PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc(SS) DEGREE EXAMINATION MAY 2024
(Fifth Semester)

Branch – **SOFTWARE SYSTEMS (five year integrated)**

DISCIPLINE SPECIFIC ELECTIVE-I: DATA MINING

Time : Three Hours Maximum: 75 Marks

Section –A (10 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** Marks (10 x 1 = 10)

1. Identify the term used to define the task of inferring a model from labeled training data.
   a) Supervised learning b) Unsupervised Learning
   c) Reinforcement Learning d) E learning

2. Under which of the following do self-organizing maps lie.
   a) Supervised Learning b) Unsupervised Learning
   c) Reinforcement Learning d) Reimbursement

3. Identify the incorrect option among the following which is not involved in data mining.
   a) Data exploration b) Knowledge extraction
   c) Data Transformation d) Data archaeology

4. Identify the correct options which are considered before investing in data mining.
   a) Vendor consideration b) Functionality
   c) Compatibility d) All of the above

5. What does OLTP stand for?
   a) Offline transaction processing b) Online Transaction Processing
   c) Outline traffic processing d) Online Text process

6. Identify the correct option which defines Data mart.
   a) a sub group of data ware house b) Another type of data warehouse
   c) Not related to data warehouse d) None

7. Where is data warehousing used?
   a) Transaction system b) Logical system
   c) Decision support system d) detection

8. Identify the options below that a data warehouse can include.
   a) Database table b) Online data
   c) Flat files d) Menus

9. ETL stands for _____
   a) Effect, transfer, and load b) Explain, transfer and load
   c) Extract, transfer, and load d) Extract, transform and load

10. Identify among the following for which system of data warehousing is mostly used.
    a) Data mining and data storage b) Data integration and data storage
    c) Reporting and data analysis d) Data cleaning and data storage

Section –B (25 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** Marks (5 x 5 = 25)

11. a) Summarize the major issues faced by data mining systems.
    (OR)
    b) Analyze the classifications of Data systems.

Cont...

12. a) Discuss on Data preprocessing.
(OR)
b) Explain about Data Reduction.

13. a) Discuss on Apriori Algorithm.
(OR)
b) Explain Principal component Analysis.

14. a) Explain Multi dimensional association rules.
(OR)
b) Illustrate the Mining quantitative association rules.

15. a) Categorize Various kinds of Association rules with examples.
(OR)
b) Write a short note on Applications of Data Mining.

SECTION -C (40 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** Marks (5 x 8 = 40)

Q. No. 16 is Compulsory

16. Discuss on Data Mining functionalities.

17. a) Write a short notes on Data Aggregation with ex's.
(OR)
b) Illustrate the concepts of Mining frequent Patterns.

18. a) Analyze the Methods of Fp tree.
(OR)
b) Discuss on Ensemble Pruning with appropriate ex.

19. a) Explain about Sequence data in Data Mining.
(OR)
b) Write down the Major applications of Clustering methods with ex's.

20. a) Elucidate the concepts of Graph Mining.
(OR)
b) Outline the concepts of Web Mining with diagrammatic representation.

Z-Z-Z END