### 首次超越LSTM:Facebook 门卷积网络新模型能否取代递归模型？

马春杰

在机器翻译中，语言模型非常关键，目前被认为最好用的是**基于长短记忆网络（LSTM）**。他与经典的n-gram相比，**可以表征大型的文本，以及长距离的依存性**。

CNN有一个优点，那就是通过共享权重，由局部到整体，实现对输入“whole picture”的建模，而LSTM是通过逐帧递推的方式来建模整体，而递推过程中引入了**门机制**来进行信息的选择。

贾磊的观点是：

Facebook的这篇论文恰恰是通过在CNN技术中引入LSTM的“门机制”来解决语言顺序依存问题，是对传统cnn技术很大的丰富和完善，文章具有很高的理论价值和实践意义。

#### 模型详情

目前语音建模的主要方法都是基于递归神经网络的，Facebook AI研究院提出了一种卷积的方法来建模，引入了一个新的**门机制（gating mechanism）**，可以释放梯度传播。

**这是此类任务中，一个非递归性的方法首次超过了强大的递归模型**。

#### 结果

模型是基于两个大型的数据集：WikiText-103和谷歌Billion Word（GBW）。并且与LSTM和RNN进行横向对比，结果如下：

测试结果

测试结果

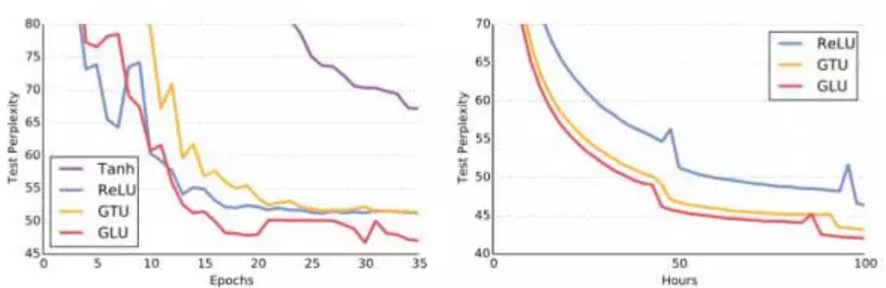


在单个GPU的情况下，GCNN的性能做到了最好。并且，据Facebook 研究者在论文中介绍，他们使用的 GCNN-13 模型拥有13层神经网络、每层包含1268个单元，而LSTM每层拥有1024个单元。

在多个GPU的情况下，只有超大型LSTM模型在性能上比GCNN好，但是超大型LSTM-2048（代表层数）使用了32个GPU，训练时间为3周，GCNN只使用1个GPU，训练时间1周。

对于一篇完整的文章来说，GCNN模型的性能也比LSTM要好得多。

测试结果



#### LSTM

LSTM 目前在行业内有着广泛的应用，范围包括但不限于：不分段连续手写识别上、自主语音识别、机器翻译等等。作为非线性模型，LSTM可作为复杂的非线性单元用于构造更大型深度神经网络。在自然语言理解上有着重要作用。

在算法模型的演进过程中，Facebook提供了一种新的思路，现在提取代还为时尚早，还是要看最终的应用效果，毕竟LSTM已经广泛应用。