PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BVoc DEGREE EXAMINATION DECEMBER 2023
(Fifth Semester)

Branch – **NETWORKING AND MOBILE APPLICATIONS**

DISCIPLINE SPECIFIC ELECTIVE – I DATA MINING

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 first step in the knowledge discovery process?
    (i) Data selection
    (ii) Missing data
    (iii) Data cleaning
    (iv) Data deletion

2 Multiple data sources may be combined is called as _____.
    (i) Data Reduction
    (ii) Data filling
    (iii) Data Integration
    (iv) Data Transformation

3 Why is the snowflake schema applied?
    (i) Transformation
    (ii) Aggregation
    (iii) Normalization
    (iv) Generalization

4. What is supervised learning uses _____.
    (i) labeled data
    (ii) unlabeled data
    (iii) specific data
    (iv) All of these

5. Which of the following is required by K-means clustering?
    (i) Defined distance metric
    (ii) Number of clusters
    (iii) Initial guess as to cluster centroid
    (iv) All of the above

SECTION - B (15 Marks)
Answer **ALL** Questions
ALL Questions Carry **EQUAL** Marks

(5 x 3 = 15)

6   a. Bring out the steps for knowledge discovery process
        OR
    b. Describe three challenges to data mining regarding user interaction issues.

7   a Discuss issues to consider during data integration.
        OR
    B Explain about attribute subset selection.

8   a Distinguish data warehouse from other data repository.
        OR
    b How measures can be categorized?

9   a Explain about how to generate Association rules from frequent item sets.
        OR
    b Outline about Learning and classification process.

10 a Explain about k- means cluster algorithm.
OR
   b Discuss about Mining Sequence Data.

SECTION -C (30 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** Marks   (5 x 6 = 30)

11 a   Outline the major challenges to data mining in diversity of database types .
OR
   b   Explain the frame work of data Warehouse in detail.

12 a   How to fill the missing values for attribute?
OR
   b   Discuss about Data reduction strategies.

13 a   Classify the three data warehouse models from the architecture point of view.
OR
   b   Explain about the OLAP operations.

   a   Outline the major steps of decision tree classification.
OR
   b   Discuss the Gain ratio in classification.

15 a   Summarize the Requirements for Cluster Analysis.
OR
   b   Explain about graph and networks in mining.

Z-Z-Z   END