# 总结报告

## 简言

本身任务比较简单，就个人而言，它本身其实对神经网络的知识要求不高，反而有点考察python编程的感觉，毕竟具体的网络更像是套公式，不需要怎么去改变什么。

个人感觉对最后的设计有些失望吧，就是磨拳搽掌了很久等到了这个，有些落空

## 设计思路

总体希望自己的程序好看一点（一直在追求，不过感觉差点意思），所以选择先构造LSTM的一个节点，然后用这个节点去构建整个LSTM。具体而言就是用nn.module去构造，尽量构造标准一点的节点和层，不过具体没有详细参考官方文档，只是参考了参数进行了设计（主要是英语还是太吃力，不到万不得已还是不打算读英语文档）然后照着PPT上的公式按班就部的把LSTMCell写好，然后就是构造出LSTM层，为了方便和torch提供的LSTM类似可以直接选择层数，多层之间节点数统一，这样顺便就可以把提高做了（小小吐槽下，确实不是提高该有的难度，不过考虑很多非计算机的专业，恐怕也是没有办法的事情）

## 分析

结果明显比学长给的直接用torch自带的LSTM跑的效果差很多，速度这个不提，毕竟是专门的框架，自己写的比专门的，哪怕它应该会给我们自己写的提供优化，但多少还是差点意思。主要是结果差点有点多，ppl直接翻了一倍，推测有两种情况，参数写的有问题，不过感觉结果不算太坏，而且继续学的话多少也会有下降，又大概检测了下，感觉应该没问题（不过没有细致检测，不太敢确定，这里偷了点懒），另一种情况是初始化的问题，毕竟是自己写的，初始化方面就会不太好，官方自带的初始化不说是最优，但确实算不错的了。

然后就是一个小问题就是，不知道哪里搞错了，网络没法全部加载到GPU,最后手动在网络内部又加了一个移动网络到目标设备的代码，这样程序封装的就不太好，不太满意（这里偷了懒，没有仔细检查问题）

## 修改内容

具体就是在开头加了两个函数，然后把源代码的LSTM替换成了MyLSTM，还有一些关于GPU的一些调整，很大是因为上面的说的问题，不然直接就分一个文件来写MyLSTM更符合要求一些。网络已经是双层的了，由于本身参数类似torch的LSTM,所以直接改一下num\_layers=2就好。

## 个人感想

ps：实在想不到写什么了，毕竟这确实没有遇到什么磕磕碰碰的，没什么感想，这里随便写点关于这个课的想法，水点字数

最开始看到这课的时候，真的是欣喜若狂，不知道什么时候NLP成为我的方向，网上教程杂乱，个人又是耐不住寂寞，静不下心学，这个课可以说是很好的机会。不过有怕内容高深，自己知识浅薄，学不会，有许多的担心。不知道古人朝圣是什么心态，但课程开始前确实给人一种源于灵魂的颤抖，闻道的气息。不过这学期确实作息乱了，第一节课，很想听，但大脑有些不太支持，勉强接受信息，但不能深入思考，有些遗憾。

实践真的很不错，确实不太喜欢大学来以后代码靠学书，而不是靠实践摸索的模式，时间久了，感觉自己许多能力再退化。不过怎么说呢，理想很丰满，现实多是不尽人意。可能是照顾到许多人薄弱的，又想能够给大家一些锻炼的机会，最后有点乱七八糟的感觉吧（不太会说话的样子）我也不能给出什么建议，但确实这种实践不太适合，跟像是在学编程，而不是神经网络的东西，像是填空，缺了点实践的意思。一个像样的实践确实需要参与其中的大量环节，不然没法很好的理解一些巧妙之处，只觉这样写没错，并不会觉得特别。不过这个难度与工作量确实不太好平衡（只能苦了老师和学长们慢慢研究了）

关于课程来说，由于早课是在太困了，没能很好吸收，不能给出一个好的评价。个人感觉总体能够把握住网络的一些思想（学会一个网络估计还是得靠实践，多写），不过感觉朦胧了些。不过毕竟任务重，但时间少，不过还是希望能够给一些推荐，能有途径更好的去深入了解，毕竟现在网络上资料虽多，但很乱。