

Data-Science 1

machine
learning

KdG Karel de Grote
Hogeschool



24/04/2024

Inhoud

- evaluatiemetrieken voor classificatie
 - welk probleem?
 - confusion matrix
 - accuracy
 - precision en recall
 - F-measure
 - andere maten

Problem

Voorbeeld

- gegeven: beslissingsboom
- vraag: hoe goed is deze beslissingsboom om waarden te voorspellen?

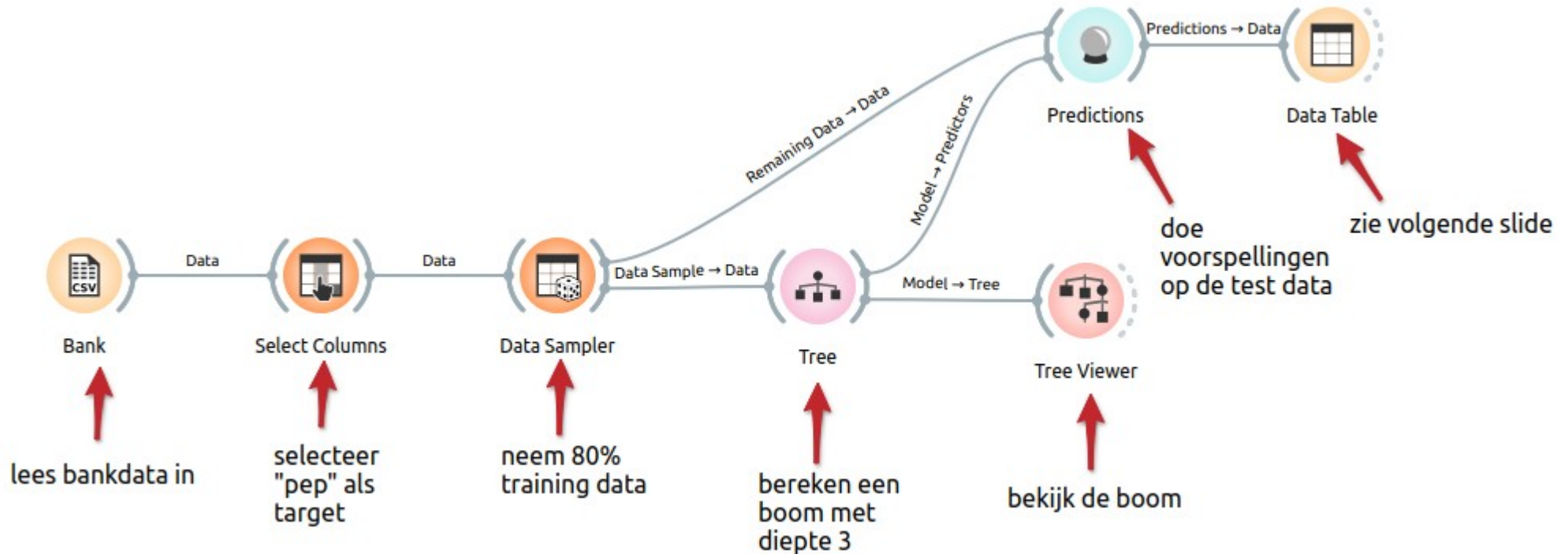
Oplossing

- splits data in
 - training dataset
 - test dataset
- maak de boom met de training dataset
- test de boom met de test dataset

Resultaat

- we bekomen nu een reeks voorspelde waarden door de boom en een reeks werkelijke waarden

Voorbeeld: bankleningen



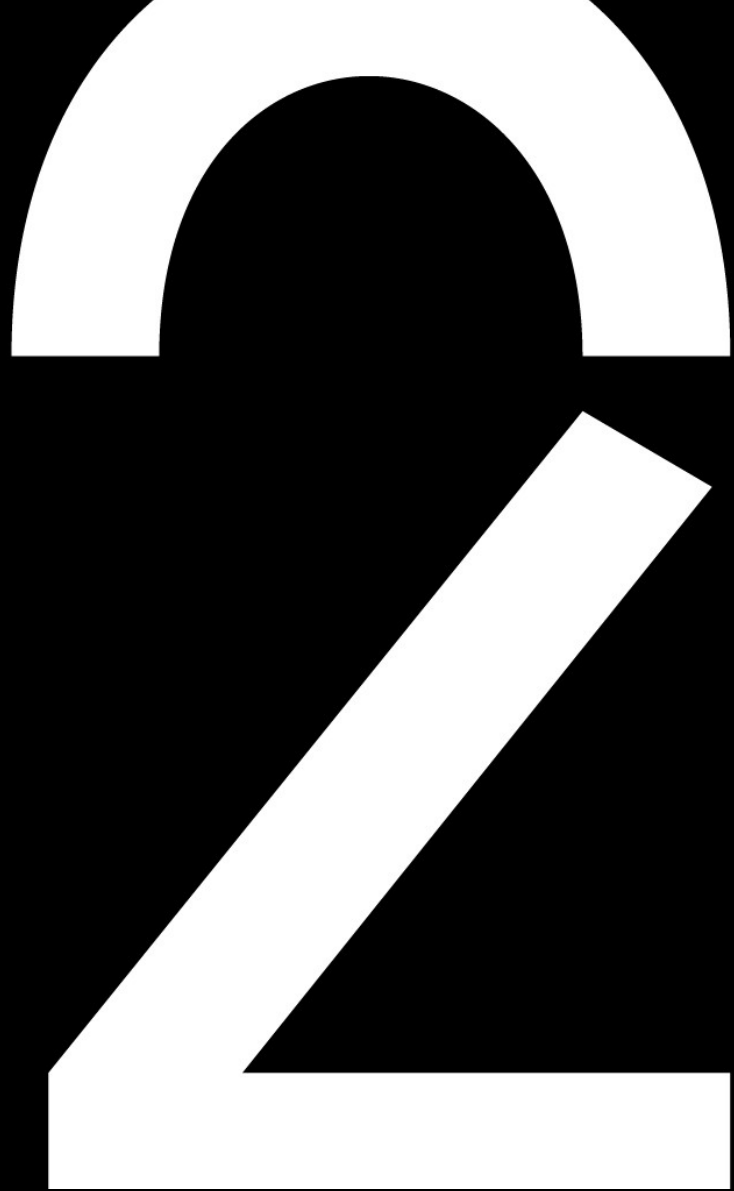
Voorbeeld: bankleningen

	pep	Tree
1	YES	YES
2	NO	NO
3	YES	YES
4	NO	NO
5	YES	YES
6	NO	NO
7	YES	YES
8	YES	NO
9	YES	YES
10	NO	YES
11	YES	YES
12	YES	NO
13	NO	YES
14	YES	NO
15	NO	NO
16	NO	NO
17	YES	NO
18	NO	NO
19	YES	NO
20	YES	NO

Resultaat

- de bekomen waarden zijn categorisch
- je kan dus geen “error” berekenen zoals bij regressie
- je kan wel tellen hoeveel keer de waarden overeenkomen en hoeveel keer niet

Confusion matrix



Confusion matrix

- Stel: er zijn 2 categorieën
- we noemen deze “positive” en “negative”
- er zijn 4 mogelijkheden
 - positive voorspeld, echte waarde is positive (True Positive)
 - positive voorspeld, echte waarde is negative (False Positive)
 - negative voorspeld, echte waarde is positive (False Negative)
 - negative voorspeld, echte waarde is negative (True Negative)

Confusion matrix

- We zetten deze in een matrix

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	true positives (TP)	false positive (FP)
	negatief	false negative (FN)	true negatives (TN)

Voorbeeld

- in foto's detecteren of er een gezicht in staat

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	876	54
	negatief	11	891

Meerdere categorieën

- als er meerdere categorieën zijn, dan heb je meer kolommen en rijen
- bv: herken dieren (kat, hond, konijn):

		Werkelijke klasse		
		kat	hond	konijn
Voorbeeld klasse	kat	echt kat	vals kat	vals kat
	hond	vals hond	echt hond	vals hond
	konijn	vals konijn	vals konijn	echt konijn

Accuracy

Accuracy

- we zoeken 1 getal dat uitdrukt hoe goed de voorspelling is
- accuracy is het aantal juiste voorspellingen gedeeld door het totaal aantal voorspellingen

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Voorbeeld

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	876	54
	negatief	11	891

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{876 + 891}{876 + 54 + 11 + 891} = 0,9645 = 96,45 \%$$

Meerdere categoriën

- Als er meer dan 2 categorieën zijn:
- bereken de accuracy per categorie (de categorie is dan “positive” en de anderen zijn “negative”)

		Werkelijke klasse	
		kat	niet-kat
Voorspelde klasse	kat	30	5+1
	niet-kat	10+0	20+2+7+25

Precision en recall



Probleem met accuracy

De test dataset bevatte heel veel “negative” waarden
Het algoritme heeft geleerd om alles negative te bestempelen...

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	0	0
	negatief	110	990

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{0 + 990}{0 + 990 + 0 + 110} = 0,9 = 90 \%$$

**vergelijk: als je niet weet of het “word” of “wordt” is, gok dan op “wordt”
want dat komt het meeste voor :-)**

Precision en recall

- men berekent 2 maten:

$$precision = \frac{TP}{TP + FP} \qquad recall = \frac{TP}{TP + FN}$$

- deze moeten beiden hoog zijn
- precision = hoeveel procent van de voorspelde positive waren ook daadwerkelijk positive?
- recall = hoeveel procent van de werkelijk positive werden ook als positive voorspeld?

Voorbeeld

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	876	54
	negatief	11	891

$$precision = \frac{TP}{TP + FP} = \frac{876}{876 + 54} = 94,19 \%$$

$$recall = \frac{TP}{TP + FN} = \frac{876}{876 + 11} = 98,76 \%$$

Voorbeeld

training set bevatte vooral negative

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	1	0
	negatief	110	990

$$precision = \frac{TP}{TP + FP} = \frac{1}{1 + 0} = 100\%$$

$$recall = \frac{TP}{TP + FN} = \frac{1}{1 + 110} = 0,9\%$$

Voorbeeld

training set bevatte vooral positive

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	110	990
	negatief	0	0

$$precision = \frac{TP}{TP + FP} = \frac{110}{110 + 990} = 10\%$$

$$recall = \frac{TP}{TP + FN} = \frac{110}{110 + 0} = 100\%$$

Meerdere categoriën

- Als er meer dan 2 categorieën zijn:
- bereken de precision en recall per categorie (de categorie is dan “positive” en de anderen zijn “negative”)

		Werkelijke klasse	
		kat	niet-kat
Voorspelde klasse	kat	30	5+1
	niet-kat	10+0	20+2+7+25

F-measure

Probleem

- precision en recall zijn 2 getallen
- ze moeten allebei zo hoog mogelijk zijn
- we willen maar 1 getal dat kan uitdrukken of ze beiden hoog zijn of niet

Oplossing

- neem het (gewogen harmonisch) gemiddelde van precision (P) en recall (R)

$$F_{\beta} = \frac{(\beta^2 + 1) \cdot P \cdot R}{\beta^2 \cdot P + R}$$

- als $\beta=1$ dan spelen beiden evenveel mee
- als $\beta>1$ dan wordt er meer belang gehecht aan recall
- als $\beta<1$ dan wordt er meer belang gehecht aan precision

Oplossing

- F_1 wordt veel gebruikt:

$$F_1 = \frac{2 \cdot P \cdot R}{P + R}$$

Voorbeeld

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	876	54
	negatief	11	891

$$precision = \frac{TP}{TP + FP} = \frac{876}{876 + 54} = 94,19 \%$$

$$recall = \frac{TP}{TP + FN} = \frac{876}{876 + 11} = 98,76 \%$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} = 96,42 \%$$

Voorbeeld

training set bevatte vooral negative

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	1	0
	negatief	110	990

$$precision = \frac{TP}{TP + FP} = \frac{1}{1 + 0} = 100\%$$

$$recall = \frac{TP}{TP + FN} = \frac{1}{1 + 110} = 0,9\%$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} = 1,79\%$$

Voorbeeld

training set bevatte vooral positive

		werkelijke klasse	
		positief	negatief
voorspelde klasse	positief	110	990
	negatief	0	0

$$precision = \frac{TP}{TP + FP} = \frac{110}{110 + 990} = 10\%$$

$$recall = \frac{TP}{TP + FN} = \frac{110}{110 + 0} = 100\%$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} = 18,18\%$$

Meerdere categoriën

- Als er meer dan 2 categorieën zijn:
- bereken de F-measure per categorie (de categorie is dan “positive” en de anderen zijn “negative”)

		Werkelijke klasse	
		kat	niet-kat
Voorspelde klasse	kat	30	5+1
	niet-kat	10+0	20+2+7+25

Andere maten



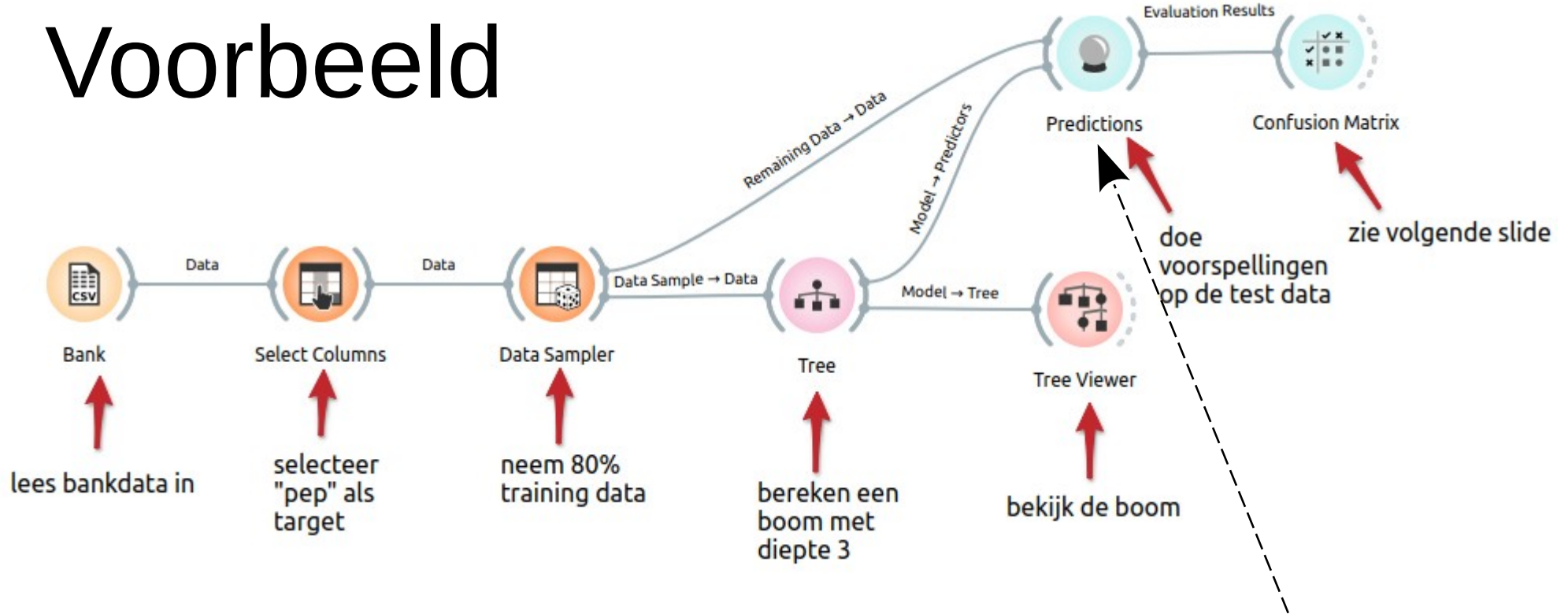
Andere maten

- er zijn nog andere maten die gebruikt worden:
 - False Positive Rate (FPR) = $FP / (TP + FN)$
 - True Positive Rate (TPR) = $TP / (TP + FN)$
 - dit is een andere benaming voor “recall”
 - sensitivity = $TP / (TP + FN)$
 - dit is een andere benaming voor “recall”
 - specificity = $TN / (TN + FP)$

The image features a solid black background. In the upper right corner, there is a white rectangular shape. A thick white diagonal line runs from the bottom right towards the top right. A horizontal white line is positioned in the middle left section of the image. Below this horizontal line, the word "Orange" is written in a white, serif font.

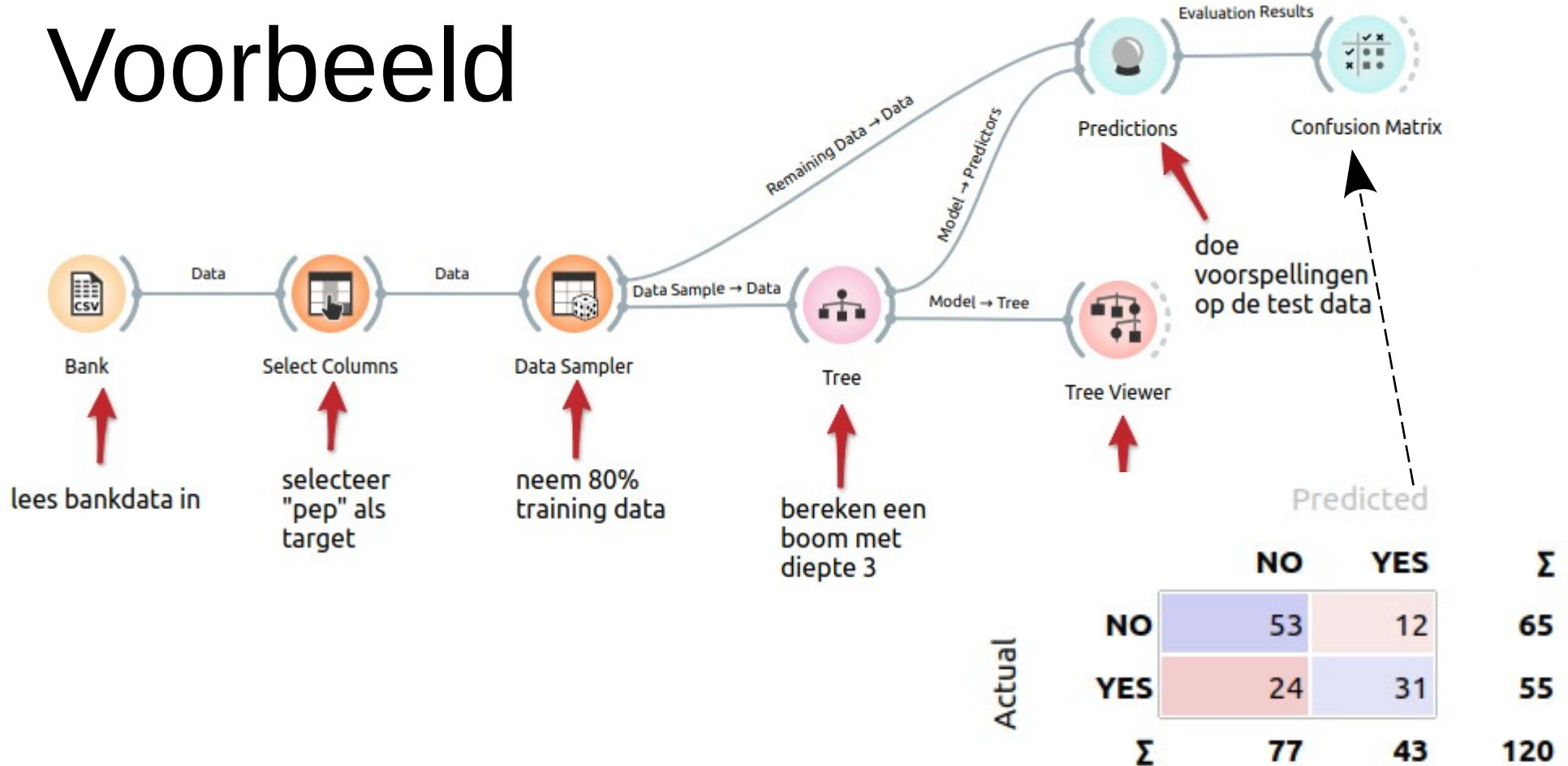
Orange

Voorbeeld

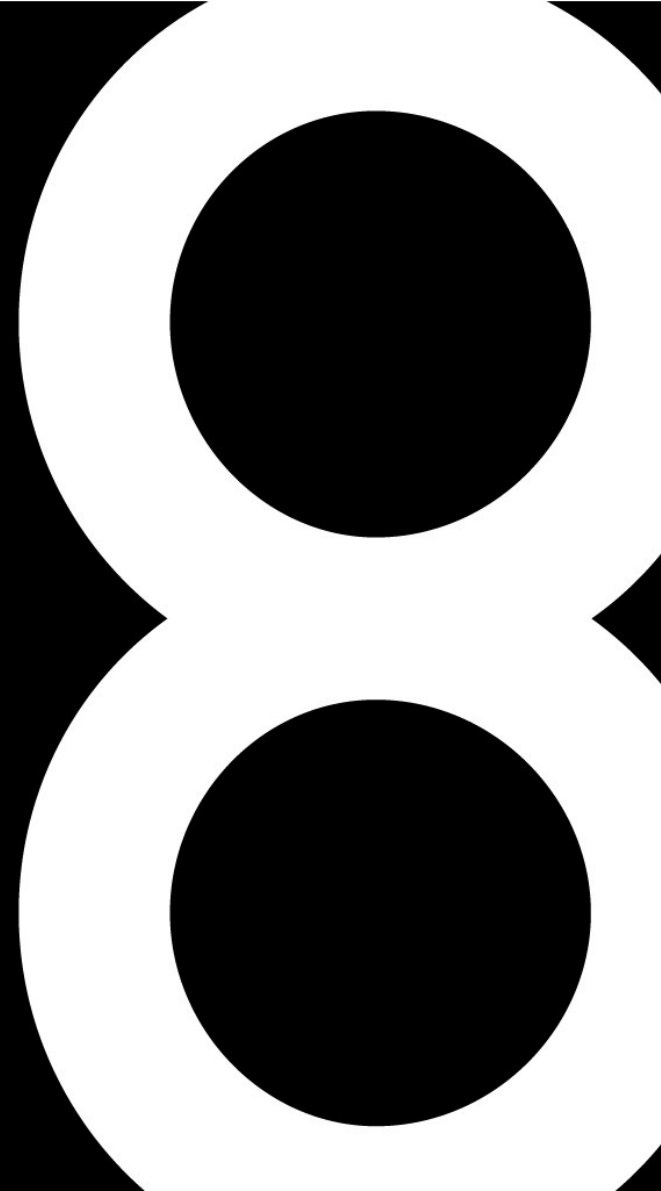


<input checked="" type="checkbox"/> Show performance scores							Target class: (Average over classes)
Model	AUC	CA	F1	Prec	Recall	MCC	
Tree	0.752	0.700	0.694	0.703	0.700	0.394	

Voorbeeld



Oefeningen



Oefeningen

- Zie Canvas
 - Foto's herkennen
 - Bank
 - Titanic
 - Music Genre Classification