

M. Tech II Semester Regular/Suppl. Examinations, July 2016

DATA MINING AND DATAWAREHOUSING

(Computer Science Engineering)

Time: 3 Hours

Max. Marks: 60

Note: Answer any **FIVE** questions. All Questions carry equal Marks.

5 × 12 = 60M

1. a) Explain motivating challenges of data mining. (6M)
 b) Discuss the issues to consider during data integration. (6M)
2. a) Write measures of similarity and dissimilarity for simple attributes. (5M)
 b) Explain about various OLAP operations. (7M)
3. a) Explain Bayesian Classification with relevant example. (7M)
 b) Write algorithm for decision tree induction. (5M)
4. Given the following transaction data:

TID	Items
1	{Jacket, Boots}
2	{Milk, Cheese, Bread, Shoes}
3	{Cloth, Bread}
4	{Milk, Bread, Shoes, Pork, Apple, Soup}
5	{Cheese, Shoes, Beef}
6	{Jacket, Bread, SkiPants}

- a) What are the supports and confidences of the following two rules? (4M)
 Rule1: Milk → Bread
 Rule2: Bread → Milk
- b) Find Frequent item sets using Apriori if minimum support is 30% and confidence is 60% (8M)
5. a) What is Hierarchical method of clustering? Explain Agglomerative Hierarchical Clustering? (8M)
 b) Write short notes on types of clustering (4M)
6. a) Write about bootstrap alternative techniques (6M)
 b) Write short notes on data preprocessing (6M)

7. a) Explain FP-Growth algorithm with suitable example. (7M)
b) Write short notes on Bayesian Belief network. (5M)
8. a) Explain K-means clustering and state its strengths and weaknesses (7M)
b) Explain compact representation of frequent itemsets. (5M)
