

Package ‘cheatR’

May 6, 2020

Type Package

Title Catch Cheaters

Description

A set of functions to compare texts for similarity, and plot a graph of similarities among the compared texts. These functions were originally developed for detection of overlap in course hand-in.

Version 1.2.1

Maintainer Mattan S. Ben-Shachar <matanshm@post.bgu.ac.il>

URL <https://mattansb.github.io/cheatR>

BugReports <https://github.com/mattansb/cheatR/issues>

Depends R (>= 3.6.0)

Imports textreadr, ngram, purrr, utils, R.utils

Suggests knitr, rmarkdown, testthat, devtools, shiny, DT, ggplot2,
tidygraph, ggraph, grid

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

VignetteBuilder knitr

NeedsCompilation no

Author Mattan S. Ben-Shachar [aut, cre]
(<<https://orcid.org/0000-0002-4287-4801>>),
Almog Simchon [aut]

Repository CRAN

Date/Publication 2020-05-06 19:20:02 UTC

R topics documented:

catch_em	2
catch_em_app	3

compare_txt	3
plot.chtrs	4
summary.chtrs	5

Index	6
--------------	----------

catch_em	<i>Match cheaters</i>
----------	-----------------------

Description

Match cheaters

Usage

```
catch_em(flist, n_grams = 10, time_lim = 1L, progress_bar = TRUE)
```

Arguments

flist	a list of documents (.doc/.docx/.pdf). A full/relative path must be provided.
n_grams	see ngram package.
time_lim	max time in seconds for each comparison. Default is 1 second, had no problem comparing documents with 50K words.
progress_bar	Should a progress bar be printed to the console?

Value

A correlation matrix of class chtrs with each cell indicating the match (0-1) between two of the documents.

Author(s)

Mattan S. Ben-Shachar

Examples

```
if (interactive()) {
  files <- choose.files()
  catch_em(files)
}
```

catch_em_app	<i>Run catch_em() with shiny</i>
--------------	----------------------------------

Description

Run catch_em() interactively.

Usage

```
catch_em_app(...)
```

Arguments

... Not used.

Value

A shiny app object.

Author(s)

Almog Simchon

Examples

```
if (interactive()) {  
  catch_em_app()  
}
```

compare_txt	<i>Match cheaters</i>
-------------	-----------------------

Description

Match cheaters

Usage

```
compare_txt(txt1, txt2, n_grams = 10, across = c("both", "txt1", "txt2"))
```

Arguments

txt1, txt2 character vectors to compare, each of length 1.
n_grams see [ngram](#) package.
across How should the percentage of overlap be computed?

Value

The percent (0-1) of overlap between the texts

Author(s)

Mattan S. Ben-Shachar

Examples

```
text1 <- "My horse is large and white, and I ride it every day."
text2 <- "My mule is large and brown, and I ride it most days."
compare_txt(text1, text2, n_grams = 3)
```

plot.chtrs

Plot cheatrs / histogram of similarity scores

Description

Requires ggraph and ggplot2 to work.

Usage

```
## S3 method for class 'chtrs'
plot(x, weight_range = c(0.4, 1), remove_lonely = TRUE, digits = 0, ...)

## S3 method for class 'chtrs'
hist(x, ...)
```

Arguments

x	output of <code>catch_em()</code> .
weight_range	range of edge values to plot
remove_lonely	should lonely nodes (not connected to any edges) be removed from the graph?
digits	Number of digits to round the percentage to.
...	passed to <code>ggraph::ggraph()</code> or <code>ggplot2::geom_histogram</code> .

Value

A ggplot2 plot.

Author(s)

Mattan S. Ben-Shachar

Examples

```
if (interactive()) {
  files <- choose.files()
  res <- catch_em(files)

  plot(res)
  hist(res)
}
```

summary.chtrs

Summarise Cheatrs

Description

Summarise Cheatrs

Usage

```
## S3 method for class 'chtrs'
summary(object, bad_files = FALSE, ...)
```

Arguments

object	output of <code>catch_em()</code> .
bad_files	logical. Instead of the result matrix, should return instead the list of bad files (that did not compare / load)? Defaults to FALSE.
...	Not used.

Value

The input chtrs matrix, or a list of bad files (when `bad_files = TRUE`).

Author(s)

Mattan S. Ben-Shachar

Examples

```
if (interactive()) {
  files <- choose.files()
  res <- catch_em(files)

  summary(res, bad_files = TRUE)
}
```

Index

`catch_em`, 2

`catch_em()`, 4, 5

`catch_em_app`, 3

`compare_txt`, 3

`ggplot2::geom_histogram`, 4

`ggraph::ggraph()`, 4

`hist.chtrs` (`plot.chtrs`), 4

`ngram`, 2, 3

`plot.chtrs`, 4

`summary.chtrs`, 5