

Roll No.

Total No. of Pages : 02

Total No. of Questions : 07

**BCA (2013 & Onward) (Sem.-5)**  
**DATA WAREHOUSING AND MINING**  
Subject Code : BSBC-501  
Paper ID : [B1154]

Time : 3 Hrs.

Max. Marks : 60

**INSTRUCTION TO CANDIDATES :**

1. SECTION-A is COMPULSORY consisting of TEN questions carrying TWO marks each.
2. SECTION-B contains SIX questions carrying TEN marks each and a student has to attempt any FOUR questions.

**SECTION-A**

**Q1. Answer briefly :**

- a) Differentiate between operational and informational data stores.
- b) What is multidimensional data? Give two examples.
- c) What is OLTP?
- d) Define Data Mining.
- e) Briefly discuss the Snowflake schema.
- f) Discuss discovery driven cube.
- g) What is a Decision Tree?
- h) What is Apriori algorithm?
- i) What are the different types of data used in cluster analysis?
- j) How can the accuracy of a classifier increased?

## SECTION-B

- Q2. Define Data warehouse. Elaborate in detail the design and construction of data warehouses.
- Q3. What do you mean by data pre-processing? Explain the various stages in the process of data pre-processing.
- Q4. Suppose that a data warehouse consists of the four dimensions date, spectator, location and game and the two measures count and charge, where charge is the fare that a spectator pays when watching a game on a given date. Spectators may be students, adults or seniors, with each category having its own charge rate.
- a) Draw a Star schema diagram for the data warehouse.
  - b) Starting with the base cuboid [date, spectator, location, game], what specific OLAP operations should one perform in order to list the total charge paid by student spectators at GM\_Place in 2015?
- Q5. What is Attribute-oriented Induction? Describe how this is implemented.
- Q6. What is Clustering? Discuss the various clustering algorithms.
- Q7. Explain data visualization with reference to data mining.