24.10.2017

Katarzyna Mlicka

Inżynieria Obliczeniowa

Wydział Inżynierii Metali i Informatyki Przemysłowej AGH

Nr indeksu 286120

SPRAWOZDANIE 2

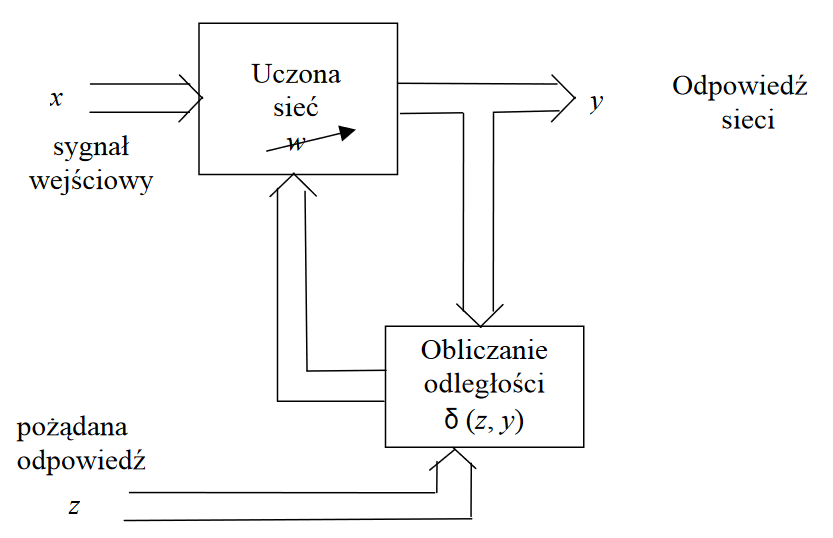
**Budowa i działanie sieci jednowarstwowej**

Celem ćwiczenia jest poznanie budowy i działania jednowarstwowych sieci neuronowych oraz uczenie rozpoznawania wielkości liter.

Sieć neuronowa jest układem neuronów odpowiednio ze sobą połączonych w warstwy, przy czym wyróżnia się zazwyczaj warstwę wejściową, warstwy ukryte i warstwę wyjściową:



Proces uczenia polega na utrwalaniu określonych zachowań na bazie przebytych doświadczeń. Jeśli chodzi o jednowarstwowe sieci neuronowe polega to na wymuszaniu na sieci określonej reakcji na określone sygnały wejściowe. Proces ten dzielimy na 3 etapy; uczenia (nadzorowanego bądź samouczenia), testowania i aplikacji.

Schemat uczenia z nauczycielem.

W tym uczeniu wagi sieci są dobierane w taki sposób, aby wyjścia z sieci zbliżyły się do wyjść żądanych.

REGUŁA DELTA

Najbardziej popularna spośród metod uczenia nadzorowanego.

Zakłada że wraz z każdym wektorem wejściowym X do neuronu podawany jest sygnał Z (wymagana odpowiedź neuronu na sygnał X) dla ciągu uczącego mającego postać:

U = <<X1, Z1>, <X2, Z2>, ... , <XN, ZN>>

W j-tym kroku neuron odpowiada na sygnał x(j) sygnałem:

y(j) = W(j)^T \* X(j)

Oraz jest określany błąd:

δ(j) = z(j) - y(j)

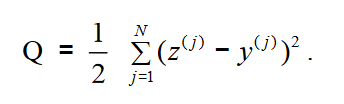
Na podstawie sygnału błędu δ(j) oraz wektora wejściowego X(j) możliwe jest takie skorygowanie wektora wag W(j) aby neuron generował sygnał bliższy zadanemu. Nowy wektor wag W(j+1) obliczany jest według reguły:

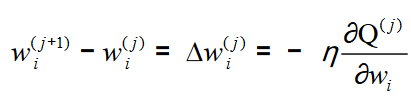
W(j+1) = W(j) + [η(j) \* δ(j) \* X(j)]

η -> współczynnik uczenia

Na początku tej metody należy dobrać losowo dobrany wektor wag W(0). Bezwarunkowo trzeba unikać przyjmowania jednakowych wartości dla różnych składowych wektora W na początku procesu uczenia. Nie dotrzymanie tego warunku prowadzi do braku postępów w początkowym etapie uczenia.

Celem tego procesu jest doprowadzenie do sytuacji w której odpowiedź neuronu y(j) będzie zgodne z wymaganymi wartościami z(j). Da sie ten proces sprowadzić do funkcji:

 Przyjmując algorytm spadku gradientu należy zmienić każdą wagę wi osobno o pewną wartość Δwi:

i -> indeks oznacza konkretną wagę neuronu

Uwzględniając że Q jest zależne od y, a y jest funkcją wektora wag W, można obliczyć pochodną funkcji złożonej dla elementów nieliniowych. Analogiczny jest sposób uczenia elementów liniowych.

W przypadku rozróżniania dużych liter od małych widać, iż nauczona sieć bardzo poprawnie rozpoznaje przezentowane wzory, nawet zniekształcone, ale także dostrzega w nich i inne litery (na przykład O w B, czy R w B). Na tej podstawie można dokładnie badać jakie cechy liter zostały uznane przez siec za istotne dla opisu litery.

Zwiększenie precyzji rozpoznawania liter czy też szerzej - znaków i symboli opartych na matrycy prezentowanej sieci można uzyskać poprzez np. powiększenie jej rozmiarów i jednocześnie odpowiednio definiując okno dyskryminatora filtrującego wyniki.

Do najczęstszych problemów uczenia należy uogólnienie, zbieżność, rozmiar sieci, zbyt duża lub mała ilość neuronów w arstwach ukrytych. Zaleca się nie stosowania więcej niż 2 warstw ukrytych. Zgodnie z zaleceniami podanymi przez Masters'a, we wstępnych pracach nad siecią dobrze jest przestrzegać 3 zasad:

1. Należy stosować tylko jedną warstwę ukrytą
2. Należy używać niewielu neuronów ukrytych
3. Należy uczyć sieć aż do granic możliwości