

Code: IT6T4

**III B.Tech - II Semester – Regular/Supplementary Examinations  
March 2020**

**DATA MINING AND DATA WAREHOUSING  
(INFORMATION TECHNOLOGY)**

Duration: 3 hours

Max. Marks: 70

**PART – A**

Answer *all* the questions. All questions carry equal marks

11x 2 = 22 M

1.

- a) What is Data mining?
- b) What is classification?
- c) What is information gain?
- d) Define nominal attribute.
- e) What is Data Cube?
- f) What is confidence?
- g) Write the formula for cosine similarity.
- h) Differentiate supervised and unsupervised learning.
- i) What is Rule Based Classification?
- j) What is closed pattern set?
- k) What is outlier?

## PART – B

Answer any *THREE* questions. All questions carry equal marks.

3 x 16 = 48 M

2. a) Write about data mining functionalities. 8 M  
b) Illustrate Data cleaning process with an example. 8 M
3. a) Differentiate data matrix versus dissimilarity matrix. 8 M  
b) Explain binary & numerical attributes with examples. 8 M
4. Illustrate Apriori algorithm with an example. 16 M
5. a) Write about Rule Based classification. 8 M  
b) Discuss metrics for evaluating classifier performance. 8 M
6. a) Write OPTICS clustering algorithm. 8 M  
b) Discuss various hierarchical clustering methods. 8 M