

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

CSDN

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

下载 (//download.csdn.net?ref=toolbar) GitChat (//gitbook.cn/?ref=csdn)

更多

6



登录 (https://passport.csdn.net/account/login?ref=toolbar) 注册 (http://passport.csdn.net/account/mobile/register?ref=toolbar&action=mobile) 写博客 (http://write.blog.csdn.net/entry?ref=toolbar) 上传资源 (http://source.csdnblog1.com/)

attention 机制



2016年08月01日 16:24:34

标签：深度学习 (http://so.csdn.net/so/search/s.do?q=深度学习&t=blog)



5211

attention 机制

什么是attention

attention机制是（非常）松散地基于人类的视觉注意机制。就是按照“高分辨率”聚焦在图片的某个特定区域并以“低分辨率”感知图像的周边区域的模式，然后不断地调整聚焦点。这个概念最早出现在认知心理学上面。用一个最通俗的例子来解释就是，当你认真的去做一件事情，你的注意力会被更多的分配在手上的事情，而忽略身边的其他事情；当你喜欢一个人，你会自然而然的关注他，忽略其他身边的异性。这就是attention机制对人生活的影响，可以说这种影响无处不在。所以人们就在想，将这个机制应用在图片上可加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！

立即体验

三星手机价格

以上内容由 百度 为你推荐



丫头片子不懂事 (http://blo...)

+ 关注

(http://blog.csdn.net/qq_26609915)

码云

原创

9

粉丝

4

喜欢

0

未开通
(https://gitee.com/qq_26609915)

他的最新文章

更多文章 (http://blog.csdn.net/qq_26609915)

word embedding的模型与测试 (http://blog.csdn.net/qq_26609915/article/details/76515816)

FastText.zip compressing text classification models (http://blog.csdn.net/qq_26609915/article/details/54584505)

tensorflow linspace (http://blog.csdn.net/qq_26609915/article/details/52758647)



内容举报



返回顶部

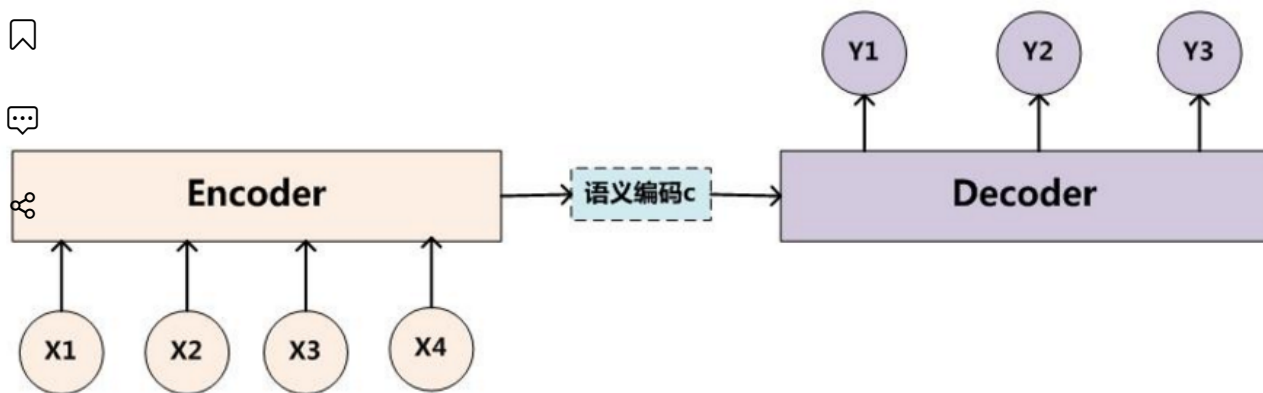


行，是否可以应用在文本处理。答案是可行的，其实当我们快速阅读或者读长文本的时候，我们的注意力是集中在关键词，事件或实体上。通过大量实验证明，将attention机制应用在机器翻译，摘要生成，阅读理解等问题上，取得的成效显著。

Attention与神经网络

Encoder-Decoder框架

Encoder-Decoder框架可以直观地理解为：一个适合处理由一个context生成一个target的通用处理模型。



例如机器翻译中，对于句子对 $\langle X, Y \rangle$ ，我们的目标是给定输入句子 X ，通过Encoder-Decoder框架来生成目标句子 Y 。 X 和 Y 可以是同一种语言，也可以是两种不同的语言。而 X 和 Y 分别由各自的单词序列构成：

$$X = \langle x_1, x_2, \dots, x_m \rangle$$

$$Y = \langle y_1, y_2, \dots, y_n \rangle$$

加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！

Generating News Headline
urrent Neural Networks (http://t.qq_26609915/article/details/1060)

强化学习RL实战篇 (http://t.qq_26609915/article/details/1060)

三星手机价格

以上内容由 百度 为你推荐

相关推荐

Attention注意力机制--原理与应用 (http://blog.csdn.net/joshuaxx316/article/details/70665388)

自然语言处理中的Attention Model：是什么及为什么 (http://blog.csdn.net/malefactor/article/details/50550211)

注意力机制（Attention Mechanism）在自然语言处理中的应用 (http://blog.csdn.net/jdbc/article/details/52948351)

Attention机制 (http://blog.csdn.net/u011531010/article/details/59123388)

内容举报

返回顶部

登录

注册

×

Encoder对输入句子X进行编码，将输入句子通过非线性变换转化为中间语义表示C：

$$C = F(x_1, x_2, \dots, x_m)$$

解码器Decoder，其任务是根据句子X的中间语义表示C和之前已经生成的历史信息 y_1, y_2, \dots, y_{i-1} 来生成i时刻要生成的单词 y_i 。

对于输入而言，Encoder-Decoder模型是没有注意力的，是注意力不集中的分心模型。在目标句子中的每个单词的生成过程：

$$y_1 = f(C), y_2 = f(C, y_1), y_3 = f(C, y_1, y_2), \dots$$



不论生成哪个单词，是 y_1, y_2 ，还是 y_3 也好，他们使用的句子X的语义编码C都是一样的，没有任何区别。

举个例子，输入的英文句子是：Cats eat mice，Encoder-Decoder框架逐步生成中文单词：“猫”，“吃”，“老鼠”。在翻译词语“猫”的时候，分心模型里面的每个英文单词对于翻译目标单词“猫”贡献是相同的，很明显这里不太合理，显然“Cats”对于翻译成“猫”更重要，但是分心模型是无法体现这一点的。这样的模型在输入句子比较短的时候问题不大，但是如果输入句子比较长，缺点就比较明显了。如图所示：

(https://passport.csdn.net/a



一点点加盟费

人脸识别算法 显卡回收

彩钢岩棉夹芯板 钢琴价格

手持式三维扫描仪 高倍显

语音识别 燃烧机 app开发

广告

三星手机价格

以上内容由 百度 为你推荐

广告



中医推拿培训



他的热门文章

attention 机制 (http://blog.csdn.net/qq_26609915/article/details/52086772) 内容举报

5187

强化学习RL实战篇 (http://blog.csdn.net/qq_26609915/article/details/52460812) 返回顶部

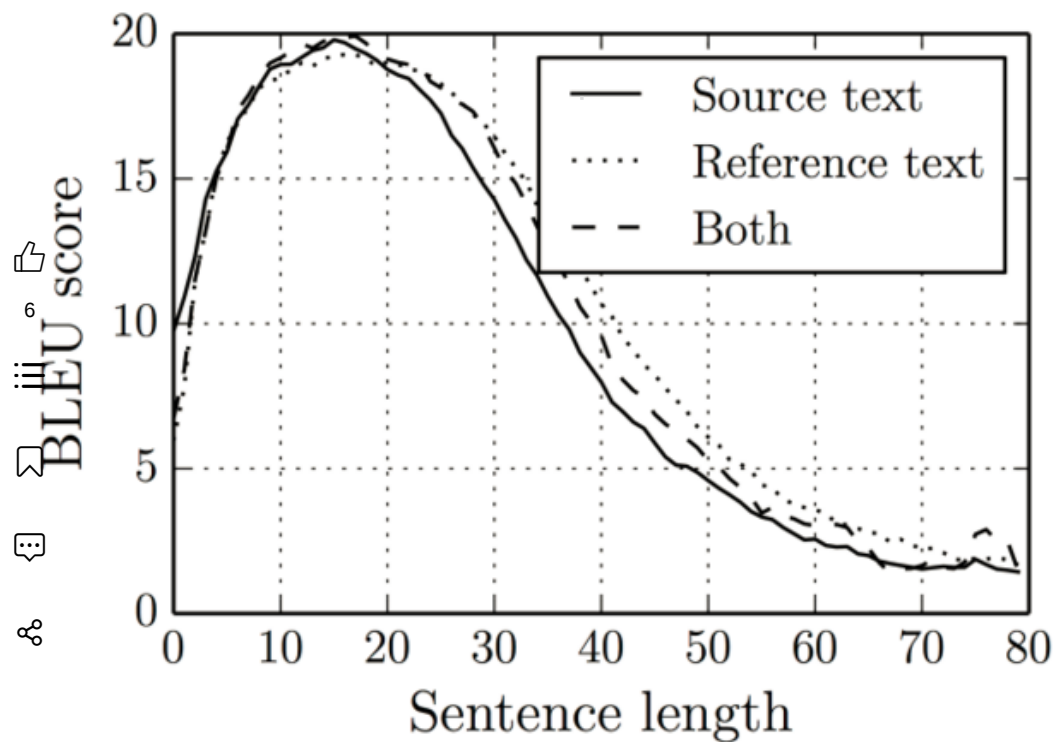
2901

登录

注册



加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！



当文本的长度超过15，Encoder-Decoder框架的结果就逐步下滑了。这是因为，无论输入包含多少信息量，最终所有语义完全通过一个中间语义向量来表示，单词自身的信息已经消失。encoder中丢失很多细节信息，decoder的结果也就会显著变差。

Attention model

Attention严格上讲，是一种机制，而不是具体的model的实现(这一点很像贪心，并不是一种算法，而是一种idea,但人们习惯上称为贪心算法)。我们将Attention引入Encoder-Decoder框架，通过前面的输入X和中间语义表示C来共同决定decoder的结果。增加了Attention model的Encoder-Decoder框架如下图所示：

加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！

机器学习之深度强化学习 (<http://passport.csdn.net/article/details/52119512>)

2801

递归神经网络 (recursive neural network) 实现句子编码 (Autoencode) 实现句子编码 (Autoencode) 实现句子编码 (Autoencode)

2236

FastText.zip compressing text on models (http://blog.csdn.net/qq_26609915/article/details/5458450)

354

三星手机价格

以上内容由 百度 为你推荐



内容举报



返回顶部

登录

注册



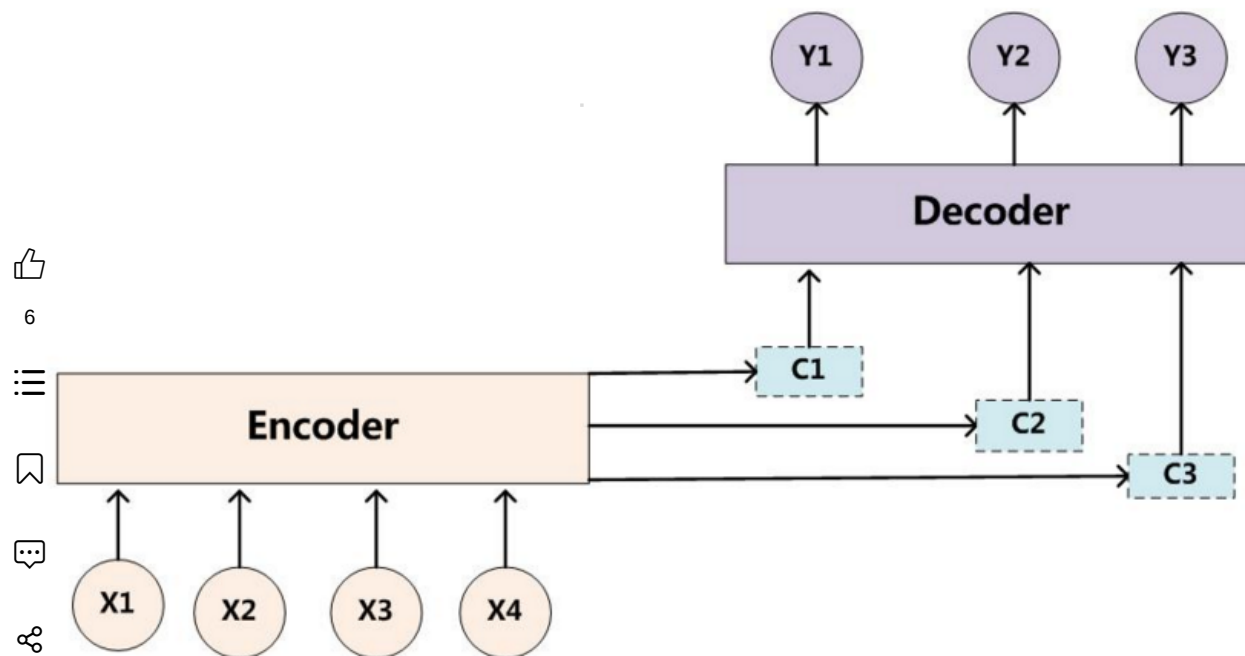
(https://passport.csdn.net/a

三星手机价格

以上内容由 百度 为你推荐



广告



目标句子中的每个单词都学会其对应的源语句子中单词的注意力分配概率信息。在生成每个单词 Y_i 的时候，原先都是相同的中间语义表示 C 会替换成根据当前生成单词而不断变化的 C_i 。即生成目标句子单词的过程成了下面的形式：

$$y_1 = f(C_1), y_2 = f(C_2, y_1), y_3 = f(C_3, y_1, y_2)$$

而每个 C_i 可能对应着不同的源语句子单词的注意力分配概率分布。比如对于上面的英汉翻译来说，其对应的信息可能如下：

$$C_1 = g(0.6 * f('Cats'), 0.2 * f('eat'), 0.2 * f('mice'))$$

加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！

[登录](#)
[注册](#)


⚠
内容举报

⬆
TOP
返回顶部

$$C_2 = g(0.2 * f('Cats'), 0.7 * f('eat'), 0.1 * f('mice'))$$

$$C_3 = g(0.3 * f('Cats'), 0.2 * f('eat'), 0.5 * f('mice'))$$

这里，计算注意力分配概率分布信息要根据不同的网络模型来选择不同的计算框架。



这是参考大牛博客写的读后感



自然语言处理中的Attention Model：是什么及为什么 (<http://blog.csdn.net/malefactor/article/details/50550211>)

Attention 应用

(<http://www.cosmosshadow.com/ml/%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C/2016/03/08/Attention.html>)

深度学习和自然语言处理中的attention和memory机制 (<http://geek.csdn.net/news/detail/50558>)

Neural Attention Model for Abstractive Sentence Summarization (<https://zhuanlan.zhihu.com/p/21388469>)

(<https://passport.csdn.net/a>

三星手机价格

以上内容由 百度 为你推荐



广告



内容举报



返回顶部



fminsearch (fminsearch) 2016-09-13 13:40

1楼

加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长

登录

注册



[回复](#) 1条回复

相关文章推荐

Attention注意力机制--原理与应用 (<http://blog.csdn.net/joshuaxx316/article/details/706653...>)



注意力机制即Attention mechanism在序列学习任务上具有巨大的提升作用，在编解码器框架内，通过在编码段加入A模型，对源数据序列进行数据加权变换，或者在解码端引入A模型，对目标数据进行加权...



joshuaxx316 (<http://blog.csdn.net/joshuaxx316>) 2017年04月24日 21:51 9268



自然语言处理中的Attention Model：是什么及为什么 (<http://blog.csdn.net/malefactor/article...>)



要是关注深度学习在自然语言处理方面的研究进展，我相信你一定听说过Attention Model（后文有时会简称AM模型）这个词。AM模型应该说是过去一年来NLP领域中的重要进展之一，在很多场景被证明有...



malefactor (<http://blog.csdn.net/malefactor>) 2016年01月20日 18:26 57417



广告

惊呆了！微博和阿里背后的数据库有多厉害？

想不到！数据库作为最关键的基础设施，渗透技术领域的方方面面，我阿里和微博的师哥们是这么分享的...

(http://www.baidu.com/cb.php?c=lgF_pyfqHmknjTzrjb0lZ0qnfK9ujYzP1nsrjD10Aw-5Hc3rHnYnHb0TAq15HfLPWRznjb0T1YdryN-mhu9Pvf4uAm1Pjc30AwY5HDdnHn3njTdPH60lgF_5y9YlZ0lQzq-uZR8mLPbUB48ugfEpZNGXy-jULNzTvRETVNzpyN1gvw-lA7GUatLPjqdIAdxTvqdThP-5yF_UvTkn0KzujYk0AFV5H00TZcqn0KdpyfqHRLPjnvnfKEpyfqHc4rj6kP0KWpyfqP1cvrHnz0AqLUWYs0ZK45HcsP6KWThnqPjR1njT)

注意力机制（Attention Mechanism）在自然语言处理中的应用 (<http://blog.csdn.net/jdbc/art...>)
加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！

[登录](#)[注册](#)

(<https://passport.csdn.net/a>)

三星手机价格

以上内容由 百度 为你推荐



广告



内容举报



返回顶部

关于attention的发展综述文章



jdbc (<http://blog.csdn.net/jdbc>) 2016年10月27日 18:17 7397

Attention机制 (<http://blog.csdn.net/u011531010/article/details/59123388>)

转载自：<http://blog.csdn.net/malefactor/article/details/50550211> Attention机制的基本思想是相当直观简洁的，是从人脑注意力模型引入的概念...



u011531010 (<http://blog.csdn.net/u011531010>) 2017年03月01日 20:09 766

深度学习方法（九）：自然语言处理中的Attention Model注意力模型 (<http://blog.csdn.net/xb...>)

上一篇博文深度学习方法（八）：Encoder-Decoder模型，基本Sequence to Sequence模型描述了基本的Encoder-Decoder模型，在作为翻译模型的时候，这种基本的Enc...



xbinworld (<http://blog.csdn.net/xbinworld>) 2017年02月04日 00:27 6405

1



3.10/盒
供应钢钉线卡 pp乐清
钢钉线厂家批发圆形钢

2



1150.0/套
原装正品 摩托罗拉防
爆对讲机GP328 专业

3



1450.00/台
原装正品摩托罗拉
GP328防爆对讲机专

Attention和增强循环神经网络 (<http://blog.csdn.net/a398942089/article/details/53888797>)

本文重点讲述的是一种被称为attention的方法，有的人将其译为“聚焦”，但觉得这种翻译将原文对神经网络拟人化的手法给扔掉了，因此保留了原来的称谓。Attention，顾名思义，就是让神经网络对部分...



a398942089 (<http://blog.csdn.net/a398942089>) 2016年12月26日 20:12 2218

加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长 **Attention and Memory in Deep Learning and NLP(深度学习和NLP中的注意和记忆机制) 阅读..**

登录

注册



(<https://passport.csdn.net/a>

三星手机价格

以上内容由 百度 为你推荐




内容举报




返回顶部

什么是Attention Mechanisms(注意机制) 神经网络中的注意机制类似人类的视觉注意机制，人类视觉本质上看图片会聚焦在“高清晰度”的区域，同时也会感知周围“低清晰度”的区域，然...


 aPYXa (<http://blog.csdn.net/aPYXa>) 2016年07月26日 20:19 1665


阅读理解任务中的Attention-over-Attention神经网络模型原理及实现 (<http://blog.csdn.net/liuchong>...

 本文是“Attention-over-Attention Neural Networks for Reading Comprehension”的阅读笔记。这篇论文所处理的任务是阅读理解里面的完形填空问...

 liuchong (<http://blog.csdn.net/liuchong>) 2017年06月06日 09:24 2531


用于文本分类的RNN-Attention网络 (http://blog.csdn.net/thriving_fcl/article/details/733812...

 这篇博客主要介绍Attention机制在文本分类任务上的作用，原理以及附带的代码实现。...

 thriving_fcl (http://blog.csdn.net/thriving_fcl) 2017年06月17日 15:51 2320

深度学习笔记(六)：Encoder-Decoder模型和Attention模型 (<http://blog.csdn.net/u01459501>...

这两天在看attention模型，看了下知乎上的几个回答，很多人都推荐了一篇文章Neural Machine Translation by Jointly Learning to Align and ...

 u014595019 (<http://blog.csdn.net/u014595019>) 2016年10月15日 23:09 19201

自然语言处理中的Attention Model：是什么及为什么 (<http://blog.csdn.net/jdbc/article/detail>...

要是关注深度学习在自然语言处理方面的研究进展，我相信你一定听说过Attention Model（后文有时会简称AM模型）这个词。AM模型应该说是过去一年来NLP领域中的重要进展之一，在很多场景被证明有...

加入CSDN，享受更精准的内容推荐，与3000万程序员共同成长

(<https://passport.csdn.net/a>

三星手机价格

以上内容由 百度 为你推荐




内容举报


返回顶部

登录


注册



<https://passport.csdn.net/a>

Tensorflow 自动文摘: 基于Seq2Seq+Attention模型的Textsum模型 (<http://blog.csdn.net/rockingdingo>)


这篇文章中我们将基于Tensorflow的Seq2Seq+Attention模型，介绍如何训练一个中文的自动生成新闻标题的模型。自动总结 (Automatic Summarization)类型的模型一直...

 rockingdingo (<http://blog.csdn.net/rockingdingo>) 2017年02月15日 21:42 13256

6

seq2seq学习笔记 (http://blog.csdn.net/Jerr__y/article/details/53749693)

@author: huangyongye1. RNN基础对于RNN，我看到讲得最通俗易懂的应该是Andrej发的博客：The Unreasonable Effectiveness of Recur...

 Jerr__y (http://blog.csdn.net/Jerr__y) 2016年12月20日 11:41 25823

Tensorflow 自动文摘: 基于Seq2Seq+Attention模型的Textsum模型 (http://blog.csdn.net/Real_Myth)

from: Tensorflow 自动文摘: 基于Seq2Seq+Attention模型的Textsum模型 Github下载完整代码 <https://github.com>...

 Real_Myth (http://blog.csdn.net/Real_Myth) 2017年04月07日 16:51 4928

seq2seq模型 (<http://blog.csdn.net/sunlylorn/article/details/50607376>)

seq2seq模型，简单点说，是一个翻译模型，把一个sequence翻译成另一个sequence，最早在SMT领域被证明。其基本思想是两个RNN LM，一个作为encoder，另一个作为decoder...

 sunlylorn (<http://blog.csdn.net/sunlylorn>) 2016年01月29日 15:08 18718

模型汇总16 各类Seq2Seq模型对比及《Attention Is All You Need》中技术详解 (http://blog.csdn.net/qq_26609915/article/details/52086772)
加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！

三星手机价格

以上内容由 百度 为你推荐



内容举报




返回顶部

登录


注册




1、已有Seq2Seq模型 Seq2Seq模型是处理序列到序列问题的利器，尤其是在神经网络翻译（NMT）方面，取得了很大的成功。Seq2Seq由一个encoder和一个decoder构成，encode...


 lqfarmer (<http://blog.csdn.net/lqfarmer>) 2017年06月20日 22:00 1505



basic_rnn_seq2seq与embedding_attention_seq2seq输入的差别 (<http://blog.csdn.net/u013713117>)

 在用embedding_attention_seq2seq 跑完数据之后，想用basic_rnn_seq2seq再跑一遍进行对比，发现报错了，错误是ValueError: Linear is expected to have shape (..., 1) but got shape (..., 2) instead.

 u013713117 (<http://blog.csdn.net/u013713117>) 2017年02月22日 16:34 1854


Tensorflow新版Seq2Seq接口使用 (http://blog.csdn.net/thriving_fcl/article/details/74165062)

 Tensorflow 1.0.0 版本以后，开发了新的seq2seq接口，弃用了原来的接口。旧的seq2seq接口也就是tf.contrib.legacy_seq2seq下的那部分，新的接口在tf.nn.

  thriving_fcl (http://blog.csdn.net/thriving_fcl) 2017年07月02日 23:50 5517

seq2seq以及Attention机制 (http://blog.csdn.net/Michael_liuyu09/article/details/77943299)

RNN 由一系列相同的网络构成（图 1 中的长方形表示一个网络），上一个词语的向量表示作为计算下一个网络的输入，如此循环。整个句子每个词计算完成，便得到了一个句子的语意向量。所谓enco...

 Michael_liuyu09 (http://blog.csdn.net/Michael_liuyu09) 2017年09月12日 11:33 193

tensorflow seq2seq模型 代码阅读分析 (http://blog.csdn.net/vincent_hbl/article/details/770...)

如果刚开始入门该模型请阅读tf官方说明：Sequence-to-Sequence Models模型应用于机器翻译的示例代码：github如果还没有看懂的translate示例代码，请先理解透彻，再加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！

(<https://passport.csdn.net/a>)

三星手机价格

以上内容由 百度 为你推荐



内容举报



返回顶部

登录

注册

