



CURSO:

**Taller de Machine Learning
para el análisis y visualización
en Power BI**

Módulo II. Métodos de Machine Learning

Laboratorio de Inteligencia Artificial

Tema 04 – Deep Learning

**Profesor: Saúl Domínguez Isidro, PhD.
Contacto: saul.dominguez@lania.edu.mx**

Objetivo

Conocer conceptos básicos del aprendizaje profundo

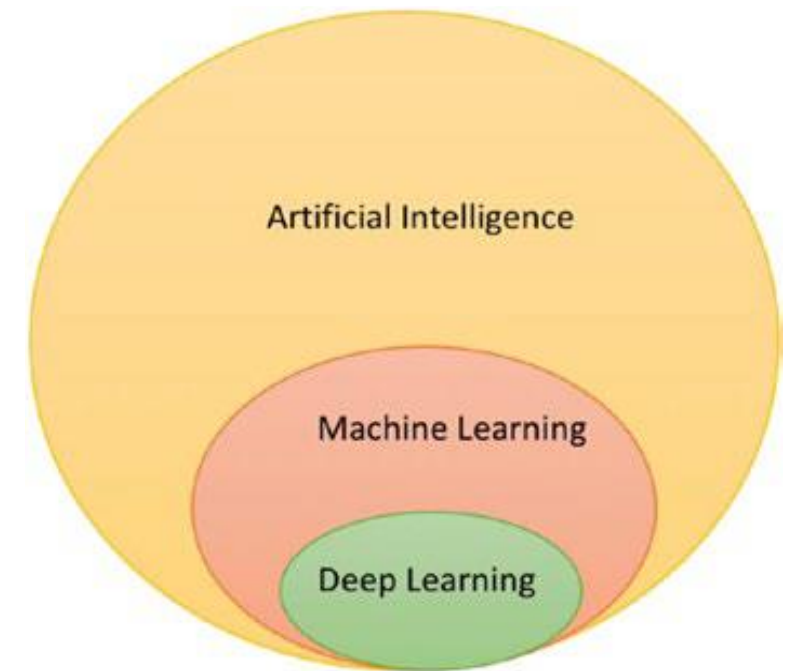


Contenido

- Introducción
- Redes neuronales artificiales
- Anatomía de una red neuronal convolucional

Aprendizaje

- El aprendizaje profundo cae bajo el paraguas del aprendizaje automático, que es un subconjunto de la inteligencia artificial (IA).
- El aprendizaje profundo es un término para las tecnologías que utilizan algoritmos de redes neuronales artificiales (ANN). Los expertos consideran que el aprendizaje profundo y las ANN son la misma cosa y usan los términos indistintamente

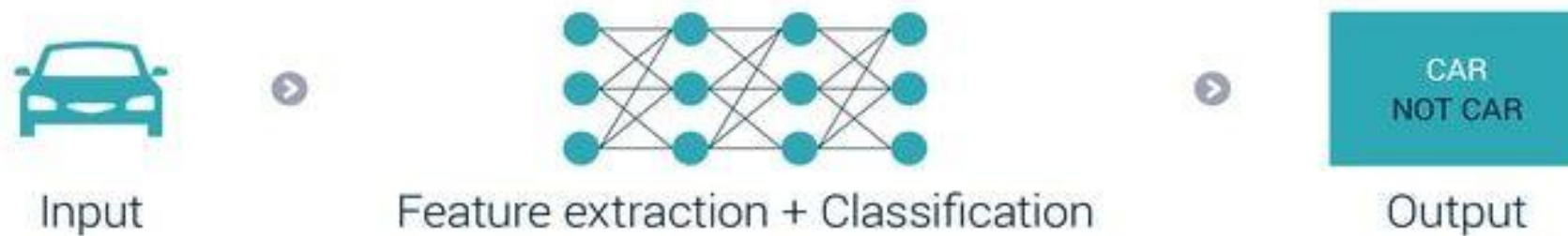


Diferencias con ML

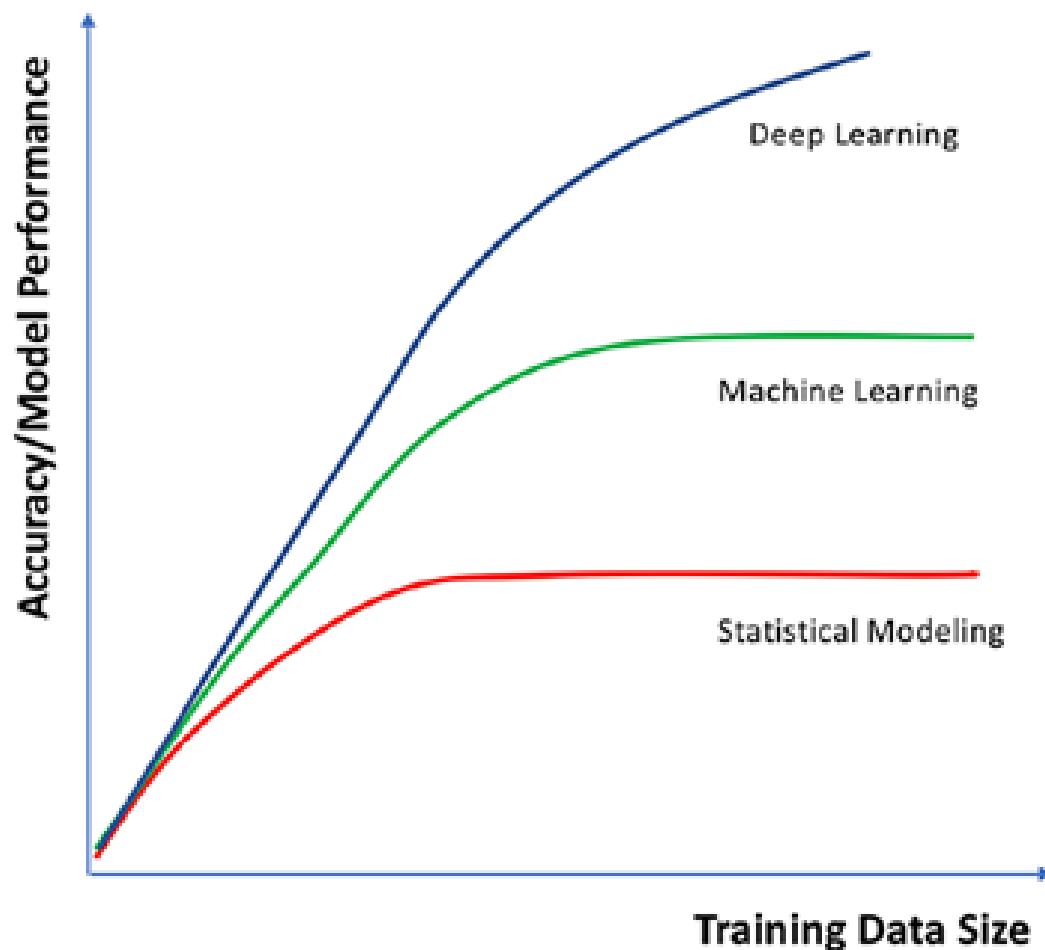
Machine Learning



Deep Learning



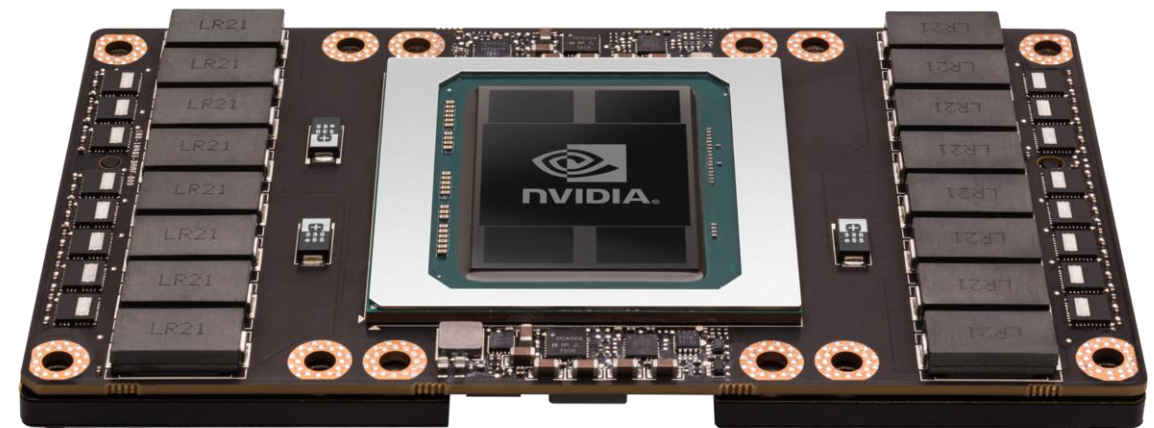
Perspectiva de DL



Tecnología a favor

- Hacer que una red neuronal funcione rápidamente es difícil. Cientos o miles de neuronas deben interactuar entre sí en paralelo. Dependiendo de la tarea, las CPU tradicionales podrían tardar semanas en generar una predicción a partir de una ANN. Con las GPU, la misma tarea que llevó semanas solo puede llevar días u horas.

Las GPU fueron construidas por primera vez por NVIDIA para manejar las operaciones masivamente paralelas que los videojuegos



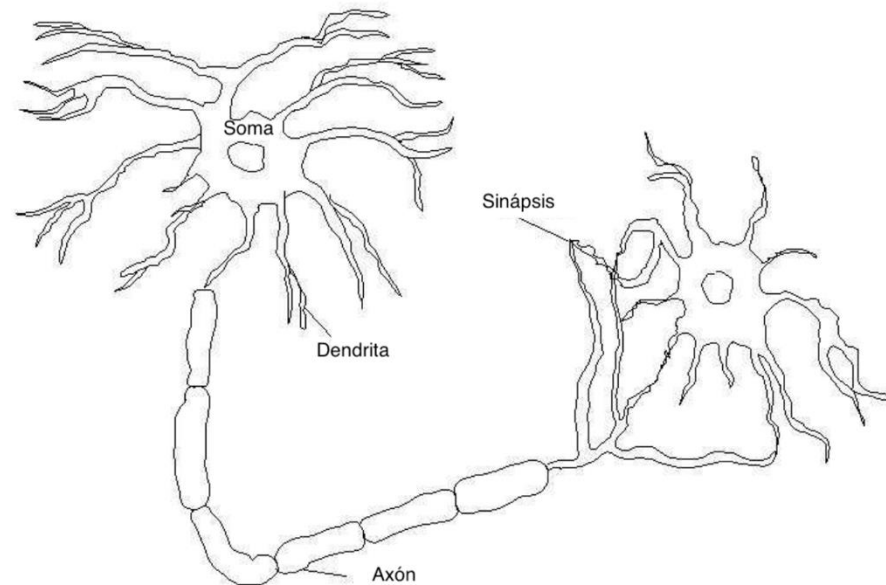
En 2009, Andrew Ng y varios otros descubrieron que podían usar GPU para el aprendizaje profundo a gran escala.

El punto de inflexión en DL

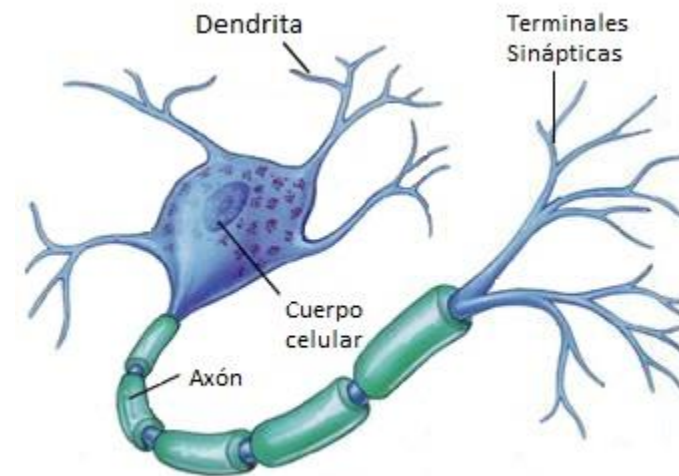
- 2006 - Geoffrey Hinton, et al. - punto de inflexión para las ANN.
 - Introdujo un algoritmo que podría afinar el procedimiento de aprendizaje utilizado para entrenar a las ANN con múltiples capas ocultas. La clave fue utilizar un algoritmo de descenso de gradiente que podría ajustar cada capa de la ANN por separado.
 - El otro descubrimiento clave optimizó la configuración inicial de los pesos. Esto permitió que los datos de alta dimensión, o datos con muchas características, se convirtieran en datos de baja dimensión, aumentando el poder predictivo.

Inspiración biológica

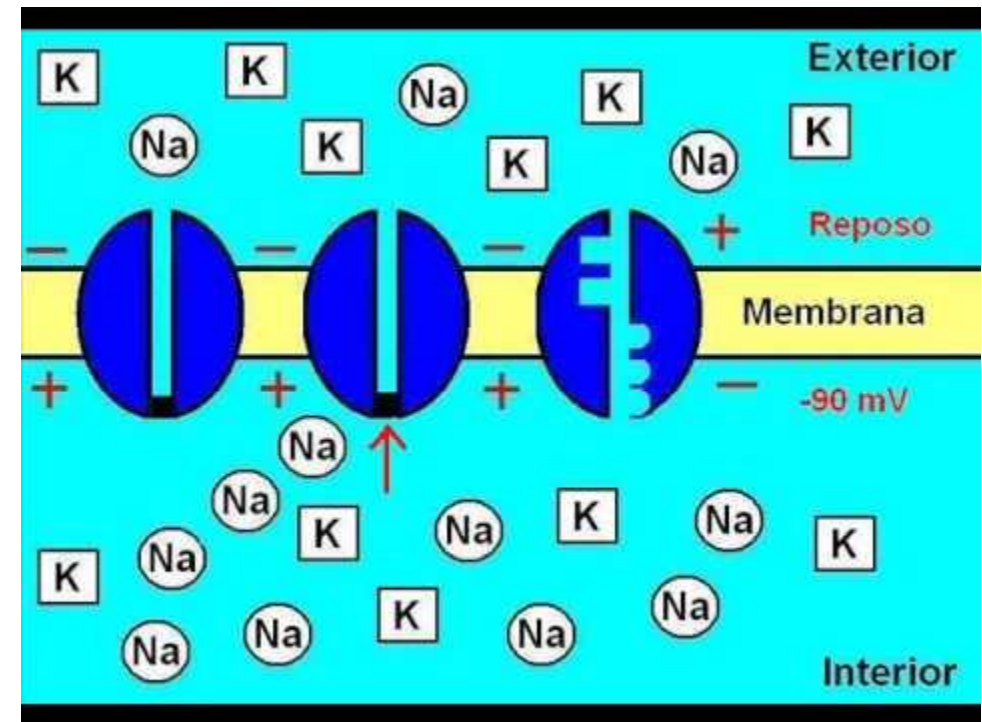
- Las bases de las Redes Neuronales Artificiales (RNA) descansan en la conceptualización de las neuronas, el cómo reciben datos, los procesan, obtienen un resultado y transmiten la información hacia otras neuronas, formando circuitos o redes



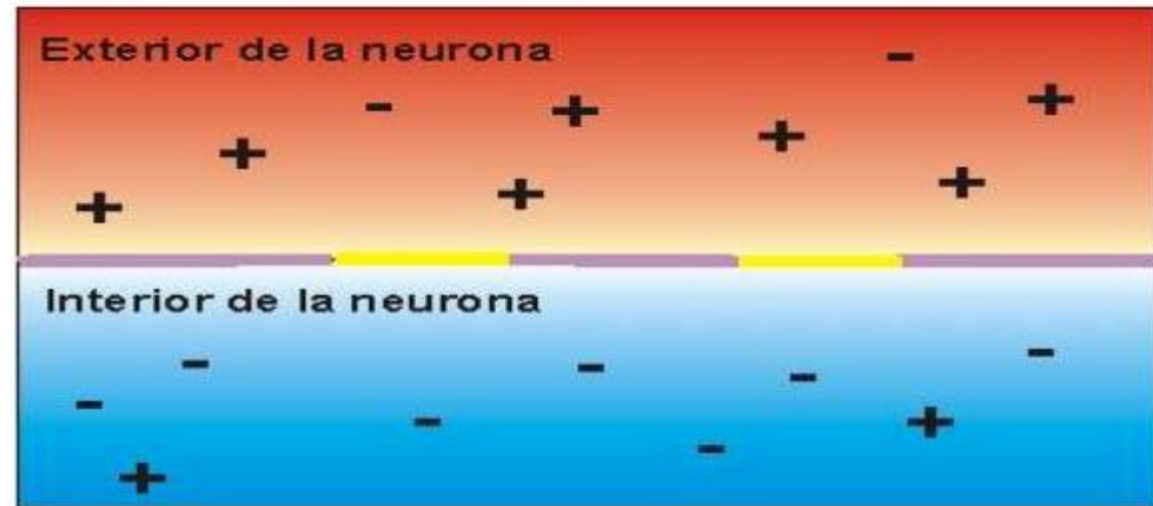
- Las dendritas reciben señales de entrada
- Estas señales son procesadas por el soma (cuerpo celular),
- El soma produce una señal de salida que es transmitida a otras neuronas mediante el axón
- Las señales que se transmiten son de naturaleza química (en el exterior) y eléctrica



- La unión entre dos neuronas mediante estas señales se conoce como sinapsis
- Cada neurona mantiene un líquido en su interior, diferente en composición al exterior
- A esta diferencia se le llama potencial de reposo

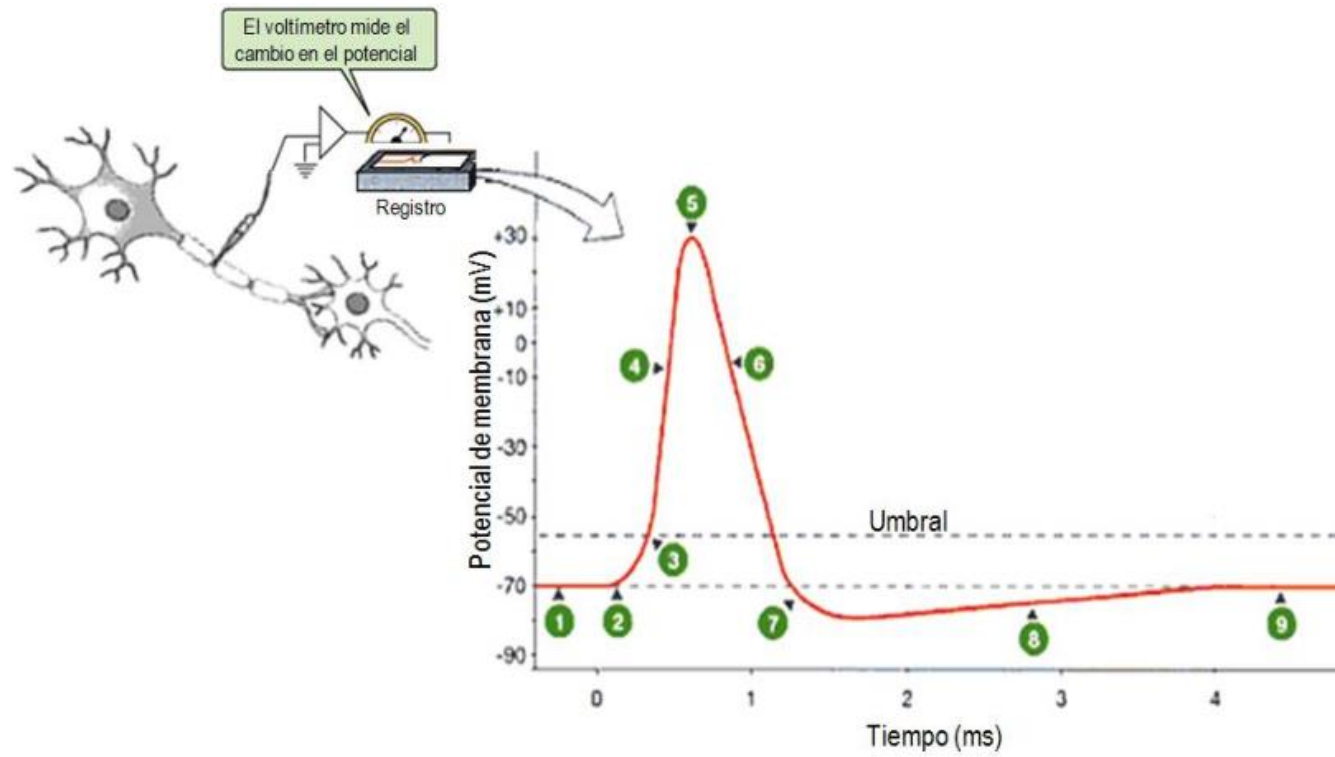


- Al llegar señales procedentes de otras neuronas a través de las dendritas, actúan de manera acumulativa, bajando ligeramente el potencial de reposo de la neurona en cuestión

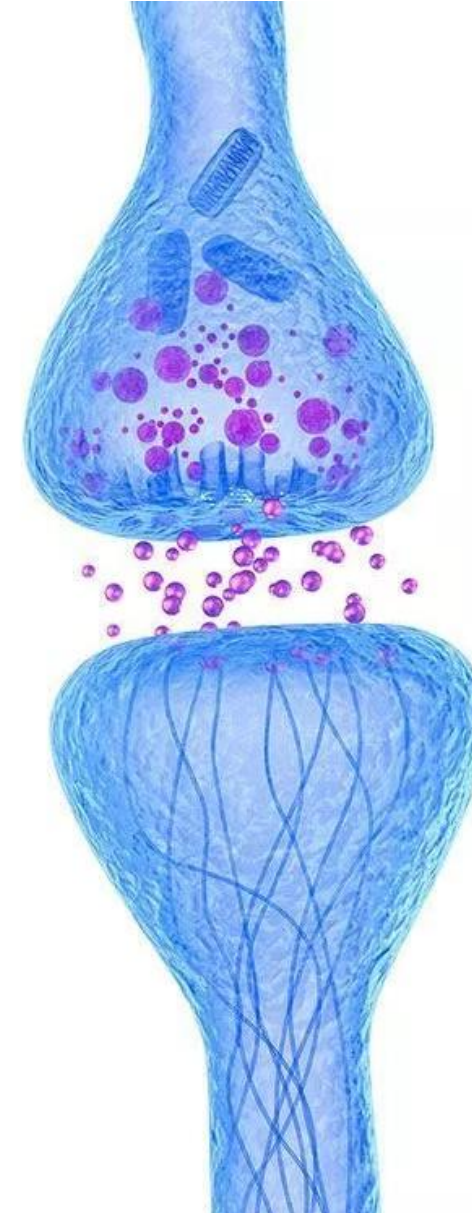


Este potencial modificado afecta a su vez la permeabilidad de la membrana

- La inversión de voltaje se conoce como potencial de acción
- El potencial de acción se propaga a lo largo del axón



- Algunas señales de entrada estarán activas y otras en reposo
- La suma de los efectos excitadores e inhibidores determina si la neurona será o no estimulada



1943

Modelo computacional de una Neurona biológica



Warren McCulloch



Walter Pitts

Elementos importantes

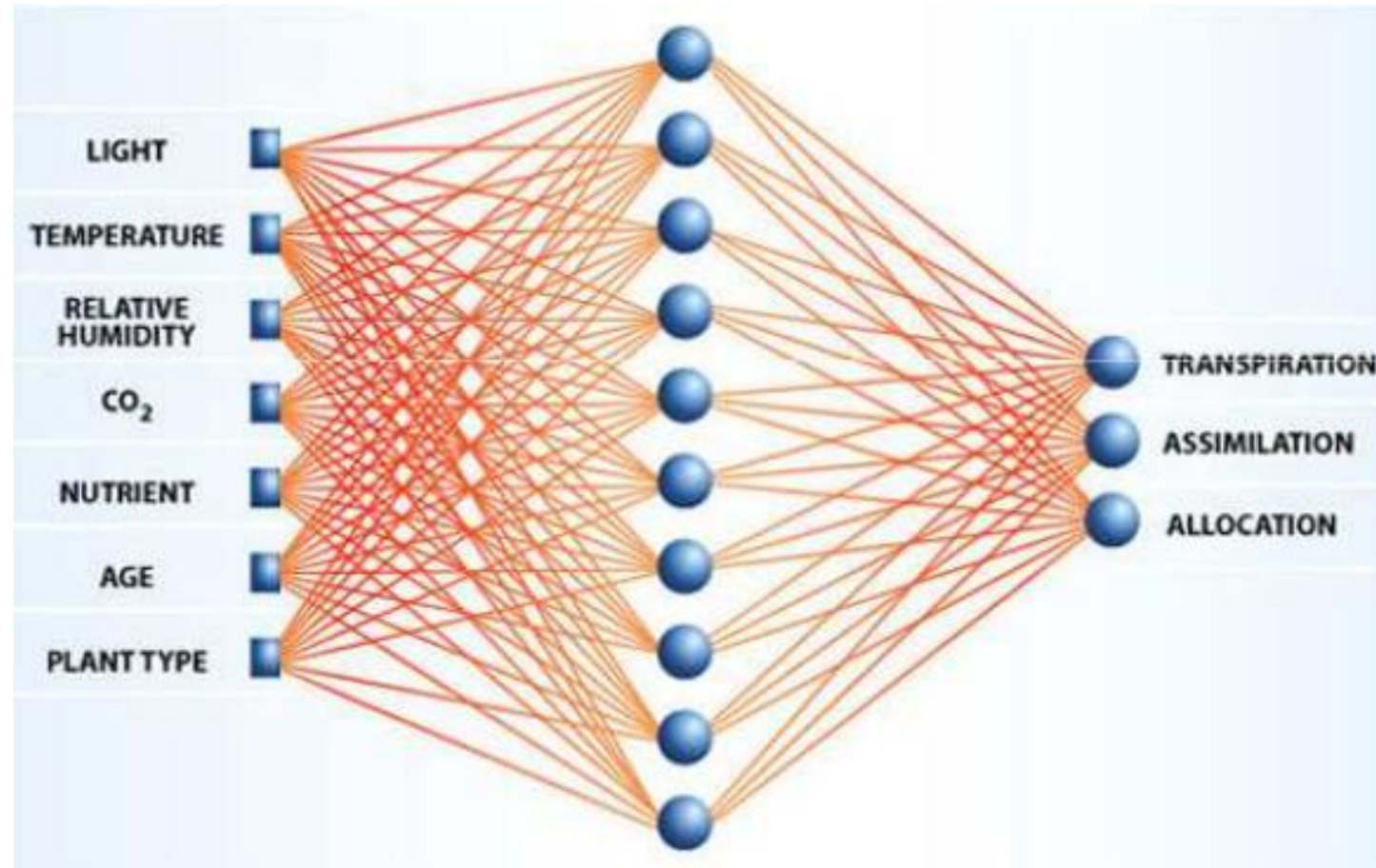
- Las señales que llegan a la sinapsis son las entradas a la neurona
- Estas señales de entrada son ponderadas (atenuadas o amplificadas) mediante un parámetro, denominado peso, asociado a la sinapsis correspondiente

- Las señales de entrada pueden
- Excitar a la neurona (sinapsis con peso positivo)
- Inhibir a la neurona (sinapsis con peso negativo)
- El efecto es la suma ponderada de las entradas
- Si la suma es igual o mayor que el umbral de la neurona, la neurona se activa

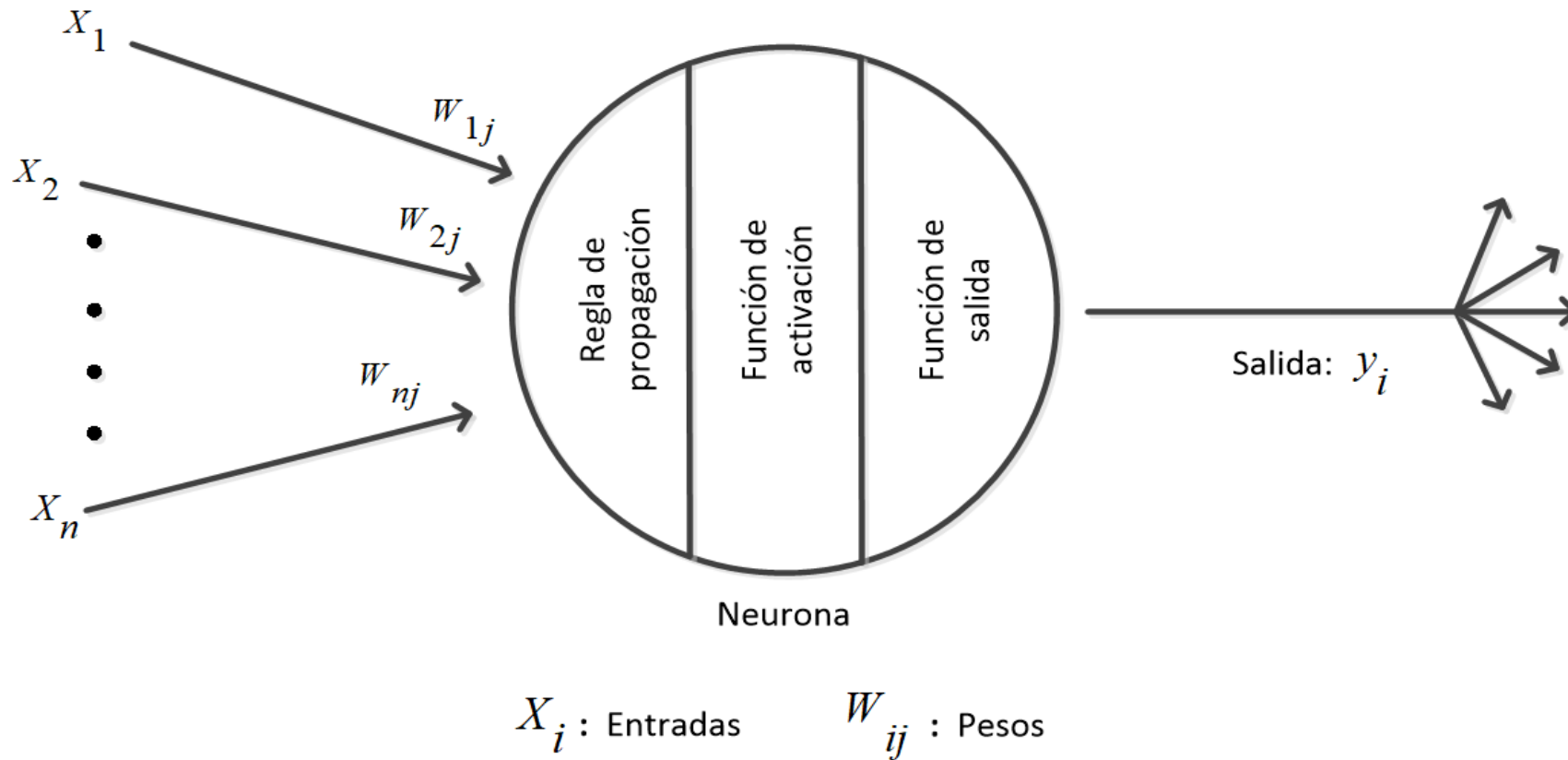
Elementos de una estructura de RNA

- Conjunto de procesadores elementales o neuronas artificiales
- Patrón de conectividad o arquitectura
- Dinámica de activaciones
- Una regla o dinámica de aprendizaje
- El entorno en que opera

- Tipos de neuronas
 - De entrada: Reciben señales del entorno
 - Ocultas: No tienen contacto con el entorno
 - De salida: Envían señales al entorno
- Una capa o nivel es un conjunto de neuronas cuyas entradas proceden de la misma fuente y cuyas salidas se dirigen al mismo destino



Elementos de una neurona artificial



Regla de propagación

- permite obtener el valor del potencial postsináptico h_i de una neurona, a partir de las entradas y los pesos

$$h_i(t) = \sigma_i \left(w_{ij}, x_j(t) \right)$$

La función más habitual es de tipo lineal, y consiste en la suma ponderada de las entradas con los pesos sinápticos, que formalmente se puede interpretar como un producto escalar de los vectores de entradas y pesos

Función lineal

$$h_i(t) = \sum_j w_{ij} x_j = W_j^T \cdot X$$

Distancia Euclidea (usada en mapas de Kohonen)

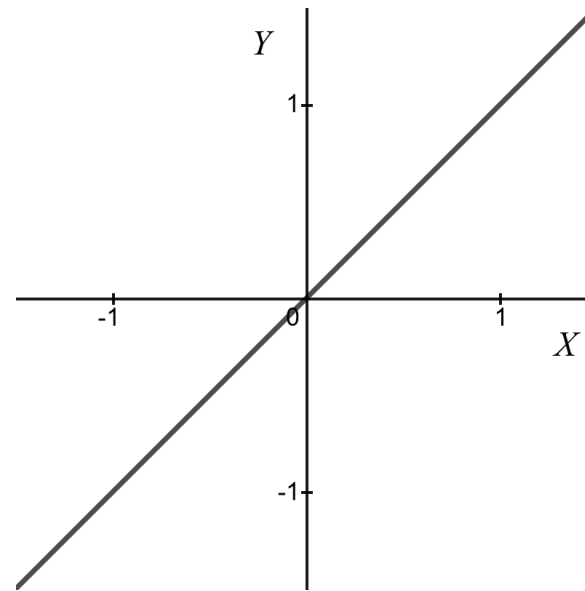
$$h_i(t) = \sqrt{\sum_j (x_j - w_{ij})^2}$$

Función de activación o transferencia

- proporciona el estado de activación actual
- La función de activación $f(.)$ se considera determinista, y en la mayor parte de los modelos es monótona creciente y continua, emulando las neuronas biológicas.

Lineal o Identidad

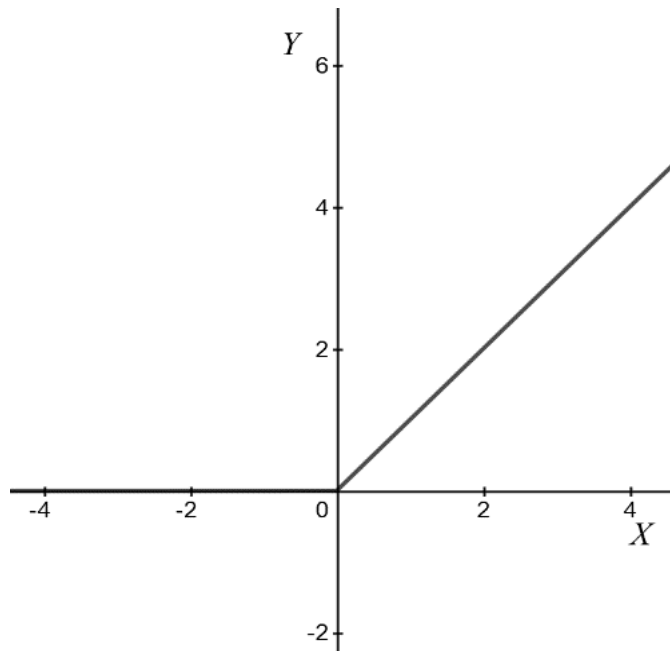
- La salida de esta función está dada por su entrada, lo que equivale a decir que no se aplica ninguna función y es poco utilizada. Su rango va de $[-\infty, \infty]$



$$f(x) = x$$

Rectified Linear Unit o ReLU

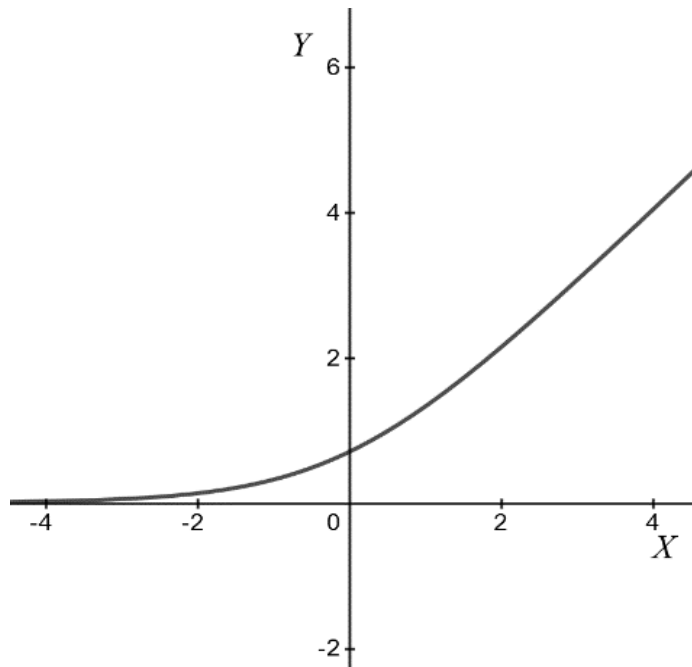
- consiste en conservar el valor de entrada siempre y cuando sea mayor o igual a cero. En otro caso suplanta el valor con cero



$$f(x) = \max(0, x)$$

Softplus

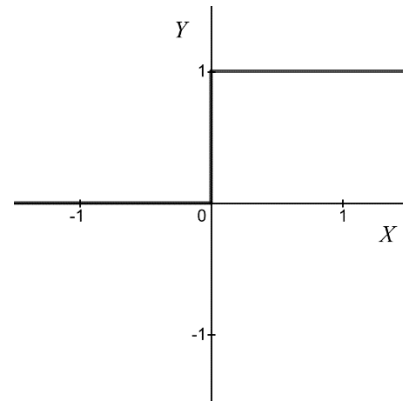
- Esta aproximación tolera valores negativos hasta cierto extremo



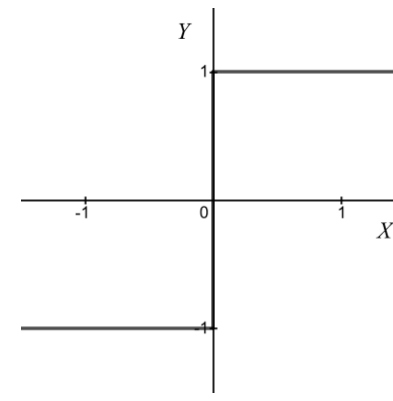
$$f(x) = \ln(1 + e^x)$$

Umbral o Escalonada

- Es utilizada únicamente cuando las salidas de la red son binarias



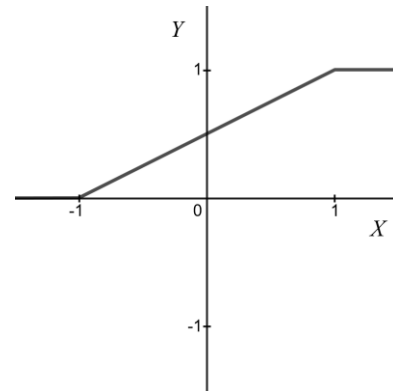
$$f(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$



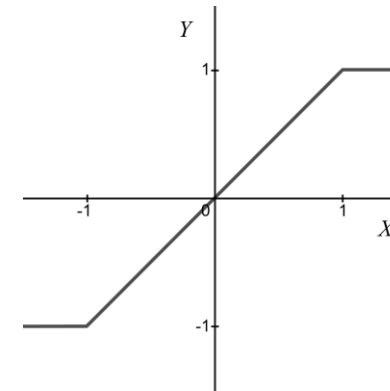
$$f(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{si } x < 0 \end{cases}$$

Lineal y Mixta (Lineal a Tramos)

- Esta función está obligada a permanecer en un rango de valores reales predeterminados y es apropiada cuando se requiere información analógica como salida



$$f(x) = \begin{cases} 0 & \text{si } x < -c \\ 1 & \text{si } x > c \\ \frac{x}{2c} + \frac{1}{2} & \text{en otro caso} \end{cases}$$

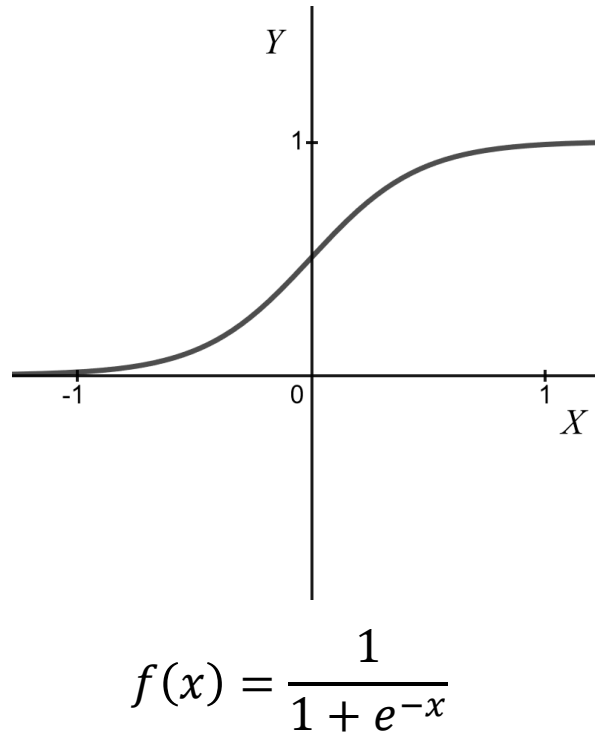


$$f(x) = \begin{cases} -1 & \text{si } x < -c \\ 1 & \text{si } x > c \\ a.x & \text{en otro caso} \end{cases}$$

Donde c es el límite superior de la suma de todas las entradas de activación y -c el límite inferior

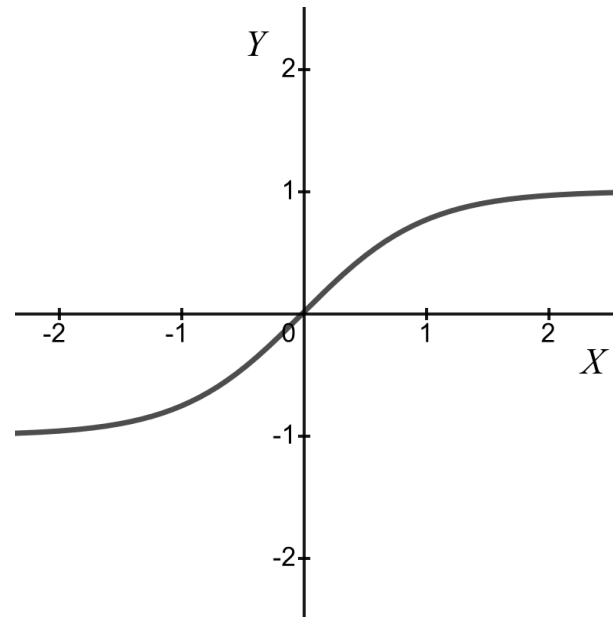
Sigmoide (logaritmo sigmoidal)

- Función más popular en las RNA debido a que es derivable



Tangente hiperbólica

- Esta función es semejante a la sigmoidea e igualmente popular



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

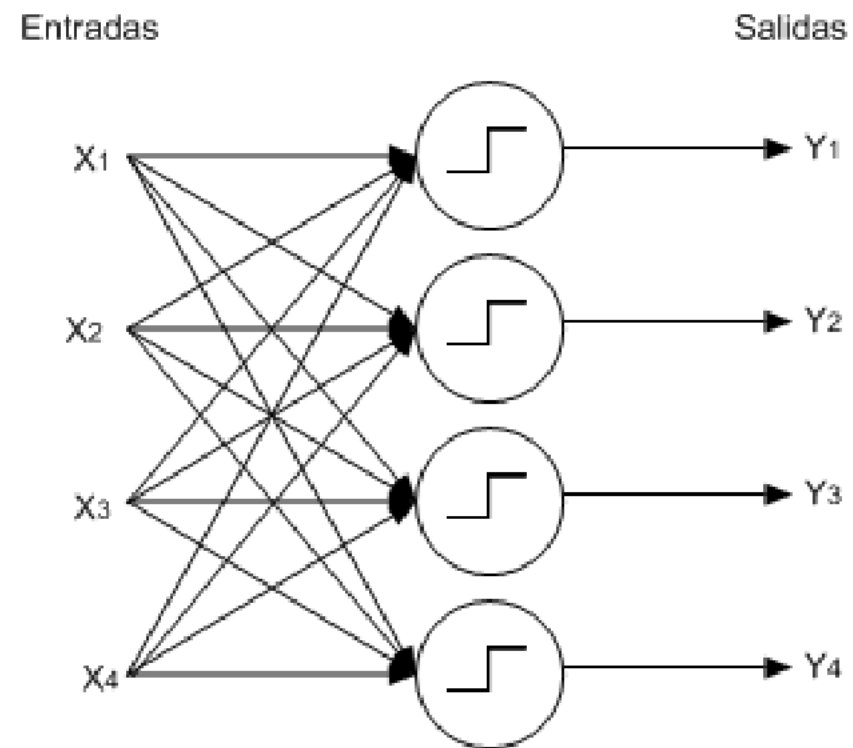
Función de salida

- Esta función proporciona la salida global de la neurona $y_i(t)$ en función de su estado de activación actual $a_i(t)$
- Usualmente en los modelos neuronales la función de salida es simplemente la identidad $f(x)=x$, de tal forma que el estado de activación de la neurona se considera como la propia salida

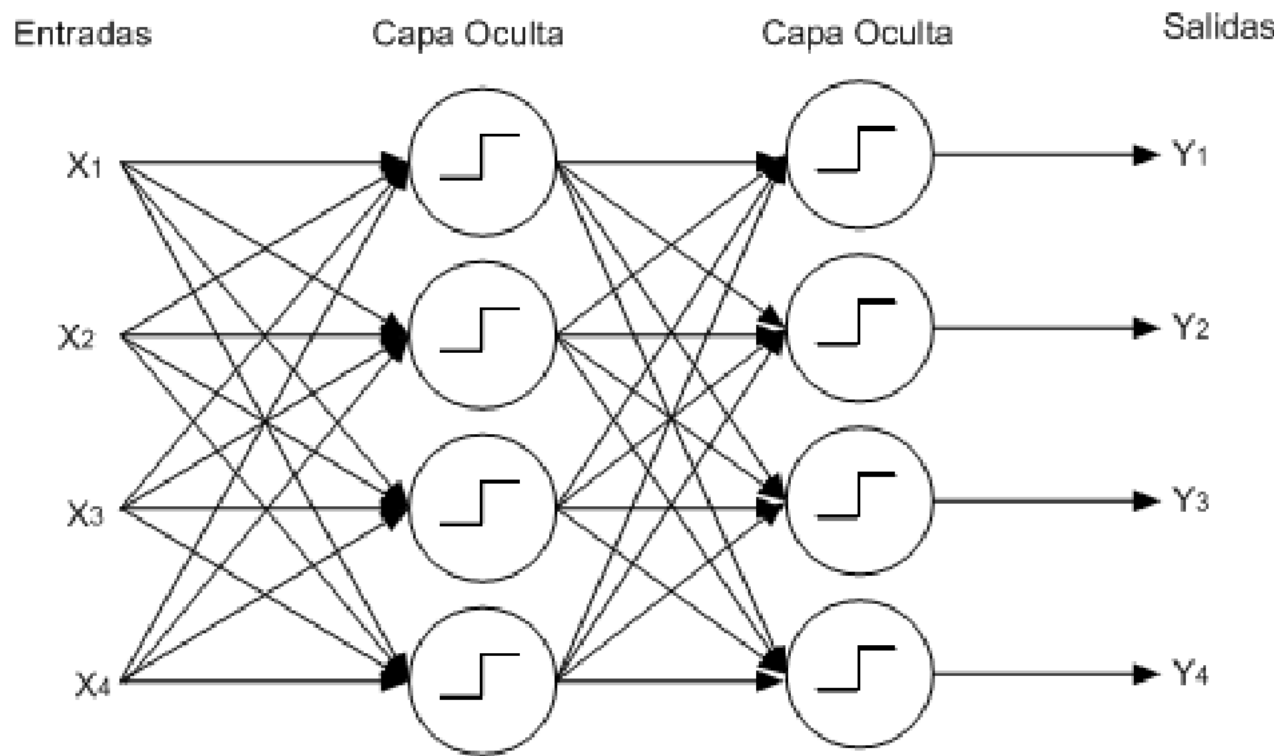
$$y_i(t) = F_i(a_i(t)) = a_i(t)$$

Esto ocurre en los modelos de RNA más comunes, tal como el perceptrón multicapa (MLP) o la adalina

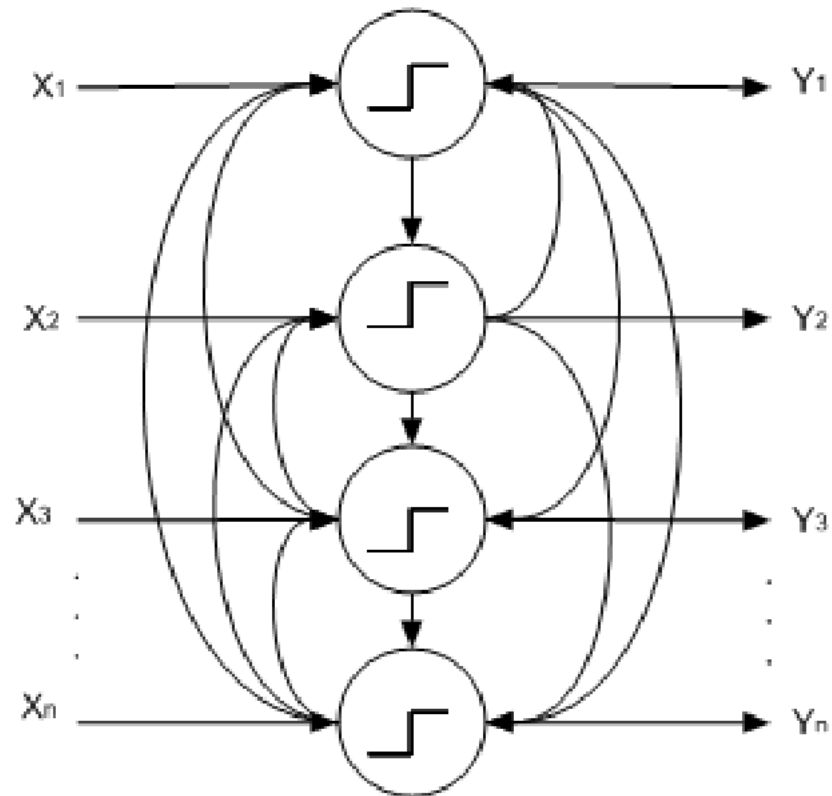
Redes monocapa



Redes multicapa



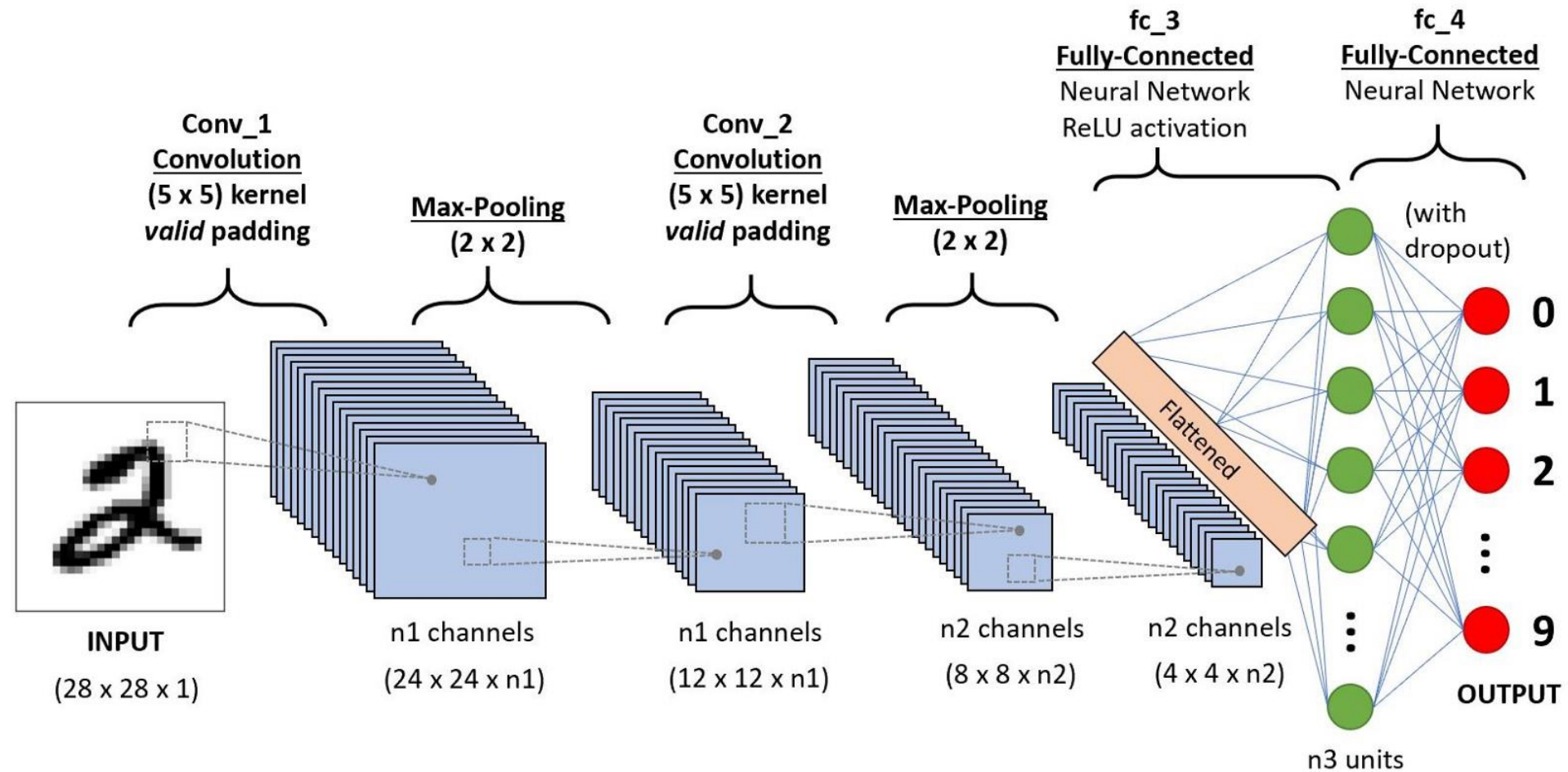
Redes recurrentes



Clasificación general

| Número de capas | Tipo de conexiones | Dirección | Redes neuronales |
|-----------------|---------------------------------|---------------------------------|--|
| Redes Monocapa | Lineales | Hacia adelante (Feedforward) | <ul style="list-style-type: none">• Perceptrón |
| | | Autorecurrentes | <ul style="list-style-type: none">• Máquina de Boltzman• Brain-State-In-A-Box |
| | Recurrentes | No Autorrecurrentes | <ul style="list-style-type: none">• Red de Hopfield |
| Redes Multicapa | Lineales (Mixtas según aplique) | Hacia adelante (Feedforward) | <ul style="list-style-type: none">• Perceptrón multicapa• ADALINE – MADALINE |
| | | Hacia atrás (Feedback) | <ul style="list-style-type: none">• Kohonen SOM (Self-Organizing Maps) |
| | | Ambos sentidos | <ul style="list-style-type: none">• ART (Adaptive Resonance Theory) |

CNN



Convolución

