

Package ‘mccf1’

August 13, 2018

Title Creates the MCC-F1 Curve and Calculates the MCC-F1 Metric and the Best Threshold

Version 1.0

Date 2018-07-31

Maintainer Chang Cao <kirin.cao@mail.utoronto.ca>

Depends R (>= 3.3.3)

Imports ROCR, ggplot2

Description The MCC-F1 analysis is a method to evaluate the performance of binary classifications. The MCC-F1 curve is more reliable than the Receiver Operating Characteristic (ROC) curve and the Precision-Recall (PR) curve under imbalanced ground truth. The MCC-F1 analysis also provides the MCC-F1 metric that integrates classifier performance over varying thresholds, and the best threshold of binary classification.

License GPL (>= 2)

URL <https://bitbucket.org/hoffmanlab/mccf1/>

BugReports <https://stackoverflow.com/questions/tagged/mccf1>

RoxygenNote 6.0.1

NeedsCompilation no

Author Chang Cao [aut, cre],
Michael Hoffman [aut],
Davide Chicco [aut]

Repository CRAN

Date/Publication 2018-08-13 14:40:06 UTC

R topics documented:

mccf1	2
plot.mccf1	2
summary.mccf1	3

Index	5
--------------	----------

mccf1	<i>MCC (Matthews correlation coefficient) - F1</i>
-------	--

Description

This function creates the "mccf1" object.

Usage

```
mccf1(response, predictor)
```

Arguments

response a vector of actual values (0 or 1) in binary classification.
 predictor a vector of predicted values (between 0 and 1) in binary classification.

Value

the "mccf1" object, which is basically a list containing a vector of thresholds, a vector of corresponding normalized MCC, and a vector of corresponding F1.

Examples

```
response <- c(rep(1, 1000), rep(0, 10000))
set.seed(2017)
predictor <- c(rbeta(300, 12, 2), rbeta(700, 3, 4), rbeta(10000, 2, 3))
x <- mccf1(response, predictor)
head(x$normalizedMCC)
# [1] NaN 0.5150763 0.5213220 0.5261152 0.5301566 0.5337177
head(x$F1)
# [1] NaN 0.001998002 0.003992016 0.005982054 0.007968127 0.009950249
head(x$thresholds)
# [1] Inf 0.9935354 0.9931493 0.9930786 0.9925507 0.9900520
```

plot.mccf1	<i>Plot the MCC-F1 curve</i>
------------	------------------------------

Description

This function plots the MCC-F1 score curve using ggplot.

Usage

```
## S3 method for class 'mccf1'
plot(x, main = "the MCC-F1 score curve", xlab = "F1 score",
     ylab = "normalized MCC", .curveFileName, ...)
```

Arguments

x	"mccf1" object resulting from the function mccf1
main	main title of the plot (default: "the MCC-F1 score curve")
xlab, ylab	x- and y- axis annotation (default: "F1 score", "normalized MCC")
.curveFileName	character string, used for naming pdf file of the plot with pdf (if this argument is present, the plot will be stored in pdf with specified file name)
...	further arguments passed to and from methods

Value

the ggplot object

Examples

```
response <- c(rep(1, 1000), rep(0, 10000))
predictor <- c(rbeta(300, 12, 2), rbeta(700, 3, 4), rbeta(10000, 2, 3))
plot(mccf1(response, predictor))
```

summary.mccf1	<i>Summarize the the performance of a binary classification using MCC-F1 metric and the best threshold</i>
---------------	--

Description

This function calculates the MCC-F1 metric and the best threshold for a binary classification.

Usage

```
## S3 method for class 'mccf1'
summary(object, digits, bins = 100, ...)
```

Arguments

object	"mccf1" object resulting from the function mccf1
digits	an integer, used for number formatting with signif
bins	an integer, will be used to divide the range of normalized MCC when calculating the MCC-F1 metric (default = 100)
...	further arguments passed to and from methods

Value

a data.frame that shows the MCC-F1 metric (between 0 and 1) and the best threshold (between 0 and 1)

Examples

```
response <- c(rep(1, 1000), rep(0, 10000))
set.seed(2017)
predictor <- c(rbeta(300, 12, 2), rbeta(700, 3, 4), rbeta(10000, 2, 3))
## Not run: summary(mccf1(response, predictor))
# mccf1_metric best_threshold
# 0.3508904 0.786905
summary(mccf1(response, predictor), bins = 50)
# mccf1_metric best_threshold
# 0.3432971 0.786905
## Not run: summary(mccf1(response, predictor), digits = 3)
# mccf1_metric best_threshold
# 0.351 0.787
```

Index

`mccf1`, [2](#)

`pdf`, [3](#)

`plot.mccf1`, [2](#)

`signif`, [3](#)

`summary.mccf1`, [3](#)