

Package ‘iForm’

April 22, 2018

Type Package

Title Forward Selection Under Marginality Principle

Version 1.0

Date 2016-03-29

Author Kirk Gosik

Maintainer Kirk Gosik <kgosik@broadinstitute.org>

Description Extended variable selection approaches to jointly model main and interaction effects from high-dimensional data.

License GPL-3

LazyData TRUE

RoxygenNote 6.0.1.9000

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

R topics documented:

| | |
|-------------|---|
| iForm | 1 |
| iformselect | 2 |

iForm

Interaction Screening for Ultra-High Dimensional Data

Description

Extended variable selection approaches to jointly model main and interaction effects from high-dimensional data originally proposed by Hao and Zhang (2014) and extended by Gosik and Wu (2016). Based on a greedy forward approach, their model can identify all possible interaction effects through two algorithms, iFORT and iFORM, which have been proved to possess sure screening property in an ultrahigh-dimensional setting.

Arguments

| | |
|--------------|--|
| formula | an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted. The details of model specification are given under 'Details'. |
| data | data.frame of your data with the response and all p predictors |
| strong | logical TRUE to use strong heredity or FALSE to use weak heredity (default TRUE) |
| higher_order | logical TRUE indicating to include order-3 interactions in the search (default FALSE) |

Details

Runs the iForm selection procedure on the dataset and returns a linear model of the final selected model.

Value

a summary of the linear model returned after the selection procedure

Author(s)

Kirk Gosik

| | |
|-------------|---|
| iformselect | <i>Inner workings for different selections under different higher-orders and strength of marginality.</i> |
|-------------|---|

Description

Extended variable selection approaches to jointly model main and interaction effects from high-dimensional data originally proposed by Hao and Zhang (2014) and extended by Gosik and Wu (2016). Based on a greedy forward approach, their model can identify all possible interaction effects through two algorithms, iFORT and iFORM, which have been proved to possess sure screening property in an ultrahigh-dimensional setting.

Arguments

| | |
|-----------|--|
| x | design matrix of predictors |
| y | response variable |
| p | the size of the predictor set |
| n | the size of the number of observations |
| candidate | the current candidate set of predictors to select from |
| solution | the current set of predictors already selected |
| model | the set of predictors to use in the final model |

| | |
|---------------------------|---|
| <code>bic</code> | the cutoff value for determining the model set |
| <code>step</code> | the step in the iteration currently on |
| <code>strong</code> | indicator of the strength of marginality to be used |
| <code>higher_order</code> | logical TRUE indicating to include order-3 interactions in the search (default FALSE) |

Details

Runs the iForm selection procedure on the dataset and returns a linear model of the final selected model.

Value

a summary of the linear model returned after the selection procedure

Author(s)

Kirk Gosik