Code No: 126VW **R15**

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
B. Tech III Year II Semester Examinations, December - 2019
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)	Define Data Cube. Give an example.	[2]
b)	What are semi-addictive measures?	[3]
c)	What is Data Binaryzation?	[2]
d)	Describe Data Normalization methods in brief.	[3]
e)	Define the measure confidence in association rule mining.	[2]
f)	Differentiate frequent subsequence and frequent substructure.	[3]
g)	What is Decision tree? Give an example.	[2]
h)	How to evaluate classifier?	[3]
i)	Define Clustering of Data.	[2]
j)	List out all partitioning methods for clustering data.	[3]

PART - B

(50 Marks)

2.a)	Describe 3-tier Architecture of Data Warehouse with a neat sketch.	
b)	Design Fact constellation table with suitable example.	[6+4]
	OR	
3.a)	Describe various OLAP operations performed on Multidimensional Data Model.	
b)	Differentiate ROLAP, MOLAP and HOLAP server functionalities.	[6+4]
4.a)	Explain about various Data Mining Tasks with appropriate examples.	
b)	Discuss about any two measures of similarity.	[6+4]
	OR	
5.a)	"Data preprocessing is necessary before data mining process". Justify your answer.	
b)	Enumerate feature subset selection methods.	[6+4]
6.a)	Illustrate FP-growth algorithm with a suitable example.	
b)	Discuss about maximal frequent Item set.	[7+3]
	OR	
7.a)	Categorize various kinds of Association Rules with examples.	
b)	How to represent Frequent Itemset in compact format?	[6+4]

8.a) Discuss about attribute selection measure ‘Information Gain’.
b) How will you solve a classification problem using Bayesian Belief Networks? [4+6]

OR

9.a) Write about Lazy Learners for classification.
b) Describe kNN Algorithm for data classification with appropriate example. [4+6]

10.a) Discuss about key issues in Hierarchical clustering.
b) Demonstrate DIANA clustering algorithm with example. [5+5]

OR

11.a) What are the advantages of PAM Method?
b) How to cluster the data sets using k-mediod clustering algorithm? [4+6]

---ooOoo---