

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

**CLASS: MBA
BRANCH: MBA**

**SEMESTER : III
SESSION : MO/2024**

SUBJECT: MT549R1 DATA MINING

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.
-

- | | | CO | BL |
|--------|--|-----|-----|
| Q.1(a) | Describe major issues in data mining. | [5] | 1 2 |
| Q.1(b) | Explain any one data pre-processing tasks in detail. | [5] | 1 2 |

- | | | | | |
|--------|---|-----|---|---|
| Q.2(a) | Explain the role of data mining for marketing and CRM. | [5] | 2 | 2 |
| Q.2(b) | Describe similarity and dissimilarity matrix. Calculate similarity matrix and dissimilarity matrix for the Grade and Progress attribute as shown in the following data: | [5] | 2 | 3 |

Obj Id	Grade	Progress	Numeric
1	A	Excellent	45
2	B	Fair	22
3	C	Good	64
4	A	Excellent	28

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

Transaction	List of items
T1	I1, I2, I3
T2	I2, I3, I4
T3	I4, I5
T4	I1, I2, I4
T5	I1, I2, I3, I5
T6	I1, I2, I3, I4

[Support threshold = 50%, Confidence = 60%]

- | | | | | |
|--------|---|-----|---|-----|
| Q.3(b) | Explain how partition algorithm overcomes the main limitation of Apriori, with the help of following example: | [5] | 3 | 3,4 |
|--------|---|-----|---|-----|

Transaction	List of items
T1	I1, I2, I4
T2	I2, I4
T3	I3, I4
T4	I1, I2, I5
T5	I1, I4
T6	I1, I3, I5
T7	I1, I2, I3, I5
T8	I1, I2, I3

[Min. support = 2]

- | | | | | |
|--------|---|-----|---|---|
| Q.4(a) | Find three clusters of the following data using K-Means clustering algorithm and Euclidean distance:
A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9) | [5] | 4 | 3 |
| Q.4(b) | Describe perceptron, architecture and working of neural network. Discuss any two types of neural network. | [5] | 4 | 2 |

PTO

- | | | | |
|--------|---|---|---|
| Q.5(a) | Explain definition, process and applications of any one advanced mining technique. [5] | 5 | 2 |
| | How does that technique differ from data mining? | | |
| Q.5(b) | Explain the process of classification with the help of decision tree algorithm. Discuss [5] | 5 | 2 |
| | any two attribute selection measures used by the decision tree algorithms. | | |

.....22/11/2024.....M