

## Metody i narzędzia BIG DATA

### Lista 5

Zad. 1 Rozwiąż zadanie redukcji wymiaru z wykorzystaniem metody PCA (ang. *principal component analysis*) dla prostych zbiorów danych dostępnych m.in. w pakiecie *sci-kit learn*:

- a) *Iris dataset*
- b) *Wine dataset*
- c) Inny, wybrany samodzielnie

Zadania:

1. Dokonaj standaryzacji danych.
2. Dokonaj redukcji każdego ze zbiorów do dwóch wymiarów. Przedstaw w formie graficznej uzyskany zbiór.

Zad. 2 Rozwiąż zadanie redukcji wymiaru z wykorzystaniem metody PCA (ang. *principal component analysis*) dla rzeczywistych zbiorów danych dostępnych m.in. w pakiecie *sci-kit learn*:

- a) *California Housing dataset*
- b) Inny, wybrany samodzielnie np. z bazy [openml.org](http://openml.org).

Zadania:

1. Dokonaj standaryzacji danych.
2. Dokonaj redukcji każdego ze zbiorów do dwóch wymiarów. Przedstaw w formie graficznej uzyskany zbiór.

Zad. 3 Rozwiąż zadanie redukcji wymiaru z wykorzystaniem metody LDA (ang. *linear discriminant analysis*) z wykorzystaniem zbiorów takich jak w zadaniu 1 i 2.

Zad. 4 Rozwiąż zadanie redukcji wymiaru z wykorzystaniem metody SVD (ang. *singular value decomposition*) z wykorzystaniem zbiorów takich jak w zadaniu 1 i 2.