

**FACULTY OF ENGINEERING**  
**B.E. 4/4 (CSE) II-Semester (Main) Examination, April / May 2013**

**Subject : Data Mining**

**Time : 3 Hours**

**Max. Marks: 75**

**Note: Answer all questions of Part - A and answer any five questions from Part-B.**

**PART – A (25 Marks)**

1. What is Data Mining? (2)
2. Give techniques for Data Transformation. (3)
3. Define support and confidence. (2)
4. Define Linear Regression (2)
5. What is bit map indexing and Join indexing methods to speed up query processing? (3)
6. Distinguish Data mart and Data Warehouse. (2)
7. Differentiate snow flake and Fact constellations schemas for multi dimensional databases. (3)
8. What is the role of Meta data repository in a Data Warehouse? (2)
9. What is binning? Smooth the following data using by bin means and by bin boundaries 4, 8, 15, 21, 24, 25, 28, 34 (3)
10. What is Web Mining? (3)

**PART – B (5x10=50 Marks)**

- 11.(a) Explain various Data Mining functionalities with examples. (5)  
 (b) Discuss various issues in Data Mining. (5)
- 12.(a) Draw three-tier data Warehouse Architecture and explain various levels in the architecture. (6)  
 (b) Differentiate OLAP and OLTP. (4)
- 13.(a) Explain various Data Mining Primitives. (7)  
 (b) Define (i) Loose coupling (ii) Semi tight coupling (iii) tight coupling architecture of data mining (3)
14. Write and explain Apriori algorithm to find all frequent item sets and strong association rules for the following database, where min\_sup=60% and min\_conf=80%. (10)

Tid	Items
T100	{K, A, D, B}
T200	{D, A, C, E, B}
T300	{C, A, B, E}
T400	{B, A, D}

- 15.(a) Explain in detail Bayesian Classifier. (7)  
 (b) Explain how classifier accuracy can be estimated. (3)
16. Explain K-means algorithm for the following example to find the cluster where k =3 and initial points are assigned to clusters as  $C_1 = \{A_1, A_2, A_3\}$ ,  $C_2 = \{A_4, A_5, A_6\}$ ,  $C_3 = \{A_7, A_8\}$  (10)

	X	Y
A <sub>1</sub>	2	10
A <sub>2</sub>	2	5
A <sub>3</sub>	8	4
A <sub>4</sub>	5	8
A <sub>5</sub>	7	5
A <sub>6</sub>	6	4
A <sub>7</sub>	1	2
A <sub>8</sub>	4	9

17. Write short notes on any two of the following: (10)
  - (a) Sequential pattern mining
  - (b) Mining multimedia databases
  - (c) Mining text databases