Roll No												

PRESIDENCY UNIVERSITY
BENGALURU
School of Computer Science Engineering & Information Science

Make Up Examinations – December 2025

Semester: MK
Course Code: CSE2021
Course Name: Data Mining
Department: CSE

Date: 26-12-2025
Time: 1.00pm to 04.00pm
Max Marks: 100
Weightage: 50%

Instructions:

(i) Read the all questions carefully and answer accordingly.
(ii) Do not write any matter on the question paper other than roll number.

PART A

Answer ALL Questions. Each question carries 2 marks. **(5Qx 2M= 10M)**

Q.No	Questions	Marks	CO	RBT
1.	Define data mining and briefly explain its importance.	2	CO1	Remember
2.	What are the key stages in the data mining process?	2	CO1	Remember
3.	What are the main types of data encountered in data mining?	2	CO1	Remember
4.	List four common data quality issues.	2	CO1	Remember
5.	What is the Apriori principle?	2	CO3	Remember
6.	What is the main advantage of the FP-Growth algorithm over Apriori?	2	CO3	Remember
7.	Define over fitting in the context of classification models.	2	CO4	Remember
8.	Explain the term "Knowledge Discovery in Databases (KDD)".	2	CO1	Understand
9.	What are the major issues in data mining?	2	CO1	Remember
10.	What are the two types of data?	2	CO1	Remember

PART B

Answer any FIVE Questions. Each question carries 10 marks. **(5Q X 10M=50M)**

Q.No	Questions	Marks	CO	RBT
11	Explain the KDD (Knowledge Discovery in Databases) process and its relationship to data mining.	**10**	CO 1	Understand
12	Describe the various tasks involved in data mining. How do these tasks contribute to achieving data mining goals?	**10**	CO 1	Understand
13	Compare and contrast data reduction techniques. How do they impact the quality of data mining results?	**10**	CO 2	**Analyze**
14	Describe the process of handling missing values in a dataset. What are the pros and cons of different approaches?	**10**	CO 2	Understand
15	Describe the Apriori algorithm for finding frequent item sets. What are its limitations?	**10**	CO 3	Understand
16	Compare and contrast the Apriori and FP-Growth algorithms in terms of efficiency and scalability.	**10**	CO 3	Analyze
17	Describe the process of decision tree induction. What are the key considerations in choosing the best attribute for splitting?	**10**	CO 4	Understand

PART C

Answer any TWO Questions. Each question carries 15 marks. **(2Q X 15M=30M)**

Q.No	Questions	Marks	CO	RBT
18	Evaluate the various data preprocessing techniques. How do they contribute to improving the efficiency and effectiveness of data mining algorithms?	**15**	CO 2	Evaluate
19	Analyze the FP-Growth algorithm in detail. Explain how it constructs the FP-tree and uses it to mine frequent patterns. Provide an example to illustrate its workings.	**15**	CO 3	**Analyze**
20	Evaluate the applications of frequent pattern mining in real-world scenarios. Discuss three different domains where it has been successfully applied, and analyze the challenges and limitations in each case.	**15**	CO 4	Evaluate

****