# Q1: Derivation

Initialization of weights

$Z_1 = W_1 x + b_1$

$a_1 = g(Z_1)$

$Z_2 = W_2 a_1 + b_2$

$a_2 = Z_2$

$\hat{y} = a_2$

$L = 0.5 \sum (\hat{y} - y)^2$

The update formula :-

$$W = W - \alpha \frac{\partial L}{\partial W}$$

$$b = b - \alpha \frac{\partial L}{\partial b}$$

Update Rule :-

$$\frac{\partial L}{\partial W_2} = (\hat{y} - y) \frac{\partial}{\partial W_2} (\hat{y} - y) = (\hat{y} - y) \frac{\partial \hat{y}}{\partial W_2}$$

$$\frac{\partial L}{\partial W_2} = (\hat{y} - y) \frac{\partial a_2}{\partial W_2} = (\hat{y} - y) a_1$$

$$\frac{\partial L}{\partial W_2} = (\hat{y} - y) a_1$$

$$\frac{\partial L}{\partial b_2} = (\hat{y} - y)$$

Back propagation with chain rule :

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a_2} \frac{\partial a_2}{\partial z_2} \frac{\partial z_2}{\partial a_1} \frac{\partial a_1}{\partial z_1} \frac{\partial z_1}{\partial w_1}$$

$$\frac{\partial L}{\partial a_2} \cdot \frac{\partial a_2}{\partial z_2} = a_2 - y$$

$$\frac{\partial z_2}{\partial a_1} = w_2$$

$$\frac{\partial a_1}{\partial z_1} = g'(z_1) = g(z_1)(1 - g(z_1))$$

$$\frac{\partial z_1}{\partial w_1} = x$$

Final update rule for layers, weight & bias :-

$$\frac{\partial L}{\partial w_1} = (a_2 - y)\, w_2\, g(z_1)(1 - g(z_1))\, x$$

$$\frac{\partial L}{\partial b_1} = (a_2 - y)\, w_2\, g(z_1)(1 - g(z_1))$$