

Sztuczna inteligencja

Ćwiczenia 2. Miary oceny klasyfikacji

Celem ćwiczenia jest zapoznanie z miarami oceny klasyfikacji danych.

Confusion matrix

Macierz prezentująca wyniki badanego klasyfikatora, określa ona ile obiektów i do jakich klas zostało przydzielonych. Na jej podstawie możemy oszacować poprawność badanego klasyfikatora.

		Klasy rzeczywiste	
		Kot	Pies
Klasy rozpoznane/ przewidziane	Kot	10	3
	Pies	2	8

Z tabeli możemy wyczytać:

- Klasyfikacja została dokonana na 23 obiektach
- Mamy 12 kotów i 11 psów
- 10 kotów i 8 psów zostało rozpoznanych prawidłowo
- 2 koty zostały rozpoznane jako psy, a 3 psy zostały rozpoznane jako koty

		correct result / classification	
		E1	E2
obtained result / classification	E1	tp (true positive)	fp (false positive)
	E2	fn (false negative)	tn (true negative)

Miary klasyfikacji

Accuracy – dokładność klasyfikatora. Stosunek ilości dobrze rozpoznanych obiektów do ilości wszystkich obiektów.

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn}$$

Precision(positive predictive value) - dokładność klasyfikacji w obrębie rozpoznanej klasy

$$PPV = TP / (TP + FP)$$

Recall (true positive rate, sensitivity, czułość) – mówi o tym na ile nasz klasyfikator potrafi rozpoznać obiekty z danej klasy

$$TPR = TP/P = TP/(TP + FN)$$

False positive rate (Fall-out)

$$FPR = FP/N = FP/(FP + TN)$$

Miara F (F-measure)

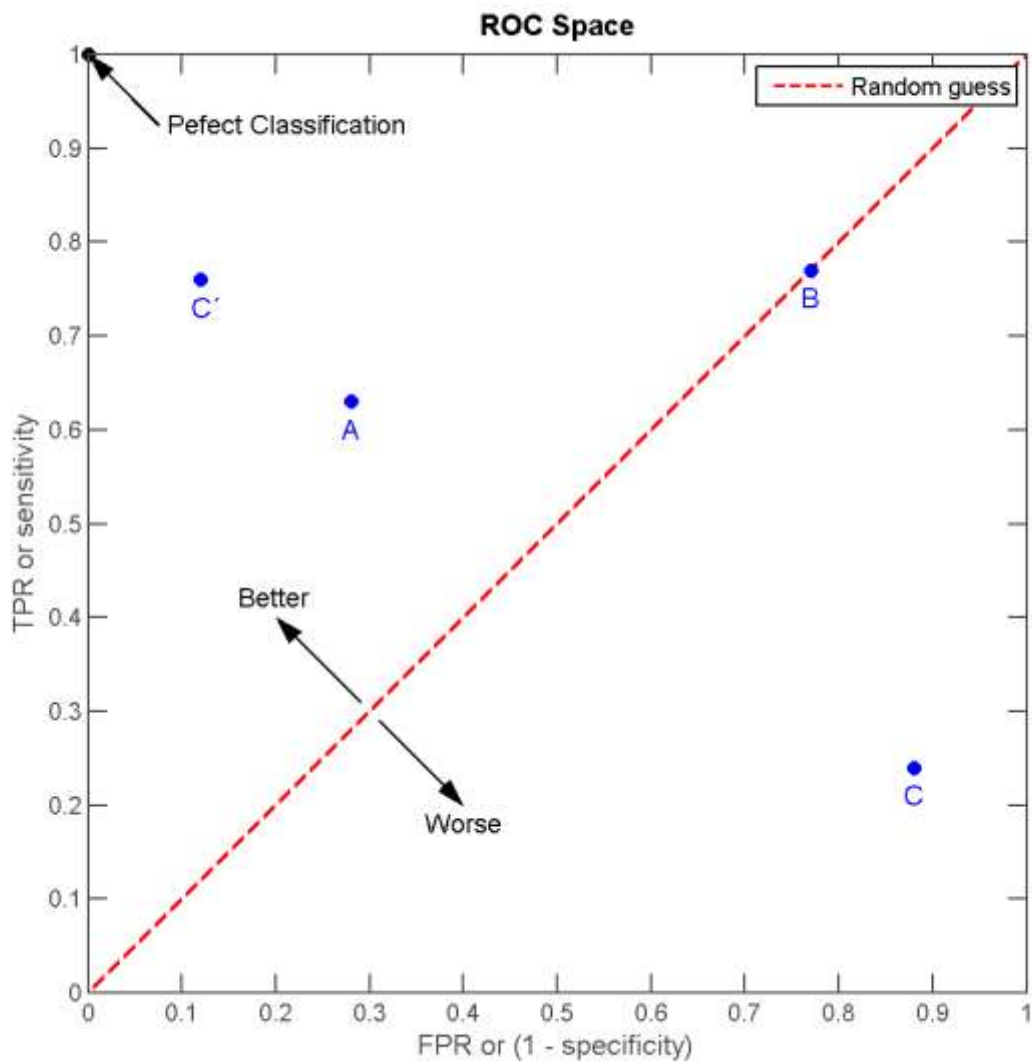
$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}}$$

Miara F1 (F1-measure, F1-score) – w powyższym wzorze $\beta=1$

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

Krzywe ROC

Aby utworzyć krzywą ROC na jednej osi odkładamy wartość recall(sensitivity, czułości), a na drugiej wartość $(1 - \text{specyficzność}) = FPR$. Krzywa ROC bada zatem stosunek wielkości TP do FP. Wysoka wrażliwość oznacza, że system prawidłowo odróżnia obiekty relewantne (np. prawidłowo rozpoznanie pacjentów chorych). Mała 1 - specyficzność mówi nam, że system niewiele obiektów nierelwantnych klasyfikuje jako relewantne (niewielu pacjentów zdrowych uznaje za chorych).



Każdy punkt na krzywej ma współrzędne (FPR, recall) lub (1-specyficzność, czułość). Krzywe ROC opisują zachowanie się danego modelu klasyfikacyjnego przy zmieniającej się wartości odcięcia klasyfikacji.

Pole pod krzywą ROC (AUC- Area Under Curve) jest miarą jakości danej metody, pole 0.5 oznacza najgorszą klasyfikację, zaś pole=1 najlepszą.

Zad1. Dla danych z tabel a) i b) policz miary oceny klasyfikacji: accuracy, precision, recall, F1. Który z klasyfikatorów działa lepiej?

a)

		Klasy rzeczywiste	
		Kot	Pies

Klasy rozpoznane/ przewidziane	Kot	140	20
	Pies	60	22

b)

		Klasy rzeczywiste	
		Kot	Pies
Klasy rozpoznane/ przewidziane	Kot	100	2
	Pies	100	40

Zad 2.

Dokonując klasyfikacji 27 zwierząt, 8 kotów, 6 psów i 13 królików, klasyfikator zwrócił następujące wyniki:

		Klasy przewidziane		
		Kot	Pies	Królik
Klasy rzeczywiste	Kot	5	3	0
	Pies	2	3	1
	Królik	0	2	11

Zad 3.

Zapoznaj się z dokumentacją funkcji „roc” i „plotroc” w środowisku matlab (dostępne w pakiecie Neural Network Toolbox). Wykreśl krzywą roc dla przykładowych danych

origLabel=y=[1 0 1 1 1 0 1 0 1 0 1 1 0 0]

output=[1 0.1 1.2 0 0.4 0.6 1.3 1 0.9 0.22 0.9 0.7 0.1 0.3]

Zad 4.

Zapoznaj się z zbiorem danych Iris z bazy UCI <http://archive.ics.uci.edu/ml/datasets/Iris>.

Wykonaj klasyfikację zbioru danych „Iris” i wykreśl krzywe dla każdej klasy ROC.

```
load iris_dataset
net = newpr(irisInputs,irisTargets,20);
net = train(net,irisInputs,irisTargets);
irisOutputs = sim(net,irisInputs);
[tpr,fpr,thresholds] = roc(irisTargets,irisOutputs)
plotroc(irisTargets,irisOutputs)
```

Oceń dla których klasy klasyfikacja jest najlepsza.