



Sztuczne sieci neuronowe

Krzysztof A. Cyran
POLITECHNIKA ŚLĄSKA
Instytut Informatyki, p. 311

Wykład 1



PLAN:

- **Jak (klasyczne) komputery oceniają filozofię Alberta Einsteina**
- **Historia i korzenie teorii sztucznych sieci neuronowych**
- **Wprowadzenie do teorii sieci neuronowych**
- **Sieci neuronowe w praktyce**
- **Jeszcze raz o Albercie Einsteinie, filozofii i sztucznej inteligencji**
- **Bibliografia**

Jakim filozofem był Albert Einstein ?

- „Naukowcy są kiepskimi filozofami” - A. Einstein, *Out of my later years*, Philosophical Library, New York 1950
- A. Einstein był wybitnym naukowcem (i zdawał sobie z tego sprawę).
- Wniosek klasycznego, inteligentnego (???) systemu ekspertowego:
<<Albert Einstein uważał się za kiepskiego filozofa>>

Biologiczne korzenie

- Mózg ludzki: objętość 1,4 l., pow. 2000cm², masa 1,5 kg
- Kora mózgowa: grubość 3 mm, 10¹⁰-10¹¹ komórek nerwowych, liczba połączeń (synaps) 10¹⁴-10¹⁵
- Impulsy komórek nerwowych: częstotliwość 1-100Hz, czas trwania 1-2 ms, napięcie 100mV
- Szybkość pracy mózgu:
 10^{15} połączeń x 100Hz = 10¹⁷ op./s

Biologiczne korzenie (cd)



- Cechy mózgu:
 - odporność na błędy
 - elastyczność
 - odporność na zaszumioną informację

Historia sztucznych sieci neuronowych (SSN)

- 1943 – matematyczny model sztucznego neuronu McCullocha i Pittsa
- 1949 – Hebb – reguła uczenia bez nadzoru (Hebba) w *Organization of Behaviour*
- 1958 – perceptron Rosenblatta – pierwsza implementacja SSN: elektroniczno-elektromechaniczny układ, warstwa wejściowa – warstwa wyjściowa, uczenie zbieżne (o ile istnieje rozwiązanie), aplikacja: do rozpoznawania znaków alfanumerycznych

Historia rozwoju SSN (cd)



- 1960 – Widrow i Hoff: Madaline
(multiple adaptive linear neurons)
- 1969 – Minsky i Papert: książka *Perceptrons* – krytyka SSN,
- zastój do początku lat 80.

Historia rozwoju SSN (cd)

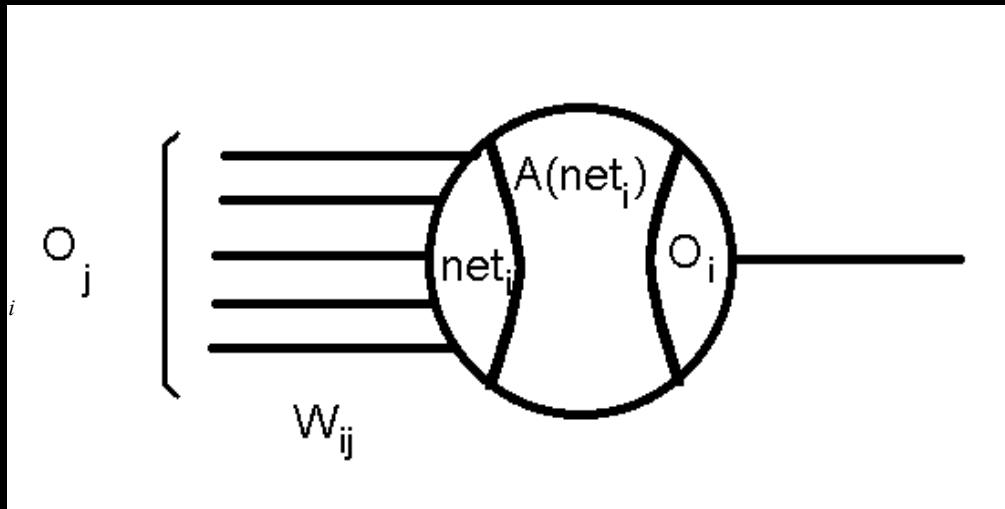
Odrodzenie SSN:

- 1982 – Hopfield buduje pamięć autoasocjacyjną, sieć rozwiązującą NP-zupełny problem komiwojażera,
- 1974, 1982, 1986 – algorytm wstecznej propagacji błędów odpowiedzią na krytykę Minsky'ego i Paperta
- Początek lat 90: matematyczny dowód na sensowność wykorzystania SSN (perceptronów wielowarstwowych) w roli klasyfikatorów w warunkach niepewności probabilistycznej

Zalety sztucznych sieci neuronowych (SSN)

- Potrafią odpowiadać w warunkach informacji niepełnej
- Nie wymagają znajomości algorytmu rozwiązania zadania (automatyczne uczenie)
- Przetwarzają informację w sposób wysoce równoległy
- Potrafią generalizować (uogólniać na przypadki nieznane)
- Są odporne na częściowe uszkodzenia
- Potrafią realizować pamięć asocjacyjną (skojarzeniową – podobnie jak działa pamięć u ludzi) w przeciwieństwie do pamięci adresowanej (charakterystycznej dla klasycznych komputerów)

Sztuczny neuron



$$n_i = \sum_j w_{ij} O_j$$

$$A_i = A(n_i)$$

$$O_i = TF(A_i)$$

czyli :

$$O_i = TF(A(\sum_j w_{ij} O_j))$$

Ogólna postać reguły uczenia

$$\Delta w_{ij} = g(n_i, d_i)h(O_j, w_{ij})$$

g, h – dowolne funkcje

Reguła uczenia Hebba

$$\Delta w_{ij}^{(p)} = \eta n_i^{(p)} O_j^{(p)}$$

Gdy: $g \sim n_i, h \sim O_j$

$$\Delta w_{ij} = g(n_i, d_i) h(O_j, w_{ij})$$

Reguła uczenia Delta (Widrowa Hoffa)

$$\Delta w_{ij}^{(p)} = \eta (d_i^{(p)} - n_i^{(p)}) O_j^{(p)}$$

Gdy:

$$g(n_i, d_i) = \eta (d_i - n_i),$$

$$h(O_j, w_{ij}) = O_j$$

$$\Delta w_{ij} = g(n_i, d_i) h(O_j, w_{ij})$$

Typy uczenia:

- Uczenie nadzorowane gdy w regule w funkcji g jako argument występuje d_i
- Uczenie nienadzorowane gdy w regule w funkcji g nie występuje d_i

$$\Delta w_{ij} = g(n_i, d_i)h(o_j, w_{ij})$$

Reguła Grossberga (1976)

$$\Delta w_{ij} = \eta n_i (O_j - w_{ij})$$

Gdy:

$$g(n_i, d_i) = \eta n_i,$$

$$h(O_j, w_{ij}) = O_j - w_{ij}$$

$$\Delta w_{ij} = g(n_i, d_i) h(O_j, w_{ij})$$

Reguła uczenia Hebb - anty Hebb

$$\Delta w_{ij} = \eta n_i (2O_j - 1)$$

Gdy: $g = \eta n_i$, $h = 2O_j - 1$

$$\Delta w_{ij} = g(n_i, d_i) h(O_j, w_{ij})$$

Reguła uczenia Hebb - anty Hebb z forsowaniem

$$\Delta w_{ij} = \eta d_i (2O_j - 1)$$

Gdy: $g = \eta d_i$, $h = 2O_j - 1$

$$\Delta w_{ij} = g(n_i, d_i) h(O_j, w_{ij})$$

Reguła uczenia Hopfielda

$$\Delta w_{ij} = \eta(2n_i - 1)(2O_j - 1)$$

Gdy: $g = \eta(2n_i - 1)$, $h = 2O_j - 1$

$$\Delta w_{ij} = g(n_i, d_i)h(O_j, w_{ij})$$

Reguła uczenia Hopfielda z forsowaniem

$$\Delta w_{ij} = \eta(2d_i - 1)(2O_j - 1)$$

Gdy: $g = \eta(2d_i - 1)$, $h = 2O_j - 1$

$$\Delta w_{ij} = g(n_i, d_i)h(O_j, w_{ij})$$

Reguła Kohonena – stosowana do modyfikacji jednego „zwycięskiego” neuronu

$$\Delta w_{ij} = \eta (O_j - w_{ij})$$

Gdy:

$$g(n_i, d_i) = \eta,$$

$$h(O_j, w_{ij}) = O_j - w_{ij}$$

$$\Delta w_{ij} = g(n_i, d_i) h(O_j, w_{ij})$$

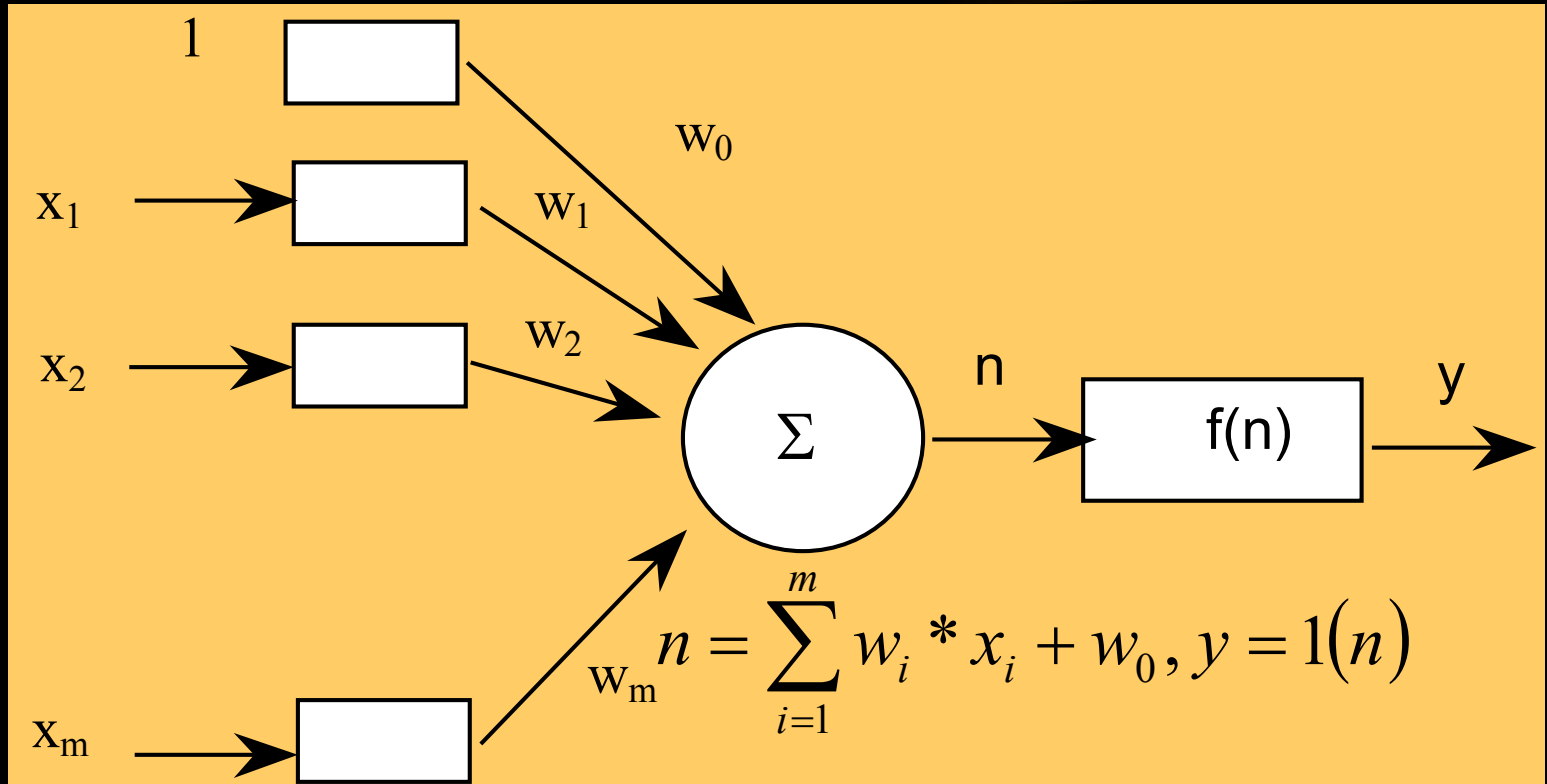
Reguła uczenia w niektórych sieciach Hopfielda oraz BAM

$$\Delta w_{ij}^{(p)} = \eta d_i^{(p)} O_j^{(p)}$$

Gdy: $g = \eta d_i$, $h = O_j$

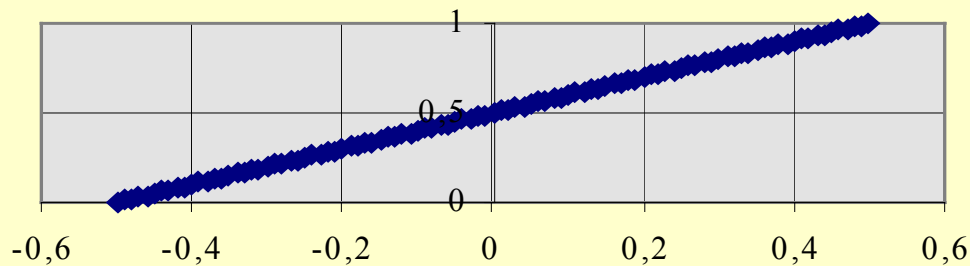
$$\Delta w_{ij} = g(n_i, d_i) h(O_j, w_{ij})$$

Model neuronu McCullocha-Pittsa



Funkcje przejścia

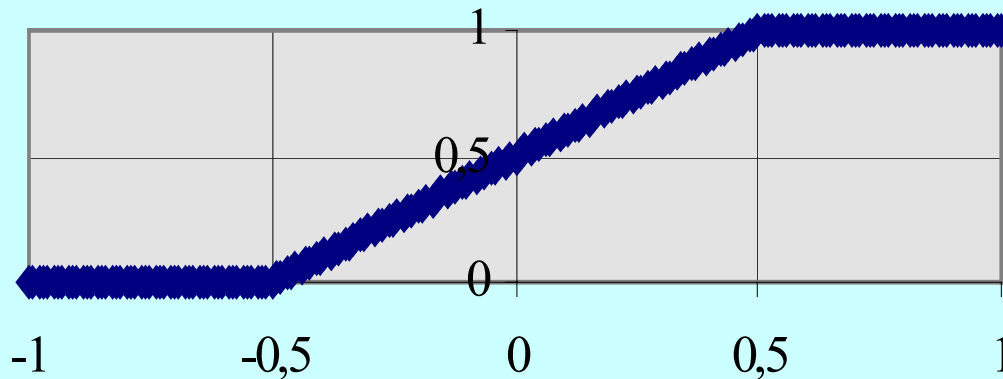
$$y(n) = kn$$



- Liniowa (liniowa, ciągła, różniczkowalna)

Funkcje przejścia (cd)

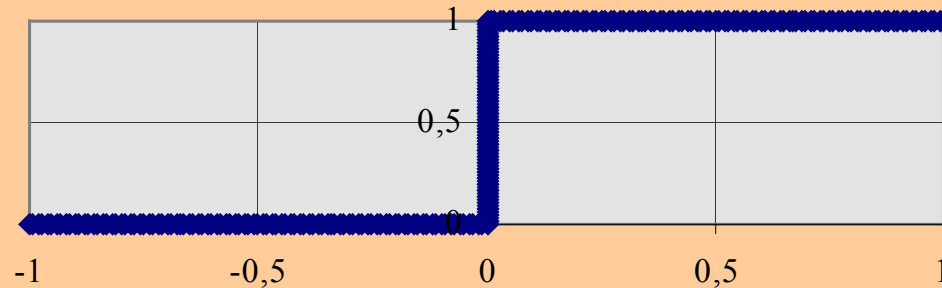
$$y(n) = \begin{cases} -1, & \text{dla } n \leq -1/k \\ kn, & \text{dla } -1/k < n \leq 1/k \\ 1, & \text{dla } n > 1/k \end{cases}$$



- Liniowa progowa
(nielinowa, ciągła, nieróżniczkowalna)

Funkcje przejścia (cd)

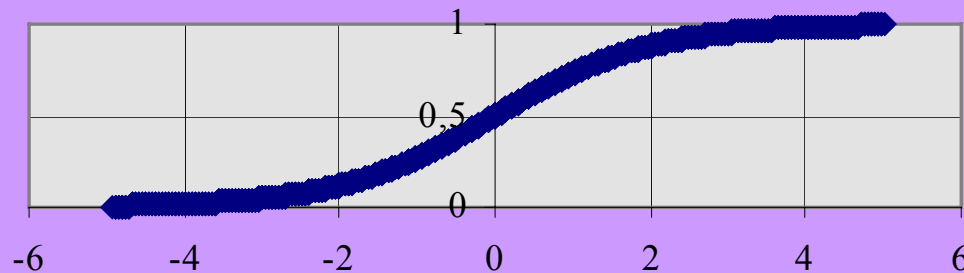
$$y(n) = \begin{cases} 1 & \text{dla } n > 0 \\ 0 & \text{dla } n \leq 0 \end{cases}$$



- Skoku jednostkowego
(nieliniowa, nieciągła, nieróżniczkowalna)

Funkcje przejścia (cd)

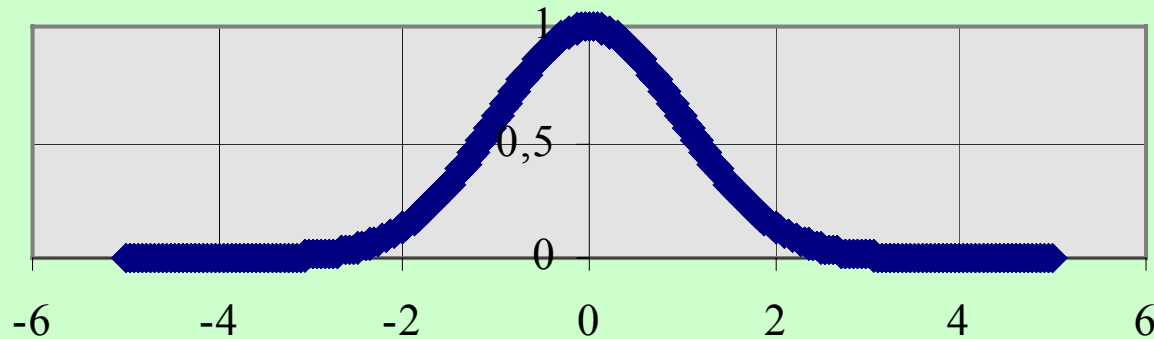
$$y(n) = \frac{1}{1 + e^{-\beta n}}$$



- Sigmoidalna
(nieliniowa, ciągła, różniczkowalna)

Funkcje przejścia (cd)

$$y(n) = e^{-n^2}$$

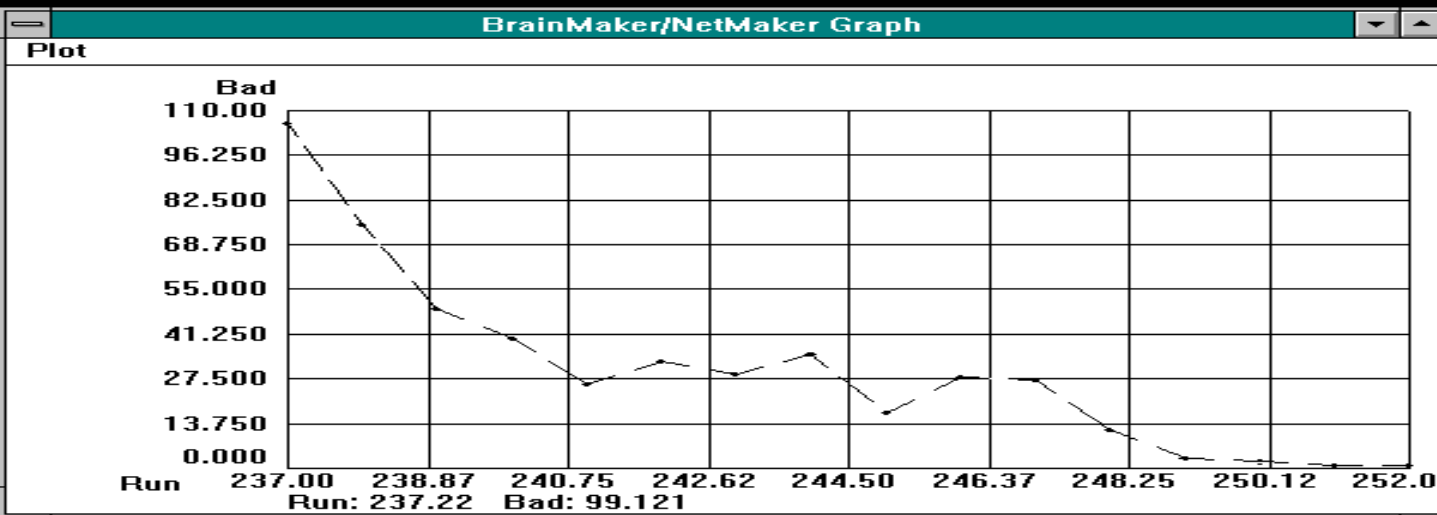
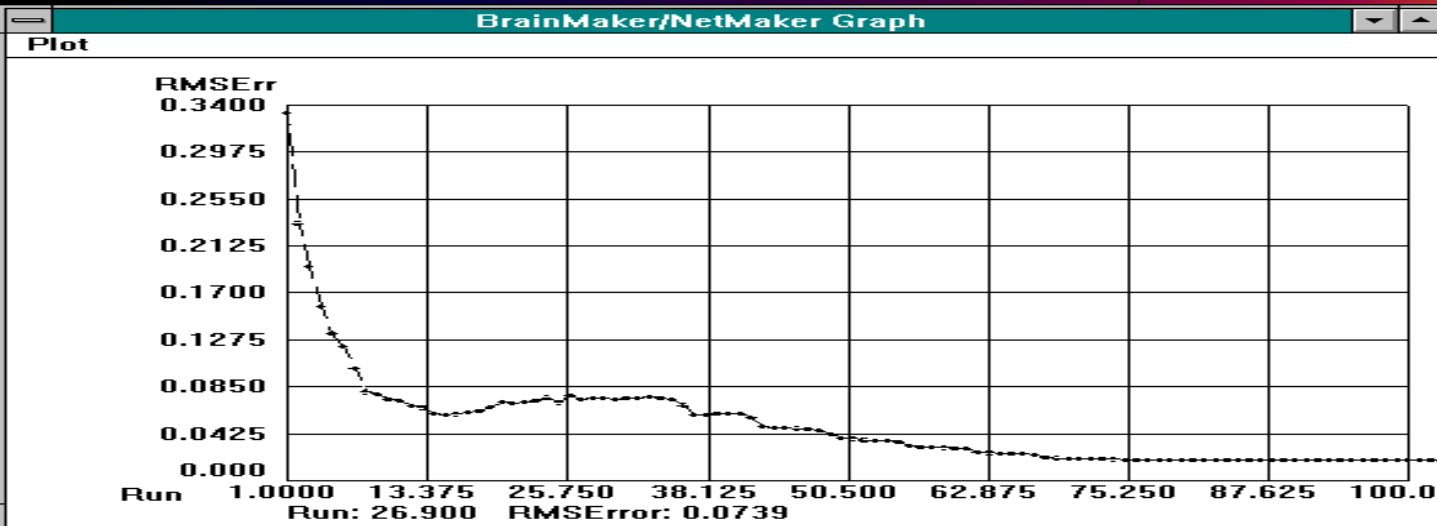


- Gaussa (nielinowa, ciągła, różniczkowalna, niemonotoniczna)

Klasyfikacja sieci neuronowych

- Ze względu na topologię:
 - ze sprzężeniami (pamięć autoasocjacyjna Hopfielda)
 - bez sprzężeń (perceptron wielowarstwowy, mapa cech Kohonena, sieć z radialnymi funkcjami bazowymi)
- Ze względu na sposób uczenia:
 - uczenie bez nadzoru (mapa cech Kohonena)
 - uczenie z nadzorem (perceptron wielowarstwowy, sieć z radialnymi funkcjami bazowymi)
- Ze względu na funkcję przejścia:
 - liniowe (Adaline)
 - nieliniowe (perceptron wielowarstwowy, sieć z radialnymi funkcjami bazowymi)

Proces uczenia z nadzorem



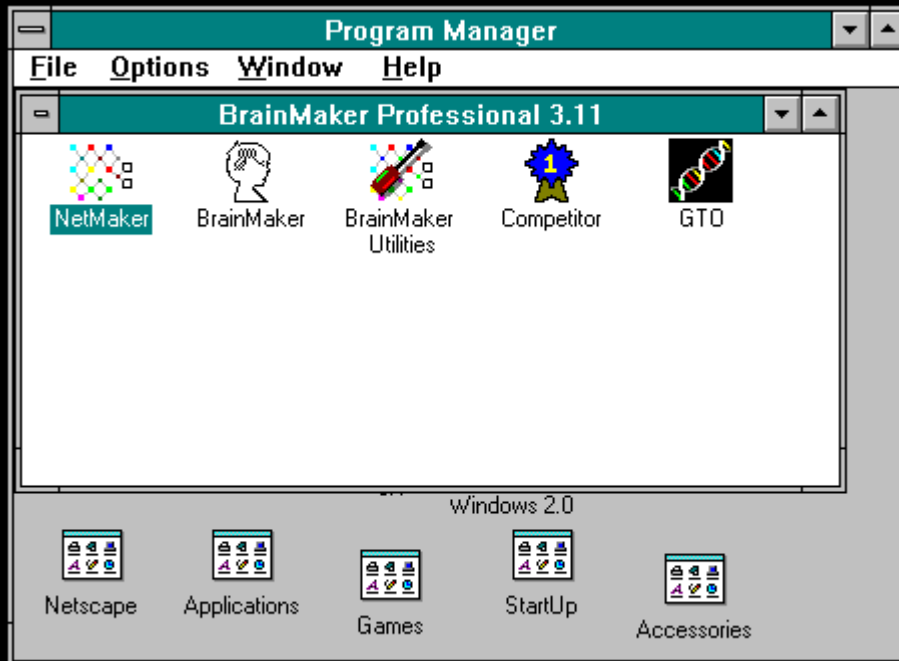
- Minimalizacja błędu średniokwadrat
- Minimalizacja ilości błędnych odpowiedzi

Zastosowania sieci neuronowych



- Rozpoznawanie obrazów
- Rozpoznawanie mowy
- Diagnostyka medyczna
- Prognozowanie w ekonomii
- Sterowanie ruchami robota
- Symulowanie zachowań struktur biologicznych neuronów

Przykładowe pakiety symulujące sieci neuronowe



-BrainMaker

-Matlab NN

-Neural ++

-Stock Prophet

- ...

Znowu o Einsteinie

- Zastrzeżenia do wyciągniętego w oparciu o ścisłe reguły logiczne wniosku: „Einstein uważał się za kiepskiego filozofa” spowodowane asocjacją nazwiska Einstein z jego zamiłowaniem do filozofii oraz z jego geniuszem.
- Konieczny dodatkowy (ukryty) kontekst wypowiedzi
- Łatwość rozumu ludzkiego (pierwowzoru SN) z radzeniem sobie z nieprecyzyjnymi (niedopowiedzonymi) stwierdzeniami poprzez odkrywanie ukrytego domyślnego kontekstu.
- Odkryty sens wypowiedzi Einsteina: „naukowcy [NA OGÓŁ] są kiepskimi filozofami”, zatem nie musiał on uważać się za kiepskiego filozofa
- Każdy naukowiec może teraz spokojnie filozofować sądząc, że to „na ogół” jego nie dotyczy...(mając wszakże na uwadze, że za sprawą lepszego samopoczucia, jego filozofia niestety nie stanie się lepsza).

Na koniec BARDZO filozoficzne pytania



- Czym jest inteligencja
 - Czym jest świadomość
 - Czym jest samoświadomość
- i wreszcie:
- Czy można POWAŻNIE mówić o sztucznej inteligencji, w sytuacji, kiedy maszyna, która z jednej strony wygrywa z mistrzem świata w szachy, z drugiej, nie ma o tym najmniejszego pojęcia?

Bibliografia

- T. Masters, *Sieci neuronowe w praktyce – programowanie w języku C++*, Warszawa 1996
- R. Tadeusiewicz, *Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami*, Warszawa 1998
- J. Hertz, A. Krough, R. G. Palmer, *Wstęp do obliczeń neuronowych*, Warszawa 1995
- J. Korbicz, A. Obuchowicz, D. Uciński, *Sztuczne sieci neuronowe – podstawy i zastosowania*, Warszawa 1993

Bibliografia (cd)

- R. Tadeusiewicz, *Sieci neuronowe*, Warszawa 1993
- S. Osowski, *Sieci neuronowe w ujęciu algorytmicznym*, Warszawa 1996
- D. Rutkowska, M. Piliński, L. Rutkowski, *Sieci neuronowe, algorytmy genetyczne i systemy rozmyte*, Warszawa 1999
- J. Lawrence, *Introduction to Neural Networks*, Nevada City 1994