Package 'ImpactEffectsize'

January 20, 2025

Type Package

Title Calculation and Visualization of the Impact Effect Size Measure

Description A non-parametric effect size measure capturing changes in central tendency or shape of data distributions. The package provides the necessary functions to calculate and plot the Impact effect size measure between two groups.

Version 0.7

Date 2024-04-14

URL <https://cran.r-project.org/package=ImpactEffectsize>

Encoding UTF-8

Author Joern Loetsch [aut,cre] (<<https://orcid.org/0000-0002-5818-6958>>),
  Alfred Ultsch [aut] (<<https://orcid.org/0000-0002-7845-3283>>)

Maintainer Jorn Lotsch <j.loetsch@em.uni-frankfurt.de>

LazyData true

Imports caTools, matrixStats, parallelDist, methods, stats, graphics,
  Rcpp

Suggests testthat

Depends R (>= 3.5.0)

License GPL-3

RoxygenNote 6.1.1

LinkingTo Rcpp

NeedsCompilation yes

Repository CRAN

Date/Publication 2024-04-14 16:20:02 UTC

Contents

ImpactEffectsize-package

ImpactEffectsize-package

Description

Calculation and visualization of the Impact effect size measure. A non-parametric effect size measure capturing changes in central tendency or shape of data distributions for feature selection preceding machine-learning. The package provides the necessary functions to calculate and plot the Impact effect size measure between two groups.

References

Lotsch, J., and Ultsch, A. (2019): Impact – An R Package for calculation and visualisation of the Impact distance and data distribution-shape based effect size measure.

BcellLymphomaCD79   *Example data of bimodal CD79 expression.*

Description

Data set of CD79 expression in patients with B lymphoma (class 1) and in controls (class 2).

Usage

```
data("BcellLymphomaCD79")
```

Details

Size 258429 x 2 , Dimensions 1, stored in BcellLymphomaCD79$Data

Classes 2, stored in BcellLymphomaCD79$Classes

Examples

```
data(BcellLymphomaCD79)
str(BcellLymphomaCD79)
```

FeatureselectionData   *Example data with two groups and the Impact effet size measure.*

Description

Dataset with 2 classes and 20 variables that allow class separation at varous degrees of difficulty.

Usage

```
data("FeatureselectionData")
```

Details

Size 2000 x 20 , Dimensions 1, stored in `FeatureselectionData$Var0001,...,FeatureselectionData$Var0020`

Classes 2, stored in `FeatureselectionData$Classes`

Examples

```
data(FeatureselectionData)
str(FeatureselectionData)
```

FlowcytometricData   *Example data of hematologic marker expression.*

Description

Data set of 8 flow cytometry-based lymphoma makers from 1,494 cells from healthy subjects (class 1) and 1,302 cells from lymphoma patients (class 2).

Usage

```
data("FlowcytometricData")
```

Details

Size 2796 x 9 , Dimensions 1, stored in `FlowcytometricData$$[CD3,CD4,CD8,CD11,CD19,CD103,CD200,IgM]`

Classes 2, stored in `FlowcytometricData$Classes`

Examples

```
data(FlowcytometricData)
str(FlowcytometricData)
```

Impact   *Impact effect size measure*

Description

Calculates the Impact effect size measure that is based on the group distance and the difference in the shape of the data distribution between two groups.

Usage

```
Impact(Data, Cls, PlotIt = FALSE, pde = TRUE,
  col = c("red","blue"), meanLines = FALSE, medianLines = FALSE, ...)
```

Arguments

Data	the data of both groups as a vector.
Cls	the class information as a vector of similar length as the data vector.
PlotIt	plots the pdf of the two groups as Pareto density estimation.
col	colors of the two lines to plot. Will be ignored if PlotIt = FALSE.
pde	plots a standard pdf insted of a non-parametric Pareto density estimation. Will be ignored if PlotIt = FALSE.
meanLines	adds perpendicular lines at the means of the two groups. Will be ignored if PlotIt = FALSE.
medianLines	adds perpendicular lines at the medians of the two groups. Will be ignored if PlotIt = FALSE.
...	more graphical parameters can be given as plot arguments if PlotIt = TRUE.

Value

Returns a list of value incuding the main result and intermediate results.

Impact	the main effect size measure.
MorphDiff	the extend of the group-difference in the shapes of the pdf.
CTDiff	the extend of the difference in the group medians.

Author(s)

Jorn Lotsch and Alfred Ultsch

References

Lotsch, J., and Ultsch, A. (2019): ImpactEffectsize – an R Package for calculation and visualisation of the Impact distance and shape based effect size measure.

Examples

```
## example 1
data("FeatureselectionData")
ImpactSize <- Impact(Data = FeatureselectionData$Var0011,
  Cls = FeatureselectionData$Classes, PlotIt = TRUE)

## example 2
data("BcellLymphomaCD79")
data("FeatureselectionData")
data("FlowcytometricData")
data("SameMeansData")
data("StocksFluctuation")
```

SameMeansData   *Example artificial data with two groups of same means but different data distribution shapes.*

Description

Dataset with 2 classes six variables were both classes have the same means but different shapes of the distribution.

Usage

```
data("SameMeansData")
```

Details

Size 2000 x 7 , Dimensions 1, stored in SameMeansData$NOchangeInMandS,...,SameMeansData$NegChi2andGauss

Classes 2, stored in SameMeansData$Classes

Examples

```
data(SameMeansData)
str(SameMeansData)
```

StocksFluctuation   *Example data of stock fluctuation.*

Description

Data set of Log ratios of daily changes of n =5,522 for 10 German stocks with low fluctuation (class 1) or high fluctuation (class 2).

Usage

```
data("StocksFluctuation")
```

Details

Size 5522 x 2 , Dimensions 1, stored in `StocksFluctuation$logFluctuation`

Classes 2, stored in `StocksFluctuation$Classes`

Examples

```
data(StocksFluctuation)
str(StocksFluctuation)
```

Index