

Package ‘ddc’

December 14, 2022

Type Package

Title Distance Density Clustering Algorithm

Version 1.0.1

Date 2022-12-13

Maintainer Ruizhe Ma <maruizhe.cs@gmail.com>

Description A distance density clustering (DDC) algorithm in R. DDC uses dynamic time warping (DTW) to compute a similarity matrix, based on which cluster centers and cluster assignments are found. DDC inherits dynamic time warping (DTW) arguments and constraints. The cluster centers are centroid points that are calculated using the DTW Barycenter Averaging (DBA) algorithm. The clustering process is divisive. At each iteration, cluster centers are updated and data is reassigned to cluster centers. Early stopping is possible. The output includes cluster centers and clustering assignment, as described in the paper (Ma et al (2017) <[doi:10.1109/ICDMW.2017.11](https://doi.org/10.1109/ICDMW.2017.11)>).

License GPL (>= 2)

Depends R (>= 4.2)

Imports dtw (>= 1.22), dtwclust (>= 5.5), parallel (>= 4.2), magrittr (>= 2.0), utils

RoxygenNote 7.2.1

Encoding UTF-8

ByteCompile yes

BuildResaveData no

NeedsCompilation yes

Suggests knitr, rmarkdown, spelling, testthat (>= 3.0.0)

Language en-US

Config/testthat/edition 3

VignetteBuilder knitr

Author Ruizhe Ma [cre, aut],
Bing Jiang [aut]

Repository CRAN

Date/Publication 2022-12-14 12:00:05 UTC

R topics documented:

createDistMatrix	2
createLabelMatrix	3
createStandardMatrix	3
ddc	4

Index	5
--------------	----------

createDistMatrix	<i>Create the dataframe of the Dissimilarity matrix</i>
------------------	---

Description

Use the DTW to generate the matrix

Usage

```
createDistMatrix(standard_matrix, output_dir = NULL, mc.cores = 1, ...)
```

Arguments

standard_matrix	the matrix generated by function 'createStandardMatrix'
output_dir	the file to save the dissimilarity matrix data
mc.cores	the number of cores would be used in parallel
...	the same parameters which would be used in 'dtw' for calculating the distances of events

Value

the matrix, which describes pairwise distinction between M objects. It is a square symmetrical ' $M \times M$ ' matrix with the (ij) th element equal to the value of a chosen measure of distinction between the (i) th and the (j) th object.

Examples

```
original_data <- data.frame("1"=c(1, 2, 1), "2"=c(5,6,7),
  "3"=c(4, 5, 8), "4"=c(3, 1, 9))
standard_matrix <- createStandardMatrix(data = original_data)
dist_matrix <- createDistMatrix(standard_matrix = standard_matrix)
```

createLabelMatrix *Create the dataframe with event names and the related labels*

Description

Create the dataframe with event names and the related labels

Usage

```
createLabelMatrix(data, output_dir = NULL)
```

Arguments

data data structure as the files in "UCR Time Series Classification Archive"
output_dir the file to save the label matrix data

Value

the dataframe, including event names and labels

Examples

```
original_data <- data.frame("1"=c(1, 2, 1), "2"=c(5,6,7),  
                          "3"=c(4, 5, 8), "4"=c(3, 1, 9))  
label_matrix <- createLabelMatrix(data = original_data)
```

createStandardMatrix *Create the dataframe, only including the event data*

Description

Create the dataframe, only including the event data

Usage

```
createStandardMatrix(data, output_dir = NULL)
```

Arguments

data data structure as the files in "UCR Time Series Classification Archive"
output_dir the file to save the standard matrix data

Value

the dataframe of event data

Examples

```
original_data <- data.frame("1"=c(1, 2, 1), "2"=c(5,6,7),
  "3"=c(4, 5, 8), "4"=c(3, 1, 9))
standard_matrix <- createStandardMatrix(data = original_data)
```

ddc

Execute DDC to cluster the dataset

Description

Execute DDC to cluster the dataset

Usage

```
ddc(dist_matrix, standard_matrix, label_matrix, end_cluster_num = NULL, ...)
```

Arguments

dist_matrix	the created dist matrix
standard_matrix	the original data matrix
label_matrix	the matrix including events and labels
end_cluster_num	the max number of cluster when the procedue ends
...	including: mc.cores(cores used in parallel), the dtw parameters like step.pattern, keep, mc.cores

Value

the cluster array as a result, including 'Centroid', 'Elements' and 'DBAValue' for each cluster

Examples

```
original_data <- data.frame("1"=c(1, 2, 1), "2"=c(5,6,7),
  "3"=c(4, 5, 8), "4"=c(3, 1, 9))
standard_matrix <- createStandardMatrix(data = original_data)
label_matrix <- createLabelMatrix(data = original_data)
dist_matrix <- createDistMatrix(standard_matrix = standard_matrix)
result <- ddc(dist_matrix=dist_matrix, standard_matrix=standard_matrix,
  label_matrix=label_matrix, end_cluster_num=2)
```

Index

`createDistMatrix`, [2](#)
`createLabelMatrix`, [3](#)
`createStandardMatrix`, [3](#)
`ddc`, [4](#)