

Code No.: AD514PE

R20

H.T.No.

8

R

CMR ENGINEERING COLLEGE : HYDERABAD
UGC AUTONOMOUS
III-B.TECH-I--Semester End Examinations (Supply) - June- 2024
DATA WAREHOUSE & DATA MINING
(AI&DS)

[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.

PART-A

(20 Marks)

1. a) Define data mining. [2M]
- b) Classify the data mining systems. [2M]
- c) State the necessity of a data warehouse. [2M]
- d) List out the OLAP operations. [2M]
- e) What are lazy learners in classification? [2M]
- f) Name the attribute selection measures. [2M]
- g) Differentiate between clustering and classification. [2M]
- h) Summarize the requirements for cluster analysis. [2M]
- i) Mention the features of streaming data. [2M]
- j) Where do you use text mining? [2M]

PART-B

(50 Marks)

2. Elaborate the functionality of data mining task primitives with suitable examples. [10M]
- OR**
- 3.a) Inspect the techniques used for data cleaning. [5M]
 - b) Analyze the concept of discretization and concept hierarchy generation. [5M]
4. Outline the schemas for multidimensional data models. [10M]
- OR**
- 5.a) Draw and explain the architecture of a data warehouse. [5M]
 - b) Differentiate between operational database systems and data warehouses. [5M]
6. Write and explain the basic algorithm for inducing a decision tree from training tuples. [10M]
- OR**
7. Apply the Naive Bayes classifier for any five instances from the below dataset. [10M]

Day	Weather	Temperature	Humidity (%)	Wind	Play Tennis
D1	Sunny	Hot	85	Weak	No
D2	Sunny	Hot	90	Strong	No
D3	Overcast	Hot	84	Weak	Yes
D4	Rainy	Warm	96	Weak	Yes
D5	Rainy	Cool	80	Weak	Yes
D6	Rainy	Cool	70	Strong	No
D7	Overcast	Cool	70	Strong	Yes
D8	Sunny	Warm	95	Weak	No
D9	Sunny	Cool	70	Weak	Yes
D10	Rainy	Warm	78	Weak	Yes
D11	Sunny	Warm	70	Strong	Yes
D12	Overcast	Warm	90	Strong	Yes
D13	Overcast	Hot	75	Weak	Yes
D14	Rainy	Warm	88	Strong	No

8. How does the k-means algorithm work? Interpret the approach of k-means partitioning algorithm. [10M]

OR

9. Define outlier. Elucidate the available methods for implementation of outlier analysis. [10M]

10. Examine the process of mining the sequential patterns in transactional databases. [10M]

OR

11. Demonstrate the mechanism for mining the time-series data. [10M]
