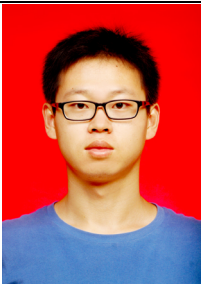


李忠阳

哈尔滨工业大学计算机科学与技术



手机：17766589134 邮箱：zyli@ir.hit.edu.cn 1992年生，预计2020博士毕业
2011/08 - 2015/07 : 哈尔滨工业大学/计算机学院/本科
2015/09 - 2016/07 : 哈尔滨工业大学/计算机学院/硕士
2016/09 - 至今 : 哈尔滨工业大学/计算机学院/博士在读
所在实验室：社会计算与信息检索研究中心SCIR，社会预测组
导师：刘挺教授

项目经历

- 1. 百度知道问答对数据过滤项目 (python, c++, shell)** : 在百度实习期间，所在团队为自然语言处理部深度问答团队。我负责百度知道以及相关网站问答对数据垃圾过滤的工作，同时负责处理一些数据，为小组其他同学提供支持。期间推出自动问答产品“小度机器人”。
- 2. SCIR-电影票房预测系统 (python)** : 对该系统原理比较熟悉，负责日常的维护与调试。
- 3. 北京语言大学词语搭配项目 (python)** : 负责从时间跨度10年的海量新闻语料中挖掘词语搭配信息，主要解决了查询时间复杂度过大的问题，针对任意时间段、任意查询词即时返回查询结果。

专业技能

语言 : python, c/c++, shell, 具有熟练的编程能力，能快速上手一种新语言

算法 : 大一大二2年的ACM比赛经历，熟悉常见算法和数据结构，在各大OJ刷题500余道

数据处理 : 能使用hadoop系统进行大数据处理任务，能熟练使用python对文本进行处理

机器学习 : 熟悉SVM等常见机器学习算法，能快速应用RNN等常见深度学习模型，了解增强学习

研究方向 : 文本数据挖掘，信息抽取，知识库构建

研究工作

- 1. 基于生成对抗网络的故事结尾推理生成** : 生成器为Seq2Seq模型，判别器是一个二元句子分类器，利用强化学习将分类器输出作为reward传给生成器。输入为4句话的故事上下文，输出为故事结尾。投IJCAI被拒（得分4484，主要质疑为实验评估），转投EMNLP（得分333，主要质疑为创新性）。
- 2. 从大规模无结构化文本数据构建事理图谱** : 事理图谱是我的博士课题，目标是从大规模无结构化数据中抽取事件之间的逻辑演化关系，形成有向图结构的事理知识库。目前在出行领域数据上进行了实验，取得了不错的效果。未来考虑挖掘金融新闻数据，构建股市涨跌的事理图谱。

所获奖励

2013年ACM黑龙江省赛一等奖，区域赛三等奖

2012年国家励志奖学金

2013年ACM东北地区四省赛一等奖

2013年优秀班干部

2015年入学国家二等奖学金

社会实践

2014年6月-9月在百度自然语言处理部实习

2016年3月至2016年5月计算机网络助教

2015年6月-10月在北京HTC实习