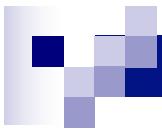


INTELIGENCIA ARTIFICIAL APLICADA AL CONTROL

Tema 5: Redes Neuronales. Introducción

Dpto.: Arquitectura de Computadores y Automática
Autor: Matilde Santos



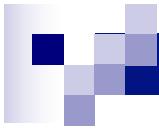
REDES NEURONALES

5.1. INTRODUCCIÓN

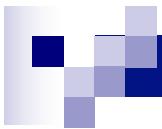
5.2. REDES ANALÓGICAS

5.3. APRENDIZAJE

5.4. NEURO CONTROL

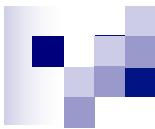


5.1 INTRODUCCIÓN A LAS REDES NEURONALES



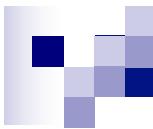
MOTIVACIÓN

- Imitar la estructura del cerebro
 - Neuronas
 - Conexiones
 - Transmisión impulsos eléctricos
- Inteligencia Artificial conexionista



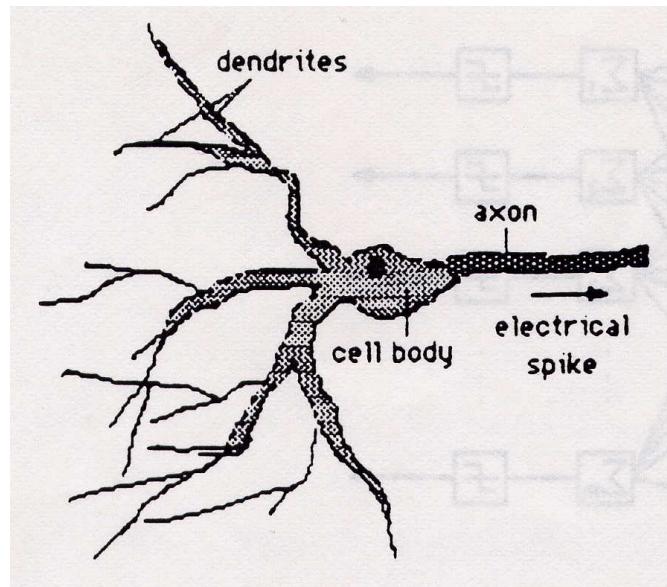
REDES DE NEURONAS ARTIFICIALES (RNA O ANN)

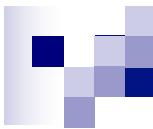
- Sistema de procesamiento de información
- Paradigma de aprendizaje



ORIGEN

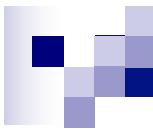
- 1950s: Minski
- Analogía biológica
- Diversidad de interpretaciones





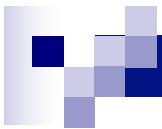
HISTORIA

- **1936**, Turing. Cerebro como modelo de computación
- **1943**, McCulloch y Pitts, modelado de una red neuronal simple mediante circuitos eléctricos
- **1957**, Rosenblatt, Perceptrón.
- **1969**, Widrow y Hoff, Adaline (Adaptive linear elements)
 - Primera aplicación a un problema real (filtro adaptativo para eliminar ecos en las líneas telefónicas)
- **1986**, desarrollo de las redes neuronales
 - Redes de segunda generación,



REDES NEURONALES

- Conjunto de autómatas conectados entre sí que generan propiedades globales emergentes
 - Sistema de interconexión de neuronas en una red que colaboran para producir un estímulo de salida
- Su funcionamiento depende:
 - El procesamiento realizado por cada unidad
 - Su conectividad (arquitectura o topología)
 - Las reglas que gobiernan la dinámica del envío de mensajes entre unidades



PROCESAMIENTO

■ LÓGICO

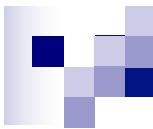
- Procesador local: función lógica

■ ANALÓGICO

- Computación local: suma + umbral

■ INFERENCIAL

- Inferencia local: reglas



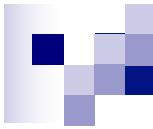
ELEMENTOS

- Elementos de procesamiento (neuronas)
- Conexiones (pesos)
 - Conocimiento simbólico
- Función de propagación o de activación
 - Procesamiento local (sumatorio de la entradas ponderadas por el peso de su conexión)
- Función de transferencia o de salida
 - Interpretación del resultado de la activación



BENEFICIOS

- Elevada tolerancia a fallos
- Auto-organización
- Información distribuida
 - Cálculos intensivos
 - Tiempo real
 - Información difusa, con ruido o inconsistencias
- Aprendizaje adaptativo
 - Procesar cualquier función o tarea
- Flexibilidad
- Fácil integración con otras técnicas



APLICACIONES

RECONOCIMIENTO
DE PATRONES

IDENTIFICACIÓN

MONITORIZACIÓN

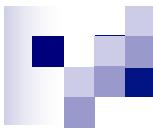
PROCESAMIENTO
DE SEÑALES

PREDICCIÓN

CONTROL DE
PROCESOS

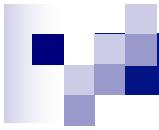
ROBÓTICA

OPTIMIZACIÓN

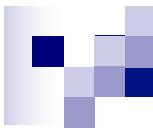


ÁMBITOS DE APLICACIÓN

- Diagnóstico:
 - De fallos (plantas industriales, nucleares, ..)
 - Médico: enfermedades, concentraciones
- Control, Identificación
- Clasificación
 - Señales de radar, sonar, detección de explosivos
 - Detección de fraudes, firmas, caracteres
- Predicción
 - Finanzas: tendencias, riesgos, evolución
 - Series temporales, demanda, el tiempo, trayectorias
- Optimización
 - Recursos



5.2 REDES NEURONALES ANALÓGICAS



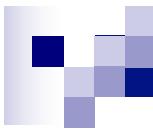
NEURONA BIOLÓGICA

- Los estímulos recibidos por el cerebro (sentidos) son transmitidos por las neuronas mediante conexiones sinápticas
- Cuando una neurona es estimulada libera una pequeña cantidad de un componente químico (neurotransmisor)
- Este viaja a través del axón para llegar a las dendritas de otras neuronas, y el proceso se repite
 - Se incrementa o disminuye la relación de las neuronas involucradas
 - Un determinado estímulo puede activar o inhibir a una neurona

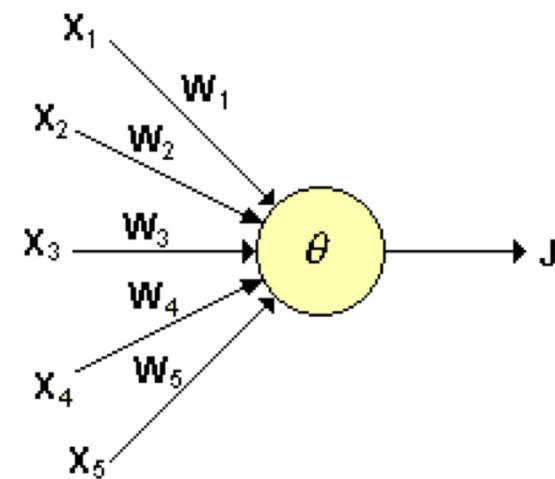
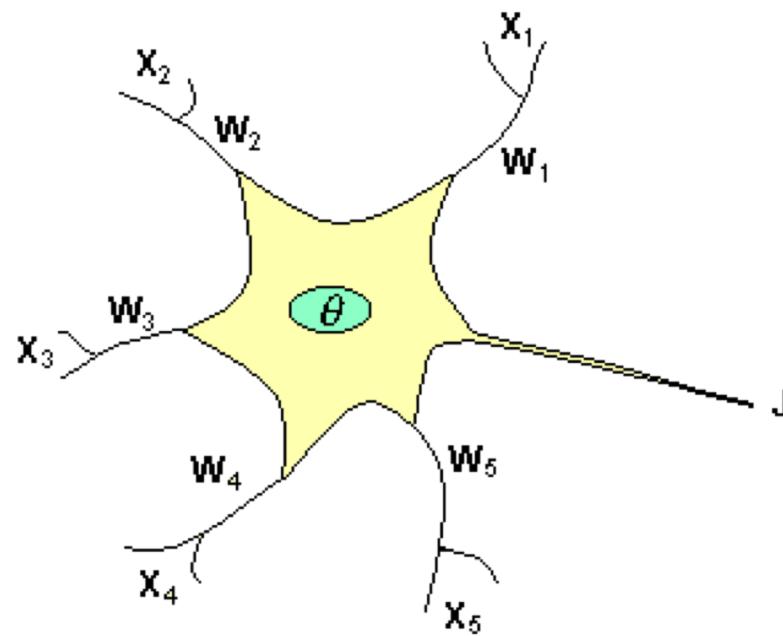
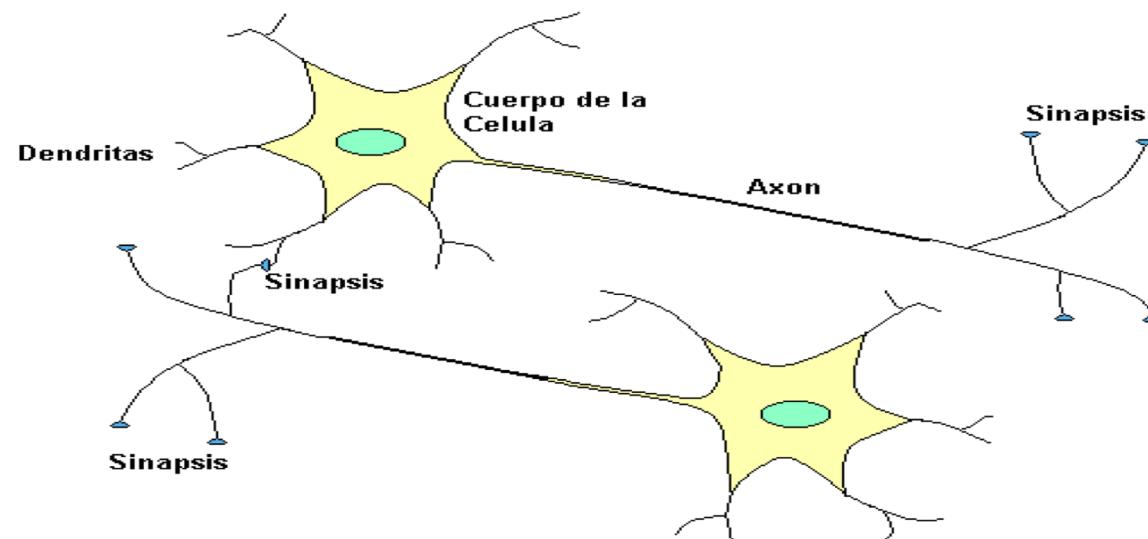
Neurona: 10^{-6} m

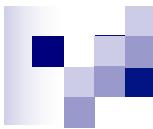
Nº total: 10^{10} - 10^{12}

Nº conex x neurona 10^0 - 10^4



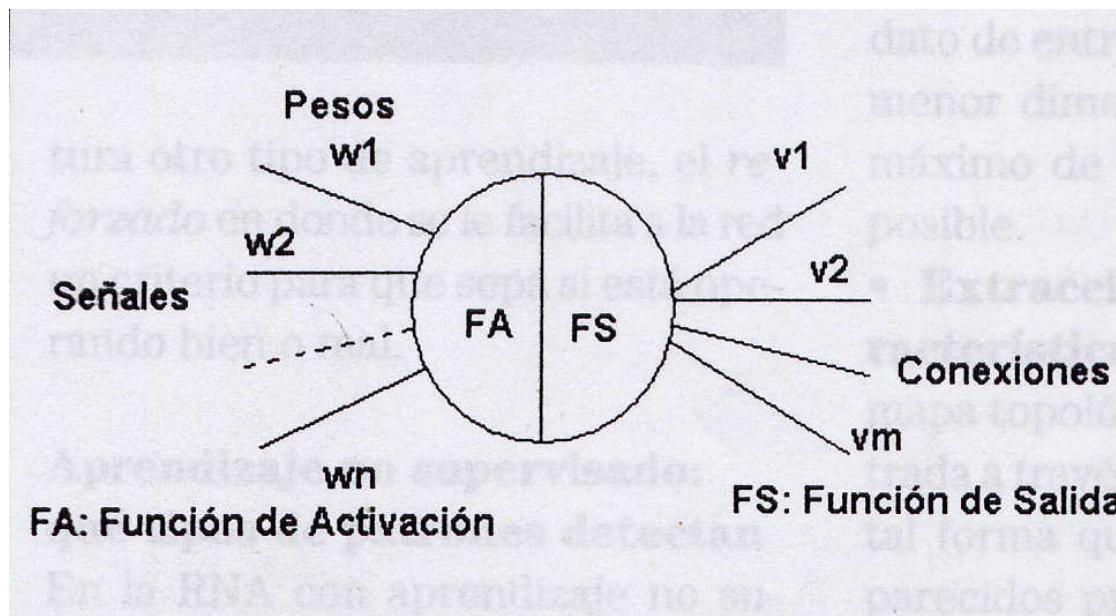
NEURONA ANALÓGICA

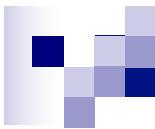




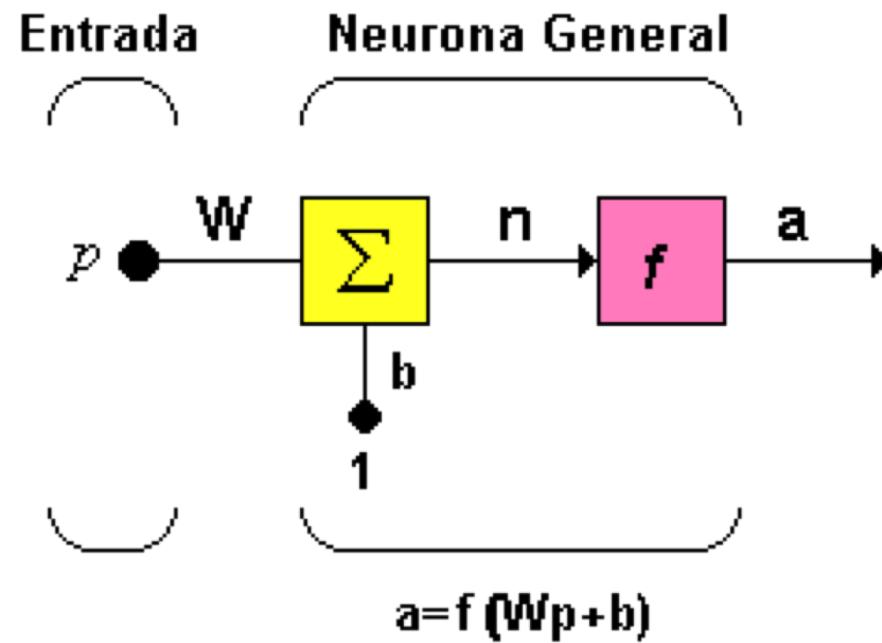
NEURONA ARTIFICIAL

Autómata sencillo que recibe entradas de otras unidades vecinas y calcula un valor de salida



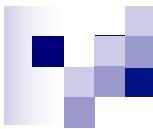


FUNCIONAMIENTO (1)

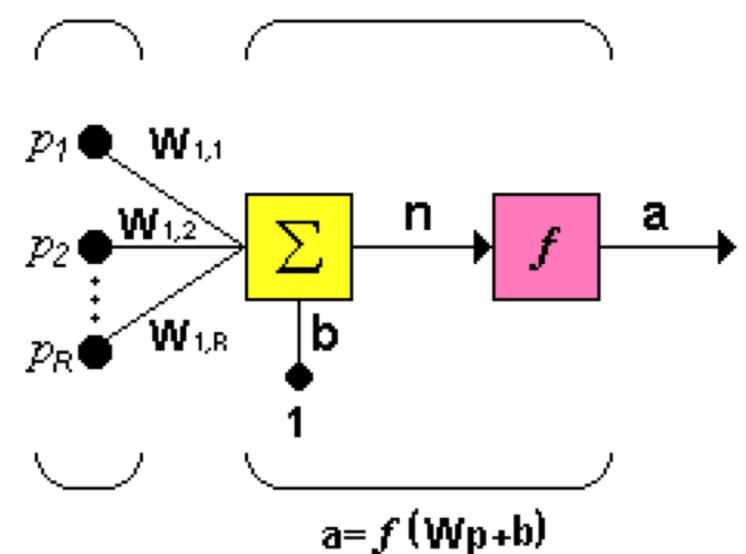
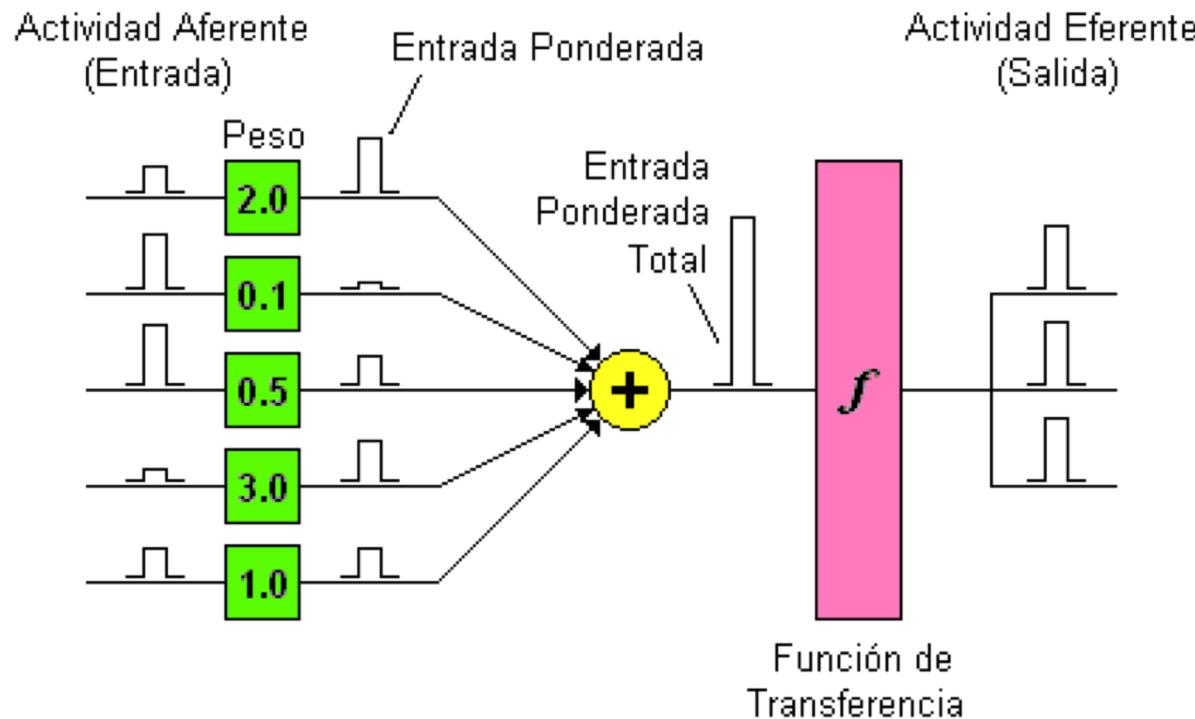


$$net_{a_i} = \sum_{i=1}^n W_i X_i = \vec{X} \vec{W}$$

$$a = f(Wp + b)$$



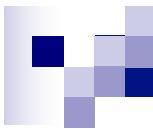
FUNCIONAMIENTO (2)



$$n = w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b$$

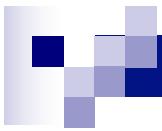
$$n = \mathbf{Wp} + b$$

$$a = f(\mathbf{Wp} + b)$$



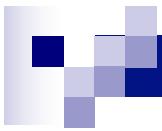
FUNCIONES DE ACTIVACIÓN

- Continuas o discretas
- Diferenciables
- Rango de las entradas (normalización)
- Estado de activación
 - Reposo (no activo) o excitación (activo)
- Salidas (normalizadas)
 - Discretas:
 - 1 -> estado activo
 - 0 -> estado pasivo, neurona en reposo
 - Continuas: $[0,1]$ ó $[-1,1]$



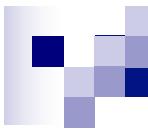
FUNCIÓN DE SALIDA

- Obtiene un valor de salida a partir del estado de activación
- Suele ser igual que la función de activación



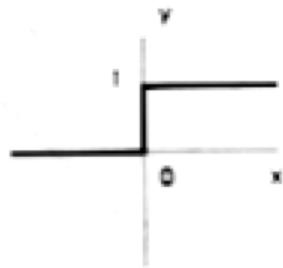
FUNCIONES DE ACTIVACIÓN/SALIDA

- ✓ Función escalón o umbral
- ✓ Función lineal
- ✓ Saturación (mixta)
- ✓ Función sigmoide
- ✓ Tangente hiperbólica
- ✓ Tangente gaussiana, ...

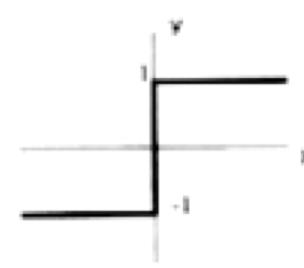


FUNCIONES DE SALIDA

Escalón o umbral

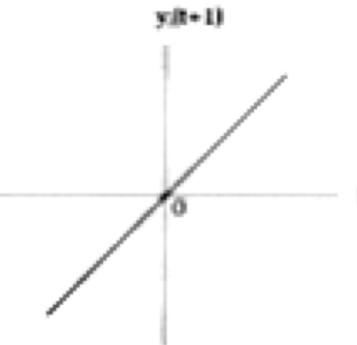


$$f(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$



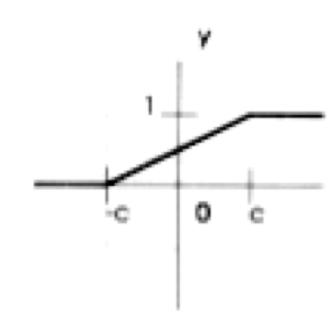
$$f(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{si } x < 0 \end{cases}$$

y(t+1)

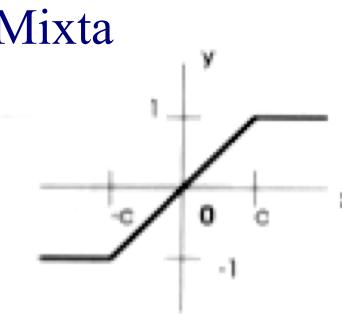


Lineal

Mixta

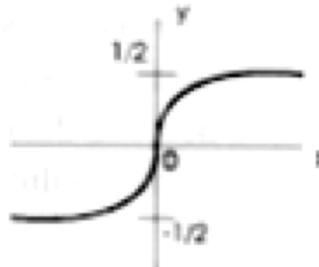
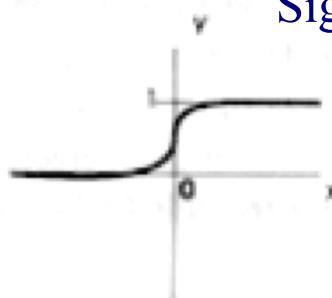


$$f(x) = \begin{cases} 0 & \text{si } x < -c \\ 1 & \text{si } x > c \\ x / (2c) + 1/2 & \text{en otro caso} \end{cases}$$



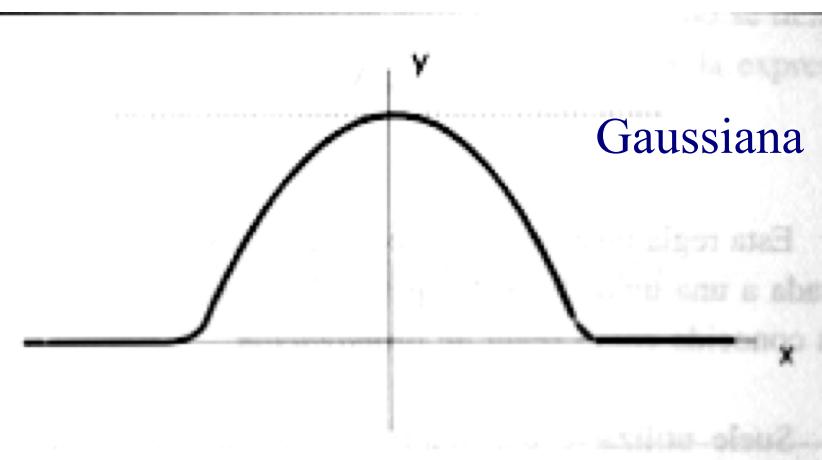
$$f(x) = \begin{cases} -1 & \text{si } x < -c \\ 1 & \text{si } x > c \\ ax & \text{en otro caso} \end{cases}$$

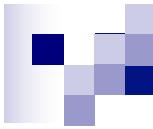
Sigmoide



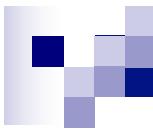
$$f(x) = 1 / (1 + e^{-x})$$

Gaussiana



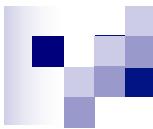


Nombre	Relación Entrada /Salida	Icono	Función
Limitador Fuerte	$a = 0 \text{ } n < 0$ $a = 1 \text{ } n \geq 0$		hardlim
Limitador Simétrico Fuerte	$a = -1 \text{ } n < 0$ $a = +1 \text{ } n \geq 0$		hardlims
Lineal Positiva	$a = 0 \text{ } n < 0$ $a = n \text{ } 0 \leq n$		poslin
Lineal	$a = n$		purelin
Lineal Saturado	$a = 0 \text{ } n < 0$ $a = n \text{ } 0 \leq n \leq 1$ $a = 1 \text{ } n > 1$		satlin
Lineal Saturado Simétrico	$a = -1 \text{ } n < -1$ $a = n \text{ } -1 \leq n \leq 1$ $a = +1 \text{ } n > 1$		satlins
Sigmoidal Logarítmico	$a = \frac{1}{1 + e^{-n}}$		logsig
Tangente Hiperbólica Sigmoidal	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
Competitiva	$a = 1 \text{ } \text{Neurona con } n \text{ max}$ $a = 0 \text{ } \text{El resto de neuronas}$		compet



REDES ANALÓGICAS

- No recurrentes
 - Monocapa (Perceptrón/Adaline)
 - Perceptrón multicapa
- Recurrentes
 - Redes de Hopfield (Memorias asociativas)
 - Máquinas de Boltzmann
 - Redes de Elman
- Otros modelos de redes



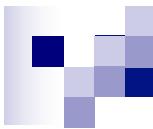
REDES ANALÓGICAS

■ Monocapa

- Conexiones laterales entre neuronas de la misma capa
 - *Red de Hopfield; Red Brain-State-In-A-Box, etc*

■ Multicapa

- Neuronas agrupadas por niveles
 - Conexiones hacia delante (*feedforward*)
 - *Perceptron, Adaline, Madaline, ...*
 - Conexiones hacia atrás (*feedback*), bicapa
 - *ART, Neocognitron, ...*

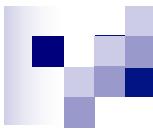


REDES NO RECURRENTES

REDES DE FLUJO DIRECTO

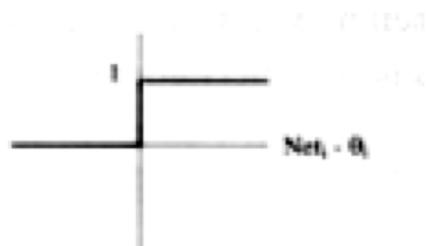
ENTRADA  SALIDA

- Sinapsis: conexión entre la neurona m la neurona n (peso w_{mn})
- Peso:
 - positivo: excitadora,
 - negativo: inhibidora
 - 0: no existe conexión
- Matriz de pesos W



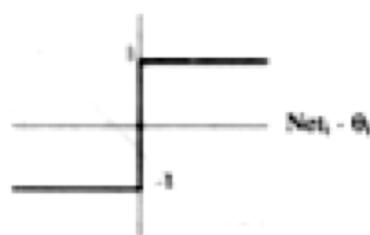
REDES DE PRIMERA GENERACIÓN

REDES FEEDFORWARD: Conexiones hacia delante



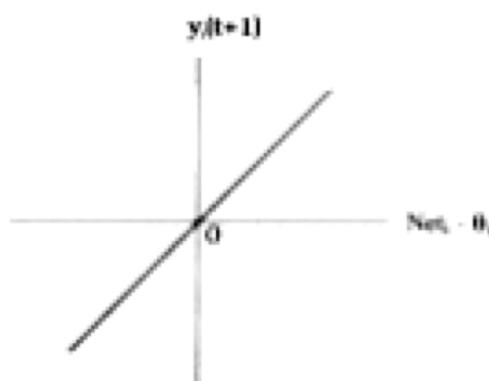
PERCEPTRON

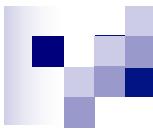
(Rosenblat, 1957)



ADALINE

(Widrow y Hoff, 1960)





PERCEPTRÓN

■ ELEMENTOS:

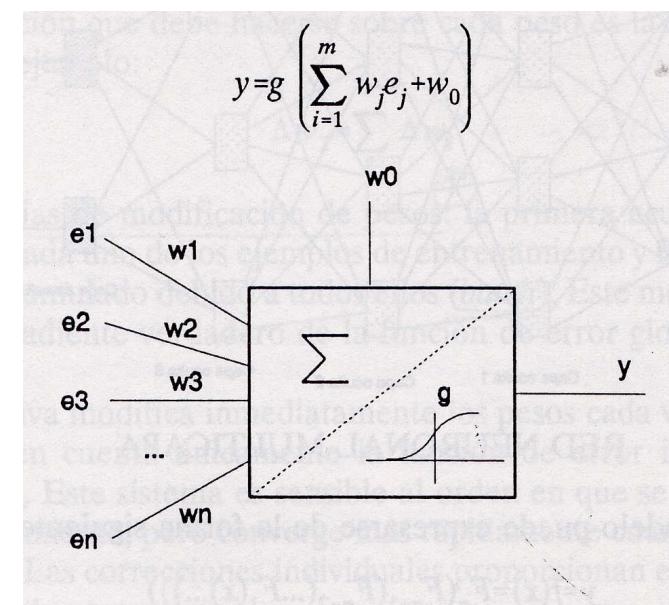
- Varias neuronas de entrada y una de salida

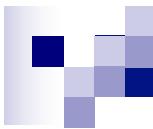
■ CONECTIVIDAD: pesos

■ PROCESAMIENTO: suma + umbral

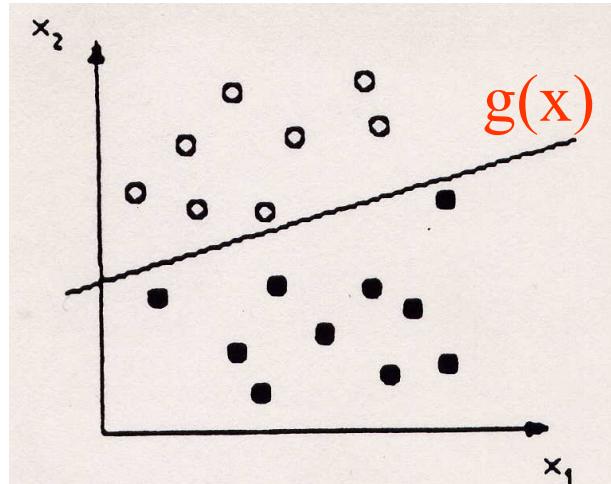
- Función de activación
- Función de salida

*Mecanismo de aprendizaje:
función criterio del perceptrón*





EJEMPLO: CLASIFICACIÓN



Entradas: x_1, x_2

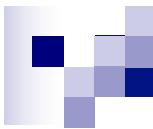
Clase 1 \Rightarrow salida 1

Clase 2 \Rightarrow salida 0

Función de activación $g(x) = \sum_{i=0}^n w_i x_i = w_1 x_1 + w_2 x_2 + w_0$

Función de salida $o(x) = \begin{cases} 1, & g(x) > 0 \\ 0, & g(x) < 0 \end{cases}$

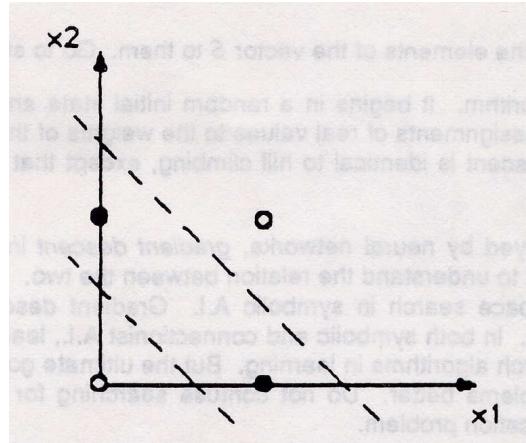
$g(x) = 0 \Rightarrow$ SUPERFICIE DE DECISIÓN



TEOREMA DE CONVERGENCIA DEL PERCEPTRÓN

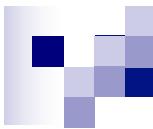
- Puede encontrar una solución siempre que ésta exista (reconocimiento de patrones lineales)
 - Superficie de Decisión

EJEMPLO: XOR



X1	X2	XOR	
0	0	0	Clase 2
0	1	1	Clase 1
1	0	1	Clase 1
1	1	0	Clase 2

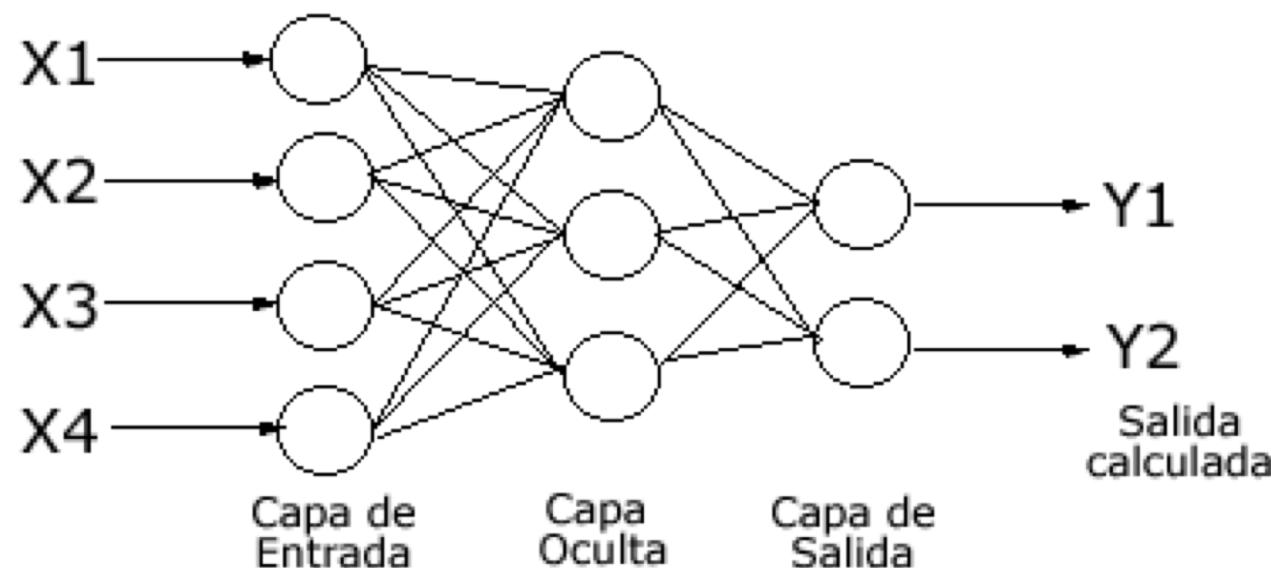
PROBLEMAS LINEALMENTE SEPARABLES

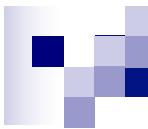


REDES DE SEGUNDA GENERACIÓN

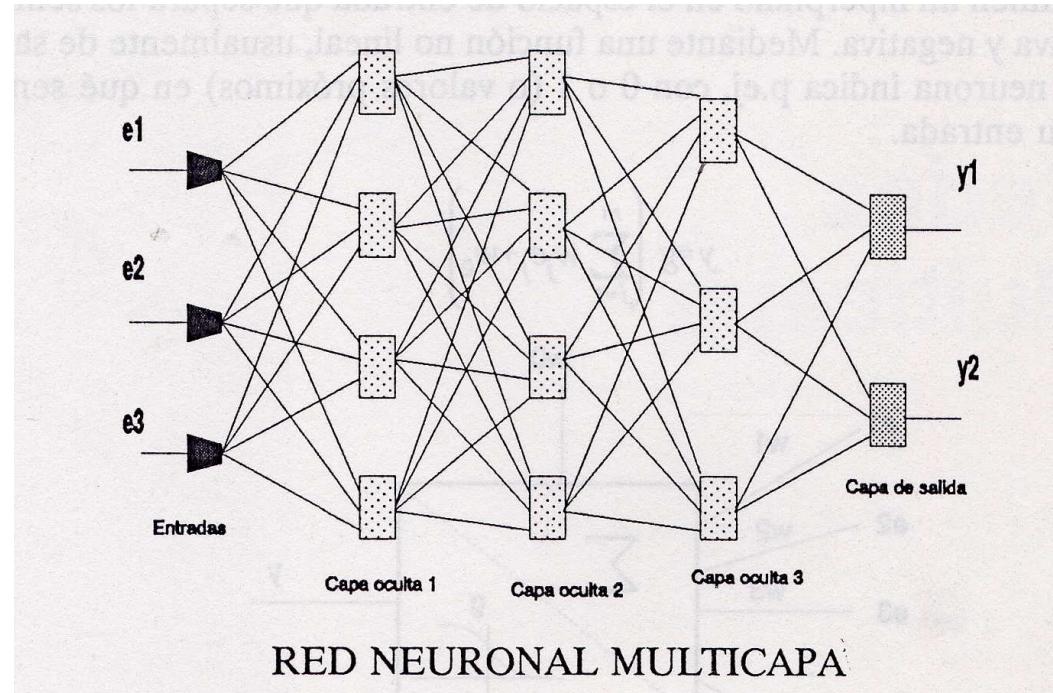
REDES MULTICAPA

Red de flujo directo con capas de unidades entre la entrada y la salida (capas ocultas)



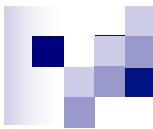


PERCEPTRÓN MULTICAPA



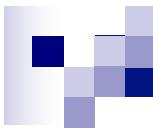
NO ASEGURADA LA CONVERGENCIA

*Mecanismo de aprendizaje:
Retropropagación (backpropagation)*



PERCEPTRÓN MULTICAPA

- Función de activación derivable
 - Aprendizaje: algoritmo de retropropagación
 - Conexiones: todas las de un nivel con todas las del siguiente
 - Valores reales
 - Con una sola capa oculta y un número arbitrario de neuronas puede representar cualquier función
- LIMITACIÓN: **no** asegurada la convergencia



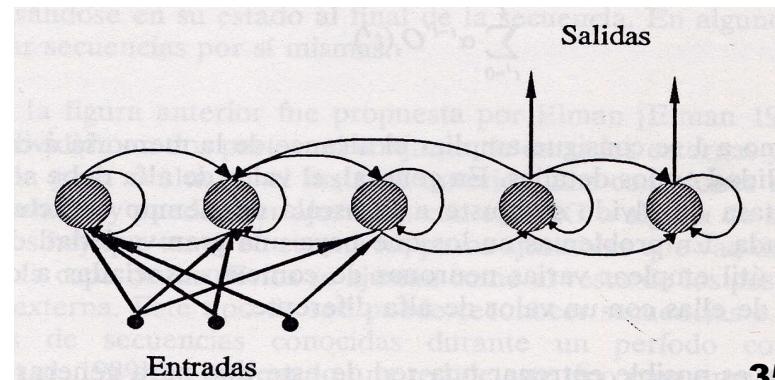
REDES RECURRENTES

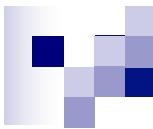
■ MONOCAPA

- Se utilizan típicamente para autoasociación
- Regenerar información de entrada incompleta o distorsionada

■ Hopfield, BSB, Learning Matrix

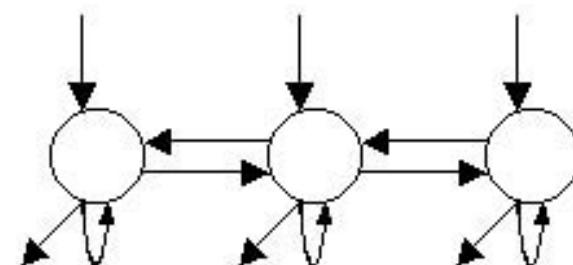
- Conexiones hacia delante (*feedforward*)
- Conexiones hacia atrás (*feedback*)
- Conexiones laterales



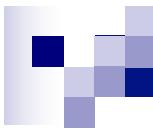


REDES RECURRENTES

- MONOCAPA
- CONECTIVIDAD (necesitan menos conexiones)
- DINÁMICAS (estados de activación varían con el tiempo)
- REALIMENTACIÓN
- ENTRADA: estado inicial de algunas unidades
- SALIDA: valor estático
- APRENDIZAJE: más lento

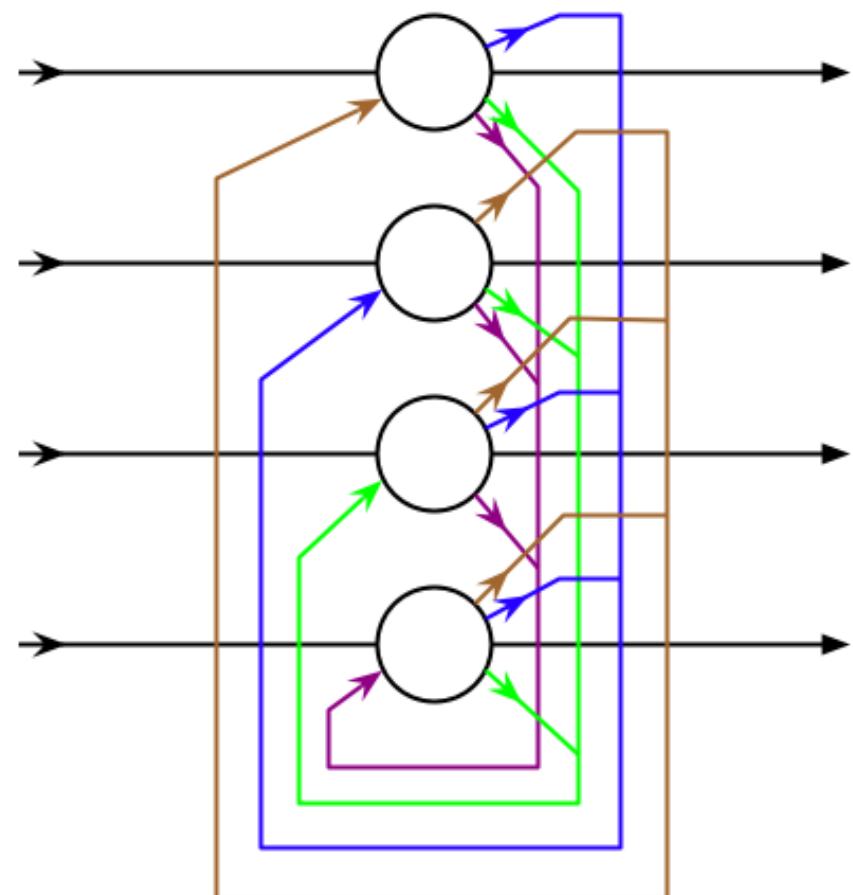


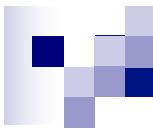
Red Recurrente Monocapa



REDES DE HOPFIELD

- Red neuronal artificial recurrente (John Hopfield)
- Se inspiran en la memorización de configuraciones en cristales de spin
- Se usan como sistemas de memoria asociativa con unidades binarias
- Diseñadas para converger a un mínimo local (estados estables)

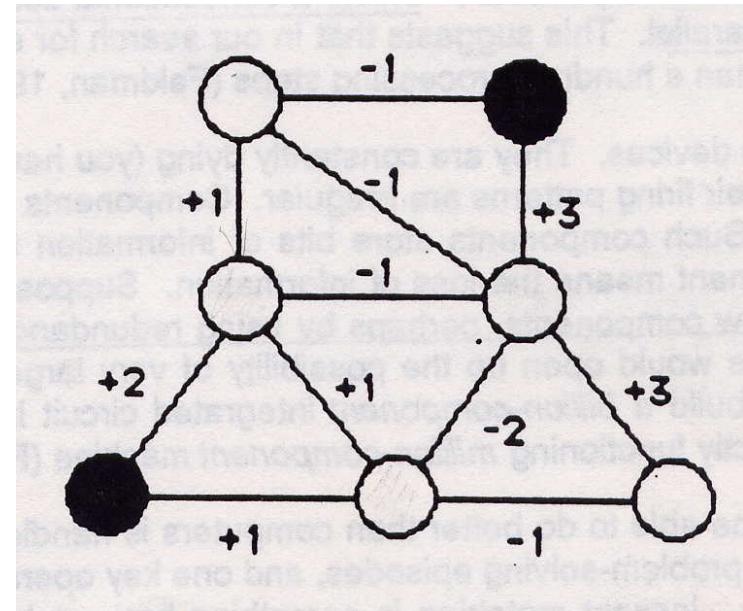


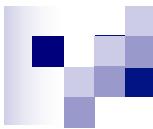


REDES DE HOPFIELD

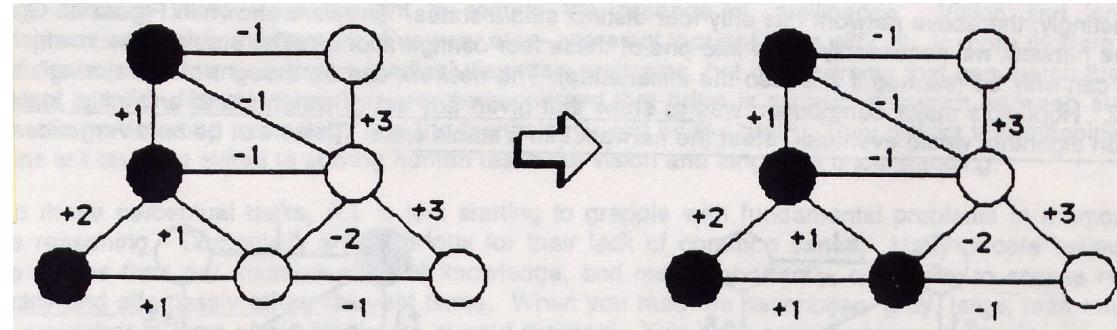
- Redes monocapa con estados binarios (0,1 ó -1,1)
- El valor se determina según las unidades superen o no un determinado umbral

- ◆ Activación de la unidad i (s_i):
 - $S_i = 1$
 - $S_i = 0$
- ◆ Conexiones: pesos W_{ij}
- ◆ Funcionamiento: relajación paralela
- ◆ Convergencia: mínimos de la función energía

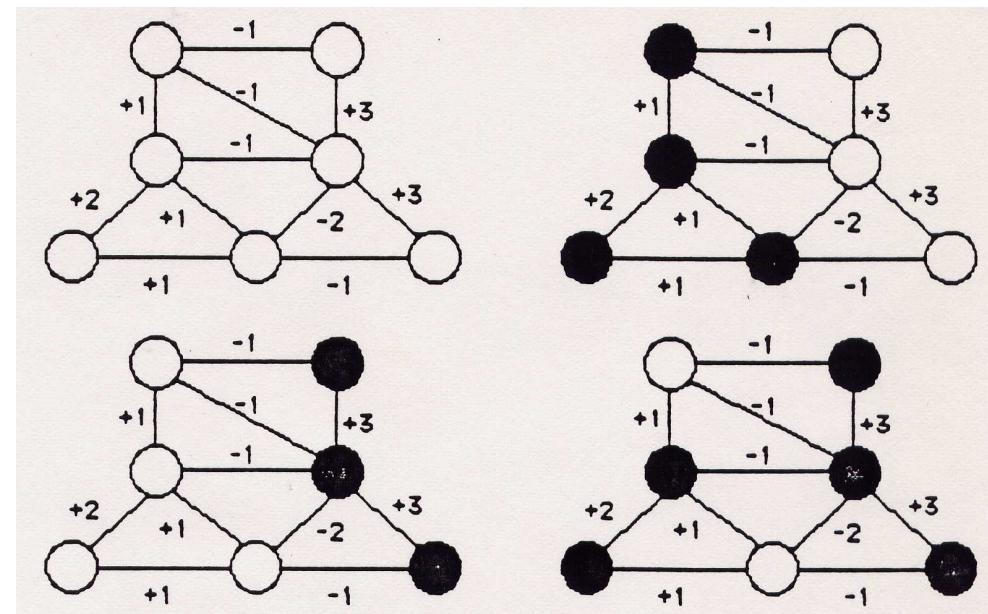


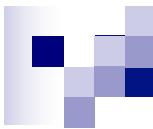


MECANISMO : RELAJACIÓN PARALELA



ATRACTORES O ESTADOS ESTABLES





ESTADOS ESTABLES

Ecuación de activación de las n neuronas

$$\frac{\partial y_i}{\partial t} = -Ay_i + \sum_{k=1}^n w_{ik}f(y_k) + aI_i, \quad i=1,2,\dots,n$$

entradas externas

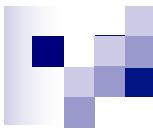
Valores de equilibrio $\frac{\partial y_i}{\partial t} = 0$

FUNCIÓN DE ENERGÍA

$$E = -\frac{1}{2} \left(\sum_i^n \sum_k^n w_{ik} y_i y_k + A \sum_i^n y_i y_i \right)$$

Mínimo locales:
⇒ estados estables

NO SOLUCIÓN GLOBAL ÓPTIMA



MÁQUINA ESTOCÁSTICA DE BOLTZMAN

REDES DE HOPFIELD
+
SIMULATED ANNEALING

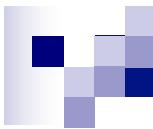
} Red neuronal estocástica

Ecuación de activación estocástica $p_k = \frac{1}{1 + e^{\Delta E_k / T}}$

p_k : probabilidad de activación de la neurona k

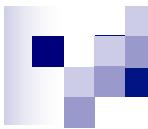
T = TEMPERATURA

- $T \uparrow$: comportamiento aleatorio
- $T \downarrow$: comportamiento red Hopfield



BRAIN STATE IN A BOX (BSB)

- J. Anderson (1977), Red de Hopfield con conexiones con valores reales
- Red no-lineal autoasociativa
- Aprendizaje auto organizado Hebbiano
 - Red “fully connected”
 - Se basa en la dinámica de atractores (minimiza función de energía de Lyapunov) hasta alcanzar un equilibrio
 - Los estados estables de la red son representados por las aristas de un hipercubo n dimensional
 - El estado de salida de cada neurona es un valor real continuo entre +1 y -1
 - Todas las neuronas se actualizan simultáneamente



BRAIN STATE IN A BOX (BSB)

La función de la neurona es una rampa continua y limitada:

$$f(\text{net}) = \min(1, \max(-1, \text{net}))$$

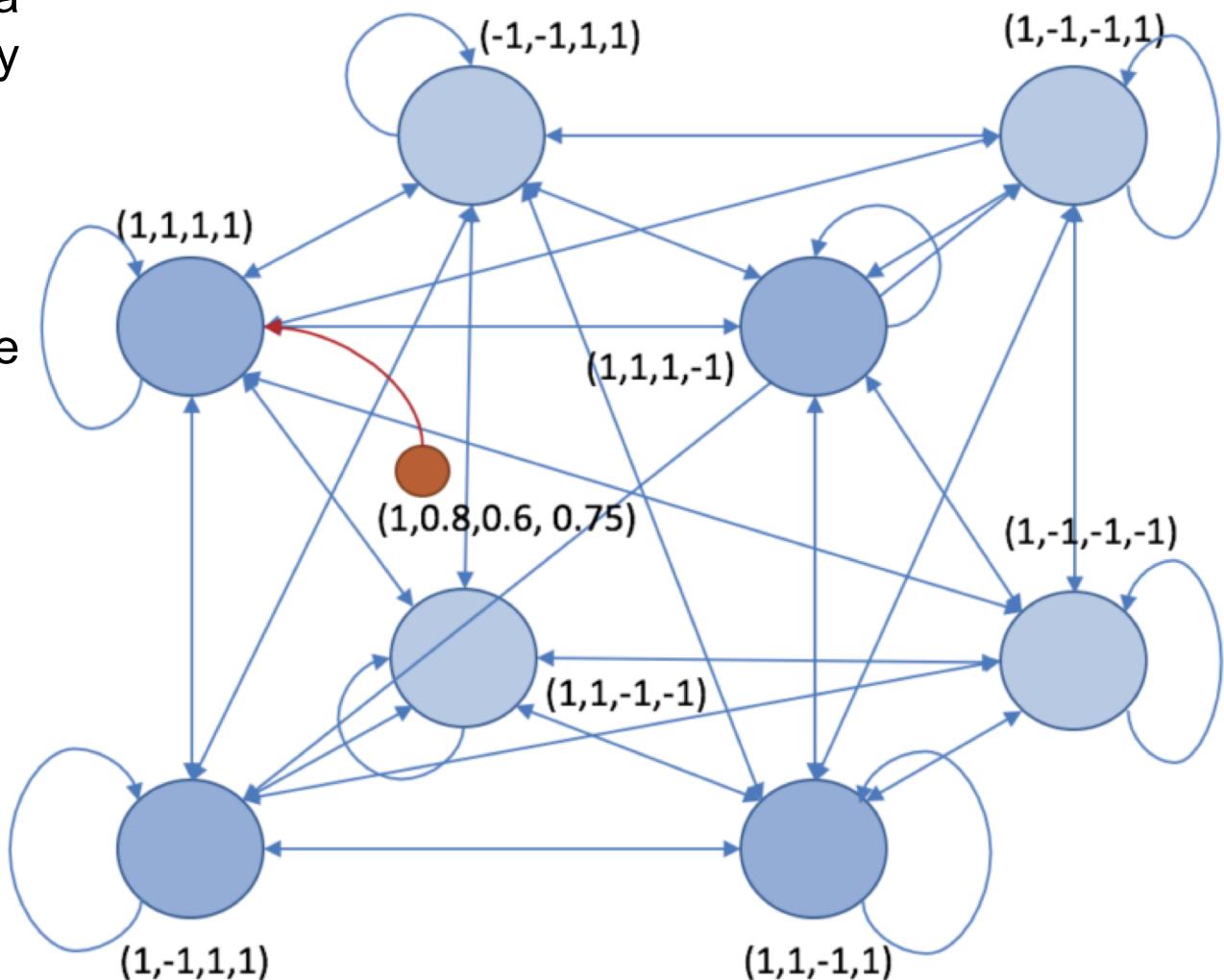
El cambio de estado de una neurona es:

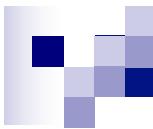
$$x_i(t+1) = f\left(\sum_{j=1,n} w_{ij} x_j(t)\right)$$

Con $x_i(t)$ el estado de la neurona i_{th} en t .

Los pesos se calculan como:

$$w_{ij} = 1/P \sum_p P(v_{p,i} v_{p,j})$$





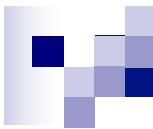
VENTAJAS E INCONVENIENTES

■ Ventajas

- Tolerantes a fallos
- Procesos distribuidos
 - Cada nodo trabaja con información local

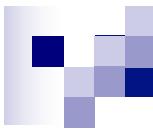
■ Inconvenientes

- El tiempo crece exponencialmente con el tamaño de la máquina y con la magnitud de las conexiones
- Mínimos locales
- Aprendizaje no muy definido
- Entradas ortogonales



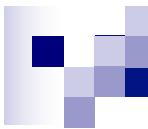
APLICACIONES

- Memorias auto-asociativas y direccionables por contenido
- Validación de hipótesis
 - Representación distribuida del conocimiento
- Regeneración de información de entrada incompleta o distorsionada



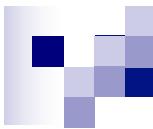
OTROS TIPOS DE REDES

- Self-Organizing Maps (SOM)
- Mapas Topológicos de Kohonen
- Art (Teoría de Resonancia Adaptiva)
- Neocognitrón
- ...
- Deep learning (Redes Convolucionales)



DEEP LEARNING (APRENDIZAJE PROFUNDO)

- **Subcampo dentro de las redes neuronales artificiales**
- Redes con gran cantidad de capas de nodos
 - A mayor cantidad de capas, mayor es la profundidad de la red
- Requisitos:
 - Gran conjunto de datos disponible, para entrenar a los sistemas
 - Capacidad y el rendimiento de los sistemas de almacenamiento
 - Potencia de procesamiento (uso de las GPU y chips específicos)
- Aplicaciones:
 - Conversión de voz a texto
 - Traducción de idiomas
 - Reconocimiento de imágenes (caras, matrículas coche, médico, ...)
 - Asistente personal (patrón de comportamiento)



REDES NEURONALES CONVOLUCIONALES (CNN)

- La CNN es un tipo de Red Neuronal Artificial con aprendizaje supervisado que procesa sus capas imitando al cortex visual del ojo humano para identificar distintas características en las entradas
- Contiene varias capas ocultas especializadas y con una jerarquía: (las primeras capas pueden detectar líneas, curvas y se van especializando hasta llegar a capas más profundas que reconocen formas complejas como un rostro o la silueta de un animal)

