CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS
IV–B.TECH–I–Semester End Examinations (Supply) - December- 2025
DATA MINING
(CSE)

[Time: 3 Hours]   **[Max. Marks: 70]**

Note: This question paper contains two parts A and B.

Part A is compulsory which carries 20 marks. Answer all questions in Part A.

Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART-A   (20 Marks)

1. a)	Define data mining.	[2M]
b)	What is missing data?	[2M]
c)	Define frequent patterns.	[2M]
d)	What is Support?	[2M]
e)	State classification.	[2M]
f)	Describe tree pruning.	[2M]
g)	What are the applications of clustering?	[2M]
h)	State K-Means method.	[2M]
i)	Define Web mining.	[2M]
j)	Define data stream mining.	[2M]

PART-B   (50 Marks)

2. List and describe the five primitives for specifying the data mining tasks. [10M]

OR

3. What Data preprocessing? Explain various forms of data preprocessing. [10M]

4. Explain the APRIORI algorithm with an example. [10M]

OR

5. Discuss in brief about,
   (a) Maximum frequent itemset [5M]
   (b) Closed frequent itemset [5M]

6. Discuss briefly about Decision tree induction algorithm. [10M]

OR

7. Describe Bayesian Belief Network with an example. [10M]

8. Explain in briefly about PAM (partition around medoids) algorithm. [10M]

OR

9. Explain the concept of hierarchical clustering. Discuss the agglomerative and divisive approaches. [10M]

10. Explain text clustering in detail. [10M]

OR

11. Discuss the process of mining the World Wide Web. Explain web structure mining. [10M]

****