

Subject Code: H0501/R13

M. Tech –II Semester Regular/ Supply Examinations, October, 2015

DATA WAREHOUSING AND DATA MINING.

(Common to CS and CS&E)

Time: 3 Hours

Max Marks: 60

Answer any FIVE questions

All questions carry EQUAL marks

1. a) What are the Steps involved in data preprocessing
b) What is the Data quality? Discuss.
2. a) What is the linear SVM? How it is used in classification?
b) Discuss the model over-fitting.
3. a) Explain the concept hierarchy.
b) Discuss the Apriori algorithm with an example.
4. a) Discuss the Naïve Bayesian Classification with an example.
b) Explain the ANN (Artificial Neural Networks) classification.
5. a) Discuss the F-P Growth Algorithm with an example
b) Discuss the cluster evaluation
6. a) Explain the Agglomerative Hierarchical clustering with an example
b) Explain the DB Scan clustering
7. a) Discuss the Proximity based outlier detection
b) Discuss the Density based outlier detection.
8. Write short notes on the following
 - a) Web usage mining
 - b) search engines,
