Package 'ScRNAIMM'

November 17, 2023

Title Performing Single-Cell RNA-Seq Imputation by Using Mean/Median
Imputation

Version 0.1

Maintainer Mohamed Soudy <Mohmedsoudy2009@gmail.com>

Description Performing single-cell imputation in a way that preserves the biological varia-
tions in the data. The package clusters the input data to do imputation for each clus-
ter, and do a distribution check using the Anderson-Darling normality test to impute dropouts us-
ing mean or median (Yazici, B., & Yolacan, S. (2007) <DOI:10.1080/10629360600678310>).

License GPL-3

Encoding UTF-8

RoxygenNote 7.2.3

Imports magrittr, dplyr, nortest, mclust, scDHA

NeedsCompilation no

Author Mohamed Soudy [aut, cre],
Sascha Jung [aut],
Antonio DEL SOL [aut]

Repository CRAN

Date/Publication 2023-11-17 20:10:10 UTC

R topics documented:

---

cluster_cells   *Perform cell clustering based on scDHA method*

Description

Cluster cells based on scDHA methods to get cluster labels

Usage

```
cluster_cells(ScRNA_filtered, Normalize = TRUE, k=NULL, n=5000)
```

Arguments

ScRNA_filtered	ScRNA-seq data set generated by filter_ScRNA function
Normalize	Boolean parameter whether to apply log10 normalization for the data or not
k	Number of clusters if there is a prior knowledge about that
n	Number of genes to keep after feature selection step

Value

a vector that contains the cell labels

Author(s)

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

---

evaluate_clustering   *Evaluate the clustering if you have the original labels*

Description

Evaluate the clustering if you have the original labels

Usage

```
evaluate_clustering(cluster_labels, original_labels)
```

Arguments

cluster_labels	Cluster labels generated by cluster_cells functions or user-defined
original_labels	Original labels of the ScRNA-seq data

Value

ARI of clustering 'a value between 0 and 1' 1 indicates best clustering

Author(s)

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

Examples

```
evaluate_clustering(c(1,1,1,1,2,2,3,3), c(1,1,1,1,3,3,3,2))
```

---

filter_ScRNA   *Remove genes which are not expressed in at least one cell*

Description

Filter out the genes that are not expressed in at least one cell type

Usage

```
filter_ScRNA(ScRNA_mat)
```

Arguments

ScRNA_mat   ScRNA-Seq matrix where genes are in rows and cells are in columns

Value

Filtered ScRNA-seq matrix that contains genes that are at least expressed in one cell type

Author(s)

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

---

prepare_dataset   *Prepare the data set for the imputation*

Description

This function aims to get the indices where all genes are zeros or genes that is expressed on only one sample per cell

Usage

```
prepare_dataset(filtered_data, cluster_labels)
```

Arguments

filtered_data   ScRNA-seq data set generate by filter_ScRNA function

cluster_labels   cell labels

Value

a data frame that contains the processed ScRNA-seq data

Author(s)

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

---

run_pipeline *Run the main pipeline for ScRNAIMM*

Description

Run the main pipeline for ScRNAIMM

Usage

```
run_pipeline(ScRNA,label=NULL,k=NULL,cells=TRUE,genes=TRUE,outdir=NULL,dataset=NULL)
```

Arguments

ScRNA	ScRNA-seq data set generated by prepare_dataset function
label	Prior knowledge about cluster labels if NULL, will use our clustering function
k	Prior knowledge about number of clusters if NULL, will use our clustering function
cells	Boolean whether to do the imputation based on cell clustering or not
genes	Boolean whether to do the imputation based on genes or not
outdir	Path to output directory to write the imputed data
dataset	Name of the data set to be the name of the output directory

Value

a data frame with the imputed values

Author(s)

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

ScRNA_imp_MM *Perform ScRNA-seq imputation using mean/Median*

Description

Perform ScRNA-seq imputation using mean/Median

Usage

```
ScRNA_imp_MM(ScRNA_filtered, cluster_labels = NULL, cells = TRUE, genes = FALSE)
```

Arguments

ScRNA_filtered	ScRNA-seq data set generated by prepare_dataset function
cluster_labels	Cluster labels generated by cluster_cells function or user-defined
cells	Boolean whether to do the imputation based on cell clustering or not
genes	Boolean whether to do the imputation based on genes or not

Value

a data frame with the imputed values

Author(s)

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

---

scRNA_MMI *Performs a distribution check for the data*

Description

Performs a distribution check for the data

Usage

```
scRNA_MMI(scRNA_dataset)
```

Arguments

scRNA_dataset	ScRNA-seq data set

Value

ScRNA-Seq Imputed data set

Author(s)

Mohamed Soudy <Mohmedsoudy2009@gmail.com>

Index