Discussion 13/July/2021

(i) Understand how math behind neural networks works
(ii) Do an example
(iii) Try predicting for $(X_1 \; ... \; X_{30})$ together
(iv) Explain how loss is taken into picture

Input

| $R_1$ | ... | $R_{30}$ | $W_1$ | ... | $W_{30}$ | C | $O_1$ | ... | $O_{30}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | 0 | | 0 |
| | | | | | | | 1 | | 1 |
| | | | | | | | 0 | | 0 |
| | | | | | | | ⋮ | | ⋮ |

learning rate $\eta = 0.001$

$$\phi_1(v) = f(v) = \begin{cases} 0 & \text{for } v < 0 \\ v & \text{for } v >= 0 \end{cases} = \phi_2(v) = \phi_3(v) = \phi_4(v)$$

$$\phi_5(v) = \frac{1}{1+e^{-v}}$$

Page No.:
Date:

Model considerations

61 : 128 : 64 : 32 : 2 : 1

For mathematical purpose, we consider smaller model.

Let us consider only 2 D2D pairs

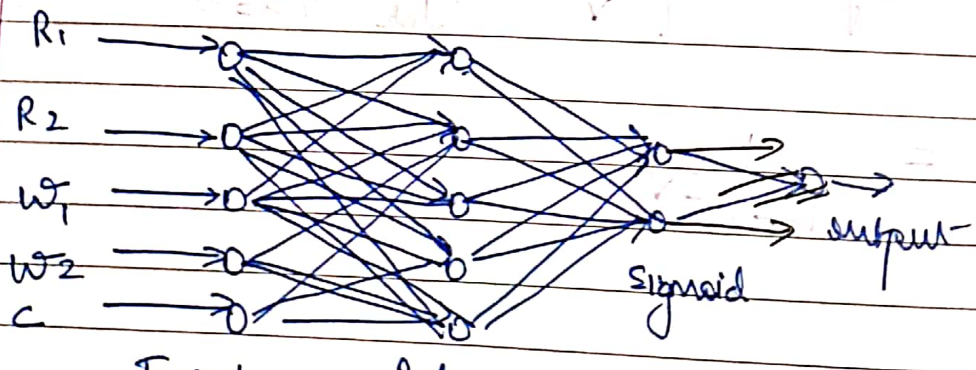
Note: Not true data

| $R_1$ | $R_2$ | $W_1$ | $W_2$ | C | $O_1$ | $O_2$ |
|---|---|---|---|---|---|---|
| 1 | 1.5 | 1 | 0.5 | 1 | 1 | 0 |
| 1.5 | 1 | 1 | 0.5 | 1 | 0 | 1 |
| 1 | 1.5 | 1.5 | 0.5 | 0 | 1 | 0 |
| 1.5 | 0.5 | 0.5 | 1.5 | 1 | 0 | 1 |

~~First~~ Let the model be

5 : 5 : 2 : 1

input → → output




$\phi_1(v) \Rightarrow v \quad \text{if} \quad v >= 0$
$\qquad\quad 0 \quad \text{else}$

$$\phi_2(v) = \frac{1}{1+e^{-v}}$$

let weights 1 be

$$W_1 = \begin{bmatrix} 0.2 & 1 & 0.2 & 1 & 0.2 & 1 \\ -0.2 & 1 & -0.2 & 1 & -0.2 & 1 \\ 0.2 & -1 & 0.2 & -1 & 0.2 & -1 \\ -0.2 & 1 & -0.2 & 1 & -0.2 & 1 \\ 0.2 & 1 & 0.2 & 1 & 0.2 & 1 \end{bmatrix}$$

5

$$\text{Input } X_1 = \begin{bmatrix} 1 \\ 1 \\ 1.5 \\ 1 \\ 0.5 \\ 1 \end{bmatrix} \quad 1 \rightarrow \text{bias } b_0$$

Forward Pass

$$V_1 = \begin{pmatrix} V_{11} \\ V_{12} \\ V_{13} \\ V_{14} \\ V_{15} \end{pmatrix} = W_1 X_1^T = \begin{bmatrix} 0.2 & 1 & 0.2 & 1 & 0.2 & 1 \\ -0.2 & 1 & -0.2 & 1 & -0.2 & 1 \\ 0.2 & -1 & 0.2 & -1 & 0.2 & -1 \\ -0.2 & 1 & -0.2 & 1 & -0.2 & 1 \\ 0.2 & 1 & 0.2 & 1 & 0.2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1.5 \\ 1 \\ 0.5 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 3.6 \\ 2.4 \\ -2.4 \\ 2.4 \\ 3.6 \end{bmatrix}$$

$$\overline{y}_1 = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{15} \end{bmatrix} = \begin{bmatrix} \text{Relu}(3.6) \\ \text{Relu}(2.4) \\ \text{''}\ (-2.4) \\ \text{''}\ (2.4) \\ \text{''}\ (3.6) \end{bmatrix} = \begin{bmatrix} 3.6 \\ 2.4 \\ 0 \\ 2.4 \\ 3.6 \end{bmatrix}$$

Scanned with CamScanner

Let weights 2 be as

$$W_2 = \begin{bmatrix} -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

Now the function is sigmoid

$$V_2 = W_2 \cdot y_1^T$$

$$= \begin{bmatrix} -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3.6 \\ 2.4 \\ 0 \\ 2.4 \\ 3.6 \end{bmatrix}$$

$$= \begin{bmatrix} 1.4 \\ -1.4 \end{bmatrix}$$

$$\bar{y_2} = \text{sigmoid}(V_2)$$

$$y_{21} = \frac{1}{1+e^{-1.4}} \qquad y_{22} = \frac{1}{1+e^{1.4}}$$

$$= 0.8022 \qquad = 0.1978$$

for output 1

Calculating Error $\Rightarrow$ squared error function

$$E = \Sigma \frac{1}{2} (\text{target} - \text{output})^2$$

$$E_{01} = \frac{1}{2}(1 - 0.8022)^2 = 0.0195$$

$$E_{02} = \frac{1}{2}(0 - 0.1978)^2 = 0.0195$$

$$E_{total} = 0.0195 + 0.0195$$
$$= 0.03912$$

Backward Pass

$\odot \rightarrow$ Hadamard Transform

$$\Delta W_2 = \eta_2 \cdot \delta_2 y_1^T \qquad \text{let } \eta_2 = 0.01$$

$$\delta_2 = E_{01} \odot \phi_2'(v_2(k))$$

$$= E_{01} \odot (\text{partial derivative of sigmoid})$$

$$= E_{01} \odot \frac{\partial(\text{sigmoid})}{\partial v_2}$$

$$= E_{01} \odot (1.4)(1 - 1.4)$$

$$= 0.0195 \times (-0.56)$$

$$= -0.01092$$

$$E_{02} \odot (-1.4)(1 + 1.4)$$

$$= \cancel{0.65}$$

$$-0.06552$$

$$\Delta W_{2_{01}} = 0.01 \times (-0.01092) \times [1 \quad 3.6 \quad 2.4 \quad 0 \quad 2.4 \quad 3.6]$$

$$= [-0.0001092 \quad -0.00039312 \quad -0.00026208 \quad 0 \quad -0.00026208 \quad -0.00039312]$$

similarly

$$\Delta W_{2_{02}} = 0.01 \times (-0.06552) \times [1 \quad 3.6 \quad 2.4 \quad 0 \quad 2.4 \quad 3.6]$$

$$= [-0.0006552 \quad -0.245 \quad -0.16 \quad 0 \quad -0.16 \quad -0.245]$$

$$W_2(K+1) = W_2(k) + \Delta W_2$$

$$W_2(2) = W_2(1) + \Delta W_2$$

$$= [\cancel{0.9998 \quad 3.5996 \quad 2.3997 \quad 0 \quad 2.3997} \quad 3.5996]$$

$$[$$

$$= \begin{bmatrix} -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix} + \begin{bmatrix} -0.0001902 & \cdots & -0.\ldots \\ \cdots & & \end{bmatrix}$$

$$W_2(2) = \begin{bmatrix} -1.0001902 & 0.9996 & -1.00026208 & 1 & -1.00026 & 0.999\ldots \\ 0.9993 & -1.245 & 0.84 & -1 & 0.84 & -1.24\ldots \end{bmatrix}$$

Page No.:
Date:

$$\delta = W_2^T \delta_2 \odot \phi'(V_1)$$

Note $\Rightarrow$ $\delta_{ij} = W_{i+1}^T \delta_{j+1} \odot \delta_j'(V_j(k))$

$$= \begin{bmatrix} -1.001902 & 0.999 \\ 0.9996 & -1.245 \\ -1.000262 & 0.84 \\ 1 & -1 \\ -1.0002 & 0.84 \\ 0.9996 & -1.245 \end{bmatrix} \begin{bmatrix} -0.1092 \\ -0.6522 \end{bmatrix} \odot \begin{bmatrix} \cancel{3.6} \\ \cancel{2.4} \\ \cancel{2.4} \\ \cancel{2.4} \\ 3.6 \end{bmatrix}$$

$[3.6 \quad 2.4 \quad 2.4 \quad 2.4 \quad 3.6]$

$$= \begin{bmatrix} 0.\cancel{109}5345 \\ -0.808 \\ -0.4386 \\ 0.53 \\ -0.4386 \\ -0.808 \end{bmatrix}_{5\times1} \underset{1\times5}{[1 \quad 3.6 \quad 2.4 \quad 2.4 \quad 2.4 \quad 3.6]}$$

$$\cancel{DW_1}\delta_1 = \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}_{5\times5}$$

$$\therefore \Delta W_1 = 0.01 \times \underset{5\times5}{\delta_1} \times \underset{1\times5}{[\quad]}$$

$$= \begin{bmatrix} \\ \\ \\ \end{bmatrix}_{5\times5}$$

$W_{K_2} = W_{K_1} + \Delta W_1 \Rightarrow$ New weights for hidden Relu layer.

Scanned with CamScanner