

Sustainability Circle

Anders Källström

Skapad 2019-08-27. Uppdaterad: 2019-08-27

Artikel 7:

## Artificiell intelligens - Herre, slav eller kollega?

Artificiell intelligens (AI) har många namn. Deep learning, machine learning, autonomous systems, mönsterigenkänning, natural language processing, computer vision är några av de begrepp befinner sig i AI-begreppets absoluta närhet och som bör få dig att associera till detta avsnitt:

För några decennier sedan såg vi på datorn som något som skulle befria oss från rutinmässig informationshantering. Självt skulle vi ägna oss åt avancerade arbetsuppgifter som intressanta analyser, bedömningar och beslutsfattande.

Nu ser vi att datorerna, med hjälp av de allt mer avancerade algoritmerna, inte bara klarar elementära och tråkiga arbetsuppgifter, utan även några av de mest avancerade. Ingenjörer, medicinare, rättsvårdare, ekonomer, marknadsförare, naturvetare, militärer och underrättelseverksamhet lägger ner mycket tid på att förstå hur algoritmerna kan användas och når allt oftare fantastiska resultat. Resultat, som med god marginal överträffar hjärnans.

Hur går det till? Det är inte alldeles lätt att förstå. Men, låt oss ändå göra ett försök. Detta är nämligen en viktig del av vår digitala framtid:

Låt oss utgå från hjärnan; Hjärnan lagrar mängder med intryck från sinnen i minnet. Medvetet och omedvetet reflekterar hjärnan därefter över dessa minnen, försöker karakterisera, klassificera och sätter namn på dem samt funderar på hur de hänger ihop. Dessa reflektioner ackumuleras i form av det vi upplever som insikter, livserfarenhet, bildning och visdom<sup>1</sup>.

Algoritmerna försöker bygga *modeller* över data som finns i big data. Förebilden är hur hjärnan fungerar. Därför talar vi om *neurala modeller*.

Algoritmerna letar, för att efterlikna hjärnan, i big data efter

---

<sup>1</sup> Kolb, D; *Experienced based learning*; 1984

korrelationer och regressioner (statistiska samband). I illustration 1 nedan har algoritmerna hittat mer eller mindre starka samband mellan faktorerna a, b, c, d och e.

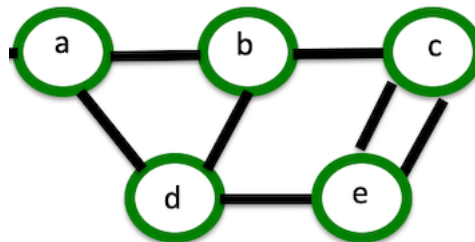


Illustration 1: Hur a, b, c, d och e korrelerar med varandra

Nästa steg är att identifiera de *kausala sambanden* mellan dessa faktorer, det vill säga vad som är orsak och vad som är *verkan*. Sambanden i illustration 1 ovan skall nu bli pilar som visar *riktningen*. Det görs genom att mäta när och därmed i vilken ordning faktorerna inträffar.

Resultatet framgår av illustration 2 nedan. Nu vet vi att det är a som påverkar b, att c och e påverkar varandra...

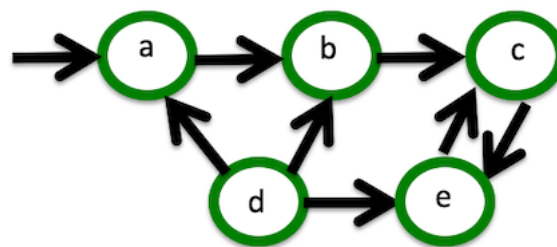


Illustration 2: Kausala samband mellan faktorer som korrelerar

Om en konstruktionsritning visar hur tekniken är tänkt att fungera så visar den neurala modellen hur den fungerar.

Modellen kan användas för att analysera big data såväl "nedströms" som "uppströms". Analyser nedströms kan användas för att förutsäga (prediktera) *vad* som kan inträffa. Analyser uppströms kan användas för att ta reda på *varför* något hände (rot-orsaks-analys). Den artificiella intelligensen erbjuder alltså därmed svaren på "vad" och "varför".

Vi har tidigare konstaterat att "data is the new oil". Det är en sanning som nu skall preciseras: De största värdena i databaserna ligger inte i enskilda data, utan i *sambanden mellan* dessa data, det vill säga i de neurala modellerna.

Algoritmernas byggande av de neurala modellerna kallas "inskolning". Inskolningen tar åtskilliga månader om big data är omfattande och sambanden komplexa.

Men, kraven på algoritmerna är större än så. Den värld som sensorerna mäter är föränderlig. Det måste få konsekvenser för de neurala modellerna. De måste förändras i takt med verkligheten för att alltid vara relevanta. Algoritmerna ägnar sig därför inte bara åt "inskolning" utan därefter även åt ett "livslångt lärande". Detta pågår outtröttligt dygnet runt, 365 dagar om året.

Digitalt lärande i form av inskolning + fortsatt livslångt lärande är en av definitionerna<sup>2</sup> på *artificiell intelligens* (AI). Begreppet *machine learning* (som används sedan 2016) är en synonym glosa.

Verklighetens neurala modeller omfattar naturligtvis miljon- och miljardtals faktorer och dynamiska samband. AI utvecklas därmed snabbt till en *black box*. En black box en modell för bearbetning av data som människan på grund av modellens komplexitet och dynamik inte kan förstå<sup>3</sup>. Algoritmerna kan exempelvis lösa ett problem i ett mycket stort antal dimensioner. Hjärnan klarar knappast av att förstå mer än fyra. Därför har Explaining AI (XAI) blivit ett nytt utvecklingsområde. När XAI inte lyckas får människan leva med att AI löste problemet utan att vi förstår *hur*.

Frågan är om AI blir vår herre, slav eller kollega?

### Praktikfall: Fortum

Fortum är ett börsnoterat energiföretag med huvudkontor i Finland. Företaget har en omfattande verksamhet i bland annat Finland och Sverige. Företaget är inriktat mot produktion av elektricitet (vattenkraft och kärnkraft) samt värme (kraftvärmeverk).

2017 beslutade Fortum, som en del av det övergripande arbetet med digitaliseringen, att testa Smart Maintenance med hjälp av AI och neurala nätverk. Detta gjordes inledningsvis på kraftvärmeverken i Finland.

Vi träffade Fortums Linda Ramstedt Marklund, redan i augusti 2017, för att få en första inblick i projektet.

---

<sup>2</sup> En allmänt fastlagd definition av AI saknas

<sup>3</sup> En *white box* är en manuellt byggd modell. Denna är givetvis mycket enklare och mer eller mindre statisk och därför, de allra flesta fallen, ointressant. Ett mellanting är en *grey box*, dvs modellens stomme är manuellt byggd, men algoritmerna har därefter utökat modellen.



Varför tar ni hjälp av algoritmer för att bli bättre på underhåll?

- Att möjliggöra prediktion av haverier, är ett viktigt område för oss, berättar Linda. Vi har arbetat under många år med att samla in data. Men vi kände inte att vi fick ut nyttan av det, på ett strukturerat vis. Nu har vi börjat leta efter mönster i dessa data.

Hur går det till?

- Vi använder oss av neurala nätverk. Med hjälp av dessa fångar vi mönster i historiska data. När dessa mönster, efter en upplärningsperiod, har "satt sig" kan vi jämföra dem med online-data. Om on-linedata avviker från mönstren får vi varningar. Varningarna används för att vidta åtgärder innan avvikelserna leder till haverier.

Var gör ni detta?

- Vi har provat de neurala modellerna på kraftvärmeverken i Finland.

Vad är vinsten?

- Vi levererar alltid värme till kunden, men om vi får ett haveri i en basanläggning så måste vi starta upp en reservanläggning, exempelvis en oljepanna, och det blir både dyrt och mindre miljövänligt.

Vad är svårt tekniskt i användningen av algoritmerna?

- Det är inte så svårt, åtminstone inte längre. Vi har gjort en resa. Vi började med att bygga upp algoritmer för små delar av anläggningen, exempelvis för en pump. Vi insåg ganska snabbt att det inte var så svårt. Men, när vi skulle skala upp det upptäckte vi att vi behövde andra algoritmer, för andra maskiner, så att vi kunde fånga in hela produktionsanläggningen. Den bygger på tiotusentals mätvärden. Då insåg vi att skalbarheten var komplicerad och att vi behövde algoritmer som bygger upp algoritmerna. Det är en black box-lösning. Vi vet alltså inte exakt hur den fungerar. Algoritmerna testar vilka algoritmer som skall användas för att bygga det neurala nätverket. Det sker automatiskt. Datorn behöver några veckor på sig för att göra jobbet.

Vad händer sedan?

- När de neurala modellerna är på plats så börjar de larma när det uppstår avvikelser mellan hur anläggningen enligt de neurala modellerna brukar fungera och hur den fungerar just nu. Och, då måste vi validera vilka larm som är sanna och vilka som är falska.

Detta är ju ett bra exempel på artificiell intelligens. Har ni full koll på hur det fungerar..?

- Nej, det har vi inte, men det är också lite av definitionen på ett neuralt nätverk. Det neurala nätverket är smartare än vad man kan förstå. Mönstren är som de mönster vi har i hjärnan men mer komplexa, för komplexa för att vi fullt ut skall förstå dem.

Hur känns det att inte fullt ut förstå vad som händer i modellerna?

- Vi stoppar in alla våra värden och får larm avseende avvikelser. När vi går ut i anläggningen så ser vi om de är relevanta eller inte. Och, då behöver inte jag förstå exakt hur datorn kommit fram till dessa

Har utveckling och implementering kostat mycket?

- Nu är framtiden här. Det finns färdiga algoritmer som kan implementeras. För bara ett år sedan kändes det som en orimlig investeringskostnad att bygga upp modeller för ett helt kraftvärmeverk. Nu finns det en lösning som gör det automatiskt. Så det är både kostnadseffektivt, går fort och kräver inte mycket av våra egna resurser.



Två år senare, i augusti 2019, träffar vi Linda för att ta del av de fortsatta erfarenheterna.

- De senaste åren har vi rullat ut en fullskalelösning, för ett prediktivt underhåll, berättar Linda. Den fungerar i korthet så att vi använder alla processdata, alla mätsignaler som finns från ett kraftvärmeverk eller ett kraftverk. Algoritmerna identifierar mönster och genererar larm redan vid tidiga avvikelser. Vi reagerar på larmen och kan därmed förebygga otillgänglighet i anläggningen. Det innebär att vi får högre tillgänglighet vilket påverkar vår lönsamhet. Vi är fortfarande nöjda med den här lösningen. Vi börjar långsamt närma oss någon sorts fullskalig implementation. Det är en stor organisation så det tar tid även om vi når snabba resultat.

Ni började med ett kraftvärmeverk i Finland.

- Precis. Vi har installerat det på kraftvärmeverken i Finland. Vi har under våren gjort en fullskalig installation på fjärrvärmeverksamheten här i Stockholm. Nu håller vi på och kikar på verksamheterna i Ryssland, Polen och i Norge. Vi tittar även på kärnkraften och vattenkraften.

Om vi pratar lite om Finland där ni har mest erfarenhet. Du var väldigt entusiastisk för två år sedan.

- Man vänjer sig snabbt ... Om vi tittar på Finland så har problematiken med otillgängligheten varierat över tiden. Vi har haft perioder när vi lyckats kapa otillgänglighetskostnaderna väldigt mycket och så har vi perioder när vi inte

lyckats lika bra. Alla åtgärder som vi gjort, baserat på larmen, följs upp så att vi lär oss om sambandet mellan larm och otillgängligheten. Det är så vi kan bedöma nyttan av systemet. Om det händer något i anläggningen som inte larmades så följer vi även upp detta. Svaret kan vara att modellen behövde tränas på nya problem eller att vi behöver fler eller nya sorters sensorer. Vi följer upp allt väldigt noga. Det är viktigt att vi följer upp både vad som hänt och effekten av de åtgärder vi gör baserat på larmen.

Det gick väldigt fort att implementera algoritmerna i Finland. Hade ni alla sensorer som ni behövde?

- Ja, kraftverk har extremt mycket sensorer. Ett kraftvärmeverk har mellan tio tusen och femtio tusen sensorer. Vattenkraftverk har färre. Dessutom har vi en pragmatisk inställning. Vi kör igång på de sensorer som vi har. Om vi behöver fler så upptäcker vi det, när vi får haverier som vi inte lyckats prediktera. Med facit i hand, det var ett fåtal nya sensorer, en handfull per anläggning, som vi behövde komplettera med. Och, det är en styrka med den nya tekniken. Den kan hantera alla mätsignaler – även från nya sensorer.

Hade ni ordning på databaserna?

- Vi har konsoliderat data så att den finns på ett ställe. Och, så streamar vi från detta. De neurala modellerna bryr sig inte om ifall data beskriver funktionen i en pump eller något annat. Algoritmerna letar bara efter mönster. Sedan är det klart att vi måste veta vart vi skall gå och titta om du får ett visst larm. Om du får data som varierar på ett konstigt sätt så kan det indikera en trasig sensor och då får vi byta ut den. Vi har ett väldigt pragmatiskt förhållningssätt. Vi börjar med det vi har.

Vi pratar om att prediktera fel. Använder ni de neurala modellerna för att hitta *orsaken* till felen?

- Vi får väldigt tidiga larm. Därmed får vi ofta också grundorsaken. Det behövs inte så mycket felsökning. Den första signalen indikerar var problemet uppstår först. Därmed åtgärdar vi något som senare skulle kunna orsakat ett haveri. Dessutom kan vi mappa larmen mot arbetsorderna i underhållssystemet. Så nästa gång som vi får ett liknande larm så är det lätt att veta vad vi skall vidta för åtgärd.

Hur avancerade är de neurala modellerna?

- Varje modell baseras på mellan hundra och tusen signaler, från lika många sensorer. Det är väldigt avancerade neurala modeller som beskriver mönstren. Om hjärnan kan hantera en handfull parametrar så hanterar de neurala nätverken alltså upp emot tusen.

Vad har ni för erfarenhet av falsklarm? Falsklarm kan ju delas in i positiva, det vill säga att man får ett larm som man inte skulle haft, och negativa, att ett fel inte larmas.

- När vi börjar implementera algoritmerna får vi en del positiva falsklarm under valideringsfasen. Det är alltså en del av systemets träning då mänsklig kunskap adderas till modellerna. När vi upptäcker att vi *inte* fått ett larm som vi borde fått så beror det oftast på att det saknas en sensor. Då sätter vi förstås in en sådan. Om det finns en sensor som fungerar så får man gå in och träna modellen på just denna händelse.

När man genomfört ett första projekt är det ju, som i Fortums fall, dags för en uppskalning. Vilka är utmaningarna?

- Det handlar som vanligt om change management, att ändra människors beteende och om viljan att lära av varandra. Ingenjörer vill ju ofta hitta på sina egna lösningar. Det kan också vara svårt att bevisa den ekonomiska effekten, det vill säga hur mycket otillgängligheten skulle ha kostat om vi *inte* hade infört de neurala modellerna. Hur hög hade kostnaden varit för det som inte hänt? Det blir snabbt filosofiskt. Där tekniken är implementerad, så är folk nöjda och ser detta som ett bra komplement. Det försöker vi sprida bland annat med studiebesök hos dem som infört de neurala modellerna.

Vi har inte haft några tekniska problem i samband med uppskalningen. Den är väldigt skalbar, tekniskt sett.

Ser ni något etiskt problem med neurala modeller som är mer komplicerade än vad ni kan förstå?

- Så länge vi inte automatiserar handlingar så ser jag inga moraliska problem. Det krävs alltid en mänsklig validering av larmen innan de leder till konkreta åtgärder. Avvikelse i mönstren från sensorerna analyseras innan vi vidtar åtgärder. Algoritmerna tar aldrig några beslut. Jag kan inte se att detta skulle innebära något moraliskt problem. Om vi får ett larm som vi inte kan verifiera så ser vi det som en träning av modellen.



De traditionella kontrollsystemen genererar förhållandevis många larm. Den neurala modellen genererar jämförelsevis väldigt få, som dessutom ofta visat sig vara relevanta, så det är inget problem att validera dem genom att gå ut och titta efter.

Vad ser ni se ert nästa steg i arbetet med AI och algoritmer?

- Hittills har vi implementerat algoritmer som hittar mönster och avvikelser i vår processdata vilket har varit enkelt att implementera eftersom dessa algoritmer inte är så beroende av hur data är strukturerad. För att komma vidare i vårt arbete med algoritmer vill vi kunna jobba med flera olika datakällor som input till prediktiva modeller på ett automatiserat och skalbart sätt.

Vad ser ni för nyttor med det?

- De neurala nätverken är bra på att detektera avvikelser och prediktera fel som kommer ske inom de närmaste dagarna till veckorna och nu vill vi fokusera på hur vi kan jobba mer datadrivet med den mer långsiktiga underhålls planering med perspektivet upp till 10 år.

Vilka lösningar ser ni kring detta?

- Vi håller på att kika på grafdatabaser som en hjälp för att koppla ihop information från olika datakällor med olika struktur. Vi har hittills bara gjort ett antal experiment kring detta men det ser mycket lovande ut. Det som tidigare var svårt att hantera i olika relationsdatabaser verkar vara lättare att koppla ihop i grafer så vi hoppas på att kunna koppla ihop processdata med underhållsdata och dokument för att få ihop all information kring våra anläggningar.



