CMR ENGINEERING COLLEGE: : HYDERABAD
UGC AUTONOMOUS
III–B.TECH–II–Semester End Examinations (Supply) - January- 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 Cleaning?	[2M]
b)	What are the challenges of KDD?	[2M]
c)	Differentiate between closed item set, closed frequent item set.	[2M]
d)	Define maximal frequent item set.	[2M]
e)	Mention the characteristics of K-nearest neighbor classification.	[2M]
f)	Write the approaches to solving a classification problem.	[2M]
g)	Define hierarchical clustering.	[2M]
h)	Define Clustering.	[2M]
i)	Explain the Purpose of Web Content Mining	[2M]
j)	Mention three main areas of web Mining	[2M]

PART-B   (50 Marks)

2. Draw and Explain the typical architecture of a data mining system.   [10M]

OR

3. Discuss data mining as a step-in knowledge discovery (KDD) process and various   [10M]
    challenges associated.

4. Explain Association Rule mining and explain Apriori algorithm.   [10M]

OR

5. How can we mine closed frequent item sets? Explain.   [10M]

6. Explain Naïve Bayes classification with a suitable example.   [10M]

OR

7. Explain with neat sketch classification process.   [10M]

8. Explain k-means clustering technique with an example.   [10M]

OR

9. What is Cluster Analysis? What are some typical applications of Clustering? [10M]

10. Elaborate Text Clustering with an example?   [10M]

OR

11. What is Web structure mining? How does it differ from Web content mining?   [10M]

****