

Code No: **R41053****R10****Set No. 1****IV B.Tech I Semester Supplementary Examinations, October/November - 2017****DATA WARE HOUSING AND DATA MINING****(Common to Computer Science and Engineering and Information Technology)****Time: 3 hours****Max. Marks: 75****Answer any FIVE Questions
All Questions carry equal marks**

- 1 a) What are the different measures to assess the quality of Data? [8]
b) What are various types of datasets? Explain. [7]
- 2 a) What are various methods for data reduction? Explain with examples. [8]
b) Mention different measures of similarity. [7]
- 3 a) Describe efficient processing of OLAP queries [8]
b) Write about Data warehouse implementation. [7]
- 4 a) Explain various methods to evaluate the performance of a classifier. [8]
b) Write an algorithm for decision tree induction. [7]
- 5 a) What is meant by Bayesian Classifier? How Bayes theorem is used for classification? [8]
b) Compare belief network and Naïve Bayes classifier. [7]
- 6 a) What is meant by frequent item set generation? where it can be used.? [8]
b) Write the apriori algorithm. [7]
- 7 a) What do you mean by clustering? Explain different types of clusters. [8]
b) Write the algorithm of bisecting K-means and how it is different from simple K-means? [7]
- 8 Write short notes on
a) DBSCAN
b) Hierarchical clustering
c) Agglomerative Clustering [15]