



Vejen til vækst og innovation går gennem AI – hvis vi husker dataetikken

Birgitte Kofod Olsen, partner, ph.d., i Carve Consulting, forperson og medstifter af DataEthics.eu, bko@carve.dk

Mads Krogh Munch Nielsen, partner, Carve Consulting

Danmark er et af de mest digitaliserede lande i Europa. Vi har årtiers erfaring med at effektivisere arbejds- og produktionsprocesser og har data om danskerne, der bygger på en århundredlang tradition for at føre registre. Vi er også det første land i verden, der har indført en forpligtelse for virksomheder til at redegøre for deres arbejde med dataetik i deres årsrapport.

Samtidig er AI et strategisk indsatsområde i EU, hvor det ses som driveren til økonomisk vækst – og ligefrem en ny teknologisk og økonomisk æra med forventninger om forbedring af produktiviteten med 40 procent. Også her er dataetik højt på dagsordenen og spiller en central rolle i strategien for et Europa, der er klar til den digitale tidsalder.

Begge dele bidrager til, at vi nu ser både offentlige myndigheder og private virksomheder kaste sig ud i at udforske løsninger og services, der er baseret på kunstig intelligens (AI) og machine learning (ML).

Denne artikel er opdelt i tre dele. Først gives en række eksempler på, hvor AI afprøves og er ved at vinde ind. Dernæst gives et overblik over de dataetiske overvejelser, som pt. drøftes, og hvor man finder anbefalinger og lovgivning. Endelig gives konkrete eksempler på, hvordan forskellige anvendelser af AI kræver inddragelse af dataetiske overvejelser.

Årsregnskabslovens § 99 d blev indføjet i loven i 2020. Paragraffen bestemmer, at store virksomheder, som har en politik for dataetik, skal supplere ledelsesberetningen med en redegørelse for virksomhedens politik.

Redegørelsen skal også indeholde oplysninger om virksomhedens arbejde med dataetiske spørgsmål.

Har virksomheden ikke en politik for dataetik, skal ledelsesberetningen indeholde en redegørelse med forklaring på baggrunden herfor.

AI kræver strategiske overvejelser

Kunstig intelligens, Artificial Intelligence (AI), og selvlærende modeller, kendt som Machine Learning (ML) er ikke længere et nyt begreb, og det anvendes i stadig flere sammenhænge. Alligevel er det for mange lidt af en "black box", og rejser spørgsmål som: Hvad kan det anvendes til? Hvilke konsekvenser har det for os som borgere og forbrugere? Og for vores samfund? Spørgsmål som disse har ført til krav om etiske og da-

tæetiske overvejelser i tilknytning til udvikling af den kunstige intelligens, når den tages i brug, og når man ser resultaterne af dens anvendelse.

AI vinder ind mange steder

AI Signaturprojekter

Regeringen har i løbet af 2020 sammen med KL og Danske Regioner søsat 15 signaturprojekter, der skal afprøve potentialet i AI i forhold til at løfte kvaliteten og kapaciteten i den offentlige sektor.

Signaturprojekterne er rettet både mod borgernære opgaver og administrative funktioner. Blandt de borgerrettede projekter ses AI-støtte til at matche jobsøgende borgere med relevante ledige jobs, til visitering af borgere til rengøring og rehabilitering, samt til forbedring af indsatser på sundhedsområdet inden for diagnosticering, behandling, forebyggelse og prædiktation af sygdomme som kræft, KOL og hjertesvigt.

De administrative processer får også hjælp fra AI til at opnå korrekt og ensartet sagsbehandling, effektiv fordeling og journalisering af mails fra borgerne og til at skaffe kortere svartid på byggeansøgninger.

Optimering og bæredygtighed

Hos de danske virksomheder er der fokus på at optimere produktion, transport og distribution gennem AI-løsninger, der fx kan forbedre udnyttelsen af energi og brændstof eller reducere tid og udgifter med dynamisk ruteplanlægning.

Servicevirksomhederne har for alvor fået øjnene op for potentialet i dataanalyse for at opnå større indsigt i kundernes ønsker, vaner og præferencer. Det kan udnyttes til forbedring af deres kundeservice, fx ved brug af chatbots, og kreditvurdering af forbrugere. I den finansielle sektor er der derudover eksempler på AI-understøttelse af risikovurderinger, efterforskning af svindel og udvikling af bæredygtige investeringer.

Advokatbranchen og andre rådgivningsvirksomheder tager også i stadig større omfang AI i brug til informationsindsamling og beslutningsstøtte.

I landbruget sparker AI nye tanker i gang om en digital andelsbevægelse, bæredygtige produktionsformer og øget fødevarerikkerhed. Potentialet i fiskeriet for at tilpasse fangsten til den daglige efterspørgsel og samtidig sikre bæredygtige fiskebestande ved hjælp af AI er også så småt ved at komme i fokus.

Dataetikken indhold

Der foreligger endnu ingen autoritative bud på, hvad dataetik er. Imidlertid er der mange anbefalinger og forslag til dataetiske principper og processer, der skal medvirke til at integrere dataetik i udvikling og drift af AI-baserede løsninger i såvel offentlig som privat sektor

Mennesket i centrum

Omdrejningspunktet i dataetikken i både EU og Danmark er, at menneskets værdighed og interesser er i centrum. Dette menneske-centriske perspektiv rammer kernen i den beskyttelse af grundrettighederne, der er indlejret i EU-traktaten og i menneskerettighederne i de internationale konventioner. Her har mennesket forrang i forhold til statens interesse i systemer og effektivisering.

Som koncept favner dataetik principper om værdighed, selvbestemmelse, ligestilling og gennemsigtighed, men også lovgivning. Et væsentligt element er den retlige beskyttelse af rettigheder og frihedsrettigheder, herunder persondatabeskyttelse og respekt for privatliv, således som det er stadfæstet i EU-traktaten og nationale forfatninger. Et tredje element, som tænkes ind i dataetikken, er robust informationssikkerhed. Denne del er knyttet til den EU-retlige regulering af netsikkerhed og cybersikkerhed. Klima- og miljømæssig bæredygtighed, samt social og økonomisk bæredygtighed, bliver også i stigende omfang tillagt betydning som dataetiske elementer. Dataetikken og de krav, den giver anledning til, er altså forankret i både lovgivning og etik samt et praktisk krav om robust sikkerhed.



EU HLEG AI, Ethics Guidelines for Trustworthy AI (April 2019).

DataEthics: Dataetiske principper

MENNESKET I CENTRUM

Menneskets interesser har altid forrang for institutionelle og kommercielle interesser. Individet har den primære gavn af databehandlingen.

INDIVIDUEL DATAKONTROL

Mennesket har individuel datakontrol og handlekraft. Det er det enkelte menneske, der har den primære kontrol over, hvad dets data bruges til.

GENNEMSKUELIGHED

Databehandling og automatiserede beslutninger skal give mening for det enkelte menneske. De skal være transparente og skal kunne forklares. Risici, sociale, etiske og samfundsmæssige konsekvenser skal stå klart.

ANSVARLIGHED

Ansvarlighed er en organisations bevidste, saglige og systematiske brug og beskyttelse af persondata. Bæredygtig persondatabehandling er indlejret i hele organisationen og sikrer etisk ansvarlighed.

LIGE VÆRDIGHED

Demokratisk databehandling bevarer magtfordelingen i samfundet. Der tages særlige hensyn til sårbare mennesker og deres risiko for diskrimination, stigmatisering og bias.

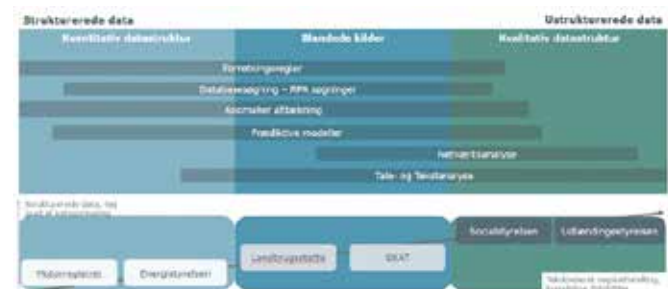
<https://dataethics.eu/da/dataetiske-principper/>

Når råmateriale er data

Data er råmateriale i den algoritme-baserede økonomi, som AI og ML er med til at drive frem. Det betyder, at lovgivningen om persondatabeskyttelse og udveksling af data skal tages med betragtning, når vi

introducerer kunstig intelligens. Som led i den dataetiske analyse er det derfor vigtigt at afdække de anvendte datatyper og den relevante lovgivning, der knytter sig til behandlingen af disse.

I AI og ML-baserede løsninger anvendes data i både struktureret form og ustruktureret form, altså fra registre, databaser og andre kategoriserede data, og til kvalitative strukturer som fx fritekstfelter, telefonsamtaler eller tekstbaserede dokumenter. Mellem disse yderpunkter ligger de blandede datasæt hos f.eks. sociale medier, som både er egnet til søgninger, til afdækning af anomalier, til prædiktioner og til tekstanalyse. Der ligger en kompleksitet i, at man anvender forskellige analytiske metoder alt efter karakteren af data. Hvis vi som eksempel betragter offentlige tilsyn, så vil de valgte metoder være meget forskellige afhængigt af, om tilsynet baseres på et kendt register eller gennemlæsningen af et stort antal tekstbaserede rapporter, journaler mv. Nedenstående figur illustrerer dette.



I datasættene gemmer sig mange forskellige former for data: persondata, produktionsdata, geografiske data, finansielle data, metadata mv. Behandlingen af disse typer af data er undergivet lovgivning, både dansk og EU-retlig regulering.

Persondata og andre data

To overordnede forordninger fra EU sætter rammerne for databehandlingen i en AI-sammenhæng, nemlig persondataforordningen GDPR (General Data Protection Regulation, 2016), som regulerer behandlingen af personoplysninger. Desuden er der forordningen om fri udveksling af andre data end personoplysninger, FFD (Regulation on the Free Flow of non-personal data, 2019).

FFD er særlig relevant i forhold til blandede datasæt, der består af både persondata og andre datatyper. Hvis de data, der ikke er personoplysninger, kan kobles til en person, som på den måde gøres identificerbar, skal oplysningerne betragtes som personoplysninger, og altså behandles under GDPR.

Det betyder, at en kvalitetskontrolrapport for en produktionslinje skal betragtes som indeholdende personoplysninger, hvis det er muligt at sætte oplysningerne i rapporten i forbindelse med specifikke medarbejdere, fx dem der står for indstillingen af produktionsparametre.

Det samme kan være tilfældet med CRM-systemer. De indeholder typisk både personoplysninger og andre data i form af kundernes købshistorik samt rapporter og analyser af de samlede salgsaktiviteter. Hvis oplysningerne anonymiseres i rapporteringen og dermed frakobles identificerbare personer, kan denne del af datasættet dog behandles efter reglerne i FDD.

Eksempler: Kunstig intelligens i praksis

Kunstig intelligens anvendes i dag til at understøtte beslutninger. AI/ML bruges her til at indsamle den information, der skal til for at træffe en

beslutning på et givet område. Informationsindsamlingen kan beskrives i tre former:

- Assisteret informationsindsamling. Anvendelse af avanceret analyse til at raffinere dataindsamling fra infrastruktur-teknologier, såsom Cloud, Datalakes mv., til brug for understøtning af manuelle beslutninger.
- Augmented informationsindsamling. Machine Learning modeller lægges oven på eksisterende informationssystemer til at understøtte og simplificere beslutninger, taget af medarbejdere.
- Autonom informationsindsamling. Her er processerne fuldt digitaliserede og automatiserede fra indsamling af information til levering af information, som maskiner, chatbots og systemer kan handle ud fra.

I det følgende gives eksempler på, hvilke dataetiske overvejelser, som kan knyttes til de tre former for informationsindsamling, der understøtter beslutninger.

Assisteret informationsindsamling

Der er store besparelser i tid og omkostninger at hente i en selv-lærende tekstanalysemodel, som kan gennemlæse alle de skriftlige dokumenter og oplysninger, der findes i et kontrolsystem eller et CRM-system, så man kan afdække mønstre og sammenligne dem over tid. Måske kan modellen også oversætte dokumenterne, for så at afslutte med at formulere et beslutningsreferat eller opstille en rapport med de væsentligste elementer, modellen har fundet i sagen.

Ved brug af sentimentanalyser kan tidsforbruget på en opgave reduceres væsentligt, fordi analyseprogrammet er udviklet, så det forstår sætningerne i et kundekort eller en kontrolrapport. Det kan således give støtte til at identificere de væsentlige informationer og finde relevante kontrakter, vurderinger eller hændelser og på den måde medvirke til en hurtig og oplyst beslutning.

I sentimentanalyser anvendes en sprogvektor. Hvis den er trænet på et irrelevant materiale i forhold til det område, opgaven eller beslutningen angår, vil sagen blive analyseret forkert.

Dataetiske overvejelser

Hvis den sprogvektor, der anvendes til kreditvurdering, fx er trænet på dokumenter, der knytter bopæl sammen med termen ghettoplan, så kunden tildeles en høj risikoscore allerede på grund af bopælen, kan modellen føre til, at kunden udelukkes fra en finansiel service. Beslutningen kan derfor ses som biased eller diskriminerende.

Der kræves derfor transparens i de anvendte analytiske teknologier og en systematisk og regelmæssig tilpasning og auditering af følsomheden i modellerne, så man undgår, at de fører til forkerte eller utilsigtede afgørelser.

Augmented informationsindsamling

Segmentering

Avanceret analyse kan også hjælpe på andre måder end ved at forstå teksten i de datakilder, der analyseres. Man kan fx få en selv-lærende ML model til at segmentere et sagsmateriale efter kompleksitet og det behov for kompetencer, der er nødvendigt for at kunne træffe en beslutning.

På den måde vil materialet blive fordelt til de rette medarbejdere med det samme, og dermed bliver der sparet ressourcer, både ved at færre opgaver skal gå i tilbageløb, og ved at de erfarne kræfter kan

koncentrere sig om avancerede opgaver.

Når en tekstanalytisk motor går i gang med at finde relaterede sager, foretager den også nogle fravalg. Den beslutter hvilke elementer, der er relevante at fremhæve som karakteristiske, og så matcher den sagen op mod sager med samme karakteristika.

Dataetiske overvejelser

Det dataetiske problem i relation til fravalget består i risikoen for forkerte afgørelser. Det kan opstå som følge af, at der er indlejret bias i den tekstforståelse, som modellen bygger på. De sager, som er relevante for en bestemt gruppe borgere eller kunder, fx personer med handicap, bliver måske fravalgt pga mangelfuld tekstanalyse. Brugen af modellen bidrager dermed til udelukkelse eller stigmatisering af en person, fx i adgang til job eller til en offentlig ydelse. ML-baserede modeller kan dog også ses som et væsentligt bidrag til at skabe større gennemskuelighed om baggrunden for sagens udfald og dens konsekvenser for den berørte borger, men det kræver, at modellen er neutral i sin bias.

Prædiktive modeller

Prædiktive modeller kan bruges til at prioritere mellem serviceopgaver, fx ved pensionsrådgivning at prioritere personlig rådgivning til de kunder, der har størst behov, fremfor dem der er mest profitable. Disse modeller anvendes ofte parallelt med segmentering. Baseret på tidligere rådgivning kan man identificere de variable, der har med den enkelte sagstype at gøre, og på den måde 'forudsige' hvilke sager, der har mange træk til fælles med de sager, der er endt med positivt resultat for kunden eller et afslag.

Man kan anvende forskellige algoritmer til at drive de prædiktive og selv-lærende modeller. Nogle af dem giver detaljeret information om, hvilke variable, der bidrager med hvilke vægte i scoringsmodellen, mens andre ikke er selvforklarende. Hvis man vælger at gøre brug af en algoritme, der ikke i sig selv forklarer hvilke variable, der vægter noget i scoringsmodellen, kan man anvende såkaldt xAI – explainable AI. xAI giver input til hvilke egenskaber i modellerne, der bidrager til at forklare modellens resultater.

Dataetiske overvejelser

Fra et dataetisk synspunkt udgør xAI et væsentligt element i at sikre ligebeholdning og undgå bias eller fordomme om personer og deres baggrund eller kontekst. Analyser af de anvendte variable og vægtningen skal derfor – som et dataetisk krav – integreres både i udviklingsfasen og i regelmæssig test af modellens resultater. Derudover bør de også underkastes revision. I den sammenhæng stilles der også et dataetisk krav til sammensætningen af udviklings- og testteams. De bør være sammensat med tanke på mangfoldig repræsentation i forhold til uddannelse, baggrund, køn, alder, livsstil og kultur for bedst muligt at kunne forhindre bias og diskriminerende output.

Autonom informationsindsamling

Automatiseret beslutningsstøtte til gennemførelse af sagsforløb udgør et vigtigt bidrag til effektivisering af kundeforhold og arbejdsprocesser. Det er typisk de manuelle beslutninger, der koster tid i processen, og det vil derfor have høj værdi, hvis vi kan identificere beslutninger, der er tilpas simple og éntydige til, at vi kan lade digitale assistenter overtage beslutningerne.

Chatbots i kundeservice

Chatbots møder man efterhånden på mange hjemmesider. Særligt forsikringsbranchen har været hurtige til at tage chatbots i brug for at skræddersy forsikringsløsninger.

Brugen af chatbots i den direkte kunde- eller borgerservice eller som en intern støttefunktion bidrager til både hurtigere sagsbehandling og højere grad af ensartethed i serviceydelsen. De informationer, chatboten er baseret på, vil være trukket fra en samlet erfaringsbase og dermed indeholde eksempler på best practices eller særlige karakteristika.

Selvlærende modeller til procesanalyse

Avanceret processtøtte er egnet til at håndtere og styre store sagskomplekser. Støtten kan baseres på selvlærende modeller, som analyserer sagsgangen. Når sagerne gennemgår den nødvendige proces med godkendelser og gennemgang af forskellige delelementer, bliver progressiviteten registreret af en avanceret mønstergenkendende algoritme. Hvis der er sagsforløb, der afviger fra normalen for den type af sager, vil sagsbehandleren blive advaret direkte af beslutningsstøtten. Det anvendes fx i sagsforløb med forældelsesfrister. Her vil advarslene gå på risiko for overskridelse af en konkret frist.

Beslutningsstøtte til kontrolfunktioner

Avanceret analyse kan også indlejres i beslutningsstøtte til funktioner med kontrolformål og efterforskning af svindel. Her vil ML-algoritmen fx kunne identificere uberettigede udbetalinger i sociale programmer eller til forsikringstagere ved bl.a. at trække på data fra sociale medier. Erfaringer fra Sverige viser, at denne type kontrol afdækker problemer, som ikke bliver fundet i manuelle kontroller.

Dataetiske overvejelser

Autonome beslutningsprocesser gennemføres uden menneskelig indblanding. Derfor er det afgørende, at borgeren, kunden eller medarbejderen bliver informeret om, at en anvendt chatbot eller kontrolfunktion ikke er styret af et menneske, og at dens anbefalinger kan udfordres over for en ledelse.

Autonome beslutninger bidrager ideelt til at fjerne præferencer og

bias fra beslutninger. I praksis kræver det dog grundige analyser i udviklings- og testfasen, der afdækker om eksisterende bias bliver reproduceret af algoritmen. Et dataetisk krav er derfor, at medarbejderne ved, hvilke læringsloops modellerne har gennemgået for at eliminere bias.

Integreret dataetik

Automatiseret beslutningsstøtte bringes i anvendelse for at skabe større indsigt, mere ensartede processer, besparelser og til at frigøre medarbejdere fra rutineprægede opgaver. Automatiseret beslutningsstøtte handler derfor ikke om at erstatte medarbejderne, men om at drage fordel af de bedste egenskaber ved både mennesker og teknologi. I den indsats skal vi huske at integrere dataetik – fra udviklingsfasen til brugtagning, test og vedligehold af algoritmerne.

Kan vi nå i mål med det, vil vi kunne understøtte effektivisering og kvalitetssikring i både offentlig og privat sektor og samtidig sikre, at menneskets interesser er i centrum. Både når vi træffer beslutning om at anvende AI til beslutningsstøtte, når vi vælger teknologi og når vi anvender de resultater, som teknologien foreslår.

Læs mere her

Digitaliseringsstyrelsen, Signaturprojekter om kunstig intelligens i kommuner og regioner <https://digst.dk/strategier/kunstig-intelligens/signaturprojekter/>

Erhvervsstyrelsen, Virksomhedsguiden om dataetik

<https://virksomhedsguiden.dk/erhvervsfremme/content/lydelse/dataetik/36ef5443-d54b-4a75-b747-9409e13d8e14/>

EU Practical guidance for businesses on how to process mixed datasets, May 2019 <https://ec.europa.eu/digital-single-market/en/news/practical-guidance-businesses-how-process-mixed-datasets>

EU Whitepaper on artificial intelligence: a European approach to excellence and trust (February 2020) https://ec.europa.eu/info/files/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

Birgitte Kofod Olsen og Mads Krogh Munch Nielsen, AI kan hjælpe pressede sagsbehandlere – men der er etiske udfordringer, ING/DATA-TECH, 12. november 2019.

<https://pro.ing.dk/datatech/artikel/tekstanalyse-kan-hjaelpe-pressede-sagsbehandlere-men-der-er-etiske-udfordringer>

Forfatter: Birgitte Kofod Olsen

Birgitte har arbejdet med implementering af menneskerettigheder i over 25 år og deltog i de første diskussioner i FN i 90-erne om virksomheders samfundsansvar. Hun er tidligere CSR-chef i Tryk forsikring og rådgiver i dag virksomheder, organisationer og myndigheder om CSR-strategi og omsætning af de internationale standarder til praksis. Siden 2015 har hun hjulpet en række virksomheder med at integrere Verdensmålene i strategiske og konkrete indsatser, og udført analyser om udviklingen på området for Erhvervsstyrelsen.



Mads Krogh Munch Nielsen

Mads har arbejdet med udvikling og implementering af avancerede dataanalysemodeller siden årtusindeskiftet, og har undervejs automatiseret beslutningsstøtte på forskellige steder som hos store internationale banker, Arbetsförmedlingen i Sverige, SKAT, FDA og Scotland Yard. Han er tidligere ansvarlig for Fraud & Intelligence området i North EMEA hos SAS Institute, og har i dag fokus på at implementere, og rådgive virksomheder og myndigheder om anvendelsen af AI og avanceret analyse i digitaliseringen af arbejdsprocesser.

