

RNN的高级应用



不会停的蜗牛 (/u/7b67af2e61b3) [+ 关注](#)

2016.09.30 10:44* 字数 1657 阅读 2955 评论 0 喜欢 2 赞赏 1

(/u/7b67af2e61b3)

本文结构：

四个问题

1. 每个问题是什么
2. 应用什么模型
3. 模型效果

CS224d-Day 11:

Recursive neural networks -- for different tasks (e.g. sentiment analysis)

课程链接 ([https://link.jianshu.com?t=https://web.archive.org/web/20160314075834/http://cs224d.stanford.edu/syllabus.ht](https://link.jianshu.com?t=https://web.archive.org/web/20160314075834/http://cs224d.stanford.edu/syllabus.html)

[ml\)](https://web.archive.org/web/20160314075834/http://cs224d.stanford.edu/syllabus.html)

视频链接 ([https://link.jianshu.com?t=https://www.youtube.com/watch?](https://link.jianshu.com?t=https://www.youtube.com/watch?v=24FQOQMcoIY&list=PLIJy-eBtNFt4CSVWYqscHDdP58M3zFHIG&index=11)

[v=24FQOQMcoIY&list=PLIJy-eBtNFt4CSVWYqscHDdP58M3zFHIG&index=11\)](https://link.jianshu.com?t=https://www.youtube.com/watch?v=24FQOQMcoIY&list=PLIJy-eBtNFt4CSVWYqscHDdP58M3zFHIG&index=11)

课件链接 ([https://link.jianshu.com?](https://link.jianshu.com?t=https://web.archive.org/web/20160313081419/https://cs224d.stanford.edu/lectures/CS224d-Lecture10.pdf)

[t=https://web.archive.org/web/20160313081419/https://cs224d.stanford.edu/lectures/CS224d-Lecture10.pdf\)](https://link.jianshu.com?t=https://web.archive.org/web/20160313081419/https://cs224d.stanford.edu/lectures/CS224d-Lecture10.pdf)

四个问题



这次课主要讲了标准的 Recursive neural networks 模型及其扩展模型在3个问题上的应用和效果，最后的 Tree LSTM 简单地介绍了模型和效果。

这3个问题分别是 Paraphrase detection , Relation classification , Sentiment Analysis。每个模型都可以应用到任意一个问题上，只不过效果不同，有些模型对一些问题表现会更优一些。

- | | |
|--------------------------------------|-------------------------|
| 1. Standard RNNs: | Paraphrase detection |
| 2. Matrix-Vector RNNs: | Relation classification |
| 3. Recursive Neural Tensor Networks: | Sentiment Analysis |
| 4. Tree LSTMs: | Phrase Similarity |
-

1.Paraphrase detection

目的是判断两句话是否具有相同的意思

用到的模型是标准的RNN

为了解决这个问题，需要思考：

- 怎样比较两个句子？

通过成对地比较两个句子的短语，这个时候可以用 standard RNN，因为它可以得到一个合理的树结构，也就是句子的短语组成结构。

- 怎样用相似度来判断两个句子的意义是一样的？



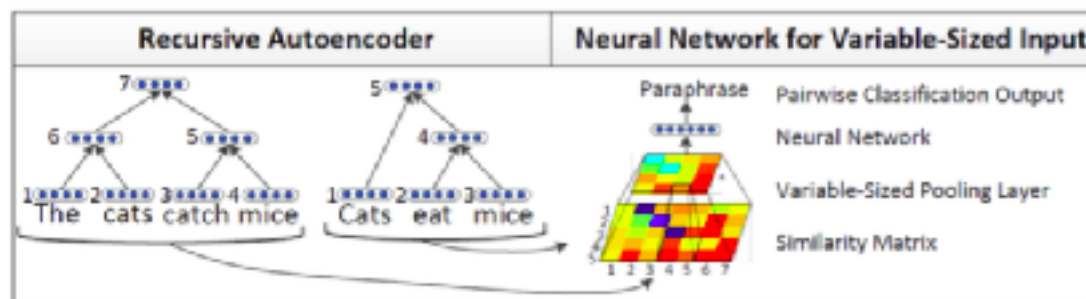
如果用两个树结构的顶点去判断，那会丢掉很多中间环节的信息。

如果只计数两个句子中相似短语的个数，那么会丢掉位置信息，即这些短语出现在什么位置。

所以用 **similar matrix** 来表示相似度。

如下图，左边是两个句子，树结构中分别有7个部分和5个部分，右边是由similar matrix 到最后结果的过程。

similar matrix 由5行7列组成，颜色深浅表示两个树结构相应部分间的相似度大小。



- 上图中，为什么不能直接把 similar matrix 直接投入神经网络中？

因为这个矩阵的大小是随着输入句子的不同而变化的。

所以需要引用一个 **pooling 层**，它可以将输入的 similar matrix 映射成维度一致的矩阵，然后再投入到 RNN 中。

最后的效果：



Method	Acc.	F1
Rus et al.(2008)	70.6	80.5
Mihalcea et al.(2006)	70.3	81.3
Islam et al.(2007)	72.6	81.3
Qiu et al.(2006)	72.0	81.6
Fernando et al.(2008)	74.1	82.4
Wan et al.(2006)	75.6	83.0

2.Relation Classification

目的是识别词语之间的关系

尤其是 very + good 这种，very 此时只是一个加强 good 的作用

用到的模型是标准的 Matrix - Vector RNN

普通的 RNN 中，每个节点都是由向量表示的，在这个 Matrix - Vector RNN 中，每个节点除了向量外自带一个矩阵，在由 left 和 right child 生成 parent 的时候，对彼此作用各自的矩阵后，再去生成 parent。



- 用向量和矩阵的区别？

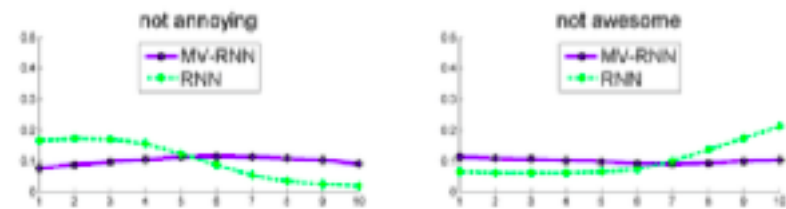
单独的向量反映不出相互作用这种层次的意义，加上矩阵作用给对方后，可以显示出 very 可以让 good 这种形容词更强的意义。

矩阵是随机初始化的，通过 Back Propagation 和 Forward Propagation 可以不断地学习出来。

最后的效果：

下图中，横轴是 1-10 星号的电影，纵轴是 not annoying 这样的词出现在相应级别中的比例。

在 not annoying，not awesome 这两个例子中，绿色的 RNN 没有蓝色的 MV - RNN 表现得好，因为 not annoying 出现在低星级的次数不应该比出现在高星级电影中的次数多。



另外一种问题是因果关系等的判断



最后的效果：

对于这个问题，用了不同的模型和feature来看效果。

在 SVM 用了好多feature，最后的效果是 82.2

POS : part of speech

wordnet 大量人工生成的数据

prefix 等其他形态学的特征

dependency parse feature 不同类型的parser

textrunner 百万的网上数据

Google n - gram 几十亿个 n - gram

单纯用神经网络模型，数据量没那么大的时候，效果不到80%

加入了 POS，WordNet，NER 数据后，变成了 82.4，优于SVM。

数据越多的话，效果越好。



3.Sentiment Analysis

目的是识别句子表达的情感

用到的模型是RNTN (Recursive Neural Tensor Network)

- 用 Bag of words 这种方法有缺陷：

一个 not 后面多个 positive 的词时，应该是否定，结果被判断成肯定。

前半句否定，后半句肯定，后半句的效果比前半句更强的时候，怎么判断出来。

解决方案，一个是更好的数据，一个是更好的模型

更好的数据：

人工标注 11,855 个句子的 215,154 个短语，每个短语由不同的人标注 3 次。

下图是标注结果的可视化，横轴是短语的长度，纵轴是各个情感类别的比例。



最后的效果：

可以发现，用新的 tree bank 的模型效果要比原来的好，肯定否定情感分类越准。

更好的模型：

RNTN (Recursive Neural Tensor Network)

这个模型可以让 word 之间有更多的 interaction，‘very good’的词向量的转置和矩阵 V 再和词向量本身作用。



最后的效果：

RNTN 作用在新的 Tree Bank 上效果可以高达 85.4.



4.Semantic Similarity

目的是识别语义相似性

用到的模型是 Tree LSTMs

Tree LSTMs 和普通的 LSTMs 的不同之处在于 Tree LSTMs 是从 tree 的结构中进行 LSTMs 的建模。

parent 的 hidden层是其 children 的 hidden 层的和，每一个 forget unit 是根据具体的某个节点来计算的，计算最终 cell 时要把所有 forget units 和对应的 cells 相乘并求和，其他部分和普通LSTMs计算方法一样。



最后的效果：



[cs224d]

Day 1. 深度学习与自然语言处理 主要概念一览

(<https://www.jianshu.com/p/6993edef96e4>)

Day 2. TensorFlow 入门 (<https://www.jianshu.com/p/6766fbc43b9>)

Day 3. word2vec 模型思想和代码实现 (<https://www.jianshu.com/p/86134284fa14>)

Day 4. 怎样做情感分析 (<https://www.jianshu.com/p/1909031bb1f2>)

Day 5. CS224d - Day 5: RNN快速入门 (<https://www.jianshu.com/p/bf9ddfb21b07>)

Day 6. 一文学会用 Tensorflow 搭建神经网络

(<https://www.jianshu.com/p/e112012a4b2d>)

Day 7. 用深度神经网络处理NER命名实体识别问题

(<https://www.jianshu.com/p/581832f2c458>)

Day 8. 用 RNN 训练语言模型生成文本 (<https://www.jianshu.com/p/b4c5ff7c450f>)

Day 9. RNN与机器翻译 (<https://www.jianshu.com/p/23b46605857e>)

Day 10. 用 Recursive Neural Networks 得到分析树

(<https://www.jianshu.com/p/403665b55cd4>)

Day 11. RNN的高级应用 (<https://www.jianshu.com/p/0e840f92b532>)


我是 不会停的蜗牛 Alice

85后全职主妇

喜欢人工智能，行动派

创造力，思考力，学习力提升修炼进行中

欢迎您的喜欢，关注和评论！

 技术博文 (/nb/5173140)

举报文章 © 著作权归作者所有

不会停的蜗牛 (/u/7b67af2e61b3)

写了 224835 字，被 3243 人关注，获得了 1969 个喜欢





我是 Alice 喜欢人工智能，行动派 创造力，思考力，学习力提升修炼进行中 欢...

(/u/7b67af2e61b3)

+ 关注

♡ 喜欢 (/sign_in?utm_source=desktop&utm_medium=not-signed-in-like-button) | 2



更多分享

(http://cwb.assets.jianshu.io/notes/images/6073913/weibo/image_2)



(/apps/download?utm_source=nbc)

被以下专题收入，发现更多相似内容



坚持写作100天 (/c/6f43264f8299?

utm_source=desktop&utm_medium=notes-included-collection)



数据科学家 (/c/0adc32d3cf07?utm_source=desktop&utm_medium=notes-

included-collection)



NLP (/c/173cbecfcfc2?utm_source=desktop&utm_medium=notes-

included-collection)

(/p/78f087804861?

utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendation)

词向量、句子向量、篇章向量的一些理解（转） (/p/78f087804861?utm_ca...



(转自) http://blog.csdn.net/sinat_26917383/article/details/54882554 另外相关

博客：词向量三部曲 在cips2016出来之前，笔者也总结过种类繁多，类似词向..



zoulala (/u/a1699b5ddfc1?



Figure 1. An illustration of the word mover's distance. 1-stop words (**bold**) of both documents are embedded in $rd2vec$ space. The distance between the two documents is minimum cumulative distance that all words in document 1 travel to exactly match document 2. (Best viewed in color.)

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

(/p/15411de409f1?



utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendation)

使用 TensorFlow 做文本情感分析 (/p/15411de409f1?utm_campaign=mal...

使用 TensorFlow 做文本情感分析 本文将通过使用TensorFlow中的LSTM神经网络方法探索高效的深度学习

方法。 作者： Adit Deshpande July 13, 2017 翻译来源： <https://www.oreilly.com/learning/perf...>



Datartisan数据工匠 (/u/ad75474d9e73?)

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

Deep Learning in NLP-词向量和语言模型 (/p/d46609947489?utm_campa...

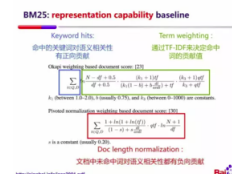
Deep Learning 算法已经在图像和音频领域取得了惊人的成果，但是在 NLP 领域中尚未见到如此激动人心的结果。关于这个原因，引一条我比较赞同的微博。@王威廉：Steve Renals算了一下icassp录取文章题目中...



stonelin3935 (/u/6c767c086b0d?

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

(/p/3a9f49834c4a?



utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendation)

浅谈智能搜索和对话式OS (/p/3a9f49834c4a?utm_campaign=maleskine&...

前面的文章主要从理论的角度介绍了自然语言人机对话系统所可能涉及到的多个领域的经典模型和基础知

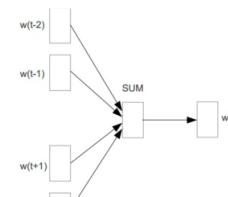
识。这篇文章，甚至之后的文章，会从更贴近业务的角度来写，侧重于介绍一些与自然语言问答业务密切相..



我偏笑_NSNirvana (/u/2293f85dc197?

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

(/p/6993edef96e4?



utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendation)

深度学习与自然语言处理 主要概念一览 (/p/6993edef96e4?utm_campaign...

CS224d - Day 1: 要开始系统地学习 NLP 课程 cs224d，今天先来一个课程概览。课程一共有16节，先对每一节中提到的模型，算法，工具有个总体的认识，知道都有什么，以及它们可以做什么事情。简介：1....



不会停的蜗牛 (/u/7b67af2e61b3?

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

(/p/7e302f881ce6?



utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendation)

你的欲望有多大，成功就会有多快（二） (/p/7e302f881ce6?utm_campaig...

之前一篇文章谈到过欲望对于人成功的重要性，主要说到的是创业需要有的心态，也就是强烈的目标性，皆下来的8篇文章，我将会对创业还需要关注的其他方面做出诠释，供各位想创业的小伙伴们参考，毕竟考虑...



比你高一个头 (/u/7097171344c1?

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

从今天起，做一个会赚钱的人（深度好文） (/p/1936041162ff?utm_campai...

你身边有这种骚里骚气的人吗？经常听得到他们说“我不爱钱”，一副对钱无所谓的态度，实际上比任何一个人都爱钱，骨子里锱铢必较，巴不得占尽全世界的小便宜。上网淘个东西可以货比100家，最终为省下的五...





九尾猫猫 (/u/b2b4827f95aa?)

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

(/p/a08db6f57fab?)



utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendation)

《岁月》 (/p/a08db6f57fab?utm_campaign=maleskine&utm_content=n...

第六十一天 《岁月》 生活的中心 我不是局外人 更不是旁观者 时间碾碎的曾经 或已被掩埋或已风干 夜幕下 星点，月光如旧 世界安静的听得见呼吸 欣赏一盏路灯的寂寞 同时也寂寞着自己 往事的碎片 拼凑出一...



shufan书凡 (/u/3b710d7df787?)

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

(/p/1daa5e46a7bb?)



utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendation)

舒颜系列 (/p/1daa5e46a7bb?utm_campaign=maleskine&utm_content=n...

肌肤亚健康 "肌肤敏感"正是"肌肤亚健康"的状态之一，存在于各种肌肤类型，是你的肌肤在面对外界的压力、刺激而发出的求助。其实造成肌肤敏感的诱因有很多，包括先天的遗传或者来自外界的压力和刺激。 ...



杨书丽 (/u/8386d7685e90?)

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

(/p/63d1e48c8529?)



utm_campaign=maleskine&utm_content=note&utm_medium=seo_notes&utm_source=recommendation)



车水马龙人生百态 (/p/63d1e48c8529?utm_campaign=maleskine&utm_c...

潍坊记



SUN喜笑成风 (/u/ae1ad20f8f83?

utm_campaign=maleskine&utm_content=user&utm_medium=seo_notes&utm_source=recommendation)

