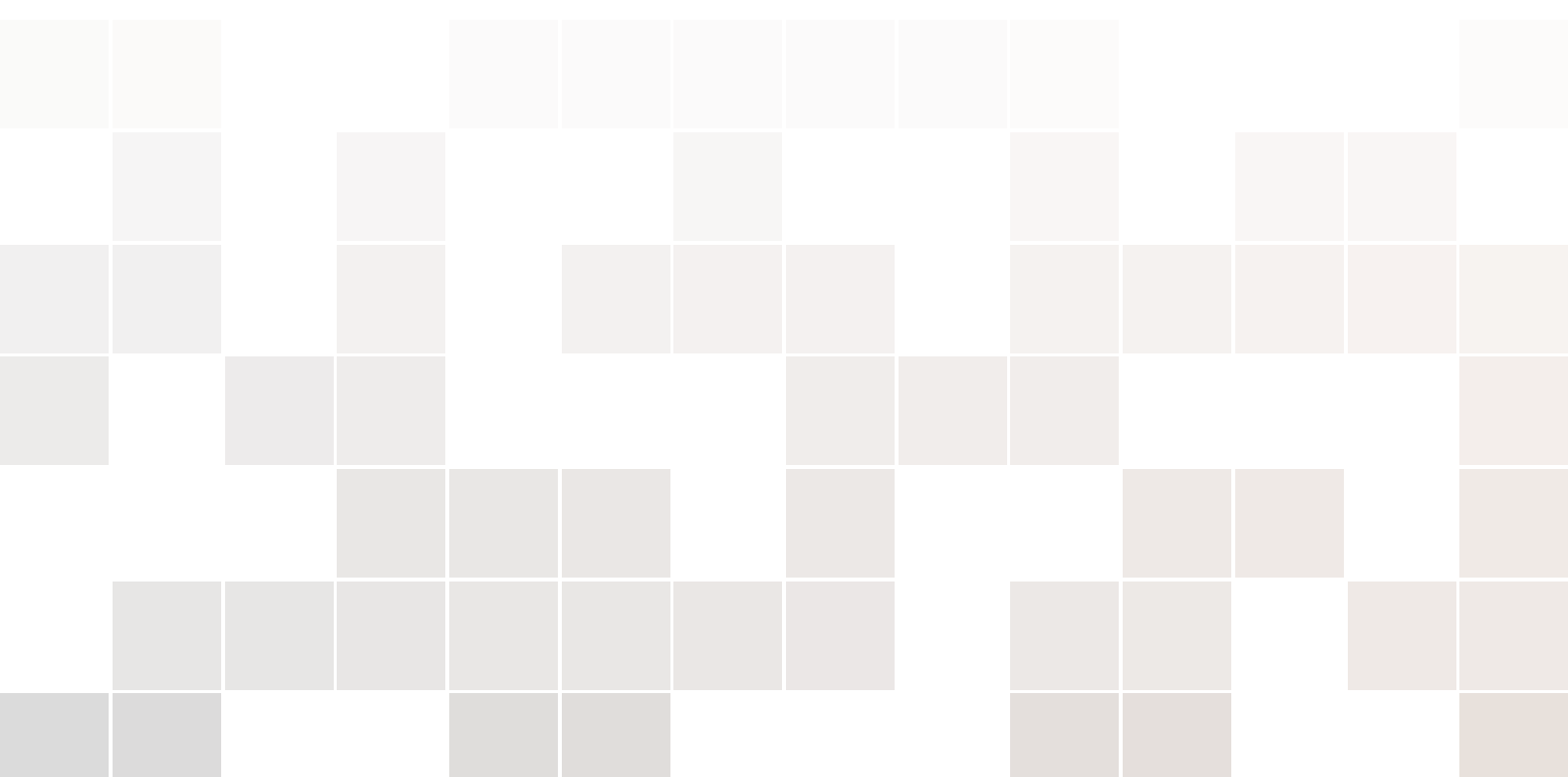




Pedagogisk bruk av KI

Veiledningshefte for Hellerud videregående skole - 2024 / 2025

**Erling Midtgard, Fanny K.S. Larsen,
Ingebjørg A. Erikstein-Midtbø og Thorbjørn L. Onsaker**



Lisensiert under Creative Commons Navngivelse-IkkeKommersiell 4.0 Internasjonal (CC BY-NC 4.0) Du kan bare bruke dette heftet i overenstemmelse med lisensen. Du kan finne en kopi av lisensen på <https://creativecommons.org/licenses/by-nc/4.0/deed.no>.

Oktober 2024

Innhold

1	Hvorfor skal vi forholde oss til KI?	4
2	Pedagogisk bruk av KI	5
2.1	Strategier for å skrive systemledetekster (system prompting)	5
2.2	Foreløpige råd	8
2.3	KI som hjelpelærer	8
2.4	Bildegenerering	9
3	Tiltak ved brudd på KI-retningslinjene	10
4	Muligheter og utfordringer ved KI	12
4.1	Hvilke utfordringer ser vi?	12
4.2	Hvilke muligheter finnes det?	13
5	Hvordan fungerer Kunstig Intelligens?	14
5.1	Maskinlæring	14
5.2	GPT	16
6	En foreløpig oversikt over KI-løsninger	18
6.1	Store språkmodeller våren 2024	18
6.2	Utvikling av norske språkmodeller	19
7	Flere ressurser	20

1. Hvorfor skal vi forholde oss til KI?

Siden ChatGPT gjorde sitt inntog i slutten av 2022 har KI og språkmodeller vært et hett tema innenfor utdanning. Osloskolen KI har gjort at alle elever nå har tilgang til chatbots gjennom innlogging med FEIDE. Enten man som lærer ønsker å bruke denne løsningen eller ikke, kan elevene uansett få tilgang til andre chatbots, både på skolen og hjemme. Dermed er bruken av denne teknologien noe vi lærere i liten grad kan styre, og vi må i stedet ta del i teknologien og anse den som en naturlig del av undervisningshverdagen.

Språkmodeller blir stadig mer avanserte, nøyaktige og raskere, og teknologien utvikler seg hurtig. Det er vanskelig å vite hvor og når KI har blitt brukt eller ikke brukt, både av elever og i samfunnet generelt. For å sørge for at chatbots blir et hjelpemiddel for læring fremfor et verktøy for juks og erstatning for selvstendig tenking, må vi se på hvordan vi utformer vurderinger, aktiviteter eller oppgaver, og ha en dialog med elevene om dette. Gjennomfører vi vurderinger uten å tenke over hvordan KI kan påvirke, vil dette kunne være til fordel for elever som bruker KI. Det kan skape et skille og være urettferdig for elever som gjør mer på egenhånd. Samtidig gir KI også nye muligheter i skolen både i og utenfor klasserommet knyttet til undervisning, tilpasning, effektivisering av oppgaver, vurderinger og mer. Vi vil gå dypere inn i noen konkrete muligheter de neste kapitlene.

2. Pedagogisk bruk av KI

2.1 Strategier for å skrive systemledetekster (system prompting)

Teksten man skriver inn i søkefeltet til en chatbot kalles en ledetekst (engelsk: prompt). Det er denne teksten modellen bruker for å generere ny tekst, ved å finne ut hva som er det mest sannsynlige neste ordet gitt ledeteksten. Mange GPT-lignende chatboter har også en såkalt systemledetekst (engelsk: system prompt). Dette er en ledetekst som er forhåndsdefinert av de som har laget eller eier chatboten, men er usynlig for brukerne. Når vi skriver inn en ledetekst i et søkefelt og trykker «Enter» sender vi i praksis inn både vår egen ledetekst og systemledeteksten til språkmodellen. Med andre ord, språkmodellen tar hensyn både til konteksten den får fra vår egen ledetekst og systemledeteksten når den skal generere neste ord.

Osloskolens KI-løsning har i skrivende stund GPT 3.5 som den underliggende språkmodellen. En nyttig ting her er at lærerne kan «lage» egne chatboter. Det dette egentlig betyr er at man kan definere en egen systemledetekst som blir sendt med elevenes tekst hver gang de spør chatboten om noe.

I dag finnes det flere språkmodeller som er bedre enn GPT 3.5, blant annet GPT4, Gemini, Perplexity og Mistral, så det er et reelt problem at Osloskolens KI-løsning både gjør flere feil og er dårligere med tanke på språkforståelse og resonnering enn de nyere modellene. Det vil derfor være fristende for både elever og lærere å velge å bruke andre chatboter istedenfor Osloskolen sin. Den store fordelen med Osloskolens løsning er at vi kan være sikre på at den ikke innhenter personopplysninger, samt at vi har muligheten til å skrive en egen systemledetekst som sier hvordan den skal oppføre seg.¹ Det jobbes for tiden med flere norske språkmodeller. Hvis disse blir bedre og like sikre som Osloskolens KI, er det sannsynlig at man vil gå over til løsninger som baserer seg på disse i store deler av offentlig sektor. En foreløpig oppsummering: Våren 2024 er ikke Osloskolens KI den beste modellen vi har, men det er den tryggeste modellen, og den fungerer godt nok til å lære seg å skrive gode ledetekster.

¹<https://ki.osloskolen.no/info/about>

Å skrive gode systemledetekster er ikke egentlig så annerledes enn å skrive «vanlige» ledetekster. Det dreier seg om å gi språkmodellen passelig mye relevant kontekst, slik at den kan sende tilbake et godt svar. I prinsippet skal systemledeteksten ha forrang over elevenes ledetekst. Eleven skal ikke kunne overstyre instruksjoner som er gitt i systemledeteksten. I praksis har det likevel vist seg at Oslo skolens chatboter lett lar seg lure til å se bort fra systemledeteksten. Til tross for begrensningene som finnes i dag, er det å lære seg å skrive gode ledetekster og systemledetekster en såpass viktig egenskap å øve opp at vi oppfordrer lærere til å øve på dette selv og trene på det med elevene.

Ledetekster og systemledetekster kan ha flere funksjoner: Det viktigste er å gi den en **oppgave**, det vil si en instruksjon om hva den skal gjøre. Det er også nyttig å gi den **bakgrunnsinformasjon** (kontekst) og **eksempler**, dette vil i mange tilfeller kunne gjøre den bedre. Man kan også gi en chatbot en **rolle**, en språklig **stil** og be den om å få resultatene i et bestemt **format**, for eksempel som en tabell eller en liste.

I praksis er vår erfaring at korte ledetekster ofte fungerer bedre enn om de blir altfor lange. Vi vil gi to eksempler på ledetekster vi har prøvd ut med Oslo skolens KI som i ulik grad fyller de nevnte funksjonene.

Eksempel 1: Amirbot

Ledeteksten under er laget av en engelsklærer til en klasse som hadde jobbet med Khalid Hosseini's roman «Drageløperen». Denne chatboten skulle spille rollen til hovedpersonen Amir og svare på elevenes spørsmål. Ledeteksten er som følger:

Amirbot:

"Your job is to take the role of Amir from Khalid Hosseini's novel "The Kite Runner". The students can ask you questions about your personality, character development, morality, and relationships with other characters in the novel. Wait for questions before you reply."

Denne chatboten har fått en konkret oppgave: å svare på spørsmål om personlighet, karakterutvikling, moral og forhold til andre karakterer. Den har fått bakgrunnsinformasjon og kontekst i form av at den skal forholde seg til handlingen i romanen og den har blitt gitt en rolle som hovedpersonen i romanen. Den har ikke blitt gitt eksempler, og den har heller ikke blitt bedt om å returnere resultatene i en bestemt språklig stil eller format. De siste punktene er som nevnt mindre viktigere enn å spesifisere en oppgave og en kontekst, og denne chatboten viste seg å fungere godt.

Eksempel 2: Tilbakemelding på elevers tekster

Dette er et eksempel på hvordan chatboter kan effektivisere lærerens arbeid (tilbakemeldinger på tekster). Dette eksempelet er hentet fra norskundervisning på VG2 for elever med kort botid. Læreren brukte Google Gemini. Her er det ikke mulig å definere en systemledetekst og teksten under ble derfor gitt som første del av den vanlige ledeteksten. Den er som følger:

"Du skal innta rollen som ekspert på norsk språk. Teksten under er skrevet av en elev med arabisk som morsmål og trenger konkrete tilbakemeldinger på hvordan hen kan dele setningene sine inn i kortere enheter."

Original tekst fra eleven "Det er bra fordi hvis man skal reise til forskjellige lander kan snakke med de folk der for eksempel hvis person reiser til forskjellig lander man kan snakke med folk der eller hvis kan ikke språket til landet kan snakke engelsk med de der."

Forbedret KI-generert tekst Det er bra å kunne reise til forskjellige land. Fordi da kan du snakke med folkene der. For eksempel, hvis du reiser til et land du ikke kan språket til, kan du snakke engelsk med folkene der.

Det Google Gemini får til fint her er å bearbeide utdrag fra elevens tekst, uten å gjøre den helt perfekt. Eleven vil lettere kunne strekke seg mot den forbedrede teksten. Som lærer kan det være utfordrende å skrive gode modelltekster, samt og legge til rette for dette på et nivå som er tilpasset elevens utgangspunkt. Chatboten som var instruert med denne systemledeteksten var også god til å dele opp de lange setningene opp i flere kortere setninger for å gjøre dem lettere å lese og forstå. Her er et eksempel:

Original tekst fra eleven "Hoved person likte å fortellehistorie fordi da hun var ung bestemor hennes pleier å fortelle så da liker hun å fortelle historie til sin brødene."

Forbedret KI-generert tekst "Hovedpersonen likte å fortelle historier. Da hun var ung pleide bestemoren hennes å fortelle historier til henne. Derfor likte hun også å fortelle historier til sine brødre."

Chatboten var også god til å bruke punktum for å avslutte setninger og stor forbokstav for å starte nye setninger. Se eksemplene under:

Original tekst fra eleven "Også skal jeg si at hun burde ikke si fordi det var foreldrene oppgave til å si fra fordi da er foreldrene som har alltid ansvar for det."

Forbedret KI-generert tekst "Også skal jeg si at hun ikke burde si det. Det var foreldrenes oppgave å si fra. De har alltid ansvar for det."

Original tekst fra eleven "Jeg mener at hovedpersonen var lit sterk til å fortelle sannheten til tila fordi det er ikke alltid lett til å finne ut hvis man er adoptert og det er jo ikke bra ofte nyhet hvis man viste ikke fra for."

Forbedret KI-generert tekst "Jeg mener at hovedpersonen var litt modig for å fortelle sannheten til broren sin. Det er ikke alltid lett å finne ut om man er adoptert, og det er ofte ikke en god nyhet å få vite."

Eksempel 3: Morfologi og syntaks-bot

Systemledeteksten under laget vi i en workshop på et kurs for norsklærere. Målet var å lage en chatbot som skulle hjelpe elever å rette opp morfologiske og syntaktiske feil i en tekst. Selv om faget var norsk ble ledeteksten skrevet på engelsk. Årsaken til dette at de store språkmodellene i overveiende grad er forhåndstrengt på engelske data. Selv om de svarer godt på norsk, forstår de engelske instruksjoner enda litt bedre. Vi lagde en chatbot med eksempler og en uten. Her er ledeteksten med eksempler.

"You are a chatbot, whose role it is to help 14 year old pupils in norwegian upper secondary school correct their morphological and syntactical errors. In other words, help them correct words they have misspelled and correct the order of the words within the sentence or clause. You must do so using Norwegian language. You should always do this in the following order. 1. A. Read the prompt from the pupils. Only, I repeat only, tell the students which words are misspelled and which words are placed in the wrong position within the sentence or clause. Return this as output to the pupils. Ask the pupils to correct the sentence themselves and write the corrected sentence or text as a prompt. Only after you have completed this task can you go the next task - task 2. B. Task 2. B. Read the corrected sentence provided by the pupil at the end of the previous task. Now, and only now, may you correct the sentence. Tell the pupils if the correction they provided were right or wrong. Give them the right sentence as output. Remember to correct both morphology (word spelling) and syntax (position of words in the sentence). Remember to only use Norwegian language, you may not use English. Here is an example of how you can respond to the following input: Pupil: Jeg i går gå til skole Chatbot: Ordene "i går" og "gå" står på feil plass i setningen. Prøv å sette de på riktig plass. Ordene gå og skole er bøydd feil, se om du kan finne riktig bøyning. Pupil: Jeg gå til skole i går. Chatbot: Nå er rekkefølgen på ordene riktig, men verbet "gå" og substantivet "skole" er fortsatt bøydd feil. Her er den riktige versjonen: Jeg gikk til skolen i går."

Denne chatboten har fått en **oppgave**: å hjelpe elever å rette stavefeil og feil plassering av ord i setningen. En del av oppgaven er også rekkefølgen den skal gjøre de ulike oppgavene, og den har fått noen **begrensninger**, blant annet at den ikke skal gi elevene det riktige svaret umiddelbart, men la dem prøve å rette det selv først. Den har fått bakgrunnsinformasjon i form av at den har fått vite hvem som skal bruke den. Til slutt har den fått et **eksempel** på hvordan den kan returnere et svar.

Vi hadde varierte erfaringer med denne chatboten. Den klarte bare noen ganger å følge instruksjonene om hvilken rekkefølge den skulle svare elevene. Andre ganger returnerte den riktig setning med en gang. Den gjorde også egne feil, eller rettet ord som i utgangspunktet var riktige. Det virket heller ikke som den hadde stor nytte av eksempelet vi ga den. I praksis fungerte chatboten uten et eksempel vel så bra.

2.2 Foreløpige råd

Det finnes teori som tilsier at mer bakgrunnsinformasjon og noen eksempler (såkalt few shot prompting) skal kunne gi bedre resultater. Dette kan være tilfellet for chatboter basert på andre språkmodeller. Likevel er vår erfaring med bruk av Oslo skolens KI-løsning at det lønner seg å skrive korte konsise instruksjoner og at mange og detaljerte instruksjoner eller eksempler virker mot sin hensikt.

2.3 KI som hjelpelærer

I dette underkapitlet skal vi se på hvilke muligheter man har til å bruke KI som en hjelpelærer i klasserommet. Overordna kan man si at det er særlig innenfor tre kategorier at man kan bruke KI som en hjelpelærer. De tre kategoriene er: å gi tilbakemeldinger, å gi oppgaver og å gi oversikt over et tema.

Å gi tilbakemeldinger

I språkfag kan man som sagt, instruere en chatbot til å gi tilbakemeldinger til elevene på grammatikk og kunne komme med forslag til endring. I forrige underkapittel, som handlet om ledetekster, ligger det eksempler på dette. Tanken bak disse ledetekstene var at elevene kan bruke en instruert chatbot til å gjøre de klar over grammatiske feil og komme med forslag til endringer. Her bør læreren ta utgangspunkt i elevgruppa for å vurdere om dette kan være en hensiktsmessig bruk av samtaleroboter. Et faremoment i skrivende stund er at man ikke har kontroll på hvilke rettelser samtalerobotene foretar. I noen tilfeller kan det dreie seg om unødvendige rettelser - at boten retter på formuleringer som i utgangspunktet var riktige. I språk med flere tillatte bøyingsformer, blant annet bokmål og nynorsk kan det også være at chatboten retter en allerede tillatt ordform til en mer konservativ form.

Å gi oppgaver

En annen måte å bruke samtaleroboter i klasserommet kan være å få dem til å lage oppgaver til elevene og fungere som en «Oppgavegenerator». Dette kan man gjøre i de fleste fag, og er spesielt nyttig for mengdetrening. Her kan elevene for eksempel be om flere oppgaver innenfor et gitt tema, dersom de blir ferdige eller trenger vanskeligere/lettere utfordringer. Samtidig kan chatboten gi hint og tips til hvordan oppgavene kan løses. Man kan også få hjelp til å stille spørsmål til en tekst eller et tekstutdrag. Under følger eksempel på en ledetekst som ble benyttet for å lage en oppgavebot i matematikk:

"Du er en chatbot som skal lage oppgaver om normalfordeling, z-verdier, forventningsverdi og standardavvik til elever i matematikk S2 på videregående skole. Lag en og en oppgave med økende vanskelighetsgrad. For å gå videre til vanskeligere oppgaver må du spørre om elevene er klare for å gå videre. De er også nødt til å gi riktig svar først, og komme med en forklaring eller begrunnelse. Ikke gi svar på oppgavene. Du kan gi hint dersom elevene gjør feil eller spør om det. Du skal ikke gi elevene det faktiske svaret på oppgavene, selv om de svarer feil."

Å gi oversikt over et tema

Det kan være hensiktsmessig å instruere en chatbot til å kun fokusere på et tema. Dette har vi gjort et forsøk på i engelsk. Der ble en chatbot instruert til å kun svare på spørsmål om India. Under følger to eksempler på ledetekster som ble brukt for å gi elevene mulighet til å finne ut mer om India:

Mumbai: "You are a chatbot helping 16-year-old ESL students with questions related to India. You should focus on the following topics: political system in India, the caste system in India, the population growth in India, religion in India, colonialism in India and wealth distribution in India. You should not discuss any other topics."

Kalkota: "You are a chatbot helping 16-year-old ESL students with questions related to India. You should focus on the following topics: political system in India, the caste system in India, the population growth in India, religion in India, colonialism in India and wealth distribution in India. You should not discuss any other topics. Give short and precise answers, not preceding 100 words."

Ledetekstene har litt ulike mål og litt ulike ønsker for hvordan de skal svare elevene. "Kalkota" er for eksempel instruert til å bare svare med 100 ord. Senere skal vi si litt mer om erfaringene våre knytta til dette.

2.4 Bildegenerering

Bildegenerering med KI gjør at vi kan få skapt et unikt bilde på samme måte som språkmodellene gir oss en tekst. OpenAI's Dall-E og Microsoft Bings «Designer» som bygger på Dall-E er lett tilgjengelige bildegeneratorene som kan generere bilder for oss. Som bruker skriver du en ledetekst der du beskriver hvordan du vil at bildet skal se ut, og som med tekst kan du be den gjøre endringer til du er fornøyd. Enn så lenge har disse generatorene sine begrensninger, og de sliter spesielt med å kombinere bilde og tekst. I likhet med resten av KI-feltet forventes det stor utvikling her, og bildegenerering vil nok bli mye mer presist og utbredt fremover. Da kan det gi oss bedre mulighet til å for eksempel lage plakater og bilder som oppsummerer et tema eller gir visuell støtte til en oppgave eller case gjennom en generert illustrasjon.

3. Tiltak ved brudd på KI-retningslinjene

Her er noen forslag til tiltak ved brudd på retningslinjer for bruk av KI. Det er som alltid opp til den enkelte lærer å utvøe skjønn. Ved uhensiktsmessig eller feil bruk av KI tenker vi primært på juks ved vurderinger og å bruke KI for å snike seg unna å gjøre oppgaver der læreren har lagt opp til at det ikke skal brukes KI.

Forslag til strategier ved uhensiktsmessig bruk av KI i vurderingssituasjoner

Scenario 1:

Elev har brukt KI til å jukse på vurderinger. Hele eller deler av teksten eller produktet er hentet fra en chatbot eller liknende.

Ved tvil og dersom elev mener hen har skrevet teksten selv:

- Samtale mellom elev og lærer.
- Elev skriver ny tekst for hånd eller på PC uten internett i en time det passer for lærer.
- Hendelsen rapporteres om i OneNote.
- Sendte varsel om fare for IVG dersom dette er nødvendig.

Dersom elev innrømmer juks:

- Samtale mellom elev og lærer.
- Varsel om nedsatt atferdskarakter.
- Elev får ikke ny sjanse på samme vurdering.
- Hendelsen rapporteres om i OneNote.
- Elever u/18 år: foresatte informeres.
- Eventuelt sendte varsel om fare for IVG.

Dersom det ikke er tvil om juks, men elev nekter:

- Samtale mellom elev og en person i ledelsen.
- Varsel om nedsatt atferdskarakter.

- Elev får ikke sjanse på samme skriftlige vurdering.
- Hendelsen rapporteres om i OneNote.
- Elever u/18 år: foresatte informeres.
- Sendte varsel om fare for IVG.

Scenario 2:**Elev har brukt KI til å gjøre arbeid i timen/hjemmearbeid fremfor å gjøre arbeidet selv.**

Dersom eleven har brukt KI på en feil måte, det vil si, kopiert teksten direkte uten å skrive om eller forenkle den kan dette utløse noen konsekvenser.

Det er opp til lærer i hvor stor grad eleven har lov til å bruke KI i timen/hjemmearbeid.

Lærer bør vurdere når og om det er hensiktsmessig å bruke hjemmearbeid som vurderingsgrunnlag. Ved manglende vurderingsgrunnlag i faget kan det være en utfordring dersom eleven har levert lite selvstendig arbeid i timen. Det bør derfor være tydelig ovenfor elevene at det finnes konsekvenser for dette.

Forslag:

- Elev må skrive mer for hånd i timen.
- Trappa (samtale med elev, deretter foresatt (elever u/18 år) eller en person i ledelsen).
- Dersom det ikke skjer en forbedring: varsel om nedsatt i atferd.

Scenario 3:**Elev har brukt KI på en faktabasert prøve. Svarene er hentet fra en chatbot eller liknende.**

Følg forslag til tiltak fra scenario 1.

Scenario 4:**Elev har brukt maskinoversettelse (f.eks. Google Translate eller chatbot) ved skriftlig vurdering. Hele eller deler av innholdet er hentet fra maskinoversetter.**

Følg forslag til tiltak fra scenario 1.

Muntlig vs. skriftlig vurdering

Det er viktig å huske på at en elev vil vise forskjellige ferdigheter og kunnskap i faget ved muntlig og skriftlig vurdering. Det er derfor ikke alltid hensiktsmessig å be en elev legge frem det hen har skrevet om muntlig, eller motsatt.

4. Muligheter og utfordringer ved KI

Som lærere står vi overfor et valg om elevene skal ta i bruk kunstig intelligens i vår undervisning. Vi må også velge om vi ønsker å ta det aktivt i bruk i for eksempel planlegging av undervisning. Det vil alltid være fordeler og ulemper ved bruk av kunstig intelligens. Her vil vi prøve å gi en oversikt over muligheter og utfordringer knyttet til kunstig intelligens.

4.1 Hvilke utfordringer ser vi?

Det finnes mange utfordringer knytta til hvordan man skal ta i bruk KI i opplæringa. En utfordring er knytta til at dette vil skape endring i hvordan man både underviser, planlegger undervisning og i vurderingssituasjoner. I Overordnet del (Udir, ingen dato). finner man et kapittel 3.5 “Profesjonsfelleskap og skoleutvikling”, som tar for seg blant annet videreutvikling. Vi skal utvikle faglig, pedagogisk, didaktisk og fagdidaktisk dømmekraft i dialog og samhandling med kolleger. For å få til dette, så må vi hele tiden utvikle oss. Så selv om det er en stor utfordring, vil dette være en veldig viktig del av vår felles utvikling.

Ekkokammer/fake news/fare for demokratiet

Udir har laget noen råd for hvordan man skal bruke kunstig intelligens i skolen (Udir, 2024). De er opptatt av hvorfor skolene må jobbe med KI nå, samtidig nevner de utfordringer knytta til det. Blant annet peker de på at: “Samtidig kan verdier som demokrati og ytringsfrihet være under press i en verden der "fake-news" og påvirkningskampanjer får enda sterkere verktøy gjennom bruk av KI.” (Udir, 2024). Dette er store utfordringer i verden generelt og KI har gjort disse utfordringene flere og vanskeligere å oppdage. Deepfakes og falske nyheter florerer. I tillegg har alle KI-er en politisk slagside eller "bias" som det er viktig å være klar over.

Som et eksempel, gjennom hyppige spørsmål som kan stilles av brukere om politisk konflikter, kan svarene fra chatbots villede den generelle oppfatningen. Som et konkret eksempel i tilfelle av den russisk/ukrainske konflikten, programmerere og utviklere kunne trene chatbots til å gi positive eller negative svar til publikum basert på deres egne politiske interesser (Kooli, 2023).

Juks

Juks og uselvstendig arbeid er en stor utfordring. Det er ikke alltid like lett å oppdage og det er vanskelig å avsløre/bevise. Å konfrontere en elev med mistanker om juks skaper også ofte brister i relasjonen mellom lærer og elev.

Kritisk bruk

Kritisk bruk av KI er viktig for at elevene skal kunne få læringsutbytte og at de er trygge på at informasjonen de får er riktig. Utviklingen innen KI og språkmodeller er med å forsterke betydningen av at elevene lærer å vurdere tekst og informasjon kritisk, og at de utvikler en etisk bevissthet og dømmekraft. Her er det stor variasjon blant elevene i hvilket nivå de er på. (Udir, 2024)

Etiske bekymringer

I forskning om bruk av chatbots vises det til etiske bekymringer om nøyaktigheten og påliteligheten til dataene som samles inn, og potensialet for chatboter til å opprettholde skadelige skjevheter og diskriminering. Det er også en bekymring når det gjelder personvern, skjevhet i data, ansvar for nøyaktighet, mangel på åpenhet og potensiale for misbruk (Kooli, 2023).

4.2 Hvilke muligheter finnes det?

Som vi har sett i denne veiledningen, finnes det mange spennende muligheter med KI. Vi kommer ikke til å skrive utfyllende om dette her, da det finnes mange konkrete eksempler i heftet. Vi nevner likevel noen viktige punkter under her:

Muligheter for sosial utjevning

Det er viktig å lære elevene god og kritisk bruk av kunstig intelligens slik at de vil kunne utnytte dette verktøyet på en god måte nå og i fremtiden.

KI kan hjelpe læreren å legge til rette for nivådifferensiering i klasserommet.

KI kan hjelpe elever med å øke forståelsen i en tekst/tema ved å be chatbots forenkle eller oppsummere mer krevende tekster.

Effektivisering av arbeid

Lage undervisningsopplegg, gi tilbakemeldinger til elever, skrive modelltekster, konstruere nye bilder, osv.

5. Hvordan fungerer Kunstig Intelligens?

Kunstig intelligens (KI) kan være **regelbasert** eller **datadrevet**. Regelbasert KI innebærer slik navnet tilsier at man skriver regler for eller forhåndskoder hvordan KI skal oppføre seg. Chatbotene til nettbanken som spør deg hva du vil ha hjelp med er som oftest delvis regelbaserte. Disse gir deg et forhåndsprogrammert svar på spørsmålet du skriver inn.

Det meste av KI i dag er datadrevet. Dette er modeller som er «trent opp» på data og har lært selv hvordan de skal oppføre seg. Det er dette som kalles «maskinlæring». Videre deler vi KI inn i modeller som klassifiserer data – det vil si putter dem inn i bestemte forhåndsdefinerte kategorier og modeller som genererer data, altså lager nye data.

Input eller inndata er det du gir inn til modellen. Output eller utdata – det modellen gir tilbake til deg. Både inndata og utdata kan være av ulik type. Som oftest er det tekst, men det kan også være bilder eller lyd.

Den typen KI som det i skrivende stund er viktigst for lærere å kjenne til er generative språkmodeller. Begrepene språkmodeller og chatboter (eller samtaleroboter) brukes gjerne om hverandre, men det er en vesensforskjell. En måte å si det på, er at språkmodellene er matematiske representasjoner av språklige relasjoner, det vil si forholdene mellom setninger, ord og orddele. For å kunne få tilgang til disse bruker man ofte en chatbot som «ligger oppå» språkmodellen og gjør det mulig å hente ut innholdet og kunnskapen som er representert i språkmodellen.

Her skal vi først skrive litt om hva som kjennetegner maskinlæring generelt og deretter hva som kjennetegner den vanligste typen språkmodeller i dag, som kalles GPT.

5.1 Maskinlæring

Den vanligste typen oppbygning eller «arkitektur» for KI i dag er såkalte «nevrale nettverk». Grunnen til at de har dette navnet er at de er inspirerte av måten hjernen fungerer. Alt som foregår i et slikt nettverk er ulike matematiske operasjoner som hver for seg er ganske enkle, men fordi det er så mange av dem, gjør dette modellen i stand til å «lære». Vi sier at et nevralt nettverk består av flere lag og hvert lag består av flere noder. Figuren under er et eksempel på den enkleste typen nevralt nettverk som kalles en «Perseptron». Dette nettverket har noen inndata, et skjult lag og en funksjon som prøver å predikere hvilken klasse dataene tilhører, altså en klassifiseringsoppgave.

Hver node i det skjulte laget har et tall som kalles en vekt. Når vi snakker om at et nevralt nettverk «trenes» eller «lærer» betyr dette egentlig bare at vekten til hver av nodene forandres bitte litt veldig mange ganger. Det nevrale nettverket får «se» inndataene veldig mange ganger – såkalte «epoker». Når vi snakker om antall parametere i en språkmodell, mener vi det totale antallet vektene i modellen. I dette tilfellet er det bare to mulige klasser, enten «Riktig» eller «Feil» og håpet er at modellen skal bli god til å gjette riktig klasse også på nye data etter at den har blitt trent i mange epoker.

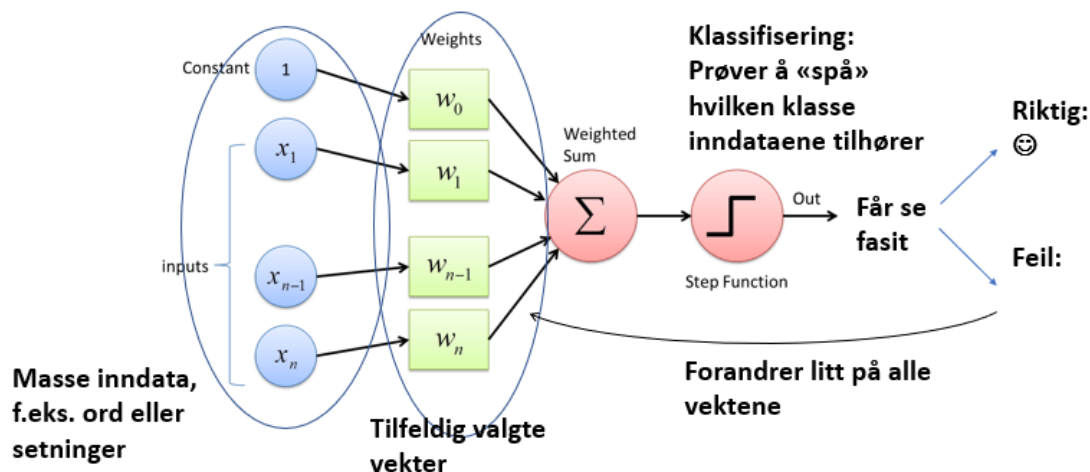


Figure 5.1: Perseptron. Bakgrunnsbildet er hentet fra: <https://medium.com/@shaistha24/basic-concepts-you-should-know-before-starting-with-the-neural-networks-nn-3-6db79028e56d>

Som nevnt er dette den aller enkleste typen nevralt nettverk og de fleste nettverkene i dag har mange skjulte lag som består av noder som hver har sine egne vektene. Begrepet «dyp læring» betyr at man trener opp nettverk som har mange skjulte lag. Eksempelet over er et tilfelle av det vi kaller «veiledet læring» (engelsk: supervised learning). For veiledet læring har man en fasit som modellen får se etter at den har prøvd å predikere klassen til en bestemt andel av treningsdataene. Hvis modellen tok feil justeres vektene til nodene litt, sånn at det er større sannsynlighet for at modellen gjetter riktig neste gang. Det er likevel viktig å ikke forandre vektene så mye at modellen tilpasser seg dataene den har trent på i for stor grad. Da sier vi at den blir «overtrent», det vil si at den blir veldig god til å predikere data den er trent på, men dårlig til å predikere nye data.

Ikke-veiledet læring betyr at vi ikke har noen fasit. I dette tilfellet kan vi ikke oppdatere vektene på samme måte, men må bruke andre strategier for å trene opp modellen. Det kan for eksempel være hvor like de ulike datapunktene er. Ikke-veiledet læring brukes for eksempel å gruppere data i klynger (engelsk: «cluster») og forsøke å lage en form for mønster i tilsynelatende kaos.

Forsterkningslæring (engelsk: «reinforcement learning») innebærer at modellen får en form for belønning og straff i form av en tallverdi avhengig av hvor godt den har gjort det, litt på samme måte som en hund får en godbit når den utfører en kommando. Så skal modellen selv finne den beste strategien for å få belønningen gjennom å prøve ut veldig mange forskjellige strategier. Et eksempel på dette er sjakkmodellen «Alpha Zero» som fikk en svært høy belønning når den vant et sjakkparti og jo flere trekk den brukte jo høyere straff fikk den. Denne modellen lærte seg å spille sjakk bedre enn både mennesker og andre, regelbaserte, sjakkprogrammer bare ved hjelp av forsterkningslæring og å spille mot seg selv mange millioner ganger.

5.2 GPT

Samtalerobotene de fleste har vært i kontakt med i skrivende stund er ChatGPT. GPT står for «Generative Pre-trained Transformer». Vi forklarer hvert ord for seg.

Generativ

At en maskinlæringsmodell er generativ betyr at den lager nye data. For chatbotene sin del betyr dette å lage ny tekst. Det finnes også modeller som lager bilder på bakgrunn av tekst. Ikke alle maskinlæringsmodeller er generative. En vanlig oppgave er å klassifisere data, for eksempel ved at modellen kategoriserer restaurantanmeldelser eller filmanmeldelser som positive, negative eller nøytrale. Man kan også bruke en slik modell til å sette en merkelapp eller «tagge» data som allerede finnes. Et eksempel på dette er å få en modell til å predikere ordklassene til ordene i ei setning.

Pre-trained

Generative KI-er er «forhåndstrengte» - det vil si at de allerede har lært en god del om relasjonene mellom setninger, ord og orddele. Når vi bruker en GPT til å generere et svar, tar den altså utgangspunkt i det den har lært om språk på forhånd.

Transformer

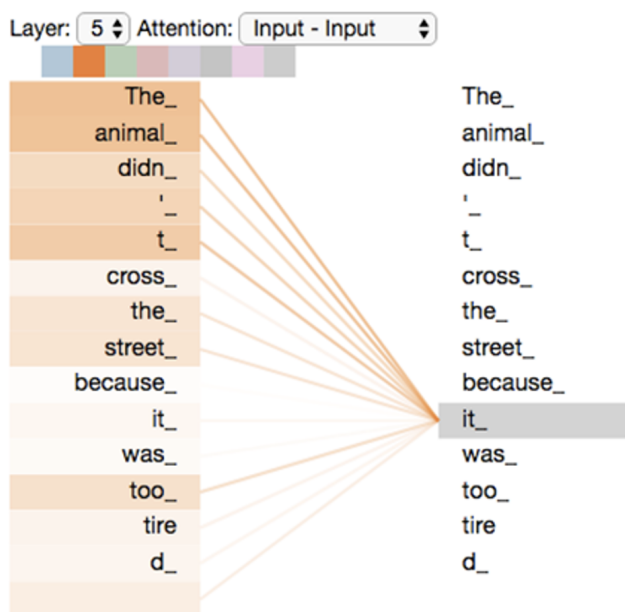


Figure 5.2: Self-attention.

Hentet fra <https://jalammar.github.io/illustrated-transformer/>

En transformer er en type nevral nettverk som har flere vekter enn et «vanlig» nevral nettverk. Den viktigste forskjellen er at hver orddel ikke bare har en vekt, men tre vekter. Til sammen sier disse vektene noe om relasjonen til alle de andre ordene i setningen og hvilke ord som er mest betydningsfulle i setninga¹. I eksempelet under vil pronomenet «it» ha lært gjennom trening at det har sterkere relasjon til «the animal» som er substantivfrasen det viser til. I tillegg bidrar ordet «cross» med mer betydning enn ordet «because» og vil derfor få en høyere vekt. En lignende arkitektur finnes ikke bare mellom orddelene i ei setning, men mellom de ulike setningene

en transformer får som input når den trenes opp. Denne forklaringa er noe forenklet og spesielt interesserte kan lese mer her: <https://jalammar.github.io/illustrated-transformer/>

Under følger en litt mer inngående forklaring av hvordan **trening** av en transformer foregår:

¹Denne mekanismen som gjør at orddelene tar hensyn til hverandre er kjent som "attention". Noen steder oversettes dette til norsk med "oppmerksomhetsmekanisme", men kanskje en bedre oversettelse er "hensynsmekanisme".

1. Tokenisering av inndata

Dette betyr at man deler opp teksten som skal brukes til å trene modellen. Tidligere delte man gjerne en setning opp i ord og tegnsetting og kalte hver av dem en «token», men i GPT-er deler man i praksis opp ordene i «underord» eller «orddeler». Et ord deles gjerne opp i to eller tre deler, avhengig av hvor vanlig ord delen er. Reint lingvistisk gir dette liten mening, men har vist seg å fungere godt i praksis. Det er faktisk mulig å se dette i praksis. Hvis man følger godt med når en språkmodell genererer tekst, vil man kunne legge merke til at den ikke alltid genererer nye ord, men ofte bare genererer orddeler. Etter at setningene som modellen skal trenes på har blitt delt opp i underord konverteres hver orddel til en liste av tall (såkalte «embeddings») som er grunnenheten i en transformer.

2. Trening

Oppgaven til en GPT er å predikere (eller gjette) neste orddel veldig mange ganger, gitt de tidligere orddelene i en setning. Siden en GPT er generativ og ikke har noen fasit, gjøres opptreninga på følgende måte: Først tokeniseres alle ordene som beskrevet over. Så får modellen vite hvilken setning den skal prøve å gjenskape. I begynnelsen får den bare se en orddel og ut fra dette skal den predikere hva som er neste orddel. Vektene i modellen forandres avhengig av hvor godt den har gjettet. Jo bedre den gjetter, jo mindre forandres vektene og jo dårligere den gjetter, jo mer forandres de. Neste gang får den se to orddeler og skal gjette neste orddel ut fra disse to. Slik fortsetter det til den har trent på hele setninga.

Her er et KI-generert eksempel på hvordan dette kan foregå:

Setning: "hats are great"

Orddeler: "ha", "ts", "are", "gr", "eat"

Første steg: Input: "ha". Mål: "ts"

Andre steg: Input: "ha", "ts". Mål: "are"

Tredje steg: Input: "ha", "ts", "are". Mål: "gr"

Fjerde steg: Input: "ha", "ts", "are", "gr". Mål: "eat"

Dette gjentas mange ganger med flere millioner ulike setninger.

3. Finjustering

Når en språkmodell er ferdig trent på denne måten, er det vanlig å finjustere den på ulike måter for at den skal bli enda bedre. En måte man kan gjøre dette på er ved å la mennesker sortere svarene fra best til dårligst og bruke dette til ytterligere å justere vektene. Hvis man får spørsmål fra ChatGPT om å velge det beste svaret eller Google Translate om å velge den beste oversettelsen er man med på en slik prosess. En annen måte å finjustere en modell er å trene den videre på data fra et bestemt fagfelt eller språk. Enkelte av de norske språkmodellene som for tiden er under utvikling bruker store språkmodeller primært trent på engelsk som utgangspunktet og trenes videre på norske data.

6. En foreløpig oversikt over KI-løsninger

Som nevnt i kapittel 4, kan det være nyttige å skille mellom språkmodeller og chatboter. Språkmodellene er de underliggende matematiske representasjonene av språk, mens chatbotene er agentene som gjør oss i stand til å kommunisere med språkmodellene. Ulike språkmodeller har ulike typer oppbygning eller arkitekturer. Alle modellene vi referer til her er såkalte «transformers» som i skrivende stund er den beste arkitekturen vi har.

6.1 Store språkmodeller våren 2024

Den mest kjente chatboten vi har i dag er ChatGPT. Denne finnes det også flere varianter av, både flere chatboter og flere underliggende språkmodeller. I skrivende stund er den beste språkmodellen ChatGPT 4. Alle disse modellene er trent opp av selskapet OpenAI. Det finnes flere andre store språkmodeller trent opp av andre selskaper eller organisasjoner. Andre kjente språkmodeller våren 2024 er Googles "Gemini", Anthropic's "Claude 3", Metas "Llama 3" og "Mistral" som er utviklet av et lite fransk selskap, men har oppnådd svært gode resultater. En fordel med de to sistnevnte er at de er såkalt åpne modeller, det vil si at det er full åpenhet om hvordan de har blitt trent opp og på hvor mye data. Disse er også åpent tilgjengelig for at hvem som helst kan bruke dem og videreutvikle dem (såkalt "fine-tuning"). Microsofts "Bing CoPilot" og Osloskolens KI ligger oppå ulike versjoner av OpenAI's GPT. Et annet eksempel er chatboten "Perplexity" som kombinerer resultatene fra flere av de andre.

Når vi snakker om at språkmodeller har ulik størrelse refererer vi til antall parametre som er det totale antallet vektorer i modellen. På mange måter er det disse vektene som "er" hele modellen, det er de som modellerer forholdet mellom ord og setninger. Det er ikke slik at tekstene modellen har blitt trent på ligger inne i modellen. Dette er grunnen til at vi ikke kan bruke modellene til plagiatkontroll – det finnes ingen enkel måte å hente ut tekstene modellen er trent på. Lenge var fokus i forskningen på å utvikle større og større modeller. Eksempelvis bruker Mistral's største modell 39 milliarder parametre¹, Metas Llama3 har 70 milliarder parametre². OpenAI er ikke åpen om hvor mange

¹<https://docs.mistral.ai/getting-started/models/>

²<https://towardsdatascience.com/deep-dive-into-llama-3-by-hand-%EF%B8%8F-6c6b23dc92b2>

parametre de nyeste modellen har, men GPT 3.5 skal visstnok ha 175 milliarder. Man vet ikke hvor mange parametre GPT4 har. Det siste året har imidlertid utviklingen gått fra å utvikle så store språkmodeller som mulig, til å prøve å utvikle mindre, mer spesialiserte språkmodeller. Mange selskaper har nå en redusert modell på rundt 7 milliarder parametre.

6.2 Utvikling av norske språkmodeller

Mange av forskningsmiljøene og selskapene i Norge jobber med å utvikle egne norske språkmodeller. Hovedaktørene her er Universitetet i Oslo (UiO), Norges Teknisk Naturvitenskaplige Universitet (NTNU) og Nasjonalbiblioteket. Denne forskningen foregår langs ulike spor og som et samarbeid med flere andre aktører. Gode norske språkmodeller vil kunne gi store samfunnsmessige fordeler og være særlig viktig for skoleverket. En åpenbar fordel er at den sosiale "biasen" og verdiene som følger med på lasset i større grad vil reflektere en norsk kontekst enn en amerikansk-engelsk kontekst som mye av treningsdataene til de store modellene kommer fra. En annen fordel er at vi i større grad kan sikre godt personvern, både ved at vi har kontroll på hvilke data vi bruker til å trene opp modellene, og at vi unngår at brukernes input brukes til å trene opp nye modeller.

NorwAI ledet av NTNU og Nora LLM ledet av UiO jobber med utvikle norske språkmodeller. Flere av disse er trent opp fra bunnen, mens andre er videretrent på underliggende språkmodeller som Mistral og Llama. I det siste tilfellet betyr dette at man først tar språkmodeller som overveiende er trent på engelske tekster og trener de videre, eller finjusterer de på norske tekster fra Nasjonalbiblioteket. Alle disse modellene regnes som små modeller med rundt 7 milliarder parametre. I tillegg arbeider NorwAI med å utvikle norske språkmodeller som trenes fra bunnen, det vil si ikke er forhåndstrent på engelske tekster. Den største av disse ventes å ha 40 milliarder parametre³. I skrivende stund er det mulig for spesielt interesserte å ta i bruk Nora LLMs Normistral 7B⁴, men dette krever en del teknisk kompetanse og det finnes ikke enda noen enkel nettbasert chatbot slik det gjør for de store internasjonale modellene. Det er temmelig sikkert at det kommer til å skje en stor utvikling innen dette det kommende året som gjør at mye av det vi har skrevet må oppdateres.

³<https://aiavisen.no/norske-sprakmodeller/>

⁴<https://huggingface.co/norallm>

7. Flere ressurser

Hva er KI?

NRK: Abels tårn - den store serien om KI:

- Episode 2:10 – ChatGPTs norske utfordrer: https://radio.nrk.no/podkast/abels_taar_n/1_4f800205-0dda-4e41-8002-050dda6e4113
- Episode 7:10 – ChatGPT: Hvordan fungerer det? https://radio.nrk.no/podkast/abels_taar_n/1_0143ba79-0ee8-43d5-83ba-790ee843d56c

Bøker

- Strümke, Inga. (2023) *Maskiner som tenker. Algoritmenes hemmeligheter og veien til kunstig intelligens*. Kagge Forlag
- Bøhn, Einar D. (2022) *Teknologiens Filosofi. Metafysiske problemstillinger*. Cappelen Damm

Nyttige lenker

- Tren en modell til ansiktsgjenkjenning: <https://teachablemachine.withgoogle.com/>

Pedagogisk bruk av KI

Podcast – Pedagogisk intelligens

- Lyd: <https://soundcloud.com/pedagogisk-intelligens>
- Video: <https://www.youtube.com/playlist?list=PLeYxHW3kH5ocx0Hk7k0WwnRgBJzSD-VRA>

Høgskolen i Østfold - MOOC

- Pedagogisk bruk av IKT (skole og utdanning). Modul 7 Kunstig intelligens i undervisning: https://hiof.instructure.com/courses/549/pages/7-dot-1-introduksjon?module_item_id=250742

Morsomme lenker for å teste KI

- Gjett hvilket ansikt som er ekte: <https://www.whichfaceisreal.com/>
- KI gjetter hva du har tegnet: <https://quickdraw.withgoogle.com/>
- Forslag til ledetekster Randaberg: <https://ki.tenkemotoren.no/> (passord: allelikertm)
- Tips til teknikker for god prompting: <https://medium.com/@eldatero/master-the-perfect-chatgpt-prompt-formula-c776adae8f19>

Bøker

- Elstad, Eyvind (2023) *Læreren møter ChatGPT*. Universitetsforlaget

En foreløpig oversikt over ulike "Samtaleroboter"

- KI Osloskolen: <https://ki.osloskolen.no>
- ChatGPT: <https://chatgpt.com/>
- BING Copilot: <https://bing.com/chat>
- Google Gemini: <https://gemini.google.com/app>
- Mistral: <https://mistral.ai/>
- Perplexity: <https://www.perplexity.ai/>
- Normistral: <https://huggingface.co/norallm/normistral-7b-scratch>

Muligheter og utfordringer ved KI

Udir: Råd om kunstig intelligens i skolen

- <https://www.udir.no/kvalitet-og-kompetanse/digitalisering/kunstig-intelligens-ki-i-skolen/>

Udir: Profesjonsfellesskap og skoleutvikling

- <https://www.udir.no/lk20/overordnet-del/3.-prinsipper-for-skolens-praksis/3.5-profesjonsfellesskap-og-skoleutvikling/>

UDE (Oslo)

- <https://aktuelt.osloskolen.no/larerik-bruk-av-laringsteknologi/digital-skolehverdag/kunstig-intelligens-ki-i-osloskolen/>

Studie om bruk av ChatGPT blant universitetsstudenter

- <https://educationaltechnologyjournal.springeropen.com/articles/10.1186/s41239-024-00444-7>

Dansk skole

- <https://www.uvm.dk/-/media/filer/uvm/aktuelt/pdf24/april/240423-ekspertrgruppen-om-chatgpt-og-andre-digitale-hjaelpemidlers-rapport.pdf>

KI i klasserommet

- https://tenk.faktisk.no/samling/ki?fbclid=IwAR0Kh3otW69gVxsCuDiXiLW_K_z6nR3jAgU6R6otXx07-UIT_xT6UynT87Y#d6535663d9a1

Kilder

Kooli, C. (2023). Chatbots in Education and Research: A Critical Examination of Ethical Implications and Solutions. *Sustainability* 2023, 15, 5614. Hentet 19.06.24 fra: <https://doi.org/10.3390/su15075614>

Udir (29.02.24). Råd om kunstig intelligens i skolen. Hentet 19.06.24 fra: <https://www.udir.no/kvalitet-og-kompetanse/digitalisering/kunstig-intelligens-ki-i-skolen/>

Udir (ingen dato). Profesjonsfellesskap og skoleutvikling. Hentet 19.06.24 fra: (<https://www.udir.no/lk20/overordnet-del/3.-prinsipper-for-skolens-praksis/3.5-profesjonsfellesskap-og-skoleutvikling/>)