Seat No. : _____

DO-104

December-2025

5th Year Integrated M.Sc. (CS), Sem.-V

Machine Learning

(Computer Science)

Time : 2:30 Hours] **[Max. Marks : 50**

Instructions : (1) Draw Diagrams neatly wherever necessary.

(2) Make Assumptions wherever necessary.

1. (a) Describe the different types of Machine Learning. Explain Supervised, Unsupervised, and Reinforcement Learning with suitable examples. **5**

1. (b) Explain the general Machine Learning workflow from data collection to model evaluation. **5**

OR

1. (a) How do we choose the right loss function for a given task? Enlist and explain types of loss functions used in linear regression. **5**

1. (b) Explain the concept of the Bias-Variance Trade-off in Machine Learning. Discuss how bias and variance affect model performance with suitable examples. **5**

2. Explain LI and L2 Regularization techniques with proper equations and comparison. **10**

OR

2. Describe the complete performance evaluation process for classification models using suitable examples. **10**

3. (a) Describe Support Vector Machines (SVM) in detail. Explain the concept of margin, hyperplane, and kernel functions. **5**

3. (b) Differentiate between KNN, SVM, and Decision Tree classifiers with suitable examples. **5**

OR

3. Consider the four points :

Q1 = (1, 1), Q2 = (1, 2), Q3 = (4, 4), Q4 = (5, 4). **10**

Run the k-means algorithm with k = 2 and the following initial centroids :

$\mu_1(0) = (1, 1), \mu_2(0) = (5, 4)$

Perform full iterations (assignment step and centroid update) until the centroids do not change. Show all steps and final clusters.

4. Explain the role of distances in clustering. Enlist different metrics used for distance for numeric, non-numeric and mix data. Explain each in detail. **10**

OR

4. Consider the five 2-D points : **10**

P1 = (1, 1), P2 = (2, 1), P3 = (1, 2), P4 = (8, 8), P5 = (9, 7).

(a) Compute the pairwise Euclidean distance matrix (show distances to two decimal places).

(b) Based on the distances, suggest a natural grouping (clusters) and list which points belong to each cluster.

(c) For each proposed cluster, compute the centroid (mean point).

(d) Compute the Sum of Squared Errors (SSE) for the clustering you proposed.

5. Answer the following : (any **10**) **10**

(1) What is Machine Learning ?

(2) Name any one type of Machine Learning.

(3) Which type of learning uses labelled data ?

(4) What is a Dataset ?

(5) Define Classification.

(6) Define Regression.

(7) What is Clustering ?

(8) What is the purpose of a cost function in Linear Regression ?

(9) Name the optimization algorithm used in Gradient Descent.

(10) What is overfitting ?

(11) What is underfitting ?

(12) What does K stand for in K-Nearest Neighbours (KNN) ?

(13) What is the full form of SVM ?

(14) What is PCA used for ?

(15) What is the main goal of Unsupervised Learning ?

———————