VI Semester M.C.A./IV Semester M.Sc. Examination, June/July 2018
(CBCS)
COMPUTER SCIENCE
MCA 6E12/4E5 : Data Mining

Time : 3 Hours
Max. Marks : 70

Instructions : Answer any five from Part – A and any four from Part – B.

PART – A

Answer **any five** questions. **Each** question carries **six** marks :
(5×6=30)

1. How does the data warehousing and data mining work together ?

2. What is data mining functionality ? Explain different types of data mining functionality with examples.

3. Explain types of OLAP Servers.

4. Explain FP tree algorithm with an example.

5. Discuss the following clustering algorithm using examples :
   a) K-means.
   b) K-medoid.

6. With a neat diagram explain the components of data warehouse.

7. Illustrate the difference between supervised, semi-supervised and unsupervised learning with examples.

8. With an example, describe snowflake and fact constellations.

PART – B

Answer **any four** questions. **Each** question carries **ten** marks :
(4×10=40)

9. a) What is Data Preprocessing ? What are the steps involved in it ? Explain the different issues in Data Preprocessing in detail. 6

   b) Write about Bayesian classification. 4

10. a) Explain a few real time applications of data mining.   4

    b) Distinguish between classification and prediction with example.   6

11. a) What do you mean by Lazy Learner Classification ?   4

    b) Apply the algorithm to discover frequent item sets on the following transaction data set.   6

TID	Items
1	{a, b}
2	{b, c, d}
3	{a, c, d, e}
4	{a, d, e}
5	{a, b, c}
6	{a, b, c, d}
7	{a}
8	{a, b, c}
9	{a, b, d}
10	{b, c, e}

12. a) How will you solve a classification problem using decision trees with example.   6

    b) What is outlier analysis ? Explain.   4

13. a) Perform a comparative study between MOLAP and ROLAP.   4

    b) Describe the issues and challenges in the implementation of data mining systems.   6

14. Write a note on :

    a) Genetic Algorithm.   5

    b) Multilayer feed-forward Neural Network.   5

**VI Semester M.C.A./IV Semester M.Sc. Examination, June 2017
(CBCS)
COMPUTER SCIENCE
MCA 6E 12/4E5 : Data Mining**

Time : 3 Hours

Max. Marks : 70

Instruction : *Answer* **any five** *questions from Part – **A** and* **any four** *questions in Part – **B**.*

PART – A

Answer **any five** questions. **Each** question carries **six** marks. **(5×6=30)**

1. How is a Data Warehouse different from a Database ? Explain the 3-tier architecture of a Data Warehouse.

2. What are the benefits of a Data Warehouse to a Business Analyst ?

3. Explain types of OLAP servers.

4. Discuss the role of a concept hierarchy in Data Mining.

5. Explain :
   i) Frequent item sets and closed item sets.
   ii) Support and confidence measures.

6. Explain multi-dimensional association rule mining with an example.

7. Illustrate the difference between supervised, semi-supervised and unsupervised learning with examples.

8. Discuss various attribute selection methods.

PART – B

Answer **any four** questions. **Each** question carries **ten** marks. **(4×10=40)**

9. a) Discuss the different types of warehouse schema. 6
   b) Differentiate :
      i) OLAP and Data Mart
      ii) Operational Database and Informational Database. 4

P.T.O.

10. a) What is a OLAP Cube ? Explain basic OLAP operations.   5

    b) How does Data Load Tuning and Query Tuning techniques improve the performance of a Data Warehouse ?   5

11. a) What is the need for data pre-processing ? Discuss various data pre-processing tasks.   6

    b) What is Bitmap indexing ? Explain with an example.   4

12. a) Write an algorithm to discover frequent item sets without candidate generation.   5

    b) Apply the algorithm to discover frequent item sets on the following transaction data set.   5

TID	Items
1	{a, b}
2	{b, c, d}
3	{a, c, d, e}
4	{a, d, e}
5	{a, b, c}
6	{a, b, c, d}
7	{a}
8	{a, b, c}
9	{a, b, d}
10	{b, c, e}

13. a) What are classification rules ? How is regression related to classification ?   5

    b) Explain Lazy learner with an example.   5

14. a) What is density based clustering ? Write DBSCAN clustering algorithm.   5

    b) What are outliers ? Discuss different types of outliers.   5

———————