2) Derivative of Tanh

The hyperbolic tangent (tanh) activation function has the following form:-

$$f(x) = \tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

$$f'(x) = \frac{d}{dx}\left[\frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}\right]$$

Applying quotient rules:-

$$\frac{d}{dx}\left[\frac{u(x)}{v(x)}\right] = \frac{\frac{d}{dx}u(x)\cdot v(x) - u(x)\cdot\frac{d}{dx}v(x)}{v(x)^2}$$

Therefore $f'(x)$ becomes:-

$$= \frac{\frac{d}{dx}[e^{x} - e^{-x}]\cdot[e^{x} + e^{-x}] - [e^{x} - e^{-x}]\cdot\frac{d}{dx}[e^{x} + e^{-x}]}{(e^{x} + e^{-x})^2}$$

$$= \frac{\left(\frac{d}{dx}[e^{x}] - \frac{d}{dx}[e^{-x}]\right)(e^{x} + e^{-x}) - (e^{x} - e^{-x})\left(\frac{d}{dx}[e^{x}] + \frac{d}{dx}[e^{-x}]\right)}{(e^{x} + e^{-x})^2}$$

Exponential rule.

$$\frac{d(e^{u(x)})}{dx} = e^{u(x)}\cdot\frac{d(u(x))}{dx}$$

$$= \frac{\left(e^{x} - e^{-x}\cdot\frac{d}{dx}(-x)\right)(e^{x} + e^{-x}) - (e^{x} - e^{-x})\left(e^{x} + e^{-x}\cdot\frac{d}{dx}(-x)\right)}{(e^{x} + e^{-x})^2}$$

$$= \frac{(e^{x} - e^{-x}\cdot 1)(e^{x} + e^{-x}) - (e^{x} - e^{-x})(e^{x} - e^{-x}\cdot 1)}{(e^{x} + e^{-x})^2}$$

$$= \frac{(e^{x} + e^{-x})^2}{(e^{x} + e^{-x})^2} - \frac{(e^{-x} - e^{-x})^2}{(e^{x} + e^{-x})^2} \Rightarrow 1 - f(x)^2$$

Hence, $f'(x) = 1 - f(x)^2$ proved#

1) Forward backward pass.

Task:-

1) $h_1 = i_1 \times w_1 + i_2 \times w_3 + i_3 \times w_5 = 0.765$

2) $h_2 = i_1 \times w_2 + i_2 \times w_4 + i_3 \times w_6 = 0.52$

$$\hat{y}_1 = h_1 \times w_7 + h_2 \times w_8 + h_3 \times w_9$$
$$= 0.765 \times 0.7 + 0.52 \times 0.25 + 1 \times (-0.1)$$
$$= 0.5655$$

2) $$MSE_1 = \frac{1}{2}(y - \hat{y}_1)^2$$
$$= \frac{1}{2}(0.5 - 0.5655)^2$$
$$= 0.002145$$

3) Gradient using back propagation.

$$w_9 = w_9 - \alpha\left(\frac{\partial Error}{\partial w_9}\right)$$

$$\frac{\partial Error}{\partial w_9} = \frac{\partial Err}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial w_9}$$
$$= \frac{1}{2} \times \frac{d(y-\hat{y})^2}{d\hat{y}} \cdot \frac{h_1 w_7 + h_2 w_8 + h_3 w_9}{d w_9}$$
$$= \frac{1}{2} \times 2(\hat{y} - y) \cdot h_3$$
$$= -h_3(y - \hat{y})$$

Therefore, $w_9 = w_9 - \alpha(-h_3(y-\hat{y}))$

Similarly,

$$w_8 = w_8 - \alpha(-h_2(y-\hat{y}))$$
$$w_7 = w_7 - \alpha(-h_1(y-\hat{y}))$$

Also,

$$w_6 = w_6 - \alpha\left(\frac{\partial Error}{\partial w_6}\right)$$

$$\frac{\partial Error}{\partial w_6} = \frac{\partial Err}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial h_2} \cdot \frac{\partial h_2}{\partial w_6}$$

$$= \frac{1}{2}\frac{d}{d\hat{y}}(y-\hat{y})^2 \cdot \frac{\partial(h_1w_7+h_2w_8+h_3w_9)}{\partial h_2} \cdot \frac{\partial(i_1w_2+i_2w_4+i_3w_6)}{\partial w_6}$$

$$= -(y-\hat{y}) \cdot w_8 \cdot i_3$$
$$= -i_3(y-\hat{y})w_8$$

$$\therefore w_6 = w_6 - \alpha(-i_3(y-\hat{y})w_8)$$

Similarly,

$$w_5 = w_5 - \alpha(-i_3(y-\hat{y})w_7)$$
$$w_4 = w_4 - \alpha(-i_2(y-\hat{y})w_8)$$
$$w_3 = w_3 - \alpha(-i_2(y-\hat{y})w_7)$$
$$w_2 = w_2 - \alpha(-i_1(y-\hat{y})w_8)$$
$$w_1 = w_1 - \alpha(-i_1(y-\hat{y})w_7)$$

4) update weights

$$w_9 = w_9 - \alpha(-h_3(y-\hat{y})) = -0.1019$$
$$w_8 = w_8 - \alpha(-h_2(y-\hat{y})) = 0.2989$$
$$w_7 = w_7 - \alpha(-h_1(y-\hat{y})) = 0.6984$$
$$w_6 = w_6 - \alpha(-i_3(y-\hat{y})w_8) = -0.1404$$
$$w_5 = w_5 - \alpha(-i_3(y-\hat{y})w_7) = 0.8986$$
$$w_4 = w_4 - \alpha(-i_2(y-\hat{y})w_8) = 0.19975$$
$$w_3 = w_3 - \alpha(-i_2(y-\hat{y})w_7) = 0.4493$$
$$w_2 = w_2 - \alpha(-i_1(y-\hat{y})w_8) = 0.7996$$
$$w_1 = w_1 - \alpha(-i_1(y-\hat{y})w_7) = -0.3009$$

6) Perform forward pass:

$h_1 = i_1 w_1 + i_2 w_3 + i_3 w_5 = 0.7626$

$h_2 = i_1 w_2 + i_2 w_4 + i_3 w_6 = 0.5191$

$\hat{y}_2 = h_1 \times w_7 + h_2 \times w_8 + h_3 \times w_9$

$= 0.7626 \times 0.698 + 0.5191 \times 0.249 + 1 \times (-0.1019)$

$= 0.5599$

7) MSE_2 calculation:

$MSE_2 = \frac{1}{2}(y - \hat{y}_2)^2$

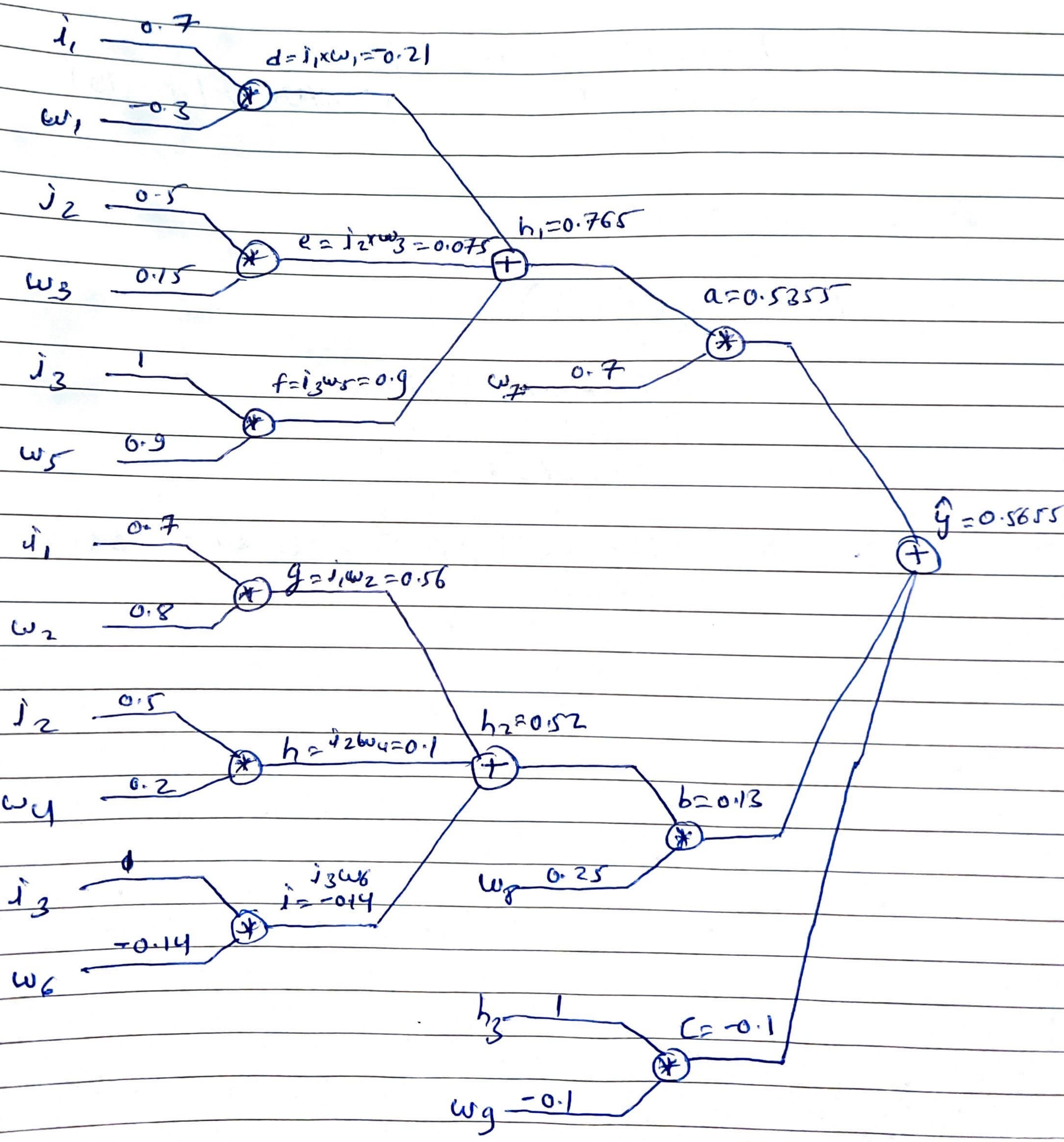
$= \frac{1}{2}(0.5 - 0.5599)^2$

$= 0.001794$

Therefore, $MSE_2 < MSE_1$. Thus error has been reduced.

5) Computation graph for forward pass and Backward.

Forward pass :-



Backward pass:

$$f = a+b+c$$

where,

$$a = h_1 \times w_7 \quad / \quad b = h_2 \times w_8 \quad / \quad c = h_3 \times w_9$$

$$\frac{\partial f}{\partial a} = \frac{\partial (a+b+c)}{\partial a} = 1$$

$$\frac{\partial f}{\partial b} = \frac{\partial (a+b+c)}{\partial b} = 1$$

$$\frac{\partial f}{\partial c} = \frac{\partial (a+b+c)}{\partial c} = 1$$

$$\frac{\partial f}{\partial w_7} = \frac{\partial f}{\partial a} \cdot \frac{\partial a}{\partial w_7} = 1 \cdot \frac{\partial h_1 w_7}{\partial w_7} = h_1 = 0.765$$

$$\frac{\partial f}{\partial h_1} = \frac{\partial f}{\partial a} \cdot \frac{\partial a}{\partial h_1} = 1 \cdot \frac{\partial h_1 w_7}{\partial h_1} = w_7 = 0.7$$

similarly,

$$\frac{\partial f}{\partial w_8} = h_2 = 0.52 \quad / \quad \frac{\partial f}{\partial h_2} = w_8 = 0.25$$

$$\frac{\partial f}{\partial w_9} = h_3 = 1 \quad / \quad \frac{\partial f}{\partial h_3} = w_9 = -0.1$$

$$h_1 = d+e+f \quad / \quad h_2 = g+h+i$$

$$\frac{\partial f}{\partial d} = \frac{\partial f}{\partial a} \cdot \frac{\partial a}{\partial h_1} \cdot \frac{\partial h_1}{\partial d}$$

$$= 1 \cdot \frac{\partial h_1 \cdot w_7}{\partial h_1} \cdot \frac{\partial (d+e+f)}{\partial d}$$

$$= w_7 \cdot 1 = w_7 = 0.765$$

Similarly,

$$\frac{\partial f}{\partial e} = w_7 = 0.765 \quad / \quad \frac{\partial f}{\partial f} = 1 \cdot w_7 \cdot 1 = w_7 = 0.765$$

$$\frac{\partial f}{\partial g} = \frac{\partial f}{\partial b} \cdot \frac{\partial b}{\partial h_2} \cdot \frac{\partial h_2}{\partial g}$$

$$= 1 \cdot w_8 \cdot 1 = w_8 = 0.52$$

$$\frac{\partial f}{\partial h} = w_8 = 0.52 \quad / \quad \frac{\partial f}{\partial i} = 0.52$$

Also,

$$\frac{\partial f}{\partial i_1} = \frac{\partial f}{\partial a} \cdot \frac{\partial a}{\partial h_1} \cdot \frac{\partial h_1}{\partial d} \cdot \frac{\partial d}{\partial i_1}$$

$$= 1 \cdot w_7 \cdot 1 \cdot w_1$$

$$= w_1 w_7$$

$$= -0.3 \times 0.765 = -0.2295$$

$$\frac{\partial f}{\partial w_1} = \frac{\partial f}{\partial a} \cdot \frac{\partial a}{\partial h_1} \cdot \frac{\partial h_1}{\partial d} \cdot \frac{\partial d}{\partial w_1}$$

$$= 1 \cdot w_7 \cdot 1 \cdot i_1$$

$$= w_7 \cdot i_1 = 0.765 \times 0.7 = 0.5355$$

For others its calculated similarly as above.