

Code No: 126EW

R13

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD

B. Tech III Year II Semester Examinations, May - 2017

DATA WAREHOUSING AND DATA MINING

(Information Technology)

Time: 3 hours

Max. Marks: 75

Note: This question paper contains two parts A and B.

Part A is compulsory which carries 25 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 10 marks and may have a, b, c as sub questions.

PART - A

(25 Marks)

- 1.a) What is the significance of Data warehousing? [2]
- b) How Data mining is different from KDD? [3]
- c) What is Star Schema? [2]
- d) Describe the features of OLTP systems. [3]
- e) What is Data Transformation? [2]
- f) What are the principles of APRIORI algorithms? [3]
- g) Describe Decision Tree construction. [2]
- h) How to measure best split of any classification? [3]
- i) What are the key issues in Hierarchical clustering? [2]
- j) Give an overview of Outlier Detection. [3]

PART - B

(50 Marks)

2. Explain in detail about 3 tier Data Warehousing architecture with a neat sketch. [10]
- OR**
3. Describe the architecture of OLTP with its operation. [10]
- 4.a) What are the challenges of KDD?
 - b) Discuss about Dimensionality Reduction. [5+5]
- OR**
5. Explain the basic Data mining tasks with example. [10]
6. Briefly discuss about different partition algorithms with an example. [10]
- OR**
- 7.a) What is Frequent Item Set Generation? Explain.
 - b) Explain the compact representation of Frequent Item Data Set. [5+5]
- 8.a) What are the general approaches to consider for solving classification problem?
 - b) Describe about different classification techniques. [5+5]
- OR**
- 9.a) How to evaluate any classifier model, which was build for classification.
 - b) Discuss about KNN classification. [5+5]
10. Explain with example how clustering can be with large databases. [10]
11. Discuss about K-means clustering algorithm step by step with an example. [10]