

**BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI  
(END SEMESTER EXAMINATION)**

**CLASS: B.TECH.  
BRANCH: CSE**

**SEMESTER : VI  
SESSION : SP/2024**

**SUBJECT: AI303 UNSUPERVISED LEARNING**

**TIME: 3 Hours**

**FULL MARKS: 50**

**INSTRUCTIONS:**

1. The question paper contains 5 questions each of 10 marks and total 50 marks.
  2. Attempt all questions.
  3. The missing data, if any, may be assumed suitably.
  4. Before attempting the question paper, be sure that you have got the correct question paper.
  5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.
- 

- |   |     |     |         |
|---|-----|-----|---------|
| Q.1(a) Explain and compare different machine learning approaches from data, computing and application perspective. Give applications for each of them, with suitable example. | [5] | CO1 | BL<br>2 |
| Q.1(b) Give examples of any three types of data used in clustering analysis. Explain the similarity/dissimilarity measures for each of them.                                  | [5] | CO1 | 2       |
| Q.2(a) Discuss the followings using suitable examples:<br>(i) Data visualization techniques.<br>(ii) Divisive v/s Agglomerative hierarchical clustering                       | [5] | CO1 | 3       |
| Q.2(b) Describe the single linkage, complete linkage and average linkage approach using an example. Apply Ward's method of clustering using following data:                   | [5] | CO2 | 3       |

ID	Gender	Age	Salary
1	F	27	19,000
2	M	51	64,000
3	M	52	100,000
4	F	33	55,000
5	M	45	45,000

- |   |     |     |   |
|---|-----|-----|---|
| Q.3(a) Describe the K-Means Algorithm for clustering. Compare and differentiate the MacQueen's Method, Forgy's Method and Jancey's Method of clustering using an example. | [5] | CO3 | 3 |
| Q.3(b) Consider a set of 1-D objects {22,9,12,15,18,27,11,36,10,3,14,32}. Apply the BIRCH algorithm and explain the step-by-step execution of clustering.                 | [5] | CO3 | 4 |
| Q.4(a) Explain the DBSCAN algorithm. Demonstrate how it is used for clustering.   | [5] | CO4 | 2 |
| Q.4(b) Use the following data set to  | [5] | CO4 | 3 |

Buying	Maintenance	Safety	Class
Low	High	Medium	Acceptable
Low	High	High	Very Good
Very High	Medium	Medium	Acceptable
Very High	Very High	Low	Unacceptable
Very High	Very High	Medium	Unacceptable
Low	Low	Medium	Good

- (i) Represent this database as a graph, based upon attributes are weighed nodes. Each tuple in the database is represented with attribute values as a node and edges are to represent connections between the attribute values for the specific tuple.
- (ii) Find the similarity between the second and fifth tuples using both formulas discussed in the section on ROCK.
- (iii) For the above data set, find the common neighbors for the second and fifth tuples where  $\Theta=0.3$

**PTO**

Q.5(a) Discuss various measures of cluster validity using an example.

[5] CO5 3

Q.5(b) Write brief notes on:

[5] CO5 2

(i) CACTUS clustering algorithm

(ii) Statistical test and confusion matrix for cluster validity

::::::01/05/2024::::::M