

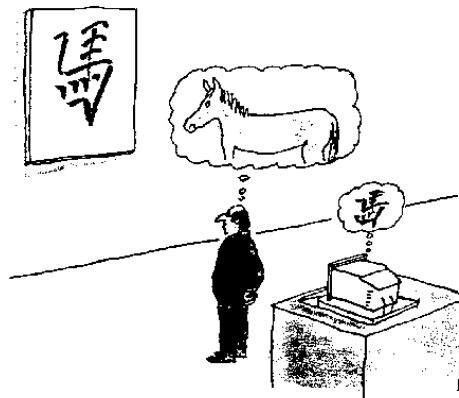
Metody prognozowania: Sztuczne sieci neuronowe

Cechy mózgu

- ✓ODPORNY NA USZKODZENIA;
- ✓ELASTYCZNY – ŁATWO DOSTOSOWUJE SIĘ DO ZMIENNEGO OTOCZENIA;
- ✓UCZY SIĘ - NIE MUSI BYĆ PROGRAMOWANY;
- ✓POTRAFI RADZIĆ SOBIE Z INFORMACJĄ ROZMYTĄ, LOSOWĄ, ZASZUMIONĄ LUB NIESPÓJNĄ;
- ✓ W WYSOKIM STOPNIU RÓWNOLEGŁY;
- ✓ MAŁY, ZUŻYWA BARDZO MAŁO ENERGII.



Cechy mózgu



Komputer widzi inaczej

Metody Prognozowania: Sztuczne sieci neuronowe

3

Cechy mózgu



Komputer widzi inaczej

Metody Prognozowania: Sztuczne sieci neuronowe

4

Cechy mózgu

- ✓ LICZBA POŁĄCZEŃ SYNAPTYCZNYCH W MÓZGU: $10^{10} - 10^{11}$;
- ✓ GĘSTOŚĆ POŁĄCZEŃ SYNAPTYCZNYCH: $\sim 10^4$ /NEURON;
- ✓ CZĘSTOTLIWOŚĆ GENERACJI SYGNAŁÓW PRZEZ NEURON: $\sim 1 - 100$ Hz;
- ✓ SZACUNKOWA SZYBKOŚĆ PRACY: $\sim 10^{18}$ OPERACJI/S
(DLA PORÓWNIANIA NAJSZYBSZE KOMPUTERY $\sim 10^{12}$ OPERACJI/S.

Metody Prognozowania: Sztuczne sieci neuronowe 5

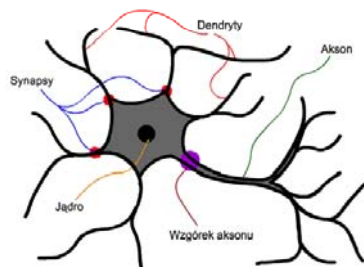
Sztuczny mózg ?

If the human brain were so simple
that we could understand it,
we would be so simple that we couldn't"

- Emerson Pugh -

Metody Prognozowania: Sztuczne sieci neuronowe 6

Inspiracje neurofizjologiczne (1)



Neuron (komórka nerwowa)

Jądro - "centrum obliczeniowe" neuronu. To tutaj zachodzą procesy kluczowe dla funkcjonowania neuronu.

Akson - "wyjście" neuronu. Za jego pośrednictwem neuron powiadamia świat zewnętrzny o swojej reakcji na dane wejściowe. Neuron ma tylko jeden akson.

Wzgórek aksonu - stąd wysyłany jest sygnał wyjściowy, który wędruje dalej poprzez akson.

Dendryt - "wejście" neuronu. Tędy trafiają do jądra sygnały mające być w nim później poddane obróbce. Dendrytów może być wiele - biologiczne neurony mają ich tysiące.

Synapsa - jeśli dendryt jest wejściem neuronu, to synapsa jest jego furtką. Może ona zmienić moc sygnału napływającego poprzez dendryt.

Metody Prognozowania: Sztuczne sieci neuronowe 7

Inspiracje neurofizjologiczne (2)

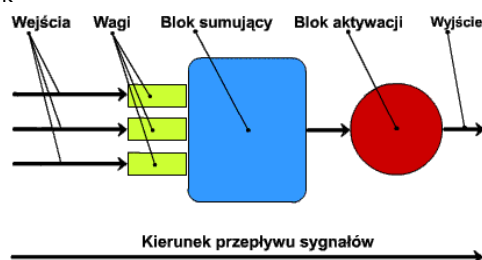
Wejścia to **dendryty**, lub ściślej: **sygnały** przez nie nadchodzące.

Wagi to cyfrowe odpowiedniki modyfikacji dokonywanych na sygnałach przez **synapsy**.

Blok sumujący to odpowiednik **jądra**,

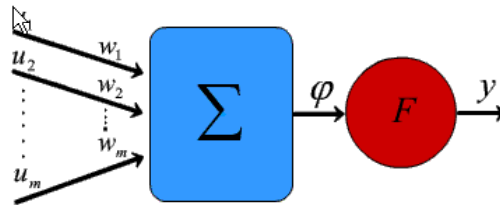
Blok aktywacji to **wzgórek aksonu**,

Wyjście - to **akson**.



Metody Prognozowania: Sztuczne sieci neuronowe 8

Model neuronu (1)



Potencjał membranowy:
$$\varphi = \sum_{i=1}^m w_i u_i = \mathbf{w}^T \mathbf{u}$$

Metody Prognozowania: Sztuczne sieci neuronowe 9

Model neuronu (2)

Blok aktywacji F , może być opisany różnymi funkcjami. Np. dla funkcji liniowej funkcja ma postać:

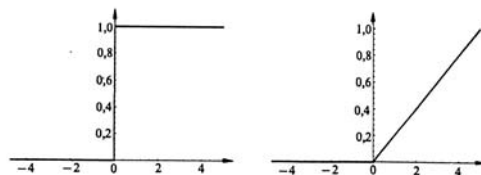
$$y = k\varphi$$

(sieci liniowe, neuron Adaline, sieci Madaline)

Inny przykład to funkcja skoku jednostkowego:

$$y = \begin{cases} 1, & \varphi > \varphi_h \\ 0, & \varphi \leq \varphi_h \end{cases}$$

φ_h – wartość progowa



Metody Prognozowania: Sztuczne sieci neuronowe 10

Model neuronu (3)

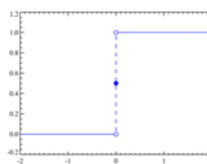
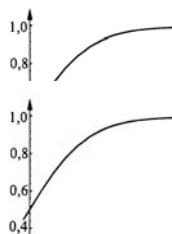
Funkcjami, które w dokładniejszy sposób opisują nieliniową charakterystykę przejścia neuronu biologicznego, są funkcje sigmoidalne (logistyczne):

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

$$\beta\varphi \rightarrow \infty \Rightarrow y \rightarrow 1$$

$$\beta\varphi \rightarrow -\infty \Rightarrow y \rightarrow 0$$

$$\beta \rightarrow \infty \Rightarrow \frac{1}{1 + e^{-\beta\varphi}} \rightarrow \mathbb{1}(\varphi)$$



Metody Prognozowania: Sztuczne sieci neuronowe 11

Model neuronu (4)

Funkcje tangensoidalne:

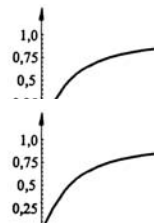
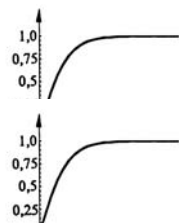
$$y = \tanh\left(\frac{\alpha\varphi}{2}\right) \frac{1 - e^{(-\alpha\varphi)}}{1 + e^{(-\alpha\varphi)}}$$

$$\alpha\varphi \rightarrow \infty \Rightarrow y \rightarrow 1$$

$$\alpha\varphi \rightarrow -\infty \Rightarrow y \rightarrow -1$$

$$\alpha \rightarrow \infty \Rightarrow \tanh\left(\frac{\alpha\varphi}{2}\right) \rightarrow \text{sgn}(\varphi)$$

$$\text{sgn}(\varphi) = \begin{cases} +1, & \varphi > 0 \\ -1, & \varphi \leq 0 \end{cases}$$



Metody Prognozowania: Sztuczne sieci neuronowe 12

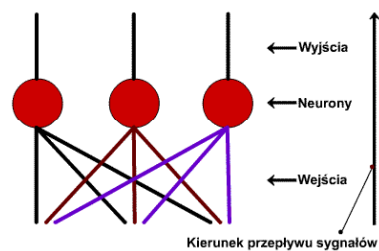
Sztuczna sieć neuronowa (1)

Już jeden neuron można nazwać siecią - jednoelementową.

Aby jednak móc użyć sieci w jakimś poważnym celu, trzeba będzie użyć większej ilości neuronów.

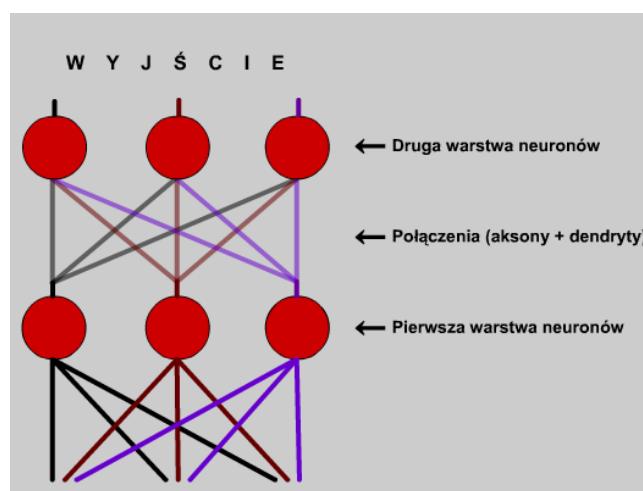
Neuron może mieć jeden lub więcej dendrytów (**wejść**) i dokładnie jeden akson (**wyjscie**).

Łączymy warstwami.



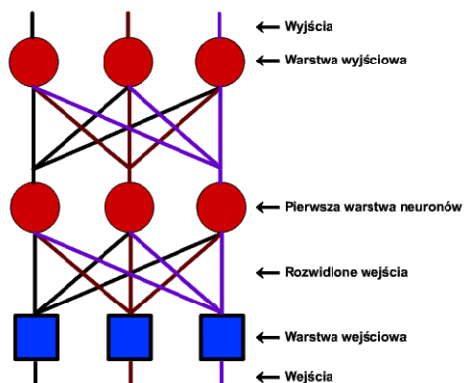
Metody Prognozowania: Sztuczne sieci neuronowe 13

Sztuczna sieć neuronowa (2)



Metody Prognozowania: Sztuczne sieci neuronowe 14

Sztuczna sieć neuronowa (3)



- warstwa (wejściowa): rozdziela sygnały wejściowe w taki sposób, aby każdy neuron pierwszej warstwy otrzymał „pełen komplet”.
- Fakt istnienia warstwy wejściowej jest tak oczywisty, że pomija się ją podając ilość warstw sieci. Oznacza to, że wyrażenie: "sieć jednowarstwowa" oznacza *de facto* sieć dwuwarstwowa, gdzie niejako "zerową" warstwą jest warstwa wejściowa, a dopiero po niej następuje pierwsza.

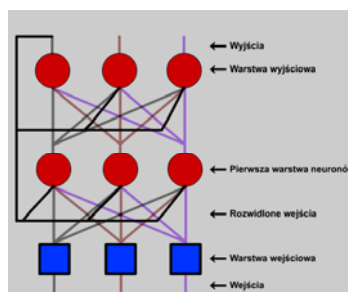
Metody Prognozowania: Sztuczne sieci neuronowe 15

Rodzaje sieci (1)

Sieć jednokierunkowa: sygnały płyną równo i spokojnie od wejścia poprzez warstwy ukryte do wyjścia.

Sieć rekurencyjna: występują tzw. **sprzężenia zwrotne**. Za ich pośrednictwem sygnały mogą po przejściu danej warstwy wracać na jej wejście, zmieniając przy tym swoje wartości, co powtarza się wiele razy, aż do osiągnięcia pewnego stanu ustalonego.

Sieć komórkowa



Metody Prognozowania: Sztuczne sieci neuronowe 16

Działanie sieci (1)

Generalnie sieci używa się na zasadzie:
„pokażmy jej coś i zobaczymy, jak zareaguje”

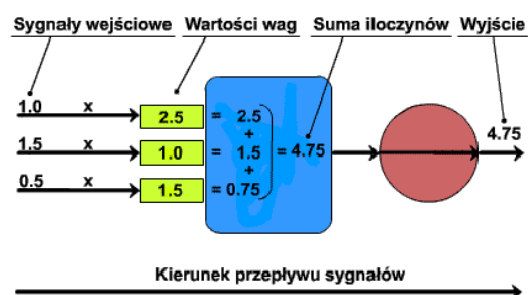


Pytanie to odpowiednio przetworzone dane (np. obraz, dźwięk, dane giełdowe, warunki pogodowe) przesyłane na wejście sieci.

Odpowiedź - znów są to pewne dane, które projektant sieci potrafi przełożyć na konkretne informacje.

Metody Prognozowania: Sztuczne sieci neuronowe 17

Działanie sieci (2)



Przykład **z pominięciem bloku aktywacji**.
(za odpowiedź neuronu uznajemy **potencjał membranowy**).

Metody Prognozowania: Sztuczne sieci neuronowe 18

Działanie sieci (3)

Waga wejścia: informacja o tym, jak duże znaczenie dla neuronu ma sygnał podany na wejście (**pamięć neuronu**).

Duża wartość dodatnia - dane wejście ma duże znaczenie "pozytywne",

Duża co do wartości bezwzględnej wartość ujemna to znak, iż wejście ma duże znaczenie "negatywne".

WAGA "NOWY"	WAGA "ŁADNY"	Co to za neuron?
5	1	Lubi nowe samochody, ich wygląd mało się liczy.
1	-1	Wszystko mu jedno...
-5	4	Samochód musi być ładny, ale stary.
4	5	Lubi nowe i piękne auta.
-4	-4	Ekstrawagancki neuron - lubi stare, brzydkie maszyny.

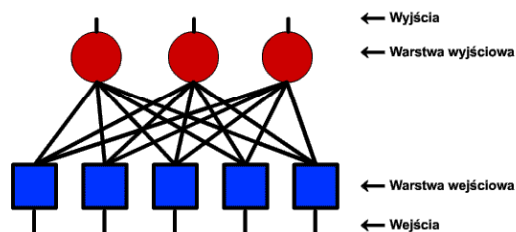
W SSN 1 - tak, 0 - nieważne, -1 - NIE

Metody Prognozowania: Sztuczne sieci neuronowe 19

Perceptron (2)

W roku 1958 Rosenblatt opracował i zbudował sztuczną sieć neuronową nazwaną **perceptronem**.

Pierwszy fizycznie skonstruowany układ symulujący sieć nerwową, który wraz ze zdolnością do uczenia się wykazywał zdolność do poprawnego działania nawet po uszkodzeniu części jego elementów. W swojej najprostszej postaci perceptron zbudowany był z dwóch oddzielnych warstw neuronów reprezentujących odpowiednio wejście i wyjście:



Metody Prognozowania: Sztuczne sieci neuronowe 20

Perceptron prosty (2)

Idea perceptronu:

1. Elementem składowym perceptronu jest sztuczny neuron, którego model matematyczny może być opisany funkcją aktywacji unipolarną lub bipolarną:

$$y = \begin{cases} 1, & \varphi > 0 \\ 0, & \varphi \leq 0 \end{cases} \quad \text{sgn}(\varphi) = \begin{cases} +1, & \varphi > 0 \\ -1, & \varphi \leq 0 \end{cases}$$

gdzie: $\varphi = \sum_{j=1}^m w_j u_j - \theta$

przy czym w_j oznacza wagę i -tego połączenia wstępującego do elementu; u_j - wartość i -tego wejścia; θ - wartość progową funkcji aktywacji.

2. Sieć perceptronową można podzielić jednoznacznie na ściśle uporządkowane i rozłączne klasy elementów zwane warstwami, wśród których wyróżnić można warstwę wejściową i wyjściową. Pozostałe noszą nazwę warstw ukrytych.

Metody Prognozowania: Sztuczne sieci neuronowe 21

Perceptron prosty (3)

3. Perceptron nie zawiera połączeń pomiędzy elementami należącymi do tej samej warstwy.
4. Połączenia pomiędzy warstwami są asymetryczne i skierowane zgodnie z ich uporządkowaniem, tzn. od warstwy wejściowej do pierwszej warstwy ukrytej, następnie od pierwszej do drugiej warstwy ukrytej, itd. aż do warstwy wyjściowej.
5. Nie ma połączeń zwrotnych.

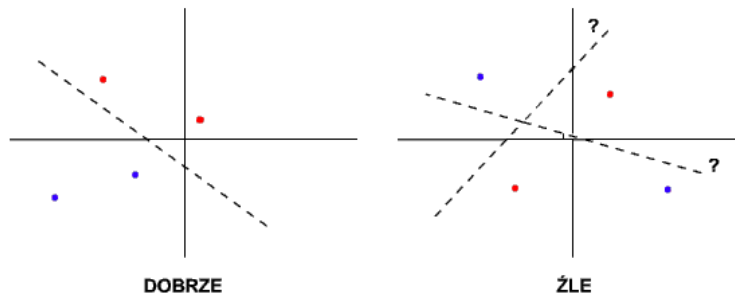
Warstwa wejściowa (**warstwa zerowa**) posiada elementy o nieco uproszczonej funkcji przejścia i jednym wejściu. Jest to swego rodzaju układ receptorów odbierających sygnały wejściowe i po wstępnym ich przetworzeniu (np. normalizacji, filtracji) przesyłających je do elementów warstwy następnej.

Perceptron zawierający jedynie warstwę wejściową i wyjściową nazywany jest **perceptronem jednowarstwowym** lub **perceptronem prostym**.

Metody Prognozowania: Sztuczne sieci neuronowe 22

Wybór struktury SSN (1)

file:///D:/!!!_pk/_Metody%20prognozowania_MP/from%20net/SSN/pliki/strukt/strukt3.html



Metody Prognozowania: Sztuczne sieci neuronowe 23

Wybór struktury SSN (2)

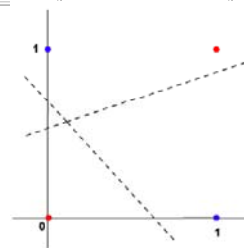
Z powodu wymogu **liniowej separowalności** zadania dla sieci jednowarstwowej nie można za jej pomocą rozwiązać na przykład zadania **XOR**, czyli **alternatywy wykluczającej**.

Zdanie A	Zdanie B	A OR B	A XOR B
0	0	0	0
0	1	1	1
1	0	1	1
1	1	1	0

XOR za pomocą sieci neuronowej:
dwa wejścia (**A i B**)
jedno wyjście (**A XOR B**).

XOR = 1 na wyjściu sygnał akceptacji (1) XOR = 0 -
sygnał odrzucenia (-1)

Sieć jednowarstwowa jest bezsilna w
konfrontacji z zadaniem XOR !!.

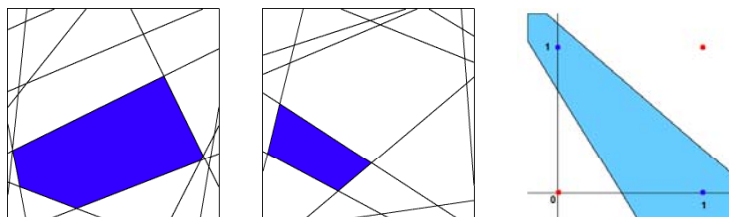


Metody Prognozowania: Sztuczne sieci neuronowe 24

Wybór struktury SSN (4)

Należy użyć sieci wielowarstwowej.

Sieć dwuwarstwowa wyznacza w przestrzeni wejść wypukły i spójny obszar, w którym znajdują się punkty odpowiadające akceptowanym przez sieć obiektom wejściowym (czyli takim, które po podaniu na wejście wywołują odpowiedź równą 1).

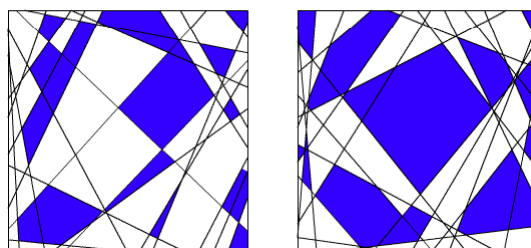


Ilość linii zależy od ilości neuronów w sieci.

Metody Prognozowania: Sztuczne sieci neuronowe 25

Wybór struktury (5)

Sieć trzywarstwowa - wyznacza ona taki obszar pozytywnej odpowiedzi, który nie musi być ani spójny, ani wypukły.



Sieć trzywarstwowa jest w stanie rozwiązać praktycznie **każdy** rodzaj zadania.

Dodam jeszcze, że takie różnice między sieciami o różnej liczbie warstw występują tylko w sieciach **nieliniowych**.

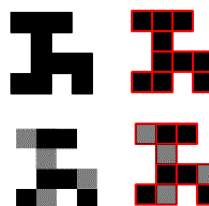
Metody Prognozowania: Sztuczne sieci neuronowe 26

Używanie SSN(1)

Wejście -> zależy od specyfikacji zadania.

Musimy ocenić, które z cech obiektu będą dla niego reprezentatywne. To zależy od tego, pod jakim względem chcemy obiekty segregować.

Wzrost człowieka	Wartość którą podamy na wejście sieci
< 140 cm	1
141 - 150 cm	2
151 - 170 cm	3
171 - 180 cm	4
> 180 cm	5



Sygnał dodatni: TAK, sygnał ujemny: NIE

Metody Prognozowania: Sztuczne sieci neuronowe 27

Używanie SSN(2)

Wyjście:

1: TAK;
-1: NIE;

Zamiana na system binarny

Co w środku ? (ilość neuronów [warstwach ukrytych](#))

W przypadku warstwy wejściowej i wyjściowej liczba neuronów jest ściśle powiązana z warunkami zadania

W warstwach ukrytych jest ona praktycznie dowolna. Przeważnie dobiera się ją drogą eksperymentów.

Metody Prognozowania: Sztuczne sieci neuronowe 28

Uczenie SSN (2)

- ✓ Struktura sieci jest ważna ale nie najważniejsza !!
- ✓ Cała wiedza sieci neuronowej tkwi w wartościach jej wag.
- ✓ Uczenie sieci polega na automatycznym (według odpowiedniego algorytmu) dobraniu takich wartości wag, przy których sieć będzie możliwie najlepiej rozwiązywała dane zadanie.

Rodzaje uczenia:

- z nauczycielem (nauczyciel zadaje pytanie i podpowiada odpowiedź);
- bez nauczyciela (uczenie nienadzorowane, sieć otrzymuje tylko obiekt wejściowy, bez narzuconego dla niego wyjścia)

Uczenie SSN (3)

Z nauczycielem:

Na wejścia podajemy kolejne liczby tego wzorca: (3, 2, 4, 5, 4).

Pierwszy neuron ma akceptować ten wzorzec, zaś drugi i trzeci - odrzucać go.

Na wyjściu zatem miały się pojawić wartości: (1, -1, -1).

Musimy teraz sprawdzić, jak sieć zareaguje na ten obiekt **przed** rozpoczęciem procesu uczenia. Np., po podaniu na wejście naszego wzorca sieć wyrzuci odpowiedź: (-1, -1, 1).

Porównując ją z odpowiedzią prawidłową widzimy, że musimy nauczyć neurony: pierwszy i trzeci. Drugi od razu odpowiedział prawidłowo. Algorytm uczenia sprawdza, na wyjściach których neuronów pojawił się błąd, a następnie poprawia wagi tych neuronów, aby ów błąd zmniejszyć (najlepiej do zera) itd ...

Uczenie SSN (4)

Bez nauczyciela:

Mamy pięć neuronów w warstwie wyjściowej, to każdy z nich jakoś zareaguje na podany sygnał.

Następnie wagi każdego z neuronów zmieniają się zgodnie z zasadą, że:

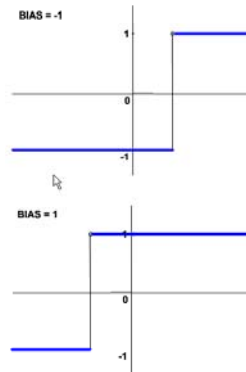
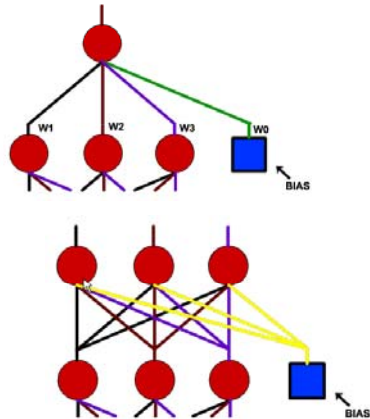
- ✓ neurony, które silnie zareagowały na sygnał wejściowy (akceptacja), zostają nauczone tak, aby reagować na ten sygnał jeszcze silniej;
- ✓ neurony, które zareagowały słabo (obojętność), nie są uczone;
- ✓ neurony, które zareagowały silną odpowiedzią ujemną (odrzućcie), są uczone tak, aby jeszcze bardziej zdecydowanie odrzucały ten sygnał.

Uczenie SSN (4)

- ✓ Obiekty podajemy na wejście w przypadkowej kolejności
- ✓ **Współczynnik uczenia:** algorytm uczenia (czy to z nauczycielem, czy też bez) zawsze musi obliczyć korekty wag, czyli zdecydować, o jaką wartość zmienić wagi neuronu, który jeszcze nie potrafi rozpoznawać danego obiektu. Korekty mogą być korygowane przez współczynnik uczenia (*learning rate*).
- ✓ **Momentum:** bezwładność procesu uczenia. (korekta wag neuronu zależy nie tylko od sygnału wejściowego i błędu, jaki neuron popełnił, ale również od tego, jak duża była korekta wag w poprzednim kroku uczenia)

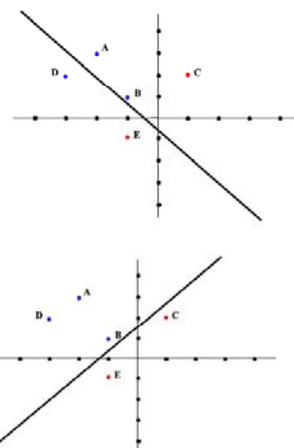
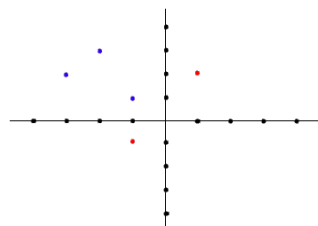
Uczenie SSN (5)

✓ **BIAS**: dodatkowe wejście neuronu ze swoją własną wartością wagi.



Metody Prognozowania: Sztuczne sieci neuronowe 33

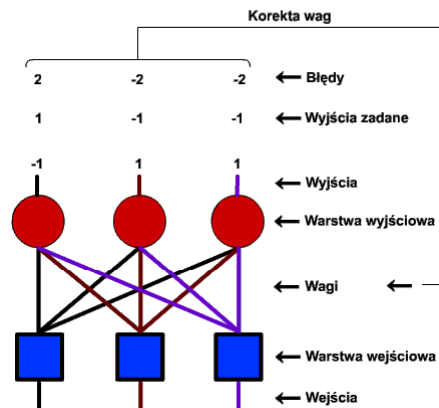
Uczenie SSN (6)



Metody Prognozowania: Sztuczne sieci neuronowe 34

Uczenie SSN (7)

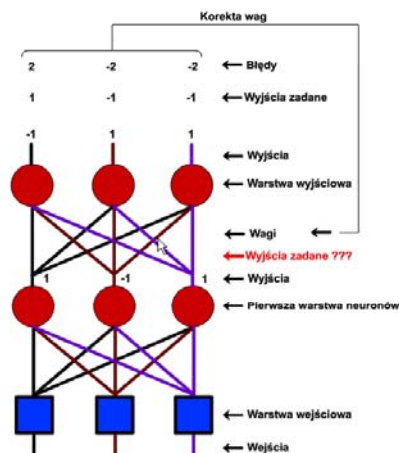
Uczenie sieci jednowarstwowej:



Metody Prognozowania: Sztuczne sieci neuronowe 35

Uczenie SSN (8)

Uczenie sieci wielowarstwowej:

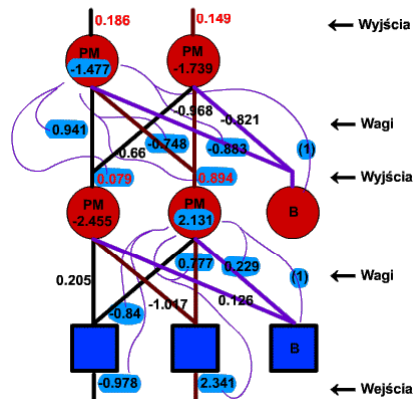


Algorytm **wstecznej propagacji błędów.**

Metody Prognozowania: Sztuczne sieci neuronowe 36

Uczenie SSN (9)

1. Algorytm wstecznej propagacji błędów (ang. *backpropagation*) wymaga, aby funkcja aktywacji była ciągła i różniczkowalna, (np. krzywa logistyczna).



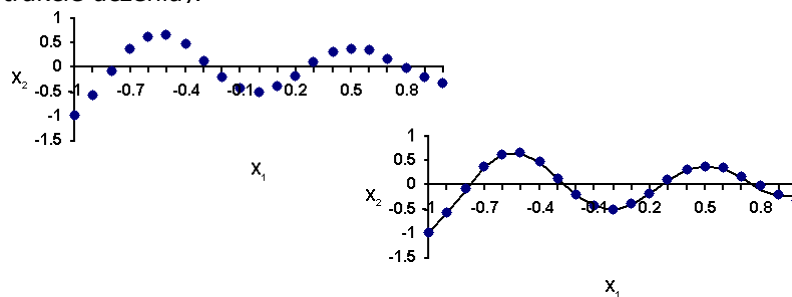
Metody Prognozowania: Sztuczne sieci neuronowe 37

Zastosowanie sieci:

Rozpoznawanie obrazów;
 Rozpoznawanie i synteza mowy;
 Analiza sygnałów radarowych;
 Kompresja obrazów;
 Prognozowanie sprzedaży;
 Prognozowanie giełdy;
 Interpretacja badań biologicznych i medycznych;
 Diagnostyka układów elektronicznych;
 Typowania w wyścigach konnych;
 Dobór pracowników;
 Selekcja celów śledztwa w kryminalistyce;
 Typowanie w wyścigach konnych.....

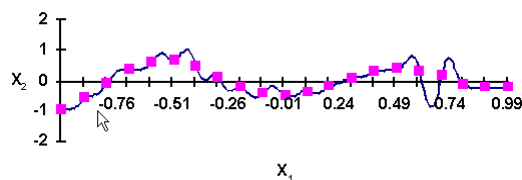
Metody Prognozowania: Sztuczne sieci neuronowe 38

Sieć skonstruowana na podstawie zbioru uczącego posiada będzie zdolność do generalizacji (która przejawiać się będzie w tym, że wyniki obliczone przez sieć będą poprawne nie tylko dla tych wartości zmiennej X_1 , które wystąpiły w zbiorze uczącym, ale również dla tych, które nie były wprowadzane na wejścia w trakcie uczenia).



Metody Prognozowania: Sztuczne sieci neuronowe 39

Jeżeli w wyniku zbytniego zwiększenia liczby neuronów ukrytych sieć utraci zdolność do generalizacji, to będzie ona prawidłowo pracować po wprowadzeniu wartości wejściowych występujących w zbiorze uczącym, lecz będzie jednocześnie generować całkowicie błędne wartości zmiennej X_2 odpowiadające tym wartościom X_1 , które **nie były reprezentowane w zgromadzonym zbiorze danych**.



Metody Prognozowania: Sztuczne sieci neuronowe 40

Sztuczne sieci neuronowe posiadają szereg cech, dzięki którym mogą stanowić przydatne narzędzie modelowania i prognozowania:

1. Zależności występujące w znacznej liczbie szeregów czasowych mają charakter nieliniowy, co stanowi podstawową przesłankę do tego, aby do ich modelowania stosować modele nieliniowe. Jednokierunkowe sieci neuronowe. Posiadają one zdolności do aproksymacji dowolnych zależności nieliniowych jak również charakteryzują się zdolnościami generalizującymi.
2. Stosowanie modeli tego typu nie wymaga znajomości postaci funkcji opisującej istniejącą prawidłowość. W związku z tym modele neuronowe mogą znaleźć zastosowanie wszędzie tam, gdzie nie jest znane dokładne prawo opisujące kształtowanie się badanych zależności.
3. Modele neuronowe mają charakter adaptacyjny. Mogą służyć do opisu zależności zmieniających się w czasie. W chwili pojawienia się nowych danych przeprowadzony może zostać proces douczenia sieci, co umożliwi uwzględnienie w tworzonym modelu informacji zawartych w najnowszych obserwacjach.

Metody Prognozowania: Sztuczne sieci neuronowe 41

Modele **nie powinny** być stosowane wtedy, gdy:

1. Badacz nie dysponuje odpowiednią liczbą obserwacji mogących stanowić podstawę do oszacowania modelu. Szacowanie sieci posiadających stosunkowo dużą liczbę parametrów na podstawie nielicznego zbioru uczącego prowadzi najczęściej do powstania modeli dopasowujących się do danych uczących, ale nie posiadających zdolności do opisu ogólnych prawidłowości istniejących w danych.
2. Trudno znaleźć uzasadnienie dla stosowania modeli neuronowych w sytuacji, w której znany jest charakter występującej prawidłowości (np. znana jest postać funkcji opisującej analizowane zjawisko). W takim przypadku oszacowanie parametrów funkcji pociąga za sobą mniejsze nakłady i prowadzi do powstania modelu łatwiejszego w interpretacji.

Metody Prognozowania: Sztuczne sieci neuronowe 42

Do najistotniejszych minusów sieci neuronowych należy zaliczyć:

1. Potrzeba odpowiedniego przygotowania danych - uzależniona od charakteru zmiennych i zastosowanego rodzaju sieci,
2. problemy związane z doбором właściwej struktury modelu neuronowego (rodzaj sieci, przyjęte modele neuronu, liczba neuronów i sposób ich połączenia),
3. konieczność wyboru właściwego algorytmu uczenia sieci,
4. wysokie nakłady czasowe związane z oszacowaniem modelu neuronowego,
5. brak (w większości przypadków) możliwości bezpośredniej interpretacji poszczególnych współczynników modelu neuronowego.