

Package ‘SQB’

July 8, 2018

Title Sequential Bagging on Regression

Description Methodology: Remove one observation. Training the rest of data that are sampled without replacement and given this observation's input, predict the response back. Replicate this N times and for each response, take a sample from these replicates with replacement. Average each responses of sample and again replicate this step N time for each observation. Approximate these N new responses by using bootstrap method and generate another N responses y'. Training these y' and predict to have N responses of each testing observation. The average of N is the final prediction. Each observation will do the same.

Version 0.4

Author Moshu Xie <mxie622@gmail.com>

Maintainer Moshu Xie <mxie622@gmail.com>

Depends R (>= 3.2.5), rpart , parallel, caret, nnet, pls

Suggests MASS

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2018-07-07 23:00:02 UTC

R topics documented:

SQB	2
SQBreg	2
Index	4

 SQB

Sequential bagging method on regression

Description

Combine bagging, bootstrap decision tree and kernel density estimate

Author(s)

Moshu Xie <mxie622@gmail.com>

SQBreg

*SQBreg***Description**

Do prediction using sequential bagging method with tree based learning algorithm

Usage

```
SQBreg(data.train, data.test, y, res, reps, cores, FunkDE, control,
        SQBalgorithm.1, SQBalgorithm.2, k, ncomp, nnet.size)
```

Arguments

data.train	Training dataset
data.test	Testing dataset
y	Numeric response variable
res	Resampling size. Could not be greater than the input data size.
reps	Replicates for the first bagging, default 100
cores	Use multi-cores, default one core, use cores='maxcores' for full use.
FunkDE	Kernel density estimate function. Use different kernel to fit, default logistic kernel.
control	Use in rpart package, rpart.control to tune the control
SQBalgorithm.1	Use for the initial training. Option: CART, lm(default), knnreg, nnet, PCR.
SQBalgorithm.2	Use for the last training. Option: CART, lm(default), knnreg, nnet, PCR.
k	The number of nearest neighbour used for knnreg
ncomp	The number of component used for PCR
nnet.size	The number of hidden layer and neuron for nnet

Value

Given testing set input, make a regression prediction

References

Breiman L., Friedman J. H., Olshen R. A., and Stone, C. J. (1984) *Classification and Regression Trees*.

Soleymani, M. and Lee S.M.S(2014). Sequential combination of weighted and nonparametric bagging for classification. *Biometrika*, 101, 2, pp. 491–498.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, 7(1):1-26.

Examples

```
data(hills, package="MASS")
rt.df <- hills[sample(nrow(hills)),]
data.train <- rt.df[1 : (length(rt.df[, 1]) - 1), ]
data.test <- rt.df[-(1 : (length(rt.df[, 1]) - 1)),]
fit <- SQBreg(data.train, data.test, reps = 30, y = "time")
fit
```

Index

SQB, [2](#)

SQB-package (SQB), [2](#)

SQBreg, [2](#)