

Code No: MC1643/R16

MCA IV Semester Regular/Supplementary Examinations, April-2019

DATA WAREHOUSING AND DATA MINING

Time: 3 Hours

Max. Marks: 60

*Answer Any FIVE Questions
All Questions Carry Equal Marks*

1. a Explain in detail Measures of Similarity and Dissimilarity.
b How to select the right tool for Preprocessing or analysis using Data exploration.
2. a Discuss general approach for solving a Classification problem in detail.
b What are the characteristics of K-Nearest neighbor algorithm?
3. Can we design a method that mines the complete set of frequent item sets without Candidate Generation? If yes, explain with example table mentioned above.
4. a Discuss in detail Unsupervised Cluster evaluation using Cohesion and Separation.
b Describe Hierarchical Clustering Algorithm.
5. What is web mining? Explain the following briefly:
 - i. Web usage mining
 - ii. Web content mining
 - iii. Web structure mining
6. a How to evaluate the Classifier Accuracy? Discuss in detail.
b What are the advantages of FP-Growth algorithm?
7. How Support Vector Machines are applied for the classification of both Linear and Nonlinear data?
8. a Clearly explain the working of the DBSCAN algorithm using appropriate example.
b Write a note on Enterprise search.
