



Stichting NIOC en de NIOC kennisbank

Stichting NIOC (www.nioc.nl) stelt zich conform zijn statuten tot doel: het realiseren van congressen over informatica onderwijs en voorts al hetgeen met een en ander rechtstreeks of zijdelings verband houdt of daartoe bevorderlijk kan zijn, alles in de ruimste zin des woords.

De stichting NIOC neemt de archivering van de resultaten van de congressen voor zijn rekening. De website www.nioc.nl ontsluit onder "Eerdere congressen" de gearchiveerde websites van eerdere congressen. De vele afzonderlijke congresbijdragen zijn opgenomen in een kennisbank die via dezelfde website onder "NIOC kennisbank" ontsloten wordt.

Op dit moment bevat de NIOC kennisbank alle bijdragen, incl. die van het laatste congres (NIOC2025, gehouden op donderdag 27 maart 2025 jl. en georganiseerd door Hogeschool Windesheim). Bij elkaar zo'n 1500 bijdragen!

We roepen je op, na het lezen van het document dat door jou is gedownload, de auteur(s) feedback te geven. Dit kan door je te registreren als gebruiker van de NIOC kennisbank. Na registratie krijg je bericht hoe in te loggen op de NIOC kennisbank.

Het eerstvolgende NIOC vindt plaats op 18 maart 2027 in Arnhem en wordt georganiseerd door HAN University of Applied Sciences.

Reacties over de NIOC kennisbank en de inhoud daarvan kun je richten aan de beheerder:

R. Smedinga kennisbank@nioc.nl.

Vermeld bij reacties jouw naam en telefoonnummer voor nader contact.



**Hanzehogeschool
Groningen**

University of Applied Sciences

Ontofylogenetische Tertiaire Retentie: Machine Learning in het HBO

dr. Johan Blok

j.blok@pl.hanze.nl

**share your talent.
move the world.**

Onderzoeker-docent Software Engineering & Data Science

Lectoraat New Business & ICT Lectoraat Personalised Digital Health

Overzicht

1. Wat is Ontofylogenetische Tertiaire Retentie?
2. Wat is Machine Learning?
3. Machine Learning in Onderwijs en Onderzoek bij HBO-ICT

1. Ontofylogenetische Tertiaire Retentie

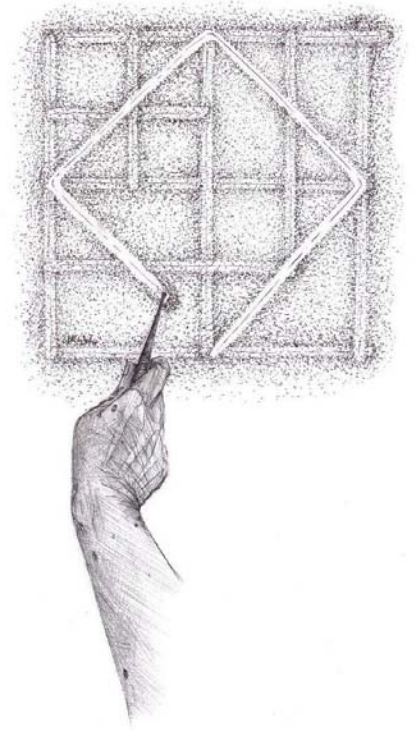
Bernard Stiegler (1952)

- Frans filosoof
- *Philosopher par accident (Per toeval filosoferen)*, 2004
- ~~Filosofie van de techniek~~
- Techniek maakt fundamenteel deel uit van 'mens zijn'.
- Maatschappijkritiek



Weten als herinneren

- Plato (Meno): weten als herinneren
- Stiegler (contra Plato): herinneren is altijd *hypomnesis* (mnemotechniek),
- Dus:
weten bestaat altijd (ten dele) uit mnemotechniek.



Ontofylogenetisch

Samentrekking van:

- *Ontogenetisch*: ontwikkeling van het individu.
- *Fylogenetisch*: ontstaan van groepen
- Menselijk geheugen is Ontofylogenetisch want product van:
 - Individuele ervaring (ontogenetisch)
 - Intergenerationele accumulatie van kennis (fylogenetisch)
- Techniek, d.w.z. technische artefacten constitueren het geheugen.

Retentie: vasthouden in het bewustzijn

Primaire versus secundaire retentie (Husserl/Brentano):

- Primaire retentie: vasthouden in het nu
- Secundaire retentie: vasthouden in het verleden (herinnering, product verbeelding)
- *Tertiaire Retentie* (Stiegler):

herhalen van temporeel object middels techniek

Wat voor stap wordt met ML gezet? Is er sprake van verder gaande vorm van retentie, namelijk:

Herhalen van het denken zelf middels ML technieken?



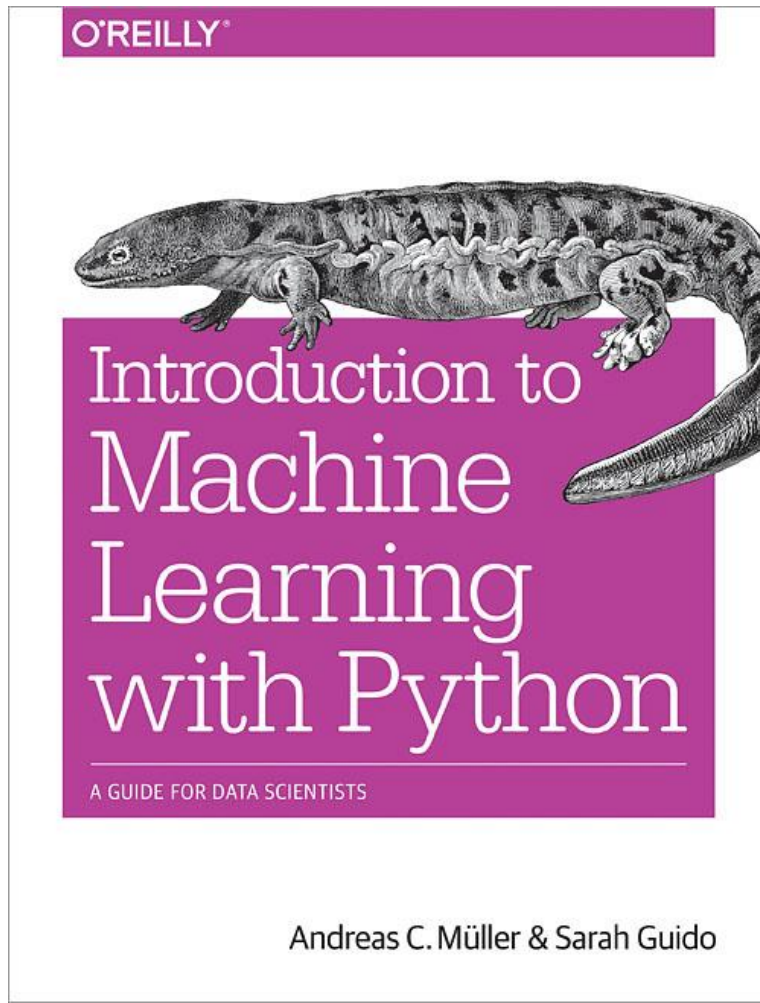
2. Machine Learning

Experimentele toepassing

- Opleving oude richtingen en ideeën uit o.a. Kunstmatige Intelligentie
- Modernisering, aanpassing, verbetering en vernieuwing

Explosief toegenomen beschikbaarheid:

- Opslag
- Rekenkracht
- Data

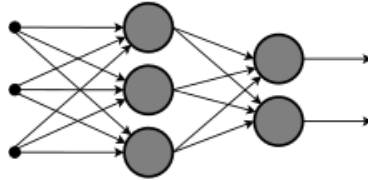


Machine Learning

Koken zonder recept

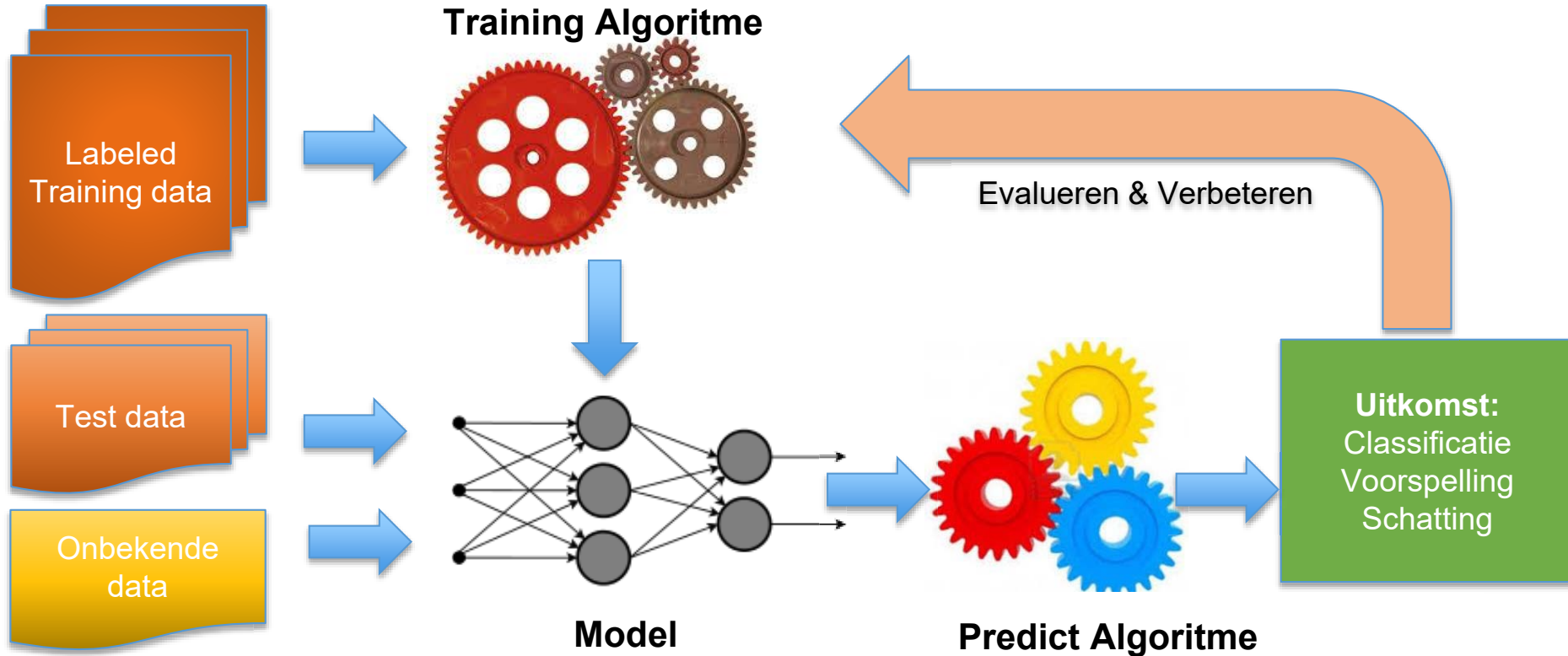


Model voor het
bakken van een
appeltaart



Machine Learning

Supervised



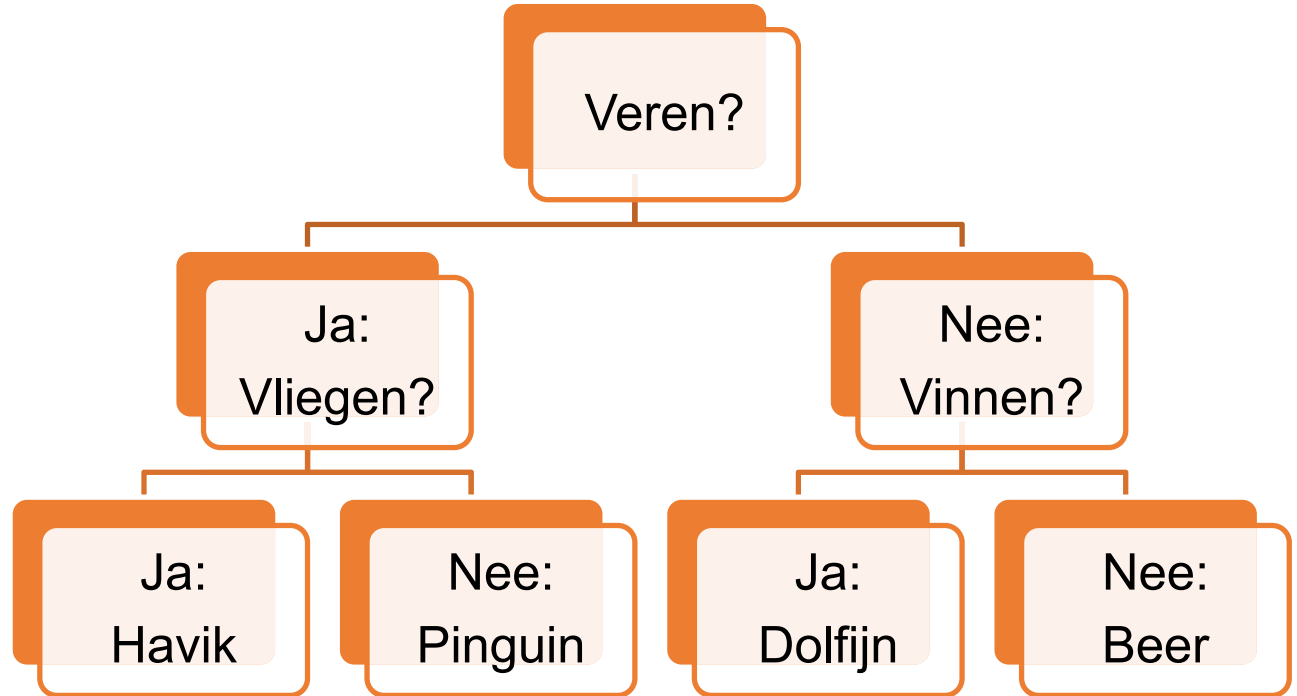
Decision Trees

Voorbeeld: classificatie in de biologie

*Dataset met dieren
en eigenschappen*

Dataset features:

Veren
Vliegen
Vinnen
Leeftijd



Voorbeeld in code

```
import numpy as np

from sklearn.model_selection import train_test_split
from sklearn.datasets import load_iris
from sklearn.tree import DecisionTreeClassifier

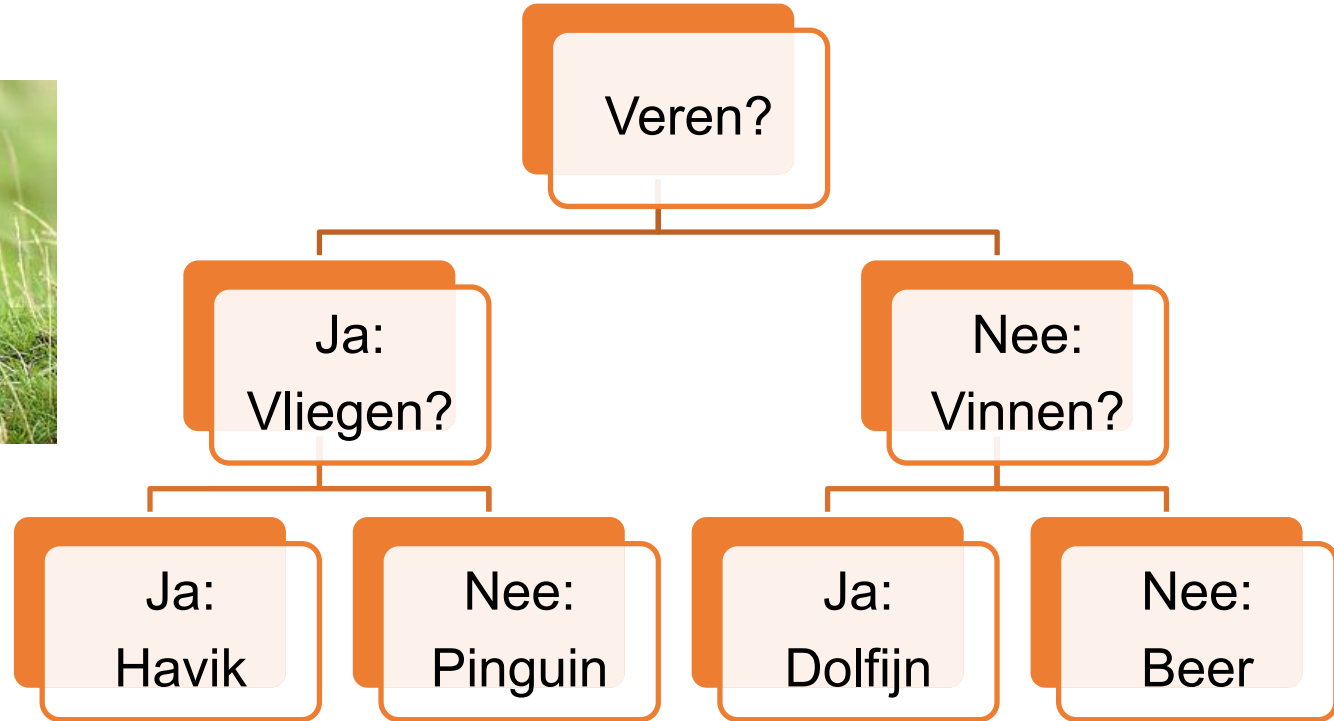
iris = load_iris()
X = iris.data
y = iris.target
X_train, X_test, y_train, y_test = train_test_split(X, y, random_state=0)

estimator = DecisionTreeClassifier(max_leaf_nodes=3, random_state=0)
estimator.fit(X_train, y_train)
```

Underfitting



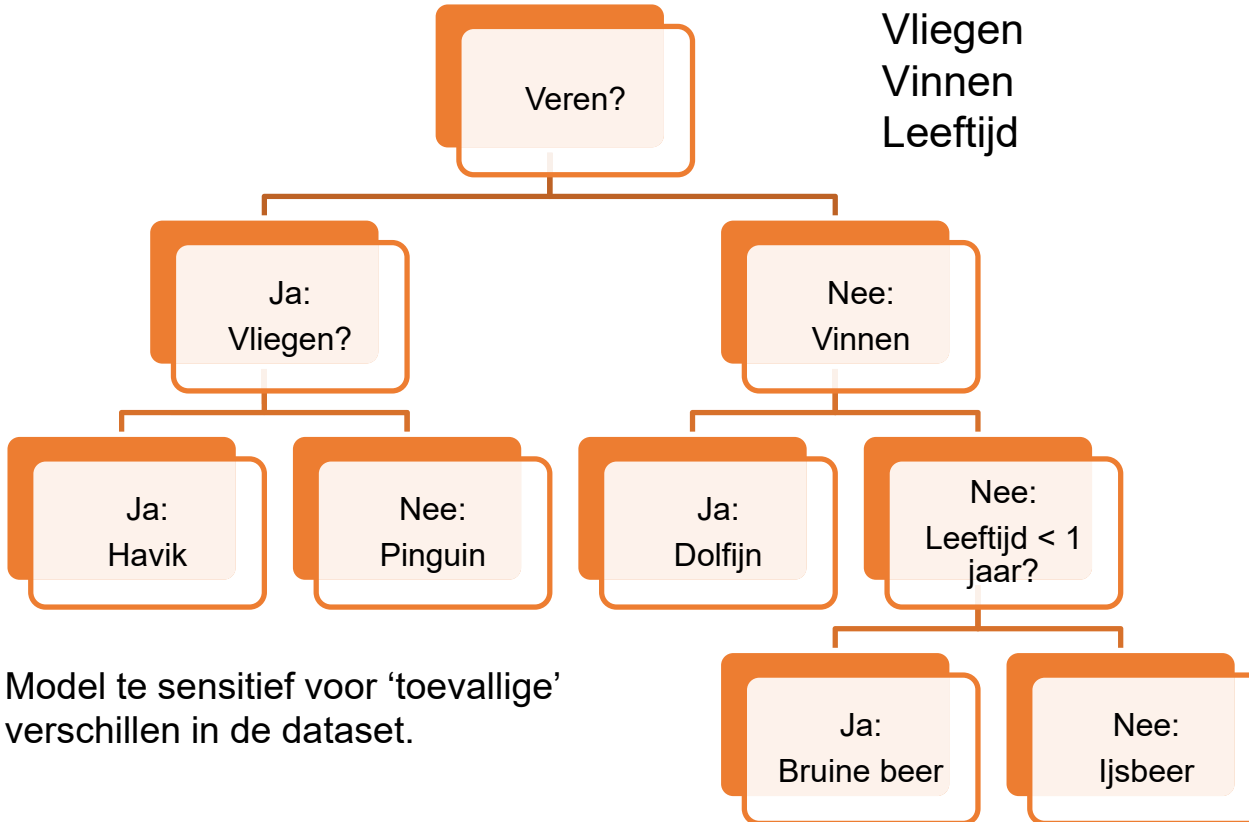
Model generaliseert te sterk en negeert relevante verschillen in de dataset.



Overfitting

Dataset features:

Veren
Vliegen
Vinnen
Leeftijd



Model te sensitief voor 'toevallige' verschillen in de dataset.



Underfitting and overfitting

Generalization

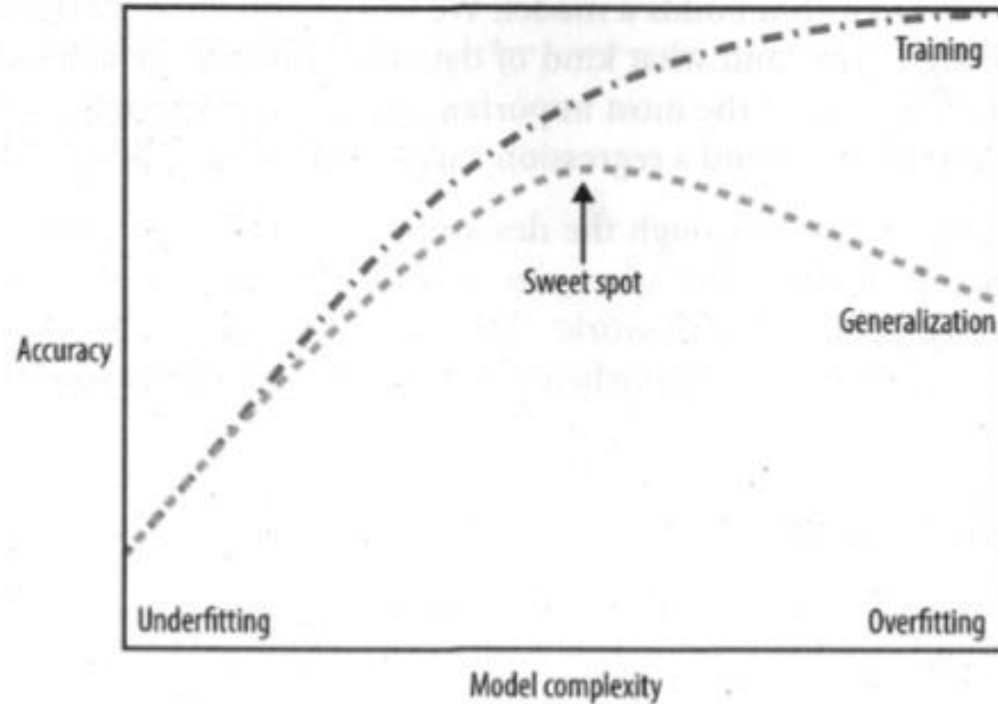
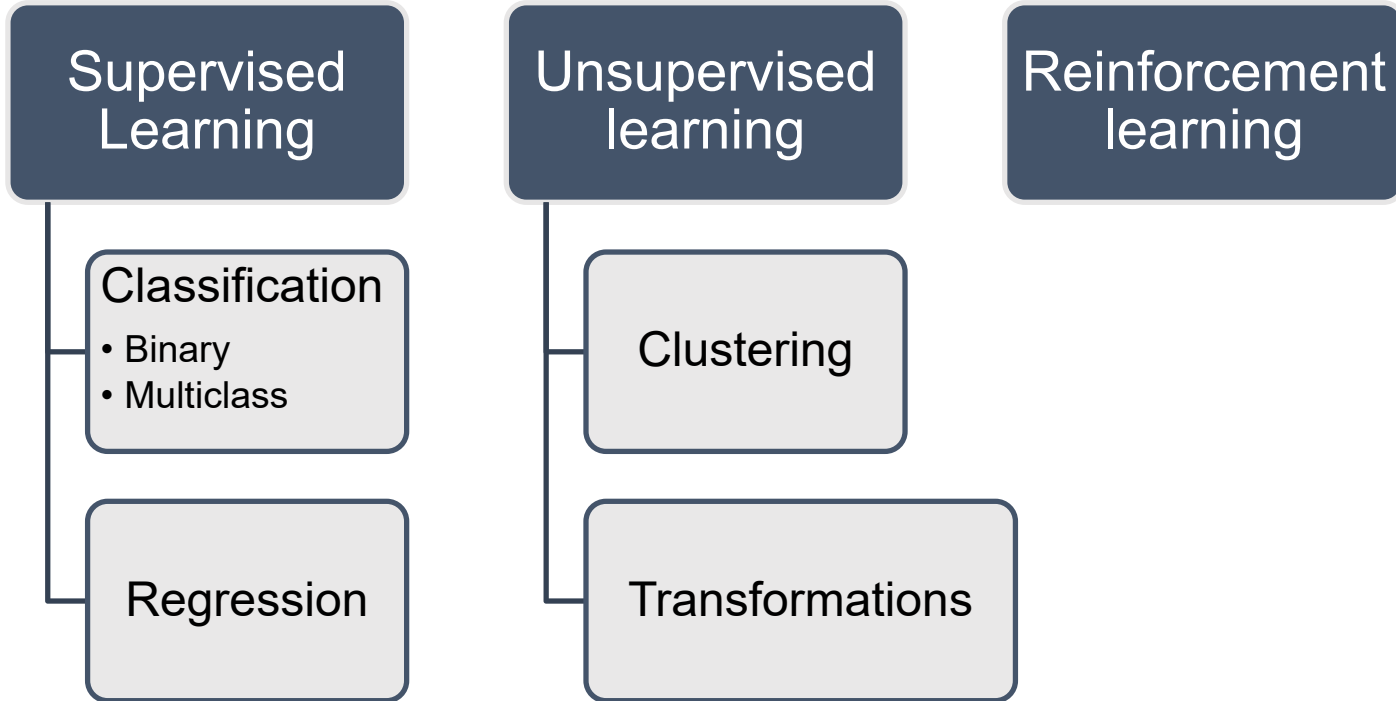


Figure 2-1. Trade-off of model complexity against training and test accuracy

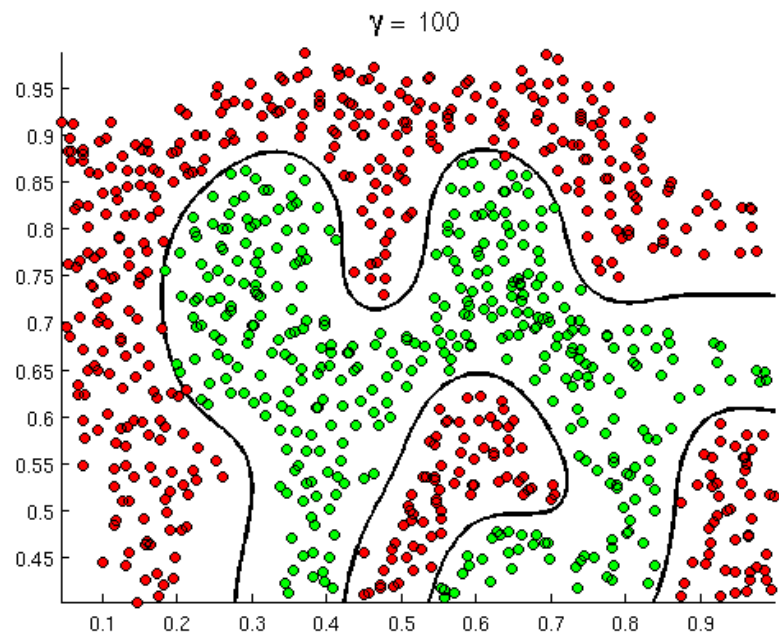
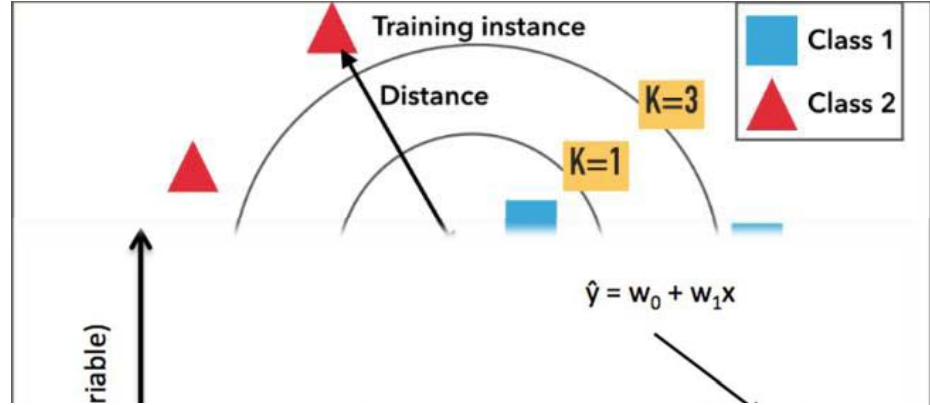
Types of machine learning



Machine Learning Algorithms

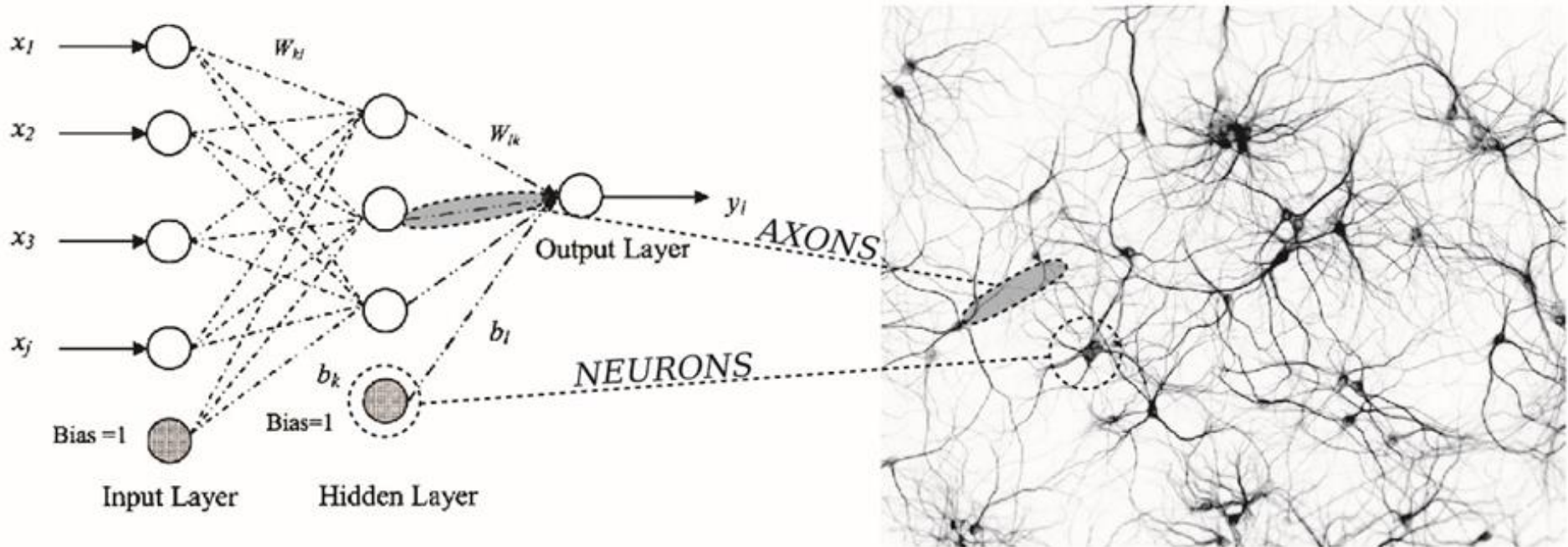
Supervised

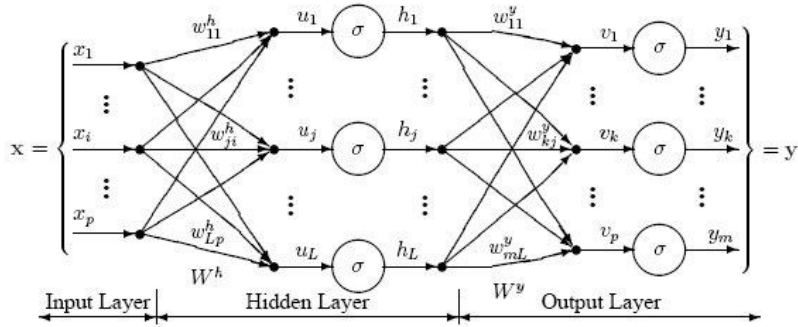
- Decision Tree
- Nearest Neighbour
- Linear Models
- Bayesian Models
- Support Vector Matrix
- Neural Networks (deep learning)



Neural Networks

‘Deep Learning’





- Complexe modellen
- Configuratie belangrijk, maar nauwelijks inzichtelijk
- CPU Intensief
- Complexe taken



bus parked on the street. a city street scene. front windshield of a bus. man walking on sidewalk. a silver car parked on the street. a city scene. a green traffic light. a building in the background. the bus has a number. a large building. a brick building. red brick building with windows. a blue sign with a white arrow. white lines on the road.

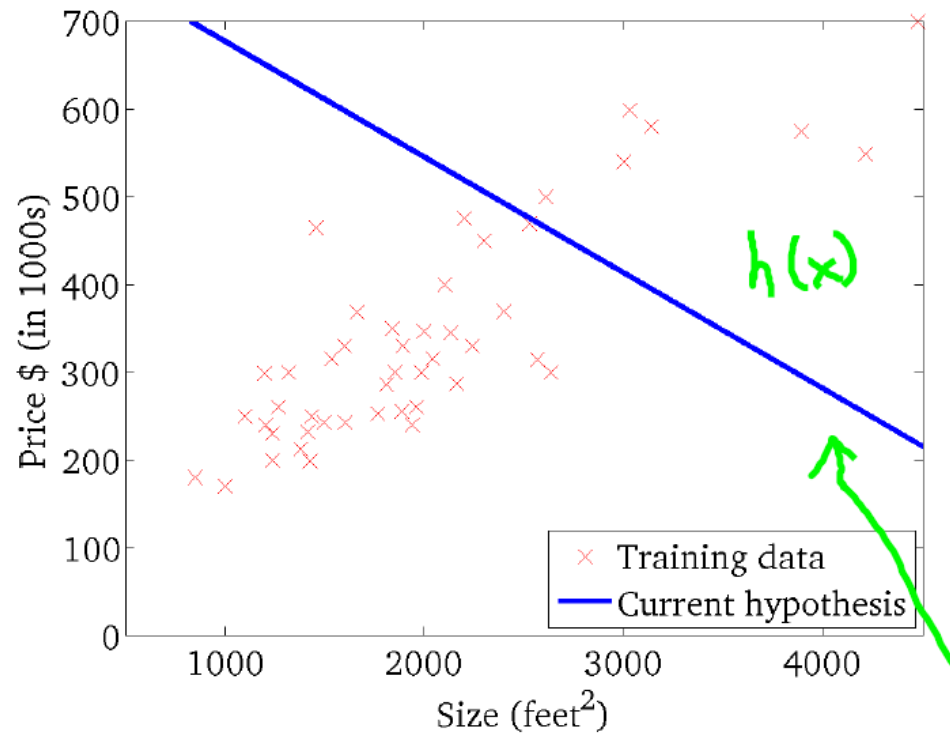
Wiskundig

Hypothesis: $h_{\theta}(x) = \theta_0 + \theta_1 x$

Parameters: θ_0, θ_1

Cost Function: $J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$

Goal: minimize $J(\theta_0, \theta_1)$
 θ_0, θ_1

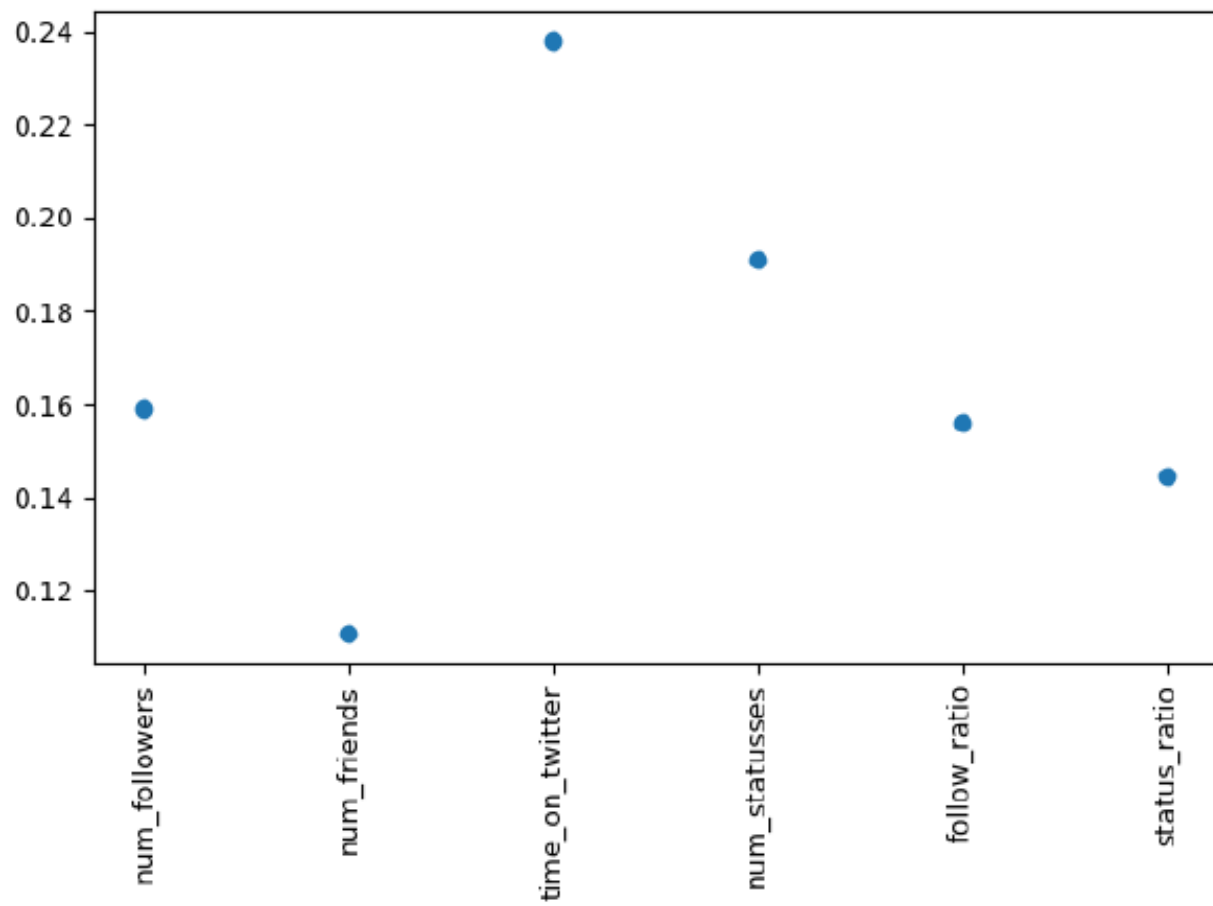


3. Machine Learning in het HBO

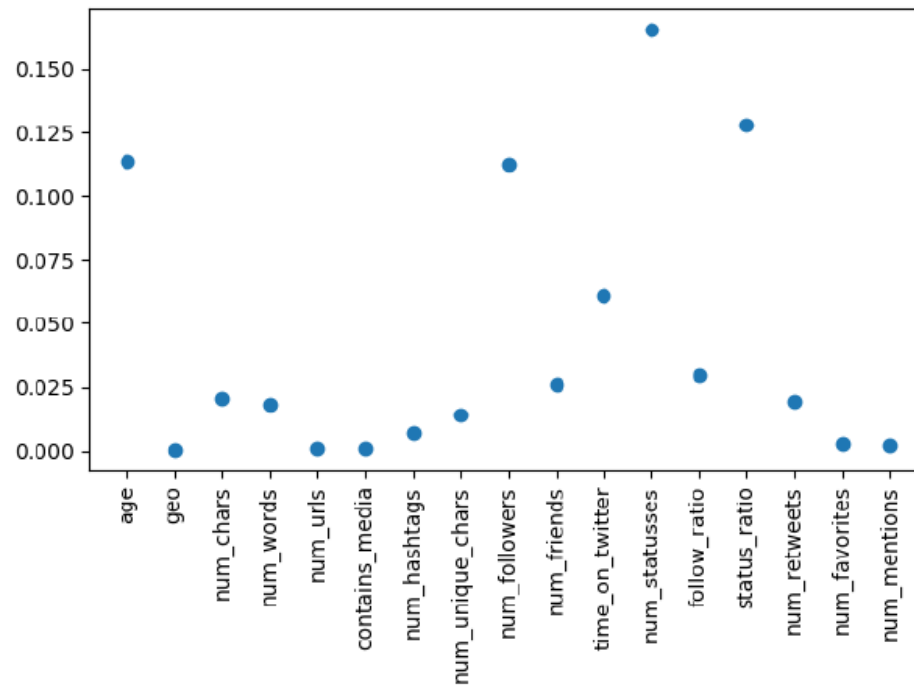
- Haalbaar?
- Zinnig?
- Op welke manier?

Vereist kennisniveau

- Basisgrip van statistiek: rol van kansen begrijpen bij opdelen in trainings- en testset.
- Conceptueel begrip van ML aanpak
- Voor- en nadelen diverse ML algoritmes (~~wiskundige essentie~~)
- Kennis van technieken als Cross-validation, Grid Search, e.d.
- Analyseren van resultaten middels Metrieken
- Beredeneerd verbeteringen doorvoeren



Figuur 4.9 Feature importance grafiek van experiment 3.



Figuur 4.6 Feature importance grafiek van experiment 2.

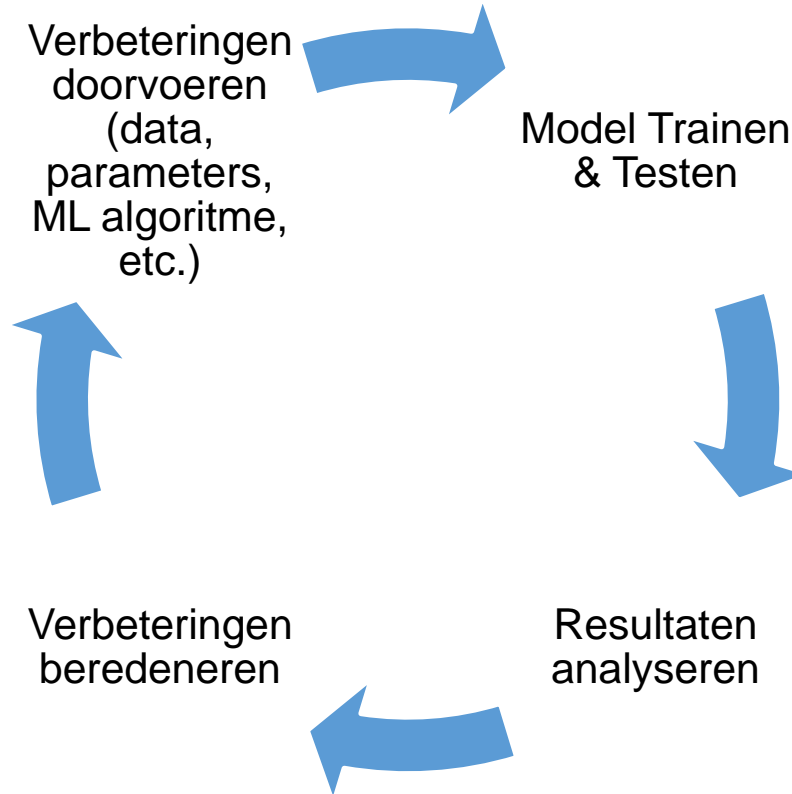
Belangrijke features	Onbelangrijke features
De tweet leeftijd	Geolocatie
Aantal volgers van de bron	Aantal urls
De leeftijd van het bron account	Bevat media
Aantal statussen door bron gepost	Aantal favorites door andere users
Leeftijd gedeeld door aantal statussen ratio	Aantal mentions van andere users

Tabel 4.2 De vijf meest en minst belangrijke features

Experiment	F ₁ -score	Precision	Recall
2 Gradient boosted met alle features	0.36	0.33	0.40
3 Gradient boosted met user features	0.20	0.21	0.20
4 Gradient boosted met selected features	0.44	0.44	0.44
5 Support vector machine met alle features	0.34	0.34	0.36
6 Support vector machine met user features	0.35	0.31	0.40

Tabel 4.3 Een overzicht van de resultaten van alle experimenten.

Cyclus



Experimenteel

Model	Resolution	Crops	Models	Top-1 error	Top-5 error
GoogLeNet ensemble	224	144	7	-	6.67%
Deep Image low-res	256	-	1	-	7.96%
Deep Image high-res	512	-	1	24.88	7.42%
Deep Image ensemble	variable	-	-	-	5.98%
BN-Inception single crop	224	1	1	25.2%	7.82%
BN-Inception multicrop	224	144	1	21.99%	5.82%
BN-Inception ensemble	224	144	6	20.1%	4.9%*

Figure 4: *Batch-Normalized Inception comparison with previous state of the art on the provided validation set comprising 50000 images. *BN-Inception ensemble has reached 4.82% top-5 error on the 100000 images of the test set of the ImageNet as reported by the test server.*

Accelerating training of deep neural networks: <https://arxiv.org/pdf/1502.03167.pdf>

Python notebook voorbeeld:

<https://github.com/rhievery/Data-Analysis-and-Machine-Learning-Projects/tree/master/example-data-science-notebook>

Praktijk

Onderwerpen:

- Detecteren van anomalie in logging
- Automatische data extractor (b.v. bedrijfsgegevens)
- Classificeren van tweets op betrouwbaar/onbetrouwbaar

Resultaten:

- Bruikbaar product vergt al snel beduidend meer tijd en kennis dan een afstudeerperiode
- Voor bedrijven kennismaking met toepassing van ML
- HBO-ICT'ers brugfuncties tussen 'data scientist' en software engineers

Machine Learning & Tertiaire Retentie

- Is ML van andere orde dan bij andere technologische stap? Is er sprake van een fundamenteel verdergaande van het ‘kennen’ en dus ‘denken’?
- ML maakt beslissingen herhaalbaar, maakt ‘leerprocessen’ herhaalbaar.
- Zorg Stiegler: massaal vermijden van ‘difference’, van verschillende interpretaties door massale consumentenindustrie die alles tot cliché’s reduceert.

ML in webwinkels: het kooppatroon van Jan lijkt op dat van Piet, dus geef ik Jan een aanbieding van een artikel dat Piet heeft gekocht.

“Denken is strijden tegen de eigen luiheid en gemakzucht”
(Stiegler, *Per toeval filosoferen*, blz. 111)