Statistik-
främjandet

TILL MINNE AV
OVE FRANK

SIDAN 24

VÄLKOMMEN PÅ
ÅRSMÖTE I MARS

SIDAN 31

Qvintensen

NR 1 2020

TEMA/

Artificiell intelligens och maskininlärning

SIDORNA 8-23

ÄR DET P-VÄRDEN VI
SKA VARA RÄDDA FÖR?

SIDAN 4

Ett axplock av Qvintensen 2009–2019.

Välkommen som skribent i Qvintensen

Vi uppskattar om du vill bidra till Qvintensen. Det är Statistikfrämjandets medlemmar som gör att vi kan ha en levande medlemstidning med intressant innehåll för just statistiker i Sverige.

Vad är Qvintensen?

Qvintensen är Statistikfrämjandets medlemstidskrift. Den ska innehålla sådant som på ett eller annat sätt kan vara av intresse för medlemmarna. Främjandets fyra delföreningar har också möjlighet att komma med information riktad speciellt till deras medlemmar.

Qvintensen lämpar sig alltså inte för information med kort varsel om kommande aktiviteter eller lediga tjänster. Sådan information skickar Statistikfrämjandet fortlöpande till medlemmarna via e-post. Genom Statistikfrämjandets hemsida har dessutom både medlemmar och icke-medlemmar tillgång till en samling artiklar och inlägg på nätet, som också går under namnet ”Qvintensen”. Men dess innehåll avviker helt från pappersversionens innehåll och berörs inte av de anvisningar som ges här. Följande anvisningar avser enbart Qvintensens pappersversion, som för närvarande utkommer med två nummer per år.

Vad kan Qvintensen innehålla?

Qvintensen är ingen vetenskaplig tidskrift, utan ett medlemsblad. Den innehåller information om Statistikfrämjandets (och delföreningarnas) verksamhet. Den ger framför allt möjlighet för medlemmar att bidra med artiklar, debattinlägg och annat som kan vara aktuellt och intressant för statistiker. Av tradition publiceras inga recensioner. Artiklar bör vara relativt populärt hållna och inte kräva avancerade specialkunskaper. Alltså inte den typ av artiklar som hör hemma i prestigeladdade vetenskapliga tidskrifter.

Att tänka på för dig som skribent

Qvintensen har mycket små redaktionella resurser. Därför uppskattar vi om du tänker på följande saker när det gäller text:

- Manuskript kan vara olika långa, men inte mer än 6 A4-sidor.
- Manuskript bör ha enkelt standardformat utan konstigheter. Använd inte avstavning, fotnoter, rak högermarginal eller dubbla spalter. Manuskript måste vara möjliga att redigera (inte PDF-format).
- Skriv på svenska om det inte finns speciella skäl att använda engelska.
- Texten bör vara lätt att förstå och du bör därför undvika formler (enstaka matematiska eller statistiska symboler är tillåtna).
- En eventuell referenslista ska i första hand göra det möjligt för läsaren att lätt hitta de referenser som det hänvisas till i artikeln. Harvard-systemet rekommenderas.

Beträffande eventuella foton, tabeller och diagram, tänk på följande:

- Färgtryck är möjligt bara på två sidor: Främre omslagets insida och bakre omslagets insida.
- Foton ska bifogas dokumenten som enskilda filer. De bör inte infogas i textdokumentet.
- Bildformatet bör vara .jpg, .eps eller .tif.
- Välj aldrig att förminska eller komprimera foton, om dator eller telefon erbjuder detta.
- Foton ska innehålla så många kb som möjligt. Om bilderna innehåller för lite information blir de små i tidningen.
- Bilder som kopieras från nätet innehåller oftast för få kb. Välj därför alltid att "ladda ner högupplöst fil" när detta erbjuds.
- Tabeller och diagram bifogas helst som separata PDF-filer.

Några källor som du kan ha nytta av när du skriver är:

- Språkrådets handbok *Svenska skrivregler*, som innehåller massor av skrivtekniska tips. (*Svenska skrivregler*, fjärde upplagan, 2017. Liber).
- Svenska Akademiens ordböcker SAOL (Svenska Akademiens ordlista) och SO (Svensk ordbok). (Båda finns på nätet: sök på ”www.svenska.se.”)

Kontakt

Är det något du undrar över, ta kontakt med redaktören eller redaktionen; se spalten med redaktionell och annan information i senaste nummer av Qvintensen.

I förra numret av Qvintensen hade vi ett antal artiklar om p-värden och problematik kring dessa. Här kommer ett inlägg inspirerat av den diskussionen.

Ur Qvintensen nr 2 2019.

Är det p-värden vi skall vara rädda för?

Låt mig först säga mina erfarenheter som statistiker är från epidemiologin, där vi försöker hitta samband mellan olika riskfaktorer (buller, rök, damm, stress ...) och olika hälsoutfall (död, hjärt- och lungsjukdom, diabetes ...). Vi studerar stickprov dragna från exempelvis bosatta i Stockholm. Givet att alla förutsättningar för ett test är uppfyllda, så gör man i genomsnitt ett falskt ”fynd” vid var 20:e test om man använder $p<0,05$ som vattendelare, och om det är sant att det inte finns någon effekt. Så långt fungerar p-värden klanderfritt. Resultat som går i hypotesens riktning och har låga p-värden är vad vi ofta letar efter. Punktskattning och 95% konfidensintervall för punktskattningen är informativa och bra. I en enskild studie kan de slumpmässiga felen vara påtagliga, vilket framgår av konfidensintervallens bredd i förhållande till de effekter vi letar efter. Eftersom epidemiologiska studier ofta upprepas av olika forskare, så är i slutändan (efter en metaanalys av de ingående studierna) de slumpmässiga felen ofta små.

Det kan finnas många metodproblem som hotar validiteten i en epidemiologisk studie. Det viktiga är att vi har en moralisk skyldighet att polemisera mot andra publicerade resultat, om vi ser att våra resultat går i en annan riktning. Om vi inte är noggranna med att publicera studier med negativt resultat studier så havererar forskningen. Ett problem med observationsstudier är att man kan hitta skensamband som t.ex. att alkohol ökar risken för lungcancer. Nej, de som dricker mycket tenderar att också röka mycket. Skensambandet uppstår om vi inte justerar fullt ut för effekten av rökning. I värsta fall därför att vi inte har med rökning eller ens någon proxy för rökning i sambandsberäkningen. Rökning är en confounder som vilseleder sambandet mellan alkohol och lungcancer. I observationsstudier (till skillnad mot experiment där individerna slumpas ut till experiment- och kontrollgrupp) måste alla faktorer som är confounders till ett samband vara kända på förhand, och vi måste ha inhämtat information om dem för att kunna ta hänsyn till dem i sambandsberäkningen. Inte lätt.

Slumpmässiga mätfel i utfallsvariabeln gör det svårare att hitta samband. Slumpmässiga mätfel i exponeringsvariabeln (alkohol, om man studerar sambandet mellan alkohol och lungcancer) leder till att sambandets styrka underskattas. Alla confoundingvariabler måste finnas med i sambandsberäkningen och får inte ha mätfel, för i så fall kan man inte justera fullt ut för den vilseledande effekten. Om rökning är en confounder så behöver vi veta hur mycket var och en av individerna i studien röker. Det duger inte att approximera faktorn rökning med hur mycket grannarna röker i det bostadsområde där individen bor. Det leder till ett mätfel i faktorn rökning med åtföljande bristande justering för confounding och i värsta fall ett vilseledande resultat.

Confounding är ett problem i observationsstudier, inte minst om man letar efter svaga effekter av exponeringsvariabeln. Ett sätt att mildra problemet, om man hittar samband med traditionell confoundingjustering, är att samla in nya data dragna från några grupper som alla är relativt homogena, vilket minskar risken för confounding inom grupp. En homogen grupp skulle kanske kunna vara sjuksköterskor födda i Sverige i åldrarna 30-50 år. Även i dessa grupper skulle vi försöka justera bort confounding på traditionellt vis. En nackdel är att slutsatserna begränsas till de grupper vi undersökt, samt att det krävs stora stickprov. En fördel är att vi får en djupare insikt om sambandet. Med ett internationellt samarbete kan de slumpmässiga felen ändå hållas små. När en studie upprepas är det viktigt att definitioner (exempelvis vad som är skräpmat eller vid vilket värde en variabel kategoriseras) görs på samma sätt i de ingående studierna. Annars öppnas dörren för ”data massage” med åtföljande missvisande låga p-värden. Här förutsätter jag att studier i olika länder är jämförbara om man bara gjort dem på samma sätt. Så är kanske inte alltid fallet vilket i dessa fall gör forskningen svårare och resultaten mindre generaliserbara.

Låt mig tillägga att man vanligen justerar för socioekonomiska faktorer genom att använda utbildning eller typ av yrke som en proxy för okända confounders. Andra proxys för okända confounders kan vara bostadsområde, eller genomsnittsinkomsten i det bostadsområde där individen bor.

För att återknyta till rubriken: Nej det är nog mer än p-värdets beräkning och formella tolkning som skapar problem. De många felkällorna i en epidemiologisk studie leder ofta till ett systematiskt fel i punktskattningen, varför även p-värdet blir fel och konfidensintervallen ligger centrerade över fel värde, där det sanna värdet kan vara både högre och lägre. Den som har en *älsklingshypotes*, och som vet vilka resultat han eller hon förväntar sig, kan övertolka det mesta oberoende av om p-värden används eller inte. Det viktiga är att studien upprepas och att även negativa resultat publiceras och inte stoppas i byrålådan. Kvarstående confounding är ett hot mot slutsatserna och här hjälper det inte att studien upprepas. Det bästa är att samla in data som det går att dra slutsatser från. Lättare sagt än gjort. Och, till sist, var ödmjuk inför resultatet.

TOMAS LIND, STATISTIKER
CENTRUM FÖR ARBETS- OCH MILJÖMEDICIN
REGION STOCKHOLM

Qvintensen

Ansvarig utgivare
Joakim Malmdin

Redaktörer
Jan Wretman, 070-781 78 35
Marcus Berg

Redaktion
Hans Alberg
Jens Malmros
Rolf Larsson
Marie Linder
Marika Wenemark

E-post wretman.jan@gmail.com

Produktion
Form och redigering: Mezzo Media AB
Tryckeri: Trydells Tryckeri AB

Annonser
Annonser i Qvintensen bokas med redaktören.
Annonsutskick per e-post bokas med Statistikfrämjandets sekreterare. Priset är 6 000 kronor för institutionsmedlem eller motsvarande, och
8 500 kronor för övriga annonsörer. På alla priser tillkommer 25 % moms.

Statistik-
främjandet

Adress
Svenska statistikfrämjandet
Sekreterare: Mattias Strandberg

E-post sekrfram@gmail.com

Webbplats www.statistikframjandet.se

»...även om ett samhälle med 100% arbetslöshet i viss mening är jämlikt, så är det inte uppenbart hur ett sådant samhälle kan organiseras.»

OLLE HÄGGSTRÖM, SIDORNA 12-15

Qvintensen
STATISTIKFRÄMJANDET
NR 1 2020

Innehåll

Våra föreningar

ISSN 2000-1819

AI-ML, ett område med många ingångar

Det händer ibland att jag behöver ringa till något servicenummer för att fråga om något. Ofta är det då en i förväg inspelad röst, som ber mig uttrycka mitt ärende. Det känns konstigt att tala med någon som inte är en levande människa, men jag försöker ändå stapplande utrycka vad jag vill veta. Det brukar bli tyst ett tag, och sedan säger datorrösten något i stil med ”Du vill alltså veta …”. Vilket inte alls stämmer med det jag frågade om. Jag försöker på nytt säga något, men utan resultat, och så småningom avbryter jag samtalet. Det här torde väl vara mina enda personliga möten med så kallad artificiell intelligens, och jag får skamset erkänna att datorn har förklarat mig underkänd. Men nog om detta.

Temat för det här numret av Qvintensen är artificiell intelligens (*AI*) och maskininlärning (*ML*). Det är termer från datavetenskapen, som man nu allt oftare stöter på även i statistikerkretsar. Det verkar även som att man inom datavetenskapen ibland använder vissa traditionella statistiska begrepp, fast under nya namn. Inför detta nummer hade vi uttryckt önskemål om bidrag rörande anknytningspunkter mellan *AI-ML* och statistik, vad gäller teori, tillämpningar och utbildning. Jag erkänner utan förbehåll att jag är mycket okunnig om dessa saker och jag har med stort intresse läst de bidrag som kommit. Jag vill tacka författarna, som medverkat med kort varsel, och jag vill också tacka Marcus Berg, som varit medredaktör för detta nummer, och som varit en värdefull diskussionspartner.

***AI-ML* är ett område** med ingångar från många olika håll. Sådant som datorteknologi, datorprogrammering, matematik, numeriska metoder, logik, sannolikhetsteori, statistisk teori, hjärnfysiologi, inlärningspsykologi, medvetandefilosofi, etik och samhällsanalys. När *AI-ML* ska presenteras på ett mer populärt sätt talas det oftast om spektakulära saker som ansikts-, röst- och textigenkänning och om självlärande robotar, som ska ha samtalsliknande kontakt med levande människor. För att inte tala om science fiction och dystopiska framtidsvisioner, som ofta kommer med i bilden. Det där verkar för en utomstående inte ha så mycket med statistik att göra. Men ett gemensamt mål både för tillämpad statistik och för vissa tillämpningar av *AI-ML* verkar vara att man utifrån insamlade data vill dra slutsatser och eventuellt fatta beslut. Eller åtminstone försöka skaffa sig en någorlunda tillförlitlig bild av hur saker och ting hänger ihop.

Hur är det nu med anknytningen mellan statistik och *AI-ML*? Den har jag nog inget grepp om. Inom traditionell statistik brukar man skilja mellan deskriptiv statistik och statistisk inferens. Den statistiska inferensen är sannolikhetsbaserad. Där används termer som sannolikhet, stokastisk variabel, sannolikhetsmodell, parameter, estimation, prediktion och hypotesprövning. I *AI-ML*-fallet däremot stöter jag inte på några sådana termer, utan bara hänvisning till avancerad datorteknologi och avancerade datorprogram. Men vilka är de övergripande tankarna inom *AI-ML*? Det är som sagt oklart för mig

I statistiska tillämpningar är intresset ofta inriktat mot säkerheten (eller snarare osäkerheten) hos de resultat som erhålls. Man resonerar om datas osäkerhet, om bias, konfidens och signifikans i samband med slutresultaten. Diskuterar man inom *AI-ML* sådant som osäkerhet och relevans hos de data som datorn utnyttjar? Och hur bedömer man tillförlitligheten hos de resultat som datorn producerar? Kommer sannolikhetsteori och inferensteori någon gång in i bilden? Det vet jag ingenting om. Har jag en föråldrad uppfattning om vad statistik är för någonting? Låt oss nu se vad de medverkande skribenterna har att säga.

Ingenting som sägs i detta nummer ska uppfattas som av Statistikfrämjandet sanktionerade åsikter, utan varje författare står själv för de åsikter som framförs.

JAN WRETMAN

Ordlista artificial intelligence och machine learning

Vilken ära att bli tillfrågad som gästredaktör till Qvintensen, man tackar för förtroendet! Jag vill nu passa på att förbereda er läsare på vad ni har att se fram emot.

Artificiell intelligens (AI) är det senaste heta buzzwordet som tangerar vår bransch och vårt yrke. Alla vill ha det, men få tycks kunna komma överens om vad begreppet verkligen inbegriper. Jag brukar säga att AI är *något artificiellt som gör något intelligent*, och att grovt förenklat så kan all form av AI placeras på en skala från Artificial Narrow Intelligence (ANI) till Artificial General Intelligence (AGI). **Narrow AI** är en modell eller algoritm som gör en sak, exempelvis läser ett tweet och graderar semantiken. Kopplar man ihop flera narrow AI till ett system så kan man få dom att åstadkomma mer komplexa utfall, exempelvis att läsa flera tweets och skicka ut en varning ifall det sägs för mycket negativt under en viss tidsperiod, men varje komponent är fortfarande narrow AI. **AGI** är all AI som vi ser i Hollywoodfilmer, robotar som tänker och resonerar och försöker åstadkomma komplexa utfall i samhället (exempelvis utplåna mänskligheten). AI är ett spektrum och all AI vi ser i samhället idag ligger närmare ANI än vad det gör AGI.

» Alla vill ha det, men få tycks kunna komma överens om vad begreppet verkligen inbegriper.»

AI försöker åstadkomma uppsatta mål med hjälp av ett stort antal olika verktyg: konkreta regler, matematiska optimeringar, natural language processing (NLP), statistiska modeller, machine learning (ML) algoritmer och/eller deep learning algoritmer. **Regler** kan vara simpla (”om du får välja färg, välj blå”) eller komplexa (”om en kvinna frågar vad du lyssnar på för musik och det är en lördagskväll, svara The Beatles”), men dom är alltid strikta. **Matematiska optimeringar** (t.ex. linjär optimering) är också regelsystem, men tar beslut baserat på vad som är optimalt, givet matematiskt uttryckta förutsättningar. **NLP** analyserar text och utläser attityder och semantik. **Statistiska modeller** (ex regression), förutsätter jag att ni alla vet vad det är.

Machine learning (ML) är matematiska algoritmer som producerar, likt statistiska modeller, numeriska skattningar eller någon form av klassificering. Man brukar dela in ML i **supervised learning** (vanliga prediktionsproblem eller klassificeringar), **unsupervised learning** (t.ex. klustringsalgoritmer som K means clustering eller dimensionalitetsreduceringsmetoder som principalkomponentanalys (PCA)) och **reinforcement learning** (iterativa algoritmer som interagerar med sin ”omgivning” och lär sig optimala beteenden för att uppfylla sitt belöningskriterium, det är till exempel vanligt att använda reinforcement learning för att lära en dator att spela ett TV-spel som Super Mario). I branschen går meningarna isär huruvida reinforcement learning ligger under ML eller ifall det är ett separat området, personligen lägger jag ingen vikt åt varken det ena eller det andra hållet.

Deep learning algoritmer är baserade på artificiella **neurala nätverk** som skattar s.k. **hidden layers** och sedan använder **activation functions** för att klassificera eller predicera. Dom neurala nätverken optimeras ofta genom s.k. **backpropagation**, men det finns andra optimeringsförfaranden. Deep learning kan vara supervised eller unsupervised. Algoritmerna kräver stora mängder data (exempelvis stora fotodatabaser) och tar väldigt lång tid att träna upp (ofta dagar eller veckor). Deep learning är en black box, och det går inte att får reda på exakt hur dom olika hidden layers räknades fram. Det är vanligt bland personer (framför allt statistiker) som håller på och lär sig om deep learning att man vill veta exakt hur algoritmen går till väga för att skatta sina hidden layers, men det går inte. Man måste helt enkelt acceptera att det är en black box.

Machine learning och statistik försöker ofta åstadkomma samma saker, men då dom ursprungligen kommer från olika discipliner så finns det mycket utbytbar terminologi områdena emellan: Det som vi inom statistiken kallar för variabler kallas inom ML för **features**, parameterskattningar kallas för **weights**, och att anpassa en modell (fitting the model) kallas för att algoritmen/modellen lär sig (**learning**). När en modell presterar bra på testdata, kallas det för att den **generaliserar bra**.

Som ni förstår var det här en högst kortfattad introduktion till AI och ML vars syfte är att sätta lite relevanta ord och begrepp på kartan snarare än att ge någon form av djupare förståelse. Med detta sagt så hoppas jag att ni nu känner er lite bättre förberedda att ta er an alla spännande artiklar i detta temanummer. Håll tillgodo!

MARCUS BERG

Statistikämnet är hotat – tack och lov

Data finns numera överallt och är en enorm drivkraft för modern industri. I denna blomstringstid för dataanalys finns paradoxalt nog en oro för statistikämnets död. Vi utmanas av närliggande ämnesområden som maskininlärning. Det är en välbehövlig katalysator till förnyelse av vårt ämne.

Jag har under åren 2011–2019 lett avdelningen för statistik och maskininlärning vid Linköpings universitet med ambitionen att integrera statistik med signalbehandling och datavetenskap, speciellt maskininlärning. I det arbetet har jag tillsammans med kollegor utvecklat och undervisat på ett stort antal kurser i maskininlärning för statistiker och ingenjörer, och forskat i gränslandet mellan statistik, maskininlärning och artificiell intelligens. Jag upplever därför inget hot från angränsande ämnen, utan ser tvärtom stora möjligheter till samarbeten och framför allt en vitalisering av statistikämnet.

Jag ska här föreslå sex områden där jag upplever att vårt ämne bör förändras om vi ska fortsätta att spela en central roll i det nya informationssamhället. Jag är väl medveten om att andra universitet, speciellt de som inte har en närhet till en teknisk högskola, kommer behöva välja en delvis annan väg än den vi valde i Linköping. Men jag vill hävda att mina sex förslag är applicerbara på alla utbildningar i statistik och matematisk statistik; de handlar om modernisering av undervisning i statistisk dataanalys och inferens. Som utgångspunkt för mitt inlägg ska jag först beskriva vad jag själv lägger i begrepp som artificiell intelligens, maskininlärning och data science.

Vad är artificiell intelligens och maskininlärning?

Artificiell intelligens (AI) handlar om att ge maskiner eller mjukvara någon form av människoliknande kognitiv förmåga. Mer konkret handlar AI om att göra det möjligt för en maskin att lära känna sin miljö och fatta beslut som maximerar sannolikheten att uppnå uppsatta mål. Moderna AI-system använder data för att lära upp dessa system med hjälp av maskininlärning. Under senare år har det skett stora framsteg inom den enklare formen av artificiell intelligens där maskiner utför mycket specialiserade uppgifter som att känna igen bilder eller förstå talat språk. Begreppen ”känna igen” och ”förstå” ska inte övertolkas här, dagens AI kan inte sägas förstå språk i någon djupare semantisk mening, men har lärt sig mönster i språket genom att tränas på enorma mängder text och ljud. Genombrott i s.k. generell AI där maskiner, likt människor, kan generalisera kunskaper till helt nya domäner har visat sig avsevärt svårare. En del AI-forskare föredrar därför att tala om utökad intelligens (eng. augmented intelligence) där maskiner (t ex mobiltelefoner) utökar en människas intelligens.

Maskininlärning (ML) är vetenskapen om metoder och algoritmer som gör det möjligt för en maskin eller mjukvara att använda data för att lära sig om världen och fatta optimala beslut under osäkerhet. Det klassiska exemplet är en robot som lär sig om omgivningen från ett antal sensorer och kontinuerligt fattar beslut för att uppnå ett mål. Ämnet överlappar med reglerteknik och signalbehandling. Både namnet maskininlärning och detta klassiska exempel målar dock upp en alltför snäv bild av ämnesområdet. Den stora majoriteten av tillämpningar av maskininlärning involverar inte fysisk hårdvara, utan handlar snarare om olika former av mjukvara med automatiserade prediktioner och beslut. Det kan t ex handla om s.k. appar som rekommenderar filmer åt filmkonsumenten, ger läkaren diagnosförslag baserat på patientmätningar, hjälper jordbrukaren med råd om gödsling från sensorer på åkern, eller utför automatiserad försäljning av aktier. ML har även börjat användas inom samhällsvetenskap och (digital) humaniora.

Under senare år har även termen *Data Science* (DS) populariserats och det finns en stor arbetsmarknad för s.k. data scientists. DS verkar betyda olika saker för olika personer. Det finns en stor spridning i examina på personer som titulerar sig som data scientists, allt från datavetare utan några större statistiska ämneskunskaper till personer med master- eller t.o.m. doktorsexamen i statistik/matematisk statistik. Det finns också alternativa titlar, som data analyst och data engineer, på olika delar av spektrumet från enkel datainsamling till avancerad dataanalys, som

Det är en balansakt att välja vilka delar av dagens kursutbud som ska bort och var betoningen i undervisningen ska läggas.

tolkas lite olika beroende vem man frågar. Kvalificerade statistiker som använder titeln data scientists eller liknande har i allmänhet större kunskaper i datahantering, programmering och algoritmer än den genomsnittlige statistikern.

I den vidaste definitionen handlar maskininlärning i princip om alla former av datadriven prediktion och beslutsfattande under osäkerhet. Och någonstans här börjar termen bli kontroversiell för oss statistiker. Det har ju vi statistiker hållit på med i minst 100 år! Jag är den förste att hålla med om att mycket inom ML är traditionella statistiska metoder under nya namn. Jag är också rätt less på den överdrivna tilltron till maskininlärning som ett slags trollstav, som kan omvandla brusiga skräpdata till precisa prediktioner. Men jag välkomnar också hur nya ämnen som ML utmanar statistikämnet med nya perspektiv, och tvingar oss åtminstone till självreflektion, men förhoppningsvis också till förändring. För även om statistik och ML arbetar med samma palett av inferens, prediktion och beslut, så är betoningen på olika delområden ofta radikalt olika. ML har mycket att lära sig från oss statistiker vad gäller modellering och inferensteori. Men vi statistiker kan också lära oss en hel del från ML om effektiva beräkningar, storskalig datahantering och om nya högintressanta tillämpningar och probabilistiska modeller för komplexa problem.

»»»

Sex områden där jag upplever att statistikämnet bör förändras

»»» **SEX OMRÅDEN DÄR JAG UPPLEVER ATT STATISTIKÄMNET BÖR FÖRÄNDRAS**

1. Prediktion och beslut

I traditionella statistikkurser hanteras prediktion ofta mycket styvmoderligt, t.ex. som en lite hastigt presenterad punktprediktion med tillhörande prediktionsintervall i regressionsanalys. Prediktion borde istället ha en central roll och genomsyra det mesta på våra kurser.

Beslutsfattande under osäkerhet är ofta den verkliga slutprodukten av en statistisk analys och borde också ha en mycket större roll i våra kurser. Inte bara för att moderna tillämpningar ofta kräver automatiserat beslutsfattande, men även för att beslutsperspektivet ger en tankeram som underlättar modellval och andra beslut under inferensfasen. Inom ML handlar det mesta om prediktion och beslut, och det får stora konsekvenser för modellering och inferens. Hypotestest används t ex i mycket mindre omfattning inom ML. Vi borde fundera på om vi ska fortsätta att ge så mycket utrymme åt hypotestest i våra kurser. Ett problembaserat beslutsperspektiv skulle automatiskt skifta fokus mot effektstorlekar.

Prediktion och beslut ger också konkreta motiverande exempel för studenter. Jag utvecklade för några år sedan en grundkurs i statistik där jag inleder kursen med att demonstrera prediktion av handskrivna siffror utifrån pixelerade bilder. Jag estimerar en prediktionsmodell i R från 50 000 handskrivna siffor och visar att modellen kan användas för att prediktera en sannolikhetsfördelning över siffrorna 0-9 för en ny pixelerad bild. Jag diskuterar automatiserat beslut utifrån modellen. Exemplet visar hela kedjan av data-sannolikhetsmodell-inferens-prediktion-beslut och förmedlar att statistik används för att lösa verkliga problem. Det finns inte en student som undrar ”vad statistik ska vara bra för” efter den demonstrationen. De exakta detaljerna kan naturligtvis inte förklaras vid första kurstillfället, med jag återkopplar regelbundet till exemplet under kursens gång.

2. Flexibla modeller och regularisering

Modeller i ML är ofta mycket mer flexibla än de modeller som vi lär ut på våra statistikkurser. Med flexibla modeller menar jag rikt parametriserade modeller som kan anpassa en stor klass av funktionella samband och fördelningar. Flexibla modeller behövs i ML där problemen ofta kan vara kraftigt olinjära och icke-Gaussiska. Modellerna är i princip alltid överparametriserade men kombineras med regularisering för att undvika överanpassning. Regularisering kan ses som en komplexitetskostnad som tas hänsyn till vid estimationen, t ex via penalized likelihood. De mest kända varianterna är L1- och L2-regularisering, som leder till Lasso respektive ridge estimatorer. Med regularisering blir den faktiska modellkomplexiteten ofta mycket lägre än antalet parametrar. ML-områdets inriktning mot prediktion och beslut minskar behovet av att kunna tolka individuella koefficienter och att beräkna deras signifikans i dessa komplexa modeller. Istället pågår forskning om hur man ska kunna förstå prediktioner och beslut från komplexa ML-system. Vid en olycka måste en självkörande bil kunna förklara *varför* den valde att väja för ett objekt på vägen, dvs vilka sensordata triggade just det beslutet? Även kausalitet har fått en alltmer framträdande roll inom ML under senare år.

Vi bör fortsätta att lära våra studenter att börja en analys med enkla modeller, som ofta fungerar tillräckligt bra. Men vi bör i mycket större utsträckning undervisa om mer komplexa modeller och regularisering, och sluta att föra vidare den överdrivna skräcken för nominell komplexitet som är utbredd bland många statistiker.

3. Programmering, simulering och andra verktyg

En data scientist har ofta en kompetens som företag lätt kan se: de har en vana att arbeta med dataanalys i moderna programmeringsspråk och datorverktyg. Det är svårare för ett företag att identifiera och förstå vikten av en *statistikers* kärnkompetens. Det är viktigt att vi även lär ut programmering och andra verktyg mer ordentligt. Datavetare pratar om *computational thinking* som en problemlösningsprocess med datorn som hjälp. Det handlar inte bara om att lära sig grunderna i t.ex. R, eller andra populära språk för datanalys, som Python, utan ett annat arbetssätt, där datorn spelar en central och helt naturlig roll. Våra kurser borde innehålla inslag av mjukvaruutveckling, olika typer av verktyg för rapportskrivning där kod integreras med text och matematik, versionshanteringssystem, felsökning, kodtestning, prestandaoptimering, länkning till andra programmeringsspråk etc. Dessa inslag kan visserligen undervisas av datavetare, men jag tror att det är viktigt att vi statistiker själva integrerar delar av detta i vår undervisning. Det finns nämligen ett intrikat samspel mellan statistiken och dessa färdigheter, och även vi lärare bör vara moderna statistiker. Vi har nyligen startat månatliga möten vid Stockholms universitet där personalen, inkl. doktorander, presenterar olika typer av datorbaserade verktyg och metoder för varandra.

Jag har många gånger imponerats över hur snabbt studenter lär sig dessa färdigheter när de väl får chansen. Datorer och programmering är naturligtvis inget substitut för matematisk analys utan ett komplement som effektiviserar inlärningen, speciellt om det kombineras med simulering. Det ska falla sig naturligt för studenter att gå hem och skriva en liten programkod som simulerar data från modellen som presenterades vid dagens föreläsning. För många studenter är det här ett utmärkt sätt att förstå sannolikhetsmodeller och deras egenskaper.

Programmeringsvana studenter öppnar också upp för datorbaserad examination. På några av mina kurser får studenterna ta med sig kod från datorlaborationer till tentamen i datasal, där de sedan löser problem genom att kombinera matematiska beräkningar på papper med datorberäkningar i ett modernt programmeringsspråk.

4. Beräkningseffektivitet och stora data

Moderna statistiska problem med stora datamängder och komplexa modeller skattas som regel med iterativa optimerings- eller simuleringsalgoritmer. I många tillämpningar sker analysen också i realtid med s.k. strömmande data som anländer sekventiellt med mycket hög frekvens. Beräkningstiden för att estimera en modell på stora, snabba datamängder är därför ofta helt avgörande för den statistiska analysen. Jag har ofta fascinerats över vilka oväntade aspekter som bestämmer om en beräkning kan genomföras inom rimlig tid. Utvecklingen av modeller och inferensmetoder i min egen forskning har ofta drivits av beräkningsaspekter. När man arbetar med stora komplexa datamängder inser man snabbt vikten av ett samspel mellan statistisk modellering och beräkningar. Datavetare och numeriska matematiker tränas hårt i att förstå komplexitet och exekveringstider för olika algoritmer och numeriska metoder. En del av detta bör också ingå som en naturlig del i statistikundervisningen.

5. Nya datatyper, brus och anomalier

Våra kurser tenderar att presentera data som en prydlig tabell där raderna är observationer på ett antal kolumnvariabler. Men många data kommer allt oftare i avsevärt konstigare former, t ex som högupplösta bilder, datorgenererade loggfiler eller komplexa sensormätningar, eller som text i elektroniska dokument på webben. System som automatgenererar stora mängder data innehåller ofta mycket svårhanterligt brus, en stor andel saknade observationer och outliers. Vi bör därför lära våra studenter hur man samlar in, tvättar och representerar data inför statistisk analys, och hur dessa aspekter påverkar modellval och estimation. Här finns också en viktig roll för försöksplaneringen. Idag samlar företag in data mer eller mindre oplanerat, ofta helt drivet av vilken typ av mätteknik de råkar ha tillgänglig. Resultat blir alltför ofta att de insamlade datamaterialen är helt fel för problemet man vill studera med ML, eller av så dålig kvalitet att de blir statistiskt värdelösa. Här är statistisk kompetens ovärderlig och vi bör träna våra studenter i att identifiera rätt data för rätt problem, och rätt modell för en given datatyp och datakvalitet. Här kommer återigen beräkningseffektivitet in som ett viktigt kriterium vid val av modell och estimationsalgoritm. Allt hänger ihop.

6. Bayes

Bayesiansk inferens spelar en mycket stor roll inom maskininlärning. Många av de mest populära läroböckerna i maskininlärning har ett bayesianskt perspektiv. Bayes ses som det naturliga sättet att hantera prediktion, beslutfattande och regularisering. Många attraheras också av möjligheten att kombinera data med apriorinformation. En stor anledning till att Bayes är så populär i ML-kretsar är att regularisering med fördel kan formuleras som en aprioriförd elning, t.ex. är L1-regularising ekvivalent med en Laplacefördelning som prior. En aprioriförd elning används ofta också för att lägga in subjektiv information om någon slags mjukhet för ett flexibelt parametriserat funktionssamband, som annars riskerar att överanpassa data. Komplexa modeller för intrikata datatyper i ML-tillämpningar är också ofta lättare att hantera bayesianskt via simulering från posteriorfördelningen med t ex Markov Chain Monte Carlo (MCMC) eller genom s k stochastic variational inference speciellt anpassat för stora datamängder. Bayesiansk inferens bör vara en integrerad del av modern statistikundervisning. Jag har själv undervisat en grundkurs i statistisk inferens utifrån ett bayesianskt perspektiv och det går alldeles utmärkt.

AVSLUTANDE ORD

Strålande tider, härliga tider!

Jag har försökt beskriva min syn på maskininlärningsområdet och hur jag tror att det kan fungera som en katalysator för förändring. I några fall handlar det om rätt små förändringar i befintliga kurser, i andra fall om rejäla omtag, men även att helt nya kurser bör ersätta andra förlegade kurser. Det här arbetet är en balansakt där vi måste vara försiktiga med att inte förlora vårt ämnes matematiska grund. Vi måste noga välja vilka delar av dagens kursutbud som ska bort och var vi ska lägga betoningen i undervisningen. Men det är hög tid att ompröva gamla sanningar nu. Kill your darlings!

Jag vill också påpeka något uppenbart: om vi statistiker ska undervisa i maskininlärning och datordriven dataanalys, så bör vi även *forska* inom dessa områden. Jag har under de senaste 2-3 åren börjat se tecken på en positiv förändring här, där ett åtminstone ett fåtal svenska statistiker och probabilister har börjat publicera sig i de ledande ML-forumen. I statistikfältet finns det också flera tidskrifter i computational statistics av hög kvalitet där även ledande ML-forskare publicerar sig.

ML-området är extremt dynamiskt och har genomgått stora förändringar under de senaste 4-5 åren. Neurala nätverk har gjort en spektakulär återkomst, speciellt inom bilder, text och ljud, och det har utvecklats en stor ingenjörsdriven subkultur inom ML för denna typ av ansats. Det kommer alltså bli viktigt att skilja ut verkliga genombrott från tillfälliga trender, men även att tillåta en viss flexibilitet i kursinnehållet.

Min prediktion är att statistikämnet kommer att genomgå stora positiva förändringar under detta nya decennium, delvis som ett resultat av det upplevda hotet från angränsande ämnen som maskininlärning. Det är strålande tider, härliga tider!

MATTIAS VILLANI,
PROFESSOR I STATSTIK VID
STOCKHOLMS OCH LINKÖPINGS UNIVERSITET

(Artikeln ovan finns inom kort även att läsa på https://statfr.blogspot.com/)

AI-utvecklingståget har under 2010-talet nått högre fart än någonsin tidigare, och kan väntas accelerera ytterligare. Ombordstigningen kan vara ett gott tillfälle att bilda sig en samlad uppfattning om hela den AI-utveckling man är med och bidrar till, anser Olle Häggström.

AI-utvecklingen och

Tänk på konsekvenserna

AI-utvecklingståget har under 2010-talet nått högre fart än någonsin tidigare, och kan väntas accelerera ytterligare. När jag ser mig omkring bland kollegor i matematisk statistik är det allt fler som valt att hoppa på tåget, i hopp om att bli delaktiga i en spännande utveckling, och få del av uppmärksamheten och miljarderna. I vissa fall handlar det mest om att sätta ny etikett på den forskning man redan bedrivit i årtionden, men ofta är det en mer uppriktigt avsedd omorientering.

En forskare är alltid skyldig att ur etisk synvinkel beakta de potentiella konsekvenserna av sin forskning – en ståndpunkt jag nyligen utvecklade i Häggström (2019). För den teoretiskt inriktade forskaren kan detta framstå som så abstrakt att det blir oöverstigligt, och reflektionen urartar ofta i ett slags koketterande över den egna verksamhetens brist på tillämpning, med matematikern G.H. Hardy som förebild, vars *A Mathematician's Apology* från 1940 förvisso är välformulerad och läsvärd, men till sitt budskap likväl helt åt skogen (Hardy 1940).

Ombordstigningen på AI-tåget kan vara ett gott tillfälle att börja ta konsekvenstänkandet och etiken på allvar. För den vars projekt handlar om någon viss AI-tillämpling behövs naturligtvis överväganden specifikt kring denna tillämpning, men det är också viktigt att bilda sig en samlad uppfattning om hela den AI-utveckling man är med och bidrar till. Det sistnämnda är precis vad jag med denna text (och inte minst dess källhänvisningar) vill hjälpa läsaren att komma igång med. Att få en samlad bild av AI-utvecklingen och dess möjliga konsekvenser är dessutom viktigt inte bara för hugade AI-forskare utan mer allmänt också för varje samhällsengagerad medborgare med ambitionen att förstå samtiden och de faktorer som kan komma att avgöra vår framtid.

Positivt och negativt

Samhället och våra liv är på väg att präglas allt mer av olika AI-system, på gott och ont.

Till de uppenbara positiva sidorna hör hur AI som beslutsverktyg inom medicinsk diagnostisering kan komma att få enorma positiva effekter på sjukvården och vår hälsa, och liknande förbättringar är att vänta i en rad andra branscher. Ekonomiska bedömare räknar med att innovationer inom AI och robotik kommer att stå för en betydande del av den ekonomiska tillväxten det närmaste årtiondet, och på längre sikt är möjligheterna i princip obegränsade.

I den negativa vågskålen ligger exempelvis hur vi idag tycks stå på randen till en mycket farlig kapprustning inom militär AI-teknik för så kallade autonoma vapen, även kända (med en något mindre artig term) som mördarrobotar. Massförstörelsepotentialen i sådan teknik har att göra med skalbarhet: 100 000 Kalashnikovs kan inte användas till att döda 100 000 gånger fler fiender än en enda med mindre än att vi har lika många soldater att sätta vapnen i händerna på. Om vi ersätter Kalashnikovs med mördarrobotar fungerar däremot motsvarande uppskalning av dödlighetskapacitet utan att på samma sätt behöva öka den militära personalstyrkan. Och om vi istället för Kalashnikovs drar en parallell till kärnvapen, som vi ju ändå lyckats leva med i över 70 år utan att förgöra oss själva, så ligger den stora faran i att medan kärnvapen lyckligtvis kräver storindustriella processer – vilket är grunden till att de ännu är ett tiotal nationalstaters exklusiva egendom – så är spridning av militär AI-teknik ojämförligt mycket svårare att förhindra.

Farligt på ett annat sätt är det slags

forskarens ansvar

AI-programvara som vi inom kort kan få i förlängningen av exempelvis den textgenerator GPT-2 som OpenAI offentliggjorde i februari 2019. GPT-2 och dess 1,5 miljarder parametrar hade tränats och anpassats med hjälp av texter på miljoner utvalda webbsidor, och fungerar på så vis att givet en inledande textsnutt, som inte behöver omfatta mer än en eller ett par rader, så utvecklar programmet en längre text som bedöms passa som fortsättning på inledningsraderna. En prompt som ”Hillary Clinton and George Soros” räcker för att det skall gå loss i vildsinta konspirationsteorier. Programmet, om än långt ifrån fulländat, visar häpnadsväckande prestanda i att identifiera och imitera olika genrer och textstilar samt generera (någorlunda) trovärdiga fortsättningar på givna textinledningar.

Tidigare hade OpenAI gjort sina olika produkter fritt tillgängliga, men denna gång bedömde de att även om GPT-2 har stor potential att användas för goda syften, så banar dess teknik också väg för så pass allvarligt missbruk av olika slag – de nämner författande av falska nyhetsartiklar och vilseledande online-imitation av specifika personer, samt automatiserad produktion av såväl fejkat innehåll på sociala medier som spam och phishing-försök – att ett mer restriktivt tillvägagångssätt var motiverat. I syfte att ge forskarsamhället och regeringar en respit för att förbereda motåtgärder valde de att släppa enbart en mindre kraftfull version av programmet [Radford m.fl. 2019]. Mer än en temporär och ganska kort respit kunde det inte rimligtvis bli eftersom andra AI-utvecklingsföretag knappast ligger långt efter, och i november 2019, blott nio månader efter den ursprungliga lanseringen, släpptes den fullskaliga versionen av programmet.

Etiska och andra samhällsproblem

Viktigt att förstå rörande AI-utvecklingen mer allmänt är att även om programvaran fungerar precis som vi hoppats, och även om vi mot förmodan lyckas undvika att den hamnar i händerna på IS-terrorister, phishing-bedragare och ryska trollfabriker, så finns likväl potentiella samhällskonsekvenser som ger upphov till komplicerad etisk problematik.

Ett ganska uppenbart fall rör automatisering och arbetsmarknad. Mycket av den AI-utveckling som ekonomer sätter sin tillit till vad gäller ekonomisk tillväxt handlar om att skapa maskiner som kan ta över arbetsuppgifter från oss människor och utföra dem snabbare, bättre och billigare än vi själva förmår. Detta kan få effekter på arbetslösheten, och i ganska omedelbar förlängning ekonomisk ojämlikhet. När en maskin tar över en människas arbete ersätts en löneinkomst (för den som tidigare gjorde arbetet) av en kapitalinkomst (för den som äger maskinen), och eftersom kapitalinkomster är mer ojämnt fördelade än löneinkomster tenderar detta att leda till ökad ekonomisk ojämlikhet – förutsatt att inte ekonomisk-politiska åtgärder sätts in för att omfördela och mildra denna effekt. Jag tror att vi lugnt kan räkna med att utvecklingen av autonoma fordon leder till en revolution av transportbranschen där alla chaufförsyrken inom ett par decennier kommer att ha försvunnit från arbetsmarknaden, och en liknande utveckling kan komma även i andra branscher.

Ändå är det inte självklart att den väntade AI-drivna automatiseringsvågen också kommer att leda till det som kallas teknologisk arbetslöshet, med vilket menas stigande arbetslöshet orsakad av rationalisering till följd av tekniska framsteg. Att människors arbetsuppgif-

ter övergår till maskiner är ju inte något nytt fenomen – tvärtom har det pågått genom hela historien och i synnerhet sedan 1800-talets industrialisering. Vid 1800-talets mitt arbetade cirka 75% av den svenska arbetskraften inom jordbruket, en siffra som idag är nere i cirka 3%, men denna utveckling har inte försatt 72% i arbetslöshet. De gick istället till andra sektorer, inledningsvis främst industrisektorn, därefter allt mer till tjänstesektorn. Något liknande skulle vi kunna tänka oss exempelvis för alla de buss-, taxi- och lastbilschaufförer som kan väntas förlora sina jobb i den kommande omvandlingen av transportsektorn.

Men kommer detta – att arbetstillfällen i nya sektorer uppstår i ungefär samma takt som anställda i gamla sektorer konkurreras ut av maskiner – att förbli fallet framöver? Denna ungefärliga jämvikt kan gott tänkas bestå åtminstone några decennier framåt, men saken är knappast självklar med tanke på ett antal nya omständigheter, som att det idag inte längre bara är fysiska och manuella arbetsuppgifter som rationaliseras bort, utan allt oftare även intellektuella. Ett ofta citerat exempel på det sistnämnda är automatiserad journalistik, och personligen har jag svårt att tro att vi i evighet skall förmå hitta nya arbetsuppgifter där den mänskliga förmågan överstiger maskinernas.

Jag vill inte diskutera frågan om automatisering och arbetsmarknad utan att också prova att vända på perspektivet. Lönearbete kanske trots allt inte är livets mening? Kanske bör vi snarare se automatiseringen som en befrielseprocess? Så låt oss försöka föreställa oss ett utopiskt samhälle där all vår tid omvandlats till fritid som vi kan ägna åt vad vi vill. På tillräckligt lång sikt ser jag inte detta som någon omöjlighet. Det kräver dock noggrann eftertanke, ty även om ett samhälle med 100% arbetslöshet i viss mening är jämlikt, så är det inte uppenbart hur ett sådant samhälle kan organiseras. Och vad kanske värre är: om det är dit vi vill så behöver vi också ha en fungerande plan för vägen dit, vilken rimligtvis går via arbetslöshetsnivåer på 20%, 50% och 90%. Hur kan vi passera sådana övergångsstadier utan att vidga de ekonomiska klyftorna ofantligt och utan att skapa social instabilitet? Det här är svåra frågor som vi inte har tydliga svar på idag.

Erodering av mänskligt beslutsfattande

Den israeliske historikern och författaren Yuval Noah Harari tar upp det senast nämnda problemkomplexet i sin bok *21 Lessons for the 21st Century* (Harari 2018), som ett led i en bredare genomgång och syntes av de ofantliga samhällsutmaningar vi står inför. Vad gäller AI-utvecklingen breddar han frågan om huruvida AI-tekniken kan väntas ta våra arbeten till den ur existentiell synvinkel kanske ännu mer skrämmande frågan om huruvida den till och med kan komma att ersätta allt mänskligt beslutsfattande.

Harari börjar med att skissera ett till synes oskyldigt tankeexperiment. Antag att en grupp vänner samlas hemma hos en av dem för att ha en trevlig filmkväll tillsammans. Vilken film skall de se? Den frågan kan ge upphov till lång diskussion kring hur de skall jämka sina respektive filmsmaker, och den film de till slut fastnar för kan visa sig vara ett lyckat eller ett mindre lyckat val. Att överlåta valet till en AI-algoritm skulle kunna erbjuda en mycket säkrare väg till en film som tillfredsställer alla i gruppen. Ett uppenbart sätt vore att var och en i gruppen ger en lista över sina favoritfilmer som indata till algoritmen, vilken härur härleder deras preferensprofiler och letar fram någon rulle ur arkiven som tycks passa dem alla. Nackdelen med en så grovhuggen algoritm är dock, som Harari påpekar, att självrapportering inte är något särskilt pålitligt sätt att få veta en persons verkliga preferenser. När jag får höra av än den ene och än den andre vilket mästerverk *Gudfadern* är, och den sedan inte faller mig på läppen, kan det bli så att jag efteråt ändå meddelar vilken fantastisk upplevelse filmen var, för jag vill ju inte gärna framstå som okultiverad, och till slut kanske jag till och med lurar mig själv.

Mycket bättre vore om AI:n fick tillgång till data om vilka filmer vi verkligen sett på TV, och sett färdigt. En kamera kopplad till TV-skärmen skulle därtill kunna läsa av våra ögonrörelser och ansiktsuttryck för att utröna vilka scener som fyller oss med känslor och vilka vi är likgiltiga inför, och algoritmen kan bli ännu mer träffsäker om den visuella informationen kompletteras med hjälp av biosensorer med uppgifter om hjärtrytm, blodtryck och hjärnaktivitet. Det går att tänka sig att en totalitär stat (i något slags förlängning av det pågående kinesiska försöket med ett socialt kreditsystem) påtvingar oss sådan intim övervakning, men mycket tyder på att vi kommer att gå in i det frivilligt, inför utsikten om bättre filmupplevelser och andra vardagliga fördelar.

Beslutet om vilken film vi skall se är förstås en trivialitet, och priset för att välja fel är inte särskilt högt. Harari påpekar dock att AI med den här sortens inblick i våra inre kan komma att hjälpa oss med långt mer livsavgörande val. Vilken universitetsutbildning skall jag gå? Skall jag försöka gå vidare efter en halvlyckad första dejt i hopp om ett mer allvarligt förhållande och i förlängningen kanske äktenskap och familj? Den här sortens beslut tas i hög grad med intuition och magkänsla, och träffsäkerheten är högst begränsad, men med AI-teknik kan vi få hjälp att göra långt bättre och mer välgrundade val. Frågan är om vi verkligen kommer att överlåta dessa beslut på våra AI-hjälpmedel, men Harari föreslår att i takt med att vi lär oss att varje försök att avvika från AI-rekommendationerna sannolikt blir som att skjuta sig själv i foten, så kommer vi allt mer slaviskt att göra som de föreslår.

En tillvaro som känns meningsfull helt

utan eget beslutsfattande har jag svårt att föreställa mig, men kanske kan vi skapa avgränsade arenor där vi har utrymme att fatta beslut och där våra felgrepp inte får så allvarliga konsekvenser. Betyder detta i så fall att vår framtid finns i datorspelen? AI-forskaren Stuart Russell vid UC Berkeley reflekterar i avslutningskapitlet till sin aktuella bok *Human Compatible: AI and the Problem of Control* (Russell 2019) kring det slags erodering av mänskligt beslutsfattande vi här talar om, och som kan bli fallet även om vi lyckas programmera AI att sätta våra intressen i högsätet. Den mänskliga kulturen har gradvis genom årtusendena ackumulerats, tack vare att varje generation genom en omsorgsfull utbildningsprocess lärt upp nästa, men nu kanske vi står inför en epok där kunskapen istället förvaltas av maskiner, och mänskligheten blir till lata och viljelösa passagerare på ett kryssningsfartyg där allt tas om hand av dessa maskiner – som i filmen *WALL-E*. Om vi vill undvika en sådan utveckling behövs ett sätt att finna en balans, där maskinerna assisterar oss lagom mycket för att våra liv skall förbättras, samtidigt som vi inte förlorar den agens och handlingskraft som behövs för att inte göra livet tomt och meningslöst. Russell avslutar:

> Varje småbarnsförälder vet vad jag talar om. Så snart barnet passerat det hjälplösa spädbarnsstadiet krävs en balans mellan att göra allt för barnet och att lämna det att lösa alla sina problem själv. Vid ett visst stadium inser barnet att föräldern är fullt kapabel att knyta dess skosnöre men väljer att avstå. Är detta mänsklighetens framtid – att behandlas som ett barn, för alltid, av överlägsna maskiner? Kanske inte, […] och troligen inte heller som djur på ett zoo. Det verkar inte finnas någon bra analogi i dagens värld till den framtida relationen mellan oss och de maskiner som verkar i vårt intresse. Hur detta skall sluta återstår att se.

En bättre introduktion än *Human Compatible* till de tekniska, etiska, politiska och sociala frågorna kring framtida AI-utveckling finns i dagsläget inte. Och Russell nöjer sig inte (så som jag mestadels gjort i denna modesta text) med att skissera problemen, utan han rullar upp skjortärmarna och gör sitt bästa att visa hur vi, om vi tar oss samman, kan lösa dem.

En ganska stor del av boken ägnar Russell åt ett problem jag här undvikit, nämligen hur vi förhindrar att maskinerna sedan de uppnått övermänsklig allmänintelligens plötsligt en dag konstaterar att vår existens är inkompatibel med deras mål (eller helt enkelt irrelevant) och gör slut på oss. Skälet till att jag undvikit det är att jag av erfarenhet vet att en stor del av läsekretsen omedelbart kommer att avfärda ett sådant scenario som science fiction, och att det krävs många sidor för att övertyga om att det i själva verket är en risk värd att ta på allvar. Russell gör det bättre och mer övertygande än någon annan gjort, åtminstone sedan filosofen Nick Bostroms inflytelserika *Superintelligence* (Bostrom, 2014), eller kanske någonsin. Bostroms bok är fortfarande läsvärd men något mer akademisk och mindre lättillgänglig än Russells.

Slutord

Jag inser att de aspekter på AI-futurologi, AI-risk och AI-etik jag här diskuterat kan tyckas överväldigande för den enskilde AI-forskaren, och att det kan vara frestande att stoppa huvudet i sanden och se sig själv som en i det stora hela betydelselös kugge i maskineriet som blott gör sitt jobb. Att ge efter för den impulsen duger emellertid inte, ty gör man det är man bara en liten lort.

OLLE HÄGGSTRÖM, CHALMERS

(Artikeln ovan finns inom kort även att läsa på https://statfr.blogspot.com/)

Referenser

Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford.

Harari, Y. N. (2018) *21 Lessons for the 21st Century*, Jonathan Cape, London.

Hardy, G.H. (1940) *A Mathematician's Apology*, Cambridge University Press, Cambridge, UK, https://www.math.ualberta.ca/mss/misc/A%20Mathematician%27s%20Apology.pdf

Häggström, O. (2019) Vetenskap på gott och ont, i *Vetenskaplig redlighet och oredlighet* (red B. Lindberg), Kungliga Vetenskaps- och Vitterhets-Samhället, Göteborg, s 71-85, http://www.math.chalmers.se/~olleh/EtikKVVS.pdf

Radford, A., Wu, J., Amodei, D., Amodei, D., Clark, J., Brundage, M. och Sutskever, I. (2019) Better language models and their implications, OpenAI, 14 februari, https://openai.com/blog/better-language-models/

Russell, S. (2019) H*uman Compatible: Artificial Intelligence and the Problem of Control*, Viking, New York.

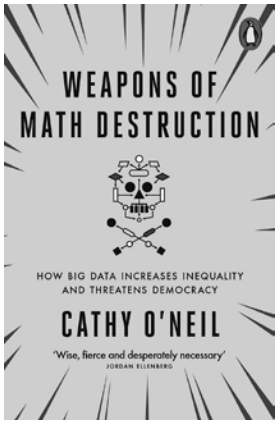
AI skapar behov av etiska riktlinjer

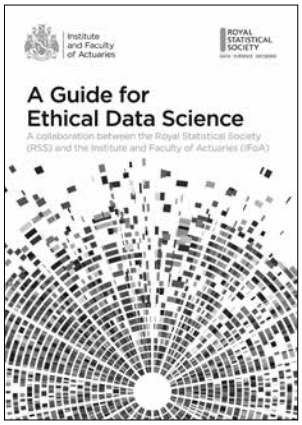
Jag heter Maria Karlsson och är universitetslektor i statistik vid Umeå universitet och för närvarande också Svenska statistikfrämjandets vice ordförande. Denna text skriver jag dock såsom privatperson och helt vanlig statistikfrämjare. Det innebär att jag inte varit så rigorös i mina efterforskningar kring källor etc. som jag vill och måste vara i min yrkesroll. Jag vill heller inte påstå att det jag skriver om är något som jag egentligen har någon speciell sakkunskap om. Det här är istället en text som mer bygger på magkänsla och de tankar som poppar upp i huvudet just nu. Jag tänker nämligen skriva lite om hur jag funderar kring hur AI, och i synnerhet maskininlärning med stora datamängder, påverkar oss statistiker och statistikfrämjare när det gäller de etiska överväganden som måste göras.

Jag tänker då inte på AI i form av eventuella framtida självstyrande vapen (”mördarrobotar”), som ofta kommer upp i media när riskerna med AI diskuteras. Nej, jag tänker mer på hur användandet av olika typer av maskininlärningsalgoritmer (men också mer ”gammeldags” statistiska modeller för prediktioner) som grund för automatiserade beslut kan få betydande konsekvenser, både positiva och negativa, för individerna de berör. Förändrar det hur vi behöver tänka kring etik? Ja, troligtvis, är mitt svar.

Många av mina tankar kring detta har sin upprinnelse från en bok jag läst och som jag varmt vill rekommendera. Boken heter *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* och är skriven av Cathy O’Neil (Penguin Books, 2017). Boken är mycket tänkvärd och samtidigt lättläst p.g.a. dess populärvetenskapliga karaktär. Författaren är en matematiker som har disputerat vid Harvard och som jobbat som data scientist bl.a. i den finansiella sektorn. Jag fick kännedom om boken genom kollegor som deltagit i ISI World Congress 2019. Många av sessionerna där behandlade statistik i gränslandet till data science och/eller maskininlärning, och på många av dessa sessioner så kom boken på tal när etiska frågeställningar i samband med den ”nya statistiken” diskuterades.

I boken ges många exempel på hur maskininlärningsalgoritmer och statistiska modeller baserade på stora datamängder används i praktiken i det amerikanska samhället. Det gäller allt ifrån individualiserad reklam och riktade politiska kampanjer till utsortering av vilka som kallas till jobbintervjuer, vilka som får kredit, vilken avgift som måste betalas för vissa tjänster, längden på fängelsestraff, lärares löner och fortsatta anställning utifrån elevers testresultat o.s.v.





Två lästips om de etiska frågeställningarna.

Ett vanligt problem, som O’Neil lyfter fram, är icke-transparensen (eng. ”opacity”) hos modellerna/algoritmerna. Det vill säga att de individer som ”drabbas” av algortimerna/modellerna inte vet vilka faktorer som gör att just de inte blir kallade på jobbintervjuer eller inte får ett lån eller får betala en väldigt hög skolavgift. De vet alltså inte vad de borde förändra för att ligga bättre till. Individer kan inte heller begära rättning av eventuella fel i informationen om sig själva, eftersom individerna inte vet vilken information som använts.

Det är också i hög grad så att modellerna/algoritmerna som används och besluten som tas baserat på dem inte ifrågasätts, eftersom de är just matematiska modeller och därför i folks ögon är för komplicerade att förstå. Matematiska modeller anses representera sanningen och därför vara rättvisa. Det finns lagar mot diskriminering, t.ex. att kreditbeslut inte får fattas baserat på en individs etnicitet. Men när algoritmerna/modellerna baseras på information som är korrelerad med etnicitet så kan ju diskriminering förekomma ändå. Fastän den då görs i den ”rättvisa” matematiska modellens namn, så att den inte märks och kan ifrågasättas av vem som helst.

Ett annat lästips som jag kan ge är A *Guide for Ethical Data Science* som publicerades av Royal Statistical Society (RSS) tillsammans med Institute and Faculty of Actuaries (IFoA) 2019. Det är ett dokument med etiska riktlinjer för data scientists (finns att läsa på nätet). Guidens riktlinjer ska ses som ett komplement till andra existerande riktlinjer och ska kunna utgöra ett praktiskt användbart verktyg för dem som jobbar med data science och de utmaningar detta medför. I dokumentet behandlas följande fem huvudprinciper:

1. Seek to enhance the value of data science for society.
2. Avoid harm.
3. Apply and maintain professional competence.
4. Seek to preserve or increase trustworthiness.
5. Maintain accountability and oversight.

Det är alltså, på sätt och vis, inget nytt under solen, men de praktiska exemplen och strategierna kopplade till huvudprinciperna gav i alla fall upphov till en del nya tankar hos mig. Det är kanske dags att se över, revidera och komplettera Svenska statistikfrämjandets etiska kod som har ganska många år på nacken.

MARIA KARLSSON
ENHETEN FÖR STATISTIK,
HANDELSHÖGSKOLAN
UMEÅ UNIVERSITET

The author of this article, Michael Sachs, collaborates with researchers from around Karolinska Institutet. His research interests include the development and evaluation of risk prediction models and biomarkers, statistical computing, causal inference in observational studies, and tools for reproducible research.

Some thoughts on AI in medicine and biostatistics

The idea of artificial intelligence has a long history. The ancient Greeks and Romans fantasized about inanimate objects and machines imbued with consciousness and intelligence through magic or clever design. Classical philosophers studied the nature of human consciousness and intelligence, some arguing that they are characterized by the logical manipulation and processing of information. In the 1950s and 60s, when computers could store and process data, that dream began to appear closer to reality. Here we are at the dawn of a new decade, after years of exponential growth in the availability of electronic data and the speed of processing thereof. How can we characterize the current state of AI and its relevance in the modern era of biomedical research?

The most basic form of AI is called an expert system. An expert system is a piece of software that encodes knowledge into a series of logical steps that responds to input in order to perform a well-defined task. These systems assume that intelligence can be characterized by a set of rules or criteria. For example, a program that implements official treatment recommendation guidelines. Or in the biostatistics field, a smart data visualization tool that suggests certain graphs based on the structure and format of the input data. Such systems can be quite useful, yet they may not be considered intelligent because they do not adaptively learn from new data.

On the next level up we have artificial narrow intelligence, which is a system that learns adaptively from training data in order to accomplish a specific task. In this class we have perceptrons, as they are called in the AI literature, that are software programs that have been trained to recognize complex input to which they ascribe meaning. Examples include handwritten character recognition, voice recognition and interpretation, and other image interpretation systems. In medicine, image recognition systems have been implemented in a variety of fields including oncology and orthopedics to augment traditional medical image interpretation.

How do these systems learn from data? This is where machine learning comes in. Machine learning is a collection of statistical and computational techniques that are designed to form predictions or classifications based on input data. The basic statistical problem they are solving is one of prediction. Compared to standard statistical prediction methods like regression, ML techniques often require much more data to be trained, and are better suited to applications with a near infinite pool of training data with a high signal to noise ratio (like image recognition). It is a misconception that they perform equally well in simple cohort or register based studies where the signal to noise ratio is low, even if there is a large sample size. Accurate predictions of a future or unobserved event is thought to be a fundamental requirement for a useful AI system, but accurate predictions alone are not enough. If there is a clearly defined space of actions in the decision making context, then a useful AI system must use the predictions and other information to make a helpful decision.

Recall that a system can be considered intelligent if it learns from data (machine learning), and then uses those learnings to achieve a specific goal through flexible adaptation. There is no shortage of reports in the medical literature of systems that have learned from data, and no shortage of statistical and machine learning methods for applying to data. What is lacking, however, are clear descriptions of the goals of the AI system, and emphasis on the methods for addressing those goals and evaluating the success of the proposed systems. Biostatisticians in Sweden are well-positioned to address these shortcomings.

In applied research, there is no shortage of tools available to do machine learning. Indeed, all biostatisticians have the tools available to them to perform machine learning, even though they may

»»» SOME THOUGHTS ON AI IN MEDICINE AND BIOSTATISTICS

not commonly be called ML methods. For instance, logistic regression can be used to train an AI system to make classifications based on input data. Many other complex ML methods are available as open source software packages, too numerous to list here. In designing an AI system for use in medicine, the main focus should be on the clinical context, and the intended use of the system. Then, if existing data can be used to inform the system, use the tools that are available and accessible to you to solve those problems. In medical research the usual goals are to diagnose disease, identify risk factors for disease, treat disease, prognosticate, prevent disease, and improve understanding of basic science. Prediction and AI systems can be used to achieve any of these aims. However, it should be noted that the scientific aim comes first, development of a prediction model in itself is not a scientific objective.

When technology companies or researchers develop AI systems to say, play a board or computer game, such as AlphaGo developed by Deepmind, knowledge about the rules of the game allow them to run as many experiments as they want to train the AI system. During repeated game trials, the AI system can try different strategies to learn what works best to achieve its objective, to win the game. In medicine, where the goal is to improve human health, it would neither be feasible nor ethical to run repeated experiments to train an AI system to optimize a treatment strategy. Instead, we often have to rely on data from completed studies to train AI systems, whether they be observational or interventional. In either case, in order for the AI system to be adopted in clinical practice, one needs to generate evidence to determine whether the suggested AI system achieves its objective. If the system is trained on existing data, that means that the treatment decisions that it suggests were not in fact used in reality, yet we need to know whether the use of those decisions would lead to clinical benefit for future individuals. This highlights the fact that causal inference is another important area of research for biostatisticians especially in the field of AI. The key question is whether the use of an AI system causes benefit on average in the population in which it is intended to be used. Short of a randomized controlled trial in which participants are randomized to either use the AI system for treatment decision making or not, careful consideration of confounding, generalizability, and methods of estimation are critical areas in applied and biostatistical research.

»This highlights the fact that causal inference is another important area of research for biostatisticians especially in the field of AI. »

What about biostatistics education? Novel and complex machine learning methods do not need to be prioritized, as these are rapidly evolving and there is no clear superior approach anyway. The current biostatistics curriculum meets the needs of predictive modeling with regression methods, and flexible extensions thereof. What does need additional emphasis are methods for evaluating predictive performance, such as split sample validation, cross-validation, and the bootstrap. Errors in predictive model validation are pervasive in the literature, and these can be tackled early on in introductory biostatistics courses. In advanced courses related to prediction, more emphasis can be placed on the decision making process using predictions, utility measures of those decision processes, and study designs for development and evaluation of predictions. In addition, causal thinking and statistical methods for causal inference are important in general, and equally if not more so in AI.

In conclusion, AI is an exciting field that is gaining in interest and applicability because of the continuing increase in computing power and availability of data. Biostatisticians need to be aware of the challenges in the field and remember that statistical thinking is just as important in AI as any other area of applied research. There are plenty of opportunities to contribute in biostatistical and applied research related to AI. The key to success is to not get lost in the jargon and excitement, and focus on what we biostatisticians know and do best.

MICHAEL SACHS
BIOSTATISTICAL RESEARCHER
THE CLINICAL EPIDEMIOLOGY UNIT
KAROLINSKA INSTITUTET

Tre tips för att lyckas med en AI-implementering

Inledning

AI (artificiell intelligens) är ett begrepp som de senaste åren diskuterats vitt och brett inom många olika branscher. Från att det varit mer av ett teoretiskt koncept går utvecklingen mot att fler och fler applikationer faktiskt blir verklighet. Media rapporterar om vikten av att följa med i den teknologiska utvecklingen, och företagen utforskar och utvecklar för att kunna dra nytta av den nya tekniken. För att lyckas med att inte bara utveckla utan även implementera finns det olika perspektiv som behöver tas i beaktande – här kommer mina bästa tips.

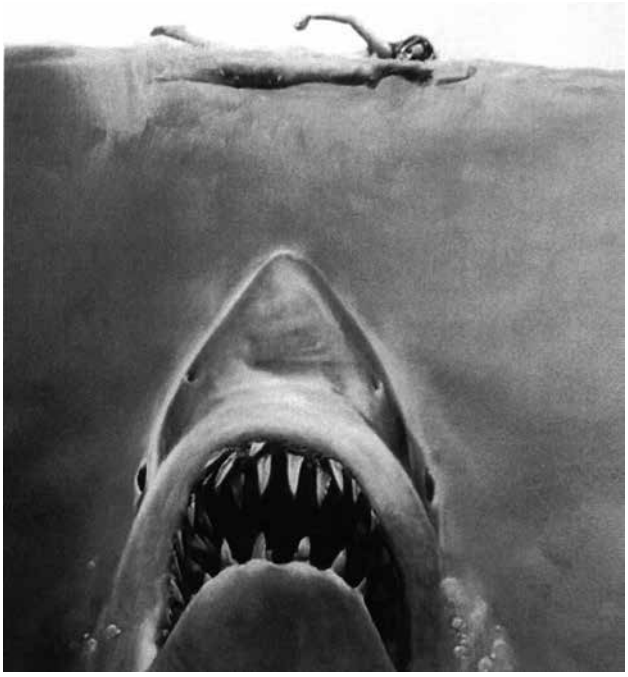


BILD: WIKIPEDIA

Om den enda bild man har av hajar är den man fått från filmen Hajen så är det inte så märkligt om man upplever rädsla för hajar. Liknande tendenser ser Josefine Hagbarth inom teknik, dvs. att känslan för tekniken baseras på icke helt rationella grunder.

1. Definiera tillämpningsproblemet

Jag har fokuserat min karriär på att arbeta med att implementera ny teknik. Jag använder begreppet ”ny teknik” just för att det, enligt min uppfattning, inte finns någon anledning att behandla *AI*-teknik på något annat sätt än annan teknik som är ny. Det som i mina ögon definierar ny teknik är att tekniken inte bara innebär införandet av ett nytt verktyg, utan att den också på något vis omdefinierar verksamheten. Då krävs det inte bara systemstöd utan även nya arbetssätt och strukturer för att den nya tekniken ska komma till nytta. I vissa fall kan det handla om förändringar i behov av arbetskraft, i andra fall kan det handla om att höja kapacitet eller kvalitet i verksamheten.

Här framträder en viktig poäng med denna typ av implementering: en verksamhet måste främst utgå från det man vill uppnå, från det problem man vill lösa, när man överväger införande av ny teknik. Den enkla anledningen till detta är förstås att olika typer av teknik löser olika typer av problem. Jämför med en anställningssituation, där är det sällan företaget uttrycker att det söker efter en ”människa” i största allmänhet. I regel preciseras att det är en kompetens i form av läkare, advokat, javautvecklare eller något annat, som företaget har behov av för att lösa de problem det är fråga om. Att tala om implementering av *AI* blir därför i många fall alltför brett för att kunna hanteras, när det gäller en implementering som ska fungera i verkligheten.

***AI* är snarare att** betrakta som ett paraplybegrepp som inkluderar ett stort antal olika tekniker och applikationer, där allt från robotgräsklippare till identifiering av potentiella kunder kan räknas in i begreppet om man så vill. Synen på *AI* är inte heller statisk, utan den tenderar att ändra sig med tiden, allteftersom utvecklingen går framåt. Ett uttryck för detta ges i det s.k. Teslers teorem från 1970, vanligen citerat som att ”*AI* is whatever hasn't been done yet”, eller fritt översatt: ”*AI* är det som vi ännu inte lyckats åstadkomma”. Teslers tanke anspelar på att vad vi tidigare föreställde oss att *AI* var inte längre stämmer med vad vi anser att det är idag. Något som tidigare ansågs kräva *AI*, som att detektera en viss typ av åkomma på en röntgenbild, kan idag avfärdas som att det ju bara är en viss typ av algoritm som kan analysera pixlar på rätt sätt – inte kan väl det vara det man menar med *AI*…? På det här sättet sker en avdramatisering av tekniken. Det blir mer konkret vad syftet är, och man ser vilket värde tekniken kan tillföra. Håller man med om Teslers synsätt blir det onekligen svårt att implementera *AI* – eftersom *AI* ur det perspektivet helt enkelt inte finns.

2. Skapa gemensamma definitioner

En annan problematik med att prata om *AI* som helhet är att varje person har en egen uppfattning om vad det betyder. Eftersom begreppet är så brett kan också uppfattningen skilja sig avsevärt mellan olika personer. Jag konstaterar att jag vid upprepade tillfällen möter både skepsis och ibland rädsla för att ”*AI* kommer ta över världen”. Min erfarenhet är att den synen är starkt representerad hos personer som inte har någon tidigare erfarenhet av att implementera den här »»»

»»» **TRE TIPS FÖR ATT LYCKAS MED EN AI-IMPLEMENTERING**

typen av teknik. Däremot har många sett science fiction-filmen Terminator. Fenomenet kan i vissa aspekter liknas med vad som uppstod efter filmen Hajen: Djuret blev inte farligare i samband med att filmen visades, men rädslan för hajar ökade kraftigt i och med den framställning som gjordes. Om den enda bild man har av hajar är den man fått från filmen så är det inte så märkligt om man upplever rädsla för hajar, trots att det finns ett antal varianter av hajar som är helt ofarliga för oss människor.

Liknande tendenser tycker jag mig se inom teknik, dvs. att känslan för tekniken baseras på icke helt rationella grunder. Detta förstärks av att definitionen av *AI* inte är entydig. När begreppet *AI* används blir risken större att den enskilda personen tänker i termer av sin subjektiva uppfattning och känsla för *AI* som helhet, istället för den specifika teknik som åsyftas. För att slippa den här typen av problematik med generell rädsla och skepsis i samband med introduktion av ny teknik är min erfarenhet att man bör börja med definitionerna och med att beskriva vad det faktiskt är man pratar om. För ingen tror väl på allvar att gräsklipparroboten som åker runt i trädgården kommer ta över världen. Samtidigt finns det viss teknik och speciellt applicering av denna som jag kan uppleva som både olustig och i vissa fall skrämmande, speciellt i samband med övervakning. Men några företags och myndigheters sätt att använda t.ex. ansiktsigenkänning kan inte tillåtas ligga till grund för den generella uppfattningen om *AI*. Då går vi miste om enorma fördelar och en fantastisk potential. Specifikation och definition av vilken teknik och vilka begrepp som avses blir därför central. För att implementering av ny teknik ska lyckas, är det viktigt med problemformulering och definition av teknik och begrepp. Det krävs också en samsyn rörande detta, eftersom den här typen av implementeringar ställer krav på flera olika delar inom organisationen.

3. Förstå teknikens syfte och potential

Det krävs engagemang och samarbete mellan olika parter för att genomföra en lyckad *AI*-implementering. Verksamhetsspecialister behöver bidra med förståelse av processer och information om vad den nya tekniken behöver uppfylla, och hur den kan sammanfogas med befintlig teknik. *IT*-specialister behöver säkerställa kapacitet och säkerhetsmässiga förutsättningar för att kunna genomföra implementeringen och inte minst behöver tekniskspecialister förstå tekniken och få utbildning för att kunna bedöma och utvärdera den tekniska potentialen. Att förstå syftet med en viss teknik blir centralt.

En typ av *AI* som många företag nu tittar på att implementera är chatbotten. Det finns flera olika typer av chatbottar på marknaden med olika avancerad nivå av teknik men de flesta använder någon form av machine learning och natural language processing för att kunna hålla en dialog med en användare. Syftet här är alltså just att kunna hålla en dialog och tolka vad användaren skriver för att kunna svara enligt vad chatbotten är ”lärd” eller programmerad.

Intelligensen ligger här i att förstå vad det är användaren skriver, om det är en fråga eller ett påstående, vilket ämne som berörs etc. Om en kommun som exempel har en chatbot som medborgarna kan ställa frågor till, skulle någon kunna skriva ”Vilken tid stänger simhallen?” till chatbotten. För att kunna förstå den här frågan behöver chatbotten göra ett antal analyser. Den behöver tex. förstå att det är en fråga, att det handlar om någon form av tidsaspekt och att det berör en speciell anläggning. I det här fallet kan man designa lösningen på olika sätt; antingen kan chatbotten presentera samtliga öppettider på samtliga simhallar i kommunen, alternativt ställa följdfrågorna: (1) ”Vilken simhall?” och (2) ”Vilken dag?”. Båda lösningarna kräver dock att det finns ett färdigt svar att hämta – i första fallet en tabell med simhallar och tider, i det andra fallet en typ av formel som skapar ett svar i stil med ”Simhallen i Rosbäcken stänger kl 19 imorgon.”. Med andra ord kräver båda lösningarna att den underliggande datan är strukturerad på ett sådant sätt att chatbotten kan presentera det för användaren.

Förståelsen för att chatbottens syfte är att hålla en dialog är här avgörande. Det kräver nämligen att den data som man vill att chatbotten ska ha tillgång till är strukturerad på rätt sätt – man kan inte bara be den söka i en ostrukturerad databas. Chatbottens syfte är inte att strukturera data och den har inte heller teknisk potential att göra det. Har verksamheten med andra ord inte strukturerade data att tillgå är det ingen större idé att investera i en chatbot som ska kunna svara på frågor med hjälp av data.

Till sist

Inget av det som jag skrivit ovan är egentligen speciellt uppseendeväckande – och till stor del är det just detta som är poängen. Det finns i mina ögon ingen anledning för företag att behandla *AI* annorlunda från hur man implementerar annan form av ny teknik. All implementering bör utgå ifrån (1) ett problem som ska lösas, (2) en av alla inblandade välkänd begreppsapparat samt (3) förståelse för tekniken. Ju större gap det är mellan den hittills använda tekniken och den nya teknik som ska implementeras, desto viktigare blir dessa tre punkter för att man ska lyckas. Den nya teknik som paraplybegreppet *AI* representerar är inte längre något teoretiskt fenomen utan finns numera överallt i vårt samhälle. Utvecklingen sker hela tiden, och aktörer bör analysera vilka utmaningar som finns framöver och vilken typ av teknik som är bäst för att lösa olika problem.

JOSEFINE HAGBARTH
KONSULT I AI-IMPLEMENTERING
QUBES

Intresset för AI har varierat över tid. Dagens hajp beror på att det nu rent tekniskt går att beräkna saker i en hastighet som inte var möjlig tidigare. Den snabba utveckling på den fronten kan vi tacka datorspelen för.

Om artificiell intelligens och statistik

Vad är artificiell intelligens?

För att förstå hur statistik och artificiell intelligens (förkortat *AI*) beror av varandra behöver man först förstå vad *AI* är för något. Dessvärre är detta en fråga som inte är lätt att besvara. Beroende på vem man frågar, så kan man få olika svar. I den här artikeln ges min subjektiva syn.

AI kan vara alltifrån ett verktyg för beslutsassistans till en självlärande algoritm. Enligt den definitionen skulle man kunna hävda att något så enkelt som en tärning – eller något annat som på ett slumpmässigt sätt hjälper en att fatta ett beslut – är *AI*. Det finns med andra ord inte någon tydlig definition, utan det är snarare upp till var och en vad som läggs in i begreppet *AI*.

Min egen definition grundar sig på den mer filosofiska frågan: vad är intelligens? Ett möjligt svar är att det handlar om förmågan att dra slutsatser, samt om möjligheten att lära nytt. Statistiken tar plats i *AI* när det gäller förmågan att dra slutsatser, medan möjligheten att lära nytt har att göra med programmeringsteknik. Att koppla ihop statistiska modeller med system som automatiskt uppdaterar algoritmer är enligt denna argumentation vad som borde kallas *AI*, eftersom uppdateringar gör att algoritmer ”lär sig” nya saker i och med den nya information som de får tillgång till. Men detta är som sagt min subjektiva definition, och du har rätt till din egen.

Varför är AI så populärt just nu?

Faktum är att ”nu” har pågått under längre tid än vad som är allmänt känt. Själva begreppet *AI* myntades av John McCarthy under en konferens 1956 i samband med frågan ”Kan maskiner tänka?”. Denna fråga hade från början ställts av Alan Turing i artikeln ”Computing Machinery and Intelligence”, publicerad i filosofitidskriften *Mind* 1950. Artikeln innehåller bl.a. följande resonemang (här förenklat) om vad som ska menas med ”tänkande” hos en dator. Vi tänker oss en situation där en person via tangentbord samtalar med en osynlig samtalspartner, som i själva verket är en dator. Men försökspersonen vet inte om samtalspartnern är en människa eller en dator. Om personen inte genom samtalet kan avgöra ifall motparten är en människa eller en dator, så skulle datorn betraktas som ”tänkande”. Detta test kom att kallas Turingtestet. Testet har ifrågasatts, och redan 1966 kom ett program, Eliza, som kunde skapa en illusion av en tänkande dator.

Ända sedan begreppet AI uppkom har intresset för ämnet varierat. Perioder med stor framtidstro, med tillhörande stora forskningsanslag, har följts av perioder med svalnat intresse och minskade anslag, när de förutspådda resultaten uteblivit. Så hur kommer det sig att det är en så stor hajp nu?

Svaret är ganska enkelt. Rent tekniskt är det nu möjligt att beräkna saker »»»

i en hastighet som inte var möjlig tidigare. Till viss del beror det på att datorernas processorer blivit snabbare, men den främsta orsaken är möjligheten till parallella processer. Att det har skett en så snabb utveckling på den fronten kan vi tacka datorspelen för. I moderna datorspel är det många moment som ska behandlas grafiskt samtidigt. Som exempel så skall varje vektor av ljus beräknas individuellt. Även om en processor är snabb, kan den inte beräkna alla dessa ljuskällor samtidigt. Datorspelsföretagens lösning var att utveckla separata grafikkort för dessa beräkningar, så kallade *graphics processing units (GPU)*. Till skillnad från en vanlig *CPU*, vilken består av ett fåtal kraftfulla processorkärnor som snabbt kan utföra komplicerade beräkningar, utgörs en *GPU* av tusentals förhållandevis långsamma processorer, som med fördel kan hantera stora volymer av enklare beräkningar.

I statistiska processer, där beräkningarna görs i följd, har detta inte så stor betydelse, men för djupa neurala nätverk – som vanligen används vid bild- och textigenkänning – så spelar detta stor roll. Om de olika lagren vilka bygger upp en pixel i en bild är oberoende av varandra, kan beräkningarna parallelliseras. Även olika bilder kan beräknas simultant. Själva tiden för beräkningen kan på detta sätt kortas ofantligt. I vissa fall till någon tusendel av vad en enkel processor hade behövt för samma procedur. Vi kan alltså tacka datorspelsutvecklingen för hajpen inom *AI* just nu.

AI i arbetslivet, en bransch under utveckling?

För ett tag sedan var jag var på en presentation rörande en mjukvaruuppdatering av en personaldatabas. Själva presentationen berörde inte alls *AI*. När den var över utbrast en av åhörarna: ”Varför är det inte någon *AI* i det här? Det är ju så enkelt med *AI* idag.” Det var nog ingen där som egentligen förstod vilken typ av *AI* som personen önskade, antagligen inte han själv heller. Som jag nämnde tidigare finns det ingen definition av vad *AI* är, vilket innebär att åhörarens fråga inte är felaktig, bara förvirrande. En personaldatabas skulle kunna innehålla massor med smarta *AI*-lösningar, men behövs dessa, och vilken funktion skulle de i så fall ha?

För att minska risken för förvirring bör *AI* brytas ner och delas in i områden som är mer deskriptiva och därmed enklare att förstå. Möjliga områden skulle kunna vara: traditionella statistiska modeller, *machine learning, deep learning, reinforced learning* etc.

Innan jag började på min nuvarande tjänst gick jag på en del jobbintervjuer. Tjänsterna som utlystes hade i mina ögon inte så mycket med *AI* att göra. Många av dem jag talade med var trots detta övertygade om att de behövde *AI*-kompetens inom sin organisation. De hade en förväntan om att *AI* skulle lösa deras problem, men eftersom de inte riktigt visste vad *AI* var, så hade de svårt att formulera sitt faktiska behov. Det var inte ovanligt att de sökte efter någon med statistisk bakgrund, medan vad de egentligen behövde var någon som kunde bygga upp databaser eller *data lakes*. Detta ligger närmare *big data*, något som numera delvis räknas som en underkategori av *AI*-kompetens.

»Definitionen av en data scientist är i mångt och mycket lika bred och förvirrande som definitionen av AI. Det är inte en skyddad yrkestitel, så vem som helst kan kalla sig för detta.»

Även om marknaden utvecklas snabbt, så finns det fortfarande en hajp kring begreppet *AI*. Som konsult kan jag se att det som efterfrågas emellanåt är sådant som statistiker traditionellt jobbar med. Men det är inte statistiker som efterfrågas, utan *data scientists*. Det som egentligen efterfrågas är personer med vissa kunskaper utöver vad den traditionella statistikern har. Själva statistiken är en del, men till detta kommer sådant som kunskap om datahantering, programmering, olika molntjänster, områdesexpertis med mera. Detta är en väldigt bred yrkesbeskrivning, varför en *data scientist* emellanåt beskrivs som en enhörning: det finns ingen med kunskaper inom allt som efterfrågas av dem.

Definitionen av en *data scientist* är i mångt och mycket lika bred och förvirrande som definitionen av *AI*. Det är inte en skyddad yrkestitel, så vem som helst kan kalla sig för detta. Det är inte ovanligt att datavetare – och i viss mån även nationalekonomer – kallar sig för *data scientists*.

Inom mitt yrke har jag träffat personer som kallar sig *data scientists*, men som inte förstår statistik. De kan programmera och de löser sina problem genom att använda modeller som oftast levererar, men som är svåra att förstå, så kallade *black box*-modeller. Jag har också träffat personer som enbart kan ekonometri, och i viss mån linjära modeller. Problemet med detta är att förvirringen kring vem som kan *AI* samt vem som utger sig för att kunna *AI* förstärker förvirringen kring vad *AI* faktiskt är. För att råda bot på detta behövs konkretisering av både *AI* och vad som krävs för att utveckla *AI*.

Kan du AI?

En vän till mig sa något i stil med: ”Jag är statistiker, inte *data scientist*! Men titeln ’data scientist’ ger mig 20 000 kronor extra i månaden.”

Detta är en överdrift då tjänsten kräver andra kunskaper, men citatet sammanfattar ändå den här artikeln ganska bra. På grund av att marknaden fortfarande utvecklas och företag inte alltid är medvetna om vad de efterfrågar, så skulle följande scenario kunna uppstå:

1. Ett företag vill ha *AI*, men man är inte säker på vad detta är och betalar därför multum för något som är en enkel statistisk modell, t.ex. en linjär regression.
2. Någon tar sig an jobbet, benämner sig *data scientist* och får i och med detta bra betalt.
3. *Data scientisten* kan vara statistiker och således kan det vara du!

I dagsläget krävs det inte mycket för att kunna kalla sig *data scientist* eller svara ja på frågan: Kan du *AI*? Men marknaden är under utveckling, och för att vi som jobbar inom området ska kunna utvecklas och utveckla *AI*, så krävs det mer än kunskaper inom statistik. Du bör kunna programmera i Python eller R, som är de två mest efterfrågade språken på marknaden idag. Du bör även ha goda kunskaper inom SQL och i någon av molntjänsterna AWS, Azure eller Google Cloud, så pass att du känner dig bekväm med att hantera data och modeller. Först då kan du enligt min mening med rätta kalla dig för *data scientist*.

PER MARCUS
KONSULT INOM DATA SCIENCE,
SOGETI SWEDEN

(Per Marcus har tidigare arbetat med utveckling av AI på Tölve, samt varit föreläsare vid Statistiska institutionen på Stockholms universitet.)

Det tekniska och det analytiska – en data scientists nya roll i det nya decenniet

Kompetensskillnader mellan olika roller

Låt oss börja med den populära myten att data scientists är enhörningar: De ska vara lika duktiga analytiker som en person med en utbildning i statistik och lika bra eller bättre programmerare som en person med en utbildning i datavetenskap. Nyexaminerade doktorer som i sin avhandling analyserat data genom att programmera i något språk som R eller Python har de senaste åren varit väldigt anställningsbara – tack vare tillräcklig kunskap inom både statistik och programmering, och dessutom med den vetgirighet som krävs för att lyckas bra som data scientist.

Då nydisputerade doktorer är en förhållandevis liten pool att anställa från blir frågan om företag ska anställa en statistiker som lärt sig programmera eller en datalog som lärt sig analysera data. För många faller det sig mer naturligt att lära en statistiker att programmera, men faktum är att datalogerna har varit snabbare på att lära sig att analysera. Den grundläggande orsaken till detta är det stora antalet *machine learning* (ML) algoritmer som finns tillgängliga på marknaden och att ML många gånger är enklare att applicera på ett problem än statistiska modeller. Statistiska modeller kräver förkunskaper om bland annat datas fördelning i urvalet och i populationen, vilken är den datagenererande funktionen etc., medan ML algoritmer endast kräver att data är i rätt form – exempelvis så blir det inte så bra ifall vi låter postnummer hanteras som numeriska variabler och datorn bestämmer sig för att räkna ut standardavvikelsen på postnumret och ger negativ effekt till alla postnummer som ligger tre standardavvikelser över medel. När man säkerställt att data är i rätt form är det fritt fram att se vad man får för resultat. Endast i fall där något ser konstigt ut eller om resultaten inte är bra nog går man tillbaka och begrundar möjliga lösningar för att förbättra modellen, då ofta genom att fundera på vad man kan lägga till för nya datakällor.

Olika arbetsbeskrivningar

Ett klassiskt statistiskt analysprojekt – hypotes, datainsamling, tester/modellering, analys, rapport – kallas ofta för ett ad hoc projekt inom data science, vilket påvisar en skillnad i synen på arbetet. Ett typiskt data science projekt går ut på att sätta upp en data pipeline från en datakälla och sedan kontinuerligt göra körningar på data som kommer in, för att sedan spotta ut en klassificering eller uppdatera scores för alla kunder i ett register (eller något liknande).

Detta är också varför Python har börjat ta över som det mest prominenta programmeringsspråket för data scientists, och varför SAS och Matlab knappt längre utgör ett intressant alternativ för många arbetsgivare som söker data scientists (Jeff Hale 2019, https://bit.ly/38e2hiw). Python och R erbjuder väldigt likartade funktioner, men det är sedan länge känt att statistiker tenderar föredra R medan dataloger föredrar Python.

Denna utveckling har i sin tur sparkat igång en motreaktion bland vissa data scientists som har börjat prata om data engineering som ett komplementyrke till data scientists, som skulle fokusera mer på den tekniska biten kring datainsamling, manipulering och produktion, medan data scientists skulle få vara metodinriktade och arbeta mer med modellering och tillämpning av modellresultaten. En artikel på *KD nuggets* föreslår att det borde gå ungefär tio DE:s på en DS (https://bit.ly/2G05y8Y). Även om förhållandet inte är exakt så är det intressant att se hur branschen börjar föreslå en sådan sned fördelning mellan antalet som arbetar med att få data på plats och antalet som faktiskt gör något med data när de är på plats.

Samtidigt som jag kan uppskatta att man bryter ut det mest tekniska ur en data scientists arbetsbeskrivning (och lägger över det på dessa nyinstiftade data engineers), så tycker jag att hela resonemanget fokuserar för mycket på det tekniska och missar en viktig del av vad som är en data scientists kompetens, nämligen områdeskunskap (fritt översatt från domain expertise som Rachel Schutt och Cathy O'Neill skrev om i sin bok *Doing data science: Straight talk from the frontline*. (O'Reilly Media, Inc., 2013)).

Avslutningsvis

Efterfrågan på personer som kan analysera och skapa värde från data har bara ökat under de senaste åren och kommer att fortsätta göra det framöver. Statistikers främsta expertis är att sätta data i rätt kontext och dra meningsfulla slutsatser. En av svagheterna med att *machine learning, reinforcement learning* och *deep learning* algoritmer är *black boxes* (så även AI, via satslogik) är att det blir svårare för beslutsfattare att syna och kritisera resultat, framför allt om man endast har ringa kunskaper om metoderna. Därför tror jag att statistikers kompetens på arbetsplatserna kommer att vara viktigare än någonsin under det kommande decenniet.

EMMA ANDERSSON
STATISTIKER

TILL MINNE AV OVE FRANK

Professor Ove Frank, Södermalm, avled den 21 oktober i en ålder av 82 år efter en tids sjukdom. Hans närmaste anhöriga är hustrun Kerstin och barnen Mikael, Camilla och Marika med familjer.

Ove föddes den 27 september 1937. Han började studera bland annat matematik och matematisk statistik vid Stockholms universitet och skrev sin licentiatavhandling i matematisk statistik. Den handlade om signalstörning och var Oves första arbete med entropimått (läs mer om entropimått i Oves artikel ”Tid för nya metoder i tillämpad statistik”, Qvintensen nr 2, 2011).

Ove fick sedan arbete som operationsanalytiker vid Försvarets forskningsanstalt. Ett av hans första bidrag var en numera klassisk lösning till problemet hur tvåvägstrafik kan ledas på en enkel järnväg utan att alla tåg samlas på ett ställe. Mycket av hans arbete på FOA var hemligstämplat, men intresset för koder, entropi och chiffer hade han kvar i hela sitt liv. Parallellt med arbetet fortsatte han forska och disputerade 1971 vid Statistiska institutionen i Stockholm med den banbrytande avhandlingen ”Statistical Inference in Graphs”. Den var minst tjugo år före sin tid. Då sågs hans område, grafer, som lite udda men idag är det ett hett område.

Redan i provföreläsningen 1974 för professuren i Lund tog Ove upp ämnet statistikundervisning. Åhörarna den gången uppfattade ämnesvalet som banalt, men senare har vi som lärare insett dess betydelse. Sitt djupa pedagogiska intresse visade han även senare som lärare, prefekt och linjenämndsordförande. Ove stannade i Lund i tio år. Under ett sabbatsår i UC Riverside skrev han tillsammans med David Strauss ett epokgörande arbete om Markovgrafer. Detta öppnade ett nytt forskningsfält inom nätverksanalys. En av hans doktorander i Lund, Krzysztof Nowicki, blev senare professor vid institutionen i Lund.

När professuren i statistik i Stockholm blev ledig, flyttade Ove och familjen tillbaka till Stockholm och statistiska institutionen. Även där satte han sin prägel på verksamheten. Under sin tid som prefekt såg han till att institutionen anordnade den första konferensen för kvinnliga statistiker. Han entusiasmerade doktorander för sin specialitet, grafer. Han arrangerade konferenser och blev en central person inom det internationella forskningsnätverket om grafer. En av dem som han samarbetade mycket med, Tom Snijders, blev senare också hedersdoktor i Stockholm. Flera av hans studenter har sedan blivit internationellt uppmärksammade. Ove var även mycket intresserad av att samarbeta tvärvetenskapligt både inom och utanför Sverige, bland annat inom psykologi, kriminologi och socialt arbete. I Qvintensen nr 3, 2011 finns en intervju med Ove där han bland annat berättar om livet som doktorand på 60- och 70-talet, sitt tvärvetenskapliga intresse och sina tankar om pedagogik.

Efter pensionen fortsatte Ove sin forskning med flera intressanta resultat. Han besökte också regelbundet institutionen tills hans hälsa gjorde det omöjligt. Han delade då frikostigt med sig av sina erfarenheter till stor glädje för många på institutionen, även för dem som arbetade inom andra områden. Ove lämnar ett stort tomrum efter sig, både på institutionen och i forskarvärlden.

GEBRENEGUS GHILAGABER, INGEGERD JANSSON, KRZYSZTOF NOWICKI,
DANIEL THORBURN, ARBETSKAMRATER OCH LÄRJUNGAR

FMS

ORDFÖRANDEN HAR ORDET

Nya visioner i föreningen

FMS samlar personer med intresse för och anknytning till medicinsk statistik. Vi representerar också svensk medicinsk statistik i internationella sammanhang, t ex i European Federation of Statisticians in the Pharmaceutical Industry (EFSPI). Under vårmötet 2019 valdes en ny styrelse bestående av: Ingeborg Waernbaum, Catrin Wessman, Jonas Häggström, Sara Ekberg, Anna Nilsson, Henrik Renlund, Fredrik Norström. Vi har också två nya representanter i EFSPI: Anna Torrång och Andreas Gustavsson.

»Tillsammans med deltagarna på höstmötet har vi identifierat tre särskilt viktiga arbetsområden»

Den 18:e oktober genomfördes FMS höstmöte, temat var observationsstudier. Presentationerna handlade bland annat om kausal inferens, medicinska tillämpningar, registerdata och etikprövning. Talarna kom från både universitet och företag från olika delar av landet. Höstmötet innehöll också en mini-workshop där vi diskuterade visioner för FMS och samlade in förslag från deltagarna.

Under året har en arbetsgrupp bestående av Ingeborg Waernbaum, Anna Torrång, Jonas Häggström och Fredrik Norström inlett ett visionsarbete för FMS. Tillsammans med deltagarna på höstmötet har vi identifierat tre särskilt viktiga arbetsområden: (i) att uppmuntra/främja utbildning i biostatistik, (ii) att understödja kontakter mellan biostatistiker och (iii) att vara synliga i samhället. Här har vi satt igång med en enkel handlingsplan och arbete i flera riktningar, som vi kommer att fortsätta med under året. Vi kommer också att arbeta med att utveckla hemsidan, som har ett nytt domännamn (fmsbiostat.se).

Under 2019 delades två stipendier för konferensdeltagande ut: till Aminata Ndiaye, MEB, Karolinska Intitutet (13 000 SEK), och till Caroline Weibull , Klinisk Epidemiologi, Karolinska Institutet (11 000 SEK).

INGEBORG WAERNBAUM

ORDFÖRANDEN HAR ORDET

»Ingen teknik kan läsa folks tankar»

Jag dansar lindy hop och har gjort så i många år. På frågan ”Vad är lindy hop?” brukar jag alltid svara ”Det är en pardans. En swingdans. Startade i Amerika på 20-talet.” Men det går ju inte att säga längre, för 20-talet är ju nu!

Så vad kommer detta glada 20-tal att ha att bjuda på, då? Lätt att gissa men svårt att säga, kan jag tycka. Men en sak är jag säker på: AI och Machine Learning kommer att fortsätta växa i popularitet, men kommer inte att ersätta statistiska metoder. Anledningen till detta är simpel: Det finns saker som AI inte kan göra, men som går att göra med statistiska metoder.

Mest tydligt blir det när det kommer till just surveyundersökningar: Det finns idag väldigt avancerad teknik för att följa folks beteenden både online (via pixlar och cookies, exempelvis) och offline (via kameror eller satelliter, exempelvis), men det finns inget sätt att med teknik läsa av en persons tankar eller dra ifrån dem deras åsikter – och tur är väl det, för vad för samhälle skulle det vara att leva i?!

Skämtsam ton åsido så tycker jag att det finns något viktigt här som vi surveystatistiker bör ha med oss när vi pratar med andra närliggande discipliner: Oavsett hur kritisk man vill vara mot branschen och dom utmaningar som finns med sjunkande svarsfrekvenser och ökande komplexitet i undersökningarna, så kvarstår det faktum att ställa frågor är det enda etablerade och väl fungerande sättet som finns för att ta reda på vad människor faktiskt tycker och tänker.

Och tycka och tänka kommer människor att fortsätta göra även i det nya decenniet.

MARCUS BERG

Vinnare av MarC Awards 2019

Torsdag den 28/11 gick undersökningsbranschens galakväll MarC Awards av stapeln i Stockholm. (”MarC” står för Market Research and Customer Insight.) MarC Awards arrangeras av SMIF (en intresseorganisation för svenska undersökningsföretag) i samarbete med Swedish Insighters, ESOMAR, Surveyföreningen och Svenska Statistikfrämjandet. Omkring 100 personer från alla delar av undersökningsbranschen hade samlats för att fira vinnarna i årets sex priskategorier. Alla vinnare samt motiveringstexterna finner du nedan (ordagrant hämtat från SMIF:s hemsida).

ÅRETS NYKOMLING
ALMASA KULENOVIC, DAPRESY
Almasa har på endast två år i undersökningsbranschen satt sig in i vitt skilda områden och hur kunderna bäst ska använda sina resultat. Genom detta har hon framgångsrikt levererat ett stort antal komplicerade svenska och internationella rapporteringslösningar, och är idag chef för Dapresys leveransorganisation.

ÅRETS OPINIONSUNDERSÖKARE
PETER SANTESSON, DEMOSKOP
Peter Santesson är doktor i statsvetenskap vid Lunds Universitet och har forskat om reformprocesser och politiskt beslutsfattande. Han har lett arbetet med att utveckla Demoskops väljarbarometer och är pragmatisk, noggrann och transparent med varje detalj i den process det innebär att göra väljarundersökningar.

ÅRETS UNDERSÖKNINGSPROJEKT
KANTAR OCH SCANDIC
A great hotel experience for the many people. Kantar har på uppdrag av Scandic satt upp en internationell Customer Experience-lösning där man genom en kombination av enkäter, textanalys och visualisering dagligen har full insyn i Scandics gästers nöjdhet för hela hotellvistelsen från incheck till utcheck.

ÅRETS UPPDRAGSGIVARE
LENNART NYBERG, PAULIG FOODS
Lennart har jobbat med undersökningar i 30 år, både som undersökare och köpare. Han är en krävande köpare som vet hur han utmanar sina leverantörer till att bli bättre. Lennart är med och utvecklar vår bransch genom sin holistiska, utmanande men också långsiktiga approach.

ÅRETS STATUSHÖJARE
INTERNETSTIFTELSEN
Genom undersökningen ”Svenskarna och internet” kartlägger Internetstiftelsen internetanvändningen i Sverige. Rapporten ger förståelse för hur användningen av internet i Sverige förändras från år till år, och deras välutvecklade sätt att presentera rapporten gör att allmänhetens intresse för statistik och undersökningar ökar.

BRANSCHENS HEDERSPRIS
SÖREN HOLMBERG, GÖTEBORGS UNIVERSITET
Sören Holmberg på Göteborgs Universitet har forskat om valprocessen och dess roll i en en representativ demokrati sedan 70-talet. Har var med och grundade SOM-institutet och har återkommande varit expertkommentator i SVT vid riksdagsvalen.

Demoskops debattartikel i DN

På DN Debatt-sidan publicerades den 15 november 2019 nedanstående artikel, skriven av personer knutna till undersökningsföretaget Demoskop. Ingress, artikelrubrik och artikeltext från DN återges här:

DN DEBATT 15/11. Under två år har vi kombinerat telefonintervjuer med en onlinerekryterad webbpanel. Visserligen blev vår barometer den mest träffsäkra, men i riksdagsvalet visade det sig att vårt kvarstående inslag av telefonintervjuer försämrade träffsäkerheten. Slutsatsen är att vi nu överger telefonintervjuer som metod i väljarbarometern, skriver företrädare för Demoskop.

Nu slutar vi mäta väljarnas partisympatier via telefon

Tillförlitligheten i opinionsundersökningar diskuteras ständigt. Resultaten kan få stor politisk betydelse. Särskilt känsliga är väljarbarometrarna som följer utvecklingen av partisympatierna. Flera faktorer gör att tillförlitligheten ifrågasätts. Allmänhetens kommunikationsvanor förändras snabbt samtidigt som viljan att delta i undersökningar i dag är historiskt låg och sjunkande. Det finns också tecken på att man ibland drar sig för att svara uppriktigt på frågor om partisympatier.

För knappt två år sedan genomförde Demoskop parallellmätningar med olika urvals- och insamlingsmetoder. Traditionella telefonintervjuer baserade på ett slumpmässigt befolkningsurval (så kallat sannolikhetsurval) metodtestades tillsammans med flera olika typer av onlinepaneler. Vi kunde samtidigt konstatera att allmänhetens vana att inte svara på samtal från okända nummer på mobiltelefoner hade tillkommit som ett problem för telefonundersökningar som allt oftare ringer på mobilnummer (DN Debatt 30/12 2017). Det var dessutom känt att telefon inte är en ideal datainsamlingsmetod om frågor kan uppfattas som känsliga. Slutsatsen då blev att telefonintervjuer behövde kompletteras med webbaserade intervjuer. Tack vare den metodrevisionen blev Demoskops väljarbarometer den som kom närmast resultatet i riksdagsvalet 2018.

Förutsättningarna för opinionsundersökningar fortsätter att förändras. Därför har Demoskop genomfört ett nytt metodtest av olika undersökningsmetoder. Två olika leverantörer av telefonintervjuer, som båda används i välkända väljarbarometrar som publiceras i riksmedier, har parallelltestats med två onlinepaneler som har olika källor. Vidare har vi uppdragit åt en branschledande leverantör av urval att försöka ta fram telefonnummer till ett slumpmässigt befolkningsurval. Vi har också upprepat 2017 års undersökning av allmänhetens benägenhet att inte svara på samtal från okänt nummer.

Bortfall kan vara en allvarlig felkälla. Vårt metodtest visar på ett betydande bortfall i telefonmätningarna. Av ett slumpmässigt befolkningsurval saknas telefonnummer till 20 procent av de utvalda personerna. Därefter är det ytterligare 34 procent av urvalet som man inte lyckas komma i kontakt med: fel nummer, ingen svarar, man når bara en telefonsvarare, i vissa fall finns språk- eller hörselproblem. Innan man har fått kontakt med urvalspersonerna har 54 procent av det ursprungliga urvalet redan fallit bort. Av de återstående 46 procent där man får kontakt är det närmare hälften, 20 procentenheter, som inte vill bli intervjuade. Av det ursprungliga urvalet är det i slutänden 26 procent där man kan genomföra en intervju (1.000 telefonintervjuer, slumpmässigt allmänhetsurval, genomförda 9–16 oktober 2019). Det som från början var ett slumpmässigt urval är det till slut inte längre om det finns systematik i bortfallet.

Bortfallssiffran kan dessutom vara i underkant. I det separata försöket att nummersätta urval, där Demoskop har haft full insyn i urvalet, var bortfallet markant större. Där saknades telefonnummer till 34 procent av urvalet, och bortfallet var särskilt stort bland yngre. I åldersgruppen 18–38 år saknades nummer till hela 46 procent av de utvalda.

Med så höga bortfall blir undersökningar känsliga för om olika grupper är systematiskt lättare eller svårare att komma i kontakt med. Särskilt känsligt blir det om det finns en åsiktssystematik i bortfallet, det vill säga ett samband mellan de åsikter man vill undersöka och sannolikheten för att man deltar i undersökningen. Demoskops analys av oviljan att svara på samtal från okänt nummer tyder på att det kan finnas en sådant samband.

Att undersöka kontaktovillighet har uppenbara metodsvårigheter och resultaten bör tolkas försiktigt. Men undersökningen indikerar att ovilja att ta samtal från okända nummer är utbredd

Av ett slumpmässigt befolkningsurval saknas telefonnummer till 20 procent av de utvalda personerna. Därefter är det ytterligare 34 procent av urvalet som man inte lyckas komma i kontakt med. Nu överger Demoskop telefonintervjuer.

och ökar. I dag uppger tre av fyra i en onlineundersökning att de sällan eller aldrig svarar på okända nummer. Andelen har ökat med 12 procentenheter de senaste två åren. 29 procent uppger att de aldrig svarar på okända nummer.

Screening av okända nummer ser ut att variera med åsikter. Det är vanligare om man står längre till höger (röstar på M, KD eller SD) eller längre till vänster (röstar på V) än när man röstar på något av januaripartierna (S, C, L och MP). Andelen som aldrig svarar på okända nummer är 10 procentenheter högre bland SD:s väljare (33 procent) än bland januaripartiernas (23 procent).

Ett sätt att utvärdera alternativa metoder för en väljarbarometer är att analysera hur väl man lyckas reproducera resultatet av det senaste riksdagsvalet i undersökningen. Om man gör en urvalsundersökning som är helt utan andra fel än det som uppstår för att man undersökt ett urval och inte hela populationen, dvs alla uppgiftslämnare medverkar, minns rätt och svarar sanningsenligt, skulle svaren på frågan om vad man röstade på i det senaste valet ge ett resultat som speglade det verkliga valresultatet.

Så ser det aldrig ut. Dessa andra fel är ofrånkomliga, och storleken på dem ger en indikation på hur stora mätfelen är med olika metoder. Vårt metodtest visar att det fortfarande finns svårig¬heter med att korrekt skatta SD-sympatier via telefonintervjuer. Efter att ha vägt för skevheter i urvalen underskattar telefonundersökningar andelen SD-väljare i föregående val med cirka 4 procentenheter. Motsvarande siffra för onlinepanel visar en relativt blygsam underskattning (–0,5 procentenheter). De sammanlagda avvikelserna mot föregående val är större i båda de testade telefonmätningarna jämfört med onlinepanel.

Bland praktiker som dagligen arbetar med telefonundersökningar är felen välkända, men de diskuteras sällan offentligt och konkret. Under två år har Demoskop kombinerat telefonintervjuer med en onlinerekryterad webbpanel. Visserligen blev vår barometer den mest träffsäkra, men i riksdagsvalet visade det sig att vårt kvarstående inslag av telefonintervjuer försämrade träffsäkerheten marginellt. Slutsatsen för Demoskops del efter denna utvärdering är att nu överge telefonintervjuer som metod i väljarbarometern, främst på grund av bortfall och vissa mätfel.

I takt med att kommunikationsvanor förändras och viljan att svara på undersökningar minskar blir metodfrågor allt svårare. Det finns också betydande kvalitetsskillnader mellan olika metoder. Det vore därför felaktigt att påstå att en metod är överlägsen en annan. Det som krävs är att fortsätta utvärdera hur dessa presterar i svenska sammanhang och att offentligt redovisa resultaten. På det sättet kan vi öppna för en konstruktiv och transparent metoddiskussion. Det här är vårt försök att bidra till den.

KARIN NELSSON, VD, DEMOSKOP
ANDERS LINDHOLM, SENIOR ADVISOR, DEMOSKOP
LARS LYBERG, SENIOR ADVISOR, DEMOSKOP
PETER SANTESSON, OPINIONSCHEF, DEMOSKOP

Information om ackreditering

1) Från FENStatS arbetsgrupp

Vi lever i en värld med större och större behov av snabbare och mer korrekt material för statistiska analyser. Behovet av ett kvalitetsmärke för statistiker har diskuterats under många år. Möjlighet till ackreditering har sedan flera år erbjudits till medlemmar i AMSTAT och RSS.

Inom FENStatS (The Federation of European National Statistical Societies) har en internationellt sammansatt arbetsgrupp utvecklat ett system för ackreditering, som nu har antagits av FENStatS styrelse. De nationella organisationerna inbjuds nu att ansluta sig till detta system och alltså erbjuda sina medlemmar möjlighet till ackreditering. Om den nationella organisationen vill ansluta sig, måste man meddela FENStatS detta, samt utse minst tre revisorer som ska granska inkomna ansökningar.

Mer information om ackreditering finns på FENStatS hemsida.

MAGNUS PETTERSSON
ORDFÖRANDE I FENSTATS ACCREDITATION COMMITTEE

2) Från Svenska statistikfrämjandet

Den internationella organisationen FENStatS har under flera år arbetat med att få till stånd en ackreditering av statistiker inom Europa, och vid årsmötet sommaren 2019 togs ett principbeslut om genomförande. Ett förslag om hur ackrediteringen skulle gå till skickades hösten 2019 till de europeiska medlemsorganisationerna. En grundtanke är att ackrediteringen ska vara frivillig, dvs. att det är upp till varje individ om han eller hon vill ackreditera sig.

Vid Svenska statistikfrämjandets styrelsemöte i december 2019 beslutade styrelsen att ansluta sig till förslaget från FENStatS, och därmed inleda arbetet med att ge svenska statistiker möjlighet att ackreditera sig. FENStatS håller nu på att ta fram nödvändig dokumentation, som ska guida arbetet, och Svenska statistikfrämjandets styrelse kommer inom kort att påbörja arbetet med att bl.a. sätta samman den grupp som ska gå igenom inkomna ansökningar.

Mer information kommer att delges medlemmarna under årsmötet den 18/3 2020.

MATTIAS STRANDBERG
SEKRETERARE, SVENSKA STATISTIKFRÄMJANDET

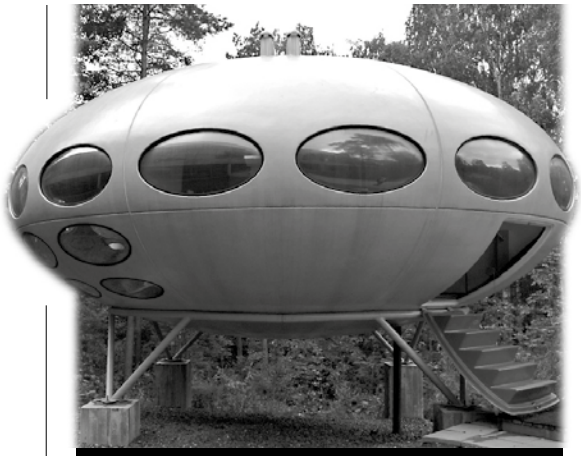
Statistikfrämjare blir hedersdoktor i juridik

Juridiska fakulteten vid Lunds universitet har utsett en medlem i Statistikfrämjandet, Anders Nordgaard, till hedersdoktor 2020. Anders disputerade i matematisk statistik vid Linköpings universitet och är docent i statistik. Han har intresserat sig för tillämpning av statistisk metodik i samband med forensisk bevisvärdering. Nu är han forensisk specialist vid Nationellt forensiskt centrum (NFC) i Linköping, och dessutom adjungerad lektor vid Avdelningen för statistik och maskininlärning (STIMA) i Linköping.

Anders har både som forskare och praktiker medverkat till utvecklandet av Bayesiansk metodik vid värdering av forensisk bevisning. Han har bland annat publicerat vetenskapliga artiklar om användningen av likelihood ratio i bevisvärdering. Han är också Editor-in-Chief för den ledande internationella tidskriften Law, Probability and Risk, som ges ut av Oxford University Press. I sin verksamhet vid NFC har han under lång tid haft en ledande roll i utvecklandet av forensisk resultatvärdering, och han har i detta arbete gjort en väsentlig insats för svensk rättssäkerhet.

Anders har under många år haft ett nära samarbete med juridiska fakulteten vid Lunds universitet, och bidragit till fakultetens verksamhet inom både forskning och undervisning. På forskningssidan medverkar han för närvarande i ett treårigt forskningsprojekt om robusthet i forensisk bevisning, som bedrivs i samarbete mellan Lunds universitet och NFC. På undervisningssidan har Nordgaard sedan många år varit verksam som gästföreläsare i bevisvärdering, och hans föreläsningar om forensisk bevisning har varit mycket uppskattade av studenterna.

Samtliga uppgifter kommer från Juridiska fakulteten vid Lunds universitet.



Har du flyttat?

Du kan själv ändra dina uppgifter genom att gå till *https://www.membit.net/m4-member/login* där du loggar in med dina personliga inloggningsuppgifter (finns på din senaste inloggningsavi). Vid frågor kontakta Mattias Strandberg på sekrfram@gmail.com.

Årsmöte i Svenska statistikfrämjandet

Onsdag den 18 mars är det dags för årsmöte i Svenska statistikfrämjandet med vidhängande konferens. Årsmötet är det främsta forumet för medlemspåverkan som föreningen har, så det är viktigt att du som medlem tar dig tid att närvara för att föreningen ska kunna fortsätta arbeta med saker som är viktiga för dig som medlem.

I samband med årsmötet håller vi vår traditionsenliga konferens, där årets tema är FRAMTID. Vi kommer bland annat få höra från inga mindre än **WALTER RADERMACHER** (Generaldirektör FENStatS, fd. Generaldirektör Eurostat) och även **JOAKIM STYMNE** (Generaldirektör SCB). Vi lovar även flera andra talare, som bland annat kommer att tala om AI och nya tekniska möjligheter i statistiskt arbete, så missa inte detta!

Vi kommer i år att vara på Stockholms universitet
- Årsmötet genomförs på förmiddagen i sal E487
- Konferensen genomförs sedan kl. 13-17 i hörsal 4

Lokalerna ligger i Södra huset på campus Frescati.

Efter årsmötet bjuder Svenska statistikfrämjandet på middag på en restaurang i närheten. Dagen efter (torsdag den 19 mars) är det som vanligt dags för medlemsföreningarnas årsmöten, som också kommer att hållas i lokaler på Stockholms universitet. Lokaler finns bokade kl. 9-12 för dessa möten.

Qvintensen

POSTTIDNING B

SVERIGE
PORTO BETALT
PORT PAYÉ

Avsändare/Sender: Svenska statistikfrämjandet
c/o Mattias Strandberg, Blåeldsvägen 16, 165 73 Hässelby

Bli medlem i Svenska statistikfrämjandet

Svenska statistikfrämjandets syfte är bland annat att främja sund användning av statistik som beslutsunderlag och att väcka och sprida intresse för statistik i samhället.

För att bli medlem, gå till http://www.statistikframjandet.se och läs mer i högerspalten under "Vill du bli medlem?". Har du frågor kontakta Mattias Strandberg på sekrfram@gmail.com.

Du får Qvintensen i brevlådan och platsannonser via e-post.

Det ställs inga krav för att bli medlem; alla som är intresserade av statistik och vill stödja statistikens roll i samhället är välkomna.

Statistik-
främjandet