Package ‘Tariff’

October 29, 2018

Type Package

Title Replicate Tariff Method for Verbal Autopsy

Version 1.0.5

Date 2018-10-23

Author Zehang Li, Tyler McCormick, Sam Clark

Maintainer Zehang Li <lizehang@uw.edu>

Description Implement the Tariff algorithm for coding cause-of-death from verbal autopsies. The Tariff method was originally proposed in James et al (2011) <DOI:10.1186/1478-7954-9-31> and later refined as Tariff 2.0 in Serina, et al. (2015) <DOI:10.1186/s12916-015-0527-9>. Note that this package was not developed by authors affiliated with the Institute for Health Metrics and Evaluation and thus unintentional discrepancies may exist between the this implementation and the implementation available from IHME.

License GPL-2

RoxygenNote 6.1.0

NeedsCompilation no

Repository CRAN

Date/Publication 2018-10-29 05:40:08 UTC

R topics documented:

plot.tariff ................................................................. 2
print.tariff_summary .................................................... 2
RandomVA3 ............................................................... 3
SampleCategory3 ......................................................... 3
summary.tariff ........................................................... 4
tariff ................................................................. 4

Index 7
plot.tariff  
*Plot CSMF of the results obtained from Tariff algorithm*

**Description**

This function plots the CSMF of the fitted results.

**Usage**

```r
## S3 method for class 'tariff'
plot(x, top = NULL, min.prob = 0, ...)
```

**Arguments**

- `x`  
fitted object from `tariff`
- `top`  
maximum causes to plot
- `min.prob`  
minimum fraction for the causes plotted
- `...`  
Arguments to be passed to/from graphic function

**Examples**

```r
data("RandomVA3")
test <- RandomVA3[1:200,]
train <- RandomVA3[201:400,]
allcauses <- unique(train$cause)
fit <- tariff(causes.train = "cause", symps.train = train, symps.test = test, causes.table = allcauses)
plot(fit, top = 10, main = "Top 5 population COD distribution")
plot(fit, min.prob = 0.05, main = "Population COD distribution (at least 5%)")
```

print.tariff_summary  
*Print method for the summary of the results obtained from Tariff algorithm*

**Description**

This function prints the summary message of the fitted results.

**Usage**

```r
## S3 method for class 'tariff_summary'
print(x, ...)
```
**RandomVA3**

**Arguments**

- `x` summary object for Tariff fit
- `...` not used

---

<table>
<thead>
<tr>
<th>RandomVA3</th>
<th>400 records of Sample Input</th>
</tr>
</thead>
</table>

**Description**

This is a dataset consisting of 400 arbitrary sample input deaths randomly sampled from cleaned PHMRC data.

**Format**

400 arbitrary input records.

**Examples**

```r
data(RandomVA3)
head(RandomVA3$train)
head(RandomVA3$test)
```

---

<table>
<thead>
<tr>
<th>SampleCategory3</th>
<th>Grouping of causes in RandomVA3</th>
</tr>
</thead>
</table>

**Description**

This is a matrix specifying a default grouping of the causes used in RandomVA3.

**Format**

17 by 2 matrix

**Examples**

```r
data(SampleCategory3)
SampleCategory3
```
Summary of the results obtained from Tariff algorithm

Description
This function prints the summary message of the fitted results.

Usage
```r
## S3 method for class 'tariff'
summary(object, top = 5, id = NULL, ...)
```

Arguments
- **object**: fitted object from `tariff`
- **top**: number of top CSMF to show
- **id**: the ID of a specific death to show
- **...**: not used

Examples
```r
data("RandomVA3")
test <- RandomVA3[1:200, ]
train <- RandomVA3[201:400, ]
allcauses <- unique(train$cause)
fit <- tariff(causes.train = "cause", symps.train = train,
smps.test = test, causes.table = allcauses)
correct <- which(fit$causes.test[2] == test$cause)
accuracy <- length(correct) / dim(test)[1]
summary(fit)
summary(fit, top = 10)
summary(fit, id = "p849", top = 3)
```

Replicate Tariff methods

Description
This function implements Tariff method.
Usage

tariff(causes.train, symps.train, symps.test, causes.table = NULL,
    use.rank = TRUE, nboot.rank = 1, use.sig = TRUE, nboot.sig = 500,
    use.top = FALSE, ntop = 40, ...)

Arguments

causes.train character vector of causes, or the column name of cause in the training data
symps.train N.train by S matrix
symps.test N.test by S matrix
causes.table list of causes in the data
use.rank logical indicator for whether using ranks instead of scores
nboot.rank number of re-sampling for baseline rank comparison. Default to 1, which re-
samples training data to have a uniform cause distribution of the same size. Set
this to 0 removes bootstrapping the training dataset.
use.sig logical indicator for whether using significant Tariff only
nboot.sig number of re-sampling for testing significance.
use.top logical indicator for whether the tariff matrix should be cleaned to have only top
symptoms
ntop number of top tariff kept for each cause
... not used

Value

score matrix of score for each cause within each death
causes.train vector of most likely causes in training data
causes.test vector of most likely causes in testing data
csmf vector of CSMF
causes.table cause list used for output, i.e., list of existing causes in the training data
use.rank logical indicator for whether using ranks instead of scores

Author(s)

Zehang Li, Tyler McCormick, Sam Clark
Maintainer: Zehang Li <lizehang@uw.edu>

References


Examples

```r
data("RandomVA3")
test <- RandomVA3[1:200, ]
train <- RandomVA3[201:400, ]
allcauses <- unique(train$cause)
fit <- tariff(cause = "cause", symps.train = train, symps.test = test, causes.table = allcauses)
correct <- which(fit$causes.test[,2] == test$cause)
accuracy <- length(correct) / dim(test)[1]
```
Index

*Topic Tariff
  print.tariff_summary, 2
  summary.tariff, 4
  tariff, 4

*Topic datasets
  RandomVA3, 3
  SampleCategory3, 3

plot.tariff, 2
print.tariff_summary, 2

RandomVA3, 3

SampleCategory3, 3
summary.tariff, 4

tariff, 2, 4, 4