

Roll No

MCA - 501

MCA. V Semester

Examination, December 2016

Data Warehousing and Mining

Time : Three Hours

Maximum Marks : 70

- Note:** i) Answer five questions. In each question part A, B, C is compulsory and D part has internal choice.
ii) All parts of each question are to be attempted at one place.
iii) All questions carry equal marks, out of which part A and B (Max. 50 words) carry 2 marks, part C (Max. 100 words) carry 3 marks, part D (Max. 400 words) carry 7 marks.
iv) Except numericals, Derivation, Design and Drawing etc.

Unit - I

1. a) What do you understand by KDD?
b) Write down the applications of data mining.
c) Describe the major challenges in data mining.
d) Discuss major issues in data mining regarding mining methodology and user interaction.

OR

What kinds of resources are typically required to undertake mining? Explain.

Unit - II

2. a) What do you mean by data warehouse backend tools and utilities.
b) Draw a snowflake schema diagram for the data warehouse.
c) How is a data warehouse different from a database? How are they similar? Explain.
d) What do you understand by the term OLAP? List the various types of OLAP operations.

MCA-501

PTO

[2]

OR

What do you understand by Data marts? Also explain dependent, independent and distributed data marts.

Unit - III

3. a) What is analytical characterization?
b) What is meant by dimensionality reduction?
c) What do you understand by discretization and concept hierarchy generation for numerical data?
d) List and describe the five primitives for specifying a data mining task.

OR

What do you understand by "data cleaning as a process"? Give the approaches for data cleaning.

Unit - IV

4. a) What are multidimensional association rule?
b) What is frequent itemset?
c) What is constraints-Based association mining? Explain.
d) What is Apriori algorithm? Give few techniques to improve the efficiency of Apriori algorithm.

OR

Explain association rule in mathematical notation. Define support and confidence in association rule mining.

Unit - V

5. a) What are major Clustering methods?
b) What do you mean by outlier?
c) Specify the five criteria for the evaluation of classification and prediction.
d) What are different classification methods? Explain.

OR

Write short notes on :

- i) Gini Index
- ii) Support Vector Machine.
- iii) Tree Induction Algorithm.

MCA-501