0112

National University of Computer and Emerging Sciences

Lahore Campus

Applied Machine Learning (CS4104)

Sessional-I Exam

Total Time (Hrs): 1

Total Marks: 20

Total Questions: 2

Date: Feb 24th 2025

Course Instructor(s)

Dr Kashif Zafar

Mr Razi Uddin

Ms Anosha Khan

[redacted] Roll No

[redacted] Section

[redacted] Student Signature

Instructions:

- **Attempt all questions on answer book**
- **Show complete working step by step**

CLO # 2: Apply supervised learning techniques to solve classification problems

Q1: **[13 points]**

A zoologist is working on a classification project to identify two different types of creatures on a newly discovered planet: **Type M** and **Type H (M= Martians (fictional creature *belongs to* Mars) & H (Human)**. The zoologist has collected data on four physical characteristics: ***Color*** (Green or Not Green), number of **Limbs** (2 or 3), **Size** (Small or Tall), and whether or not the creature is **Odorous** (Yes or No). The data collected is as follows:

	Creature Type	Color	Limbs	Size	Odorous
1	M	No	3	Small	Yes
2	M	Yes	2	Tall	No
3	M	Yes	3	Tall	No
4	M	No	2	Small	Yes
5	M	Yes	3	Tall	No
6	H	No	2	Tall	Yes
7	H	No	2	Small	No
8	H	No	2	Tall	No
9	H	Yes	2	Small	No
10	H	No	2	Tall	Yes

a) Which attribute would the ID3 algorithm choose to use for the root of the tree (no pruning)? **[2 points]**

b) Draw the full decision tree that would be learned for this data. **[8 points]**

c) Suppose we have a validation set as follows. What will be the prediction of the tree for each test instance. **[3 points]**

National University of Computer and Emerging Sciences

Lahore Campus

Color	Limbs	Size	Odorous
Yes	2	Tall	No
No	3	Small	Yes
Yes	3	Small	Yes

CLO # 2: Apply supervised learning techniques to solve classification problems

Q2: **[7 points]**

You are given the following dataset, where the goal is to predict the **"Marks (%)"** based on the features **"Study Method"** and **"Attendance Level"** using a regression tree model:

Study Method	Attendance	Marks (%)
Online	High	90
Offline	Low	60
Online	Medium	75
Offline	High	85
Hybrid	Low	65

a) Build **Regression Tree** for the above dataset? Stop splitting further if less than 3 training examples left in a branch. **(5 points)**

b) Based on your model, what would be the predicted **marks for a student who studies offline with medium attendance? (2 points)**