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

10IS74

Seventh Semester B.E. Degree Examination, June/July 2018

Data Warehousing and Data Mining

Time: 3 hrs. Max. Marks:100

Note: ***Answer FIVE full questions, selecting at least TWO questions from each part.***

PART – A

1 a. Define data warehouse. With neat diagram, explain implementation of operational data store. **(10 Marks)**

b. With suitable example, explain star schema for multidimensional database. **(10 Marks)**

2 a. i) Explain the characteristics of OLAP systems. **(05 Marks)**

ii) Give the differences between OLTP and OLAP systems. **(05 Marks)**

b. Explain cube operations with suitable examples. **(10 Marks)**

3 a. Explain KDD process. Highlight the importance of four core data mining tasks. **(10 Marks)**

b. Write a note on different types of data sets. **(10 Marks)**

4 a. With suitable example illustrate the Apriori principle. **(10 Marks)**

b. Explain frequent internet generation in FP-growth algorithm. **(10 Marks)**

PART – B

5 a. Write a decision tree for the mammal classification problem. Explain how to generate decision tree using Hunt's algorithm. **(10 Marks)**

b. What is rule-based classifier? With example explain how a rule based classifier works. **(10 Marks)**

6 a. What are Bayesian classifiers? State Baye's theorem and explain how Baye's theorem is used in the Naïve Bayesian classifier with example. **(10 Marks)**

b. What is cluster analysis? List out the difference between cluster analysis and classification. Mention desired features of cluster analysis. **(10 Marks)**

7 a. Explain different methods of computing distance. **(10 Marks)**

b. Give the taxonomy of cluster analysis methods. **(10 Marks)**

8 a. With an example, explain K-means clustering method. **(10 Marks)**

b. i) What is web data mining? **(02 Marks)**

ii) List the major differences between conventional searching and web searching. **(08 Marks)**

* * * * *