

Code No: 137BQ

R16

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech IV Year I Semester Examinations, March - 2021

DATA MINING

(Common to CSE, IT)

Time: 3 Hours

Max. Marks: 75

Answer any Five Questions
All Questions Carry Equal Marks

--

- 1.a) How to handle redundancy in data integration?
b) Explain principal component analysis as a method of dimensionality reduction. [7+8]
2. How can we mine closed frequent item sets? Explain. [15]

3. Explain market basket analysis and its relevance to association rule. Explain the Apriori algorithm using the following transactional data assuming that the support count is 22%. Illustrate with an example.

TID	LIST OF ITEMS
001	milk, dal, sugar, bread
002	Dal, sugar, wheat, jam
003	Milk, bread, curd, paneer
004	Wheat, paneer, dal, sugar
005	Milk, paneer, bread
006	Wheat, dal, paneer, bread.

[15]

4. Discuss K- Nearest neighbor classification-Algorithm and Characteristics. [15]
5. How Neural Networks can be used for Data classification? Which algorithm is suitable? Explain them with example? [15]

6. Explain various issues and challenges in data mining. [15]

- 7.a) Describe web usage mining.
b) Explain about Text Clustering with an illustrative example. [7+8]

- 8.a) Write and explain about the k-medoids algorithm.
b) Describe distance based outlier detection. [8+7]

--ooOoo--