

李博Garvin的专栏

阿里云机器学习PD

统计

主要讲述算法和业务的结合，适合初学者

京东地址

作者公众号：凡人机器学习

厂家直销

86型万能底盒

PVC 环保材质 家装 品质保障

¥0.40/个

个人资料

李博Garvin

博客专家

原创	粉丝	喜欢	评论
237	1047	471	464

等级：博客 7

访问量：85万+

积分：1万+

排名：1682

[置顶] 20行代码实现电影评论情感分析

标签：TensorFlow 深度学习 机器学习

2018年03月09日 09:33:17

2339人阅读

评论(4)

收藏

0

分类：机器学习 (39)

版权声明：本文为博主原创文章，未经博主允许不得转载。
http://blog.csdn.net/gshengod/article/details/79492975

背景

情感分析有很多的应用场景，比如做一个电商网站，卖家需要时刻关心用户对于商品的评论是否是正面的。再比如做一个电影的宣传和策划，电影在键盘侠们中的口碑也至关重要。互联网上关于任何一个事件或物品都有可能产生成千上万的文本评论，如何定义每一个文本的情绪是正面或是负面的，是一个很有挑战的事情。挑战体现在以下几个方面，区别于结构化数据，评论数据的长短不一，很难限定到固定的维度。另外很难通过某个词判断用户的情绪，举个极端的例子，fu*k通常被认为是贬义词，但是如果一条评论是“傲海 is fu*kinghandsome”，那么其实就是一个正向的意义。本文我要使用IMDB数据集，通过20行的Tensorflow代码实现一个电影评论预测模型，准确率可以超过百分之九十五，里面涉及到一些词袋模型以及embedding的概念，这些我等下一篇文章再讲，这一次我们就讲实践。

数据

介绍下本文用到的训练数据IMDB，完全开源的一个电影评价数据集，有好几万条真实的电影评论数据。

每一个数据都被储存为txt文件存放，其中pos文件夹就存放正面评论的数据，neg文件夹存放负面情绪的数据。这些数据可以通过一些简单的代码提取出来，并且标记。数据集下载地址：
<http://ai.stanford.edu/~amaas/data/sentiment/>

简单来说，在数据预处理阶段，需要把这些文本按照正向和负向打标，并且把文本向量化，比如“Aohai is fuc*inggentle”这种话要怎么变成数值向量，下一篇文章会具体说明。


0





阅读全文

• 上一篇 聊聊互联网实习相关的事儿



一点点加盟多



餐饮收银系统



8G内存条价格



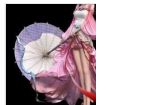
迈巴赫s600报



炒股软件 排行



非全日制研究



网络游戏

您还没有登录,请[\[登录\]](#)或[\[注册\]](#)

查看评论



xhtmlmy

2楼 2018-03-09 22:04发表

感谢分享，学习一下



SOAFR

1楼 2018-03-09 14:20发表

你文件里面的DS.store是什么

查看更多评论

电影评论人名抽取与情感分析



yinglish_ 2017年09月16日 20:59 1526

电影评论的数据挖掘是我一直以来的项目。最近的刚完成的一个工作于是分析影评中表露出来的对演员角色的情感倾向。主要的工作便是识别评论中人名+情感分析。评论人名抽取与情感分析影迷用户的电影评论中往往包含大量...

kaggle之电影文本情感分类



lijingpengchina 2016年08月19日 14:40 1500

电影文本情感分类 Github地址 这个任务主要是对电影评论文本进行情感分类，主要分为正面评论和负面评论，所以是一个二分类问题，二分类模型我们可以选取一些常见的模型比如贝叶斯、逻辑回归等，这里挑战之...

文章搜索

博客专栏



Cocos2d实例教程

文章：8篇
阅读：23555



android-tips

文章：20篇
阅读：100311



机器学习算法-python实现

文章：14篇
阅读：129353



git学习笔记

文章：5篇
阅读：8021



LeetCode从零单排

文章：31篇
阅读：39438



机器学习实践

文章：12篇
阅读：27217

文章分类

- linux (11)
- c语言 (2)
- java (49)
- c# (12)
- 百度地图api (5)
- 学习 笔记 (63)
- web 互联网 (3)
- android 开发 (25)
- DataMining (28)
- Cocos2d实例教程 (8)
- J2EE-ssh (2)
- 算法与数据结构 (47)
- JDBC (3)
- 开源夏令营 (13)
- python (16)
- git (5)
- 面试 (6)
- SQL (3)
- Hadoop (1)
- 分布式计算 (3)
- shell (1)
- 机器学习 (40)

阅读排行

【机器学习算法-python实现】...	(20320)
Android系统截屏的实现（附...	(17915)
android tesseract-ocr实例教程...	(17472)
【机器学习算法-python实现】...	(17158)
【机器学习算法-python实现】...	(16853)
【机器学习算法-python实现】...	(13389)
【机器学习算法-python实现】...	(12218)
c#中WebBrowser控件的使用方法	(11969)
【android4.3】记一次完整的a...	(11668)
新闻个性化推荐系统(python)-...	(11544)

文章存档

2018年3月

(1)

2018年2月

(4)

2018年1月

(3)

2017年12月

(4)

2017年11月

(3)

展开

车展



联系我们



请扫描二维码联系客服
webmaster@csdn.net
400-660-0108
QQ客服 客服论坛

关于 招聘 广告服务 百度

©1999-2018 CSDN版权所有
京ICP证09002463号

经营性网站备案信息

网络110报警服务

中国互联网举报中心

北京互联网违法和不良信息举报中心

程序员不会英语怎么行？

老司机教你一个数学公式秒懂天下英语

酒店情感分析

Nicholas_Liu2017 2017年06月27日 20:49 对于预料集，可以从这个数据堂拿数据：http://www.datatang.com/ 其中关于酒店情感数据分析部分：http://www.datatang.com/data/shop-page.ht...

基于R语言对用户评论进行情感分析

m0_37553060 2017年08月18日 16:29 利用机器学习可以很方便的做情感分析。本篇文章将介绍在R语言中如何利用机器学习方法来做情感分析。在R语言中，由TimothyP.Jurka开发的情感分析以及更一般的文本挖掘包已经得到了很好的发展。你可以...

【python 自然语言处理】对胡歌【猎场】电视剧评论进行情感值分析

主要内容1、什么是SnowNLP? 2、如何进行情绪判断? 3、猎场豆瓣热门短评抓取。4、猎场热门短评情感分析。什么是SnowNLP? snowNLP，可以分词，标注，还可以进行情绪分析。Sn...

恭喜：一个公式教你秒懂天下英语

老司机教你一个数学公式秒懂天下英语



电影评论情感分析 keras

u012969412 2017年05月02日 15:19 1198 #-*- coding:utf-8 -*- import re import numpy as np from keras.models import Sequential from keras.la...

阅读：面向电子商务的评论文本情感分析研究

评论文本的情感分析（倾向性分析、观点识别、意见挖掘、情感分类）两种方法：基于语义，基于机器学习。基于语义：国外：（支持向量机（SVM）和贝叶斯算法具有较好的情感分类效果）。国内：hownet词...

【深度学习与Theano】LSTM网络-情感分析

UncleMuses 2016年01月21日 23:46 4944 概述本文提供一个如何使用Theano实现使用LSTM结构的RNN例子。模型被用来通过电影综述来进行情感分析，综述数据来源于Large Movie Review Dataset，即IMDB。在这个...

商品评论中的实体情感分析

chixujohnny 2017年01月04日 00:03 1405 在之前的文章中我已经将商品评论中的评价实体，也就是特征（feature）给挖掘出来了，我看了一下，挖掘的效果还是不错的。原文链接：http://blog.csdn.net/chixujohnny/ar...

CNN情感分析（文本分类）

qq575379110 2017年05月10日 22:58 3215

这篇文章翻译至denny britz的博客，本来想自己用TensorFlow实现CNN情感分析，看过denny的github后，就决定不浪费时间了。当然，会在他的基础上做一些改进。 一、数据预处理 ...

IMBD影评数据的情感分析



PKU_ZZY

2017年01月30日 12:32

📖 986

IMBD影评数据的情感分析



0



Python语言编程入门

python教程

百度广告

Kaggle情感分析(Bag of Words Meets Bags of Popcorn)

Kaggle上有两个情感分析的任务，这是其中的一个Bag of Words Meets Bags of Popcorn。任务虽然已经结束了，但是

还有点研究的意义。 <https://www.kaggl...>



taneijia

2016年01月08日 14:29

📖 4043

Python微博评论进行情感分析



hzp666

2018年01月04日 11:31

📖 809

利用500W条微博语料对评论进行情感分析 布道 3 个月前 本文已在CSDN，CSDN微博， CSDN公众号， IT技术之家 等平台转发。最近身边的人都谈论一件事：10...

基于电商消费者评论的商品属性抽取及情感分析

基于电商消费者评论的商品属性抽取及情感分析 Chao Liu (lch614730@163.com)



NUSTM

2015年01月22日 20:49

📖 868

kaggle 电影评论情感分析 贝叶斯分类



zn505119020

2017年09月12日 18:10

📖 120

```
import pandas as pd from sklearn.feature_extraction.text import CountVectorizer from sklearn.naive_b...
```

程序员不会英语怎么行？

老司机教你一个数学公式秒懂天下英语



电商产品评论数据情感分析代码详解



appleyuchi

2017年06月16日 17:49

📖 2009

本章代码建议在linux下面运行 下面代码的意思是从评论数据中抽取品牌是美的的数据 (15-1) #-*- coding: utf-8 -*- import pandas as pd inp...

中文情感分析语料库



noter16

2017年07月19日 10:32

📖 4914

中文情感分析语料库

情感分析系统的预处理工作



ccg159

2013年10月07日 14:11

📖 935

老板给了个情感分析系统的任务，一开始各种纠结不知道怎么入手。再三拖延后看了篇论文，知道了有个SVMLight的现成系统可以用，但是要把文件处理成它规定的格式，于是就是这篇文章的主要内容了。主要工作...

小白理解 js 20行代码实现贪吃蛇



chen_mo

2017年11月18日 21:54

328

小白理解 js 20行代码实现贪吃蛇 偶然发现网上很多20行js实现贪吃蛇游戏的代码，但多数没有注释，作为一个新手非常痛苦，于是想试着理解，经历了数个小时后，终于摸到了一丝皮毛，故分享出来。 ...



0



[代码的艺术]20行代码的贪吃蛇



qq_33094497

2016年09月21日 10:33

18

[代码的艺术]20行代码的贪吃蛇

