

## Laboratorium Podstaw Sztucznej Inteligencji

### Ćwiczenie 2.

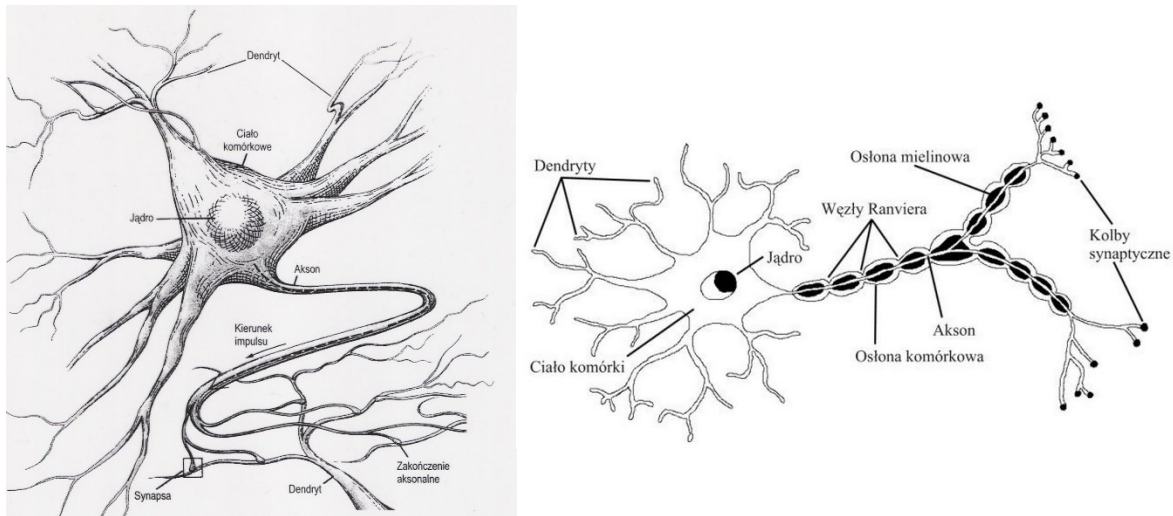
Rozwiązanie zagadnienia klasyfikacji za pomocą pojedynczego neuronu  
uczonego metodą bezgradientową.

Opracował:

Dr inż. Piotr Urbanek.

## 1. WPROWADZENIE – BUDOWA NEURONU

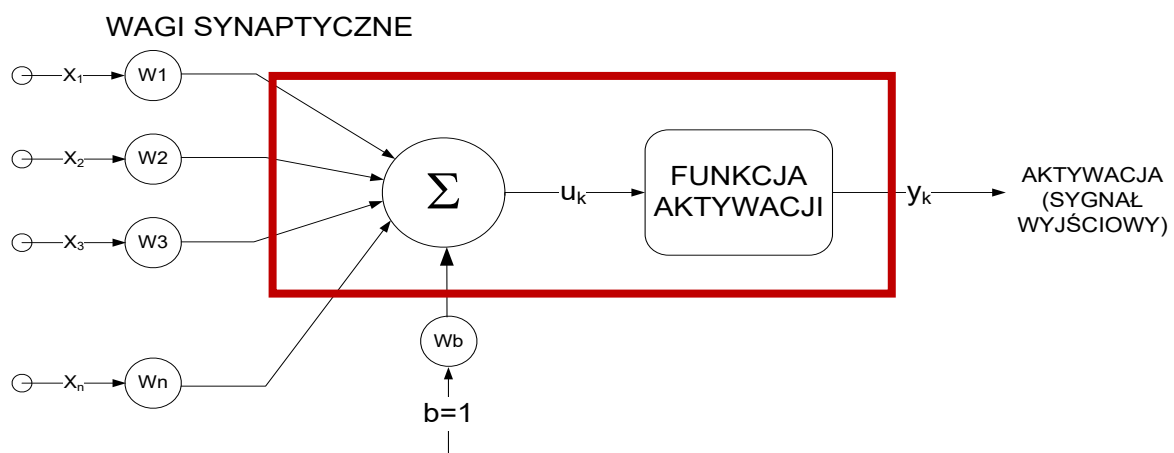
Wygląd pojedynczego neuronu, z którego składa się najbardziej skomplikowana znana nam struktura we wszechświecie, zwana mózgiem przedstawia rys. 1.



Rys. 1. Uproszczony schemat struktury pojedynczego neuronu.

Neuron taki składa się z jądra komórkowego, do którego dochodzą wejścia zwane dendrytami oraz jednego wyjścia zwanego aksonem, który poprzez połączenia synaptyczne łączy się z dendrytami kolejnego neuronu tworząc około  $10^{15}$  połączeń. Synapsy pełnią rolę tzw. bariery progowych, które przepuszczają sygnał elektryczny wtedy, gdy jest on większy od bariery potencjału elektrycznego występującymi w synapsach.

W 1943 roku dwaj uczeni amerykańscy McCulloch i Pitts zaproponowali matematyczny model takiego neuronu.



$$u_k = \sum_i w_{ki} x_i + w_b \cdot 1$$

Rys. 2. Model matematyczny neuronu.

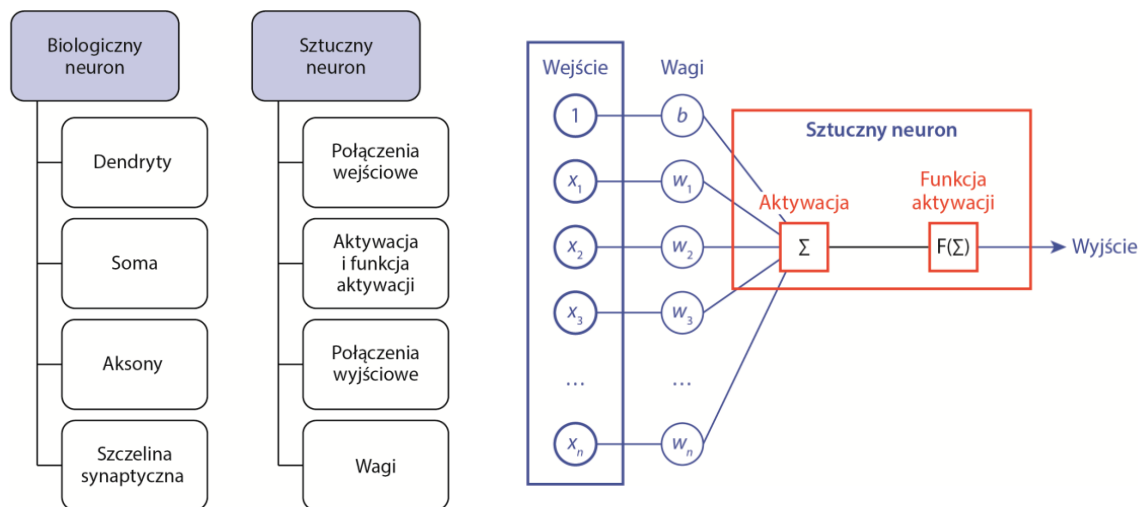
Funkcja aktywacji modeluje istnienie połączeń synaptycznych, a w wagach (mnożnikach)  $w_1, w_2, \dots, w_n$  oraz  $w_b$  przechowywana jest wiedza, jak neuron powinien reagować na wprowadzane sygnały wejściowe.

Funkcję aktywacji da się zamodelować jako tzw. funkcję Heaviside'a daną jednym z dwóch wzorów:

$$f(x) = \begin{cases} -1 & \text{dla } x \leq 0 \\ 1 & \text{dla } x > 0 \end{cases} \quad (1)$$

$$f(x) = \begin{cases} 0 & \text{dla } x \leq 0 \\ 1 & \text{dla } x > 0 \end{cases} \quad (2)$$

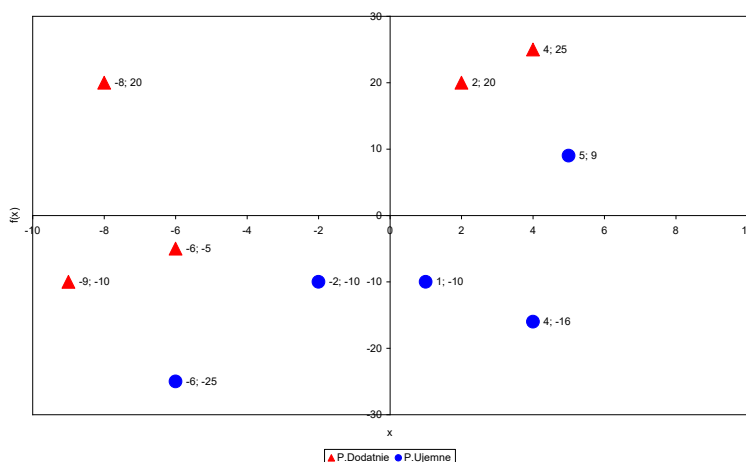
Podsumowując, na rys. 3. Przedstawiono podstawowe podobieństwa pomiędzy neuronem biologicznym a neuronem matematycznym<sup>1</sup>:



Rys. 3. Porównanie neuronu biologicznego ze sztucznym

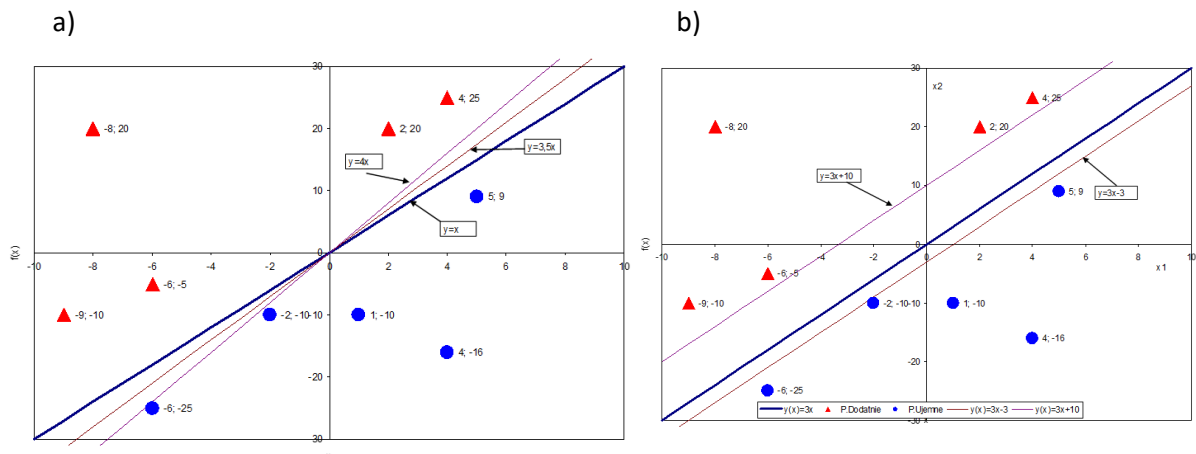
## 2. ZASTOSOWANIE POJEDYNCZEGO NEURONU W OBLICZENIACH

Najprostszym zastosowaniem obliczeniowym pojedynczego neuronu jest rozwiązanie zagadnienia klasyfikacji dwóch zbiorów zdefiniowanych, np. na płaszczyźnie kartezjańskiej.



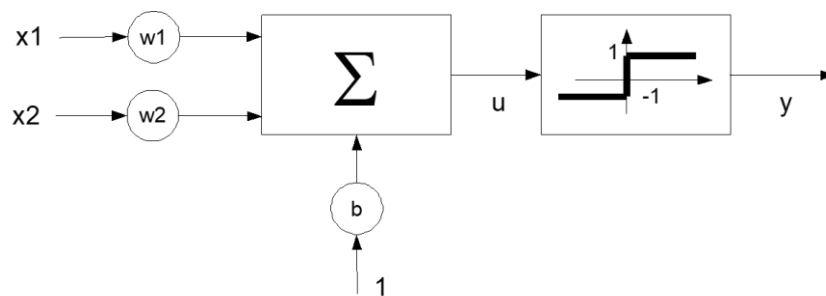
Rys. 4. Płaszczyzna kartezjańska ze zdefiniowanymi punktami na płaszczyźnie.

Warunkiem poprawnego działania takiego neuronu jest to, żeby punkty te były liniowo separowalne, co oznacza możliwość poprowadzenia między nimi linii prostej, jak na rys. 5 a i b



Rys. 5. Proste separujące punkty na płaszczyźnie kartezjańskiej. a)  $y=ax$  i b)  $y=ax+b$

Proste separujące mogą być zamodelowane strukturą pojedynczego neuronu.



Rys. 6. Model pojedynczego neuronu reprezentujący na płaszczyźnie kartezjańskiej prostą separującą.

Model ten może być opisany równaniem 3:

$$y = f(w_1 \cdot x_1 + w_2 \cdot x_2 + b) \quad (3)$$

Co w zapisie macierzowym przyjmuje postać:

$$y = f\left([w_1 \quad w_2] * \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + b\right) \quad (4)$$

Można udowodnić, że wektor wag  $[w_1 \ w_2]$  reprezentuje współczynnik kierunkowy prostej, a  $b$  przesunięcie prostej względem układu współrzędnych.

### 3. UCZENIE POJEDYNCZEGO NEURONU METODĄ BEZGRADIENTOWĄ.

W przypadku dwóch wejść  $[x_1 \ x_2]^T$  perceptron dzieli płaszczyznę na dwie części. Zatem może klasyfikować punkty należące do dwóch umownych klas, np. "dodatnich" i "ujemnych". Podział ten wyznacza prosta o równaniu:

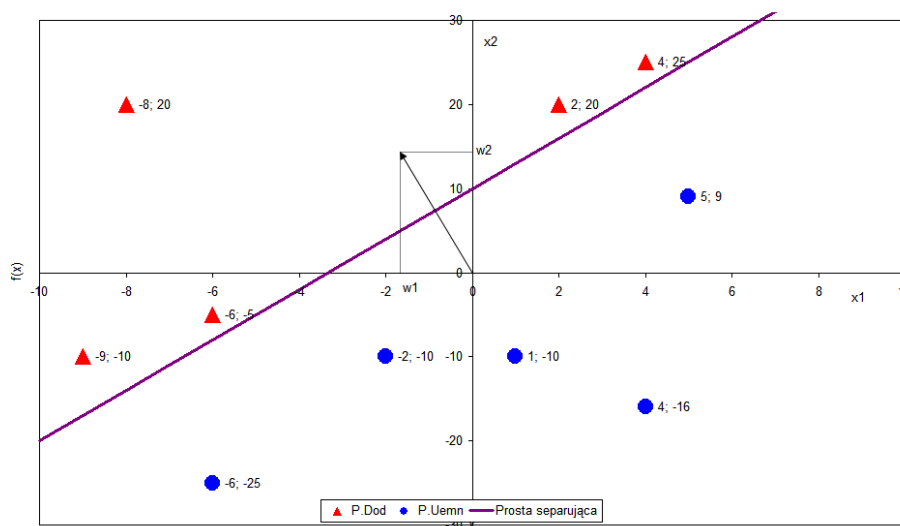
$$w_1 \cdot x_1 + w_2 \cdot x_2 + b = 0 \quad (5)$$

Zatem dwa charakterystyczne punkty, przez które przechodzi prosta o równaniu (5) wynoszą:

$$x_2 = -b/w_2 \text{ dla } x_1 = 0 \quad (6a)$$

$$x_1 = -b/w_1 \text{ dla } x_2 = 0 \quad (6b)$$

co daje się przedstawić na rysunku 7

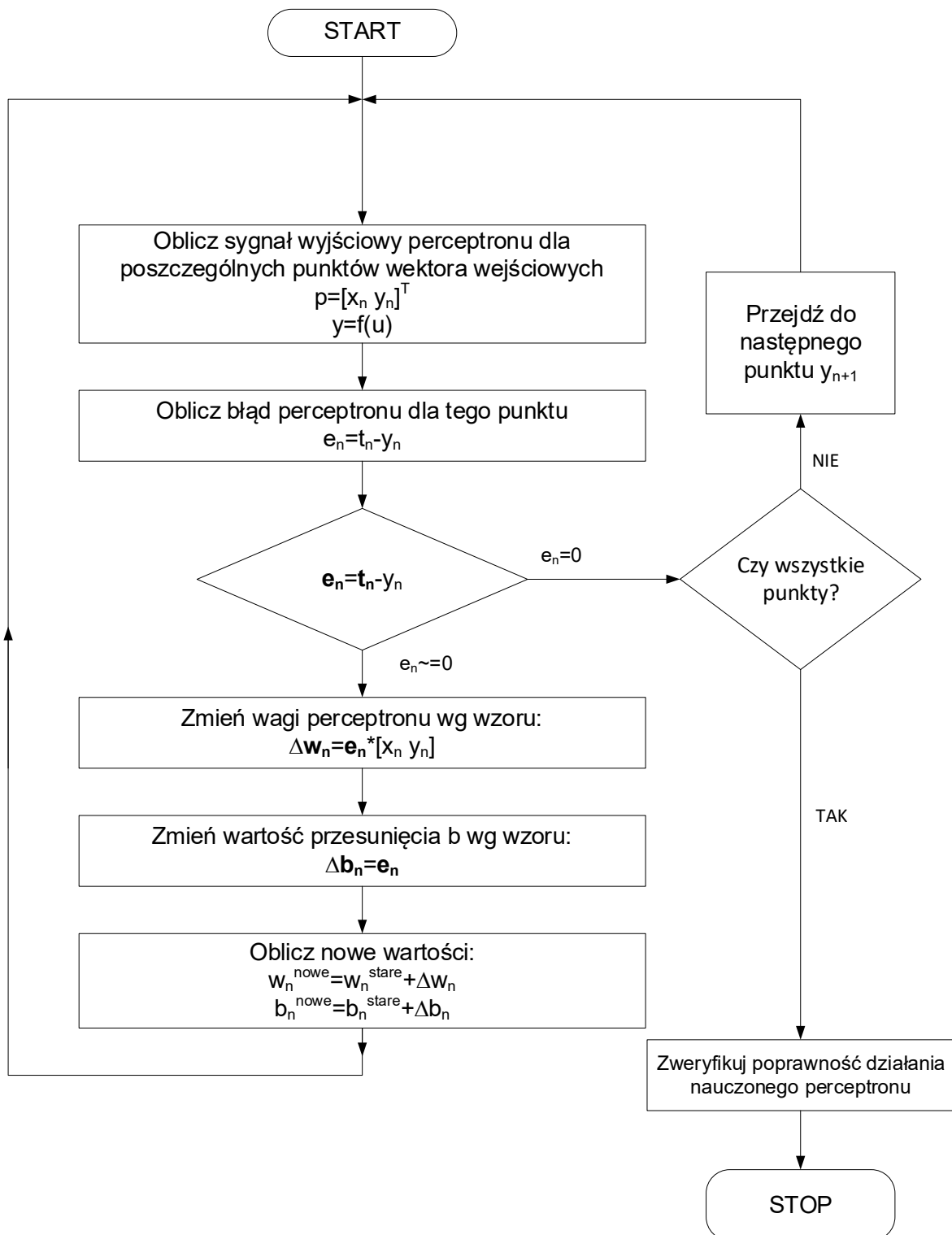


Rys. 7. Położenie prostej separującej dwa zbiory na płaszczyźnie kartezjańskiej.

Warto zauważyć, że w teorii takich prostych może być nieskończenie wiele, zatem istnieje nieskończenie wiele par  $[w_1 \ w_2]$  i  $b$ , które spełniają zadanie klasyfikacji.

Wyznaczanie wartości  $[w_1 \ w_2]$  i  $b$  może odbywać się dwiema metodami: metodą bezgradientową oraz wykorzystującą gradient błędów perceptronu. Jednak ta druga metoda stosowana jest głównie w sieciach o większej liczbie neuronów ułożonych w struktury wielowarstwowe. Będzie o nich mowa w dalszych ćwiczeniach w ramach obecnego laboratorium.

Sieć działań uczenia pojedynczego neuronu ze skokową funkcją aktywacji przedstawia rys. 8.



Rys. 8. Sieć działań nauki pojedynczego perceptronu metodą bezgradientową.

Opisać go można w następujący sposób:

1. Przygotowanie dwóch ciągów: uczącego i weryfikującego.  
W skład ciągu uczącego wchodzi wektor wejściowy oraz wektor wyjściowy.
2. Ustanowienie początkowych wartości wag (wartości przypadkowe).
3. Po przetworzeniu wektora wejściowego nauczyciel porównuje wartości otrzymane z wartościami oczekiwanymi informując sieć o błędzie odpowiedzi.

4. Dla danego punktu uczącego następuje korekcja wag, tak aby błąd odpowiedzi uzyskany przy powtórным przetworzeniu wektora wejściowego równał się zero.
5. Czynności 1-4 powtarza się aż dla każdego punktu uczącego wyjście perceptronu będzie zgadzać się z wartością ze zbioru nauczyciela.

W metodzie tej nie chodzi o wyznaczenie wartości błędu całkowitego dla wszystkich klasyfikowanych punktów, lecz o osiągnięcie błędu o wartości zero dla każdego punktu oddzielnie.

#### **4. ZADANIE.**

1. Napisać aplikację modelującą działanie pojedynczego perceptronu użytego do rozwiązania zagadnienia klasyfikacji dwóch zbiorów.
2. Po jej wykonaniu przetestować proces uczenia wykonanej struktury dla różnych liniowo separowalnych zbiorów punktów uczących.
3. Po nauczaniu neuronu sprawdzić poprawność jego działania wprowadzając punkty spoza zbioru uczącego. Wszystkie wprowadzone punkty powinny być sklasyfikowane poprawnie.