Package 'VRPM'

July 11, 2017

Type Package

Title Visualizing Risk Prediction Models

Version 1.2

Date 2017-07-11

Author Vanya Van Belle

Maintainer Vanya Van Belle <vanya.vanbelle@esat.kuleuven.be>

Description This is a package to visualize risk prediction models. For each predictor, a color bar represents the contribution to the linear predictor or latent variable. A conversion from the linear predictor to the estimated risk or survival is also given. (Cumulative) contribution charts enable to visualize how the estimated risk for one particular observation is obtained by the model. Several options allow to choose different color maps, and to select the zero level of the contributions. The package is able to deal with 'glm', 'coxph', 'mfp', 'multinom' and 'ksvm' objects. For 'ksvm' objects, the visualization is not always exact. Functions providing tools to indicate the accuracy of the approximation are provided in addition to the visualization.

Imports Hmisc, survival, fields, R2HTML, viridis, kernlab, ROCR, ggplot2, shiny, methods, stats, grDevices, graphics, utils

Suggests mfp, VGAMdata, nnet

License GPL-3

RoxygenNote 5.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2017-07-11 10:42:14 UTC

R topics documented:

`ccchart` *Cumulative contribution chart.*

Description

Display a graph explaining how the risk prediction for a new observation is obtained from the risk prediction model. All contributions are added in a cumulative way to end up with the linear predictor that is transformed into a risk estimate.

Usage

```
ccchart(x, obs, filename, zerolevel = "zero", risklabel, riskcutoff = 0.1,
  type, sorted = FALSE, time, xmin, xmax)
```

Arguments

`x`	`glm`, `coxph`, `mfp`, `multinom` or `ksvm` object.
`obs`	A data.frame containing the predictor values of the observation for which the chart should be made.
`filename`	The name of the resulting file (default: ccchart).
`zerolevel`	The value of the contributions that should be put to zero. If "zero", the contributions are represented as they are. If "min", for each predictor or set of predictors contributing to an interaction, the minimal observed value of the contribution in the training data is substracted from the contribution to ensure that the contribution is always positive. If "median" or "mean", the median or mean value is substracted from the contributions, respectively (default="zero"). See below for more details.
`risklabel`	A character string representing the label for the represented risk. For multinomial logistic regression models, a vector of risk labels should be provided. See the examples for an illustration of the approach.
`riskcutoff`	A value between 0 and 1, indicating the change in color for the represented score and risk bars (default: 0.1). A risk lower than the riskcutoff is visualized in green, a risk higher than the riskcutoff is visualized in red.
`type`	A string specifying the type of plot: "logistic" for logistic regression models, and "survival" for cox proportional hazard models. This should generally not be provided by the user.
`sorted`	logical. If TRUE the contributions are sorted in increasing order (default=FALSE).
`time`	The time at which the estimated survival should be calculated. As default, the estimated survival at median survival time is reported. If the median survival time can not be calculated, the estimated survival at the latest event time is reported. For objects that are not a member of the `coxph` class, this is redundant.

xmin	Minimal values of input variables to be represented on the visualization. These values only have an influence on continuous input variables.
xmax	Maximal values of input variables to be represented on the visualization. These values only have an influence on continuous input variables.

Details

The cumulative contribution chart constitutes from different bars, representing the contribution of the predictors to the score (translated linear predictor) in a cumulative way. Depending on the value of zerolevel, the visualized contributions are slightly different. If zerolevel="zero", the contribution for variable x^p is $\beta_p f_p(x^p)$, with β_p the model coefficient corresponding to this predictor and $f_p(x^p)$ a (possible) transformation of x^p. If zerolevel is "min", "median" or "mean", a value equal to the minimum, median and mean of the contribution $\beta_p f_p(x^p)$ in the training data, respectively, is substracted from the contribution. See the references for more information. At the bottom of the chart, the score (i.e. the translated linear predictor or latant variable) of this observation is reported next to the maximal score observed in the training data. The risk corresponding to both of these scores are represented as well. Risk predictions above the riskcutoff result in red bars, risk predictions below this cutoff result in green bars.

Note

This graph can not be used for cox proportional hazard regression including strata.

For coxph models, it is necessary to include model=TRUE in the model fit.

For multinom models, it is necessary to include model=TRUE in the model fit.

For multinom models, more than one output file is generated. A first series of plots visualizes how the linear predictors are obtained. The files are named "filename_outcome_level_cchart", with "outcome_level" the name of the outcome level for which the linear predictor is visualized. A second series of plots visualizes how the linear predictors are transformed into a risk prediction for each outcome level. The files are named "filename_p_outcome_level_cchart".

For multinom models, a vector of risk labels needs to be made and provided to the ccchart() function. See the examples for an illustration of the approach.

For ksvm models, it is necessary to include prob.model=TRUE in the model fit.

The plot is not shown in a graphical window but saved in the current working directory.

Author(s)

Vanya Van Belle

References

Van Belle V., Van Calster B., *Visualizing risk prediction models*, PLoS ONE, 10(7):e0132614. doi:10.1371/journal.pone.0132614 (2015).

Van Belle V., Van Calster B., Suykens J.A.K., Van Huffel S. and Lisboa P., *Explaining support vector machines: a color based nomogram*, Internal Report 16-27, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

Van Belle V., Van Huffel S., Timmerman D., Froyman W., Bourne T. and Van Calster B., *A color based nomogram for Multinomial Logistic Regression*, Internal Report 16-28, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

See Also

colplot, cchart

Examples

```
#### logistic regression
mydata <- read.csv("https://stats.idre.ucla.edu/stat/data/binary.csv")
mydata$rank <- factor(mydata$rank)
fit <- glm(admit ~ gre + gpa + rank, data = mydata, family = "binomial")
patient1=data.frame(gre=386,gpa=3.58,rank=3)
ccchart(fit,obs=patient1,filename="ccchart1")

#### cox proportional hazard regression
library(mfp)
data(GBSG)
fit<-coxph(Surv(rfst, cens) ~ age+tumsize+posnodal+prm+esm+menostat+tumgrad, data = GBSG,
model=TRUE)
patient1=data.frame(age=52,tumsize=45,posnodal=23,prm=1025,esm=562,menostat=2,tumgrad=3)
ccchart(fit,patient1,time=500,filename="ccchart2")

#### multinomial logistic regression model
library(nnet)
library(VGAMdata)
data(xs.nz)
marital.nz <- xs.nz[,c("marital","sex","age","height","weight")]
mydata <- marital.nz[complete.cases(marital.nz),]
fit <- multinom(marital ~ sex + age + height + weight, data = mydata,model=TRUE)
# for multinimial logistic regression, a vector of risk labels needs to be made
# and provided to the colplot function
outnames=colnames(fitted(fit))
labels=c(paste("Linear predictor for",outnames[-1]),paste
("Predicted chance of being",outnames))
patient1=data.frame(sex="F",age=27,height=1.68,weight=58.6)
ccchart(fit,obs=patient1,filename="multinom_ccchart",risklabel=labels,zerolevel="zero")

#### Support Vector Machine classifier
## Not run:
library(kernlab)
data(iris)
levels(iris$Species)[levels(iris$Species)=="setosa"] <- "other"
levels(iris$Species)[levels(iris$Species)=="virginica"] <- "other"
names(iris)=c("SL","SW","PL","PW","Species")
set.seed(100)
model <-ksvm(Species ~ ., data = iris,prob.model=TRUE,kpar=list(0.03),C=10)
obs1=data.frame(SL=5.2,SW=3.0,PL=1.5,PW=0.3)
# The plot should be based on all training data, so the following code should be used:
newmodel=preplotperf(model,iris,indy=5,zerolevel="min")
```

```
ccchart(newmodel,obs=obs1,filename="iris_ccchart",risklabel="Chance on versicolor",
zerolevel="median")

## End(Not run)
```

---

cchart *Contribution chart.*

Description

Display a graph explaining how the risk prediction of a new observation is obtained from the model. All contributions to the linear predictor or latent variable are visualized and summed to obtain the linear predictor. The latter is then transformed into a risk estimate.

Usage

```
cchart(x, obs, filename, zerolevel = "zero", risklabel, sorted = FALSE,
  time, xmin, xmax)
```

Arguments

x	glm, coxph, mfp, multinom or ksvm object.
obs	A data.frame containing the predictor values of the observation for which the chart should be made.
filename	The name of the resulting file (default: cchart).
zerolevel	The value of the contributions that should be put to zero. If "zero", the contributions are represented as they are. If "min", for each predictor or set of predictors contributing to an interaction, the minimal observed value of the contribution in the training data is substracted from the contribution to ensure that the contribution is always positive. If "median" or "mean", the median or mean value is substracted from the contributions, respectively (default="zero"). See below for more details.
risklabel	A character string representing the label for the represented risk. For multinomial logistic regression models, a vector of risk labels should be provided. See the examples for an illustration of the approach.
sorted	logical. If TRUE the contributions are sorted in increasing order (default=FALSE).
time	The time at which the estimated survival should be calculated. As default, the estimated survival at median survival time is reported. If the median survival time can not be calculated, the estimated survival at the latest event time is reported. For objects that are not a member of the coxph class, this is redundant.
xmin	Minimal values of input variables to be represented on the visualization. These values only have an influence on continuous input variables.
xmax	Maximal values of input variables to be represented on the visualization. These values only have an influence on continuous input variables.

Details

The contribution chart is a bar plot representing the contribution of each predictor or set of predictors to the score (translated linear predictor) by means of bars. Depending on the value of `zerolevel`, the visualized contributions are slightly different. If `zerolevel="zero"`, the contribution for variable x^p is $\beta_p f_p(x^p)$, with β_p the model coefficient corresponding to this predictor and $f_p(x^p)$ a (possible) transformation of x^p. If `zerolevel` is "min", "median" or "mean", a value equal to the minimum, median and mean of the contribution $\beta_p f_p(x^p)$ in the training data, respectively, is substracted from the contribution. See the references for more information. The sum of all the contributions is the score (i.e. translated linear predictor or latent variable) which is transformed into the risk estimate. The range of all contributions in the training set are visualized by means of black horizontal lines.

Note

This graph can not be used for cox proportional hazard regression including strata.

For `coxph` models, it is necessary to include `model=TRUE` in the model fit.

For `multinom` models, it is necessary to include `model=TRUE` in the model fit.

For `multinom` models, more than one output file is generated. A first series of plots visualizes how the linear predictors are obtained. The files are named "filename_outcome_level_cchart", with "outcome_level" the name of the outcome level for which the linear predictor is visualized. A second series of plots visualizes how the linear predictors are transformed into a risk prediction for each outcome level. The files are named "filename_p_outcome_level_cchart".

For `multinom` models, a vector of risk labels needs to be made and provided to the `cchart()` function. See the examples for an illustration of the approach.

For `ksvm` models, it is necessary to include `prob.model=TRUE` in the model fit.

The plot is not shown in a graphical window but saved in the current working directory.

In case `zerolevel="min"` and `xmin` and `xmax` are provided by the user, it is possible to have negative contributions within the training data.

Author(s)

Vanya Van Belle

References

Van Belle V., Van Calster B., *Visualizing risk prediction models*, PLoS ONE, 10(7):e0132614. doi:10.1371/journal.pone.0132614 (2015).

Van Belle V., Van Calster B., Suykens J.A.K., Van Huffel S. and Lisboa P., *Explaining support vector machines: a color based nomogram*, Internal Report 16-27, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

Van Belle V., Van Huffel S., Timmerman D., Froyman W., Bourne T. and Van Calster B., *A color based nomogram for Multinomial Logistic Regression*, Internal Report 16-28, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

See Also

colplot, ccchart

Examples

```
#### logistic regression
mydata <- read.csv("https://stats.idre.ucla.edu/stat/data/binary.csv")
mydata$rank <- factor(mydata$rank)
fit <- glm(admit ~ gre + gpa + rank, data = mydata, family = "binomial")
patient1=data.frame(gre=386,gpa=3.58,rank=3)
cchart(fit,obs=patient1,filename="cchart1")

#### cox proportional hazard regression
library(mfp)
data(GBSG)
fit<-coxph(Surv(rfst, cens) ~ age+tumsize+posnodal+prm+esm+menostat+tumgrad, data = GBSG,
model=TRUE)
patient1=data.frame(age=52,tumsize=45,posnodal=23,prm=1025,esm=562,menostat=2,tumgrad=3)
# Indicate the risk estimate at 500 days
cchart(fit,patient1,time=500,filename="cchart6")


#### multinomial logistic regression model
library(nnet)
library(VGAMdata)
data(xs.nz)
marital.nz <- xs.nz[,c("marital","sex","age","height","weight")]
mydata <- marital.nz[complete.cases(marital.nz),]
fit <- multinom(marital ~ sex + age + height + weight, data = mydata,model=TRUE)
# for multinimial logistic regression, a vector of risk labels needs to be made
# and provided to the colplot function
outnames=colnames(fitted(fit))
labels=c(paste("Linear predictor for",outnames[-1]),paste
("Predicted chance of being",outnames))
patient1=data.frame(sex="F",age=27,height=1.68,weight=58.6)
cchart(fit,obs=patient1,filename="multinom",risklabel=labels,zerolevel="zero")

#### Support Vector Machine classifier
## Not run:
library(kernlab)
data(iris)
levels(iris$Species)[levels(iris$Species)=="setosa"] <- "other"
levels(iris$Species)[levels(iris$Species)=="virginica"] <- "other"
names(iris)=c("SL","SW","PL","PW","Species")
set.seed(100)
model <-ksvm(Species ~ ., data = iris,prob.model=TRUE,kpar=list(0.03),C=10)
obs1=data.frame(SL=5.2,SW=3.0,PL=1.5,PW=0.3)
# The plot should be based on all training data, so the following code should be used:
newmodel=preplotperf(model,iris,indy=5,zerolevel="median")
cchart(newmodel,obs=obs1,filename="iris2",zerolevel="median")

## End(Not run)
```

`colplot` *Visualize a risk prediction model by means of colored bars.*

Description

Display a graph (color based nomogram) in which the contributions of each predictor or set of predictors is represented in a colored bar. The color indicates the value of the contribution.

Usage

```
colplot(x, filename, coloroptions = 2, zerolevel = "zero", risklabel, xmin,
  xmax, adverse, obs, q5, q95, time)
```

Arguments

`x`	`glm`, `coxph`, `mfp`, `multinom` or `ksvm` object.
`filename`	The name of the resulting file (default: colplot).
`coloroptions`	If 1, the rainbow color map is used. If 2, a sequential color map is used. If 3, a diverging color map is used. If 4, a black-and-white color map is used. If 5, the viridis color map is used. (default=2)
`zerolevel`	The value of the contributions that should be put to zero. If "zero", the contributions are represented as they are. If "min", for each predictor or set of predictors contributing to an interaction, the minimal observed value of the contribution in the training data is substracted from the contribution to ensure that the contribution is always positive. If "median" or "mean", the median or mean value is substracted from the contributions, respectively (default="zero"). See below for more details.
`risklabel`	A character string representing the label for the represented risk. For multinomial logistic regression models, a vector of risk labels should be provided. See the examples for an illustration of the approach.
`xmin`	Minimal values of input variables to be represented on the visualization. These values only have an influence on continuous input variables.
`xmax`	Maximal values of input variables to be represented on the visualization. These values only have an influence on continuous input variables.
`adverse`	A logical indicating whether the score and risk range in the adverse direction (default=FALSE, i.e. high score corresponds to a high risk).
`obs`	A data.frame containing the predictor values of the observation that should be added to the plot.
`q5`	A data.frame containing the predictor values of the 5th percentiles of the predictors that should be added to the plot. This only impacts the plot for continous variables.
`q95`	A data.frame containing the predictor values of the 95th percentiles of the predictors that should be added to the plot. This only impacts the plot for continous variables.

`time` The time at which the estimated survival should be calculated. As default, the estimated survival at median survival time is reported. If the median survival time can not be calculated, the estimated survival at the latest event time is reported. For objects that are not a member of the `coxph` class, this is redundant.

Details

The colorplot is a chart (similar to a nomogam) that visualizes the contribution of a predictor or a set of predictors by means of colored bars. Depending on the value of `zerolevel`, the visualized contributions are slightly different. If `zerolevel="zero"`, the contribution for variable x^p is $\beta_p f_p(x^p)$, with β_p the model coefficient corresponding to this predictor and $f_p(x^p)$ a (possible) transformation of x^p. If `zerolevel` is "min", "median" or "mean", a value equal to the minimum, median and mean of the contribution $\beta_p f_p(x^p)$ in the training data, respectively, is substracted from the contribution. See the references for more information. `coloroptions` enables to choose between diffent color maps. It is recommended to use the sequential or the viridis color map when `zerolevel="min"` and a diverging color map when `zerolevel` is "median" or "mean". For the latter case, a white color will correspond to zero points. In the color bar converting the score to a risk, white will correspond to the median or mean observed risk in the training data, respectively.

Note

This graph can not be used for cox proportional hazard regression including strata.

For `coxph` models, it is necessary to include `model=TRUE` in the model fit.

For `multinom` models, it is necessary to include `model=TRUE` in the model fit.

For `multinom` models, more than one output file is generated. A first series of plots visualizes how the linear predictors are obtained. The files are named "filename_outcome_level", with "outcome_level" the name of the outcome level for which the linear predictor is visualized. A second series of plots visualizes how the linear predictors are transformed into a risk prediction for each outcome level. The files are named "filename_p_outcome_level". A third series of plots uses an alternative way to represent the calculation of the risk of the non-reference outcome levels. These plots are named "filename_outcome_level_wing".

For `multinom` models, a vector of risk labels needs to be made and provided to the `colplot()` function. See the examples for an illustration of the approach.

For `ksvm` models, it is necessary to include `prob.model=TRUE` in the model fit.

The plot is not shown in a graphical window but saved in the current working directory.

Author(s)

Vanya Van Belle

References

Van Belle V., Van Calster B., *Visualizing risk prediction models*, PLoS ONE, 10(7):e0132614. doi:10.1371/journal.pone.0132614 (2015).

Van Belle V., Van Calster B., Suykens J.A.K., Van Huffel S. and Lisboa P., *Explaining support vector machines: a color based nomogram*, Internal Report 16-27, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

Van Belle V., Van Huffel S., Timmerman D., Froyman W., Bourne T. and Van Calster B., *A color based nomogram for Multinomial Logistic Regression*, Internal Report 16-28, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

See Also

cchart, ccchart

Examples

```
# default options
mydata <- read.csv("https://stats.idre.ucla.edu/stat/data/binary.csv")
mydata$rank <- factor(mydata$rank)
fit <- glm(admit ~ gre + gpa + rank, data = mydata, family = "binomial")
colplot(fit)

#### cox proportional hazard regression
library(mfp)
data(GBSG)
fit<-coxph(Surv(rfst, cens) ~ age+tumsize+posnodal+prm+esm+menostat+tumgrad, data = GBSG,
model=TRUE)
colplot(fit)

#### multinomial logistic regression model
library(nnet)
library(VGAMdata)
data(xs.nz)
marital.nz <- xs.nz[,c("marital","sex","age","height","weight")]
mydata <- marital.nz[complete.cases(marital.nz),]
fit <- multinom(marital ~ sex + age + height + weight, data = mydata,model=TRUE)
# for multinimial logistic regression, a vector of risk labels needs to be made
# and provided to the colplot function
outnames=colnames(fitted(fit))
labels=c(paste("Linear predictor for",outnames[-1]),paste
("Predicted chance of being",outnames))
# visualize the model: more than one plot is generated in the current directory
colplot(fit,coloroptions=3,risklabel=labels,filename="div")


#### Support Vector Machine classifier
## Not run:
library(kernlab)
data(iris)
levels(iris$Species)[levels(iris$Species)=="setosa"] <- "other"
levels(iris$Species)[levels(iris$Species)=="virginica"] <- "other"
names(iris)=c("SL","SW","PL","PW","Species")
# RBF kernel
model <-ksvm(Species ~ ., data = iris,prob.model=TRUE,kpar=list(0.03),C=10)
# The plot should be based on all training data, so the following code should be used:
newmodel=preplotperf(model,iris,indy=5,zerolevel="min")
colplot(newmodel,filename="IRIS2",zerolevel="min",coloroptions=5)
```

```
## End(Not run)
```

HTMLsummary *Summarize the risk prediction plots.*

Description

Create an HTML page that summarizes all the plots that have been created to visualize the risk prediction. When a patient was given, a patient summary will be given as well. This function is only applicable for `multinom` objects.

Usage

```
HTMLsummary(fit, summaryfile, filename, title)
```

Arguments

`fit`	A `multinom` object or a character vector. In the latter case the vector consists of the names of the outcome levels.
`summaryfile`	Name of the resulting HTML file
`filename`	The filename that was given to the `colplot`, `cchart` or `ccchart` commands.
`title`	Title of the summary.

Author(s)

Vanya Van Belle

References

Van Belle V., Van Huffel S., Timmerman D., Froyman W., Bourne T. and Van Calster B., *A color based nomogram for Multinomial Logistic Regression*, Internal Report 16-28, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

Examples

```
# only applicable for multinom objects
library(nnet)
library(R2HTML)
library(VGAMdata)
data(xs.nz)
marital.nz <- xs.nz[,c("marital","sex","age","height","weight")]
mydata <- marital.nz[complete.cases(marital.nz),]
fit <- multinom(marital ~ sex + age + height + weight, data = mydata,model=TRUE)
# for multinimial logistic regression, a vector of risk labels needs to be made
# and provided to the colplot function
outnames=colnames(fitted(fit))
labels=c(paste("Linear predictor for",outnames[-1]),paste
```

```
("Predicted chance of being",outnames))
patient1=data.frame(sex="F",age=27,height=1.68,weight=58.6)
colplot(fit,coloroptions=3,risklabel=labels,filename="multinom")
# this generates a file mysummary.html containing all plots provided by the above
# colplot statement.
HTMLsummary(fit=fit, summaryfile="mysummary", filename="multinom",
title="Global summary of a multinomial logistic regression model estimating
the risk on different marital statuses.")
# this generates a file mysummary1.html containing all plots provided by the below
# colplot statement.  Since an observation is provided, a patient summary will also
# be given.
colplot(fit,obs=patient1,coloroptions=3,risklabel=labels,filename="multinom")
HTMLsummary(fit=fit, summaryfile="mysummary1", filename="multinom",
title="Global summary of a multinomial logistic regression model estimating
the risk on different marital statuses.")
```

`plotperf` *Performance plots for the approximation of an SVM model.*

Description

Generate performance plots for the approximation of an SVM model.

Usage

```
plotperf(mymodel, mydata, indy, mytestdata, type = "all", filename)
```

Arguments

`mymodel`	Element of class `ksvm`.
`mydata`	Data on which `mymodel` was trained on.
`indy`	Column number of the outcome in `mydata`.
`mytestdata`	Data on which to evaluate `mymodel`. (Optional)
`type`	Type of performance plot (c="all","lp","probs","outcomes","contributions", "ROC","corrplot"). See details for more information.
`filename`	Name of the resulting graph.

Details

Different types of plots are possible. When `type="all"`, all the options are generated. When `type="lp"`, the latent variables of the approximation and the original SVM model are plotted against eachother. When `type="probs"` the estimated probabilities of the approximation and the SVM model are plotted against eachother. When `type="outcomes"` a bubble plot indicating the agreement between the approximation and the SVM model is generated. When `type="contributions"`, the range of the contributions within the approximation, the range of the rest term and the range of the latent variable of the SVM model are represented by means of boxplots. All of these are shifted to have a median equal to zero. When `type="ROC"`, ROC curves for the approximation and the SVM model are plotted. When mytestdata is non-empty, ROC curves for the test set are also provided.

Author(s)

Vanya Van Belle

References

Van Belle V., Van Calster B., Suykens J.A.K., Van Huffel S. and Lisboa P., *Explaining support vector machines: a color based nomogram*, Internal Report 16-27, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

Examples

```
#### Support Vector Machine classifier
library(kernlab)
data(iris)
levels(iris$Species)[levels(iris$Species)=="setosa"] <- "other"
levels(iris$Species)[levels(iris$Species)=="virginica"] <- "other"
names(iris)=c("SL","SW","PL","PW","Species")
# good model
model <-ksvm(Species ~ ., data = iris,prob.model=TRUE,kpar=list(0.03),C=10)
# bad model
model2 <-ksvm(Species ~ ., data = iris,prob.model=TRUE,kpar=list(10),C=10)
# plot latent variables of approximation and SVM
plotperf(model,iris,indy=5,type="lp",filename="iris")
plotperf(model2,iris,indy=5,type="lp",filename="iris2")
# plot contributions of approximation and SVM
# good model: rest term is small in comparison with other contributions and lpmodel
# (latent variable of SVM)
plotperf(model,iris,indy=5,type="contributions",filename="iris")
# bad model: rest term is large in comparison with other contributions and lpmodel
# (latent variable of SVM)
plotperf(model2,iris,indy=5,type="contributions",filename="iris2")
# plot latent variables of approximation and SVM
plotperf(model,iris,indy=5,type="outcomes",filename="iris")
plotperf(model2,iris,indy=5,type="outcomes",filename="iris2")
```

---

preplotperf *Preprocess a* ksvm *object*

Description

Performs necessary preprocessing of a ksvm object when the plots should be generated based on all training data and not only on the support vectors.

Usage

```
preplotperf(model, mydata, indy, mytestdata, zerolevel = "zero",
  risklabel = "Estimated risk", adverse = FALSE)
```

Arguments

model	Object of class ksvm
mydata	Data on which mymodel was trained on.
indy	Column number of the outcome in mydata.
mytestdata	Data on which to evaluate mymodel. (Optional)
zerolevel	The value of the contributions that should be put to zero. If "zero", the contributions are represented as they are. If "min", for each predictor or set of predictors contributing to an interaction, the minimal observed value of the contribution in the training data is substracted from the contribution to ensure that the contribution is always positive. If "median" or "mean", the median or mean value is substracted from the contributions, respectively (default="zero"). See below for more details.
risklabel	A character string representing the label for the represented risk. For multinomial logistic regression models, a vector of risk labels should be provided. See the examples for an illustration of the approach.
adverse	A logical indicating whether the score and risk range in the adverse direction (default=FALSE, i.e. high score corresponds to a high risk).

Details

Depending on the value of zerolevel, the visualized contributions are slightly different. If zerolevel="zero", the contribution for variable x^p is $\beta_p f_p(x^p)$, with β_p the model coefficient corresponding to this predictor and $f_p(x^p)$ a (possible) transformation of x^p. If zerolevel is "min", "median" or "mean", a value equal to the minimum, median and mean of the contribution $\beta_p f_p(x^p)$ in the training data, respectively, is substracted from the contribution. See the references for more information.

Value

List object

Author(s)

Vanya Van Belle

References

Van Belle V., Van Calster B., Suykens J.A.K., Van Huffel S. and Lisboa P., *Explaining support vector machines: a color based nomogram*, Internal Report 16-27, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

See Also

colplot, cchart, ccchart

`runVRPMexample` *Run R Shiny app*

Description

Run a selected R Shiny application to illustrate the working of the VRPM package.

Usage

```
runVRPMexample(example)
```

Arguments

`example`	The name of the application that should be loaded. `example` should be one of "IntermittentClaudication", "Iris", "Pima" or "xsnz".

Details

Different applications are possible. To illustrate the visualization of a logistic regression model, an application on the Intermittent Claudication model (see references) can be loaded using `example=` "IntermittentClaudication". Two illustrations for the visualization of support vector machine classifiers can be loaded using `example="Iris"` and `example="Pima"`. To illustrate the possibilities for visualizing multinomial logistic regression models, use `example="xsnz"`.

Author(s)

Vanya Van Belle

References

Van Belle V., Van Calster B., *Visualizing risk prediction models*, PLoS ONE, 10(7):e0132614. doi:10.1371/journal.pone.0132614 (2015).

Van Belle V., Van Calster B., Suykens J.A.K., Van Huffel S. and Lisboa P., *Explaining support vector machines: a color based nomogram*, Internal Report 16-27, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

Van Belle V., Van Huffel S., Timmerman D., Froyman W., Bourne T. and Van Calster B., *A color based nomogram for Multinomial Logistic Regression*, Internal Report 16-28, ESAT-Stadius, KU Leuven (Leuven, Belgium), 2016

Examples

```
## Not run:
#### Logistic regression model
runVRPMexample("IntermittentClaudication")
#### Multinomial logistic regression model
runVRPMexample("xsnz")
#### Support Vector Machine classifier
runVRPMexample("Iris")
```

```
runVRPMexample("Pima")
## End(Not run)
```

Index