18CMP03

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc DEGREE EXAMINATION MAY 2022
(First Semester)

Branch – **COMPUTER SCIENCE**

DATA MINING AND ITS APPLICATIONS

Time: Three Hours Maximum: 50 Marks

SECTION-A (5 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** marks (5 x 1 = 5)

1. Which of the following also used as the fifth step in the knowledge discovery process?
   a. Data transformation b. Data mining
   c. Data selection d. Data cleaning
2. Metadata are created for the data names and definitions of the given _____.
   a. warehouse b. data warehouse
   c. data d. data about data
3. Find the CART _____.
   a. Classification and Regression Trees b. Class and Regression Trees
   c. Cube and Regression Trees d. Classification Array Regression Trees
4. The agglomerative approach, also called the_____, starts with each object forming a separate group.
   a. bottom-up approach b. Top-down approach
   c. middle level approach d. hierarchical approach
5. Finger vein biometrics, also called vein matching or vascular technology, is a technique for biometric _____that analyzes the patterns of blood vessels visible from the surface of the skin of fingers.
   a. application b. domain
   c. authentication d. security

SECTION - B (15 Marks)
Answer **ALL** Questions
ALL Questions Carry **EQUAL** Marks (5 x 3 = 15)

6 a What are major tasks in data preprocessing.
OR
b Explain the data cleaning.

7 a What is a data cube?
OR
b Explain the role of concept hierarchies

8 a Write a short notes on decision tree induction.
OR
b Explain the case - based reasoning.

9 a What is cluster analysis?
OR
b Find the quality and scalability of CLARA?

10 a Write short notes on DNA.
OR
b Explain the data mining in biometrics with examples.

Cont...

SECTION -C (30 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** Marks (5 x 6 = 30)

11 a Discuss about the steps in the process of knowledge discovery in data mining.

OR

b What are major issues in data mining?

12 a Differences between operational database systems and data warehouses.

OR

b Explain the three tiger data warehouse architecture with example.

13 a Explain the Apriori algorithm with example.

OR

b Explain the *k*-Nearest-Neighbor classifiers in data mining.

14 a What are the requirements for cluster analysis?

OR

b Explain the hierarchical methods in data mining.

15 a Discuss the data mining for financial data analysis.

OR

b Explain the data mining for telecommunication industry.

Z-Z-Z END