Max. Marks:100

Note: Answer any FIVE full questions, selecting atleast TWO questions from each part.

## PART - A

- What is Operational Data Store (ODS)? Explain the ODS structure with a neat diagram. 1 (07 Marks)
  - What are the differences between ODS and Data warehouse? (07 Marks)
  - (06 Marks) Explain the implementation steps in Data warehouse.
- Explain Codds OLAP characteristics. (07 Marks)
  - (06 Marks) Explain ROLAP and MOLAP.
    - c. Explain Data cube operations. (07 Marks)
- What are the challenges that motivated the development of data mining? (07 Marks)
  - (06 Marks) b. Explain the core data mining tasks.
  - c. What are the different strategies that are used in Data preprocessing? Explain any four of (07 Marks)
- (06 Marks) Explain Frequent itemset generation of the Apriori Algorithm. a.
  - What are the methods that are used for generating frequent itemsets? Explain them in detail. (07 Marks)

Construct a FP – tree for the following Data set.

Constitute and to the total and the total an										
TID	1	2	3	4	5	6	7	₹ 8	9	10
Items	{a,b}	{b,c,d}	{a,c,d,e}	{a,d,e}	{a,b,c}	{a,b,c,d}	{a}	√{a,b,c}	{a,b,d}	{b,c,e}

(07 Marks)

## PART - B

- What are the uses of classification model? Explain the general approach to solving a 5 (06 Marks) classification problem.
  - b. Explain Algorithm for Decision Tree Induction. (07 Marks)
  - Explain K nearest neighbor classification Algorithm. Summarize the characteristics of nearest - neighbor classifier. (07 Marks)
- a. Explain Bayes Theorem with a relevant example to solve a prediction problem. (06 Marks)
  - What are the ways that are used to construct the ensemble of classifiers? Explain the general (07 Marks) procedure for ensemble method.
  - c. Explain Multiclass problem.

- (07 Marks)
- a. What is Cluster Analysis? Explain different types of clusters. (06 Marks)
  - b. Explain Basic Agglomerative Hierarchical Clustering Algorithm. Give its strength & (07 Marks) weakness.
  - c. Explain DBSCAN Algorithm. What are the strength and weakness of DBSCAN algorithm? (07 Marks)
- Write short notes on:
  - Web content mining.
  - Text mining.
  - Text clustering.
  - d. Mining Temporal Databases.

CMRIT LIBRARY BANGALORE - 560 037

(20 Marks)

\*\*\*\*