CBCS SCHEME

USN | | | | | | | | | |

16/17MCA442

Fourth Semester MCA Degree Examination, June/July 2019
Data Warehousing and Data Mining

Time: 3 hrs. Max. Marks: 80

Note: ***Answer FIVE full questions, choosing one full question from each module.***

Module-1

1 a. Define a Data warehouse. List and briefly explain the key features of data warehouse. **(06 Marks)**

b. Discuss briefly about star, snow flake and fact constellation schema with necessary diagram. **(10 Marks)**

OR

2 a. Differentiate between OLAP and OLTP. **(06 Marks)**

b. Explain the following terms :
i) Data Mart ii) Virtual warehouse iii) Data Cube iv) ROLAP and MOLAP. **(10 Marks)**

Module-2

3 a. Define Data Mining. With a neat diagram, explain the process of knowledge discovery in databases. **(08 Marks)**

b. List and explain the Data mining applications. **(08 Marks)**

OR

4 a. List and explain the attributes respective to qualitative and quantitative measurement. **(08 Marks)**

b. Mention the data mining techniques used during data preprocessing explain any 2 in detail. **(08 Marks)**

Module-3

5 a. State Apriori principle for generating Item sets. Write and explain the pseudo code for the frequent item set generation part of apriori algorithm. **(08 Marks)**

b. Construct item set Lattice for the item set I = {a, b, c, d, e} and list all the item subsets. **(08 Marks)**

OR

6 a. Discuss briefly the Alternative methods for generating Frequent Item sets. **(08 Marks)**

b. Construct FP- tree for the following Transaction data set

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}

(08 Marks)

Module-4

7 a. What is a decision tree? Write Hunt's algorithm for Decision tree induction. **(06 Marks)**
b. Discuss K- nearest neighbor classification algorithm with characteristics of nearest neighbor classifiers with necessary diagram. **(10 Marks)**

OR

8 a. Discuss briefly the various techniques for improving the accuracy of classification methods. **(08 Marks)**
b. Discuss the multi class problem with example. **(08 Marks)**

Module-5

9 a. What is cluster analysis? Briefly explain the cluster analysis methods. **(10 Marks)**
b. Describe DBSCAN method briefly. **(06 Marks)**

OR

10 a. List and explain the applications of clustering techniques. **(08 Marks)**
b. List and explain different types of clusters. **(08 Marks)**

* * * * *