

Package ‘rpart.utils’

May 19, 2014

Title Tools for parsing and manipulating rpart objects, including generating machine readable rules.

Version 0.5

Author Craig Varrichio <canthony427@gmail.com>

Maintainer Craig Varrichio <canthony427@gmail.com>

Description This package contains additional tools for working with rpart objects. Most importantly, it includes methods for converting rpart rules into a series of structured tables sufficient for executing the decision tree completely in SQL.

Depends R (>= 3.1.0),rpart

Suggests RODBC

License GPL-3

LazyData true

R topics documented:

rpart.lists	1
rpart.rules	2
rpart.rules.push	2
rpart.rules.table	4
rpart.subrules.table	5
Index	6

rpart.lists	<i>Creates lists of variable values (factor levels) associated with each rule in an rpart object.</i>
-------------	--

Description

Creates lists of variable values (factor levels) associated with each rule in an **rpart** object.

Usage

rpart.lists(object)

Arguments

object an rpart object

Value

a list of lists

Examples

```
library(rpart)
fit<-rpart(Reliability~.,data=car90)
rpart.lists(fit)
```

rpart.rules	<i>Returns a list of strings summarizing the branch path to each node.</i>
-------------	--

Description

Returns a list of strings summarizing the branch path to each node.

Usage

```
rpart.rules(object)
```

Arguments

object an rpart object

Examples

```
library(rpart)
fit<-rpart(Reliability~.,data=car90)
rpart.rules(fit)
```

rpart.rules.push	<i>Writes rule tables required to process rpart rules in SQL to an open RODBC connection.</i>
------------------	---

Description

This function handles the process of pushing tabular versions of **rpart** rules to an RODBC connection. The entire process of generation and writing is completed with a single call, with all necessary subcalls handled within this function.

Usage

```
rpart.rules.push(object, connection, rulePrefix = "", tablePrefix = "")
```

Arguments

object	an rpart object
connection	and open RODBC connection
rulePrefix	A character string to prepend to each rule name to allow for multiple rule sets
tablePrefix	A character string to prepend to each table name to allow for multiple rule sets

Details

Once the tables have been pushed to the database, unpivoted source data can be processed using the rpart model with SQL code similar to the following:

```
WITH SOURCE AS
(
  SELECT
    ID,
    TYPE,
    VALUE
  FROM DATA
  UNPIVOT
  (
    VALUE FOR TYPE IN (FIELD1, FIELD2, FIELD3)
  )UNPVT
),
MATCHES AS
(
  SELECT
    ID
  ,Subrule
  ,Variable
  ,SR.Value
  ,Less
  ,Greater
  FROM
    SOURCE S
  LEFT JOIN SUBRULES SR
  ON
    TYPE = VARIABLE
  AND (
    S.value = SR.value
    OR S.value < SR.Less
    OR S.value > SR.Greater
  )
),
MATCHED_SUBRULES
AS (
  SELECT
    Subrule
  ,ID
  FROM
    MATCHES M
  GROUP BY
```

```

        Subrule
        ,ID
        ),
    MATCHED_RULES
    AS (
        SELECT
            R.[Rule]
            ,MS.*
        FROM
            RULES AS R
        LEFT JOIN MATCHED_SUBRULES MS
        ON R.SUBRULE=MS.SUBRULE AND Leaf='TRUE '
    )
    ,
    COUNTS AS
    (
        SELECT
            [RULE]
            ,ID
            ,MATCH_COUNT=COUNT(DISTINCT SUBRULE)
            ,NEEDED_COUNT=(SELECT COUNT(DISTINCT SUBRULE) FROM RULES R WHERE R.[RULE]=MR.[RULE])
        FROM
            MATCHED_RULES MR
        GROUP BY
            [RULE]
            ,ID
    )
    SELECT
        RULE
        ,ID
    FROM COUNTS
    WHERE
        MATCH_COUNT=NEEDED_COUNT

```

The frame is also passed to the database which allows extracting the estimates generated by the rpart model.

rpart.rules.table	<i>Returns an unpivoted table of branch paths (subrules) associated with each node.</i>
-------------------	---

Description

Returns an unpivoted table of branch paths (subrules) associated with each node.

Usage

```
rpart.rules.table(object)
```

Arguments

object	an rpart object
--------	-----------------

Examples

```
library(rpart)
fit<-rpart(Reliability~.,data=car90)
rpart.rules.table(fit)
```

rpart.subrules.table	<i>Returns an unpivoted table of variable values (factor levels) associated with each branch.</i>
----------------------	---

Description

Returns an unpivoted table of variable values (factor levels) associated with each branch.

Usage

```
rpart.subrules.table(object)
```

Arguments

object	an rpart object
--------	-----------------

Examples

```
library(rpart)
fit<-rpart(Reliability~.,data=car90)
rpart.subrules.table(fit)
```

Index

rpart.lists, [1](#)
rpart.rules, [2](#)
rpart.rules.push, [2](#)
rpart.rules.table, [4](#)
rpart.subrules.table, [5](#)