USN ☐☐☐☐☐☐☐☐☐☐

10IS74

Seventh Semester B.E. Degree Examination, June/July 2019
Data Warehousing and Data Mining

Time: 3 hrs.   Max. Marks:100

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

PART - A

1   a. What is Operational Data Store (ODS)? Explain the ODS structure with a neat diagram.
                                                                                  (07 Marks)
    b. What are the differences between ODS and Data warehouse?   **(07 Marks)**
    c. Explain the implementation steps in Data warehouse.   **(06 Marks)**

2   a. Explain Codds OLAP characteristics.   **(07 Marks)**
    b. Explain ROLAP and MOLAP.   **(06 Marks)**
    c. Explain Data cube operations.   **(07 Marks)**

3   a. What are the challenges that motivated the development of data mining?   **(07 Marks)**
    b. Explain the core data mining tasks.   **(06 Marks)**
    c. What are the different strategies that are used in Data preprocessing? Explain any four of
       them.   **(07 Marks)**

4   a. Explain Frequent itemset generation of the Apriori Algorithm.   **(06 Marks)**
    b. What are the methods that are used for generating frequent itemsets? Explain them in detail.
                                                                                  (07 Marks)
    c. Construct a FP – tree for the following Data set.

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

5   a. What are the uses of classification model? Explain the general approach to solving a
       classification problem.   **(06 Marks)**
    b. Explain Algorithm for Decision Tree Induction.   **(07 Marks)**
    c. Explain K – nearest neighbor classification Algorithm. Summarize the characteristics of
       nearest – neighbor classifier.   **(07 Marks)**

6   a. Explain Bayes Theorem with a relevant example to solve a prediction problem.   **(06 Marks)**
    b. What are the ways that are used to construct the ensemble of classifiers? Explain the general
       procedure for ensemble method.   **(07 Marks)**
    c. Explain Multiclass problem.   **(07 Marks)**

7   a. What is Cluster Analysis? Explain different types of clusters.   **(06 Marks)**
    b. Explain Basic Agglomerative Hierarchical Clustering Algorithm. Give its strength &
       weakness.   **(07 Marks)**
    c. Explain DBSCAN Algorithm. What are the strength and weakness of DBSCAN algorithm?
                                                                                  (07 Marks)

8   Write short notes on :
    a. Web content mining.
    b. Text mining.
    c. Text clustering.
    d. Mining Temporal Databases.   **(20 Marks)**

*****