

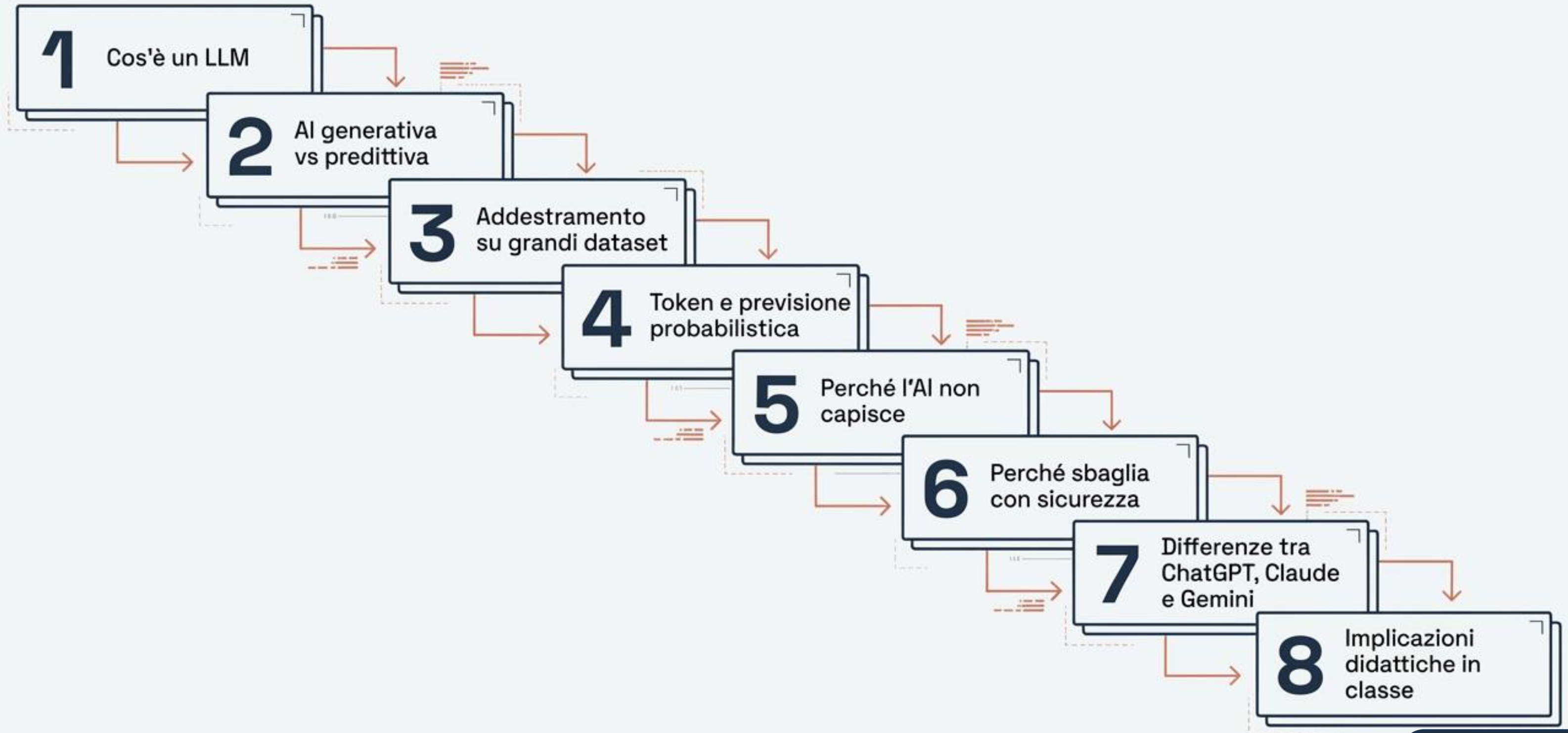


AI, CHE SCUOLA!

LEZIONE 3: Come funziona l'AI: LLM e machine learning

I.I.S. Tonino Guerra – Cervia
Prof. Giuseppe Chierchia

L'Agenda della Lezione (60 min)



Cos'è un LLM (Il Motore dell'AI)

LLM = Large Language Model.
Un modello matematico addestrato per svolgere compiti sul linguaggio naturale basato su miliardi di parametri.

Un LLM **non è un database di risposte**. È un sistema che impara schemi statistici dal linguaggio umano e li usa per generare testo nuovo, parola dopo parola.



AI (GPT-4): ~1.000 miliardi di parametri (connessioni matematiche)



Cervello Umano:
~100.000 miliardi di sinapsi (funzionamento radicalmente diverso)

AI Predittiva vs. AI Generativa

AI Predittiva



Cosa fa: Classifica, prevede un output noto (una categoria o un numero).

Esempi: Filtro antispam, rilevamento frodi, Netflix.

In classe: Analisi dei voti, suggerimenti adattivi.

AI Generativa



Cosa fa: Crea contenuti nuovi.

Esempi: ChatGPT, Claude, Midjourney.

In classe: Scrittura, spiegazioni, materiali didattici.

A scuola ci riferiamo quasi sempre all'AI generativa.

Come “Impara” l’AI: L’Addestramento

Pre-training: Analizza centinaia di miliardi di parole su cluster di GPU per imparare a prevedere la parola successiva.
(Dati grezzi: Wikipedia, Reddit, libri).

Fine-tuning: Affinamento su dati specifici (es. conversazioni utente-assistente) per compiti concreti.

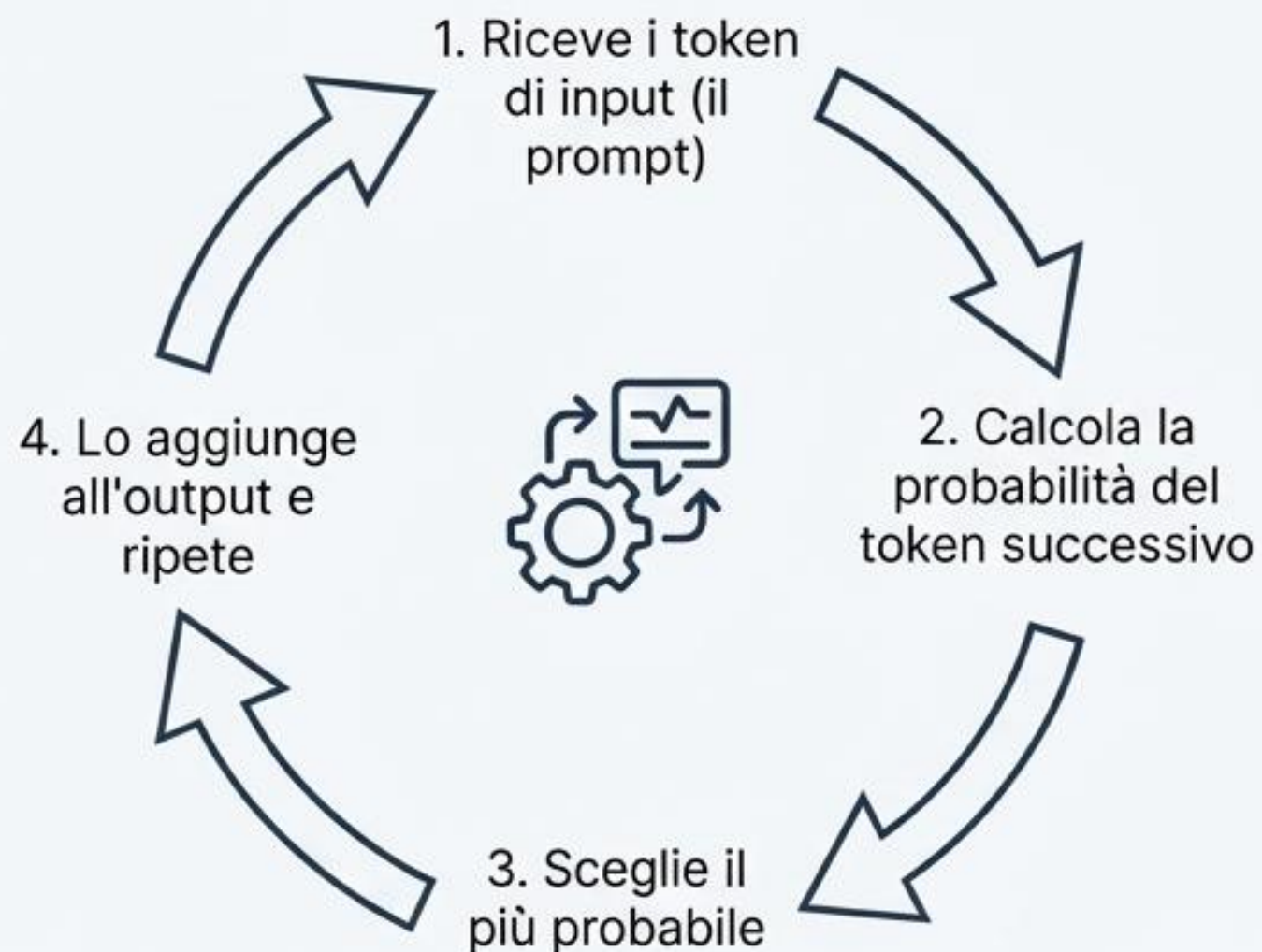
RLHF: Reinforcement Learning from Human Feedback. Valutatori umani guidano il modello verso output sicuri e utili.



Se addestrato su testi inglesi scritti da uomini adulti occidentali, cosa ci aspettiamo riguardo ai suoi bias?

Token e Previsione Probabilistica

Un **Token**: Un pezzo di parola (circa 3-4 caratteri in italiano).



Non lo "sa", è solo statisticamente probabile.

L'Illusione della Comprensione (La “Stanza Cinese”)

L'AI manipola simboli senza capirne il significato (John Searle).



Cosa sembra fare



Cosa fa davvero

Capire la domanda

Elaborare statisticamente i token

Ragionare

Generare token probabilisticamente coerenti

Ricordare

Usare la finestra di conversazione

Sapere se ha ragione

Nessuna verifica esterna automatica

Avere opinioni

Produrre testo stilisticamente coerente

Perché Sbaglia con Sicurezza (Le Allucinazioni)

L'AI sbaglia con un tono autorevole. Non c'è alcun segnale visivo di errore.

Priorità alla Fluidità

Addestrato per coerenza linguistica, non per accuratezza fattuale.

Nessuna Incertezza Nativa

Manca un meccanismo per dire "non lo so"; genera sempre qualcosa.

Dati Imperfetti

Assorbe errori dai dati di training senza filtro critico.

Amnesia di Contesto

"Dimentica" informazioni nelle conversazioni lunghe.

Takeaway: Non è una "AI cattiva", è una proprietà strutturale. Il pensiero critico è fondamentale.

La Scelta dello Strumento (Non esiste il "miglior" modello)

ChatGPT (OpenAI)



- + Punti di forza:** Versatilità, DALL-E, ecosistema plugin.
- Punti deboli:** Allucinazioni frequenti, dati non trasparenti.

Claude (Anthropic)



- + Punti di forza:** Scrittura, ragionamento complesso, meno allucinazioni.
- Punti deboli:** No generazione immagini, ecosistema ridotto.

Gemini (Google)



- + Punti di forza:** Integrazione Google Workspace, multimodale, accesso web reale.
- Punti deboli:** Meno preciso nel ragionamento logico complesso.

Implicazioni Didattiche: Cosa Cambia in Classe



Fact-checking Obbligatorio:
L'AI sbaglia con sicurezza → la verifica è imprescindibile.



Il Prompt come Competenza:
La qualità dell'input determina la qualità dell'output.



Stop all'Antropomorfizzazione:
L'AI non "capisce", non "sa", non "vuole".



Errori come Regola:
Un errore AI non è un'eccezione, è strutturale.



L'AI come Specchio Critico:
Confrontare le risposte, non accettarle passivamente.



Riflessione sui Dati:
Chi ha scritto i testi di addestramento? Quali voci mancano?

Attività Pratiche Suggerite (20 min in classe)

Obiettivo: Sviluppare un rapporto epistemologico, non solo tecnico.



Riepilogo e Spunti di Discussione

Summary Highlights

- ✓ LLM = Motore statistico, non database.
- ✓ Generativa = Crea testo nuovo.
- ✓ Token = Costruzione pezzo per pezzo.
- ✓ Nessuna comprensione reale.
- ✓ Allucinazioni strutturali → Fact-checking.

Come descrivereste l'AI a uno studente di 13 anni?

Qual è il rischio di usarla senza conoscerne il funzionamento?

Come cambia il valore del "pensiero critico" nell'era degli LLM?

Quale di questi concetti è più urgente trasmettere in classe?

Prossima lezione →

Lezione 4: Errori dell'AI: allucinazioni, bias e fake news