Enrollment No.: _____

Darshan UNIVERSITY
योगः कर्मसु कौशलम्

Darshan Institute of Engineering & Technology
B.Tech. | Sem-5 | Summer-2025

Course Code	:	2101CS521	**Date**	:	07-05-2025
Course Name	:	Data Mining	**Duration**	:	150 Minutes
			Total Marks	:	70

Instructions:
1. Attempt all the questions.
2. Figures to the right indicates maximum marks.
3. Make suitable assumptions wherever necessary.

Q.1 **(A)** Are All Patterns Interesting? Justify your answer. **4**

(B) Explain methods to find dissimilarity of numeric data with an example. **3**

OR

Differentiate Symmetric vs Skewed Data.

(C) Explain any seven issues of data mining. **7**

OR

Explain five number summaries with a boxplot.

Q.2 **(A)** Explain different ways to fill missing data. **4**

(B) Explain Equal width binning. **3**

OR

Explain Equal depth binning.

(C) Explain min-max, z-score, decimal scaling with example. **7**

OR

Explain sampling and attribute selection methods of data reduction.

Q.3 **(A)** Explain Maximal and Closed Frequent Itemsets with an example. **4**

(B) Explain support and confidence with example. **3**

OR

Explain Frequent Itemset and Frequent Subsequence with example.

(C) Consider a transactional database. **7**

TID	**Items**
T1	1, 3, 4
T2	2, 3, 5
T3	1, 2, 3, 5
T4	2, 5

Suppose the minimum support count is 2 and minimum confidence threshold is 90%. Find all frequent itemsets using Apriori Algorithm and generate association rules.

OR

Consider a transactional database.

TID	Items
100	M, O, N, K, E, Y
200	D, O, N, K, E, Y
300	M, A, K, E
400	M, U, C, K, Y
500	C, O, O, K, I, E

Suppose the minimum support count is 3. Find all frequent itemsets using FP-growth algorithm.

Q.4 **(A)** Explain Precision and recall for classification model. **4**

(B) Explain ROC curve with an example. **3**

OR

Explain Bootstrap method with an example.

(C) Explain Decision Tree algorithm with an example. **7**

OR

Explain Naive Bayesian classification with an example.

Q.5 **(A)** Differentiate k-mean and k-medoids. **4**

(B) Explain dendrogram with an example. **3**

OR

Differentiate supervised and unsupervised learning.

(C) Explain DBSCAN with algorithmic steps and example. **7**

OR

What is an outlier? Explain outlier detection method with an example.

***