

# [译] 理解 LSTM 网络



朱小虎Neil (/u/696dc6c6f01c) + 关注

2015.08.28 12:36\* 字数 3736 阅读 290000 评论 81 喜欢 523 赞赏 27  
(/u/696dc6c6f01c)

Neil Zhu, 简书ID Not\_GOD, University AI 创始人 & Chief Scientist, 致力于推进世界人工智能化进程。制定并实施 UAI 中长期增长战略和目标, 带领团队快速成长为人工智能领域最专业的力量。

作为行业领导者, 他和UAI一起在2014年创建了TASA (中国最早的人工智能社团), DL Center (深度学习知识中心全球价值网络), AI growth (行业智库培训) 等, 为中国的人工智能人才建设输送了大量的血液和养分。此外, 他还参与或者举办过各类国际性的人工智能峰会和活动, 产生了巨大的影响力, 书写了60万字的人工智能精品技术内容, 生产翻译了全球第一本深度学习入门书《神经网络与深度学习》, 生产的内容被大量的专业垂直公众号和媒体转载与连载。曾经受邀为国内顶尖大学制定人工智能学习规划和教授人工智能前沿课程, 均受学生和老师好评。

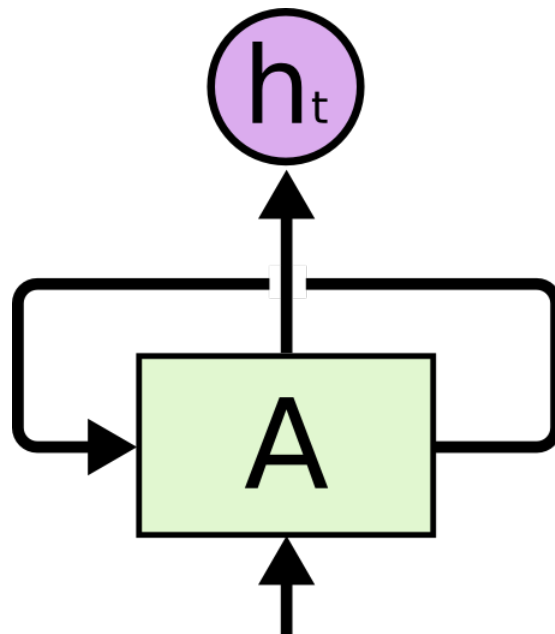
本文译自 Christopher Olah 的博文 (<https://link.jianshu.com?t=http://colah.github.io/posts/2015-08-Understanding-LSTMs/>)

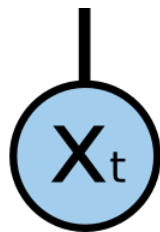
## Recurrent Neural Networks

人类并不是每时每刻都从一片空白的大脑开始他们的思考。在你阅读这篇文章时候, 你都是基于自己已经拥有的对先前所见词的理解来推断当前词的真实含义。我们不会将所有的东西都全部丢弃, 然后用空白的大脑进行思考。我们的思想拥有持久性。

传统的神经网络并不能做到这点, 看起来也像是一种巨大的弊端。例如, 假设你希望对电影中的每个时间点的时间类型进行分类。传统的神经网络应该很难来处理这个问题——使用电影中先前的事件推断后续的事件。

RNN 解决了这个问题。RNN 是包含循环的网络, 允许信息的持久化。

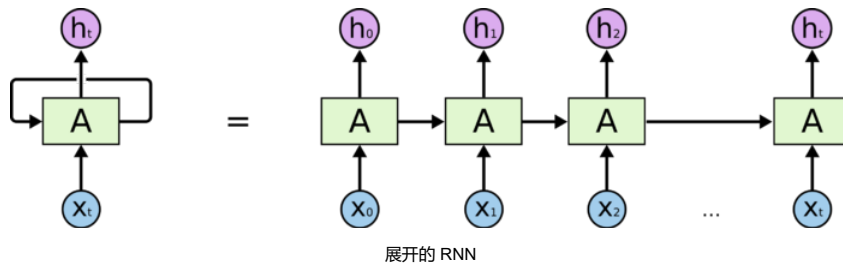




RNN 包含循环

在上面的示例图中，神经网络的模块，A，正在读取某个输入  $x_i$ ，并输出一个值  $h_i$ 。  
循环可以使得信息可以从当前步传递到下一步。

这些循环使得 RNN 看起来非常神秘。然而，如果你仔细想想，这样也不比一个正常的神经网络难于理解。RNN 可以被看做是同一神经网络的多次复制，每个神经网络模块会把消息传递给下一个。所以，如果我们将这个循环展开：



展开的 RNN

链式的特征揭示了 RNN 本质上是与序列和列表相关的。他们对于这类数据的最自然的神经网络架构。

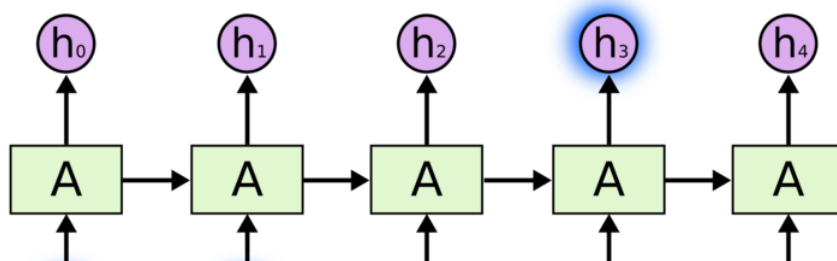
并且 RNN 也已经被人们应用了！在过去几年中，应用 RNN 在语音识别，语言建模，翻译，图片描述等问题上已经取得一定成功，并且这个列表还在增长。我建议大家参考 Andrej Karpathy 的博客文章——The Unreasonable Effectiveness of Recurrent Neural Networks (<https://link.jianshu.com?t=http://karpathy.github.io/2015/05/21/rnn-effectiveness/>) 来看看更丰富有趣的 RNN 的成功应用。

而这些成功应用的关键之处就是 LSTM 的使用，这是一种特别的 RNN，比标准的 RNN 在很多的任务上都表现得更好。几乎所有的令人振奋的关于 RNN 的结果都是通过 LSTM 达到的。这篇博文也会就 LSTM 进行展开。

## 长期依赖（Long-Term Dependencies）问题

RNN 的关键点之一就是他们可以用来连接先前的信息到当前的任务上，例如使用过去的视频段来推测对当前段的理解。如果 RNN 可以做到这个，他们就变得非常有用。但是真的可以么？答案是，还有很多依赖因素。

有时候，我们仅仅需要知道先前的信息来执行当前的任务。例如，我们有一个语言模型用来基于先前的词来预测下一个词。如果我们试着预测 “the clouds are in the sky” 最后的词，我们并不需要任何其他的上下文 —— 因此下一个词很显然就应该是 sky。在这样的场景中，相关的信息和预测的词位置之间的间隔是非常小的，RNN 可以学会使用先前的信息。

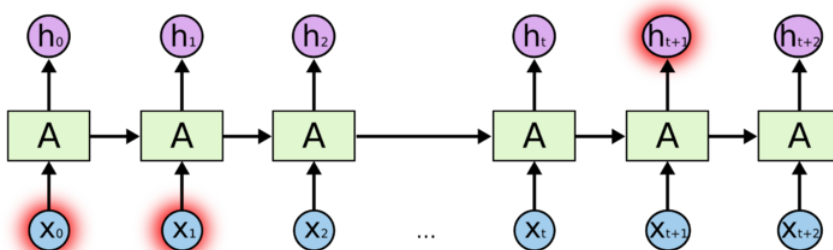




不太长的相关信息和位置间隔

但是同样会有一些更加复杂的场景。假设我们试着去预测“I grew up in France... I speak fluent French”最后的词。当前的信息建议下一个词可能是一种语言的名字，但是如果我们h需要弄清楚是什么语言，我们是需要先前提到的离当前位置很远的 France 的上下文的。这说明相关信息和当前预测位置之间的间隔就肯定变得相当的大。

不幸的是，在这个间隔不断增大时，RNN 会丧失学习到连接如此远的信息的能力。



相当长的相关信息和位置间隔

在理论上，RNN 绝对可以处理这样的 长期依赖 问题。人们可以仔细挑选参数来解决这类问题中的最初级形式，但在实践中，RNN 肯定不能够成功学习到这些知识。Bengio, et al. (1994) (<https://link.jianshu.com?t=http://www-dsi.ing.unifi.it/~paolo/ps/tnn-94-gradient.pdf>)等人对该问题进行了深入的研究，他们发现一些使训练 RNN 变得非常困难的相当根本的原因。

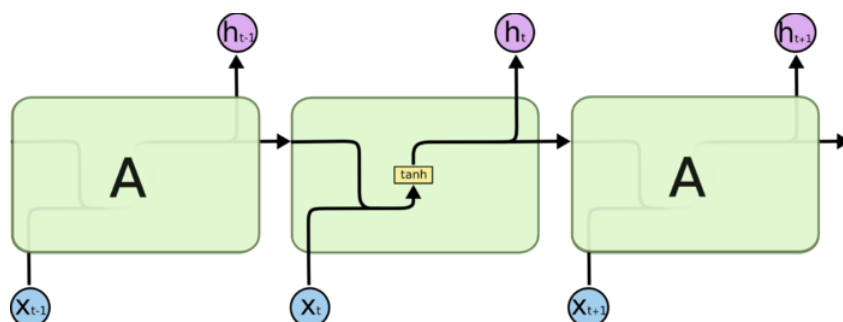
然而，幸运的是，LSTM 并没有这个问题！

## LSTM 网络

Long Short Term 网络——一般就叫做 LSTM ——是一种 RNN 特殊的类型，可以学习长期依赖信息。LSTM 由Hochreiter & Schmidhuber (1997) ([https://link.jianshu.com?t=http://deeplearning.cs.cmu.edu/pdfs/Hochreiter97\\_lstm.pdf](https://link.jianshu.com?t=http://deeplearning.cs.cmu.edu/pdfs/Hochreiter97_lstm.pdf))提出，并在近期被Alex Graves (<https://link.jianshu.com?t=https://scholar.google.com/citations?user=DaFHynwAAAAJ&hl=en>)进行了改良和推广。在很多问题，LSTM 都取得相当巨大的成功，并得到了广泛的使用。

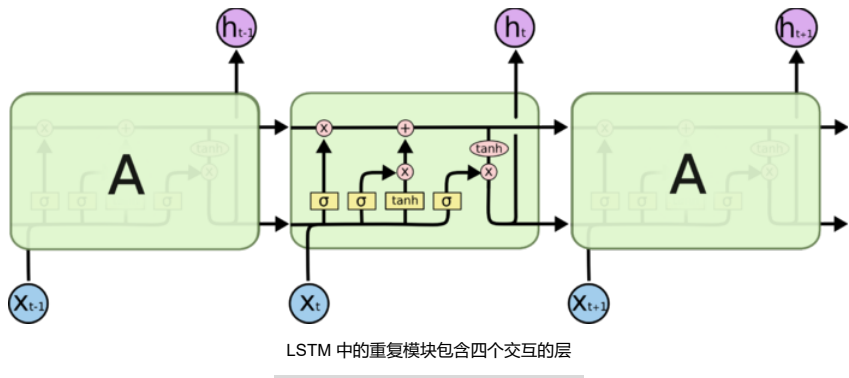
LSTM 通过刻意的设计来避免长期依赖问题。记住长期的信息在实践中是 LSTM 的默认行为，而非需要付出很大代价才能获得的能力！

所有 RNN 都具有一种重复神经网络模块的链式的形式。在标准的 RNN 中，这个重复的模块只有一个非常简单的结构，例如一个 tanh 层。



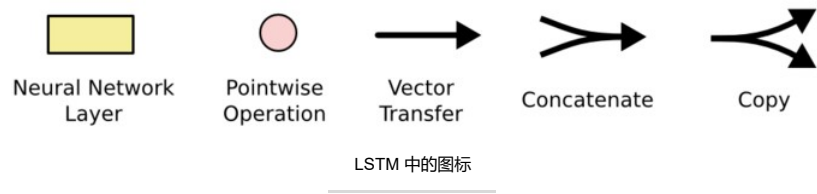
标准 RNN 中的重复模块包含单一的层

LSTM 同样是这样的结构，但是重复的模块拥有一个不同的结构。不同于 单一神经网络层，这里是有四个，以一种非常特殊的方式进行交互。



LSTM 中的重复模块包含四个交互的层

不必担心这里的细节。我们会一步一步地剖析 LSTM 解析图。现在，我们先来熟悉一下图中使用的各种元素的图标。

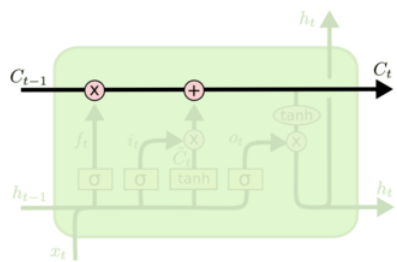


在上面的图例中，每一条黑线传输着一整个向量，从一个节点的输出到其他节点的输入。粉色的圈代表 pointwise 的操作，诸如向量的和，而黄色的矩阵就是学习到的神经网络层。合在一起的线表示向量的连接，分开的线表示内容被复制，然后分发到不同的位置。

LSTM 的核心思想

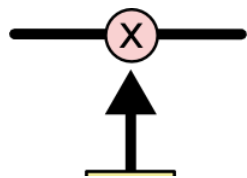
LSTM 的关键就是细胞状态，水平线在图上方贯穿运行。

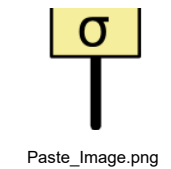
细胞状态类似于传送带。直接在整个链上运行，只有一些少量的线性交互。信息在上面流传保持不变会很容易。



Paste\_Image.png

LSTM 有通过精心设计的称作为“门”的结构来去除或者增加信息到细胞状态的能力。门是一种让信息选择式通过的方法。他们包含一个 sigmoid 神经网络层和一个 pointwise 乘法操作。





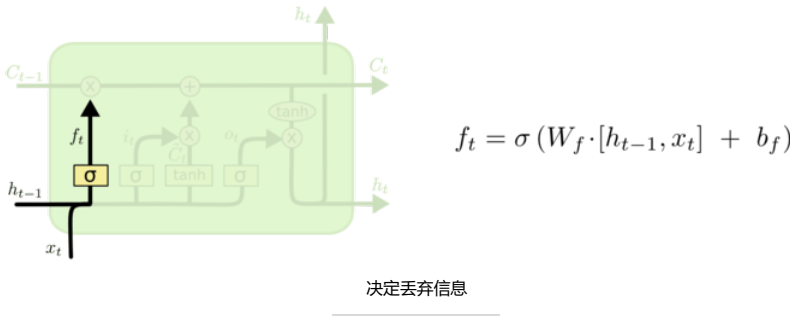
Sigmoid 层输出 0 到 1 之间的数值，描述每个部分有多少量可以通过。0 代表“不许任何量通过”，1 就指“允许任意量通过”！

LSTM 拥有三个门，来保护和控制细胞状态。

### 逐步理解 LSTM

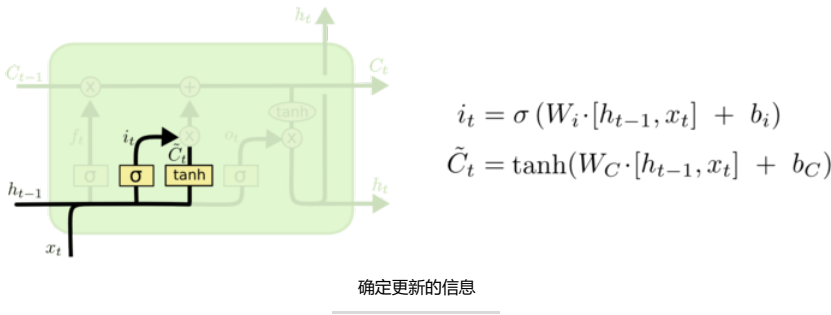
在我们 LSTM 中的第一步是决定我们会从细胞状态中丢弃什么信息。这个决定通过一个称为**忘记门层**完成。该门会读取  $h_{t-1}$  和  $x_t$ ，输出一个在 0 到 1 之间的数值给每个在细胞状态  $c_{t-1}$  中的数字。1 表示“完全保留”，0 表示“完全舍弃”。

让我们回到语言模型的例子中来基于已经看到的预测下一个词。在这个问题中，**细胞状态可能包含当前主语的性别，因此正确的代词可以被选择出来。当我们看到新的主语，我们希望忘记旧的主语。**



下一步是确定什么样的新信息被存放在细胞状态中。这里包含两个部分。第一，sigmoid 层称“输入门层”决定什么值我们将要更新。然后，一个 tanh 层创建一个新的候选值向量， $\tilde{c}_t$ ，会被加入到状态中。下一步，我们会讲这两个信息来产生对状态的更新。

在我们语言模型的例子中，我们希望增加新的主语的性别到细胞状态中，来替代旧的需要忘记的主语。

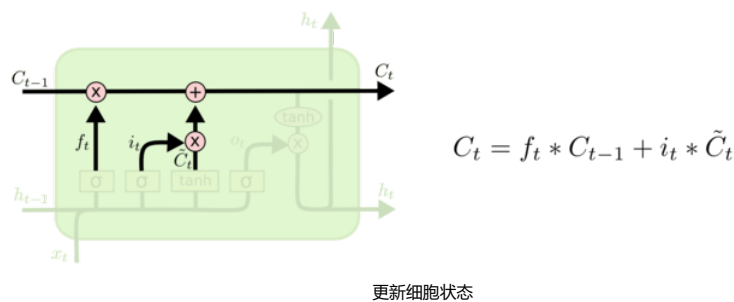


现在是更新旧细胞状态的时间了， $c_{t-1}$  更新为  $c_t$ 。前面的步骤已经决定了将会做什么，我们现在就是实际去完成。

我们把旧状态与  $f_t$  相乘，丢弃掉我们确定需要丢弃的信息。接着加上  $i_t * \tilde{c}_t$ 。这就是新的候选值，根据我们决定更新每个状态的程度进行变化。

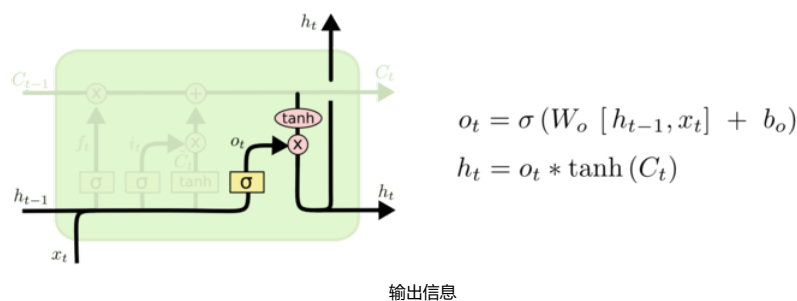
在语言模型的例子中，这就是我们实际根据前面确定的目标，丢弃旧代词的性别信息并添加新的信息的地方。





最终，我们需要确定输出什么值。这个输出将会基于我们的细胞状态，但是也是一个过滤后的版本。首先，我们运行一个 sigmoid 层来确定细胞状态的哪个部分将输出出去。接着，我们把细胞状态通过 tanh 进行处理（得到一个在 -1 到 1 之间的值）并将它和 sigmoid 门的输出相乘，最终我们仅仅会输出我们确定输出的那部分。

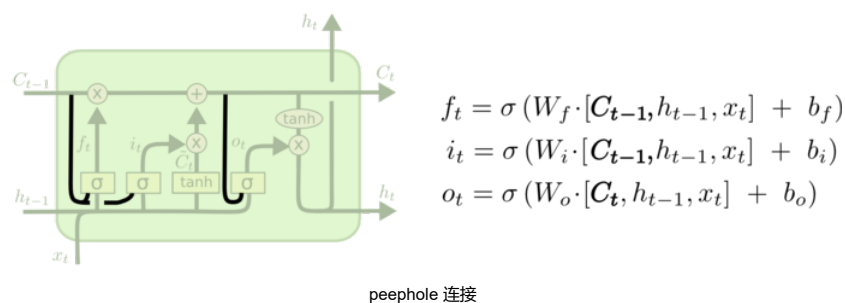
在语言模型的例子中，因为他就看到了一个 **代词**，可能需要输出与一个 **动词** 相关的信息。例如，可能输出是否代词是单数还是复数，这样如果是动词的话，我们也知道动词需要进行的词形变化。



## LSTM 的变体

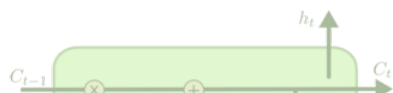
我们到目前为止都还在介绍正常的 LSTM。但是不是所有的 LSTM 都长成一个样子的。实际上，几乎所有包含 LSTM 的论文都采用了微小的变体。差异非常小，但是也值得拿出来讲一下。

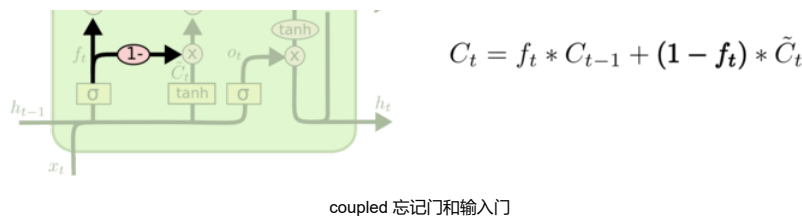
其中一个流形的 LSTM 变体，就是由 Gers & Schmidhuber (2000) (<https://link.jianshu.com?t=ftp://ftp.idsia.ch/pub/juergen/TimeCount-IJCNN2000.pdf>) 提出的，增加了“peephole connection”。是说，我们让门层也会接受细胞状态的输入。



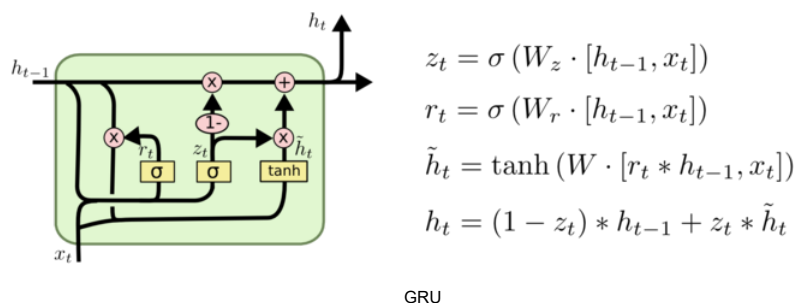
上面的图例中，我们增加了 peephole 到每个门上，但是许多论文会加入部分的 peephole 而非所有都加。

另一个变体是通过使用 coupled 忘记和输入门。不同于之前是分开确定什么忘记和需要添加什么新的信息，这里是一同做出决定。我们仅仅会当我们将要输入在当前位置时忘记。我们仅仅输入新的值到那些我们已经忘记旧的信息的那些状态。





另一个改动较大的变体是 Gated Recurrent Unit (GRU), 这是由 Cho, et al. (2014) (<http://link.jianshu.com?t=http://arxiv.org/pdf/1406.1078v3.pdf>) 提出。它将忘记门和输入门合成了一个单一的 更新门。同样还混合了细胞状态和隐藏状态, 和其他一些改动。最终的模型比标准的 LSTM 模型要简单, 也是非常流行的变体。



这里只是部分流行的 LSTM 变体。当然还有很多其他的, 如 Yao, et al. (2015) (<https://link.jianshu.com?t=http://arxiv.org/pdf/1508.03790v2.pdf>) 提出的 Depth Gated RNN。还有用一些完全不同的观点来解决长期依赖的问题, 如 Koutnik, et al. (2014) (<https://link.jianshu.com?t=http://arxiv.org/pdf/1402.3511v1.pdf>) 提出的 Clockwork RNN。

要问哪个变体是最好的? 其中的差异性真的重要吗? Greff, et al. (2015) (<https://link.jianshu.com?t=http://arxiv.org/pdf/1503.04069.pdf>) 给出了流行变体的比较, 结论是他们基本上是一样的。Jozefowicz, et al. (2015) (<https://link.jianshu.com?t=http://jmlr.org/proceedings/papers/v37/jozefowicz15.pdf>) 则在超过 1 万种 RNN 架构上进行了测试, 发现一些架构在某些任务上也取得了比 LSTM 更好的结果。

### 3.4. Some Statistics

We evaluated 10,000 different architectures and 1,000 of them made it past the initial filtering stage on the memorization problem. Each such architecture has been evaluated on 220 hyperparameter settings on average. These figures suggest that every architecture that we evaluated had a reasonable chance at achieving its highest performance. Thus we evaluated 230,000 hyperparameter configurations in total.

Jozefowicz等人论文截图

## 结论

刚开始, 我提到通过 RNN 得到重要的结果。本质上所有这些都可以使用 LSTM 完成。对于大多数任务确实展示了更好的性能!

由于 LSTM 一般是通过一系列的方程表示的, 使得 LSTM 有一点令人费解。然而本文中一步一步地解释让这种困惑消除了不少。

LSTM 是我们在 RNN 中获得的重要成功。很自然地, 我们也会考虑: 哪里会有更加重大的突破呢? 在研究人员间普遍的观点是: “Yes! 下一步已经有了——那就是**注意力**!” 这个想法是让 RNN 的每一步都从更加大的信息集中挑选信息。例如, 如果你使用 RNN 来产生一个图片的描述, 可能会选择图片的一个部分, 根据这部分信息来产生输出的词。实际上, Xu, et al. (2015) (<https://link.jianshu.com?t=http://arxiv.org/pdf/1502.03044v2.pdf>) 已经这么做了——如果你希望深入探索**注意力**可能这就是一个有趣的起点! 还有一些使用注意力的相当振奋人心的研究成果, 看起来有更多的东西亟待探索……



注意力也不是 RNN 研究领域中的唯一的发展方向。例如，Kalchbrenner, *et al.* (2015) (<https://link.jianshu.com?t=http://arxiv.org/pdf/1507.01526v1.pdf>) 提出的 Grid LSTM 看起来也是很有前途。使用生成模型的 RNN，诸如Gregor, *et al.* (2015) (<https://link.jianshu.com?t=http://arxiv.org/pdf/1502.04623.pdf>) Chung, *et al.* (2015) (<https://link.jianshu.com?t=http://arxiv.org/pdf/1506.02216v3.pdf>) 和 Bayer & Osendorfer (2015) (<https://link.jianshu.com?t=http://arxiv.org/pdf/1411.7610v3.pdf>) 提出的模型同样很有趣。在过去几年中，RNN 的研究已经相当的燃，而研究成果当然也会更加丰富！

## 致谢

I'm grateful to a number of people for helping me better understand LSTMs, commenting on the visualizations, and providing feedback on this post.

I'm very grateful to my colleagues at Google for their helpful feedback, especially Oriol Vinyals (<https://link.jianshu.com?t=http://research.google.com/pubs/OriolVinyals.html>), Greg Corrado (<https://link.jianshu.com?t=http://research.google.com/pubs/GregCorrado.html>), Jon Shlens (<https://link.jianshu.com?t=http://research.google.com/pubs/JonathonShlens.html>), Luke Vilnis (<https://link.jianshu.com?t=http://people.cs.umass.edu/~luke/>), and Ilya Sutskever (<https://link.jianshu.com?t=http://www.cs.toronto.edu/~ilya/>). I'm also thankful to many other friends and colleagues for taking the time to help me, including Dario Amodei (<https://link.jianshu.com?t=https://www.linkedin.com/pub/dario-amodei/4/493/393>), and Jacob Steinhardt (<https://link.jianshu.com?t=http://cs.stanford.edu/~jsteinhardt/>). I'm especially thankful to Kyunghyun Cho (<https://link.jianshu.com?t=http://www.kyunghyuncho.me/>) for extremely thoughtful correspondence about my diagrams. Before this post, I practiced explaining LSTMs during two seminar series I taught on neural networks. Thanks to everyone who participated in those for their patience with me, and for their feedback.

人工智能时代每个人都将面临挑战，想要了解更多相关知识和实践经验，请关注公众号“UniversityAI”。

UAI 人工智能

您的支持可以鼓励作者写出更多的文章。

赞赏支持



deep learning (/nb/339523)

举报文章 © 著作权归作者所有

朱小虎Neil (/u/696dc6c6f01c) ♂

+ 关注








写了 476724 字，被 3480 人关注，获得了 1929 个喜欢

(/u/696dc6c6f01c)

我是 Neil 朱小虎，University AI 创始人 & Chief Scientist，UniversityAI-AI-Unconference Meetup 组织者，...

喜欢 | 523



更多分享

(http://cwb.assets.jianshu.io/notes/images/1936343



下载简书 App ▶  
随时随地发现和创作内容



(/apps/download?utm\_source=nbc)




登录 (/sign-in?utm\_source=desktop&utm\_medium=not-signed-in-comment-form)

81条评论


只看作者


按喜欢排序 按时间正序 按时间倒序




friskit (/u/bdca2bda41b1)  
20楼 · 2017.01.18 01:28

(/u/bdca2bda41b1)  
感谢作者的翻译，不过这里有个疑问，是针对olah的文章的。原文中LSTM的这个单元，如果两条线汇聚代表“concatenate”的话，最初 $x_t$ 和 $h_{t-1}$ 两者连接之后，后面的维度都会加倍，这样 $h_t$ 的维度其实就是 $h_{t-1}$ 的二倍，如此一来，维度越扩越大啊。相反，看了Google在Udacity上的Deep Learning教程的关于LSTM的例子，他们的做法不是concatenate，而是把二者相加，保持维度相等。所以感觉olah的这个写法是不是有些问题？


6人赞  回复

friskit (/u/bdca2bda41b1): 把公式整理了一下之后发现我差不多明白了。。主要的秘诀在于权值矩阵。比如说， $x$ 和 $h$ 的维度是 $[m, 1]$ ，两者concatenate之后维度是 $[2m, 1]$ ，所以只需要 $w$ 的维度是 $[m, 2m]$ ， $b$ 的维度是 $[m, 1]$ 就可以使得运算后的维度是 $[m, 1]$ 了。  
2017.01.18 01:41  回复


-----


jianshuhu (/u/c4fbd798203e): @friskit (/users/bdca2bda41b1) 原来是这样，谢谢！  
2017.07.04 10:52  回复

-----

ApacheCN\_飞龙 (/u/b508a6aa98eb): 假设你有个运算是  $AB + CD$ ，四个都是矩阵，那么等价于  $AC$  左右拼接，乘上  $BD$  上下拼接。  
2018.01.11 23:11  回复


-----


 添加新评论



大姨夫司机 (/u/d986d284aade)  
2楼 · 2015.08.28 14:10

(/u/d986d284aade)  
不是我吹逼呀 这排版很舒服了

4人赞  回复

朱小虎Neil (/u/696dc6c6f01c): @大姨夫斯基 (/users/d986d284aade) 就第一幅图有点大。。。  
2015.08.28 14:46  回复



DediWho (/u/d4e584126d8d): @Not\_GOD (/users/696dc6c6f01c) 图文并貌好文章。不过就是有些图因为上面的字小，想点开放大看，但因为放大后背景是黑的，所以一些模块的黑色连接线看不到了

2015.08.29 06:18 回复

添加新评论



zgkk (/u/e51265bc0f3f)

15楼 · 2016.08.12 20:47

(/u/e51265bc0f3f)

“让我们回到语言模型的例子中来基于已经看到的预测下一个词。在这个问题中，细胞状态可能包含当前主语的类别，因此正确的代词可以被选择出来。当我们看到新的代词，我们希望忘记旧的代词。”是不是应该是“当我们看到新的主语”吧

3人赞 回复

熊15 (/u/d12fd6e50d91): @zgkk (/users/e51265bc0f3f) 作者把原文的“subject”都翻译成了“代词”，个人觉着 翻译成“主语”更好！赞同！

2016.08.21 10:25 回复

朱小虎Neil (/u/696dc6c6f01c): @熊15 (/users/d12fd6e50d91) @zgkk (/users/e51265bc0f3f) 感谢两位指出，已经改好。~~ 😊

2016.08.21 16:25 回复

熊15 (/u/d12fd6e50d91): @Not\_GOD (/users/696dc6c6f01c) 感谢大神的分享！

2016.08.21 16:58 回复

添加新评论 | 还有1条评论，展开查看



hualitlc (/u/de1799e4fd76)

7楼 · 2016.01.18 19:46

(/u/de1799e4fd76)

翻译的很好，很有用的干货，辛苦了。有一个小错误，“在我们语言模型的例子中，我们希望增加新的代词的类别到细胞状态中，来替代旧的需要忘记的代词”下边的配图第二个公式应该是  $C_t = \tanh(W_c [h_{t-1}, x_t] + b_c)$ 。

1人赞 回复

朱小虎Neil (/u/696dc6c6f01c): @hualitlc (/users/de1799e4fd76) 谢谢指出，已经改了。

2016.03.28 12:17 回复

添加新评论



fighting41love (/u/69d0f878050c)

19楼 · 2017.01.05 06:12

(/u/69d0f878050c)

首先，非常感谢作者能够花时间翻译此文。辛苦了！

不过这里有一小问题需要讨论：

文中这一部分：前主语的类别，因此正确的代词可以被选择出来。当我们看到新的主语，我们希望忘记旧的主语。

这句翻译的有点问题。我猜作者是说 主语的性别，后面找正确的代词，she or he的意思，应该不是类别，是性别。

1人赞 回复

朱小虎Neil (/u/696dc6c6f01c): 感谢，已经修改。

2017.01.05 10:49 回复

添加新评论





我\_a249 (/u/680aefadf62b)

25楼 · 2017.03.02 11:44

(/u/680aefadf62b)

这里的 $W[h,x]$ 是什么意思，讲 $h$ 与 $x$ 拼接成一个更长的向量？

1人赞 回复

朱小虎Neil (/u/696dc6c6f01c): 对的

2017.03.02 20:41 回复

朱小虎Neil (/u/696dc6c6f01c): 本文中将  $h$  和  $x$  的权重矩阵整合在一个了，可能和其他一些地方的论文写得不同：)

你可以展开一个 example 的公式看看

2017.03.02 21:38 回复

我\_a249 (/u/680aefadf62b): @Not\_GOD (/users/696dc6c6f01c) 晓得了，谢谢

😊 (我没有团队，一个人做这个循环网络，理解起来比较费劲，一个细节就够我折腾一天的，有好的教程可以推荐一下么) 🙄

2017.03.18 17:30 回复

添加新评论



DediWho (/u/d4e584126d8d)

4楼 · 2015.08.29 05:59

(/u/d4e584126d8d)

已在微信打赏

赞 回复

朱小虎Neil (/u/696dc6c6f01c): @DediWho (/users/d4e584126d8d) 🙏

2015.08.29 23:51 回复

添加新评论



jiangnan\_hugo (/u/97a7c51e9f7b)

5楼 · 2015.11.23 17:21

(/u/97a7c51e9f7b)

翻译的好。

赞 回复

朱小虎Neil (/u/696dc6c6f01c): @jiangnan\_hugo (/users/97a7c51e9f7b) 谢谢

2015.11.23 18:29 回复

添加新评论



benpum (/u/2c4a7cc61d16)

6楼 · 2016.01.04 12:35

(/u/2c4a7cc61d16)

m

赞 回复



Soldier\_Gary (/u/6c47a7eb6b4e)

8楼 · 2016.04.26 15:59

(/u/6c47a7eb6b4e)

翻译的很棒! 👍

赞 回复



被以下专题收入，发现更多相似内容



关于搜索，我们聊聊 (/c/ba90746aebc2?

utm\_source=desktop&utm\_medium=notes-included-collection)



今日看点 (/c/3sT4qY?utm\_source=desktop&utm\_medium=notes-included-

collection)



iOS学习笔记 (/c/bb62bc7671d2?

utm\_source=desktop&utm\_medium=notes-included-collection)



好东西 (/c/e77ce8b808e2?utm\_source=desktop&utm\_medium=notes-

included-collection)



码源 (/c/9150acdad06d?utm\_source=desktop&utm\_medium=notes-

included-collection)



已打赏 (/c/109142bca086?utm\_source=desktop&utm\_medium=notes-

included-collection)



疯狂的学习 (/c/87fb8a4db6f7?utm\_source=desktop&utm\_medium=notes-

included-collection)

展开更多 ∨

推荐阅读

更多精彩内容 > (/)

**苏幕遮·放怀 (/p/2cf5564c7f0e?utm\_campaign=male...**

(/p/2cf5564c7f0e?

青层山，绿池水，长风过面，面香暖花开。桥连远山云行路，草木濡沫，定处远山间。西子美，明夜至。登船再览，常忆故游时。十月秋霜裹枫露，卷开心门，远胜茶花浓。

utm\_campaign=maleskine&utm\_content=note&utr

朱小虎Neil (/u/696dc6c6f01c?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=pc\_all\_hots&utm\_source=recommendation)

**月迹 (/p/5a70b7decc33?utm\_campaign=maleskine&utm\_content=note...**

红玉寻不见但知身何处移步楼宇隙静候待其出刹间露牙白渐渐吐新尘满轮光晕耀大愿终落成

朱小虎Neil (/u/696dc6c6f01c?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=pc\_all\_hots&utm\_source=recommendation)

**在体制内熬到30岁才想跳槽的人，后来到底怎么样了？ ...**

(/p/a222ead54a36?

文|赵晓璃 写在前面的话：借用钱钟书先生那句经典的话：“城里的人想出去，城外的人想进来”。用来形容体制内外再妥帖不过。随着时代迅猛的发展，一方面人们比任何时候都渴望稳定，另一方面很多稳定的单位正渐渐失去原有的优势。不论是出于自我觉醒还是被迫选择，越来越多的体制内从业人员开始思考一个问题，那就是，如果一旦“铁饭碗”被打破，我该如何生存？今年年初，我陆续陆续收到了体制内转行成功的咨询者反馈，选取了两则典型案例加工整理，试图给读者呈现一些真实的经历；并在此基础上做了些归纳与总结，希望能给更多的职场人一些启发和帮助。一、“我发现这是一片沼泽，并且越陷越深”因为家人的强烈要求，咨询...

utm\_campaign=maleskine&utm\_content=note&utr

赵晓璃 (/u/8c97c482455f?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=pc\_all\_hots&utm\_source=recommendation)

**这才是向生活致敬的最好方式！ | 1月早餐合辑 (/p/d06...**

(/p/d06bb27c86e9?

电影《蒂凡尼的早餐》中奥黛丽·赫本饰演的女主虽然生活乱糟糟，但对于早餐却有独特的仪式感。每当心神不宁、迷茫惶恐时，她都会打扮精致，穿上小黑裙，打车到蒂凡尼珠宝店，一边吃着面包，一边欣赏珠宝。将早餐吃出了自己的仪式感，给这个迷茫的追梦者多了一些勇气。对于仪式感，以前我们家也做的不好，这种对于生活无所谓或者怕麻烦的心态，让我错失了很多值得被记录的日子，不怕你们笑话，二十几年来我们家从来没有一起过生日，我长这么大也没真

utm\_campaign=maleskine&utm\_content=note&utr



正的吃过自己的生日蛋糕。这就是我原生家庭带给我的，然而，想要改变只需要立刻去做就可以！ 电影《小王子》里小王子说：“仪式感，就是使某一天与其他日子不同，使某一时刻与其他时刻不同。...

潇逸霏 (/u/cf4fa3c04bb9?  
utm\_campaign=maleskine&utm\_content=user&utm\_medium=pc\_all\_hots&utm\_source=recommendation)

简宝玉分享会 | 如何通过写作开启斜杠人生，赚取写作... (/p/6a83c0245661?  
utm\_campaign=maleskine&utm\_content=note&utr

-0- 嘿，大家好，我是用时间酿酒。 本周日（3.25）晚8:00——9:00，我将联合简书举行一场关于写作的免费分享。主题就是文章的标题：如何通过写作开启斜杠人生，赚取写作的第一桶金？ 详细的分享大纲和参与方式已经放在文末，大家可以拉到最后报名。本次分享已经结束，文字稿已经更新简书，点击直接查看：分享干货版 | 如何通过写作开启斜杠人生，赚取写作的第一桶金？ -1- 现在想跟大家“闲聊”几句，谈谈我对写作的看法。一直以来我都在问自己一个问题：写作到底带给了我什么？ 一开始我的答案是写作督促我更高质量地阅读。我的专业是统计学，大学四年，主要是和数字打交道。所以听到我写文，身边的朋友都不...

用时间酿酒 (/u/0fe9f776f37a?  
utm\_campaign=maleskine&utm\_content=user&utm\_medium=pc\_all\_hots&utm\_source=recommendation)

(/p/87fd419c94cc?



utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)

多图万字文 | 从神经元到CNN、RNN、GAN...神经网络看本文绝对够了 (/p/...

作者 | FJODOR VAN VEEN 编译 | AI100 (ID: rgznai100) 在深度学习十分火热的今天，不时会涌现出各种新型的人工神经网络，想要实时了解这些新型神经网络的架构还真是不容易。光是知道各式各样的神经网络模型缩写（如：DCIGN、BiLSTM、DCG...

AI科技大本营 (/u/da69420ec62d?  
utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

(/p/f054f8daec68?



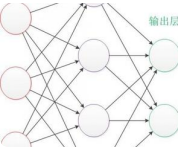
utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)

多图 | 从神经元到CNN、RNN、GAN...神经网络看本文绝对够了 (/p/f054f8...

作者 | FJODOR VAN VEEN 编译 | AI100 (ID: rgznai100) 在深度学习十分火热的今天，不时会涌现出各种新型的人工神经网络，想要实时了解这些新型神经网络的架构还真是不容易。光是知道各式各样的神经网络模型缩写（如：DCIGN、BiLSTM、DCG...

AI科技大本营 (/u/da69420ec62d?  
utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

(/p/77e69ff5cfae?



utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)

重磅！神经网络浅讲：从神经元到深度学习 (/p/77e69ff5cfae?utm\_campai...

神经网络是一门重要的机器学习技术。它是目前最为火热的研究方向-深度学习的基础。学习神经网络不仅可以让你掌握一门强大的机器学习方法，同时也可以更好地帮助你理解深度学习技术。本文以一种简单的，循序的方式讲解神经网络。适合对神经网络了解不多的同学。本文对阅读没有一定的前提要求，但...

BURIBURI\_ZAEMON (/u/00d1ed2b53ae?  
utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)



(/p/e834a5e22028?

utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)

## 深度学习笔记 - 理解LSTM网络 (/p/e834a5e22028?utm\_campaign=males...

转载：《理解 LSTM 网络 | 我爱计算机》Recurrent Neural Networks 人类对事物的理解并不是从一片空白开始的，往往都是基于对先前所见所闻的理解来对看到的进行推断；我们的思想有持久性，对市场历史的理解对将来的走势也具有一定的参考和借鉴。传统的神经...



Kimichen7764 (/u/a068f7c95c7b?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

## 译 理解RNN和LSTM网络 (/p/f0bcd862eef1?utm\_campaign=maleskine&...

循环神经网络(RNN) 人们的每次思考并不都是从零开始的。比如说你在阅读这篇文章时，你基于对前面的文字的理解来理解你目前阅读到的文字，而不是每读到一个文字时，都抛弃掉前面的思考，从头开始。你的记忆是有持久性的。传统的神经网络并不能如此，这似乎是一个主要的缺点。例如，假设你...



云时之间 (/u/7ebb78d8c38a?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

## 【0418 晨读感悟 大红花】先定一个小小目标 (/p/47303a3aa1cf?utm\_ca...

一个习惯的养成平均要66天，不同的习惯需要的时间相差很大，从18天到254天不等。有时候踏出第一步，比坚持下去还要难。要么拖延，要么没有勇气，迟迟不肯尝试。当然，多数人还是坚持不下去。定了一个目标，过了“鸡血期”就想放弃了，而且总有理由说服自己。其实，有时候是被自己定...



我奉你为神 (/u/8aab287f8b02?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

(/p/bb3212f83bc3?

utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)

## 三十，四十 (21) (/p/bb3212f83bc3?utm\_campaign=maleskine&utm\_c...

微信订阅号有个《夜听》，每晚十点和大家见面，昨晚的话题是别让爱成为一种伤害，大体内容是：人生中有许多种爱，但别让爱成为一种伤害。如果真诚是一种伤害，我选择谎言；如果谎言是一种伤害，我选则沉默；如果沉默是一种伤害，我选择离开。如果束缚是苦 你会不会选择结束 如果强...



小泥壶 (/u/c58333aef966?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

(/p/cb2b8c2a3384?

utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)

## 岁月缝花，给你一股人情味儿 (/p/cb2b8c2a3384?utm\_campaign=malesk...

01. 你没有过这样的体验，身边比我们年长的人告诉我们：你还没经历过社会的黑暗，所以你还不懂某某事、某某人.....又或是，你还没体会过人心险恶，还太单纯。诸如此类，貌似过来人对你的忠告。可是，慢慢的你长大了，你觉得事情真的是这样吗？社会的确黑暗，人心当然隔肚皮，但你依然长这...



成全某日有因的干哭 (/u/b60babf19c7e?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

(/p/6cd6be3b4781?

utm\_campaign=maleskine&utm\_content=note&utm\_medium=seo\_notes&utm\_source=recommendation)

## 曾志伟当他配角, 周星驰为他配音, 今卖保险为生! (/p/6cd6be3b4781?utm\_...

9岁时便被发掘出道，先是在香港电台拍摄了多部电视剧的演出，接到涉足大银幕，成为红极一时的童星 邓一君，曾用名邓浩君，1976年10月24日出生于香港。一张娃娃脸的邓一君，在无线剧集中常演一些主角的弟弟或是主角的少年。1993年，邓一君15岁时凭借《记得香蕉成熟时》扮演男主...



孕育讲堂 (/u/ee2ae7bea2e7?

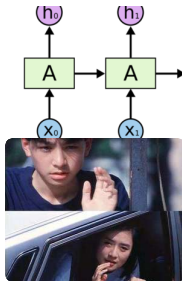
utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

## 真心话 (/p/eca90fb263c8?utm\_campaign=maleskine&utm\_content=not...

六年前，一群生在操场上玩游戏，唯独M孤身一人，老师问其原因，他支支吾吾回答“没人带我一起。”六年后，老师让学生分学习小组，还是那个少年孤身一人，老师又问其原因，他高傲的回答“我怕其他人拖我的后腿。”六年来，他对抗孤立的唯一办法就是让自己高傲起来。他成为了学院的异类，能力...



密云阿炳



(/u/4274023e57eb?

utm\_campaign=maleskine&utm\_content=user&utm\_medium=seo\_notes&utm\_source=recommendation)

