Code: 13A05603

R13

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.

*****

Code: 9A05706 **R09**

B.Tech IV Year I Semester (R09) Supplementary Examinations June 2016

DATA WAREHOUSING & DATA MINING

(Computer Science & Engineering)

Time: 3 hours Max. Marks: 70

Answer any FIVE questions

All questions carry equal marks

*****

1 (a) How is data warehouse differs from data base? How are they similar?
(b) Explain the role of sampling in data processing.

2 (a) Give brief description about the different types of OLAP servers.
(b) Write about the general strategies for cube computation.

3 (a) Describe the method to mine the frequent itemsets by using data format technique.
(b) Explain the various steps involved in ARCS.

4 Discuss in detail about the issues regarding preprocessing the data for classification and prediction.

5 Given two objects represented by the tuples (22, 1, 42, 10) and (20, 0, 36, 8):
(a) Compute Minkowski distance between the two object using q = 3.
(b) Compute the Euclidean distance between the two objects.
(c) Compute the Manhattan distance between the two objects.

6 Write in detail about the similarity search in time series analysis.

7 Give brief description about the spatial data mining.

8 Write short notes on the following:
(a) Data mining privacy and data security.
(b) Data mining and collaborative filtering.

*****

Code: 13A05603 **R13**

B.Tech III Year II Semester (R13) Regular Examinations May/June 2016

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.
(b) What is data mart?
(c) What is FP growth?
(d) Compare clustering with classification.
(e) What is Data purging?
(f) What is the difference between OLTP and OLAP?
(g) What is Fact Table?
(h) Define Slice and Dice operation.
(i) Define Association Rule Mining.
(j) What is the use of the knowledge base?

PART – B

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

UNIT – I

2 Explain the steps in knowledge discovery data base.

OR

3 Describe OLAP operations in multidimensional data model.

UNIT – II

4 Explain with example the various steps in Decision tree induction.

OR

5 Explain about evaluating the performance of classifier.

UNIT – III

6 (a) Explain Rule based classifiers.
(b) Explain the Nearest neighbor classification.

OR

7 State Bayes theorem and discuss how Bayesian classifiers work.

UNIT – IV

8 Explain how the efficiency of apriori algorithm is improved.

OR

9 Explain the mining single dimensional Boolean associated rules from transactional data bases.

UNIT – V

10 Explain hierarchical methods of clustering.

OR

11 Discuss different types of clustering methods.

*****