



Culture

INTELLIGENZA ARTIFICIALE Parla lo studioso Nello Cristianini domani al «Festival della mente» di Sarzana

Andrea Capocci pagina 12

Da oggi a domenica la nuova edizione del «Festival della mente», dedicato alla nascita delle idee

Sfide di intelligenza artificiale

Intervista a Nello Cristianini, esperto del tema e docente alla Università di Bath. Domani sarà ospite a Sarzana

ANDREA CAPOCCI

■ ■ ■ Quando il «festival della mente» di Sarzana è nato, gli ideatori probabilmente pensavano alla mente umana. Oggi però è difficile parlare di «mente» senza tirare in ballo l'intelligenza artificiale. E infatti all'edizione 2024, in programma da oggi a domenica 1 settembre, intervverrà anche Nello Cristianini, che insegna Intelligenza artificiale all'università di Bath (Regno Unito) e ha pubblicato da poco il saggio *Machina sapiens. L'algoritmo che ci ha rubato il segreto della conoscenza* per il Mulino (pp. 160, euro 15). «Questo - spiega Cristianini, che ne discuterà domani, alle 11 e alle 14.15 - è un momento speciale: il test di Turing è stato superato. Significa che oggi è difficile capire se stiamo conversando con una persona o con una macchina. E le domande che attendono risposta sono ancora molte».

Sarà il tema del suo intervento al festival di Sarzana?

A Sarzana parlerò degli ultimi scritti di Alan Turing, il fondatore dell'informatica, in cui anticipava alcune delle cose che si stanno realizzando in questi giorni: dagli agenti in grado di conversare al *machine learning*. Sarà divertente collegare gli eventi degli ultimi anni con i sogni e le previsioni degli anni Cinquanta.

Con il suo test, Turing diede una definizione di «intelligenza»?

Solo in un certo senso: superare il test di Turing è una condizione sufficiente ma non necessaria per dimostrare intelligenza. Un gatto o una lumaca non supererebbero il test di Turing ma sono certamente esseri intelligenti. Per non impelagarsi con pro-

blemi filosofici, con quel test Turing stabilì un criterio molto esigente per distinguere ciò che è sicuramente intelligente.

E la sua definizione di intelligenza qual è?

La mia definizione è quella più accettata dalla scienza moderna: un sistema è intelligente quando è in grado di comportarsi in modo efficiente in situazioni mai viste. È anche quella adottata dal regolamento europeo sull'intelligenza artificiale, l'«AI act».

Cosa sappiamo su come vede il mondo o ragiona l'intelligenza artificiale?

Non si sa molto di quello che avviene all'interno dei modelli di linguaggio, come ChatGPT, Gemini o Claude, ma si sta lavorando molto intensamente per studiarlo. In un articolo recente pubblicato da Anthropic, l'azienda che produce Claude, si mostrano le rappresentazioni interne che quel modello ha della realtà. Per esempio, al suo interno esiste un simbolo che rappresenta il Golden Gate e viene attivato a ogni menzione, o descrizione indiretta, di quel ponte, in qualsiasi lingua, o anche se gli si mostra una semplice foto del Golden Gate. E ci sono migliaia di altri simboli dello stesso tipo all'interno di Claude. Insomma, stiamo iniziando a capire qualcosa, in un meccanismo che ha miliardi di componenti. Ma è solo l'inizio.

L'esperto di tecnologia Chris Anderson ha scritto che teorie e modelli non servono più, perché oggi possiamo sfruttare le correlazioni statistiche tra fenomeni senza interrogarsi sui nessi causali. ChatGPT sembra la realizzazione di quella profe-

zia. Lei però sostiene che abbiamo un urgente bisogno di una teoria che spieghi come ragiona ChatGPT. Dunque la teoria serve ancora?

Questo è uno dei punti più affascinanti del momento presente: le teorie linguistiche che abbiamo creato negli ultimi decenni non producono buoni metodi di traduzione automatica, la statistica sì. Allo stesso modo, non abbiamo una teoria di cosa induce un dato cliente a comperare un certo libro o a guardare un certo video, ma la statistica può produrre previsioni utili per il marketing. Insomma: se il gioco è fare previsioni, sembra davvero che le teorie siano solo uno degli approcci possibili e che l'intelligenza artificiale ne preferisca un altro. Ma questo non vuole dire che un giorno non troveremo delle buone spiegazioni per il comportamento dei nostri agenti intelligenti. Per il caso di ChatGPT: penso che ci sia davvero molto che possiamo capire, prima di arrenderci, certo io ci proverò.

Il modo con cui ChatGPT impari ci sta facendo cambiare idea anche su come impariamo noi?

Forse anche noi usiamo le correlazioni statistiche più che la logica deduttiva.

È vero, ma non è sorprendente: forse stupirebbe solo i matematici, non certo un biologo. Quando attraverso la strada e vedo arrivare un autobus, non mi metto a dedurre dagli assiomi: mi sposto e basta. Questo fanno anche gli altri animali. Perché non una macchina?

I «Large Language Model» come ChatGPT imparano dai dati reali e dalle persone fisiche però possono dare vita a errori e «allucinazioni». Sappiamo per-

ché succede?

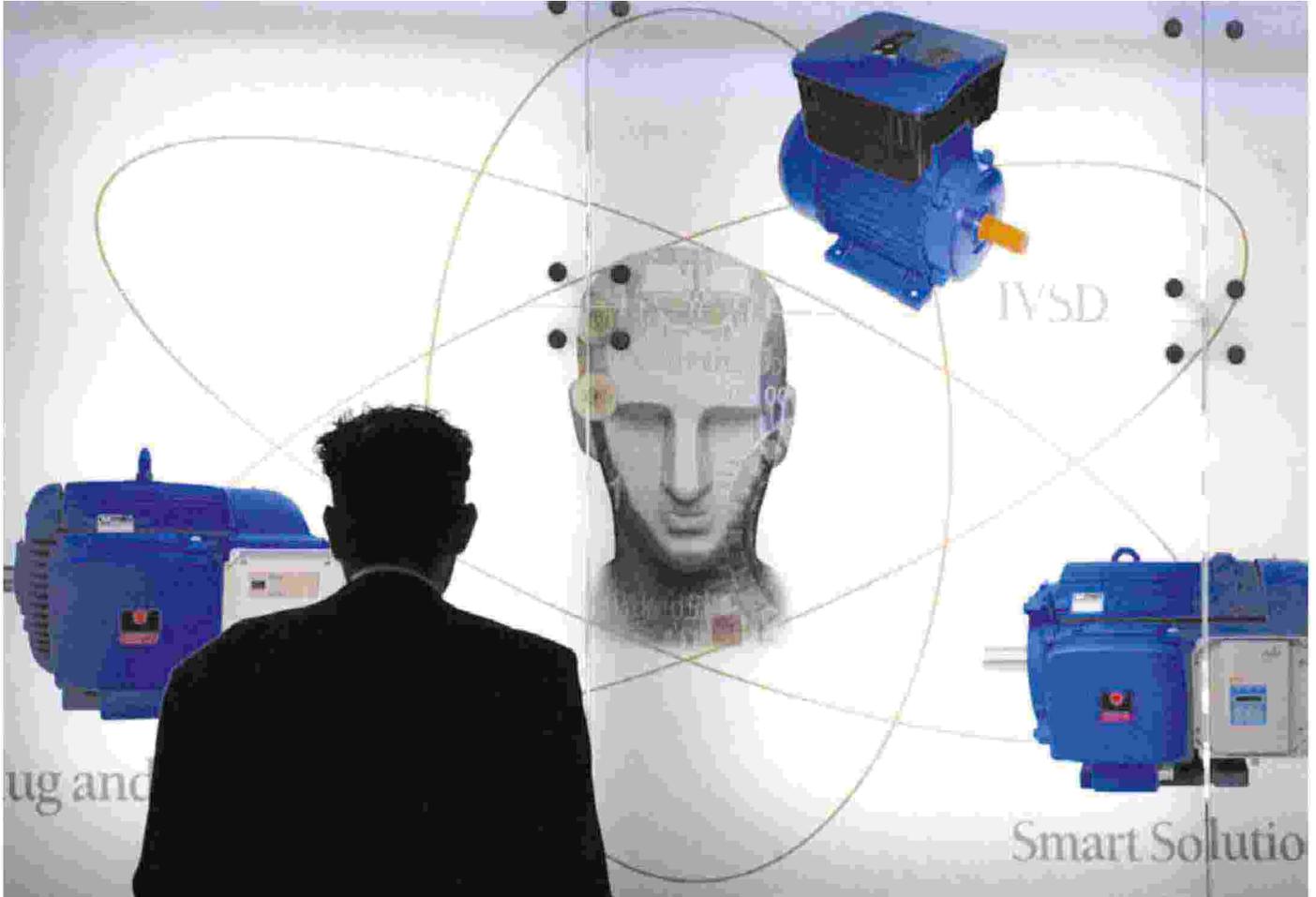
Qualunque conoscenza si trovi all'interno di un *Large Language Model*, proviene da due fonti: i dati di «pre-addestramento», basati su documenti di testo, e i segnali di «fine-tuning», giudizi di tipo «sì/no» forniti da addestratori umani durante il collaudo finale. I problemi emergono dalla prima fase: o i dati contengono informazioni incorrette, oppure vengono collegati tra loro in modo incorretto. La parte di «fine-tuning», tenta di risolvere questo problema, o meglio di mitigarlo. Gli errori insidiosi, le cosiddette allucinazioni, nascono da collegamenti troppo arditi tra informazioni distinte.

Le leggi che regolamentano l'intelligenza artificiale bastano a tenerne sotto controllo i rischi?

L'AI è una tecnologia potente, e quindi è necessario regolamentarne l'uso, come facciamo da sempre anche con le semplici automobili: sarebbe strano il contrario. Controllarla è una questione diversa: posso immaginare molte situazioni in cui qualcosa non va come vorremmo. Avremo bisogno di qualche anno per imparare a controllare questa tecnologia, così com'è avvenuto in molti altri casi. Ma sono certo che ci arriveremo. Nel frattempo si tratta di usare prudenza, buon senso e integrità. Le leggi saranno sempre una parte necessaria della soluzione, e spero che continueremo a raffinarle mentre impariamo a conoscere questa nuova tecnologia.

Ritaglio stampa ad uso esclusivo del destinatario, non riproducibile.

074898



Una pubblicità della fiera della tecnologia di Hannover in Germania foto Ap

*Avremo ancora bisogno di qualche anno
per imparare a controllare questa tecnologia.
Nel frattempo si tratta di usare prudenza*



Ritaglio stampa ad uso esclusivo del destinatario, non riproducibile.

074898