

Sieci neuronowe

Sebastian Bożek

<http://www.biofizyk.pl>

Plan prezentacji

1. Metoda obliczeń i metoda skojarzeń
2. Neuron biologiczny
3. Model sztucznego neuronu
4. Funkcja aktywacji – rodzaje sztucznych neuronów
5. Sztuczna sieć neuronowa - zapis wektorowo-macierzowy
6. Metody uczenia sieci
7. Przykłady zastosowań
 - Rozpoznawanie liter
 - Analiza widm promieniowanie
 - Prognozowanie parametrów betonu

Koncepcja sztucznych sieci neuronowych

Komputer

Pamięć

1.5740e+001
6.0000e+000
3.5000e+000
4.2484e+000
5.3758e+000
4.7091e+000
4.4983e+000
6.5373e+001

Dokładność

Błędy komputerów
to błędy programistów /
architektów

Szybkość obliczeń



<http://www.wikipedia.pl>

Brak zmęczenia

serwery

Człowiek

Inteligencja

?

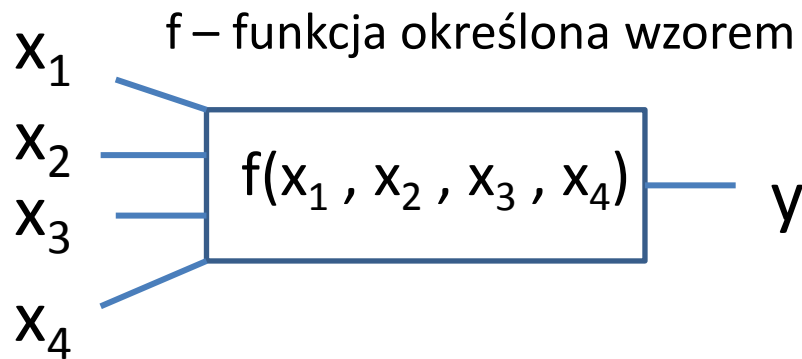
Maszyna wykonuje kod programu – algorytm.

Sztuczna inteligencja – personifikacja maszyn – rozszerzenie ich działania o obszar do tej pory zarezerwowany tylko dla organizmów żywych – kojarzenia, rozpoznawania, klasyfikacji i przewidywania na podstawie skojarzeń.

Obliczenia vs kojarzenie

Metoda obliczeniowa

- 1) Pomiary lub obliczenia
- 2) Stworzenie wzoru
- 3) Podstawianie danych do wzoru



Metoda skojarzeń

DOŚWIADCZENIE OPARTE NA
WIELU PODOBNYCH
PRZYKŁADACH



W pewnych sytuacjach bardziej efektywna wydaje się metoda skojarzeń

Przykład: kierowca rajdowy

Z jaką prędkością wejść w ten zakręt?

Metoda obliczeniowa

Dane:

- x1 – masa pojazdu
- x2 – współczynnik tarcia
 - x21 rodzaj nawierzchni
 - x22 wilgotność itd..
- x3 – docisk spoileru
- x4 – promień krzywizny zakrętu
- x5 – moc silnika
- itd..

Szukane:

Prędkość

$$V = f(x_1, x_2, x_3 \dots)$$

Metoda skojarzeń

Kierowca przejechał **WIELE**
PODOBNYCH ZAKRĘTÓW w
PODOBNYCH WARUNKACH.

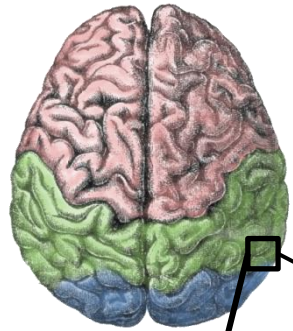
Jego wiedza bazuje głównie na

DOŚWIADCZENIU

które uwzględni dane.

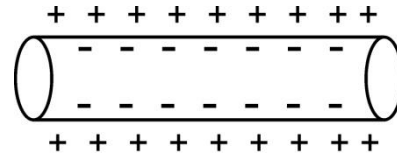
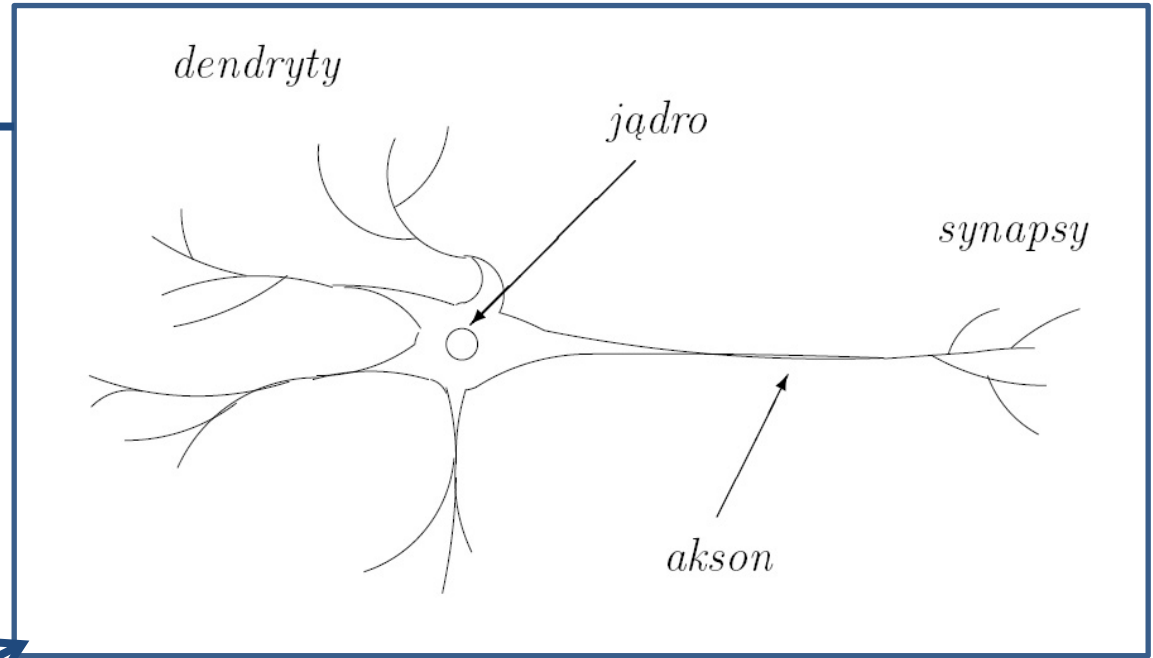
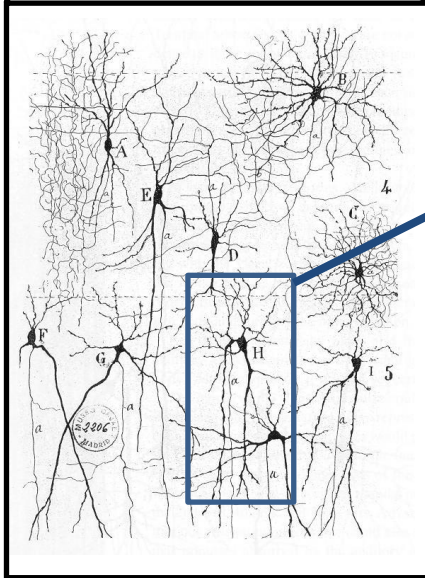


Komórka budująca mózg - NEURON

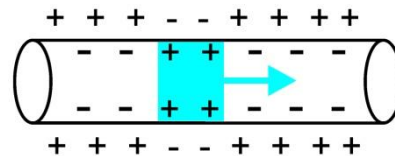


* $\sim 10^{11}$

10^{15} połączeń

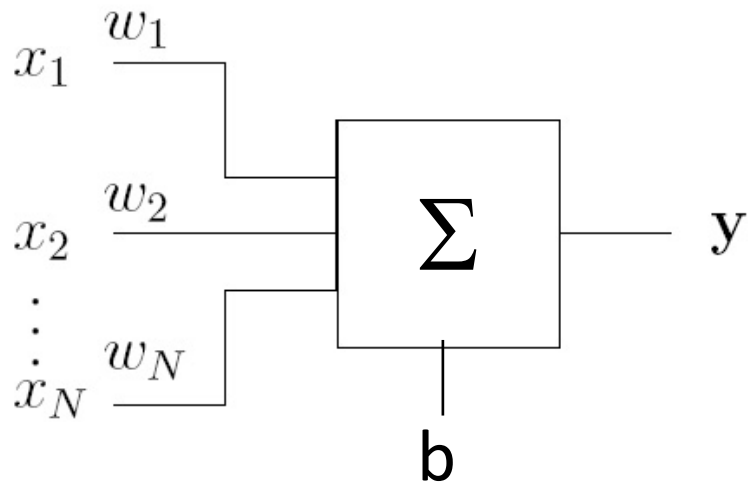
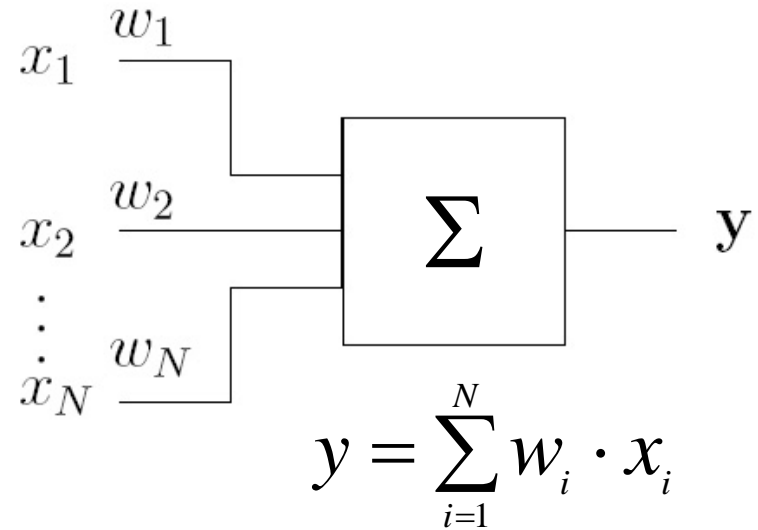
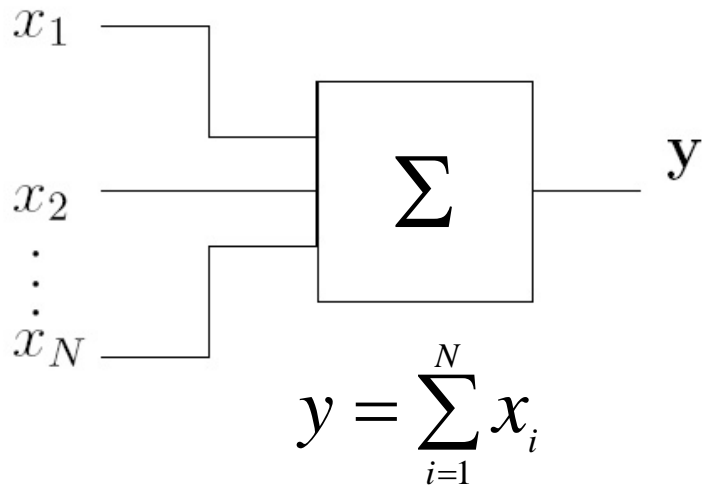


Impuls nerwowy



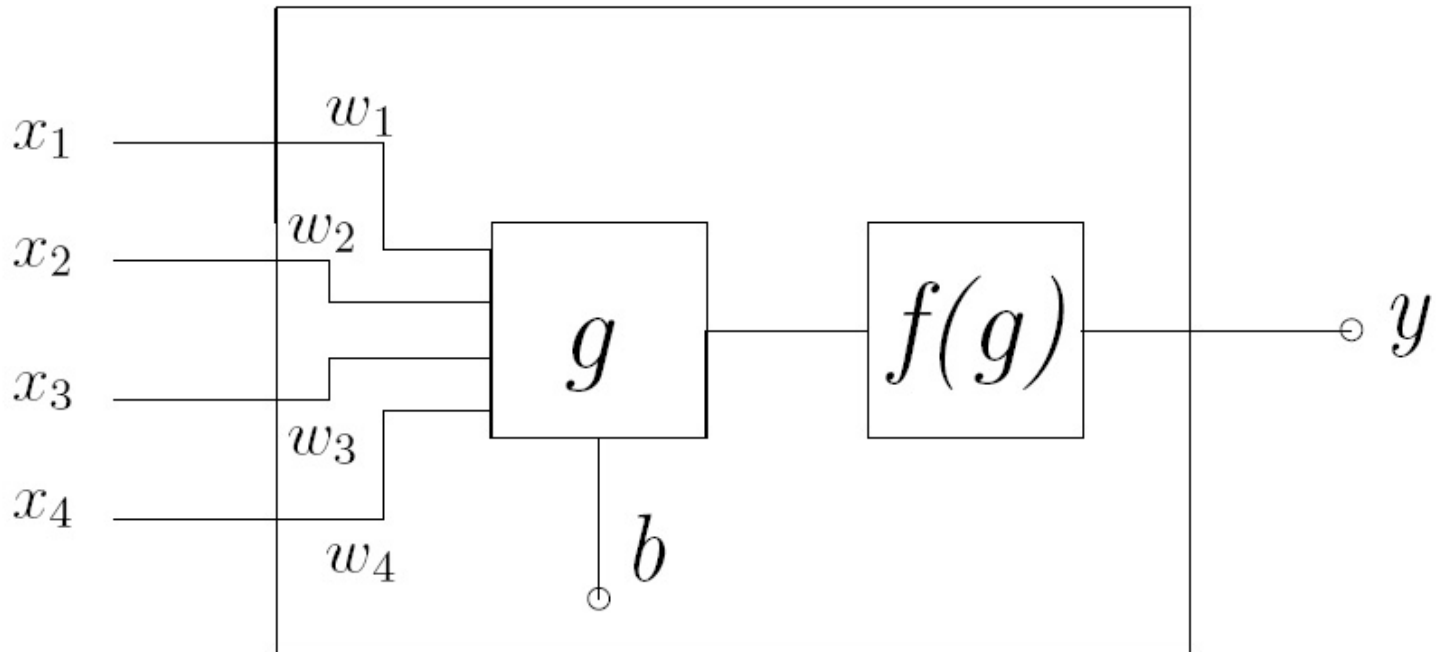
**Układ
elektroniczny ?**

Prosty model neuronu



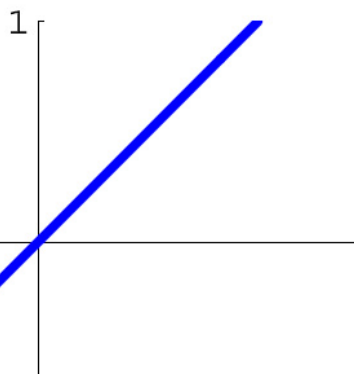
$$y = \sum_{i=1}^N w_i \cdot x_i - b$$

Funkcja aktywacji



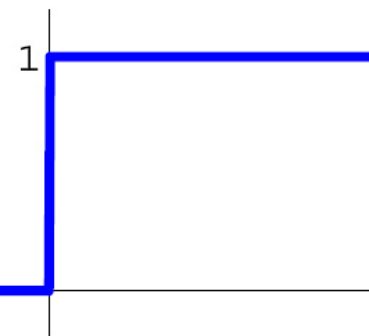
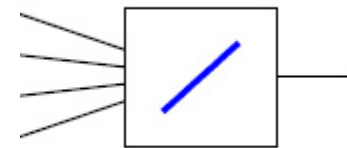
$$g = \sum_{i=1}^N w_i \cdot x_i - b \quad y = f(g)$$

Rodzaje sztucznych neuronów



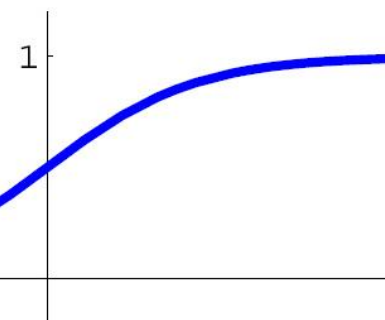
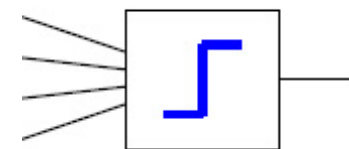
Neuron liniowy

$$f(g) = g$$



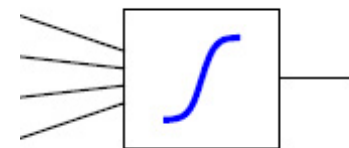
Perceptron

$$f(g) = \begin{cases} 0, & g \leq 0 \\ 1, & g > 0 \end{cases}$$



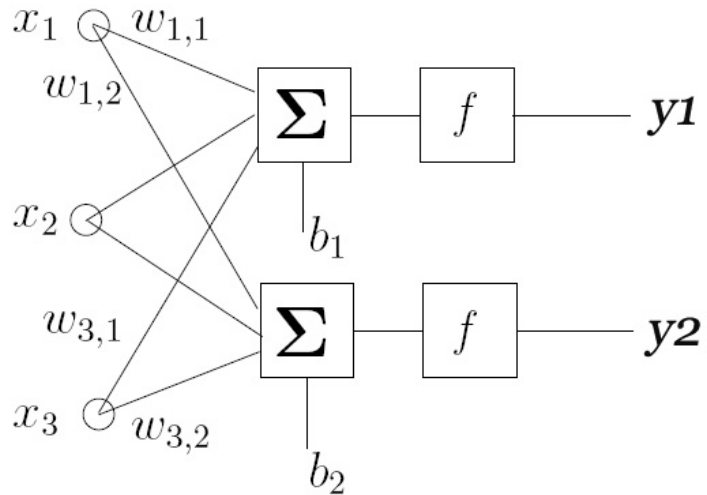
Neuron sigmoidalny (funkcja logistyczna)

$$f(g) = \frac{1}{1 + \lambda \cdot e^{-g}}$$

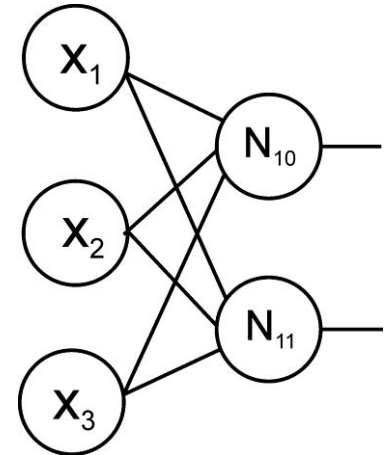


Sztuczna sieć neuronowa - SSN

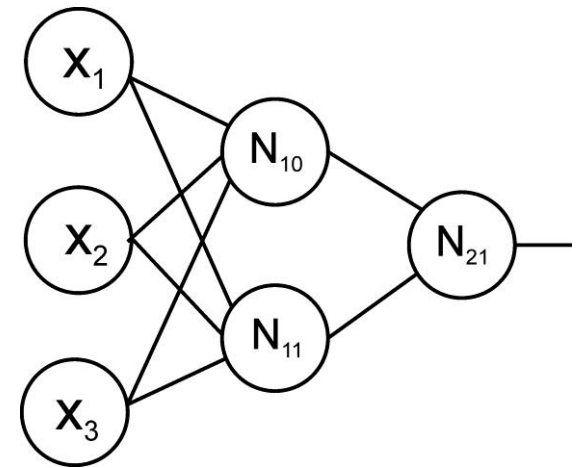
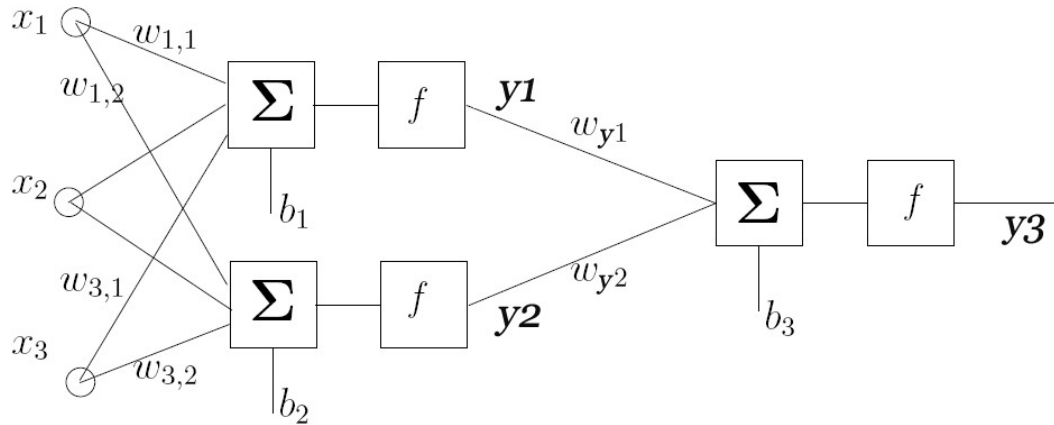
ANN – Artificial Neural Network



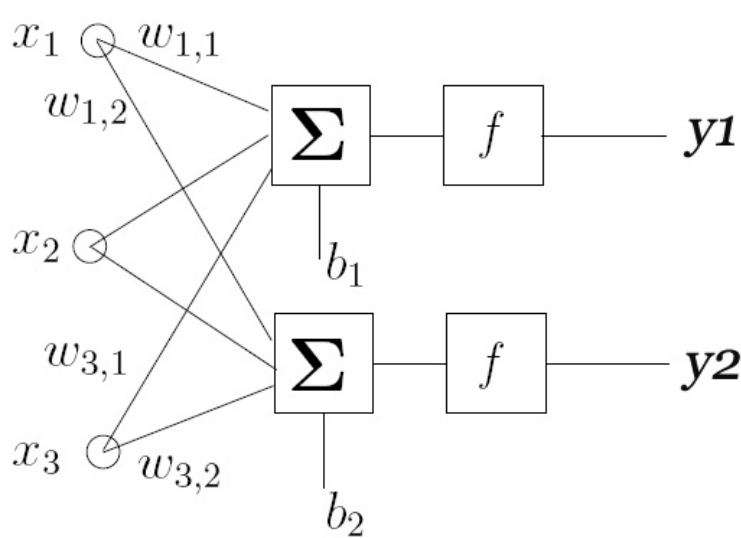
Jednowarstwowa
(single layer)



Wielowarstwowa
(multilayer)

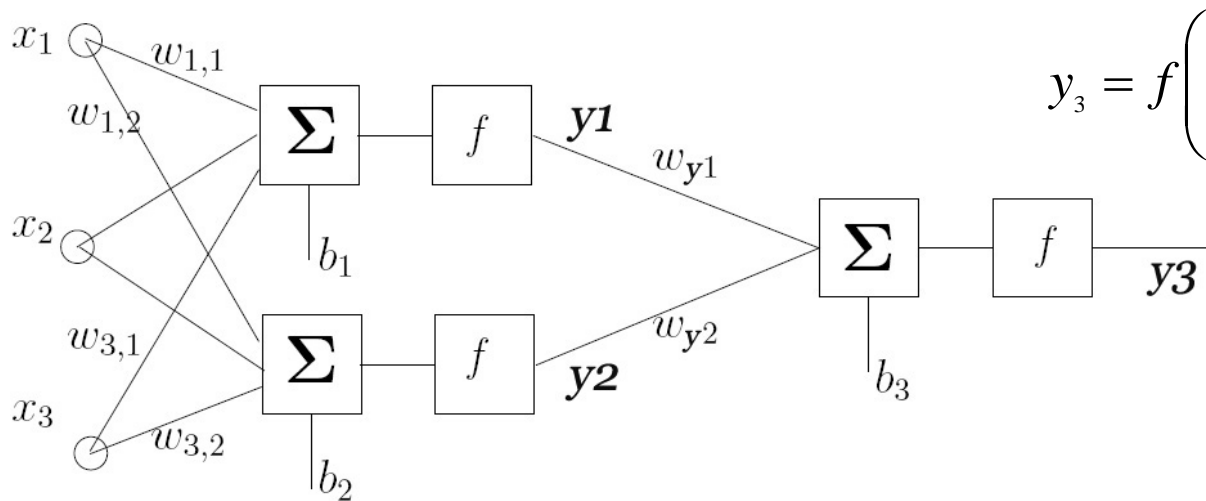


Zapis wektorowo-macierzowy



$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} \\ w_{2,1} & w_{2,2} \\ w_{3,1} & w_{3,2} \end{bmatrix} \quad \mathbf{y} = f(\mathbf{W}^T \cdot \mathbf{x} + \mathbf{b})$$



$$y_3 = f\left(\begin{bmatrix} w_{y1} & w_{y2} \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + b_3\right)$$

Rodzaje uczenia sieci neuronowych

Uczenie – przystosowywanie się sieci do zadanego zbioru (ciągu) uczącego

Uczenie z nauczycielem

Znana jest pożądana odpowiedź sieci na przedstawiony sygnał wejściowy.

Ciąg uczący to ciąg danych wejściowych (inputs) i skojarzonych z nimi pożądanymi odpowiedziami (targets).

Przykład: wychowanie

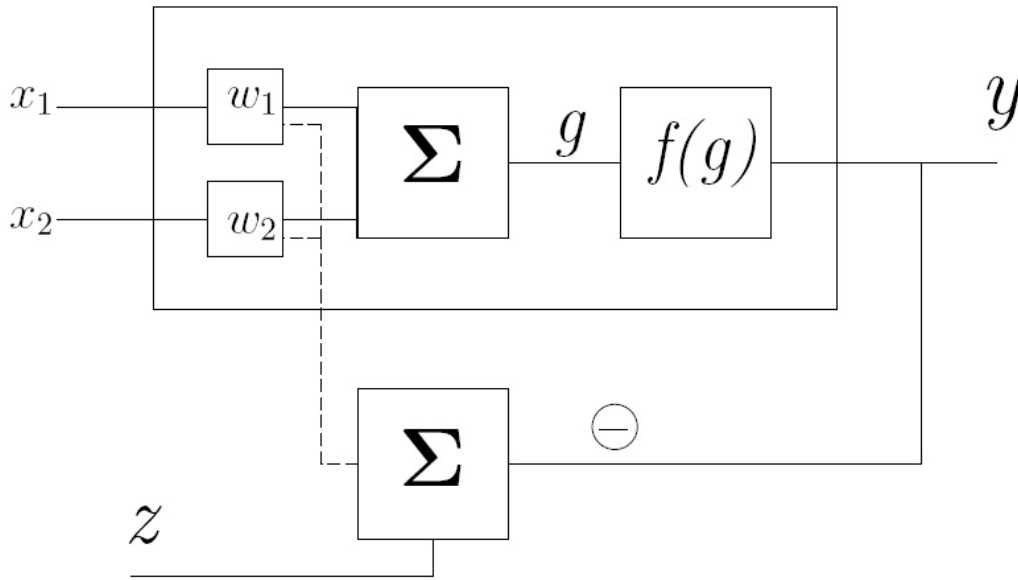
Uczenie bez nauczyciela

Pożądana odpowiedź (targets) nie jest znana. Samouczenie polega na utrwalaniu charakterystycznych cech sieci.

Przykład:

„Krystalizowanie się”
światopoglądu.

Uczenie z nauczycielem – reguła „Delta”



Ciąg uczący $n=1,2,\dots,L$

$$(\mathbf{x}_n, z_n) \quad y_n$$

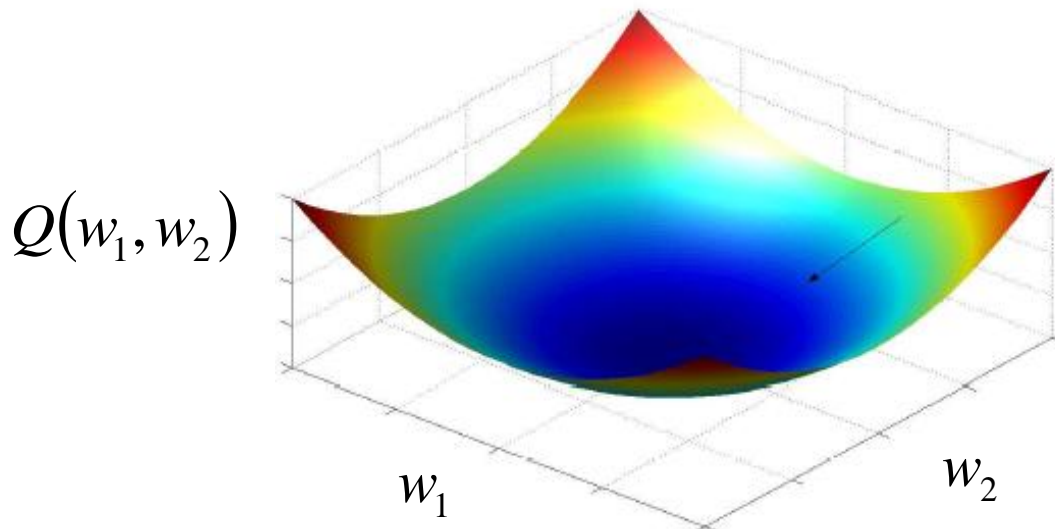
Funkcja błędu

$$Q = \frac{1}{2} \sum_{j=1}^L (z_j - y_j)^2$$

$$Q_j = \frac{1}{2} (z_j - y_j)^2$$

$$\Delta w_{ij} = -\eta \frac{\partial Q_j}{\partial w_{ij}}$$

η - stała uczenia



Metoda momentum

$$\eta \frac{\partial Q_j}{\partial w_{ij}} \left\{ \begin{array}{l} \text{Za małe – powolne uczenie} \\ \text{Za duże – oscylacje wokół minimum} \end{array} \right.$$

Metoda **momentum** – korekta wag według reguły

$$\Delta w_{ij} = -\eta \frac{\partial Q_j}{\partial w_{ij}} + \alpha \cdot \Delta w_{i(j-1)}$$

Nadanie punktowi reprezentującemu wagi pewnego momentu bezwładności. Gdy kolejne gradienty wskazują ten sam kierunek, ich działanie się kumuluje i przyrosty wag stają się coraz większe. Gdy gradienty są przeciwne, ruch punktu jest hamowany.

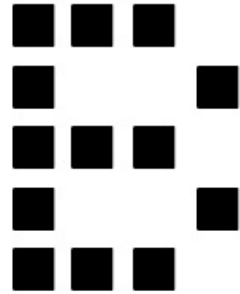
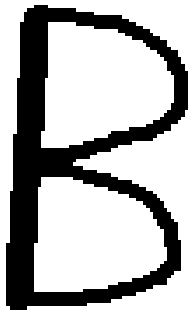
Przykład 1

Rozpoznawanie i klasyfikacja

Sieć rozpoznająca litery alfabetu

R. Tadeusiewicz, *Sieci neuronowe*, Akademicka Oficyna Wydawnicza, 1993

Wektorowy zapis obrazu litery

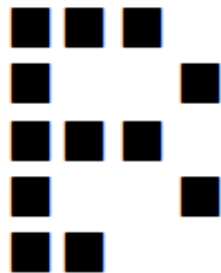


1	1	1	0
1	0	0	1
1	1	1	0
1	0	0	1
1	1	1	0
1	1	1	0



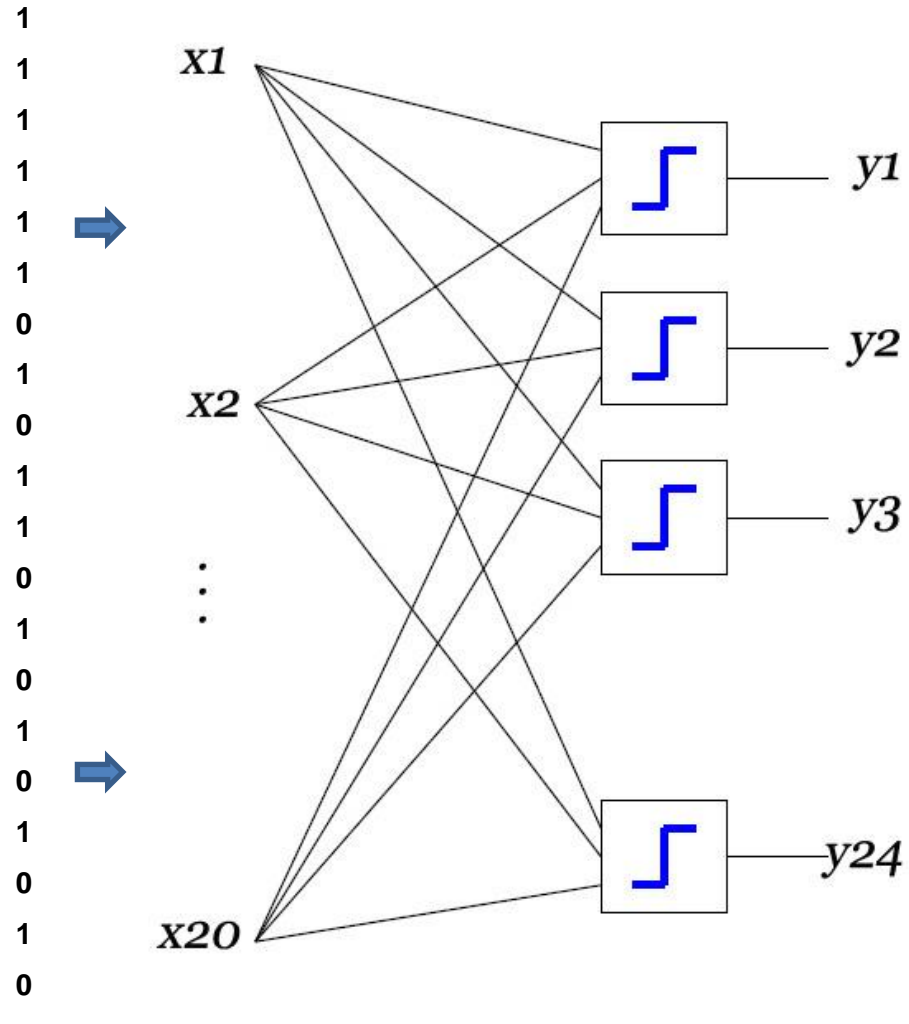
$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

$$\mathbf{B} = [1 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 0]^T$$



?

B



$\mathbf{Y}(\mathbf{B})$

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,24} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,24} \\ \vdots & \vdots & & \vdots \\ w_{20,1} & w_{20,2} & \cdots & w_{20,24} \end{bmatrix}$$

$$\mathbf{Y}(\mathbf{B}) = f(\mathbf{W}^T \cdot \mathbf{B})$$

$$f(g) = \begin{cases} 0, & g \leq 0 \\ 1, & g > 0 \end{cases}$$

Przykład 2:

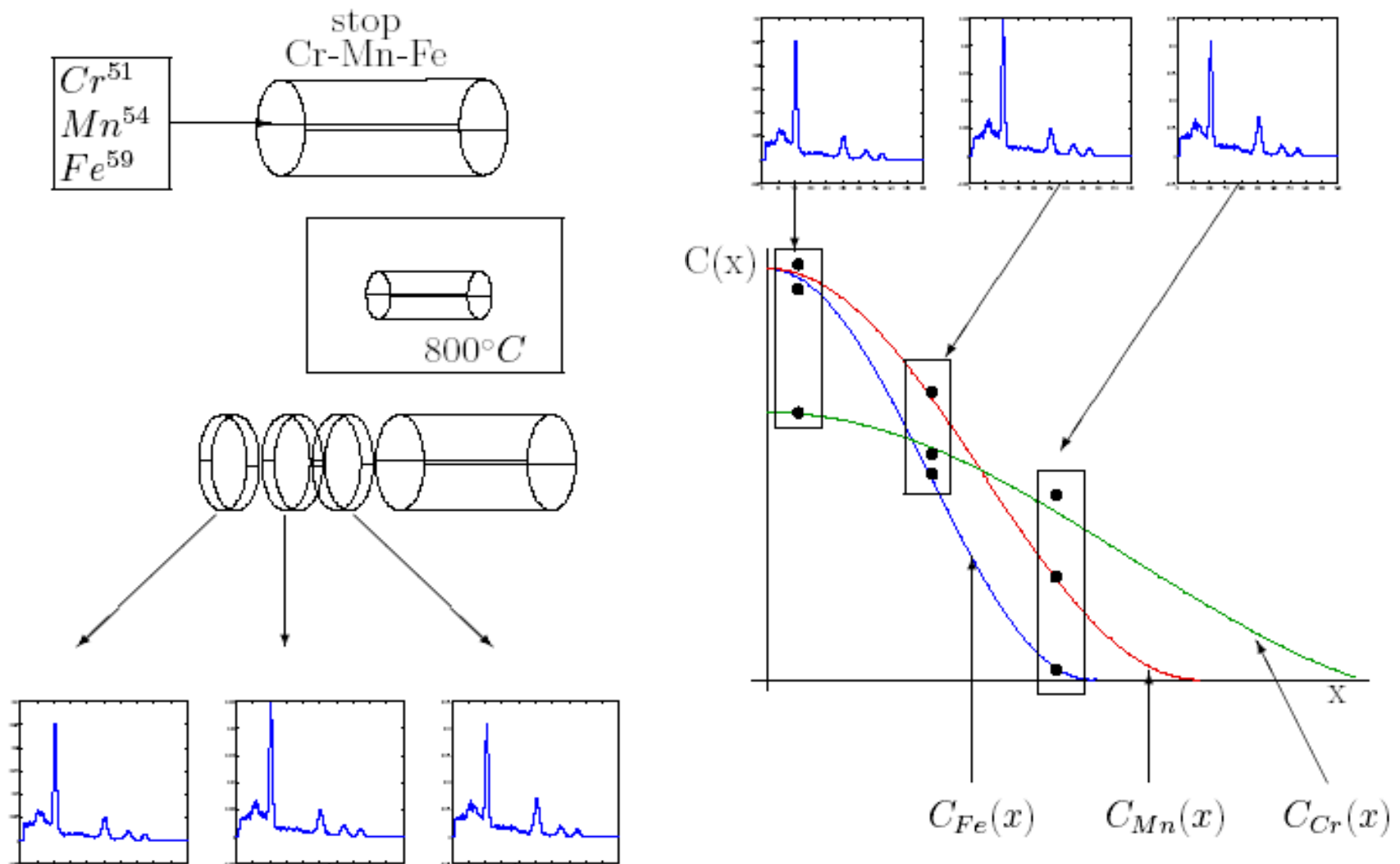
Prognozowanie

Sieć wyznaczająca procentowe udziały poszczególnych izotopów promieniotwórczych na podstawie widm próbek zmierzonych detektorem scyntylacyjnym

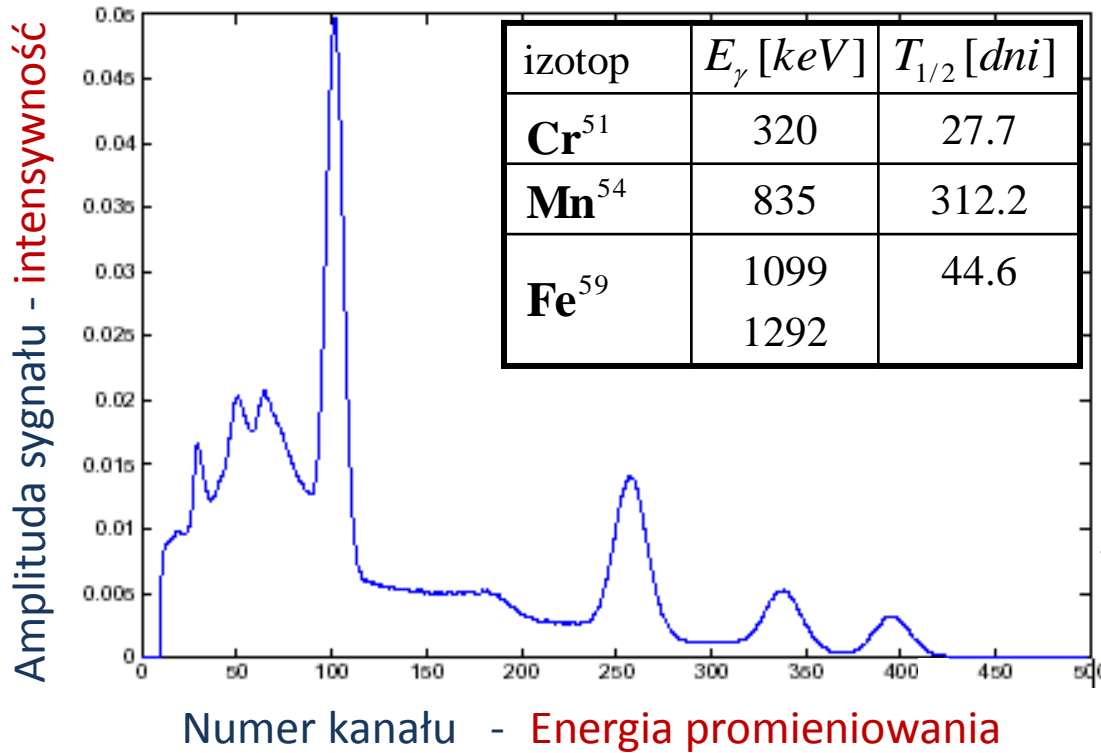
S.Bożek, Oznaczanie składu gamma promieniotwórczych izotopów metodą sztucznych sieci neuronowych, praca magisterska, AGH 2003

Dyfuzja znaczników promieniotwórczych

Badanie dyfuzji poszczególnych pierwiastków w stali manganowo-chromowej z wykorzystaniem znaczników promieniotwórczych

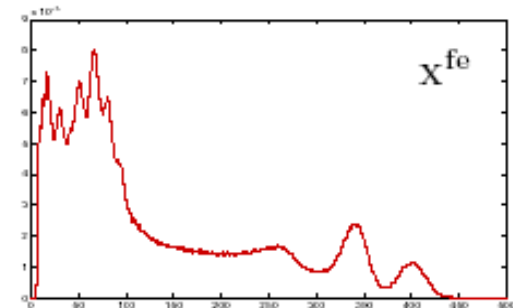
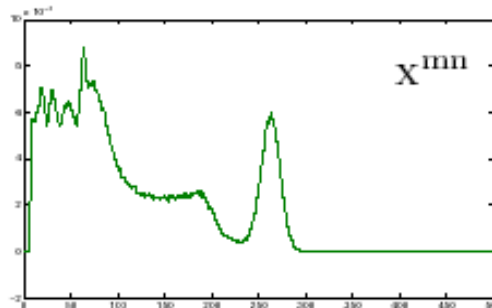
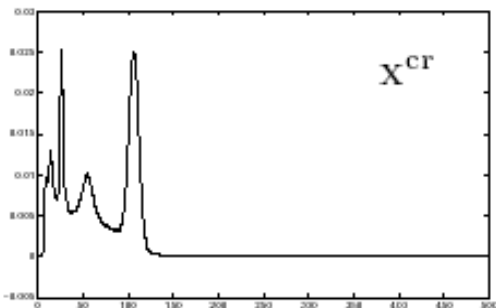


Widmo mierzone detektorem scyntylacyjnym

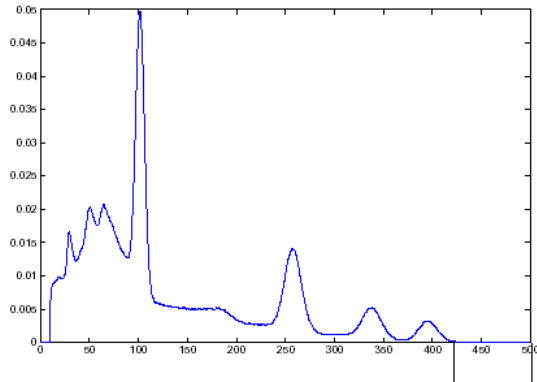


Detektor scyntylacyjny cechuje się słabą zdolnością rozdzielczą, jednak jest tańszy od dokładniejszych detektorów półprzewodnikowych

Widmo całości jest pewną kombinacją widm składowych



Wektorowy opis widma



k kanałów detektora

plik.txt

0.009475
0.009320
0.009602
0.009707
0.010257
0.011660
0.013819
⋮

kolumna
k liczb

Wektor w **k**-wymiarowej przestrzeni

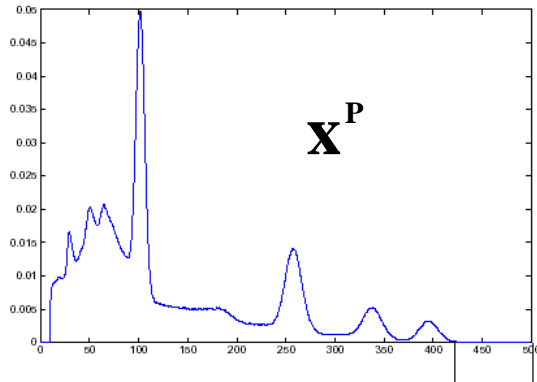
$$\mathbf{X}^P = (X_1^P, X_2^P, \dots, X_k^P)$$

Normalizacja

$$\mathbf{X}^P = \frac{\mathbf{X}^P}{\sum_{i=1}^k X_i^P}$$

$$\int_0^{\infty} f(x) dx = 1$$

Wersory widmowe



$$\mathbf{X}^{Cr} = \frac{\mathbf{X}^{Cr}}{\sum_{i=1}^k X_i^{Cr}}$$

$$\mathbf{X}^{Mn} = \frac{\mathbf{X}^{Mn}}{\sum_{i=1}^k X_i^{Mn}}$$

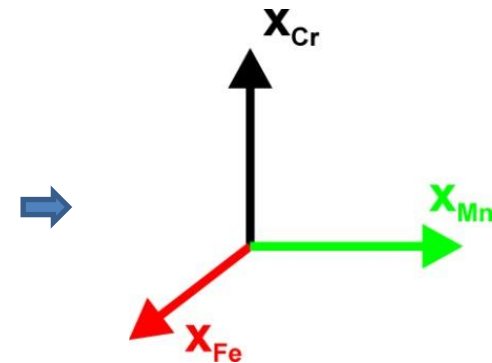
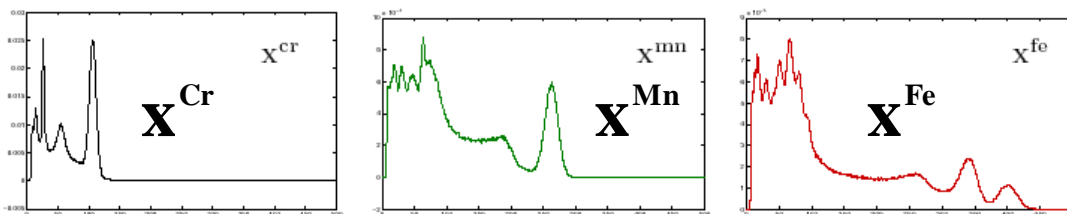
$$\mathbf{X}^{Fe} = \frac{\mathbf{X}^{Cr}}{\sum_{i=1}^k X_i^{Cr}}$$

c_{Cr}, c_{Mn}, c_{Fe} – stężenia %

$$c_{Cr} + c_{Mn} + c_{Fe} = 1$$

$$\mathbf{X}^P = c_{Cr} \mathbf{X}^{Cr} + c_{Mn} \mathbf{X}^{Mn} + c_{Fe} \mathbf{X}^{Fe}$$

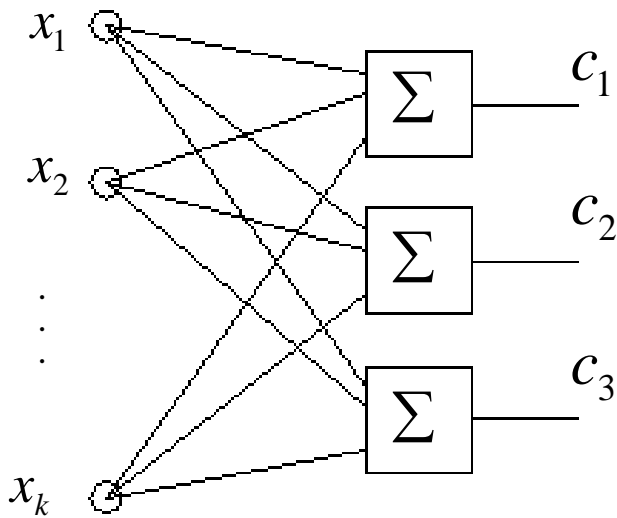
Wektor widmowy próbki jest sumą wektorów widmowych składowych izotopów



Sieć jednowarstwowa

Dla licznika posiadającego k kanałów sieć ma k wejść (inputs)

Liczba neuronów, i tym samym liczba wyjść (outputs), odpowiada liczbie analizowanych izotopów

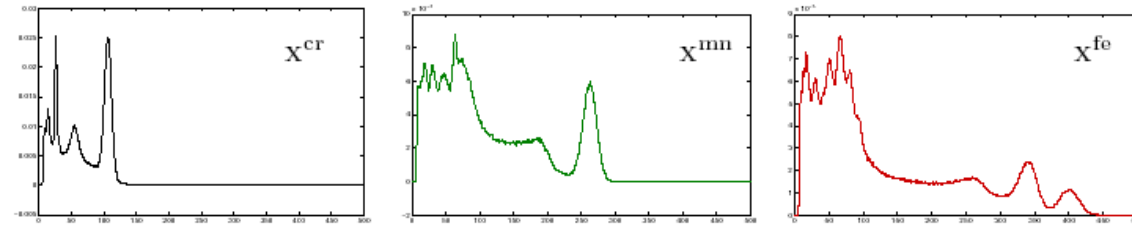


$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ x_k \end{bmatrix}$$

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & w_{31} \\ w_{21} & w_{22} & w_{32} \\ \cdot & \cdot & \cdot \\ w_{k1} & w_{k2} & w_{k3} \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

Ciąg uczący



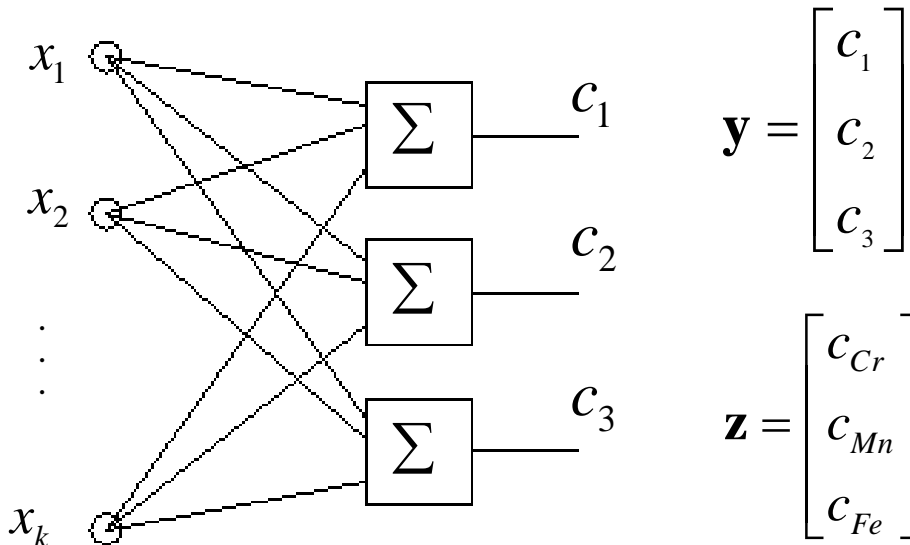
$$c_{cr} = rand$$

$$c_{mn} = (1 - c_{cr}) rand$$

$$c_{fe} = 1 - c_{cr} - c_{mn}$$

Współczynniki stężeń w ciągu uczącym są generowane przy użyciu generatora liczb losowych

$$\mathbf{X}^P = c_{Cr} \mathbf{X}^{Cr} + c_{Mn} \mathbf{X}^{Mn} + c_{Fe} \mathbf{X}^{Fe}$$



$$Q_j = \frac{1}{2} (z_j - y_j)^2$$

$$\Delta w_{ij} = -\eta \frac{\partial Q_j}{\partial w_{ij}}$$

Przykład 3:

Prognozowanie

Sieć prognozująca fizyczne właściwości betonu

Wspomagane komputerowo prognozowanie fizycznych właściwości betonu o wysokiej wytrzymałości, zawierającego nanocząstki Fe_2O_3 , Cement, wapno beton, wrzesień-październik 2012 r., Nr 5, www.cementwapnobeton.pl

Computed-aided prediction of physical and mechanical properties of high strength concrete containing Fe_2O_3 nanoparticles

Wytrzymałość betonu

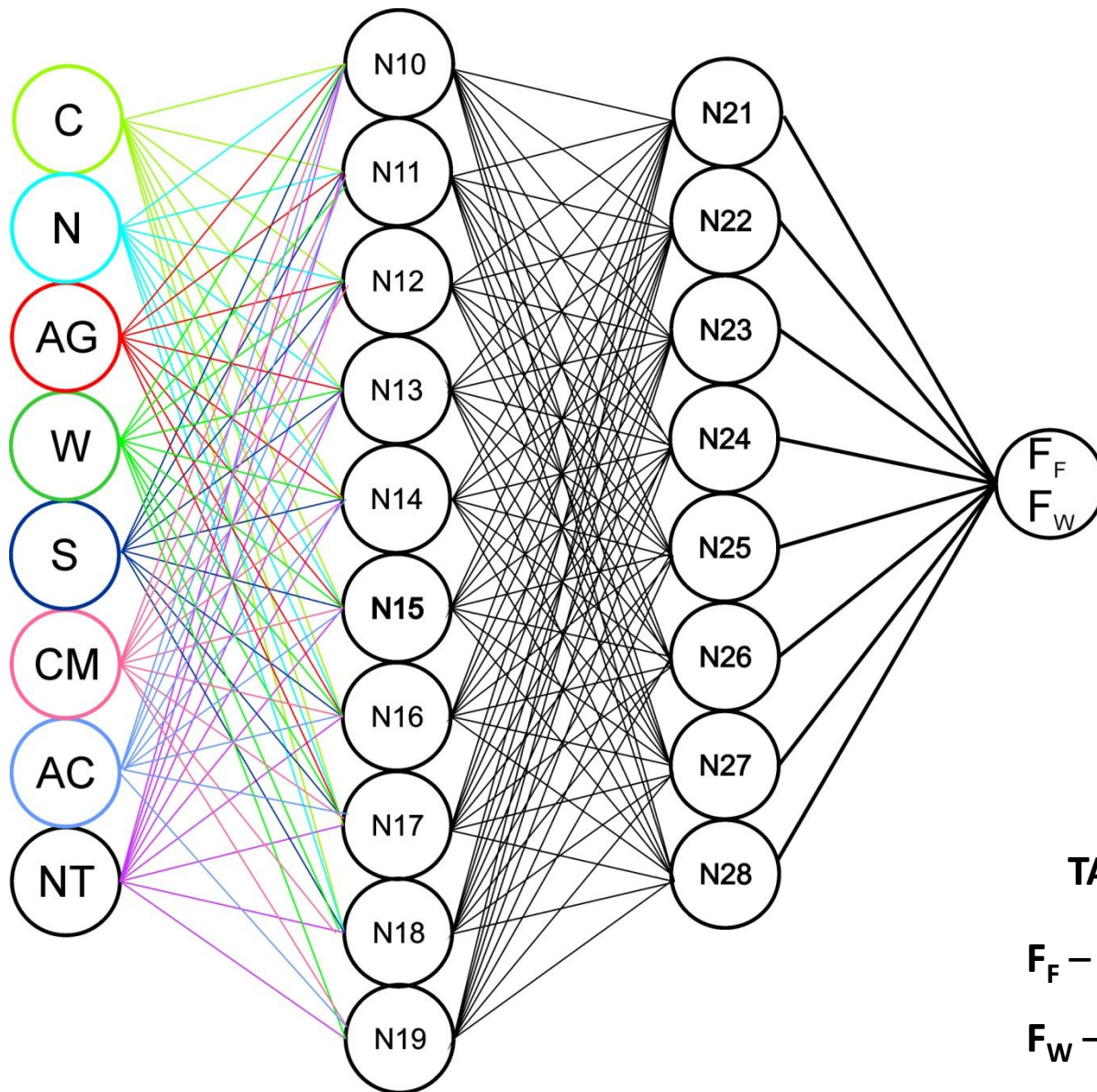
Wytrzymałość – najważniejsza cecha betonu.

Mierzona jest zazwyczaj po normowym czasie dojrzewania.

Wytrzymałość betonu zależy od wielu czynników.

Do prognozowania wytrzymałości na zgniecie oraz nasiąkliwości betonu zastosowano 2 warstwową sieć neuronową z sigmoidalną funkcją aktywacji neuronów.

Architektura sieci



INPUTS

- C** – zawartość cementu
- N** – zawartość nanocząstek
- AG** – rodzaj kruszywa
- W** – zawartość wody
- S** – ilość superplastyfikatora
- CM** – rodzaj środowiska dojrzewania
- AC** – czas dojrzewania
- NT** – liczba pokazów

TARGETS

- F_F** – wytrzymałość na zginanie
- F_W** – nasiąkliwość

Ciąg uczący / testujący

I N P U T S								T A R G E T S	
zawartość cementu [kg/m ³]	zawartość nanocząstek [kg/m ³]	rodzaj kruszywa	zawartość wody [kg/m ³]	zawartość superplastyfikatora [kg/m ³]	środowisko dojrzewania	czas dojrzewania	liczba pokazów	wytrzymałość na zginanie	nasiąkliwość
450	0	3	0	18	1	7	3	4,6	3,6
450	0	3	0	18	1	28	3	4,6	5,9
447,8	2,25	3	0	18	1	7	2	4,7	3,9
447,8	2,25	3	0	18	1	28	3	5,1	2,3
447,8	2,25	3	0	18	1	90	3	5,9	0,9
445,5	4,5	3	0	18	1	7	3	5,6	4,6
445,5	4,5	3	0	18	1	28	1	5,1	2,5
445,5	4,5	3	0	18	1	90	2	5,8	1,1
443,3	6,75	3	0	18	1	7	2	5,1	4,7
443,3	6,75	3	0	18	1	28	1	5,1	2,8
441	9	3	0	18	1	7	1	3,9	4,7
441	9	3	0	18	1	90	1	4,8	1,5
450	0	3	0	18	2	90	2	4,4	5,1
447,8	2,25	3	0	18	2	7	2	5,4	6,3
447,8	2,25	3	0	18	2	28	1	5,7	1,5
445,5	4,5	3	0	18	2	7	3	5,9	6,7
443,3	6,75	3	0	18	2	90	1	5,9	1,1
441	9	3	0	18	2	28	2	6,9	1,9
450	0	4	0,18	17,82	1	7	2	4,1	4,7
445,5	4,5	4	0,18	17,82	1	7	2	4	1,7
445,5	4,5	4	0,18	17,82	1	28	1	4,3	1,6
441	9	4	0,18	17,82	1	7	2	4,6	1,8
441	9	4	0,18	17,82	1	28	3	6,2	1,7
437,5	13,5	4	0,18	17,82	1	28	1	6,3	1,2
432	18	4	0,18	17,82	1	7	2	5,7	1,4
427,5	22,5	4	0,18	17,82	1	2	2	3,7	4,4
427,5	22,5	4	0,18	17,82	1	28	1	6,6	1,1

Wyniki

WARTOŚCI EKSPERYMENTALNE I ODPOWIEDZI SIECI NEURONOWEJ					
Wytrzymałość na zginanie			Nasiąkliwość		
Eksperyment	SSN-1	SSN-2	Eksperyment	SSN-1	SSN-2
4,6	4,6	4,7	3,6	3,7	3,6
4,6	4,8	4,5	5,9	5,9	5,9
4,7	5,4	4,8	3,9	4,1	4,0
5,1	5,2	5,2	2,3	2,6	2,4
5,9	6,5	5,9	0,9	1,0	1,0
5,6	5,6	5,6	4,6	4,8	4,6
5,1	5,3	5,3	2,5	2,3	2,6
5,8	5,9	5,8	1,1	0,6	1,1
5,1	5,1	5,0	4,7	4,8	4,9
5,1	5,0	5,2	2,8	2,8	2,8
3,9	3,8	4,0	4,7	4,8	4,6
4,8	4,8	4,8	1,5	1,6	1,8
4,4	4,2	4,4	5,1	5,1	5,0
5,4	5,3	5,0	6,3	6,3	6,1
5,7	5,6	5,3	1,5	2,5	1,6
5,9	5,8	5,8	6,7	6,9	7,2
5,9	5,9	6,0	1,1	1,2	1,1
6,9	6,9	6,9	1,9	1,8	1,9
4,1	3,9	4,0	4,7	4,3	4,3
4,0	4,2	4,1	1,7	1,8	1,8
4,3	4,4	4,4	1,6	1,6	1,7
4,6	4,4	4,7	1,8	1,7	1,7
6,2	5,9	6,1	1,7	1,7	1,4
6,3	6,4	6,4	1,2	1,2	1,2
5,7	5,6	5,6	1,4	1,6	1,4
3,7	3,8	3,7	4,4	4,9	4,5
6,6	6,7	7,3	1,1	1,1	1,2

Eksperyment

Wartości uzyskane doświadczalnie

SSN-1

Uczenie sieci neuronowej zostało zakończone dla zadanej liczby powtórzeń ciągu uczącego (1000)

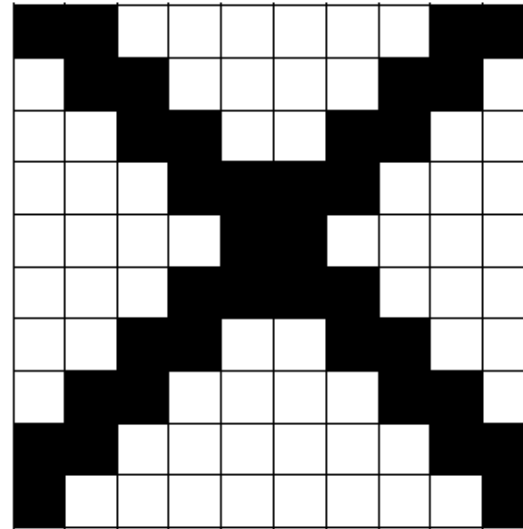
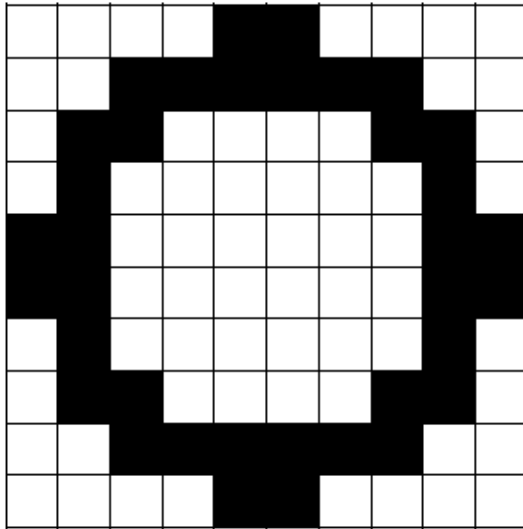
SSN-2

Uczenie sieci neuronowej zostało zakończone po osiągnięciu ustalonego minimalnego błędu sieci.

- J. Żurada, M. Barski, W. Jędruch, ***Sztuczne sieci neuronowe***, PWN 1996
- S. Ossowski, ***Sieci neuronowe do przetwarzania informacji***, Wyd. Politechniki Warszawskiej 2006
- R. Tadeusiewicz, ***Sieci neuronowe***, Akademicka Oficyna Wydawnicza, 1993
- <http://www.cs.put.poznan.pl/rklaus/assn/neuron.htm>

Zagadnienia do egzaminu – część 1/3

1. Wymień i scharakteryzuj rodzaje sztucznych neuronów.
2. Twój zespół opracowuje komputerowy program do gry w kółko i krzyżyk. Znaki kółka i krzyżyka w rozdzielczości 10x10 (rysunek) grający będzie rysował na tablecie. Twoim zadaniem jest zbudowanie sieci neuronowej odróżniającej kółko od krzyżyka. Zaproponuj najprostszы model sieci neuronowej realizującej takie zadanie. Dołącz rysunek przedstawiający architekturę tej sieci.



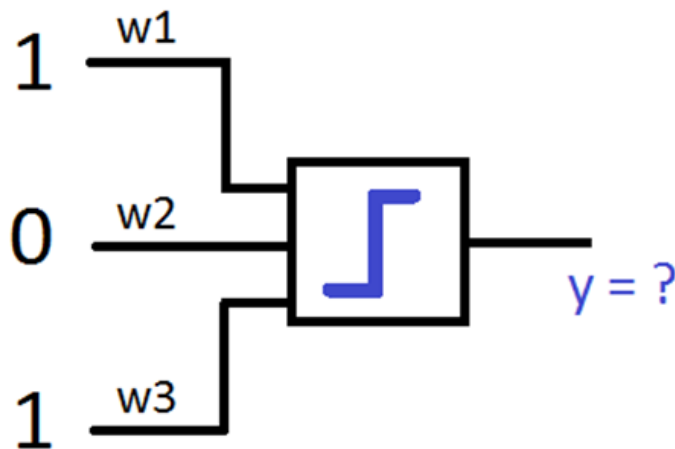
Zagadnienia do egzaminu – część 2/3

3. Na wejście perceptronu (rysunek) trafia sygnał $[1 \ 0 \ 1]^T$

Współczynniki wagowe neuronu wynoszą odpowiednio:

$w_1 = 0.5$, $w_2 = 0.3$, $w_3 = -0.2$.

Oblicz odpowiedź y perceptronu.



4. Przedstaw klasyczny schemat uczenia neuronu (regułę „delta”) oraz metodę „momentum”.

5. W kontekście sztucznych sieci neuronowych wyjaśnij następujące pojęcia:

- a) sygnały wejściowe (dane), (input variable)
- b) warstwa wejściowa, (input layer)
- c) warstwa ukryta, (hidden layer)
- d) warstwa wyjściowa, (output layer)
- e) sygnały wyjściowe (wyniki), (output)
- f) wartość progowa, (bias)
- g) zadana (wymagana) odpowiedź neuronu na sygnał, (target)
- h) ciąg uczący, (training set)
- i) ciąg testujący, (testing set)
- j) epoka, (epoch)
- k) waga neuronu, (weight)
- l) szybkość uczenia, (learning rate)
- m) momentum, (momentum factor)