Enrolment No.: 524CSE00771	
BENNETT UNIVERSITY THE TIMES GROUP	**School:** School of Computer Science Engineering & Technology

MID-TERM EXAMINATION, ODD SEMESTER OCTOBER 2025

Programme: BTech	**Semester: III**
Course Code: CSET 211	**Course Name: Statistical Machine Learning**
Time: 01 Hour	**Max. Marks:** 20

Instructions:

1. *Attempt all the sections.*
2. *Do not write anything on the question paper except enrolment number.*
3. *Assume missing data suitably, if any.*

CO No.	Course Outcome Statements	Bloom's Taxonomy Level
CO1	To understand key features and methods of Statistical Machine Learning (SML).	L2
CO2	To apply statistical machine learning methods on the available datasets.	L3
CO3	To evaluate statistical machine learning techniques.	L5
CO4	To design solutions to real world problems using various machine learning techniques.	L6

SECTION A
(3Q x 2M = 06 Marks)

ALL QUESTIONS ARE COMPULSORY

S. No.	Approximate time to attempt Sections: Section A: 3 Questions: 10 min Section B: 2 Questions: 20 min Section C: 2 Question: 30 min	Marks	Course Outcome (Please mention CO1 or CO2 etc.)	BTL Level (Please mention L1 or L2 or etc.)
Q1	Differentiate between supervised learning, unsupervised learning, and reinforcement learning with one example each.	2	CO1	L2
Q2	A dataset has values [10, 15, 20, 25, 30]. Compute mean, median, and variance.	2	CO2	L3
Q3	Differeniate between label encoding and one-hot encoding? Provide an example.	2	CO1	L2

SECTION B
(2Q x 3M = 06 Marks)

ATTEMPT ANY TWO QUESTIONS

Q4	The marks of 6 students are [12, 15, 20, 18, 25, 30]. Find standard deviation, range and quartiles(Q1, Q2, Q3). Discuss on the spread of data.	3	CO3	L4
Q5	A company collected the following house data: **Size (sq. ft) / Price (₹ in lakhs)** 1000 / 50 1500 / 70 2000 / 90 Find a simple linear regression equation of Price on Size (use least squares method)	3	CO3	L4
Q6	Discuss the following: (i) Confusion Matrix (ii) Precision vs Recall (iii) Dummy variable trap	3	CO2	L3

SECTION-C
(2Q x 4M = 08 Marks)

ATTEMPT ANY TWO QUESTIONS

Q7	A dataset has the following points: (2,3), (4,5), (6,7), (8,9) with class labels [A, A, B, B]. Identify the class of the point (5,6) using KNN with k=3 (Euclidean distance) and discuss all steps.	4	CO3	L4
Q8	Consider the dataset below for binary classification: **Actual / Predicted** 1 / 1 TP 0 / 1 1 / 1 TP 0 / 0 TN 1 / 0 0 / 0 TN Construct the Confusion matrix and calculate Accuracy, Precision, Recall, and F1-score.	4	CO3	L3
Q9	You are analyzing car sales data with features: Price, Engine Size, Fuel Type, Mileage. Illustrate the steps of exploratory data analysis and preprocessing before applying multiple regression.	4	CO3	L4