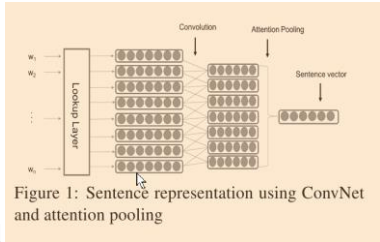
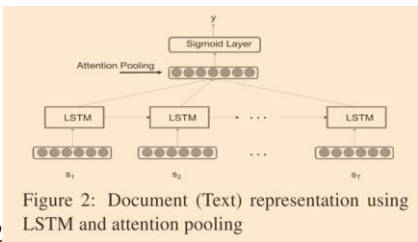


第_十_周周记

周一	
完成内容	看关于罚金预测的代码，查找资料学习相关代码以及资料。
内容描述	1、看其中的两个程序：predict.py 和 train.py。 2、运行 predict.py，并 debug predict.py，了解该程序各部分的运行顺序。 发现预测类别有八类。 3、Debug read_test.py，发现该程序是对词频的统计
未解决问题	还有部分代码不理解

周二	
完成内容	学习罚金预测的代码。
内容描述	1、python json http://www.runoob.com/python/python-json.html 2、h5py http://docs.h5py.org/en/latest/quick.html 3、numpy.argmax() http://blog.csdn.net/u013713117/article/details/53965572 4、enumerate()使用 http://blog.csdn.net/churximi/article/details/51648388 5、python 数据持久存储：pickle 模块的基本使用 https://www.cnblogs.com/pzxbc/archive/2012/03/18/2404715.html 6、cut_sentence.py 切分句子，以 “。！？” 为分隔符进行分割句子，最终得到以单个句子为一个单位的词典。 7、Build_vocab.py 是建立词典，统计出现的词的频率，将频率小于 5 的词去除，可减小程序的计算量。 8、Read_test.py 是统计词频
未解决问题	无

周三	
完成内容	看论文：徐芳.罚金预测；徐芳. 作文自动评分综述_0104；
内容描述	1、了解罚金预测程序整体流程，用 CRNN+Attention+RF 来分类。罚金类别是 8 种，文本预处理提取关键词，提取文本的浅层特征，用 one-hot 文本分布式表示。 2、了解基于分类的作文自动评分方法，朴素贝叶斯分类（便是为当前文本寻找最合适的类别，也就是最大化作文属于某一类别的对数概率 $\log P(c Doc)$ ）；K-近邻分类（挑出与待测试样本最相似的 k 个样本，若该 k 个样本的大部分属于某个类别，则该测试样本也属于该类别。）；基于神经网络的分类是罚金预测的分类方法（对文档中的句子进行建模，然后进行基于句子级别的文本分类）。
未解决问题	无

周四	
完成内容	看论文: Dong F, Zhang Y, Yang J. Attention-based Recurrent Convolutional Neural Network for Automatic Essay Scoring[C]// Conference on Computational Natural Language Learning. 2017:153-162.
内容描述	<p>1、Convolutional layer 卷积层</p> <p>2、Deep Learning 最优化方法之 RMSProp http://blog.csdn.net/bvl10101111/article/details/72616378 http://blog.csdn.net/u010089444/article/details/76725843</p> <p>3、卷积神经网络中图像池化操作全解析 http://blog.csdn.net/u010402786/article/details/51541465</p> <p>4、评价标准: QWK</p> <p>5、模型</p> <div style="display: flex; justify-content: space-around;">   </div> <p>5.1 ; 5.2</p> <p>6、使用 MSE (mean square error) 均方误差</p>
未解决问题	无

周五	
完成内容	看罚金预测程序
内容描述	<p>1、在对该程序有更深入的了解之后,发现之前改的程序想法有误。训练数据是已经将刑期提取出来的文本,而预测的文本则应该式没有刑期提取的原始文本,然后开始预测。原先罚金种类有 8 种,所以应将那段程序给为相应的刑期类别。</p> <p>2、刑期类别大的分类应该分为五类,即拘役、管制、有期徒刑、无期徒刑和死刑。其中有有期徒刑可以分为 7 类即: 1 年、2 年、3 年、5 年、7 年、10 年、15 年。</p>
未解决问题	无

周末	
完成内容	看 Automatic Features for Essay Scoring – An Empirical Study - EMNLP 2016
内容描述	无
未解决问题	

工程汇总	
任务描述	无
代码量	

未解决问题	无
-------	---

论文汇总	
论文列表	<p>[1] Dong F, Zhang Y. Automatic Features for Essay Scoring – An Empirical Study[C]// Conference on Empirical Methods in Natural Language Processing. 2016.</p> <p>[2]徐芳.罚金预测</p> <p>[3]徐芳. 作文自动评分综述_0104</p> <p>[4] Dong F, Zhang Y, Yang J. Attention-based Recurrent Convolutional Neural Network for Automatic Essay Scoring[C]// Conference on Computational Natural Language Learning. 2017:153-162.</p>
论文摘要	
未解决问题	无

下周任务	
工作	看相关的代码，准备语料，实现模型
论文	
其他	
汇总	

日期:18/3/4 - 18/3/10