(NO. OF PAGES: 2)

(FOR THE CANDIDATES ADMITTED DURING THE ACADEMIC YEAR 2024 - 2026 ONLY)

24PCS206

REG.NO.

N.G.M.COLLEGE (AUTONOMOUS) : POLLACHI

END-OF-SEMESTER EXAMINATIONS : May 2025

M.Sc. Computer Science(SF) **MAXIMUM MARKS: 75**

SEMESTER: II **TIME : 3 HOURS**

24PCS206-DATA MINING USING R TOOL

SECTION - A (10 X 1 = 10 MARKS)

ANSWER THE FOLLOWING QUESTIONS.
MULTIPLE CHOICE QUESTIONS.

1) Which of the following is not an issue in Data Mining? **(K1)**
a) High dimensionality b) Shortage of data c) Overfitting d) Outliers

2) Which of the following is not a basic Data Mining task? **(K1)**
a) Classification b) Prediction c) Spooling d) Clustering

3) The basic building block of a data set is _____**(K1)**
a) Data mining b) Data object c) Cluster d) Association rules

4) Which of the following is not important in determining data quality? **(K1)**
a) Accuracy b) Consistency c) Completeness d) Database

5) Data can be updated in _____environment. **(K1)**
a) data warehouse. b) data mining. c) operational. d) informational

ANSWER THE FOLLOWING IN ONE (OR) TWO SENTENCES

6) Write a note on Data Integration. **(K2)**
7) Define: "Tree Pruning". **(K2)**
8) What is cluster analysis? **(K2)**
9) Comment on Data Cube. **(K2)**
10) Explain the R application in governmental use. **(K2)**

SECTION – B (5 X 5 = 25 MARKS)
ANSWER EITHER (a) OR (b) IN EACH OF THE FOLLOWING QUESTIONS.

11) a) Distinguish between the Relational Databases and Transactional Databases. **(K3)**
(OR)
b) Examine the major issues in Data Mining. **(K3)**
12) a) Discover the generating association rules from Frequent Itemsets. **(K3)**
(OR)
b) Evaluate the Attribute selection measures in Decision tree induction. **(K3)**

Ethical paper

(CONT…2)

13) a) Summarize the classical partitioning methods in *k*-Means and *k*-Medoids. **(K3)**

(OR)

b) Estimate the requirements of STatistical INformation Grid (STING). **(K3)**

14) a) Show the typical OLAP operations in Data Warehousing. **(K3)**

(OR)

b) Outline the differences between MOLAP and HOLAP. **(K3)**

15) a) Illustrate the recent trends in Data Mining. **(K3)**

(OR)

b) Describe the R application in E-commerce. **(K3)**

Section-C (5 X 8 = 40 MARKS)

ANSWER EITHER (a) OR (b) IN EACH OF THE FOLLOWING QUESTIONS.

(K4 (Or) K5)

16) a) Compare the Cluster analysis and Outlier analysis. **(K4)**

(OR)

b) Analyze the noisy data in Data Cleaning. **(K5)**

17) a) Demonstrate the Apriori algorithm to finding Frequent Itemsets using candidate generation. **(K5)**

(OR)

b) Assess the implementation of Naïve Bayesian classification. **(K4)**

18) a) Identify the Agglomerative and Divisive Hierarchical Clustering. **(K5)**

(OR)

b) Elucidate the Hierarchical Clustering Algorithm for categorical attributes using ROCK. **(K4)**

19) a) Determine the Business Analysis Framework for Data Warehouse Design. **(K4)**

(OR)

b) Elaborate the efficient processing of OLAP queries. **(K5)**

20) a) Integrate the R application in Social media. **(K4)**

(OR)

b) Enumerate the R application in Banking system. **(K5)**

****

Ethical paper