

R18

Code No: 155ET

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

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

DATA WAREHOUSING AND DATA MINING

(Common to CSBS, CSIT, CSE(CS))

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 the differences between star and snowflake schema. [2]
- b) What is data warehouse metadata? [3]
- c) What are the application areas of data mining? [2]
- d) What is descriptive and predictive data mining? [3]
- e) Explain whether association rule mining is supervised or unsupervised type of learning. [2]
- f) Define Confidence measure. [3]
- g) What is Decision tree? [2]
- h) What are the advantage of Bayesian classification? [3]
- i) What is outlier detection? [2]
- j) What is the goal of clustering? [3]

PART – B

(50 Marks)

- 2.a) List out the OLAP operations in multidimensional data model and explain with example.
- b) Explain multidimensional data model with a neat diagram. [5+5]

OR

- 3.a) Compare OLTP and OLAP.
- b) Explain the types of OLAP server in detail. [5+5]
- 4.a) What are the key issues in data Mining? Explain.
- b) Explain different data mining tasks. [5+5]

OR

- 5.a) Describe different data cleaning approaches.
- b) How can we handle missing values? Explain in detail. [5+5]
6. Write and explain the algorithm for mining frequent item sets without candidate generation. Give relevant example. [10]

OR

7. Apply the Apriori algorithm for discovering frequent item sets of the following. Use 3 for minimum support count. [10]

TID	Items purchased
101	milk,bread,eggs
102	milk,juice
103	juice,butter
104	milk,bread,eggs
105	coffee,eggs
106	coffee
107	coffee,juice
108	milk,bread,cookies,eggs
109	cookies,butter
110	milk,bread

8. Explain classification by Decision tree induction algorithm with example. [10]

OR

- 9.a) Explain briefly the test conditions for different types of attributes.
b) Explain k- nearest neighbor classifier with example. [5+5]

- 10.a) Define Clustering. Explain about Types of Data in Cluster Analysis.
b) Classify various Clustering methods. [5+5]

OR

11. Describe in detail about hierarchical clustering methods. [10]

---ooOoo---