

Code No: 156EZ

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD**B. Tech III Year II Semester Examinations, August/September - 2024****DATA MINING****(CSE (AI&ML), AI&ML, AI&DS)****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) List the functionalities of data mining. [2]
- b) Mention the task primitives used in data mining. [3]
- c) How do you measure the efficiency of APRIORI algorithm? [2]
- d) Name the steps in association rule mining. [3]
- e) What is the role of IF-THEN rule in classification? [2]
- f) List the advantages of using decision trees. [3]
- g) Define clustering. [2]
- h) State the principle of hierarchical methods. [3]
- i) Distinguish Web mining and text mining. [2]
- j) What are the features of unstructured text mining? [3]

PART – B**(50 Marks)**

- 2.a) Describe the steps involved in data mining when viewed as knowledge discovery process.
- b) Explain how the evolutions of database technology lead to data mining. [5+5]

OR

- 3.a) Explain the need of data preprocessing and various forms of preprocessing.
- b) Discuss the characteristics of Data warehousing and steps in building a data warehousing architecture. [5+5]

- 4.a) How associations and correlations are used for mining the data?
- b) Explain item merging concept for mining closed frequent item sets. [5+5]

OR

- 5.a) Brief about Graph pattern mining.
- b) Describe various kinds of association rules. [5+5]

- 6.a) Discuss decision tree over fitting and pruning techniques.
- b) State classification problem and discuss a general approach to solve classification problem. [5+5]

OR

- 7.a) With an example, describe Naïve Bayesian Classification method.
- b) How to solve a classification problem using k-nearest neighbor algorithm? [5+5]

- 8.a) How clustering is useful? What are its applications?
b) Discuss about the outliers. Explain the weakness and strengths in hierarchical clustering methods. [5+5]

OR

- 9.a) What is the purpose of Density based methods? Discuss.
b) Using an example, give the advantage of Grid based clustering method. [5+5]
- 10.a) Give the basic concepts in mining data streams.
b) What activities are involved in mining spatial data? [5+5]

OR

- 11.a) Explain about hierarchy of categories in text mining.
b) Write short note on Mining the World Wide Web. [5+5]

---ooOoo---

paper Aug-2024