UCLL

# CONCEPTS OF DATA & ANALYTICS

Een algemene introductie

Les x

# LERENDE MACHINES

- **Zijn computers echt slim of voeren ze enkel geprogrammeerde instructies uit?**
- **Hoe slim is kunstmatige intelligentie?**
- **Hoe leren computers?**

*In de vorige les kwam je al te weten wat algoritmes zijn en hoe ze werken. Je maakte ook kennis met een aantal veel voorkomende algoritmes.*

Kunstmatige intelligentie. Het klinkt velen angstaanjagend in de oren. Zullen computers en robots ons leven gaan domineren? Is er nog wel een toekomst weggelegd voor ons? Heeft het nog wel zin om te studeren als computers en robots ons werk overnemen? Zal AI straks slimmer zijn dan de mens of is dit nu al het geval?

In deze module kom je te weten hoe ver het staat met de ontwikkeling van kunstmatige intelligentie. Hoe werkt AI nu eigenlijk? Kan je kunstmatige intelligentie nu eigenlijk wel vergelijken met menselijke intelligentie? Werken die beide vormen van intelligentie op dezelfde manier?

Zijn er terreinen waarop AI de mens nooit zal voorbijsteken?

# Hoe leren mensen en computers?

Computers volgen stappenplannen (**algoritmes**) om invoer tot de gewenste uitvoer te verwerken. Hoe slim die stappenplannen soms ook wel niet zijn, toch kunnen ze niet tippen aan menselijke intelligentie. Bovendien zijn al die algoritmes eerst door mensen bedacht. Computers slagen er namelijk niet in om zelf stappenplannen te bedenken of om de best mogelijke oplossingen te bedenken voor problemen.

Mensen kunnen dit wel. Ook al maak je in je leven behoorlijk wat foute keuzes, toch leer je in veel gevallen uit je fouten. We leren uiteraard niet alles door eerst fouten te maken. De meeste kennis en vaardigheden leren we door ervaring of door naar voorbeelden te kijken. Herhaling helpt daarbij, dat weten we allemaal.

# Kunstmatig leren

*Kunstmatige intelligentie betekent dat computersystemen zelf leren tot oplossingen te komen*. We programmeren software om zelf te leren door ervaring of aan de hand van voorbeelden, net zoals mensen en zelfs dieren dit doen.

Computerwetenschappers proberen hiervoor de werking van neuronen in de menselijke hersenen te simuleren met software. We spreken in dit geval van een **neuraal netwerk**.

# Leren uit voorbeelden

De meeste dingen in ons leven leren we door voorbeelden van anderen. Zelfs op school moeten we leren uit onze fouten.

> Stel dat je een kat als huisdier hebt en je zegt telkens tegen je zoontje van twee: “Dit is de kat”. Je zoontje heeft echter nog nooit een hond gezien. De kans bestaat dat hij bij zijn eerste ontmoeting met een hond zegt: “Kijk mama, kat!” Jij zal dan zeggen: “Nee hoor, dat is een hond”.
>
> Door dit vaak te herhalen, zal je zoontje snel het verschil leren tussen een hond en een kat. Hij zal geleidelijk beide leren onderscheiden en even zeer begrijpen dat de schapen van de buren al evenmin honden of katten zijn. Kinderen (ook op school) leren door “supervised learning”. Ze leren uit hun fouten en uit voorbeelden omdat iemand hen daar op wijst.

Door neurale netwerken in software te programmeren, kunnen computers zelf leren (machine learning). We voeren de AI-software eerst met **trainingssets**. Dat kunnen bijvoorbeeld een reeks foto's van honden zijn. Door een massa foto's van honden als trainingssets aan de software te voeren, kan het neuraal netwerk stilaan zelf honden op andere foto's leren herkennen.

## Gewicht geven

In ons leven moeten we voortdurend beslissingen nemen. Hierbij wegen we tal van factoren af. Welke kleren zullen we vandaag dragen? Welk geschenk ga ik kopen? Wat ga ik vanavond eten? Voor wie zal ik stemmen bij de verkiezingen? Hierbij maak je heel wat afwegingen en sommige factoren wegen sterker door in je eindbeslissing.

Uiteindelijk tel je al die factoren met elkaar op. Sommige zullen zwaarder doorwegen.

Een neuraal netwerk (of het nu gaat om een stel neuronen in onze hersenen of om een kunstmatig neuraal netwerk) krijgt op die manier heel wat "inputs" (afwegingen) die elk hun eigen gewicht krijgen. Het netwerk "berekent" een beslissing en stuur die uit als "uitvoer".

Als je twijfelt of je de fiets of de auto zou nemen om naar je werk te gaan, dan weeg je heel wat factoren af: het weer, de afstand, vermoeidheid, kostprijs, of je nog naar de winkel moet… De ene dag neem je de auto, de andere dag weer niet. Je uiteindelijke beslissing (output) hangt van heel veel inputs af. De neuronen in je hersenen maken heel wat (statistische) afwegingen. Misschien neemt je collega wel de fiets, terwijl jij de auto neemt. Hij/zij heeft misschien een andere afweging gemaakt. Als je echter de ene dag strontnat werd door een fikse regenbui, leer je in de toekomst eerst het weerbericht te bekijken voor je op je fiets stapt.

1. Scan de QR-code
2. Bekijk de films
3. Maak de bijhorende opdracht.

# Leren onder toezicht

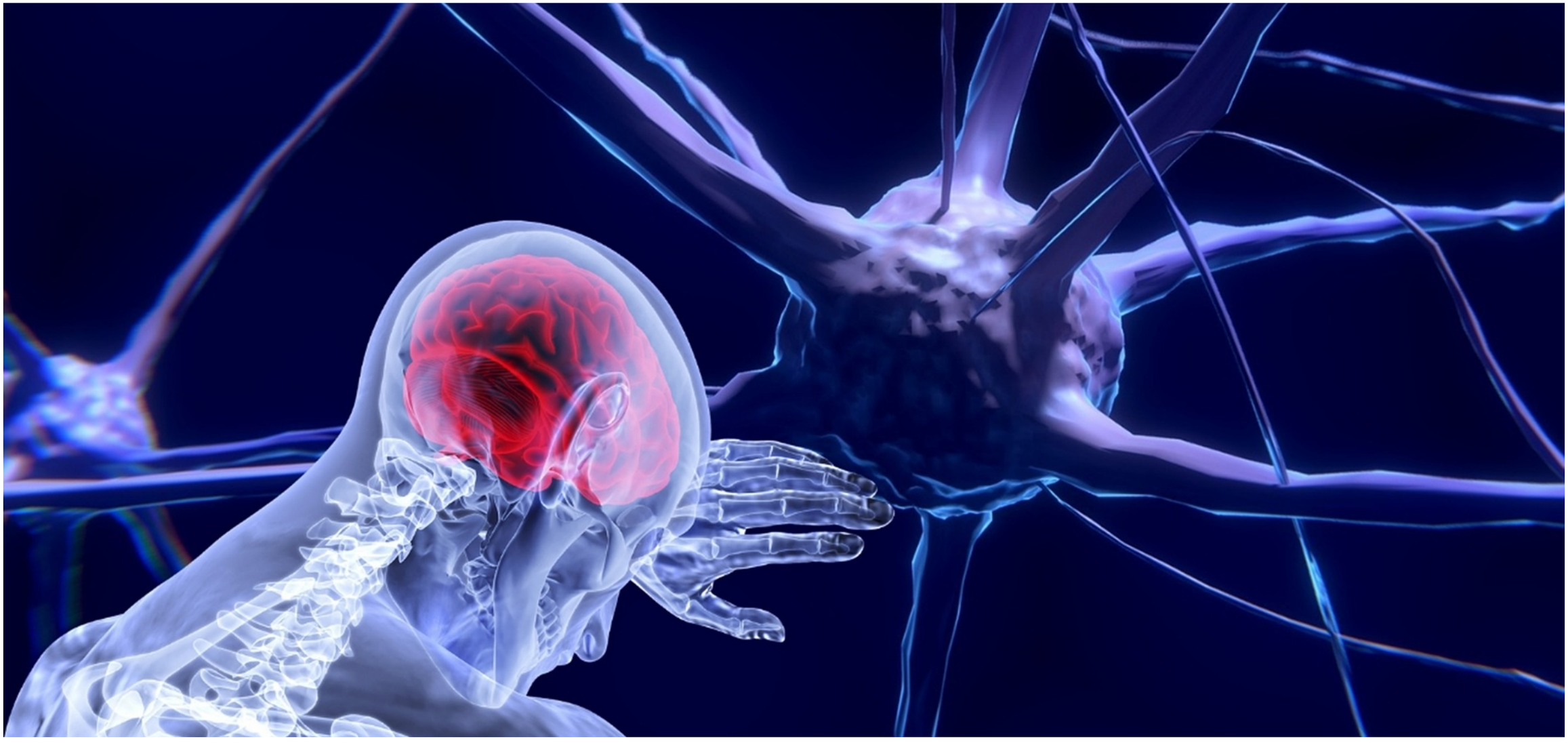
Het is niet omdat je een uitgebreide dataset/trainingset aan je neurale netwerk voedt, dat het vanaf dat moment foutloos werkt. Een neuraal netwerk leert eveneens uit zijn fouten. In deze fase is menselijke tussenkomst nog van belang. De mens kan zelf nog “controleren” of de beslissing van de het neurale netwerk ok is of niet.

Een spamfilter is een bekend voorbeeld van AI-software op basis van een neuraal netwerk. Heel wat mails die het woord “casino” en “penisverlenging” bevatten, zullen spamberichten zijn. Maar stel dat je vriend je een mail stuurt met de melding “Ga je vanavond mee naar het casino?” dan wil je niet dat dit bericht in de spam terechtkomt. Of als je echt problemen hebt en je huisdokter stuurt je een mailtje over je voortplantingsorgaan met daarin het zinnetje “we gaan je ziekteverlof verlengen”, dan wil je niet dat de combinatie van “verlengen” en “penis” in hetzelfde bericht tot “spam” leidt. Het neuraal netwerk moet heel wat afwegingen maken om te besluiten of een bericht al dan niet als spam gemarkeerd moet worden.

Zo’n spamfilter bouw je door een reeks van verdachte woorden samen te stellen. Berichten waarin een of meer van die woorden voorkomen, moeten al dan niet als spam worden herkend. Die verdachte woorden krijgen in het ene bericht de classificatie “true negative” (geen spam), in het andere “true positive” (wel degelijk spam).

Andere termen kunnen leiden tot de classificatie “false positive” of “false negative”. Dit gebeurt wanneer een spambericht onverwacht toch in je gewone mailbox terechtkomt of een goed bericht in de spam. In dat geval kan je een bericht zelf classificeren. Op dat moment leert de spamfilter dat hij een fout heeft gemaakt. Daar zal hij in de toekomst rekening mee houden. Spamfilters worden op die manier steeds beter en maken minder en minder fouten.

Daarom kan je zeggen dat AI-software letterlijk voorspellingen doet op basis van wat de software eerder is geleerd. Die voorspellingen kunnen ok zijn (true positive, true negative), maar even zeer fout (false positive of true negative). In dit laatste geval moet het algoritme bijgestuurd worden.



**Kennen**

1. Wat is supervised learning?
2. Wat is het nut van trainingssets?
3. Waarvoor dienen kunstmatige neurale netwerken?

**Kunnen**

1. Bepalen wanneer het zinvol zou kunnen zijn om een kunstmatig neuraal netwerk in te zetten.
2. Zelf een planmatige aanpak bedenken om een supervised neuraal network uit te bouwen.