A

Question Paper Code: 55802

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

Fifth Semester

Information Technology

15UIT502 - DATA WAREHOUSING AND DATA MINING

(Regulation 2015)

Duration: Three hours   Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 1 = 10 Marks)

1. _____ is the heart of the warehouse.   CO1- R

   (a) Data mining database servers

   (b) Data warehouse database servers

   (c) Data mart database servers..

   (d) Relational data base servers.

2. The data is stored, retrieved & updated in _____   CO1- R

   (a) OLAB   (b) OLTP   (c) FTP   (d) SMTP

3. A _____ model identifies patterns or relationships   CO2- R

   (a) Descriptive.   (b) Predictive.

   (c) Regression.   (d) Time series analysis.

4. The various aspects of data mining methodologies   CO2- R
    is/are_____

   (i) Mining various and new kinds of knowledge

   (ii) Mining knowledge in multidimensional space

   (iii) Pattern evaluation and pattern or constraint-guided mining.

   (iv) Handling uncertainty, noise, or incompleteness of data

   (a) i, ii and iv only   (b) ii, iii and iv only   (c) i, ii and iii only   (d) All i, ii, iii and iv

5. Support Vector Machines (SVMs, also support vector networks) are   CO3- R

   (a) Supervised learning   (b) Unsupervised learning

   (c) Both a and b   (d) None of the above

6. _____ is used to map a data item to a real valued prediction   CO3- R
   variable.

   (a) Regression   (b) Time series analysis.

   (c) Prediction.   (d) Classification

7. An Outlier is a   CO4- R

   (a) Rare chance of occurrence within a given data set

   (b) Observation point that is distant from other observations.

   (c) Both a and b

   (d) None of the above

8. Incorrect or invalid data is known as _____..   CO4- R

   (a) changing data.   (b) noisy data.   (c) outliers.   (d) missing data

9. Temporal data mining refers to extract _____ information   CO5- R
   from large collections of temporal data

   (a) Useful abstract   (b) Non-trivial   (c) Implicit   (d) All the above

10. Web content mining describes the discovery of useful information   CO5- R
    from the _____contents.

    (a) Text   (b) Web   (c) Page   (d) Level

PART – B (5 x 2= 10Marks)

11. Compare OLTP and OLAP Systems.   CO1- R

12. Define Data mining.   CO2- R

13. How are association rules mined from large databases?   CO3-R

14. List the requirements of clustering in data mining.   CO4- R

15. What is meant by Spatial mining?   CO5- R

PART – C (5 x 16= 80 Marks)

16. (a) Define data warehouse. Explain its features. Diagrammatically   CO1- U   (16)
         illustrate and discuss the data warehouses architecture.

Or

55802

(b)	With a neat sketch, Explain three tier architecture and different models of data ware housing.	CO1- U	(16)

17. (a) Discuss about different types of data and functionalities. CO2 -U (16)

Or

(b) Analyze the various tasks and issues involved to Integrate a data Mining system with a Data Warehouse. CO2- Ana (16)

18. (a) What are Bayesian belief networks? Explain in detail about Bayesian Theorem. Give an example.. CO3- App (16)

Or

(b) Elaborate in detail about to mine closed frequent data item sets and Constraint Based Association Mining. CO3- Ana (16)

19. (a) Explain K-means clustering algorithm and Write the key issue in hierarchical clustering algorithm. CO4- U (16)

Or

(b) Explain outlier analysis in detail with an example. Discuss the use of outlier analysis. CO4- App (16)

20. (a) Explain the need and applications of Web mining and Web content mining in detail. CO5- U (16)

Or

(b) Explain in detail about Spatial clustering algorithm (SD). Give an example CO5- U (16)

55802

55802