**《****自然语言处理》大作业**

**自实现双层LSTM模型**

学 院： 计算机科学与工程

班 级： 人工智能2001

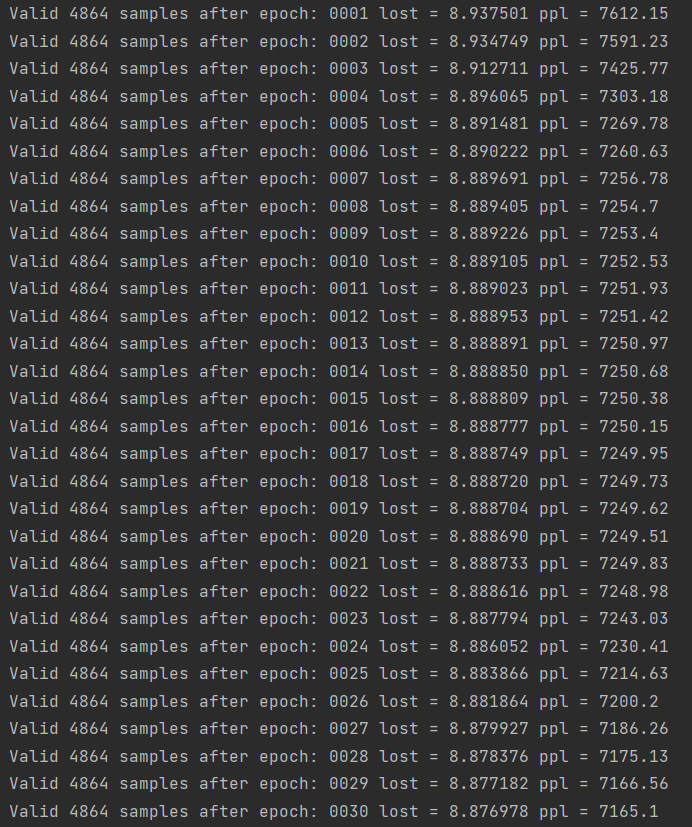
姓 名： 许子强

学 号： 20201111

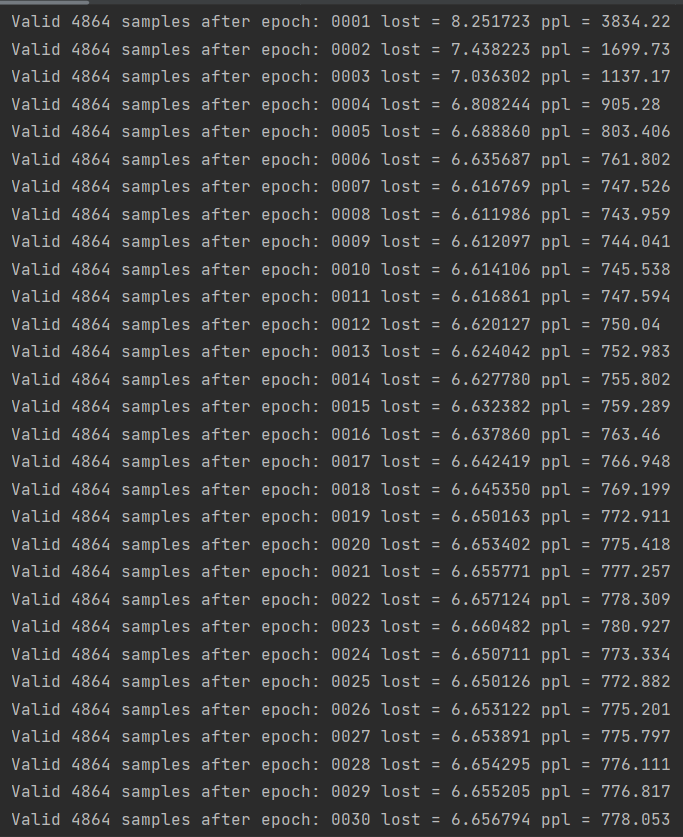
1. 系统设计
2. 实验结果

**2.1 有Softmax的结果（单层LSTM）**

**2.1 有Softmax的结果（双层LSTM）**



**2.2 无Softmax的结果（双层LSTM）**



1. 实验分析
   1. 维度分析

在实践课上听学长说要看懂参数的维度，当时不以为意。然而在我刚开始做实践作业1时，就因为维度问题头疼了很久。在本次大作业期间也出现过因为tensor的维度不匹配而报错，当即明白了——弄明白参数的维度很重要。

* 1. 参数不宜过多

如果每层的参数都较大且层数较多，如多个隐藏层维度设置成n\_class（7613），则训练时显存不够，无法训练。

* 1. 模型输出是否使用Softmax

在训练时发现个有意思的现象——如果在模型的输出层上，先用线性函数拟合输出概率、再用Softmax归一化，那么loss、ppl会非常缓慢地减小；但如果只用线性函数拟合而不用Softmax，那么loss、ppl会先迅速减小、再缓慢增大。很遗憾目前不清楚原因。

1. 总结感悟

为期八周的NLP课程已经结束了，期间有痛苦也有收获。