Red neuronal: Actualización de pesos

García Rivera Bogdan Kaleb
MIA-2

May 2, 2025

El problema a resolver es realizar los cálculos necesarios para poder actualizar los pesos. En la figura 1 se muestra la red neuronal a utilizar.

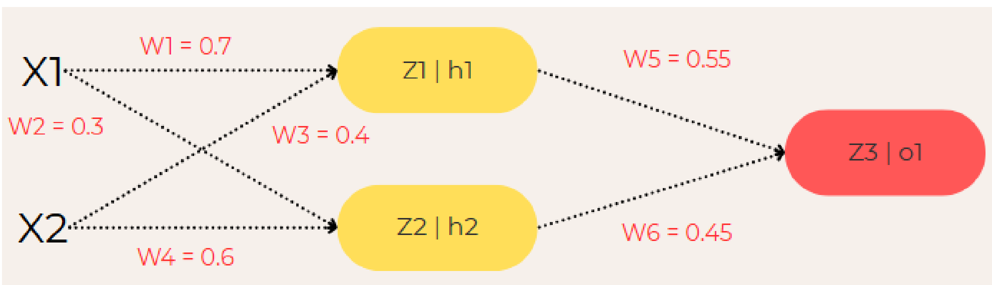

Figure 1: Arquitectura de la red neuronal a analizar

1 Forward propagation

Suponiendo que las entradas de X_1 y X_2 son 0.5 y 0.3 respectivamente, se tiene lo siguiente:

$$Z_1 = X_1\omega_1 + X_2\omega_3 = (0.5)(0.7) + (0.3)(0.4) = 0.47 \tag{1}$$

Se realiza lo mismo con Z_2:

$$Z_2 = X_1\omega_2 + X_2\omega_4 = (0.5)(0.3) + (0.3)(0.6) = 0.33 \tag{2}$$

Posteriormente se calculan las salidas:

$$h_1 = \frac{1}{1+e^{-z1}} = \frac{1}{1+e^{-0.47}} \approx 0.615 \tag{3}$$

$$h_2 = \frac{1}{1+e^{-z2}} = \frac{1}{1+e^{-0.33}} \approx 0.581 \tag{4}$$

Para la salida de Z_3 se calcula como:

$$Z_3 = h_1\omega_5 + h_2\omega_6 = (0.615)(0.55) + (0.581)(0.45) \approx 0.6 \tag{5}$$

Por lo tanto, la salida de la red neuronal estará dada por:

$$salida = \frac{1}{1+e^{-z3}} = \frac{1}{1+e^{-0.6}} \approx 0.646 \tag{6}$$

En esta red, la salida esperada debería ser igual a 1, por lo que se debe calcular el error existente entre la salida esperada con respecto a la real. Para este proceso se hace uso del error cuadrático medio (MSE).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i}^2) = (1-0.646)^2 \approx 0.125 \tag{7}$$

2 Back propagation

Para este caso se debe de hacer los cálculos de atrás hacia adeltante. Se puede comenzar con ω_5:

$$\omega_5^* = \omega_5 - \eta\frac{\partial C}{\partial \omega_5} \tag{8}$$

Este cálculo no es directo, por lo que se lleva a cabo a través de ciertas derivadas parciales:

$$\frac{\partial C}{\partial \omega_5} = \frac{\partial C}{\partial o_1}\cdot\frac{\partial o_1}{\partial z_3}\cdot\frac{\partial Z_3}{\partial \omega_5} \tag{9}$$

Para el caso de $\frac{\partial C}{\partial o_1}$:

$$C = (y - o_1)^2$$

$$\frac{\partial C}{\partial o_1} = \frac{\partial}{\partial o_1}(y-o_1)^2 = -2(y-o_1)$$

Sustituyendo o_1 y $y = 1$:

$$\frac{\partial C}{\partial o_1} = -2(1-0.646) \approx -0.71 \tag{10}$$

De igual manera se hacen los cálculos posteriores. Recordando que $o = \frac{1}{1+e^{-z}}$:

$$\frac{\partial o_1}{\partial Z_3} = \frac{1}{1+e^{-z_3}}\cdot(1-\frac{1}{1+e^{-z_3}}) = 0.646(1-0.646) = 0.229 \tag{11}$$

Posteriormente se hace el cálculo del último término:

$$\frac{\partial Z_3}{\partial \omega_5} = \frac{\partial}{\partial \omega_5}\cdot(h_1\omega_5 + h_2\omega_6) = h_1 = 0.615 \tag{12}$$

Finalmente:

$$\frac{\partial C}{\partial \omega_5} = (-0.71) * (0.229) * (0.615) \approx -0.1$$

Así entonces, el peso actualizado es el siguiente con $\eta = 0.1$:

$$w_5^* = 0.55 - \eta(-0.1) = 0.56$$

Para el peso de ω_6 es un proceso parecido:

$$\frac{\partial C}{\partial \omega_6} = \frac{\partial C}{\partial o_1} \cdot \frac{\partial o_1}{\partial z_3} \cdot \frac{\partial Z_3}{\partial \omega_6} \tag{13}$$

Esto es entonces:

$$\frac{\partial C}{\partial \omega_6} = (0.229)(-0.71)h_2 = (0.229)(-0.71)(0.581) \approx -0.094$$

Por lo tanto:

$$\omega_6^* = 0.45 - (0.1)(-0.094) \approx 0.4594$$

Para w_1 :

$$\frac{\partial C}{\partial w_1} = \frac{\partial C}{\partial o_1} \cdot \frac{\partial o_1}{\partial z_3} \cdot \frac{\partial z_3}{\partial h_1} \cdot \frac{\partial h_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_1} \tag{14}$$

Para este caso se tiene lo siguiente:

$$\frac{\partial z_3}{\partial h_1} = w_5 = 0.55$$

$$\frac{\partial h_1}{\partial z_1} = h_1 \cdot (1 - h_1) \approx 0.615 \cdot 0.385 \approx 0.236$$

$$\frac{\partial z_1}{\partial w_1} = X_1 = 0.5$$

Así entonces:

$$\frac{\partial C}{\partial w_1} = (-0.71)(0.229)(0.55)(0.236)(0.5) \approx -0.01$$

Por lo que el nuevo peso para w_1 con $\eta = 0.1$:

$$w_1^* = 0.7 - 0.1(-0.01) = 0.701$$

Para w_2:

$$\frac{\partial C}{\partial w_2} = \frac{\partial C}{\partial o_1} \cdot \frac{\partial o_1}{\partial z_3} \cdot \frac{\partial z_3}{\partial h_2} \cdot \frac{\partial h_2}{\partial z_2} \cdot \frac{\partial z_2}{\partial w_2} \tag{15}$$

Para este caso se tiene lo siguiente:

$$\frac{\partial z_3}{\partial h_2} = w_6 = 0.45$$

$$\frac{\partial h_2}{\partial z_2} = h_2 \cdot (1 - h_2) \approx 0.581 \cdot 0.419 \approx 0.243$$

$$\frac{\partial z_2}{\partial w_2} = X_1 = 0.5$$

Así entonces:

$$\frac{\partial C}{\partial w_1} = (-0.71)(0.229)(0.45)(0.243)(0.5) \approx -0.00889$$

Por lo que el nuevo peso para w_2 con $\eta = 0.1$:

$$w_2^* = 0.3 - 0.1(-0.00889) = 0.3008$$

Para w_3:

$$\frac{\partial C}{\partial w_3} = \frac{\partial C}{\partial o_1} \cdot \frac{\partial o_1}{\partial z_3} \cdot \frac{\partial z_3}{\partial h_1} \cdot \frac{\partial h_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial w_3} \tag{16}$$

Para este caso se tiene lo siguiente:

$$\frac{\partial z_3}{\partial h_1} = w_5 = 0.55$$

$$\frac{\partial h_1}{\partial z_1} = h_1 \cdot (1 - h_1) \approx 0.615 \cdot 0.385 \approx 0.236$$

$$\frac{\partial z_2}{\partial w_2} = X_2 = 0.3$$

Así entonces:

$$\frac{\partial C}{\partial w_1} = (-0.71)(0.229)(0.55)(0.236)(0.3) \approx -0.00633$$

Por lo que el nuevo peso para w_3 con $\eta = 0.1$:

$$w_3^* = 0.4 - 0.1(-0.00633) = 0.400633$$

Para w_4:

$$\frac{\partial C}{\partial w_4} = \frac{\partial C}{\partial o_1} \cdot \frac{\partial o_1}{\partial z_3} \cdot \frac{\partial z_3}{\partial h_2} \cdot \frac{\partial h_2}{\partial z_2} \cdot \frac{\partial z_2}{\partial w_4} \tag{17}$$

Para este caso se tiene lo siguiente:

$$\frac{\partial z_3}{\partial h_2} = w_6 = 0.45$$

$$\frac{\partial h_2}{\partial z_2} \approx 0.243$$

$$\frac{\partial z_2}{\partial w_4} = X_2 = 0.3$$

Así entonces:

$$\frac{\partial C}{\partial w_4} = (-0.71)(0.229)(0.45)(0.243)(0.3) \approx -0.00533$$

Por lo que el nuevo peso para w_4 con $\eta = 0.1$:

$$w_4^* = 0.6 - 0.1(-0.00533) = 0.605533$$

Finalmente la tabla de pesos actualizados se visualiza en la tabla 1.

Table 1: Actualización de Pesos

Peso	**Valor Antiguo**	**Valor Nuevo**
w_1	0.7	0.701
w_2	0.3	0.3008
w_3	0.4	0.400633
w_4	0.6	0.605533
w_5	0.55	0.56
w_6	0.45	0.4594

Como es de observarse, los pesos en una primera iteración no cambian drasticamente.

Para verificar si la actualización de pesos ha acercado el valor de salida a 1, se puede repetir el proceso inicial de cálculo.

3 Forward propagation con los nuevos pesos

Nuevamente se supone a X_1 y X_2 son 0.5 y 0.3 respectivamente, se tiene lo siguiente:

$$Z_1 = X_1\omega_1 + X_2\omega_3 = (0.5)(0.701) + (0.3)(0.400633) \approx 0.4706899 \tag{18}$$

Se realiza lo mismo con Z_2:

$$Z_2 = X_1\omega_2 + X_2\omega_4 = (0.5)(0.3008) + (0.3)(0.605533) \approx 0.3320599 \tag{19}$$

Posteriormente se calculan las salidas:

$$h_1 = \frac{1}{1+e^{-z1}} = \frac{1}{1+e^{-0.}} \approx 0.615547 \tag{20}$$

$$h_2 = \frac{1}{1+e^{-z2}} = \frac{1}{1+e^{-0.33}} \approx 0.582260 \tag{21}$$

Para la salida de Z_3 se calcula como:

$$Z_3 = h_1\omega_5 + h_2\omega_6 = (0.615547)(0.55) + (0.582260)(0.45) \approx 0.6005 \tag{22}$$

Por lo tanto, la salida de la red neuronal estará dada por:

$$salida = \frac{1}{1+e^{-z3}} = \frac{1}{1+e^{-0.6005}} \approx 0.6457 \tag{23}$$

Idealmente, la salida debería ser nuevamente 1; sin embargo, con los pesos actualizados, el resultado no muestra una mejora, sino un ligero empeoramiento.