Sentiment analysis for Yelp review classification

<https://medium.com/tensorist/classifying-yelp-reviews-using-nltk-and-scikit-learn-c58e71e962d9>

github

<https://github.com/AbhiAgarwal/yelp-nlp>

# [Does sentiment analysis work? A tidy analysis of Yelp reviews](http://varianceexplained.org/r/yelp-sentiment/)

<http://varianceexplained.org/r/yelp-sentiment/>

<https://github.com/patelparth30j/yelp-sentiment-analysis>

bag of words:

<https://github.com/patelparth30j/yelp-sentiment-analysis/blob/master/yelp_03bagOfWords.ipynb>

word2vec

https://github.com/patelparth30j/yelp-sentiment-analysis/blob/master/yelp\_03bagOfWords.ipynb

# Python NLTK自然语言处理：词干、词形与MaxMatch算法

# <https://zhuanlan.zhihu.com/p/24522344>

# GLM, Random Forest(best), Linear model (rough, no details)

# http://rstudio-pubs-static.s3.amazonaws.com/130466\_792c7b61df694c859a2e037df8c7dc8f.html#/

# useful!!! Install problems??

# http://www.developintelligence.com/blog/2017/03/predicting-yelp-star-ratings-review-text-python/

filter

condition\_latitude = (df\_merged\_business\_review\_user.business\_latitude < 48.987386) \

& (df\_merged\_business\_review\_user.business\_latitude > 18.005611)

condition\_longitude = (df\_merged\_business\_review\_user.business\_longitude > -124.626080) \

& (df\_merged\_business\_review\_user.business\_longitude < -62.361014)

df\_merged\_business\_review\_user\_only\_us = df\_merged\_business\_review\_user.loc[condition\_latitude & condition\_longitude]

统计词频代码

<http://blog.csdn.net/spynao/article/details/50187419>