Q. Multi layer perceptron for solve non-linear seperable.

Theory:

Single layer perceptron fail for X-OR problem because—
The X-OR function outputs 1 when two binary input differ otherwise 0. For X-OR—

0 0 → 0
0 1 → 1
1 0 → 1
1 1 → 0

If we plot these points the (1,0) are on the (0,1) and (1,0) corners and (0,0) are on the (0,0) and (1,1) So No single straight line can seperate all 1's output from all 0's output.

This means the X-OR function is not linear seperable so a single layer perceptron cannot solve it. A multi layer perceptron overcomes this by transforming the data in a hidden layer, making seperation possible.

Training data

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

we use this particular equation, $y = x_1\bar{x}_2 + \bar{x}_1 x_2$

$y = z_1 + z_2$

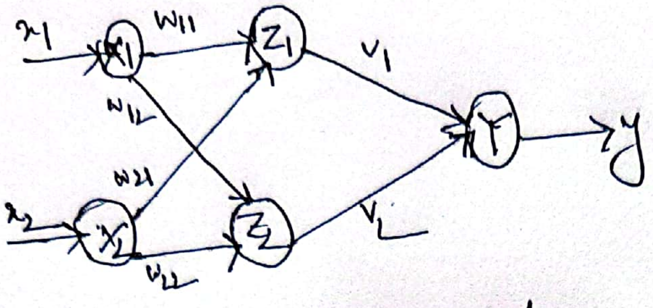
∴ so, $y = z_1 + z_2$

where, $z_1 = x_1\bar{x}_2$

$z_2 = \bar{x}_1 x_2$

$y = z_1 \text{ (OR) } z_2$

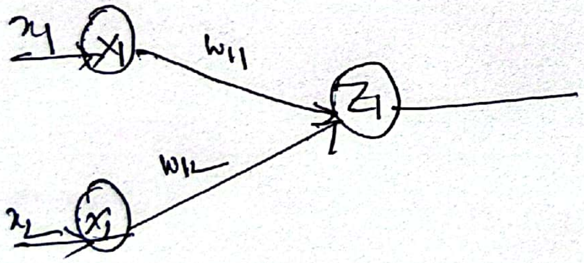
So,
Network diagram,



Here, x_1, x_2 are two input
Z_1 and Z_2 are intermediatary perceptron
y is the output

Activation ~~step~~ Function

$$f(y_{in}) = \begin{cases} 1 \; ; \text{ if } y_{in} \geq \theta \\ 0 \; ; \text{ if } y_{in} < \theta \end{cases}$$. Here θ is threshold



Learning Rule

(i) compute weighted sum $= \sum w_i x_i$
(ii) Activation function used . and find perceptron output.
(iii) If target = output ; No ~~cap~~ weight update else
$w_i = w_i + \eta (t - y) x_i$

Here, η = learning rate
t = target
y = perceptron output
x_i = input
w_i = old weight

Example:

Assume,
initial weight are [1,1]
Threshold $= 1$ and
learning rate $= 1.5$ so,

and

x_1	x_2	y	z_1
0	0	0	0
0	1	10	0
1	0	1	1
1	1	0	0

So, $z_1 = x_1 \bar{x}_2$

Epoch-1

(i) input (0,0) target $= 0$
weighted sum $= 1\times0 + 1\times0 = 0$
step function $= 0 < 1$ so, output $= 0$
So, No weighted update (output = target)

(ii) input (0,1) target $= 0$
weighted sum $= 1\times0 + 1\times1 = 1$
step function $1 = 1$ so output $= 1$
output $\neq$ target so
$w_{11} = 1 + 1.5\,(0-1) * 0 = 1$
$w_{21} = 1 + 1.5\,(0-1) * 1 = -0.5$

(iii) input (1,0) target $= 1$
weighted sum $= 1*1 + (-0.5)*0 = 1$
No update

(iv) input (1,1) target $= 0$
weighted sum $= 1*1 + (-0.5)*1 = 1 - 0.5 = 0.5$
step function $=$ output $= 0$
No update

~~Epoch-2~~

$$\boxed{\begin{aligned} \text{So, } w_{11} &= 1 \\ w_{21} &= -0.5 \end{aligned}}$$

Again for 2nd Function,

$z_2 = \bar{x}_1 x_2$

x_1	x_2	y	z_2
0	0	0	0
0	1	1	1
0	1	1	0
1	0	1	0
1	1	0	0

And Assume weight $= [1, 1]$
Threshold $= 1$ and
learning rate $= 1.5$

So,

Epoch-1

(i) input (0, 0), ~~target~~ target = 0
weighted sum $= 1*0 + 1*0 = 0$
step function $= 0 < 1$ so output $= 0$
No update (target = output)

(ii) input (0, 1) target = 1
So, weighted sum $= 1*0 + 1*1 = 1$
Step function $= 1 \leq 1$ So output $= 1$
No update (target = output)

(iii) input (1, 0) target = 0
weighted sum $= 1*1 + 0*1 = 1$
step function $= 1 = 1$ so output $= 1$
update (target ≠ output)
$w_{12} = 1 + 1.5 (0-1)* = -0.5$
$w_{22} = 1 + 1.5 (0-1)*0 = 1$

(iv) input (1, 1) target = 0
weighted sum $= (-0.5)*1 + 1*1 = 0.5$
step function $= 0.5 < 1$ so output $= 0$
No update

$$\boxed{\text{So, } w_{12} = -0.5 \quad w_{22} = 1}$$

Now,

x_1	x_2	z_1	z_2	y
0	0	0	0	0
0	1	0	1	1
1	0	1	0	1
1	1	0	0	0

because,

$$y = z_1 \text{ (OR) } z_2$$

Assume initial weight $[1,1] = [v_1, v_2]$

Threshold $= 1$ and

learning rate $= 1.5$

Epoch - 1

(i) input (0, 0) target = 0

so, weighted sum = $1*0 + 1*0 = 0$

step function $= 0 < 1$ so output 0

No update

(ii) input (0, 1) target = 1

weighted sum = $1*0 + 1*0 = 1$

step function $= 1 = 1$ so output = 1

No update

(iii) input (1, 0) target = 1

weighted sum = $1*1 + 1*0 = 1$

step function = $1 = 1$ so output = 1

No update

(iv) input (0, 0) target = 1

weighed sum = $1*0 + 1*0 = 0$

step function $= 0 < 1$ so output = 0

No update

So, $v_1 = v_2 = 1$