



B.Tech V Semester Regular/Supplementary Examinations, December 2021

DATA WAREHOUSING & DATA MINING
(Information Technology)

Maximum Marks: 70

Date: 05.01.2022 Duration: 3 hours

- Note:
1. This question paper contains two parts A and B.
 2. Part A is compulsory which carries 20 marks. Answer all questions in Part A.
 3. Part B consists of 5 Units. Answer any one full question from each unit.
 4. Each question carries 10 marks and may have a, b, c, d as sub questions.

Part-A

All the following questions carry equal marks

- 1 Differentiate Online Analytical Processing (OLAP) and Online Transaction Processing (OLTP). (10x2M=20 Marks)
- 2 Define snowflake schema.
- 3 Justify the importance data mining.
- 4 What is dimensionality reduction?
- 5 Discuss about support and confidence with respect to association rule mining .
- 6 Explain frequent item set mining.
- 7 Define classification.
- 8 How to select a best classification algorithm for your dataset?
- 9 List down the hierarchical clustering algorithms.
- 10 Describe about outliers.

Part-B

Answer All the following questions.

(5X10M=50Marks)

- 11 Explain data warehouse components with neat diagram. (10M)
OR
- 12 a) Write about the working of ROLAP. (5M)
b) Compare MOLAP and HOLAP. (5M)
- 13 a) Show the order of knowledge discovery process. (5M)
b) Examine the purpose of data preprocessing in data mining. (5M)
OR
- 14 a) Analyze the role of feature selection in data mining. (5M)
b) Discuss about any two similarity measures technique. (5M)
- 15 a) What are the applications of association rule mining? (2M)
b) Explain Apriori algorithm with an example. (8M)
OR
- 16 Write about the steps and example for frequent pattern growth algorithm. (10M)

- 17 a) What are the basic requirements for classification? (4M)
b) Demonstrate Decision tree algorithm with an example. (6M)
OR
- 18 Discuss the working of naïve Bayes algorithm with an example. (10M)
- 19 a) How to evaluate the performance of clustering algorithms. (4M)
b) Write notes on divisive clustering algorithm. (6M)
OR
- 20 Explain K-means and PAM clustering algorithms in detail. (10M)