

8.2 Report

1. Teacher Forcing:

1.1 一般地，我们会随机从预料中取一个句子的开头作为第一个 token / 采用EOS / SOS作为第一个input，以这个句子的每个 token 作为 ground-truth，来进行 loss 的计算。

问题：如果我们一个词生成的不是很好，会导致后面生成的词以及句子效果不好，因此采用ground-truth进行纠正，而不是根据模型自己之前生成的词，这样子，即使中途生成的词语不对，也可以work

1.2 eacher Force 也有一个问题，就是 Train 和 Inference 阶段的不同：Train 阶段是有 ground-truth 给你纠偏，但是 Inference阶段没有这个，因此我们可以采用概率的方式进行train

2. Seq2Seq beam search

在inference阶段，由于采用的argmax会导致高频的生成（默认已经softmax），因此在需要对这个进行改进，所以采用beam search。

具体的请见：http://gluon-nlp.mxnet.io/api/notes/beam_search_generation.html

3.coherence

思考 明日实现

4.今日结果：

```
943
944 epoch: 50
945 [input]腿 搁 在 办公 桌 上
946 [output]我会努力
947
948 [expect]这 五六 分钟 我 只 关心 说唱
949
950 [input]这 五六 分钟 我 只 关心 说唱
951 [output]我会努力
952
953 [expect]大脑 像 硬币 投进 的 湖 里 的 波浪 激荡
954
955 [input]大脑 像 硬币 投进 的 湖 里 的 波浪 激荡
956 [output]我不会像曾经一样为别人而脱衣
957
958 [expect]如果 你们 智商 170 和 我 一样
959
960 [input]耳边 的 铃铛 晶晶 亮 轻轻 唱 interesting 的 声音 记 在 了 心上
961 [output]嘻嘻哈哈爸爸妈妈不喜欢嘻哈不喜欢我说脏话
962
963 [expect]倒 掉 这 鸡 汤 错 的 路 走 几 趟 开心 难过 换 来 换 去 像 是 在 打 乒乓
964
965
```