Package 'textanalyzer'

January 29, 2025

Type Package

Title 'textanalyzer', an R Package to Analyze Text

Version 0.2.0

Description
It analyzes text to create a count of top n-grams, including tokens (one-word), bigrams(two-word), and trigrams (three-word), while removing all stopwords. It also plots the n-grams and corresponding counts as a bar chart.

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.2

Depends tidytext, tidyr, dplyr, ggplot2, utils, stats

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

Config/testthat/edition 3

VignetteBuilder knitr

NeedsCompilation no

Author Pushker Ravindra [aut, cre]

Maintainer Pushker Ravindra <pushker@gmail.com>

Repository CRAN

Date/Publication 2025-01-29 17:20:02 UTC

Contents

analyze_bigrams *Analyze Bigrams*

Description

Analyze text with ngram=2 (bigrams).

Usage

```
analyze_bigrams(in_text, top_rows = 25)
```

Arguments

in_text	a character vector. Text to be analyzed as a character vector.
top_rows	a numeric vector of length 1. Number of top rows to be returned.

Details

analyze_bigrams

Value

A data.frame with two columns - bigram (character vector) and count (numeric vector).

Author(s)

Ravindra Pushker

Examples

```
analyze_bigrams(in_text=c("The quick brown fox jumps over the lazy dog."))
```

analyze_ngrams *Analyze NGrams*

Description

Analyze text with ngram among 1, 2 or 3.

Usage

```
analyze_ngrams(in_text, ngram = 1, top_rows = 25)
```

Arguments

in_text	a character vector. Text to be analyzed as a character vector.
ngram	a numeric_vector of length 1. Ngram = 1, 2 or 3.
top_rows	a numeric vector of length 1. Number of top rows to be returned.

Details

analyze_ngrams

Value

A data.frame with two columns - word/bigram/trigram (character vector) and count (integer vector).

Author(s)

Ravindra Pushker

Examples

```
analyze_ngrams(in_text=c("The quick brown fox jumps over the lazy dog."))
```

---

analyze_tokens *Analyze Tokens*

Description

Analyze text with ngram=1

Usage

```
analyze_tokens(in_text, top_rows = 25)
```

Arguments

in_text	a character vector. Text to be analyzed as a character vector.
top_rows	a numeric vector of length 1. Number of top rows to be returned.

Details

analyze_tokens

Value

A data.frame with two columns - word (character vector) and count (numeric vector).

Author(s)

Ravindra Pushker

Examples

```
analyze_tokens(in_text=c("The quick brown fox jumps over the lazy dog."))
```

`analyze_trigrams` *Analyze Trigrams*

Description

Analyze text with ngram=3 (trigrams).

Usage

```
analyze_trigrams(in_text, top_rows = 25)
```

Arguments

`in_text`	a character vector. Text to be analyzed as a character vector.
`top_rows`	a numeric vector of length 1. Number of top rows to be returned.

Details

analyze_trigrams

Value

A data.frame with two columns - trigram (character vector) and count (numeric vector).

Author(s)

Ravindra Pushker

Examples

```
analyze_trigrams(in_text=c("The quick brown fox jumps over the lazy dog."))
```

plot_ngrams *Plot Ngrams*

Description

Plot ngrams - Word(s) vs. Count.

Usage

```
plot_ngrams(ngrams_data, top_rows = 25, plot_nrows = 25)
```

Arguments

ngrams_data	a data.frame containing word and n columns.
top_rows	a numeric vector of length 1. Number of top rows to be returned.
plot_nrows	a numeric vector of length 1. Number of rows to be plotted.

Details

plot_ngrams

Value

A ggplot plot object of bar chart with words and their counts.

Author(s)

Ravindra Pushker

Examples

```
plot_ngrams(data.frame(word=c("test1", "test2"), n=c(25, 30)))
```

Index