

**B.Tech. – VIEP – COMPUTER SCIENCE AND  
ENGINEERING (BTCSVI)**

**Term-End Examination**

**June, 2014**

00551

**BICS-020 : DATA WAREHOUSING AND MINING**

*Time : 3 hours*

*Maximum Marks : 70*

---

**Note :** *Attempt any seven questions. All questions carry equal marks. Assume suitable missing data, if any.*

---

1. (a) What do you mean by data mining ? Define data, information and knowledge. 2+3=5  
(b) Explain KDD process. 5
  
2. (a) Differentiate between OLTP and Data warehouse systems. 5  
(b) What do you mean by subject oriented, integrated, time-variant and non-volatile collection of data in Data Warehouse ? 5
  
3. (a) Explain the architecture of Data Warehouse with neat diagram. 5  
(b) Explain metadata and its importance. 5

4. (a) Differentiate between Star schema and Snowflake schema with example. 5
- (b) Explain multidimensional model of a Data Warehouse. 5
5. (a) What are the different methods of data transformation ?  
Suppose that the mean and standard deviation of the values for the attribute income are ₹ 54,000 and ₹ 16,000 respectively. Apply Z-score normalization, to transform a value of ₹ 73,600 for income. 3+2=5
- (b) Explain the various strategies of data reduction. 5
6. (a) What are the different primitives by which data mining query can be defined ? 5
- (b) What is interestingness measures ? Mention some interestingness measures for association rules. 2+3=5
7. (a) What are the major differences between methods of learning from examples and data mining methods ? 5
- (b) Explain Market Basket Analysis. 5

8. (a) Trace the result of using Apriori algorithm on the dataset given below with support 50% and confidence 60% : 5

<u>TID</u>	<u>Items</u>		
10	1	3	4
20	2	3	5
30	1	2	3 5
40	2	5	

- (b) What are the different types of data that often occur in cluster analysis? 5
9. (a) What is multilevel association rules? Write down the different approaches to mining multilevel association rules. 2+3=5
- (b) What is data classification? How is prediction different from classification? 5
10. Write short notes on any *two* : 5+5=10
- (a) Spatial mining
  - (b) Testing of Data Warehouse
  - (c) ETL process of Data Warehouse
  - (d) World Wide Web
-