R18

Code No: 155EP

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech III Year I Semester Examinations, March - 2024

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 is a data warehouse?	[2]
b)	What are the characteristics of an interesting pattern?	[3]
c)	Define confidence.	[2]
d)	Give examples for multi-level association rule.	[3]
e)	What is meant by rule accuracy?	[2]
f)	What are the merits of lazy learner?	[3]
g)	What is the need of Jaccard measure?	[2]
h)	What are the basic requirements of clustering?	[3]
i)	Define recall.	[2]
j)	Give examples for non-spatial data in spatial database.	[3]

PART – B

(50 Marks)

2. Describe the five data mining task primitives with illustrations. [10]

OR

3. How to integrate a data mining system with a data warehouse? Explain with example. [10]

4. Explain Apriori algorithm for finding frequent item sets of a transactional database. [10]

OR

5. How to apply association rule mining to graph databases? Illustrate graph mining. [10]

6.a) Discuss the importance of information gain in decision tree induction.
b) With suitable data, demonstrate the working of k nearest neighbor classifier. [5+5]

OR

7. Explain Bayesian classification with an example. [10]

8. Explain the concept of hierarchical clustering and complete linkage algorithm for clustering. [10]

OR

9. What is the need of detecting outliers? Explain the applications of outlier analysis. [10]

10. What is a data stream? Explain the characteristics and challenges in mining data streams. [10]

OR

11. Explain the features of multimedia database and the significance of similarity search in multimedia data. [10]

---ooOoo---