Homework 6

Tora Moralos

1- Feature dimension: $p = 2$

2- $h_1 = \sigma(z_1)$, $\sigma(z) = \frac{1}{1+e^{-z}}$ $\quad z_1 = 0.15 \cdot 0.05 + 0.20 \cdot 0.10 + 0.35$

$h_1 = 0.593$

$h_2 = \sigma(z_2)$, $z_2 = 0.25 \cdot 0.05 + 0.30 \cdot 0.10 + 0.35$

$h_2 = 0.597$

$o_1 = \sigma(h_1 \cdot 0.40 + h_2 \cdot 0.45 + 0.60) = 0.7513$

$o_2 = \sigma(h_1 \cdot 0.50 + h_2 \cdot 0.55 + 0.60) = 0.7729$

3. Squared error $= (0.01 - o_1)^2 + (0.99 - o_2)^2 = 0.597$

Squared root $(e) = \sqrt{0.597} = 0.7725$

4- $f = (o_1 - \hat{o}_1)^2 + (o_2 - \hat{o}_2)^2$ $\qquad n^{(in)} \rightarrow \bigcirc \rightarrow n^{(out)}$

$$\frac{\partial f}{\partial w_5} = \frac{\partial f}{\partial o_1^{out}} \cdot \frac{\partial o_1^{out}}{\partial o_1^{in}} \cdot \frac{\partial o_1^{in}}{\partial w_5}$$

$$= -2(o_1 - \hat{o}_1^{out})\, \hat{o}_1^{out}(1 - \hat{o}_1^{out})\, h_1^{out} = 0.2770 \cdot h_1^{out} = 0.1643$$

$$\frac{\partial f}{\partial w_6} = -2(o_1 - \hat{o}_1^{out})\, \hat{o}_1^{out}(1 - \hat{o}_1^{out})\, h_2^{out} = 0.2770 \cdot h_2^{out} = 0.1653$$

$$\frac{\partial f}{\partial w_7} = \frac{\partial f}{\partial o_2^{out}} \cdot \frac{\partial o_2^{out}}{\partial o_2^{in}} \cdot \frac{\partial o_2^{in}}{\partial w_7}$$

$$= -2(o_2 - \hat{o}_2^{out})\, \hat{o}_2^{out}(1 - \hat{o}_2^{out})\, h_1^{out} = -0.0762\, h_1^{out} = -0.0452$$

$$\frac{\partial f}{\partial w_8} = -2(o_2 - \hat{o}_2^{out})\, \hat{o}_2^{out}(1 - \hat{o}_2^{out})\, h_2^{out} = -0.0762\, h_2^{out} = -0.045$$

$$\frac{\partial f}{\partial w_1} = \frac{\partial f^{out}}{\partial o_1^{out}} \cdot \frac{\partial o_1^{out}}{\partial o_1^{in}} \cdot \frac{\partial o_1^{in}}{\partial h_1^{out}} \cdot \frac{\partial h_1^{out}}{\partial h_1^{in}} \cdot \frac{\partial h_1^{in}}{\partial w_1}$$

$$+ \frac{\partial f}{\partial o_2^{out}} \cdot \frac{\partial o_2^{out}}{\partial o_2^{in}} \cdot \frac{\partial o_2^{in}}{\partial h_1^{out}} \cdot \frac{\partial h_1^{out}}{\partial h_1^{in}} \cdot \frac{\partial h_1^{in}}{\partial w_1}$$

$$= -2(o_1 - \hat{o}_1^{out})\, \hat{o}_1^{out}(1 - \hat{o}_1^{out}) \cdot w_5\, h_1^{out}(1 - h_1^{out})\, i_1$$

$$+ (-2(o_2 - \hat{o}_2^{out})\, \hat{o}_2^{out}(1 - \hat{o}_2^{out})\, w_7\, h_1^{out}(1 - h_1^{out})\, i_1)$$

$$= 0.000877$$

$$\frac{\partial f}{\partial w_2} = \frac{\partial o_1^{out}}{\partial o_1^{in}} \cdot \frac{\partial o_1^{in}}{\partial h_1^{out}} \cdot \frac{\partial h_1^{out}}{\partial h_1^{in}} \cdot \frac{\partial h_1^{in}}{\partial w_2} + \frac{\partial o_2^{out}}{\partial o_2^{in}} \cdot \frac{\partial o_2^{in}}{\partial h_1^{out}} \cdot \frac{\partial h_1^{out}}{\partial h_1^{in}} \cdot \frac{\partial h_1^{in}}{\partial w_2}$$

$$= -2(o_1 - \hat{o}_1^{out}) \cdot \hat{o}_1^{out}(1 - o_1^{out}) \cdot w_5 \, h_1^{out}(1 - h_1^{out}) \, i_2$$

$$+ (-2(o_2 - \hat{o}_2^{out}) \hat{o}_2^{out}(1 - \hat{o}_2^{out}) \, w_7 \, h_1^{out}(1 - h_1^{out}) \, i_2)$$

$$= \boxed{0.0017}$$

$$\frac{\partial f}{\partial w_3} = \frac{\partial o_1^{out}}{\partial o_1^{in}} \cdot \frac{\partial o_1^{in}}{\partial h_2^{out}} \cdot \frac{\partial h_2^{out}}{\partial h_2^{in}} \cdot \frac{\partial h_2^{in}}{\partial w_3} + \frac{\partial o_2^{out}}{\partial o_2^{in}} \cdot \frac{\partial o_2^{in}}{\partial h_2^{out}} \cdot \frac{\partial h_2^{out}}{\partial h_2^{in}} \cdot \frac{\partial h_2^{in}}{\partial w_3}$$

$$= -2(o_1 - \hat{o}^{out}) \hat{o}_1^{out}(1 - \hat{o}_1^{out}) \cdot w_6 \, h_2^{out}(1 - h_2^{out}) \, i_1$$

$$+ (-2(o_2 - \hat{o}_2^{out}) \hat{o}_2^{out}(1 - \hat{o}_2^{out}) \, w_8 \cdot h_2^{out}(1 - h_2^{out}) \, i_1)$$

$$= \boxed{0.000995}$$

$$\frac{\partial f}{\partial w_4} = \frac{\partial o_1^{out}}{\partial o_1^{in}} \cdot \frac{\partial o_1^{in}}{\partial h_2^{out}} \cdot \frac{\partial h_2^{out}}{\partial h_2^{in}} \cdot \frac{\partial h_2^{in}}{\partial w_4} + \frac{\partial o_2^{out}}{\partial o_2^{in}} \cdot \frac{\partial o_2^{in}}{\partial h_2^{out}} \cdot \frac{\partial h_2^{out}}{\partial h_2^{in}} \cdot \frac{\partial h_2^{in}}{\partial w_4}$$

$$= -2(o_1 - \hat{o}_1^{out}) \hat{o}_1^{out}(1 - \hat{o}_1^{out}) \, w_6 \cdot h_2^{out}(1 - h_2^{out}) \, i_2$$

$$+ (-2(o_2 - \hat{o}_2^{out}) \hat{o}_2^{out}(1 - \hat{o}_2^{out}) \, w_8 \, h_2^{out}(1 - h_2^{out}) \, i_2)$$

$$= \boxed{0.00199}$$

5- $$\boxed{w_i^{(1)} = w_i^{(0)} - \tau \nabla \frac{\partial f}{\partial w_i}}$$

$w_1^{(1)} = 0.15 - \tau(0.0008)$

$w_2^{(1)} = 0.20 - \tau(0.0017)$

$w_3^{(1)} = 0.25 - \tau(0.00099)$

$w_4^{(1)} = 0.30 - \tau(0.00199)$

$w_5^{(1)} = 0.40 - \tau(0.1648)$

$w_6^{(1)} = 0.45 - \tau(0.1653)$

$w_7^{(1)} = 0.50 + \tau(0.0452)$

$w_8^{(1)} = 0.55 + \tau(0.0455)$