

Modul 2 - gruppe 4

IN5480 - Interaction with AI



IN5480 - Interaction with AI	0
1. Beskrivelse av gruppen	2
2. Interesseområde	2
3. Bakgrunn	3
4. Forskningsspørsmål	4
5. Metoder	4
6. Funn	5
7. Prototype - Kompis	5
8. Appendix 1	6
8.1 Design av chatbot	6
8.2 Refleksjon	7
9. Appendix 2	7
9.1 Resultater fra maskinlæringsmodellen	7
9.2 Refleksjon	8
Litteraturliste	9

1. Beskrivelse av gruppen

Vi er en gruppe med seks designstudenter på førsteåret master. Gruppen består av:

Mathias Demeer Strøm (mathiads@uio.no)

Tonje Løfqvist (tonjlo@uio.no)

Jakob Føllesø Egeland (jakobfe@uio.no)

Ida Ødegård (idaode@uio.no)

Eivind Gisholt (eivindgi@uio.no)

Stian Dolmseth (stiandol@uio.no)

Fire av oss har en bachelorgrad i “Informatikk: Design, bruk og interaksjon” ved Universitetet i Oslo, mens to av oss har en bachelorgrad i “Medie- og interaksjonsdesign” fra Universitetet i Bergen.

2. Interesseområde

Vi er interessert i å finne ut hvordan man kan implementere AI i skolen/universitetet. I undervisningssammenheng? I grunnskole eller høyere utdanning? For enkeltstudenter eller for flere/et klasserom? For lærere/forelesere? Etter et lite søk fant vi fort ut at det var mange muligheter for å bruke AI i utdanning. Vi fant følgende sitat fra Unesco:

“Artificial Intelligence (AI) has the potential to address some of the biggest challenges in education today, innovate teaching and learning practices, and ultimately accelerate the progress towards SDG 4” (Unesco, u.d)

Etter en runde med brainstorming rundt dette temaet ble vi enige om at teknologier som vi er interesserte i å utforske er chatbots i form av en personlig assistent. Vi tror dette kan være et interessant tema ettersom samtlige i gruppen uttrykte frustrasjon i forbindelse med planlegging og organisering av studiene. Vi vil derfor utforske om det er mulig å bruke AI i form av en personlig hjelper som kan assistere studenter i studiehverdagen. Slik som Luckin et al. skriver er AI stort og interdisiplinært, og omfatter mange forskningsfelt (Luckin et al., s.14).

3. Bakgrunn

Følstad og Brandtzæg snakker om hvordan vi er på vei i en retning hvor vi ikke bare interagerer med maskiner i form av sveiping, klikking og scrolling - men også gjennom tekst som representerer vårt daglige språk (Følstad og Brandtzæg, 2017, s.38). Store bedrifter som Domino's og Taco Bell har for eksempel prøvd å implementere chatbots som håndterer ordrer istedenfor å bruke de ordinære grensesnittene sine.

Firmaet K12 lanserte i 2020 en chatbot som fungerer for sine skoler i USA. Chatboten har som mål å forenkle kommunikasjonen mellom foreldre, deres barn og lærerne. Den lover å være nyttig for dem som ikke kan være fysisk til stede for å motta informasjon om for eksempel retningslinjer for Covid-19, måltider eller foreldremøter. Under den første implementeringen av chatboten svarte den på over 400 spørsmål over 30 dager fra 200 deltakere - med en gjennomsnittlig kundetilfredshet på 9,6 av 10 (Forbes, 2020). Det kan tenkes at slik teknologi også kan fungere som et godt verktøy for studenter på UiO som lurere på alt fra administrative oppgaver til tips om hvordan man kan stresse mindre i hverdagen.

Bratteteig og Verne skriver i sin artikkel "Does AI make PD obsolete?" at folk ofte endrer sin atferd for at teknologien skal prestere bedre. De har enklere språk i samtale med en chatbot enn med et menneske (Bratteteig og Verne, 2018, s. 4). I vår case er dette interessant da vi ønsker å snakke med studentene om deres hverdag. Vi må imidlertid være klar over at chatboten vår ikke kan erstatte menneskelig kontakt og at den derfor ikke fungerer som en samtalepartner, men mer som en motivasjonshjelp i hverdagen.

Amershi et al., presenterer 18 retningslinjer for design av menneske-AI interaksjon og guideline 5 representerer Bratteteig og Verne sitt argument godt. Denne lyder: "Match relevant social norms. Ensure the experience is delivered in a way that users would expect, given their social and cultural context." (Amershi et al., 2019). Vi må passe på at studentene er klare over hva vår chatbot er i stand til slik at den ikke blir brukt på feil grunnlag eller at studentene blir frustrert over den.

4. Forskningsspørsmål

Vi tror det sosiale aspektet rundt AI er en viktig faktor i undervisningssammenheng, og vil se nærmere på dette i prosjektet. Vi har derfor utformet følgende to spørsmål:

- Hvordan kan kunstig intelligens hjelpe og motivere studenter med å organisere hverdagen bedre?
- Hvordan kan kunstig intelligens bidra til økt trivsel hos studentene?

5. Metoder

For å utforske om AI har en positiv effekt på studentenes hverdag skrev vi i den første prosjektplanen at vi ønsket å bruke flere kvalitative metoder som fokusgrupper, litteratursøk, brukertester. Etter diskusjon innad i gruppa ble vi enige om å gjennomføre litteratursøk for å få bedre forståelse for AI (kap. 3), i tillegg til en dagbokstudie. Ved å gjennomføre en slik studie ville vi få bedre innsikt i studentenes hverdag og dypere forståelse for hvordan de har det. Dagbokstudien la også til rette for refleksjon rundt hva som fungerer og ikke i studentenes hverdag.

Vi har samarbeidet med fem studenter fra Universitetet i Oslo som skulle svare på forhåndsdefinerte spørsmål fra oss i en periode på en uke. Spørsmålene skulle besvares hver kveld. Spørsmålene vi stilte var:

- Hvor lite/mye motivert har du vært for å jobbe med skolearbeid i dag? Hvorfor (ikke)?
- Hvor godt har du trives med studiene i dag? Hvorfor (ikke)?
- Var det noe som var vanskelig idag?

Vi jobber iterativt i denne oppgaven, så til neste modul ønsker vi å gjennomføre brukertester for å forme løsningen til å bli best mulig tilpasset brukernes behov.

6. Funn

Gjennom dagbøkene som studentene skrev, fikk vi verdifull innsikt om deres hverdag og hva de preges av. På generell basis var studentene positive til dagbokskrivning, da det ga dem muligheten til å strukturere tankene sine og reflektere over egen hverdag. De veiledende dagbokspørsmålene hjalp studentene med å bli mer bevisst på hvordan de har de i hverdagen, og samtidig også hva som gjør at dagene deres er mer eller mindre gode. Chatboten vår inneholder derfor flere spørsmål som ligner de som stod i dagbokstudien, slik at studentene får reflektert over hvordan de har det i sin hverdag.

Det kom tydelig fram i resultatene at studiehverdagen og studentenes motivasjon til å arbeide med skolearbeid i stor grad blir preget av andre ting enn skole. Trivsel og motivasjon ble i stor grad vektlagt av studentene. I tillegg til arbeid tilknyttet studiene, vektla studentene søvn, trening og mat som avgjørende faktorer i hvor gode eller dårlige dagene deres var. Dersom studentene har dårlige dager forteller de at det av og til ikke skal mer enn en enkel oppmuntring til for å snu humøret, noe vi ønsker å ta med oss videre inn i vår løsning.

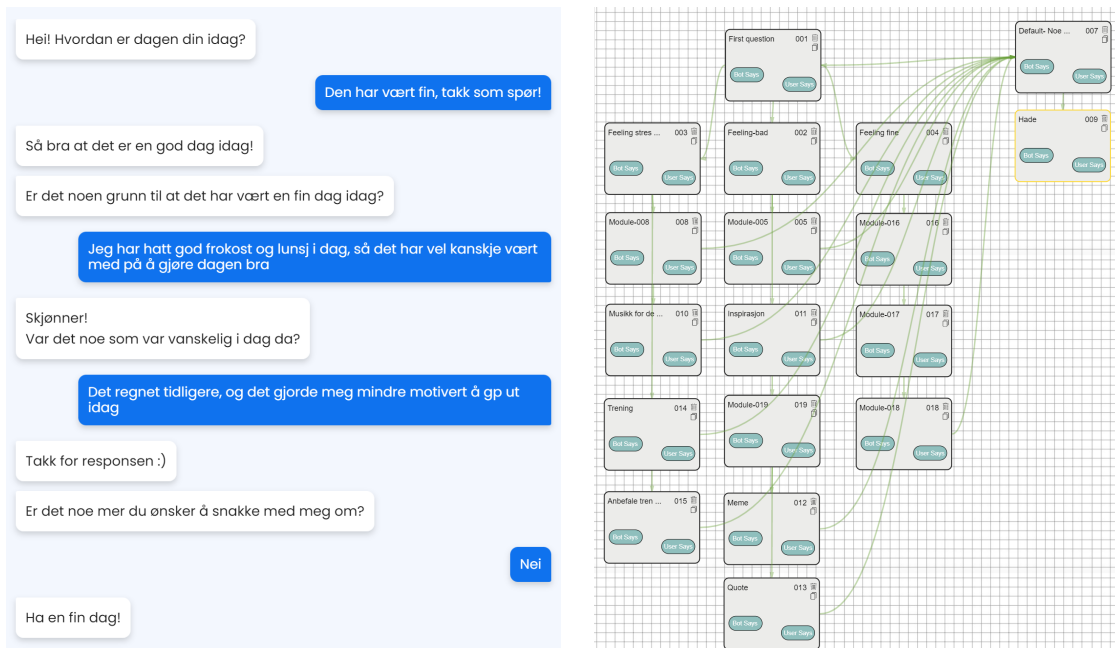
7. Prototype - Kompis

Vi har laget en chatbot ved navn Kompis som vår prototype. Hensikten med prototypen er at chatboten skal gi studenten motivasjon og tips. Prototypen er bygget med Chatteron og flyten er basert på svaret fra en bruker som har tre forskjellige utfall i første møte. Ut ifra hvilken flyt chatboten velger blir det presentert ulik data og innhold basert på brukernes svar. Når brukeren kommer på “slutten av flyten” vil chatboten stille et oppfølgingsspørsmål som enten tar deg tilbake til start eller avslutter sekvensen. Videre finner du en beskrivelse av den overordnede flyten til chatboten:

- Om dagen har vært bra vil chatboten spørre om hvorfor dagen har vært som den har vært. Dette er for å få brukeren til å reflektere over dagen sin.
- Om dagen har vært dårlig vil chatboten spørre om hvorfor dagen har vært som den har vært, for deretter å komme med mulige tiltak til å forbedre dagen til studenten.
- Hvis dagen har vært helt grei og studenten er sliten eller ikke har klart å komme helt i gang med dagen, vil chatboten spørre om studenten trenger motivasjon i form av en quote eller morsomme “memes”.

Ved den første prototypen vår baserte vi oss mer på knapper som brukeren måtte bruke for å kunne kommunisere med chatboten. I iterasjon 2 - Kompis - av den opprinnelige prototypen, har vi gjort den mer avansert til å klare å skille ut forskjellige ord, som brukes for å ta videre valg.

Dette gjør at lengre setninger også er akseptert, og prototypen vår fungerer mer som tenkt.



8. Appendix 1

8.1 Design av chatbot

I denne delen ønsket vi å utforme en chatbot som kunne svare til deler av problemstillingen vår om hvordan kunstig intelligens kan hjelpe studenter med å organisere og øke trivselen i hverdagen. Ved hjelp av chatteron.io skrudde vi sammen en enkel prototype som kunne initiere til en samtale om hvordan mennesket på den andre siden har det. I hovedsak er chatboten designet med studenter som hovedmålgruppe, men den kan fint benyttes av andre som rett og slett ønsker å dele noe fra dagen sin og få oppmuntring eller støtte.

Interaksjon med boten starter med at den spør hvordan dagen går, og basert på om brukeren velger "stresset", "dårlig" eller "bra", reagerer boten deretter. Hvis brukeren indikerer at en føler seg dårlig eller er stresset, vil boten forsøke å finne ut hva som plager brukeren ved å spørre om det f.eks. gjelder studier eller fritid. Interaksjonen består av både predefinerte

knapper brukeren kan trykke samt noe fritekst, og boten kan foreslå ulike løsninger dersom brukeren er sliten, sulten eller demotivert - for å nevne noe.

8.2 Refleksjon

Det å utforme en chatbot har vært en lærerik og spennende sosio-teknisk utfordring. Fra et teknisk perspektiv, har vi erfart at det kan være krevende å utforme dialog som åpner opp for et utvalg av input fra bruker, men som samtidig begrenser skopet for mulige samtaleemner, da vi ikke har kapasitet til å implementere for mange ulike interaksjonsmuligheter.

Fra et sosialt perspektiv har vi fått inntrykk av at samtaler over fritekst gir en mer medmenneskelig opplevelse enn ved bruk av knapper, selv om knappene gir oss designere bedre kontroll over samtalen mellom bot og bruker. Gjennom godt utformede setninger og iterativ testing mener vi at fritekst vil skape en bedre brukeropplevelse til tross for større risiko for at boten ikke kan svare på alt brukeren forespør.

9. Appendix 2

Under maskinlæringsoppgaven i appendix 2 var målet at vi skulle oppnå forståelse for hvordan kunstig intelligens og chatboter fungerer i praksis. Vi brukte et eksisterende chatbotprogram og endret variablene i koden for å sammenligne “loss” og “accuracy” i outputen. Vi forsøkte å finne en sammenheng mellom mengden nevroner, antall treningsiterasjoner og mengde input.

9.1 Resultater fra maskinlæringsmodellen

Dersom vi kjører programmet med de opprinnelige verdiene hvor **batch_size = 32**, **max_words = 1000**, **epochs = 2**, får vi dette som resultat:

```
Epoch 1/2
29/29 [=====] - 1s 10ms/step - loss: 3.2232 - accuracy: 0.1244
Epoch 2/2
29/29 [=====] - 0s 5ms/step - loss: 2.8171 - accuracy: 0.1722 -
```

Dersom vi endrer **max_words = 10 000**, øker tiden for hvert steg den utfører - men modellen er fortsatt like nøyaktig i siste epoch:

```
Epoch 1/2
29/29 [=====] - 1s 39ms/step - loss: 3.2675 - accuracy: 0.1500 -
Epoch 2/2
29/29 [=====] - 1s 36ms/step - loss: 2.8652 - accuracy: 0.1722 -
```


Dersom **batch_size = 320**, **max_words = 1000**, **epochs = 2**, så tar modellen lenger tid per steg, den blir betydelig mindre nøyaktig og har noe større “loss”.

```
Epoch 1/2
3/3 [=====] - 1s 85ms/step - loss: 3.3857 - accuracy: 0.0578
Epoch 2/2
3/3 [=====] - 0s 27ms/step - loss: 3.3450 - accuracy: 0.1556
```

Dersom **batch_size = 32**, **max_words = 1000**, **epochs = 200**, vil de mange iterasjonene føre til relativt stabile og jevne resultater med et noe lavere “loss”.

```
Epoch 198/200
29/29 [=====] - 0s 7ms/step - loss: 2.6261 - accuracy: 0.1722
Epoch 199/200
29/29 [=====] - 0s 7ms/step - loss: 2.6264 - accuracy: 0.1478
Epoch 200/200
29/29 [=====] - 0s 7ms/step - loss: 2.6262 - accuracy: 0.1722
```

9.2 Refleksjon

Det var tidkrevende og vanskelig å sette seg inn i oppgaven da vi har lite erfaring med Python og maskinlæring. Vi forsøkte først å kjøre koden lokalt men måtte gi opp og bruke Google Colab istedenfor fordi vi fikk problemer med å importere NumPy.

Vi så utvikling i programmet når vi kjørte flere treningsiterasjoner, men vi er fremdeles litt usikre på hvordan man skal tolke all dataen. Vi opplevde at det var utydelig hvilke variabler vi skulle eksperimentere med og hvilke endringer vi skulle lete etter, og mye av tiden gikk derfor med til å opparbeide teknisk forståelse av oppgaven. Outputen vi leste av ga ikke veldig stor variasjon til tross for endring av variabler, og det skapte usikkerhet om hvorvidt vi justerte de riktige variablene. Det var også vanskelig å se en sammenheng mellom feilmeldingene som dukket opp, da det av og til kunne komme feilmelding da vi prøvde å snakke norsk til chatboten, og andre ganger ikke. Det var også vanskelig å skjønne om den lærte av å snakke med oss eller om svarene alltid var helt tilfeldige. Til tross for noen tekniske utfordringer, var dette en spennende og lærerik oppgave å arbeide med, og den har gitt oss et innblikk i hvordan AI og maskinlæring fungerer.

Litteraturliste

Amershi, S. Weld, D. Vorvoreanu, M. Fourney, A. Nushi, B. Collisson, P. Teevan, J. (2019). *Guidelines for human-AI interaction*. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM. Lest: 15.10.2021

Bratteteig, Tone. Verne, Guri. (2018). *Does AI make PD Obsolete? Exploring challenges from Artificial Intelligence to Participatory Design*. In Proceedings of PDC 2018, Belgium. Lest: 20.10.2021

Forbes. *Company Launches AI Chatbot Designed For K-12 School Districts*. Hentet fra: <https://www.forbes.com/sites/simonchandler/2020/10/30/company-launches-ai-chatbot-designed-for-k-12-school-districts/?sh=7b0a698b4b26>

Følstad, A. Brandtzæg, P. B. (2017). *Chatbots and the new world of HCI*. *interactions*, 24(4), 38-42. Lest: 19.10.2021

Unesco. Artificial intelligence in education. Hentet fra: <https://en.unesco.org/artificial-intelligence/education> Lest 09.09.21

Luckin, Rose; Holmes, Wayne; Griffiths, Mark and Forcier, Laurie B. (2016). *Intelligence Unleashed: An argument for AI in Education*. *Pearson Education, London*. Hentet fra: <http://oro.open.ac.uk/50104/1/Luckin%20et%20al.%20-%202016%20-%20Intelligence%20Unleashed.%20An%20argument%20for%20AI%20in%20Educ.pdf>