Package 'RobustMetrics'

September 2, 2025

Type Package

Title Calculates Robust Performance Metrics for Imbalanced Classification Problems

Version 0.1.1

Maintainer Bernhard Klar <bernhard.klar@kit.edu>

Description Calculates robust Matthews Correlation Coefficient (MCC) and robust F-Beta Scores, as introduced by Holzmann and Klar (2024) <doi:10.48550/arXiv.2404.07661>. These performance metrics are designed for imbalanced classification problems. Plots the receiver operating characteristic curve (ROC curve) together with the recall / 1-precision curve.

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

URL https://github.com/BernhardKlar/RobustMetrics

BugReports https://github.com/BernhardKlar/RobustMetrics/issues

NeedsCompilation no

Author Bernhard Klar [aut, cre], Hajo Holzmann [aut]

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2025-09-02 05:30:07 UTC

Contents

`FScore` *F-Beta Score*

Description

Compute the F-Beta Score.

Usage

```
FScore(
  actual = NULL,
  predicted = NULL,
  TP = NULL,
  FN = NULL,
  FP = NULL,
  TN = NULL,
  beta = 1
)
```

Arguments

`actual`	A vector of actual values (1/0 or TRUE/FALSE)
`predicted`	A vector of prediction values (1/0 or TRUE/FALSE)
`TP`	Count of true positives (correctly predicted 1/TRUE)
`FN`	Count of false negatives (predicted 0/FALSE, but actually 1/TRUE)
`FP`	Count of false positives (predicted 1/TRUE, but actually 0/FALSE)
`TN`	Count of true negatives (correctly predicted 0/FALSE)
`beta`	Beta squared is the weight of recall in harmonic mean

Details

Calculate the F-Beta Score. Provide either:

- `actual` and `predicted` or
- `TP`, `FN`, `FP` and `TN`.

Value

F-Beta Score.

References

Holzmann, H., Klar, B. (2024). Robust performance metrics for imbalanced classification problems. arXiv:2404.07661. LINK

Examples

```
actual <-    c(1,1,1,1,1,1,0,0,0,0)
predicted <- c(1,1,1,1,0,0,1,0,0,0)
FScore(actual, predicted)
FScore(TP=4, FN=2, FP=1, TN=3)
```

MCC *Matthews correlation coefficient*

Description

Compute Matthews correlation coefficient.

Usage

```
MCC(
  actual = NULL,
  predicted = NULL,
  TP = NULL,
  FN = NULL,
  FP = NULL,
  TN = NULL
)
```

Arguments

actual	A vector of actual values (1/0 or TRUE/FALSE)
predicted	A vector of prediction values (1/0 or TRUE/FALSE)
TP	Count of true positives (correctly predicted 1/TRUE)
FN	Count of false negatives (predicted 0/FALSE, but actually 1/TRUE)
FP	Count of false positives (predicted 1/TRUE, but actually 0/FALSE)
TN	Count of true negatives (correctly predicted 0/FALSE)

Details

Calculate Matthews correlation coefficient. Provide either:

- actual and predicted or
- TP, FN, FP and TN.

Value

Matthews correlation coefficient.

References

Holzmann, H., Klar, B. (2024). Robust performance metrics for imbalanced classification problems. arXiv:2404.07661. LINK

Examples

```
actual <-    c(1,1,1,1,1,1,0,0,0,0)
predicted <- c(1,1,1,1,0,0,1,0,0,0)
MCC(actual, predicted)
MCC(TP=4, FN=2, FP=1, TN=3)
```

---

rf.data *Example Random Forest Data*

Description

This dataset contains example data from a Random Forest model.

Usage

```
rf.data
```

Format

A data frame with 2 columns:

actual Actual values

predicted Predicted probabilities

Source

Full test data set using random forest classifier, see Section 6 in Reference.

References

Holzmann, H., Klar, B. (2024). Robust performance metrics for imbalanced classification problems. arXiv:2404.07661. LINK

`robFScore` *Robust F-Beta Score*

Description

Compute a robust version of the F-Beta Score.

Usage

```
robFScore(
  actual = NULL,
  predicted = NULL,
  TP = NULL,
  FN = NULL,
  FP = NULL,
  TN = NULL,
  beta = 1,
  d0 = 0.1
)
```

Arguments

`actual`	A vector of actual values (1/0 or TRUE/FALSE)
`predicted`	A vector of prediction values (1/0 or TRUE/FALSE)
`TP`	Count of true positives (correctly predicted 1/TRUE)
`FN`	Count of false negatives (predicted 0/FALSE, but actually 1/TRUE)
`FP`	Count of false positives (predicted 1/TRUE, but actually 0/FALSE)
`TN`	Count of true negatives (correctly predicted 0/FALSE)
`beta`	Beta squared is the weight of recall in the harmonic mean
`d0`	Weight of the estimated true positive probability in the harmonic mean

Details

Calculate the robust F-Beta Score F_{β,d_0} with two parameters. Provide either:

- `actual` and `predicted` or
- `TP`, `FN`, `FP` and `TN`.

If $d_0 = 0$, the robust F-Beta Score coincides with the F-Beta Score.

Value

robust F-Beta Score.

References

Holzmann, H., Klar, B. (2024). Robust performance metrics for imbalanced classification problems. arXiv:2404.07661. LINK

Examples

```
actual <-    c(1,1,1,1,1,1,0,0,0,0)
predicted <- c(1,1,1,1,0,0,1,0,0,0)
robFScore(actual, predicted, beta=1, d0=0.1)
robFScore(TP=4, FN=2, FP=1, TN=3, beta=1, d0=1)
```

---

robFScore2 *General robust F-Beta Score*

Description

Compute a robust version of the F-Beta Score with two additional parameters.

Usage

```
robFScore2(
  actual = NULL,
  predicted = NULL,
  TP = NULL,
  FN = NULL,
  FP = NULL,
  TN = NULL,
  d1 = 1,
  d0 = 0.1,
  c = 1
)
```

Arguments

actual	A vector of actual values (1/0 or TRUE/FALSE)
predicted	A vector of prediction values (1/0 or TRUE/FALSE)
TP	Count of true positives (correctly predicted 1/TRUE)
FN	Count of false negatives (predicted 0/FALSE, but actually 1/TRUE)
FP	Count of false positives (predicted 1/TRUE, but actually 0/FALSE)
TN	Count of true negatives (correctly predicted 0/FALSE)
d1	Weight of recall in the harmonic mean (corresponds to beta squared)
d0	Weight of the estimated true positive probability in the harmonic mean
c	Additional parameter in numerator

Details

Calculate the robust F-Beta Score F_{rb} with two additional parameters. Provide either:

- `actual` and `predicted` or
- `TP`, `FN`, `FP` and `TN`.

If $d_1 = \beta^2, d_0 = c = 0$, the robust F-Beta Score coincides with the F-Beta Score.

Value

robust F-Beta Score with two additional parameters.

References

Holzmann, H., Klar, B. (2024). Robust performance metrics for imbalanced classification problems. arXiv:2404.07661. LINK

Examples

```
actual <-    c(1,1,1,1,1,1,0,0,0,0)
predicted <- c(1,1,1,1,0,0,1,0,0,0)
robFScore2(actual, predicted, d0 = 0.1, c = 0.1)
robFScore2(TP=4, FN=2, FP=1, TN=3, d0 = 0.1, c = 1)
```

`robMCC` *Robust Matthews correlation coefficient*

Description

Compute a robust version of Matthews correlation coefficient (MCC).

Usage

```
robMCC(
  actual = NULL,
  predicted = NULL,
  TP = NULL,
  FN = NULL,
  FP = NULL,
  TN = NULL,
  d = 0.1
)
```

Arguments

actual	A vector of actual values (1/0 or TRUE/FALSE)
predicted	A vector of prediction values (1/0 or TRUE/FALSE)
TP	Count of true positives (correctly predicted 1/TRUE)
FN	Count of false negatives (predicted 0/FALSE, but actually 1/TRUE)
FP	Count of false positives (predicted 1/TRUE, but actually 0/FALSE)
TN	Count of true negatives (correctly predicted 0/FALSE)
d	Parameter of the robust MCC

Details

Calculate the robust MCC. Provide either:

- actual and predicted or
- TP, FN, FP and TN.

If $d = 0$, the robust MCC coincides with the MCC.

Value

robust MCC.

References

Holzmann, H., Klar, B. (2024). Robust performance metrics for imbalanced classification problems. arXiv:2404.07661. LINK

Examples

```
actual <- c(1,1,1,1,1,1,0,0,0,0)
predicted <- c(1,1,1,1,0,0,1,0,0,0)
robMCC(actual, predicted, d=0.05)
robMCC(TP=4, FN=2, FP=1, TN=3, d=0.05)
```

ROC_curve *ROC curve*

Description

Plot ROC curve together with recall / 1-precision curve.

Usage

```
ROC_curve(actual, predicted, d = c(0.01, 0.05, 0.1, 0.5))
```

Arguments

actual	A vector of actual values (1/0 or TRUE/FALSE)
predicted	A vector of predicted probabilities (numeric values in $[0, 1]$)
d	A vector of length 4

Details

Instead of a precision-recall curve, a recall / 1-precision curve is plotted in the same coordinate system as the ROC curve.

Grey circles show the corresponding MCC optimal points; black symbols show points optimal with respect to the robust MCC for different values of d.

Value

ROC curve.

References

Holzmann, H., Klar, B. (2024). Robust performance metrics for imbalanced classification problems. arXiv:2404.07661. LINK

Examples

```
actual <- rf.data[, 1]
predicted <- rf.data[, 2]
ROC_curve(actual, predicted, d=c(0.01,0.02,0.1,0.5))
```

Index