

B.Tech III Year II Semester (R13) Regular & Supplementary Examinations May/June 2017

**DATA MINING**

(Common to CSE and IT)

Time: 3 hours

Max. Marks: 70

**PART – A**

(Compulsory Question)

\*\*\*\*\*

- 1 Answer the following: (10 X 02 = 20 Marks)
- (a) Explain the terms descriptive and predictive data mining.
  - (b) Describe various challenges in data mining.
  - (c) Explain the term 'Over fitting'.
  - (d) Explain the significance of decision trees in data mining process.
  - (e) Explain in brief about multiclass problem.
  - (f) List and explain in brief the attributes that are used to compare classification and prediction methods.
  - (g) Explain in brief confidence based pruning.
  - (h) Explain in brief the process of mining temporal pattern in data stream.
  - (i) List the advantages and disadvantages hierarchical clustering algorithm.
  - (j) Explain in brief scalable clustering.

**PART – B**

(Answer all five units, 5 X 10 = 50 Marks)

**UNIT – I**

- 2 List and explain various data mining tasks.

**OR**

- 3 Define OLAP. Explain in detail various OLAP operations.

**UNIT – II**

- 4 Describe in detail the process of classification by using Decision tree induction.

**OR**

- 5 Explain in detail hold out method for evaluating classifier.

**UNIT – III**

- 6 Explain how Rule Based method is used for classification in data mining process.

**OR**

- 7 Explain in detail the methods to handle class imbalance problem.

**UNIT – IV**

- 8 Discuss in detail various methods that improve the efficiency of Apriori algorithm.

**OR**

- 9 Illustrate with example in detail, the process of generating association rules from frequent itemsets.

**UNIT – V**

- 10 Explain in detail DBSCAN clustering algorithm.

**OR**

- 11 Define cluster analysis. List and explain applications of cluster analysis.

\*\*\*\*\*