Homework 7

(1.1) $W = [1, 1, -1, 0.5, 1, 2]^T$

ReLU hidden layer activation f-n
Sigmoid output layer act. f-n

$x = 4$, $y = 0$

$w_1 = 1$, $w_2 = 1$, $w_3 = -1$, $w_4 = 0.5$, $w_5 = 1$, $w_6 = 2$

a)

$z_1 = x w_1 = 4 \times 1 = 4$

$z_2 = x w_2 = 4 \times 1 = 4$

$z_3 = x w_3 = 4 \times (-1) = -4 \Rightarrow 0$ because $z_3 < 0$

network input $z_0 = w^T x \Rightarrow$ in this case the output of the z's will be used for the network input $\Rightarrow 0 = w_4 z_1 + w_5 z_2 + w_6 z_3 =$

$= (0.5) \times 4 + 1 \times 4 + 2 \times 0 = 2 + 4 = 6$

$\Rightarrow$ output $\rightarrow \sigma(w^T x) = \sigma(6)$

$$\sigma(0) = \frac{1}{1 + e^{-0}} \Rightarrow \frac{1}{1 + e^{-6}} = \boxed{0.998}$$

b) $E_{total} = \sum \frac{1}{2} (target - output)^2 =$

$= \frac{1}{2} (0 - 0.998)^2 = \boxed{0.498}$

c) $\frac{\partial E}{\partial w_4} = \frac{\partial E}{\partial o_{out}} \times \frac{\partial_{out} O}{\partial O} \times \frac{\partial O}{\partial w_4}$

$\frac{\partial E}{\partial o_{out}} = -2(y - \hat{y}) = -2(0 - 0.998) = \boxed{1.996}$

$\frac{\partial_{out} O}{\partial O} = \frac{e^{-O}}{(1+e^{-O})^2} = \sigma(O)(1 - \sigma(O)) =$

$= 0.998(1 - 0.998) = \boxed{0.002}$

$\frac{\partial O}{\partial w_4} = z_{1_{out}} = \boxed{4}$

$\Rightarrow \frac{\partial E}{\partial w_4} = 1.996 \times 0.002 \times 4 = \boxed{0.016}$

$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial o_{out}} \times \frac{\partial_{out} O}{\partial O} \times \frac{\partial O}{\partial w_5}$

$\frac{\partial O}{\partial w_5} = z_{2_{out}} = 4$

$\Rightarrow \frac{\partial E}{\partial w_5} = \boxed{0.016}$

$\frac{\partial E}{\partial w_6} = \frac{\partial E}{\partial o_{out}} \times \frac{\partial_{out} O}{\partial O} \times \frac{\partial O}{\partial w_6}$

$\frac{\partial O}{\partial w_6} = z_{3_{out}} = 0$

$\Rightarrow \frac{\partial E}{\partial w_6} = \boxed{0}$

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial out\,z_1} \times \frac{\partial out\,z_1}{\partial z_1} \times \frac{\partial z_1}{\partial w_1}$$

$$\frac{\partial E}{\partial out\,z_1} = \frac{\partial E}{\partial out\,O} \cdot \frac{\partial out\,O}{\partial O} \times \frac{\partial O}{\partial out\,z_1}$$

$= 1.996 \times 0.002 \times \frac{1}{2} = \boxed{0.002}$

$\frac{\partial out\,z_1}{\partial z_1} = \boxed{1}$ $\quad \frac{\partial z_1}{\partial w_1} = \boxed{4}$

$\Rightarrow \frac{\partial E}{\partial w_1} = 4 \times 1 \times 0.002 = \boxed{0.008}$

$$\frac{\partial E}{\partial w_2} = \frac{\partial E}{\partial out\,z_2} \times \frac{\partial out\,z_2}{\partial z_1} \times \frac{\partial z_2}{\partial w_2}$$

$$\frac{\partial E}{\partial out\,z_2} = \frac{\partial E}{\partial out\,O} \times \frac{\partial out\,O}{\partial O} \times \frac{\partial O}{\partial out\,z_2} =$$

$= 1.996 \times 0.002 \times 1 = \boxed{0.004}$ $\quad \frac{\partial out\,z_2}{\partial z_2} = \boxed{1}$

$\Rightarrow \frac{\partial E}{\partial w_2} = 0.004 \times 4 \times 1 = \boxed{0.016}$ $\quad \frac{\partial z_2}{\partial w_2} = \boxed{4}$

$$\frac{\partial E}{\partial w_3} = \frac{\partial E}{\partial out\,z_3} \times \frac{\partial out\,z_3}{\partial z_1}^{\to 0} \cdot \frac{\partial z_3}{\partial w_3}^{\to 4} = \boxed{0}$$

$1.996 \times 0.002 \times 2$

$\Rightarrow \boxed{Ew = [0.008, 0.016, 0.00, 0.016, 0.016, 0.00]}$

HW7 continuation

(1.1) d) $w_1 = w_1 - \eta \times \frac{\partial E_{total}}{\partial w_1} \Rightarrow \eta = 1$

$\Rightarrow$
$w_1 = 1 - 0.008 = 0.992$
$w_2 = 1 - 0.016 = 0.984$
$w_3 = -1 - 0 = -1$
$w_4 = 0.5 - 0.016 = 0.484$
$w_5 = 1 - 0.016 = 0.984$
$w_6 = 2 - 0 = 2$

updated weights

$z_1 = 4 \times 0.992 = 3.968$
$z_2 = 4 \times 0.984 = 3.936$
$z_3 = 4 \times (-1) = -4 \rightarrow z_{3_{out}} = 0$

$\rightarrow 0 = 0.484 \times 3.968 + 0.984 \times 3.936 + 2 \times 0 =$
$= \boxed{5.794}$

$\Rightarrow$ output $= O_{out} = \sigma(5.794) = \frac{1}{1 + e^{-5.794}} =$
$= \boxed{0.997}$

$E_{total} = \frac{1}{2}(0 - 0.997)^2 = \boxed{0.497}$

e) We see that the new weights improved this NN as the loss decreased $\rightarrow E_2 < E_1$.