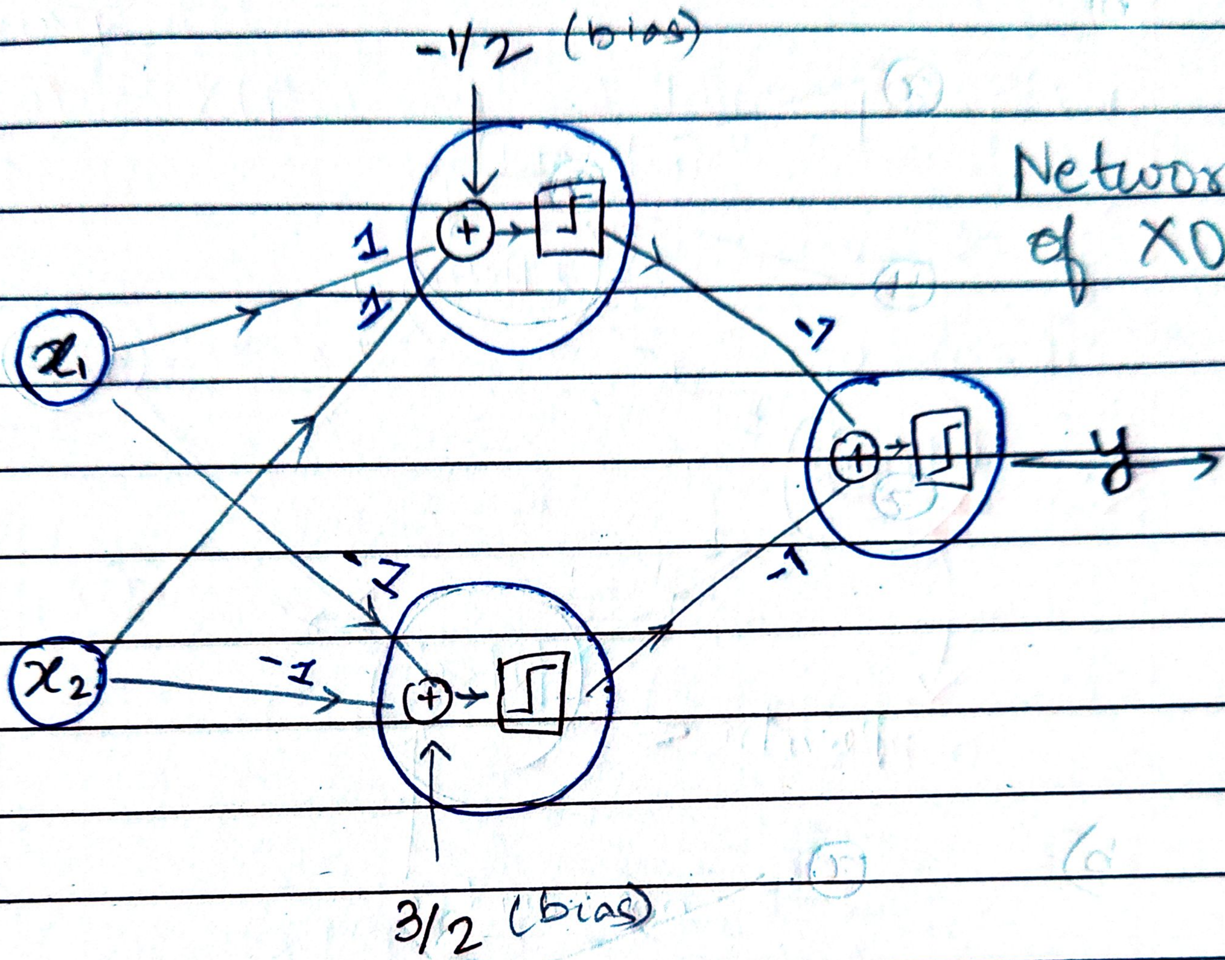
AI20BTECH11027

Date

Page

1) a)



Network of XOR

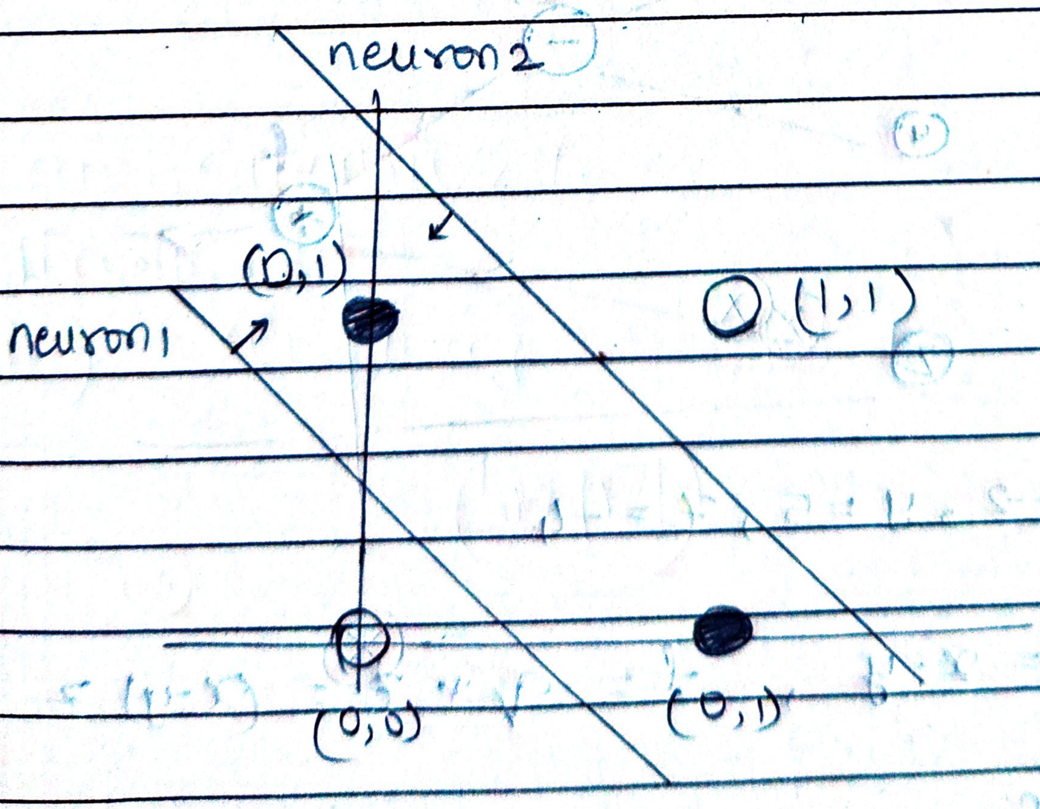
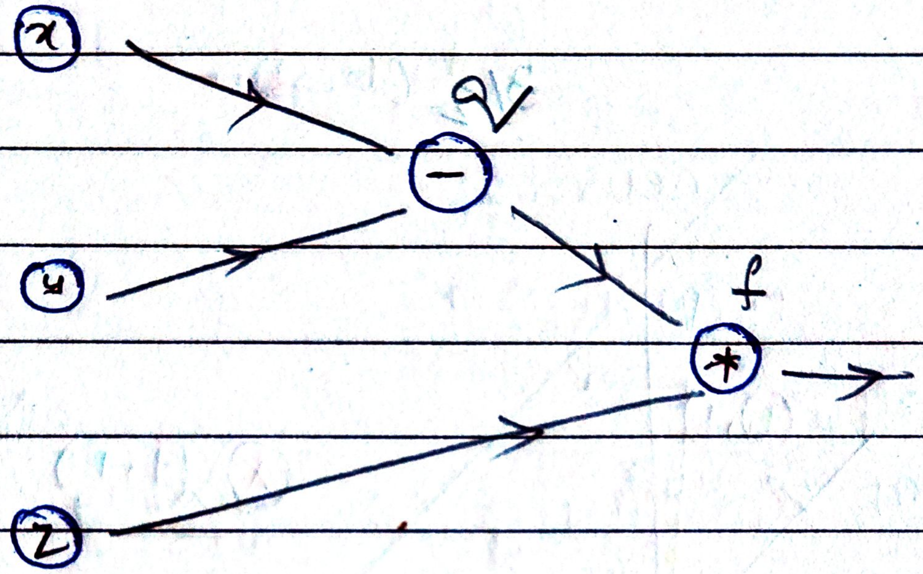


figure of XOR

b)



$x = -2,\ y = 5,\ z = -4$

$q = x - y,\quad f = q * z = (x-y)\,z$

$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}[xz - yz] = z = -4$

$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(xz - yz) = -z = 4$

$\frac{\partial f}{\partial z} = \frac{\partial}{\partial z}(xz - yz) = x - y = -7$

2) Given,
the cross entropy error function for multiclass classification problem

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

where, $y_k(x,w) = \dfrac{\exp(a_k(x,w))}{\sum_j \exp(a_k(x,w))}$

To find $\dfrac{\partial E}{\partial a_k}$. For this, we'll need $\dfrac{\partial y_i}{\partial a_j}$

When $i=j$:

$$\frac{\partial y_i}{\partial a_j} = \frac{\partial}{\partial a_j}\left[\frac{e^{a_j}}{\sum e^{a_k}}\right]$$

$$= \frac{e^{a_j}\left(\sum e^{a_k}\right) - e^{a_j} e^{a_j}}{\left(\sum e^{a_k}\right)^2}$$

$$= \frac{e^{a_j}}{\sum e^{a_k}} - \left(\frac{e^{a_j}}{\sum e^{a_k}}\right)^2 = y_j - y_j^2$$

$$= y_j(1-y_j)$$

1

When $i \neq j$:

$$\frac{\partial y_i}{\partial a_j} = \frac{\partial}{\partial a_j}\left[\frac{e^{a_i}}{\sum e^{a_k}}\right]$$

$$= \frac{0 - e^{a_j} e^{a_i}}{\left(\sum e^{a_k}\right)^2} = -y_i y_j \quad \text{—} \; (2)$$

from (1), (2)

$$\frac{\partial E}{\partial a_k} = -\sum_j \frac{\partial t_j \ln(y_j)}{\partial a_k}$$

$$= -\sum_j t_j \frac{\partial \ln(y_j)}{\partial a_k} = -\sum_j \frac{t_j}{y_j} \frac{\partial y_j}{\partial a_k}$$

$$= \frac{-t_k}{y_k} \frac{\partial y_k}{\partial a_k} - \sum_{j \neq k} \frac{t_j}{y_j} \frac{\partial y_j}{\partial a_k}$$

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

$$= -t_k + \sum_j t_j y_k = -t_k + y_k$$

$$\therefore \frac{\partial E}{\partial a_k} = y_k - t_k$$

3) Given,

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

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

To prove: $E_{ENS} \leq E_{AV}$

Substituting,

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

$$E_x\left[\left(\frac{\sum y_m(x)}{M}\right)^2 + f(x)^2 - \frac{2 f(x) \sum y_m(x)}{M}\right]$$

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

$\because E(aX + bY) = a E(X) + b E(Y)$.

$$\frac{1}{M^2} E_x\left[\left(\sum (y_m(x))\right)^2\right] + \cancel{E_x\left[f(x)^2\right]} - \frac{2}{M} E_x\left[f(x) \sum y_m(x)\right]$$

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

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

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

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

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

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

To prove: $E_x\left[\left(\frac{\sum_{m=1}^{M} y_m(x)}{M}\right)^2\right] \leq \frac{\sum_{m=1}^{M} E_x\left(y_m^2(x)\right)}{M}$ — 1

Jensen's inequality: If $h(x)$ is convex, & $a_i \geq 0, \forall i$, $\ni \sum a_i = 1$

$h(\sum a_i x_i) \leq \sum a_i h(x_i)$ — 2

Comparing 1, 2.
we get $a_i = \frac{1}{M}$, $h(x) = E_x(x^2)$.

So if $E_x(x^2)$ is convex, using 2,
we can prove 1.

Page

But $E_x(x^2)$ is obviously convex, as x^2 is convex, and

$$\frac{\partial^2}{\partial x^2}\left(E_x(x^2)\right) = 2\sum_{i=1}^{n} p_i > 0, \forall x$$

$$\left(\text{if } E_x(x^2) = \sum_{i=1}^{n} p_i x_i^2\right) \quad (\because p_i \geq 0, \forall i)$$

$\therefore$ the inequality holds, using Jensen's inequality.

By a similar argument, for any general error function $E(y)$, if its convex, by Jensen's inequality, it will satisfy the inequality

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