

Algoritmes leiden niet automatisch tot eerlijkere selectie

Arbeid & Economie Wetenschap en innovatie

Steeds meer werkgevers zetten algoritmes in voor werving en selectie. Bij Nike en Unilever is het niet meer ongevoel om pas in de derde ronde van de sollicitatieprocedure in gesprek te komen met een menselijke medewerker. Vooroordelen en ongelijkheid kunnen zo meer ruimte krijgen, volgens Annemarie Hiemstra en Isabelle Nevels.

Door **Annemarie Hiemstra, Isabelle Nevels**
3 december 2018

Het aantal aanbieders van geautomatiseerde en digitale selectiemethoden neemt toe. Zo voorspelt het algoritme van het Nederlands-Chinese bedrijf Seedlink op basis van taalgebruik van sollicitanten wie er het meest geschikt is voor een functie en maakt het Amerikaanse HireVue gebruik van video-opnames om tot een automatisch gegenereerde beoordeling te komen.

Dit sluit nauw aan bij het huidige debat. Is het nog wel nodig om klassieke kennis, karaktereigenschappen en competenties, typisch gemeten met psychologische vragenlijsten en assessments, in kaart te brengen? Wellicht dat meer impliciete patronen, die via algoritmes herkend worden op basis van grote volumes data (zoals geschreven tekst, of video's van sollicitanten), even goede of zelfs betere resultaten kunnen behalen bij het selecteren van de beste kandidaat.

Organisaties hebben slechts de beschikking over hun eigen inschatting van de meerwaarde en eerlijkheid van selectie via computerondersteuning. Wetenschappelijke bewijs voor het gebruik van technologieën is er echter nog nauwelijks. Hoe kunnen we ervoor zorgen dat computersystemen op een verantwoorde manier worden ingezet?

Meer inclusiviteit dankzij de computer?

Voorstanders van het gebruik van artificiële intelligentie voor personeelsvraagstukken, zoals algoritmes om kandidaten te screenen, stellen dat het kan leiden tot meer inclusieve werkomgevingen. Verschil in baankansen en promotie voor werkzoekenden en werknemers vanuit diverse minderheidsgroepen is immers nog steeds een veelvoorkomend probleem.

In veel Europese landen worden immigranten harder getroffen door werkloosheid. In Nederland is het werkloosheidspercentage van inwoners met een migratieachtergrond meer dan twee keer zo groot als dat van autochtonen. Daarnaast werken veel etnische minderheden momenteel onder hun opleidingsniveau. Dit heeft veel te maken met het feit dat minderheidsgroepen gemiddeld slechter scoren op de meest valide selectie-instrumenten.

Computer heeft geen raciale voorkeur

Voor veel organisaties is het dus van groot belang om gebruik van discriminerende en oneerlijke selectiemethodes te verminderen, terwijl kwaliteit van de wijze van selectie wel gewaarborgd wordt. De komst van selectiealgoritmes lijkt hiervoor een veelbelovende oplossing.

In tegenstelling tot menselijke beoordelaars kan een computer namelijk geen slechte dag hebben. Bovendien heeft een computer meer capaciteit om informatie te verwerken in een fractie van de tijd die een recruiter eraan kwijt zou zijn. Waar een menselijke beoordelaar een sollicitant die veel op hem of haar lijkt zal prefereren, wordt een computer niet geleid door eigen voorkeuren. Dit kan bijdragen aan meer gelijke kansen voor kandidaten van verschillende achtergronden.

Algoritmes kunnen vooroordelen bevatten

Tegelijkertijd is er een groeiend besef, onder psychologen en computerwetenschappers, dat algoritmes voor selectie gemaakt zijn door mensen en dus gebaseerd zijn op menselijke input. Het algoritme zelf is niet bevooroordeeld en behandelt alle sollicitanten gelijk, maar als deze input vertekend is (*biased*), dan leert het algoritme ook om al deze vertekeningen mee te nemen. Zelfs als de persoon die het algoritme maakt of implementeert de beste bedoelingen heeft, kan een algoritme biased zijn.

Een goede illustratie hiervan is de door Microsoft uitgebrachte chatbot, genaamd Tay. Het onderliggende algoritme van Tay was zo ontwikkeld dat het kon leren van online conversaties met gebruikers op verschillende sociale media platformen. Al binnen 1 dag maakte de robot echter zulke racistische en seksistische opmerkingen en posts, dat Tay offline moest worden gehaald.

Een ander, subtieler voorval van algoritmische discriminatie op basis van geslacht, werd in 2015 aan het licht gebracht door wetenschappers aan de Carnegie Mellon University. Deze ontdekten dat advertenties op Google voor hoogbetaalde banen zes keer minder vaak getoond werden aan vrouwen dan aan mannen.

Selectiealgoritmes kunnen sociale ongelijkheden vergroten

In de organisatiecontext worden selectiealgoritmes vaak gebaseerd op historische data of op kenmerken van personen die succesvol zijn in een organisatie. Hierbij worden scores van sollicitanten vergeleken met uitslagen van de best presterende werknemers in een bedrijf.

Wanneer deze groep bestaande werknemers niet representatief is voor de brede pool aan sollicitanten, dan kan dit nadelig zijn voor sollicitanten uit achtergestelde groepen. Dit geldt zowel voor selectieprocedures, als voor beslissingen promoties van bestaande werknemers.

Dit toont aan hoe machines accurate correlaties kunnen vinden in de data van een bedrijf, die desondanks discriminerende resultaten laten zien. In plaats van vermindering van sociale ongelijkheid kan computer-gebaseerde beslisondersteuning dus juist toetreding tot de arbeidsmarkt bemoeilijken en sociale ongelijkheden vergroten.

Onduidelijk welke factoren meegewogen worden

Een ander probleem bij digitale selectiemethoden is dat er vaak wel honderden variabelen worden meegenomen in de formules, waardoor ze te gecompliceerd zijn om geïnterpreteerd te kunnen worden. Dit brengt een aantal ethische vraagstukken met zich mee.

Als we niet precies weten welke factoren zijn meegenomen in de voorspellingen van zo'n algoritme over sollicitanten, hoe kunnen we er dan van verzekerd zijn dat het algoritme op een eerlijke manier heeft besloten welke kandidaat het meest geschikt is voor een baan? Hoe kunnen we waken voor de gevaren van het overmoedig inzetten van deze systemen en zorgen voor verantwoord gebruik van deze technologieën?

Algoritmische controles bij bedrijven

De eerste stap hierbij is bewustwording. Het is belangrijk om kritisch te blijven ten opzichte van deze snel evoluerende digitale methodes. Op dit moment worden de eerste methodes al ontwikkeld om algoritmische bias te detecteren en te voorkomen.

[Recent onderzoek](#) kijkt naar de mogelijkheid om data voor de besluitvorming zo te modificeren dat een algoritme minder snel discriminerende voorspellingen maakt. Tevens pleiten sommige onderzoekers voor het [houden van algoritmische controles bij bedrijven](#), waarbij er door behulp van [speciale technieken](#) vastgesteld kan worden of er sprake is van oneerlijkheid in softwaresystemen. Omdat dit onderzoeksgebied pas ontstaan is, bestaat er nog geen duidelijke overeenstemming over de beste manier om het functioneren van algoritmen te meten.

Sociale- en computerwetenschappers moeten kennis bundelen

Daarnaast hoeven klassieke en digitale manieren van assessment elkaar niet uitsluiten. Het is belangrijker om te verkennen op welke manier de sterke kanten van psychologische en technologische benaderingen kunnen worden gecombineerd, met als doel eerlijkere en verbeterde personeelsbeslissingen te nemen.

Een essentiële factor om dit te kunnen verwezenlijken is wederzijds begrip tussen sociale wetenschappers die geen IT-achtergrond hebben en computerwetenschappers. Het verantwoord gebruiken van big data en artificiële intelligentie vraagt om gecombineerde kennis en vaardigheden tussen de beoefenaars van deze twee wetenschappen.

Om dit in de toekomst te bereiken, zijn enkele universiteiten en organisaties, waaronder de Erasmus Universiteit Rotterdam en de Technische Universiteit Delft, een [Strategisch Partnerschap](#) aangegaan, genaamd Big Data in Psychological Assessment (BDPA). Deze samenwerking dient om beter in te spelen op deze technologische ontwikkelingen en dit te integreren in nieuw en verbeterd onderwijsaanbod voor aankomend organisatiepsychologen en computerwetenschappers.

Dit soort interdisciplinaire samenwerkingen tussen computerwetenschappers en sociaal wetenschappers is ook de eerste stap in het bereiken van een consensus over de betrokken procedures bij het eerlijk inzetten van algoritmische technieken binnen de werving- en selectie.

Annemarie Hiemstra is universitair docent organisatiepsychologie aan de Erasmus Universiteit Rotterdam en verbonden aan [BDPA](#). Isabelle Nevels is Master student Arbeids- en Organisationspsychologie en Business Information Management aan de Erasmus Universiteit Rotterdam.

Foto: [Leigh Anthony DEHANEY](#) (Flickr Creative Commons)