

Reg. No.

--	--	--	--	--	--	--	--	--	--



**MCA 503**

**V Semester M.C.A. Degree Examination, December 2018**  
**COMPUTER APPLICATIONS**  
**Data Mining Techniques**

Time : 3 Hours

Max. Marks : 75

**Instructions :** Answer **any five** of the following questions. **All** questions carry **equal** marks.

1. a) Briefly describe various database on which data mining could be applied. **(7+8)**  
b) Differentiate data characterization and data discrimination.
2. a) Describe the characteristics of normally distributed data and skew data. **(7+8)**  
b) Suppose that a set of data is grouped into intervals. The intervals and corresponding frequencies are as follows. Compute the approximate median value for the data.

<b>Age</b>	<b>Frequency</b>
1 – 5	200
6 – 15	450
16 – 20	300
21 – 50	1500
51 – 80	700
81 – 110	44

3. a) Distinguish between Distributive measures and Holistic measures. **(7+8)**  
b) What is data discretization ? How numeric data sets  $D : [0, \dots, N_s]$ , where  $N_s$  is the total number of seconds in a year (365 days), can be represented with different concept hierarchies ?
4. a) Define Data warehouse. Describe its characteristics, benefits and limitation. **(7+8)**  
b) Explain the following concepts with examples :
  - i) Snowflake schema
  - ii) Fact constellation schema.

P.T.O.



5. a) Describe various operations that are associated with OLAP Model. **(7+8)**  
b) Briefly explain the data mining query language primitives with an example in detail.
6. a) Write a neat diagram and explain the three-tier data warehouse architecture. **(7+8)**  
b) Explain how data warehousing and OLAP relate to data mining using the integrated OLAM and OLAP architecture.
7. a) Briefly outline the major steps of decision tree classification. **(7+8)**  
b) Explain the back propagation algorithm for classification.
8. a) What is a prediction ? Distinguish between linear regression and nonlinear regression with examples. **(7+8)**  
b) What is clustering ? Explain the hierarchical clustering method in details.
9. Write short notes on the following : **(7+8)**  
a) Distance metrics  
b) Predictor error measures.
-