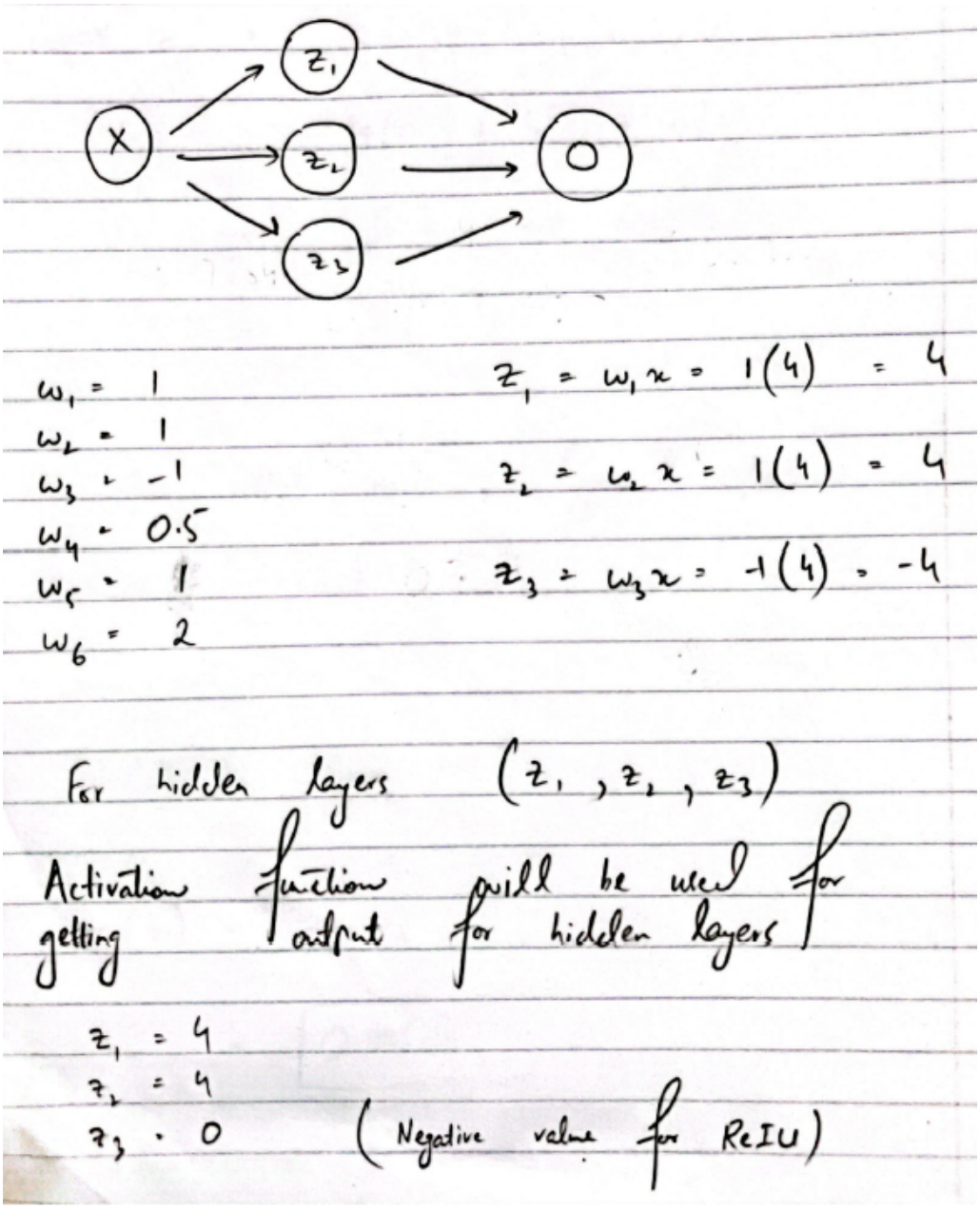
CS 422 - Homework 7

1. Exercises

1.1

(a)



Output $z = z_1 w_4 + z_2 w_5 + z_6 w_6$

$= 4(0.5) + 4(1) + 0(2)$

$= 2 + 4 + 0$

$= \underline{\underline{6}}$

For the output node, Sigmoid activation function is used. Hence

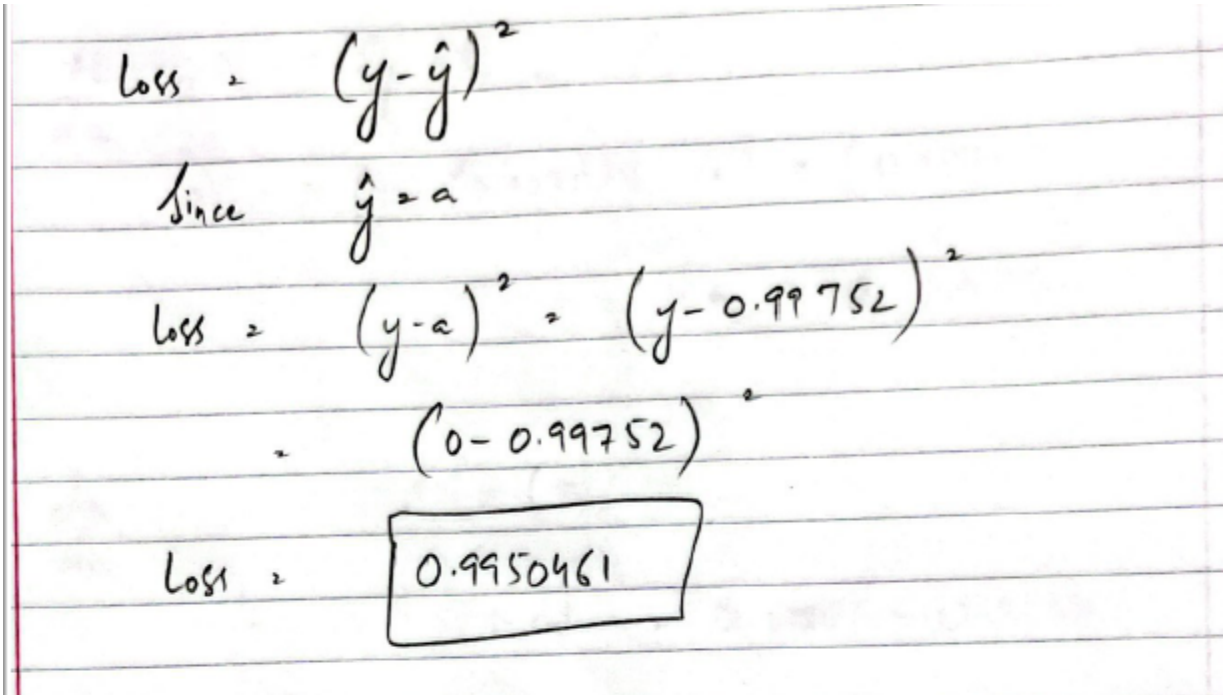
$a = \sigma(z)$

$\sigma(x) = \frac{1}{1+e^{-x}}$

$\sigma(z) = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-6}}$

$= \boxed{0.9950461}$

(b)


Loss = $(y-\hat{y})^2$
Since $\hat{y} = a$
Loss = $(y-a)^2 = (y-0.99752)^2$
$= (0-0.99752)^2$
Loss = 0.9950461

(c)

$$\frac{dL}{da} = -z(y-a)$$

$$= -z(0-(0.9975)) = z(0.9975)$$

$$= @.\ 1.9950$$

$$\frac{da}{dz} = d(\sigma(z))$$

$$= a(1-a) = 0.9975(1-0.9975)$$

$$= 0.0024$$

$$\frac{dL}{dz} = \frac{dL}{dz} \times \frac{dz}{dw_5}$$

$$\frac{dL}{dz} = \frac{dL}{da} \times \frac{da}{dz} = 1.9950 \times 0.0025$$

$$= \cancel{0.0048}\ 0.0050$$

$$\frac{dL}{dw_4} = \frac{dL}{dz} \times \frac{dz}{dw_4} = 0.005_0 (z_1')$$

$$= 0.005_0 \times 4$$

$$= 0.02$$

$$\frac{dL}{dw_5} = \frac{dL}{dz} \times \frac{dz}{dw_5} = 0.005_0 (z_2') \cdot 0.005_0 (4)$$

$$= 0.02$$

$$\frac{dL}{dw_6} = \frac{dL}{dz} \times \frac{dz}{dw_6} = 0.005_0 (z_3')$$

$$= 0$$

$$z = \underbrace{z_1' w_4}_{z_2''} + \underbrace{z_2' w_5}_{z_3''} + \underbrace{z_3' w_6}_{z_4''}$$

$$\frac{dL}{dz_1'} = w_4 \qquad \frac{dL}{dz_2'} = dw_5$$

$$\frac{dL}{dz_3'} = w_6$$

$$\frac{dL}{dz_1} = \frac{dL}{dz} \times \frac{dz}{dz_1} = 0.0050 \times 0.5 = 0.0025$$

$$\frac{dL}{dz_2} = \frac{dL}{dz} \times \frac{dz}{dz_2} = 0.0050\ (1) = 0.005$$

$$\frac{dL}{dz_3} = \frac{dL}{dz} \times \frac{dz}{dz_3} = 0.0050\ (2) = 0.01$$

$z_1' = ReLU(z_1)$ $\quad \frac{dz_1'}{dz} = 1$

$z_2' = ReLU(z_2)$ $\quad \frac{dz_2'}{dz} = 1$

$z_3' = ReLU(z_3)$ $\quad \frac{dz_3'}{dz} = 1$

Similarly $\frac{x}{2z_2} = 0.0050$, $\frac{dL}{dz_3} = 0.01$

$z_1 = w_1 x$

$z_2 = w_2 x$

$z_3 = w_3 x$

$\frac{\partial z_1}{\partial w_1} = x$

$\frac{\partial z_2}{\partial w_2} = x$

$\frac{\partial z_3}{\partial w_3} = x$

$\frac{dL}{dw_1} = \frac{dL}{dz_1} \times \frac{dz_1}{dw_1} = 0.0025\,(4) = 0.005$

$\frac{dL}{dw_2} = \frac{dL}{dz_2} \times \frac{dz_2}{dw_2} = 0.0050\,(4) = 0.002$

$\frac{dL}{dw_3} = \frac{dL}{dz_3} \times \frac{dz_3}{dw_3} = 0.01\,(4) = 0.04$

Now for updation

$w_1 = w_1 - \alpha \frac{dL}{dw_1}$ α represents learning rate

$\alpha = 0.1$

$w_1 = 1 - (0.1)(0.01) = 1 - 0.001 = 0.999$

$w_2 = 1 - (0.1)(0.02) = 0.998$

$w_3 = -1 - (0.1)(0.04) = -1.004$

$w_4 = 0.5 - (0.1)(0.02) = 0.498$

$w_5 = 1 - (0.1)(0.02) = 0.998$

$w_6 = 2 - (0.1)(0) = 2$

Forward Computation

$z_1 = xw_1 = 4(0.999) = 3.996$

$z_2 = xw_2 = 4(0.998) = 3.992$

$z_3 = xw_3 = 4(-1.004) = -4.016$

$z_4 = xw_4 = 1.992$

$z_1' = 0.3996$

$z_2' = \frac{x}{\sigma(z_2)} = 3.992$

$z_3 = 0$

$z = z_1' w_4 + z_2' w_5 + z_3' w_6$

$= 3.996(0.5) + 3.992(1) + 0(2)$

$= \boxed{5.99}$

(d)

$a = \sigma(z) = \frac{1}{1+e^{-z}}$

$= \boxed{0.9975026}$

$L(y, a) = (y-a)^2 = (0 - 0.9975026)^2$

$= \underline{\underline{0.995001}}$

(e)

First output : 0.99752

Output after update : 0.9975026

$Loss_1$: 0.99505

$Loss_2$: 0.99501

Since $Loss_1 > Loss_2$, output after update is closer to target (0).