Package ‘banter’

July 10, 2018

Type Package

Title BioAcoustic EVENT Classifier


Version 0.9.3

Depends magrittr

Imports dplyr, ggplot2, gridExtra, methods, parallel, plyr, randomForest, ranger, rfPermute, rlang, stats, tibble, tidyr

License GNU General Public License

RoxygenNote 6.0.1

NeedsCompilation no

Author Eric Archer [aut, cre], Taiki Sakai [aut]

Maintainer Eric Archer <eric.archer@noaa.gov>

Repository CRAN

Date/Publication 2018-07-10 15:20:06 UTC

R topics documented:

banter-package .................................................. 2
addBanterDetector ............................................. 2
getBanterModel ............................................... 3
getBanterModelData ......................................... 4
getDetectorNames ............................................ 5
getSampSize ................................................... 6
initBanterModel .............................................. 7
modelPctCorrect ............................................. 8
numCalls ..................................................... 8
numEvents ................................................... 9
plotDetectorTrace .......................................... 10
predict ...................................................... 11
runBanterModel ............................................. 12
addBanterDetector

Description

banter

Add a detector model to a BANTER classifier.

Usage

addBanterDetector(x, data, name, ntree, sampsize = 1, importance = FALSE, num.cores = NULL)

removeBanterDetector(x, name)

Arguments

x       a banter_model object.
data     detector data.frame or named list of detector data.frames. If a data.frame, then name must be provided.
name     detector name.
ntree    number of trees.
sampsize number or fraction of samples to use in each tree.
importance retain importance scores in model? Defaults to FALSE and will be ignored if num.cores > 1.
num.cores number of cores to use for Random Forest model.

Value

a banter_model object with the detector model added or removed.

Author(s)

Eric Archer <eric.archer@noaa.gov>
Examples

data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add the 'bp' (burst pulse) detector model
bant.mdl <- addBanterDetector(
    x = bant.mdl,
    data = train.data$detectors$bp,
    name = "bp",
    ntree = 50, sampsize = 1, num.cores = 1
)
bant.mdl

# remove the 'bp' detector model
bant.mdl <- removeBanterDetector(bant.mdl, "bp")
bant.mdl

def getBanterModel

Description

Extract BANTER event or detector Random Forest model.

Usage

getBanterModel(x, model = "event")

Arguments

x a banter_model object.
model name of model to extract. Default is "event" to summarize the event-level model. Can also be name of a detector.

Value

a randomForest model object.

Author(s)

Eric Archer <eric.archer@noaa.gov>
Examples

data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
    bant.mdl, train.data$detectors,
    ntree = 50, sampsize = 1, num.cores = 1
)
# run BANTER event model
bant.mdl <- runBanterModel(bant.mdl, ntree = 1000, sampsize = 1)

# extract the event randomForest model
library(randomForest)
event.rf <- getBanterModel(bant.mdl)
event.rf

# extract the burst pulse (bp) detector model
bp.rf <- getBanterModel(bant.mdl, "bp")
bp.rf

getBanterModelData  Extract Random Forest Model Data

Description

Extract BANTER event data used for the Random Forest model.

Usage

getBanterModelData(x)

Arguments

x  a banter_model object.

Value

a randomForest model object.

Author(s)

Eric Archer <eric.archer@noaa.gov>
getDetectorNames

Examples

data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
    bant.mdl, train.data$detectors,
    ntree = 50, sampsize = 1, num.cores = 1
)
# run BANTER event model
bant.mdl <- runBanterModel(bant.mdl, ntree = 1000, sampsize = 1)

event.df <- getBanterModelData(bant.mdl)
head(event.df)

class(getDetectorNames) <- c("banter_model")

getDetectorNames x

Arguments

x a banter_model object.

Value

a vector of names.

Author(s)

Eric Archer <eric.archer@noaa.gov>

Examples

data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
    bant.mdl, train.data$detectors,
    ntree = 50, sampsize = 1, num.cores = 1
)
getSampSize

getSampSize(bant.mdl)

---

### Description

Return sample size used for a BANTER model.

### Usage

`getSampSize(x, model = "event")`

### Arguments

- `x` : a `banter_model` object.
- `model` : name of model to extract. Default is "event" to summarize the event-level model. Can also be name of a detector.

### Value

a vector of sample sizes.

### Author(s)

Eric Archer <eric.archer@noaa.gov>

### Examples

```r
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
  bant.mdl, train.data$detectors,
  ntree = 50, sampsize = 2, num.cores = 1
)
# run BANTER event model
bant.mdl <- runBanterModel(bant.mdl, ntree = 1000, sampsize = 1)

# sample size for the event model
getsampSize(bant.mdl)

# sample size for the burst pulse (bp) detector model
getsampSize(bant.mdl, "bp")
```
Description

Initialize a BANTER model with event data.

Usage

`initBanterModel(x)`

Arguments

- `x` a data.frame of events. Every row is a unique event. Must have columns named `event.id` and `species`. All other columns will be used as predictor variables for the BANTER event classifier model.

Value

- a `banter_model` object without any detector models.

Note

Values in the column `species` are passed through the `make.names` function on creation to ensure they don’t include invalid characters.

Author(s)

Eric Archer <eric.archer@noaa.gov>

Examples

```r
data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
bant.mdl
```
numCalls

Number and Proportion of Calls

Description

Number and proportion of calls in BANTER detector models.

modelPctCorrect

Model Percent Correct

Description

Extract percent correctly classified by species for detector and event models.

Usage

modelPctCorrect(x)

Arguments

x

a banter_model object.

Value

a data.frame with the percent correctly classified for each model in x.

Author(s)

Eric Archer <eric.archer@noaa.gov>

Examples

data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
  bant.mdl, train.data$detectors,
  ntree = 50, sampsize = 1, num.cores = 1
)
# run BANTER event model
bant.mdl <- runBanterModel(bant.mdl, ntree = 1000, sampsize = 1)
modelPctCorrect(bant.mdl)
numEvents

Usage

numCalls(x, by = c("species", "event"))

propCalls(x, by = c("species", "event"))

Arguments

x               a banter_model object.
by              return summary by "species" or "event".

Author(s)

Eric Archer <eric.archer@noaa.gov>

Examples

data(train.data)
# initialize BANTER model with event data
bant.mdl1 <- initBanterModel(train.data$events)
# add all detector models
bant.mdl1 <- addBanterDetector(
    bant.mdl1, train.data$detectors,
    ntree = 50, sampsize = 1, num.cores = 1
)
# run BANTER event model
bant.mdl1 <- runBanterModel(bant.mdl1, ntree = 1000, sampsize = 1)

# number of calls by species and event
numCalls(bant.mdl1, "species")
numCalls(bant.mdl1, "event")

# proportion of calls by species and event
propCalls(bant.mdl1, "species")
propCalls(bant.mdl1, "event")

-----------

numEvents Number of Events
-----------

Description

Number of events in BANTER model by species.

Usage

numEvents(x, model = "event")
Arguments

x  a banter_model object.
model  name of model to extract. Default is "event" to summarize the event-level model. Can also be name of a detector.

Value

a data.frame giving the number of events available for each species.

Author(s)

Eric Archer <eric.archer@noaa.gov>

Examples

data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
  bant.mdl, train.data$detectors,
  ntree = 50, sampsize = 1, num.cores = 1
)
# run BANTER event model
bant.mdl <- runBanterModel(bant.mdl, ntree = 1000, sampsize = 1)

# number of events in event model
numEvents(bant.mdl)

# number of events in burst pulse (bp) detector model
numEvents(bant.mdl, "bp")
**predict**

**Predict BANTER events**

**Description**

Predict species of events for novel data from a BANTER model.

**Usage**

```r
## S3 method for class 'banter_model'
predict(object, new.data, ...)
```

**Arguments**

- `object` a `banter_model` object.
- `new.data` a list of event and detector data that has the same predictors as in the `banter_model`. It must contain elements called `events` and `detectors`. The `events` element must be a data.frame that has a column called `event.id` and the same predictor columns as the event data used to initialize the banter model (see `initBanterModel`). The `detectors` element must be a named list with the same detectors used to build the model (see `addBanterDetector`).
- `...` unused.

**Examples**

```r
data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
  bant.mdl, train.data$detectors,
  ntree = 50, sampsize = 1, num.cores = 1
)

plotDetectorTrace(bant.mdl)
```
Value

A list with the following elements:

- **events** the data frame used in the event model for predictions.
- **predict.df** data.frame of predicted species and assignment probabilities for each event.

Author(s)

Eric Archer <eric.archer@noaa.gov>

Examples

data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
  bant.mdl, train.data$detectors,
  ntree = 50, sampsize = 2, num.cores = 1
)
# run BANTER event model
bant.mdl <- runBanterModel(bant.mdl, ntree = 1000, sampsize = 1)

# predict test data
data(test.data)
test.pred <- predict(bant.mdl, test.data)
test.pred

---

**runBanterModel**

**Run BANTER Model**

Description

Build full event classifier model

Usage

runBanterModel(x, ntree, sampsize = 1)

Arguments

- **x** a banter_model object.
- **ntree** number of trees.
- **sampsize** number or fraction of samples to use in each tree.

Value

a banter_model object with the complete BANTER model.
Examples

```r
data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
    bant.mdl, train.data$detectors,
    ntree = 50, sampsize = 1, num.cores = 1
)
# run BANTER event model
bant.mdl <- runBanterModel(bant.mdl, ntree = 1000, sampsize = 1)
summary(bant.mdl)
```

---

**BANTER Classifier Model Summary**

**Description**

Display summaries for event and detector models

**Usage**

```r
## S3 method for class 'banter_model'
summary(object, model = "event", n = 0.1, bins = 20, ...)
```

**Arguments**

- `object`: a `banter_model` object.
- `model`: name of model to summarize. Default is "event" to summarize the event-level model. Can also be name of a detector.
- `n`: number of final iterations to summarize OOB error rate for. If between 0 and 1 is taken as a proportion of chain.
- `bins`: number of bins in inbag histogram.
- `...`: ignored.

**Value**

In the plot that is created, the upper panel shows the trace of the Random Forest model OOB rate across sequential trees in the forest. The lower plot shows a frequency histogram of the number of times each sample was inbag (used as training data in a tree in the forest). The vertical red lines indicate the expected inbag rate for samples of each species.
Author(s)

Eric Archer <eric.archer@noaa.gov>

Examples

data(train.data)
# initialize BANTER model with event data
bant.mdl <- initBanterModel(train.data$events)
# add all detector models
bant.mdl <- addBanterDetector(
  bant.mdl, train.data$detectors,
  ntree = 50, sampsize = 1, num.cores = 1
)
# run BANTER event model
bant.mdl <- runBanterModel(bant.mdl, ntree = 1000, sampsize = 1)
summary(bant.mdl)

Description

A list of events and call data from detectors for testing BANTER model

Usage

data(test.data)

Format

list

train.data

Training events and detectors

Description

A list of events and call data from detectors for training BANTER model

Usage

data(train.data)

Format

list
Index

*Topic **datasets**
  - test.data, 14
  - train.data, 14

*Topic **package**
  - banter-package, 2

  addBanterDetector, 2, 11

  banter (banter-package), 2
  banter-package, 2
  banter_model, 2–13

  getBanterModel, 3
  getBanterModelData, 4
  getDetectorNames, 5
  getSampSize, 6

  initBanterModel, 7, 11

  make.names, 7
  modelPctCorrect, 8

  numCalls, 8
  numEvents, 9

  plotDetectorTrace, 10
  plotRFtrace, 11

  predict, 11
  predict, banter_model-method (predict), 11
  predict.banter_model (predict), 11
  propCalls (numCalls), 8

  randomForest, 3, 4
  removeBanterDetector
    (addBanterDetector), 2
  runBanterModel, 12

  summary, 13
  summary, banter_model-method (summary), 13