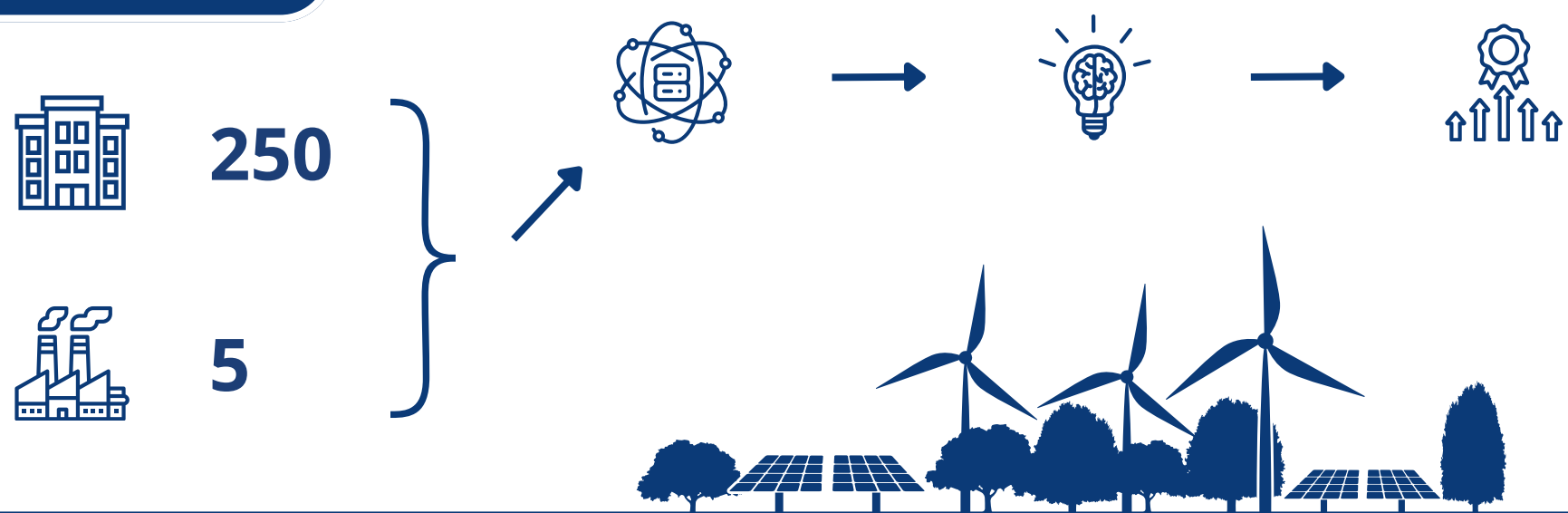


Van URL's naar Inzichten:

AI-gedreven Klantensegmentatie en Duurzaamheidsrangschikking

Situering



Onderzoeksdoelen

- 1 Herkennen van de **industrie**
- 2 Ranking op **duurzaamheid**
- 3 Relatie tussen **industrie** en **duurzaamheid**



Methodologie

67

36



25

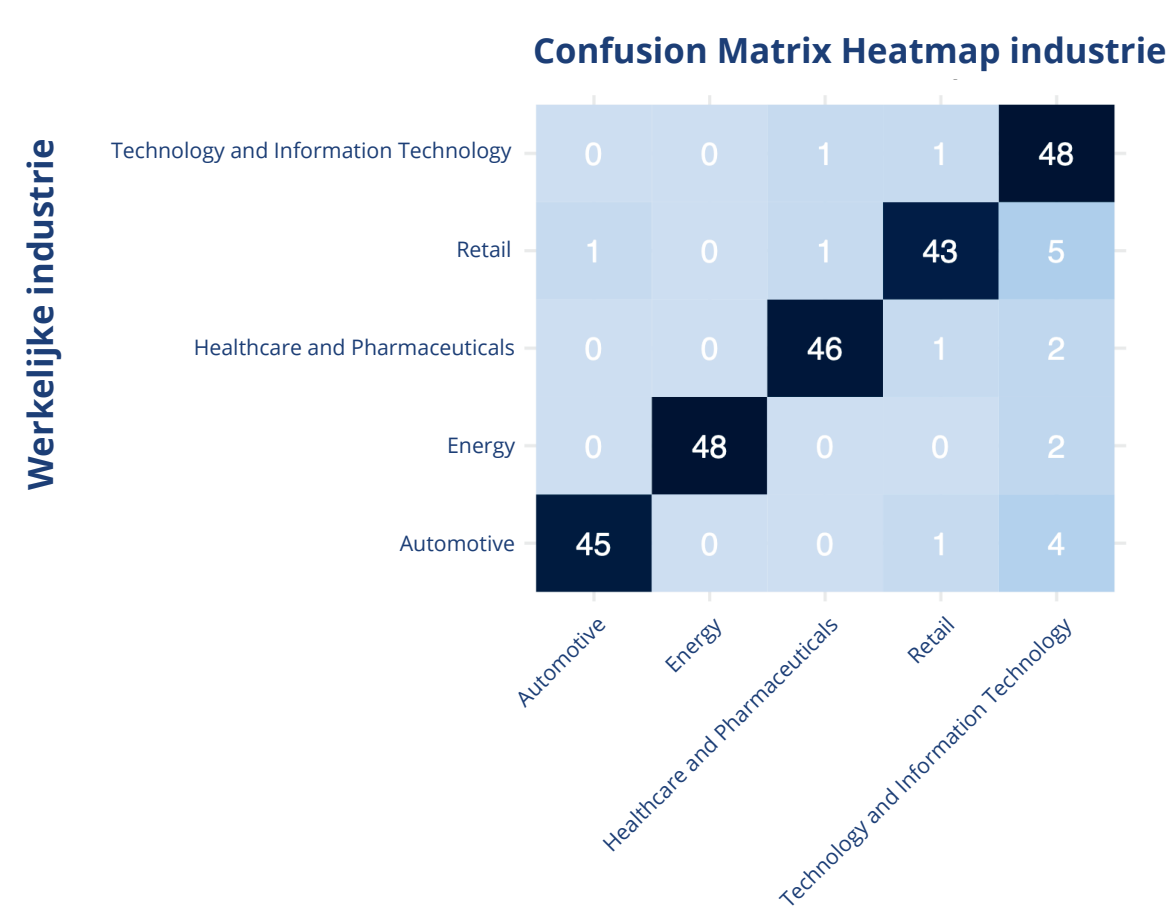
5

34

Resultaten

1

Herkennen van de industrie



Figuur 1: Confusion Matrix Heatmap industrie

Industrie	Precision	Recall	F1-score
Automotive	0,9783	0,9000	0,9375
Energy	1,0000	0,9600	0,9796
Healthcare and Pharmaceuticals	0,9583	0,9388	0,9485
Retail	0,9348	0,8600	0,8958
Technology and Information Technology	0,7869	0,9600	0,8649

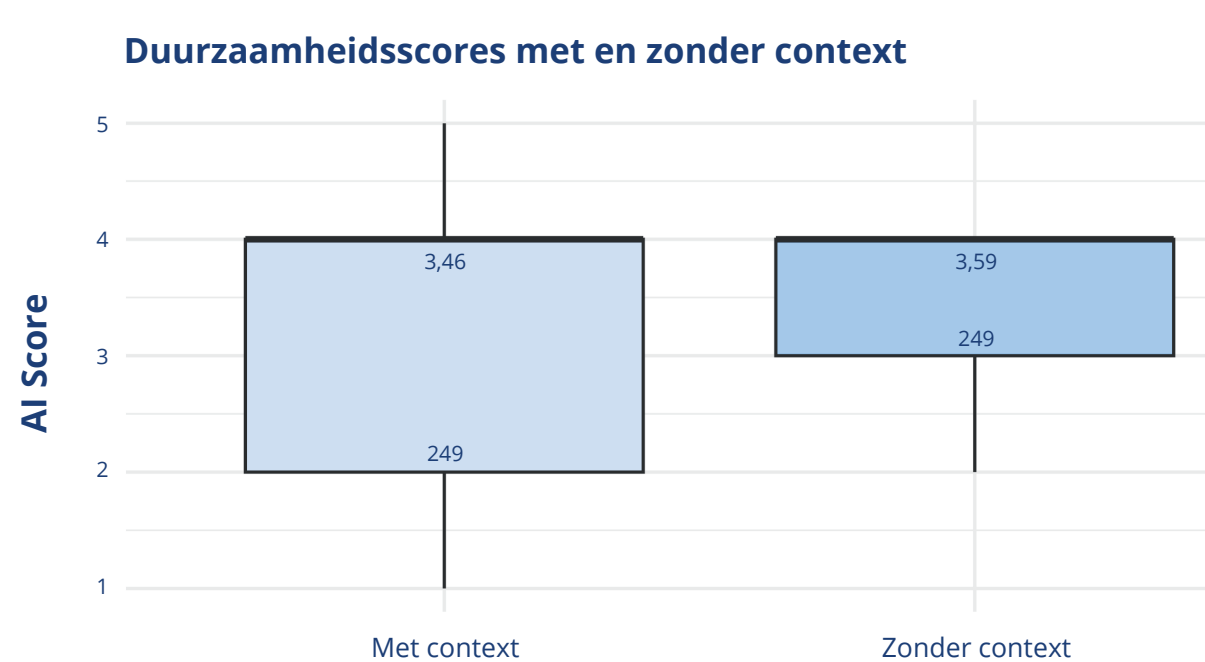
Tabel 2: Precision, Recall en F1-score per industrie

Algemene statistieken	Waarde
Accuracy	0,9237
95% CI	[0,8834;0,9534]
No Information Rate (NIR)	0,9583
P-Value [Acc > NIR]	0,9348
Kappa	0,7869

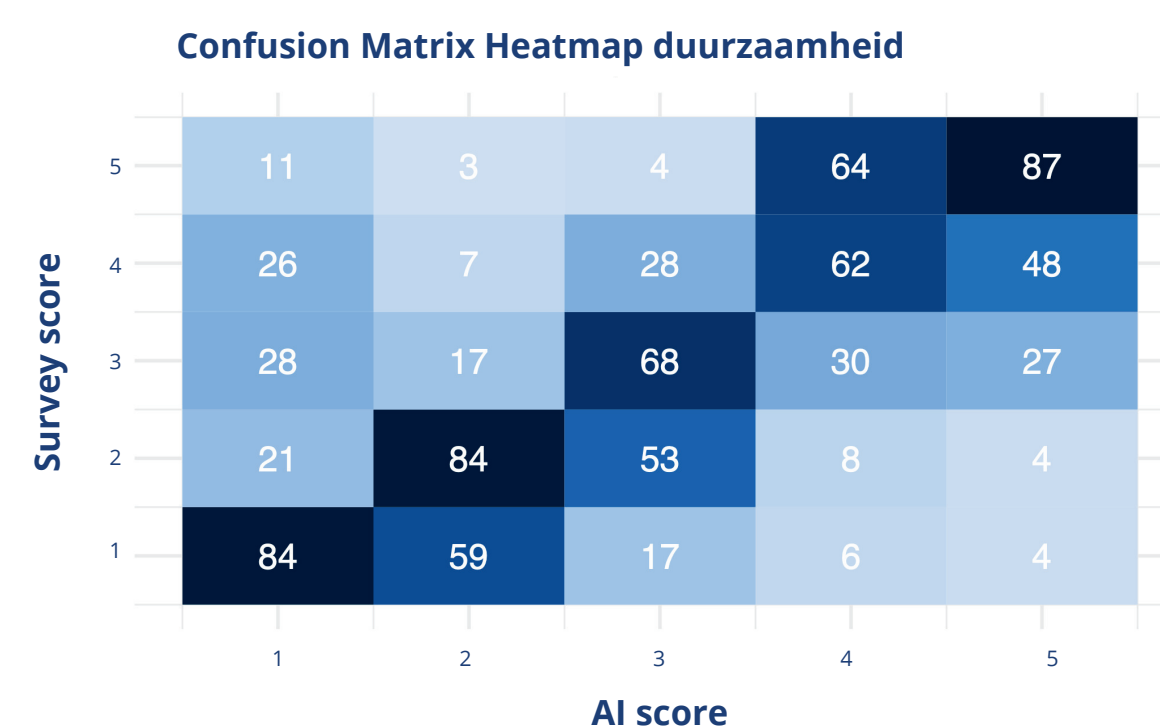
Tabel 3: Algemene statistieken industrie-toekenning

2

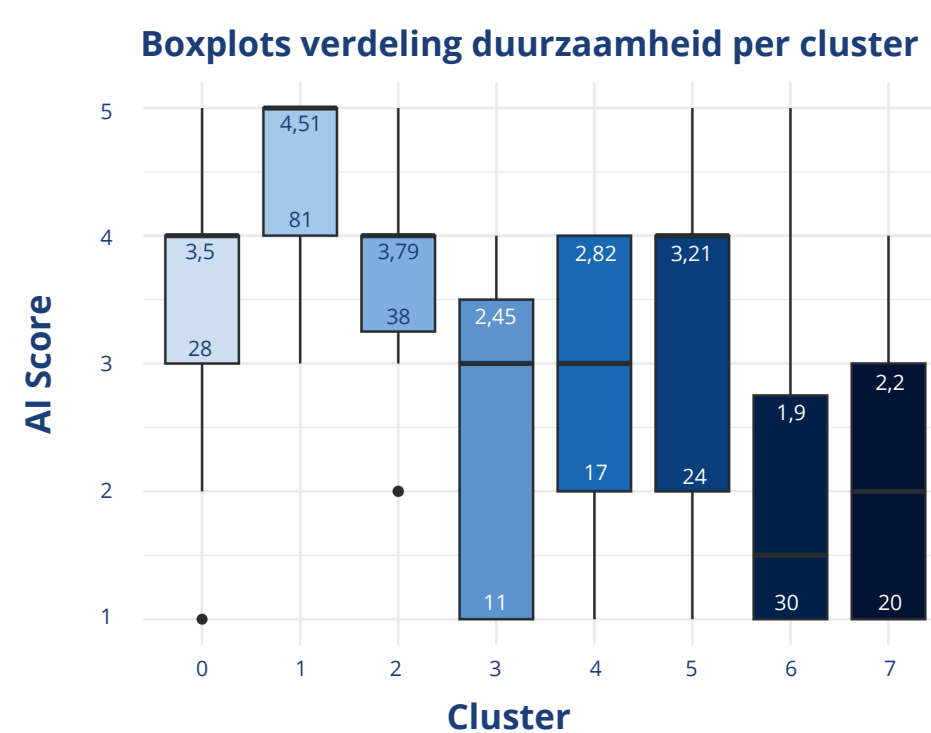
Ranking op duurzaamheid



Figuur 2: Boxplots duurzaamheidsscores met en zonder context



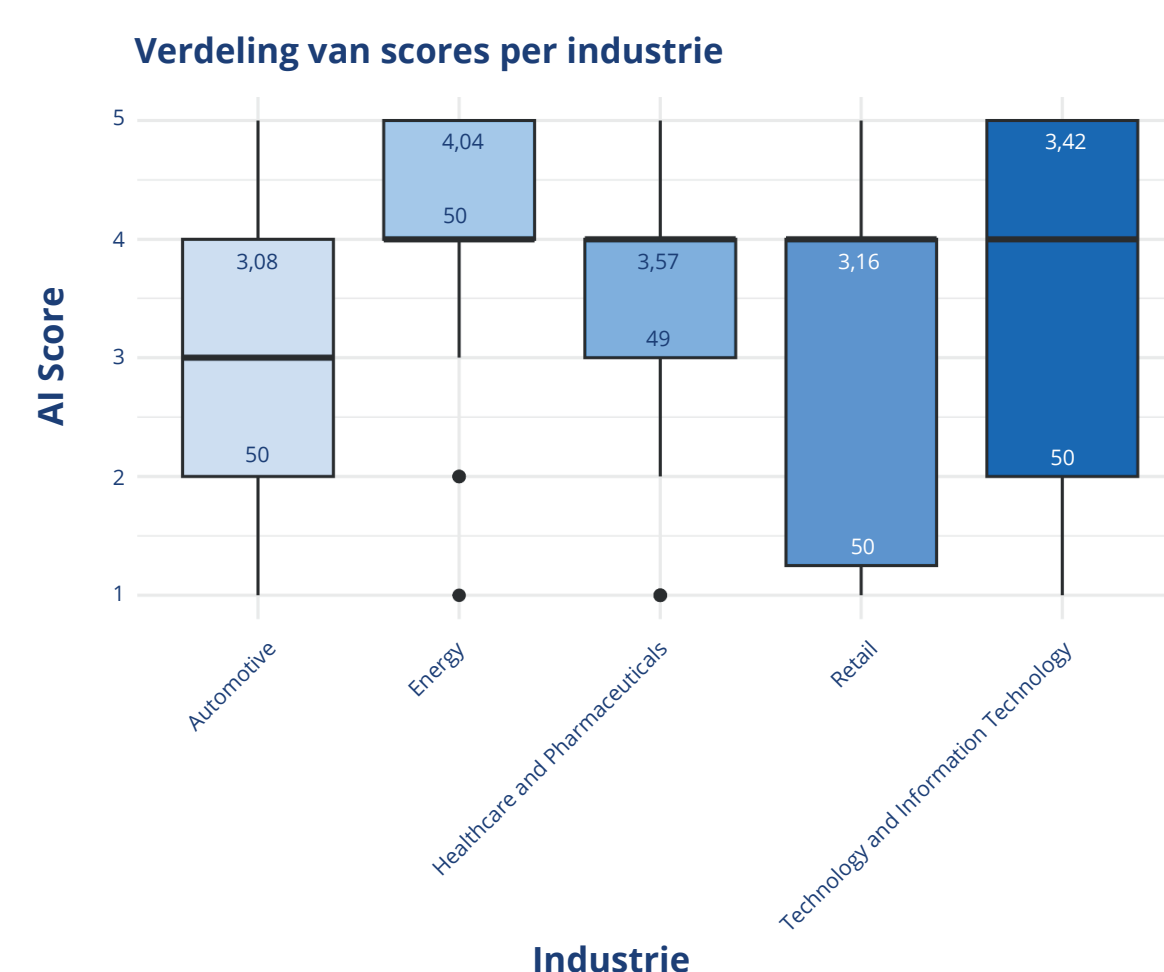
Figuur 3: Confusion Matrix Heatmap duurzaamheid



Figuur 4: Boxplots verdeling duurzaamheid per cluster

3

Relatie tussen industrie en duurzaamheid



Figuur 5: Confusion Matrix Heatmap industrie

Source	Df	Sum Sq	Mean Sq	F-value	Pr(>F)
Industrie	4	29,2	7,305	4,406	0,00186 **
Residuals	244	404,5	1,658		
Signif. codes	0'***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1				

Tabel 3: Precision, Recall en F1-score per industrie

Multiple R-Squared
0,06737

Tabel 4: Multiple R-Squared

Conclusie

- 1 Herkennen van de **industrie**
- 2 Ranking op **duurzaamheid**
- 3 Relatie tussen **industrie** en **duurzaamheid**

Accuracy 92,37%

88% afwijking ≤ 1

F-waarde 4,406

P-waarde $<2,2e-16$

MAE 0,61

P-waarde 0,00186

Correlatie 74,32%

Multiple R-Squared 6,74%