

Data Mining (CSE542)

Homework 08

ID: ____ Name: _조원석_ Date: _2023.05.22_

Task-1

Given the graph in Figure 16.7,

Cluster the graph into two clusters using ratio cut and normalized cut.

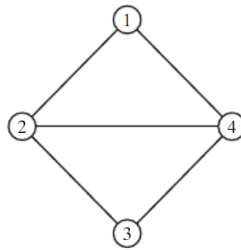


Figure 16.7. Graph

1. Adjacency Weight Matrix – A

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

2. Degree Matrix – D

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}$$

3. Laplacian Matrix – L (Normalized Laplacian Matrix)

$$L = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$$

4. Normalized Asymmetric Laplacian L^a (Normalized Symmetric Laplacian)

$$L^a = \begin{bmatrix} 1 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{3} & 1 & -\frac{1}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 1 \end{bmatrix}$$

For the ratio cut

Task-2 Consider the contingency table,

	iris-setosa	iris-versicolor	iris-virginica	
	T_1	T_2	T_3	n_i
C_1 (squares)	0	47	14	61
C_2 (circles)	50	0	0	50
C_3 (triangles)	0	3	36	39
m_j	50	50	50	$n = 150$

Compute the following measures for each cluster.

1. precision
2. recall
3. F-Measure
4. Jaccard Coefficient

Cluster C1:

$$\text{Precision} = T_1 / (T_1 + T_2 + T_3) = 0 / (0 + 47 + 14) = 0$$

$$\text{Recall} = T_1 / N_i = 50 / 61$$

$$\text{F-Measure} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) = (2 * 0 * 50/61) / (0 + 50/61) = 0$$

$$\text{Jaccard Coefficient} = T_1 / (T_1 + T_2 + T_3) = 0 / (0 + 47 + 14) = 0$$

Cluster C2:

$$\text{Precision} = T_2 / (T_1 + T_2 + T_3) = 0 / (50 + 0 + 0) = 0$$

$$\text{Recall} = T_2 / N_i = 0 / 50 = 0$$

$$\text{F-Measure} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) = (2 * 0 * 0) / (0 + 0) = \text{undefined}$$

$$\text{Jaccard Coefficient} = T_2 / (T_1 + T_2 + T_3) = 0 / (50 + 0 + 0) = 0$$

Cluster C3:

$$\text{Precision} = T3 / (T1 + T2 + T3) = 36 / (0 + 3 + 36) = 36 / 39$$

$$\text{Recall} = T3 / Ni = 36 / 39$$

$$\text{F-Measure} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$= (2 * 36 / 39 * 36 / 39) / (36 / 39 + 36 / 39) = 36$$

$$\text{Jaccard Coefficient} = T3 / (T1 + T2 + T3) = 36 / (0 + 3 + 36) = 36 / 39$$

	Precision	Recall	F-Measure	Jaccard Coefficient
C1	0	$\frac{36}{39}$	0	0
C2	0	0	Undefined	0
C3	$\frac{36}{39}$	$\frac{36}{39}$	36	$\frac{36}{39}$