

立即体验

验证

CSDN

博客 (<http://blog.csdn.net/>) 学院 (<http://edu.csdn.net/?ref=toolbar>) 下载 (<http://download.csdn.net/?ref=toolbar>) GitChat (<http://gitbook.cn/?ref=csdn>) 更多 ▾

0

0

0

0

0

0

0

0

word2vec中的bin文件转换为txt 文件

原创

2016年11月29日 15:58:59

标签 : [neural network \(<http://so.csdn.net/so/search/s.do?q=neural network&t=blog>\)](http://so.csdn.net/so/search/s.do?q=neural network&t=blog)

3037

[python]

1. **from** gensim.models **import** word2vec

2.

3. model = word2vec.Word2Vec.load_word2vec_format(' /home/ubuntu/word2vec/PubMed-w2v.bin', binary=True)

4. model.save_word2vec_format(' /home/ubuntu/word2vec/PubMed-w2v.txt', binary=False)

运行代码时提醒错误 :

/home/ubuntu/workspace/pythondemo/env/bin/python /home/ubuntu/workspace/BioEvent/mainFile/test.py

/home/ubuntu/workspace/pythondemo/env/lib/python3.5/site-packages/gensim/utils.py:1015: UserWarning: Pattern library is not installed, lemmatization won't be available.

warnings.warn("Pattern library is not installed, lemmatization won't be available.")

http://blog.csdn.net/

Process finished with exit code 0

作为轻度强迫症的我，看到这个UserWarning 极不爽快，于是就安装 Pattern, 也是各种安装不成功，网上搜索原因的时候发现：python2.x 才支持Pattern ，而我用的是python3.5.2, 不能因为个Warning退回Python2.x 于是我忍了。果然只是轻度强迫症

第二种方法（与第一种大同小异，不过也记录一下）

[python]

1. **from** gensim.models **import** word2vec

2.

3. model = word2vec.Word2Vec.load_word2vec_format('Path/to/GoogleNews-vectors-negative300.bin', binary=True)

4. model.save("file.txt")

第三种方法（其实都一样啦）

[python]

1. **import** codecs

2. **from** gensim.models **import** Word2Vec

3.

4. **def** main():

5. path_to_model = 'GoogleNews-vectors-negative300.bin'

6. output_file = 'GoogleNews-vectors-negative300_test.txt'

7. export_to_file(path_to_model, output_file)

8.

9.

weixin_3506... (<http://my.csdn.net/?ref=toolbar>)

写博客 (<http://write.blog.csdn.net/plogedit?ref=toolbar>)

source=csdnblog

小妖精Fsky (<http://blog.csdn.net/u014221266>)

+ 关注

码云

未开通

(<https://github.com>)

原创

275

粉丝

48

喜欢

0

未开通

(<https://github.com>)

utm_source

他的最新文章

更多文章 (<http://blog.csdn.net/u014221266>)

network中的weight initialization (<http://blog.csdn.net/u014221266/article/details/78806874>)

[ICML2015]Batch Normalization:Accelerating Deep Network Training by Reducing Internal Covariate Shift (<http://blog.csdn.net/u014221266/article/details/78788409>)

[ACL2015]A Dependency-Based Neural Network for Relation Classification (<http://blog.csdn.net/u014221266/article/details/78778555>)

过拟合函数中系数大小的理解 (<http://blog.csdn.net/u014221266/article/details/78765799>)

pytorch版的biLstm+crf实现sequence label (<http://blog.csdn.net/u014221266/article/details/78664824>)

相关推荐

Delphi7高级应用开发随书源码 (<http://download.csdn.net/download/chenxh3>)

7076

BIN和TXT转换 (<http://blog.csdn.net/hadao/article/details/5598676>)

Delphi7高级应用开发随书源码 (<http://download.csdn.net/download/chenxh3>)

7076

<http://blog.csdn.net/appleml/article/details/53393790>

1/5



Delphi7高级应用开发随书源码 (<http://download.csdn.net/download/chenx...>)

[/http://download/](#) 2003年04月30日 00:00 676KB [下载](#)

[e/details/44995203](#)
[11990](#)

正则表达式匹配连续多个空格或tab空格
(<http://blog.csdn.net/u014221266/article/d>
[etails/41891825](#))
[9643](#)


word2vec学习笔记[转] (<http://blog.csdn.net/zhoub1668/article/details/23269705>)

谷歌已经使用Deep Learning技术开发了许多新方法来解析语言，目前，谷歌开源了一款基于Deep Learning的学习工具——word2vec，这是首款面向大众的Deep Learning学习...

 [zhoub1668](http://blog.csdn.net/zhoub1668) (<http://blog.csdn.net/zhoub1668>) 2014年04月09日 14:50 [35548](#)

Gensim Word2vec 使用教程 (http://blog.csdn.net/Star_Bob/article/details/47808499)

本文主要基于Radim Rehurek的Word2vec Tutorial.**准备输入**Gensim的word2vec的输入是句子的序列. 每个句子是一个单词列表代码块例如：>>> # impor...

 [Star_Bob](http://blog.csdn.net/Star_Bob) (http://blog.csdn.net/Star_Bob) 2015年08月20日 15:26 [25357](#)




程序员跨越式成长指南

完成第一次跨越，你会成为具有一技之长的开发者，月薪可能翻上几番；完成第二次跨越，你将成为拥有局部优势或行业优势的专业人士，获得个人内在价值的有效提升和外在收入的大幅跃迁.....

(http://www.baidu.com/cb.php?c=lgF_pyfqHmknjzrjD0lZ0qnK9ujYzP1f4PjnY0Aw-5Hc4nj6vPjm0TAq15Hf4rjn1n1b0T1YYPvRsmwywbPW7bnvP-mhzmz0AwY5HDdnHndPH0LPjD0IgF_5y9YIZ0IQzqMpgwBUvqoQhP8QvIGIAPCmgfEmvq_lyd8Q1R4uWc4uHf3uAckPHRkPWN9PhcsmW9huWqdlAdxTvqdThP-5HDknWFBmhkEusKzujYk0AFV5H00TZcqN0KdpyfqHRLPjnvnfKEpyfqHnsnj0YnsKWpyfqP1cvrHnz0AqLUWYs0ZK45HcsP6KWThnqPWn4PWm))


word2vec使用指导 (<http://blog.csdn.net/zhoub1668/article/details/24314769>)

word2vec是一个将单词转换成向量形式的工具。可以把对文本内容的处理简化为向量空间中的向量运算，计算出向量空间上的相似度，来表示文本语义上的相似度。 一、理论概述（主要来源于<http://l...>

 [zhoub1668](http://blog.csdn.net/zhoub1668) (<http://blog.csdn.net/zhoub1668>) 2014年04月22日 16:34 [131512](#)


word2vec使用说明 (<http://blog.csdn.net/Yelbosh/article/details/49721569>)

Google的word2vec官网：<https://code.google.com/p/word2vec/> 下载下来的Demo源码文件共有如下几个： word2vec – Revision 4...

 [Yelbosh](http://blog.csdn.net/Yelbosh) (<http://blog.csdn.net/Yelbosh>) 2015年11月08日 20:53 [10768](#)


【python gensim使用】word2vec词向量处理中文语料 (<http://blog.csdn.net/churximi/articl...>)

word2vec介绍word2vec官网：<https://code.google.com/p/word2vec/> word2vec是google的一个开源工具，能够根据输入的词的集合计算出词与词之间的...

 [churximi](http://blog.csdn.net/churximi) (<http://blog.csdn.net/churximi>) 2016年05月21日 20:57 [25189](#)


word2vec词向量训练及中文文本相似度计算 (<http://blog.csdn.net/Eastmount/article/details/...>)

本文是讲述如何使用word2vec的基础教程，文章比较基础，希望对你有帮助！ 官网C语言下载地址：http://word2vec.googlecode.com/svn/trunk/ Word2vec...

 Eastmount (http://blog.csdn.net/Eastmount) 2016年02月18日 00:35 046215


Java版本word2vec跑中文数据使用心得 (http://blog.csdn.net/czs1130/article/details/53355...

笔者目前在研究中文情感分析，其中考虑到使用word2vec来对词进行量化，从而达到计算词相似度或者之后拓展其他用法的目的。由于项目是Java版的，所以研究的是java版本的word2vec在此写下...

 czs1130 (http://blog.csdn.net/czs1130) 2016年11月26日 21:37 05825


python 环境下gensim中的word2vec的使用笔记 (http://blog.csdn.net/philosophyatmath/ar...

centos 7, python2.7, gensim (0.13.1)语料：http://211.136.8.18/files/1094000015A9F94/mattmahoney.net/dc...

 philosophyatmath (http://blog.csdn.net/philosophyatmath) 2016年08月29日 16:57 016971


词向量之加载word2vec和glove (http://blog.csdn.net/u010041824/article/details/70832295)

1 Google用word2vec预训练了300维的新闻语料的词向量googlenews-vecctors-negative300.bin，解压后3.39个G。 可以用gensim加载进来，...

 u010041824 (http://blog.csdn.net/u010041824) 2017年04月26日 20:57 03907


word2vec——高效word特征求取 (http://blog.csdn.net/abcjennifer/article/details/46397829)

继上次分享了经典统计语言模型，最近公众号中有很多做NLP朋友问到了关于word2vec的相关内容， 本文就在这里整理一下做以分享。 本文分为 概括word2vec 相关工作 模型结构 Count...

 abcjennifer (http://blog.csdn.net/abcjennifer) 2015年06月07日 16:07 023385


深度学习：使用 word2vec 和 gensim (http://blog.csdn.net/chivalrouslil/article/details/5047...

gensim 官网：http://radimrehurek.com/gensim/index.html Word2vec: the good, the bad (and the fast) ...

 chivalrouslil (http://blog.csdn.net/chivalrouslil) 2016年01月06日 17:23 02524


gensim中使用word2vec (http://blog.csdn.net/a1368783069/article/details/52025764)

训练语料由于自己在csdn的上传空间不够，暂时将语料放在百度云上 链接: https://pan.baidu.com/s/1qYKRXOo 密码: 4psr 文件名是 text8 或者在参考文章...

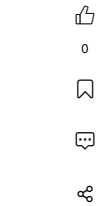
 a1368783069 (http://blog.csdn.net/a1368783069) 2016年07月25日 17:34 06149

深度学习笔记——基于Word2vec和Doc2vec的句子对匹配方法 (http://blog.csdn.net/mpk_no...

句子对匹配（Sentence Pair Matching）问题是NLP中非常常见的一类问题，所谓“句子对匹配”，就是说给定两个句子S1和S2，任务是判断这两个句子是否具备某种类型的关系。本文是一个...

 mpk_no1 (http://blog.csdn.net/mpk_no1) 2017年06月01日 22:20 01278





[mpk_no1 \(http://blog.csdn.net/mpk_no1\)](http://blog.csdn.net/mpk_no1)

2017年06月01日 22:39

11370

ANSJ调用word2vec model文件 (<http://blog.csdn.net/chivalrousli/article/details/49996073>)

在使用word2vec对文本进行分析时能产生一个bin文件，可以用ANSJ调用该文件得到相近词，等同于替换 ./distance vectors. bin命令。代码如下：im...

[chivalrousli \(http://blog.csdn.net/chivalrousli\)](http://blog.csdn.net/chivalrousli)

2015年11月23日 15:46

1105

ANSJ调用word2vec model文件 (<http://blog.csdn.net/jj12345jj198999/article/details/11640...>)

在使用word2vec对文本进行分析时能产生一个bin文件，可以用ANSJ调用该文件得到相近词，等同于替换 ./distance vectors. bin命令。代码如下：import...

[jj12345jj198999 \(http://blog.csdn.net/jj12345jj198999\)](http://blog.csdn.net/jj12345jj198999)

2013年09月13日 13:55

17407

用Word2vec训练中文wiki，构造词向量并做词聚类 (http://blog.csdn.net/accumulate_zhang...)

I利用word2vec训练中文wiki,构造词向量,并搞搞词聚类。

[accumulate_zhang \(http://blog.csdn.net/accumulate_zhang\)](http://blog.csdn.net/accumulate_zhang)

2016年09月25日 14:57

5575

机器学习实验（十）：基于WiFi fingerprints用自编码器(Autoencoders)和神经网络(Neural N...

Autoencoders and Neural Network for Place recognition with WiFi fingerprints 本文来源于Michal Nowicki • ...

[u013719780 \(http://blog.csdn.net/u013719780\)](http://blog.csdn.net/u013719780)

2016年11月21日 17:08

4351
