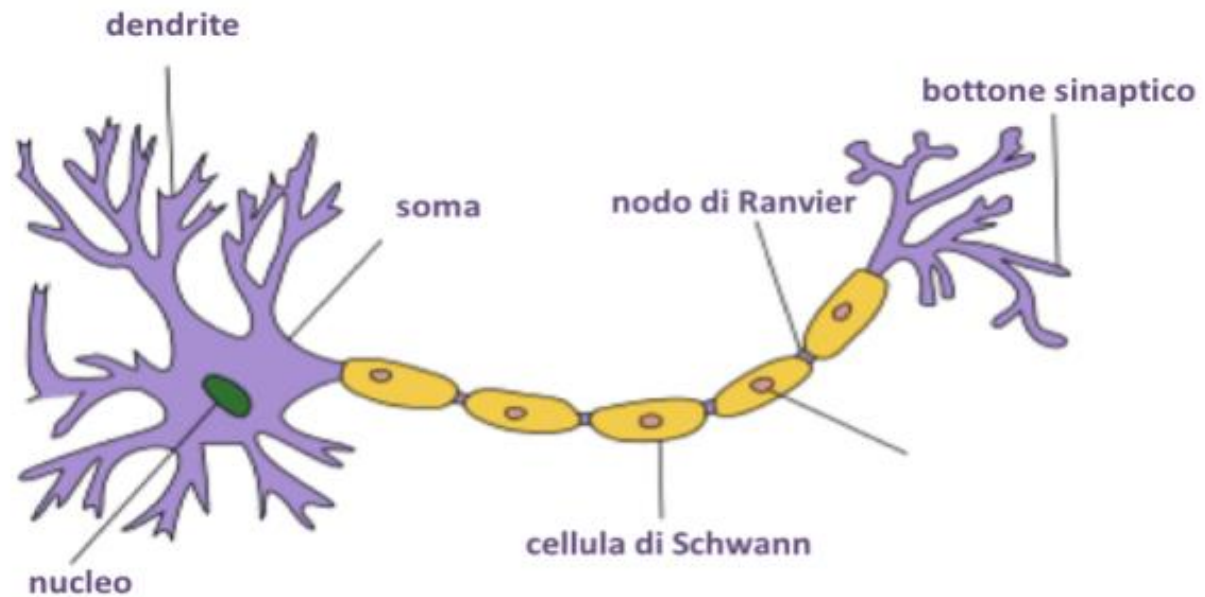


RETI NEURALI ARTIFICIALI NLP, LANGUAGE MODEL E GPT-3

N.I.L. 894855
J. F. S. 953198
Y.Y.L.979685

PRINCIPI E MODELLI DELLA
PERCEZIONE



COS'È UN NEURONE?

- unità fondamentali del cervello e del sistema nervoso, in grado di trasmettere impulsi di tipo elettrico, che permettono la trasmissione e l'elaborazione delle informazioni da una parte all'altra del corpo .

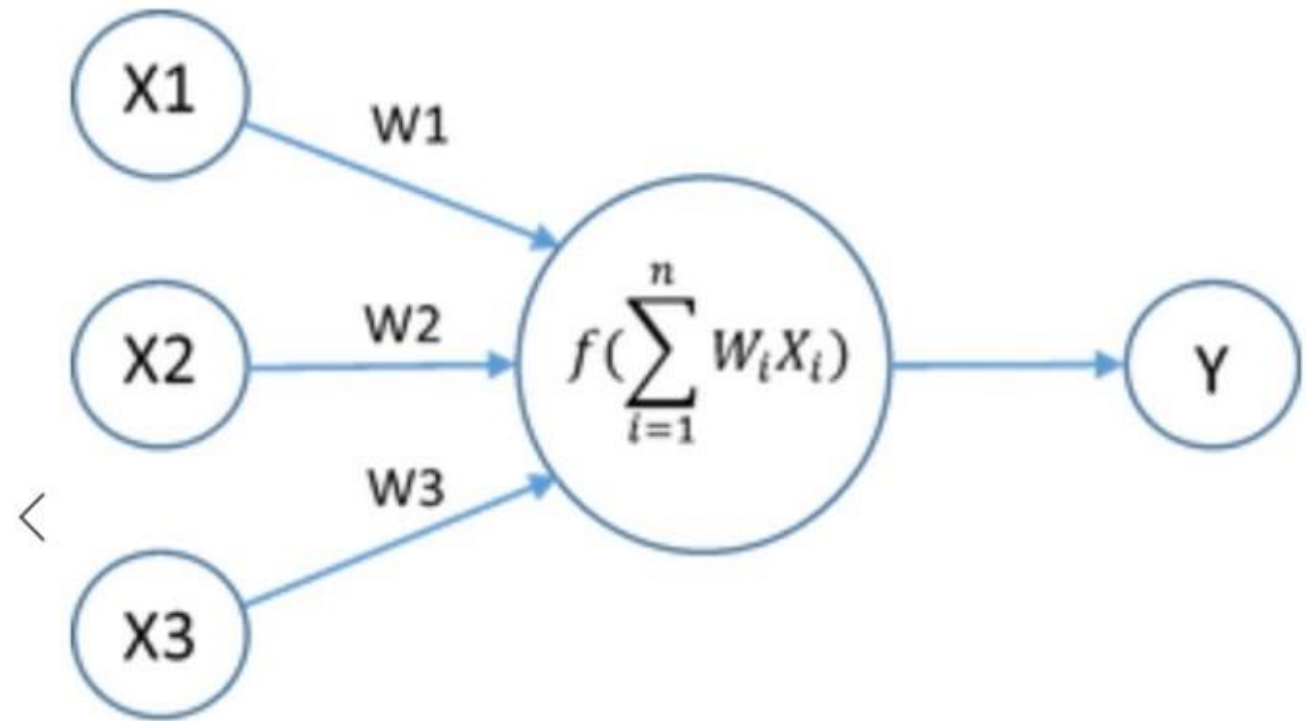


COS'È UNA RETE NEURALE?

- è il sistema di connessioni di neuroni e sinapsi che costituiscono il cervello degli esseri viventi.

NEURONE COME FUNZIONE MATEMATICA

- x_1, x_2, x_3 sono input e w_1, w_2, w_3 sono i loro pesi e li ranghiamo in un neurone e y è l'output dato dal neurone.



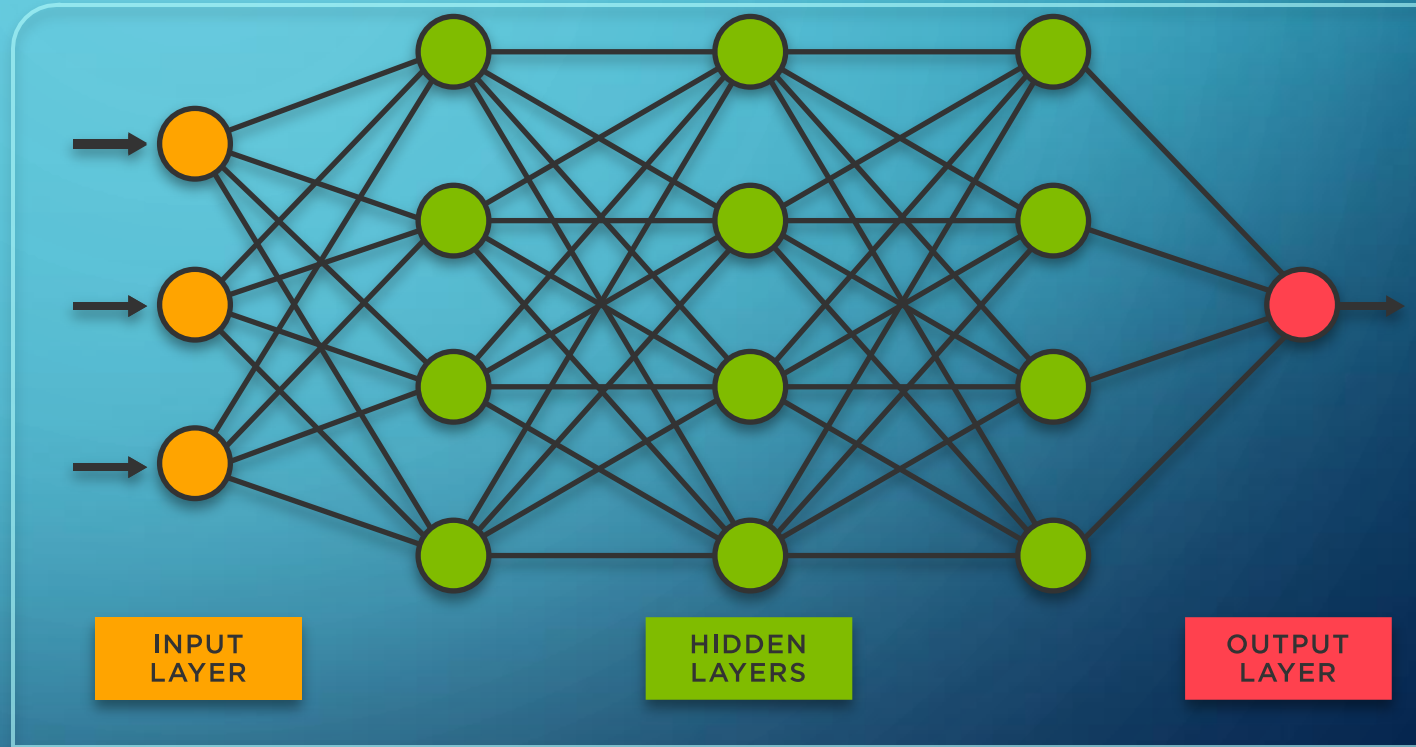


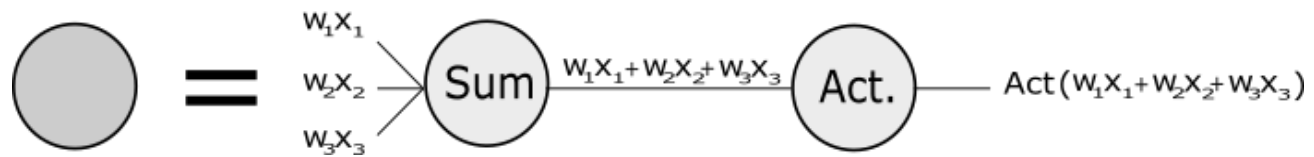
RETE NEURALE ARTIFICIALE

- un tipo di modello di intelligenza artificiale che si ispira sulla struttura e sul funzionamento delle reti neurali.
- Si basa su un insieme di algoritmi matematici e modelli computazionali che consentono di analizzare e riconoscere pattern complessi nei dati.

COME FUNZIONANO LE RETI NEURALI?

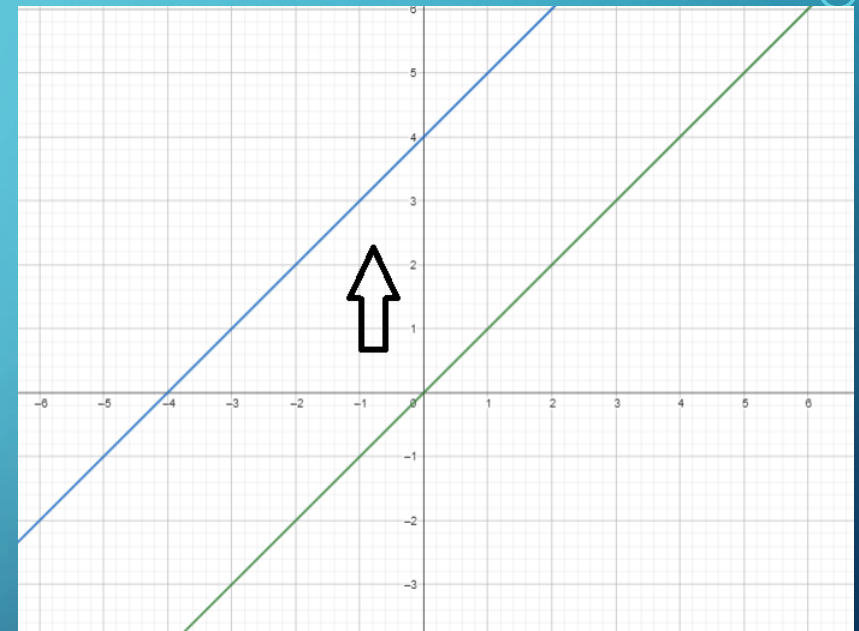
- Le reti neurali artificiali (ANN), sono composte da livelli di nodi che contengono un livello di input, uno o più livelli nascosti e un livello di output.





$$\sum_{i=1}^m w_i x_i + \text{bias} = w_1 x_1 + w_2 x_2 + w_3 x_3 + \text{bias}$$

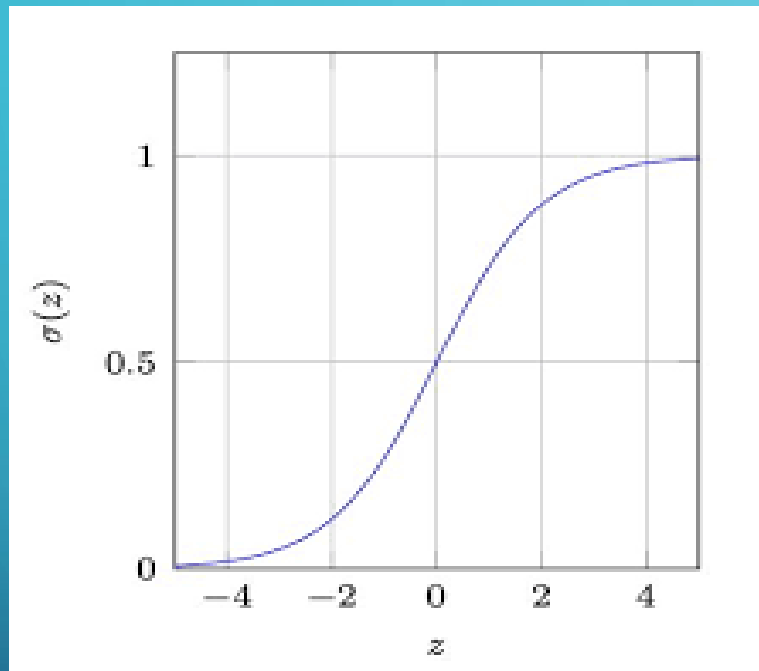
$$\text{output} = f(x) = \begin{cases} 1 & \text{if } \sum w_i x_i + b \geq 0 \\ 0 & \text{if } \sum w_i x_i + b < 0 \end{cases}$$



BIAS

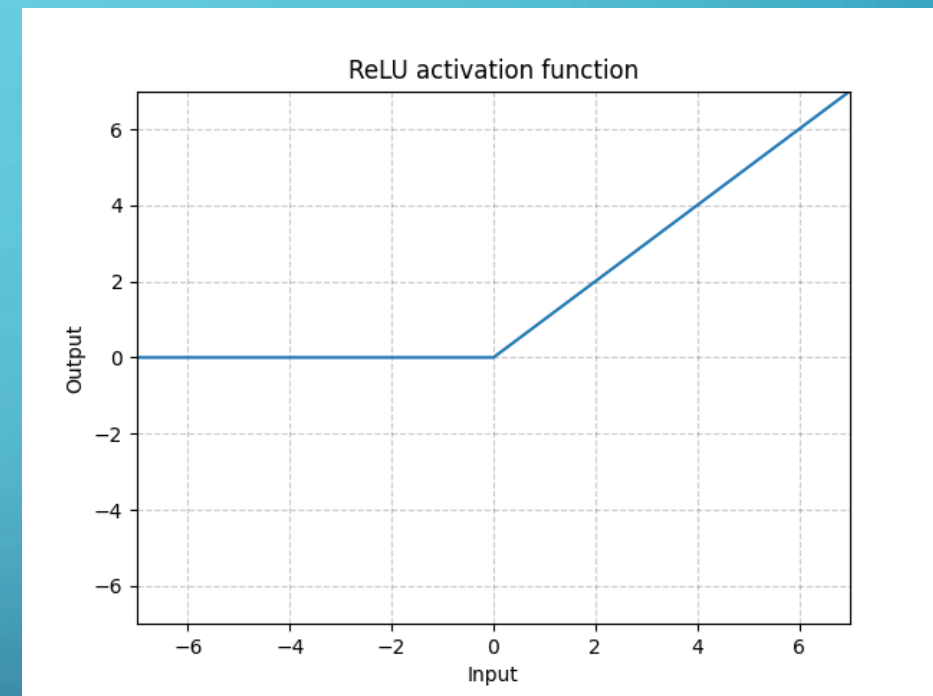
FUNZIONE DI ATTIVAZIONE

1. Funzione Sigmoidea



$$f(x) = \frac{1}{1 + e^{-x}}$$

2. Funzione ReLU



$$f(x) = \max(0, x)$$

LANGUAGE MODEL

Modello statistico in
NLP.

[breve introduzione di cos'è
il language modell]

Comprende e genera
testo in linguaggio
naturale.



L'IMPORTANZA DEI LANGUAGE MODEL

[qua parlo del perché è importante il language model]

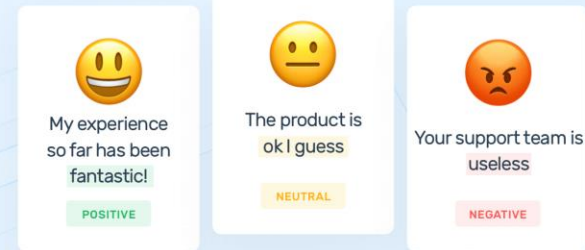


L'IMPORTANZA DEI LANGUAGE MODEL PT.2

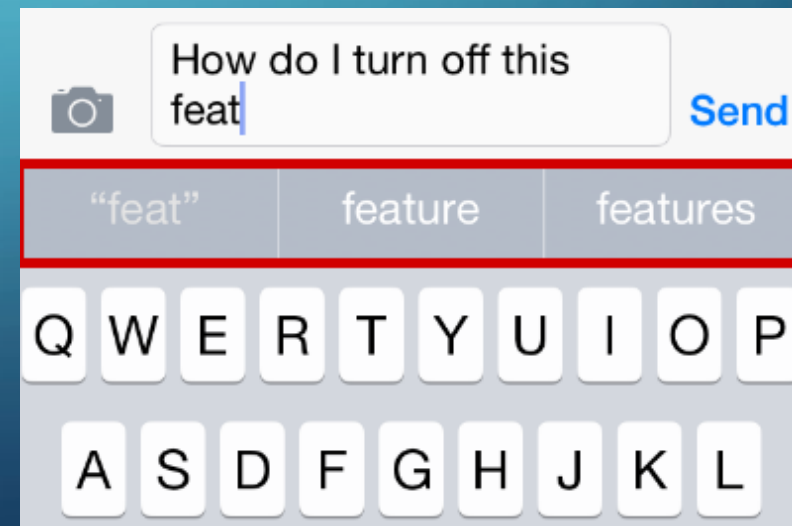
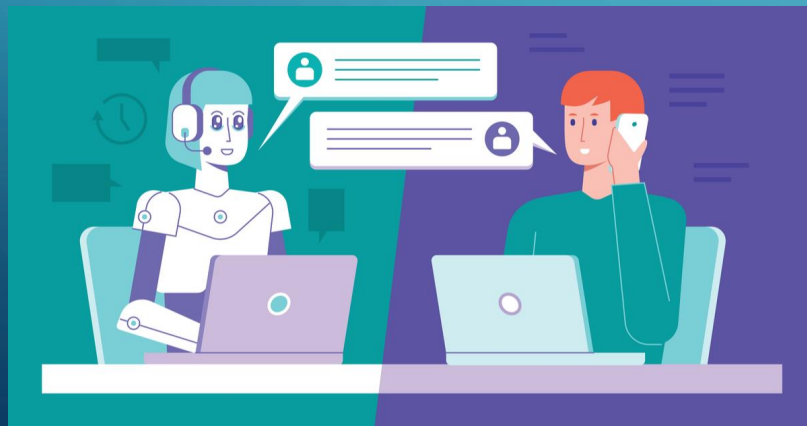
MACHINE TRANSLATION



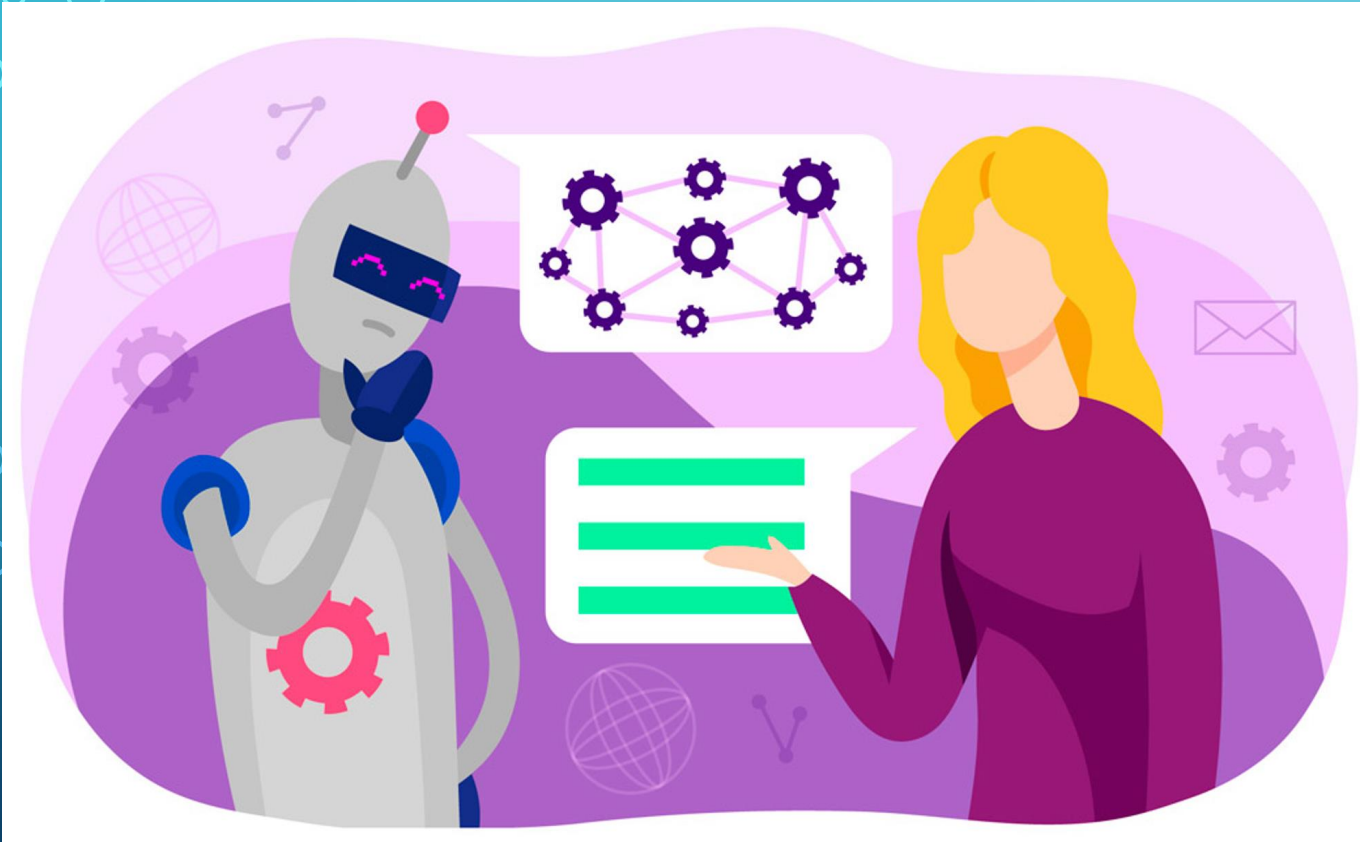
Sentiment Analysis



MonkeyLearn



"NON PARLIAMO LA STESSA LINGUA VERO?"



- Generare e comprendere testo in modo autonomo, come un essere umano.
- Non basta generare frasi che siano grammaticalmente corrette.
- Devono avere un senso.
- [qua parlo della differenza tra il linguaggio naturale e quello delle macchine e quindi ci sono differenza nella codifica]
- A machine needs a language model, which estimates the probabilities of sentences. If a language model is good, it will assign a larger probability to a correct option.

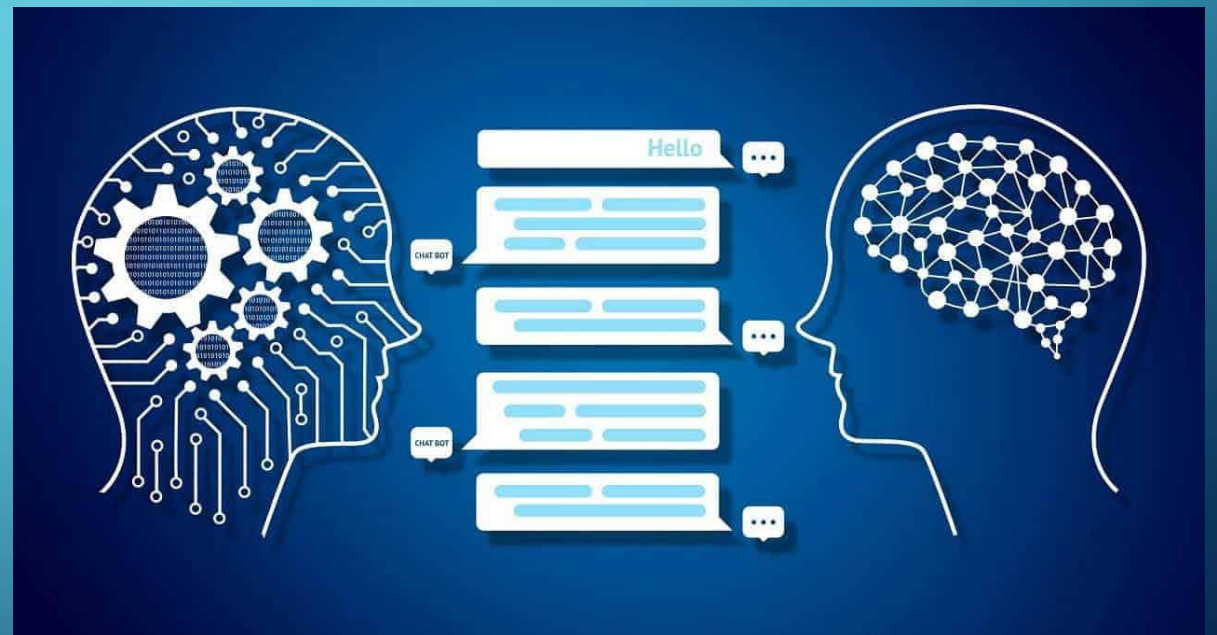
COME FUNZIONA?

- come fa riesce a predire la prossima parola?
- fase di training
- algoritmi probabilistici

TIPI DI MODEL LANGUAGE

- Modelli di linguaggio basati su:

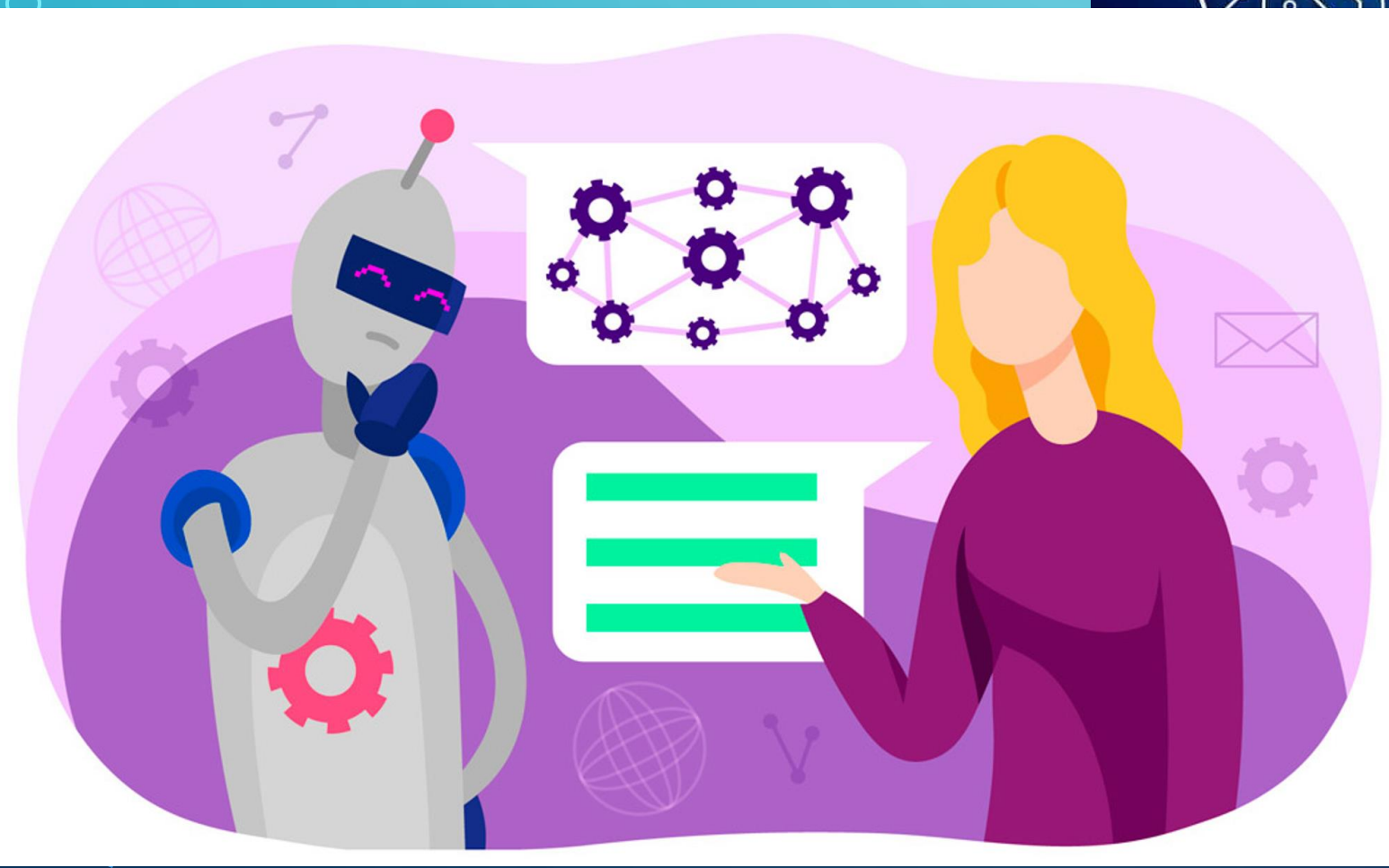
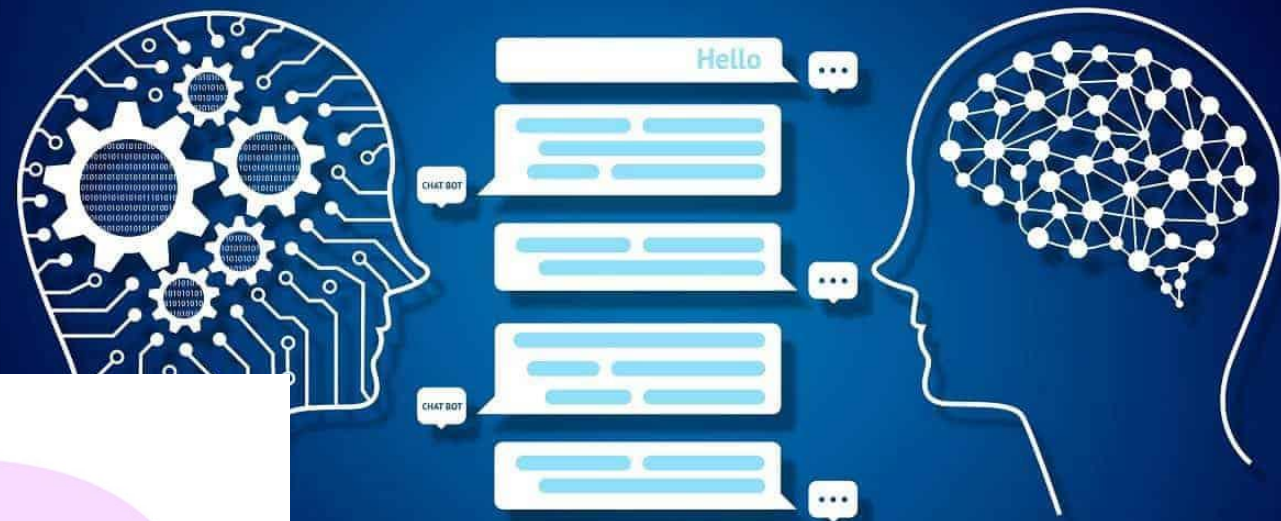
- Statistiche
- Reti neurali

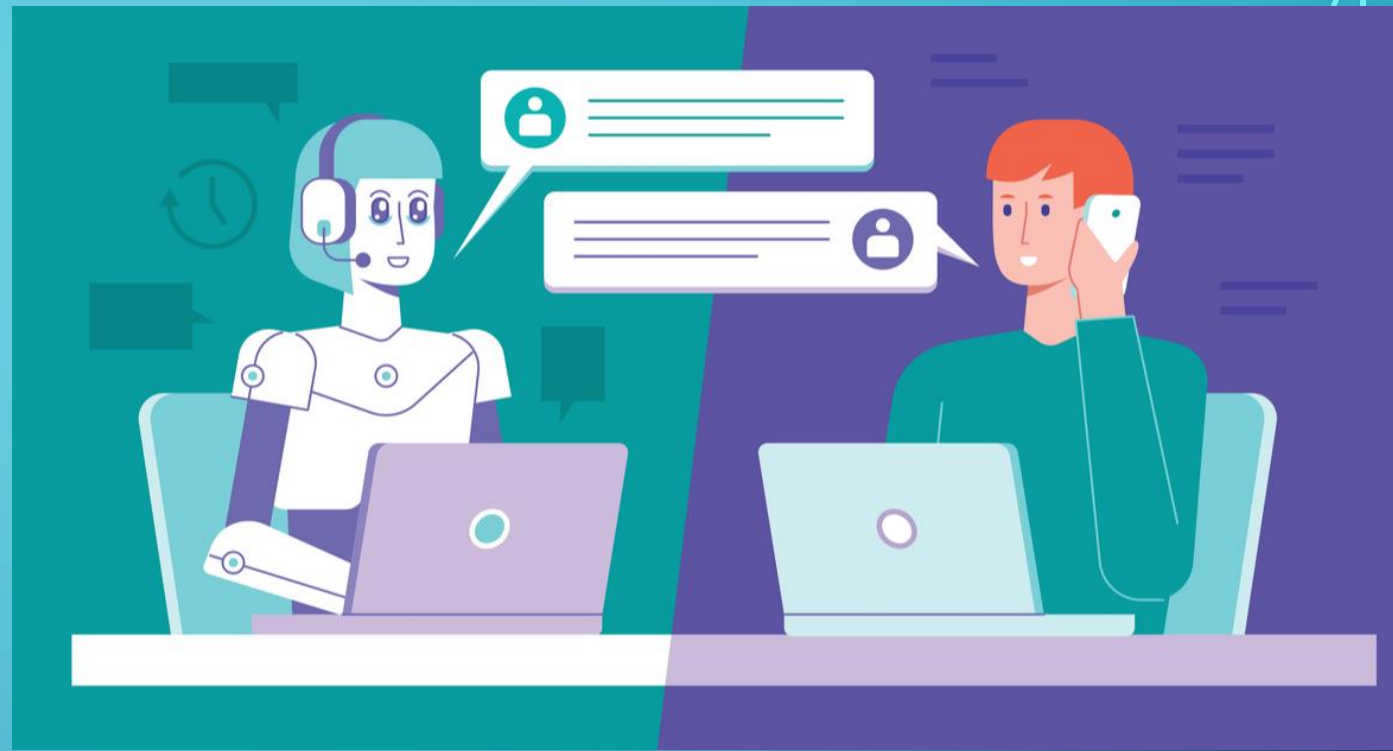


N-GRAM

- A machine needs a language model, which estimates the probabilities of sentences. If a language model is good, it will assign a larger probability to a correct option.

IMMAGINI CHE POTREBB





Can you please come **here** ?



History



Word being predicted

The background is a blue gradient with decorative white circuit-like lines in the corners. The lines consist of straight segments and small circles, resembling a stylized electronic circuit or neural network structure.

GPT-3

(Generative Pre-trained Transformer 3)

- Deep artificial neural network
- Large Language Model

Sviluppata da OpenAI (2020)



INPUT

Descrizione task
in linguaggio naturale

GPT-3



OUTPUT

Risultato sotto forma di un
qualsiasi tipo di testo
(testi altamente simili a quelli umani)

Svariati task:

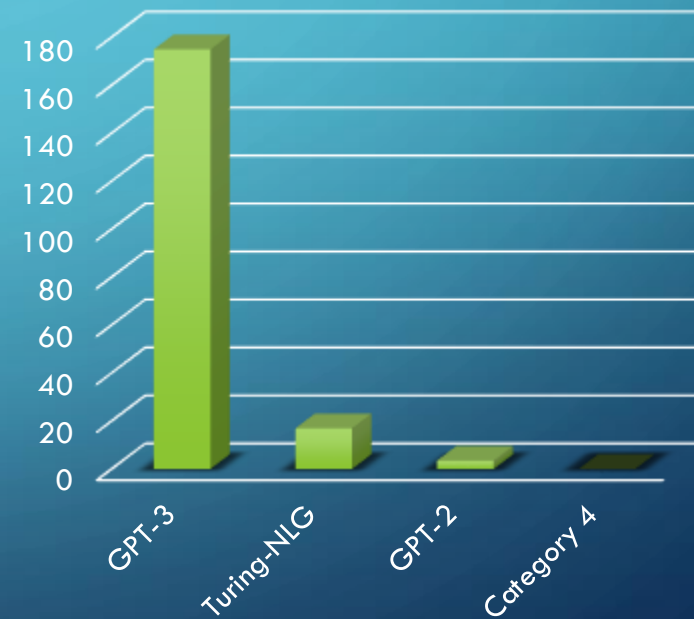
- Scrittura articoli, poesie, post social media
- Traduzione testi
- Generazione riassunti
- Chatbot
- Sentiment analysis
- Traduzione testo in codice (viceversa)
- Debugging



Caratteristiche

175 miliardi di parametri (pesi)

- permette alla rete neurale di produrre output **estremamente accurati**
- fase di training **molto onerosa**



45 TB di dati per il training

Dataset	Quantità (token)	Peso nel mix di training
Common Crawl	410 miliardi	60%
WebText2	19 miliardi	22%
Books1	12 miliardi	8%
Books2	55 miliardi	8%
Wikipedia	3 miliardi	3%

Common Crawl: archivio di dati di pagine web.

WebText2: testi di pagine web linkate in post di Reddit con almeno 3+ upvote.

Books1, Books2: raccolta di libri disponibili in rete

Pre-training

- Acquisizione **elementi generali di un linguaggio**:
 - Dizionario
 - Sintassi
 - Espressioni di senso figurato
 - Argomentazioni
- Training non specifico ad un particolare task
- Apprendimento **non supervisionato**



Language model in grado di esprimere scelte lessicali, pattern grammaticali e argomentazioni in modo **predittivo**

(slide su assenza fine-tuning)

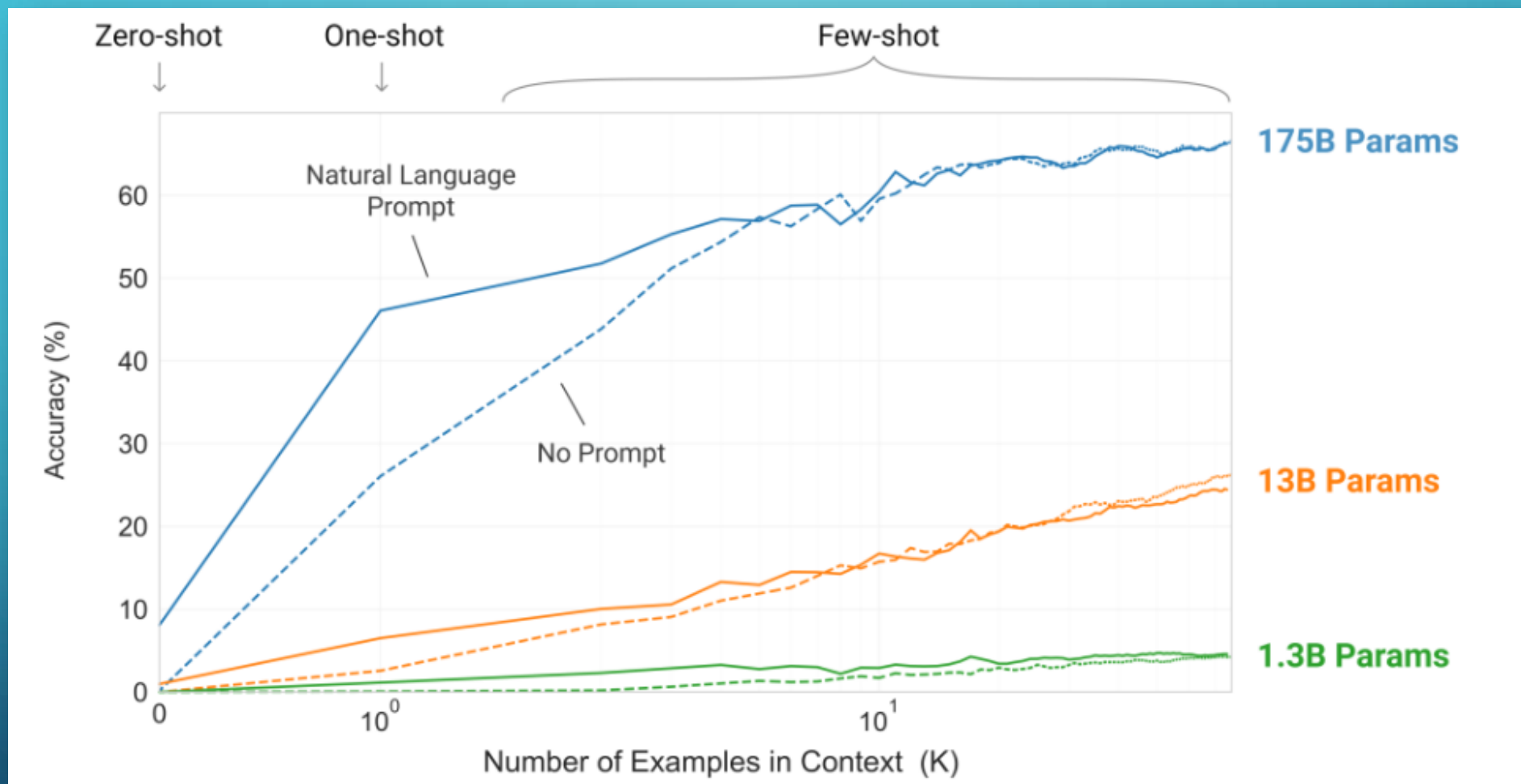
«Language models are Few-Shot Learners»

- Paradigma di apprendimento Few-Shot
- Maggiore capacità del language model ne aumenta l'abilità al **riconoscimento di pattern** in una **serie di esempi**.
- **All'aumentare del numero di esempi** forniti, il modello impara più dimostrazioni (contesti) riuscendo a migliorare l'accuratezza delle previsioni



(«Language models are Few-Shot Learners», 2020)

Performance sul task di rimuovere simboli estranei da una parola



(«Language models are Few-Shot Learners», 2020)

PRO

- Numerose applicazioni
- **Generalizzazione** dei task
- **No fine-tuning**

CONTRO

- **Non continua ad imparare** costantemente (pre-training)
- È un black box model
- Output può contenere **bias** (dipende dai dataset di training)

IMPLEMENTAZIONI

- ChatGPT
- Dall-E
- OpenAI Codex