CS19BTECH11020

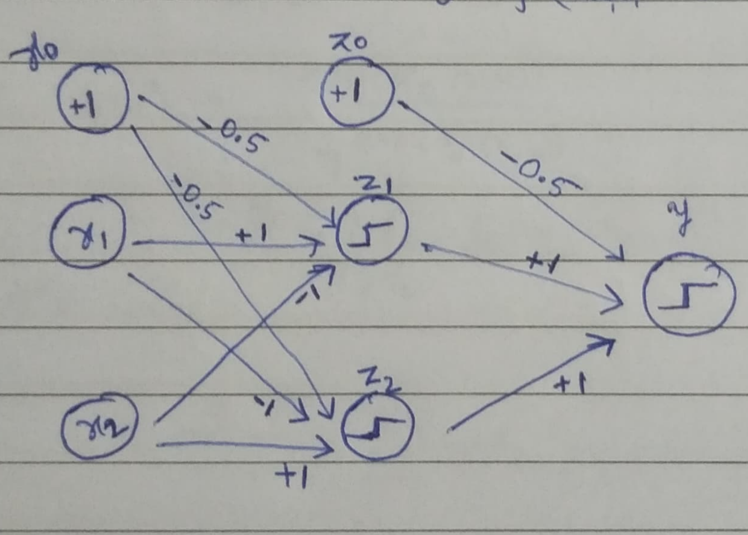
Assignment 3

1) (a) XOR

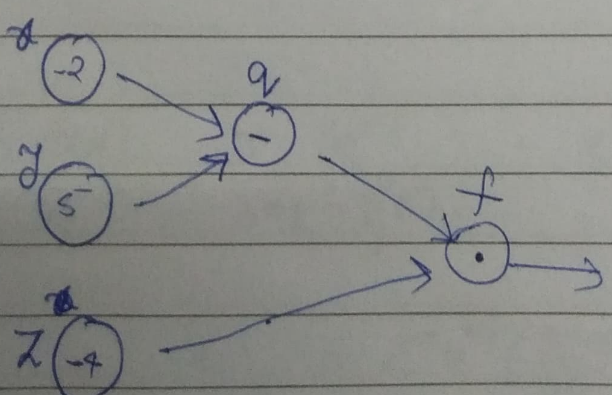
x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

2 layer perceptron $\quad z_i = f\left(\sum_{i=1}^{2} w_i x_i + 0.5\, x_0\right)$, f = hardlimiting f^n

$y = f\left(\sum_{i=1}^{2} w_i z_i + 0.5\, z_0\right)$, $f(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$



(b) $x = -2$, $y = 5$, $z = -4$, $\quad q = x - y$, $f = q * z$



$$\frac{\partial f}{\partial q} = z, \quad \frac{\partial f}{\partial z} = q$$

$$\frac{\partial f}{\partial x} = \frac{\partial q}{\partial x} * \frac{\partial f}{\partial q} = \frac{\partial q}{\partial x} * z = z = -4$$

$$\frac{\partial f}{\partial y} = \frac{\partial q}{\partial y} * \frac{\partial f}{\partial q} = \frac{\partial q}{\partial y} * z = -z = 4$$

$$\frac{\partial f}{\partial z} = q = x - y = -7$$

Hence shown.

(2) $E(w) = -\sum_{n=1}^{N}\sum_{k=1}^{K} t_{kn} \ln y_k(x_n, w)$

$$y_k(x, w) = \frac{\exp(\alpha_k(x, w))}{\sum_j \exp(\alpha_j(x, w))}$$

$$\Rightarrow \frac{\partial E}{\partial \alpha_k} = \sum_{n=1}^{N}\sum_{k=1}^{K} \frac{\partial}{\partial \alpha_k} t_{kn} \ln(y_k(x_n, w)))$$

$$= -\sum_{n=1}^{N}\sum_{k=1}^{K} \frac{\partial}{\partial \alpha_k} t_{nk} \ln\left(\frac{e^{(\alpha_k(x,w))}}{\sum_j e^{(\alpha_j(x,w))}}\right)$$

$$= -\sum_{n=1}^{N}\sum_{k=1}^{K}\left[\frac{\partial}{\partial \alpha_k}(t_{nk}\alpha_k) - \frac{\partial}{\partial \alpha_k}\left(t_{nk} \ln \sum_j e^{\alpha_j}\right)\right.$$

$$\frac{\partial y_k(x, w)}{\partial \alpha_k} = \frac{\partial}{\partial \alpha_k}\left(\frac{e^{\alpha_k}}{\sum_j e^{\alpha_j}}\right) = \frac{e^{\alpha_k}\sum e^{\alpha_j} - (e^{\alpha_k})^2}{(\sum_j e^{\alpha_j})^2} = y_k - y_k^2$$

$$\frac{\partial y_i}{\partial \alpha_k} = \frac{\partial}{\partial \alpha_k}\frac{e^{\alpha_k}}{\sum_j e^{\alpha_j}} = \frac{\sum_j e^{\alpha_j}\left(\frac{\partial}{\partial \alpha_k} e^{\alpha_k}\right) - e^{\alpha_k}\frac{\partial}{\partial \alpha}\left(\sum_j e^{\alpha_j}\right)}{(\sum_j e^{\alpha_j})^2}$$

_/_/_

$$= \frac{0 - e^{\alpha_j} e^{\alpha_k}}{(\sum_j e^{\alpha_i})^2} = - y_j y_k$$

$$\frac{\partial E}{\partial \alpha_k} = -\frac{t_k}{y_k} y_k (1-y_k) - \sum_{j \neq k} \frac{t_j}{y_j} (-y_j y_k)$$

$$= -t_k + y_k \sum_{j=1}^{k} t_j$$

$$= -t_k + y_k$$

$$= y_k - t_k$$

Hence proved.

(3) Given, a convex f^n $f(x) = x^2$

$$E_{AV} = \frac{1}{M} \sum_{m=1}^{M} E_x [(y_m(x) - f(x))^2]$$

$$E_{ENS} = E_x \left[\left(\frac{1}{M} \sum_{m=1}^{M} y_m(x) - f(x) \right)^2 \right]$$

WKT Jensen's equality for a function h is

$$h\left(\frac{\sum a_i x_i}{\sum a_i}\right) \leq \left(\frac{\sum a_i h(x_i)}{\sum a_i}\right)$$

Let us assume all the weights to be $= 1$

Hence sum of weights $= M$ ie $\sum a_i = M$

Let $x_i = [y_m(x) - f(x)^2]$

$h = E_x$

So substituting in Jensen's equality

$$E_x\left(\frac{\sum_{m=1}^{M} y_m(x) - f(x)^2}{M}\right) \leq \frac{\sum_{m=1}^{M} E_x\left[y_m(x) - f(x)^2\right]}{M}$$

LHS = E_{ENS} , RHS = E_{AV}

$$\Rightarrow E_{ENS} \leq E_{AV}$$

Hence proved.