

APÉNDICE B

Algoritmo de Retro-Propagación

Este algoritmo es uno de los más populares para entrenar redes de perceptrones multicapa y consiste de los siguientes pasos.

Sean:

W_{ji} El peso (valor) de la conexión que va hacia el neurón j desde el neurón i .

E_p Error del patrón p , el cual está dado por:

$$E_p = \frac{1}{2} \sum_j (sd_j - sr_j)^2$$

donde:

sd_j y sr_j son la salida deseada y la salida de la red en el neurón j , respectivamente.

ET Error total del conjunto de entrenamiento, el cual está dado por:

$$ET = \sum_p E_p$$

Op_j Es la salida del j -ésimo neurón del patrón p y está dado por:

$$Op_j = F(\text{NET}_{pj})$$

$F(\text{NET}_{pj})$ Es la función de activación. En este caso utilizaremos como función de activación a la función sigmoide, la cual está dado por:

$$F(\text{NET}_{pj}) = \frac{1}{1 + e^{-\text{NET}_{pj}}}$$

donde: NET_{pj} está dado por:

$$\text{NET}_{pj} = \sum_i W_{ji} Op_i$$

donde:

Op_i es la entrada de la red si i está en la capa de entrada de la red,

Op_i es la salida del i -ésimo neurón de la capa anterior, si i no está en la capa de entrada.

1. Inicialice todos los pesos y umbrales de la red con valores aleatorios pequeños
2. Dé como entrada a la red el primer patrón p del conjunto de entrenamiento e inicialice el ET a cero.
3. Realice la evaluación de la red
4. Obtenga el error de la red E_p y acumule este error en ET
5. Ajuste los pesos (W_{ji}) de la red aplicando la siguiente regla:(regla delta generalizada).

$$W_{ji}(t+1) = W_{ji}(t) + \Delta p W_{ji}$$

donde:

$$\Delta p W_{ji} = \eta \delta_{pj} O_{pi}$$

y η es el coeficiente de aprendizaje el cual debe de ser inicializado con algún valor entre 0 y 1.

El valor de δ se calcula como sigue:

- a) Si j corresponde a la capa de salida de la red, entonces δ_{pj} está dada por:

$$\delta_{pj} = (s_{dj} - s_{rj}) F'(\text{NET}_{pj})$$

que es equivalente a:

$$\delta_{pj} = (s_{dj} - s_{rj}) [F(\text{NET}_{pj})] [1 - F(\text{NET}_{pj})]$$

- b) Si j no corresponde a la capa de salida, entonces δ es calculada como sigue:

$$\delta_{pj} = [F(\text{NET}_{pj})] [1 - F(\text{NET}_{pj})] \sum_k \delta_{pk} W_{kj}$$

6. Alimente a la red con el siguiente patrón del conjunto de entrenamiento y regrese al paso 3.
7. Una vez que a alimentado a la red con todos los patrones del conjunto de entrenamiento, regrese al paso 2.

Repita los pasos del 2 al 6 hasta que el ET sea suficientemente pequeño.