

DATA MINING & WARE HOUSING

Time : 3 Hours

Min. Passing Marks : 24

Maximum Marks : 80

Instruction to Candidates :

Attempt any five questions, selecting one question from each unit. All questions carry equal marks. (Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used/calculated must be stated clearly.)

< Unit-I >**OR**

1. (a) What do you mean by data mining describe it and differentiate between data ware house and data mining. [8]
- (b) What do you mean by data processing? and write short notes on the following
 - (i) Regression
 - (ii) Clustering

[8]

OR

1. (a) What do you mean by data reduction and what are the different process of data reduction? [8]
- (b) Explain data discretization and concept hierarchy generation. [8]

< Unit-II >

2. Write short notes on

- (i) Quartiles
- (ii) Range
- (iii) Outliers
- (iv) Boxplots

[16]

OR

2. Write short notes on

- (i) Generalized association rules
- (ii) Multilevel association rules

[16]

< Unit-III >

3. (a) What are the different categories of classification rules. [8]
- (b) What do you mean by information gain and how it is calculated. [8]

3. (a) What are the different classification techniques. [8]
- (b) What are advantages and disadvantages of decision tree approach over other approaches of data mining. [8]

< Unit-IV >

4. (a) What is data warehousing? And explain the architecture of data warehousing? [8]
- (b) What are the different data processing models. [8]

OR

4. (a) Explain multidimensional model of a data warehouse. [8]
- (b) Write short notes on : [8]
 - (i) Star schema
 - (ii) Snone flake schema

[8]

< Unit-V >

5. What do you mean by aggregation? and explain different aggregate functions. [16]

[16]

OR

5. (a) Explain in brief how the OLAP handles aggregation [8]
- (b) Write short note on [8]
 - (i) ROLAP
 - (ii) OLAP
 - (iii) MOLAP
 - (iv) HOLAP

[8]

[8]