Package 'EfficientOASW'

October 16, 2023

Title Efficient Clustering Algorithms for Optimizing the Average Silhouette Width

Version 0.0.0.9000

Author Minh Long Nguyen <edelweiss611428@gmail.com>

Maintainer Minh Long Nguyen <edelweiss611428@gmail.com>

Description This package implements the original Optimum Silhouette (OSil) clustering algorithm and provides an O(N) faster implementation of the exact OSil algorithm.

License GPL (>= 3)

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.3

LinkingTo Rcpp

Imports Rcpp, cluster, stats

R topics documented:

ASW	*The Average Silhouette Width*

Description

This function computes the Average Silhouette Width (ASW).

Usage

```
ASW(C, dist)
```

Arguments

C	A clustering solution. It must be an integer vector of k unique values 1,2,...,k.
dist	A "dist" object, which can be obtained by the "dist" function.

Value

The ASW of the clustering C with respect to the distance matrix dist.

Author(s)

Minh Long Nguyen <edelweiss611428@gmail.com>

References

Rousseeuw, P.J. (1987) Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math., 20, 53–65.

---

effOSil   *Efficient Optimum Silhouette Clustering Algorithm*

Description

effOSil implements the exact Optimum Silhouette (OSil) algorithm. It is O(N) times faster than the original OSil algorithm at the cost of storing O(N) additional values.

Usage

```
effOSil(dist, k, initClustering = NULL, initMethod = "average")
```

Arguments

dist	A "dist" object, which can be obtained by the "dist" function.
k	The number of clusters.
initClustering	A user-specified initialized clustering. It must be an integer vector of k unique values 1,2,...,k. By defaualt, initClustering is set to NULL. If initClustering is NULL, initMethod is used; otherwise, initClustering is used.
initMethod	An initialization method. By default, effOSil uses average-linkage clustering ("average") for initialization. Other options include well-established clustering algorithms such as Partition Around Medoids ("pam"), single-linkage clustering ("single"), complete-linkage clustering ("complete"), the Ward's method ("ward.D"), etc.

Value

Clustering The OSil clustering solution.

ASW The ASW associated with the OSil clustering.

nIter The number of iterations needed for convergence.

Author(s)

Minh Long Nguyen <edelweiss611428@gmail.com>

OSil   *Original Optimum Silhouette Clustering Algorithm*

Description

OSil implements the original Optimum Silhouette algorithm.

Usage

```
OSil(dist, k, initClustering = NULL, initMethod = "average")
```

Arguments

dist	A "dist" object, which can be obtained by the "dist" function.
k	The number of clusters.
initClustering	A user-specified initialized clustering. It must be an integer vector of k unique values 1,2,...,k. By defaualt, initClustering is set to NULL. If initClustering is NULL, initMethod is used; otherwise, initClustering is used.
initMethod	An initialization method. By default, effOSil uses average-linkage clustering ("average") for initialization. Other options include well-established clustering algorithms such as Partition Around Medoids ("pam"), single-linkage clustering ("single"), complete-linkage clustering ("complete"), the Ward's method ("ward.D"), etc.

Value

Clustering The OSil clustering solution.

ASW The ASW associated with the OSil clustering.

nIter The number of iterations needed for convergence.

Author(s)

Minh Long Nguyen <edelweiss611428@gmail.com>

References

Batool F. (2019). Optimum average silhouette width clustering. *PhD Thesis*, University College London.

Index