

第七周周记

周一	
完成内容	阅读 2017 年浙江大学谢志宁的硕士论文《中文命名实体识别算法研究》
内容描述	1. 了解基于单一模型的中文命名实体识别 2. 了解基于混合模型的中文命名实体识别 3. 了解引入字词向量与样本生成的中文命名实体识别
未解决问题	无

周二	
完成内容	阅读 2017 年浙江大学隋臣的硕士论文《基于深度学习的中文命名实体识别研究》
内容描述	1. 了解循环神经网络和长短期记忆模型 2. 了解隐马尔可夫模型、条件随机场、支持向量机三种模型的比较 3. 了解神经网络的命名实体识别模型实现
未解决问题	无

周三	
完成内容	阅读 2017 年北京化工大学徐梓豪的硕士论文《基于统计模型的中文命名实体识别方法研究及应用》
内容描述	1. 了解隐马尔可夫模型的命名实体识别过程 2. 了解基于规则和词典的病历关键信息抽取
未解决问题	无

周四	
完成内容	观看视频《神经网络：梯度下降 (Gradient Descent in Neural Nets)》
内容描述	了解误差方程的基本知识
未解决问题	无

周五	
完成内容	阅读论文《基于条件随机场模型的中文分词系统研究与实现》
内容描述	1. 了解中文分词的几种主要算法的知识，如基于匹配的分词方法 2. 了解最大熵模型的约束条件和原则 3. 了解基于条件随机场模型的中文分词系统研究与实现
未解决问题	无

周末	
完成内容	观看莫烦 python 视频《Linux 简易教学》
内容描述	学习了一些 Linux 基本指令和文件权限等知识。
未解决问题	无

工程汇总	
完成任务	如上所述

任务描述	如上所述
代码量	无
未解决问题	无

论文汇总	
论文列表	[1] 《中文命名实体识别算法研究》 [2] 《基于深度学习的中文命名实体识别研究》 [3] 《基于统计模型的中文命名实体识别方法研究及应用》 [4] 《基于条件随机场模型的中文分词系统研究与实现》
论文摘要	<p>命名实体识别(Name Entity Recognition, NER)是指识别文本中具有特定意义的实体，主要包括人名、地名、组织机构名等，是将非结构化数据转为结构化数据的一个重要技术手段，是计算机正确理解文本信息的关键步骤，也是信息抽取、情感分析、问答系统等多个自然语言处理应用的基础任务，因此命名实体识别的研究存在着重要意义。但由于中文语言自身的特点，中文命名实体仍存在许多难点，其主要难点包括：</p> <p>(1)中文命名实体识别通常是基于单一模型的识别，这些模型具有各自的优缺点和局限性。</p> <p>(2)中文命名实体识别通常是基于词序列的识别，需要借助中文分词技术，中文命名实体识别的效果往往依赖于中文分词的准确率。</p>
未解决问题	无

下周任务	
工作	继续查找相关的论文和继续阅读相关文献。
论文	继续寻找与中文分词和命名实体识别相关的论文
其他	无
汇总	了解更多与我的课题相关的知识

日期:2018/02/12 -

2018/02/18