

Package ‘TextMiningGUI’

October 12, 2022

Version 0.3

Date 2021-04-16

Title Text Mining GUI Interface

Maintainer Conrado Reyes <coreyes@gmail.com>

URL <https://c0reyes.github.io/TextMiningGUI/>

Depends R(>= 3.6), syuzhet

Imports tcltk, dplyr, tidyr, tidytext, tibble, tm, slam, ggplot2,
ggwordcloud, RColorBrewer

Suggests tkrplot, SnowballC, readxl, jsonlite, parallel, ggrepel,
ggpubr, igraph, ggraph, ape, topicmodels, ca, corrr,
data.table, knitr, rmarkdown

Description Graphic interface for text analysis, implement a few methods such as biplots, correspondance analysis, co-occurrence, clustering, topic models, correlations and sentiments.

License GPL (>= 2)

LazyData true

NeedsCompilation no

Author Conrado Reyes [aut, cre],
Purificacion Galindo [tch]

Repository CRAN

Date/Publication 2021-04-17 16:20:02 UTC

R topics documented:

chistes	2
jokes	2
TextMiningGUI	2

Index	5
--------------	----------

chistes

chistes

Description

Data from: <https://github.com/liopic/chistes-nlp>

jokes

jokes

Description

Data from: <https://github.com/taivop/joke-dataset>

TextMiningGUI

TextMiningGUI

Description

Graphic interface for text analysis, implement a few methods such as biplots, correspondence analysis, co-occurrence, clustering, topic models, correlations and sentiments.

File Menu:

- Can import files: csv, excel, json or RData.
- Save project.
- Set work directory.

Data Menu:

- Converter Columns
- Transform
- Slice
- View Data
- View Lexical Table
- View Clean Data

Analysis Menu:

- Statistics
- Most common words
- Word Group

- Word Cloud
- Co-ocurrence
- Cluster
- Correlation
- Correlation Between Two Groups
- AFC
- HJ-Biplot
- Emotions & Sentiments
- Topic Models

Usage

```
TextMiningGUI(seed = 0)
```

Arguments

seed the seed of internal function.

References

- Becue, M. B. (1992) El análisis estadístico de datos textuales. La lectura según los escolares de la enseñanza primaria.
- Becker, R. A., Chambers, J. M. and Wilks, A. R. (1988). The New S Language. Wadsworth & Brooks/Cole. (S version.)
- Benzècri, J. P. (1973). L'Analyse des Données: L'Analyse des correspondances. Paris: Dunod.
- Blei, D. B., Ng, A. Y. N., Jordan, M. I. J. (2003) Latent Dirichlet Allocation.
- Caballero, D. C. (2014) Grupos de Discusión y HJ-Biplot: Una Nueva Forma de Análisis Textual.
- Caballero, D.C. (2011). El HJ-Biplot como Herramienta en el Análisis de Grupos de Discusión. Salamanca: Universidad de Salamanca.
- Collins, M. J. C. (1996) A new Statistical Parser Based on Bigram Lexical Dependencies.
- Diaz-Faes, A. D. (2013) HJ-Biplot como herramienta de inspección de matrices de datos bibliométricos.
- Feinerer, I., Hornik, K. (2019). tm: Text Mining Package. R package version 0.7-7. <https://CRAN.R-project.org/package=tm>
- Feinerer, I., Hornik, K., Meyer, D. (2008). Text Mining Infrastructure in R. Journal of Statistical Software 25(5): 1-54. <http://www.jstatsoft.org/v25/i05/>
- Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency vs interpretability of classifications. Biometrics, 21, 768–769.
- Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. Biometrika 58, 3 , 453-467
- Galindo, M. P. (1985). Contribuciones a la Representación Simultánea de Datos Multidimensionales. Tesis Doctoral. Salamanca: Universidad de Salamanca.

- Galindo, P. (1986). Una alternativa de representaci3n simultanea: HJ-Biplot. *Qüestii3* ,13-23.
- Hartigan, J. A. and Wong, M. A. (1979). Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics*, 28, 100–108. doi: 10.2307/2346830.
- Jockers, ML. (2015). Syuzhet: Extract Sentiment and Plot Arcs from Text. <https://github.com/mjockers/syuzhet>
- Kuhn, M., Jackson, S., Cimentada, J. (2020). corrr: Correlations in R. R package version 0.4.2. <https://CRAN.R-project.org/package=corrr>
- Lloyd, S. P. (1957, 1982). Least squares quantization in PCM. Technical Note, Bell Laboratories. Published in 1982 in *IEEE Transactions on Information Theory*, 28, 128–137.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, eds L. M. Le Cam & J. Neyman, 1, pp. 281–297. Berkeley, CA: University of California Press.
- Müller, K., Wickham, H. (2020). tibble: Simple Data Frames. R package version 3.0.1. <https://CRAN.R-project.org/package=tibble>
- Nenadic, O., Greenacre, M. (2007) Correspondence Analysis in R, with two- and three-dimensional graphics: The ca package. *Journal of Statistical Software* 20(3):1-13.
- Silge, J., Robinson, D. (2016). “tidytext: Text Mining and Analysis Using Tidy Data Principles in R. <https://doi.org/10.21105/joss.00037>
- Osuna, Z. (2006). *Contribuciones al An3lisis de Datos Textuales*.
- Osuna, Z. O. (2004) An3lisis estadístico de datos textuales. Aplicaci3n al estudio de las declaraciones del Libertador Sim3n Bolívar
- Robertson, S. R. (2004) Understanding Inverse Document Frequency: On theoretical arguments for IDF.
- Vicente-Villard3n, J. L. (2017). MultBiplotR: Multivariate Analysis using Biplots. R package version 0.1.0. <http://biplot.dep.usal.es/multbiplot/multbiplot-in-r/>
- Ward, J. H., Jr. (1963), "Hierarchical Grouping to Optimize an Objective Function", *Journal of the American Statistical Association*, 58, 236–244.
- Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wickham, H., Fran3ois, R., Henry, L., Müller, K. (2020). dplyr: A Grammar of Data Manipulation. R package version 0.8.5. <https://CRAN.R-project.org/package=dplyr>
- Wickham, H., Lionel Henry, L. (2020). tidy: Tidy Messy Data. R package version 1.0.2. <https://CRAN.R-project.org/package=tidy>
- NRC Word-Emotion Association Lexicon. (2010). <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>
- ProgGUIinR: support package for «Programming Graphical User Interfaces in R». (2014). <https://rdr.io/cran/ProgGUIinR/>

Examples

```
library(TextMiningGUI)
if(TextMiningGUI()){}
```

Index

chistes, [2](#)

jokes, [2](#)

TextMiningGUI, [2](#)