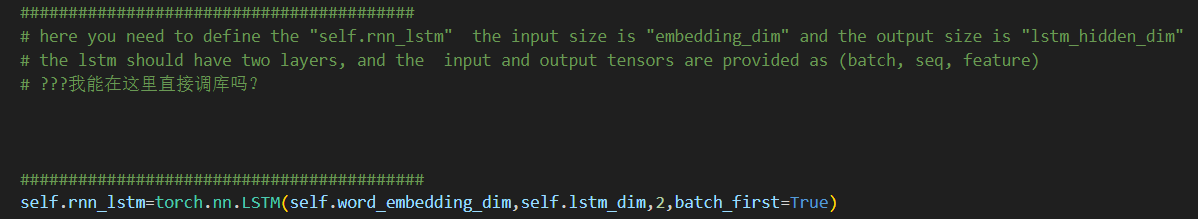
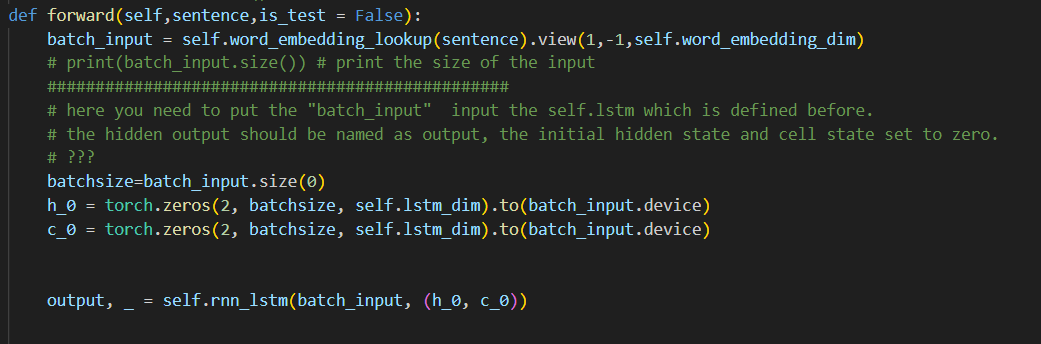
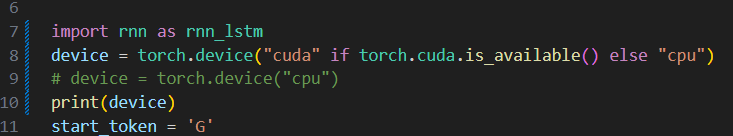
# 使用LSTM实现诗歌生成

一、补全程序



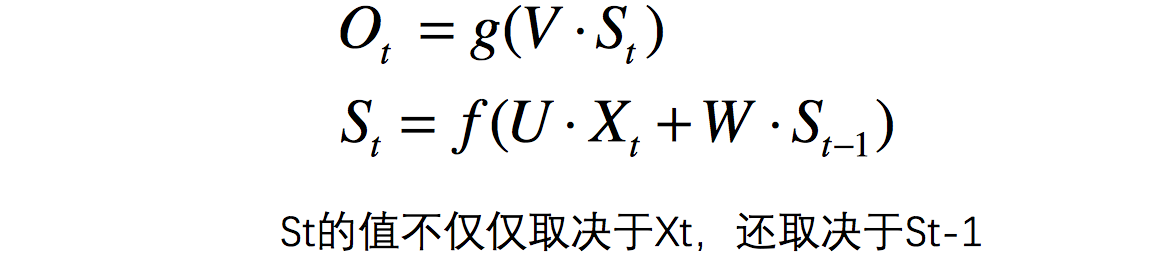


使用gpu训练



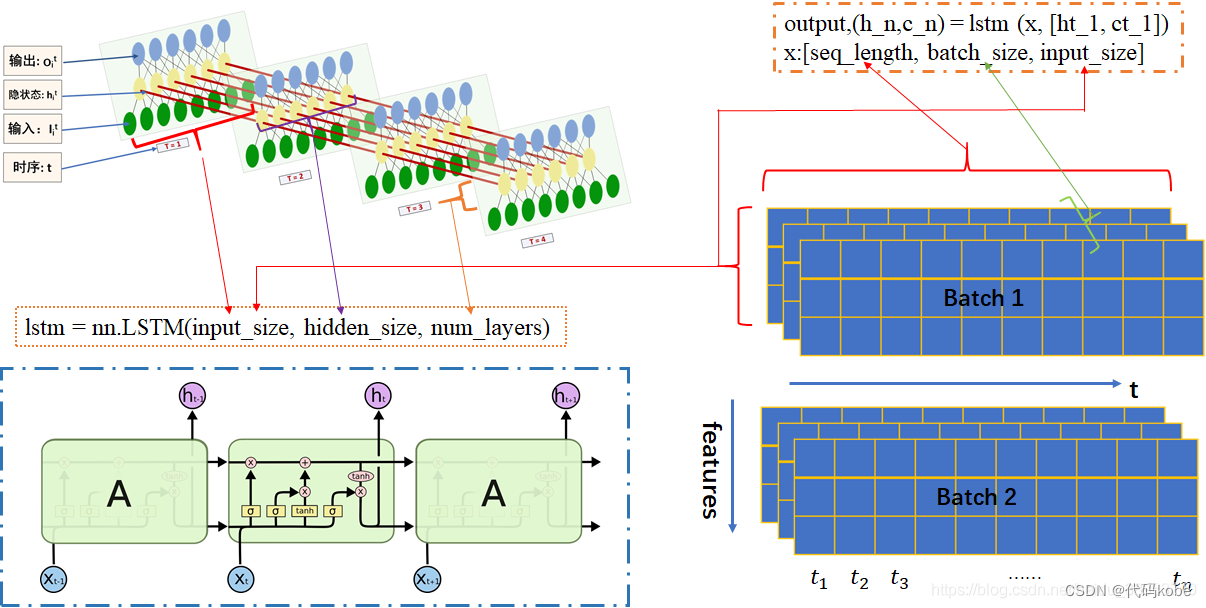
二、模型解释  
循环神经网络（RNN）

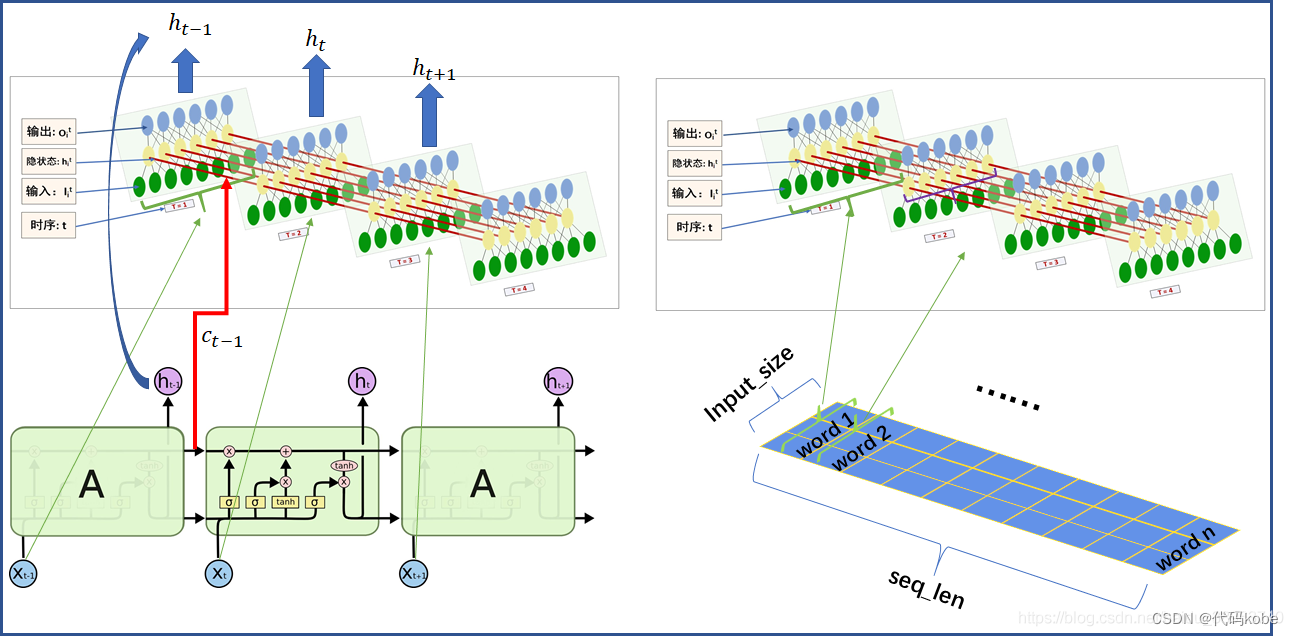
RNN之所以可以解决序列问题，是因为它可以记住每一时刻的信息，每一时刻的隐藏层不仅由该时刻的输入层决定，还由上一时刻的隐藏层决定。我认为理解循环神经网络的关键之一是如下的函数：



长短时记忆（LSTM）

LSTM是一种常用于处理序列数据的深度学习模型，与传统的 RNN（循环神经网络）相比，LSTM引入了三个门（ 输入门、遗忘门、输出门，如下图所示）和一个 细胞状态（cell state），这些机制使得LSTM能够更好地处理序列中的长期依赖关系。下面这两张图很好的说明了LSTM的数据流程和输入输出格式。





门控循环单元（GRU）

是一种循环神经网络（Recurrent Neural Network, RNN）变体，旨在处理序列数据。GRU在LSTM（Long Short-Term Memory，长短期记忆网络）的基础上进行了简化，引入了更少的参数量和结构复杂度。GRU通过使用门控机制有效解决了传统RNN存在的梯度消失和梯度爆炸问题，尤其适合处理长时间依赖的数据。GRU的设计目的是在保持计算效率的同时，拥有较高的性能，适用于广泛的序列处理任务。

三、诗歌生成过程

生成诗歌的过程是一个逐步预测的过程：给定一个起始字（如“日”），将其转换为ID并作为输入传入RNN模型。模型根据输入的字序列预测下一个字的概率分布，并选择概率最高的字作为下一个字。将预测的字添加到已有的诗歌序列中，继续预测下一个字，直到达到结束标记 end\_token 或者生成的诗歌长度超过30。

四、诗歌与训练

