

## 第 9 周周记

周一	
完成内容	
内容描述	
未解决问题	

周二	
完成内容	阅读论文《音乐领域典型事件抽取技术的研究》宋凡
内容描述	基于 pattern 的信息抽取，主要的工作原理简单说就是从已标注语料或者 Web 资源上通过统计的方法，首先获得对某种事件常见表述模式的一个模式库，然后利用这些已有的模式资源来和待识别的文本进行匹配，如果匹配成功就利用模式库中相关的槽信息，进行相关元素（一般为实体）的抽取工作
未解决问题	

周三	
完成内容	阅读论文《基于 LDA 模型和 AP 聚类的主题事件抽取技术》
内容描述	采用 LDA 主题模型实现文本主题层次信息抽取，主题提取时采用 k-means 算法进行聚类。 采用 ap 聚类算法（仿射传播 Affinity Propagation）对主题下文本信息统计提取。
未解决问题	在法律文档中，主题事件可能不好提取，不过分层提取可以考虑

周四	
完成内容	阅读论文《基于结构化学习的事件抽取研究》
内容描述	基于卷积神经网络的文本分类：卷积神经网络的工作方式是，首先选用固定窗口大小的卷积滤波器在原始句子上面滑动，然后得到对应于该卷积滤波器的特征图，接着对特征图采用最大池化的方式，得到对应特征值。 跳窗卷积神经网络：利用卷积神经网络的优势，主要针对传统卷积神经网络的输入层进行特殊的修改，并且引入位置特征向量，最终能够很好的得到词级别的全局结构信息
未解决问题	

周五	
完成内容	阅读论文《基于结构化学习的事件抽取研究》
内容描述	结构感知器可以看作是感知器模型的一个变体，采用的是维特比译码的方式来训练模型的，其本质上就是一个利用动态规划的方式求解最优路径的过程。

	Joint 事件抽取模型（联结型的事件抽取模型）把事件抽取任务看成是一个结构化学习的任务，采用了结构感知器来训练事件抽取模型。 基于 RNNs（循环神经网络）的结构化学习模型先通过其他的模型把每个标签训练成向量，采用同相对位置特征向量一样的方式，再计算最终得分。
未解决问题	

周末	
完成内容	整理标注语料
内容描述	对已标注的语料进行整理，准备下一步工作。
未解决问题	

工程汇总	
完成任务	阅读论文，了解结构化学习的相关知识。
任务描述	
代码量	
未解决问题	

论文汇总	
论文列表	[1]张建恒,黄蔚,胡国超.基于 LDA 模型和 AP 聚类的主题事件抽取技术[J/OL].计算机与现代化,2017(12):77-81+87[2018-03-03]. <a href="http://kns.cnki.net/kcms/detail/36.1137.TP.20180123.1707.030.html">http://kns.cnki.net/kcms/detail/36.1137.TP.20180123.1707.030.html</a> . [2]张正宽. 基于结构化学习的事件抽取研究[D].北京邮电大学,2017.
论文摘要	
未解决问题	

下周任务	
工作	标注语料，阅读论文，完成开题报告。
论文	
其他	
汇总	

日期:2018/02/25-2018/03/03