

<b>7E4243</b>	Roll No. : _____	Total Printed Pages : <b>3</b>
	<b>7E4243</b>	
B. Tech. (Sem. VII) (Main/Back) Examination, December-2012		
Computer Engg.		
7CS6.2 Data Mining & Ware Housing		

Time : 3 Hours]

[Total Marks : 80  
[Min. Passing Marks : 24

*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.*

Use of following supporting material is permitted during examination.  
(Mentioned in form No. 205)

1. \_\_\_\_\_ Nil \_\_\_\_\_

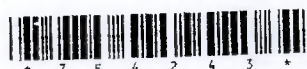
2. \_\_\_\_\_ Nil \_\_\_\_\_

**UNIT - I**

- 1 (a) What are the different phases of the knowledge discovery from database ? 8
- (b) Write short notes on :
- (i) Data cleaning
- (ii) Noisy data. 4+4

**OR**

- 1 (a) What is data integration and transformation ? 6
- (b) Differentiate between data mining and data warehouse. 4
- (c) What is data cube aggregation ? Explain. 6



UNIT - II

- 2 (a) What do you understand by the measurement of central tendency ? Explain Apriori algorithm. 8
- (b) What are data mining association rules ? Explain. 8

OR

- 2 (a) Describe any one association rules algorithm with example. 10
- (b) What do you mean by Data Generalization ? Explain. 6

UNIT - III

- 3 (a) Describe Neural Network Techniques for Data Mining. Also explain the relation between decision trees and classification rules. 10
- (b) Write a short note on Hierarchical clustering. 6

OR

- 3 (a) What do you mean by Genetic Algorithm ? How hypothesis testing and refinement task can be done using Genetic Algorithm. 8
- (b) Write short notes on :  
 (i) STING  
 (ii) CLIQUE. 4+4

UNIT - IV

- 4 (a) What do you mean by schema ? Discuss various schema used in data ware housing. 8
- (b) What are data cubes and snow flakes ? Explain. 8

OR





- 4 Write short notes on :
- (i) 3-Tier Architecture
  - (ii) Multi-Dimensional Data model
  - (iii) Stars
  - (iv) Delivery Process.

4×4=16

UNIT - V

- 5 (a) What is OLAP ? Describe different ways to control Aggregation problem.
- (b) What is Regression ? Explain the linear regression model.

8

8

OR

- 5 Explain briefly :
- (i) Aggregate Tables
  - (ii) Good fit query tools
  - (iii) MOLAP
  - (iv) Data warehouse tuning.

4×4=16

---