



Term Evaluation (Odd) Semester Examination September 2025

Roll no. 2220029

Name of the Course: B.Tech C.S.E
Semester: VII
Name of the Paper: Data Warehousing and Data Mining
Paper Code: TCS 722
Time: 1.5 hour

Maximum Marks: 50

Note:

- (i) Answer all the questions by choosing any one of the sub-questions
- (ii) Each question carries 10 marks.

Q1. (10 Marks)
a. "Data mining is not just extracting data, but extracting useful knowledge." Discuss this statement in detail by comparing data mining with information retrieval and traditional query processing. (CO1)

OR

b. Explain in detail the functional areas of data mining such as classification, clustering, prediction, and outlier analysis. Illustrate each with suitable examples. (CO1)

Q2. (10 Marks)
a. Explain the various forms of data preprocessing in detail. Why is preprocessing considered the most critical step in data mining? Support your answer with a real-world case study. (CO1)

OR

b. Explain the step-by-step process of detecting and handling outliers in a dataset. Provide an example using binning and clustering techniques. (CO1)

Q3. (10 Marks)
a. Explain the process of mining class comparisons. Differentiate between discriminant description and characteristic description using examples. (CO2)

OR

b. A supermarket stores sales data in a data cube with dimensions: Time (Month), Location (City), and Product (Category). Illustrate how data cube aggregation can be performed to summarize sales at the state level instead of city level. (CO1)

Q4. (10 Marks)
a. Describe concept description. Explain data generalization and analytical characterization in the context of data mining. Give an example of how these can be applied in a retail business database. (CO2)

OR

b. Differentiate between measures of central tendency and measures of dispersion. For the dataset {40, 45, 50, 55, 55, 60, 65, 70}, compute mean, mode, median, variance, and standard deviation. (CO2)

Q5. (10 Marks)
a. Write short notes on the following with examples:
i. Multi-level association rules
ii. Multi-dimensional association rules (CO2)

OR

b. Write the step-by-step working procedure of Apriori algorithm. Highlight its strengths, weaknesses, and suitable applications. (CO2)