ENDGAME(Team-8)
Nonidh Singh(20171203)
Sarthak Singhal(20171091)

BAYESIAN NETWORKS

OVERVIEW (Project 14)

You are the coordinator of the UG2k17 batch trip and you want to plan a short trip for around 2-3 days. Planning involves deciding which place to go, what time of the year, how to go etc. Model a Bayes' net that takes into account the factors that affect the planning of a trip, and assign conditional probabilities appropriately.

STRUCTURAL CONSTRAINTS

1. Your Bayes' net should have 8-16 nodes, and so, you may have to rule out a few factors, and take only the top k factors that you feel are the most significant for your topic.
2. There should be a minimum of 3 layers in the net i.e. at least a set of three nodes related to each other and in different layers. A 4/5 layer network is advised. In other words, nodes in such types of nets tend to have in the range of 2-5 parent nodes. You need to take into consideration the fact that your CPT should be neither too small nor too big.
3. If you feel that the number of dependencies is very high, you could prune some of them with an appropriate justification for the same. (However, this does not imply that you can do the same to make an easy net and evade calculations. :))
4. If there are too many rows in a Conditional Probability Table (CPT), try to remove the less significant ones. A CPT with 40 rows is considered big enough, for most of the cases we have provided, and so feel free to prune the rest of it, but do give a justification for the same. Note that your CPTs may be much smaller or slightly larger than this. They should always contain enough information to model reality to a good extent, and abstract out details, which you feel are unnecessary, and make.

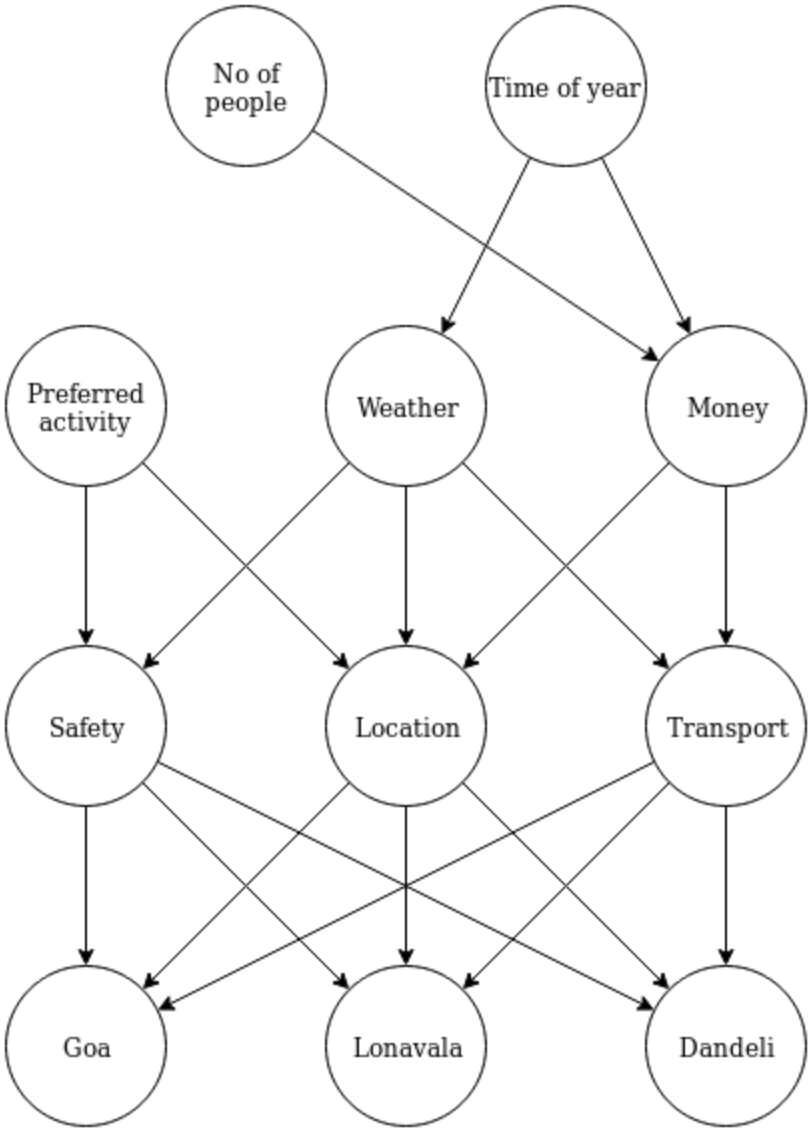
DELIVERABLES

1. A diagram of the Bayes' net
2. Conditional Probability Tables for all dependencies
3. A justification behind the reasons for your choosing different probabilities for different nodes.

4. A solved query of the following type(where p(X) refers to the parent of X):
    a. P(X | p(X), p(p(X)))
    b. P(p(p(X)) | X)
    c. P(p(X) | X, p(p(X)))

PROPOSED SOLUTION

DIAGRAM OF THE BAYES' NET:



KEY:

SYMBOL	VARIABLE	VALUES
NOP	No of people	Less, Medium. High
TOY	Time of year	Summer, Winter, Monsoon
PA	Preferred activity	Camping, Water sports, Trekking, Sun bath, Adventure sports
W	Weather	Sunny, Cool, Other
M	Money	Less, Sufficient
S	Safety	Yes, No
LOC	Location	Beach, Hills, Historical
T	Transport	Bus, Train, Plane
G	Goa	Yes, No
L	Lonavala	Yes, No
D	Dandeli	Yes, No

CPT'S FOR ALL DEPENDENCIES:

No of people:

NOP	Less	Medium	High
P(NOP)	.15	.75	.1

➔ More number of people will opt to go for a trip amidst extreme academic pressure.

➔ Probability of people being very less or very high is low.

.

Time of year

TOY	Summer	Winter	Monsoon
P(TOY)	.34	.33	.33

➔ All three outcomes are equally likely.

Preferred activity

PA	Camping	Water sports	Trekking	Sun bath	Adventure sports
P(PA)	.2	.2	.2	.2	.2

➔ All activities equally likely.

Weather

TOY	Sunny	Cool	Other
Summer	.7	.1	.2
Winter	.1	.7	.2
Monsoon	.2	.3	.5

➔ Weathers vary according to the season.

➔ Refers to the weather at the destination place.

➔ We have a rough estimate of destination weather as all places are closely located with not much variation.

➔ In monsoon season sunny days are less expected.

Money

NOP	TOY	Less
Less	Summer	.8
Less	Winter	.6
Less	Monsoon	.7
Medium	Summer	.4
Medium	Winter	.3
Medium	Monsoon	.35
High	Summer	.2
High	Winter	.05
High	Monsoon	.1

➔ Money represents total cash in hand beforehand after initial expenses.

➔ Assumptions made:

- Cost per person is more in summers so we'll have less cash for other things.
- Cost per person is less in winters so we'll have more cash for other things.
- Cost per person is medium in monsoons so we'll have medium cash for other things.

➔ TOY will have less impact on money as NOP will have.

➔ Clearly if NOP is more money will be high corresponding to any TOY. Implies that if NOP is high, money would be more irrespective of TOY. P(M | NOP=high)>P(M | NOP!=high) for all seasons.

Safety

PA	W	Yes
Camping	Sunny	.8
Camping	Cool	.75
Camping	Other	.7
Water sports	Sunny	.7
Water sports	Cool	.65
Water sports	Other	.6
Trekking	Sunny	.66
Trekking	Cool	.61
Trekking	Other	.55
Sun bath	Sunny	.9
Sun bath	Cool	.9
Sun bath	Other	.85
Adventure sports	Sunny	.5
Adventure sports	Cool	.48
Adventure sports	Other	.4

➔ Other may include rain, snow etc. so probability of being safe would be less.

➔ Cool may imply later part of the day so daylight might reduce so less safe.

➔ Sun bath considered more safe than any other activity.

➔ Adventure sports considered more riskier than other activities.

Location

PA	W	M	Beach	Hills	Historical
Camping	Sunny	Less	.1	.6	.3
Camping	Sunny	Sufficient	.09	.65	.26
Camping	Cool	Less	.1	.55	.35
Camping	Cool	Sufficient	.09	.6	.31
Camping	Other	Less	.1	.6	.3
Camping	Other	Sufficient	.09	.65	.26
Water sports	Sunny	Less	.6	.3	.1
Water sports	Sunny	Sufficient	.59	.32	.09
Water sports	Cool	Less	.61	.29	.1
Water sports	Cool	Sufficient	.6	.31	.09
Water sports	Other	Less	.7	.2	.1
Water sports	Other	Sufficient	.65	.25	.09
Trekking	Sunny	Less	0	.7	.3
Trekking	Sunny	Sufficient	0	.75	.25
Trekking	Cool	Less	0	.6	.4
Trekking	Cool	Sufficient	0	.65	.35
Trekking	Other	Less	0	.5	.5
Trekking	Other	Sufficient	0	.53	.47
Sun bath	Sunny	Less	.8	0	.2
Sun bath	Sunny	Sufficient	.8	0	.2
Sun bath	Cool	Less	.82	0	.18
Sun bath	Cool	Sufficient	.81	0	.19
Sun bath	Other	Less	.6	0	.4

Sun bath	Other	Sufficient	.6	0	.4
Adventure sports	Sunny	Less	.5	.4	.1
Adventure sports	Sunny	Sufficient	.49	.41	.1
Adventure sports	Cool	Less	.45	.45	.1
Adventure sports	Cool	Sufficient	.4	.5	.1
Adventure sports	Other	Less	.7	.2	.1
Adventure sports	Other	Sufficient	.66	.26	.08

- ➔ PA would play major role in deciding the location.
- ➔ Historic place refers to a place famous for many historical monuments. It doesn't mean any other things are not there.
- ➔ If PA is camping then we would not want to visit a place having beaches as not many camping sites would be available. (P(LOC=Beach)<P(LOC=Historical)<P(LOC=Hills))
- ➔ Camping in hilly areas would require more money so increases with money.
- ➔ If we have less money then we can prefer beach a bit more as not much cost is required.
- ➔ If weather is cool, probability of going to hilly areas for camping may decrease a bit.
- ➔ In rainy weather camping at hilly areas may not be advised.
- ➔ If PA is water sports, beach would be preferred followed by hills(hills may have waterfalls).
- ➔ Historic places may be preferred because of low cost.
- ➔ Not much change if weather is sunny or cool.
- ➔ If PA is trekking hills may be preferred followed by historic places.
- ➔ No one would go to place having beaches for trekking.
- ➔ If PA is sunbath beaches would be preferred.
- ➔ No one would go to hilly areas for having sun bath.
- ➔ Money is less significant if activity is sunbath.
- ➔ If PA is adventure sports, beaches and hilly areas would have almost same probability.

Transport

W	M	Bus	Train	Plane
Sunny	Less	.4	.5	.1
Sunny	Sufficient	.1	.6	.3
Cool	Less	.4	.5	.1
Cool	Sufficient	.1	.6	.3

Other	Less	.2	.7	.1
Other	Sufficient	.03	.65	.32

- ➔ If money is high then we'll prefer plane as travel time would be much less.
- ➔ Sunny or cool wouldn't affect transportation as during the travel time the weather would keep on fluctuating between sunny and cool.
- ➔ If there is rain or snow then we might want to travel by train.
- ➔ Goa distance:650km
- ➔ Dandeli 600km
- ➔ Lonavala 600km

Goa

S	LOC	T	Yes
No	Beach	Bus	.5
No	Beach	Train	.51
No	Beach	Plane	.6
No	Hills	Bus	.2
No	Hills	Train	.24
No	Hills	Plane	.3
No	Historical	Bus	.5
No	Historical	Train	.52
No	Historical	Plane	.53
Yes	Beach	Bus	.75
Yes	Beach	Train	.76
Yes	Beach	Plane	.85
Yes	Hills	Bus	.27
Yes	Hills	Train	.31
Yes	Hills	Plane	.37
Yes	Historical	Bus	.5
Yes	Historical	Train	.52

Yes	Historical	Plane	.53

- ➔ If LOC=beach, then high chances of going to Goa.
- ➔ If less safety then we may not prefer a bus. Eg in extreme foggy weather we may prefer a plane. But transport will play only slight role.
- ➔ If safety=less Goa would be preferred more as it is a safe place.
- ➔ Historic places almost independent of safety.

Lonavala

S	LOC	T	Yes
No	Beach	Bus	.4
No	Beach	Train	.41
No	Beach	Plane	.5
No	Hills	Bus	.7
No	Hills	Train	.71
No	Hills	Plane	.72
No	Historical	Bus	.8
No	Historical	Train	.82
No	Historical	Plane	.84
Yes	Beach	Bus	.55
Yes	Beach	Train	.56
Yes	Beach	Plane	.65
Yes	Hills	Bus	.77
Yes	Hills	Train	.79
Yes	Hills	Plane	.82
Yes	Historical	Bus	.81
Yes	Historical	Train	.84
Yes	Historical	Plane	.85

- ➔ If safety increases, then chances of going to lonavala would increase as it is more difficult to handle the batch at lonavala than goa.

➔ Going to beach in Lonavala is less probable than in Goa.

Dandeli

S	LOC	T	Yes
No	Beach	Bus	.5
No	Beach	Train	.52
No	Beach	Plane	.55
No	Hills	Bus	.73
No	Hills	Train	.74
No	Hills	Plane	.75
No	Historical	Bus	.4
No	Historical	Train	.42
No	Historical	Plane	.44
Yes	Beach	Bus	.55
Yes	Beach	Train	.58
Yes	Beach	Plane	.63
Yes	Hills	Bus	.76
Yes	Hills	Train	.77
Yes	Hills	Plane	.79
Yes	Historical	Bus	.45
Yes	Historical	Train	.46
Yes	Historical	Plane	.47

➔ All activities can be enjoyed at Dandeli

SAMPLE QUERY:

Find P' = P(Goa=Yes | LOC=Beach, T=Train, PA=Adventure sports, W=Sunny)

Thus we have to find:

P(Goa, S, LOC, T, PA, W, M, NOP, TOY)

According to net's property:

P' = P(Goa | S, LOC, T) X P(S | PA, W) X P(LOC | PA, W, M) X P(T | W, M) X P(PA) X P(W | TOY) X P(M | NOP, TOY) X P(NOP) X P(TOY)
P' = P(Goa=Yes | S, LOC=Beach, T=Train) X P(S | PA=Adventure sports, W=Sunny) X P(LOC=Beach | PA=Adventure sports, W=Sunny, M) X P(T=Train | W=Sunny, M) X P(PA=Adventure sports) X P(W=Sunny | TOY) X P(M | NOP, TOY) X P(NOP) X P(TOY)
P' = ∑ S ∑ M ∑ NOP ∑ TOY P(Goa=Yes | S, LOC=Beach, T=Train) X P(S | PA=Adventure sports, W=Sunny) X P(LOC=Beach | PA=Adventure sports, W=Sunny, M) X P(T=Train | W=Sunny, M) X P(PA=Adventure sports) X P(W=Sunny | TOY) X P(M | NOP, TOY) X P(NOP) X P(TOY)
P' = .2 X ∑ S ∑ M ∑ NOP ∑ TOY P(Goa=Yes | S, LOC=Beach, T=Train) X P(S | PA=Adventure sports, W=Sunny) X P(LOC=Beach | PA=Adventure sports, W=Sunny, M) X P(T=Train | W=Sunny, M) X P(W=Sunny | TOY) X P(M | NOP, TOY) X P(NOP) X P(TOY)
= .2 X { ∑ S P(Goa=Yes | S, LOC=Beach, T=Train) X P(S | PA=Adventure sports, W=Sunny) } X { ∑ M ∑ NOP ∑ TOY P(LOC=Beach | PA=Adventure sports, W=Sunny, M) X P(T=Train | W=Sunny, M) X P(W=Sunny | TOY) X P(M | NOP, TOY) X P(NOP) X P(TOY) }
= .2 X P1 X P2 where: P1 = { ∑ S P(Goa=Yes | S, LOC=Beach, T=Train) X P(S | PA=Adventure sports, W=Sunny) } P2 = { ∑ M ∑ NOP ∑ TOY P(LOC=Beach | PA=Adventure sports, W=Sunny, M) X P(T=Train | W=Sunny, M) X P(W=Sunny | TOY) X P(M | NOP, TOY) X P(NOP) X P(TOY) }
Now P1 = P(Goa=Yes | S=Yes, LOC=Beach, T=Train) X P(S=Yes | PA=Adventure sports, W=Sunny) +

P(Goa=Yes | S=No, LOC=Beach, T=Train) X P(S=No | PA=Adventure sports, W=Sunny)

P1 = (.76 X .5) + (.51 X .5)
P1 = .635

P' = .2 X .635 X P2 where:

P2 = { ∑ M ∑ NOP ∑ TOY P(LOC=Beach | PA=Adventure sports, W=Sunny, M) X P(T=Train | W=Sunny, M) X P(W=Sunny | TOY) X P(M | NOP, TOY) X P(NOP) X P(TOY) }

P2 =
∑ NOP ∑ TOY P(LOC=Beach | PA=Adventure sports, W=Sunny, M=Less) X P(T=Train | W=Sunny, M=Less) X P(W=Sunny | TOY) X P(M=Less | NOP, TOY) X P(NOP) X P(TOY)
+
∑ NOP ∑ TOY P(LOC=Beach | PA=Adventure sports, W=Sunny, M=Sufficient) X P(T=Train | W=Sunny, M=Sufficient) X P(W=Sunny | TOY) X P(M=Sufficient | NOP, TOY) X P(NOP) X P(TOY)

P2 =
P(LOC=Beach | PA=Adventure sports, W=Sunny, M=Less) X P(T=Train | W=Sunny, M=Less) X ∑ NOP ∑ TOY P(W=Sunny | TOY) X P(M=Less | NOP, TOY) X P(NOP) X P(TOY)
+
P(LOC=Beach | PA=Adventure sports, W=Sunny, M=Sufficient) X P(T=Train | W=Sunny, M=Sufficient) X ∑ NOP ∑ TOY P(W=Sunny | TOY) X P(M=Sufficient | NOP, TOY) X P(NOP) X P(TOY)

P2 =
.5 X .5 X ∑ NOP ∑ TOY P(W=Sunny | TOY) X P(M=Less | NOP, TOY) X P(NOP) X P(TOY)
+
.49 X .6 X ∑ NOP ∑ TOY P(W=Sunny | TOY) X P(M=Sufficient | NOP, TOY) X P(NOP) X P(TOY)

P2 =
.25 X ∑ NOP ∑ TOY P(W=Sunny | TOY) X P(M=Less | NOP, TOY) X P(NOP) X P(TOY)
+
.294 X ∑ NOP ∑ TOY P(W=Sunny | TOY) X P(M=Sufficient | NOP, TOY) X P(NOP) X P(TOY)

Now P2 = .25 X P3 + .294 X P4 where
P3 = ∑ NOP ∑ TOY P(W=Sunny | TOY) X P(M=Less | NOP, TOY) X P(NOP) X P(TOY)
P4 = ∑ NOP ∑ TOY P(W=Sunny | TOY) X P(M=Sufficient | NOP, TOY) X P(NOP) X P(TOY)

P3 =
∑ TOY P(W=Sunny | TOY) X P(M=Less | NOP=Less, TOY) X P(NOP=Less) X P(TOY)
+
∑ TOY P(W=Sunny | TOY) X P(M=Less | NOP=Medium, TOY) X P(NOP=Medium) X P(TOY)
+
∑ TOY P(W=Sunny | TOY) X P(M=Less | NOP=High, TOY) X P(NOP=High) X P(TOY)

P3 =
P(W=Sunny | TOY=Summer) X P(M=Less | NOP=Less, TOY=Summer) X P(NOP=Less) X

P(TOY=Summer)
+
P(W=Sunny | TOY=Summer) X P(M=Less | NOP=Medium, TOY=Summer) X P(NOP=Medium) X P(TOY=Summer)
+
P(W=Sunny | TOY=Summer) X P(M=Less | NOP=High, TOY=Summer) X P(NOP=High) X P(TOY=Summer)

+

P(W=Sunny | TOY=Winter) X P(M=Less | NOP=Less, TOY=Winter) X P(NOP=Less) X P(TOY=Winter)
+
P(W=Sunny | TOY=Winter) X P(M=Less | NOP=Medium, TOY=Winter) X P(NOP=Medium) X P(TOY=Winter)
+
P(W=Sunny | TOY=Winter) X P(M=Less | NOP=High, TOY=Winter) X P(NOP=High) X P(TOY=Winter)

+

P(W=Sunny | TOY=Monsoon) X P(M=Less | NOP=Less, TOY=Monsoon) X P(NOP=Less) X P(TOY=Monsoon)
+
P(W=Sunny | TOY=Monsoon) X P(M=Less | NOP=Medium, TOY=Monsoon) X P(NOP=Medium) X P(TOY=Monsoon)
+
P(W=Sunny | TOY=Monsoon) X P(M=Less | NOP=High, TOY=Monsoon) X P(NOP=High) X P(TOY=Monsoon)

P3 =
.7 X .8 X .15 X .34
+
.7 X .4 X .75 X .34
+
.7 X .2 X .1 X .34
+
.1 X .6 X .15 X .33
+
.1 X .3 X .75 X .33
+
.1 X .05 X .1 X .33
+
.2 X .7 X .15 X .33
+
.2 X .35 X .75 X .33
+
.2 X .1 X .1 X .33

P3 = 0.140195

Now P2 = .25 X .140195 + .294 X P4 where
P4 = ∑ NOP ∑ TOY P(W=Sunny | TOY) X P(M=Sufficient | NOP, TOY) X P(NOP) X P(TOY)

P4 =
∑ TOY P(W=Sunny | TOY) X P(M=Sufficient | NOP=Less, TOY) X P(NOP=Less) X P(TOY)
+
∑ TOY P(W=Sunny | TOY) X P(M=Sufficient | NOP=Medium, TOY) X P(NOP=Medium) X P(TOY)
+
∑ TOY P(W=Sunny | TOY) X P(M=Sufficient | NOP=High, TOY) X P(NOP=High) X P(TOY)

P4 =
P(W=Sunny | TOY=Summer) X P(M=Sufficient | NOP=Less, TOY=Summer) X P(NOP=Less) X P(TOY=Summer)
+
P(W=Sunny | TOY=Summer) X P(M=Sufficient | NOP=Medium, TOY=Summer) X P(NOP=Medium) X P(TOY=Summer)
+
P(W=Sunny | TOY=Summer) X P(M=Sufficient | NOP=High, TOY=Summer) X P(NOP=High) X P(TOY=Summer)

+

P(W=Sunny | TOY=Winter) X P(M=Sufficient | NOP=Less, TOY=Winter) X P(NOP=Less) X P(TOY=Winter)
+
P(W=Sunny | TOY=Winter) X P(M=Sufficient | NOP=Medium, TOY=Winter) X P(NOP=Medium) X P(TOY=Winter)
+
P(W=Sunny | TOY=Winter) X P(M=Sufficient | NOP=High, TOY=Winter) X P(NOP=High) X P(TOY=Winter)

+

P(W=Sunny | TOY=Monsoon) X P(M=Sufficient | NOP=Less, TOY=Monsoon) X P(NOP=Less) X P(TOY=Monsoon)
+
P(W=Sunny | TOY=Monsoon) X P(M=Sufficient | NOP=Medium, TOY=Monsoon) X P(NOP=Medium) X P(TOY=Monsoon)
+
P(W=Sunny | TOY=Monsoon) X P(M=Sufficient | NOP=High, TOY=Monsoon) X P(NOP=High) X P(TOY=Monsoon)

P4 =
.7 X .2 X .15 X .34
+
.7 X .6 X .75 X .34
+

.7 X .8 X .1 X .34 + .1 X .4 X .15 X .33 + .1 X .7 X .75 X .33 + .1 X .95 X .1 X .33 + .2 X .3 X .15 X .33 + .2 X .65 X .75 X .33 + .2 X .9 X .1 X .33
P4 = .196805
Now P2 = .25 X .140195 + .294 X .196805
P2 = 0.09290942
Now P’ = .2 X .635 X .09290942
Hence P’ = 0.0117995 = **0.018 approx**