

# Package ‘BootValidation’

November 14, 2017

**Type** Package

**Title** Adjusting for Optimism in 'glmnet' Regression using Bootstrapping

**Version** 0.1.3

**Author** Antonio Jose Canada Martinez

**Maintainer** Antonio Jose Canada Martinez <ancamar2@gmail.com>

## Description

Main objective of a predictive model is to provide accurated predictions of a new observations. Unfortunately we don't know how well the model performs. In addition, at the current era of omic data where  $p \gg n$ , is not reasonable applying internal validation using data-splitting. Under this background a good method to assessing model performance is applying internal bootstrap validation (Harrell Jr, Frank E (2015) <doi:10.1007/978-1-4757-3462-1>.) This package provides bootstrap validation for the linear and logistic 'glmnet' models.

**License** GPL (>= 2)

**Encoding** UTF-8

**LazyData** true

**Imports** glmnet, pbapply, pROC, parallel

**RdMacros** Rdpack

**RoxygenNote** 6.0.1

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2017-11-14 15:30:31 UTC

## R topics documented:

vboot	2
vboot.elnet	3
vboot.lognet	3

<b>Index</b>	<b>4</b>
--------------	----------

---

`vboot`*Generic function for bootstrap validation*

---

## Description

Validate 'glmnet' linear or logistic regression using bootstrap.

## Usage

```
vboot(glmnet_fit, x, y, s, nfolds = 5, B = 200, cv_replicates = 100,  
      n_cores = max(1, parallel::detectCores() - 1))
```

## Arguments

<code>glmnet_fit</code>	Object from glmnet fit
<code>x</code>	A matrix of the predictors, each row is an observation vector.
<code>y</code>	A vector of response variable. It can be numeric or factor with two levels
<code>s</code>	Value of the penalty parameter "lambda" selected from the original cv.glmnet
<code>nfolds</code>	Number of folds for cross validation as in cv.glmnet
<code>B</code>	Number of bootstrap samples
<code>cv_replicates</code>	Number of replicates for the cross-validation step
<code>n_cores</code>	number of cores to use in parallel. Default detectCores()-1

## References

Jerome Friedman, Trevor Hastie, Robert Tibshirani (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33(1), 1-22. URL <http://www.jstatsoft.org/v33/i01/>.

Frank Harrell (2015). Harrell Jr, F. E. (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer.

## Examples

```
# Create the data  
set.seed(25)  
x <- matrix(rnorm(80),ncol=4)  
y <- x[,4]*0.8+x[,3]*0.4+rnorm(20)  
# Fit glmnet model  
fit_enet <- glmnet::glmnet(x, y, alpha = 0.5)  
# Bootstrap validation  
vboot(fit_enet, x, y, nfolds = 3, B = 2, s = 0.5, cv_replicates = 5, n_cores = 1)
```

---

vboot.elnet	<i>Internal bootstrapping validation lineal glmnet model</i>
-------------	--

---

**Description**

Validate glmnet linear regression using bootstrap.

**Usage**

```
## S3 method for class 'elnet'
vboot(glmnet_fit, x, y, s, nfolds, B, cv_replicates, n_cores)
```

**Arguments**

glmnet_fit	Object from glmnet fit
x	A matrix of the predictors, each row is an observation vector.
y	A vector of response variable. Should be numeric
s	Value of the penalty parameter "lambda" selected from the original 'cv.glmnet'
nfolds	Number of folds for cross validation as in 'cv.glmnet'
B	Number of bootstrap samples
cv_replicates	Number of replicates for the cross-validation step
n_cores	number of cores to use in parallel. Default detectCores()-1

---

vboot.lognet	<i>Internal bootstrapping validation</i>
--------------	--

---

**Description**

Validate glmnet logistic regression using bootstrap.

**Usage**

```
## S3 method for class 'lognet'
vboot(glmnet_fit, x, y, s, nfolds, B, cv_replicates, n_cores)
```

**Arguments**

glmnet_fit	Object from glmnet fit
x	A matrix of the predictors, each row is an observation vector.
y	A vector of response variable. Should be a factor with two levels
s	Value of the penalty parameter "lambda" selected from the original 'cv.glmnet'
nfolds	Number of folds for cross validation as in cv.glmnet
B	Number of bootstrap samples
cv_replicates	Number of replicates for the cross-validation step in 'cv.glmnet'
n_cores	number of cores to use in parallel. Default detectCores()-1

# Index

vboot, [2](#)  
vboot.elnet, [3](#)  
vboot.lognet, [3](#)