

Package ‘PLMplusPMMLpkg’

May 14, 2015

Type Package

Title Read json files, create/predict lm, save in rds,pmml

Version 1.0

Date 2015-03-12

Author Georgia Tsiliki

Maintainer Georgia Tsiliki <gtsiliki@central.ntua.gr>

Depends RCurl, pmml, jsonlite

Description

This package demonstrates how to build/predict lm models and save in rds and pmml format

License GPL-2

R topics documented:

| | |
|----------------------------------|----|
| PLMplusPMMLpkg-package | 1 |
| dat1 | 2 |
| dat1i | 3 |
| dat1m | 4 |
| dat1p | 5 |
| lm.funct | 6 |
| pred.funct | 7 |
| read.in.json | 8 |
| read.in.json.for.pred | 10 |

| | |
|--------------|-----------|
| Index | 12 |
|--------------|-----------|

PLMplusPMMLpkg-package

Reads in json files, create/predict lm, save in serialized and pmml format

Description

This package demonstrates how to build/predict lm models and save in serialized and pmml format

Details

Package: PLMplusPMMLpkg
Type: Package
Version: 1.0
Date: 2015-03-12
License: GPL-2

Author(s)

Georgia Tsiliki

Maintainer: Georgia Tsiliki <gtsiliki@central.ntua.gr>

References

Package source is in the help file for package.skeleton

Examples

```
data("dat1p")
data("dat1i")
data("dat1m")

data.file<- read.in.json.for.pred(dat1p, dat1m, dat1i)
```

| | |
|------|-----------------------------|
| dat1 | <i>A sample data object</i> |
|------|-----------------------------|

Description

The dataset for this test is a data frame

Usage

```
data("dat1")
```

Format

A list of two objects

datasetURI a character vector- ambit data set uri

dataEntry a data frame containing two columns: compound and values. Compound is a character vector with all compound anbit uris, and calues is a data frame with all numeric values of the data set (compounds by features)

Details

There are no more details

Source

The source of this function is in the

References

There are no references

Examples

```
data(dat1)
## maybe str(dat1) ; plot(dat1) ...
```

| | |
|-------|--|
| dat1i | <i>A list including names of independent variables in model dat1m and predicted feature's name</i> |
|-------|--|

Description

For lm model: A list including a data frame of two columns, the 1st named 'ModelCoef' giving the dummy coefficient names produced for independent features in the model, and 2nd 'RealFeatureNames' which are ambit's feature names. Also included is the PredictedFeatures, a character string indicating the name of the predicted feature.

Usage

```
data("dat1i")
```

Format

A list with two objects.

indepFeatures data frame with 4 observations on the following 2 variables: ModelCoef, RealFeatureNames. The first is a character vector, giving the dummy coefficient names produced for independent features in the model, and the second a character vector, which are ambit's feature names (uris).

predictedFeatures A character string giving the predicted features name (ambit uri plus the word 'Predicted').

Details

Example dataset to predict with dat1m

Source

The source of this function is in the

References

There are no references

Examples

```
data(dat1i)
## maybe str(dat1i) ; plot(dat1i) ...
```

| | |
|-------|---------------------------------|
| dat1m | <i>Serialized lm model file</i> |
|-------|---------------------------------|

Description

A character string for a serialized lm model

Usage

```
data("dat1m")
```

Format

A character string

Details

Example lm model based on dat1

Source

The source of this function is in the

References

There are no references

Examples

```
data(dat1m)
## maybe str(dat1m) ; plot(dat1m) ...
```

| | |
|-------|-----------------------------|
| dat1p | <i>A sample data object</i> |
|-------|-----------------------------|

Description

The dataset for this test is a data frame

Usage

```
data("dat1p")
```

Format

A list of two objects

datasetURI a character vector- ambit data set uri

dataEntry a data frame containing two columns: compound and values. Compound is a character vector with all compound anbit uris, and values is a data frame with all numeric values of the data set (compounds by features)

Details

Data set for prediction with dat1m

Source

The source of this function is in the

References

There are no references

Examples

```
data(dat1p)
## maybe str(dat1p) ; plot(dat1p) ...
```

| | |
|----------|---------------------|
| lm.funct | <i>Linear model</i> |
|----------|---------------------|

Description

lm.funct perform lm and is expecting data in json format

Usage

lm.funct(dataset, predictionFeature, parameters)

Arguments

- dataset list of 2 objects, datasetURI:= character string, code name of dataset, dataEntry:= data frame with 2 columns
- predictionFeature character string specifying which is the prediction feature in dataEntry
- parameters list with parameter values

Details

No details required

Value

- A List
 - rawModel A serialized lm object is (class raw)
 - pmmlModel A pmml lm object
 - predictedFeatures
 - A character string givingth the predicted feature name (ambit)
 - independentFeatures
 - A list with all independent features included in the lm model
 - additionalInfo For lm: A list including a data frame with all independent features in the lm model and their dummy name in the model, and predicted feature’s name

Note

No notes for this function

Author(s)

Georgia Tsiliki

References

The help file of skeleton

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==> Define data, use random,
##-- or do help(data=index) for the standard data sets.

data("dat1")

x1<- "https://apps.ideaconsult.net/ambit2/feature/22137"

x2<- list('theta'=1,'lambda'=2)

lm.mod<- lm.funct(dat1,x1,x2)
```

pred.funct

lm prediction

Description

predict.lm function expecting list from fromJSON

Usage

```
pred.funct(dataset, rawModel, additionalInfo)
```

Arguments

| | |
|----------------|---|
| dataset | Data for prediction. A list of two objects: datasetURI (a character string), dataEntry (a data frame). |
| rawModel | R model serialized. |
| additionalInfo | Any additional information needed for rawModel. For lm: A list including a data frame showing the independent variables of the original data included in the model, and a character vector showing the predicted feature's name (as produced by lm.funct). The data frame consists of two columns, the 1st named 'ModelCoef' giving the dummy coefficient names produced for independent features in the model, and 2nd 'RealFeatureNames' which are ambit's feature names. |

Details

The function returns an error if the supplied dataset does not include all the independent features in the raw model.

Value

A data frame for prediction vector

Note

No notes for this function

Author(s)

Georgia Tsiliki

References

the help file of skeleton

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==> Define data, use random,
##--or do help(data=index) for the standard data sets.

data("dat1p")
data("dat1i")
data("dat1m")

pred.res<- pred.funct(dat1p, dat1m, dat1i)
```

read.in.json

Read in function for json files

Description

This function reads in a json data file and produces a list with y and x variables

Usage

```
read.in.json(dataset, predictionFeature, parameters)
```

Arguments

| | |
|-------------------|---|
| dataset | list of 2 objects, datasetURI:= character string, code name of dataset, dataEntry:= data frame with 2 columns |
| predictionFeature | character string specifying which is the prediction feature in dataEntry |
| parameters | list with parameter values |

Details

json file includes uris, data, indicator of y variable, model's parameters list

Value

A list is returned

| | |
|-------------------------|---|
| <code>x.mat</code> | data frame with independent variables |
| <code>y.mat</code> | singular data frame with dependent variable |
| <code>data.names</code> | names of independent features |
| <code>par.list</code> | list of parameters |

Note

No notes for this function

Author(s)

Georgia Tsiliki

References

The help file of skeleton

Examples

```
##---- Should be DIRECTLY executable !! ----
##-- ==> Define data, use random,
##-- or do help(data=index) for the standard data sets.

data("dat1")

x1<- "https://apps.ideaconsult.net/ambit2/feature/22137"

x2<- list('theta'=1,'lambda'=2)

r.dat1<- read.in.json(dat1,x1,x2)
```

`read.in.json.for.pred` *Read in function for json files for prediction*

Description

This function reads in a json data file and produces a list with y, x variables, and additionalInfo

Usage

```
read.in.json.for.pred(dataset, rawModel, additionalInfo)
```

Arguments

| | |
|-----------------------------|---|
| <code>dataset</code> | Data for prediction. A list of two objects: <code>datasetURI</code> (a character string), <code>dataEntry</code> (a data frame). |
| <code>rawModel</code> | R model serialized. |
| <code>additionalInfo</code> | Any additional information needed for <code>rawModel</code> . For <code>lm</code> : A list including a data frame showing the independent variables of the original data included in the model, and a character vector showing the predicted feature's name (as produced by <code>lm.funct</code>). The data frame consists of two columns, the 1st named 'ModelCoef' giving the dummy coefficient names produced for independent features in the model, and 2nd 'RealFeatureNames' which are ambit's feature names. |

Details

The function returns an error if the supplied dataset does not include all the independent features in the raw model.

Value

A List including:

| | |
|-----------------------------|---|
| <code>x.mat</code> | data frame with independent variables |
| <code>model</code> | R model |
| <code>additionalInfo</code> | Any additional information needed for <code>rawModel</code> . For <code>lm</code> : the dependent variables included in the model and also the new data set, and the predicted feature's name |

Note

No notes for this function

Author(s)

Georgia Tsiliki

References

The help file of skeleton

Examples

```
##---- Should be DIRECTLY executable !! ----  
##-- ==> Define data, use random,  
##-- or do help(data=index) for the standard data sets.  
  
data("dat1p")  
data("dat1i")  
data("dat1m")  
  
data.file<- read.in.json.for.pred(dat1p, dat1m, dat1i)
```

Index

- *Topic **datasetsi**
 - dat1i, [3](#)
- *Topic **datasetsm**
 - dat1m, [4](#)
- *Topic **datasetsp**
 - dat1p, [5](#)
- *Topic **datasets**
 - dat1, [2](#)
- *Topic **package**
 - PLMplusPMMLpkg-package, [1](#)
- *Topic **pmmlLM**
 - lm.funct, [6](#)
- *Topic **predlmjson**
 - pred.funct, [7](#)
- *Topic **readjsonp**
 - read.in.json.for.pred, [10](#)
- *Topic **readjson**
 - read.in.json, [8](#)

- dat1, [2](#)
- dat1i, [3](#)
- dat1m, [4](#)
- dat1p, [5](#)

- lm.funct, [6](#)

- PLMplusPMMLpkg
 - (PLMplusPMMLpkg-package), [1](#)
- PLMplusPMMLpkg-package, [1](#)
- pred.funct, [7](#)

- read.in.json, [8](#)
- read.in.json.for.pred, [10](#)