Package ‘SLModels’

October 12, 2022

Type Package

Title Stepwise Linear Models for Binary Classification Problems under Youden Index Optimisation

Version 0.1.2

Depends stats, ROCR

Maintainer Rocio Aznar-Gimeno <raznar@itainnova.es>

Description Stepwise models for the optimal linear combination of continuous variables in binary classification problems under Youden Index optimisation. Information on the models implemented can be found at Aznar-Gimeno et al. (2021) <doi:10.3390/math9192497>.

License GPL-3

Encoding UTF-8

NeedsCompilation no

Author Rocio Aznar-Gimeno [aut, cre] (<https://orcid.org/0000-0003-1415-146X>), Luis Mariano Esteban [aut] (<https://orcid.org/0000-0002-3007-302X>), Gerardo Sanz [aut] (<https://orcid.org/0000-0002-6474-2252>), Rafael del Hoyo-Alonso [aut] (<https://orcid.org/0000-0003-2755-5500>)

Repository CRAN

Date/Publication 2022-02-03 14:30:10 UTC

R topics documented:

SLModels .......................................................... 1

Index 4

SLModels

Stepwise Linear Models for Binary Classification Problems under Youden Index Optimisation
Description

Stepwise models for the optimal linear combination of continuous variables in binary classification problems under Youden Index optimisation. Information on the models implemented can be found at Aznar-Gimeno et al. (2021) <doi:10.3390/math9192497>.

Usage

SLModels(data, algorithm="stepwise", scaling=FALSE)

Arguments

data Data frame containing the input variables and the binary output variable. The last column must be reserved for the output variable.

algorithm string; Stepwise linear model to be applied. The options are: "stepwise", "min-max", "minmaxmedian", "minmaxiqr"; default value: "stepwise".

scaling boolean; if TRUE, the Min-Max Scaling is applied; if FALSE, no normalisation is applied to the input variables; default value: FALSE.

Details

The "stepwise" algorithm refers to our proposed stepwise algorithm on the original variables which is the adaptation for the maximisation of the Youden index of the one proposed by Esteban et al. (2011) <doi:10.1080/02664761003692373>. The general idea of this approach, as suggested by Pepe and Thompson (2000) <doi:10.1093/biostatistics/1.2.123>, is to follow a step by step algorithm that includes a new variable in each step, selecting the best combination (or combinations) of two variables, in terms of maximising the Youden index.

The "minmax" algorithm refers to the distribution-free min–max approach proposed by Liu et al. (2011) <doi:10.1002/sim.4238>. The idea is to reduce the order of the linear combination beforehand by considering only two markers (maximum and minimum values of all the variables/biomarkers). This algorithm was adapted in order to maximise the Youden index.

The "minmaxmedian" algorithm refers to our proposed algorithm that considers the linear combination of the following three variables: the minimum, maximum and median values of the original variables.

The "minmaxiqr" algorithm refers to our proposed algorithm that considers the linear combination of the following three variables: the minimum, maximum and interquartile range (IQR) values of the original variables.

More information on the implemented algorithms can be found in Aznar-Gimeno et al. (2021) <doi:10.3390/math9192497>.

Value

Optimal linear combination that maximises the Youden index. Specifically, the function returns the coefficients for each variable, optimal cut-off point and Youden Index achieved.

Note

The "stepwise" algorithm becomes a computationally intensive problem when the number of variables exceeds 4.
Author(s)
Rocío Aznar-Gimeno, Luis Mariano Esteban, Gerardo Sanz, Rafael del Hoyo-Alonso

References


Examples
```r
#Create dataframe
x1<-rnorm(100, sd =1)
x2<-rnorm(100, sd =2)
x3<-rnorm(100, sd =3)
x4<-rnorm(100, sd =4)
z <- rep(c(1,0), c(50,50))
DT<-data.frame(cbind(x1,x2,x3,x4))
data<-cbind(DT,z)

#Example 1#
SLModels(data) #default values: algorithm="stepwise", scaling=FALSE
#Example 2#
SLModels(data, algorithm="minmax") #scaling=FALSE, default value
#Example 3#
SLModels(data, algorithm="minmax", scaling=TRUE)
#Example 4#
SLModels(data, algorithm="minmaxmedian", scaling=TRUE)
#Example 5#
SLModels(data, algorithm="minmaxiqr", scaling=TRUE)
```
Index

SLModels, 1