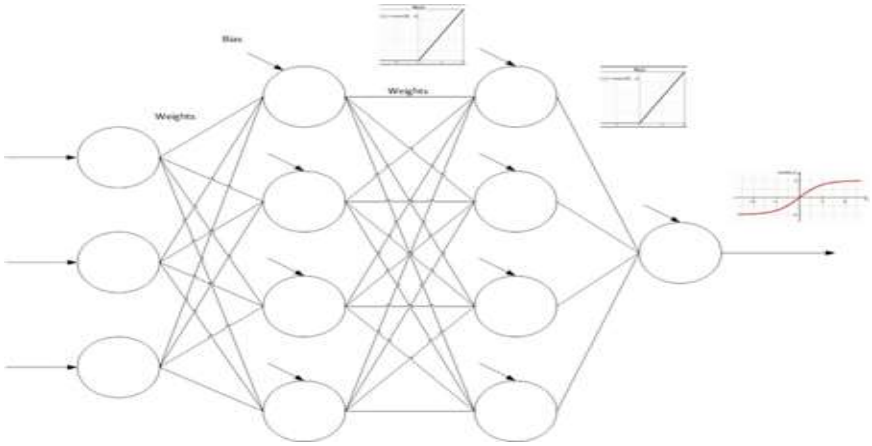
SOLUCION PREGUNTA 1

$$Xs = \begin{bmatrix} 2.5 & 3.5 & -0.5 \\ 4 & -1 & 0.5 \\ 0.5 & 1.5 & 1 \\ 3 & 2 & -1.5 \end{bmatrix}$$

$$Ys = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$



Estructura

Entrada:
3 Neuronas que reciben los valores de entrada (en este caso, 3 características o valores).

Capa Oculta 1:
4 Neuronas.
Función de activación ReLU: Aplica la activación ReLU

Capa Oculta 2:
4 Neuronas.
Función de activación ReLU.

Capa de Salida:
1 Neurona
Función de activación Tangente Hiperbólica (Tanh): (-1, 1)

Paso 1
Inicialización de Pesos y Sesgos

W $\sim \mathcal{U}\left(-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\right)$ Se toma como referencia la inicialización de Xavier

Capa Oculta 1 (4 neuronas, 3 entradas)

Los pesos de la primera capa W1 deben moverse en el rango $(-0.577, 0.577)$

Pesos W1 =

$$\begin{bmatrix} 0.1 & 0.5 & -0.2 \\ -0.3 & 0.4 & 0.1 \\ 0.2 & -0.1 & -0.5 \\ 0.05 & 0.3 & 0.2 \end{bmatrix}$$

Sesgo b1 =

$$\begin{bmatrix} 0.05 \\ -0.07 \\ 0.02 \\ -0.1 \end{bmatrix}$$

Capa Oculta 2 (4 neuronas, 3 entradas)

Los pesos de la primera capa W2 deben moverse en el rango $(-0.5, 0.5)$

Pesos W2 =

$$\begin{bmatrix} 0.1 & 0.4 & -0.1 & 0.2 \\ -0.2 & 0.3 & 0.2 & 0.4 \\ 0.3 & -0.1 & -0.4 & -0.1 \\ 0.45 & 0.2 & 0.35 & 0.3 \end{bmatrix}$$

Sesgo b2 =

$$\begin{bmatrix} 0.05 \\ 0.1 \\ 0.3 \\ -0.1 \end{bmatrix}$$

Capa Salida (1 neuronas, 4 entradas)

Los pesos de la primera capa W1 deben moverse en el rango $(-0.577, 0.577)$

Pesos Wout =

$$\begin{bmatrix} 0.5 & -0.4 & 0.4 & -0.2 \end{bmatrix}$$

Sesgo bout = 0.1

Forward Propagation

Primera entrada Tomemos la primera fila de Xs (entrada $[2.5, 3.5, -0.5]$):

Capa Oculta 1:

La salida de la primera capa oculta, Z1 , se calcula como:

$Z1 = W1 \cdot Xs + b1$

Tomemos la primera fila de Xs (entrada [2.5,3.5, −0.5]):

$$Z1 = \begin{bmatrix} 0.1 & 0.5 & -0.2 \\ -0.3 & 0.4 & 0.1 \\ 0.2 & -0.1 & -0.5 \\ 0.05 & 0.3 & 0.2 \end{bmatrix} * \begin{bmatrix} 2.5 \\ 3.5 \\ -0.5 \end{bmatrix} + \begin{bmatrix} 0.05 \\ -0.07 \\ 0.02 \\ -0.1 \end{bmatrix}$$

$$\begin{bmatrix} 2.1 \\ 0.6 \\ 0.4 \\ 1.075 \end{bmatrix} + \begin{bmatrix} 0.05 \\ -0.07 \\ 0.02 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 2.15 \\ 0.53 \\ 0.42 \\ 0.975 \end{bmatrix}$$

Z1 = [2.15 , 0.53 , 0.42, 0.975]

ReLU: f(x) = max (0,x):

A1= ReLU (Z1) = [2.15 , 0.53 , 0.42, 0.975]

Capa Oculta 2:

La salida de la segunda capa oculta, Z2 , se calcula como:

$Z2 = W2 \cdot A1 + b2$

$$Z2 = \begin{bmatrix} 0.1 & 0.4 & -0.1 & 0.2 \\ -0.2 & 0.3 & 0.2 & 0.4 \\ 0.3 & -0.1 & -0.4 & -0.1 \\ 0.45 & 0.2 & 0.35 & 0.3 \end{bmatrix} * \begin{bmatrix} 2.15 \\ 0.53 \\ 0.42 \\ 0.975 \end{bmatrix} + \begin{bmatrix} 0.05 \\ 0.1 \\ 0.3 \\ -0.1 \end{bmatrix}$$

$$\begin{bmatrix} 0.58 \\ 0.20 \\ 0.33 \\ 1.51 \end{bmatrix} + \begin{bmatrix} 0.05 \\ 0.1 \\ 0.3 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 0.63 \\ 0.30 \\ 0.63 \\ 1.41 \end{bmatrix}$$

Z2= [0.63 , 0.303 , 0.6265, 1.413]

ReLU: f(x) = max (0,x):

A2= ReLU (Z2) = [0.63 , 0.303 , 0.6265, 1.413]

Capa Salida:

Zout = Wout * A2 + b out

Z out = $\begin{bmatrix} 0.5 & -0.4 & 0.4 & -0.2 \end{bmatrix} * \begin{bmatrix} 0.63 \\ 0.303 \\ 0.6265 \\ 1.413 \end{bmatrix} + 0.1$

Tanh

Z out = 0.2618 0.2559

La activación final es:

y1 = 0.2559

Segunda entrada Tomemos la segunda fila de Xs (entrada [4, -1 , 0.5]:

Capa Oculta 1:

La salida de la primera capa oculta, Z1 , se calcula como:

$Z1=W1 \cdot Xs+b1$

Tomemos la segunda fila de Xs (entrada [4, -1, 0.5]):

$$Z1 \begin{bmatrix} 0.1 & 0.5 & -0.2 \\ -0.3 & 0.4 & 0.1 \\ 0.2 & -0.1 & -0.5 \\ 0.05 & 0.3 & 0.2 \end{bmatrix} * \begin{bmatrix} 4 \\ -1 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 0.05 \\ -0.07 \\ 0.02 \\ -0.1 \end{bmatrix}$$

$$\begin{bmatrix} -0.2 \\ -1.55 \\ 0.65 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.05 \\ -0.07 \\ 0.02 \\ -0.1 \end{bmatrix} = \begin{bmatrix} -0.15 \\ -1.62 \\ 0.67 \\ -0.1 \end{bmatrix}$$

Z1 = [-0.15 , -1.62 , 0.67, -0.1]

ReLU: f(x) = max (0,x):

A1= ReLU (Z1) = [0, 0, 0.67, 0]

Capa Oculta 2:

La salida de la segunda capa oculta, Z2 , se calcula como:

$Z2=W2 \cdot A1+b2$

$$Z2= \begin{bmatrix} 0.1 & 0.4 & -0.1 & 0.2 \\ -0.2 & 0.3 & 0.2 & 0.4 \\ 0.3 & -0.1 & -0.4 & -0.1 \\ 0.45 & 0.2 & 0.35 & 0.3 \end{bmatrix} * \begin{bmatrix} 0 \\ 0 \\ 0.67 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.05 \\ 0.1 \\ 0.3 \\ -0.1 \end{bmatrix}$$

$$\begin{bmatrix} -0.067 \\ 0.134 \\ -0.268 \\ 0.2345 \end{bmatrix} + \begin{bmatrix} 0.05 \\ 0.1 \\ 0.3 \\ -0.1 \end{bmatrix} = \begin{bmatrix} -0.017 \\ 0.234 \\ 0.032 \\ 0.1345 \end{bmatrix}$$

Z1 = [-0.017 , 0.234, 0.032, 0.1345]

ReLU: f(x) = max (0,x):

A2= ReLU (Z2) = [0 , 0.234 , 0.032, 0.1345]

Capa Salida:

Zout = Wout * A2 + b out

$$Z\ out = \begin{bmatrix} 0.5 & -0.4 & 0.4 & -0.2 \end{bmatrix} * \begin{bmatrix} 0 \\ 0.234 \\ 0.032 \\ 0.1345 \end{bmatrix} + 0.1$$

Tanh

Z out = -0.0077 0.0077

La activación final es:

y2 = 0.0077

Tercera entrada Tomemos la tercera fila de Xs (entrada [0.5, 1.5, 1.0]:

Capa Oculta 1:

La salida de la primera capa oculta, Z1 , se calcula como:

Z1=W1·Xs+b1

Tomemos la tercera fila de Xs (entrada [0.5, 1.5, 1.0]):

$$Z1 \begin{bmatrix} 0.1 & 0.5 & -0.2 \\ -0.3 & 0.4 & 0.1 \\ 0.2 & -0.1 & -0.5 \\ 0.05 & 0.3 & 0.2 \end{bmatrix} * \begin{bmatrix} 0.5 \\ 1.5 \\ 1 \end{bmatrix} + \begin{bmatrix} 0.05 \\ -0.07 \\ 0.02 \\ -0.1 \end{bmatrix}$$

$$\begin{bmatrix} 0.6 \\ 0.55 \\ -0.55 \\ 0.675 \end{bmatrix} + \begin{bmatrix} 0.05 \\ -0.07 \\ 0.02 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 0.65 \\ 0.48 \\ -0.53 \\ 0.575 \end{bmatrix}$$

Z1 = [0.65 , 0.48 ,-0.53, 0.575]

ReLU: f(x) = max (0,x):

A1= ReLU (Z1) = [0.65, 0.48, 0.0, 0.575]

Capa Oculta 2:

La salida de la segunda capa oculta, Z2 , se calcula como:

Z2=W2·A1+b2

$$Z2= \begin{bmatrix} 0.1 & 0.4 & -0.1 & 0.2 \\ -0.2 & 0.3 & 0.2 & 0.4 \\ 0.3 & -0.1 & -0.4 & -0.1 \\ 0.45 & 0.2 & 0.35 & 0.3 \end{bmatrix} * \begin{bmatrix} 0.65 \\ 0.48 \\ 0 \\ 0.575 \end{bmatrix} + \begin{bmatrix} 0.05 \\ 0.1 \\ 0.3 \\ -0.1 \end{bmatrix}$$

$$\begin{bmatrix} 0.372 \\ 0.244 \\ 0.0895 \\ 0.561 \end{bmatrix} + \begin{bmatrix} 0.05 \\ 0.1 \\ 0.3 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 0.422 \\ 0.344 \\ 0.3895 \\ 0.461 \end{bmatrix}$$

Z2 = [0.422, 0.344, 0.3895, 0.461]

ReLU: f(x) = max (0,x):

A2= ReLU (Z2) = [0.422 , 0.344 , 0.3895, 0.461]

Capa Salida:

Zout = Wout * A2 + b out

$$Z\ out = \begin{bmatrix} 0.5 & -0.4 & 0.4 & -0.2 \end{bmatrix} * \begin{bmatrix} 0.422 \\ 0.344 \\ 0.3895 \\ 0.461 \end{bmatrix} + 0.1$$

Tanh

Z out = 0.237 0.2327

La activación final es:

y3 = 0.2327

Cuarta entrada Tomemos la cuarta fila de Xs (entrada [3.0, 2.0, -1.5]:

Capa Oculta 1:

La salida de la primera capa oculta, Z1 , se calcula como:

$Z1 = W1 \cdot Xs + b1$

Tomemos la cuarta fila de Xs (entrada [3, 2, -1.5]):

$$Z1 = \begin{bmatrix} 0.1 & 0.5 & -0.2 \\ -0.3 & 0.4 & 0.1 \\ 0.2 & -0.1 & -0.5 \\ 0.05 & 0.3 & 0.2 \end{bmatrix} * \begin{bmatrix} 3 \\ 2 \\ -1.5 \end{bmatrix} + \begin{bmatrix} 0.05 \\ -0.07 \\ 0.02 \\ -0.1 \end{bmatrix}$$

$$\begin{bmatrix} 1.6 \\ -0.25 \\ 1.15 \\ 0.45 \end{bmatrix} + \begin{bmatrix} 0.05 \\ -0.07 \\ 0.02 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 1.65 \\ -0.32 \\ 1.17 \\ 0.35 \end{bmatrix}$$

Z1 = [1.65 ,-0.32 , 1.17, 0.35]

ReLU: f(x) = max (0,x):

A1= ReLU (Z1) = [1.65, 0, 1.17, 0.35]

Capa Oculta 2:

La salida de la segunda capa oculta, Z2 , se calcula como:

$Z2 = W2 \cdot A1 + b2$

$$Z2 = \begin{bmatrix} 0.1 & 0.4 & -0.1 & 0.2 \\ -0.2 & 0.3 & 0.2 & 0.4 \\ 0.3 & -0.1 & -0.4 & -0.1 \\ 0.45 & 0.2 & 0.35 & 0.3 \end{bmatrix} * \begin{bmatrix} 1.65 \\ 0 \\ 1.17 \\ 0.35 \end{bmatrix} + \begin{bmatrix} 0.05 \\ 0.1 \\ 0.3 \\ -0.1 \end{bmatrix}$$

$$\begin{bmatrix} 0.118 \\ 0.044 \\ -0.008 \\ 1.257 \end{bmatrix} + \begin{bmatrix} 0.05 \\ 0.1 \\ 0.3 \\ -0.1 \end{bmatrix} = \begin{bmatrix} 0.168 \\ 0.144 \\ 0.292 \\ 1.157 \end{bmatrix}$$

Z2 = [0.168, 0.144, 0.292, 1.157]

ReLU: f(x) = max (0,x):

A2= ReLU (Z2) = [0.168, 0.144, 0.292, 1.157]

Capa Salida:

Zout = Wout * A2 + b out

$$Z_{out} = \begin{bmatrix} 0.5 & -0.4 & 0.4 & -0.2 \end{bmatrix} * \begin{bmatrix} 0.168 \\ 0.144 \\ 0.292 \\ 1.157 \end{bmatrix} + 0.1$$

		Tanh
Z out =	0.0118	0.0118

La activación final es:

y4 = 0.0118

Backward Propagation

Calculo del error de cada salida

	errori	=	yi´	-	yi		
Entrada 1	error 1	=	0.2559	-	1	=	-0.7441
Entrada 2	error 2	=	0.0077	-	-1	=	1.0077
Entrada 3	error 3	=	0.2327	-	-1	=	1.2327
Entrada 4	error 4	=	0.0118	-	1	=	-0.9882

Paso 1: Calculo del Error (Pérdida)

Función de pérdida

L = (1/4) * ((0.2559 -1)^2 + (0.0077-(-1))^2+(0.2327-(-1))^2+(0.0118-1)^2)

L = 1.016308158

Primera Entrada

Paso 2: **Backward Propagation desde la Capa de Salida hacia la Segunda Capa Oculta**

1 Gradiente en la salida

(∂L / ∂Y out) .= (y out - y) = (0-2559 - 1) = -0.7441

2 Derivada de la función tanh aplicada en la salida

(∂L / ∂ Z out) = (∂L / ∂ y out) * tanh ´(Zout)

tanh ´(Zout) = 1- tanh^2(Zout) = 1 - (0.2559)^2 = 0.93452

(∂L / ∂ Z out) = .=-0.7441*0.9345 = -0.6954

3 Gradiente de los pesos W3 y las bias b3 en la capa de salida

Gradiente de los pesos Wout

(∂L / ∂W out) = (∂L / ∂ Zout) * A2

$$(\partial L / \partial W \text{ out}) = -0.69536145 * \begin{bmatrix} 0.63 \\ 0.303 \\ 0.6265 \\ 1.413 \end{bmatrix} = \begin{bmatrix} -0.4381 \\ -0.2107 \\ -0.4356 \\ -0.9825 \end{bmatrix}$$

Gradiente de bias b out

(∂L / ∂b out3) = (∂L / ∂ Z out) = -0.695

Paso 3: Backward Propagation hacia la Segunda Capa Oculta

1 Cálculo del gradiente de la salida A2 de la segunda capa

(∂L / ∂A2) = (∂L / Zout) * Wout

(∂L / Zout) = -0.69536145

Wout = [0.5, -0.4, 0.4, -0.2]

$$(\partial L / \partial A2) = -0.69536145 * \begin{bmatrix} 0.5 \\ -0.4 \\ 0.4 \\ -0.2 \end{bmatrix} = \begin{bmatrix} -0.3477 \\ 0.27814 \\ -0.2781 \\ 0.13907 \end{bmatrix}$$

2 Derivada de la función tanh aplicada en Z2

(∂L / ∂ Z2) = (∂L / ∂ A2) * tanh ' (Z 2)

tanh ´(x) = 1 - tanh^2 (x)

$$\tanh'(Z2) = \begin{bmatrix} (1-(0.63)^2) \\ (1-(0.303)^2) \\ (1-(0.6265)^2) \\ (1-(1.413)^2) \end{bmatrix} = \begin{bmatrix} 0.603 \\ 0.908 \\ 0.607 \\ -0.997 \end{bmatrix}$$

$$(\partial L / \partial Z2) = \begin{bmatrix} -0.348 \\ 0.278 \\ -0.278 \\ 0.139 \end{bmatrix} * \begin{bmatrix} 0.603 \\ 0.908 \\ 0.607 \\ -0.997 \end{bmatrix} = \begin{bmatrix} -0.210 \\ 0.253 \\ -0.169 \\ -0.139 \end{bmatrix}$$

3 Gradiente respecto a los pesos W2 y bias 2

Gradiente de los pesos W2

$(\partial L / \partial W2) = (\partial L / \partial Z2) * A1T$

A1 = [2.15 , 0.53 , 0.42, 0.975]

$$(\partial L / \partial W2) = \begin{bmatrix} -0.210 \\ 0.253 \\ -0.169 \\ -0.139 \end{bmatrix} * \begin{bmatrix} 2.15 , 0.53 , 0.42, 0.975 \end{bmatrix} =$$

$$= \begin{bmatrix} -0.452 & -0.1113 & -0.0882 & -0.20475 \\ 0.54395 & 0.13409 & 0.10626 & 0.246675 \\ -0.36335 & -0.08957 & -0.07098 & -0.164775 \\ -0.29885 & -0.07367 & -0.05838 & -0.135525 \end{bmatrix}$$

Gradiente bias b2

$$(\partial L / \partial b2) = \begin{bmatrix} -0.210 \\ 0.253 \\ -0.169 \\ -0.139 \end{bmatrix}$$

Paso 4: Backward Propagation hacia la Primera Capa Oculta

1 Gradiente de A1

$(\partial L / \partial A1) = (\partial L / \partial Z2) * W2T$

2 Propagado gradiente hacia atrás

$$(\partial L / \partial Z2) = \begin{bmatrix} -0.210 \\ 0.253 \\ -0.169 \\ -0.139 \end{bmatrix}$$

La Matriz de pesos W2

$$W2 = \begin{bmatrix} 0.1 & 0.4 & -0.1 & 0.2 \\ -0.2 & 0.3 & 0.2 & 0.4 \\ 0.3 & -0.1 & -0.4 & -0.1 \\ 0.45 & 0.2 & 0.35 & 0.3 \end{bmatrix}$$

$$(\partial L / \partial A1) = W2T * (\partial L / \partial Z2)$$

$$(\partial L / \partial A1) = \begin{bmatrix} 0.1 & -0.2 & 0.3 & 0.45 \\ 0.4 & 0.3 & -0.1 & 0.2 \\ -0.1 & 0.2 & -0.4 & 0.35 \\ 0.2 & 0.4 & -0.1 & 0.3 \end{bmatrix} * \begin{bmatrix} -0.210 \\ 0.253 \\ -0.169 \\ -0.139 \end{bmatrix}$$

$$(\partial L / \partial A1) = \begin{bmatrix} -0.185 \\ -0.019 \\ 0.091 \\ 0.034 \end{bmatrix}$$

3 Gradiente de Z1 en la primera capa oculta

$$(\partial L / \partial Z1) = (\partial L / \partial A1) * \tanh'(Z1)$$

$$\tanh'(Z1) = 1 - \tanh^2(Z1)$$

$$\tanh'(Z1) = \begin{bmatrix} (1-(2.15)^2) \\ (1-(0.53)^2) \\ (1-(0.42)^2) \\ (1-(1.975)^2) \end{bmatrix} = \begin{bmatrix} -3.623 \\ 0.719 \\ 0.824 \\ -2.901 \end{bmatrix}$$

$$(\partial L / \partial Z1) = \begin{bmatrix} -0.185 \\ -0.019 \\ 0.091 \\ 0.034 \end{bmatrix} * \begin{bmatrix} -3.623 \\ 0.719 \\ 0.824 \\ -2.901 \end{bmatrix} = \begin{bmatrix} 0.670 \\ -0.014 \\ 0.075 \\ -0.100 \end{bmatrix}$$

4 Gradiente de W1

$(\partial L / \partial W1) =$ $(\partial L / \partial Z1) * XT$

Primera entrada de X1 = [2.5, 3.5 , -0.5]

$$(\partial L / \partial W1) = \begin{bmatrix} 0.670 \\ -0.014 \\ 0.075 \\ -0.100 \end{bmatrix} * \begin{bmatrix} 2.5, 3.5, -0.5 \end{bmatrix}$$

$$(\partial L / \partial W1) = \begin{bmatrix} 1.675 & 2.345 & -0.335 \\ -0.035 & -0.049 & 0.007 \\ 0.1875 & 0.2625 & -0.0375 \\ -0.25 & -0.35 & 0.05 \end{bmatrix}$$

Paso 5: **Gradiente de las bias b1 - Primera Capa Oculta**

$$(\partial L / \partial b1) = (\partial L / \partial Z1) = \begin{bmatrix} 0.670 \\ -0.014 \\ 0.075 \\ -0.100 \end{bmatrix}$$

Resumen de los Gradientes Finales: Primera Entrada

Gradiente W3

$$(\partial L / \partial Wout) = \begin{bmatrix} -0.438 \\ -0.211 \\ -0.436 \\ -0.983 \end{bmatrix}$$

Gradiente b3

$(\partial L / \partial b\ out) =$ -0.695

Gradiente W2

$$(\partial L / \partial W2) = \begin{bmatrix} -0.452 & -0.111 & -0.088 & -0.205 \\ 0.544 & 0.134 & 0.106 & 0.247 \\ -0.363 & -0.090 & -0.071 & -0.165 \\ -0.299 & -0.074 & -0.058 & -0.136 \end{bmatrix}$$

Gradiente b2

$$(\partial L / \partial b2) = \begin{bmatrix} -0.210 \\ 0.253 \\ -0.169 \\ -0.139 \end{bmatrix}$$

Gradiente W1

$$(\partial L / \partial W1) = \begin{bmatrix} 1.675 & 2.345 & -0.335 \\ -0.035 & -0.049 & 0.007 \\ 0.188 & 0.263 & -0.038 \\ -0.250 & -0.350 & 0.050 \end{bmatrix}$$

Gradiente b1

$$(\partial L / \partial b1) = \begin{bmatrix} 0.670 \\ -0.014 \\ 0.075 \\ -0.100 \end{bmatrix}$$

Actualización de pesos y bias con Descenso de Gradiente

$$W_{nuevo} = W_{viejo} - n * ((\partial L / \partial W)$$

$$b_{nuevo} = b_{viejo} - n * ((\partial L / \partial b)$$

Primera Entrada

Revisar los Gradientes

Gradiente W out

$$(\partial L / \partial W_{out}) = \begin{bmatrix} -0.438 \\ -0.211 \\ -0.436 \\ -0.983 \end{bmatrix}$$

Gradiente b out

$$(\partial L / \partial b_{out}) = -0.695$$

Tasa de Aprendizaje (n)

$$n = 0.01$$

Actualziar los pesos de W out

$$W_{out} = \begin{bmatrix} 0.5 \\ -0.4 \\ 0.4 \\ -0.2 \end{bmatrix} - 0.01 * \begin{bmatrix} -0.438 \\ -0.211 \\ -0.436 \\ -0.983 \end{bmatrix}$$

$$W_{out} = \begin{bmatrix} 0.504 \\ -0.398 \\ 0.404 \end{bmatrix}$$

$$\begin{bmatrix} -0.190 \end{bmatrix}$$

Actualizar los pesos de b out

b out = 0.107

Gradiente W 2

$$(\partial L / \partial W\,2) = \begin{bmatrix} -0.452 & -0.111 & -0.088 & -0.205 \\ 0.544 & 0.134 & 0.106 & 0.247 \\ -0.363 & -0.090 & -0.071 & -0.165 \\ -0.299 & -0.074 & -0.058 & -0.136 \end{bmatrix}$$

Gradiente b 2

$$(\partial L / \partial b\,2) = \begin{bmatrix} -0.210 \\ 0.253 \\ -0.169 \\ -0.139 \end{bmatrix}$$

Tasa de Aprendizaje (n)

n = 0.01

Actualziar los pesos de W 2

$$W\,2 = \begin{bmatrix} 0.105 & 0.401 & -0.099 & 0.202 \\ -0.205 & 0.299 & 0.199 & 0.398 \\ 0.304 & -0.099 & -0.399 & -0.098 \\ 0.453 & 0.201 & 0.351 & 0.301 \end{bmatrix}$$

Actualizar los pesos de b 2

$$b\,2 = \begin{bmatrix} 0.052 \\ 0.100 \\ 0.300 \\ -0.100 \end{bmatrix}$$

Gradiente W 1

$$(\partial L / \partial W\,1) = \begin{bmatrix} 1.675 & 2.345 & -0.335 \\ -0.035 & -0.049 & 0.007 \\ 0.188 & 0.263 & -0.038 \\ -0.250 & -0.350 & 0.050 \end{bmatrix}$$

Gradiente b 1

$$(\partial L / \partial b\ 1) = \begin{bmatrix} 0.670 \\ -0.014 \\ 0.075 \\ -0.100 \end{bmatrix}$$

Tasa de Aprendizaje (n)

n = 0.01

Actualziar los pesos de W 1

$$W\ 1 = \begin{bmatrix} 0.083 & 0.477 & -0.197 \\ -0.300 & 0.400 & 0.100 \\ 0.198 & -0.103 & -0.500 \\ 0.053 & 0.304 & 0.200 \end{bmatrix}$$

Actualizar los pesos de b 1

$$b\ 1 = \begin{bmatrix} 0.043 \\ -0.070 \\ 0.019 \\ -0.099 \end{bmatrix}$$

Forward Propagation Nuevos Pesos y Bias

Primera entrada

Tomemos la primera fila de Xs (entrada [2.5,3.5, −0.5]):

Capa Oculta 1:

La salida de la primera capa oculta, Z1 , se calcula como:

$Z1 = W1 \cdot Xs + b1$

Tomemos la primera fila de Xs (entrada [2.5,3.5, −0.5]):

$$Z1 = \begin{bmatrix} 0.083 & 0.477 & -0.197 \\ -0.300 & 0.400 & 0.100 \\ 0.198 & -0.103 & -0.500 \\ 0.053 & 0.304 & 0.200 \end{bmatrix} * \begin{bmatrix} 2.5 \\ 3.5 \\ -0.5 \end{bmatrix} + \begin{bmatrix} 0.043 \\ -0.070 \\ 0.019 \\ -0.099 \end{bmatrix}$$

$$\begin{bmatrix} 1.9755 \\ 0.6 \\ 0.3845 \end{bmatrix} + \begin{bmatrix} 0.043 \\ -0.070 \\ 0.019 \end{bmatrix} = \begin{bmatrix} 2.019 \\ 0.530 \\ 0.404 \end{bmatrix}$$

1.0965	-0.099	0.998

Z1 = [2.019 , 0.53 , 0.404, 0.998]

ReLU: f(x) = max (0,x):

A1= ReLU (Z1) = [2.019 , 0.53 , 0.404, 0.998]

Capa Oculta 2:

La salida de la segunda capa oculta, Z2 , se calcula como:

Z2=W2·A1+b2

$$Z2=\begin{bmatrix} 0.105 & 0.401 & -0.099 & 0.202 \\ -0.205 & 0.299 & 0.199 & 0.398 \\ 0.304 & -0.099 & -0.399 & -0.098 \\ 0.453 & 0.201 & 0.351 & 0.301 \end{bmatrix} * \begin{bmatrix} 2.019 \\ 0.53 \\ 0.404 \\ 0.998 \end{bmatrix} + \begin{bmatrix} 0.052 \\ 0.100 \\ 0.300 \\ -0.100 \end{bmatrix}$$

$$\begin{bmatrix} 0.59 \\ 0.22 \\ 0.30 \\ 1.46 \end{bmatrix} + \begin{bmatrix} 0.052 \\ 0.100 \\ 0.300 \\ -0.100 \end{bmatrix} = \begin{bmatrix} 0.64 \\ 0.32 \\ 0.60 \\ 1.36 \end{bmatrix}$$

Z2= [0.64 , 0.32 , 0.60, 1.36]

ReLU: f(x) = max (0,x):

A2= ReLU (Z2) = [0.64 , 0.32 , 0.60, 1.36]

Capa Salida:

Zout = Wout * A2 + b out

$$Z\,out = \begin{bmatrix} 0.5 & -0.4 & 0.4 & -0.2 \end{bmatrix} * \begin{bmatrix} 0.64 \\ 0.32 \\ 0.6 \\ 1.36 \end{bmatrix} + 0.1$$

Tanh

Z out = 0.26 0.254

La activación final es:

y1 = 0.254