Code No. 3168

FACULTY OF INFORMATICS

B. E. 4/4 (IT) I-Semester (**Supplementary**) Examination, June / July 2011

Subject : **Data Watehousing & Data Mining** (Elective-II)

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 are various stages of KDD process? 3
2. What does a data warehouse provide for business analysis? 3
3. Distinguish between Fact table and Dimension table. 3
4. What is meta data? 2
5. What are various kinds of association rules? 3
6. What are the disadvantages of apriori algorithm? 2
7. What is outlier data? 2
8. Write the advantages of Grid based clustering methods. 3
9. What is meant by accuracy of a classifier? 2
10. Draw the star schema of a multidimensional data model. 2

Part – B (5 x 10 = 50 Marks)

11.a) Breifly explain the architecture of a typical data mining system. 5
b) Explain about data cleaning process. 5

12.a) Explain about three-tier architecture of a data warehouse. 5
b) Explain about various kinds of OLAP operations with an example. 5

13. Write the FP growth algorithm. Using this algorithm compute the frequent patterns for the following transational database. 10

TID	Item-list
T100	C, M, S, B
T200	S, M, A, P
T300	A, M, B, P
T400	B, M, S

Min-support = 60%

14. Explain DBSCAN and BIRCH algorithms. 10
15. Explain back propagation algorithm for classification. 10

16.a) Explain K-Means method for partitioning. 5
b) Explain Text Mining. 5

17. Write short notes on : 10
a) Snow flake and fact constellation schemas
b) Decision tree pruning