

李俊毅

13603542466 ljyduke@gmail.com dukeenglish.github.io 1994.09

教育经历

爱丁堡大学 人工智能硕士（NLP方向）	2016.09 – 2017.11
湖南大学 信息安全本科	2012.09 – 2016.06

工作经历

字节跳动	2020.08 – 至今
内容理解与AIGC负责人。曾负责房产AIGC，用户增长算法方向，商业化广告直播高光。	
电子科学研究院	2019.01 – 2020.07
负责 NLP 平台与问答系统。对 NLP 相关工作做全面负责，并撰写相关需求文档。协助培养一名研究生（知识图谱方向，主要辅导信息抽取）。	
美团点评	2018.01 – 2019.01
负责 NLP 平台设计开发工作。算法通道新员工评定TOP2（不区分一二名）	

部分项目经历

评价UGC内容生态负责人	2024.02 – 至今
方向负责人 3人，统筹多团队协作 生活服务	
梳理业务与方向，内容结构化理解与增长负责人。完成内容画风优化，高质量内容增长以及内容标签化建设。	
<ul style="list-style-type: none">• 内容理解：多业务多体裁内容理解。实现人机协同标注与自动训练平台，提升标注效率与训练效率，解放算法人力• 内容生成：基于LLM进行自动化生成功能，通过生成方式优化，偏好对齐等手段，推动总结/推荐语生成，并完成POI/商品等标签化建设• 内容增长：设计实现uplift模型，完成线上激励增长技术建设，相比人工AB收益2%• 工程优化：设计实现DAG工程架构，支持线上Node敏捷安全开发。	
垂直业务AIGC负责人	2023.05 – 2023.12
方向负责人 11人，统筹多团队协作 垂直业务-房家	
预研大模型技术，推动主导业务立项，并实现从0到1推动LLM应用上线。牵头推动各个团队合理分工与配合，制定在线与离线评估指标，最终回答效果人工评估达标率80%，优秀率70%。线上数据LT累计增加1.43%，人均对话2.1次，高质量线索占比28%，内容物料外放SEO获客1.9w。	
<ul style="list-style-type: none">• 团队方向设计与梳理：整体团队架构分为大模型、业务和公用基础建设三部分<ul style="list-style-type: none">– 大模型包括：知识图谱、大模型优化、AIGC与对话引擎（3人）+ 实习协助数据处理– 业务包括：幸福里对话机器人，住小帮对话机器人、General AIGC（3人）+ 实习协助GAI– 公用基础建设：房源库建设、通用NER与意图理解、AIGC审核安全（3人）– 外部依赖：房源推荐（1人）+ 其他线索变现等合作• 产品建设：推动协同产品组织多次内部测试、用户调研与竞品调研、数据与指标建设，分析优化产品迭代思路• 内容理解与知识库建设：引入LLM进行内容理解与内容生产，设计知识库并实现其自动更新（人机协同），以加强业务信息与大模型的协同应用• 产研协同：梳理从NLU、RAG到Prompt的整体优化方案，引入产运进行意图分类、检索知识库与prompt协作优化与调试，释放研发的调试人力	

- 整体设计检索-生成结合对话系统，主要从模型、NLU与决策、Prompt Engine、安全性后处理、插件、反馈优化等几个技术方面进行优化
 - 模型优化：完成模型数据、训练、评估三个环节的建设。
 - * 建立benchmark，涉及通用评估和垂直评估
 - * 设计实现模型自动化评估和选型pipeline，主要基于benchmark进行
 - * 重点通过SFT在猜你想问，query改写上优化分别将准确率提升了20个百分点左右
 - * 建设Reward模型作为LLM的自动化评估，并为未来优化做储备
 - AIGC与对话引擎：
 - * 架构上设计这几个模块：NLU (NER+意图识别, 双90+), Prompt Engine (RAG-Graph, RAG-embedding, prompt构建), NLG (RAG & LLM), 安全机审 (影响NLU、NLG以及知识库构建), 插件调用
 - * 同步生成、异步生成 (队列)、流式生成 (轮询, sse)
 - NLU与决策：设计多级意图与状态转移机制并实现自动化决策 (PE、检索、plugin等)，实现LLM+状态机落地NLU，覆盖意图与改写评估均在90%以上；支持反问、线索引导等多业务需求，同时探索直接基于LLM的决策能力。
 - Prompt Engine：基于COT思想，利用NLU与状态机结果判断调用plugin/retrieval，并完成prompt线上自动化构建，最终引导LLM进行生产
 - 安全性后处理与工程优化：设计流式机审后处理并落地，设计实现在线与离线幻觉问题解决方案；基于流式生产需求，协同前后端一同调研并设计优化工程架构，并完成落地
 - 基于用户反馈的优化：设计AB实验机制支持线上迭代；完善线上反馈机制，引入用户反馈来优化模型性能
- 通用生产建设：基于LLM能力，开发prompt调试平台，由产品进行Prompt调试以实现RD人力释放，支持线上标题改写、生成，标签/摘要生产与内容理解、问答生产以及增长物料生产等多业务需求，均取得正向效果。

增长算法方向负责人

2021.09 – 2023.02

方向负责人 3人，统筹多团队协作

房产

以技术身份切入，贯穿产运研角色以推动获取整体项目收益。AB实验收益在3%-30%不等

- 自动化投放：针对以关键词作为定向条件的搜索广告，搭建关键词自动更新全链路，构建召回与排序模块
- 创意优化：协调多团队，探索实验新的创意样式 (VR互动创意、动态词包等)，获取显著差异化优势
- SEO：针对搜索引擎进行网站优化获量，打通前后端产运研协作优化
- RTA 投放：挖掘人群包实现人群定向，建立用户价值预估模型，精细化增长

直播引流视频生成与投放

2020.08 – 2021.09

项目负责人 4人，统筹多团队协作

商业化广告

负责直播视频生成的架构设计、模块分工以及整体统筹协调工作，包括：直播视频数据流截取，计划与创意编排模块、视频加工与剪辑生成模块、其他视频元素生成 (文案、视频组件、模版等) 以及视频优选模块。

- 负责直播视频生成，实验获得20%的收益，当前已基于此上线产品
- 设计并沉淀直播视频生成方法论，将广告创意拆解为业务产运经验和创意各组件要素进行分析，并从时空两个维度对视频进行理解与生成
- 负责视频生成平台从 0-1 的设计编码，支持业务包括：广告诊断、直播广告视频生成、派生和自助产品

美团 NLP 平台建设

2018.05 – 2018.12

项目负责人

美团点评

负责美团用户平台下 NLP 平台规划设计，负责项目推进。在开发中负责NLP平台 python 框架设计实现、服务封装和部分算法 (词法分析) 设计实现

- 平台支持酒旅评论分析、景点问题测评等UGC 处理，跑腿标签生成，POI相似度、营业预测以及证店一致性

等相关业务，业务响应时间缩短50%以上

- 设计实现基于 BiLSTM+CRF 和 BERT 的词法分析服务，在内部测试集准确率和召回率以及 f1 值均由 94%左右提升至 96%左右。
- 校招即设计实现 poi 营业状态预测算法，准召均在90%以上，并就此申请专利
- 针对美团跑腿业务基于历史搜索内容实现标签提取与推荐，在不同城市的CTR和提单率均有所提升，CTR 提升最大达 3.41%，提单率提升最大达1.76%
- 设计实现基于 UGC 的标签关键词生成算法并设计标签体系

基于 NLP 平台的问答系统

2019.01 – 2020.08

项目负责人

电子科学研究院

负责设计实现研究院对话系统，作为重要演示成果，参与多项目竞标。

- 在 TPU 上实现基于 bert 的模型预训练，使用业务相关的语料与通用预料一同混合训练，并使用小的词表，以降低参数量（使用bert-base 结构可以降低9.8%，RoBERTa-tiny 结构可以降低 92.6%）
- 实现基于自训练 bert 的文本相似度计算，文本相似度的准确率由 78.5%提升至 88%左右，top5 的召回率由83%提升到93%左右

社会实践与学术活动

论文

- CLUE: A Chinese Language Understanding Evaluation Benchmark (COLING2020)
- Light Pre-Trained Chinese Language Model for NLP Tasks (NLPCC2020)
- Few-shot Learning for Chinese NLP tasks (NLPCC2021)

相关活动与应用实践

- ACL/EMNLP/NLPCC/AACL 等 NLP 相关会议审稿人，累计申请十二项专利，十一项第一作者
- 组建 11 人团队，参与公司内比赛获奖，成员属性横跨产运研设计与战略
- Coursera 课程 translator&reviewer，翻译出版《Python 忍者秘籍》，字节跳动技术团队公众号发文
- CLUE 中文语言理解测评基准执行委员会、NLPCC2020 会议 Shared Task 轻量级模型竞赛负责人
- 掘金受邀演讲两次，知乎科技互联网万粉作者

分享

- 知道自己在每个时间点想要什么。
- 兴趣特长：喜欢火影忍者，漫威系列。爱好篮球、吉他，主持经验丰富。
- 获评湖南大学优秀毕业生，优秀学生干部，辩论赛亚军。任学生会主席，辩论队领队，图书馆学生馆长。