

R18

Code No:157BC

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B.Tech IV Year I Semester Examinations, December-2023/January-2024

DATA MINING

(Common to CSE, IT, ECM, ITE)

Time: 3 Hours

Max. Marks: 75

Note: i) Question paper consists of Part A, Part B.

ii) Part A is compulsory, which carries 25 marks. In Part A, Answer all questions.

iii) In Part B, Answer any one question from each unit. Each question carries 10 marks and may have a, b as sub questions.

PART – A

(25 Marks)

- 1.a) What is a data warehouse? [2]
- b) How does constraint-based method influence the discovery of association rules? [3]
- c) What is the advantage of rule-based classification systems? [2]
- d) List the major clustering methods. [3]
- e) Define time-series data mining. [2]
- f) What is the purpose of data preprocessing in data mining? [3]
- g) What is association rule mining? [2]
- h) How does a Bayesian classifier handle uncertainty? [3]
- i) Mention the applications where clustering is commonly used. [2]
- j) Briefly describe multimedia data mining. [3]

PART – B

(50 Marks)

- 2.a) How do task primitives contribute to the overall effectiveness of data mining? Explain. [5+5]
 - b) Explain the significance of data preprocessing in data mining. [5+5]
- OR**
- 3.a) Compare data mining with data warehouse. [5+5]
 - b) Provide a step-by-step explanation of how data mining can be applied to classify data. [5+5]
4. Explore the different mining methods used in association rule mining and discuss their advantages and limitations. [10]
- OR**
- 5.a) Explain the process of mining frequent patterns. [5+5]
 - b) Describe the role of correlation analysis in identifying relationships within datasets. [5+5]
6. Provide a step-by-step explanation of how a lazy learning algorithm, such as k-Nearest Neighbors (k-NN), makes predictions. [10]
- OR**
7. Compare and contrast rule-based classification systems with decision tree induction. [10]

QA QA QA QA QA QA QA G

- 8.a) Discuss the relationship between density-based and grid-based clustering methods.
- b) Describe the steps involved in a grid-based clustering approach. [5+5]

OR

QA QA QA QA QA QA QA G

- 9. How do partitioning methods work in cluster analysis, and what factors should be considered when choosing an appropriate method? Explain. [10]

- 10. Explain the following terms.
 - a) Mining the World Wide Web.
 - b) Spatial data mining.

[5+5]

OR

QA QA QA QA QA QA QA G

- 11. Elaborate on the process of mining sequence patterns in transactional databases, outlining the steps and potential challenges involved. [10]

---ooOoo---

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G

QA QA QA QA QA QA QA G