B.E./B.Tech. DEGREE EXAMINATION, NOVEMBER/DECEMBER 2014.

Seventh Semester

Computer Science and Engineering

CS 2032/CS 701/10144 CSE 32 — DATA WAREHOUSING AND DATA MINING

(Common to Sixth Semester Information Technology)

(Regulation 2008/2010)

(Common to PTCS 2032/10144 CSE 32 – Data Warehousing and Data Mining for B.E. (Part-Time) Sixth Semester – Computer Science and Engineering – Regulation 2009/2010)

Time : Three hours Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. What is Star-schema?
2. Define Metadata with an example.
3. Define the categories of tools in business analysis.
4. What is Virtual warehouse?
5. What are the types of data?
6. What is Meta learning?
7. Define the term interestingness of patterns.
8. Define Lazy learners.
9. What are the Tree pruning methods?
10. Define Divisive Hierarchical Clustering.

PART B — (5 × 16 = 80 marks)

11. (a) (i) Explain the evolution of Database technology. (8)

(ii) Explain the process of mapping the data warehouse to multiprocessor architecture. (8)

Or

(b) (i) Explain the different types of data repositories on which mining can be performed? (8)

(ii) Explain the architecture of data warehouse. (8)

12. (a) (i) Design multi-dimensional data model for hospital data warehouse , consist three dimensions time, doctor, and patient and the two measures count and charge, where charge is a fee that a doctor charges a patients for a visit. (3+4)

(1) Enumerate three classes of schema that are popularly used for modeling data warehouses.

(2) Draw a schema diagram for the above data warehouse using all of the schema classes listed in (1).

(ii) How to reduce the size of the fact table? Explain with an example. (9)

Or

(b) (i) List out the OLAP operations in multidimensional data model and explain with an example (8)

(ii) Differentiate OLAP and OLTP (8)

13. (a) Explain the various data mining issues and functionalities in detail.

Or

(b) State and explain the various classification of data mining systems with example.

14. (a) Explain the apriori algorithm for finding frequent item sets with suitable example.

Or

(b) Explain Bayesian classification and Rule based classification. Give example for any one classification and explain in detail.

15. (a) Explain Hierarchical method and density based method of classifications with example.

Or

(b) (i) Explain the types of data in cluster analysis in detail with example. (8)

(ii) Explain Outlier analysis with example. (8)