

REDES NEURONALES ARTIFICIALES

Lección 3

EL PERCEPTRÓN MULTICAPA

Diseño

El diseño de la arquitectura de la red es el primer paso para la elaboración de un perceptrón multicapa, esto implica determinar la función de activación que se utilizara, así como el numero de neuronas y capas de la red.

- El numero de neuronas en las capas de entrada y de salida, vienen dados por las variables que definen el problema. En caso de no conocer con exactitud el numero de variables relevantes se aconseja realizar un análisis(correlación, sensibilidad, etc.) para determinar este numero, tarea que puede llegar a ser un poco complicada.
- El numero de neuronas en las capas ocultas, así como la cantidad de capas ocultas son elegidos por el diseñador, al no existir un método o regla que permita determinar el numero optimo. En la mayoría de los casos se realiza mediante prueba y error.

Propagación de valores de entrada

En esta sección veremos como se propagan los valores de entrada a través de las múltiples capas de un perceptrón. Nota: fíjese bien en las nomenclaturas ya que seguirán apareciendo en esta lección.

Sea un Perceptron multicapa con C capas - $C - 2$ capas ocultas- y n_c neuronas en la capa c , para $c = 1, 2, \dots, C$. Sea $W^c = (w_{ij}^c)$ la matriz de pesos donde w_{ij}^c representa el peso de la conexión de la neurona i de la capa c para $c = 2, \dots, C$. Denotaremos a_i^c a la activación de la neurona i de la capa c . Estas activaciones se calculan del siguiente modo:

- La capa de entrada transmite las señales recibidas desde el exterior por lo tanto:

$$a_i^1 = x_i \text{ para } i = 1, 2, \dots, n_1$$

Propagación de valores de entrada

- Las capas ocultas procesan la información recibida de la capa anterior, aplicando la función de activación f a la suma de los productos de las entradas con sus respectivos pesos:

$$a_i^c = f \left(\sum_{j=1}^{n_{c-1}} w_{ji}^{c-1} a_j^{c-1} + u_i^c \right) \text{ para } i = 1, 2, \dots, n_c \text{ y } c = 2, 3, \dots, C-1$$

- La capa de salida, al igual que la oculta, aplica la función de activación f a la suma de los productos de las entradas por sus respectivos pesos:

$$y_i = a_i^C = f \left(\sum_{j=1}^{n_{C-1}} w_{ji}^{C-1} a_j^{C-1} + u_i^C \right) \text{ para } i = 1, 2, \dots, n_c$$

Propagación de valores de entrada

- La función de activación mencionada anteriormente puede ser escogida de una larga lista, pero para los perceptrones multicapa las mas comunes son la función sigmoideal(Rango [0,1]) y la función tangente hiperbólica(Rango [-1,1]):

- Función sigmoideal:

$$f_{sigm}(x) = \frac{1}{1 + e^{-x}}$$

- Función tangente hiperbólica:

$$f_{thip}(x) = \frac{1 - e^{-x}}{1 + e^{-x}}$$

