



Handelshøyskolen BI - campus Bergen

BTH 36301

Bacheloroppgave i markedsføring

Bacheloroppgave

Kunstig intelligens i rekruttering: Norske tech-ansattes reaksjoner på diskriminering

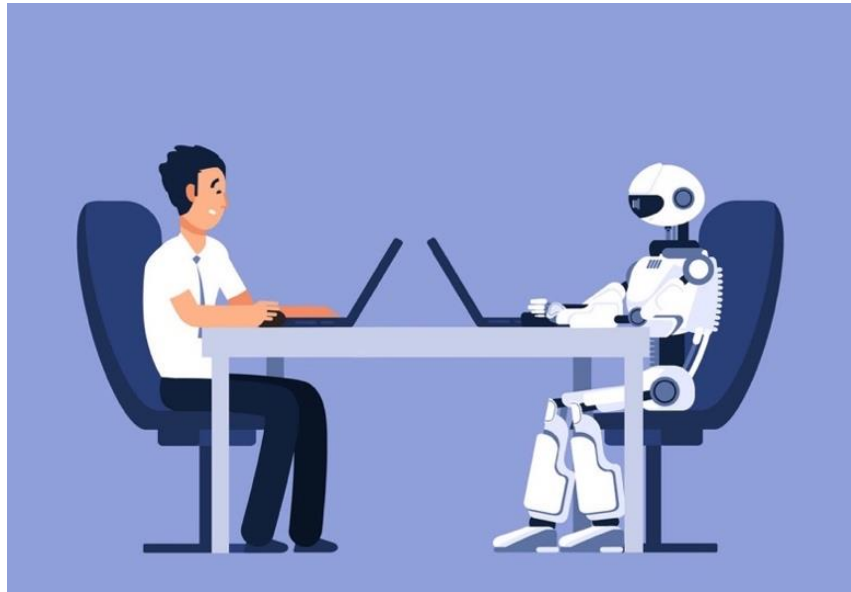
Navn: Adrian Sætrang, Selin Alexandra Ilic

Utlevering: 06.01.2020 09.00

Innlevering: 03.06.2020 12.00

KUNSTIG INTELLIGENS I REKRUTTERING:

Norske tech-ansattes reaksjoner på diskriminering



(<https://www.vectorstock.com/royalty-free-vector/robot-and-businessman-robots-vs-human-future-vector-21303527>)

Bacheloroppgave

Eksamenskode: BTH 36301

6. semester, Handelshøyskolen BI campus Bergen

“Denne oppgaven er gjennomført som en del av studiet ved Handelshøyskolen BI. Dette innebærer ikke at Handelshøyskolen BI går god for de metoder som er anvendt, de resultater som er fremkommet, eller de konklusjoner som er trukket.”

Antall ord: 11.290

Sammendrag

Hensikten med denne oppgaven er å se nærmere på bruken av kunstig intelligens i rekruttering, hvor vårt hovedfokus ligger på reaksjoner knyttet til kjønnsdiskriminering. Vi valgte å undersøke et utvalg med kunnskap om teknologi, og derfor ble vårt forskningsspørsmål:

«Hvordan reagerer norske tech-ansatte på diskriminering i rekrutteringsprosesser når denne utføres av en kunstig intelligens?»

For å besvare dette spørsmålet ble det formulert en hypotese hvor vi påstår at det blir målt mindre moralsk forargelse når en diskriminering blir gjort av en kunstig intelligens sammenlignet med et menneske. Studien ble gjennomført som et eksperiment, hvor vi ønsket å måle forskjell i moralsk forargelse blant respondentene. De ble delt i to grupper med hver sin tilhørende case som omhandlet diskriminering av kvinner, hvor forskjellen var hvem som utførte diskrimineringen – en HR-spesialist eller en algoritme.

Resultatene viser at tech-ansatte blir mindre moralsk forarget ved diskriminering gjort av kunstig intelligens. De ser også løsninger for hvordan feilen kan rettes opp fremfor å skrote hele algoritmen. Dette kan knyttes opp mot deres kunnskap om teknologi. Vi kan konkludere med at det med fordel kan benyttes kunstig intelligens i rekrutteringsprosesser og at man på denne måten kan redusere diskriminering. Bruk av menneskelige ressurser og etiske problemstillinger som vi presenterer i oppgaven må likevel tas i betraktning for å oppnå rettferdig rekruttering.

Forord

Denne bacheloroppgaven er vår avslutning på tre givende år ved Handelshøyskolen BI Bergen. Det har vært et annerledes og krevende semester på mange måter, men vi avslutter det med et positivt sinn og stolthet over alt vi har fått utrettet.

Vi har gjennom disse årene vært innom mange spennende tema, og med tiden utviklet en ekstra interesse for feltet organisasjon og ledelse. Dagens digitale utvikling har gjort at vi også har sett verdien i å sette oss mer inn i temaer som teknologi, og problemstillinger knyttet til dette. Det fanget derfor vår oppmerksomhet da førsteamanuensis Mads Nordmo Arnestad presenterte spennende eksperimenter for bacheloroppgaver denne våren, hvor nettopp disse temaene var blant dem.

Prosessen med å skrive og samarbeide om oppgaven ble naturligvis annerledes enn vi hadde sett for oss, men vi har tatt lærdom av utfordringene og kommet sterkere ut av det. Heldigvis er dagens teknologiske løsninger gode, så ingenting har stått i veien for svært hjelpsomme veiledninger, noe som har vært helt avgjørende for oppgaven vår. Temaene var på mange måter helt nye for oss, og det var mye å sette seg inn i for å kunne ha et godt grunnlag for forskningen.

Vi vil først og fremst takke alle bedriftene som tok seg tid til å gjennomføre undersøkelsen vår og gitt oss interessante og relevante data. Takk til alle forelesere og administrasjonen på BI, som står på for at vi skal kunne tilegne oss lærdom og inspirasjon. Til slutt vil vi gi en stor takk til vår veileder Mads Nordmo Arnestad som introduserte oss for dette spennende eksperimentet, og for hans løftende engasjement for oppgaven.

Bergen, 01.06.20

Innholdsfortegnelse

SAMMENDRAG	II
FORORD	III
1.0 INNLEDNING	6
1.1 DISKRIMINERING.....	6
1.1.2 <i>Mannsdominans i tech-bransjen</i>	7
1.2 FORSKNINGSSPØRSMÅL.....	8
1.3 REKRUTTERING.....	8
1.4 KUNSTIG INTELLIGENS I REKRUTTERING.....	10
2.0 TEORI OG HYPOTESE	12
2.1 FREMVEKSTEN AV KUNSTIG INTELLIGENS.....	12
2.1.1 <i>Maskinlæring</i>	13
2.1.2 <i>Etiske utfordringer</i>	13
2.1.3 <i>Nasjonal strategi</i>	14
2.1.4 <i>Psykologiske konsekvenser</i>	14
2.2 MORALSK FORARGELSE.....	15
2.3 MOTIVASJON.....	16
2.4 HYPOTESE.....	16
3.0 METODE	17
3.1 FORSKNINGSDESIGN.....	17
3.1.1 <i>Kvantitativ og kvalitativ metode</i>	17
3.1.2 <i>Vignetteметоден</i>	18
3.1.3 <i>Primær og sekundærdata</i>	18
3.2 REKRUTTERING OG UTVALG.....	19
3.3 EKSPERIMENTDESIGN.....	20
3.4 UTFALLSMÅL.....	21
3.5 VALIDITET OG RELIABILITET.....	22
4.0 RESULTAT	23
4.1 DESKRIPTIVE DATA.....	23
4.2 KUNNSKAP OM TEMAET.....	24
4.3 SAMMENHENG MELLOM VARIABLENE.....	24
4.4 TEST AV MELLOMGRUPPEFORSKJELLER.....	25
4.4.1 <i>T-test</i>	25
4.5 RESPONDENTENES KOMMENTARER.....	25
5.0 DISKUSJON	27
5.1 TEORETISKE IMPLIKASJONER.....	27
5.2 PRAKTISKE IMPLIKASJONER.....	28
5.3 BEGRENSNINGER OG ANBEFALINGER.....	30
6.0 KONKLUSJON	32
7.0 REFERANSER	33
8.0 VEDLEGG	40

I en verden som er under stadig endring er det viktig for selskaper å være tilpasningsdyktige. En sentral forutsetning for å være det er blant annet å ha de riktige menneskene i organisasjonen. I tillegg må bedriften være attraktiv nok for å kunne tiltrekke seg nye og dyktige medarbeidere, som er utfordringer enhver bedrift står overfor (Goldstein et al., 2017; Kuvaas & Dysvik, 2017). Studier fra McKinsey viser at selv for jobber med lav kompleksitet så vil en toppkvalifisert ansatt være 50% mer produktiv enn en gjennomsnittlig medarbeider (Keller & Meaney, 2017). For jobber med høy kompleksitet vil de yte 125% mer, og er det en svært komplisert jobb vil forskjellen i produktivitet være hele 800%. Selskaper som derfor kan finne og tiltrekke seg den beste arbeidskraften vil ha et stort konkurransefortrinn foran andre i samme bransje (Goldstein et al., 2017, s. 3). Tidligere var de mest verdifulle globale selskapene industrigiganter som Exxon og General Motors (Johnston, 2018), mens det nå preges av teknologiselskaper som Apple, Google og Microsoft. På 60-tallet hadde General Motors rundt 600 000 ansatte med inntekter tilsvarende 12 milliarder dollar årlig (Gulliford & Parker Dixon, 2019). Til sammenligning var Apple i 2017 verdens mest verdifulle selskap (Interbrand, 2017) og oppnådde inntekter på 800 milliarder dollar i USA alene, med bare 116 000 ansatte (Gulliford & Parker Dixon, 2019). Den teknologiske utviklingen har gjort at kvaliteten på de ansatte derfor har stor betydning for veksten og inntjeningen til selskapet.

Det kreves mye ressurser for å finne og tiltrekke seg de beste kandidatene. Mange selskaper mottar ofte flere hundre søknader på hver stilling (Black & van Esch, 2020), og det å gå gjennom hver enkelt søknad og CV blir svært tidkrevende. Store selskaper kan oppleve å få mange flere enn dette også, for eksempel får Google hvert år rundt to millioner søknader til sine kontorer i USA (Torres, 2017). At det går an å levere CV og søknad digitalt har ført til at mellom 75% og 88% av kandidatene egentlig er underkvalifiserte (Ideal, 2020), men likevel velger å søke fordi det bare krever noen tastetrykk. Skal selskapene da klare å finne frem til de beste hodene for å holde seg konkurransedyktige, trengs det i utgangspunktet enorme HR-avdelinger for å gå gjennom alle CV-ene. Teknologi innen kunstig intelligens (KI) vil derimot kunne gjøre denne jobben veldig mye raskere, i tillegg til å kunne opptre helt fordomsfritt og rettferdig. At en algoritme kan vurdere kandidaters bakgrunn helt

subjektivt, fritt for diskriminering, uten noen form for påvirkning fra andre høres i utgangspunktet ut som en god idé. Det finnes likevel flere eksempler på at dette ikke alltid er tilfelle.

I juli 2017 sendte en Google-ansatt ut en intern e-post hvor han hevdet at kjønnsulikheter i teknologibransjen kunne skyldes biologiske forskjeller. Dette dokumentet spredte seg raskt og førte til en offentlig debatt som til slutt endte med at den ansatte fikk sparken fra selskapet en måned senere (Wakabayashi, 2017). Bare ett år etterpå kom en annen historie om diskriminering fra tech-bransjen frem. Det ble kjent at en algoritme, basert på maskinlæring utviklet av Amazon for å gjennomgå kandidaters CV-er til utviklerstillinger, systematisk valgte bort kvinner. Den ga en negativ vektning på CV-er som inneholdt ord som for eksempel «kaptein for damelaget i sjakk» og vitnemål fra ulike kvinneuniversiteter. Denne algoritmen ble senere skrotet av Amazon, som på sin side hevdet at den aldri hadde blitt brukt helt uavhengig uten tilsyn fra mennesker. Meningen med algoritmen var at det bare var for å supplere arbeidet til de ansatte som gikk gjennom søknadene (Dastin, 2018). Denne saken om diskriminering fra Amazon fikk mye mindre oppmerksomhet i media enn det Google opplevde da den interne e-posten ble kjent. Vi ble dermed nysgjerrige på rekrutteringsprosesser gjort av en KI, og stilte på bakgrunn av disse hendelsene spørsmål ved hvorfor det var så ulike reaksjoner på disse diskrimineringene. Den eneste forskjellen var jo i prinsippet bare hvem som utførte handlingene.

1.0 Innledning

1.1 Diskriminering

Blir noen behandlet mindre gunstig enn andre snakker vi om diskriminering (Ikdahl, 2018). Diskriminering kan oppleves i ulike former, men i denne oppgaven fokuserer vi på kjønn. For eksempel kan kvinner med like eller bedre kvalifikasjoner enn menn bli mindre foretrukket basert på kjønn, og da er det nærmere bestemt kjønnsdiskriminering (Aas & Ikdahl, 2019). Dette er ikke noe nytt, og er heller ikke

begrenset til uhell knyttet til KI. Gjennom historien har vi sett at jobber som blir ansett som viktigere eller bedre betalt, har en tendens til å sile ut kvinner. Derav også årsaken til at det er et behov for et beskyttende lovverk som sikrer likestilling (Hicks, 2018). Det finnes i dag en rekke lover som skal sørge for at diskriminering ikke skjer og at personopplysninger håndteres på en forsvarlig måte (Iversen, 2015, s. 30–31). Til tross for dette og at kvinners posisjon i arbeidslivet har endret seg opp gjennom årene, eksisterer det fortsatt betydelige ulikheter. Kjønsdiskriminering antas å være en sannsynlig årsak til dette (Heilman & Caleo, 2018).

1.1.2 Mannsdominans i tech-bransjen

Flere rapporter viser til at andelen kvinner i tech-selskaper er lav verden over. Blant annet en undersøkelse utarbeidet av Kantar TNS for Oda Nettverk, et ledende nettverk for kvinner i tech, viser at det i Norge i 2018 var en total andel av kvinner i bransjen på 28%. Dette er den høyeste andelen som er målt og viser at utviklingen går i riktig retning, men at det likevel er mye som gjenstår (Kantar TNS, 2018). En annen rapport vi fant spesielt interessant er skrevet av Plan International Norway under et samarbeid med Telenor, kalt Plan Telenor Global Partnership. Denne rapporten handler om kjønnsgapet i teknologibransjen i Skandinavia, og viser at vi har et av de største kjønnsgapene i hele verden når det kommer til valg av utdanning og karriere innen teknologi eller STEM (Science, Technology, Engineering, Mathematics). I en undersøkelse de foretok ble det spurt hvilke utfordringer jenter og gutter møter hvis de vil ha en utdanning eller karriere innenfor teknologi/STEM, og 34,88% av 172 kvinnelige respondenter svarte at kjønnsbasert diskriminering er en nøkkelfaktor som begrenser unge kvinner å søke en fremtid innen tech. Nesten 80% mener at det finnes stereotypiske holdninger om hva jenter og gutter kan og burde gjøre (Plan International Norway, 2019). Dette er oppsiktsvekkende når vi vet at de nordiske landene ofte blir sett på som de mest fremoverlente og kjønnsbalanserte landene i verden (Erlandsson, 2019). Også i den britiske teknologibransjen rapporterer halvparten av kvinnene at de har opplevd diskriminering, som igjen påvirker antall kvinner som velger en karriere innen STEM (P. Agarwal, 2020).

1.2 Forskningsspørsmål

Funnene over gjør det interessant å se hvordan de som jobber i nettopp den norske tech-bransjen stiller seg til denne type kjønnsdiskriminering, og leder oss inn på vårt forskningsspørsmål:

«Hvordan reagerer norske tech-ansatte på diskriminering i rekrutteringsprosesser når denne utføres av en kunstig intelligens?»

Grunnen til relevansen er først og fremst den sterke veksten i bruk av teknologi og kunstig intelligens i samfunnet. Vi ser at områder i næringslivet, som for eksempel rekruttering, i større grad blir preget av denne utviklingen. Alt dette, kombinert med et vedvarende fokus på likestilling og hvordan vi skal takle denne type utfordringer, gjør at vi anser temaet som dagsaktuelt og viktig. Det er allerede gjort en studie på hvor moralsk forarget folk blir dersom diskriminering skjer, og hvilke forskjeller som slår ut dersom diskrimineringen er gjort av en KI sammenlignet med et menneske. Det studien viser er at vanlige folk blir *mer* forarget av den samme handlingen dersom det er et menneske som har gjort det (Bigman et al., 2020). Vi ønsket derfor å se om funnene lot seg replikere i et utvalg bestående av norske tech-ansatte. I tillegg til at den målgruppen gjerne har mer kunnskap om teknologi enn folk flest, er de også som nevnt en del av en bransje som er kjent for å være svært mannsdominert, noe som gjør at de er en interessant gruppe å studere nærmere.

1.3 Rekruttering

Rekruttering dreier seg om å finne den kandidaten som er best skikket til en ledig stilling basert på analyser av bedriftens personalbehov, strategi og ressurser. Samtidig er dette en prosess som handler om å anskaffe et mangfold av kompetente kandidater til å søke om å bli ansatt i selskapet (Eriksen, 1996). Rekruttering og ansettelse av nye medarbeidere er således en av de viktigste oppgavene en leder har, og det er kritisk både for lederen og bedriftens konkurranseevne å tiltrekke seg de beste kandidatene (Iversen, 2015, s. 9–10). For å vite hvordan man kan tiltrekke seg den

beste arbeidskraften er det viktig at man forstår hvilken type arbeidskraft som faktisk trengs (Yu & Cable, 2012). Da blir det sentralt å definere hvilke stillinger som vil være strategisk viktige for å oppnå et konkurransefortrinn som vil skape vekst. Dette er typisk stillinger som krever mye spesifikk kunnskap og ekspertise (Boudreau & Ramstad, 2005).

Om selskapet velger å rekruttere en kandidat med helt gjennomsnittlig grad av kunnskap mister de muligheten til et konkurransefortrinn. Velger de i stedet å gå grundigere til verks ved å sette krav til ferdigheter hos kandidaten og overvåke rekrutteringen mer nøye, vil utfallet kunne gi større gevinst over tid (Yu & Cable, 2012). Ansetter bedriften for eksempel en god leder kan dette få en hel avdeling til å blomstre, mens en dårlig leder kan virke ødeleggende for de ansatte som er involvert. Kostnaden ved en feilansettelse vil også potensielt være høyere enn den ekstra kostnaden en mer kvalifisert ansatt vil kreve av ressurser (Bartlett & Ghoshal, 2002). Dette grunnet tapt tid, manglende effektivitet og tapte lønnskostnader i tillegg til menneskelig- og arbeidsmiljømessige konsekvenser (Eriksen, 1996). Det er også viktig å være klar over at bedriftens omdømme og attraktivitet gir føringer for søkermassen. Dette gjelder søkermassens omfang men også kvaliteten av søkerne. Et godt omdømme vil gi flere og bedre søkere, og er således noe bedriften burde fokusere på (Kuvaas & Dysvik, 2017, s. 91–92).

Sortering av søknad og CV er bare starten av den innledende delen av en rekrutteringsprosess, som viderefølges av et intervju (Iversen, 2015, s. 71). Forventningene ved et jobbintervju er at det kan avdekke andre forhold enn det som kommer frem gjennom søknad og CV, og er den mest anvendte rekrutteringsmetoden i norske virksomheter. Det blir likevel stilt spørsmål ved jobbintervjuets validitet, da særlig med tanke på at det her ligger en sårbarhet for stereotypier og feilslutninger (Rogstad & Sterri, 2018). I arbeidet med rekrutteringsprosesser kan det oppstå bevisst eller ubevisst favorisering og fordommer hos den som rekrutterer, og det hevdes at dette er den største feilkilden under et intervju (Iversen, 2015, s. 76). For eksempel kan kvinner bli valgt bort av menn fordi de blir sett på som mindre produktive om de

allerede er, eller kanskje skal bli mødre. Dette kan da være en bevisst fordom som følge av en tankegang om at kvinner ofte kutter ned på arbeidsmengden når de går inn i morsrollen. En mer ubevisst fordom er det som kalles likhetseffekten. Dette går ut på at man har lettere for å like en som har en lignende bakgrunn og personlighet som seg selv, og forskning viser at dette ofte påvirker utfallet i en rekrutteringsprosess (Erlandsson, 2019; Fatemi, 2019; Iversen, 2015, s. 79). Studier viser også at det ofte diskrimineres til fordel for medlemmene innad i gruppen, hvor kjønn er en av de viktige grupperingene (Brewer & Kramer, 1985).

1.4 Kunstig intelligens i rekruttering

Det finnes en rekke forskjellige definisjoner på hva kunstig intelligens faktisk er, men en populær definisjon er at det dreier seg om utviklingen av maskiner som kan handle på menneskelignende måter innenfor områder som læring, resonnement og selvkorreksjon (Kok et al., 2002). Som et resultat av dette kan utstrakt bruk av denne type teknologi være med på å løse utfordringene knyttet til blant annet effektivitet og diskriminering i rekrutteringsprosesser (Black & van Esch, 2020).

I en rekrutteringsprosess er gjennomgangen av søknader det mest tidkrevende. Verdens største kosmetikkprodusent L’Oreal tok i bruk kunstig intelligente programmer for å gjennomgå CV-er, og opplevde at tiden det tok å gå gjennom en CV og innkalle til intervju gikk fra 45 til 4 minutter, en reduksjon på over 90% (Sharma & Ltd, 2018). Hotellkjeden Hilton Hotels & Resorts tok i bruk et lignende program, og så at tiden det tok å ansette kandidater gikk fra 42 til 5 dager, en 88% nedgang (Mclaren, 2018). Dette viser at om et selskap bruker kortere tid på en rekrutteringsprosess kan det få et konkurransefortrinn ved å kunne tilby de beste kandidatene jobb tidligere enn det konkurrentene klarer, samtidig som de bruker mindre ressurser. Kunstig intelligente verktøy har i tillegg gått fra å bare fange opp nøkkelord i søknader og CV-er, til å kunne oppdage ferdigheter som ikke nødvendigvis står skrevet i søknadsteksten. Algoritmene kan eksempelvis antyde personlighetskarakteristikken «standhaftighet» gjennom setninger som omhandler at

kandidaten ikke har gitt seg når han/hun har møtt motstand (Black & van Esch, 2020).

Det finnes også eksempler på kunstig intelligente rekrutteringsverktøy som faktisk har bidratt til at kvinneandelen i selskapet har gått opp. Evry lanserte en rekrutteringsrobot som målte kandidatenes personlige egenskaper knyttet til blant annet kontinuerlig utvikling, læring, matematikk og logisk resonnering. Hver kandidat var helt anonym, så grunnlaget for om de kom videre til neste steg i prosessen var utelukkende basert på objektive kriterier (Gulbrandsen, 2017). Ved bruken av dette verktøyet var ikke kjønn og bakgrunn noe som kom frem før slutten av prosessen, hvor deltakerne allerede hadde bevist sine ferdigheter. Dette var med på å gjøre hele rekrutteringen mye mer rettferdig.

Andre områder hvor KI kan utgjøre en forskjell i rekrutteringsprosessen er for eksempel når en kandidat først har kommet til intervju. Intervjuet kan gjøres over video hvor KI kan analysere ansiktsuttrykk, kroppsspråk, talemåte og tilstedeværelse. Dette brukes for å identifisere løgn, kognitive egenskaper og profesjonalitet (Bersin, 2018; Linas, 2019). Denne utviklingen kalles Natural Learning Processing (NLP) (Laurent et al., 2017). Slike systemer kan i tillegg vurdere kandidater etter hvordan de presterer i intervjuet sammenlignet med suksessfulle og gjennomsnittlige ansatte som har hatt stillingen tidligere (Black & van Esch, 2020). Ved flere hundre søkere på en stilling vil det også være vanskelig å gi alle personlige tilbakemeldinger underveis i prosessen, og en chatbot som genererer automatiske tilbakemeldinger vil derfor være en stor ressurs i dette arbeidet. Kunstig intelligens kan i tillegg brukes for å markedsføre riktige stillinger til aktuelle kandidater i de riktige kanalene, analysere hva som blir svart i ulike tester og avtale videre intervjuer (D. Agarwal et al., 2018; Tyfting, 2019).

For at et selskap skal kunne finne de beste menneskene er det også nødt til å vurdere flere enn de kandidatene som har sendt inn søknad (Guinan et al., 2014). Det finnes

kunstig intelligente verktøy som kan samle data fra sosiale medier, ulike forum og nettsider for å finne mennesker som ikke i utgangspunktet er på utkikk etter ny jobb. Disse har likevel vist frem ferdigheter og kunnskap, og kan med det vise seg aktuelle for ledige stillinger (Campbell et al., 2020). Undersøkelser viser at 80% av de som ikke er på utkikk etter ny jobb, likevel vil vurdere et passende jobbtilbud om de får det presentert. Det kan være med på å påvirke grunnen til at det er cirka tre ganger så mange passive kandidater som aktive (Black & van Esch, 2020; Fatemi, 2019). Dette viser at om selskaper får kjennskap til passive kandidater som kan være interesserte i stillinger, har de mulighet til å sikre seg enda bedre kvalifiserte ansatte enn de som opprinnelig søkte på stillingen. Dette vil øke sjansen for en enda større avkastning fra denne ansettelsen.

Den økte betydningen av kvaliteten på ansatte i organisasjoner, gjør at rekruttering vil bli satt enda høyere på agendaen i tiden fremover (Tudow et al., 2019). Når antall søknader på hver stilling bare blir høyere, er det viktig å få behandlet disse på mest mulig effektiv, god og rettferdig måte. Likevel er det bare 38% av de største selskapene som aktivt bruker KI i rekrutteringen (D. Agarwal et al., 2018), noe som tyder på at det fortsatt er mange bedrifter som har mye å gå på i effektiviseringen av rekrutteringsarbeidet.

2.0 Teori og hypotese

2.1 Fremveksten av kunstig intelligens

De siste tiårene har vi opplevd en markant økning i integreringen av autonome maskiner og algoritmer i samfunnet (Zhao, 2018). Mange oppgaver som tidligere ble utført av mennesker blir nå gjort av disse algoritmene og maskinene. Dette er noe som øker for hver dag som går. Teknologien vil frigjøre tid for de ansatte til å prioritere oppgaver som trenger mer kognitive egenskaper, og som ikke KI er avansert nok til å ta over (enda). Kunstig intelligens har til nå vært brukt på en rekke områder i næringslivet og har ved flere anledninger vist seg å være mer effektivt enn mennesker. Dette har vært på oppgaver som for eksempel styring av inventar

(Cárdenas-Barrón et al., 2012), kundeservice (Cheung et al., 2003), selvbetjente kasser i matbutikker (Aquilina & Saliba, 2019) og rekrutteringsavgjørelser (Faliagka et al., 2012).

2.1.1 Maskinlæring

Maskinlæring har fasilitert for denne utviklingen, og er en egen gren innenfor kunstig intelligens. Teknologien skaper antakelser basert på innsamlet data på en måte som ofte har vært bedre enn det mennesker har evnet å komme opp med. Dette handler om datasystemers evne til å kunne forbedre seg selv, lære mønster og ta beslutninger, uten at det er avhengig av forhåndsprogrammerte instruksjoner fra mennesker (SAS, 2020). Når systemet kjenner til ulike mønstre av store mengder data, kan dette brukes til å predikere forskjellige scenarier og utfallet av disse. Predikeringene vil bli mer nøyaktige desto mer data systemet analyserer, og er derfor avhengig av at systemet kontinuerlig innhenter informasjon (Murphy, 2012). Maskinlæringen er bygd opp av algoritmer som er fullstendige og nøyaktige beskrivelser av fremgangsmåten for løsninger av ulike oppgaver eller beregningsoppgaver (Hovde & Grønmo, 2020). Selv om veksten av algoritmer basert på maskinlæring fører med seg en rekke spennende muligheter, er det også bekymringer knyttet til dette.

2.1.2 Etske utfordringer

Et sett av bekymringer knyttet til KI er de etiske problemstillingene som oppstår. Etikk er et sentralt tema for mange som jobber med tech, og den svenske KI-eksperten Fredrik Heintz som er forsker på Linköpings Universitet, mener at etikken bak er like viktig som teknologien i seg selv. Han sier at det trolig trengs like mye forskning på det filosofiske og juridiske plan som på teknologien for at den skal fungere i samfunnet (Dustin, 2018). Et spørsmål kan for eksempel være om vi skal tillate algoritmer å ta avgjørelser i det moralske domenet (Bigman & Gray, 2018; Conitzer et al., 2017; Young & Monroe, 2019). Et annet spørsmål tar for seg om det skal programmeres etiske retningslinjer inn i algoritmene (Awad et al., 2018; Conitzer et al., 2017; Malle et al., 2016). I tillegg ser vi at fordelene ved KI også

bringer med seg skyldspørsmål når KI forårsaker skade (Bigman et al., 2019; Malle et al., 2019; Shank, DeSanti, et al., 2019).

Ifølge KI-forskere er det ingen distinksjoner mellom sinnet og maskiner. Både det og det å tilskrive maskiner mentale egenskaper bringer opp moralske spørsmål (Nath & Sahu, 2020). Teknologi gir imidlertid muligheten til å programmere roboter slik at de holder seg til bare tall og fakta. Som et resultat av dette viser flere til en oppfatning om at bruk av mekaniske sinn kan redusere korrupsjon, bestikkelser og annen umoralsk oppførsel som kan oppstå i virksomheten. En sentral etisk utfordring blir dermed hvordan man skal implementere beslutningsprosedyrer som også er etisk forsvarlige (Kvalnes, 2015).

2.1.3 Nasjonal strategi

I januar 2020 la regjeringen frem en nasjonal strategi for kunstig intelligens og bruken av denne i samfunnet. Her kommer de med forslag til hvordan algoritmer kan kontrolleres, blant annet ved å åpne dem for innsyn eller revisjon, eller bygge inn personvern og etiske hensyn helt fra start. Da vil det for eksempel være viktig å se på om algoritmen kan føre til diskriminering og om den lar seg manipulere (Kommunal- og moderniseringsdepartementet, 2020). Dette indikerer at det blir lagt opp til et økt fokus på etikk i bruken av KI, og ikke minst at man må være særlig oppmerksom på tilfeller av umoralske handlinger som diskriminering og at dette unngås.

2.1.4 Psykologiske konsekvenser

Ser vi forbi disse etikkspørsmålene som vil være med på å bestemme hvordan KI vil endre samfunnet vi lever i, er det viktig å forstå de psykologiske konsekvensene av den utstrakte bruken av KI. Hvordan vil folk reagere når en algoritme tar avgjørelser som vi mennesker ser på som umoralske? Et eksempel på en umoralsk handling er når KI diskriminerer folk på bakgrunn av kjønn og rase. Det er bevist at algoritmer diskriminerte kjønn i anbefalinger om prøveløslatelse (Angwin et al., 2016), helseforsikring (Obermeyer et al., 2019) og kredittvurderinger (Stankiewicz, 2019).

For å finne ut mer om hvordan mennesker reagerer på tilfeller hvor algoritmer diskriminerer basert på faktorer som rase og kjønn, er vi nødt til å forstå hvordan folk reagerer på diskriminering generelt.

2.2 Moralsk forargelse

Et kraftig verktøy vi har innenfor menneskelig psykologi, som brukes til å bekjempe diskriminering og andre typer av umoralske handlinger, er moralsk forargelse (Batson et al., 2007; Russell & Giner-Sorolla, 2011; Spring et al., 2018). Moralsk forargelse mobiliserer mennesker til å straffe urettferdig behandling (Fiske & Tetlock, 1997; Gummerum et al., 2016; Nelissen & Zeelenberg, 2009), boikotte uetiske selskaper (Lindenmeier et al., 2012) og fremmer kollektiv handling (Martin et al., 1984; Miller et al., 2011). Moralsk forargelse utviklet seg for å svare på menneskelig handling, men ettersom stadig flere avgjørelser som kan resultere i diskriminering gjøres av algoritmer, er det viktig og interessant å undersøke om folk vil være like forarget når det er maskiner som diskriminerer. Vi foreslår at forargelse ved diskriminering av algoritmer vil avhenge av de psykologiske årsakene til moralsk forargelse.

Vi vet at diskriminering er skadelig og urettferdig for dem som blir diskriminert. Disse menneskene sitter mer i fengsel, har vanskeligere for å få gode jobber og får dessuten mindre betalt for dem. De får heller ikke den hjelpen de fortjener fra helsevesenet og bankene ser på de med en høyere risiko (Bigman et al., 2020). Dersom den eneste grunnen til at folk opplever forargelse over diskriminering er denne skaden og urettferdigheten, ville vi ikke forvente at identiteten til den som utfører den diskriminerende handlingen har noe å si. Mennesker bør dermed være like forarget over diskriminering gjort av mennesker og diskriminering gjort av algoritmer. Det kan likevel ha seg slik at denne skaden ikke er den eneste grunnen til at man kjenner på en forargelse over diskriminering (Goodwin et al., 2014; Uhlmann et al., 2015). Når man vurderer slike hendelser er det ikke bare utfallet man tar stilling til, men også handlingspersonens karakter. Derfor kan man kjenne på forargelse når en person har avslørt en feilaktig moralsk holdning, da spesielt om den

diskriminerende handlingen viser at personen er motivert av fordommer (Bigman et al., 2020).

2.3 Motivasjon

Forskning innen moralpsykologi viser at når man dømmer en beslutningstaker, tar man også i betraktning motivasjonen som ligger bak beslutningstakerens oppførsel (Bigman & Tamir, 2016; Levine & Schweitzer, 2014; Reeder et al., 2002). Vi foreslår at motivasjon på samme måte kan påvirke forargelsen man opplever ved diskriminering. Når en som diskriminerer er motivert av fordommer, kan altså folk oppleve å bli mer forarget enn når det ikke foreligger en motivasjon.

Mennesker oppfatter de mentale tilstandene til algoritmer på en annen måte enn de mentale tilstandene til mennesker (Bigman & Gray, 2018; K. Gray & Wegner, 2012; Malle et al., 2016). De blir sett på som i mindre stand til å tenke rasjonelt og planlegge sine handlinger, mindre i stand til å ha følelser (Bigman & Gray, 2018; H. Gray et al., 2007) og har mindre fri vilje enn mennesker (Weisman et al., 2017). Måten folk oppfatter roboter på, påvirker hvor mye de stoler på dem (Gogoll & Uhl, 2018), klandrer dem (Malle et al., 2016; Shank & DeSanti, 2018), og vil at de skal ta beslutninger (Bigman & Gray, 2018; Young & Monroe, 2019).

2.4 Hypotese

Basert på teoriene og forskningen presentert foreslår vi at også tech-ansatte vil ha mindre sannsynlighet for å tilskrive en forhånds dømt motivasjon til en diskriminerende algoritme enn et diskriminerende menneske. Dette gjør at når diskriminering blir utført av en algoritme blir de mindre forarget. På bakgrunn av dette lyder vår hypotese slik:

H1: Norske tech-ansatte blir mindre moralsk forarget når diskriminering blir utført av en algoritme i stedet for et menneske.

3.0 Metode

I denne delen av oppgaven vil vi beskrive hvordan vi gjennomførte vår studie, ved å blant annet gå gjennom fremgangsmåten for rekrutteringen, utvalg og eksperimentdesignet. En metode vil si en planmessig fremgangsmåte for å fremskaffe kunnskap om hvordan vi ønsker å gå frem i en forskningsprosess. Grunnen til at man foretar en undersøkelse vil alltid være at man ønsker å få svar på et spørsmål, og få bekreftet eller avkrefte en antakelse (Gripsrud et al., 2016, s. 15).

3.1 Forskningsdesign

Vi ville utforske om det var forskjell i reaksjon hos tech-ansatte når en KI og et menneske utfører diskriminering, og dermed valgte vi et kausalt design for oppgaven. Denne typen design er relevant når man skal se på sammenhenger mellom variabler (Gripsrud et al., 2016, s. 54). Ved et slikt design benyttes en form for eksperiment, hvor meningen er å manipulere de uavhengige variablene. Basert på dette kan vi se om de har en effekt på den avhengige variabelen og trekke en konklusjon.

3.1.1 Kvantitativ og kvalitativ metode

Det finnes to ulike måter å samle inn data til forskningsoppgaven, henholdsvis kvalitativ og kvantitativ metode. Kvalitative data blir innhentet når man ønsker å gå i dybden for å virkelig forstå fenomenet man forsker på. Teknikker ved kvalitativ metode kan være dybdeintervjuer og fokusgrupper hvor man plukker ut representanter fra utvalget for å lære mer om forskningsspørsmålet (Gripsrud et al., 2016, s. 103). I denne oppgaven ble det bestemt å gå for en kvantitativ tilnærming ved å opprette en spørreundersøkelse. Dette brukes for å forklare fenomener og sammenhenger mellom dem (Fangen & Sellerberg, 2011, s. 71).

Data som innsamles i kvantitativ metode kan kategoriseres ettersom de er tallfestet. For å samle inn dette valgte vi å dele undersøkelsen elektronisk via Qualtrics. Når like spørsmål blir stilt til mange personer innenfor målgruppen sikrer vi oss bredde, i tillegg til at svarene blir kvantifiserbare slik at vi kan trekke statistiske sammenhenger

(Gripsrud et al., 2016). Vi ønsket å måle grad av moralsk forargelse og var derfor avhengig av data som var kvantifiserbare.

3.1.2 Vignetteметоден

Spørreundersøkelsen ble gjennomført ved et randomisert, kontrollert eksperiment. Dette var et survey eksperiment basert på vignetteметоден, siden deltakerne ble presentert to scenarioer hvor alt var likt bortsett fra én detalj. Ved bruk av vignette blir det presentert korte historier om hypotetiske karakterer som deltakerne videre skal respondere på (Finch, 1987), og det er med på å gjøre de presenterte scenarioene mer realistiske (Atzmüller & Steiner, 2010). Eksperimentet gikk ut på å teste om respondentene ble mer moralsk forarget når en algoritme drev diskriminering i rekrutteringsbeslutningene, sammenlignet med når et menneske gjorde det samme. Siden bare detaljen om hvem som utførte diskrimineringen er forskjellig, vil en forskjell i gjennomsnitt i grad av moralsk forargelse skyldes at det reageres annerledes basert på denne manipulerede variabelen.

3.1.3 Primær og sekundærdata

Vi har i denne oppgaven tatt i bruk både primær- og sekundærdata. Primærdata blir brukt for å få en mer detaljert forståelse av temaet, og er data som blir samlet inn til et spesifikt formål eller for å besvare undersøkelsesspørsmål (Gripsrud et al., 2016, s. 69). Våre primærdata består utelukkende av svarene vi hentet inn fra spørreundersøkelsen. Dette ble samlet inn fordi vi ønsket svar fra én spesifikk gruppe om et bestemt fenomen. Vi har også benyttet oss av sekundærdata som er data samlet inn av andre, og med andre formål (Gripsrud et al., 2016, s. 68). Sekundærdataene i denne oppgaven er samlet inn fra tidligere forskning på områder som er relevante for vårt forskningsspørsmål.

3.2 Rekruttering og utvalg

I studien vår var det som nevnt ønskelig å kartlegge tech-ansattes reaksjoner på diskriminering. Siden vi begrenset oss til tech-ansatte grunnet kunnskapen deres om teknologi, betegnes dette som et vurderingsutvalg. Denne type utvalg består av personer som innehar bestemte egenskaper man ønsker skal være representert i utvalget (Gripsrud et al., 2016, s. 176). Som studenter er dette et krevende utvalg å få tak i, og av naturlige årsaker valgte vi å begrense oss til den norske tech-bransjen.

Ved hjelp av et allerede etablert nettverk og en gjennomgang av bedrifter som kunne ha en tech-avdeling, ble det sendt ut både meldinger på LinkedIn og e-mailer hvor vi delte undersøkelsen.¹ I invitasjonen til å delta i studien skrev vi kort at det handlet om kjønnsdiskriminering i teknologibedrifter, og at vi håpte mottakeren av invitasjonen kunne dele den med sine kollegaer. Sektorene vi baserte oss på var hovedsakelig martech, fintech og IT-konsulentselskaper, med god variasjon i størrelsen på virksomheten og antall ansatte. Målet var å nå ut til de som anså seg selv som tech-medarbeidere, og få disse til å svare på spørsmålene våre. Vi unngikk derfor å dele undersøkelsen med tilfeldige personer eller på sosiale medier, for å sikre at svarene var representative for studien og utvalget vi ønsket å fokusere på. Derfor fikk deltakerne dette presentert i starten av spørreskjemaet:

“Vi ønsker kun at mennesker som jobber med teknologi skal delta i denne studien. Vi definerer dette bredt, som i at man har som del av sine arbeidsoppgaver å utvikle, teste, evaluere, kjøpe eller selge teknologi. Vi håper hver deltaker selv vil bruke eget skjønn for å bedømme hvorvidt man kan sies å arbeide med teknologi.”

¹ Siden denne studien er en del av en større internasjonal studie og fordi rekrutteringskravet både var større og mer spesifikt enn det som er vanlig på en bacheloroppgave, er dette datasettet samlet inn i samarbeid mellom to grupper - dette i samråd med veileder og er forhåndsgodkjent.

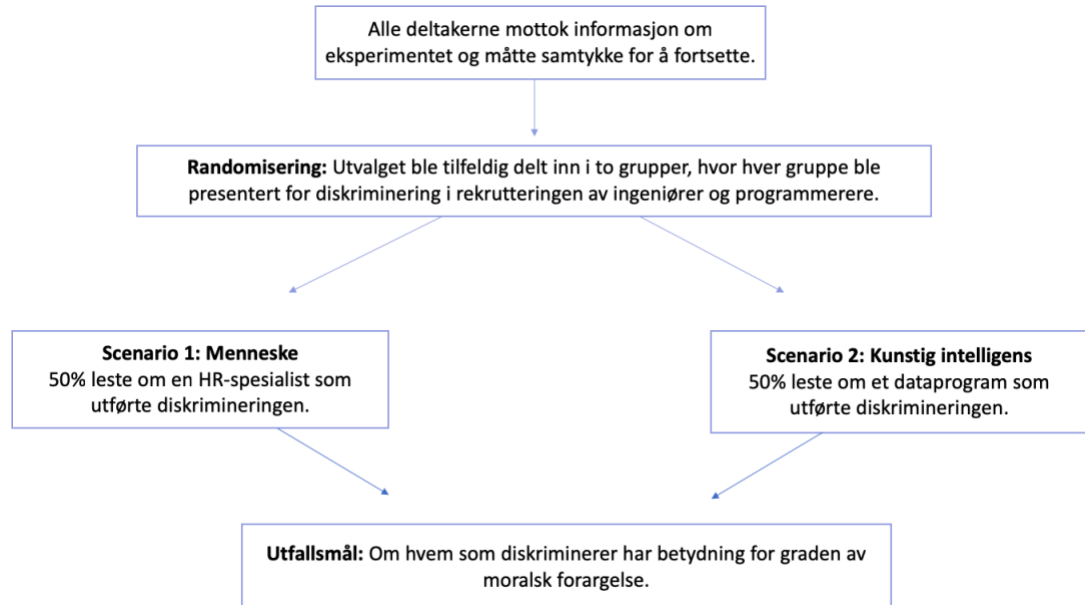
3.3 Eksperimentdesign

Før deltakerne begynte på selve undersøkelsen ble de først møtt av en side med all relevant informasjon om studien. De ble deretter bedt om å bekrefte at de jobbet med teknologi. Dersom de bekreftet dette ble de sendt videre til undersøkelsen, som bestod av 16 spørsmål hvor den første delen var randomisert. Det vil si at helt tilfeldig fikk cirka halvparten av respondentene en annen undersøkelse enn den andre halvparten. Felles for begge var at casen de fikk presentert handlet om en rekrutteringsprosess hvor det forekom diskriminering av kvinnelige søkere, og den som rekrutterte hadde full råderett over alle ansettelser. I den ene delen var rekrutteringen utført av en HR-spesialist og i den andre ble den utført av et dataprogram, altså en algoritme basert på kunstig intelligens. Der noen leste at HR-spesialisten herr Akselsen la på en negativ vektning på alle CV-er som hadde ordet «kvinne» i seg, leste andre at dataprogrammet COMPNET gjorde det samme.

Spørsmålene som ble stilt og utfallsvariablene var dermed identiske, sett bort i fra tvisten om *hvem* som utførte handlingene. Deltakerne visste ingenting om at den undersøkelsen de havnet i var annerledes enn den en annen deltaker kunne få. Undersøkelsen hadde som mål å få frem hvordan deltakeren reagerte på situasjonen de fikk presentert, og hvor galt og umoralsk det var av den som rekrutterte. Deltakerne ble blant annet bedt om å rangere i hvilken grad de var enig i at det var galt på en skala fra 1 til 7, der 1 = svært uenig og 7 = svært enig. Lignende skalaer ble brukt da de ble spurt om å rangere grad av bekymring og hvor foruroliget de ble dersom den samme diskrimineringen hadde skjedd i deres egne selskap, om den/det som utførte handlingen burde bli erstattet/avskjediget og hvorvidt selskapet burde gjøre en innsats for å hindre diskriminering og ansette flere kvinner.

For å teste om respondentene forstod hva de ble spurt om og hvem som fattet ansettelsesbeslutningene i casen de leste, ble det i slutten av denne delen foretatt en oppmerksomhetstest. Denne testen bestod av et kontrollspørsmål om den manipulererte variabelen, og spurte om de hadde lest om et menneske eller en algoritme. Det er sentralt at de forstod dette for at vi skal kunne hente ut korrekte analyser. Deretter ble

det spurt generelle spørsmål om deltakeren, blant annet alder, kjønn og hvor mye kunnskap de har om kunstig intelligens og algoritmer sammenliknet med gjennomsnittspersonen i Norge. Figur 1 viser en illustrasjon av hele eksperimentdesignet.



Figur 1: Eksperimentdesign

3.4 Utfallsmål

For å besvare forskningsspørsmålet ble det utarbeidet to avhengige variabler. Disse var herr Akselsen og COMPNET. Vi ønsket ved hjelp av spørsmålene i vignetten å se om disse variablene målte gruppeforskjeller i moralsk forargelse. Siden det kan være vanskelig for folk å koble moralsk forargelse direkte til en hypotetisk karakter, ønsket vi å teste reaksjonen deres på selve handlingen med noen uavhengige variabler. Disse uavhengige variablene resulterte i følgende ledd:

- Jeg ville blitt moralsk forarget av de diskriminerende handlingene til COMPNET/Akselsen.
- De diskriminerende handlingene til COMPNET/Akselsen var urettferdige.
- De diskriminerende handlingene til COMPNET/Akselsen var umoralske.
- COMPNET/Akselsen sin diskriminering av kvinnelige søkere var galt.

Vi opprettet deretter to sammensatte variabler, en for herr Akselsen og en for COMPNET, som tok gjennomsnittet av disse fire leddene for å kunne måle den moralske forargelsen. Det ble i tillegg opprettet sammensatte variabler for å måle bekymring og forurolighet over diskrimineringen gjort av COMPNET/Akselsen. Den siste variabelen vi målte var hvordan deltakerne mente selskapet burde reagere på feilen som ble begått i begge tilfellene.

3.5 Validitet og reliabilitet

Begrepene validitet og reliabilitet er sentrale når man skal være kritisk til egen forskning. Når en ser på hvor godt man har målt det som var hensikten å måle, snakker man om validitet (Gripsrud et al., 2016, s. 61). Primærdataene ble innsamlet ved hjelp av et randomisert, kontrollert eksperiment av tech-ansatte. For å unngå at resultatene ble påvirket av personer som ikke hadde tilstrekkelig kunnskap om teknologi delte vi som nevnt ikke spørreskjemaet åpent i sosiale medier. Vi har utelukkende delt undersøkelsen med bedrifter og personer vi anså hadde en tilknytning til tech slik at svarene ble mest mulig representative for bransjen. Dette har styrket validiteten til resultatene. Det var likevel opp til respondenten selv å avgjøre hvor mye de kunne om teknologi. For å sikre god nok kvalitet på sekundærdataene fokuserte vi på å bare hente data fra anerkjente tidsskrifter så godt det lot seg gjøre.

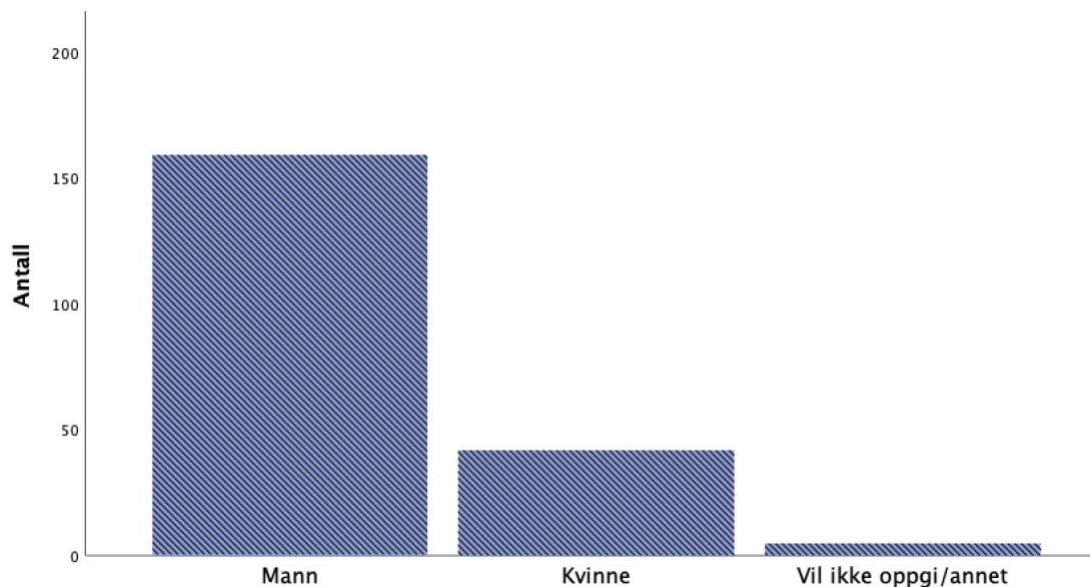
Reliabilitet handler om i hvilken grad en kan stole på at resultatene man har fått fra undersøkelsen er pålitelige. Altså at resultatet blir det samme om eksperimentet blir gjentatt enten ved den samme eller andre metoder (Gripsrud et al., 2016, s. 61). I sekundærdataene fant vi forskning der det var målt forskjell i moralsk forargelse ved diskriminering gjort av kunstig intelligens (Bigman et al., 2020). Dette ble videreført hvor den eneste forskjellen var at utvalget bestod av tech-ansatte i stedet for et tilfeldig utvalg mennesker.

4.0 Resultat

Det var 298 respondenter som startet på undersøkelsen, av disse var det 206 respondenter som fullførte hele. Videre var det 51 respondenter som ikke bestod oppmerksomhetstesten “*Hvem fattet ansettelsesbeslutningen i historien du leste?*”. Disse ble rensset fra datasettet før vi gjennomførte analysen, men er likevel inkludert i de deskriptive dataene da de er med på å gi et bilde av hvem som deltok i undersøkelsen.

4.1 Deskriptive data

Vi begynner med å presentere generell informasjon om de 206 respondentene som fullførte. 77,2% av respondentene er menn, 20,4% er kvinner og 2,4% ville ikke oppgi kjønn. Alderen på respondentene strekker seg fra 21 til 63 år, med en gjennomsnittsalder på 35,19. Spredningen på alder er stor med et standardavvik på 9,58.

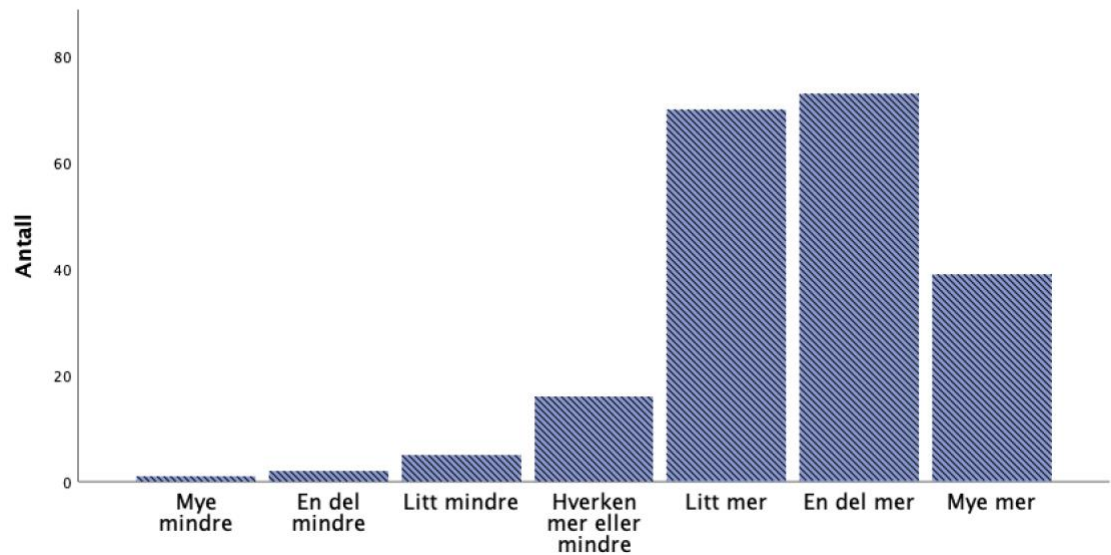


Figur 2: Kjønn

4.2 Kunnskap om temaet

Det var også interessant å se på hva respondentene mente om kunnskapen de har om KI, og om den skilte seg fra kunnskapen til den gjennomsnittlige nordmann.

Deltakerne kunne rangere oppfatningen av sin egen kunnskap på en skala fra 1-7, hvor 1 = mye mindre og 7 = mye mer enn gjennomsnittlig. Resultatet viser at en substansiell andel vet mer enn gjennomsnittspersonen, bare 3,9% av respondentene mener de vet mindre.



Figur 3: Kunnskap

4.3 Sammenheng mellom variablene

I tillegg til å bli spurt om hvor moralsk forarget de ble av diskrimineringen, ble deltakerne også stilt spørsmål om andre faktorer ved casen de fikk presentert. De ble spurt om hvor bekymret de ble av slik diskriminering ($r = 0,61$, $p < 0,001$), i hvilken grad de mente at mennesket eller algoritmen skulle bli sparket/skrotet, i hvilken grad de mente at selskapet skulle komme med en offentlig unnskyldning, gjøre en intern evaluering og gjøre en innsats for å ansette flere kvinner (Cronbach's $\alpha = 0,69$). Vi fant ingen signifikant forskjell mellom faktorene bak disse variablene ($ps > 0,11$) og vil ikke diskutere dette videre.

4.4 Test av mellomgruppeskjeller

Vi analyserte enkelte variabler ved å bruke frekvensanalyser, og sammenhenger mellom variablene ble analysert i form av t-tester. Logikken i en t-test for én stikkprøve fra en populasjon er å teste i hvor stor grad stikkprøvens gjennomsnittsverdi avviker fra den hypotetiske verdien relativt til variabiliteten eller usikkerheten i estimatet (Gripsrud et al., 2016, s. 251).

4.4.1 T-test

Ved å gjennomføre en uavhengig t-test så avdekket vi, som vi forutså, at deltakerne var mindre moralsk forarget når diskrimineringen ble gjort av en algoritme ($M = 6,16$, $SD = 1,55$) enn når diskrimineringen ble utført av et menneske ($M = 6,60$, $SD = 0,98$), $t(153) = 2,15$, $p = 0,33$, Cohen's $d = 0,34$). Disse resultatene gir støtte til vår hypotese om at tech-ansatte blir mindre moralsk forarget når diskriminering blir gjort av en algoritme i motsetning til et menneske.

4.5 Respondentenes kommentarer

Respondentene fikk mulighet til å utdype med egne ord hvor enig/uenig de var i at dataprogrammet/HR-spesialisten burde bli forkastet/avskjediget. I tekstsvarene til de som fikk scenarioet om algoritmen går det igjen at de mener verktøyet kan utbedres, og flere kommer med konkrete forslag på hvordan dette kan gjøres. Svarene fra de som fikk presentert et menneske som diskriminerer bærer preg av at det er flere som mener han burde blitt fratatt ansvaret for ansettelser, få andre oppgaver og at han faktisk bryter loven og burde bli sparket. Tabellene under viser et lite utvalg av disse kommentarene.

COMPNET (algoritme)	
Respondent	Kommentar
Ukjent (56)	«Moderat enig fordi, vektleggingen på ordet «kvinne» kan forkastes eller programmet kan oppdateres sånn at det ikke ser negativt på ordet i en CV. Burde være en enkelt og fikse.»
Kvinne (42)	«Dersom compnet fungerer med RIKTIGE variabler, kan disse justeres slik at kjønn ikke vektet i prosessen.»
Mann (37)	«Om løsningen generelt sett er god ville jeg først forsøkt å anbefale å endre algoritmen og/eller få trent om løsningen på nytt.»
Mann (36)	«Kan være avhengig om det er inkompetanse som ligger til grunn (de burde tydeligvis testet det for så åpenbare biaser vi vet om) eller om det var en bias de aksepterte. Ingen av delene er formildende, men det ene er uansvarlighet og inkompetanse og det andre er uforenlige verdier.»
Mann (38)	«Programvare kan alltid fikses.»

Tabell 1: COMPNET

Herr Akselsen (HR-spesialist)	
Respondent	Kommentar
Mann (31)	«Han burde miste full råderett over ansettelser.»
Ukjent (51)	«Han følger ikke loven og har dermed ikke noe i en slik stilling å gjøre.»
Kvinne (31)	«Han bør få en mulighet til å endre retning, men er det mulig å endre slike holdninger?»
Mann (28)	«Akselsen bør irrettesettes og om han ikke viser forståelse for hvorfor handlingene han har gjort er gale bør han avskjediges.»
Kvinne (63)	«Herr akselsen burde blitt forklart hva han har gjort galt og få andre oppgaver. Vi har stillingsvern i Norge, det gjelder også ham.»

Tabell 2: Herr Akselsen

5.0 Diskusjon

I diskusjonsdelen tar vi med oss resultatet som har kommet frem i studien vår, og diskuterer dette med teoretiske og praktiske implikasjoner. De teoretiske implikasjonene baserer seg på om hypotesen kan bekreftes å bli støttet fra teorien som vi har gått gjennom tidligere i oppgaven. De praktiske implikasjonene viser hva resultatene fra vår studie kan bety, og hva som med det blir våre råd til de lederne og organisasjonene som eventuelt vil benytte seg av algoritmer i rekrutteringsprosesser.

5.1 Teoretiske implikasjoner

Resultatene bekreftet vår hypotese om at tech-ansatte, som vanlige folk, blir mindre moralsk forarget av diskriminering gjort av en algoritme enn når det er gjort av et menneske. Moralsk forargelse blir brukt som en reaksjon på diskriminering (Spring et al., 2018) og graden av hvor forarget man blir avhenger av måten diskrimineringen skjer på (Goodwin et al., 2014). Det er viktig å presisere at selv om deltakerne blir mindre forarget av algoritmen så viser undersøkelsen at graden av forargelse fortsatt er høy. Dette tyder på at diskriminering aldri er akseptabelt, og det er noe som trolig uansett vil provosere i en eller annen grad uavhengig av hvem som gjør det. Grunnen til at de blir noe mer forarget over diskrimineringen gjort av den HR-ansatte kan henge sammen med at herr Akselsen er mer anlagt for å være forutinntatt og motivert av fordommer enn det en algoritme er (Bigman & Tamir, 2016). Dette støtter vårt forslag om at forargelse ved diskriminering gjort av algoritmer vil avhenge av de psykologiske årsakene. I tekstsvarene kommer det frem at tech-ansatte tror grunnen til diskrimineringen kan være at han mener kvinnene er mindre kvalifiserte eller at han ønsker å ansette de kandidatene som ligner seg selv. Dette stemmer overens med det forskning om likhetseffekten sier om at en rekrutterer har lettere for å velge de med lik bakgrunn og personlighetstrekk som seg selv (Fatemi, 2019).

Det som imidlertid skiller de tech-ansatte fra andre er at de vet mer om teknologien bak det kunstig intelligente dataprogrammet. Av respondentene var det bare 3,9% som mente de kunne mindre enn en gjennomsnittlig nordmann om kunstig intelligens. Mange av de som jobber innen teknologi vet dermed at algoritmen er bygget opp av

innhenting av store mengder data, at den finner mønster og baserer beslutningene sine på dette. Dette er noe som videre kan resultere i at algoritmen foretrekker menn over kvinner på grunn av at det for eksempel allerede er flere vellykkede menn i selskapet (J.D, 2019). Det kan da tenkes at studiens deltakere vil bli mindre forarget siden de tilsynelatende vet at diskrimineringen kan skyldes måten den er programmert, og at det ikke nødvendigvis foreligger noen faktiske personlige motiv bak. På den andre siden kan dataene algoritmen er basert på være påvirket av at det faktisk foreligger en diskriminerende kultur i selskapet. I tekstsvarene fremkommer det at denne ukulturen kan være grunnen til at feilen med algoritmen i det hele tatt eksisterer, og den høye graden av forargelse er mer rettet mot at selskapet har akseptert denne type holdninger.

5.2 Praktiske implikasjoner

Ved innføring av kunstig intelligente programmer i rekruttering, er det flere faktorer som vil være hensiktsmessige for ledelsen å ta hensyn til. En av dem er at KI i rekruttering kan redusere diskriminering av kvinner, men at det har noen forutsetninger. En algoritme vil slik teknologien er i dag, ikke kunne vurdere noe annet enn dokumentert kunnskap, finne mønstre og forhåpentligvis fjerne bias i vurderingen av kandidater (Clark, 2017). For å best utnytte dette må teknologien derfor være programmert på en forsvarlig måte og jevnlig oppfølges. Da kan man sikre seg at maskinlæringen ikke tar over og uten oppsyn lærer seg å utføre diskriminerende handlinger, men dette kan være lettere sagt enn gjort. Et menneske vil dog i tillegg til å se på disse objektive kriteriene, også kunne vurdere basert på moral, sunn fornuft og andre ting som ikke lar seg dokumentere rent objektivt.

Dersom et verktøy skal erstatte flere av oppgavene til en HR-ansatt, er det viktig å belyse hvordan det heller kan brukes som et supplement og hjelpemiddel fremfor å ta over og få full råderett over rekrutteringsprosessene. Slik kan den ansatte fokusere fullt på de tingene som krever menneskelig intelligens og kritisk tenkning, og det reduserer omfanget av etiske problemstillinger knyttet til når KI får fritt spillerom. Det vil derfor være hensiktsmessig å bruke en kombinasjon av menneskelig og

kunstig intelligens i rekrutteringen, for å maksimere fordelene av begge metodene (Kuncel et al., 2014). Ved en slik kombinasjon kan menneskene i organisasjonen fokusere på oppgaver som krever menneskelig dømmekraft og vurderingsevne, og la teknologien ta seg av de administrative oppgavene (Kolbjørnsrud et al., 2016).

En annen praktisk implikasjon er at tech-ansatte lettere ser løsninger for utfordringer knyttet til diskriminering når det er en algoritme som står bak, enn når det er et menneske. Dette underbygges av tekstsvarene i undersøkelsen. Her er det flere som foreslår konkret hvilke endringer som skal til, noe som også blir gjenspeilet i hvor mange som har svart at de vet mer om KI enn en gjennomsnittlig nordmann. Det som har vist seg å gå igjen i svarene er at dataprogrammet ofte bare kan trenge en enkel justering for å ikke begå samme feil igjen. For herr Akselsen trengs det derimot mer omfattende ressurser for å endre hans holdninger til kvinner, om det i det hele tatt er mulig. Terskelen er også lavere for å sparke herr Akselsen eller gi han andre oppgaver, enn å skrote dataprogrammet. Det kan derfor tenkes at hvis man har kjennskap til hva som er galt med algoritmen og hva som skal til for å rette opp i den, vil man gjerne være mer forståelsesfull for eventuelle feil den gjør.

I tillegg blir det viktig å ta i betraktning hva en rettferdig rekrutteringsprosess har å si for selskapets omdømme. De som fremstår som attraktive arbeidsgivere får flere søkere de kan velge mellom. Det positive omdømmet gjenspeiles i selskapets faktiske HR-praksiser hvor det aktivt fremkommer at ansatte er en viktig ressurs (Kuvaas & Dysvik, 2017, s. 92). Kandidater som blir ansatt på slutten av en urettferdig rekrutteringsprosess utvikler også ofte lavere tilknytning til organisasjonen, og vil ha større sannsynlighet for å slutte tidligere (Hausknecht et al., 2004). Om det er kjent at en bedrift benytter seg av rekrutteringsverktøy som ikke bedømmer kandidater på likt grunnlag, kan dette hindre dyktige potensielle kandidater fra å søke. Organisasjonen vil da gå glipp av det som opprinnelig kunne vært et viktig konkurransefortrinn.

5.3 Begrensninger og anbefalinger

Til tross for at studien vår har lagt frem noen interessante funn, har den også noen begrensninger. Det hadde vært ønskelig at respondentenes kjønn var mer likt fordelt, da det kan tenkes at kvinner i større grad kan kjenne seg igjen i den fiktive historien de fikk presentert. Dersom andelen kvinnelige respondenter var høyere kunne dette påvirket resultatene, og grad av moralsk forargelse ville muligens vært annerledes og gjerne enda sterkere. Det er likevel hverken overraskende eller til å unngå at en stor andel av utvalget vårt er menn. Som nevnt innledningsvis, og som resultatene viser ved at bare 20,4% av respondentene våre er kvinner, bærer tech-bransjen preg av et mannsdominert arbeidsmiljø. En anbefaling til videre forskning er å gå mer målrettet frem for å hente inn flere kvinner fra denne bransjen slik at andelen mellom kjønnene blir så lik som mulig og flere synsvinkler kan bli belyst.

En annen begrensning ligger i at vi fokuserte på norske tech-ansatte. Dette gjør at våre svar ikke nødvendigvis lar seg replikere om vi overfører dem til bedrifter i andre land. Norge er kjent for å være langt fremme på likestilling og kjønnsbalanse, og kan dermed være annerledes fra andre kulturer på mange måter. Det er derfor grunn til å tro at en lik undersøkelse i et annet land som ikke har kommet like langt vil gi ulikt resultat. Fremtidig forskning burde derfor undersøke det samme, men i andre land og kulturer. Dette kan bekrefte våre funn og/eller avdekke andre nyanser av temaet enn det vi har kommet frem til.

Det finnes også begrensninger knyttet til selve undersøkelsen. Siden oppmerksomhetstesten bare var et spørsmål med to svaralternativer finnes muligheten for at noen av deltakerne svarte riktig uten å egentlig ha satt seg godt nok inn i undersøkelsen. Dette kan ha skjedd ved ren tilfeldighet og at de dermed bestod testen, noe som da vil svekke validiteten til studien vår. Det var også en hel del som hoppet ut av undersøkelsen når de måtte ønske. Grunnet dette er det noen svar med forskjellig antall respondenter.

Det er verdt å nevne at vi på ingen måte kan garantere at alle respondentene våre faktisk jobber innen tech. Siden undersøkelsen baserer seg på en online survey er det opp til hver enkelt deltaker å vurdere i starten om de anser seg selv som tech-ansatte eller ikke. Som nevnt tidligere begrenset vi denne muligheten så godt det lot seg gjøre. Utover dette stoler vi blindt på respondenten. Til videre forskning anbefales det derfor å utføre en lignende studie som et felteksperiment, og gå ut til utvalgte tech-bedrifter. På denne måten er det enklere å kontrollere flere av de begrensingene vi har nevnt. Da sikrer man at de som deltar i studien faktisk kjenner til teknologi, at det er nok kvinnelige deltakere og at hele eksperimentet blir fullført av hver deltaker. Et felteksperiment vil i tillegg teste deltakerne i et mer naturlig miljø enn det som er tilfelle med et spørreskjema (Gripsrud et al., 2016, s. 58). Det kan tenkes at mange vil reagere på en annen måte når de befinner seg i en mer realistisk situasjon enn når de besvarer et anonymt spørreskjema på nett. Om ledere og kollegaer er vitne til reaksjonen man uttrykker, er det rimelig å anta at en kanskje vil ha vanskeligere for å eksempelvis stå opp mot diskrimineringen dersom selskapet man jobber i er svært mannsdominert. Dette er med på å svekke den ytre validiteten til vår forskning.

Til slutt vil vi nevne at undersøkelsen vår ikke belyser diskriminering som skjer under selve intervjuprosessen, kun den tidlige utvelgelsen basert på CV og søknad. Av den grunn vil vi anbefale til videre forskning å undersøke bredden i vårt forskningsspørsmål. Dette kan gjøres ved se på diskriminering som kan forekomme også senere i rekrutteringsprosessen, for eksempel reaksjoner på diskriminering der man har latt KI analysere et videointervju. Problemstillinger knyttet til at KI kan samle inn data fra personer som ikke har søkt en spesifikk stilling er også interessante.

6.0 Konklusjon

Bruken av KI i rekruttering har ført med seg flere suksesshistorier, men også noen alvorlige feil. Vi ønsket å belyse hvordan disse feilene ble reagert på av et utvalg vi antok hadde mer kunnskap om teknologien enn andre. Våre resultater viser at tech-ansatte reagerer likt som de uten tech-bakgrunn, men at de på grunn av sin kunnskap om temaet skiller seg noe ut. Ut fra hypotesen fikk vi bekreftet at de blir mindre moralsk forarget når en algoritme diskriminerer enn når et menneske gjør det samme, og at de videre ser løsninger for hvordan problemet kan håndteres. De reagerer også med en høy moralsk forargelse når et menneske diskriminerer, men ser ikke løsninger på samme måte som de gjør med algoritmen.

Vi kan med dette konkludere at kunstig intelligens med fordel kan benyttes for å redusere diskriminering, men med måte. Ethiske problemstillinger er høyaktuelle og må tas i betraktning både når teknologien programmeres og når den er satt i arbeid. På denne måten kan kunstig intelligens kombineres med bruk av menneskelige ressurser, og man får det beste av to verdener.

7.0 Referanser

- Agarwal, D., Bersin, J., Lahiri, G., Schwartz, J., & Volini, E. (2018). *The rise of the social enterprise* (s. 93). Deloitte.
https://www2.deloitte.com/content/dam/insights/us/articles/HCTrends2018/2018-HCTrends_Rise-of-the-social-enterprise.pdf
- Agarwal, P. (2020, mars 4). *Gender Bias In STEM: Women In Tech Still Facing Discrimination*. Forbes.
<https://www.forbes.com/sites/pragyaagarwaleurope/2020/03/04/gender-bias-in-stem-women-in-tech-report-facing-discrimination/>
- Angwin, J., Larson, J., Surya, M., & Lauren, K. (2016). *Machine Bias*. ProPublica.
<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Aquilina, Y., & Saliba, M. A. (2019). An automated supermarket checkout system utilizing a SCARA robot: Preliminary prototype development. *Procedia Manufacturing*, 38, 1558–1565. <https://doi.org/10.1016/j.promfg.2020.01.130>
- Atzmüller, C., & Steiner, P. (2010). Experimental Vignette Studies in Survey Research. *Methodology: European Journal of Research Methods for The Behavioral and Social Sciences*, 6, 128–138. <https://doi.org/10.1027/1614-2241/a000014>
- Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The Moral Machine experiment. *Nature*, 563(7729), 59–64.
<https://doi.org/10.1038/s41586-018-0637-6>
- Bartlett, C. A., & Ghoshal, S. (2002). Building competitive advantage through people. *MIT Sloan Management Review; Cambridge*, 43(2), 34–41.
- Batson, C. D., Kennedy, C. L., Nord, L. A., Stocks, E. L., Fleming, D. A., Marzette, C. M., Lishner, D. A., Hayes, R. E., Kolchinsky, L. M., & Zenger, T. (2007). Anger at unfairness: Is it moral outrage? *European Journal of Social Psychology*, 37(6), 1272–1285.
<https://doi.org/10.1002/ejsp.434>
- Bersin, J. (2018, november 2). AI Comes To Recruiting: Will Interviews Go The Way Of The Dinosaur? *JOSH BERSIN*. <https://joshbersin.com/2018/11/ai-comes-to-recruiting-will-interviews-go-the-way-of-the-dinosaur/>
- Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition*, 181, 21–34. <https://doi.org/10.1016/j.cognition.2018.08.003>
- Bigman, Y. E., & Tamir, M. (2016). The road to heaven is paved with effort: Perceived effort amplifies moral judgment. *Journal of Experimental Psychology: General*, 145(12), 1654–1669. <https://doi.org/10.1037/xge0000230>
- Bigman, Y. E., Waytz, A., Alterovitz, R., & Gray, K. (2019). Holding Robots Responsible: The Elements of Machine Morality. *Trends in Cognitive Sciences*, 23(5), 365–368.
<https://doi.org/10.1016/j.tics.2019.02.008>
- Bigman, Y. E., Wilson, D., Arnestad, M. N., Waytz, A., & Gray, K. (2020). People are less

Morally Outraged with Algorithm Discrimination. *Under utarbeidelse*.

Black, J. S., & van Esch, P. (2020). AI-enabled recruiting: What is it and how should a manager use it? *Business Horizons*, 63(2), 215–226.
<https://doi.org/10.1016/j.bushor.2019.12.001>

Boudreau, J. W., & Ramstad, P. M. (2005). Where's Your Pivotal Talent? *Harvard Business Review*, 83(4), 23–24.

Brewer, M. B., & Kramer, R. M. (1985). The Psychology of Intergroup Attitudes and Behavior. *Annual Review of Psychology*, 36(1), 219.
<https://doi.org/10.1146/annurev.ps.36.020185.001251>

Campbell, C., Sands, S., Ferraro, C., Tsao, H.-Y. (Jody), & Mavrommatis, A. (2020). From data to action: How marketers can leverage AI. *Business Horizons*, 63(2), 227–243.
<https://doi.org/10.1016/j.bushor.2019.12.002>

Cárdenas-Barrón, L. E., Treviño-Garza, G., & Wee, H. M. (2012). A simple and better algorithm to solve the vendor managed inventory control system of multi-product multi-constraint economic order quantity model. *Expert Systems with Applications*, 39(3), 3888–3895. <https://doi.org/10.1016/j.eswa.2011.09.057>

Cheung, C. F., Lee, W. B., Wang, W. M., Chu, K. F., & To, S. (2003). A multi-perspective knowledge-based system for customer service management. *Expert Systems with Applications*, 24(4), 457–470. [https://doi.org/10.1016/S0957-4174\(02\)00193-8](https://doi.org/10.1016/S0957-4174(02)00193-8)

Clark, J. (2017, mars 21). *How cognitive solutions are transforming HR: InterConnect 2017*. Business Operations. <https://www.ibm.com/blogs/internet-of-things/hr-cognitive-solutions/>

Conitzer, V., Sinnott-Armstrong, W., Borg, J. S., Deng, Y., & Kramer, M. (2017). Moral Decision Making Frameworks for Artificial Intelligence. *ISAIM*.

Dastin, J. (2018, oktober 10). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>

Dustin. (2018, februar 13). *AI-eksperten: «Etikk er like viktig som teknologi»*. <https://www.dustin.no/tjenester/kunnskapsbanken/archive/ai-eksperten-etikk-er-like-viktig-som-teknologi/>

Eriksen, S. I. (1996). Rekruttering og lederutvelgelse. I *Norbok*. TANO.
https://urn.nb.no/URN:NBN:no-nb_digibok_2008071004066

Erlandsson, A. (2019). Do Men Favor Men in Recruitment? A Field Experiment in the Swedish Labor Market: *Work and Occupations*. <https://doi.org/10.1177/0730888419849467>

Faliagka, E., Ramantas, K., Tsakalidis, A., & Tzimas, G. (2012). Application of learning algorithms to online recruitment systems.pdf. *ICIW 2012 : The Seventh International Conference on Internet and Web Applications and Services Application, c*, 215–220.

Fangen, K., & Sellerberg, A.-M. (2011). *Mange ulike metoder* (1.). Gyldendal.

- Fatemi, F. (2019, oktober 21). *How AI Is Uprooting Recruiting*. <https://www.forbes.com/sites/falonfatemi/2019/10/31/how-ai-is-uprooting-recruiting/#7bf593ae46ce>
- Finch, J. (1987). THE VIGNETTE TECHNIQUE IN SURVEY RESEARCH. *Sociology*, 21(1), 105–114. JSTOR.
- Fiske, A. P., & Tetlock, P. E. (1997). Taboo Trade-offs: Reactions to Transactions That Transgress the Spheres of Justice. *Political Psychology*, 18(2), 255–297. <https://doi.org/10.1111/0162-895X.00058>
- Gogoll, J., & Uhl, M. (2018). Rage against the machine: Automation in the moral domain. *Journal of Behavioral and Experimental Economics*, 74, 97–103. <https://doi.org/10.1016/j.socec.2018.04.003>
- Goldstein, H. W., Pulakos, E. D., Semedo, C., & Passmore, J. (2017). *The Wiley Blackwell Handbook of the Psychology of Recruitment, Selection and Employee Retention*. John Wiley & Sons, Incorporated. <http://ebookcentral.proquest.com/lib/bilibrary/detail.action?docID=4860506>
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106(1), 148–168. <https://doi.org/10.1037/a0034726>
- Gray, H., Gray, K., & Wegner, D. (2007). Dimensions of Mind Perception. *Science (New York, N.Y.)*, 315, 619. <https://doi.org/10.1126/science.1134475>
- Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125(1), 125–130. <https://doi.org/10.1016/j.cognition.2012.06.007>
- Gripsrud, G., Olsson, U. H., & Silkoset, R. (2016). *Metode og dataanalyse* (3.). Cappelen Damm Akademisk.
- Guinan, P. J., Parise, S., & Rollag, K. (2014). Jumpstarting the use of social technologies in your organization. *Business Horizons*, 57(3), 337–347. <https://doi.org/10.1016/j.bushor.2013.12.005>
- Gulbrandsen, C. (2017, juli 12). Rekrutterer ved hjelp av robot – ansetter flere kvinner | DN. *Dagens Næringsliv*. <https://www.dn.no/arbeidsliv/robotteknologi/utdannelse/evry-evry/rekrutterer-ved-hjelp-av-robot-ansetter-flere-kvinner/2-1-227334>
- Gulliford, F., & Parker Dixon, A. (2019). AI: The HR revolution. *Strategic HR Review*, 18(2), 52–55. <https://doi.org/10.1108/SHR-12-2018-0104>
- Gummerum, M., Van Dillen, L. F., Van Dijk, E., & López-Pérez, B. (2016). Costly third-party interventions: The role of incidental anger and attention focus in punishment of the perpetrator and compensation of the victim. *Journal of Experimental Social Psychology*, 65, 94–104. <https://doi.org/10.1016/j.jesp.2016.04.004>
- Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant Reactions to Selection

- Procedures: An Updated Model and Meta-Analysis. *Personnel Psychology*, 57(3), 639–683.
<https://doi.org/10.1111/j.1744-6570.2004.00003.x>
- Heilman, M. E., & Caleo, S. (2018, januar 28). *Gender Discrimination in the Workplace*. The Oxford Handbook of Workplace Discrimination.
<https://doi.org/10.1093/oxfordhb/9780199363643.013.7>
- Hicks, M. (2018, oktober 12). Why tech's gender problem is nothing new. *The Guardian*.
<https://www.theguardian.com/technology/2018/oct/11/tech-gender-problem-amazon-facebook-bias-women>
- Hovde, K.-O., & Grønmo, S. (2020). Algoritme. I *Store norske leksikon*.
<http://snl.no/algoritme>
- Ideal. (2020, mai 21). AI For Recruiting: A Definitive Guide For HR Professionals. *Ideal*.
<https://ideal.com/ai-recruiting/>
- Ikdahl, I. (2018). Diskriminering. I *Store norske leksikon*. <http://snl.no/diskriminering>
- Interbrand. (2017). *Best Global Brands 2017 Rankings*. Interbrand.
<https://www.interbrand.com/best-brands/best-global-brands/2017/ranking/>
- Iversen, O. I. (2015). *Rekrutterings- og intervjueteknikk*. Fagbokforlaget.
- J.D, A. S. (2019, desember 12). *AI: Discriminatory Data In, Discrimination Out*. SHRM.
<https://www.shrm.org/resourcesandtools/legal-and-compliance/employment-law/pages/artificial-intelligence-discriminatory-data.aspx>
- Johnston, S. (2018, januar 31). Largest companies 2008 vs. 2018, a lot has changed. *Milford Asset*. <https://milfordasset.com/insights/largest-companies-2008-vs-2018-lot-changed>
- Kantar TNS. (2018). *Andelen kvinner i norsk IT-bransje*. <https://www.ikt-norge.no/wp-content/uploads/2019/10/kvinner-i-itbransjenoda2018-kantartns.pdf>
- Keller, S., & Meaney, M. (2017). I *Leading Organizations: Ten Timeless Truths* (s. 14). Bloomsbury Publishing Plc.
<http://ebookcentral.proquest.com/lib/bilibrary/detail.action?docID=5240791>
- Kok, J. N., Boers, Egbert J.W., Kusters, Walter A., Putten, Peter van der, & Poel, Mannes. (2002). Artificial Intelligence: Definition, Trends, Techniques and Cases. *Knowledge for Sustainable Development: An Insight into the Encyclopedia of Life Support Systems*, 1095–1107.
- Kolbjørnsrud, V., Amico, R., & Thomas, R. J. (2016, november 2). How Artificial Intelligence Will Redefine Management. *Harvard Business Review*.
<https://hbr.org/2016/11/how-artificial-intelligence-will-redefine-management>
- Kommunal- og moderniseringsdepartementet. (2020). *Nasjonal strategi for kunstig intelligens*.
<https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/no/pdfs/ki-strategi.pdf>

- Kuncel, N. R., Ones, D. S., & Klieger, D. M. (2014, mai 1). In Hiring, Algorithms Beat Instinct. *Harvard Business Review*, May 2014. <https://hbr.org/2014/05/in-hiring-algorithms-beat-instinct>
- Kuvaas, B., & Dysvik, A. (2017). *Lønnsomhet gjennom menneskelige ressurser* (3.). Fagbokforlaget.
- Kvalnes, Øyvind. (2015). *Moral Reasoning at Work*. Palgrave Macmillan UK. <https://doi.org/10.1057/9781137532619>
- Laurent, P., Chollet, T., & Herzberg, E. (2017). *Deloitte—Intelligent automation entering the business world* [Economy & Finance]. <https://www.slideshare.net/NickMeijvander/deloitte-intelligent-automation-entering-the-business-world>
- Levine, E. E., & Schweitzer, M. E. (2014). Are liars ethical? On the tension between benevolence and honesty. *Journal of Experimental Social Psychology*, 53, 107–117. <https://doi.org/10.1016/j.jesp.2014.03.005>
- Linás, P. (2019, januar 3). Five key recruitment technology trends to watch in 2019. *Information Age*. <https://www.information-age.com/recruitment-technology-2019-123477849/>
- Lindenmeier, J., Schleer, C., & Pricl, D. (2012). Consumer outrage: Emotional reactions to unethical corporate behavior. I *Journal of Business Research* (Bd. 65, Nummer 9, s. 1364–1373). <https://doi.org/10.1016/j.jbusres.2011.09.022>
- Malle, B. F., Magar, S. T., & Scheutz, M. (2019). AI in the Sky: How People Morally Evaluate Human and Machine Decisions in a Lethal Strike Dilemma. I M. I. Aldinhas Ferreira, J. Silva Sequeira, G. Singh Virk, M. O. Tokhi, & E. E. Kadar (Red.), *Robotics and Well-Being* (s. 111–133). Springer International Publishing. https://doi.org/10.1007/978-3-030-12524-0_11
- Malle, B. F., Scheutz, M., Forlizzi, J., & Voiklis, J. (2016). Which robot am I thinking about? The impact of action and appearance on people's evaluations of a moral robot. *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 125–132. <https://doi.org/10.1109/HRI.2016.7451743>
- Martin, J., Brickman, P., & Murray, A. (1984). Moral outrage and pragmatism: Explanations for collective action. *Journal of Experimental Social Psychology*, 20(5), 484–496. [https://doi.org/10.1016/0022-1031\(84\)90039-8](https://doi.org/10.1016/0022-1031(84)90039-8)
- Mclaren, S. (2018). *How Hilton, Google, and More Have Dramatically Reduced Their Time to Hire*. <https://business.linkedin.com/talent-solutions/blog/recruiting-strategy/2018/how-4-companies-reduced-time-to-hire>
- Miller, D. T., Effron, D. A., & Zak, S. V. (2011). From moral outrage to social protest: The role of psychological standing. *The Psychology of Justice and Legitimacy*, November, 103–124. <https://doi.org/10.4324/9780203837658>
- Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.

- Nath, R., & Sahu, V. (2020). The problem of machine ethics in artificial intelligence. *AI & SOCIETY*, 35(1), 103–111. <https://doi.org/10.1007/s00146-017-0768-6>
- Nelissen, R. M. A., & Zeelenberg, M. (2009). Moral emotions as determinants of third-party punishment: Anger, guilt, and the functions of altruistic sanctions. *Judgment and Decision Making*, 4(7), 543–553.
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
- Plan International Norway. (2019). *The Gender Gap in Technology in Scandinavia*. https://www.telenor.com/wp-content/uploads/2019/09/The-Gender-Gap-in-Technology-in-Scandinavia_Full-report.pdf
- Reeder, G. D., Kumar, S., Hesson-McInnis, M. S., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology*, 83(4), 789–803. <https://doi.org/10.1037/0022-3514.83.4.789>
- Rogstad, J., & Sterri, E. B. (2018). Passe inn og passe til. *Tidsskrift for samfunnsforskning*, 59(01), 41–65. <https://doi.org/10.18261/issn.1504-291X-2018-01-03>
- Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger, but not moral disgust, responds to intentionality. *Emotion (Washington, D.C.)*, 11(2), 233–240. <https://doi.org/10.1037/a0022598>
- SAS. (2020, mai 28). *Machine Learning: What it is and why it matters*. https://www.sas.com/en_us/insights/analytics/machine-learning.html
- Shank, D. B., & DeSanti, A. (2018). Attributions of morality and mind to artificial intelligence after real-world moral violations. *Computers in Human Behavior*, 86, 401–411. <https://doi.org/10.1016/j.chb.2018.05.014>
- Shank, D. B., DeSanti, A., & Maninger, T. (2019). When are artificial intelligence versus human agents faulted for wrongdoing? Moral attributions after individual and joint decisions. *Information, Communication & Society*, 22(5), 648–663. <https://doi.org/10.1080/1369118X.2019.1568515>
- Sharma, A., & Ltd, P. M. M. P. (2018, august 16). How AI reinvented hiring practice at L'Oréal. *People Matters*. https://www.peoplesmatters.in/article/technology/how-the-worlds-largest-cosmetic-company-transformed-its-hiring-practice-with-ai-19006?media_type=article&subcat=techhr-2018&title=how-the-worlds-largest-cosmetic-company-transformed-its-hiring-practice-with-ai&id=19006
- Spring, V. L., Cameron, C. D., & Cikara, M. (2018a). The Upside of Outrage. *Trends in Cognitive Sciences*, 22(12), 1067–1069. <https://doi.org/10.1016/j.tics.2018.09.006>
- Spring, V. L., Cameron, C. D., & Cikara, M. (2018b). The Upside of Outrage. *Trends in Cognitive Sciences*, 22(12), 1067–1069. <https://doi.org/10.1016/j.tics.2018.09.006>
- Stankiewicz, K. (2019). *Twitter complainer says Apple is to blame for credit card issues*.

CNBC. <https://www.cnn.com/2019/11/11/apple-shouldnt-pass-the-blame-on-gender-bias-says-complainant.html>

Sunstein, C. R., Kahneman, D., & Schkade, D. (1998). Assessing punitive damages. *Yale Law Journal*, *107*(50), 2071–2153.

Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review*, *109*(3), 451–471.
<https://doi.org/10.1037//0033-295X.109.3.451>

Torres, M. (2017, juli 26). 2 million job candidates are desperate to work for Google. Why? *Ladders | Business News & Career Advice*. <http://www.theladders.com/career-advice/2-million-job-candidates-google>

Tudow, A., Zanni, T., & George, R. (2019). *The future of HR in the technology sector* (s. 4). KPMG. <https://assets.kpmg/content/dam/kpmg/uk/pdf/2019/07/the-future-of-hr-in-the-technology-sector.pdf>

Tyfting, M. (2019, desember 9). Council Post: How Technology Can Enhance Modern Recruitment. *Forbes*.
<https://www.forbes.com/sites/forbeshumanresourcescouncil/2019/09/12/how-technology-can-enhance-modern-recruitment/>

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A Person-Centered Approach to Moral Judgment: *Perspectives on Psychological Science*.
<https://journals.sagepub.com/doi/10.1177/1745691614556679>

Wakabayashi, D. (2017, august 7). Google Fires Engineer Who Wrote Memo Questioning Women in Tech. *The New York Times*.
<https://www.nytimes.com/2017/08/07/business/google-women-engineer-fired-memo.html>

Weisman, K., Dweck, C. S., & Markman, E. M. (2017). Rethinking people's conceptions of mental life. *Proceedings of the National Academy of Sciences*, *114*(43), 11374–11379.
<https://doi.org/10.1073/pnas.1704347114>

Young, A. D., & Monroe, A. E. (2019). Autonomous morals: Inferences of mind predict acceptance of AI behavior in sacrificial moral dilemmas. *Journal of Experimental Social Psychology*, *85*, 103870. <https://doi.org/10.1016/j.jesp.2019.103870>

Yu, K. Y. T., & Cable, D. M. (2012). Recruitment and Competitive Advantage: A Brand Equity Perspective. I *The Oxford Handbook of Organizational Psychology, Volume 1* (Bd. 1, s. 212). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199928309.013.0007>

Zhao, A. (2018, oktober 12). The Rise of Artificial Intelligence | IMMpress Magazine. *Immpress Magazine*. <http://www.immpressmagazine.com/the-rise-of-artificial-intelligence/>

Aas, K. N., & Ikdahl, I. (2019). Kjønnsdiskriminering. I *Store norske leksikon*.
<http://snl.no/kj%C3%B8nnsdiskriminering>

8.0 Vedlegg

Case 1:

Se for deg det følgende:

Du har jobbet i et høyteknologiselskap de siste fem årene.

For to år siden fikk COMPNET, et dataprogram basert på kunstig intelligens, full råderett over ansettelse av ingeniører og programmerere i ditt selskap.

Det ble nylig oppdaget at COMPNET var forutinntatt mot kvinner når den evaluerte søkeres CV'er. COMPNET la på en negativ vektning av alle CV'er som hadde ordet "kvinne" i seg, som i "kaptein på kvinnelaget i sjakk".

Dette forhindret mange talentfulle og kvalifiserte kvinnelige ingeniører fra å få godt betalte jobber i ditt selskap.

Case 2:

Se for deg det følgende:

Du har jobbet i et høyteknologiselskap de siste fem årene.

For to år siden fikk herr Akselsen, en HR spesialist, full råderett over ansettelse av ingeniører og programmerere i ditt selskap.

Det ble nylig oppdaget at Akselsen var forutinntatt mot kvinner når han evaluerte søkeres CV'er. Akselsen la på en negativ vektning av alle CV'er som hadde ordet "kvinne" i seg, som i "kaptein på kvinnelaget i sjakk".

Dette forhindret mange talentfulle og kvalifiserte kvinnelige ingeniører fra å få godt betalte jobber i ditt selskap.