



# United International University (UIU)

## Dept. of Computer Science & Engineering (CSE)

### Final Exam: Spring 2023

Course Code: CSE 3811, Course Title: Artificial Intelligence

Total Marks: 40

Duration: 2 hours

**Answer all questions.** Marks are indicated in 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. Suppose you have been given the following joint distribution table over the random variables A, B and C.

A	B	C	P(A, B, C)
+a	+b	+c	0.1
+a	+b	-c	0.15
+a	-b	+c	0.05
+a	-b	-c	0.25
-a	+b	+c	0.05
-a	+b	-c	0.1
-a	-b	+c	0.25
-a	-b	-c	0.05

Find the following probability values from the table:

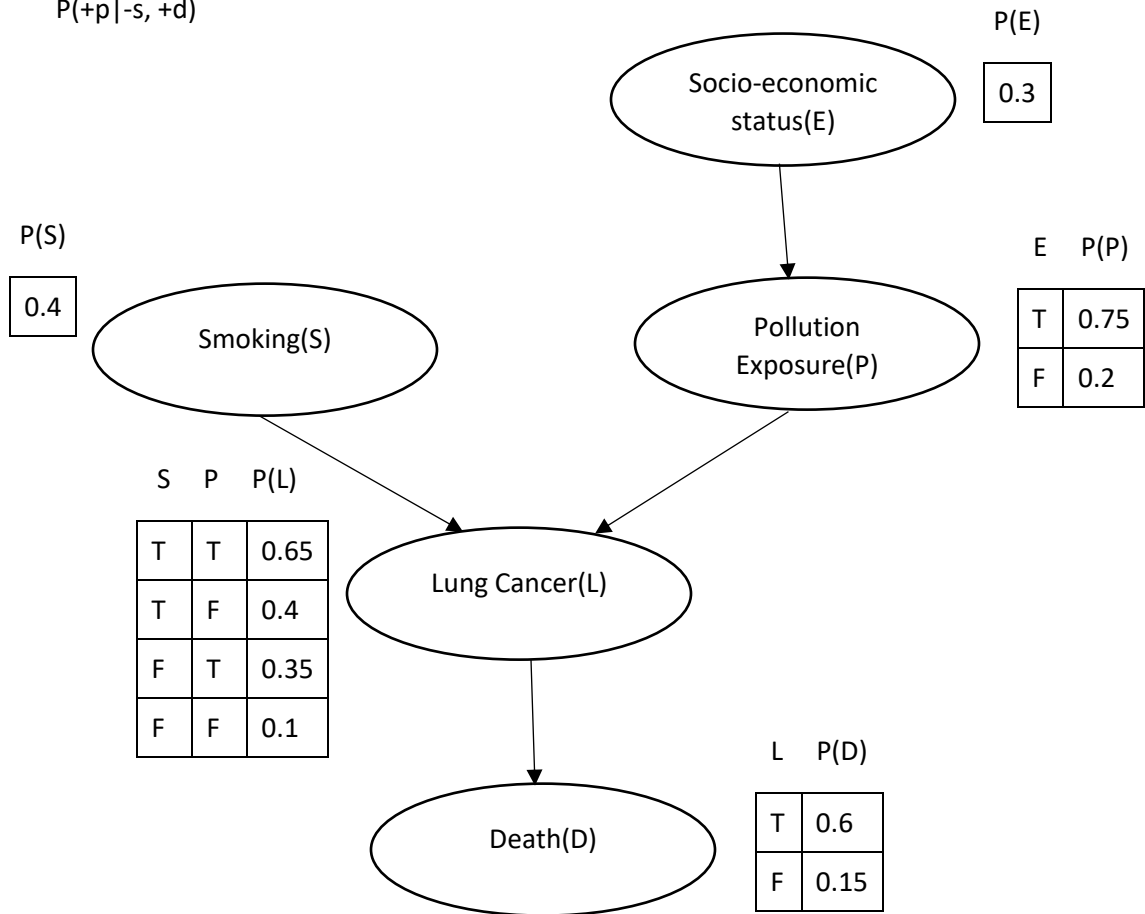
[1+2]

- $P(+b)$
- $P(A|C=+c)$

2. A test for a certain disease gives a positive result 95% of the time if the person actually carries the disease. However, the test also gives a positive result 2% of the time when the individual is not carrying the disease. It is known that 15% of the population carries the disease. If a person tests positive, what is the probability that he or she has the disease? [3]
3. You have done a survey on 200 residents of a city to analyze their travel preferences. 60 of the residents prefer summer season for travel, 40 prefer rainy season and the rest prefer winter. Among the summer loving people 30% choose mountain as destination, 50% choose sea and the rest prefer forest. Among the people travelling in the rainy season 20% choose mountain as destination, 30% choose sea and the rest prefer forest. Among the winter loving people 40% choose mountain as destination, 40% choose sea and the rest prefer forest.
- Based on this data, construct a full joint distribution among the two random variables Season(S) and Location(L). [2]
  - Calculate the following probabilities from your table:
    - Probability of a resident preferring sea. [1.5]
    - Probability of a mountain loving resident preferring winter season. [1.5]
4. a. Following traffic rules and maintaining vehicle fitness contribute to less traffic jam. Vehicle fitness is also impactful in reducing road accidents. Traffic jam affects economic growth, education and timely treatment for patients needing hospitals. Find appropriate random variables from this scenario and construct a bayes net. How many probability entries will be required for this bayes net? Explain your calculation. [2+1]

b. Determine the following probability information from the given bayes net. At each query, +x denotes X = true, and -x denotes X = false. [2+4]

- i.  $P(+e, -s, -p, -l, +d)$
- ii.  $P(+p|-s, +d)$



5. You are analyzing the trend of fuel type migration for energy source of countries. You have calculated the transition probability of how the countries change energy source yearly shown in the table below.

Next Year → Current Year ↓	Wind	Petrol	Nuclear
Wind	0.55	0.2	0.25
Petrol	0.05	0.6	0.35
Nuclear	0.05	0.1	0.85

- a. Given this year (Year 1) a country uses "Petrol" as its fuel source:
  - i. What is the probability of the country will use **Petrol** energy source in Year 3? [1.5]
  - ii. What is the probability of the country will use **Nuclear** energy source in Year 4? [1.5]
- b. Determine the probability of the used energy source being Wind, Petrol and Nuclear in the long run (stationary distribution) [3]

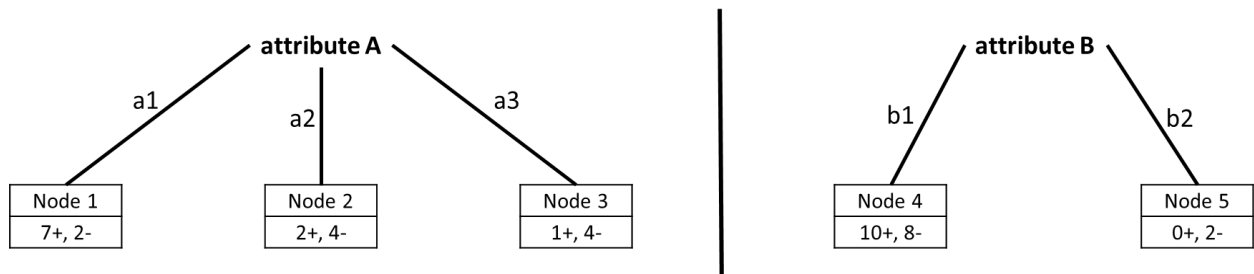
6. Consider a **binary** classification dataset with 20 samples has equal number of positive and negative samples but only two attributes: **A** and **B**.

Attribute **A** has three unique values: (a1, a2, a3)

Attribute **B** has two unique values: (b1, b2)

Target/Label Y has two class: yes represented by + and negative represented as -

The figure below displays two partitionings of the dataset that resemble the ID3 decision tree algorithm at the top level. The partitioning on the left is based on the values of attribute A, while the partitioning on the right is based on the values of attribute B.



Now answer the following questions:

[3+3+1]

- Calculate **Entropy** values for each of the five nodes shown and the parent node
- Calculate **Information Gain** values for both attributes A and B
- Which **attribute** will ID3 decision tree select first to partition the dataset at this level?

7. Consider the following dataset which contains four features and a label of some road accident cases. The features are **Weather condition**, **Road condition**, **Traffic condition**, **Engine problem** and the label is **Accident**, which represents whether there has been an accident or not.

SNo.	Weather condition	Road condition	Traffic condition	Engine problem	Accident
1	Rain	bad	high	no	yes
2	snow	average	normal	yes	yes
3	clear	bad	light	no	no
4	clear	good	light	yes	yes
5	snow	good	normal	no	no
6	rain	average	light	no	no
7	rain	good	normal	no	no
8	snow	bad	high	no	yes
9	clear	good	high	yes	no
10	clear	bad	high	yes	yes

Use Naive Bayes classifier to determine the probability of having an accident with the features <Rain, good, normal, no>. **Show details calculation.**

[7]