

CompLex



Senter for rettsinformatikk / Avdeling for forvaltningsinformatikk

Kjetil Løyland

En litteraturstudie om maskinlæring og jus

1/2018



UiO **Det juridiske fakultet**

Henvendelser om denne bok kan gjøres til:

Senter for rettsinformatikk

Postboks 6706 St. Olavs plass

0130 Oslo

Tlf. 22 85 01 01

<http://www.jus.uio.no/ifp/om/organisasjon/seri/>

ISBN 978-82-72261-65-7

ISSN 0806-1912

Grafisk produksjon: 07 Media AS – 07.no

Innhold

1	Maskinlæring og jus – en innledning	5
1.1	Om studien	5
1.2	Grunnleggende om maskinlæring og jus	7
1.3	Oversikt over den videre studien	8
2	Forskning på metadatabasert maskinlæring innen jus	10
2.1	Generelt	10
2.2	Særlig relevant forskningslitteratur om metadatabasert maskinlæring	10
2.3	Avsluttende kommentar – metadata mot fritekst	13
3	Forskning på fritekstbasert maskinlæring innen jus	14
3.1	Generelt	14
3.2	Særlig relevant forskningslitteratur om fritekstbasert maskinlæring	16
3.2.1	Innledning	16
3.2.2	Fritekstbaserte maskinlæringsmodeller anvendt på rettsavgjørelser	16
3.2.3	Fritekstbaserte maskinlæringsmodeller anvendt på lovbestemmelser	19
3.2.4	Fritekstbaserte maskinlæringsmodeller anvendt på argumenter utledet fra rettsavgjørelser	20
3.3	Avsluttende kommentar – begrensinger i fritekstbasert analyse ...	24
4	Kommersielle anvendelser av maskinlæring innen jus	27
4.1	Metadatabasert maskinlæring i kommersiell virksomhet	27
4.2	Fritekstbasert maskinlæring i kommersiell virksomhet	29
5	Forventet bruk av maskinlæring innen jus	31
5.1	Fremtidig bruk av maskinlæring innen jus generelt	32
5.1.1	Begrensninger i teknologien slik den er i dag	32
5.1.2	Nye analyseverktøy	33
5.1.3	Jusstudiets og rettsvitenskapens fremtid	34

5.2	Fremtidig bruk av maskinl�ring i Norge	35
5.2.1	Det norske rettssamfunnets potensielle digitale infrastruktur	35
5.2.2	En konservativ norsk juriststand?	37
5.2.3	S�rtilfelle: Den norske rettskildel�ren	37
5.3	�nskede og u�nskede konsekvenser av � anvende maskinl�ring p� det juridiske omr�det	39
6	Kildeliste	42

1 Maskinlæring og jus – en innledning

1.1 Om studien

Litteraturstudien er skrevet på oppdrag fra Senter for rettsinformatikk (SERI) ved Institutt for privatrett ved Universitetet i Oslo. Bakgrunnen for studien var et ønske om å opplyse relevante miljøer om forskningen på et lite utforsket område i norsk sammenheng.

Formålet med litteraturstudien er å gi en systematisk og oversiktlig gjennomgang av faglitteratur for å identifisere ulike typer anvendelser av maskinlæring innen det juridiske feltet. Studien omfatter altså ikke rettsdogmatiske og rettspolitiske drøftelser mv. av maskinlæring, men er begrenset til kartlegging av anvendelser slik situasjonen var ved årsskriftet 2017/2018.

Jeg har løst oppdraget ved å gjøre litteratursøk i biblioteksdatabaseen Oria, tilknyttet Universitetet i Oslo. Jeg har ikke funnet litteratur om emnet på norsk. Siktemålet har vært å kartlegge og redegjøre for den sentrale internasjonale forskningslitteraturen om emnet. Siden mine språkkunnskaper begrenser seg til engelsk, har jeg ikke gjennomgått eventuell litteratur om emnet skrevet på andre språk.

Gjennomgående for litteratursøkene på dette området er at de nødvendigvis inneholder søkeord som er sterkt assosiert med andre fagområder enn jus. Søkeordet «machine learning» er for eksempel et kjernebegrep innen informatisk litteratur, og søkeordet «prediction» er på lignende måte et kjernebegrep innen statistisk litteratur. Samtlige søk gjort i forbindelse med studien ga mange treff, men relevant litteratur om maskinlæring og jus var som regel gjemt mellom lange lister med kilder fra ren informatisk og statistisk litteratur. Jeg gjennomgikk treffene med litteratur til det punktet jeg mente var rimelig. Jeg brukte vanlig øyesyn og undersøkte titler som potensielt kunne ha noe med det juridiske området å gjøre. Jeg understøtte titlene nærmere ved både å se på den oppgitte beskrivelsen av artikkelen på Oria, og i noen tilfeller ved å påbegynne lesing av artikkelen før jeg innså at det ikke var relevant materiale for studien.

Jeg søkte først på «machine learning and law» som ga 109 345 treff på Oria. Artikkelen «Machine Learning and Law» (2014) av Harry Surden ble rangert som mest relevant av søkemotoren. Dette er en innføringsartikkel om emnet.

Nedover listene fant jeg «Argument Based Machine Learning» (2009) av Martin Mozina m.fl., «Can Artificial Intelligence and Online Dispute Resolution enhance efficiency and effectiveness in Courts» (2017) av John Zeleznikow og «The Law Machine» (2013) av Ted Harbert som jeg anså relevante for studien.

Deretter søkte jeg på «legal machine learning». Dette søket ga 63 845 treff. Artikkelen «Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective» (2016) av Tsarapatsanis m.fl. ble funnet høyt oppe på listene.

Søket «legal analytics» som problemområdet noen ganger kalles ga ingen relevante treff. Søkeordet «legal prediction» ga 123 206 treff, deriblant sentrale «Quantitative legal prediction» (2013) av Daniel Martin Katz.

Litteratursøkene gjort i søkemotoren til Oria ga begrensede resultater. Litteratursøket ble derfor supplert med en fremgangsmåte som bestod i å undersøke henvisninger i den litteraturen jeg hadde funnet, for så å søke opp disse enkeltvis i Oria. Dette ga betydelig bedre resultater. For å illustrere dette kan jeg nevne at fra «Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective» (2016) av Tsarapatsanis, D mfl, fant jeg henvisninger til «Automatically Classifying Case Texts and Predicting Outcomes» (2009) av Kevin Ashley og Stephanie Bruninghaus. Fra denne artikkelen fant jeg henvisninger til «Predicting Securities fraud settlements and amounts: a hierarchical Bayesian model of federal securities class» (2012), BB McShane mfl.

Jeg har også gjort personsøk i Oria med navnene til de sentrale forfatterne i den litteraturen jeg hadde gjennomgått. Jeg søkte først på «Daniel Martin Katz». Dette ga 17 289 treff og mange artikler som ikke knyttet seg til emnet. Men jeg fant også «A general approach for predicting the behaviour of the Supreme Court of the United States» (2017) av Michael Bommarito, Josh Blackman og Daniel Martin Katz.

Deretter søkte jeg på «Kevin Ashley». Dette ga 12 818 treff. Søket ga også en rekke artikler han har skrevet om emnet gjennom sin yrkeskarriere. I stedet for å la hans artikler være gjengangere i litteraturstudien, lot jeg samleverket «Artificial Intelligence and Legal Analytics» (2017) av Kevin Ashley være styrende for litteraturstudien, fremfor hans enkelte artikler. Fra denne boken fant jeg dessuten en rekke kilder som jeg har brukt i litteraturstudien.

Nesten all litteraturen som gjennomgås er fra fagfellevurderte tidsskrifter eller fra bøker som er gitt ut av seriøse forlag. Det lille som ikke er fra fagfellevurderte tidsskrifter er internettsider som forteller om ulike produkter i skjæringspunktet mellom maskinlæring og jus som ikke typisk dekkes i artikkelformat. Litteraturen jeg gjennomgår i det følgende representerer etter mitt skjønn den mest sentrale og relevante litteraturen om anvendelser av maskinlæring innen det juridiske feltet.

1.2 Grunnleggende om maskinlæring og jus

Med «jus» i denne litteraturstudien mener jeg alle områder som kan forbindes med juridisk aktivitet. Det kan strekke seg fra dogmatiske rettsspørsmål, til rettskildelæren, og til utøvelse av juridisk virksomhet som jusstudent eller jurist. Med «maskinlæring» mener jeg en undergruppe av kunstig intelligens som omhandler dataprogrammer som kan lære fra erfaring og på den måten utvikle seg over tid.¹

Selv om teknikkene man benytter seg av kan variere, er det stort sett enighet i informatikklitteraturen om definisjonen av maskinlæring. Når ordet «læring» brukes i definisjonen er det viktig å understreke at det ikke er snakk om kognitiv læring, men *funksjonell* læring der maskinen kan lære å utføre bestemte operasjoner på en bedre (mer funksjonell) måte. Selv om maskinlæring er en undergruppe av det man kaller kunstig intelligens, er maskinlæringsteknikker kun intelligente hvis man anlegger et funksjonelt perspektiv. Teknologien er ikke intelligent i betydningen kognitiv eller menneskelig, men den kan *fremstå* intelligent fordi den gjør oppgaver vi forbinder med menneskelig aktivitet. Surden skriver at forskningen på kunstig intelligens gjennomgikk et skifte på 1980-tallet da det lenge mislykkede prosjektet om å replisere menneskelig intelligens ble erstattet av forsøk på funksjonell læring via blant annet maskinlæring.²

Surden forklarer hvordan tilsynelatende menneskelige oppgaver kan bli løst ved bruk av maskinlæring, ved å vise til et eksempel om automatiske spamfiltre i e-post³: Et automatisk spamfilter kan lages ved at en person gir en maskinlæringsalgoritme eksempler på e-post med spam/ikke-spam. Dette er eksempel på veiledet læring, altså der en person mater maskinlæringsalgoritmen. Algoritmen vil deretter forsøke å finne fellestrekk (felles attributter) mellom e-postene markert som spam på den ene siden, og e-postene som er markert som ikke-

1 Surden, H. (2014) s. 89.

2 *Op. cit.* s. 98.

3 *Op. cit.* s. 96-97.

spam på den andre siden. Algoritmen vil så forsøke å konstruere generelle regler for hvilke e-poster som skal kategoriseres som spam.⁴ Algoritmen kan for eksempel oppdage at utsagnet «earn extra cash» ofte finnes i e-postene personen har kategorisert som spam, mens det ikke finnes i e-postene personen har kategorisert som ikke-spam. Algoritmen kan da danne regelen om at e-post med utsagnet «earn extra cash» tenderer statistisk å være spam mer enn ikke-spam. Etter å ha dannet en rekke regler tester personen algoritmen på et nytt sett med e-poster. Algoritmen skal da automatisk klare å filtrere ut e-post med spam.

Algoritmene kan bruke tekst som datagrunnlag. Den ovennevnte regelen om at e-post med «earn extra cash» statistisk tenderer å være spam, benytter fritekst som datagrunnlag. Fritekst er lite strukturert tekst. I rettslig sammenheng kan fritekst for eksempel være lover og domsavgjørrelser. En maskinlæringsalgoritme kan også bruke metadata som datagrunnlag. Metadata er kontekstoppbyggende informasjon. Det kan for eksempel være tall, navn, klokkeslett, person. Dersom spam-algoritmen som er nevnt ovenfor kommer frem til en regel om at e-post med Hviterussland som opphavsland mest sannsynlig er spam, vil denne regelen benytte seg av metadata som datagrunnlag. De fleste maskinlæringsalgoritmer vil bruke både metadata og fritekst som datagrunnlag. I det følgende vil jeg likevel begrepsmessig skille mellom eksempler på maskinlæring som hovedsakelig bygger på metadata og eksempler på maskinlæring som hovedsakelig bygger på fritekst. Bakgrunnen for dette skillet er dels at maskinlæring som hovedsakelig bygger på fritekst må prosesseres gjennom språkteknologi.⁵ Dels vil henholdsvis metadatabasert og maskinlæring basert på fritekst ofte ha så forskjellige mål, at et begrepsmessig skille anses nødvendig.⁶

1.3 Oversikt over den videre studien

Det finnes både akademisk forskning på problemområdet jus og maskinlæring og kommersiell bruk av maskinlæring i jus. Den akademiske forskningen knytter seg normalt til utviklingen av modeller som kan anvendes på det juridiske området, mens den kommersielle bruken av maskinlæring i jus typisk er hjelpemerketøy som brukes av forskjellige juridiske aktører. Både den akademiske forskningen og den kommersielle bruken av maskinlæring i jus er beskrevet i forskningslitteraturen gjennomgått i forbindelse med denne studien.

4 En algoritme er en beskrivelse av en serie operasjoner som skal utføres for å løse et problem.

5 Natural Language Processing (NLP).

6 Det begrepsmessige skillet følger også av systematikken i det sentrale verket til Ashley (2017) som brukes aktivt i hele litteraturstudien.

I kapittel 2 og 3 vil jeg gjengi hovedinnholdet i den akademiske forskningen på henholdsvis metadatabasert og fritekstbasert maskinlæring innen jus. I kapittel 4 beskriver jeg den juridisk-kommersielle bruken av henholdsvis metadatabasert maskinlæring i 4.1 og fritekstbasert maskinlæring i 4.2. Kapittel 2, 3 og 4 er en litteraturgjennomgang, og en deskriptiv gjengivelse av den konkrete anvendelsen av maskinlæring i jus i dag, slik bruken fremgår av den tilgjengelige utvalgte litteraturen. I kapittel 5 presenterer jeg en hypotese om fremtidig bruk av maskinlæring innen jus. Jeg reflekterer også over ønskede og uønskede effekter av en slik mulig utvikling.

2 Forskning på metadatabasert maskinlæring innen jus

2.1 Generelt

Daniel Martin Katz skrev i 2013 et omfattende arbeid som forsøkte å introdusere metadatabasert maskinlæring på det juridiske området.⁷ Katz mener det først og fremst er prediksjonspotensialet som gjør anvendelsen av maskinlæring innen jus verdifullt.⁸ Interessen for prediksjoner speiles også i den øvrige forskningslitteraturen om maskinlæring innen jus. Katz' poeng er først og fremst at mange advokater vil kunne tjene penger på å komme med raske og gjennomtenkte prediksjoner til klienter i rådgivningsøyemed. En maskinlæringsbasert prediksjonsalgoritme vil ifølge Katz ha større og klarere minne enn en advokat. Det er også et poeng at algoritmene vil søke å gi en nøytral prediksjon uavhengig av advokatens utenforliggende interesser.

Prinsippene bak maskinlæringsbasert prediksjon er enkle: Man bruker en algoritme som er trent opp (enten manuelt eller automatisk) på et gitt datasett (eldre saker), og bruker det på et nytt datasett (nye saker). Algoritmen skal da kunne forutsi resultatet i en ny sak. Hvorvidt algoritmen kommer til riktig resultat, er avhengig av hvor godt datasettet med eldre saker er. Siden denne delen omtaler metadatabasert maskinlæring, er interessen knyttet til de algoritmene som hovedsakelig bruker metadata som datagrunnlag som for eksempel informasjon om hvordan dommere har dømt tidligere eller informasjon om hvilket resultat advokatene i saken har fått til i tidligere saker.

2.2 Særlig relevant forskningslitteratur om metadatabasert maskinlæring

Michael Bommarito, Josh Blackman og Daniel Martin Katz publiserte i april 2017 resultatene fra et forsøk om prediksjoner av avgjørelser fra den amerikanske føderale høyesterett i «A general approach for predicting the behavior of the Supreme Court of the United States».⁹ Prediksjoner om hva den amerikanske føderale høyesterett vil gjøre er en øvelse som har opptatt amerikanske juridiske

⁷ Katz, Daniel M. (2013).

⁸ *Ot. cit.* s. 912.

⁹ Bommarito, M., Blackman, J., Katz, D.M. (2017).

og politiske kommentatorer i mange tiår. Forfatterne skriver at prediksjoner av avgjørelsene ikke bare er viktige for aktørene i saken, men også for samfunns- livet for øvrig. Et eksempel er det tilfellet hvor et børsnotert selskap er involvert i en sak i domstolen. Forfatterne ønsket å finne ut om maskinlæring er et frukt- bart alternativ til etablerte prediksjonsmetoder, for eksempel ekspertkommen- tatorers vurderinger.

Forskerne ønsket å undersøke om maskinlæring kunne brukes til to predik- sjonsoppgaver. For det første undersøkte de om modellen de konstruerte kunne forutsi resultatet i saken bedre enn tilfeldig gjetning.¹⁰

For det andre undersøkte de om modellen de konstruerte kunne forutsi den enkelte dommerstemmen bedre enn tilfeldig gjetning.

Forskerne brukte metadata som datagrunnlag. Som nevnt tidligere er dette kon- tekstuelle variabler, og ikke variabler basert på innholdet i domsteksten. Fors- kerne hentet domsavgjørelser fra en sentral digital database for avgjørelser fra den amerikanske føderale høyesterett. Modellen de konstruerte brukte blant annet følgende variabler hentet fra disse dokumentene:

- Navn på dommer
- Årstall
- Måned
- Navn på ankedomstolen saken kommer fra
- Navn på ankende part
- Navn på ankemotparten

Forskerne matet også modellen manuelt med andre variabler de mente kunne være fruktbare. Dette omfattet blant annet om muntlig høring var satt opp for saken eller ikke, og tiden fra muntlig høring til domstolens avgjørelse. Disse tilleggsopplysningene var kvalitative observasjoner som manuelt ble satt inn i modellen.

Modellen skulle forutsi om en anket sak kom til å bli tatt til følge eller bli stående. Dette er en binær klassifikasjonsoppgave. Selv om den amerikanske føderale høyesterett primært er en ankedomstol, kan domstolen noen ganger fungere som opphavsdомstol for føderale rettsspørsmål. I disse sakene vil naturligvis ikke resultatet være om en anke blir tatt til følge eller bli stående, men heller om saksøker eller saksøkte får medhold. Disse sakene passer ikke med det binære utgangspunktet for forsøket, og ble derfor ekskludert fra resultatene.

10 Det engelske statistiske begrepet «null model».

Modellen forutså riktig resultat i saken 70,2 % av gangene, og forutså 71.9 % av de enkelte dommerstemmene. Artikkelforfatterne hevder det er en mer eller mindre etablert sannhet i det amerikanske juridiske miljøet at ankene til den føderale høyesterett blir tatt til følge i 63 % av sakene og at 57 % av dommerstemmene tar ankene til følge basert på de siste 35 terminene. Sammenligningsmodellen tilfeldig gjetning ble derfor satt til sannsynligheten for å treffe riktig ved kun å satse på at anken ble tatt til følge. Tilfeldig gjetning-modellen ble på den måten satt til å treffe 63 % riktig i resultater i saken og 57 % riktig i dommerstemme. Modellen til artikkelforfatterne gjorde det bedre enn tilfeldig gjetning.

I «Predicting Securities fraud settlements and amounts: a hierarchical Bayesian model of federal securities class action lawsuits» utforsker BB McShane mfl mulighetene for bruk av metadatabasert maskinlæring til å analysere utfall av investeringsbedragerisaker.¹¹ Dette er typisk gruppesøksmål ført av investorer som har kjøpt eller solgt verdipapir og opplevd økonomisk tap som følge av brudd på verdipapirlovgivningen. Etter å ha opparbeidet seg en database med gruppesøksmål, ønsket forskerne å anvende et bayesiansk nettverk. Bruk av bayesiansk nettverk tar sikte på å representere et sett av tilfeldige variabler og deres betingede avhengigheter fremstilt ved hjelp av en graf. Teknikken brukes gjerne til å finne kausale forhold. Et bayesiansk nettverk kan for eksempel brukes til å utrede eventuelle sammenhenger mellom sykdommer og symptomer. I dette forsøket ble teknikken brukt til å undersøke hvilke kontekstuelle variabler som spiller inn på resultatet i investeringsbedragerisøksmål.

Variablene brukt i dette forsøket omfattet blant annet (1) om sakene ble forlikt eller avvist, (2) hvor mange verdipapirhandler som var knyttet til saken, (3) tiden det tok før gruppen gikk til søksmål, (4) om grupperepresentanten for saksøker var et individ eller en institusjon, (5) om det involverte selskapet var børsnotert, (6) den eventuelle børskursen, (7) om innsidehandel var påstått, og (8) om det forelå brudd på aktuelle revisjonsregelverk (her «GAAP»¹²).

Forskerne fant at variabler som kausalt påvirket at et søksmål ble forlikt, blant annet kunne være (1) at det var mange verdipapirhandler knyttet til saken, (2) at aktuelt revisjonsregelverk var brutt (her «GAAP») og (3) at grupperepresentanten for søksmålet var et individ i motsetning til en institusjon. Variabler som talte mot at søksmålet ble forlikt var blant annet (1) at gruppen brukte lang tid på å gå til søksmål, (2) at børskursen hadde økt og (3) at det ikke var oppført noen grupperepresentant for saksøker.

11 McShane, B.B. m.fl. (2012).

12 «Generally Accepted Accounting Principles».

2.3 Avsluttende kommentar – metadata mot fritekst

Gjennomgående for metadatabasert maskinlæring innen jus er at modellene ikke befatter seg med fritekst, for eksempel i form av begrunnelsen i rettsavgjørelser. Prediksjonene vil derfor ikke være basert på alminnelige krav til juridisk argumentasjon, i Norge tydeliggjort ved rettskildelæren. Ashley skriver om metadatabasert maskinlæring at «its rules may not necessarily seem reasonable to humans».¹³

Prediksjoner basert på metadata kan være et verdifullt supplement for politiske og juridiske kommentatorer som skal gi spådommer til en bred offentlighet. Men når prediksjonen ikke er basert på juridisk argumentasjon, reiser det et spørsmål om hvor hensiktsmessig denne type prediksjoner er som grunnlag for juridisk rådgivning.

Investeringsbedragerisaksforsøket til McShane mfl kan antas å ha verdi på det juridiske området. Ved å anvende det bayesianske nettverket vil man i større grad kunne avdekke faktiske forhold som er av betydning i slike saker. Gruppesøksmål kan ofte ende med forlik. Siden forlik ofte kan være begrunnet i utenrettslige og selskapsstrategiske vurderinger som for eksempel effekten av dårlig PR, kan avdekking av kausale faktiske forhold være nyttig. Hvis man med denne type undersøkelser klarer å fastsette grensen for når et selskap anser seg tjent med et forlik eller ikke, kan det i større grad være hensiktsmessig å begrunne et gruppesøksmål med forhold avdekket i denne typer undersøkelser. I saker som involverer forlik kan undersøkelser basert på metadata være en verdifull tjeneste i forlengelsen av typiske rettsråd.

Prediksjoner basert på metadata kan altså ha en misjon innen det juridiske området. Samtidig er det vanskelig å se hvordan juridisk analyse forankret i krav om juridisk argumentasjon kan forbedres ved å anvende denne type modeller. For at maskinlæring skal kunne virke inn på denne delen av jussen, bør den hovedsakelig basere seg på fritekst som datagrunnlag.

13 Ashley, K. (2017) s. 111.

3 Forskning på fritekstbasert maskinlæring innen jus

3.1 Generelt

Det jeg kaller fritekstbasert maskinlæring er maskinlæringsalgoritmer som hovedsakelig bruker fritekst som grunnlag. Fritekst er som nevnt lite strukturert tekst som i rettslig sammenheng kan være for eksempel lover og domsavgjørelser. For at friteksten skal kunne bli behandlet av en algoritme må den prosesseres gjennom metoder fra språkteknologien¹⁴.

Siden jeg tar utgangspunkt i et skille mellom metadatabasert maskinlæring og fritekstbasert maskinlæring blant annet på grunn av prosessering av tekst, kan det være nyttig å gå gjennom grunnprinsippene for prosessering.

Savelka og Grabmair har forsøkt å gi en oppskrift på hvordan man kan gå frem for å programmere maskinlæringsalgoritmer basert på fritekst, ved bruk av metoder fra språkteknologien. Jeg skal gjengi denne i kortversjon her, og den fulle oppskriften kan leses i Ashleys samleverk.¹⁵

Prosessering:

Språkteknologi kan defineres som teknikker for å transformere menneskelig språk til interne representasjoner som dataprogrammer kan veilede til å gjøre ulike oppgaver.¹⁶ Prosessen består i å gjennomføre denne transformasjonen. I oppskriften til Savelka og Grabmair er målet å transformere et sett med ulike tekstmasser til det man kaller en støttevektormaskinmodell som er en metode for veiledet maskinlæring.

Ulike tekstmasser kan være domstekster innenfor et spesifikt rettsområde, for eksempel alle avgjørelsene fra Høyesterett om den ulovfestede obligasjonsrettslige læren om *condictio indebiti*.

Støttevektormaskinmodeller brukes ofte til å klassifisere ulike elementer eller til regresjonsanalyser. Det er denne maskinlæringsteknikken jeg ovenfor har beskrevet i avsnitt 1.2 om e-poster med spam/ikke-spam. Hvis man mater algoritmen med treningseksempler på noe som er markert å tilhøre én av to katego-

¹⁴ Natural Language Processing (NLP).

¹⁵ Ashley, K (2017), s. 236-240.

¹⁶ Dale, R (2000): «Systems that map human language into internal representations that computer programs can manipulate appropriately to perform a task».

rier, skal treningsalgoritmen for støttevektormaskinen lage en modell som kan avgjøre om et nytt eksempel faller inn under den ene kategorien eller ikke.

Steg 1: Finne data

Det første steget er å finne teksten man skal prosessere. I en norsk juridisk sammenheng kan dette for eksempel være Høyesteretts avgjørelser om *condictio indebiti* hentet fra Lovdata eller Rettsdata. Deretter skal disse prosesseres.

Steg 2: Språkteknologisk prosessering – Stemming, leksikalsk analyse og annotasjon

Stemming¹⁷: Ved å «stemme» ordene i teksten, avgjør man den morfologiske roten til et ord for å kunne eliminere overflatiske varianter. Det kan være den morfologiske roten i form av et ords bøyingsform eller avledning. For eksempel vil et egenavn «Hagstrøm» bli til «hagstrøm», og bøyingsformene «skriver» og «skrevet» blir til «skrive».

Leksikalsk analyse¹⁸: Leksikalsk analyse er blant annet ordseparering. Setningen «*Condictio indebiti* er en ulovfestet lære» blir til «*condictio*», «*indebiti*», «er», «en», «ulovfestet», «lære». En annen bruk av leksikalsk analyse er å konvertere en sekvens med tegn til en sekvens med token, altså en tekststreng med en identifiserbar mening. Denne meningen kan for eksempel være «ord» i motsetning til «tall». Leksikalsk analyse går også gjerne ut på å eliminere stoppord, det vil si ord som brukes mye i tekst, men som har liten informasjonsverdi. Eksempler er determinativer som «en» og «et» eller preposisjoner som «i» og «på».

Leksikalsk analyse gjøres gjerne sammen med en parsing. Parsing går ut på å konvertere en streng med symboler i henhold til formell grammatikk. Formell grammatikk er ulike formasjonsregler som definerer hvilke strenger som er syntaktisk gyldige.

Annotasjon¹⁹: Man bruker annotasjon til å gi mening til ord i teksten som kan være uklare. Et eksempel kan være ordklassetaggning som er å kategorisere ord i ulike klasser, for eksempel substantiv, verb, adjektiv eller adverb for å gjøre det enklere for algoritmen å gi mening til teksten.

Etter at teksten i datasettet er prosessert blir teksten representert som attributtvektorer. En attributtvektor er en nummerisk representasjon av teksten slik at den skal kunne bli behandlet av algoritmen. Hvert vektorelement har en nume-

17 Det engelske informatiske begrepet «Word stemming».

18 Det engelske informatiske begrepet «Tokenize».

19 Det engelske informatiske begrepet «Annotation».

risk verdi som representerer en egenskap eller et trekk ved teksten. Attributtvektoren kan for eksempel representere bruken av ordet «condictio» i en tekst.

3.2 Særlig relevant forskningslitteratur om fritekstbasert maskinlæring

3.2.1 Innledning

I den følgende oversikten over relevant forskningslitteratur om fritekstbasert maskinlæring, har jeg delt aktuelle forsøk inn i tre typer modeller. Den første modelltypen jeg skal gjennomgå er modeller som baserer seg på fritekst fra rettsavgjørelser, den andre gjelder tekst fra lovbestemmelser, og den tredje knytter seg til argumenter utledet fra rettsavgjørelser.

3.2.2 Fritekstbaserte maskinlæringsmodeller anvendt på rettsavgjørelser

Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective – Tsarapatsanis og Alteras (2016)

Forskerne ønsket å benytte fritekstbasert maskinlæring til å utvikle et verktøy som kan støtte analyse av potensielle brudd på Den europeiske menneskerettighetskonvensjonen (EMK). Målet var å danne en prediksjonsmodell som kunne forutsi brudd på EMK i Den europeiske menneskerettighetsdomstolen (EMD).

Forskerne understreker at tidligere arbeid om prediksjon av rettsavgjørelser i stor grad har vært basert på metadata, jf. avsnitt 2.2. Forfatterne understreker videre at deres forsøk representerer den første systematiske studien om prediksjon av rettsavgjørelser fra en stor internasjonal domstol, gjort ved hjelp av maskinlæringsbasert fritekstanalyse.

Ved å sammenligne likheter i tekst fra rettsavgjørelser fra EMD og saksdokumenter i en ny sak, skal man ifølge forskerne kunne finne indikasjoner på resultatet i den nye saken. Forfatterne forklarer at de aller helst skulle sammenlignet saksdokumentene i tidligere saker og saksdokumenter i ny sak. Men siden de ikke har hatt tilgang til saksdokumenter i tidligere saker, måtte de gjøre bruk av EMD-avgjørelsene i stedet. Forskerne understreker at dette forbeholdet må tas i betraktning når man vurderer forsøket.

Forskerne definerer problemet brudd/ikke-brudd på en artikkel i EMK som en binær klassifikasjonsoppgave. Tekstmassen ble hentet fra HUDOC som er den elektroniske databasen for avgjørelser fra EMD. Før artikkelforfatterne går gjennom metoden de har brukt, problematiserer de kort om man kan stole på

teksten om faktum som dommerne skriver i domstekstene. Denne problematiseringen er også relevant i en norsk kontekst, jf. for eksempel Hans Petter Gravers arbeid om retorikk og jus. Graver fremhever blant annet at faktumtekster som er brukt i høyesterettsavgjørelser speiler det avsluttende resultatet i saken.²⁰ Dette gjelder enten det er bevisst eller underbevisst. Dette gir grobunn for tanken om at man ikke kan bruke det oppgitte faktum som trainingssett i en modell som skal forutsi et resultat. Forskerne forutsetter likevel i forsøket at man må kunne stole på det oppgitte faktum i tidligere avgjørelser. For det første har EMD få muligheter til å utrede faktum, og må stole på faktum slik nasjonale domstoler har presentert det. For det andre mener forfatterne at det ville ha blitt oppdaget om EMD åpenlyst presenterte en skjev faktumbeskrivelse i rettsavgjørelser. Dessuten finnes det få eksempler på at parter har klaget på EMDs faktumbeskrivelser.

I forsøket bruker forskerne EMK artikkel 3, 6 og 8 og avgjørelser tilhørende disse artiklene. Begrunnelsen for utvalget er at disse bestemmelsene inngår i mange avgjørelser, noe som gir en stor tekstmasse. Forskerne utelot å bruke 10 % av rettsavgjørelsene i tekstmassen, slik at disse i stedet kunne brukes for å teste modellens prediksjoner.

Etter å ha samlet tekstene knyttet til artiklene, ble teksten i avgjørelsene delt inn i flere deler som tilsvarende avgjørelsesstrukturen i EMD. Først kommer prosesshistorien («procedure») som angir hvordan saken kom til EMD. Deretter kommer delen om faktum («the facts») som er all informasjon som ikke er domstolens drøftelse, herunder både relevante faktiske omstendigheter i saken («The circumstances of the case») og relevant nasjonal jus som skal prøves («relevant law»). Til slutt kommer kategorien jus («law») som er domstolens drøftelse hvor både det påståtte bruddet på EMK nevnes («alleged violation of article X»), partenes anførsler («Parties submissions») og domstolens drøftelse («merits»).

Forskerne tok utgangspunkt i rettsavgjørelser knyttet til hver enkelt artikkel. Alle rettsavgjørelsene ble parsert inn i den avgjørelsesstrukturen jeg har gjenomgått ovenfor. Forskerne fjernet ordsekvenser som nevner konklusjonen i de ovennevnte delene. Dette gjorde de ved å bruke metoden *regulære uttrykk* slik at de ulike delene ikke anga konklusjonen. Ved regulære uttrykk kan man programmere eliminering av gitte ordsekvenser i tekstmassen. Deretter preprosesserer teksten med språkteknologi. Forskerne fjernet blant annet stoppord i tekstmassen.

20 Graver, H.P (2013).

Fra tekstmassen med EMD-avgjørelser utledet forskerne flere n-gram. Et n-gram er en sekvens med ett eller flere elementer, gjerne ord fra tekstmassen. Et n-gram bestående av ett ord kalles unigram og et n-gram bestående av to ord kalles bigram. I dette forsøket utledet forskerne flere unigram, og satt dem inn i en «bag-of-words»-modell (BOW). I denne modellen blir tekstmassen representert som en «bag» – en sekk – med flere unigram. Her fjernes all syntaktisk og semantisk sammenheng mellom ordene.

For hvert sett med avgjørelser i tekstmassen, satt programmet opp de topp 2000 mest brukte ordsekvensene. Så konstruerer forskerne «topics» for hver EMK-artikkel ved å sette sammen ordsekvenser som er semantisk like. Forskerne følger det de kaller en «distributional hypothesis». Dette er tanken om at semantisk like ord også opptrer i like kontekster. For artikkel 3 i EMK om tortur var en topic «Treatment by state officials». Ofte brukte ordsekvenser i rettsavgjørelser som anga brudd på artikkelen var «police», «officer» og «force». Dersom nye saksdokumenter også inneholder de samme ordsekvensene som går igjen i rettsavgjørelsen, antar modellen at den nye saken er et brudd på EMK.

Etter å ha brukt 90 % av tekstmassen på å komme frem til modellen, tester forskerne modellens prediksjoner med de urørte 10 % av tekstmassen. Forskerne kommer deretter frem til at modellene i gjennomsnitt kan forutsi med 79 % sikkerhet riktig resultat i nye saker.

Ettersom forskerne delte alle avgjørelser i flere deler, kunne de også meddele at delen som beskriver faktum er den delen av en EMD-avgjørelse som gir mest treffsikre prediksjoner. Det er altså faktumlikhet forskerne mener er den mest sentrale variabelen når en skal forutsi om det er brudd på en av de nevnte ovenfor nevnte artiklene i EMK eller ikke.

Forfatterne går så inn i en diskusjon om funnet knyttet til faktumlikhet og forutsigelse løser en klassisk rettsteoretisk debatt mellom det som kalles rettsformalister («legal formalists») og rettsrealister («legal realists»). Med rettsformalister mener forfatterne teoretikere som tror på en viss juridisk formalisme som for eksempel argumentasjon med rettskildefaktorer. Med rettsrealister mener forfatterne teoretikere som mener realiteten i saken (faktum) er det helt sentrale for dommeren, og at den juridiske begrunnelsen kun er et etterfølgende skalleskjul. Forfatterne slutter at forsøket kan brukes som argument for at rettsrealistene har rett. Forskerne fremsetter denne konklusjonen med forebehold grunnet datausikkerheten og forenklingen av debatten.

3.2.3 Fritekstbaserte maskinlæringsmodeller anvendt på lovbestemmelser

Interaktivt søkeverktøy for lover – Savelka mfl. (2015)

Savelka mfl. har forsøkt å lage et interaktivt verktøy som bruker juridiske kilde-søk fra databaser. Et norsk eksempel er Lovdatas lov- og forskriftsbase. Formålet til modellene i denne sammenhengen er å optimalisere kildesøket til å treffe mer presise og riktige lovbestemmelser.

Dersom man søker etter en lovbestemmelse som kan passe til et gitt faktum, søker man gjerne på ord som er forbundet med faktum. Dette kan gi en rekke resultater, men et problem som ofte oppstår er at den riktige lovbestemmelsen ikke anses mest relevant av søkemotoren. Det er denne utfordringen Savelka vil forsøke å løse med maskinlæring. I verktøyet blir søkerens valg av lovbestemmelse lagret og programmert inn i en modell. Modellen blir altså matet med positive og negative eksempler som gjør at den kan se fellestrekk og danne generelle regler. Disse reglene anvendes når man står overfor et nytt søk etter lover med en annen faktumbeskrivelse. Modellen oppdateres, ser etter nye forbindelser og utvikles ytterligere. Det blir et samspill mellom manuelle søk etter lovbestemmelser og modellens automatiske rutiner. Men i motsetning til eksempelet med e-poster i avsnitt 1.2 ovenfor, er det tredjemenn (brukere) som står for den manuelle jobben slik at verktøyet inngår naturlig i kildesøktjenesten.

Savelka beskriver resultatene som gode. Verktøyet klarte å klassifisere riktig i 80 % av tilfellene.

DALOS – Francesconi mfl (2010)

Ashley skriver om en ontologi som heter DALOS laget av Francesconi.²¹ DALOS står for the Drafting Legislation with Ontology base Support. En ontologi er innen informasjonsvitenskapen en formell representasjon av et sett med begreper innenfor et kunnskapsområde. Denne ontologien ble laget for å gjøre lovfor-slag i medlemsstater i EU om rettsområdet forbrukerbeskyttelse enklere. En ontologi skal gjenspeile menneskelig ekspertkunnskap om et område. Det sentrale med DALOS er at den anvendte maskinlæring til å identifisere begreper og forhold mellom begrepene automatisk slik at ontologien ble enklere å sette sammen.

Målet med ontologien var å tilby en taksonomi som kunne være juridisk relevant. En taksonomi er en systematisering og klassifisering av noe. DALOS forsøkte å tilby en taksonomi for typer av normative bestemmelser som knytter seg

21 Ashley, K (2017), s. 178-180.

til forbrukerbeskyttelse, men også et begrepsspråk på engelsk og italiensk for å beskrive typiske situasjoner relatert til forbrukerbeskyttelse.

Forskerne utledet begreper fra en tekstmasse med dokumenter om forbrukerbeskyttelse (kontrakter, lover osv) på engelsk og italiensk ved bruk av språkteknologi kombinert med maskinlæringsteknologi.

Resultatet modellen gav kom til syne i det som ble kalt det leksikalske nivået («the Lexical Layer»). Her er uttrykkene lenket til noen typer underliggende lingvistiske forhold: hyponyms, equivalents og fuzzynyms. Uttrykket «hyponym» refererer til et ord med mer spesifikk mening enn et generelt begrep som kan anvendes på ordet. Det engelske uttrykket «supplier» er lenket til dets hyponym, «supplier of goods», mens dets «equivalent» på italiensk er «fornitore». Et «fuzzynym» er ord som hører sammen siden de deler underliggende mening uten nødvendigvis å være semantisk like. Statistisk ble «supplier» og «consumer» brukt mye sammen i dokumentene uten at de likner semantisk, og de regnes da som fuzzynyms.

Det ontologiske nivået (Ontological Level) er den faktiske sammenhengen begrepene inngår i, mens det leksikalske nivået (Lexical Layer) er der modellen har plassert begrepene automatisk. Etter prosessen måtte en menneskelig kontrollør manuelt flytte begrepene fra det leksikalske nivået til det ontologiske nivået. Dette var en slags bekreftelsesøvelse. For eksempel er det semantiske forholdet mellom fuzzynyms som «supplier» og «consumer» også riktig på det ontologiske nivået. Forsøket var altså avhengig av en menneskelig kontrollør, men arbeidet skal ha blitt langt enklere med maskinlæring. Selve resultatet av den automatiserte kategoriseringen i det leksikalske nivået blir ikke oppgitt i Ashley (2017).

3.2.4 Fritekstbaserte maskinlæringsmodeller anvendt på argumenter utledet fra rettsavgjørelser

Legal argument mining – Mochales and Moens (2011)

I Mochales og Moens arbeid med automatisk å utvinne juridiske argumenter fra tekst, forsøkte forskerne å utlede grunnleggende bestanddeler i juridiske argumenter. Tanken var at automatisk kategorisering kan gjøre juridisk analyse enklere. Forskerne begynte med å definere et juridisk argument og dets sentrale bestanddeler.

Et argument ble definert som et sett med påstander som til sammen utgjorde premisser, unntatt det som måtte antas å være en konklusjon.

Forskerne ønsket å bruke maskinlæring til automatisk (1) å kategorisere setninger som påstander i et juridisk argument (eller ikke), og (2) å kategorisere påstander som premiss eller konklusjon i et juridisk argument.

1. Å kategorisere setninger som påstander i et juridisk argument eller ikke

Setninger ble brukt som attributtvektorer. For å lære opp modellen til å identifisere setninger som påstander i et argument, ble setningene representert som områdegenerelle attributter. Forfatterne ga eksempler på påstander i juridisk argumentasjon. Man anvendte deretter maskinlæring på treningssettet for å kunne utlede generelle regler om hva som utgjør en påstand i juridisk argumentasjon.

Man så blant annet på hvert enkelt ord og sammensetningen av ord i de identifiserte påstandene. Man så på ordklassetaggingen, altså sammensetningen av adverb, verb og modalitet. Påstander inneholdt for eksempel ofte modaliteter som «may», «must», «shall» eller «should». Modellen så også på ulike nøkkelord som indikerte påstand som «but», «consequently» eller «because of». Den så dessuten på tegnsetningsmønstre, syntaktisk analyse og visse typer tekststatistikk. Tekststatistikk kan være setningslengde, gjennomsnittlig ordlengde og antall punktum.

Modellen som skulle identifisere juridiske argumenter i EMD-tekstmassen traff riktig i 80 % av tilfellene.

2. Å kategorisere påstander som premiss eller konklusjon i et juridisk argument

Forskerne matet en modell med eksempler på premisser og konklusjoner. Modellen fant blant annet at variabler som setningslengde og setningenes plassering i dokumentet var viktige. At setningen kom mot slutten av dokumentet, kunne for eksempel indikere at man sto overfor en konklusjon. Bruken av verb-bøyning og verbtype var også viktig. Nøkkelordsøk som «see», «mutatis mutandis», «having reached this conclusion» og «by a majority» indikerte for eksempel at man sto overfor en konklusjon. Et annet eksempel er at setninger som siterte en lovbestemmelse eller inneholdt en rettslig definisjon gjerne indikerte premisser.

Resultatet for å kategorisere påstander som premisser eller konklusjoner i tekstmassen var på henholdsvis 68 % treffsikkerhet for premisser og 74 % treffsikkerhet for konklusjoner.

Automatically classifying case text (SMILE og IBP) – Ashley og Bruminghaus (2009)

Kevin Ashley og Stefanie Bruminghaus beskriver utviklingen av programmene SMILE og IBP. Til sammen skulle disse to programmene kunne gjøre automatiske prediksjoner av rettsavgjørelser med grunnlag i argumenter fra tidligere praksis.

Det ene programmet SMILE (Smart index learner) skulle lære automatisk å identifisere rettslige faktorer²² i korte tekstlige beskrivelser av saker. Uttrykket rettslige faktorer refererer til et sett med klassifikasjoner av momenter som beviselig har betydning i visse rettsspørsmål basert på tidligere rettsavgjørelser. I dette forsøket var det faktorer som har betydning i rettsavgjørelser om forretningshemmeligheter.²³ Etter at SMILE hadde identifisert ulike rettslige faktorer i teksten, ble disse sendt over til programmet IBP (Issue Based Prediction). Forskerne hadde programmert inn en rekke typiske rettsspørsmål som faktorene skulle gi et svar på. For hvert rettsspørsmål som ble reist, fant IBP ut hvordan de ulike rettslige faktorene hadde blitt veid i tidligere saker. Med grunnlag i hvordan spørsmålene ble besvart av faktorene, skulle IBP være i stand til å gi en prediksjon for hele saken. IBP gav også et forsøk på forklaring av avveiningen mellom de ulike rettslige faktorene som er forståelig for brukerne.

Den sentrale ideen bak forsøket var å forene praksisbasert resonnering og tekstrepresentasjon ved bruk av språkteknologi. Når det gjelder utviklingen av SMILE var det sentrale spørsmålet hvordan man kan representere tekst best mulig for å støtte automatisk ekstraksjon av rettslige faktorer. SMILE ble gitt et treningssett med setninger som var manuelt markert som positive og negative eksempler på rettslige faktorer. Forskerne brukte tre representasjonsteknikker for å teste hvilken som mest effektiv. Først testet de en bag-of-words-modell (BOW) der ordene ble representert helt uavhengig av den tekstlige sammenhengen de inngår i. I tillegg eliminerte de blant annet numre og tegnsetting. Forfatterne viser et eksempel på en prosessert representert setning i bag-of-words-modellen: «a agreement Dickman nondisclosure signed».

Deretter brukte de en teknikk kalt «Roles replaced». Denne teknikken ligner på bag-of-words-modellen ved at også denne fjerner sammenhengen ordene inngår i. Men med denne teknikken erstatter man de saksspesifikke opplysningene med standardiserte roller man typisk har i rettsavgjørelser, for eksempel en sak-søker eller saksøkt. Forfatterne illustrerer teknikken ved å bruke den samme eksempelsetningen: «a agreement defendant nondisclosure signed». Til slutt

22 Det engelske rettsinformatiske begrepet «Legal factors».

23 Det amerikanske rettsområdet «Trade Secret Law».

bruker forskerne en teknikk kalt «Prepositional phrases». Istedenfor å representere ord fokuserer teknikken på preposisjonsmønstre i setningen basert på fire predefinerte preposisjonskategorier: subjekt-verb, verb-objekt, verb-preposisjon og verb-adjektiv. Forskerne illustrerer det ved å vise den samme eksempelsetningen i representert form: «(defendant sign) (sign nondisclosure_agreement)».

Modellen bruker en nærmeste nabo-algoritme for å finne de setningene i nye saker som ligner de positive og negative eksemplene på rettslige faktorer. I dette forsøket brukte forskerne særegne notatbeskrivelser av faktiske forhold i nye saker, og ikke faktumbeskrivelser i rettsavgjørelser. Dette forholdet gjør at teknikken at resultatene ikke kan godtas uten forbehold.

IBP mottok altså et sett med rettslige faktorer fra SMILE. Deretter identifiserte IBP preprogrammerte rettsspørsmål som faktorene skulle løse. Et eksempel på et slikt rettsspørsmål er om den aktuelle informasjonen det tvistes om i saken er verdifull. Hvert rettsspørsmål ble så løst av faktorene etter hvilken vekt faktorene hadde hatt i tidligere rettspraksis. Eksempler på faktorer som taler for saksøkers side er at man får et konkurransefortrinn og at det er et unikt produkt. Etter at de forskjellige rettsspørsmålene blir løst, skulle programmet også kunne gi et svar på hvem som vinner saken. Dersom informasjonen er ansett verdifull taler dette for eksempel for at saksøker skal vinne frem med sin påstand om brudd på regler om forretningshemmeligheter.

Resultatene fra IBP var sterke. Den kom frem til riktig resultat 91,8 % av forsøkene. Det vil si at IBP kom frem til riktig resultat i 169 saker, og kun bommet i 14 saker. SMILEs resultater for automatisk å utlede rettslige faktorer var derimot vesentlig svakere på rundt 25 %. Forfatterne regner likevel forsøket som en suksess. Forfatterne trekker frem at det er første gang et program kan resonnerer automatisk med grunnlag i fritekst fra rettsavgjørelser.

LUIMA – Grabmair mfl. (2015)

LUIMA er forsøk på å lage et automatisk typesystem for juridiske tekster. Et typesystem fungerer som en struktur for markering av tekst. Et typesystem er som en slags ontologi for tekstanalyse ved at den definerer typer av markeringer, begreper og forhold som gjerne dukker opp i juridiske tekster. LUIMA brukte et etablert rammeverk kalt UIMA og tilpasset det til juridiske tekster. LUIMA fokuserte på begreper, forhold og annotering for å identifisere setningers argumentative rolle i rettsavgjørelser.

LUIMA har fire nivåer: «Sentence Level types», «Formulation types», «Mention types», og «Term types».

«Sentence Level types» i LUIMA fanger opp ni sentrale roller setninger kan ha i juridisk argumentasjon, for eksempel å etablere en regel, etablere et vilkår, å finne noe faktisk eller en konklusjon som finner noe bevist.

«Formulation types» fanger typiske uttrykk i juridisk argumentasjon og typiske måter dommere uttrykker seg på i rettsavgjørelser. I en norsk-rettslig sammenheng kan et eksempel være «(j)eg har kommet til...» som innleder votumets konklusjon.

«Mention types» fanger måter begreper i juridisk argumentasjon blir referert til i juridiske tekster. I norsk sammenheng kan man for eksempel vise til at Høyesterett ofte viser til ulovfestet rett.

«Term types» representerer grunnleggende terminologi brukt i juridisk argumentasjon, så vel som begreper som ofte brukes på et spesifikt rettsområde. Et eksempel på grunnleggende terminologi i norsk juridisk argumentasjon kan være «saksøker» og et eksempel på et begrep som brukes på spesifikt rettsområde kan være «skyld» som kan ha betydning i saker om *condictio indebiti*.

Ideelt skal mennesker som markerer tekstene kunne kategorisere avgjørelser etter disse typene etablert i LUIMA-strukturen. Dette blir et treningssett. Deretter skal modellen automatisk kunne identifisere markeringen i nye rettsavgjørelser basert på markeringen gjort av mennesker. Dette er mulig fordi modellen har fått en rekke positive og negative eksempler på riktig markering. Grabmair hevder den automatiske markeringen nesten blir feilfri. Siden begreper og forhold mellom begreper kan uttrykkes på mange forskjellige måter, vil også modellen kunne identifisere disse alternative uttrykksmåtene på en automatisk eller semi-automatisk måte. Ved en semi-automatisk fremgangsmåte vil modellen identifisere uttrykk, og menneskelige redaktører godta eller ikke godta kategoriseringen.

3.3 Avsluttende kommentar – begrensinger i fritekstbasert analyse

I stedet for å fokusere på lovbestemmelser eller rettsavgjørelser, er det flere forfattere som tar til orde for at den mest presise bruken av maskinlæring på juridisk tekst skjer gjennom setningers rolle i juridisk argumentasjon.

I forsøket til Tsarapatsanis og Alteras om EMD-prediksjoner tar ikke modellen de konstruerer hensyn til juridiske argumenter, men forsøker heller å gjøre prediksjoner på grunnlag av hvilke ord som brukes i tekstmassen. Som nevnt i gjennomgangen over indikerer blant annet bruk av ordsekvenser som «police», «officer» og «force» i nye saksdokumenter brudd på EMK artikkel 3. En innvending mot denne fremgangsmåten er at bruken av akkurat disse tre ordene ikke har noe juridisk argumentasjonsverdi. Bruk av disse ordene gir ingen veiledning i spørsmål om handlinger utført av statlige ansatte skal anses som tortur etter EMK artikkel 3. Ordsekvensene indikerer kanskje at statlige tilsatte har utført en handling, men det må anses å være en forutsetning for at spørsmålet i det hele tatt skal diskuteres i EMD. Et eksempel for å beskrive hvor outrert modellen kan være når den ikke fokuserer på argumenter, finner man også i EMD-forsøket. Blant topp 5-ordsekvenser som indikerte brudd på EMK artikkel 8 om rett til privatliv og familieliv fant man ordsekvensen «russian». Dersom nye saksdokumenter inneholdt ordsekvensen «russian» indikerte altså dette brudd på artikkel 8.

Martin Mozina mfl. argumenterer også for at tekstbasert maskinlæring bør befatte seg med argumenter istedenfor hele tekster. Dette er dels fordi Mozina antar det er mer effektivt å forholde seg til argumenter hvis tekstmassen er forholdsvis liten.²⁴ Selv om det finnes mange rettsavgjørelser i verden, kan det ikke sammenlignes med andre typer populære datagrunnlag i andre bransjer som for eksempel markedsføringsbyråer som gjerne henter datagrunnlaget de bruker fra sosiale medier. Dels trekker Mozina frem at det er positivt at resultatene ved argumentbasert maskinlæring kan forstås og tolkes av mennesker. At mennesker kan forstå avveiningene som er gjort blir også trukket frem som et stort fremskritt for forskerne som sto bak verktøyet SMILE og IBP nevnt over.

Interessen for argumenter i litteraturen må ses i sammenheng med en diskusjon av hvor fruktbare black box-algoritmer er på deler av det juridiske området. En black box-algoritme kjennetegnes gjerne ved at den kommer frem til et resultat uten å forklare hvordan. Kjente maskinlæringsbaserte programmer som IBM Watson er bygget på en denne type teknologi. Selv om den er programmert til for eksempel å finne rettsregler vil den ikke kunne vise hvordan den kom frem til regelen, fordi den ikke kan resonnerer med relevante regler og begreper.²⁵ Dette vil være godt nok for applikasjoner som tar sikte på å utføre tjenester, for eksempel å bedre søkefunksjonen i lovdata-baser. Men dersom man skal bruke maskiner til å finne frem til noe substansielt mener jeg det er en fordel at maskinen kan vise hvordan den frem til resultatet. Man kan forestille seg at noen har

24 Mozina, M (2009) s. 54.

25 Ashley, K (2017) s. 17.

konstruert et juridisk «black box»-program som klarer å forutsi domsresultater nært opp mot 100 % uten at man forstår hvordan programmet kommer frem til konklusjonene. Dette programmet kan være verdifullt for en analytiker som vedder på domsutfall, men ikke stort mer enn det. Hvis ikke programmet klarer å vise hvordan den kom til resultatet, vil det verken hjelpe aktørene i saken eller kunne si noe mer om jussen som sådan. Jeg mener det er grunn til å slutte seg til forskerne som hevder at maskinlæring innen jus i nær fremtid først og fremst vil bli viktig dersom man fokuserer på juridiske argumenter.

4 Kommersielle anvendelser av maskinlæring innen jus

4.1 Metadatabasert maskinlæring i kommersiell virksomhet

Advokatrekruttering

Katz skriver om advokatrekruttering.²⁶ Han trekker paralleller til ideen om Moneyball som har vakt oppsikt i sportsverdenen, og viser til et selskap som heter Lawyer Metrics. Dette firmaet utvikler modeller som skal gjøre prediksjoner om en aktuell kandidat vil være en god ansettelse for et advokatfirma. Modellene tar først og fremst sikte på å vurdere nyutdannede kandidater som ikke har mye erfaring med tidligere arbeid som advokat.

Tanken bak Lawyer Metrics og denne type HR-firmaer er å fjerne personlig bias i ansettelses- og evalueringsprosessene. De bruker maskinlæringsteknologi til å lage modeller for å finne ut hvilke karakteristiske trekk eller egenskaper som typisk gir en lønnsom advokat. Dette gjør de etter å ha undersøkt hvilke advokater som typisk er lønnsomme og de dataene disse advokatene kan måles etter. Dette er trekk som kan bli observert i kandidatens CV og vitnemål. Katz beskriver typiske kriterier som er aktuelle. Blant tradisjonelle variabler nevner Katz karakterer, rangering av institusjonen kandidaten er utdannet fra og praktikantopphold. Lawyer Metrics mener dessuten ifølge Katz at utradisjonelle viktige variabler som ofte blir oversett er arbeidserfaring fra lavinntektsjobber, andre universitetsgrader, deltakelse i lagsport og publisering av innhold i tidsskrifter.

Risikoanalyse – «Risk assesement»-verktøy

Private foretak utvikler algoritmer med metadatabasert maskinlæring knyttet til å analysere og håndtere ulike former for risiko. De mest kjente verktøyene knytter seg til gjentakelsessfare i strafferetten som aktualiserer spørsmål om både kausjonssummer («bond amounts») og straffeutmåling. New York Times opplyser dessuten at disse verktøyene blir brukt hos politiet i Kansas til å forutsi hvor det mest sannsynlig vil skje kriminalitet.²⁷ På denne måten kan politiet benytte seg av ressursene sine mer effektivt. Slike algoritmer blir også brukt til å gi myndighetene råd om domfeltes soningsforhold og eventuelle prøveløslatelse.²⁸

26 Katz, Daniel Martin (2013) s. 935.

27 Smith (2016).

28 *Ibid.*

Algoritmene har fått mye oppmerksomhet i amerikanske medier, særlig algoritmen Compas som er utviklet av Northpointe Inc. En mann i Wisconsin gikk til søksmål mot Northpointe Inc. da han ikke fikk innsyn i risikoanalysen som lå til grunn i straffutmålingsvurderingen mot ham. Northpointe Inc. argumenterte med at dette var en forretningshemmelighet som skulle forbli hemmelig. Northpointe Inc. fikk medhold i Wisconsins statlige høyesterett og anken til den føderale høyesterett ble avslått sommeren 2017.²⁹

Lex Machina

Lex Machina er et selskap som ble etablert av tidligere ansatte ved Stanford University. De har bakgrunn både fra jus og informatikk. Forretningsmodellen til Lex Machina baserer seg på å forutsi utfall av patentsøksmål basert på metadata som kan utledes fra domsdokumenter. Eksempler på denne type metadata kan være utfallet av saken, hvilke advokater som har deltatt i saken og hvilke dommere som har deltatt i saken.

Ved automatisk å inkludere disse dataene fra dommer inn i algoritmene har Lex Machina til enhver tid oppdatert informasjon over hvilke IP-advokater som normalt inngår forlik i søksmål, hvilke dommere som normalt gir en partstype medhold («patent trolls») osv.³⁰

Nøkkelen til suksessen for Lex Machina ligger ifølge Harbert i at de har fullstendig oversikt over metadataen i slike saker (Lex Machina-databasen). Produktet er attraktivt fordi brukere kan orientere seg tidlig i saksgangen på grunnlag av Lex Machinas metadatabaserte prediksjoner. Lex Machina tilbyr årlige abonnementer til ulike advokatfirmaer for rundt 50 000 dollar (2013).

Ravel Law

Ravel ble grunnlagt av nyutdannede fra Stanford Law School.³¹ Sammen med biblioteket til Harvard Law School har de forsøkt å scanne deler av amerikanske rettsavgjørelser på statlig nivå, og gjøre disse tilgjengelige i et digitalt format. I tillegg tilbyr Ravel visuelle kart som viser hvordan en dom siterer en annen dom i forbindelse med domsbaserte juridiske begreper.³² Et eksempel kan være et tenkt juridisk begrep som regulerer tilbakebetalingsplikt ved feilbetaling som ligner *condictio indebiti* som har blitt utviklet i amerikansk rettspraksis. Ravels digitale visuelle kart viser hvordan en rettsavgjørelse bruker andre rettsavgjørelser når den redegjør for den tenkte domsbaserte læren. Dette gjør den ved

²⁹ *Ibid.*

³⁰ Harbert, T (2013).

³¹ Ravel Law, 2018.

³² Ashley, K (2017), s. 354.

visuelle kartmodeller. Slike kart er altså pedagogiske verktøy som skal gjøre det mer intuitivt for jurister å analysere begreper.

Ravel leverer også analyse av tidligere saker basert på dommerhistorie og advokathistorie på samme måte som Lex Machina.

4.2 Fritekstbasert maskinlæring i kommersiell virksomhet

E-discovery

E-discovery er en applikasjon for å samle, utveksle og analysere elektronisk lagret informasjon som kan brukes som bevis før en rettssak. I mange saker kan det være snakk om millioner av elektroniske dokumenter. I E-discovery kalles maskinlæringsteknologien prediktiv koding.³³ Ashley beskriver hvordan en advokat går frem for å navigere i et slikt verktøy.³⁴

Ved å bruke generelle søkeverktøy kan en advokat gjennomgå et stort antall elektroniske dokumenter med ulike former for søk. Etter at advokaten har gjort sine søk, vil modellen beregne relevansen av ulike dokumenter. Dette etablerer en forbedret prediktiv modell som kan anvendes på resten av dokumentene.

Grossman og Cormack skrev i 2011 en artikkel der de hevdet at E-discovery i tillegg til å være vesentlig kostnadsbesparende, også kan være bedre til å analysere store tekstmasser enn mennesker.³⁵ Dette resultatet kom de frem til etter et forsøk som sammenlignet mennesker mot E-discovery i oppgaver om minne og presisjon.

Ross

Ashley skriver om Ross som er et skybasert juridisk spørsmål/svar-verktøy. Ross bruker teknologien til IBM Watson.³⁶ Det sentrale med Ross er at den aksepterer spørsmål på naturlig språk,³⁷ og gir svar basert på lovgivning, rettsavgjørelser og andre rettskilder. Svaret er normalt ikke et sammenhengende juridisk argument, men en henvisning til en rekke kilder som har henvisninger og drøftelser av de temaene som etterspørres. Det oppgis ikke på nettsidene hvor treffsikker Ross er.

33 Det engelske begrepet «Predictive coding».

34 Ashley, K (2017), s. 241.

35 Grossman og Cormack (2011).

36 Ashley, K (2017). s. 351.

37 «Natural language».

Split-Up

Split-Up er en australsk applikasjon som gir råd om bofordeling etter en skilsmisse.³⁸ For å utvikle Split-Up identifisererte 94 australske familierettseksperter faktorer som er viktige i saker om bofordeling etter skilsmisse. Ekspertene måtte utlede faktorene fra tidligere rettsavgjørelser slik at en maskinlæringsalgoritme kunne lære av variablene. På den måten lærte Split-Up hvordan dommere vektet faktorene i tidligere saker.

Annet

Eksempler på applikasjoner som ikke er berørt i forskningslitteraturen om maskinlæring og jus er Kira Systems³⁹, Luminance⁴⁰ og Neota Logic⁴¹. Dette er applikasjoner som blant annet tilbyr tjenester knyttet til automatisk kontraktsanalyse. Applikasjonene gjør det enklere å navigere i en stor kontrakts-tekstmasse blant annet ved å kategorisere kontraktene etter egne temaer. Et eksempel kan være temaet force majeure. Ved å navigere ut fra temaet force majeure skal man automatisk få opp de relevante delene av kontraktene som er lastet opp som knytter seg til force majeure. Det opplyses ikke på nettsidene hvor effektive disse applikasjonene er.

38 Zeleznikow, J (2017).

39 Kira Systema (2018).

40 Luminance (2018).

41 Neota Logic (2018).

5 Forventet bruk av maskinlæring innen jus

Katz skrev i 2013 at man de siste årene har opplevd en økt interesse for problemområdet maskinlæring og jus.⁴² Dette er en utvikling som også indirekte går frem av litteraturen nevnt i denne studien ettersom det tidligste arbeidet rapportert om er fra 2009. Katz skriver at utviklingen dels skyldes utvikling i datakraft⁴³, og dels at datalagring⁴⁴ har blitt vesentlig billigere.

Selv om artikkelforfatterne som er nevnt i denne litteraturstudien er positive til utviklingen videre, finnes det også kommentatorer som er skeptiske til en fremtid med utstrakt bruk av maskinlæring innen jus. Katz illustrerer denne type jurister med utsagnet «(y)ou cannot replace what I do with a computer».⁴⁵ Når man tenker på hva slags tankeøvelser juridisk argumentasjon krever, er det ikke så vanskelig å kjenne seg igjen i det utsagnet. Et eksempel fra norsk rett er argumentasjon basert på verdier eller reelle hensyn. For at et program skal kunne argumentere effektivt med slike flytende størrelser, er det grunn til å tro at programmet må ha kognitive evner lik et menneske.

Katz mener nyere teknologiske innovasjoner motbeviser skeptikeren, og viser til flere eksempler på etablerte sannheter som har vist seg å være uriktige.⁴⁶ Katz trekker frem IBM Watsons seier over en spillmester i Jeopardy i 2011 og fremveksten av selvkjørende biler. Både mestring av Jeopardy og bilkjøring er noe man tradisjonelt har forbundet med kognitive menneskelige evner, men dette har altså blitt ettertrykkelig motbevist. Etter å ha gått gjennom et utvalg forsøk med maskinlæring anvendt på jus, hevder Ashley det er urealistisk at fremtidige juridiske applikasjoner vil kunne bedrive dyp læring eller på egen hånd kunne formulere nye juridiske argumenter.⁴⁷ Jeg tolker denne tilsynelatende motsetningen som en skinnuenighet. Jeg forstår ikke Katz' utsagn som en påstand om at fremtidens applikasjoner vil kunne bedrive dyp læring, men heller at maskinlæring vil kunne håndtere det området av jussen som vi normalt forbinder med kognitive egenskaper. Katz gir derimot ingen konkrete anbefalinger om hvordan maskinlæring skal kunne løse disse oppgavene.

42 Katz, D.M (2013) s. 913-917.

43 Det engelske uttrykket «Computing power».

44 Det engelske uttrykket «Data storage».

45 Katz, D.M (2013) s. 922-928.

46 Ibid.

47 Ashley, K (2017), s. 381.

På dette stadiet i utviklingen kan man kanskje også si at diskusjonen om et data-programs evne til kognitiv juridisk argumentasjon virker litt unyansert. Selv om Ashley muligens har rett i at et dataprogram ikke kan bedrive dyp læring eller formulere nye juridiske argumenter, er det deler av jussen som i dag er egnet for maskinlæring. For det første finnes det områder i jussen hvor man langt på vei klarer seg uten den type skjønn som betinger dyp læring. Et eksempel fra norsk rett kan være trafikkssaker hvor man langt på vei er avhengig av helt faste regler uten innslag av skjønn. For det andre finnes det en rekke arbeidsoppgaver forbundet med juridisk virksomhet som kan forbedres, for eksempel kildesøk eller selskapsgjennomganger.⁴⁸

I den følgende fremstillingen vil jeg skille mellom antatt fremtidig bruk av maskinlæring innen jus generelt (5.1) og fremtidig bruk av maskinlæring innen jus i Norge (5.2). Den fremtidige bruken av maskinlæring generelt vil muligens sammenfalle med den fremtidige bruken i Norge. Når jeg likevel behandler Norge spesielt, er det fordi jeg da vil ha mulighet til å si noe om det jeg oppfatter som særnorske juridiske forhold.

Avslutningsvis vil jeg vurdere ønskede og uønskede konsekvenser av å anvende maskinlæring på det juridiske området (5.3).

5.1 Fremtidig bruk av maskinlæring innen jus generelt

5.1.1 Begrensninger i teknologien slik den er i dag

For å kunne gi en noenlunde presis hypotese om hvordan maskinlæring og jus vil se ut i fremtiden, er det nyttig å se på noen av begrensningene som knytter seg til dagens teknologi.

Surden skriver om begrensningene til maskinlæringsbaserte juridiske prediksjonsmodeller.⁴⁹ For det første må det man ønsker å forutsi gjelde de samme forhold som algoritmen bygger på. Dette begrenser universaliteten til modellene. Dette gjør at man antakelig må konstruere modeller for hvert enkelt område hvor man ønsker prediksjoner. Dette rigide trekket ved teknologien gjør den kostbar. For det andre må algoritmen bygge på en omfattende tekstmasse før den er robust nok til å kunne gi gode prediksjoner. Prediksjonen kan bli upresis hvis den mangler data om forhold som kan være viktig for saken.

⁴⁸ Det engelske uttrykket «Due dilligence».

⁴⁹ Surden, H (2014) s. 105-107.

Et tredje problem er generalisering når materialet er så begrenset at det ikke tillater generalisering. Dette er et annet ord for bias, altså at man forutsetter at datasettet sier alt om verden når det er mulighet for at det ikke gjør det. Vi kan gå tilbake til spamfiltereksempelet fra avsnitt 1.2. Gitt at alle e-postene fra datasettet man har brukt er fra Hviterussland. Da kan modellen anta at spam fra et annet land ikke er spam, og derfor komme med feilslutninger dersom spam sendes fra Frankrike.

Et fjerde problem Surden trekker frem, er at selv om en skulle bygge på fritekst-baserte algoritmer, er det ikke sikkert teksten sier alt om jussen. I Norge er denne problematikken kjent gjennom kamuflasjetesen i rettskildelæren; nemlig at dommere dekker over sin egentlige verdiargumentasjon med en rettslig begrunnelse.⁵⁰ En annen side av dette er at dommere i prinsippet står fritt til å dømme slik de ønsker uten henvisning til tidligere praksis som modellen forutsetter.

5.1.2 Nye analyseverktøy

Ashley mener at dersom man klarer å kombinere digitaliseringen av juridiske tekster, enten lover eller domsavgjørelser, med maskinlæringsteknologi, vil fremtidens applikasjoner kunne gjøre helt nye oppgaver.⁵¹ Han mener at digitaliserte modeller⁵² for juridisk argumentasjon vil være en bro mellom tekstmasser med juridiske tekster og mennesker som løser realistiske juridiske spørsmål.

Ashley beskriver flere utfordringer man må løse for at hypotesen om endring blir en realitet. Den mest sentrale utfordringen knytter seg til hvorvidt det vil være nok manuelt markert treningsdata for maskinlæring og automatisk markering. Ashley har noen forslag til hvordan man kan løse dette problemet. Det første forslaget involverer crowdsourcing av markering av juridiske tekster. Han viser til Breauz og Schaub (2014) som gjorde nettopp dette ved å lage et enkelt rammeverk for markering som åpner for at også ikke-jurister kan bidra.⁵³ Ashley nevner også at flere av forsøkene jeg har gjennomgått i avsnitt 3.2, brukte relativt utrente jusstudenter som markerte tekster, for eksempel i LUIMA. Zywic og Gomex (2008) går et steg videre og mener at manuell markering av domstekst hjelper studenter å engasjere seg i vitenskapelig argumentasjon og foreslår det blir brukt som et læringsverktøy på jusstudiet.⁵⁴ Det samme gjelder det internettbaserte verktøyet Classroom Salon som gjorde at studenter i en klasse kunne

50 Bergo, K (2002) s. 31.

51 Ashley, K (2017), s. 350.

52 Det engelske rettsinformatiske begrepet »Computationalized models«.

53 Ashley, K (2017), s. 374.

54 *Op. cit.* s. 375.

lese sammen, markere sammen og svare på spørsmål sammen (Blecking, 2014).⁵⁵ I etterkant konkluderte forskere med at studentene beviselig ble bedre til å lese (Herring and Lynch, 2014).⁵⁶

Ashley skriver at fremtidige juridiske applikasjoner ikke vil kunne formulere helt nye juridiske argumenter, selv om man løser disse problemene.⁵⁷ Ashley mener likevel fremtidens juridiske applikasjoner kan være nyttige. Han mener en mer realistisk fremtidshypotese er applikasjoner som ved enkle søk kan oppdrive juridiske argumenter fra tidligere praksis som er relevante i en ny sak man står overfor. Selv om applikasjonene ikke kan formulere juridiske argumenter, vil de kunne gjøre den juridiske argumentasjonsoppgaven vesentlig enklere for mennesker.

5.1.3 Jusstudiets og rettsvitenskapens fremtid

Katz forsøker å beskrive fremtiden til jusstudiet med bruk av maskinlæring.⁵⁸ Hypotesene til Katz er for det meste generelle, men en klar anbefaling er at det juridiske utdannelsesstilbudet må endres. Resonnementet til Katz er at man ikke kan påkoste mennesker en så lang og dyr utdanning hvis den ikke er relevant for det markedet som venter den nyutdannede kandidaten. Jusstudiets fremtid vil altså speile fremtiden til rettssamfunnet for øvrig.

Katz begrunner troen på maskinlæring i rettsvitenskapen blant annet ved å vise til den kjente amerikanske juristen Oliver Wendell Holmes Jr.⁵⁹:

*«For the rational study of the law the blackletter man may be the man of the present, but the man of the future is the man of statistics and the master of economics».*⁶⁰

Holmes var i 1897 opptatt av prediksjonens sentrale funksjon i jussen, og mente allerede da at statistikk ville kunne spille en viktig rolle i rettsvitenskapen. Fra skandinavisk rettslitteratur kan man også vise til Alf Ross forutsigelsestese om gyldig rett som involverer en fremtredende statistisk komponent.⁶¹

I forsøket til Tsarapatsanis mfl (2016), som gikk ut på å forutsi avgjørelser fra EMD nevnt i avsnitt 3.2, forsøker forfatterne å plassere funnene sine i en klas-sisk rettsvitenskapelig debatt mellom rettsformalister og rettsrealister. Poenget

⁵⁵ *Ibid.*

⁵⁶ *Ibid.*

⁵⁷ *Op. cit. s. 381.*

⁵⁸ Katz, D.M (2013) s. 965-966.

⁵⁹ *Op. cit. s. 936*

⁶⁰ Holmes, O. W (1897).

⁶¹ Ross, A (1953).

mitt i denne sammenhengen er ikke å vurdere om de rettsteoretiske drøftelsene til artikkelforfatterne gir gode eller dårlige svar. Poenget mitt er å illustrere tendensen om at stadig flere som opererer i skjæringspunktet mellom jus og maskinlæring ønsker å bruke de nye mulighetene som oppstår til å si noe om de klassiske tekstene fra rettsvitenskapen som har lagt grunnlaget for hvordan vi forstår retten.

Et eksempel som ikke er hentet fra noe artikkellitteratur, er den London-baserte oppstrartsbedriften CaseCrunch. Ved siden av å tilby tjenester knyttet til prediksjoner, har de et offisielt oppdrag, nemlig å løse spørsmålet: «*is there an objective dimension to legal knowledge?*»⁶²

Jeg mener det er grunn til å tro at flere vil benytte maskinlæring i forsøk på å utvikle nye innfallsvinkler til å analysere rettsvitenskapelige grunnlagsspørsmål.

5.2 Fremtidig bruk av maskinlæring i Norge

Dette avsnittet gjelder min personlige hypotese om i hvilken grad, og på hvilken måte, maskinlæringsverktøy vil bli anvendt på det juridiske området i fremtiden.

Hypotesen er basert på inntrykket fra litteraturen som er gjennomgått i forbindelse med studien, personlige samtaler med aktører i det norske rettssamfunnet og mitt personlige inntrykk av den norske juriststanden.

5.2.1 Det norske rettssamfunnets potensielle digitale infrastruktur

Som gjennomgangen av selskapene Ravel Law og Lex Machina i avsnitt 4.1 antyder, er det verdifullt med tilgang til flest mulig digitale rettsavgjørelser. Arbeidet i Norge med Lovdata og Rettsdata har gjort at de fleste avgjørelser av betydning allerede er tilgjengelig digitalt. Dette gjelder særlig avgjørelser fra Høyesterett. Dessuten er elektroniske databaser som Lovdata og Rettsdata inkorporert som en vanlig del av en jurists arbeidshverdag; advokater så vel som andre jurister. Initiativet med heldigital eksamen ved Det juridiske fakultet ved Universitetet i Oslo gjør antakelig at troen på digitale hjelpemidler bare vil øke med fremtidens jurister.

Et annet interessant aspekt med norske rettsavgjørelser i denne sammenhengen er prejudikatsfunksjonen til Høyesterett. Høyesterettsavgjørelser er på de aller

62 Official Mission – Case-crunch.com (2018).

fleste norske rettsområder de eneste avgjørelsene med rettskildemessig verdi ut over den konkrete sak. Dette gjør at de fleste undersøkelser av norsk rett kan ta utgangspunkt i en enhetlig samling av tekster som følger samme struktur og kan hentes fra samme sted. De fleste undersøkelser av amerikansk rett vil til sammenligning ikke kunne ta utgangspunkt i avgjørelser fra den amerikanske føderale høyesterett. Denne domstolen løser kun et fåtall juridiske problemer knyttet til føderale rettsspørsmål.

Dette er forhold som i utgangspunktet gjør norske rettsforhold godt disponert for bruk av maskinlæring.

Man kan se for seg et forsøk som ligner på SMILE plus IBP som undersøker alle avgjørelsene på et velavgrenset rettsområde med grunnlag i rettspraksis fra Høyesterett. Der Ashley og Bruninghaus tok utgangspunkt i forretningshemmeligheter, kunne et norsk forsøk eksempelvis undersøke alle avgjørelsene som omhandlet negativ rettskraft etter tvl. § 19-15 (3) og gamle tvistemålsloven § 163. Siden professor Anne Robberstad allerede har forsøkt å gjennomføre en slik systematisk gjennomgang i sin monografi om emnet⁶³, har man både veiledning til å gjøre ekspertvurderinger og et godt sammenligningsmateriale.

En sentral utfordring når man skal undersøke norsk jus med maskinlæringsverktøy, er det norske språket. Som det fremgår av litteraturgjennomgangen er alle forsøkene gjennomført med en engelsk tekstmasse. Selv i de engelskspråklige forsøkene har det oppstått problemer knyttet til engelsk juridisk terminologi. Forsøk i Norge måtte først håndtert problemet knyttet til norsk språk, der nest norsk-juridisk språk. Ettersom utviklere på andre områder i Norge har klart å bruke språkteknologi på norsk eksemplifisert gjennom kommersiell bruk av chatboter, er det grunn til å tro at det først og fremst er de norsk-juridiske uttrykk som byr på problemer.

Nøkkelen til å løse dette problemet er å finne tilstrekkelig datagrunnlag bestående av norsk-juridisk tekst. Eksempler på norsk-juridisk språk kan være bruk av uttrykk som «herunder», «rettskilde», «avveining» for å nevne noen eksempler på ord som ikke er mye brukt i dagligtalen. Norske avgjørelser fra alle domstolsinstanser må blant bli brukt for å trene opp algoritmen til å forstå slike uttrykk. En annen kilde kunne vært elektronisk lagrede oppgavebesvarelser fra studenter. Hvorvidt man kan klarer å bruke språkteknologi effektivt på norsk- juridisk språk er for øyeblikket noe usikkert.

63 Robberstad, A (2006).

Med unntak av usikkerheten som knytter seg til bruk av norsk-juridisk språk, har det norske rettssamfunnet trekk som gjør det godt disponert for hjelp med maskinlæring. Verktøy utviklet på grunnlag av metadatabasert maskinlæring som for eksempel risikoanalyseverktøy nevnt i kapittel 4 vil derimot ikke være avhengig av norsk språk eller juridisk-norsk språk. Det avgjørende for disse verktøyene er at det finnes tilstrekkelig med tilgjengelige statistiske data som muligens allerede finnes. Eksempler er blant annet kriminalstatistikk. Det spørs om kanskje ikke dette er en del av den første utviklingen i skjæringspunktet maskinlæring og jus man vil se i Norge.

5.2.2 En konservativ norsk juriststand?

Utsagn om at norske jurister er konservative dukker ofte opp når det snakk om teknologisk utvikling og jus. Jeg skal ikke gjøre en sosiologisk undersøkelse, men ønsker å påpeke noen forhold av betydning.

At det tar lang veldig lang tid å implementere lyd- og videoutstyr i domstolene, til tross for sterk politisk vilje, er et argument som støtter opp om beskyldningen om en konservativ norsk juriststand. Blant norske advokater kan dessuten både advokatmonopolet og partnerstrukturen i firmaene vitne om at tradisjonelle tanker fortsatt holdes i hevd.

Selv om man skulle mene at den norske jurisstanden er konservativ, er det likevel mulig å skimte en vilje til endring. På Universitetet i Oslo innføres heldigital eksamen og blant norske advokater avholdes det jevnlig konferanser for å diskutere fremtidens hjelpeverktøy. Det dukker opp nye selskaper som tilbyr automatiserte juridiske tjenester, om enn ikke ennå basert på maskinlæringsteknologi.⁶⁴ Norske advokatfirmaer abonnerer i stor utstrekning på engelske E-discovery-tjenester, og noen advokatfirmaer har til og med forsøkt å lage egne verktøy tilpasset norske forhold.

Det viktigste poenget er kanskje at utviklingen av juridiske verktøy basert på maskinlæring er antatt å være vesentlig kostnadsbesparende for advokater. Utviklingen av modeller med maskinlæring vil i startfasen være betydelig ressurskrevende. Men på den andre siden kan man risikere å ende opp med å miste livsgrunnlaget dersom man ikke gjør investeringen. Selv om advokatstanden skulle være konservativ, kan det være så vesentlig kostnadsbesparende å legge om virksomheten digitalt, at man ikke har noe valg.

64 Lawbotics (2018).

5.2.3 Særtilfelle: Den norske rettskildelæren

Jeg vil avslutte delen om fremtidig bruk av maskinlæring på jus i Norge med å skissere et mulig forsøk knyttet til den norske rettskildelæren. Dette er et eksempel på hvordan man kan anvende teknologien som er gjennomgått i litteraturen i kapittel 2, 3 og 4 på norske forhold. Det er også et eksempel på å bruke maskinlæring til å si noe om rettsvitenskapen som nevnt i avsnitt 5.1.3. Dessuten kan eksempelet beskrive nærmere hvordan norsk juridisk infrastruktur er godt disponert for maskinlæringsverktøy nevnt i innværende del over.

Helt siden den danske professoren Alf Ross fikk innpass i norsk juridisk teori, har norsk rettskildeanalyse dreiet seg om å undersøke hvordan Høyesterett bruker rettskildene for å kunne si noe om rettskildenes interne plassering. Rettskildenes interne plassering måles tradisjonelt etter en relevans-, slutning- og vekt-modell der man undersøker om Høyesterett bruker en rettskilde, hvordan Høyesterett slutter fra en rettskilde, og hvor mye vekt Høyesterett tillegger en rettskilde.

Dette er en (norm-)deskriptiv analyse som i dag gjøres manuelt av de teoretikerne som driver med emnet. Det har ikke vært noen forsøk på å gjøre en systematisk analyse av rettskildebruk i alle saker fra Høyesterett. Jeg mener dette ville åpnet for verdifulle historiske og nåtidige analyser som ville beriket disiplinen. Problemet er ikke at det er umulig å analysere all rettskildebruk fra Høyesterett, men at det er veldig tidkrevende. Min hypotese er at man ved bruk av slik maskinlæringsteknologi som er gjennomgått i denne litteraturstudien, kan gjøre en slik systematisk analyse av all praksis fra Høyesterett med en rimelig ressursbruk.

Den mest realistiske fremgangsmåten for å oppnå en slik fullstendig analyse, vil antagelig være å utvikle et verktøy til å analysere ett avgrenset rettsområde, for så å la andre bruke verktøyet på andre områder. Altså en slags ekspertbasert crowdsourcing. Dette kunne vært gjort på følgende måte:

Steg 1: Finne et spesifikt område innen jussen med omfattende rettspraksis, for eksempel strafferett. Målet med undersøkelsen blir å undersøke rettskildebruken. I strafferetten utkrystalliseres en spesiell rettskildemessig problemstilling om legalitetsprinsippet; setter Høyesterett loven foran andre rettskilder?

Steg 2: Finne alle avgjørelser fra Høyesterett om strafferett fra et gitt årstall og frem til i dag – avhengig av ønskene for analysene.

Steg 3: Bruke et program man har utviklet til automatisk/semi-automatisk å identifisere: Rettskildemarkering, rettskildeslutninger, rettskildekonflikter og vektingen mellom dem.

Dette programmet kan utvikles med bakgrunn i teknologien nevnt i LUIMA i avsnitt 3.2. Ekspertene kan markere en del rettsavgjørelser for å danne mønstre. I stedet for å markere setninger knyttet til tvistens spørsmål og eventuelle bevis-spørsmål som i LUIMA, skal ekspertene kun markere opp rettskildefaktorene og deres interne forhold (slutning, konflikt, og vekt). Markeringen vil så implementeres i et program som bruker maskinlæringsalgoritmer som kan anvendes på nye saker for å teste og lære. Programmet skal da etter hvert klare å identifisere de nevnte kategoriene, og gjøre den påfølgende analysen vesentlig enklere, mer presis og mindre tidkrevende.

LUIMA har vært effektiv og det på mye mer komplekse forhold enn rettskilde-markering. Samtidig kan det dukke opp særspørsmål knyttet til rettskildefaktoren reelle hensyn og dens uttrykk i avgjørelsestekster. I prinsippet skal programmet kunne identifisere alternative uttrykksmåter så lenge det nevnes i teksten. Et problem for utvikleren av programmet vil antakelig være de situasjonene der reelle hensyn som har vært med i vurderingen, ikke engang er nevnt i teksten man analyserer (kamouflasjetesen). Samtidig kan man hevde at dette ikke vil være et så stort problem all den tid verktøyet har gjort den øvrige jobben mindre tidkrevende. Den siste manuelle reelle hensyn-analysen vil derfor ikke utelukke mulighetene for fullstendige rettskildemessige analyser på et rettsområde.

Poenget med dette tenkte forsøket er å vise at bruk av en maskinlæringsbasert fremgangsmåte potensielt kan gjøre analyse av all rettskildebruk fra Høyesterett mye enklere, og med det potensielt endre rettskildedisiplinens praktiske og teoretiske forutsetninger fullstendig.

5.3 Ønskede og uønskede konsekvenser av å anvende maskinlærings på det juridiske området

Avslutningsvis kan det være nyttig med en kort refleksjon over det som er gjennomgått i denne rapporten. Jeg ønsker å legge vekt på at teknologien slik den er i dag, er avhengig av menneskelig styring. Det er dermed uinteressant å vurdere konsekvensen av et futuristisk AI/robot-samfunn. Det er for eksempel mange viktige spørsmål som aktualiseres dersom dommere erstattes fullstendig med programmer basert på kunstig intelligens. Dette er i skrivende stund veldig van-

skelig å forestille seg, dels fordi slik teknologi ikke eksisterer, og dels på grunn av de samfunnsomveltninger en slik situasjon ville forutsatt.

En rapport fra McKinsey Global Institute fra 2017 om fremtidens arbeidsmarked med automatisering antar at 23 % av juristers jobber vil forsvinne.⁶⁵ Det er grunn til å tro at en del arbeid som tidligere ble gjort av uerfarne jurister, enten saksbehandlere i offentlig forvaltning eller advokatfullmektiger i advokatfirmaer, snart kan gjøres av en eller to personer som kontrollerer et automatisert verktøy. Man antar også at noen av jobbene kan gjøres av ikke-jurister.⁶⁶

De jobbene som skapes som følge av dette vil blant annet knytte seg til problemområdet maskinlæring og jus, et tilbud som ikke finnes ved noen juridiske fakultet i Norge slik det er nå. Dessuten mener flere kommentatorer at arbeidsoppgavene i større grad vil spisses mot oppgaver som innebærer mellommenneskelige relasjoner som for eksempel forhandlingsvirksomhet.⁶⁷

Jeg syns det er vanskelig å rubrisere et annerledes og potensielt tøffere arbeidsmarked for nyutdannede jurister som en uønsket konsekvens. Det er en bransje i utvikling, på lik linje med andre bransjer som har gjennomgått en automatiseringsutvikling. Forhåpentligvis vil de juridiske fakultetene i tråd med utviklingen tilby fagtilbud som knytter IT aktivt inn i jussen for å speile arbeidsmarkedet slik Katz ønsker.

Hvis man forutsetter en utstrakt bruk av maskinlæring innen jus kan man peke på problemer dels knyttet til teknologien og dels knyttet til bruken av teknologien. Teknologien i seg selv er basert på algoritmer som enten er trent opp av mennesker eller semi-selvlærende. Algoritmen er som nevnt helt avhengig av datagrunnlaget det er trent på. I denne sammenheng kan man peke på eksempler fra USA der slike algoritmer knyttet til risikoanalyse omtalt i kapittel 4 er beskyldt for å være ubalansert, særlig når det gjelder rase.

Propublica har gjort en gjennomgang av risikoalgoritmene som ble brukt i staten Florida.⁶⁸ Propublica konkluderer med at algoritmene er svake og kun kommer til riktig prediksjon i 20% av tilfellene. Propublica går også gjennom et enkelttilfelle der de sammenligner vurderingen av en hvit mann og en svart kvinne. Algoritmen viste seg å ta feil ettersom verktøyet plasserte lavere risiko hos den hvite mannen som utførte et grovt innbrudd rett etterpå. Det er vanske-

65 Chui, M mfl. (2017).

66 Lohr, S (2017).

67 *Ibid.*

68 Angwin mfl. (2016).

lig å gi enkelthistorier stor vekt, men det fungerer som illustrasjon på at algoritmene ikke er feilfrie.

Hvis man ikke etterprøver algoritmene tilstrekkelig eller har et rammeverk som kan gi innsyn i algoritmene, kan det dessuten oppleves urettferdig for den personen algoritmen skal bestemme over. Sånn sett vil det kunne være interessant å reise spørsmålet om hvordan innsynsspørsmålet som ble behandlet i Wisconsin nevnt i kapittel 4, hadde blitt behandlet i norske domstoler.

Det er grunn til å tro at de juridiske verktøyene som lages på grunnlag av maskinlæring vil gi grobunn for en demokratisering av retten. Det finnes eksempler på apper som utvikles i Storbritannia som gjør det enklere for alle å gå til søksmål.⁶⁹ En slik demokratisering vil de aller fleste anse som ønsket. I Norge har det blitt rettet kritikk mot både høye advokatsalærer og unødvendig vanskelige prosesshandlinger. Dette lå blant annet bak innføringen av ny tvistelov. Disse bekymringene vil utviklingen av verktøy med maskinlæring bidra til å redusere.

69 Winick (2017).

6 Kildeliste

Angwin, Julia., m.fl. (2016). Machine Bias. Prorepublica. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (besøkt 22. januar 2018).

Ashley, Kevin D. Artificial Intelligence and Legal Analytics. Cambridge, UK: Cambridge University Press. 2017.

Ashley, Kevin D., og Bruninghaus, Stefanie. Automatically classifying case texts and predicting outcomes. Artificial Intelligence and Law, 17 (2), 125-65. (2009).

Bergo, Knut. En realistisk rettskildelære – et urealistisk prosjekt. Retfærd 2002/4 nr 99. (2002) Nettlenke: http://retfaerd.org/wp-content/uploads/2014/06/Retfaerd_99_2002_4_s19_32.pdf (Sist besøkt: 19. februar 2018)

Bommarito, Michael., Katz, Daniel M., og Blackman, Josh. A general approach for predicting the behaviour of the Supreme Court of the United States. PLoS ONE, 12 (4), 1-18 (2017).

Case Crunch. Official mission (2018) – <https://www.case-crunch.com> (besøkt: 22. januar 2018).

Chui, Michael., Manyika, James., Miremadi, Mehdi., Bughin, Jacques., George, Katy., Wilmott, Paul., Dewhurst, Martin. A Future That Works: Automation, Employment, And Productivity. McKinsey Global Institute (2017) -https://www.mckinsey.com/~media/McKinsey/Global%20Themes/Digital%20Disruption/Harnessing%20automation%20for%20a%20future%20that%20works/MGI-A-future-that-works_Full-report.ashx (besøkt: 28. januar 2018).

Dale, Robert. Handbook of Natural Language Processing. Marcel Dekker, Inc. New York. 2000.

Francesconi, Enrico., Montemagni, Simonetta., Peters, Wim, og Tiscornia, Daniela. Integrating a bottom-up and top-down methodology for building semantic resources for the multilingual legal domain. Sider 95-121 fra: Francesconi, Enrico, Montemagni, Simonetta, Peters, Wim og Tiscornia, Daniela, Semantic Processing of Legal Texts. Berlin, Heidelberg: Springer-Verlag. 2010.

Grabmair, Matthias., Ashley, Kevin D., Chen, Ran mfl. Introducing LUIIMA: an experiment in legal conceptual retrieval of vaccine injury decisions using a UIMA type system and tools. Sider 1-10 fra: Proceedings of the 15th International Conference on Artificial Intelligence and Law. ICAIL 2015. New York, NY: ACM (2015).

Graver, Hans Petter, og Bakken, Jonas. Juss, språk og retorikk. Fagbokforlaget. 2013.

Grossman, Maura R, Cormack, Gordon V. Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient than Exhaustive Manual Review. Richmond Journal of Law and Technology (Vol. XVII, Issue 3) (2011).

Harbert, Tad. The Law Machine. IEEE Spectrum, 50 (11), 31-54 (2013).

Holmes, Oliver Wendell Jr. The Path of the Law. Harvard Law Review Volume 10 s. 457-478 (1897).

Katz, Daniel M. Quantitative Legal Prediction – or – How I learned to stop worrying and start preparing for the Data-Driven Future of the Legal Services Industry. Emory Law Journal, 62 (4), 909-66 (2013).

Kira Systems. (2018), <https://www.kirasystems.com/how-it-works/contract-analysis> (besøkt 12. februar 2018).

Lohr, Steve. (2017). A.I. Is Doing Legal Work. But It Wont Replace Lawyers, Yet. New York Times – <https://www.nytimes.com/2017/03/19/technology/lawyers-artificial-intelligence.html> (besøkt 27. januar 2018).

Lawbotics. (2018), <http://lawbotics.io/index.html> (besøkt 12 februar 2018).

Luminance. (2018), <https://www.luminance.com> (besøkt 12. februar 2018).

McShane, BB, Watson, OP, Baker, T, Griffiths, SJ. Predicting Securities fraud settlements and amounts: a hierarchial Bayesian model of federal securities class action lawsuits. Journal of Empirical Legal Studies, 9 (3), 482-510 (2012).

Mochales, Raquel og Moens, Marie-Francine. Argumentation mining. Artificial Intelligence and Law, 19 (1), 1-22 (2011).

Mozina, Martin., Bratko, I., Zabkar, J. Argument Based Machine Learning. Argumentation in Artificial Intelligence, 463-482 (2009).

Neota Logic. (2018), <https://www.neotalogic.com/solution/intelligent-document-automation/> (besøkt 12. februar 2018).

Ravel Law. Ravel: Data Driven Research (2018), www.ravellaw.com (besøkt 22. januar 2018).

Robberstad, Anne. Rettskraft. Fagbokforlaget. 2006.

Ross, Alf. Om ret og rætfærdighed. Nyt nordisk forlag. 1953.

Ross Intelligence. Ross: Your Brand New Super Intelligent Attorney (2018), www.rossintelligence.com/ (besøkt: 22. januar 2018).

Savelka, Jaromir., Trivedi, Gaurav, og Ashley, Kevin D. Applying an interactive machine learning approach to statutory analysis. Pages 101-10 fra: Legal Knowledge and Information Systems: JURIX 2015: The Twenty-Eights Annual Conference. Amsterdam: IOS Press (2015).

Savelka, Jaromir og Grabmair, Matthias. (Brief) Introduction to (Selected Aspects of) Natural Language Processing and Machine Learning, Tutorial at Workshop on Automated Detection, Extraction and Analysis of Semantic Information in Legal Texts (2015).

Smith, Mitch. (2016). In Wisconsin, a Backlash Against Using Data to Foretell Defendants Future. New York Times – <https://www.nytimes.com/2016/06/23/us/backlash-in-wisconsin-against-using-data-to-foretell-defendants-futures.html> (besøkt 22. januar 2018).

Surden, Harry. Machine Learning and Law. Washington Law Review, 89 (1), 87-115 (2014).

Tsarapatsanis, D., Prejotiu-Pietro, D., Lampos, V. og Alteras, N. (2016). Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective. PeerJ Computer Science, 2016, 2:e93 <https://doi.org/10.7717/peerj-cs.93>.

Winick, Erin. (2017). This Chatbot Will Help You Sue Anyone. MIT Technology Review – <https://www.technologyreview.com/s/609418/this-chatbot-will-help-you-sue-anyone/> (besøkt 22. januar 2018).

Zelevnikow, John. (2017). Can Artificial Intelligence and Online Dispute Resolution enhance efficiency and effectiveness in Courts? International Journal for Court Administration, 8 (2), 30-45 (2017).