

# Leren door verwarring

Alhoewel verwarring vaak kan leiden tot vervelende situaties, zijn er ook situaties waarin verwarring leidt tot meer begrip. In dat geval is het misschien beter om te spreken van consistentie. 'Als dit zo is, dan moet dat toch zus zijn?', is dan de redenatie. En als die logica niet op gaat – oftewel, de redenatie zorgt voor verwarring – is het misschien beter om de redenatie te veranderen. Men zoekt op die manier naar consistentie tussen verschillende stukjes kennis. Voor het leerproces kan verwarring dus essentieel zijn. Door continu te proberen deze verwarring weg te nemen is er een duidelijk doel: als de verwarring weg is, heeft u het begrepen. Dat dit principe ook gebruikt kan worden door computers, is onlangs aangetoond in een publicatie waarin faseovergangen in de natuurkunde geïdentificeerd worden via machinaal leren. Evert van Nieuwenburg

280

## Machinaal leren en vastestoffysica

Om een computer iets te leren aan de hand van voorbeelden, is het essentieel dat er grote hoeveelheden data beschikbaar zijn. Tegenwoordig worden er ook in de natuurkunde grote hoe-

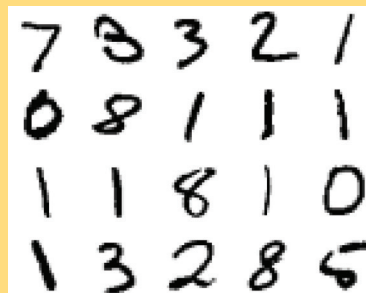
veelheden data geproduceerd; denk bijvoorbeeld aan de simulaties en metingen van deeltjesversnellers. Machinaal leren wordt daarbij al langer toegepast om data te interpreteren en te verwerken. Nieuw is het gebruik van machinaal leren voor vastestoffysica.

Daar staat het concept van faseovergangen centraal en het onderliggende idee is dan om een computer te leren verschillende fasen te herkennen. Volgens is het dan de vraag of de faseovergangen tussen die verschillende fasen correct voorspeld worden.

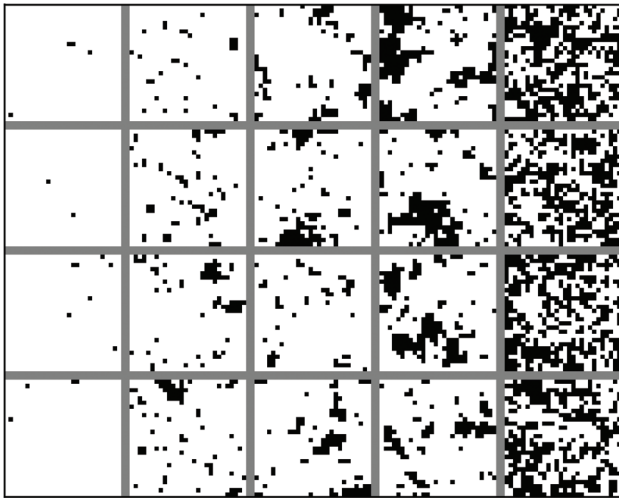
Als voorbeeld kunnen we kijken naar het Ising-model in twee dimensies. In het Ising-model worden kleine magneten beschreven die de twee polarisatietoestanden met aanduiding  $-1$  of  $+1$  kunnen aannemen en vastzitten op een vierkant rooster. Deze waarden stellen de spin van elektronen voor die in twee verschillende richtingen kunnen wijzen. Magnetten die zich (horizontaal of verticaal) naast elkaar bevinden wisselwerken met elkaar en willen graag in dezelfde richting wijzen. Dat is ook precies wat er gebeurt bij lage temperaturen; de magneten wijzen in grote meerderheid dezelfde kant op en het systeem vormt een ferromagneet. Bij hoge temperaturen echter is er genoeg energie in het systeem zodat de magneten willekeurig van  $-1$  naar  $+1$  heen en weer fluctueren en zijn er geen geordende patronen zoals clusters te herkennen. Dit is de paramagnetische fase van het

## Machinaal leren

Bij machinaal leren gaat het er om een computer een taak te laten volbrengen, zonder deze precies te vertellen welk algoritme uitgevoerd moet worden. In plaats daarvan wordt de computer getraind aan de hand van heel veel voorbeelden, een proces dat *supervised learning* heet (leren onder toezicht, ook wel gecontroleerd leren genoemd in het Nederlands). Het bekendste voorbeeld daarvan is het trainen van een zogenoemd neurale netwerk voor het herkennen van handgeschreven getallen. Het neurale netwerk krijgt voorbeelden te zien van plaatjes van handgeschreven getallen tussen 0 en 9 (zoals te zien in figuur 1) en wordt ook verteld welk getal er te zien is in elk plaatje. Aan de hand van deze voorbeelden (en nog ongeveer 60.000 meer) leert de computer handgeschreven getallen te herkennen met meer dan 99% betrouwbaarheid.



**Figuur 1** Voorbeelden van handgeschreven getallen die een computer leert herkennen. Elk getal is een plaatje van 28 bij 28 zwart-wit pixels.



**Figuur 2** Voorbeelden van configuraties in het tweedimensionale Ising-model op een rooster van 28 bij 28 pixels (vergelijk figuur 1). Elke pixel stelt een magneet voor die de waarde +1 heeft als de pixel zwart is en -1 voor wit. De plaatjes in de meest linker kolom zijn ver onder de kritieke temperatuur en die in de meest rechtse zijn er ver boven. De middelste drie kolommen zijn dicht bij de kritieke temperatuur en zijn moeilijk te onderscheiden.

systeem. Bij een kritieke temperatuur vindt er dus een faseovergang plaats tussen de ferromagnetische fase en de paramagnetische fase. In figuur 2 is een aantal van dit soort configuraties afgebeeld. Bij erg lage en hoge temperaturen zijn de verschillen duidelijk te zien, maar rond de faseovergang is het erg moeilijk voor een mens te zeggen of het systeem een ferro- of een paramagneet is.

Om de computer deze configuraties te leren, moeten we de configuraties van te voren al labelen als ‘ferromagneet’ of ‘paramagneet’. In het beste geval hoeven we dat alleen te doen voor de configuraties waarvoor het overduidelijk is (namelijk die ver onder en boven de kritieke temperatuur). In het ergste geval moeten we ze allemaal labelen en dat is natuurlijk eigenlijk valsspelen. Het gaat er namelijk om dat de computer zelf uitvindt dat er twee fasen zijn met een faseovergang daartussen. Dit laatste probleem is het focuspunt van *unsupervised learning* (ongecontroleerd leren). Als we de fasen al kenden, hoefden we immers niet een computer te leren om ze te vinden!

*Unsupervised learning* is in het algemeen een veel moeilijker taak, maar dit is het punt waar de verwarring en de consistentie uit de inleiding van pas komen. Met behulp van die concepten kunnen we het *unsupervised* probleem effectief in een *supervised* probleem omzetten.

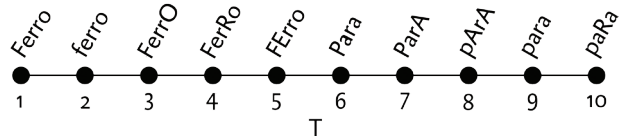
### Verwarring

De situatie is nu zo, dat we een grote hoeveelheid data hebben die (nog) niet gelabeld is. Laten we er voor het gemak even vanuit gaan dat deze data verkregen zijn uit een experiment of simulatie en dat deze geordend zijn aan de hand van een parameter  $T$  (zoals de temperatuur voor het Ising-model). We gaan er ook vanuit dat er een faseovergang plaatsvindt als functie van deze parameter en dat de faseovergang zich bevindt bij de waarde  $T_c$ , maar deze waarde is voor ons onbekend.

Ons doel is nu om een labeling te vinden van de data, zodat alle datapunten met  $T < T_c$  het label 0 krijgen en alle andere datapunten het label 1. Hoe? Door slim te gokken en gebruik te maken van *supervised learning*. Dit kunnen we aan de hand van een voorbeeld demonstreren.

In figuur 3 zijn onze fictieve datapunten afgebeeld, bestaande uit verschillende spellingen van de woorden ‘ferro’ en ‘para’. De datapunten zijn genummerd en ze zijn geordend volgens een fictieve parameter  $h$ . Dit zijn onze fasen en de variatie in de spelling is ter indicatie dat niet alle data voor een bepaalde fase hetzelfde zijn. In plaats van deze woorden worden voor de echte toepassing uiteraard configuraties zoals in figuur 2 genomen.

Het is voor ons duidelijk dat de faseovergang zich tussen datapunt 5



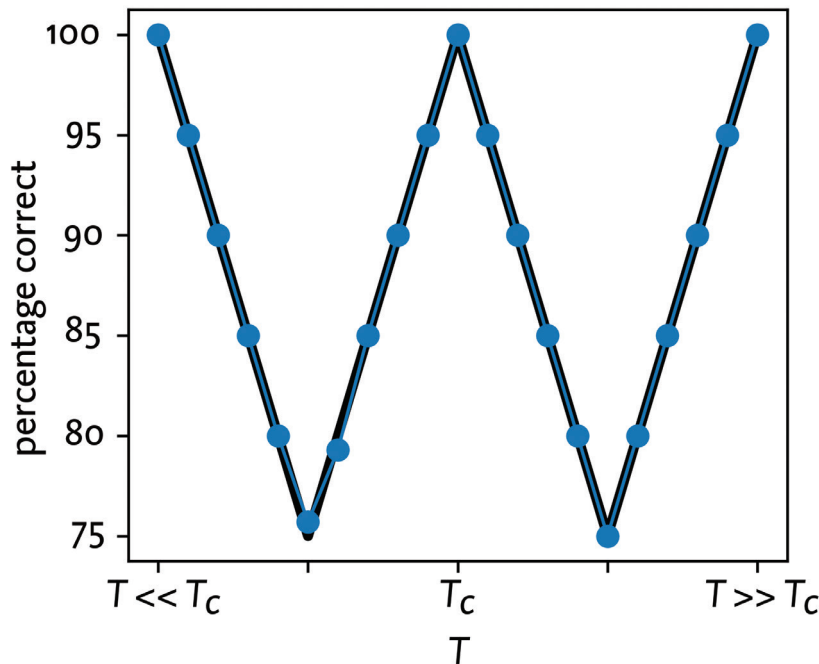
**Figuur 3** Fictieve data (verschillende spelling van de woorden ‘ferro’ en ‘para’) ter demonstratie van het verkrijgen van consistente labels. Voor ons als mens is het simpel te zien dat de overgang tussen datapunten 5 en 6 ligt. Om dat met een computer te vinden wordt de overgang eerst gegokt en dan wordt er gekeken of dit een consistente labeling oplevert.

en datapunt 6 bevindt, maar we doen alsof we dat niet weten. We gokken eerst dat de faseovergang zich tussen datapunten 1 en 2 bevindt en kennen dus label 0 toe aan alles links van deze gegokte faseovergang (in dit geval alleen aan datapunt 1) en label 1 aan de rest. We hebben nu dus data met een verkeerde labeling, maar we weten nog niet dat deze labeling verkeerd is. Om uit te vinden of deze labeling verkeerd is of niet, proberen we deze gelabelde data nu aan een computer te leren. Vier keer ziet de computer ‘ferro’ met label 1 en maar een keer met label 0. De computer denkt dus dat ‘ferro’ label 1 krijgt. Als we vervolgens vragen wat de computer denkt dat het eerste datapunt voor label zou krijgen, krijgen we als antwoord label 1. De eerste ‘ferro’ hebben we echter als 0 gelabeld en de computer is dus incorrect. In totaal zijn dan 9 van de 10 punten goed, een score van 90%. Gokken we de faseovergang tussen plaatjes 2 en 3, dan behaalt de com-

Evert van Nieuwenburg studeerde natuurkunde in Leiden en promoveerde vervolgens aan de ETH Zürich in Zwitserland. Hij werkt nu als postdoc bij het Californian Institute of Technology.



evert.v.nieuwenburg@gmail.com



**Figuur 4** Het percentage correct gelabelde datapunten als functie van de gegokte labeling. De labeling is gedefinieerd door het punt  $T$ , dat de overgang tussen labels 0 en labels 1 aangeeft. De middelste piek kan nu worden geïdentificeerd als de kritieke temperatuur  $T_c$ .

puter een score van 80%. En mochten we de faseovergang op de juiste plek

gokken, dan kunnen alle datapunten consistent gelabeld worden en beha-

len we een score van 100%! Merk ook op dat als we alle data hetzelfde label geven, we eveneens een score van 100% halen; maar deze maxima zijn triviaal en vertellen ons niets over het overgangspunt.

De score als functie van de gegokte faseovergang laat dus een maximum zien waar de faseovergang op de juiste plek gegokt wordt, zoals te zien in figuur 4 voor het Ising-model. Als functie van de gegokte kritische temperatuur  $T$  is er een duidelijke piek te zien die de echte kritische temperatuur  $T_c$  identificeert. Dat dit ook werkt voor gecompliceerdere data is beschreven in de publicatie waarin deze methode ook gebruikt wordt om een topologische faseovergang en een faseovergang in de eigentoestanden van een quantummechanisch systeem met wanorde te vinden.

### Referenties

- 1 E.P.L. van Nieuwenburg, Y.-H. Liu en S.D. Huber, *Nature Physics* **13**, 435-439.
- 2 J. Carrasquilla en R.G. Melko, *Nature Physics* **13**, 431-434.

## Boekbespreking



# Samen met de CIA

Op het *Journal* en in de krant [1] was er aandacht voor dit boek, maar vooral vanwege aspecten waar de ondertitel van dit boek op slaat: *Operaties achter het IJzeren Gordijn*. Voor natuurkundigen is juist het deel van het boek dat gaat over wat zich voor het IJzeren Gordijn afspeelt interessant. Het langste hoofdstuk gaat namelijk over, zoals de auteur Cees Wiebes zegt, “De moeder van alle geheime kwesties: nucleaire spionage”. Ik beperk mij tot de bespreking van dit deel van het boek. De aanprijzingen van de media vindt u in de genoemde artikelen.

Terwijl de Amerikanen nog nauwgezet hun atoomgeheimen bewaakten,

de eerste spion uit het Manhattanproject, Klaus Fuchs, ontmaskerd werd en de dan nog niet zo bekende Bruno Pontecorvo overliep naar de Russen, bouwden de Noren en de Nederlanders samen een kernreactor. De Noren hadden het zware water ( $D_2O$ ) en de Nederlanders het uranium. Wander de Haas had dat nog net voor de Tweede Wereldoorlog aangeschaft, vrijwel direct nadat kernsplijting was ontdekt (begin 1939). Na de oorlog wil de belangrijkste natuurkundige van Nederland uit die tijd, Hendrik Kramers, de opgelopen achterstand van de Nederlandse natuurkunde inhalen. Een zo snel mogelijke ontwikkeling van de kernfysica is daarvoor gewenst. De re-

actor wordt gebouwd in Kjeller, Noorwegen. Buiten de grootmachten om is er opeens een plek waar iedereen mee mag doen aan de kernfysica en dat terwijl de Koude Oorlog vol op stoom is. Niet gek dus dat er spionnen rondlopen. En het toppunt van nucleaire spionage is natuurlijk spionage met dubbelspionnen en dat is waar Cees Wiebes verslag van doet.

Deze roemruchte periode van de Nederlandse kernfysica in Kjeller is uitgebreid beschreven door bijvoorbeeld Jacob Goedkoop [2], de eerste directeur van de reactor in Petten, maar het woord spionage komt bij hem niet voor. Wiebes' boek is daarom een welkome aanvulling op de geschiedenis