BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(MID SEMESTER EXAMINATION SP/2025)

CLASS: BTECH **SEMESTER : VI**
BRANCH: AIML **SESSION : SP/2025**

SUBJECT: AI303 UNSUPERVISED LEARNING

TIME:02 Hours **FULL MARKS: 25**

INSTRUCTIONS:
1. The question paper contains 5 questions each of 5 marks and total 25 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
***4*. Tables/Data handbook/Graph paper etc., if applicable, will be supplied to the candidates**

				CO	BL
Q.1(a)	For the given data matrix, discretize the feature F4 as follows: 0-5: 0; 6-10: 1, and rewrite the data matrix.	[2]	2	3
Q.1(b)	Treat F4 as nominal and compute the dissimilarity matrix, using the formula for distance computation for mixed attributes. All other features are numeric.	[3]	2	3
Q.2(a)	Write a short note on data visualization.	[2]	1	2
Q.2(b)	What is cluster analysis? Describe with the help of examples.	[3]	1	4
Q.3(a)	Write the nearest neighbor algorithm for agglomerative clustering.	[2]	2	3
Q.3(b)	Illustrate the above algorithm for single link distance computation with help of an example.	[3]	2	3
Q.4(a)	How is the average link distance between two clusters computed?	[2]	2	2
Q.4(b)	Given the data points: {(1,2), (2,2), (2,4), (5,2), (6,1)}, perform Average link clustering to obtain 3 clusters. Update the distance matrix in each iteration and draw the final dendrogram. (Use Manhattan distance for the initial distance computation.)	[3]	2	3
Q.5(a)	Describe the key concepts of Ward's method.	[2]	2	4
Q.5(b)	Apply the Ward's method on the following datapoints and also draw the dendrogram. A (1,2), B(3,3), C(6,5), D(8,8), E(12, 10)	[3]	2	3

	F1	F2	F3	F4	F5
X1	15	6	8	0	7
X2	6	0	1	5	3
X3	8	20	0	8	25
X4	2	5	10	7	4

::::::24/02/2025::::::E