Code No. 2403 / N

FACULTY OF INFORMATICS

B.E. 3/4 (IT) II – Semester (New) (Main) Examination, May 2013

Subject : Data Warehousing and Data Mining (Elective – I)

Time : 3 hours **Max. Marks : 75**

Note : Answer all questions from Part-A. Answer any FIVE questions from Part-B.

PART – A (25 Marks)

1. What is a data warehouse? And what is its use? 2
2. State the purpose of data cleaning address the two issues in data cleaning. 3
3. What is pattern-interestingness? 2
4. Define frequent item set. How do you determine support and confidence? 3
5. What is decision tree induction? 2
6. What is a decision boundary in support vector machine? 3
7. Define any two similarity measures used in cluster analysis. 2
8. What is an outlier? Mention one use of an outlier. 3
9. Define precision, recall and F-score. 3
10. What is dimensionality reduction? 2

PART – B (50 Marks)

11. Describe knowledge discovery process.
12. Explain the use of wavelet transforms in dimensionality reduction.
13. Sketch an integrated OLAM and OLAP architecture and explain its working.
14. Explain induction of a decision tree using information gain.
15. State k-means clustering algorithm and explain the process of clustering.
16. What are data streams? Explain random sampling method.
17. Write short notes on :
    a) Spatial data mining b) Text mining

******