

Code No.: R22CS504PC

R22

H.T.No.

8

R

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

**III–B.TECH–I–Semester End Examinations (Regular) - December- 2025**

**DATA MINING**

**(CSE)**

**[Time: 3 Hours]**

**[Max. Marks: 60]**

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

Part A is compulsory which carries 10 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.

**PART-A**

**(10 Marks)**

1. a) What is Discretization? [1M]
- b) What is data cleaning? [1M]
- c) Define frequent item set. [1M]
- d) Define Association Rule. [1M]
- e) What is a Decision Tree? [1M]
- f) Define Classification. [1M]
- g) Differentiate between Classification and Clustering. [1M]
- h) What is Outlier Detection? [1M]
- i) List any two types of Web Mining. [1M]
- j) What is Episode Rule Discovery in Text Mining? [1M]

**PART-B**

**(50 Marks)**

2. Explain the KDD process with a neat diagram. [10M]
- OR**
3. Explain Data Preprocessing techniques with suitable examples. [10M]
4. Explain the APRIORI algorithm with an example. [10M]
- OR**
5. Explain in detail about Closed Frequent Item set With an example. [10M]
6. Compare and Contrast KNN and Decision Tree classifiers. [10M]
- OR**
7. Discuss various attribute selection measures used in Decision Tree Construction. [10M]
8. Explain the K-Means clustering algorithm with a suitable example. [10M]
- OR**
9. Explain the steps involved in Agglomerative Hierarchical Clustering. [10M]
10. Explain the architecture and process of Web Mining with suitable examples. [10M]
- OR**
11. Illustrate the challenges and solutions in handling Unstructured Text during Text Mining. [10M]

\*\*\*\*\*