Assignment - 4

Shruthi Reddy Pingili
1001716021.

1) We have 11 variables.

Q1; (a) A has 5 values.

$B_1, B_2, B_3, \ldots, B_{10}$ has 7 values.

So, total $A \times B_1 \times B_2 \times \cdots \times B_{10}$.

Hence $\underbrace{5 \times 7 \times 7 \times 7 \times \cdots \times 7}_{10 \text{ times}}$.

5×7^{10} values of $5 \times 7^{10} - 1$ numbers

So, this can be done in $5 \times 7^{10} - 1$ ways.

(b) $P(A, B_1, B_2, \ldots, B_{10})$.

$= P(B1/A)\, P(B2/A)$.

Hence, $P(B_i/A)$ needs $(7-1) \times 5 = 30$ values.

$P(A)$ needs $5 - 1 = 4$ values.

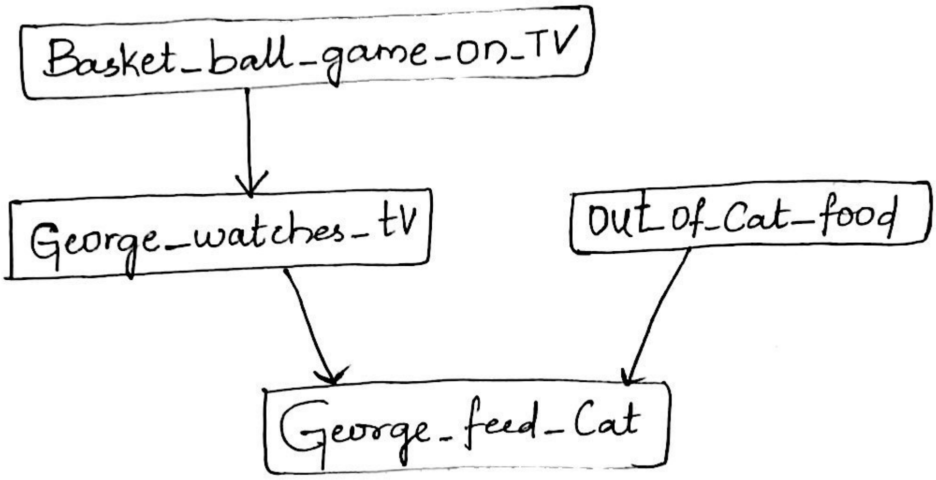
So Total: $30 \times 10 + 4$

$300 + 4$

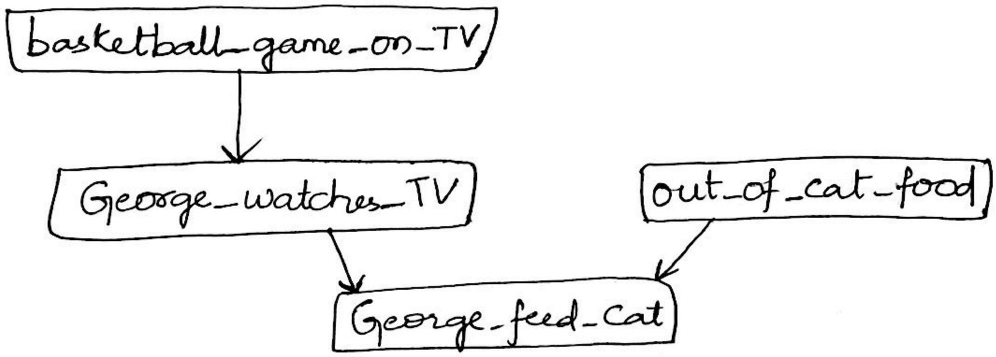
$= 304$ values.

$\therefore$ Total 304 values.

Task-3



Task-4



P(B)
0.30411

P(0.41)
0.1698

B	P(TV)
T	0.9279
F	0.1181

TV	Out	P(C=T)	P(C=F)
T	T	0.0416	0.9583
T	F	0.7064	0.2935
F	T	0.3157	0.6842
F	F	0.9587	0.04124

Here,

TV: George watches TV

Out: Out of CAT food.

5) a) Markovian Blanket of N

Sol^n:

Children of N : R, S

Parents of Children of node N : M, O.

Parent of N : I.

Markovian Blanket : R, S, M, O, I.

b) $P(A, F) = P(F/A) \times P(A)$.

$= 0.8 \times 0.8$

$= 0.64$.

c) $P(N, \text{not}(D)/I) = \dfrac{P(N, D', I)}{P(I)}$.

$$= \frac{P(N/I)\, P(I/D')\, P(D')}{P(I/D)\, P(D) + P(I/D')\, P(D')}$$

$$= \frac{0.1 \times 0.2 \times 0.5}{0.5 \times 0.5 + 0.2 \times 0.5}$$

$$= \frac{0.01}{0.35}$$

$$= \frac{1}{35}$$

$\therefore P(N, \text{not}(D)/I) = 1/35$.

6) Consider 1st case with A as root node.

Sol^n:

Scanned with CamScanner

A

1

Class	A=1	B	C
X		2	1
X		3	3
Y		2	2

Total = 3
X = 3
y = 0

2

Class	A=2	B	C
X		1	2
Y		1	1
Y		1	2
Y		1	1

x = 1
y = 3
Total = 4

3

Class	A=3	B	C
X		2	2
Y		1	1
Y		3	1

Total = 3
X = 1
y = 2

① $H(\epsilon) = 1.$

② $H(\epsilon 1) = 0.$

③ $H(\epsilon_3) = \frac{1}{4}\log_2\left(\frac{1}{4}\right) - \frac{3}{4}\log_2\left(\frac{3}{4}\right)$

$= 0.81127.$

④ $H(\epsilon_3) = \frac{-1}{3}\log_2\left(\frac{1}{3}\right) - \frac{2}{3}\log_2\left(\frac{2}{3}\right)$

$= 0.5278 - 0.3956$

$= 0.9234.$

InfoGain(A).

$= H(\epsilon) - \frac{4}{10}H(\epsilon_2) - \frac{3}{10}H(\epsilon_3).$

$= 0.3858.$

In the second case with B as root node.

$H(\epsilon) = 1$

$H(\epsilon_1) = \frac{-1}{4}\log_2\left(\frac{1}{4}\right) - \frac{3}{4}\log_2\left(\frac{3}{4}\right).$

$= 0.81127.$

$$H(E_2) = \frac{-3}{4}\log_2\left(\frac{3}{4}\right) - \frac{1}{4}\log_2\left(\frac{1}{4}\right).$$

$= 0.821127.$

$H(E_3) = 1.$

$$\text{InfoGain}(B) = 1 - \frac{4}{10}H(E_1) - \frac{4}{10}H(E_2) - \frac{2}{10}H(E_3)$$

$= 0.1512.$

Considering 3rd case with c as root node,

$H(E) = 1$

$$H(E_1) = \frac{-1}{5}\log_2\left(\frac{1}{5}\right) - \frac{4}{5}\log_2\left(\frac{4}{5}\right).$$

$= 0.7218.$

$$H(E_2) = \frac{-3}{4}\log_2\left(\frac{3}{4}\right) - \frac{1}{4}\log_2\left(\frac{1}{4}\right).$$

$= 0.81127.$

$H(E_3) = 0.$

$$\text{InfoGain}(C) = H(C) - \frac{5}{10}H(E_1) - \frac{4}{10}H(E_2) - \frac{1}{10}H(E_3).$$

$$= 1 - \left(\frac{1}{2} \times 0.7218\right) - \left(\frac{2}{5}\,(2 \times 0.8 + 1.29)\right).$$

$= 0.3146.$

InfoGain(A) > InfoGain(C) > InfoGain(B).

$0.3858 > 0.3146 > 0.1512.$

$\therefore$ A receives highest info Gain.

7)
Sol^n^

a) Entropy $H(A) = H\left(\frac{80}{100}, \frac{20}{100}\right)$.

$$H(A) = \frac{-80}{100}\log_2\left(\frac{80}{100}\right) - \frac{20}{100}\log_2\left(\frac{20}{100}\right).$$

$$H(A) = 0.2575 + 0.4643.$$

$$H(A) = 0.7218.$$

b) Info Given

$$= H(A) - \frac{35}{100} \times H\left(\frac{20}{85}, \frac{15}{35}\right) - \frac{65}{100} \times H\left(\frac{5}{65}, \frac{60}{65}\right).$$

$$= 0.7218 - \frac{35}{100}\left(\frac{-20}{35}\log_2\left(\frac{20}{35}\right) - \frac{15}{35}\log_2\left(\frac{15}{35}\right)\right)$$

$$- \frac{65}{100}\left(\frac{-60}{65}\log\left(\frac{60}{65}\right) - \frac{5}{65}\log\left(\frac{5}{65}\right)\right)$$

$$= 0.7218 - 0.344 - 0.2543.$$

Info given $= 0.1229$.

c) Info going would be 0
It is seperated.

∴ No, change would be observed.

d) $A \longrightarrow B \longrightarrow D$.

deafnode $\longrightarrow D$.

output: will wait.