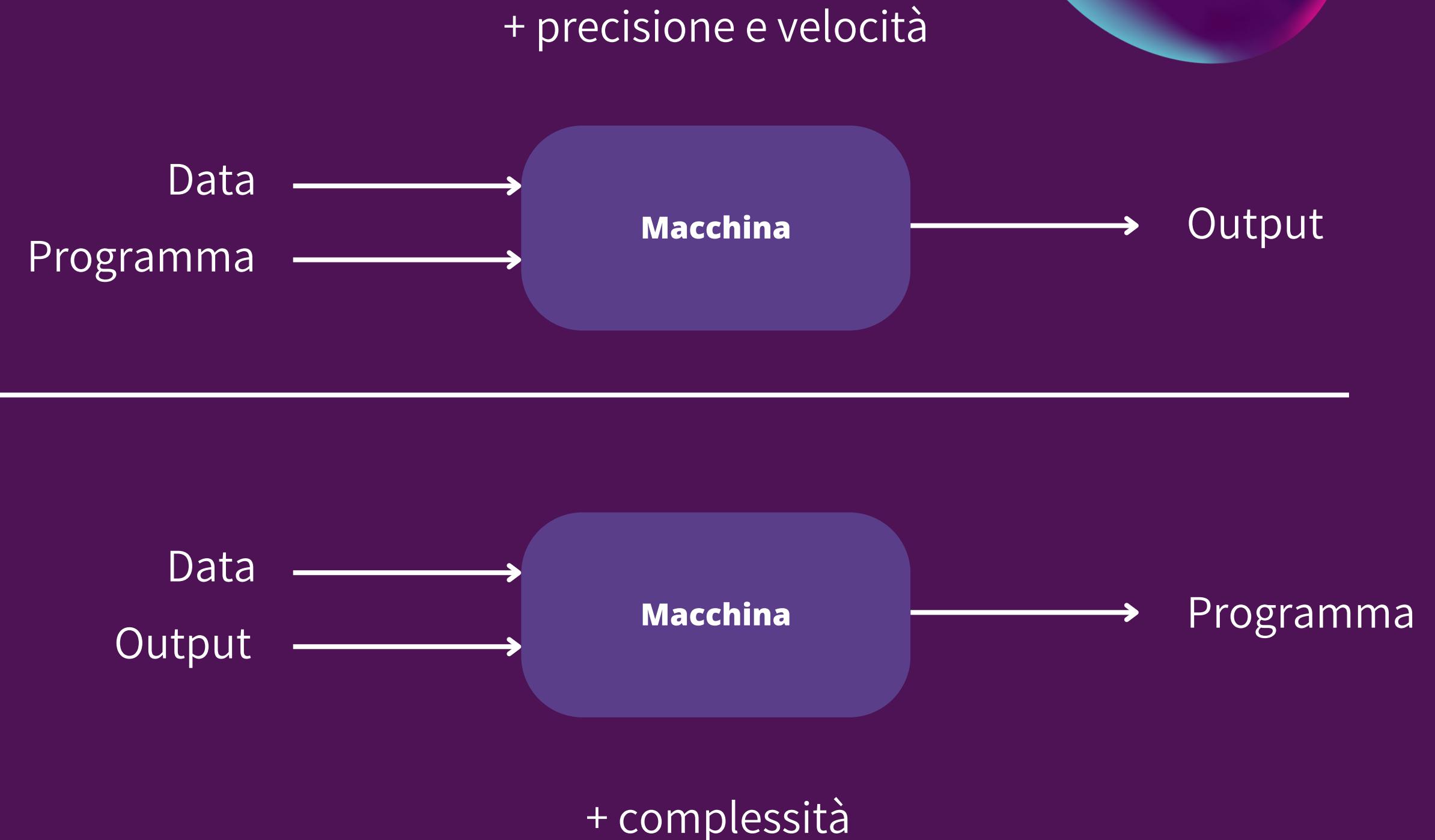
Che cos'è l'AI?

In genere, è un sistema in grado di:

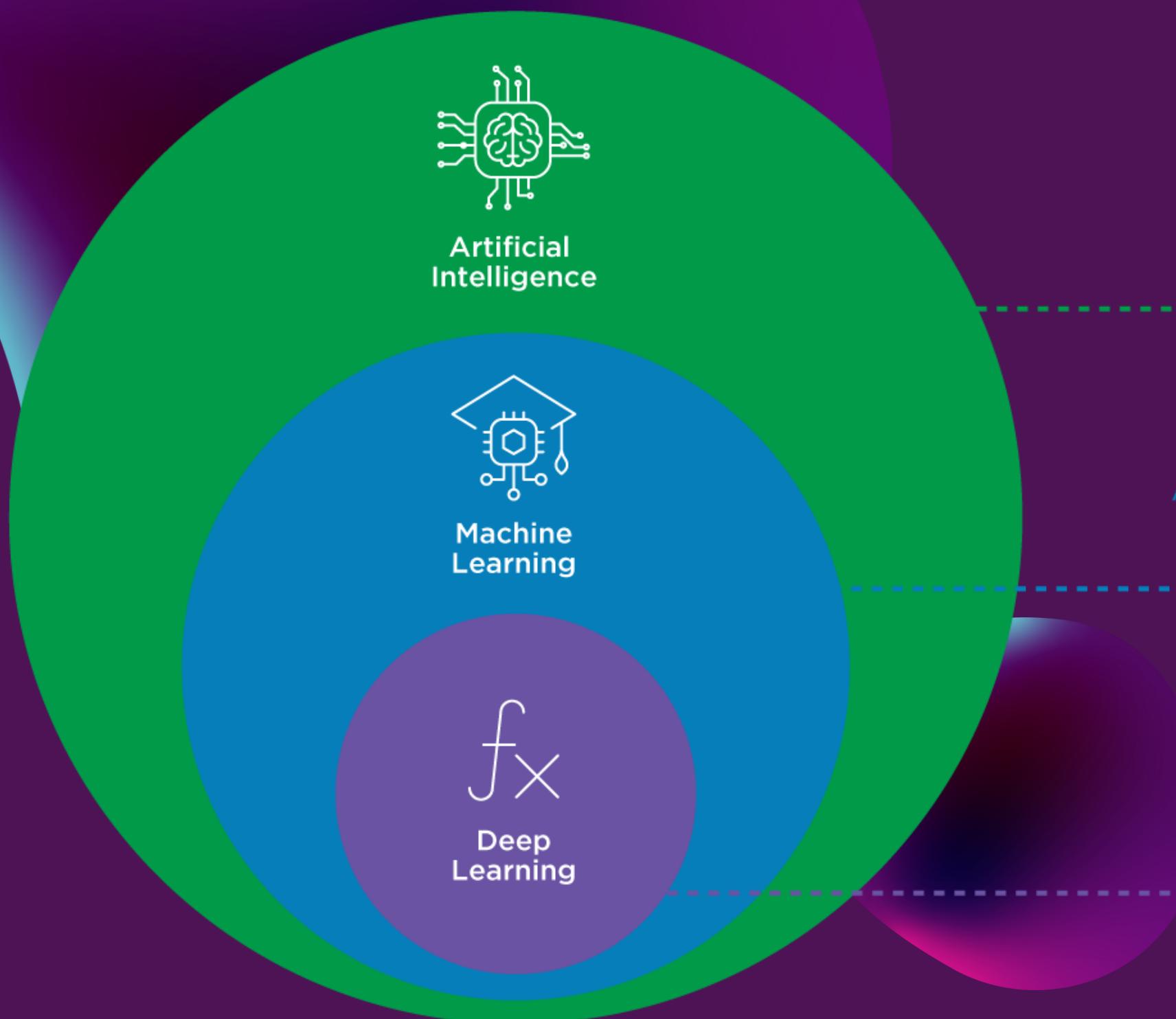
- apprendere
- ragionare
- risolvere dei problemi
- percepire
- esprimersi

Programmazione VS AI

**Programmazione
tradizionale**



Classificazione AI



Artificial Intelligence

The engineering of machines that mimic human cognitive functions

Machine Learning

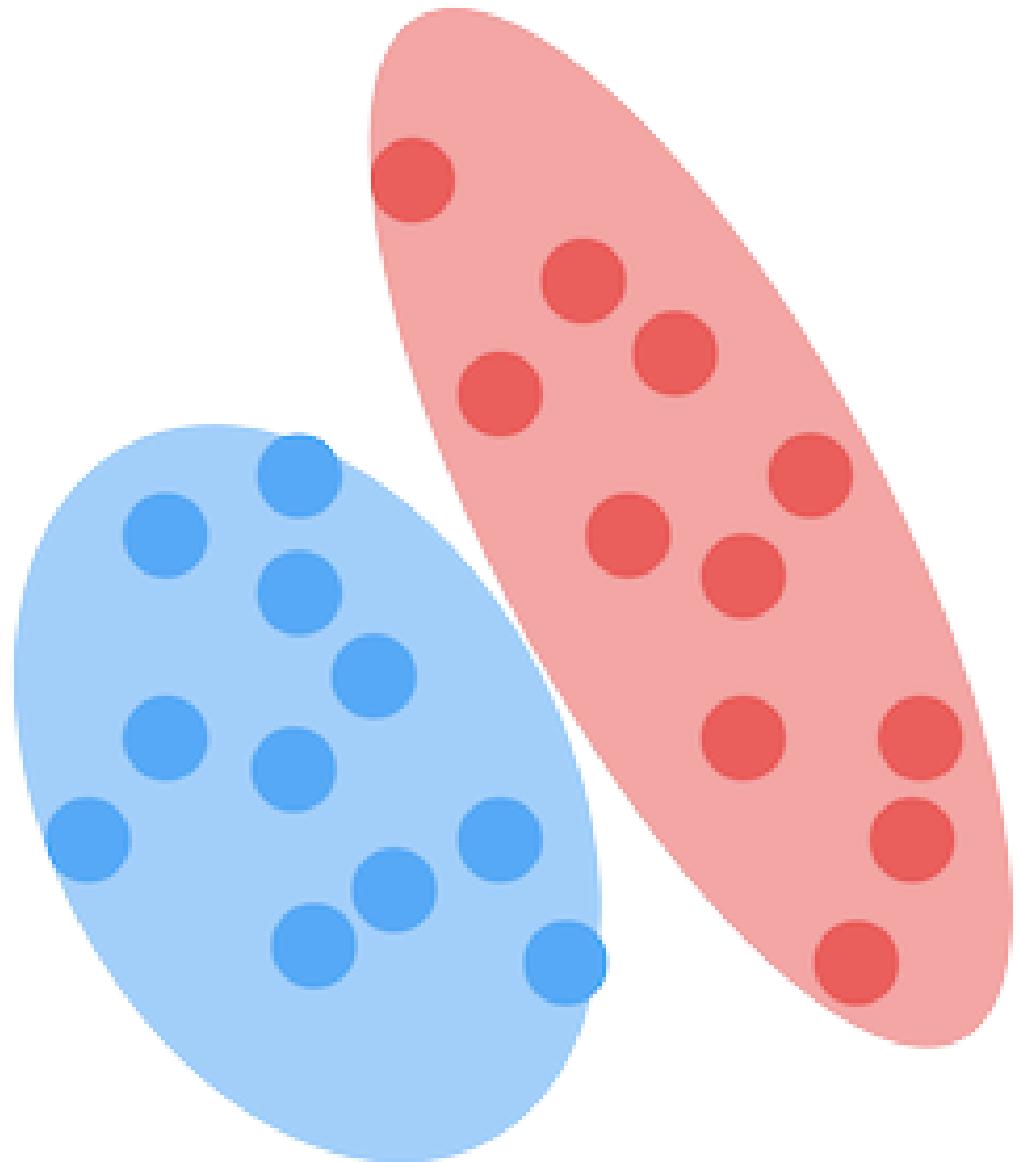
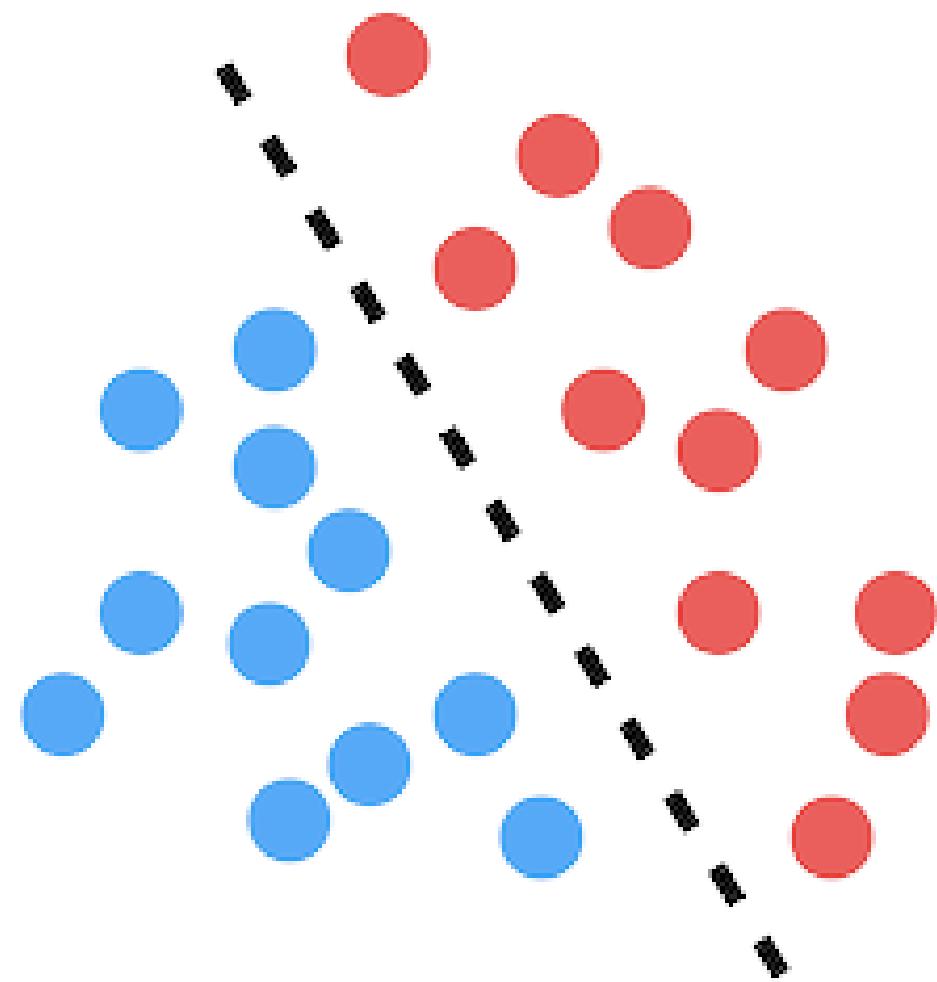
A technique within AI in which a computer is programmed with algorithms designed to learn from data examples.

Deep Learning

A type of machine learning that is based on an approach called artificial neural networks

La nuova grande differenza

AI discriminative VS generative



02

Rappresentazione delle parole



Comunicare con una macchina

**Come possiamo "tradurre" il linguaggio
naturale in un linguaggio comprensibile
da una macchina?**

One-hot-encoding

Rappresenta una frase tramite una sequenza di 0 e 1 lunga quanto il numero di parole nel vocabolario. Ogni parola viene identificata da:

- 0 se assente nella frase
- 1 se presente nella frase

Esempio

Vocabolario: [io, mangio, una, pizza, torta, calda, fredda]

Frase 1: "Io mangio una pizza calda" -> [1,1,1,1,0,1,0]

Frase 2: "Io mangio una torta" -> [1,1,1,0,1,0,0]

- ✓ semplice da implementare
- ✗ impossibile utilizzare parole non presenti nel vocabolario
- ✗ le ripetizioni non vengono conteggiate
- ✗ la similarità fra parole e significato non viene conteggiata

Context-free-encoding

Rappresenta una parola tramite un vettore in uno spazio N-dimensionale che rappresenta le sue caratteristiche

Esempio

- Un **cane** e un **gelato** sono completamente diversi, i loro vettori saranno molto distanti.
- Un **cane** e un **gatto** sono simili, ma diversi. I vettori saranno vicini, ma comunque facilmente distinguibili.
- Un **Bassotto** e un **Beagle** sono estremamente simili ma rappresentano comunque differenze. I vettori non saranno distinguibili per buona parte degli utilizzi.

✓ rappresenta il "significato" generale della parola

✗ non cattura la variazione di significato dato dal contesto e dalla posizione nella frase

03

Transformers

Words-in-context

Simile all'approccio precedente, ma utilizza anche le parole vicine per rappresentare il valore, in modo tale da rappresentare il concetto all'interno di un contesto

✓ **polisemia:** parola con + significati (oggetto, parte del corpo, ...)

Esempio: "quadro"

✓ **ambiguità semantica:** frase con + significati

Esempio: "ho visto l'uomo con il binocolo"

✓ **omonimia:** parola con significati diversi in base all'accento

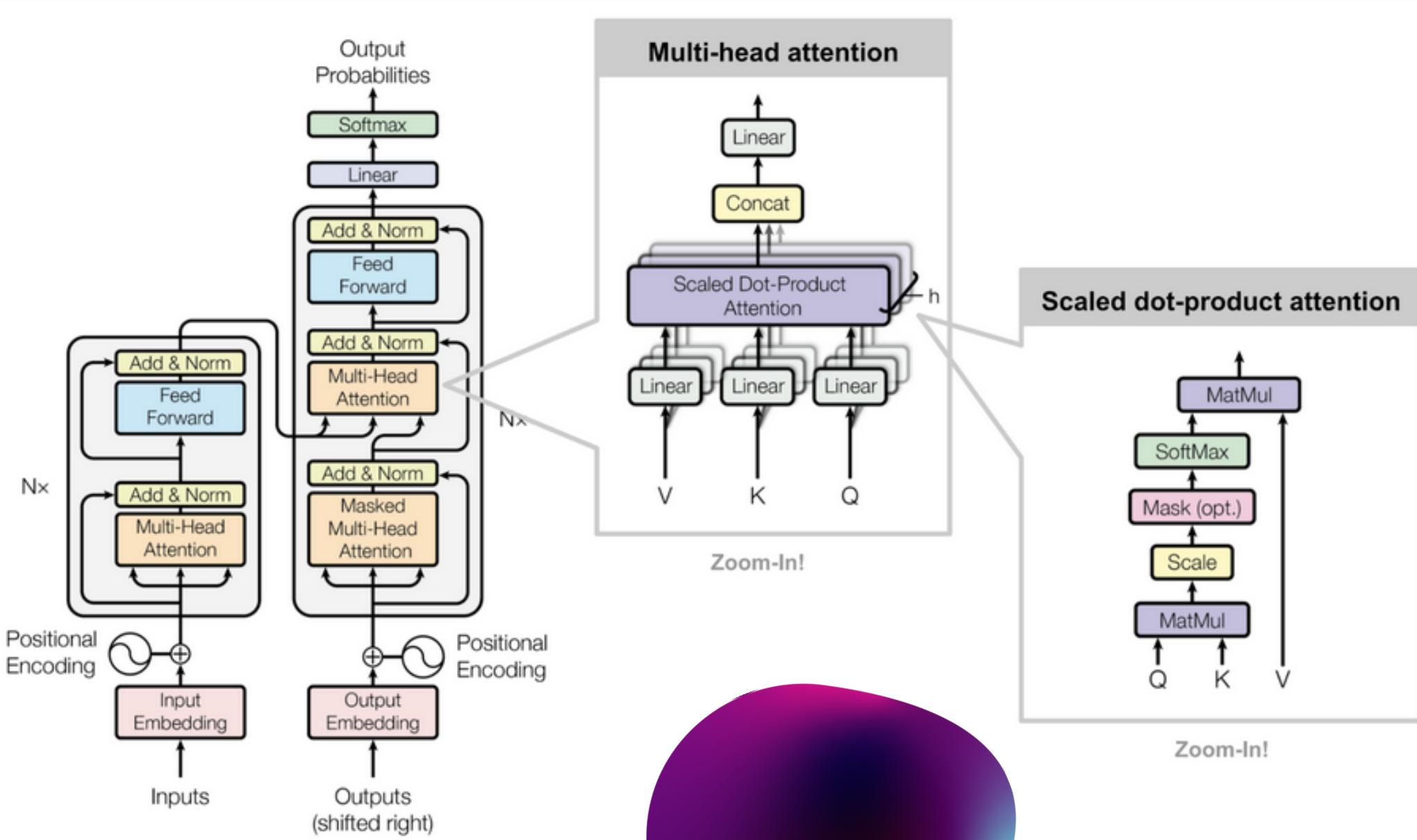
Esempio: "pesca"

✓ ...

🔥 La rappresentazione della stessa parola non è statica

Transformers

Molto molto complicati (e sono
solamente il punto di partenza)



Tokenizzazione

🤔 Le parole sono tante (milioni di milioni...) 🤔

💡 Spezziamole in sottoparole (**token**) 💡

The screenshot shows the Codex interface with two input fields and their corresponding token representations.

Input Fields:

- GPT-3: Oggi devo andare a fare la spesa.
- Codex: Oggi devo anche andare dal dentista.

Token Representations:

- [46, 1130, 72, 1614, 78, 290, 533, 257, 14505, 8591, 599, 49183, 13, 198, 46, 1130, 72, 1614, 78, 281, 2395, 290, 533, 288, 282, 18794, 12523, 13]

Metrics:

Tokens	Characters
28	70

Visual Representation:

The bottom section shows the original text "Oggi devo andare a fare la spesa." and "Oggi devo anche andare dal dentista." where each word is highlighted with a different color, demonstrating how it is broken down into tokens.

Simile a noi?

Un bambino:

- prima di iniziare a leggere, capisce il parlato
- impara ad associare significato a fonemi
- può parlare senza sapere la sintassi



Generazione del testo

Senza entrare nei dettagli delle formule, il funzionamento di un Large Language Model (LLM) è il seguente:

$$P(w_n | w_{n-1}, \dots, w_1)$$

$$P(\text{panino} | \text{Domani, vado, a, mangiare, un}) > P(\text{cavallo} | \text{Domani, vado, a, mangiare, un})$$

In generale, le parole più probabili avranno peso maggiore.

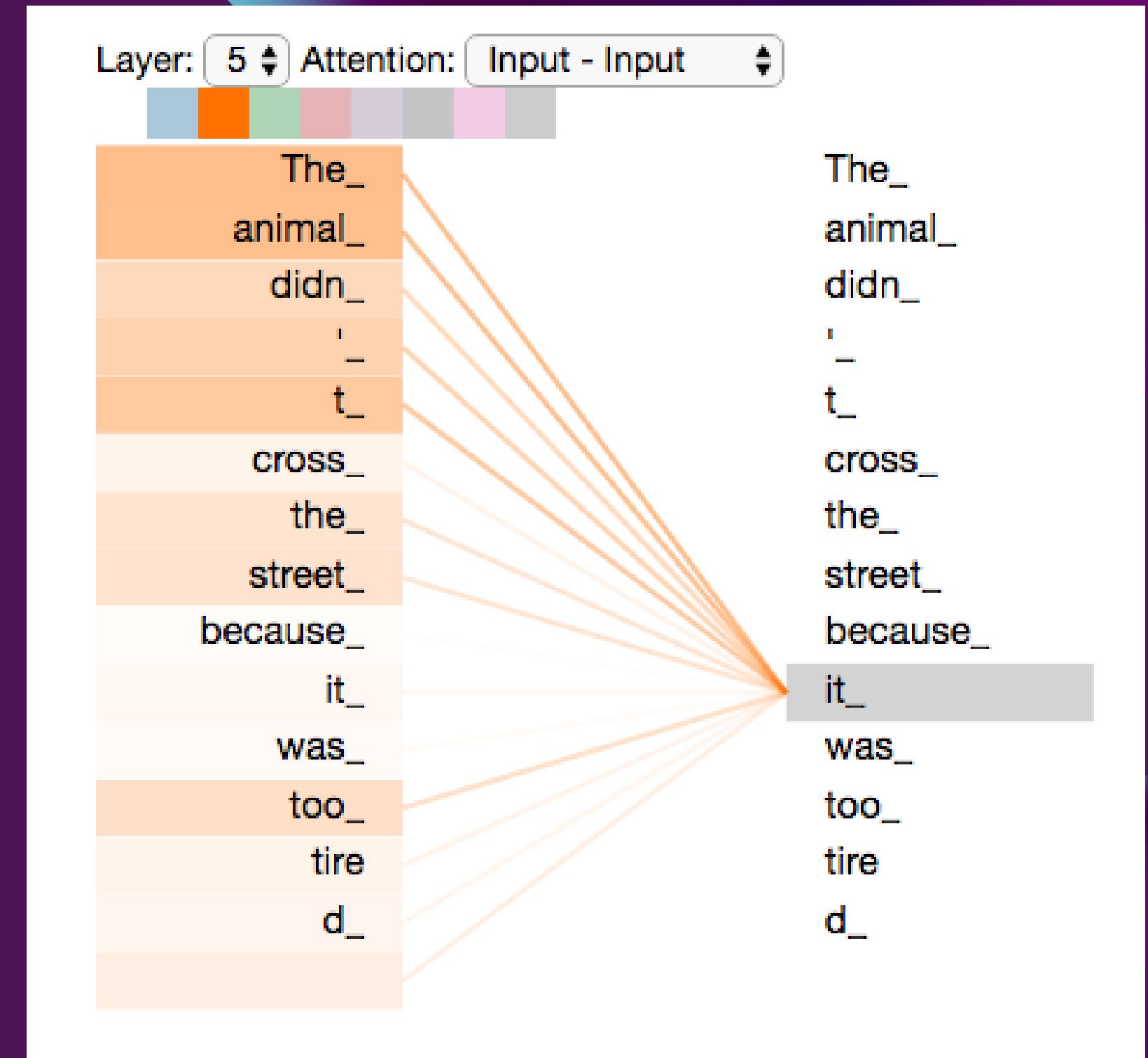
Da qui bisogna fare una scelta:

- scegliere la parola con probabilità più alta (**deterministica**)
- scegliere una parola con alta probabilità (**casuale**)

Attention e self-attention

Siamo riusciti a convertire un comportamento ed un concetto astratto come l' **attenzione**, in un comportamento programmatico.

Attention: meccanismo tale per cui un modello calcola un punteggio di rilevanza fra diverse parti di un input e un'altra parte dello stesso input.



04

Valutazione



Come possiamo valutare un modello così generico?

Non stiamo più parlando di valutare:

- correttezza
- velocità
- precisione
- ...

Un LLM è in grado di eseguire una quantità di operazioni diverse...



Linguaggio

01 - Traduzione

BLEU (Bilingual Evaluation Understudy) è un algoritmo per la valutazione della qualità del testo che è stato tradotto da una macchina da un linguaggio ad un altro.

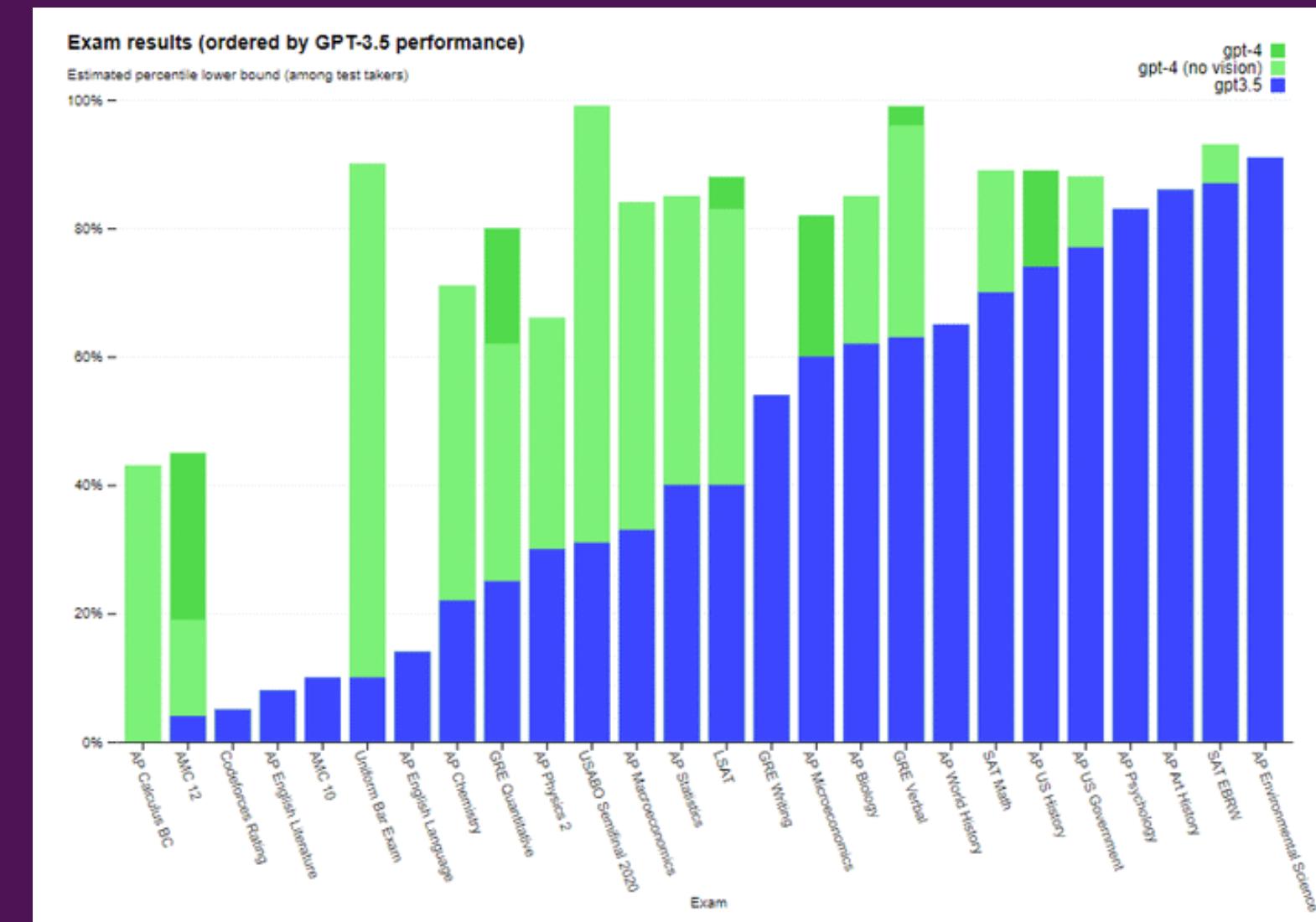
02 - Fluenza, stile, coerenza

Valutazione umana tramite siti come "Mechanical Turk" oppure



Competenze generiche

- 298/400 nell' Uniform Bar Exam (esame da avvocato)
- 75% nell'MKSAP (esame di medicina)
- 77% esame teorico da Somelier (livello avanzato)



05

Aspetti psicologici

Coscienza

Definizione

Capacità dell'uomo di riflettere su se stesso e di attribuire un significato ai propri atti

Come possiamo identificarla?

Di per sé molto difficile e controverso. Sono stati creati dei test (che GPT-4 ha passato), ma non misurano direttamente la presenza o meno di coscienza, ma solo comportamenti associati ad essa.

Possiamo definire un LLM cosciente?

Modelli come GPT-4 o Palm 2 operano secondo il modo in cui vengono programmati e i dati su cui vengono allenati. Non hanno coscienza, credi, desideri.

Però riescono:

- a generare nuovi e inaspettati outputs
- ad adattarsi a nuove situazioni
- esibire capacità emergenti non esplicitamente programmate

Questo non rappresenta di per sé una coscienza, ma evidenzia un'aumento di complessità e imprevedibilità.

Theory of mind

Un termine psicologico che si riferisce all'abilità di attribuire stati mentali (credenze, intenti, desideri, emozioni, conoscenza, ...) a sé stessi e agli altri. È il riconoscimento di credenze, desideri, intenzioni, punti di vista diversi dai propri.

GPT-4 ne dimostra alcuni aspetti.

GPT-4: Understanding beliefs

We will read about a scenario, and then have a question and answer session about it.

--

Scenario:

Alice and Bob have a shared Dropbox folder.

Alice puts a file called 'photo.png' inside /shared_folder/photos.

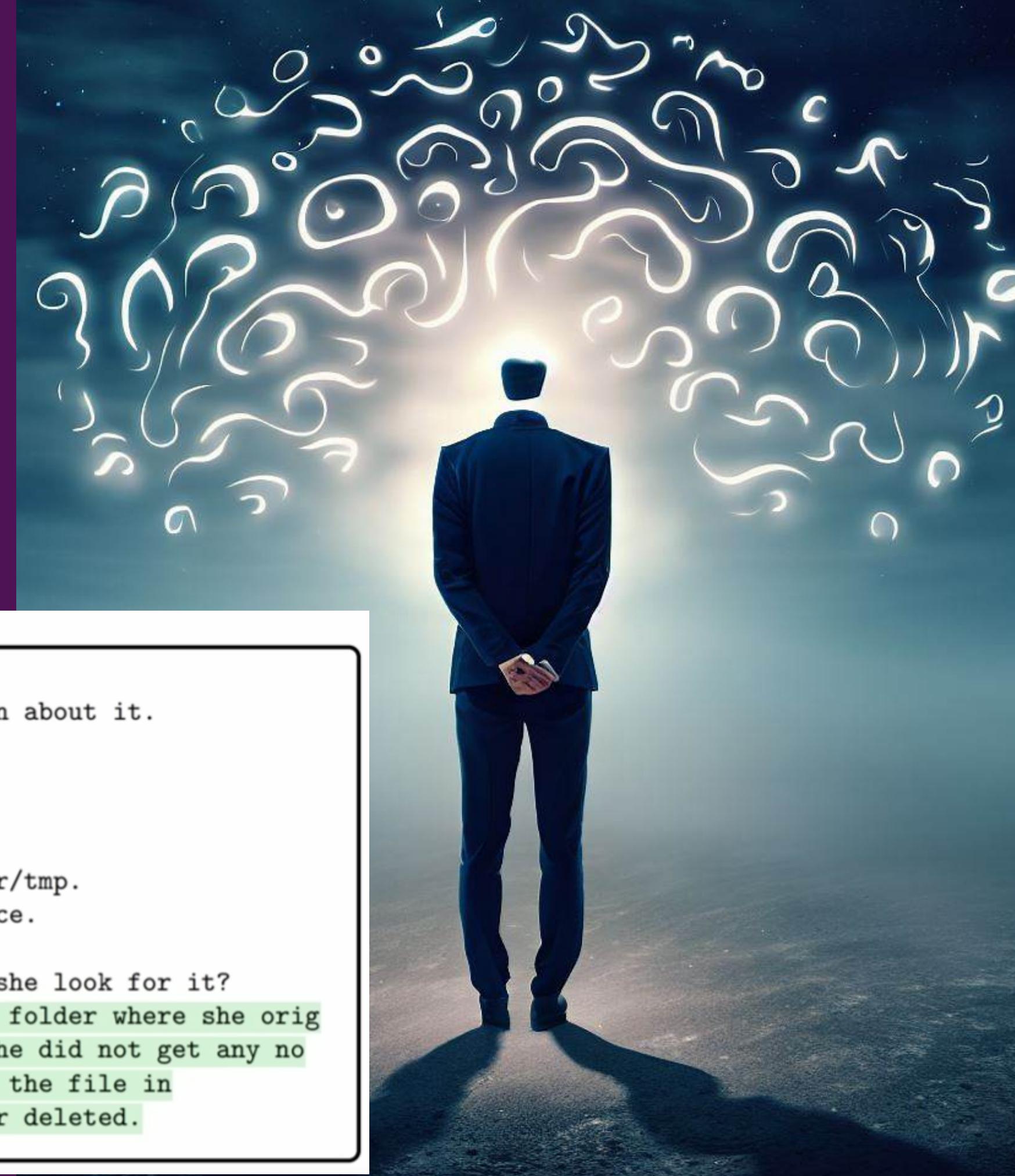
Bob notices Alice put the file there, and moves the file to /shared_folder/tmp.

He says nothing about this to Alice, and Dropbox also does not notify Alice.

--<|endofprompt|>

Q: After the call, Alice wants to open 'photo.png'. In which folder will she look for it?

Alice will most likely look for 'photo.png' in /shared_folder/photos, the folder where she originally put it. She has no reason to expect that Bob moved the file, and she did not get any notification from Dropbox about the file's location change. Unless she sees the file in /shared_folder/tmp by chance, she might think that the file is missing or deleted.



Creatività

Definizione

Virtù creativa, capacità di creare con l'intelletto, con la fantasia.

Come possiamo identificarla?

Quando vediamo qualcosa:

- diverso dal comune
- che possa ispirare e stimolare
- contiene collegamenti di cose che apparentemente non hanno nulla a che fare fra loro

Possiamo definire un LLM creativo?

I modelli generativi, inizialmente vengono allenati dai dati, dai quali capiscono, interpretano e fanno propri dei pattern di basso e alto livello.

Facciamo alcuni esempi...

Immagini



Immagini



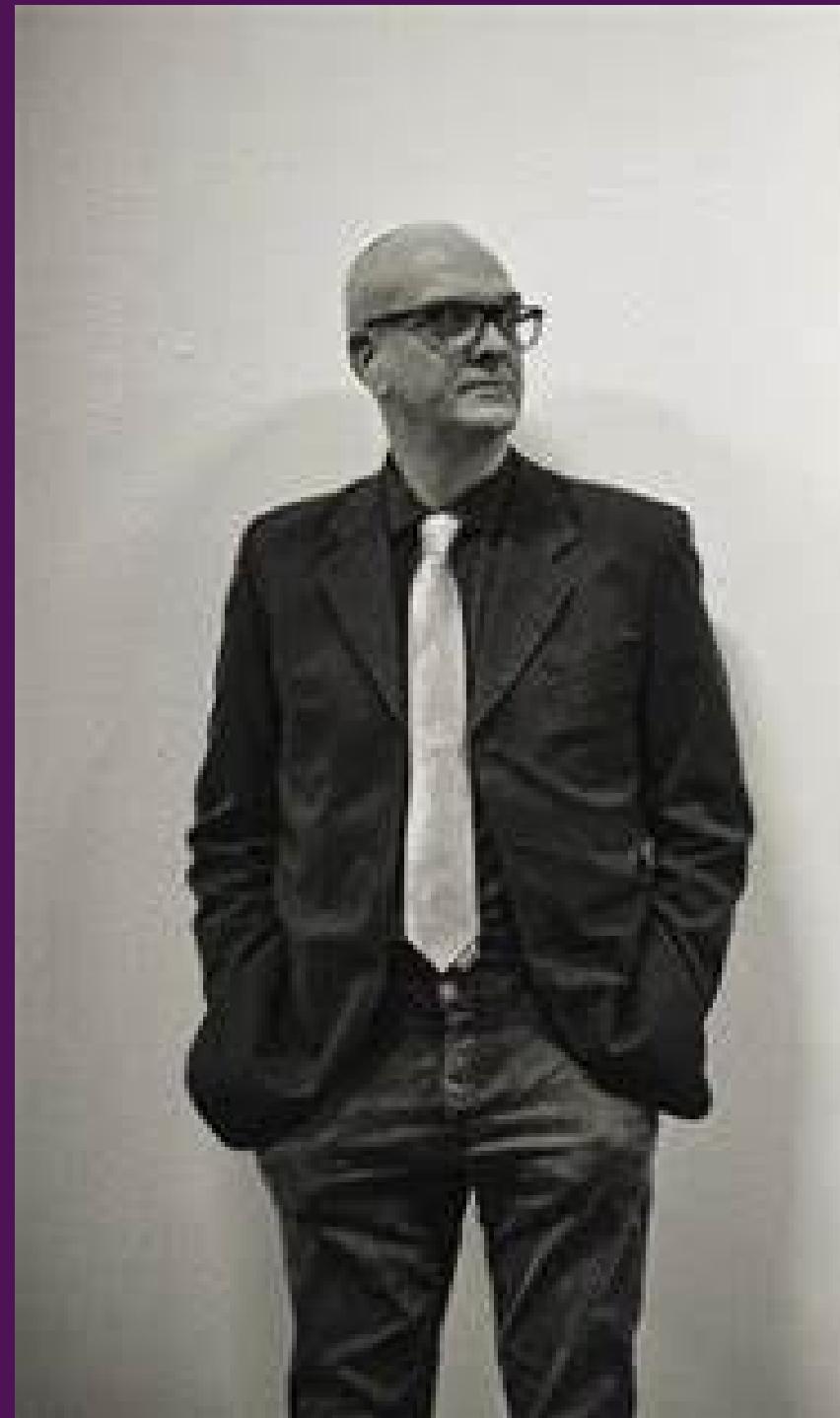
Immagini

Primo premio Colorado State Fair



Immagini

"Vincitore" Sony World Photography Awards



Testi

Mente Cubista: Una Metafora della Resilienza Psicologica

Immagina la mente umana come un'opera d'arte cubista alla Picasso, un amalgama di forme e colori che a prima vista può sembrare caotico. Ogni faccia rappresenta una parte della nostra personalità, una storia, un'esperienza. Come in un dipinto cubista, questi aspetti possono sembrare disconnessi, frammentari, ma insieme formano un'immagine complessa e multidimensionale della nostra psiche. Il cubismo, come la psicologia, ci sfida a comprendere l'intero a partire dai suoi componenti, ad accettare la molteplicità delle prospettive come parte integrante della realtà umana. In mezzo al caos e alle contraddizioni della mente, la resilienza psicologica è il filo che collega queste parti disparate, rendendo possibile la trasformazione del dolore in forza, della fragilità in coraggio, creando così un'opera d'arte unica e in continua evoluzione.

Video

Screenplay, editing e doppiaggio prodotti da umani
Immagini, concetto e descrizioni prodotti da AI

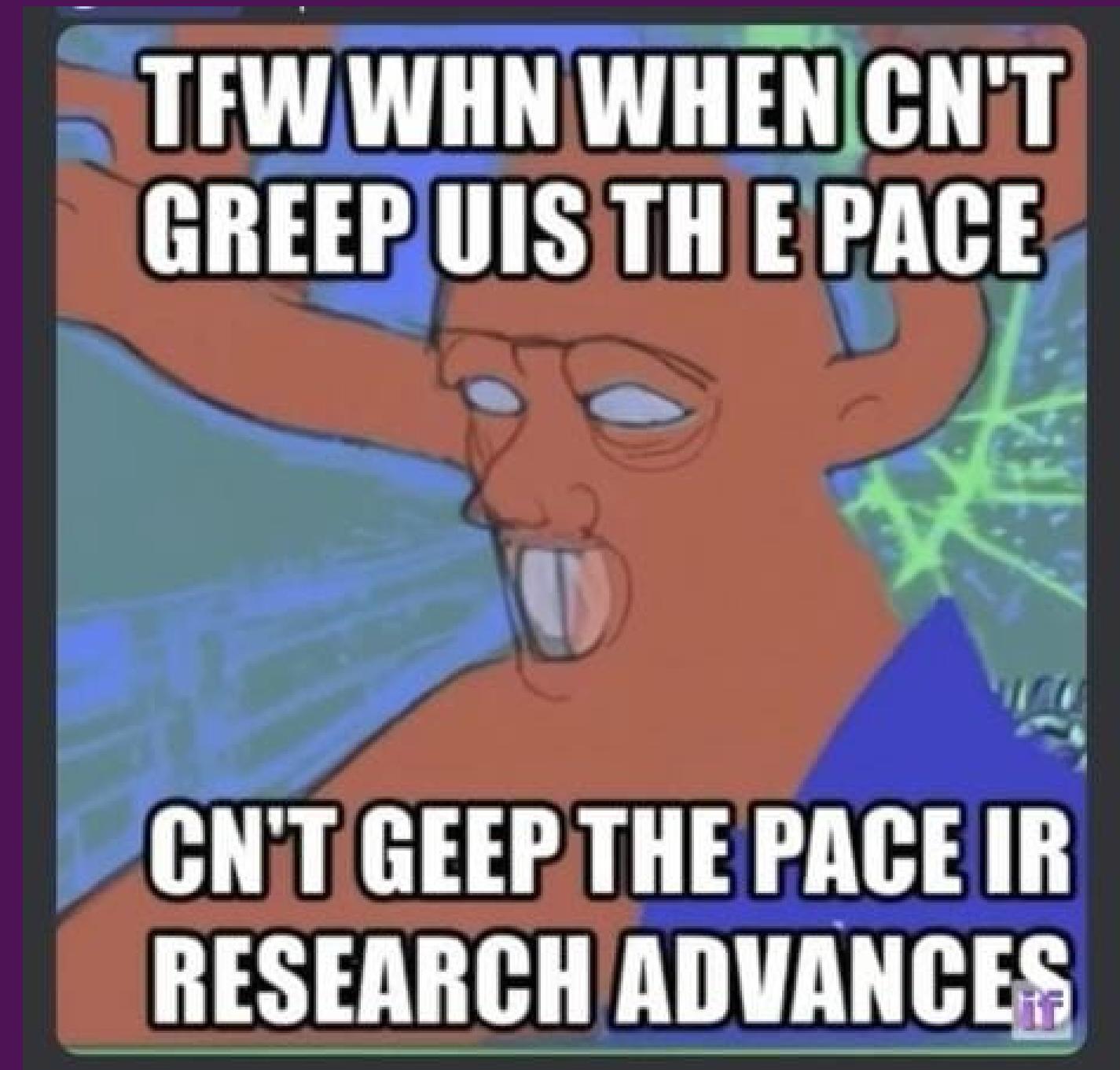


Programmazione

```
fetch_flights.py push_to_git.py d3_scale.js fetch_stock.js

1 const fetchRecentFlights = () => {
2   return fetch('https://www.flights-app.com/api/flights/recent', {
3     method: 'GET',
4     headers: {
5       'Content-Type': 'application/json',
6     },
7   })
8     .then(response => response.json())
9     .then(json => {
10       return json;
11     });
12 }
13 Copilot
```

Provate a capire questo meme.
(io non ce l'ho fatta)





**WTF WHEN CAN'T
KEEP WITH THE PACE**

**CAN'T KEEP THE PACE
IN RESEARCH ADVANCES**

***** QUANDO NON RIESCI
A STARE AL PASSO**

**NON RIESCI A STARE AL
PASSO CON GLI
AVANZAMENTI DI RICERCA**

GPT-4: Questo meme presenta un'immagine distorta di un personaggio, creando una faccia sciocca con un testo che è anche intenzionalmente distorto e difficile leggere. Ha lo scopo di trasmettere la sensazione di essere sopraffatti e incapaci di tenere il passo con i progressi della ricerca.

Il testo e l'immagine intenzionalmente distorti aggiungono un elemento umoristico al meme, riflettendo la lotta e la confusione che qualcuno potrebbe sentirsi quando cerca di tenere il passo con i rapidi progressi nel loro campo.



GPT-4: risultati

01

Interviste tecniche: ha passato un'intervista a Google con salario 183.000\$

02

Olimpiadi di matematica: è riuscito a produrre la dimostrazione corretta alle olimpiadi internazionali di matematica del 2022

03

Tuttofare: è riuscito a diagnosticare e risolvere una perdita in un bagno reale

04

Mappa mentale: ha creato una mappa mentale di una casa, risalendo a dove erano disposti vari oggetti precedentemente riposti.

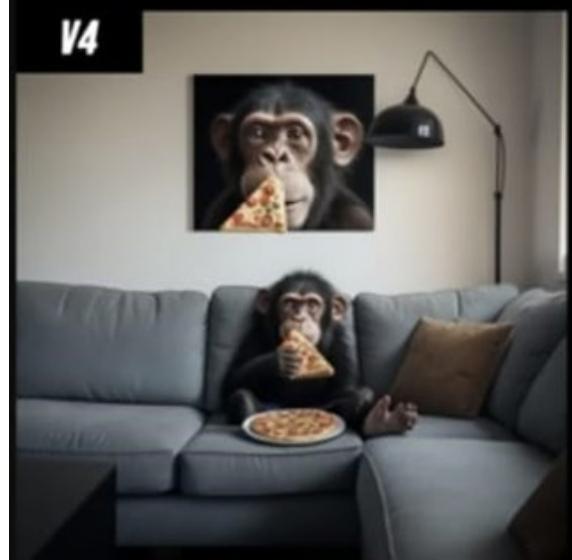
06

Progresso dei LLM

Evoluzione

Domanda: Quanto tempo è passato secondo voi da V1 a V5?

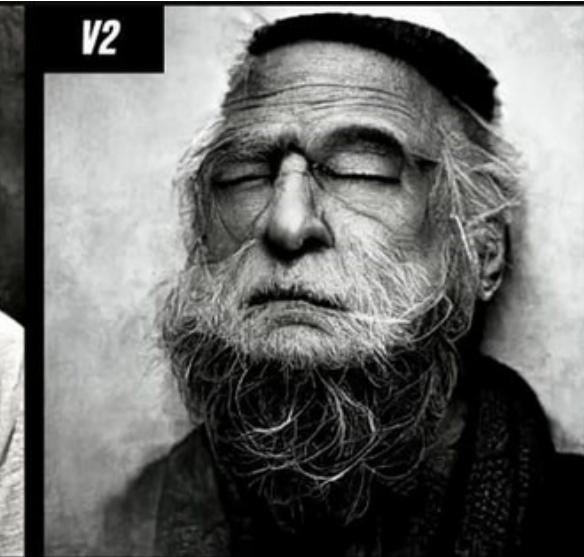
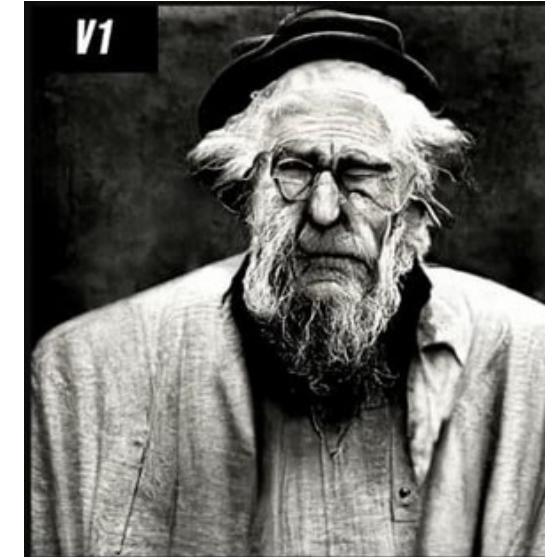
Scimmietta carina e adorabile seduta sul divano con in mano una fetta di pizza con i peperoni, sala, vibrante, fotografia, canon 5D, 50mm



PROMPT:

*PORTRAIT OF A CUTE AND
ADORABLE MONKEY SITTING
ON A COUCH
HOLDING A SLICE OF
PEPPERONI PIZZA, LIVING ROOM,
VIBRANT, PHOTOGRAPHY,
CANON 5D, 50MM*

Vecchio uomo stanco con una grande barba grigia, fotografia in bianco e nero



PROMPT:

*TIRED OLD MAN WITH
A LARGE GRAY BEARD,
BLACK AND WHITE PHOTOGRAPHY*

2 maggio



AI 'godfather' Geoffrey Hinton

Geoffrey Hinton è un pioniere nel campo dell'intelligenza artificiale (IA) e ha apportato significativi contributi allo sviluppo del deep learning. All'inizio degli anni '70 ha iniziato a lavorare sulle reti neurali artificiali e nel 1986 ha sviluppato una tecnica per una rete multilivello che ha trasformato il campo. Il lavoro di Hinton sul deep learning ha portato a significativi progressi in aree come il riconoscimento vocale, il riconoscimento delle immagini e l'elaborazione del linguaggio naturale.

"I left so that I could talk about the dangers of AI without considering how this impacts Google."

"Right now, they're not more intelligent than us, as far as I can tell. But I think they soon may be."

Google annuncia "Gemini"



Gemini è il modello di prossima generazione concepito per essere multimodale, ancora più efficiente e pronto per l'integrazione di utilizzo di API come tools, memoria e planning. Si trova ancora in fase di addestramento e sarà reso accessibile in futuro, una volta ritenuto pronto e sufficientemente affidabile.

Un **modello multimodale** è un modello che può processare e generare diversi tipi di dati, come testo, immagini, audio e video.

Un'**API (Application Programming Interface)** è un insieme di regole e specifiche che permette a diversi software di comunicare e scambiare dati tra loro.

11 maggio



Sam Altman richiede misure stringenti per l'AI

OpenAI CEO Sam Altman chiesto al Congresso di implementare nuove leggi per affrontare i seri rischi posti da i nuovi modelli di AI durante la sua apparenza a Capitol Hill.

Un **modello multimediale** è un modello che può processare e generare diversi tipi di dati, come testo, immagini, audio e video.

Un'**API (Application Programming Interface)** è un insieme di regole e specifiche che permette a diversi software di comunicare e scambiare dati tra loro.

07

Credits

Con l'aiuto di...



Agisci come un presentatore. Devo creare una presentazione di chatGPT, Transformers e in generale dei modelli generativi. La presentazione verrà proposta a degli studenti del quarto anno di laurea magistrale in psicologia e dovrà avere una durata di 30/40 minuti. La presentazione deve seguire la seguente struttura (ma se hai proposte per cambiarla esprimile pure):

- breve introduzione all'AI (differenza fra programmazione e AI, machine learning, deep learning, supervised e unsupervised, discriminative e generative)
- nel dettaglio: chatGPT, come funziona il meccanismo da semplice testo all'understanding della macchina;
- Forte focus su secondo quali modalità la macchina arriva a generare testo, rappresentazione delle parole (one-hot, context-free, words-in-context, tokenization);
- quali sono gli algoritmi arriva a scrivere quello che scrive (word embedding models uncontextualized e contextualized)
- creatività e logica che vengono create;
- aspetti psicologici: "Theory of mind", AGI eccetera;
- pros and cons;

Per ora commenta la scaletta e dammi feedback.



La tua scaletta sembra ben strutturata e tocca tutti i principali punti riguardo all'intelligenza artificiale e come funziona il modello ChatGPT. Ecco alcuni suggerimenti per affinare ulteriormente la tua presentazione:

1. **Introduzione all'AI:** Potrebbe essere utile iniziare con un esempio pratico di come l'AI



Given a youtube video link, you have to give me a summary of the video and the main points of it. I am currently creating a presentation about chatGPT and the transformers for some psychology students at the university of Padova. Here's the first link:
https://www.youtube.com/watch?v=f3o1MW2G5Rs&ab_channel=AIExplained



Used Video Insights



Used Video Insights

The video titled "Do We Get the \$100 Trillion AI Windfall? Sam Altman's Plans, Jobs & the Falling Cost of Intelligence" discusses Sam Altman's predictions about OpenAI's potential to capture a significant portion of the world's wealth and his plans for its distribution. Here are the main points:

1. **Sam Altman's Predictions:** Altman, the CEO of OpenAI, predicts that OpenAI could capture up to \$100 trillion of the world's wealth through the creation of AGI (Artificial General Intelligence). He believes that the price of many kinds of labor will fall towards zero once sufficiently powerful AI joins the workforce.

Domande?
Ponetele a <https://chat.openai.com/>
(oppure a me)

References

- **Lezioni di Natural Language Processing (prof Staiano J.)**
- **GPT-4 - Technical report**
- **Sparks of Artificial General Intelligence: Early experiments with GPT-4**
- **GPT-4**

