Code: 9A05706

1

B.Tech IV Year II Semester (R09) Regular Examinations, March/April 2013

DATA WAREHOUSING & DATA MINING

(Electronics & Computer Engineering)

Time: 3 hours Max. Marks: 70

Answer any FIVE questions.
All questions carry equal marks.

*****

1 (a) Give the architecture of a typical data mining system.
(b) Explain data transformation and integration.

2 (a) What is a data warehouse? Discuss the role of metadata repository in a data warehouse.
(b) Discuss star-cubing algorithm.

3 Explain Apriori algorithm with an example database and discuss the approaches to improve the efficiency of Apriori.

4 (a) Describe the issues regarding classification and prediction.
(b) Write about support vector machines as classifiers.

5 (a) Discuss K-Medoids approach. Compare it with hierarchical methods.
(b) What is an outlier? What is the need for outlier detection? Discuss applications.

6 What is a social network? Explain its characteristics and social network analysis.

7 (a) Explain web structure mining.
(b) What data mining functionalities are applicable to text databases?

8 Write short notes on:
(a) Data mining for intrusion detection.
(b) Ubiquotous data mining.

*****

Code: 9A05706

2

B.Tech IV Year II Semester (R09) Regular Examinations, March/April 2013

DATA WAREHOUSING & DATA MINING

(Electronics & Computer Engineering)

Time: 3 hours Max. Marks: 70

Answer any FIVE questions.

All questions carry equal marks.

*****

1 (a) Discuss the major issues in data mining.
(b) What is the need for preprocessing of data? List the various forms of preprocessing.

2 (a) Define data warehouse. How is it different from a database system?
(b) Explain OLAP operations in multidimensional data model.

3 What is a frequent itemset? What is a closed frequent itemset? How to find these itemsets? Explain any one approach for each type of itemsets.

4 (a) Discuss the significance of attribute selection measures in decision tree induction.
(b) Explain associative classification techniques.

5 (a) Describe constraint based cluster analysis.
(b) Discuss AGNES and DIANA approaches.
(c) What is the necessity of outlier analysis?

6 What is multirelational data mining? Explain in detail inductive logic programming approach to multirelational classification.

7 (a) Discuss generalization-based mining of plan databases.
(b) What is web usage mining?

8 (a) In which aspects data mining may contribute to biological data analysis. Explain them.
(b) Discuss theoretical foundations of data mining.

*****

Code: 9A05706

3

B.Tech IV Year II Semester (R09) Regular Examinations, March/April 2013

DATA WAREHOUSING & DATA MINING

(Electronics & Computer Engineering)

Time: 3 hours Max. Marks: 70

Answer any FIVE questions.
All questions carry equal marks.

*****

1 (a) Give the classification of data mining systems.
(b) How to integrate a data mining system with a database?
(c) Discuss various data normalization approaches.

2 What is meant by concept description? Explain attribute oriented induction and analytical characterization.

3 (a) Discuss market basket analysis.
(b) Explain frequent pattern mining with vertical data format.

4 (a) Write the basic algorithm for inducing a decision tree from training tuples.
(b) Describe genetic algorithms for classification.

5 (a) Explain a hierarchical clustering algorithm using dynamic modeling.
(b) Discuss STING clustering.

6 (a) What is periodicity analysis? Explain with examples.
(b) Discuss hidden Markov model for biological sequence analysis.

7 (a) Explain dimensionality reduction for text.
(b) Describe the process flow of vision-based page segmentation algorithm.

8 Write short note on:
(a) Commercial data mining systems.
(b) Visual and audio data mining.
(c) Invisible data mining.

*****

Code: 9A05706

4

B.Tech IV Year II Semester (R09) Regular Examinations, March/April 2013

DATA WAREHOUSING & DATA MINING

(Electronics & Computer Engineering)

Time: 3 hours Max. Marks: 70

Answer any FIVE questions.
All questions carry equal marks.

*****

1 (a) Explain the data mining task primitives.
(b) How do handle missing values in data mining?
(c) Briefly discuss the major issues in data mining.

2 What is a data warehouse? In what ways it is different from a database. Discuss various data models used for data warehouse.

3 (a) Discuss Apriori algorithm with illustrations.
(b) Explain mining various kinds of association rules.

4 (a) Describe induction of a decision tree using information gain.
(b) Explain basic sequential covering algorithm with example.

5 (a) Discuss briefly grid based methods for clustering.
(b) Explain deviation based outlier detection.

6 (a) How does the lossy counting algorithm find frequent item?
(b) Explain sequential pattern mining.

7 What is a spatial database? What is spatial data mining? Briefly discuss spatial OLAP.

8 Write short notes on:
(a) Web-wide tracking.
(b) Collaborative filtering.
(c) Data mining applications.

*****