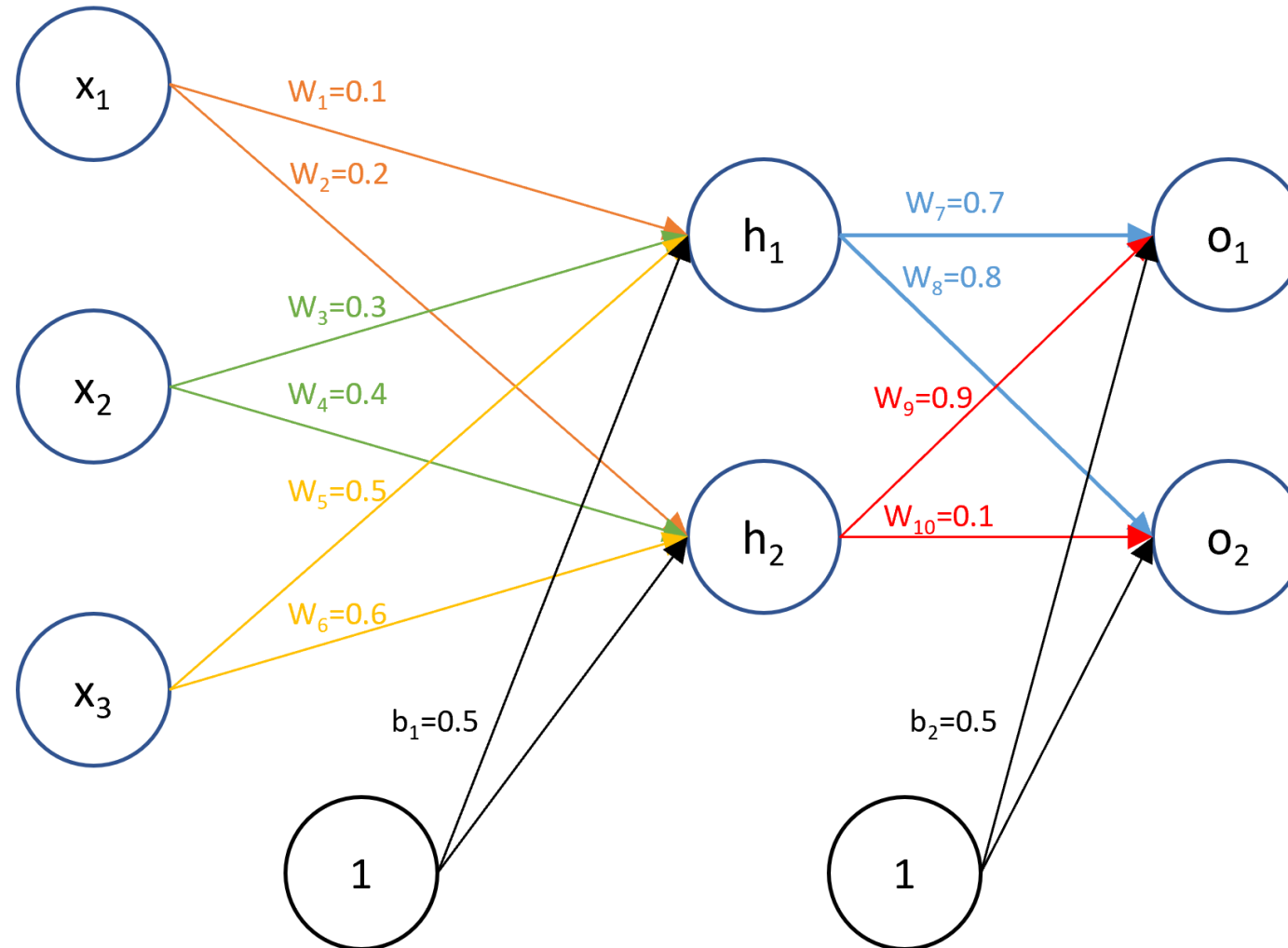


Backpropagation Example

Forward Propagation Example



Input to hidden layer

$$w_1x_1 + w_3x_2 + w_5x_3 + b_1 = z_{h_1}$$

$$w_2x_1 + w_4x_2 + w_6x_3 + b_1 = z_{h_2}$$

$$h_1 = \sigma(z_{h_1})$$

$$h_2 = \sigma(z_{h_2})$$

Hidden layer to output layer

$$w_7h_1 + w_9h_2 + b_2 = z_{o_1}$$

$$w_8h_1 + w_{10}h_2 + b_2 = z_{o_2}$$

$$o_1 = \sigma(z_{o_1})$$

$$o_2 = \sigma(z_{o_2})$$

Z stands for net sum of a neuron

Using sigmoidal activation function

$$w_1x_1 + w_3x_2 + w_5x_3 + b_1 = z_{h_1} = 0.1(1) + 0.3(4) + 0.5(5) + 0.5 = 4.3$$

$$h_1 = \sigma(z_{h_1}) = \sigma(4.3) = 0.9866$$

$$w_2x_1 + w_4x_2 + w_6x_3 + b_1 = z_{h_2} = 0.2(1) + 0.4(4) + 0.6(5) + 0.5 = 5.3$$

$$h_2 = \sigma(z_{h_2}) = \sigma(5.3) = 0.9950$$

$$w_7h_1 + w_9h_2 + b_2 = z_{o_1} = 0.7(0.9866) + 0.9(0.9950) + 0.5 = 2.0862$$

$$o_1 = \sigma(z_{o_1}) = \sigma(2.0862) = 0.8896$$

$$w_8h_1 + w_{10}h_2 + b_2 = z_{o_2} = 0.8(0.9866) + 0.1(0.9950) + 0.5 = 1.3888$$

$$o_2 = \sigma(z_{o_2}) = \sigma(1.3888) = 0.8004$$

Backpropagation Calculations

We now define the sum of squares error using the target values and the results from the last layer from forward propagation.

$$E = \frac{1}{2}[(o_1 - t_1)^2 + (o_2 - t_2)^2]$$

$$\frac{dE}{do_1} = o_1 - t_1$$

O_1 actual output; t_1 desired output

$$\frac{dE}{do_2} = o_2 - t_2$$

$$\frac{dE}{dw_7} = \frac{dE}{do_1} \frac{do_1}{dz_{o_1}} \frac{dz_{o_1}}{dw_7}$$

$$\frac{dE}{dw_7} = (o_1 - t_1)(o_1(1 - o_1))h_1$$

$$\frac{dE}{dw_7} = (0.8896 - 0.1)(0.8896(1 - 0.8896))(0.9866)$$

$$\frac{dE}{dw_7} = 0.0765$$

Backpropagation Calculations

$$\frac{dE}{dw_8} = \frac{dE}{do_2} \frac{do_2}{dz_{o2}} \frac{dz_{o2}}{dw_8}$$

$$\frac{dE}{dw_8} = (0.7504)(0.1598)(0.9866)$$

$$\frac{dE}{dw_8} = 0.1183$$

$$\frac{dE}{dw_9} = \frac{dE}{do_1} \frac{do_1}{dz_{o1}} \frac{dz_{o1}}{dw_9}$$

$$\frac{dE}{dw_9} = (0.7896)(0.0983)(0.9950)$$

$$\frac{dE}{dw_9} = 0.0772$$

$$\frac{dE}{dw_{10}} = \frac{dE}{do_2} \frac{do_2}{dz_{o2}} \frac{dz_{o2}}{dw_{10}}$$

$$\frac{dE}{dw_{10}} = (0.7504)(0.1598)(0.9950)$$

$$\frac{dE}{dw_{10}} = 0.1193$$

$$\frac{dE}{db_2} = \frac{dE}{do_1} \frac{do_1}{dz_{o1}} \frac{dz_{o1}}{db_2} + \frac{dE}{do_2} \frac{do_2}{dz_{o2}} \frac{dz_{o2}}{db_2}$$

$$\frac{dE}{db_2} = (0.7896)(0.0983)(1) + (0.7504)(0.1598)(1)$$

$$\frac{dE}{db_2} = 0.1975$$

Backpropagation Calculations

$$\frac{dE}{dw_1} = \frac{dE}{dh_1} \frac{dh_1}{dz_{h_1}} \frac{dz_{h_1}}{dw_1}$$

$$\frac{dE}{dh_1} = \frac{dE}{do_1} \frac{do_1}{dz_{o_1}} \frac{dz_{o_1}}{dh_1} + \frac{dE}{do_2} \frac{do_2}{dz_{o_2}} \frac{dz_{o_2}}{dh_1}$$

$$= (0.7896)(0.0983)(0.7) + (0.7504)(0.1598)(0.8) = 0.1502$$

Backpropagation Calculations

$$\frac{dE}{dw_1} = (0.1502)(0.0132)(1) = 0.0020$$

The calculations for $\frac{dE}{dw_3}$ and $\frac{dE}{dw_5}$ are below

$$\frac{dE}{dw_3} = \frac{dE}{dh_1} \frac{dh_1}{dz_{h_1}} \frac{dz_{h_1}}{dw_3}$$

$$\frac{dE}{dw_3} = (0.1502)(0.0132)(4) = 0.0079$$

$$\frac{dE}{dw_5} = \frac{dE}{dh_1} \frac{dh_1}{dz_{h_1}} \frac{dz_{h_1}}{dw_5}$$

$$\frac{dE}{dw_5} = (0.1502)(0.0132)(5) = 0.0099$$

$$\frac{dE}{dw_2} = \frac{dE}{dh_2} \frac{dh_2}{dz_{h_2}} \frac{dz_{h_2}}{dw_2}$$

$$\frac{dE}{dh_2} = \frac{dE}{do_1} \frac{do_1}{dz_{o_1}} \frac{dz_{o_1}}{dh_2} + \frac{dE}{do_2} \frac{do_2}{dz_{o_2}} \frac{dz_{o_2}}{dh_2}$$

$$\frac{dE}{dh_2} = (0.7896)(0.0983)(0.9) + (0.7504)(0.1598)(0.1) = 0.0818$$

Plugging the above into the formula for $\frac{dE}{dw_2}$, we get

$$\frac{dE}{dw_2} = (0.0818)(0.0049)(1) = 0.0004$$

The calculations for $\frac{dE}{dw_4}$ and $\frac{dE}{dw_6}$ are below

$$\frac{dE}{dw_4} = \frac{dE}{dh_2} \frac{dh_2}{dz_{h_2}} \frac{dz_{h_2}}{dw_4}$$

$$\frac{dE}{dw_4} = (0.0818)(0.0049)(4) = 0.0016$$

$$\frac{dE}{dw_6} = \frac{dE}{dh_2} \frac{dh_2}{dz_{h_2}} \frac{dz_{h_2}}{dw_6}$$

$$\frac{dE}{dw_6} = (0.0818)(0.0049)(5) = 0.0020$$

The final error derivative we have to calculate is $\frac{dE}{db_1}$, which is done next

$$\frac{dE}{db_1} = \frac{dE}{do_1} \frac{do_1}{dz_{o_1}} \frac{dz_{o_1}}{dh_1} \frac{dh_1}{dz_{h_1}} \frac{dz_{h_1}}{db_1} + \frac{dE}{do_2} \frac{do_2}{dz_{o_2}} \frac{dz_{o_2}}{dh_2} \frac{dh_2}{dz_{h_2}} \frac{dz_{h_2}}{db_1}$$

$$\frac{dE}{db_1} = (0.7896)(0.0983)(0.7)(0.0132)(1) + (0.7504)(0.1598)(0.1)(0.0049)(1) = 0.0008$$

$$w_1 := w_1 - \alpha \frac{dE}{dw_1} = 0.1 - (0.01)(0.0020) = 0.1000$$

$$w_2 := w_2 - \alpha \frac{dE}{dw_2} = 0.2 - (0.01)(0.0004) = 0.2000$$

$$w_3 := w_3 - \alpha \frac{dE}{dw_3} = 0.3 - (0.01)(0.0079) = 0.2999$$

$$w_4 := w_4 - \alpha \frac{dE}{dw_4} = 0.4 - (0.01)(0.0016) = 0.4000$$

$$w_5 := w_5 - \alpha \frac{dE}{dw_5} = 0.5 - (0.01)(0.0099) = 0.4999$$

$$w_6 := w_6 - \alpha \frac{dE}{dw_6} = 0.6 - (0.01)(0.0020) = 0.6000$$

$$w_7 := w_7 - \alpha \frac{dE}{dw_7} = 0.7 - (0.01)(0.0765) = 0.6992$$

$$w_8 := w_8 - \alpha \frac{dE}{dw_8} = 0.8 - (0.01)(0.1183) = 0.7988$$

$$w_9 := w_9 - \alpha \frac{dE}{dw_9} = 0.9 - (0.01)(0.0772) = 0.8992$$

$$w_{10} := w_{10} - \alpha \frac{dE}{dw_{10}} = 0.1 - (0.01)(0.1193) = 0.0988$$

$$b_1 := b_1 - \alpha \frac{dE}{db_1} = 0.5 - (0.01)(0.0008) = 0.5000$$

$$b_2 := b_2 - \alpha \frac{dE}{db_2} = 0.5 - (0.01)(0.1975) = 0.4980$$

Learning rate =0.01