Exercise 2:

Given, $E(w, b) = \frac{1}{2} \sum_{k \in N_0} (O_k - t_k)^2$

We take,

output of input Layer $= O_i^I$

output of hidden Layer $= O_j^H$

Bias of hidden Layer $= b_i^H$

Input of hidden Layer

$$z_j^H = \sum_{i=1}^{N^I} w_{ji}^H O_i^I + b_i^H$$

output of output layer $= O_k^O$

Bias of output Layer $= b_k^O$

Input of output Layer

$$z_k^O = \sum_{j=1}^{N^H} w_{kj}^O O_j^H + b_k^O$$

Activation function $= Sig(t)$

and $Sig'(t) = Sig(t)\,(1 - Sig(t))$

So, $O_j^H = Sig(z_j^H)$

and $O_k^O = Sig(z_k^O)$

$$\frac{\partial E}{\partial O_k^o} = \frac{1}{2} \cdot 2 \sum_{k=1}^{N^o} (O_k^o - t_k) = \sum_{k=1}^{N^o} (O_k^o - t_k)$$

$$\frac{\partial O_k^o}{\partial Z_k^o} = \text{sig}'(Z_k^o)$$
$$= \text{sig}(Z_k^o)\left(1 - \text{sig}(Z_k^o)\right)$$

$$\frac{\partial Z_k^o}{\partial O_j^H} = \omega_{kj}^O$$

$$\frac{\partial O_j^H}{\partial Z_j^H} = \text{sig}'(Z_j^H)$$
$$= \text{sig}(Z_j^H)\left(1 - \text{sig}(Z_j^H)\right)$$

$$\frac{\partial Z_j^H}{\partial \omega_{ji}^H} = \omega_{ji}^H$$

Now,

$$\frac{\partial E}{\partial \omega_{ji}^H} = \frac{\partial E}{\partial O_k^o} \cdot \frac{\partial O_k^o}{\partial Z_k^o} \cdot \frac{\partial Z_k^o}{\partial O_j^H} \cdot \frac{\partial O_j^H}{\partial Z_j^H} \cdot \frac{\partial Z_j^H}{\partial \omega_{ji}^H}$$
$$= \sum_{k=1}^{N^o} (O_k^o - t_k) \cdot \text{Sig}(Z_k^o)\left(1 - \text{Sig}(Z_k^o)\right)$$
$$\cdot\, \omega_{kj}^O \cdot \text{sig}(Z_j^H)\left(1 - \text{sig}(Z_j^H)\right)$$
$$\cdot\, \omega_{ji}^H$$

Extra:

$$\frac{\partial E}{\partial w_{kj}^{o}} = \frac{\partial E}{\partial O_k^{o}} \cdot \frac{\partial O_k^{o}}{\partial z_k^{o}} \cdot \frac{\partial z_k^{o}}{\partial w_{kj}^{o}}$$

$$= (O_k^{o} - t_k) \cdot sig(z_k^{o})\,(1 - sig(z_k^{o})) \cdot O_j^{H}$$

$$\frac{\partial E}{\partial O_j^{H}} = \frac{\partial E}{\partial O_k^{o}} \cdot \frac{\partial O_k^{o}}{\partial z_k^{o}} \cdot \frac{\partial z_k^{o}}{\partial O_j^{H}}$$

$$= \sum_{k=1}^{N^{o}} (O_k^{o} - t_k) \cdot sig(z_k^{o})\,(1 - sig(z_k^{o})) \cdot w_{kj}^{o} .$$

$$E = \frac{1}{2}\left(y(x_n, w) - t_n\right)^2$$

$$\boxed{y_k = b_k} \longrightarrow \sum_j w'_{kj} z_j$$

Hidden layer

output: $z_j = h(a_j)$

Input: $a_j = \sum_i w_{ji} x_i$ → Input.

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial z_j} \cdot \frac{\partial z_j}{\partial a_j} \cdot \frac{\partial a_j}{\partial w_{ji}}$$

$$= (y_k - t_n) \cdot \frac{\partial}{\partial z_j}(y_k) \cdot h'(a_j) \quad x_i$$

$$= \sum_k \underbrace{(y_k - t_n)}_{= \delta_k \text{ (given)}} w'_{kj}\, h'(a_j)\, x_i$$

$$= \left(h'(a_j) \sum_k \delta_k w'_{kj}\right) x_i$$