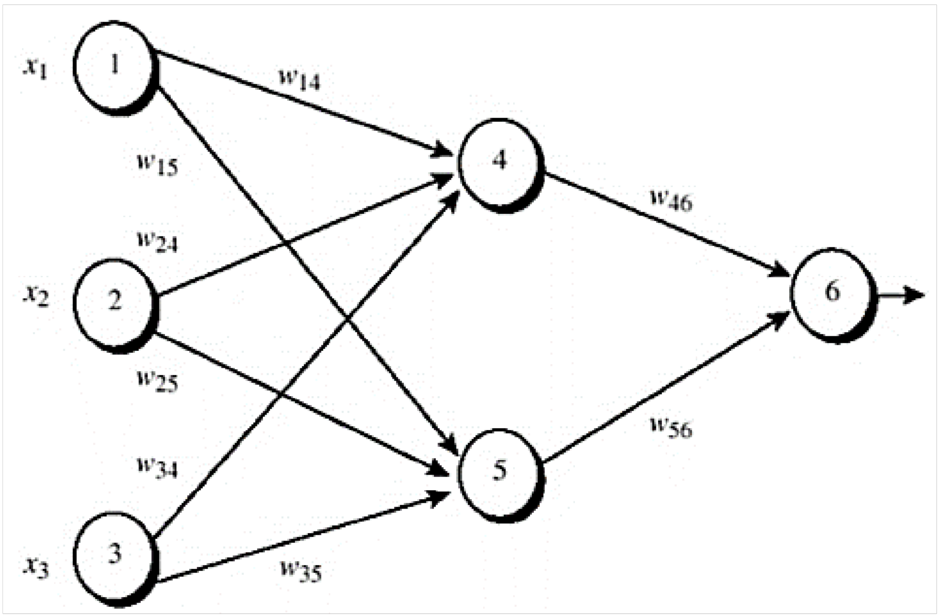
Quiz 8

Problem. Consider the following neuron network, which includes 3 input neurons, 2 hidden neurons and 1 output neurons.



Initial input, weight and bias values are

x_1	x_2	x_3	w_{14}	w_{15}	w_{24}	w_{25}	w_{34}	w_{35}	w_{46}	w_{56}	θ_4	θ_5	θ_6
1	0	1	0.2	−0.3	0.4	0.1	−0.5	0.2	−0.3	−0.2	−0.4	0.2	0.1

The expected output value is 1. The learning rate is 0.9

Knowing that the actual output at some neuron *j* is calculated as follows.

$$y_j(p) = \text{sigmoid}[\sum_{i=1}^{n} x_i(p) \times w_{ij}(p) + \theta_j]$$

where *n* is the number of inputs of neuron *j*, w_{ij} is the corresponding link from a neuron *i* in the previous layer to neuron *j*, and θ_j is the bias at neuron *j*.

Present all calculations required to perform the backpropagation once (i.e., one forward pass and one backward pass) on the given neural network in the following cases

Step 1 – Initialization

Step 2 – Activation

At iteration p:

- Calculate the actual output from n=3 inputs of neuron j in the hidden layer:

$$y_j(p) = sigmoid\left(\sum_{i=1}^{n} x_i(p) w_{ij}(p) + \theta_j\right)$$

- Calculate the actual output from k=2 inputs of neuron m in the hidden layer:

$$y_k(p) = sigmoid(\sum_{j=1}^{2} y_j(p) w_{jk}(p) + \theta_k$$

Step 3 – Weight Training:

The output layer:

- Calculate the error gradient for neuron k in the output layer:

$$\delta_k(p) = y_k(p) * [1 - y_k(p)] * [y_{d,k}(p) - y_k(p)]$$

- Calculate the weight corrections:

$$\Delta w_{jk}(p) = \alpha * y_j(p) * \delta_k(p)$$

- Update the weights at the output neurons:

$$w_{jk}(p+1) = w_{jk}(p) + \Delta w_{jk}(p)$$

The hidden layer:

- Calculate the error gradient for neuron j in the hidden layer:

$$\delta_j(p) = y_j(p) * [1 - y_j(p)] * \sum_{k=1}^{l} \delta_k w_{jk}(p)$$

- Calculate the weight corrections:

$$\Delta w_{ij}(p) = \alpha * x_i(p) * \delta_j(p)$$

- Update the weights at the hidden neurons:

$$w_{ij}(p+1) = w_{ij}(p) + \Delta w_{ij}(p)$$

a) Ignore all biases *(precision to 3 decimal places).*

Ignore all biases – Forward

Output at neuron 4	0.426
Output at neuron 5	0.475
Output at neuron 6	0.527

Ignore all biases – Backward

Error gradient at neuron 6	0.118
Error gradient at neuron 5	-0.006

Error gradient at neuron 4	-0.009
Update w46	-0.255
Update w56	-0.150
Update w14	0.192
Update w15	-0.305
Update w24	0.400
Update w25	0.100
Update w34	-0.508
Update w35	0.195

b) Consider all biases such that each bias is treated as a neuron and thus it will be also updated *(precision to 3 decimal places).*

(Consider all biases – Forward

Output at neuron 4	0.332
Output at neuron 5	0.525
Output at neuron 6	0.552

Consider all biases – Backward

Error gradient at neuron 6	0.111
Error gradient at neuron 5	-0.006
Error gradient at neuron 4	-0.007
Update w46	-0.267
Update w56	-0.148
Update w14	0.193
Update w15	-0.305
Update w24	0.400
Update w25	0.100
Update w34	-0.507
Update w35	0.195
Update bias 6	-0.407
Update bias 5	0.195
Update bias 4	0.200