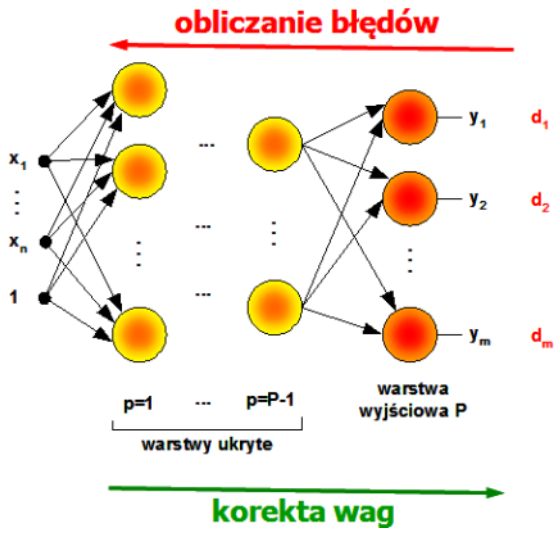
30 marca 2010
20:16
By TeWu

Metoda Propagacji wstecznej błędu



Najpierw obliczamy błędy neuronów zaczynając od nauronów w warstwie wyjściowej, potem koregujemy wagi w przeciwnej kolejności.

Wszystkie neurony muszą posiadać ciągłą funkcje aktywacji by można było je uczyć metodą propagacji wstecznej.

POCHODNE CIĄGŁYCH FUNKCJI AKTYWACJI

Sigmoidalna:

$$f'(x) = \alpha * f(x) * (1 - f(x))$$

Tangensoidalna:

$$f'(x) = \frac{\alpha}{2}(1 - f(x)^2)$$

Warstwa Ukryta:

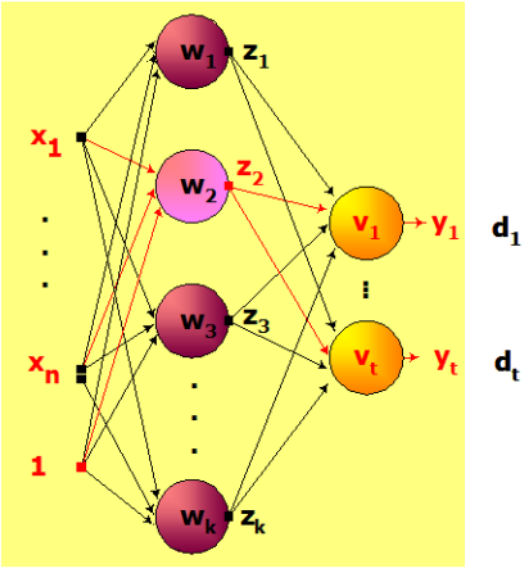
$$\Delta W_{ij}^{l} = \eta * error_i^{l} * x_{ij}^{l}$$

$$error_i^{l} = \sum_{n=1}^{nCount(l+1)} (error_n^{l+1} * W_{ni}^{l+1}) * f'(NET_i^{l})$$

Warstwa Wyjściowa:

$$\Delta W_{ij}^{l} = \eta * error_i^{l} * x_{ij}^{l}$$

$$error_i^{l} = (d_i - y_i) * f'(NET_i^{l})$$



Legenda: dawno dawno temu....

η – współczynnik uczenia

x_{ij}^{l} - wartość na j-tym wejściu i-tego nauronu w warstwie l

$error_i^{l}$ – błąd i-tego neuronu w warstwie l

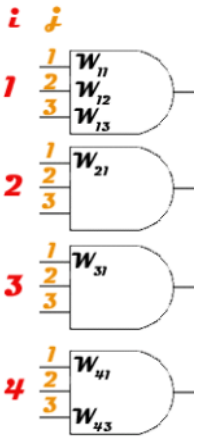
$nCount(l)$ – ilość neuronów w warstwie l

W_{ij}^{l} – j-ta waga i-tego neuronu w warstwie l

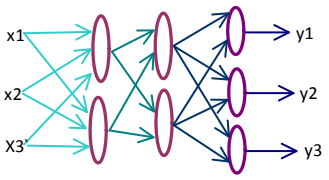
$f'(NET_i^{l})$ – wartość pochodnej funkcji aktywacji i-tego nauronu w warstwie l

y_i – warność UZYSKANA na wyjściu i-tego neuronu warstwy wyjściowej

d_i – warność OCZEKIWANA na wyjściu i-tego neuronu warstwy wyjściowej



Wszystkie wzory zakładają, że sieć skonstruowana jest tak, że j-te wejście wchodzi w j-tą wagę każdego neuronu, a j-te wyjście z warstwy l wchodzi w j-tą wagę każdego nauronu w warstwie l+1



A dlaczego nie potraktować sieci wielowarstowych metodą perceptronową ??

ODP

Wiemy czego się spodziewamy na wyjściu sieci, ale nie wiemy czego powinniśmy się podziewać na wyjściu warstwy ukrytej