

Code: IT5T1

III B.Tech - I Semester – Regular Examinations - November 2014

**DATA WARE HOUSING & DATA MINING
(INFORMATION TECHNOLOGY)**

Duration: 3 hours

Marks: 5x14=70

Answer any FIVE questions. All questions carry equal marks

1. a) Define Data Mining? Explain in brief about the issues in Data Mining Systems? 7 M
b) Why do we need Data Preprocessing? Explain in brief about major tasks in Data Preprocessing? 7 M
2. a) What is a Data warehouse? Explain in brief about Snow Flake Schema? 7 M
b) Explain in brief about Data cube? 7 M
3. a) Illustrate the different Schemas for Multi Dimensional Data base? 7 M
b) Describe the various descriptive statistical measures for Data Mining? 7 M
4. a) Explain in brief about Apriori Algorithm? 7 M
b) Explain in brief about Constraint based Association Mining? 7 M

5. a) What is Classification? Explain how classification is done using Decision Tree Induction? 7 M
- b) Explain in brief about Rule based Classification method? 7 M
6. What is K-means algorithm? Suppose that the data mining task is to cluster the following eight points in to three clusters $A_1(2,10)$ $A_2(2,5)$ $A_3(8,4)$ $B_1(5,8)$ $B_2(7,5)$ $B_3(6,4)$ $C_1(1,2)$ $C_2(4,9)$.
The distance function is Euclidean distance. Suppose initially A_1 , B_1 , C_1 are assigned as the center of each cluster, respectively. Use K means algorithms to show only
- i) The Three cluster centers after the first round of execution and
 - ii) The final Three clusters 14 M
7. a) What is Outlier? Explain in brief about Outlier detection methods? 7 M
- b) Explain in brief about Grid based methods? 7 M
8. a) Explain in brief about Mining Complex Data types? 7 M
- b) Explain in brief about the Application of Data mining for Intrusion detection? 7 M