**5 神经语言模型**

1. 目的
2. 掌握循环神经网络结构，包括LSTM等；
3. 使用Tensorflow构建循环神经网络模型；
4. 要求
5. 建立一个规范的LSTM网络；
6. 在PTB（Penn Treebank）语料库上进行语言模型的训练和评估。
7. Trained model和数据集下载地址。https://pan.baidu.com/s/1sZvLoYFV4ILQh05Y6okMZA
8. 步骤

在训练过程中，可以通过tensorboard观察参数的变化情况

1. 整体框架（main()）

①获取数据，并分为训练集、验证集、测试集

②计算单个epoch中需要训练批次的数量，获得每个epoch迭代的次数

③定义计算图

定义初始化参数，初始化权重和偏置

定义网络结构

定义训练模型

定义验证、测试模型

④在会话session中执行计算图，运行模型

初始化计算图的值

tensorboard的可视化

参数综合

设置保存的路径

将训练集、测试集、验证集转换为可以迭代的小批次数据

迭代训练模型，迭代次数num\_epoch

按批次batch进行训练，并定期检查整个验证集的perplexity

存储Tensorboard的参数

训练完成后，测试测试集的perplexity

保存模型

1. 功能函数的实现

①PTB数据处理

读取数据并返回训练集、验证集、测试集(数组中的元素为每个单词在字典中对应的id)；输入处理好的数据集，迭代生成批次数据

②LSTM模型class PTBModel(object)

定义输入层，输入层维度为batch\_size \* num\_steps ；

定义正确的输出 ；

定义lstm结构 ；

生成单词向量embedding，单词总数为10000，单词向量维度为hidden\_size200；

输出层 ；

交叉熵损失；

梯度下降优化 ；

③在单个epoch中按批次对数据进行迭代run\_epoch()

1. 结果

1.训练参数

一个训练序列的长度 40

单词向量维度 200

初始学习率 1，按照0.9的指数次方下降

迭代次数 20epoch

批次大小 60

Dropout 0.5

梯度裁剪 5

LSTM 3层

2.结果

评价指标：perplexity；越小越好

Validation perplexity：130.834

Training perplexity：133.866

Testing perplexity：125.196

