R13

Code No: 814BD

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
MCA IV Semester Examinations, December - 2019
DATA WAREHOUSING AND DATA MINING

Time: 3 Hours **Max. Marks: 60**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 20 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 8 marks and may have a, b, c as sub questions.

PART - A

5 × 4 Marks = 20

1.a) What is data preprocessing? What are the various forms of it. [4]
b) Explain the architecture of data warehouse. [4]
c) Distinguish between Apriori algorithm and Frequent pattern growth algorithm. [4]
d) What are the steps involved in preparing the data for classification? [4]
e) What are the advantages and disadvantages of k-means algorithm. [4]

PART - B

5 × 8 Marks = 40

2.a) Explain about the Integration of a Data mining system with a Database or Data warehouse system.
b) List the techniques for Data cleaning. [4+4]

OR

3.a) What are the issues in Data mining?
b) Explain about the classification of data mining systems. [4+4]

4. Explain in detail about the implementation of a data warehouse. [8]

OR

5. Illustrate OLAP operations in the Multidimensional Data Model. [8]

6. Discuss the FP-growth algorithm. Explain with an example. [8]

OR

7. Discuss about mining multilevel association rules from transaction databases in detail. [8]

8. Explain in detail about classification of Back propagation algorithm. [8]

OR

9. Explain linear regression and non-linear regression with example. [8]

10. Write algorithms for k-medoid clustering and Explain with simple data. [8]

OR

11. Explain about the Statistical-based outlier detection. [8]

---ooOoo---