

**P.E.S. College of Engineering, Mandya - 571 401***(An Autonomous Institution affiliated to VTU, Belagavi)***Second Semester, Master of Computer Applications (MCA)****Semester End Examination; October - 2023****Data Warehousing and Data Mining***Time: 3 hrs**Max. Marks: 100***Course Outcomes***The Students will be able to:**CO1: Understand the basic concepts of Data Warehousing and OLAP Implementation.**CO2: Discuss the basic concepts, techniques, and applications of data mining.**CO3: Illustrate the association rules to extract appropriate pattern in massive data.**CO4: Describe the different classification techniques.**CO5: Illustrate the clustering techniques and outlier analysis in detail for better organization and retrieval of data.***Note: I) Answer any FIVE full questions, selecting ONE full question from each unit.****II) Any THREE units will have internal choice and remaining TWO unit questions are compulsory.****III) Each unit carries 20 marks.**

Q. No.	Questions	Marks	BLs	COs	POs
<b>UNIT - I</b>		<b>20</b>			
1 a.	Define Data Warehousing. Explain the steps for DW implementation.	9	L1	CO1	PO1,2
b.	Explain the FASMI characteristics of OLAP systems.	5	L2	CO1	PO1
c.	List the Codd's OLAP characteristics.	6	L1	CO1	PO1
<b>UNIT - II</b>		<b>20</b>			
2 a.	Define Data Mining. Discuss the different types of dataset with example.	10	L1,2,6	CO2	PO1,2
b.	List the Data Preprocessing steps. Explain any two preprocessing steps with an example.	10	L1,L2	CO2	PO1,2
<b>UNIT - III</b>		<b>20</b>			
3 a.	Illustrate with example, the Apriori algorithm for finding a frequent itemset.	10	L2	CO3	PO1,2,3
b.	Define Maximal, Closed frequent Itemset with an example.	5	L1	CO3	PO1,2
c.	What are all the factors affects to complexity of association rule discovery, explain?	5	L1,L2	CO3	PO1,2
<b>OR</b>					
d.	Explain methods for generating candidate Itemset with example.	8	L2	CO3	PO1,2,3
e.	Explain the different steps involved in subsequent extension of FP tree with example.	12	L3	CO3	PO1,2,3

Contd....2

**UNIT - IV****20**

- |      |  |    |    |     |         |
|------|--|----|----|-----|---------|
| 4 a. | Summarize the decision tree induction algorithm characteristics.   | 10 | L2 | CO4 | PO1,2   |
| b.   | How to build a Rule based classifier using Indirect method for rule extraction? List the characteristics of Rule based classifier. | 10 | L1 | CO4 | PO1,2,3 |

**OR**

- |    |   |    |    |     |           |
|----|---|----|----|-----|-----------|
| d. | Explain the following:                                |    |    |     |           |
|    | i) Web robot detection                                | 10 | L2 | CO4 | PO1,2,3   |
|    | ii) Bayesian classifier                               |    |    |     |           |
| e. | Discuss the design issues of decision tree induction. | 10 | L2 | CO4 | PO1,2,3,5 |

**UNIT - V****20**

- |      |  |    |    |     |         |
|------|--|----|----|-----|---------|
| 5 a. | List and discuss the different types of Data in Clustering.                                | 6  | L2 | CO5 | PO1,2   |
| b.   | List the properties of Computing distance in Clustering.                                   | 4  | L1 | CO5 | PO1,2   |
| c.   | Explain agglomerative cluster algorithm and methods defining proximity between clustering. | 10 | L2 | CO5 | PO1,2,3 |

**OR**

- |    |  |   |    |     |       |
|----|--|---|----|-----|-------|
| d. | Explain Divisive Hierarchical Clustering algorithm with example. | 8 | L2 | CO5 | PO1,2 |
| e. | How to find the validity of cluster analysis methods.            | 8 | L1 | CO5 | PO1,2 |
| f. | List the quality of Cluster analysis method.                     | 4 | L1 | CO5 | PO1,2 |

\* \* \* \*