



När felaktig data ska

Innan jag började att driva ett företag där vi designade IoT-sensorer hade jag inte tänkt så mycket på varför en sensor skulle monteras på ett visst sätt. En temperatur var en temperatur, så länge rätt data kom in så var allt lugnt, eller?

Nu när AI gör entré så har jag efteråt fått ett kvitto på att det var bra att vi var noga från början. För de flesta av oss har någon gång stött på ett AI-svar som känts fel eller rent av nedslående. En hallucinerad sammanfattning, en felaktig översättning, ett svar som ignorerade det uppenbara, eller ett smart system som tog ett dumt beslut. Ibland kan det vara underhållande, och säkert ett bra skämt för kommande AW eller kundlunch. Men när AI-beslut börjar användas i fysiska system – för att styra enheter, reagera vid nödsituationer eller påverka utfall inom fastigheter, vård eller industri – då är det inte längre roligt med hallucinerad data, då kan det bli stökigt på riktigt.

Den här artikeln utforskar en särskild sårbarhet i AI som ofta förbises: bias – och hur den smyger sig in genom sensorer och annan IoT-infrastruktur. När datainsamling kombineras med edge computing och maskin-inlärning kan även välmenande AI-logik hamna snett. Om din AI är tränad, validerad eller implementerad på grundval av snedvriden data kommer inte slutsatserna att vara neutrala. I värsta fall kan slutsatserna både legitimera redan felaktiga antaganden eller förstärka dem.

Hur bias påverkar resultatet när AI och IoT samverkar

Till skillnad från traditionella IT-system samlar IoT-system in stora mängder realtidsdata från fysiska sensorer i komplexa och varierade miljöer. Denna data kan användas till och utgöra grunden för att träna AI-modeller – och det gör att den i förlängningen används som bas för beslut. Om insamlad data inte speglar verkligheten korrekt, riskerar alla komponenter i AI-pipelinen att byggas på fel

som byggs på fel.

När bias tar sig in i IoT-system riskerar det att förstärka befintliga ojämlikheter – vilket kan skapa nackdelar för exempelvis en grupp eller ett område. Det handlar inte bara om rent teknisk noggrannhet, utan konsekvenserna av bias sträcker sig till områden som förtroende, etik och efterlevnad av regler. Diskriminerande AI-algoritmer kan på olika marknader även leda till juridiska och ekonomiska konsekvenser.

Var bias uppstår i IoT-data

Bias i IoT-baserade system kan ha många källor – både tekniska och sociotekniska. Sensorer kan generera felaktiga data på grund av hårdvarufel, dålig kalibrering, avvikelser eller saknade värden över tid. Sensorfel som ligger utanför förväntad data, degradering på grund av ålder eller miljöpåverkan, jamen alla typer av fel spelar in. Några konkreta exempel: luftkvalitetssensorer kan driva iväg i sina värden över tid, det kan handla om något så enkelt som en rengöring som missats. En luftfuktighetssensor kan påverkas vid väderväxlingar eller när temperaturen är nära noll grader. Den här typen av sensorer kräver rätt implementation i hårdvaran och ibland även tillsyn för att leverera kvalitativ data.

Det vanligaste problemet är dock inte hårdvara, utan människor. Sensorer monteras fel och ger felaktiga värden. En temperatursensor som ska mäta rumstemperatur som placeras i direkt solsken kommer, i jämförelse med andra korrekt monterade sensorer, att introducera bias. Hårdvarufel eller tillfälliga uppkopplingsproblem kan också leda till ofullständiga tids-

Av Ulf Seijmer, Induo och AKKR8



Ulf Seijmer är innovationschef på Induo samt CTO och medgrundare till hårdvaruföretaget AKKR8. Han är också styrelsemedlem i EUTECHs IoT Alliance. År 2012 skrev han en av de första artiklarna om 5G, en artikel som publicerades i Elektroniktidningen. Ulf är en frekvent anlitad föreläsare inom trådlös teknik för LPWAN och 5G.

serier. Här ligger en hel del på designerns bord. I mitt företag AKKR8 utvecklar vi sensorer som tar hänsyn till detta. De kan lagra tusentals datapunkter lokalt vid behov om de tappar kontakt med omvärlden och de filtrerar bort osannolika värden, vi kan till och med filtrera bort data utanför vissa "ointressanta" gränser. Detta är en bra startpunkt men samtidigt en risk för ytterligare bias om man begränsar filtreringen för mycket.

Obalanserad datainsamling i IoT

Samplingsbias är när data samlas in oproportionerligt under vissa tider, på vissa platser eller från specifika användargrupper. Algoritmer som tränas på statistiskt skev indata kan därför generera ogiltiga eller icke-generaliserbara resultat när de tillämpas på andra grupper eller i andra kontexter.

IoT-system samlar av naturliga skäl ofta in data från begränsade miljöer eller urval av befolkningen. Men, det krävs att begränsningen är genomtänkt, för ibland blir det fel. Ett sådant exempel är smarta städer, där sensorer oftare placeras i centrala områden medan områden utanför storstäderna eller stads-kärnorna inte täcks på samma sätt. Detta fenomen kallas ibland "sensoröknar" – zoner där lite eller ingen data samlas in. Resultatet blir att hela områden, levnadsmiljöer eller grupper blir osynliga i datamängden, vilket leder till snedvridna analyser och feltolkningar.

När data från sensorer med begränsade dataset används för trafikstyrning, luftkvalitetsövervakning eller nyttjandegrad av olika områden vid olika tider, riskerar beslut att baseras på

information från vissa områden – och därmed introducera geografiska och sociala skevheter i det offentliga beslutsfattandet. Sensorer på bara vissa platser visar inte på samhällets faktiska behov, sensoröknar kan spegla samhället på fel sätt.

Bias – det sociotekniska lagret

Låt oss återvända till temperatursensorn som monterades för varmt. Det är ett mänskligt misstag som formar utfallet som blir missvisande. Men det finns fler mänskliga beslut som formar arkitekturen i ett IoT-system. Om hårdvara och algoritmer utvecklas utan mångfald i åtanke kan bias smyga sig in på andra sätt än bara via sensordata. Till exempel förekommer det att dataset som används för ansiktigenkänning eller rösttolkning är snedfördelade – det förekommer överrepresentation av vita eller manliga användare. Här riskerar alltså välutvecklade IoT-lösningar att förstärka ojämlikheter. Ett annat exempel är bärbar teknik som smarta klockor, dessa använder ofta optiska sensorer för att mäta biometrisk data. Flera studier har visat att pulsmätning med grönt LED-ljus är betydligt mindre noggrann för personer med mörkare hudtoner samt för individer med fetma, de senare kanske den grupp som kan behöva tekniken allra mest.

Röstassistenter och användargränssnitt

Brittiska studier har visat att röstassistenter missförstår användare med icke-amerikansk accent nästan dubbelt så ofta som de med perfekt uttal. Från min egen erfarenhet av att använda Googles motsvarighet Home-

forma smarta beslut



Kommer någon av karaktärerna behöva åka tillbaka till kontoret för att hämta ett mindre batteri? Detaljer som AI missar.

assistant misslyckas jag med att prata med den för det mesta, tyvärr, men det kan bero på min svengelska. Felmarginalen ökar ytterligare när användare pratar dialektalt eller med informella uttryck. Dessutom är dessa assistenter ofta "feminiserade" i sin design – kritiker menar att det återspeglar medvetna eller omedvetna val som förstärker könsstereotyper och bidrar till att upprätthålla problematiska sociala normer.

Industri 4.0 och tillverkning

Min vision av prediktivt underhåll är att sensorer i framtiden skall lära sig att anpassa sig till maskinen de sitter på. I industriella miljöer används IoT-sensorer brett för att övervaka maskiners hälsa, genom att gå från kalenderbaserat underhåll, till underhåll när det behövs. I industrin används också sensorer för att optimera produktionsflöden. Men prediktivt underhåll som bygger på AI-modeller tränade med data från en viss typ av maskiner eller driftförhållanden kan ge felaktiga resultat när de appliceras på andra maskiner eller i andra miljöer. Jag kom för några år sedan i kontakt med en kund som tillverkade material-

transportörer för sågverksindustrin, sådana som flyttade sågspån från en plats till en annan. De berättade att när många kunder startade sin maskin så tänkte de inte på temperaturen. Med fuktigt spån eller material och en temperatur nära nollan så förkortades maskinens slitdelars livslängd betydligt, då materialet frös. Hade vi tränat en AI-modell på detta men glömt fästa fokus på temperaturen så hade vi missat en viktig variabel för korrelation.

Kvalitetskontroll

Om AI-algoritmer tränas enbart på en viss typ av data, data rensat från anomalier, finns risken att AI felklassificerar data som avviker från eller inte förekommer i träningsdata. I teorin är detta fullt möjligt i IoT-drivna tillverkningsmiljöer – särskilt när träningsdata är ofullständig eller snedfördelad. AI-modeller tenderar att föredra mellanmjölk framför avvikelserna.

Ett AI-system som tränats på snedvriden data kommer att prestera sämre och vara mindre tillförlitligt – särskilt i situationer där underrepresenterade användare ingår. Det här speglar ett djupare problem: IoT-system

lärs sig av världen såsom den har dokumenterats, inte såsom den faktiskt är. Och när dataset speglar samhällets befintliga ojämlikheter gör teknologin det också. Dåligt utformade tekniska system återspeglar inte bara bias – risken är att de förstärker den.

Skadat förtroende

Företag som ignorerar bias idag kan bli lätta byten och falla offer för kundkritik, negativ publicitet eller regulatorisk granskning. Trots att AI-tjänster är lätta att använda kräver de omsorg – det handlar om att undvika genvägar och istället arbeta noggrant.

Ett exempel är en AI-genererad bild jag fick skapad, föreställande en batteribytsituation i en IoT-sensor. Kommer någon av karaktärerna behöva åka tillbaka till kontoret för att hämta ett mindre batteri? Detaljer som AI missar.

Processer och medvetenhet

För att hantera bias måste IoT-företag aktivt analysera och justera sina data- och AI-flöden, det kan handla om såväl dataanalys och revision som kalibrering och säkring av datakvalitet.

Utöver tekniska åtgärder är

det avgörande att främja inkluderande designprinciper och jobba med expertis på vissa domäner under utvecklingen. Tvärfunktionella team – med ingenjörer, ämnesexperter och affärsanalytiker – kan förhindra att man förlitar sig på historiskt partisk data. Företag bör utfärda bias-riskdeklarationer och följa branschstandarder, så vi skapar AI-modeller med hög hygienfaktor.

Avslutande reflektioner: designa för rättvisa

Bias i IoT-data är ett komplext problem som kräver insikt i både tekniska och mänskliga faktorer. Tekniska orsaker (som sensorfel eller urvalsbias) och socioteknisk beslutsfattning (till exempel dataval, gränssnittsdesign) kan samverka och förstärka varandra. AI-integrering kräver att rättvisa byggs in i systemarkitekturen från början. Bias är inte bara en biprodukt av felaktig data – det är en systemdesignutmaning som måste hanteras från sensornivå till AI-modellen och vidare till beslutens logik.

Bias är redan ett hot i molnbaserade modeller som tränats på internetdata. Men när AI tar steget in i fysiska miljöer förändras spelplanen. I en AI-värld måste rättvisa designas in – inte hoppas på. Ingen data är neutral. Ingen sensor är helt opartisk. Varje gång vi mäter världen gör vi val.

Att begränsa sig för mycket kan innebära en risk för bias. Att redan på designstadiet begränsa målgruppen eller användningsområden kan innebära att man designar för ett bias. Därtill läggs nu ett ansvar på exempelvis sensortillverkare att enheter inte skall kunna generera data som är osannolik, exempelvis 500 graders rumstemperatur. Så där finns ett regulatoriskt krav att filtrera bort data som produkttillverkaren anser vara osannolik. Helt klart är detta en balansgång som börjar med en enkel fråga: Bygger vi system som verkligen ser alla – eller bara de som redan är överrepresenterade? ■