

LABORATORIUM 5

Redukowanie wymiarowości

Zestawy danych

<https://data.gov/> to portal otwartych danych rządu Stanów Zjednoczonych.

<https://www.ncdc.noaa.gov/cag/> - zapewnia portal klimatyczny NOAA

<https://www.esrl.noaa.gov/psd/data/timeseries/> - Portal badań Ziemi administracja laboratoriów (ESRL, Earth System Research Laboratory) NOAA dostarcza miesięczne i sezonowe z danymi klimatycznymi.

<https://www.quandl.com/search> - Quandl udostępnia setki bezpłatnych szeregów czasowych z danymi finansowymi, a także zestawy danych z płatnym dostępem.

<https://datamarket.com/data/list/?q=provider:tsdl> - **biblioteka danych TSDL** (biblioteka danych szeregów czasowych) zawiera linki do setek tymczasowych zbiorów danych rządu w wielu obszarach przemysłowych.

<http://archive.ics.uci.edu/ml/datasets.html> - repozytorium uczenia maszynowego Uniwersytetu Kalifornijskiego w Irvine (UCI) zawiera dziesiątki zestawów danych szeregów czasowych z różnych obszarów.

ZADANIE

1. Należy napisać program (za pomocą biblioteki **scikit-learn**), który wykonuje redukowanie wymiarowości zestawu danych **za pomocą wyodrębniania i wyboru cech**
Sprawdzić, jak wpłynęły powyższe sposoby na wyniki algorytmów klasyfikacji (otrzymany w laboratorium 3-4):
 - regresji logistycznej
 - maszyny wektorów nośnych
 - jądra SVM
 - Uczenia drzew decyzyjnych
 - Algorytm k-najbliższych sąsiadów

Pracujemy z zestawem danych, który Państwo wybrali na laboratorium 4 i sprawdzamy na nim algorytmy klasyfikacji.

Na ocenę 3 proszę o wykonanie redukowania wymiarowości zestawu danych za pomocą wyodrębniania i wyboru cech i zaprogramowanie dowolnego (jednego!) algorytmu lub powtórzyć kod z wykładu.

na 4 proszę o wykonanie redukowania wymiarowości zestawu danych za pomocą wyodrębniania i wyboru cech i zaprogramowanie dowolne dwa algorytmy(metodu)

na 5 proszę o wykonanie redukcji wymiarowości zestawu danych za pomocą wyodrębniania i wyboru cech i zaprogramowanie dowolnych trzech algorytmów (metod)

Zrób podsumowanie: jak zmieniły się wyniki w klasyfikacji, dlaczego