SUBJECT CODE: CS/ME/DM54

B.C.A. DEGREE EXAMINATION – NOVEMBER 2010
FIFTH SEMESTER

REG. NO. : _____

COURSE   : MAJOR ELECTIVE
PAPER   : DATA MINING AND WAREHOUSING
TIME   : 20 MINUTES   MAX. MARKS: 20

SECTION – A   **(20X1=20)**

I. Choose the correct answer:

1. A common source for data is _____.
    a. Data mart   b. Dataset   c. Raw Data

2. _____ is a typical method belonging to Density – based approach.
    a. CLARANS   b. Diana   c. DenClue

3. Which analysis is used to find similarities between data?
    a. Genetics   b. Cluster   c. Pattern

4. _____ method belongs to Clustering algorithms.
    a. Partitioning Method   b. CLARANS method   c. DenClue method

5. Which Algorithm is for mining distance – based outliers?
    a. Nested loop algorithm
    b. Nearest Neighbor Algorithm
    c. Cluster Algorithm

6. What does Naïve Bayes Algorithm generate?
    a. Mining Models   b. Leaf Nodes   c. Decision Nodes

7. MOLAP data is stored in form of _____.
    a. Multidimensional Cubes   b. Relational Database   c. None

8. Which one is customer-oriented and used for data analysis and querying by clerks,
    clients and IT professionals?
    a. OLTP   b. OLAP   c. SCD

9. In which scheme all dimensions will be linked directly with a fat table.
    a. Star schema   b. Snow schema   c. None of These

10. Schemas in Data Warehouse are _____
    a. De Normalized   b. Normalized   c. All the above

..2

II. Fill in the blanks:

11. In preprocessing, vectors are divided into two sets, the _____ and the _____.

12. Instead of taking the mean value of the object in a cluster as a reference point, _____ can be used.

13. Spatial data mining is the application to find pattern in _____.

14. Expand OLAP_____

15. _____ groups objects based on their similarity and has wide applications.

III. State True or False

16. Data mining is seen as an increasingly important tool by modern business to transform data.

17. In Partitioning approach one of the typical methods is Diana.

18. Semi-tight Coupling is one of the data mining architecture.

19. Outlier detection and analysis are very useful for fraud detection.

20. A GIS is a computer system capable of capturing, storing, analyzing, and displaying geographically referenced information.

****

B.C.A. DEGREE EXAMINATION – NOVEMBER 2010
FIFTH SEMESTER

COURSE   : MAJOR ELECTIVE
PAPER   : DATA MINING AND WAREHOUSING
TIME   : 2 HOURS & 40 MINUTES   MAX. MARKS: 80

SECTION – B   (8X5=40)

Answer any eight questions:

1. Explain Data mining task primitives with example.

2. Give an overview of performance issues in Data mining?

3. Briefly explain Genetic Algorithms.

4. Describe Data Warehouse Architecture with neat diagram.

5. Briefly explain issues regarding classification and prediction.

6. Write about Mining associations.

7. Explain the types of data in cluster analysis.

8. Briefly explain about text mining.

9. Explain the use of Data mining intrusion detection.

10. Give an account on social impacts of Data mining.

SECTION – C   (4X10=40)

Answer any four questions:

11. Explain the following:-

    a.   Data Cleaning

    b.   Data integration and transformation

12. Explain the various schemes for multi dimensional databases.

13. Compare OLAP and OLTP Systems.

14. What is prediction? Discuss the issues regarding prediction.

15. Write a note on Cluster Analysis.

16. Explain Data mining Applications for Financial data analysis.

****