Package 'rpart.LAD'

May 20, 2020

Type Package

Title Least Absolute Deviation Regression Trees

Version 0.1.2

Date 2020-05-20

Author Stephan Dlugosz [aut,cre]

Maintainer Stephan Dlugosz <stephan.dlugosz@googlemail.com>

Description Recursive partitioning for least absolute deviation regression trees. Another algorithm from the 1984 book by
Breiman, Friedman, Olshen and Stone in addition to the 'rpart' package (Breiman, Friedman, Olshen, Stone (1984, ISBN:9780412048418).

License GPL-3

NeedsCompilation yes

Depends R (>= 3.0.2), Rcpp (>= 0.12.3), rpart (>= 3.1.0)

LinkingTo Rcpp (>= 0.12.3)

SystemRequirements C++11

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

Repository CRAN

Date/Publication 2020-05-20 05:10:02 UTC

R topics documented:

LAD — *'rpart'-method: List of required functions for inducing 'rpart'-like LAD regression trees*

Description

'rpart'-method: List of required functions for inducing 'rpart'-like LAD regression trees

Usage

```
LAD
```

Format

An object of class `list` of length 4.

Examples

```
mystate <- data.frame(state.x77, region = state.region)
names(mystate) <- casefold(names(mystate)) #remove mixed case

fit <-  rpart(murder ~ ., data = mystate, minsplit = 10, method = LAD)
plot(fit); text(fit)
```

rpart.LAD — *Least Absolute Deviation Regression Trees*

Description

Recursive partitioning for least absolute deviation regression trees. Another algorithm from the 1984 book by Breiman, Friedman, Olshen and Stone in addition to the rpart package.

Details

The package provides a list LAD which can be used as method parameter to rpart.

According to continuous regressors, this implementation directly follows the description in the 1984 book by Breiman, Friedman, Olshen and Stone, but the mentioned "updating" procedure for median computation has been replaced by an sort-and-search approach, which efficiently allows to compute all weighted medians for potential splits in $O(n \log n)$. Computation of the LAD however is $O(n^2)$ complex.

According to discrete regressors, the algorithm makes use of the (rpart-specific) simplification heuristic which takes just those splits into account, which are in line with the ordering of the categories with respect to their medians.

Index