

DRZEWA DECYZYJNE

LAB 4 – Uczenie z nadzorem

Zadanie 1. (1 pkt.)

Zbiór danych: <https://archive.ics.uci.edu/ml/datasets/Smartphone-Based+Recognition+of+Human+Activities+and+Postural+Transitions>

- Treningowy: X_train.txt / y_train.txt
- Testowy: X_test.txt / y_test.txt

Zadanie 2. (2 pkt.)

Zbuduj modele predykcyjne na podstawie następujących algorytmów zakładając domyślne parametry:

- SVM
- kNN
- Decision Tree
- Random Forest

Zadanie 3. (1 pkt.)

Na podstawie zdefiniowanych metryk oceny klasyfikatorów oceń skuteczność klasyfikacji algorytmów z zadania 2

- Confusion matrix: TP / TN / FP / FN
- ACC
- Recall
- F1
- AUC

Zadanie 3. (2 pkt.)

Dokonaj wyboru najlepszego algorytmu klasyfikacji na podstawie kros-walidacji (CV) dla 5 prób. Jako rezultat zwróć następujące parametry:

- Wartości średniej z wyników klasyfikacji
- Średniego odchylenia standardowego z wyników klasyfikacji

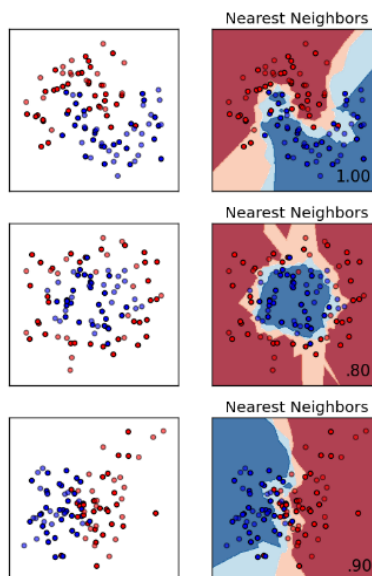
Zadanie 4. (2 pkt.)

Znajdź optymalne wartości parametrów klasyfikacji (możesz korzystać z dowolnych technik) dla wykorzystywanych klasyfikatorów tak aby skuteczność trenowania była maksymalna. Następnie na podstawie wyselekcjonowanego algorytmu wraz ze zdefiniowanymi parametrami zwróć wynik testowania algorytmu na zbiorze testowym.

Zadanie 5. (2 pkt.)

Zaprezentuj graficznie wynik klasyfikacji i testowania wyselekcjonowanego algorytmu:

- Wykres typu scatterplot
- Etykieta klasy -> unikatowy kolor i kształt
- Wykres powinien zawierać dla danego klasyfikatora następujące podwykresy:
 - Rozkład próbek treningowych / testowych z podziałem na klasy
 - Wynik trenowania modelu z podziałem na klasy
 - Wynik testowania modelu z podziałem na klasy



poglądowe wykresy