

Code No: R20A1206

MALLA REDDY COLLEGE OF ENGINEERING & TECHNOLOGY

(Autonomous Institution – UGC, Govt. of India)

III B.Tech II Semester Regular/Supplementary Examinations, June 2024**Data Warehousing and Data Mining****(IT, CSE-AIML & B.Tech-AIDS)**

Roll No									
----------------	--	--	--	--	--	--	--	--	--

Time: 3 hours**Max. Marks: 70**

Note: This question paper Consists of 5 Sections. Answer **FIVE** Questions, Choosing **ONE** Question from each SECTION and each Question carries 14 marks.

SECTION-I

- | | | | BCLL | CO(s) | Marks |
|---|---|--|-------------|--------------|--------------|
| 1 | A | Explain about Three-tier architecture of a Datawarehouse. | L2 | CO-I | [7M] |
| | B | Compare and contrast operational database systems and data warehouses in terms of their purpose, structure, and usage. | L2 | CO-I | [7M] |

OR

- | | | | | | |
|---|---|---|-----------|-------------|-------------|
| 2 | A | Describe what a fact constellation is and how it differs from a star schema. | L4 | CO-I | [7M] |
| | B | Compare and contrast ROLAP (Relational OLAP), MOLAP (Multidimensional OLAP), and HOLAP (Hybrid OLAP) architectures. | L4 | CO-I | [7M] |

SECTION-II

- | | | | | | |
|---|---|--|-----------|--------------|-------------|
| 3 | A | Describe the main functionalities of data mining. | L2 | CO-II | [7M] |
| | B | Explain about the classification of data mining systems. | L2 | CO-II | [7M] |

OR

- | | | | | | |
|---|---|--|-----------|--------------|-------------|
| 4 | A | Define data reduction and explain its importance in reducing the complexity of large datasets. | L4 | CO-II | [7M] |
| | B | Provide examples of common data mining task primitives and explain how they are used in real-world data mining projects. | L4 | CO-II | [7M] |

SECTION-III

- | | | | | | |
|---|---|--|-----------|---------------|-------------|
| 5 | A | Discuss the Apriori principle and how it is used to generate frequent item sets efficiently. | L3 | CO-III | [7M] |
| | B | Define support and confidence measures in the context of association rules. | L1 | CO-III | [7M] |

OR

- | | | | | | |
|---|---|--|-----------|---------------|-------------|
| 6 | A | Find the frequent itemsets using FP-Growth algorithm. Consider min_support=2,min_confidence=60%. | L3 | CO-III | [7M] |
|---|---|--|-----------|---------------|-------------|

<i>TID</i>	<i>List of item IDs</i>
T100	I1, I2, I5
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3
T600	I2, I3
T700	I1, I3
T800	I1, I2, I3, I5
T900	I1, I2, I3

B Define closed frequent item sets and explain their importance in summarizing association rules. **L2** **CO-III** [7M]

SECTION-IV

7 A Describe the general approaches used to solve a classification problem. **L2** **CO-IV** [7M]

B Provide an overview of various classification techniques, including decision trees, Naive Bayes, Bayesian Belief Networks, and K-Nearest Neighbor classification. **L4** **CO-IV** [7M]

OR

8 A Describe the algorithm for decision tree induction. **L2** **CO-IV** [7M]

B Explain about Naïve Bayesian Classification. **L2** **CO-IV** [7M]

SECTION-V

9 A Describe the major categories of clustering methods. **L2** **CO-V** [7M]

B Discuss the initialization methods, distance measures, and convergence criteria used in the K-Means algorithm. **L4** **CO-V** [7M]

OR

10 A Explain the concepts of agglomerative and divisive hierarchical clustering methods. **L2** **CO-V** [7M]

B Provide examples of popular partitioning methods such as K-Means and PAM. **L3** **CO-V** [7M]
