



DE OPINIE VAN ERIC POSTMA

'AI wordt niet superieur aan de mens'

De grootste valkuil van kunstmatige intelligentie (AI) is dat mensen zich laten intimideren door verhalen over het gevaar dat AI de mensheid gaat overnemen. "Dat is een verkeerde én gevaarlijke beeldvorming", zegt AI-professor Eric Postma. "Onderschat je eigen kracht niet."

Definities te over. Maar wat is kunstmatige intelligentie, ofwel Artificial Intelligence (AI) volgens Eric Postma? Jeugdige, vlote, informele en snelle formulering: "Ruwweg proberen met computertechnieken iets te doen of zelfs beter te doen dan wat mensen kunnen. Terug in de tijd: schaken. Computers kunnen heel snel vooruitdenken. Stop dus eerst de regels van het schaakspel en alle mogelijke schaakvarianten in de computer. Vervolgens weet de computer de beste zet te doen als de tegenstander een schaakstuk heeft verschoven. De computer heeft namelijk veel meer geheugen en rekenkracht. Daardoor kan hij beter schaken dan de mens."

"Een computer weet echter niet wat een schaakstuk of schaakbord is. Hij werkt in een abstracte wereld en weet alleen wat wij de formalisatie noemen: de regels, wat wel en niet mag met de stukken. Als je dat in de gaten houdt, snap je ook dat daar de kracht ligt van AI: snel en diep zoeken dankzij computerkracht. Dat is niks nieuws, want we hebben al heel lang de calculator die de mens de baas is."

Hoe verklaar je dan de huidige hyperigheid rond AI?

"Automatisch leren aan de hand van data is bewezen technologie waarvan we niet meer opkijken. De hype is echter gebaseerd op technologie die *deep learning* heet en al dertig jaar oud is. De laatste tien jaar is er echter een doorbraak, een soort *technology push*, dankzij vooral sterk toegenomen computerkracht en internet. Daardoor kun je systemen trainen als je maar voldoende data hebt."

WERELDKENNIS

De grote uitdaging voor AI is omgaan met dingen waar mensen beter in zijn en die we eigenlijk voor de hand vinden liggen, vervolgt Postma. "Wij zijn heel goed in patronen herkennen, zoals die stoel daar. Dat vinden we niet bijzonder. Maar men is vijftig jaar bezig geweest om een computer dat te leren. Pas een jaar of tien geleden is hierin die doorbraak gemaakt. Veel van wat je nu leest over AI, heeft daarmee te maken. Het herkennen van patronen in signalen, zoals geluid en spraak, of in beeld en tekst."

"Als je tekst analyseert met de computer, begrijpt deze de tekst niet zoals mensen dat doen. Maar hij kan er wel een soort semantiek – vage betekenis – uit halen. In de trant van: dit is een tekst over een bos of over huizen. Maar de details, die een beroep doen op wat ik noem wereldkennis, daar is een computer slecht in." →



ERIC POSTMA: "LAAT JE NIET INTIMIDEREN DOOR HET GEVAAR VAN AI."



Wat vind je een mooi voorbeeld van AI?

“Dat zit in het medische domein. Een patholoog kan huidkanker herkennen op basis van talrijke voorbeelden. Daarbij hoeft je niks te weten van de patiënt en slechts te kijken naar een stukje huid en visuele afwijkingen. Daarin is AI superieur aan de menselijke patholoog.”

“Het gevaar van dergelijke uitspraken is – en dat zie je steeds in de media – dat er een generalisatie van wordt gemaakt. Namelijk dat computers superieur worden aan de mens in het algemeen. Dat is niet zo, gelukkig zou ik bijna zeggen. De computer kan huidkanker in bepaalde vormen en gradaties perfect signaleren, maar weet niet dat het om een patiënt gaat en welke behandeling deze nodig heeft.”

Dat laatste is een kwestie van tijd, zou ik zeggen.

“Dat verwachten mensen. Maar ‘dat komt nog’ is veel complexer dan men denkt. Waar we het over hebben wordt wel ‘*narrow AI*’ genoemd. AI is ontzettend sterk in het kleine domein huidkanker herkennen. Maar neem nou een actuaaris. Deze kijkt bijvoorbeeld naar een patroon in financiële gegevens en documenten, en herkent daar een afwijking in. Dat kan een computer ook. Maar die actuaaris weet dat deze afwijking, ik noem maar wat, wordt veroorzaakt omdat het Sinterklaas is of een bepaalde regeling van kracht is. Natuurlijk kun je tegenwerpen dat je dergelijke informatie ook in de computer kunt stoppen. Maar dát is niet zo eenvoudig. Zeker niet omdat die informatie regelmatig verandert en omdat vaak een beroep wordt gedaan op algemene wereldkennis die niet in de krant of op internet staat. Kennis die je bezit omdat je in een financiële cultuur bent opgegroeid. Daar zit

de harde uitdaging van de AI. Dat noemen ze Artificial General Intelligence (AGI). Dat gebied schiet niet op. Het staat bijna stil. Er wordt wel voorspeld dat over dertig jaar machines dezelfde capaciteit hebben als het menselijk brein. Maar de redenering daarvoor laat zien hoe weinig de complexiteit van het menselijk brein wordt doorgrond.”

AUTORIJDEN IS GEEN MENSELIJKE INTELLIGENTIE MAAR EEN MENSELIJKE VAARDIGHEID

Wat is de huidige state of the art?

“De mens moet AI-leersystemen voorzien van trainingsmateriaal, zoals duizenden financiële rapportages, en die een classificatie meegeven, wat goed en fout is. Het getrainde systeem kan vervolgens in nieuwe documenten die het nooit eerder heeft gezien diezelfde classificaties genereren met een bepaalde betrouwbaarheid. Bijvoorbeeld een analysemethode om de essentie uit een tekst te halen zodat wetenschappers duizenden studies naar een bepaalde aandoening relatief snel kunnen screenen. Het gaat hierbij om een algemene semantische classificatie: waar gaat het over? Maar het gaat dus niet over details. Want die vereisen vaak specifieke kennis van wat die onderzoekers in hun studies hebben gestopt.”

Gaat bijvoorbeeld de zelfrijdende auto niet verder, namelijk van waarnemen en classificeren naar handelen?

“Autorijden is geen menselijke intelligentie maar een menselijke vaardigheid. Een vorm van *narrow AI*. Om auto te rijden, hoef je niet te weten hoe de wereld in elkaar steekt. Je hoeft alleen maar te weten wat de verkeersregels zijn en hoe je moet omgaan met andere weggebruikers. Dat laatste is het moeilijke aspect.”

“Zelfrijdende auto's zijn statistisch gezien nu al veiliger dan de mens. Op zich een reden om ze toe te laten. Ze zijn vooral betrouwbaar op snelwegen, waar vrij voorspelbaar verkeer rijdt. Het probleem zit 'm nog in de fijnmazigheid van met name stadsverkeer. Daar heb je te maken met talrijke subtiele regels. De hoeveelheid data daarover is te beperkt om die AI-systemen te kunnen trainen.”

“Stel dat je op een kruispunt zonder stoplicht staat. Zal ik eerst gaan of die ander...? Dat schat je in. Of neem iemand bij een zebrapad die aanstalten maakt om over te steken. De automobilist anticipeert daarop door te remmen. Maar als de persoon treuzelt en op zijn smartphone kijkt, trekt de automobilist weer op. Plotseling steekt de persoon toch over. Dan stop je weer. Dat zijn ontzettend ingewikkelde processen. Dat kan de AI wel leren op den duur, maar het is geen triviaal klusje.”

Wat kan AI betekenen voor het actuaariaat?

“De actuaaris kan AI onder meer gebruiken om patronen te herkennen of statische kansen in te schatten. Wat is bijvoorbeeld het gevaar van cybercrime als je daarvoor een verzekering ontwikkelt? Op basis van data over gevallen van cybercrime in meerdere branches kun je de machine voeden met leeralgoritmen, met kennis die in potentie hoogwaardiger is dan die van mensen om te voorspellen hoe groot het gevaar is. Daarin zie ik AI als een hulpmiddel en niet als gehele of gedeeltelijke vervanging.”

AI kan niet de beslissing nemen.

“Dat kan wel, maar dat zou ik liever niet doen. Want heel veel beslissingen doen een beroep op menselijke kennis die verder reikt dan de data zelf. Ook een actuaaris gebruikt deze ‘wereldkennis’, ervaringen in de maatschappij, in zijn of haar beoordeling. Om die kennis te vangen, moet je ontzettend veel meer data in de systemen stoppen. Hoeveel data zijn er niet in mensen gestopt vanaf hun geboorte? Wij vinden dat allemaal gewoon en eenvoudig, maar dat is geenszins het geval. We zitten niet voor niets zo lang op school en hebben heel veel ervaring om te kunnen bewegen in onze maatschappij. Dat is niet makkelijk te vervangen door een systeem dat wordt getraind op informatie uit databases. Het is niet onmogelijk, maar praktisch gezien heel erg lastig. Het duurt nog heel lang voordat we daar in de buurt van komen. Nee, over die termijn doe ik geen uitspraak.”

Wat zie je als belangrijkste ethische kant aan AI?

“De wisselwerking tussen technologie en mens is een van de belangrijkste. Beide moeten zich aanpassen zodat je tot een hybride symbiose komt. Mensen, zoals de actuaaris, blijven een expert die gebruikmaakt van nieuwe tools. Maar deze kun je pas goed gebruiken als je ze – grotendeels – begrijpt. Wat zijn hun sterke en zwakke kanten? Als je dat eenmaal leert, kan er een goede wisselwerking komen.”

“De vraag daarbij is in hoeverre je vertrouwt op het oordeel van zo'n algoritme en in hoeverre op het oordeel van jezelf. Dit geldt voor elk domein. In domeinen waar je dat vrij makkelijk kunt beoordelen, zoals van medici, zie je dat de systemen veel beter zijn dan een arts. Maar voor weerbarstige problemen die te maken hebben met wereldkennis, wordt het lastiger.”

“Machine learning is een dure naam voor een statistisch model dat je traint op verzamelde data. Machines kunnen met die statistiek liegen door verkeerde samples van data waarin een bias zit. Zo stond in de krant bijvoorbeeld dat AI seksistisch is. In werkelijkheid gaat het om een bias in de door mensen gegenereerde data.”

MAAR DE DETAILS, DIE EEN BEROEP DOEN OP WAT IK NOEM WERELDKENNIS, DAAR IS EEN COMPUTER SLECHT IN

“Een voorbeeld van zo'n seksistisch AI systeem is Google Translate. Gesuggereerde vertalingen van woorden in zogenaamde ‘gender neutrale’ talen zoals het Turks, geven de voorkeur aan de mannelijke variant. Zo is ‘arts’ in het Turks ambigu op geslacht: mannelijk of vrouwelijk. Maar de Engelse vertaling kiest meestal de mannelijke variant omdat die het meest voorkomt. Dat is een statisch goed te verklaren bias. Dat wordt dan opgepikt door de pers, die er een racistische interpretatie aan verbindt. Maar het is in feite een bias in de data, voorkomend uit de ongelijke behandeling van mannen en vrouwen in onze samenleving. Kortom, het gaat erom dat er, als de actuaaris een belangrijk besluit moet nemen en zijn daarbij ondersteunende systeem meer nadruk legt op bepaalde afwijkingen of onregelmatigheden, een bias ontstaat.”



Hoe raak je bewust van zo'n bias?

"Dat ligt nou aan de vaardigheid van een actuaaris. Je hebt met AI niet zomaar een tool, je moet deze leren kennen en de beperkingen ervan weten. Dan kun je twee dingen doen. De tool verbeteren door te trainen op betere data. Of je beslissing aanpassen door rekening met de bias te houden."

"Een en ander kun je uitsluitend oplossen door met zo'n tool aan de gang te gaan. Ga gewoon eens spelen met wat data en doe eenvoudige experimentjes met machine learning. Of ga een cursus volgen."

Het interview vindt plaats in het voormalige klooster (en kweekschool) van de congregatie der Zusters van de Heilige Maagd Maria in 's-Hertogenbosch. Tegenwoordig is er de Jheronimus Academie voor Data Science (JADS) gevestigd, waaraan Postma verbonden is. "Bij JADS hebben we bijvoorbeeld de executive education. Dat is een cursusprogramma voor bedrijven en instellingen om met data science om te gaan. Dan oefen je met algoritmen op je eigen data en leer je heel snel de beperkingen en kracht kennen van wat je kunt doen."

Voor AI heb je veel knappe koppen nodig. Zijn die er wel?

"Nee, maar die komen er wel dankzij de opleidingen voor data scientist en AI-onderzoeker. Zo leidt JADS tijdens een masterprogramma mensen twee jaar lang op om data science te combineren met ondernemerschap. Om die te vermarkten en in te zetten bij bedrijven."

"Aan Tilburg University volgen bachelors met verschillende achtergronden, zoals business en legal, een eenjarige master *Data Science for Society* waar ze alle basistechnieken uit de data science leren, inclusief de juridische en economische aspecten. Daarmee kunnen ze een linking pin vormen tussen allerlei afdelingen. Zo ging de afstudeerscriptie van een student politieke wetenschappen over hoe je data science kunt toepassen om te voorspellen wanneer conflicten tussen naburige naties ontstaan. Hij had daartoe een omvangrijke database van nieuwsberichten uit twee landen over elkaar. Hij gebruikt zijn bachelorkennis in het domein van de data science."

Wat is de belangrijkste valkuil voor een actuaaris?

"Dezelfde als voor alle professionals. Dat je je laat intimideren door verhalen over het gevaar van AI dat de mensheid gaat overnemen, zoals Elon Musk en wijlen Stephen Hawking suggereerden. Dat is een verkeerde én gevaarlijke beeldvorming omdat ze suggereren dat de technieken veel slimmer zijn dan ze werkelijk zijn. Natuurlijk, AI kan de kracht om iets te doen enorm verhogen. Maar onderschat je eigen kracht niet, gebouwd op jarenlange intuïtie en ervaring. Die kun je enorm verrijken door AI-technieken te gebruiken. Maar je kunt deze niet vervangen." ■



Eric Postma (1961) is professor in Artificial Intelligence (AI) aan het Cognitive Science & AI departement van Tilburg University (2008) en de Jheronimus Academie voor Data Science (JADS) in 's-Hertogenbosch (2016). Hij is gespecialiseerd in beeldverwerking, signaalverwerking, affectieve computerverwerking en deep learning. Zijn bekendste project betreft de automatische analyse van schilderijen van Van Gogh. Met behulp van geavanceerde beeldanalysemethoden ontwikkelde Postma systemen voor het bepalen van de auteur van kunstwerken.

Postma heeft, na eerst een jaar natuurkunde, cognitieve wetenschappen gestudeerd. Daarna heeft hij onder meer studies in neurowetenschappen, informatica, AI en neurale netwerken gevolgd. In 1994 is hij gepromoveerd aan de Universiteit Maastricht.

"De rode draad in mijn carrière is altijd mijn interesse in mensen. Ik gebruik de computer om mensen te begrijpen. Vandaar dat ik de wisselwerking tussen mens en technologie zo interessant vind. De meeste voorspellingen die je hoort over AI zijn óf van technologen die de psychologie niet begrijpen, óf van psychologen die de technologie niet begrijpen. Ik zit in het midden."