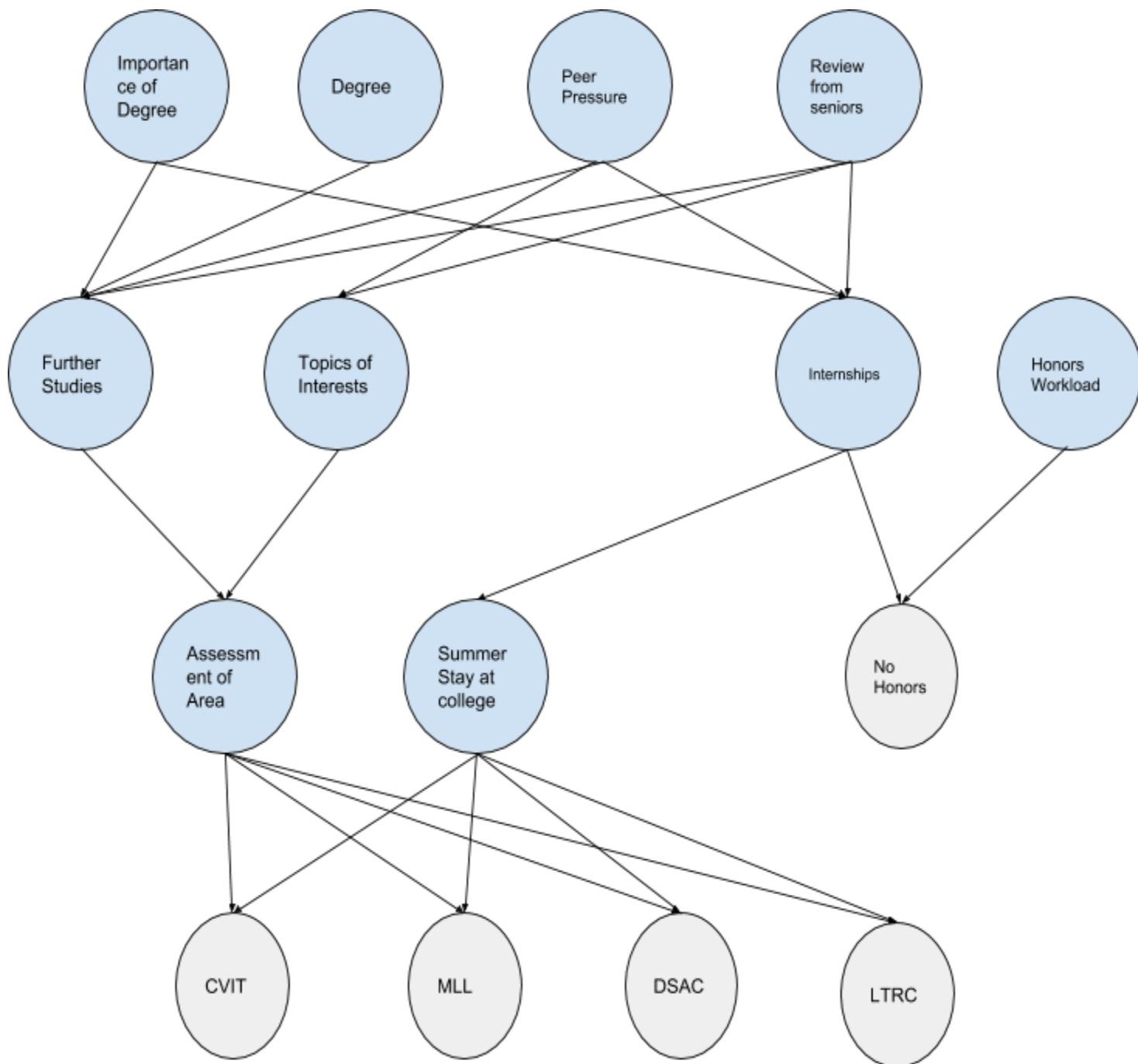


Assignment - 3 Bayesian Network

Vani Sancheti 20171179

Ishita Vohra 20171054



Key:

HD	Honours Degree	Important, Not Important
Deg	Degree	B.Tech , B.Tech+M.S
Rev	Review From Seniors with respect to honours	Good, Average
Pp	Peer Pressure with respect to doing research	Affected, No affected
Pf	Plans for further studies	Yes, No
In	Plans for Internships	Yes, No
Toi	Interest in Topics	Machine Learning, Data Mining, NLP, Computer Vision
Sum	Stay in summers	0 month, 1 month, 2 month
Hwl	Honour's Work Load	High, Medium, Low
ASA	Assessment of Area {Future development in that area}	Excellent, Good, Average
Hon	Honours Decision	Yes, No
CVIT	CVIT	Yes, No
MLL	MLL	Yes, No
DSAC	DSAC	Yes, No
LTRC	LTRC	Yes, No

CONDITIONAL PROBABILITY TABLES:

Honours Degree:

HD	Important	Not Important
P(HD)	0.6	0.4

Branch:

Deg	B.Tech.	BTech+MS
P(Deg)	0.45	0.55

Peer Pressure:

Pp	Affected	Not Affected
P(Pp)	0.7	0.3

Review from Seniors:

Rev	Good	Average
P(Rev)	0.8	0.2

Honours Workload:

Hwl	High	Medium	Low
P(Hwl)	0.5	0.3	0.2

Further Studies:

HD	Deg	Pp	Rev	Pf.No	Pf.Yes
Important	BTech	Affected	Good	0.2	0.8
Not Important	BTech	Affected	Good	0.3	0.7
Important	BTech+MS	Affected	Good	0.1	0.9
Not Important	BTech+MS	Affected	Good	0.5	0.5
Important	BTech	Not Affected	Good	0.05	0.95
Not Important	BTech	Not Affected	Good	0.6	0.4
Important	BTech+MS	Not Affected	Good	0.2	0.8
Not Important	BTech+MS	Not Affected	Good	0.7	0.3
Important	BTech	Affected	Average	0.4	0.6
Not Important	BTech	Affected	Average	0.3	0.7
Important	BTech+MS	Affected	Average	0.3	0.7
Not Important	BTech+MS	Affected	Average	0.4	0.6
Important	BTech	Not Affected	Average	0.5	0.5
Not Important	BTech	Not Affected	Average	0.7	0.3
Important	BTech+MS	Not Affected	Average	0.35	0.65
Not Important	BTech+MS	Not Affected	Average	0.45	0.55

Topics of Interest:

Pp	Rev	MLL	Data Mining	NLP	Computer Vision
<i>Affected</i>	<i>Good</i>	0.35	0.1	0.2	0.35
<i>Affected</i>	<i>Average</i>	0.25	0.2	0.3	0.25
<i>Not Affected</i>	<i>Good</i>	0.3	0.15	0.25	0.3
<i>Not Affected</i>	<i>Average</i>	0.25	0.25	0.25	0.25

Internships:

Rev	Pp	HD	In.Yes	In.No
<i>Good</i>	<i>Affected</i>	<i>Important</i>	0.4	0.6
<i>Average</i>	<i>Affected</i>	<i>Not Important</i>	0.8	0.2
<i>Good</i>	<i>Affected</i>	<i>Not Important</i>	0.6	0.4
<i>Average</i>	<i>Affected</i>	<i>Important</i>	0.45	0.55
<i>Good</i>	<i>Not Affected</i>	<i>Important</i>	0.3	0.7
<i>Average</i>	<i>Not Affected</i>	<i>Not Important</i>	0.9	0.1
<i>Good</i>	<i>Not Affected</i>	<i>Not Important</i>	0.8	0.2
<i>Average</i>	<i>Not Affected</i>	<i>Important</i>	0.4	0.6

Assessment of Areas

<i>Pf</i>	<i>Toi</i>	<i>Excellent</i>	<i>Good</i>	<i>Average</i>
Yes	<i>Machine Learning</i>	0.55	0.3	0.15
Yes	<i>Data Mining</i>	0.35	0.35	0.3
Yes	<i>NLP</i>	0.5	0.3	0.2
Yes	<i>Computer Vision</i>	0.52	0.34	0.14
No	<i>Machine Learning</i>	0.5	0.27	0.23
No	<i>Data Mining</i>	0.33	0.32	0.35
No	<i>NLP</i>	0.45	0.25	0.3
No	<i>Computer Vision</i>	0.5	0.3	0.2

Summer Stay at College

<i>In</i>	<i>0 Month</i>	<i>1 Month</i>	<i>2 Month</i>
Yes	0.7	0.25	0.05
No	0.3	0.3	0.4

No Honours

<i>In</i>	<i>Hwl</i>	<i>Yes</i>	<i>No</i>
Yes	<i>High</i>	0.65	0.35
Yes	<i>Medium</i>	0.4	0.6
Yes	<i>Low</i>	0.3	0.7
No	<i>High</i>	0.55	0.45
No	<i>Medium</i>	0.4	0.6
No	<i>Low</i>	0.2	0.8

CVIT

ASA	Sum	P(Yes)	P(No)
<i>Excellent</i>	0 Months	0	1
<i>Good</i>	0 Months	0	1
<i>Average</i>	0 Months	0	1
<i>Excellent</i>	1 Months	0.3	0.7
<i>Good</i>	1 Months	0.4	0.6
<i>Average</i>	1 Months	0.2	0.8
<i>Excellent</i>	2 Months	0.8	0.2
<i>Good</i>	2 Months	0.75	0.25
<i>Average</i>	2 Months	0.3	0.7

MLL

ASA	Sum	P(Yes)	P(No)
<i>Excellent</i>	0 Months	0.55	0.45
<i>Good</i>	0 Months	0.35	0.75
<i>Average</i>	0 Months	0.3	0.7
<i>Excellent</i>	1 Months	0.65	0.35
<i>Good</i>	1 Months	0.45	0.55
<i>Average</i>	1 Months	0.4	0.6
<i>Excellent</i>	2 Months	0.8	0.2
<i>Good</i>	2 Months	0.75	0.25
<i>Average</i>	2 Months	0.7	0.3

DSAC

ASA	Sum	P(Yes)	P(No)
<i>Excellent</i>	0 Months	0.8	0.2
<i>Good</i>	0 Months	0.6	0.4
<i>Average</i>	0 Months	0.4	0.6
<i>Excellent</i>	1 Months	0.9	0.1
<i>Good</i>	1 Months	0.8	0.2
<i>Average</i>	1 Months	0.65	0.35
<i>Excellent</i>	2 Months	0.95	0.05
<i>Good</i>	2 Months	0.85	0.15
<i>Average</i>	2 Months	0.7	0.3

LTRC

ASA	Sum	P(Yes)	P(No)
<i>Excellent</i>	0 Months	0.85	0.15
<i>Good</i>	0 Months	0.65	0.35
<i>Average</i>	0 Months	0.55	0.45
<i>Excellent</i>	1 Months	0.9	0.1
<i>Good</i>	1 Months	0.75	0.25
<i>Average</i>	1 Months	0.6	0.4
<i>Excellent</i>	2 Months	0.95	0.05
<i>Good</i>	2 Months	0.8	0.2
<i>Average</i>	2 Months	0.7	0.3

Justification

1. For Topics of Interest if a person is Affected by peer pressure then the probability of him choosing major growing areas are more. This has been taken into consideration.
2. For Summer Stay, as it is dependent on Internship for a student it is more likely that he may do the internship for all 3 months.
3. “No honours” option is taken only for B.Tech students as dual degree students have to choose their labs for research.
4. If a student is a dual degree then his/her chances for further studies would be higher.
5. For Assessment of areas, major growing areas are considered excellent. It highly depends on the scope that the area has in the future.
6. For different labs, the allowance for summer stay is taken into consideration while deciding the probabilities for each lab.(eg: In case of CVIT I will stay in summer so the probability of selecting CVIT if I am going home will be zero).
7. The probability for choosing a particular lab is more for the case when I am staying for summer with respect to the case when I am not staying for summer or staying respectively for less number of months because most of the lab's want us to stay in summer.
8. Peer pressure is taken into consideration with respect to peers taking honours.
9. Review of seniors is taken into consideration with respect to honours.
10. Honours Workload is considered to be more because you have to do a project every semester. But it also depends on each lab whether the workload is high, medium, low. But most of the time load is considered to be high.
11. Assessment of area is with respect to a student so it depends on their topics of interest, their plans for further studies. It also involves the future scope of that particular area

Query

Query taken : $P(X | p(X), p(p(X)))$

$P(\text{MLL} = \text{Yes} | \text{Assessment of area} = \text{Excellent}, \text{Topics of Interest} = \text{Machine Learning}) = P(\text{MLL} = \text{Yes} | \text{Assessment of Area} = \text{Excellent})$

$$P(\text{MLL} = \text{Yes} | \text{Assessment of Area} = \text{Excellent}) = 0.55x(P(\text{sum} = 0 \text{ Months})) + 0.65x(P(\text{sum} = 1 \text{ Months})) + 0.8x(P(\text{sum} = 2 \text{ Months}))$$

$$P(\text{sum} = 0 \text{ Months}) = 0.7x(P(\text{In} = \text{Yes})) + 0.3x(P(\text{In} = \text{No}))$$

$$P(\text{sum} = 1 \text{ Months}) = 0.25x(P(\text{In} = \text{Yes})) + 0.3x(P(\text{In} = \text{No}))$$

$$P(\text{sum} = 2 \text{ Months}) = 0.05x(P(\text{In} = \text{Yes})) + 0.4x(P(\text{In} = \text{No}))$$

$$\begin{aligned} P(\text{In} = \text{Yes}) = & 0.4x(P(\text{Rev} = \text{Good}))x(P(\text{Pp} = \text{Affected}))x(P(\text{HD} = \text{Important})) + \\ & 0.8x(P(\text{Rev} = \text{Average}))x(P(\text{Pp} = \text{Affected}))x(P(\text{HD} = \text{Not Important})) + \\ & 0.6x(P(\text{Rev} = \text{Good}))x(P(\text{Pp} = \text{Affected}))x(P(\text{HD} = \text{Not Important})) + \\ & 0.45x(P(\text{Rev} = \text{Average}))x(P(\text{Pp} = \text{Affected}))x(P(\text{HD} = \text{Important})) + \\ & 0.3x(P(\text{Rev} = \text{Good}))x(P(\text{Pp} = \text{Not Affected}))x(P(\text{HD} = \text{Important})) + \\ & 0.9x(P(\text{Rev} = \text{Average}))x(P(\text{Pp} = \text{Not Affected}))x(P(\text{HD} = \text{Not Important})) + \\ & 0.8x(P(\text{Rev} = \text{Good}))x(P(\text{Pp} = \text{Not Affected}))x(P(\text{HD} = \text{Not Important})) + \\ & 0.4x(P(\text{Rev} = \text{Average}))x(P(\text{Pp} = \text{Not Affected}))x(P(\text{HD} = \text{Important})) \end{aligned}$$

$$\begin{aligned} = & 0.4x0.8x0.7x0.6 + \\ & 0.8x0.2x0.7x0.4 + \\ & 0.6x0.8x0.7x0.4 + \\ & 0.45x0.2x0.7x0.6 + \\ & 0.3x0.8x0.3x0.6 + \\ & 0.9x0.2x0.3x0.4 + \\ & 0.8x0.8x0.3x0.4 + \end{aligned}$$

$$0.4 \times 0.2 \times 0.3 \times 0.6$$

$$= 0.5074$$

$$P(\text{In} = \text{No}) = 0.6 \times (P(\text{Rev} = \text{Good})) \times (P(\text{Pp} = \text{Affected})) \times (P(\text{HD} = \text{Important})) +$$

$$0.2 \times (P(\text{Rev} = \text{Average})) \times (P(\text{Pp} = \text{Affected})) \times (P(\text{HD} = \text{Not Important})) +$$

$$0.4 \times (P(\text{Rev} = \text{Good})) \times (P(\text{Pp} = \text{Affected})) \times (P(\text{HD} = \text{Not Important})) +$$

$$0.55 \times (P(\text{Rev} = \text{Average})) \times (P(\text{Pp} = \text{Affected})) \times (P(\text{HD} = \text{Important})) +$$

$$0.7 \times (P(\text{Rev} = \text{Good})) \times (P(\text{Pp} = \text{Not Affected})) \times (P(\text{HD} = \text{Important})) +$$

$$0.1 \times (P(\text{Rev} = \text{Average})) \times (P(\text{Pp} = \text{Not Affected})) \times (P(\text{HD} = \text{Not Important})) +$$

$$0.2 \times (P(\text{Rev} = \text{Good})) \times (P(\text{Pp} = \text{Not Affected})) \times (P(\text{HD} = \text{Not Important})) +$$

$$0.6 \times (P(\text{Rev} = \text{Average})) \times (P(\text{Pp} = \text{Not Affected})) \times (P(\text{HD} = \text{Important}))$$

$$= 0.6 \times 0.8 \times 0.7 \times 0.6 +$$

$$0.2 \times 0.2 \times 0.7 \times 0.4 +$$

$$0.4 \times 0.8 \times 0.7 \times 0.4 +$$

$$0.55 \times 0.2 \times 0.7 \times 0.6 +$$

$$0.7 \times 0.8 \times 0.3 \times 0.6 +$$

$$0.1 \times 0.2 \times 0.3 \times 0.4 +$$

$$0.2 \times 0.8 \times 0.3 \times 0.4 +$$

$$0.6 \times 0.2 \times 0.3 \times 0.6$$

$$= 0.2016 + 0.0112 + 0.0896 + 0.0462 + 0.1008 + 0.0024 + 0.0192 + 0.0216$$

$$= 0.4926$$

$$P(\text{sum} = 0 \text{ Months}) = 0.7 \times 0.5074 + 0.3 \times 0.4926$$

$$= 0.50296$$

$$P(\text{sum} = 1 \text{ Months}) = 0.25 \times 0.5074 + 0.3 \times 0.4926$$

$$= 0.27463$$

$$P(\text{sum} = 2 \text{ Months}) = 0.05 \times 0.5074 + 0.4 \times 0.4926$$

$$= 0.22241$$

$$P(\text{MLL}=\text{Yes}|\text{Assessment of Area}=\text{Excellent})=0.55 \times 0.50296 + 0.65 \times 0.27463 + 0.8 \times 0.22241 = 0.6330655$$

$$\begin{aligned} P(\text{MLL} = \text{Yes} \mid \text{Assessment of area} = \text{Excellent}, \text{Topics of Interest} = \text{Machine Learning}) \\ = P(\text{MLL} = \text{Yes} \mid \text{Assessment of Area} = \text{Excellent}) \\ = 0.6330655 \end{aligned}$$

Since MLL , Summer Stay and Internship is independent of Topic of Interest, therefore, the above probabilities have been reduced to its equivalent probabilities.

Formula Used:

$$\begin{aligned} & P(\text{Mll} = \text{Yes} , \text{ASA} = \text{Excellent} , \text{Topics of Interest} = \text{Machine Learning}) / \\ & \quad P(\text{ASA} = \text{Excellent} , \text{Topics of Interest} = \text{Machine Learning}) = \\ & = P(\text{MLL} \mid \text{ASA} = \text{Excellent}) * P(\text{ASA} = \text{Excellent} \mid \text{Topics of Interest} = \text{Machine Learning}) * \\ & P(\text{Topics Of Interest} = \text{Machine Learning}) / \\ & \quad P(\text{ASA} = \text{Excellent} \mid \text{Topics of Interest} = \text{Machine Learning}) * P(\text{Topics of Interest} = \text{Machine Learning}) \\ & = P(\text{MLL} \mid \text{ASA} = \text{Excellent}) \end{aligned}$$