Package 'cbar'

October 24, 2017

Type Package

Title Contextual Bayesian Anomaly Detection in R

Version 0.1.3

Date 2017-11-23

Description Detect contextual anomalies in time-series data with Bayesian data analysis. It focuses on determining a normal range of target value, and provides simple-to-use functions to abstract the outcome.

Author Kim Seonghyun <shyeon.kim@scipi.net>

Maintainer Kim Seonghyun <shyeon.kim@scipi.net>

URL https://github.com/zedoul/cbar

BugReports https://github.com/zedoul/cbar/issues

Depends R (>= 3.3.0)

Imports Boom, bsts, dplyr, magrittr, ggplot2, stats

Suggests datasets, knitr, testthat

License Apache License 2.0 | file LICENSE

LazyData true

RoxygenNote 6.0.1

NeedsCompilation no

Repository CRAN

Date/Publication 2017-10-24 13:20:22 UTC

R topics documented:

`bsts_model` *Create bsts model*

Description

Create bsts model

Usage

```
bsts_model(.data, spec = NULL, ...)
```

Arguments

`.data`	training set
`spec`	`cbar.model.spec` object
`...`	params for `bsts_spec_static`

Value

`bsts` which is a bsts model

References

Scott, S. L., & Varian, H. R. (2014). Predicting the present with bayesian structural time series. International Journal of Mathematical Modelling and Numerical Optimisation, 5(1-2), 4-23.

bsts_spec_static *Specify bsts model for static linear regression*

Description

Specify bsts model for static linear regression

Usage

```
bsts_spec_static(.data, sigma_guess = NULL, upper_limit = NULL,
  sd_prior_sample_size = 32, expected_model_size = 3, expected_r2 = 0.8,
  prior_df = 50, niter = 1000, ping = 0, model_options = NULL, ...)
```

Arguments

.data	time-series data to be trained
sigma_guess	an argument for bsts::bsts
upper_limit	an argument for bsts::bsts
sd_prior_sample_size	an argument for bsts::bsts
expected_model_size	an argument for bsts::bsts
expected_r2	an argument for bsts::bsts
prior_df	an argument for bsts::bsts
niter	an argument for bsts::bsts
ping	an argument for bsts::bsts
model_options	an argument for bsts::bsts
...	params for bsts_model

Value

cbar.model.spec object for model specification

cbar cbar *package*

Description

cbar: Contextual Bayesian Anomaly Detection in R

This function generates `cbar` object to detect contextual anomaly and to abstract analysis output.

Usage

```
cbar(.data, ref_period, mea_period, apply_standardized = T, interval = 0.95,
  ...)
```

Arguments

`.data`	data table with datetime, y, and predictors
`ref_period`	performance reference period
`mea_period`	performance measurement period
`apply_standardized`	whether it will standardized data or not
`interval`	credible interval. 0.95 by default.
`...`	params for `bsts_model`

Details

See the README on Github

For the input `.data`, note that you should use `datetime` for the first column name. Also, you should use `numeric` type for other columns.

Examples

```
library(cbar)

.data <- mtcars
rownames(.data) <- NULL
datetime <- seq(from = Sys.time(), length.out = nrow(.data), by = "mins")
.data <- cbind(datetime = datetime, .data)

ref_session <- 1:16
mea_session <- 17:nrow(.data)

obj <- cbar(.data, ref_session, mea_session)
```

destandardized *Destandardize a vector*

Description

Destandardize a vector

Usage

```
destandardized(y_hat, y_mu, y_sd)
```

Arguments

y_hat	standardized numeric vector
y_mu	a mean value of unstandarized vector
y_sd	a standard deviation of unstandarized vector

inference *Infer from predictive posetrior prediction of bsts model*

Description

Infer from predictive posetrior prediction of bsts model

Usage

```
inference(.model, alpha = 0.05)
```

Arguments

.model	bsts model
alpha	percentile for anomaly

Value

data.frame with observations and predictions

plot_error *Print estimation error plot*

Description

Print estimation error plot

Usage

```
plot_error(.cbar, xlab = "", ylab = "Estimation error", method = "diff",
  ...)
```

Arguments

.cbar	cbar object
xlab	a label for x-axis
ylab	a label for y-ayis
method	diff
...	params for boxplot

Value

boxplot object

plot_error_ *Print estimation error plot*

Description

Print estimation error plot

Usage

```
plot_error_(.error, xlab = "", ylab = "Estimation error", method = "diff",
  ...)
```

Arguments

.error	error data frame
xlab	a label for x-axis
ylab	a label for y-ayis
method	diff
...	params for boxplot

`plot_incprob` *Print inclusion probablity plot*

Description

Print inclusion probablity plot

Usage

```
plot_incprob(.cbar, threshold = 0.1, horiz = T, cex.names = 0.5,
  xlab = "Inclusion probability (%)", las = 1, ...)
```

Arguments

`.cbar`	cbar object
`threshold`	a threhold for inclusion probablity
`horiz`	horiz
`cex.names`	cex.names
`xlab`	xlab
`las`	las
`...`	params for barplot

Value

`boxplot` object

`plot_incprob_` *Print inclusion probablity plot*

Description

Print inclusion probablity plot

Usage

```
plot_incprob_(.incprob, threshold = 0.1, horiz = T, cex.names = 0.5,
  xlab = "Inclusion probability (%)", las = 1, ...)
```

Arguments

.incprob	data frame
threshold	a threhold for inclusion probablity
horiz	horiz
cex.names	cex.names
xlab	xlab
las	las
...	params for barplot

Value

boxplot object

---

plot_ts *Print time-series plot*

Description

Print time-series plot

Usage

```
plot_ts(.cbar, x_label = "", y_label = "", seq_by = NULL)
```

Arguments

.cbar	cbar object
x_label	a label for x-axis
y_label	a label for y-ayis
seq_by	increment of the sequence, which is NULL by default

Value

ggplot object

plot_ts_ *Print time-series plot*

Description

Print time-series plot

Usage

```
plot_ts_(target_data, x_label = "", y_label = "", seq_by = NULL)
```

Arguments

target_data	data frame
x_label	a label for x-axis
y_label	a label for y-ayis
seq_by	increment of the sequence, which is NULL by default

Value

ggplot object

point_prediction *Get point prediction from posterior means and response trajectories*

Description

Get point prediction from posterior means and response trajectories

Usage

```
point_prediction(y_hat, .posterior_mean, alpha = 0.05)
```

Arguments

y_hat	response trajectories
.posterior_mean	posterior mean values
alpha	alpha

Value

data.frame for predicted values

posterior_mean *Generate posterior mean of the response variable*

Description

This one is used for point prediction based one predictive posterior distribution

Usage

```
posterior_mean(.model)
```

Arguments

.model	bsts_model

Value

vector that contains poseterior means

print.cbar *Print cbar object*

Description

Print cbar object

Usage

```
## S3 method for class 'cbar'
print(x, ...)
```

Arguments

x	cbar object to print
...	further arguments passed to or from other methods

response_trajectory *Generate trajectories of the response variable*

Description

Note that `posterior_state_samples` returns posterior mean, whereas, this one returns posterior mean + noise

Usage

```
response_trajectory(.model)
```

Arguments

`.model`	`bsts_model`

Details

This one is used for lower and upper bounds

Value

`data.frame` that contains predicted value `y_hat`

standardized *Standardize a vector*

Description

Standardize a vector

Usage

```
standardized(y)
```

Arguments

`y`	numeric vector

summarise_anomaly *Summarise anomaly detection result*

Description

Summarise anomaly detection result

Usage

```
summarise_anomaly(.cbar, .session = NULL)
```

Arguments

.cbar	cbar object
.session	names of sessions, which is NULL by default

Value

data.frame that summarises input data with anomaly label

summarise_incprob *Summarise inclusion probability of model*

Description

Summarise inclusion probability of model

Usage

```
summarise_incprob(.cbar, threshold = 0.1)
```

Arguments

.cbar	cbar object
threshold	threshold of inclusion probability, which is .1 by default

Value

vector that summarises inclusion probabilities for each MCMC samples

summarise_pred_error *Summarise prediction error of model*

Description

This function uses absolute difference and mean absolute percentage error for summarising prediction errors

Usage

```
summarise_pred_error(.cbar, .session = "measurement")
```

Arguments

.cbar	cbar object
.session	names of sessions, which is NULL by default

Value

data.frame with prediction errors

summarise_session *Summarise anomaly in session*

Description

Summarise anomaly in session

Usage

```
summarise_session(.cbar, .session = NULL)
```

Arguments

.cbar	cbar object
.session	names of sessions, which is NULL by default

Value

data.frame that summarises outcome for each sessoin

Index