NOTE:
• Question 1 is compulsory and carries 28 marks. Answer any FOUR questions from the rest. Marks are indicated against each question.
• Parts of a question should be answered at the same place.

Q.1
a. Define data warehouse in detail.

b. Write and explain a 3-D data cube representation of the data in any table, according to the dimensions time, item and location.

c. Define Bayes’ theorem.

d. What are the compositions of data transformation during data pre-processing?

e. What are the different partitioning methods?

f. Show a three-tier data warehousing architecture, with the help of a suitable diagram.

g. What are the different strategies for data reduction during data pre-processing? (7×4)

Q.2
a. Describe OLAP operations in the multidimensional data models. (9)

b. Explain data warehouse implementation in detail. (9)

Q.3
a. Compare classification and prediction methods. (9)

b. Write an algorithm for inducing a decision tree from training tuples. (9)

Q.4
a. What are the key features of a data warehouse? Discuss. (9)

b. What are the major issues in data mining? (9)

Q.5
a. Write the Apriori algorithm for discovering frequent itemsets for mining Boolean association rules. (9)

b. Explain rule induction using a sequential covering algorithm in rule based classification. (9)

Q.6
a. Explain a hierarchical clustering algorithm for categorical attributes. (9)

b. Write a note on OLAP servers. (9)
Q.7

a. What are the requirements of clustering in data mining? (9)

b. Explain various data mining applications in detail. (9)