

Package ‘RJcluster’

September 23, 2020

Title RJ Clustering Algorithm

Version 0.1.0

Description

Clustering algorithm for high dimensional data. This algorithm is ideal for data where $n \ll p$.

License GPL (≥ 2)

Encoding UTF-8

LazyData true

Imports Rcpp ($\geq 1.0.2$), matrixStats, infotheo, rlang, stats,
graphics, profvis, mclust, doParallel, foreach, parallel

LinkingTo Rcpp, RcppArmadillo

Suggests testthat ($\geq 2.1.0$), knitr, rmarkdown

RoxygenNote 7.1.1

VignetteBuilder knitr

Depends R (≥ 2.10)

NeedsCompilation yes

Author Rachael Shudde [aut, cre],
Shahina Rahman [aut],
Valen Johnson [aut]

Maintainer Rachael Shudde <rachael.shudde@gmail.com>

Repository CRAN

Date/Publication 2020-09-23 08:00:06 UTC

R topics documented:

f_rez	2
generateSimulationData	2
RJclust	3
scaleRJ	4

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

f_rez *Calculate ami and nmi*

Description

Calculate ami and nmi

Usage

```
f_rez(v1, v2)
```

Arguments

v1 first cluster vector
v2 second cluster vector

Value

Returns ami and nmi values (a value between 0 and 1) that measures how close the classification between the two clusters is. A value closer to 1 means the clusters are more similar across v1 and v2.

Examples

```
cluster1 <- sample(1:5, size = 10, replace = TRUE)  
cluster2 <- sample(1:2, size = 10, replace = TRUE)  
f_rez( cluster1, cluster2 )
```

generateSimulationData
generateSimulationData

Description

generateSimulationData

Usage

```
generateSimulationData(sigma = 1, sparsity = 0.02, seed = 1234)
```

Arguments

sigma noise level - default is 1
sparsity percentage of relevant features - default is 0.02
seed Random seed - default is 1234

Value

Returns simulation data - a 21000x1000 sparse matrix with 4 clusters (2 clusters are n = 10,000 and 2 clusters are n = 500)

Examples

```
X = generateSimulationData()
```

RJclust

RJclust

Description

RJclust

Usage

```
RJclust(X, num_cut = NULL, seed = 1)
```

Arguments

X	Data input
num_cut	Number of cuts for RJ algorithm (suggested sqrt(p))
seed	Seed (default = 1)

Value

Returns RJ algorithm result

Examples

```
X = generateSimulationData()
X = X$X
clust = RJclust(X, 100)
```

scaleRJ	<i>scaleRJ</i>
---------	----------------

Description

scaleRJ

Usage

```
scaleRJ(X, medians = FALSE)
```

Arguments

X	Numeric matrix to be scaled
medians	A logical, if TRUE the medians of the columns are used to scaled

Value

The scaled matrix

Examples

```
data = generateSimulationData()  
Xscaled = scaleRJ(data$X)
```

Index

f_rez, [2](#)

generateSimulationData, [2](#)

RJclust, [3](#)

scaleRJ, [4](#)