

# Laboratorium: SVM

March 24, 2022

## 1 Cel/Zakres

- Klasyfikacja SVM.
- Skalowanie.
- Budowanie potoków.

## 2 Przygotowanie danych

Użyj testowego zbioru danych do klasyfikacji występowania raka piersi.

```
from sklearn import datasets

data = datasets.load_breast_cancer()
print(data['DESCR'])
```

## 3 Klasyfikacja

1. Podziel zbiór danych na uczący i testujący w proporcjach 80/20.
2. Wykonaj klasyfikację SVM dla cech: area, smoothness tj.:
  1. LinearSVC, z funkcją straty “hinge”,
  2. LinearSVC, z funkcją straty “hinge”, po uprzednim automatycznym skalowaniu wartości cech.
3. Policz dokładność (accuracy) dla w/w klasyfikacji osobno na zbiorze uczącym i testującym, zapisz wartości na liście w kolejności: zbiór uczący bez skalowania, zbiór testujący bez skalowania, zbiór uczący ze skalowaniem, zbiór testujący ze skalowaniem. W.w. listę zapisz w pliku Pickle `bc_acc.pkl`.

4 pkt.

4. Czy skalowanie coś dało?
5. Ekperyment powtórz dla poniższego zbioru danych dla cech: długość i szerokość płatk.
6. Policz dokładność (accuracy) dla w/w klasyfikacji osobno na zbiorze uczącym i testującym, zapisz wartości na liście w kolejności: zbiór uczący bez skalowania, zbiór testujący bez skalowania, zbiór uczący ze skalowaniem, zbiór testujący ze skalowaniem. W.w. listę zapisz w pliku Pickle `iris_acc.pkl`.

4 pkt.

7. Czy skalowanie coś dało?

```
data = datasets.load_iris()
print(data['DESCR'])
```

## 4 Prześlij raport

Prześlij plik o nazwie `lab4.py` realizujący ww. ćwiczenia.

Sprawdzone będzie, czy skrypt Pythona tworzy wszystkie wymagane pliki oraz czy ich zawartość jest poprawna.