Package 'ehaGoF'

May 8, 2026

Type Package

Title Calculates Goodness of Fit Statistics

Version 0.1.1

Author Alper Gulbe [cre],
Ecevit Eyduran [aut]

Maintainer Alper Gulbe <alper.gulbe@igdir.edu.tr>

Description
Calculates 15 different goodness of fit criteria. These are; standard deviation ratio (SDR), coefficient of variation (CV), relative root mean square error (RRMSE), Pearson's correlation coefficients (PC), root mean square error (RMSE), performance index (PI), mean error (ME), global relative approximation error (RAE), mean relative approximation error (MRAE), mean absolute percentage error (MAPE), mean absolute deviation (MAD), coefficient of determination (R-squared), adjusted coefficient of determination (adjusted R-squared), Akaike's information criterion (AIC), corrected Akaike's information criterion (CAIC), Mean Square Error (MSE), Bayesian Information Criterion (BIC) and Normalized Mean Square Error (NMSE).

Imports graphics, stats

License GPL-2

Encoding UTF-8

LazyData FALSE

RoxygenNote 7.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2020-08-11 14:10:15 UTC

Contents

Goodness of Fit 2
Goodness of Fit - Coefficient of Variation 4
Goodness of Fit - Mean Absolute Percentage Error 5
Goodness of Fit - Pearson's Correlation Coefficients 7
Goodness of fit - Relative Root Mean Square Error 8

Goodness of Fit - Root Mean Square Error 10
Goodness of Fit - Standard Deviation Ratio 11
Goodness of Fit : Adjusted Coefficient of Determination (Adjusted R-Squared) 12
Goodness of Fit : Adjusted R-Squared 14
Goodness of Fit : Akaike's Information Criterion 15
Goodness of Fit : Coefficient of Determination (R-Squared) 17
Goodness of Fit : Corrected Akaike's Information Criterion 18
Goodness of Fit : Global Relative Approximation Error 19
Goodness of Fit : Mean Absolute Deviation 21
Goodness of Fit : Mean Error 22
Goodness of Fit : Mean Relative Approximation Error 23
Goodness of Fit : Performance Index 25
Goodness of Fit : R-Squared 26

Index **28**

`Goodness of Fit` *Goodness of Fit*

Description

Tests predicted and observed values for the goodness of fit with various criteria. The goodness of fit tests are used to test how well the model fits. Measures of goodness of fit typically summarize the argument between targets or observed values and the values expected or predicted under the model in question.

Usage

```
GoF(Observations, Predicts,
                nTermInAppr = 2,
                ndigit = 3,
                RMSE = TRUE,
                RRMSE = TRUE,
                SDR = TRUE,
                CV = TRUE,
                PC = TRUE,
                PI = TRUE,
                ME = TRUE,
                RAE = TRUE,
                MRAE = TRUE,
                MAPE = TRUE,
                MAD = TRUE,
                RSq = TRUE,
                ARSq = TRUE,
                AIC = TRUE,
                CAIC = TRUE)
```

Arguments

`Observations`	Observed values or target vector.
`Predicts`	Predicted values. Values produced or fitted by approximation or regression.
`nTermInAppr`	Number of terms used in approximation or regression model. Generally 2 for simple linear model. Default is 2.
`ndigit`	Number of digits in decimal places. Default is 3.
`RMSE`	Whether to show Root Mean Square Error statistics. Default is TRUE.
`RRMSE`	Whether to show Relative Root Mean Square Error statistics. Default is TRUE.
`SDR`	Whether to show Standard Deviation Ratio statistics. Default is TRUE.
`CV`	Whether to show Coefficient of Variance statistics. Default is TRUE.
`PC`	Whether to show Pearson's Correlation Coefficients statistics. Default is TRUE.
`PI`	Whether to show Performance Index statistics. Default is TRUE.
`ME`	Whether to show Mean Error statistics. Default is TRUE.
`RAE`	Whether to show Global Relative Approximation Error statistics. Default is TRUE.
`MRAE`	Whether to show Modified Relative Approximation Error statistics. Default is TRUE.
`MAPE`	Whether to show Mean Absolute Percentage Error statistics. Default is TRUE.
`MAD`	Whether to show Mean Absolute Deviation statistics. Default is TRUE.
`RSq`	Whether to show Coefficient of Determination (R-Squared) statistics. Default is TRUE.
`ARSq`	Whether to show Adjusted Coefficient of Determination (Adjusted R-Squared) statistics. Default is TRUE. Warning: nTermInAppr must be supplied.
`AIC`	Whether to show Akaike's Information Criterion statistics. Default is TRUE. Warning: nTermInAppr must be supplied.
`CAIC`	Whether to show Corrected Akaike's Information Criterion statistics. Default is TRUE. Warning: nTermInAppr must be supplied.

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan, Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas and Wilhelm Grzesiak, Pakistan J. Zool., vol. 49(1), pp 257-265, 2017.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 9
inputs <- 0:9

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*2 + rnorm(10)

# linear regression model
model<-lm(targets~inputs)

# About the model
summary(model)

# Number of Terms
n = length(model$coefficients)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit statistics
GoF(targets, predicted, nTermInAppr=n)
```

Goodness of Fit - Coefficient of Variation

Coefficient of Variation.

Description

Calculates and returns goodness of fit - coefficient of variation (CV).

Usage

```
gofCV(Obs, Prd, dgt=3)
```

Arguments

Obs	Observed or measured values or target vector.
Prd	Predicted or fitted values by the model. Values produced by approximation or regression.
dgt	Number of digits in decimal places. Default is 3.

Value

CoeficientOfVariation

Goodness of fit - coefficient of variation (CV).

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of Different Data Mining Algorithms for Prediction of Body Weight From Several Morphological Measurements in Dogs - S Celik, O Yilmaz

Examples

```
# dummy inputs, independent variable
# integers from 0 to 19
inputs <- 0:19

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(20)

# linear regression model
model<-lm(targets~inputs)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit - coefficient of variation (CV)
gofCV(targets, predicted)
```

Goodness of Fit - Mean Absolute Percentage Error

Mean Absolute Percentage Error

Description

Calculates and returns the goodness of fit criterion: mean absolute percentage error (MAPE), a.k.a. mean absolute percentage deviation or MAPD.

Usage

```
gofMAPE(Obs, Prd, dgt=3)
```

Arguments

Obs	Observed values or targets.
Prd	Predicted or expre-ected values produced by the model.
dgt	Number of digits in decimal places. Default is 3.

Details

Mean absolute percentage error (MAPE) is a measure of prediction accuracy of a forecasting method in statistics. It is commonly used as a loss function for regression problems and in model evaluation, for its very intuitive interpretation in terms of relative error. It usually expresses accuracy as a percentage.

Value

MAPE	Mean absolute percentage error (MAPE) of given set.

Note

For more information look at these papers:

Rob J. Hyndman, Anne B. Koehler, Another look at measures of forecast accuracy, International Journal of Forecasting, Volume 22, Issue 4, 2006, Pages 679-688, ISSN 0169-2070,

Arnaud de Myttenaere, Boris Golden, Bénédicte Le Grand, Fabrice Rossi, Mean Absolute Percentage Error for regression models, Neurocomputing, Volume 192, 2016, Pages 38-48, ISSN 0925-2312,

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan, Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas and Wilhelm Grzesiak, Pakistan J. Zool., vol. 49(1), pp 257-265, 2017.

Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms - Daniel Zaborski, Muhammad Ali, Ecevit Eyduran, Wilhelm Grzesiak, Mohammad Masood Tariq, Ferhat Abbas, Abdul Waheed, Cem Tirink - Pakistan journal of zoology, 2019

Comparison of Different Data Mining Algorithms for Prediction of Body Weight From Several Morphological Measurements in Dogs - S Celik, O Yilmaz

Examples

```
# dummy inputs, independent variable
# integers from 0 to 9
inputs <- 0:9

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*2 + rnorm(10)

# linear regression model
model<-lm(targets~inputs)
```

```
# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit - mean absolute percentage error statistics
gofMAPE(targets, predicted)
```

Goodness of Fit - Pearson's Correlation Coefficients

Pearson's Correlation Coefficients

Description

Calculates and returns Pearson's correlation coefficients (PC).

Usage

```
gofPC(Obs, Prd, dgt = 3)
```

Arguments

Obs	Observed or measured values or target vector.
Prd	Predicted or fitted values by the model. Values produced by approximation or regression.
dgt	Number of digits in decimal places. Default is 3.

Value

PearsonCorrelation

Pearson's correlation coefficients (PC)

Author(s)

Prof. Dr. Ecevit EYDURAN, TA. Alper GULBE

References

OBILOR Esezi Isaac, AMADI Eric Chikweru, Test for Significance of Pearson's Correlation Coefficient, International Journal of Innovative Mathematics, Statistics & Energy Policies 6(1):11-23, Jan-Mar, 2018.

Reza Soleimani, Amir Hossein Saeedi Dehaghani, Alireza Bahadori, A new decision tree based algorithm for prediction of hydrogen sulfide solubility in various ionic liquids, Journal of Molecular Liquids, Volume 242, 2017, Pages 701-713, ISSN 0167-7322, https://doi.org/10.1016/j.molliq.2017.07.075. (http://www.sciencedirect.com/science/article/pii/S0167732217305123)

Examples

```
# dummy inputs, independent variable
# integers from 0 to 19
inputs <- 0:19

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(20)

# linear regression model
model<-lm(targets~inputs)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit - Pearson's correlation coefficient
gofPC(targets, predicted)
```

Goodness of fit - Relative Root Mean Square Error

Relative Root Mean Square Error

Description

Calculates and returns relative root mean square error (RRMSE) of the model. The ratio of the mean of square root of residuals squared to the mean of observed values.

Usage

```
gofRRMSE(Obs, Prd, dgt = 3)
```

Arguments

Obs	Observed values or target vector.
Prd	Predicted values. Values produced by approximation or regression.
dgt	Number of digits in decimal places. Default is 3.

Details

RRMSE is calculated by dividing RMSE by the mean of observed values.

Value

RelativeRootMeanSquareError

Relative root mean square error (RRMSE) of given set.

Note

For more information: Despotovic, M., Nedic, V., Despotovic, D., & Cvetanovic, S., Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation, Renewable and Sustainable Energy Reviews, Volume 56, April 2016, Pages 246-260, ISSN 1364-0321, http://dx.doi.org/10.1016/j.rser.2015.11.058.

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan, Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas and Wilhelm Grzesiak, Pakistan J. Zool., vol. 49(1), pp 257-265, 2017.

Indirect Estimation of Structural Parameters in South African Forests Using MISR-HR and LiDAR Remote Sensing Data - Precious Nokuthula Wistebaar Mahlangu, Renaud Mathieu, Konrad Wessels, Laven Naidoo, Michel M Verstraete, Gregory P Asner, Russell Main

Examples

```
# Input values, independent variable
input <- 0:4

# Target vector, observed values, dependent variable
target <- c(1.9, 4.1, 5.89, 7.9, 10.01)

# Simple linear regression, target across input like: target = a * input + b,
# where a and b are coefficients.
model <- lm(target~input)

# Information about the model
summary(model)

# Values predicted by the model
predicted <- predict(model)

# using library ehaGoF for goodness of fit
library(ehaGoF)

# Goodness of fit - relative root mean square error (RRMSE)
gofRRMSE(target, predicted)
```

Goodness of Fit - Root Mean Square Error
Root Mean Square Error

Description

Calculates and returns root mean square error (RMSE).

Usage

```
gofRMSE(Obs, Prd, dgt = 3)
```

Arguments

Obs	Observed or measured values or target vector.
Prd	Predicted or fitted values by the model. Values produced by approximation or regression.
dgt	Number of digits in decimal places. Default is 3.

Value

RootMeanSquareError
Root mean square error (RMSE)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan - Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas, Wilhelm Grzesiak

Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms - Daniel Zaborski, Muhammad Ali, Ecevit Eyduran, Wilhelm Grzesiak, Mohammad Masood Tariq, Ferhat Abbas, Abdul Waheed, Cem Tirink - Pakistan journal of zoology, 2019

Indirect Estimation of Structural Parameters in South African Forests Using MISR-HR and LiDAR Remote Sensing Data - Precious Nokuthula Wistebaar Mahlangu, Renaud Mathieu, Konrad Wessels, Laven Naidoo, Michel M Verstraete, Gregory P Asner, Russell Main, Remote Sens. 2018, 10, 1537 ; doi:10.3390/rs10101537.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 19
inputs <- 0:19

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(20)

# linear regression model
model<-lm(targets~inputs)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit - root mean square error (RMSE)
gofRMSE(targets, predicted)
```

Goodness of Fit - Standard Deviation Ratio

Standard Deviation Ratio

Description

Calculates and returns standard deviation ratio (SDR).

Usage

```
gofSDR(Obs, Prd, dgt=3)
```

Arguments

`Obs`	Observed values or target vector.
`Prd`	Predicted values. Values produced by approximation or regression.
`dgt`	Number of digits in decimal places. Default is 3.

Value

`StandardDeviationRatio`
: Standard deviation ratio (SDR) of given set.

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gülbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan, Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas and Wilhelm Grzesiak, Pakistan J. Zool., vol. 49(1), pp 257-265, 2017.

Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms - Daniel Zaborski, Muhammad Ali, Ecevit Eyduran, Wilhelm Grzesiak, Mohammad Masood Tariq, Ferhat Abbas, Abdul Waheed, Cem Tirink - Pakistan journal of zoology, 2019

Comparison of Different Data Mining Algorithms for Prediction of Body Weight From Several Morphological Measurements in Dogs - S Celik, O Yilmaz

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==>  Define data, use random,
##--or do  help(data=index)  for the standard data sets.

# Input values, independent variable
input <- 0:4

# Target vector, observed values
target <- c(1.9, 4.1, 5.89, 7.9, 10.01)

# Simple linear regression target across input like target = a * input + b,
# where a and b are coefficients.
model <- lm(target~input)

# Information about the model
summary(model)

# Values predicted by the model
predicted <- predict(model)

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit - standard deviation ratio (SDR)
gofSDR(target, predicted)
```

`Goodness of Fit : Adjusted Coefficient of Determination (Adjusted R-Squared)`

Adjusted Coefficient of Determination (Adjusted R-Squared)

Description

Calculates and returns adjusted coefficient of determination (adjusted R-squared).

Usage

```
gofACoD(Obs, Prd, nTermInAppr = 2, dgt = 3)
```

Arguments

`Obs`	Observed or measured values or target vector.
`Prd`	Predicted or fitted values by the model. Values produced by approximation or regression.
`nTermInAppr`	Number of terms in approximation or regression models formula, interception included. For simple linear regression with one independent variable is simply 2. Default is 2.
`dgt`	Number of digits in decimal places. Default is 3.

Value

`AdjustedCoefficientofDetermination`
: Goodness of fit - adjusted coefficient of determination (adjusted R-squared)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of Different Data Mining Algorithms for Prediction of Body Weight From Several Morphological Measurements in Dogs - S Celik, O Yilmaz.

A new decision tree based algorithm for prediction of hydrogen sulfide solubility in various ionic liquids - Reza Soleimani, Amir Hossein Saeedi Dehaghani, Alireza Bahadori.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 19
inputs <- 0:19

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(20)

# linear regression model
model<-lm(targets~inputs)

# About the model
summary(model)

# Number of Terms
n = length(model$coefficients)

# model's predicted values against targets
predicted<-model$fitted.values
```

```
# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit : adjusted coefficient of determination (adjusted R-squared)
gofACoD(targets, predicted, dgt=4,nTermInAppr=n)
```

`Goodness of Fit : Adjusted R-Squared`

Adjusted R-Squared (Adjusted Coefficient of Determination)

Description

Caclulates and returns adjusted coefficient of determination (adjusted R-squared).

Usage

```
gofARSq(Obs, Prd, nTermInAppr = 2, dgt = 3)
```

Arguments

`Obs`	Observed or measured values or target vector.
`Prd`	Predicted or fitted values by the model. Values produced by approximation or regression.
`nTermInAppr`	Number of terms in approximation or regression models formula, interception included. For simple linear regression with one independent variable is simply 2. Default is 2.
`dgt`	Number of digits in decimal places. Default is 3.

Value

`ARsquared`	Goodness of fit - adjusted coefficient of determination (adjusted R-squared)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of Different Data Mining Algorithms for Prediction of Body Weight From Several Morphological Measurements in Dogs - S Celik, O Yilmaz.

A new decision tree based algorithm for prediction of hydrogen sulfide solubility in various ionic liquids - Reza Soleimani, Amir Hossein Saeedi Dehaghani, Alireza Bahadori.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 99
inputs <- 0:99

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(100)

# linear regression model
model<-lm(targets~inputs)

# About the model
summary(model)

# Number of Terms
n = length(model$coefficients)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit : adjusted R-squared
gofARSq(targets, predicted, dgt=4, nTermInAppr=n)
```

Goodness of Fit : Akaike's Information Criterion

Akaike's Information Criterion

Description

Calculates and returns Akaike's information criterion (AIC).

Usage

```
gofAIC(Obs, Prd, nTermInAppr=2, dgt = 3)
```

Arguments

`Obs`	Observed or measured values or target vector.
`Prd`	Predicted or fitted values by the model. Values produced by approximation or regression.
`nTermInAppr`	Number of terms in approximation or regression models formula, including interception. For simple linear regression with one independent variable is simply 2. Default is 2.
`dgt`	Number of digits in decimal places. Default is 3.

Value

`AkaikesInformationCriterion`
: Akaike's information criterion (AIC)

Note

When `n/k` is not greater than 40, where `n` is the number of observations and `k` is the number of terms in approximation, Corrected Akaike's Information Criterion (`gofCAIC`) is used.

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan - Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas, Wilhelm Grzesiak.

Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms - Daniel Zaborski, Muhammad Ali, Ecevit Eyduran, Wilhelm Grzesiak, Mohammad Masood Tariq, Ferhat Abbas, Abdul Waheed, Cem Tirink - Pakistan journal of zoology, 2019.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 99
inputs <- 0:99

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(100)

# linear regression model
model<-lm(targets~inputs)

# About the model
summary(model)

# Number of Terms
n = length(model$coefficients)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit : Akaike's information criterion (AIC)
gofAIC(targets, predicted, dgt=4, nTermInAppr=n)
```

`Goodness of Fit : Coefficient of Determination (R-Squared)`
Coefficient of Determination (R-Squared)

Description

Calculates and returns coefficient of determination (R-squared).

Usage

```
gofCoD(Obs, Prd, dgt = 3)
```

Arguments

`Obs`	Observed or measured values or target vector.
`Prd`	Predicted or fitted values by the model. Values produced by approximation or regression.
`dgt`	Number of digits in decimal places. Default is 3.

Value

`CoefficientofDetermination`
Goodness of fit - coefficient of determination (R-squared)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of Different Data Mining Algorithms for Prediction of Body Weight From Several Morphological Measurements in Dogs - S Celik, O Yilmaz.

A new decision tree based algorithm for prediction of hydrogen sulfide solubility in various ionic liquids - Reza Soleimani, Amir Hossein Saeedi Dehaghani, Alireza Bahadori.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 99
inputs <- 0:99

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(100)

# linear regression model
model<-lm(targets~inputs)
```

```
# About the model
summary(model)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit : coefficient of determination (R-squared)
gofCoD(targets, predicted)
```

`Goodness of Fit : Corrected Akaike's Information Criterion`
Corrected Akaike's Information Criterion

Description

Calculates and returns corrected Akaike's information criterion.

Usage

```
gofCAIC(Obs, Prd, nTermInAppr = 2, dgt = 3)
```

Arguments

`Obs`	Observed or measured values or target vector.
`Prd`	Predicted or fitted values by the model. Values produced by approximation or regression.
`nTermInAppr`	Number of terms in approximation or regression models formula, interception included. For simple linear regression with one independent variable is simply 2. Default is 2.
`dgt`	Number of digits in decimal places. Default is 3.

Value

`CorrectedAkaikesInformationCriterion`
Goodness of fit - corrected Akaike's information criterion (cAIC)

Note

When `n/k` is greater than 40, where `n` is the umber of observations and `k` is the number of terms in approximation, Akaike's Information Criterion (`gofAIC`) is used.

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan - Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas, Wilhelm Grzesiak.

Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms - Daniel Zaborski, Muhammad Ali, Ecevit Eyduran, Wilhelm Grzesiak, Mohammad Masood Tariq, Ferhat Abbas, Abdul Waheed, Cem Tirink - Pakistan journal of zoology, 2019.

Examples

```
# dummy inputs, independent variable
# integers fron 0 to 79
inputs <- 0:79

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(80)

# linear regression model
model<-lm(targets~inputs)

# About the model
summary(model)

# Number of Terms
n = length(model$coefficients)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for Goodness of Fit function (GoF)
library(ehaGoF)

# Goodness of Fit : Corrected Akaike's Information Criterion
gofCAIC(targets, predicted, dgt=4, nTermInAppr=n)
```

Goodness of Fit : Global Relative Approximation Error

Global Relative Approximation Error

Description

Calculates and returns global relative approximation error (RAE).

Usage

```
gofRAE(Obs, Prd, dgt = 3)
```

Arguments

Obs	Observed or measured values or target vector.
Prd	Predicted or fitted values by the model. Values produced by approximation or regression.
dgt	Number of digits in decimal places. Default is 3.

Value

RelativeApproximationError
: Global relative approximation error (RAE)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan - Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas, Wilhelm Grzesiak.

Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms - Daniel Zaborski, Muhammad Ali, Ecevit Eyduran, Wilhelm Grzesiak, Mohammad Masood Tariq, Ferhat Abbas, Abdul Waheed, Cem Tirink - Pakistan journal of zoology, 2019.

Comparison of Different Data Mining Algorithms for Prediction of Body Weight From Several Morphological Measurements in Dogs - S Celik, O Yilmaz.

The Connection Dependent Threshold Model for Finite Sources -A Generalization of the Engset Multirate Loss Model - Ioannis D. Moscholios and Michael D. Logothetis.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 99
inputs <- 0:99

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(100)

# linear regression model
model<-lm(targets~inputs)

# About the model
summary(model)

# model's predicted values against targets
predicted<-model$fitted.values
```

```
# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit : global relative approximation error (RAE)
gofRAE(targets, predicted)
```

`Goodness of Fit : Mean Absolute Deviation`
Mean Absolute Deviation

Description

Calculates and returns mean absolute deviation (MAD).

Usage

```
gofMAD(Obs, Prd, dgt = 3)
```

Arguments

`Obs`	Observed or measured values or target vector.
`Prd`	Predicted or fitted values by the model. Values produced by approximation or regression.
`dgt`	Number of digits in decimal places. Default is 3.

Value

`MeanAbsoluteDeviation`
Goodness of fit - mean absolute deviation (MAD)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan - Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas, Wilhelm Grzesiak.

Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms - Daniel Zaborski, Muhammad Ali, Ecevit Eyduran, Wilhelm Grzesiak, Mohammad Masood Tariq, Ferhat Abbas, Abdul Waheed, Cem Tirink - Pakistan journal of zoology, 2019.

Comparison of Different Data Mining Algorithms for Prediction of Body Weight From Several Morphological Measurements in Dogs - S Celik, O Yilmaz.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 99
inputs <- 0:99

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(100)

# linear regression model
model<-lm(targets~inputs)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for Goodness of Fit function (GoF)
library(ehaGoF)

# Goodness of Fit : Mean Absolute Deviation
gofMAD(targets, predicted, dgt=4)
```

Goodness of Fit : Mean Error

Mean Error

Description

Calculates and returns mean error (ME).

Usage

```
gofME(Obs, Prd, dgt = 3)
```

Arguments

Obs	Observed or measured values or target vector.
Prd	Predicted or fitted values by the model. Values produced by approximation or regression.
dgt	Number of digits in decimal places. Default is 3.

Value

MeanError	Goodness of fit - mean error (ME)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of the Predictive Capabilities of Several Data Mining Algorithms and Multiple Linear Regression in the Prediction of Body Weight by Means of Body Measurements in the Indigenous Beetal Goat of Pakistan - Ecevit Eyduran, Daniel Zaborski, Abdul Waheed, Senol Celik, Koksal Karadas, Wilhelm Grzesiak.

Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms - Daniel Zaborski, Muhammad Ali, Ecevit Eyduran, Wilhelm Grzesiak, Mohammad Masood Tariq, Ferhat Abbas, Abdul Waheed, Cem Tirink - Pakistan journal of zoology, 2019.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 19
inputs <- 0:19

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(20)

# linear regression model
model<-lm(targets~inputs)

# About the model
summary(model)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit : mean error (ME)
gofME(targets, predicted)
```

Goodness of Fit : Mean Relative Approximation Error

Mean Relative Approximation Error

Description

Calculates and returns mean relative approximation error (MRAE).

Usage

```
gofMRAE(Obs, Prd, dgt = 3)
```

Arguments

Obs	Observed values or target vector.
Prd	Predicted values. Values produced by approximation or regression.
dgt	Number of digits in decimal places. Default is 3.

Value

MeanRelativeApproximationError
: Goodness of fit - mean relative approximation error (MRAE)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

The Connection Dependent Threshold Model for Finite Sources -A Generalization of the Engset Multirate Loss Model - Ioannis D. Moscholios and Michael D. Logothetis.

Competitive adsorption equilibrium modeling of volatile organic compound (VOC) and water vapor onto activated carbon - Imranul I. Laskara, Zaher Hashishoa,∗, John H. Phillipsb, James E. Andersonc, Mark Nichols.

A new decision tree based algorithm for prediction of hydrogen sulfide solubility in various ionic liquids - Reza Soleimani, Amir Hossein Saeedi Dehaghani, Alireza Bahadori.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 19
inputs <- 0:19

# dummy targets/observed values, dependent variable
# a product of 2 times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(20)

# linear regression model
model<-lm(targets~inputs)

# About the model
summary(model)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit : mean relative approximation error (MRAE)
gofMRAE(targets, predicted)
```

`Goodness of Fit : Performance Index`
Performance Index

Description

Calculates and returns performance index (PI).

Usage

```
gofPI(Obs, Prd, dgt = 3)
```

Arguments

`Obs`	Observed or measured values or target vector.
`Prd`	Predicted or fitted values by the model. Values produced by approximation or regression.
`dgt`	Number of digits in decimal places. Default is 3.

Value

`PerformanceIndex`
Goodness of fit: performance index (PI)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Prediction of Selected Reproductive Traits of Indigenous Harnai Sheep under the Farm Management System via various Data Mining Algorithms - Daniel Zaborski, Muhammad Ali, Ecevit Eyduran, Wilhelm Grzesiak, Mohammad Masood Tariq, Ferhat Abbas, Abdul Waheed, Cem Tirink - Pakistan journal of zoology, 2019

Examples

```
# dummy inputs, independent variable
# integers from 0 to 19
inputs <- 0:19

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(20)

# linear regression model
model<-lm(targets~inputs)
```

```
# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit - performance index (PI)
gofPI(targets, predicted)
```

Goodness of Fit : R-Squared

R-Squared (Coefficient of Determination)

Description

Calculates and returns R-squared (coefficient of determination).

Usage

```
gofRSq(Obs, Prd, dgt = 3)
```

Arguments

Obs	Observed or measured values or target vector.
Prd	Predicted or fitted values by the model. Values produced by approximation or regression.
dgt	Number of digits in decimal places. Default is 3.

Value

RSquared	Goodness of fit - coefficient of determination (R-squared)

Author(s)

Prof. Dr. Ecevit Eyduran, TA. Alper Gulbe

References

Comparison of Different Data Mining Algorithms for Prediction of Body Weight From Several Morphological Measurements in Dogs - S Celik, O Yilmaz.

A new decision tree based algorithm for prediction of hydrogen sulfide solubility in various ionic liquids - Reza Soleimani, Amir Hossein Saeedi Dehaghani, Alireza Bahadori.

Examples

```
# dummy inputs, independent variable
# integers from 0 to 99
inputs <- 0:99

# dummy targets/observed values, dependent variable
# a product of 2*times inputs minus 5 with some normal noise
targets <- -5 + inputs*1.2 + rnorm(100)

# linear regression model
model<-lm(targets~inputs)

# About the model
summary(model)

# model's predicted values against targets
predicted<-model$fitted.values

# using library ehaGoF for goodness of fit.
library(ehaGoF)

# Goodness of fit : coefficient of determination (R-squared)
gofRSq(targets, predicted)
```

Index

GoF *(Goodness of Fit)*, 2
gofACoD *(Goodness of Fit : Adjusted Coefficient of Determination (Adjusted R-Squared))*, 12
gofAIC *(Goodness of Fit : Akaike's Information Criterion)*, 15
gofARSq *(Goodness of Fit : Adjusted R-Squared)*, 14
gofCAIC *(Goodness of Fit : Corrected Akaike's Information Criterion)*, 18
gofCoD *(Goodness of Fit : Coefficient of Determination (R-Squared))*, 17
gofCV *(Goodness of Fit - Coefficient of Variation)*, 4
gofMAD *(Goodness of Fit : Mean Absolute Deviation)*, 21
gofMAPE *(Goodness of Fit - Mean Absolute Percentage Error)*, 5
gofME *(Goodness of Fit : Mean Error)*, 22
gofMRAE *(Goodness of Fit : Mean Relative Approximation Error)*, 23
gofPC *(Goodness of Fit - Pearson's Correlation Coefficients)*, 7
gofPI *(Goodness of Fit : Performance Index)*, 25
gofRAE *(Goodness of Fit : Global Relative Approximation Error)*, 19
gofRMSE *(Goodness of Fit - Root Mean Square Error)*, 10
gofRRMSE *(Goodness of fit - Relative Root Mean Square Error)*, 8
gofRSq *(Goodness of Fit : R-Squared)*, 26
gofSDR *(Goodness of Fit - Standard Deviation Ratio)*, 11
Goodness of Fit, 2
Goodness of Fit - Coefficient of Variation, 4
Goodness of Fit - Mean Absolute Percentage Error, 5
Goodness of Fit - Pearson's Correlation Coefficients, 7
Goodness of fit - Relative Root Mean Square Error, 8
Goodness of Fit - Root Mean Square Error, 10
Goodness of Fit - Standard Deviation Ratio, 11
Goodness of Fit : Adjusted Coefficient of Determination (Adjusted R-Squared), 12
Goodness of Fit : Adjusted R-Squared, 14
Goodness of Fit : Akaike's Information Criterion, 15
Goodness of Fit : Coefficient of Determination (R-Squared), 17
Goodness of Fit : Corrected Akaike's Information Criterion, 18
Goodness of Fit : Global Relative Approximation Error, 19
Goodness of Fit : Mean Absolute Deviation, 21
Goodness of Fit : Mean Error, 22
Goodness of Fit : Mean Relative Approximation Error, 23
Goodness of Fit : Performance Index, 25
Goodness of Fit : R-Squared, 26