Anche le super Intelligenze Artificiali sbagliano: ecco perché

Data 10 gennaio 2021
Categoria Medicinadigitale

Solo il controllo sistematico ci può salvare

Negli ultimi anni i nuovi sistemi di intelligenza artificiale (IA) basati sull'apprendimento automatico non supervisionato (Deep Learning automatico) hanno destato forti entusiasmi per i risultati che continuano ad ottenere.
Tuttavia, come già avvenuto diverse volte nella storia dell'umanità, l'uomo si innamora spesso delle proprie scoperte e finisce con il sopravvalutarle, ignorandone i limiti e sottovalutandone i pericoli; questo fenomeno psicologico è ben noto agli studiosi di psicologia cognitiva che ne hanno ben descritto le conseguenze: è il noto bias della "overconfidence" la sopravvalutazione inconsapevole (1).
A ciò dobbiamo aggiungere il fatto che una quota significativa della letteratura tecnico-scientifica sulla intelligenza artificiale è prodotta da ricercatori retribuiti dalle multinazionali della informatica: a chi avesse perplessità in merito ricordiamo che un secolo di esperienza nella ricerca farmacologica ci ha dimostrato oltre ogni ragionevole dubbio che i conflitti di interesse condizionano sempre pesantemente il giudizio dei ricercatori, spesso in maniera inconsapevole (2).
Uno dei limiti delle soluzioni fornite dai sistemi di intelligenza artificiale è che queste non si basano solo sulla potenza delle reti neurali ma anche e soprattutto sulla qualità, quantità e modalità di introduzione dei dati.

A questo proposito è importante ricordare che:

1) I dati non parlano da soli: la modalità di ordinamento e di immissione delle informazioni influenzano le conclusioni che ne vengono tratte, sia nei sistemi controllati dagli uomini che, ancor più, nei sistemi non supervisionati: ad esempio dati raccolti in intervalli temporali diversi non sono sempre confrontabili perché le variabili e le dinamiche possono essere molto differenti (3)

2) Le variabili devono essere tutte ben note e definite così come dovrebbero essere note e definite le correlazioni tra le varie variabili: i dispositivi di intelligenza artificiale non sono spesso infatti in grado di riconoscere od escludere tali interferenze autonomamente.

3) E' molto importante conoscere approfonditamente il problema e definire chiaramente gli obiettivi tenendo ben presente il contesto ed i fattori influenzanti. E' famoso il simpatico episodio del premio IG Nobel per la medicina conferito nel 2019 a Silvano Gallus, geniale ricercatore dell'Istituto Mario Negri, per una ricerca sulla pizza. Una parte consistente della stampa medica divulgativa interpretò lo studio (correttamente descritto dal ricercatore!) come una dimostrazione dell'effetto protettivo della pizza verso vari tipi di cancro e malattie cardiovascolari. Gallus chiarì che l'effetto protettivo non era tanto nella pizza ma nella dieta mediterranea alla quale il frequente consumo di pizza era strettamente correlato (4)

4) Una quota importante degli errori dei sistemi di intelligenza artificiale sono dovuti al fatto che l'uomo utilizza vari tipi di memoria (episodica, semantica, operativa-procedurale ecc),nonchè conoscenze implicite utilizzate all'occorrenza, mentre nei sistemi di intelligenza artificiale i processi di ragionamento devono essere accuratamente e dettagliatamente descritti fin nei minimi particolari dagli algoritmi. L'uomo, di fronte ad un problema inaspettato, non si mette a consultare algoritmi ma attinge a risorse di cui egli stesso non è consapevole, mentre l'intelligenza artificiale segue meccanicamente l'algoritmo.

Questo dato può forse spiegare tragici eventi come le tragedie dei Boeing schiantatisi al suolo per incapacità dei sistemi di guida automatica di modificare in tempi brevi la traiettoria, ma più spesso per fortuna hanno risvolti comici come quei sistemi super intelligenti che di fronte al problema di far passare un tavolo per una porta stretta anziché voltarlo od inclinarlo come avrebbe fatto un bambino di 10 anni, proposero di segare le gambe (5).

Ecco uno schema che può aiutare a comprendere I limiti della IA:

[b]Modello tradizionale: esempio di ricerca basata su procedure di regressione logistica[/b]
1) Si raccolgono i dati sui pazienti
2) Si estraggono i fattori di rischio (ad esempio età, familiarità, abitudini voluttuarie precedenti malattie, etc)
3) Si assegna un "peso" a questi fattori e si calcola quindi il rischio totale ponderato.

[b]Modello basato sul Deep Learning non supervisionato[/b]
1)Introduzione di grandi quantità di dati
2) Ricerca di correlazioni di ogni genere, anche casuali e spurie, tra i vari dati utilizzando in genere milioni di simulazioni matematiche
3)Si ottiene un risultato preliminare, si testa questo primo risultato e si valuta se sia coerente con le aspettative iniziali

4) Si ripete il processo per un numero elevato di volte fino ad ottenere le soluzioni desiderate.

[b]Nell'approccio tradizionale la funzione critica ed autocritica di controllo è costantemente attiva: più precisamente vi è un controllo di struttura, di processo, e di risultato.[/b]
Nel Deep Learning supervisionato dovrebbe esservi altrettanta accuratezza, mentre in quello non supervisionato, molto più semplice e meno faticoso per gli umani, la incontrollabilità del processo porta ad accettare facilmente tutti quei risultati che siano compatibili con le nostre aspettative (si cade quindi spesso nel bias della aspettativa, ascertainment bias)
[b]Il controllo è pertanto limitato alla struttura ed al risultato, il processo è incontrollabile[/b].

Il problema degli sconosciuti limiti della IA, decisamente sottovalutato anche nel nostro Paese, è invece seriamente considerato dalla elite medica USA che recentemente nel prestigioso Jama ha proposto una vera e propria guida alla lettura critica degli studi che usano la IA in medicina (6).

Ne proponiamo qui i punti salienti:

[b]Criteri di verifica di studi basati sul Deep Learning non supervisionato[/b]
1) Le conoscenze apportate dalla ricerca presentata sono coerenti con quanto già noto e dimostrato da altre ricerche?
2) Lo studio apporta nuove conoscenze?
3)Lo studio comporta modifiche nella consueta pratica diagnostica o terapeutica?
4) Tali innovazioni possono migliorare la qualità della diagnostica e della terapia? Ed a quale prezzo?
5)Tali innovazioni possono applicarsi ai nostri pazienti nella pratica quotidiana?
6)I benefici attesi sono chiaramente superiori ai costi economici sociali ed esistenziali?

Continua in una prossima pillola…

Riccardo De Gobbi e Giampaolo Collecchia

Bibliografia

1) Pat Croskerry: Bias: a normal operating characteristic of the diagnosing brain Diagnosis 2014; 1(1): 23–27 DOI 10.1515/dx-2013-0028
2) Camilla H Nejstgaard et al.: Association between conflicts of interest and favourable recommendations in clinical guidelines, advisory committee reports, opinion pieces, and narrative reviews: systematic review BMJ 2020;371:m4234 |doi:10.1136/bmj.m4234
3) Martin Bland: Statistica Medica Maggiolo Ed.2019
4)https://www.marionegri.it/magazine/dieta-mediterranea
5) John Pavlus : Common Sense Comes Closer to Computers Quanta Magazine https://www.quantamagazine.org/common-sense-comes-to-computers-20200430/
6) Yun Liu, Po-Hsuan Cameron Che net Al.: How to Read Articles That Use Machine Learning Users' Guides to the Medical Literature JAMA. 2019;322(18):1806-1816. doi:10.1001/jama.2019.16489

Perapprofondire:

Collecchia G. De Gobbi R.: Intelligenza Artificiale e Medicina Digitale Il Pensiero Scientifico Ed. Roma 2020
pensiero.it/catalogo/libri/pubblico/intelligenza-artificiale-e-medicina-digitale