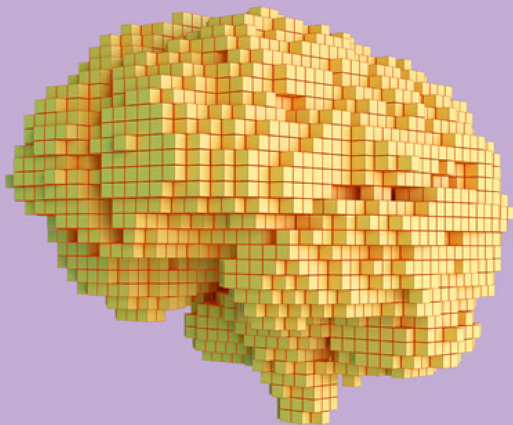


Remy Gieling

Starten en groeien met kunstmatige intelligentie



ONTDEK DE GROEI- KANSEN VAN AI

Lessen van
Picnic, Salesforce,
ParkBee en andere
succesformules

VANDUUREN
MANAGEMENT

mt/sprout

Starten en groeien met kunstmatige intelligentie

ONTDEK DE GROEI- KANSEN VAN AI

Lessen van Picnic, Salesforce, Parkbee en
andere succesformules

VANDUUREN
MANAGEMENT

ISBN 978 90 8965 509 7
NUR 801
Thema KJH, KJS, KJC
Trefwoord Artificial Intelligence, innovatie, groeikansen, toekomst

Auteur: Remy Gieling
Eindredacteur: Philip Bueters
Artdirection en omslagontwerp: Atomo Design, Amsterdam
Omslagillustratie: Afiat Sukmaraga
Fotografie: Jesaja Hizkia
Vormgeving binnenwerk: Holland Graphics, Amsterdam
Redactie: Josien Moerman, Vertaal- en redactie bureau Am Rhein, Zwolle
Drukker: Balmedia, Schiedam

1^e druk, december 2020

©2020 Van Duuren Management en MT/Sprout

Niets uit dit boek mag worden overgenomen zonder voorafgaande schriftelijke toestemming van de uitgever. Wil je iets overnemen, neem dan contact op met de uitgeverij: info@vanduurenmanagement.nl.

Ondanks alle zorg die we aan dit boek besteed hebben, kan noch de auteur, noch de redactie, noch de uitgever enige aansprakelijkheid aanvaarden voor de schade die het gevolg is van eventuele fouten in deze uitgave.



**ALS WE NIET
OPPASSEN
MISSEN WE
DE REVOLUTIE
IN ARTIFICIAL
INTELLIGENCE**

Constantijn van Oranje



INHOUDS OPGAVE

VOORWOORD Hans de Jong Philips	7
INLEIDING Wat je gaat leren	12
01 WAT IS AI? Het slimste meisje van de klas	19
GROEICASE SEEDLINK	42
02 HET IDEALE AI-PROJECT Droom groots, begin klein	51
GROEICASE PICNIC	70
03 DATA Zo delf je de ultieme brandstof	79
GROEICASE SALESFORCE	101
04 DE KEERZIJDE VAN ALGORITMES Ethische vraagstukken die je niet kunt negeren	109
GROEICASE GEMEENTE AMSTERDAM	131
05 WELKE BANEN GAAT AI VERVANGEN? Een onvermoeibaar arbeidsleger	139
GROEICASE SLIMMER AI	154
06 DE AI VAN OVERMORGEN Aan de kant, homo sapiens	163
GROEICASE BLUETICK	179
07 IS EEN ALGORITME OOK CREATIEF? De kracht van nieuwsgierigheid	187
GROEICASE PARKBEE	198
08 AI IN BRANCHES EN SECTOREN Nog een paar praktische voorbeelden om het af te leren	205
GROEICASE KEPLER VISION	226
09 AI-ECOSYSTEEM Nederland zet de schouders eronder	233
NAWOORD	246
MUST-READS Nog meer kennis en inspiratie	248
MUST-WATCH E-learnings voor verdiepende kennis	254
MUST-USE Handige AI-tools voor marketing en sales	257
REGISTER	260

INLEIDING

Wat je gaat leren

Als klein jochie speelde ik op de MS-DOS-computer van mijn ouders voor het eerst het spel *Prince of Persia*. Floppydisk erin, een paar toetscombinaties en op het monochrome scherm kwam de virtuele wereld tot leven. Doel van het spel was om in zestig minuten de prinses te redden van een gruwelijke dood. Onderweg waren er talloze hindernissen: metalen grijpers, dodelijke drankjes, maar het meest fascinerende vond ik toch wel de ‘bewakers’ die verslagen moesten worden door tegen ze te zwaardvechten. Mijn kinderbrein kon maar moeilijk bevatten dat deze – toen nog – hyperrealistische tegenstanders niet door mensen werden bestuurd, maar door de computer. Later zou ik begrijpen dat dit de eerste stappen waren op het gebied van wat ze in computergames ‘AI’ noemen: *artificial intelligence*.

Inmiddels heeft artificial intelligence of – in goed Nederlands – kunstmatige intelligentie zich uit de wereld van de computerspellen losgeweekt en is zich aan het nestelen in alle facetten van ons digitale én analoge leven. Van de gepersonaliseerde aanbevelingen op Netflix en bol.com, geautomatiseerde beeldherkenning bij de politie,

medicijnontwikkeling en Apple's Siri tot aan de 'autopilot'-functie in een Tesla.

Niet voor niets noemt Sundar Pichai, de baas van Google en moederbedrijf Alphabet, de ontwikkelingen rond AI belangrijker voor de mensheid dan de ontdekking van elektriciteit en vuur: 'Het is de belangrijkste technologie waar ooit aan gewerkt is.' Volgens onderzoekers van PwC zal de toepassing van AI een wereldwijde economische groei van 15,3 biljoen dollar (twaalf nullen) opleveren. Belangrijke kanttekening: PwC verwacht dat het gros van die groei ten goede zal komen aan China en de VS, landen waar bedrijven en overheden fors investeren in AI-ontwikkeling. Willen we niet achterblijven en slechts de kruimels oppikken, dan zullen we ook in Nederland snel uit de AI-startblokken moeten komen.

Gelukkig is kunstmatige intelligentie niet meer iets wat alleen door wetenschappers valt te doorgronden. Bovendien hoeft je geen Amazon, Facebook, Google of Apple te heten om je bedrijfsvoering te verbeteren met *machine learning*. Zoals je verderop in dit boek zult lezen, zijn er veel handige tools beschikbaar die je het zware rekenwerk uit handen nemen. Zelfs een kleine organisatie met een paar medewerkers kan op deze manier AI gebruiken om te concurreren met de ontelbare handen en breinen achter een *corporate*.

Geen zorgen, dit boek is niet geschreven door *geeks* en techneuten. Voordat we straks een sprong in het diepe nemen, zullen we beginnen bij de basis. Wat is kunstmatige intelligentie? Hoe weet je welk product of proces je met AI kunt automatiseren of versnellen? En waar vind je de brandstof voor AI: data? We schrijven over de uitdagingen die AI met zich meebrengt op het gebied van ethiek, veiligheid, het vinden van talent en de impact op banen, bedrijven en de maatschappij.

Veel mensen krijgen bij het horen van artificial intelligence spontaan de beelden van *The Terminator* of *The Matrix* op hun netvlies. Gelukkig zijn we daar nog lang niet. Met de techniek die nu beschikbaar is, zijn algoritmes geschikt om grote hoeveelheden data snel te verwerken, maar altijd met een beperkte opdracht. Toch zijn er wel degelijk grote denkers die waarschuwen voor de mogelijke gevolgen van AI op de lange termijn: Bill Gates, Stephen Hawking en Elon Musk. Verderop in dit boek wordt hun argumentatie uiteengezet.

Voor dit boek heb ik met veel experts, ondernemers, leiders en denkers gesproken en veel zelfstudie gedaan rond het vraagstuk: hoe kunnen Nederlandse bedrijven meesurfen op die aankomende monstergolf genaamd AI? Mijn doel is te laten zien dat AI wel technisch klinkt, maar in de praktijk betrekkelijk eenvoudig te implementeren is.

Bovenal wil ik je laten zien hoe andere bedrijven AI nu al in praktijk brengen. Van veelbelovende startups, snelgroeiende scaleups tot grote organisaties. Denk aan Bluetick, dat een slimme zoekmachine voor advocaten bouwt, online supermarkt Picnic, die AI door de hele organisatie heeft gevlochten. Ook NS en zelfs de gemeente Amsterdam optimaliseren hun bedrijfsvoering en verbeteren de klantervaring, allemaal dankzij die wonderlijke technologie.

Omdat dit geen vuistdikke encyclopedie moest worden, komen diverse onderwerpen gesimplificeerd aan bod. Goed ingelezen AI-puristen zal dit een doorn in het oog zijn, maar ik hoop dat jij het ervaart als een wijnproeverij aan ideeën: springt één reserva er voor jou uit, dan is het een goed startpunt om je er verder in te verdiepen. Dat is nodig, want dit is hét moment om met AI te starten. Ondernemen met algoritmes heeft een sneeuwbal effect: zelflerende systemen worden nu eenmaal steeds slimmer naarmate je ze voedt met meer data. Hoe eerder je begint, hoe

groter je voorsprong op termijn. Het omgekeerde geldt ook: hoe langer je wacht ten opzichte van een concurrent die wel investeert in machine learning, hoe sneller je achterstand zich opbouwt.

AI heeft voor elk bedrijf – van groot tot klein – het potentieel om de onderneming te versterken met de kennis van duizenden knappe koppen. Die zitten niet in de kantine, maar draaien ergens in de cloud om jouw bedrijf superslim te maken. Ik hoop dat je de zoektocht naar de toepassingen van AI net zo fascinerend vindt als ik toen ik dit boek schreef. Ik start nu Prince of Persia nog eenmaal op, om die verdraaide AI te laten zien wie – voorlopig nog – de baas is: de mens.

Remy Gieling

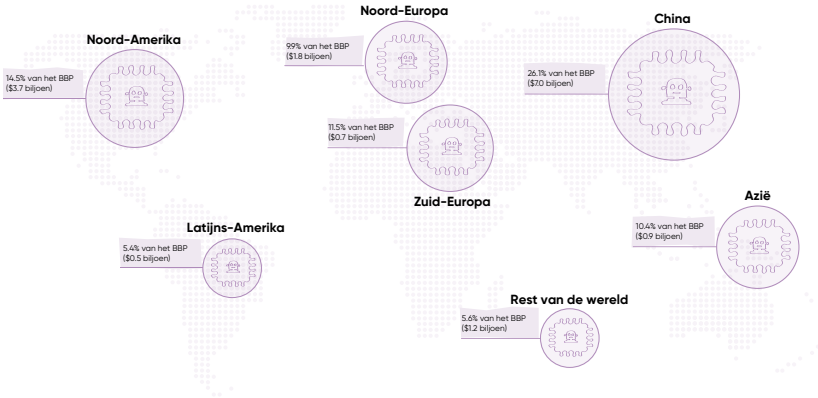
Hoofdredacteur MT/Sprout

Founder ai.nl.

P.S. Wil je wekelijks op de hoogte blijven van het belangrijkste nieuws rondom AI voor Nederlandse bedrijven? Meld je dan vooral aan voor mijn gratis nieuwsbrief op www.ai.nl.

Ontdek de groeikansen van AI

AI levert 15,3 biljoen (15 duizend miljard) dollar aan waarde aan de wereldwijde economie in 2030, maar welke regio het meest profiteert van AI?



Bron: PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution (<https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>)



**SOFTWARE
IS EATING
THE WORLD,
BUT AI IS
GOING TO EAT
SOFTWARE**

Jensen Huang, ceo Nvidia



01

WAT IS AI?

Het slimste meisje van de klas

De eigenaar van de kapsalon moet vreemd hebben opgekeken. Nietsvermoedend neemt ze de telefoon op en krijgt de assistent van een klant aan de lijn. Of die een afspraak kan krijgen tussen 10 en 12 uur? Ja hoor. Welke behandeling de klant wil? Wassen en knippen. Welke voornaam kan ze noteren? Lisa. De afspraak is gemaakt en de kapster wenst de assistent een prettige dag. Wat ze niet wist, is dat ze enkele dagen later wereldnieuws zou worden. Google-CEO Sundar Pichai presenteerde een tijdje later Google Duplex: een computerprogramma dat via kunstmatige intelligentie een menselijke stem levensecht imiteert en reserveringen voor je maakt bij restaurants, voorstellingen en... kapsalons. Het algoritme vult zelfs stiltes op met een bevestigend 'hmm-hmm', waardoor het voor degene aan de andere kant van de lijn amper te onderscheiden is van een gesprekspartner van vlees en bloed. Het systeem kan klanten ontzorgen, zegt Pichai: 60 procent van de kleine bedrijven in de VS gebruikt geen online boekingssysteem. Google Duplex slaagt met gemak voor de befaamde Turingtest: hoe weet je nog met zekerheid of je communiceert met een mens of een robot?

Een sterk staaltje AI. Maar wie tien experts vraagt wat kunstmatige intelligentie oftewel artificial intelligence in essentie is, krijgt tien verschillende antwoorden. Het is lastig om een eenduidige, allesomvattende definitie te geven, maar we vatten het begrip als volgt samen: artificial intelligence bestaat uit machines die intelligent gedrag vertonen. Dat kan bestaan uit software die razendsnel personen op een foto herkent, software die uit een enorme dataset accurate voorspellingen doet over je omzetcijfers voor het komende kwartaal, software die in een chatgesprek op menselijke wijze een conversatie voert, Google Home dat vliegensvlug antwoord geeft op de vraag: 'Wat wordt het weer vandaag?' tot aan robothond Spot, die zonder enige moeite de trap op en af kan lopen.

Intelligentie

Als het aan filosoof Bas Haring ligt, begint de zoektocht naar de betekenis van kunstmatige intelligentie bij de vraag: wat is intelligentie? 'Voor mij is dat het vermogen om te redeneren – op een correcte manier, snel, logische conclusies trekken. Bij intelligentie denk ik aan het slimste meisje van de klas, dat zonder probleem overal een antwoord op heeft. Kunstmatige intelligentie is dus het vermogen van een niet-biologisch systeem om precies dat te doen.' Haring, inmiddels bekend van boeken, tv-programma's en lezingen, was een van de eerste Nederlanders die afstudeerden op het vak kunstmatige intelligentie. 'De studie heette Cognitieve Kunstmatige Intelligentie en viel onder de faculteit Filosofie. Ik was de derde die de opleiding afrondde en was dus op een haar na Nederlands eerste *kunstmatige-intelligenticus*. In die tijd had je twee paradigma's: ten eerste de 'expertsystemen' waarbij je kennis van de mens in een computersysteem goot. Bijvoorbeeld hoe een arts tot een bepaalde diagnose kwam. De computer moest die redentatie volgen en

zou vervolgens ook tot dezelfde conclusie moeten komen. Het andere paradigma vormden de lerende systemen, tegenwoordig neurale netwerken, al was er voor die tijd eigenlijk te weinig data en rekenkracht om ze in praktijk te brengen. Ik hield mij bezig met de vraag: hoe kunnen computers leren? Toen ik op de opleiding zat, had de schaakgrootmeester Kasparov nog niet verloren van Deep Blue. Er zaten toen oude rotten in het vak die dachten dat een computer nooit zou kunnen winnen van een menselijke speler. Dat zou nooit ofte nimmer gebeuren. Nog voor mijn afstuderen was het zo ver. Zo was het eerst met schaken, nu met zelfrijdende auto's en straks met iets anders. Computers zijn hartstikke intelligent, alleen komt een machine op andere manieren tot antwoorden dan hoe onze hersenen informatie aan elkaar knopen.'

Het driedeurenprobleem

Een goed voorbeeld van redeneervermogen is volgens Bas Haring te vinden in het driedeurenprobleem, ook wel bekend als het *Monty Hall problem*. Dat luidt als volgt: 'Stel dat je deelneemt aan een spelprogramma en je mag kiezen uit drie deuren: achter een van de deuren staat een auto, achter de andere twee staan geiten. Je kiest een willekeurige deur. De presentator, die weet wat er achter de deuren staat, opent vervolgens een andere deur en laat zien dat daarachter een geit staat. Hij geeft je vervolgens de mogelijkheid om ten aanzien van de twee overgebleven deuren van keuze te veranderen. Is het wisselen van deur in jouw voordeel? Veel mensen zouden hier ontkennend op antwoorden. Je zou immers kunnen denken dat de kans 1 op 3 was dat je de juiste deur kiest en dat die nu 1 op 2 is. Toch is dat niet het geval. Blijf je bij je keuze, dan is de kans 1 op 3, wissel je dan verhoog je die naar 2 op 3. Moeilijk te geloven? Googel het voor de grap eens.

Het driedeurenprobleem is een voorbeeld dat mooi laat zien hoe moeilijk het voor – de meeste – menselijke breinen is om abstracte statistische informatie te verwerken. Computers kunnen dat als geen ander.

Data en rekenkracht

Artificial intelligence is dus geen nieuwe technologie. Wel heeft het vakgebied de afgelopen jaren bovengemiddeld veel interesse gekregen van bedrijven, overheden en wetenschap. Dat komt doordat kunstmatige intelligentie drie ingrediënten nodig heeft om succesvol te zijn: knappe koppen die algoritmes en modellen kunnen programmeren, een flinke hoeveelheid rekenkracht en een nog veel grotere hoeveelheid data. De afgelopen zeventig jaar zijn er ruim voldoende briljante wiskundigen en programmeurs geweest, maar pas sinds een paar jaar zijn computers snel genoeg om de complexe berekeningen te doen die nodig zijn om algoritmes in de praktijk te laten werken. Nu ons leven zich meer en meer online afspeelt, is er eindelijk voldoende data beschikbaar om computersystemen te laten leren. Online is er immers data in overvloed: elke *like* die iemand geeft, de tijd die hij op een webpagina doorbrengt, tot hoe ver hij scrollt en de aankopen die hij doet: slimme software kan dergelijke datapunten aan elkaar knopen om te voorspellen of en wanneer een volgende klant op jouw koopknop drukt.

Ook het 'fysieke' leven is steeds vaker gekoppeld aan de digitale wereld: de deurbel die je telefoon een seintje geeft als de pakketbezorger voor de deur staat, de Philips Hue-lampen die aangaan als je Netflix opstart, de thermostaat die automatisch de verwarming opstookt als je vanaf je werk naar huis rijdt, maar ook het schermpje dat op Schiphol vraagt of je toiletbezoek naar tevredenheid was. Geschat wordt dat elke dag 2,5

quintiljoen bytes (een 1 met 30 nullen erachter) aan nieuwe data wordt opgeslagen. In China weten ze inmiddels niet anders: een Chinees haalt zijn fiets van het slot met zijn smartphone, zet zijn wasmachine aan met behulp van een qr-code en betaalt in de winkel met gezichtsherkenning.

Algoritmes

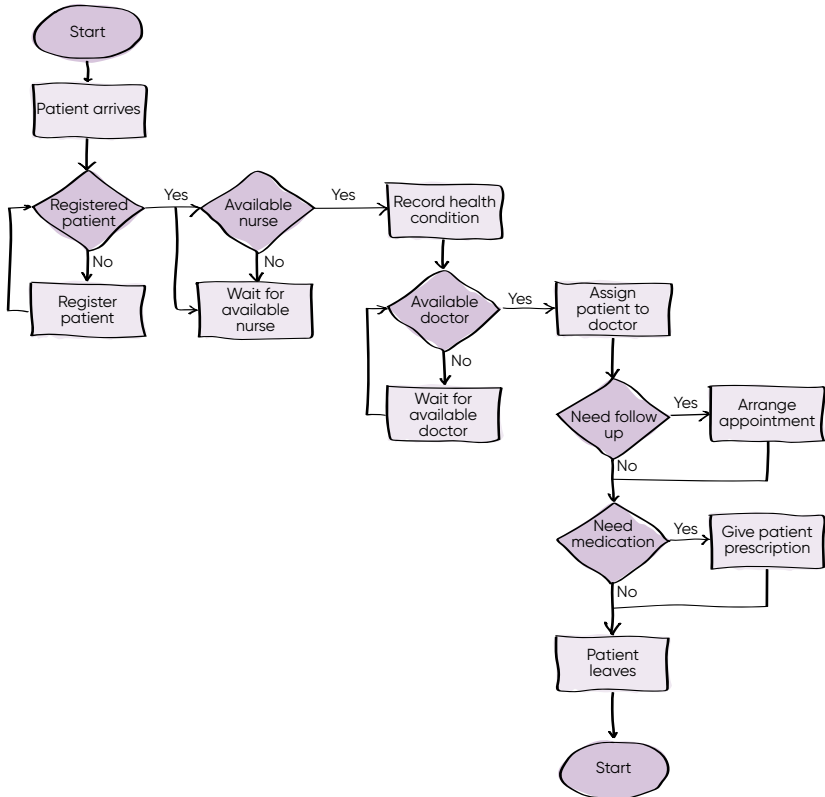
Aan de basis van kunstmatige intelligentie staan algoritmes. Algoritmes zijn instructies voor een computer om een bepaald probleem op te lossen. Zoals je een recept in een kookboek volgt om een mooie Indiase curry te maken, zou je ook in een stappenplan kunnen opschrijven hoe bepaalde bedrijfsprocessen geregeld zijn. Een voorbeeld: hoe bepaal je of een medewerker van de salesafdeling presteert? Belt hij genoeg? Maakt hij genoeg afspraken in de maand? Communiceert hij voldoende? Blijft er onderaan de streep meer over dan eruit gaat? Gooi er een weging tegenaan en je kunt er een competentiescore uit berekenen. Je kunt dus zelf de kern van algoritmes schrijven. Dat begint bij heel secuur nadenken hoe beslissingen tot stand komen en ze opschrijven in een *flowchart*.

Omdat computers niet zo goed zijn met taal, maar erg goed in razendsnelle berekeningen, werken algoritmes dus niet via een geschreven stappenplan, maar met wiskundige formules. Zijn de stappen eenmaal uitgeschreven, dan kan een wiskundig genie, datadeskundige of programmeur ermee aan de slag om de doorvertaling te maken naar computercode.

Sommige formules zijn betrekkelijk eenvoudig, andere – voor complexe toepassingen – zijn enkel te doorgronden door mensen met een doctoraat in geavanceerde statistiek. De formules zijn in de loop der jaren weinig veranderd, maar de rekenkracht die ervoor nodig is om ze op te lossen, was voorheen enkel beschikbaar in onderzoekslaboratoria

Ontdek de groeikansen van AI

van universiteiten of achter de hoge muren van grote techbedrijven. Met de komst van de *cloud* heeft iedereen ineens toegang tot de krachtige CPU's (processors) en GPU's (grafische kaarten) die nodig zijn voor het ontwikkelen van toepassingen via algoritmes.



Flowchart van een theoretisch algoritme van een ziekenhuisbezoek

AI om ons heen

Hoewel je het niet altijd beseft, is kunstmatige intelligentie niet meer weg te denken in de wereld om ons heen. Immens complexe rekenmodellen

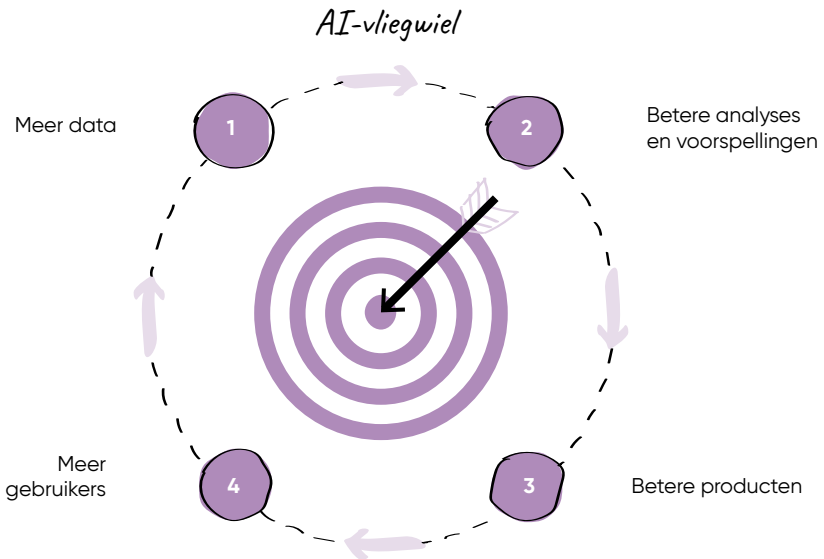
zorgen er in een oogwenk voor dat je uit miljarden webpagina's het beste resultaat bovenaan in Google te zien krijgt (net onder degene die het meest betaalt). Een arts laat jouw CT-scan misschien wel eerst door een algoritme checken, en niet de bankmedewerker beoordeelt jouw aanvraag voor een lening, maar dit wordt op de achtergrond door een computerprogramma gedaan.

Startups ontwikkelen slimme tools die volledig gebouwd zijn rondom AI – denk aan de juridische zoekmachine Bluetick of het Rotterdamse Ciphix, dat processen voor grote klanten als KLM en Unilever automatiseert. Corporates vormen in een rap tempo AI-afdelingen om bestaande producten en diensten te optimaliseren. Philips, KPN, ING en NS zoeken actief samenwerking met de wetenschap om niet alleen bestaande technieken toe te passen, maar ook onderzoek naar grote uitdagingen mogelijk te maken. Daarnaast gebruikt de overheid de technologie meer en meer. Zo kan een camera-auto van de gemeente Amsterdam zwerfafval herkennen en de locatie doorgeven aan de reinigingsdienst. Maar de techgiganten, die zijn pas echt spekkoper. Met de enorme hoeveelheid informatie over onze interesses (Facebook), ons koopgedrag (Alibaba), onze vragen (Google), contacten (Tencent) en het zakenleven (Microsoft) worden ze almaar slimmer.

Ook Amazon.com is heer en meester over het algoritme. Het bedrijf werkt volgens het principe van de zogeheten *AI flywheel*, zo valt te lezen in het boek *Bezonomics* van Brian Dumaine. Geen beslissing wordt genomen zonder dat algoritmes door de data-schatkamer zijn geploeterd om advies uit te brengen. Sterker nog, algoritmes zijn bijna geheel verantwoordelijk voor de inkoopafdeling geworden: ze bepalen welk product, wanneer, waar en tegen welke prijs ingekocht moet worden om voldoende voorraad voor de webshop te behouden en plaatsen in veel gevallen zelfstandig de orders. Ook op de werkvloer regeert het algoritme: halverwege 2020

Ontdek de groeikansen van AI

had het bedrijf meer dan een miljoen medewerkers. Voor 'eenvoudige functies' bleek het inefficiënt dat een mens sollicitanten behandelt. Wie bijvoorbeeld orderpicker in een distributiecentrum van Amazon wil worden, wordt online aangenomen (of afgewezen) door een computer zonder dat er een mens aan te pas komt. Ben je eenmaal werkzaam in het magazijn, dan houden slimme sensoren je productiviteit secuur in de gaten. Werk je te langzaam? Dan laat een computer weten dat je bent ontslagen.



Containerbegrip

Artificial intelligence is dus een containerbegrip voor alle intelligente niet-biologische systemen, van simpele algoritmes tot Deep Thought, dat in de *The Hitchhiker's Guide to the Galaxy* het antwoord geeft op de ultieme vraag over het leven, het universum en alles (*spoiler alert*: dat is het getal 42).

Veel mensen vragen zich ook af wat AI níét is. De meeste voorbeelden hierboven illustreren wat we *narrow AI* (of ‘zwakke AI’) noemen: computerprogramma’s die goed zijn in het oplossen van één betrekkelijk simpele taak. Zelfs een zelfrijdende auto besturen valt hieronder: die kan je hopelijk veilig van A naar B brengen, maar is niet in staat om je huis te stofzuigen, een discussie over de zin van het leven te voeren of prijswinnende literatuur te schrijven.

In populaire sciencefiction heeft AI vaak bovenmenselijke capaciteiten; denk aan Deep Thought, HAL 9000 in *2001: A Space Odyssey* en de T-800 uit *The Terminator*. Die zitten de mensheid dikwijls flink in de weg. Dat dit overigens niet altijd het geval is, bewees de schattige opruimrobot WALL•E in de gelijknamige Pixar-film. Kunstmatige intelligentie die het menselijk brein evenaart, wordt *general AI* (of ‘sterke AI’) genoemd. Sommige wetenschappers denken dat computers al rond 2040 een dergelijke intelligentie hebben ontwikkeld. Voorlopig komen computers zelfs met de meest geavanceerde technieken niet in de buurt van dit toekomstbeeld. In dit boek focussen we ons liever op wat de techniek al wel kan.

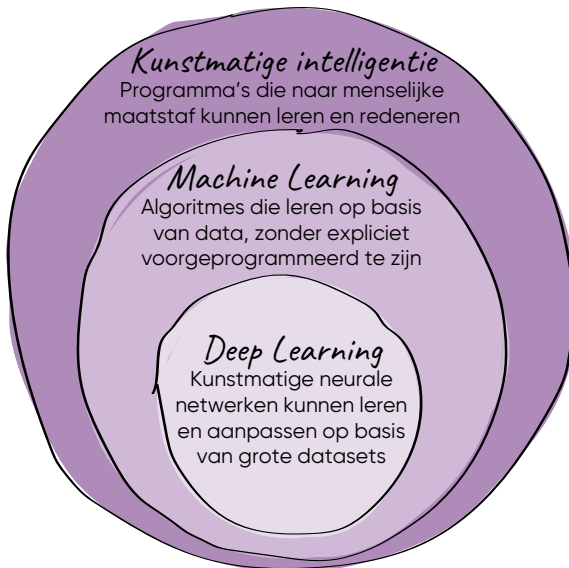
Machine learning

Een veelvoorkomend subgebied van kunstmatige intelligentie is *machine learning*. Daarbij volgt de computer niet langer een set voorgeprogrammeerde regels, maar leert hij door middel van ‘ervaring’. Wie bijvoorbeeld een programma wilde ontwerpen om spam te filteren, moest vroeger een uitgebreid handboek schrijven: is de mail afkomstig van een bonafide afzender? (dan mag hij door), staat erin dat Jort Kelder bitcoins wil verkopen? (waarschijnlijk spam), mailt je schoonmoeder? (twijfelgeval). Spammers werden echter steeds slimmer en wisten keer op keer de vaste regelset van de spamfilters om de tuin te leiden. Bovendien

Ontdek de groeikansen van AI

is het bijna onmogelijk om alle kenmerken van een bonafide mail of spammail handmatig te programmeren.

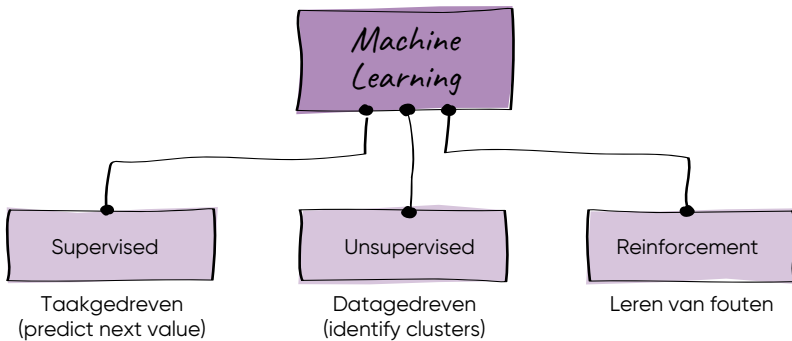
Machine learning gaat andersom te werk: je voedt het systeem met miljoenen e-mails (de 'trainingsdata'), waarvan je – handmatig – aangeeft welke wel spam zijn en welke niet. De software kan vervolgens via wiskundige berekeningen (het algoritme) zelf patronen herkennen in de mails van dubieuze afzenders. Zo kan met grote zekerheid worden voorspeld of de volgende mail die binnenkomt echt van tante Bep is die je een goede verjaardag wenst, of van een Russische hacker die jouw bedrijfsgeheimen wil achterhalen. Ook als je wilt voorspellen wanneer het beste moment is om een marketing-e-mail te verzenden, zul je heel veel voorbeelden van soortgelijke e-mails moeten verzamelen, met daarbij het aantal keer dat de mail werd geopend.



Bron: Data Catchup

Met of zonder toezicht

Omdat je bij bovenstaande voorbeelden de computer expliciet moet vertellen waar hij naar op zoek is (spammails of data voor de *ideal open rate*), spreken we hierbij van *supervised learning*. Je bent actief betrokken bij het trainingsproces, omdat je op elk ingevoerd voorbeeld een label hebt geplakt. Bepaalde algoritmes zijn echter ook in staat om verbanden te achterhalen in grote datasets, waar geen labels aan hangen. We spreken in dat geval van *unsupervised learning*. Een voorbeeld: je wilt weten welke patronen er zitten in de aankopen van je klanten, zodat je klantprofielen kunt opstellen voor gerichte advertenties. De laatste variant, *reinforcement learning*, is pas recent in zwang gekomen. Hierbij draait de computer zelfstandig miljoenen simulaties, waarin hij door middel van straffen en beloning (via een score, je kunt software immers moeilijk een tik geven) steeds probeert op een betere uitkomst te komen. Dat kan een robot zijn die moet leren een trap op te lopen zonder te vallen (een trede omhoog is een punt erbij, vallen is minpunten), of een computer die leert om het spel Pong, Mario of Flappy Bird te spelen. Reinforcement learning lijkt op de digitale doorvertaling van de evolutietheorie: alleen de ‘fitste’ algoritmes, die een hoge score behalen, worden gebruikt in de simulatie.



Taken

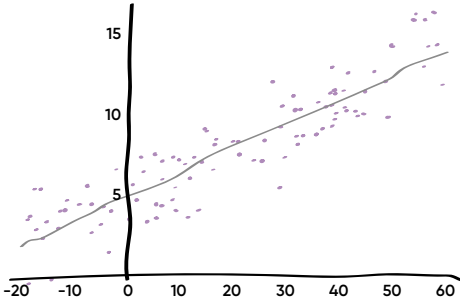
Plat gezegd zijn algoritmes bedreven in het verrichten van een drietal taken:

1. objecten classificeren
2. voorspellen
3. clusteren van groepen data.

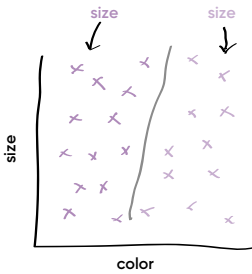
Hiervoor hebben we het gehad over het classificeren van objecten; dat kan gaan om het herkennen van spammail, maar ook om de vraag of een medewerker salarisverhoging moet krijgen of dat een leningaanvraag moet worden goedgekeurd of afgewezen. Naast classificeren kunnen algoritmes op basis van historische data ook goed voorspellingen doen over de toekomst. Stel, een autodealer houdt tien jaar lang bij welk type klant welk type auto wanneer koopt, hoe vaak die daarvoor is langs geweest, hoelang hij in de showroom loopt en welk weer het was op de dag van aankoop. Dan kan de dealer een algoritme laten voorspellen hoe groot de kans is dat de eerstvolgende persoon die binnen komt lopen, vertrekt met een nieuwe auto. Tot slot zijn er algoritmes die in grote datasets zelfstandig overeenkomstigheden kunnen herkennen. Deze informatie wordt geclusterd en inzichtelijk gemaakt in een grafiek. Zo kan een algoritme alle huizenkopers onderverdelen in 'bakjes' van leeftijd, sekse, opleiding, salaris, postcodegebied en gezinsgrootte, maar ook lengte, oogkleur, muzieksmaak, merkvoorkeuren en zoekgedrag, om hier onderlinge dwarsverbanden in te vinden.

Machine learning is goed in drie taken

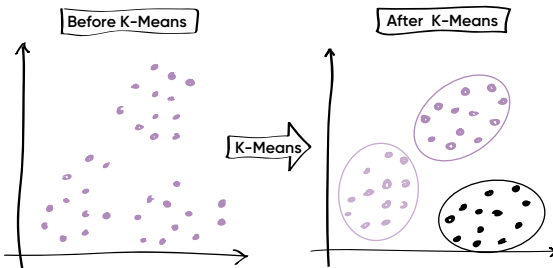
01 Voorspellen



02 Classificeren



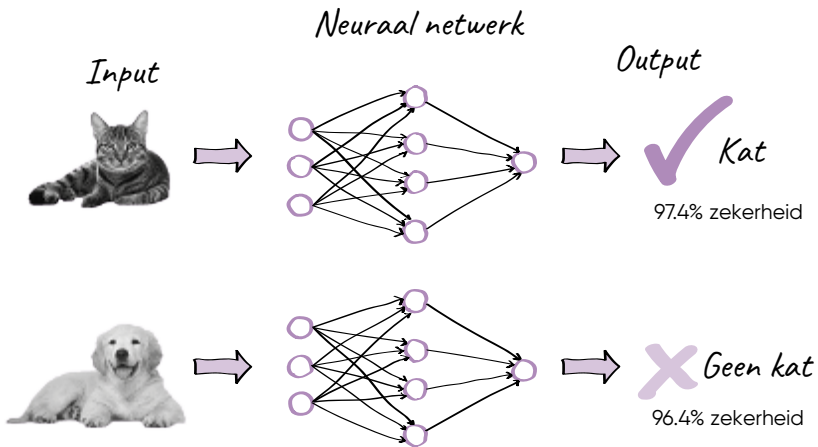
03 Patroonherkenning



Neurale netwerken

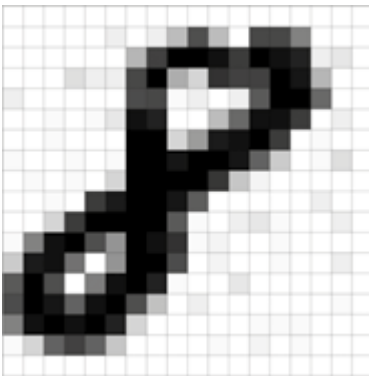
Waar veel methodes voor machine learning redelijk eenvoudig te begrijpen zijn, is er één geavanceerde techniek die de afgelopen jaren bovengemiddeld veel interesse heeft gekregen: neurale netwerken. Geïnspireerd door het menselijk brein hebben wetenschappers de *perceptron* gemaakt: een digitale variant van het neuron, de complexe zenuwcel. Al in 1958 is de perceptron bedacht, maar de toepassing ervan is pas kortgeleden van de grond gekomen. De meeste neurale netwerken bestaan uit meerdere lagen perceptrons. Elke perceptron bevat een eigen rekenkundige formule, past de formule toe op ingevoerde data en geeft de uitkomst door aan de volgende laag met perceptrons, die weer een nieuwe formule op de gegevens loslaten, en zo verder.

Hoe meer lagen, hoe complexer het model. Een neuraal netwerk begint met een laag van één of meer invoerwaarden, wat kan variëren van een getal – zeg de temperatuur – tot de grijswaarde van een pixel in een foto. Zijn de neurale netwerken geavanceerd en bestaan ze uit veel lagen, dan spreken we niet langer alleen van machine learning, maar ook van *deep learning*.



Bron: ARM

Neurale netwerken zijn bij uitstek geschikt voor het doorgronden van *ongestructureerde data*. Ongestructureerde data kun je niet zoals tekst en getallen makkelijk doorzoeken. Voorbeelden zijn geluidsfragmenten, pdf's, video's en foto's. Een neurale netwerk kan bijvoorbeeld voorspellen wat er op een foto te zien is. Het knipt daarvoor de foto's op in (groepen van) individuele pixels. Die pixels (groepen) zijn donker of licht en hebben daarmee een bepaalde waarde van 0 (puur wit) tot 255 (puur zwart). Het neurale netwerk gaat op zoek naar herkenbare patronen in die pixels. Daarvoor moet het wel gevoed zijn met heel veel trainingsfoto's waarbij vermeld is wat er op het beeld te zien is. Een neurale netwerk heeft immers geen begrip van het concept 'kat', of zoals hieronder het getekende cijfer '8', maar als het voldoende voorbeelden heeft gezien, kan het met een bepaalde zekerheid zeggen dat er op basis van de pixels hoogstwaarschijnlijk een 8 te zien is.



```

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 12 0 11 39 137 37 0 152 147 84 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 1 0 0 0 0 41 168 250 255 235 162 255 238 206 11 13 0 0 0 0 0 0 0 0 0 0
0 0 0 16 9 9 150 251 45 21 184 159 154 255 233 40 0 0 0 0 0 0 0 0 0 0 0 0
10 0 0 0 0 0 0 145 146 3 10 0 11 124 253 255 187 0 0 0 0 0 0 0 0 0 0 0 0
0 0 3 0 4 15 236 216 0 0 38 189 247 248 169 0 11 0 0 0 0 0 0 0 0 0 0 0
1 0 2 0 0 0 0 253 253 23 62 224 241 255 164 0 5 0 0 0 0 0 0 0 0 0 0 0
6 0 0 4 0 3 252 250 228 255 255 234 112 28 0 2 17 0 0 0 0 0 0 0 0 0 0 0
0 2 1 4 0 21 255 253 251 255 172 31 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 4 0 163 225 251 255 229 120 0 0 0 0 0 0 11 0 0 0 0 0 0 0 0 0 0 0
0 0 21 162 255 255 254 255 126 6 10 14 6 0 0 0 9 0 0 0 0 0 0 0 0 0 0
3 79 242 255 141 66 255 245 189 7 8 0 0 5 0 0 0 0 0 0 0 0 0 0 0 0 0
26 221 237 90 0 67 251 255 144 0 0 0 0 0 7 0 0 11 0 0 0 0 0 0 0 0 0
125 255 141 0 87 244 255 200 3 0 0 13 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0
145 248 228 116 235 255 141 34 0 11 0 1 0 0 0 1 3 0 0 0 0 0 0 0 0 0
85 237 253 246 255 210 21 1 0 1 0 0 6 2 4 0 0 0 0 0 0 0 0 0 0 0 0 0
6 23 112 157 114 32 0 0 0 0 2 0 0 0 7 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

```

Bron: Towards Data Science

Hoe een computer een handgeschreven getal herkent: de lichtintensiteit van de pixels.

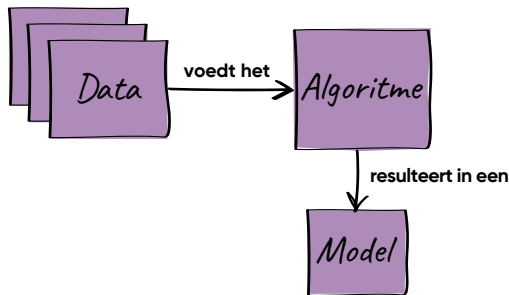
Een goede trainingsset aanleggen is een flinke klus. Hoe je dat doet en waar je bestaande trainingssets vandaan haalt, lees je in hoofdstuk 3 over het vinden, verzamelen en toepassen van data.

Voor de meeste toepassingen is het niet nodig om je eigen deep learning software te ontwikkelen, maar gebruik je een opensource-framework als PyTorch of TensorFlow. Omdat de invoer van informatie door talloze virtuele neuronen gaat voordat er een uitkomst is, is de werking van een neuraal netwerk vaak lastig te achterhalen. Mocht je een systeem ontwikkelen dat besluit of iemand wel of geen lening krijgt toegekend, dan is deze techniek niet het meest geschikt: je wilt nog kunnen uitleggen waarom een aanvraag wordt af- of toegewezen.

Het model

Wanneer je een algoritme hebt getraind, krijg je uiteindelijk een model. Dit model zorgt ervoor dat de data die je invoert, op basis van de achterliggende berekeningen een uniforme uitkomst oplevert. Bevat de trainingsset alleen maar rode appels en groene peren, dan zal het model een groene appel hoogstwaarschijnlijk als peer herkennen. Sommige systemen kunnen wel leren van nieuwe data. Zeg je in dat geval tegen het systeem 'nee, dit is geen peer, maar een appel', dan zal het algoritme een nieuwe rekenmethode verzinnen om in de toekomst een betere voorspelling te doen. Deze nieuwe formule levert een nieuw model op. Wanneer een organisatie algoritmes inzet, zijn die dus niet per definitie zelflerend. In dat geval moet geregeld worden gecheckt of de modellen nog wel aansluiten op de actualiteit. De komst van het coronavirus is hier een mooi voorbeeld van: talloze algoritmes hielden rekening met bepaalde patronen die ze kenden uit het verleden: wegen zijn om acht uur 's ochtends overvol, mensen gaan in juli op vakantie, 's avonds zweten mensen in de sportschool en een gemiddeld gezin koopt per week één pak wc-papier. Je snapt het: alle voorspellingen konden in één keer bij het grofvuil. Daarnaast is het zaak dat je je data up-to-date houdt. Wie voor een brillenketen zoekt naar de ideale locatie voor een nieuwe vestiging,

mist kansen als hij naar demografische gegevens uit 2010 kijkt. Misschien is er ergens een verzorgingshuis gebouwd of zijn er ondertussen een tiental concurrenten bij gekomen.



Bouwpakket

Algoritmes en modellen hoef je niet altijd zelf te programmeren. Gelukkig maar, niet elk bedrijf heeft budget voor een klein leger datadeskundigen. Online zijn er voldoende digitale gereedschappen die het voor grote én kleinere organisaties mogelijk maken om de kracht van algoritmes te gebruiken, zonder de rekenkundige fundering erachter te hoeven begrijpen. Het zijn vaak de grote techbedrijven die belangrijke raamwerken voor het toepassen van kunstmatige intelligentie vrij beschikbaar hebben gesteld als opensourcesoftware. Google doneerde TensorFlow aan de gemeenschap, Facebook deed hetzelfde met PyTorch. Ondertussen regelde Microsoft zijn Azure-cloud in om het *crunchen* van data zo eenvoudig mogelijk te maken en deed Amazon hetzelfde met het immense serverpark van Amazon Web Services.

Deze en vele andere tools maken het inmiddels ook voor een organisatie met betrekkelijk weinig mensen mogelijk om de kracht van artificial intelligence te benutten. Zoals AI-expert Jim Stolze het omschrijft:

‘Vroeger maakte je websites door html-codes achter elkaar te tikken, tegenwoordig kan zelfs de grootste leek via Wix.com of WordPress een pagina in elkaar zetten. Datzelfde gebeurt langzaam maar zeker ook met kunstmatige intelligentie. Mensen met een beetje technische handigheid kunnen zo AI-modellen in elkaar zetten.’

In den beginne ...

Als geboorteplaats van de moderne kunstmatige intelligentie geldt de gerenommeerde universiteit Dartmouth in de Amerikaanse staat New Hampshire. Hier kwamen in 1956 een twintigtal wetenschappers samen om acht weken lang te praten over de potentie van artificial intelligence. Tot die tijd werd het theoretische werkveld vaak aangeduid als *thinking machine*.

Wie in Nederland alles weet over de Dartmouth Conferentie, zoals de historische conferentie inmiddels genoemd wordt, is professor Emile Aarts, ooit rector magnificus van Tilburg University en initiator van de Jheronimus Academy of Data Science (JADS) in Den Bosch, een universiteit die zich richt op het toepassen van artificial intelligence in de praktijk. Aarts verzorgt nog steeds regelmatig colleges en heeft de ontwikkelingen van de techniek van dichtbij kunnen volgen: ‘Ik was te jong om het bewust mee te maken, maar wat was ik graag bij de Dartmouth Conferentie geweest! Daar is het fundament gelegd voor een technologie die van invloed is op zoveel facetten van ons huidige leven en een ongekende rol speelt in de toekomst van de mensheid. Mijn fascinatie voor AI is ontstaan door Alan Turing, door het ongelooflijk knappe brein van de man. En toch ook wel door de tragiek van zijn bestaan – Turing is nadat hij de sleutel voor de Enigma-codeermachine van de nazi’s had gekraakt, vervolgd wegens

homoseksualiteit. Dat heeft er naar alle waarschijnlijkheid toe geleid dat hij een einde aan zijn leven maakte.

Zijn Turingtest, die de vraag stelt of iemand weet of hij communiceert met een mens of machine, is nog steeds een belangrijke leidraad in de ontwikkeling van AI. Na de conferentie van Dartmouth zijn er diverse belangrijke momenten geweest in de geschiedenis van kunstmatige intelligentie, met twee AI-zomers waarin er veel optimisme was over de mogelijkheden en er veel geld beschikbaar was bij universiteiten, bedrijven en overheden. Tussendoor zijn er twee winters geweest, waarin vrijwel geen onderzoek meer plaatsvond en de investeringen wegvielen. Ik krijg geregeld de vraag of er een nieuwe winter aankomt, maar die kans is nihil. Destijds was er onvoldoende rekenkracht, data en kennis beschikbaar om kunstmatige intelligentie succesvol toe te passen. Inmiddels is dat er allemaal in overvloed en is AI niet meer weg te denken. *AI is here to stay* en zal invloed hebben op elk denkbaar aspect van ons bestaan. Daar is niets engs aan, het zal ons leven alleen maar makkelijker en leuker maken.'

De keuze voor de term *artificial intelligence* is voor AI-expert Jim Stolze nog steeds tegen het zere been: 'Wij, de nerds, zijn schuldig aan de angst die bij veel mensen leeft voor het algoritme. De naam is daar de grote boosdoener achter. Die zegt eigenlijk: "Mensen, jullie hebben wel intelligentie, maar kijk eens: we hebben nu iets ontwikkeld dat hetzelfde kan – en misschien nog wel beter." Ik vind dat AI beter zou kunnen staan voor *augmented intelligence*. Het geeft ons mensen gereedschap om betere beslissingen te nemen. Zo hoeft een ambtenaar op een ministerie helemaal niet te vrezen voor zijn baan, maar kan hij zich dankzij de technologie focussen op het oplossen van complexe vraagstukken in plaats van het verwerken van informatie in verschillende computersystemen.'

Met alleen een stel tools ben je er echter niet. Of een AI-project slaagt, hangt af van een hoop factoren. Hoe je het best kunt starten, waar het laaghangend fruit hangt en waar de valkuilen zich bevinden, lees je in het volgende hoofdstuk.

Tijdslijn

- 1642** De Franse uitvinder Blaise Pascal bouwt de eerste mechanische rekenmachine.
- 1763** Thomas Bayes ontwikkelt een raamwerk om te redeneren over de waarschijnlijkheid van gebeurtenissen, de basis voor algoritmes.
- 1837** Wiskundige Ada Lovelace publiceert een ontwerp voor een programmeerbare machine.
- 1921** De term 'robot' wordt voor het eerst gebruikt door de Tsjechische schrijver Karel Čapek.
- 1943** Walter McCulloch bedenkt het theoretische concept van neurale netwerken.
- 1950** Isaac Asimov publiceert zijn drie wetten van robotica in zijn boek *I, Robot*.
- 1950** De Britse wiskundige Alan Turing publiceert de Turingtest als methode om de intelligentie van een machine te meten.
- 1955** De term 'artificial intelligence' wordt voor het eerst gebruikt door de Amerikaanse professor John McCarthy tijdens de Dartmouth Conferentie.
- 1957** Uitvinding van de perceptron, de digitale versie van een neuron en basis voor neurale netwerken.
- 1958** De AI-programmeertaal LISP komt op de markt.
- 1958** Het eerste schaakalgoritme is Bernsteins Chess AI.
- 1959** Arthur Samuel bedenkt het begrip 'machine learning'.

- 1960** Het eerste neurale netwerk ontwikkeld, met slechts één perceptron-laag.
- 1965** Intel-oprichter Gordon Moore voorspelt dat het aantal transistors in een chip elke twee jaar zal verdubbelen: *Moore's Law*.
- 1965** Chatbot ELIZA is in staat om gesprekken te voeren.
- 1967** Het eerste algoritme verslaat een menselijke speler in een schaakwedstrijd.
- 1972** De eerste 'intelligente' robot: de WABOT-1.
- 1980** Edward Feigenbaum legt de basis voor expertsystemen – de voorloper van geavanceerde algoritmes.
- 1981** De interesse verschuift van 'general intelligence', die het menselijk brein evenaart, naar 'narrow AI': computersystemen die kleine, praktische problemen oplossen.
- 1984** Onderzoekers van Carnegie Mellon University bouwen een zelfrijdende auto.
- 1986** Het eerste geavanceerde neurale netwerk met meerdere lagen perceptrons.
- 1989** Yann LeCun (inmiddels Hoofd AI bij Facebook) ontwikkelt het eerste Convolutional Neural Network en legt de basis voor *computer vision*.
- 1990** Stemherkenning in opkomst.
- 1991** De Nederlander Guido van Rossum ontwikkelt de programmeertaal Python, inmiddels de standaard voor machine learning.
- 1997** IBM's Deep Blue verslaat schaakkampioen Garry Kasparov.
- 1998** Onderzoekers van MIT tonen KISmet, een robot die reageert op menselijke emotie.
- 1999** Sony brengt robothond AIBO op de markt.
- 2002** iRobot brengt de Roomba uit als eerste autonoom opererende stofzuigrobot voor particulieren.

- 2006** Deep Neural Networks doen hun intrede.
- 2009** Stanford-onderzoekers publiceren ImageNet, met 14 miljoen gelabelde afbeeldingen waarop algoritmes kunnen trainen.
- 2011** IBM's Watson verslaat de kampioen in het tv-spel *Jeopardy!*.
- 2011** Slimme speakers als Google Home, Amazon Alexa en Apple's Siri doen hun intrede.
- 2014** Generative Adversarial Networks (GAN's) doen hun intrede.
- 2014** Reinforcement learning raakt in zwang.
- 2014** De eerste chatbot, genaamd 'Eugene Goostman', doorstaat de Turingtest.
- 2015** Google stelt AI-framework TensorFlow ter beschikking.
- 2015** Elon Musk initieert OpenAI, een stichting die onderzoek naar *general artificial intelligence* mogelijk moet maken, zodat de kennis en beslissingen over superintelligentie niet enkel in handen van commerciële conglomeraten komen.
- 2015** Stephen Hawking en andere prominente AI-denkers pleiten in een open brief voor een verbod op autonome wapens.
- 2016** Facebook stelt AI-framework PyTorch algemeen beschikbaar.
- 2016** Microsofts chatbot Tay leert snel van sociale media en wordt racistisch scheldkanon voordat het bedrijf de software offline haalt.
- 2016** De Britse Google-dochter DeepMind verslaat met Alpha Go de Zuid-Koreaanse wereldkampioen in Go, waarbij meer zetten mogelijk zijn dan atomen in het bekende universum.
- 2017** Amazon introduceert SageMaker voor machine learning modellen op zijn AWS-plaform.
- 2017** Google DeepMind ontwikkelt Alpha Go Zero, een verbeterde versie die voorganger Alpha Go keer op keer verslaat.
- 2018** Een door een algoritme gegenereerd kunstwerk wordt voor 432.500 dollar verkocht bij veilinghuis Christie's.

- 2019** Google Duplex laat een computerstem telefonisch reserveringen maken voor Android-gebruikers.
- 2020** OpenAI publiceert tekstverwerker GPT-3, getraind op miljarden webpagina's om zelf teksten te genereren.