

Code No: 155EP

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech III Year I Semester Examinations, January/February - 2023

DATA MINING

(Computer Science and Engineering – Data Science)

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 are the applications of Data Mining? [2]
- b) What do you mean by Data transformation? Give an example. [3]
- c) List out various kinds of Association rules with examples. [2]
- d) How to mine closed frequent item sets? [3]
- e) What is Baye's theorem? [2]
- f) What are the benefits of tree pruning? [3]
- g) List out various clustering methods. [2]
- h) Write about significance of Outlier Analysis. [3]
- i) What is Multimedia Data Mining? [2]
- j) List out the applications of web mining. [3]

**PART – B****(50 Marks)**

- 2.a) Explain about the challenging issues in Data mining.
- b) Discuss about the classification of Data Mining System. [5+5]

**OR**

- 3.a) Explain about the Data Cleaning techniques in detail.
- b) Summarize the data mining task primitives and why primitives play important role in Data mining process. [5+5]

4. How to find all frequent item sets using Apriori algorithm for the given data where  $\text{min-sup} = 2$ . [10]

Transactional Data for an *AllElectronics* Branch

TID	List of item IDs
T100	11, 12, 15
T200	12, 14
T300	12, 13
T400	11, 12, 14
T500	11, 13
T600	12, 13
T700	11, 13
T800	11, 12, 13, 15
T900	11, 12, 13

QA QA QA QA QA QA QA G

**OR**

5.a) How to find frequent item sets using Vertical Data Format? Explain.

b) Discuss about support and confidence.

[6+4]

6. Explain about Decision Tree Induction method for classification with suitable example.

[10]

**OR**

7. How will you solve a classification problem using Bayesian Belief Networks? Explain with an example.

[10]

8.a) Demonstrate Agglomerative clustering algorithm with example.

b) How to cluster the data sets using PAM clustering algorithm?

[5+5]

**OR**

9. Discuss about DBSCAN clustering Algorithm with suitable example.

[10]

10.a) Discuss the web content mining in detail.

b) Explain about unstructured text mining.

[5+5]

11. Explain about the following:

a) Spatial Data Mining

b) Mining Data Streams.

[5+5]

QA QA QA QA QA QA QA G

---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