

李涵宇

✉ l-hy12@outlook.com · ☎ (+86) 188-1013-7007 · Birthdate: 2002/01/15

教育背景

| | |
|-----------------------------------|-------------|
| 清华大学, 北京 | 2020 – 至今 |
| 在读博士研究生 计算机科学与技术, 预计 2025 年 6 月毕业 | |
| 纽约州立大学布法罗分校, 纽约州, 美国 | 2019.7 |
| 访问学者 (导师: Junsong Yuan) | |
| 清华大学, 北京 | 2016 – 2020 |
| 学士 数理基础科学 | |

研究兴趣

推荐系统, 序列、重排序、跨域推荐模型, 扩散模型等技术在推荐中的应用

项目经历 (以下成果均为第一作者)

凝胶电泳图像中异常单克隆蛋白检测 纽约州立大学布法罗分校 2019.7-2021.3

- 电泳法检测单克隆蛋白对多种疾病的诊断至关重要。该工作用高斯混合模型表示电泳图像, 并用峰值检测方法识别异常视觉特征。该分类器具有良好的可解释性, 并且在小训练集上表现也较好。
- 该成果已发表在计算机科学会议视觉通信与图像处理 (VCIP) [代码链接] [论文链接]。

考虑用户隐私的远程线上实验工具 清华 2020.9-2021.7

- 为了促进信息检索领域用户实地实验的开展, 减少研究人员的开发工作量, 消除平台设计与开发壁垒, 并在数据的自然收集的过程中最大程度保护用户隐私, 我们设计、开发了一个考虑用户隐私的开源远程用户行为实验工具 RUS-toolkit。
- 该成果已发表在顶级计算机科学会议国际信息检索大会 (SIGIR, CCF A) [代码链接] [论文链接]。

考虑动态多样性需求的重排序 清华-腾讯 2021.7-2022.7

- 大多数多样性感知研究都认为, 提供更多多样化的结果总能提高用户满意度。然而, 用户对多样性有不同程度的需求, 这种需求在不同的交互会话中会发生动态变化。
- 我们通过对大规模真实推荐数据集的广泛分析验证了这一说法。然后, 我们提出了一种新颖的重排序方法来满足用户的动态多样性需求, 即考虑动态多样性的重排序 (DDAR) 模型。
- 该方法在多个数据集上取得显著提升, 目前在投。[代码链接]

过滤协同信息的跨域推荐框架 清华-腾讯 2022.7-2024.1

- 尽管目前关于跨域推荐的研究已经相当丰富, 但大部分工作主要集中在让重叠用户在不同领域的建模更具表现力, 以传递领域间的信息, 而在监督信号层面上解决负迁移问题的研究相对较少。此外, 现有的方法大多是整体的跨域模型, 无法跟上最新的单域推荐进展, 整体架构不够灵活。
- 针对上述研究难题和现有模型的不足, 我们提出了一种新的框架——协同信号正则化的用户转换 (CUT)。该可扩展的跨域推荐框架通过过滤源域中无关信息, 利用更多的有用知识, 以提高目标领域的性能。同时, 它还可以使用各种单领域推荐系统作为基础, 并将其扩展到跨域推荐任务。
- 该成果已被顶级计算机科学会议国际信息检索大会 (SIGIR, CCF A) 录用[代码链接]。

技术能力

编程语言 Python, C, C++, Java · 开发平台 Pytorch, Tensorflow, Django · 英语 TOEFL: 109, GRE: 329

学术活动与个人荣誉

| | |
|-------------------------|---------------|
| 清华大学《信息检索的前沿研究》课程助教 | 2022 年 3 月-至今 |
| M 奖 (前 7%), 美国大学生数学建模竞赛 | 2019 年 1 月 |