



United International University (UIU)

Dept. of Computer Science & Engineering (CSE)

Final Exam: Summer 2024

Course Code: CSE 3811, Course Title: Artificial Intelligence

Total Marks: 50

Duration: 2 hours

Answer all questions. Marks are indicated on the right side of each question.

[Any examinee found adopting unfair means will be expelled from the trimester/program as per UIU disciplinary rules.]

1. Answer all the questions:

- Write the generalized formula for the Chain Rule. [1]
- A doctor is called to see a sick child. The doctor has prior information that 90% of sick children in that neighborhood have the flu, while the other 10% are sick with measles. A well-known symptom of measles is rash. The probability of having a rash if one has measles is 0.95. However, occasionally children with flu also develop rash, and the probability of a rash if one has flu is 0.08. Upon examining the child, the doctor finds a rash. **What is the probability that the child has measles?** [2]
- Suppose we have 3 cards identical in form except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side is colored black. The 3 cards are mixed up in a hat, and one is randomly selected and put down on the ground. **If the upper side of the chosen card is red, what is the probability that the other side is black?** [3]
- In a study, physicians were asked what the odds of lung cancer would be in a person who was initially thought to have a 1% risk of cancer but who ended up with a positive test result (a test accurately classifies about 80% of cancerous tumors and 90% of harmless tumors.) **95 out of hundred physicians estimated the probability of cancer to be about 75%. Do you agree?** [4]

2. (a) A hospital is interested in modeling the probability of a patient having a disease (D). The disease causes two symptoms: fever (F) and fatigue (T). A positive blood test (B) is highly correlated with the presence of the disease, but not directly with the symptoms. Fatigue (T) can also be caused by lack of sleep (S), independent of the disease.

Each random variable has two possible events:

- Disease (D): {Yes, No}
- Fever (F): {Yes, No}
- Fatigue (T): {Yes, No}
- Blood test (B): {Positive, Negative}
- Sleep (S): {Sufficient, Insufficient}

(i) **Construct a Bayes net** based on this description, clearly indicating the relationships between the variables D (Disease), F (Fever), T (Fatigue), B (Blood test), and S (Sleep). [4]

(ii) Compare the memory complexity of storing the full joint probability distribution table with the memory required to store the Bayes net's conditional probability tables (CPTs). **Calculate the number of rows needed in each case.** [2]

(b) Consider the following Bayes net:

- A (Rain) has two possible values: {True, False}.
- B (Sprinkler) has two possible values: {True, False}.
- C (Wet Grass) has two possible values: {True, False}.

The conditional probability tables (CPTs) are provided as follows:

Table 1: P(A)	
A	P(A)
True	0.3
False	0.7

Table 2: P(B A)		
A	B	P(B A)
True	True	0.4
True	False	0.6
False	True	0.6
False	False	0.4

Table 3: P(C A, B)			
A	B	C	P(C A, B)
True	True	True	0.9
True	True	False	0.1
True	False	True	0.8
True	False	False	0.2
False	True	True	0.7
False	True	False	0.3
False	False	True	0.1
False	False	False	0.9

- (i) Compute the joint probability $P(A = \text{True}, B = \text{True}, C = \text{True})$. [3]
(ii) Compute $P(A = \text{True} \mid C = \text{True})$ using Bayes' rule. [3]

3. The following dataset contains four features related to football players: **Height**, **Weight**, **Speed**, and **Passing Accuracy**. The label, **Position**, indicates the playing position of each player.

Player Id	Height	Weight	Speed	Passing Accuracy	Playing Position
1	Tall	Fit	Average	Average	Defender
2	Medium	Fit	Fast	Great	Midfielder
3	Tall	Overweight	Average	Poor	Defender
4	Short	Lightweight	Fast	Great	Striker
5	Medium	Fit	Fast	Average	Midfielder
6	Tall	Fit	Average	Poor	Defender
7	Medium	Lightweight	Fast	Great	Midfielder
8	Medium	Fit	Fast	Average	Midfielder
9	Short	Lightweight	Fast	Great	Striker
10	Tall	Overweight	Average	Poor	Defender
11	Medium	Fit	Fast	Great	Midfielder
12	Medium	Fit	Fast	Great	Striker
13	Tall	Fit	Average	Poor	Defender
14	Short	Lightweight	Fast	Great	Striker
15	Medium	Fit	Fast	Average	Midfielder

Use Naive Bayes classifier to determine the probability of a Player with the features <Tall, Lightweight, Fast, Great> being a **Midfielder**. Show details calculation. [12]

4. In a retail system, we are analyzing customer purchasing behavior based on three attributes: **Product Category, Price Range, and Customer Rating**. The goal is to determine how these attributes influence the decision to buy a product (Yes/No).

Product Category	Price Range	Customer Rating	Buy
Electronics	High	5	Yes
Fashion	Low	4	No
Electronics	Low	5	Yes
Home & Kitchen	Low	3	No
Fashion	High	4	Yes
Electronics	High	3	No
Fashion	Low	5	Yes
Home & Kitchen	Low	4	No
Electronics	High	4	Yes
Fashion	High	3	No
Home & Kitchen	Low	5	Yes
Electronics	Low	3	No
Fashion	High	4	Yes
Home & Kitchen	High	3	No
Electronics	Low	5	Yes

Your task is to learn a Decision tree based on this data to predict whether a particular customer will buy the product or not. **Which should be the attribute in the root node of the decision tree?** Just find the attribute at the root node. Show detailed calculations. [8]

5. A smartphone company wants to predict the brand loyalty of its users based on recent surveys. They have recently conducted a survey among 120 people for two consecutive years (2017,2018) that yielded the following trends:

Initially, 50 were iPhone users, 40 were Google Pixel users, and the others were Samsung users. After introducing new features and pricing strategies, the following transitions occurred:

- Among the iPhone users, 10 switched to Google Pixel, 5 changed to Samsung, while the others remained iPhone users.
- Seven(7) Google Pixel users switched to iPhones, eight(8) switched to Samsung, and the rest stayed with Google Pixel.
- Six(6) Samsung users switched to iPhones, four(4) switched to Google Pixels, and the others remained with Samsung.

Based on this survey:

- (a) The company wants to document these trends in a table. **Help them find out the Transition Trend in a tabular format.** [3]
- (b) Help the company predict **what percentage of their customers will be iPhone users in 2022** [3]
- (c) The company wants to predict the long-run probability of a customer being in any of the 3 categories mentioned above (iPhone, Google Pixel, or Samsung users). Help them calculate these probabilities. [2]