Tribhuvan University
Institute of Science and Technology
2076
✡

Bachelor Level / Forth Year /Eighth Semester/Science
Computer Science and Information Technology-(CSc. 451)
(Data warehousing and Data mining)

Full Marks: 60
Pass Marks: 24
Time: 3 hours.

Candidates are required to give their answers in their own words as for as practicable.
The figures in the margin indicate full marks.

Group A

Attempt any two questions.

[2×10=20]

1. Discuss the types of web mining. Explain why K-Means is sensitive to outlier and how does K-Medoid minimize this issue. [4 + 6]

2. Do pattern and information refer to same aspect? Justify. Differentiate between data warehouse and operational database. [3 + 7]

3. List the problems of Apriori algorithm with its possible solutions. Consider the following transaction dataset. [3 + 7]

Transaction_ID	Item_List
T1	{K, A, D, B}
T2	{D, A, C, E, B}
T3	{C, A, B, E}
T4	{B, A, D}

   What association rules can be found in this set, if the minimum support is 3 and the minimum confidence is 80%.

Group B

Attempt any eight questions.
Question No. 13 is compulsory.

[8×5=40]

4. How classification plays significance role in data mining? Explain. [5]

5. Are the information given by data mining is always useful? What are the issues in data warehousing and data mining? [2 + 3]

6. Explain the four characteristics of data warehouse. [5]

7. Explain the optimization techniques in data cube computation. [5]

8. How multidimensional data model helps in retrieving information? Explain with suitable example. [5]

9. Compare the OLAP servers, ROLAP, MOLAP and HOLAP. [5]

10. Give a syntax and example of data mining query language. [5]

11. Differentiate between KDD and data mining. [5]

12. What does data warehouse tuning mean? Describe the parameters. [5]

13. Write short notes on (**Any Two**) [2.5 + 2.5]
    a. Evolution analysis
    b. Decision trees
    c. Text mining
    d. Classification using Regression