

赞同 145

分享

## 2024快手：如何将LLM的embedding应用到推荐系统中？



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

145 人赞同了该文章

### Introduction

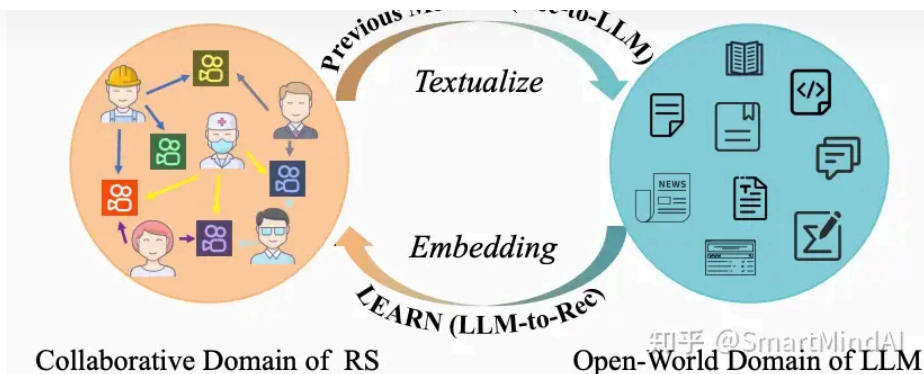
推荐系统（RS）利用ID嵌入策略处理用户和内容，通过交互学习建立关联。然而，它忽视了内容描述中的语义信息<sup>+</sup>，影响新用户和长尾用户推荐的精准度。因此，如何利用这些信息以改进RS是研究重点。

借助LLMs<sup>+</sup>在NLP的强大能力，尤其是其对开放世界知识<sup>+</sup>的掌握，研究人员正在探索如何将LLM应用于推荐系统，以增强其对内容描述的理解。目标是利用LLM的预训练内容嵌入，以补充ID嵌入法在新用户和长尾推荐上的不足。

首先，由于LLM的通用知识与推荐系统的交互数据（协作知识）不同，需要克服通用知识与推荐领域间的鸿沟。一种策略是冻结LLM参数，通过构建对话式任务，让LLM解析用户-内容交互，生成推荐。另一种方法是“微调LLM”，在特定推荐领域数据上训练，让它通过预测任务理解用户和内容关系。这两种方法都旨在通过LLM的逻辑推理，提升推荐系统的准确性和泛化能力。在实际工业环境中，直接应用推理的LLM处理全量用户交互数据困难。

鉴于LLM的输入限制通常为2K到4K个令牌，处理成千上万商品的频繁交互是个巨大的挑战。随着序列长度<sup>+</sup>的增加，计算复杂度会以指数级增长，这使得一次性加载所有历史交互成为不可能。因此，需要设计更有效的分块或近似策略，以降低对LLM资源的直接消耗。

同时，在推荐系统中微调LLM往往导致灾难性遗忘和性能退化。尽管在理论上可行，但在大量推荐数据上对大型语言模型<sup>+</sup>（LLMs）进行全面参数微调通常会导致对开放世界知识的灾难性遗忘，从而导致性能下降。即使是部分参数调整，如LoRA<sup>+</sup>，也往往无法达到满意的效果。造成这些问题的关键因素有两个。一是领域差距，即LLMs的通用知识与推荐系统的专有知识间存在鸿沟；二是训练目标不匹配，LLMs的预训练目标与推荐任务对信息的需求不一致。



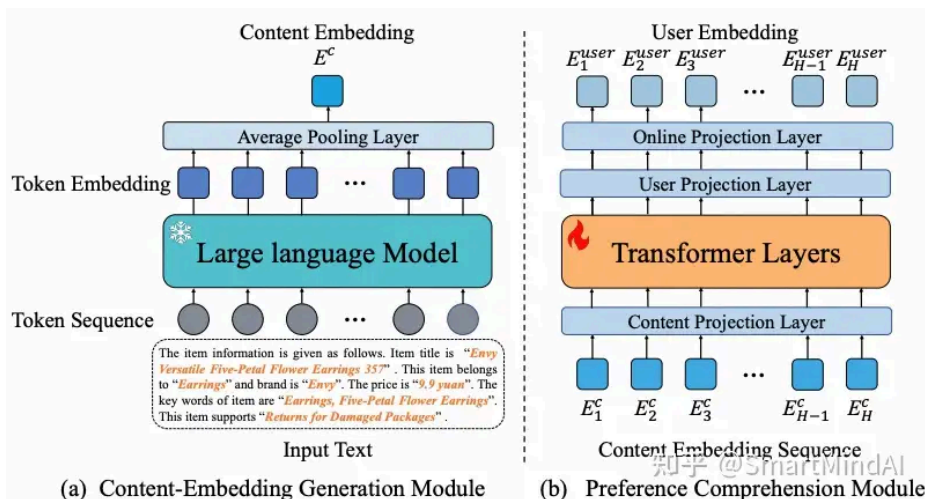
为解决这些问题，我们提出LEARN框架，其核心是LLM-to-Rec，而非从推荐到LLM的路径。我们利用LLM作为内容提取器，以推荐任务为导向进行微调，以保持对开放世界知识的利用，同时适应推荐系统的具体需求。这种方法在处理大规模数据和应对计算挑战时表现出优势，避免了全面微调可能导致的问题，从而提高了推荐系统的性能。

## Method

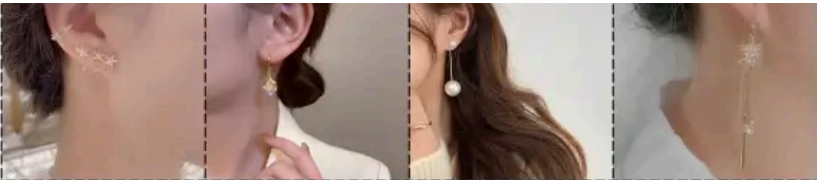












我们首先证实了利用LLM生成的物品内容嵌入具有区分性，证明了其在推荐系统中的应用潜力。随后，我们引入了LEARN框架，这是一种创新的双塔结构<sup>+</sup>设计。LEARN通过结合LLM的开放式知识和推荐任务需求，通过LLM-to-Rec途径，实现了对LLMs知识的有效利用，以提升推荐系统的准确性和适应性，特别是在处理新用户和长尾推荐问题上。

## Content Embedding of LLM

内容嵌入在当前推荐系统（RS）的研究和实践备受关注，尤其是得益于LLM的强大内容理解和推理。我们主张将LLM生成的嵌入融入RS，通过巧妙地处理文本输入，LLMs能捕捉用户兴趣，提取产品特征，并关联相似产品。然而，如何将这些文本信息转化为RS能用的形式是个挑战。



因此，我们不直接使用LLM预测的词汇，而是提议利用LLM的输出嵌入作为输入。考虑到对LLM嵌入效果的探索相对有限，尤其是针对主要基于解码<sup>+</sup>的模型，我们设计了一个基于物品对（I2I）的实证研究，使用包含100万样本的大规模数据集来评估其实际价值。

														
<b>Title:</b>	Envy Versatile Five-Petal Flower Earrings 357	Envy New Style Versatile Elegant Earrings 242	Envy Simple Double-Wear Fringe Earrings 351 Earrings	Envy Snowflake Ear Thread Earrings 203 Earrings										
<b>Category:</b>	Earrings	Earrings	Earrings	Earrings										
<b>Brand:</b>	Envy	Envy	Envy	Envy										
<b>Price:</b>	9.9 yuan	6.9 yuan	5.99 yuan	9.8 yuan										
<b>Keywords:</b>	Earrings, Five-Petal Flower Earrings	Earrings	Fringe Earrings, Earrings	Earrings, Ear Thread, Earrings										
<b>Attribute:</b>	Returns for Damaged Packages	Returns for Damaged Packages	Returns for Damaged Packages	Returns for Damaged Packages										
(a) Retrieval results of an "Earring".														
<table><tr><td></td><td>Query</td><td>Top 1</td><td>Top 2</td><td>Top 3</td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr></table>						Query	Top 1	Top 2	Top 3					
	Query	Top 1	Top 2	Top 3										
														
<b>Title:</b>	Children's Digital Building Blocks (1-5 Years Early Education and Intelligence)	Children's Educational Toy Intelligence Digital Building Blocks Assembly Puzzle Baby Brain Development Boy Girl Early Education	Infant Toddler Educational Developmental Puzzle Multifunctional Early Learning Building Blocks 1-2-3 Years Old Boys Girls Baby Toys	Educational Shape and Number Learning Box										
<b>Category:</b>	Building Block Puzzle	Building Block Puzzle	Building Block Puzzle	Other Early Education										
<b>Brand:</b>	—	Wooden Ball	Wooden Ball	—										
<b>Price:</b>	39.0 yuan	27.9 yuan	19.9 yuan	16.9 yuan										
<b>Keywords:</b>	Digital Building Blocks, Building Blocks	Toy, Children's Toy	Baby Toys, Building Blocks	Number Learning Box, Learning Box										
<b>Attribute:</b>	Returns for Damaged Packages, Flash Sale/Welfare Purchase, Pre-sale, Buy Now Pay Later	Returns for Damaged Packages, Buy Now Pay Later	Returns for Damaged Packages, Buy Now Pay Later	Returns for Damaged Packages, Buy Now Pay Later										
(b) Retrieval results of a "Building Block Puzzle".														

Model Architecture

我们的目标是建立一个推荐模型，能通过分析用户过去的购买或互动行为，理解他们的兴趣。我们构建了两个时间序列数据集

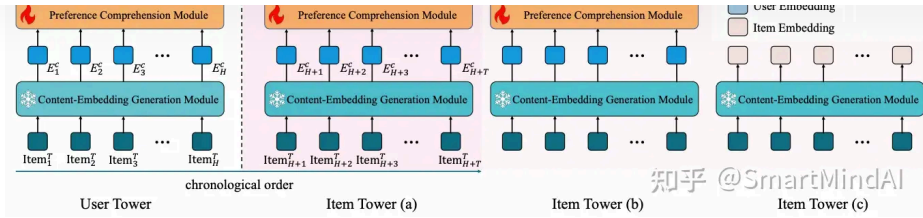
User Tower

用户塔由两个模块组成，分别是内容嵌入生成（CEG）和偏好理解（PCH）。CEG模块利用预训练LLM作为基础网络，冻结参数，避免直接使用线性分类器。它通过LLM的解码器隐藏状态生成商品的令牌嵌入，这些嵌入通过平均池化聚合为内容嵌入 $emb^c$ 。所有商品共享相同的CEG模块结构。

PCH模块则根据商品内容嵌入创建用户偏好嵌入。它利用用户历史交互序列 $U^{hist}$ 的每个商品描述生成相应内容嵌入序列 $U^{hist}$ ，然后通过Content Projection层进行维度调整，接着是一个12层自定义的Transformer，类似BERT-base，但采用因果注意力机制，以捕获商品关系和用户时间顺序偏好。

Item Tower

物品塔负责处理商品内容描述，生成针对推荐的个性化物品嵌入。我们设计了三种变体：



1. 变体1沿用用户塔结构，使用相同模型并忽略用户历史交互  $U^{hist}$ ，而是使用用户目标交互  $U^{tar}$ 。通过因果注意力增强用户和物品嵌入的关联，性能优秀，是默认选择。
2. 变体2对比，采用自我注意力，每个商品仅考虑自身内容，虽然能进行内容到协作域的映射，但性能低于使用因果注意力的变体1。
3. 变体3独创，直接将LLM的全局内容嵌入直接用于推荐，省去对齐步骤。在训练阶段，变体1处理用户目标交互，变体2和3独立处理每个商品，推理时则仅依据单个商品描述。

## Training Target of Recommendation Domain

为解决LLM的开放世界内容与推荐系统中用户和商品的协作域之间的差距，我们设定推荐任务作为训练目标。在线推荐系统通过比较用户嵌入与商品库中所有商品嵌入，找出最相关的K个商品。为了适应在线环境，我们采用自监督对比学习，通过最大化用户对感兴趣商品的相似性，同时最小化与其他商品的相似性来建模用户偏好。

LEARN框架使用InfoNCE损失进行训练。用户嵌入  $E^{user}$  源于用户历史交互序列  $\hat{U}^{hist}$ ，通过用户塔计算；商品嵌入  $E^{item}$  来自用户目标交互通过物品塔获得。正样本包括目标交互中用户和商品的嵌入对，负样本则包含不同用户间以及不同用户目标交互下的用户和商品嵌入对。InfoNCE损失函数的具体形式如下：

$$L = - \sum_{(u,i) \in Pos} \log \frac{e^{sim(E_u^{user}, E_i^{item})}}{\sum_{j=1}^N e^{sim(E_u^{user}, E_j^{item})} + \sum_{k=1, k \neq i}^K e^{sim(E_u^{user}, E_k^{item})}} + \sum_{j=1}^N \log \frac{1}{\sum_{k=1}^K e^{sim(E_u^{user}, E_k^{item})}}$$

其中  $sim()$  表示相似度计算  $Pos$  表示正样本对  $N$  是负样本的数量  $K$  是正样本的大小。

在具体操作中，用户第  $i$  的历史交互嵌入  $E_{i,j}^{user}$  来自前  $H$  个交互，目标交互嵌入  $E_{i,k}^{item}$  来自第  $T$  个目标交互。整个训练集有  $N_u$  个用户。我们生成  $H$  个用户嵌入和  $T$  个商品嵌入，形成  $N_{hist} \times N_{tar}$  个正样本对，对应于用户历史与目标交互的匹配，以及

$$N_{hist} \times (bs - 1) \times N_{tar}$$

个负样本对，这些负样本是与不同用户的历史交互及非目标交互的组合，其中  $bs$  是批大小。

对于每一个正样本，我们应用损失函数，作为全面的、密集型的损失来优化模型。这样，LEARN框架通过这样的对比学习策略，既能缩小LLM内容与实际推荐环境的差距，又能精确地学习到用户的个性化偏好。

## EXPERIMENTS

### Offline Dataset

我们基于快手电商平台的数据，历时10个月，从2022年6月至2023年4月，共收集了1200万真实用户和3100万件商品。剔除无效和异常，只保留了1200万有效用户。商品行为按照季度分类，如购买、搜索、浏览、点击，每种行为的样本数量依据80%分位数确定。我们从六个维度收集商品详情，包括标题、类别、品牌、价格、关键词和属性，特别是包含服务类型如“花呗分期”、“退换货”和“先买后付”等的属性。整个用户交互序列被划分为9个月的历史交互序列  $U^{hist}$  和1个月的目标交互序列  $U^{tar}$ 。

数据集被分为训练集和测试集，其中训练集包含1200万用户，总计2500万历史交互商品和1200万目标交互商品；测试集则有24万用户，500万历史交互商品，以及一个独立的20万商品展示集用



Offline Experimental Settings

Implementation Details

我们利用Baichuan2-7B作为中文理解的主力模型，因其强大的中文处理能力。在训练初期，我们冻结参数以保持稳定性。我们选择Adam优化器<sup>+</sup>，批大小为30。采用余弦学习率，初始学习率为5e-5，warm-up阶段设置为0.1。整个训练过程共10个epoch<sup>+</sup>。实验在8台NVIDIA V100 GPU上并行执行。在评估时，我们采用hit rate (H@50, H@100) 和recall<sup>+</sup> (R@50, R@100) 作为标准，以衡量模型性能。默认情况下，用户历史交互和目标交互的长度分别为80和40。

Offline Results

Comparison of embedding types (RQ1)

为了检验内容嵌入的有效性，我们尝试将传统推荐系统中常用的ID嵌入替换为内容嵌入。结果显示，利用LLM生成的内容嵌入显著提升了性能，H@100指标从0.0370提升至0.0701，增长率高达89.46%。

Table 1: Performance comparison of input embedding in the LEARN framework on the offline dataset

Input Embedding	H@50	R@50	H@100	R@100
ID	0.0244	0.0533	0.0370	0.0769
BERT	0.0357	0.0552	0.0576	0.0843
LLM (Ours)	<b>0.0440</b>	<b>0.0610</b>	<b>0.0701</b>	<b>0.0905</b>

为了更深入地证明这一点，我们还用BERT的CLS token替换了LLM嵌入。与BERT生成的内容相比，LLM嵌入的表现更好，达到0.0701，这表明LLM嵌入能提供更强的表达能力。这种内容嵌入源于LLM在大规模预训练中处理了大量文本，因此具备更高的信息量和更强的文本信息解读能力。

Comparison with previous Rec-to-LLM method (RQ2)

为了解决RS与LLM在协同域与开放世界域间差异的问题，我们设计了两个基准方法。首先，Baseline-v1（冻结LLM参数）利用预训练但固定的LLM，将用户历史与所有内容内容融合成单个文本，通过平均池化提取用户嵌入。每内容单独输入LLM进行计算。评估策略与之前保持一致。其次，Baseline-v2（调整所有LLM参数）允许所有参数更新，它也模仿Rec-to-LLM，将推荐数据转化为自然语言对话，通过预测下一个词优化模型。这两种基线分别代表了冻结和调整LLM的不同策略，旨在通过观察LLM在不同条件下的表现，来揭示两者在处理推荐任务中的异同。

Table 2: The quantitative metrics results of LLM-based generative recommendation model.

Model	LLM Weights	H@50	R@50	H@100	R@100
Baseline-v1	frozen	0.0069	0.0154	0.0101	0.0210
Baseline-v2	learnable	0.0134	0.0208	0.0180	0.0262
LEARN (Ours)	frozen	<b>0.0440</b>	<b>0.0610</b>	<b>0.0701</b>	<b>0.0905</b>

实验数据显示，Baseline-v1，由于其使用预训练LLM的冻结参数处理推荐数据，表现出明显的性能瓶颈，主要原因是缺乏针对推荐领域的专业知识，无法有效利用协同信息。相比之下，Baseline-v2虽然在H@100指标上有轻微提升，但在微调LLM以适应推荐任务后，其性能并未达到LEARN框架。我们观察到，这种差异源于协同领域与开放世界知识的鸿沟，尤其是在处理包含大量噪声的行业推荐数据时，开放世界知识的处理效果灾难性地下降。因此，我们的LEARN框架能够更好地平衡这两者，有效克服了这一问题，这也是其优越性的一个体现。

和Beauty) 上, 这些数据集规模小且可能无法全面反映真实工业环境, 如我们的数据集, 包含3100万商品和1200万用户, 远超样本量。其次, 由于数据质量高且规模较大, 微调LLM并未造成所谓的"灾难性遗忘", 反而LLM的全局理解和文本解析能力为传统推荐模型带来了额外价值, 通过将文本信息输入LLM并获取用户偏好描述, 提升了Rec-to-LLM方法的表现。

Comparison with SOTA models proposed for industry scenarios (RQ3)

在将LEARN应用于实际工业环境的比较中, 我们选择Amazon Book Reviews 2014作为基准, 该数据集包含800万用户对200万本书的2200万条评论。为确保公正性, 我们遵循HSTU的处理和评估标准。结果显示, LEARN在H@50和H@200上的提升分别为2.92%和9.72%, N@50和N@200分别高出10.38%和6.08%, 但在H@10和N@10指标上略逊一筹。这种差异源于模型特性: ID嵌入擅长高频互动和热门商品预测, 而内容嵌入则擅长理解和推荐非高频但内容相关物品, 尤其在大推荐列表(如前50或100)上表现突出。与SASRec和HSTU这类基于ID的模型相比, LEARN通过LLM的丰富内容表示实现了显著性能提升, 验证了其有效性。

Ablation studies of modules within the LEARN framework (RQ4)

我们构建了三个变体来验证塔式结构的必要性。实验数据显示, 变体3, 当直接使用LLMs生成的内容嵌入作为项嵌入时, 模型无法正常训练, 原因在于用户塔和项塔间架构的显著不匹配, 以及LLM内容与推荐任务所需的项嵌入在领域上的差距。相比之下, 尽管带自注意力的变体2表现稍弱, 带因果注意力的变体1通过整合前一项嵌入特征显著增强了项嵌入的表现。基于这些结果, 我们决定采用图(a)中的变体1作为LEARN框架的标准配置。

Table 4: Ablation studies of item tower in the LEARN framework on the offline dataset. Item tower v1 is adopted as the default settings.

Item Tower	H@50	R@50	H@100	R@100
Variant 3	NaN	NaN	NaN	NaN
Variant 2	0.0313	0.0488	0.0505	0.0675
Variant 1 (Ours)	0.0440	0.0610	0.0701	0.0905

LoRA通过添加可训练的新特征, 结合原始参数, 生成混合特征, 作为微调模型的输入。然而, 由于LLM的多数参数冻结, 原始特征占主导, 导致混合特征质量不高。同时, LLM的开放世界知识与LoRA的协作知识在领域上存在鸿沟, 这导致混合特征在推荐任务中的性能不理想。

Table 5: Ablation studies of the PCH module.

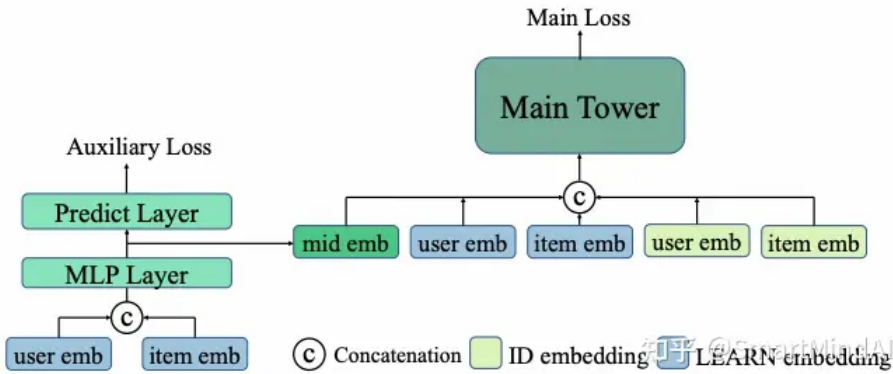
Backbone	Parameter Initialization	Training mode	Trainable parameters	H@100	R@100
LLM	Pre-train	LoRA	134M	0.0376	0.0560
			286M	0.0504	0.0709
			572M	0.0513	0.0720
Transformer	Random Init.	Finetuning	100M	0.0701	0.0905

Ranking A/B Experiments

我们实际在推荐系统的A/B测试中采用了LEARN架构, 并从2024年1月开始, 在快手应用的短视频广告推送环境中实施了这项技术的部署。

Model Structure

为了将LEARN框架生成的LLM嵌入与ID嵌入有效融合, 我们在现在线性模型\*中增加了辅助网络。这个附加部分旨在优化不同嵌入类型的融合, 以提升系统性能。辅助网络通过定义CVR损失, 引导



AUC+ Evaluation

我们在KuaiShou电商平台一个包含3亿用户交互和1千万商品的十亿级离线数据集上进行评估，数据于2023年12月7日采集，分为训练集和测试集，后者于前者后一天获取。评估指标选用UAUC（用户平均AUC）和WUAUC（加权用户平均AUC），而非仅AUC，因为它们能更全面反映用户排名的性能，体现用户体验。UAUC考虑均匀权重，尤其关注长尾用户的表现。

Table 6: The AUC results on KuaiShou App data.

Method	UAUC	WUAUC
Baseline	0.6885	0.7002
LEARN (Ours)	0.6969 (+0.84pp)	0.7078 (+0.76pp)

如表所见，我们的方法在UAUC和WUAUC上的提升分别达到了0.84%和0.76%，这归功于LEARN框架的优异泛化能力，它显著提升了对长尾用户兴趣的捕捉。为确认这一推断，我们进行了线上A/B测试以进行更深入的验证。

Online Revenue Improvement

我们在快手App的热门短视频流媒体平台进行了在线A/B测试，该应用日活跃用户达3亿。我们将20%的流量分配给我们的改进模型（LEARN）和基线模型。这项为期9天的实时测试显示，收益和点击率<sup>+</sup>（CVR）稳定且明显提升。即使只有2%的增长，在高用户基数下也体现出显著的实际收益，因为在线推荐模型的收入动辄以百万计。

Table 7: A/B test results for different user and item types. "Proportion" indicates the percentage of users (items) within each category, relative to the total number of users (items).

Level	Type	Proportion	Revenue	AUC
User	cold-start	7.16%	+1.56%	+0.17pp
	long-tail	27.54%	+5.79%	+0.68pp
	others	65.30%	+0.32%	+0.021pp
Item	cold-start	3.15%	+8.77%	+0.29pp
	long-tail	26.47%	+4.63%	+0.21pp
	others	70.38%	+0.35%	+0.01pp

我们依据用户过去的六个月交互数据（点击、印象和购买频次）分类，定义了三种类型：冷启动用户（无交互）、长尾用户（1-35次）和其他用户。研究表明，LEARN模型在所有用户群体中，特别是在冷启动和长尾用户中，展现出了显著的性能提升。此外，我们按购买历史将产品分为冷启动（无购买记录）、长尾（少于30次）和常购（超过30次）。实验结果显示，LEARN显著提升了冷启动和长尾产品的收入和AUC指标，这归功于其对购买历史稀疏产品表示的高效能力。

编辑于 2024-05-31 07:20 · IP 属地北京

推荐系统   快手   LLM



理性发言，友善互动

2 条评论

默认   最新



sin

机翻太严重了, 翻成中文反而看不懂🙄

07-15 · 浙江

回复   喜欢



温良儒雅张世豪

+1, 不懂这种机翻的有什么意义。。。我要看机翻不如直接去arxiv开个网页端翻译

08-12 · 美国

回复   喜欢

推荐阅读

百面LLM-44

提问：在LLM中选择像传统RL中 value network和policy network共享底座会有问题吗？如果有解释一下为什么？ 回答：这种做法是有问题的。但甚至在主流的TRL库中使用的就是value network和polic...

swthe...   发表于大模型面试...

2024.6 W2 商用LLM API进展简评

0、前言本期是“商用LLM API进展扫描”系列的第3篇了，前序文章：2024Q1商用LLM API进展简评 【2024.4 W1】2023.10-2024.1 商用LLM API进展简评 V3本期主要讨论从2024.4-2024.6的商用LLM...

孔某人   发表于孔某人的低...



LLM从0到1之PEFT

Alex   发表于自然语言处...



给LLM悲观或反对者们的建议

MoonCancer