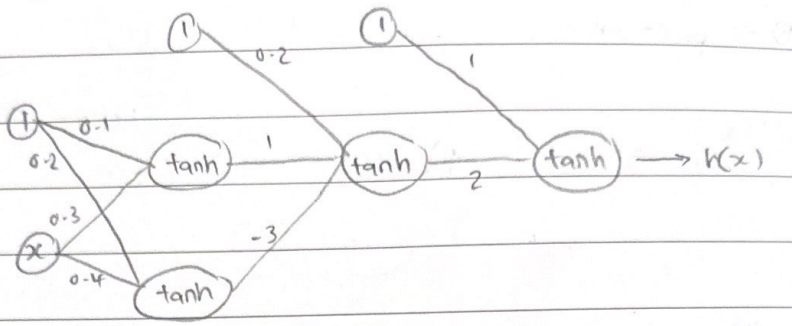
Class activity 3 - Neural Network Example

$x = 2,\ y = 1$



Forward - propagation

$$W^{(1)}_{2\times 2} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix} \qquad W^{(2)} = \begin{bmatrix} 0.2 \\ 1 \\ -3 \end{bmatrix} \qquad W^{(3)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$x^{(0)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$s^{(1)} = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 0.7 \\ 1.0 \end{bmatrix}$$

$$x^{(1)} = \begin{bmatrix} 1 \\ \tanh(0.7) \\ \tanh(1.0) \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 0.6 \\ 0.76 \end{bmatrix}$$

$$s^{(2)} = w^{(2)^T}(x^{(1)})$$

$$= \begin{bmatrix} 0.2 & 1 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 0.6 \\ 0.76 \end{bmatrix}$$

$$= -1.48$$

$$x^{(2)} = \begin{bmatrix} 1 \\ \tanh(-1.48) \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ -0.9 \end{bmatrix}$$

$$s^{(3)} = w^{(3)^T}(x^{(2)})$$

$$= \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -0.9 \end{bmatrix}$$

$$= -0.8$$

$$x^{(3)} = \begin{bmatrix} 1 \\ \tanh(-0.8) \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ -0.664 \end{bmatrix}$$

Back-propagation

$$\frac{\partial E(x^{(L)})}{\partial x^{(L)}} = 2(x^{(L)} - y) \qquad \frac{\partial \tanh(s^{(L)})}{\partial (s^{(L)})} = 1 - \tanh^2(s^{(L)})$$

$$\delta^{(3)} = 2(-0.664 - 1)(1 - \tanh^2(-0.8))$$
$$= 2(-1.664)(1 - 0.441)$$
$$= -1.86$$

$$\delta^{(2)} = (1 - \tanh^2(-1.48)) \otimes (2 \cdot \delta^{(3)})$$
$$= (1 - 0.813) \cdot 2(-1.86)$$
$$= -0.68$$

$$\delta^{(1)} = \begin{bmatrix} 1 - \tanh^2(0.7) \\ 1 - \tanh^2(1.0) \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -3 \end{bmatrix} (-0.68) \qquad (w^{(2)}, \delta^{(2)})$$

$$= \begin{bmatrix} 1 - 0.365 \\ 1 - 0.58 \end{bmatrix} \begin{bmatrix} -0.68 \\ 2.04 \end{bmatrix}$$

$$= \begin{bmatrix} -0.4318 \\ 0.8568 \end{bmatrix}$$

$$G^{(3)} = x^{(2)} \delta^{(3)}$$
$$= \begin{bmatrix} 1 \\ 0.9 \end{bmatrix} -1.86$$
$$= \begin{bmatrix} -1.86 \\ -1.674 \end{bmatrix}$$

$$G^{(2)} = x^{(1)} \delta^{(2)}$$
$$= \begin{bmatrix} 1 \\ 0.6 \\ 0.76 \end{bmatrix} -0.68$$
$$= \begin{bmatrix} -0.68 \\ -0.408 \\ -0.5168 \end{bmatrix}$$

$$G^{(1)} = x^{(0)} \delta^{(1)}$$
$$= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} -0.4318 & 0.8568 \end{bmatrix}$$
$$= \begin{bmatrix} -0.4318 & 0.8568 \\ -0.8636 & 1.7136 \end{bmatrix}$$