Package 'SaturnCoefficient'

December 10, 2024

Encoding UTF-8

Title Statistical Evaluation of a UMAP Dimensionality Reduction

Version 1.0

Description
A function that reads an input matrix, its dimensionality reduction produced by UMAP, and evaluates the quality of this dimensionality reduction by producing a real value in the [0; 1] interval. We call this real value Saturn coefficient. A higher value means better dimensionality reduction; a lower value means worse dimensionality reduction.

BugReports https://github.com/davidechicco/SaturnCoefficient/issues

Depends R (>= 4.0.0)

License GPL-3

URL https://github.com/davidechicco/SaturnCoefficient

Imports dplyr, MatrixCorrelation, ProjectionBasedClustering, stats, qpdf, umap

Suggests knitr, rmarkdown

VignetteBuilder knitr

RoxygenNote 7.3.1

NeedsCompilation no

Author Davide Chicco [aut, cre] (<https://orcid.org/0000-0001-9655-7142>)

Maintainer Davide Chicco <davidechicco@davidechicco.it>

R topics documented:

`continuity_score` *Function that calculates the continuity score of a UMAP dimensionality reduction*

Description

Function that calculates the continuity score of a UMAP dimensionality reduction

Usage

```
continuity_score(original_matrix, umap_output_layout, VERBOSE)
```

Arguments

`original_matrix`	input matrix
`umap_output_layout`	output matrix of UMAP applied to original_matrix
`VERBOSE`	prints some intermediate message to standard output or not

Value

a real value containing the continuity score

Examples

```
this_nrows <- 200
this_ncols <- 100
this_min <- 0
this_max <- 10000
noise_random_matrix <- matrix(runif(n = this_nrows * this_ncols,
     min = this_min, max = this_max), nrow = this_nrows)
input_matrix <- as.matrix(noise_random_matrix)
these_nearest_neighbors <- 15
this_min_dist <- 0.05

library("umap")
custom.settings <- umap::umap.defaults
custom.settings$"n_neighbors" <- these_nearest_neighbors
custom.settings$"min_dist" <- this_min_dist

x_umap <- umap::umap(input_matrix, config=custom.settings)

this_verbose <- FALSE
thisCon <- continuity_score(input_matrix, x_umap$"layout",  this_verbose)
cat("continuity = ", thisCon, "\n", sep="")
```

Saturn_coefficient *Function that calculates the Saturn coefficient to quantify the quality of a UMAP dimensionality reduction*

Description

Function that calculates the Saturn coefficient to quantify the quality of a UMAP dimensionality reduction

Usage

```
Saturn_coefficient(original_matrix, umap_output_layout, VERBOSE)
```

Arguments

original_matrix	input matrix
umap_output_layout	output matrix of UMAP applied to original_matrix
VERBOSE	prints some intermediate message to standard output or not

Value

a real value containing the Saturn coefficient

Examples

```
this_nrows <- 200
this_ncols <- 100
this_min <- 0
this_max <- 10000
noise_random_matrix <- matrix(runif(n = this_nrows * this_ncols,
     min = this_min, max = this_max), nrow = this_nrows)
input_matrix <- as.matrix(noise_random_matrix)
these_nearest_neighbors <- 15
this_min_dist <- 0.05

library("umap")
custom.settings <- umap::umap.defaults
custom.settings$"n_neighbors" <- these_nearest_neighbors
custom.settings$"min_dist" <- this_min_dist

x_umap <- umap::umap(input_matrix, config=custom.settings)

this_verbose <- FALSE
thisSaturn <- Saturn_coefficient(input_matrix, x_umap$"layout",  this_verbose)
cat("Saturn coefficient = ", thisSaturn, "\n", sep="")
```

three_metrics *Function that calculates the Saturn coefficient, trustworthiness score, and the continuity score of a UMAP dimensionality reduction*

Description

Function that calculates the Saturn coefficient, trustworthiness score, and the continuity score of a UMAP dimensionality reduction

Usage

```
three_metrics(original_matrix, umap_output_layout, VERBOSE)
```

Arguments

original_matrix	input matrix
umap_output_layout	output matrix of UMAP applied to original_matrix
VERBOSE	prints some intermediate message to standard output or not

Value

a dataframe containing the Saturn coefficient, the trustworthiness score, and the continuity score

Examples

```
this_nrows <- 200
this_ncols <- 100
this_min <- 0
this_max <- 10000
noise_random_matrix <- matrix(runif(n = this_nrows * this_ncols,
     min = this_min, max = this_max), nrow = this_nrows)
input_matrix <- as.matrix(noise_random_matrix)
these_nearest_neighbors <- 15
this_min_dist <- 0.05

library("umap")
custom.settings <- umap::umap.defaults
custom.settings$"n_neighbors" <- these_nearest_neighbors
custom.settings$"min_dist" <- this_min_dist

x_umap <- umap::umap(input_matrix, config=custom.settings)

this_verbose <- FALSE
theseThreeMetrics <- three_metrics(input_matrix, x_umap$"layout",  this_verbose)
print(theseThreeMetrics)
```

`trustworthiness_score` *Function that calculates the trustworthiness score of a UMAP dimensionality reduction*

Description

Function that calculates the trustworthiness score of a UMAP dimensionality reduction

Usage

```
trustworthiness_score(original_matrix, umap_output_layout, VERBOSE)
```

Arguments

`original_matrix`	input matrix
`umap_output_layout`	output matrix of UMAP applied to original_matrix
`VERBOSE`	prints some intermediate message to standard output or not

Value

a real value containing the trustworthiness score

Examples

```
this_nrows <- 200
this_ncols <- 100
this_min <- 0
this_max <- 10000
noise_random_matrix <- matrix(runif(n = this_nrows * this_ncols,
     min = this_min, max = this_max), nrow = this_nrows)
input_matrix <- as.matrix(noise_random_matrix)
these_nearest_neighbors <- 15
this_min_dist <- 0.05

library("umap")
custom.settings <- umap::umap.defaults
custom.settings$"n_neighbors" <- these_nearest_neighbors
custom.settings$"min_dist" <- this_min_dist

x_umap <- umap(input_matrix, config=custom.settings)

this_verbose <- FALSE
thisTW <- trustworthiness_score(input_matrix, x_umap$"layout",  this_verbose)
cat("trustworthiness = ", thisTW, "\n", sep="")
```

Index