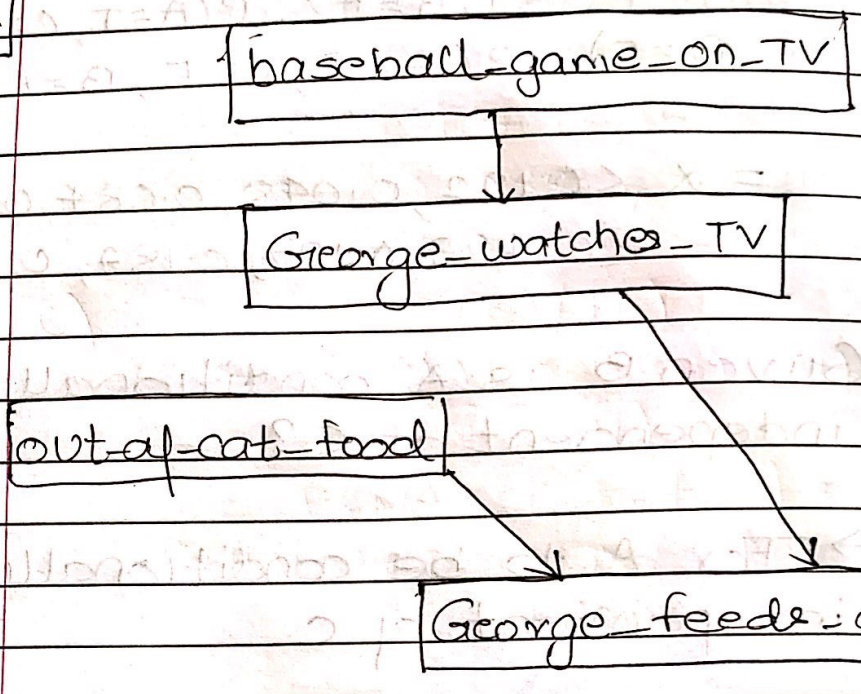
NAME :- PRATIK PRAMOD SINGH
NET ID :- PPS0417
UTA ID :- 1001670417.

Page No.
Date

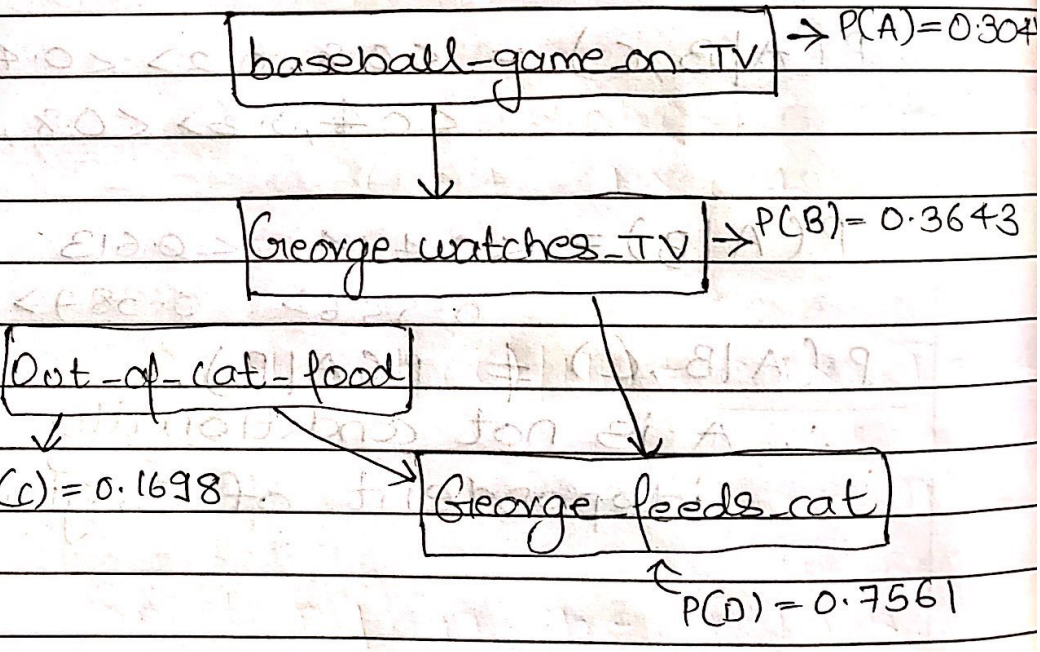
AI ASSIGNMENT - 9.

Q.1.



⇒ Bayesian Network Design:-

Q.2.



A = P(T) of B
T = 0.9279
F = 0.1181

Page No.
Date

Now,
T = George watches TV
F = George out cat-food
P(A) = Probability of baseball game on TV
P(B) = Probability of out of cat food
P(G1) = Probability of George feeds cat
P(T) = Probability of George watches TV

George watches TV	George out cat food	P(G1)	P(G1)'
T	T	0.04166	0.9583
T	F	0.7064	0.2935
F	T	0.3157	6.6842
F	F	0.9587	0.0412

Note:- Excel sheet: Task 2.xlsx.

Q.3] a) Markovian Blanket of Node L.
parent of L : G1
children of L : P, Q
Other parents of children P & Q are :- P = K & L
Q = L & M

b) P(H,C) = P(H,C,M) + P(H,C,~M)
= [P(C) + P(H/C) + P(M|H)] + [P(C) + P(H/C) + P(~M|H)]

$= [0.6 \times 0.6 \times 0.1] + [0.6 \times 0.6 \times (0.1 + 0.6)]$

$= 0.036 + 0.252$

$= 0.288$ //

c) $P(M, \text{not}(C) / H)$

$= \frac{P(M), \text{not}(C), H)}{P(H)}$

$= \frac{P(M) * P(\text{not}(C)) * P(H)}{P(H,M,C) + P(H,M,\text{not}(C)) + P(H,\text{not}(M),C) + P(H,\text{not}(M),\text{not}(C))}$

$= \frac{P(M/H) \cdot P(H/\text{not}(C)) \cdot P(\text{not}(C))}{P(H/C) \cdot P(C) + P(H/\text{not}(C)) \cdot P(\text{not}(C))}$

$= 0.6 \times 0.6 + 0.1 \times (1 - 0.6)$

$= 0.36 + 0.1 \times 0.4$

$= 0.40$ //

$\therefore P(M, \text{not}(C) / H) = \frac{0.1 \times 0.1 \times 0.4}{0.4}$

$= 0.01$ //

Q4] bnet.py in the folder.

Page No.
Date

Q.5] a) Initial Entropy

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

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

$$= -0.8\log_2(0.8) - 0.2\log_2(0.2)$$

$$= -0.8(-0.321) - 0.2(-2.321)$$

$$= 0.2575 + 0.4643$$

$$H(A) = 0.7218 \text{ //}$$

b) Information Gain:

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

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

$$\left. - \frac{65}{100} \times \left(\frac{-60}{65}\log_2\left(\frac{60}{65}\right) - \frac{5}{65}\log_2\left(\frac{5}{65}\right)\right)\right)$$

$$= 0.7218 - (0.35 \times (-0.571 \times -0.80735)$$

$$- (0.4285 \times -1.2239) - 0.65$$

$$(-0.923 \times (-0.1154) - 0.0769 \times (-3.70)))$$

$$= 0.7218 - 0.344 - 0.2543$$

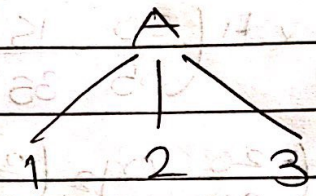
$$= 0.1229$$

Info Gain = 0.1229 //

c) Info gain would be 0, as it's repeated.
Hence, there was no change observed.

d) leaf node : F
A → C → F
O/P : Will wait.

Q.6. With Root node = 'A'.



class A=1	B	C	class A=2	B	C	Class A=3	B	C
x	2	1	x	1	2	x	2	2
x	3	3	y	1	1	y	1	1
x	2	2	y	1	2	y	3	1
			y	1	1			

Total 3 Total 4 Total 3

x = 3 x = 1 x = 1

y = 0 y = 3 y = 2

1. $H(E) = 1$ 2. $H(E_1) = 0$ 3. $H(E_2) = -\frac{1}{4}\log_2(1/4)$

$= 0.5 + 0.3112$

$= 0.81127$

4. $H(\epsilon_3) = -1/3 \log_2(1/3) - 2/3 \log_2(2/3)$

$= -0.5278 - 0.3956$

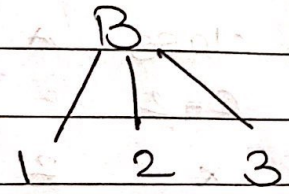
$= 0.9234$

Info gain of $A = H(\epsilon) - 4/10 H(\epsilon_2) - 3/10 H(\epsilon_3)$

$= 1 - 0.4 \times 0.811 - 0.3 \times 0.9234$

$H(A) = 0.3985$ //

• With Root node = 'B'



Class C		Class A C		Class A C	
B = 1		B = 2		B = 3	
x	3 2	x	1 1	x	1 3
y	2 1	x	3 2	y	3 1
y	3 1	x	1 2		
y	2 1	y	2 2		
Total = 4		Total = 4		Total = 2	
x = 1		x = 3		x = 1	
y = 3		y = 1.		y = 1.	

1. $H(\epsilon) = 1$
2. $H(\epsilon_1) = -1/4 \log_2(1/4) - 3/4 \log_2(3/4) = 0.81127$
3. $H(\epsilon_2) = -3/4 \log_2(3/4) - 1/4 \log_2(1/4) = 0.81127$
4. $H(\epsilon_3) = 1$

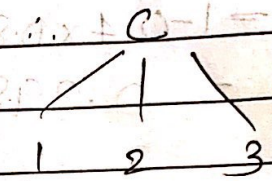
Page No.
Date

∴ Info gain = 1 − 4/10 (H(E1) − 4/10 H(E2) − 2/10 · H(E3)

= 1 − 0.4 × 0.811 − 0.4 × 0.8 − 0.2 ...

H(B) = 0.1812...

• With Root Node 'C'



class	A	B	class	A	B	class	A	B
C=1			C=2			C=3		
x	1	2	x	2	1	x	1	3
y	2	1	x	3	2			
y	3	1	x	1	2			
y	2	3	y	2	2			
y	2	1						
Total = 5			Total = 4			Total = 1		
x = 1			x = 3			x = 1		
y = 4			y = 1					

H(E) = 1

H(E1) = (−1/5 log2(1/5)) − 4/5 (log2(4/5))

= 0.2 log2(0.2) − 0.8 log2(0.8)

= 0.4643 + 0.2575

H(E1) = 0.7218...

H(E2) = −3/4 log2(3/4) − (1/4) log2(1/4)

= 0.75 log2(0.25) − 0.25 log2(0.25)

H(E2) = 0.811...

$H(E_3) = 0.$

$\therefore$ Info Gain $= H(E) - 5/10\, H(E_1) - 4/10 \cdot H(E_2) - 1/10\, H(E_3)$

$= 1 - 0.3609 - 0.3245$

$H(C) = 0.3146.$

Since, Information Gain of root node A is greater than others i.e. $0.3985 > H(C) = 0.3146 > H(B) = 0.1512.$

So, we choose Attribute 'A'. //