USN | | | | | | | | | | |

15CS651

Sixth Semester B.E. Degree Examination, Jan./Feb. 2023
Data Mining and Data Warehousing

Time: 3 hrs.

Max. Marks: 80

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

1 a. What is data warehouse and explain its features? **(04 Marks)**
   b. Explain the measures and how they are categorized? **(04 Marks)**
   c. Consider sales example and explain star schema and snowflakes schema and compare them. **(08 Marks)**

OR

2 a. What is concept hierarchy, describe any one type of concept hierarchy? **(04 Marks)**
   b. Write a note on Enterprise warehouse and Datamart. **(04 Marks)**
   c. Explain OLAP operations in multidimensional data model with example. **(08 Marks)**

Module-2

3 a. Describe a 3 tier data warehousing architecture with neat diagram. **(08 Marks)**
   b. What is datamining? Explain motivating challenges in its development? **(08 Marks)**

OR

4 a. How to index OLAP data using bitmap indexing? Explain with an example. **(06 Marks)**
   b. Explain how data mining tasks are categorized. Explain the core datamining tasks with examples. **(10 Marks)**

Module-3

5 a. Describe the generation of candidate item set and frequent item sets where the minimum support count is 2 for the following transactional data. **(10 Marks)**

TID	List of item IDs
T100	11, 12, 15
T200	12, 14
T300	12, 13
T400	11, 12, 14
T500	11, 13
T600	12, 13
T700	11, 13
T800	11, 12, 13, 15
T900	11, 12, 13

   b. Write Apriori algorithm. **(06 Marks)**

OR

6 a. Explain the alternative methods for generating frequency item sets. **(10 Marks)**

b. Define the following with an example each:
   i) Association rule
   ii) Support of a rule
   iii) Confidence of a rule. **(06 Marks)**

Module-4

7 a. Explain classification process. Write algorithm for decision tree induction technique. **(10 Marks)**

b. Write an algorithm for K-nearest neighbor classifier and explain. **(06 Marks)**

OR

8 a. Discuss in detail the Bayesian classification methods with suitable examples. **(10 Marks)**

b. Explain rule based classifier and ripper algorithm. **(06 Marks)**

Module-5

9 a. What is cluster analysis? What are the applications of cluster analysis? **(08 Marks)**

b. Describe the density based clustering with algorithm. **(08 Marks)**

OR

10 a. Explain the different types of clusters. **(08 Marks)**

b. Explain agglomerative hierarchical clustering. **(08 Marks)**

* * * * *