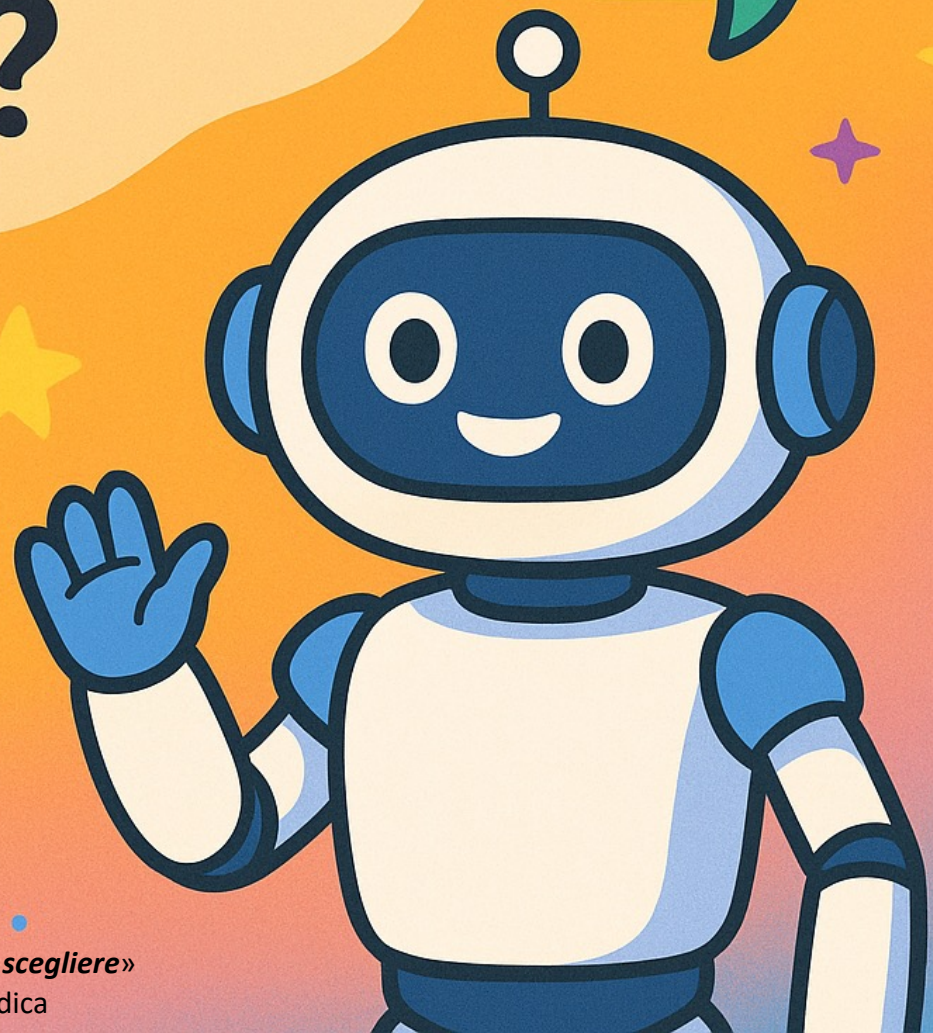


Come funziona ChatGPT?



Mini-lezione su
modelli linguistici,
potenzialità e limiti



Liceo Scientifico Statale «S. Cannizzaro» - Palermo

Anno Scolastico 2025/2026

Progetto di Educazione Civica

«*ChatGPT tra creatività, funzionamento ed etica: capire per scegliere*»

Docenti: Prof.ssa Roberta Accardi – Prof.re Erasmo Modica

Che cos'è un modello linguistico?

Immagina un programma capace di prevedere la parola **più probabile** per continuare una frase, basandosi su un'analisi di miliardi di testi già scritti e su cui è addestrato.

- **Come funziona?** Pensa al completamento automatico del tuo smartphone, ma con una potenza inimmaginabile.
- **Cosa non fa?** Non ragiona, non prova emozioni e non ha una vera comprensione o intenzioni. Si limita a imitare gli schemi linguistici che ha imparato.

Un esempio semplificato

Sul tuo smartphone scrivi:

"Domani vado al..."

Il suggerimento automatico potrebbe proporre:

"mare", "cinema", "lavoro", "scuola"

Un modello come ChatGPT fa lo stesso — ma su scala colossale e con una memoria contestuale molto più profonda.

Frase da completare

"I gatti amano dormire sul..."

Immagina che il modello debba decidere la **prossima parola**.

Ecco cosa fa.

Il modello ha "letto" miliardi di testi, e ha visto spesso frasi come:

- "I gatti amano dormire sul divano"
- "I gatti amano dormire sul letto"
- "I gatti amano dormire sul cuscino"

Quindi **calcola le probabilità** delle possibili parole successive.

Probabilità della parola successiva

"I gatti amano dormire sul..."

Parola successiva	Probabilità stimata
letto	41%
divano	28%
cuscino	15%
pavimento	6%
tetto	3%
banana	0,00001%

In base a come è configurato, il modello potrebbe:

- ☐ **Scegliere la più probabile** ("letto")
- ☐ **Mescolare un po' le carte** per essere più creativo (magari sceglie "divano")
- ☐ **Inserire casualità controllata** per generare risposte meno prevedibili



Ma ricordati:

Tutto questo avviene **senza che il modello "sappia" cosa sia un gatto, un letto o il sonno**. Sta solo **imitando lo stile e gli schemi** del linguaggio umano, basandosi su cosa ha "letto" più spesso.

La «palestra» di un'I.A.

ChatGPT impara attraverso un processo di addestramento continuo e supervisionato.

- 1) **Allenamento Iniziale:** Viene "nutrito" con un'enorme quantità di dati testuali, come libri, articoli e siti web, per apprendere le regole del linguaggio.
- 2) **Fine-Tuning:** Gli sviluppatori lo specializzano e correggono per affinare le sue capacità e renderlo più utile per compiti specifici.
- 3) **RLHF (Reinforcement Learning from Human Feedback):** Dei "docenti" umani lo guidano, valutando le sue risposte per "educarlo" a fornire risultati più sicuri e pertinenti.

Un esempio di RLHF

Domanda posta al modello: *Perché il cielo è blu?*

Il modello genera **due risposte possibili**:

 **Risposta A:** *"Il cielo appare blu a causa della rifrazione della luce nei gas atmosferici."*

 **Risposta B:** *"Il cielo appare blu a causa della **diffusione di Rayleigh**, un fenomeno fisico in cui le molecole dell'atmosfera diffondono la luce solare. La luce blu viene diffusa più di altri colori perché ha una lunghezza d'onda più corta."*



Feedback umano

Un esperto o annotatore umano guarda entrambe le risposte e valuta quale sia:

- ☐ **Più corretta**
- ☐ **Più chiara**
- ☐ **Più utile per l'utente**

In questo caso, vince la risposta B, perché fornisce il nome esatto del fenomeno (diffusione di Rayleigh), spiega meglio il meccanismo e usa un linguaggio preciso ma accessibile.



Cosa succede dopo

- ☐ Il modello riceve un **“premio”** (matematico) per aver generato la risposta B.
- ☐ Nei cicli successivi di addestramento, imparerà a **dare più peso a risposte simili a quella B**.
- ☐ Col tempo, questo aiuta il modello a **diventare più utile e preciso** nelle sue risposte.

Punti di forza di ChatGPT

- Un **assistente** sempre disponibile.
- **Sintesi imbattibile:** trasforma testi lunghi e complessi in riassunti chiari e concisi.
- **Riformulazione creativa:** spiega lo stesso concetto con parole e stili diversi, perfetto per studiare o trovare nuove prospettive.
- **Generatore di idee:** usalo come partner per sessioni di *brainstorming*, per scrivere racconti o per trovare ispirazione.

I limiti di ChatGPT



1. Allucinazioni



Fatti inventati o imprecisi

Può fornire informazioni sbagliate, citazioni false o numeri inesistenti. Verifica sempre le fonti.



2. Pregiudizi nascosti (bias)



Riflette stereotipi presenti nei dati

Come ogni strumento addestrato su testi umani, può riportare pregiudizi culturali, sociali o linguistici.



3. Sensibilità al prompt



Dipende da come chiedi

La qualità della risposta cambia molto se riformuli la domanda. Spesso servono più tentativi.



4. Mancanza di autenticità



Imita, non crea davvero

Non ha idee proprie, ma rielabora ciò che ha “visto” nei testi. Nessuna creatività o pensiero originale.

Come scrivere un buon prompt

 Scopo	Cosa vuoi ottenere?	"Fammi un riassunto"
 Pubblico	A chi è destinata la risposta?	"Spiegalo a un ragazzo di 15 anni"
 Formato	Come deve essere scritta?	"In un elenco puntato"
 Lunghezza	Quanto deve essere lunga?	"Massimo 10 righe"
 Stile	Che tono deve usare?	"Stile universitario, formale"
 Fonti	Vuoi dati verificabili?	"Cita almeno 2 fonti attendibili"
 Passaggi	Vuoi un processo guidato?	"Spiegamelo passo per passo"

✅ Esempio completo di prompt ben formulato:

"Fammi un riassunto in 10 righe, con tono divulgativo, spiegato per un quindicenne, in elenco puntato, e cita 2 fonti attendibili."

Uno strumento, non un sostituto

 **ChatGPT è un alleato, non una mente alternativa.**

Un assistente potente che può **amplificare le tue capacità**, aiutarti a capire, riassumere, esplorare.

Ma non potrà **mai sostituire il pensiero critico**, la creatività vera, il giudizio umano.

 **La vera sfida non è artificiale, ma umana.**

Saper **collaborare con l'IA** vuol dire:

- ✓ Conoscerne i limiti
- ✓ Usarla con consapevolezza
- ✓ Farne uno **strumento di apprendimento**, non di dipendenza

 **La vera intelligenza sta nel fare le domande giuste.**