

Code: IT6T4

**III B.Tech-II Semester–Regular/Supplementary Examinations–March 2019**

**DATA MINING & 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) List out the any four functionality of data mining?
- b) What are the different steps in data transformation?
- c) Compare OLTP and OLAP Systems.
- d) What is mean by numeric attribute?
- e) Formulate the principle frequent itemset and closed itemset.
- f) Give the difference between Boolean association rule and quantitative association rule.
- g) What is the difference between Naive Bayes and a Bayes theorem?
- h) What is over sampling and under sampling?
- i) State the role of cluster analysis.
- j) What is the objective function of K-means algorithm?
- k) List the some applications of data mining.

## PART – B

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

3 x 16 = 48 M

2. a) List and discuss the classification of data mining systems. 8 M
- b) Describe in detail about various data transformation techniques. 8 M
3. a) Explain the data warehouse ETL Process. 12 M
- b) Briefly outline how to compute the dissimilarity between objects described by the Nominal attributes and Asymmetric binary attributes. 4 M
4. a) Explain and Apply the Apriori algorithm for discovering frequent item sets of Transactional Data for an All Electronics Branch table.  
Use 2 for the minimum support value. Illustrate each step of the Apriori Algorithm. 12 M

<b>TID</b>	<b>LIST OF ITEMS</b>
T100	I1,I2,I5
T200	I2,I4
T300	I2,I3
T400	I1,I2,I4
T500	I1,I3
T600	I2,I3
T700	I1,I3
T800	I1,I2,I3,I5
T900	I1,I2,I3

- b) Define infrequent pattern and negative pattern. 4 M
5. a) What is Classification? What are the features of Bayesian classification? Explain in detail with an example. 12 M
- b) What approach would you use to apply decision tree induction? 4 M
6. a) How agglomerative hierarchical clustering works? Explain with an example. 8 M
- b) Discuss in detail about the various detection techniques in outlier. 8 M