References and resources: machine learning for NLP

* ["rtweet":](https://github.com/ropensci/rtweet) downdloading tweets to a DataFrame with R
  + a set of [slides](https://mkearney.github.io/nicar_tworkshop/#1) by the package author
* [Online statistical tests: StatsKingdom](https://www.statskingdom.com/)
* MultiLabel classification: [Enron dataset's labels](https://bailando.berkeley.edu/enron/enron_categories.txt)
* ["The Big Bad NLP dataset"](https://datasets.quantumstat.com/): a compilation of modern NLP tasks, with noteBooks.
* [*"*Real or not: NLP with disaster tweets"](https://www.kaggle.com/c/nlp-getting-started/data). Kaggle competition.
* Kaggle "sentiment analysis competition winners", interview how they did: https://www.youtube.com/playlist?list=PLqFaTIg4myu8Iz-n8nGZzfsa4ykvjXnR-
* [ParallelDots, online text analysis APIs for](https://www.paralleldots.com/text-analysis-apis) several tasks: sentiment analysis, tags' prediction, keyword generator, entity extraction, comparing similarity of texts, different emotions analysis, intent analysis, abusive text prediction, etc.
* [Syuzhet package](https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html) for sentiment-analysis categorization: accessing the sentiment-analysis tool of NLP Stanford
* [sentiment140](http://help.sentiment140.com/): an interesting project for automatic sentiment categorization of tweets
  + [description of the corpus, algorithms, etc.](http://help.sentiment140.com/for-students)
  + [R package](https://github.com/okugami79/sentiment140): an R package to use the "sentiment140" from R, including instruction for installation
* [Tidy text mining with R](http://tidytextmining.com/) (an alternative to "tm" R package for NLP preprocessing)
* [Top-10 data mining algorithms (explained)](https://rayli.net/blog/data/top-10-data-mining-algorithms-in-plain-english/)
* [Stanford TreeBank project.](http://nlp.stanford.edu/sentiment/treebank.html) "Recursive deep models for semantic compositionality over a semantic treebank". Using deep learning for online sentiment analysis classification. For learning the models, the project has a large annotated dataset. The project is expalined in a divulgative style in [this article](https://engineering.stanford.edu/news/stanford-algorithm-analyzes-sentence-sentiment-advances-machine-learning). The deep learning methodology of the project is based on the [following paper](http://nlp.stanford.edu/~socherr/EMNLP2013_RNTN.pdf).
* From [RDataMining](http://www.rdatamining.com/home) portal: many applications on the use of R for different data mining tasks
  + [Text mining with R](http://www.rdatamining.com/docs/text-mining-with-r): Twitter data analysis
  + [Twitter data analysis with R](http://www.rdatamining.com/docs/twitter-analysis-with-r): text mining and social network analysis
* [Awesome sentiment analysis](https://github.com/xiamx/awesome-sentiment-analysis): A curated list of Sentiment Analysis methods, implementations and misc.
* ["5 things you need to know about sentiment analysis and classification"](https://www.kdnuggets.com/2018/03/5-things-sentiment-analysis-classification.html)
* Bing Liu's [website on "Opinion mining, sentiment analysis and opinion spam detection: the machine learning approach"](https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html). A key website on the topic
* [18 NLP key terms, explained](https://www.kdnuggets.com/2017/02/natural-language-processing-key-terms-explained.html) for ML practitioners and NLP novices
* F. Hemmatian, M.K. Sorabi (2019). ["A survey on classification techniques for opinion mining and sentiment analysis".](https://link.springer.com/article/10.1007/s10462-017-9599-6) Artificial Intelligence Review, 52, 1495-1545

Última modificación: miércoles, 15 de diciembre de 2021, 12:22