

M.Sc. Part-II Semester-III (CBCS) Examination

COMPUTER SCIENCE

3MCS1 : DATA MINING AND DATA WAREHOUSING

Time : Three Hours]

[Maximum Marks : 80

N.B.:— (1) Assume suitable data wherever necessary.

(2) Illustrate your answer with the help of neat sketches.

1. (a) What are the different data mining functionalities ? Explain. 8
- (b) Give the difference between following with example : 6
 - (i) Discrimination and classification.
 - (ii) Characterization and clustering.

OR

2. (a) In real life data tuple with missing values for some attributes are a common occurrence. Describe the method of handling this problem. 6
- (b) Describe the issues of challenges in data mining regarding data mining methodology. 8
3. (a) Explain the three tier Architecture of the data warehouse. 7
- (b) Describe the multidimensional data model. 6

OR

4. (a) How cube computation method is used in dimensionality ?
 - (i) Compute a dense full cube for low dimensionality
 - (ii) Compute a sparse iceberg cube of high dimensionality. 7
- (b) Explain in brief ROLAP, MOLAP and HOLAP. How materialization refers to the computation of all of the cuboids in a lattice defining a data cube ? 6
5. (a) Describe a method for generating association rules from frequent itemset. 7
- (b) Explain Apriori Algorithm for finding frequent itemset with suitable example. 6

OR

6. (a) Explain Association Analysis to Correlation Analysis. What measures are considered for comparison of pattern evaluation measures. 7
- (b) How mining is performed in multilevel, multidimensional space ? Give example. 6
7. (a) Explain following terms with respect to attribute selection measures :
 - (i) Information Gain
 - (ii) Gain Ratio
 - (iii) Gini index. 6

- (b) Explain why tree pruning is useful in decision tree induction. What is drawback of using a separate set of tuple to evaluate pruning ? 7

OR

8. (a) Explain how accuracy and errors can be measured. 6
(b) How to calculate the worst-case computational complexity of the decision tree algorithm. What is the role of computational cost of growing tree. 7
9. (a) Discuss in detail partitioning method. Write algorithm for K-means centroid base technique. 8
(b) Explain following giving example of each :
(i) Agglomerative Hierarchical clustering
(ii) Divisive Hierarchical clustering. 6

OR

10. (a) Briefly describe and give example of each of following approaches of clustering :
(i) Partitioning method
(ii) Hierarchical method
(iii) Density Base method
(iv) Grid Base Method. 8
(b) Explain the following :
(i) Mining Time Series Data
(ii) Mining Sequence Pattern. 6
11. (a) Explain following data mining applications :
(i) Mining for financial data
(ii) Mining for Telecommunication industry data. 6
(b) Name and describe the main theoretical foundation that have been proposed for data mining. 7

OR

12. (a) Explain how clustering and classification are used in graph mining. 6
(b) State and explain implementation method for audio data mining. Is it possible to integrate audio and video data mining method ? Explain. 7