

Basavarajeswari Group of Institutions

BALLARI INSTITUTE OF TECHNOLOGY & MANAGEMENT

(Autonomous Institute under Visvesvaraya Technological University, Belagavi)

2024 SCHEME

USN

--	--	--	--	--	--	--	--	--

Course Code

M	M	C	B	3	1	1	A
---	---	---	---	---	---	---	---

Third Semester MCA Degree Examinations, January 2026

DATA MINING & DATA WAREHOUSING

Duration: 3 hrs

Max. Marks: 100

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. Missing data, if any, may be suitably assumed

<u>Q. No</u>	<u>Question</u>	<u>Marks</u>	<u>(RBTl:CO: PI)</u>
<u>Module-1</u>			
1.	a. Explain the schemas of multidimensional data models.	10	(2:1: 1.2.1)
	b. Describe the OLAP Server Architectures: ROLAP, MOLAP, and HOLAP. Highlight their advantages and disadvantages?	10	(2:1: 1.2.2)
(OR)			
2.	a. Explain the characteristics of dimension tables. How do they differ from fact tables?	10	(2:2: 3.2.1)
	b. Explain how OLTP systems differ from OLAP systems in terms of functionality with example.	10	(2:2: 2.2.2)
<u>Module-2</u>			
3.	a. Define data mining and explain the steps involved in the Knowledge Discovery in Databases (KDD) process.	10	(2:2: 1.2.2)
	b. Describe various measures of similarity and dissimilarity used in data mining?	10	(2:2: 2.2.1)
(OR)			
4.	a. Explain data pre-processing. Write a short note on aggregation, sampling and dimensionality reduction with example.	10	(2:2: 2.3.1)
	b. Discuss the major challenges faced in data mining. Provide suitable examples.	10	(2:2: 1.2.3)
<u>Module-3</u>			
5.	a. Generate frequent item sets for the following transactional database using the FP-Growth algorithm with a minimum support count = 2.	10	(3:3: 1.2.1)

Transaction ID	Items
T1	{A, B, D, E}
T2	{B, C, E}
T3	{A, B, C, E}
T4	{A, C, D}
T5	{B, C, D, E}

- b. Consider the transaction data set. 10 (3:3: 1.2.2)

IID	1	2	3	4	5	6
Items	{a, b}	{b, c, d}	{a, c, d, e}	{a, d, e}	{a, b, c}	{a, b, c, d}

7	8	9	10
{a}	{a, b, c}	{a, b, d}	{b, c, e}

Construct the FP tree by showing the trees separately after reading each transaction. Also find the frequent item set ordered by their corresponding suffixes.

(OR)

6. a. Illustrate frequent item set generation of Apriori algorithm with suitable example. 10 (2:3: 2.1.3)
- b. Briefly explain alternative methods for generating frequent item sets. 10 (2:3: 2.1.1)

Module-4

7. a. Write Hunt's algorithm and apply it on the following dataset to construct a decision tree. 10 (2:4: 1.1.2)

Tid	Home	Marital	Annual	Defaulted
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

- b. Explain the different ways of implementing Rule-Ordering schemes. 10 (2:4: 3.2.1)

(OR)

8. a. How does decision tree induction algorithm works? Explain with an example. 10 (2:4: 2.2.1)
- b. Discuss the concept of Bayesian classifiers with suitable example. 10 (2:4: 2.2.2)

Module-5

9. a. Compare K-Means algorithm and DBSCAN algorithms. 10 (2:5: 2.1.3)
- b. What is K-Means clustering? Explain its algorithmic procedure and discuss the drawbacks of the K-Means approach. 10 (2:5: 1.2.1)

(OR)

- 10 a. Explain basic agglomerative hierarchical clustering algorithm. Give its strength and weakness. 10 (2:5: 1.2.1)
- b. Explain the concept of cluster analysis. Discuss various clustering techniques and cluster types. 10 (2:5: 1.1.3)

** ** *