# CoRAL: Collaborative Retrieval-Augmented Large Language Models Improve Long-tail Recommendation

PDF (/pdf?id=NH1XUvCJJ3)

*Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Tong Yu (/profile?id=~Tong_Yu3), Zhankui He (/profile?id=~Zhankui_He1), Jianing Wang (/profile?id=~Jianing_Wang4), Yupeng Hou (/profile?id=~Yupeng_Hou1), Julian McAuley (/profile?id=~Julian_McAuley1)*

Published: 17 May 2024, Last Modified: 17 May 2024 KDD2024 - Research Research Track, Senior Area Chairs, Area Chairs, Reviewers, Authors Revisions (/revisions?id=NH1XUvCJJ3) BibTeX 

**Keywords:** Large language models, Collaborative Filtering, Long-tail Recommendation

**Abstract:**
The long-tail recommendation is a challenging task for traditional recommender systems, due to data sparsity and data imbalance issues. The recent development of large language models (LLMs) has shown their abilities in complex reasoning, which can help to deduce users' preferences based on very few previous interactions. However, since most LLM-based systems rely on items' semantic meaning as the sole evidence for reasoning, the collaborative information of user-item interactions is neglected, which can cause the LLM's reasoning to be misaligned with task-specific collaborative information of the dataset. To further align LLMs' reasoning to task-specific user-item interaction knowledge, we introduce collaborative retrieval-augmented LLMs, \textbf{CoRAL}, which directly incorporate collaborative evidence into the prompts. Based on the retrieved user-item interactions, the LLM can analyze shared and distinct preferences among users, and summarize the patterns indicating which types of users would be attracted by certain items. The retrieved collaborative evidence prompts the LLM to align its reasoning with the user-item interaction patterns in the dataset. However, since the capacity of the input prompt is limited, finding the minimally-sufficient collaborative information for a certain recommendation task can be challenging. We propose to find the optimal interaction set through a sequential decision-making process and develop a retrieval policy learned through a reinforcement learning (RL) framework, \textbf{CoRAL}. Based on our experimental results, we find CoRAL can significantly improve LLMs' reasoning abilities on specific recommendation tasks. Our analysis also reveals that CoRAL can more efficiently explore collaborative information through reinforcement learning.

**Submission Guidelines Blind:** Yes
**Submission Guidelines Format:** Yes
**Submission Guidelines Limit:** Yes
**Submission Guidelines Authorship:** Yes
**Student Author:** Yes
**Serve As Reviewer:** juw069@ucsd.edu
**Submission Number:** 1359

Filter by reply type | Filter by author | Search keywords...

Sort: Newest First | - | = | ≡

Everyone | Program Chairs | Submission1359... | Submission1359 Area... | Submission1359... | Submission1359 Authors | Submission1359... | Submission1359... | Submission1359... | Submission1359... | Submission1359... | Submission1359... | Submission1359... | ✖

*19 / 19 replies shown*

## Paper Decision

Decision | Program Chairs | 07 May 2024, 12:38 (modified: 17 May 2024, 00:12)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers, Authors | Revisions (/revisions?id=LN28GgyTaP)

**Decision:** Accept

**Comment:**

SAC: Hady Lauw AC: Liangjie Hong The paper tackles the problem of long-tail recommendation in conventional recommender systems, emphasizing imbalance and data sparsity concerns. CoRAL, as proposed by the authors, incorporates collaborative evidence directly into LLM prompts. The authors utilize the reinforcement learning (RL) paradigm in CoRAL to demonstrate a sequential decision-making process and create a retrieval approach. The research shows notable gains in LLMs' reasoning abilities for particular recommendation tasks through experimental results.

Overall, the paper meets the bar for publication. The submission can be further improved from addressing issues raised from reviewers. The author(s) can use detailed reviews from each reviewer to improve the manuscript.

## Official Review of Submission1359 by Reviewer R6GZ

Official Review | Reviewer R6GZ | 30 Mar 2024, 10:20 (modified: 05 Apr 2024, 00:07)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Reviewer R6GZ, Authors

Revisions (/revisions?id=Y7LS8B0dK8)

**Review:**

The paper tackles the problem of long-tail recommendation in conventional recommender systems, emphasizing imbalance and data sparsity concerns. CoRAL, as proposed by the authors, incorporates collaborative evidence directly into LLM prompts. The authors utilize the reinforcement learning (RL) paradigm in CoRAL to demonstrate a sequential decision-making process and create a retrieval approach. The research shows notable gains in LLMs' reasoning abilities for particular recommendation tasks through experimental results.

Pros:

- Integrating collaborative evidence directly into the LLM enhances the model's ability to understand user preferences and align its recommendations accordingly.
- The use of reinforcement learning (RL) to develop a retrieval policy that allows CoRAL to adaptively select relevant user-item interactions, improving recommendation quality.

Cons:

- The integration of RL for developing retrieval policies adds complexity to the CoRAL framework.
- Most of baselines are quite dated.
- Limit the capacity of input prompts in incorporating collaborative information. Finding the minimally-sufficient interaction set for a recommendation task can be challenging, potentially limiting the model's ability to capture all relevant information.

**Questions:**

- Have you considered the complexity of the model? Is it much slower than other recent methods?
- In the experiments, have you considered SOTA method for Collaborative Filtering or Popularity debiasing?

**Ethics Review Flag:** No

**Ethics Review Description:** No
**Scope:** 3: The work is somewhat relevant to the Research track of KDD and is of narrow interest to a sub-community
**Novelty:** 4: Average
**Technical Quality:** 5: Above Average
**Presentation Quality:** 2: Average (it needs some effort to understand, but it should be ok after some editing)
**Reproducibility:** 2: Average (some information is missing, but that could be easily fixed in the camera-ready version)
**Reviewer Confidence:** 3: The reviewer is confident but not certain that the evaluation is correct

- = ≡

## Response to Reviewer

Official Comment
Authors (Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Yupeng Hou (/profile?id=~Yupeng_Hou1), Zhankui He (/profile?id=~Zhankui_He1), +3 more (/group/info?id=KDD.org/2024/Research_Track/Submission1359/Authors))
12 Apr 2024, 04:12 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**
**Response to Cons 1 and Question 1** During inference, our method requires the same inference time as the baseline LLM-based recommender system [e], since both our method and [a] require only one prompting and generation of the LLM. Concerning the model training, our model only requires the training of a lightweight retrieval policy model, whose parameter size is comparable to the size of traditional collaborative filtering models. We will add such discussion to our paper as suggested.

**Response to Cons 2 and Question 2** It can be arguable which methods for collaborative filtering and popularity debiasing are the exact SOTA methods suitable for the comparisons, considering that there are various recommendation settings. However, please note most of the baselines included in our paper are still regarded as competitive methods in very recent papers [b,c] in the year 2023. Besides, our approach can work as a plugin and is compatible with a wide range of collaborative filtering methods to further improve these methods. Specifically, the experimental results in Table 1 demonstrate the consistent performance of variants of CoRAL with different backbone collaborative filtering methods (i.e., DFM, WDL, AFM, and DCN), which demonstrates the effectiveness and compatibility of CoRAL on various collaborative filtering methods.

**Response to Cons 3** The limited capacity of input prompts is the limitation for all retrieval-augmented approaches [d,e,f], in which identifying the most relevant and effective information is crucial. Due to the limitation of restricted prompt length of LLMs, we develop the retrieval policy to find the minimally-sufficient interaction set, that can maximally capture the relevant information, as validated in Section 5.2.


[a] Sanner, Scott, et al. "Large language models are competitive near cold-start recommenders for language-and item-based preferences." Proceedings of the 17th ACM conference on recommender systems. 2023.

[b] Chang, Jianxin, et al. "Pepnet: Parameter and embedding personalized network for infusing with personalized prior information." Proceedings of the 29th ACM SIGKDD. 2023.

[c] Xu, Yue, et al. "Multi-factor sequential re-ranking with perception-aware diversification." Proceedings of the 29th ACM SIGKDD. 2023.

[d] Xu, Fangyuan, Weijia Shi, and Eunsol Choi. "RECOMP: Improving retrieval-augmented LMs with context compression and selective augmentation." The Twelfth ICLR. 2023.

[e] Wang, Zhiruo, et al. "Learning to filter context for retrieval-augmented generation." arXiv preprint arXiv:2311.08377 (2023).

[f] Chevalier, Alexis, et al. "Adapting language models to compress contexts." arXiv preprint arXiv:2305.14788 (2023).


- = ≡

*Replying to Response to Reviewer*

## Official Comment by Reviewer R6GZ

Official Comment Reviewer R6GZ 18 Apr 2024, 05:01

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**

Thanks for your response. I have read your rebuttal, and I would like to keep my score.

*Replying to Official Comment by Reviewer R6GZ*

## Response to Reviewer

Official Comment

Authors (Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Yupeng Hou (/profile?id=~Yupeng_Hou1), Zhankui He (/profile?id=~Zhankui_He1), +3 more (/group/info?id=KDD.org/2024/Research_Track/Submission1359/Authors))

18 Apr 2024, 23:54 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**

We greatly appreciate your time in reviewing our response. Thanks!

## Official Review of Submission1359 by Reviewer nJNJ

Official Review Reviewer nJNJ 21 Mar 2024, 21:56 (modified: 05 Apr 2024, 00:07)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Reviewer nJNJ, Authors

Revisions (/revisions?id=5y7BUDJS3o)

**Review:**

Quality: This work addresses the issue of input constraints when using LLM for recommendation systems, adopting reinforcement learning to select effective users and items. The article explores RAG+LLM enhanced recommendation systems, which is very novel, with thorough and professional experiments.

Clarity: The structure is clear, the expressions are clear, easy to follow, and it provides pseudocode.

Originality: The keywords of this work are very innovative, LLM+RAG+Long-tail for recommendation. However, the focus should be more on LLM+RAG, as the overall impression is very much like traditional reinforcement learning work. Reinforcement learning seems to act more like a sampler, with the remaining part about LLM being very concise.

Significance: This work addresses the issue of limited prompt input when using LLM for recommendation, where it's not possible to input as many items as possible for interaction, which is a very meaningful starting point. LLM+RAG is also very appealing.

Pros:

- This work addresses the constrained prompt input issue in using LLM for recommendations. RAG+LLM is highly appealing.
- The article's structure is logical and easy to follow. Pseudocode is provided.
- There are ample implementation details provided.
- The experiments on "Actor-critic Learning Curves" are highly persuasive.

Cons:

- There are errors in the details of Fig1. Moreover, collaborative signals can also be provided using conventional methods, without the necessity of reinforcement learning.
- There is very limited description and experimentation regarding LLM.

**Questions:**

Questions：

- Only allowing the model to output yes or no, does this waste the capability of RLHF, and does this approach waste the capability of LLM as well?

- Does this article provide code?
- "4.2.2 User-item Retrieval." can also be achieved using conventional top-k methods. Why introduce reinforcement learning? How to ensure the stability and effectiveness of reinforcement learning? Discuss the advantages and disadvantages of using reinforcement learning.
- What LLM was used, is it GPT-4?
- Is this work related to RAG?
- For the y Initialized Policy, does "learning from the popular items" already provide good results on its own, without the need for reinforcement learning?
- What is the duration of reinforcement learning training? Where does the computational resource consumption mainly occur after introducing reinforcement learning, and how much is it?
- Did this work specifically design a module to address the long-tail problem, or was the long-tail only adopted in the experimental setup?

**Ethics Review Flag:** No
**Ethics Review Description:** No.
**Scope:** 4: The work is relevant to the Research track of KDD and is of broad interest to the community
**Novelty:** 4: Average
**Technical Quality:** 5: Above Average
**Presentation Quality:** 2: Average (it needs some effort to understand, but it should be ok after some editing)
**Reproducibility:** 2: Average (some information is missing, but that could be easily fixed in the camera-ready version)
**Reviewer Confidence:** 3: The reviewer is confident but not certain that the evaluation is correct

## Response to Reviewer

Official Comment

Authors (Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Yupeng Hou (/profile?id=~Yupeng_Hou1), Zhankui He (/profile?id=~Zhankui_He1), +3 more (/group/info?id=KDD.org/2024/Research_Track/Submission1359/Authors))

12 Apr 2024, 04:13 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**
**To Cons 1 and Q3**

1. We apologize for the incorrect color we used for "User #1 #3" in Figure 1(a). We will correct this information in the updated version.
2. Concerning the necessity of the RL component in CoRAL, due to the limited prompt capacity in LLMs, selecting the minimally-sufficient collaborative filtering information is critical, which requires a balance between the exploration and exploitation for the retrieval policy. Since a large number of similar users and similar items exist in the datasets, simply exploiting without exploration will leverage the data of very similar users or similar items in the model training, which is data-inefficient and motivates us to consider RL-based methods. In Table 1, the comparison between CoRAL and the ablation baseline "CoRAL-random", which uses the top-k methods based on Jaccard similarity, demonstrates the necessity of considering such balance via the RL framework.
3. We are further motivated to use RL methods for the optimization of non-differentiable objectives (i.e., the final prediction accuracy). One of the technical constraints in optimizing collaborative filtering information is that actions from the retrieval policy need to be serialized by a prompt template, which makes the model not end-to-end differentiable. In addition, for the generalizability of our proposed method in any LLMs, we do not have to assume the LLM itself to be differentiable either.
4. In addition, we would like to also mention that reinforcement learning methods are not unprecedented in recommendation but on the contrary widely adopted by many recommender systems [a,b,c].

**To Cons 2 Q2,4** As described in "5.1.4 Implementation Details.", we use the GPT-4 backbone model. We also plan to release the code if the paper is accepted, to ensure that the readers get all the implementation details.

**To Q1** We would like to first clarify that RLHF techniques are not mentioned or used in our paper. Instead, the goal of reinforcement learning in our paper is to find minimally-sufficient collaborative filtering information that fits in the limited prompt length of LLMs. In addition, using LLMs for binary question-answering is also widely explored in multiple NLP tasks [d,e,f].

**To Q5** Our approach is related to RAG. However, since the collaborative filtering information is latent and needs to be derived from the interactions between users and items, the retrieval cannot be directly achieved through semantic similarity. To address this challenge, we propose the RL retrieval policy to learn to achieve long-term rewards. In addition, since collaborative filtering information is represented differently from conventional language context, we specifically design the collaborative prompt to better incorporate such collaborative filtering information in LLMs.

**To Q6** Without the RL alignment, the method by learning from the popular items cannot directly use collaborative filtering methods for retrieval augmentation. When we conducted the experiments in Table 2, by the approaches CoRAL w/ init., we observed that our model can achieve improved performance, by further updating the initial model (learned from the popular items) by reinforcement learning to align the collaborative filtering policy and LLMs. We will provide additional details in the updated version.

**To Q7** We report our learning curves in Figure 2, in which each method is trained for around 1K steps with a batch size of 16 (details in ''5.1.4 Implementation Details''). During RL training, since we are only adding a lightweight retrieval policy, the additional computational resource consumption is comparable to conventional collaborative filtering methods. After RL training, the inference cost of our method is similar to the LLM-based recommender system baseline [40] in our paper.

**To Q8** We first identify the long-tail problem in recommendation can mostly benefit from LLM's zero-shot generation ability. Besides, we specifically use the short-head knowledge in our reinforcement learning approach (detailed in Section 5.3) to address the long-tail problems, whose effectiveness is validated in Table 2, Figure 2, and Section 5.3.

[a] Afsar, et al. "Reinforcement learning based recommender systems: A survey." ACM Computing Surveys.

[b] Chen, Xiaocong, et al. "A survey of deep reinforcement learning in recommender systems: A systematic review and future directions." arXiv.

[c] Chen, Xiaocong, et al. "Deep reinforcement learning in recommender systems: A survey and new perspectives." Knowledge-Based Systems 2023.

[d] Yu, Fangyi, et al. "Exploring the effectiveness of prompt engineering for legal reasoning tasks." Findings of ACL 2023.

[e] Tripathi, Yogesh, et al. "InSaAF: Incorporating Safety through Accuracy and Fairness| Are LLMs ready for the Indian Legal Domain?." arXiv

[f] Verma, et al. "Theory of Mind abilities of Large Language Models in Human-Robot Interaction: An Illusion?." arXiv

## Official Review of Submission1359 by Reviewer Yraq

Official Review Reviewer Yraq 21 Mar 2024, 18:46 (modified: 05 Apr 2024, 00:07)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Reviewer Yraq, Authors

Revisions (/revisions?id=AyjkwaNQYQ)

**Review:**
This paper proposes CoRAL to address long-tail issue in recommendation, i.e., most of user/item have few interactions. This paper have the following advantage that I appreciate: It is interesting to see the interaction information into the prompt, and augment it with a reinforcement learning based framework to find the minimal interactions that are sufficient to learn good user interests. Experiments seem to demonstrate the effectiveness of the proposed methods.

I also have the following question for the answer:

For the RL part, the author basically adopts an offline setting where the collected data are used as the "replay buffer". However, this exclude any "exploration" and rely all on exploitation of the past. I'm wondering would the bias of the dataset be a problem?

In addition, there are some existing work on incorporating collaborative information with RecSys, e.g., [1,2], it good to see the authors discuss them in the paper.

[1] Representation learning with large language models for recommendation
[2] Collaborative large language model for recommender systems
**Questions:**
Please refer to my main review.

**Ethics Review Flag:** No
**Ethics Review Description:** N/A
**Scope:** 4: The work is relevant to the Research track of KDD and is of broad interest to the community
**Novelty:** 4: Average
**Technical Quality:** 5: Above Average
**Presentation Quality:** 2: Average (it needs some effort to understand, but it should be ok after some editing)
**Reproducibility:** 3: Excellent (it provides sufficient details, and the code and data are accessible)
**Reviewer Confidence:** 3: The reviewer is confident but not certain that the evaluation is correct

-
=
≡

## Response to Reviewer

Official Comment
Authors (Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Yupeng Hou (/profile?id=~Yupeng_Hou1), Zhankui He (/profile?id=~Zhankui_He1), +3 more (/group/info?id=KDD.org/2024/Research_Track/Submission1359/Authors))
12 Apr 2024, 04:14 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**
**Response to Question 1** We adopt the most conventional setting of RL in recommendation [a,b,c], in which some works also design and identify their exploration method in a similar way [c,d]. Understanding the impact of bias in exploration was similarly not investigated in [a,b,c,d]. We acknowledge that understanding the impact of bias in exploration is very interesting. However, this is beyond the scope of our paper, and we leave this as a future work.

**Response to Question 2** Thanks for the suggestions, we will include more discussions about these works in our related works section. In the suggested work [1], they use LLMs for semantic context extraction, which augments existing recommendation methods with user and item profiles extracted from LLMs. Whereas, we propose to directly use LLMs' reasoning abilities for recommendation. Specifically, we design the collaborative retrieval method to augment collaborative filtering information in the LLM's prompt as part of the reasoning evidence. As discussed in Line 237 - Line 240, methods like [2] require much more cost in fine-tuning the LLM and assume that abundant data exists, which is not feasible in long-tail recommendations.

[a] Xue, Wanqi, et al. "Prefrec: Recommender systems with human preferences for reinforcing long-term user engagement." Proceedings of the 29th ACM SIGKDD. 2023.

[b] Cai, Tianchi, et al. "Model-free Reinforcement Learning with Stochastic Reward Stabilization for Recommender Systems." Proceedings of the 46th International ACM SIGIR. 2023.

[c] Liu, Shuchang, et al. "Exploration and regularization of the latent action space in recommendation." Proceedings of the ACM Web Conference 2023. 2023.

[d] Zhang, Changshuo, et al. "UOEP: User-Oriented Exploration Policy for Enhancing Long-Term User Experiences in Recommender Systems." arXiv preprint arXiv:2401.09034 (2024).

-
=
≡

## Official Review of Submission1359 by Reviewer 38in

Official Review Reviewer 38in 20 Mar 2024, 23:42 (modified: 19 Apr 2024, 10:12)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Reviewer 38in, Authors

Revisions (/revisions?id=J0o3oeO2Mf)

**Review:**

This paper presents CoRAL, a novel method to combine collaborative filtering and large language models for long-tail recommendation. In CoRAL, collaborative information is used for training a retrieval model that retrieves most relevant user-item interactions for retrieval augmented generation in the long-tail recommendation task. CoRAL uses reinforcement learning to optimize for minimal-sufficient collaborative information in the prompt. On various categories of Amazon Product Reviews dataset, CoRAL outperforms state of the art long-tail recommendation methods.

Pros:

1. The paper is well-structured and the proposed methods are well-presented.
2. The proposed method does not require finetuning the LLMs, which is computationally expensive.
3. The retrieval process is optimized for reducing the number of collaborative user-item interactions, which reduces the inference cost of the LLMs.

Cons:

1. This work does not include LLM-based baselines.
2. The use of item index instead of item name in LLMs is not well-justified.
3. There are multiple components in CoRAL, but the implementation details are not mentioned.

[1] Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5). RecSys 2022.

**Questions:**

1. CoRAL uses user and item indexes instead of using item names. Why isn't item names considered in this work?
2. There are other LLM-based methods for recommendation tasks, such as P5 [1]. Although P5 focuses on several other recommendation tasks, it is intuitive to train the P5 model for the long-tail recommendation task. Are there reasons why LLM-based methods like P5 is not included as baselines?

[1] Recommendation as Language Processing (RLP): A Unified Pretrain, Personalized Prompt & Predict Paradigm (P5). RecSys 2022.

**Ethics Review Flag:** No
**Ethics Review Description:** None.
**Scope:** 4: The work is relevant to the Research track of KDD and is of broad interest to the community
**Novelty:** 5: Above Average
**Technical Quality:** 5: Above Average
**Presentation Quality:** 3: Excellent (it is a pleasure to read and easy to understand, it has good figures and insightful explanations)
**Reproducibility:** 1: Poor (no code/data are given and important information is missing)
**Reviewer Confidence:** 4: The reviewer is certain that the evaluation is correct and very familiar with the relevant literature

-
=
≡

## Response to Reviewer

Official Comment

Authors ( Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Yupeng Hou (/profile?id=~Yupeng_Hou1), Zhankui He (/profile?id=~Zhankui_He1), +3 more (/group/info?id=KDD.org/2024/Research_Track/Submission1359/Authors))

12 Apr 2024, 04:15 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**

**Response to Cons 1** To clarify, we actually include the LLM-based baseline (LLM-Language, CoRAL-random) as discussed in ''5.1.3 Baselines''.

**Response to Cons 2 and Question 1** As described in ''5.1.1 Dataset'', we actually use item names if they are available in the dataset. Only when the item names are not available, the item descriptions are used for item identification.

**Response to Cons 3** We have thoroughly discussed relevant implementation details in ''5.1.4 Implementation Details''. If anything requires further clarification, we will add more details. Besides, we also plan to release the code if the paper is accepted, to ensure that the readers understand all the implementation details.

**Response to Question 2** As discussed in Line 237 - Line 240, suggested methods like P5 [1] require much more cost and assume that abundant data exists, which is not feasible in long-tail recommendations. Whereas, we propose to directly use LLMs' reasoning abilities for recommendation and retrieve relevant collaborative filtering information as the reasoning evidence. In our case, a much smaller number of recommendation data samples is needed to train our retrieval policy model, which makes our approach more suitable in our studied long-tail recommendation setting. In addition, our paper uncovers the exceptional potentials of LLM's capability in long-tail recommendation, offering valuable insights regarding its capabilities and limitations. Since we mainly focus on the exploration of LLM's capability in long-tail recommendation, we think P5 may not be the most relevant baseline we can fairly compare with. Thanks for pointing out this line of work. We will add more discussion about this work in our related works.

## Official Review of Submission1359 by Reviewer 6q5C

Official Review Reviewer 6q5C 20 Mar 2024, 19:12 (modified: 16 Apr 2024, 21:41)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Reviewer 6q5C, Authors

Revisions (/revisions?id=tdp5PLgrM5)

**Review:**
This work proposes the CoRAL framework to enhance the inference capabilities of LLM for recommendation tasks via utilizing reinforcement learning to incorporate minimally-sufficient collaborative evidence into the prompts. The main contributions of the paper include:

- Retrieve additional user-item interactions as collaborative cues for LLM-based recommendation.
- Model the retrieval process as a continuous decision task and introduce to adopt reinforcement learning to find minimally-sufficient collaborative information.

Pros:

- The idea is easy to follow.
- The improvements are quite amazing (Table 1).

Cons:

- Some key parts are not described, making it difficult to reproduce the result.
- The technical novelty is not high.
- Prevalent LLM-based RS baselines are missing in the experiments, making it hard to judge the performance of this work.

The novelty mainly lies in modeling the sequential retrieval process a MDP and adopting Eq. 3 as the reward, while the prompt design and the retrieval policy network (DDPG) are standard techniques. Hence the technical contributions are not significant. Besides, the baselines only include one LLM-based RS [40], while representative LLM-based RS like TALLRec are not compared.

**Questions:**
- Line 316-line 321, Line 419-line 424, Line 671: p_t indicates the generation probability of the token "Yes". And the authors mention the backbone is GPT-4. I think GPT-4 is not open-source. Do you use ChatGPT's API? Can it tell the generation probability or it only generates tokens? If the generation probability is not provided via API, how can you get p_t? Or do you use one alternative LLM so you can directly access the generation probability? This part is not well described and needs more clarification.
- The baselines only include one LLM-based RS [40], while representative baselines are not compared. LLM-based RS are not new in 2024 [1]. It is better to compare some of them [2][3] or give the explanation why they cannot be compared. Particularly, the performance of traditional deep learning based RS and LLM-language is far behind CoRAL. F1 of CoRAL

is almost close to 0.9 in some cases, meaning that for most cases, CoRAL recommends correctly. I am very curious about how such amazing improvements are achieved because no additional data source (collaborative information is considered in all DL based methods) is included and all CoRAL relies is the power of LLM. If other LLM-based methods also show similar performance as LLM-Language, then LLM is not the main reason for the promising performance of CoRAL.

[1] A Survey on Large Language Models for Recommendation. Arxiv.

[2] TALLRec: An Effective and Efficient Tuning Framework to Align Large Language Model with Recommendation. RecSys 2024.

[3] GenRec: Large Language Model for Generative Recommendation. Arxiv.

**Ethics Review Flag:** No
**Ethics Review Description:** Nil.
**Scope:** 4: The work is relevant to the Research track of KDD and is of broad interest to the community
**Novelty:** 4: Average
**Technical Quality:** 5: Above Average
**Presentation Quality:** 2: Average (it needs some effort to understand, but it should be ok after some editing)
**Reproducibility:** 2: Average (some information is missing, but that could be easily fixed in the camera-ready version)
**Reviewer Confidence:** 4: The reviewer is certain that the evaluation is correct and very familiar with the relevant literature

-
=
≡

## Response to Reviewer

Official Comment

Authors ( Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Yupeng Hou (/profile?id=~Yupeng_Hou1), Zhankui He (/profile?id=~Zhankui_He1), +3 more (/group/info?id=KDD.org/2024/Research_Track/Submission1359/Authors))

12 Apr 2024, 04:16 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**

**Response to Cons 1 and Question 1 regarding implementation details** Our backbone is GPT-4 as described in line 671. We do use OpenAI API for GPT-4 calls. Extracting probability for each generation token is enabled as an option of GPT-4 API (detailed description in https://platform.openai.com/docs/api-reference/chat/create#chat-create-logprobs (https://platform.openai.com/docs/api-reference/chat/create#chat-create-logprobs)). We also plan to release the code if the paper is accepted, to ensure that the readers get all the implementation details.

**Response to Cons 3,5 and Question 2 regarding LLM-based baselines**

1. In Line 242 - Line 246, we have discussed such relevant challenges, especially in the long-tail recommendation scenario. Due to limited user-item interaction data in long-tail recommendations, fine-tuning or instruction tuning of an LLM-based recommender system (e.g., [2,3]) is infeasible.
2. We also want to clarify that the major research question in our paper is how to efficiently incorporate collaborative filtering information in LLM-based recommendation models. Since the suggested relevant works [2,3] are not directly focusing on injecting collaborative filtering information, but efficient tuning methods in LLM with recommendations instead, we may consider such works not the most direct and closest baselines to our method.
3. To understand the model performance, we want to first demonstrate that the backbone LLM without collaborative filtering information can perform significantly better than traditional recommendation methods (e.g., comparison between the ''LLM-Language'' baseline and traditional recommendation baselines on Gift Cards subset, detailed in Table 1), which validates the advantages of leveraging the power of LLMs. However, we also observe that sometimes only relying on LLMs is not enough. For example, ''LLM-Language'' achieves undesirable performance on the Prime Pantry subset, which we attribute to the lack of collaborative filtering information in LLMs. Therefore, we are motivated to explore injecting collaborative information, which is not naturally compatible with the LLM generation paradigm, via retrieval-augmented methods.

**Response to Cons 2 regarding technical novelty** We would like to emphasize our contributions in the following.

1. We recognize the LLM's power in long-tail recommendations and for the first time identify the misalignment problem of LLM in recommendations due to the lack of collaborative filtering information.

2. We develop a lightweight auxiliary model to inject collaborative filtering information, which is one of the most important kinds of information in the recommendation, in LLM-based long-tail recommendation. We consider the research question, how to incorporate collaborative filtering information in LLM-based long-tail recommendation, both novel and challenging.
3. We develop an RL framework to solve the collaborative retrieval augmentation problem, which can effectively find the minimal-sufficient collaborative filtering information that fits in the limited prompt length of LLMs.
4. We propose a novel prompt design, which enables a more compact information representation of the retrieved user-item interaction history.

In response to the reviewer's concern about the technical novelty of the RL framework, we would like to mention that similar to the previous works [d,e,f] where DDPG is widely used in recommender systems, we use DDPG as an instantiation of our proposed framework. In our paper, we did not claim the contribution that we develop a new reinforcement learning algorithm.

[a] Hou, Yupeng, et al. "Large language models are zero-shot rankers for recommender systems." European Conference on Information Retrieval. Cham: Springer Nature Switzerland, 2024.

[b] He, Zhankui, et al. "Large language models as zero-shot conversational recommenders." Proceedings of the 32nd ACM CIKM. 2023.

[c] Sanner, Scott, et al. "Large language models are competitive near cold-start recommenders for language-and item-based preferences." Proceedings of the 17th ACM conference on RecSys. 2023.

[d] Xue, Wanqi, et al. "Prefrec: Recommender systems with human preferences for reinforcing long-term user engagement." Proceedings of the 29th ACM SIGKDD. 2023.

[e] Cai, Tianchi, et al. "Model-free Reinforcement Learning with Stochastic Reward Stabilization for Recommender Systems." Proceedings of the 46th International ACM SIGIR. 2023.

[f] Liu, Shuchang, et al. "Exploration and regularization of the latent action space in recommendation." Proceedings of the ACM Web Conference 2023.

-
=
≡

*Replying to Response to Reviewer*

## Official Comment by Reviewer 6q5C

Official Comment Reviewer 6q5C 16 Apr 2024, 21:41

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**
Thanks for your clarification which addresses my main concerns. I have raised the score accordingly.

For `Response to Cons 1 and Question 1 regarding implementation details`, I think this is a key part for readers to understand how to implement the method. I suggest you adding the description to the revised paper.

-
=
≡

*Replying to Official Comment by Reviewer 6q5C*

## Response to Reviewer

Official Comment

Authors ( Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Yupeng Hou (/profile?id=~Yupeng_Hou1), Zhankui He (/profile?id=~Zhankui_He1), +3 more (/group/info?id=KDD.org/2024/Research_Track/Submission1359/Authors))

18 Apr 2024, 23:55 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**
We greatly appreciate your time in reviewing our response and providing additional valuable feedback. We will add the implementation details in the updated version of our paper. Thanks!

- = ≡

# Official Review of Submission1359 by Reviewer 5d6u

Official Review Reviewer 5d6u 18 Mar 2024, 15:21 (modified: 12 Apr 2024, 10:06)

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Reviewer 5d6u, Authors

Revisions (/revisions?id=8pNlO4iN5d)

**Review:**

**Summary**

This paper introduces a retrieval policy learnt through RLHF for retrieving optimal user-item interactions for LLM prediction for improving long-tail item prediction. They show that their model is rather superior when compared to other baselines.

**Strengths**

The paper has a good idea to train some collaborative retrieval function to augment improve the prompt engineering in LLMs for recommender systems. This work explores the topic well and shows some interesting design decisions. Overall I think the problem they bring up is quite important as often it is not feasible to feed the entire history of a user to the LLM's context let alone the histories of other users.

**Weaknesses**

The paper is a bit unclear on some sections as well as the motivation for specifically using RL to obtain the retrieval policy is not super clear. The experimental section leaves much to be desired, an analysis into why this retrieval policy works so much better, i.e how often is the policy just retrieving the closes user in embedding space, making the policy necessary? Some further analysis other than training performance would add a lot to the paper.

**Questions:**

How often is the policy just retrieving the closes user in embedding space, making the policy necessary?

How does the model perform on traditional ranking metrics ndcg/recall if this hurts performance too much on regular tasks it may not be worth it.

**Ethics Review Flag:** No
**Ethics Review Description:** n/a
**Scope:** 3: The work is somewhat relevant to the Research track of KDD and is of narrow interest to a sub-community
**Novelty:** 4: Average
**Technical Quality:** 5: Above Average
**Presentation Quality:** 2: Average (it needs some effort to understand, but it should be ok after some editing)
**Reproducibility:** 2: Average (some information is missing, but that could be easily fixed in the camera-ready version)
**Reviewer Confidence:** 3: The reviewer is confident but not certain that the evaluation is correct

- = ≡

## Response to Reviewer

Official Comment

Authors (Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Yupeng Hou (/profile?id=~Yupeng_Hou1), Zhankui He (/profile?id=~Zhankui_He1), +3 more (/group/info?id=KDD.org/2024/Research_Track/Submission1359/Authors))

12 Apr 2024, 04:17 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**

**Response to Weaknesses 1 and Question 1 regarding retrieving the closest user and motivation for RL** Our RL-based approach is motivated by efficient collaborative filtering information retrieval and injection for LLMs.

1. In Table 1, we observe that CoRAL outperforms the ablation baseline "CoRAL-random" which retrieves the closest user based on Jaccard similarity. Besides, in Table 1, when comparing a variant of our approach without exploration (i.e., CoRAL-random) to LLM-Language, although we have observed some consistent performance improvement on the AUC metric, the F1 metric is relatively inferior.
2. We observe that a large number of similar users and similar items exist in the datasets, and simply exploiting without exploration will leverage the data of very similar users or similar items in the model training. As a

result, the model training is data-inefficient.
3. We are further motivated to use RL methods for the optimization of non-differentiable objectives (i.e., the final prediction accuracy). One of the technical constraints in optimizing collaborative filtering information is that actions from the retrieval policy need to be serialized by a prompt template, which makes the model not end-to-end differentiable. In addition, for the generalizability of our proposed method in any LLMs, we do not have to assume the LLM itself to be differentiable either.

These considerations motivate us to consider the balance between exploration and exploitation via the RL framework. We will add such discussion to our paper as suggested.

**Response to Question 2 regarding metrics like NDCG/Recall in the ranking setting** We would like to mention that click-through prediction is one of the major tasks in recommendations, which has been studied by traditional recommender systems [a,b,c], conventional language model-based recommender systems [d,e], and recommendation foundation models [f,g]. The metric AUC is used in all these works and is widely used as a standard evaluation metric. Considering that these previous works [a,b,c,d,e,f,g] optimize the same metric that is used in our paper, we respectfully disagree with the reviewer's comment suggesting such works as unworthy.

[a] Hou, Xuyang, et al. "Deep Context Interest Network for Click-Through Rate Prediction." Proceedings of the 32nd ACM CIKM. 2023.

[b] Qin, Jiarui, et al. "Learning to retrieve user behaviors for click-through rate estimation." ACM Transactions on Information Systems 41.4 (2023): 1-31.

[c] Lyu, Fuyuan, et al. "Optimizing feature set for click-through rate prediction." Proceedings of the ACM Web Conference 2023. 2023.

[d] Yang, Hao, et al. "Practice on Effectively Extracting NLP Features for Click-Through Rate Prediction." Proceedings of the 32nd ACM CIKM. 2023.

[e] Wang, Dong, et al. "BERT4CTR: An Efficient Framework to Combine Pre-trained Language Model with Non-textual Features for CTR Prediction." Proceedings of the 29th ACM SIGKDD. 2023.

[f] Gong, Yuqi, et al. "An Unified Search and Recommendation Foundation Model for Cold-Start Scenario." Proceedings of the 32nd ACM CIKM. 2023.

[g] Lin, Jianghao, et al. "MAP: A Model-agnostic Pretraining Framework for Click-through Rate Prediction." Proceedings of the 29th ACM SIGKDD. 2023.

*Replying to Response to Reviewer*

## Official Comment by Reviewer 5d6u

Official Comment Reviewer 5d6u 12 Apr 2024, 10:08

Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**

Thank you to the authors for the promptly response!

I have raised my score as you have addressed my main question. I would just like this to be clarified in the paper

```
we observe that CoRAL outperforms the ablation baseline ''CoRAL-random'' which retriev
es the closest user based on Jaccard similarity.
```

As in the paper it states

```
CoRAL-random: The LLM is also augmented by collaborative information. However, the ret
rieval policy is just a
rule-based model which randomly retrieves items.
```

This is where my main confusion came from and my primary issues with the paper.

*Replying to Official Comment by Reviewer 5d6u*

## Response to Reviewer

Official Comment

Authors ( Junda Wu (/profile?id=~Junda_Wu1), Eric Chang (/profile?id=~Eric_Chang3), Yupeng Hou (/profile?id=~Yupeng_Hou1), Zhankui He (/profile?id=~Zhankui_He1), +3 more (/group/info?id=KDD.org/2024/Research_Track/Submission1359/Authors))

18 Apr 2024, 23:55 Program Chairs, Senior Area Chairs, Area Chairs, Reviewers Submitted, Authors

**Comment:**
We greatly appreciate your time in reviewing our response and providing additional valuable feedback. We will add more details and clarification about the rule-based model in CoRAL-random in the updated version of our paper. Thanks!