

Code No: 872AD

R20

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

MCA II Semester Examinations, March - 2023

DATA MINING

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) Explain various data smoothing techniques. [5]
- b) What are maximal and closed frequent itemsets? [5]
- c) What is decision tree? Explain with example. [5]
- d) Elaborate on outlier detection. [5]
- e) Explain about web mining in brief. [5]

PART – B

(50 Marks)

2. Explain various methods for dimensionality reduction. [10]
- OR**
3. Elaborate on discretization and binarization. [10]
4. Explain apriori algorithm with an example. [10]
- OR**
5. What are the compact representations of frequent item set? Explain with example. [10]
6. What are various measures for selecting the best split? Explain them. [10]
- OR**
7. Explain Bayesian belief networks. [10]
8. Explain PAM algorithm in detail. [10]
- OR**
- 9.a) What are disadvantages of k-means algorithm? How can they be overcome? [5+5]
- b) How to detect outliers? Explain with an example.
10. Explain web usage mining. [10]
- OR**
11. Explain text mining in detail. [10]

---ooOoo---