

Zadania dzień 2 – NumPy

Piotr Pasza Storożenko

1. Niech \mathbf{x} będzie wektorem.

- Wypisz wszystkie wartości z przedziału $[-2, -1] \cup [1, 2]$.
- Wypisz liczbę oraz proporcję ujemnych elementów \mathbf{x} .
- Policz średnią z wartości bezwzględnych z \mathbf{x} .
- Znajdź element najbardziej oddalony od 0.
- Znajdź element najbardziej oddalony od 2.
- Wypisz część ułamkową każdej z liczb.
- Stwórz wektor \mathbf{y} , taki, że $\mathbf{y}[\mathbf{i}]$ równa się nieujemny dla $\mathbf{x}[\mathbf{i}] \geq 0$ oraz ujemny w przeciwnym przypadku.
- Stwórz wektor \mathbf{z} , taki, że $\mathbf{z}[\mathbf{i}]$ równa się mały dla $\mathbf{x}[\mathbf{i}] < 1$, duży dla $\mathbf{x}[\mathbf{i}] > 1$ i średni w przeciwnym przypadku. Wektor wygeneruj następująco:

```
import numpy as np
np.random.seed(123)
x = np.round(np.random.normal(size = 20), 2)
```

2. Napisz funkcję `clamp(x)`, która elementy mniejsze od 0, ustawi na 0, a elementy większe od 1 na 1.

3. Napisz funkcję `err(x, y)`, która policzy proporcję źle sklasyfikowanych elementów –

$$\text{ERR}(\mathbf{w}, \mathbf{y}) = \frac{1}{|\mathbf{w}|} \sum_{j=1}^{|\mathbf{w}|} 1(w_j = y_j), \text{ gdzie } 1(a = b) = 1, \text{ a } 1(a \neq b) = 0.$$

3. Napisz funkcję `mse(x, y)`, która policzy błąd średniokwadratowy –

$$\text{MSE}(\mathbf{w}, \mathbf{y}) = \frac{1}{|\mathbf{w}|} \sum_{j=1}^{|\mathbf{w}|} |w_j - y_j|^2$$

4. Napisz funkcję `rmse(x, y)`, która policzy pierwiastek z błędu średniokwadratowego –

$$\text{RMSE}(\mathbf{w}, \mathbf{y}) = \frac{1}{|\mathbf{w}|} \sqrt{\sum_{j=1}^{|\mathbf{w}|} (|w_j - y_j|^2)}$$

5. Napisz funkcję `mad(x, y)`, która policzy błąd bezwzględny $\text{MAD}(\mathbf{w}, \mathbf{y}) = \frac{1}{|\mathbf{w}|} \sum_{j=1}^{|\mathbf{w}|} |w_j - y_j|$

6. Napisz funkcję `approx_pi(n)`, która przybliży liczbę π metodą Monte Carlo.

7. Napisz funkcję `min_max_col(X)`. Niech $X \in \mathbb{R}^{n \times d}$. Każda z n kolumn macierzy reprezentuje oddzielny punkt z d wymiarowej przestrzeni. Wyznacz macierz $B \in \mathbb{R}^{2 \times d}$, gdzie $b_{1,j} = \min_i x_{i,j}$ oraz $b_{2,j} = \max_i x_{i,j}$. Jest to macierz zawierająca informację o rozpiętości danych w każdym z d wymiarów.

8. Napisz funkcję `standarize(X)`, która wyskaluje każdą kolumnę macierzy (oddzielnie). Proces skalowania polega na:

- Policzeniu średniej z kolumny;
- Policzeniu odchylenia standardowego z kolumny;
- Odjęciu wartości średniej od kolumny;
- Podzieleniu kolumny przez odchylenie standardowe. \

Wyskalowanie kolumny sprawia, że jej wartość średnia wynosi 0, a odchylenie standardowe 1. Jest to procedura często stosowana w ML.