

Code No.: IT621PE

R20

H.T.No.

8 R

CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS
III-B.TECH-II-Semester End Examinations (Regular) - June- 2024
DATA MINING
(IT)

[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) What is data preprocessing? [2M]
- b) List out different dimensionality reduction techniques? [2M]
- c) Define support and confidence in Association rule mining. [2M]
- d) What is a closed item set? [2M]
- e) List out the characteristics of KNN classifier? [2M]
- f) What are the four values of a confusion matrix? [2M]
- g) What is the role of dendrogram in clustering? [2M]
- h) Discuss the two approaches to improve quality of hierarchical clustering. [2M]
- i) Give the taxonomy of web mining? [2M]
- j) Illustrate the tasks of text mining? [2M]

PART-B

(50 Marks)

2. Define data cleaning? Express the different techniques for handling missing values? [10M]
- OR**
3. Briefly describe the four stages of knowledge discovery (KDD)? [10M]
4. Explain the APRIORI algorithm with an example? [10M]
- OR**
5. Describe FP-Growth algorithm with an example? [10M]
6. Discuss the methods for expressing attribute test conditions? [10M]
- OR**
7. Describe Bayesian Belief Network with an example [10M]
8. a) Explain about the agglomerative and divisive hierarchical methods? [05M]
- b) Discuss about Outlier Detection? [05M]
- OR**
9. Contrast k-means clustering with PAM clustering approach? [10M]
10. Explain about hierarchy of categories in text mining? [10M]
- OR**
11. Discuss various kinds of patterns to be mined from web/server logs in web usage mining. [10M]
