ASSIGNMEN-2

Theoretical Part:

(1) Given,

$$O = w_0 + w_1(x_1 + x_1^2) + \ldots\ldots + w_n(x_n + x_n^2)$$

$x_1, x_2, \ldots x_n$ - inputs

$w_1, w_2, \ldots w_n$ - weights, $w_0 \rightarrow$ bias weight

Activation function: $f(x) = x$

$$E_d = \frac{1}{2}\sum_d (t_d - O_d)^2$$

$$\frac{\partial E}{\partial w_i} = \frac{\partial}{\partial w}\frac{1}{2}\sum_d (t_d - O_d)$$

$$= \frac{1}{2}\sum_d \frac{\partial}{\partial w_i}(t_d - O_d)^2$$

$$= \frac{1}{2}\sum_d 2(t_d - O_d)\frac{\partial}{\partial w_i}(t_d - O_d)$$

$$= \sum_d (t_d - O_d)\frac{\partial}{\partial w_i}(t_d - (w_0 + w_i(x_1 + x_i^2) + \ldots + w_n(x_n + x_n^2)))$$

$$= \sum_d (t_d - O_d)(0 - \frac{\partial}{\partial w_i}(w_i(x_i + x_i^2)))$$

$$\frac{\partial E}{\partial w_i} = -\sum_d (t_d - O_d)(x_i + x_i^2)$$

$$\Delta w_i = -\eta\frac{\partial E}{\partial w_i}$$

$$\Delta w_i = \eta\sum_{d\in D}(t_d - O_d)(x_i + x_i^2)$$

η : Learning rate

D : set of all training samples

t_d : target o/p for d^{th} training sample

O_d : actual o/p for d^{th} training sample

x_{id} : value of i^{th} attribute for the d^{th} training sample

$$\boxed{w_i^{new} = w_i + \Delta w_i}$$

(2) i/p layer $\rightarrow f(x) = x$ activation

hidden and o/p layer $\rightarrow h(x)$ activation

(a) y_5 in terms of weight, i/ps & $h(x)$

$$net_1 = f(x_1) = x_1$$

$$net_2 = f(x_2) = x_2$$

$$net_3 = h[w_{31} net_1 + w_{32} net_2] = h_3(w_{31}x_1 + w_{32}x_2)$$

$$net_4 = h[w_{41} net_1 + w_{42} net_2] = h(w_{41}x_1 + w_{42}x_2)$$

$$y_5 = h(w_{53} net_3 + w_{54} net_4)$$

$$\Downarrow$$

$$y_5 = h(w_{53}(h(w_{31}x_1 + w_{32}x_2)) + w_{54} h(w_{41}x_1 + w_{42}x_2))$$

(b) $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ $\quad w^{(1)} = \begin{pmatrix} w_{3,1} & w_{3,2} \\ w_{4,1} & w_{4,2} \end{pmatrix}$

$$w^{(2)} = (w_{5,3} \quad w_{5,4})$$

$net_1 = x_1$, $net_2 = x_2$

$$H_L = \begin{pmatrix} net_3 \\ net_4 \end{pmatrix} = h\left(\begin{pmatrix} w_{3,1} & w_{3,2} \\ w_{4,1} & w_{4,2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = h(w^{(1)} \cdot X)$$

$$y_5 = h((w_{5,3} \quad w_{5,4}) \cdot H_L)$$

$$y_5 = h(w^{(2)} \cdot h(w^{(1)} \cdot X))$$

(c) $h_s(x) = \dfrac{1}{1+e^{-x}}$

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

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

$$= \frac{2}{1+e^{-2x}} - \left(\frac{\cancel{1+e^{-2x}}}{\cancel{1+e^{-2x}}}\right)$$

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

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

∴ In both, the parameters differ only by linear transformations and constants.

→ The two activation functions can generate the same o/p functions.