

[🏠](#) > [综合](#) > NLP+2vec | 认识多种多样的2vec向量化模型

NLP+2vec | 认识多种多样的2vec向量化模型

[综合](#)[来源:sinat_26917383](#)[🕒 2017-04-08 13:53](#)[👁 558°C](#)[💬 0评论](#)[目录\(?\)](#)[\[+\]](#)

来自于github一位博主的整理，好多都没有看到过+还有我一直期待去研究的。
github：<https://github.com/MaxwellRebo/awesome-2vec>

1、word2vec

耳熟能详的NLP向量化模型。

Paper: <https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf>

Java: <http://deeplearning4j.org/word2vec>

C++: <https://github.com/jdeng/word2vec>

Python:

<https://radimrehurek.com/gensim/models/word2vec.html>

<https://github.com/danielfrg/word2vec>

2、doc2vec

[IT问道推荐](#)[关闭](#)

银行贷款频频被拒？

“Dr信用牛牛”让你远离信用污点 国内首家信用健康管理平台免费为你提供信用修复方案

[首页](#) [问答](#) [面试题](#) [程序人生](#) [开源代码库](#) [关于](#)

word2vec模型对词向量进行平均处理，我们仍然忽略了单词之间的排列顺序对情感分析的影响。即上述的word2vec只是基于词的维度进行“语义分析”的，而并不具有上下文的“语义分析”能力。

作为一个处理可变长度文本的总结性方法，Quoc Le 和 Tomas Mikolov 提出了 Doc2Vec方法。除了增加一个段落向量以外，这个方法几乎等同于 Word2Vec。和 Word2Vec 一样，该模型也存在两种方法：Distributed Memory(DM) 和 Distributed Bag of Words(DBOW)。DM 试图在给定上下文和段落向量的情况下预测单词的概率。在一个句子或者文档的训练过程中，段落 ID 保持不变，共享着同一个段落向量。DBOW 则在仅给定段落向量的情况下预测段落中一组随机单词的概率。

(参考：[Algorithm & NLP\] 文本深度表示模型——word2vec&doc2vec词向量模型](#))

3、tweet2vec

Paper: <https://arxiv.org/abs/1605.03481>

Python: <https://github.com/bdhingra/tweet2vec>

一些社交文本中的语言结构跟书面语大不相同，所以作者别出心裁的特意做了一个基于字符组合的模型，其可以基于整个微学习到一种向量化的表达方式。

4、batter-pitcher-2vec

Python: <https://github.com/airalcorn2/batter-pitcher-2vec>

5、illustration-2vec

Python: <https://github.com/rezoo/illustration2vec>

6、lda2vec

IT问道推荐

[关闭](#)



银行贷款频频被拒？

“Dr信用牛牛”让你远离信用污点 国内首家信用健康管理平台免费为你提供信用修复方案

[首页](#) [问答](#) [面试题](#) [程序人生](#) [开源代码库](#) [关于](#)

Python: <https://github.com/cemoody/lda2vec>

将LDA与word2vec相结合，LDA是我比较在意的研究方向，这个模型很早提出，不过应用地倒是不多。

IT问道
www.itwendao.com

lda2vec the topics can be ‘supervised’ and forced to predict another target.

7、sentence2vec

Python: <https://github.com/klb3713/sentence2vec>

8、wiki2vec

Java/Scala: <https://github.com/idio/wiki2vec>

IT问道推荐

[关闭](#)



银行贷款频频被拒？

“Dr信用牛牛”让你远离信用污点 国内首家信用健康管理平台免费为你提供信用修复方案

[首页](#) [问答](#) [面试题](#) [程序人生](#) [开源代码库](#) [关于](#)

Paper: <http://nlp.cs.singhua.edu.cn/~jun/topic-embedding.pdf>

Python: <https://github.com/askerlee/topicvec>

10、entity2vec

Python: <https://github.com/ot/entity2vec>

Paper: http://www.di.unipi.it/~ottavian/files/wsdm15_fel.pdf

11、str2vec

Python: <https://github.com/pengli09/str2vec>

12、node2vec

Paper: <https://arxiv.org/abs/1607.00653>

Page: <https://snap.stanford.edu/node2vec/>

Python: <https://github.com/aditya-grover/node2vec>

可见：笔记 | 基于网络节点的node2vec、论文、算法python实现

13、item2vec

Paper: <https://arxiv.org/abs/1603.04259>

14、author2vec

IT问道推荐

关闭



银行贷款频频被拒？

“Dr信用牛牛”让你远离信用污点 国内首家信用健康管理平台免费为你提供信用修复方案



[首页](#) [问答](#) [面试题](#) [程序人生](#) [开源代码库](#) [关于](#)

15、 playlist2vec

Python: <https://github.com/mattddenewitz/playlist-to-vec>

16、 sense2vec

Paper: <https://arxiv.org/abs/1511.06388>

17、 med2vec

Paper: <https://arxiv.org/abs/1602.05568>

Python: <https://github.com/mp2893/med2vec>

SCODE Word Embeddings using Substitute Words

Python: <https://github.com/ai-ku/wvec>

18、 game2vec

Python: <https://github.com/warchildmd/game2vec>

19、 paper2vec

Paper: <https://arxiv.org/pdf/1703.06587.pdf>

IT问道推荐

关闭



银行贷款频频被拒？

“Dr信用牛牛”让你远离信用污点 国内首家信用健康管理平台免费为你提供信用修复方案

[首页](#) [问答](#) [面试题](#) [程序人生](#) [开源代码库](#) [关于](#)[杭州房产网](#)[完美动力](#)

相关推荐

- opencv+Recorder | OpenCV 中的 Canny 边界检测+轮廓
- python | gensim训练word2vec及相关函数与功能理解
- opencv | opencv中实现行人检测：HOG+SVM（二）
- DeepFashion | 衣物时尚元素关键点定位+时尚元素对齐技术
- Recycleview封装Adapter增加头部尾部
- 那些年刷过的题 - Runtime
- View draw流程分析
- 自定义UIDatePicker，以后不用第三方的，想怎么改就怎么改~~~
- Kotlin基本语法

- opencv | 图片与视频的读入、显示、写出与基本绘图函数介绍
- opencv | HOG描述符介绍+opencv中HOG函数介绍（一）
- SSD+caffe | Single Shot MultiBox Detector 目标检测（一）
- IQA+不懂 | 图像清洗：图像质量评估
- 博弈论游戏 - 信任的进化
- 制作表白动画
- 自定义一个和京东一样的效果的UIAlertView
- 从源码角度分析UIScrollView嵌套ListView显示

[登录](#)

留下您的评论和问题吧

IT问道推荐[关闭](#)

“Dr信用牛牛”让你远离信用
信用健康管理平台免费为你提供信用修复服务

[首页](#) [问答](#) [面试题](#) [程序人生](#) [开源代码库](#) [关于](#)

IT问道内正在发布内容



热评话题

横向列表——点击左右滚动实现 - 互联网 - IT问道

如何在 Laravel 项目中处理 Excel 文件 - 互联网 - IT问道

位运算专题 - 编程语言 - IT问道

JSON-RPC (jsonrpc4j) 使用demo - 互联网 - IT问道

Visual Studio 使用正则表达式查找替换 - 编程语言 - IT问道

我在BAT学到的技术工具 - valgrind - 互联网 - IT问道

今天我收到浙江卫视中国新歌声发来的一条短信说我手机中奖了是真...

版权所有 © 广州市禄志信息科技有限公司 粤ICP备16049987号 [商务合作QQ：2948173213（请注明来意）](#) 站长统计

IT问道

大图



银行贷款频频被拒？

“Dr信用牛牛”让你远离信用污点 国内首家信用健康管理平台免费为你提供信用修复方案

