

END TERM EXAMINATION

FOURTH SEMESTER [MCA] MAY-JUNE 2018

Paper Code: MCA-204

Subject: Data Warehousing and Data Mining

Time: 3 Hours

Maximum Marks: 75

Note: Attempt five questions in all including Q no.1 which is compulsory.
Select one question from each unit.

- Q1 Answer **any ten** of the following:- (10x2.5=25)
- (a) What is Knowledge Discovery in Database? How does it related to Data Mining?
 - (b) Provide example for incomplete and inconsistent data.
 - (c) What is the difference between classification and prediction?
 - (d) What are the data mining tasks performed on a text database?
 - (e) List out any two major strengths of decision tree method.
 - (f) Discuss various data mining task primitive.
 - (g) What is meant by 'Association Rule Mining'?
 - (h) What is the difference between fact table and dimensional table?
 - (i) What are the requirements of cluster analysis?
 - (j) What is the difference between data warehouse and data mart.
 - (k) What do you mean by Cluster Analysis?
 - (l) Give some alternative terms for data mining.
 - (m) What are the steps involved in KDD process.
 - (n) List out the component of start schema.
 - (o) What is roll-up operation?

UNIT-I

- Q2 (a) Explain the three-tier data warehouse architecture. (6.25)
- (b) What is OLAP? Explain the OLAP operations on multidimensional data with example. (6.25)
- Q3 (a) What are the decision support applications? Why are the traditional DBMS inadequate for decision making? (6.25)
- (b) ~~Perform Quick sort dry run on following data. Show the various steps involved in the quick sorting of the data.~~ What is data warehousing and why it is important for decision support? (6.25)

UNIT-II

- Q4 (a) What are the main considerations that are involved in the designing of the data Warehouse? Also discuss the methods for the implementation of Data Warehousing systems? (6.25)
- (b) With an example, describe any two schema (star/snowflake/fact constellation) definition using DMQL statements. (6.25)
- Q5 (a) What are the drawbacks of virtual data warehouse? Explain in detail. (6.25)
- (b) Explain different type of OLAP servers? (6.25)

P.T.O.

MCA-204
P/2

UNIT-III

- Q6 (a) Describe the data mining functionalities and examine what kind of patterns can be mined. (6.25)
(b) What is the necessity of data preprocessing of the data? How can data preprocessing techniques improve the quality of data. (6.25)
- Q7 (a) List any one tool for data mining and show how that tool will be useful for a ration shop system. (6.25)
(b) What are the various data mining tasks supported by the DB miner system? (6.25)

UNIT-IV

- Q8 (a) What is cluster analysis? Explain the classification of major clustering and partitioning methods. (6.25)
(b) What is decision tree? Elaborate on classification done using Decision tree induction. (6.25)
- Q9 (a) Write k-means and k-medoids algorithm and explain with example. (6.25)
(b) What is the difference between density based methods and grid based methods? Explain with suitable example. (6.25)
