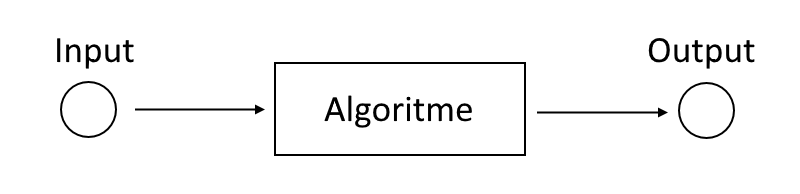
Wat zijn zelflerende systemen?

# Machine learning

Een zelflerend systeem is een systeem dat werkt met een of meerdere algoritmes gebaseerd op machine learning. Machine learning werd door Arthur Samuel, een pionier op dit gebied, gedefinieerd als: "A field of study that gives computers the ability to learn without being explicitly programmed.”[1]. In tegenstelling tot de eerder genoemde algoritmes is een zelflerend systeem in staat zichzelf te verbeteren. Hierdoor kan het taken uitvoeren waarbij reguliere algoritmes tekort schieten. Welke taken dit betreft, zullen we in de derde deelvraag behandelen.

Figuur Zelflerend systeem systematisch

In figuur \*\*\*\* is een schematische weergave van een zelflerend systeem afgebeeld. Bepaalde input data gaat het systeem in en bepaalde output data komt het systeem uit. De input en output data bestaat uit één of meerdere getallen. Als de input simpelweg een getal betreft, zal dit getal direct gebruikt kunnen worden als input. Als de input echter bijvoorbeeld bestaat uit een plaatje of een ander datatype dan cijfers, zal dit omgezet moeten worden in een reeks getallen om dit in een systeem te kunnen gebruiken. Het algoritme zal deze getallen bewerken tot de gewenste output. Deze output wordt eveneens in getallen gegeven. Waar nodig zullen deze getallen dus weer moeten worden omgezet tot het gewenste datatype.

Er zijn vele manieren waarop het systeem zichzelf kan verbeteren. Er is een verschil tussen de manier van leren en het gebruikte zelflerende algoritme. De manier van leren (Zie ook onder het kopje *Trainen*) bepaald de manier waarop het systeem de input data verwerkt………….

Training

Een zelflerend systeem begint in de meeste gevallen zonder enige kennis van de data. Om de gewenste output te kunnen produceren is het dus nodig om het systeem eerst input data te geven zodat het kan leren. Dit proces wordt het “trainen” genoemd. Voor het trainen van een zelflerend systeem is training data nodig. Deze data moet gelijk of gelijkwaardig zijn aan de “echte” data. De training data kan in veel verschillende vormen voorkomen en de manier van trainen is afhankelijk van de vorm van de (training) data. Er zijn drie prominente manieren waarop een zelflerend systeem getraind kan worden: supervised, unsupervised en reinforcement learning.  
  
Supervised Learning

In het geval van supervised learning heb je te maken met *labeled training data*. Anders gezegd: van een bepaalde input is de gewenste output al bekend. Een klassiek voorbeeld van een *labeled dataset* is een dataset van huisprijzen en huiseigenschappen (zie figuur xx)

|  |  |  |  |
| --- | --- | --- | --- |
| Huisprijs (output) | Huiseigenschappen (input) | | |
|  | Woonoppervlakte | Perceeloppervlakte | Aantal kamers |
| € 519.000 | 124 m² | 311 m² | 4 |
| € 569.000 | 133 m² | 309 m² | 5 |
| € 569.500 | 170 m² | 310 m² | 6 |

Figuur xx Labeled dataset   
Bron: http://www.funda.nl/koop/huizen/   
  
Bij de training dataset van figuur \*\*\*\* is de gegeven input de huiseigenschappen en de gewenste output de huisprijs. Het systeem wordt met deze dataset getraind. Hierdoor leert het een output te produceren die steeds dichter bij de gewenste output ligt. Als er een verband bestaat tussen de huiseigenschappen en de huisprijs, wat waarschijnlijk het geval is, zal het zelflerende systeem na genoeg trainen in staat zijn zelf bij nieuwe huiseigenschappen een huisprijs te voorspellen. [2]

## Unsupervised Learning

Unsupervised learning kan gebruikt worden bij een *unlabeled dataset* ofwel, een dataset waarbij de data niet geclassificeerd is en er geen gewenste output bekend is. Als je een dataset hebt van heel veel niet-geordende foto’s is het niet mogelijk om dit te classificeren. Als een deel van de dataset gelabeld wordt, zal met behulp van supervised learning de rest van de dataset geclassificeerd kunnen worden. Dit is echter in veel gevallen niet mogelijk, bijvoorbeeld doordat de dataset enorm groot is of er zodanig veel verschillende groepen bestaan dat het menselijk niet mogelijk is ook maar een deel te labelen. Ook kan het zo zijn dat men niet weet of er een verband aanwezig is.   
Kortom: unsupervised learning wordt gebruikt voor het classificeren van data, zonder dat er groepen vooraf gedefinieerd zijn. Met behulp van deze vorm van training zullen in een grote dataset verbanden kunnen worden ontdekt, die men misschien niet zonder hulp had kunnen achterhalen.[3]

## Reinforcement Learning

Reinforcement learning zorgt ervoor dat een “agent” een actie onderneemt in een bepaalde context.

Om dit te bereiken zijn er verschillend methoden, waarvan wij er een aantal nader zullen toelichten.

-niet alles bedacht

-learning over time (over data)  
-classification/regression  
-supervized 🡪 already catogorized  
-unsupervisid 🡪 geen labels  
-reinforcement 🡪 geen labels, geen goed/slecht  
-back propergation  
  
Arteficial Neural networks

-convenutional

-recurrent

-deep

# K Nearest Neighbours

# Support vector machine

# Classification vs Regression

# Evolutionary systems

deep learning 🡪 meerdere lagen

# Back propergation

# Bronnen

[1] Bron: <https://www.cims.nyu.edu/~munoz/files/ml_optimization.pdf>Geraadpleegd op: 21-5-2017  
Laatst gewijzigd op: niet bekend

[2] Bron: <https://www.coursera.org/learn/machine-learning/lecture/1VkCb/supervised-learning>  
Geraadpleegd op: 21-5-2017  
Laatst gewijzigd op: niet bekend

[3] Bron: https://www.coursera.org/learn/machine-learning/lecture/olRZo/unsupervised-learning  
Geraadpleegd op: 21-5-2017  
Laatst gewijzigd op: niet bekend