DOPPIOZERO

La superintelligenza artificiale

Pino Donghi
9 Aprile 2025

“È possibile che una macchina sia più intelligente di un essere umano?”

Si apre con questa domanda *SOVRUMANO. Oltre i limiti della nostra intelligenza*, l’ultimo libro di Nello Cristianini edito da il Mulino in questo 2025, a completare una trilogia inaugurata con *La scorciatoia* (2023) e seguita da *Machina sapiens* (2024): e c’è da dire che titolo e sottotitolo lasciano supporre la più che probabile risposta. Al pari dei due precedenti saggi, a colpire è l’uso del passato remoto così da ripercorrere i passaggi di un’avventura di ricerca che – scrive Cristianini nel primo capitolo dal titolo “Il giorno dopo” – è cominciata nel 2012, e dove nel 16° e conclusivo, “L’ultimo esame dell’umanità” facciamo la conoscenza di Dan Hendrycks e Alex Wang, due veterani valutatori dei sistemi di Intelligenza Artificiale (IA), che “alla fine del 2024 avevano deciso di chiudere una volta per tutte la questione di quali compiti cognitivi fossero inaccessibili alle macchine”. *Alla fine del 2024…*, dire “ieri” non rende nemmeno l’idea.

È stato sufficiente che l’anno del Drago del 2012 si riaffacciasse nel 2024, dopo un giro completo del calendario astrale cinese, per assistere all’affermazione dell’era delle macchine pensanti e, con loro, all’assegnazione dei premi Nobel per la fisica e la chimica ad alcuni dei protagonisti della “storia” raccontata da Cristianini. Se vogliamo capire cosa sta per accadere, bisogna tornare a ciò che accadde… *ben dodici anni fa!*

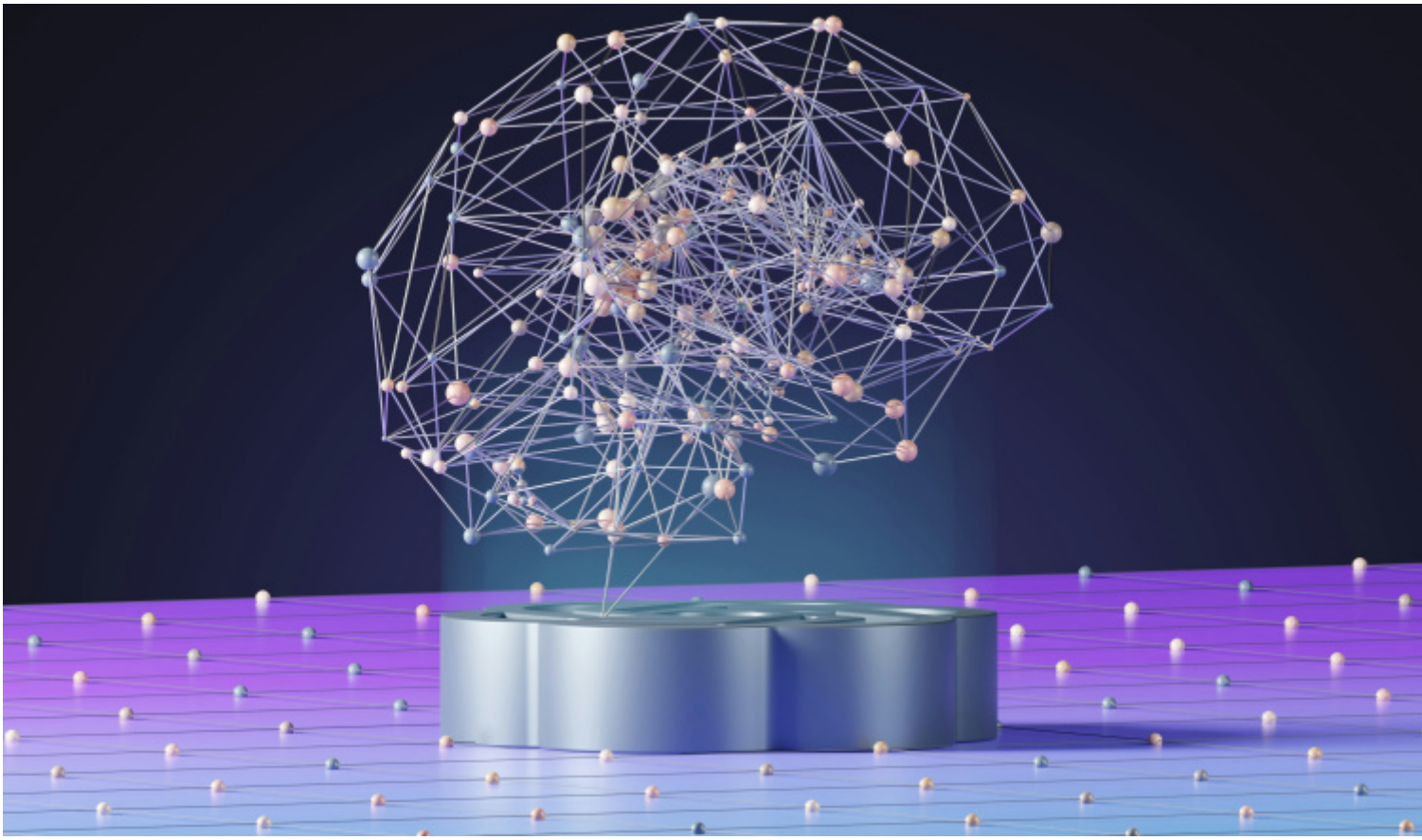
In principio fu il riconoscimento visivo degli oggetti, uno di quei compiti che i critici-scettici dello sviluppo tecnologico periodicamente considerano irrisolvibili e che puntualmente vengono invece risolti. Non solo. Come è stato già per il confronto *uomo* (Lee Sedol) – *macchina* (AlphaGo), nell’antico gioco cinese del Go, la macchina prima “sorpassa” le prestazioni umane, quindi genera “incredulità”, infine vince per “distacco”. Non meraviglia, in una tradizione culturale come quella del campione sud coreano, la dichiarazione di Lee Sedol il 12 Marzo del 2016, dopo la terza sconfitta contro AlphaGo: “Mi scuso di essere così debole”. Del resto, aveva incontrato il “sovrumano, ovvero ciò che trascende la possibilità e i limiti della natura umana”.

Perturbante, si potrebbe commentare, per alcuni disturbante ma era già successo a Gary Kasparov nel 1999, superato a scacchi da DeepBlu e, ancora prima nel 1962, quando Arthur Samuel e il suo programma IBM 704 per giocare a Dama, avevano sconfitto Robert Nealey, un giocatore cieco del Connecticut che quattro anni dopo sarebbe diventato campione di quello Stato. “*Sì però…*” si tratta pur sempre di giochi o, per dirla in termini più tecnici, di compiti singoli e ben definiti, ciò che in psicologia è conosciuto come *narrow ability*. Sicché, anche il riconoscimento visivo di oggetti è un’abilità ristretta, o specifica che dir si voglia, ma la sua complessità, ancora nel 1988, portava Hans Moravec, un pioniere della robotica, a scrivere: “è relativamente facile ottenere prestazioni paragonabili a un adulto nei test di intelligenza o nella dama, ma difficile o impossibile avere abilità di un bambino di un anno quando si parla di percezione”. Una situazione che, per decenni, è stata battezzata “il paradosso di Moravec”. Ma poi, nel 2012 appunto, e precisamente il 13 Ottobre a Firenze, nel corso della terza edizione dell’*ImageNet Large Scale Visual Recognition Challenge* (ILSVRC), esplode il Big Bang, con il nome di AlexNet, un algoritmo costruito dal padre delle reti neurali, Geoff Hinton (sì, proprio lui, il premio Nobel per la Fisica del 2024) e da Ilya Sutskever, che poi avrebbe co-inventato GPT. Se fino a qualche tempo prima, in competizioni come ILSVRC, i “partecipanti” dovevano indovinare a quale categoria, tra le 1000 proposte (del tipo, oggetti, animali, scene di vario tipo), appartenessero 100mila

immagini, diverse ma simili al milione e duecentomila immagini utilizzate per addestrare l'algoritmo, e migliorando le performance in frazioni di percentuale ogni anno… ebbene: in un colpo solo ALexNet migliorò di quasi dieci punti il record dell'anno precedente, correttamente classificando l'83,6% delle immagini. Meglio di un essere umano? "*Sì, però…*".

La scienza è fatta di prove e confronti. L'eroe di questa disciplina – racconta Cristianini – si chiama Andrej Karpathy che nel 2014 decise di sacrificarsi provando a classificare un campione casuale di 1500 immagini tratte da quelle usate per le competizioni tra algoritmi. "È stata dura", aveva commentato, ma ne era valsa la pena: immagini correttamente classificate al 94,9% (5,1 di errore) contro l'83,6 di AlexNet… un buon algoritmo ma nulla di sovrumano. Peccato (!) che già l'anno dopo ResNet raggiungeva il 96,4%; nel 2017 SENet il 97,5 e "… oggi le macchine ci superano in compiti di riconoscimento visivo, e non solo nell'accuratezza, ma anche nella quantità di categorie che possono imparare". Sorpasso, incredulità, distacco: come per Lee Sedol.

Sconcerto? Preoccupazione? Beh, nel 2020, la rivista Nature chiarisce come possiamo leggere questi numeri con uno sguardo, invece, di positiva meraviglia: "Qui presentiamo un sistema di Intelligenza Artificiale in grado di superare gli esperti umani nella previsione del cancro al seno […]. Mostriamo una riduzione assoluta del 5,7 e dell'1,2 (rispettivamente negli Stati Uniti e nel Regno Unito) nei falsi positivi e del 9,4% e del 2,7% nei falsi negativi […] In uno studio di sei radiologi, il sistema di Intelligenza Artificiale ha superato tutti i lettori umani…" Non basta? Dal 1988, ogni due anni, biologi molecolari di tutto il mondo si sfidano con l'obiettivo di prevedere la forma di una data proteina (la competizione si chiama CASP: un rompicapo matematico); per anni i vincitori hanno guadagnato il premio con punteggi tra il 50 e il 60 (in percentuale indica quanto si avvicina la struttura prevista a quella reale): nel 2018 il vincitore è stato AlphaFold, un algoritmo creato da DeepMind (la stessa azienda creatrice di AlphaGo), che ha ottenuto 80 punti: il suo creatore, Demis Hassabis, CEO e co-fondatore di DeepMind, ha ottenuto per questo il Nobel per la Chimica 2024. Sorpasso, incredulità, distacco.



"*Sì, però…*". Cristianini ricorda che la tipica risposta "sì, però", quella che incontriamo in molti pensosi editoriali, nell'espressione scettica di tanti autorevoli studiosi, spesso anche a tavola tra amici, in famiglia…

anche a quel “sì, però” aveva già pensato Alan Turing nel celeberrimo articolo del 1950 che rappresenta, non c’è dubbio, *il verbo* al principio di tutta questa storia. *Computing Machinery and Intelligence*, laddove Turing anticipa pure l’argomento basato su varie disabilità, ovvero l’idea che *sì, però*… ci sono cose che una macchina non potrà mai fare: “*A questo riguardo vengono suggerite numerose caratteristiche X. Ne offro una selezione: essere gentile, intraprendente, bello, amichevole, avere iniziativa, avere senso dell’umorismo, distinguere il bene dal male, commettere errori, innamorarsi, godersi fragole e panna, fare innamorare qualcuno, imparare dall’esperienza, usare le parole in modo appropriato, essere oggetto del proprio pensiero, avere la stessa varietà di comportamento di un uomo, fare qualcosa di veramente nuovo. Di solito non viene offerto alcun supporto per queste affermazioni*”. O, come scrive in altre parole Cristianini: “Ammetto che puoi far fare alle macchine tutte le cose che hai detto, ma non sarai mai in grado di crearne una che faccia X”: tipo un’automobile a guida automatica che se la cavi nel traffico di Los Angeles o – si parva licet – sul Grande Raccordo Anulare di Roma. *Si però...*, si però, per molte della abilità elencate da Turing, nel frattempo si è arrivati a una possibile misura e conseguente trasformazione in test, per esempio *imparare dall’esperienza, usare le parole in modo appropriato, fare qualcosa di veramente nuovo*. E se l’eventuale vita emotiva di una macchina non è osservabile “in principio” (come in principio non possiamo sapere come ci si sente ad essere un pipistrello), intanto: “Comprensione del linguaggio? Superato. Cultura generale? Superato. Matematica elementare? Superato. Scienza, matematica avanzata, riconoscimento di immagini? Tutti superati, talvolta con punteggi superiori a quelli umani”. Come efficacemente commenta l’autore, “Siamo una strana e affascinante specie: mentre spendiamo miliardi per costruire una macchina in grado di pensare, allo stesso momento ci rifiutiamo di accettare che questo sia possibile”.

La domanda che “in principio” fu di Turing, se le macchine possono essere intelligenti, è una domanda dell’altro secolo: oggi che il test omonimo è diventato obsoleto, la domanda diventa se intelligenti le macchine possono esserlo come e più di noi. Se oltre a sorpassarci e distaccarci (dopo un primo momento di incredulità), con riferimento alle abilità specifiche di cui abbiamo principalmente riassunto a partire dalla prima parte di *Sovrumano*, quella dedicata all’l’***a****rtificial* ***n****arrow* ***i****ntelligence* (ANI), sapremo farle andare oltre. Con la seconda parte, AGI, e la terza, ASI, Cristianini ci introduce all’***a****rtificial* ***g****eneral* ***i****ntelligence* e alla sfida finale dell’***a****rtificial* ***s****uper* ***i****ntelligence*. E faremmo un torto al lettore di questa recensione se ne illustrassimo i contenuti: che oltre ad essere straordinariamente interessante, *Sovrumano* è ben scritto e va letto tutto e con attenzione. Sul nostro presente e più immediato futuro non è ammessa ignoranza. A mo’ di indice, nella parte dedicata all’AGI, ossia dell’IA capace di eseguire “tutti” i compiti cognitivi umani, si incontrano le figure dei valutatori dei sistemi intelligenti, impegnati a creare test sempre più difficili, e quelle degli addestratori, impegnati a preparare macchine che li superino (ripassando la psicometria e facendo più vicina conoscenza di GPT, Gemini, Llama e Claude); nella parte dedicata a ASI, l’IA sovrumana, la forma ancora ipotetica che supererebbe quella umana in ogni aspetto cognitivo, comprese creatività, risoluzione dei problemi e conoscenza generale, si incontra lo “spazio da esplorare”, oltre i compiti e i livelli di prestazione per noi inattingibili. Il problema e la sfida non riguarda solo l’eventualità che le macchine ci superino, e di gran lunga, ma che possano svolgere compiti che noi non arriviamo nemmeno a contemplare, problemi cognitivi completamente estranei alla natura umana. Possibile arrivare a queste macchine super intelligenti? E cosa contemplerebbero, cosa scoprirebbero? Qui l’autore torna a un argomento, il primo, che aveva già affrontato in *La scorciatoia*: ogni intelligenza ha dei limiti e quindi ci sono sempre dei territori che le sono inaccessibili, mentre sono facilmente frequentabili da altre intelligenze. La gatta di Cristianini non comprende la meccanica quantistica (e del resto, Feynman sosteneva che non la capiva nessuno), ma al contrario del suo “padrone” lei sente arrivare il postino e vede i topolini in fondo al giardino; così come ci sono macchine che già oggi hanno una visione iperspettrale e che “annusano” la presenza di reti wi-fi; e ci sono anche algoritmi in grado di “ingannare” loro simili inducendoli in errore. Che effetto ci farà, si chiede Cristianini, vivere a contatto con macchine che sapranno fare un passo oltre il limite di ciò che possiamo conoscere?

Per l’ *Humanity’s Last Exam*, anticipavamo, per l’ultimo esame dell’umanità, Dan Hendrycks e Alex Wang vogliono confrontare l’IA con gli individui migliori al mondo sulle domande più difficili che è possibile concepire, e per la costruzione di questo test hanno chiesto aiuto a tutti coloro in grado di contribuire così da

arrivare ad almeno 1000 domande, escludendo tutte quelle facili e per le quali non si trovino già le risposte in rete. La scadenza per l'invio era il 1° Novembre 2024, e le domande rimarranno segrete fino a quando, durante questo 2025, l'IA sosterrà l'esame. Dire che attendiamo con curiosità il risultato, è dire poco.

Certo, poi, quello che sarà, sarà comunque intelligenza, super quanto si vuole ma non coscienza, emozione, libero arbitrio, volontà… poco male: che non siamo insuperabili, nemmeno in intelligenza, in fondo lo abbiamo sempre saputo. Se e quando incontreremo qualcosa di superiore a noi, "lo affronteremo con la stessa speranza, fede, dignitosa resistenza. E gli diremo fieri, *Sì, però…*"

Non a caso, Cristianini chiude il sipario con le parole dell'Amleto, atto I, scena V:

Ci sono più cose in cielo e in terra, Orazio, di quante ne sogni la tua filosofia.

Leggi anche:
Riccardo Manzotti | L'IA pensa. E noi?
Riccardo Manzotti | Se l'IA sa tutto, perché imparare?
Riccardo Manzotti | Intelligenza artificiale: proprio come noi?
Tiziano Bonini | Può l'intelligenza artificiale essere etica?
Stefano Bartezzaghi | Prompt, Chi parla?
Luigi Bonfante | L'arte alla sfida con l'IA

Se continuiamo a tenere vivo questo spazio è grazie a te. Anche un solo euro per noi significa molto.
Torna presto a leggerci e SOSTIENI DOPPIOZERO

NELLO
CRISTIANINI
SOVRUMANO
OLTRE I LIMITI
DELLA NOSTRA INTELLIGENZA