



邓嘉鑫

+86 13260479736 | dengjiaxin2022@ia.ac.cn | jiaxindeng.xyz | 汉族 | 中共党员



教育经历

中国科学院自动化研究所 | 模式识别与智能系统, NLPR | 学术型硕士研究生 2022-09 – 2025-06
GPA: 3.62/4.00, 模式识别国家重点实验室, 主要研究方向为多模态学习, 连续学习, 在实时推荐系统领域方面有一定的研究和工程经验。2025年应届生。

电子科技大学 | 计算机科学与技术, 计算机科学与工程学院 | 工学学士 2018-08 – 2022-06
GPA: 3.99/4.00, 专业排名 3/254(前 1.2%), 四川省优秀大学毕业生, 获优秀奖学金多次, 全国大学生数学竞赛二等奖。

主干课程

- 本科阶段: 数据结构与算法(91)、数字逻辑与微处理器系统设计(挑战性课程)(90)、线性代数与空间解析几何(95)、电子电路基础(挑战性课程)(94)、Java 程序设计(94)、软件工程(94)、人工智能(92)、数学分析(97)、离散数学(97)、计算机操作系统(89)、随机数学(92)、高级语言程序设计(93)、计算机组成原理(90)
- 硕士阶段: 计算机算法设计与分析(94)、风险投资(91)、算法中的最优化方法(94)、自然语言处理、模式识别、高级机器学习

获奖情况

2024.05	中国科学院大学 2023-2024 学年度优秀三好学生	Top 70
2023.06	中国科学院大学 2022-2023 学年度优秀三好学生	Top 15%
2023.05	ACM SIGMM ICMR 2023 最佳论文提名	Top 5
2023.03	教育部-华为“智能基座”未来之星	Top 15
2022.03	2022 届四川省优秀大学毕业生	Top 409
2021.04	美国大学生数学建模竞赛 Honorable Mention	Top 24%
2020.12	电子科技大学“大学生创新创业训练计划”	优秀结题
2020.10	电子科技大学优秀学生奖学金(连续三年)	一等奖
2019.11	第十一届全国大学生数学竞赛(非数学类)	省级二等奖
2019.05	电子科技大学第二十九届数学竞赛(非数学类)	二等奖

技术成果

2024.05	Jiaxin Deng, Shiyao Wang, et al. MMBee: Live Streaming Gift-Sending Recommendations via Multi-Modal Fusion and Behaviour Expansion. In KDD 2024. (CCF-A, Acceptance Rate: 20%)	
2024.03	Jiaxin Deng, Shiyao Wang, et al. A Multimodal Transformer for Live Streaming Highlight Prediction. In ICME 2024. (CCF-B, Poster)	
2023.04	Jiaxin Deng, Dong Shen, et al. A Unified Model for Video Understanding and Knowledge Embedding with Heterogeneous Knowledge Graph Dataset. In ICMR 2023. (CCF-B, Oral presentation)	
2024.01	《基于多模态 Transformer 的帧级别直播点击率预测方案》发明专利	实质审查
2023.01	《基于异构知识图谱数据集的多模态视频特征提取方案》发明专利	专利公开
2020.10	《一类中值问题的一般解法》论文录用(《数学学习与研究》期刊)	第一作者

实习经历

快手科技多媒体内容理解组(MMU) | 算法实习生 2022.03 – 2022.10

- 基于异构知识图谱数据集的多模态视频表征学习: 从事关于多模态视频理解大模型与知识图谱相结合的研究探索工作, 利用知识图谱中普遍存在的常识性知识使得多模态视频理解大模型具备常识性推理的能力, 并进一步提高大模型的视频表征能力。在多模态任务中引入知识图谱信息, 并通过网络结构和损失函数的优化, 使得模型在朴素的内容理解能力之外, 具备了常识性推理能力, 进而提升内容理解(视频检索视频)和知识

图谱(视频推理标签, 标签推理标签)的相关指标。在视频推理标签, 标签推理标签的指标上, HitS@10 分别提高了 **42.36%** 和 **17.73%**, 取得了 SOTA 的效果。该方案已形成论文, 题目为 **A Unified Model for Video Understanding and Knowledge Embedding with Heterogeneous Knowledge Graph Dataset**, 目前已被 ICMR 2023 接收并获得了 **Best Paper Candidate** 称号, 同时基于此论文申请了题为《基于异构知识图谱数据集的多模态视频特征提取方案》的发明专利, 已进入实质审查阶段。同时本人在业务方面辅助完成了快手河图视频标签框架 v3 的迭代升级, 负责了 3 个垂类(情感、美食、三农)的二三级下探工作, 使其准确率与覆盖率符合上线需求, 目前本人负责的新的垂类已上线。

快手科技多媒体内容理解组(MMU) | 算法实习生

2023.01 – 2023.06

- **基于多模态的直播精彩时刻预测方案:** 直播间在互动的过程中, 其长播率/点击率等消费指标曲线通常会呈现非平稳状态, 而其中的一些峰值点往往对应着直播内容的高光时刻, 如主播跳舞/唱歌等。如果算法能够感知到这些精彩片段的即将发生, 并对该直播间的分发力度加以扶持, 则可以放大这些精彩内容带来的收益; 对主播来说, 主播的精彩时刻有更多的观众, 正向激励主播进行更多优质内容的呈现; 对用户来说, 看到的内容“正值巅峰”, 用户体验会更好; 对平台来说, 时长/点击的提升意味着进入直播间的人更多, 时长和打赏也有更大提升的空间; 而目前系统遇到的问题是还不能及时地感知到“精彩时刻”的即将发生(即系统的预测结果趋于平稳)。因此我们希望通过直播实时内容的建模, 利用多模态视频理解大模型的优势, 为快手用户分发精彩直播片段, 从而增长快手直播页面的自然流量时长以及直播观看的 DAU。此方案利用多模态预训练大模型以 20 个历史连续直播片段的视频帧, 主播 ASR, 观众弹幕以及主播的垂类 embedding 作为输入, 模型输出下一个直播片段的 CTR 以达到预测直播片段精彩时刻的目的。损失函数方面采用经典的排序学习中的 listwise 损失与带差分约束的 pairwise 损失相结合的方式监督模型预测 CTR 的学习过程。同时采用了 Dynamic Time Warping 的对比学习方式缓解了各个模态之间的对齐问题。模型采用流式训练架构, 每天迭代 1.4 万+个直播间进行在线学习以适应数据分布的缓慢变化。通过两天的 A/B 实验表明本方案在快手极速版中页面自然流量时长增长达 **6.015%**, 直播观看大于 1 分钟的 DAU 增长达 **1.772%**。该方案形成的题为 **A Multimodal Transformer for Live Streaming Highlight Prediction** 的论文已被 ICME 2024 接收。

快手科技直播推荐部/实时内容推荐组 | 算法实习生

2023.08 – 2023.09

- **基于 GNN 行为拓展与多模态特征的打赏预估方案:** 在现有的基于 MMOE 的多目标预估模型的基础上, 通过对主播侧内容与用户侧行为序列进行建模, 增强直播间实时内容分发与用户兴趣的关联性, 从而提升我们流量分发的策略的准确性。对于主播侧方面, 对当前切片下的多模态视觉特征/评论特征/音频特征进行三路正交融合, 并作为 gtr 预估的一路输入, 建立多模态内容和打赏目标的关系。在模型上线过程中发现推理性能存在问题, 经过性能优化(fp16, 维度更改为 8 的倍数, batch 操作等)后, 线上极速版 A/B 测试送礼观众数 **+4.215%**, 直播曝光次数 **+1.982%**, 关注主播次数 **+2.642%** 上述工作虽然显著提升了链路对实时内容变化的感知能力, 然而当前模型缺少用户侧内容信息, 使得模型对用户兴趣的感知能力较弱, 模型的个性化推荐能力不强。同时由于用户打赏行为较为稀疏, 同一个用户与同一个主播的 <user, author>pair 对多次出现的频率极低。因此, 对于用户打赏历史行为的学习及 <用户, 主播>pair 关系的学习是不充分的。由于图序列是更能反应用户个性化的特征而适用于排序场景。图序列抽取的目的是基于用户或目标广告的同构或异构多跳序列, 进行序列扩展, 帮助感知用户自身行为及与社区行为, 反应用户在各场景各粒度上的兴趣。相比用户历史行为序列, 可以帮助解决用户行为稀疏的问题, 主动学习多跳下的相似行为与聚合兴趣, 因此我们将图序列本身作为用户的扩展兴趣序列或者主播的高相关序列直接作为信息输入到排序模型。首先采用已有的打赏行为与 swing 相似度构建了包含 u2a/a2u/a2a 边的图存储服务, 此服务以 hive 表产出作为输出天级更新; 在模型训练时, 用户侧以 uid 作为触发, 在图中挖掘多跳下的信息, 主要包括: u2a2u(目标用户的相似用户), u2a2u2a(目标用户的潜在兴趣主播), u2a2a(目标用户的相似主播); 主播侧通过 aid 作为触发, 在图中挖掘多跳下的信息, 主要包括: a2a(目标主播的相似主播), a2u2a(目标主播的潜在兴趣主播); 得到上述序列后对其进行 pooling 后输入 MMOE 模型中。加入行为拓展序列后离线测试模型 gtr AUC **+1.0029pp**, UAUC **+1.1667pp**, WUAUC **+1.3072pp**。同时该方案已经形成论文, 题目为 **MMBee: Live Streaming Gift-Sending Recommendations via Multi-Modal Fusion and Behaviour Expansion**, 该论文已被 KDD 2024 接收。

个人总结

- 本人乐观开朗、在校成绩优异、自驱能力强, 具有良好的沟通能力和团队合作精神。
- 可以使用英语进行工作交流(四级成绩 628, 六级成绩 590), 平时有阅读英文书籍和口语练习的习惯。
- 有较强的自主学习能力和动手实践能力, 本人硕士期间各类课程成绩优异, 同时对新事物与新知识有着强烈的好奇心与学习接受能力, 有较强的动手能力。