CONTINUOUS INTERNAL EVALUATION- 2

Dept: CSE	Sem / Div:6A&B	Sub: Data Mining and Data Ware housing	S Code:18CS641
Date:25/06/2021	Time:3.00-4.30 PM	Max Marks: 50	Elective: Yes
Note: Answer any 2 full questions, choosing one full question from each part.			

-

QN		Questions	Marks	RBT	COs
		PART A			
1	a	Explain various data mining task with examples.	10	L2	CO2
	b	For the following vectors X, Y, calculate the cosine, Correlation, simple matching coefficient, Hamming distances, Jaccord coefficient. X=(1,1,0,1,0,1) Y=(1,1,1,0,0,1)	10	L3	CO2
	c	What is data mining? With a neat diagram explain KDD process in detail.	5	L2	CO2
		OR			
2	a	Describe various types of attributes and data sets.	10	L2	CO2
	b	Define data preprocessing? Mention the steps involved in it. Explain any two steps in detail.	10	L2	CO2
	c	Consider the following 2 binary vectors, X=(1,0,0,0,0,0,0,0,0,0) Y=(0,0,0,0,0,0,1,0,0,1). Find any two similarity and dissimilarity measures.	5	L2	CO2
		PART B			
3	a	What is frequent item set generation? Explain frequent item-sets generation using Apriori principle.	10	L2	CO3
	b	Generate frequent item set and strong association rules for the given data with support-50% and confidence-70% using Apriori algorithm.	15	L3	CO3
		OR			
4	a	What is association analysis? Explain association rule, confidence and support.	10	L2	CO3
	b	Generate frequent item set and strong association rules for the given data with support-2 and confidence-80% using Apriori algorithm.	15	L3	CO3

TID	10	20	30	40
Items	{1,3,4}	{2,3,5}	{1,2,3,5}	{2,5}

TID	100	200	300	400	500	600	700	800	900
Items	I1,I2, I5	I2,I4	I2,I3	I1,I2, I4	I1,I3	I2,I3	I1,I3	I1,I2, I3,I5	I1,I2, I3

Prepared by: Prof. Thapaswini P S

HOD