

## Redes Neuronales

### Sistema Neuronal

Neurona Biológica (Ramón y Cajal, 1888) es la unidad básica de la red nerviosa del cuerpo humano, como procesadores de información. Se compone de: las dendritas como canal de entrada de la información, el soma es el órgano de cómputo, el axón como canal de salida.

### Sistema Neuronal Artificial

Las neuronas artificiales (Rumelhart y McClelland, 1986) emulan la estructura de las redes neuronales biológicas con el fin de alcanzar una funcionalidad similar:

- Procesamiento paralelo:  
Derivado de la interconexión
- Memoria distribuida:  
Redundancia en el almacenamiento
- Adaptabilidad al entorno:  
Aprender de la experiencia

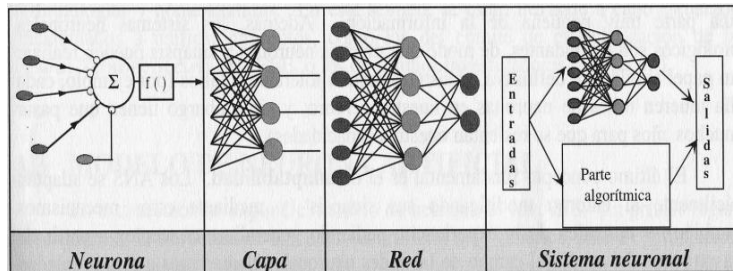


Fig. 1 Sistema global de proceso de una red neuronal

Composición:

- Un conjunto de entradas  $x_j$  y unos pesos sinápticos  $w_{ij}$ , con  $j = 1, \dots, n$
- Una regla de propagación  $h_i$  definida a partir del conjunto de entradas y los pesos

$$h_i(x_1, \dots, x_n, w_{i1}, \dots, w_{in})$$

- El umbral  $\theta_i$
- Función de activación: Representa la salida de la neurona y su estado de activación.

$$y_i = f_i(h_i) = f_i\left(\sum_{j=0}^n w_{ij}x_j\right)$$

Ejemplos de funciones de activación

i) Neuronas All-In:

Las funciones de propagación son escalonadas (salida digital).

ii) Neuronas continuas sigmoidea

Función habitual para obtener salidas continuas.

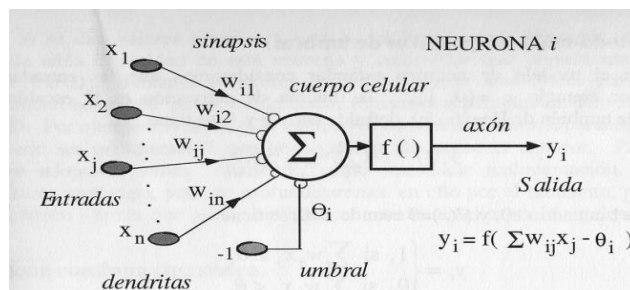


Fig.2 Modelo de neurona artificial standard

### \*) Arquitecturas de redes neuronales

Patrón de conexión de una red neuronal.

Los nodos se conectan por medio de sinapsis, estando el comportamiento de la red determinado por la estructura de conexiones sinápticas.

Estas conexiones sinápticas son direccionales (se agrupan en estructuras llamadas capas).

Tipos de capas según estructura:

De entrada, o sensorial: Compuesta por neuronas que reciben datos o señales procedentes del entorno.

Una capa de salida: Se compone de neuronas que proporcionan la respuesta de la red neuronal.

Una capa oculta: No tiene una conexión directa con el entorno. Es capaz de representar características del entorno que trata de modelar.

Tipos de capas según número:

Monocapa: Una única capa de neuronas

Multicapa: Varias capas

Tipos de capas según el flujo de datos

Redes unidireccionales (feed forward): No presenta bucles cerrados o conexiones.

Redes recurrentes o realimentadas (feed back): El flujo de información puede tener un bucle de atrás hacia adelante.

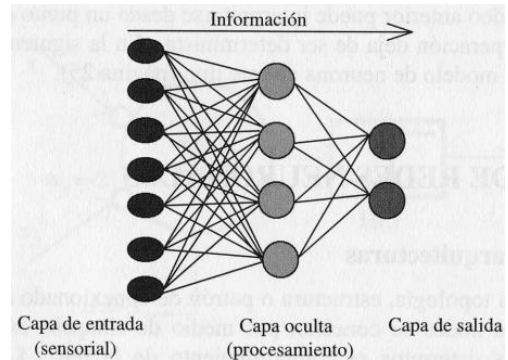


Fig.3 Arquitectura unidireccional con tres capas de neuronas: una capa de entrada, una capa oculta y una capa de salida

Tipos de relación según actualización de estados

Dinámica síncrona (todas las neuronas pertenecientes a una misma capa se actualizan a la vez)

Dinámica asíncrona (cada neurona actualiza su estado sin atender a las demás neuronas)

### Red Neuronal Artificial

Propiedades del grafo dirigido conocido como red Neuronal Artificial:

A cada nodo (neurona)  $i$  se le asocia una variable de estado  $X_i$

A cada conexión  $(i, j)$  se le asocia un peso  $w_{ij} \in \mathbb{R}$

A cada nodo (neurona)  $i$  se le asocia un umbral  $\theta_i \in \mathbb{R}$

Para cada nodo  $i$ , se define una función que depende de los pesos de sus conexiones, del umbral y de los estados de los nodos  $j$  que estén conectados con el nodo  $i$ . El valor de esta función proporciona el nuevo estado del nodo.

### El Asociador Lineal - Aprendizaje Hebbiano

Modelos de redes neuronales unidireccionales organizados en capas, cuyas finalidades sean tratar un problema de aprendizaje supervisado. Las neuronas de salida son continuas.

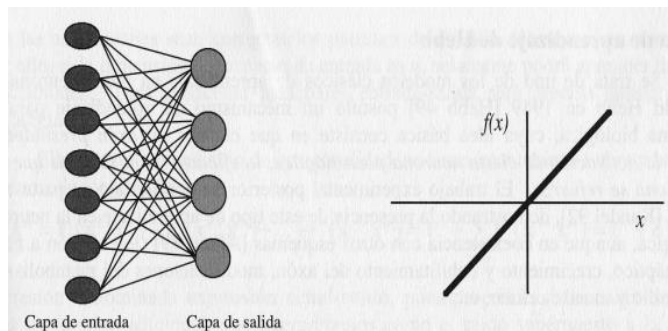


Fig.4 Arquitectura y función de activación del asociador lineal

Entradas:  $x_1, \dots, x_n$

Salidas:  $y_1, \dots, y_m$

Matriz de pesos sinápticos  $W \in M(m, n), \forall w_{ij} \in W$

Operación:  $y_i = \sum_{j=1}^n w_{ij}x_j$ , con  $i = 1, \dots, m$

El asociador lineal (consta únicamente de una capa de neuronas) debe aprender a asociar  $N$  pares entrada-salida ajustando la matriz de pesos  $W$ , de tal manera que ante entradas similares a  $x^\tau$  responda con salidas similares a  $y^\tau$ :

$$D = \{(x^\tau, y^\tau), \tau = 1, \dots, N\}$$

### Modelo clásico de aprendizaje de Hebb

Un axón presináptico causa la activación de cierta neurona pos-sináptica, la eficacia de la sinapsis que las relaciona se refuerza. Fue un modelo pionero tanto en neurociencias como en neurocomputación.

Regla de actualización de pesos:

$$\Delta w_{ij} = \varepsilon y_i x_i, \quad 0 \leq \varepsilon \leq 1$$

$$w_{ij}^{new} = w_{ij}^{old} + \Delta w_{ij}^{\tau}$$

### El Perceptrón Simple

La operación en un perceptrón simple que consta de  $n$  neuronas de entrada y  $m$  neuronas de salida se puede expresar como:

$$y_i = f\left(\sum_{j=1}^n w_{ij} x_j - \theta_i\right)$$

Las neuronas de entrada son discretas y la función de activación de las neuronas de la capa de salida es de tipo escalón. El perceptrón simple puede utilizarse como clasificador.

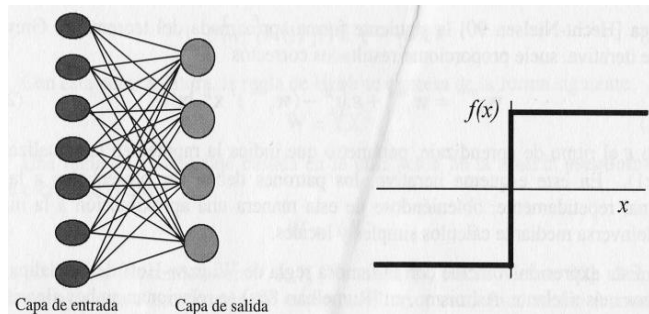


Fig.5 Arquitectura y función de transferencia de un perceptrón simple

### La Adalina

$$y_i(t) = \sum_{j=1}^n w_{ij} x_j - \theta_i \quad \text{con } i = 1, \dots, m$$

Se utiliza la regla de aprendizaje Widrow-Hoff, también conocida como LMS (Least Mean Square) o regla de mínimos cuadrados. Esta regla permite actualizaciones de los pesos proporcionales al error cometido por la neurona.

El método de optimización utilizado por la regla LMS es el denominado descenso por el gradiente, el cual puede ser visto como un optimizador local en un espacio de búsqueda continuo. El objetivo del aprendizaje consiste en encontrar la configuración de los mismos que corresponde al mínimo global de la función de error o función de coste definida. Con frecuencia en una red neuronal genérica nos vamos a tener que conformar con un mínimo local suficientemente bueno.

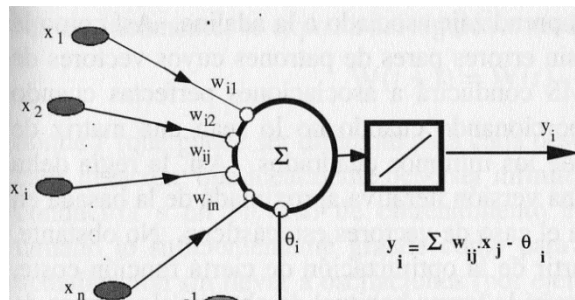


Fig.6 Neurona lineal de la Adalina

### El Perceptrón Multicapa

El perceptrón multicapa con una única capa de neuronas ocultas puede discriminar regiones convexas. Por otra parte, el perceptrón multicapa con dos capas de neuronas ocultas es capaz de discriminar regiones de forma arbitraria.

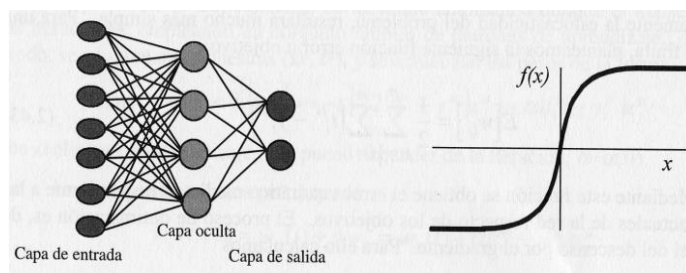


Fig.7 Arquitectura y función de activación para el perceptrón multicapa

El perceptrón multicapa o MLP (Multi-Layer Perceptron) se suele entrenar por medio de un algoritmo de retropropagación de errores o BP (Back Propagation) de ahí que dicha arquitectura se conozca también bajo el nombre de red de retropropagación.