

2023腾讯：AgentCF-利用多智能体进行协同学习，提升推荐系统效果



SmartMindAI

专注搜索、广告、推荐、大模型和人工智能最新技术，欢迎关注我

已关注

4 人赞同了该文章

Introduction

随着LLMs在决策中的出色表现，它们正被用于模拟包括日常生活和软件开发在内的复杂社会行为⁺。尽管已有研究主要关注对话，但用户在推荐系统中非言语行为⁺，如商品选择，揭示了更多个性信息。然而，当前方法通过解析文本可能无法完全捕捉到这些行为背后的深层模式，因为通用模型与个性化行为建模⁺之间存在差距，如购买尿布与啤酒的例子。因此，关键任务是研发新技术，以更精确地在LLM驱动模拟中描绘人类行为。

在本文研究中，我们以推荐系统⁺为例，探究如何利用LLM驱动的代理来模拟用户与内容间的交互。尽管已有如RecAgent这样的方法通过用户代理进行，但它们往往过于依赖通用LLMs，未充分考虑物品侧的个性化建模。推荐系统的成功依赖于全面理解用户对每个物品的独特喜好。我们的目标是不仅模仿用户行为，而且通过双向建模，即同时考虑用户和内容代理的动态互动，来优化推荐效果。这涉及到精细地刻画用户与物品间的双向关系，因为用户的选择不仅是基于自身偏好，也受物品吸引力的影响。

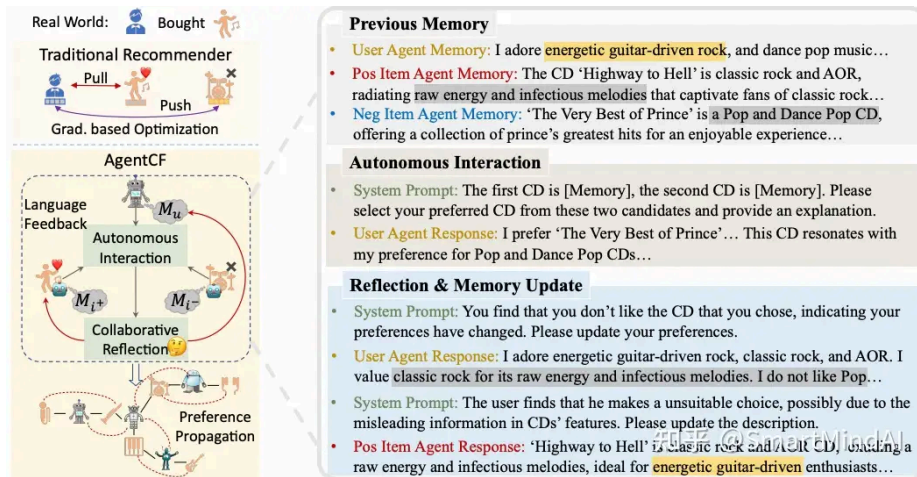
模拟用户代理和内容代理的互动，而非单纯依赖文本，来理解用户-内容关系。我们在优化过程中，让两者参数同时更新，以适应交互数据。

我们不仅关注用户行为，还重视物品的吸引力，通过双向建模来提升推荐精度。与以往自我探索优化不同，我们推崇的是以整体互动为中心的协作学习，以全面理解和建模用户和内容的动态交互。虽然Reflexion和Self-Refine在优化通用代理上有一定贡献，但我们的方法更侧重于跨代理的协作，以实现更精准的个性化推荐。

因为，我们提出了AgentCF，这是一种基于代理的协同过滤⁺模型，打破了仅关注用户行为的传统，同时考虑物品的特性。我们通过赋予内容代理记忆功能，模拟潜在用户的喜好。核心是通过协作优化两个代理，利用LLMs的反思能力，让它们在交互中学习和调整。

我们设计了协作反思机制，允许代理间相互影响，以互动方式内化协同过滤概念。在实证研究⁺中，AgentCF在推荐任务上表现优异，与经典模型和LLM推荐器相当。实验还展示了模拟的个性化交互，用户和内容代理展现出多样交互类型，为用户行为模拟提供了新的视角。这项工作的重要贡献包括：将用户和内容都作为代理，发展双向互动建模，以及通过协作反思实现记忆更新。

Methodology



在传统的推荐系统中，我们处理用户集合 \mathcal{U} 和内容集合 \mathcal{I} ，以及它们的交互历史 \mathcal{D} ，这些数据按时间顺序排列。目标是创建 $f(u, i)$ 函数，表示用户对内容的喜好度，特别强调与用户互动过的内容（正项 \mathcal{D}_+ ）比非互动的内容（负项 \mathcal{D}_- ）应得到更高评分。BPR通过用户对两项的相对排名建模，而NCF则运用神经网络来学习用户与内容间的关联。

我们通过这样的方式将学习用户偏好的过程转化为机器学习中的梯度优化问题，利用训练数据集 \mathcal{D} 进行优化。这正是我们'AgentCF'模型设计的基础，它借鉴了上述合作学习方法，通过用户代理和内容代理的协作互动来提升推荐的准确性。在这个新型的LLM驱动的任务中，我们定义了一个函数

$$f_{LLM}(u, \{c_1, \dots, c_n\})$$

它负责对候选内容列表 $\{c_1, \dots, c_n\}$ 进行排序。与传统推荐系统中动态调整的偏好函数相比，LLM在这个过程中保持恒定，不参与训练或优化过程。LLM作为一个固定的指导者，提供一种预测用户行为的方式，而不是通过学习优化来改进其推荐。这种设计使得LLM能够提供稳定且个性化的排序建议。

AgentCF概览 在处理代理应用时，记忆和反思机制在不同任务中已有所研究，但多数侧重于任务或用户层面。然而，针对推荐系统中物品这一重要交互对象，我们认识到建立用户-内容互动模型的重要性，尤其是理解和利用相似者相吸的原则。为此，我们提出AgentCF，它聚焦于物品代理，以增强协同过滤的核心逻辑。我们的两大技术创新如下：

1. 协同记忆优化：我们创新性地通过用户与物品代理的互动而非仅依赖历史交互，进行非梯度的联合优化。在每个迭代中，代理通过互动后反思并调整自己的状态，以反映实际交互与预期行为的偏差，进而动态适应真实环境。

的相似用户选择，从而通过偏好自我传播来影响未来的决策。这种方法从用户角度强化了协同过滤，通过实际互动强化了用户对物品的认知，增强了推荐的个性化。

Collaborative Agent Optimization

我们通过用户和内容代理的互动，更新他们的记忆，模仿隐式偏好传播。不同于传统方法，我们的模型允许学习新信息，克服冷启动问题，且记忆是动态的，随交互更新。这种动态协作强化了个性化推荐，结合了传统和现代AI的优势，提供了更全面、动态的推荐策略。

Memory Design

为了将LLM代理专用于推荐，我们设计了用户和物品代理各自的内存结构。用户代理的'内存模块'负责存储用户偏好，包括短期记忆 M_u^s （反映近期更新的偏好，初始为通用偏好）和长期记忆 M_u^l （记录用户偏好的演变）。后者是一个历史偏好集合，便于用户在新互动中获取相关历史信息。

物品代理的'内存模块'则是可调的，主要用于存储物品的固有特征和被用户偏好认可的信息。尽管每个物品有一个统一的 M_i ，它会基于身份信息（如标题和类别）初始化，并在与用户互动中随用户偏好的反馈而更新，这有助于提取物品全局特性，促进协同过滤中的信息传播。

Autonomous Interactions for Contrastive Item Selection

$$i^o = f_{LLM}(M_u; M_{i^-}; M_{i^+})$$

$$y_{exp} = \text{Prompt}_{LLM}(i^o; M_u; M_{i^-}; M_{i^+})$$

Collaborative Reflection and Memory Update

我们通过上述交互模拟，使代理学会模仿真实交互，然后通过比较代理决策与真实数据来获取反馈，进行协作优化。不同于传统推荐的反向更新，我们不直接调整LLMs，而是专注于用户和内容代理的记忆更新。

每个决策（ i^o ）和解释（ y_{exp} ）都是一个反馈点，引导用户和内容代理进行协作反思。当代理做出符合实际行为的选择，我们就强化正面记忆；若选择错误，我们会引导他们修正错误，以提高模拟的准确性。这个反思过程基于用户和内容记忆，不同于任务特定或用户导向的反思，它强调了双方在理解偏好和特征上的合作学习。

$$M_u^{s'} \leftarrow \text{Reflection}^u(i^o; y_{exp}; M_u; M_{i^-}; M_{i^+}),$$

$$M_i' \leftarrow \text{Reflection}^i(i^o; y_{exp}; M_u; M_{i^-}; M_{i^+})$$

$$M_u^{l'} \leftarrow \text{Append}(M_u^l; M_u^s),$$

我们使用' $M_u^{s'}$ '表示用户 u 的反射短期记忆，用' M_i' '代表内容的反射记忆。用户 u 的更新长期记忆 $M_u^{l'}$ 通过结合短期记忆和历史信息，以便在新交互时检索关键偏好信息。

特别地，我们保持负向内容代理的记忆不变，因为LLMs可能过于关注其缺陷，而忽略了其对其他潜在用户的吸引力。用户和内容代理的更新遵循公式，这里的反思是基于用户 u 、正向内容 i^+ 和负向内容 i^- 的记忆，这有助于区分正负实例，深化它们对交互关系的理解。优化过程迭代进行，每次包括选择步骤和协作反思。目标是使用户代理决策接近真实用户的行为，或至多达到设定的最大迭代次数。在协同优化中，模拟用户和内容代理模仿真实个体的决策路径，用户代理个性化增强，内容代理通过学习交互来更好地适应接受者需求，从而提升模拟的真实性和交互质量。

Connection with Classical Recommendation Models

在推荐系统中，BPR和NCF等模型通过嵌入参数来表达用户偏好和内容特性，遵循向前偏好评估和反向参数更新的两步优化策略。我们的方法借鉴了这一思路，将用户和内容记忆视为语言层面的参数，模拟推荐系统的过程。选择过程对应前向评估，而协作反思则扮演后向更新的角色。不同于传统的梯度优化，我们利用反思机制，类似于语义梯度，引导用户代理自我调整，而不仅仅是调整参数。通过协作，用户和内容代理相互影响，形成偏好传递，这在用户-内容交互中实现了协同过滤的核心理念。

- 1. 联合偏好评估：LLM利用用户和内容代理的交互信息，对潜在内容进行评分，仿照实际用户兴趣度。
- 2. 协作学习调整：虽然不采用传统梯度更新，LLM通过对用户和内容记忆的反思，间接调整特征表示，以揭示潜在互动趋势。
- 3. 偏好传递：在后续互动中，LLM通过代理间互动将已学习的偏好知识传递，增强相似用户和内容的匹配，实现无梯度优化下的协同过滤效果。

$$\mathcal{R}_B = f_{LLM}(M_u^s; \{M_{c_1}, \dots, M_{c_n}\})$$

- 1. 领域偏好深化：LLM利用用户长期记忆，不仅当前兴趣，还能拓展到用户特定领域的偏好，从而提供更专业且精确的个性化建议。
- 2. 历史行为洞察：面对稀疏数据，LLM通过分析用户历史交互，预测他们对内容未来可能的兴趣，确保推荐序列的连贯性和相关性。

通过这些高级策略，我们的方法增强了LLM的个性化推荐效能，同时在数据有限时也能提供连续的用户体验。

$$\mathcal{R}_{B+R} = f_{LLM}(M_u^t; M_u^s; \{M_{c_1}, \dots, M_{c_n}\})$$
$$\mathcal{R}_{B+H} = f_{LLM}(M_u^t; \{M_{i_1}, \dots, M_{i_m}\}; \{M_{c_1}, \dots, M_{c_n}\})$$

M_u^t 代表从用户代理+u的长期记忆中获取的特定偏好，通过查询内容代理记忆来获取，这反映了用户在特定领域的兴趣。用户与m个内容的交互历史记录，如{ M_{i_1}, \dots, M_{i_m} }构成用户行为历史的集合。

通过优化后的用户和内容代理，我们系统能更准确地模拟现实生活中的交互模式，包括个体互动、社会行为及推荐系统整体动态。关键在于，我们把物品转化为代理，这促进了非生物对象间创新且吸引人的交互探讨，特别是在新内容初次接触时，内容间代理的交互可以自然地传递用户偏好，帮助新内容快速融入系统。

Experiments

Datasets

我们利用亚马逊评论的文本丰富子集进行了实证研究。

Evaluation Metrics

为了衡量性能，我们采用了NDCG@K (K=1, 5, 10) 作为标准，这源于先前的研究。我们采用留一法进行评估，即以每个用户最近的未考虑项作为实际目标。我们将模型当作一个排序器，与九个随机内容一起对目标项进行排列。为减少偶然性，我们对每个测试样本做了三次重复，并报告了平均结果。

Overall Performance

我们将在四个数据集上对我们的方法与基线方法进行比较，并在表中展示结果。

Method	CDs _{sparse}			CDs _{dense}			Office _{sparse}			Office _{dense}		
	N@1	N@5	N@10	N@1	N@5	N@10	N@1	N@5	N@10	N@1	N@5	N@10
BPR _{full}	0.1900	0.4902	0.5619	0.3900	0.6784	0.7089	0.1600	0.3548	0.4983	0.5600	0.7218	0.7625
SASRec _{full}	0.3300	0.5680	0.6381	0.5800	0.7618	0.7925	0.2500	0.4106	0.5467	0.4700	0.6226	0.6959
BPR _{sample}	0.1300	0.3597	0.4907	0.1300	0.3485	0.4812	0.0100	0.2709	0.4118	0.1200	0.2705	0.4576
SASRec _{sample}	0.1900	0.3948	0.5308	0.1300	0.3151	0.4676	0.0700	0.2775	0.4437	0.3600	0.5027	0.6137
Pop	0.1100	0.2802	0.4562	0.0400	0.1504	0.3743	0.1100	0.2553	0.4413	0.0700	0.2273	0.4137
BM25	0.0800	0.3066	0.4584	0.0600	0.2624	0.4325	0.1200	0.2915	0.4693	0.0600	0.3357	0.4540
LLMRank	0.1367	0.3109	0.4715	0.1333	0.3689	0.4946	0.1750	0.3340	0.4728	0.2067	0.3881	0.4928
AgentCF _B	0.1900	0.3466	0.5019	0.2067	0.4078	0.5328	0.1650	0.3359	0.4781	0.2067	0.4217	0.5335
AgentCF _{B+R}	0.2300	0.4373	0.5403	0.2333	0.4142	0.5405	0.1900	0.3589	0.5022	0.3359	0.4591	0.5247
AgentCF _{B+H}	0.1500	0.4004	0.5115	0.2100	0.4164	0.5198	0.2133	0.4379	0.5076	0.1600	0.3986	0.5147

Further Model Analyses

1 无自主交互：我们不提供用户的真实交互，而是直接将数据交给代理，观察模拟与真实交互的偏差，以此评估代理反思的有效性，结果显示这有助于识别和改进模型。(2) 去除非用户代理：我们移除用户代理优化，仅用用户历史交互，结果显示，单纯依靠这些记录无法充分捕捉用户深层偏好，突显了赋予代理自我学习和记忆的能力以适应用户行为的重要性。(3) 无内容代理

Variants	CDs _{dense}		Office _{dense}	
	N@1	N@10	N@1	N@10
AgentCF _B	0.2067	0.5328	0.2067	0.5335
w/o Auto. Interaction	0.1200	0.4964	0.1733	0.5031
w/o User Agent	0.1100	0.4693	0.2200	0.5145
w/o Item Agent	<u>0.1767</u>	<u>0.5128</u>	<u>0.1800</u>	<u>0.5169</u>

Performance Copmarison w.r.t. Position Bias and Popularity Bias

我们通过检验代理对位置和流行度偏见的敏感性来衡量其个性化模拟能力。结果显示，使用用户历史交互的LLMRank模型存在明显的流行度和位置偏好，倾向于推荐热门和排名高的内容。然而，尽管如此，我们的方法尽管不能完全消除这些偏见，但在稳定性上有所改善，表明它有效地模拟了个性化的用户代理，使其能够基于个人偏好而非通用常识进行评价。这展示了我们在不牺牲个性化的同时，如何克服通用模型+的局限性。

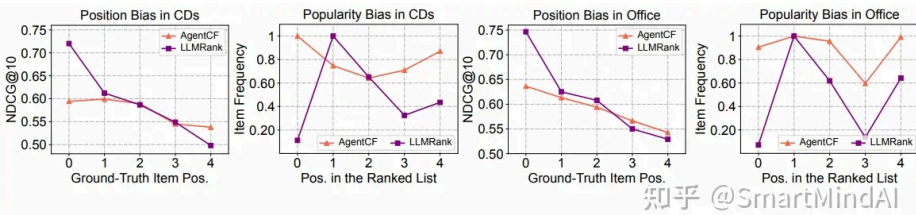
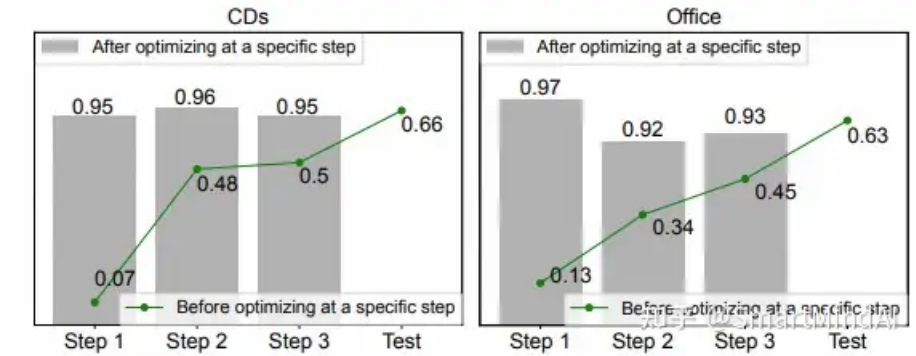


Figure 2: Analysis of whether our approach can simulate personalized agents to mitigate position bias and popularity bias.

Effectiveness of Collaborative Reflection

我们研究了协作反思对代理优化的效果，通过观察代理与真实用户的匹配度随优化过程的变化。我们采用三次最近交互作为优化输入，这称为三个优化步骤。每次优化后，我们检验未优化和优化后的代理能否准确识别正面内容。结果显示，优化过程导致用户代理的偏好与真实用户越来越接近，优化后的代理在首次尝试（未优化时）正确选择的比例显著提升，达到约95%，这证实了这种方法的有效性。

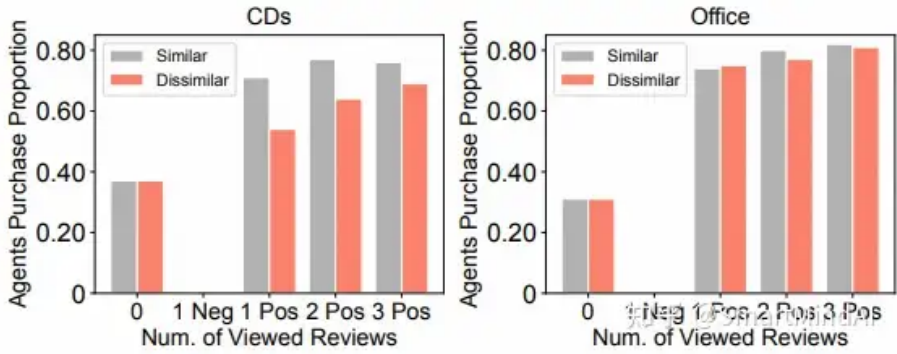


Simulations on Other Types of Interactions

User-user Interaction Simulation

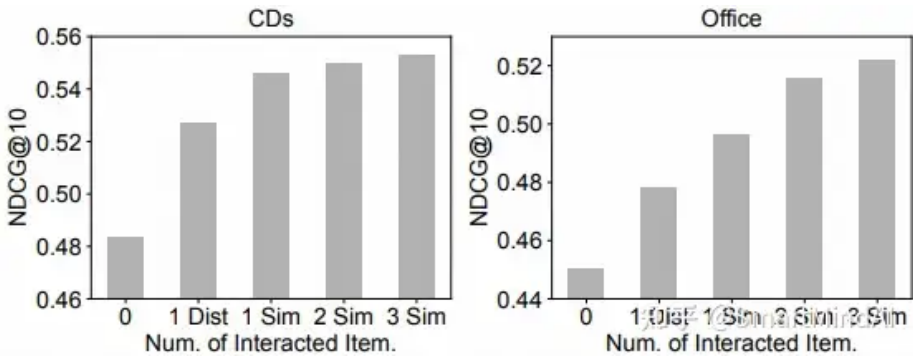
在实际场景中，用户在尝试新产品时通常会参考他人意见。为此，我们通过AgentCF模型模拟用户间的互动，让用户代理不仅浏览评论，还能生成自己的评价。具体操作是，当面对未与商品互动过的用户时，先让已互动过的代理编写评论。接着，我们将这些评论呈现给测试用户，观察他们在

变以及对商品的讨论，这些细节见附录以供深入分析。



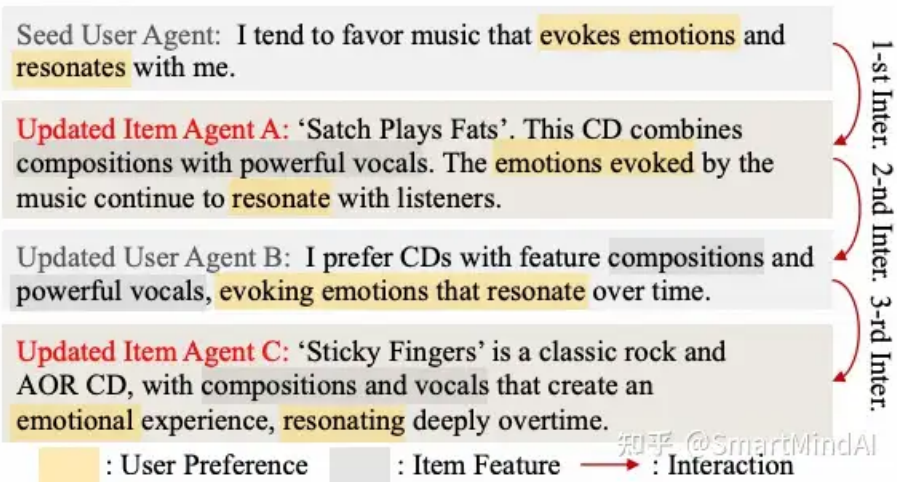
Item-item Interaction Simulation

我们通过模拟新物品和热门物品的交互来解决推荐系统中的“冷启动”问题。我们促使新物品代理与已熟知热门物品的代理互动，利用后者丰富的交互经验来预热冷启动的代理，通过估计潜在用户的喜好调整其记忆。然后，用户代理对九个训练好的物品进行排序，对比使用原始冷启动记忆和经过调整后的记忆结果。实验显示，这种交互增强了用户偏好的传播，显著提升推荐性能。重要的是，即使不同身份的新物品代理也能通过理解身份信息与个性记忆的关系，部分缓解了冷启动问题，说明交互策略对于激活物品间的互动至关重要。因此，AgentCF的这一设计有助于改进推荐系统的表现。



Process of Preference Propagation

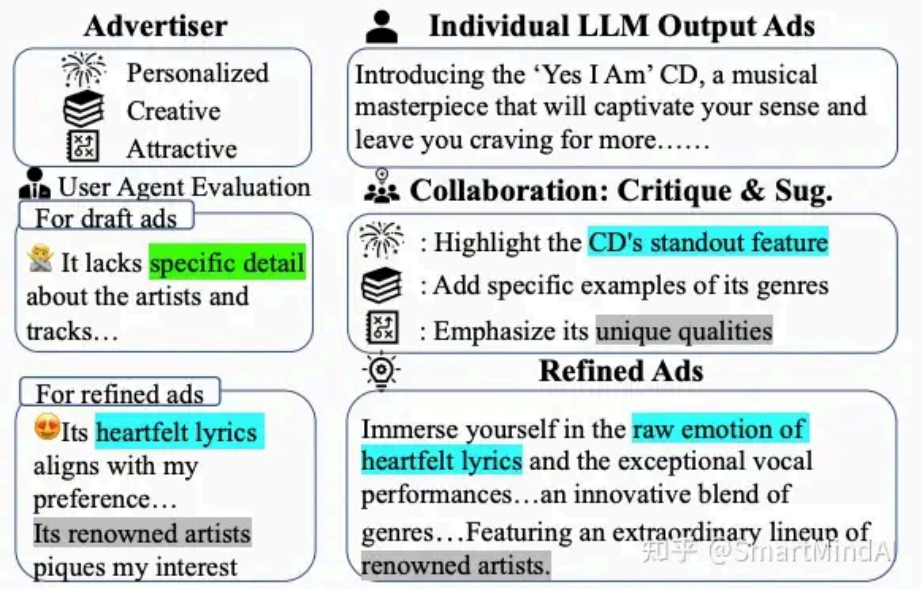
我们通过协作优化来促进偏好在用户和物品代理间的传播，以解决推荐系统的冷启动问题。首先，我们初始化种子用户代理具有非典型偏好的记忆。接下来，通过用户和物品代理的自主交互，记忆得以优化。



图中展示了优化后的记忆示例，显示出偏好在代理间通过互动传播。我们通过比较优化前后用户代理对种子偏好的模仿情况来评估传播效率。结果显示，随着交互增多，相似行为代理表达相似偏好

Collaborative Advertisements Creation

我们利用集体智慧理论，研究大型语言模型+协作在推荐系统中的应用。在广告创作的模拟环境中，我们设计了多元化的广告代理，分别擅长个性化、创意和吸引力。针对单一语言模型生成的广告草案，我们引导这些代理进行互动，提出批评和建议，以优化广告质量。



结果显示，通过这种方法，我们能有效地利用各代理的专业特长，生成更受用户欢迎的广告，体现了集体智慧在推荐系统中的增强效应。

原文《AgentCF: Collaborative Learning with Autonomous Language Agents for Recommender Systems》

发布于 2024-06-17 11:01 · IP 属地北京

多智能体 推荐系统 腾讯



理性发言，友善互动



发布



还没有评论，发表第一个评论吧