



- (d) The point with maximum distance from others
- 1-e. Which hierarchical approach typically produces more accurate hierarchies but is computationally expensive? (CO4 K2) 1
- (a) Agglomerative (AGNES)
- (b) Divisive (DIANA)
- (c) Single-linkage
- (d) Multiple-linkage

2. Attempt all parts:-

- 2.a. Briefly explain the term intelligent agent with an example. (CO1 K2) 2
- 2.b. Differentiate numerical and categorical features with example.(CO2 K2) 2
- 2.c. Briefly explain the term hyperplane, soft margin and hard margin in the context of support vector machine. (CO3 K2) 2
- 2.d. . Give a brief explanation of univariate and multivariate regression. (CO3 K2) 2
- 2.e. State the main purpose of a Gaussian Mixture Model in clustering. (CO4 K2) 2

### **SECTION-B**

15

3. Answer any three of the following:-

- 3-a. Explain the Hill Climbing and Simulated Annealing algorithms. Analyze their working, advantages, and limitations, and compare how each algorithm handles local maxima in optimization problems. Illustrate your answer with a diagram. (CO1 K4) 5
- 3-b. Given the dataset with two numerical features: 5  
 $X = \{2, 4, 6, 8\}$  and  $Y = \{3, 7, 9, 12\}$ ,  
 Using the above dataset, perform principal component analysis by doing the following analytical steps:  
 1. Compute the Mean of each feature.  
 2. Calculate the Covariance matrix.  
 3. Determine the Eigenvalue and Eigenvector of the covariance matrix. (CO2 K4)
- 3-c. Explain the concept of dimensionality reduction. Analyze why reducing the number of dimensions is beneficial when working with high-dimensional datasets. Illustrate your answer with examples. (CO2 K4) 5
- 3.d. Explain L1 and L2 regularization techniques. Analyze their working, differences, and impact on model parameters. Illustrate your answer with examples showing how each method affects model weights. (CO3 K4) 5
- 3.e. Explain the K- means clustering algorithm in detail. Apply the K-means algorithm to the dataset provided and demonstrate the formation of 2 clusters. where  $K = 2$ , initial centroids: **A (1,2)** and **C (4,5)** 5

Point	x	y
A	1	2
B	2	1
C	4	5
D	5	4
E	3	3

**SECTION-C**

4. Answer any five of the following:-

- 4-a. Apply the A\* search algorithm to the graph shown below to find the optimal path from node 'S' (Start) to node 'G' (Goal). Show all intermediate steps including open and closed lists at each stage, and finally, present the path and total cost. 4

Graph:

Edges:

S → A(3), S → B(4)

A → C(5)

B → D(3)

C → G(4)

D → G(3)

Heuristic h(n):

h(S)=6, h(A)=5, h(B)=4, h(C)=3, h(D)=2, h(G)=0. (CO1 K4)

- 4-b. Analyze BFS and DFS with an example graph, highlighting the differences in node traversal and use-cases. (CO1 K4) 4

- 4-c. Discuss how eigenvalues and eigenvectors are computed and used in PCA. Explain how they determine the direction and importance of principal components, and illustrate the process with an example dataset. (CO2 K4) 4

- 4-d. Analyze the role of feature engineering in machine learning and explain how selection, extraction, and transformation affect models, with a real-world example. (CO2 K4) 4

- 4-e. Explain the Support Vector Machine (SVM) algorithm. Describe how it separates classes using a hyperplane, the role of support vectors, and margin maximization. Illustrate your explanation with a simple 2D example. (CO3 K4) 4

- 4-f. Consider the following dataset: 4

Point	X	Y	Class
A	1	2	Red
B	2	3	Red
C	3	3	Blue
D	6	5	Blue
E	7	8	Blue

A new point P(4,4) needs to be classified using KNN with K = 3.

Tasks:

1. Compute the distances from P to all points.
2. Identify the 3 nearest neighbors.
3. Predict the class of P. (CO3 K4)

- 4-g. Apply DBSCAN on the dataset given below using  $\epsilon = 2$  and MinPts = 3. 4

Points: (1,1), (1,2), (2,1), (8,8), (8,9)

Identify core points, border points, and noise points. (CO4 K4)

- 4-h. Explain the hierarchical clustering algorithm, covering both agglomerative and divisive approaches. Illustrate how clusters are formed using a dendrogram. (CO4 4

K4)

REG\_JULY\_DEC\_2025