IT-840 B.E. VIII Semester Examination, June 2015

Data Mining and Warehousing Elective-IV

Time: Three Hours http://www.rgpvonline.com Maximum Marks: 70

Note: Attempt all questions. All questions carry equal marks.

1. a) How is a data warehouse different from a database? How are they similar to each other ? 7

b) Describe three data warehouse models - the enterprise warehouse, the data mart and the virtual warehouse. 7

OR

2. a) Discuss system development life cycle of a data warehouse. What factors should be considered while designing a data warehouse? 7

b) Describe star schema and snowflake schema with examples. 7

3. a) Why most data warehouse system support index structures? Discuss methods to index OLAP data. 7

b) Discuss typical OLAP operations in brief. 7

OR

4. a) Define OLAP. What are the four different types of OLAP server from implementation point of view ? Explain briefly. 7

b) How the different types of models apply to the architected environment? 7

5. a) Discuss various issues in data mining ? 7

b) What do you mean by data reduction? What are the strategies of the data reduction. 7

OR

6. Explain briefly: 14

i) Text mining ii) Web usage mining iii) Spatial mining iv) Web structure mining

7. a) What is the 'Apriori property'? How is it used by the APRTORI algorithm? What are the drawbacks of the apriori algorithm? - 7

b) What do you mean by association rule mining? Give an example of market basket analysis from the real world. 7

OR

8. a) Explain Fp-growth algorithm with an example. 7 b) Discuss the latest trends in association rule mining? 7

9. a) Suppose we have the following points: (1,1), (2,4), (3,4), (5,8), (6,2), (7,8). Use k - means algorithm (k = 2) to find two cluster. The distance function is Euclidean distance. Find 2 cluster using k - mers clustering algorithm. Use (1, 1) and (2, 4) to form the initial clusters. 7 b) What are the requirements of clustering in data mining? 7

OR

10. a) Why is decision tree induction popular? Discuss over fitting of on induced tree and two approaches to avoid over - fitting using suitable examples/ diagrams. 7

b) What is Hierarchical clustering? Differentiate agglomerative and Divisive Hierarchical clustering. 7