

2013.12 - 2015.12

360 - 商业产品事业部 - 数据挖掘组

算法工程师

核心开发

◆ 联盟广告项目高商词触发与排序

- ✧ 高商词触发: 负责 content targeting 和 query targeting 策略。
- ✧ 负责高商词排序: 先在流量、地域、时间三维度上研发带回退策略的一跳统计 CTR 和二跳统计 CPM, 后融合一跳 CTR 和二跳 CPM 得到高商词的 rank_score。
- ✧ 策略效果: 大站的 CTR1 提升超过 180%, CPM1 提升超过 120%。
- ✧ 项目副产品: 基于 MapReduce 的大规模稀疏矩阵相乘。

◆ 展示广告项目广告触发

独立负责

- ✧ 目的: 旧版广告引擎直接建立从 tag 到 ad 的倒排索引, 该方案存在倒排索引平均长度长, 在检索时需要硬剪枝, 造成 CTR 等投放效果不好。为此建立从 conjunction 到 ad 的倒排索引。
- ✧ 具体策略: conjunction 包括流量维度、定向维度(包括兴趣定向、host 重定向、关键词重定向)、广告粒度维度(包括广告、广告组、广告主、广告类目)。针对<ad, conjunction> pair 先在这三维度上研发带回退策略的统计 CTR 模块, 然后结合 bid price 计算<ad, conjunction> pair 的 rank_score。
- ✧ 策略效果: CTR1 提升 30%, CPM1 提升 10%。

2012.7 - 2013.11

人民搜索 - 舆情部 - 社交网络挖掘小组

算法工程师

- ◆ 为监控微博舆情, 独立负责了地域识别、情感正负面分析、识别博主所属行业等策略模块以及流式处理微博数据等工程模块; 并负责设计 social business 产品, 其功能包含社会化媒体管理、社交网络监听、社交账号分析和潜在客户挖掘等功能。

2011.6 - 2011.8

Hulu 北京 - 视频检索组

研发工程师(实习)

- ◆ 优化 query understanding 模块和拼写错误纠正模块。错误纠正的准确率从 75% 提高至 78%。

科研经历

2011.9 - 2012.5 词语对齐快速增量式训练方法(863 项目“多语言互联网语言翻译关键技术”子课题) 独立完成

- ◆ 词语对齐是翻译模型构建流程的瓶颈, 对基于无监督学习的词语对齐问题, 当有新语料时, 应用迭代训练收敛速度更快的 online EM 算法, 以替换通常所用的 batch EM 算法的方法, 实现增量式训练。
- ◆ 实验效果: 模型训练速度提高 5 倍, 词语对齐和机器翻译的质量与 batch EM 方式的效果持平。

2010.5 - 2010.9

2010 年国际口语机器翻译评测 (IWSLT 2010)

核心成员

- ◆ 研发汉英、英汉双方向基于规则的命名实体翻译系统; 优化基于层次短语模型的统计机器翻译解码器。
- ◆ 评测成绩: 总共 8 项评价指标中 6 项第一和 2 项第二(全球 11 支科研队伍参赛)。

论文 & 专利

- ◆ 在 360 工作期间, 以第一作者身份提交 16 篇专利申请, 其中法律状态为“有权-授权”的有 1 篇, 状态为“审中-实审”的有 3 篇, 状态为“审中-公开”的有 5 篇。(截止 2018 年 1 月)
- ◆ 罗维. 词语对齐的快速增量式训练方法研究. 北京大学学报(自然科学版), 2013, 49(1).
- ◆ Hao Xiong, Jun Xie, Hui Yu, Kai Liu, Wei Luo, Haitao Mi, Yang Liu, Yajuan Lv and Qun Liu. The ICT Statistical Machine Translation Systems for the IWSLT 2010. In Proceedings of IWSLT 2010.
- ◆ 罗维, 吉宗诚, 吕雅娟, 刘群. 一种改进词语对齐的新方法. 第五届全国青年计算语言学研讨会(YWCL2010), 武汉. 2010 年 10 月.

荣誉 & 奖励

- ◆ 在 360 工作期间: 2017 年度公司优秀个人、2016 年第一季度季度之星; 2017 年全年绩效位于 top 5%、多次绩效评估位于 top 20%。
- ◆ 在中科院计算所读书期间: 连续 3 年获得中国科学院研究生院三好学生。
- ◆ 在华中科技大学读书期间: 湖北省 2009 年优秀学士学位论文(2009.7)、华中科技大学 2009 届优秀毕业生、1 次国家奖学金、2 次校三好学生、1 次校级特优生、1 次校级优秀学生干部。