

Package ‘protoshiny’

May 18, 2022

Type Package

Title Interactive Dendrograms for Visualizing Hierarchical Clusters
with Prototypes

Version 0.1.0

Description Shiny app to interactively visualize hierarchical clustering with prototypes.
For details on hierarchical clustering with prototypes, see
Bien and Tibshirani (2011) <[doi:10.1198/jasa.2011.tm10183](https://doi.org/10.1198/jasa.2011.tm10183)>. This package currently launches the application.

License GPL-3

Encoding UTF-8

Imports DT, dynamicTreeCut, methods, protoclust, rare, shiny,
shinycssloaders, shinythemes, tools

Depends R (>= 4.0.0)

RoxygenNote 7.1.2

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Andee Kaplan [aut, cre],
Jacob Bien [aut],
Mike Bostock [cph] (Author of included d3.js library),
Brian Reavis [cph] (Author of included selectize.js library)

Maintainer Andee Kaplan <andee.kaplan@colostate.edu>

Repository CRAN

Date/Publication 2022-05-18 16:30:08 UTC

R topics documented:

as.protoclust	2
covid19	3
faces	3
get_clusters	3

get_cut_from_merge_id	4
get_paths_from_cut	4
get_server	5
get_ui	5
movies	5
seafrow	5
visualize_hc	6
words	6
Index	7

as.protoclust	<i>Convert a generic hclust object to a protoclust object.</i>
---------------	--

Description

Convert a generic hclust object to a protoclust object.

Usage

```
as.protoclust(hc, protos)
```

Arguments

hc	A generic hclust object using any linkage.
protos	A vector of prototype label indices for each branch in the dendrogram. The i-th element is the index of the prototype corresponding to the cluster formed on the i-th merge.

Value

An object of class `protoclust` with a user specified vector of representative prototypes for each cluster. This object is just like `hclust` but has an identical element:

merge, height, order	identical to the values returned by <code>hclust</code>
protos	a vector of length n - 1. The i-th element is the index of the prototype corresponding to the cluster formed on the i-th merge.

covid19	<i>Dataset covid19</i>
---------	------------------------

Description

This is a protoclust object for use with protoshiny. It is the hierarchical clustering of US counties using mobility data (proportion of people staying home at least 8 hours and proportion of people going into a restaurant).

Source

<https://cmu-delphi.github.io/covidcast/covidcastR/>

faces	<i>Dataset faces</i>
-------	----------------------

Description

This is a protoclust object for use with protoshiny. It is the hierarchical clustering of 400 grayscale images of faces that correspond to 10 images of 40 individuals.

Source

<https://www.kaggle.com/kasikrit/att-database-of-faces>

get_clusters	<i>Extract cluster labels</i>
--------------	-------------------------------

Description

Extract cluster labels

Usage

```
get_clusters(hc)
```

Arguments

hc A protoshiny object downloaded from the web interface.

Value

A vector of saved cluster assignments resulting from interaction in the protoshiny application.

get_cut_from_merge_id *Get clustering of leaves from merge_id*

Description

This function expects a set of node id's of a hierarchical clustering object that would define a partition of the leaves. It returns a vector giving the implied clustering of the leaves (analogous to the output of 'stats::cutree').

Usage

```
get_cut_from_merge_id(hc, merge_id)
```

Arguments

hc	An object of class 'hclust'
merge_id	A numeric vector. Indexing of nodes should match that of the 'merge' matrix in an 'hclust' object. In particular, positive numbers specify interior nodes and negative numbers specify leaves. These nodes must define a partition of the leaves, meaning no node in 'merge_id' should be a descendant of another and every leaf should either be included directly or be the descendant of one of the interior nodes referenced by 'merge_id'.

get_paths_from_cut *Given a clustering get the root-to-node paths for visible nodes*

Description

This function returns the binary path from root to node for each node that should be displayed in order to return the input clustering. Note that this will only be possible for a clustering 'cl' in which each cluster corresponds to a branch of 'hc'.

Usage

```
get_paths_from_cut(hc, cl)
```

Arguments

hc	An object of class 'hclust'
cl	A numeric vector in the format of the output of 'stats::cutree' that gives the cluster memberships of each leaf.

get_server	<i>Server function</i>
------------	------------------------

Description

Server function

Usage

```
get_server()
```

get_ui	<i>UI function</i>
--------	--------------------

Description

UI function

Usage

```
get_ui()
```

movies	<i>Dataset movies</i>
--------	-----------------------

Description

This is a protoclust object for use with protoshiny. It is the hierarchical clustering of 13,816 movies based on users' ratings and taggings of movies.

Source

<https://grouplens.org/datasets/movielens/25m/>

seafLOW	<i>Dataset seafLOW</i>
---------	------------------------

Description

This is a protoclust object for use with protoshiny. It is the hierarchical clustering of 6,336 scatter-plots resulting from continuous-time flow cytometry measurements of phytoplankton.

Source

Provided by Mattias Rolf Cape and Francois Ribalet.

visualize_hc	<i>Run a local instance of protoshiny</i>
--------------	---

Description

Run a local instance of protoshiny

Usage

```
visualize_hc()
```

Value

No return value, launches the protoshiny Shiny application.

Examples

```
## Not run:  
  visualize_hc()  
  
## End(Not run)
```

words	<i>Dataset words</i>
-------	----------------------

Description

This is a protoclust object for use with protoshiny. It is the hierarchical clustering of 19,599 of the most common words from Grolier's Encyclopedia.

Source

<https://cs.nyu.edu/~roweis/data.html>

Index

`as.protoclust`, 2

`covid19`, 3

`faces`, 3

`get_clusters`, 3

`get_cut_from_merge_id`, 4

`get_paths_from_cut`, 4

`get_server`, 5

`get_ui`, 5

`hclust`, 2

`movies`, 5

`protoclust`, 2

`seafLOW`, 5

`visualize_hc`, 6

`words`, 6