(⌒)

B.Tech.

EIGHTH SEMESTER EXAMINATION, 2005-2006

DATA MINING AND WAREHOUSING

Time : 3 Hours   *Total Marks : 100*

Note : (i)   *Attempt ALL questions.*

(ii)   *All questions carry equal marks.*

(iii) *In case of numerical problems assume data wherever not provided.*

(iv) *Be precise in your answer.*

002534

1.   Attempt *any four* parts of the following :   (5×4=20)

(a)   What is a Data Warehouse ? How does it differ from a Database ?

(b)   What steps are involved in the acquisition of data for a Data Warehouse ?

(c)   What are the difficulties in implementing Data Warehouses ?

(d)   What is the multidimensional data model ? How is it used in Data Warehousing ?

(e)   Define the terms :
   (i)   OLAP   (ii)   ROLAP   (iii)   MOLAP
   (iv) DSS   (v)   Datamarts

(f)   Describe the characteristics of Data Warehouse. How is the concept of relational view related to Data Warehouse ?

2. Attempt *any four* parts of the following : (5x4=20)

   (a) Describe the benefits and drawbacks of a source-driven architecture for gathering of data at a data-warehouse as compared to a destination-driven architecture.

   (b) What are data warehouse schemes ?

   (c) What do you understand by Data cleaning and Backflushing in context of Data warehousir ?

   (d) What are the typical functionalities of a Data warehouse ?

   (e) List the open issues and research problems in Data warehousing.

   (f) How would you differentiate between a Data warehouse and Views ?

3. Attempt *any two* parts of the following : (10x2=20)

   (a) What is Business Analysis ? How would you do so for an organization ?

   (b) Describe the capabilities of Online Analytical Processing.

   (c) What is the role of Artificial Intelligence in Data Mining ?

4. Attempt *any two* parts of the following : (10x2=20)

   (a) Explain the concept of Datamining. What are the goals or tasks it attempts to facilitate ?

   (b) What are the different phases of knowledge discovery from databases ? Describe a complete application scenario in which new knowledge may be mined from an existing database of transactions ?

   (c) Explain the process of Clustering in Datamining.

5. Attempt *any two* parts of the following : (10x2=20)

(a) What are the mistakes that Data warehousing managers should avoid ?

(b) What are Dr. E.F. Codd's 12 guidelines for OLAP ?

(c) What do you understand by Data Visualization ? What are the common techniques of Data Visualization ? What are the characteristics of good Visualization ?

- o O o -