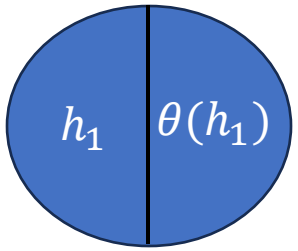
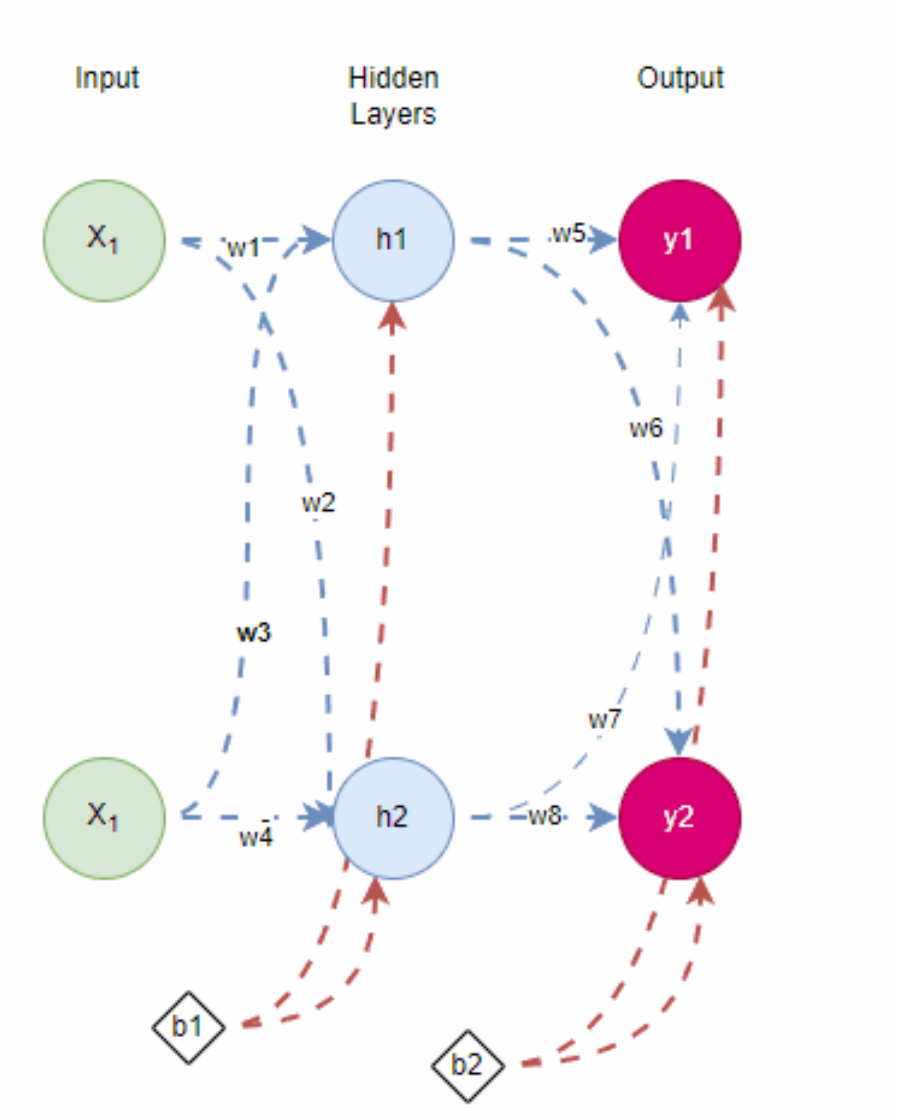
Backpropagation with Example

Network Training





$$h_1 = x_1 * w_1 + x_2 * w_3 + b_1$$

$$sigmoid = \frac{1}{1 + e^{-x}}$$

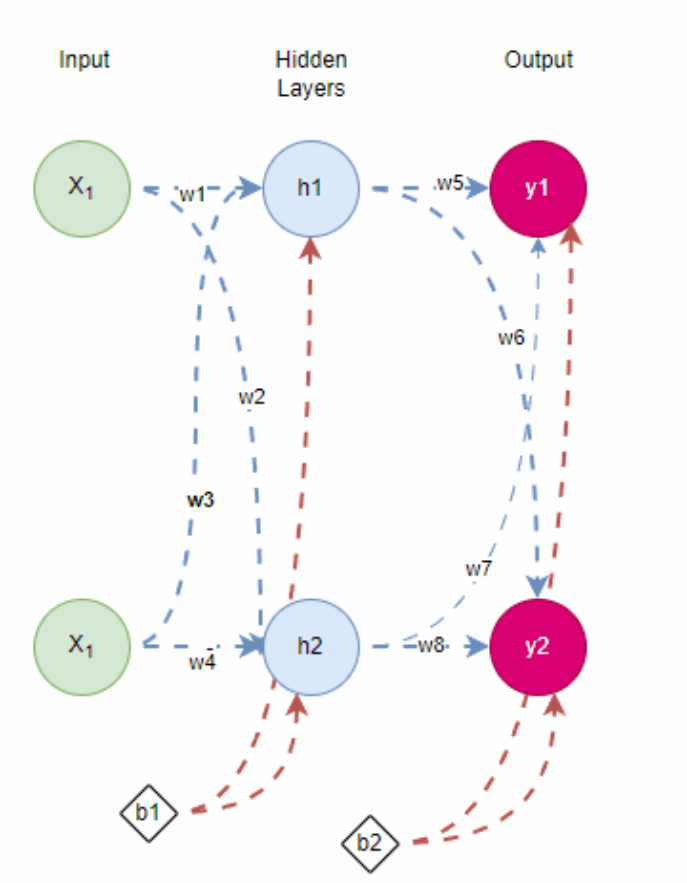
$$\theta(h_1) = \frac{1}{1+e^{-h_1}}$$

Input& weights

X1	**0.09**
X2	**0.1**
W1	0.11
W2	0.12
W3	0.13
W4	0.14
W5	0.15
W6	0.16
W7	0.17
W8	0.18
B1	0.19
b2	0.2

Target values

T1	0.01
T2	0.99

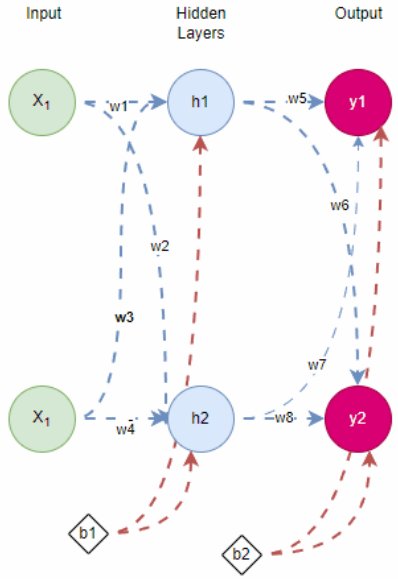


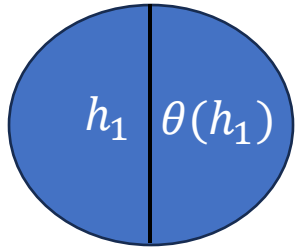
Input& weights

X1	**0.09**
X2	**0.1**
W1	0.11
W2	0.12
W3	0.13
W4	0.14
W5	0.15
W6	0.16
W7	0.17
W8	0.18
B1	0.19
b2	0.2

Target values

T1	0.01
T2	0.99





$h_1 = x_1 * w_1 + x_2 * w_3 + b_1$

h_1 =0.09*.11+0.10*0.13+0.19= 0.2129

$\theta(h_1) = \frac{1}{1+e^{-h_1}}$ = 0.553

$h_2 = x_1 * w_2 + x_2 * w_4 + b_1$

h_2 =0.09*.12+0.10*0.14+0.19= 0.2248

$\theta(h_2) = \frac{1}{1+e^{-h_2}}$ = 0.555

Input& weights

X1	**0.09**
X2	**0.1**
W1	0.11
W2	0.12
W3	0.13
W4	0.14
W5	0.15
W6	0.16
W7	0.17
W8	0.18
B1	0.19
b2	0.2

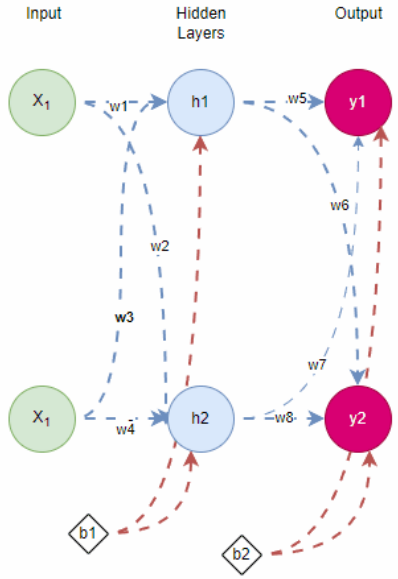
Target values

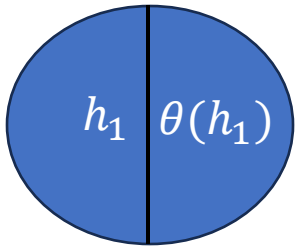
T1	0.01
T2	0.99

$\theta(h_1) = 0.553$

$\theta(h_2) = 0.555$

0.575





$$y_1 = \theta(h_1) * w_5 + \theta(h_2) * w_7 + b_2$$

$y_1 = 0.553 * 0.15 + 0.555 * 0.17 + 0.2 = 0.382$

$\theta(y_1) = 0.594$

$$y_2 = \theta(h_1) * w_6 + \theta(h_2) * w_8 + b_2$$

$y_2 = 0.553 * 0.16 + 0.0.555 * 0.18 + 0.2 = 0.388$

$\theta(y_2) = 0.595$

Loss function

Target values

T1	0.01
T2	0.99

$$Loss = \sum \frac{1}{2}(target - output)^2$$

$\theta(y_1) = 0.594$

$\theta(y_2) = 0.595$

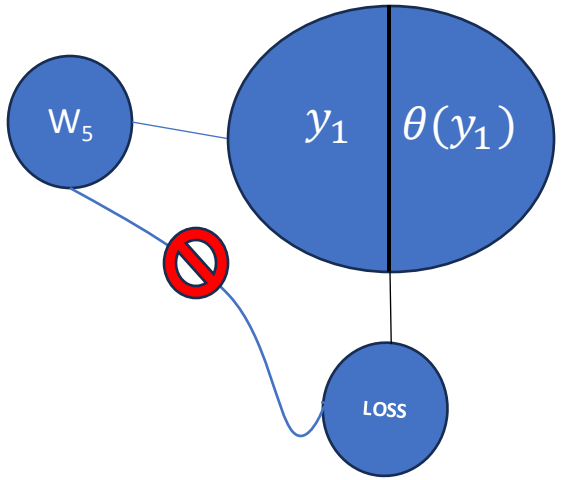
$$Loss = \frac{1}{2}(T1 - outy1)^2 + \frac{1}{2}(T2 - outy2)^2$$

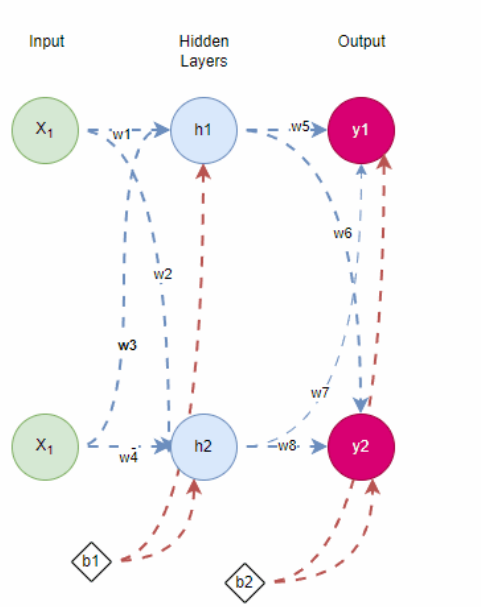
$$Loss = \frac{1}{2}(0.01 - 0.594)^2 + \frac{1}{2}(0.99 - .595)^2$$

$Loss = 0.1676 + 0.0861125 = 0.1676$

Backpropagation- Chain rule - Partial derivative

Error at $W_5 = \frac{\partial_{loss}}{\partial_{w_5}}$





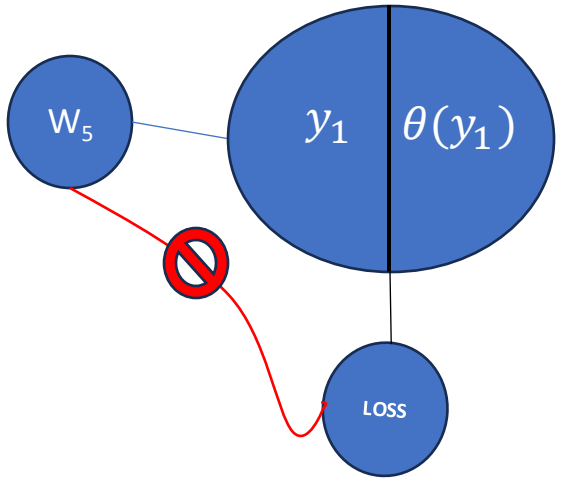
Error at $W_5 = \frac{\partial_{loss}}{\partial_{\theta(y1)}} * \frac{\partial_{\theta(y1)}}{\partial_{y1}} * \frac{\partial_{y1)}}{\partial_{w5}}$

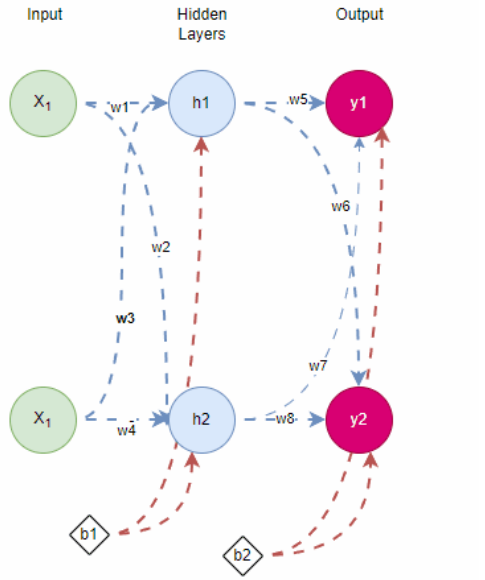
Backpropagation- Chain rule - Partial derivative

Error at W_5= $\frac{\partial_{loss}}{\partial_{\theta(y1)}} * \frac{\partial_{\theta(y1)}}{\partial_{y1}} * \frac{\partial_{y1)}}{\partial_{w5}}$

$$\frac{\partial_{loss}}{\partial_{\theta(y1)}}$$

$$Loss = \frac{1}{2}(T1 - \theta(y1))^2 + \frac{1}{2}(T2 - \theta(y2))^2$$





$$\frac{\partial_{loss}}{\partial_{\theta(y1)}} = 2 * \frac{1}{2}(T1 - \theta(y1))^{2-1} * -1 + 0$$

$$\frac{\partial_{loss}}{\partial_{\theta(y1)}} = -(T1 - outy1)$$

$$\frac{\partial_{loss}}{\partial_{\theta(y1)}} = -(0.01 - 0.594) = 0.584$$

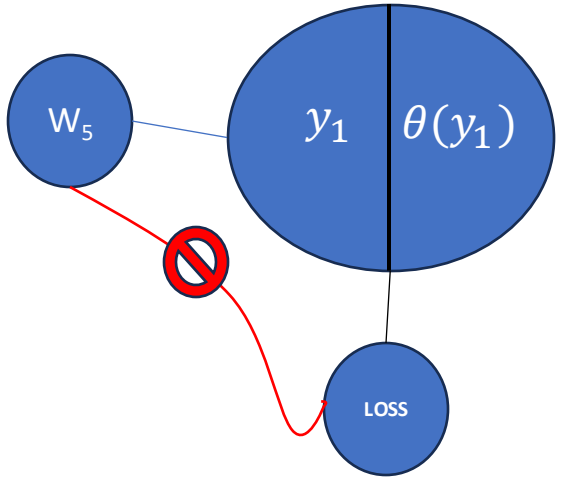
1/2 $[\![(T2 - outy2)]\!]$^2

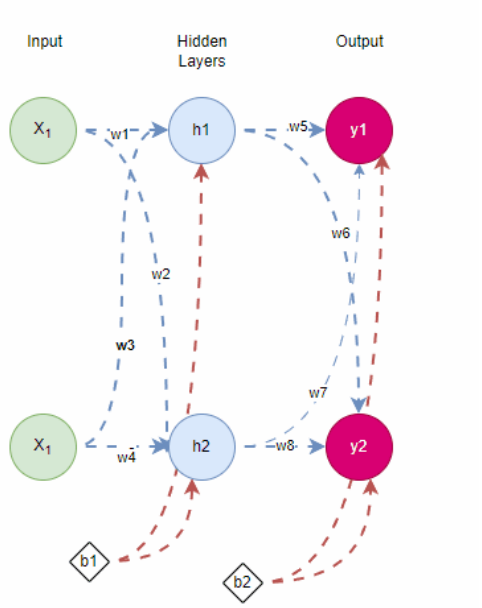
T1	0.01
T2	0.99

$\theta(y_1) = 0.594$

Backpropagation- Chain rule - Partial derivative

Error at $W_5 = \frac{\partial_{loss}}{\partial_{\theta(y1)}} * \frac{\partial_{\theta(y1)}}{\partial_{y1}} * \frac{\partial_{y1)}}{\partial_{w5}}$





$$\theta(y_1) = \frac{1}{1+e^{-y_1}}$$

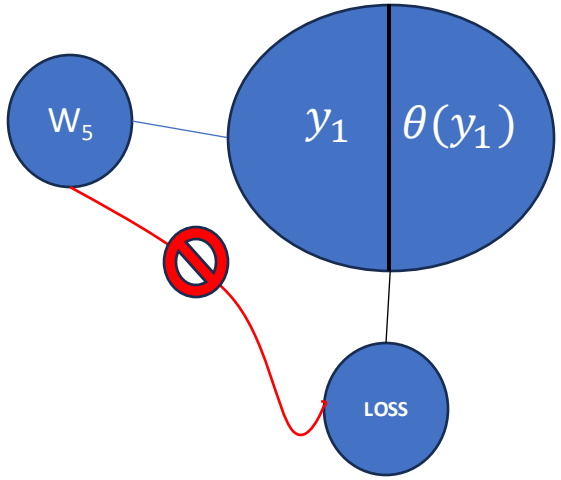
$$\frac{\partial_{\theta(y1)}}{\partial_{y1}} = \theta(y1) * \big(1 - \theta(y1)\big) = 0.594 * (1 - 0.594) = 0.241$$

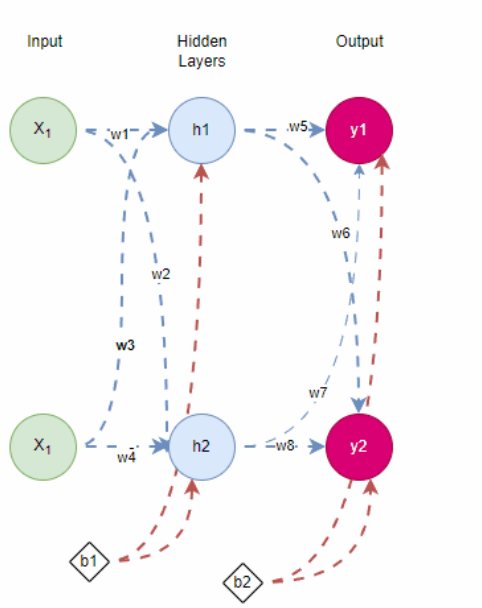
$$y_1 = \theta(h_1) * w_5 + \theta(h_2) * w_7 + b_2$$

$$\frac{\partial_{y1}}{\partial_{w5}} = \theta(h_1) * 1 + 0 + 0 = 0.553$$

Backpropagation- Chain rule - Partial derivative

Error at $W_5 = \frac{\partial_{loss}}{\partial_{\theta(y1)}} * \frac{\partial_{\theta(y1)}}{\partial_{y1}} * \frac{\partial_{y1)}}{\partial_{w5}}$





$\frac{\partial_{loss}}{\partial_{\theta(y1)}} = 0.584$ $\frac{\partial_{\theta(y1)}}{\partial_{y1}} = 0.241$ $\frac{\partial_{y1}}{\partial_{w5}} = 0.553$

Error at $W_5 = 0.584 * 0.241 * 0.553 = 0.077$

$W_{5new} = W_5 - \alpha * 0.77$

α=0.5

$W_{5new} = 0.15 - 0.5 * 0.77 = 0.111$

W5	**0.15**

Input& weights

X1	**0.09**
X2	**0.1**
W1	
W2	
W3	
W4	
W5	**0.11**
W6	
W7	
W8	
B1	
b2	

Target values

T1	0.01
T2	0.99