JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -1 EXAMINATION- 2026

M.Tech-I Semester (CSE-Data Science)

COURSE CODE (CREDITS):22M1WCI133 (3) MAX. MARKS: 15

COURSE NAME: Introduction To Statistical Learning

COURSE INSTRUCTOR: Prof. Pardeep Kumar MAX. TIME: 1 Hour

Note: *(a) All questions are compulsory.*

(b) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

(c) Use of calculators is allowed

JUIT EXAM SEPTEMBER 2026

<table>
<tr><th>Q.No</th><th>Question</th><th>CO</th><th>Marks</th></tr>
<tr><td>Q1</td><td>Consider the following student data:
<table><tr><th>Student_Id</th><th>Student_Study_Hours (X)</th><th>Attendance Pecentage (Y)</th><th>Marks (Z)</th></tr><tr><td>1</td><td>2</td><td>60</td><td>55</td></tr><tr><td>2</td><td>3</td><td>65</td><td>60</td></tr><tr><td>3</td><td>4</td><td>70</td><td>65</td></tr><tr><td>4</td><td>5</td><td>75</td><td>72</td></tr><tr><td>5</td><td>6</td><td>80</td><td>78</td></tr><tr><td>6</td><td>7</td><td>85</td><td>85</td></tr></table>Build the multiple linear regression model for the above given data to predict marks obtained by a student on the basis of study hours and attendance. What would be the marks obtained by a student who is spending 10 hours on study and secure 80 % attendance in a department?</td><td>2</td><td>3+1</td></tr>
<tr><td>Q2</td><td>A survey of 300 students records their academic performance and study method:
<table><tr><th>Study Method</th><th>Poor</th><th>Average</th><th>Good</th><th>Row Total</th></tr><tr><td>Self-study</td><td>20</td><td>50</td><td>30</td><td>100</td></tr><tr><td>Coaching</td><td>15</td><td>35</td><td>50</td><td>100</td></tr><tr><td>Online</td><td>25</td><td>40</td><td>35</td><td>100</td></tr><tr><td>Column Total</td><td>60</td><td>125</td><td>115</td><td>300</td></tr></table>Test whether study method and academic performance are independent of each other. Also compute the degree of freedom. For your calculated degree of freedom, tabular value for chi square is 9.488 at level of significance 0.05.</td><td>2</td><td>4</td></tr>
</table>

190

Q3	A Logistic Regression model is used to predict whether patients have a particular condition. For 10 patients, the actual and predicted classes are:	1	4

Patient	Actual	Predicted
1	1	1
2	1	1
3	1	0
4	0	1
5	0	0
6	1	1
7	0	0
8	0	1
9	1	0
10	0	0

Calculate:

1. TP, TN, FP and FN
2. Accuracy
3. Precision
4. Recall
5. F1-score

Q4	The following data represents **house size and house price:**	2	3

House	Area (sq.ft.)	Price (₹ lakh)
H1	1000	50
H2	1200	55
H3	1500	65
H4	1800	72
H5	2000	80
H6	2200	90

Using **K = 3**, predict the price of a house having an area of **1600 sq.ft.**

MID SEMESTER EXAMINATION- SEPTEMBER-2026