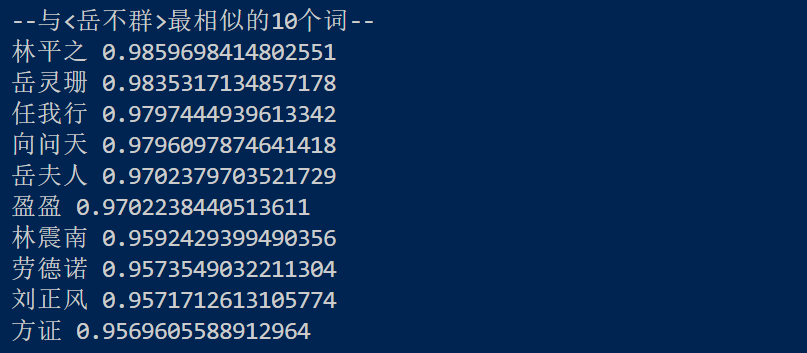
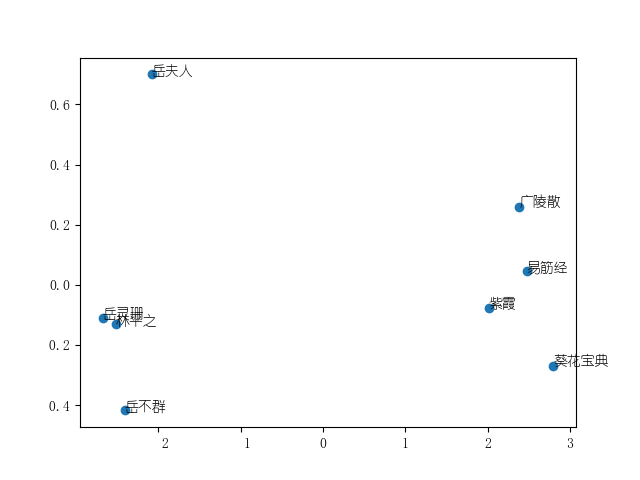
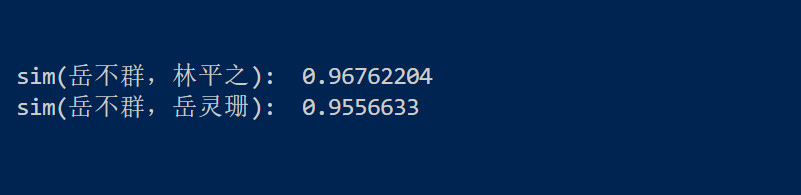
文本表示学习实验

1. word2vec词表示
2. 截图与岳不群最相近的十个词、截图人物与功夫在空间中的相对位置



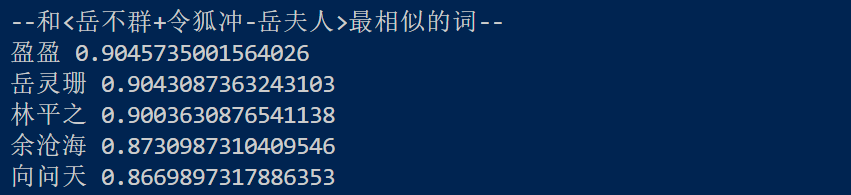


1. 给出重新训练后下面两个词的相似度sim(岳不群，林平之)和 sim(岳不群，岳林珊)

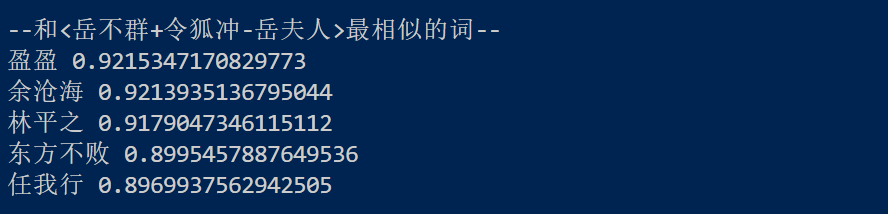


1. 截图给出与岳不群 + 令狐冲 - 岳夫人最相似的5个人物

训练前：

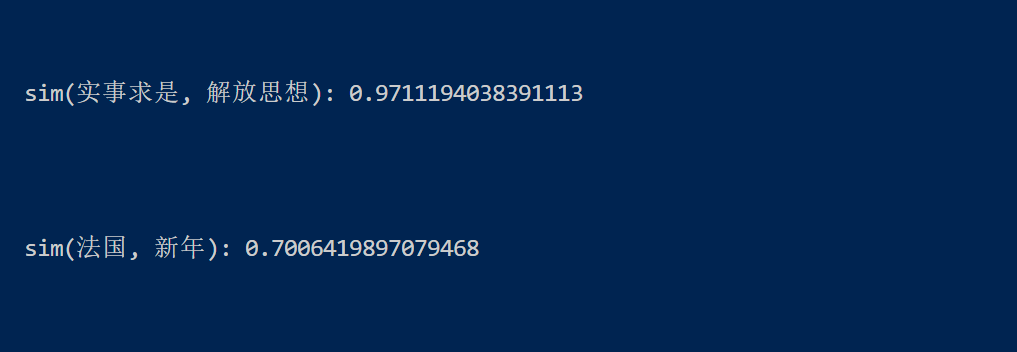


训练后：

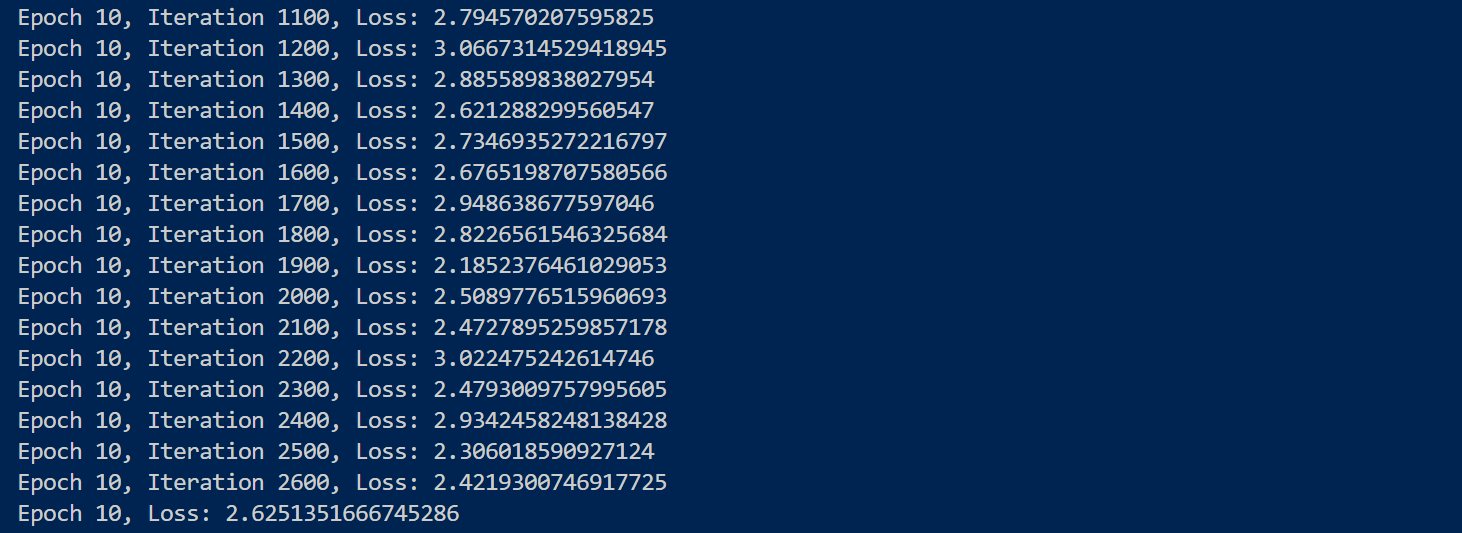


1. 给两对相似词的相似度/两对不相似词的相似度
   1. 如sim(岳不群，令狐冲)=0.95 sim(岳不群，吸星大法)=0.5

这里使用人民日报语料进行训练，结果如下：



1. LSTM搭建与训练
2. 提交rnn.py，其中应该完整包括完整的数据处理与模型代码以及训练代码
3. 截图训练的loss



1. 基于BERT的词表示与句子表示
2. 给出BERT的参数计算过程，如

"hidden\_size": 768,

"intermediate\_size": 3072,

"max\_position\_embeddings": 512,

"num\_attention\_heads": 12,

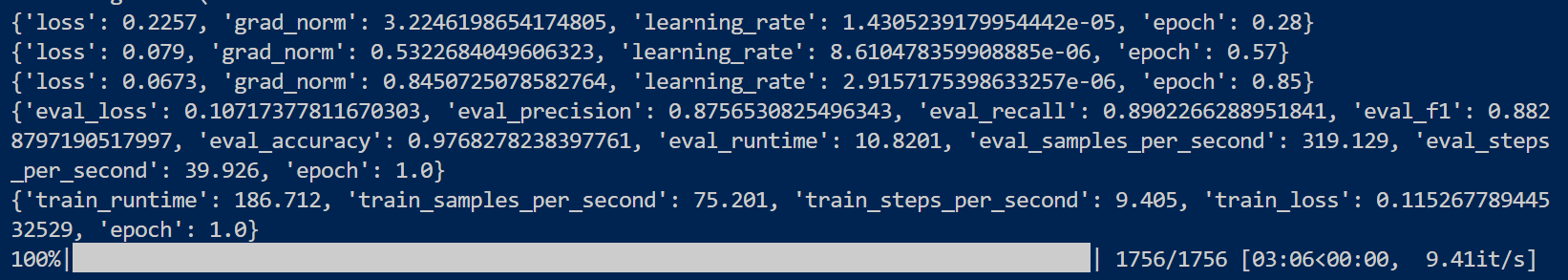
"num\_hidden\_layers": 12,

"position\_embedding\_type": "absolute",

"vocab\_size":

* 1. 词向量为30522\*768，总参数为23,440,896
  2. 位置编码为512\*768，总参数为393,216
  3. 一层transformers参数计算：
     1. 多头自注意力：(768 \* 768 \* 12 / 12 \* 3 + 768 \* 768) = 2,359,296
     2. 前馈神经网络：768 \* 3072 \* 2= 4,718,592
     3. 层归一化：hidden\_size\*2\*2=768\*2\*2=3072；
  4. 总参数为：23,440,896+393,216+12\*（2,359,296+4,718,592+3072）=108,805,632

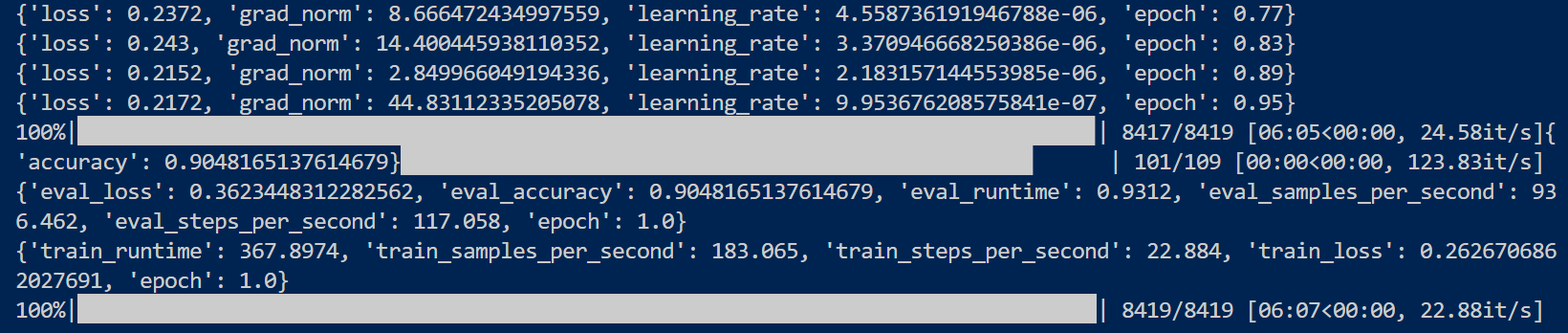
1. 截图训练结束之后的评估结果



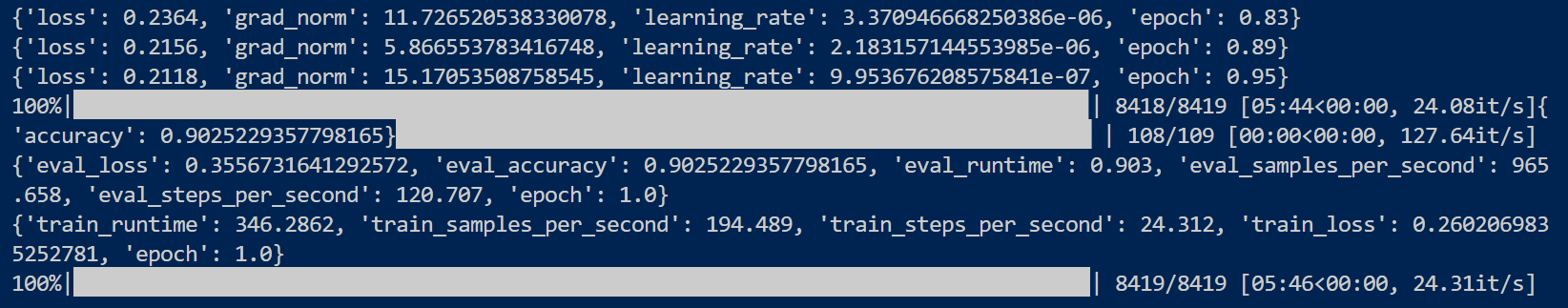
1. 提交train\_cls.py，其中应该包括完整的训练代码。提交BERT.py，其中包括完整的mean pooling代码与max pooling代码，仅提交一个BERT.py文件，其中不同的pooling代码注释掉。补充下面的不同方法的性能的表格，其中使用CLS的方法accuracy需要大于0.9：

|  |  |
| --- | --- |
| 方法 | Accuracy |
| CLS | 0.9048165137614679 |
| Mean pooling | 0.9025229357798165 |
| Max pooling | 0.8956422018348624 |

CLS ：



Mean pooling：



Max pooling:

