Graphic Era HILL UNIVERSITY
Established by an Act of the State Legislature of Uttarakhand (Adhiniyam Sankhya 12 of 2011)
University under section 2(f) of UGC Act, 1956

Mid Term (Odd) Semester Examination October 2024

Roll no 22 9 4 0 3 8

Name of the Course and semester: B.Tech.(CSE) & V
Name of the Paper: Machine Learning
Course Code: TCS 509

Time: 1.5 Hour | Maximum Marks: 50

Note

(i) All Questions are compulsory.
(ii) Each question carries 10 marks.

Q.1 **(10 Marks)** **CO1**

(a) Explain five potential applications of Machine Learning. Establish technical flow of machine learning.

OR

(b) What are different challenges to achieve Machine Learning? Provide proper example for the same.

Q.2 **(10 Marks)** **CO1**

(a) Given a list of numbers: [4, 6, 2, 8, 4, 10, 4, 6, 2], write Python code to calculate the mean and mode of the dataset. If there are multiple modes, display all of them.

OR

(b) Write Python code to calculate the standard deviation of the following dataset: [12, 15, 14, 10, 18, 20, 16, 19, 17]. Explain what the standard deviation tells you about the data.

Q.3 **(10 Marks)** **CO1**

(a) Explain the Reinforcement learning with the help of proper example.

OR

(b) What is the difference between machine learning and deep learning? Explain with example.

Q.4 **(10 Marks)** **CO2**

(a) Given a dataset with a date column, write Python code to extract meaningful features from the date, such as the year, month, day, day of the week, and whether it's a weekend. Also, create a new feature that represents the difference in days between two date columns (if available).

OR

(b) Load a dataset and calculate the correlation matrix between the numerical columns. Then, create a heatmap to visualize the correlations and identify which features are highly correlated.

Q.5 **(10 Marks)** **CO2**

(a) You are given a dataset containing missing values in various columns. Write a Python function to identify the number of missing values in each column, and then handle them by either removing the rows with missing values or filling them with the mean/median/mode, based on the data type of the column.

OR

(b) Write Python code to detect outliers in the dataset using the Interquartile Range (IQR) method. Once detected, either remove or replace the outliers, and visualize the data before and after the outlier treatment.