Printed Pages—2   **ECS075**

B. Tech.
(SEM. VII) THEORY EXAMINATION 2011-12
DATA MINING AND DATA WAREHOUSING

Time : 3 Hours   *Total Marks : 100*

Note :– Attempt **all** questions.

1. Attempt any **two** parts of the following :   **(10×2=20)**

    (a) Describe in brief the important steps of data mining and data mining functionalities.

    (b) Write and describe important types of difficulties in data mining process.

    (c) (i) Describe in brief the process of Data Integration and Transformation.

        (ii) Write and explain the characteristics of operational data.

2. Attempt any **two** parts of the following :   **(10×2=20)**

    (a) Explain the market basket analysis. Describe the basic concepts of association rule mining.

    (b) Describe the Apriori algorithm for FIM (Frequent Itemset Mining) and verify it through a suitable example.

    (c) Why is the task of mining frequent itemsets difficult ? Explain the reasons.

ECS075/KIH-26390   *[Turn Over*

3. Attempt any **two** parts of the following :   (10×2=20)

    (a) Describe classification. Briefly outline the major ideas of Basiyan classification.

    (b) What is clustering ? How is this different than classification ? Explain any one approach for clustering.

    (c) Explain the types of data that often occur in cluster analysis and briefly explain how to preprocess that **data** for clustering.

4. Attempt any **two** parts of the following :   (10×2=20)

    (a) Briefly explain important approaches to build the data warehouse.

    (b) Describe various schemas of multidimensional data models.

    (c) What are the differences between the three main types of data warehouse usage : information processing, analytical processing and data mining ? Briefly explain.

5. Attempt any **two** parts of the following :   (10×2=20)

    (a) Define and describe the basic similarities and differences among ROLAP, MOLAP, and HOLAP.

    (b) Discuss various OLAP operations. Explain how query performance can be improved by cascading the operations.

    (c) Describe the following in brief :

        (i) Data mining interface

        (ii) Testing of data warehouse.