

Classification analysis/Analiza klasyfikacji

Discriminant Analysis/Analiza dyskryminacyjna

Pattern Recognition/Rozpoznawanie wzorów

Supervised Learning/Uczenie maszynowe pod nadzorem

Próba $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \dots, (X_{N-1}, Y_{N-1}), (X_N, Y_N)$ z K populacji



Learning sample/Próba ucząca

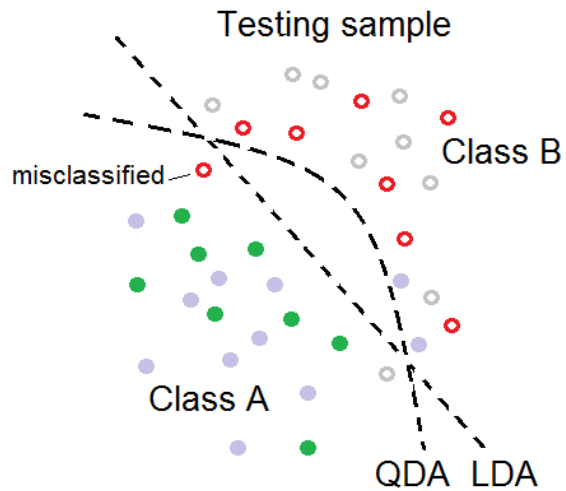
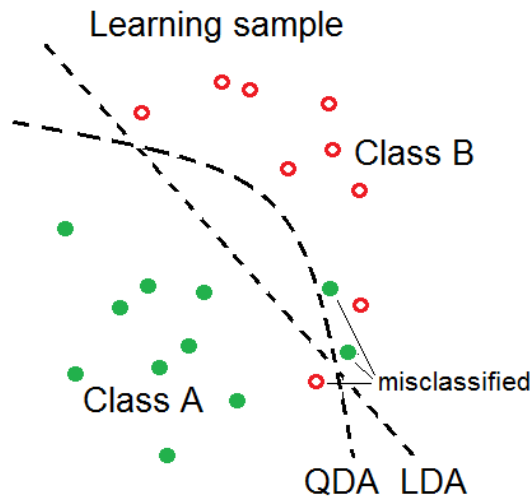
$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$

(na jej podstawie budujemy model)

Testing sample/Próba testowa

$(X_{n+1}, Y_{n+1}), (X_{n+2}, Y_{n+2}), \dots, (X_N, Y_N)$

(jeśli Y_{n+1}, \dots, Y_N są nieznane, to predykcja,
jeśli Y_{n+1}, \dots, Y_N są znane, to weryfikacja modelu)



QDA – z kwadratową funkcją klasyfikacyjną (klasyfikatorem), LDA - z liniową funkcją klasyfikacyjną

	Sklassyfikowane jako A	Sklassyfikowane jako B
Z populacji A	8	2
Z populacji B	1	9

	Sklassyfikowane jako A	Sklassyfikowane jako B
Z populacji A	7	0
Z populacji B	1	8

Estimacja błędu klasyfikacji dla próby uczącej

$$\hat{e}(\hat{d}) = \frac{3}{20} = 0,15$$

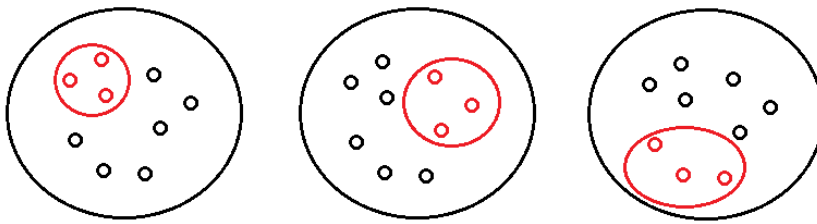
Estimacja błędu klasyfikacji dla próby testowej

$$\hat{e}(\hat{d}) = \frac{1}{16} = 0,0625$$

Błąd resubstytucji – ponownego podstawienia (testing sample= learning sample)

$$\hat{e}_R = 0,15$$

Cross-validation/sprawdzanie krzyżowe



Learning sample

Testing sample

Bootstrap – klonowanie próbki

Próbka (3,4,5)

$B = 10$ (10 spośród 27 możliwych wariacji z powtórzeniami)

Próby bootstrapowe (losowanie z powtórzeniami 3 elementów z próby pierwotnej)

(3,3,5), (4,5,5), (3,3,4), (3,4,5), (4,4,4), (3,4,5), (3,3,5), (3,4,4), (3,4,5), (5,5,5)