论文笔记《neural paraphrase generation with stacked residual LSTM networks》

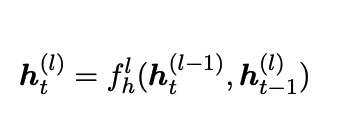
论文来源:2016 COLING

论文主要内容：

这是第一篇采用神经网络的方法做paraphrase generation。

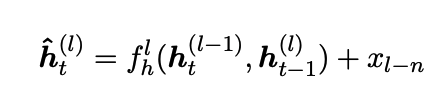
本文主要提出了a stacked residual LSTM network，给LSTM增加了residual connections。并在三个数据集PPDB，WikiAnswers，MSCOCO上进行评估，评价结果显示模型在BLEU，METEOR，TER，embedding-based similarity metric上评价均优于seq2seq，attention-based和双向LSTM。

首先需要了解下stacked LSTM的结构，第l层的隐层状态获取方式如下：

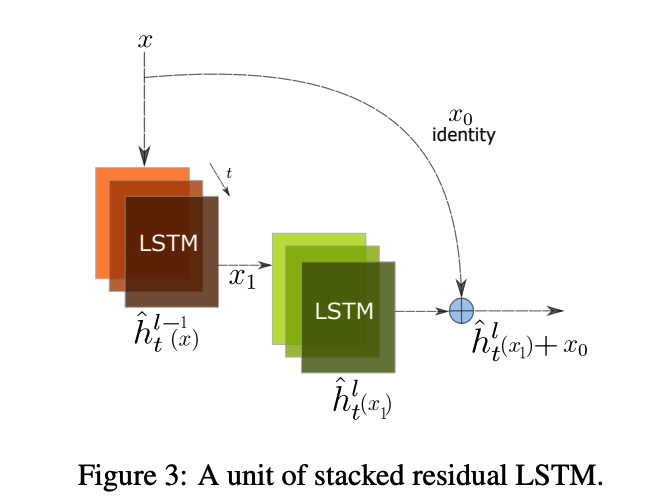


其中表示第l层在t时刻的隐层向量，表示第l层在t-1时刻的隐层向量，表示第l-1层在t时刻的隐层向量.

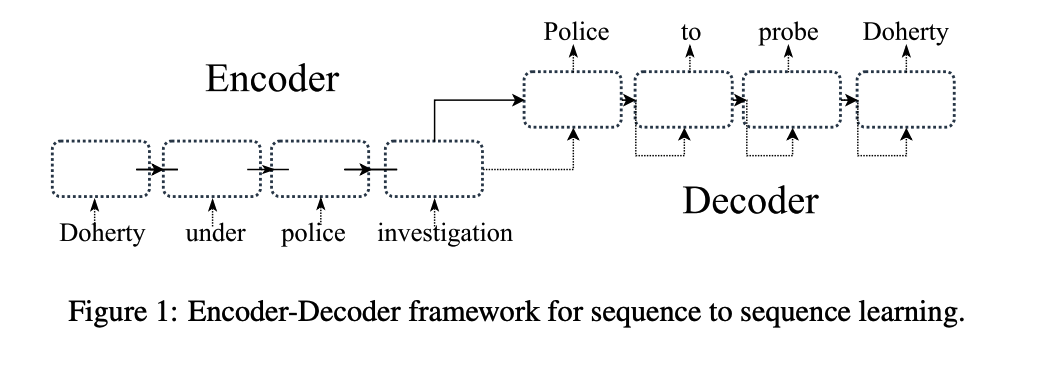
接下来的a stacked residual LSTM networks就是将残差引入多层LSTM中，



其中的表示第i+1层的输入，公式中用的是，所以在第n层之后的lstm都加了残差。具体的stacked residual LSTM结构如下：



作者的训练框架采用的是encoder-decoder框架



实验结果如下：

