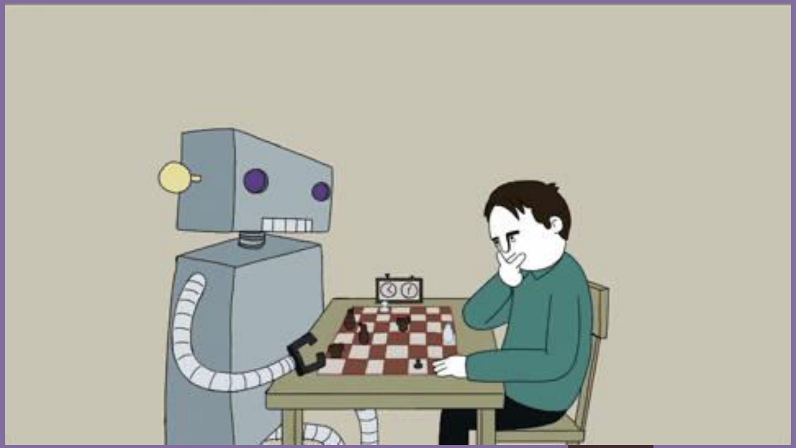
CSE4006 DEEP LEARNING



Dr K G Suma
Associate Professor
School of Computer Science and Engineering

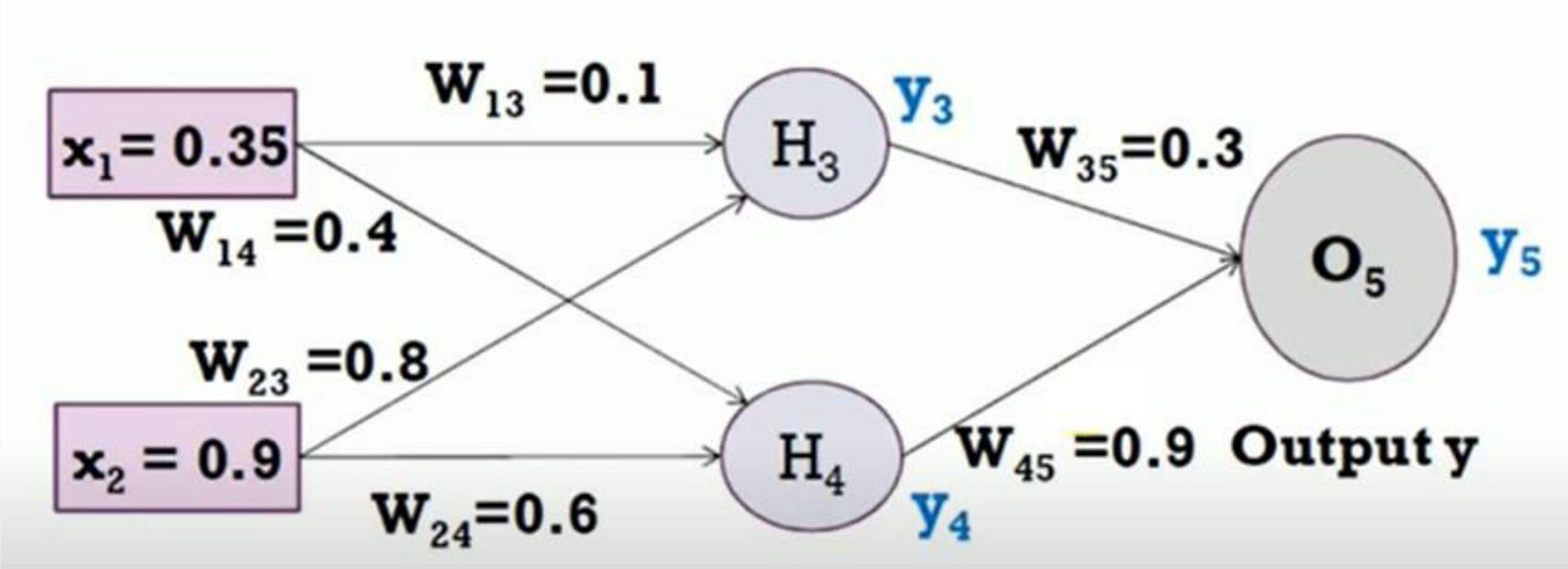
Module No. 2
Practical Deep Networks
8 Hours

- Multilayer Perceptron
- Gradient based Learning
- Backpropagation Algorithm
- Regularization for Deep Learning
- Optimization for training deep models

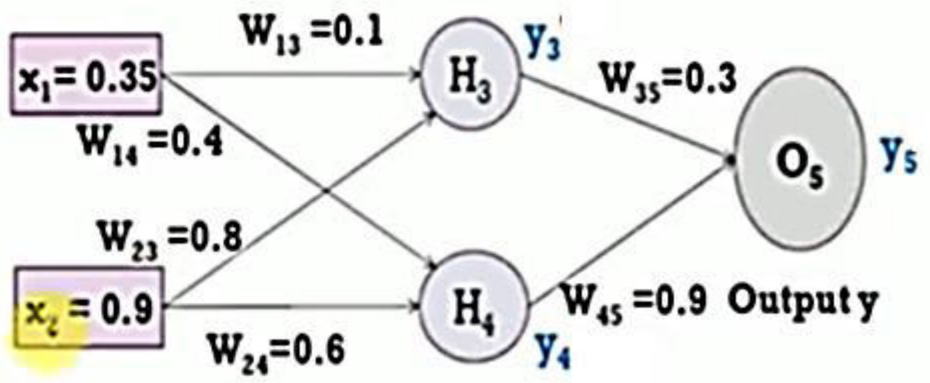
BACKPROPAGATION SOLVED EXERCISE - 1

Backpropagation Algorithm Exercise - 1

- Assume that the neurons have a Sigmoid Activation function, perform a forward pass and a backward pass on the network. Assume that the actual output of y is 0.5 and learning rate is 1. Perform another forward pass.



Backpropagation Algorithm Exercise - 1



- Forward Pass: Compute output for y3, y4 and y5.

$$a_j = \sum_j (w_{i,j} * x_i) \qquad yj = F(aj) = \frac{1}{1 + e^{-a_j}}$$

$$a_1 = (w_{13} * x_1) + (w_{23} * x_2)$$
$$= (0.1 * 0.35) + (0.8 * 0.9) = 0.755$$
$$y_3 = f(a_1) = 1/(1 + e^{-0.755}) = 0.68$$

$$a_2 = (w_{14} * x_1) + (w_{24} * x_2)$$
$$= (0.4 * 0.35) + (0.6 * 0.9) = 0.68$$
$$y_4 = f(a_2) = 1/(1 + e^{-0.68}) = 0.6637$$

$$a_3 = (w_{35} * y_3) + (w_{45} * y_4)$$
$$= (0.3 * 0.68) + (0.9 * 0.6637) = 0.801$$
$$y_5 = f(a_3) = 1/(1 + e^{-0.801}) = \mathbf{0.69} \text{ (Network Output)}$$

$$\text{Error} = y_{target} - y_5 = -0.19$$

0.5 - 0.69

Backpropagation Algorithm Exercise - 1

- Each weight changed by:

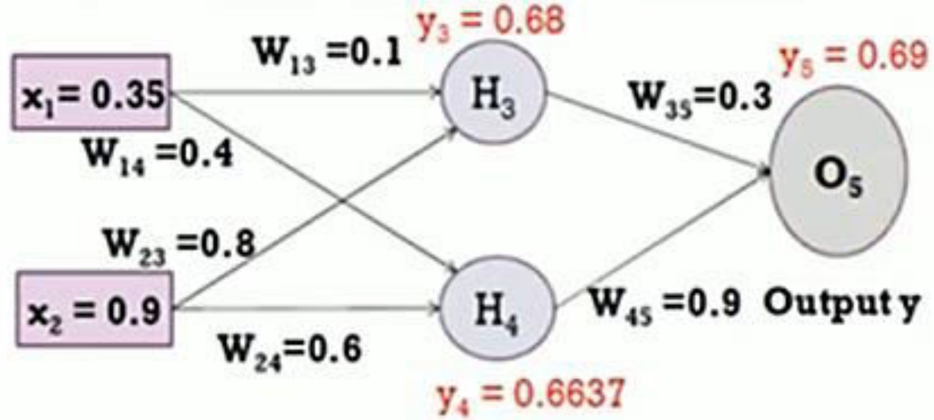
$$\Delta w_{ji} = \eta \delta_j o_i$$

$$\delta_j = o_j(1 - o_j)(t_j - o_j) \qquad \text{if } j \text{ is an output unit}$$

$$\delta_j = o_j(1 - o_j)\sum_k \delta_k w_{kj} \qquad \text{if } j \text{ is a hidden unit}$$

- where η is a constant called the learning rate
- tj is the correct teacher output for unit j
- δj is the error measure for unit j

Backpropagation Algorithm Exercise - 1



- Backward Pass: Compute δ3, δ4 and δ5.

For output unit:

$\delta_5 = y(1-y)\,(y_{target} - y)$
$= 0.69*(1-0.69)*(0.5-0.69) = -0.0406$

For hidden unit:

$\delta_3 = y_3(1-y_3)\, w_{35} * \delta_5$
$= 0.68*(1 - 0.68)*(0.3 * -0.0406) = -0.00265$

$\delta_4 = y_4(1-y_4) w_{45} * \delta_5$
$= 0.6637*(1 - 0.6637)* (0.9 * -0.0406) = -0.0082$

$$\Delta w_{ji} = \eta \delta_j o_i$$

$$\delta_j = o_j(1-o_j)(t_j - o_j) \qquad \text{if } j \text{ is an output unit}$$

$$\delta_j = o_j(1-o_j)\sum_k \delta_k w_{kj} \qquad \text{if } j \text{ is a hidden unit}$$

Backpropagation Algorithm Exercise - 1

Compute new weights

$$\Delta w_{ji} = \eta \delta_j o_i$$

$\Delta w_{45} = \eta\ \delta_5\ y_4 = 1 * -0.0406 * 0.6637 = -0.0269$
$w_{45}\ (\text{new}) = \Delta w_{45} + w_{45}(\text{old}) = -0.0269 + (0.9) = 0.8731$

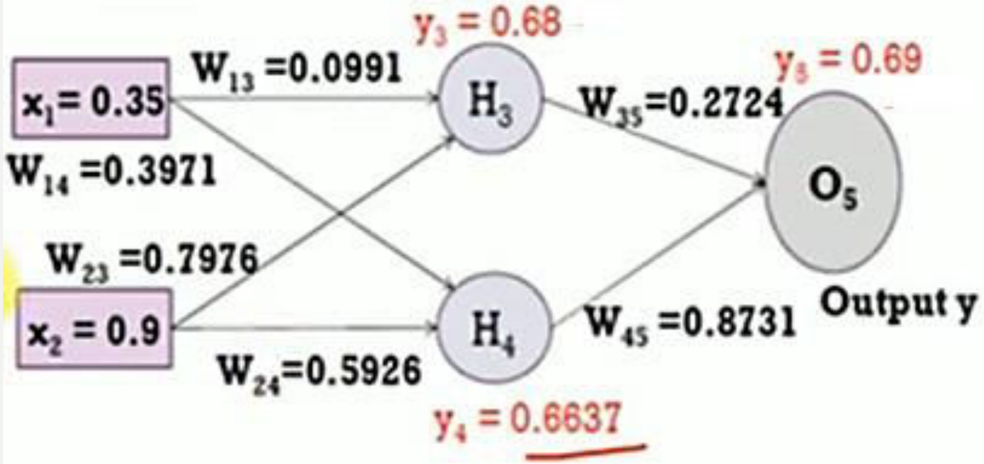
$\Delta w_{14} = \eta\ \delta_4\ x_1 = 1 * -0.0082 * 0.35 = -0.00287$
$w_{14}\ (\text{new}) = \Delta w_{14} + w_{14}(\text{old}) = -0.00287 + 0.4 = 0.3971$

Backpropagation Algorithm Exercise - 1

- Similarly, update all other weights

i	j	w_{ij}	δ_j	x_i	η	Updated w_{ij}
1	3	0.1	-0.00265	0.35	1	0.0991
2	3	0.8	-0.00265	0.9	1	0.7976
1	4	0.4	-0.0082	0.35	1	0.3971
2	4	0.6	-0.0082	0.9	1	0.5926
3	5	0.3	-0.0406	0.68	1	0.2724
4	5	0.9	-0.0406	0.6637	1	0.8731

Backpropagation Algorithm Exercise - 1



Error $= y_{target} - y_5 = -0.182$

- Forward Pass: Compute output for y3, y4 and y5.

$$a_j = \sum_j (w_{i,j} * x_i) \qquad yj = F(aj) = \frac{1}{1 + e^{-a_j}}$$

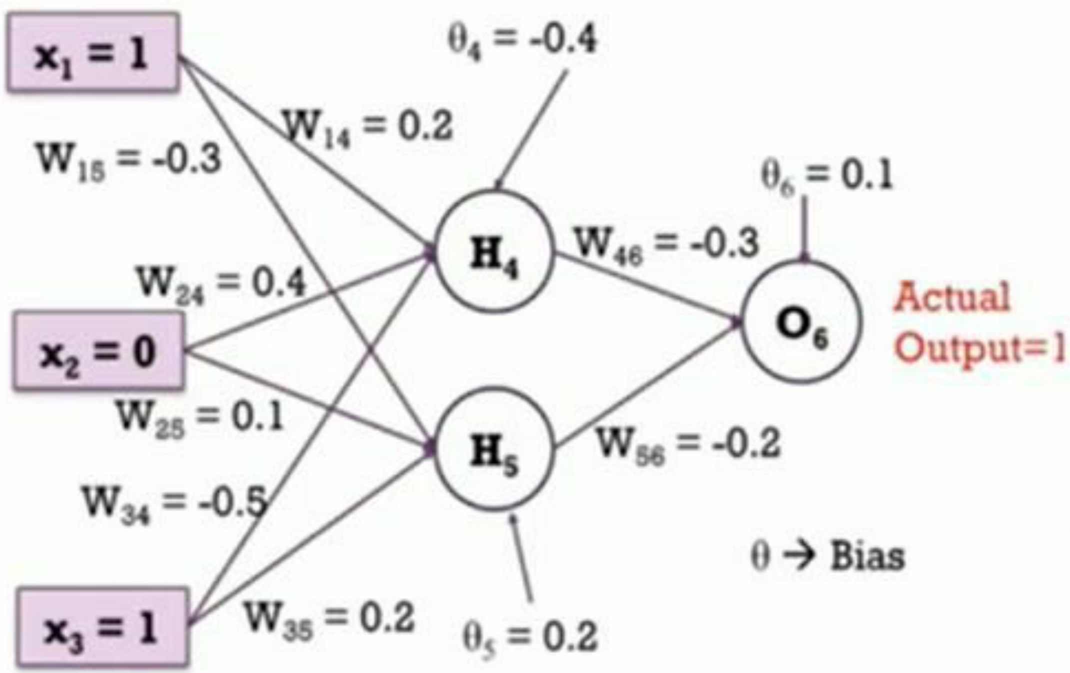
$$a_1 = (w_{13} * x_1) + (w_{23} * x_2)$$
$$= (0.0991 * 0.35) + (0.7976 * 0.9) = 0.7525$$
$$y_3 = f(a_1) = 1/\,(1 + e^{-0.7525}) = 0.6797$$

$$a_2 = (w_{14} * x_1) + (w_{24} * x_2)$$
$$= (0.3971 * 0.35) + (0.5926 * 0.9) = 0.6723$$
$$y_4 = f(a_2) = 1/\,(1 + e^{-0.6723}) = 0.6620$$

$$a_3 = (w_{35} * y_3) + (w_{45} * y_4)$$
$$= (0.2724 * 0.6797) + (0.8731 * 0.6620) = 0.7631$$
$$y_5 = f(a_3) = 1/\,(1 + e^{-0.7631}) = \mathbf{0.6820} \text{ (Network Output)}$$

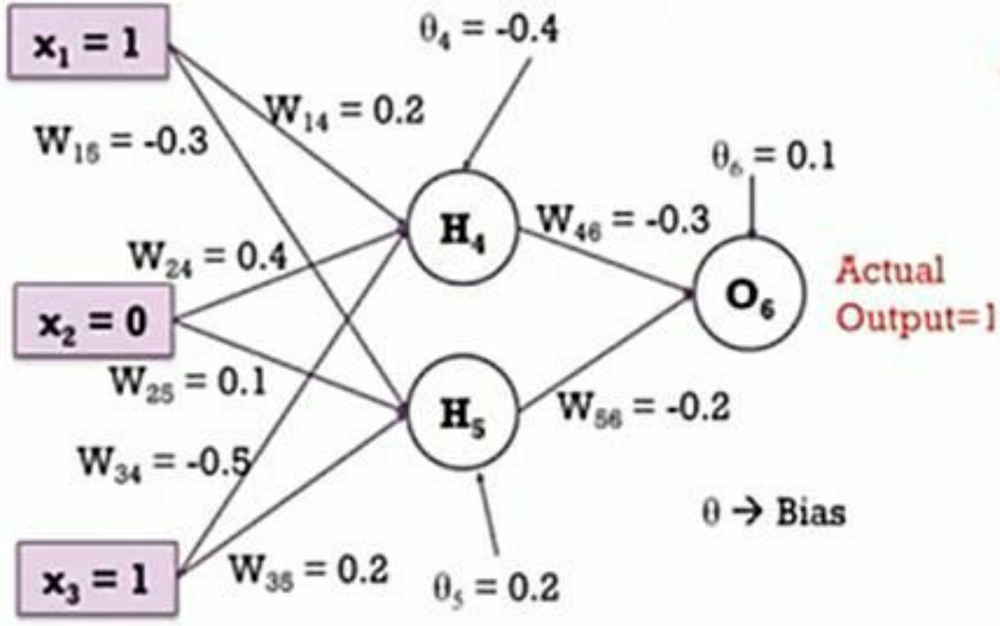
BACKPROPAGATION SOLVED EXERCISE- 2

Backpropagation Algorithm Exercise - 2



Assume that the neurons have a sigmoid activation function, perform a forward pass and a backward pass on the network. Assume that the actual output of y is 1 and learning rate is 0.9. Perform another forward pass.

Backpropagation Algorithm Exercise - 2



Error $= y_{target} - y_6 = 0.526$

- Forward Pass: Compute output for y4, y5 and y6.

$$a_j = \sum_j (w_{i,j} * x_i) \qquad yj = F(aj) = \frac{1}{1 + e^{-a_j}}$$

$$a_4 = (w_{14} * x_1) + (w_{24} * x_2) + (w_{34} * x_3) + \theta_4$$
$$= (0.2 * 1) + (0.4 * 0) + (-0.5 * 1) + (-0.4) = -0.7$$
$$O(H_4) = y_4 = f(a_4) = 1/(1 + e^{0.7}) = 0.332$$

$$a_5 = (w_{15} * x_1) + (w_{25} * x_2) + (w_{35} * x_3) + \theta_5$$
$$= (-0.3 * 1) + (0.1 * 0) + (0.2 * 1) + (0.2) = 0.1$$
$$O(H_5) = y_5 = f(a_5) = 1/(1 + e^{-0.1}) = 0.525$$

$$a_6 = (w_{46} * H_4) + (w_{56} * H_5) + \theta_6$$
$$= (-0.3 * 0.332) + (-0.2 * 0.525) + 0.1 = -0.105$$
$$O(O_6) = y_6 = f(a_6) = 1/(1 + e^{0.105}) = \mathbf{0.474}$$

Backpropagation Algorithm Exercise - 2

- Each weight changed by:

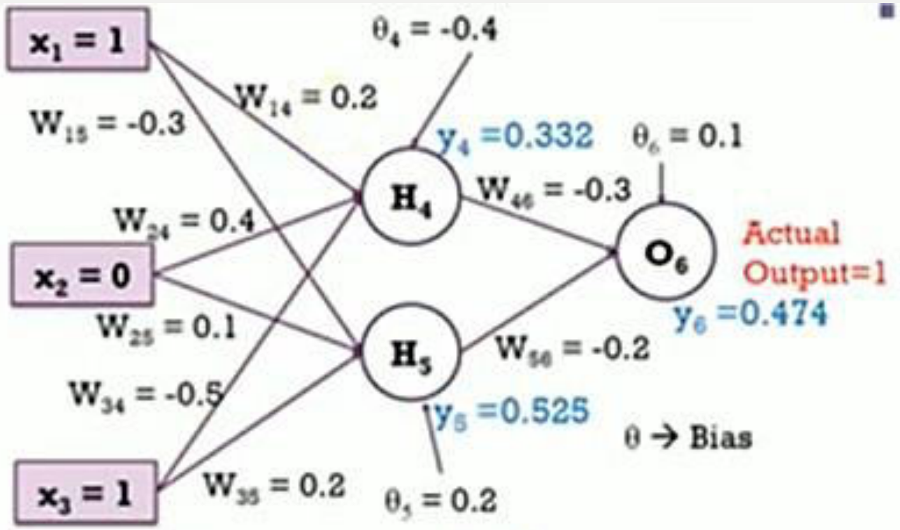
$$\Delta w_{ji} = \eta \delta_j o_i$$

$$\delta_j = o_j(1 - o_j)(t_j - o_j) \qquad \text{if } j \text{ is an output unit}$$

$$\delta_j = o_j(1 - o_j)\sum_k \delta_k w_{kj} \qquad \text{if } j \text{ is a hidden unit}$$

- where η is a constant called the learning rate
- tj is the correct teacher output for unit j
- δj is the error measure for unit j

Backpropagation Algorithm Exercise - 2



- Backward Pass: Compute δ4, δ5 and δ6.

For output unit:

$\delta_6 = y_6(1-y_6)\,(y_{target} - y_6)$

$= 0.474*(1-0.474)*(1-0.474) = 0.1311$

For hidden unit:

$\delta_5 = y_5(1-y_5)\, w_{56} * \delta_6$

$= 0.525*(1 - 0.525)*(-0.2 * 0.1311) = -0.0065$

$= y_4(1-y_4)\, w_{46} * \delta_6$

$= 0.332*(1 - 0.332)* (-0.3 * 0.1331) = -0.0087$

Compute new weights

$$\Delta w_{ji} = \eta \delta_j o_i$$

$\Delta w_{46} = \eta\, \delta_6\, y_4 = 0.9 * 0.1311 * 0.332 = 0.03917$

$w_{46}\ (new) = \Delta w_{46} + w_{46}(old) = 0.03917 + (-0.3) = -0.261$

$\Delta w_{14} = \eta\, \delta_4\, x_1 = 0.9 * -0.0087 * 1 = -0.0078$

$w_{14}\ (new) = \Delta w_{14} + w_{14}(old) = -0.0078 + 0.2 = 0.192$

Backpropagation Algorithm Exercise - 2

Similarly, update all other weights

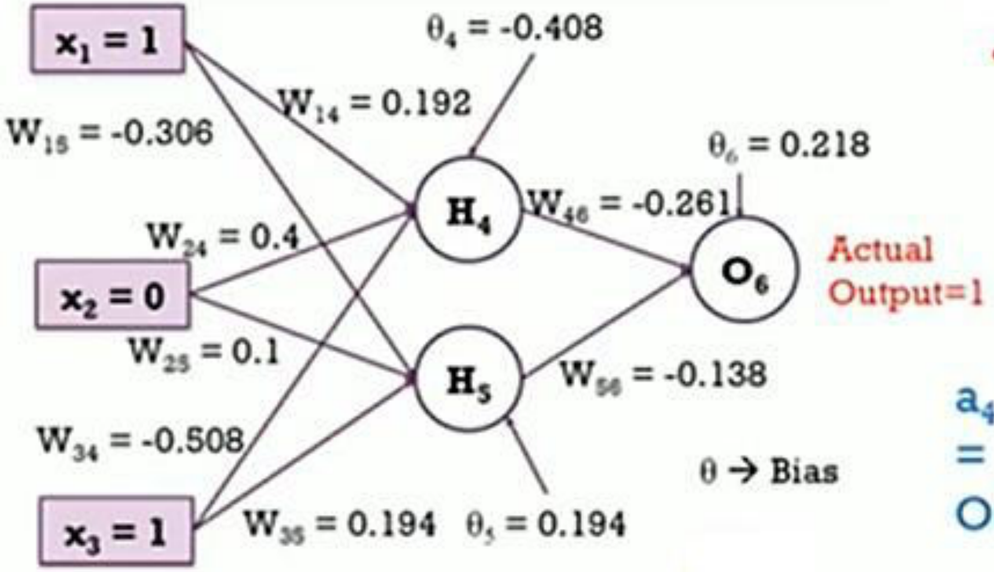
i	j	w_{ij}	δ_j	x_i	η	Updated w_{ij}
4	6	-0.3	0.1311	0.332	0.9	-0.261
5	6	-0.2	0.1311	0.525	0.9	-0.138
1	4	0.2	-0.0087	1	0.9	0.192
1	5	-0.3	-0.0065	1	0.9	-0.306
2	4	0.4	-0.0087	0	0.9	0.4
2	5	0.1	-0.0065	0	0.9	0.1
3	4	-0.5	-0.0087	1	0.9	-0.508
3	5	0.2	-0.0065	1	0.9	0.194

Backpropagation Algorithm Exercise - 2

- Similarly, update bais weights

θ_j	Previous θ_j	δ_j	η	Updated θ_j
Θ_6	0.1	0.1311	0.9	0.218
Θ_5	0.2	-0.0065	0.9	0.194
Θ_4	-0.4	-0.0087	0.9	-0.408

Backpropagation Algorithm Exercise - 2



$\text{Error} = y_{target} - y_6 = 0.485$

- Forward Pass: Compute output for y4, y5 and y6.

$$a_j = \sum_j (w_{i,j} * x_i) \qquad yj = F(aj) = \frac{1}{1 + e^{-a_j}}$$

$a_4 = (w_{14} * x_1) + (w_{24} * x_2) + (w_{34} * x_3) + \theta_4$
$= (0.192 * 1) + (0.4 * 0) + (-0.508 * 1) + (-0.408) = -0.724$
$O(H_4) = y_4 = f(a_4) = 1/\,(1 + e^{0.724}) = 0.327$

$a_5 = (w_{15} * x_1) + (w_{25} * x_2) + (w_{35} * x_3) + \theta_5$
$= (-0.306 * 1) + (0.1 * 0) + (0.194 * 1) + (0.194) = 0.082$
$O(H_5) = y_5 = f(a_5) = 1/\,(1 + e^{-0.082}) = 0.520$

$a_6 = (w_{46} * H_4) + (w_{56} * H_5) + \theta_6$
$= (-0.261 * 0.327) + (-0.138 * 0.520) + 0.218 = 0.061$
$O(O_6) = y_6 = f(a_6) = 1/\,(1 + e^{-0.061}) =$ **0.515** (Network Output)