



Introducción al Deep Learning

Dr. Ing. Gabriel Hermosilla Vigneau

Redes Neuronales Artificiales

Parte 1

Redes neuronales artificiales

- Una red neuronal es un procesador masivamente paralelo y distribuido compuesto por unidades procesadoras elementales (neuronas). Ésta tiene una tendencia natural a almacenar conocimiento experimental y a dejarlo disponible para su uso (Aleksander and Morton 1990).
- Se asemeja al cerebro humano en dos aspectos:
 - Adquiere conocimiento de su medio ambiente a través de un proceso de aprendizaje.
 - Las interconexiones entre neuronas, conocidas como pesos sinápticos, se utilizan para almacenar el conocimiento adquirido.

Redes neuronales artificiales

Algunas características:

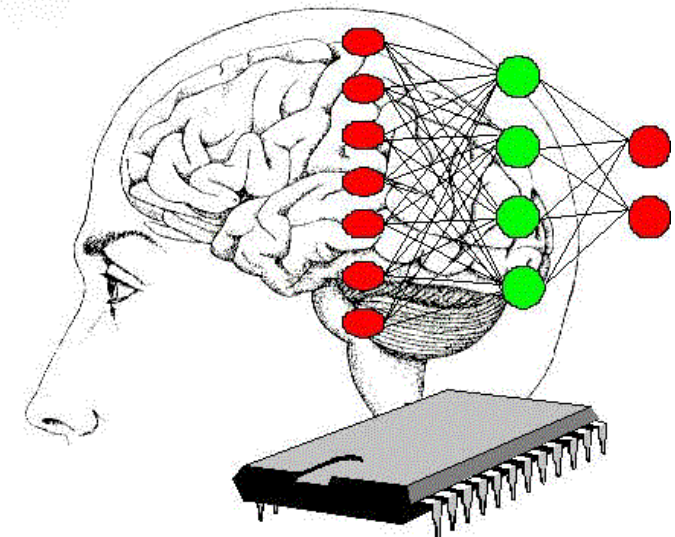
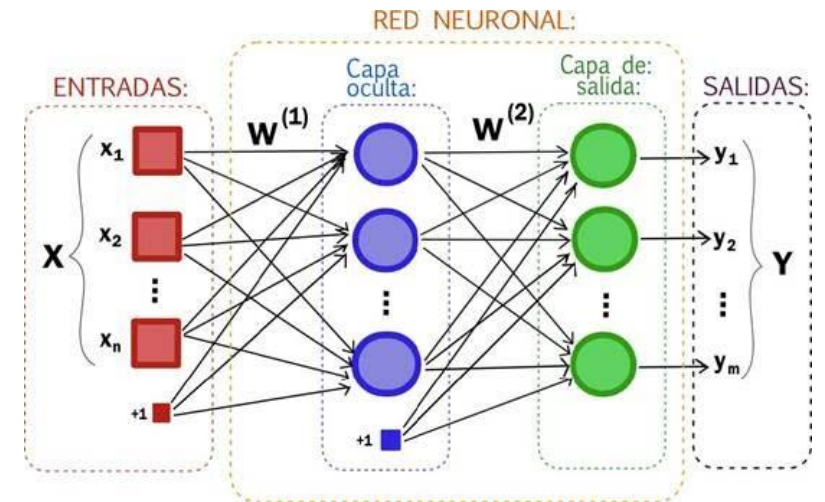
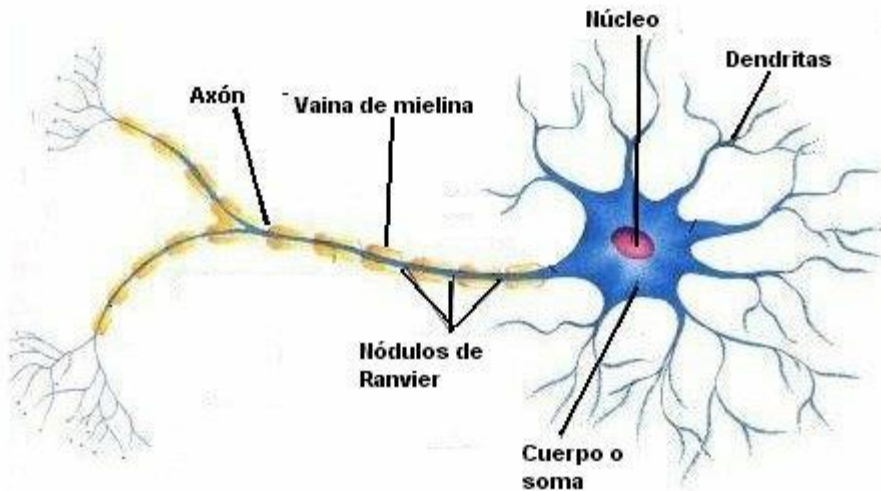
- Algoritmo de Aprendizaje:
 - Procedimiento usado para realizar el proceso de aprendizaje.
 - Su función es modificar los pesos sinápticos de la red con el objetivo de obtener un comportamiento deseado.
- Topología de la red:
 - Es posible modificar la topología de la red (número de neuronas y sus interconexiones).
- Habilidad de aprender y capacidad de generalizar:
 - Después del aprendizaje la red debe responder adecuadamente a nuevas entradas (no utilizadas en el aprendizaje).

Propiedades de las Redes neuronales artificiales

- No-linealidad
- Mapa entrada-salida
 - La red aprende a partir de ejemplos, construyendo un mapa de entrada-salida
- Adaptiva
 - Una red entrenada para operar en un ambiente específico puede ser re-entrenada
- Confiabilidad de la respuesta
 - Una red neuronal puede ser entrenada para seleccionar ciertos patrones, y para establecer el grado de confianza en la decisión
- Información de contexto
 - El conocimiento se representa por la estructura y el estado de activación de la red

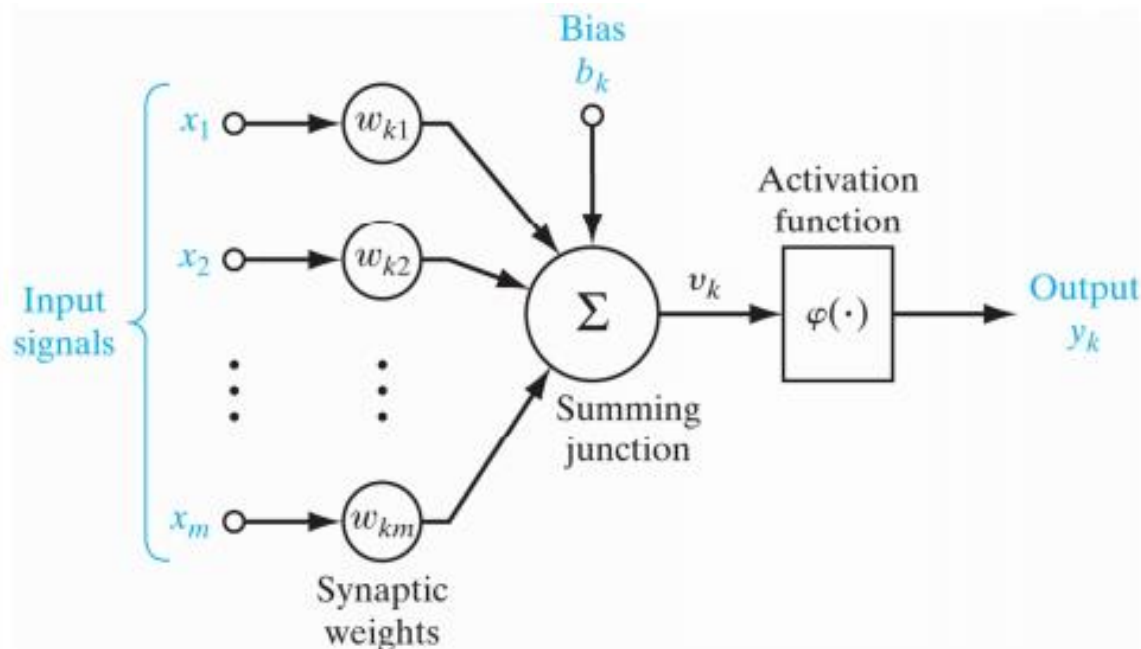
La Neurona

- Soma: Cuerpo de la célula
- Axón: Líneas de transmisión
- Dendritas: Zonas de recepción
- Sinapsis: Interacción entre neuronas



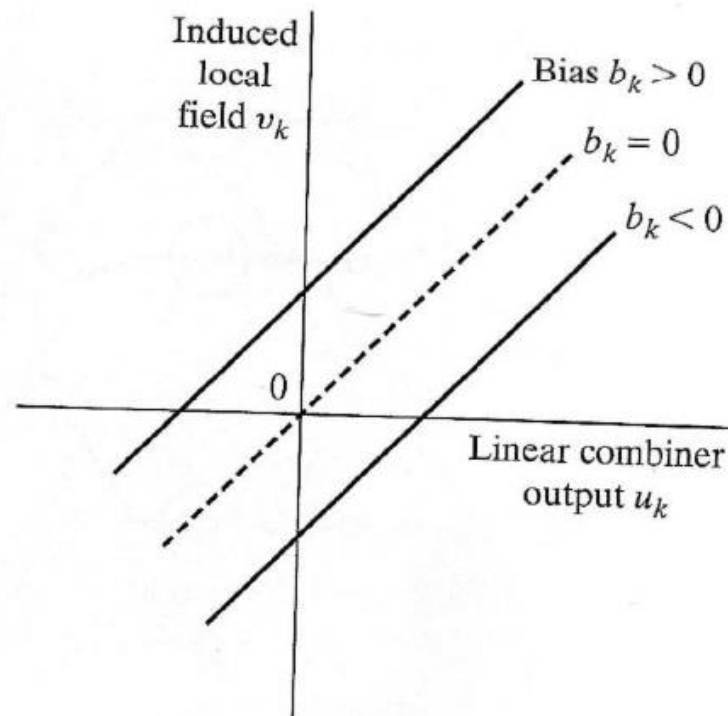
Modelo de una neurona

- Señales de entrada x_j
- Pesos sinápticos w_{kj}
- Sumador o combinador lineal $v_k = \sum_{j=1}^m w_{kj} x_j$
- Función de activación $y_k = \varphi(u_k + b_k)$, donde b_k es una constante (bias-sesgo)



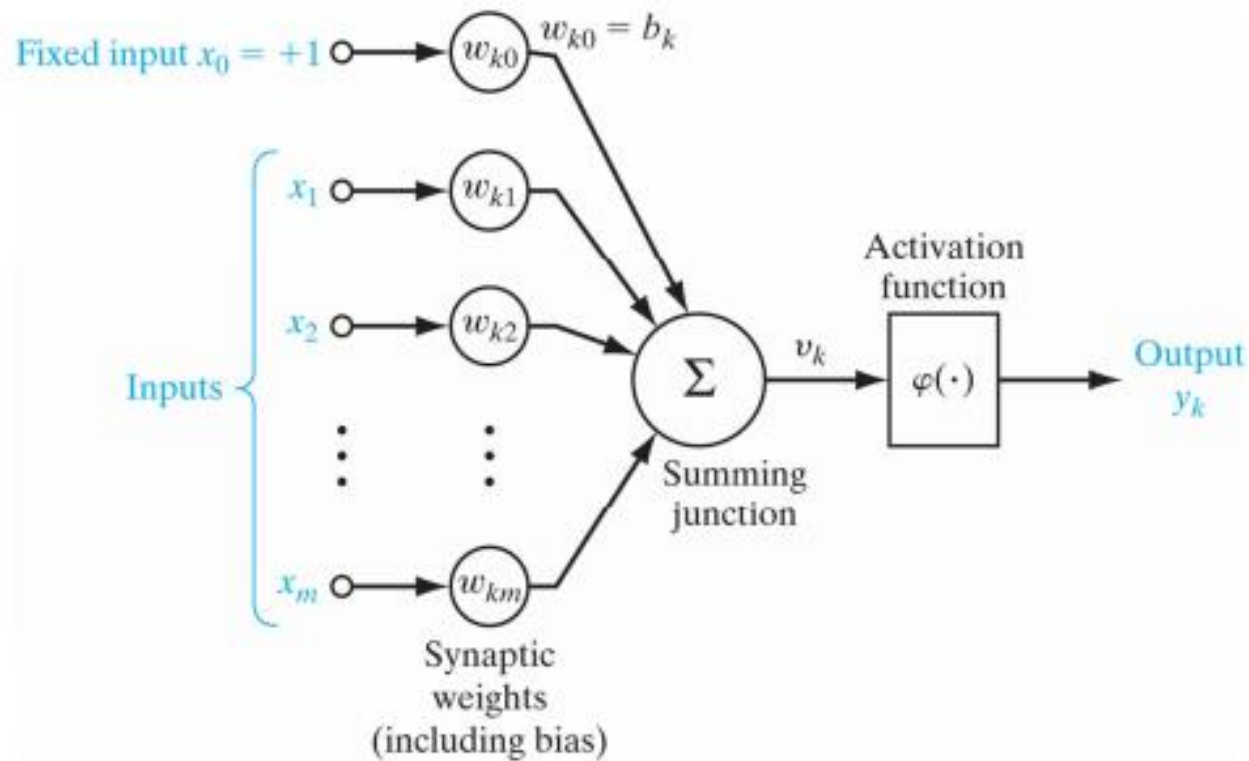
Modelo de una neurona

- La constante b_k tiene el efecto de aplicar una transformación afín a la salida del combinador lineal.
- Como resultado de esta transformación, el gráfico v_k no pasa por el origen



Modelo de una neurona

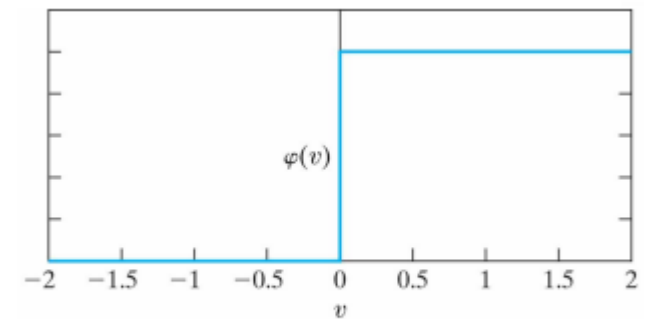
- Modelo no lineal de una neurona artificial ($w_{k0} = b_k$)



Funciones de activación

- Función umbral o de Heaviside

$$\varphi(v) = \begin{cases} 1 & \text{si } v \geq 0 \\ 0 & \text{si } v < 0 \end{cases}$$



Neuronas binarias

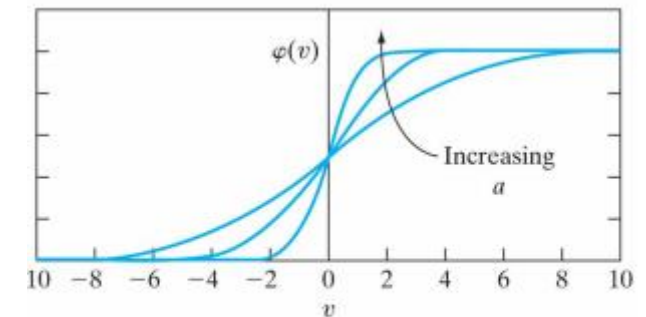
- Función sigmoïdal

- Función logística [0, 1]

$$\varphi(v) = \frac{1}{1 + e^{-av}}$$

- Tangente hiperbólica [-1, 1]

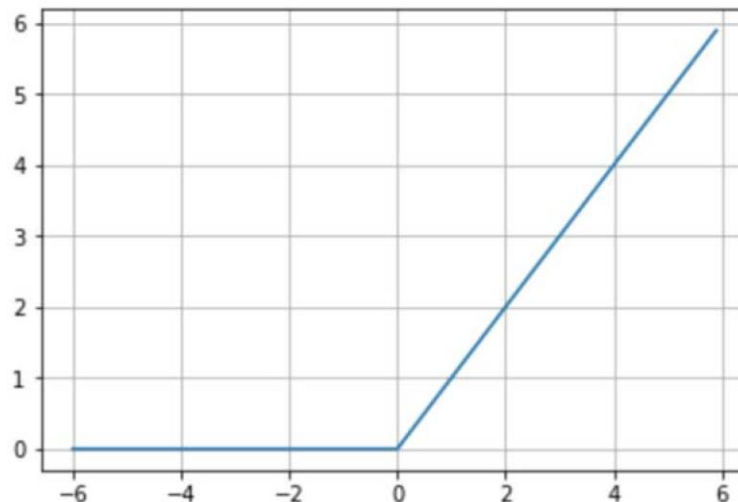
$$\varphi(v) = \tanh v = \frac{e^v - e^{-v}}{e^v + e^{-v}}$$



Neuronas sigmoïdales

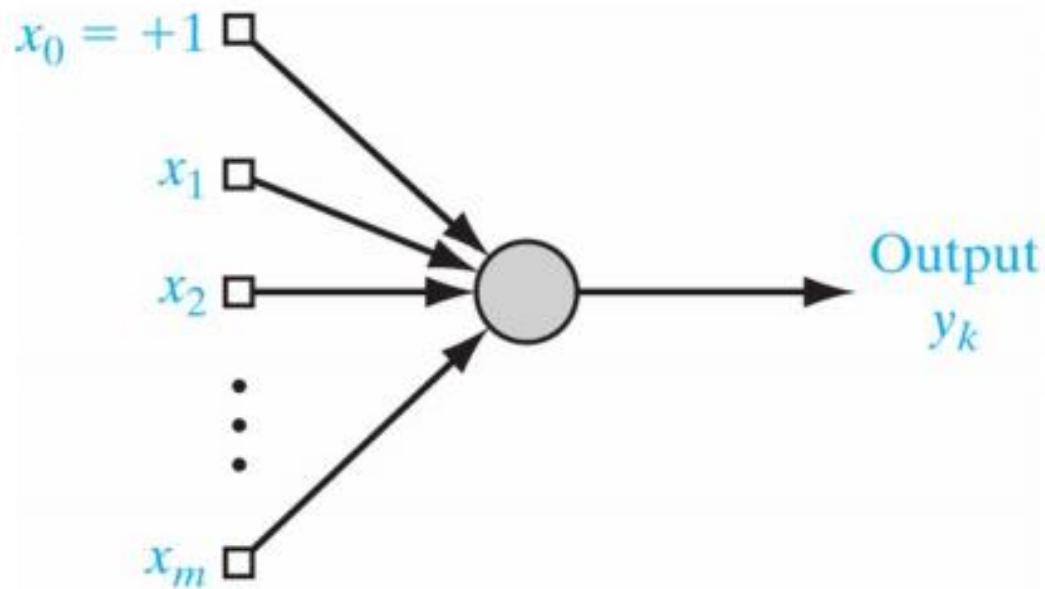
Funciones de activación: ReLU

- La función de activación Rectified Linear Unit (ReLU) es una transformación muy interesante que activa un solo nodo si la entrada está por encima de cierto umbral.
- El comportamiento por defecto y más habitual es que mientras la entrada tenga un valor por debajo de cero, la salida será cero, pero cuando la entrada se eleva por encima, la salida es una relación lineal con la variable de entrada de la forma $f(x) = x$.
- La función de activación ReLU ha demostrado funcionar en situaciones diferentes, y actualmente es muy usada.



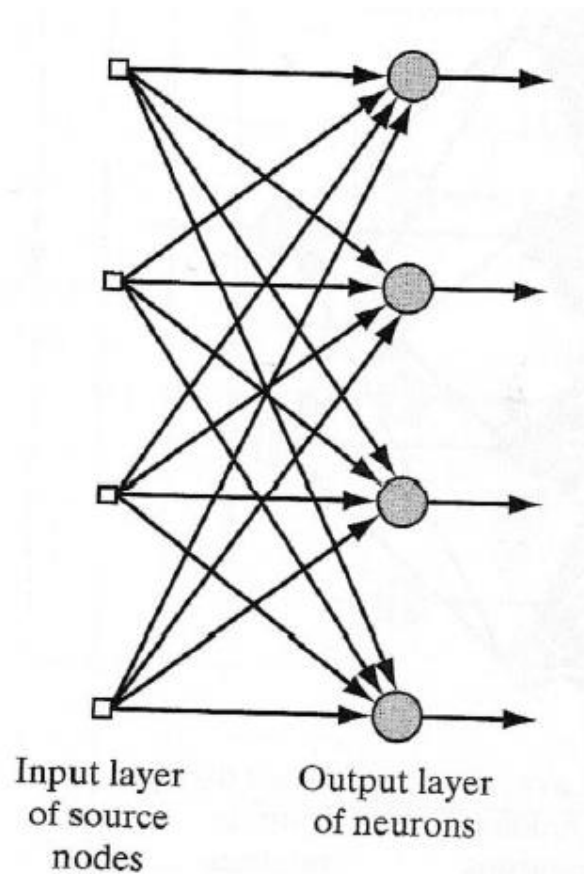
Arquitecturas de Redes Neuronales

- Perceptrón



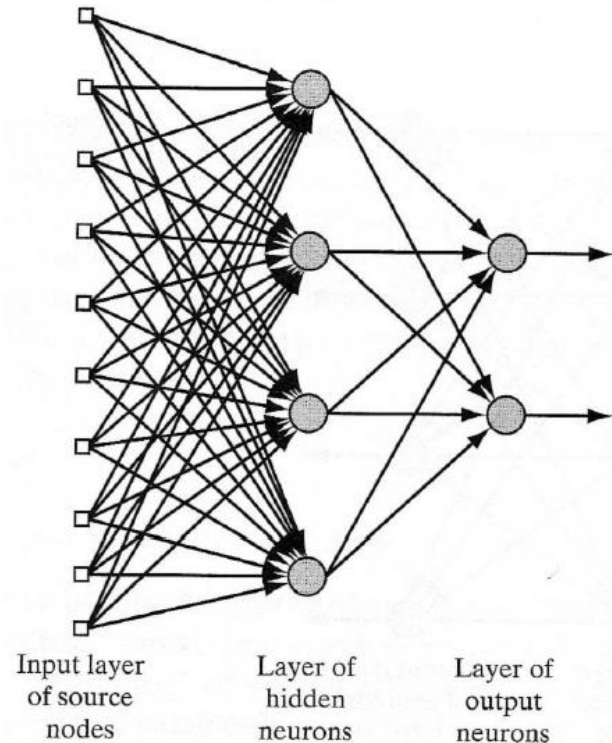
Arquitecturas de Redes Neuronales

- Redes pre-alimentadas de una capa (Single-layer feedforward networks)



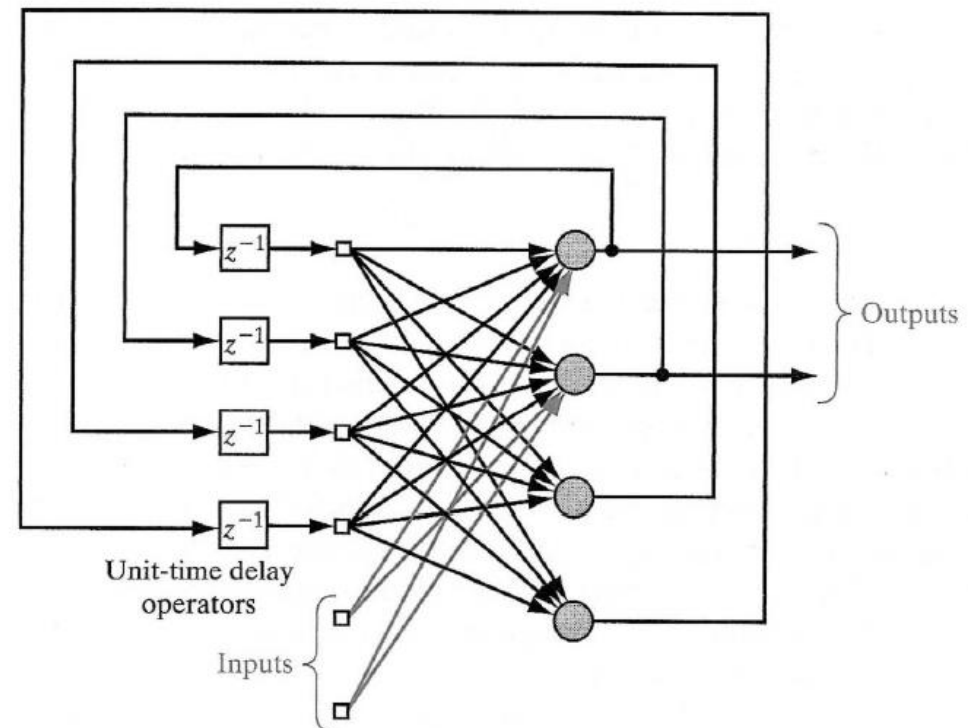
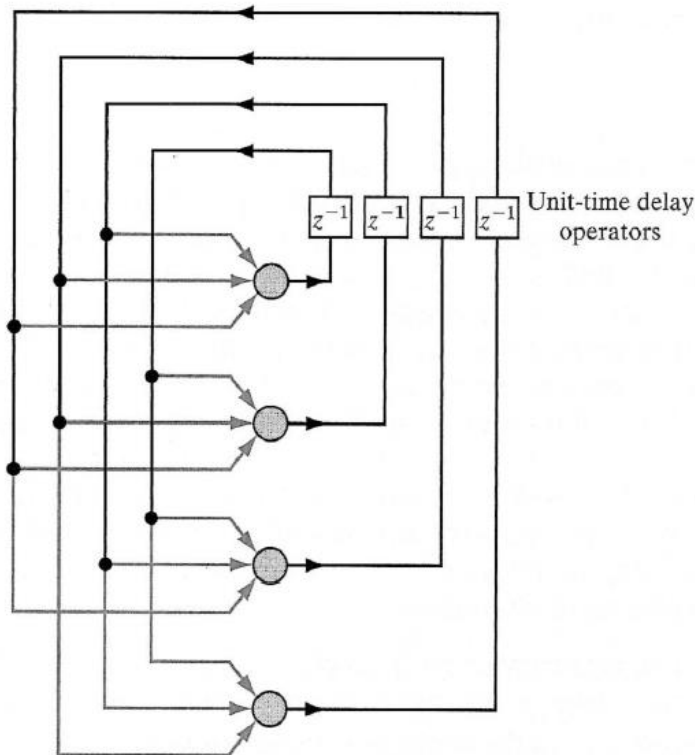
Arquitecturas de Redes Neuronales

- Redes Multicapas (Multi-layer feedforward networks)
 - Poseen muchas capas ocultas. El término “oculta” se refiere a que no es visible como las capas de entrada y salida. La figura muestra una red 10-4-2, 10 nodos de entrada, 4 nodos ocultos y 2 nodos de salida.



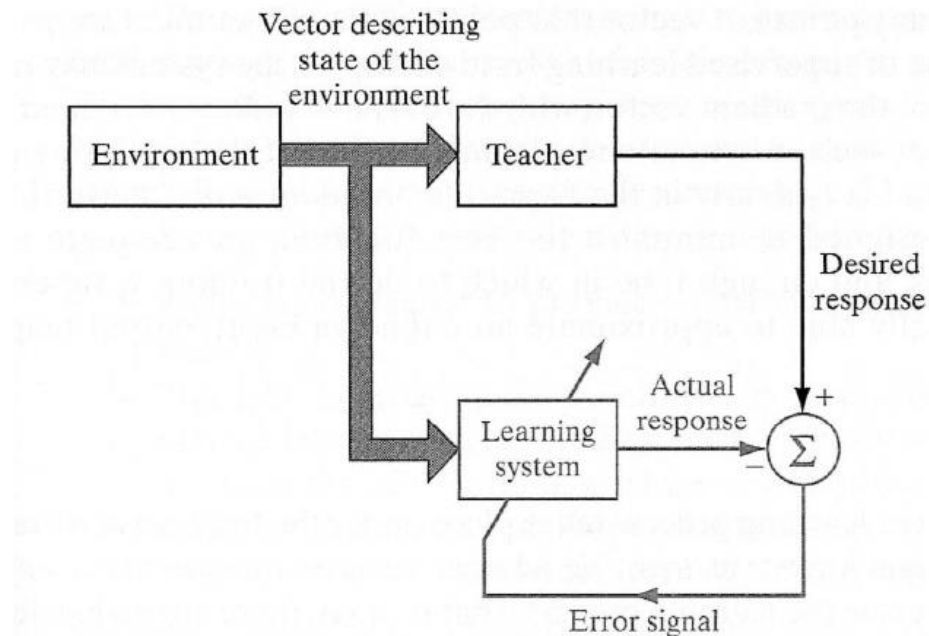
Arquitecturas de Redes Neuronales

- Redes Recurrentes (Recurrent networks)
 - Se distinguen por tener al menos un lazo cerrado. En la figura de la izquierda se observa una red sin unidades ocultas ni auto-realimentación. La figura de la derecha posea unidades ocultas.



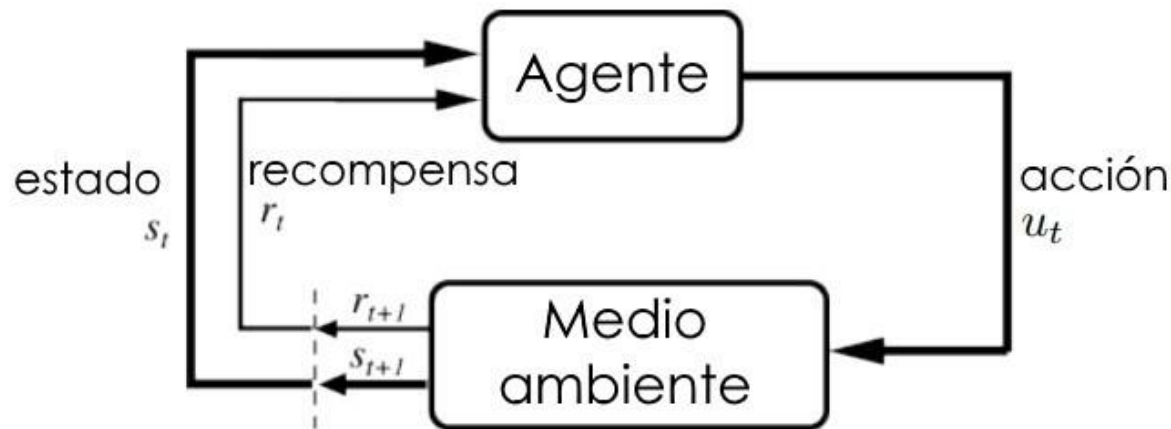
Formas de aprendizaje de redes neuronales

- Aprendizaje supervisado (con instructor)
 - El instructor tiene conocimiento del ambiente en la forma de ejemplos de entrada-salida (deseada).
 - Generalmente se supervisa la salida mediante etiquetas de los ejemplos.
 - La relación de entrada-salida se aprende a través de minimizar un indicador de desempeño (función de costo).



Formas de aprendizaje de redes neuronales

- Aprendizaje reforzado
 - Un agente genera acciones sobre un ambiente, dependiendo de estas acciones se cambian estados y se generan recompensas al agente.
 - Al principio, el agente no entiende la política que debe maximizar. Al continuar realizando acciones en el ambiente, el agente comienza a obtener una recompensa acumulada, es decir comienza a aprender.
 - Para optimizar la política deberá cambiar sus estados, mediante una exploración a nuevos estados y explotando (maximizando) su recompensa.



[Figure source: Sutton & Barto, 1998]

Formas de aprendizaje de redes neuronales

- Aprendizaje no supervisado
 - Aprendizaje auto-organizado (sin instructor)
 - Los parámetros libres se optimizan con respecto a una medida independiente de la tarea a realizar.
 - Clustering

