

中文译文：深度学习、自然语言处理和表征方法

<http://blog.jobbole.com/77709/>

英文原文：Deep Learning, NLP, and Representations

<http://colah.github.io/posts/2014-07-NLP-RNNs-Representations/>

### 总结：

这篇文章中主要提到了单层神经网络，单词嵌入（word embeddings），表征这几个概念，结合具体的实例，写的是通俗易懂，在引用参考文献的位置都给出了对应的链接，一些参考文献中的工作还是挺有意思的。

关于单层神经网络，介绍的浅显易懂，把神经网络比作查询表，很直观。

关于单词嵌入，文中介绍的意思是把一个词转换为一个多维向量，还用 t-SNE 工具直观的给出了单词嵌入空间的图，加上表格中的例子，更加易懂了。比较有意思的就是相同的词在单词嵌入空间中的距离是很近的，这个特点很有利用价值，利用这一点，提到了一些具体的应用场景，例如语法错误检查，性别类中不同代词（男-女，叔叔-阿姨，国王-王后等）之间的距离也是很相似的。关于这些应用，原文的作者的评价是，这些都是在研究方法的过程中所得的副产品。

关于表征 representation，单词嵌入应该就算是一种吧，文中重点介绍了共同表征/共享嵌入（shared representation），将多个不同空间的嵌入信息映射到同一个空间，并介绍了两个很好的应用场景——双语单词嵌入和图像-文字嵌入。

双语单词嵌入，把两个语言中的单词嵌入空间用图像直观的看的话，形状是相似的，相似的词在图像中所处的位置是临近的。

图像-文字嵌入，也是相同概念的信息比较接近，例如猫的图像距离文字“猫”就很近，汽车的图像距离“汽车”就很近。这里引用了斯坦福一个小组和谷歌一个小组的工作，感觉挺有意思。

最后介绍了递归神经网络以及为什么适用于 NLP。

整体来看，这篇文章更像是一篇不错的科普文章，个人认为读完之后还是有收获的，尤其是 Word Embedding 这个概念。

昨天搜了一下关于深度学习的相关博客，感觉很有难度。

感觉自己还是只知其一不知其二，不知道如何在 NLP 中使用 DL。

下面是英文原文中一些个人认为不错的概念和句子。

- 1: It's true, essentially, because the hidden layer can be used as a lookup table.
- 2: word embeddings;
- 3: It seems natural for a network to make words with similar meanings have similar vectors.
- 4: You've seen all the words that you understand before, but you haven't seen all the sentences that you understand before. So too with neural networks.
- 5: Word embeddings exhibit an even more remarkable property: analogies between words seem to be encoded in the difference vectors between words.
- 6: This general tactic – learning a good representation on a task A and then using it on a task B – is one of the major tricks in the Deep Learning toolbox. It goes by different names depending on the details: pretraining, transfer learning, and multi-task learning. One of the great strengths of this approach is that it allows the representation to learn from more than one kind of data. There's a counterpart to this trick. Instead of learning a way to represent one kind of data and using it to perform multiple kinds of tasks, we can learn a way to map multiple kinds of data into a single representation!

## 7: Shared Representations

- (1) Bilingual Word Embeddings;
- (2) Embed images and words in a single representation;

8: By merging sequences of words, A takes us from representing words to representing phrases or even representing whole *sentences*! And because we can merge together different numbers of words, we don't have to have a fixed number of inputs.