ANNA UNIVERSITY COIMBATORE

B.E. / B.TECH. DEGREE EXAMINATIONS : MAY / JUNE 2010

REGULATIONS – 2007

SIXTH SEMESTER

070230060 - DATA WAREHOUSING AND MINING

(COMMON TO CSE / IT)

TIME : 3 Hours Max.Marks : 100

PART – A

(20 x 2 = 40 MARKS)

ANSWER ALL QUESTIONS

1. What is a data warehouse?
2. Differentiate fact table and dimension table.
3. Briefly discuss the schemas for Multidimensional Databases.
4. Compare OLTP and OLAP.
5. List the issues to be considered during Data Integration.
6. Write the strategies for data reduction.
7. Why is it important to have data mining query language?
8. Write the syntax for characterization.
9. List the techniques to improve the efficiency of Apriroi algorithm.
10. Define support and confidence.
11. What is FP growth?
12. How meta rules are useful in Constraint-based association mining.
13. What is Bayesian theorem?
14. Why is tree pruning useful in decision tree induction?
15. Give the difference between agglomerative and divisive hierarchical clustering.
16. Define Outliers. List various outlier detection approaches.
17. List out the methods for information retrieval.
18. What is Web usage mining?

19. Give some applications of data mining.
20. What is a time-series database?

PART – B

(5 x 12 = 60 MARKS)

ANSWER ANY FIVE QUESTIONS

21. Describe the data warehouse architecture with a neat diagram.

22. Explain the steps involved in Attribute-Oriented Induction for Data Characterization.

23. Discuss how Apriori algorithm is used to find frequent itemsets with an example.

24. How does data classification work? Discuss the major steps of back propagation algorithm with an example.

25. What is clustering? List the types of clustering techniques and discuss the Partitioning methods in detail.

26. Describe the various data mining primitives for specifying a data mining task.

27. a) List out the various issues to be considered in data mining. 6

    b) Describe how to evaluate the accuracy of the classifier and increase the classifier accuracy. 6

28. What is Web mining? Explain how to identify authoritative web pages and automatic classification of web documents.

*****THE END*****