Package 'tashu'

September 4, 2019

Type Package

Title Analysis and Prediction of Bicycle Rental Amount

Version 0.1.0

Maintainer Jiwon Min <miw5281@gmail.com>

Description Provides functions for analyzing citizens' bicycle usage pattern and predicting rental amount on specific conditions.
Functions on this package interacts with data on 'tashudata' package, a 'drat' repository.
'tashudata' package contains rental/return history on public bicycle system('Tashu'), weather for 3 years and bicycle station information.
To install this data package, see the instructions at <https://github.com/zeee1/Tashu_Rpackage>.
top10_stations(), top10_paths() function visualizes image showing the most used top 10 stations and paths.
daily_bike_rental() and monthly_bike_rental() shows daily, monthly amount of bicycle rental.
create_train_dataset(), create_test_dataset() is data processing function for prediction.
Bicycle rental history from 2013 to 2014 is used to create training dataset and that on 2015 is for test dataset.
Users can make random-forest prediction model by using create_train_model()
and predict amount of bicycle rental in 2015 by using predict_bike_rental().

License GPL (>= 2)

Encoding UTF-8

LazyData true

Imports ggplot2, lubridate, dplyr, randomForest, plyr, reshape2, RColorBrewer, drat

Suggests knitr, rmarkdown, tashudata

Additional_repositories https://zeee1.github.io/drat

VignetteBuilder knitr

RoxygenNote 6.1.1

Depends R (>= 3.5.0)

NeedsCompilation no

Author Jiwon Min [aut, cre]

Repository CRAN

Date/Publication 2019-09-04 09:30:02 UTC

R topics documented:

`create_test_dataset` *Create training dataset on specific station for prediction*

Description

A function to create training dataset on 'station_number' bicycle station by preprocessing bicycle rental history and weather data from 2013 to 2014.

Usage

```
create_test_dataset(station_number)
```

Arguments

`station_number`	number that means the number of each station.(1 ~ 144)

Value

a dataset containing feature and rental count data on 'station_number' station, 2013 ~ 2014

Examples

```
## Not run: test_dataset <- create_test_dataset(1)
```

create_train_dataset *Create training dataset on specific station for prediction*

Description

A function to create training dataset on 'station_number' bicycle station by preprocessing bicycle rental history and weather data from 2013 to 2014.

Usage

```
create_train_dataset(station_number)
```

Arguments

station_number	number that means the number of each station.(1 ~ 144)

Value

a dataset containing feature and rental count data on 'station_number' station, 2013 ~ 2014

Examples

```
## Not run: train_dataset <- create_train_dataset(1)
```

create_train_model *Create random-forest training model for bicycle rental prediction.*

Description

Create random-forest training model for bicycle rental prediction.

Usage

```
create_train_model(train_dataset)
```

Arguments

train_dataset	Training dataset created by create_train_dataset()

Value

random forest training model

Examples

```
## Not run: train_dataset <- create_train_dataset(3)
rf_model <- create_train_model(train_dataset)
## End(Not run)
```

daily_bicycle_rental *Visualize amount of bicycle rental at each day of week.*

Description

A function analyzing bicycle rental pattern on each day of week and visualizing analyzed result.

Usage

```
daily_bicycle_rental()
```

Examples

```
## Not run: daily_bicycle_rental()
```

extract_features *Extract feature columns from train/test dataset*

Description

Extract feature columns from train/test dataset

Usage

```
extract_features(data)
```

Arguments

data	data with feature columns and others

Value

data containing only feature columns

monthly_bicycle_rental

Visualize the change of bicycle rental amount by temperature and each month.

Description

A function drawing a plot that shows change of temperature and bicycle rental ratio in each month.

Usage

```
monthly_bicycle_rental()
```

Examples

```
## Not run: monthly_bicycle_rental()
```

predict_bicycle_rental

Predict hourly Demand of bicycle in 2015.

Description

predict hourly amount of bicycle rental in 2015 using random forest algorithm. Create prediction model using 'train_dataset' and forecast demand of bicycle rental according to the condition of 'test_dataset'

Usage

```
predict_bicycle_rental(rf_model, test_dataset)
```

Arguments

rf_model	random forest prediction model create by create_train_model()
test_dataset	testing dataset

Value

test_dataset with predictive result.

Examples

```
## Not run: train_dataset <- create_train_dataset(3)
test_dataset <- create_test_dataset(3)
rf_model <- create_train_model(train_dataset)
test_dataset <- predict_bicycle_rental(rf_model, test_dataset)
## End(Not run)
```

top10_paths *Visualize Top 10 Pathes that were most used from 2013 to 2015.*

Description

Visualize Top 10 Pathes that were most used from 2013 to 2015.

Usage

```
top10_paths()
```

Examples

```
## Not run: top10_paths()
```

top10_stations *Visualize top 10 stations that were most used from 2013 to 2015.*

Description

Draw a plot that visualized most used top 10 stations on barchart.

Usage

```
top10_stations()
```

Value

Data frame that contains top 10 most used stations from 2013 to 2015

Examples

```
## Not run: top10_stations()
```

Index