C

Question Paper Code: 55802

B.E. / B.Tech. DEGREE EXAMINATION, NOV 2019

Fifth Semester

Information Technology

15UIT502 - DATA WAREHOUSING AND DATA MINING

(Regulation 2015)

Duration: Three hours   Maximum: 100 Marks

Answer ALL Questions

PART A - (5 x 1 = 5 Marks)

1. Which schema contains normalization process   CO1- R

    (a) Star   (b) Snowflakes   (c) Fact Constellation   (d) Ring

2. Identify the factor is used in data preprocessing   CO2- R

    (a) Incomplete   (b) Inconsistent   (c) Noisy   (d) All of the above

3. Mention the approach to correlate two attributes   CO3- R

    (a) Chi-Square   (b) Pearson's Product (c) Cosine   (d) All of the above

4. The object doesn't comply with general properties is called as_____   CO4- R

    (a) Gird model   (b) Outlier   (c) Density model   (d) Cluster

5. Which data refers the sequence of primary type   CO5- R

    (a) Spatial   (b) Temporal   (c) Relational   (d) Entity

PART – B (5 x 3= 15 Marks)

6. Differentiate star and snowflake schema.   CO1- R

7. Mention the role of five number summaries.   CO2- R

8.   Define decision tree.   CO3- R

9. What do you mean by types of data in cluster analysis?   CO4- R

10. Write the role of temporal association role.   CO5- R

PART – C (5 x 16= 80 Marks)

11. (a) Explain the multidimensional data model operations with the   CO1- U   (16)
following dimensions. Employee-name,duration,location

Or

(b) (i) Compare OLTP and OLAP.   CO1-U   (6)

(ii) Discuss the dataware house three tier architecture with neat   CO1-U   (10)
diagram.

12. (a) (i) Write short notes on data mining functionalities.   CO2- U   (8)

(ii) Illustrate how the data mining system is integrated with the   CO2- U   (8)
data ware house system.

Or

(b) Explain the data preprocessing steps of data integration and data   CO2- U   (16)
reduction with suitable examples.

13. (a) Find all frequent item set using without candidate generation   CO3- App   (16)
method for the following data set. Let min_sup=30%

TID	Items
1	E,A,D,B
2	D,A,C,E,B
3	C,A,B,E
4	B,A,D
5	D
6	D,B
7	A,D,E
8	B,C

Or

(b) Apply back propagation classification algorithm with the three   CO3- App   (16)
input layer,two hidden layer and one output layer for multi layer
feed neural network. And also show net input, net output and
updated weight with the sample weight values.

14. (a) Explain k-means partition clustering analysis approach with   CO4- App   (16)
suitable examples and algorithms.

Or

(b) Discuss the gird based clustering method with the student   CO4- App   (16)
database.

2

55802

15. (a) Discuss the web content mining with suitable examples. CO5- U (16)

Or

(b) Explain spatial clustering algorithm with suitable examples. CO5- U (16)

55802