Part 1.

Q1.

1.

X	Y	P(X,Y)	probability rules used
0	0	0.09	P(X,Y) = P(X) * P(Y | X) = 0.3 * 0.3 = 0.09
0	1	0.24	P(X,Y) = P(X) * P(Y | X) = 0.3 * 0.7 = 0.21
1	0	0.49	P(X,Y) = P(X) * P(Y | X) = 0.7 * 0.8 = 0.56
1	1	0.14	P(X,Y) = P(X) * P(Y | X) = 0.7 * 0.2 = 0.14

2.

X	Y	Z	P(X,Y,Z)
0	0	0	0.054
0	0	1	0.036
0	1	0	0.168
0	1	1	0.042
1	0	0	0.336
1	0	1	0.224
1	1	0	0.112
1	1	1	0.028

probability rules used: P(X,Y,Z) = P(X) * P(Y|X) * P(Z|X,Y)

Cause Z is independent from X given Y.
P(X, Y, Z) = P(X) * P(Y|X) * P(Z|Y)

X = 0,Y=0, Z=0: P(X, Y, Z) = 0.3 * 0.3 * 0.6 = 0.054
X = 1,Y=0, Z=1: P(X, Y, Z) = 0.7 * 0.8 * 0.4 = 0.224
X = 1,Y=1, Z=0: P(X, Y, Z) = 0.7 * 0.2 * 0.8 = 0.112
X = 1,Y=1, Z=1: P(X, Y, Z) = 0.7 * 0.2 * 0.2 = 0.028

3.

(1)probability rules used:

- The sum rule
  - $P(X = x) = \sum_{y \in \Omega} P(X = x, Y = y)$
  - $P(X_1 = x_1, \ldots X_n = x_n) = \sum_{y_1 \in \Omega_1, \ldots, y_m = \Omega_m} P(X_1 = x_1, \ldots X_n = x_n, Y_1 = y_1, \ldots, Y_m = y_m)$

P(Z=0)
= P(X=1, Y=1, Z=0)+ P(X = 1, Y = 0, Z = 0) + P(X=0,Y=0, Z=0) + P(X=0, Y=1, Z=0)
= 0.112 +0.336 +0.168+0.054
= 0.67

P(X=0,Z=0)
= P(X = 0, Y = 1, Z = 0) + P(X=0, Y=0, Z=0)
=0.168+0.054 = 0.222

(2)probability rules used:

If A and B are independent of each other, then
- $A \perp B$
- P(A | B) = P(A)
- P(B | A) = P(B)
- P(A, B) = P(A) * P(B)

If X and Z are independent to each other, P(X,Z) should equals P(X) * P(Z)

P(X=0,Z=0) = 0.222 but P(X=0) * P(Z=0) = 0.3 * 0.67 = 0.201

So X and Z are not independent to each other.

4.

(1)

$P(Z = 1)$

$= P(X = 0, Y = 0, Z = 1)+ P(X = 0, Y = 1, Z = 1)+ P(X=1, Y=0, Z=1)+ P(X=1, Y=1, Z=1)$

$=0.036+0.042+0.224+0.028 =0.33$

$P(X = 1, Y = 0|Z = 1)$

$= P(X = 1, Y = 0, Z = 1) / P(Z = 1)$

$=0.224/0.33 = 0.679$

(2)

$P(Y = 0, Z = 0)$

$= P(X=0, Y=0, Z=0)+ P(X = 1, Y = 0, Z = 0)$

$=0.054+0.336 =0.39$

$P(X = 0 \mid Y = 0, Z = 0)$

$= P(X = 0, Y = 0, Z = 0) / P(Y = 0, Z =0)$

$= 0.054/0.39=0.138$

Q2.

1.

(i)

$P(B,C)$

$= P(C) * P(B|C)$

$= 0.4 * 0.2$

$=0.08$

(ii)

$P(\neg A|B)$

$=1-P(A|B)$

$=0.7$

(iii)

Because $A \perp B|C$,

$P(A, B|C)$

$= P(A|C) * P(B|C)$

$=0.5 * 0.2$

=0.1

(iv)
Because A ⊥ B|C,
P(A|B,C)
= P(A|C)
=0.5

(v)
P(A, B, C)
=P(A,B) * P(C|B)
= P(B) * P(A | B) * (P(C) * P(B | C))/P(B)
=0.7 * 0.3 * ((0.4 *0.2)/0.7)
=0.024

Q3.

P(C) = 0.997 (Common DF)
P(R) = 0.003 (Rare DF)
P(W|C) = 0.001(Common DF mutate to have extra wings)
P(W|R) = 1 (Rare DF have extra wings)
We need to figure out the probability of rare species when we see a DF have extra wings.
So we need to figure out P(R|W)
P(R|W) = (P(W|R) * P(R))/ P(W)
But we don't know P(W)
P(W) = P(W,R) + P(W,C) (sum rule)
= P(R) * P(W|R) +P(C) * P(W|C)
= 0.003 * 1 + 0.997 * 0.001
= 3997/1000000
P(R|W) = (1 * 0.003)/(3997/1000000)
= 3000/3997
= 0.7506
So the probability is 0.7506

Part 2.

Q1.

P(F0 = 0| C = 0) = 0.6442953020134228
P(F0 = 0| C = 1) = 0.35570469798657717
P(F0 = 1| C = 0) = 0.3333333333333333
P(F0 = 1| C = 1) = 0.6666666666666666

P(F1 = 0| C = 0) = 0.4228187919463087
P(F1 = 0| C = 1) = 0.5771812080536913
P(F1 = 1| C = 0) = 0.4117647058823529
P(F1 = 1| C = 1) = 0.5882352941176471

P(F2 = 0| C = 0) = 0.6577181208053692
P(F2 = 0| C = 1) = 0.3422818791946309
P(F2 = 1| C = 0) = 0.5490196078431373
P(F2 = 1| C = 1) = 0.45098039215686275

P(F3 = 0| C = 0) = 0.6040268456375839
P(F3 = 0| C = 1) = 0.3959731543624161
P(F3 = 1| C = 0) = 0.39215686274509803
P(F3 = 1| C = 1) = 0.6078431372549019

P(F4 = 0| C = 0) = 0.6644295302013423
P(F4 = 0| C = 1) = 0.33557046979865773
P(F4 = 1| C = 0) = 0.5098039215686274
P(F4 = 1| C = 1) = 0.49019607843137253

P(F5 = 0| C = 0) = 0.5302013422818792

P(F5 = 0| C = 1) = 0.4697986577181208
P(F5 = 1| C = 0) = 0.6470588235294118
P(F5 = 1| C = 1) = 0.35294117647058826

P(F6 = 0| C = 0) = 0.4966442953020134
P(F6 = 0| C = 1) = 0.5033557046979866
P(F6 = 1| C = 0) = 0.21568627450980393
P(F6 = 1| C = 1) = 0.7843137254901961

P(F7 = 0| C = 0) = 0.6510067114093959
P(F7 = 0| C = 1) = 0.348993288590604
P(F7 = 1| C = 0) = 0.23529411764705882
P(F7 = 1| C = 1) = 0.7647058823529411

P(F8 = 0| C = 0) = 0.7583892617449665
P(F8 = 0| C = 1) = 0.24161073825503357
P(F8 = 1| C = 0) = 0.6666666666666666
P(F8 = 1| C = 1) = 0.3333333333333333

P(F9 = 0| C = 0) = 0.7114093959731543
P(F9 = 0| C = 1) = 0.28859060402684567
P(F9 = 1| C = 0) = 0.3333333333333333
P(F9 = 1| C = 1) = 0.6666666666666666

P(F10 = 0| C = 0) = 0.4161073825503356
P(F10 = 0| C = 1) = 0.5838926174496645
P(F10 = 1| C = 0) = 0.3333333333333333
P(F10 = 1| C = 1) = 0.6666666666666666

P(F11 = 0| C = 0) = 0.6644295302013423
P(F11 = 0| C = 1) = 0.33557046979865773
P(F11 = 1| C = 0) = 0.21568627450980393
P(F11 = 1| C = 1) = 0.7843137254901961

Q2.

Probability for Spam is 3.020244874387394e-06,Probability for non-spam is 0.0004620049715764379.

Instance : 1 is non_spam

Probability for Spam is 5.5140976761978446e-05,Probability for non-spam is 4.0855635930579417e-05.

Instance : 2 is spam

Probability for Spam is 0.00018644455371759411,Probability for non-spam is 0.00012776774190121569.

Instance : 3 is spam

Probability for Spam is 5.2350911156048155e-06,Probability for non-spam is 0.0006037954762596702.

Instance : 4 is non_spam

Probability for Spam is 5.863981931440459e-05,Probability for non-spam is 9.134498979293801e-05.

Instance : 5 is non_spam

Probability for Spam is 5.5933366115278225e-05,Probability for non-spam is 4.531325026841299e-05.

Instance : 6 is spam

Probability for Spam is 3.43552854461566e-06,Probability for non-spam is 0.00032863644196655 1.

Instance : 7 is non_spam

Probability for Spam is 6.190253957422096e-05,Probability for non-spam is 0.00039404283148337113.

Instance : 8 is non_spam

Probability for Spam is 0.00018644455371759411,Probability for non-spam is 3.6936543039323476e-05.

Instance : 9 is spam

Probability for Spam is 2.0416855350858785e-05,Probability for non-spam is 0.0006881308235775 48.

Instance : 10 is non_spam

Q3.

In naive Bayes algorithm, we assume that attributes are conditionally independent, but in real life, attributes will affect each other in most cases. For this dataset, we use naive Bayes based on the assumption, but email from an invalid reply address is more likely to contain an amount of upper case text, which means that these attributes are not independent

Part 3.

Q1.

$P(A,B,C,D,E,F,G) = P(A) * P(B) * P(C|A) * P(F|B) * P(D|B,C) * P(E|C) * P(G|E,D,F)$

Q2.

(i) A and B are independent with each other
(ii)A and D are conditionally independent given C
(iii)B and G are conditionally independent given D and F
(iv)D and F are conditionally independent given B
(v)C and G are conditionally independent given D and E

Q3.

1. node E

2. node D

$P(D|E) = P(D)$? No, $E \to D$

3. node F

$P(F|D,E) = P(F)$? No.

$P(F|D,E) = P(F|D)$? No

$P(F|D,E) = P(F|E)$? No, $D \to F$, $E \to F$

4. node G

$P(G|D,E,F) = P(G)$? Yes, $F \to G$, ~~no link from E or D~~ $D \to G$, $E \to G$

5. node B

$P(B|D,E,F,G) = P(B)$? No

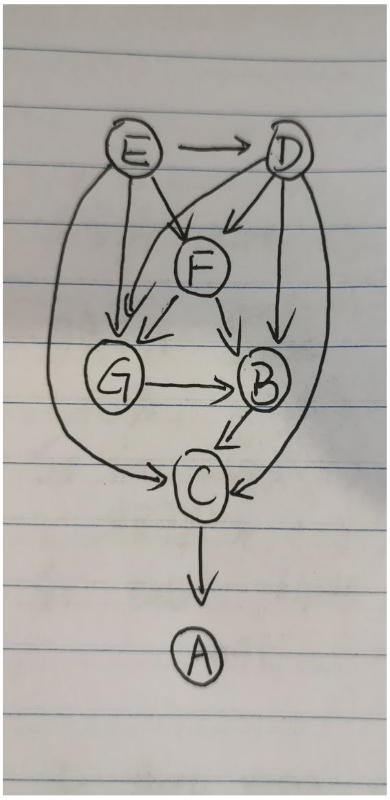
$P(B|D,E,F,G) = P(B|D)$? No

$P(B|D,E,F,G) = P(B|D,E)$? No

$P(B|D,E,F,G) =$ ~~$P(B|D,F,G)$~~ Yes, $D \to B$, $F \to B$, $G \to B$

$P(B|D,F,G)$?

6. node C



$P(C|B,D,E,F,G) = P(C)$? No

$P(C|B,D,E,F,G) = P(C|B,E,D)$? Yes, $B \to C$, $E \to C$, $D \to C$.

7. node A

$P(A|B,C,D,E,F,G) = P(A)$? No

$P(A|B,C,D,E,F,G) = P(A|C)$? Yes, $C \to A$.

Q4.

7. node A

$P(A|B,C,D,E,F,G) = P(A)$? No

$P(A|B,C,D,E,F,G) = P(A|C)$? Yes, $C \rightarrow A$.

$P(C) = P(C,A) + P(C,\neg A)$ (sum rule)

$= P(A) * P(C|A) + P(\neg A) * P(C|\neg A)$

$= 0.7 * 0.6 + (1-0.7) * 0.3$ (normalization rule)

$= 0.51$

$P(\neg B) = 1 - P(B) = 0.8$

$P(\neg C) = 1 - P(C) = 0.51$

$P(D) = P(D,B,C) + P(D,\neg B,C) + P(D,B,\neg C) + P(D,\neg B,\neg C)$ (sum rule)

$= (P(B) * P(C) * P(D|B,C)) + (P(\neg B) * P(C) * P(D|\neg B,C)) +$
$(P(B) * P(\neg C) * P(D|B,\neg C)) + (P(\neg B) * P(\neg C) * P(D|\neg B,\neg C))$ (product rule)

$= (0.2 * 0.51 * 0.7) + (0.8 * 0.51 * 0.5) + (0.2 * 0.49 * 0.6) + (0.8 \times 0.49 \times 0.2)$

$= 0.0714 + 0.204 + 0.0588 + 0.0784$

$= 0.4126$

So, $P(D)$ is 0.4126

Part 4.

Q1.

M: Has Meeting
LC: Has Lecture
O: Office
L: Light
C: Computer

M	P(M)
1	0.7
0	0.3

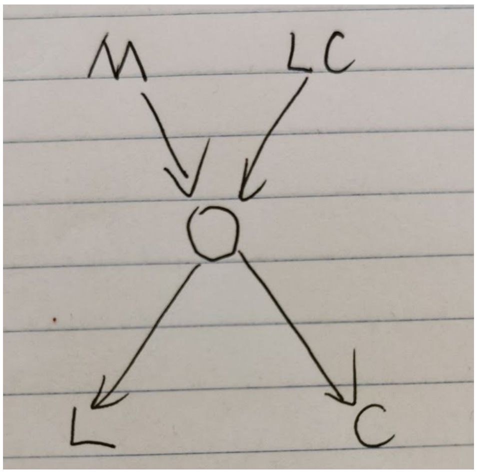
LC	P(LC)
1	0.6
0	0.4

M	LC	O	P(O|M,LC)
1	1	1	0.95
1	1	0	0.05
1	0	1	0.75

1	0	0	0.25
0	1	1	0.8
0	1	0	0.2
0	0	1	0.06
0	0	0	0.94

L	O	P(L|O)
1	1	0.5
0	1	0.5
1	0	0.02
0	0	0.98

C	O	P(C|O)
1	1	0.8
0	1	0.2
1	0	0.2
0	0	0.8



Q2.

2/2 +2/2+8/2+4/4+4/2 =10

Q3.

P(M=0) * P(LT=1) * P(O=1|M=0, LT=1)* P(L=0|O=1)* P(C=1|O=1)

=0.3*0.6*0.8*0.5*0.8 = 0.0576

Q4.

P(O|M=1, T=1) + P(O|M=1,T=0) + P(O|M=0,T=1) + P(O|M=0,T=0)

=0.95*0.7*0.6+0.75*0.7*0.4+0.8*0.3*0.6+0.06*0.3*0.4

=0.7602

Q5.

P(C=1|O=1)*P(L=0|O=1)

= 0.8*0.5

= 0.4

Q6.

Light and Computer are independent is given Office, but office is not given here. So there is no effect on the students' belief that Rachels' light is on.