COMP307 Assignment 3

Part 1

X	P(X)
0	0.3
1	0.7

Y	X	P(Y|X)
0	0	0.3
1	0	0.7
0	1	0.8
1	1	0.2

Z	Y	P(Z|Y)	P(Z|Y)
0	0	0.6	0.3
1	0	0.4	0.2
0	1	0.8	0.4
1	1	0.2	0.1

1.

X	Y	P(X)	P(Y|X)	P(X,Y)
0	0	0.3	0.3	0.09
1	0	0.7	0.8	0.56
0	1	0.3	0.7	0.21
1	1	0.7	0.2	0.14

Used product rule: P(X,Y) = P(X) * P(Y|X)

2.

X	Y	Z	P(X,Y)	P(Z|Y)	P(X,Y, Z)
0	0	0	0.09	0.6	0.054
1	0	0	0.56	0.6	0.336
0	1	0	0.21	0.4	0.168
1	1	0	0.14	0.4	0.112
0	0	1	0.09	0.8	0.036
1	0	1	0.56	0.8	0.224
0	1	1	0.21	0.2	0.042
1	1	1	0.14	0.2	0.028

Rules Used:

- Product rule: P(X,Y,Z) = P(Z|X,Y)*P(X,Y)
- Z is independent from P(X|Y) so P(Z|X,Y) = P(Z|Y)
- Conditionally independance so P(Z,X,Y) = P(Z|Y)*P(X,Y)

3.

i)

P(Z=0)	P(X=0,Z=0)
0.67	0.222

Using numbers from section 2.
P(Z=0) = P(X=0,Y=0,Z=0) + P(X=1,Y=0,Z=0) + P(X=0,Y=1,Z=0) + P(X=1,Y=1,Z=0) = 0.67
P(X=0,Z=0) = P(X=0,Y=0,Z=0) + P(X=0,Y=1,Z=0) = 0.222

ii)

Are X and Z independent of each other?
if X and Z are independent P(X,Z) = P(X)*P(Z)
P(X=0,Z=0) = P(X=0)*P(Z=0)
0.3*0.67 = 0.201
P(X=0,Z=0) = 0.222

0.201 != 0.222
Therefore X and Z are not independent of each other

4.

i)

Steps	**Rules**
P(X=1, Y=0|Z=1)	
=P(Z|X,Y)*P(X,Y)/P(Z)	Bayes rule
P(Z|X,Y) = P(Z|Y)	Conditionally independent
P(X,Y|Z) = P(Z|Y)*P(X,Y)/P(Z)	
P(X=1,Y=0|Z=1) = P(Z=1|Y=0)*P(X=1,Y=0)/P(Z=1)	
P(Z=1) = 1-P(Z=0)=1-0.67=0.33	Normalisation
P(X=1,Y=0|Z=1) = P(Z=1|Y=0)*P(X=1,Y=0)/P(Z=1) = 0.4*0.56/0.33 = 0.679	

ii)

P(X=0|Y=0, Z=0)
P(X=0|Y=0, Z=0) = P(X=0,Y=0, Z=0)/(P(X=0,Y=0,Z=0) + P(X=1,Y=0,Z=0))
= 0.054/(0.054+0.336)
= 0.161

Part 2

1.

Feature	value	P(value|not-spam)	P(value|spam)
1	0	0.6423841059602649	0.33962264150943394
	1	0.3576158940397351	0.660377358490566
2	0	0.423841059602649	0.41509433962264153
	1	0.5761589403973509	0.5849056603773585
3	0	0.6556291390728477	0.5471698113207547
	1	0.3443708609271523	0.4528301886792453
4	0	0.6026490066225165	0.39622641509433965
	1	0.3973509933774834	0.6037735849056604
5	0	0.6622516556291391	0.5094339622641509
	1	0.33774834437086093	0.49056603773584906
6	0	0.5298013245033113	0.6415094339622641
	1	0.47019867549668876	0.3584905660377358
7	0	0.4966887417218543	0.22641509433962265
	1	0.5033112582781457	0.7735849056603774
8	0	0.6490066225165563	0.24528301886792453
	1	0.3509933774834437	0.7547169811320755
9	0	0.7549668874172185	0.660377358490566
	1	0.24503311258278146	0.33962264150943394
10	0	0.7086092715231788	0.33962264150943394
	1	0.2913907284768212	0.660377358490566
11	0	0.41721854304635764	0.33962264150943394

	1	0.5827814569536424	0.660377358490566
12	0	0.6622516556291391	0.22641509433962265
	1	0.33774834437086093	0.7735849056603774

2.

Instance 1:
Input: [1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0]
P(S|D): 3.6646132342666684e-06
P(!S|D): 0.00045599287219218876
Class: Not Spam

Instance 2:
Input: [0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1]
P(S|D): 5.7900134716171085e-05
P(!S|D): 4.1822057115519664e-05
Class: Spam

Instance 3:
Input: [1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1]
P(S|D): 0.0001877146421019266
P(!S|D): 0.00012844947310385982
Class: Spam

Instance 4:
Input: [0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0]
P(S|D): 6.153970500877269e-06
P(!S|D): 0.0005951777569292127
Class: Not Spam

Instance 5:
Input: [1, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1]
P(S|D): 6.19949306356931e-05
P(!S|D): 9.227590129769324e-05
Class: Not Spam

Instance 6:
Input: [1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 1]
P(S|D): 5.969882209363039e-05
P(!S|D): 4.626123885304098e-05

Class: Spam

Instance 7:
Input: [0, 0, 0, 0, 1, 1, 0, 1, 0, 0, 0, 0]
P(S|D): 4.115382745890395e-06
P(!S|D): 0.00032587162314003077
Class: Not Spam

Instance 8:
Input: [0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 1, 1]
P(S|D): 6.51874586963203e-05
P(!S|D): 0.00038986164799395464
Class: Not Spam

Instance 9:
Input: [1, 1, 1, 1, 1, 0, 1, 0, 0, 1, 0, 1]
P(S|D): 0.00018771464210192657
P(!S|D): 3.781456451188396e-05
Class: Spam

Instance 10:
Input: [1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0]
P(S|D): 2.2768307076804742e-05
P(!S|D): 0.0006754291417850848
Class: Not Spam

3.

Naive Bayes assumes all features are independent such that P(X,Y) = P(X)*P(Y)
If these features are not independent then P(X,Y) != P(X)*P(Y)
This can vastly affect the probability calculations of the classifier.

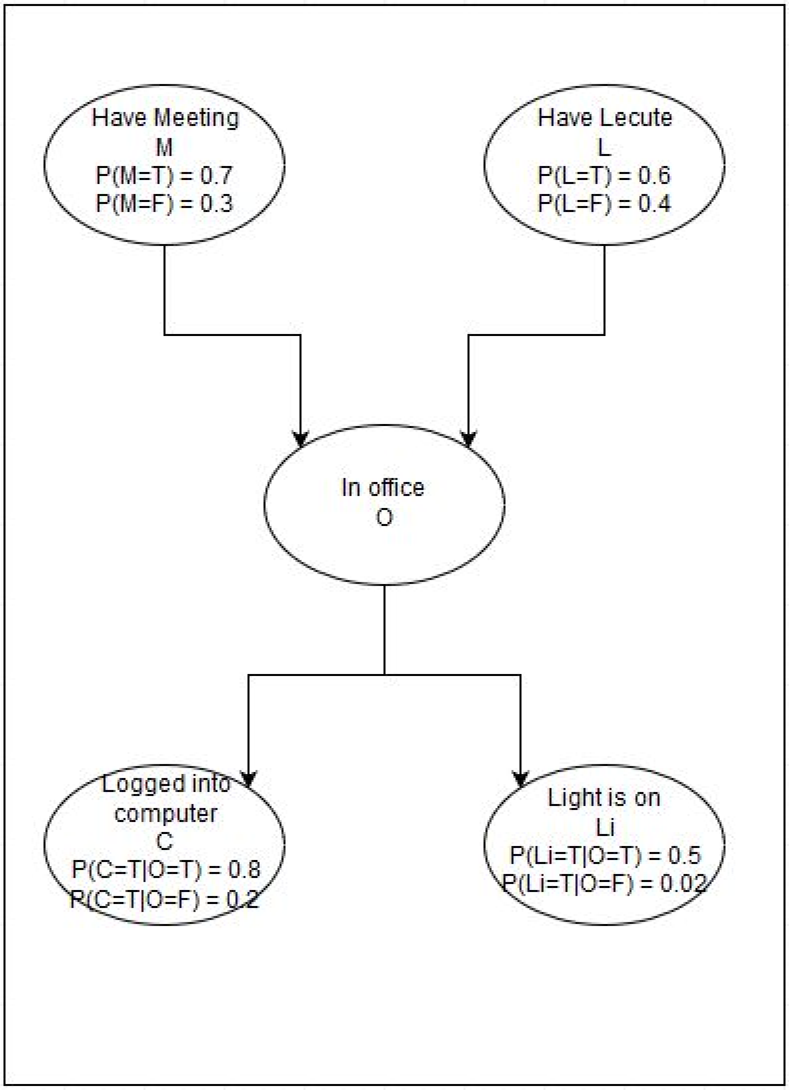
“ Each email is specified by 12 binary attributes, indicating the presence of features such as “Viagra”, “MILLION DOLLARS”, significant amounts of text in CAPS, an invalid reply-to address, and so on.”

It is likely that the emails features are not independent. I believe this because if an email contains “MILLION DOLLARS” it is more likely to contain significant amounts of text in caps which means that these features are not independent.

Part 3

1.

Bayesian network:



O node probabilities:

M	L	P(O=T|M, L)
T	T	0.95
F	T	0.80

T	F	0.75
F	F	0.06

2.

P(M,L,O,C,Li)=P(M)*P(L)*(O|M,L) *P(C|O)*P(Li|O) 24+24+26+27+27=10
There are 10 free parameters in this Bayesian network.

3.

Rachel has lectures(L), has no meetings(M), she is in her office(O) and logged on her computer(C) but with lights(Li) off.
= P(L=T,M=F,O=T,C=T,Li=F)
=P(M=F)*P(L=T)*P(O=T|M=F,L=T)*P(C=T|O=T)*P(Li=F|O=T)
=0.3*0.6*0.8&0.8*(1-0.5)
=0.0576

4.

Probability rachel is in her office.
= P(O=T)
=P(O=T,M=T,L=T)+P(O=T,M=F,L=T)+P(O=T,M=T,L=F)+P(O=T,M=F,L=F)
=P(O=T|M=T,L=T)*P(M=T,L=T)+P(O=T|M=F,L=T)*P(M=F,L=T)+P(O=T|M=T,L=F)*P(M=T,L=F)+
P(O=T|M=F,L=F)*P(M=F,L=F)

Because L and M are independent.

P(O=T)
= P(O=T|M=T,L=T)*P(M=T)*P(L=T) + P(O=T|M=F,L=T)*P(M=F)*P(L=T) +
P(O=T|M=T,L=F)*P(M=T)*P(L=F) + P(O=T|M=F,L=F)*P(M=F)*P(L=F)
= 0.95*0.7*0.6 + 0.8*0.3*0.6 + 0.75*0.7*0.4 + 0.06*0.3*0.4
= 0.399 + 0.144 + 0.21 + 0.0072
= 0.7602

5.

If rachel is in her office(O=T), what is the probability that she is logged on(C=T), but her lights are off(Li=F)

P(C=T,Li=F|O=T)
= P(C=T,Li=F|O=T)/P(O=T)
= P(O=T)*P(C=T|O=T)*P(Li=F|O=T)/P(O=T)
= P(C=T|O=T)*P(Li=F|O=T)
= 0.8*(1-0.5)
= 0.4

6.

If rachel is logged on(C=T), how does this effect how likely her light is on (P(Li))?
P(C=T|O=T) = 0.8
P(C=T|O=F) = 0.2

Probability that Rachel is in the office given that she is logged into her computer:
P(O=T|C=T) = P(C=T|O=T)*P(O=T)/P(C=T)
P(C=T) = P(C=T|O=T)*P(O=T) + P(C=T|O=F)*P(O=F)
= 0.8*0.7602 + 0.2*(1-0.7602) = 0.65612
P(O=T|C=T) = 0.8*0.7602/0.65612 = 0.927
P(O=F|C=T) = 1-0.927 = 0.073

Probability that the light is on given that Rachel is logged in:
P(Li=T|C=T) = P(Li=T,O=T|C=T)+P(Li=T,P=F|C=T)
= P(Li=T|O=T)*P(O=T|C=T) + P(Li=T|O=F)*P(O=F|C=T)
= 0.5*0.927 + 0.02*0.073 = 0.46496

Part 4

1.

P(P=t|X=t)

i)

Evidence variables: X
Hidden variables: S,C,D
Query variables: P

ii)

Join **P(X|C)** and **P(C|P,S)**
This gives **P(X,C|P,S)**
Eliminate C to get **P(X|P,S)**

Join **P(S)** and **P(X|P,S)**
This gives **P(X,S|P)**
Eliminate S gives **P(X|P)**

Join **P(X|C)** and **P(P)**
This gives **P(X,P)**
Eliminate P gives **P(X)**
Find **P(P|X)** from P(X,P) and P(X)

iii)

Using the steps shown above

P(X|C)

X	**C**	**P(X|C)**
1	1	0.9
1	0	0.2
0	1	0.1
0	0	0.8

P(C|P,S)

C	P	S	P(C|P,S)
1	1	1	0.05
0	1	1	0.95
1	0	1	0.03
0	0	1	0.97
1	1	0	0.02
0	1	0	0.98
1	0	0	0.001
0	0	0	0.999

P(X,C|P,S)

X	C	P	S	P(X,C|P,S)
1	1	1	1	0.045
0	1	1	1	0.005
1	0	1	1	0.19
0	0	1	1	0.76
1	1	0	1	0.027
0	1	0	1	0.003
1	0	0	1	0.194
0	0	0	1	0.776
1	1	1	0	0.018
0	1	1	0	0.002
1	0	1	0	0.196
0	0	1	0	0.784
1	1	0	0	0.0009
0	1	0	0	0.0001
1	0	0	0	0.1998
0	0	0	0	0.7992

P(X|P,S)

X	P	S	P(X|P,S)
1	1	1	0.235
0	1	1	0.765
1	0	1	0.221
0	0	1	0.779
1	1	0	0.214
0	1	0	0.786
1	0	0	0.2007
0	0	0	0.7993

P(S)

S	P(S)
1	0.3
0	0.7

P(X,S|P)

X	S	P	P(X,S|P)
1	1	1	0.0705
0	1	1	0.2295
1	0	1	0.1498
0	0	1	0.5502
1	1	0	0.0663
0	1	0	0.2337
1	0	0	0.14049
0	0	0	0.55951

P(X|P)

X	P	P(X|P)
1	1	0.2203

0	1	0.7797
1	0	0.20679
0	0	0.79321

P(P)

P	**P(P)**
1	0.9
0	0.1

P(X,P)

X	**P**	**P(X,P)**
1	1	0.19827
0	1	0.70173
1	0	0.020679
0	0	0.079321

P(X)

X	**P(X)**
1	0.218949
0	0.781051

P(P|X)

P	**X**	**P(P|X)**
1	1	0.905553
0	1	0.094447
1	0	0.898443
0	0	0.101557

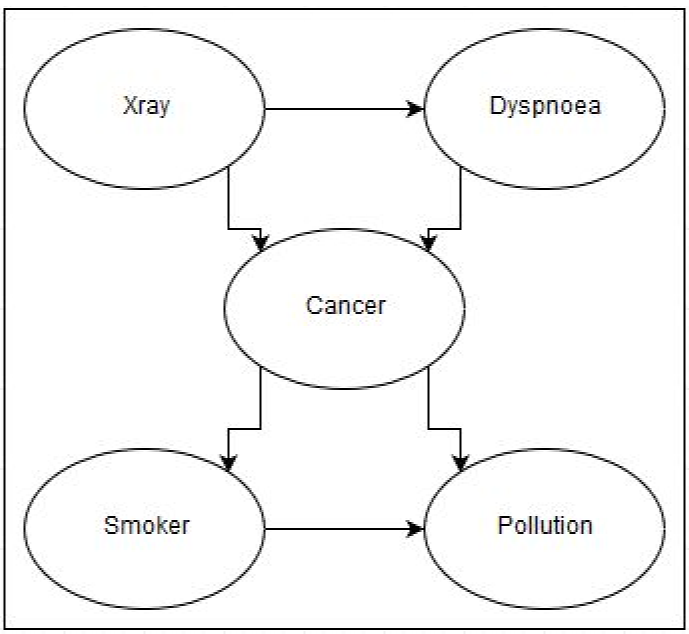
Finally P(P=1|P=1) = 0.905553

2.

Variables that are independent of each other or conditionally independent given another variable:

1. P(P) and P(S) are independent
2. P(X|C) and P(D|C) are conditionally independent
3. P(X|C) and P(P|C) are conditionally independent

3.



Nodes:
Given the order xray, dyspnoea cancer, smoker, pollution, the nodes have been added as such giving this structure.

Connections:
Xray -> Dyspnoea:
Connected due to their common cause cause cancer.

Xray -> Cancer:
Cancer can be the reason someone would get an xray

Dyspnoea -> Cancer:
Cancer can be the reason someone has dyspnoea

Cancer -> Smoker:
Smoking can be the reason someone has cancer

Cancer -> Pollution:
Pollution can be the reason someone has cancer

Smoker -> Pollution:
Smoking and pollution share the effect of causing cancer.