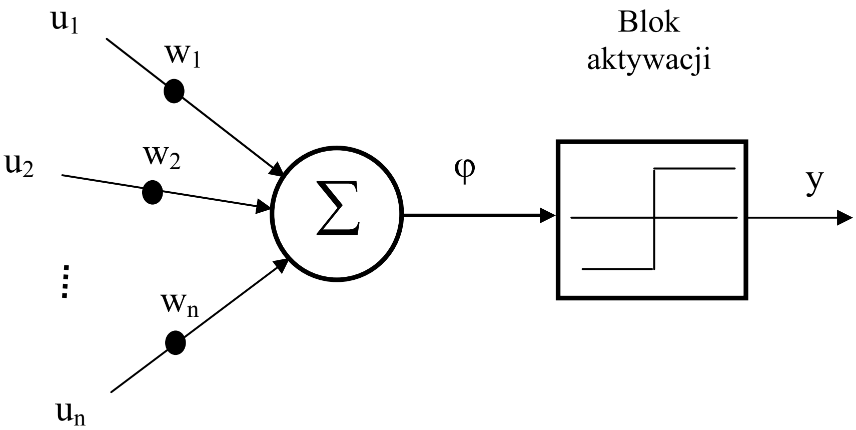
Laboratorium systemów obliczeń inteligentnych

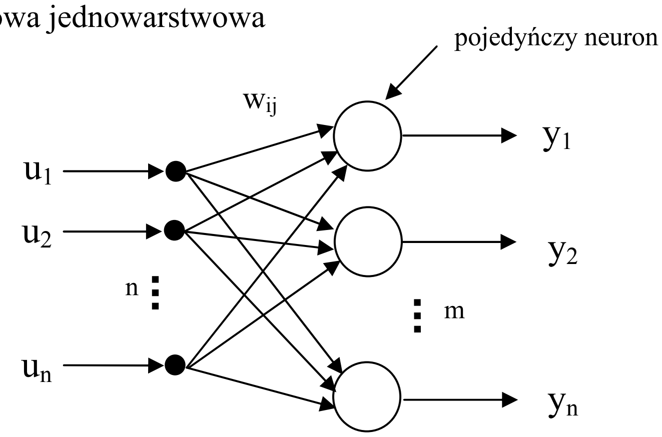
Modele neuronów. Bezgradientowe metody uczenia sieci neuronowych.

1. Perceptron - Matematyczny model neuronu



Gdzie **u_1, u_2, ..., u_n** - sygnały wejściowe, **w_1, w_2, ..., w_n** - wartości wag, **y** - sygnał wyjściowy

2. Sieć perceptronowa jednowarstwowa



Gdzie **u_1, u_2, ..., u_n** - sygnały wejściowe, **w_{ij}** - wartości wag, **y_1, y_2, ...,y_m** - sygnały wyjściowe

$$\varphi_j = \sum_{i=0}^{n} u_i w_{ij} \quad , \quad y_j = f\left(\varphi_j\right)$$

3. Algorytm uczenia perceptronu

a) podanie danych na wejścia neuronu $\mathbf{u_1, u_2, ..., u_n}$ i wyznaczenie wartości na wyjściu neuronu $\mathbf{y_1, y_2, ..., y_m}$
b) porównanie odpowiedzi neuronu $\mathbf{y_j}$ z odpowiedziami oczekiwanymi $\mathbf{yz_j}$
c) obliczenie korekt wag $\mathbf{w_{ij}}$ według następującego wzoru (j - nr wyjścia):

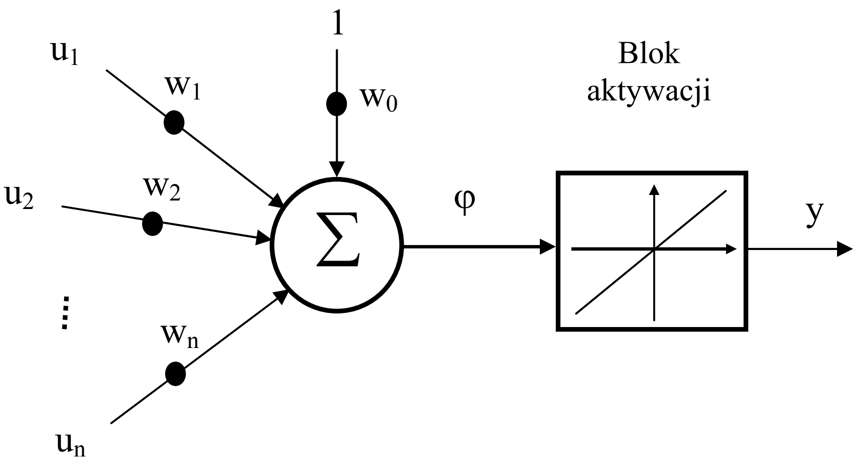
$$\Delta w_{ij}^t = \begin{cases} 0 & gdy \quad y_j^t = yz_j^t \\ u_i^t & gdy \quad y_j^t = 0 \ i \ yz_j^t = 1 \\ -u_i^t & gdy \quad y_j^t = 1 \ i \ yz_j^t = 0 \end{cases}$$

$$\Delta w_{ij} = \sum_{t=1}^{N} \Delta w_{ij}^t$$

d) skorygowanie wag według następującego wzoru

$$w_{ij}^{nowe} = w_{ij}^{stare} + \eta \Delta w_{ij}$$

4. Model neuronu Adaline



Gdzie $\mathbf{u_1, u_2, ..., u_n}$ - sygnały wejściowe, $\mathbf{w_1, w_2, ..., w_n}$ - wartości wag, $\mathbf{y}$ - sygnał wyjściowy

$$\varphi = \sum_{i=0}^{n} u_i w_i \quad , \quad y = \varphi$$

5. Algorytm uczenia sieci Madaline metodą Widrowa-Hoffa (reguła delty)

e) podanie danych na wejścia sieci $\mathbf{u_1, u_2, ..., u_n}$ i wyznaczenie wartości na wyjściach sieci $\mathbf{y_i}$
f) porównanie odpowiedzi sieci $\mathbf{y_i}$ z odpowiedziami oczekiwanymi $\mathbf{y_{zi}}$ i wyznaczenie błędu według następującego wzoru

$$\delta_i^t = y_{zi}^t - y_i^t$$

g) obliczenie korekt wag $\mathbf{w_0, w_1, ..., w_n}$ według następującego wzoru:

$$\Delta w_i^t = -2\delta_i^t u_i^t$$

$$\Delta w_i = \sum_{t=1}^{P} \Delta w_i^t$$

h) skorygowanie wag według następującego wzoru

$$w_i^{nowe} = w_i^{stare} - \eta \Delta w_i$$

i) Kroki od a do d powtarzane są aż suma kwadratów błędów ξ nie stanie się wystarczająco mała

$$\xi = \sum_{t=1}^{P} \sum_{i=1}^{m} \left(y_{zi}^t - y_i^t \right)^2$$

gdzie **P** to liczba zbiorów uczących a **m**. to liczba wyjść sieci

Zadania - Perceptron

1. Wyznaczyć wektor wag $\mathbf{[w_0\ w_1\ w_2]}$, dla których neuron będzie klasyfikował elementy dwóch zbiorów $\mathbf{A=\{a_1,a_2\}}$ i $\mathbf{B=\{b_1,b_2\}}$ gdzie $\mathbf{a_1=(1,1), a_2=(2,1), b_1=(4,3), b_2=(1,5).}$
2. Wyznaczyć parametrycznie granicę decyzyjną neuronu o dwóch wejściach.
3. Dokonać symulacji uczenia i testowania neuronu przy użyciu następujących instrukcji z toolboxu *neural networks: HARDLIM, LEARNP, TRAINP, RANDS* dla następujących danych

a) wykorzystać dane z zadania 1
b) $A=\{a_1,a_2,a_3\}$, $B=\{b_1,b_2\}$ gdzie $a_1=\{-0.5,-0.5)$ $a_2=(0.5,-1)$ $a_3=(-1,1)$, $b_1=(1,2)$ $b_2=(-0.5,2.5)$

4. Zbadać wpływ kroku uczenia η na szybkość oraz przebieg uczenia neuronu

Zadania - Adaline i Madaline

1. Zbadać czy poniższe zadania klasyfikacji możliwe są do realizacji przy użyciu perceptronu (jeżeli zadanie jest możliwe do realizacji wyznaczyć wektor wag **w,** dla których neuron będzie klasyfikował elementy dwóch zbiorów **A** i **B**

a) $A=\{a_1,a_2,a_3\}$, $B=\{b_1,b_2\}$ gdzie $a_1=\{-0.5,-0.5)$ $a_2=(0.5,-1)$ $a_3=(-1,1)$, $b_1=(1,2)$ $b_2=(-0.5,2.5)$
b) $A=\{a_1,a_2\}$, $B=\{b_1,b_2\}$ gdzie $a_1=\{-0.5,-0.5)$ $a_2=(0.5,0.5)$, $b_1=(0.3,-0.5)$ $b_2=(0,1)$
c) $A=\{a_1,a_2\}$, $B=\{b_1,b_2,b_3\}$ gdzie $a_1=\{-0.5,-0.5)$ $a_2=(-0.5,0.5)$ $b_1=(0.3,-0.5)$, $b_2=(-0.1,1)$ $b_3=(-0.9,0)$
d) $A=\{a_1,a_2\}$, $B=\{b_1,b_2\}$ gdzie $a_1=\{-1,-1)$ $a_2=(5,2)$, $b_1=(0,3)$ $b_2=(2,-2)$

2. Zbadać czy poniższe zadanie klasyfikacji możliwe są do realizacji przy użyciu sieci jednowarstwowej (jeżeli zadanie jest możliwe do realizacji wyznaczyć wagi **w,** dla których sieć będzie klasyfikował elementy czterech zbiorów **A,B,C,D**) $A=\{a_1,a_2,a_3\}$, $B=\{b_1,b_2\}$, $C=\{c_1,c_2,c_3\}$, $D=\{d_1,d_2\}$, gdzie $a_1=\{1,12)$ $a_2=(7,18)$ $a_3=(8,16)$, $b_1=(8,6)$ $b_2=(10,8)$, $c_1=\{3,5)$ $c_2=(0,2)$ $c_3=(-3,8)$, $d_1=(-5,-15)$ $d_2=(-15,-13)$

3. Wyznaczyć 10 wzorców uczących liniowo separowalnych, podzielić je na 4 klasy oraz

a) stworzyć perceptron składający się z 2 neuronów i nauczyć go klasyfikować dane wzorce do poszczególnych klas
b) zbadać zdolności uogólniające perceptronu poprzez podawanie sygnałów zbliżonych do sygnałów uczących i obserwację odpowiedzi sieci na te sygnały

4. Dane są dwa zbiory $A=\{a_1,a_2,a_3\}$, $B=\{b_1,b_2\}$ gdzie $a_1=\{-0.5,-0.5)$ $a_2=(-0.5,0.5)$ $a_3=(-80,100)$, $b_1=(0.3,-0.5)$ $b_2=(-0.1,1)$, przeprowadzić na ich podstawie uczenie perceptronu. Jakie wnioski nasuwają się po analizie procesu uczenia.