

# PAD 04

## Praca Domowa 4 – Numpy

### Zadanie 1 (4 pkt)

Wgraj zawartość pliku Zadanie\_1.csv. Stwórz tablice w numpy w oparciu o wczytane dane.

Podaj:

- Ile komórek liczy sobie macierz
- Ile ma wierszy a ile kolumn
- Oblicz następujące statystyki opisowe dla macierzy: średnia/mediana/wariancja
- Powtórz czynność z podpunktu b), ale tym razem usuń z macierzy braki danych.

### Zadanie 2 (2 pkt)

Wgraj dane z pliku Zadanie\_2.csv.

- Znajdź wektory własne, oraz wartości własne dla zawartej w pliku macierzy
- Oblicz macierz odwrotną dla macierzy z pliku

### Zadanie 3 (4 pkt)

Wgraj dane z pliku Zadanie\_3.csv

Wykorzystaj następujący wzór:

$$\frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}},$$

Gdzie A i B to wektory reprezentujące kolejne kolumny.

Aby stworzyć macierz podobieństwa w oparciu o macierz z pliku.

$$A(3,3) = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$
$$B(5,3) = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \\ b_{51} & b_{52} & b_{53} \end{bmatrix}$$

WSKAZÓWKA:

$$\text{Cosine Similarity Matrix (3,5)} = \begin{vmatrix} \frac{\vec{a}_1 \vec{b}_1}{\|\vec{a}_1\| \|\vec{b}_1\|} & \frac{\vec{a}_1 \vec{b}_2}{\|\vec{a}_1\| \|\vec{b}_2\|} & \dots & \frac{\vec{a}_1 \vec{b}_5}{\|\vec{a}_1\| \|\vec{b}_5\|} \\ \frac{\vec{a}_2 \vec{b}_1}{\|\vec{a}_2\| \|\vec{b}_1\|} & \frac{\vec{a}_2 \vec{b}_2}{\|\vec{a}_2\| \|\vec{b}_2\|} & \dots & \frac{\vec{a}_2 \vec{b}_5}{\|\vec{a}_2\| \|\vec{b}_5\|} \\ \frac{\vec{a}_3 \vec{b}_1}{\|\vec{a}_3\| \|\vec{b}_1\|} & \frac{\vec{a}_3 \vec{b}_2}{\|\vec{a}_3\| \|\vec{b}_2\|} & \dots & \frac{\vec{a}_3 \vec{b}_5}{\|\vec{a}_3\| \|\vec{b}_5\|} \end{vmatrix}$$

#### Zadanie 4\* (Nieobowiązkowe – 3 punkty dodatkowe)

W pliku Zadanie\_4.csv zawarte są dane o stanie dostępności gabinetu lekarskiego w danej minucie.

Napisz program, który przekształci dane w taki sposób, żeby w jednej linijce zapisany był początek i koniec przedziału czasowego, kiedy lekarz jest dostępny

Dane zapisz do pliku CSV.

*Podpowiedź: Wykorzystaj do tego zadania bibliotekę pandas i metodę groupBy()*