

Roll No.

Total No. of Pages : 02

Total No. of Questions : 09

B.Tech. (CSE) (Sem.-6)

**MACHINE LEARNING**

Subject Code : BTCS618-18

M.Code : 79257

Date of Examination : 10-12-2025

Time : 3 Hrs.

Max. Marks : 60

**INSTRUCTIONS TO CANDIDATES :**

1. SECTION-A is COMPULSORY consisting of TEN questions carrying TWO marks each.
2. SECTION-B contains FIVE questions carrying FIVE marks each and students have to attempt any FOUR questions.
3. SECTION-C contains THREE questions carrying TEN marks each and students have to attempt any TWO questions.

**SECTION-A**

**1. Write briefly :**

- a) Why is model evaluation and validation important in machine learning and what methods are commonly used for these processes?
- b) What is the significance of a loss function in training machine learning models?
- c) How would you define a "dependent variable" in regression analysis and what role does it play in the modeling process?
- d) What is missing data in the context of machine learning and how can it affect model performance?
- e) How does a residual plot help in evaluating the goodness of fit of a regression model?
- f) What is the purpose of data transformation in machine learning? Give example.
- g) Explain the concept of the F1-score in classification tasks.
- h) What is classification in machine learning and how does it fundamentally differ from regression? Give appropriate example of each.
- i) Why is clustering important in machine learning? What are some practical applications where clustering is beneficial?
- j) How would you define specificity in the context of classification models and why is it an important metric?

## SECTION-B

2. What challenges arise when working with large datasets in association rule mining and what strategies can be implemented to overcome these challenges?
3. Provide an explanation of polynomial regression with a detailed example.
4. Explain the purpose of a residual plot in linear regression. What characteristics define a "good" residual plot and how does it contribute to evaluating model performance?
5. Define the concept of a "population" in genetic algorithms. How does the population evolve over successive generations?
6. In a binary classification scenario, a model exhibits high precision but low recall. Evaluate the real-world implications of this performance and suggest approaches to achieve a balanced trade-off between precision and recall for specific application needs.

## SECTION-C

7. Describe different data cleaning methods in detail, providing an example for each.
8. Examine the influence of various data types, including categorical and continuous variables, on the choice of split algorithms and the criteria used to evaluate splits.
9. Discuss real-world scenarios or problem, domains where mutation and elitism play a crucial role in the effectiveness of Genetic Algorithms. Provide examples illustrating the significance of these components.

**NOTE : Disclosure of Identity by writing Mobile No. or Making of passing request on any page of Answer Sheet will lead to UMC against the Student.**