

BETYDNINGEN AF BIG DATA OG MACHINE LEARNING (MASKININDLÆRING) FOR ET EFFEKTIVT ARBEJDSMILJØTILSYN I FREMTIDEN

Indledning

Tilsyn er formentlig det vigtigste politiske instrument, arbejdstilsynsmyndigheder anvender til at sikre, at virksomheder tager de nødvendige skridt til at overholde bestemmelser om sundhed og sikkerhed på arbejdspladsen. Effekten af tilsyn afhænger dog af en række forskellige faktorer. En grundlæggende faktor er den proces, hvorefter der udvælges tilsynsobjekter, dvs. virksomheder eller steder, der skal føres tilsyn med. I princippet er der mindst tre forskellige tilgange til udvælgelsen. Den første tilgang er at føre tilsyn med alle virksomheder uanset potentiel risiko, virksomhedsstørrelse, branche eller andre kriterier. Den anden tilgang går ud på at udvælge virksomheder på grundlag af stikprøver, hvor alle virksomheder, uanset karakteristika, har lige stor sandsynlighed for at blive udvalgt. Hvad angår forebyggelse og økonomi, anses begge disse metoder normalt for at være ineffektive (Blanc, 2013). Derfor udvælger de fleste arbejdstilsynsmyndigheder objekter på grundlag af den tredje tilgang, den risikobaserede tilgang. Den risikobaserede tilgang går i korte træk ud på at udvælge tilsynsobjekter på grundlag af risikoniveau.

Den risikobaserede tilgang er ganske vist et væsentligt princip for de fleste moderne arbejdstilsynsmyndigheder, men dens anvendelse i praksis er forbundet med store udfordringer. Den primære årsag til det er, at der mangler tilstrækkeligt finkornede metoder til risikoanalyse (Mischke et al., 2013). Uden egnede metoder til at gøre risikobaseret prioritering mulig er der fare for, at den risikobaserede tilgang bliver en politisk erklæring fra myndighedens side uden konkrete praktiske konsekvenser. Der er derfor behov for at udvikle metoder, der gør det muligt at udvælge højrisikovirksomheder (Weil, 2008).

De fleste arbejdstilsynsmyndigheder indsamler og lagrer enorme mængder data om deres tilsynsobjekter og tilsynsaktiviteter. Arbejdstilsynsmyndigheder råder derfor potentielt over store og hurtigt voksende datamængder – i vore dage kaldet 'big data'. Big data, i kombination med machine learning-teknologi, anvendes i stigende grad til forskellige prædiktive formål ved at lære af skjulte tendenser i dataene. For eksempel testes big data og machine learning-teknikkens prædiktive værdi på områder så forskellige som kræftprognoser og patientresultater, prognoser for antal konkurser, fremskrivninger af olieprisen, afsløring af skattesvig, prognoser for kriminalitetsudviklingen og prognoser for kursudviklingen på aktiemarkedet. Det grundlæggende spørgsmål, der behandles i dette oplæg, er dog, om brugen af big data og machine learning-teknologi til at udvælge tilsynsobjekter med høj risiko er en hensigtsmæssig kurs at slå ind på for arbejdstilsynsmyndigheder.

Risikobaseret udvælgelse

Ifølge principperne for bedste praksis inden for udmøntning af lovgivning, skitseret af Organisationen for Økonomisk Samarbejde og Udvikling (OECD, 2014), bør risikoanalyse og risikovurdering være udgangspunktet, når arbejdstilsynsmyndigheder udvælger objekter til tilsyn. Det betyder, at virksomheder bør udvælges til tilsyn på grundlag af vurderinger af sandsynligheden for og konsekvenserne af risikoelementer som ulykker, eksponering for skadelige stoffer og ulovlige arbejdsvilkår. Grundlaget for risikobaseret udvælgelse er anerkendelsen af, at det som følge af begrænsede tilsynsressourcer ikke er muligt at kontrollere alle risikoområder og alle risikoobjekter. Når det gælder arbejdstilsynsmyndigheders tilsyn med sundhed og sikkerhed på

arbejdspladsen, betyder det, at nogle problemområder skal prioriteres højere end andre. Desuden skal nogle virksomheder prioriteres højere til tilsyn end andre.

Princippet bag risikobaseret udvælgelse er ikke nyt. For snart 50 år siden påpegede Robens-komiteén i sin evaluering af det britiske system for arbejdstilsyn, at den risikobaserede tilgang (kombineret med selvregulering) er idealet, når myndighedernes tilsyn skal moderniseres (Robens, 1972). For at sikre omkostningseffektiv brug af tilsynsressourcer anbefalede Robens-komiteén, at tilsynsmyndigheden koncentrerede sine ressourcer selektivt på de mest alvorlige problemområder og prioriterede virksomheder og problemer, der var blevet identificeret gennem den systematiske analyse af alle tilgængelige data om sundhed og sikkerhed, f.eks. ulykkesstatistikker, tekniske oplysninger og tilsynsmyndighedens lokalkendskab.

Anbefalingerne i Robens-rapporten er i vid udstrækning blevet indarbejdet af arbejdstilsynsmyndigheder internationalt og i EU-medlemsstaterne. Udbredelsen af den risikobaserede tilgang betyder, at de fleste moderne arbejdstilsynsmyndigheder har taget idéen om at tage ressourcer fra lavrisikoobjekter og koncentrere flere tilsynsressourcer på objekter med de højeste risici til sig. Det kræver en form for dataanalyse. Udviklingen af analysemetoder til at identificere højrisikobrancher og risikoeksponerede arbejdstagergrupper er vidt fremskredet. Sådanne risikobaserede analyser er typisk baseret på nationale statistikker for, for eksempel, erhvervs sygdomme, arbejdsulykker og erhvervs mæssige eksponeringer. Analyserne udgør grundlaget for tilsynskampagner, strategiplaner og nationale – og internationale – prioritetsområder.

Langt mindre almindelige end de brede risikobaserede analyser er metoder, der gør det muligt at foretage prioritering på tværs af virksomheder inden for en branche. Når arbejdstilsynsmyndigheder udvælger konkrete risikoeksponerede virksomheder, er det en almindelig tilgang at tage udgangspunkt i tilsynsførendes lokalkendskab. Nogle arbejdstilsynsmyndigheder, blandt andet den danske og den svenske, har undersøgt brugbarheden af systemer, hvor risikoen rangordnes på grundlag af additive modeller. Ved at anvende additive modeller får den enkelte virksomhed tildelt risikoscorer ud fra en række virksomhedskarakteristika (f.eks. størrelse, branche, antal registrerede ulykker), som sammenlægges til en samlet score. Virksomhederne med de højeste samlede scorer prioriteres til tilsyn. Problemet ved at anvende sådanne additive modeller er dog, at deres prædiktive validitet ligger på et relativt lavt niveau, dvs. at scoren ikke er særligt godt egnet til at skille højrisikovirksomheder fra lavrisikovirksomheder.

Big data og machine learning

Processen med at prioritere mellem virksomheder er som at finde nåle i en høstak. I dette tilfælde består høstakken potentielt af hundredtusindvis af mulige tilsynsobjekter, hvor kun et bestemt antal af disse objekter er nåle, dvs. har en risiko på et uacceptabelt niveau. At finde nåle i en høstak er langt hen ad vejen det, big data og machine learning handler om.

Det vigtigste formål med machine learning-algoritmer er at tilvejebringe en statistisk model, der kan udnyttes til at udarbejde prognoser, klassificeringer, skøn og lignende opgaver. Når det for eksempel gælder prognoser på kræftområdet, har forskerne i nu mere end tre årtier brugt machine learning-algoritmer til at beregne disponering for kræft, kræfttilbagefald og kræftoverlevelse. Tematisk er prognoser på kræftområdet langt fra den risikobaserede udvælgelse af tilsynsobjekter. Begge er imidlertid eksempler på prædiktive udfordringer – eller nål i en høstak-problemer.

De to mest almindelige typer machine learning-algoritmer er supervised learning og unsupervised learning. I supervised learning består algoritmen af en afhængig variabel (f.eks. risikoniveau), som skal beregnes ud fra et sæt uafhængige variabler. Præcise beregninger kræver naturligvis en høj grad af korrelation mellem de uafhængige variabler og den afhængige variabel. I unsupervised learning er der ingen afhængig variabel at beregne, men formålet med algoritmen er at samle

dataene i grupper (f.eks. forskellige risikogrupper) efter ensartethed. I modsætning til additive modeller – for eksempel dem, der anvendes af arbejdstilsynsmyndighederne i Danmark og Sverige – forbedrer de algoritmer, der bruges inden for machine learning, deres beregninger progressivt, især gennem forsøg eller fejl. Det betyder, at maskinen lærer af tidligere succeser (korrekte beregninger) og fejl (forkerte beregninger) og forsøger at bruge denne viden til at gøre sine beregninger mere præcise ud fra den modtagne feedback.

Udnyttelse af big data og machine learning til at udvælge tilsynsobjekter

Supervised og unsupervised learning-algoritmer kræver en tilstrækkelig mængde data, når det gælder både antal observationer og antal variabler – normalt kaldet 'features'. Som allerede nævnt indsamler og lagrer de fleste arbejdstilsynsmyndigheder enorme mængder data om deres tilsynsobjekter og tilsynsaktiviteter. Dataene vedrører typisk virksomhedsspecifikke features som antal medarbejdere, virksomhedsalder, branchegruppering, antal tidligere tilsyn, resultater af tidligere tilsyn og anmeldelser af ulykker. Desuden stiger datamængden dag for dag, efterhånden som resultaterne fra nye tilsyn kommer til. I princippet bør håndteringen af udfordringen med at udvælge højrisikovirksomheder ved at udnytte big data dermed – i hvert fald umiddelbart betragtet – passe godt til machine learning-algoritmer. Alligevel har der kun været få sådanne forsøg. Der er dog nogle bemærkelsesværdige undtagelser, som alle viser, at big data og machine learning kan være særdeles relevant for arbejdstilsynsmyndigheder, når de skal løse udfordringen med at udvælge tilsynsobjekter med høj risiko.

Det første eksempel er en videnskabelig undersøgelse, der undersøgte machine learning-metoders egnethed til at forudberegne arbejdsulykker, eller mere konkret fald på gulv (Matías et al., 2008). Trods sine relativt præcise beregninger er svagheden ved denne undersøgelse, at de features, der indgår i algoritmerne, ikke er den type data, arbejdstilsynsmyndigheder normalt har til rådighed (f.eks. brug af personlige værnemidler og administrativ praksis). Desuden udgør fald på gulv kun en meget lille del af de risici på arbejdspladsen, som arbejdstilsynsmyndighederne beskæftiger sig med.

Det andet eksempel er også en videnskabelig undersøgelse (Hajakbari og Minaei-Bidgoli, 2014). I denne undersøgelse blev der udviklet et scoringsystem for beregning af risikoen for arbejdsulykker. I undersøgelsen blev det konkluderet, at det er muligt at beregne risikoen for forskellige typer arbejdsulykker relativt præcist på grundlag af en række generelle virksomhedskarakteristika (en virksomheds hovedaktivitet, kønsfordeling, antal medarbejdere osv.). Det blev endvidere konkluderet, at algoritmen kunne udnyttes til at identificere arbejdspladser med behov for tilbagevendende arbejdstilsyn. De data, der blev brugt i denne undersøgelse, blev hentet i en arbejdstilsynsmyndigheds database. Også denne undersøgelse har dog den svaghed, at arbejdsulykker kun udgør én ud af de mange risici på arbejdspladsen, som arbejdstilsynsmyndigheder beskæftiger sig med. Derudover er der det særlige problem ved brug af ulykkesstatistikker, at man ved, at der er høj risiko for underrapportering.

Det tredje eksempel er et værktøj udviklet af den norske arbejdstilsynsmyndighed, som skal hjælpe tilsynsførende med at udvælge virksomheder på grundlag af risiko (Dahl et al., 2018). Værktøjet ved navn Risk Group Prediction Tool (RGPT) inddeler virksomheder i fire grupper ud fra forudberegnet risiko: virksomheder med laveste risiko, lav risiko, høj risiko og højeste risiko. Jo højere risikogrupper en given virksomhed er indplaceret i, desto større er sandsynligheden for, at et fremtidigt tilsyn i denne virksomhed vil identificere alvorlige afvigelser, når det gælder overholdelse af arbejdsmiljølovgivningen. Tilsynsførende kan se, hvilken gruppe en virksomhed er indplaceret i, via arbejdstilsynsmyndighedens interne webbaserede brugergrænseflade. Når der skal udvælges virksomheder til tilsyn, underrettes tilsynsførende derfor om virksomhedernes risikogrupper og er dermed i stand til at udvælge virksomheder på grundlag af viden om risici.

RGPT blev opbygget på grundlag af prædiktiv modellering ved hjælp af en machine learning-algoritme, der anvender såkaldt binær logistisk regressionsanalyse. På grundlag af regressionsmodellen indplaceres alle virksomheder i Norge (omkring 230 000) i en af de fire risikogrupper. Det sker i to trin. I det første trin beregner regressionsmodellen sandsynligheden for, at et fremtidigt tilsyn identificerer alvorlige afvigelser, når det gælder overholdelse af sundheds- og sikkerhedsbestemmelser. I det andet trin bruger modellen den beregnede p-værdi til at indplacere virksomheden i en risikogruppe.

Til at begynde med blev værktøjet udviklet på grundlag af registreringer fra omkring 35 000 sundheds- og sikkerhedstilsyn gennemført af arbejdstilsynsmyndigheden. Værktøjets beregninger bliver imidlertid gradvist og automatisk mere præcise, efterhånden som antallet af tilsyn stiger. Det betyder, at algoritmen justerer sig selv på baggrund af den feedback (korrekte eller forkerte beregninger), den modtager, når der gennemføres nye tilsyn, der registreres i arbejdstilsynsmyndighedens database.

RGPT-værktøjet hører til gruppen af supervised learning-algoritmer, hvor arbejdsmiljøtilsyn, der resulterer i alvorlige afvigelser (afhængig variabel), skal beregnes ud fra en række virksomhedskarakteristika (features). De features, RGPT-værktøjet anvender, er generelle virksomhedskarakteristika som virksomhedsstørrelse, branchegruppe, antal tidligere tilsyn, resultater fra tidligere tilsyn, virksomhedsalder, geografisk placering og anmeldte ulykker. Værktøjets prædiktive validitet kontrolleres hver måned, og erfaringen er indtil videre (efter omkring 18 måneders test), at algoritmen er i stand til at udvælge virksomheder med høj risiko ekstremt præcist. Det betyder, at der er få falske positiver og få falske negativt, dvs. at få tilsyn inden for gruppen med laveste risiko fører til, at der identificeres alvorlige afvigelser, mens langt hovedparten af tilsyn inden for gruppen med højeste risiko fører til, at der identificeres alvorlige afvigelser. Grupperne med lav og høj risiko ligger mellem de to ekstremer.

Resultaterne fra brugen af det værktøj, den norske arbejdstilsynsmyndighed har udviklet, viser, at det er muligt at udvælge tilsynsobjekter ved at udnytte big data og machine learning. Lignende machine learning-tilgange er også blevet testet af mindst to andre europæiske arbejdstilsynsmyndigheder med lovende resultater: den svenske arbejdstilsynsmyndighed (Ridemar, 2018) og den nederlandske arbejdstilsynsmyndighed SZW (Jacobusse og Veenman, 2016). Det norske værktøj kan dog ikke nødvendigvis overføres til andre arbejdstilsynsmyndigheder, da brugen af det afhænger af, hvordan data lagres, og af datakvalitet, dataadgang og databasestruktur. Desuden indebærer udvælgelse af virksomheder på grundlag af værktøjet en accept af måden, risiko defineres og operationaliseres på i algoritmen. Som beskrevet er værktøjet baseret på en definition af risiko, som indebærer, at jo højere risikogruppe en given virksomhed er indplaceret i, desto større er sandsynligheden for, at et tilsyn i denne virksomhed vil identificere alvorlige afvigelser, når det gælder overholdelse af arbejdsmiljølovgivningen. Det betyder, at værktøjet primært har at gøre med risici forbundet med styring og kontrol, ikke iboende risici. Mens risici forbundet med styring og kontrol opstår på grundlag af en virksomheds evne og vilje til at styre risici (f.eks. ved at overholde relevant lovgivning), handler iboende risici om de risici, der opstår ud fra arten af en virksomheds aktiviteter (f.eks. fald fra højde, eksponering for kemiske agenser og belastning af bevægeapparatet).

I praksis er der en sammenhæng mellem risici forbundet med styring og kontrol på den ene side og iboende risici på den anden side. Det betyder dog ikke, at der nødvendigvis er stor empirisk korrelation mellem de to typer risici. Hvis man derfor blindt forlader sig på værktøjer, hvor udvælgelsen er baseret på én type risici, kan det betyde, at en anden type ikke tages i betragtning. Denne udfordring tackles inden for de norske regler ved at lægge vægt på iboende risici, når prioriterede områder, risikoeksponerede grupper af arbejdstagere og højrisikobrancher identificeres, mens der lægges vægt på risici forbundet med styring og kontrol, når der konkret udvælges virksomheder.

Udfordringer

Det faktum, at der ikke nødvendigvis er empirisk korrelation mellem risici forbundet med styring og kontrol på den ene side og iboende risici på den anden side, fører os frem til endnu en – nok endnu større – udfordring ved at anvende big data og machine learning-algoritmer i forbindelse med risikobaseret udvælgelse. De tre ovenstående eksempler på machine learning-værktøjer er alle eksempler på endimensionel udvælgelse, dvs. udvælgelse baseret på én bestemt definition og operationalisering af risiko. Risici på arbejdsmarkedet er imidlertid ikke kun af én bestemt type. De tilsynsførende myndigheder beskæftiger sig derfor med mange forskellige typer risici, f.eks. ulykker, eksponering for kemiske agenser, eksponering for biologiske agenser, psykosociale trusler, risikofaktorer relateret til bevægeapparatet samt social dumping. Inden for disse risikotyper er der en række undertyper. Det er en stor udfordring at udvikle risikomodeller, der tager højde for alle disse risikotyper, fordi der ikke nødvendigvis er korrelation mellem de forskellige typer. At tage højde for alle disse risikotyper er derfor ganske forskelligt fra at beregne en bestemt risikotypes sandsynlighed (Dahl et al., 2018).

En anden, men dermed forbundet, udfordring gør opgaven med risikobaseret udvælgelse endnu mere kompleks. Det er den såkaldte politiske faldgrube (Black, 2010). Selv om machine learning-algoritmer er dynamiske, i den forstand at de kan lære af og justere sig i forhold til succeser og fejl, kan de ikke tage højde for forskellige politiske synspunkter. For det første er den politiske kontekst uberegnelig. Det betyder, at risikotyper, som det giver mening at prioritere i dag, ikke nødvendigvis giver mening at prioritere i morgen. For det andet har den politiske kontekst mange facetter. Således har forskellige interessenter, f.eks. politikere, arbejdsgivere, arbejdstagere, medierne og offentligheden, forskellige holdninger til, hvilke typer risici det giver mening at prioritere. Det viser, at risiko på arbejdsmarkedet ikke nødvendigvis er en objektiv enhed, men en social konstruktion.

En tredje udfordring, som det er værd at tage i betragtning, har at gøre med det faktum, at selv om arbeidstilsynsmyndigheder har rådighed over enorme mængder data om deres tilsynsobjekter, vedrører disse data normalt virksomhedsniveauet, og data på virksomhedsniveau er ikke nødvendigvis de mest egnede data at tage i betragtning (se for eksempel Gunningham og Sinclair, 2007). I en database identificeres en unik virksomhed typisk ved hjælp af en unik identifikator såsom et virksomhedsregistreringsnummer. En machine learning-algoritmes evne til at tildele en given virksomhed en given beregnet risikoværdi beror på unikke identifikatorer. Det er dog ikke alle potentielle tilsynsobjekter, der automatisk kan identificeres ved hjælp af en unik identifikator. For eksempel er det i byggebranchen ikke nødvendigvis en konkret virksomhed, der udvælges til tilsyn, men en midlertidig byggeplads. Der er mindst to udfordringer forbundet med et tilsynsobjekt af midlertidig karakter. For det første kan byggepladser og andre midlertidige arbejdssteder ikke nødvendigvis identificeres ved hjælp af en unik identifikator. For det andet betyder deres midlertidige karakter, at selv hvis de kunne identificeres, ville en machine learning-algoritme ikke nødvendigvis få chancen for at lære af sine beregningssucceser og -fejl, før byggepladsen nedlægges, og de virksomheder, der befandt sig på byggepladsen, er flyttet til nye konstellationer på en ny byggeplads.

Afsluttende bemærkninger

De ovenfor beskrevne udfordringer viser, at der er en række væsentlige problemer forbundet med at udvælge tilsynsobjekter med høj risiko ved at udnytte big data og machine learning-teknikker. Disse udfordringer betyder dog på ingen måde, at sådanne teknikker ikke er nyttige inden for en risikobaseret tilgang. Faktisk viser udfordringerne, at det næppe er en fordel for risikobaseret udvælgelse at basere den fuldt og helt på brug af machine learning-algoritmer. Det viser ovenstående eksempel fra Norge. I stedet for at lade algoritmen vælge og vrage direkte mellem objekter træffer de tilsynsførende beslutninger på grundlag af viden om risici og ud fra algoritmens beregninger. Det indebærer en kombination af kunstig og menneskelig intelligens, som hver især

komplementerer den andens styrker. Når det gælder forudberegninger af komplekse sociale hændelser generelt, er en kombination af de to typer intelligens formentlig en nødvendighed.

Referencer

- Black, J., 2010. Risk-based regulation: Choices, practices and lessons being learnt. I OECD 2010 (red.), Risk and regulatory policy: Improving the governance of risk. OECD, Paris.
- Blanc, F., 2013. Inspection reforms: why, how, and with what results. OECD, Paris.
- Dahl, Ø., Kilskar, S. S., Skarholt, K., Rosness, R., 2018. Risikobasert tilsyn i de nordiske arbeidstilsynene. [Risikobaserede arbeidstilsyn i de nordiske lande.] Nordisk Ministerråd, København.
- Gunningham, N., Sinclair, D., 2007. Factors impinging on the effectiveness of the Mines Inspectorate. The Australian National University, Canberra.
- Hajakbari, M. S., Minaei-Bidgoli, B., 2014. A new scoring system for assessing the risk of occupational accidents: A case study using data mining techniques with Iran's Ministry of Labor data. Journal of Loss Prevention in the Process Industries 32, s. 443-453.
- Jacobusse, G., Veenman, C., 2016. On selection bias with imbalanced classes. International Conference on Discovery Science, Bari, s. 325-340.
- Matías, J. M., Rivas, T., Martín, J. E., Taboada, J., 2008. A machine learning methodology for the analysis of workplace accidents. International Journal of Computer Mathematics 85, s. 559-578.
- Mischke, C., Verbeek, J. H., Job, J., Morata, T. C., Alvesalo-Kuusi, A., Neuvonen, K., Clarke, S., Pedlow, R. I., 2013. Occupational safety and health enforcement tools for preventing occupational diseases and injuries. Cochrane Database of Systematic Reviews 2013, udgave 8, artikel nr. CD010183.
- OECD, 2014. OECD Best practice principles for regulatory policy: Regulatory enforcement and inspections. Organisationen for Økonomisk Samarbejde og Udvikling, Paris.
- Ridemar, A., 2018. Decision support for SWEA inspections. KTH Royal Institute of Technology, Stockholm.
- Robens, L., 1972. Safety and Health at Work: Report of the Committee 1970-1972. Her Majesty's Stationery Office, London.
- Weil, D., 2008. A strategic approach to labour inspection. International Labour Review 147, s. 349-375.

Forfattere: Øyvind Dahl, Senior Researcher, SINTEF Digital — Safety and Reliability, Trondheim, Norge

Annick Starren, Det Europæiske Arbejdsmiljøagentur (EU-OSHA)