

Software bespaart wetenschappers duizenden uren lezen per jaar

Universiteit Utrecht zet algoritme in om eerdere onderzoeken te beoordelen

Jan Fred van Wijnen
Amsterdam

Kunstmatige intelligentie kan een belangrijk deel van het onderzoek naar ziekten en psychische stoornissen overnemen. Nieuwe software van onderzoekers aan de Universiteit Utrecht kan zelf studies selecteren, die worden gebruikt om behandelingen te verbeteren.

Hiermee wordt wetenschappers tijd-rovend werk uit handen genomen. Het gaat in Nederland om tienduizenden werkdagen per jaar, stellen de Utrechtse wetenschappers. De loonkosten hiervoor lopen in de miljoenen.

De software is ontwikkeld door Rens van de Schoot, pas benoemd als hoogleraar statistiek voor 'kleine databestanden'. Het algoritme — het slimme hart van het programma — selecteert studies die worden gebruikt in zogeheten 'systematische overzichtsartikelen' of reviews.

Bij veel ziekten en mentale stoornissen wordt het aantal studies naar thera-

pieën te groot om te overzien. Daarom schrijven wetenschappers regelmatig overzichtsartikelen, waarin ze alle publicaties over een behandeling vergelijken. Zo worden nieuwe richtlijnen voor artsen en therapeuten ontwikkeld.

Maar het aantal studies naar therapieën wordt zo groot, zegt hoogleraar klinische psychologie Pim Cuijpers aan de Vrije Universiteit in Amsterdam, 'dat ook de analyses van de studies steeds omvangrijker worden'.

Zelf is Cuijpers betrokken bij een onderzoek naar biologische en psychologische verklaringen voor depressie. Daarvoor lezen twee wetenschappers elk 150.000 samenvattingen van studies. De

gemiddelde onderzoeker leest er 43 in een uur. Bij een achturige werkdag waarin ze non-stop lezen, zouden ze met deze klus anderhalf jaar bezig zijn.

Dit onderzoek is wel een uitschieter, zegt Cuijpers. Meestal beperkt het werk zich tot het lezen van enkele duizenden samenvattingen. Dat is 'enkele maanden' werk voor twee onafhankelijke onderzoekers, zegt hij. Voor een gemiddeld onderzoek in de psychologie rekent hij met 15.000 samenvattingen (leestijd: bijna tachtig dagen voor twee onderzoekers samen). Die worden teruggebracht tot 2200 studies om volledig te lezen.

Onderzoekers lezen zoveel samenvattingen omdat het selecteren nog een groot proces is. Ze typen zoektermen in een digitale bibliotheek, en krijgen vervolgens vele duizenden samenvattingen. De meeste zijn niet relevant, maar dat weten ze pas als ze alles hebben gelezen.

'Je wilt niet één artikel missen dat relevant is voor je onderzoek', zegt Cuijpers. 'Daarom kost het zoveel tijd.' Als het al-

goritme van de Utrechtse projectgroep 100% betrouwbaar is, zegt hij, 'dan gaat dat ons heel veel tijd schelen'.

Het algoritme, dat nog verder wordt ontwikkeld, reduceert het aantal samenvattingen tot nu toe met bijna 80%, door automatisch alles te filteren wat niet relevant is. Bij die score wordt maximaal één studie ten onrechte geschrapt.

De Universiteit Utrecht registreerde dit jaar 283 onderzoeken als systematische overzichtsartikelen. Als elk onderzoek begint met 6000 samenvattingen (een aanname van de groep-Van de Schoot), dan zaten Utrechtse wetenschappers dit jaar ruim 11.000 werkdagen te lezen. 'Onze software kan dat binnen één dag', zegt Van de Schoot.

Wetenschappers zijn nu nog veel tijd kwijt aan het inventariseren en beoordelen van bestaande publicaties



'Zesduizend uittreksels doorspitten?'
Pagina 6