



## Term Evaluation (Odd) Semester Examination September 2025

Roll no.....

Name of the Course: B. TECH(CSE)

Semester: V

Name of the Paper: Machine Learning

Paper Code: TCS 509

Time: 1.5-hour

Maximum Marks: 50

**Note:**

- (i) Answer all the questions by choosing any one of the sub-questions
- (ii) Each question carries 10 marks.

Q1.	(10 Marks)	
a. What is meant by an outlier? How do outliers affect the mean and standard deviation? Can two different datasets have the same mean but different standard deviations? Give an example.	CO1	
OR		
b. For the Given dataset: 5, 8, 10, 15, 18, 22, 34, 46 (i) Find the mean, median, and mode. (ii) Compute the average deviation from mean. (iii) Compute the standard deviation and total sum of squares		
Q2.	(10 Marks)	CO2
a. Explain the concept of Machine Learning. Discuss the types of Machine Learning with examples. How do Supervised and Unsupervised Learning differ in their approach and applications?	CO2	
OR		
b. Differentiate between supervised, unsupervised, and reinforcement learning with suitable examples. Why is reinforcement learning referred to as learning by trial and error?		
Q3.	(10 Marks)	CO1
a. Given a dataset with missing and inconsistent values, demonstrate the steps you would take to clean and transform the data.	CO1	
OR		
b. Explain any five types of data Visualization used in Exploratory Data Analysis (EDA) using Matplotlib.		
Q4.	(10 Marks)	CO2
a. Describe the basics of dictionaries. Define a Python function 'orangecap(d)' which reads a dictionary 'd' of the following form and identifies the player with the highest total score. The function should return a pair (playername, topscc.e), where playername is the name of the player with the highest score and topscore is the total runs scored by the player. Input: orangecap( { 'test1': { 'Dhoni':74, 'Kohli':150 }, 'test2': { 'Dhoni':29, 'Pujara':42 } })	CO2	
OR		
b. Read the file into a string and split it to obtain all the words in the file. Build a dictionary that maps each word to the number of its occurrences in file. Print the dictionary and the word that has the most occurrences in the file.		
Q5.	(10 Marks)	CO3
a. What is exploratory data analysis (EDA), and why is it necessary?		
OR		
a. Explain the following: (i) Numpy (ii) Scikit-learn (iii) Seaborn (iv) Pandas		