

谭聪

求职意向：自然语言处理算法工程师

教育背景

- 2013--至今 硕士(计算机科学与技术), 哈尔滨工业大学深圳研究生院, GPA: 3.516 专业前 15%(15/118).
导师: 王晓龙教授
研究方向: 问答系统, 信息检索, 机器学习, 自然语言处理
毕业时间: 2015 年 12 月
- 2009--2013 学士(计算机科学与技术), 杭州电子科技大学, 成绩: 专业前 15%.

研究项目

- NTCIR-11 Temporalia 时间信息检索评测: 在搜索引擎的搜索过程中, 考虑查询的时间属性和结果文档中的时间信息, 使得搜索结果既内容相关又符合时间要求。我们的评测结果获得国际第二, 并发表论文一篇 [3]。
主要工作: 使用 Solr 对文档建立索引和检索, 抽取问句和搜索结果的特征, 包括文本特征和时间特征, 使用 Learning to Rank 算法对搜索结果进行重排序。
- SemEval-2015 task3 社区问答中答案选择任务 (Answer Selection in Community Question Answering), 是指在社区问答中, 对每个问题的答案质量进行判定。该任务包含两种语言: Arabic 和 English。在 Arabic 子任务中, 我们组 (HITSZ-ICRC) 取得了第三名, 在 English 子任务中, 获得了第二名 [1]。
主要工作: 抽取文本特征, 比如 BOW、tf-idf、共现词等, 也尝试过结合 word2vec 词向量特征, 使用集成方法进行分类。
- SemEval-2015 task5 QA TempEval 是指以问答的方式来评价对文本中时间信息的理解程度。主要涉及三个方面: 1. 文本中时间的抽取和归一化, 2. 事件的抽取, 3. 事件和时间、事件与事件之间关系的抽取 [2]。
- 开源 Github 为 Stanford CoreNLP, timeml-qa, caevo 等项目提交过代码, 主要是修复一些 bug 等。

专业技能

- 🔧 熟悉 C++、Java、Python 等编程语言;
- 🔧 熟悉主要的数据结构和算法。本科期间获得浙江省第八届程序设计竞赛铜牌。
- 🔧 了解常用机器学习算法、自然语言处理等相关技术; 熟练使用机器学习工具: libsvm, scikit-learn, weka 等。
- 🔧 熟悉 Linux 环境, 掌握 Linux 下常用命令及开发调试方法, 熟悉 git 等版本管理工具

所获奖励和证书

- 2014-10 哈尔滨工业大学深圳研究生院一等奖学金
2014-10 哈尔滨工业大学深圳研究生院三好学生
2013-09 哈尔滨工业大学深圳研究生院特等奖学金

校园实践

- 2014.12.5-9 NLPCC 2014 和 ADL 52 期志愿者。
- 2014-01-2014.12 深圳大学城微软技术俱乐部技术部部长, 在担任技术部部长期间, 组织了多次俱乐部活动, 还获邀参加了 2014 年微软夏令营。
- 2013-09-2014.09 哈工大深圳研究生院研究生会学术部成员, 组织过多项校内活动, 比如辩论赛、科技展、主持人大赛等, 并主持了主持人大赛。

论文

- [1] Yongshuai Hou, Cong Tan, Qingcai Chen, and Xiaolong Wang. Hitsz-icrc: An integration approach for qa tempeval challenge. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 398--402, Denver, Colorado, June 2015. Association for Computational Linguistics.
- [2] Yongshuai Hou, Cong Tan, Xiaolong Wang, Yaoyun Zhang, Jun Xu, and Qingcai Chen. Hitsz-icrc: Exploiting classification approach for answer selection in community question answering. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pages 403--409, Denver, Colorado, June 2015. Association for Computational Linguistics.
- [3] Yongshuai Hou, Cong Tan, Jun Xu, Youcheng Pan, Qingcai Chen, and Xiaolong Wang. Hitsz-icrc at ntcir-11 temporalia task. In *Proceedings of the 11th NTCIR Conference on Evaluation of Information Access Technologies*, pages 468--473, 2014.
- [4] Yongshuai Hou, Xiaolong Wang, Qingcai Chen, Man Li, and Cong Tan. User input classification for chinese question answering system. In *Machine Learning and Cybernetics*, volume 481, pages 52--59. Springer Berlin Heidelberg, 2014.