



```
26. A = dict.fromkeys(text_labels,0) #预测正确的各个类的数目
27. B = dict.fromkeys(text_labels,0) #测试数据集中各个类的数目
28. C = dict.fromkeys(text_predict_labels,0) #预测结果中各个类的数目
29. for i in range(0,len(labels_right)):
30.     B[labels_right[i]] += 1
31.     C[labels_predict[i]] += 1
32.     if labels_right[i] == labels_predict[i]:
33.         A[labels_right[i]] += 1
34. print(A )
35. print(B)
36. print( C)
37. #计算准确率，召回率，F值
38. for key in B:
39.     p = float(A[key]) / float(B[key])
40.     r = float(A[key]) / float(C[key])
41.     f = p * r * 2 / (p + r)
42.     print ("%s:\t p:%f\t r:%f\t f:" % (key,p,r,f))
```

FastText 文本分类使用心得 (http://diog.cs  
dn.net/thriving\_fc/article/details/5323985  
6)



java版本情况：

github上下载地址：

- 1. [https://github.com/ivanhk/fastText\\_java](https://github.com/ivanhk/fastText_java)

看了下sh脚本的使用方法，自己简单了些个text的方法，正好有用，后面会拿xgboost进行对比，看看效果，效果可以的写成service进行上线：

```
[java]
1. package test;
2. import java.util.List;
3.
4. import fasttext.FastText;
5. import fasttext.Main;
6. import fasttext.Pair;
7.
8. public class Test {
9.     public static void main(String[] args) throws Exception {
10.
11.         String[] text = {
12.             "supervised",
13.             "-input",
14.             "/Users/shuubiasahi/Documents/python/fasttext/news_fasttext_train.txt",
15.             "-output", "/Users/shuubiasahi/Documents/faste.model", "-dim",
16.             "10", "-lr", "0.1", "-wordNgrams", "2", "-minCount", "1",
17.             "-bucket", "10000000", "-epoch", "5", "-thread", "4" };
18.         Main op = new Main();
19.         op.train(text);
20.         FastText fastText = new FastText();
21.         String[] test = { "就读", "科技", "学生", "学生", "学生"};
22.         fastText.loadModel("/Users/shuubiasahi/Documents/faste.model.bin");
23.         List<Pair<Float, String>> list = fastText.predict(test, 6); //得到最大可能的六个预测概
24.         率
25.         for (Pair<Float, String> parir : list) {
26.             System.out.println("key is:" + parir.getKey() + "    value is:"
27.                 + parir.getValue());
28.         }
29.     }
30. }
```

在线课程

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

[http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-](http://www.baidu.com/cb.php?c=lgF_pyfqHmknjmsnj0iZ0qnfK9ujYzP1mznWR10Aw-)

java中输入一行字符，分别统计英文字  
母、空格、数字的个数 (http://blog.csdn.n  
et/luoyexuge/article/details/49531737)

利用tensorflow来做验证码识别 (http://blo  
g.csdn.net/luoyexuge/article/details/5413  
0765)

```
28.         System.out.println(Math.exp(list.get(0).getKey())); //得到最大预测概率
29.
30.     }
31.
32. }
```

oZ00J

🔖 7656

python中pandas库学习笔记 (http://blog.csdn.net/luoyexuge/article/details/49104583)

🔖 7556

这里设置bucket不适用设置过大，过大会产生OOM，而且模型保存太大，上面的设置模型保存就有1个g，-wordNgrams可以设置为2比设置为1能提高模型分类的准确性，

结果情况:

key is:0.0 value is: \_\_label\_\_edu  
key is:-17.75125 value is: \_\_label\_\_affairs  
key is:-17.75125 value is: \_\_label\_\_economic  
key is:-17.75125 value is: \_\_label\_\_ent  
key is:-17.75125 value is: \_\_label\_\_fashion  
key is:-17.75125 value is: \_\_label\_\_game

1.0

注意fasttext对输入格式有要求，label标签使用 “\_\_label\_\_”+实际标签的形式， over有问题联系我




1



2016年5月26 我的模型已经上线了 效果还不错

版权声明：本文为博主原创文章，未经博主允许不得转载。


 发表你的评论

(http://my.csdn.net/weixin\_35068028)

 neruda1991 (/neruda1991) 2017-12-02 12:10 4楼


(/neruda1991) 旭哥应该也是经常用java来实现工程的吧

回复 1条回复

 TANG85YAN (/TANG85YAN) 2017-10-30 18:11 3楼

(/TANG85YAN) 请问为什么会出：The kernel appears to have died. It will restart automatically.的情况？  
我使用的是Anaconda3-5.0.0-Windows-x86\_64，python版本Python 3.6.2

回复

 yuzhengfei7 (/yuzhengfei7) 2017-09-17 20:22 2楼

(/yuzhengfei7) 作为一个fastText初学者小白，我有些疑问可以帮解决下么！


回复 1条回复

[查看 11 条热评](#)

## 相关文章推荐

**文本分类（六）：使用fastText对文本进行分类--小插曲 (<http://blog.csdn.net/lxg0807/article/...>)**

测试facebook开源的基于深度学习的对文本分类的fastText模型 fasttext python包的安装.pip install fasttext第一步获取分类文本，文本直接用的清华大学的新...

 lxg0807 (<http://blog.csdn.net/lxg0807>) 2016年10月28日 21:44 [评论](#) 14555

**fastText原理及应用 (<http://blog.csdn.net/chivalrouslil/article/details/76218099>)**

本文转载自: [http://blog.csdn.net/yick\\_liao/article/details/62222153](http://blog.csdn.net/yick_liao/article/details/62222153) 简介 fastText是一种Facebook AI R...

 chivalrouslil (<http://blog.csdn.net/chivalrouslil>) 2017年07月27日 19:50 [评论](#) 723


**2017年前端报告：程序员薪酬上涨70%！**

前端程序员的薪酬曝光，2017年，平均上涨70%，月薪20的人最为常见！以下为详细数据.....

([http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknj0dP1f0lZ0qnK9ujYzP1nYPH0k0Aw-5Hc3rHnYnHb0TAq15HfLPWRznjb0T1YYryf1uARkmvc3mWKbuAfY0AwY5HDdnHnYn1DzPHD0lgF\\_5y9YIZ0lQzquZR8mLPbUB48ugfElAqspynElvNBnHqdlAdxTvqdThP-5yF\\_UvTkn0KzujYk0AFV5H00TZcqN0KdpyfqHRLPjnvnfKEpyfqHc4rj6kP0KWpyfqP1cvrHnz0AqLUWYs0ZK45HcsP6KWThnqPHbdPHD](http://www.baidu.com/cb.php?c=lgF_pyfqHmknj0dP1f0lZ0qnK9ujYzP1nYPH0k0Aw-5Hc3rHnYnHb0TAq15HfLPWRznjb0T1YYryf1uARkmvc3mWKbuAfY0AwY5HDdnHnYn1DzPHD0lgF_5y9YIZ0lQzquZR8mLPbUB48ugfElAqspynElvNBnHqdlAdxTvqdThP-5yF_UvTkn0KzujYk0AFV5H00TZcqN0KdpyfqHRLPjnvnfKEpyfqHc4rj6kP0KWpyfqP1cvrHnz0AqLUWYs0ZK45HcsP6KWThnqPHbdPHD))


**NLP | 高级词向量表达（二）——FastText（简述、学习笔记） ([http://blog.csdn.net/sinat\\_2...](http://blog.csdn.net/sinat_2...))**

FastText是Facebook开发的一款快速文本分类器，提供简单而高效的文本分类和表征学习的方法，不过这个项目其实是有两部分组成的，一部分是这篇文章介绍的 fastText 文本分类（pap...

 sinat\_26917383 ([http://blog.csdn.net/sinat\\_26917383](http://blog.csdn.net/sinat_26917383)) 2017年02月03日 23:20 [评论](#) 14455


**FastText 文本分类使用心得 ([http://blog.csdn.net/thriving\\_fcl/article/details/53239856](http://blog.csdn.net/thriving_fcl/article/details/53239856))**

fasttext文本分类

 thriving\_fcl ([http://blog.csdn.net/thriving\\_fcl](http://blog.csdn.net/thriving_fcl)) 2016年11月20日 12:01 [评论](#) 12187

**fasttext使用笔记 ([http://blog.csdn.net/m0\\_37306360/article/details/72832606](http://blog.csdn.net/m0_37306360/article/details/72832606))**

这里记录使用fastText训练word vector笔记 github地址：<https://github.com/facebookresearch/fastText> 下载到本机: ...

 m0\_37306360 ([http://blog.csdn.net/m0\\_37306360](http://blog.csdn.net/m0_37306360)) 2017年06月01日 16:32 [评论](#) 1526

**一学就会的 WordPress 实战课**

学习完本课程可以掌握基本的 WordPress 的开发能力，后续可以根据需要开发适合自己的主题、插件，打造最个性的 WordPress 站点。



1



([http://www.baidu.com/cb.php?c=lgF\\_pyfqhHmknjfvP1m0lZ0qnfK9ujYzP1f4PjnZ0Aw-5Hc4nj6vPjm0TAq15Hf4rjn1n1b0T1YvmWIhmy79nvfsuAfvHfz0AwY5HDdnHnYn1DzPHD0lgF\\_5y9YlZ0lQzqMpgwBUvqoQhP8QviGIAPCmgfEmvq\\_lyd8Q1N9nHmvnj7hnHPWnjFhPAD1Pyn4uW99ujqdlAdxTvqdThP-5HDknWw9mhhEusKzujYk0AFV5H00TZcq0KdpyfqhHRLPjnvnfKEpyfqhHnsnj0YnsKWpyfqP1cwrHnz0AqLUWYs0ZK45HcsP6KWThnqPH0vpJt](http://www.baidu.com/cb.php?c=lgF_pyfqhHmknjfvP1m0lZ0qnfK9ujYzP1f4PjnZ0Aw-5Hc4nj6vPjm0TAq15Hf4rjn1n1b0T1YvmWIhmy79nvfsuAfvHfz0AwY5HDdnHnYn1DzPHD0lgF_5y9YlZ0lQzqMpgwBUvqoQhP8QviGIAPCmgfEmvq_lyd8Q1N9nHmvnj7hnHPWnjFhPAD1Pyn4uW99ujqdlAdxTvqdThP-5HDknWw9mhhEusKzujYk0AFV5H00TZcq0KdpyfqhHRLPjnvnfKEpyfqhHnsnj0YnsKWpyfqP1cwrHnz0AqLUWYs0ZK45HcsP6KWThnqPH0vpJt))

**FastText (<http://blog.csdn.net/Dinosoft/article/details/53928409>)**

听闻FastText是word2vec的加速版，吓得我赶紧找点资料看看，免得被人给忽悠了。《PaperWeekly 第五期——从Word2Vec到FastText》这篇文章不错，对word2vec...

Dinosoft (<http://blog.csdn.net/Dinosoft>) 2016年12月29日 15:20 1968

**fastText具体代码实现过程 ([http://blog.csdn.net/qz\\_28743951/article/details/74251856](http://blog.csdn.net/qz_28743951/article/details/74251856))**

标签传播算法（Label Propagation）及Python实现zouxy09@qq.com<http://blog.csdn.net/zouxy09> 众所周知，机器学习可以大体分为三大类：监督...

qq\_28743951 ([http://blog.csdn.net/qz\\_28743951](http://blog.csdn.net/qz_28743951)) 2017年07月03日 18:57 379

**A Python interface for Facebook fastText library ([http://blog.csdn.net/qz\\_28743951/article/details/74251856](http://blog.csdn.net/qz_28743951/article/details/74251856))**

Latest Version: 0.8.3fasttext is a Python interface for Facebook fastText

qq\_28743951 ([http://blog.csdn.net/qz\\_28743951](http://blog.csdn.net/qz_28743951)) 2017年07月04日 10:20 454

**fastText原理及应用 ([http://blog.csdn.net/Yick\\_Liao/article/details/62222153](http://blog.csdn.net/Yick_Liao/article/details/62222153))**

简介fastText是一种Facebook AI Research在16年开源的一个文本分类的。其特点就是fast。相对于其它文本分类模型，如SVM，Logistic Regression和neural network...

Yick\_Liao ([http://blog.csdn.net/Yick\\_Liao](http://blog.csdn.net/Yick_Liao)) 2017年03月15日 13:49 6058

**fastText Ngram 的处理过程 (<http://blog.csdn.net/LaineGates/article/details/77839847>)**

最近小研究了下fastText过程，挺有收获，这里单讲下fastText的ngram处理过程，其余训练一类和word2vec很像，有兴趣的请移步 fastText 源码分析。基础信息首先说下（1...

LaineGates (<http://blog.csdn.net/LaineGates>) 2017年09月04日 17:36 546

**fasttext (<http://blog.csdn.net/phoeny0201/article/details/52329477>)**

同样的作者

phoeny0201 (<http://blog.csdn.net/phoeny0201>) 2016年08月26日 16:12 1770

**Facebook : FastText 理解和在query意图识别的应用 ([http://blog.csdn.net/hero\\_fantao/article/details/72677186](http://blog.csdn.net/hero_fantao/article/details/72677186))**

Facebook 在2016年第一次对外公开FastText算法时候，应该是引起很大一番讨论，因为论文提到他以更快的速度达到和DNN类似的效果。这里不再争论这点。当时吸引我一点的是他在大规模数据集上的...

hero\_fantao ([http://blog.csdn.net/hero\\_fantao](http://blog.csdn.net/hero_fantao)) 2017年04月06日 22:55 1710



1




[转]Facebook 开源的快速文本分类器 FastText (http://blog.csdn.net/heiyeshuwu/article/det...

Facebook声称fastText比其他学习方法要快得多，能够训练模型“在使用标准多核CPU的情况下10分钟内处理超过10亿个词汇”，特别是与深度模型对比，fastText能将训练时间由数天缩短到几...

 heiyeshuwu (http://blog.csdn.net/heiyeshuwu) 2017年03月16日 22:21 669


FastText 词向量与文本分类 (http://blog.csdn.net/chuchus/article/details/77771128)

1.简介FastText, 一种技术, 也是 An NLP library by Facebook.2.原理3.FastText libraryGitHub: fastText 它由两部分组成: w...

 chuchus (http://blog.csdn.net/chuchus) 2017年09月01日 11:38 669


python3安装fasttext (http://blog.csdn.net/appleychi/article/details/77765637)

Collecting fasttext Downloading https://pypi.doubanio.com/packages/a4/86/ff826211bc9e28d4c371668b...

 appleychi (http://blog.csdn.net/appleychi) 2017年09月01日 09:46 661


文本分类（六）：使用fastText对文本进行分类--小插曲 (http://blog.csdn.net/haluoluo211/ar...

http://blog.csdn.net/lxg0807/article/details/52960072 环境说明：python2.7、Linux 自己打自己脸，目前官方的包只能在li...

 haluoluo211 (http://blog.csdn.net/haluoluo211) 2017年06月28日 15:59 501


使用fasttext完成文本处理及文本预测 (http://blog.csdn.net/Season\_For\_Lin/article/details/7...

使用fasttext实现文本处理及文本预测

 Season\_For\_Lin (http://blog.csdn.net/Season\_For\_Lin) 2017年11月18日 15:25 16548


使用Clion调试fastText源代码 (http://blog.csdn.net/sinat\_33455447/article/details/74505671)

前言 最近在导师的带领下开始研究fastText，fastText是由Facebook开源的快速文本分类器，比深度学习快了很多，一般深度学习要4，5个小时才能跑出来的大文本分类，fastText只要...

 sinat\_33455447 (http://blog.csdn.net/sinat\_33455447) 2017年07月06日 12:25 413

centos 下 gcc c++ 版本更新 Fasttext Glove (http://blog.csdn.net/dongfangxiaozi\_/article/...

centos 默认的gcc版本为4.4.7，对于GitHub上的很多项目都没法正常编译。为了编译Glove词向量工具，遇到以下错误：当然很多项目都不能正常编译。在centos下更新gc...

 dongfangxiaozi\_ (http://blog.csdn.net/dongfangxiaozi\_) 2017年03月14日 09:22 752



社团划分——Fast Unfolding算法 (http://blog.csdn.net/google19890102/article/details/486...

社团划分——Fast Unfolding算法一、社区划分问题1、社区以及社区划分在社交网络中，用户相当于每一个点，用户之间通过互相的关注关系构成了整个网络的结构，在这样的网络中，有的用户之间的连接较为...



1



 google19890102 (<http://blog.csdn.net/google19890102>) 2015年09月22日 20:03  13275

---