



PAPER ID-311663

Roll No:

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

**MTECH**  
**(SEM I) THEORY EXAMINATION 2023-24**  
**DATA WAREHOUSING & DATA MINING**

**TIME: 3HRS**

**M.MARKS: 70**

**Note: 1.** Attempt all Sections. If require any missing data; then choose suitably.

**SECTION A**

**1. Attempt all questions in brief. 2 x 7 = 14**

a.	Define Data Warehousing.
b.	What factors lead to the mining of data?
c.	Are all the patterns interesting?
d.	Differentiate between binary symmetric attributes and asymmetric attributes.
e.	Find the mode of the following dataset: 12,13,34,32,21,29,40,11,39,23. What is the advantage of mode over mean and median?
f.	Given two objects represented by the tuples (22, 2, 45, 10) and (20, 10, 26, 2): Compute the Manhattan distance between these two objects.
g.	What do you mean by Temporal Mining?

**SECTION B**

**2. Attempt any three of the following: 7 x 3 = 21**

a.	Write short notes on: i. Steps of Knowledge Discovery in data ii. Handling Missing values in database.
b.	What is Data Mart? What are the different types of Data Marts?
c.	Draw the box-and-whisker plot of the following dataset: 4.3, 5.1, 3.9, 4.5, 4.4, 4.9,5.0, 4.7, 4.1, 4.6, 4.4, 4.3, 4.8, 4.4, 4.2, 4.5, 4.4.
d.	Cluster the following dataset with points (2,4), (6,8), (1,2), (4,5), (3,5) into two clusters using K-Means algorithm (using Euclidean distance algorithm only).
e.	Explain ROLAP, MOLAP and HOLAP in detail.

**SECTION C**

**3. Attempt any one part of the following: 7 x 1 = 7**

a.	How mapping a 2D table into multidimensional data model? Explain with suitable example.
b.	Write short notes on: i. Data Characterization and Data Discrimination ii. Binning

**4. Attempt any one part of the following: 7 x 1 = 7**

a.	Differentiate between: (i) Min-Max and Z-score Normalization with examples (ii) Binary data variables and Nominal data variables with examples
b.	Explain the major components of Data Mining Architecture.

**5. Attempt any one part of the following: 7 x 1 = 7**

a.	Discuss Decision tree-based classifiers in detail.
----	--



PAPER ID-311663

Roll No:

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

**MTECH**  
**(SEM I) THEORY EXAMINATION 2023-24**  
**DATA WAREHOUSING & DATA MINING**

TIME: 3HRS

M.MARKS: 70

b.	Classify the tuple X = (age = youth, income = medium, student= yes, credit rating = fair) using Bayes Theorem.					
	<i>RID</i>	<i>age</i>	<i>income</i>	<i>student</i>	<i>credit_rating</i>	<i>Class: buys_computer</i>
	1	youth	high	no	fair	no
	2	youth	high	no	excellent	no
	3	middle_aged	high	no	fair	yes
	4	senior	medium	no	fair	yes
	5	senior	low	yes	fair	yes
	6	senior	low	yes	excellent	no

6. Attempt any one part of the following: 7 x 1 = 7

a.	Explain various types of clustering methods. Discuss any one partitioning clustering algorithm.
b.	Discuss DBSCAN clustering algorithm with suitable example.

7. Attempt any one part of the following: 7 x 1 = 7

a.	Differentiate between (a) OLAP and OLTP in detail. (b) Slice and Dice operations with an example.
b.	Define Spatial Data. How mining of spatial data is done?