ITESO

Universidad Jesuita de Guadalajara

LEARNING STAGE OF A MULTILAYER PERCEPTRON

Let N be the number of inputs, L the number of hidden neurons and M the number of outputs. Suppose you have Q different inputs with known outputs for the learning stage.

w^h is a $L \times N$ matrix and w^o is a $M \times L$ matrix filled originally with random real numbers between -1 and 1

x is a $Q \times N$ matrix with the values of the inputs of the known patterns.

d is a $Q \times M$ matrix with the values of the outputs of the known patterns.

f is a sigmoid function (in our case $f(x) = \frac{1}{1+e^{-ax}}$)

You have to repeat the folowing procedure for every row j of x and d until E is small enough

FORWARD

$$
\begin{aligned}
net^h &= w^h x_j^t \\
y^h &= f\left(net^h\right) \\
net^o &= w^o y^h \\
y &= f\left(net^o\right)
\end{aligned}
$$

BACKWARD

$$
\begin{aligned}
\delta_i^0 &= \left(\left(d_j^t\right)_i - y_i\right) y_i \left(1 - y_i\right) \\
\delta_i^h &= y_i^h \left(1 - y_i^h\right) \left[\left(w^o\right)^t \delta^o\right]_i \\
\Delta w^o &= \alpha \delta^o \left(y^h\right)^t \\
\Delta w^h &= \alpha \delta^h x_j
\end{aligned}
$$

ERROR

$$E = \|\delta^o\|$$