



UiT Norges arktiske universitet

Det juridiske fakultet

(U)tilsiktet diskriminering ved bruk av kunstig intelligens for utvelgelse av fysiske personer i et kontrollperspektiv

Johanne Grov Eilertsen

Masteroppgave i rettsvitenskap JUR-3902 mai 2022

Innholdsfortegnelse

(U)tilsiktet diskriminering ved bruk av kunstig intelligens for utvelgelse av fysiske personer i et kontrollperspektiv	1
1 Innledning.....	1
1.1 Tema og problemstilling	1
1.2 Aktualitet.....	1
1.3 Rettskilder og metode.....	3
1.3.1 Anvendte rettskilder	3
1.3.2 Metode.....	5
1.4 Avgrensing av avhandlingen	6
1.5 Videre fremstilling	6
2 Profilerings til kontroll som diskrimineringsrettslig problem	7
2.1 Innledning.....	7
2.2 Eksempler på kontroll av personer på bakgrunn av etnisitet	8
2.2.1 Konklusjon	9
3 Nærmere om kunstig intelligens og underkategorier som maskinlæring.....	11
3.1 Innledning.....	11
3.2 Ytterligere begreper innenfor domenet «kunstig intelligens»	11
4 Forutinntatt kunstig intelligens.....	15
4.1 Innledning.....	15
4.1.1 Harms of allocations.....	16
4.1.2 Harms of representation	17
5 Nasjonal lovgivning innenfor diskrimineringsretten.....	19
5.1 Innledning.....	19
5.2 Positive plikter og negative forbud	19
5.3 Grunnlovens diskrimineringsvern	20
5.4 Aktivitetsplikten for offentlige myndigheter.....	22

5.5	Diskrimineringsvernet etter likestillings- og diskrimineringsloven.....	23
5.5.1	Direkte forskjellsbehandling	24
5.5.2	Indirekte forskjellsbehandling.....	25
5.6	Den diskrimineringsrettslige bevisbyrderegelen	26
6	Saker som har anvendt maskinl�ring for kontroll	27
6.1	Innledning.....	27
6.2	Bokontroll til L�nekassen i samarbeid med Riksrevisjonen	28
6.3	Veneziakommisjonens rapport om AWIR-saken fra Nederland.....	29
6.4	SyRI-saken fra Nederland	31
7	Diskrimineringsvernets anvendelse p� overnevnte saker.....	34
7.1	Innledning.....	34
7.2	Foreligger det forskjellsbehandling?	34
7.3	«Noe mer enn en ubetydelig ulempe»	37
7.4	Rettferdiggj�ring	41
7.5	Konklusjon	43
	Referanseliste	46
	Lovregister	46
	Forarbeider, Norges offentlige utredninger og forskrift.....	47
	H�yesterettspraksis.....	48
	Avgj�relser fra Den europeiske menneskerettighetsdomstolen	48
	Forvaltningspraksis, avgj�relse fra Nederland og Veneziakommisjonen	49
	Direktiv, forordninger og konvensjoner	49
	Litteratur.....	50

1 Innledning

1.1 Tema og problemstilling

Tema for denne avhandlingen er kunstig intelligens ved maskinlæring som brukes til å velge ut personer for etterkontroll av mottatte velferdsgoder. Problemstillingen er om det kan oppstå brudd på diskrimineringsforbudet ved anvendelse av maskinlæring i kontroll av borgere som har mottatt goder gitt av velferdsstaten. Med kontroll menes det tilsyn eller en form for overvåkning for å sjekke om det er gitt korrekte opplysninger eller at det ikke foreligger ulovligheter.

Diskrimineringsvernet genererer flere underproblemstillinger som må behandles for å kunne konkludere med at det foreligger ulovlig forskjellsbehandling. For det første må det tas stilling til om det foreligger forskjellsbehandling. For det andre må denne forskjellsbehandlingen behandle eller stille noen dårligere enn andre. Til sist må det vurderes om forskjellsbehandlingen likevel er lovlig på grunn av at den har et saklig formål, forskjellsbehandlingen er nødvendig for å oppnå formålet, og den ikke er uforholdsmessig inngripende overfor den som utsettes for forskjellsbehandlingen.

Diskrimineringsretten er spesielt viktig når det kommer til digitalisering. Dette fordi maskinlæringsprosesser kan forsterke tidligere feilaktige avgjørelser, og dermed anses å være forutinntatt. Det å være forutinntatt vil si at meningene er farget av tidligere historikk. Dette kan føre til at positive samfunnsendringer, som for eksempel likestilling, vil stagnere. Videre kan forutinntatte avgjørelser føre til at individer i samfunnet ikke får den hjelpen de trenger, eller at enkelte blir urimelig forskjellsbehandlet.

1.2 Aktualitet

Både privat og offentlig sektor har gjennomgått, og gjennomgår en digitaliseringsprosess der kunstig intelligens er en av teknologiene som benyttes. Digitalisering er et stort felt som har mange underkategorier. En underkategori til digitalisering er kunstig intelligens. En ytterligere underkategori av kunstig intelligens er maskinlæring.

I 2020 publiserte regjeringen en nasjonal strategi for kunstig intelligens.¹ Det skrives innledningsvis i den nasjonale strategien for kunstig intelligens at «Regjeringen vil at Norge

¹ Nasjonal strategi for kunstig intelligens, Kommunal og moderniseringsdepartementet, januar 2020.

skal gå foran i utvikling og bruk av kunstig intelligens med respekt for den enkeltes rettigheter og friheter».² Dersom Regjeringen vil at Norge skal være verdensledende på kunstig intelligens er det behov for at den rettslige reguleringen må være i tråd med dagens diskrimineringsrett slik at maskiner ikke kan diskriminere mennesker.

Menneskers rettigheter omhandler blant annet retten til å ikke bli diskriminert. Dette har fått stor oppslutning i samfunnsdebatten når det kommer til bias i maskinlæring. Selv om vi kan fastslå av resultatet at noen blir forskjellsbehandlet, så er det imidlertid viktig å vurdere hvordan man skal håndtere slike saker i fremtiden.

I dagens digitale verden hvor endringer skjer raskt, er det viktig å være forut for uønskede hendelser. Ved hjelp av kunstig intelligens kan en maskin utføre tidkrevende og kjedelige arbeidsoppgaver, som vi mennesker ikke vil utføre, og dermed frigjøre tid til å utføre andre arbeidsoppgaver. Før i tiden gikk forskjellige servicetilbud via telefon eller samtale med mennesker. I dag er disse servicetilbudene utviklet slik at flere av kundeserviceoppgavene nå behandles av forskjellige «chatbotter». Dette har spart den enkelte servicemedarbeideren for mye tid, som nå kan bli brukt på andre oppgaver.

Maskiner har muligheten til å prosessere store mengder data på svært kort tid ved hjelp av algoritmer, noe som tilsier at bruken av kunstig intelligens vil være både tidsbesparende og kostnadseffektivt. En algoritme er et sett av regler som benyttes på et datasett for å oppnå et ønsket resultat. Dersom man ser på bruken i et samfunnsmessig perspektiv vil man se at kunstig intelligens er noe vi allerede bruker i hverdagen selv om vi ikke alltid tenker over at dette er kunstig intelligens.

Ta for eksempel søknadsprosessen for studiestøtte fra Lånekassen, som bruker automatiserte vedtakssystemer. Gjennom søknadsprosessen får du flere forskjellige bakgrunnsspørsmål, som for eksempel om man har ansvar for barn, eller om man bor hjemme hos sine foreldre. Ut fra svarene man gir har algoritmen lært seg at denne studenten oppfyller vilkårene til å motta en viss sum i studiestøtte. Hele prosessen gjøres digitalt hvor en ikke er i kontakt med en menneskelig saksbehandler. Hvis man så oppfyller alle vilkårene, får man vedtaket innen

² Nasjonal strategi for kunstig intelligens, Kommunal og moderniseringsdepartementet, januar 2020.

få minutter. Denne algoritmen har dermed blitt programmert til å innvilge søknader om studiestøtte utfra en modell som man kan knytte til en «hvis x, så y-modell».

Forvaltningspraksis hvor man tidligere har hatt saksbehandlere til å utøve lovfestet forvaltningsrett kan bli optimalisert ved hjelp av kunstig intelligens. Dette handler imidlertid mer om hvordan teknologien brukes enn hva slags teknologi det er.

Automatiserte vedtakssystemer står i motsetning til beslutningsstøttende systemer der vedtaket fortsatt fattes av et menneske, men teknologi brukes underveis. I beslutningsstøttende systemer vil datasystemet først velge ut individer som må kontrolleres, for så å bli fulgt opp av en menneskelig saksbehandler.³

Kunstig intelligens kan være hjelpelig til å optimalisere en prosess, men det kan også medføre uønskede konsekvenser for enkeltpersoner eller grupper av mennesker.⁴ Et eksempel er personer som blir feilaktig tatt ut i etterkontroll av velferdsgoder, og som følge av dette mister støtte på flere tusen kroner.

Noen ganger er det ikke mulig å forklare utfallet på grunn av svarte bokser som ikke gir videre forklaring av datagrunnlaget som ble matet til modellen. Det å forklare et resultat eller en handling er svært viktig i diskrimineringsretten fordi det er nødvendig i prosessen når man skal vurdere om en eventuell forskjellsbehandling er lovlig eller ikke.⁵ Svarte bokser kan føre til at uønskede resultater ikke kan forklares. Dette vil kunne føre til en uheldig virkning av blant annet fremtidig treningsdata. Dersom personer har blitt satt i grupper hvor det er umulig å se hvilke kjennetegn ved personen som gjør at de plasseres i gruppene, vil dette kunne forårsake kontroll- og forklaringsproblemer.

1.3 Rettskilder og metode

1.3.1 Anvendte rettskilder

Kongeriket Norges Grunnlov av 17. mai 1814 oppstiller et generelt diskrimineringsforbud som gir et grunnlovsmessig vern. Den sentrale lovfestede rettskilden for oppgaven vil imidlertid være lov 16 juni nr. 51 om likestilling og forbud mot diskriminering (likestillings-

³ Kunstig intelligens og personvern, Rapport fra Datatilsynet januar 2018 s. 17.

⁴ Se avhandlingens kapittel 6.

⁵ Se avhandlingens 5.6.

og diskrimineringsloven heretter ldl.). Loven gir både rettigheter for menneskene i Norge, og plikter for det offentlige og andre virksomheter. Eksempelvis inkluderer loven retten til å ikke bli ulovlig forskjellsbehandlet og aktivitetsplikten for offentlige myndigheter og virksomheter.

Flere menneskerettighetskonvensjoner har særlig betydning på likestillings- og diskrimineringsfeltet. Blant annet FNs konvensjon 21. desember 1965 om avskaffelse av alle former for rasediskriminering (Rasediskrimineringskonvensjonen/RDK) og FNs konvensjon 18. desember 1979 om avskaffelse av alle former for diskriminering av kvinner (Kvinnediskrimineringskonvensjonen/KDK). Etter menneskerettsloven §§ 2 og 3 blir KDK inkorporert og gitt forrang etter norsk rett. Ldl. § 5 inkorporerer RDK som norsk lov.

Nasjonale domstoler har per dags dato ikke behandlet saker om diskriminering gjort av maskinlæringssystemer. Det vil imidlertid bli vist til relevante dommer fra Høyesterett og avgjørelser fra Likestillings- og Diskrimineringsombudet (heretter LDO) eller Diskrimineringsnemnda gjennom drøftelsen av de nasjonale rettskildene.⁶ LDO behandlet klagesaker innenfor diskrimineringsretten fram til 2018. I dag blir imidlertid klagesaker behandlet av Diskrimineringsnemnda. Nemnda er et «uavhengig forvaltningsorgan», og kan ikke instrueres i sin faglige virksomhet.⁷

Den rettskildemessige vekten av nemndas eller ombudets avgjørelser stilles på lik linje som annen forvaltningspraksis.⁸ Det vises imidlertid til at nemndas avgjørelser ofte viser tilbake til sin egen praksis, og at en del avgjørelser har vært bidragsytende til at nasjonal lovgivning og dens ordlyd blir tilpasset til å passe med nemndas praksis⁹. Nemnda skal blant annet være spesialister på likestillings- og diskrimineringsfeltet, og kan bli omtalt som et ekspertorgan.

⁶ Diskrimineringsombudet behandlet klagesaker innenfor diskrimineringsretten fram til 2018, i dag blir klagesaker behandlet av Diskrimineringsnemnda.

⁷ Lov 16. juni 2017 nr. 50 Likestillings- og diskrimineringsombudet og Diskrimineringsnemnda (Diskrimineringsombudsloven diskrol.)

⁸ Likestillings- og diskrimineringsrett, Midlertid kompedieutgave for studenter av Anne Hellum og Vibeke Blaker Strand 2021, s. 50-51

⁹ Likestillings- og diskrimineringsrett, Midlertid kompedieutgave for studenter av Anne Hellum og Vibeke Blaker Strand 2021, s. 51

Dette trekker i retning av at deres uttalelser har rettskildemessig vekt, men at dette må vurderes på bakgrunn av kvaliteten på vedtak eller uttalelser.¹⁰

Videre er det tatt utgangspunkt i mange rapporter og strategier utarbeidet av ekspertgrupper som er oppnevnt på vegne av staten eller Den europeiske union.¹¹ Rapportene gir bakgrunnsinformasjon av relevant litteratur på diskrimineringsrett- og digitaliseringsområdet. Disse kildene kan betegnes som sekundære rettskilder og faller inn under betegnelsen «juridisk litteratur».¹²

I kapittel 3, 4 og 5 vil det bli benyttet juridisk litteratur fra nasjonale og internasjonale kilder. Teorien kan i enkelte tilfeller ha stor innflytelse på diskrimineringsretten. Dette fordi forvaltningen viser til teorien, og den kan bli brukt under tilsyn av virksomheter. De internasjonale og nasjonale kildene er blant annet Datatilsynet og ekspertgruppen fra Europakommisjonen samt juridiske professorer som er kjent innenfor likestillings- og diskrimineringsretten.

1.3.2 Metode

I fravær av nasjonale rettsavgjørelser og forarbeider er det tatt utgangspunkt i sakene som behandles i avhandlingens kapittel 6. Et sentralt punkt i denne avhandlingen er den kunstige intelligensen og algoritmene som brukes. I de aktuelle sakene som gjennomgås i kapittel 6, vil det ikke alltid være mulig å få innsyn i hvordan algoritmene fungerer eller hvordan de har kommet frem til resultatet. Dette på grunn av systemene er basert på svart bokser.

Man kan imidlertid på bakgrunn av sakene se noen fellestrekk som viser hvordan forvaltningen innenfor diskrimineringsretten møter nye hindringer når de bruker nye verktøy som maskinlæring for å effektivisere en prosess. De tre sakene i avhandlingens kapittel 6 viser hvordan maskinlæring kan bli brukt i kontrollering av mennesker som har fått et gode av staten. Ved å vise til tre saker kan man enklere finne likheter samt oppdage tendenser som kan oppstå når man bruker maskinlæring.

¹⁰ Likestillings- og diskrimineringsrett, Midlertid kompedieutgave for studenter av Anne Hellum og Vibeke Blaker Strand 2021, s. 51

¹¹ Se for eksempel avhandlingens kapittel 3 for definisjoner.

¹² Rett og rettsanvendelse av Jens Edvin A. Skoghøy, Universitetsforlaget 2018 s. 215.

1.4 Avgrensning av avhandlingen

Det er naturlig å avgrense denne avhandlingen grunnet alle områdene som kunstig intelligens og diskrimineringsretten treffer. Diskrimineringsretten er et stort felt og i sammenheng med kunstig intelligens oppstilles det en rekke problemstillinger som er interessant å fordype seg i.

Eksempelvis er det ofte spørsmål innenfor diskrimineringsretten om hvem som er ansvarlig for diskrimineringen, eller hvem som skal stå ansvarlig for den som har diskriminert. Hvilke av den behandlingsansvarlige, databehandleren eller en tredjepart¹³ av algoritmene som er ansvarlig for diskriminering gjort av maskinlæring er et spørsmål som kunne ha vært aktuelt tema. Overnevnte eksempler vil imidlertid ikke bli behandlet noe nærmere grunnet plasshensyn.

En sentral del av oppgaven omhandler kunstig intelligens herunder maskinlæring. EU-kommisjonens forslag til ny forordning om ansvarlig kunstig intelligens vil ikke være bli omtalt i avhandlingens hoveddel fordi forordningen ikke er vedtatt. Forordningens Del III som omhandler KI-systemer som innebærer høy risiko for personers helse, sikkerhet eller grunnleggende rettigheter kunne vært en relevant problemstilling, men siden det sentrale i oppgaven allerede er vedtatt i lover blir disse brukt for å besvare oppgavens problemstilling..

Det tas utgangspunkt i kontroller gjennomført av offentlige myndigheter i Nederland og Norge, men det er kun nasjonal diskrimineringsrett som blir anvendt med sider til den internasjonale likestillings- og diskrimineringsretten. De to sakene fra Nederland er valgt fordi sakene er relevante når det kommer til hvordan maskinlæring kan bli brukt i et kontrollperspektiv. I tillegg er diskrimineringsretten et felleseuropeisk regelverk med like bestemmelser som våre nasjonale rett slik at man enklere kan anvende reglene på andre lands saker.

1.5 Videre fremstilling

Avhandlingens overordnede tema er det offentliges kontroll med borgerne i et velferds- og diskrimineringsperspektiv. Først vil profilering som ledd i kontroll bli gjennomgått, og hvorfor slik profilering kan anses som forskjellsbehandling i diskrimineringsretten.

¹³ Personopplysningsloven § 1, jf. personvernforordningen artikkel 4 nummer 7,8 og 10. Definisjon på «behandlingsansvarlige», «databehandleren» og «tredjepart».

Deretter skal det redegjøres for kunstig intelligens og underkategorier som maskinlæring. For å kunne vurdere diskrimineringsretten opp mot kunstig intelligens og maskinlæring er det nødvendig å forstå hvorvidt og i hvilken utstrekning kunstig intelligens og maskinlæring kan resultere i forutinntatte resultater.

Det vil ikke gis en full fremstilling av diskrimineringsretten, men det legges vekt på noen temaer som er særlig relevant for oppgavens problemstilling. Diskrimineringsretten inneholder svært mange beskyttende regler og plikter for ansvarlige, som det er nødvendig å få en oversikt over før man vurderer hvordan disse reglene skal anvendes på kunstig intelligens.

I kapittel 6 av avhandlingen vil det gis eksempler på det offentlige bruk av kunstig intelligens i forvaltningen, og statens kontroll av velferdsgoder som befolkningen mottar fra staten. Det vises her til en sak fra Norge, samt to saker fra Nederland. Sakene som er fra utlandet, vil ha en annen oppbygging av rettsreglene enn man har i Norge. Ved å vise til sakene og bruke nasjonal diskrimineringsrett, vil man kunne vise til hvilke diskriminerings spørsmål som oppstår i saker som omhandler kontroll av borgerne ved bruk av maskinlæring.

Avslutningsvis vil det gis avsluttende bemerkninger på hvordan veien videre kan se ut for kunstig intelligens innenfor diskrimineringsretten, og trekkes en konklusjon for sakene.

2 Profilerings til kontroll som diskrimineringsrettslig problem

2.1 Innledning

Både politiet og andre offentlige myndigheter bruker kontroll av borgere for å oppdage uregelmessigheter, og for å forebygge fremtidige uregelmessigheter. Det å foreta forebygging i det offentlige er en av politiets hovedstrategier.¹⁴ Diskriminering på bakgrunn av etnisk profilering og digital utvelgelse av personer for kontroll har en rekke kontaktflater. Som for eksempel at det er etnisiteten til personer som blir tyngdepunktet i avgjørelsen for utvelgelse. I avhandlingens eksempler vil det imidlertid være kunstig intelligens, nærmere bestemt maskinlæringsprogrammer, som velger ut personer på bakgrunn av parametere som tilsier at

¹⁴ I forkant av kriminaliteten, Forebygging som politiets hovedstrategi (2021–2025), Politiet, s. 3

de er en høyere risiko enn andre grupper. Kontrollen av velferdsgoder kan sammenlignes av den kontrollen politiet gjør av personer når de skal sjekke deres identitet i forbindelse med oppdrag.

Etter internasjonale regler oppstilles det et vern mot stereotypisering. Etter KDK artikkel 5 skal stereotypering på grunnlag av manns- eller kvinneroller ikke forekomme så lenge de forsterker fordommer i samfunnet. Stereotypisering kan trekke paralleller til raseprofilering hvor stereotypering på grunnlag av rase/etnisitet ikke skal skje. Stereotypisering er når man har en generell tilnærming til grupper som nødvendigvis ikke har forankring i virkeligheten

I plikten til å fremme likestilling og forhindre diskriminering innebærer dette «en plikt til å motarbeide de underliggende årsakene til diskriminering, for eksempel stereotyper eller fordommer som regner noen personer eller grupper som mindreverdige på grunn av deres personlige kjennetegn».¹⁵

Kontroll av enkeltborgere er ikke en ny problemstilling i diskrimineringsretten.

Raseprofilering kan oppstå under en kontroll om kontrollen blir gjort på feilaktig grunnlag som stereotypisering. Den nasjonale diskrimineringsretten anvender ikke begrepet «rase» som et diskrimineringsgrunnlag, men etter internasjonale kilder vil dette være nært knyttet til diskrimineringsgrunnlaget «etnisitet».¹⁶ Raseprofilering er imidlertid et begrepet som er kjent på folkemunne, og kan også bli omtalt som etnisk profilering.

I tilfeller hvor personer med annen etnisitet blir kontrollert kan det være uklart om diskrimineringsvernet kommer til anvendelse eller ikke. For å finne ut om det foreligger diskriminering, må man i hver enkelt sak foreta en konkret vurdering på bakgrunn av dens faktum.

2.2 Eksempler på kontroll av personer på bakgrunn av etnisitet

I en sak fra 2013 kom Diskrimineringsombudet frem til at det ikke forelå diskriminering på grunn av etnisitet.¹⁷ En kvinne og hennes venninner ble stoppet av politiet på vei ut av et

¹⁵ Prop.81 L (2016–2017) punkt 2.5.3.1 s. 251.

¹⁶ Se for eksempel FNs rasediskrimineringskonvensjon artikkel 1 nummer 1 hvor «rase» og «etnisitet» faller inn under det samme diskrimineringsvernet.

¹⁷ Diskrimineringsombudet sak 13/615.

utested på for at politiet skulle foreta en identitetskontroll av kvinnen. Grunnlaget for å foreta en identitetskontroll var at politiet hadde mottatt to separate tips fra publikum om at kvinnen solgte seksuelle tjenester til utestedets kunder. Kvinnen hadde en annen etnisk opprinnelse enn norsk, og opplevde kontrollen som svært fornærmende og nedverdiggende.

På vegne av kvinnen anmeldte en mann politiet for diskriminering på grunnlag av etnisitet/hudfarge. Mannen mente at politiet urettmessig hadde stoppet kvinnen og diskriminert henne på bakgrunn av hennes etnisitet. Han anførte at identitetskontrollen ikke ville ha blitt utført om hun var en hvit etnisk norsk kvinne.

Ombudet kom frem til at det ikke forelå diskriminering på bakgrunn av etnisitet i denne saken, men la til grunn at det kan foreligge diskriminering fra offentlig myndighet om en identitetskontroll gjør at noen «blir behandlet dårligere enn andre».¹⁸

I en sak fra 2020 kom Diskrimineringsnemnda frem til at det forelå diskriminering av en person på bakgrunn av hans etnisitet.¹⁹ Personen hadde blitt stoppet av politiet for kontroll ti ganger i perioden 2015-2020. Politiet hadde kun loggført tre av ti kontroller av mannen. På grunn av den manglende loggføringen var det vanskelig å etterprøve hvorfor mannen ble stoppet.

Politidistriktet vedkjente at mannen var stoppet flere ganger i forhold til andre i befolkningen, og at dette vanskelig kunne forklares på annen måte enn at det var mannens etnisitet som var grunnlaget for kontrollene.²⁰ Nemnda kom frem til at det både forelå direkte og indirekte diskriminering av mannen fra politiet på grunn av kontrollene. Mannens etnisitet var det som var avgjørende for at han ble stoppet av politiet.

2.2.1 Konklusjon

Spørsmålet om det er forskjellsbehandling i de tilfellene noen behandles eller stilles dårligere må besvares med en konkret vurdering i hver enkelt sak.

I den første saken fra 2013 viste det seg at politiet allerede hadde stoppet rundt 50 personer i området den kvelden i forbindelse med forebygging av prostitusjon. I tillegg viste politiet til

¹⁸ Se sak 13/615 under «ombudets vurdering».

¹⁹ DIN-2020-175.

²⁰ DIN-2020-175 under nemndas vurdering: Direkte diskriminering ved hyppige politikontroller.

at de kunne utføre identitetskontroll av antatte utenlandske statsborgere på spontan basis hvis det forelå en grunn til å sjekke personen.²¹ Kontrollen hadde også foregått utenfor utestedet ved stengetid. Med andre ord var det få personer i nærheten da kontrollen tok sted. Samlet sett vurderte ombudet at antall personer som var blitt kontrollert, utlendingslovens § 21 og tidspunktet for kontrollen medførte at det ikke forelå ulovlig forskjellsbehandling av kvinnen og hennes venninner.

I den andre saken fra 2020 la nemda vekt på at hyppigheten, og mangelen på loggføringen, av kontrollene medførte at personen ble diskriminert på grunn av hans etnisitet. Det er selve hyppigheten og omfanget som her utgjorde at personen ble «behandlet dårligere» og at det dermed var en klar belastning for personen.²²

Konklusjonen er dermed at hyppigheten av kontroller kan medføre en større belastning for personen som blir kontrollert. For å konkludere at det foreligger diskriminering kan man ikke bare se på resultatet. Det er også nødvendig å se på prosessen som førte til resultatet. Sett i sammenheng med avhandlingen vil svart boks-problematikken vanskeliggjøre denne prosessvurderingen.

Et forslag som kan være behjelpelig for å motvirke raseprofilering er en kvitteringsordning. Ved å opprette en kvitteringsordning vil de som blir stoppet i kontroll av politiet motta en skriftlig dokumentasjon på at de er tatt ut i kontroll og på hvilket grunnlag dette er gjort. Dette vil være med på å skape tillit overfor politiet, men også bidra til notoritet i kriminalitetsbekjempelsesarbeidet som politiet gjør. Byrådet i Oslo kommune har for eksempel gått inn for at politiet i Oslo skal begynne med kvitteringsordning.²³

²¹ Lov 15. mai 2008 nr. 25 om utlendingers adgang til riket og deres opphold her (utlendingsloven) § 21.

²² DIN-2020-175 under nemdas vurdering: Indirekte diskriminering ved at politiet bare unntaksvis loggfører kontrollene.

²³ Plattform for byrådssamarbeid mellom Arbeiderpartiet, Miljøpartiet De Grønne og Sosialistisk Venstreparti i Oslo 2019-2023 s. 48.

3 Nærmere om kunstig intelligens og underkategorier som maskinlæring

3.1 Innledning

Det finnes ingen universal definisjon av kunstig intelligens. Forklaringen kan være at det er et stort område som stadig er i endring. Som følge av at området utvikler seg, vil også definisjonen endre seg. En ekspertgruppe nedsatt av EU har definert kunstig intelligens til å være «systemer [som] utfører handlinger, fysisk eller digitalt, basert på tolkning og behandling av strukturerte eller ustrukturerte data, i den hensikt å oppnå et gitt mål. Enkelte KI-systemer kan også tilpasse seg gjennom å analysere og ta hensyn til hvordan tidligere handlinger har påvirket omgivelsene».²⁴ Denne definisjonen er lagt til grunn i den nasjonale strategien og vil også bli lagt til grunn som definisjonen av kunstig intelligens i avhandlingen.²⁵

3.2 Ytterligere begreper innenfor domenet «kunstig intelligens»

Som nevnt ovenfor er det mange ulike typer definisjoner av kunstig intelligens. Kunstig intelligens kan derfor tenkes som en paraplybetegnelse for andre datasystemer som faller inn under betegnelsen. Dette vil for eksempel være systemer som jobber med maskinlæring, regelbaserte algoritmer, eller dyplærende algoritmer.

Med regelbaserte algoritmer menes systemer som følger ett sett med instruksjoner. Dette kan for eksempel forklares med hvis x, gjør y.²⁶ Fordelen med regelbaserte algoritmer er at de ofte er forutsigbare, ettersom de følger et mønster som finnes i datagrunnlaget gitt av mennesker.²⁷ Systemet vil ikke selv søke etter mer datagrunnlag for å kunne videreutvikle seg. Systemet vil

²⁴ Independent High Level Expert Group set up by the European Commission (2018): A definition of AI: Main capabilities and disciplines. Oversettelse er publisert i Nasjonal strategi for kunstig intelligens av regjeringen fra 2020.

²⁵ Nasjonal strategi for kunstig intelligens, Kommunal og moderniseringsdepartementet, januar 2020, s. 9.

²⁶ Nasjonal strategi for kunstig intelligens, Kommunal og moderniseringsdepartementet, januar 2020, s. 10.

²⁷ Nasjonal strategi for kunstig intelligens, Kommunal og moderniseringsdepartementet, januar 2020, s. 11.

kun fokusere og arbeide med de parameter som er blitt gitt.²⁸ Regelbaserte algoritmer vil være til stor hjelp i arbeid som behandler og sorterer store mengder med data.

Maskinlæring kan defineres som «et sett teknikker og verktøy som lar maskiner «tenke» ved å lage matematiske algoritmer basert på akkumulert data».²⁹ Det karakteristiske ved maskinlæring er at det er datagrunnlaget som bidrar til å utvikle systemet. Ved å lære seg datagrunnlaget, tilpasse seg dette og utvikle seg, vil maskinlærende kunstige intelligens utvikle seg selv ut fra det datamaterialet det blir tilbudt.

Innenfor maskinlæring kan algoritmen enten lære seg å klassifisere datagrunnlaget, sortere i klynger, foreta regresjonsanalyser eller foreta assosiasjonsanalyser.³⁰ Når data blir klassifisert settes datagrunnlaget i grupper for å se hvem eller hva som hører til hver kategori.

I klyngeteknikken vil algoritmen identifisere sterke fellestrekk eller sammenhenger mellom tilsynelatende svært forskjellig data. Et eksempel på klyngeteknikk er når en kundegruppe til en matvarekjede anbefaler billige matvareprodukter til de kundene som har lagt inn informasjon om at de er studenter.³¹ I regresjonsteknikken vil algoritmen finne en sammenheng mellom en lik eller flere ulike parametere som er satt. For eksempel dersom man ønsker å forutsi en prisøkning, forutsi lovbrudd eller forutsi etterspørsel på en vare.³²

Assosiasjonsanalyser gjelder data som kan knyttes opp til hverandre på en eller annen måte. Algoritmen vil da kunne anta hva som er neste steg i prosessen. For eksempel dersom du ser på flybilletter til Alicante, vil en assosiasjonsanalyse gjøre at man får opp forslag til hoteller eller leiebiler i samme by. Dette er mye brukt av reiseselskaper som for eksempel Norwegian eller SAS.

²⁸ Kunstig intelligens og personvern, Rapport fra Datatilsynet januar 2018 s. 6.

²⁹ Kunstig intelligens og personvern, Rapport fra Datatilsynet januar 2018 s. 5.

³⁰ Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law. Rapport av Janneke Gerards (Utrecht University) og Raphaële Xenidis (University of Edinburgh and University of Copenhagen) 2020, punkt 1.2.2.1 s. 33 og 34.

³¹ Elements of AI under punktet Maskinlæring, forskjellige typer maskinlæring
<https://course.elementsofai.com/no/4/1>

³² Elements of AI under punktet Maskinlæring, regresjon <https://course.elementsofai.com/no/4/3>

Innenfor både klyngeteknikken, regresjonsteknikken og assosiasjonsanalyser vil maskinlæring lære og utvikle seg på forskjellig måte. Systemet kan enten lære etter en veiledet læring, uveiledet læring eller forsterkende læring.³³ Hvilke type maskinlæring man skal anvende kommer an på hvilket datagrunnlag man skal sortere og hvilket resultat man ønsker.

Med veiledet læring menes det at systemet har blitt matet med data som er valgt ut og kategorisert på et tidligere stadium. Datasettet vil bli delt opp i to deler hvor den ene delen blir brukt til å lære opp modellen og den andre delen blir brukt for å kontrollere hvor god modellen er på ny data. Den første delen blir omtalt som treningsdata og den andre delen som testdata.³⁴

Modellen skal kunne anvende regelen som den har lært seg fra treningsdataen til testdataen som den blir matet med. For å vurdere hvor god systemet er på testdataen, så kan ikke treningsdataen bli benyttet lenger. Dersom det er et stort avvik på testdataen fra treningsdataen, kan man konkludere med at det ikke er en god modell som er blitt laget. Om systemet derimot viser seg å stemme overens både med trenings og testdataen, så vil dette være en indikasjon på en god modell.

Et eksempel på test og treningsdata kan være å sortere transportmiddelet «bil» og «sykkel» hver for seg. Sykler har to hjul og biler har fire hjul. Biler har et ratt som er rundt, og slike kjennetegn kan gjøre det enklere å sortere datagrunnlaget. Det kan imidlertid bli vanskelig for maskinen når man får opp el-sykler med motor, slik som biler også har. Da er det andre kjennetegn som modellen må lære seg for å kunne skille de to kategoriene fra hverandre.

I bruk av uveiledet læring skal systemet selv klare å gruppere og kategorisere datagrunnlaget som den blir matet med.³⁵ I dette tilfellet vil algoritmen ikke få testdata som viser den riktige fasiten slik som den vil under veiledet læring. Dyplærende algoritmer er en slik type maskinlærende algoritme som er uveiledet, og som tar i bruk store mengder data for å lære seg opp. En type dyplærende algoritme er ansiktsgjenkjenning hvor systemet selv skal kunne se at det er et menneske ved å gjenkjenne nese, ører, øyne og munn. Ansiktsgjenkjenning

³³ Kunstig intelligens og personvern, Rapport fra Datatilsynet januar 2018 s. 7-10.

³⁴ Kunstig intelligens og personvern, Rapport fra Datatilsynet januar 2018 s. 7.

³⁵ Kunstig intelligens og personvern, Rapport fra Datatilsynet januar 2018 s. 8.

fungere imidlertid ikke optimalt for alle brukere. Mørkhudede kvinner har for eksempel en lavere sjans for å bli oppdaget av maskinen enn andre.³⁶ Dette kan være på grunn av at det er få mørkhudete kvinner som ble brukt i datagrunnlaget.

Forsterkende læring gjelder de modellene som bruker uveiledet læring for å bygge opp en beslutning basert på prøv og feil-metoden. Et menneske ser over beslutningen for så å se om den er riktig eller feil.³⁷ Modellen vil endre og forbedre seg til det formålet den skal bli brukt til som følge av hva mennesket i loopen bestemmer er riktig.

Maskinlærende algoritmer anses å være dynamisk og i en stadig endring, mens de regelbaserte algoritmene er statiske.³⁸ Maskinlæring kan i enkelte tilfeller være problematisk grunnet overtilpasning eller undertilpasning. Overtilpasning skjer når algoritmen presterer bra sammen med treningsdataen, men dårlig sammen med ny data fordi den ikke har lært seg å lage en god nok generell regel den kan følge. Undertilpasning er når algoritmen verken er god på treningsdataen eller ny data.

Svart boks, er ikke en type maskinlæring, men en betegnelse på at det ikke er mulig å forklare hvordan maskinlæringsalgoritmen har kommet fram til svaret. Man kan sammenligne en svart boks med dyplærende algoritmer hvor det ikke er mulig å få et klart svar på hvordan algoritmen kom frem til det svaret som det gjorde, eller hvilke parametere som har blitt lagt vekt på.³⁹ Det er ikke mulig å få et klart svar fordi man ikke har innsikt i hvordan algoritmen fungerer. Dyplærende algoritmer blir noen ganger formulert som om den etterligner en menneskehjerne med informasjonsstrøm ut og inn, uten å vite hvordan det påvirker systemet.

Selv om man tester resultatet til en maskinlærende algoritme med treningsdata, vil man ikke kunne åpne den svarte boksen for å få en forklaring på grunn av modellens oppsett.

Datatilsynet har vist til Skatteetaten sin prediksjonsmodell som plukker ut personer som

³⁶ Najibi, Alex «Racial Discrimination in Face Recognition Technology», Harvard University Blog, science policy, special edition: science policy and social justice, 24. oktober 2020.

³⁷ Kunstig intelligens og personvern, Rapport fra Datatilsynet januar 2018 s. 9.

³⁸ Kunstig intelligens og personvern, Rapport fra Datatilsynet januar 2018 s. 9.

³⁹ Kunstig intelligens og personvern, Rapport fra Datatilsynet januar 2018 s. 12.

systemet mener er en høy risiko for å ha oppgitt feil informasjon eller jukset på selvangivelsen sin.⁴⁰ I forbindelse med modellen sin uttaler Skatteetaten:

«Når vi bygger modellen på denne måten, vet vi ikke nødvendigvis hva det er som gjør at en bestemt skattyter blir rangert til å ha stor risiko for feil. Rangeringen er et resultat av komplekse sammensetninger av dataene i modellen».⁴¹

Dette er et eksempel på svart boks problemet hvor resultater ikke kan forklares.

4 Forutinntatt kunstig intelligens

4.1 Innledning

Når man har forstått hvilke typer kunstig intelligens som finnes, og hvordan de fungerer, så er neste steg å sette dette i kontekst med diskrimineringsretten. Videre må man da se hvordan kunstig intelligens kan bli forutinntatt og muligens diskriminerer. Dette er ofte omtalt som «bias» på engelsk. Det å være biased/forutinntatt betyr at man forhåndsdommer på bakgrunn av forutbestemte forestillinger eller fordommer som ikke har noe grunnlag i fakta. Det at kunstig intelligens blir forutinntatt handler om de historiske dataene som datasytemet blir matet med. Dette fordi de historiske dataene som skal være med på å trene opp og teste systemet er innhentet fra reelle opplysninger fra virkeligheten.

Selv om samfunnet har en mer tolerant holdning, og setter søkelys på ikke-diskriminering og mangfold, så forhåndsdommer fortsatt mennesker på bakgrunn av sine fordommer. Hvis resultatet av disse meningene senere blir brukt som historisk data vil datasettet forsterke tidligere menneskeskapte holdninger, handlinger og praksis.

Studier om diskriminering gjort av kunstig intelligens er i USA et relevant tema. I Europa og i Norge har ikke problematikken blitt like mye satset på. Dette er imidlertid på vei til å endres. Som nevnt i innledningen har både Europarådets ekspertgruppe om kunstig intelligens og den nasjonale regjeringens strategiplan kommet de siste årene.⁴²

⁴⁰ Skatteetatens Analysenytt mars 2016.

⁴¹ Skatteetatens Analysenytt mars 2016 s. 14.

⁴² Nasjonal strategi for kunstig intelligens, Kommunal og moderniseringsdepartementet, januar 2020.

USA er hjemlandet til store teknologigiganter som Google, Meta (Facebook Inc), Amazon, Apple og Microsoft. Alle disse selskapene bruker enorme ressurser på å drive frem den digitale tidsalder. USA vil med andre ord muligens ha et større omfang av saker som involverer kunstig intelligens og forskjellsbehandling både nå og i fremtiden.

Et eksempel fra USA er maskinlæringsystemet som Amazon anvendte for å ansette jobbsøkere. Systemet viste at mannlige jobbsøkere ble foretrukket fremfor kvinner.⁴³ Datasystemet som skulle være behjelpelig med å finne toppkandidatene blant alle søkerne ble bygget opp ved hjelp av firmaets ansettelseshistorikk 10 år tilbake i tid. Innenfor teknologibransjen arbeider det svært mange hvite og asiatiske menn, og når de historiske dataene ble anvendt på nye kandidater var parameteren «kvinne» noe som ikke gikk overens med tidligere ansettelsler.⁴⁴

I forsøket på å forstå hvilke typer diskriminering det er snakk om når det blir utført av kunstig intelligens, har Kate Crawford fastslått at kunstig intelligens kan skape ulemper på to forskjellige måter. Dette har hun klassifisert under «harms of allocations» og «harms of representation».⁴⁵

4.1.1 Harms of allocations

«Harms of allocations» eller skaden ved fordeling på norsk, omhandler hvordan systemer tildeler eller fratrukker en person en mulighet eller ressurser.⁴⁶ Ofte er denne typen forutinntatthet sett i sammenheng med økonomiske perspektiv. For eksempel hvem som får et gode fra staten, hvem som blir tildelt lån fra banken eller at et forsikringselskap bestemmer at kvinner må betale en høyere andel i bilforsikring fordi de angivelig er dårligere enn menn å kjøre bil ifølge systemet.

⁴³ Dastin, Jeffrey «Amazon scraps secret AI recruiting tool that showed bias against women», Reuters, 11. oktober 2018, lest 3. april 2022.

⁴⁴ Levitsky, Allison “For the first time, white men weren't the largest group of U.S. hires at Google this year”, Silicon Valley Business Journal, 5. Mai 2020. Lest 1. Mai 2022.

⁴⁵ The trouble with bias, Kate Crawford taler under Conference on Neural Information Processing Systems den 29. oktober 2017, sett 25. mars 2022.

⁴⁶ The trouble with bias, Kate Crawford taler under Conference on Neural Information Processing Systems den 29. oktober 2017, sett 25. mars 2022.

Konsekvensene av fordeling er umiddelbare som følge av at det er systemet som enten gir et positivt eller negativt svar alt ettersom hvem du er som person. I tillegg er det enklere å oppdage tilfeller hvor noen blir tildelt eller fratatt noe som helt åpenbart føles feil.

Kvinner og bilkjøring er ett eksempel som kan havne inn under skaden ved feilaktig fordeling. I 2020 var det 415 menn som ble hardt skadd mot 212 kvinner.⁴⁷ Hvis man tok en kvinne og en mann som har hatt samme kjørehistorikk, men kvinnen fikk en høyere månedlig forsikringspolise på forsikringen så ville dette føltes åpenbart feil for kvinnen. Spørsmålet man kan stille er hvilke parameter som har blitt lagt vekt på for å komme frem til svaret.

4.1.2 Harms of representation

«Harms of representations» omhandler skaden av representasjon, eller mangelen på representasjon. Denne problematikken knytter seg til de tilfellene hvor systemene gjør en utskilling av informasjonen i tråd med hva systemet har lært er riktig.⁴⁸ Ofte er denne type forutinntatthet sett i sammenheng med et kulturelt perspektiv og de fordommene eller oppfatningene man måtte ha i et samfunn.

Dette kan for eksempel være at de historiske dataene viser at det er et høyt antall menn som er ansatt som sivilingeniører i et firma. Hvis systemet dermed lærer seg at parameteren «mann» og «sivilingeniør» henger sammen, så vil dette ha utslag i sluttresultatet. Hvis disse parameterne for eksempel blir brukt i en ansettelsesprosess vil dette kunne medføre at menn som er sivilingeniører oftere blir foretrukket av programvaren enn kvinner som er sivilingeniører.

Kate Crawford legger til grunn at det finnes fem forskjellige typer av «harms of representations»:⁴⁹

Den første skadegjøringen er stereotypering. For eksempel «menn er leger og kvinner er sykepleiere». Den andre skadegjøringen er gjenkjenning. For eksempel hvor kvinner med mørk hud ikke blir gjenkjent av ansiktsgjenkjennelsesprogrammer. Den tredje skadegjøringen

⁴⁷ Statistisk sentralbyrå, trafikkulykker med personskade 2020, oppdatert 14. mars 2022.

⁴⁸ The trouble with bias, Kate Crawford foredragsholder under Conference on Neural Information Processing Systems den 29. oktober 2017, sett 25. mars 2022.

⁴⁹ The trouble with bias, Kate Crawford foredragsholder under Conference on Neural Information Processing Systems den 29. oktober 2017, sett 25. mars 2022.

er nedverdiggelse. Dette er tilfeller hvor personer eller grupper blir ydmyket på grunn av en gruppe de måtte tilhøre. For eksempel hvis man skriver inn «samer er» i søkefeltet til Google, og de umiddelbare forslagene som kommer opp er eksempler på at samer er hårsåre eller late.

Den fjerde skadegjøringen er underrepresentasjon. I dette tilfellet viser ikke datasystemet et representativt bilde av samfunnet. For eksempel hvor Google som oftest viser bilder av hvite menn når man søker på ordet «direktør». Den femte skadegjøringen kalles «ex-nomination» som ikke har fått et tilsvarende navn på norsk, men direkte oversatt betyr eks-nominasjon. Dette er de tilfellene hvor typetilfeller ikke blir tildelt i sin korrekte kategori eller ikke blir navngitt i det hele tatt. Herunder «lege» og «kvinnelig lege». Ved å bruke «kvinnelig lege» vil dette tilsi at det er en annen kategori enn «lege» og dermed ikke falle inn under korrekt definisjon.

Kate Crawford anser det som vanskelig å oppdage skadene som følger av representasjon. Dette fordi skaden defineres å være av en langvarig karakter, de er vanskelig å kunne gjengi og diffus i sin opptreden.⁵⁰

Et spennende aktuelt tilfelle som viser eksempelet på skaden som representasjon kan gjøre er testen NRK⁵¹ har gjort ved å opprette to brukere på sosiale medier appen TikTok.⁵² Ved hjelp av to maskiner opprettet NRK to brukere på appen med IP-adresser som er tilknyttet en by i Ukraina og en i Russland. I tillegg fremstår maskinene som om de er 19 år gamle menn. IP-adressen, alder og kjønn gir informasjon til algoritmene som TikTok bruker at man ønsker videoer som enten er populære i det området, eller som vanligvis passer for 19 åringer. I dette tilfellet ønsket NRK å se hvilke videoer to menn i militæralder fikk.

Den ukrainske brukeren fikk i løpet av noen minutter flere videoer av krigen i Ukraina, og menn som var klar for å kjempe eller andre patriotiske videoer. Den russiske brukeren fikk

⁵⁰ The trouble with bias, Kate Crawford foredragsholder under Conference on Neural Information Processing Systems den 29. oktober 2017, sett 25. mars 2022.

⁵¹ Bøe og Bjørke, «TikTok viser ikke krigen i Ukraina til brukere i Russland», publisert 03.04.2022, sist lest 13. april 2022.

⁵² En mobilapplikasjon som tillater brukerne å dele videoer på opptil 3 minutter med andre brukere. Appen har en startside som heter «for you» hvor algoritmene skal tilpasse hvilke videoer fra andre brukere som passer perfekt til deg på bakgrunn av andre videoer du liker, ser på, søkeord du bruker osv.

opp fire videoer om Ukraina. Ingen omhandlet krigen, men viste videoer av dansing og maling.

Om den russiske brukeren prøver å søke på videoer om krigen så får den ikke opp noe. I spørsmål fra NRK til TikTok om hvorfor brukerne får opp forskjellige videoer vises det til Russlands nye «fake news»-lov.

Tilfellet gir et eksempel på skaden ved underrepresentasjon hvor algoritmene bestemmer hva du får se eller ikke se. Eller for eksempel stereotypisering hvis det blir antatt at russerne ikke er interessert i krigen.

5 Nasjonal lovgivning innenfor diskrimineringsretten

5.1 Innledning

Følgende kapittel gir ikke et komplett bilde av diskrimineringsretten, men de viktigste som kan være innenfor problemstillingen maskinlæring brukt til kontroll. Det vil være nødvendig å få en oversikt over diskrimineringsrettslige begrep og vilkår som kan være aktuell å benytte i saker hvor fysiske personer har blitt kontrollert ved hjelp av maskinlæring. Dette fordi for å kunne vurdere en eventuell forskjellsbehandling oppstiller flere underspørsmål som behandlet i innledningen.

5.2 Positive plikter og negative forbud

I norsk rett er det ingen konkrete norske rettskilder som omhandler kunstig intelligens og diskriminering. Dette er muligens fordi det er et nytt område innenfor jussen, eller fordi den lovgivende makt sett at dette muligens ikke er et behov enda. Det er også vanskelig å oppdage (u)tilsiktet diskriminering,⁵³ noe som kan være grunnen for at vi ikke har flere tilfeller nasjonalt. Norge har imidlertid over flere år jobbet kontinuerlig med å utvikle diskrimineringsretten. Dagens gjeldende lovgivning er utarbeidet i takt med de internasjonale rettskildene.⁵⁴

⁵³ Se avhandlingens kapittel 4.1.2 om «harms of representation».

⁵⁴ Per dags dato er det syv menneskerettighetskonvensjoner som forbyr og regulerer tiltak mot diskriminering. Norge har inkorporert fem av dem gjennom menneskerettsloven § 2. Dette gjelder EMK, SPR ØSK, BK, og KDK. I tillegg kommer rasediskrimineringskonvensjonen og Konvensjonen mot diskriminering av personer med funksjonshemming hvor Norge har ratifisert sistnevnte.

Innenfor diskrimineringsretten, herunder både nasjonalt og internasjonalt, arbeider man med et skille mellom at diskriminering som er forbudt, og bestemmelser som er utformet som plikter for subjektene i diskrimineringsretten. Dette omtales ofte som diskrimineringsvernets individuelle side og vernets strukturelle side.⁵⁵ Vernets strukturelle side utgjør positive plikter for staten, og er forankret i menneskerettighetskonvensjonene ØSK, SP, BK og EMK.

Vernets individuelle side er også omtalt som den negative siden. Den individuelle siden finner man i de internasjonale konvensjonene, men gjør seg spesielt gjeldende i nasjonal lovgivning hvor forbudene er klart linket opp til diskrimineringsgrunnlag. Dette vil bli behandlet i diskrimineringsvernet etter likestilling- og diskrimineringsloven.

Norge har flere bestemmelser i nasjonal lovgivning som pålegger staten eller ansvarlige plikter for å fremme likestilling. Noen av disse bestemmelsene vil bli gjennomgått i kommende underkapitler.

5.3 Grunnlovens diskrimineringsvern

Grunnlovens § 98 oppstiller et vern for likebehandling og mot diskriminering. Bestemmelsen ble vedtatt etter grunnlovsrevideringen i 2014, og før denne tid inneholdt ikke Grunnloven noen generell bestemmelse som omhandlet likhet for loven og ikke-diskriminering.⁵⁶

Bestemmelsen lyder som følger:

«Alle er like for loven.

Intet menneske må utsettes for usaklig eller uforholdsmessig forskjellsbehandling.»

Ordlyden tilsier at det ikke skal foretas noen favorisering av borgerne, og at lik behandling er normen. Etter første ledd følger det et likhetsprinsipp for mennesker i Norge.

Menneskerettighetsutvalget la blant annet vekt på at likhetsprinsippet i Grunnloven skulle synliggjøres og fremheves. Dette er prinsipper rettssystemet vårt er bygget på, og lovgiver har

⁵⁵ Likestillings- og diskrimineringsrett, Midlertid kompedieutgave for studenter av Anne Hellum og Vibeke Blaker Strand 2021, s. 62

⁵⁶ Grunnloven har imidlertid inneholdt bestemmelser som har signalisert tanker om likhet som de tidligere bestemmelsene i §§ 95, 103 og 108.

ikke «adgang til å gi privilegier til enkelte personer eller grupper i samfunnet».⁵⁷ Utvalget bemerket også at likhetsprinsippet er et prinsipp som er forankret i politikere, jurister og i store deler av samfunnet allerede, og dermed ikke endrer rettstilstander på noe særlig.⁵⁸ Selv om tidligere versjoner av Grunnloven ikke inneholdt likhetsprinsipp eller et diskrimineringsforbud, anser man det som en grunnleggende sedvane i vårt rettssystem.

Likhetsprinsippet er i tillegg sterk tilknyttet prinsippet om ikke-diskriminering etter andre ledd. Menneskerettighetsutvalget bemerket at likhetsprinsippet vil fungere som et tolkningsprinsipp for domstolene, og at diskrimineringsforbudet vil kunne være gjenstand for domstolsprøving.⁵⁹

Etter annet ledd skal ingen bli utsatt for «usaklig» eller «uforholdsmessig» forskjellsbehandling. Vurderingen om forskjellsbehandlingen er «usaklig» eller «uforholdsmessig» er en sammenligningssituasjon av de som er negativt behandlet og de som ikke er det. En naturlig språklig forståelse av ordlyden tilsier at det kun er den negative forskjellsbehandlingen som er ilagt et grunnlovsværn.

Subjektet som ikke skal forskjellsbehandles er «menneske». Dette vil tilsi at bestemmelsen ikke favoriserer noen spesielle grupper i samfunnet eller ønsker å verne noen sterkere enn andre. Vernet avgrenses mot selskaper og juridiske personer.⁶⁰

Ved å være kjønnsnøytral vil bestemmelsen være dynamisk slik at den alltid vil verne om ethvert menneske i staten og følge rettstilstanden som ligger til grunn. Hvor kvinner tradisjonelt har vært det kjønn som har blitt forskjellsbehandlet mest gjennom tiden, så kan dette endre seg i fremtiden. Uansett hvilket kjønn som blir undertrykket så vil kvinner, menn og ikke-binære være beskyttet av bestemmelsen siden alle faller inn under ordlyden «menneske».

⁵⁷ Dok. 16 (2011–2012) kapittel 26.6.1.1 s. 144.

⁵⁸ Dok. 16 (2011–2012) kapittel 26.6.1.1 s. 144.

⁵⁹ Dok. 16 (2011–2012) kapittel 26.6.1.1 s. 144.

⁶⁰ Dok. 16 (2011–2012) kapittel 26.6.2.8 s. 152.

5.4 Aktivitetsplikten for offentlige myndigheter

Aktivitetsplikten følger av likestillings- og diskrimineringsloven kapittel 4, og pålegger offentlige og private aktører til å aktivt, målrettet og planmessig arbeide for likestillingsarbeid.⁶¹ Arbeidet for å hindre diskriminering og fremme arbeidet med likestilling er et arbeid som enklere kan nås ved å pålegge en sterkere part, som offentlig myndighet, plikter som de må følge.⁶² Herunder faller aktivitetsplikten. Dette vil si at bedriftene ikke kan fremstå som at de fremmer likestilling, men de må utføre handlinger, endre praksis eller på annet vis foreta seg noe som viser at de arbeider aktivt for likestilling.⁶³

Den generelle aktiviseringsplikten for offentlige myndigheter følger av ldl. § 24 første ledd. «Offentlige myndigheter» skal «i all sin virksomhet» jobbe «aktivt» for å fremme likestilling. Med «offentlig myndighet» må det forstås «statlige, kommunale og fylkeskommunale forvaltningsorganer som utøver offentlig myndighet».⁶⁴ Med «all sin virksomhet» må det forstås at det gjelder alle ledd av arbeidet som utføres. Forarbeidene legger til grunn at aktivitetsplikten «omfatter således all myndighetsutøvelse og all tjenesteyting».⁶⁵

En naturlig språklig forståelse av «aktivt» tilsier at dette er et kontinuerlig arbeid som hele tiden skal være et arbeidskrav for den ansvarlige. Etter forarbeidene menes «aktivt» at offentlige myndigheter både skal starte tiltak som kan forhindre diskriminering, men også fremme likestilling.⁶⁶

Videre må virksomheten jobbe «målrettet». Dette vil si at arbeidet for likestilling må være konkretisert og målbart. Dette kan for eksempel være at flere kvinner har blitt ansatt eller at flere kvinner har fått tilbud om fast stilling. Forarbeidene legger også til grunn at det skal være en person som er ansvarlig for at offentlig myndigheter når målet.⁶⁷

⁶¹ Likestillings- og diskrimineringsrett, Midlertid kompedieutgave for studenter av Anne Hellum og Vibeke Blaker Strand 2021, s. 311.

⁶² Prop.81 L (2016–2017) punkt 24.2.9 s. 231.

⁶³ Likestillings- og diskrimineringsrett, Midlertid kompedieutgave for studenter av Anne Hellum og Vibeke Blaker Strand 2021, s. 312.

⁶⁴ Prop.88 L (2012–2013) kapittel 20.1 s. 162.

⁶⁵ Prop.63 L (2018–2019) kapittel 3.10.1.5 s. 93.

⁶⁶ Prop.81 L (2016–2017) punkt 24.2.9 s. 231.

⁶⁷ Prop.81 L (2016–2017) punkt 24.2.9 s. 231.

Til sist må offentlige myndigheter arbeide «planmessig» med likestillingsarbeidet. Dette betyr at offentlig myndighet må være organisert og ryddig i sitt arbeid. Forarbeidene viser til at dette kan gjøres ved å utarbeide og følge en strategi for likestillingsarbeidet.⁶⁸ Det gis ingen føringer for hvordan strategi det skal anvendes, men at dette må gjøres på bakgrunn av hvilke situasjoner og utfordringer hvert enkelt organ i den offentlige myndighet har.⁶⁹

Etter forarbeidene understreker departementet at aktivitetsplikten omfatter aktive tiltak for likestilling på alle samfunnsområder, og dette «vil innebære å avskaffe alle former for diskriminering i lover, forskrifter, rundskriv, vedtak mv.».⁷⁰

Herunder må det forsås at aktivitetsplikten vil være der samfunnet befinner seg, enten det er digitalt eller fysisk. Offentlige myndigheters ansvar for likestillingsarbeidet vil bety at de skal jobbe for å hindre at noen blir ulovlig diskriminert. Dette må også gjelde i de tilfellene der myndighetene selv utfører kontroll av borgerne.

5.5 Diskrimineringsvernet etter likestillings- og diskrimineringsloven

Diskrimineringsgrunnlagene følger av ldl. § 6, og ordlyden uttrykker at ethvert forhold som er et resultat av direkte og indirekte forskjellsbehandling er forbudt.

Forbudet mot direkte og indirekte diskriminering omtales gjerne som diskrimineringsvernets individuelle side, hvor vernet må anvendes i individuelle saker. For at en kan konstatere at det foreligger ulovlig forskjellsbehandling må vilkårene være oppfylt.

For det første må det enten foreligge direkte eller indirekte forskjellsbehandling etter ldl. §§ 7 eller 8. For det andre må denne forskjellsbehandlingen ha et bestemt diskrimineringsgrunnlag som følger av ldl. § 6. For det tredje kan ikke forskjellsbehandlingen være saklig begrunnet og dermed lovlig etter ldl. §§ 9, 10 eller 11.

⁶⁸ Prop.81 L (2016–2017) punkt 24.2.9 s. 231.

⁶⁹ Prop.81 L (2016–2017) punkt 24.2.9 s. 231.

⁷⁰ Prop.81 L (2016–2017) punkt 24.2.9 s. 231.

5.5.1 Direkte forskjellsbehandling

Direkte forskjellsbehandling etter § 7 omhandler de tilfellene hvor en person «behandles dårligere» enn en annen person som befinner seg på det samme samfunnsområdet, og som har en direkte sammenheng med et vernet diskrimineringsgrunnlag etter ldl. § 6.

Vilkåret etter § 7 er at noen skal bli «behandlet dårligere» enn andre. At noen blir behandlet «dårlig» vil si at de blir behandlet på en negativ måte, og ubetydelige ulemper ikke faller inn under bestemmelsen.⁷¹ Behandling er både handlinger og unnlater. Med «handling» menes «alle former for praksis, betingelser, vedtak, bestemmelser, avtaler osv.».⁷²

Bestemmelsen fastslår ikke noen tidsbegrensning for når den negative handlingen eller unnlatsen skal ha funnet sted.

Forskjellsbehandling etter § 7 må føre til skade eller ulempe for den som forskjellsbehandles. Forskjellsbehandlingen må eksempelvis føre til tap av fordeler, økonomisk tap eller færre muligheter sammenlignet med andre i en tilsvarende situasjon.⁷³ Det er ikke et krav om at samtlige personer innenfor et diskrimineringsgrunnlag mener det er forskjellsbehandling eller at alle må bli forskjellsbehandlet før det kan anses som direkte forskjellsbehandling.⁷⁴

Det trenger ikke ha vært en hensikt til å diskriminere. Selv om vedkommende ikke har ment å diskriminere den utsatte personen, kan det likevel bli ansett som forskjellsbehandling. Etter forarbeidene omtales dette som at «det foreligger et diskriminerende element i selve handlingen».⁷⁵ Dette følges også opp av Høyesterett. I Rt. 2012 s. 424 uttalte Høyesterett at diskrimineringsforbudet «ikke er begrenset til tilfeller der det foreligger diskrimineringshensikt».⁷⁶ Den utilsiktete eller ubevisste diskriminering rammes også,

⁷¹ Prop.81 L (2016–2017) s. 314.

⁷² Prop.88 L (2012–2013) kapittel 2.1.2 s. 16.

⁷³ Prop.81 L (2016–2017) s. 314, jf. også pkt. 12.2.4 s. 101-103.

⁷⁴ Lovkommentar til Likestillings- og diskrimineringsloven, kommentarutgave av Anne Jorun Bolken Ballangrud og Margrethe Søbstad, punkt 7.2.2.

⁷⁵ Prop.81 L (2016–2017) kapittel 12.2.2.1 s. 99 med videre henvisning til NOU 2009: 14 kapittel 3.2.5 s. 42.

⁷⁶ Rt. 2012 s. 424 A avsnitt 32.

uavhengig av skyld hos arbeidsgiver. Videre uttalte Høyesterett i Rt. 2014 s. 402 at også «ubevisst – det vil si uaktsom – diskriminering rammes».⁷⁷

5.5.2 Indirekte forskjellsbehandling

Etter ldl. § 8 må en «tilsynelatende nøytral» bestemmelse, betingelse, praksis, handling eller unnlattelse «stille personer dårligere». Kravet er ikke knyttet til spesifikke personer, og det kan dermed være mulig å se hen til hypotetiske tilfeller. En fortolkning av ordet «dårligere» vil være det samme som etter ldl. § 7.

Med «tilsynelatende nøytral» må man forstå hva som ved første øyekast ser ut som en objektiv og upartisk bestemmelse, betingelse, praksis, handling eller unnlattelse. For en del personer vil ikke denne forskjellsbehandlingen slå feil ut, men kravet er at det er noen som kan bli «dårligere stilt enn andre». Det er tilstrekkelig at den dårlige behandlingen rammer «langt flere» diskrimineringsvernede personer enn andre.

Forarbeidende til § 8 omtaler indirekte forskjellsbehandling som «kjennetegnes ved at handlingen mv. er nøytral i den forstand at den eksempelvis ikke er motivert av eller direkte knyttet til forholdene nevnt i § 6 første ledd. Det kan være praksiser som er iverksatt med det formål å ivareta legitime hensyn, men som faktisk vil ramme en bestemt gruppe».⁷⁸

Videre må det bemerkes at det er «personer» som må stilles dårligere. Dette vil altså tilsi at det ikke kun er enkeltpersoner, men også individer av grupper som kan rammes. Dette blir fulgt opp av forarbeidene, hvor det uttales at en gruppetilnærming må være det klare utgangspunktet, og at man må ta en sammenligning mellom grupper i samfunnet når man vurderer om det foreligger indirekte forskjellsbehandling.⁷⁹ Ifølge lovgiver kan dette utgangspunktet fravikes om det foreligger «typiskhet» i problemstillingen. For eksempel at diskrimineringsgrunnlaget er kjønn, men at det kun er en kvinne som blir indirekte forskjellsbehandlet. Det typiske for denne situasjonen er at det ofte er kvinner på det aktuelle

⁷⁷ Rt. 2014 s. 402 A avsnitt 59.

⁷⁸ Prop.81 L (2016–2017) s. 314-315.

⁷⁹ Prop.81 L (2016–2017) punkt 12.2.5.2 s. 104.

samfunnsområdet som blir diskriminert.⁸⁰ Dette er også lagt til grunn av Høyesterett i en nyere avgjørelse, HR-2022-350-A.⁸¹

5.6 Den diskrimineringsrettslige bevisbyrderegelen

Bevisbyrdereglene følger av ldl. § 37 og første ledd lyder som følger:

«Diskriminering skal anses å ha skjedd hvis det foreligger omstendigheter som gir grunn til å tro at diskriminering har skjedd, og den ansvarlige ikke sannsynliggjør at diskriminering likevel ikke har skjedd».

En naturlig språklig forståelse av «grunn til å tro» tilsier at klageren må vise til de omstendighetene som antas å falle inn under diskrimineringsreglene. Terskelen for sannsynliggjøring er imidlertid lav som følge av at det kun skal være en mistanke om at det foreligger diskriminering.

Innenfor diskrimineringsretten må klageren eller saksøkeren selv bringe frem presumsjonen om at diskriminering har skjedd.⁸² Dette kan for eksempel være i form av hendelser som har skjedd eller dokumentasjon på antatt diskriminering. Deretter må den som er anklagd eller saksøkt motbevise at det ikke foreligger diskriminering. Dersom den innklagde/saksøkte ikke kan motbevise at det ikke foreligger diskriminering, eller i det minste vise at det er lovlig forskjellsbehandling, vil dette tale i fordel for klager/saksøker.

Bevisbyrdereglene gir muligheten for å spre bevisreglene utover, og effektivisere vernet mot diskriminering.⁸³ I vanlige rettslige saker praktiserer man med en sannsynlighetsovervekt når det kommer til bevis. Om det er mer enn 50 % sannsynlig at noe er sant så vil dette tale i favør for klager.⁸⁴ I diskrimineringsretten vil presumsjonen av at det foreligger diskriminering være nok til at det er den anklagde som må føre bevis for at det ikke foreligger diskriminering.

⁸⁰ Prop.81 L (2016–2017) punkt 12.2.5.2 s. 104.

⁸¹ HR-2022-350-A avsnitt 48.

⁸² Likestillings- og diskrimineringsrett, Midlertid kompedieutgave for studenter av Anne Hellum og Vibeke Blaker Strand 2021, s. 292.

⁸³ Ot.prp nr. 77 (2000–2001) punkt 12.2.3. s. 90 og NOU 2009:14 s. 272.

⁸⁴ Pedersen, Jussi Erik og Skoghøy, Jens Edvin A, «Grunnleggende sivilprosess», s. 288.

For det kan sies at det foreligger «grunn til å tro» at diskriminering har skjedd, må personen(e) behandles eller stilles dårligere etter ldl. §§ 7 eller 8. Deretter må forskjellsbehandlingen være knyttet til et diskrimineringsgrunnlag i ldl. § 6.

Dersom klager/saksøker ikke oppnår å etablere en presumsjon om at det foreligger diskriminering, kan ikke saken gå videre. Dette gjelder også i saker hvor klager mener seg diskriminert på bakgrunn av forhold som ikke faller inn under diskrimineringsvernet. Som for eksempel hvor en eier av dieselbil føler seg diskriminert som følge av de fordelene el-bileiere får av staten.⁸⁵

Høyesterett har lagt til grunn at det er den alminnelige bevisbyrderegelen som skal anvendes, og det mest sannsynlige beviset skal legges til grunn.⁸⁶ Videre slås det fast at bevisene som er mest tidsnære hendelsen har størst vekt, og at begge parter plikter å legge frem de typer bevis de måtte ha i en sak.⁸⁷

Videre uttalte Høyesterett at om det foreligger en sjanse for at det er en absolutt bevistvil så skal dette gå ut over «den ansvarlige».⁸⁸ Dette betyr at hvis noe er like sannsynlig som det er usannsynlig, så er det den som er anklaget for å diskriminere som må bære ansvaret for å fremlegge bevis på at forskjellsbehandling ikke har skjedd eventuelt hvorfor denne forskjellsbehandlingen i så fall var lovlig. Høyesterett bemerker at slike tilfeller er sjelden i praksis, men dette utelukker ikke problemstillingen.⁸⁹

6 Saker som har anvendt maskinlæring for kontroll

6.1 Innledning

Sakene under har anvendt kunstig intelligens, herunder maskinlæring, for å utføre kontroll av borgere som har mottatt et samfunnsmessig gode. Bakgrunnen for å bruke maskinlæring er at man ønsker å effektivisere en prosess som vanligvis er tids- og ressurskrevende. Slik bruk av

⁸⁵ DIN-2021-262.

⁸⁶ HR-2020-2476-A avsnitt 74.

⁸⁷ HR-2020-2476-A avsnitt 74.

⁸⁸ HR-2020-2476-A avsnitt 75.

⁸⁹ HR-2020-2476-A avsnitt 75.

maskinlæring er klassiske eksempler på hvordan man tenker at man skal ta i bruk datasystemer for å optimalisere velferdsstaten.

Hensikten med å vise til disse tre eksemplene er at dette er reelle situasjoner hvor man kan ta utgangspunkter for å diskutere bruken av diskrimineringsrettslige begrep og vilkår.

6.2 Bokkontroll til Lånekassen i samarbeid med Riksrevisjonen

Elever eller studenter som søker lån hos Statens lånekasse for utdanning (Lånekassen) kan få omgjort deler av sitt lån til stipend så lenge de er borteboere. Med en borteboer betyr det at man ikke bor i sine foreldres hus sammen med dem. Studenter som har fått innvilget stipend godkjenner at de er borteboere tillater at Lånekassen kan foreta en kontroll av dem.⁹⁰

Lånekassen utfører titusenvis av kontroller hvert år. Tidligere har dette blitt gjort av saksbehandlere, men Lånekassen har vært en av de første offentlige instansene som har tatt i bruk maskinlæring på et så stort nivå som bokkontrollen er. I 2008, 2011 og 2012 var det 1000 studenter som ble tatt ut i kontroll, og disse ble vurdert av saksbehandlere. Fra og med 2015 ble det gjennomført en fullskalakontroll hvor 48 000 studenter ble kontrollert for det forgående året. Kunnskapsdepartementet anslo at 72 millioner kroner kunne ha blitt omgjort til lån hvis man foretok en kontroll av elevene og studentene.⁹¹

Riksrevisjonen har i samarbeid med Lånekassen foretatt en undersøkelse av hvorvidt Lånekassen kan bruke maskinlæring i sin tjeneste. Formålet med undersøkelsen «har vært å teste ut hvilke typer problemstillinger som kan være fornuftige for fremtidige «revisjoner av algoritmer» mht. metodikk, tilnærming fokus og avgrensning».⁹² Det omtales også at dette ikke er en vanlig revisjon hvor det blir foretatt en kontroll av forvaltningen, men et utviklingsprosjekt.⁹³

⁹⁰ FOR-2019-03-15-254, Forskrift av tildeling av utdanningsstøtte for undervisningsåret 2019-2020 § 12-2.

⁹¹ Bokkontroll basert på maskinlæring, En vurdering ved Riksrevisjonen i samarbeid med Statens lånekasse for utdanning punkt 2.1 s. 7.

⁹² Bokkontroll basert på maskinlæring, punkt 1.1 s. 4.

⁹³ Bokkontroll basert på maskinlæring, punkt 1.2 s. 5, fotnote 4.

Modellen til Lånekassen som det er tatt utgangspunkt i er fra 2018, og har de siste årene blitt utviklet videre. Det er imidlertid tatt utgangspunkt i 2018-modellen i vurderingen.⁹⁴

Datagrunnlaget ble bygget på bakgrunn av fullskalakontrollene fra studentene og elevene som hadde mottatt støtte i 2014.

Poenget med å bruke maskinlæringsmodellen var at Lånekassen skulle ha mulighet til å kontrollere alle søkere, men på bakgrunn av predikasjoner gjort av maskinen så kunne de raskere finne misligholdere ved hjelp av færre ressurser.⁹⁵ Det er rundt 60 000 personer som mottar støtte fra Lånekassen. 10 000 av disse ble tatt ut i tilfeldig kontroll av boplass, og deretter ble 15 000 personer fra det resterende antallet plukket ut på bakgrunn av høyest predikert sannsynlighet for å ha oppgitt feil bostatus.⁹⁶

Modellen til Lånekassen viser at det er høyere sannsynlighet for å bli plukket ut i bokkontroll om du er mann. En av grunnene for at menn blir plukket ut oftere er at de historiske dataene viser at det er oftere menn som oppgir feil bostatus enn det kvinner gjør. De historiske dataene viser at det er 4 % kvinner som oppgir feil bostatus, mot 7 % av mennene. Det er imidlertid 2,27 ganger så stor sannsynlighet for at menn blir trukket ut i bokkontroll enn kvinner.⁹⁷

6.3 Veneziakommisjonens rapport om AWIR-saken fra Nederland

I 2005 tok Nederland i bruk trygdeprogrammet «General Act on Income-Related Schemes or General Act on Means-Tested Scheme (the AWIR)».⁹⁸ Programmet gikk ut på at foreldre av barn som var under skolealder kunne få utbetalt en godtgjørelse hvis de måtte ta i bruk barnehager eller dagmammaer/barnepassere for å kunne være i jobb. Godtgjørelsen ble regnet ut på bakgrunn av foreldrenes lønn og skulle dekke deler av kostnaden.⁹⁹ Forskjellig fra

⁹⁴ Bokkontroll basert på maskinlæring, punkt 3.1 s. 9.

⁹⁵ Bokkontroll basert på maskinlæring, punkt 4 s. 14.

⁹⁶ Bokkontroll basert på maskinlæring, punkt 4.1 s. 15.

⁹⁷ Bokkontroll basert på maskinlæring, punkt 5.2 s. 19 og 20.

⁹⁸ The Netherlands opinion on the legal protection of citizens, CDL-AD (2021)031 s. 4, punkt 7.

Oversatt fra nederlandsk av Veneziakommisjonen. I Nederland ble dette omtalt som Harmonisatie van inkomensafhankelijke regelingen (Algemene wet inkomensafhankelijke regelingen).

⁹⁹ CDL-AD (2021)031 s. 4, punkt 7.

vanlig barnetrygd, skulle stønaden være behovsprøv. Man kunne få støtte fra staten så lenge personene kunne bevise at de hadde et behov.

Fra og med 2011 brukte Nederland programmet maskinlæring for å kunne sjekke innbyggerne som hadde mottatt støtte. Den selvlærende algoritmen hadde mottatt data som skulle vise korrekte og ukorrekte søknader som resulterte i at noen innbyggere ble ansett som en høy risiko for å svindle til seg støtte de ikke hadde rett på. En av parameterne som indikerte om man var en høy risiko var statsborgerskap, som kan vise at man er en annen etnisitet. Hvis borgeren var utenlandsk ga systemet beskjed om at det måtte foretas en omstendelig undersøkelse av søkeren.¹⁰⁰ Med andre ord ble man ansett som høyrisiko om man ikke var nederlandsk fra før.

I 2013 oppdaget staten en storstilt kriminell gruppe som utnyttet ordningen. Ringlederen var i et annet land, men sendte andre individer til Nederland for å søke om trygdeytelser selv om de ikke oppfylte vilkårene for ytelsen. Etter individene hadde mottatt trygdeytelsene tok de ut pengene for så å dra tilbake til sitt hjemland. AWIR-systemet var lagt opp til at man kunne søke om støtte på forhånd og dermed få utbetalt stønaden før det gjorde seg gjeldende slik at fremtidige utgifter for barnepass kunne bli betalt med en gang.¹⁰¹

For å få bukt med utnyttelsen av ordningen skulle det innføres kontroller av de som hadde mottatt ytelse. Frem til januar 2021 hadde AWIR-programmet en klausul om at hvis det blir foretatt en revidering av godtgjørelsen eller en revisjon av forskuddsbetalingen, så må søkeren betale tilbake hele beløpet.¹⁰² Dette medførte at de søkerne som ikke hadde tilstrekkelig dokumentasjon på at barnepass måtte tilbakebetale stønaden de hadde fått. Foreldrene måtte tilbakebetale hele beløpet selv om det var en minimal forskjell på det de hadde fått utbetalt og kostnaden av barnehagen/barnepasseren. For noen familier utgjorde dette rundt 30 000 euro.¹⁰³

AWIR-systemet var høyst sentralisert, og de lokale trygdekontorene kunne ikke gi et godt nok svar på hvorfor hele beløpet måtte tilbakebetales. Via et internnotat fra 2013 ble det gitt

¹⁰⁰ CDL-AD (2021)031 s. 4, punkt 11.

¹⁰¹ CDL-AD (2021)031 s. 4, punkt 12.

¹⁰² CDL-AD (2021)031 s. 4, punkt 13.

¹⁰³ CDL-AD (2021)031 s. 4, punkt 15.

instruksjoner om at ingen assistanse eller forklaring skulle bli gitt til foreldrene som ble rammet.¹⁰⁴

Foruten om kritikk angående dårlig deling av informasjon og oppklaring av praksis mellom forvaltningen og staten retter kommisjonen også kritikk mot faren for diskriminering gjort av kunstig intelligens.¹⁰⁵ Kommisjonen legger blant annet vekt på at uveiledet maskinlæring vil utgjøre nærmest en umulig oppgave i å kunne oppdage *bias*.¹⁰⁶ Om man i så tilfelle finner ut at maskinlæringssystemet er forutinntatt må systemet læres opp på nytt til at de parameterne som diskriminerer vil bli endret.¹⁰⁷

I denne saken ble det altså konkret rettet kritikk mot at maskinlæringssystemet gjorde en fare for diskriminering. De maskinlæringssystemene som er uveiledet vil særlig utgjøre en fare siden det ikke vil være et menneske i loopen som kan kontrollere utfallet av algoritmen. Det vises også imidlertid til at maskinlæringen er basert på historisk data hvor mennesker tidligere har diskriminert.¹⁰⁸ Dette kan bety at når man først får det på svart på hvitt at noen parametere utgjør en større sjanse for forskjellsbehandling, så kan det være mulig å oppdage tidligere uoppdagede diskrimineringer. Dette vil imidlertid bero på at man har et maskinlæringssystem som gir en forklaring på resultatet.

6.4 SyRI-saken fra Nederland

For å avdekke borgere som hadde en høy risiko for å begå trygdesvindler introduserte Nederlands regjering et system ved navn System Risico Indicate (SyRI). Ved hjelp av algoritmer skulle SyRI finne de individene som ville utnytte velferdssystemet, og det var sagt at systemet ville gi et effektivt og grundig kontrollprogram for myndighetene.¹⁰⁹ Algoritmene til SyRI skulle behandle datagrunnlag som staten allerede hadde fått tilgang på via offentlige instanser som hadde avtalt om å samarbeide og utveksle informasjon med hverandre. Dette gjaldt blant annet den nederlandske skatteetaten, arbeids- og sosialdepartementet, politiet,

¹⁰⁴ CDL-AD (2021)031 s. 4, punkt 17.

¹⁰⁵ CDL-AD (2021)031 s. 19, punkt E.

¹⁰⁶ Se avhandlingens kapittel 4 vedrørende forutinntatt kunstig intelligens.

¹⁰⁷ CDL-AD (2021)031 s. 20, punkt 97 og 98.

¹⁰⁸ CDL-AD (2021)031 s. 20, punkt 97

¹⁰⁹ ECLI:NL: RBDHA:2020:1878 punkt 3.2. Engelsk oversettelse av SyRI-dommen.

påtalemyndighetene og andre offentlige instanser som er en del av velferdssystemet i Nederland.

De overnevnte instansene besitter personlig data som arbeid, informasjon om tidligere gitte sanksjoneringer fra staten, skatteinformasjon, informasjon om eierskap av eiendom og ting, offentlige ytelser mottatt eller nektet, navn, adresse, kjønn, alder, postadresse, helsedata, utdanning, gjeldsoversikt og historisk oversikt av overholdelse av lover og regler.¹¹⁰ Alt dette datagrunnlaget skulle programmet SyRI ta i bruk for å kunne identifisere individer med forhøyet risiko for misbruk av ordningen. Ut fra denne listen av personlig informasjon kan man gjerne si at man har hele livshistorien til et menneske og deres kontakt med velferdssamfunnet.

Myndighetene i Nederland la blant annet vekt på nødvendigheten av kontroll i samfunnet da de la frem forklaringen på hvordan SyRI fungerer:

«SyRI includes the technical infrastructure and associated procedures through which data can be linked and thereafter anonymously analyzed in a secure environment, in order to generate risk notifications. A risk notification means that a legal person or natural person is considered worth investigating in connection with possible fraud, unlawful use of and noncompliance with legislation. This method leads to a more effective and efficient use of the control instrument».¹¹¹

I høringsnotatet har myndighetene lagt vekt på at det er «anonymiserte data» som blir brukt. Det kommer imidlertid frem av dommen at datagrunnlaget ikke er anonymisert, men pseudonymisert.¹¹² Det vil si at personopplysninger ble aidentifisert slik at de ikke automatisk kan identifiseres med en person eller et selskap, ved for eksempel å erstatte navn eller personnummer med en kode. Etter personvernforordningen regnes ikke pseudonymisering som anonymisering da man kan finne tilbake til personen med den

¹¹⁰ SyRI-dommen, punkt 4.17 i dommen.

¹¹¹ Oversettelse av høringsnotat til den nederlandske Wet SUWI:3. Oversatt av forfatterne av «Digital welfare fraud detection and the Dutch SyRI judgement» av Marvin van Bekkum og Frederik Zuiderveen Borgesius. European Journal of Social Security, 2021, Vol 23 (4) s. 325.

¹¹² SyRI-dommen, punkt 4.28 og 4.29.

kodifiserte informasjonen.¹¹³ Dette vil tilsi at datagrunnlaget ikke var randomisert, og at parameterne som ble lagt vekt på var utslagsgivende for å bekrefte om en person var en høyere risiko.

Domstolen bemerker at trygd er en bærebjelkene i det nederlandske samfunnet, og det å kjempe mot trygdesvindler er viktig, og at det gir mening at staten prøver ny teknologi for å bekjempe svindel.¹¹⁴ Retten kom imidlertid frem til at SyRI var i strid med Den europeiske menneskerettskonvensjon artikkel 8.¹¹⁵ Artikkel 8 går kort ut på at enhver har rett på privatliv, og for at inngrep i denne rettigheten skal være lovlig så må den være i samsvar med loven, nødvendig i et demokratisk samfunn og ha et legitimt formål.

Retten slår tidlig fast at SyRI hadde et legitimt formål, og uttaler:

«The court is of the opinion that the SyRI legislation is in the interest of economic wellbeing and thereby serves a legitimate purpose as adequate verification as regards the accuracy and completeness of data based on which citizens are awarded entitlements is vitally important».¹¹⁶

For å kunne se om SyRI er nødvendig i et demokratisk samfunn la retten vekt på den manglende muligheten for innsyn for borgerne og prinsippet om åpenhet.¹¹⁷ Lovgiveren kunne ikke forklare eller gi informasjon om hvordan datagrunnlaget hadde kommet til konklusjonen om hvorfor det forelå en høyere risiko for noen borgere enn andre. Dette knytter seg til svart boks-problematikk.

Risikoen for diskriminering, eller om noen hadde blitt diskriminert, gikk ikke domstolen videre inn på. Her ville det imidlertid vært anledning for domstolen til å peke på at forskjellsbehandling kan forekomme når man bruker maskinlæring for kontroll.

¹¹³ Personopplysningsloven § 1, jf. personvernforordningen artikkel 4 nummer 5.

¹¹⁴ SyRI-dommen, punkt 6.3.

¹¹⁵ Den europeiske menneskerettskonvensjon fra 3. september 1953, artikkel 8.

¹¹⁶ SyRI-dommen, punkt 6.4.

¹¹⁷ SyRI-dommen, punkt 6.90, 6.86 og 6.87

For eksempel ble SyRI kun brukt i distrikt eller soner som ble betegnet som «problemdistrikt».¹¹⁸ Det vil si områder som opplever et høyere antall kriminalitet, trygdesvindler eller annen uønsket oppførsel som et velferdssamfunn ønsker seg. SyRI-systemet var en bidragsyter til at grupper med en lavere sosioøkonomiske status, eller de med innvandrerbakgrunn som hadde adresse i problemdistrikt, ble ansett som en høyere risiko for trygdesvindler enn andre i samfunnet.

7 Diskrimineringsvernets anvendelse på overnevnte saker

7.1 Innledning

Det individuelle diskrimineringsvernet er vanskelig å anvende på saker der forskjellsbehandling er forårsaket av maskinlæringsprogrammer. Dette kan det være flere grunner for, og dermed kan det være behov for mer oppmerksomhet rundt de strukturelle pliktene for de som tar i bruk systemene.¹¹⁹

For det første kan det være vanskelig å anvende diskrimineringsvernet i de sakene hvor man er usikker på om det foreligger forskjellsbehandling på grunn av liten eller ingen innsikt eller forklaring av algoritmen som følge av svart boks-problematikken. For det andre er det et spørsmål om når man kan si at det foreligger forskjellsbehandling, når den involverte part ikke vet hva som er grunnlaget eller begrunnelsen for å bli plukket ut i kontroll. Hvordan kan man da vurdere om systemet er «nøytralt» eller direkte forskjellsbehandler?

7.2 Foreligger det forskjellsbehandling?

Som behandlet i avhandlingen punkt 5.5 har vi to typer forskjellsbehandling i Norge som er vernet av diskrimineringslovgivningen, direkte og indirekte forskjellsbehandling. Hvis forhold havner inn under disse bestemmelsene må det foretas en helhetsvurdering om det foreligger ulovlig forskjellsbehandling som enten er direkte eller indirekte.¹²⁰

¹¹⁸ SyRI-dommen, punkt 6.93.

¹¹⁹ Se avhandlingens kapittel 4.1-4.3.

¹²⁰ Se avhandlingens kapittel 5.5.

Det er ikke alltid like enkelt å foreta et klart skille mellom direkte og indirekte forskjellsbehandling.¹²¹ I noen tilfeller vil personer kunne bli direkte forskjellsbehandlet, men de samtidig kan bli indirekte forskjellsbehandlet på grunn av andre forhold i saken. For eksempel tilfeller som både omfavner graviditet og kjønn.

Spørsmålet her vil være om behandlingen som har blitt gjort i de overnevnte sakene kan anses som forskjellsbehandling etter nasjonal rett. Forskjellsbehandlingen må resultere i at en person blir «behandlet dårligere» eller «stilt dårligere» enn «andre personer».¹²²

En naturlig språklig forståelse av ordet forskjellsbehandling er at man blir satt i en mindre gunstig posisjon enn andre i tilsvarende situasjon. Det er ikke nødvendig at det er en faktisk person, det kan også dreie seg om lignende eller hypotetiske tilfeller. Hvis du blir forskjellsbehandlet til fordel av andre vil du ikke behandles eller stilles dårligere, og dermed er det kun snakk om negative virkninger.

Som hovedregel skal det foretas en objektiv vurdering om en handling eller en unnlattelse er negativ forskjellsbehandling eller ikke.¹²³ Om det ikke er mulig å ta utgangspunkt i den objektive vurderingen så den diskriminerende part sin subjektive forståelse av hendelsene vurderes.¹²⁴

Diskrimineringsvernet sett i sammenheng med den første saken om Lånekassen, er det naturlig å knytte dette opp til diskrimineringsgrunnlaget «kjønn».¹²⁵ Med kjønn menes det biologiske kjønn man eller kvinne som man blir tilegnet ved fødselen som følge av at man har XX eller XY kromosom eller hvis man har tatt kjønnskorrigerende operasjon for å endre sitt kjønn.¹²⁶

¹²¹ Ligestillings- og diskrimineringsrett, Midlertid kompedieutgave for studenter av Anne Hellum og Vibeke Blaker Strand 2021, s. 181.

¹²² Ldl. §§ 7 og 8. Se avhandlingens kapittel 5.5.1 og 5.5.2 for en innføring i direkte og indirekte diskriminering.

¹²³ Lovkommentar til Ligestillings- og diskrimineringsloven, kommentarutgave av Anne Jorun Bolken Ballangrud og Margrethe Søbstad, punkt 7.6.4.

¹²⁴ Prop.81 L (2016–2017) punkt 12.2.4.2 s. 102.

¹²⁵ Se ldl. § 6 første ledd.

¹²⁶ Prop.81 L (2016–2017) punkt 11. 2.2.1 s. 81.

Kriteriene som må vurderes er om mennene «stilles eller behandles dårligere» enn andre på grunn av at de er menn, som er diskrimineringsgrunnlaget. Forskjellsbehandlingen må resultere i at mennene stilles eller behandles dårligere enn kvinner, og det må være en årsakssammenheng mellom behandlingen og situasjonen. Det er helt klart at parameteren «mann» har et høyere utslagsgivende effekt til å bli kontrollert enn «kvinne». På bakgrunn av rapporten er det hele 2,27 ganger mer sannsynlig at menn blir plukket ut enn kvinner til å være sannsynlige misligholdere.¹²⁷ Dette omtales som «false positive rate». Modellen gir utslag for at det er flere misligholdere enn det egentlig er på grunn av parameterne som er lagt inn.

Kvinner oppga feil bostatus 4 % av gangene kontra mennene sine 7 %. Dette vil tilsi at hvis programmet kun skulle valgt ut på bakgrunn av den reelle andelen basert på historisk treningsdata, ville menn uansett blitt plukket oftere ut som sannsynlig misligholder. Det er imidlertid graden av false positive rate som utgjør at menn blir «stilt dårligere» enn kvinner. Hvis modellen skulle likebehandlet gruppene så skulle det kun vært 3 % flere menn som ble plukket ut enn kvinner.

I den andre saken om rapporten fra Venezia-kommisjonen henvises det til at AWIR-systemet tok i bruk svært mange parametere for å kunne klassifisere personer som høy risiko eller ikke. Det er imidlertid kun parameteren «nasjonalitet» som har blitt nevnt konkret til å bli vurdert til å være en høy risiko. Områder med høy kriminalitet ga også systemet grunn til å tro at man var en høyere risiko, men dette er ikke beskyttet av et diskrimineringsgrunnlag. Hvis man imidlertid er av annen nasjonalitet og i tillegg bodde i et problemområde, ble man ansett som en enda høyere risiko for trygdesvindler.

I den tredje saken, SyRI, var det også personer med annen etnisk opprinnelse som ble valgt ut av datasystemet, og antatt at de har en høyere risiko for trygdesvindler. I tillegg ble områder med høy kriminalitet ansett om problemområder som medførte høyere risiko. Akkurat som i AWIR-saken så var det en stor andel personer med annen etnisitet som bodde i disse områdene.

¹²⁷ Bokkontroll basert på maskinlæring, En vurdering ved Riksrevisjonen i samarbeid med Statens lånekasse for utdanning punkt 5.2 s. 19.

Spørsmålet er om personene i saken ble «stilt eller behandlet dårligere» enn andre som følge av diskrimineringsgrunnlaget «etnisitet». Etter ldl. § 6 første ledd siste punktum omfattes nasjonal opprinnelse inn under etnisitet. Nasjonal opprinnelse vil omfatte hvor man har blitt født eller hvor man har hatt sin oppvekst eller sin bakgrunn fra.¹²⁸ Nasjonal opprinnelse omfatter ikke nasjonalitet eller statsborgerskap, men forskjellsbehandling etter statsborgerskap i seg selv kan likevel resultere i diskriminering og dermed være forbudt etter likestillings- og diskrimineringsloven.¹²⁹

Sakene fra Nederland gir ingen indikasjon på hvor mye større grad de med annen etnisitet ble valgt ut i motsetning til nederlandske borgere. Det er imidlertid bemerket at det er parameteren «etnisitet» som utgjorde at man var en høyere risiko for kontroll. I begge sakene fra Nederland viser det at det er allerede utsatte grupper i samfunnet som blir hardt kontrollert og forskjellsbehandlet.

Sett samlet, på bakgrunn av de overnevnte sakene, er det grunnlag for å si at det skjer forskjellsbehandling på grunn av forskjellige parametere som maskinlæringen legger vekt på. I tillegg er både «kjønn» og «etnisitet» vernede diskrimineringsgrunnlag.

En kan slutte at det foreligger forskjellsbehandling på grunnlag av vernede diskrimineringsgrunnlag.

7.3 «Noe mer enn en ubetydelig ulempe»

Vilkåret for at det foreligger ulovlig forskjellsbehandling er om noen behandles eller stilles dårligere enn andre. Forarbeidene har som nevnt lagt vekt på at dette er noe mer enn en ubetydelig ulempe.¹³⁰ Hvis ulempen ikke er tilstrekkelig stor vil dette medføre at det ikke er forskjellsbehandling. Herunder kan en se hen til problematikken «harms of allocations», hvor personer for eksempel blir fratatt muligheter eller ressurser.¹³¹

¹²⁸ Prop.81 L (2016–2017) punkt 11. 2.3.3 s. 82 som viser til Ot.prp.nr.33 (2004–2005) punkt 10.1.8.3 s. 88-89.

¹²⁹ Prop.81 L (2016–2017) punkt 11. 2.3.3 s. 82 og Ot.prp.nr.33 (2004–2005) punkt 10.1.8.3 s. 88-89.

¹³⁰ Se avhandlingens kapittel 5.5 for redegjørelse av direkte og indirekte diskriminering

¹³¹ Se avhandlingens punkt 4.1.1.

Det er nødvendig å se på om forskjellsbehandlingen er en ulempe, og en ulempe i den grad av at den ikke er ubetydelig. Med ubetydelig må det forstås at den er alminnelig i den forstand at det ikke kan gjøres gjeldende.

For å anse noe som en ulempe kan man enten se på det med et smalt blikk, hvor man ser på ulempen personen eller gruppen blir påført i den enkelte sak. Ulempen kan ellers vurderes i et bredere perspektiv, hvor man ser på ulempen i en helhetlig samfunnskontekst.

Det å se på ulempen i en samfunnskontekst vil være i tråd med likestillings- og diskrimineringslovens formålsbestemmelse i § 1, hvor etter andre ledd skal likestilling bety likeverd, like muligheter og like rettigheter. Etter forarbeidene presiseres det at likeverd innebærer «at enhver skal bli respektert for den man er, uavhengig av biologiske, sosiale og kulturelle forskjeller».¹³²

I hver enkelt sak vil dette stille seg forskjellig fra person til person. Ett eksempel er saken til Lånekassen. Ulempen i denne saken vil knytte seg til at du må innhente bekreftelse på din bosituasjon og sende dette inn til Lånekassen. Det er mye mulig at ikke alle har like god kontroll på sine personlige anliggende saker, hvor for eksempel de ikke har husleiekontrakt eller annen dokumentasjon på at de bor utfor barndomshjemmet.¹³³

Dette kan i verste fall medføre at noen studenter/elever ikke kan dokumentere at de bor utenfor barndomshjemmet. Videre vil dette medføre at de enten ikke mottar mer studiestøtte eller i verstefall at de ikke får muligheten til å søke om studiestøtte på et senere tidspunkt, på grunn av at de tidligere har oppgitt feil opplysning.

Det er imidlertid bemerket at Lånekassen er lempelige på kravene som stilles til dokumentasjonen som sendes inn da de vet at det er forskjellig praksis for hvordan man betaler husleie.¹³⁴ Dette vil trekke i retning av at så lenge man sender inn noen form for

¹³² Prop.81 L (2016–2017) kapittel 8.8.3 s. 57.

¹³³ Bokkontroll basert på maskinlæring, En vurdering ved Riksrevisjonen i samarbeid med Statens lånekasse for utdanning punkt 3.1 s. 10.

¹³⁴ Bokkontroll basert på maskinlæring, En vurdering ved Riksrevisjonen i samarbeid med Statens lånekasse for utdanning punkt 3.1 s. 10.

dokumentasjon for at man bor utenfor barndomshjemmet så vil man fortsatt kunne motta støtte.

Den konkrete ulempen i AWIR-saken knytter seg til at man muligens mister støtten til barnepass. Som nevnt tidligere kunne dette utgjøre opp mot 30 000 euro.¹³⁵ Et beløp på denne størrelsen kan ikke regnes som en beskjedent sum noe som tilsier at det ikke bare er en ubetydelig ulempe å miste støtten man har søkt om.

Hvis man ser bort fra beløpet som personene muligens må tilbakebetale, så er også barnepassaspektet et viktig moment. Hele poenget med AWIR var å hjelpe de småbarnsfamiliene som ikke hadde mulighet til å ta seg av barna sine i arbeidstiden. Hvis personen ikke har råd til å kunne betale for barnepass av småbarn så vil dette gå ut over familiens økonomi på en annen måte, for eksempel ved at en av foreldrene må endre sin arbeidssituasjon.

Hvis man ser dette i sammenheng med den nasjonale diskrimineringsretten så er ett av formålene å styrke kvinners og minoriteters stilling i samfunnet.¹³⁶ Normen i samfunnet er at hvis det ikke er mulig å sende barna til barnepassere, så er det en av foreldrene som må være hjemme, som oftest mor. Hvis det i tillegg er de med annen nasjonalitet som oftere blir trukket ut i kontroll, vil dette også bety at kvinnelige minoriteter oftere vil måtte tilbakebetale stønaden de hadde mottatt. Det vil være de kvinnelige minoritetene som faller utenfor samfunnet på grunn av de ikke har mulighet til å få barnepass. Dette vil både gå ut over hver enkeltperson som en konkret ulempe, men det vil også være en ulempe i en bredere samfunnskontekst. Vi ønsker at immigranter skal integreres i samfunnet fordi de er en viktig samfunnsressurs som arbeidskraft, men også fordi at de kanskje ikke har behov for like mye velferdsytelser om de er i varig arbeid.¹³⁷

Ulempen i SyRI-saken vil også treffe forskjellig hvis man ser på det med et smalt eller vidt blikk. På den ene siden så vil ulempen for den enkelte som blir tatt ut i kontroll være tyngende. For det første så var det kun personer i problemdistrikt som ble tatt ut i kontroll. Dette vil trekke i retning av at personene allerede opplever tyngende forhold i sitt privatliv.

¹³⁵ Se avhandlingens punkt 4.2.

¹³⁶ Likestillings og diskrimineringsloven § 1 siste ledd.

¹³⁷ NOU 2011:7 s. 9 og 12

Kontrollering av offentlige myndigheter vil kunne forsterke stereotypingen om at problemdistrikt er gettoer, og at de som bor der er en byrde for samfunnet. Dette vil kunne forsterke en tanke om oss mot dem, hvor de som er en høy risiko kontinuerlig vil havne på kanten med velferdssamfunnet. Som vist til i kapittel 2 av avhandlingen er stereotyping og gjentatt kontroll en ulempe i seg selv.

På den andre siden må det ses hen til den ulempen kontrolleringen i SyRI-saken kan medføre kontroll i andre saker. Dette gjelder også for generell utvelgelse for kontroll ved maskinlæring. Som vist til tidligere er maskinlæringssystemer basert på historiske data. Et viktig moment innenfor dette er at de som tidligere har blitt ansett som en høy risiko vil kunne fortsette å bli ansett som en høy risiko i fremtiden på grunn av datasettet. Dette vil medføre at de også senere vil være de som må utlevere seg selv og alltid sende inn mer dokumentasjon for å kunne motta en eller annen form for støtte.

Herunder vil man måtte se direkte koblinger til ulempen personer som blir kontrollert av politiet opplever.¹³⁸ De som opplever å bli stoppet av politiet i kontroll kan føle seg utpekt og mistenkeliggjort på bakgrunn av deres etnisitet eller andre diskrimineringsgrunnlag. Hvis personer med annen etnisitet alltid skal være de som blir ansett som «høy risiko» så vil dette tilsi at man alltid vil diskriminere dem fordi de alltid må motbevise samfunnet at de ikke er noen som får uriktig stønad eller at de er kriminelle på andre området. Diskrimineringsnemda kom for eksempel frem til at en hyppig kontroll av borgere helt klart medførte en ulempe.¹³⁹

Både internasjonale regler og nasjonale lover stiller strenge krav til hjemmel for utlevering av persondata. Vi har både et nasjonalt og et internasjonalt regelverk som regulerer retten til privatliv, familieliv og dens hjem.¹⁴⁰ Med dette må det også forstås at det å bli kontrollert hver gang man er i kontakt med det offentlige er tyngende og inngripende i en persons private sfære. Det må imidlertid bemerkes at retten til privatliv og familie ikke gir en person rett på

¹³⁸ Se avhandlingens kapittel 2.

¹³⁹ Se avhandlingens kapittel 2.1 og 2.1.1 hvor Diskrimineringsnemda konkluderte i DIN-2020-175

¹⁴⁰ Den europeiske menneskerettighetskonvensjon artikkel 8.

sosial stønad på noe som helst vis eller en type levestandard. Dette har blant annet blitt bemerket av Den europeiske menneskerettsdomstolen (EMD).¹⁴¹

I tillegg til retten til privatliv har man retten til et familieliv. Det å motta stønad til sin familie kan være en måte for stater å demonstrere at de anerkjenner og respekterer familielivet.¹⁴²

Hvis disse familiene skal kontinuerlig bli kontrollert på grunn av at de har et større antall familiemedlemmer vil dette kunne oppfattes som tyngende for disse familiene noe som igjen kan tilsi at de ikke søker på stønad som de har behov for.

7.4 Rettferdiggjøring

Ett av de viktigste spørsmålene man må stille når det gjelder saker som kan forskjellsbehandle er hvorfor akkurat denne gruppen skal kontrolleres, og om denne forskjellsbehandlingen kan rettferdiggjøres. Dette vil knytte seg opp til unntaksbestemmelsene for forskjellsbehandling som man finner i ldl. §§ 9, 10 og 11.

Forskjellsbehandling som har et «saklig formål», er «nødvendig for å oppnå formålet» og ellers ikke er «uforholdsmessig inngripende overfor den som er forskjellsbehandlet», rammes ikke av diskrimineringsforbudet, jf. ldl. § 9.

Hvis man finner i en sak at det er forskjellsbehandling av et diskrimineringsgrunnlag, og at dette utgjør en ulempe for personen(e), kan det likevel være nødvendig, saklig og forholdsmessig?

Kravet om rettferdiggjøring krever at man skal gå nærmere inn i dybden hver enkelt sak hvor det er terskelen for den enkelte sak som er avgjørende. Dette vil medføre at man ikke har anledningen til å ta en helhetlig drøftelse under rettferdighet på grunn av at den vil bli for vag. Som nevnt i punkt 5.5 må diskrimineringsvernet bli vurdert i individuelle saker, dermed kan man ikke kan vurdere hvert enkelt tilfelle konkret. Vurderingen blir gjort etter lovens vilkår om hva som kunne blitt lagt vekt på i en individuell sak.

¹⁴¹ Di Trizio v. Switzerland, no. 7186/09, § 59, 2. februar 2016, og Jeunesse v. the Netherlands [GC], no. 12738/10, § 131, 3. oktober 2014.

¹⁴² Dhahbi v. Italy, no. 17120/09, premiss 41.

Spørsmålet under rettferdiggjøring må på grunn av dette for det første knytte seg til hvilket «saklig formål» sakene ønsket å oppnå med sitt maskinlæringsprogram.¹⁴³

Hvis en ser det Lånekassen så var målet å kunne kontrollere flere personer som borteboere, men at maskinlæringsmodellen skulle bli så presis at det ikke var nødvendig å foreta for mange kontroller. Hvor det tidligere kun har vært tusen stykker som har blitt kontrollert blir nå ti tusener kontrollert.

Målet med AWIR-saken var formålet behovsprøving av inntekstrelaterte ordninger og annen stønad til befolkningen. Dette kan sammenlignes med folketrygden i Norge. AWIR-systemet skulle i sin opprinnelige form være en form for å gi økonomisk trygghet til de som var særlig utsatt i samfunnet. Videre skulle programmet være en kontroll av de som oppholdt seg ulovlig i landet og som mottok støtte de ikke hadde krav på. Formålet med SyRI-saken var å bekjempe velferdssvindler ved å overvåke og utføre kontroller av innbyggerne.

Spørsmålet under rettferdiggjøring må for det andre vise til at forskjellsbehandlingen måtte «være nødvendig for å oppnå formålet» til sakene.¹⁴⁴

Hvis man faktisk ser en tendens til at flere menn jukser med bostøtten sin, så kan dette tilsi at man burde kontrollere flere menn. Det samme gjelder for de i befolkningen med annen etnisitet. Hvis de oftere begår velferdssvindler, burde man dermed kontrollere flere av dem. Det burde imidlertid kun bli kontrollert like mange som den historiske dataen viser at jukser. Henholdsvis 4 % av kvinnene og 7 % av mennene. Den falske positive raten har gjort det slik at maskinlæringsystemet trukket ut flere menn enn det som er nødvendig for kontroll. Hvis maskinlæringsystemet i tillegg kun trekker ut personer fra problemdistrikt med annen etnisitet vil ikke dette være «nødvendig for å oppnå formålet» siden personer fra alle samfunnsklasser begår kriminalitet.

Spørsmålet under rettferdiggjøring må for det tredje knytte seg til om forskjellsbehandlingen var «uforholdsmessig inngripende for den eller de som forskjellsbehandles».¹⁴⁵

¹⁴³ Ldl. § 9 første ledd bokstav a

¹⁴⁴ Ldl. § 9 første ledd bokstav b

¹⁴⁵ Ldl. § 9 første ledd bokstav c

Forholdsmessighetskravet i avhandlingens sakseksempler må ses i sammenheng med de som blir stoppet av politiet i kontroll. Det at man jevnlig blir mistenkeliggjort på grunn av personlig informasjon som man ikke kan endre kan oppleves som inngripende i en persons sfære. I DIN-2020-175 konkluderte Diskrimineringsnemda med at det var uforholdsmessig inngripende for personen å bli kontrollert gjentatte ganger av politiet. Nemda går ikke noe videre inn på hvor den forholdsmessige/uforholdsmessige grensen går ved kontroll. Om nemda videre skal håndtere saker hvor personer har blitt kontrollert jevnlig så burde de uttrykkelig konkretisere hvilke ulemper personen(e) har opplevd, og i hvilken grad dette opplevdes uforholdsmessig inngripende for den enkelte.

En kontroll på bakgrunn av etnisitet være uforholdsmessig tyngende fordi det legges opp til en presumsjon at kun etniske personer er lovlige eller normalen i samfunnet. Hvis man hele tiden blir ansett som et avvik vil ikke dette anses som forholdsmessig.

7.5 Konklusjon

De overnevnte sakene i kapittel 6 viser flere vanskelige vurderingsmomenter som man må ta stilling til innenfor diskrimineringsretten. For det første er det vanskelig for personer i det hele tatt forstå eller oppdage at det foreligger forskjellsbehandling. For det andre kan det være vanskelig å konkludere om det er forskjellsbehandling om man ikke har innsyn i algoritmene. Dette kan gjøre det vanskelig å oppdage om det vil oppstå en ulempe for de kontrollerte personene forut for kontrollen. Til sist kan det oppstå vanskeligheter med vurderingen av rettferdiggjøringen på grunn av at vurderingen blir for vag.

Gjennomgangen av sakene viser at dagens lovgivning er passende til å anvende på disse sakene. Det må imidlertid bemerkes at resultatet må være forklarbar slik at man vet hva man skal anvende diskrimineringsvernet på. Forklaringen for at man kan ta et klart standpunkt om dagens rettslige begreper er egnet til slike situasjoner er todelt.

For det første så er likestillings- og diskrimineringsloven er nøytralt i den forstand at vernet gjelder på alle samfunnsarenaer. Dette gjelder både teknologisk og fysisk. Det er mulig å fastslå at maskinlæringssystemer diskriminerer på bakgrunn av resultatene, så det vil for det første være mulig å bruke diskrimineringsretten på teknologiske arenaer.

For det andre så har vi allerede aktive plikter som kan benyttes i diskrimineringsspørsmål utført av kunstig intelligens. Hvis en ser hen til aktivitetsplikten som offentlig myndighet må følge, så kan denne benyttes for det digitale likestillingsarbeidet. For eksempel om man er

Skatteetaten eller Lånekassen, så må man påse at rutinene innad i virksomheten og den type velferdsgoder man tilbyr ikke diskriminerer.

Aktivitetsplikten gir plikt til både de ansatte, men også virksomheten i sin helhet til å motarbeide diskriminering og fremme likestilling. I tillegg vil en slik aktivitetsplikt være hjelpelig med bevisføringen for den påståtte diskriminerte og virksomheten på grunn av at de må arbeide «aktivt», «målrettet» og «planmessig» med sitt likestillingsarbeid. I dette må det legges en plikt til å kunne vise til de systemene som blir anvendt av den offentlige virksomheten.

Den viktigste årsaken til hvorfor man har forutinntatte algoritmer er at vi mennesker diskriminerer. Maskiner har ingen moral. For å kunne hindre at diskriminering skjer er det et behov for et kraftig kunnskapsløft for de som anvender kunstig intelligens. Ansatte både innenfor det offentlige og det private må kunne være oppmerksom på diskrimineringsmuligheten, og burde dermed ha en aktiv plikt til å forhindre diskriminering.

Hvis den som håndterer eller jobber sammen med algoritmen stoler fullt og helt på den, vil dette kunne medføre at mennesket i loopen automatisk aksepterer ethvert resultat fra prediksjonsmodellen. For å kunne minske risikoen for at kunstig intelligens diskriminerer må det være en menneskelig inngripen i beslutningsprosessen. Dette mennesket må se over datamaskinens arbeid og godkjenne det.

Videre må dette mennesket ikke være for selvsikker på sine kunnskaper om kunstig intelligens. Et mulig problem med å bruke algoritmer er at mennesket som er i loopen tror at han eller hun har en forståelse for programvaren og dermed godkjenner resultatet uten videre. Dette vil kunne medføre tjenestefeil.

Et forslag for å redusere risikoen av automatisert aksept fra saksbehandleren er at det skal være gode rutiner og opplæring av veiledere. Informasjonen må være forståelig og gi veilederne muligheten til å kunne sammenligne datamaskinens vurderinger opp mot andre momenter som for eksempel lovgitte regler man må holde seg innenfor.

Et annet forslag er kvitteringsordninger som gir personen et varsel om at det har blitt tatt en vurdering som indikerer at det kan foreligge en høyere risiko for å bli utsatt en kontroll. Dette vil medføre at man har notoritet på hvor mange kontroller som har blitt varslet om, og muligens de som også blir utført. Denne kvitteringsordningen burde også oppgi hvorfor

personen har blitt ansett som en høyere risiko enn andre. For eksempel slik som at personer kan få innsyn i sin sykejournal. Dette beror imidlertid på at det må være mulig å forklare utfallet av algoritmen.

Et siste forslag på hvordan man kan unngå at maskinlærende programmer gir oss uforklarlige resultater er at man legger ned et krav om at det ikke er lov til å bruke svart boks-modeller. Mangel på forklarbarhet om hvorfor algoritmen kom til det resultatet som den gjorde vil være et problem når man skal legge frem bevis i en sak. I stedet for svart boks burde man ta i bruk en hvit boks som gir en forklaring på hvordan man kom frem til det resultatet som man kom frem til. En hvit boks vil være behjelpelig hvis man kan konkludere med at det foreligger forskjellsbehandling og at bevisbyrden er overført til den som er ansvarlig.

Som Skatteetaten har vist til så vet de ikke hvorfor algoritmen kommer frem til det resultatet det har gjort på grunn av de har tatt i bruk systemer som ikke oppgir forklaringer.¹⁴⁶ Hvis maskinlæringsmodellen til Skatteetaten hadde forskjellsbehandlet noen, så ville ikke Skatteetaten vært i stand til å legge frem bevis etter ldl. § 37. Så selv om deres forskjellsbehandling kunne blitt godkjent som lovlig, så vil maskinlæringsmodellen være et hinder for bevisskravet. Dette vil dermed tale i favør for den som har blitt forskjellsbehandlet selv om det var et saklig formål, som var nødvendig for å oppnå formålet og ikke var uforholdsmessig inngripende for personen.

Det første forslaget vil gi de som anvender maskinlæringen et ansvar for å hindre diskriminering. Det andre forslaget vil være behjelpelig med å opplyse om de ulempene en person kan oppleve som følge av gjentatte kontroller av det offentlige. Det tredje forslaget går ut på hvordan maskinlæringssystemet er blitt laget.

Konklusjonen er at det kan oppstå brudd på diskrimineringsforbudet ved anvendelse av maskinlæring i kontroll av borgere som har mottatt goder gitt av velferdsstaten. De tre mulige løsningsforslagene over vil både være en jobb for de som utvikler maskinlæringssystemet hvor de må kunne forklare resultatet, og for den lovgivende makt som må påse at kunstig intelligens er i samsvar med dagens regler og rettigheter. Forslag nummer tre gir også

¹⁴⁶ Se avhandlingens punkt 3.2 om forklaring av «svart boks».

mulighet for at det offentlige må konkretisere at de kun vil ta i bruk algoritmer som er forklarbar.

Referanseliste

Lovregister

- | | |
|------|--|
| 1814 | Lov 17. mai 1814 Kongerikets Norges Grunnlov |
| 1999 | Lov 21. mai 1999 nr. 30 om styrking av menneskerettighetenes stilling i norsk rett (menneskerettsloven mrl.) |

2008	Lov 15. mai 2008 nr. 25 om utlendingers adgang til riket og deres opphold her (utlendingsloven)
2017	Lov 16. juni 2017 nr. 50 likestillings- og diskrimineringsombudet og Diskrimineringsnemnda (Diskrimineringsombudsloven diskrol.)
2018	Lov 15. juni nr. 38 om behandling av personopplysninger (personopplysningsloven)

Forarbeider, Norges offentlige utredninger og forskrift

Ot.prp.nr.77 (2000–2001)	Om lov om endringer i likestillingsloven mv. (plikt til å arbeide for likestilling, skjerping av forbudet mot forskjellsbehandling på grunn av kjønn, forbud mot seksuell trakassering mv.)
Ot.prp.nr.33 (2004–2005)	Om lov om forbud mot diskriminering på grunn av etnisitet, religion mv. (diskrimineringsloven)
NOU 2009:14	Et helhetlig diskrimineringsvern Diskrimineringslovutvalgets utredning om en samlet diskrimineringslov, grunnlovsvern og ratifikasjon av tilleggsprotokoll nr. 12 til EMK
NOU 2011:7	Velferd og migrasjon Den norske modellens framtid
Dok. 16 (2011–2012)	Rapport fra Menneskerettighetsutvalget om menneskerettigheter i Grunnloven

Prop.88 L (2012–2013)	Diskrimineringslovgivning (diskrimineringsloven om seksuell orientering, likestillingsloven, diskrimineringsloven om etnisitet, diskriminerings- og tilgjengelighetsloven)
Prop.81 L (2016–2017)	Lov om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven)
Prop.63 L (2018–2019)	Endringer i diskrimineringsombudsloven og likestillings- og diskrimineringsloven (etablering av et lavterskeltilbud for behandling av saker om seksuell trakassering og en styrking av aktivitets- og redegjøringsplikten)
FOR-2019-03-15-254	Forskrift om tildeling av utdanningsstøtte for undervisningsåret 2019–2020

Høyesterettspraksis

Rt. 2012 s. 424 A

Rt. 2014 s. 402 A

HR-2020-2476-A

HR-2022-350-A

Avgjørelser fra Den europeiske menneskerettighetsdomstolen

Jeunesse v. the Netherlands

Jeunesse v. the Netherlands [GC],

ECHR:2014:1003JUD001273810

Dhahbi v. Italy

Dhahbi v. Italy [J],

ECHR:2014:0408JUD001712009

Di Trizio v. Switzerland

Di Trizio v. Switzerland [J],

ECHR:2016:0202JUD000718609

Forvaltningspraksis, avgjørelse fra Nederland og Veneziakommisjonen

Diskrimineringsombudet (2013)

13/615

Diskrimineringsnemnda (2020)

DIN-2020-175

Rechtbank Den Haag (2020)

ECLI:NL:RBDHA:2020:1878, 5. februar 2020

Engelsk oversettelse av SyRI-avgjørelsen.

Veneziakommisjonen (2021)

The Netherlands opinion on the legal protection of citizens, CDL-AD (2021)031

Diskrimineringsnemnda (2021)

DIN-2021-262

Direktiv, forordninger og konvensjoner

Den europeiske
menneskerettskonvensjon

Konvensjon om beskyttelse av
menneskerettighetene og de grunnleggende
friheter, Roma 4. november 1950.

Kvinnediskrimineringskonvensjonen

De forente nasjoners internasjonale konvensjon 3.
september 1981 om avskaffelse av alle former for
diskriminering av kvinner med tilleggsprotokoll 6.
oktober 1999
(Kvinnediskrimineringskonvensjonen/KDK)

Personvernforordningen	Europaparlaments- og rådsforordning (EU) 2016/679 av 27.april 2016 om vern av fysiske personer i forbindelse med behandling av personopplysninger og om fri utveksling av slike opplysninger samt om oppheving av direktiv 95/46/EF (generell personvernforordning) (PVF, GDPR)
Rasediskrimineringskonvensjonen	De forente nasjoners internasjonale konvensjon 21. desember 1965 om avskaffelse av alle former for rasediskriminering (Rasediskrimineringskonvensjonen/RDK)

Litteratur

Ballangrud og Søbstad (2021)	Ballangrud, Anne Jorun Bolken og Søbstad, Margrethe «Likestillings- og diskrimineringsloven - lov 16. juni 2017 nr. 51 om likestilling og forbud mot diskriminering- lovkommentar»
Byrådet Oslo kommune (2019)	Plattform for byrådssamarbeid mellom Arbeiderpartiet, Miljøpartiet De Grønne og Sosialistisk Venstreparti i Oslo 2019-2023, 22. oktober 2019. https://oslo.arbeiderpartiet.no/nyheter/byradsplattform-2019-2023/
Bøe og Bjørke (2022)	Bøe, Henrik og Bjørke, Christian Nicolai, «TikTok viser ikke krigen i Ukraina til brukere i Russland», publisert 03.04.2022, sist lest 13. april 2022 https://www.nrk.no/vestfoldogtelemark/xl/tiktok-viser-ikke-krigen-i-ukraina-til-brukere-i-russland-1.15898864 .
Crawford (2017)	Crawford, Kate, The trouble with bias, Conference on Neural Information Processing Systems, 29. oktober 2017.

- https://www.youtube.com/watch?v=fMym_BKWQzk
(Sett 25. mars 2022)
- Dastin (2018) Jeffrey Dastin, «Amazon scraps secret AI recruiting tool that showed bias against women», Reuters, 11. oktober 2018
<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>
(lest 3. April 2022)
- Datatilsynet (2018) Kunstig intelligens og personvern, rapport fra Datatilsynet, 11. januar 2018.
<https://www.datatilsynet.no/regelverk-og-verktoy/rapporter-og-utredninger/kunstig-intelligens/>
- Europakommisjonen (2018) Independent High Level Expert Group set up by the European Commission: A definition of AI: Main capabilities and disciplines, 18. desember 2018
- Gerards og Xenidis (2020) A Gerards, Janneke (Utrecht University) og Xenidis, Raphaële (University of Edinburgh and University of Copenhagen) algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law. Rapport utviklet av Europakommisjonen, 2020.
<https://www.equalitylaw.eu/downloads/5361-algorithmic-discrimination-in-europe-pdf-1-975>
- Hellum og Strand (2021) Hellum, Anne og Strand, Vibeke Blaker *Likestillings- og diskrimineringsrett, Midlertid kompedieutgave for studenter*, 1. utgave, Gyldendal 2021.
- Kommunal og moderniseringsdepartementet (2020) Nasjonal strategi for kunstig intelligens, Kommunal og moderniseringsdepartementet, 14. januar 2020.

- <https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/>
- Levitsky, Allison (2020) “For the first time, white men weren't the largest group of U.S. hires at Google this year”, Silicon Valley Business Journal, 5. mai 2020. Lest 1. mai 2022.
<https://www.bizjournals.com/sanjose/news/2020/05/05/for-the-first-time-white-men-werent-the-largest.html>
- Najibi, Alex (2020) «Racial Discrimination in Face Recognition Technology», Harvard University, 24. oktober 2020, (lest 30. april 2022).
<https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/>
- NTNU, Helsingfors
Universitet, Reaktor mfl. Nettbasert studie. Elements of AI.
<https://course.elementsofai.com>
- Pedersen og Skoghøy (2020) Pedersen, Jussi Erik og Skoghøy, Jens Edvin A. «Grunnleggende sivilprosess», Oslo, 1. utg. Universitetsforlaget 2020
- Politiet (2018) I forkant av kriminaliteten, Forebygging som politiets hovedstrategi (2021-2025).
<https://www.politiet.no/globalassets/05-om-oss/03-strategier-og-planer/i-forkant-av-kriminaliteten.pdf>
- Riksrevisjonen (2019) Bokkontroll basert på maskinlæring, En vurdering ved Riksrevisjonen i samarbeid med Statens lånekasse for utdanning, 2019. Saksnr. 2019/01316
- Skatteetaten (2016) Skatteetatens Analysenytt, mars 2016
<https://www.skatteetaten.no/globalassets/om-skatteetaten/analyse-og-rapporter/analysenytt/analysenytt2016-1.pdf>

- Skoghøy (2015) Skoghøy, Jens Edvin A., *Rett og rettsanvendelse*, Oslo, Universitetsforlaget 2018.
- Statistisk sentralbyrå (2022) Trafikkulykker med personskade, oppdatert 13. april 2022
<https://www.ssb.no/transport-og-reiseliv/landtransport/statistikk/trafikkulykker-med-personskade>
- van Bekkum og Borgesius (2021) van Bekkum, Marvin og Borgesius, Frederik Zuiderveen «Digital welfare fraud detection and the Dutch SyRI judgemen», *European Journal of Social Security*, 2021, Vol 23 (4)

