Data Warehouse And Data Mining Question Bank

UNIT –I

1. What is data warehouse?
2. What is the benefits of data warehouse?
3. What is the difference between OLTP and OLAP?
4. Briefly state different between data ware house & data mart?
5. What are the characteristics of data warehouse?
6. Define data warehouse?
7. What is the need of data warehouses?
8. Define OLAP?
9. Define multidimensional data model?
10. What is a data cube?
11. Define dimensions?
12. What are facts?
13. Define OLTP?
14. Define OLAP?
15. Define dimension table?
16. Define fact table?
17. What are lattice of cuboids?
18. What is apex of cuboid?
19. List out the various OLAP operations?
20. Give the names of warehouse schemas?
21. Define star schema?
22. Define snowflake schema?
23. Draw a neat diagram of data warehouse architecture?
24. Define data mart?
25. Define metadata?
26. What are the applications of metadata?
27. List out the types of metadata?
28. What are the phases present in development cycle of a data warehouse?
29. Give the differences between a database and a data warehouse?
30. Explain multidimensional data model with a neat diagram?
31. List out the OLAP operations and explain the same with an example?
32. Describe about dimension modeling in detail?
33. Explain the various schemas of a data warehouse?
34. Define data warehouse. Draw the architecture of data warehouse and explain the three tiers in detail?
35. Explain in detail about the implementation of a data warehousing?
36. Define metadata and explain the types of metadata?
37. Discuss the development lifecycle of a data warehouse?
38. What is multidimensional data model? Discuss the schemas for multidimensional data?

UNIT-2

1. Define Data mining?
2. What are the other terminologies referring to data mining?
3. List out the applications of data mining?

4. Differentiate data mining tools and query tools?
5. What is meant by machine learning?
6. List out the data mining processing steps?
7. What are the techniques used in data mining?
8. Explain the knowledge discovery phases?
9. Name some of the data mining applications?
10. What is data mining? Explain the steps in data mining process?
11. Explain major requirements and challenges in data mining?
12. What are different types of Data Mining Techniques? Explain any one in detail?.
13. Explain the data mining functionalities?
14. Explain the contrast between data mining tools and query tools?
15. Give in detail about the data mining techniques?
16. What are the processes being carried out in backend of data warehouse?
17. Explain the processes taking place in the backend of a data warehouse?
18. Explain in detail about the implementation of a data warehousing?

UNIT-3

1. Write notes on k-means algorithm?
2. Define association rule mining?
3. Define a concept hierarchy?
4. What is Decision tree?
5. Define the centroid of the cluster?
6. Discuss the classification by decision tree induction?
7. Explain density based clustering methods in detail?
8. Consider a database, D, consisting of 9 transactions. Suppose min.support count required is 2 and let min.confidence required is 70%. Use the apriori algorithm to generate all the frequent candidate itemsets Ci and frequent itemsets Li.

TID	**List of Items**
T100	I1, I2, I5
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3
T600	I2, I3
T700	I1, I3
T800	I1, I2 ,I3, I5
T900	I1, I2 , I5

9. Explain about the partitioning methods?
10. Discuss about model based clustering methods?
11. Explain in details each one of these steps.
    i. Decision Support System

ii. Association rules
iii. The Apriori algorithm Key Concepts

12. How the K-Mean Clustering algorithm works?

13. Suppose that we have the following table of a database of transactions D, depending on these transactions determine Support and Confidence values for the following items I.

$X \Rightarrow Y$
Bread ⇒ PeanutButter
PeanutButter ⇒ Bread
Beer ⇒ Bread
PeanutButter ⇒ Jelly
Jelly ⇒ PeanutButter
Jelly ⇒ Milk

items I

Transaction	Items
t_1	Bread,Jelly,PeanutButter
t_2	Bread,PeanutButter
t_3	Bread,Milk,PeanutButter
t_4	Beer,Bread
t_5	Beer,Milk

a database of transactions D

14. Consider a database, D, consisting of 4 transactions. Suppose min.support count required is 2 and let min.confidence required is 70%. Use the apriori algorithm to generate all the frequent candidate itemsets Ci and frequent itemsets Li.

TID	ITEM
100	1 ,3, 4
200	2, 3, 5
300	1 ,2, 3, 5
400	2 ,5

15. What is the purpose of Apriori Algorithm?
16 How to generate association rules from frequent item sets?
17. Define clustering? Why clustering is important in Data Mining? Write its uses?
18. What is association rule mining? Explain Apriori algorithm in data mining?

UNIT-4

1. Define web mining?
2. What is a multimedia database?
3. Define web content mining?
4. What is spatial mining?
5. What is time series analysis?
6. Define sequence mining?
7. What are the applications of data mining?
8. Explain the process of mining the World Wide Web?
9. Explain the various types of web mining?
10. Explain spatial mining and time series mining?