Yilong Wang 483728
Linda Li 458029

Task 1.

We can give z_k into the expression for $\hat{y}$ from (2):

$$\hat{y} = w_0^{(2)} + \sum_{k=1}^{K} w_k^{(2)} \left(w_{0k}^{(1)} + \sum_{n=1}^{N} w_{nk}^{(1)} x_n\right)$$
$$= \left(w_0^{(2)} + \sum_{k=1}^{K} w_k^{(2)} w_{0k}^{(1)}\right) + \sum_{n=1}^{N} \left(\sum_{k=1}^{K} w_k^{(2)} w_{nk}^{(1)}\right) x_n$$

We define new biases $v_0 = w_0^{(2)} + \sum_{k=1}^{K} w_k^{(2)} w_{0k}^{(1)}$ and new weights $v_n = \sum_{k=1}^{K} w_k^{(2)} w_{nk}^{(1)}$

Therefore, the output $\hat{y}$ of the MLP without hidden layers is: $\hat{y} = v_0 + \sum_{n=1}^{N} v_n x_n$

Task 2.

a) $a_k = w_{0k}^{(1)} + \sum_{n=1}^{N} w_{nk}^{(1)} x_n$

b) $\hat{y}_j = w_{0j}^{(2)} + \sum_{k=1}^{K} w_{kj}^{(2)} z_k = w_{0k}^{(2)} + \sum_{k=1}^{K} w_{kj}^{(2)} \cdot \tanh(a_k)$

Task 3.

$$E_n = \frac{1}{2} \sum_{j=1}^{J} (\hat{y}_j - y_j)^2$$

a)

$$\frac{\partial E_n}{\partial w_{kj}^{(2)}} = \frac{\partial E_n}{\partial \hat{y}_j} \cdot \frac{\partial \hat{y}_j}{\partial w_{kj}^{(2)}} = \frac{1}{2}\, 2 \cdot (\hat{y}_j - y_j) \cdot \tanh(a_k)$$

b)

$$\frac{\partial E_n}{\partial w_{nk}^{(1)}} = \sum_{j=1}^{J} \frac{\partial E_n}{\partial \hat{y}_j} \cdot \frac{\partial \hat{y}_j}{\partial z_k} \cdot \frac{\partial z_k}{\partial a_k} \cdot \frac{\partial a_k}{\partial w_{nk}^{(1)}}$$

$$= \sum_{j=1}^{J} (\hat{y}_j - y_j) \cdot w_{kj}^{(2)} \cdot (1 - \tanh^2(a_k)) \cdot x_n$$