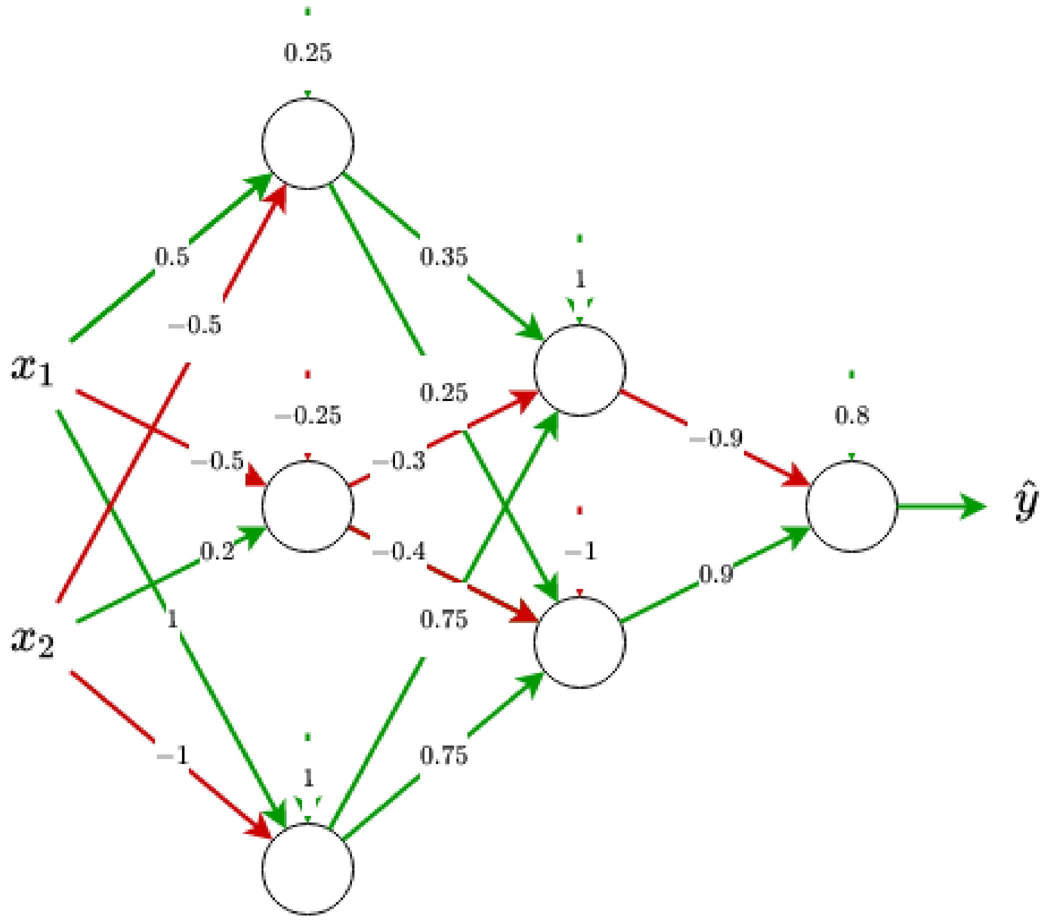
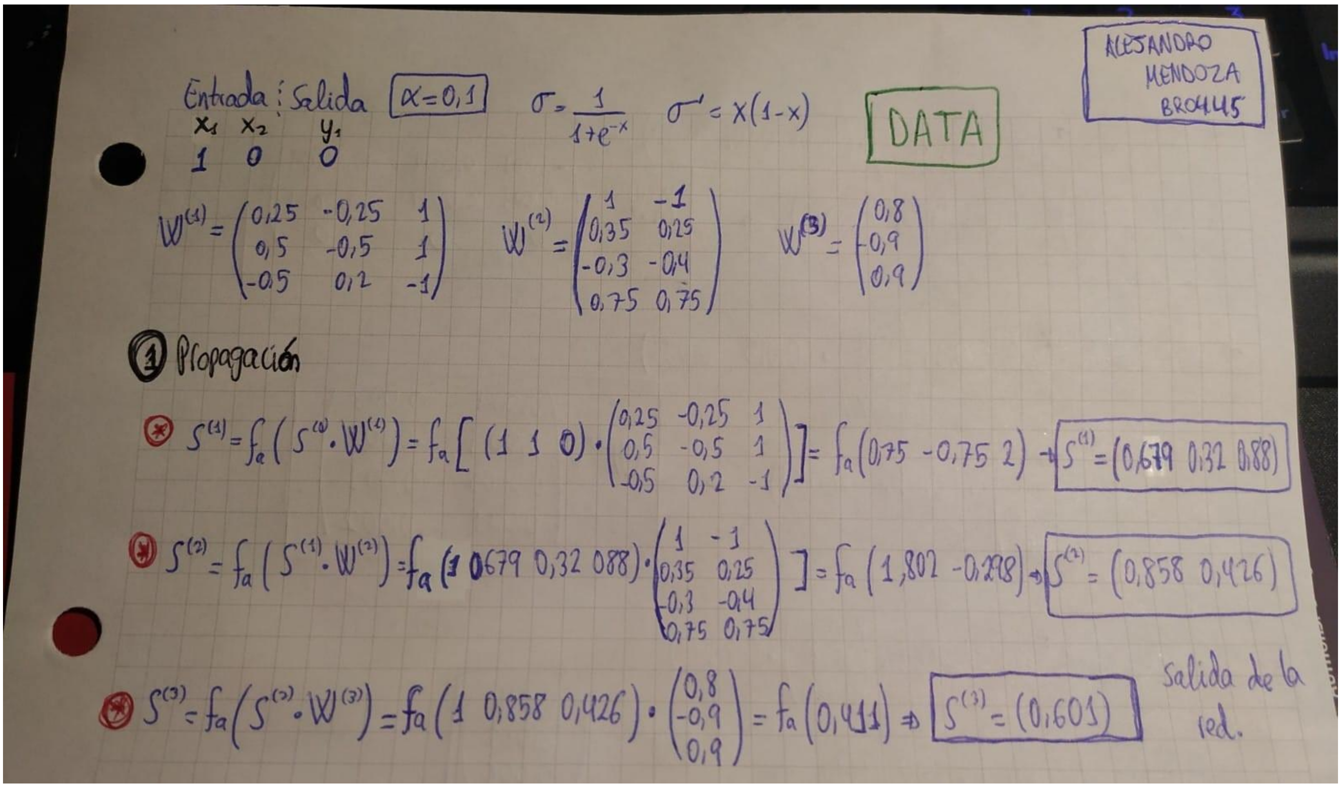
PERCEPTRÓN MULTICAPA – IA

Tenemos la siguiente red de neuronas:



Empezamos analizando los datos de partida, así como los vectores de pesos de cada una de las capas, para después hacer la propagación y acabar sacando la salida de la red:



ALEJANDRO MENDOZA BR0445

Entrada : Salida $\boxed{\alpha = 0{,}1}$ $\sigma = \frac{1}{1+e^{-x}}$ $\sigma' = x(1-x)$ DATA

x_1	x_2	y_1
1	0	0

$$W^{(1)} = \begin{pmatrix} 0{,}25 & -0{,}25 & 1 \\ 0{,}5 & -0{,}5 & 1 \\ -0{,}5 & 0{,}2 & -1 \end{pmatrix} \quad W^{(2)} = \begin{pmatrix} 1 & -1 \\ 0{,}35 & 0{,}25 \\ -0{,}3 & -0{,}4 \\ 0{,}75 & 0{,}75 \end{pmatrix} \quad W^{(3)} = \begin{pmatrix} 0{,}8 \\ -0{,}9 \\ 0{,}9 \end{pmatrix}$$

1 Propagación

$$S^{(1)} = f_a\left(S^{(0)} \cdot W^{(1)}\right) = f_a\left[\begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0{,}25 & -0{,}25 & 1 \\ 0{,}5 & -0{,}5 & 1 \\ -0{,}5 & 0{,}2 & -1 \end{pmatrix}\right] = f_a\begin{pmatrix} 0{,}75 & -0{,}75 & 2 \end{pmatrix} \rightarrow \boxed{S^{(1)} = \begin{pmatrix} 0{,}679 & 0{,}32 & 0{,}88 \end{pmatrix}}$$

$$S^{(2)} = f_a\left(S^{(1)} \cdot W^{(2)}\right) = f_a\left[\begin{pmatrix} 1 & 0{,}679 & 0{,}32 & 0{,}88 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 0{,}35 & 0{,}25 \\ -0{,}3 & -0{,}4 \\ 0{,}75 & 0{,}75 \end{pmatrix}\right] = f_a\begin{pmatrix} 1{,}802 & -0{,}278 \end{pmatrix} \rightarrow \boxed{S^{(2)} = \begin{pmatrix} 0{,}858 & 0{,}426 \end{pmatrix}}$$

$$S^{(3)} = f_a\left(S^{(2)} \cdot W^{(3)}\right) = f_a\begin{pmatrix} 1 & 0{,}858 & 0{,}426 \end{pmatrix} \cdot \begin{pmatrix} 0{,}8 \\ -0{,}9 \\ 0{,}9 \end{pmatrix} = f_a\begin{pmatrix} 0{,}411 \end{pmatrix} \rightarrow \boxed{S^{(3)} = \begin{pmatrix} 0{,}601 \end{pmatrix}}$$

Salida de la red.

Para realizar la retropropagación, primero calculamos para cada capa, con la ayuda de la salida de la capa posterior, el error cometido, a excepción de la última capa, en la cual calculamos el error basándonos en la salida deseada y en la salida de la red:

② Retropropagación

$$\delta^{(3)} = (d-s) \odot f'(s^{(3)}) = (0 - 0{,}601) \odot [(0{,}601) \odot (1-(0{,}601))]$$

$$\delta^{(3)} = (-0{,}601) \odot [(0{,}601) \odot (0{,}399)] = (-0{,}601) \odot (0{,}239) \Rightarrow \boxed{\delta^{(3)} = (-0{,}143)}$$

$$\delta^{(2)} = (\delta^{(3)} \cdot W^{(3)T}) \cdot f'(s^{(2)}) = [(-0{,}143) \cdot (-0{,}9 \;\; 0{,}9)] \odot [(0{,}858 \;\; 0{,}426) \odot (1-(0{,}858 \;\; 0{,}426))]$$

$$\delta^{(2)} = (0{,}128 \;\; -0{,}128) \odot [(0{,}858 \;\; 0{,}426) \odot (0{,}142 \;\; 0{,}574)]$$

$$\delta^{(2)} = (0{,}128 \;\; -0{,}128) \odot (0{,}121 \;\; 0{,}244) \Rightarrow \boxed{\delta^{(2)} = (0{,}015 \;\; -0{,}031)}$$

$$\delta^{(1)} = (\delta^{(2)} \cdot W^{(2)T}) \cdot f'(s^{(1)}) = \left[(0{,}015 \;\; -0{,}031) \cdot \begin{pmatrix} 0{,}35 & -0{,}3 & 0{,}75 \\ 0{,}25 & -0{,}4 & 0{,}75 \end{pmatrix}\right] \odot [(0{,}679 \;\; 0{,}32 \;\; 0{,}88) \odot (1-(0{,}679 \;\; 0{,}32 \;\; 0{,}88))]$$

$$\delta^{(1)} = (-0{,}003 \;\; +0{,}008 \;\; -0{,}012) \odot [(0{,}679 \;\; 0{,}32 \;\; 0{,}88) \odot (0{,}321 \;\; 0{,}68 \;\; 0{,}12)]$$

$$\delta^{(1)} = (-0{,}003 \;\; 0{,}008 \;\; -0{,}012) \odot (0{,}217 \;\; 0{,}217 \;\; 0{,}105) \Rightarrow \boxed{\delta^{(1)} = (-0{,}001 \;\; 0{,}002 \;\; -0{,}001)}$$

$$\Longrightarrow$$

Obtenidos los errores de cada capa, calcularemos las variaciones de pesos de cada capa, para saber cuánto tienen que cambiar para el siguiente intervalo de tiempo:

$$\Delta W^{(1)} = \alpha\, S^{(0)T} \cdot \delta^{(1)} = 0{,}1 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -0{,}001 & 0{,}002 & -0{,}001 \end{pmatrix}$$

$$\Delta W^{(1)} = \begin{pmatrix} 0{,}1 \\ 0{,}1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -0{,}001 & 0{,}002 & -0{,}001 \end{pmatrix} \Rightarrow \boxed{\Delta W^{(1)} = \begin{pmatrix} -0{,}0001 & 0{,}0002 & -0{,}0001 \\ -0{,}0001 & 0{,}0002 & -0{,}0001 \\ 0 & 0 & 0 \end{pmatrix}}$$

$$\Delta W^{(2)} = \alpha \cdot S^{(1)T} \cdot \delta^{(2)} = 0{,}1 \cdot \begin{pmatrix} 1 \\ 0{,}679 \\ 0{,}32 \\ 0{,}88 \end{pmatrix} \cdot \begin{pmatrix} 0{,}015 & -0{,}031 \end{pmatrix}$$

$$\Delta W^{(2)} = \begin{pmatrix} 0{,}1 \\ 0{,}067 \\ 0{,}032 \\ 0{,}088 \end{pmatrix} \cdot \begin{pmatrix} 0{,}015 & -0{,}031 \end{pmatrix} \Rightarrow \boxed{\Delta W^{(2)} = \begin{pmatrix} 0{,}001 & -0{,}003 \\ 0{,}001 & -0{,}002 \\ 0{,}0005 & -0{,}001 \\ 0{,}001 & -0{,}002 \end{pmatrix}}$$

$$\Delta W^{(3)} = \alpha \cdot S^{(2)T} \cdot \delta^{(3)} = 0{,}1 \cdot \begin{pmatrix} 1 \\ 0{,}858 \\ 0{,}426 \end{pmatrix} \cdot (-0{,}143)$$

$$\Delta W^{(3)} = \begin{pmatrix} 0{,}1 \\ 0{,}085 \\ 0{,}042 \end{pmatrix} \cdot (-0{,}143) \Rightarrow \boxed{\Delta W^{(3)} = \begin{pmatrix} -0{,}014 \\ -0{,}012 \\ -0{,}006 \end{pmatrix}}$$

Por último, aplicaremos cada variación a su peso correspondiente, para obtener los nuevos pesos:

$$* \; W^{(1)}_{(t+1)} = W^{(1)}_{(t)} + \Delta W^{(1)} = \begin{pmatrix} 0,25 & -0,25 & 1 \\ 0,5 & -0,5 & 1 \\ -0,5 & 0,2 & -1 \end{pmatrix} + \begin{pmatrix} -0,0001 & 0,0002 & -0,0001 \\ -0,0001 & 0,0002 & -0,0001 \\ 0 & 0 & 0 \end{pmatrix} \rightarrow \boxed{W^{(1)}_{(t+1)} = \begin{pmatrix} 0,249 & -0,249 & 0,999 \\ 0,499 & -0,499 & 0,999 \\ -0,5 & 0,2 & -1 \end{pmatrix}}$$

$$* \; W^{(2)}_{(t+1)} = W^{(2)}_{(t)} + \Delta W^{(2)} = \begin{pmatrix} 1 & -1 \\ 0,35 & 0,25 \\ -0,3 & -0,4 \\ 0,75 & 0,75 \end{pmatrix} + \begin{pmatrix} 0,001 & -0,003 \\ 0,001 & -0,002 \\ 0,005 & -0,001 \\ 0,001 & -0,002 \end{pmatrix} \rightarrow \boxed{W^{(2)}_{(t+1)} = \begin{pmatrix} 1,001 & -1,001 \\ 0,351 & 0,247 \\ -0,299 & -0,401 \\ 0,751 & 0,747 \end{pmatrix}}$$

$$* \; W^{(3)}_{(t+1)} = W^{(3)}_{(t)} + \Delta W^{(3)} = \begin{pmatrix} 0,8 \\ -0,9 \\ 0,9 \end{pmatrix} + \begin{pmatrix} -0,014 \\ -0,012 \\ -0,006 \end{pmatrix} \Rightarrow \boxed{W^{(3)}_{(t+1)} = \begin{pmatrix} 0,085 \\ 0,072 \\ 0,036 \end{pmatrix}}$$