

END TERM EXAMINATION

FOURTH SEMESTER [MCA]MAY-JUNE 2013

Paper Code: MCA 204

Subject: Data Warehousing & Data Mining

Time:3 Hours

Maximum Marks:60

Note: Q.No.1 is compulsory. Attempt one question from each unit.

- 1 (i) Data warehouse data is non-volatile. Explain
- (ii) What is data mart? When is it appropriate?
- (iii) What do you mean by web enabled data warehouse?
- (iv) What is factless fact table? Explain with example.
- (v) Define HOLAP?
- (vi) Explain slice and dice in context of OLAP.
- (vii) What are data mining primitives?
- (viii) What are the difference between classification and clustering?
- (ix) How data mining is the primary step in knowledge discovery in data?
- (x) What is DMQL?

2x10=20

Unit-I

2. (a) Draw and explain basic building blocks of data warehouse
- (b) What is strategic information? What are different characteristics of strategic information? For an insurance business, name five types of strategic information.

(5, 5)

3. (a) What is information package diagram (IPD)? How it helps in dimensional analysis? Make an IPD of sale analysis system.
- (b) What do you mean by informational system? How is it different from operational system?

(7, 3)

Unit-II

4. (a) Differentiate between fully additive and semi-additive facts with examples?
- (b) What is STAR Schema? Explain by taking suitable example. A dimension table is wide and fact table is deep. Justify.

(3, 7)

P.T.O

[-2-]

5. (a) Explain aggregate fact table. What are various ways of forming aggregate fact table?

(b) What are hyper cubes? How do they apply in an OLAP system?

(5, 5)

Unit-III

6. (a) What do you mean by data mining functionalities?

(b) What is data mining? Explain Knowledge Discovery In Data ((KDD) process in details.

(5, 5)

7. (a) What is data integration and transformation? How are these performed?

(b) How data mining is different from OLAP? Explain briefly.

(5, 5)

Unit-IV

8. (a) Write short note on

(i) Memory Based Reasoning

(ii) Neural Network

(b) What is Association rule mining? Describe Apriori method with example.

(5, 5)

9. (a) Suppose that the data mining task is to cluster the following eight points, with (x, y) representing location, into 3 clusters.

A1(2, 10); A2(2, 5); A3(8, 4); B1(5, 8); B2(7, 5); B3(6, 4); C1(1, 2); C2(4, 9);

The distance function is Euclidean distance. Suppose initially, we assign A1,

B1 and C1 as the center of each cluster respectively. Use the k-means algorithm to show

(a) the three cluster centers after the first round execution

(b) the final three clusters.

(b) Discuss various applications of data mining.

(5, 5)
