

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) Define data mining. Give significance of data mining.
 - (b) Explain the term OLAP.
 - (c) Explain about decision tree. Explain its role in data mining process.
 - (d) List various types of outcomes corresponding to various test conditions expressing the attributes.
 - (e) Explain the Ensemble Methods in brief.
 - (f) List and explain in brief the attributes that are used to compare classification and prediction methods.
 - (g) Explain Apriori principle in brief.
 - (h) Explain in brief the process of mining temporal pattern in data stream.
 - (i) Explain the term Clustering. List the limitations of K-means clustering algorithm.
 - (j) Explain in brief density based clusters and conceptual clusters.

PART – B

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

UNIT – I

- 2 List and explain various challenges in data mining.
- OR**
- 3 List and explain the processing steps that may be applied to the data for data mining.

UNIT – II

- 4 Explain the process of tree induction. Describe its implementation using Hunt's algorithm.
- OR**
- 5 Explain the term over fitting. Discuss how this over fitting can be addressed.

UNIT – III

- 6 Explain how Bayesian method is used for classification in data mining process.
- OR**
- 7 Explain in detail neural network learning for classification using back propagation algorithm.

UNIT – IV

- 8 Explain the use of Dynamic FP-Tree in data representation.
- OR**
- 9 Explain in detail Attribute oriented induction.

UNIT – V

- 10 Explain in detail hierarchical clustering algorithm.
- OR**
- 11 Define cluster analysis. List and explain the applications of cluster analysis. Also explain various types of clustering.
