CMPT 310 Assignment 4 – Part I **Shan Ying (Felicity) Yang – 301305759**

April 18, 2022

a. $E_n = \frac{1}{2}(y(x_n, w) - t_n)^2 = \frac{1}{2}\left(z_1^{(4)} - t_n\right)^2$

$z_1^{(4)} = h\left(a_1^{(4)}\right), recall: h(a) = a, \frac{\partial h(a)}{\partial(a)} = 1$

$\frac{\partial E_n}{\partial z_1^{(4)}} = \frac{\partial}{\partial z_1^{(4)}}\left(\frac{1}{2}\left(z_1^{(4)} - t_n\right)^2\right) = \frac{1}{2}\frac{\partial}{\partial z_1^{(4)}}\left(z_1^{(4)} - t_n\right)^2 = \frac{1}{2}2\left(z_1^{(4)} - t_n\right)(1)$

$\frac{\partial E_n}{\partial z_1^{(4)}} = z_1^{(4)} - t_n$

$\frac{\partial E_n}{\partial a_1^{(4)}} = \frac{\partial z_1^{(4)}}{\partial a_1^{(4)}}\frac{\partial E_n}{\partial z_1^{(4)}} = \frac{\partial h\left(a_1^{(4)}\right)}{\partial\left(a_1^{(4)}\right)}\left(z_1^{(4)} - t_n\right) = (1)\left(z_1^{(4)} - t_n\right)$

$\therefore \boldsymbol{\frac{\partial E_n}{\partial a_1^{(4)}} = z_1^{(4)} - t_n \equiv \delta_1^{(4)}}$

b. $\frac{\partial E_n(w)}{\partial w_{12}^{(3)}} = \frac{\partial E_n(w)}{\partial a_1^{(4)}}\frac{\partial a_1^{(4)}}{\partial w_{12}^{(3)}} = \delta_1^{(4)} \cdot \frac{\partial a_1^{(4)}}{\partial w_{12}^{(3)}} = \delta_1^{(4)} \cdot \frac{\partial\left[w_{11}^{(3)}z_1^{(3)} + w_{12}^{(3)}z_2^{(3)} + w_{13}^{(3)}z_3^{(3)}\right]}{\partial w_{12}^{(3)}}$

$\therefore \boldsymbol{\frac{\partial E_n(w)}{\partial w_{12}^{(3)}} = \delta_1^{(4)} \cdot z_2^{(3)}}$

c. $\frac{\partial E_n(w)}{\partial a_1^{(3)}} = \frac{\partial E_n(w)}{\partial a_1^{(4)}}\frac{\partial a_1^{(4)}}{\partial a_1^{(3)}} = \delta_1^{(4)} \cdot \frac{\partial a_1^{(4)}}{\partial z_1^{(3)}} \cdot \frac{\partial z_1^{(3)}}{\partial a_1^{(4)}}$

$\frac{\partial E_n(w)}{\partial a_1^{(3)}} = \delta_1^{(4)} \cdot \frac{\partial\left[w_{11}^{(3)}z_1^{(3)} + w_{12}^{(3)}z_2^{(3)} + w_{13}^{(3)}z_3^{(3)}\right]}{\partial z_1^{(3)}} \cdot \frac{\partial z_1^{(3)}}{\partial a_1^{(4)}} = \delta_1^{(4)} \cdot w_{11}^{(3)} \cdot \frac{\partial z_1^{(3)}}{\partial a_1^{(4)}}$

$recall\ that\ h(a)\ for\ hidden\ layers\ is: \frac{1}{1+e^{-x}}, \frac{\partial h(a)}{\partial(a)} = h(a)\big(1 - h(a)\big)$

$\frac{\partial E_n(w)}{\partial a_1^{(3)}} = \delta_1^{(4)} \cdot w_{11}^{(3)} \cdot \frac{\partial h\left(a_1^{(3)}\right)}{\partial a_1^{(3)}} = \delta_1^{(4)} \cdot w_{11}^{(3)} \cdot \left(h\left(a_1^{(3)}\right)\left(1 - h\left(a_1^{(3)}\right)\right)\right)$

$\because \frac{\partial h\left(a_1^{(3)}\right)}{\partial\left(a_1^{(3)}\right)} = h\left(a_1^{(3)}\right)\left(1 - h\left(a_1^{(3)}\right)\right) = \begin{cases} 0 & for\ x = 0 \\ h\left(a_1^{(3)}\right)\left(1 - h\left(a_1^{(3)}\right)\right) & otherwise \end{cases}$

$\therefore \boldsymbol{\frac{\partial E_n(w)}{\partial a_1^{(3)}} = \delta_1^{(4)} \cdot w_{11}^{(3)} \cdot \left(h\left(a_1^{(3)}\right)\left(1 - h\left(a_1^{(3)}\right)\right)\right) = \delta_1^{(3)}}$

d. $\frac{\partial E_n(w)}{\partial w_{11}^{(2)}} = \frac{\partial E_n(w)}{\partial a_1^{(3)}} \cdot \frac{\partial a_1^{(3)}}{\partial w_{11}^{(2)}} = \delta_1^{(3)} \cdot \frac{\partial a_1^{(3)}}{\partial w_{11}^{(2)}}$

$\frac{\partial E_n(w)}{\partial w_{11}^{(2)}} = \frac{\partial E_n(w)}{\partial a_1^{(3)}} \cdot \frac{\partial\left[w_{11}^{(2)}z_1^{(2)} + w_{12}^{(2)}z_2^{(2)} + w_{13}^{(2)}z_3^{(2)}\right]}{\partial w_{11}^{(2)}} = \frac{\partial E_n(w)}{\partial a_1^{(3)}} \cdot z_1^{(2)}$

$\therefore \boldsymbol{\frac{\partial E_n}{\partial w_{11}^{(2)}} = \delta_1^{(4)} \cdot w_{11}^{(3)} \cdot \left(h\left(a_1^{(3)}\right)\left(1 - h\left(a_1^{(3)}\right)\right)\right) \cdot z_1^{(2)}}$

e. $\frac{\partial E_n(w)}{\partial a_1^{(2)}} = \delta_1^{(2)} = \sum_k \frac{\partial E_n(w)}{\partial a_k^{(3)}} \cdot \frac{\partial a_k^{(3)}}{\partial a_1^{(2)}} = \frac{\partial E_n(w)}{\partial a_1^{(3)}} \cdot \frac{\partial a_1^{(3)}}{\partial z_1^{(2)}} \cdot \frac{\partial z_1^{(2)}}{\partial a_1^{(2)}} + \frac{\partial E_n(w)}{\partial a_2^{(3)}} \cdot \frac{\partial a_2^{(3)}}{\partial z_2^{(2)}} \cdot \frac{\partial z_2^{(2)}}{\partial a_2^{(2)}} + \frac{\partial E_n(w)}{\partial a_3^{(3)}} \cdot \frac{\partial a_3^{(3)}}{\partial z_3^{(2)}} \cdot \frac{\partial z_3^{(2)}}{\partial a_3^{(2)}}$

$\frac{\partial E_n(w)}{\partial a_1^{(2)}} = \delta_1^{(3)} \cdot \frac{\partial\left[w_{11}^{(2)}z_1^{(2)} + w_{12}^{(2)}z_2^{(2)} + w_{13}^{(2)}z_3^{(2)}\right]}{\partial z_1^{(3)}} + \delta_2^{(3)} \cdot \frac{\partial\left[w_{21}^{(2)}z_1^{(2)} + w_{22}^{(2)}z_2^{(2)} + w_{23}^{(2)}z_3^{(2)}\right]}{\partial z_1^{(3)}} + \delta_3^{(3)} \cdot \frac{\partial\left[w_{31}^{(2)}z_1^{(2)} + w_{32}^{(2)}z_2^{(2)} + w_{33}^{(2)}z_3^{(2)}\right]}{\partial z_1^{(3)}}$

$\frac{\partial E_n(w)}{\partial a_1^{(2)}} = \delta_1^{(3)} \cdot w_{11}^{(2)} + \delta_2^{(3)} \cdot w_{21}^{(2)} + \delta_3^{(3)} \cdot w_{31}^{(2)}$

$\therefore \boldsymbol{\frac{\partial E_n}{\partial a_1^{(2)}} = \delta_1^{(2)} = \sum_{k=1}^{3} \delta_k^{(3)} \cdot w_{k1}^{(2)} = \sum_k \delta_k^{(3)} \frac{\partial a_k^{(3)}}{\partial a_1^{(2)}}}$

f. $\frac{\partial E_n(w)}{\partial w_{11}^{(1)}} = \frac{\partial E_n(w)}{\partial a_1^{(2)}} \cdot \frac{\partial a_1^{(2)}}{\partial w_{11}^{(1)}} = \delta_1^{(2)} \cdot \frac{\partial h\left(w_{11}^{(1)}z_1^{(1)} + w_{12}^{(1)}z_2^{(1)} + w_{13}^{(1)}z_3^{(1)}\right)}{\partial w_{11}^{(1)}}$

$\boldsymbol{\frac{\partial E_n(w)}{\partial w_{11}^{(1)}} = \delta_1^{(2)} \cdot z_1^{(1)} \cdot h\left(w_{11}^{(1)}z_1^{(1)} + w_{12}^{(1)}z_2^{(1)} + w_{13}^{(1)}z_3^{(1)}\right)\left(1 - h\left(w_{11}^{(1)}z_1^{(1)} + w_{12}^{(1)}z_2^{(1)} + w_{13}^{(1)}z_3^{(1)}\right)\right)}$

$where\ h(a)\ is\ the\ logistic\ function, h(a) = \frac{1}{1+e^{-x}}$