Roll No.											

PRESIDENCY UNIVERSITY

BENGALURU

End - Term Examinations – December 2025	
Date: 08-12- 2025	**Time:** 01:00pm – 04:00pm

School: SOCSE,SOIS	**Program:** B.TECH	
Course Code :CSD3408	**Course Name:** DATA MINING AND WAREHOUSING	
Semester: V	**Max Marks**:100	**Weightage**: 50%

CO - Levels	CO1	CO2	CO3	CO4	CO5
Marks	26	24	26	24	

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. **10Q x 2M=20M**

1.	Define outlier analysis.	2 Marks	L1	CO1
2.	What is concept hierarchy in multidimensional modeling?	2 Marks	L2	CO1
3.	What is data cleaning in the context of KDD?	2 Marks	L2	CO1
4.	List any two characteristics of OLTP systems.	2 Marks	L1	CO2
5.	Give an example of a data transformation applied in ETL.	2 Marks	L2	CO2
6.	What is the main objective of partitioning-based clustering?	2 Marks	L2	CO3
7.	What is the main purpose of the CLARA algorithm in clustering?	2 Marks	L2	CO3
8.	Define clustering in data mining.	2 Marks	L1	CO3
9.	Identify a challenge in mining unstructured web content.	2 Marks	L2	CO4
10.	What is tokenization in text mining?	2 Marks	L1	CO4

Part B

Answer the Questions. **Total Marks 80M**

11.	**a.**	Explain the architecture and key steps involved in Knowledge Discovery in Databases (KDD).	**10 Marks**	**L2**	**CO1**
	b.	Discuss different types of data and explain how data type affects the choice of mining technique.	**10 Marks**	**L1**	**CO1**
Or					
12.	**a.**	Explain the steps of data preprocessing and discuss their significance in preparing data for mining.	**10 Marks**	**L2**	**CO1**
	b.	Explain the main functionalities of data mining and illustrate with real-world examples.	**10 Marks**	**L2**	**CO1**

13.	**a.**	Explain the architecture of a Data Warehouse with a neat diagram, highlighting the functions of each component.	**10 Marks**	**L2**	**CO2**
	b.	Explain the different types of Data Warehouse schemas with diagrams and describe the purpose of each.	**10 Marks**	**L2**	**CO2**
Or					
14.	**a.**	Compare and contrast OLAP operations with examples and diagrams.	**10 Marks**	**L5**	**CO2**
	b.	Explain the key building blocks of a Data Warehouse with a neat diagram and describe the function of each block.	**10 Marks**	**L2**	**CO2**

15.	**a.**	Explain the different clustering paradigms with suitable examples.	**10 Marks**	**L2**	**CO3**
	b.	Explain the K-means algorithm with a step-by-step example. Discuss its advantages and limitations.	**10 Marks**	**L2**	**CO3**
Or					
16.	**a.**	Describe the phases of the BIRCH algorithm and discuss its advantages for large datasets.	**10 Marks**	**L1**	**CO3**
	b.	Explain the DBSCAN algorithm with an example	**10 Marks**	**L2**	**CO3**

17.	**a.**	Elaborate the components of Web Content Mining, Web Structure Mining, and Web Usage Mining. Compare their techniques and applications in real-world scenarios.	**10 Marks**	**L4**	**CO4**
	b.	Explain the knowledge discovery process in Visual Data Mining with suitable examples.	**10 Marks**	**L2**	**CO4**
Or					
18.	**a.**	Discuss the steps involved in Text Mining from preprocessing to knowledge discovery. Explain how text clustering helps in organizing large textual repositories.	**10 Marks**	**L1**	**CO4**
	b.	Evaluate different visualization techniques used in data mining	**10 Marks**	**L5**	**CO4**