

- Q.18 What is data reduction ? Discuss its techniques in brief? (CO4)
- Q.19 What is data cleaning ? Why it is useful ? (CO3)
- Q.20 What is frequent itemset and closed itemset ?(CO4)
- Q.21 Explain Apriori algorithm in brief. (CO3)
- Q.22 Define cluster analysis ? What are requirements of cluster analysis ? (CO4)

#### **SECTION-D**

**Note:** Long answer type questions. Attempt any two questions out of three questions. (2x8=16)

- Q.23 Explain the difference between operational database systems and data warehouse in detail. (CO1)
- Q.24 Explain various tasks involved in data mining in detail . Also discuss its benefits and functionalities in detail. (CO4)
- Q.25 Discuss the KDD (Knowledge Discovery in Databases ) process with suitable diagram in detail (CO4)

No. of Printed Pages : 4 223864C  
Roll No. ....

#### **6th Sem. / Artificial Intelligence & Machine Learning**

#### **Subject : Data Warehousing and Data Mining**

Time : 3 Hrs. M.M. : 60

#### **SECTION-A**

**Note:** Multiple choice questions. All questions are compulsory (6x1=6)

- Q.1 Data Mining is also referred to as \_\_\_\_\_. (Co1)  
 a) Knowledge discovery in databases  
 b) Data Cleaning  
 c) Data extraction  
 d) Data management
- Q.2 Which of the following is a subset of data warehouse focused on a specific functional area ? (CO1)  
 a) Data mart                  b) Association rules  
 c) Flat files                  d) Database

Q.3 OLAP stands for . (CO2)

- a) On-Line Application Processing
- b) On-Line Analytical Processing
- c) On-Line Ability Processing
- d) None of the above

Q.4 Who is called the Father of Data Warehouse ?(CO1)

- a) Charles Babbage b) Ralph kimball
- c) Bill Inmon d) Fritz Bauer

Q.5 Which of the following is not an issue in data mining?  
(CO3)

- a) High \dimensionality b) Shortage of data
- c) Overfitting d) Outliers

Q.6 \_\_\_\_\_ servers support multidimensional views of data through array-based multidimensional storage engines. (CO4)

- a) ROLAP b) MOLAP
- c) Data warehouse d) Database

## SECTION-B

**Note:** Objective/ Completion type questions. All questions are compulsory. (6x1=6)

Q.7 What is the full form of ROLAP, MOLAP and HOLAP? (CO2)

Q.8 Define a virtual warehouse . (CO1)

Q.9 What is a fact table ? (CO3)

Q.10 What is the issue of outliers in data mining ?(CO4)

Q.11 What do you mean by discretization ? (CO4)

Q.12 Define data preprocessing. (CO3)

## SECTION-C

**Note:** Short answer type questions. Attempt any eight questions out of ten questions. (8x4=32)

Q.13 Discuss major characteristics of data warehouse . (CO1)

Q.14 Explain any four applications of data warehouse (CO1).

Q.15 Explain star schema and snow-flake schema in brief. (CO1)

Q.16 Explain OLAP server architecture in brief. (CO2)

Q.17 What are the major issue in data mining ? (CO4)