



(<http://blog.csdn.net/u014595019/article/details/52571966>)

深度学习笔记(五)：LSTM (<http://blog.csdn.net/u014595019/article/details/52605693>)

深度学习笔记(六)：Encoder-Decoder模型和Attention模型

(<http://blog.csdn.net/u014595019/article/details/52826423>)



7



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 Translate (<https://arxiv.org/abs/1409.0473>) 我看了下，感觉非常的不错，里面还大概阐述了encoder-decoder(编码)模型的概念，以及传统的RNN实现。然后还阐述了自己的attention模型。我看了一下，自己做了一些摘录，写在下面

## 1.Encoder-Decoder模型及RNN的实现

所谓**encoder-decoder模型**，又叫做**编码-解码模型**。这是一种应用于seq2seq问题的模型。

那么seq2seq又是什么呢？简单的说，就是根据一个输入序列 $x$ ，来生成另一个输出序列 $y$ 。seq2seq有很多的应用，例如翻译，文档摘取，问答系统等等。在翻译中，输入序列是待翻译的文本，输出序列是翻译后的文本；在问答系统中，输入序列是提出的问题，而输出序列是答案。

为了解决seq2seq问题，有人提出了encoder-decoder模型，也就是编码-解码模型。所谓编码，就是将输入序列转化成一个固定长度的向量；解码，就是将之前生成的固定向量再转化成输出序列。

ubuntu16.04下安装CUDA(<https://myasport.csdn.net/account>)

tensorflow-gpu版本  
n.net/u014595019/  
015)

electron开发记录(王  
strap并对其一些用  
g.csdn.net/u014595  
53610731)



专业清洗油烟机



牙膏去黑头



近视手术后遗症



洗瓶机  
百度一下

相关推荐



深度学习(城市)blog.csdn自然语言处理中的  
Encoder/Decoder模型40基本Sequence t

encoder-decoder ([http://blog.csdn.net/sinat\\_17451213/article/details/55102873](http://blog.csdn.net/sinat_17451213/article/details/55102873))

使用Encoder-Decoder模型自动生成对联的思路 (<http://blog.csdn.net/malefactor/article/details/51124732>)

深度学习笔记——Attention Model (注意力模型) 学习总结 ([http://blog.csdn.net/mpk\\_no1/article/details/72862348](http://blog.csdn.net/mpk_no1/article/details/72862348))



内容举报



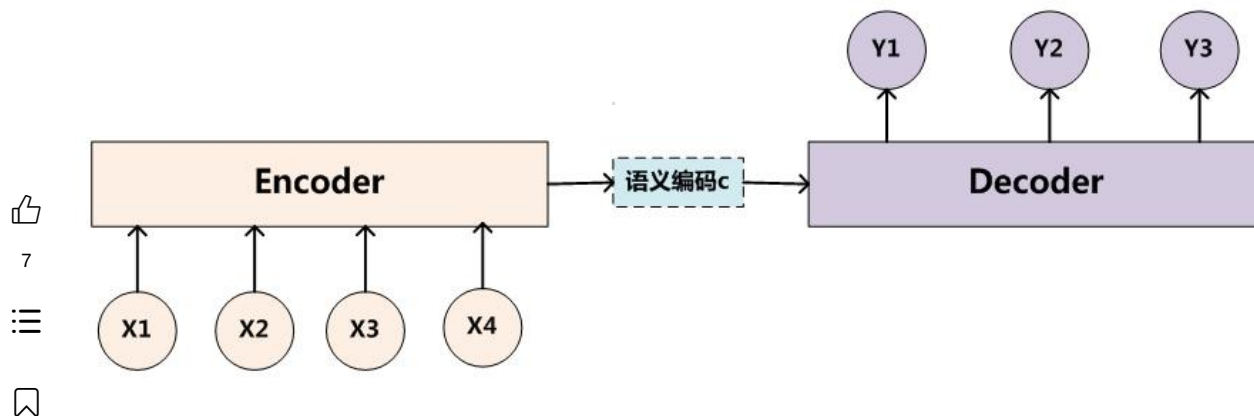
返回顶部

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





当然了，这个只是大概的思想，具体实现的时候，编码器和解码器都不是固定的，可选的有CNN/RNN/BiRNN/GRU/LSTM等等，你可以自由组合。比如说，你在编码时使用BiRNN，解码时使用RNN，或者在编码时使用RNN，解码时使用LSTM等等。

这边为了方便阐述，选取了编码和解码都是RNN的组合。在RNN中，当前时间的隐藏状态是由上一时间的状态和当前时间输入决定的，也就是

$$h_t = f(h_{t-1}, x_t)$$

获得了各个时间段的隐藏层以后，再将隐藏层的信息汇总，生成最后的语义向量

$$C = q(h_1, h_2, h_3, \dots, h_{T_x})$$

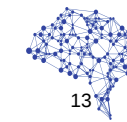
一种简单的方法是将最后的隐藏层作为语义向量C，即

$$C = q(h_1, h_2, h_3, \dots, h_{T_x}) = h_{T_x}$$

加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！



### 博主专栏



深度学习

&tensorflow笔记

(http://blog.csdn.net/column/details/13414.html)

(http://blog.csdn.net/column/details/13414.html)



自然语言处理

(http://blog.csdn.net/column/details/13461.html)

119492

(http://blog.csdn.net/column/details/13461.html)



内容举报



返回顶部

登录

注册



解码阶段可以看做编码的逆过程。这个阶段，我们要根据给定的语义向量C和之前已经生成的输出序列  $y_1, y_2, \dots, y_{t-1}$  来预测下一个输出的单词  $y_t$ ，即

$$y_t = \operatorname{argmax} P(y_t) = \prod_{t=1}^T p(y_t | \{y_1, \dots, y_{t-1}\}, C)$$

也可以写作

7

≡

$$y_t = g(\{y_1, \dots, y_{t-1}\}, C)$$

🔖

而在RNN中，上式又可以简化成

💬

$$y_t = g(y_{t-1}, s_t, C)$$

🔗

其中  $s$  是输出RNN中的隐藏层，C代表之前提过的语义向量， $y_{t-1}$ 表示上个时间段的输出，反过来作为这个时间段的输入。而  $g$  则可以是一个非线性的多层的神经网络，产生词典中各个词语属于  $y_t$  的概率。

**encoder-decoder模型虽然非常经典，但是局限性也非常大。**最大的局限性就在于编码和解码之间的唯一联系就是一个固定长度的语义向量C。也就是说，编码器要将整个序列的信息压缩进一个固定长度的向量中去。但是这样做有两个弊端，一是语义向量无法完全表示整个序列的信息，还有就是先输入的内容携带的信息会被后输入的信息稀释掉，或者说，被覆盖了。输入序列越长，这个现象就越严重。这就使得在解码的时候一开始就没有获得输入序列足够的信息，那么解码的准确度自然也要打个折扣了

## 2.Attention模型

为了解决这个问题，作者提出了Attention模型，或者说注意力模型。简单的说，这种模型在产生输出的时候，还会产生一个“注意力范围”表示接下来输出的时候要重点关注输入序列中的哪些部分，然后根据关注的区域来产生下一个输出，如此往复。模型的大概示意图如下所示

(https://passport.csdn.net/account)

专业清洗油烟机 牙膏去黑头 洗瓶机 百度一下 洗瓶机

### 他的热门文章

tensorflow笔记：多层LSTM代码分析 (http://blog.csdn.net/u014595019/article/details/52759104)

🔖 59802

tensorflow笔记：常用函数说明 (http://blog.csdn.net/u014595019/article/details/52805444)

🔖 45080

自己动手写word2vec (四):CBOW和skip-gram模型 (http://blog.csdn.net/u014595019/article/details/51943428)

🔖 34414

tensorflow笔记：流程，概念和简单代码注释 (http://blog.csdn.net/u014595019/article/details/52677412)

🔖 33318

深度学习笔记(三)：激活函数和损失函数 (http://blog.csdn.net/u014595019/article/details/52562159)

🔖 26392

登录

注册



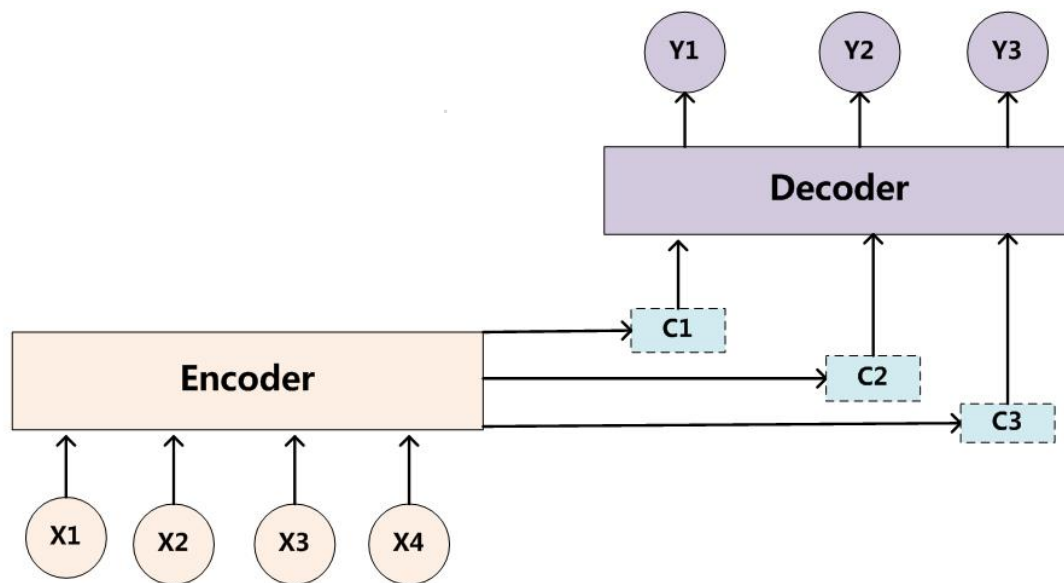
内容举报



返回顶部



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



相比于之前的encoder-decoder模型，attention模型最大的区别就在于它不在要求编码器将所有输入信息都编码进一个固定长度的向量之中。相反，此时编码器需要将输入编码成一个向量的序列，而在解码的时候，每一步都会选择性的从向量序列中挑选一个子集进行进一步处理。这样，在产生每一个输出的时候，都能够做到充分利用输入序列携带的信息。而且这种方法在翻译任务中取得了非常不错的成果。

在这篇文章中，作者提出了一个用于翻译任务的结构。解码部分使用了attention模型，而在编码部分，则使用了BiRNN(bidirectional RNN,双向RNN)

## 2.1 解码

我们先来看看解码。解码部分使用了attention模型。类似的，我们可以将之前定义的条件概率写作

加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！  

$$p(y_i | y_1, \dots, y_{i-1}, X) = g(y_{i-1}, s_i, c_i)$$



内容举报

返回顶部

登录

注册

×



上式 $s_i$ 表示解码器 $i$ 时刻的隐藏状态。计算公式是

$$s_i = f(s_{i-1}, y_{i-1}, c_i)$$

注意这里的条件概率与每个目标输出 $y_i$ 相对应的内容向量 $c_i$ 有关。而在传统的方式中，只有一个内容向量 $C$ 。那么这里的内容向量 $c_i$ 又该怎么算呢？其实 $c_i$ 是由编码时的隐藏向量序列 $(h_1, \dots, h_{T_x})$ 按权重相加得到的。

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j$$

由于编码使用了双向RNN，因此可以认为 $h_i$ 中包含了输入序列中第 $i$ 个词以及前后一些词的信息。将隐藏向量序列按权重相加，表示在生成第 $j$ 个输出的时候的注意力分配是不同的。 $\alpha_{ij}$ 的值越高，表示第 $i$ 个输出在第 $j$ 个输入上分配的注意力越多，在生成第 $i$ 个输出的时候受第 $j$ 个输入的影响也就越大。那么现在我们又有新问题， $\alpha_{ij}$ 又是怎么得到的呢？这个其实是由第 $i-1$ 个输出隐藏状态 $s_{i-1}$ 和输入中各个隐藏状态共同决定的。也即是

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})}$$

$$e_{ij} = a(s_{i-1}, h_j)$$

也就是说， $s_{i-1}$ 先跟每个 $h$ 分别计算得到一个数值，然后使用softmax得到 $i$ 时刻的输出在 $T_x$ 个输入隐藏状态中的注意力分配向量。这个分配向量也就是计算 $c_i$ 的权重。我们现在再把公式按照执行顺序汇总一下：

(https://passport.csdn.net/account)



近视手术后遗症

洗瓶机

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



返回顶部

$$e_{ij} = a(s_{i-1}, h_j)$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})}$$

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j$$

$$s_i = f(s_{i-1}, y_{i-1}, c_i)$$

$$y_i = g(y_{i-1}, s_i, c_i)$$



7



上面这些公式就是解码器在第*i*个时间段内要做的事情。作者还给了一个示意图：

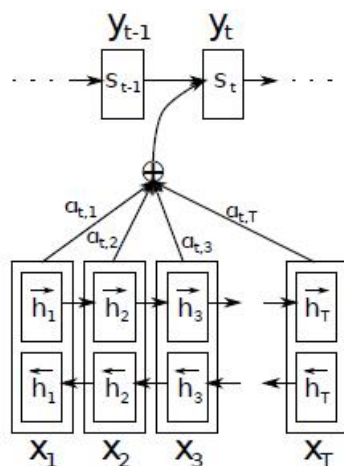


Figure 1: The graphical illustration of the proposed model trying to generate the  $t$ -th target word  $y_t$  given a source sentence  $(x_1, x_2, \dots, x_T)$ .

## 2.2 编码

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/account>)



内容举报

返回顶部

登录

注册



相比于上面解码的创新，这边的编码就比较普通了，只是传统的单向的RNN中，数据是按顺序输入的，因此第 $j$ 个隐藏状态 $\vec{h}_j$ 只能携带第 $j$ 个单词本身以及之前的一些信息；而如果逆序输入，则 $\overleftarrow{h}_j$ 包含第 $j$ 个单词及之后的一些信息。如果把这两个结合起来， $h_j = [\vec{h}_j, \overleftarrow{h}_j]$ 就包含了第 $j$ 个输入和前后的信息。



### 3.实验结果



为了检验性能，作者分别使用传统模型和attention模型在英语-法语的翻译数据集上进行了测验。



传统模型的编码器和解码器各有1000个隐藏单元。编码器中还有一个多层神经网络用于实现从隐藏状态到单词的映射。在优化方面，使用了SGD(minibatch stochastic gradient descent)以及Adadelta,前者负责采样，后者负责优化下降方向。



得到的结果如下：

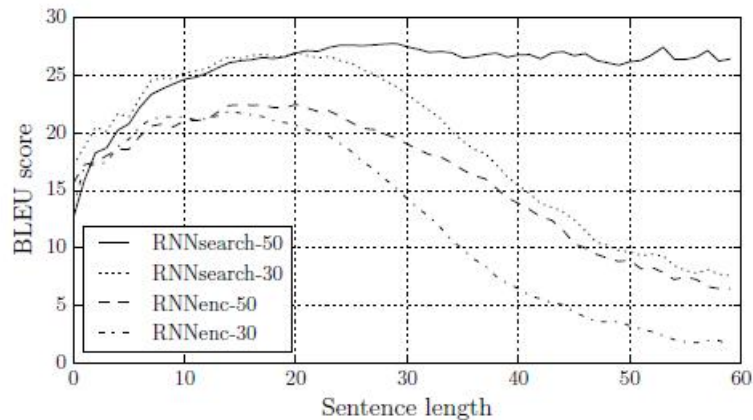


Figure 2: The BLEU scores of the generated translations on the test set with respect to the lengths of the sentences. The results are on the full test set which includes sentences having unknown words to the models.

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



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



返回顶部

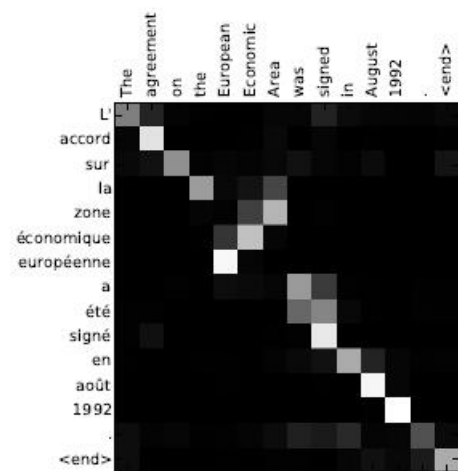


图中RNNenc表示传统的结构，而RNNsearch表示attention模型。后面的数字表示序列的长度。可以看到，不论序列长度，attention模型的性能均优于传统的编码-解码模型。而RNNsearch-50甚至在长文本上的性能也非常优异

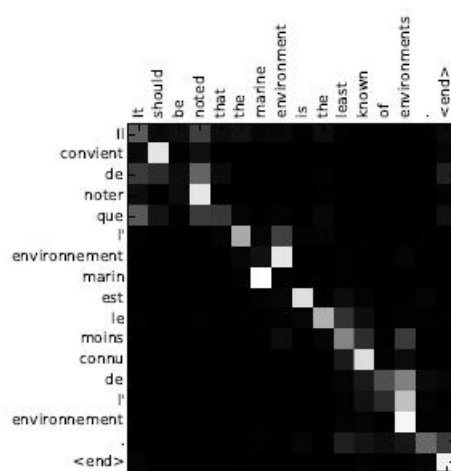
除了准确度之外，还有一个很值得关注的东西：注意力矩阵。之前已经提过，每个输出都有一个长为 $T_x$ 的注意力向量，那么将这些向量合起来看，就是一个矩阵。对其进行可视化，得到如下结果



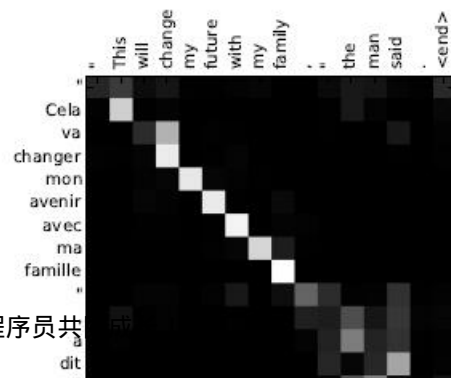
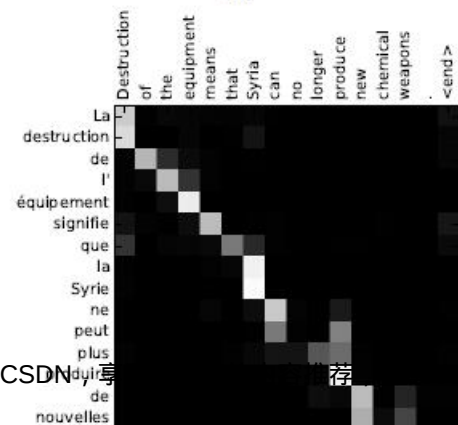
7



(a)



(b)



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



专业清洗油烟机



牙膏去黑头



近视手术后遗症



洗瓶机  
百度一下  
洗瓶机

广告

加入CSDN大家庭

推荐

万程序员共

登录

注册

⚠  
内容举报

⏮  
返回顶部



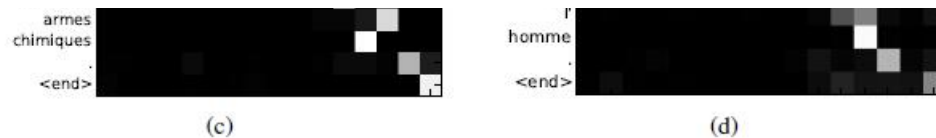


Figure 3: Four sample alignments found by RNNsearch-50. The x-axis and y-axis of each plot correspond to the words in the source sentence (English) and the generated translation (French), respectively. Each pixel shows the weight  $\alpha_{ij}$  of the annotation of the  $j$ -th source word for the  $i$ -th target word (see Eq. (6)), in grayscale (0: black, 1: white). (a) an arbitrary sentence. (b–d) three randomly selected samples among the sentences without any unknown words and of length between 10 and 20 words from the test set.

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



其中x轴表示待翻译的句子中的单词(英语), y轴表示翻译以后的句子中的单词(法语)。可以看到尽管从英语到法语的过程中,有些单词的顺序发生了变化,但是attention模型仍然很好的找到了合适的位置。换句话说,就是两种语言下的单词“对齐”了。因此,也有人把注意力模型叫做对齐(alignment)模型。而且像比于用语言学实现的硬对齐,这种基于概率的软对齐更加优雅,因为能够更全面的考虑到上下文的语境。

版权声明：本文为博主原创文章，未经博主允许不得转载。

本文已收录于以下专栏：深度学习&tensorflow笔记 (<http://blog.csdn.net/column/details/13414.html>)



wonder\_wgf (/wonder\_wgf) 2017-12-07 22:00

1楼

(/wonder\_wgf) 确定你的博客编码和解码没写反？

加入CSDN会员享受更精准的内容推荐，与5000万程序员共同成长！



内容举报



返回顶部

登录

注册



## 相关文章推荐

### 深度学习方法（八）：自然语言处理中的Encoder-Decoder模型，基本Sequence to Sequenc...



Encoder-Decoder（编码-解码）是深度学习中非常常见的一个模型框架，比如无监督算法的auto-encoding就是用编码-解码的结构设计并训练的；比如这两年比较热的image captio...

7



xbinworld (<http://blog.csdn.net/xbinworld>) 2017年01月19日 00:07 6837



### encoder-decoder ([http://blog.csdn.net/sinat\\_17451213/article/details/55102873](http://blog.csdn.net/sinat_17451213/article/details/55102873))



Encoder-decoder框架 Encoder-decoder框架最抽象的一种表示：即由一个句子生成另一个句子的通过处理模型。对于句子对,目标是给定输入句子x，期待通过encoder-dec...



sinat\_17451213 ([http://blog.csdn.net/sinat\\_17451213](http://blog.csdn.net/sinat_17451213)) 2017年02月14日 17:48 917



### 【前端逆袭记】我是怎么从月薪4k到40k的！

谨以此篇文章献给我奋斗过的程序人生！我第一次编码是在我大一的时候....

([http://www.baidu.com/cb.php?c=lgF\\_pyfqHmknj0dP1f0IZ0qnfK9ujYzP1ndPWb10Aw-5Hc3rHnYnHb0TAq15HfLPWRznjb0T1d-m1DkPyRvuAnsPWDknAmk0AwY5HDdnHn3njm4PHb0IgF\\_5y9YIZ0IQzq-uZR8mLPbUB48ugfEIAqspynEivNBnHqdlAdxTvqdThP-5yF\\_UvTkn0KzujYk0AFV5H00TZcqn0KdpyfqHRLPjnvnfKEpyfqHc4rj6kP0KWpyfqP1civrHnz0AqLUWYs0ZK45HcsP6KWThnqnH6knHD](http://www.baidu.com/cb.php?c=lgF_pyfqHmknj0dP1f0IZ0qnfK9ujYzP1ndPWb10Aw-5Hc3rHnYnHb0TAq15HfLPWRznjb0T1d-m1DkPyRvuAnsPWDknAmk0AwY5HDdnHn3njm4PHb0IgF_5y9YIZ0IQzq-uZR8mLPbUB48ugfEIAqspynEivNBnHqdlAdxTvqdThP-5yF_UvTkn0KzujYk0AFV5H00TZcqn0KdpyfqHRLPjnvnfKEpyfqHc4rj6kP0KWpyfqP1civrHnz0AqLUWYs0ZK45HcsP6KWThnqnH6knHD))

### 使用Encoder-Decoder模型自动生成对联的思路 (<http://blog.csdn.net/malefactor/article/det...>)

本文介绍了如何利用Encoder-Decoder框架来建立对联自动生成系统的主要思路。  
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/account>)



专业清洗油烟机



牙膏去黑头



近视手术后遗症



洗瓶机

百度一下

洗瓶机



内容举报



返回顶部

登录

注册





malefactor (<http://blog.csdn.net/malefactor>) 2016年04月13日 18:21 21339

## 深度学习笔记——Attention Model (注意力模型) 学习总结 ([http://blog.csdn.net/mpk\\_no1/a...](http://blog.csdn.net/mpk_no1/a...))

Attention Model (注意力模型) 学习总结, 包括soft Attention Model, Global Attention Model和Local Attention Model, 静态AM, ...



7



mpk\_no1 ([http://blog.csdn.net/mpk\\_no1](http://blog.csdn.net/mpk_no1)) 2017年08月06日 21:49 7067



## 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 6404



7

0.20/个

批量改价, 行车记录仪  
线扣. 记录仪线夹. 布线

8

120.0/台

对讲机迷你超薄微型  
商务酒店发廊美容院民

9

370.00/件

专业供应钢钉线卡 圆  
形方型线卡 固定夹电

## 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 9267

## attention model (<http://blog.csdn.net/apsvvfb/article/details/59110169>)

How did I select papers?First, I tried to search for "attention" in CVPR2014-2016, ICCV2009-2015 an...

加入CSDN, 享受更精准的内容推荐, 与5000万程序员共同成长! 505

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



专业清洗油烟机



牙膏去黑头



近视手术后遗症



洗瓶机  
百度一下

洗瓶机

广告



内容举报



返回顶部



登录

注册



## 王小草【深度学习】笔记第七弹--RNN与应用案例：注意力模型与机器翻译 (<http://blog.csdn.n...>)



标签（空格分隔）：王小草深度学习笔记1. 注意力模型1.2 注意力模型概述注意力模型（attention model）是一种用于做图像描述的模型。在笔记6中讲过RNN去做图像描述，但是精准度可能差强...

 sinat\_33761963 ([http://blog.csdn.net/sinat\\_33761963](http://blog.csdn.net/sinat_33761963)) 2016年12月08日 17:12  3012

7

## 视觉注意力的循环神经网络模型 ([http://blog.csdn.net/Leo\\_Xu06/article/details/53491400](http://blog.csdn.net/Leo_Xu06/article/details/53491400))



视觉注意力的循环神经网络模型

 Leo\_Xu06 ([http://blog.csdn.net/Leo\\_Xu06](http://blog.csdn.net/Leo_Xu06)) 2016年12月07日 20:43  3926





## LSTM的Encoder-Decoder模式 (<http://blog.csdn.net/niuniuyuh/article/details/59108795>)

深度学习笔记(一)：logistic分类 深度学习笔记(二)：简单神经网络，后向传播算法及实现 深度学习笔记(三)：激活函数和损失函数 深度学习笔记：优化方法总结(BGD,SGD,Momentu...

 niuniuyuh (<http://blog.csdn.net/niuniuyuh>) 2017年03月01日 15:20  1906

## java netty之decoder与encoder的使用 ([http://blog.csdn.net/fjs\\_cloud/article/details/93421...](http://blog.csdn.net/fjs_cloud/article/details/93421...))

在前面的文章中，对netty的使用都是基于byte类型的，但是在实际情况，直接处理byte数据是很少的，一般都是将这些数据转换成自己定义的一些类型。也就是说在实际情况decoder与encode...

 fjs\_cloud ([http://blog.csdn.net/fjs\\_cloud](http://blog.csdn.net/fjs_cloud)) 2013年07月16日 14:29  9114

## 使用 URLDecoder 和 URLEncoder 对中文字符进行编码和解码 (<http://blog.csdn.net/justlov...>)

URLDecoder 和 URLEncoder 用于完成普通字符串 和 application/x-www-form-urlencoded MIME 字符串之间的相互转换。在加入CSDN，享受更精准的内容推荐，与5000万程序员共同成长！本文中，我们可以使用URLD...

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



 内容举报

 返回顶部

登录

注册

×



 justloveyou\_ ([http://blog.csdn.net/justloveyou\\_](http://blog.csdn.net/justloveyou_)) 2017年02月25日 23:17 9632

## 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...



7

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




## 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




## LSTM源码分析 (<http://blog.csdn.net/u011274209/article/details/53329082>)

代码来源：LSTM Networks for Sentiment Analysis ps：markdown真难用啊

 u011274209 (<http://blog.csdn.net/u011274209>) 2016年11月25日 01:09 2116

## 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...

 jdbc (<http://blog.csdn.net/jdbc>) 2016年01月26日 11:16 2328

## 深度学习中的Attention模型介绍及其进展 (<http://blog.csdn.net/jteng/article/details/52864401>)

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/account>)



专业清洗油烟机



牙膏去黑头



近视手术后遗症

广告



洗瓶机  
百度一下

洗瓶机



内容举报



返回顶部

登录

注册



近期对深度学习中的Attention模型进行了深入研究，该模型在图像识别、语音识别和自然语言处理三大深度学习的热门领域均有广泛的使用，是2014和2015年深度学习领域的重要进展。现对其原理、主要应用...



jteng (<http://blog.csdn.net/jteng>) 2016年10月20日 15:29 6980

## Attention model 的一些概念 (<http://blog.csdn.net/u014221266/article/details/49132041>)



自去年在机器翻译领域得到有效验证之后，基于Attention的神经网络模型在今年会议上（2015EMNLP）大放异彩 <http://www.epistemics.co.uk/staff/nmilton>...



u014221266 (<http://blog.csdn.net/u014221266>) 2015年10月14日 22:03 1367



## 以Attention Model为例谈谈两种研究创新模式 (<http://blog.csdn.net/malefactor/article/detail...>)



各位观众朋友好，也许此刻您刚打开电梯.....上了年纪的读者估计能看出上面一句是引用了韩乔生老先生的名言，我写东西就喜欢用名人名言开头，这好习惯这么多年怎么也改不了。您问韩乔生是谁？恭喜您，作为90后您以...



malefactor (<http://blog.csdn.net/malefactor>) 2016年01月26日 00:18 21284

## RNN的四种代表性扩展—Attention and Augmented Recurrent Neural Networks（一）(<http://blog.csdn.net/jdbc/article/details/52826423>)

Recurrent neural networks are one of the staples of deep learning, allowing neural networks to work ...



jdbc (<http://blog.csdn.net/jdbc>) 2017年02月23日 23:47 561

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



专业清洗油烟机



牙膏去黑头



近视手术后遗症



洗瓶机  
百度一下

洗瓶机

广告



内容举报



返回顶部

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

