



Kunstig intelligens i videregående skole: en undersøkelse av kunnskap og bruk

Forfattere: Ronja Lem Hetland og Max Folstad Andersen, Ullern videregående skole

Kunstig intelligens (KI) har blitt en viktig del av samfunnet og kommer til å få større og større innflytelse i fremtiden. Tidligere forskning har sett på konsekvenser og bruksområder for KI innen utdanningssektoren. I denne studien ønsket vi å kartlegge kunnskap og hyppighet av bruk i videregående skoler i Oslo med utgangspunkt i Ullern videregående skole. Det ble brukt en poengbasert kunnskapstest med flervalgsspørsmål og spørsmål om bruk av KI. Svarene til elever og lærere ble sammenlignet. Vi fant at lærerne i gjennomsnitt hadde mer kunnskap enn elevene. I tillegg hadde både elever og lærere lite kjennskap til UDIRs regler om KI. Studien kan gi et innblikk i hyppighet av bruk og kunnskap hos elever og lærere i Oslo, som kan gi viktig kunnskap til utdanningssektoren i møte med KI. Samtidig har resultatet begrensninger grunnet størrelsen på studien og det kreves videre forskning på temaet.

Introduksjon

Kunstig intelligens (KI) har raskt utviklet seg til en globalt anvendt teknologi. Det spås at KI-utviklingen vil prege de kommende årene (Wehus, 2024). KI-løsninger, som språkmodeller, har blitt tatt i bruk flere steder i arbeidslivet, og har potensiale til å effektivisere de fleste sektorer. Måten språkmodellene, som ChatGPT og Bard, virker på, er ved å predikere det neste ordet basert på teksten den allerede har predikert og brukerens ledetekst. Predikasjonene skjer på bakgrunn av datagrunnlaget språkmodellen er trent på, og omfatter ofte millioner av artikler, chatter og tekster, fra plattformer som Wikipedia, artikler og bøker (Shreya, 2023). OpenAI sin ChatGPT ble lansert 30. november 2022, og lanseringen har ført til at teknologien er blitt gratis tilgjengelig for allmenheten. Dette har ført til diskusjon knyttet til troverdigheten av informasjonen ChatGPT gir. Det har blitt funnet at ChatGPT blant annet er mer pålitelig som inspirasjonskilde enn til å generere selvstendig innhold (Melteig, 2023; Kaste, 2024; Raj, Singh, Kumar, & Verma, 2023). Det er også stilt spørsmål ved den politiske vinklingen til ChatGPT og hvordan den kan påvirke brukerne (Motoki, Neto, & Rodrigues, 2023).

Kort tid etter lanseringen av ChatGPT begynte elever å ta i bruk teknologien. Den økte bruken har ført til utfordringer knyttet til juks med KI, samt læreres bekymringer rundt den nye teknologien (Molnes, 2024). Som respons kom Utdanningsdirektoratet (UDIR) med et totalforbud mot bruk av KI på vurderinger og eksamen før sommeren 2023. Til tross for forbudet, har KI gradvis blitt tatt i bruk i utdanningssektoren,

spesielt gjennom lanseringen av KI for Osloskolen i slutten av 2023. KI for Osloskolen er en snakkebot som bygger på ChatGPT, men som er i henhold til skolens retningslinjer for personvern (Osloskolen 2023).

Siden lanseringen av ChatGPT har det kommet noen studier knyttet til bruk av KI i skolesystemet, men det omfatter i hovedsak mindre studier, og få er norske. En relativt ny studie, som analyserte bruk av KI blant norske studenter, viste at majoriteten av studentene bruker KI (Hauge & Øygarden, 2024). En annen undersøkelse, av Enjolras og Solheim (2023), fant at av elever som går på videregående yrkesfag i Norge, så bruker 18 % av elevene KI i liten grad og 1 % av elevene bruker KI i stor grad. Her var det mulige kategoriene: ingen kjennskap, noe kjennskap, bruker i liten grad og bruker i stor grad. De fant også at av elevene som går videregående allmennfag i Norge, så bruker 23 % KI i liten grad og 1 % bruker KI i stor grad. Det er viktig at elever lærer om KI for å forberede dem på arbeidslivet og samfunnet. For å muliggjøre det må lærere ha den nødvendige kompetansen til å lære dem opp. Økt kunnskap blant lærere er også nødvendig for å forhindre juks med KI. Studien vår baserte seg på Ullern videregående skole i Norge. Hensikten var å undersøke kunnskap om og hyppighet av bruk av KI hos elever og lærere på videregående skoler i Oslo Vest. En poengbasert kunnskapstest med flervalgsspørsmål og spørsmål om bruk av KI ble brukt for å innhente data. Forventede funn var god kjennskap til regler, mer bruk blant elever enn lærere, og en korrelasjon mellom bruk og kunnskap hos begge gruppene.

Metode

Hensikten med forsøket var å undersøke bruk av og kunnskap om KI til lærere og elever på videregående skoler i Oslo Vest. En Microsoft Forms ble sendt ut 8. januar 2024, til kontaktlærere som delte undersøkelsen med tilhørende elevgrupper på alle tre trinn i kontaktlærertid. Undersøkelsen bestod av 14 flervalgsspørsmål totalt (vedlegg 1). Av disse var det fem ja/nei spørsmål, åtte spørsmål med fire alternativer, og ett spørsmål om hyppighet av bruk, delt inn i seks kategorier. De 13 første spørsmålene ble brukt for å kartlegge kunnskap og ga derfor ett poeng hver. Spørsmålet om bruk ga ikke poeng. For lærerne ble det sendt ut en tilsvarende undersøkelse 5. januar, som hadde to tillegsspørsmål om bruk i undervisning. De hadde fire kategorier hver, og ga ikke poeng. Begge undersøkelsene var anonyme, og hver bruker kunne svare én gang.

450 elever og 36 lærere svarte på undersøkelsen. Spørsmålet «Hvordan burde man IKKE stille spørsmål til ChatGPT» med svaralternativene «Stille detaljerte spørsmål», «Stille kortest mulige spørsmål», «Gi bakgrunnsinformasjon med spørsmålet» og «Stille lengst mulig spørsmål», ble fjernet grunnet upresis formulering som skapte uenighet om korrekt svar. Datanalyse ble gjort med Python 3.12.0 (amd64, Windows 10), kode vedlagt (vedlegg 3). I dataanalysen ble kategoriene til spørsmålet: «Til slutt en undersøkelse: Hvor mye har du brukt KI?» gjort om til tall fra 0-5. Der 0 vil si ingen bruk og 5 vil si daglig bruk. En tosidig t-test ble utført mellom poengsummene for elever og lærere. Resultatene fra spørsmål om bruk var ikke normalfordelt, så en Mann-Whitney U test ble utført mellom hyppighet av bruk for elever og lærere. P-verdiene ble sammenlignet med et signifikansnivå på 0,05.

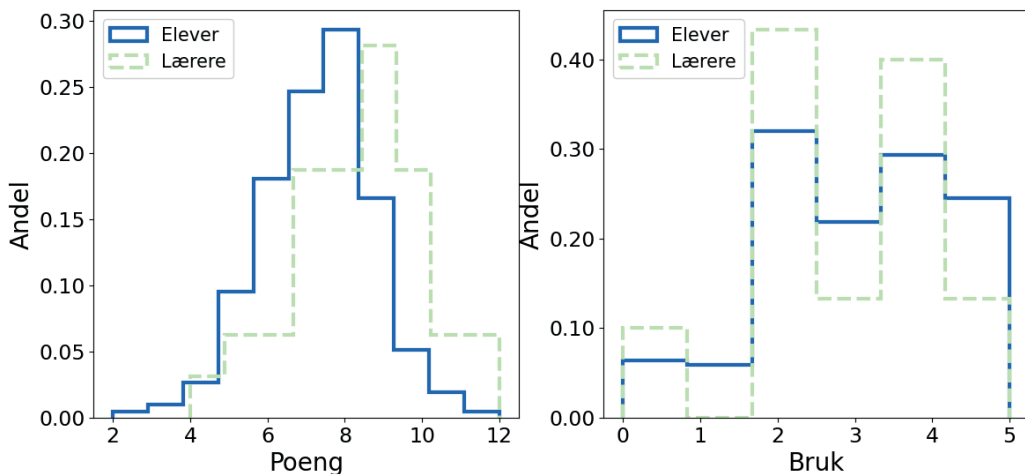
Resultater

Resultatene viste at gjennomsnitt med standardavvik for elever og lærere var henholdsvis 7,33 (1,60) og 8,42 (Det var en 9,1 % forskjell i poeng mellom elever og lærere (T-test, $p=0,00013$) (figur 2). En t-test kunne benyttes ettersom dataen var normalfordelt (figur 1).

Det ble også funnet at kjennskapen til skolens regler for bruk av KI på vurdering var lav både hos elever og lærere, med henholdsvis 17,3 % og 11,4 % korrekt (tabell 1).

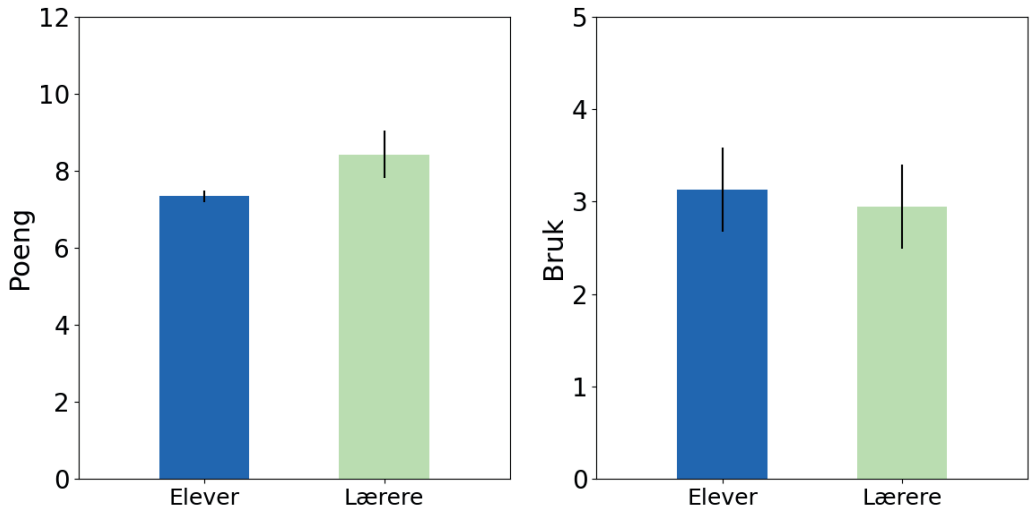
Tabell 1. Spørsmålene i undersøkelsen med andel elever og lærere som har svart galt og riktig.

Spørsmål	Elever %		Lærere %	
	Feil	Korrekt	Feil	Korrekt
Hva står KI for?	1,1	98,9	0,0	100,0
Hva er ChatGPT best på?	21,8	78,2	16,7	83,3
Hva er skolens regler for bruk av KI på vurderinger?	82,7	17,3	88,6	11,4
Er ChatGPT politisk nøytral?	79,7	20,3	47,2	52,8
Har ChatGPT følelser?	4,7	95,3	2,8	97,2
Hva er fordelen med "KI for Osloskolen"?	62,8	37,2	16,7	83,3
Kan tidligere spørsmål påvirke svarene fra ChatGPT?	53,8	46,2	63,9	36,1
Hvilket spørsmål vil ChatGPT svare MINST korrekt på?	59,5	40,5	40,0	60,0
Er bildet KI generert?	5,6	94,4	0,0	100,0
Kan ChatGPT svare bra på: Hvor mange elever går det på Ullern videregående skole?	13,0	87,0	25,0	75,0
Kan ChatGPT svare bra på: Hvem er president i USA i 2023?	65,0	35,0	47,2	52,8
Kan ChatGPT svare bra på: Hvordan kan jeg strukturere en fagartikkel om demokratiet i Norge?	12,8	87,2	8,3	91,7



Figur 1. Histogram av fordelingen av total poengsum på kunnskapsspørsmål (0-12) og fordeling av hyppighet i bruk (0-5) for elever og lærere. For bruk tilsvarer 0 ingen bruk og 5 ukentlig bruk.

Det ble ikke funnet noen statistisk signifikant forskjell i bruk mellom elever og lærere (Mann-Whitney U test, $p=0,44$) (figur 2 og tabell 2). Det var en forskjell på 3,7 % i bruk mellom elever og lærere med gjennomsnitt henholdsvis 3,16 (1,39) og 2,94 (1,39). Verdien tre for bruk tilsvarer månedlig bruk.



Figur 2. Gjennomsnittlig poeng på kunnskapsspørsmål (0-12) og hyppighet av bruk (0-5) for elever og lærere med 95 % konfidensintervall. For bruk tilsvare 0 ingen bruk og 5 ukentlig bruk.

Tabell 2. Andel svar i hver kategori til spørsmålet: «Til slutt en undersøkelse: Hvor mye har du brukt KI?» for elever og lærere.

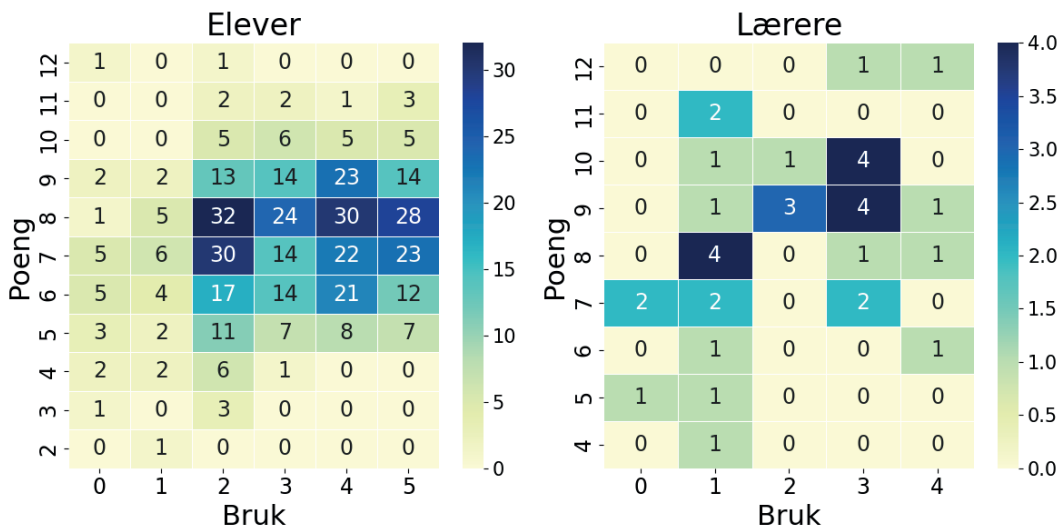
Gruppe	Kategori %					
	[0] Ingenting	[1] Har testet en gang	[2] Har testet litt	[3] Har brukt det månedlig	[4] Har brukt det ukentlig	[5] Har brukt det flere ganger i uken
Lærere	8,3	0,0	36,1	11,1	33,3	11,1
Elever	4,5	4,9	26,9	18,4	24,7	20,6

Rundt halvparten av lærerne (52,8 %) bruker KI i undervisningsplanleggingen og noe færre (30,6 %) bruker det i selve undervisningen (tabell 3).

Tabell 3. Hvor ofte lærere bruker KI i undervisning og undervisningsplanlegging.

Spørsmål	Kategorier %			
	Null ganger i uka	Én gang i uka	Flere ganger i uka	Daglig
Hvor ofte bruker du KI i undervisning?	63,9	30,6	5,6	0,0
Hvor ofte bruker du KI i undervisningsplanleggingen?	47,2	52,8	0,0	0,0

Det ble ikke funnet noen korrelasjon mellom hyppighet av bruk og kunnskap (poeng), se figur 3.



Figur 3. Heatmap som viser sammenheng mellom hyppighet av bruk (0-5) og poeng (0-12) på kunnskapsspørsmål for elever og lærere. Kategorier der det ikke var noen tilfeller er ikke vist. Dette vil si kategori fem om bruk for lærere, samt kategori 0-3 om poeng for lærere og kategori 0-1 om poeng for elever.

Diskusjon

I undersøkelsen ble bruken og kunnskapen til elever og lærere på Ullern videregående skole undersøkt. Forskjellen i poeng på kunnskapsspørsmålene mellom elever og lærere var på 9,1 % som tilsvarer rundt ett poeng, altså en liten effekt til tross for at den er statistisk signifikant. I tillegg vil kunnskap om KI trolig endre seg raskt. Forskjellen kan skyldes at lærerne har vært på seminar med kurs om kunstig intelligens og hvordan det kan brukes i skolen.

Et interessant funn var at få elever og lærere hadde kjennskap til skolens regler for bruk av KI på vurderinger. 17,4 % av elevene og 11,4 % av lærerne svarte korrekt. Kunnskapshullet kan føre til at elever blir tatt for juks og får nedsatt orden. Det er viktig at det settes inn tiltak for å informere elever og lærere for å unngå negative konsekvenser ved eksamen.

Resultatene viste ingen statistisk signifikant forskjell i bruk mellom elever og lærere. Det kan skyldes at få lærere svarte på undersøkelsen. Det kan også skyldes at kunnskapsgapet mellom de to gruppene ikke er så stort. Det er viktig at lærerne har tilstrekkelig kunnskap om KI for å kunne avsløre juks og lære elever hvordan verktøy brukes riktig.

Det ble ikke funnet noen klar korrelasjon mellom hyppighet av bruk og kunnskap, vist i figur 3. Dette kan skyldes at det var få kategorier for bruk og at økt bruk ikke nødvendigvis fører til økt kunnskap. Det er tendenser til en korrelasjon mellom hyppighet av bruk og kunnskap for lærerne, så et større datagrunnlag vil kunne gi en sammenheng. Det ble ikke stilt spørsmål om hvordan deltagerne bruker kunstig intelligens eller hva det brukes til. Det ble heller ikke skilt mellom god og dårlig bruk. Selv om man bruker KI ofte, betyr ikke det at man bruker teknologien på en god måte. For eksempel er det ikke sikkert at hyppig bruk av KI til juksing vil bidra til økt kunnskap.

Det var noen svakheter ved måten undersøkelsen ble gjennomført. At flervalgsspørsmål ble brukt kunne også hatt en innvirkning med at noen deltakere kan være flinkere til å eliminere alternativer. I tillegg var

det kun tolv spørsmål som kun gir et overfladisk inntrykk av kunnskap. Spørsmål tar kun utgangspunkt i noen få KI-modeller med hovedvekt på ChatGPT så det viser ikke breddekunnskap. Svarene ble automatisk vist ved fullført undersøkelse, som gjorde det mulig å dele riktig svar med andre deltagere. Dermed kunne elver og lærere ha kommunisert underveis. En mulig forbedring er å ta testen i et klasserom, i stillhet, der en person følger med.

Resultatene viser at 30,6 % av lærerne bruker KI i undervisning ukentlig, og at majoriteten bruker KI i undervisningsplanleggingen. Kategoriene tar ikke i betraktning månedlig bruk, som kan gjøre at resultatene ikke viser den fulle bredden, men gir et inntrykk av hyppig og sjeldnere bruk av KI i undervisning og undervisningsplanlegging. Hvor hyppig bruken av KI er i undervisningen kan ha en påvirkning på kunnskapen og bruken til elevene. Det kan føre til forskjeller mellom klasser som kan gi ulikt utgangspunkt i møte med KI i framtiden.

Overførbarheten til studien er trolig begrenset til Oslo. Ettersom det er fritt skolevalg i Oslo og KI for osloskolen er tilgjengelig for elever i offentlig skole, vil studiens funn knyttet til elever trolig ha en overføringsverdi til videregående skoler i resten av byen, i hvert fall i Oslo vest. Når det gjelder resultatene til lærerne gjør lite datagrunnlag og kursing i KI at overføringsverdien kan være begrenset.

Kunnskap om KI kommer til å bli viktig i framtiden og det er viktig for at skolen sørger for at elevene har den nødvendige kunnskapen. Det er ennå lite kunnskap om hvordan KI har blitt en del av skolehverdagen, men dette endrer seg raskt. Videre forskning er nødvendig for å kartlegge bruk av og kunnskap om KI i norske skoler, som er viktig bakgrunnsinformasjon for en eventuell integrering av teknologien i utdanningssektoren. Denne studien kan gi et innblikk i den nåværende situasjonen, men det kan være interessant å gjenta om noen år. Fremtidig forskning kan også se på korrelasjon mellom mengde opplæring lærere har fått og kunnskapen til elever, eller forskjeller mellom kjønn.

Referanser

- Enjolras, B. & Solheim, Ø. B. (2023). Notat om nordmenns bruk av kunstig intelligens. Senter for forskning på sivilsamfunn og frivillig sektor
- Hauge, M. S. & Øygarden, K. F. (2024). Studiebarometeret 2023 – Hovedtendenser. (4-2024). NOKUT. <https://www.nokut.no/studiebarometeret/analyser>
- Kaste, A. M. (2024, 1. januar). – ChatGPT passer godt til å lage struktur, men ikke innhold, mener forsker. Forskning. <https://www.forskning.no/chatgpt-kunstig-intelligens/chatgpt-passer-godt-til-a-lage-struktur-men-ikke-innhold-mener-forsker/2296347>
- Melteig, E. (2023, 5. juni). Slik gjør ChatGPT forskerne mer effektive. Titan.uio. <https://www.titan.uio.no/teknologi/2023/slik-gjor-chatgpt-forskerne-mer-effektive.html>
- Molnes, G. (2024, 17. januar). Seks av ti lærere har tatt elever i KI-juks. Utdanningsnytt. <https://www.utdanningsnytt.no/chatgpt-juks-kunstig-intelligens/seks-av-ti-laerere-har-tatt-elever-i-ki-juks/386624>
- Motoki, F., Neto, V. P. & Rodrigues, V. (2023). More human than human: measuring ChatGPT political bias. Public Choice. <https://doi.org/10.1007/s11127-023-01097-2>
- Osloskolen. (2024, 11. mars). Kunstig intelligens (KI) i Osloskolen. Osloskolen. <https://aktuelt.osloskolen.no/larerik-bruk-av-laringsteknologi/digital-skolehverdag/kunstig-intelligens-ki-i-osloskolen/>
- Raj, R., Singh, A., Kumar, V., & Verma, P. (2023). Analyzing the potential benefits and use cases of ChatGPT as a tool for improving the efficiency and effectiveness of business operations. Bench-Council Transactions on Benchmarks, Standards and Evaluations, 3(3). <https://doi.org/10.1016/j.tbench.2023.100140>
- Shreya, J. (2023, 6. juni). The Making of ChatGPT: From Data to Dialogue. Harvard. <https://sitn.hms.harvard.edu/flash/2023/the-making-of-chatgpt-from-data-to-dialogue/>

Wehus, W. N. (2024, 14. januar). Forskere spår 10 måter kunstig intelligens vil prege de neste årene:
– Jeg tror ikke eksamen vil finnes om fem år. Forskning. <https://www.forskning.no/arbeid-kunstig-intelligens-partner/forskere-spar-10-mater-kunstig-intelligens-vil-prege-de-neste-arene-jeg-tror-ikke-eksamen-vil-finnes-om-fem-ar/2308453>

