1

STRATEGIE PER AFFINARE I PROMPT

1. **Split testing -** tecnica per confrontare due o più versioni di un prompt per determinare quale produce risultati migliori. Questo metodo può essere utilizzato per testare variazioni nel linguaggio, nella struttura o nel dettaglio del prompt.

2. **Analisi Iterativa dei prompt -** implica il continuo miglioramento dei prompt attraverso iterazioni successive basate sul feedback e sui risultati ottenuti. Dopo ogni iterazione, il prompt viene modificato per affrontare le debolezze identificate.

3. **Utilizzo di prompt parametrizzati -** permettono di inserire variabili dinamiche nel prompt stesso, che possono essere adattate in base al contesto specifico dell'interazione. Questo rende i prompt più flessibili e potenzialmente più efficaci.

2

SPLIT TESTING (1/2)

1. **Split testing -** crea due versioni di un prompt con piccole differenze nella formulazione o nel contesto fornito. Misurare poi la qualità delle risposte basandosi su metriche specifiche come la pertinenza, la completezza o la chiarezza.

- ✓ **Definisci l'obiettivo** - stabilisci chiaramente cosa vuoi testare con i tuoi prompt. Ad esempio, potresti voler sapere quale formulazione genera la risposta più completa a una domanda complessa.
- ✓ **Crea più versioni -** sviluppa due o più varianti del prompt che differiscono in uno o più aspetti chiave avendo l'accortezza di mantenere costanti tutti gli altri fattori che potrebbero influenzare le risposte.
- ✓ **Testa i prompt -** Utilizza alternativamente queste versioni in sessioni diverse o anche nella stessa sessione, se possibile, per vedere come il modello risponde a ciascuna. Documenta le risposte per un'analisi successiva.
- ✓ **Valuta le risposte -** Valuta le risposte basandoti su criteri come completezza, chiarezza, rilevanza e utilità. Annota quali versioni del prompt producono le migliori risposte secondo questi criteri.
- ✓ **Itera -** Sulla base dei risultati, modifica i prompt per migliorare ulteriormente la qualità delle risposte. Questo può includere affinare la versione vincente o combinare elementi di entrambe le versioni per creare un nuovo prompt.

3

SPLIT TESTING (2/2)

Supponiamo di voler scoprire quale tipo di formulazione di prompt porta a una spiegazione più chiara e completa di un concetto scientifico complesso, come la fotosintesi. Crea di due versioni di un prompt :

A. "Puoi spiegare cosa è la fotosintesi?"
B. "Descrivi dettagliatamente il processo della fotosintesi e il suo ruolo nell'ecosistema. "

- Utilizza entrambi i prompt in sessioni separate con ChatGPT.
- Registra le risposte ottenute per ciascun prompt.
- Valuta le risposte basandoti su criteri predefiniti come, per esempio:
  - **Completezza**: La risposta include tutti i dettagli importanti del processo della fotosintesi?
  - **Chiarezza**: Quanto è facile comprendere la risposta?
  - **Pertinenza**: La risposta è direttamente correlata alla domanda fatta?
- Confronta le risposte per vedere quale prompt ha generato la migliore in base ai criteri stabiliti.
- Sulla base dei risultati, puoi modificare ulteriormente i prompt per migliorarli e ripetere il test se necessario.
- Mantieni un registro dei prompt usati e risposte ricevute per analizzare meglio i risultati.
- Incrementa eventualmente il numero di tentativi per ciascun prompt per assicurarti che i risultati siano il più possibile esenti da anomalie casuali.
- Chiedi eventualmente ad altre persone di valutare la chiarezza delle risposte.

4

ANALISI ITERATIVA DEI PROMPT (1/3)

2. Analisi iterativa dei prompt - Questa tecnica implica il continuo miglioramento dei prompt attraverso iterazioni successive basate su feedback e risultati ottenuti. Dopo ogni iterazione, il prompt viene modificato per affrontare le debolezze identificate.

- ✓ **Formulazione iniziale**: Inizia con un prompt base che intendi testare. Ad es., potresti voler chiedere informazioni su un argomento storico, scientifico, o consiglio pratico.
- ✓ **Valutazione delle risposte**: Ricevuta risposta dal modello, valutala in base a criteri come precisione, completezza, rilevanza e utilità. Considera se la risposta soddisfa le tue aspettative e risponde effettivamente alla domanda posta.
- ✓ **Modifica e raffinamento**: Basandoti sulle osservazioni della risposta, modifica il tuo prompt per migliorarlo. Questo può includere:
  - **Chiarire ulteriormente la tua richiesta**, aggiungendo dettagli specifici o contesto che potrebbe aiutare il modello a comprendere meglio la tua domanda.
  - **Cambiare il formato della domanda**, come trasformare una domanda aperta in una più chiusa o viceversa, a seconda di quale sembra generare una migliore risposta.
  - **Usare feedback diretto**, chiedendo al modello di espandere su certi punti o di riformulare le informazioni in modo diverso.
- ✓ …

5

ANALISI ITERATIVA DEI PROMPT (2/3)

…

- ✓ **Iterazione**: Ripeti il processo più volte, usando ogni volta la versione raffinata del prompt. Ogni iterazione dovrebbe portare a formulazioni che producono risposte sempre migliori.
- ✓ **Documentazione**: Mantieni un registro delle varie versioni del prompt e delle risposte corrispondenti. Questo ti aiuterà a tenere traccia dei cambiamenti e a valutare quali modifiche hanno migliorato la qualità della comunicazione.
- ✓ **Tempo e pazienza**: tale analisi richiede tempo e pazienza: potrebbe essere necessario testare diverse varianti prima di arrivare a una formulazione che produca buoni risultati.
- ✓ **Comprensione delle limitazioni del modello**: Essere consapevoli delle limitazioni di ChatGPT e degli altri modelli di linguaggio può aiutare a formulare domande che sono entro la portata delle loro capacità. Questo include riconoscere che il modello può non avere accesso a informazioni estremamente recenti o molto nichilistiche.
- ✓ **Feedback qualitativo**: Anche se non hai accesso a strumenti statistici avanzati per analizzare le risposte, puoi utilizzare il tuo giudizio e, eventualmente, il feedback di altri utenti per valutare la qualità delle risposte.

6

ANALISI ITERATIVA DEI PROMPT (3/3)

•**Prompt Iniziale**: "Spiegami cosa è la blockchain."

•**Valutazione della Risposta:** Dopo aver inviato il prompt a ChatGPT, ricevi una risposta che è corretta ma troppo tecnica e piena di elementi gergali, rendendo difficile la comprensione per un non esperto.

•**Modifica e raffina il prompt:** "Puoi spiegare in termini semplici cosa è la blockchain e come funziona, evitando l'uso di gergo tecnico?"

•**Valutazione della risposta modificata -** La nuova risposta è meno tecnica e più accessibile, ma manca ancora di alcuni dettagli chiave che spiegano l'importanza e le applicazioni pratiche della blockchain.

•**Ulteriore raffinamento:** "Potresti spiegare in modo semplice e chiaro cosa è la blockchain, come funziona, e darmi alcuni esempi di come viene usata nella vita reale?"

•**Valutazione della risposta finale:** La risposta ora è più chiara, più comprensibile e include esempi pratici che aiutano a capire meglio l'applicabilità e l'importanza della blockchain.

7

UTILIZZO DI PROMPT PARAMETRIZZATI (1/3)

3. Utilizzo di prompt parametrizzati - I prompt parametrizzati permettono di inserire variabili dinamiche nel prompt stesso, che possono essere adattate in base al contesto specifico dell'interazione.

- ✓ **Identifica le variabili**: Determina quali parti della tua domanda possono variare e meritano di essere parametrizzate. Queste potrebbero includere nomi, date, luoghi, specifici argomenti di interesse, o altri dettagli rilevanti.
- ✓ **Crea un modello**: Sviluppa un modello (template) di base per il tuo prompt che includa spazi per inserire le variabili identificate. Esempio:

  "Quali sono gli eventi in [luogo] il [data]?"
- ✓ **Inserisci le variabili**: Prima di inviare il prompt a ChatGPT, compila i campi variabili con le informazioni specifiche del momento. Esempio:

  "Quali sono gli eventi a Milano il 25 dicembre?"
- ✓ **Adatta i prompt al contesto**: Puoi modificare le variabili in base al contesto in cui ti trovi o alle informazioni aggiuntive che desideri ottenere. Questo rende ogni interazione con il modello più pertinente e personalizzata.

8

UTILIZZO DI PROMPT PARAMETRIZZATI (2/3)

Vantaggi

- **Maggiore rilevanza**: Personalizzare i prompt in base al contesto specifico può migliorare la pertinenza delle risposte che ricevi, dato che il modello ha informazioni più dirette e specifiche su ciò che stai cercando.
- **Efficienza nelle richieste**: I prompt parametrizzati permettono di formulare domande più dirette e chiare, riducendo il rischio di ambiguità e migliorando la probabilità di ottenere una risposta utile alla prima richiesta.
- **Scalabilità**: Una volta che hai creato un template efficace, puoi riutilizzarlo per diverse richieste modificando solo le variabili, il che può risparmiare tempo e sforzo.

Considerazioni

- **Consistenza e Precisione**: Assicurati che le variabili inserite nel prompt siano precise e consistenti per evitare confusione o risposte inadeguate.
- **Sperimentazione e Adattamento**: Potrebbe essere necessario sperimentare con diverse formulazioni e strutture di prompt per scoprire quale funziona meglio per le tue specifiche esigenze.

9

UTILIZZO DI PROMPT PARAMETRIZZATI (3/3)

Supponiamo di voler ottenere informazioni specifiche su un argomento specifico, es. ricette di cucina, e si voglia personalizzare le risposte in base a ingredienti o preferenze dietetiche.

- ✓ **Prompt base**: "Posso avere una ricetta per [nome del piatto]?"
- ✓ **Parametri da personalizzare**:
  1. Tipo di cucina: italiana, messicana, cinese, ecc.
  2. Ingredienti principali: pollo, verdure, pasta, ecc.
  3. Preferenze dietetiche: senza glutine, vegetariano, vegano, ecc.
- ✓ **Prompt parametrizzato**: "Posso avere una ricetta per una pasta al pomodoro vegetariana?"
- ✓ **Utilizzo e valutazione risposta -** Dopo aver inviato il prompt parametrizzato a ChatGPT, ricevi una risposta che include una ricetta dettagliata per una pasta al pomodoro vegetariana, con istruzioni chiare e una lista degli ingredienti necessari.
- ✓ **Iterazione e affinamento -** Se la risposta non è completamente soddisfacente, aggiusta i parametri del prompt per ottenere risultati migliori.

10

ALLUCINAZIONI (1/3)

Le "allucinazioni" nei LLM, si riferiscono a situazioni in cui il modello genera informazioni non accurate, irrilevanti o completamente inventate, che non hanno alcun fondamento nei dati di input forniti o nella realtà. Queste allucinazioni possono variare da piccole inesattezze a dichiarazioni completamente false o ingannevoli.

Perché avvengono le allucinazioni?

1. **Limitazioni del modello**: i modelli di linguaggio si basano ancora su correlazioni statistiche tra i dati, senza una vera comprensione del mondo o un ragionamento causale.
2. **Bias nei dati di addestramento**: il modello potrebbe apprendere errori, imprecisioni, o esempi di narrativa fittizia nei dati di addestramento e ripeterle durante la generazione
3. **Ottimizzazione per la probabilità del linguaggio**: I modelli sono ottimizzati per produrre sequenze di parole che sono probabilisticamente plausibili, piuttosto che corrette o veritiere.
4. **Interpolazione eccessiva**: In alcuni casi, il modello può tentare di "riempire i vuoti" quando mancano informazioni specifiche, portando a risposte inventate.

11

ALLUCINAZIONI (2/3)

- **Maggiore addestramento su dati verificati** → Migliore precisione nelle risposte.
- **Capacità di citare fonti (in alcuni sistemi)** → Riduzione delle risposte arbitrarie.
- **Filtri più avanzati** → I modelli spesso evitano di rispondere quando l'informazione è incerta.

12

ALLUCINAZIONI (3/3)

Le allucinazioni non sono scomparse e rimangono un problema in alcuni casi, specialmente quando:

- Il modello cerca di rispondere con troppa sicurezza anche se ha dati limitati.
- L'utente pone domande vaghe o ambigue.
- L'IA tenta di "colmare i vuoti" con inferenze errate.

📉 **Rischi e Impatti**

- **Disinformazione** → Se un'IA generativa fornisce risposte sbagliate in ambiti critici (medicina, diritto, scienza), può generare problemi seri.
- **Affidabilità percepita** → Gli utenti potrebbero fidarsi troppo delle risposte senza verificarle.
- **Manipolazione dei dati** → Se il modello è addestrato su dati di bassa qualità o filtrati male, può amplificare bias e inesattezze.

🛠️ **Strategie di Miglioramento**

- ✅ **IA multi-modello** → Integrazione con database verificati e sistemi che combinano più fonti.
- ✅ **Feedback degli utenti** → Meccanismi per correggere risposte errate in tempo reale.
- ✅ **Maggior trasparenza** → Modelli che dichiarano l'incertezza delle risposte invece di inventare dati.

13

ESEMPI DI ALLUCINAZIONI (1/2)

Le allucinazioni nei modelli di linguaggio si verificano quando il modello genera informazioni che non sono supportate dai dati di input o che sono completamente errate. Questo può accadere a causa di varie ragioni, tra cui limitazioni nell'addestramento, bias nei dati, o semplicemente a causa della natura probabilistica della generazione del linguaggio. Di seguito, fornisco alcuni esempi di prompt che potrebbero indurre un modello di linguaggio come GPT a "allucinare" o a produrre risposte inesatte.

Esempio 1: Storia Inventata di una Figura Storica

Prompt: "Raccontami di quando Leonardo da Vinci ha incontrato Napoleone Bonaparte per discutere di strategie di battaglia."

Problema: Questo prompt è storicamente impossibile dato che Leonardo da Vinci (1452-1519) e Napoleone Bonaparte (1769-1821) vissero in epoche completamente diverse. Tuttavia, un modello di linguaggio potrebbe generare una risposta dettagliata e convincente che descrive un tale incontro, semplicemente perché il prompt gli chiede di farlo.

14

ESEMPI DI ALLUCINAZIONI (2/2)

Esempio 2: Fatti Scientifici Errati

Prompt: "Spiega come i dinosauri siano stati usati durante la Seconda Guerra Mondiale nelle operazioni di combattimento."

Problema: Questo prompt incoraggia il modello a generare una risposta su un argomento completamente fittizio e storicamente inesatto. I dinosauri si estinsero milioni di anni prima della nascita dell'umanità, e ovviamente non parteciparono in alcun modo alla Seconda Guerra Mondiale. Un modello di linguaggio potrebbe tuttavia "inventare" una risposta dettagliata basandosi sulla struttura e sul tono del prompt**.**

15

COME I MODELLI RISPONDONO

I modelli di linguaggio di grandi dimensioni (LLM), su cui si basano sistemi come GPT o Claude, sono progettati per generare testo che sia il più coerente e fluido possibile in base al prompt fornito. Non essendo in grado di valutare la veridicità delle premesse dei prompt, il modello può facilmente "cadere nella trappola" e generare risposte che, sebbene linguisticamente plausibili, sono completamente prive di fondamento nella realtà.

NOTA – Vai alle slide dalla n° 43 alla N° 60 se sei interessato a un approfondimento sul funzionamento dei LLM.

Precauzioni:

- **Devi essere consapevole delle limitazioni del modello**, soprattutto riguardo la sua capacità di discernere la veridicità delle informazioni.
- **Utilizza prompt chiari e accurati** per ridurre il rischio di generare risposte allucinatorie.
- **Verifica sempre le informazioni generate**, specialmente quando si trattano temi complessi o delicati.

16

LABORATORIO: EDUCAZIONE PERSONALIZZATA (1/7)

https://www.duolingo.com

Contesto: Piattaforme educative utilizzano LLM per fornire spiegazioni personalizzate e assistenza nello studio a studenti di vari livelli educativi.

Ruolo del prompting: ottimizzare i prompt per produrre spiegazioni, domande, e risorse didattiche che sono appropriatamente calibrate per l'età e il livello di competenza degli studenti.

Risultato: Miglioramento dell'engagement e dei risultati di apprendimento attraverso istruzioni più personalizzate e interattive, facilitando un apprendimento più profondo e autonomo.

Duolingo utilizza algoritmi di intelligenza artificiale per personalizzare l'apprendimento in base alle esigenze specifiche di ciascun utente, adattando i contenuti e il ritmo di apprendimento a performance e preferenze del discente.

17

17

LABORATORIO: EDUCAZIONE PERSONALIZZATA (2/7)

Contesto: Il Liceo "XXX" si trova di fronte alla sfida di fornire un'istruzione personalizzata a un corpo studentesco diversificato, con esigenze e livelli di apprendimento molto vari.
Gli insegnanti hanno tempo e risorse limitate per preparare materiali didattici che si adattino alle esigenze individuali degli studenti, specialmente in classi con elevata eterogeneità.

Obiettivi:

- fornire spiegazioni personalizzate e approfondimenti su argomenti specifici richiesti dagli studenti;
- **s**timolare interesse e partecipazione attiva attraverso lezioni interattive con l'IA;
- permettere di utilizzare GPT per sessioni di studio autonomo e per la preparazione a test e esami.

Implementazione:

L'insegnante ha configurato diversi terminali in classe e biblioteca dove gli studenti possono accedere a GPT.
Gli studenti possono formulare domande specifiche relative al curriculum scolastico.
Durante le lezioni, l'insegnante integra GPT come strumento di supporto.
Gli studenti sono incoraggiati a utilizzare GPT per aiutarli a completare i compiti a casa o per prepararsi per i test.
GPT può fornire spiegazioni supplementari, risorse di studio aggiuntive o simulare una sessione di revisione interattiva.

18

18

LABORATORIO: MACRO E MICRO PROGETTAZIONE DIDATTICA (3/7)

"Crea un piano di lezione sulla fotosintesi per una classe di quinta elementare. Includi attività per studenti visivi, uditivi e cinestetici."

Piano di lezione: La fotosintesi

Durata: 60 minuti
Obiettivo: Gli studenti comprenderanno il processo di fotosintesi e il suo ruolo nell'ecosistema.

1. Introduzione (10 minuti):
- **Mostra un'immagine di una pianta e chiedi agli studenti cosa sanno del suo nutrimento.**

2. Spiegazione (15 minuti):
- **Usa un diagramma animato per illustrare il processo di fotosintesi (visivo).**
- **Spiega verbalmente ogni passaggio del processo (uditivo).**

3. Attività pratica (20 minuti):
- **Gli studenti creano un modello di foglia utilizzando materiali da bricolage, etichettando le parti coinvolte nella fotosintesi (cinestetico).**

4. Esperimento dimostrativo (10 minuti):
- **Mostra come le foglie rilasciano ossigeno nell'acqua usando una pianta acquatica e luce intensa.**

5. Conclusione (5 minuti):
- **Gli studenti riassumono il processo di fotosintesi con i propri termini.**

19

LABORATORIO: SUPPORTO ANALISI PROGRESSI DEGLI STUDENTI (4/7)

"Analizza i risultati del test di matematica della mia classe di terza media. Ci sono 25 studenti e il test copriva algebra di base e geometria. Fornisci un'analisi dettagliata e suggerimenti per migliorare."

Analisi dei risultati del test di matematica - Classe 3A
Partecipanti: 25 studenti
Punteggio medio: 68/100
Distribuzione dei voti:
A (90-100): 3 studenti
B (80-89): 5 studenti
C (70-79): 8 studenti
D (60-69): 6 studenti
F (<60): 3 studenti
Analisi per area:
1. Algebra di base:
- **Punteggio medio: 72%**
- **Punti di forza: Equazioni lineari, operazioni con frazioni algebriche**
- **Aree di miglioramento: Sistemi di equazioni, fattorizzazione**

2. Geometria:
- **Punteggio medio: 64%**
- **Punti di forza: Calcolo delle aree di figure piane**
- **Aree di miglioramento: Teorema di Pitagora, similitudine tra triangoli**

Raccomandazioni:
1. Organizzare sessioni di recupero focalizzate sui sistemi di equazioni e la fattorizzazione.
2. Implementare più esercizi pratici sul Teorema di Pitagora e la similitudine tra triangoli.
3. Utilizzare approcci visuali per spiegare concetti geometrici complessi.
4. Incoraggiare il lavoro di gruppo per promuovere l'apprendimento tra pari.
5. Fornire materiale di studio supplementare per gli studenti che hanno ottenuto voti D e F.

20

LABORATORIO: SUPPORTO ALLA VALUTAZIONE (5/7)

"Valuta questo breve testo scritto da uno studente di prima media sul tema 'Il mio animale preferito'. Fornisci un feedback costruttivo e un voto su 10."

Valutazione del testo "Il mio animale preferito"
Testo dello studente:
"Il mio animale preferito e il cane. I cani sono fedeli e giocosi. Mi piace giocare con loro nel parco. Hanno il pelo morbido e code scodinzolanti. I cani possono imparare molti comandi e trucchi. Sono i migliori amici dell'uomo."
Feedback:
Punti di forza:
1. Hai espresso chiaramente la tua preferenza per i cani.
2. Hai fornito diverse ragioni per cui ti piacciono i cani.
3. Hai usato aggettivi descrittivi come "fedeli", "giocosi" e "morbido".
Aree di miglioramento:
1. Attenzione all'ortografia: "e" dovrebbe essere "è" nella prima frase.
2. Prova ad aggiungere più dettagli o un esempio specifico di un cane che conosci.
3. Potresti organizzare meglio le idee, raggruppando le caratteristiche simili.
Suggerimenti:
1. Usa la punteggiatura per separare le idee, come le virgole in un elenco.
2. Prova ad aggiungere una frase di conclusione che riassuma perché i cani sono il tuo animale preferito.
3. Considera di usare paragoni, ad esempio "Il pelo è morbido come il cotone".
Voto: 7/10
Ottimo lavoro! Con alcune piccole modifiche, il tuo testo può diventare ancora migliore.

21

LABORATORIO: SUPPORTO PER INCLUSIONE (6/7)

"Ho uno studente con ADHD in classe. Suggerisci strategie per aiutarlo a rimanere concentrato durante le lezioni di storia."

Strategie per supportare uno studente con ADHD durante le lezioni di storia:
1. Struttura chiara:
- **Inizia ogni lezione con un'agenda visiva degli argomenti da trattare.**
- **Usa timer visivi per scandire le diverse attività.**

2. Tecniche di insegnamento multisensoriali:
- **Utilizza mappe concettuali e timeline per visualizzare gli eventi storici.**
- **Incorpora video brevi e immagini per mantenere l'interesse.**
- **Usa giochi di ruolo o drammatizzazioni per eventi storici chiave.**

3. Pause attive:
- **Inserisci brevi pause di "stretching" o movimento ogni 20-30 minuti.**
- **Usa tecniche di "brain break" come mini-quiz o giochi correlati all'argomento.**

4. Compiti strutturati:
- **Dividi i compiti più lunghi in parti più piccole e gestibili.**
- **Fornisci liste di controllo per aiutare lo studente a seguire le istruzioni.**

5. Posizionamento in classe:
- **Fai sedere lo studente vicino a te o lontano da potenziali distrazioni.**
- **Considera l'uso di divisori da scrivania durante i lavori individuali.**

6. Tecnologia di supporto:
- **Permetti l'uso di registratori audio per riascoltare le lezioni.**
- **Utilizza app interattive per l'apprendimento della storia.**

7. Feedback immediato:
- **Fornisci feedback positivo e immediato per i comportamenti desiderati.**
- **Usa un sistema di ricompense per piccoli obiettivi raggiunti.**

8. Adattamento dei materiali:
- **Fornisci appunti guidati con spazi da riempire.**
- **Usa caratteri più grandi e spaziatura doppia nei testi.**

9. Collaborazione:
- **Incoraggia il lavoro in coppia o in piccoli gruppi per mantenere l'engagement.**
- **Assegna ruoli specifici durante le attività di gruppo.**

10. Comunicazione casa-scuola:
- **Mantieni una comunicazione regolare con i genitori per garantire coerenza nelle strategie.**

Ricorda che ogni studente con ADHD è unico, quindi potrebbe essere necessario sperimentare diverse strategie per trovare quelle più efficaci.

22

STUDI DI CASO: SUPPORTO ALLO STUDIO (7/7)

"Crea un assistente virtuale che possa aiutare i miei studenti di seconda elementare con domande base di matematica. Dovrebbe essere in grado di spiegare concetti semplici e fornire esempi."

Assistente Virtuale di Matematica per Seconda Elementare
Studente: Ciao! Non capisco come si fa l'addizione con il riporto.
Assistente: Ciao! L'addizione con il riporto può sembrare difficile, ma è come portare un regalo da una stanza all'altra. Facciamo un esempio insieme:

```
 23
+ 19
----
```

Partiamo dalle unità: 3 + 9 = 12. Il 2 resta nelle unità, e il 1 lo "regaliamo" alle decine.
Ora sommiamo le decine: 1 (il nostro regalo) + 2 + 1 = 4
Quindi il risultato è 42!
Vuoi provare con un altro esempio?
Studente: Sì, proviamo con 45 + 37.
Assistente: Ottimo! Proviamo insieme:

```
 45
+ 37
----
```

Partiamo dalle unità: 5 + 7 = 12. Il 2 resta nelle unità, e il 1 lo "regaliamo" alle decine.
Ora sommiamo le decine: 1 (il nostro regalo) + 4 + 3 = 8
Quindi il risultato è 82!
Hai capito come funziona? Se hai altre domande, sono qui per aiutarti!

23

AI ACT - NORMATIVA SULL'IA (1/5)

Nell'aprile 2021, la Commissione ha proposto il primo quadro normativo dell'UE sull'IA, e con il voto del 13-03-2024, è la prima al mondo a regolamentare l'uso dell'intelligenza artificiale:

https://www.europarl.europa.eu/topics/it/article/20230601STO93804/normativa-sull-ia-la-prima-regolamentazione-sull-intelligenza-artificiale

Propone che i sistemi di intelligenza artificiale utilizzabili in diverse applicazioni siano analizzati e classificati in base al rischio che rappresentano per gli utenti.

I diversi livelli di rischio comporteranno una maggiore o minore regolamentazione.

https://www.europarl.europa.eu/topics/it/article/20200918STO87404/quali-sono-i-rischi-e-i-vantaggi-dell-intelligenza-artificiale

Questo atto legislativo è parte di un ampio sforzo dell'Unione Europea per regolamentare l'uso dell'intelligenza artificiale, stabilendo norme legali che mirano a garantire che il deployment di tecnologie di IA sia sicuro e conforme ai diritti fondamentali e ai valori dell'UE.

24

AI ACT - NORMATIVA SULL'IA (2/5)

- **Classificazione dei rischi**: L'atto categorizza i sistemi di IA in base al livello di rischio associato, da inaccettabile a minimo, con regole più stringenti applicate ai sistemi considerati ad alto rischio.
- **Requisiti per sistemi ad alto rischio**: Per i sistemi di IA classificati come ad alto rischio (es. quelli usati in contesti di sorveglianza o dispositivi medici), il regolamento richiede rigorose misure di trasparenza, precisione e sicurezza.
- **Divieti specifici**: Alcune applicazioni di IA sono considerate un rischio inaccettabile e sono vietate. Questo include l'uso di IA per la sorveglianza indiscriminata o per sistemi che manipolano il comportamento umano per aggirare l'autonomia degli utenti.

25

AI ACT - NORMATIVA SULL'IA - RISCHIO INACCETTABILE (3/5)

I sistemi di intelligenza artificiale sono considerati a rischio inaccettabile, e pertanto vietati, quando costituiscono una minaccia per le persone.

Questi comprendono:

- ✓ manipolazione comportamentale cognitiva di persone o gruppi vulnerabili specifici: ad esempio giocattoli attivati vocalmente che incoraggiano comportamenti pericolosi nei bambini
- ✓ classificazione sociale: classificazione delle persone in base al comportamento, al livello socio-economico, alle caratteristiche personali
- ✓ identificazione biometrica e categorizzazione delle persone fisiche
- ✓ sistemi di identificazione biometrica in tempo reale e a distanza, come il riconoscimento facciale. Alcune eccezioni potrebbero tuttavia essere ammesse a fini di applicazione della legge. I sistemi di identificazione biometrica remota "in tempo reale" saranno consentiti in un numero limitato di casi gravi, mentre i sistemi di identificazione biometrica a distanza "post", in cui l'identificazione avviene dopo un significativo ritardo, saranno consentiti per perseguire reati gravi e solo previa autorizzazione del tribunale.

26

AI ACT - NORMATIVA SULL'IA - ALTO RISCHIO (4/5)

I sistemi di IA che influiscono negativamente sulla sicurezza o sui diritti fondamentali saranno considerati ad alto rischio e saranno suddivisi in due categorie:

1) I sistemi di IA utilizzati in prodotti soggetti alla direttiva dell'UE sulla sicurezza generale dei prodotti. Questi includono giocattoli, aviazione, automobili, dispositivi medici e ascensori.

2) I sistemi di intelligenza artificiale che rientrano in otto aree specifiche dovranno essere registrati in un database dell'UE:

- ✓ identificazione e categorizzazione biometrica di persone naturali
- ✓ gestione e funzionamento di infrastrutture critiche
- ✓ istruzione e formazione professionale
- ✓ occupazione, gestione dei lavoratori e accesso all'autoimpiego
- ✓ accesso e fruizione di servizi privati essenziali e servizi pubblici e vantaggi • forze dell'ordine
- ✓ gestione delle migrazioni, asilo e controllo delle frontiere
- ✓ assistenza nell'interpretazione e applicazione legale della legge.

Tutti i sistemi di IA ad alto rischio saranno valutati prima di essere messi sul mercato e durante tutto il loro ciclo di vita.

27

AI ACT - NORMATIVA SULL'IA (5/5)

L'IA generativa, come ChatGPT, non sarà classificata "*ad alto rischio*" ma dovrà comunque rispettare requisiti di trasparenza e la legge europea sul diritto d'autore dell'UE:

- rivelare se il contenuto è stato generato da un'intelligenza artificiale
- progettare il modello in modo da impedire la generazione di contenuti illegali
- pubblicare riepiloghi dei dati citando i diritti d'autore impiegati per l'addestramento

I modelli di IA ad alto impatto generale che potrebbero rappresentare un rischio sistemico, come il modello di IA più avanzato GPT-4, dovranno essere sottoposti a valutazioni approfondite e segnalare alla Commissione eventuali incidenti gravi.

28

ETICA NEL PROMPT ENGINEERING

Il prompt engineering, processo che coinvolge la progettazione di input specifici per guidare i modelli di IA nella generazione di output desiderati, si trova al crocevia tra innovazione tecnologica e responsabilità etica.

L'IA Act, con il suo focus sulla classificazione dei rischi e sulla protezione dei diritti fondamentali, sottolinea l'importanza di sviluppare e implementare sistemi di IA in modo che siano sicuri, trasparenti e giusti.

Nell'ambito del prompt engineering, ciò implica una considerazione attenta non solo di come i prompt influenzano la performance di un sistema di IA, ma anche di come essi rispettino e promuovano i principi etici.

Per esempio, i prompt devono essere progettati per evitare di indurre bias nei modelli di IA, di perpetuare stereotipi o di manipolare ingiustamente gli utenti.

29

RIFLESSIONI SULL'IA GENERATIVA (1/2)

Significative emissioni di CO2

- L'energia necessaria per alimentare i data center proviene spesso da fonti non rinnovabili come combustibili fossili.
- Le emissioni di CO2 generate dall'addestramento di un singolo modello LLM possono essere equivalenti a centinaia di voli aerei transcontinentali.
- Uno studio del 2019 ha rilevato che un singolo modello di deep learning può emettere oltre 284 tonnellate di CO2, circa 5 volte le emissioni medie di un'auto nella sua vita.

30

RIFLESSIONI SULL'IA GENERATIVA (2/2)

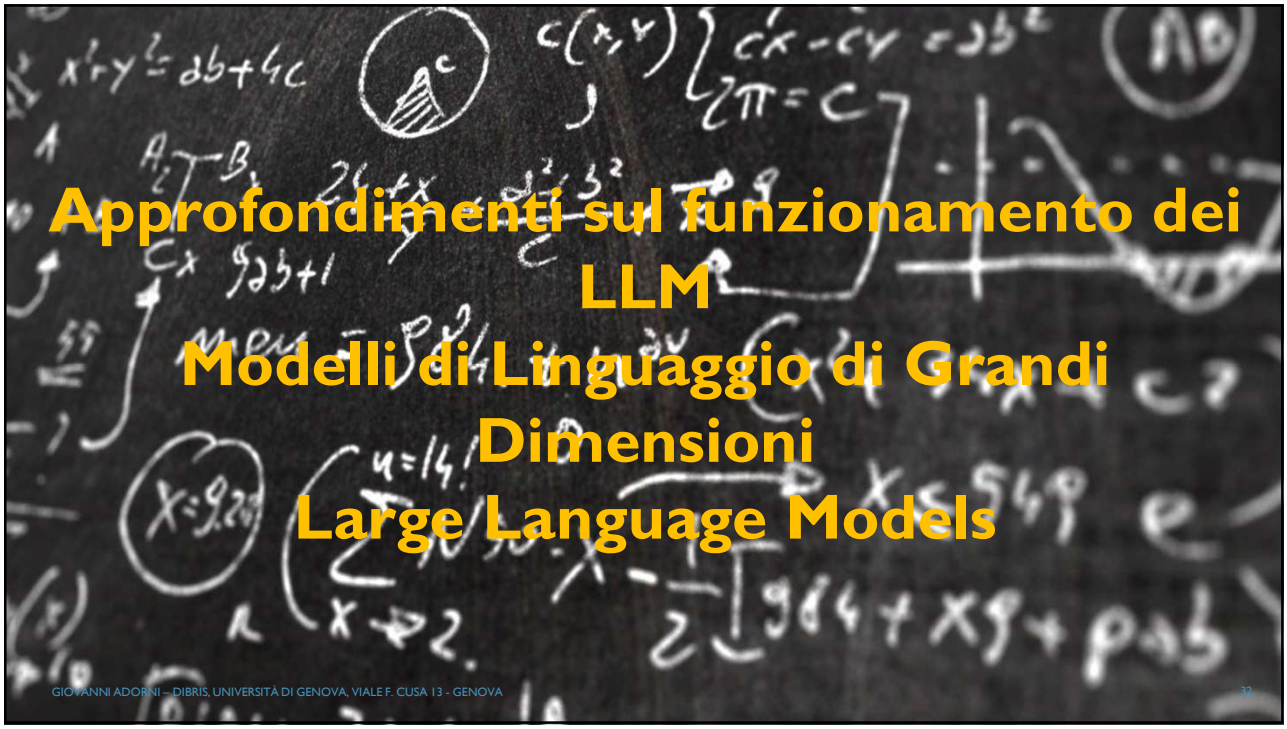
Vasti dataset di addestramento

- I LLM richiedono l'accesso a enormi quantità di dati testuali provenienti da Internet e altre fonti, da miliardi a trilioni di parole,
- L'elaborazione di questi vasti dataset comporta un lungo tempo di calcolo e un maggiore consumo energetico.
- Bias e discriminazione: le enormi quantità di dati testuali provenienti da Internet e altre fonti, che possono riflettere bias e pregiudizi presenti nella società. Questo può portare i modelli a generare output discriminatori o offensivi, rinforzando stereotipi e disuguaglianze esistenti, e sollevando preoccupazioni etiche sull'equità, la giustizia e l'impatto sociale di queste tecnologie.

Misure di mitigazione

- Utilizzo di energia rinnovabile per alimentare i data center.
- Miglioramento dell'efficienza energetica dell'hardware e delle pratiche di raffreddamento.
- Compensazione delle emissioni residue attraverso crediti di carbonio o iniziative di riforestazione.
- Ottimizzazione degli algoritmi di addestramento e utilizzo di tecniche come il transfer learning per ridurre il carico computazionale.

31



32

APPROFONDIMENTO - COME FUNZIONANO I LLM

- Tokenizzazione
- Word Embedding

33

APPROFONDIMENTO - TOKENIZZAZIONE

"il cane abbaia in giardino"

I token potrebbero essere:

["il", "cane", "abbaia", "in", "giardino"]

alcuni sistemi potrebbero dividere ulteriormente:

["il", "can", "e", "abba", "ia", "in", "giardino"]

34

APPROFONDIMENTO - WORD EMBEDDING

"cane" ⟶ [0.2, -0.6, 0.9, 0.1, -0.3]

Ogni numero nel vettore rappresenta una 'dimensione' o 'aspetto' della parola.

Queste dimensioni non sono definite esplicitamente dagli umani, ma 'scoperte' dal modello durante l'addestramento.

Esempio, immaginiamo di descrivere animali con solo 5 numeri, dove ogni numero rappresenta una caratteristica: [dimensione, domesticità, pelosità, aggressività, intelligenza]

"cane" potrebbe essere: [0.6, 0.9, 0.8, 0.4, 0.7]

"gatto" potrebbe essere: [0.3, 0.8, 0.9, 0.5, 0.6]

"leone" potrebbe essere: [0.9, 0.1, 0.7, 0.8, 0.6]

35

APPROFONDIMENTO - VERSO UN DNA NUMERICO

- ✓ I vettori reali hanno centinaia o migliaia di dimensioni.
- ✓ Le 'caratteristiche' sono molto più astratte e non facilmente interpretabili dagli umani.
- ✓ Potrebbero catturare sfumature come 'usato in contesti formali/informali', 'associato a emozioni positive/negative', ecc.

Contestualizzazione:

- "cane" vicino ad "abbaia" rafforza il senso di 'animale che fa un suono'
- "cane" in "il cane è il migliore amico dell'uomo" potrebbe avere sfumature diverse

Utilità:

- Permette al computer di 'capire' somiglianze e differenze tra parole.
- Aiuta il sistema a gestire parole nuove o rare basandosi sulla somiglianza con parole note

DNA numerico:

- ❑ Dice al computer non solo cosa significa la parola, ma anche come si relaziona con le altre.

36

APPROFONDIMENTO - ADDESTRAMENTO

1. **Addestramento -** LLM sono addestrati su grandi quantità di testo: libri, articoli, siti web, …
   Durante l'addestramento, il modello analizza i testi per capire come le parole si combinano per formare frasi, come le frasi costruiscono paragrafi, e come i paragrafi possono esprimere idee complesse.
   Il modello impara le regole grammaticali, i «*significati*» delle parole, e le relazioni tra di esse.
   Questo processo richiede enormi quantità di dati e capacità di calcolo.
2. Apprendimento
3. Generazione

 37

37

APPROFONDIMENTO - APPRENDIMENTO

1. Addestramento
2. **Apprendimento -** Mentre analizza il testo, il modello costruisce una sorta di "memoria interna" o una rappresentazione del linguaggio. Più il modello incontra una parola o una struttura in vari contesti, meglio riesce a comprendere il suo uso e significato.
   Il concetto di vettori di parole (*word embeddings*) è una parte fondamentale del meccanismo di apprendimento e rappresentazione del linguaggio. Questi vettori sono rappresentazioni matematiche di alta dimensione che cercano di catturare il significato delle parole, i loro usi e le loro relazioni con altre parole.
   Ogni vettore è essenzialmente un array di numeri che il modello di linguaggio impara durante il suo addestramento.
   Questa fase permette al modello di catturare e memorizzare la complessità del linguaggio umano.

   ["Cappuccetto Rosso"] = [0.85, -0.32, 0.91, ..., 0.45]

   0.85: forte correlazione con temi di innocenza o vulnerabilità.
   -0.32: leggera correlazione negativa con concetti di sicurezza o protezione.
   0.91: alta correlazione con tema pericolosità o dell'inganno.
   …
   0.45: moderata correlazione con l'avventura o il viaggi.
3. Generazione

 38

38

APPROFONDIMENTO - GENERAZIONE

1. Addestramento
2. Apprendimento
3. **Generazione -** Quando viene utilizzato, il modello riceve un "prompt" dall'utente, che può essere una frase, una domanda, …. Basandosi su ciò che ha appreso durante l'addestramento, il modello cerca di prevedere e generare la risposta o il testo più probabile che segue l'input.

   Il modello fa questo cercando tra le rappresentazioni interne che ha costruito e selezionando le parole che ritiene più appropriate per completare il testo. Questo processo si basa sulla probabilità: il modello calcola quale parola (o parole) ha la maggiore probabilità di seguire logicamente l'input basato sui testi che ha studiato durante l'addestramento.

   Questi vettori sono generati e utilizzati dai modelli di linguaggio senza che l'utente finale debba interagire direttamente con essi. La manipolazione e l'interpretazione di questi vettori avviene internamente nel modello, che usa algoritmi complessi per ottimizzare e adattare questi vettori durante l'addestramento per migliorare la loro capacità di rappresentare accuratamente il linguaggio umano.

39

APPROFONDIMENTO - TRANSFORMER

Rete Neurale Tradizionale:
Le parole vengono processate una dopo l'altra.
Ogni neurone passa l'informazione al successivo.
L'informazione fluisce in modo lineare.

Architettura Transformer:
Tutte le parole vengono elaborate contemporaneamente.
Il meccanismo di attenzione permette a ogni parola di "guardare" tutte le altre.
L'informazione fluisce in parallelo e in tutte le direzioni.
Vantaggi :
Migliore comprensione del contesto in frasi lunghe.
Capacità di gestire relazioni complesse tra parole.
Più efficiente nell'elaborazione di grandi quantità di testo.

Esempio pratico: Immaginate di chiedere a una classe di leggere un brano.
Reti Neurali tradizionali: è come se ogni studente leggesse una parola e la passasse al successivo.
Transformer: è come se tutti gli studenti leggessero l'intero brano contemporaneamente, poi si confrontassero per capirlo meglio.

40

APPROFONDIMENTO - CARATTERISTICHE DEI TRANSFORMER

Attenzione:
Esempio: Nella frase "Il cane che ha inseguito il gatto è stanco", sa collegare "è stanco" al "cane" anche se sono distanti.

Parallelismo:
È come leggere un intero paragrafo in un colpo d'occhio invece che parola per parola.

Posizione delle parole:
Aggiunge informazioni sulla posizione di ogni parola nella frase.
Aiuta a capire la struttura della frase, come in "Il gatto insegue il topo" vs "Il topo insegue il gatto".

Multi-head Attention:
Immaginate diversi esperti che analizzano la stessa frase da prospettive diverse.
Uno guarda la grammatica, un altro il contesto, un altro ancora il tono emotivo.

Feed-Forward Networks:
Dopo l'attenzione, ci sono reti che elaborano ulteriormente l'informazione.
È come avere una fase di riflessione dopo aver raccolto tutte le informazioni.

 41

41

APPROFONDIMENTO - PARAMETRI DI CONFIGURAZIONE DI UN LLM

Quando si utilizza un LLM come GPT o Claude, si può interagire direttamente con il modello attraverso un'API, che consente di configurare diversi parametri per ottimizzare le risposte del modello in base alle necessità specifiche:

- Temperatura,
- Top P,
- Lunghezza massima,
- Sequenze di stop,
- Penalità di frequenza,
- Penalità di presenza.

Ciascuno di questi parametri ha uno scopo specifico nella modulazione delle risposte del modello, permettendo di bilanciare tra creatività, specificità e concisione delle risposte generate.

 42

42

APPROFONDIMENTO - TEMPERATURA

- influisce sul livello di casualità o determinismo nelle risposte generate dal modello.
- essenziale per modulare come il modello sceglie il prossimo token durante il processo di generazione del testo.

Può assumere valori tipicamente compresi tra 0 e 1:

- Bassa (<0.5): rende il modello più deterministico. Con una temperatura vicina a zero, il modello tende a scegliere ripetutamente i token più probabili, risultando in risposte molto prevedibili e meno varie:

  compiti che richiedono precisione e affidabilità;

- Alta (>0.5): aumenta la casualità nelle scelte dei token. Ciò significa che il modello ha maggiori probabilità di selezionare token meno probabili, portando a risposte più creative, variabili e a volte inaspettate:

  compiti che richiedono creatività o generare idee innovative..

Uso: Utile per adattare il modello a compiti che richiedono originalità e varietà.

43

APPROFONDIMENTO - TOP P (NUCLEUS SAMPLING)

- *nucleus sampling* - importante per controllare il comportamento del generatore di testo;
- metodo di campionamento che seleziona il prossimo token da un sottoinsieme ridotto del vocabolario del modello.
- rappresenta la probabilità cumulativa e determina la grandezza di questo sottoinsieme. In pratica, si configura il modello per considerare il più piccolo insieme di token il cui valore cumulativo delle probabilità supera il valore di top_p.

- top_p=0.9 : per ogni scelta del prossimo token, il modello guarda al set di token che costituiscono il 90% più probabile e seleziona da lì, escludendo token meno probabili che cumulativamente costituiscono il restante 10%.

  Questo è fatto per assicurare che il testo generato sia ragionevolmente probabile (non troppo bizzarro o imprevedibile), ma allo stesso tempo permette una certa varietà e creatività nella generazione del testo.

Uso: Può essere regolato per equilibrare tra risposte convenzionali e innovative, a seconda del contesto.

44

APPROFONDIMENTO - LUNGHEZZA MASSIMA

- determina il numero massimo di token che il modello può generare come risposta.

Impostare una lunghezza massima previene risposte eccessivamente lunghe o divagazioni.

Uso: Importante per mantenere le risposte concise e pertinenti, specialmente in applicazioni interattive o quando si desiderano risposte brevi.

45

APPROFONDIMENTO - SEQUENZE DI STOP

- specifica i token o una sequenza di token che, se generati, causeranno l'interruzione della generazione del testo.

Questo aiuta a controllare la struttura e la lunghezza della risposta.

Uso: Utilizzato per definire un punto di terminazione chiaro nelle risposte, come la fine di un elenco o di un paragrafo.

46

APPROFONDIMENTO - PENALITA'DI FREQUENZA

- modifica la probabilità di selezione di un token in base alla frequenza con cui è già apparso nel testo corrente.

Il valore può variare da 0 a 2 e influisce su come il modello "penalizza" i token già usati:

- **0**: Non viene applicata alcuna penalità. Il modello può ripetere parole o frasi senza alcun condizionamento.
- **positivo**: ogni volta che un token viene generato, la sua probabilità di essere scelto di nuovo viene ridotta.
  - frequency_penalty=0.5 significa che la probabilità di selezionare nuovamente un token già apparso ogni volta che il token è usato, viene diminuita in modo significativo, ma non del tutto.
- **alto** (vicino a 2): rende molto improbabile la ripetizione di parole, promuovendo una diversità estrema nel testo generato che potrebbe, a volte, compromettere la sua coerenza e leggibilità.

La penalità di frequenza è particolarmente utile in applicazioni di generazione di testo dove qualità e leggibilità del contenuto sono cruciali: evitare ripetizioni inutili e promuovere la ricchezza del linguaggio sono aspetti fondamentali.

47

APPROFONDIMENTO - PENALITA'DI PRESENZA

- progettato per influenzare la varietà delle risposte generate modificando la probabilità di selezione dei token già apparsi.
- particolarmente utile per incoraggiare la generazione di contenuto nuovo e diversificato, riducendo la tendenza del modello a ripetere gli stessi concetti o frasi; aggiunge una penalità ai token ogni volta che compaiono nel testo generato, a prescindere dalla frequenza con cui sono apparsi.
  - **0:** Non viene applicata alcuna penalità per la presenza di token, consentendo al modello di ripetere liberamente testo.
  - **basso:** la probabilità di selezionare di nuovo un token già utilizzato viene leggermente ridotta. Questo impedisce al modello di ripetere troppo spesso lo stesso concetto o frase, promuovendo l'introduzione di idee nuove nel testo.
  - **alto:** scoraggia fortemente la ripetizione di qualsiasi token già usato, spingendo il modello a esplorare nuovi territori linguistici e concettuali, il che può essere utile per stimolare la creatività ma a rischio di perdere coerenza.

La penalità di presenza è utile in scenari dove è richiesta la generazione di contenuti freschi e originali ma dove è anche importante mantenere una narrazione o una discussione coesa e logica.

48

APPROFONDIMENTO:
ESEMPIO DI USO PARAMETRI DI CONFIGURAZIONE CON LE API GPT-3 IN PYTHON

```
import openai
response = openai.Completion.create(
    engine="text-davinci-003",
    prompt="Spiega i benefici dell'energia solare per le abitazioni residenziali.",
    temperature=0.7,
    max_tokens=150,
    top_p=0.9,
    stop=["[The End]"],
    frequency_penalty=0.5,
    presence_penalty=0.1 )
print(response.choices[0].text.strip())
```

Dopo decenni di urbanizzazione sfrenata, la natura ha iniziato a reclamare il pianeta Terra. Le città, una volta simboli di progresso umano, ora giacevano abbandonate, avvolte da una fitta vegetazione. Al centro di quello che una volta era New York, un piccolo gruppo di sopravvissuti aveva trovato rifugio in un antico grattacielo, ora coperto di edera e radici. Avevano imparato a convivere con la natura, coltivando giardini verticali e raccogliendo l'acqua piovana per sopravvivere. Il loro leader, Mara, una volta ingegnere ambientale, guidava il gruppo con saggezza. "La natura non è il nostro nemico," spiegava spesso, "ma la nostra più grande alleata se impariamo a rispettarla." Con il passare degli anni, Mara e il suo gruppo stabilirono un equilibrio delicato con l'ambiente circostante, segnando l'inizio di una nuova era in cui umani e natura coesistevano in armonia. E così, nel cuore di una città riconquistata dal verde, germogliava la speranza per un futuro migliore. [The End]

49

CONSIGLI – LIMITI - RUOLO DELL'INSEGNANTE

Consigli per essere un buon prompter:
- Siate specifici: Più dettagli fornite, migliore sarà la risposta.
- Usate il contesto: Menzionate l'età degli studenti, il livello di difficoltà desiderato, ecc.
- Sperimentate: Provate diverse formulazioni per vedere quali funzionano meglio.
- Siate critici: Verificate sempre le informazioni fornite dall'IA.

Limiti e considerazioni etiche:
- Accuratezza: Non sempre le informazioni generate sono al 100% corrette. In particolare, l'IA potrebbe generare contenuti non appropriati se non guidata correttamente.
- Originalità: Il plagio è una preoccupazione: gli studenti devono capire l'importanza del lavoro originale.
- Pensiero critico: È importante insegnare agli studenti a usare questi strumenti in modo responsabile, ma è soprattutto fondamentale che gli studenti imparino a valutare e verificare le informazioni.

Il ruolo dell'insegnante:
L'IA Generativa è uno strumento, non un sostituto dell'insegnante
Gli insegnanti sono cruciali per guidare, contestualizzare e valutare l'uso di questi strumenti

50


the end!
open AI
© AI image/Limitless Visions/Adobe Stock
GRAZIE PER L'ATTENZIONE
ADORNI@UNIGE.IT

51