☆ Spring 2016 Midterm

Q1] Entropy $X_1 = \frac{4}{10}\left[-\frac{1}{4}\log\frac{1}{4} - \frac{3}{4}\log\frac{3}{4}\right]$

$+\frac{6}{10}\left[-\frac{1}{3}\log\frac{1}{3} - \frac{2}{3}\log\frac{2}{3}\right]$

Entropy $X_2 = \frac{3}{10}[0] + \frac{7}{10}\left[-\frac{2}{7}\log\frac{2}{7}\right.$

$\left.-\frac{5}{7}\log\frac{5}{7}\right]$

Entropy $X_3 = \frac{6}{10}[1] + \frac{4}{10}[1]$

Entropy $X_4 = \frac{4}{10}[1] + \frac{6}{10}[1]$

X_1	X_3	X_4	Y
1	1	0	1
0	0	1	1
0	1	1	-1
0	1	0	-1
0	0	0	-1
1		0	-1
0	1	1	-1

Choose X_2,

Level 2,

Entropy X_1 is least

Choose X_1

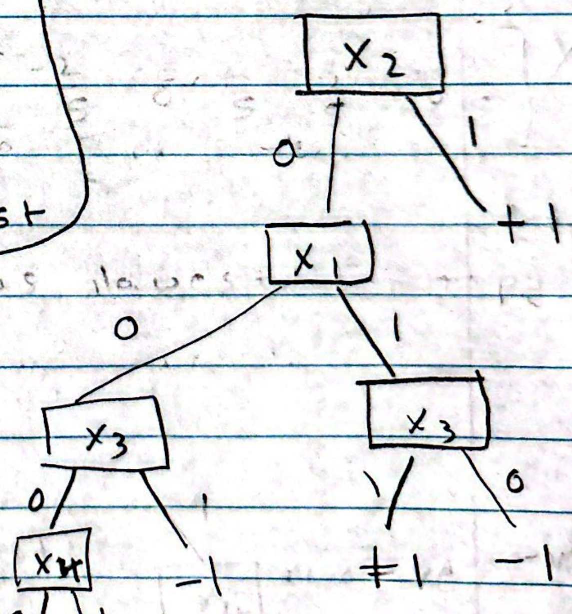
Level 3,

$X_1 = 1$, $X_1 = 0$

X_3	X_4	Y
1	0	1
0	0	-1

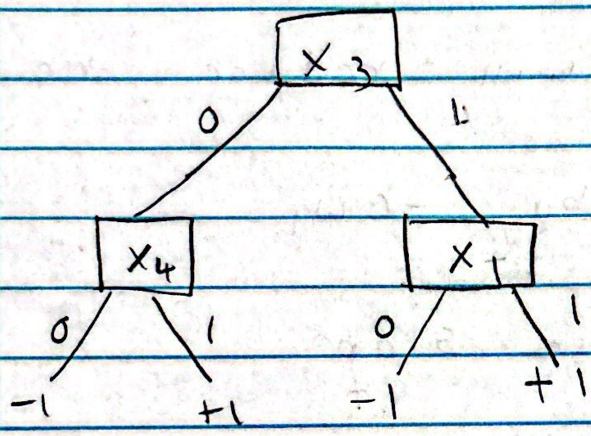
X_3	X_4	Y
0	1	1
1	1	-1
0	0	-1
1	0	-1
1	1	-1

Choose X_3



b] How to find solution?
↳ Find 2 or 3 attribute that do not have contradictory Y output

X_1, X_3 & X_4 satisfy this condition



c] we will prefer b) tree because it has smaller height so it is likely to be optimal. Also tree with more nodes repersenting the same concept are more likely to overfit the training examples

---

Question 2]

a] Refer to Spring 2015 midterm Q2] a]

b] $w_1 x_1 + w_2 x_2 = 0$

slope $= \frac{-w_1}{w_2} = -\left(\frac{w_1}{w_2}\right)$

* slope increases when w_1 derearases

$-\frac{c}{2}(w_1)^2$ will reduce the value

of w_1 by $-cw_1$

* will increase slope

- L2 -> Yes, is possible if we select medium large value of c slope will increase from L1. So it is possible to have L2

- L3 -> No, for L3 to be possible ~~slop~~. c has to be negative

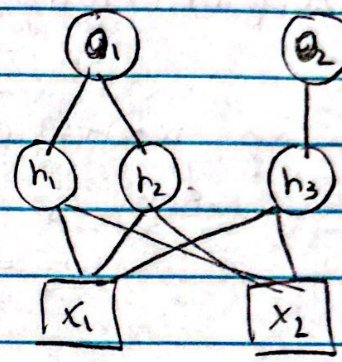
- L4 -> Yes, but the value of c has to be very large to increase slope by that much. It is not complete vertical, so it is therotically possible

Q3] $y_1 = XOR = (x_1 \wedge x_2) \vee (-x_1 \wedge -x_2)$

$y_2 = (x_1 \wedge -x_2)$

$h_1 = x_1 \wedge x_2$

$w_0 = -1.5,\ w_1 = 1,\ w_2 = 1$



$h_2 = -x_1 \wedge -x_2$

$w_0 = 0.5,\ w_1 = -1,\ w_2 = -1$

$O_1 = h_1 \vee h_2$

$w_0 = 1.5,\ w_1 = 1,\ w_2 = 1$

$h_3 = x_1 \wedge -x_2$

$w_0 = -0.5,\ w_1 = 1,\ w_2 = -1$

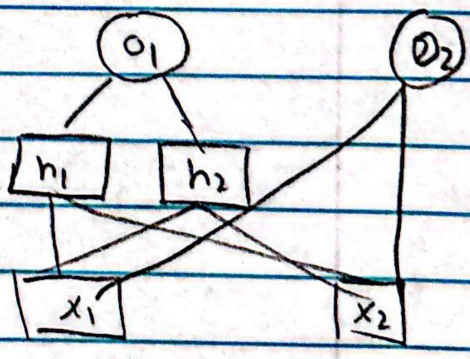
$o_2 = \cancel{w} \cdot h_3$

$w_0 = 0;\ w_1 = 1$

All units ~~o~~ h_1, h_2, h_3, o_1, o_2 are signed units i.e if $f(x) > 0$, 1 eles 0

Alternate solution,

Skip h_3 if allowed



b]

$$\frac{\partial E}{\partial w_j} = -\sum_{i=1}^{m} (y_i - o_i)^2 (x_j + x_j^{1.5})$$

Algorithm)

Loop until convergence

for j in features n

$$w_j = w_j - \sum_{d=1}^{m} (y_d - o_d)^2 (x_{jd} + x_{jd}^{1.5})$$

end for

end loop

equivalent to $x_i d$ if i represents i^{th} feature from 1 to n