

## Wat is kunstmatige intelligentie

Voor veel mensen is 'kunstmatige intelligentie', of AI, een bekend begrip, maar tegelijkertijd weten de meesten niet wat het precies is, wat het inmiddels kan, laat staan hoe het werkt. Innovatie en adoptie van nieuwe technologie leunt echter voor een deel op begrip van die technologie. Daarom geven we hier een uitleg van wat kunstmatige intelligentie is en hoe het werkt, zonder in de technische details te treden.

### Een overzicht van het werkveld

**Kunstmatige Intelligentie (AI)** is het werkveld dat zich bezighoudt met het vermogen van computers om problemen op te lossen, zoals natuurlijke intelligentie dat zou doen. Het bestaat uit verschillende vakgebieden zoals Machine Learning en Robotica.

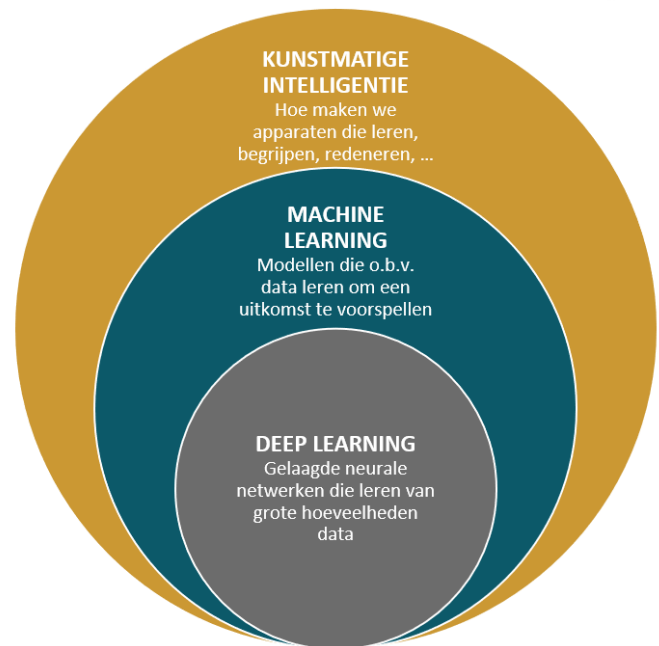
**Machine Learning (ML)** is een belangrijk vakgebied binnen AI. Machine Learning refereert naar digitale modellen die leren door 'ervaring' in de vorm van data. Hoe meer data je aanbiedt, hoe beter het model wordt in het voorspellen van de juiste uitkomst. De modellen zijn veelal gebaseerd op statistiek, denk bijvoorbeeld aan regressie.

**Deep Learning** is een heel interessant soort model binnen Machine Learning. Het verwijst naar kunstmatige neurale netwerken, die relatief veel weghebben van hoe het menselijk brein informatie verwerkt.

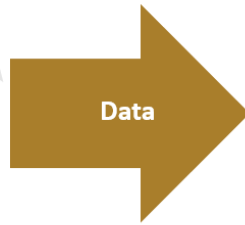
Een neuraal netwerk bestaat uit enorm veel verbindingen tussen knooppunten, zoals neuronen in de hersenen met elkaar verbonden zijn. Het netwerk leert door telkens de sterkte van onderlinge verbinding aan te passen op basis van de data die het krijgt over de buitenwereld.

### Meer over Machine Learning

We gebruiken allemaal software, applicaties die iets voor ons doen, zoals Word en Excel. Deze applicaties zijn geprogrammeerd. Bij machine learning gaat dat anders. Er wordt geen software geprogrammeerd, maar een model getraind. Het model zelf en de data die nodig is om het model te trainen, worden vaak wel bewerkt met een programmeertaal. Maar in de kern draait het om een lerend model.



Hoe meer (correcte) data,  
hoe beter.



Classificaties of clusters

Een model moet eerst trainen.

Model

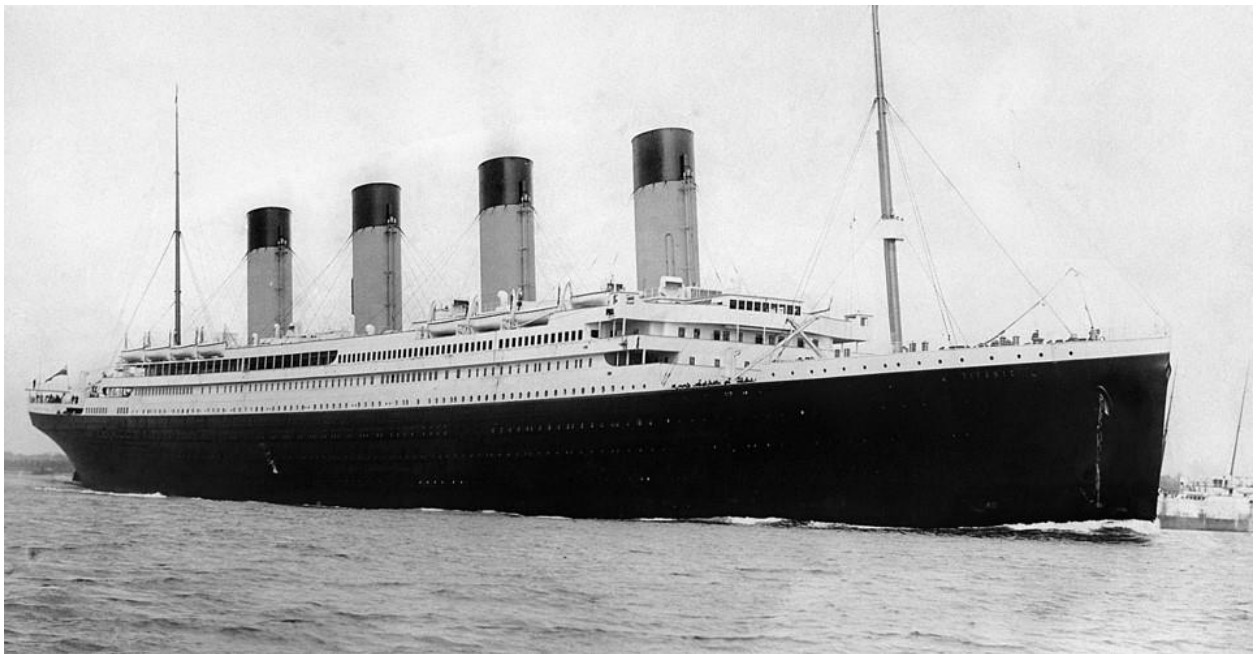
Regressie, een beslisboom  
of een neurale netwerk.

### Wat kan een Machine Learning model?

Een Machine Learning kan, als je het plat slaat, twee verschillende dingen, die in eerste instantie niet heel spannend lijken. Een getraind model kan een grote hoeveelheid data in logische groepen verdelen (clusteren), of een model kan voorspellen tot welke groep een item behoort (classificatie). Dit klinkt nog abstract. Een voorbeeld om het te verduidelijken.

### Wat kan een model voorspellen?

Zou je van de passagiers van de Titanic kunnen voorspellen wie de ramp overleeft en wie niet? Dus stel je krijgt de lijst met namen van de mensen die aan boord waren. Welke gegevens zou je dan nog meer willen hebben om te voorspellen: [naam] heeft het **overleefd** / **niet overleefd**.



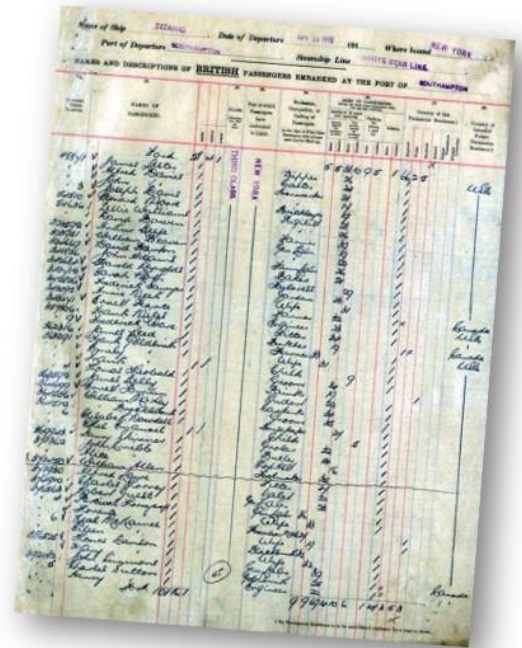
Je ziet misschien al dat dit een classificatie-probleem is. Maar zou het kunnen denk je? Laten we het eens proberen, wat heb je nodig?

Of de passagier een man of een vrouw is, zou dat uitmaken in de overlevingskans? En hoe oud degene is? Een kind van 3 maakt minder kans dan een man van 25!

En in welke klasse de passagier zat, eerste, tweede of derde klasse, zou dat uitmaken? En waar diegene is opgestapt?

Al deze gegeven helpen inderdaad mee bij het juist voorspellen van de uitkomst: overleefd / niet overleefd. Stel dat je een model hebt die je al deze gegevens voert, van alle passagiers op 10 na, die houd je achter. En je geeft het model ook of de betreffende persoon het heeft overleefd of niet.

Zou het model van die laatste 10 kunnen voorspellen wie het wel en niet hebben overleefd? Ja, dat kan zeker! En behoorlijk nauwkeurig ook!



Passagiersgegevens van de Titanic

---

*'How I got 98% prediction accuracy with Kaggles Titanic Competition'*

---



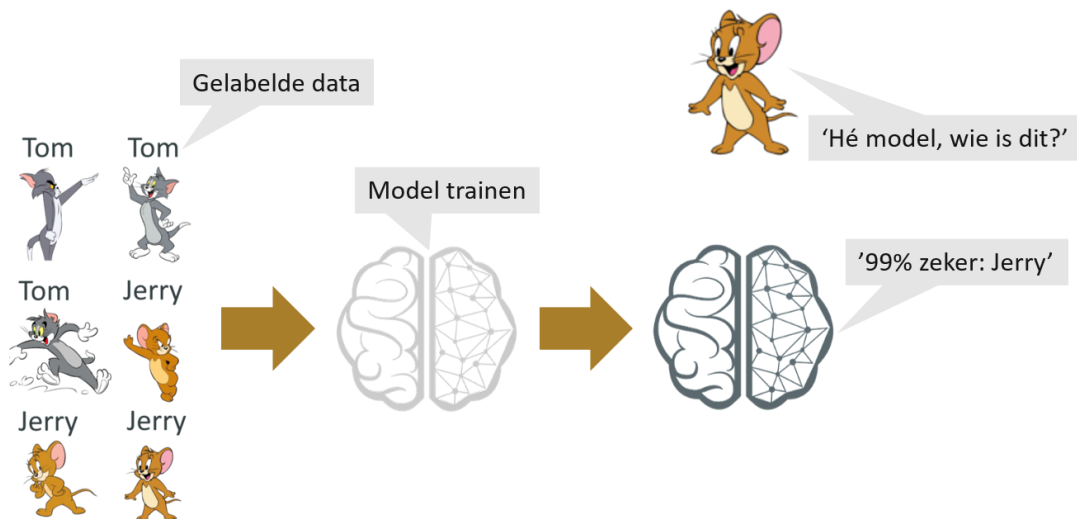
## Leren in 3 smaken

Modellen kunnen leren van data op 3 verschillende manieren:

- Supervised learning
- Unsupervised learning
- Reinforcement learning (deze laten we hier verder even buiten beschouwing.)

### Supervised learning

Bij supervised learning is de data die het model in gaat gelabeld. Het model kan zo tijdens de training continu toetsen hoe goed zijn voorspelling is.



Enmaal getraind kan het model met (een bepaalde zekerheid) voorspellen, welk 'label' er op nieuwe data geplakt moet worden. Is het label een bepaalde klasse, bijvoorbeeld temperatuurklassen: 'ijskoud', 'koud', 'warm', 'heet', dan noemen we dit classificeren. Is de voorspelling een specifieke waarde, zoals: 'het wordt morgen 13,8 graden', dan noemen we dit regressie.

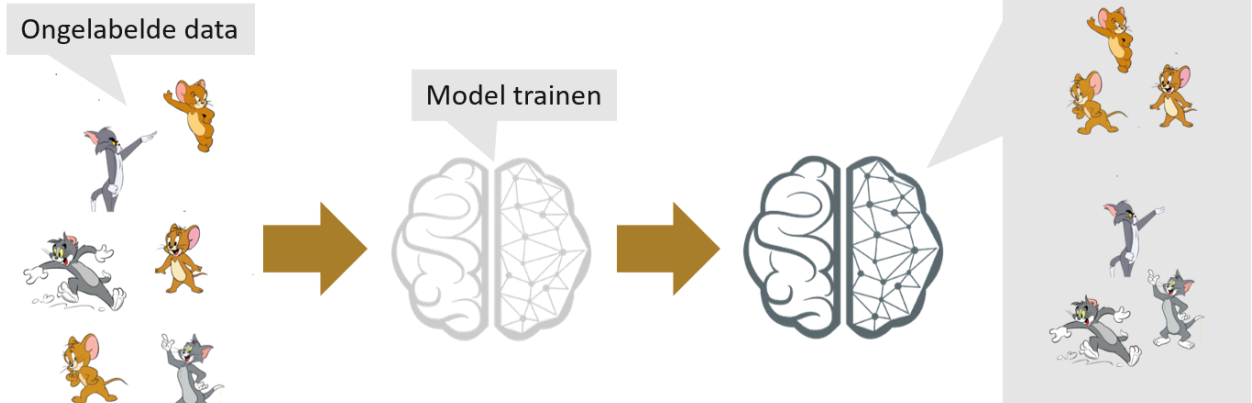
De Titanic-opgave is dus typisch supervised learning met als uitkomst een classificatie.

Een ander mooi voorbeeld van een supervised learning model is 'Deep Patient'. Dit model analyseert de medische voorgeschiedenis van een (toekomstige) patiënt om daarmee de kans op circa 80 aandoeningen te voorspellen, tot een jaar *voordat* de eerste symptomen zich bij de patiënt voordoen!

### Unsupervised learning

Bij unsupervised learning trainen we een model met ongelabelde data. De uitkomst is dan ook geen label, maar een voorspelling van logische groepen, ofwel clusters.





Het laten opdelen van data in groepen lijkt in eerste instantie misschien niet zo spannend, maar de toepassingen zijn heel spectaculair. Denk bijvoorbeeld aan het clusteren van alle banktransacties door een bank. Zodra een transactie niet in een cluster past - bijvoorbeeld omdat het om een miljoen dollar gaat dat 's nachts naar Zwitserland wordt overgemaakt - krijgt de beveiliging een seintje en kunnen ze de transactie onderzoeken. Je kunt zo dus ook afwijkingen (zgn. anomaly detection) in gegevens opsporen.

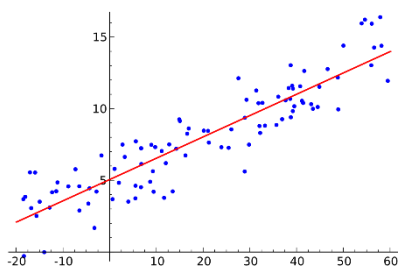
Of denk aan NETFLIX die je advies heeft over films die je waarschijnlijk ook leuk vind. Dat zijn clusters van films die op elkaar lijken!

## Verschillende technieken om een machine te laten leren

Naast de verschillende soorten leren, zoals supervised en unsupervised, zijn er verschillende technieken die als model kunnen dienen. We bespreken er 3.

### 1. Regressie

Zoals gezegd speelt statistiek een belangrijke rol in Machine Learning. Aan de basis staat regressie; het voorspellen van uitkomsten door de bestpassende wiskundige functie te vinden. (Bijvoorbeeld lineaire regressie  $Y = aX + b$ )

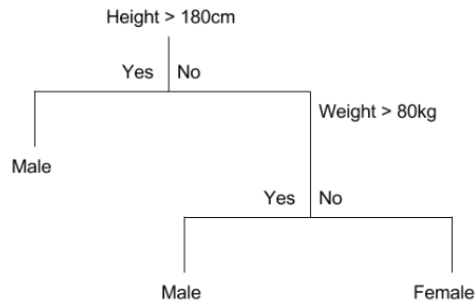


Hebben we de juiste lijn gevonden, dan kunnen we elke uitslag voorspellen, simpelweg een kwestie van x invullen en berekenen wat y wordt. Het vinden van die meest passende lijn is een truc die we hier even buiten beschouwing laten.



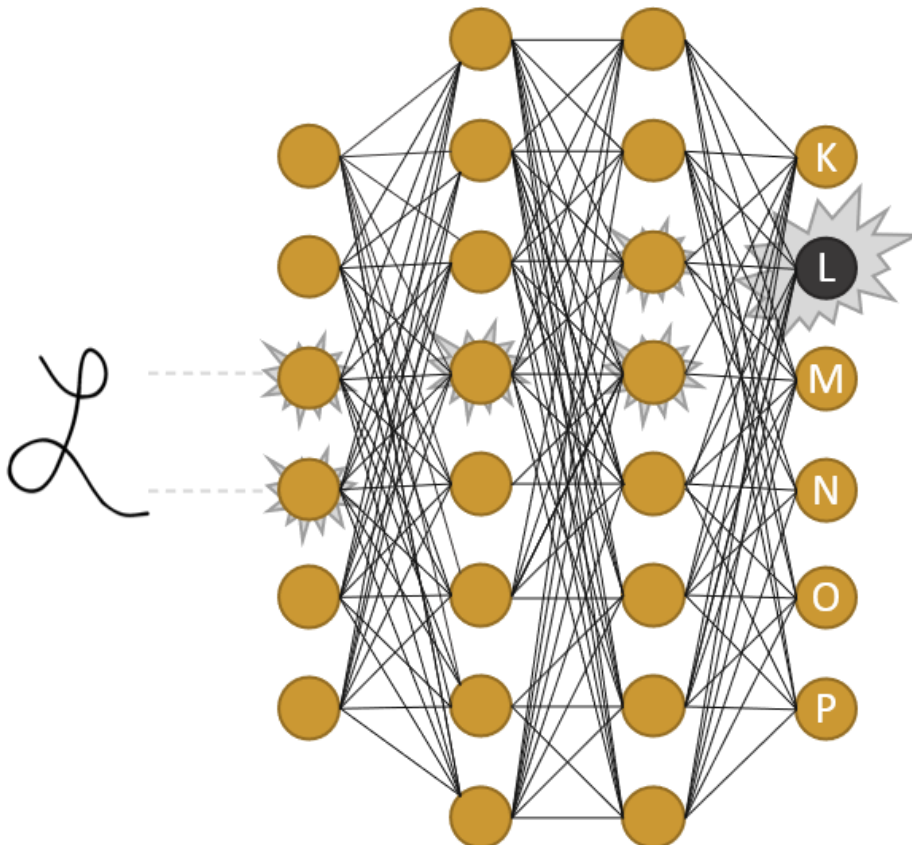
## 2. Een beslisboom

Een andere manier om te leren is om een beslisboom op te stellen en de vragen en antwoorden bij elke vertakking zo aan te passen, dat de uitkomst steeds accurater wordt.



## 3. Een neuraal netwerk

Een neuraal netwerk bestaat uit lagen met knooppunten die met elkaar verbonden, zoals neuronen in een menselijk brein. De knooppunten 'vuren' naar elkaar via verbindingen. Tijdens het trainen past het netwerk de verbindingen (weegfactoren) aan en zo herkent een netwerk bijvoorbeeld een handgeschreven 'L' in een automatische postsorteermachine.





Een model trainen is net Master Mind spelen  
 Hoe een model wordt getraind is een technisch verhaal, maar  
 het principe is vergelijkbaar met hoe wij Master Mind spelen;  
 slimme trial-and-error.

Master Mind spelen	Neuraal Netwerk
1. Je begint met 4 willekeurige kleuren	1. De verbindingen tussen de cellen hebben een willekeurige weegfactor.
2. Je leert van elke poging door de feedback	2. Het netwerk leert van elke trainingsronde
3. Je past iets aan en leert zo meer	3. Het netwerk past weegfactoren aan, leert meer
4. Tot je wint!	4. Tot het model de uitkomst goed kan voorspellen!

Nog een mooi voorbeeld van wat er met machine learning mogelijk is! Klik op de afbeelding:

