CLASS: MBA   SEMESTER : III
BRANCH: MBA   SESSION : MO/2023

SUBJECT: MT549 DATA MINING

TIME: 3 Hours   FULL MARKS: 50

INSTRUCTIONS:
1. The question paper contains 5 questions each of 10 marks and total 50 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
4. Before attempting the question paper, be sure that you have got the correct question paper.
5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.

			CO	BL
Q.1(a)	Explain the data mining issues and challenges.	[5]	1	2
Q.1(b)	Describe the various steps involved in data pre-processing.	[5]	1	2
Q.2(a)	Discuss the role of Data Mining in CRM.	[5]	2	2
Q.2(b)	i) Describe KDD process.	[2+3]	2	3
	ii) Calculate accuracy, precision and recall for the given matrix.			

n=165	Predicted: NO	Predicted: YES	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

			CO	BL
Q.3(a)	Define frequent itemset generation and association rule. Explain the working of partition algorithm and how it improves Apriori with the help of an example.	[5]	3	2
Q.3(b)	Apply Apriori on the following dataset to find frequent itemsets and generate association rules [Min. support count =2 and min. confidence is 60%]	[5]	3	3

TID	Items
T1	I1, I2, I5
T2	I2, I4
T3	I2, I3
T4	I1, I2, I4
T5	I1, I3
T6	I2, I3
T7	I1, I3
T8	I1, I2, I3, I5
T9	I1, I2, I3

			CO	BL
Q.4(a)	Using the K-means and Euclidean distance, find two clusters in the following data: {(2,4), (3,5), (4,3), (4,5), (6,7), (5,6), (7,5)}.	[5]	4	3
Q.4(b)	Explain classification. Consider the car robbery data as follows:	[5]	4	3

S#	Color	Type	Origin	Stolen? (Class)
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Solve the above to identify the query for Red Domestic SUV using Naive Bayes classifier.

Q.5(a)	Discuss any two advanced mining techniques.	[5]	5	2
Q.5(b)	Describe the concept of perceptron and hidden layers in a neural network architecture. Explain any 2 types of neural network.	[5]	5	2