Package: BASiNET (via r-universe)

October 23, 2026

Title Classification of RNA Sequences using Complex Network Theory

Version 0.0.5

Description It makes the creation of networks from sequences of RNA, with this is done the abstraction of characteristics of these networks with a methodology of threshold for the purpose of making a classification between the classes of the sequences. There are four data present in the 'BASiNET' package, ``sequences'', ``sequences2'', ``sequences-predict'' and ``sequences2-predict'' with 11, 10, 11 and 11 sequences respectively. These sequences were taken from the data set used in the article (LI, Aimin; ZHANG, Junying; ZHOU, Zhongyin, 2014) <doi:10.1186/1471-2105-15-311>, these sequences are used to run examples. The BASiNET was published on Nucleic Acids Research, (ITO, Eric; KATAHIRA, Isaque; VICENTE, Fábio; PEREIRA, Felipe; LOPES, Fabrício, 2018) <doi:10.1093/nar/gky462>.

Depends R (>= 3.4.0)

License GPL-3

Encoding UTF-8

Imports igraph, Biostrings, RWeka, randomForest, rmcfs, grDevices, graphics, stats, rJava

biocViews Software, BiologicalQuestion, GenePrediction

RoxygenNote 6.1.0

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Eric Augusto Ito [aut] (<https://orcid.org/0000-0002-0030-6881>), Fabricio Martins Lopes [aut, cre] (<https://orcid.org/0000-0002-8786-3313>)

Maintainer Fabricio Martins Lopes <fabricio@utfpr.edu.br>

Config/pak/sysreqs zlib1g-dev

Repository https://cran.r-universe.dev

Date/Publication 2022-08-19 07:50:25 UTC

RemoteUrl https://github.com/cran/BASiNET

RemoteRef HEAD

RemoteSha 774d97a18f5622d3c8009c6bb452976a55654756

Contents

classification 2
createGraph2D 3
createNet 4
measures 5
minMax 5
reschedule 6
threshold 7

Index **8**

`classification` *Performs the classification methodology using complex network theory*

Description

Given two distinct data sets, one of mRNA and one of lncRNA. The classification of the data is done from the structure of the networks formed by the sequences. After this is done classifying with the J48 classifier and randomForest. Can be also created in the current directory a file of type arff called' result 'with all values so that it can be used later. There is also the graphic parameter that when TRUE generates graphs based on the results of each measure. Using the J48 classifier it is possible to generate a tree based on the dataset and then save this tree so that it can be used to predict other RNA sequences

Usage

```
classification(mRNA, lncRNA, word = 3, step = 1, sncRNA, graphic,
  classifier = c("J48", "RF"), load, save)
```

Arguments

`mRNA`	Directory where the file .FASTA lies with the mRNA sequences
`lncRNA`	Directory where the file .FASTA lies with the lncRNA sequences
`word`	Integer that defines the size of the word to parse. By default the word parameter is set to 3
`step`	Integer that determines the distance that will be traversed in the sequences for creating a new connection. By default the step parameter is set to 1
`sncRNA`	Directory where the file .FASTA lies with the sncRNA sequences (OPTIONAL)

graphic	Parameter of the logical type, TRUE or FALSE for graphics generation. As default graphic gets FALSE
classifier	Character Parameter. By default the classifier is J48, but the user can choose to use randomForest by configuring as classifier = "RF". The prediction with a model passed by the param load only works with the classifier J48.
load	When defined this parameter will be loaded the file which is the model previously saved in the current directory with the name entered in this parameter. No file is loaded by default
save	when set, this parameter saves a .arff file with the results of the features in the current directory and also saves the tree created by the J48 classifier so that it can be used to predict RNA sequences. This parameter sets the file name. No file is created by default

Value

Results with cross-validation or the prediction result

Author(s)

Eric Augusto Ito

Examples

```
# Classification - cross validation
library(BASiNET)
arqSeqMRNA <- system.file("extdata", "sequences2.fasta", package = "BASiNET")
arqSeqLNCRNA <- system.file("extdata", "sequences.fasta", package = "BASiNET")
classification(mRNA=arqSeqMRNA,lncRNA=arqSeqLNCRNA)
classification(mRNA=arqSeqMRNA,lncRNA=arqSeqLNCRNA, save="example") #Save Tree to Predict Sequences
# Prediction
mRNApredict <- system.file("extdata", "sequences2-predict.fasta", package = "BASiNET")
lncRNApredict <- system.file("extdata", "sequences-predict.fasta", package = "BASiNET")
modelPredict <- system.file("extdata", "modelPredict.dat", package = "BASiNET")
classification(mRNApredict,lncRNApredict,load=modelPredict)
```

---

createGraph2D *Creates a two-dimensional graph between a measure and the threshold*

Description

For an analysis of each measure, the createGraph2D () function was created in order to visualize the behavior of each measurement in relation to the threshold. This function creates a graph (Measure x Threshold) from an array, mRNA sequences are given the blue color, the lncRNA sequences are given a red color. In cases where there is a third class this will be given the green color

Usage

```
createGraph2D(matrix, numSeqMRNA, numSeqLNCRNA, nameMeasure)
```

Arguments

matrix	matrix of the measure for the creation of two-dimensional graph
numSeqMRNA	Integer number of mRNA sequences
numSeqLNCRNA	Integer number of lncRNA sequences
nameMeasure	Character Parameter that defines the name of the measure to put in the title of the graph

Author(s)

Eric Augusto Ito

---

createNet *Creates an untargeted graph from a biological sequence*

Description

A function that from a biological sequence generates a graph not addressed having as words vertices, this being able to have its size parameter set by the' word 'parameter. The connections between words depend of the' step 'parameter that indicates the next connection to be formed

Usage

```
createNet(word, step, sequence)
```

Arguments

word	This integer parameter decides the size of the word that will be formed
step	It is the integer parameter that decides the step that will be taken to make a new connection
sequence	It is a vector that represents the sequence

Value

Returns the non-directed graph formed through the sequence

Author(s)

Eric Augusto Ito

measures *Abstracting Characteristics on Network Structure*

Description

Given a graph, it is made up of several features on the graph structure and returns a vector with the data obtained

Usage

```
measures(graph)
```

Arguments

graph	The complex network that will be measured

Value

Return a vector with the results of the measurements in order: Average shortest path length, clustering Coefficient, degree, assortativity, betweenness, standard deviation, maximum, minimum, number of motifs size 3 and number of motifs of size 4

Author(s)

Eric Augusto Ito

minMax *Minimum and maximum*

Description

Verifies the minimum and maximum values of the results.

Usage

```
minMax(matrix, mRNA, lncRNA, sncRNA, rangeMinMax)
```

Arguments

matrix	Array with results numerics
mRNA	Integer number of mRNA sequences
lncRNA	Integer number of lncRNA sequences
sncRNA	Integer number of sncRNA sequences
rangeMinMax	Vector that will be returned with the minimum and maximum values

Value

Returns the vector with the minimum and maximum values for the scale

Author(s)

Eric Augusto Ito

`reschedule` *Rescales the results between values from 0 to 1*

Description

Given the results the data is rescaled for values between 0 and 1, so that the length of the sequences does not influence the results. The rescaling of the mRNA and lncRNA are made separately

Usage

```
reschedule(matrix, mRNA, lncRNA, sncRNA, rangeMinMax)
```

Arguments

`matrix`	Array with results numerics
`mRNA`	Integer number of mRNA sequences
`lncRNA`	Integer number of lncRNA sequences
`sncRNA`	Integer number of sncRNA sequences
`rangeMinMax`	Vector with the minimum and maximum values for the scale

Value

Returns the array with the rescaled values

Author(s)

Eric Augusto Ito

`threshold` *Applies threshold on the network from a value*

Description

Given an integer value X, a cut, that is, edges that are cut will be assigned zero. This cut will be done in the network where the edges have a weight less than the value of X.

Usage

```
threshold(x, net)
```

Arguments

`x`	Integer value that would limit the edges
`net`	Complex network where the edges will be cut

Value

Returns the complex network with the cuts already made

Author(s)

Eric Augusto Ito

Index

classification, 2
createGraph2D, 3
createNet, 4

measures, 5
minMax, 5

reschedule, 6

threshold, 7