


19.1，依次类推。因此，根据Tfidf文档表示方法和余弦相似度方法，与我们的查询文档vec最相似的文档为3号文档，相似度达到82.1%。注意：4-8号文档与vec没有任何公共的属性，因此相似度为0.0。

了解详细请看《相似度查询》(<http://radimrehurek.com/gensim/tut3.html>)教程。

[1] (<http://radimrehurek.com/gensim/tutorial.html#id1>) This is the same corpus as used in **Deerwester et al. (1990): Indexing by Latent Semantic Analysis** (http://www.cs.bham.ac.uk/~pxt/IDA/lsa_ind.pdf), Table 2.



 发表你的评论

(http://my.csdn.net/weixin_35068028)

相关文章推荐



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) 2015年08月20日 15:26  25271

python | gensim训练word2vec及相关函数与功能理解 (http://blog.csdn.net/sinat_26917383...)

一、gensim介绍gensim是一款强大的自然语言处理工具，里面包括N多常见模型： - 基本的语料处理工具 - LSI - LDA - HDP - DTM - DIM - ...

 sinat_26917383 (http://blog.csdn.net/sinat_26917383) 2017年04月09日 11:23  5561



 **霸气！重磅改革！吴恩达说：女儿识字后就教她学Python！**

Python的火爆最近越来越挡不住了，连身边多年工作经验的朋友都开始学Python了！他是这么说的....

(http://www.baidu.com/cb.php?c=lgF_pyfqhHmknjnvPjc0IZ0qnfK9ujYzP1ndPWb10Aw-5Hc3rHnYnHb0TAq15HfLPWRznpjb0T1dbrHw9mvm3ujDzn1NbmlN90AwY5HDdnHnYnjcYPHm0lgF_5y9YIZ0lQzq-uZR8mLPbUB48ugfEIAqspynETZ-YpAq8nWqdlAdxTvqdThP-5yF_UvTkn0KzujYk0AFV5H00TZcq0KdpyfqhHRLPjnvnfKEpyfqhHc4rj6kP0KWpyfqP1cvrHnz0AqLUWYs0ZK45HcsP6KWThnqnHn3PW0)



Gensim入门教程 (http://blog.csdn.net/Together_CZ/article/details/7288505)

转自：http://www.cnblogs.com/iloveai/pl/gensim_tutorial.html What is Gensim? Gensim是一款开源的第三方Py...

 Together_CZ (http://blog.csdn.net/Together_CZ) 2017年06月06日 22:08  348


【gensim中文教程】开始使用gensim (<http://blog.csdn.net/DuinoDu/article/details/766186...>)


原文链接介绍了基本概念，以及理解和使用gensim的基本元素，并提供了一个简单的例子。...

 DuinoDu (<http://blog.csdn.net/DuinoDu>) 2017年08月03日 14:08  417



gensim 教程 -Part1 (http://blog.csdn.net/Thinking_boy1992/article/details/53285592)

本文翻译自 Gensim使用Python的标准日志模型，在不同的优先级中来记录各种东西；为了激活日志，运行： >>> import logging >>> logging.basicConfig

 内容举报

 返回顶部

的集合计算出词与词之间的...

 chunximi (<http://blog.csdn.net/chunximi>) 2016年05月21日 20:57  25106

 返回顶部


Python版的Word2Vec -- gensim 学习手札 中文词语相似性度量 V1.1 (<http://blog.csdn.net/M...>)

前言相关内容链接： 第一节：Google Word2vec 学习手札 昨天好不容易试用了一下Google自己提供的Word2Vector的源代码，花了好长时间训练数据，结果发现似乎Python并不能...

 MebiuW (<http://blog.csdn.net/MebiuW>) 2016年08月24日 20:10  9398

Gensim做中文主题模型（LDA） (<http://blog.csdn.net/kunlong0909/article/details/48029111>)

中文语料来自<http://www.sogou.com/labs/dl/c.html> 的精简版（tar.gz格式） 24M jerry@hq:/u01/jerry/Reduced\$ ls C0000...

 kunlong0909 (<http://blog.csdn.net/kunlong0909>) 2015年08月27日 16:03  2378



Python与自然语言处理（二）基于Gensim的Word2Vec (<http://blog.csdn.net/monkey13149...>)

Java调用NLPIC实现分词和标注工作，Python调用Word2Vec实现词向量相似度计算


 monkey131499 (<http://blog.csdn.net/monkey131499>) 2016年04月12日 10:13  8524

如何安装python gensim (http://blog.csdn.net/Star_Bob/article/details/47835217)

本文前提是已经安装了python的scipy和numpy模块，针对eclipse 编辑器的使用首先，下载python gensim包进入下载页面：<http://download.csdn.net/d...>



 Star_Bob (http://blog.csdn.net/Star_Bob) 2015年08月21日 15:25  3754

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

 <http://download.csdn.net/download/chenx...> 2003年04月30日 00:00 676KB [下载](#)

主题模型TopicModel：通过gensim实现LDA (<http://blog.csdn.net/pipisorry/article/details/4...>)



<http://blog.csdn.net/pipisorry/article/details/46447561>使用python gensim轻松实现lda模型。gensim简介Gensim是一个相当专...

 pipisorry (<http://blog.csdn.net/pipisorry>) 2015年06月10日 22:27  8194

 内容举报

python中用gensim做wiki的中文数据word2vector处理 (<http://blog.csdn.net/lixintong1992/...>)



gensim 用conda install gensim 与pip install gensim 安装是不同的 提示C编译器会更快，装了 MinGW...

 lixintong1992 (<http://blog.csdn.net/lixintong1992>) 2015年12月23日 14:54  5941

 返回顶部

利用Python gensim基于中文语料建立LSA隐性语义模型 (http://blog.csdn.net/lo_cima/articl...)

传统词向量空间模型 对语料库中每个文档进行分词后，将生成一个词典。每个文档对应一个与词典长度一致的向量。若文档中出现某个词，则向量中相应的项目非零。下标index通过将词语输入Hashing函数获得...

 lo_cima (http://blog.csdn.net/lo_cima) 2016年07月07日 16:46  1699

 28   

 28  