

MSR-Rec: Multi-Step Reasoning-Enhanced LLM for Sequential Recommendation

Tuo Wang¹, Meng Jian¹, Ge Shi^{2*}, Lifang Wu¹, Yashen Wang³

¹School of Information Science and Technology, Beijing University of Technology, Beijing, 100124, China

²School of Computer Science and Technology, Beijing Institute of Technology, Beijing, 100081, China

³Artificial Intelligence Institute of CETC, Beijing, 100049, China

{wangtuo1221, jianmeng648}@163.com, tinkersxy@gmail.com, lfwu@bjut.edu.cn, yswang.arthur@gmail.com

Abstract

Sequential recommendation has become indispensable in modern digital services. Prevalent recommendation techniques formulate the recommendation task with a language instruction fed into large language models (LLMs) to generate recommendations. However, the implicit interaction scenario of recommendation task cannot provide explicit reasoning supervision to activate LLM’s multi-step reasoning capability. Besides, the manner of reasoning for enhancing recommendation is still underexplored. Therefore, we investigate activating multi-step reasoning with users’ interactions and propose a multi-step reasoning-enhanced LLM (MSR-Rec), which tightly integrates reasoning with recommendation from designing reasoning chain to reasoning-based recommendation. A task-decomposed reasoning chain is elaborately designed to imitate users’ thinking process, seamlessly involving reasoning into recommendation. Following the reasoning chain, MSR-Rec synthesizes reasoning supervision and fine-tunes LLM to adapt for task-specific reasoning. In inference, bidirectional reasoning is implemented from user and item sides, performing a closed-loop reasoning for recommendation. Comprehensive experiments demonstrate that MSR-Rec achieves the state-of-the-art performance in both recommendation quality and reasoning interpretability, advancing the integration of reasoning and recommendation in LLM-based systems.

Introduction

Sequential recommendation becomes increasingly pivotal in modern information services in e-commerce, streaming, and social platforms (Sun et al. 2019; Zhou et al. 2020). The core lies in predicting the next interaction of users by modeling their historical behaviors as sequences of chronologically ordered items, capturing evolving preferences and temporal dynamics. Beyond conventional collaborative filtering (CF) learning (Jian et al. 2025; Wang et al. 2025b), large language models (LLMs) emerge as a transformative technique to infer users’ interests for sequential recommendation, benefiting from their extensive knowledge and sophisticated reasoning ability (Zhang et al. 2025c, 2024b). The fundamental challenge is to activate the reasoning capability of LLM to

*Corresponding author.

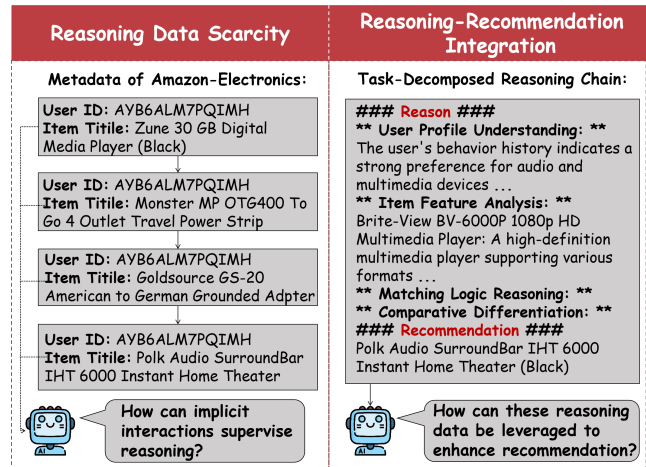


Figure 1: From behavior data to multi-step reasoning depicted decision-making process: transforming implicit user-item interactions into explicit multi-step reasoning for recommendation.

correct cognitive biases and enhance the reliability and interpretability of recommendations.

Contemporary LLM for recommendation (LLM4Rec) formulates the recommendation task with a language instruction to guide LLM for understanding the task and predicting users’ behaviors. Early works (Dai et al. 2023) employ prompt engineering to utilize the instruction-following and inherent reasoning abilities of pre-trained LLMs for interest prediction. Subsequent studies like TALLRec (Bao et al. 2023) and LLaRA (Liao et al. 2024) have explored supervised fine-tuning (SFT) with LoRA (Hu et al. 2022) to enhance domain-adapted recommendation knowledge in LLM. Despite effectiveness, the implicit single-step reasoning relies on the fundamental pattern matching paradigm of the recommendation objective, hindering the recommendation performance. Indeed, user interests reflect a diverse and complex cognitive logic, necessitating the implementation of deep reasoning analysis. Explicit multi-step reasoning can shift learning from intuitive perception to structural cognition, which shows particular promise to enhance LLMs for recommendation.

To activate explicit multi-step reasoning in LLM4Rec, sequential recommendation scenarios encounter two critical limitations: reasoning data scarcity and reasoning-recommendation integration in LLM, as illustrated in Figure 1. 1) **Reasoning data scarcity:** Reasoning supervision poses a significant obstacle to fine-tune LLMs (Guo et al. 2025; Yang et al. 2025). Unlike general domains, e.g., mathematics and coding, recommendation datasets typically consist solely of implicit interaction records as shown in the left of Figure 1, lacking explicit an multi-step reasoning chain to depict users’ decision-making process. The absence of annotated reasoning data prevents LLMs from learning behavioral reasoning logic, potentially leading to biases when recommending items. 2) **Reasoning-recommendation integration:** Reasoning generation task prioritizes generating logically coherent behavioral descriptions, whereas recommendation task focuses on effectively predicting user behavior. Their distinct objectives may lead to optimization conflicts. The manner of reasoning to enhance recommendation needs to be elaborately designed.

To address these limitations, we systematically investigate how to activate explicit multi-step reasoning for enhancing LLM4Rec and introduce MSR-Rec, a multi-step reasoning-enhanced LLM that systematically constructs and leverages reasoning chains throughout the entire recommendation pipeline. The proposed MSR-Rec designs a structural reasoning chain by simulating users’ decision-making process progressively, which works as a task-decomposed chain-of-thought (CoT) prompt for a large LLM to synthesize reasoning data. The generated reasoning data and behavior data are fed jointly to fine-tune a small LLM to activate its multi-step reasoning ability for recommendation. During the inference stage, our bidirectional reasoning strategy promotes the model to understand users’ interactions and adaptively fuses predictions based on reasoning. The main contributions of the proposed MSR-Rec are summarized as follows:

- We delve deeply into the users’ decision-making process, breaking down the recommendation task into a reasoning chain of preference recognition, attribute analysis, and logic matching to prompt LLMs for interest-related thinking.
- We propose a multi-step reasoning-enhanced LLM (MSR-Rec) to synthesize reasoning data and activate the multi-step reasoning capability of LLM in recommending items, achieving explicit reasoning supervision on LLM4Rec.
- A bidirectional reasoning-enhanced recommendation prompt enables closed-loop reasoning-augmented recommendations in LLMs.

Related Work

Sequential Recommendation

Sequential recommendation aims to predict users’ next interaction by learning their interests based on their historical item sequences (Wu et al. 2019; Xie et al. 2022). Traditional methods focus on developing sophisticated interest

encoders using various neural architectures to infer implicit collaborative relationships and extract user interests (Wang et al. 2025a; Wang 2025). GRU4Rec (Hidasi et al. 2015) employs recurrent neural networks to capture temporal dependencies within the sequence of items. Caser (Tang and Wang 2018) conceptualizes sequences as behavioral patterns and uses convolutional neural networks to learn multi-scale user representations. SASRec (Kang and McAuley 2018) introduces self-attention mechanisms to effectively model long-range behavioral dependencies. Despite these advances, the recommendation capabilities of these methods are still fundamentally limited by the core pattern matching paradigm of intuitive perception on the recommendation objective.

LLMs for Sequential Recommendation

Recently, LLMs have been introduced to improve recommendations with their general knowledge, semantic understanding and contextual reasoning ability (Zhang et al. 2024a, 2025a; Tang et al. 2025; Zhang et al. 2025b). Dai et al. (2023) designed domain-specific prompts to investigate ChatGPT’s zero-shot ranking performance, which primarily utilizes the instruction-following and inherent reasoning abilities of LLMs through prompting engineering for interest prediction. Bao et al. (2023) constructed a two-stage instruction-tuning approach to adapt Llama in few-shot settings, which fine-tunes LLM to enhance domain-adaptation to recommendation scenarios. Hou et al. (2025) also deployed a two-stage fine-tuning framework that incorporates multi-domain knowledge to support domain-specific interest learning for recommendation. Liao et al. (2024) developed hybrid prompts that integrate contextual text with collaborative embeddings of items for instruction-tuning, achieving semantic-behavior collaboration. However, these implicit reasoning methods primarily rely on the fundamental pattern matching paradigm of the recommendation objective, neglecting the diverse and complex cognitive logic behind user behaviors, which inevitably hinders the recommendation performance. Users’ complex decision-making processes require sophisticated reasoning to avoid cognitive biases in recommending items.

Reasoning-Enhanced LLMs

LLMs’ reasoning ability has been extensively leveraged to improve the performance of diverse tasks, enabling a transparent process (Fang et al. 2025). Wei et al. (2022) employed CoT prompting to empower LLMs to perform explicit reasoning steps in dealing with complex data. RecSAVER (Tsai et al. 2024) instructs LLM to generate reasoning for rating prediction through specific prompts. Wang et al. (2022) and Zhou et al. (2022) boosted performance through self-consistency sampling and least-to-most prompting, though the burden of prompt engineering remains substantial. Pre-trained models often require lengthy and meticulously designed prompts due to limited domain knowledge. To address this limitation, recent efforts explore reasoning-enhanced fine-tuning techniques to adapt LLMs to the various domains. Zelikman et al. (2022) proposed an iterative self-training framework using generated reasoning

paths, while Yao (2024) developed contrastive learning objectives to optimize reasoning steps. These advances inspire us to delve deeply into users’ behavioral logic and develop a comprehensive task-oriented reasoning chain to activate multi-step reasoning capability in LLM4Rec.

Methodology

This work systematically explores task-oriented explicit reasoning for enhancing LLM4Rec and proposes a multi-step reasoning-enhanced LLM (MSR-Rec) model for sequential recommendation. As shown in Figure 2, MSR-Rec consists of three key modules: structural reasoning data generation, reasoning-enhanced fine-tuning, and bidirectional reasoning-based recommendation. The problem definition and the details of each module are presented in this section.

Problem Definition

Let U and I denote the sets of users and items, respectively. For each user $u \in U$, their chronologically ordered interaction sequence is represented as $S_u = \{i_1, i_2, \dots, i_N\}$, where $i_n \in I$ represents the item that user u interacted with at the n -th time step, $n = 1, 2, \dots, N$. The paradigm of LLM4Rec feeds task instruction and textual content T_u of the historical sequence S_u into LLMs to predict the next item t_{N+1} from the candidate item set C_u , where $T_u = \{t_1, t_2, \dots, t_N\}$, and t_n represents the title of item i_n . Beyond intuitive pattern matching for next item prediction, the proposed MSR-Rec implements sophisticated multi-step reasoning for a progressive prediction to derive both recommendations and interpretations.

Structural Reasoning Data Generation

The recommendation scenario lacks a clear theory for decomposing reasoning steps in predicting users’ interests. To activate LLMs’ explicit multi-step reasoning capabilities, we design a comprehensive task-oriented reasoning chain and synthesize reasoning data through CoT prompting to address the reasoning data scarcity limitation in recommendation scenarios.

1) Structural Reasoning Extraction via Large LLM.

To reason along users’ behavioral logic, we decompose their decision-making processes and model the recommendation task with a multi-step progressive reasoning chain to imitate human-level interest-related thinking, which enables the model to think step-by-step before recommending items, as illustrated in Figure 3. In the instruction input, [dataset], [history] and [candidates] represent placeholders, which are filled by the target domain, the user’s behavior sequence T_u , and the candidate item set C_u , respectively. Prior to generating recommendations, the LLM follows the instruction (preference recognition, attribute analysis, and logic matching) to perform fine-grained preference-candidate alignment. At the final step, the model is instructed to summarize the rationale for the derived recommendation. This structural reasoning design ensures that the model implements a transparent recommendation process. A large LLM, DeepSeek-R1 with 671B parameters, is instructed with the reasoning chain to synthesize high-quality

reasoning paths. The task-decomposed reasoning chain and the stable large LLM guarantee the generation of reliable and insightful reasoning paths, which are fed into a small LLM as supervision signals to activate the model’s explicit reasoning capability.

2) Reasoning Data Pruning. To further polish the reasoning data, we perform data pruning to filter out incorrect or incomplete reasoning paths. For each generated reasoning path, we evaluate the prediction consistency between the recommended item in the final step 5 and the ground-truth item. The validated paths are further assessed by step completeness, ensuring all the designed reasoning steps are generated. Paths satisfying these criteria are retained as effective reasoning data. To improve generation efficiency, we set a pruning threshold τ to control the reasoning path generation. During each iteration, the generated path is immediately validated, and the reasoning process is terminated once the number of valid paths reaches τ . The filtered high-quality paths serve as annotated reasoning data to fine-tune the small LLM for recommendation.

Reasoning-Enhanced Fine-Tuning

Due to the lack of task-specific knowledge, the pre-trained LLM requires parameter fine-tuning to adapt to the complex recommendation scenario.

1) Knowledge Transfer. MSR-Rec simultaneously injects the synthesized reasoning data and behavior data as supervision into a standard pre-trained LLM, Llama3-8B, to activate explicit reasoning ability and fine-tune the model for recommendation. This approach enables effective domain knowledge transfer while enhancing the model’s ability to understand and predict user interests. MSR-Rec reformulates the reasoning data and behavior data into instruction-tuning format, comprising instruction input and output, as shown in the example in Figure 4. The instruction input contains the task scenario, user’s behavior sequence, and candidate item set, while instructing the model to provide detailed analysis and recommendations. The instruction output fills the placeholders with corresponding reasoning paths and ground truth items, demarcated by special tokens [reasoning_start] and [reasoning_end] to separate reasoning paths from recommendation decisions.

2) LLM Optimization. Formally, let x denote the instruction input tokens, y^{rea} represent the reasoning tokens between special tokens, and y^{rec} indicate recommendation tokens from [reasoning_end] to the last token, as illustrated in Figure 4. The proposed MSR-Rec measures cross-entropy losses on reasoning data and behavior data, respectively, through the negative log-likelihood of conditional language modeling in Equation (1). The reasoning loss quantifies the discrepancy between synthesized reasoning paths and the model-generated texts, while the recommendation loss constrains the predicted item to align with the ground-truth item.

$$\begin{aligned} loss_{rea} &= -\frac{1}{|y^{rea}|} \sum_{t=1}^{|y^{rea}|} \log(P_{\Theta+\Phi}(y_t^{rea} | x, y_{<t}^{rea})), \\ loss_{rec} &= -\frac{1}{|y^{rec}|} \sum_{t=1}^{|y^{rec}|} \log(P_{\Theta+\Phi}(y_t^{rec} | x, y^{rea}, y_{<t}^{rec})), \end{aligned} \quad (1)$$

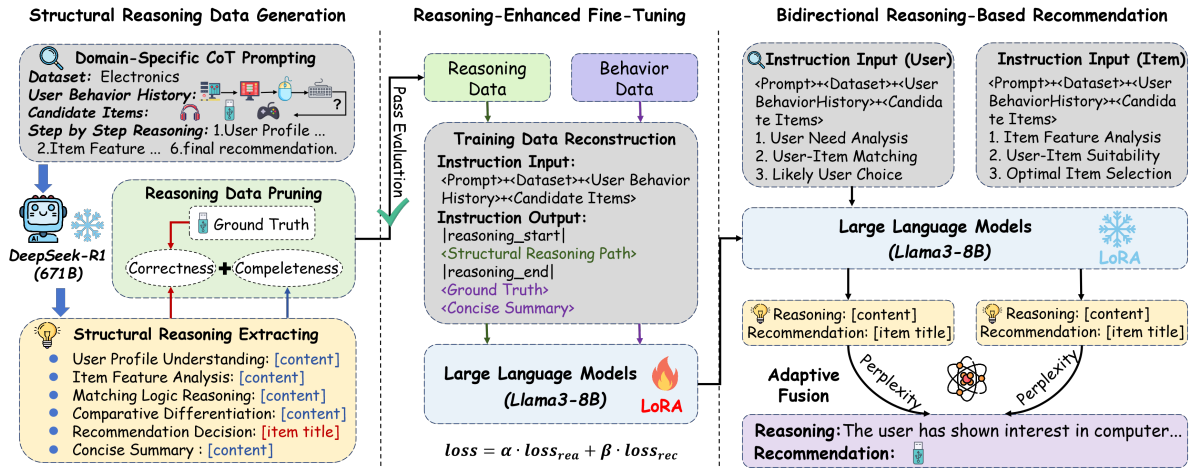


Figure 2: Framework of the proposed multi-step reasoning-enhanced LLM (MSR-Rec) model for sequential recommendation.

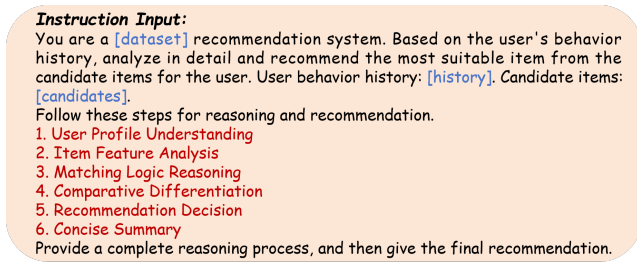


Figure 3: Reasoning chain for structural reasoning data generation.

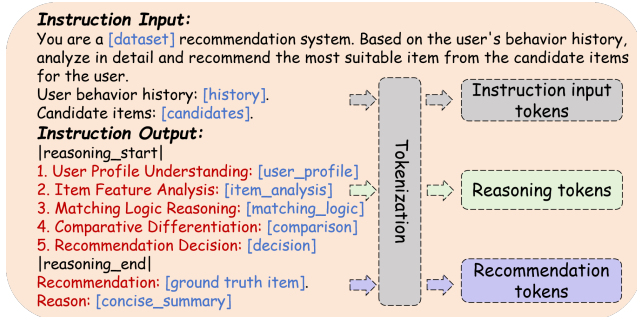


Figure 4: The format of training sample in reasoning-enhanced fine-tuning.

where y_t represents the t -th token, $y_{<t}$ represents the tokens before y_t , and Θ and Φ are respectively parameters to be fine-tuned and fixed in optimizing LLM. We employ LoRA (Hu et al. 2022) to minimize training memory cost, updating only LoRA parameters Θ while keeping LLM parameters Φ fixed. To jointly optimize reasoning and prediction for recommendation, we minimize the weighted combination of $loss_{rea}$ and $loss_{rec}$.

$$loss = \alpha \cdot loss_{rea} + \beta \cdot loss_{rec}. \quad (2)$$

Integrating progressive reasoning enhances LLM fine-tuning with behavioral knowledge and activates its multi-

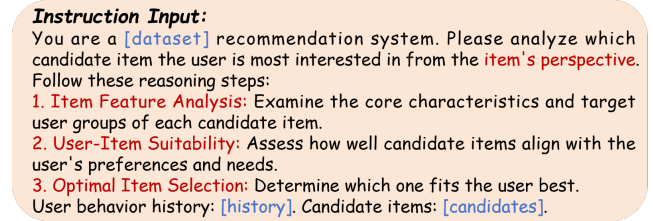
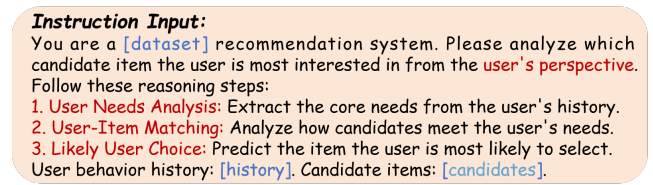


Figure 5: Reasoning chain for bidirectional reasoning-based recommendation from user and item perspectives, respectively.

step reasoning ability for recommendation, enabling the proposed MSR-Rec to generate reasoning paths for reliable and interpretable recommendations.

Bidirectional Reasoning-Based Recommendation

In addition to activating multi-step reasoning capabilities to correct cognitive biases, we also employed reasoning prompts to generate recommendations. This empowers the model to transition from expanding supervised data during the fine-tuning phase to enhancing the recommendation process during the prediction phase, thereby achieving closed-loop reasoning participation in LLM4Rec. Although lengthy and complex reasoning chains help improve reasoning quality, they are also prone to misleading LLMs, increasing the difficulty of understanding the recommendation task (Wang et al. 2024). Therefore, we design a bidirectional reasoning-based recommendation strategy for inference. Specifically, the reasoning chain is simplified to be more concise from the perspectives of both users and items to prompt the LLM to generate reasoning text and recommendations. As illus-

trated in Figure 5, the user-side prompt directs the model to summarize user preferences and match them with candidate items, whereas the item-side prompt analyzes item features and aligns them with user interests. Since the synthesized reasoning data and behavior data have been injected into the LLM at the fine-tuning stage, the model can generate coherent reasoning with understanding users’ behavior logic to aid recommendation.

To generate robust predictions, the proposed MSR-Rec adaptively fuses reasoning and recommendations from user and item perspectives. The perplexity of the generated reasoning texts are evaluated by $\text{PPL}(y) = \exp\left(-\frac{1}{|y|} \sum_{t=1}^{|y|} \log P_{\Theta+\Phi}(y_t | y_{<t})\right)$ through the well-trained LLM. When bidirectional recommendations are consistent, we select the output exhibiting lower perplexity for more fluent reasoning. In case of conflicting recommendations, higher perplexity signals imply more distinctive user preferences and item characteristics in reasoning texts, thus the corresponding output is prioritized. The adaptive fusion is implemented as

$$y = \begin{cases} \arg \min_{y_u, y_i} (\text{PPL}(y_u), \text{PPL}(y_i)), & y_u^{rec} = y_i^{rec}, \\ \arg \max_{y_u, y_i} (\text{PPL}(y_u), \text{PPL}(y_i)), & y_u^{rec} \neq y_i^{rec}, \end{cases} \quad (3)$$

where $y_u = \{y_u^{rea}, y_u^{rec}\}$ and $y_i = \{y_i^{rea}, y_i^{rec}\}$ represent reasoning-based recommendation outputs on the user and item sides, respectively. The perplexity is computationally efficient, supporting lightweight fusion of the bidirectional reasoning.

Experiments

Experimental Settings

Datasets. To evaluate MSR-Rec, we conduct experiments on three category datasets derived from the Amazon Review (McAuley et al. 2015), including Electronics, Home & Kitchen (Home), and Health & Personal Care (Health). In the datasets, implicit feedback is recorded on user-item interactions. We apply the 5-core setting, retaining users and items with at least 5 interactions to ensure data quality. The statistical details of the pre-processed datasets are summarized in Table 1. For data splitting, we adopt the leave-one-out strategy, which designates the most recent interaction in each sequence for test, the second-most recent for validation, and the remaining records for training.

Dataset	#Users	#Items	#Interactions	Density
Electronics	1,500	6,944	9,755	0.094%
Home	1,500	6,481	9,518	0.098%
Health	1,500	6,007	9,477	0.105%

Table 1: Statistics of the experimental datasets.

Evaluation Metrics. For each sequence, we pair the ground-truth item with five unobserved items to construct a candidate set for test. The target is to predict the next item from the candidate set. To evaluate the effectiveness

quantitatively, we use HitRatio@1 and ValidRatio metrics to measure recommendation quality (Liao et al. 2024). On the quality of generated structural reasoning, METEOR (Banerjee and Lavie 2005), BERTScore (Zhang et al. 2019), and GPTScore (Fu et al. 2023) metrics are compared among MSR-Rec and LLM-based baselines. A higher value of the metrics indicates better recommendation performance and reasoning quality. We report the average evaluations of the metrics for all the users.

Baselines. We compare MSR-Rec with the state-of-the-art recommendation models, including 1) Traditional sequential recommendation models: Caser (Tang and Wang 2018), GRU4Rec (Hidasi et al. 2015), and SASRec (Kang and McAuley 2018), 2) General-purpose LLM: Llama3 (Dubey et al. 2024) and Deepseek-R1 (Guo et al. 2025), 3) LLM-based recommendation models: RecSAVER (Tsai et al. 2024), TALLRec (Bao et al. 2023), and LLaRA (Liao et al. 2024).

Implementation Details. In fine-tuning LLM, we set the maximum epoch to 5 and the batch size to 32. The strengths of reasoning and recommendation losses are set as 0.7 and 0.3, respectively, by default. For the LoRA optimization, we set the rank as 8, alpha as 32. For inference, we set the temperature to 0.7 and the maximum new tokens to 1024. MSR-Rec uses DeepSeek-R1 API for reasoning data generation and Llama3-8B for the recommendation task. For traditional recommendation baselines, we truncate the maximum sequence length to 10. We implement SASRec using its official code, while Caser and GRU4Rec are adapted from the open-source library RecBole (Zhao et al. 2021). For general-purpose LLMs, we design prompts to instruct the models in generating both structural reasoning and recommendations. For LLM-based recommendation models, we follow the original papers’ hyper-parameter settings. Additionally, since RecSAVER and TALLRec are originally designed for binary recommendation, we modify the prompt to align with the task in this work.

Performance Comparison

Sequential Recommendation. We conduct experiments to compare the proposed MSR-Rec with baseline models on the Electronics, Home, and Health datasets. Table 2 provides performance comparison in terms of HitRatio@1 and ValidRatio. The results indicate that

- MSR-Rec outperforms the baselines, achieving 30.97%, 25.11%, 6.94% improvement by HitRatio over the strongest baseline on the datasets. This demonstrates that activating the multi-step reasoning capability of LLM enhances recommendations. In addition, on the sparse Electronics and Home datasets, MSR-Rec shows significant improvements, proving its advantage in inferring user interests under sparse interactions.
- MSR-Rec, TALLRec, and LLaRA surpass traditional Caser, GRU4Rec, and SASRec. This implies that LLM-based methods exceed the traditional models benefiting from semantic understanding and general knowledge.

	Electronics		Home		Health	
	HitRatio@1	ValidRatio	HitRatio@1	ValidRatio	HitRatio@1	ValidRatio
Caser	0.1427	1.0000	0.1580	1.0000	0.1840	1.0000
GRU4Rec	0.1840	1.0000	0.1680	1.0000	0.1827	1.0000
SASRec	0.1760	1.0000	0.1693	1.0000	0.1840	1.0000
Llama3 (8B)	0.1500	0.3300	0.1333	0.3093	0.1706	0.3687
DeepSeek-R1 (671B)	0.2653	0.9800	0.2760	0.9960	0.3873	0.9907
RecSAVER (671B)	0.2953	0.9833	0.2640	0.9953	0.3633	0.9860
TALLRec (7B)	0.3313	0.9800	0.2680	0.9700	0.4087	0.9787
LLaRA (7B)	0.3400	0.9987	0.2507	0.9953	0.4220	0.9933
MSR-Rec (8B)	0.4453	0.9653	0.3453	0.9367	0.4513	0.9740
Impro%	30.97	-	25.11	-	6.94	-

Table 2: Performance comparison of recommendations by HitRatio@1 and ValidRatio.

	Electronics			Home			Health		
	METEOR	BERTScore ^{F1}	GPTScore	METEOR	BERTScore ^{F1}	GPTScore	METEOR	BERTScore ^{F1}	GPTScore
Llama3 (8B)	0.0303	0.4776	22.35	0.0208	0.4613	17.15	0.0323	0.4816	36.35
RecSAVER (671B)	0.2251	0.6271	90.62	0.2194	0.6300	91.29	0.2345	0.6432	92.21
TALLRec (7B)	0.0239	0.4720	60.21	0.0181	0.4462	63.83	0.0249	0.4621	73.63
LLaRA (7B)	0.0175	0.4477	51.61	0.0100	0.4079	49.08	0.0183	0.4253	65.85
MSR-Rec (8B)	<u>0.0812</u>	<u>0.5574</u>	<u>82.35</u>	<u>0.0762</u>	<u>0.5390</u>	<u>80.99</u>	<u>0.0698</u>	<u>0.5349</u>	<u>86.52</u>

Table 3: Performance comparison of generated reasoning paths by METEOR, BERTScore, and GPTScore.

	Electronics			Home			Health		
	HitRatio@1	METEOR	GPTScore	HitRatio@1	METEOR	GPTScore	HitRatio@1	METEOR	GPTScore
MSR-Rec w/o reason	0.3613	0.0256	65.55	0.2980	0.0183	68.14	0.4393	0.0257	76.97
MSR-Rec (full)	0.4220	0.0878	85.25	0.3287	0.0775	83.18	0.4380	0.0819	88.07
MSR-Rec (user)	0.4407	0.0634	81.95	0.3267	0.0689	81.57	0.4433	0.0460	86.66
MSR-Rec (item)	<u>0.4433</u>	0.0966	86.86	<u>0.3373</u>	0.0824	83.61	0.4413	<u>0.0796</u>	<u>87.96</u>
MSR-Rec	0.4453	0.0812	82.35	0.3453	0.0762	80.99	0.4513	0.0698	86.52

Table 4: Ablation study on the role of reasoning in promoting recommendations.

- MSR-Rec, TALLRec, and LLaRA perform better than the general-purpose LLMs (Llama3 and DeepSeek-R1) in most cases, highlighting the importance of domain-specific adaptation. Furthermore, MSR-Rec’s superiority over TALLRec and LLaRA suggests that the progressive reasoning supports recommendation more than merely fine-tuning the recommendation ability.
- As a reasoning-enhanced recommendation model, RecSAVER underperforms MSR-Rec. Compared to the simple instruction in RecSAVER, the task-decomposed reasoning chain in MSR-Rec generates a more reliable reasoning path to recommend items progressively and reasoning-enhanced fine-tuning also strengthens the multi-step reasoning ability of LLM.
- MSR-Rec achieves relatively competitive ValidRatio, whereas its backbone LLM, Llama3, generates more invalid items. This also confirms that injecting reasoning supervision into the fine-tuning stage improves the recommendation capability of LLM.

Structural Reasoning. To verify the effectiveness of reasoning, we use the correct recommendation and its corresponding reasoning text extracted by DeepSeek-R1 as the reference text, and compare MSR-Rec and the LLM-based baselines with the reference by METEOR, BERTScore, and GPTScore metrics. Table 3 compares the performance of structural reasoning generated by LLMs-based models.

Although MSR-Rec performs better for recommendation, MSR-Rec’s GPTScore and BERTScore are slightly lower than RecSAVER, because small LLM are inferior in reasoning capability compared to large LLM. MSR-Rec consistently outperforms TALLRec, LLaRA, and its backbone Llama3 on the datasets, confirming its ability to provide coherent and logical reasoning interpretability. This superiority can be attributed to reasoning ability improvement from the construction of a task-decomposed reasoning chain and reasoning-enhanced fine-tuning.

Ablation Study

To verify the role of reasoning in MSR-Rec, we conduct ablation experiments on fine-tuning and inference with reasoning, comparing MSR-Rec with multiple variants. MSR-Rec w/o reason removes the synthesized reasoning data in LLM fine-tuning, relying solely on recommendation data as supervision. MSR-Rec (full) retains the entire task-decomposed reasoning chain from the LLM fine-tuning stage to the inference stage for recommendation. MSR-Rec (user/item) removes adaptive fusion and performs reasoning-based recommendation from the single side of users or items. Table 4 shows the ablation study in MSR-Rec, from which the following observations can be made.

- MSR-Rec outperforms MSR-Rec w/o reason by HitRatio@1, demonstrating that involving reasoning data as

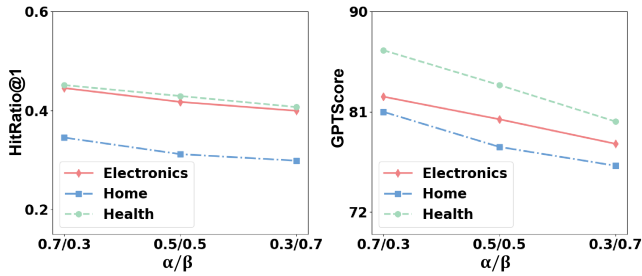


Figure 6: Recommendation performance of MSR-Rec with varying α and β .

supervision can alleviate LLM’s cognitive bias and improve recommendation ability by fine-tuning LLM. This variant also performs worse at METEOR and GPTScore, since it focuses on optimizing LLM on pattern matching for predicting items, therefore, it fails to activate reasoning ability of LLM, leading to suboptimal reasoning generation.

- MSR-Rec consistently yields better HitRatio@1 than MSR-Rec (full), but MSR-Rec (full) demonstrates superior METEOR and GPTScore over MSR-Rec. It suggests that the complex reasoning chain prompts LLM to generate a more coherent and logical reasoning path and simultaneously elevates the difficulty to understand the recommendation task. Besides, MSR-Rec exhibits slightly lower METEOR and GPTScore metrics compared with the MSR-Rec (item). This discrepancy arises because the item perspective facilitates more detailed feature-based reasoning, while our adaptive fusion in MSR-Rec inherently balances recommendation accuracy with reasoning quality.

Hyper-Parameter Analysis

We investigate the joint optimization of reasoning and recommendation with weights α and β on $loss_{rea}$ and $loss_{rec}$. Figure 6 shows HitRatio@1 and GPTScore of MSR-Rec with varying ratios of α and β on the datasets. The results indicate that as the ratio decreases, MSR-Rec’s performance consistently declines on the datasets. It implies that reasoning ability is significant to enhance LLMs for promoting recommendation, leading to more accurate reasoning and high-quality recommendations. With empirical experiments, the 0.7/0.3 ratio optimally integrates reasoning for recommendation.

Case Study

We also provide case studies with examples from the Electronics dataset to show the generated reasoning texts for recommending items as Figure 7. The reasoning outputs by MSR-Rec from item side are compared with its backbone Llama3 (8B). These cases show that MSR-Rec generates logically coherent reasoning, producing reliable and interpretable recommendations. Unfortunately, Llama3 (8B) fails to generate valid reasoning and prediction and even exhibits apparent hallucination. The comparison demonstrates that

Instruction Output: MSR-Rec (Item)

- Item Feature Analysis:** Boost ZB545 SOHO: Cellular signal booster designed for home/office use. Boston Acoustics Horizon MM226: Audio system for multimedia enhancement. Cooler Master RC-430-KWN1: PC case for building/modifying computers. Genius SW-G2.1 1250: Stereo speaker system for audio output. Laptop Battery for HP/Compaq: Replacement battery for specific laptops. AmazonBasics 4-Port USB Hub: Connectivity accessory for PCs/laptops.
- User-Item Suitability:** The user has previously purchased several computer-related products, including multiple laptops and networking equipment. This suggests they have an interest in technology and likely own or operate one or more desktop/computer setups. The laptop battery may be useful if the user owns compatible models, while other items like speakers/system upgrades would enhance overall experience.
- Optimal Item Selection:** Given the user’s history, recommending a connectivity hub (USB), stereo sound upgrade (speakers/soundbar), or PC component (case/battery replacement) makes sense as potential additions to their setup.

Among those, the **Genius SW-G2.1** stands out due to its direct relevance to audio enhancements, matching the user’s past purchase of a wireless printer server.

Instruction Output: Llama3 (8B)

```

user_profile = {"user_id": "12345", "user_name": "John Doe",
"user_behavior_history": [
  "Sony PS-LX250H Automatic Belt-Drive Turntable (Discontinued by Manufacturer)",
  "Creative Labs Muvo NX 256 MB MP3 Player",
  "Sony DVP-NS975V Up-Scaling DVD/CD/SACD Player",
  "Crosley CR49-TA Traveler 3-Speed Turntable with Stereo Speakers and Adjustable Tone Control (Tan)",
  "HP Pavilion zd8230us 17" Laptop (Intel Pentium 4 Processor 650 (H-T), 1024 MB RAM, 100 GB Hard Drive, LightScribe DVD /-RW and CD-RW Combo Drive)",
  "Zune 30 GB Digital Media Player (Black)",
  "Monster MP OT6400 BK Outlets To Go 4 Outlet Travel Power Strip (Black)",
  "Goldsource GS-20 American to German Grounded Plug Adapter",
  "Polk Audio SurroundBar IHT 6000 Instant Home Theater (Black)"]}

```

Figure 7: Case studies of multi-step reasoning and recommendation comparing MSR-Rec (item) and Llama3 (8B).

MSR-Rec successfully applies LLM’s reasoning ability to promote recommendations with both reasoning and recommendation supervision.

Conclusion

We have proposed a multi-step reasoning-enhanced LLM (MSR-Rec) model to activate LLM’s progressive reasoning ability for promoting sequential recommendation. MSR-Rec tightly integrates reasoning with recommendation at every stage, from designing a reasoning chain to fine-tuning LLM with reasoning, and to reasoning-based recommendation predictions. The superior performance demonstrates the effectiveness of the multi-step reasoning in MSR-Rec to improve recommendation and provide interpretability. The ablation study verifies the capability of reasoning-enhanced fine-tuning in alleviating LLM’s cognitive bias and adaptive bidirectional reasoning in enabling closed-loop reasoning-augmented recommendation. The case study also illustrates the advantage of MSR-Rec in generating rational reasoning paths and valid recommendation compared with its backbone LLM. In future work, we would further explore task-oriented reasoning with supervised fine-tuning and reinforcement supervision for enhancing LLM4Rec.

Acknowledgments

We would like to thank all anonymous reviewers for their valuable comments. This work was supported by the National Natural Science Foundation of China under Grant NO. 62176011, the Beijing Natural Science Foundation under Grant NO. L241053, and the Technical Field Foundation under Grant NO. 2023-JCJQ-JJ-0747.

References

- Banerjee, S.; and Lavie, A. 2005. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In *Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization*, 65–72.
- Bao, K.; Zhang, J.; Zhang, Y.; Wang, W.; Feng, F.; and He, X. 2023. Tallrec: An effective and efficient tuning framework to align large language model with recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems*, 1007–1014.
- Dai, S.; Shao, N.; Zhao, H.; Yu, W.; Si, Z.; Xu, C.; Sun, Z.; Zhang, X.; and Xu, J. 2023. Uncovering chatgpt’s capabilities in recommender systems. In *Proceedings of the 17th ACM Conference on Recommender Systems*, 1126–1132.
- Dubey, A.; Jauhri, A.; Pandey, A.; Kadian, A.; Al-Dahle, A.; Letman, A.; Mathur, A.; Schelten, A.; Yang, A.; Fan, A.; et al. 2024. The llama 3 herd of models. *arXiv e-prints*, arXiv–2407.
- Fang, Y.; Wang, W.; Zhang, Y.; Zhu, F.; Wang, Q.; Feng, F.; and He, X. 2025. Reason4Rec: Large Language Models for Recommendation with Deliberative User Preference Alignment. *arXiv preprint arXiv:2502.02061*.
- Fu, J.; Ng, S.-K.; Jiang, Z.; and Liu, P. 2023. Gptscore: Evaluate as you desire. *arXiv preprint arXiv:2302.04166*.
- Guo, D.; Yang, D.; Zhang, H.; Song, J.; Zhang, R.; Xu, R.; Zhu, Q.; Ma, S.; Wang, P.; Bi, X.; et al. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.
- Hidasi, B.; Karatzoglou, A.; Baltrunas, L.; and Tikk, D. 2015. Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939*.
- Hou, M.; Bai, C.; Wu, L.; Liu, H.; Zhang, K.; Zhang, K.; Hong, R.; and Wang, M. 2025. MoLoRec: A Generalizable and Efficient Framework for LLM-Based Recommendation. *arXiv preprint arXiv:2502.08271*.
- Hu, E. J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W.; et al. 2022. Lora: Low-rank adaptation of large language models. *ICLR*, 1(2): 3.
- Jian, M.; Wang, T.; Xia, Z.; Shi, G.; Hong, R.; and Wu, L. 2025. Geometric-Augmented Self-Distillation for Graph-Based Recommendation. *ACM Trans. Inf. Syst.*, 43(4).
- Kang, W.-C.; and McAuley, J. 2018. Self-attentive sequential recommendation. In *2018 IEEE international conference on data mining (ICDM)*, 197–206.
- Liao, J.; Li, S.; Yang, Z.; Wu, J.; Yuan, Y.; Wang, X.; and He, X. 2024. Llara: Large language-recommendation assistant. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1785–1795.
- McAuley, J.; Targett, C.; Shi, Q.; and van den Hengel, A. 2015. Image-Based Recommendations on Styles and Substitutes. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 43–52.
- Sun, F.; Liu, J.; Wu, J.; Pei, C.; Lin, X.; Ou, W.; and Jiang, P. 2019. BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 1441–1450.
- Tang, J.; Dai, S.; Shi, T.; Xu, J.; Chen, X.; Chen, W.; Jian, W.; and Jiang, Y. 2025. Think Before Recommend: Unleashing the Latent Reasoning Power for Sequential Recommendation. *arXiv preprint arXiv:2503.22675*.
- Tang, J.; and Wang, K. 2018. Personalized top-n sequential recommendation via convolutional sequence embedding. In *Proceedings of the eleventh ACM international conference on web search and data mining*, 565–573.
- Tsai, A. Y.; Kraft, A.; Jin, L.; Cai, C.; Hosseini, A.; Xu, T.; Zhang, Z.; Hong, L.; Chi, E. H.; and Yi, X. 2024. Leveraging llm reasoning enhances personalized recommender systems. *arXiv preprint arXiv:2408.00802*.
- Wang, M. 2025. SimProF: A Simple Probabilistic Framework for Unsupervised Domain Adaptation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 21153–21161.
- Wang, M.; Ren, W.; Zhang, Y.; Fan, Y.; Shi, D.; Jing, L.; and Yin, N. 2025a. Gaussian Mixture Model for Graph Domain Adaptation. *International Joint Conference on Artificial Intelligence*, 1963–1972.
- Wang, T.; Jian, M.; Xu, X.; and Wu, L. 2025b. Hierarchical Intent-Based Interest Disentanglement for Personalized Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 1–12.
- Wang, X.; Wei, J.; Schuurmans, D.; Le, Q.; Chi, E.; Narang, S.; Chowdhery, A.; and Zhou, D. 2022. Self-consistency improves chain of thought reasoning in language models. *arXiv preprint arXiv:2203.11171*.
- Wang, Y.; Tian, C.; Hu, B.; Yu, Y.; Liu, Z.; Zhang, Z.; Zhou, J.; Pang, L.; and Wang, X. 2024. Can Small Language Models be Good Reasoners for Sequential Recommendation? In *Proceedings of the ACM Web Conference 2024*, 3876–3887.
- Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Xia, F.; Chi, E.; Le, Q. V.; Zhou, D.; et al. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35: 24824–24837.
- Wu, S.; Tang, Y.; Zhu, Y.; Wang, L.; Xie, X.; and Tan, T. 2019. Session-based recommendation with graph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, 346–353.
- Xie, X.; Sun, F.; Liu, Z.; Wu, S.; Gao, J.; Zhang, J.; Ding, B.; and Cui, B. 2022. Contrastive Learning for Sequential Recommendation. In *2022 IEEE 38th International Conference on Data Engineering (ICDE)*, 1259–1273.

Yang, A.; Li, A.; Yang, B.; Zhang, B.; Hui, B.; Zheng, B.; Yu, B.; Gao, C.; Huang, C.; Lv, C.; et al. 2025. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*.

Yao, L. 2024. Large Language Models are Contrastive Reasoners. *arXiv preprint arXiv:2403.08211*.

Zelikman, E.; Wu, Y.; Mu, J.; and Goodman, N. 2022. Star: Bootstrapping reasoning with reasoning. *Advances in Neural Information Processing Systems*, 35: 15476–15488.

Zhang, C.; Wu, S.; Zhang, H.; Xu, T.; Gao, Y.; Hu, Y.; and Chen, E. 2024a. NoteLLM: A Retrievable Large Language Model for Note Recommendation. In *Companion Proceedings of the ACM Web Conference 2024*, 170–179.

Zhang, H.; Zhang, W.; Qu, H.; and Liu, J. 2025a. Enhancing human-centered dynamic scene understanding via multiple llms collaborated reasoning. *Visual Intelligence*, 3(1): 3.

Zhang, J.; Zhang, B.; Sun, W.; Lu, H.; Zhao, W. X.; Chen, Y.; and Wen, J.-R. 2025b. Slow Thinking for Sequential Recommendation. *arXiv preprint arXiv:2504.09627*.

Zhang, T.; Kishore, V.; Wu, F.; Weinberger, K. Q.; and Artzi, Y. 2019. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*.

Zhang, Y.; Bao, K.; Yan, M.; Wang, W.; Feng, F.; and He, X. 2024b. Text-like encoding of collaborative information in large language models for recommendation. *arXiv preprint arXiv:2406.03210*.

Zhang, Y.; Feng, F.; Zhang, J.; Bao, K.; Wang, Q.; and He, X. 2025c. CoLLM: Integrating Collaborative Embeddings Into Large Language Models for Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 37(5): 2329–2340.

Zhao, W. X.; Mu, S.; Hou, Y.; Lin, Z.; Chen, Y.; Pan, X.; Li, K.; Lu, Y.; Wang, H.; Tian, C.; Min, Y.; Feng, Z.; Fan, X.; Chen, X.; Wang, P.; Ji, W.; Li, Y.; Wang, X.; and Wen, J.-R. 2021. RecBole: Towards a Unified, Comprehensive and Efficient Framework for Recommendation Algorithms. 4653–4664.

Zhou, D.; Schärli, N.; Hou, L.; Wei, J.; Scales, N.; Wang, X.; Schuurmans, D.; Cui, C.; Bousquet, O.; Le, Q.; et al. 2022. Least-to-most prompting enables complex reasoning in large language models. *arXiv preprint arXiv:2205.10625*.

Zhou, K.; Wang, H.; Zhao, W. X.; Zhu, Y.; Wang, S.; Zhang, F.; Wang, Z.; and Wen, J.-R. 2020. S3-Rec: Self-Supervised Learning for Sequential Recommendation with Mutual Information Maximization. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 1893–1902.