Reg. No.



A Constituent unit of MAHE. Manipal

VII SEMESTER B.TECH. (COMPUTER SCIENCE & ENGINEERING) END SEMESTER EXAMINATIONS, MAY 2021

SUBJECT: DATA WAREHOUSE AND DATA MINING [CSE 4060]

REVISED CREDIT SYSTEM (--/05/2022)

Time: 3 Hours

MAX. MARKS: 50

Instructions to Candidates:

- ✤ Answer ALL FIVE questions.
- ✤ Missing data may be suitably assumed.

		Μ	CO	BL
1A.	Describe how User Interaction is an issue in Data Mining.	4	1	2
1 B .	Explain any four architectural types of Data Warehouses.	4	1	2
1C.	With the help of a diagram explain the Star Schema Data Model	2	2	2
2A.	With the help of an example describe the general principles and method of application of Type 2 changes to Data Warehouses	5	2	1
2B.	With the help of a diagram describe the steps involved while using replication to capture changes to source data	3	2	1
2C.	Describe any 4 strategies for Data Transformation	2	2	1
JA.	itemsets using the FP-Growth algorithm. Show the detailed steps by constructing the Conditional (Sub-)Pattern Bases and also show the conditional FP-tree associated with the conditional node I3 using pictorial representation.		5	U
	TID List of item_JDs			
	T100I1, I2, I5T200I2, I4T300I2, I3T400I1, I2, I4T500I1, I3T600I2, I3T700I1, I3T800I1, I2, I3, I5T900I1, I2, I3	5		
3B.	For the dataset shown in question 3A find all frequent itemsets using Vertical Data Format	3	3	6
3 C.	With the help of an example describe any 2 pruning strategies involved in mining Closed Itemsets	2	3	1

4A.	Consider the following figure showing a multilayer feed-forward neural network. Let the learning rate be 0.9. The initial weight and bias values of the network are given in Table 1, Classify the tuple, $X = (1, 0, 1)$ with a class label of 1 using Backpropagation algorithm. Show all steps in detail for the first iteration. x_1 y_{15} y_{26} $y_$	5	4	4
	$\frac{x_1 x_2 x_3 w_{14} w_{15} w_{24} w_{25} w_{34} w_{35} w_{46} w_{56} \theta_4 \theta_5 \theta_6}{\theta_5 \theta_6 \theta_6 \theta_6 \theta_7 \theta_8 \theta_8$			
	1 0 1 0.2 -0.3 0.4 0.1 -0.5 0.2 -0.3 -0.2 -0.4 0.2 0.1			
4B.	Describe how Sampling and Dynamic Itemset Counting techniques can help in improving the efficiency of Apriori Algorithm	3	3	1
4 C.	Write an algorithm for classifying tuples using Backpropagation algorithm	2	4	4
5A.	With the help of a diagram explain how clustering can be carried out using k-medoids clustering	4	5	2
5B.	With the help of a figure explain the working CHAMELEON for clustering data.	4	5	2
5C.	With the help of a diagram describe how Support Vector Machines classify data when data are linearly separable	2	4	4