Package 'forestControl'

February 13, 2018

February 13, 2018
Type Package
Title Approximate False Positive Rate Control in Selection Frequency for Random Forest
Version 0.1.1
Date 2018-02-13
Author Tom Wilson <tpw2@aber.ac.uk>, Jasen Finch <jsf9@aer.ac.uk></jsf9@aer.ac.uk></tpw2@aber.ac.uk>
Maintainer Tom Wilson <pre><pre><pre><pre><pre><pre><pre><pre></pre></pre></pre></pre></pre></pre></pre></pre>
Description Approximate False Positve Rate Control in Selection Frequecy for Random Forest using the methods described by Ender Konukoglu and Melanie Ganz. Methods for calculating the selection frequency threshold at false positive rate and selection frequency false positive rate feature selection.
Imports Rcpp
Suggests testthat, randomForest, ranger
License GPL (>= 3)
Encoding UTF-8
LazyData true
<pre>URL https://github.com/wilsontom/forestControl</pre>
<pre>BugReports https://github.com/wilsontom/forestControl/issues</pre>
RoxygenNote 6.0.1
LinkingTo Rcpp
Roxygen list(markdown = TRUE)
NeedsCompilation yes
R topics documented:
forestControl-package extract_params fpr_fs selection_freqs sft

2 extract_params

Index 5

forestControl-package False Positive Rate Control in Selection Frequency for Random Forest

Description

This package is an implementation of the methods described by Ender Konukoglu and Melanie Ganz in *Konukoglu, E. and Ganz, M., 2014. Approximate false positive rate control in selection frequency for random forest. arXiv preprint arXiv:1410.2838* https://arxiv.org/abs/1410.2838.

extract_params

Extract forest parameters

Description

For a randoForest or ranger classification object, extract the parameters needed to calculate an approximate selection frequency threshold

Usage

```
extract_params(x)
```

Arguments

Х

a randomForest or ranger object

Value

a list of four elements

Fn The number of features considered at each internal node (mtry)

Ft The total number of features in the data set

K The average number of binary tests/internal nodes across the enitre forest

Tr The total number of trees in the forest

Author(s)

fpr_fs 3

fpr_fs

False Postivie Rate Feature Selection

Description

Calculate the False Positive Rate (FPR) for each feature using it's selection frequency

Usage

```
fpr_fs(x)
```

Arguments

Х

a randomForest or ranger object

Value

a data.frame of selection frequencies and their false positive rate

Author(s)

Jasen Finch <jsf9@aber.ac.uk>

selection_freqs

Variable Selection Frequencies

Description

Extract variable selection frequencies from randomForest and ranger model objects

Usage

```
selection_freqs(x)
```

Arguments

Х

a randomForest or ranger object

Value

data.frame of variable selection frequencies

4 sft

sft

Selection Frequency Threshold

Description

Selection Frequency Threshold

Usage

```
sft(x, alpha)
```

Arguments

x a randomForest or ranger object alpha a false positive rate (ie, 0.01)

Value

a list of two elements

sft the selection frequency thresholdprobs_atsft the esimated false positive rate

Author(s)

Tom Wilson couk>

Index

```
extract_params, 2
forestControl-package, 2
fpr_fs, 3
selection_freqs, 3
sft, 4
```