STRUMENTI DEL CORSO

Indice degli argomenti

Le parole del corso in ordine alfabetico, con i punti in cui sono spiegate.

Spiegato in: i passi dei moduli, le sezioni di approfondimento dei moduli e le altre pagine del corso (tra «virgolette») in cui l'argomento è trattato davvero. Conviene partire da qui.

Citato anche in: i punti in cui la parola compare di passaggio, utile per vederla usata in altri contesti.

Riassunto grafico: le pagine del PDF «Il percorso della frase» in cui la parola compare.

I colori: il pallino accanto a ogni rimando dice che tipo di pagina è. Percorso base
Per capire Esempi Approfondimento

L'indice copre le parti del corso già pubblicate: moduli 1–2, 10–17, le altre pagine del corso e il riassunto grafico. Si aggiornerà con i nuovi moduli.

Versione web, con i collegamenti: https://www.cultor.net/corso-ia/indice/

A

adattatore

Il file prodotto da una rifinitura con LoRA: contiene solo le tabelle piccole addestrate e si aggiunge al modello di base quando serve.

SPIEGATO IN

Modulo 16, passo 7 LoRA: modificare un modello senza riscriverlo
Modulo 16, passo 11 Che cosa si perde

CITATO ANCHE IN

Modulo 16: approfondimento «Perché imparare fa dimenticare»; approfondimento «Provalo a casa»

addestramento

La fase in cui il modello impara: legge enormi quantità di testo, a gruppi di esempi, e usa l'errore di previsione misurato su ciascun gruppo per aggiornare poco alla volta i suoi parametri.

SPIEGATO IN

Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, passo 7 Numeri imparati, non scritti
Modulo 11, passo 3 Domanda, etichetta, contenuto
Modulo 12, passo 12 Imparare prevedendo
Modulo 12, passo 13 Quanto calcolo richiede imparare
Modulo 12, approfondimento «Sorpresa, bit e perplessità»
Modulo 13, passo 7 Imparare a rispondere
Modulo 16, passo 1 Dopo l'addestramento, i parametri si fermano
Modulo 16, passo 6 Rifinitura: riprendere l'addestramento
Modulo 16, passo 10 Imparare a memoria
Modulo 16, passo 12 Dalle istruzioni alle preferenze
Modulo 17, passo 4 Dove sbaglia più spesso
«In parole semplici» Richiamare ciò che sa di ogni pezzo
«Una frase dentro Minerva» I token si guardano: l'attenzione
«Le operazioni passo per passo» Matematica semplice, quantità enorme
«Come nascono gli embedding» Nessun numero è «giusto»
«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 2: passo 5 · Modulo 10: passi 2, 8 · Modulo 11: passi 2, 5 · Modulo 12: passi 7, 14; approfondimento «Le leggi di scala»; approfondimento «Da modello di base ad assistente»; approfondimento «Il filo dei pregiudizi · 2. Nella previsione» · Modulo 13: passi 1, 11, 13, 14, 17, 19; approfondimento «Perché un modello sbaglia con sicurezza»; approfondimento «Il filo dei pregiudizi · 3. Nelle risposte» · Modulo 14: passi 4, 6, 7, 10; approfondimento «Il filo dei pregiudizi · Nelle immagini e nella voce» · Modulo 15: passi 1, 2, 5, 10; approfondimento «I modelli specializzati per la ricerca, nel dettaglio» · Modulo 16: passi 5, 8, 9, 11, 13; approfondimento «LoRA nei dettagli»; approfondimento «Stile, testi e diritti» · Modulo 17: passi 2, 5, 8, 9, 13 · «Una domanda, tre assistenti»

RIASSUNTO GRAFICO

pagine 1, 2, 3, 4, 5, 6, 7, 9, 10

algoritmo

Procedimento descritto passo per passo, in modo così preciso che un computer lo può eseguire.

SPIEGATO IN

Modulo 1, passo 12 Le parole che si confondono

CITATO ANCHE IN

«Le operazioni passo per passo»

allucinazione

Produzione di un'informazione falsa, inventata o non adeguatamente sostenuta dalle fonti, presentata come se fosse corretta.

SPIEGATO IN

Modulo 17, passo 2 Le «allucinazioni»

CITATO ANCHE IN

Modulo 13: passo 11; approfondimento «Perché un modello sbaglia con sicurezza» · Modulo 15: passo 1

analogia tra vettori

Metodo per completare «a sta a b come c sta a ?»: si calcola b − a + c e si cerca la parola il cui vettore è più vicino al risultato.

SPIEGATO IN

Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, approfondimento «Le analogie tra vettori»
«Come nascono gli embedding» Come nasce la tabella degli embedding

CITATO ANCHE IN

Modulo 1: passo 10 · Modulo 10: approfondimento «Il filo dei pregiudizi (bias) · 1. Nei vettori» · Modulo 14: passo 3 · Modulo 17: passo 6

apprendimento automatico (machine learning)

La parte dell'intelligenza artificiale in cui la macchina ricava le regole dai dati, invece di riceverle scritte da una persona: spesso da molti esempi già classificati, ma non sempre.

SPIEGATO IN

Modulo 1, passo 6 Due strade: regole o esempi
Modulo 1, passo 12 Le parole che si confondono
Modulo 2, passo 3 Macchine che giocano

CITATO ANCHE IN

Modulo 1: passo 11 · Modulo 10: passo 5

apprendimento nel contesto

La capacità di un modello di seguire esempi o istruzioni messi nella richiesta, senza cambiare i parametri.

CITATO ANCHE IN

Modulo 16: passo 3

apprendimento profondo (deep learning)

La parte dell'apprendimento automatico che usa reti neurali con molti strati.

SPIEGATO IN

Modulo 1, passo 12 Le parole che si confondono

CITATO ANCHE IN

Modulo 2: passo 12

attenzione

Meccanismo attraverso il quale un elemento della sequenza confronta la propria rappresentazione numerica con quelle di altri elementi e combina le informazioni ricevute, attribuendo loro pesi diversi.

SPIEGATO IN

Modulo 11, passo 2 L'idea: ogni token utilizza le informazioni di quelli precedenti
Modulo 11, passo 3 Domanda, etichetta, contenuto
Modulo 11, passo 5 Dai punteggi alle percentuali
Modulo 11, passo 7 Molte teste, molte domande
Modulo 11, passo 12 Uno strato, e poi altri quindici
Modulo 11, passo 13 Quando l'attenzione si concentra su «orologio»
Modulo 11, approfondimento «Il parcheggio e le etichette condivise»
Modulo 11, approfondimento «Che cosa l'attenzione non dimostra»
Modulo 13, passo 15 Mentre scrive, la richiesta resta presente
Modulo 13, passo 19 Il viaggio di una richiesta
Modulo 14, passo 5 Le tessere si guardano tra loro
Modulo 14, passo 11 Ascoltare e poi scrivere
Modulo 17, passo 4 Dove sbaglia più spesso
«In parole semplici» Dividere l'attenzione
«Una frase dentro Minerva» I token si guardano: l'attenzione
«Le operazioni passo per passo» Gli stessi passi nei modelli veri
«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 2: passo 13 · Modulo 10: passi 5, 10 · Modulo 11: passi 1, 4, 6, 8, 9, 10; approfondimento «Perché si divide per la radice di 72» · Modulo 12: passi 3, 4, 5, 11; approfondimento «Il filo dei pregiudizi · 2. Nella previsione» · Modulo 13: passo 14 · Modulo 14: passi 1, 4, 6 · Modulo 15: passi 5, 7 · Modulo 16: passo 7; approfondimento «LoRA nei dettagli»; approfondimento «Stile, testi e diritti» · Modulo 17: passo 7 · «Passo passo dentro Minerva»

RIASSUNTO GRAFICO

pagine 2, 6, 7, 9, 10, 14, 15

B

bit

Cifra binaria: vale soltanto 0 o 1, come un interruttore spento o acceso.

SPIEGATO IN

Modulo 10, passo 1 Una frase, ancora muta
Modulo 12, passo 5 La coda lunga
Modulo 12, passo 12 Imparare prevedendo
Modulo 12, approfondimento «Quattro modi di scegliere»
Modulo 12, approfondimento «Sorpresa, bit e perplessità»
Modulo 13, passo 20 Che cos'è uno stile, per un modello
Modulo 15, passo 5 Cercare per significato
Modulo 15, approfondimento «Il punteggio per significato, nei dettagli»
«Una frase dentro Minerva» I voti diventano probabilità
«Come nascono gli embedding» Le parole trovano il loro posto
«Le probabilità dietro i modelli linguistici» Le basi teoriche

CITATO ANCHE IN

Modulo 13: passo 21 · Modulo 14: passo 12 · Modulo 15: approfondimento «Come pesano le parole i motori di ricerca» · Modulo 16: passi 3, 8, 9, 10, 11 · Modulo 17: passo 3

BM25

Una formula usata dai motori di ricerca per dare un punteggio ai documenti: conta le parole della domanda presenti nel documento, dando più peso a quelle rare.

CITATO ANCHE IN

Modulo 15: passo 4; approfondimento «Come pesano le parole i motori di ricerca»

BPE (Byte Pair Encoding)

Metodo per costruire il vocabolario dei token: si parte dai singoli byte e si fondono, una alla volta, le coppie che compaiono più spesso nei testi.

CITATO ANCHE IN

Modulo 10: approfondimento «Come nasce un vocabolario di token»

byte

Unità di base della memoria di un computer: un gruppo di 8 bit adiacenti, cioè 8 cifre binarie (0 o 1), che può assumere 256 valori diversi, da 0 a 255.

SPIEGATO IN

Modulo 10, passo 1 Una frase, ancora muta
Modulo 10, passo 3 Perché pezzi, e non parole
«Come fanno le parole a diventare numeri?»

CITATO ANCHE IN

Modulo 10: approfondimento «Modelli diversi, numeri diversi»; approfondimento «Come nasce un vocabolario di token» · «Una frase dentro Minerva» · «Le probabilità dietro i modelli linguistici»

C

campionamento

Nella generazione del testo, scelta casuale di un token in proporzione alla sua probabilità: un token al 30% esce in media 3 volte su 10; è un'alternativa alla scelta del più probabile.

SPIEGATO IN

Modulo 12, passo 8 Tirare a sorte
Modulo 12, approfondimento «Quattro modi di scegliere»
Modulo 13, passo 18 Ogni volta diverso, non sempre vero

CITATO ANCHE IN

Modulo 12: passo 10 · Modulo 14: passo 9

campionamento a nucleo

Ordina i token per probabilità e tiene il più piccolo gruppo iniziale che raggiunge almeno una certa soglia p, per esempio il 90%.

CITATO ANCHE IN

Modulo 12: passo 8; approfondimento «Quattro modi di scegliere»

catena di Markov

Un modello che prevede il pezzo successivo guardando solo gli ultimi pezzi (uno nel bigramma, due nel trigramma).

SPIEGATO IN

«Le probabilità dietro i modelli linguistici» Le basi teoriche

CITATO ANCHE IN

Modulo 2: passo 7

chatbot

Un programma con cui si conversa scrivendo, o parlando.

CITATO ANCHE IN

Modulo 1: passo 12

componente

Uno dei numeri che formano un vettore.

SPIEGATO IN

«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 10: passi 5, 7 · Modulo 12: approfondimento «Come si fa un esperimento su un modello» · Modulo 14: passo 10 · Modulo 15: approfondimento «Citare le fonti nella risposta» · Modulo 17: passo 4

RIASSUNTO GRAFICO

pagine 9, 10

connessione residua

Collegamento che mantiene un percorso diretto per il vettore precedente e gli somma il risultato dell'elaborazione, invece di sostituirlo.

CITATO ANCHE IN

Modulo 11: passi 6, 12 · «Una frase dentro Minerva»

RIASSUNTO GRAFICO

pagine 2, 15

coseno

Funzione della trigonometria: il coseno di un angolo vale 1 per direzioni uguali, 0 per direzioni perpendicolari e −1 per direzioni opposte.

SPIEGATO IN

Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, passo 8 Confrontare due frecce
Modulo 10, approfondimento «Come si calcola la somiglianza del coseno»

CITATO ANCHE IN

Modulo 10: approfondimento «Il filo dei pregiudizi (bias) · 1. Nei vettori» · Modulo 11: passo 1 · Modulo 12: passo 1 · Modulo 14: passo 6; approfondimento «Sigmoide e softmax» · Modulo 15: approfondimento «I modelli specializzati per la ricerca, nel dettaglio» · Modulo 17: passo 6 · «Come nascono gli embedding»

RIASSUNTO GRAFICO

pagina 5

D

dimensione

Per un vettore, il numero delle sue componenti.

SPIEGATO IN

Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, passo 6 Che cos'è un vettore
Modulo 17, passo 4 Dove sbaglia più spesso
«Le operazioni passo per passo» Gli stessi passi nei modelli veri
«Come nascono gli embedding» Come nasce la tabella degli embedding
«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 2: passo 14 · Modulo 10: passo 10 · Modulo 11: approfondimento «Perché si divide per la radice di 72» · Modulo 12: approfondimento «Le leggi di scala» · Modulo 13: passo 7 · «Una frase dentro Minerva»

RIASSUNTO GRAFICO

pagine 2, 4, 5, 6, 7, 10, 13, 14, 15

dimenticanza catastrofica

Il peggioramento di un modello su ciò che sapeva fare prima, dopo una rifinitura su dati nuovi: i parametri sono condivisi e spostarli verso un compito danneggia gli altri (Kirkpatrick et al., 2017).

SPIEGATO IN

Modulo 16, passo 11 Che cosa si perde

discesa del gradiente

Il metodo con cui un modello corregge i propri parametri durante l'addestramento: li sposta di poco nella direzione che riduce l'errore.

CITATO ANCHE IN

Modulo 10: passo 7 · Modulo 12: passo 12 · Modulo 16: approfondimento «Perché imparare fa dimenticare» · «Le probabilità dietro i modelli linguistici»

E

effetto ELIZA

La tendenza ad attribuire comprensione, intenzioni o sentimenti a un programma che produce frasi adatte alla conversazione.

SPIEGATO IN

Modulo 1, passo 8 Simulare o capire

CITATO ANCHE IN

Modulo 2: passo 6

effetto IA

La tendenza a non considerare più intelligenza artificiale ciò che le macchine hanno imparato a fare: scacchi, navigatori, correttori ortografici.

SPIEGATO IN

Modulo 1, passo 9 L'«effetto IA»

CITATO ANCHE IN

Modulo 1: approfondimento «Due versioni della stessa frase» · Modulo 2: passo 11

ELIZA

Il programma di conversazione scritto da Joseph Weizenbaum negli anni Sessanta, che simulava uno psicoterapeuta ripetendo e riformulando le frasi dell'utente.

SPIEGATO IN

Modulo 1, passo 8 Simulare o capire
Modulo 2, passo 6 ELIZA

CITATO ANCHE IN

Modulo 1: passo 4

embedding

Vettore di numeri che rappresenta un token.

SPIEGATO IN

Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, passo 7 Numeri imparati, non scritti
Modulo 15, approfondimento «I modelli specializzati per la ricerca, nel dettaglio»
«Come fanno le parole a diventare numeri?»
«Una frase dentro Minerva» Ogni pezzo diventa un numero
«Le operazioni passo per passo» Dalla parola al numero
«Come nascono gli embedding» Come nasce la tabella degli embedding
«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 2: passo 12 · Modulo 10: passo 12; approfondimento «Il filo dei pregiudizi (bias) · 1. Nei vettori» · Modulo 11: passi 1, 3, 10 · Modulo 12: passi 1, 12 · Modulo 13: passi 14, 19 · Modulo 15: passi 5, 6 · «In parole semplici»

RIASSUNTO GRAFICO

pagine 2, 3, 4, 6, 12, 14, 15

entropia

Misura, in bit, di quanto è incerta una scelta prima di farla: descrive l'intera distribuzione delle probabilità.

SPIEGATO IN

Modulo 12, passo 5 La coda lunga
«Le probabilità dietro i modelli linguistici» Le basi teoriche

CITATO ANCHE IN

Modulo 12: approfondimento «Quattro modi di scegliere»; approfondimento «Sorpresa, bit e perplessità» · Modulo 17: passo 3

entropia incrociata

La misura dell'errore usata per addestrare i modelli linguistici: la sorpresa media del modello davanti ai token veri di un testo.

SPIEGATO IN

«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 12: approfondimento «Sorpresa, bit e perplessità»

F

finestra di contesto

La quantità massima di testo, misurata in token, che un modello può considerare tutta insieme.

CITATO ANCHE IN

Modulo 13: passo 6 · Modulo 16: passi 2, 5 · Modulo 17: passo 4

FLOP

Un'operazione aritmetica su numeri con la virgola, come una somma o una moltiplicazione.

SPIEGATO IN

Modulo 12, approfondimento «Le leggi di scala»

CITATO ANCHE IN

Modulo 12: passo 13

fonte di prima mano

Una testimonianza o un documento direttamente collegato al fatto: il diario di un testimone, un verbale dell'epoca, una fotografia originale, una delibera che registra una decisione.

CITATO ANCHE IN

Modulo 1: approfondimento «Due versioni della stessa frase» · Modulo 15: passo 7

G

GPT-2

Modello linguistico pubblicato da OpenAI nel 2019.

CITATO ANCHE IN

Modulo 2: approfondimento «Quanto sono cresciuti i modelli» · Modulo 10: approfondimento «Modelli diversi, numeri diversi» · «Le operazioni passo per passo»

RIASSUNTO GRAFICO

pagine 4, 10, 11, 12

I

IA generale (AGI)

Un sistema, per ora ipotetico, capace di affrontare con successo quasi ogni compito intellettuale che sa svolgere un essere umano, compresi quelli mai incontrati.

SPIEGATO IN

Modulo 1, passo 7 IA ristretta e IA generale

IA generativa

I sistemi di intelligenza artificiale che producono contenuti nuovi, come testi, immagini, suoni o video, a partire da una richiesta.

SPIEGATO IN

Modulo 1, passo 12 Le parole che si confondono

IA ristretta

Un sistema di intelligenza artificiale che sa svolgere bene un compito o un gruppo di compiti simili, come giocare a scacchi o riconoscere la voce, ma non altri.

SPIEGATO IN

Modulo 1, passo 7 IA ristretta e IA generale

IA simbolica

Il modo di costruire l'intelligenza artificiale in cui le regole per svolgere un compito sono scritte da persone, come istruzioni esplicite su simboli: parole, oggetti, relazioni («se... allora...»).

CITATO ANCHE IN

Modulo 1: passo 6

IA statistica

Il modo di costruire l'intelligenza artificiale in cui la macchina apprende regolarità dai dati, invece di ricevere regole scritte da una persona.

CITATO ANCHE IN

Modulo 1: passo 6

identificatore (ID)

Il numero d'ordine di un token nel vocabolario.

SPIEGATO IN

Modulo 10, passo 4 Il numero è un indirizzo

Modulo 10, passo 5 Dall'indirizzo alla riga della tabella **«Come fanno le parole a diventare numeri?»**

«Le operazioni passo per passo» Dal numero alla fila di numeri **«Come nascono gli embedding»**

CITATO ANCHE IN

Modulo 10: passo 12 · Modulo 12: passo 9 · Modulo 14: passo 3 · «Una frase dentro Minerva» · «Le probabilità dietro i modelli linguistici»

RIASSUNTO GRAFICO

pagine 2, 3, 4, 14, 15

iniezione di istruzioni

Il tentativo di far passare un ordine nascosto dentro un testo letto dall'assistente, come una pagina web o un documento, per un'istruzione da eseguire.

CITATO ANCHE IN

Modulo 13: passo 5; approfondimento «Testi che contengono istruzioni» · Modulo 15: passo 12

inverno dell'IA

Un periodo di delusione per l'intelligenza artificiale, con meno finanziamenti e meno interesse dopo promesse non mantenute.

CITATO ANCHE IN

Modulo 2: passi 8, 9

istruzioni di sistema

Le indicazioni, scritte da chi offre un servizio di IA, che descrivono come l'assistente deve comportarsi in ogni conversazione.

SPIEGATO IN

Modulo 13, passo 5 Le istruzioni di sistema

CITATO ANCHE IN

Modulo 1: passo 12 · Modulo 13: approfondimento «Che aspetto ha il testo che riceve un assistente»; approfondimento «Testi che contengono istruzioni» · Modulo 17: passi 9, 10

K

key

Vettore con cui un token «si presenta» agli altri: viene confrontato con le query.

SPIEGATO IN

Modulo 11, passo 3 Domanda, etichetta, contenuto

Modulo 11, approfondimento «Il parcheggio e le etichette condivise»

CITATO ANCHE IN

Modulo 11: passi 4, 9, 10, 11 · Modulo 12: passo 9 · Modulo 13: passo 19 · Modulo 14: passo 4 · «Una frase dentro Minerva» · «Le operazioni passo per passo»

RIASSUNTO GRAFICO

pagine 6, 7, 14

L

leggi di scala

Relazioni osservate tra la grandezza di un modello, la quantità di dati, il calcolo usato e l'errore finale: crescendo in modo equilibrato, l'errore scende in modo prevedibile (Kaplan et al., 2020; Hoffmann et al., 2022).

SPIEGATO IN

Modulo 12, approfondimento «Le leggi di scala»

CITATO ANCHE IN

Modulo 12: passo 13 · Modulo 17: passo 4

lente sui logit

Tecnica di analisi: al vettore di un token a metà del modello si applicano la normalizzazione finale e la matrice di uscita, come se fosse l'ultimo strato.

CITATO ANCHE IN

Modulo 12: passo 3 · «Una frase dentro Minerva»

logaritmo

L'operazione inversa dell'elevamento a potenza: $\log_2 8 = 3$ perché $2^3 = 8$.

SPIEGATO IN

«Le probabilità dietro i modelli linguistici» Le basi teoriche

CITATO ANCHE IN

Modulo 12: approfondimento «Sorpresa, bit e perplessità»; approfondimento «Come si fa un esperimento su un modello» · Modulo 15: approfondimento «Come pesano le parole i motori di ricerca» · «Come nascono gli embedding»

logit

Il punteggio grezzo che il modello assegna a ogni token del vocabolario prima della softmax.

SPIEGATO IN

«Una frase dentro Minerva» Strato dopo strato

CITATO ANCHE IN

Modulo 12: passi 1, 2, 3; approfondimento «Come si fa un esperimento su un modello»

RIASSUNTO GRAFICO

pagine 2, 12

LoRA

Low-Rank Adaptation: modo leggero di fare una rifinitura.

SPIEGATO IN

Modulo 16, passo 7 LoRA: modificare un modello senza riscriverlo

Modulo 16, approfondimento «LoRA nei dettagli»

CITATO ANCHE IN

Modulo 16: passi 6, 11; approfondimento «Perché imparare fa dimenticare»

M

maschera causale

Regola che impedisce a ogni token di guardare i token che vengono dopo: il modello legge da sinistra a destra e non può sbirciare il futuro.

CITATO ANCHE IN

Modulo 11: passi 2, 9 · «In parole semplici»

RIASSUNTO GRAFICO

pagine 2, 7, 15

matrice

Tabella rettangolare di numeri.

SPIEGATO IN

Modulo 12, passo 1 Dal vettore ai punteggi

«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 11: passo 3

RIASSUNTO GRAFICO

pagine 3, 6, 7, 8, 10, 11, 12, 14

matrice di uscita

La tabella di parametri che trasforma il vettore finale in punteggi: ha una riga per ogni token del vocabolario, e il punteggio di un token è il prodotto scalare tra il vettore finale e la sua riga.

SPIEGATO IN

Modulo 12, passo 1 Dal vettore ai punteggi

CITATO ANCHE IN

«Le probabilità dietro i modelli linguistici»

memoria esterna

Informazioni conservate da un programma fuori dai parametri del modello e recuperate quando servono: appunti, riassunti, documenti archiviati.

CITATO ANCHE IN

Modulo 16: passi 2, 13

Minerva

Famiglia di modelli linguistici addestrati da zero su testi italiani e inglesi dall'Università La Sapienza di Roma (2024).

SPIEGATO IN

Modulo 2, passo 14 I modelli linguistici

Modulo 2, approfondimento «Quanto sono cresciuti i modelli»

Modulo 10, passo 2 Un token è un pezzo di testo

Modulo 10, passo 5 Dall'indirizzo alla riga della tabella

Modulo 10, approfondimento «Modelli diversi, numeri diversi»

Modulo 11, passo 3 Domanda, etichetta, contenuto

Modulo 12, passo 12 Imparare prevedendo

Modulo 12, passo 13 Quanto calcolo richiede imparare

Modulo 12, passo 14 Dal modello di base all'assistente

Modulo 12, approfondimento «Le leggi di scala»

Modulo 12, approfondimento «Il filo dei pregiudizi · 2. Nella previsione»

Modulo 13, passo 2 Rispondere è continuare

Modulo 13, passo 3 Sapere quando fermarsi

Modulo 13, passo 4 Gli esempi guidano, e possono ingannare

Modulo 13, passo 5 Le istruzioni di sistema

Modulo 13, passo 11 Plausibile non vuol dire vero

Modulo 13, passo 13 Una richiesta vera

Modulo 13, passo 20 Che cos'è uno stile, per un modello

Modulo 13, passo 21 Un nome non basta

Modulo 13, passo 22 Tono e stile

Modulo 13, approfondimento «Una data e una somma: due domande a Minerva»

Modulo 14, passo 1 Altri modelli, stesse idee

Modulo 14, passo 12 Che cosa ha sentito

Modulo 15, passo 1 Ciò che il modello non sa

Modulo 15, passo 5 Cercare per significato

Modulo 15, passo 6 Tre classifiche a confronto

Modulo 15, passo 11 Quando la risposta tradisce la fonte

Modulo 15, passo 12 Istruzioni nascoste nelle fonti

Modulo 16, passo 7 LoRA: modificare un modello senza riscriverlo

Modulo 16, passo 8 La prova: Minerva legge Manzoni

Modulo 17, passo 4 Dove sbaglia più spesso

Modulo 17, approfondimento «Misurare un pregiudizio senza ingannarsi»

«Una frase dentro Minerva» La prova del calcolo

«Passo passo dentro Minerva»

«Le operazioni passo per passo» Gli stessi passi nei modelli veri

«Come nascono gli embedding» Nei modelli veri: la riga di «orologio» in Minerva

CITATO ANCHE IN

Modulo 1: passi 6, 8, 10, 12 · Modulo 2: passi 2, 13; approfondimento «Perché proprio le schede grafiche» · Modulo 10: passi 3, 6, 7, 8, 10, 12; approfondimento «Come nasce un vocabolario di token»; approfondimento «Le analogie tra vettori»; approfondimento «Il filo dei pregiudizi (bias) · 1. Nei vettori» · Modulo 11: passi 2, 7, 8, 9, 10, 12; approfondimento «Dalla somma alla rotazione»; approfondimento «Il parcheggio e le etichette condivise» · Modulo 12: passi 1, 8, 10; approfondimento «Quattro modi di scegliere»; approfondimento «Da modello di base ad assistente» · Modulo 13: passi 1, 6, 8, 9, 14, 15, 16, 17, 18, 19; approfondimento «Il filo dei pregiudizi · 3. Nelle risposte» · Modulo 14: passi 5, 11 · Modulo 15: passi 2, 8, 9, 10 · Modulo 16: passi 1, 2, 3, 5, 6, 9, 10; approfondimento «LoRA nei dettagli»; approfondimento «Provalo a casa» · Modulo 17: passi 3, 6, 7, 9 · «In parole semplici» · «Una domanda, tre assistenti»

RIASSUNTO GRAFICO

pagine 1, 2, 3, 4, 12, 14, 15

Mistral NeMo

Modello linguistico aperto da 12 miliardi di parametri, pubblicato da Mistral AI con NVIDIA nel luglio 2024.

CITATO ANCHE IN

Modulo 10: approfondimento «Modelli diversi, numeri diversi»

modello di base

Un modello appena uscito dal pre-addestramento: è addestrato soprattutto a continuare testi e non è ottimizzato per seguire istruzioni in modo affidabile, anche se a volte risponde correttamente.

SPIEGATO IN

Modulo 12, passo 14 Dal modello di base all'assistente

Modulo 12, approfondimento «Da modello di base ad assistente»

Modulo 13, passo 7 Imparare a rispondere

CITATO ANCHE IN

Modulo 2: passo 14 · Modulo 13: passi 2, 5, 13, 21 · Modulo 15: passi 1, 12 · Modulo 16: passi 7, 12 · Modulo 17: passi 3, 4 · «Passo passo dentro Minerva»

modello linguistico

Programma che, dato un testo, calcola quale token è probabile che venga dopo.

SPIEGATO IN
Modulo 1, passo 10 Il facile è difficile
Modulo 1, passo 12 Le parole che si confondono
Modulo 2, passo 14 I modelli linguistici
Modulo 10, passo 1 Una frase, ancora muta
Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 14, passo 13 Assistenti che vedono e ascoltano
«Come fanno le parole a diventare numeri?» Dalle dita ai numeri
«Il percorso della frase, animato» Come un modello linguistico sceglie la parola successiva
«Le probabilità dietro i modelli linguistici»
CITATO ANCHE IN

Modulo 1: passo 7 · Modulo 2: passi 1, 7, 10, 13 · Modulo 12: approfondimento «Da modello di base ad assistente» · Modulo 13: passi 5, 19 · Modulo 14: passo 5 · Modulo 15: passi 2, 5, 12; approfondimento «Il punteggio per significato, nei dettagli» · Modulo 16: passo 13 · Modulo 17: passi 1, 4 · «In parole semplici» · «Le operazioni passo per passo»
RIASSUNTO GRAFICO
pagina 1

multimodale
Detto di un modello o di un assistente che lavora con più tipi di materiale insieme: testo, immagini, suono, video.
CITATO ANCHE IN
Modulo 14: passo 13

N

neurone artificiale
Un piccolo elemento di calcolo ispirato, molto alla lontana, al neurone del cervello: riceve dei numeri, li moltiplica per i suoi pesi, li somma e decide che cosa trasmettere.
CITATO ANCHE IN
Modulo 2: passi 2, 5

normalizzazione
Operazione che regola la scala delle componenti di un vettore prima del calcolo successivo.
SPIEGATO IN
«Una frase dentro Minerva» Strato dopo strato
«Le operazioni passo per passo» 243 operazioni per una parola
CITATO ANCHE IN
Modulo 11: passi 3, 12 · Modulo 12: passi 1, 3 · «Le probabilità dietro i modelli linguistici»
RIASSUNTO GRAFICO
pagine 2, 11, 13, 15

P

paradosso di Moravec
L'osservazione del ricercatore di robotica Hans Moravec (anni Ottanta): per un computer è relativamente facile ciò che per noi è difficile, come il calcolo o gli scacchi, e difficilissimo ciò che per noi è banale, come riconoscere oggetti o muoversi in una stanza.
SPIEGATO IN
Modulo 1, passo 10 Il facile è difficile

parametri
I numeri che il modello impara durante l'addestramento e poi tiene fissi quando viene usato.
SPIEGATO IN
Modulo 2, approfondimento «Quanto sono cresciuti i modelli»
Modulo 12, passo 13 Quanto calcolo richiede imparare
Modulo 12, approfondimento «Le leggi di scala»
Modulo 15, passo 2 L'idea: prima cercare, poi rispondere
Modulo 16, passo 1 Dopo l'addestramento, i parametri si fermano
Modulo 16, passo 2 Tre modi di adattare un modello
Modulo 16, passo 6 Rifinitura: riprendere l'addestramento
Modulo 16, passo 7 LoRA: modificare un modello senza riscriverlo
Modulo 16, passo 13 Che cosa non impara mai da sola
Modulo 17, passo 4 Dove sbaglia più spesso
«Le probabilità dietro i modelli linguistici» Come il modello impara
CITATO ANCHE IN
Modulo 2: passi 11, 12, 14 · Modulo 10: approfondimento «Modelli diversi, numeri diversi» · Modulo 11: passo 12 · Modulo 12: passi 1, 12, 14; approfondimento «Sorpresa, bit e perplessità»; approfondimento «Da modello di base ad assistente»; approfondimento «Come si fa un esperimento su un modello» · Modulo 14: passi 7, 12 · Modulo 15: approfondimento «I modelli specializzati per la ricerca, nel dettaglio» · Modulo 16: passi 4, 5, 11, 12; approfondimento «Perché imparare fa dimenticare» · Modulo 17: passi 9, 15
RIASSUNTO GRAFICO
pagina 3

parole vuote
Parole molto frequenti e povere di significato proprio, come articoli, preposizioni e congiunzioni, che la ricerca per parole di solito ignora.
CITATO ANCHE IN
Modulo 15: passo 4

perceptron
Il primo neurone artificiale capace di imparare, presentato da Frank Rosenblatt nel 1958: quando sbaglia una risposta corregge un poco i propri pesi.
SPIEGATO IN
Modulo 2, passo 5 Il perceptron
Modulo 2, passo 8 Il primo inverno
Modulo 2, passo 10 Le reti tornano
Modulo 2, approfondimento «Il perceptron a mano»
CITATO ANCHE IN
Modulo 2: passo 2

perdita
La misura dell'errore durante l'addestramento: per un token è −ln p, la sorpresa davanti al token vero misurata in nat.
SPIEGATO IN
«Le probabilità dietro i modelli linguistici» Come il modello impara
CITATO ANCHE IN
Modulo 11: approfondimento «Il parcheggio e le etichette condivise» · Modulo 12: approfondimento «Sorpresa, bit e perplessità» · Modulo 16: passo 11

perplessità
Misura di quanto un testo sorprende un modello: è come se, a ogni token, il modello esitasse tra quel numero di alternative equiprobabili.

SPIEGATO IN
Modulo 12, approfondimento «Sorpresa, bit e perplessità»
«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN
Modulo 12: passi 5, 12

peso di attenzione
La quota dell'attenzione che un token dedica a un altro.

SPIEGATO IN
Modulo 11, passo 5 Dai punteggi alle percentuali

CITATO ANCHE IN
Modulo 11: approfondimento «Che cosa l'attenzione non dimostra» · «Una frase dentro Minerva»

pixel
Il puntino più piccolo di un'immagine digitale.

SPIEGATO IN
Modulo 14, passo 2 Un'immagine è una griglia di numeri
Modulo 14, passo 3 Le tessere sono i token dell'immagine
Modulo 14, passo 4 Dalla tessera al vettore

CITATO ANCHE IN
Modulo 14: passo 1

pre-addestramento
La prima e più lunga fase dell'addestramento: il modello legge enormi quantità di testo e impara a prevederne ogni token successivo.

CITATO ANCHE IN
Modulo 12: passo 13; approfondimento «Da modello di base ad assistente» · Modulo 13: passo 7 · Modulo 16: passo 6

pregiudizio (bias)
Distorsione sistematica: un errore che va sempre nella stessa direzione, invece di distribuirsi a caso.

SPIEGATO IN
Modulo 10, approfondimento «Il filo dei pregiudizi (bias) · 1. Nei vettori»
Modulo 12, approfondimento «Il filo dei pregiudizi · 2. Nella previsione»
Modulo 13, approfondimento «Il filo dei pregiudizi · 3. Nelle risposte»
Modulo 14, approfondimento «Il filo dei pregiudizi · Nelle immagini e nella voce»
Modulo 17, passo 5 Da dove vengono i pregiudizi
Modulo 17, passo 6 Pregiudizi nelle parole
Modulo 17, passo 7 Pregiudizi nelle probabilità
Modulo 17, passo 8 I pregiudizi meno visibili
Modulo 17, passo 9 Correggere i pregiudizi, e i limiti della correzione
Modulo 17, approfondimento «Misurare un pregiudizio senza ingannarsi»

CITATO ANCHE IN
Modulo 1: passo 2 · Modulo 10: passo 9; approfondimento «Le analogie tra vettori» · Modulo 12: passo 4 · Modulo 13: passo 14 · Modulo 14: approfondimento «Generare immagini e suoni» · Modulo 17: approfondimento «Compiacenza: perché gli assistenti danno ragione» · «Come nascono gli embedding»

RIASSUNTO GRAFICO
pagina 10

probabilità condizionata
La probabilità di un evento sapendo che ne è avvenuto un altro: P(A | B).

SPIEGATO IN
«Le probabilità dietro i modelli linguistici» Quattro domande

prodotto scalare
Si moltiplicano le componenti di due vettori nella stessa posizione e si sommano i risultati.

SPIEGATO IN
Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 11, passo 4 Il confronto: un punteggio per token
«Le operazioni passo per passo» La parola guarda le altre

CITATO ANCHE IN
Modulo 11: passo 3; approfondimento «Perché si divide per la radice di 72» · Modulo 12: passo 1 · «Una frase dentro Minerva» · «Come nascono gli embedding»

RIASSUNTO GRAFICO
pagine 2, 5, 7, 13

proiezione
Trasformazione matematica che permette di passare da una rappresentazione numerica a un'altra.

CITATO ANCHE IN
Modulo 10: passo 10 · Modulo 11: passi 3, 6 · Modulo 12: approfondimento «Come si fa un esperimento su un modello» · Modulo 14: passo 13

prompt
Il testo che si scrive all'assistente per chiedergli qualcosa: domanda, istruzioni, esempi.

CITATO ANCHE IN
Modulo 13: passo 5

pubblico dominio
Condizione di un'opera i cui diritti economici d'autore sono scaduti: in Italia, dal 1° gennaio del settantunesimo anno dopo la morte dell'autore.

CITATO ANCHE IN
Modulo 13: passo 24 · Modulo 16: approfondimento «Stile, testi e diritti»

Q

query
Vettore che rappresenta ciò che un token «cerca» negli altri.

SPIEGATO IN
Modulo 11, passo 3 Domanda, etichetta, contenuto

CITATO ANCHE IN
Modulo 11: passi 4, 9, 10, 11; approfondimento «Il parcheggio e le etichette condivise» · Modulo 13: passo 19 · Modulo 14: passo 4 · Modulo 15: passo 5 · «Le operazioni passo per passo»

RIASSUNTO GRAFICO
pagine 6, 7, 14

R

RAG

Retrieval-Augmented Generation, generazione aumentata dal recupero: un programma cerca dei documenti e li scrive prima della domanda, così il modello risponde leggendoli.

CITATO ANCHE IN

Modulo 15: passo 2

rete in avanti

La seconda metà di ogni strato: una piccola rete che rielabora ogni token da solo, senza guardare gli altri.

SPIEGATO IN

Modulo 11, passo 12 Uno strato, e poi altri quindici
«Una frase dentro Minerva» Strato dopo strato
«Le operazioni passo per passo» Gli stessi passi nei modelli veri

CITATO ANCHE IN

Modulo 14: passo 5 · Modulo 16: passo 7 · «In parole semplici»

RIASSUNTO GRAFICO

pagine 2, 15

rete neurale

Un modello fatto di molte unità di calcolo semplici, i neuroni artificiali, collegate fra loro e organizzate in strati.

SPIEGATO IN

Modulo 1, passo 12 Le parole che si confondono
Modulo 2, passo 11 Deep Blue

CITATO ANCHE IN

Modulo 2: passi 2, 5, 8, 10, 12, 13; approfondimento «Perché proprio le schede grafiche» · Modulo 13: passi 8, 9, 14 · Modulo 14: passo 10 · «Come nascono gli embedding»

RIASSUNTO GRAFICO

pagina 1

retropropagazione

Il metodo per correggere i pesi di una rete neurale con molti strati: si misura l'errore all'uscita e si distribuisce la correzione all'indietro, strato dopo strato.

SPIEGATO IN

Modulo 2, passo 10 Le reti tornano
«Come nascono gli embedding» Come nasce la tabella degli embedding

CITATO ANCHE IN

«Le probabilità dietro i modelli linguistici»

ricerca a fascio

Metodo di generazione che tiene aperto un numero limitato di continuazioni in parallelo e sceglie, fra quelle esplorate, quella con il punteggio migliore, a volte corretto per la lunghezza.

CITATO ANCHE IN

Modulo 12: approfondimento «Quattro modi di scegliere»

rifinitura

Addestramento supplementare, su dati scelti per uno scopo e di quantità variabile, di un modello già pre-addestrato.

SPIEGATO IN

Modulo 16, passo 2 Tre modi di adattare un modello
Modulo 16, passo 6 Rifinitura: riprendere l'addestramento

CITATO ANCHE IN

Modulo 12: passo 14; approfondimento «Da modello di base ad assistente» · Modulo 13: passo 7 · Modulo 16: passi 8, 9, 11, 12, 13 · Modulo 17: passi 4, 5, 10, 12 · «Come nascono gli embedding»

RLHF

Apprendimento per rinforzo dal giudizio umano (Reinforcement Learning from Human Feedback): persone confrontano coppie di risposte; le loro preferenze addestrano un modello di ricompensa, che dà un voto alle risposte, e il modello linguistico viene poi ottimizzato per ottenere voti migliori (Ouyang et al., 2022).

CITATO ANCHE IN

Modulo 12: passo 14; approfondimento «Da modello di base ad assistente» · Modulo 13: passo 7 · Modulo 16: passo 12

RoPE (rotazione di posizione)

Il modo in cui Minerva e Mistral segnano la posizione dei token: ruotano query e key di angoli che crescono con la posizione.

CITATO ANCHE IN

Modulo 11: passo 9 · Modulo 13: passo 19 · «Le operazioni passo per passo»

S

scelta golosa

Il modo più semplice di generare un testo: a ogni passo si prende il token con la probabilità più alta.

SPIEGATO IN

Modulo 12, passo 6 Scegliere il più probabile
«Una frase dentro Minerva» Si sceglie il token

CITATO ANCHE IN

Modulo 12: passi 7, 9, 10; approfondimento «Quattro modi di scegliere» · Modulo 13: passi 18, 21; approfondimento «Una data e una somma: due domande a Minerva»

scheda grafica (GPU)

Il processore nato per disegnare le immagini dei videogiochi: ha migliaia di piccole unità che fanno moltiplicazioni e somme tutte insieme.

SPIEGATO IN

Modulo 2, passo 12 Dati e schede grafiche
Modulo 2, approfondimento «Perché proprio le schede grafiche»

CITATO ANCHE IN

Modulo 2: passo 13

segnale di fine turno

Un token speciale che un assistente impara a scrivere quando la sua risposta è finita.

CITATO ANCHE IN

Modulo 13: passi 3, 7, 12, 17; approfondimento «Che aspetto ha il testo che riceve un assistente»

segnale di inizio

Token speciale, scritto <s>, che negli esempi del corso segnala l'inizio del testo.

SPIEGATO IN

Modulo 11, passo 7 Molte teste, molte domande
«Una frase dentro Minerva» La frase si spezza in pezzi

CITATO ANCHE IN

Modulo 10: passo 2; approfondimento «Modelli diversi, numeri diversi» · Modulo 11: passo 8; approfondimento «Il parcheggio e le etichette condivise»

sigmoide

Funzione che trasforma un singolo punteggio in una percentuale da 0 a 100%: un punteggio 0 diventa il 50%.

SPIEGATO IN

Modulo 14, approfondimento «Sigmoide e softmax»

CITATO ANCHE IN

Modulo 14: passi 6, 7

sistema di IA

Nel regolamento europeo sull'intelligenza artificiale (AI Act, art.

SPIEGATO IN

Modulo 1, passo 11 Le definizioni ufficiali

CITATO ANCHE IN

Modulo 1: approfondimento «La definizione europea, elemento per elemento»

sistema esperto

Un programma che raccoglie la conoscenza di specialisti sotto forma di centinaia o migliaia di regole «se... allora...» scritte a mano, come MYCIN per gli antibiotici o XCON per configurare i computer.

SPIEGATO IN

Modulo 2, passo 9 I sistemi esperti

CITATO ANCHE IN

Modulo 2: passo 10

softmax

Funzione che trasforma una fila di punteggi qualsiasi in quote positive che sommano a 1, cioè al 100%, dando molto più peso ai punteggi più alti.

SPIEGATO IN

Modulo 11, passo 5 Dai punteggi alle percentuali
Modulo 11, approfondimento «Perché si divide per la radice di 72»
Modulo 12, passo 4 Dai punteggi alle probabilità
Modulo 14, approfondimento «Sigmoide e softmax»
«Una frase dentro Minerva» I token si guardano: l'attenzione
«Le operazioni passo per passo» La parola scelta

CITATO ANCHE IN

Modulo 11: passi 3, 4, 7; approfondimento «Il parcheggio e le etichette condivise» · Modulo 13: passo 19 · Modulo 14: passi 6, 7 · «In parole semplici» · «Come nascono gli embedding»

RIASSUNTO GRAFICO

pagine 2, 8, 11, 13, 14, 15

somiglianza del coseno

Misura di quanto due vettori puntano nella stessa direzione: vale 1 se hanno la stessa direzione, 0 se sono perpendicolari, −1 se hanno direzioni opposte.

SPIEGATO IN

Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, approfondimento «Come si calcola la somiglianza del coseno»

CITATO ANCHE IN

Modulo 10: passo 8 · Modulo 11: passo 1 · Modulo 12: passo 1

sorpresa

Quanto era improbabile un evento accaduto: $-\log_2 p$, misurata in bit.

SPIEGATO IN

Modulo 12, passo 5 La coda lunga
Modulo 12, passo 12 Imparare prevedendo
Modulo 12, approfondimento «Sorpresa, bit e perplessità»
Modulo 13, passo 20 Che cos'è uno stile, per un modello
Modulo 15, passo 5 Cercare per significato
Modulo 16, passo 8 La prova: Minerva legge Manzoni
«Una frase dentro Minerva»
«Come nascono gli embedding» «l'orologio suona», con tutti i conti
«Le probabilità dietro i modelli linguistici» Le basi teoriche

CITATO ANCHE IN

Modulo 10: passo 5 · Modulo 14: passo 12 · Modulo 15: passo 6; approfondimento «Il punteggio per significato, nei dettagli»; approfondimento «I modelli specializzati per la ricerca, nel dettaglio» · Modulo 16: passi 3, 10, 11

sovradattamento

Quando un modello si adatta troppo ai dati usati nell'addestramento e funziona peggio su dati nuovi.

CITATO ANCHE IN

Modulo 16: passo 10

spaCy

Libreria gratuita per l'analisi del linguaggio, sviluppata dall'azienda Explosion.

SPIEGATO IN

Modulo 10, approfondimento «Le analogie tra vettori»
Modulo 17, passo 6 Pregiudizi nelle parole

CITATO ANCHE IN

Modulo 10: passo 5; approfondimento «Modelli diversi, numeri diversi»; approfondimento «Il filo dei pregiudizi (bias) · 1. Nei vettori» · Modulo 17: approfondimento «Provalo a casa»

spettrogramma

Rappresentazione di un suono come una specie di immagine: per ogni istante (in orizzontale) indica quanta energia c'è in ogni fascia di frequenza (in verticale).

SPIEGATO IN

Modulo 14, passo 10 Dall'onda allo spettrogramma

CITATO ANCHE IN

Modulo 14: passi 1, 11

stanza cinese

L'esperimento mentale del filosofo John Searle (1980): una persona che non conosce il cinese risponde perfettamente a domande in cinese seguendo un libro di istruzioni, senza capire nulla.

SPIEGATO IN

Modulo 1, passo 8 Simulare o capire

stile

Il modo di scrivere di un autore o di un testo: scelte di parole, ritmo, costruzione delle frasi.

SPIEGATO IN

Modulo 13, passo 20 Che cos'è uno stile, per un modello
Modulo 13, passo 22 Tono e stile
Modulo 13, passo 23 Misurare un'imitazione
Modulo 13, passo 24 Imitare non è copiare
Modulo 16, approfondimento «Stile, testi e diritti»
«Una domanda, tre assistenti» Lo stile: ognuno ha la sua battuta

CITATO ANCHE IN

Modulo 13: passi 16, 21 · Modulo 16: passi 6, 8 · Modulo 17: passi 4, 10, 11

stilometria

Studio dello stile con le misure: lunghezza delle frasi, punteggiatura, frequenza delle parole più comuni.

CITATO ANCHE IN

Modulo 13: passo 23

strato

Uno dei blocchi di calcolo che il modello attraversa in sequenza, ciascuno con attenzione e rete in avanti.

SPIEGATO IN

Modulo 2, passo 8 Il primo inverno
Modulo 2, passo 10 Le reti tornano
Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, passo 12 Il contesto cambia lo stato
Modulo 11, passo 3 Domanda, etichetta, contenuto
Modulo 11, passo 8 La mappa completa di una testa
Modulo 11, passo 12 Uno strato, e poi altri quindici
Modulo 11, passo 13 Quando l'attenzione si concentra su «orologio»
Modulo 12, passo 3 Che cosa direbbe a metà strada
Modulo 12, approfondimento «Come si fa un esperimento su un modello»
Modulo 13, passo 15 Mentre scrive, la richiesta resta presente
Modulo 13, passo 19 Il viaggio di una richiesta
Modulo 17, passo 4 Dove sbaglia più spesso
«In parole semplici» Ripetere molte volte
«Una frase dentro Minerva» Strato dopo strato
«Le operazioni passo per passo» Si ripete, strato dopo strato
«Come nascono gli embedding» Nei modelli veri: la riga di «orologio» in Minerva
«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 1: passo 12 · Modulo 2: approfondimento «Il perceptron a mano» · Modulo 11: passi 2, 7; approfondimento «Perché si divide per la radice di 72»; approfondimento «Il parcheggio e le etichette condivise» · Modulo 12: passi 1, 11 · Modulo 13: passo 6 · Modulo 14: passi 1, 4, 5 · Modulo 16: passo 7

RIASSUNTO GRAFICO

pagine 2, 6, 10, 11, 13, 15

T

t-SNE

Metodo per disegnare su un piano punti che stanno in molte dimensioni, conservando soprattutto chi è vicino a chi (van der Maaten e Hinton, 2008).

CITATO ANCHE IN

Modulo 10: passo 10

Tekken

Il tokenizzatore di Mistral NeMo, costruito su testi di oltre cento lingue.

CITATO ANCHE IN

Modulo 10: approfondimento «Modelli diversi, numeri diversi»; approfondimento «Come nasce un vocabolario di token»

temperatura

Numero che divide i punteggi prima della softmax, quando si tira a sorte.

SPIEGATO IN

Modulo 12, passo 7 La temperatura
Modulo 12, approfondimento «Quattro modi di scegliere»
«Le operazioni passo per passo» La parola scelta
«Come nascono gli embedding» La manopola della doccia

CITATO ANCHE IN

Modulo 10: passi 5, 9 · Modulo 12: passi 1, 2, 4, 5, 8 · Modulo 13: passi 14, 18 · Modulo 17: passo 10

RIASSUNTO GRAFICO

pagine 2, 13, 15

teorema di Bayes

La regola che aggiorna la probabilità di un'ipotesi quando arriva un indizio.

SPIEGATO IN

«Le probabilità dietro i modelli linguistici» Le basi teoriche

tessera

Uno dei quadratini in cui un modello per le immagini divide una foto, per esempio di 16 × 16 pixel.

SPIEGATO IN

Modulo 14, passo 3 Le tessere sono i token dell'immagine
Modulo 14, passo 4 Dalla tessera al vettore
Modulo 14, passo 5 Le tessere si guardano tra loro

CITATO ANCHE IN

Modulo 14: passi 1, 2

test di Turing

La prova proposta da Alan Turing nel 1950, che lui chiamava gioco dell'imitazione: una persona conversa per iscritto con un'altra persona e con una macchina, senza vederle, e deve capire chi è chi.

SPIEGATO IN

Modulo 1, passo 4 Turing e il cambio di domanda
Modulo 1, approfondimento «Il gioco dell'imitazione nella versione di Turing»

CITATO ANCHE IN

Modulo 1: passo 5

testa di attenzione

Una delle unità che calcolano l'attenzione in parallelo, ciascuna con una propria query, cioè una propria «domanda».

SPIEGATO IN

Modulo 11, passo 2 L'idea: ogni token utilizza le informazioni di quelli precedenti
Modulo 11, passo 3 Domanda, etichetta, contenuto
Modulo 11, passo 6 Il messaggio che arriva a «fermò»
Modulo 11, passo 7 Molte teste, molte domande
Modulo 11, passo 8 La mappa completa di una testa
Modulo 11, passo 13 Quando l'attenzione si concentra su «orologio»
Modulo 11, approfondimento «Perché si divide per la radice di 72»
Modulo 11, approfondimento «Il parcheggio e le etichette condivise»
Modulo 11, approfondimento «Che cosa l'attenzione non dimostra»
Modulo 12, passo 11 La prova: spegnere tre teste
Modulo 12, approfondimento «Come si fa un esperimento su un modello»
Modulo 13, passo 15 Mentre scrive, la richiesta resta presente
Modulo 17, passo 4 Dove sbaglia più spesso
«In parole semplici» Più punti di vista insieme
«Una frase dentro Minerva» I token si guardano: l'attenzione
«Le operazioni passo per passo» Il modellino: tutti i conti a mano
«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 11: passi 4, 5; approfondimento «Dalla somma alla rotazione»

RIASSUNTO GRAFICO

pagine 2, 6, 10, 11, 13, 14, 15

token

Pezzo di testo che il modello tratta come unità: una parola, una parte di parola, un segno di punteggiatura o un singolo byte.

SPIEGATO IN

Modulo 10, passo 1 Una frase, ancora muta
Modulo 10, passo 2 Un token è un pezzo di testo
Modulo 10, passo 4 Il numero è un indirizzo
Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, passo 7 Numeri imparati, non scritti
Modulo 10, passo 11 Una riga sola, in qualunque frase
Modulo 10, approfondimento «Modelli diversi, numeri diversi»
Modulo 10, approfondimento «Come nasce un vocabolario di token»
Modulo 11, passo 1 Parole ancora isolate
Modulo 11, passo 2 L'idea: ogni token utilizza le informazioni di quelli precedenti
Modulo 11, passo 3 Domanda, etichetta, contenuto
Modulo 11, passo 4 Il confronto: un punteggio per token
Modulo 11, passo 5 Dai punteggi alle percentuali
Modulo 11, passo 6 Il messaggio che arriva a «fermò»
Modulo 11, passo 8 La mappa completa di una testa
Modulo 11, passo 9 Il problema dell'ordine
Modulo 11, passo 10 Le lancette della posizione
Modulo 11, passo 11 Conta la distanza, non il posto
Modulo 11, passo 12 Uno strato, e poi altri quindici
Modulo 12, passo 1 Dal vettore ai punteggi
Modulo 12, passo 2 32.768 punteggi
Modulo 12, passo 4 Dai punteggi alle probabilità
Modulo 12, passo 5 La coda lunga
Modulo 12, passo 6 Scegliere il più probabile
Modulo 12, passo 7 La temperatura
Modulo 12, passo 8 Tirare a sorte
Modulo 12, passo 9 Aggiungere e ricominciare
Modulo 12, passo 12 Imparare prevedendo
Modulo 12, passo 13 Quanto calcolo richiede imparare
Modulo 12, approfondimento «Quattro modi di scegliere»
Modulo 12, approfondimento «Sorpresa, bit e perplessità»
Modulo 12, approfondimento «Le leggi di scala»
Modulo 13, passo 1 Una chat diventa un unico testo
Modulo 13, passo 2 Rispondere è continuare
Modulo 13, passo 6 Il limite di ciò che vede
Modulo 13, passo 8 Una brutta copia prima della risposta
Modulo 13, passo 11 Plausibile non vuol dire vero
Modulo 13, passo 12 Il ciclo completo di una risposta
Modulo 13, passo 13 Una richiesta vera
Modulo 13, passo 15 Mentre scrive, la richiesta resta presente
Modulo 13, passo 17 Cento parole: contare e correggere
Modulo 13, passo 19 Il viaggio di una richiesta
Modulo 13, passo 20 Che cos'è uno stile, per un modello
Modulo 13, approfondimento «Una data e una somma: due domande a Minerva»
Modulo 14, passo 3 Le tessere sono i token dell'immagine
Modulo 14, passo 5 Le tessere si guardano tra loro
Modulo 14, passo 11 Ascoltare e poi scrivere
Modulo 14, passo 12 Che cosa ha sentito
Modulo 14, passo 13 Assistenti che vedono e ascoltano
Modulo 15, passo 5 Cercare per significato
Modulo 15, passo 10 Le fonti entrano nel testo
Modulo 16, passo 3 Gli esempi nella richiesta
Modulo 16, passo 6 Rifinitura: riprendere l'addestramento
Modulo 16, passo 8 La prova: Minerva legge Manzoni
Modulo 16, passo 10 Imparare a memoria
Modulo 17, passo 3 Perché non dice «non lo so»
Modulo 17, passo 4 Dove sbaglia più spesso
Modulo 17, approfondimento «Misurare un pregiudizio senza ingannarsi»
«Come fanno le parole a diventare numeri?»
«Una frase dentro Minerva» I token si guardano: l'attenzione
«Il percorso della frase, animato» L'idea di base
«Le operazioni passo per passo» Gli stessi passi nei modelli veri
«Come nascono gli embedding» Nei modelli veri: la riga di «orologio» in Minerva
«Le probabilità dietro i modelli linguistici» Come il modello impara

CITATO ANCHE IN

Modulo 10: passi 3, 12; approfondimento «Le analogie tra vettori» · Modulo 11: passo 7; approfondimento «Dalla somma alla rotazione»; approfondimento «Il parcheggio e le etichette condivise» · Modulo 12: passi 10, 14; approfondimento «Come si fa un esperimento su un modello» · Modulo 13: passi 3, 14, 18, 21; approfondimento «Che aspetto ha il testo che riceve un assistente»; approfondimento «Perché un modello sbaglia con sicurezza»; approfondimento «Il filo dei pregiudizi · 3. Nelle risposte» · Modulo 14: passi 1, 6; approfondimento «Generare immagini e suoni» · Modulo 15: approfondimento «Il punteggio per significato, nei dettagli»; approfondimento «Citare le fonti nella risposta» · Modulo 16: passi 5, 9, 11 · «In parole semplici» · «Una domanda, tre assistenti»

RIASSUNTO GRAFICO

pagine 1, 2, 3, 4, 12, 13, 14, 15

tokenizzatore

Programma che spezza il testo in token e assegna a ciascuno il suo numero di vocabolario.

SPIEGATO IN

Modulo 10, passo 1 Una frase, ancora muta
Modulo 10, approfondimento «Modelli diversi, numeri diversi»

CITATO ANCHE IN

Modulo 10: passi 2, 3; approfondimento «Come nasce un vocabolario di token» · Modulo 13: passi 13, 17 · Modulo 17: passo 10 · «Passo passo dentro Minerva»

tono

L'atteggiamento che un testo esprime: triste, ironico, drammatico, formale…

SPIEGATO IN

Modulo 13, passo 11 Plausibile non vuol dire vero
Modulo 13, passo 13 Una richiesta vera
Modulo 13, passo 15 Mentre scrive, la richiesta resta presente
Modulo 13, passo 16 Il tono sposta le probabilità
Modulo 13, passo 22 Tono e stile
«Una domanda, tre assistenti» Come l'abbiamo fatta

CITATO ANCHE IN

Modulo 1: passo 8 · Modulo 13: passo 19; approfondimento «Perché un modello sbaglia con sicurezza» · Modulo 17: passi 1, 3, 11 · «In parole semplici»

top-k

Modo di tirare a sorte che tiene solo i k token più probabili; le loro probabilità vengono riportate a sommare 100% prima dell'estrazione.

CITATO ANCHE IN

Modulo 12: passo 8; approfondimento «Quattro modi di scegliere»

transformer

Architettura di rete neurale presentata nel 2017 (Vaswani et al.) e alla base di quasi tutti i modelli linguistici di oggi.

SPIEGATO IN

Modulo 2, passo 13 Dal Go al transformer
«Le operazioni passo per passo» Dove si trova la parola

CITATO ANCHE IN

Modulo 2: passo 14 · Modulo 10: passo 5 · Modulo 11: passo 9; approfondimento «Perché si divide per la radice di 72»; approfondimento «Dalla somma alla rotazione» · Modulo 12: passo 11 · Modulo 13: passo 19 · Modulo 14: passi 1, 3, 4 · Modulo 17: passo 10

RIASSUNTO GRAFICO

pagina 2

U

UTF-8

Il modo più diffuso di scrivere il testo in byte.

SPIEGATO IN

Modulo 10, passo 1 Una frase, ancora muta

V

value

Vettore con il contenuto che un token «consegna» a chi lo guarda.

SPIEGATO IN

Modulo 11, passo 3 Domanda, etichetta, contenuto
Modulo 11, passo 6 Il messaggio che arriva a «fermò»
Modulo 11, approfondimento «Il parcheggio e le etichette condivise»

CITATO ANCHE IN

Modulo 12: passo 9 · Modulo 13: passo 19 · Modulo 14: passo 4 · «Una frase dentro Minerva» · «Le operazioni passo per passo»

RIASSUNTO GRAFICO

pagine 6, 9, 14

verosimiglianza della domanda

Metodo che valuta quanto un documento rende prevedibili i token di una domanda.

CITATO ANCHE IN

Modulo 15: passo 5

vettore

Fila ordinata di numeri.

SPIEGATO IN

Modulo 2, passo 12 Dati e schede grafiche
Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, passo 6 Che cos'è un vettore
Modulo 10, approfondimento «Modelli diversi, numeri diversi»
Modulo 10, approfondimento «Come si calcola la somiglianza del coseno»
Modulo 10, approfondimento «Le analogie tra vettori»
Modulo 10, approfondimento «Il filo dei pregiudizi (bias) · 1. Nei vettori»
Modulo 11, passo 3 Domanda, etichetta, contenuto
Modulo 11, passo 6 Il messaggio che arriva a «fermò»
Modulo 11, passo 11 Conta la distanza, non il posto
Modulo 11, passo 12 Uno strato, e poi altri quindici
Modulo 12, passo 1 Dal vettore ai punteggi
Modulo 12, passo 3 Che cosa direbbe a metà strada
Modulo 14, passo 1 Altri modelli, stesse idee
Modulo 14, passo 4 Dalla tessera al vettore
Modulo 14, passo 5 Le tessere si guardano tra loro
Modulo 14, passo 6 Testo e immagine nello stesso spazio
Modulo 14, passo 11 Ascoltare e poi scrivere
Modulo 15, passo 5 Cercare per significato
Modulo 15, approfondimento «I modelli specializzati per la ricerca, nel dettaglio»
Modulo 17, passo 4 Dove sbaglia più spesso
Modulo 17, passo 6 Pregiudizi nelle parole
«Una frase dentro Minerva» I token si guardano: l'attenzione
«Le operazioni passo per passo» Nove formule, dette a parole
«Come nascono gli embedding» Come nasce la tabella degli embedding

CITATO ANCHE IN

Modulo 10: passi 1, 7, 8, 9, 10, 11, 12 · Modulo 11: passi 1, 4, 9; approfondimento «Perché si divide per la radice di 72»; approfondimento «Dalla somma alla rotazione» · Modulo

12: approfondimento «Il filo dei pregiudizi · 2. Nella previsione» · Modulo 14: passo 13 · Modulo 15: passo 6 · Modulo 16: approfondimento «LoRA nei dettagli» · Modulo 17: approfondimento «Provalo a casa» · «Come fanno le parole a diventare numeri?» · «Le probabilità dietro i modelli linguistici»
RIASSUNTO GRAFICO
pagine 2, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15

vocabolario

L'elenco di tutti i token che un modello conosce, ciascuno con il suo numero.

SPIEGATO IN
Modulo 10, passo 3 Perché pezzi, e non parole
Modulo 10, passo 5 Dall'indirizzo alla riga della tabella
Modulo 10, approfondimento «Come nasce un vocabolario di token»
Modulo 14, passo 3 Le tessere sono i token dell'immagine
Modulo 17, passo 4 Dove sbaglia più spesso
«In parole semplici» Ritagliare la frase in pezzi
«Come fanno le parole a diventare numeri?» Dalle dita ai numeri
«Una frase dentro Minerva» Ogni pezzo diventa un numero
«Le operazioni passo per passo» Il modellino: tutti i conti a mano
«Come nascono gli embedding» Nei modelli veri: la riga di «orologio» in Minerva
«Le probabilità dietro i modelli linguistici» Come il modello impara
CITATO ANCHE IN
Modulo 10: passo 4; approfondimento «Modelli diversi, numeri diversi» · Modulo 12: passi 1, 4, 7, 11; approfondimento «Come si fa un esperimento su un modello» · Modulo 14: passi 4, 6, 11
RIASSUNTO GRAFICO
pagine 2, 3, 4, 6, 10, 12, 13, 14, 15

W

word2vec

Un metodo presentato da Google nel 2013 che trasforma ogni parola in un vettore di numeri ricavato da come la parola si usa nei testi; famoso l'esempio re − uomo + donna ≈ regina.

CITATO ANCHE IN
Modulo 2: passo 12 · Modulo 10: passo 5