Package 'tradeoffaucdim'

May 2, 2025

Type Package

Title Plotting Trade-Off AUC-Dimensionality

Version 0.1.0

Depends SuperLearner, R (>= 3.5)

Description Perform and Runtime statistical comparisons between models. This package aims at choosing the best model for a particular dataset, regarding its discriminant power and runtime.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Suggests spelling, testthat (>= 3.0.0)

Config/testthat/edition 3

RoxygenNote 7.3.2

Imports dplyr, speedglm, magrittr, purrr, rsample, stringr, tibble, tidyr, ROCR, caret, ez, fastDummies, fuzzySim, ggplot2

URL https://github.com/luisgarcez11/tradeoffaucdim

BugReports https://github.com/luisgarcez11/tradeoffaucdim/issues

Language en-US

NeedsCompilation no

Author Garcez Luis [aut, cre]

Maintainer Garcez Luis <luisgarcez1@gmail.com>

Repository CRAN

Date/Publication 2025-05-02 09:40:02 UTC

Contents

`apply_model` *Apply Model*

Description

Apply model and create column with fit

Usage

```
apply_model(
  obj,
  models = c("SL.glm", "SL.rpart"),
  test_partition_prop = 0.2,
  perf_measure = "auc"
)
```

Arguments

`obj`	object returned from `define_indepvars_outcome`
`models`	models to be analyzed
`test_partition_prop`	test proportion
`perf_measure`	performance measure

Value

list with fit models and parameters

Examples

```
apply_model(obj2)
```

bananaquality *Banana Quality*

Description

Banana quality dataset

Usage

```
bananaquality
```

Format

An object of class `data.frame` with 8000 rows and 8 columns.

bananaquality_sample *Banana Quality Subset*

Description

Banana quality dataset subset

Usage

```
bananaquality_sample
```

Format

An object of class `data.frame` with 50 rows and 8 columns.

bootstrap_data *Bootstrap data*

Description

Create a list with bootstrap samples

Usage

```
bootstrap_data(
  data,
  outcome = "Quality",
  indep_vars = c("Size", "Weight", "Sweetness", "Softness", "HarvestTime", "Ripeness",
    "Acidity"),
  n_samples = 50,
  n_maximum_dim = 5
)
```

Arguments

data	a dataframe to be analyzed
outcome	a string representing the outcome variable
indep_vars	a vector of strings to be considered
n_samples	number of bootstrap samples
n_maximum_dim	maximum number of variables to be considered

Value

list

Examples

```
bootstrap_data(bananaquality_sample)
```

---

compare_test *Compare test*

Description

Performs statistical tests to compare performance and runtime.

Usage

```
compare_test(obj, x_label_offset = 1, y_label_offset = 10)
```

Arguments

obj	object returned by plot_curve
x_label_offset	x coordinate to plot p-value
y_label_offset	y coordinate to plot p-value

Value

list with statistical tests performed

Examples

```
compare_test(obj5)
```

define_indepvars *Define independent variables*

Description

Define independent variables to be tested

Usage

```
define_indepvars(obj, p_in = 0.5, p_out = 0.6)
```

Arguments

obj	object returned by bootstrap_data
p_in	entry p-value used to determine variable order
p_out	removal p-value used to determine variable order

Value

list

Examples

```
define_indepvars(obj1)
```

obj1 *Example Object returned from* bootstrap_data

Description

obj1

Usage

```
obj1
```

Format

An object of class list of length 5.

obj2 *Example Object returned from* define_indepvars_outcome

Description

obj2

Usage

```
obj2
```

Format

An object of class list of length 7.

obj3 *Example Object returned from* apply_model

Description

obj3

Usage

```
obj3
```

Format

An object of class list of length 10.

obj4 *Example Object returned from* summary_statistics

Description

obj4

Usage

```
obj4
```

Format

An object of class list of length 11.

obj5 *Example Object returned from* plot_curve

Description

obj5

Usage

```
obj5
```

Format

An object of class list of length 15.

obj6 *Example Object returned from* compare_test

Description

obj6

Usage

```
obj6
```

Format

An object of class list of length 16.

plot_curve *Plot curve*

Description

Return plot features.

Usage

```
plot_curve(obj)
```

Arguments

obj	object returned by summary_statistics

Value

list with graphical features

Examples

```
plot_curve(obj4)
```

`summary_stats` *Summary Stats*

Description

Return summary statistics

Usage

```
summary_stats(obj)
```

Arguments

`obj`	object returned from `apply_model`

Value

list with summary statistics and bootstrap confidence intervals

Examples

```
summary_stats(obj3)
```

`wrapper_aucdim` *Wrap all pipeline*

Description

Wrap all pipeline

Usage

```
wrapper_aucdim(
  data,
  outcome,
  indep_vars,
  n_samples = 100,
  n_maximum_dim = 5,
  p_in = 0.5,
  p_out = 0.6,
  models = c("SL.glm"),
  test_partition_prop = 0.2,
  perf_measure = "auc",
  x_label_offset = 1,
  y_label_offset = 10
)
```

Arguments

`data`	a dataframe to be analyzed
`outcome`	a string representing the outcome variable
`indep_vars`	a vector of strings to be considered
`n_samples`	number of bootstrap samples
`n_maximum_dim`	maximum number of variables
`p_in`	entry p-value for choosing variable order
`p_out`	exclusion p-value for choosing variable order
`models`	a string representing the models to compare
`test_partition_prop`	test partition proportion
`perf_measure`	performance measure to be considered
`x_label_offset`	x coordinate for plotting
`y_label_offset`	y coordinate for plotting

Value

a list with the final object

Index