Package ‘ggscidca’

March 21, 2024

Type Package
Title Plotting Decision Curve Analysis with Coloured Bars
Version 0.2.0
Maintainer Qiang Liu <dege857@163.com>
Description Decision curve analysis is a method for evaluating and comparing prediction models that incorporates clinical consequences, requires only the data set on which the models are tested, and can be applied to models that have either continuous or dichotomous results. The 'ggscidca' package adds coloured bars of discriminant relevance to the traditional decision curve. Improved practicality and aesthetics. This method was described by Balachandran VP (2015) <doi:10.1016/S1470-2045(14)71116-7>.
License GPL-3
Encoding UTF-8
LazyData true
Imports cmprsk, ggplot2, randomForest, reshape2, survival
Depends R (>= 2.10)
RoxygenNote 7.2.1
NeedsCompilation no
Author Qiang Liu [aut, cre]
Repository CRAN
Date/Publication 2024-03-21 03:40:07 UTC

R topics documented:

Breastcancer .................................................. 2
cmprskstdca ................................................. 2
dca .......................................................... 3
df_surv ....................................................... 4
LIRI .......................................................... 5
newcrr ....................................................... 5
scidca ....................................................... 6
scidca.coxph .............................................. 9
Breastcancer  

*Description*  

A survival data on breast cancer.

*Usage*  

data(Breastcancer)  

*Format*  

An object of class `data.frame` with 660 rows and 12 columns.

*Examples*  

data(Breastcancer)  

---  

*Description*  

Generate data for plotting survival analysis decision curves.

*Arguments*  

data  
a data frame containing the variables in the model.

outcome  
the outcome, response variable. Must be a variable contained within the data frame specified in `data=.`.

predictors  
the predictor variable(s). Must be a variable(s) contained within the data frame specified in `data=.`.

probability  
specifies whether or not each of the independent variables are probabilities. The default is `TRUE`.

xstart  
starting value for x-axis (threshold probability) between 0 and 1. The default is 0.01.
xstop stopping value for x-axis (threshold probability) between 0 and 1. The default is 0.99.
xby increment for threshold probability. The default is 0.01.
ymin minimum bound for graph.
harm specifies the harm(s) associated with the independent variable(s). The default is none.
graph specifies whether or not to display graph of net benefits. The default is TRUE.
intervention specifies whether or not to display graph of net benefits. The default is TRUE.
terventionper number of net reduction in interventions per integer. The default is 100
loess.span specifies the degree of smoothing. The default is 0.10.
timepoint specifies the time point at which the decision curve analysis is performed.
cmprsk if evaluating outcome in presence of a competing risk. The default is FALSE
smooth specifies whether or not to smooth net benefit curve. The default is FALSE.
ttoutcome Enter the time variable in your data.

Details
This function was created and written by Dr Andrew Vickers to generate decision curve data.

Value
Returns a data for plotting a decision curve.

Arguments
data a data frame containing the variables in the model.
outcome the outcome, response variable. Must be a variable contained within the data frame specified in data=.
predictors the predictor variable(s). Must be a variable(s) contained within the data frame specified in data=.
probability specifies whether or not each of the independent variables are probabilities. The default is TRUE.
xstart starting value for x-axis (threshold probability) between 0 and 1. The default is 0.01.
df_surv

xstop  stopping value for x-axis (threshold probability) between 0 and 1. The default is 0.99.
xby    increment for threshold probability. The default is 0.01.
ymin   minimum bound for graph.
harm   specifies the harm(s) associated with the independent variable(s). The default is none.
graph  specifies whether or not to display graph of net benefits. The default is TRUE.
intervention plot net reduction in interventions
interventionper number of net reduction in interventions per integer. The default is 100
loess.span specifies the degree of smoothing. The default is 0.10.
smooth  specifies whether or not to smooth net benefit curve. The default is FALSE.

Details

This function was created and written by Dr Andrew Vickers to generate decision curve data.

Value

Returns a data for plotting a decision curve.

<table>
<thead>
<tr>
<th>df_surv</th>
<th>A data for competitive risk modelling.</th>
</tr>
</thead>
</table>

Description

A data for competitive risk modelling.

Usage

data(df_surv)

Format

An object of class tbl_df (inherits from tbl, data.frame) with 750 rows and 9 columns.

Examples

data(df_surv)
**LIRI**

*A data for random forest analysis.*

---

**Description**

A data for random forest analysis.

**Usage**

```r
data(LIRI)
```

**Format**

An object of class `data.frame` with 232 rows and 6 columns.

**Examples**

```r
data(LIRI)
```

---

**newcrr**

**Description**

Types of transformation of survival analysis models into competitive risk models.

**Usage**

```r
newcrr(fit, cencode = 0, failcode = 1)
```

**Arguments**

- `fit`: Modelling for Survival Analysis.
- `cencode`: Censor status, default is 0.
- `failcode`: Events of interest, default is 1.

**Value**

A list of competing risk model formats.
Description

You can use it to generate a decision curve with coloured bars.

Usage

```r
scidca(
  fit,
  newdata = NULL,
  timepoint = "median",
  cmprsk = FALSE,
  modelnames = NULL,
  merge = FALSE,
  y.min = NULL,
  xstop = NULL,
  y.max = NULL,
  pyh = NULL,
  relcol = "#c01e35",
  irrelcol = "#0151a2",
  relabel = "Nomogram relevant",
  irrellabel = "Nomogram irrelevant",
  text.size = 4.5,
  text.col = "green",
  colbar = TRUE,
  threshold.text = FALSE,
  threshold.line = FALSE,
  nudge_x = 0,
  nudge_y = 0,
  threshold.linetype = 2,
  threshold.linewidth = 1.2,
  threshold.linecol = "black",
  po.text.size = 4,
  po.text.col = "black",
  po.text.fill = "white",
  liftpoc = NULL,
  rightpec = NULL
)
```

Arguments

- `fit` Fill in the model you want to analyze. Support survival analysis and logistic regression.

- `newdata` If the decision curve of the validation set is to be analysed. Fill in the validation set data here.
timepoint  
If it is a survival analysis, fill in the point in time you need to study. The default is the median time.

cmprsk  
If it is a competitive risk model, select TRUE here.

modelnames  
Defines the name of the generated image model.

merge  
If true is selected it will merge the two long zones.

y.min  
The maximum value of the negative part of the picture. Generally defaults to positive values multiplied by 0.4.

xstop  
The maximum value of the X-axis of the picture.

y.max  
The maximum value of the Y-axis. The default value is the maximum net benefit.

pyh  
The height at which the bars are plotted cannot exceed y.min.

relcol  
The colour of the relevant part of the bar. The default is red.

irrelcol  
The colour of the irrelevant part of the bar. The default is blue.

relabel  
Relevance Tags.

irrellabel  
No relevant tags.

text.size  
Font size.

text.col  
The colour of the font.

colbar  
The default is true, and if false is selected, bar plotting is cancelled.

threshold.text  
The default is FALSE, if TRUE is selected, a text message for the threshold will be added.

threshold.line  
The default is FALSE, and if TRUE is selected, lines for the threshold will be added.

nudge.x  
Used to adjust the x-axis position of the point where the threshold is located.

nudge.y  
Used to adjust the y-axis position of the point where the threshold is located.

threshold.linetype  
The line shape of the threshold line.

threshold.linewidth  
The line width of the threshold line.

threshold.linecol  
The colour of the threshold line.

po.text.size  
The size of the threshold point text.

po.text.col  
The colour of the threshold point text.

po.text.fill  
The background of the threshold point text.

liftpec  
Threshold point left displacement.

rightpec  
Threshold point right displacement.

Details
Table 1 represents the relationship between the baseline values of the data. This function can be easily done. Creates ‘Table 1’, i.e., description of baseline patient characteristics, which is essential in every medical research. Supports both continuous and categorical variables, as well as p-values and standardized mean differences.
Value

A picture.

Examples

library(survival)
library(reshape2)
library(ggplot2)

##Import the internal data of the R package
bc<-Breastcancer
##Categorical variables converted to factors
bc$histgrad<-as.factor(bc$histgrad)
bc$er<-as.factor(bc$er)
bc$pr<-as.factor(bc$pr)
bc$ln_yesno<-as.factor(bc$ln_yesno)

##Generate Survival Analysis Model
f1<-coxph(Surv(time,status)~er+histgrad+pr+age+ln_yesno,bc)

##Draw decision curve
scidca(f1)
sclidca(f1,threshold.line = TRUE,threshold.text = TRUE)

##logistic regression model
fit<-glm(status~er+histgrad+pr+age+ln_yesno,family = binomial(link = "logit"),data=bc)

##Draw decision curve
scidca(f1)
sclidca(f1,threshold.line = TRUE,threshold.text = TRUE)

##random forest model
library(randomForest)
LIRI<-LIRI
set.seed(1)
index <- sample(2,nrow(LIRI),replace = TRUE,prob=c(0.7,0.3))
traindata <- LIRI[index==1,]
testdata <- LIRI[index==2,]
traindata$status<-as.factor(traindata$status)
#Modelling random forests
fit<-randomForest(status ~ANLN+CENPA+GPR182+BCO2 ,data=traindata,ntree=500,
important=TRUE,proximity=TRUE)
scidca(fit,newdata = traindata)
sclidca(fit,newdata = testdata)
sclidca(fit,newdata = testdata ,threshold.line = TRUE,threshold.text = TRUE)

#Competitive risk modelling
library("cmprsk")

#Import data
df_surv<-df_surv
#Convert the ending variable to the form 0,1,2.
df_surv$cancer_cr<-ifelse(df_surv$cancer_cr=="diagnosed with cancer",1,
ifelse(df_surv$cancer_cr=="dead other causes",2,0))
#Establishment of a survival analysis model
cox_model <- coxph(Surv(ttcancer, cancer_cr==1) ~ age + famhistory + marker, data = df_surv)
#Conversion to competitive risk modelling format
cox_model1<-newcrr(cox_model)
#Plotting decision curves
scidca.coxph

scidca(cox_model1, timepoint=1.5, newdata = df_surv)

Description

scidca.coxph

Usage

## S3 method for class 'cox'
scidca(
  fit,
  newdata = NULL,
  timepoint = "median",
  cmprsk = FALSE,
  modelnames = NULL,
  merge = FALSE,
  y.min = NULL,
  xstop = NULL,
  y.max = NULL,
  pyh = NULL,
  relcol = "#c01e35",
  irrecol = "#0151a2",
  relabel = "Nomogram relevant",
  irrelabel = "Nomogram irrelevant",
  text.size = 4.5,
  text.col = "green",
  colbar = TRUE,
  threshold.text = FALSE,
  threshold.line = FALSE,
  nudge_x = 0,
  nudge_y = 0,
  threshold.linetype = 2,
  threshold.linewidth = 1.2,
  threshold.linecol = "black",
  po.text.size = 4,
  po.text.col = "black",
  po.text.fill = "white",
  liftpec = NULL,
  rightpec = NULL
)

Arguments

fit Fill in the model you want to analyze. Support survival analysis and logistic regression.
newdata  If the decision curve of the validation set is to be analysed. Fill in the validation set data here.

timepoint  If it is a survival analysis, fill in the point in time you need to study. The default is the median time.

cmprsk  If it is a competitive risk model, select TRUE here.

modelnames  Defines the name of the generated image model.

merge  If true is selected it will merge the two long zones.

y.min  The maximum value of the negative part of the picture. Generally defaults to positive values multiplied by 0.4.

xstop  The maximum value of the X-axis of the picture.

y.max  The maximum value of the Y-axis. The default value is the maximum net benefit.

pyh  The height at which the bars are plotted cannot exceed y.min.

relcol  The colour of the relevant part of the bar. The default is red.

irrelcol  The colour of the irrelevant part of the bar. The default is blue.

relabel  Relevance Tags.

irrelabel  No relevant tags.

text.size  Font size.

text.col  The colour of the font.

colbar  The default is true, and if false is selected, bar plotting is cancelled.

threshold.text  The default is FALSE, if TRUE is selected, a text message for the threshold will be added.

threshold.line  The default is FALSE, and if TRUE is selected, lines for the threshold will be added.

nudge_x  Used to adjust the x-axis position of the point where the threshold is located.

nudge_y  Used to adjust the y-axis position of the point where the threshold is located.

threshold.linetype  The line shape of the threshold line.

threshold.linewidth  The line width of the threshold line.

threshold.linecol  The colour of the threshold line.

po.text.size  The size of the threshold point text.

po.text.col  The colour of the threshold point text.

po.text.fill  The background of the threshold point text.

liftpec  Threshold point left displacement.

righptec  Threshold point right displacement.

Value

A picture.
Description
scidca.crr

Usage
### S3 method for class 'crr'
scidca(
  fit,
  newdata = NULL,
  timepoint = "median",
  cmprsk = FALSE,
  modelnames = NULL,
  merge = FALSE,
  y.min = NULL,
  xstop = NULL,
  y.max = NULL,
  pyh = NULL,
  relcol = "#c01e35",
  irrelcol = "#0151a2",
  relabel = "Nomogram relevant",
  irrellabel = "Nomogram irrelevant",
  text.size = 4.5,
  text.col = "green",
  colbar = TRUE,
  threshold.text = FALSE,
  threshold.line = FALSE,
  nudge_x = 0,
  nudge_y = 0,
  threshold.linetype = 2,
  threshold.linewidth = 1.2,
  threshold.linecol = "black",
  po.text.size = 4,
  po.text.col = "black",
  po.text.fill = "white",
  liftpec = NULL,
  rightpec = NULL
)

Arguments

| fit | Fill in the model you want to analyze. Support survival analysis and logistic regression. |
newdata If the decision curve of the validation set is to be analysed. Fill in the validation set data here.
timepoint If it is a survival analysis, fill in the point in time you need to study. The default is the median time.
cmprsk If it is a competitive risk model, select TRUE here.
modelnames Defines the name of the generated image model.
merge If true is selected it will merge the two long zones.
y.min The maximum value of the negative part of the picture. Generally defaults to positive values multiplied by 0.4.
xstop The maximum value of the X-axis of the picture.
y.max The maximum value of the Y-axis. The default value is the maximum net benefit.
pyth The height at which the bars are plotted cannot exceed y.min.
relcol The colour of the relevant part of the bar. The default is red.
irrelcol The colour of the irrelevant part of the bar. The default is blue.
relabel Relevance Tags.
irrellabel No relevant tags.
text.size Font size.
text.col The colour of the font.
colbar The default is true, and if false is selected, bar plotting is cancelled.
threshold.text The default is FALSE, if TRUE is selected, a text message for the threshold will be added.
threshold.line The default is FALSE, and if TRUE is selected, lines for the threshold will be added.
nudge_x Used to adjust the x-axis position of the point where the threshold is located.
nudge_y Used to adjust the y-axis position of the point where the threshold is located.
threshold.linetype The line shape of the threshold line.
threshold.linewidth The line width of the threshold line.
threshold.linecol The colour of the threshold line.
po.text.size The size of the threshold point text.
po.text.col The colour of the threshold point text.
po.text.fill The background of the threshold point text.
liftpec Threshold point left displacement.
rightpec Threshold point right displacement.

Value

A picture.
**Usage**

```r
## S3 method for class 'glm'
scidca(
  fit,
  newdata = NULL,
  timepoint = "median",
  cmprsk = FALSE,
  modelnames = NULL,
  merge = FALSE,
  y.min = NULL,
  xstop = NULL,
  y.max = NULL,
  pyh = NULL,
  relcol = "#c01e35",
  irrelcol = "#0151a2",
  relabel = "Nomogram relevant",
  irrellabel = "Nomogram irrelevant",
  text.size = 4.5,
  text.col = "green",
  colbar = TRUE,
  threshold.text = FALSE,
  threshold.line = FALSE,
  nudge_x = 0,
  nudge_y = 0,
  threshold.linetype = 2,
  threshold.linewidth = 1.2,
  threshold.linecol = "black",
  po.text.size = 4,
  po.text.col = "black",
  po.text.fill = "white",
  liftpec = NULL,
  rightpec = NULL
)
```

**Arguments**

- `fit` Fill in the model you want to analyze. Support survival analysis and logistic regression.
newdata If the decision curve of the validation set is to be analysed. Fill in the validation set data here.

timepoint If it is a survival analysis, fill in the point in time you need to study. The default is the median time.

cmprsk If it is a competitive risk model, select TRUE here.

modelnames Defines the name of the generated image model.

merge If true is selected it will merge the two long zones.

y.min The maximum value of the negative part of the picture. Generally defaults to positive values multiplied by 0.4.

xstop The maximum value of the X-axis of the picture.

y.max The maximum value of the Y-axis. The default value is the maximum net benefit.

pyh The height at which the bars are plotted cannot exceed y.min.

relcol The colour of the relevant part of the bar. The default is red.

irrelcol The colour of the irrelevant part of the bar. The default is blue.

relabel Relevance Tags.

irrellabel No relevant tags.

text.size Font size.

text.col The colour of the font.

colbar The default is true, and if false is selected, bar plotting is cancelled.

threshold.text The default is FALSE, if TRUE is selected, a text message for the threshold will be added.

threshold.line The default is FALSE, and if TRUE is selected, lines for the threshold will be added.

nudge_x Used to adjust the x-axis position of the point where the threshold is located.

nudge_y Used to adjust the y-axis position of the point where the threshold is located.

threshold.linetype The line shape of the threshold line.

threshold.linewidth The line width of the threshold line.

threshold.linecol The colour of the threshold line.

po.text.size The size of the threshold point text.

po.text.col The colour of the threshold point text.

po.text.fill The background of the threshold point text.

liftpec Threshold point left displacement.

rightpec Threshold point right displacement.

Value A picture.
Description

scidca.randomForest

Usage

```r
## S3 method for class 'randomForest'
scidca(
  fit,
  newdata = NULL,
  timepoint = "median",
  cmprsk = FALSE,
  modelnames = NULL,
  merge = FALSE,
  y.min = NULL,
  xstop = NULL,
  y.max = NULL,
  pyh = NULL,
  relcol = "#c01e35",
  irrelcol = "#0151a2",
  relabel = "Nomogram relevant",
  irrellabel = "Nomogram irrelevant",
  text.size = 4.5,
  text.col = "green",
  colbar = TRUE,
  threshold.text = FALSE,
  threshold.line = FALSE,
  nudge_x = 0,
  nudge_y = 0,
  threshold.linetype = 2,
  threshold.linewidth = 1.2,
  threshold.linecol = "black",
  po.text.size = 4,
  po.text.col = "black",
  po.text.fill = "white",
  liftpec = NULL,
  rightpec = NULL
)
```

Arguments

- **fit**: Fill in the model you want to analyze. Support survival analysis and logistic regression.
newdata This parameter is indispensable in the random forest decision curve. Fill in your data.
timepoint If it is a survival analysis, fill in the point in time you need to study. The default is the median time.
cmprsk If it is a competitive risk model, select TRUE here.
modelnames Defines the name of the generated image model.
merge If true is selected it will merge the two long zones.
y.min The maximum value of the negative part of the picture. Generally defaults to positive values multiplied by 0.4.
xstop The maximum value of the X-axis of the picture.
y.max The maximum value of the Y-axis. The default value is the maximum net benefit.
pyh The height at which the bars are plotted cannot exceed y.min.
relcol The colour of the relevant part of the bar. The default is red.
irrelcol The colour of the irrelevant part of the bar. The default is blue.
relabel Relevance Tags.
irrellabel No relevant tags.
text.size Font size.
text.col The colour of the font.
colbar The default is true, and if false is selected, bar plotting is cancelled.
threshold.text The default is FALSE, if TRUE is selected, a text message for the threshold will be added.
threshold.line The default is FALSE, and if TRUE is selected, lines for the threshold will be added.
nudge_x Used to adjust the x-axis position of the point where the threshold is located.
nudge_y Used to adjust the y-axis position of the point where the threshold is located.
threshold.linetype The line shape of the threshold line.
threshold.linewidth The line width of the threshold line.
threshold.linecol The colour of the threshold line.
po.text.size The size of the threshold point text.
po.text.col The colour of the threshold point text.
po.text.fill The background of the threshold point text.
liftpec Threshold point left displacement.
rightpec Threshold point right displacement.

Value
A picture.
Description

Generate data for plotting survival analysis decision curves.

Arguments

data    a data frame containing the variables in the model.
outcome the outcome, response variable. Must be a variable contained within the data frame specified in data=.
predictors the predictor variable(s). Must be a variable(s) contained within the data frame specified in data=.
probability specifies whether or not each of the independent variables are probabilities. The default is TRUE.
xstart starting value for x-axis (threshold probability) between 0 and 1. The default is 0.01.
xstop stopping value for x-axis (threshold probability) between 0 and 1. The default is 0.99.
xby increment for threshold probability. The default is 0.01.
ymin minimum bound for graph.
harm specifies the harm(s) associated with the independent variable(s). The default is none.
graph specifies whether or not to display graph of net benefits. The default is TRUE.
intervention plot net reduction in interventions
interventionper number of net reduction in interventions per interger. The default is 100
loess.span specifies the degree of smoothing. The default is 0.10.
timepoint specifies the time point at which the decision curve analysis is performed.
cmprsk if evaluating outcome in presence of a competing risk. The default is FALSE
smooth specifies whether or not to smooth net benefit curve. The default is FALSE.
ttoutcome Enter the time variable in your data.

Details

This function was created and written by Dr Andrew Vickers to generate decision curve data.

Value

Returns a data for plotting a decision curve.
Index

* LIRI
  LIRI, 5
* datasets
  Breastcancer, 2
* df_surv
  df_surv, 4
Breastcancer, 2
comprskstdca, 2
dca, 3
df_surv, 4
LIRI, 5
newcrr, 5
scidca, 6
scidca.coxph, 9
scidca.crr, 11
scidca.glm, 13
scidca.randomForest, 15
stdca, 17