

Metody inteligencji obliczeniowej w analizie danych

Sieci neuronowe I

Jan Karwowski

Wydział matematyki i Nauk Informacyjnych PW

28 lutego 2022



Fundusze Europejskie
Wiedza Edukacja Rozwój



Rzeczpospolita
Polska

**Politechnika
Warszawska**

Unia Europejska
Europejski Fundusz Społeczny



Zadanie 10 pn. „Modyfikacja programów studiów na kierunkach prowadzonych przez Wydział Matematyki i Nauk Informatycznych” realizowane jest w ramach projektu „NERW 2 PW. Nauka – Edukacja – Rozwój – Współpraca” współfinansowanego ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego

- 1 Informacje organizacyjne
- 2 Inteligencja obliczeniowa
- 3 Inspiracje biologiczne sieci neuronowych
- 4 Model sztucznej sieci neuronowej

- 15 wykładów
 - 1, 2, 7, 8, 12, 13, 14, 15 – Janek Karwowski
 - 3, 4, 5, 6, 9, 10, 11 – Adam Żychowski
- Laboratoria (co tydzień)
 - Mikołaj Małkiński
 - Maciej Żelaszczyk
 - Adam Żychowski
- Egzamin (JK, AŻ)

Sztuczna Inteligencja

Inteligencja Obliczeniowa (Computational Intelligence)

Gałąź sztucznej inteligencji zajmująca się projektowaniem i analizą systemów umiających naśladować inteligentne zachowania w złożonych i zmiennych systemach. Często są inspirowane mechanizmami znanymi z biologii.

Także: Soft computing

- ❶ Sieci neuronowe
- ❷ Algorytmy ewolucyjne
- ❸ Logika rozmyta
- ❹ Podejścia hybrydowe – neuroewolucja, wnioskowanie neuronowo-rozmyte
- ❺ Algorytmy rojowe
- ❻ Rozwiązywanie gier i dynamicznych procesów sterowania

- Klasyfikacja
- Regresja
- Analiza skupień
- Wnioskowanie

• Optymalizacja

Wszystkie tematy

Andries P Engelbrecht. *Computational intelligence: an introduction*. John Wiley & Sons, 2007

Stanisław Osowski. *Sieci neuronowe do przetwarzania informacji*. Oficyna Wydawnicza Politechniki Warszawskiej, 2004

Robert Kosiński. *Sztuczne sieci neuronowe: dynamika nieliniowa i chaos*. Wydawnictwa Naukowo-Techniczne, 2008

Zbigniew Michalewicz. *Genetic algorithms+ data structures= evolution programs*. Springer Science & Business Media, 2013

Jarosław Arabas. *Wykłady z algorytmów ewolucyjnych*. Wydawnictwa Naukowo-Techniczne, 2001

James Kennedy et al. *Swarm intelligence*. Morgan Kaufmann, 2001

Trudności

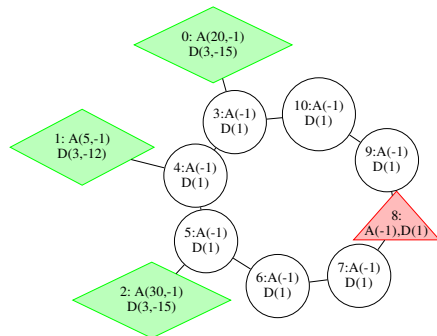
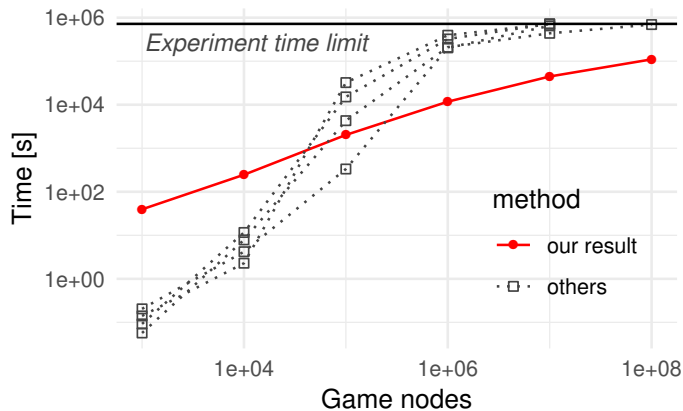
- Skomplikowany opis
- Brak dowodów zbieżności (przy realistycznych założeniach)
-

Zalety

- Skalują się do bardzo dużych rozmiarów problemów
- Można łatwo sterować dokładnością przybliżenia
-

Problem Komiwojażera (travelling salesman) – planowanie trasy.

Przykład2



Heurystyka

Sposób (algorytm, strategia) rozwiązania problemu oparty na intuicji, przeczuciu i obserwacji zagadnienia, który nie jest poparty matematycznym dowodem otrzymania optymalnego rozwiązania problemu.

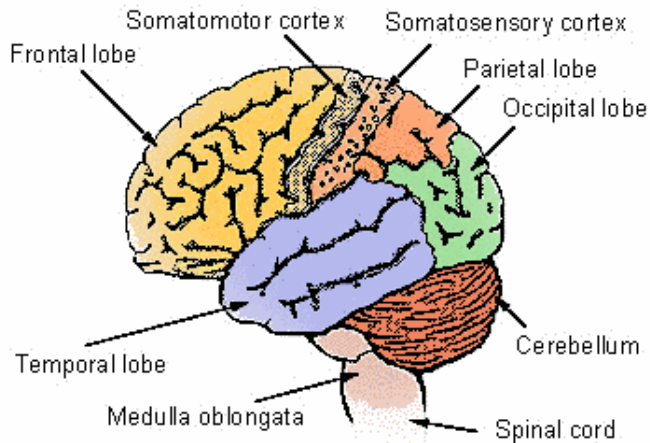
Często stosowana w problemach, których dokładne rozwiązanie jest czasochłonne lub niemożliwe.

Metaheurystyka

Sposób rozwiązania *dużej* klasy problemów (np. problemy optymalizacyjne, poszukiwanie optymalnej strategii w dowolnej grze) oparty na intuicji, przeczuciu, obserwacji, który nie jest poparty matematycznym dowodem otrzymania optymalnego rozwiązania problemu.

Heurystyka niewykorzystująca cech konkretnego zadania.

Naśladowanie struktury mózgu



Lobes of the cerebrum

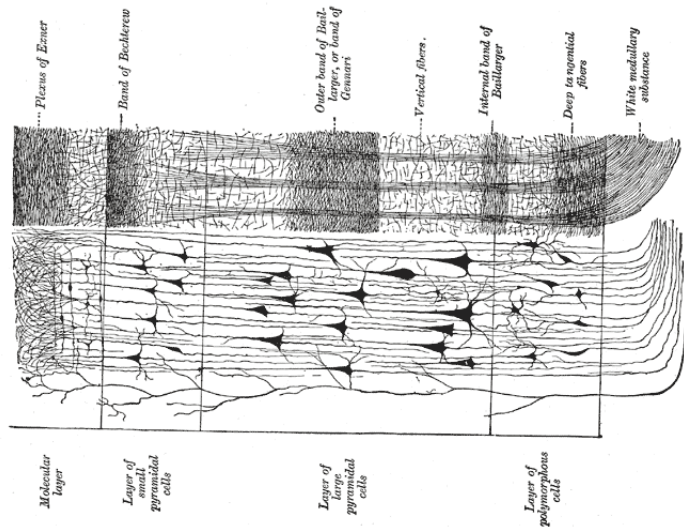


Rysunki:

https://wpclipart.com/medical/anatomy/brain/brain_3/cerebrum_lobes.png.html

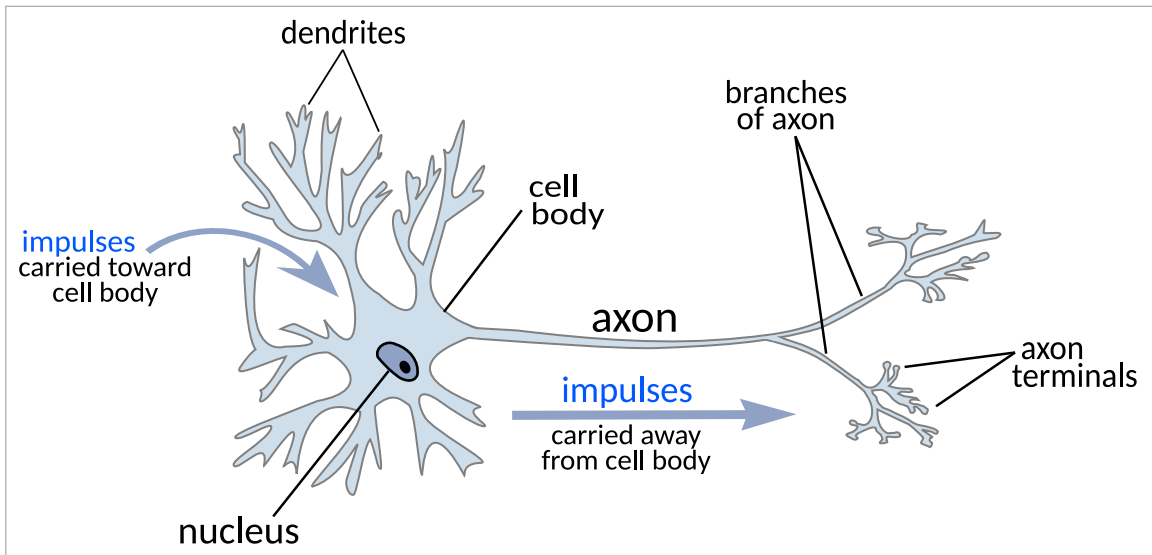
https://commons.wikimedia.org/wiki/File:IBM_704_mainframe.gif

Budowa mózgu

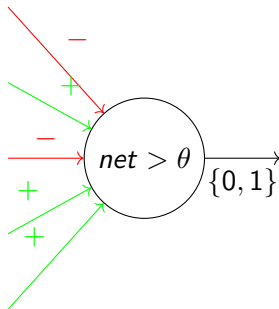


Źródło: <https://commons.wikimedia.org/wiki/File:Gray754.png>

Neuron



Źródło: https://wpclipart.com/medical/anatomy/nervous_system/neuron/



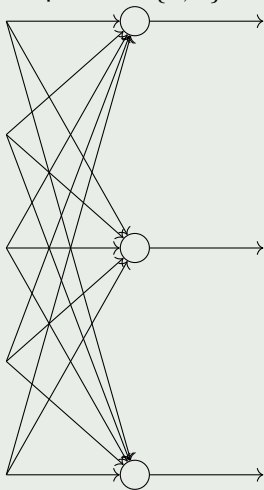
McCulloch and Pitts 1943

- wejścia $z = [z_1, z_2, \dots, z_n] \in \{0, 1\}^n$
- typ synapsy $v = [v_1, v_2, \dots, v_n] \in \{-1, 1\}^n$
- wyjście $o \in \{0, 1\}$
- próg $\theta \in \mathbb{R}$

$$o = \begin{cases} 0 & \sum_{i=1 \dots n} v_i z_i \leq \theta \\ 1 & \text{w p.p.} \end{cases}$$

Preceptron

Mapowanie $\{0, 1\}^n \rightarrow \{0, 1\}^m$



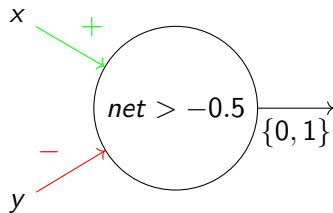
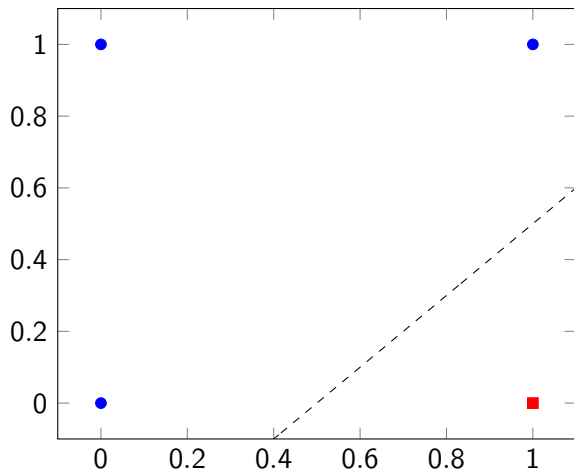
Inne architektury

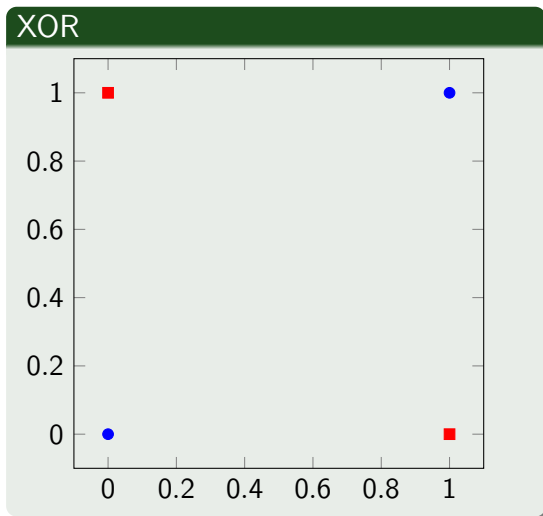
- Niepełne połączenia
- Cykle (sieci rekurencyjne)
- Połączenia w obrębie jednej warstwy

- **Klasyfikacja**
- **Regresja**
- Predykcja szeregów czasowych
- Rekonstrukcja wzorców
- **Analiza skupień**
- Kompresja/Ekstrakcja cech
- Optymalizacja

Różne sposoby połączeń między neuronami

Klasyfikacja z użyciem perceptronu



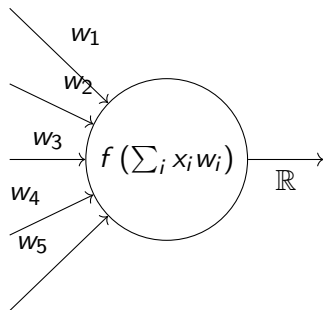


Z użyciem perceptronu nie da się rozwiązać problemu XOR. Minsky and Papert 1969
Nastąpił znaczący spadek liczby publikacji w obszarze sieci neuronowych.

- ➊ Rozwój 1943-69
- ➋ Stagnacja
- ➌ Rozwój lata '80, początek '90 – dostępność szybszych komputerów, bardziej złożone sieci
- ➍ Stagnacja
- ➎ Rozwój 2006-2012 – sieci głębokie
- ➏ Obecnie – ???

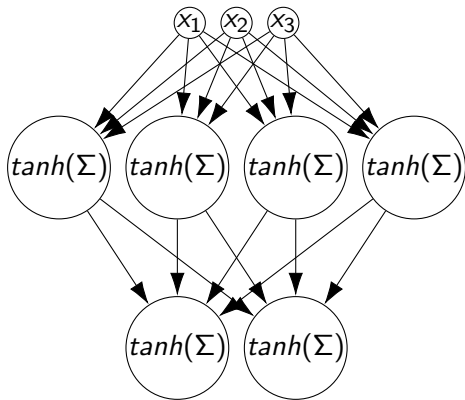
Współczesny model neuronu

(Najpopularniejszy)

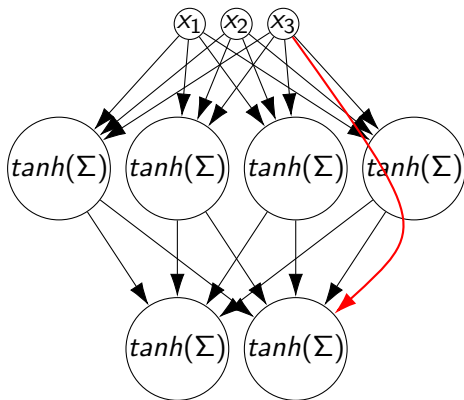


- wejścia $x = [x_1, x_2, \dots, x_n] \in \mathbb{R}^n$
- waga wejścia $w = [w_1, w_2, \dots, w_n] \in \mathbb{R}^n$
- wyjście $y \in \mathbb{R}$
- funkcja aktywacji f
 - monotoniczna
 - różniczkowalna

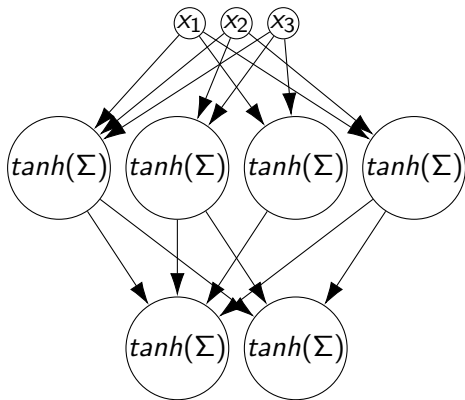
Chwilowo przyjmijmy $f(a) = \frac{e^a}{e^a + 1}$



- Propagacja informacji tylko w jednym kierunku
- Brak pętli
- Wyliczenie w znanej z góry liczbie kroków
- Zwykle zorganizowane w warstwy
- Typowo połączenia wewnątrz jednej warstwy

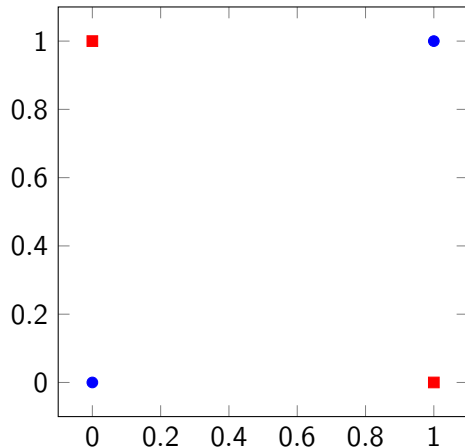


- Propagacja informacji tylko w jednym kierunku
- Brak pętli
- Wyliczenie w znanej z góry liczbie kroków
- Zwykle zorganizowane w warstwy
- Połączenia między niesąsiednimi warstwami



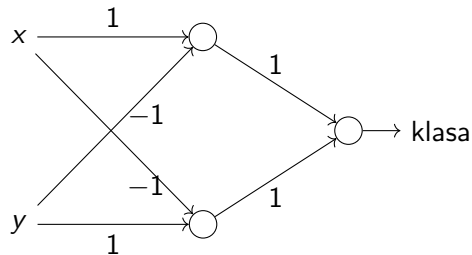
- Propagacja informacji tylko w jednym kierunku
- Brak pętli
- Wyliczenie w znanej z góry liczbie kroków
- Zwykle zorganizowane w warstwy
- Typowo połączenia wewnątrz jednej warstwy, mogą być niepełne

Sieć rozwiązująca XOR

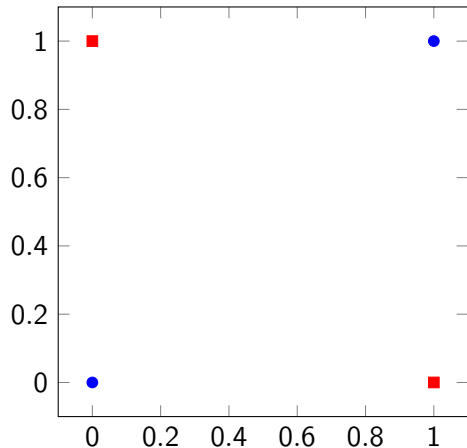


Niech ■ – klasa 1, ● – klasa 0.

Funkcja aktywacji: $f(x) = 1$, jeśli $x > 0$,
 $f(x) = 0$ w przeciwnym przypadku.

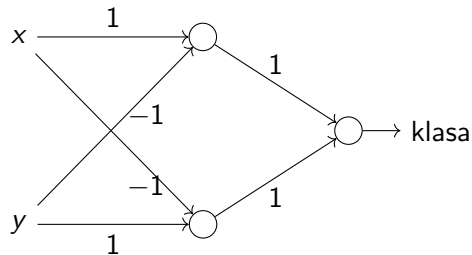


Sieć rozwiązująca XOR

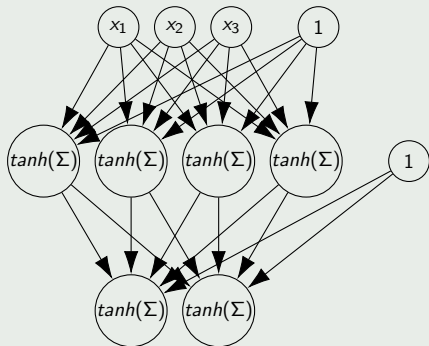


A teraz niech ■ – klasa 0, ● – klasa 1.

Funkcja aktywacji: $f(x) = 1$, jeśli $x > 0$,
 $f(x) = 0$ w przeciwnym przypadku.

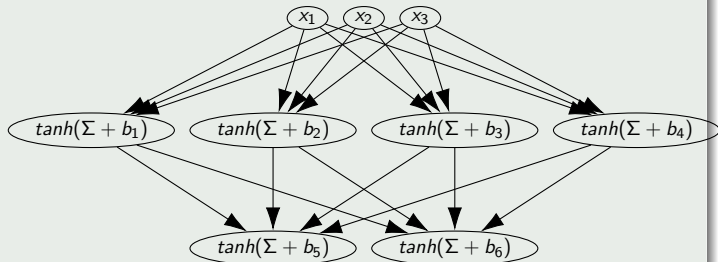


Dodatkowy neuron, który jest zawsze aktywny z wartością 1



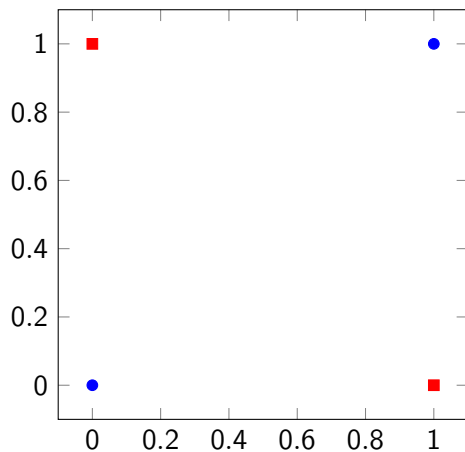
$$y_i = f \left(\sum_{n=1}^n w_n x_n + w_b x_b \right)$$

Przesunięcie argumentu funkcji aktywacji o wartość



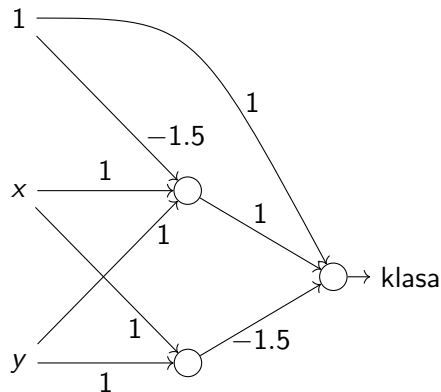
$$y_i = f \left(\sum_{n=1}^n w_n x_n + b \right)$$

Sieć rozwiązująca XOR



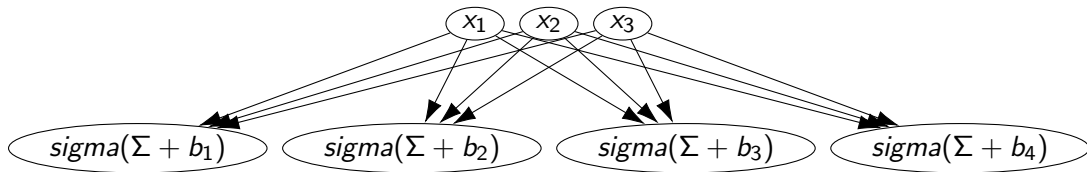
Niech ■ – klasa 0, ● – klasa 1.

Funkcja aktywacji: $f(x) = 1$, jeśli $x > 0$,
 $f(x) = 0$ w przeciwnym przypadku.



Perceptron z sigmoidalną aktywacją

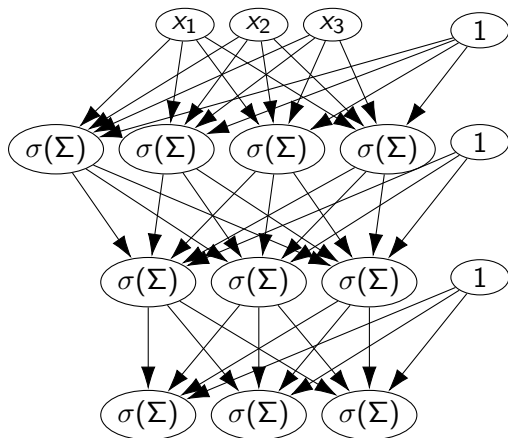
$$\sigma(x) = \frac{e^x}{e^x + 1}$$



Perceptron wielowarstwowy

Multilayer Perceptron (MLP)

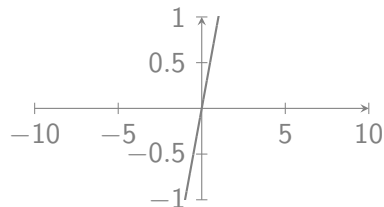
- Warstwy
- Pełne połączenia między warstwami (zazwyczaj)



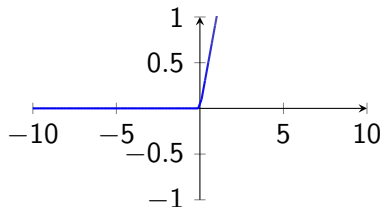
- Sieć Hopfielda, Maszyna Boltzmanna – połączenia każdy z każdym
- Sieci rekurencyjne

Popularne funkcje aktywacji

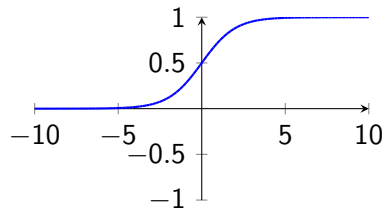
Liniowa



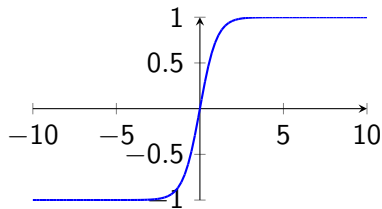
$\text{ReLU}(x) = \max\{x, 0\}$



Sigmoid $\sigma(x) = \frac{e^x}{1+e^x}$



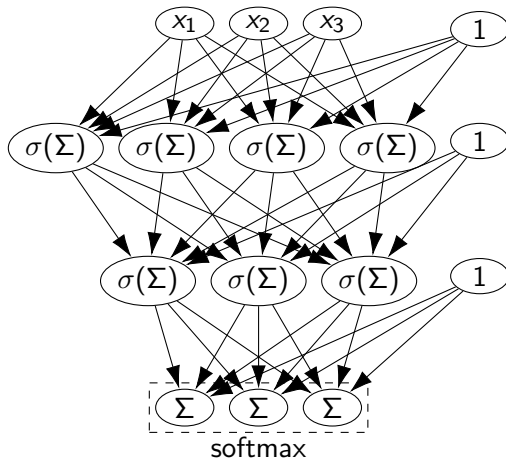
$\tanh(x) = \frac{e^{2x}-1}{1+e^{2x}} =$

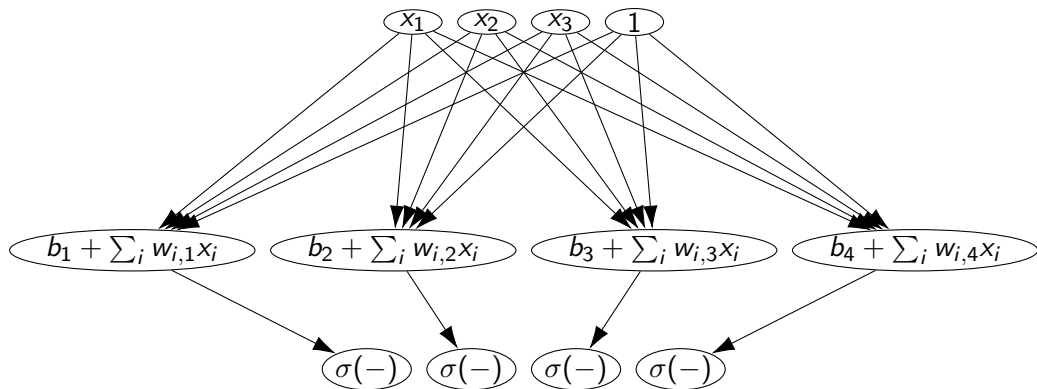


- Sieć feedforward: $N : \mathbb{R}^n \rightarrow \mathbb{R}^m$
- k klas
- Label encoding: jedna wartość $\in [0, 1]$. i -ta klasa: $\frac{i-1}{k-1}$
- One-hot encoding: wektor $\{0, 1\}^k$, i -ta klasa: jedynka na i -tej pozycji, zera na pozostałych

- Sieć feedforward: $N : \mathbb{R}^n \rightarrow (0, 1)^m$
- k klas
- Label encoding: jedna wartość $\in [0, 1]$. i -ta klasa: $\frac{i-1}{k-1}$
- One-hot encoding: wektor $\{0, 1\}^k$, i -ta klasa: jedynka na i -tej pozycji, zera na pozostałych

- Zadanie klasyfikacji
- Prawdopodobieństwo przynależności do klasy
- Zazwyczaj na warstwie wyjściowej
- W przeciwieństwie do innych funkcji aktywacji liczona na podstawie całej warstwy, a nie tylko jednego neuronu.
- $$y_i = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}}$$



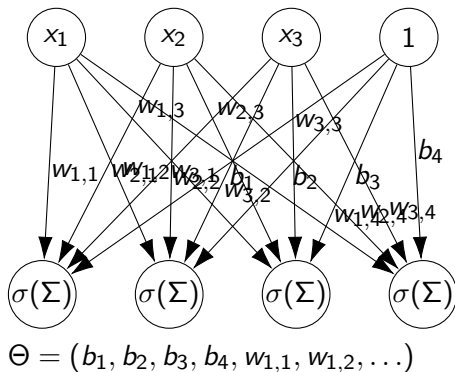


$$y_j = \sigma\left(\sum_i w_{i,j} x_i + b_j\right)$$

- Inny model neuronu
- Czas ciągły
- Neuron akumuluje pobudzenia w czasie
- Neuron wysyła impuls po przekroczeniu progu
- Hodgkin, Huxley: Hodgkin and Huxley 1952
- Izhikevich: Izhikevich 2003

Sieć neuronowa w zadaniu regresji

- Wektor cech: $X \in [0, 1]^N$,
wektor odpowiedzi $Y \in [0, 1]^M$
- k -elementowy zbiór testowy
 $\subset X \times Y$
- Model (sieć) N i parametry
(wagi) Θ
- $MAE = \sum_i \frac{|N(x_i, \Theta) - y_i|}{k}$
- Skąd wziąć wektor Θ ?



- [1] Jarosław Arabas. *Wykłady z algorytmów ewolucyjnych*. Wydawnictwa Naukowo-Techniczne, 2001.
- [2] Andries P Engelbrecht. *Computational intelligence: an introduction*. John Wiley & Sons, 2007.
- [3] Alan L Hodgkin and Andrew F Huxley. "A quantitative description of membrane current and its application to conduction and excitation in nerve". In: *The Journal of physiology* 117.4 (1952), pp. 500–544.
- [4] E.M. Izhikevich. "Simple model of spiking neurons". In: *IEEE Transactions on Neural Networks* 14.6 (2003), 1569–1572. ISSN: 1045-9227. DOI: 10.1109/tnn.2003.820440. URL: <http://dx.doi.org/10.1109/tnn.2003.820440>.
- [5] James Kennedy et al. *Swarm intelligence*. Morgan Kaufmann, 2001.
- [6] Robert Kosiński. *Sztuczne sieci neuronowe: dynamika nieliniowa i chaos*. Wydawnictwa Naukowo-Techniczne, 2008.

- [7] Warren S. McCulloch and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity". In: *Bulletin of Mathematical Biology* 52.1-2 (1943), 99–115. ISSN: 1522-9602. DOI: 10.1007/bf02459570. URL: <http://dx.doi.org/10.1007/bf02459570>.
- [8] Zbigniew Michalewicz. *Genetic algorithms+ data structures= evolution programs*. Springer Science & Business Media, 2013.
- [9] Marvin Minsky and Seymour Papert. *Perceptrons - an introduction to computational geometry*. MIT Press, 1969.
- [10] Stanisław Osowski. *Sieci neuronowe do przetwarzania informacji*. Oficyna Wydawnicza Politechniki Warszawskiej, 2004.
- [11] F. Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain.". In: *Psychological Review* 65.6 (1958), 386–408. ISSN: 0033-295X. DOI: 10.1037/h0042519. URL: <http://dx.doi.org/10.1037/h0042519>.



Fundusze Europejskie
Wiedza Edukacja Rozwój



Rzeczpospolita
Polska

**Politechnika
Warszawska**

Unia Europejska
Europejski Fundusz Społeczny



Zadanie 10 pn. „Modyfikacja programów studiów na kierunkach prowadzonych przez Wydział Matematyki i Nauk Informatycznych” realizowane jest w ramach projektu „NERW 2 PW. Nauka – Edukacja – Rozwój – Współpraca” współfinansowanego ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego