

Data-Science 1

machine
learning

KdG Karel de Grote
Hogeschool

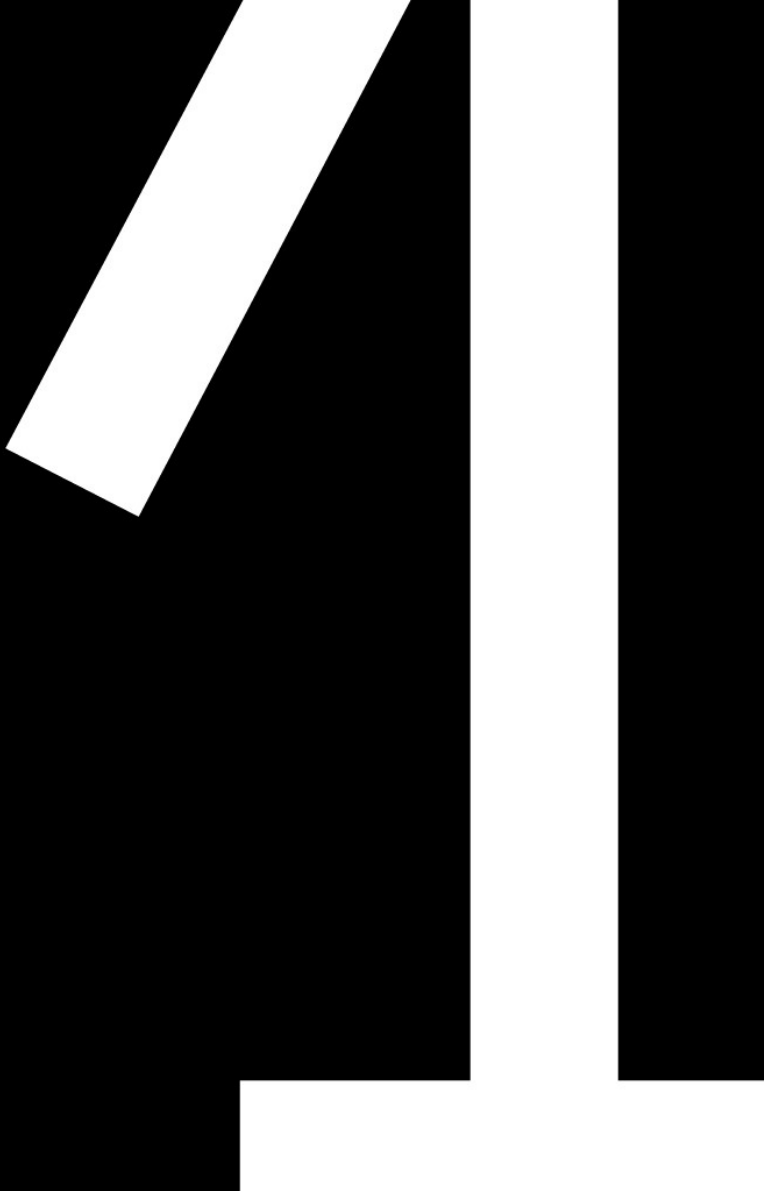


04/04/2024

Inhoud

- machine learning
 - wat is het?
 - supervised, unsupervised, reinforcement
 - regressie versus classificatie

Wat is machine
learning?

Abstract white geometric shapes on a black background, including a diagonal line, a vertical line, and a horizontal line.

WAAROM

Waarom doen we dit?

Wat willen we bereiken?

Welk resultaat verwachten we?

DATA ANALYSIS (OBJECTIVES)

DESCRIPTIVE
ANALYSIS

EXPLORATIVE
ANALYSIS

CONFIRMATORY
ANALYSIS

PREDICTIVE
ANALYSIS

PRESCRIPTIVE
ANALYSIS

OPERATIONS
RESEARCH

DATA ANALYSIS (TECHNIQUES)

SUMMERISATION

VISUALISATION

STATISTICS

MACHINE
LEARNING

EXPLAINABLE AI

OPTIMISATION

HOE

Hoe doen we dit?

Welke stappen nemen we?

Welke technieken gebruiken we?

WAAROM

Waarom doen we dit?

Wat willen we bereiken?

Welk resultaat verwachten we?

DATA ANALYSIS (OBJECTIVES)

DESCRIPTIVE
ANALYSIS

EXPLORATIVE
ANALYSIS

CONFIRMATORY
ANALYSIS

PREDICTIVE
ANALYSIS

PRESCRIPTIVE
ANALYSIS

OPERATIONS
RESEARCH

DATA ANALYSIS (TECHNIQUES)

Wat gebeurt er?

Waarom gebeurt het?

Zijn we zeker dat het gebeurt?

Wat zal er gebeuren?

Hoe kunnen we het laten gebeuren?

Heuristische optimalisatie

HOE

Hoe doen we dit?

Welke stappen nemen we?

Welke technieken gebruiken we?

**DATA ANALYSIS
(OBJECTIVES)**

**DATA ANALYSIS
(TECHNIQUES)**

Beschrijf de variatie in de data

SUMMERISATION

Visualiseer de patronen

VISUALISATION

Vergelijk de gemeten data met de verwachte data

STATISTICS

Zoek automatisch naar patronen in de data

**MACHINE
LEARNING**

Verklaar de werking van het 'black box' algoritme

EXPLAINABLE AI

Vind heuristieken om de oplossing te
optimaliseren

OPTIMISATION

HOE

Hoe doen we
dit?

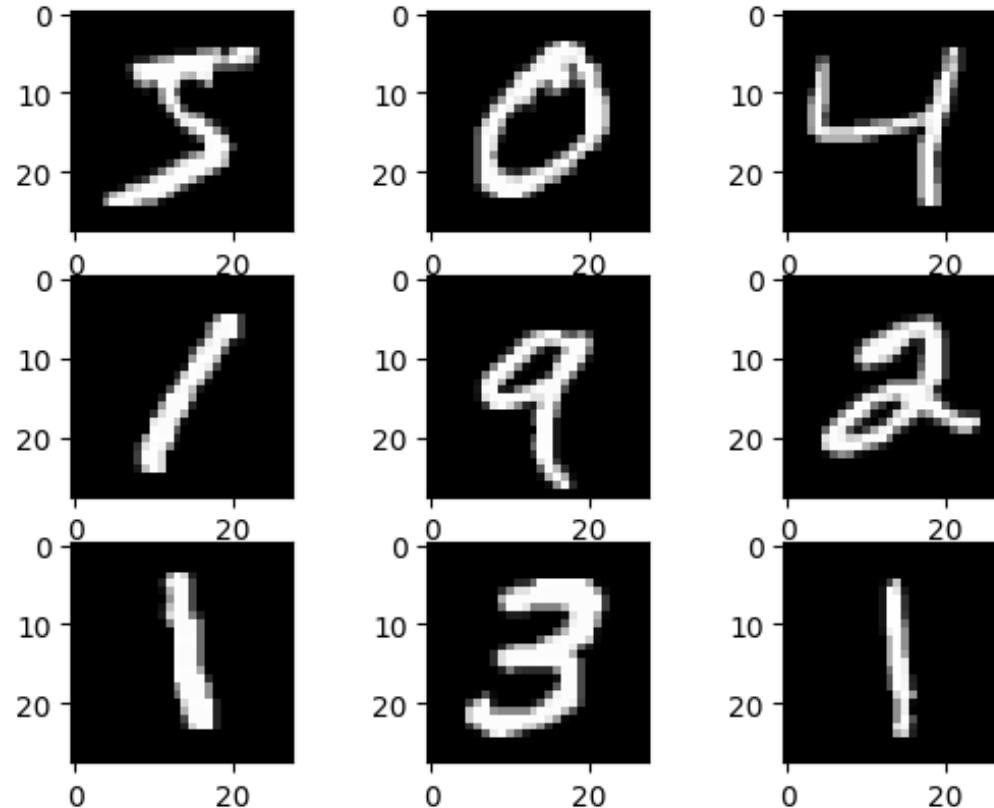
Welke stappen
nemen we?

Welke
technieken
gebruiken we?

Welk probleem?

- gegeven: data uit het verleden
- gevraagd:
 - zoek structuren in de data (dit is eigenlijk nog exploratieve analyse)
 - gebruik die structuren om een model te creëren dat toekomstige waarden kan voorspellen

Voorbeeld: cijfers herkennen



Voorbeeld: cijfers herkennen

- MNIST databank: 70000 beeldjes van handgeschreven cijfers
- vraag: kan je een afbeelding herkennen en het juiste label geven?
- ieder beeld is 28x28 pixels
- iedere pixel is 1 byte(grijswaarde)
- tabel met 70000 lijnen en 785 kolommen (1 kolom met label)

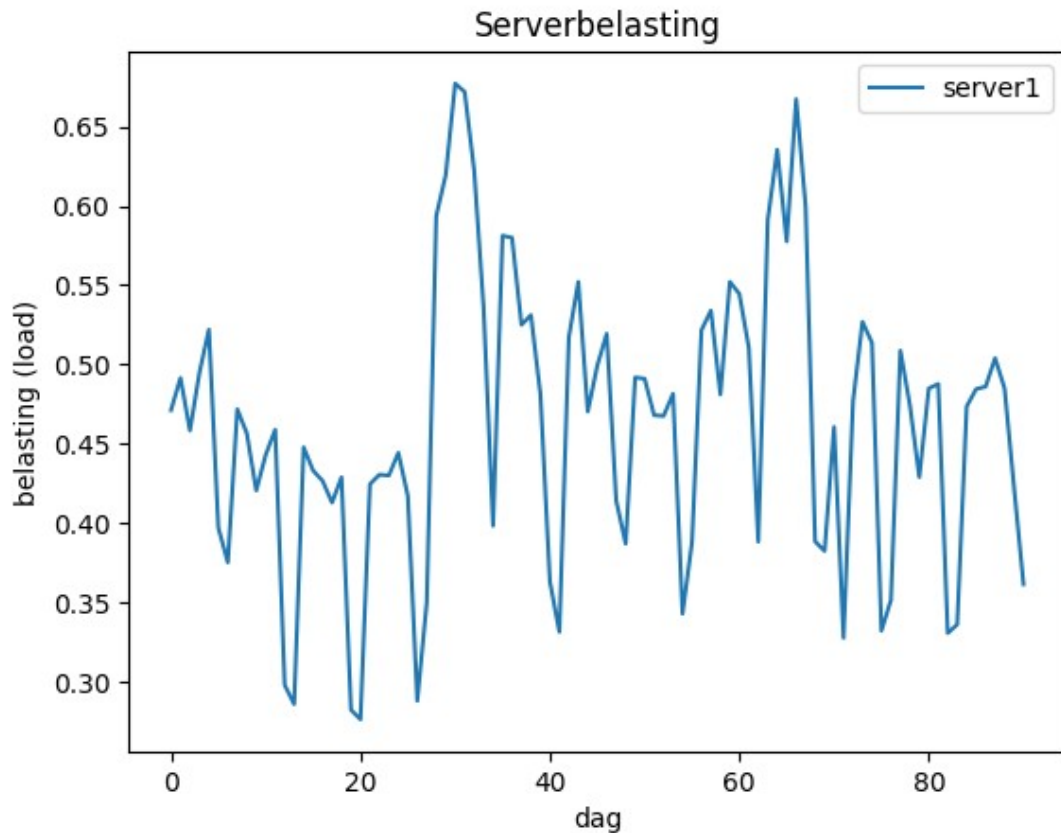
Voorbeeld: lening toewijzen

	age	sex	region	income	married	children	car	save_act	current_act	mortgage	pep
1	48	FEMALE	INNER_CITY	17546	NO	1	NO	NO	NO	NO	YES
2	40	MALE	TOWN	30085.1	YES	3	YES	NO	YES	YES	NO
3	51	FEMALE	INNER_CITY	16575.4	YES	0	YES	YES	YES	NO	NO
4	23	FEMALE	TOWN	20375.4	YES	3	NO	NO	YES	NO	NO
5	57	FEMALE	RURAL	50576.3	YES	0	NO	YES	NO	NO	NO
6	57	FEMALE	TOWN	37869.6	YES	2	NO	YES	YES	NO	YES
7	22	MALE	RURAL	8877.07	NO	0	NO	NO	YES	NO	YES
8	58	MALE	TOWN	24946.6	YES	0	YES	YES	YES	NO	NO
9	37	FEMALE	SUBURBAN	25304.3	YES	2	YES	NO	NO	NO	NO
10	54	MALE	TOWN	24212.1	YES	2	YES	YES	YES	NO	NO
11	66	FEMALE	TOWN	59803.9	YES	0	NO	YES	YES	NO	NO
12	52	FEMALE	INNER_CITY	36659.9	NO	0	YES	YES	YES	YES	NO

Voorbeeld: lening toewijzen

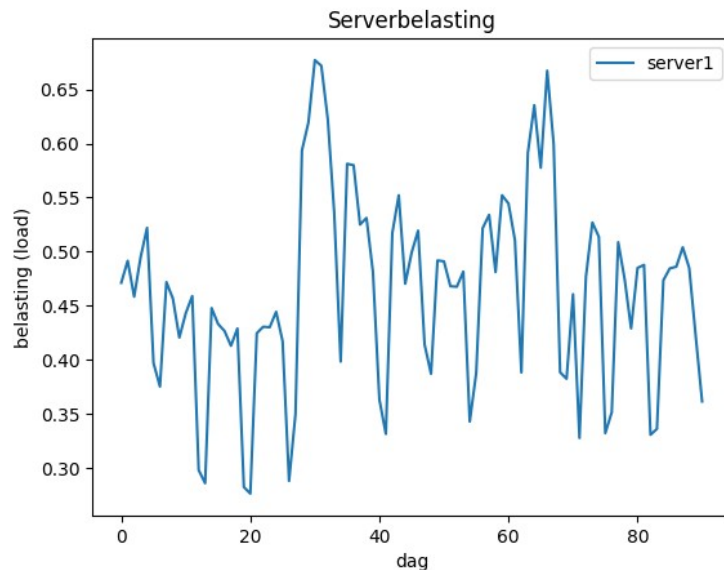
- gegeven: tabel met gegevens van vorige klanten en indicatie of ze hun lening konden terug betalen of niet (zie bank-data.csv)
- gevraagd: kan je voorspellen of een nieuwe klant de lening kan afbetalen of niet? Van wat hangt dit af?

Voorbeeld: serverbelasting



Voorbeeld: serverbelasting

- gegeven: serverbelasting in functie van de tijd
- gevraagd: kan je voorspellen wat de serverbelasting zal zijn op een bepaald moment in de toekomst?








Voorbeeld: muziek


Recommended Songs ^

Based on the songs in this playlist

REFRESH

<button>ADD</button>	Unstoppable	The Score	ATLAS
<button>ADD</button>	Rule the World	Zayde Wolf	Modern Alchemy
<button>ADD</button>	Blood // Water	grandson	Blood // Water
<button>ADD</button>	Legend	The Score	ATLAS
<button>ADD</button>	Legends Are Made	Sam Tinnesz	Babel
<button>ADD</button>	Stronger	The Score	Carry On
<button>ADD</button>	New Kings	Sleeping Wolf	The Silent Ones
<button>ADD</button>	Gladiator	Zayde Wolf	Modern Alchemy
<button>ADD</button>	The Heat	The Score	ATLAS (Deluxe)
<button>ADD</button>	Lions Inside	Valley Of Wolves	Out For Blood

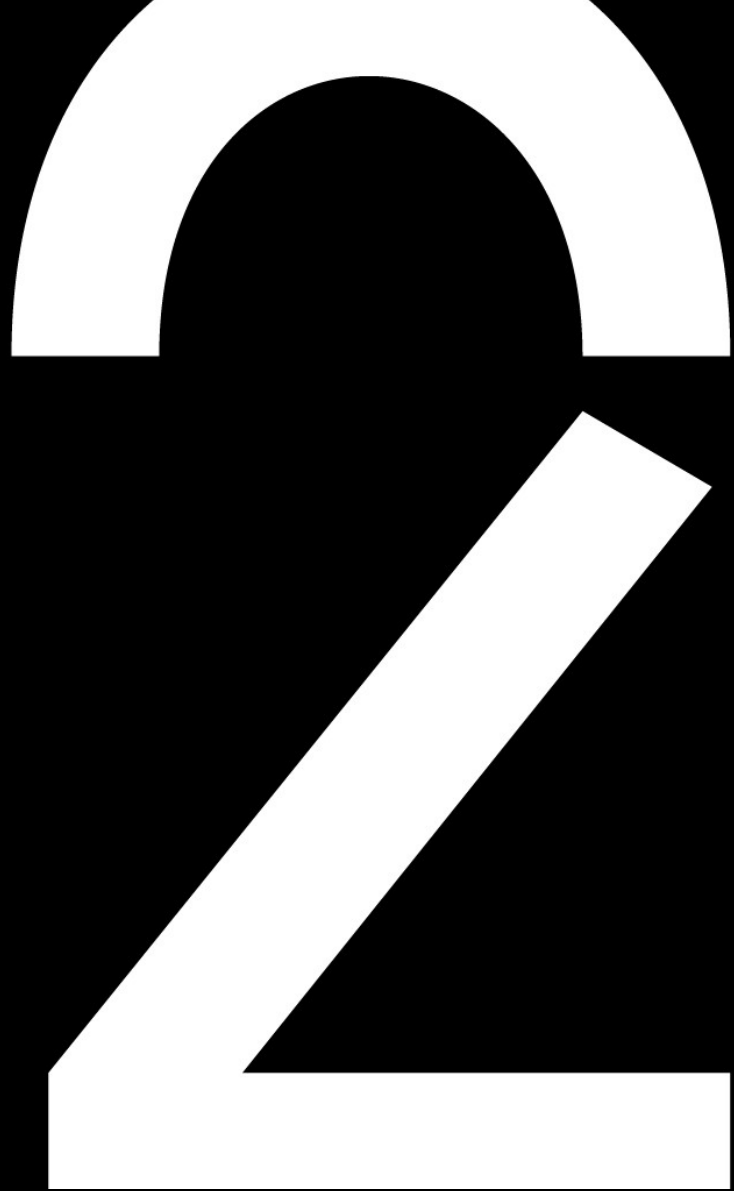


2:20  2:51

Voorbeeld: muziek

- gegeven: klanten luisteren naar bepaalde muziek
- gevraagd: welke muziek zou interessant kunnen zijn voor een bepaalde klant?
- m.a.w.: als een klant naar een nummer luistert, welke andere nummers hebben dan veel kans om ook beluisterd te worden?

Supervised
Unsupervised
Reinforcement



Supervised learning

- voorbeeld: cijfers herkennen
- gegeven: afbeeldingen van handgeschreven cijfers, iedere afbeelding is gelabeld
- de computer leert uit deze voorbeelden en kan zichzelf corrigeren a.d.h. van de labels

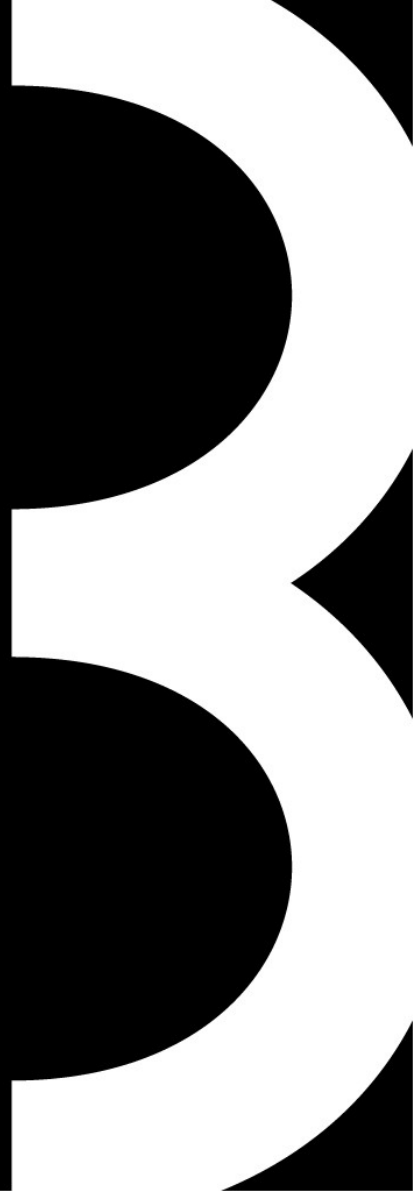
Unsupervised learning

- voorbeeld: clustering
- gegeven: afbeeldingen van verkeersborden
- de afbeeldingen zijn niet gelabeld
- computer zoekt naar gelijkaardige afbeeldingen en groepeerde deze

Reinforcement learning

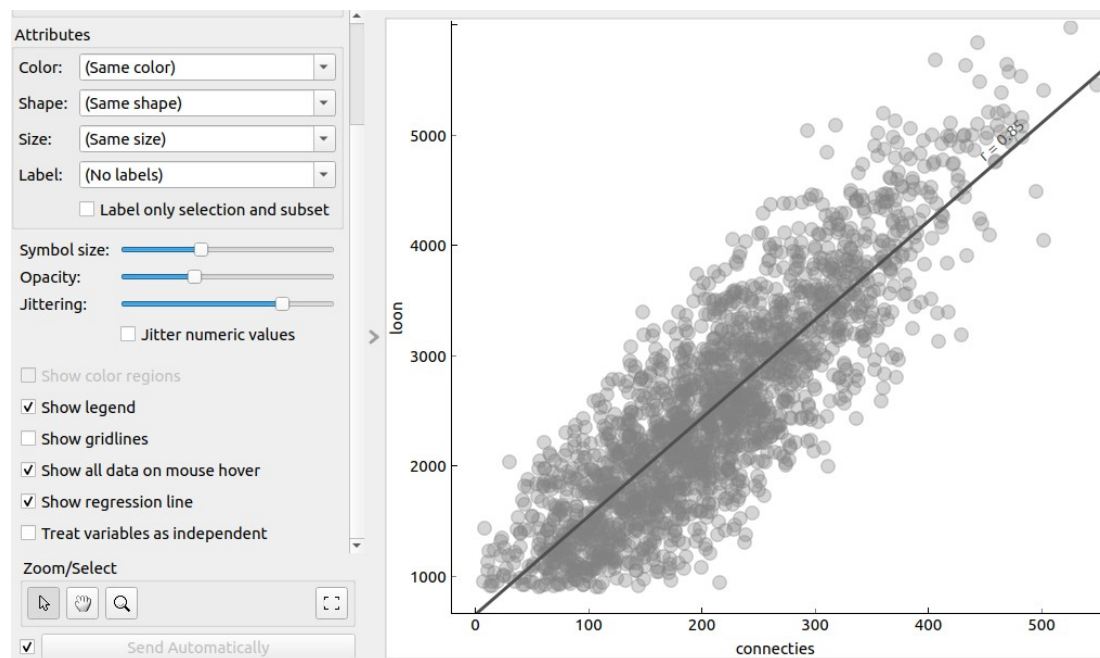
- voorbeeld: robot moet kunnen wandelen
- computer begint met willekeurige bewegingen
- als de robot valt krijgt de computer een sein dat er iets moet verbeteren
- computer zoekt zelf naar een manier om minder fouten te maken

Regressie vs classificatie



Regressie

- wat was regressie ook al weer?



Regressie

- regressie is een vorm van supervised machine learning!
 - je geeft voorbeelden (puntenwolk)
 - de lijn laat toe om voorspellingen te maken

Regressie

- wat is het meetniveau van het resultaat van een regressie?
- is dit een discrete of continue variabele?

Regressie

- vanaf nu: alle machine learning technieken die een continue waarde voorspellen, noemen we een regressietechniek

Classificatie

- wat als je een discrete variabele wil voorspellen?
- voorbeeld: karakters herkennen
 - wat is de input?
 - wat is de output?
- dit is een “classificatie”

Overzicht



Overzicht technieken in deze cursus

	supervised	unsupervised
regressie	<ul style="list-style-type: none">• lineaire regressie• neurale netwerken	<ul style="list-style-type: none">• association rule mining
classificatie	<ul style="list-style-type: none">• beslissingsbomen• neurale netwerken	<ul style="list-style-type: none">• clustering