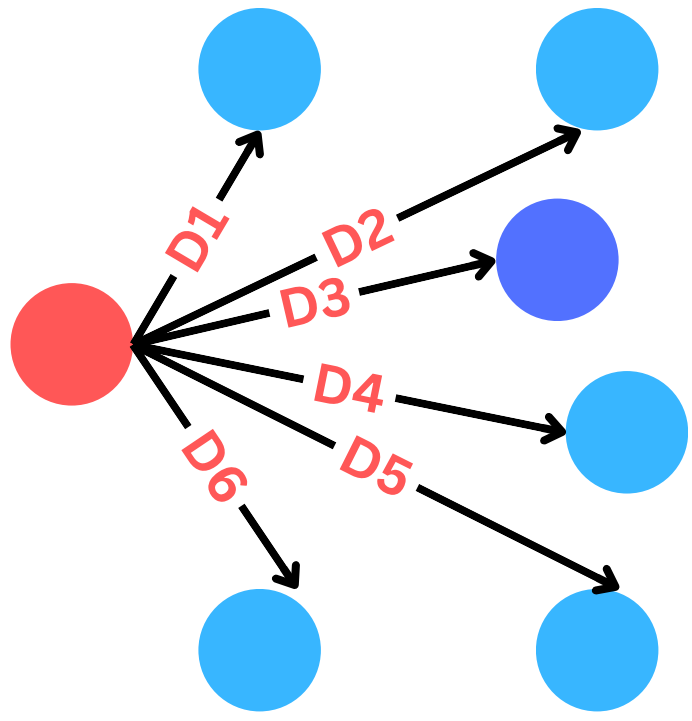
Silhouette Score

In Unsupervised Learning



D: Distance

Point

Centroid

Point (i)

$a(i)$ Dissimilarity of Point (i) from *CI*

$$a(i) = \frac{1}{|C_I| - 1} \sum_{j \in C_I, i \neq j} d(i, j)$$

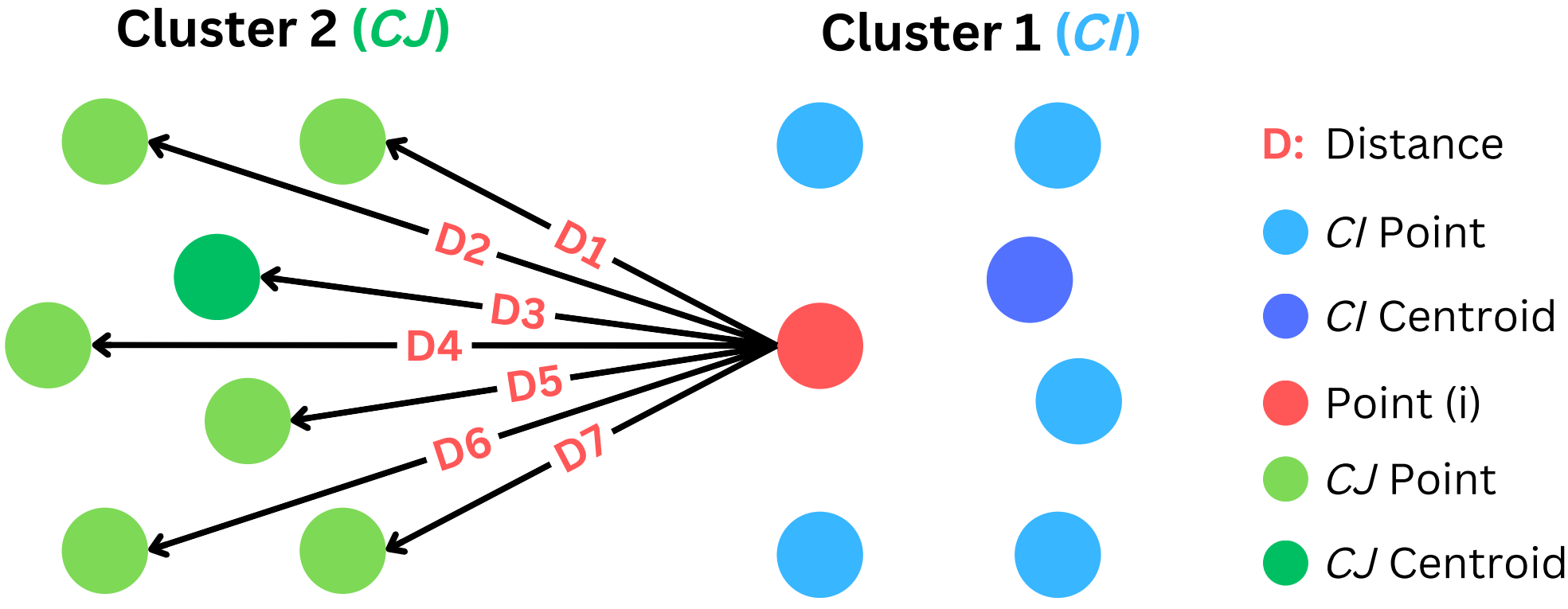
Dissimilarity of Point (i) from *CI* = **Mean Distance** from **Point (i)** to all points in the cluster ***CI***

CODANICS | KAS | CAMPUSX | KRISH NAIK

Arsalan Ali
@arslanchaos

Silhouette Score

In Unsupervised Learning



$b(i)$ Dissimilarity of Point (i) from *CJ*

$$b(i) = \min_{J \neq I} \frac{1}{|C_J|} \sum_{j \in C_J} d(i, j)$$

Dissimilarity of Point (i) from *CJ* = **Mean Distance** from **Point (i)** to all points in the cluster ***CJ***

CODANICS | KAS | CAMPUSX | KRISH NAIK

Arsalan Ali
@arslanchaos

Silhouette Score

In Unsupervised Learning

$a(i)$: Dissimilarity of **Point (i)** from all points of Cluster 1 (***CJ***)

$b(i)$: Dissimilarity of **Point (i)** from all points of Cluster 2 (***CJ***)

$$s(i) = \begin{cases} 1 - a(i)/b(i), & \text{if } a(i) < b(i) \\ 0, & \text{if } a(i) = b(i) \\ b(i)/a(i) - 1, & \text{if } a(i) > b(i) \end{cases}$$

$$-1 \leq s(i) \leq 1$$

Silhouette Score ranges from -1 to 1

High s(i) : If **Point (i)** is close to **1** then it means its properly clustered. It belongs in ***CI***

Low s(i) : If **Point (i)** is close to **-1** then it means its not properly clustered. It belongs in ***CJ***

CODANICS | KAS | CAMPUSX | KRISH NAIK

Arsalan Ali
@arslanchaos