

# Sluttrapport - Gruppe 4

IN5480 - Interaction with AI



<b>IN5480 - Interaction with AI</b>	<b>0</b>
1. Beskrivelse av gruppen	2
2. Interesseområde	2
3. Bakgrunn	3
4. Forskningsspørsmål	4
5. Metode	4
5.1 Dagbokstudie	5
5.2 Research through design	5
6. Funn	6
7. Prototype - Kompis	7
8. Brukertesting	8
8.1 Gjennomføring	8
8.2 Funn	9
9. Diskusjon	10
10. Konklusjon	12
11. Appendix 1	13
9.1 Design av chatbot	13
9.2 Refleksjon	13
12. Appendix 2	14
10.1 Resultater fra maskinlæringsmodellen	14
10.2 Refleksjon Appendix 2	15
Litteraturliste	16

# 1. Beskrivelse av gruppen

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

Mathias Demeer Strøm	<a href="mailto:mathiads@uio.no">mathiads@uio.no</a>
Tonje Løfqvist	<a href="mailto:tonjlo@uio.no">tonjlo@uio.no</a>
Jakob Føllesø Egeland	<a href="mailto:jakobfe@uio.no">jakobfe@uio.no</a>
Ida Ødegård	<a href="mailto:idaode@uio.no">idaode@uio.no</a>
Eivind Gisholt	<a href="mailto:eivindgi@uio.no">eivindgi@uio.no</a>
Stian Dolmseth	<a href="mailto:stiandol@uio.no">stiandol@uio.no</a>

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

Som studenter selv er vi interessert i å finne ut hvordan man kan implementere AI på universitetet. I forkant av studiet har vi innad i gruppen diskutert relevante bruksområder, og kommet frem til at det er studentenes hverdag vi ønsker å se dypere på. Vi vet at livet til en student kan være alt fra hektisk til gøy og ønsker å finne ut av i hvilken setting AI kunne vært relevant. Vi fant følgende sitat fra Unesco, FNs organisasjon for utdanning, vitenskap, kultur og kommunikasjon:

*“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).*

Sitatet adresserer at AI har stort potensiale i å adressere utfordringer i utdanning i dag og det er dette vi ønsker å finne noen svar på. Etter litteratursøk og diskusjoner ble vi enige om at vi er interessert i å utforske en chatbot 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, og at dette også kunne ha påvirkning på motivasjonen. 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,

interdisiplinært og omfatter mange forskningsfelt (Luckin et al., s.14), men vi ønsker å undersøke chatboter for å bli bedre kjent med muligheter og begrensninger innen slik teknologi.

### **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 (2018) 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 (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.

Kocielnik et al. (2018) viser til sitt prosjekt hvor de designet en samtaleagent som blant annet brukte chat-basert interaksjon for å støtte selvlæring og refleksjon på en arbeidsplass. Prosjektet viser til både fordeler og ulemper med chat-basert interaksjon, samt hvordan det kan brukes til å bevisstgjøre brukere på sin egen hverdag og forbedre både produktivitet og arbeidstilfredshet. Deltakerne i prosjektet følte at det gikk raskt å lese spørsmål i chat og at

det ga dem tid til å tenke seg om før de svarte i sitt eget tempo. Noen av deltakerne nevnte også at kommunikasjon gjennom chat føles lite personlig. Vi ønsker å se om vi kan finne likheter og forskjeller mellom Kocielnik et al. (2018) sitt prosjekt og vår case knyttet til studenter.

Amershi et al. (2019) 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. Metode

I denne delen vil vi redegjøre for valg av datainnsamlingsmetode og gjennomgå utførelsen av valgt metode. I tillegg vil vi redegjøre for hvorfor vi mener at prosjektet kan betraktes som et *Research through design*-prosjekt. Som nevnt i del 3 er vi interessert i å undersøke og grave i mulighetene for hvordan kunstig intelligens kan assistere og motivere studenter i en hektisk og kanskje stressende studiehverdag.

Basert på forskningsspørsmålene våre ønsker vi å komme tett innpå målgruppa for å kunne skape forståelse om hvordan studenter faktisk opplever sin egen hverdag. Valg av metode ble derfor basert på to kriterier:

1. Metoden må gjøre det mulig å samle inn kvalitative data med et klart fokus på motivasjon og trivsel.
2. Metoden må kunne tilrettelegge for at deltakerne skal kunne dele ærlige og oppriktige tanker om egen studiehverdag, uten å føle at de trenger å holde noe tilbake.

Disse to kriteriene la grunnlag for at teamet ønsket et dagbokstudie, da dette ga oss muligheten til å utforme predefinerte, men relativt åpne spørsmål som studentene selv kunne svare på når det passet dem best. Håpet var også at metoden tilrettela for at deltakerne ikke følte seg observert under datainnsamlingen, og at de sto fritt til å dele sine innerste tanker.

## 5.1 Dagbokstudie

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 og brukertester. Etter diskusjon innad i gruppa ble vi enige om å gjennomføre litteratursøk for å få bedre forståelse for AI (som beskrevet i kapittel 3), i tillegg til et dagbokstudie som datainnsamlingsmetode.

Vi har brukt en brukersentrert designprosess som tilnærming der brukernes ønsker og behov er i fokus. Vårt mål er at brukernes behov skal gi grunnlag for prototyping og gjennom dagbokstudie vil brukernes stemme alltid stå i sentrum. Denne datainnsamlingsmetoden gir oss også bedre innsikt i studentenes hverdag og kan forhåpentligvis gi en dypere forståelse for hvordan de opplever studiehverdagen. Studien la også til rette for at hver enkelt student kunne reflektere rundt hva som fungerer og ikke i en travel studiehverdag.

Gjennom prosjektets gang har vi samarbeidet med fem studenter fra Universitetet i Oslo som skulle skrive innlegg i hver sin dagbok med forhåndsdefinerte spørsmål i en periode på en uke. Spørsmålene skulle besvares hver kveld, og lød som følger:

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

## 5.2 Research through design

Zimmermann et al. (2007) foreslår i sin artikkel en modell for forskning innen interaksjonsdesign basert på konseptet *research through design*. Denne modellen har som mål at forskere innen design skal fokusere på å skape den *rette* løsningen; en løsning som skal transformere verden fra sin nåværende tilstand til en mer ønskelig tilstand (2007, s. 5). Forfatterne presenterer fire kriterier for å evaluere det akademiske bidraget til et forskningsprosjekt innen interaksjonsdesign: *Process* (prosess), *invention* (oppfinnelse), *relevance* (relevans), og *extensibility* (utvidbarhet). I kapittel 9 diskuterer vi om prosjektet vårt kan anses som et *research through design*-prosjekt.

En sentral del av *research through design* er å utforme en artefakt og bruke den aktivt i forskningen ved å måle effekt den har på omverdenen, samt å oppdage uforutsette konsekvenser. Gjennom prosjektet har vi utformet en chatbot-prototype og testet med et sett brukere, noe som vi går nærmere inn på i kapittel 7 og 8.

## 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 det 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

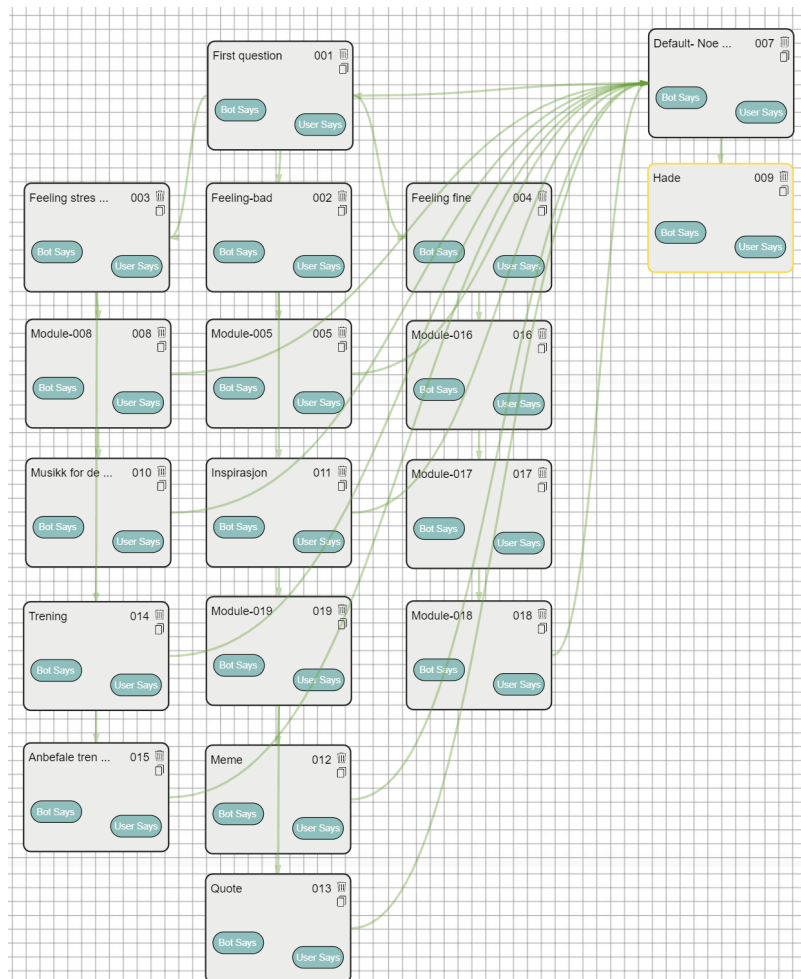


Figur 1: Demo av prototypen

Vi har laget en chatbot ved navn Kompis som vår prototype (Se figur 1). 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 forskjellig utfall i første møte (Se figur 2). 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”.





Figur 2: Flyten til chatboten i Chatteron.

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. Brukertesting

### 8.1 Gjennomføring

Vi gjennomførte tre brukertester med medstudenter vi hentet inn fra gangen. I brukertestene våre ønsket vi å se hvordan en student interagerer med chatboten vår, Kompis, og hvilke tanker studentene hadde rundt chatboten. Vi gjennomførte tre tester hvor hver student fikk en oppgave om å spille ut et humør. Dette gjorde vi for å teste de tre forskjellige flyt, som var mulig å teste chatboten med. Den første studenten fikk beskjed om å være glad, men ikke noe

mer. På denne måten satte vi visse rammer, men igjen la til rette for egne valg. Student nummer to fikk beskjed om å være negativ og sur. Dette tenkte vi kunne bli særlig interessant da vi ikke visste hvordan studenten ville interagere “sint” med chatboten, hvordan er det egentlig mulig? Siste og tredje student fikk beskjed om å være stresset. Dette tenkte vi ville være en lett mindset for en student å sette seg inn i, men det viste seg egentlig til å bli litt vanskelig da det ikke var et så klart skille mellom å være sur og stresset. Vi avsluttet hver test med et kort semi-strukturert intervju, hvor vi snakket om observasjoner vi hadde gjort oss, og om eventuelle tanker som studenten selv satt igjen med.

De planlagte spørsmålene vi stilte deltakerne etter observasjonene var:

- Hvordan synes du det var å chatte med chatboten?
- Gitt humøret du skulle spille, tror du chatboten kunne endret humøret ditt?
  - Hvorfor (ikke)?
- Tror du en mer avansert versjon av en slik chatbot vil kunne hjelpe deg i studiene?
- Er det noe som kunne gjort interaksjonen med chatboten bedre?

## 8.2 Funn

Etter brukertestene og de tre intervjuene satt vi igjen med noen tanker. Basert på det vi hadde erfart og lært virket det som at chatbot i seg selv var en bra tanke, og at to av tre studenter kunne tenkt seg å få litt mental oppfølging gjennom studiesemesteret. En av deltakerne sa følgende:

*“Det kan jo være litt kleint... eller vanskelig å snakke med noen andre om hvordan man har det osv. og da kan kanskje en chatbot være et alternativ kanskje.”*

Likevel skjønnte vi at det fortsatt var en lang vei å gå med tanke på interaksjon mellom menneske og chatbot. En ting som gikk igjen var at deltakerne følte på en form for distanse til interaksjonene fordi de var bevisst på at den ikke var “ekte”. Det virket som dette kunne gjøre motivasjonen mindre betydningsfull. Vi fikk blant annet følgende sitat:

*“Bryr meg på en måte ikke like mye om hva den (Kompis) sier når jeg vet at det er programmert.”*

Vi så også at det var alt for få valgmuligheter for å kunne føre en “normal” samtale, og det gikk for fort fra samtalen var startet og til at chatboten, Kompis, kom med forslag til hvordan hverdagen kunne bli litt bedre.

Vi satt igjen med en følelse av at retningen vi gikk i var riktig og at det finnes et behov for en slik type løsning, men og en følelse av at vi trenger betydelig større mengden data for å kunne lage en chatbot som faktisk støttet deg som en ekte Kompis.

## 9. Diskusjon

Dersom vi følger tankegangen til Følstad og Brandtzæg (2017) så er vi på vei inn i en tid hvor vi vil interagere med maskiner ved bruk av vårt daglige språk. Det er derfor naturlig å utforske muligheter og begrensninger innen teknologi som støtter slik funksjonalitet, og hvordan slike systemer kan nyttes på en positiv måte. Vi har selv undersøkt hvordan chatboter kan støtte studenter i deres hverdag i forhold til trivsel og praktiske gjøremål ved hjelp av dagbokstudie og brukertesting av en chatbot-prototype. Prototypen var veldig begrenset og skulle ideelt sett hatt et større datasett for å gi en mer realistisk opplevelse. Under brukertesten ble også deltakerne nokså begrenset ettersom de ble fortalt hvordan de skulle oppføre seg grunnet funksjonaliteten til prototypen, men det viktige var at vi fikk stimulert frem en samtale om hvorvidt en slik chatbot kunne fungert for dem eller ikke.

Resultatene våre viser at to av tre deltakere gjerne kunne tenkt seg mental oppfølging gjennom semesteret, men det ble ikke spesifisert om dette gjaldt menneskelig eller maskinstyrt støtte. Det kan diskuteres om chatboter er riktig måte å håndtere studenter sin psykiske helse på, da en av deltakerne nevnte at en maskin kunne få hen til å føle seg enda mer alene fordi man egentlig snakker med ingenting. Samtidig var det en deltaker som nevnte at det kan det være lettere å åpne seg når man ikke snakker med en ekte person.

Chatboten til K12 viser at det er stort potensiale i å benytte seg av chatbot-teknologi i utdanningssammenheng - denne baserer seg derimot på mer organisatoriske oppgaver heller enn emner om psykisk helse. Under Kocielnik et al. (2018) sitt samtaleagent-prosjekt trekkes det frem at deltakerne opplevde kommunikasjon gjennom chat lite personlig. Brukere av chatboter forenkler også ofte språket sitt fordi de forstår at funksjonaliteten er begrenset.

Sammensatt kan dette kan indikere at det vil være vanskelig å skape en genuin løsning som hjelper studenter på et dypere plan. Dagbøkene avdekket derimot at studentene ofte ikke trengte mer enn en enkel oppmuntring i hverdagen for å bli motivert, og det kan tenkes at en maskinstyrt løsning kan klare å generere en slik oppmuntring.

Dersom en chatbot skal vurderes å erstatte menneskelig kontakt bør den følge Amershi et al. (2019) sine retningslinjer for design av menneske-AI interaksjon for å unngå at det blir en ubehagelig opplevelse for brukerne. Dette er spesielt viktig for at brukerne skal føle seg trygge på at systemet behandler intim informasjon på en god måte, og for at de skal kunne vite hva tjenesten går ut på slik at det ikke blir skuffet eller frustrert over interaksjonen.

Basert på kriteriene til Zimmermann et al. (2007), mener vi at prosjektet til en viss grad kan betraktes som et *research through design*-prosjekt. Prosessen vår underveis har vært basert på en brukersentrert, iterativ designprosess, som vi mener er fullt mulig å reprodusere. Når det kommer til oppfinnelse kan det diskuteres om prosjektet har bidratt med en 'betydelig oppfinnelse', slik kravene til Zimmermann et al. (2007, s. 7) lyder, men vi mener at vi har valgt en ganske unik vinkling for den endelige prototypen, da det finnes få chatboter som er spisset mot studenter og deres studiehverdag. I følge Zimmermann et al. er ikke endemålet med et *research through design*-prosjekt å validere mot hypoteser, men heller å måle forskningens relevans opp mot virkeligheten. Vi har for dette prosjektet ikke satt noen konkrete mål for hvilken tilstand vi ønsker å oppnå med designet vårt, men vi mener det vil være fullt oppnåelig å sette forskningsspørsmålene våre opp mot en enda mer høyopløselig prototype. Når det gjelder krav om utvidbarhet mener vi at prosjektet delvis svarer til kriteriet. Det burde være mulig å basere videre eller annen forskning på prosjektet, men gitt mer tid til utforsking og videre bygging av prototype, kunne vi presentert mer håndfaste resultater som hadde skapt bedre grobunn for videre arbeid.

## 10. Konklusjon

I denne oppgaven har vi ønsket å undersøke de følgende forskningsspørsmålene:

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

Vi ser at det finnes potensiale i å bruke kunstig intelligens innen utdanning på bakgrunn av at andre har fått det til, men også basert på resultater fra vårt eget forskningsprosjekt hvor vi gjennomførte dagbokstudie og testet en chatbot-prototype med studenter. Resultatene viser at det er behov for mental oppfølging gjennom studiene, og at chatboter kan være nyttig som en veldig enkel samtalepartner som kan oppmuntre studenter i hverdagen. Samtidig kan det være vanskelig å lage et system som snakker med studentene på et dypere nivå. For at chatboter skal oppleves som mest mulig brukervennlig er det hensiktsmessig å følge Amershi et al. (2019) sine retningslinjer for design av menneske-AI interaksjon.

Ettersom prototypen vår er veldig begrenset vil det være verdt å teste en lignende løsning med et langt større datasett i en mer realistisk setting for å nærmere undersøke om en chatbot kan fungere like bra eller bedre enn menneskelig oppfølging under utdanning.

## 11. Appendix 1

### 9.1 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.

### 9.2 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.

## 12. 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.

### 10.1 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
```

## 10.2 Refleksjon Appendix 2

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
- Kocielnik, Rafael; Avrahami, Daniel; Marlow, Jennifer; Lu, Di; Hsieh, Gary. (2018). *Designing for Workplace Reflection: A Chat and Voice-Based Conversational Agent*. Hentet fra: <https://dl.acm.org/doi/abs/10.1145/3196709.3196784> Lest: 27.10.2021
- 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>
- Unesco. Artificial intelligence in education. Hentet fra: <https://en.unesco.org/artificial-intelligence/education> Lest 09.09.21
- Zimmerman, John; Forlizzi, Jodi; og Evenson, Shelley. (2007) *Research Through Design as a Method for Interaction Design Research in HCI*. Human-Computer Interaction Institute. Paper 41. Hentet fra: <http://repository.cmu.edu/hcii/41>