

## 第 2 周周记

周一	
完成内容	NER 命名实体识别问题相关
内容描述	查找 NER 相关文章, 阅读《用深度神经网络处理 NER 命名实体识别问题》, 地址: <a href="https://www.jianshu.com/p/581832f2c458">https://www.jianshu.com/p/581832f2c458</a>
未解决问题	具体到课题相关的话, NER 是否只需要套用已完成的代码实现。

周二	
完成内容	阅读《中文事件抽取技术研究》
内容描述	在事件类别识别阶段, 本文采用了一种基于触发词扩展和二元分类相结合的方法;在事件元素识别阶段, 本文采用了基于最大熵的多元分类的方法。对事件类别识别和事件元素识别进行了研究。
未解决问题	

周三	
完成内容	阅读《音乐领域典型事件抽取方法研究》
内容描述	通过领域事件词聚类的方法自动发现音乐领域具有代表性的事件, 然后采用基于关键词与触发词相结合的过滤方法简化了事件类型的识别过程。在事件元素识别中, 该文采用了基于最大熵的事件元素识别方法。
未解决问题	

周四	
完成内容	DVE 算法和 ETDA 算法的了解
内容描述	领域事件词抽取算法 DVE (Domain Verb Extraction), 事件类型发现算法 ETDA(The Event Type Discovery Algorithm), 可能可以在法律案例事件抽取中使用, 先记录下来。
未解决问题	

周五	
完成内容	了解 ACE2005
内容描述	阅读 ACE ( Automatic Content Extraction ) Chinese Annotation Guidelines for Entities, 来源 <a href="http://www.doc88.com/p-180734447445.html">http://www.doc88.com/p-180734447445.html</a>
未解决问题	全文有一定的长度, 加上全英文要适应一下语法, 未全部读完。

周末	
完成内容	了解 ACE2005

内容描述	继续阅读 ACE ( Automatic Content Extraction ) Chinese Annotation Guidelines for Entities
未解决问题	通篇读完，大概了解内容，但未细读，准备在应用时继续研究。

工程汇总	
完成任务	阅读相关论文，继续了解事件抽取的相关内容。
任务描述	基于最大熵的事件元素识别方法有一定可行性，ace2005 的大致阅读。
代码量	暂无
未解决问题	

论文汇总	
论文列表	[1] 赵妍妍, 秦兵, 车万翔, 刘挺. 中文事件抽取技术研究 [2] 丁效, 宋凡, 秦兵, 刘挺. 音乐领域典型事件抽取方法研究
论文摘要	[1] 事件抽取是信息抽取领域一个重要的研究方向, 本文对事件抽取的两项关键技术——事件类别识别以及事件元素识别进行了深入研究。在事件类别识别阶段, 本文采用了一种基于触发词扩展和二元分类相结合的方法;在事件元素识别阶段, 本文采用了基于最大熵的多元分类的方法。这些方法很好的解决了事件抽取中训练实例正反例不平衡以及数据稀疏问题, 取得了较好的系统性能。 [2] 事件抽取是信息抽取领域一个重要的研究方向。该文从音乐领域的事件抽取出发, 通过领域事件词聚类的方法自动发现音乐领域具有代表性的事件, 然后采用基于关键词与触发词相结合的过滤方法简化了事件类型的识别过程。在事件元素识别中, 该文采用了基于最大熵的事件元素识别方法。在该文构建的语料库下, 最终事件类型识别的平均 F 值达到 82.82%, 事件元素识别的平均 F 值达到 75.79%。
未解决问题	

下周任务	
工作	继续阅读相关资料，进行事件抽取的准备工作
论文	
其他	
汇总	

日期:2018/1/7-2018/1/13