Anna University Exams Nov / Dec 2017 – Regulation 2013
Rejinpaul.com Unique Important Questions – 7th Semester BE/BTECH
IT6702 Data Warehousing and Data Mining
Unit I-V

1. Explain the mapping of data warehouse to multiprocessor architecture
2. Discuss about the data warehouse Metadata?
3. With a neat diagram describe the various stages of building a data warehouse.
4. What is a data warehouse? Diagrammatically illustrate and discuss the data warehousing architecture?
5. What is a data warehouse? With the help of a neat sketch explain the various components in a data warehousing system?
6. Explain the different types of OLAP tools
7. Write the difference between multi-dimensional OLAP and multi relational OLAP?
8. Diagrammatically illustrate and discuss the architecture of MOLAP and ROLAP?
9. (a) List and discuss the basic features that are provided by reporting and query tools used for business analysis. (b) Explain the diagrammatic illustration managed query environment (MQE) architecture.
10. With diagrammatic illustration discuss data mining as a confluence of multiple disciplines.(State and explain the various classification of data mining systems with example.
11. List and discuss the data mining task primitives.
12. Discuss the following schemas used for integration of a data mining system with a database or data warehouse system: (List and discuss the steps for integrating a data mining system with a data warehouse) a.No coupling b.Loose coupling c. Semi tight coupling d.Tight coupling
13. Explain the various data mining issues and functionalities in detail
14. What is the use of data mining task? what are the basic types of data mining tasks
15.Giving a concrete example explain a method that performs frequent itemsets mining by using the prior knowledge of frequent itemset properties.
16. Explain how the Bayesian belief networks are trained to perform classification.
17. What is classification? With an example explain how support vector machines can be used for classification
18. Discuss in detail the constraint based association mining.
19. What is decision tree. Explain how classification is done using decision tree induction.
20. What is classification? With an example explain how support vector machines can be used for classification.
21. Discuss the Apriori Algorithm for discovering frequent item sets. Apply the Apriori algorithm to the follwing data set.

TransID	Items Purchased
101	Strawberry, litching, oranges
102	Strawberry, butter furit
103	butter fruit, vanilla
104	strawberry, litchi, oranges
105	banana, oranges
106	banana
107	banana, butter fruit
108	strawbeery, litchi, apple, oranges
109	apple, vanilla
110	strawberry, litchi

The set of items is {strawberry, litchi, apple, oranges, vanilla, banana, butter fruit} Use 0.3 for the minimum support value
22. Explain hierarchical method and density based method of classification with example.
23. Explain the types of data in cluster analysis in detail with example
24. Explain outlier analysis with example
25. Why is outlier mining important? Briefly describe the different approaches behind statistical based outlier detection,distance based outlier detection, and deviation based outlier detection.

26. What is hierarchical clustering? With an example discuss dendrogram representation for hierarchical clustering of data objects.

27. Consider five points { X1, X2, X3, X4, X5} with the following coordinates as a two dimensional sample for clustering: X1-(0,2); X2-(0,0); X3-(1.5,0); X4-(5,0); X5-(5,2) Illustrate the K-means partitioning algorithm (clustering algorithm) using the above data set.

Questions Are Expected for University Exams This May or may Not Be Asked for Exams