

Roll No. ....

Subject Code—2091

**M.C.A. (5th Year) EXAMINATION**

(5 Years Integrated Course)

MCA-505

**DATA MINING AND DATA WAREHOUSING**

*Time : 3 Hours*

*Maximum Marks : 100*

**Note :** Attempt any *Five* questions. All questions carry equal marks.

1. What is Data Mining ? Explain how the evolution of database technology led to data mining. Describe the steps involved in data mining when viewed as a process of knowledge discovery. 20
2. Write the Apriori algorithm for discovering frequent item sets for mining single dimensional Boolean Association Rule and discuss various approaches to improve its efficiency. 20

3. Why is Outlier Mining Important ? Briefly describe the different approaches behind statistical based outlier detection, distance-based outlier detection and deviation-based outlier detection. 20
  
4. Explain the different types of classifications such as Bayesian and classification by back propagation. 20
  
5. (a) Discuss the social impacts of data mining. 10
  
- (b) Explain application of data mining in financial sector. 10
  
6. Briefly compare the following concepts. Use an example to explain your points :
  - (a) Snowflake schema, fact constellation, starnet query model 7
  
  - (b) Data cleaning, data transformation, refresh 7
  
  - (c) Discovery driven cube, multifeature cube, virtual warehouse. 6

7. (a) Discuss the essential features of a typical data mining query language like DMQL.
- (b) List and describe any four primitives for specifying a data mining task. 10+10
8. Write short notes on the following :
- (a) Data Reduction
- (b) Data Integration
- (c) Architecture of Data Mining Systems. 5+5+10