Package 'lilikoi'

July 30, 2018

Title Metabolomics Personalized Pathway Analysis Tool

Version 0.1.0

Description Computes the pathway deregulation score for a given set of metabolites, selects the pathways with the highest mutual information and then uses them to build a classifier. F. Alakwaa, S. Huang, and L. Garmire (2018) <doi:10.1101/283408>.

Imports caret, corrplot, devtools, dplyr, e1071, gbm, ggplot2, glmnet, hash, Hmisc, infotheo, Matrix, pamr, R.oo, princurve, pROC, randomForest, reshape2, RWeka, stringr

Depends R (>= 3.4.0)

License GPL-2

Encoding UTF-8

LazyData true

URL https://github.com/lanagarmire/lilikoi

BugReports https://github.com/lanagarmire/lilikoi/issues

RoxygenNote 6.0.1.9000

Suggests testthat

NeedsCompilation no

Author Fadhl Alakwaa [aut, cre], Sijia Huang [aut]

Maintainer Fadhl Alakwaa <falakwaa@hawaii.edu>

Repository CRAN

Date/Publication 2018-07-30 11:10:03 UTC

R topics documented:

`data.hmdb` *data.hmdb*

Description

data.hmdb

Usage

```
data.hmdb
```

Format

data.frame $ Accession : Factor w/ 53287 levels 'HMDB0000001',..: 1 1 2 2 2 2 3 3 3 3 ... $ Name : Factor w/ 53287 levels '(±)-Glycerol 1,2-diacetate',..: 45 45 55 55 55 55 146 146 146 146 ... $ Pathway_Name: Factor w/ 48607 levels '11-beta-hydroxylase deficiency (CYP11B1)',..: 46317 46318 98 23433 46265 48580 19 23491 46174 46176 ...

Source

http://www.hmdb.ca/system/downloads/current/hmdb_metabolites.zip

---

`data.metaboAnalyst` *data.metaboAnalyst*

Description

data.metaboAnalyst

Usage

```
data.metaboAnalyst
```

Format

data.frame Classes 'tbl_df', 'tbl' and 'data.frame': 183937 obs. of 8 variables: $ hmdb_id : chr 'HMDB0000001' 'HMDB0000001' 'HMDB0000001' 'HMDB0000001' ... $ name : chr '1-Methylhistidine' '1-Methylhistidine' '1-Methylhistidine' '1-Methylhistidine' ... $ synonym : chr '(2S)-2-amino-3-(1-Methyl-1H-imidazol-4-yl)propanoic acid' ... $ pubchem_id: int 92105 92105 92105 92105 92105 92105 92105 92105 92105 ... $ chebi_id : int 50599 50599 50599 50599 50599 50599 50599 50599 50599 50599 ... $ kegg_id : chr 'C01152' 'C01152' 'C01152' 'C01152' ... $ metlin_id : int 3741 3741 3741 3741 3741 3741 3741 3741 3741 3741 ...

Source

http://www.metaboanalyst.ca/MetaboAnalyst/faces/docs/Resources.xhtml

---

`data.smpdb` *data.smpdb*

Description

data.smpdb

Usage

```
data.smpdb
```

Format

data.frame 'data.frame': 19245 obs. of 2 variables: $ pathway: Factor w/ 705 levels '11-beta-hydroxylase deficiency (CYP11B1)',..: 39 39 39 39 39 39 39 39 39 39 ... $ hmdb_id: chr 'HMDB0000538' 'HMDB0000161' 'HMDB0000045' 'HMDB0000250' ...

Source

http://smpdb.ca/downloads/smpdb_metabolites.csv.zip

`lilikoi.adjust_model` *Model Adjustment method using clinical factors*

Description

This function can adjust the best performing model using additional clinical factors specified by the user It plots ROC for three models: a model building using only the selected pathways, a model built using the clinical factors and a model built using both the selected pathways and the clinical factors.

Usage

```
lilikoi.adjust_model(mlResults, PDSmatrix, selected_Pathways_Weka,
  metaboliteMeasurements, clinical_factors_data, factors)
```

Arguments

mlResults	The models and other results generated from machine_learning method
PDSmatrix	the PDS matrix generated using lilikoi.get_pd_scores function
selected_Pathways_Weka	selected pathway using WEKA algorithm generated from lilikoi.select_pathways function
metaboliteMeasurements	the metaboliteMeasurements for the samples
clinical_factors_data	Metadata
factors	which the users want to add to the model

---

lilikoi.get_pd_scores *A lilikoi.get_pd_scores Function*

Description

This function allows you to compute Pathway Deregulation Score deriving make sure that you have the below database for the metabolites and pathway list: meta_path.RData

Usage

```
lilikoi.get_pd_scores(metaboliteMeasurements, metabolitePathwayTable,
  pathwayMetaboliteDatabase = lilikoi::data.smpdb, maxit = 200)
```

Arguments

metaboliteMeasurements	Metabolite levels
metabolitePathwayTable	This is the metabolitePathwayTable from lilikoi.metab_to_pathway function. This table includes the metabolite ids and their corresponding HMDB ids
pathwayMetaboliteDatabase	An external database of metabolites and their associated pathways.
maxit	Max iterations for princurve in pathifier algorithm. Lowering increases speed but lowers performance

Examples

```
filename <- system.file("extdata", "plasma_breast_cancer.csv", package = "lilikoi")
data <- read.csv(file = filename, check.names = FALSE, row.names = 1)[, 1:20]
metaboliteNames <- colnames(data)[-1]
matches <- lilikoi.metab_to_pathway(metaboliteNames, "name")
PDSmatrix <- lilikoi.get_pd_scores(data, matches, lilikoi::data.smpdb[1:23,], maxit = 1)
```

lilikoi.machine_learning

A machine learning Function

Description

This function for classification using 7 different machine learning algorithms and it plots the ROC curves and the AUC, SEN, and specificity

Usage

```
lilikoi.machine_learning(PDSmatrix, measurementLabels, significantPathways)
```

Arguments

PDSmatrix	from lilikoi.get_pd_scores
measurementLabels	Cancer or not
significantPathways	Pathways with high significance from lilikoi.select_pathways function

lilikoi.metab_to_pathway

A lilikoi.metab_to_pathway Function

Description

This function allows you to convert your metabolites id such as names, kegg ids, pubchem ids. into pathways. Metabolites which have no pathways will be excluded from any downstream analysis.

Usage

```
lilikoi.metab_to_pathway(metaboliteNames, searchType)
```

Arguments

metaboliteNames	Your metabolite data
searchType	The type of the metabolites id such as 'name', 'kegg', 'hmdb','pubchem'

Examples

```
matches = lilikoi.metab_to_pathway(c("Asparagine", "Hypotaurine", "5-Oxoproline"), "name")

matches = lilikoi.metab_to_pathway(c("C00152", "C00519", "C01879"), "kegg")
```

`lilikoi.select_pathways` *This function will return the significant pathways that can be used for training machine learning models.*

Description

This function can reduce the pathway dimension using method specified.

Usage

```
lilikoi.select_pathways(PDSmatrix, metaboliteMeasurements, threshold = 0.5,
  method = "info")
```

Arguments

`PDSmatrix`	from lilikoi.get_pd_scores function
`metaboliteMeasurements`	Levels
`threshold`	Cutoff for significance
`method`	Which gain method to use

Index