

Videnskabelig Leder

Vejledning eller vildledning af kunstig intelligens i medicinsk forskning?

Andreas Pihl

Ugeskr Læger 2024;186:V205183. doi: 10.61409/V205183

Hvis du arbejder med klinisk forskning, kan du formentlig spare tid og penge ved at bruge de store sprogmodeller (LLM) som ChatGPT eller Gemini. Det kan du i øvrigt potentielt også, hvis du skriver instrukser eller guidelines. Det er nemt at komme i gang, men der er rigtig meget at vinde ved at dygtiggøre sig i brugen af sprogmodellerne, ved f.eks. at blive god til at »prompte«/spørge og at bruge alle de tilgængelige plugins.

I en kronik i efteråret benævnte jeg områder, hvor disse kunstig intelligens (AI)-modeller kan have relevans i sundhedsvæsenet [1], og det forventes, at beslutningsstøtteværktøjer, der bygger på generativ AI, er på vej, og måske du en dag bruger dem uden at vide det. Så må vi bare håbe, at de vejleder og ikke vildleder os. Noget tyder på, at de ikke kun yder beslutningsstøtte, men også beslutningstagen [2]. Det forkromede mål må være sømløs integration, der løfter vores arbejdsdag og ikke gør det mere besværligt – som vi har set med andre teknologier.

Hvis du gerne vil dykke lidt mere ned i emnet, elaborerer *Mojadeddi & Rosenberg* emnet flot i deres statusartikel [3]. De beskriver, hvordan AI og LLM har udvist bemærkelsesværdige evner til at assistere forskere gennem alle faser af forskningsprocessen – fra idégenerering til publikation, hvilket muliggør øget effektivitet og nye muligheder i forskning. Dog påpeger de betydelige forbehold, der nødvendiggør en forsigtig og velovervejet anvendelse af teknologien.

Potentialet for AI i medicinsk forskning er betydeligt og kan på nogle måder sammenlignes med den transformation, digitaliseringen har medført. Disse værktøjer kan forbedre dataindsamling, dataanalyse og endda skrivningen af videnskabelige artikler, hvilket fremskynder processen og kan øge kvaliteten. Forestil dig LLM som en meget velkvalificeret – og billig – assistent, der kan hjælpe med dataindsamling, sparring, korrekturlæsning, skrivning mv. Og i øvrigt hjælpe dig inden for alle fagområder: juridisk, medicinsk, videnskabeligt, sprogligt mv.

I statusartiklen understreges vigtigheden af at nærme sig AI med forsigtighed. To væsentlige

udfordringer er AI's »hallucinationsevne«, hvor falsk information genereres særdeles overbevisende, samt spørgsmålet om sikker databehandling. Dette er særligt relevant i medicinsk forskning, hvor nøjagtighed og datasikkerhed er yderst vigtigt. Aktuelt hallucinerer de bedste modeller ca. 2-3%, altså ca. 98% af deres output er sandt [4], og man fristes til at spørge, om det er godt nok, og hvor meget læger konfabulerer ...

Det toneangivende tidsskrift *New England Journal of Medicine* lancerede i starten af 2024 et AI-tidsskrift (NEJM AI), hvilket markerer en milepæl inden for den medicinske AI-forskning. Med en dedikeret indsats for at udforske og integrere AI i medicinsk forskning, klinik og uddannelse sætter NEJM AI spørgsmålstegn ved, hvordan sundhedsprofessionelle bør navigere i teknologien. NEJM AI understøtter og opfordrer ligefrem forskere til brug af LLM i arbejdet, med forbehold for at forfatterne fuldt ud tager ansvar for indholdet og korrekt anerkender brugen af LLM'er [5].

I den sammenhæng vil jeg opfordre dig til at overveje, om din afdeling/forskningsgruppe bør opfordre aktivt til brug af teknologien eller i det mindste tage et standpunkt. Til trods for at alle forskellene mellem NEJM AI og Ugeskrift for Læger fristes jeg lidt retorisk til at spørge, om Ugeskrift for Læger ligeledes skal opfordre til brug af LLM'er i forskningsbidrag under lignende forudsætninger.

Det er naturligvis vigtigt ikke at kaste sig ukritisk hovedkulds ind i brugen, da der er flere faldgruber, og området i det store og hele er terra incognita. Det betyder, at den nødvendige evidens er sparsom, og de regulatoriske rammer ikke er udfoldet fyldestgørende. Der er dog ingen tvivl om, at sprogmodellerne er meget potente og kan hjælpe med flere opgaver i vores hverdag.

ChatGPT selv siger uredigeret: »Læger bør bruge ChatGPT i medicinsk forskning for at krydre dataanalysen med en smule kunstig intelligens—den perfekte opskrift på skarphed uden ekstra koffein!« Men siger samtidig, at »læger ikke bruge ChatGPT og lignende AI-modeller i medicinsk forskning, da de kan introducere fejlagtige data og forudindtagede resultater, hvilket kan kompromittere patientplejens kvalitet og forskningens integritet«.

Så kan du selv vælge, hvilket udsagn der passer dig.

Korrespondance *Andreas Pihl*. E-mail: afpihl@gmail.com

Interessekonflikter Der er anført potentielle interessekonflikter. Forfatterens ICMJE-formular er tilgængelig sammen med lederen på ugeskriftet.dk

Open Access under Creative Commons License [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)

REFERENCER

1. Pihl A. Ai bliver en større del af din hverdag, 2023. <https://ugeskriftet.dk/debat/ai-bliver-en-storre-del-af-din-hverdag> (6. maj 2024).
2. Chen S, Guevara M, Moinigi S et al. The effect of using a large language model to respond to patient

- messages. Lancet Digit Health. 2024;6(6):e379-e381. [https://doi.org/10.1016/s2589-7500\(24\)00060-8](https://doi.org/10.1016/s2589-7500(24)00060-8)
3. Mojadeddi ZM, Rosenberg J. Kunstig intelligens i forskning. Ugeskr Læger. 2024;186(16):V08230532. <https://doi.org/10.61409/V08230532>
 4. Vectara/hallucination-leaderboard. Hallucination leaderboard, 2024. <https://github.com/vectara/hallucination-leaderboard> (6. maj 2024).
 5. Koller D, Beam A, Manrai A et al. Why we support and encourage the use of large language models in NEJM AI submissions, 2023. NEJM AI, 1(1). <https://doi.org/10.1056/aie2300128>