22621

12425

3 Hours / 70 Marks Seat No. | | | | | | | | |

Instructions : (1) All Questions are *compulsory.*

(2) Answer each next main Question on a new page.

(3) Illustrate your answers with neat sketches wherever necessary.

(4) Figures to the right indicate full marks.

(5) Assume suitable data, if necessary.

(6) Use of Non-programmable Electronic Pocket Calculator is permissible.

(7) Mobile Phone, Pager and any other Electronic Communication devices are not permissible in Examination Hall.

Marks

1. Attempt any FIVE of the following : **10**

(a) Define data warehousing. List any two benefits.

(b) What is data cube ? Give the example.

(c) Define the term HOLAP.

(d) List any four kinds of data used in data mining.

(e) Define frequent and closed item sets.

(f) List four characteristics of data warehousing.

(g) Define base and apex cuboids.

2. Attempt any THREE of the following : **12**

(a) Explain star schema with diagram.

(b) Explain bitmap index for OLAP with example.

(c) What is data preprocessing ? Explain the need of data preprocessing.

(d) Explain market basket analysis with example.

3. Attempt any THREE of the following : **12**

(a) Differentiate between OLTP and OLAP (any four points).

(b) Explain multidimensional data model.

(c) Explain efficient data cube computation using full materialization method.

(d) Explain different mining steps in the process of KDD.

4. Attempt any THREE of the following : **12**

(a) Explain three-tier architecture of data warehouse with diagram.

(b) Describe any two data cube computation strategies.

(c) Explain two major issues in data mining.

(d) What is cluster analysis ? Explain the requirements of cluster analysis.

(e) Differentiate between ROLAP and MOLAP (any four points).

5. Attempt any TWO of the following : **12**

(a) Explain following data warehouse design approaches with diagram :

(i) Top-Down Approach

(ii) Bottom-Up Approach

(b) Explain different types of attributes with example.

(c) Explain K-means clustering algorithm with example.

6. Attempt any TWO of the following : **12**

(a) Explain any two OLAP operations with example.

(b) Define data cleaning. Explain different data cleaning methods with example.

(c) Consider the given database (D).

TID	Items
1	a c d
2	b c e
3	a b c e
4	b e

Apply the Apriori algorithm and find frequent item sets with 50% support and 70% confidence.