

(2 ½ Hours)

[Total Marks: 75]

- N.B. 1) All questions are compulsory.
 2) Figures to the right indicate marks.
 3) Illustrations, in-depth answers and diagrams will be appreciated.
 4) Mixing of sub-questions is not allowed.

Q. 1 Attempt ANY FOUR from the following: (20M)

- (a) Write a note on granularity of facts.
 (b) How would you describe Dimensional Model. Also explain steps to design Dimensional Model.
 (c) Compare Virtual and Hybrid Data Warehouses
 (d) Elaborate parts of Data Warehouse .
 (e) Design a Star Schema for Data Warehouse of Library Management System.
 (f) When do you prefer Virtual Data Warehouse model? Justify your answer.

Q. 2 Attempt ANY FOUR from the following: (20M)

- (a) What are various social implications of data mining?
 (b) What kind of data can be mined form Data Mining?
 (c) Discuss pre-processing of data in the data mining .
 (d) Explain FP-growth algorithm.
 (e) Explain the terms Support ,Confidence, and Lift in Association Rule metrics.
 (f) For the given following data, apply Apriori algorithm. (Min. Support= 50%, Confidence threshold = 60%)

TID	Items
T1	I1,I2,I3
T2	I2, I3,I4
T3	I4,I5
T4	I1,I2,I4
T5	I1,I2,I3,I5
T6	I1,I2,I3,I4

Dataset D

Q. 3 Attempt ANY FOUR from the following: (20M)

- (a) Discuss various classification models in detail.
 (b) Explain logistic regression with example.
 (c) Discuss why identifying outliers is important in datasets?
 (d) What is clustering analysis in data mining process?
 (e) How can you measure performance of classifier? Explain in brief.
 (f) What are some real-world applications where clustering algorithms play a significant role in different industries and domains?

Q. 4 Attempt ANY FIVE from the following: (15M)

- (a) Differentiate between OLTP and OLAP
 (b) Write a note in Data Mart.
 (c) How prediction and classification differ from each other.
 (d) Discuss Functional dependency of data.
 (e) Define attribute and enlist its types.
 (f) State the applications of Classification.