

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

CLASS: BCA  
BRANCH: BCA

SEMESTER : VI  
SESSION : SP/2025

**SUBJECT: CA353 DATA MINING & WAREHOUSING**

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 briefly about tasks and challenges associated with data mining. [5] CO 1 BL 2
- Q.1(b) Calculate similarity matrix and dissimilarity matrix for all the attributes separately using the following data: [5] CO 1 BL 3

Obj Id	Test 1	Test II	Test III
1	A	Low Priority	45
2	B	Urgent	93
3	B	High Priority	65
4	C	High Priority	74
5	A	Low Priority	23

- Q.2(a) Find frequent itemsets and generate association rules using Apriori algorithm for the following dataset: [5] CO 2 BL 3

TID	Items
T1	I1, I2, I5
T2	I2, I4
T3	I2, I3
T4	I1, I2, I4
T5	I1, I3
T6	I2, I3
T7	I1, I3
T8	I1, I2, I3, I5
T9	I1, I2, I3

[Min. Support count = 2, Min. Confidence = 60%]

- Q.2(b) Discuss Association analysis? Explain the concept of market basket analysis with the help of an example. [5] CO 2 BL 2
- Q.3(a) Explain different types of clustering with examples. Point the key differences between partitioning and hierarchical clustering. [5] CO 3 BL 2
- Q.3(b) Find two clusters of the following data using K-Means clustering algorithm and Euclidean distance: [5] CO 3 BL 3  
 $D = \{(2,3), (3,4), (2,4), (4,5), (7,8), (6,7), (8,9)\}$

PTO

- Q.4(a) Explain decision tree classifier. Discuss the basic algorithm to build a decision tree. [5] 4 2
- Q.4(b) Construct confusion matrix for the following data: [5] 4 3
- A model classifies 2 kinds of fruits: apples and mango. Out of 15 samples, 8 belong to Apples and 7 belong to the Mango class. The actual class for the 15 samples is represented as: [1,1,1,1,1,1,1,1,0,0,0,0,0,0,0]
- The classifier makes the following predictions:
- Out of 8 apples, it will classify 5 correctly as apples and wrongly predict 3 as mangoes.
  - Out of 7 mangoes, it will classify 5 correctly as mangoes and wrongly predicts 2 as apples.
- Prediction = [1,0,0,0,1,1,1,1,0,0,0,0,0,1,1]
- Calculate accuracy, precision, recall and F1-score.
- Q.5(a) Define Data Warehouse. Explain its salient characteristics along with the issues. [5] 5 2
- Describe Day 1 to Day n phenomenon of building a data warehouse.
- Q.5(b) Discuss the two major design issues of the data warehouse: granularity and partitioning of data. [5] 5 3

::::::26/04/2025::::::M