

# lstm

wanghao

2018 年 11 月 21 日

## 1 lstm 相关论文

### 1.1 An Empirical Exploration of Recurrent Network Architectures

作者希望找到一种最优的 RNN 结构，或者证明 lstm 是最好的 RNN 网络。作者做了大量的实验，然而并没有找到这种网络结构。

作者在三种任务上评估了多种 RNN 结构，找到了三种表现最好的网络结构，分别为 MUT1,MUT2,MUT3, 这三种结构都跟 GRU 很相似。本文得到了以下几个结论：

1. GRU 除了语言模型在所有的任务上都表现得比 LSTM 好。
2. 在 forget 门使用较大偏移的 LSTM 在大多数任务上表现得比 GRU 好。
3. 三种门的重要性：遗忘门 > 输入门 > 输出门, (但在音乐数据集中移除输入门和输出门表现最好)。
4. 给遗忘门添加一个很大的偏移能够有效提高 LSTM 的效果。