

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -1 EXAMINATION- 2023

B.Tech-VII Semester (CSE/IT/ECE/CE)

COURSE CODE(CREDITS): 19B1WCI731 (2)

MAX. MARKS: 15

COURSE NAME: Computational Data Analysis

COURSE INSTRUCTORS: Dr. Nishant Sharma

MAX. TIME: 1 Hour

Note: (a) All questions are compulsory.

(b) Marks are indicated against each question in square brackets.

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

Q1. What is semi-supervised learning, and when is it applicable? Explain the advantages and disadvantages of using both labeled and unlabeled data in the learning process? How are model-based learning algorithms used to make predictions?

[2 marks] [CO-1]

Q2. Suppose you are a real estate agent and want to predict the selling prices of houses based on their square footage. You have collected data on various houses, including their square footage (in square feet) and selling prices (in thousands of dollars). You want to build a linear regression model to predict house prices based on square footage. Here's a small sample of your data:

Square footage (x)	Price (in lakhs)
1000	32
1050	40
1150	45
1400	62
1700	83

Use linear regression equation to predict selling price of a house with square footage of 1550 feet.

[3 marks] [CO-1]

Q3. Based on dataset provided below, use normal equation to find coefficients of linear regression model.

X =	1	1450	Y =	210
	1	1600		230
	1	1700		245
	1	1875		275
	1	1162		180

[4 marks] [CO-2]

Q4. Explain gradient descent in context of machine learning. What are different challenges associated when selecting learning rate in gradient descent. How does feature scaling affect performance of gradient descent?

[4 marks] [CO-2]

Q5. What are the limitations faced by logistic regression? What is Linear Discriminant Analysis (LDA) and how does it overcome limitations faced by logistic regression?

[2 marks] [CO-2]