

**CMR ENGINEERING COLLEGE: : HYDERABAD  
UGC AUTONOMOUS**

**IV-B.TECH-I-Semester End Examinations (Regular) - November- 2024  
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) Name two strategies to handle missing data. [2M]
- c) What is the association rule problem? [2M]
- d) How does FP-Growth address the limitations of the APRIORI algorithm? [2M]
- e) What is the difference between supervised and unsupervised classification? [2M]
- f) What is the Naive Bayer classifier? [2M]
- g) What is the main objective of Clustering? [2M]
- h) What is outlier detection in the context of clustering? [2M]
- i) What is web mining? [2M]
- j) What is the hierarchy of categories in text mining? [2M]

**PART-B**

**(50 Marks)**

2. Describe the process of KDD and its stages. [10M]
- OR**
3. Demonstrate various similarity and dissimilarity measures and explain their application in clustering. [10M]
  4. Explain the process of Frequent item set generation and its role in association rule mining. [10M]
- OR**
5. Illustrate the working of the APRIORI algorithm step by step with an example. [10M]
  6. Explain the process of decision tree construction, including the formation of nodes and branches. [10M]
- OR**
7. Explain the K-Nearest Neighbor algorithm in detail with an example. [10M]
  8. Discuss the types of clustering and the criteria used to evaluate the effectiveness of clustering results. [10M]
- OR**
9. Explain the basic steps in agglomerative hierarchical clustering algorithm with examples. [10M]
  10. Explain the concept of web and text mining and their significance in data analysis. [10M]
- OR**
11. Demonstrate episode rule discovery and its application in analyzing sequential patterns in text. [10M]

\*\*\*\*\*