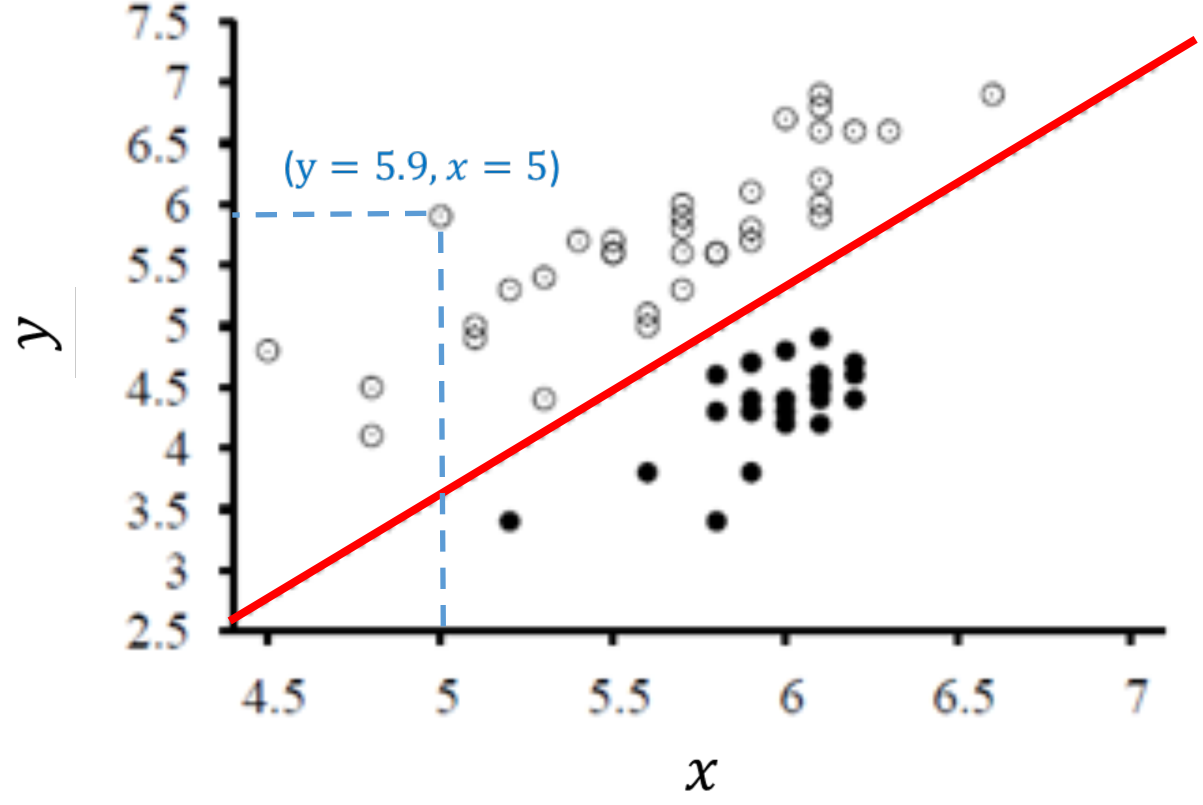
# Redes neuronales artificiales

Tema 1: El perceptrón.



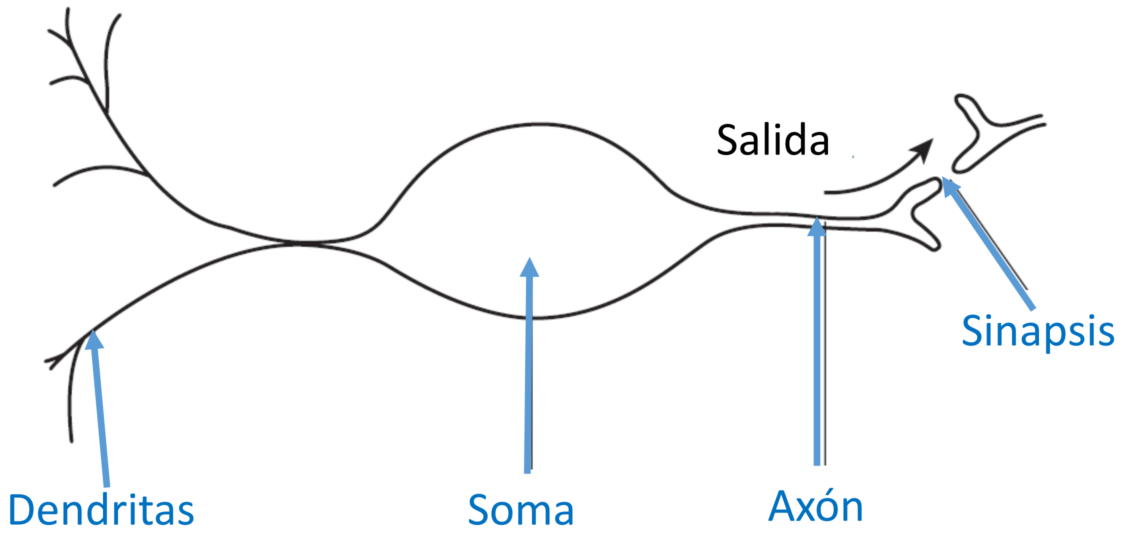
(y = 5.9, x = 5)
7.5
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
y
4.5
5
5.5
6
6.5
7
x

# Índice:

- La neurona biológica.
- La neurona artificial.
- Funciones de activación.
- El perceptrón como clasificador.

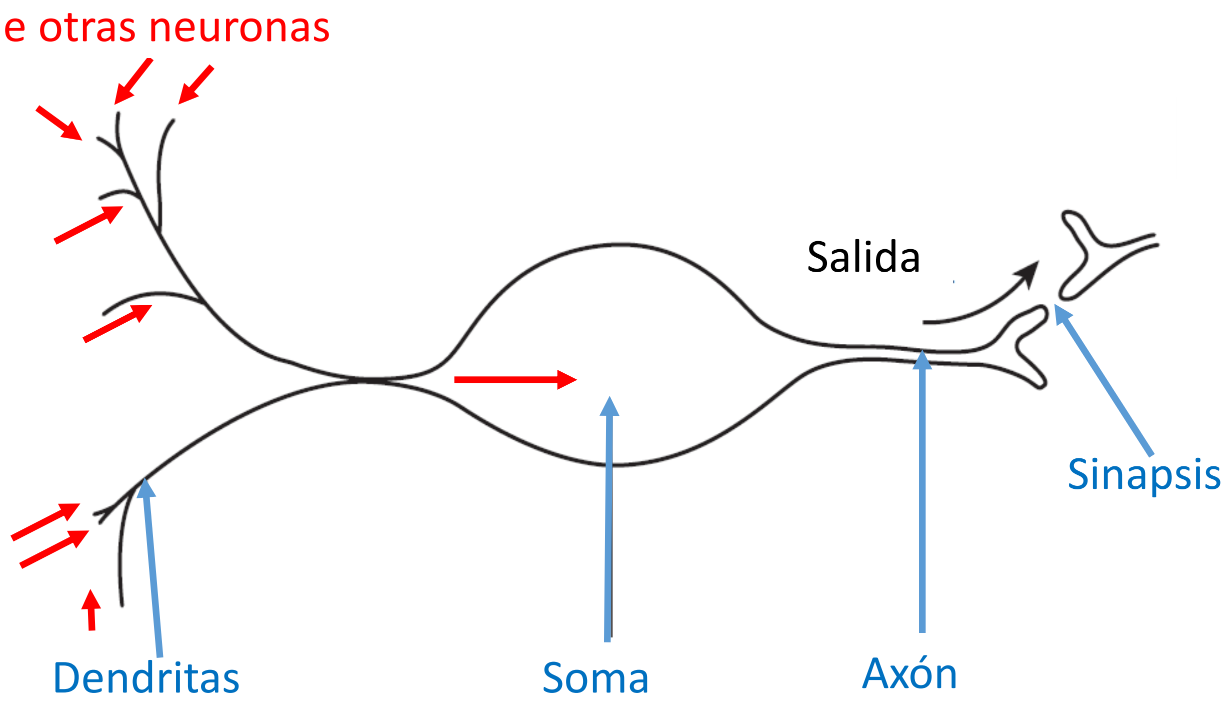
# La neurona biológica:

- Una neurona es una célula excitable eléctricamente que se comunica con otras células a través de conexiones llamadas sinapsis.
  - Neuronas sensoriales.
  - Neuronas motoras.
  - Inter-neuronas.
- Cada neurona es una unidad sencilla de procesamiento que contiene:
  - Un soma, o cuerpo de la neurona
  - Un axón.
  - Un número de dendritas.



# La neurona biológica:

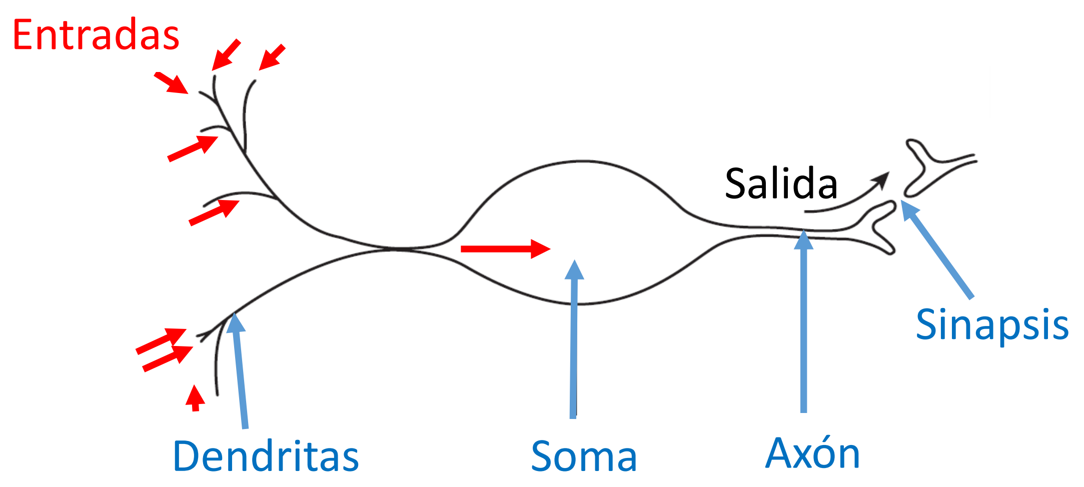
- Cuando la entrada de la neurona supera un cierto umbral, se "activa" la salida.
- Se produce una reacción química que provoca un pulso de salida llamado potencial de acción que se propaga a través del axón hacia la sinapsis.



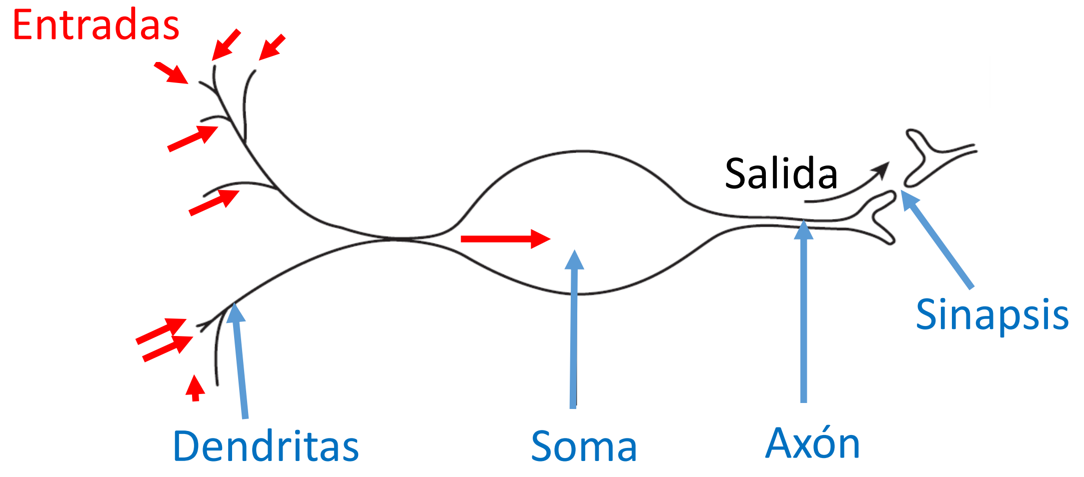
# Redes neuronales artificiales:

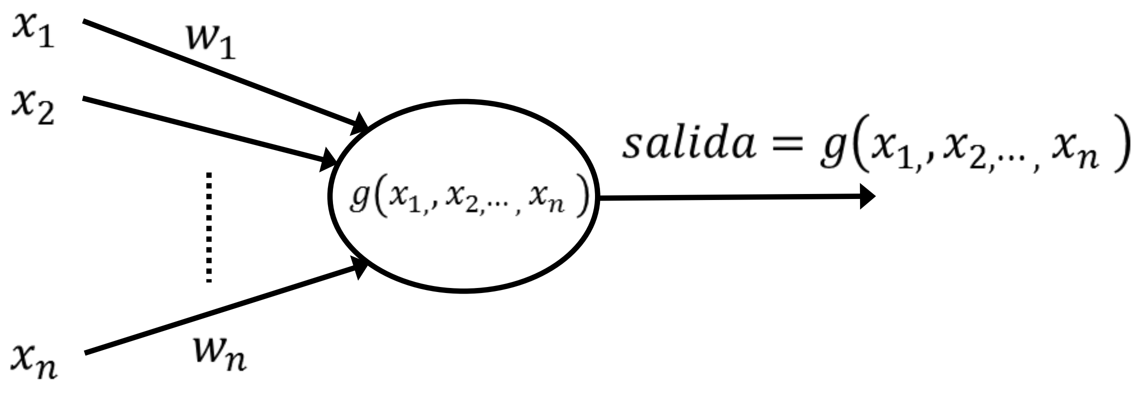
- Modelan el cerebro humano y consisten en una red de neuronas artificiales.
- Las redes neuronales artificiales y las neuronas artificiales son más sencillas que las biológicas.
  - Las redes tienen menos neuronas.
  - Las neuronas tienen menos conexiones.
- Cada neurona va a ser un nodo en la red y va a recibir un número de entradas de otras neuronas.
  - A estas entradas se les va a aplicar una función de activación que dará como salida un nivel de activación.

# Neurona biológica vs neurona artificial:

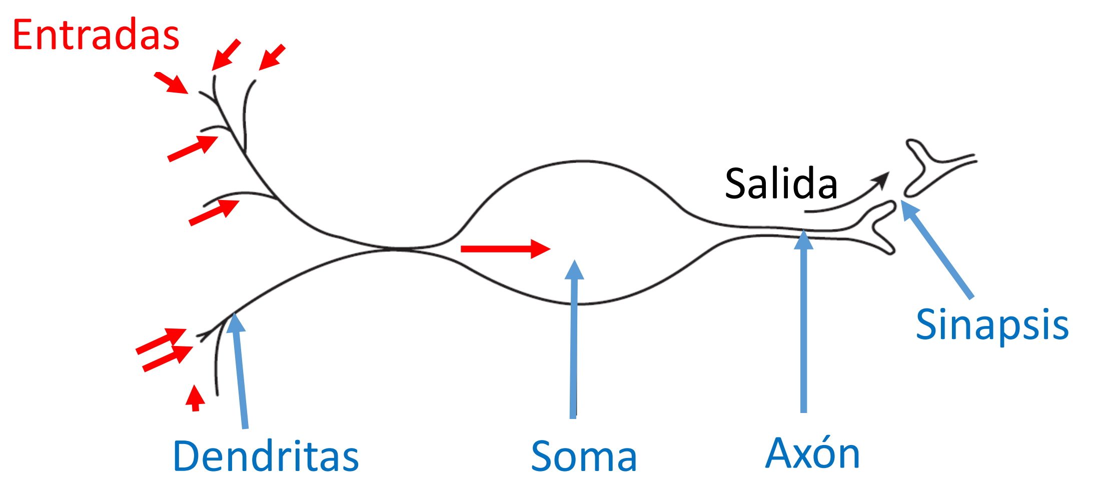


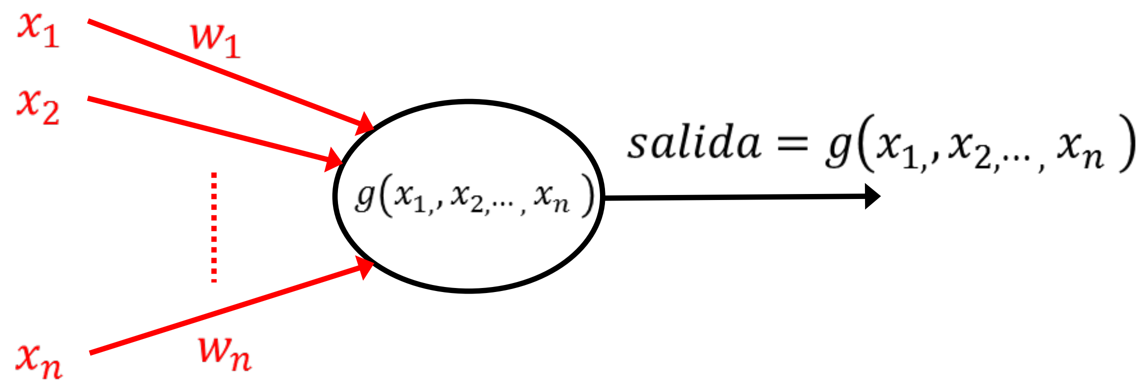
# Neurona biológica vs neurona artificial:



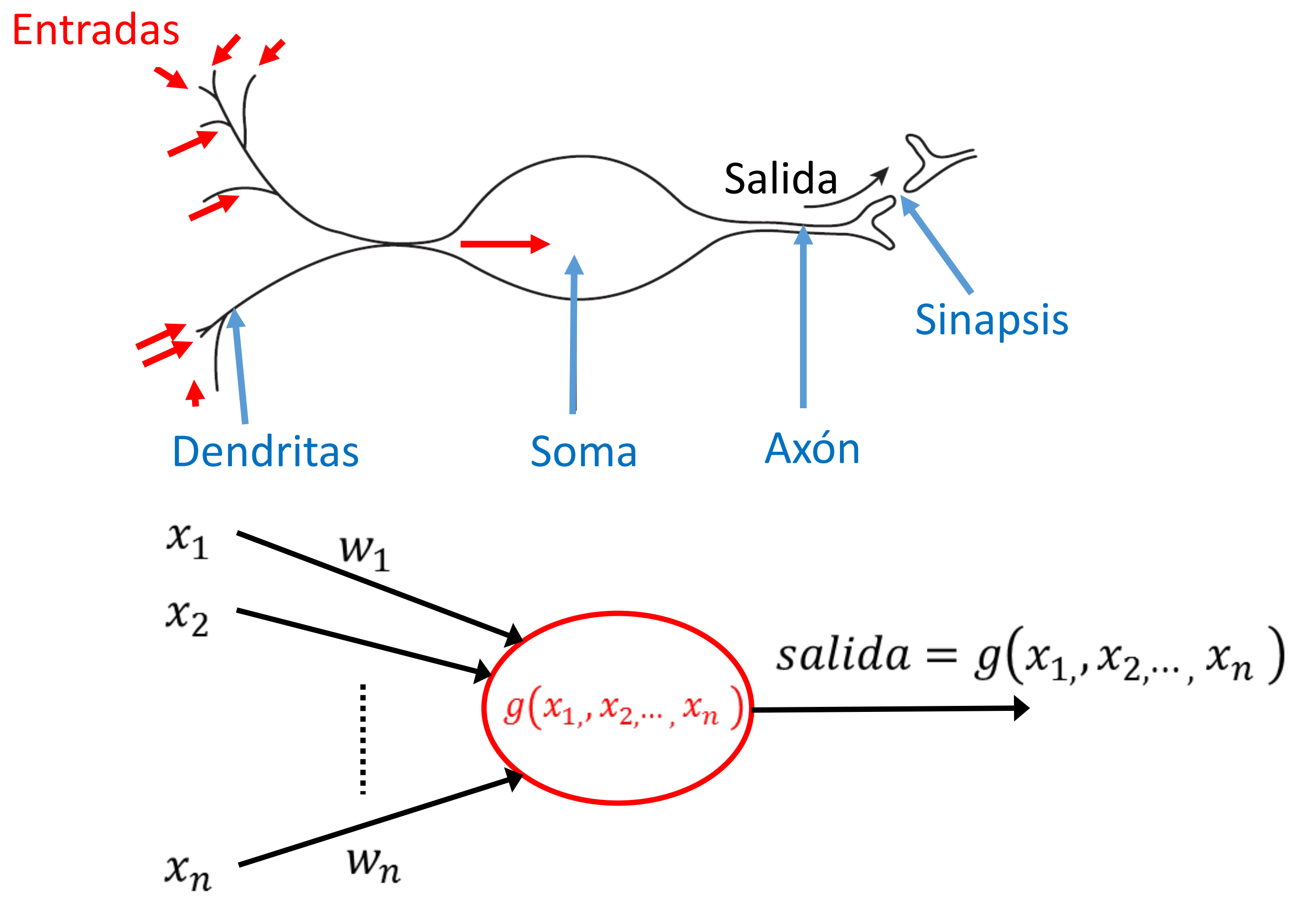



# Neurona biológica vs neurona artificial:

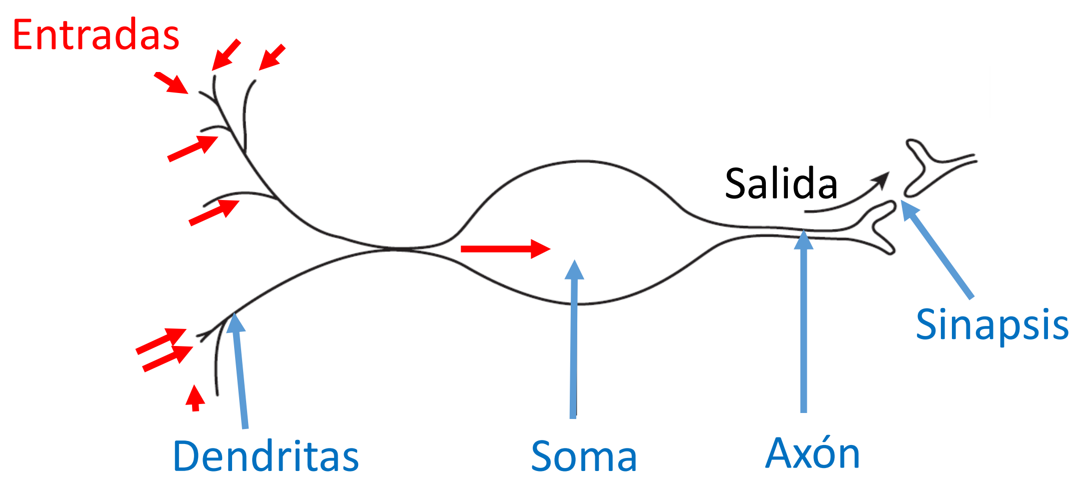


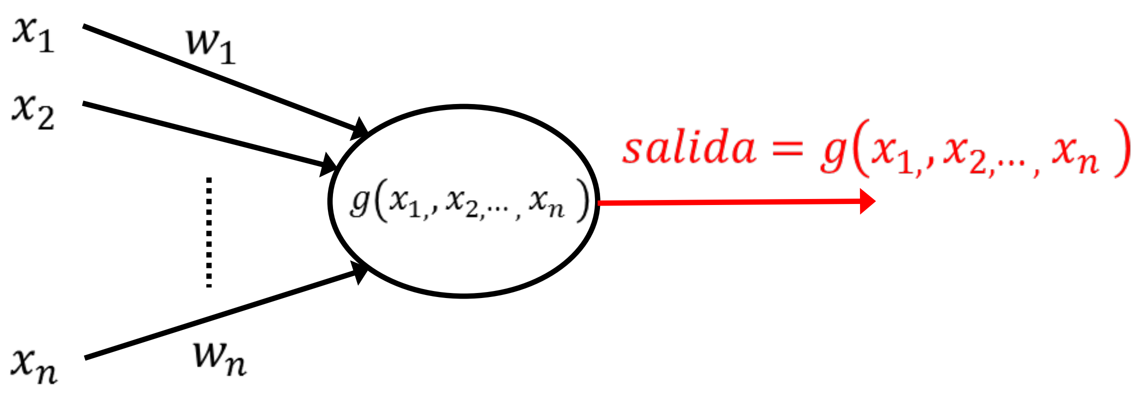



# Neurona biológica vs neurona artificial:

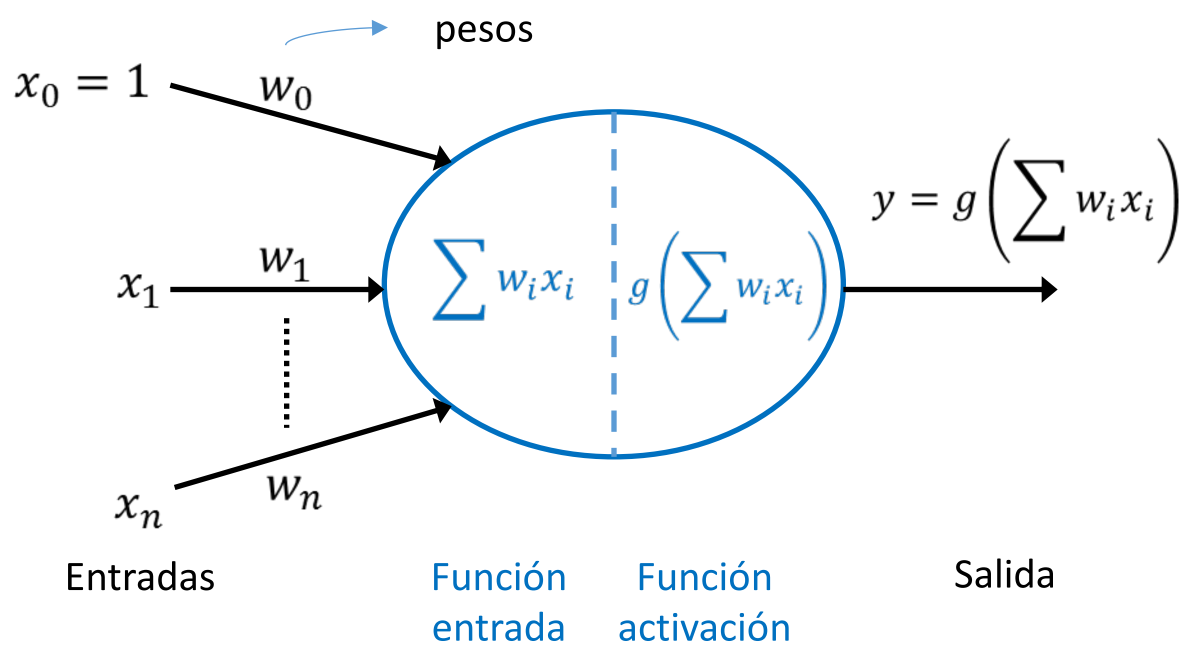


# Neurona biológica vs neurona artificial:



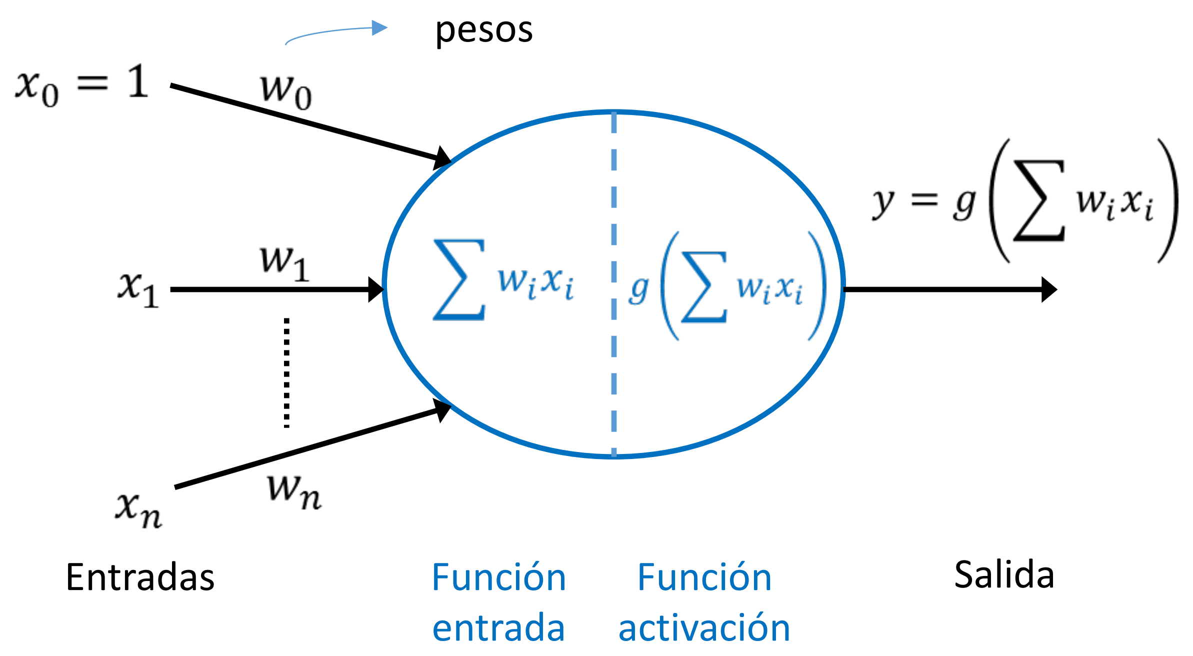



# La neurona artificial:




- Las redes neuronales están compuestas por nodos conectados mediante enlaces direccionados:
  - Cada enlace tiene asociado un peso $w_i$ que determina la fuerza y signo de la conexión.
  - Cada nodo va a tener una entrada “dummy” en la figura la $x_0$.

# La neurona artificial:

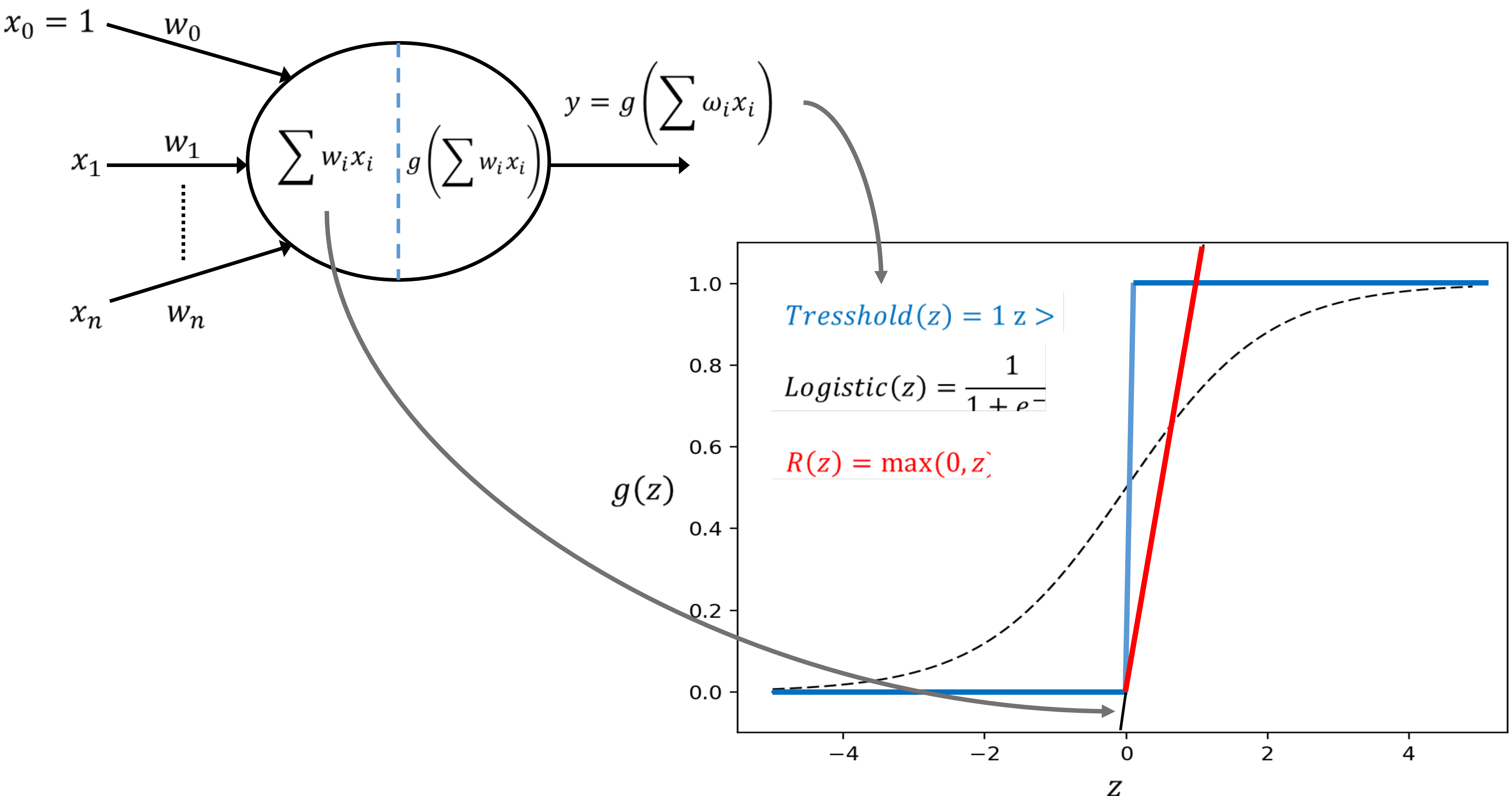



- La neurona calcula una suma ponderada de las entradas de acuerdo con los pesos $\sum_{i=0}^{n} w_i x_i$.
- Después a esta suma se le aplica una función de activación $g(\sum w_i x_i)$.

# Funciones de activación:

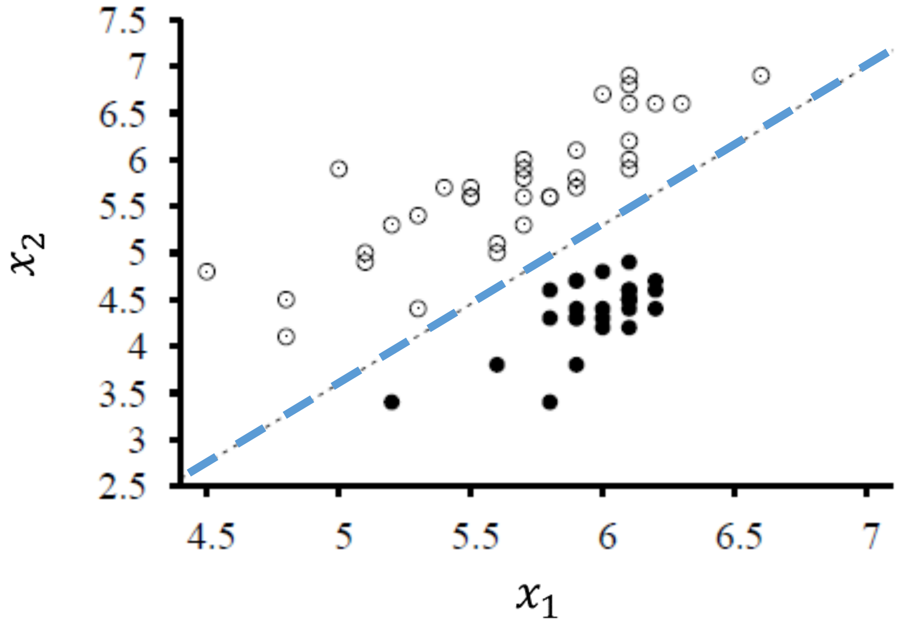
- Las funciones de activación $g(\sum w_i \cdot x_i) = g(z)$ pueden ser:
  - Un hard tresshold o función step (perceptron):
    - $Step(z) = 1 \text{ z} > 0.$
  - Una función logística (logistic) (perceptron sigmoide):
    - $Logistic(z) = \frac{1}{1+e^{-z}}.$
  - Una función rectificadora (ReLU):
    - $R(z) = \max(0, z).$

# Funciones de activación:



# El perceptrón como clasificador:

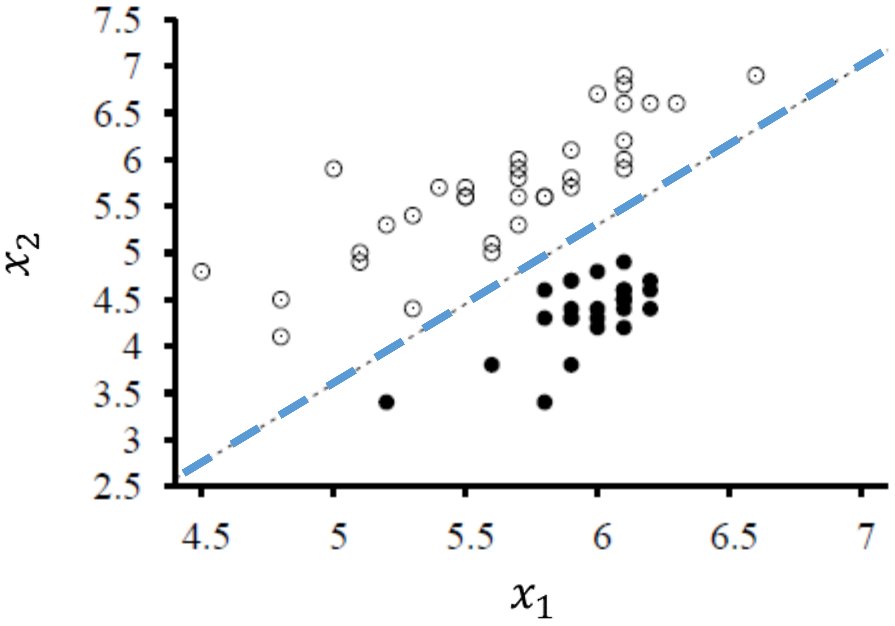
- Clasificadores lineales con hard tresshold (perceptrón):
  - Ejemplo: Terremotos ‘o’ y explosiones nucleares ‘•’’



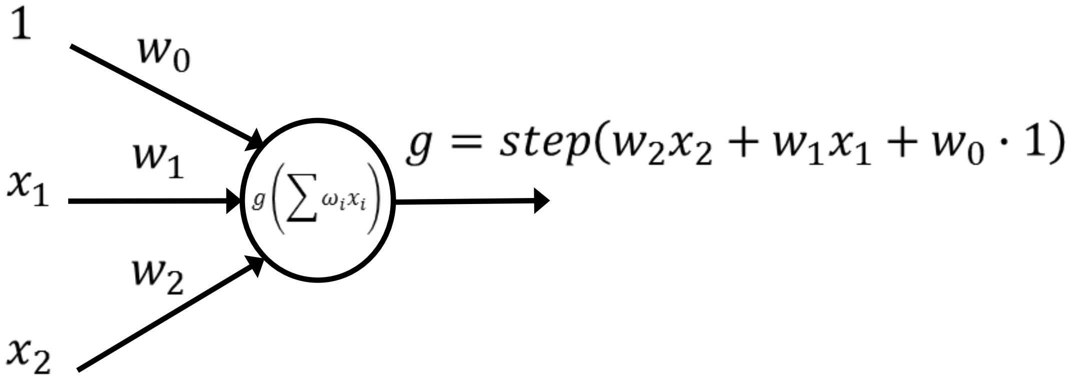

Explosiones y terremotos separables

# El perceptrón como clasificador:

- Clasificadores lineales con hard tresshold (perceptrón):
  - Ejemplo: Terremotos ‘o’ y explosiones nucleares ‘•´´

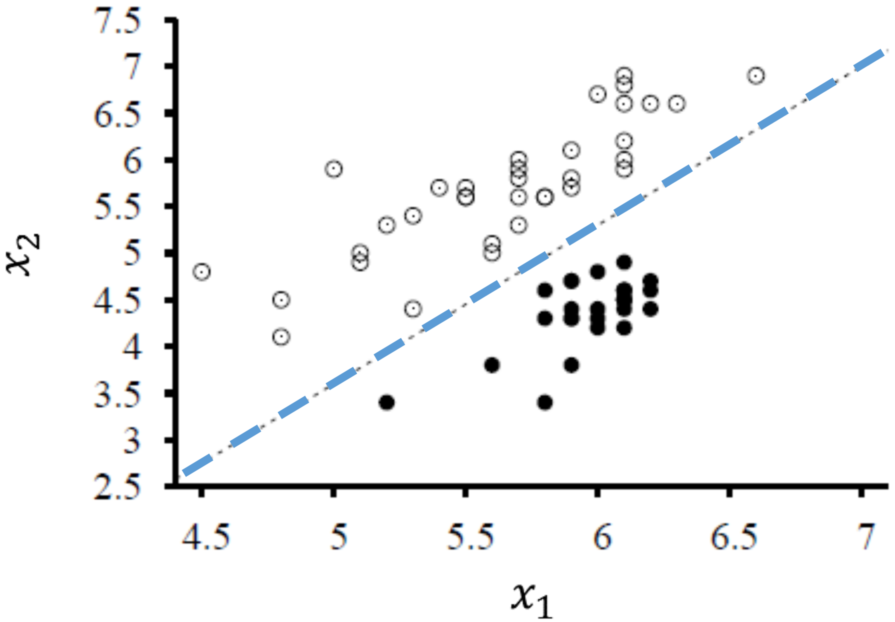



Explosiones y terremotos separables

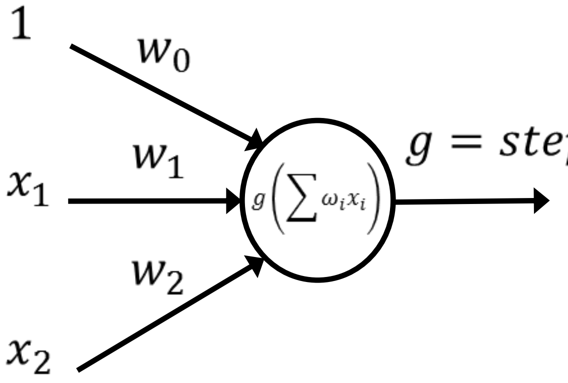


# El perceptrón como clasificador:

- Clasificadores lineales con hard tresshold (perceptrón):
  - Ejemplo: Terremotos ‘o’ y explosiones nucleares ‘•’’




Explosiones y terremotos separables




$$g = step\underbrace{(w_2 x_2 + w_1 x_1 + w_0 \cdot 1)}_{\mathbf{W} \cdot \mathbf{x}^{\mathbf{T}}}$$

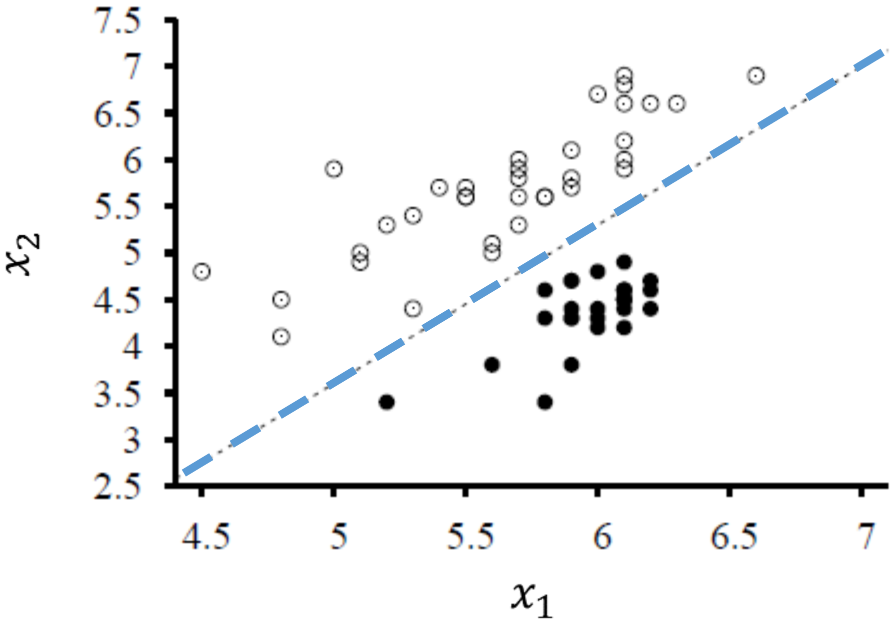
$$\downarrow$$

$$\mathbf{W} = [w_2, w_1, w_0]$$

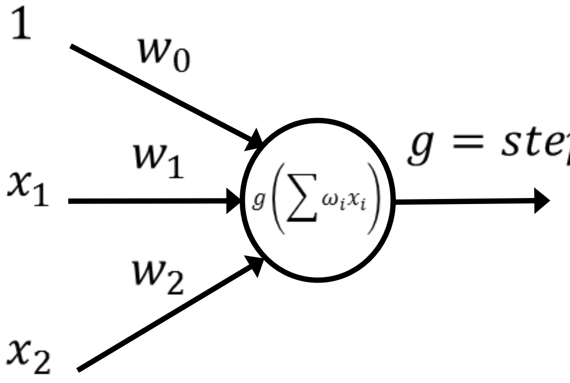
$$\mathbf{x} = [x_2, x_1, (x_0 = 1)]$$

# El perceptrón como clasificador:

- Clasificadores lineales con hard tresshold (perceptrón):
  - Ejemplo: Terremotos ‘o’ y explosiones nucleares ‘•’




Explosiones y terremotos separables




$$g = step(\underbrace{w_2 x_2 + w_1 x_1 + w_0 \cdot 1}_{\mathbf{W} \cdot \mathbf{x}^{\mathrm{T}}})$$

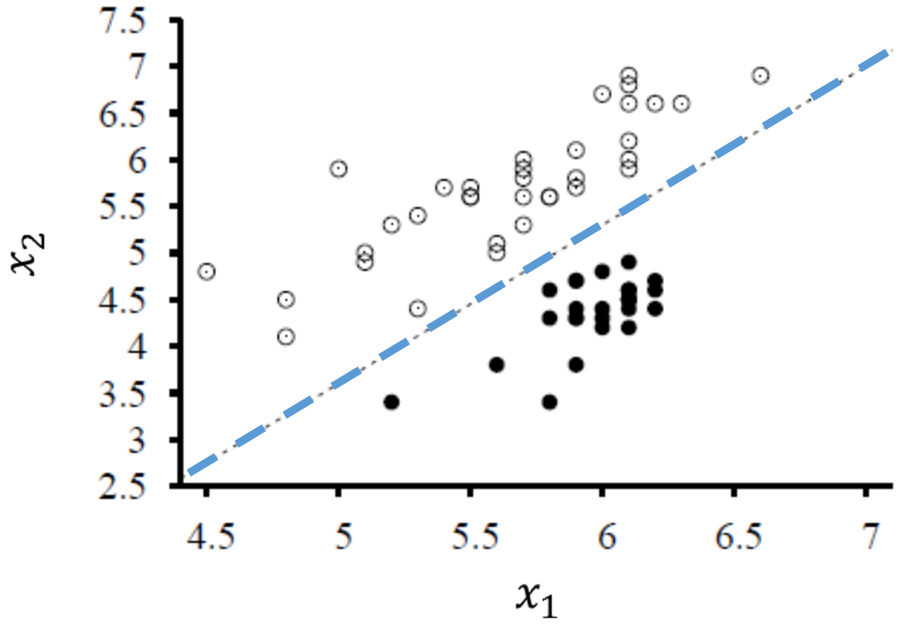
$$\mathbf{W} = [w_2, w_1, w_0]$$

$$\mathbf{x} = [x_2, x_1, (x_0 = 1)]$$

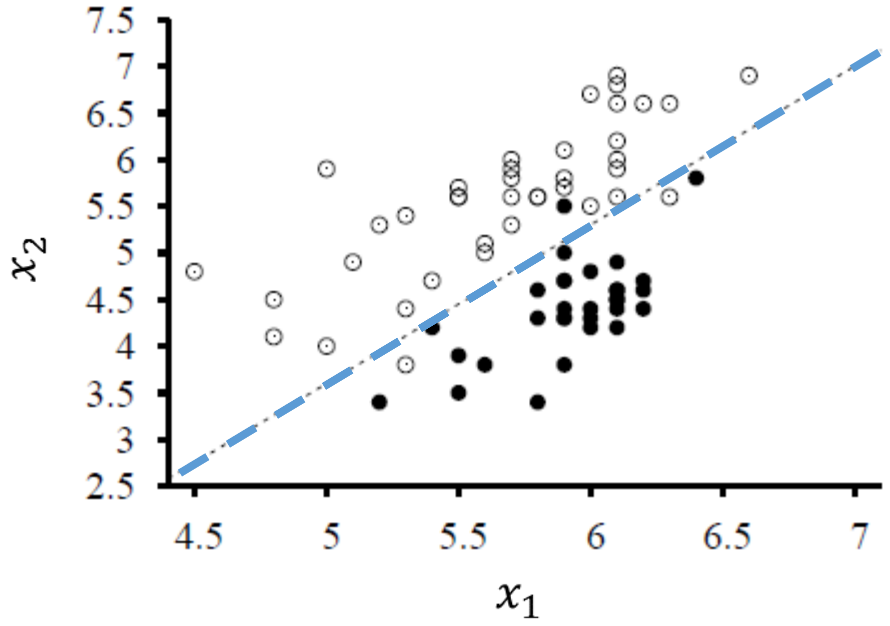
$$x_2 = 1.7x_1 - 4.9 \Rightarrow \underbrace{1.7x_1 - 4.9 - x_2}_{\mathbf{W} \cdot \mathbf{x}^{\mathrm{T}}} = 0 \quad \begin{cases} g = 1 \text{ si } \mathbf{W} \cdot \mathbf{x}^{\mathrm{T}} \geq \mathbf{0} \\ g \leq 0 \text{ si } \mathbf{W} \cdot \mathbf{x}^{\mathrm{T}} < \mathbf{0} \end{cases}$$

# El perceptrón como clasificador:

- Clasificadores lineales con hard tresshold (perceptrón):
  - Ejemplo: Terremotos ‘o’ y explosiones nucleares ‘•’’




Explosiones y terremotos separables

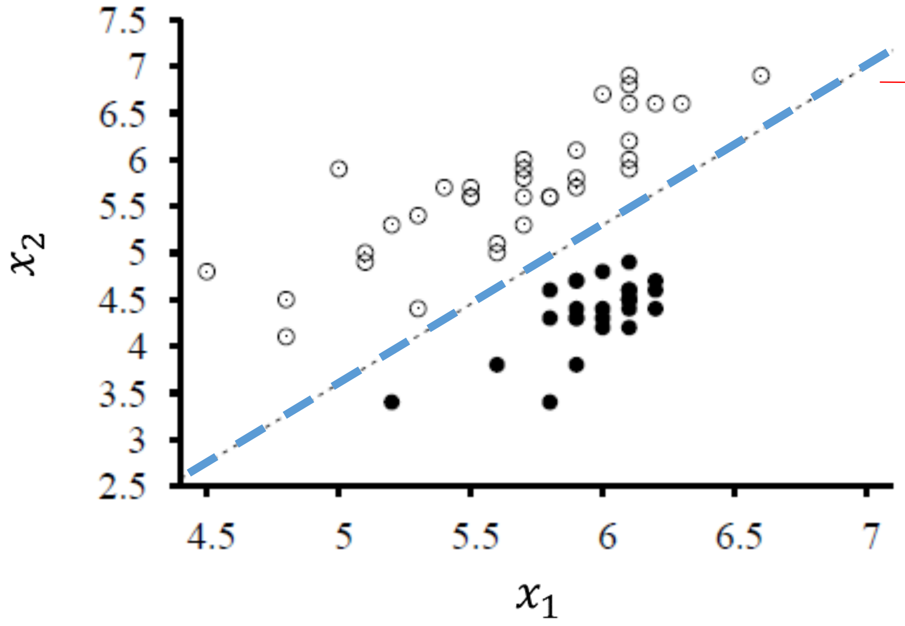



Explosiones y terremotos no separables

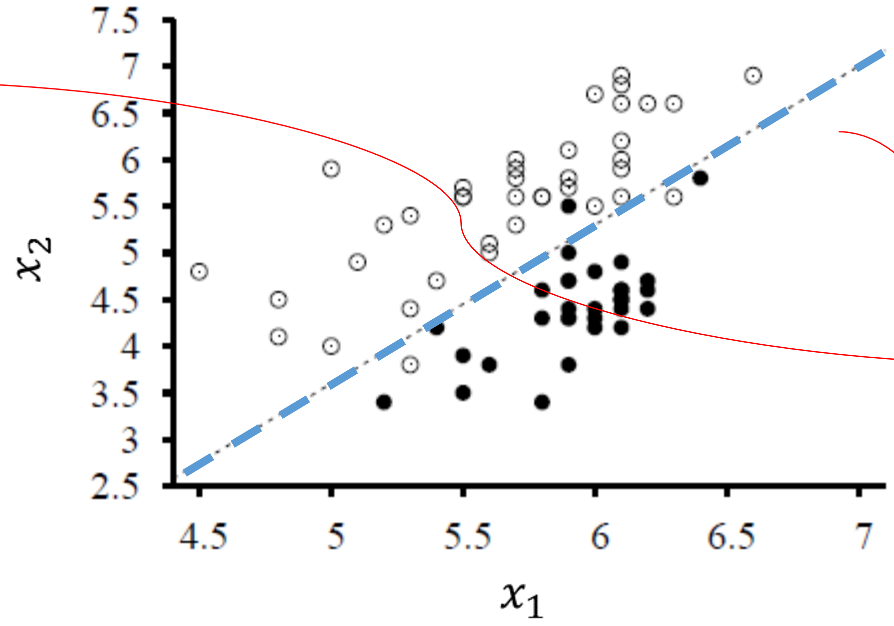
$$x_2 = 1.7x_1 - 4.9 \Rightarrow \underbrace{1.7x_1 - 4.9 - x_2}_{\mathbf{W}\cdot\mathbf{x}^\mathbf{T}}=0 \quad \begin{cases} g = 1 \text{ si } \mathbf{W}\cdot\mathbf{x}^\mathbf{T} \geq \mathbf{0} \\ g \leq 0 \text{ si } \mathbf{W}\cdot\mathbf{x}^\mathbf{T} < \mathbf{0} \end{cases}$$

# El perceptrón como clasificador:

- Clasificadores lineales con hard tresshold (perceptrón):
  - Ejemplo: Terremotos ‘o’ y explosiones nucleares ‘•´’




Explosiones y terremotos separables




Explosiones y terremotos no separables

Frontera de decisión, línea o superficie cuando hay más dimensiones

$$x_2 = 1.7x_1 - 4.9 \Rightarrow \underbrace{1.7x_1 - 4.9 - x_2}_{\mathbf{W}\cdot\mathbf{x}^{\mathbf{T}}}=0 \quad \begin{cases} g = 1 \text{ si } \mathbf{W}\cdot\mathbf{x}^{\mathbf{T}} \geq \mathbf{0} \\ g \leq 0 \text{ si } \mathbf{W}\cdot\mathbf{x}^{\mathbf{T}} < \mathbf{0} \end{cases}$$

# Redes neuronales artificiales:

- Ejemplo: dada una neurona artificial con 2 entradas $x_1 = 0.7$ y $x_2 = 0.9$ con los pesos asociados $\omega_1 = 0.8$ y $\omega_2$=0.4 y una salida. ¿Cuándo va a activar la salida si aplicamos las funciones de activación $step()$, $Logistic()$ y $R()$?
  - Función step():
    - $Step(\mathrm{x} \cdot \mathrm{W}^{\mathrm{T}}) = 1 \text{ si } \mathrm{z} > 0 \Rightarrow \mathrm{x} \cdot \mathrm{W}^{\mathrm{T}} = 0.7 \cdot 0.8 + 0.9 \cdot 0.4 = 0.92. \; Step(\mathrm{x} \cdot \mathrm{W}^{\mathrm{T}}) = 1$
  - Función sigmoid():
    - $Logistic(\mathrm{x} \cdot \mathrm{W}^{\mathrm{T}}) = \frac{1}{1+e^{-\mathrm{x} \cdot \mathrm{W}^{\mathrm{T}}}} = \frac{1}{1+e^{-0.92}} = 0.715.$
  - Función ReLU():
    - $R(z) = \max(0, \mathrm{x} \cdot \mathrm{W}^{\mathrm{T}}) = 0.92.$

# Redes neuronales artificiales:

- Ejemplo: dada una neurona artificial con 2 entradas $x_1 = 0.7$ y $x_2 = 0.9$ con los pesos asociados $\omega_1 = 0.8$ y $\omega_2$=0.4 y una salida. ¿Cuándo va a activar la salida si aplicamos las funciones de activación $Step()$, $Logistic()$ y $R()$?
  - Función $Step()$:
    - $Step(\mathbf{x} \cdot \mathbf{W}^{\mathrm{T}}) = 1 \text{ si z} > 0 \Rightarrow \mathbf{x} \cdot \mathbf{W}^{\mathrm{T}} = 0.7 \cdot 0.8 + 0.9 \cdot 0.4 = 0.92 \Rightarrow Step(0.92) = 1$
  - Función sigmoid():
    - $Logistic(\mathrm{x} \cdot \mathrm{W}^{\mathrm{T}}) = \frac{1}{1+e^{-\mathrm{x} \cdot \mathrm{W}^{\mathrm{T}}}} = \frac{1}{1+e^{-0.92}} = 0.715.$
  - Función ReLU():
    - $R(z) = \max(0, \mathrm{x} \cdot \mathrm{W}^{\mathrm{T}}) = 0.92.$

# Redes neuronales artificiales:

- Ejemplo: dada una neurona artificial con 2 entradas $x_1 = 0.7$ y $x_2 = 0.9$ con los pesos asociados $\omega_1 = 0.8$ y $\omega_2$=0.4 y una salida. ¿Cuándo va a activar la salida si aplicamos las funciones de activación $Step()$, $Logistic()$ y $R()$?
  - Función $Step()$:
    - $Step(\mathbf{x} \cdot \mathbf{W}^{\mathbf{T}}) = 1 \text{ si z} > 0 \Rightarrow \mathbf{x} \cdot \mathbf{W}^{\mathbf{T}} = 0.7 \cdot 0.8 + 0.9 \cdot 0.4 = 0.92 \Rightarrow Step(0.92) = 1$
  - Función $Logistic()$:
    - $Logistic(\mathbf{x} \cdot \mathbf{W}^{\mathbf{T}}) = \frac{1}{1+e^{-\mathbf{x} \cdot \mathbf{W}^{\mathbf{T}}}} = \frac{1}{1+e^{-0.92}} = 0.715.$
  - Función ReLU():
    - $R(z) = \max(0, \mathbf{x} \cdot \mathbf{W}^{\mathbf{T}}) = 0.92.$

# Redes neuronales artificiales:

- Ejemplo: dada una neurona artificial con 2 entradas $x_1 = 0.7$ y $x_2 = 0.9$ con los pesos asociados $\omega_1 = 0.8$ y $\omega_2$=0.4 y una salida. ¿Cuándo va a activar la salida si aplicamos las funciones de activación $Step()$, $Logistic()$ y $R()$?
  - Función $Step()$:
    - $Step(\mathbf{x} \cdot \mathbf{W}^{\mathbf{T}}) = 1 \text{ si z} > 0 \Rightarrow \mathbf{x} \cdot \mathbf{W}^{\mathbf{T}} = 0.7 \cdot 0.8 + 0.9 \cdot 0.4 = 0.92 \Rightarrow Step(0.92) = 1$
  - Función $Logistic()$:
    - $Logistic(\mathbf{x} \cdot \mathbf{W}^{\mathbf{T}}) = \frac{1}{1+e^{-\mathbf{x} \cdot \mathbf{W}^{\mathbf{T}}}} = \frac{1}{1+e^{-0.92}} = 0.715.$
  - Función $R()$:
    - $R(z) = \max(0, \mathbf{x} \cdot \mathbf{W}^{\mathbf{T}}) = 0.92.$

# Bibliografía:


1. Redes Neuronales Artificiales. Fundamentos, modelos y aplicaciones. José Ramón Hilera y Victor José Martinez Hernando. Editorial Rama.
2. Artificial Intelligence A Modern Approach. Stuart Russell and Peter Norvig. Third Edition. Editorial Pearson.
3. Artificial Intelligence Illuminated. Ben Coppin. First Edition. Editorial Jones and Bartlett.