

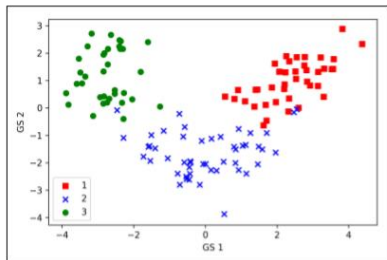
PROJEKT GRUPOWY

MACHINE LEARNING

1. Wstępne przetwarzania:
 - a. Zbiór danych adults.tab
 - b. Dla zadanego zbioru danych wykonaj następujące czynności:
 - i. Oczyszczyć zbiór danych i wprowadzić go do tabelki bazodanowej
 - ii. Podzielić losowo zbiór na część treningową i część walidacyjną
 - iii. Przygotuj zbiór do procesu uczenia maszynowego
2. Na podstawie poniższych algorytmów wykonaj
 - a. Trenowanie modeli:
 - i. SVM
 - ii. kNN
 - iii. Decision Tree
 - iv. Random Forest
 - b. Dokonaj wyboru trzech algorytmów z optymalnymi ustawieniami na podstawie kros-walidacji (CV) dla 5 prób
 - c. Wypisz standardowe wskaźniki ocen procesu trenowania:
 - i. Macierz konfuzji
 - ii. Dokładność klasyfikacji
 - iii. Członość
 - iv. Specyficzność
 - v. Pole pod krzywą AUC
 - vi. F1
 - vii. Precyzja
 - d. Wykonaj testowanie modelu na podstawie zbioru walidacyjnego, który nie brał udziału w procesie trenowania.
 - e. Wypisz standardowe wskaźniki ocen procesu testowania:
 - i. Macierz konfuzji
 - ii. Dokładność klasyfikacji
 - iii. Członość
 - iv. Specyficzność
 - v. Pole pod krzywą AUC
 - vi. F1
 - vii. Precyzja

3. Zaprezentuj graficznie wynik klasyfikacji i testowania:

- a. Wykres typu scatterplot
- b. Etykieta klasy -> unikatowy kolor i kształt



- c. Wykres powinien zawierać dla danego klasyfikatora następujące podwykresy:
 - i. Rozkład próbek treningowych / testowych z podziałem na klasy
 - ii. Wynik trenowania modelu z podziałem na klasy
 - iii. Wynik testowania modelu z podziałem na klasy

4. Wykonaj uczenie zespołowe na 3 wytypowanych klasyfikatorach

- a. Ocena jest składową 3 ocen a decyzja o przynależności klasy jest podejmowana na podstawie większości głosów
- b. Wykonaj trenowanie i testowanie modelu uczenia zespołowego
- c. Przygotuj model do przeprowadzenia testowania na podstawie danych w formacie .tab
- d. Wynik zapisz do pliku txt w postaci wektora ocen.