$E(\text{Offer Taken}) = -P(Y)\lg_2 P(Y) - P(N)\lg_2 P(N)$

$= -\frac{7}{12}\lg_2\frac{7}{12} - \frac{5}{12}\lg_2\frac{5}{12}$

$= 0.98$

$E(\text{Job Type} = \text{Engineer}) = -P(Y|E)\lg_2 P(Y|E) - P(N|E)\lg_2 P(N|E)$

$= -\frac{3}{5}\lg_2\frac{3}{5} - \frac{2}{5}\lg_2\frac{2}{5}$

$= 0.97$

$E(\text{Job Type} = \text{Doctor}) = -P(Y|D)\lg_2 P(Y|D) - P(N|D)\lg_2 P(N|D)$

$= -\frac{1}{3}\lg_2\frac{1}{3} - \frac{2}{3}\lg_2\frac{2}{3}$

$= 0.92$

$E(\text{Job Type} = \text{Teacher}) = -P(Y|T)\lg_2 P(Y|T) - P(N|T)\lg_2 P(N|T)$

$= -\frac{3}{4}\lg_2\frac{3}{4} - \frac{1}{4}\lg_2\frac{1}{4}$

$= 0.81$

$IG(\text{Job Type}) = 0.98 - \frac{5}{12}(0.97) - \frac{3}{12}(0.92) - \frac{4}{12}(0.81)$

$= 0.076$

$E(\text{Income Level} = \text{High}) = -P(Y|\text{High}) \, Lg_2 \, P(Y|\text{High}) - P(N|\text{High}) \, Lg_2$

$= -\frac{3}{8} Lg_2 \frac{3}{8} - \frac{5}{8} Lg_2 \frac{5}{7}$

$= 0.95$

$E(\text{Income Level} = \text{Medium}) = -P(Y|M) \, Lg_2 \, P(Y|M) - P(N|M) \, Lg_2 \, P(N|M)$

$= -\frac{4}{4} Lg_2 \frac{4}{4} - \frac{0}{4} Lg_2 \frac{0}{4}$

$= 0$

$IG(\text{Income}) = 0.98 - \frac{8}{12}(0.95) - \frac{4}{12}(0)$

$= 0.35$

$E(\text{Likes to Hangout} = \text{yes}) = -P(Y|y) \, Lg_2 \, P(Y|y) - P(N|y) \, Lg_2 \, P(N|y)$

$= -\frac{3}{5} Lg_2 \frac{3}{5} - \frac{2}{5} Lg_2 \frac{2}{5}$

$= 0.970$

$E(\text{Likes to Hangout} = \text{no}) = -P(Y|\text{no}) \, Lg_2 \, P(Y|\text{no}) - P(N|\text{no}) \, Lg_2 \, P(N|\text{no}) \, Lg_2$

$= -\frac{4}{7} Lg_2 \frac{4}{7} - \frac{3}{7} Lg_2 \frac{3}{7}$

$= 0.985$

$IG(\text{Likes to Hangout}) = \cancel{1-0} \; 0.98 - \frac{5}{12}(0.970) - \frac{7}{12}(0.985)$

$= \cancel{0.002} \; 0.00125$

$E(\text{Tours per yr} = 1) = -P(Y|1)\log_2 P(Y|1) - P(N|1)\log_2 P(N|1)$

$= -\frac{0}{4}\log_2\frac{0}{4} - \frac{3}{4}\log_2\frac{3}{4}$

$= 0.311$

$E(\text{Tours per yr} = 2) = -P(Y|2)\log_2 P(Y|2) - P(N|2)\log_2 P(N|2)$

$= -\frac{4}{5}\log_2\frac{4}{5} - \frac{1}{5}\log_2\frac{1}{5}$

$= 0.722$

$E(\text{Tours/yr} = 3) = -P(Y|3)\log_2 P(Y|3) - P(N|3)\log_2 P(N|3)$

$= -\frac{3}{3}\log_2\frac{3}{3} - \frac{0}{3}\log_2\frac{0}{3}$

$= 0$

$IG(\text{Tours/yr}) = 0.980 - \frac{4}{12}(0.311) - \frac{5}{12}(0.722) - \frac{3}{12}(0)$

$= 0.576$

When Tours/yr = 2

$E(\text{Job Type} = \text{Engineer}) = -P(Y|E)\log_2 P(Y|E) - P(N|E)\log_2 P(N|E)$

$= -\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3}$

$= 0.918$

$E(\text{Job Type} = \text{Teacher}) = -P(Y|T)\log_2 P(Y|T) - P(N|T)\log_2 P(N|T)$

$= -\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2}$

$= 0$

$IG(\text{Job Type}) = 0.980 - \frac{3}{5}(0.981) - \frac{2}{5}(0)$

$= 0.3914$

$E(\text{Income} = \text{High}) = -P(Y|H)\log_2 \overset{P(Y|H)}{} - P(N|H)\log_2 P(N|H)$

$$= -\frac{2}{3}\log_2\left(\frac{2}{3}\right) - \frac{1}{3}\log_2\frac{1}{3}$$

$$= \cancel{0.918}\ 0.918$$

$E(\text{Income} = \text{Medium}) = -P(Y|M)\log_2 P(Y|M) - P(N|M)\log_2 P(N|M)$

$$= -\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2}$$

$$= 0$$

$IG(\text{Income}) = \cancel{0.970 - \frac{3}{5}(0.98}$

$$= 0.970 - \frac{3}{5}(0.918) - \frac{2}{5}(0)$$

$$= 0.4292$$

$E(\text{Likes to Hangout} = \text{Yes}) = -P(Y|Yes)\log_2 P(Y|Yes) - P(N|Yes)\log_2 P(N|Yes)$

$$= -\frac{1}{1}\log_2\frac{1}{1} - \frac{0}{1}\log_2\frac{0}{1}$$

$$= 0$$

$E(\text{Likes to Hangout} = \text{No}) = -P(Y|No)\log_2 P(Y|No) - P(N|No)\log_2 P(N|No)$

$$= -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4}$$

$$= 0.811$$

$IG(\text{Likes to Hangout}) = 0.970 - \frac{1}{5}(0) - \frac{4}{5}(0.811)$

$$= 0.3312$$

...en Income Level = High given Tours/yr = 2

$E(\text{Job Type} = \text{Engineer}) = -P(Y|E)\log_2 P(Y|E) - P(N|E)\log_2 P(N|E)$

$$= -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2}$$

$$= 1$$

$E(\text{Job Type} = \text{Teacher}) = -P(Y|T)\log_2 P(Y|T) - P(N|T)\log_2 P(N|T)$

$$= -\frac{1}{1}\log_2\frac{1}{1} - \frac{0}{1}\log_2\frac{0}{1}$$

$$= 0$$

$E(\text{Job Type}) = 0.980 - \frac{2}{3}(1) - \frac{1}{3}(0)$

$$= 0.313$$

$E(\text{Likes to Hangout} = \text{Yes}) = -P(Y|Yes)\log_2 P(Y|Yes) - P(N|Yes)\log_2 P(N|Yes)$

$$= -\frac{1}{1}\log_2\frac{1}{1} - \frac{0}{1}\log_2\frac{0}{1}$$

$$= 0$$

$E(\text{Likes to Hangout} = \text{No}) = -P(Y|No)\log_2 P(Y|No) - P(N|No)\log_2 P(N|No)$

$$= -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2}$$

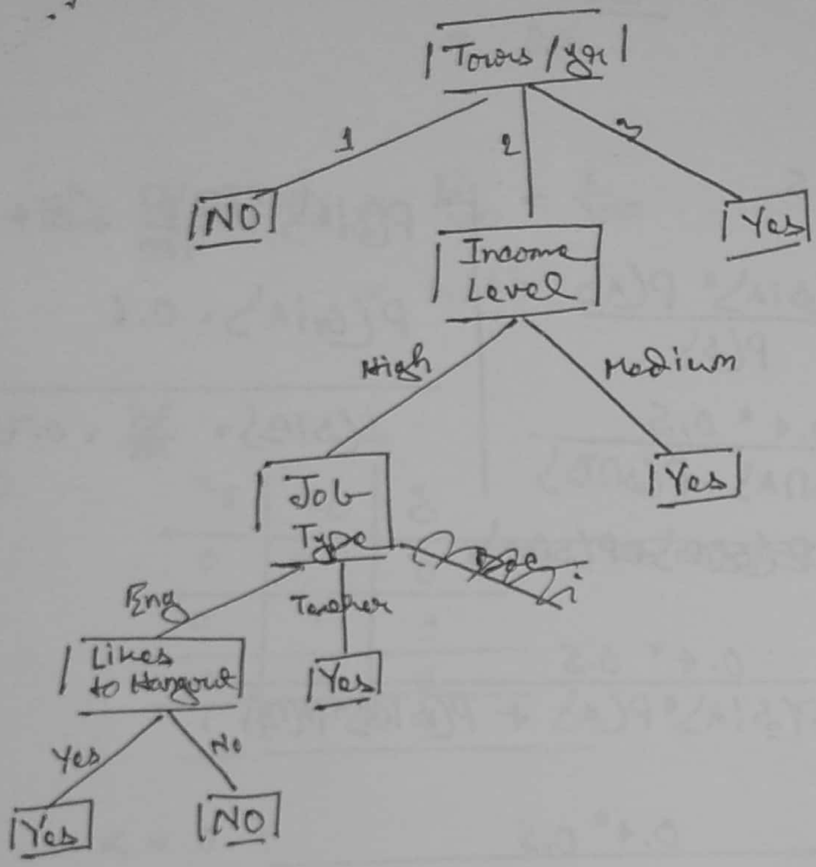
$$= 1$$

$IG(\text{Likes to Hangout}) = 0.980 - \frac{1}{3}(0) - \frac{2}{3}(1)$

$$= 0.313$$

~~Income~~

Income Level is a better attribute as it has high I.



Tours / yr
1
2
3
NO
Yes
Income Level
High
Medium
Yes
Job Type
Eng
Teacher
Likes to Hangout
Yes
Yes
No
Yes
NO

2)

A
60 G
40 S

B
25 G
75 S

$P(A) = P(B) = 0.5$

$P(S|A) = \frac{40}{100} = 0.4$

$P(G|A) = 0.6$

$S(S|B) = \frac{75}{100} = 0.75$

$$P(A|S) = \frac{P(S|A) * P(A)}{P(S)}$$

$$= \frac{0.4 * 0.5}{P(S \cap A) + P(S \cap B)}$$

$\cancel{+P(S \cap B) + P(S \cap A)}$

$$= \frac{0.4 * 0.5}{P(S|A) * P(A) + P(S|B) * P(B)}$$

$$= \frac{0.4 * 0.5}{0.4 * 0.5 + 0.75 * 0.5}$$

$= 0.348$

---

3)

200 Male
- -> 80 Juniors -> 50% interested
- -> 120 Seniors -> 70% interested

100 Female
- -> 70 Juniors -> 60%
- -> 30 Seniors -> 80%

(a)

	Junior		Senior	
Gender (G)	H	¬H	H	¬H
Male	$\frac{40}{300}$ = 0.13	$\frac{40}{300}$ = 0.13	84/300 = 0.28	$\frac{36}{300}$ = 0.12
Female	$\frac{42}{300}$ = 0.14	$\frac{28}{300}$ = 0.093	$\frac{\cancel{42}\,24}{300}$ = 0.08	$\frac{6}{300}$ = 0.02

P(Being Junior) = $\frac{40}{300} + \frac{40}{300} + \frac{42}{300} + \frac{28}{300}$

= 0.5

(ii) $P(F \cap \neg H) = \frac{28}{300} + \frac{6}{300}$

= 0.113

4)

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

$\alpha = 0.1$

Eqⁿ => $0.5x_1 + 0.5x_2 - 1.25$

$w.x = 0.5(0) + 0.5(0) - 1.25$
$= -1.25$

$h_w(x) = \text{Threshold}(-1.25)$
$= 0$

$y = h_w(x)$
So unchanged

$g(w.x) = \frac{1}{1+e^{-w.x}}$
$= \frac{1}{1+e^{-(-1.25)}}$
$= 0.223$

$w.x = 0.5(0) + 0.5(1) - 1.25$
$= -0.75$

$h_w(-0.75) = \text{Threshold}(-0.75)$
$= 0$

$y \neq h_w(x)$

$g(-0.75) = \frac{1}{1+e^{0.75}}$
$= 0.321$

~~$w_1 = 0.5 + 0.1 * (1-0) * 0 = 0.5$~~
~~$w_2 = 0.5 + 0.1 * (1-0) * 0 = 0.5$~~
~~$w_3 = 0 - 1.25 + 0.1 * (1-0) * 1 = -1.15$~~

~~∴ Updated eqⁿ:~~
~~$0.5x_1 + 0.5x_2 - 1.15 = 0$~~

$w.x = 0.5(1) + 0.5(0) - 1.15$
$= -0.65$

$hw(-0.65) = \text{Threshold}(-0.65)$
$= 0$

$g(w.x) = \left(\frac{1}{1+e^{0.65}}\right)$
$=$

$w_1 = 0.5 + 0.1 * (1-0) * 0.321 * (1-0.321) * 0 = 0.5$

$w_2 = 0.5 + 0.1 * (1-0) * 0.321 * (1-0.321) * 1 = \cancel{0.502}\ 0.52$

$w_3 = -1.25 + 0.1 * (1-0) * 0.321 * (1-0.321) * 1 = -1.23$

$\therefore$ Updated eqn:

$$hw(x) = 0.5x_1 + 0.5x_2 - 1.23 = 0$$

$w.x = 0.5(1) + 0.5(0) - 1.23$
$= -0.73$

$hw(-0.73) = \text{Threshold}(-0.73)$
$= 0$

$g(-0.73) = \frac{1}{1+e^{0.73}}$
$= 0.325$

$y \neq hw(x)$

$w_1 = 0.5 + 0.1 * (1-0) * 0.325 * (1-0.325) * 1 = 0.52$

$w_2 = 0.52 + 0.1 * (1-0) * 0.325 * (1-0.325) * 0 = 0.52$

$w_3 = -1.23 + 0.1 * (1-0) * 0.325 * (1-0.325) * 1 = -1.21$

$\therefore$ Updated eqn:

$$hw(x) = 0.52x_1 + 0.52x_2 - 1.21 = 0$$

$= 0.52(1) + 0.52(1) - 1.21$

$= -0.17$

$h_w(-0.17) = \text{Threshold}(-0.17)$

$= 0$

$g(-0.17) = \frac{1}{1+e^{0.17}}$

$= \cancel{0.5}\ 0.458$

$y \neq h_w(x)$

$w_1 = 0.52 + 0.1 * (1-0) * 0.458 * (1-0.458) * 1 = 0.54$

$w_2 = 0.52 + 0.1 * (1-0) * 0.458 * (1-0.458) * 1 = 0.54$

$w_3 = -1.21 + 0.1 * (1-0) * 0.458 * (1-0.458) * 1 = -1.19$

$\therefore$ Updated eq^n : $0.54x_1 + 0.54x_2 - 1.19 = 0$

5)

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

$\alpha = 0.1$

$w \cdot x = 0.75(0) + 0.75(0) - 1.25$
$= -1.25$

$h_w(-1.25) = \text{Threshold}(-1.25)$
$= 0$

$y = h_w(x)$

$\therefore$ unchanged

$w \cdot x = 0.75(0) + 0.75(1) - 1.25$
$= -0.5$

$h_w(-0.5) = \text{Threshold}(-0.5)$
$= 0$

$y = h_w(x)$

$\therefore$ unchanged

$w \cdot x = 0.75(1) + 0.75(0) - 1.25$
$= -0.5$

$h_w(-0.5) = \text{Threshold}(-0.5)$
$= 0$

$y = h_w(x)$

$\therefore$ unchanged

$w \cdot x = 0.75(1) + 0.75(1) - 1.25$
$= 0.25$

$h_w(0.25) = \text{Threshold}(0.25)$
$= 1$

$y = h_w(x)$

$\therefore$ unchanged

So, the eqn will remain unchanged.

P(Play Tennis = Yes | Outlook = Sunny, Temperature = 20, Humidity = Normal, Wind = Strong) = ?

P(Play Tennis = No | Outlook = Sunny, Temp = 20, Humidity = Normal, Wind = Strong) = ?

$P(Y | S \cap 20 \cap N \cap Str) = P(Sunny \cap 20 \cap N \cap Str | Y) * P(Y)$

$= P(Sunny|Y) * P(20|Y) * P(N|Y) * P(Str|Y) * P(Y)$

$= \frac{2}{9} * P(20|Y) * \frac{6}{9} * \frac{4}{9} * \frac{9}{14}$

$= \frac{2}{9} * 0.0073 * \frac{6}{9} * \frac{4}{9} * \frac{9}{14}$

$= 0.00031$

$P(N | S \cap 20 \cap N \cap Str) = P(S \cap 20 \cap N \cap Str | N) * P(N)$

$= P(S|N) * P(20|N) * P(Normal|N) * P(Str|N) * P(N)$

$= \frac{3}{5} * 0.000063 * \frac{1}{5} * \frac{5}{5} * \frac{5}{14}$

$= 0.0000027$

So, the prediction is more likely to be yes.

$$u_Y = \frac{1}{N} \sum_{n=1}^{N} x_n$$

$$= \frac{1}{9} \left[42 + 28 + 15 + 12 + 16 + 27 + 25 + 29 + 38 \right]$$

$$= 25.78$$

$$u_N = \frac{1}{5} \left[40 + 45 + 18 + 26 + 24 \right]$$

$$= 30.6$$

$$\sigma_Y^2 = \frac{1}{N} \sum_{n=1}^{N} (x_n - u)^2$$

$$= \frac{1}{9} \Big[(42 - 25.78)^2 + (28 - 25.78)^2 + (15 - 25.78)^2 + (12 - 25.78)^2 + (16 - 25.78)^2 + (27 - 25.78)^2 + (25 - 25.78)^2 + (29 - 25.78)^2 + (38 - 25.78)^2 \Big]$$

$$= 92.40$$

$$\therefore \sigma_Y = 9.61$$

$$\sigma_N^2 = \frac{1}{5} \Big[(40 - 30.6)^2 + (45 - 30.6)^2 + (18 - 30.6)^2 + (26 - 30.6)^2 + (24 - 30.6)^2 \Big]$$

$$= 103.84$$

$$\therefore \sigma_N = 10.19$$

$$...|Y) = \frac{1}{\sqrt{2\pi * 9.61}} e^{\left(-\frac{(20-25.78)^2}{2 \times 9.61}\right)}$$

$$= \cancel{0.00}\ 0.0073$$

$$P(20|N) = \frac{1}{2\pi * 10.19} e^{\left(-\frac{(20-30.6)^2}{2 \times 10.19}\right)}$$

$$= 0.000063$$