B.E. (Computer Science & Engineering) Seventh Semester (C.B.S.)

Data Warehousing & Mining

P. Pages : 2 **TKN/KS/16/7574**

Time : Three Hours * 0 8 8 5 * 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 whenever necessary.
10. Illustrate your answers wherever necessary with the help of neat sketches.
11. Use of non-programmable calculator is permitted.

1. a) What are the different data Mining Functionalities? **7**

b) Discuss the Major issues in Data mining. **7**

OR

2. a) Give the classification of data mining system. Explain in detail. **7**

b) Explain the different techniques for data reduction. **7**

3. a) What is data warehouse? Explain architecture of data warehouse. **7**

b) Enumerate three classes of Schemas that are popularly used for modeling data Warehouse. Write features of each Schema. **7**

OR

4. a) Write the difference between OLAP & OLTP. **4**

b) What is OLAP? What are the different OLAP operations that can be performed on multidimensional data model. **7**

c) Write short note on ROLAP model. **3**

5. a) Consider following transactional dataset. find frequent item sets and association rules using apriori algorithm. With support = 30% & confidence = 70% **9**

TID	List of Items IDS
T100	I1, I2 I5
T200	I2, I4
T300	I2, I3.
T400	I1, I2, I4.
T500	I1, I3,
T600	I2, I3
T700	I1, I3
T800	I1, I2, I3, I5
T900	I1, I2, I3

b) Write in brief about constraint Based Association Mining. **4**

OR

6. a) Explain in brief market Basket Analysis. **5**

b) Define the following terms. **6**
i) Frequent Item sets.
ii) Closed item sets.
iii) Association rules.

c) What is correlation Analysis? **2**

7. a) What are the different issues regarding classification and prediction. **6**

b) Write short note on. **7**
i) Support Vector Machine (SVM)
ii) Classification by Back propagation.

OR

8. a) Explain classification by Decision Tree Induction with example. **6**

b) What are the different measures for Accuracy and error in classification or prediction. **5**

c) What do you mean by Lazy Learners? **2**

9. a) How the clustering methods are categorize? **6**

b) Illustrate and explain partitioning method for clustering. **7**

OR

10. a) What do you mean by hierarchical clustering approach? Explain agglomerative and divisive hierarchical clustering. **7**

b) What is outlier? Why outlier analysis is important? **3**

c) Write short note on "Constraint- Based Cluster Analysis." **3**

11. a) Explain the technique for mining time series Data? **7**

b) Define following terms **6**
i) Data streams.
ii) Time series Data
iii) Sequence Data

OR

12. a) Write short on **any three**. **13**

i) Graph mining

ii) Mining sequence pattern in Biological Data.

iii) Social Network Analysis.

iv) Multirelational Data mining.

v) Link Mining.

****