

Printed Pages – 4

Roll No. :

576464(76)

676664(76)

M. B. A. (Fourth Semester) Examination, 2020

APR-MAY

(New Scheme)

(Management Branch)

(Specialization : Systems Management)

DATA WAREHOUSING and DATA MINING (New)

Time Allowed : Three hours

Maximum Marks : 80

Minimum Pass Marks : 32

Note : Answer any two parts of each question. All questions carry equal marks.

Unit-I

1. (a) Describe the knowledge discovery process used in databases. List essential stages of KDD.

- (b) Why do we need data mining, discuss? Describe various types of databases and scope of data mining therein.
- (c) Differentiate between :
 - (i) discrimination and characterization
 - (ii) classification and clustering

Unit-II

- 2. (a) Explain the typical Data Warehouse architecture in detail.
- (b) Describe the characteristic features of a data warehouse with examples.
- (c) Explain the star, snowflake and fact constellation schemes for representing multidimensional databases taking suitable examples.

Unit-III

- 3. (a) Discuss issues to consider during data integration.
- (b) List reasons why you think data quality is critical in a data warehouse? Why preprocessing is needed?

(c) Describe various techniques for Data Cleaning.

Unit - IV

4. (a) What is Market basket analysis?
 (b) Given are the following five transactions on items
 { M, O, N, K, E, Y, D, A, U, C, I } :

Transaction ID	Items Bought
T1	{ M, O, N, K, E, Y }
T2	{ D, O, N, K, E, Y }
T3	{ M, A, K, E }
T4	{ M, U, C, K, Y }
T5	{ C, O, O, K, I, E }

Use Apriori algorithm to compute all frequent item sets and their support, with minimum support as 3.

- (c) Describe different types of variables used in clustering.

Unit-V

5. (a) What are different applications and trends in Data Mining?

[4]

(b) Explain the various Categories of Web Mining along with their real time applications.

(c) Write short notes on :

(i) Spatial Data Mining

(ii) Multimedia Data Mining