

# Redes Neuronales

Elementos básicos de las redes neuronales  
carlos.andres.delgado@correounivalle.edu.co

Carlos Andrés Delgado S.

Universidad San Buenaventura, Cali

Junio de 2021

# Contenido

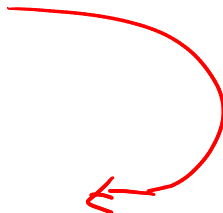
1 Elementos básicos

2 Tipos de aprendizaje

3 Arquitecturas de red Neuronal

$O(n^3)$

$O(n)$



# Contenido

**1** Elementos básicos

2 Tipos de aprendizaje

3 Arquitecturas de red Neuronal

# Elementos básicos



## Propiedades redes neuronales

- Aprendizaje adaptativo
- Generalización
- Naturaleza para propósito no-lineal
- Auto-organización
- Paralelismo masivo
- Robustez y tolerancia a ruido

$$1 + 3 = 4$$

$$3^{96} + 8^{57} = ?$$

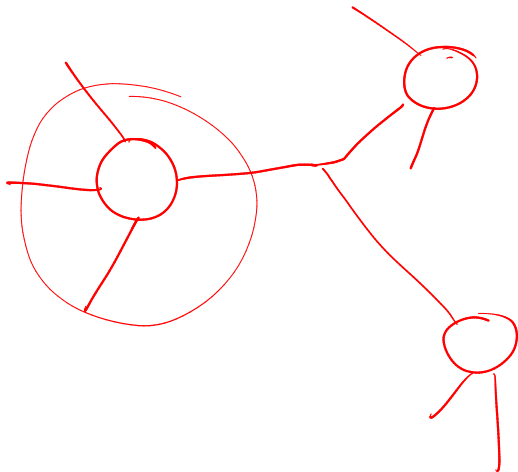
IA débil

IA fuerte

# Modelo de una neurona

## Modelo no lineal

- 1 Cada neurona recibe un conjunto de señales discretas o continuas
  - 2 Estas señales se ponderan o integran
  - 3 Cada conexión tiene un peso sináptico
  - 4 Los pesos **representan el conocimiento**
  - 5 Estos pesos se ajustan con **algoritmos de aprendizaje**
-



# Modelo de una neurona

## Modelo no lineal

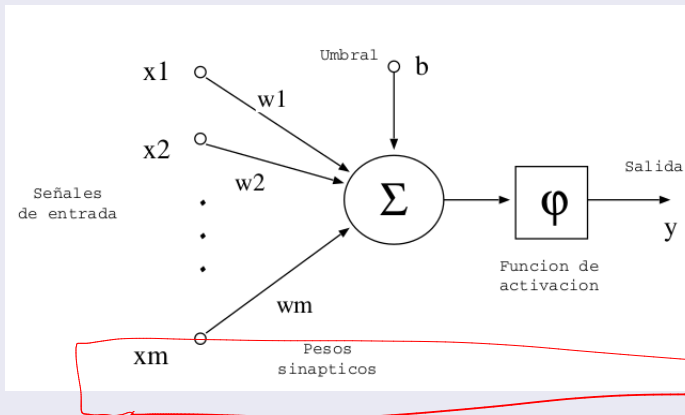



Figura: Modelo no lineal. Tomado de: [Pérez Ortiz, 1999]

# Modelo de una neurona

## Modelo no lineal

Una red neuronal tiene:

- 1 Un conjunto  $m$  de señales de entrada 
- 2 Un conjunto de sinapsis  $w_{ji}$ , donde  $i$  indica la  $i$ -ésima entrada de la neurona  $j$
- 3 Un umbral o sesgo  $b$ , puede ser positivo o negativo
- 4 Las entradas son sumadas o integradas, tomando en cuenta sus respectivos pesos
- 5 Se tiene una función de activación  $\sigma$  que describe el funcionamiento de la neurona



# Modelo de una neurona



Modelo no lineal

Este modelo lo podemos describir así:

$$z = \varphi\left(\sum_{i=1}^n mw_i x_i + b\right)$$

En forma vectorial:

$$z = \varphi(wx^T + b)$$



# Modelo de una neurona

## Funciones de activación

Con una función:

1 Función lineal: Suele variar entre 0 y 1 o -1 y 1.

2 Función escalón. Salida bivaluada  $\varphi(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x \geq 0 \end{cases}$

3 Función sigmoidea. Transformación no lineal de la entrada

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

Suele utilizarse  $a = 1$

# Modelo de una neurona

## Funciones de activación

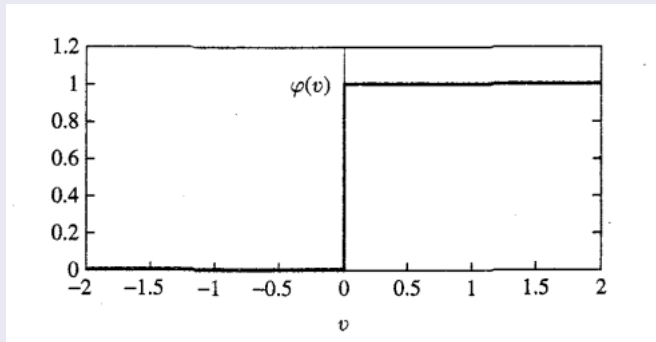


Figura: Función escalón. Tomado de: [Haykin, 1998]

# Modelo de una neurona

## Funciones de activación

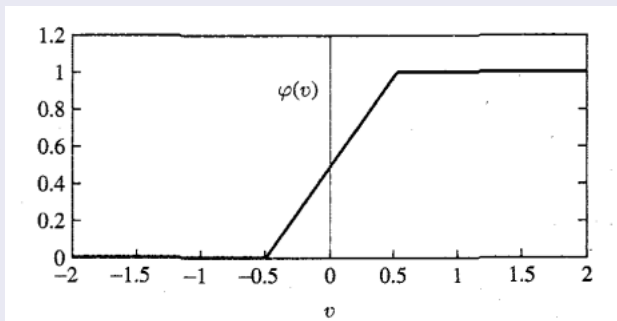


Figura: Función lineal. Tomado de: [Haykin, 1998]



## Modelo de una neurona

## Funciones de activación

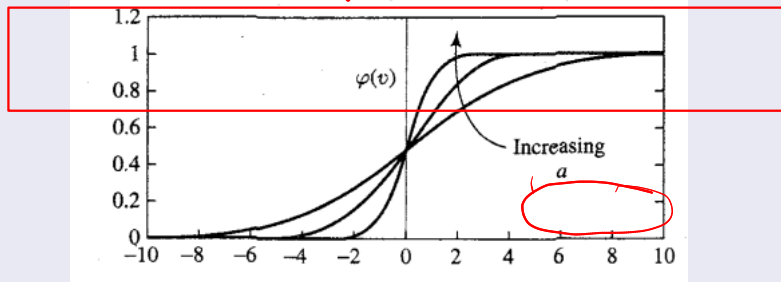


Figura: Función sigmoide. Tomado de: [Haykin, 1998]

# Modelo de una neurona

## Funciones de activación

Modelo estocástico, dada una distribución de probabilidad  $P(v)$

$$x = \begin{cases} 1 & \text{con } P(v) \\ -1 & \text{con } 1 - P(v) \end{cases}$$

# Contenido

1 Elementos básicos

2 Tipos de aprendizaje

3 Arquitecturas de red Neuronal

# Tipos de aprendizaje

## El aprendizaje

El aprendizaje en las redes neuronales se puede modelar así.

$$w(t+1) = w(t) + \Delta w(t)$$



# Tipos de aprendizaje

## Aprendizaje supervisado

- Basado en la comparación entre la salida actual y la deseada
- Los pesos se ajustan de acuerdo a patrón de entrenamiento de acuerdo
- Existe un criterio de parada para el proceso de aprendizaje de acuerdo a la medida del error

$$E = \frac{1}{N} \sum_{p=1}^N (y_d - y_c)^2$$

# Tipos de aprendizaje

## Aprendizaje no supervisado


- No hay valores objetivos
- Está basado en las correlaciones entre la entrada y patrones significantes que ayuden en el aprendizaje
- Se requiere un método de parada

# Tipos de aprendizaje

## Aprendizaje por refuerzo

- Es un caso especial de aprendizaje supervisado
- La salida deseada es desconocida
- Se castiga una mala salida y se premia una buena salida

# Tipos de aprendizaje

**Pesos** → 

**Aprendizaje evolutivo**

- Se utilizan algoritmos evolutivos para ajustar los pesos
- Se tienen funciones de evaluación de la salida de la red

$$X^* = X_1 * w_1 + X_2 * w_2 + \dots + X_n * w_n$$

$$X = \sum_{i=1}^n x_i w_i$$

$$F(x^*)$$

Función de activación

$$\left\{ \begin{array}{l} F(x) = x \\ F(x) = \tanh(x) \\ F(x) = \frac{1}{1 + e^{-x}} \end{array} \right.$$

# Contenido

1 Elementos básicos

2 Tipos de aprendizaje

3 Arquitecturas de red Neuronal

# Arquitecturas de red Neuronal

$(1, 0)$   $\boxed{\times}$   $1$   
 $(1, 0.1)$   $\nearrow$

## Clases de arquitecturas

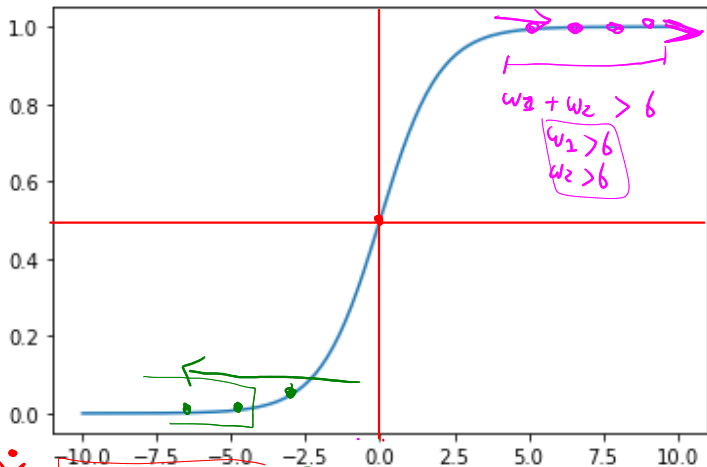
### Redes de una capa sin ciclos

- Es la forma más simple
- Consiste en una capa que recibe las entrada y emite una o más salidas

~~Pesos~~

e/

Solid



$$x_1 = 0$$
$$x_2 = 1$$

$\dot{x} = \begin{bmatrix} w_1 x_1 + w_2 x_2 \end{bmatrix} - 6$

$b = 0$   
 $x_1 = 0$   
 $x_2 = 0$

OR



$$\dot{X} = X_1 \cdot w_1 + X_2 w_2 + X_3 w_3 - b$$

$$\hat{y} = f(\dot{X})$$



# Arquitecturas de red Neuronal

## Red de una capa sin ciclos

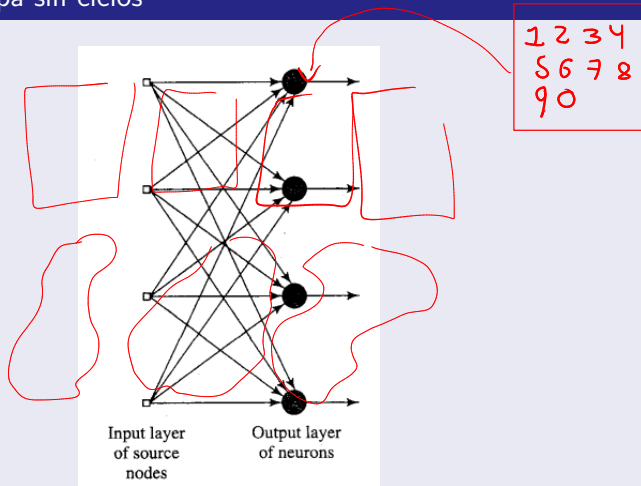


Figura: Esquema red de una capa. Tomado de: [Haykin, 1998]

# Arquitecturas de red Neuronal



## Multicapa sin ciclos

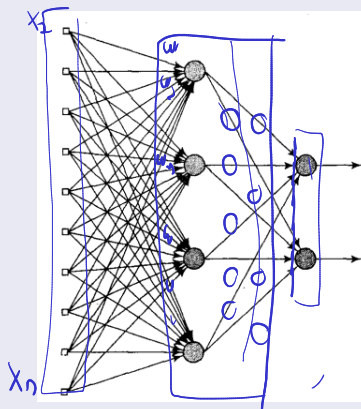
- Tiene una capa de entrada
- Tiene capas ocultas
- Tiene capas de salida

← Pesos y aprendizaje

3 0 4  
oculta

# Arquitecturas de red Neuronal

## Multicapa sin ciclos



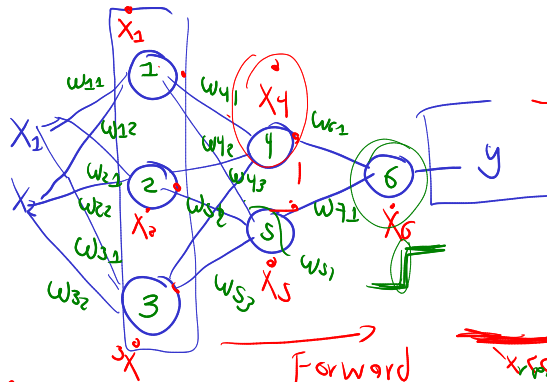
Numero de salidas

1. Salidas  $\begin{cases} 0 \\ 1 \end{cases}$

2. Salidas  $\begin{cases} 01 \\ 10 \\ 11 \\ 00 \end{cases}$

3. Salidas  $\begin{cases} \text{Perro } 10 \\ \text{Gato } 01 \\ \text{Lobo } 00 \end{cases}$

Figura: Esquema red multicapa. Tomado de: [Haykin, 1998]



tanh Sigmoid  
|neg|

$F(\dot{X}_6)$

Function activation



Softmax

$F$

$$\dot{X}_1 = w_{11} X_1 + w_{12} X_2 - b_1$$

$$S_1 = F(\dot{X}_1)$$

$$\dot{X}_2 = w_{21} X_1 + w_{22} X_2 - b_2$$

$$S_2 = F(\dot{X}_2)$$

$$\dot{X}_3 = w_{31} X_1 + w_{32} X_2 - b_3$$

$$S_3 = F(\dot{X}_3)$$

$$\dot{X}_4 = w_{41} S_1 + w_{42} S_2 + w_{43} S_3 - b_4$$

$$S_4 = F(\dot{X}_4)$$

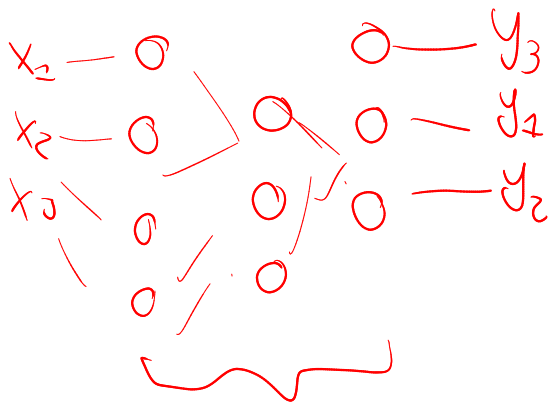
$$\dot{X}_5 = w_{51} S_4 + w_{52} S_2 + w_{53} S_3 - b_5$$

$$S_5 = F(\dot{X}_5)$$

$$\dot{X}_6 = w_{61} S_4 + w_{62} S_5 = b_6$$

$$y = F(\dot{X}_6)$$

3 entrada



$$\begin{aligned} \dot{X}_1 &= w_{11} X_1 + w_{12} X_2 - b_1 \\ \dot{X}_2 &= w_{21} X_1 + w_{22} X_2 - b_2 \\ \dot{X}_3 &= w_{31} X_1 + w_{32} X_2 - b_3 \end{aligned}$$

$$\begin{aligned} S_1 &= f(\dot{X}_1) \\ S_2 &= f(\dot{X}_2) \\ S_3 &= f(\dot{X}_3) \end{aligned}$$

$$\dot{X}_4 = w_{41} S_1 + w_{42} S_2 + w_{43} S_3 - b_4 \quad S_4 = f(\dot{X}_4)$$

$$\dot{X}_5 = w_{51} S_4 + w_{52} S_2 + w_{53} S_3 - b_5 \quad S_5 = f(\dot{X}_5)$$

$$\dot{X}_6 = w_{61} S_4 + w_{62} S_5 = b_6$$

$$y = f(\dot{X}_6)$$

or

$$\begin{matrix} \text{X}_1 \\ \text{X}_2 \end{matrix} \quad \begin{matrix} \text{N}_1 \\ \text{N}_2 \\ \text{N}_3 \end{matrix} \quad \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix} \quad \begin{matrix} \text{X}_1 \\ \text{X}_2 \end{matrix} \quad \begin{matrix} \text{pesos } 3 \times 2 \\ 2 \times 1 \end{matrix} \quad \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \begin{matrix} \text{X}_1 \\ \text{X}_2 \\ \text{X}_3 \end{matrix} \quad \begin{matrix} 3 \times 1 \end{matrix}$$

$$f\left(\begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \\ \dot{X}_3 \end{bmatrix}\right) = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} \quad \text{Segundo de la primera capa}$$

$$\begin{aligned} \dot{X}_y &= w_{y1} s_1 + w_{y2} s_2 + w_{y3} s_3 - b_y & s_y &= F(\dot{X}_y) \\ \dot{X}_s &= w_{s1} s_1 + w_{s2} s_2 + w_{s3} s_3 - b_s & s_s &= F(\dot{X}_s) \\ \dot{X}_G &= w_{G1} s_1 + w_{G2} s_2 = b_G & \mathbf{y} &= \mathbf{F}(\mathbf{X}_G) \end{aligned}$$

$$S_{a1} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}_{3 \times 1}$$

$$w_{c2} = \begin{bmatrix} w_{y1} & w_{y2} & w_{y3} \\ w_{s1} & w_{s2} & w_{s3} \end{bmatrix}_{2 \times 3} \begin{matrix} \leftarrow N_y \\ \leftarrow N_s \end{matrix}$$

$$\begin{bmatrix} w_{y1} & w_{y2} & w_{y3} \\ w_{s1} & w_{s2} & w_{s3} \end{bmatrix}_{2 \times 3} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}_{3 \times 1} = \begin{bmatrix} b_y \\ b_s \end{bmatrix}_{2 \times 1} = \begin{bmatrix} \dot{X}_y \\ \dot{X}_s \end{bmatrix}_{2 \times 1}$$

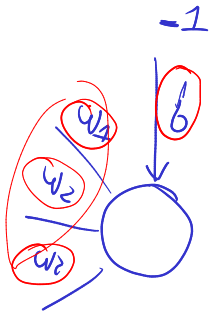
$S = F(\dot{X}_{c2})$

$$X_6 = W_{61} S_4 + W_{62} S_5 = b_6 \quad \sigma_n \quad \boxed{y = F(X_6)}$$

$$\begin{bmatrix} W_{61} & W_{62} \end{bmatrix} \begin{bmatrix} S_4 \\ S_5 \end{bmatrix} - b_6 = X_6$$

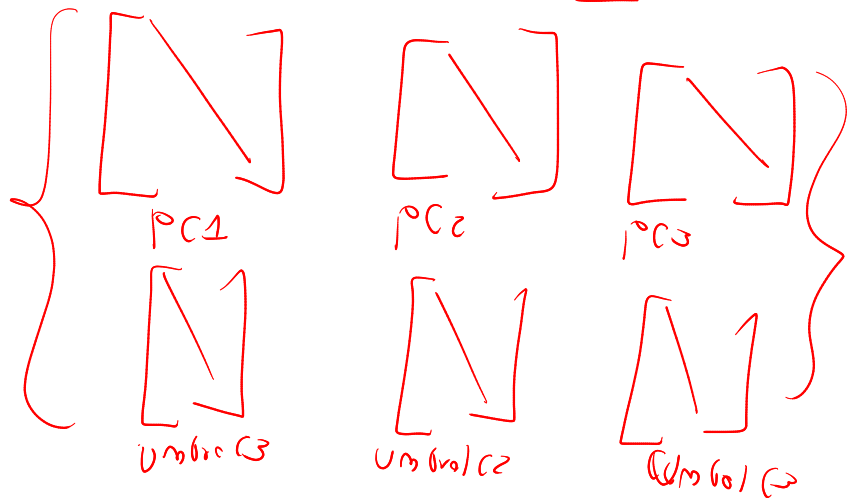
$1 \times 2 \quad 2 \times 1$

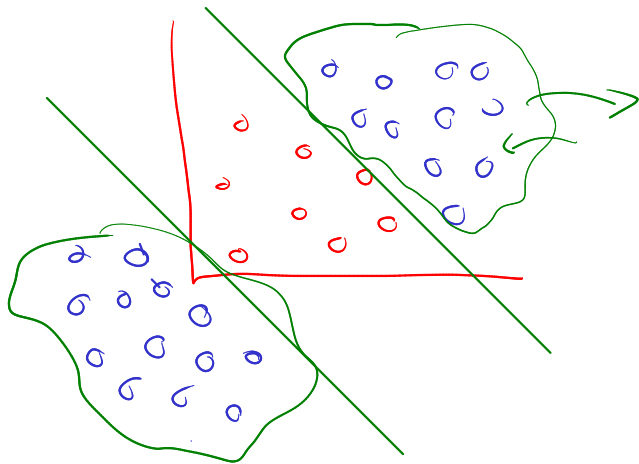
$$y = F(X_6)$$





Entrenamiento (horas) Modelo



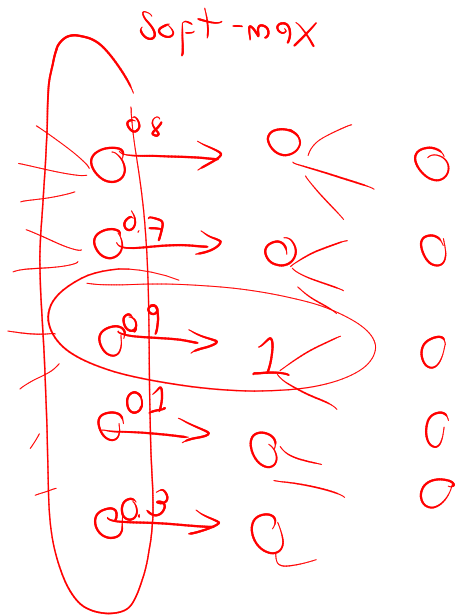
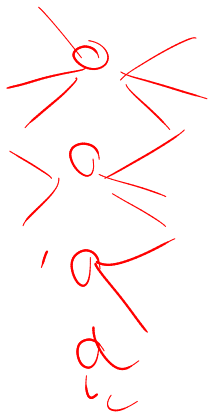


# Arquitecturas de red Neuronal




## Redes recurrentes

- Tienen estructura monocapa o multicapa
- La salidas se conectan a las entradas, pero estas tienen un retardo





# Referencias I

-  Du, K. and Swamy, M. (2006).  
*Neural Networks in a Softcomputing Framework*.  
Springer-Verlag.
-  Haykin, S. (1998).  
*Neural Networks: A Comprehensive Foundation (2nd Edition)*.  
Prentice Hall.
-  Pérez Ortiz, J. A. (1999).  
Clasificación con discriminantes: Un enfoque neuronal.  
[http:  
//www.dlsi.ua.es/~japerez/pub/pdf/cden1999.pdf](http://www.dlsi.ua.es/~japerez/pub/pdf/cden1999.pdf).  
Material de clase, Accessed: Ago-2017.

# ¿Preguntas?

Próximo tema:  
Perceptrón y adeline





→ clasificación de documentos







