Package ‘aCRM’

February 19, 2015

Type Package

Title Convenience functions for analytical Customer Relationship Management

Version 0.1.1

Date 2014-03-28

Imports dummies, randomForest, kernelFactory, ada

Author Dirk Van den Poel, Michel Ballings, Andrey Volkov, Jeroen D’haen, Michiel Van Herwegen

Maintainer Michel Ballings <michel BALLINGS@GMail.com>

Description Convenience functions for data preparation and modeling often used in aCRM.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2014-03-28 11:56:43

R topics documented:

aCRM-package ................................................................. 2
aCRMNews ................................................................. 2
Aggregate ................................................................. 3
cocktailEnsemble ........................................................ 4
Credit ................................................................. 6
imputeMissings ............................................................ 6
predict.cocktailEnsemble ............................................. 7

Index 9
Description

This package provides functions that are often used in aCRM such as missing values imputation, aggregation and dummy creation, and ensemble modeling.

Details

| Package: | aCRM |
| Type: | Package |
| Version: | 0.1.0 |
| Date: | 2013-03-24 |
| License: | GPL (>= 2) |

Currently provided functions are: `imputeMissings`, `Aggregate`, `cocktailEnsemble`, `predict.cocktailEnsemble`

Author(s)

Dirk Van den Poel, Michel Ballings, Andrey Volkov, Jeroen D"haen, Michiel Van Herwegen
Maintainer: Michel Ballings <Michel.Ballings@GMail.com>

References


See Also

`imputeMissings`, `Aggregate`, `cocktailEnsemble`, `predict.cocktailEnsemble`

Description

`aCRMNews` shows the NEWS file of the aCRM package.

Usage

`aCRMNews()`
Aggregate

Value
None.

---

**Aggregate**

**Aggregate numeric and categorical variables by an ID**

**Description**

The *Aggregate* function (not to be confounded with *aggregate*) prepares a data frame for merging by computing the sum, mean and variance of all continuous (integer and numeric) variables by a given ID variable. It also creates dummies for all categorical variables (character and factor) and subsequently computes the sum by a given ID variable. This function aims at maximal information extraction with a minimal amount of code.

**Usage**

*Aggregate*(x, by)

**Arguments**

- **x**: A data frame without the ID. Categorical variables have to be of type character or factor and continuous variables have to be of type integer or numeric.
- **by**: A vector containing ID’s.

**Value**

A data frame with the aforementioned variables aggregated by the given ID variables.

**Author(s)**

Dirk Van den Poel, Michel Ballings, Andrey Volkov, Jeroen D’haen, Michiel Van Herwegen

Maintainer: Michel Ballings <Michel.Ballings@GMail.com>

**References**


**See Also**

Other functions in this package: imputeMissings, Aggregate, cocktailEnsemble, predict.cocktailEnsemble
Examples

# Create some data
data <- data.frame(V1=as.factor(c('yes','no','yes','yes','yes','no')),
                   V2=as.character(c(1,2,3,4,4,4)),
                   V3=as.character(c(1,1,1,2,2,2)),
                   ID=as.character(c(1,1,1,2,2,2)),
# Demonstrate function
Aggregate(x=data,by=ID)

cocktailEnsemble  Cocktail Ensemble: build a model consisting of multiple classifiers.

Description

cocktailEnsemble is a classification algorithm. It builds four models by calling glm (logit),
kernelfactory, randomforest, and ada.

Usage

cocktailEnsemble(x, y)

Arguments

x  A data frame containing the predictors.
y  The response vector.

Value

An object of type cocktailEnsemble containing the four aforementioned models.

Author(s)

Dirk Van den Poel, Michel Ballings, Andrey Volkov, Jeroen D’haen, Michiel Van Herwegen
Maintainer: Michel Ballings <Michel.Ballings@GMail.com>

References

Van den Poel, D., Ballings, M., Volkov, A., D’haen, J., Vanherwegen, M., Predictive Analytics for

glm:

  Hall.
  Hall.

randomForest:

kernelFactory:

ada:

See Also
Other functions in this package: imputeMissings, Aggregate, cocktail Ensemble, predict.cocktail Ensemble

Examples

```r
#Credit Approval data available at UCI Machine Learning Repository
data(Credit)

#Create training set (take a small subset for demonstration purposes)
Credit <- data.frame(Credit[order(runif(nrow(Credit))),][1:100, c('V2', 'V3', 'V6', 'V11', 'V14', 'V15', 'Response')])
trainingset <- Credit[1:1:floor(0.50*nrow(Credit)),]

#Create test set
testset <- Credit[(floor(0.50*nrow(Credit))+1):nrow(Credit),]

#Train Cocktail Ensemble on training data
cE <- cocktailEnsemble(x=trainingset[,names(trainingset)!="Response"], y=trainingset$Response)

#Deploy Kernel Factory to predict response for test data
#pred <- predict(cE, testset[,names(testset)!="Response"])
```
Credit approval (Frank and Asuncion, 2010)

Description
Credit contains credit card applications. The dataset has a good mix of continuous and categorical features.

Usage
data(Credit)

Format
A data frame with 653 observations, 15 predictors and a binary criterion variable called Response

Details
All observations with missing values are deleted.

Source

References
The original dataset can be downloaded at http://archive.ics.uci.edu/ml/datasets/Credit+Approval

Examples
data(Credit)
str(Credit)
table(Credit$Response)

imputeMissings
Impute missing values with the median or mode.

Description
Character vectors and factors are imputed with the mode. Numeric and integer vectors are imputed with the median.

Usage
imputeMissings(data)
predict.cocktailEnsemble

Arguments

data  A data frame.

Value

A data frame.

Author(s)

Dirk Van den Poel, Michel Ballings, Andrey Volkov, Jeroen D’haen, Michiel Van Herwegen
Maintainer: Michel Ballings <Michel.Ballings@GMail.com>

References


See Also

Other functions in this package: imputeMissings, Aggregate, cocktailEnsemble, predict.cocktailEnsemble

Examples

# create some data
data <- data.frame(V1=as.factor(c('yes','no','no',NA,'yes','yes','yes')), V2=as.character(c(1,2,3,4,4,NA)), V3=c(1:6,NA), V4=as.numeric(c(1:6,NA)))

# demonstrate function
imputeMissings(data)

predict.cocktailEnsemble

*Predict method for cocktailEnsemble objects.*

Description

Prediction of new data using cocktailEnsemble. predict combines the predictions from all models contained in the cocktail by taking the mean of the probabilities.

Usage

## S3 method for class 'cocktailEnsemble'
predict(object, newdata=NULL,...)
Arguments

object  An object of class cocktailEnsemble, as created by the function cocktailEnsemble.
newdata  A data frame with the same predictors as in the training data.
...  Not used currently.

Value

A vector containing the probabilities of the predicted event.

Author(s)

Dirk Van den Poel, Michel Ballings, Andrey Volkov, Jeroen D’haen, Michiel Van Herwegen
Maintainer: Michel Ballings <Michel.Ballings@GMail.com>

References


See Also

Other functions in this package: imputeMissings, Aggregate, cocktailEnsemble, predict.cocktailEnsemble

Examples

#Credit Approval data available at UCI Machine Learning Repository
data(Credit)

#Create training set (take a small subset for demonstration purposes)
Credit <- data.frame(Credit[order(runif(nrow(Credit))),][1:100, c('V2','V3','V8','V11','V14','V15','Response')]
trainingset <- Credit[1:1:floor(0.50*nrow(Credit)),]
#Create test set
testset <- Credit[(floor(0.50*nrow(Credit))+1):nrow(Credit),]

#Train Cocktail Ensemble on training data
cE <- cocktailEnsemble(x=trainingset[,names(trainingset)!= "Response"],y=trainingset$Response)

#Deploy Kernel Factory to predict response for test data
pred <- predict(cE,testset[,names(testset)!= "Response"])

Index

*Topic **Data Manipulation**
   - Aggregate, 3
   - imputeMissings, 6

*Topic **classification**
   - cocktailEnsemble, 4
   - predict.cocktailEnsemble, 7

*Topic **datasets**
   - Credit, 6

*Topic **package**
   - aCRM-package, 2

   aCRM (aCRM-package), 2
   aCRM-package, 2
   aCRMNews, 2
   Aggregate, 2, 3, 3, 5, 7, 8
   cocktailEnsemble, 2, 3, 4, 5, 7, 8
   Credit, 6
   imputeMissings, 2, 3, 5, 6, 7, 8
   predict.cocktailEnsemble, 2, 3, 5, 7, 7, 8