

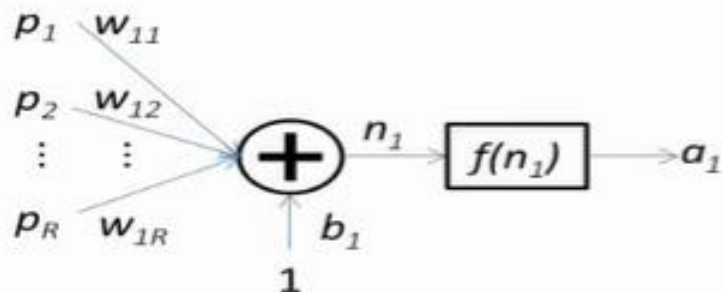


INICIO GRABACIÓN



SANJOSÉ
FUNDACIÓN DE EDUCACIÓN SUPERIOR

Neurona Artificial



p_i , entradas o patrones

w_{ij} , pesos sinápticos

b_i , polarización

n_i , entrada neta

a_i , salida (axon)

f , función de activación

$$n_1 = p_1 w_{11} + \dots + p_R w_{1R} + b_1$$

$$n_1 = \mathbf{w}_1^T \mathbf{p} + b_1$$

$$a_1 = f(\mathbf{w}_1^T \mathbf{p} + b_1)$$

$$\mathbf{w}_1 = \begin{bmatrix} w_{11} \\ w_{12} \\ \vdots \\ w_{1R} \end{bmatrix} \quad \mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_R \end{bmatrix}$$


$$n_1 = p_1 w_{11} + \dots + p_R w_{1R} + b_1$$

$$n_1 = \mathbf{w}_1^T \mathbf{p} + b_1$$

$$a_1 = f(\mathbf{w}_1^T \mathbf{p} + b_1)$$



Algoritmo del perceptrón

Inicialización aleatoria de w y b

Desde épocas = 1 a N repetir

Desde $q=1$ a Q repetir

$$a_q = \text{función escalón}(wP_q + b)$$

$$e_q = t_q - a_q$$

$$W = w + e_q * P_q^T$$

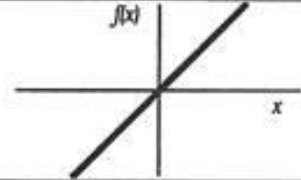
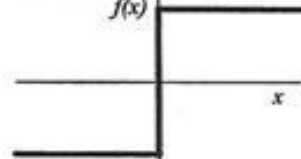
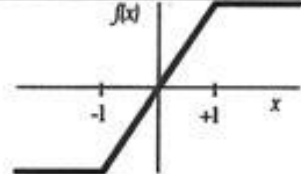
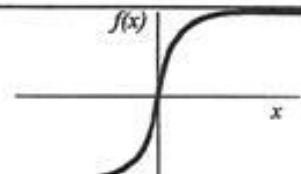
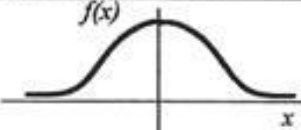
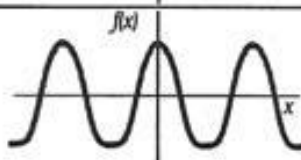
$$b = b + e_q$$

fin

fin





	Función	Rango	Gráfica
Identidad	$y = x$	$[-\infty, +\infty]$	
Escalón	$y = \text{sign}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$	
Lineal a tramos	$y = \begin{cases} -1, & \text{si } x < -l \\ x, & \text{si } -l \leq x \leq +l \\ +1, & \text{si } x > +l \end{cases}$	$[-1, +1]$	
Sigmoidea	$y = \frac{1}{1 + e^{-x}}$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$	
Gaussiana	$y = Ae^{-Bx^2}$	$[0, +1]$	
Sinusoidal	$y = A \text{sen}(\omega x + \varphi)$	$[-1, +1]$	




CARACTERISTICAS

Las Redes ADALINE son redes muy similares al Perceptrón con la diferencia de que su función de activación es lineal en lugar de ser un limitador fuerte como es el caso del Perceptrón, estas presentan la misma limitación del Perceptrón respecto al tipo de problemas que pueden resolver, ya que ambas redes solo pueden resolver problemas linealmente separables.

Son redes de aprendizaje supervisado que usan la regla de Widrow – Hoff para dicho aprendizaje o también denominada regla Delta. El algoritmo que estas usan es el LMS (Least Mean Square) siendo este más eficiente que la regla de aprendizaje del Perceptrón puesto que minimiza el error medio cuadrático.

Adaline está limitada a una única neurona de salida, un vector x como su entrada y un número real y como su salida.





ALGORITMO DE APRENDIZAJE:

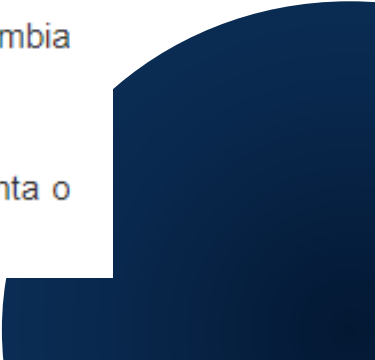
- 1.- Inicializar los pesos en forma aleatoria
- 2.- Introducir un patrón de entrada
- 3.- Calcular la salida (y), compararla con la deseada (d) y obtener la diferencia ($dp - y_p$)
- 4.- Multiplicar el resultado del paso anterior por la entrada correspondiente a cada uno de los pesos y ponderarla por la tasa de aprendizaje.
- 5.- Actualizar los pesos, sumando al valor antiguo la cantidad obtenida en el paso anterior
- 6.- Si no se ha cumplido el criterio de parada, regresar al paso 2, si se ha acabado todos los patrones repetir el algoritmo.

Criterios de parada:

Criterio 1: Fijar un número de ciclos máximo. Dicho número debe garantizar que el error cuadrático para los patrones de entrenamiento se haya estabilizado.

Criterio 2: Cuando el error cuadrático sobre los patrones de entrenamiento no cambia durante x ciclos.

Criterio 3: Cuando el error cuadrático sobre los patrones de validación no aumenta o se mantiene estable a lo largo de x ciclos.





La Regla Delta



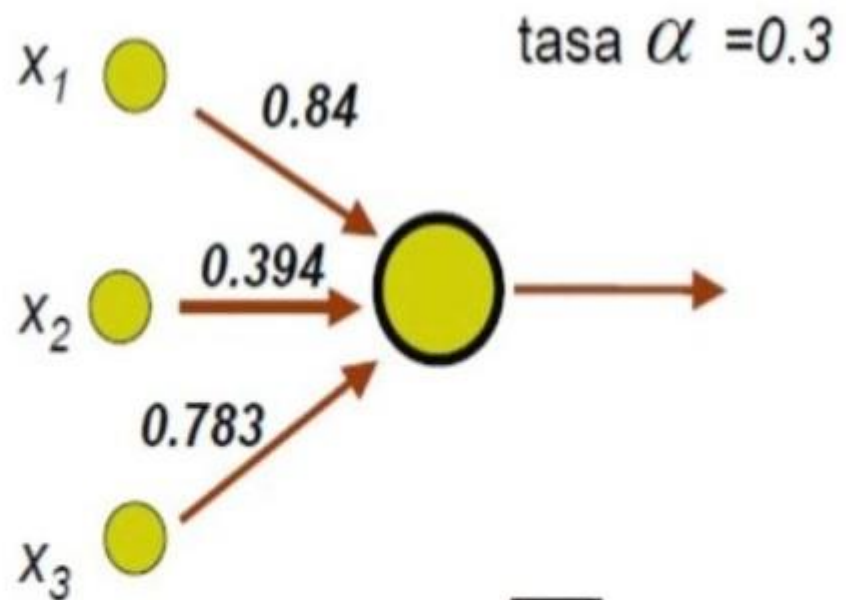
No es posible conseguir una salida exacta, pero sí minimizar la desviación cometida por la red, esto se hace minimizando el error cometido por la red para la totalidad del conjunto de patrones de aprendizaje. Hay que evaluar globalmente el error cometido por la red para todos los patrones, una medida de error global.

Medida de Error Global

La medida de error global utilizada es el **error cuadrático medio**

$$E = \sum_{p=1}^m E^p = \frac{1}{2} \sum_{p=1}^m (d^p - y^p)^2 \quad (6)$$

0	0	1	[1]
0	1	0	[2]
0	1	1	[3]
1	0	0	[4]
1	0	1	[5]
1	1	0	[6]
1	1	1	[7]



$$y = \sum_{\forall i} w_i x_i$$

0	0	1	[1]
0	1	0	[2]
0	1	1	[3]
1	0	0	[4]
1	0	1	[5]
1	1	0	[6]
1	1	1	[7]

x_1

x_2

x_3

0...

0...

0...

tasa $\alpha = 0.3$

$y = 3.09 \cdot 1 + 1.966 \cdot 1 + 1.825 \cdot 1 = 6.881$

$E = \sum_{\forall p} (d_p - y_p) = 7 - 6.881 = 0.12$

**Resultado
despues de la
primera iteración
del
entrenamiento**

$w_1 = w_1 + \alpha E \cdot x_1 = 3.09 + 0.3 \cdot 0.12 = 3.126$

$w_2 = w_2 + \alpha E \cdot x_2 = 1.966 + 0.3 \cdot 0.12 = 2.002$

$w_3 = w_3 + \alpha E \cdot x_3 = 1.825 + 0.3 \cdot 0.12 = 1.861$



FUNDACIÓN DE EDUCACIÓN SUPERIOR

SAN JOSÉ

INSTITUCIÓN TECNOLÓGICA

FIN DE
GRABACIÓN