AI ASSIGNMENT 3
Date of Submission :

Total Marks : 25 Marks
Theory section : 10 Marks
Computational section : 15 marks

Instructions:

1. Assignments are to be attempted individually.
2. Submit the assignment in a zip folder with name AI3_Rollnumber
3. Submit a file named AI3_Rollnumber_Result.pdf and combine the answers/results for the theory and computational question.
4. Submit a file named AI3_Rollnumber_Codes.py file for the python codes.
5. If any of the file is missing a penalty of -5 marks.
6. **Programming Language :** Python
7. **Extension and Penalty clause:**
   - Even a 1 minute late submission on google classroom will be considered as late. Please turn-in your submissions atleast 5 minutes before the deadline.
   - Not explaining the answers properly will lead to zero marks.

Theory

1. (6 marks) Consider the statements below and answer the questions. For question (i) 0.5 for each statement and wrong identification of random variable -0.5 marks.
   (i) (2.5 marks) Identify the random variables in the statements below , and write each of the following statements using symbols for random variables, logical connectives where necessary, and conditional probability notation.
   (ii) (1 mark) Verify that these propositions create a valid probability distribution. List the set of axioms that they satisfy.
   (iii) (1 mark) Populate the full joint probability distribution table.
   (iv) (1.5 marks) Use the joint distribution table and check for conditional independence between all the random variables that you have identified.

   (a) About 82.5 % people have travelled and have caught either corona or other diseases.

	Corona		∼Corona		Total
	travel	∼travel	travel	∼travel	
mild	0.15	0.046	0.06	0.044	0.3
severe	0.22	0.025	0.1	0.055	0.4
death	0.055	0.004	0.24	0.001	0.3
Total	0.425	0.075	0.4	0.1	1

(b) Of the people who had travelled 15 % have mild and 22 percentage have severe cases of corona, respectively.

(c) Given that a person travelled the chance they caught a disease other than corona is 0.485 rounded to 3 decimal places.

(d) About 24 % of people died of diseases other than corona after travelling.

(e) There is 0.025 probability that a person has not travelled and has severe case of corona.

(f) Given a person has not travelled the probability that the person is severely sick is about 0.457 rounded to 3 decimal places.

(g) The probability that a person has died and did have corona is 0.059.

(h) About 70 % people had mild or severe cases of any disease.

(i) There is 80 % chance that a person has travelled given that he is severely sick.

(j) There is 50 % chance a person had corona whether they travelled or not.

Answer Part (i)
Given 3 random variables : $C-$ corona:-{True,False}, $T-$ Travel:-{True,False} , $Severarity-$: takes three forms $S-$ severe ,$M-$ mild and $D-$ died.
Given :- $P(T) = 0.825$, $P(T \cap M \cap C) = 0.15$, $P(T \cap S \cap C) = 0.22$, $P(\sim C|T) = 0.485$, $P(\sim C \cap T \cap D) = 0.24$, $P(\sim T \cap C \cap S) = 0.025$, $P(S| \sim T) = 0.457$, $P(D \cap C) = 0.059$, $P(M \cup S) = 0.7$, $P(T|S) = 0.8$, $P(C) = 0.5$
Part(ii) and (iii) : Yes it does make a marginal probability distribution as in the above table. The table is also present in the recommended text book but with different values. All the values in the table can be derived from the given values.
Solving for the values : $P(\sim T) = 1 - P(T)$
$P(\sim C) = 1 - P(C)$
$P(M \cap \sim T \cap C) = P(C) - [(P(T \cap M \cap C) + P(T \cap S \cap C) + P(\sim T \cap C \cap S) + P(D \cap C))]$
$P(D) = 1 - (P(M) + P(S))$
$P(D \cap \sim C \cap \sim T) = P(D) - [P(D \cap C) + P(\sim C \cap T \cap D)]$
$P(\sim C|T) * P(T) = P(\sim C \cap T) = 0.400125 \approx 0.4$ rounded
$P(C \cap T) = P(T) - P(\sim C \cap T)$
$P(D \cap C \cap T) = [P(C \cap T) - (P(T \cap M \cap C) + P(T \cap S \cap C))]$
$P(D \cap \sim T \cap C) = P(D \cap C) - P(D \cap C \cap T)$
Here $P(T \cap S) + P(\sim T \cap S) = P(S) \Leftrightarrow P(\sim T|S) = 1 - P(T|S)$
$P(S) = P(S| \sim T) * P(\sim T)/(1 - P(T|S)) = 0.399875 \approx 0.4$ rounded
$P(M) = P(M \cup S) - P(S)$

$P(T \cap S) = P(T|S) * P(S)$
$P(S \cap T \cap \sim C) = P(T \cap S) - P(S \cap T \cap C)$
$P(S \cap \sim T \cap \sim C) = P(\sim T \cap S) - P(S \cap \sim T \cap C)$
$P(M \cap \sim C \cap \sim T) = P(\sim C \cap \sim T) - [P(S \cap \sim C \cap \sim T) + P(D \cap \sim C \cap \sim T)]$
$P(M \cap \sim C \cap T) = P(\sim C \cap T) - [P(S \cap \sim C \cap T) + P(D \cap \sim C \cap T)]$

(ii) and (iii) binary marks or 0.5 for a very good attempt.
(iv) if they have done the check $P(a, b|c) = P(a|c)P(b|c)$ for at-least 6 (0.25 each) values from the table otherwise 0.

2. (5 marks) You are playing a game where there are rooms, and you need to escape from one room to go to the next level. Each room shows you three doors. Behind one door is the key to the next room, and if you choose the other two doors, you lose a life. You pick one door. A person is standing waiting for you to choose your door. After you decide, he opens another door, revealing that you lost your life.

   (i) (1 mark) Should you switch your choice to the other unopened door to maximize your chance of winning the key?
   Answer Yes.Switch
   The probability of winning the key is higher if you switch doors.
   (a) Initially, there's a 1/3 chance that you picked the correct door and a 2/3 chance that you picked a losing door.
   (b) Man, who knows where the key is, will always reveal one of the losing doors (2/3 chance) after your initial choice. This doesn't change the fact that there's a 2/3 chance that your initial choice was wrong.
   (c) When you switch your choice, you are effectively betting on the remaining unopened door, which still has a 2/3 chance of hiding the key. So, by switching, you increase your chances of winning to 2/3, whereas if you stick with your initial choice, your chances remain at 1/3.

   (ii) (2 marks) However, if he occasionally makes a mistake, he reveals the loss of life with a probability of 1/3 and correctly reveals the key with a probability of 2/3. In this scenario, should you switch your choice to maximize your chances of winning the key?
   Answer Yes, Switch
   In this modified scenario, where the man occasionally makes a mistake, the probability of success is different.Initially, you have a 1/3 chance of picking the correct door (the one with the key) and a 2/3 chance of picking a losing door.
   (a) If the man reveals a loss of life behind one of the doors, there's a 1/3 probability that he made a mistake (revealing a loss of life) and a 2/3 probability that he correctly revealed the key.The answer depends on the probability of the man making a mistake and revealing a losing door.
   (b) If the man makes a mistake and reveals a losing door with a probability of 1/3, and he correctly reveals the key with a probability of 2/3, it's still better to switch because:

i. If you stick with your initial choice, you have a $1/3$ chance of winning (if you initially chose the door with the key) and a $2/3$ chance of losing (if you initially chose a losing door).

ii. If you switch, and the man reveals a loss of life (a $1/3$ chance due to his mistake), then you will win the key.

iii. If you switch, and the man correctly reveals the key (a $2/3$ chance), then you will lose.

So, if you switch, you will win the key with a probability of $1/3$ (Mans mistake) $+ 0$ (Man correctly reveals the key) $= 1/3$. This is still better than sticking with your initial choice, which gives you a $1/3$ chance of winning. So, in this scenario, you should switch your choice to maximize your chances of winning the key.

(iii) (1 mark) If you choose to switch, what is the conditional probability that you win the key if the man has mistakenly revealed the door that shows life lost?

Answer

Let A: You initially picked the correct door (probability $= 1/3$). B: The person mistakenly reveals the door showing the loss of life. We want to find P(A — B), the probability of winning the key given that the person has mistakenly revealed the door showing the loss of life. Bayes' Theorem states: $P(A/B) = P(B/A)P(A)/\ P(B)$

$P(B)=P(B/A)P(A)+P(B/A)P(A)$

Here, A is the event that you initially picked the wrong door, and $P(A)=2/3$ Given that you initially picked the correct door (A), the probability that the person mistakenly reveals the door showing the loss of life is $P(B/A)=1/3$ (as per the scenario). Now, let's calculate: $P(B) = 1/3 * 1/3 + 2/3*2/3$

$P(A/B) = (1/3*1/3)/(1/3*1/3+2/3*2/3) = 0.2$

(iv) (1 mark) Additionally, what is the conditional expectation of your prize (key/life lost) based on your choice to switch or stick, considering both possible scenarios? Would you choose to switch or stick based on the conditional expectation?

Answer

Scenario 1: The host correctly reveals a life lost. If you initially picked the door with the key(probability $= 1/3$), and you switch, you'll end up losing your life ($E = 0$). If you initially picked the door with lost life (probability $= 2/3$), and you switch, you'll end up with a key ($E = 1$). So, the conditional expectation of your prize if the man correctly reveals a lost life and you switch is $(1/3) * 0 + (2/3) * 1 = 2/3$.

Scenario 2: The man mistakenly reveals a lost life door. If you initially picked the key (probability $= 1/3$), and you switch, you'll end up with a lost life ($E = 0$). If you initially picked a door with lost life (probability $= 2/3$), and you switch, you'll end up with winning the key ($E = 1$). The conditional expectation of your prize if the host mistakenly reveals a lost life and you switch is $(1/3) * 0 + (2/3) * 1 = 2/3$.

Now, to find the overall conditional expectation of your prize if you switch, you need to consider both scenarios:

Conditional Expectation of Prize if You Switch = (Probability of Scenario 1) * (Conditional Expectation in Scenario 1) + (Probability of Scenario 2) * (Conditional Expectation in Scenario 2) = (1/3) * (2/3) + (2/3) * (2/3) = 2/9 + 4/9 = 6/9 = 2/3
So, the conditional expectation of your prize if you switch is 2/3, which means that, on average, if you switch, you can expect to win the key in 2 out of 3 trials.

Computational : - Some viva questions and possible answers :5 marks

1. Which of your model has best performance? why ?
2. Do we need feature selection methods in Bayesian network or is pruning better ? (Ans : Pruning in bayesian models helps in feature selection. Bayesian network can be used for feature selection.)
3. What is the difficulty in using continuous data in Bayesian network?
4. Is pre-processing a necessary process in Bayesian networks? (No)
5. What is the theory on which Bayesian network are built on ? (Bayes theorem)
6. Is bayesian network a classifier ? Explain. (probabilistic generative model)
7. What methods can be used to improve the bayesian models ?
8. How interpret-able is your model compared to others (ML models) ?
9. Why is Naive Bayes model called "Naive" ? It ignores prior distribution of parameters and assume independence of all features and all rows.