Package: featureCorMatrix (via r-universe)

August 25, 2024

Type Package

Title Measurement Level Independent Feature Correlation Matrix

Version 0.4.0

Maintainer Guido Moeser <guido.moeser@masem.de>

Description Uses three different correlation coefficients to calculate measurement-level adequate correlations in a feature matrix: Pearson product-moment correlation coefficient, Intraclass correlation and Cramer's V.

License GPL (>= 2)

Encoding UTF-8

LazyData true

Imports stats

RoxygenNote 7.1.0

NeedsCompilation no

Author Guido Moeser [aut, cre], Ilja Muhl [aut]

Date/Publication 2020-05-27 10:30:02 UTC

Repository https://masem-research.r-universe.dev

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

RemoteRef HEAD

RemoteSha 82472b3d73727c251bd97d333ecef1c61a558740

Contents

cv.test *Calculates Cramer's V Correlation Coefficient*

Description

cv.test returns the Cramer's V correlation coefficient

Usage

```
cv.test(x, y)
```

Arguments

x	a vector (categorical or numerical values)
y	a vector (categorical or numerical values)

Details

The function calculates Cramer's V based on the results of an Chi-Square-Test of Independence between two categorical variables

Value

Cramer's V

Examples

```
cv.test(x = iris$Species, iris$Sepal.Length)
```

featureCorMatrix *Calculates the Feature Correlation Matrix*

Description

featureCorMatrix returns a correlation matrix between all features

Usage

```
featureCorMatrix(dataframe, absoluteValues = FALSE)
```

Arguments

dataframe	A data.frame
absoluteValues	A flag stating if only positive correlations should be returned

Details

The function selects automatically the appropriate correlation coefficient regarding the storage type of both variables - If both variable are numerical ones, the Pearson product-moment correlation coefficient will be chosen - If both variables are categorical, Cramer's V will be used - If one variable is a numerical and the other a categorical one, the Intraclass correlation will be calculated

Value

A correlation matrix

Examples

```
featureCorMatrix(dataframe = iris, absoluteValues = TRUE)
```

`GermanCredit` *Statlog (German Credit Data) Data Set*

Description

This dataset classifies people described by a set of attributes as good or bad credit risks.

The variables are as follows:

- Credit. Target variable
- balance_credit_acc. Status of existing checking account
- duration. Duration in month
- moral. Credit history
- verw. Purpose
- hoehe. Credit amount
- sparkont. Savings account/bonds
- beszeit. Present employment since
- rate. Installment rate in percentage of disposable income
- famges. Personal status and sex
- buerge. Other debtors / guarantors
- wohnzeit. Present residence since
- verm. Property
- alter. Age in years
- weitkred. Other installment plans
- wohn. Housing
- bishkred. Number of existing credits at this bank
- beruf. Job
- pers. Number of people being liable to provide maintenance for
- telef. Telephone
- gastarb. Foreign worker

Usage

```
data(GermanCredit)
```

Format

A data frame with 1000 rows and 21 variables

Source

UCI Repository, https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data)

---

icc *Calculates the Intraclass correlation*

Description

The function calculates the Intraclass correlation based on the results of the ‘aov‘ function

Usage

```
icc(depvar, indvar)
```

Arguments

depvar	dependent variable, must be numeric
indvar	independent variable, must be categorical

Value

returns the Intraclass correlation

Examples

```
icc(depvar = iris$Sepal.Length, indvar = iris$Species)
```

Index