

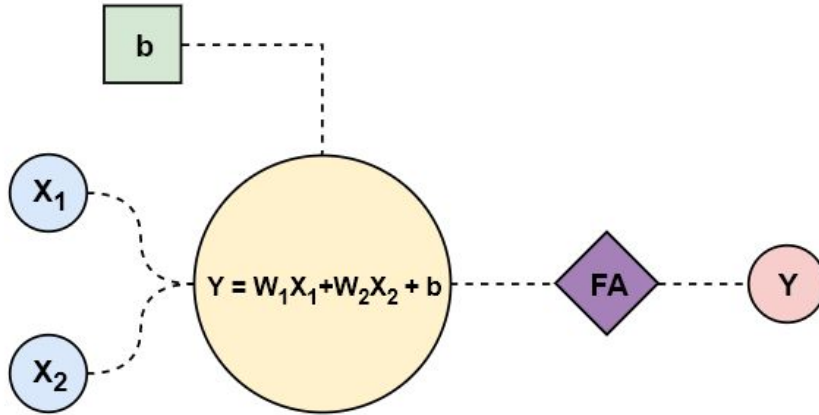
Introducción a Redes Neuronales

Edgar Rangel
Python Bucaramanga
Mayo 30, 2019

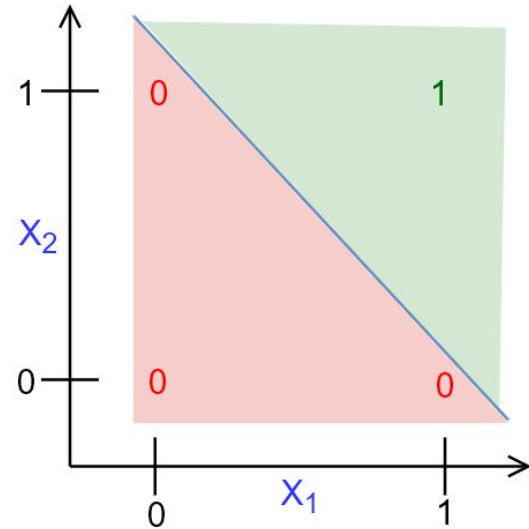


Compuerta AND

Vamos a modelar una compuerta AND

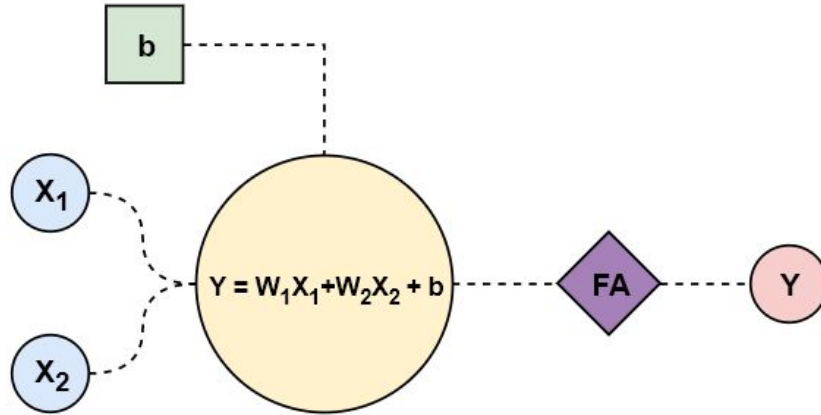


$$Y = 1X_1 + 1X_2 - 1$$

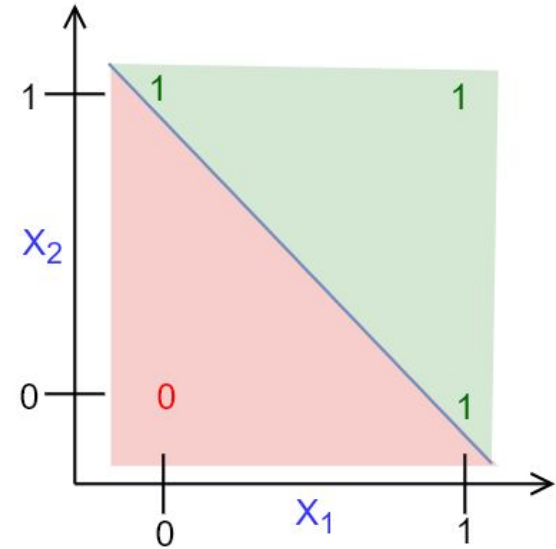


Compuerta OR

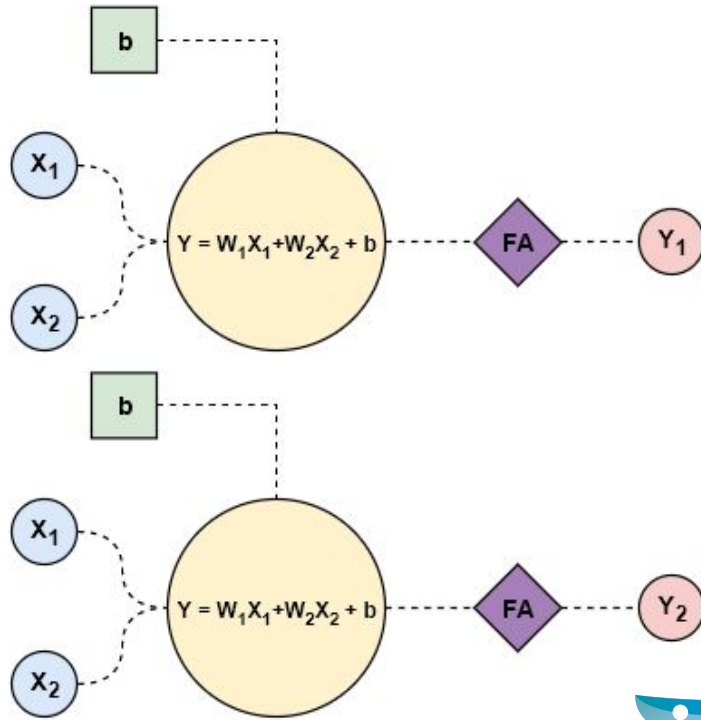
Vamos a modelar una compuerta OR



$$Y = 1X_1 + 1X_2 + 0$$

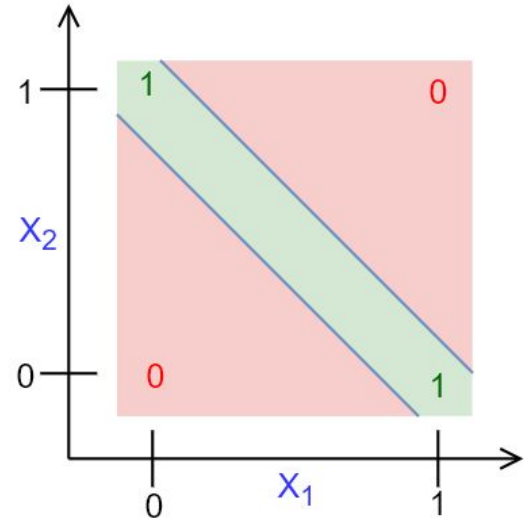


Compuerta XOR



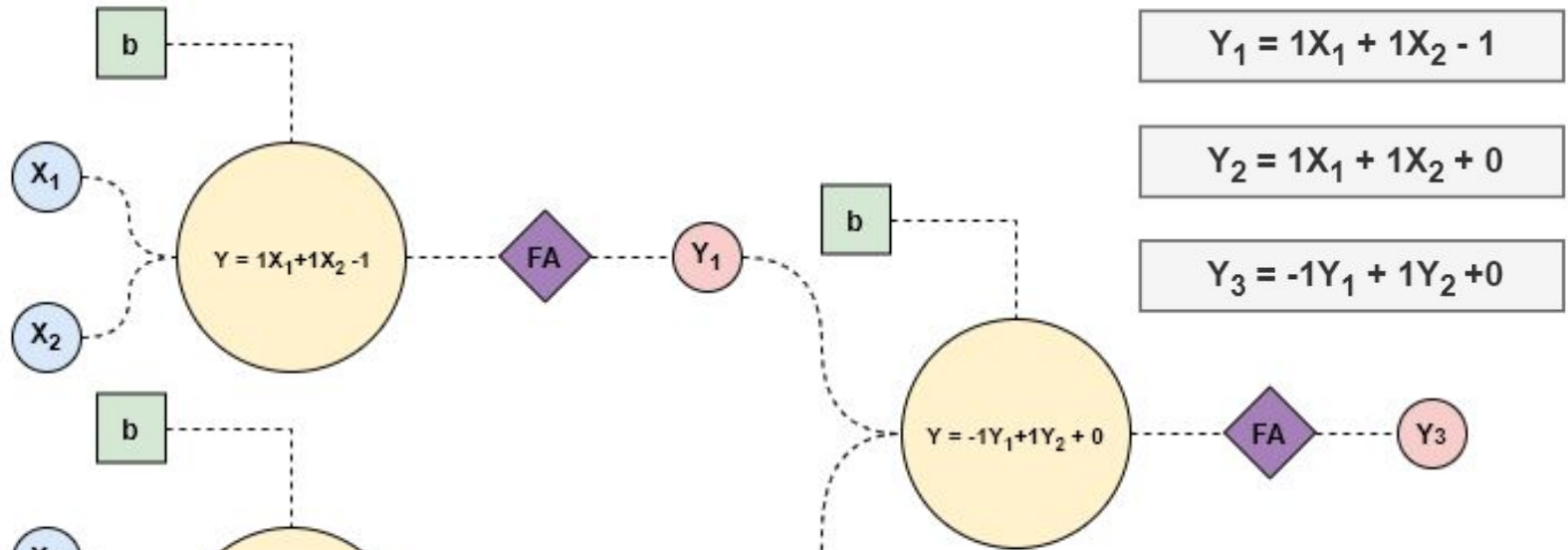
$$Y_1 = 1X_1 + 1X_2 - 1$$

$$Y_2 = 1X_1 + 1X_2 + 0$$

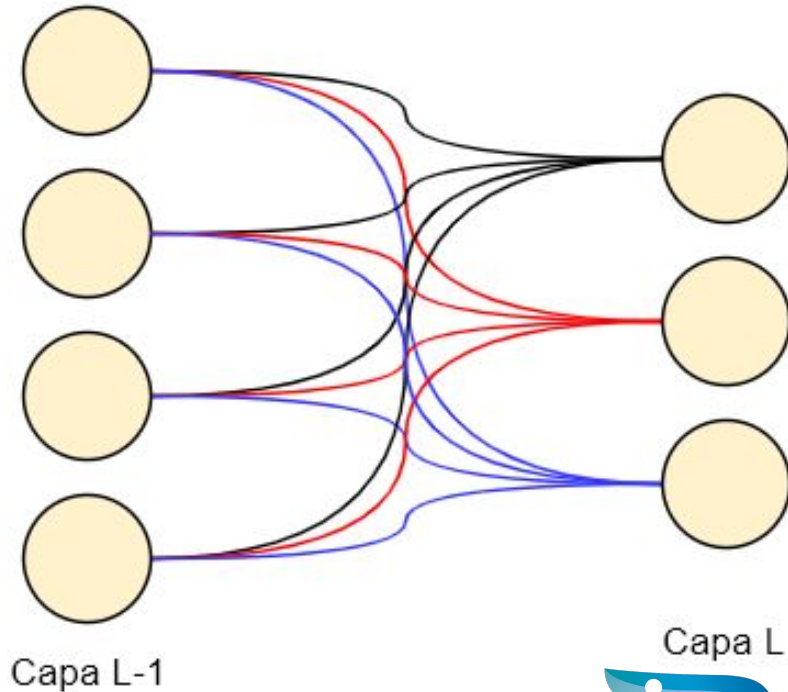


Relacionar Neuronas

¿Cómo relacionarlas?



Nuestra propia Red Neuronal



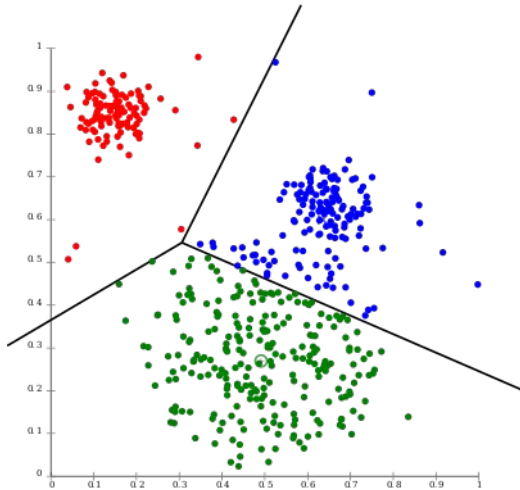
$$\begin{bmatrix} W_{11} & W_{12} & \dots & W_{1n} \\ W_{21} & W_{22} & \dots & W_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ W_{m1} & W_{m2} & \dots & W_{mn} \end{bmatrix}$$

$W_{[\text{Entrada}][\text{Neurona}]}$

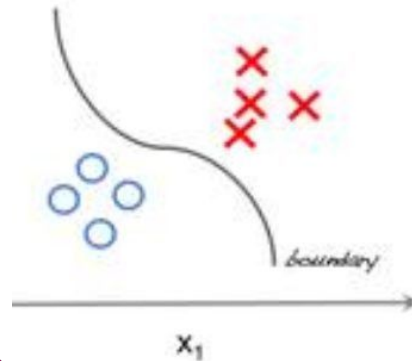
Aprendizaje Automático

Datos Reales

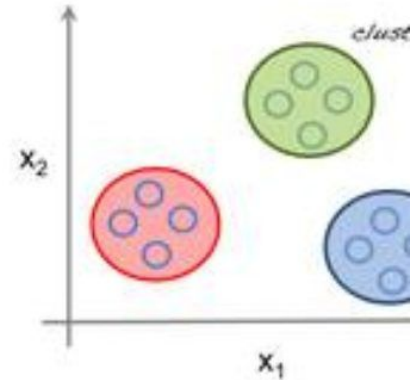
En muchos casos para que nuestra Red Neuronal aprende necesitamos información y etiquetas de esa información para entregarlos a la Red (Aprendizaje supervisado) y otros casos donde simplemente le damos la información y que la Red aprenda por sí misma (Aprendizaje No Supervisado)



Supervised learning



Unsupervised learning

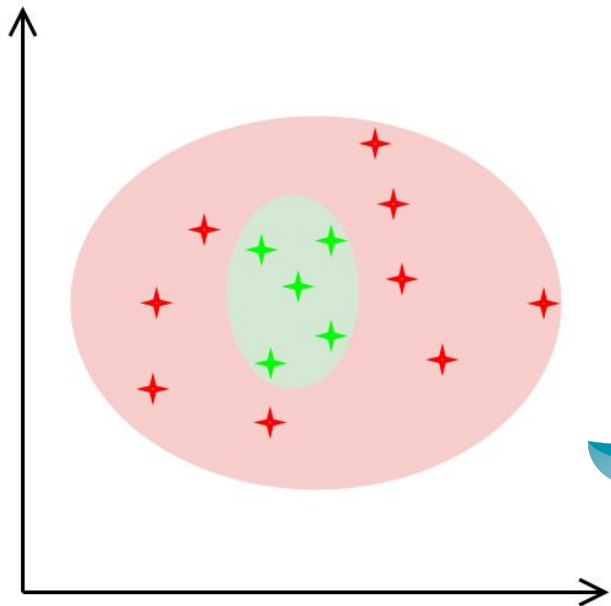


Error o función de costo

Necesitamos decirle a nuestra red si esta bien o está equivocada con los valores que nos calcula pero... ¿Como le decimos esto?

Un ejemplo clásico es comparar los resultados que nos entrega con los resultados que ya conocemos.

$$|Y_{\text{calculado}} - Y_{\text{correcto}}|$$



Regla de la Cadena

Es una regla matemática que se aplica a las derivadas en una función de tal forma que si esta función depende o está compuesta de otra función que tiene otras dependencias se pueda derivar



$$\begin{aligned}f(x) &= 20 \cdot g(x) \\ g(x) &= x^2\end{aligned}$$

$$\frac{df(x)}{dg(x)} = 20 \quad \frac{dg(x)}{dx} = 2x$$

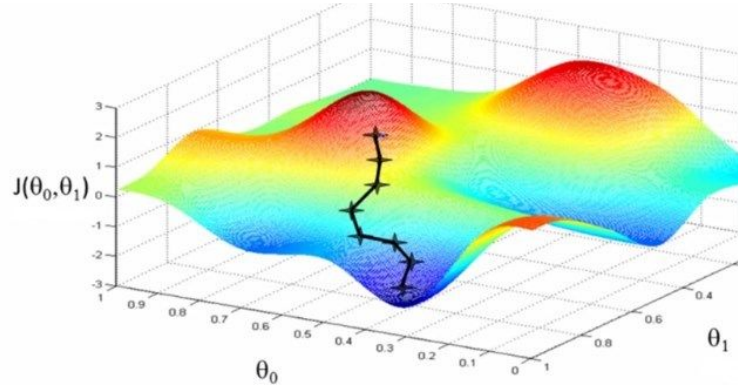
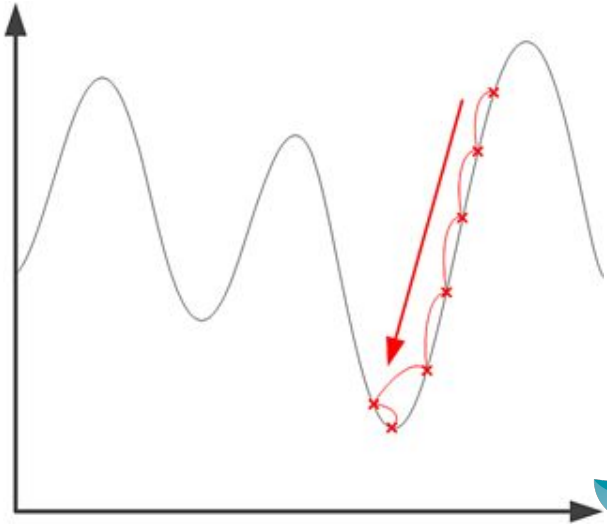
$$\frac{df(x)}{dx} = \frac{df(x)}{dg(x)} \cdot \frac{dg(x)}{dx}$$

$$\frac{df(x)}{dx} = 20 \cdot 2x$$



Descenso del gradiente

Un algoritmo clave dentro del Machine Learning donde el ejemplo clásico es el hecho de como llegar al punto más bajo de la colina únicamente con la sensación del tacto del pie
¿Como lo harías?



$$X = X - dX$$

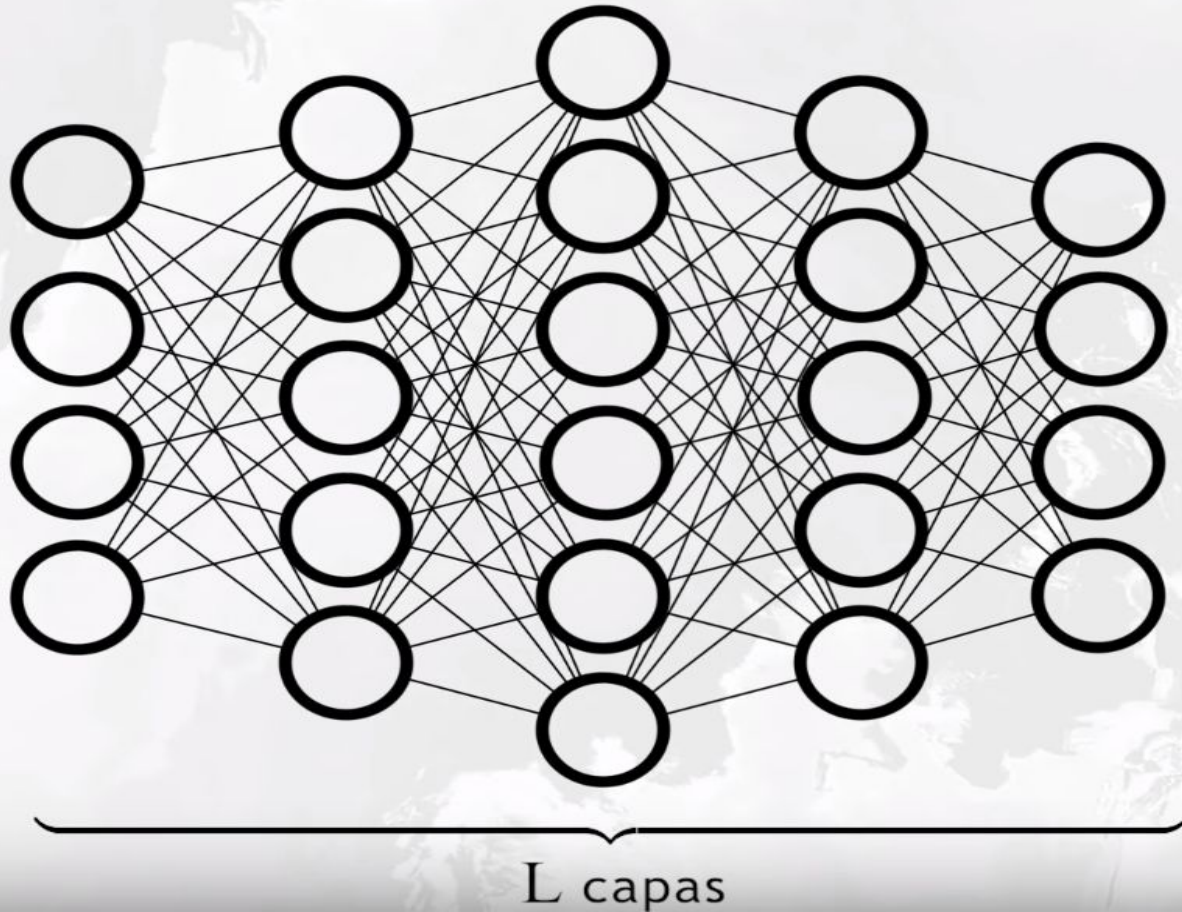
Backpropagation o Retropropagación

¿COMO VARIA EL COSTE ANTE
UN CAMBIO DEL PARAMETRO w ?

$$\frac{\partial C}{\partial w}$$

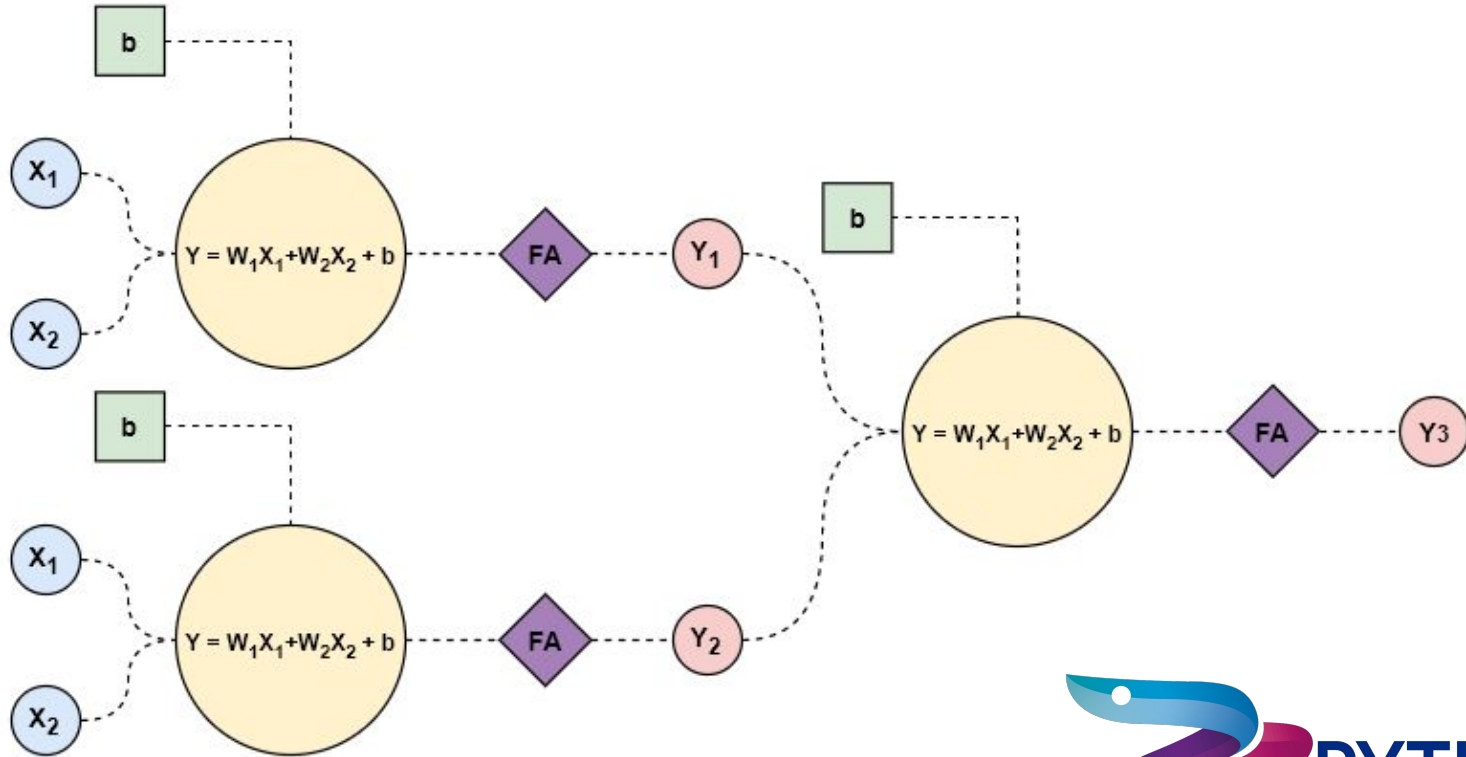


Backpropagation o Retropropagación



$$\frac{\partial C}{\partial b^L} \quad \frac{\partial C}{\partial w^L}$$

Backpropagation o Retropropagación



Backpropagation o Retropropagación

$$\frac{\partial C}{\partial w^L} = \frac{\partial C}{\partial a^L} \cdot \frac{\partial a^L}{\partial z^L} \cdot \frac{\partial z^L}{\partial w^L}$$

$$\frac{\partial C}{\partial b^L} = \frac{\partial C}{\partial a^L} \cdot \frac{\partial a^L}{\partial z^L} \cdot \frac{\partial z^L}{\partial b^L}$$

$$Z^L = W^L a^{L-1} + b^L \quad C(a^L(Z^L))$$

¿Cómo transmitir el error?

$$\begin{array}{c}
 \delta^L \\
 \frac{\partial C}{\partial w^{L-1}} = \frac{\partial C}{\partial a^L} \cdot \frac{\partial a^L}{\partial z^L} \cdot \frac{\partial z^L}{\partial a^{L-1}} \cdot \frac{\partial a^{L-1}}{\partial z^{L-1}} \cdot \frac{\partial z^{L-1}}{\partial w^{L-1}} \\
 \frac{\partial C}{\partial b^{L-1}} = \frac{\partial C}{\partial a^L} \cdot \frac{\partial a^L}{\partial z^L} \cdot \frac{\partial z^L}{\partial a^{L-1}} \cdot \frac{\partial a^{L-1}}{\partial z^{L-1}} \cdot \frac{\partial z^{L-1}}{\partial b^{L-1}}
 \end{array}$$

δ^L W^L DERIVADA DE LA FUNCION DE ACT. a^{L-2}

APLICAMOS LA CHAIN RULE A ESTA COMPOSICION

$$C(a^L(W^L a^{L-1}(W^{L-1} a^{L-2} + b^{L-1}) + b^L))$$

¿Cómo transmitir el error?

$$\begin{array}{c}
 \delta^L \qquad \qquad \qquad W^L \qquad \text{DERIVADA DE LA FUNCIÓN DE ACT.} \qquad a^{L-2} \\
 \\
 \frac{\partial C}{\partial w^{L-1}} = \frac{\partial C}{\partial a^L} \cdot \frac{\partial a^L}{\partial z^L} \cdot \frac{\partial z^L}{\partial a^{L-1}} \cdot \frac{\partial a^{L-1}}{\partial z^{L-1}} \cdot \frac{\partial z^{L-1}}{\partial w^{L-1}} \\
 \\
 \frac{\partial C}{\partial b^{L-1}} = \frac{\partial C}{\partial a^L} \cdot \frac{\partial a^L}{\partial z^L} \cdot \frac{\partial z^L}{\partial a^{L-1}} \cdot \frac{\partial a^{L-1}}{\partial z^{L-1}} \cdot \frac{\partial z^{L-1}}{\partial b^{L-1}} \\
 \\
 \frac{\partial C}{\partial z^{L-1}} = \delta^{L-1}
 \end{array}$$

¿Cómo transmitir el error?

1. COMPUTO DEL ERROR DE LA ULTIMA CAPA

$$\delta^L = \frac{\partial C}{\partial a^L} \cdot \frac{\partial a^L}{\partial z^L}$$

2. RETROPROPAGAMOS EL ERROR A LA CAPA ANTERIOR

$$\delta^{l-1} = W^l \delta^l \cdot \frac{\partial a^{l-1}}{\partial z^{l-1}}$$

3. CALCULAMOS LAS DERIVADAS DE LA CAPA USANDO EL ERROR

$$\frac{\partial C}{\partial b^{l-1}} = \delta^{l-1} \quad \frac{\partial C}{\partial w^{l-1}} = \delta^{l-1} a^{l-2}$$

Gracias por la Atención