Package 'trtf'

April 23, 2018

Title Transformation Trees and Forests

Version 0.3-1

Date 2018-04-20

Description Recursive partytioning of transformation models with
corresponding random forest for conditional transformation models
as described in 'Transformation Forests' (Hothorn and Zeileis, 2017, <arXiv:1701.02110>)
and 'Top-Down Transformation Choice' (Hothorn, 2018, <DOI:10.1177/1471082X17748081>).

Depends mlt (>= 1.0-1), partykit (>= 1.2-1)

Imports Formula, sandwich, grid, stats, variables, libcoin, utils

Suggests survival, TH.data, coin

License GPL-2

NeedsCompilation no

Author Torsten Hothorn [aut, cre] (<https://orcid.org/0000-0001-8301-0471>)

Maintainer Torsten Hothorn <Torsten.Hothorn@R-project.org>

Repository CRAN

Date/Publication 2018-04-23 11:51:50 UTC

R topics documented:

trtf-package *General Information on the* **trtf** *Package*

Description

The **trtf** package implements transformation trees and transformation forests as described in Hothorn and Zeileis (2017).

Example applications of transformation trees and forests can be replicated using `demo("applications")` and `demo("BMI")`. Figure 1 in Hothorn and Zeileis (2017) can be reproduced by `demo("QRF")`. Source code of simulation experiments is available in directory `trtf/inst/sim`.

A short talk introducing transformation trees and forests is available from https://channel9.msdn.com/Events/useR-international-R-User-conferences/useR-International-R-User-2017-Conference/Transformation-Forests.

Author(s)

This package is authored by Torsten Hothorn <Torsten.Hothorn@R-project.org>.

References

Torsten Hothorn and Achim Zeileis (2017). Transformation Forests. https://arxiv.org/abs/1701.02110.

traforest *Transformation Forests*

Description

Partitioned and aggregated transformation models

Usage

```
traforest(object, parm = 1:length(coef(object)), mltargs = list(maxit = 10000),
          update = TRUE, ...)
```

Arguments

`object`	an object of class `ctm` or `mlt` specifying the abstract model to be partitioned.
`parm`	parameters of `object` those corresponding score is used for finding partitions.
`mltargs`	arguments to `mlt` for fitting the transformation models.
`update`	logical, if `TRUE`, models and thus scores are updated in every node. If `FALSE`, the model and scores are computed once in the root node. The latter option is faster but less accurate.
`...`	arguments to `cforest`, at least `formula` and `data`.

Details

Conditional inference trees are used for partitioning likelihood-based transformation models as described in Hothorn and Zeileis (2017). The method can be seen in action in Hothorn (2018) and the corresponding code is available as `demo("BMI")`.

Value

An object of class `traforest` with corresponding `logLik` and `predict` methods.

References

Torsten Hothorn and Achim Zeileis (2017). Transformation Forests. https://arxiv.org/abs/1701.02110.

Torsten Hothorn (2018). Top-Down Transformation Choice. *Statistical Modelling*, https://arxiv.org/abs/1706.08269.

Examples

```
### Example: Personalised Medicine Using Partitioned and Aggregated Cox-Models
### A combination of <DOI:10.1177/0962280217693034> and <arXiv:1701.02110>
### based on infrastructure in the mlt R add-on package described in
### https://cran.r-project.org/web/packages/mlt.docreg/vignettes/mlt.pdf

library("trtf")
library("survival")
### German Breast Cancer Study Group 2 data set
data("GBSG2", package = "TH.data")

### set-up Cox model with overall treatment effect in hormonal therapy
yvar <- numeric_var("y", support = c(100, 2000), bounds = c(0, Inf))
By <- Bernstein_basis(yvar, order = 5, ui = "incre")
m <- ctm(response = By, shifting = ~ horTh, todistr = "MinExt", data = GBSG2)
GBSG2$y <- with(GBSG2, Surv(time, cens))

### overall log-hazard ratio
coef(cmod <- mlt(m, data = GBSG2))["horThyes"]
### roughly the same as
coef(coxph(y ~ horTh, data = GBSG2))

## Not run:

### estimate age-dependent Cox models (here ignoring all other covariates)
ctrl <- ctree_control(minsplit = 50, minbucket = 20, mincriterion = 0)
set.seed(290875)
tf_cmod <- traforest(m, formula = y ~ horTh | age, control = ctrl,
                     ntree = 50, mtry = 1, trace = TRUE, data = GBSG2)

### plot age-dependent treatment effects vs. overall treatment effect
nd <- data.frame(age = 30:70)
cf <- predict(tf_cmod, newdata = nd, type = "coef")
```

```
nd$logHR <- sapply(cf, function(x) x["horThyes"])
plot(logHR ~ age, data = nd, pch = 19, xlab = "Age", ylab = "log-Hazard Ratio")
abline(h = coef(cmod <- mlt(m, data = GBSG2))["horThyes"])
### treatment most beneficial in very young patients
### NOTE: scale of log-hazard ratios depends on
### corresponding baseline hazard function which  _differs_
### across age; interpretation of positive / negative treatment effect is,
### however, save.

## End(Not run)
```

`trafotree` *Transformation Trees*

Description

Partitioned transformation models

Usage

```
trafotree(object, parm = 1:length(coef(object)), mltargs = list(maxit = 10000), ...)
```

Arguments

`object`	an object of class `ctm` or `mlt` specifying the abstract model to be partitioned.
`parm`	parameters of `object` those corresponding score is used for finding partitions.
`mltargs`	arguments to `mlt` for fitting the transformation models.
`...`	arguments to `ctree`, at least `formula` and `data`.

Details

Conditional inference trees are used for partitioning likelihood-based transformation models as described in Hothorn and Zeileis (2017). The method can be seen in action in Hothorn (2018) and the corresponding code is available as `demo("BMI")`. `demo("applications")` performs transformation tree analyses for some standard benchmarking problems.

Value

An object of class `trafotree` with corresponding `plot`, `logLik` and `predict` methods.

References

Torsten Hothorn and Achim Zeileis (2017). Transformation Forests. https://arxiv.org/abs/1701.02110.

Torsten Hothorn (2018). Top-Down Transformation Choice. *Statistical Modelling*, https://arxiv.org/abs/1706.08269.

Examples

```
### Example: Stratified Medicine Using Partitioned Cox-Models
### A combination of <DOI:10.1515/ijb-2015-0032> and <arXiv:1701.02110>
### based on infrastructure in the mlt R add-on package described in
### https://cran.r-project.org/web/packages/mlt.docreg/vignettes/mlt.pdf

library("trtf")
library("survival")
### German Breast Cancer Study Group 2 data set
data("GBSG2", package = "TH.data")

### set-up Cox model with overall treatment effect in hormonal therapy
yvar <- numeric_var("y", support = c(100, 2000), bounds = c(0, Inf))
By <- Bernstein_basis(yvar, order = 5, ui = "incre")
m <- ctm(response = By, shifting = ~ horTh, todistr = "MinExt", data = GBSG2)
GBSG2$y <- with(GBSG2, Surv(time, cens))

### overall log-hazard ratio
coef(cmod <- mlt(m, data = GBSG2))["horThyes"]
### roughly the same as
coef(coxph(y ~ horTh, data = GBSG2))

### partition the model, ie both the baseline hazard function AND the
### treatment effect
(part_cmod <- trafotree(m, formula = y ~ horTh | age + menostat + tsize +
    tgrade + pnodes + progrec + estrec, data = GBSG2))

### compare the log-likelihoods
logLik(cmod)
logLik(part_cmod)

### stronger effects in nodes 2 and 4 and no effect in node 5
coef(part_cmod)[, "horThyes"]

### plot the conditional survivor functions; blue is untreated
### and green is hormonal therapy
nd <- data.frame(horTh = sort(unique(GBSG2$horTh)))
plot(part_cmod, newdata = nd,
     tp_args = list(type = "survivor", col = c("cadetblue3", "chartreuse4")))
```

Index