Politechnika Łódzka Instytut Informatyki Stosowanej

Laboratorium Sztucznej Inteligencji

Ćwiczenie 1.

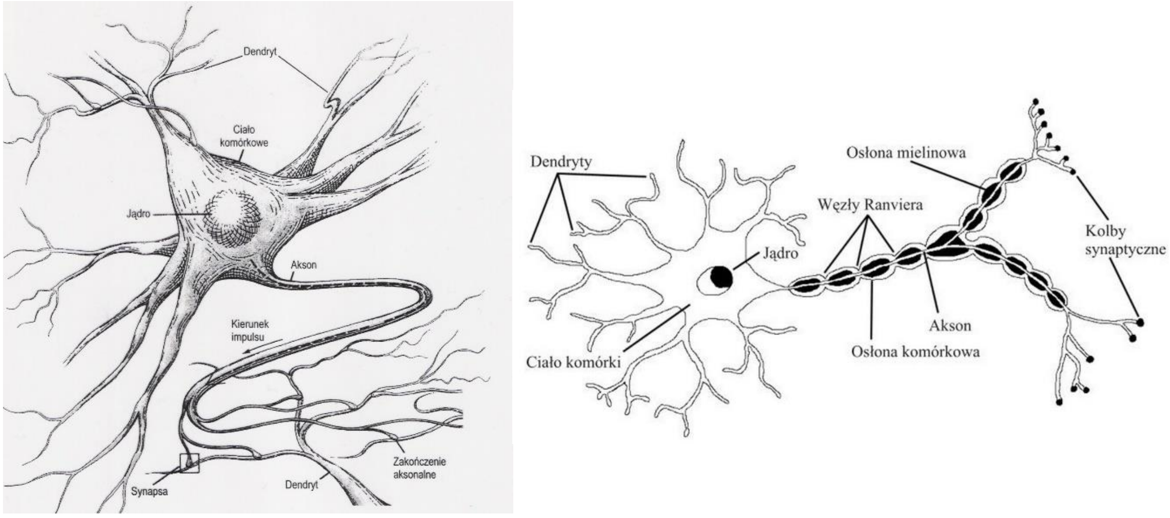
Rozwiązanie zagadnienia klasyfikacji za pomocą pojedynczego neuronu uczonego metodą bezgradientową.

Opracował:

Dr inż. Piotr Urbanek.

1. Wprowadzenie – budowa neuronu

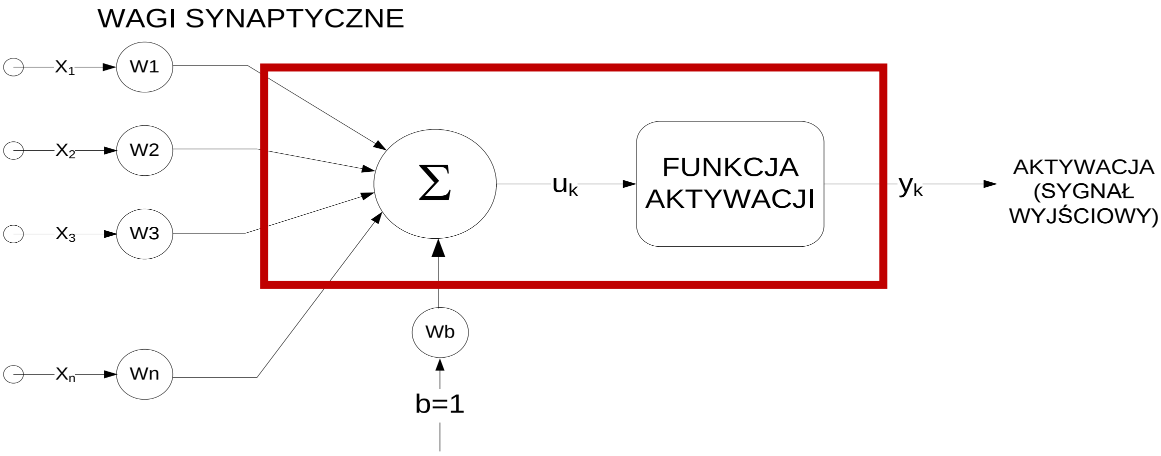
Wygląd pojedynczego neuronu, z którego składa się najbardziej skomplikowana znana nam struktura we wszechświecie, zwana mózgiem przedstawia rys. 1.



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. barier 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.

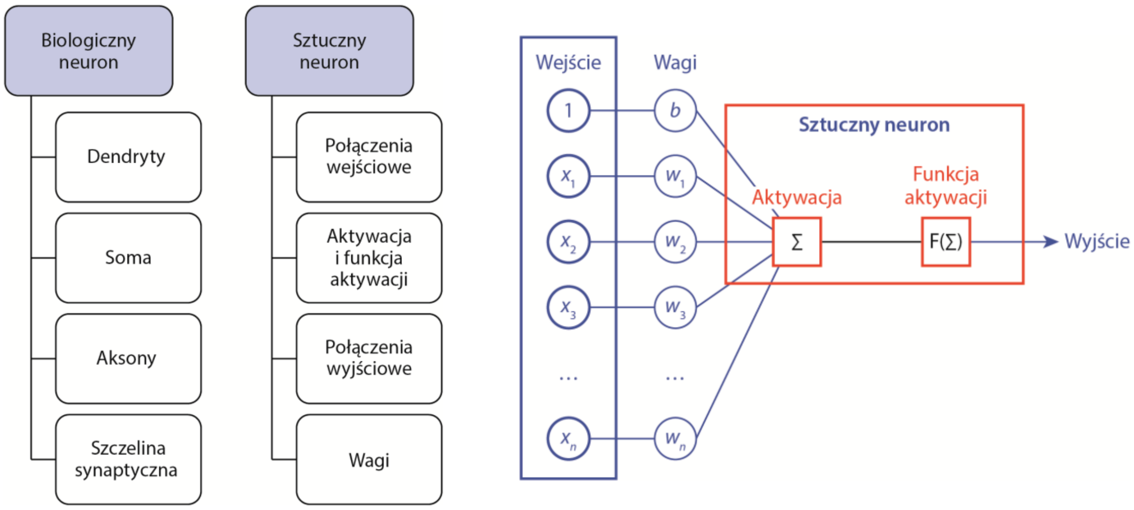
Funkcja aktywacji modeluje istnienie połączeń synaptycznych, a w wagach (mnożnikach) w_1, w_2, …. , w_n oraz w_b przechowywana jest wiedza, jak neuron powinien reagować na wprowadzane sygnały wejściowe.

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

$$f(x) = \begin{cases} 1 \; dla \; x > 0 \\ -1 \; dla \; x \leq 0 \end{cases} \tag{1}$$

$$f(x) = \begin{cases} 1 \; dla \; x > 0 \\ 0 \; dla \; x \leq 0 \end{cases} \tag{2}$$

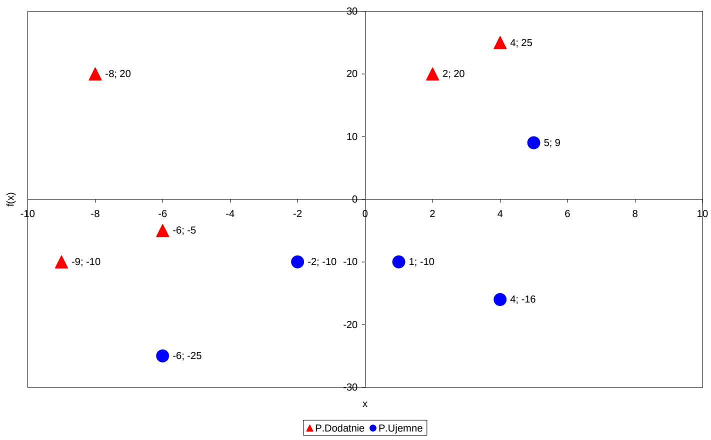
Podsumowując, na rys. 3. Przedstawiono podstawowe podobieństwa pomiędzy neuronem biologicznym a neuronem matematycznym[1]:



Rys. 3. Porównanie neuronu biologicznego ze sztucznym

2. Zastosowanie pojedynczego neuronu w obliczeniach

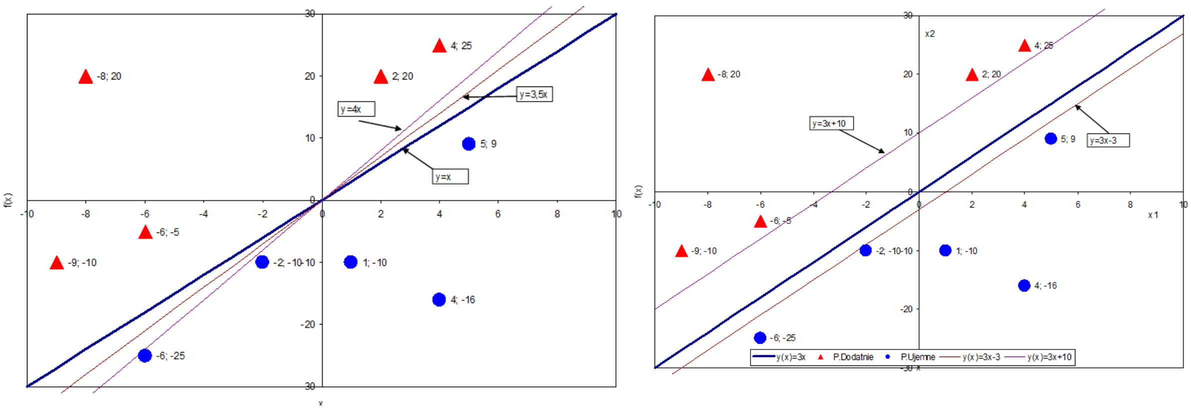
Najprostszym zastosowaniem obliczeniowym pojedynczego neuronu jest rozwiązanie zagadnienia klasyfikacji dwóch zbiorów zdefiniowanych, np. na płaszczyźnie kartezjańskiej.



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

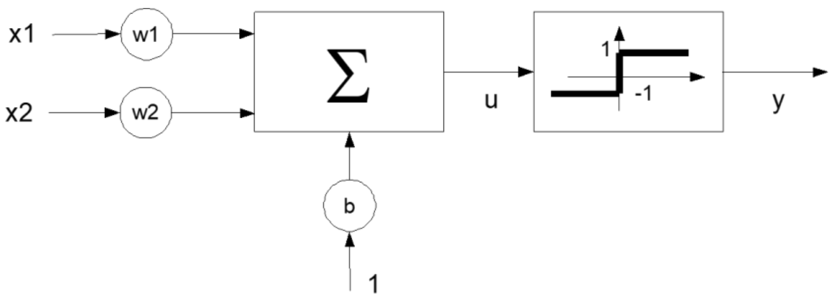
Warunkiem poprawnego działania takiego neuronu jest to, żeby punkty te były liniowo separowalne, co oznacza możliwość poprowadzenia między nimi linii prostej, jak na rys. 5 a i b

a) 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:

$$y = f(w_1 \cdot x_1 + w_2 \cdot x_2 + b) \tag{3}$$

Co w zapisie macierzowym przyjmuje postać:

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

Gdzie f oznacza funkcję Heavisade'a daną wzorem (1) lub (2). **W przypadku obecnego ćwiczenia polecany jest wzór (2).**

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

3. Uczenie pojedynczego neuronu metodą bezgradientową.

W przypadku dwóch wejść $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 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:

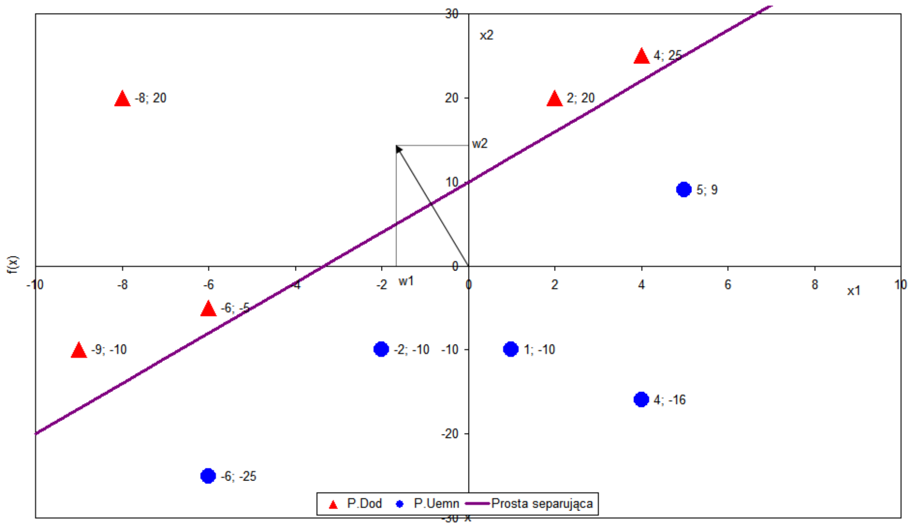
$$w_1 \cdot x_1 + w_2 \cdot x_2 + b = 0 \qquad (5)$$

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

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

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

co daje się przedstawić na rysunku 7



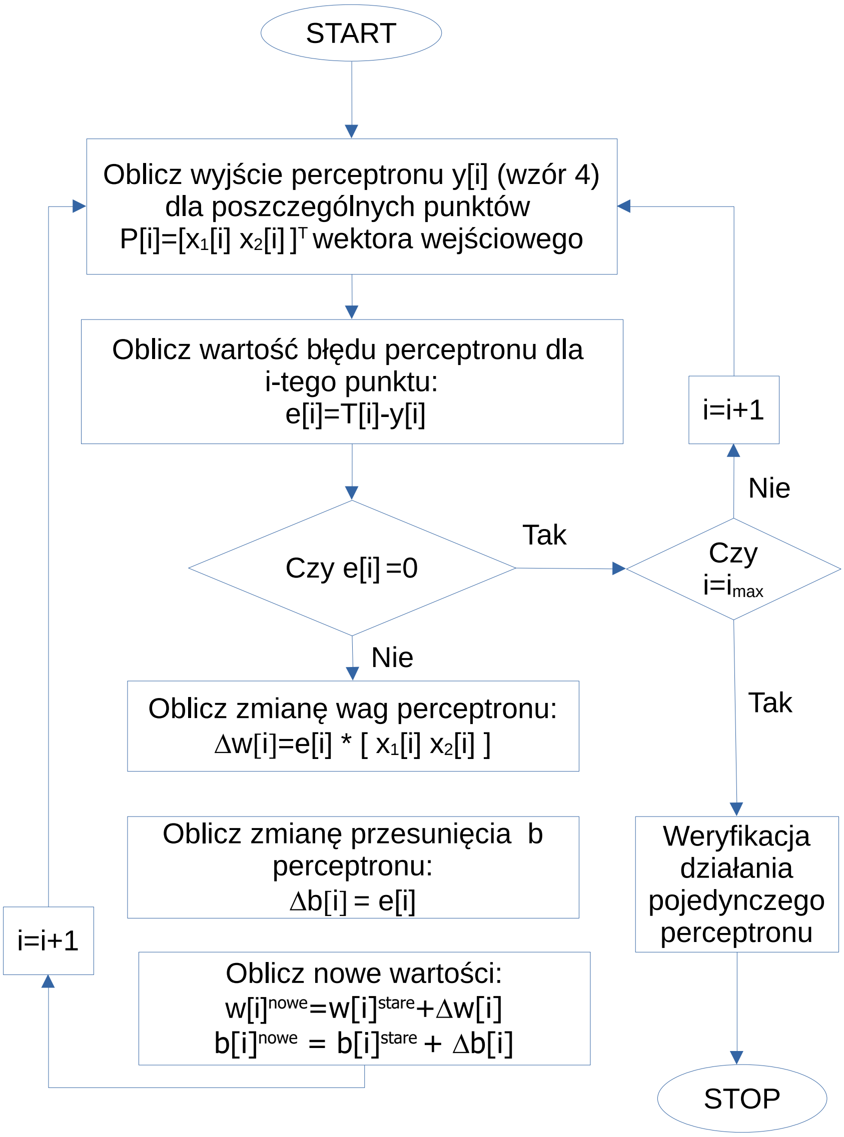
Rys. 7. Położenie prostej separującej dwa zbiory na płaszczyźnie kartezjańskiej.

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

Wyznaczanie wartości [w1 w2] i b może odbywać się dwiema metodami: metodą bezgradientową oraz wykorzystującą gradient błędu 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.

Metoda bezgradientowa uczenia pojedynczego perceptronu rozwiązującego zagadnienie klasyfikacji dwóch zbiorów.



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 powtórnym przetworzeniu wektora wejściowego równał się zeru.
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 nauczeniu neuronu sprawdzić poprawność jego działania wprowadzając punkty spoza zbioru uczącego. Wszystkie wprowadzone punkty powinny być sklasyfikowane poprawnie.