**RNN作业报告**

**2253372 康嘉玮**

1. 网络构建

题目要求在pytorch和tensorflow两个版本中任意选择一个填空。此处选择的是tensorflow版。代码部分详见提交的代码文件。

1. 名词解释

2.1 RNN

指循环神经网络。与前馈网络的区别是增加了记忆能力，部分神经元的输出不仅与当前输入有关，还与自身的前一次或几次输出有关。

2.2 LSTM

指长短期记忆神经网络。与普通的RNN网络相比，增加了门控机制并引入一个新的内部状态（记忆单元），通过输入门、遗忘门和输出门控制记忆单元的信息传递。

2.3 GRU

指门控循环单元网络。与普通RNN网络和LSTM网络相比，它不引入新的记忆单元，通过一个更新门控制保留的信息。

1. RNN网络生成诗歌的过程

由于需要给出的是诗歌生成的过程，因此不叙述数据处理与训练过程，直接讨论通过训练好的既有模型进行生成的过程。

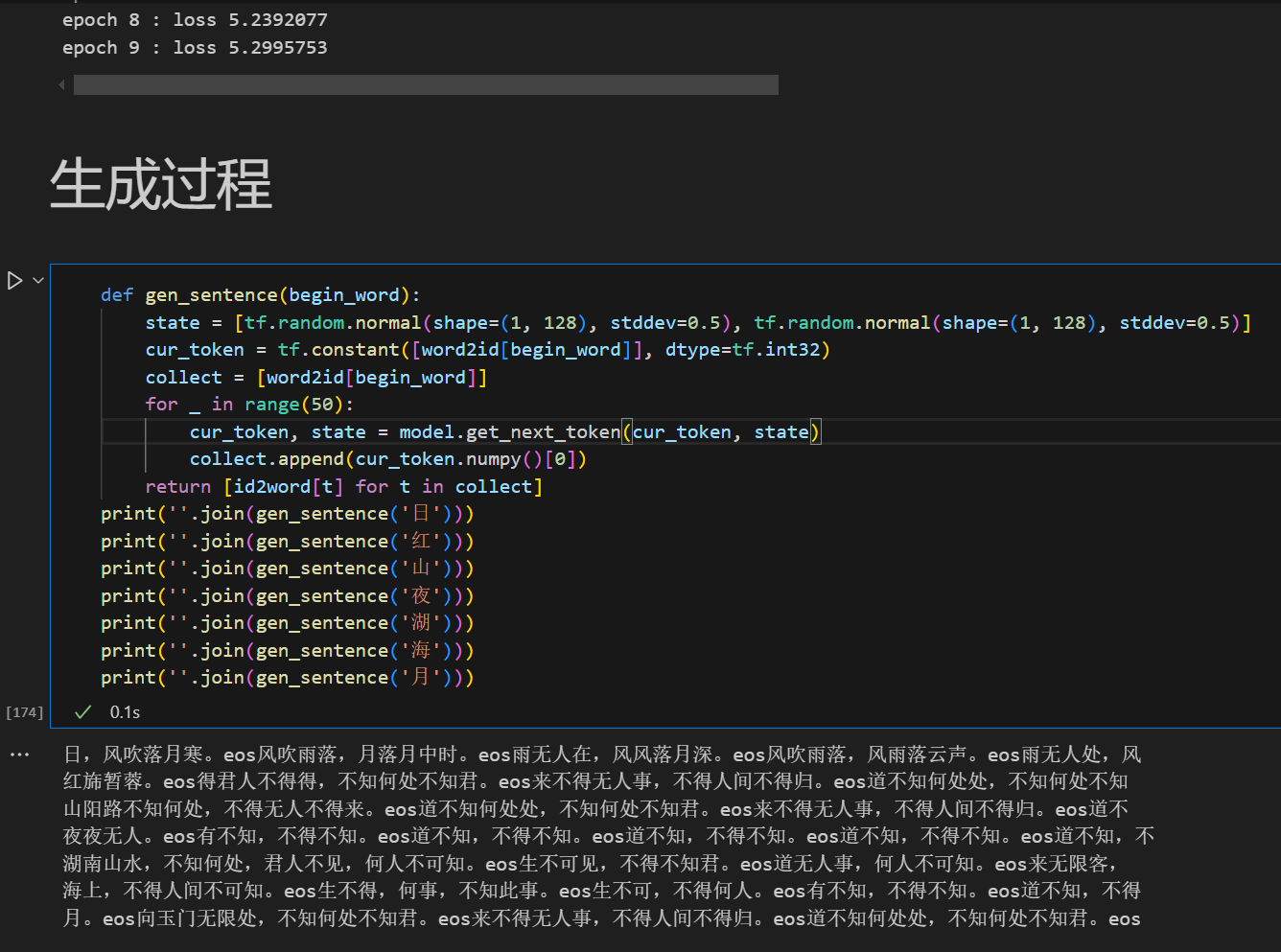
先定义一个状态和只有初始符的token，再定义一个collect用于收集由神经网络生成的字。

然后进行一定次数的循环，每次循环都调用一次训练好的模型，从现有的状态和token生成一个新的token并将其加入collect，同时更新token和状态。生成新token的方法是，将已经给定的token和状态输入到神经网络，得到更新过后的状态和logits，然后将最大logits对应的token作为新的token。

最后，collect中的字就是生成的诗中的字，将它们连接为一个字符串输出即可。

1. 诗歌生成与实验总结

略微修改了生成函数，将原本从起始符开始生成修改为从指定字开始生成，并增加了输出。按照要求，以日、红、山、夜、湖、海、月等字作为begin word，进行生成的截图如下。



上述是用RNN网络进行诗歌生成的例子。从输出可以看出，一旦生成了已经生成过的字就容易出现循环，这可能是因为每次生成字时都是固定采用最大logits值对应的字。另外，从生成的诗中大量出现的位置不合理的终结符和标点符号推测对终结符和标点符号等的处理可能也不够合理。这些地方都有可能可以进行改进。

通过本次实验，我接触到了RNN神经网络，也对它有了一些初步的了解。希望这些知识在以后的学习或者工作中可以用上。