SPRING END SEMESTER EXAMINATION-2013

8th Semester B.Tech

DATA MINING & DATA WAREHOUSING IT-811

[Regular-2009 Admitted Batch]

Full Marks: 60 Time: 3 Hours

Answer any SIX questions including Question No.1 which is compulsory.
The figures in the margin indicate full marks.
Candidates are required to give their answers in their own words as far as practicable and all parts of a question should be answered at one place only.

1. (a) Describe challenges to data mining regarding data mining methodology and user interaction issues. [2 × 5

   (b) Compute the Euclidean distance between two data objects represented by the tuples (22, 1, 42, 10) and (20, 0, 36, 8).

   (c) Define support and confidence in Association rule mining.

   (d) Differentiate clustering and classification.

   (e) Describe Tree pruning methods. [5

   (f) Define anti- monotone property.

   (g) Differentiate Agglomerative and Divisive Hierarchical Clustering. [5

(h) Write the steps of preprocessing .

(i) What are the factors involved while choosing data mining system?

(j) Write and explain characteristics of interesting pattern.

2. (a) Explain the evolution of Database technology. [4

(b) Explain the taxonomy of data mining tasks. [4

3. (a) Explain statistical perspective in data mining. [4

(b) What is incremental association rule. How it is handled? [4

4. (a) Define k-means algorithm with suitable example. [4

(b) Discuss about the components of data warehouse. [4

5. (a) Explain classification by Decision tree induction. Explain major steps of Decision Tree Induction. [4

(b) Explain Neural Network based classifications. [4

6. (a) Explain Outlier analysis. [4

(b) Differentiate OLTP and OLAP. [4

7. A database has four transactions. Let min sup = 60% and min conf = 80%. [8

TID	Items bought
T100	{ANN, CC, TC, CG}
T200	{CC, D, GG}
T300	{ANN, CC, TC, CG}
T400	{ANN, CC, D, CG}
T500	{ANN, CC, D, TC, CG}
T600	{CC, D, TC}

Find frequent item set using Apriori and FP growth algorithm.

8. Answer any two of the following [4 × 2

(a) Define text mining

(b) What does web mining mean?

(c) Define spatial data mining.

(d) Explain Genetic Algorithm.

xxxxx