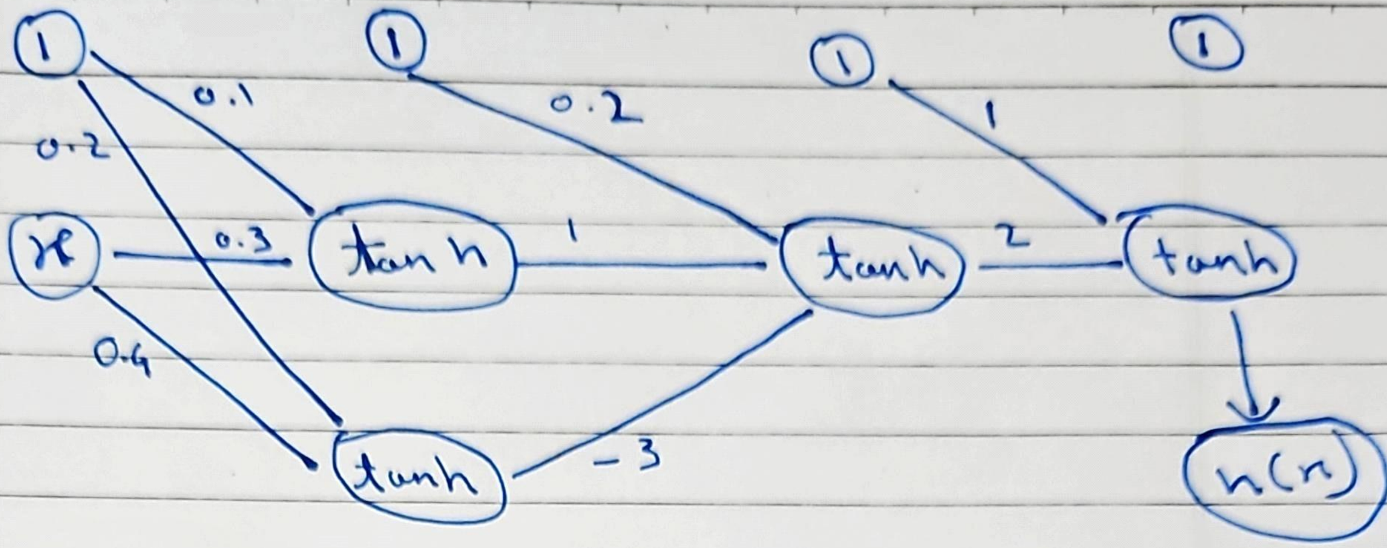


$x = 2$ $\qquad$ $y = 1$

$$x_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad w_1 = \begin{bmatrix} 0.1 \\ 0.3 \end{bmatrix} \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix}$$

$$s^1 = (w^1)^T x^0$$

$$\underline{\underline{s^1}} = \begin{bmatrix} 0.1 & 0.3 \\ 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 1 \end{bmatrix}$$

$$\underline{\underline{x^1}} = \begin{bmatrix} 1 \\ \tanh(0.7) \\ \tanh(1) \end{bmatrix} = \begin{bmatrix} 1 \\ 0.60 \\ 0.76 \end{bmatrix}$$

$$w^2 x^1 = \begin{bmatrix} 0.2 & 1 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 0.60 \\ 0.76 \end{bmatrix}$$

$$= 0.2 + 0.60 \cdot -3 \times 76$$

$$\underline{\underline{s^2}} = -1.48$$

PAGE No.
DATE / /

$$\underline{x}^2 = \begin{bmatrix} 1 \\ \tanh(-1.48) \end{bmatrix} = \begin{bmatrix} 1 \\ -0.90 \end{bmatrix}$$

$$\underline{s}^3 = \begin{bmatrix} \cdot 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -0.90 \end{bmatrix} = -0.8$$

$$x^3 = \tanh(-0.8) = -0.664 = h(x)$$

$$\boxed{h(x) \neq y \quad | \quad -0.664 \neq 1}$$

Back Propagation

$$\delta^3 = 2(x^3 - y) \cdot \theta'(s^3)$$

Derivative

$$\delta^3 = 2(-0.664 - 1) \cdot (\tanh(-0.8))$$

$$= -3.328 \cdot (1 - \tanh^2(-0.8))$$

$$\boxed{\delta^3 = -3.328 \cdot (1 - 0.440) = -1.86}$$

$$\delta^2 = (1 - \tanh^2(s^2))(w^3 \delta^3)$$

$$= (1 - \tanh^2(-1.48))(2 \cdot (-1.86))$$

$$= -0.1887 \times 3.72$$

$$\boxed{\delta^2 = -0.699}$$

PAGE No.

DATE / /

$$\delta^1 = \begin{bmatrix} 1 - \tanh^2(s^1_1) \\ 1 - \tanh^2(s^1_2) \end{bmatrix} \otimes \begin{bmatrix} w^2 \ \delta^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 - \tanh^2(0.7) \\ 1 - \tanh^2(1) \end{bmatrix} \otimes \begin{bmatrix} 1 \\ -3 \end{bmatrix} \times 0.649$$

$$= \begin{bmatrix} 0.634 \\ 0.419 \end{bmatrix} \otimes \begin{bmatrix} -0.649 \\ 2.097 \end{bmatrix}$$

$$\boxed{\delta^1 = \begin{bmatrix} -0.443 \\ 0.878 \end{bmatrix}}$$

Gradient

$$G^1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \times \begin{bmatrix} -0.443 \\ 0.878 \end{bmatrix}^T$$

$$G^1 = \begin{bmatrix} 0.443 & 0.878 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\boxed{G^1 = 2.199}$$

$G^2 =$

Gradient

$$G^1 = x^0 \cdot (\delta^1)^T$$

$$= \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} -0.443 & 0.878 \end{bmatrix}$$

$$= \begin{bmatrix} -0.443 & 0.878 \\ -0.886 & 1.756 \end{bmatrix}$$

$$G^2 = x^1 (\delta^2)^T$$

$$= \begin{bmatrix} 1 \\ 0.6 \\ 0.76 \end{bmatrix} \begin{bmatrix} -0.69 \end{bmatrix}$$

$$= \begin{bmatrix} -0.69 \\ -0.414 \\ -0.524 \end{bmatrix}$$

$$G^3 = x^2 (\delta^3)^T$$

$$= \begin{bmatrix} 1 \\ -0.9 \end{bmatrix} \begin{bmatrix} -1.86 \end{bmatrix}$$

$$= \begin{bmatrix} -1.86 \\ 1.67 \end{bmatrix}$$

$\eta = 1$

PAGE No.

DATE / /

1) $$W^1 = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.4 \end{bmatrix} - \begin{bmatrix} -0.443 & 0.878 \\ -0.886 & 1.756 \end{bmatrix}$$

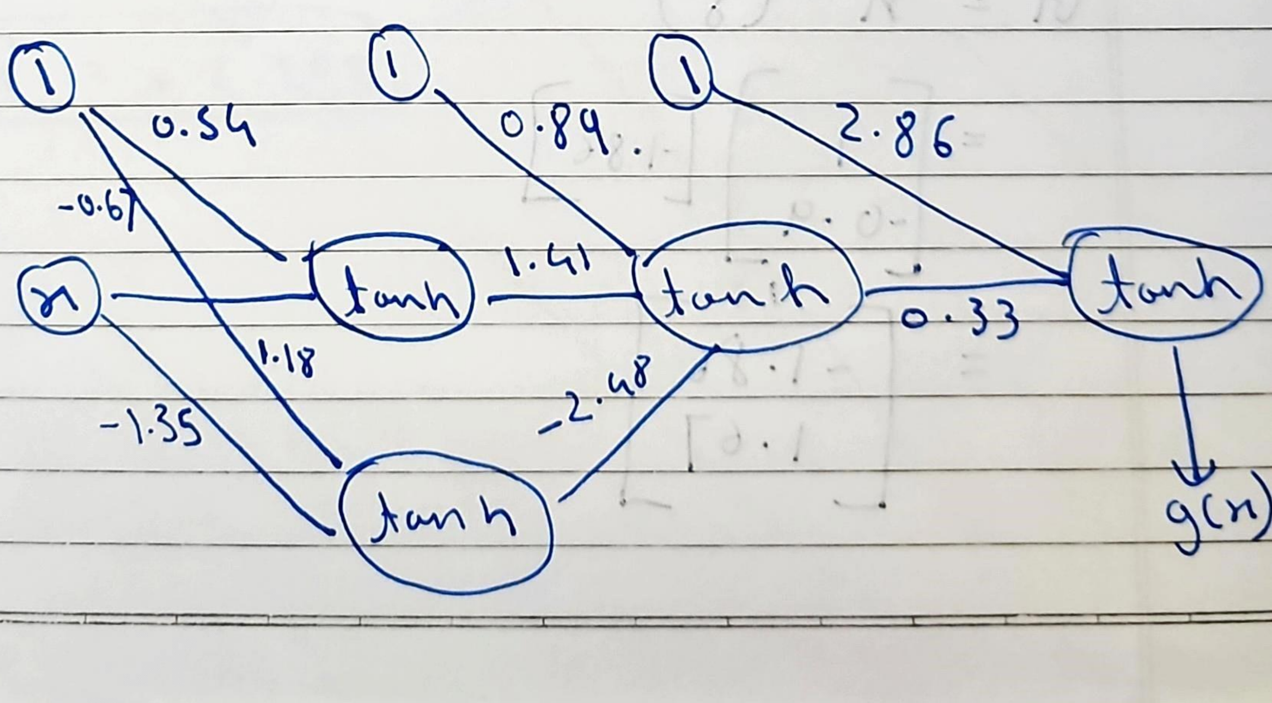
$$= \begin{bmatrix} 0.1 + 0.443 & 0.2 - 0.878 \\ 0.3 + 0.886 & 0.4 - 1.756 \end{bmatrix}$$

$$= \begin{bmatrix} 0.543 & -0.678 \\ 1.186 & -1.356 \end{bmatrix}$$

2) $$W^2 = \begin{bmatrix} 0.2 \\ 1 \\ -3 \end{bmatrix} - \begin{bmatrix} -0.69 \\ -0.41 \\ -0.52 \end{bmatrix} = \begin{bmatrix} 0.89 \\ 1.41 \\ -2.48 \end{bmatrix}$$

3) $$W^3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} -1.86 \\ 1.67 \end{bmatrix} = \begin{bmatrix} 2.86 \\ 0.33 \end{bmatrix}$$

New Network will be



Weights and gradient:

A positive gradient for a weight implies a decrease is needed, and a negative gradient suggests an increase.

1. First Layer:
    a. For the first neuron, the weights must be increased to reduce the error
    b. For the second neuron, the weights must be decreased to reduce the error
2. Second Layer:
    a. All weights must be increased to reduce the error
3. Third Layer:
    a. For the first unit, the weight must be increased to reduce the error
    b. For the second unit, the weight must be decreased to reduce the error