

Package ‘spooky’

April 13, 2022

Type Package

Title Time Feature Extrapolation Using Spectral Analysis and Jack-Knife Resampling

Version 1.1.0

Author Giancarlo Vercellino

Maintainer Giancarlo Vercellino <giancarlo.vercellino@gmail.com>

Description Proposes application of spectral analysis and jack-knife resampling for multivariate sequence forecasting. The application allows for a fast random search in a compact space of hyper-parameters composed by Sequence Length and Jack-Knife Leave-N-Out.

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Depends R (>= 3.6)

Imports purrr (>= 0.3.4), ggplot2 (>= 3.3.5), readr (>= 2.1.2), lubridate (>= 1.7.10), imputeTS (>= 3.2), fANCOVA (>= 0.6-1), scales (>= 1.1.1), tictoc (>= 1.0.1), modeest (>= 2.4.0), moments (>= 0.14), greybox (>= 1.0.1)

URL https://rpubs.com/giancarlo_vercellino/spooky

NeedsCompilation no

Repository CRAN

Date/Publication 2022-04-13 07:30:02 UTC

R topics documented:

spooky	2
time_features	3
windower	4

Index

spooky	<i>spooky</i>
--------	---------------

Description

Automatic jack-knife of spectral analysis for time feature extrapolation

Usage

```
spooky(
  df,
  seq_len = NULL,
  lno = NULL,
  n_samp = 30,
  n_windows = 3,
  ci = 0.8,
  smoother = FALSE,
  dates = NULL,
  error_scale = "naive",
  error_benchmark = "naive",
  seed = 42
)
```

Arguments

df	A data frame with time features on columns
seq_len	Positive integer. Time-step number of the forecasting sequence. Default: NULL (automatic selection between 1 and the square root of full length).
lno	Positive integer. Number of data points to leave out for resampling (using jack-knife approach). Default: NULL (automatic selection between 1 and the square root of full length).
n_samp	Positive integer. Number of samples for random search. Default: 30.
n_windows	Positive integer. Number of validation windows to test prediction error. Default: 10.
ci	Confidence interval for prediction. Default: 0.8
smoother	Logical. Flag to TRUE for loess smoothing. Default: FALSE.
dates	Date. Vector with dates for time features.
error_scale	String. Scale for the scaled error metrics. Two options: "naive" (average of naive one-step absolute error for the historical series) or "deviation" (standard error of the historical series). Default: "naive".
error_benchmark	String. Benchmark for the relative error metrics. Two options: "naive" (sequential extension of last value) or "average" (mean value of true sequence). Default: "naive".
seed	Positive integer. Random seed. Default: 42.

Value

This function returns a list including:

- exploration: list of all not-null models, complete with predictions, test metrics, prediction stats and plot
- history: a table with the sampled models, hyper-parameters, validation errors, weighted average rank
- best_model: results for the best selected model according to the weighted average rank, including:
 - testing_errors: testing errors for each time feature for the best selected model (me, mae, mse, rmsse, mpe, mape, rmae, rrmse, rame, mase, smse, sce, gmrae)
 - preds: min, max, q25, q50, q75, quantiles at selected ci, mean, sd, mode, skewness, kurtosis, IQR to range, risk ratio, upside probability and divergence for each point for predicted sequences
 - plots: standard plot with confidence interval for each time feature
- time_log

Author(s)

Giancarlo Vercellino <giancarlo.vercellino@gmail.com>

See Also

Useful links:

- https://rpubs.com/giancarlo_vercellino/spooky

Examples

```
spooky(time_features, seq_len = c(10, 30), lno = c(1, 30), n_samp = 1)
```

time_features

time features example: IBM and Microsoft Close Prices

Description

A data frame with daily prices for IBM and Microsoft since March 2017.

Usage

```
time_features
```

Format

A data frame with 2 columns and 1324 rows.

Source

finance.yahoo.com

windower	<i>support functions for spooky</i>
----------	-------------------------------------

Description

support functions for spooky

Usage

```
windower(
  df,
  seq_len,
  lno = 1,
  n_windows = 5,
  ci = 0.8,
  dates = NULL,
  error_scale,
  error_benchmark
)
```

Arguments

df	A data frame with time features on columns
seq_len	Positive integer. Time-step number of the forecasting sequence. Default: NULL (automatic selection between 1 and the square root of full length).
lno	Positive integer. Number of data points to leave out for resampling (using jackknife approach). Default: NULL (automatic selection between 1 and the square root of full length).
n_windows	Positive integer. Number of validation windows to test prediction error. Default: 10.
ci	Confidence interval for prediction. Default: 0.8
dates	Date. Vector with dates for time features.
error_scale	String. Scale for the scaled error metrics. Two options: "naive" (average of naive one-step absolute error for the historical series) or "deviation" (standard error of the historical series). Default: "naive".
error_benchmark	String. Benchmark for the relative error metrics. Two options: "naive" (sequential extension of last value) or "average" (mean value of true sequence). Default: "naive".

Author(s)

Giancarlo Vercellino <giancarlo.vercellino@gmail.com>

Index

* datasets

time_features, [3](#)

spooky, [2](#)

spooky-package (spooky), [2](#)

time_features, [3](#)

windower, [4](#)