Code: 20IT4501E

III B.Tech - I Semester – Regular / Supplementary Examinations NOVEMBER 2023

DATA MINING
(INFORMATION TECHNOLOGY)

Duration: 3 hours   Max. Marks: 70

Note: 1. This paper contains questions from 5 units of Syllabus. Each unit carries 14 marks and have an internal choice of Questions.
2. All parts of Question must be answered in one place.

BL – Blooms Level   CO – Course Outcome

			BL	CO	Max. Marks
		UNIT-I			
1	a)	Explain why Data mining is a single step in KDD Process.	L2	CO1	7 M
	b)	Describe about various Data Mining functionalities in brief.	L2	CO1	7 M
		OR			
2	a)	Define Data mining and explain about data mining on what kind of data?	L2	CO1	7 M
	b)	Explain about the major issues in data mining.	L2	CO1	7 M
		UNIT-II			
3	a)	Explain about measuring the dispersion of data.	L3	CO2	4 M
	b)	What is preprocessing? Explain the forms of Pre-processing.	L2	CO1	10 M

		OR			
4	a)	Interpret basic Statistical Descriptions of data.	L3	CO2	7 M
	b)	List and Explain Data Cleaning techniques.	L2	CO1	7 M

UNIT-III

5	a)	Define the terms Frequent Itemset, Closed Frequent Itemset, Support and Confidence.	L3	CO3	4 M
	b)	Illustrate the procedure of FP-Growth algorithm with example.	L3	CO3	10 M
		OR			
6		Interpret the procedure of Apriori algorithm. Find out the Frequent item sets with candidate generation with min-sup=2 and min_conf=70%. (Database from Table1) Table1: Transactional Database from All Electronics	L3	CO3	14 M

TID	List of Items
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
T900	I1, I2, I3

		UNIT-IV			
7	a)	Explain how prediction is different from classifications and illustrate classification with examples.	L3	CO4	7 M
	b)	Describe Decision Tree Induction algorithm with suitable example.	L3	CO4	7 M
		OR			
8	a)	Explain the algorithm of Naive Bayes Classification with example.	L3	CO4	7 M
	b)	With the help of diagram explain Bagging and Boosting.	L3	CO4	7 M
		UNIT-V			
9	a)	Compare and contrast the differences between Agglomerative Vs Divisive Clustering.	L4	CO5	7 M
	b)	Illustrate partitioning methods with an example.	L4	CO5	7 M
		OR			
10		Suppose that the data mining task is to cluster points (with (x,y) representing location) into three clusters, where points are A1(2,10), A2(2,5), A3(8,4), B1(5,8), B2(7,5), B3(6,4), C1(1,2), C2(4,9). Use K-means algorithm to form three clusters.	L4	CO5	14 M