

Laboratorium: SVM

March 24, 2022

1 Cel/Zakres

- Klasyfikacja SVM.
- Skalowanie.
- Budowanie potoków.

2 Przygotowanie danych

Załaduj zbiory danych, które będą używane w klasyfikacji.

```
from sklearn import datasets
```

Pierwszy zbiór zawiera dane obrazów przypadków nowotworów piersi:

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

Drugi zawiera „klasyczny” zbiór parametrów irysów:

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

Podpowiedź: funkcje `load_...` domyślnie zwracają obiekty `numpy`, ale jeżeli prześlemy im argument `as_frame=True`, elementy `data` oraz `target` będą strukturami `pandas`, a dodatkowo dostępny będzie element `frame`, który zawiera `data` połączone z `target`.

3 Klasyfikacja

1. Podziel zbiór danych na uczący i testujący w proporcjach 80/20.
2. Zbuduj modele klasyfikacji SVM dla średnich (*mean*) wartości cech `area` oraz `smoothness`; stwórz dwa modele:
 1. `LinearSVC`, z funkcją straty “hinge”,
 2. `LinearSVC`, z funkcją straty “hinge”, po uprzednim automatycznym skalowaniu wartości cech.
3. Policz dokładność (*accuracy*) dla ww. 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. Listę zapisz w pliku `Pickle bc_acc.pkl`.

4 pkt.

4. Czy skalowanie coś dało?
5. Ekperyment powtórz dla zbioru irysów; zbuduj model wykrywający, czy dany przypadek jest gatunku *Virginica* na podstawie cech: długość i szerokość płatka.
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?

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.