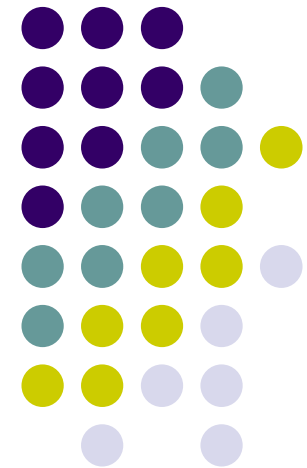


Seminarium magisterskie

Dyskusja nad tematem pracy
magisterskiej pisanej
pod kierunkiem pani
Dr hab. Małgorzaty Doman





Plan wystąpienia

- Ogólnie o sztucznych sieciach neuronowych
- Temat pracy magisterskiej z dziedziny SSN
- Alternatywne tematy
- Dyskusja: za i przeciw



Sieci neuronowe - istota

- Naśladowanie pracy ludzkiego mózgu
- System wzajemnie połączonych prostych elementów przetwarzających informacje (neuronów) – do połączeń przyporządkowane są współczynniki wagowe, określające siłę powiązań i tworzące zbiór parametrów.
- Struktura wewnętrzna – jednostki grupowane w większe zespoły zwane warstwami
- Cała wiedza sieci – w wewnętrznych odwzorowaniach, definiowanych w przez wartości wag – przywołana w procesie reakcji

Czym są...



- System uczący się,
- Bardzo dobrze nadają się do:
 - zadań związanych z rozpoznawaniem obrazów – rekonstrukcja sygnału
 - zadań optymalizacyjno - decyzyjnych,
 - szybkiego przeszukiwania dużych baz danych.
- Własność tolerowania i rozpoznawania błędów,
- Przetwarzanie równoległe = efektywne.

Czym sieci nie są...

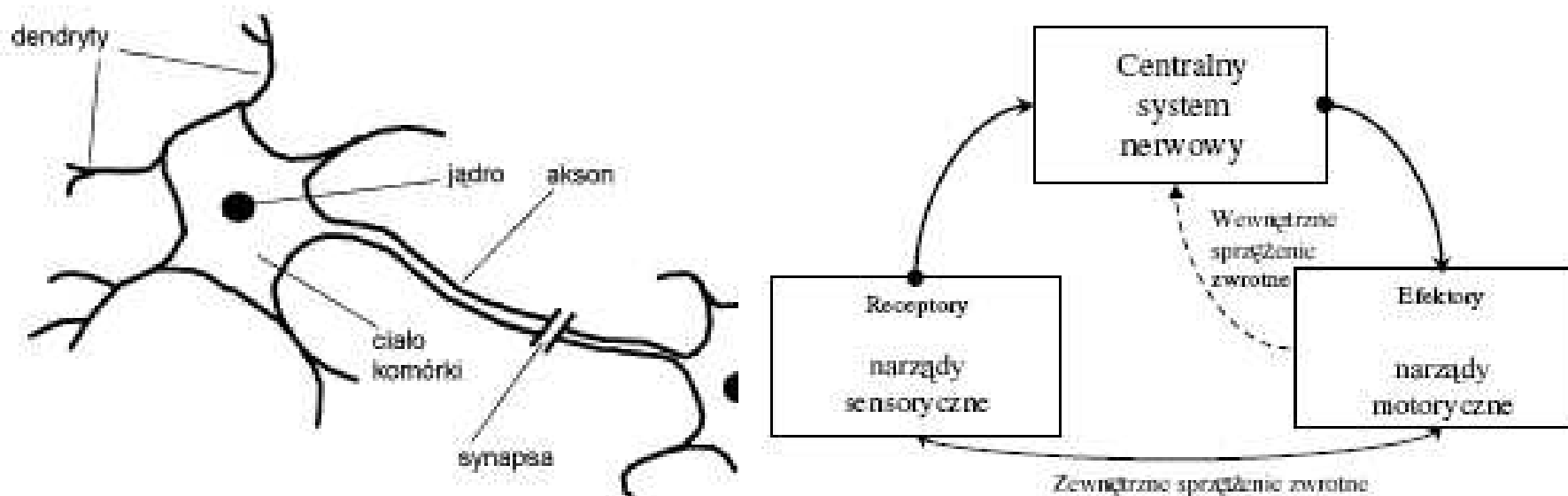


- Sieci nie uczą się algorytmów, uczą się przez przykłady,
- Są słabymi maszynami matematycznymi (mała precyzja) i słabo nadają się do typowego przetwarzania opartego o algorytmy.



Biologiczny model neuronu

- Ludzki mózg jako niedościgły wzór



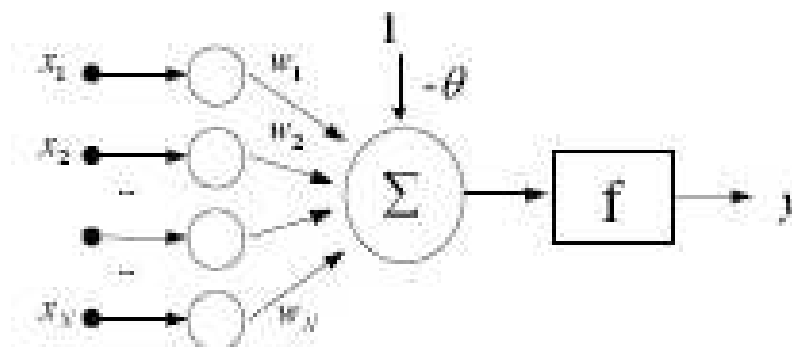
- Inspiracje biologiczne – ogólne zasady funkcjonowania SSN

Źródło: Wykład Małgorzaty Krętkowskiej



Matematyczny model neuronu

- model *McCullocha-Pittsa* (1943)



Przesyła do wielu,
ale dla wszystkich
informacja taka sama

- Gdzie:

- $x = [1, x_1, \dots, x_N]^T$ - wektor wejściowy
- $w = [-\theta, w_1, \dots, w_N]^T$ - wektor wag ($w_0 = -\theta$)
- f - funkcja aktywacji
- y - sygnał wyjściowy

Iloczyn skalarny = net
łączone pobudzenie

$$y = f\left(\sum_{i=0}^N w_i x_i\right) = f(w^T x) = f(\text{net})$$

Źródło: Bohdan Macukow i Maciej Grzenda, „Sieci neuronowe w różnych dziedzinach nauki”, Instytut Matematyki, Politechnika Warszawska,

Ewa Wołoszko

Informatyczny model neuronu



- Mózg = równoległy system złożony z około 10^{11} procesorów



- Każdy procesor ma bardzo prosty program i oblicza sumę ważoną danych WE pochodzących od innych procesów, a funkcja zwraca jedną liczbę
- Liczba przesyłana do innych procesorów i przez nie przetwarzana w podobny sposób

Źródło: Wykład Małgorzaty Krętkowskiej

Zadania sieci neuronowych



- Klasyfikacja – podział zbioru wejściowego na klasy
- Asocjacja – skojarzenie ze wzorcem
- Heteroasocjacja – skojarzenie między parą wzorców
- Optymalizacja - w tym rozwiązywanie liniowych i nieliniowych równań
- Estymacja – aproksymacja, interpolacja, filtrowanie, predykcja, prognozowanie
- Wykrywanie grup i wydobywanie cech (semantyczne sieci neuronowe)



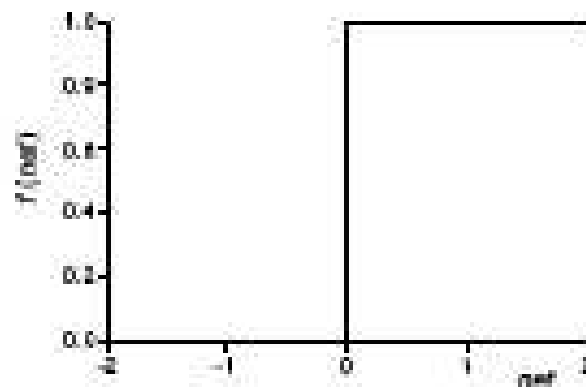
Rodzaje SNN ze względu na:

- Typ funkcji aktywacji
 - Unipolarne
 - Dyskretne – funkcja progowa
 - Ciągłe – f. logistyczna
 - Bipolarne
 - Dyskretne - funkcja Signum
 - Ciągłe – tangens hiperboliczny
 - Budowę sieci:
 - Jednokierunkowa
 - Jednowarstwowa
 - Wielowarstwowa
 - Rekurencyjna – ze sprzężeniem zwrotnym, występują cykle
 - Komórkowa – pojęcie sąsiedztwa
- Sigmoidalne funkcje aktywacji

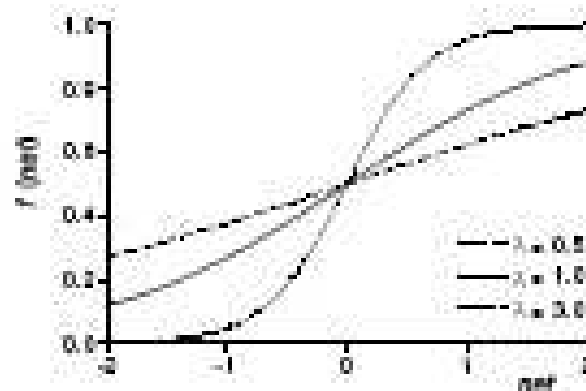
Funkcje aktywacji - unipolarne



$$f(\text{net}) = \begin{cases} 1 & \text{gdy } \text{net} \geq 0 \\ 0 & \text{gdy } \text{net} < 0 \end{cases}$$



$$f(\text{net}) = \frac{1}{1 + \exp(-\lambda \text{net})}, \quad \lambda > 0$$

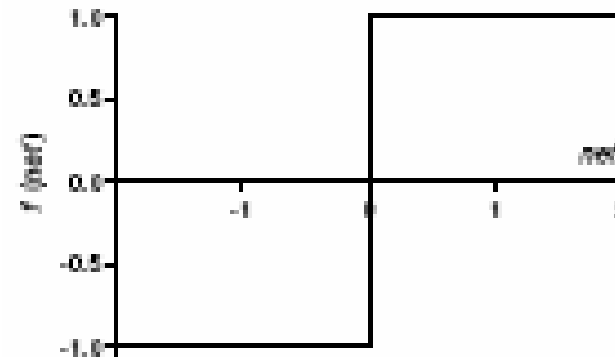


Uwzględnianie progu – uzasadnienie biologiczne

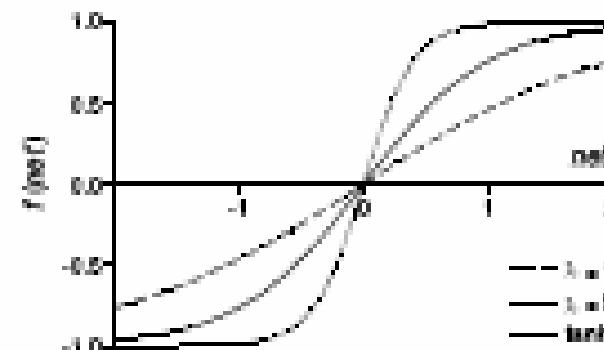
Funkcje aktywacji - bipolarne



$$f(net) = \text{sgn}(net) = \begin{cases} 1 & \text{gdzy } net \geq 0 \\ -1 & \text{gdzy } net < 0 \end{cases}$$



$$\tanh(net) = \frac{\exp(\lambda net) - \exp(-\lambda net)}{\exp(\lambda net) + \exp(-\lambda net)}, \quad \lambda > 0$$



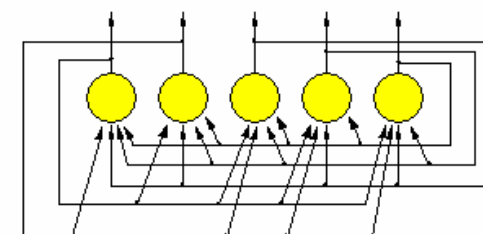
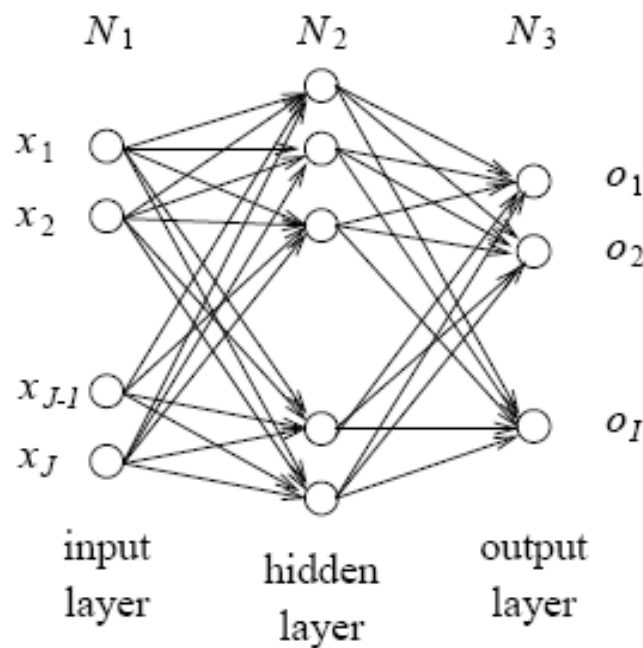
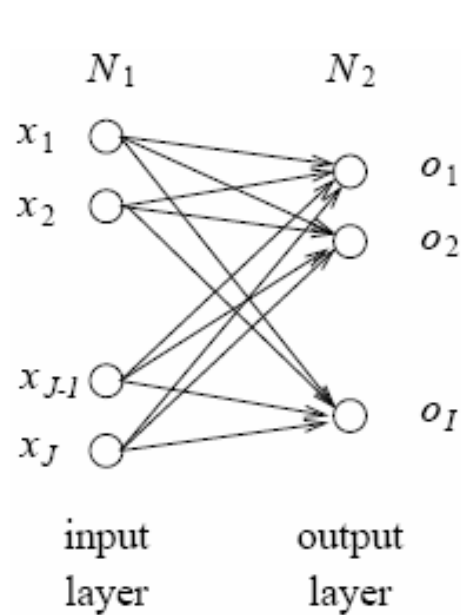
Uwzględnianie progu – uzasadnienie biologiczne



Typy sieci neuronowych

- Warstwowe sieci perceptronowe
- Sieci Kohonena
 - Sieci SOM (*Self Organising Maps*)
 - Sieci LVQ (*Learning Vector Quantization*)
- Sieci rekurencyjne
 - Częściowo rekurencyjne
 - Sieć Hopfield'a
 - Sieci BAM (*Bi-directional Associative Memory*)

Przykładowe architektury



Sieć rekurencyjna

Przygotowanie danych treningowych



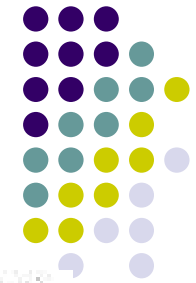
- Dobór zmiennych i wstępne przetworzenie – zależy od rozwiązywanego problemu
- Szeregi czasowe – normalizacja i detrendyzacja
- Wykorzystanie sieci Kohonena – klasyfikatory do wybrania danych treningowych
- Stosowanie surowych danych o różnych przedziałach zmienności może spowodować nieprawidłowe działanie sieci
 - Nasycenie sigmoid dla dużych wartości – hamowanie procesu uczenia – normalizacja do przedziału: $[-1, 1]$, $[0, 1]$
 - Standaryzacja – zerowa średnia i jednostkowe odchylenie



Uczenie sieci neuronowej

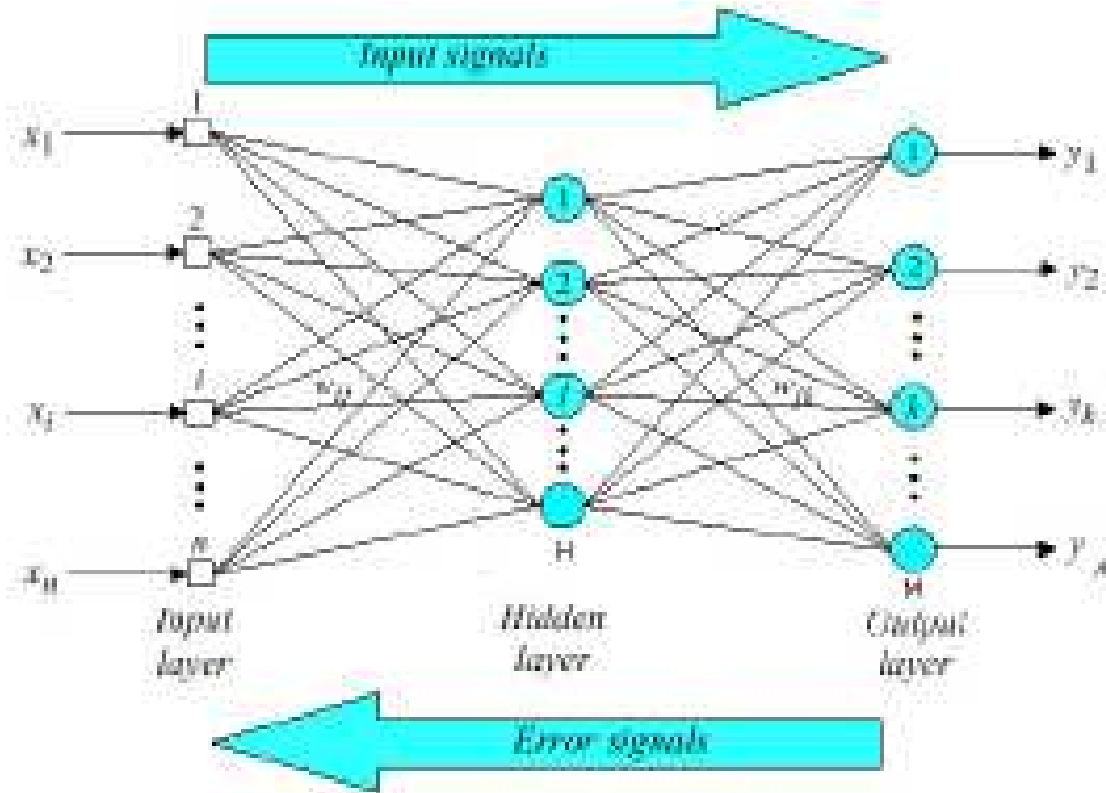
- Modyfikacje wag $W_{ij}(k+1) = W_{ij}(k) + \Delta W_{ij}(k)$
- Warianty:
 - Z nauczycielem
 - Dane treningowe = sygnał WE + poprawne reakcje
 - Bez nauczyciela (autoasocjacja, samoorganizacja)
 - Bez informacji o pożądanej reakcji
- Faza treningowa – warunek stopu
- Faza testowania – na danych testowych, nie uwzględnionych we wcześniejszym procesie uczenia
- Przetrenowanie – zbyt dokładne przystosowanie do danych treningowych, generalizacja ważniejsza od minimalizacji błędu
- Przygotowanie danych:
 - Treningowych
 - Testowych – obserwacja generalizacji
 - Walidacyjnych – ostateczna weryfikacja i ocena

Metoda wstecznej propagacji błędów

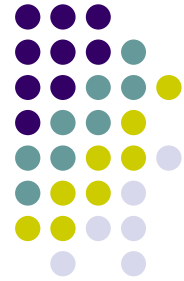


- Reguła delta – min błędu średniokwadratowego
- BPM - uogólnienie reguły delta dla sieci wielowarstwowej

$$E = \sum_{k=1}^P (y_i(k) - d_i(k))^2$$



Wybór optymalnej architektury



- Przyjęte modele neuronów
- Typ sieci: z nauczycielem, bez nauczyciela
- Wartości współczynników wagowych neuronów – inicjalizowane losowo
- Liczba warstw sieci M
 - Obszary decyzyjne formowane przez sieci perceptronowe – dodanie sieci ukrytej – bardziej złożone obszary decyzyjne
- Liczba warstw ukrytych i liczba neuronów w tych warstwach
 - dobierane eksperymentalnie:
 - za mało – niezdolność do gromadzenia wiedzy,
 - za dużo – słaba generalizacja, kłopoty z uogólnieniem
 - Sieć przeznaczona do samouczenia musi być większa
- Rozmiary warstw WE i WY zdeterminowane przez model sieci
- Liczba neuronów w poszczególnych warstwach n_i
- Przyjęty sposób połączeń neuronów

Metody budowy sieci zbliżonych do optymalnych

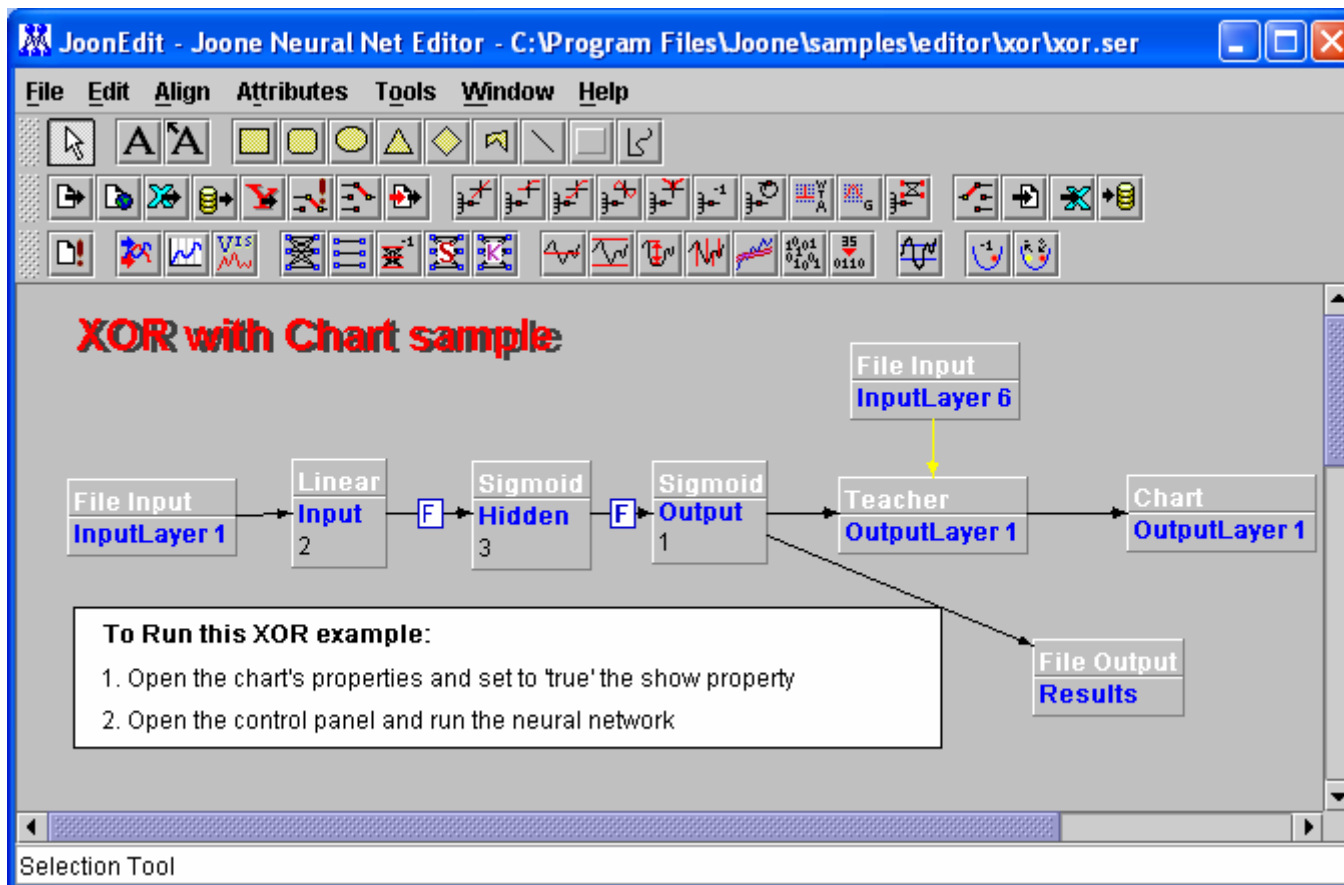


- Oszacowania analityczne (liczba połączeń w sieci nie powinna przekraczać $1/10$ liczby wzorców treningowych)
- Techniki konstrukcyjne – budowa przyrostowa
- Techniki destrukcyjne – obcinanie nadmiarowych połączeń, usuwanie mniej istotnych neuronów
- Algorytmy genetyczne – przeszukiwanie przestrzeni dopuszczalnych struktur sieci
- Wykorzystanie wiedzy od eksperta – w sieciach ekspertowych



XOR – przykład w programie Joone

- Tools => Control Panel => Teach / Use



- Pechowa funkcja XOR – Minsky i Papert (1969)



Systemy hybrydowe

- Sieci neuronowe + systemy ekspertowe
- Sieci neuronowe o logice rozmytej
- Algorytmy genetyczne do optymalizacji sieci neuronowych
- Uczenie SSN z wykorzystaniem metod ewolucyjnych

Zastosowania sieci neuronowych



- Określanie profilu klienta, *credit scoring*
- Analiza danych – grupowanie przedsiębiorstw, mapy cech banków
- Prognozowanie cen na rynkach kapitałowych - na podstawie szeregów czasowych
 - Dane historyczne + inne czynniki techniczne
 - Prognoza kursu wymiany; cen dnia: maksymalnej, zamknięcia i minimalnej
 - Modele prognozy o dłuższych horyzontach powinny odwoływać się w większej mierze do czynników fundamentalnych (CPI, INF...)
 - Model nie musi być znany, budowany w procesie uczenia
- Decyzje kupna lub sprzedaży na rynku kapitałowym – wspomaganie decyzji inwestycyjnych – sieć SOM
- Rozpoznawanie typowych wzorców zachowania rynku – identyfikacja formacji
 - Liczba neuronów WY = liczba formacji
 - Sieć Kohonena lub zestaw sieci perceptronowych + algorytm wyboru



Dlaczego sieci neuronowe?

- Interdyscyplinarne
- Temat
 - prognozowanie,
 - klasyfikowanie,
 - generowanie strategii
- wymaga bardzo dobrego zrozumienia badanego zagadnienia, obróbka danych wejściowych oparta na analizie ekonometrycznej
- Charakter badawczy – uczenie sieci, złe wyniki także dają wnioski
- Zbudowanie modułu w Javie do programu Joone
- Roboczy temat dotyczący SNN: „Zastosowanie sieci neuronowych do opracowywania strategii inwestycyjnych”

Alternatywne tematy pracy magisterskiej



- Kalkulator opcyjny – stworzenie programu na wzór:
 - <http://www.hoadley.net/options/options.htm>
- Wykorzystanie pakietu Statistica do pracy przy sieciach neuronowych
 - Dobrze opracowane przykłady zastosowań w ekonomii – StatSoft:
 - <http://www.statsoft.pl/czytelnia/neuron/wprowadoprak.html>



Za i przeciw...

- Dostępność literatury i źródeł, grono pasjonatów
- Sieci neuronowe – moda czy metoda?
- Kalkulator opcyjny – rzecz bardzo praktyczna

Moje dotychczasowe podstawowe źródła informacji



- Jerzy S. Zieliński „Inteligentne systemy w zarządzaniu – teoria i praktyka”
- Jerzy Tadeusiewicz: „Sieci neuronowe”
<http://winntbg.bg.agh.edu.pl/skrypty/0001/>
 - bossa.pl – od tego artykułu wszystko się zaczęło:
http://bossa.pl/analizy/techniczna/elementarz/sieci_neuronowe/
 - <http://www.jooneworld.com/wiki/tiki-index.php>
 - <http://www.hoadley.net/options/options.htm>
 - http://www.statsoft.pl/textbook/stathome_stat.html?http%3A%2F%2Fwww.statsoft.pl%2Ftextbook%2Fstdisfit.html

Dziękuję za uwagę!:)

