

# Het bewijs zit in de pudding

Martin Heijnsbroek  
zaterdag 18 juni 2016

## Geen enkel datamodel heeft een eeuwig leven

Of je nu als voedingswetenschapper werkt of als ‘chief data scientist’ in een bedrijf, je wilt graag verbanden in de data vinden. Want daarmee ga jij het verschil maken, en jouw visie er doorheen krijgen. En zo kun je jezelf als autoriteit in de wetenschap of in je organisatie neerzetten.

Maar heb je echt een andere kijk op de data dan de rest, dan mag je de mouwen opstropen. De gevestigde orde zal de nieuwe data snel als uitzondering beoordelen, die je niet serieus hoeft mee te nemen. En als je doorzet, en een bestaand model omver dreigt te gaan werpen, dan word je op hun autoriteit of expertise aangevallen of, nog erger, doodgezwegen.

Maar soms is er geen houden aan: het paradigma is niet meer te verdedigen en gaat schuiven. Zoals nu in de voedingswetenschappen, waar het meer dan vijftig jaar oude paradigma over cholesterol en verzadigde vetten onder vuur ligt. Zelfs bankiers adviseren investeerders nu dat het perspectief aan het kantelen is. Het onderzoeksteam van Credit Suisse kwam, na een jaar lang ruim 400 recente wetenschappelijke artikelen en boeken bestudeerd te hebben, tot duidelijke conclusies: het eten van cholesterol heeft geen invloed op de cholesterolwaarde in het bloed, en de link tussen verzadigde vetten en cardiovasculaire ziekten wordt door de wetenschap niet aangetoond.

Ook in de wereld van big data bestaat het risico van het verkramppt omarmen van het verkeerde beeld, vaak op basis van complexe modellen. Want hoe sterker jouw data scientist zich identificeert met zijn of haar superieure modellen, hoe sterker de afwijzing van een goed bedoeld idee om het anders te doen. En als je bijvoorbeeld een echt geavanceerd video-analysemodel — gebaseerd op kunstmatige intelligentie heeft zo'n model meer dan een miljard parameters — ter discussie wilt stellen, dan moet je wiskundig van zeer goeden huize komen.

Dus hoe voorkom je die weerstand om te blijven vernieuwen met modellen, en zorg je dat je niet vijftig jaar lang het verkeerde model omarmt? Als eerste moet je een cultuur proberen te bouwen waarin ruimte is voor sceptische gedachten die soms tegen de stroom in gaan. Net zoals bij LinkedIn, waar het sceptisch denken als een kwaliteit wordt omarmd. En als bestuurder moet je daarin het voortouw nemen, door afstand te houden van het soms religieuze geloof in modellen.

Maar je moet je vooral niets aantrekken van al die statistische logica waarom de resultaten echt significant zijn. Investeer in experimenten om uit te vinden of de uitkomsten van je model in lijn zijn met wat je waarneemt. Het liefst niet eenmalig, maar als onderdeel van een continu proces. Want het bewijs vind je niet in de statistiek, maar gewoon in de pudding — ook al is die niet zo gezond.

*Investeer in experimenten om te zien of de uitkomsten van je model in lijn zijn met wat je waarneemt.*

*Martin Heijnsbroek*

*Dus jij gelooft pas dat vette friet ongezond is als je er heel veel van hebt gegeten?*



*Jan Fred van Wijnen, chef Morgen*