

7. What is Data Pre-processing ? What are the various Steps of Data Pre-processing ? Discuss about data cleaning techniques. **16**

(181-EL-404321)

Roll No. ....

**PAPER ID—20068**

**Unit IV**

8. What is the process of Classification ? Compare it with prediction. Explain induction tree method of classification. **16**

9. Explain the following :  
(a) Hierarchical Clustering  
(b) Spatial Data Mining. **16**

**M. Sc. (Computer Science**

**EXAMINATION, 2023**

(Fourth Semester)

**DATA WAREHOUSE AND DATA MINING**

**Code : 17MCS24DA1**

*Time : 3 Hours*

*Maximum Marks : 80*

---

Before answering the question-paper candidates should ensure that they have been supplied to correct and complete question-paper. No complaint, in this regard, will be entertained after the examination.

---

**Note :** Attempt *Five* questions in all, selecting *one* question from each Unit. Q. No. 1 is compulsory. All questions carry equal marks.

1. (a) Define data warehouse and its characteristics.

- (b) What is the ETL process in warehouse construction ?
- (c) What is Curse of Dimensionality Problem ?
- (d) What is an outlier ?
- (e) Define confidence of an association rule.
- (f) How binning can handle noisy data ?
- (g) Define classifier accuracy.
- (h) How is prediction different from classification ?

**16**

### **Unit I**

2. Explain about various types of data warehouse in detail-Enterprise Data Warehouse, Virtual warehouse and Data Marts. Also differentiate between Enterprise Data warehouse and Virtual Warehouse.
3. Explain about various warehouse schemas-- star, snowflake, galaxy with the help of diagrammatical notation in detail. Out of these which one is the best and why.

**16**

### **Unit II**

4. (a) Explain various OLAP operations with the help of illustrative examples. **8**
- (b) Contrast the differences between OLAP and OLTP systems. **8**
5. (a) What is a Data Cube and a lattice ? Explain with the help of example. What is efficient computation of Data Cubes ? **8**
- (b) What do you understand by ROLAP servers ? How do you compare it with MOLAP and HOLAP Servers along with justifications ? **8**

### **Unit III**

6. (a) Describe the various phases in knowledge discovery process with a neat diagram. **8**
- (b) What are the various methods of mining Multi-level Association rule mining from relational databases and data warehouse ? **8**