

Evaluatiemetrieën - Oefeningen

Vraag 1 – Accuracy, Precision, Recall, F-measure

In de tabel hieronder vind je de confusion matrix van een binaire classifier.

	Predicted YES	Predicted NO	
Actual YES	100	5	105
Actual No	10	50	60
	110	55	165

Beantwoord nu de volgende vragen:

- Wat zijn de waarden voor TP, TN, FP, FN?
- Bereken nu de accuracy, precision en recall.
- Bereken de F_1 en $F_{1.5}$ -measures.
- Wat is de TPR en FPR? Vergelijk met je recall. Wat stel je vast?

Schrijf functies in Python die bovenstaande metrieken kunnen berekenen uit een confusion matrix. Let op: dit is niet mogelijk met bestaande functies van `sklearn.metrics`, omdat deze werken op de ruwe data en niet op een reeds bestaande verwarringsmatrix.

- Maak van bovenstaande tabel een (confusion) matrix in Python.
- Schrijf een aparte functie voor elk metriek en controleer je oplossing. Hint: kijk even naar de volgende Python functie: `np.diag`.

Vraag 2 – Accuracy, Precision, Recall, F-measure

In de tabel hieronder vind je de confusion matrix van een binaire classifier.

	Predicted A	Predicted B	
Actual A	100	0	100
Actual B	50	5	55
	150	5	155

- Wat zijn de waarden voor TP, TN, FP, FN?
- Bereken met je zelfgeschreven Python-functies de accuracy, precision en recall, en F_1 .
- Is dit een goede classifier?

Vraag 3 – Accuracy, Precision, Recall, F-measure

In de tabel hieronder vind je de confusion matrix van een multiclass classifier.

		Voorspelde klasse					
		Asfalt	Beton	Gras	Boom	Gebouw	
Werkelijke klasse	Asfalt	2385	0	0	0	12	2397
	Beton	4	332	1	0	0	337
	Gras	0	0	908	0	0	908
	Boom	1	0	8	1084	6	1099
	Gebouw	4	1	0	9	2053	20679
		2394	333	917	1093	2071	6808

Beantwoord nu de volgende vragen:

- a) Wat zijn de waarden voor TP, FP, FN voor de klasse Gras?

Pas je functies voor alle evaluatiemetrieken aan zodat ze de waarde per klasse teruggeven.

- b) Bereken nu de accuracy, precision en recall per klasse met deze functies.
- c) Is dit een goede classifier?

Vraag 4 – ROC Curves – The Simpsons

Lees de data in het bestand "simpsons_roc1".

- Gebruik de plot_roc functie (zie slides) om de ROC-curve te tekenen.
- Wat is de beste threshold-waarde?
- Wat is de AUC? Wat betekent het?
- Pas de gevonden threshold-waarde (zie b.) toe op de kolom y_score (deze kolom bevat de resultaten van een neurale netwerk predict_proba-functie -zie later-) om een nieuwe voorspelde kolom te maken.
- Voer de stappen a. tot en met d. uit met de dataset in het bestand "simpsons_roc2". Merk je verschillen?

Vraag 5 – ROC Curves – Infertility

Twee lineaire discriminante analysemodellen werden gebouwd met de dataset-infert -zie bestand `infert.csv`- De resultaten van de voorspellingen zijn te vinden in de bestanden `education_roc1.csv` en `education_roc2.csv`.

- a. Laad beide datasets
- b. Gebruik de `plot_roc` functie (zie slides) om beide ROC-curven te tekenen.
Gebruik voor beide datasets kolom '6-11yrs' als y-score en transformeer de kolom 'y_true' in 1 (voor de waarden '6-11yrs' en een 0 voor de andere waarden).
- c. Wat is de beste threshold-waarde voor elk van de ROC-curves?
- d. Welk model heeft uw voorkeur?