USN | | | | | | | | | | |

10IS74

# Seventh Semester B.E. Degree Examination, June/July 2017
## Data Warehousing and 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 Data warehouse? Explain in detail the different key features of warehouse. **(10 Marks)**

b. Explain in detail the difference between ODS and warehouse. **(05 Marks)**

c. What is Data Mart? **(05 Marks)**

2 a. What are the different types of OLAP operations? Explain them with suitable example. **(10 Marks)**

b. What is Data cube? With figure, explain different structure of data cubes. **(10 Marks)**

3 a. Explain in detail different types of data preprocessing techniques. **(10 Marks)**

b. With suitable example, explain Minkowski distance metric. **(10 Marks)**

4 a. Consider data transaction ID : **(10 Marks)**

| 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} |

Apply FP growth algorithm to find frequent itemset ending in 'e'.

b. Write a procedure in Apriori – gen function, which merges a pair of frequent item set. Explain with example. **(10 Marks)**

### PART – B

5 a. Construct decision tree for a mamal classification problem. Discuss design issues of decision tree. **(10 Marks)**

b. Write an algorithm for skeleton decision tree and describe different functions used in the algorithm. **(10 Marks)**

6 a. Estimate conditional probabilities of continuous attribute by Naïve Baye's classifier. **(10 Marks)**

b. Explain in detail Bagging and Boosting accuracy of classifier. **(10 Marks)**

7 a. Briefly outline how to compute dissimilarity between object described by following types of variables in cluster :
i) Interval scaled variable ii) Binary variable. **(10 Marks)**

b. What is Clustering? Describe the following approaches to clustering method :
i) Partitioning method ii) Hierarchical methods.
Give example in each case. **(10 Marks)**

8 Write short notes on :
a. Multiclass problem.
b. Mining raster database.
c. Automatic classification of web document.
d. Construction of multilayered web information base. **(20 Marks)**

*****