

Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
B.Tech Degree S6 (S,FE) Examination December 2025 (2019 Scheme)

Course Code: CDT304
Course Name: MACHINE LEARNING CONCEPTS

Max. Marks: 100

Duration: 3 Hours

PART A*Answer all questions, each carries 3 marks.*

Marks

- 1 Distinguish between classification and regression, highlighting their differences with relevant examples. (3)
- 2 Consider a scenario with three binary input variables a1, a2, and a3, and one binary output variable z. What is the size of the hypothesis space for this setup? (3)

a1	a2	a3	z
0	1	1	0
1	0	1	1
0	0	0	0
1	1	1	1

- 3 What is back propagation learning? (3)
- 4 When is RBF preferable to MLP for classification tasks? (3)
- 5 Differentiate bagging and boosting algorithms. (3)
- 6 Explain Gaussian Mixture Models. (3)
- 7 What are the four operators used in GA? (3)
- 8 List the common dimensionality reduction techniques. (3)
- 9 Explain the significance of graphical models used in machine learning? (3)
- 10 What are the main types of clustering methods used in machine learning? (3)

PART B*Answer one question from each module, each carries 14 marks.***Module I**

- 11 a) Compare reinforcement learning with supervised and unsupervised learning (7)
- b) Describe the concept of association rule learning and name two algorithms (7)

commonly used for generating association rules.

OR

- 12 a) Explain Vapnik-Chervonenkis Dimension. (7)
- b) Consider a dataset for deciding whether to play tennis based on attributes such as "Outlook," "Temperature," and "Humidity," with binary outputs indicating "Play" (1 for yes, 0 for no). Given the sample data: (7)

Outlook	Temperature	Humidity	Play
Sunny	Hot	High	0
Overcast	Hot	High	1
Rain	Mild	Normal	1
Sunny	Cool	Normal	1

- a) Determine the size of the hypothesis space for this problem.
- b) Define the concept of version space in this context.
- c) Explain how version space helps in concept learning.

Module II

- 13 a) Illustrate the architecture of a Multilayer Perceptron (MLP) with an appropriate example. (7)
- b) Explain backward propagation of error? (7)

OR

- 14 Explain SVM, its advantages and list any four real life examples of SVM? (14)

Module III

- 15 Explain the concepts of entropy and information gain. How are these used in building decision trees? Additionally, define gain ratio. (14)

OR

- 16
- | Gender | Age Group | Membership | Product Preference | Purchased |
|--------|-----------|------------|--------------------|-----------|
| Male | Young | Gold | Electronics | Yes |
| Female | Middle | Silver | Clothing | No |
| Female | Young | Gold | Books | Yes |
| Male | Senior | None | Electronics | No |
| Male | Middle | Gold | Clothing | Yes |
| Female | Senior | Silver | Books | No |
| Male | Young | Gold | Clothing | Yes |
| Female | Middle | None | Electronics | No |
- (14)

Male	Senior	Silver	Books	Yes
Female	Young	Gold	Clothing	Yes

- Using the ID3 algorithm, calculate the entropy of the dataset.
- Calculate the information gain for each attribute: Gender, Age Group, Membership, and Product Preference.
- Identify which attribute should be used for the first split in building the decision tree and justify your choice.

Module IV

- 17 Explain the concept of a genetic algorithm and describe its main components. (14)
Write detailed notes on (i) crossover, (ii) mutation, and (iii) selection.

OR

- Define the steps of Principal Components Analysis Algorithm. (7)
- Describe the algorithm of Independent Component Analysis. (7)

Module V

- Use the K-means algorithm to cluster the points B1(3, 9), B2(4, 6), B3(9, 3), B4(6, 7), B5(8, 6), B6(7, 3), B7(2, 1), B8(5, 10) into three clusters. Assume initial cluster centers as B1(3, 9), B4(6, 7), and B7(2, 1). (9)
- Describe EM algorithm for Gaussian mixtures (5)

OR

- Explain Hidden Markov model with real life examples. (7)
- What are Bayesian Networks? Explain using example. What do they predict? (7)
