

CSDN新首上线啦，邀请你来立即体验！(http://blog.csdn.net/)

立即体验

CSDN

博客 (http://blog.csdn.net/?ref=toolbar) 学院 (http://edu.csdn.net/?ref=toolbar)

下载 (http://download.csdn.net/?ref=toolbar) 更多

搜索

登录 (https://passport.csdn.net/account/login?ref=toolbar) 注册 (http://passport.csdn.net/account/mobile/register?ref=toolbar&action=mobileRegister) (activity?utm_source=csdnblog1)

1

1

1

1

1

SyntaxNet安装与使用详解

原创 2016年12月02日 19:45:30

标签: [syntaxnxt \(http://so.csdn.net/so/search/s.do?q=syntaxnxt&t=blog\) /](#) [dependency \(http://so.csdn.net/so/search/s.do?q=dependency&t=blog\) /](#) [parser \(http://so.csdn.net/so/search/s.do?q=parser&t=blog\) /](#) [postag \(http://so.csdn.net/so/search/s.do?q=postag&t=blog\)](#)

1699

安装SyntaxNet的工作环境，遇到很多麻烦，到后来发现如果不走弯路，其实并不麻烦。

1 我最开始是安装anaconda。（python2.7版本）
[https://www.continuum.io/downloads \(https://www.continuum.io/downloads\)](#)
这个包安装完后python pip numpy都解决了，而且还不用改环境变量。

2 然后用pip 安装mock和asciitree和protobuf
用apt-get安装swig

3 当然tensorflow
pip install [https://storage.googleapis.com/tensorflow/mac/tensorflow-0.5.0-py2-none-any.whl \(https://storage.googleapis.com/tensorflow/mac/tensorflow-0.5.0-py2-none-any.whl\)](#)

上面这些都很easy
其中比较难解决的是bazel，
具体看这个[https://bazel.build/versions/master/docs/install.html \(https://bazel.build/versions/master/docs/install.html\)](#)

先去github把它的安装文件下载下来
[https://github.com/bazelbuild/bazel/releases \(https://github.com/bazelbuild/bazel/releases\)](#)
1. Install JDK 8
`sudo add-apt-repository ppa:webupd8team/java` `sudo apt-get update`
`$ sudo apt-get install oracle-java8-installer`
Note: You might need to `sudo apt-get install software-properties-common` if you don't have the `add-apt-repository` command. See here.

Ubuntu Wily (15.10). To install OpenJDK 8:
`$ sudo apt-get install openjdk-8-jdk`

1. Install other required packages
`$ sudo apt-get install pkg-config zip g++ zlib1g-dev unzip`

2. Download Bazel
Download the Bazel installer for your operating system.

3. Run the installer
Run the installer:

`chmod +x bazel-version-installer-os.sh` `./bazel-version-installer-os.sh -user`
The `-user` flag installs Bazel to the `HOME/bin` directory on your system and sets the `bazelrc` path to `HOME/.bazelrc`. Use the `-help` command to see additional installation options.

1. Set up your environment

If you ran the Bazel installer with the `-user` flag as above, the Bazel executable is installed in your `$HOME/bin` directory. It's a good idea to add this directory to your default paths, as follows:

`export PATH="$PATH:$HOME/bin"`

trayfour (http://blog.csdn.net/u014422406)

码云未开通

32 13 1

他的最新文章

更多文章 (http://blog.csdn.net/u014422406)

统计不同时间段在线人数，给定若干（进入，离开）时间对 (http://blog.csdn.net/u014422406/article/details/78089376)

SVM一些问题及思考 (http://blog.csdn.net/u014422406/article/details/77931999)

pytorch版本回退，从0.2降到0.1.12 (http://blog.csdn.net/u014422406/article/details/77918609)

Unable to Connect

The Proxy was unable to connect to the remote site. responding to requests. If you feel you have reached please submit a ticket via the link provided below.

URL: [http://pos.baidu.com/s?hei=250&wid=300&di=u%2Fblog.csdn.net%2Fu014422406%2Farticle%2Fdetail%2F77918609](#)

在线课程

在线课程

问题引导系列之detail

WebAssembly中的实践 (http://edu.csdn.net/huiyiCourse/detail/602?utm_source=blog9)

Apache WebAssembly 移动研

发的进阶之路 (http://edu.csdn.net/huiyiCourse/detail/602?utm_source=blog9)

第1页 共5页

2017/11/22 下午6:03

You can also add this command to your ~/.bashrc file.

```
最后就clone工程
git clone --recursive https://github.com/tensorflow/models.git (https://github.com/tensorflow
/models.git)
cd models/syntaxnet/tensorflow
./configure
cd ..
bazel test syntaxnet/... util/utf8/...
```

保持联网的情况下，10几分钟之后会出现 pass 17test。（ps
我是在windows下连的服务器，不能联网所以用了ccproxy，而且里面
python，bazel，swig，numpy，numpy都装好了，我只安装了anaconda，覆盖服务器的python版本）

然后说一下使用的一些经历。
数据的话http://cs.stanford.edu/~danqi/data/PTB_SD_3_3_0.zip (http://cs.stanford.edu/~danqi/data/PTB_SD_3_3_0.zip)这个ok
不过里面的gold可以用，不带gold的那组数据空行中有个制表符，所以会报错Check failed: fields.size() >= 8 (2 vs. 8)Every line has to have at least 8 tab separated fields.
，所以如果你要用不带gold的那组数据的话，要把每句话之间的空行中的\t去掉。

有了这个数据
1 先跑pos tag
而syntaxnet有个context.pbtxt这个文件你可以把它理解成，训练过程中的数据的路径名字都被写在这里，
你只需要把这些路径和名字对上号，在下面输入commad的时候，直接用名字，
所以我把context.pbtxt中的training-corpus，和tuning-corpus的所在路径改成上面下载数据所在的路径。
所以下面的commad都不用变。最后那个-params=128.什么的表示训练的参数会保存到
128-0.08-3600-0.9-0 的文件夹
bazel-bin/syntaxnet/parser_trainer --task_context=syntaxnet/context.pbtxt --arg_prefix=brain_pos
--compute_lexicon --graph_builder=greedy --training_corpus=training-corpus --tuning_corpus=tuning-
corpus --output_path=models --batch_size=32 --decay_steps=3600 --hidden_layer_sizes=128
--learning_rate=0.08 --momentum=0.9 --seed=0 --params=128-0.08-3600-0.9-0

这个程序
task_context=syntaxnet/context.pbtxt 表示训练所需的数据的路径都被写在这个文件中
arg_prefix=brain_pos 我训练的是pos tag，如果是parse则要改为brain_parser
-compute_lexicon 应该是表示根据数据统计词频和词典的意思
-graph_builder=greedy 训练时greedy还是beam
output_path=models训练中的输出都会保存到models下面的brain_pos/greedy/128-0.08-3600-0.9-0 中，
包括词典 什么index 2 word word 2 index。

这个pos tag的训练过程1个小时左右，最后的正确率96%多

然后用这个pos tag为数据打上tag，但其实我们的数据已经有tag了。
这一步是用那个bazell eval命令

2 所以直接可以用这个数据来做dependency parsing。为什么不能直接训练parsing model呢？
因为在训练parsing的过程中需要用到前面训练pos tag的中间文件，这些文件都被保存在
models/brain_pos/greedy/128-0.08-3600-0.9-0 中就是一些label index的映射文件。
ps这个目录下还有个context是后面会用到的，它和前面那个context.pbtxt是一个东西的，只是他把这些中
间文件的路径也写入到context的文件中，也就是后面我们用这个context就可以找到postag所生成的中间文件
的位置。

接下来local training
bazel-bin/syntaxnet/parser_trainer \
--arg_prefix=brain_parser \
--batch_size=32 \
--projectivize_training_set \
--decay_steps=4400 \
--graph_builder=greedy \
--hidden_layer_sizes=200,200 \
--learning_rate=0.08 \
--momentum=0.85 \
--output_path=models \
--task_context=models/brain_pos/greedy/128-0.08-3600-0.9-0/context \
--seed=4 \
--training_corpus=training-corpus \

热门文章

sigmoid和softmax总结 (<http://blog.csdn.net/u014422406/article/details/52805924>)
👤 43531

推荐给初学LSTM或者懂个大概却不完全懂的人 (<http://blog.csdn.net/u014422406/article/details/52806430>)
👤 3040

SyntaxNet安装与使用详解 (<http://blog.csdn.net/u014422406/article/details/53438389>)
👤 1679

知其然知其所以然之——word2vector (<http://blog.csdn.net/u014422406/article/details/52815377>)
👤 1080

word2vector 一些问题及思考 (<http://blog.csdn.net/u014422406/article/details/71036327>)
👤 982



```
-tuning_corpus=tuning-corpus \  
-params=200x200-0.08-4400-0.85-4
```

```
-task_context=models/brain_pos/greedy/128-0.08-3600-0.9-0/context  
这句就表示我要用postag生成的context作为我的文档接口，通过它我能找到需要的文件  
-training_corpus=training-corpus \  
-tuning_corpus=tuning-corpus \  
之所以不是tagged_training_corpus=training-corpus \  
是因为前面说过我们的数据是带tag的所以不用再用postag 标记  
训练出的model会保存到models/brain_parser/greedy/200x200-0.08-4400-0.85-4
```

```
3 gobal training  
bazel-bin/syntaxnet/parser_trainer \  
-arg_prefix=brain_parser \  
-batch_size=32\ 记得beam改成32  
-decay_steps=100 \  
-graph_builder=structured \ 这个参数具体作用不太清楚，但是models/parser下会有个  
structured的文件，这和greedy是等价的，它里面有gobal的model  
-hidden_layer_sizes=200,200 \  
-learning_rate=0.02 \  
-momentum=0.9 \  
-output_path=models \  
-task_context=models/brain_parser/greedy/200x200-0.08-4400-0.85-4/context \  
//会用到loacl train生成的文件 所以用loacl train生成的context，和前面道理一样  
-seed=0 \  
-training_corpus=projectivized-training-corpus \ //这个是loacl生成的文件，上面的上面那行context中有  
他的具体路径  
-tuning_corpus=tuning-corpus \ //我们的tuning文件也tag过，不用再加入个tagged文件  
-params=200x200-0.02-100-0.9-0 \  
-pretrained_params=models/brain_parser/greedy/200x200-0.08-4400-0.85-4/model \  
-pretrained_params_names=\  
embedding_matrix_0,embedding_matrix_1,embedding_matrix_2,\  
bias_0,weights_0,bias_1,weights_1
```

如此训练完成。

此外如果你想修改这个代码的某一部分。
其中还是在你运行commad的文件下，也就是bazel-bin所在的文件夹下面的syntaxnet中包含了其中的大部分代码。有c++有python，如果运行有错，可以找到有错的文件，修改后。
其中有个很重要的BUILD文件
其中有 cc libraries
cc tests
py graph builder and trainer等阻块。
比如你修改了cc文件，
使用bazel build : parser_ops.so（冒号后面的表示BUILD文件中的那些name表示你要重新build的组块名称）
它就会把cc这些相关的文件重新编译，其中每个name都有个deps与他对应，表示重新编译这部分的内容所依赖的那些文件。

bazel build相当于以前的make makefile编译。
优点是你只需重新编译你修改的那个文件所依赖的那些文件而不是整个工程。

大致就这么多了，希望各位少走弯路，对你能够起到帮助，笔者比较懒，所以博客组织比较乱，望见谅。


版权声明：本文为博主原创文章，未经博主允许不得转载。



相关文章推荐


SyntaxNet是否可以实现中文语义分析 (http://blog.csdn.net/xu_erni/article/details/54405...)

最近查看了很多SyntaxNet方面的知识，在GitHub上也下载了中文语料库，并成功训练了自己的中文模型，因为官网没有关于中文模型的训练方法，于是便参照别人训练英文的方式进行中文的训练。在生成的模型...

 xu_erni (http://blog.csdn.net/xu_erni) 2017年01月13日 09:39 12857

如何使用 SyntaxNet进行中文模型的使用 (http://blog.csdn.net/xu_erni/article/details/539...

使用Chinese模型 1.在网上下载Chinese模型文件，网址：http://download.tensorflow.org/models/parsey_universal/Chinese.zi...

 xu_erni (http://blog.csdn.net/xu_erni) 2016年12月29日 10:29 3611



成为前端高手的路上，我避开了哪些坑？


大学毕业后我成为前端开发者，从一开始的小白到现在的“高手”，我把一些感想记录下来...

广告

(http://www.baidu.com/cb.php?c=lgF_pyfqHmknj0dP1f0IZ0qnfK9ujYzP1nsrjD10Aw-5Hc3rHnYnHb0TAq15HfLPWRznjb0T1YyuckP1fLmH6vujT1n1IW0AwY5HDdnHD1Pjf4nWT0lgF_5y9YIZ0lQzq-uZR8mLPbUB48ugfElAqspynElvNBnHqdlAdxTvqdThP-5yF_UvTkn0KZujY4rHb0mhYqn0KsTWYs0ZNGujYkPHTYn1mk0AqGujYknWb3rjDY0APGujYLnWm4n1c0ULI85H00TZbqnW0v0APzm1Y1PWcd)

如何使用 SyntaxNet进行中文模型的使用 (http://blog.csdn.net/jiangjingxuan/article/deta...

使用Chinese模型 1.在网上下载Chinese模型文件，网址：http://download.tensorflow.org/models/parsey_universal/Chinese.zi...

 jiangjingxuan (http://blog.csdn.net/jiangjingxuan) 2017年02月07日 09:30 874


syntaxnet python调用 (http://blog.csdn.net/hjmce/article/details/52457031)

1、编译安装syntaxnet 2、把models/syntaxnet/syntaxnet下的model模型文件复制到syntaxnet/bazel-bin里面的一个文件夹parser_eval.r...

 hjmce (http://blog.csdn.net/hjmce) 2016年09月07日 10:21 2371

NLP初级选手 ubuntu 下安装 google SyntaxNet (http://blog.csdn.net/u012507864/article/...

SyntaxNet: Neural Models of SyntaxNLP初级选手ubuntu 下安装google SyntaxNet初级选手安装走了很多弯路，但是也收获不少经验，特别是熟悉了环境搭建...

 u012507864 (http://blog.csdn.net/u012507864) 2016年05月23日 00:28 3241




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_pyfqHmknjfrjD0lZ0qnfK9ujYzP1f4PjnY0Aw-5Hc4nj6vPjm0TAq15Hf4rjn1n1b0T1YkPjD1PWf3rHTsuWf1PWIB0AwY5HDdnHD1Pjf4nWT0lgF_5y9YIZ0lQzqMpgwBUvqoQhP8QvIGIAPCmgfEmvq_lyd8Q1R4uWc4uHf3uAckPHRkPWN9Phc:5HDknWFBmhkEusKujY4rHb0mhYqn0KsTWYs0ZNGujYkPHTYn1mk0AqGujYkn10snjf10APGujYLnWm4n1c0ULI85H00TZbqnW0v0APzm1Yzrjvnvf)


Tensorflow Syntaxnet模型安装测试方法 (http://blog.csdn.net/huxuanlai/article/details/6...

#install dev env apt-get update apt-get install build-essential vim -y #installjdk8 apt-get insta...

 huxuanlai (http://blog.csdn.net/huxuanlai) 2017年03月13日 11:21 722


基于 Syntaxnet的中文分词问题 (http://blog.csdn.net/Nancy_xz/article/details/52981364)

第一次开通博客，尤其最近发现基于Syntaxnet的中文分词资料还挺少的，开个博客，大家可以互相交流交流。...

 Nancy_xz (http://blog.csdn.net/Nancy_xz) 2016年10月31日 12:23 1067


Install tensorflow/SyntaxNet on Ubuntu VM (http://blog.csdn.net/u010157717/article/det...

1 install Ubuntu 14.04 by virtualbox 1.1 install virtualbox https://www.virtualbox.org/wiki/Downl...

 u010157717 (http://blog.csdn.net/u010157717) 2016年06月25日 02:26 971

自然语言处理（NLP）常用开源工具总结----不定期更新 (http://blog.csdn.net/yunzhong...

学习自然语言这一段时间以来接触和听说了好多开源的自然语言处理工具，在这里做一下汇总方便自己以后学习，其中有自己使用过的也有了解不是很多的，对于不甚了解的工具以后学习熟悉了会做更新的。1.iKAna...

 yunzhongfeiniaio (http://blog.csdn.net/yunzhongfeiniaio) 2013年11月14日 20:07 17594



跟着猫哥学Golang[6] - 结构体 (http://blog.csdn.net/972301/article/details/49421227)

Golang变量声明，两种办法： 1. 使用前声明：方法： var 变量名 变量类型，例如： var i,j int package main import "fmt" var x, y, z...

972301 (http://blog.csdn.net/972301) 2015年10月26日 10:57 344

深度解读谷歌开源的最精确自然语言解析器SyntaxNet (/starzhou/article/details/514082...

深度解读谷歌开源的最精确自然语言解析器SyntaxNet 谷歌开源了SyntaxNet，也发布了针对英语的预训练解析程序 Pars ey McParseface...

starzhou (http://blog.csdn.net/starzhou) 2016-05-14 21:54 6464

Announcing SyntaxNet: The World's Most Accurate Parser Goes Open Source (/yeshen...

SyntaxNet github: https://github.com/tensorflow/models/tree/master/syntaxnet from: http://googleres...

yeshennet (http://blog.csdn.net/yeshennet) 2016-05-14 10:40 350



AjaxControlToolkit的安装与使用详解 (http://download.csdn.net/detail/...

http://download... 2013-12-09 08:34 51KB



MySQL安装配置和使用详解 (http://download.csdn.net/detail/taomanm...

http://download... 2017-05-13 09:18 3.64MB

LVS系列—LVS安装使用详解（ipvsadm命令） (/u012758088/article/details/76039493)

一、简介 LVS是Linux Virtual Server的简称，也就是Linux虚拟服务器,是一个由章文嵩博士发起的自由软件项目，它的官方网站是www.linuxvirtualserver.or...

u012758088 (http://blog.csdn.net/u012758088) 2017-07-24 23:28 180



EZ Publish安装包的使用详解 (http://download.csdn.net/detail/zzbaikec...

http://download... 2011-07-20 17:20 3.41MB



详解oracle9i Statspack安装&使用 (http://download.csdn.net/detail...

http://download... 2010-09-14 19:05 30KB

Docker的安装配置及使用详解 (/fgf00/article/details/51893771)

基本概念Docker 包括三个基本概念镜像（Image）容器（Container）仓库（Repository）先理解了这三个概念，就理解了 Docker 的整个生命周期。1、docker安装与...

fgf00 (http://blog.csdn.net/fgf00) 2016-07-12 23:58 50719



VSS的安装使用详解 (http://download.csdn.net/detail/cn_player/2673975)

http://download... 2010-09-06 11:32 8KB



PC-LINT安装与使用详解 (http://download.csdn.net/detail/xiaoshun123/...

http://download... 2010-07-21 16:33 1.31MB