Sub Name Code: DMDW
Subject Code: IT-4037
Program Name: B.Tech
Semester:VII (Regular)
Year - 2019

AUTUMN MID-SEMESTER - 2019
KIIT, Deemed to be University, Bhubaneswar-24
Data Mining and Data Warehousing
IT 4037

Time: 1 1/2 Hours **Full Mark: 20**

The figures in the margin indicate full marks. Candidates are required to give their answers in their own words as far as practicable and all parts of a question should be answered at one place only.

(Answer any four questions including question No.1 which is compulsory)

Q1.	Answer all the following questions. Provide appropriate example, if necessary.	[1×5 =5]
(a)	Describe the steps of the knowledge discovery process in data mining.	CO1
(b)	Define temporal, sequence and time-series databases.	CO1
(c)	What is data warehouse? List out OLAP operations.	CO3
(d)	What are the methods used to improve the Apriori's Efficiency?	CO2
(e)	What are the techniques to handle missing data in any database?	CO1
Q2.(a)	What is data mining? What are the application and issues of data mining. What is the difference between data base and data mining.	[3] CO1
(b)	Given two objects represented by the tuples (22, 1, 42, 10) and (20,0,36,8), compute the following distance i. Euclidean distance ii. Manhattan distance	[2] CO2
Q3. (a)	Consider the age :23, 23, 27, 27, 39, 41, 47, 49, 50. Use the following normalization to transform the age value 39	[3] CO2
(b)	Illustrate the dimensionality reduction techniques and it's importance.	CO2 [2]
Q4.(a)	Demonstrate the major tasks in data pre-processing.	CO2 [3]
(b)	What are the issues to be considered while data integration?	CO2 [2]
Q5. (a)	Consider the transactional data base with minimum support 22% and minimum confidence 70% . Find out	[3] CO4
(b)	Construct the FP growth tree for the above given transactional database.	CO4 [2]

TID	List of item_IDs
T100	I1, I2, I5
T200	I2, I4
T300	I2, I3
T400	I1, I2, I4
T500	I1, I3
T600	I2,I3
T700	I1, I3
T800	I1,I2,I3,I5

----------ALL THE BEST----------