Parameter Update 1

Chowdhury Mofizur Rahman

August 6, 2025

At first start with the output nodes. Define loss function and differentiate loss function with respect to all the output nodes.

$$Loss = \frac{1}{2}\{(Ob_5 - Y_5)^2 + (Ob_6 - Y_6)^2 + (Ob_7 - Y_7)^2\}$$
$$\frac{dLoss}{dY_5} = -\frac{1}{2}.2.(Ob_5 - Y_5) = -Err_5$$
$$\frac{dLoss}{dY_6} = -\frac{1}{2}.2.(Ob_6 - Y_6) = -Err_6$$
$$\frac{dLoss}{dY_7} = -\frac{1}{2}.2.(Ob_7 - Y_7) = -Err_7$$

Write the expressions for output nodes

$$Y_5 = Relu(W_{35}Y_3 + W_{45}Y_4 + b_5)$$
$$Y_6 = Relu(W_{36}Y_3 + W_{46}Y_4 + b_6)$$
$$Y_7 = Relu(W_{37}Y_3 + W_{47}Y_4 + b_7)$$

Now you can differentiate loss function with respect to all the output layer parameters $W_{35}, W_{45}, W_{36}, W_{46}, W_{37}, W_{47}, b_5, b_6, b_7$ using chain rule

$$\frac{dLoss}{dW_{35}} = \frac{dloss}{dY_5}.\frac{dY_5}{dW_{35}} = -Err_5.Y_3.Relu'(W_{35}Y_3 + W_{45}Y_4 + b_5)$$

$$\frac{dLoss}{dW_{45}} = \frac{dloss}{dY_5}.\frac{dY_5}{dW_{45}} = -Err_5.Y_4.Relu'(W_{35}Y_3 + W_{45}Y_4 + b_5)$$

$$\frac{dLoss}{dW_{36}} = \frac{dloss}{dY_6}.\frac{dY_6}{dW_{36}} = -Err_6.Y_3.Relu'(W_{36}Y_3 + W_{46}Y_4 + b_6)$$

$$\frac{dLoss}{dW_{46}} = \frac{dloss}{dY_6}.\frac{dY_6}{dW_{46}} = -Err_6.Y_4.Relu'(W_{36}Y_3 + W_{46}Y_4 + b_6)$$

$$\frac{dLoss}{dW_{37}} = \frac{dloss}{dY_7}.\frac{dY_7}{dW_{37}} = -Err_7.Y_3.Relu'(W_{37}Y_3 + W_{47}Y_4 + b_7)$$

$$\frac{dLoss}{dW_{47}} = \frac{dloss}{dY_7}.\frac{dY_7}{dW_{47}} = -Err_7.Y_4.Relu'(W_{37}Y_3 + W_{47}Y_4 + b_7)$$

$$\frac{dLoss}{db_5} = \frac{dloss}{dY_5}.\frac{dY_5}{db_5} = -Err_5.Relu'(W_{35}Y_3 + W_{45}Y_4 + b_5)$$

$$\frac{dLoss}{db_6} = \frac{dloss}{dY_6}.\frac{dY_6}{db_6} = -Err_6.Relu'(W_{36}Y_3 + W_{46}Y_4 + b_6)$$

$$\frac{dLoss}{db_7} = \frac{dloss}{dY_7}.\frac{dY_7}{db_7} = -Err_7.Relu'(W_{37}Y_3 + W_{47}Y_4 + b_7)$$

Now using gradient descent approach the update equations for all the output layer parameters are:

$$W_{35} = W_{35} + \eta Y_3.Err_5.Relu'(W_{35}Y_3 + W_{45}Y_4 + b_5) = \eta.Y_3.\delta_5$$
$$W_{45} = W_{45} + \eta Y_4.Err_5.Relu'(W_{35}Y_3 + W_{45}Y_4 + b_5) = \eta.Y_4.\delta_5$$

$$W_{36} = W_{36} + \eta Y_3.Err_6.Relu'(W_{36}Y_3 + W_{46}Y_4 + b_6) = \eta.Y3.\delta_6$$
$$W_{46} = W_{46} + \eta Y_4.Err_6.Relu'(W_{36}Y_3 + W_{46}Y_4 + b_6) = \eta.Y_4.\delta_6$$

$$W_{37} = W_{37} + \eta Y_3.Err_7.Relu'(W_{37}Y_3 + W_{47}Y_4 + b_7) = \eta.Y_3.\delta_7$$
$$W_{47} = W_{47} + \eta Y_4.Err_7.Relu'(W_{37}Y_3 + W_{47}Y_4 + b_7) = \eta.Y_4.\delta_7$$

$$b_5 = b_5 + \eta Err_5.Relu'(W_{35}Y_3 + W_{45}Y_4 + b_5) = \eta.\delta_5$$
$$b_6 = b_6 + \eta Err_6.Relu'(W_{36}Y_3 + W_{46}Y_4 + b_6) = \eta.\delta_6$$
$$b_7 = b_7 + \eta Err_7.Relu'(W_{37}Y_3 + W_{47}Y_4 + b_7) = \eta.\delta_7$$

Now consider the hidden nodes 3 and 4. Write the expressions for Y_3 and Y_4.

$$Y_3 = Relu\{W_{13}.X_1 + W_{23}.X_2 + b_3\}$$
$$Y_4 = Relu\{W_{14}.X_1 + W_{24}.X_2 + b_4\}$$

Now differentiate the Loss function with respect to hidden outputs Y_3 and Y_4. Both nodes 3 and 4 have parents 5, 6 and 7. Following the chain rule

$$\frac{dLoss}{dY_3} = \frac{dloss}{dY_5}.\frac{dY_5}{dY_3} + \frac{dloss}{dY_6}.\frac{dY_6}{dY_3} + \frac{dloss}{dY_7}.\frac{dY_7}{dY_3} = (-\delta_5).W_{35} + (-\delta_6).W_{36} + (-\delta_7).W_{37}$$

$$\frac{dLoss}{dY_4} = \frac{dloss}{dY_5}.\frac{dY_5}{dY_4} + \frac{dloss}{dY_6}.\frac{dY_6}{dY_4} + \frac{dloss}{dY_7}.\frac{dY_7}{dY_4} = (-\delta_5).W_{45} + (-\delta_6).W_{46} + (-\delta_7).W_{47}$$

Now we can differentiate Loss function with respect to all the hidden layer parameters W_{13}, W_{23}, W_{14}, W_{24}, b_3 and b_4

$$\frac{dLoss}{dW_{13}} = \frac{dLoss}{dY_3}.\frac{dY_3}{dW_{13}} = \{(-\delta_5).W_{35}+(-\delta_6).W_{36}+(-\delta_7).W_{37}\}.Relu^{'}(W_{13}.X_1+W_{23}.X_2+b_3).X_1$$

$$\frac{dLoss}{dW_{23}} = \frac{dLoss}{dY_3}.\frac{dY_3}{dW_{23}} = \{(-\delta_5).W_{35}+(-\delta_6).W_{36}+(-\delta_7).W_{37}\}.Relu^{'}(W_{13}.X_1+W_{23}.X_2+b_3).X_1$$

$$\frac{dLoss}{dW_{14}} = \frac{dLoss}{dY_4}.\frac{dY_4}{dW_{14}} = \{(-\delta_5).W_{45}+(-\delta_6).W_{46}+(-\delta_7).W_{47}\}.Relu^{'}(W_{14}.X_1+W_{24}.X_2+b_4).X_1$$

$$\frac{dLoss}{dW_{24}} = \frac{dLoss}{dY_4}.\frac{dY_4}{dW_{24}} = \{(-\delta_5).W_{45}+(-\delta_6).W_{46}+(-\delta_7).W_{47}\}.Relu^{'}(W_{14}.X_1+W_{24}.X_2+b_4).X_2$$

$$\frac{dLoss}{db_3} = \frac{dLoss}{dY_3}.\frac{dY_3}{db_3} = \{(-\delta_5).W_{35}+(-\delta_6).W_{36}+(-\delta_7).W_{37}\}.Relu^{'}(W_{13}.X_1+W_{23}.X_2+b_3)$$

$$\frac{dLoss}{db_4} = \frac{dLoss}{dY_4}.\frac{dY_4}{db_4} = \{(-\delta_5).W_{45}+(-\delta_6).W_{46}+(-\delta_7).W_{47}\}.Relu^{'}(W_{14}.X_1+W_{24}.X_2+b_4)$$

Therefore the updated equations for all the hidden layer parameters are as follows;

$$W_{13} = W_{13} + \eta.(\delta_5.W_{35} + \delta_6.W_{36} + \delta_7.W_{37}\}.Relu^{'}(W_{13}.X_1 + W_{23}.X_2 + b_3).X_1$$

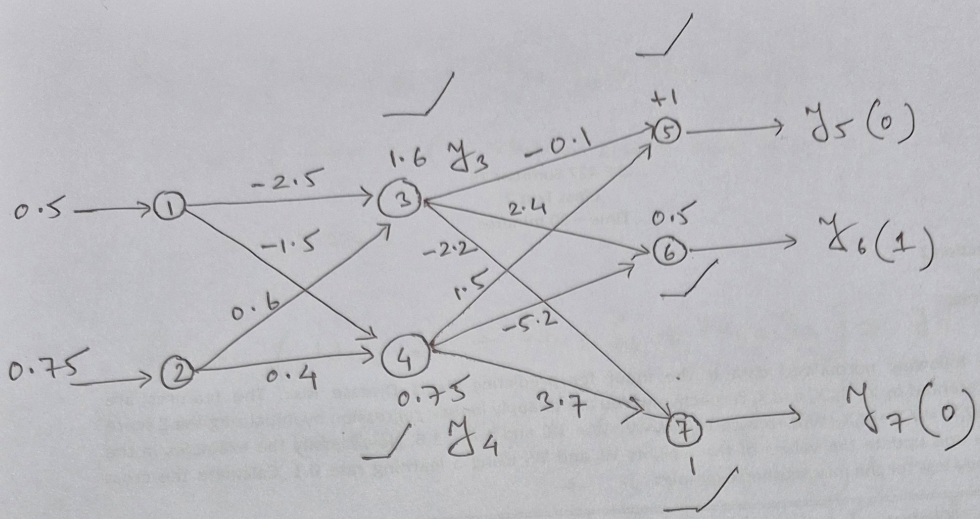
$$W_{23} = W_{23} + \eta.(\delta_5.W_{35} + \delta_6.W_{36} + \delta_7.W_{37}).Relu^{'}(W_{13}.X_1 + W_{23}.X_2 + b_3).X_2$$

$$W_{14} = W_{14} + \eta.(\delta_5.W_{45} + \delta_6.W_{46} + \delta_7.W_{47}).Relu^{'}(W_{14}.X_1 + W_{24}.X_2 + b_4).X_1$$

$$W_{24} = dW_{24} + \eta.(\delta_5.W_{45} + \delta_6.W_{46} + \delta_7.W_{47}).Relu^{'}(W_{14}.X_1 + W_{24}.X_2 + b_4).X_2$$

$$b_3 = b_3 + \eta.(\delta_5.W_{35} + \delta_6.W_{36} + \delta_7.W_{37}).Relu^{'}(W_{13}.X_1 + W_{23}.X_2 + b_3)$$

$$b_4 = b_4 + \eta.(\delta_5.W_{45} + \delta_6.W_{46} + \delta_7.W_{47}).Relu^{'}(W_{14}.X_1 + W_{24}.X_2 + b_4)$$



$$y_3 = \text{Relu}(-0.5 \times 2.5 + 0.75 \times 0.6 + 1.6) = 0.8$$

$$y_4 = \text{Relu}(-0.5 \times 1.5 + 0.75 \times 0.4 + 0.7) = 0.3$$

$$y_5 = \text{Relu}(-0.1 \times 0.8 + 0.3 \times 1.5 + 1) = 1.37$$

$$y_6 = \text{Relu}(0.8 \times 2.4 - 0.3 \times 5.2 + 0.5) = 0.86$$

$$y_7 = \text{Relu}(-2.2 \times 0.8 + 0.3 \times 3.7 + 1) = 0.35$$

$$\delta_5 = \text{Err}_5 \cdot y_5' = (0 - 1.37) \times 1 = -1.37$$

$$\delta_6 = \text{Err}_6 \cdot y_6' = (1 - 0.86) \times 1 = 0.14$$

$$\delta_7 = \text{Err}_7 \cdot y_7' = (0 - 0.35) \times 1 = -0.$$

$$\delta_3 = (w_{35}\delta_5 + w_{36}\delta_6 + w_{37}\delta_7) y_3'$$

$$= (-0.1 \times (-1.37) + 2.4 \times 0.14 + (-2.2) \times (-0.35)) \times 1$$

$$= 1.243$$

$$\delta_4 = (w_{45}\delta_5 + w_{46}\delta_6 + w_{47}\delta_7) \times y_4'$$

$$= (1.5 \times (-1.37) + (-5.2) \times 0.14 + 3.7 \times (-0.35)) \times 1$$

$$= 4.078$$

$$w_{35} = w_{35} + \eta y_3 \delta_5 = -0.1 + 0.1 \times 0.8 \times (-1.37)$$

$$w_{36} = w_{36} + \eta y_3 \delta_6 = 2.4 + 0.1 \times 0.8 \times 0.14$$

$$w_{45} = w_{45} + \eta y_4 \delta_5 \qquad w_{37} = w_{37} + \eta y_3 \delta_7$$

$$w_{46} = w_{46} + \eta y_4 \delta_6 \qquad w_{47} = w_{47} + \eta y_4 \delta_7$$

$$b_5 = b_5 + \eta \delta_5 \qquad b_6 = b_6 + \eta \delta_6 \qquad b_7 = b_7 + \eta \delta_7$$

$$w_{13} = w_{13} + \eta x_1 \delta_3 = -2.5 + 0.1 \times 0.5 \times 1.243 = -2.4$$

$$w_{14} = w_{14} + \eta x_1 \delta_4 = -1.5 + 0.1 \times 0.5 \times 4.078 = -1.3$$

$$w_{23} = w_{23} + \eta x_2 \delta_3 \qquad w_{24} = w_{24} + \eta x_2 \delta_4$$

$$b_3 = b_3 + \eta \delta_3 \qquad b_4 = b_4 + \eta \delta_4$$