

(Following Paper ID and Roll No. to be filled in your Answer Book)

PAPER ID : 2712

Roll No.

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

B.Tech.

(SEM. VII) ODD SEMESTER THEORY EXAMINATION 2012-13

DATA MINING & DATA WAREHOUSING

Time : 3 Hours

Total Marks : 100

Note : Attempt all questions.

1. Attempt any **FOUR** parts of the following : **(4×5=20)**
 - (a) Define KDD. Identify and describe the phases in KDD process.
 - (b) Explain Attribute subset selection method for data reduction with examples.
 - (c) Describe the difference between the following approaches for the integration of data mining system with database or data warehouse systems : no coupling, loose coupling, semi tight coupling, tight coupling.
 - (d) Explain Principal Component Analysis (PCA) in detail.
 - (e) What are Outliers ? How outliers analysis can be done ?
 - (f) Distinguish between Dimensionality reduction and Numerosity reduction.
2. Attempt any **TWO** parts of the following : **(2×10=20)**
 - (a) What is Association rule mining ? Explain the Apriori algorithm to find the frequent item sets.

- (b) Discuss why Analytical characterization and Attribute relevance analysis are needed.
 - (c) (i) Explain mining multilevel association rules from Transactional databases.
 - (ii) Describe statistical measures in large databases.
3. Attempt any **TWO** parts of the following : **(2×10=20)**
- (a) Describe classification and prediction. Also discuss a method regarding classification.
 - (b) Write short notes on :
 - (i) CLIQUE
 - (ii) STING.
 - (c) Describe the role of Genetic Algorithm in data mining.
4. Attempt any **TWO** parts of the following : **(2×10=20)**
- (a) Why Data warehouse is maintained separately from Database ? Differentiate between OLTP and OLAP.
 - (b) Explain with diagram the stars, snowflake and fact constellation schemas for multidimensional data bases. Also write their advantages and disadvantages.
 - (c) Describe the following :
 - (i) Concept hierarchy
 - (ii) Data Mart.

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

(a) Describe the various types of OLAP servers.

(b) Explain the following :

(i) Data mining interfaces

(ii) Testing data warehouse.

(c) Explain the following :

(i) Backup and Recovery

(ii) Security issues in data warehouse.