Code: CT75   **Subject: DATA WAREHOUSING AND DATA MINING**

ALCCS

Time: 3 Hours   JUNE 2015   **Max. Marks: 100**

PLEASE WRITE YOUR ROLL NO. AT THE SPACE PROVIDED ON EACH PAGE IMMEDIATELY AFTER RECEIVING THE QUESTION PAPER.

NOTE:
- **Question 1 is compulsory and carries 28 marks. Answer any FOUR questions from the rest. Marks are indicated against each question.**
- **Parts of a question should be answered at the same place.**

Q.1   a. Define a Data warehouse. Compare OLTP and OLAP systems.

   b. Discuss OLAP operations in the multidimensional data model.

   c. Briefly explain agglomerative and divisive Hierarchical clustering methods.

   d. "Data mining tools may improve telecommunication services". Explain.

   e. What do you understand by Noisy data? Briefly, explain any two smoothing techniques.

   f. State any four criteria on which classification and prediction methods can be compared and evaluated.

   g. Briefly explain Apriori algorithm.   **(7×4)**

Q.2   Briefly compare the following concepts. You may use an example to explain your point.
   (i) *Snowflake* schema, *fact constellation*, *starnet* query model
   (ii) Data cleaning, data transformation, data refresh
   (iii) Enterprise warehouse, data mart, virtual warehouse   **(3×6)**

Q.3   a. Define Data mining. Draw and explain the architecture of typical data mining system.
                                                                            (10)
   b. Describe three challenges to data mining regarding data mining methodology and User-interaction issues.   **(8)**

Q.4   a. How does data mining relate to information processing and online analytical processing?   **(10)**

   b. What is metadata? Explain metadata repository.   **(8)**

Q.5   a. Explain how Rule-based classification is used in data mining.   **(10)**

   b. Describe Classification based on Multiple Association Rules (CMAR) with an example.   **(8)**

Q.6   a. What is cluster analysis? Explain major categorization of clustering Methods.   **(10)**

   b. What do you mean by a web-enabled data-warehouse? Describe three of its functional features   **(8)**

Q.7   Write short note on   **(3×6)**
      (i) Data Mining for the Retail Industry
      (ii) Statistical Data Mining
      (iii) Strategies of data reduction