

**AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY**

WYDZIAŁ FIZYKI I INFORMATYKI STOSOWANEJ  
KIERUNEK STUDIÓW: INFORMATYKA STOSOWANA



# **METODY INTELIGENCJI OBLICZENIOWEJ**

---

## **Laboratorium 2**

**Klasyfikacja z użyciem sztucznych neuronów**

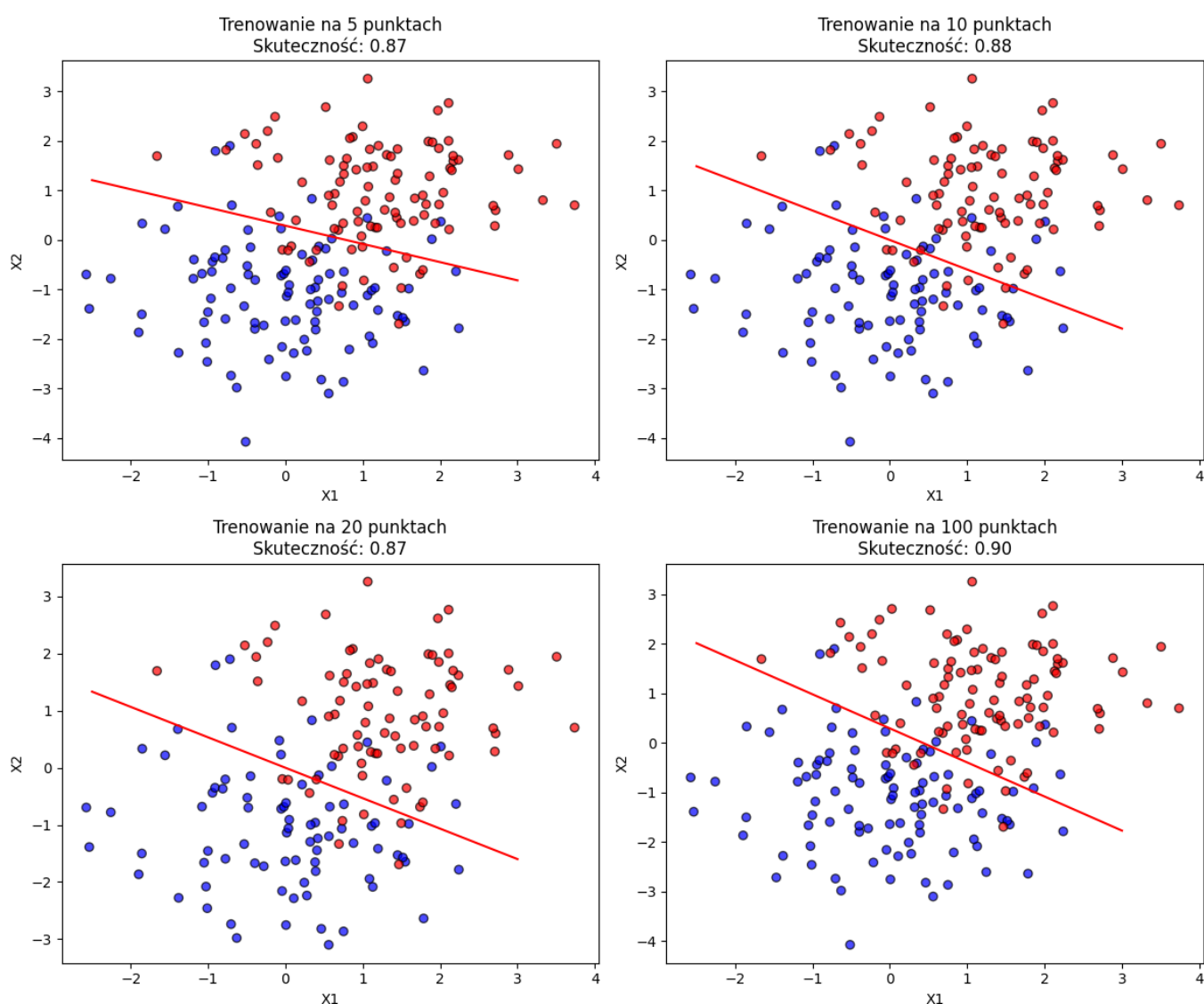
---

**zrealizował**  
Przemysław Ryś

Kraków, 11 Marzec 2025

## Zadanie 1

Zadanie polegało na wygenerowaniu zbioru punktów należących do dwóch klas: K1 i K2. Punkty z klasy K1 zostały wylosowane z rozkładu normalnego o średniej  $\vec{\mu}_1 = [0, -1]$  i wariancji  $\sigma_1^2 = 1$ , natomiast punkty z klasy K2 pochodziły z rozkładu normalnego o średniej  $\vec{\mu}_2 = [1, 1]$  i wariancji  $\sigma_2^2 = 1$ . Cały zbiór zawierał 200 punktów, po 100 dla każdej klasy. Następnie dokonano podziału danych na zbiory uczące o licznosci 5, 10, 20 oraz 100 próbek. Dla każdego podziału wyznaczono równanie prostej, która najlepiej oddziela klasy K1 i K2. Ostateczne wyniki zostały przedstawione na rysunku 1. Można tu zaobserwować rosnącą skuteczność poprawnego przypisania danych do odpowiednich klas w zależności od ilości wykorzystanych do uczenia danych (w stosunku do danych testowych).



Rys. 1: Wykresy skuteczności modeli perceptronu w zależności od ilości danych trenujących.

## Zadanie 2

Przeprowadzono analizę próbek benzyny, które przypisano do dwóch klas czystości: A i B (oznaczonych jako 0 i 1). Dane pochodziły z pliku fuel.txt, gdzie pierwsze trzy kolumny zawierały właściwości fizykochemiczne, a czwarta kolumna klasę czystości. Sprawdzono skuteczność klasyfikacji za pomocą pojedynczego neuronu, przeprowadzając pięciokrotne uczenie na całym zbiorze.

Iteracja 1: 0.9700  
 Iteracja 2: 0.9700  
 Iteracja 3: 0.9700  
 Iteracja 4: 0.9700  
 Iteracja 5: 0.9700

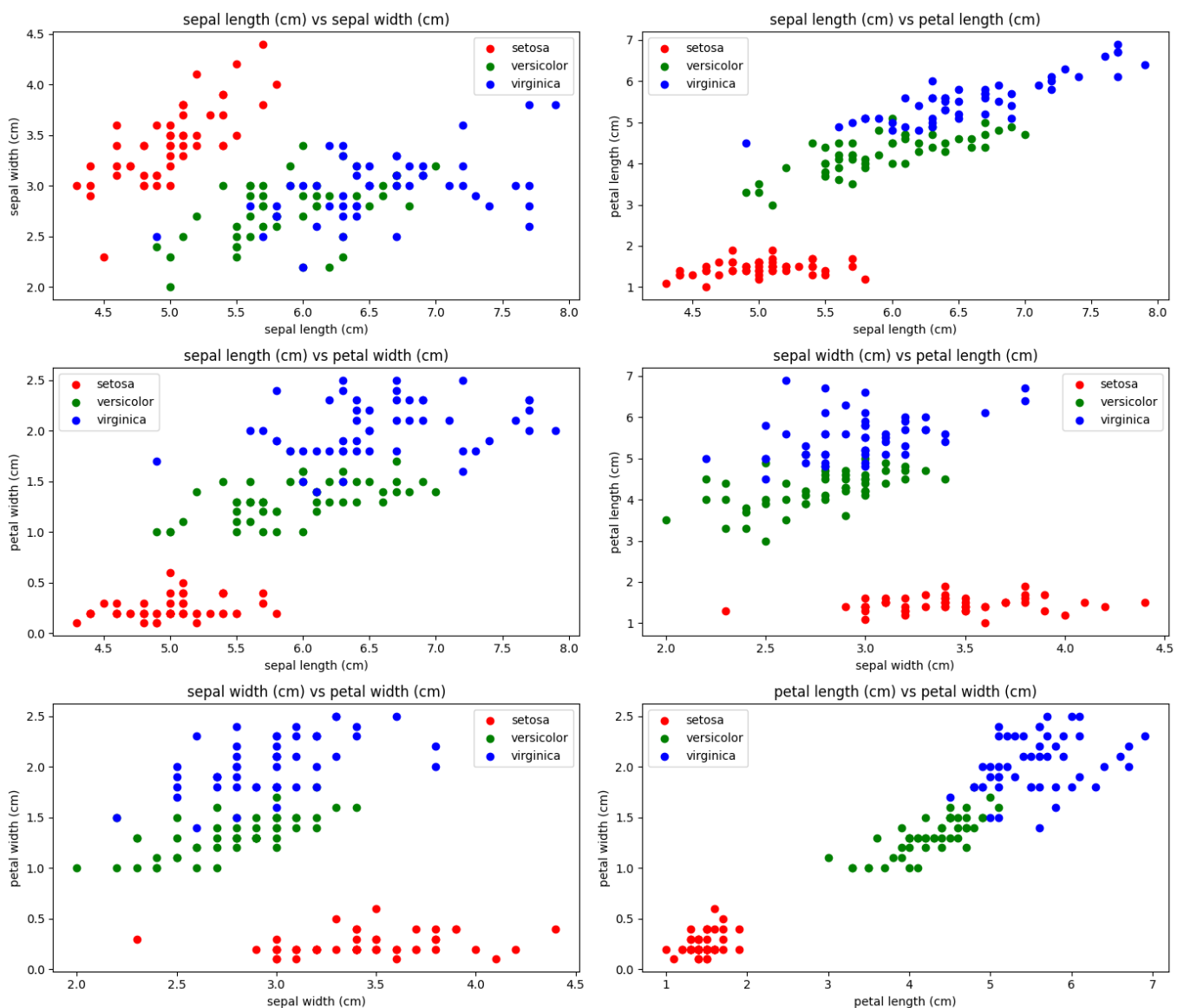
Macierz błędów była identyczna dla każdej iteracji i miała postać:

$$\begin{bmatrix} 41 & 1 \\ 2 & 56 \end{bmatrix}$$

co na 100 punktów daje 3 które są błędnie przypisane (pozadiagonalne). Macierze pomyłek na elementach diagonalnych natomiast prezentują poprawne przyporządkowania do klas.

### Zadanie 3

Wykorzystano zbiór Iris dataset dostępny w `sklearn.datasets.load_iris()`. Podzielono dane na zbiór uczący (80%) i testowy (20%). Następnie zbudowano sieć z pojedynczą warstwą perceptronów, której zadaniem była klasyfikacja gatunków irysów.



Rys. 2: Wykresy danych 3 różnych klas irysów w różnych zestawieniach osi.

**Dokładności** klasyfikacji dla kolejnych uruchomień były następujące:

0.9333, 0.9333, 0.7000, 0.7667, 0.6667

**Średnia dokładność: 0.8000**

**Macierze pomyłek:**

$$\begin{bmatrix} 11 & 0 & 0 \\ 0 & 11 & 0 \\ 0 & 2 & 6 \end{bmatrix}, \begin{bmatrix} 15 & 0 & 0 \\ 0 & 8 & 1 \\ 0 & 1 & 5 \end{bmatrix}, \begin{bmatrix} 8 & 0 & 0 \\ 7 & 3 & 0 \\ 0 & 2 & 10 \end{bmatrix}, \begin{bmatrix} 7 & 1 & 0 \\ 0 & 10 & 0 \\ 0 & 6 & 6 \end{bmatrix}, \begin{bmatrix} 11 & 0 & 0 \\ 5 & 0 & 5 \\ 0 & 0 & 9 \end{bmatrix}$$

Oznacza to, że model poprawnie klasyfikuje około 80% przypadków w uśrednionym wyniku, choć wyniki poszczególnych uruchomień wykazują pewne rozbieżności.

## Zadanie 4

Dane z Iris dataset podzielono na zbiór uczący i testowy na trzy różne sposoby: 80/20, 70/30, 60/40 (uczące/testowe). Poniżej przedstawiono wpływ różnych podziałów na skuteczność klasyfikacji.

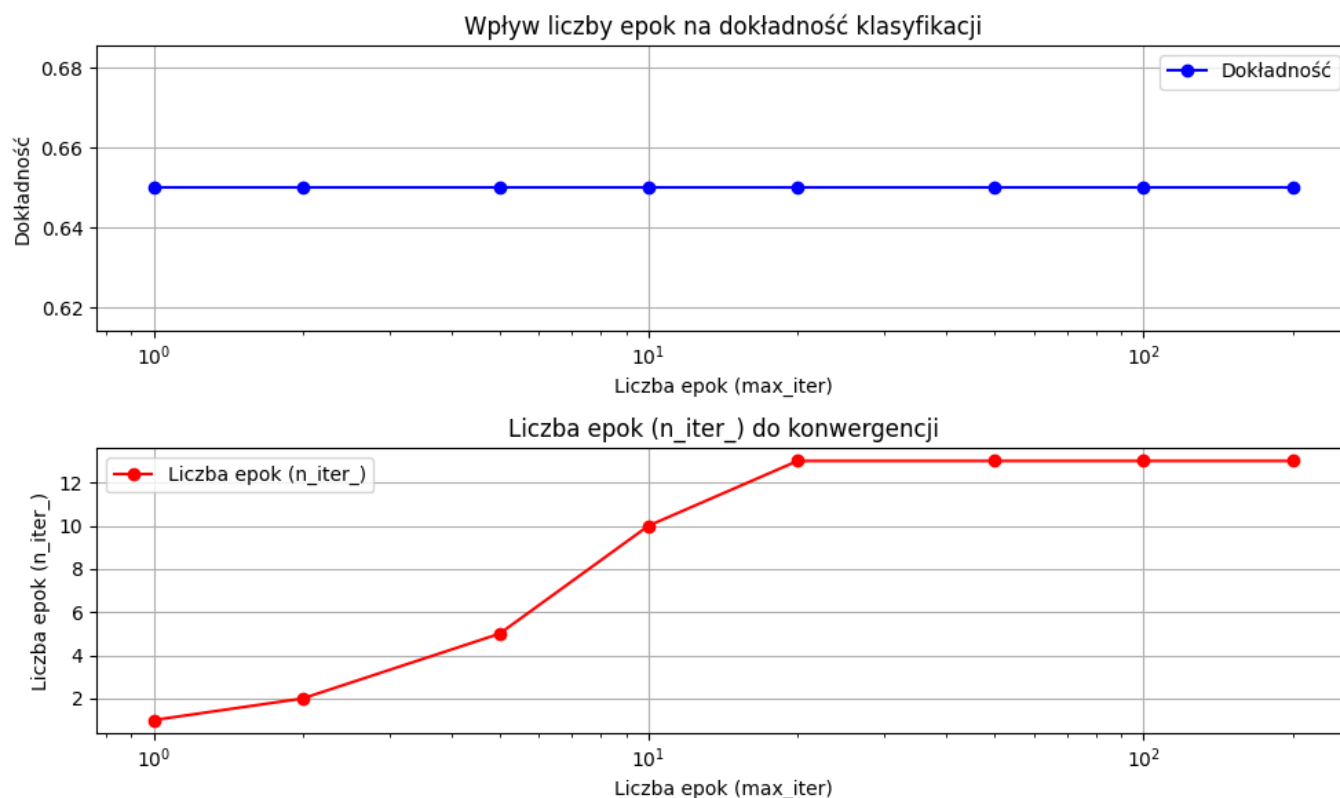
80/20 : 0.6207 || 70/30 : 0.7111 || 60/40 : 0.6500

**Średnia dokładność dla różnych podziałów: 0.6606**

## Zadanie 5

Zbadano wpływ liczby epok na trafność klasyfikacji zbioru irysów, testując różne wartości. Aby zapobiec przedwczesnemu zatrzymaniu treningu, ustawiono **tol** na małą wartość oraz **early\_stopping=False**. Dodatkowo przeanalizowano rzeczywistą liczbę epok treningowych (**n\_iter\_**).

Wyniki przedstawiono na wykresie 3



Rys. 3: Wykresy dokładności klasyfikacji oraz liczby epok do uzyskania konwergencji (zbieżności) w zależności od ilości epok.