

【LLM论文阅读】LLMRanker：利用LLM对候选物品集合进行zero-shot排序

方方的算法花园 2024年10月11日 16:15 北京

论文标题：

Large Language Models are Zero-Shot Rankers for Recommender Systems

论文链接：<https://arxiv.org/pdf/2305.08845>

论文作者所在机构：中国人民大学、加利福尼亚大学圣迭戈分校、腾讯微信

Github链接：

<https://github.com/RUCAIBox/LLMRank>

一句话概括：将推荐问题形式化为条件排序任务，精心设计多种prompt策略，深入研究 LLMs 的零采样排序能力及影响因素。

创新点

1. 将推荐问题形式化为条件排序任务

把顺序交互历史作为条件，其他检索到的项目作为候选，利用 LLMs 的内在知识对候选项目进行排序，为研究 LLMs 在推荐系统中的排序能力提供了新的视角和方法框架。

2. 精心设计多种提示策略

针对历史交互顺序感知

- **顺序提示 (Sequential prompting)**：按时间顺序排列历史交互，如 “I’ve watched the following movies in the past in order: ‘0. Multiplicity’, ‘1. Jurassic Park’,...”
- **近因聚焦提示 (Recency-focused prompting)**：除了顺序交互记录，还添加强调最近交互的句子，例如 “I’ve watched the following movies in the past in order: ‘0. Multiplicity’, ‘1. Jurassic Park’,... Note that my most recently watched movie is Dead Presidents...”
- **上下文学习 (In - context learning, ICL)**：通过扩充输入交互序列本身引入示例，如 “If I’ve watched the following movies in the past in order: ‘0. Multiplicity’, ‘1. Jurassic Park’,..., then you should recommend Dead Presidents to me and now that I’ve watched Dead Presidents, then...”。这些策略可以触发 LLMs 感知历史交互的顺序，提高排序性能。

针对偏差缓解

- **针对位置偏差**：采用位置偏差缓解的自举策略 (Bootstrapping for position bias)，通过多次随机打乱候选集并重复排序，然后合并结果来缓解 LLMs 对候选项目在提示中的顺序敏感问题，即位置偏差。例如对候选集重复排序 3 次，每次随机打乱候选项目顺序。
- **针对流行度偏差**：让 LLMs 关注历史交互可以减少流行度偏差，即当 LLMs 更多地关注历史交互时，推荐的项目流行度得分会降低。

3. 深入研究 LLMs 的零采样排序能力及影响因素

- 通过在两个广泛使用的数据集上进行大量实验，研究发现 LLMs 具有有前景的zero-shot排序能力。
- 分析了 LLMs 排序时依赖的因素，包括项目流行度、文本特征以及用户行为等。例如在不同数据集上，LLMs 对包含流行项目的候选集排序表现不同，在 ML - 1M 数据集上基于 LLM 的方法对包含流行项目的候选集排序不佳，而在 Games 数据集上表现出项目流行度和文本特征对 LLMs 排序贡献相似。同时，LLMs 的排序性能也受基于交互的候选生成模型检索到的困难负例影响，但不如一些传统基于交互的排序器严重。

关键结论

- (1) LLM难以感知给定顺序交互历史的顺序。通过采用专门设计的prompt，可以触发LLM来感知顺序，从而提高排名性能。
- (2) LLM在排名时存在位置偏差和流行度偏差，bootstrapping或专门设计的激励策略可以缓解这种偏差。
- (3) LLM显示出优秀的zero-shot排名能力，尤其是在具有不同策略的多个候选生成模型检索的候选上。

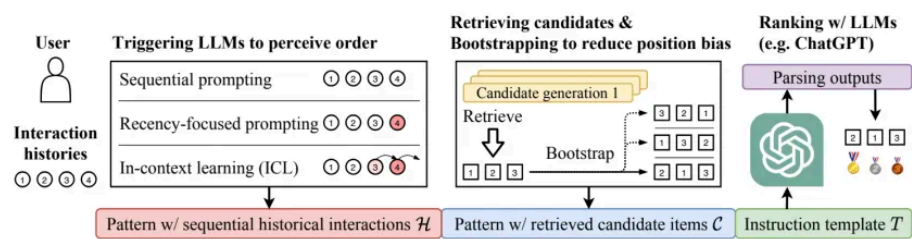


Fig. 1: An overview of the proposed LLM-based zero-shot ranking method.