

Roll No .....

**CS - 8203****B.E. VIII Semester**

Examination, June 2015

**Data Mining and Knowledge Discovery****Time : Three Hours****Maximum Marks : 70****Note :** i) Attempt one question from each unit.

ii) All questions carry equal marks.

**Unit - I**

1. a) Describe various methods used for cleaning data before it is loaded in the data warehouse.
- b) What is data marts? write the difference between Independents, dependent and distributed data marts.

Or

2. a) What is a data warehouse? discuss most important issues in data warehouse implementation.
- b) Explain and compare the following by suitable example.
  - i) Fact constellations
  - ii) Star schema.

**Unit - II**

3. a) Explain OLAP operations in multidimensional data model. Differentiate ROLAP, MOLAP and HOLAP.
- b) Differentiate between OLTP and OLAP.

Or

4. a) What is multidimensional data model? Discuss 2-D, 3-D and 4-D representation of data cube.
- b) How is data warehouse different from a database? How they are similar?

**Unit - III**

5. a) What is Knowledge discovery in data bases? Why it is important and what are its applications?
- b) Explain the various strategies of data reduction.

Or

6. a) Describe the issues to be considered during data integration.
- b) What do you mean by data preprocessing? Why it is needed?

**Unit - IV**

7. a) How do association rules differ from Traditional production rules? explain.
- b) Explain Apriori algorithm with examples.

Or

8. a) Write an algorithm for discovering itemsets without candidate generation.
- b) What is frequent itemset? Explain maximal itemset and close itemset with example.

**Unit - V**

9. a) Explain decision tree method for data classification give the suitable example.
- b) How many types of clustering methods are there? Explain any one partitioning clustering algorithm.

Or

10. Write short notes on (any two):

- i) Gini Index
- ii) Naive Bayes Method
- iii) Density - Based Methods.