Reg. No. : □□□□□□□□□□

Question Paper Code: 45025

B.E. / B.Tech. DEGREE EXAMINATION, NOV 2017

Fifth Semester

Computer Science and Engineering

14UCS505 - DATA WAREHOUSING AND DATA MINING

(Regulation 2014)

Duration: Three hours   Maximum: 100 Marks

Answer ALL Questions

PART A - (10 x 1 = 10 Marks)

1. Metadata contains

    (a) Migration tools
    (b) Tram formations details
    (c) Location and description of warehouse systems and data components
    (d) Cleanup information

2. The process of removing the deficiencies and loopholes in the data is called as

    (a) Aggregation of data   (b) Extracting of data
    (c) Cleaning up of data   (d) Loading of data

3. Which of the following tools a business intelligence system will have

    (a) OLAP tool   (b) Data mining tool
    (c) Reporting tool   (d) Both (a) and (b)

4. _____ has the ability to arrange rows, columns and cells in a fashion that facilitates analysis by intuitive visual presentation of analytical reports must exist.

    (a) Flexible reporting   (b) Generic dimensionality
    (c) Dynamic sparse martin handling   (d) Intuitive data manipulation

5. What are the criteria used for comparing classification and prediction

    (a) Predictive accuracy   (b) Robustness

    (c) Scalability and interoperability   (d) All the above

6. Spot the preprocessing technique that converts the data into appropriate forms of mining

    (a) Data cleaning   (b) Data transformation

    (c) Data reduction   (d) Data integration

7. The another name of agglomerative approach is

    (a) Bottom-up approach   (b) Top- down approach

    (c) Both (a) and (b)   (d) None of these

8. _____ attempts to form patterns that permit to to predict the next events given the available input data.

    (a) Classification   (b) Prediction

    (c) Correlation   (d) Association

9. Which of the following is/are applications of data mining?

    (a) Financial Data Analysis   (b) Retail Industry

    (c) Telecommunication Industry   (d) All the above

10. _____is a collection of data objects.

    (a) Cluster   (b) Supervised learning

    (c) Unsupervised learning   (d) Learning by observation

PART - B (5 x 2 = 10 Marks)

11. Define slice and dice operation.

12. List the various access types to the data stored in the data warehouse.

13. Define data mining and KKD.

14. What is tree pruning?

15. Explain the properties of clustering algorithm.

45025

PART - C (5 x 16 = 80 Marks)

16. (a) Explain in detail about metadata. Classify metadata and explain the same.   (16)

Or

(b) Explain about database architectures for parallel processing.   (16)

17. (a) Discuss Multidimensional Online Analytical Processing (MOLAP) and Multi Relational Online Analytical Processing (ROLAP) with relevant example.   (16)

Or

(b) Explain in detail about different types of OLAP tools.   (16)

18. (a) How data mining systems are classified? Discuss each classification with an example.   (16)

Or

(b) List and explain the primitive for specifying a data mining task. Also explain the various data pre-processing methods with suitable example.   (16)

19. (a) Explain in detail about constraint based association mining.   (16)

Or

(b) Explain the Naïve Bayes algorithm for solving classification problems. State the possible steps and conditions for effective analysis.   (16)

20. (a) Explain outlier analysis in detail with an example. Discuss the use of outlier analysis.   (16)

Or

(b) Describe in detail about outlier analysis.   (16)

_____

45025

45025