

Ćwiczenia 9

1. (S) Używając zbioru `painters` (subiektywne oceny malarzy) z pakietu `MASS` skonstruuj model LDA (klasy stanowią różne szkoły malarstwa). Ile wynosi błąd ponownego podstawiania i błąd CV LOO. Wyświetl i zinterpretuj macierz pomyłek. Czy skonstruowany model jest lepszy od modelu losowego?
2. Wykonaj liniową oraz kwadratową analizę dyskryminacyjną w celu rozróżnienia miejsca pochodzenia liści w zbiorze danych `leafshape` (szerokość, długość oraz ogonek liści z różnych miejsc Australii) z pakietu `DAAG`. Która z metod daje mniejszy błąd ponownego podstawiania, a która CV LOO?
3. (S) Przeprowadzono badanie mające na celu sprawdzenie wpływu stosowanej terapii psychologicznej na zredukowanie negatywnych efektów dojrzewania. W tym celu wybrano trzy grupy po 5 osób. Pierwsza grupa została poddana terapii rodzinnej, druga terapii indywidualnej, a trzecia nie otrzymała żadnej pomocy. Po sześciu miesiącach zbadano poziom depresji, niepokoju oraz chaosu społecznego. Otrzymano następujące wyniki (im większa wartość tym wyższy poziom cechy): Grupa 1: (6, 8, 9), (4, 3, 3), (0, 2, 8), (4, 1, 6), (0, 8, 4); Grupa 2: (11, 9, 8), (11, 6, 6), (5, 7, 4), (8, 6, 5), (4, 9, 4); Grupa 3: (12, 11, 6), (8, 8, 5), (9, 6, 7), (8, 10, 8), (11, 4, 3). Skonstruuj model LDA, QDA oraz naiwny model bayesowski (w wersji z rozkładami normalnymi). Która metoda daje mniejszy błąd CV LOO?
4. Załaduj dane:
`https://hastie.su.domains/ElemStatLearn/datasets/SAheart.data`
(choroby serca w RPA). Wyznacz błąd resubstytucji oraz błąd bootstrappowy (100 powtórzeń) dla metody 1-NN oraz RF (dwie cechy w każdym węźle). W przypadku modelu RF wyświetl macierz pomyłek i zinterpretuj. Który model można polecić dla tego problemu?