BRAC University
Department of Computer Science and Engineering

CSE 422: Artificial Intelligence (C)

Quiz 04: Spring 2025 Time: 30 Minutes Marks: 10

Name		ID		Section	

1. You build a decision tree classifier on a dataset with 5,000 labeled records. You test it with different `max_depth` values and observe the following. [3+2]

Max Depth	Training Accuracy	Test Accuracy
2	72%	70%
5	85%	83%
10	98%	84%
15	100%	76%

   (a) Identify the depths where **underfitting** and **overfitting** occur. Explain briefly.

   (b) Which tree depth provides the **best generalization**? Justify your answer in terms of the *bias-variance tradeoff*.

2. Consider a small dataset of 8 samples used to predict whether a customer will purchase a product based on two features: Weather (Rainy or Sunny) and Temperature (Cold or Hot). The class label is Purchase (Yes or No). The dataset is as follows:

Weather	Temperature	Purchase
Rainy	Cold	Yes
Rainy	Hot	No
Sunny	Cold	No
Sunny	Hot	Yes
Rainy	Hot	Yes
Rainy	Cold	Yes
Sunny	Hot	No
Sunny	Cold	No

Based on Information Gain, which attribute should be chosen as the root of the decision tree? [5]