
Fastlegen om bruk av kunstig intelligens: – Fått mer tid til å puste

REPORTASJE

CAROLINE ULVIN JOHANSSON

caroline.ulvin.johansson@tidsskriftet.no

Tidsskriftet

Flere og flere fastleger tar i bruk kunstig intelligente assistenter som transkriberer under pasientkonsultasjoner. En av disse er spesialist i allmennmedisin Benedicte Wardemann.



SKRIVEFRIE KONSULTASJONER: Fastlege Benedicte Wardemann ved Vest Helse og Trening i Bærum er en av dem som har tatt i bruk kunstig intelligens under sine pasientkonsultasjoner. Foto: Leikny Havik Skjørseth

– Hvis du vil, kan jeg spille en pasient og legge frem et problem? Så kan vi se hvordan programmet fanger opp de ulike delene?

Fastlege Benedicte Wardemann ved Vest Helse og Trening i Bærum er i ferd med å demonstrere hvordan hennes kunstig intelligente assistent fungerer. Hun trykker på knappen «ny konsultasjon» på PC-en før hun begynner å snakke:

– Hei, jeg skulle gjerne hatt hjelp til å gå ned i vekt. Jeg har prøvd på dette før og har forsøkt mange ulike dietter. Jeg klarer kanskje å gå ned fem kilo før jeg merker at det blir for vanskelig å holde det. Da sprekker jeg og går opp i vekt igjen. Dette er selvfølgelig skamfullt for meg, og jeg føler at jeg ikke mestrer noe alle andre tar for gitt.

Mens Wardemann snakker, kan man se at ordene dikteres fortløpende på PC-skjermen – via en liten, rund, svart mikrofon på skrivebordet. Navnet på programmet er Noteless, en kunstig intelligent assistent som skal hjelpe helsepersonell med dokumentasjonsjobben. Wardemann fortsetter å spille en typisk pasient, før hun deretter later som hun er legen i denne fiktive konsultasjonen.

Systematisert journalnotat

Når samtalen er ferdig, gjør fastlegen et par trykk på PC-en. Resultatet? Et utkast til et splitter ferdig journalnotat. Hun leser oppsummeringen høyt.

– Aktuelt: Ønsker hjelp til vektreduksjon. Har forsøkt ulike dietter tidligere med forbigående effekt. Klarer å gå ned ca. fem kilo før det blir for vanskelig å opprettholde vekttapet. Går deretter opp i vekt igjen. Opplever dette som psykisk belastende med følelse av skam og manglende mestring.

Programmet systematiserer altså innholdet i pasientkonsultasjonen etter punktene Aktuelt, Funn, Vurdering og Tiltak. Og notatet blir ifølge Wardemann en god del lenger enn hva hun selv ville ha skrevet.

– Jeg synes det er fascinerende at den bruker så gode setninger. Skulle jeg skrevet dette, hadde jeg brukt én eller to linjer – maks. Jeg ville nok heller ikke fått med meg at hun opplever denne skamfølelsen og manglende mestringen, som også er en del av kompleksiteten med det å gå ned i vekt.

Etter å ha sett over at alt stemmer i notatutkastet, kopierer fastlegen teksten over i pasientens faktiske journal, og sparer med dette verdifulle minutter.

KI-forskning i vinden

Det at Wardemanns KI-baserte journalnotat blir lenger, mer utfyllende og får med seg flere nyanser enn hun selv ville prioritert å skrive ned, er interessant å ha i bakhodet når man ser nærmere på forskningen som er gjort på ulike språkmodeller. I flere studier finner man nemlig at språkmodellers svar på helsespørsmål ser ut til å oppfattes som mer empatiske enn svar fra faktiske leger.

Tidsskriftet publiserte nylig en studie med tittelen Kunstig intelligens og legers svar på helsespørsmål [\(1\)](#). Her ble 192 helserelevante spørsmål med tilhørende svar fra leger hentet fra nettstedet Studenterspør.no. Deretter ble språkmodellen GPT-4 benyttet til å lage et nytt sett med svar på de samme spørsmålene. I en blindtest ble begge svarene – fra både legene og GPT-4 – vurdert av en gruppe respondenter med helsefaglig bakgrunn. Og funnene? Svar fra den kunstig intelligente språkmodellen ble oppfattet som mer empatiske, kunnskapsrike og hjelpsomme enn svar fra faktiske leger.

«Svar fra den kunstig intelligente språkmodellen ble oppfattet som mer empatiske, kunnskapsrike og hjelpsomme enn svar fra faktiske leger»

En av forskerne bak studien er Ib Jammer, som er ph.d. og anestesilege ved Haukeland universitetssjukehus. Han forteller om bakgrunnen for studien: I en av sine forelesninger la han nemlig frem en lignende studie fra USA [\(2\)](#), der man undersøkte nettopp hvordan ChatGPT besvarte helsespørsmål.

– Funnene som da ble publisert, viste at svarene fra ChatGPT ofte ble vurdert som mye bedre enn svar fra mennesker. Det syntes vi var interessant. Skal virkelig en datamaskin være bedre enn oss? Hvordan er egentlig dette? Og fungerer det også på norsk?



KI-ASSISTENT: Den lille, runde mikrofonen registrerer lyden i rommet slik at KI-programmet kan lage et utkast til journalnotat. Foto: Leikny Havik Skjærseth

Med dette som bakteppe ble altså studien til Jammer og medarbeidere til. Men til tross for at man i flere studier finner at språkmodellers svar på helsespørsmål faktisk oppfattes som mer kunnskapsrike, er det også viktig å være klar over modellenes svakheter.

Må reguleres

- Det finnes eksempler på språkmodeller som har sagt ting de ikke burde si, forteller Iege med doktorgrad i kunstig intelligens Ishita Barua.
- Jeg tror det var en mann i Belgia som tok sitt eget liv etter råd fra en språkmodell. Det er nettopp denne typen uønsket atferd man vil unngå. Dette må reguleres, og dette er grunnen til at språkmodeller er vanskelige – fordi kommunikasjon er veldig dynamisk. Dette er vanskelig å regulere per i dag, det kommer til å være vanskelig i fremtiden også. Men folk må ta innover seg at det er snakk om en språkmodell – det er ikke faktisk empati.

Barua tror språkmodeller kan være et nyttig verktøy både for leger og pasienter så lenge de blir forstått og brukt riktig. Men det finnes mange aspekter ved bruken vi må ta høyde for, og som vi kanskje heller ikke er klar over enda, sier hun.

«Det er mange mennesker som er ensomme, som ikke har noen å prate med og som finner god trøst ved å bruke språkmodeller»

- Det er mange mennesker som er ensomme, som ikke har noen å prate med og som finner god trøst ved å bruke språkmodeller. Vi klarer for eksempel ikke å imøtekomme de behovene i psykiatrien. Så gitt at dette er kommet for å bli, så må vi gjøre så godt vi kan for å sørge for at disse modellene er så trygge som mulig. Men igjen: Jeg vet ikke om det går an å gjennomregulere dette.

KI bedre enn leger – eller?

Studien publisert i Tidsskriftet føyer seg inn i rekken av mange studier som den siste tiden har blitt gjort på nettopp språkmodeller innen helsevesenet. I flere av disse finner man også at språkmodellene presterer bedre enn legene selv, forteller Barua.

- De siste seks månedene har det kommet en god del studier som tester ut språkmodeller på medisinske spørsmål og caser. Det er typisk ChatGPT man har testet, og flere av resultatene viser at den presterer bedre enn legene alene og legene som til og med bruker språkmodeller. Det siste er noe av det som har vært litt oppsiktsvekkende.

I en svensk studie [\(3\)](#) nylig publisert i BMJ Open har man derimot gjort funn som går i motsatt retning. Barua understreker at denne studien er mer sammenlignbar med norske forhold, ettersom den er gjort i vårt naboland og at den i større grad bruker kasuistikker som ligner ekte kliniske situasjoner.

- Der fant man at legene presterte bedre enn språkmodellene. Altså stikk motsatt resultat. Det må man kanskje vektlegge mye mer – at man har reelle kliniske situasjoner hvor språkmodellene ikke klarer å prestere på høyde med leger. Det er veldig interessant at vi nå har studier med sprikende resultater.

Hoppe på toget

Varierende funn til tross – at kunstig intelligens er noe som har kommet for å bli, både innen helsevesenet, men også i mange andre deler av samfunnet, er det liten tvil om. Mange har kanskje gjort seg kjent med de ulike verktøyene som finnes, mens andre har muligens enn litt mer skeptisk tilnærming. For noen kan denne kraftfulle teknologien nærmest oppleves som en trussel.

– Jeg tror ikke helsepersonell blir erstattet av KI, men vi kan bli erstattet av folk som vet hvordan vi kan nyttiggjøre oss av verktøyene som KI gir oss. Jeg pleier alltid å si at vi ikke blir overflødige, vi må bare finne oss nye nisjer vi kan jobbe med. Og det hjelper ikke å motsette oss bruken. KI kommer. Vi kan velge å være med, hoppe på toget – eller bli etterlatt på stasjonen, sier Jammer.

Barua sammenligner innføringen av dagens KI-teknologi med innføringen av internett i sin tid.

– Dette er noe som kommer til å være så gjennomgripende for alle deler av samfunnet. Det er litt som å sammenligne med hvorfor man skulle lære seg å bruke internett. Vi trenger at alle har en forståelse av hva dette grunnleggende sett er. At når du får anbefalinger fra kunstig intelligente modeller som lege, så må du stå trygt og ha den faglige integriteten som kreves. Du må forstå hva denne teknologien egentlig består av og når den er i et grenseland som gjør at du må gripe inn. Du må kunne forutse hva som kan gå galt, hvilke fallgruver som finnes, før det inntreffer.

De feltene man har kommet lengst med kunstig intelligens i helsevesenet er områder der man bruker mye bilder og mønstergjenkjenning, mener Barua. Eksempler er radiologi, kardiologi, dermatologi og gastroenterologi.

– Så er det i tillegg også mye oppmerksomhet rundt generativ KI og språkmodeller, forteller hun.

Fallgruver

Tilbake på fastlegekontoret til Benedicte Wardemann ved Vest Helse og Trening i Bærum henger det lapper rundt om i lokalet med informasjon om at fastlegene bruker kunstig intelligens under konsultasjonene. Ifølge Wardemann har ingen pasienter så langt reagert, selv om de kan se på PC-skjermen at ordene de sier dikteres forløpende.

Hun presiserer at mikrofonen ikke tar opp selve samtalen, men transkriberer det som blir sagt. Den KI-baserte løsningen er uavhengig av selve journalsystemet.

– Det er ikke naturlig å si for eksempel navnet eller fødselsnummeret til pasienten høyt når man snakker med pasienten. Derfor er den helt uavhengig av persondata.

Og når pasientkonsultasjonen er ferdig, forteller Wardemann at hun bruker omtrent ett minutt på å se over journalnotatet for å sjekke at alt stemmer.

– Man må alltid sjekke. KI-modellen er en *pleaser*. Du merker at den alltid er positiv og ikke ser noen begrensning i kunnskapsgrunnlaget.

Nettopp det, det å vite om de viktigste fallgruvene, er helt essensielt ifølge Barua.

– Jeg tror at automatiseringsbias er en av de største fallgruvene. Vi har overdreven tiltro til maskiner, så vi begynner å lene oss for mye på beslutninger og anbefalinger fra KI-modeller. Nummer to er dette med algoritmisk diskriminering. At du har KI-modeller som kanskje ikke er trent på et mangfoldig nok datagrunnlag.

«Vi har overdreven tiltro til maskiner, så vi begynner å lene oss for mye på beslutninger og anbefalinger fra KI-modeller»

Når det gjelder tidsbesparing, poengterer Jammer at man får mer ut av språkmodellene hvis man lærer seg å skrive gode instruksjoner.

– Hvis vi i vår studie ikke hadde laget noen instruks, bare presentert et spørsmål, så ville vi fått et svar det ville tatt lang tid å redigere, sier han.

– Legene er fremdeles sjefen

For fastlege Wardemann, som har en liste på 1 150 pasienter, har den kunstig intelligente teknologien gitt henne mer rom for å puste i hverdagen.

– Hvor mye tid tror du at du sparer i løpet av en snittarbeidsdag?

– Over en time. Og det har ikke gjort at jeg fyller inn dagen med flere pasienttimer – det har gjort at jeg faktisk puster. Jeg har det bedre med meg selv.

Hun tror man må tørre å være fremoverlent.

– Jeg tenker at vi må tørre å bruke ny teknologi som skal hjelpe oss i hverdagen. Jeg føler at jeg tar tilbake tiden min og at jeg får brukt mer tid på legerollen. Så mitt råd til kollegaer er at de bør prøve, være positive til ny teknologi, men selvfølgelig: være klar over at det fremdeles er vi som er sjefen.

REFERENCES

1. Mork TE, Mjøs HG, Nilsen HG et al. Kunstig intelligens og legers svar på helsespørsmål. Tidsskr Nor Legeforen 2025; 145. doi: 10.4045/tidsskr.24.0402. [PubMed][CrossRef]

2. Ayers JW, Poliak A, Dredze M et al. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Intern Med* 2023; 183: 589–96. [PubMed][CrossRef]
 3. Arvidsson R, Gunnarsson R, Entezarjou A et al. ChatGPT (GPT-4) versus doctors on complex cases of the Swedish family medicine specialist examination: an observational comparative study. *BMJ Open* 2024; 14: e086148. [PubMed][CrossRef]
-

Publisert: 5. mars 2025. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.25.0138

Opphavsrett: © Tidsskriftet 2025 Lastet ned fra tidsskriftet.no 25. desember 2025.