

BRAC UNIVERSITY
Department of Computer Science and Engineering

Examination: Semester Final
 Duration: 1 Hour 40 Minutes

Semester: Fall 2022
 Full Marks: 40

CSE 422: Artificial Intelligence

Answer 4 out of 5 from the following questions.
 Figures in the right margin indicate marks.

Name:	ID:	Section:
-------	-----	----------

1. CO6

You have been given a dataset containing 8 rows and four features (“unique_id”, “colour”, “size”, and “tail_length”). Each row represents a rat, and you are interested in finding the value of the label “has_disease”. Your goal is to build a decision tree from the table below:

unique_id	colour	size	tail_length	has_disease
#1	black	large	5.6	NO
#28	white	large	2.2	YES
#3	black	small	3.8	YES
#34	black	small	4.2	YES
#26	black	large	1.2	NO
#11	white	small	1.4	NO
#32	black	small	2.3	YES
#13	white	large	3.5	NO

- Among the columns “colour”, and “size”, which would you choose as the root node of your decision tree if you used Information Gain (IG) to construct the tree? **Construct** a decision tree using these two variables. 4
- Why is “unique_id” a bad choice for root node? **Explain.** 2
- If you want to use “tail_length” as a node while building the decision tree, what must you do beforehand? **Explain.** 2
- Given two variables $X = \{\text{outcome of an unbiased dice that can be rolled to obtain an integer value between 1 and 6 with equal probability}\}$ and $Y = \{\text{outcome of an unbiased coin that can be tossed with an equal probability of heads and tails}\}$, **Identify** which is larger: entropy of X or entropy of Y ? 2

2. CO5

- In the table below, you are given a dataset containing 9 rows and 3 features X_1, X_2, X_3 . Y is the label. Using naive bayes classifier, **Determine** the most likely value of Y if $X_1 = 1, X_2 = a, X_3 = q$. You don’t need to use any kind of smoothing or normal distribution. Just derive the probabilities from frequencies. 10

X_1	X_2	X_3	Y
1	a	p	0
2	b	r	1
3	b	p	1
3	c	q	1
2	c	r	0
1	b	q	1
2	a	p	0
3	a	r	1
3	b	q	0

3. CO5
- Suppose X is a discrete random variable whose domain is exhaustive and mutually exclusive. Now the domain of $X = \{A, B, C\}$. Assume $P(A) = 0.5$ and $P(B) = 0.3$, then **determine** (i) $P(C)$ and (ii) $P(A \cup B)$. 3
 - Suppose two coins are tossed simultaneously. Assume Event A = the 1st coin coming up heads and Event B = the 2nd coin coming up tails. Now **determine** the value of $P(A \cap B)$. 3

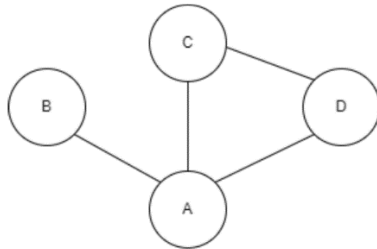
c.

	A		A'	
	B	B'	B	B'
C	0.1	0.2	0.2	Y
C'	X	0.1	0.1	0.1

Using the given table answer the following questions:

- Assume the events A and B showcase absolute independence and $P(B|A) = 0.5$. Now **determine** the value of X 1
- Using the ans obtained from (I), **determine** the value of Y 1
- Using the ans obtained from (I) and (II), **determine** the value of $P(A|B \cap C)$ 2

4. CO3



Consider the constraint graph of a problem above, where each region has to be filled up with either 1, 2, or 3. No two adjacent regions can have the same digit.

- Formulate** the variable, domain, constraint, and the goal of the problem 2
- Based on the variable ordering procedure, mention the order of variables to be assigned with digit. **Provide** adequate explanation for your ordering. 4
- Consider that node B already has digit 1 and all the other nodes are empty. If you are to provide digit to node D next, which digit should you pick? **Identify** your choice based on value selection procedure. 2
- If node B has digit 1, node D has digit 2, and rest of the nodes are unassigned, does the constraint graph remain arc consistent? Why or why not? **Explain.** 2

5. CO6

x	y
2	41
4	64
3	57

- Considering the data points above, let's say the hypothesis of a regression is $y = 10x + 20$. In this case, **Estimate** the amount of error based on an appropriate error function. 3
- What procedure you can follow to reduce the amount of error produced? **Discuss** in brief. 4
- How is logistic regression different than regular linear regression? **Discuss.** 3