ChatGPT nelle Diagnosi più Difficili

Data 06 agosto 2023
Categoria Medicinadigitale

La realizzazione di ChatGPT(CGPT) e la sua disponibilità gratuita e potenzialmente illimitata per utenti dei più svariati livelli culturali e delle competenze più differenziate ha aperto una nuova era.
Le enormi potenzialità di questa forma di intelligenza artificiale non sono ancora adeguatamente valutate: le riflessioni ed il dibattito sul suo ruolo sono purtroppo fino ad ora limitate prevalentemente agli esperti del settore, ed a pochi utenti qualificati.
È invece quanto mai importante seguire questo travolgente settore della informatica, in quanto in tempi brevi il colosso Google lancerà a livello planetario la sua nuova intelligenza artificiale ed altre IA saranno rese disponibili da altri colossi mondiali nei prossimi mesi.
In questo contesto gli utenti dovrebbero quanto meno essere in grado di valutare la appropriatezza delle risposte delle intelligenze artificiali ai problemi formulati sotto forma di domande.
Pochissimi utenti tuttavia hanno le competenze minimali richieste per verificare le prestazioni di questi sistemi e solo una minoranza di questi utenti ha colto la importanza di una verifica sistematica di tutti i sistemi di intelligenza artificiale.

Pertanto una recente ricerca pubblicata su Jama, attinente le capacità diagnostiche di CGPT suscita grande interesse e costituisce un importante punto di riferimento (1).
Vediamo in sintesi cosa hanno rilevato I ricercatori; alcuni medici ricercatori di Boston hanno addestrato CGPT utilizzando come riferimento le prime sette conferenze clinico- patologiche del New England Journal of medicine del 2023 in modo che CGPT apprendesse la strutturazione dei casi.
Successivamente i ricercatori hanno sottoposto a CGPT tutti i casi clinici pubblicati dal New England tra il gennaio 2021 ed il dicembre 2022 omettendo le ipotesi diagnostiche formulate, la discussione e le conclusioni.

L'obiettivo primario era quello di valutare in quanti casi la diagnosi finale di CGPT coincidesse con quella del NEJM.
L'obiettivo secondario era di valutare se la diagnosi, seppure non corretta, si avvicinava a quella reale o se comunque la diagnosi reale era registrata nell'elenco delle diagnosi differenziali.
Più precisamente I ricercatori hanno utilizzato una scala numerica, già utilizzata in analoghe ricerche, attribuendo un punteggio di 5 a CGPT quando la diagnosi corretta era formulata nel breve elenco di diagnosi differenziali, di 4 se la diagnosi era vicina a quella reale ma non del tutto corretta, di 3 se la ipotesi diagnostica pur non essendo corretta poteva essere utile per giungere alla vera diagnosi. I punteggi tra 2 e 0 venivano attribuiti a diagnosi scorrette che non avevano alcuna utilità nel percorso diagnostico. La media ottenuta da CGPT è stata di 4,2 con una deviazione standard di1,3.
Le prestazioni diagnostiche differenziali CGPT, considerata la complessità dei casi presentati nel New England (ad esempio encefalite in corso di morbo Behcet, miocardite da listeria monocytogenes, angiopatia amiloide cerebrale con fenomeni infiammatori eccetera) possono essere considerate di buon livello: nel 64% dei casi la diagnosi corretta era compresa nell'elenco ristretto delle diagnosi differenziali e nel 39% dei casi la prima ipotesi di CGPT era la diagnosi corretta.

Concludendo I ricercatori sottolineano come le prestazioni di CGPT, seppure ancora non ottimali competono su un piano paritetico con i migliori software di diagnosi differenziale.
Riflettendo al di qua dell'oceano potremmo concludere che questa ricerca ci porta nuove buone notizie ma ci propone il problema, sempre più urgente e rilevante, di come interagire in maniera ottimale con questo strumento di enormi capacità e potenzialità e quindi con immense ed imprevedibili potenziali ricadute.

Riccardo De Gobbi e Giampaolo Collecchia

Bibliografia

1)Kanjee Z, Crowe B, Rodman Adam: Accuracy of a Generative Artificial Intelligence Model in a Complex Diagnostic Challenge Research Letter JAMA Published online June 15, 2023 doi:10.1001/jama.2023.8288

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