B.E. (Information Technology) Semester Seventh (C.B.S.)

Data Warehousing & Mining

P. Pages : 2 **KNT/KW/16/7498**

Time : Three Hours * 0 8 8 6 * Max. Marks : 80

Notes :
1. All questions carry marks as indicated.
2. Solve Question 1 OR Questions No. 2.
3. Solve Question 3 OR Questions No. 4.
4. Solve Question 5 OR Questions No. 6.
5. Solve Question 7 OR Questions No. 8.
6. Solve Question 9 OR Questions No. 10.
7. Solve Question 11 OR Questions No. 12.
8. Due credit will be given to neatness and adequate dimensions.
9. Assume suitable data wherever necessary.
10. Illustrate your answers wherever necessary with the help of neat sketches.

1. a) Define Decision support system (DSS)? What are the failures of Past DSS. **5**

b) Differentiate between. **8**

i) Data warehouse and Data mart.
ii) Operational Database system and Decision support system.

OR

2. a) Explain three tier architecture of Dataware house with neat diagram. **7**

b) Explain Life Cycle of Dataware house with neat sketch. **6**

3. a) What are the major task of data processing? Explain in brief. **6**

b) Find **7**
i) Mean
ii) Standard deviation
iii) Min-max normalization
iv) Z-score normalization
for the following data set.
30, 36, 47, 52, 52, 56, 60, 63, 70, 70, 110

OR

4. a) What is data reduction? Explain different methods of data reduction. **7**

b) How data cleaning can be handled in Preprocessing? **6**

5. a) What is data cube? Explain the operations of OLAP with example. **6**

b) What is the need of multidimensional analysis? **4**

c) Write any six characteristics of OLAP. **4**

OR

6. a) Differentiate between OLTP and OLAP. **6**

b) Describe STAR and snow & lave schema with example. **8**

7. a) Describe the steps involved in knowledge discovery in database. **6**

b) Explain the applications of data mining in **any two.** **8**

i) Financial Institute
ii) Retail industry
iii) Telecommunication Industry.

OR

8. a) What are the major issues of data mining? **7**

b) With the schematic diagram, describe the architecture of data mining system. **7**

9. a) Explain Apriori algorithm for frequent itemset using candidate generation for the following transactional database. **9**

T I D	List of Items
T 100	A, B, D, E
T 200	B, C, E
T 300	A, B, D, E
T 400	A, B, C, E
T 500	A, B, C, D, E
T 600	B, C, D

Minimum-Support = 30%, minimum-confidence threshold = 60%, Also generate the association rules for frequent itemset.

b) Explain the procedure to generate association rules from frequent itemset. **4**

10. a) Explain FP-Growth in detail with example. How FP-Growth is efficient than apriori algorithm. **7**

b) Explain market basket analysis with example. **6**

11. a) Define BI. What are the important factors of BI? **6**

b) Explain the two Approaches of development of BI system. **7**

OR

12. a) What is Bigdata? What are the challenges and characteristics of Bigdata? **5**

b) Describe architecture of Hadoop ecosystem. **8**

****