# ZUTNLP-Text 第9周周报

2018-2019-2 学期， 2018/11/05 – 2018/11/10

|  |  |  |
| --- | --- | --- |
|  |  | |
| 指导讨论记录： | 2018-11-05 13:00- 13:30 刘小明   1. 评审上周工作完成情况； 2. 布置本周工作内容；   2018-11-05 刘凤华 | |
| 上周工作总结 | 团队工作总结 | |
| 1. 对词性标注、地名实体识别的数据进行了训练和测试  2. 运用维特比算法。  3. 学习学长的项目，能够提取关键字和给出权重  4. 优化前端数据存储问题，给词性标注添加了UI云标签  5. 学习Python相关的知识 | |
| 团队成员工作总结： | |
| 曾辉祥 | ①对词性标注的数据进行了训练和测试  ②对地名实体识别的数据进行训练、测试，对地名能够识别  ③优化前端数据存储问题，给词性标注添加了UI云标签 |
| 卢思童 | ①对于gitlab，操作不是很熟练，犯了一堆错，在多次尝试之后，大概了解了流程。  ②掌握并运用维特比算法。  ③学习Python相关的知识。 |
| 赵向阳 | ①网上搜索学习摘要关键词的方法，学习摘要关键词的操作流程，跑了一些语料。  ②按照学长的给的例子学习摘要关键词，学习别人的方法和代码。  ③尝试用自己小组的中文分词基础上进行关键词的摘要。  ④学习python |
| 陈金鑫 | ①训练了crf的model，对项目进行版本控制，能够利用crf运行分词  ②对crf的使用进行的总结  ③调用ltp实现语义决策分析 |
| 支友盟 |  |
| 房晔 | ①下载配置Jekins。  ②配置学习使用Gitlab。  ③学习使用Crf++。 |
| 张欣欣 |  |
| 本周存在问题 | 团队共性问题 | |
| 1. 对文本的分析还不是很智能  2. 词云展示方法  3. 对gitlab、github的上传时对一些文件上传不了 | |
| 团队成员问题 | |
| 曾辉祥 | ①对词云的展示方法还不是太明白  ②对数据的测试还不是很完善，比如对地名实体的识别就有一些提取不了 |
| 卢思童 |  |
| 赵向阳 | ①Python语言还是的还是不够，只了解到了一些python基础，许多语法结构都不会。  ②看不懂TextRank算法的原理和使用方法。 |
| 陈金鑫 | ①对项目进行github版本控制，建立版本库，对项目进行上传，有些文件扫描不到，没有上传上去，还有些由于文件太大上传不了，需要下载Git Lfs插件  ②Crf跑的时候合并说是超过了List的range  ③下载了ltp，在网上大概找了几种种方法1.安装Python的扩展包pylyp，Pip报错，Vcexit status 2。安装whl文件，又提示python版本不匹配3.直接调用编译好的ltp可执行文件，又出现乱码 |
| 支友盟 |  |
| 房晔 |  |
| 张欣欣 |  |
| 本周工作计划 | 团队总体计划 | |
| 1. 训练和优化分词、词性标注、地名实体识别model  2. 训练出命名实体识别model，能够跑自己的数据  3. 训练出依存句法分析model，并完成数据的测试  4. 训练出语义决策model，并完成数据的测试  5. 对比之前的nlp方法并在前端展示 | |
| 成员个人计划 | |
| 曾辉祥 | ①不断优化model，对template有一个更深的认识  ②掌握文本词云的展示方法  ③将crf在java中实现，和之前的结果进行对比 |
| 卢思童 | ①训练出依存句法分析model，并完成数据测试。  ②深度学习HMM模型+维特比算法。 |
| 赵向阳 | ①训练出命名实体识别模块  ②模块能跑自己的数据 |
| 支友盟 |
| 陈金鑫 | ①掌握ltp中srl的预料和model  ②训练出自己的语义决策model，并完成数据的测试 |
| 房晔 | ①完成crf++的具体使用。  ②了解学习spring boot，spring boot如何整合vue项目。 |
| 张欣欣 | ①跟着其他组员学习crf  ②学习java、python基础 |