Foundations of Machine learning
Assignment - 3

Rollno: ES19BTECH11003
Hemanth.K

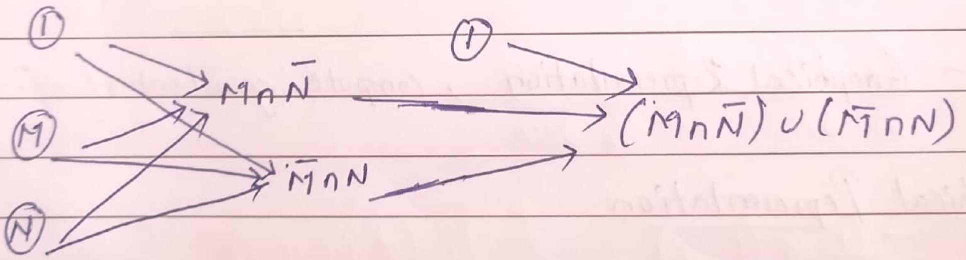
1) a) Required:-
Network diagram of two-Layer perceptron with 1 hidden Layer

We know $M \text{ XOR } N = (M \cap \bar{N}) \cup (\bar{M} \cap N)$

by XOR function

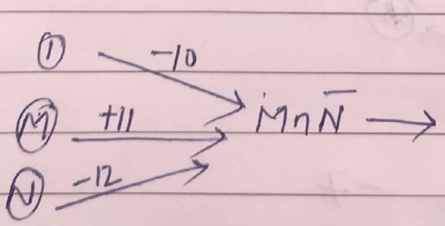
M XOR N

M	N	$M \cap \bar{N}$	$\bar{M} \cap N$	M XOR N
0	0	0	0	0
0	1	0	1	1
1	0	1	0	1
1	1	0	0	0



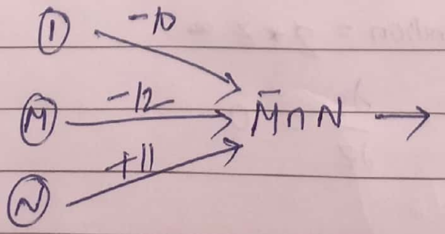
by the above figure designed.

$M \cap \bar{N}$



M	N	$\bar{N}$	$M \cap \bar{N}$
0	0	1	0
0	1	0	0
1	0	1	1
1	1	0	0

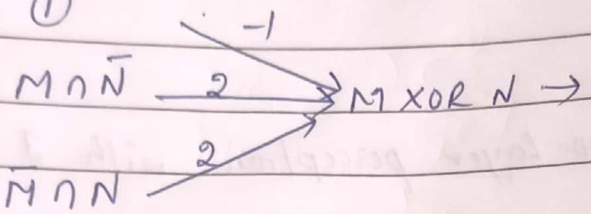
$\bar{M} \cap N$



M	N	$\bar{M}$	$\bar{M} \cap N$
0	0	1	0
0	1	1	1
1	0	0	0
1	1	0	0

M XOR N

①



M	N	M XOR N
0	0	0
0	1	1
1	0	1
1	1	0

by observing above table & diagram we can see in all cases $\sum w_i x_i \geq 0$ it is taken as 1 & $\sum w_i x_i < 0$ is taken as 0

b)

Given,

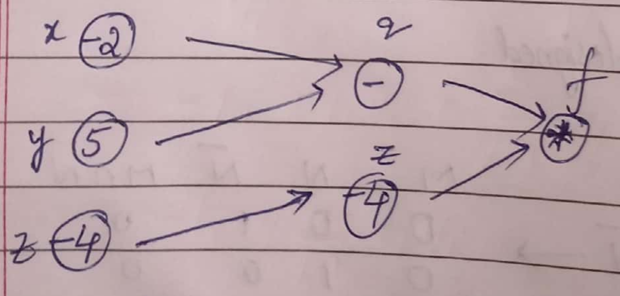
x, y, z are inputs with values $-2, 5$ & -4

Neuron q function $= x - y$

Neuron f function $= q * z$

To show:- Graphical Representation, compute gradient of f

Graphical Representation



here $q = -2-5 = -7$

q function $= x - y$, f function $= q * z$

$\frac{\partial q}{\partial x} = 1$, $\frac{\partial q}{\partial y} = -1$, $\frac{\partial q}{\partial z} = 0$

Scanned with CamScanner

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

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q} \cdot \frac{\partial q}{\partial x} = z(1) = -4$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial q} \frac{\partial q}{\partial y} = (z)(-1) = -(-4) = 4$$

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial q} \cdot \frac{\partial q}{\partial z} = (z)(0) = 0$$

2) Given, Extension of cross entropy error function

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

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

satisfies $0 \le y_k \le 1$, $\sum_k y_k = 1$

to show:- ~~Derivative~~ $\dfrac{\partial E}{\partial a_k} = y_k - t_k$

$y_k(x, w) = \dfrac{e^{a_k(x,w)}}{\sum_j e^{a_k(x,w)}}$ partially differentiate y wrt a

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

$$= \frac{e^{a_k}}{\sum_j e^{a_k}} - \left(\frac{e^{a_k}}{\sum_j e^{a_k}} \right)^2 = \frac{e^{a_k}}{\sum_j e^{a_k}} \left(1 - \frac{e^{a_k}}{\sum_j e^{a_k}} \right)$$

$$\frac{\partial y_k}{\partial a_k} = y_k(1-y_k) \quad \text{(1)}$$

for a_k, $$\frac{\partial y_i}{\partial a_k} = \frac{\partial \frac{e^{a_j}}{\sum_j e^{a_j}}}{\partial a_k} = \frac{-e^{a_i} e^{a_k}}{\sum_j^2 e^{a_j}}$$

$$= \frac{-e^{a_i}}{\sum_j e^{a_j}} \cdot \frac{e^{a_k}}{\sum_j e^{a_j}} = -y_i y_k \quad \text{(2)}$$

Now, $$\frac{\partial E}{\partial a_i} = -\sum_{j=1}^{k} \frac{\partial t_j \ln y_j}{\partial a_i} = -\sum_{j=1}^{k} t_j \frac{\partial \ln(y_j)}{\partial a_i}$$

$$= -\sum_{j=1}^{k} t_j \frac{1}{y_j} \frac{\partial y_j}{\partial a_i}$$

$$\cancel{\frac{\partial E}{\partial a_i} = -\sum_{j=1}^{k} \partial t_j \ln}$$

$$\frac{\partial E}{\partial a_i} = -\frac{t_i}{y_i}\frac{\partial y_i}{\partial a_i} - \sum_{j \neq i}^{k} \frac{t_j}{y_j}\frac{\partial y_j}{\partial a_i}$$

$$= -\frac{t_i}{\cancel{y_i}} \cancel{y_i}(1-y_i) - \sum_{j \neq i}^{k} \frac{t_j}{\cancel{y_j}} (\cancel{y_j}(1-y_i))$$

$$= -t_i + t_i y_i + \sum_{j \neq i}^{k} t_j y_i$$

$$= -t_i + \sum_{j=1}^{k} t_j y_i$$

$$= -t_i + y_i \sum_{j=1}^{k} t_j$$

$$= y_i - t_i$$

$$\boxed{\text{Hence, } \frac{\partial E}{\partial a_k} = y_k - t_k}$$

3) Given, Convex function $f(x) = x^2$;

to show:- Average expected sum of squares error $E_{AV} = \frac{1}{M} \sum_{m=1}^{M} E_x[(y_m(x) - f(x))^2]$

Expected Error $E_{ENS} = \mathbb{E}_x[(\frac{1}{M} \sum_{m=1}^{M} y_m(x) - f(x))^2]$ to satisfy $E_{ENS} \leq E_{AV}$

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

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

by use of Jensen's Equality, for a function h is

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

if we consider all the weights to be 1 $\sum^{M} a_i = M$

Let us assume $x_i = [y_m(x) - f(x)^2]$
$h = E_x$

Substitute x_i & h in convex function in equality

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

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

Hence with E_{AV} & E_{ENS} expressions we get $E_{ENS} \leq E_{AV}$

In above proof, we didn't use function value at anywhere so above statement is valid for all the error functions.