

# MV-LLMRec: Multi-View Representation Learning with Large Language Models for Recommendation (Student Abstract)

Minkyung Song, Soyoung Park, Sungsu Lim\*

Department of Computer Science and Engineering, Chungnam National University, Daejeon, Republic of Korea  
{kyung, syoung1452}@o.cnu.ac.kr, sungsu@cnu.ac.kr

## Abstract

Traditional recommenders often fail to disentangle the motivations behind user choices. To address this, we propose **MV-LLMRec**, a framework that models interactions through three views: Structural, Intent, and Conformity. MV-LLMRec leverages LLMs to generate rich semantic representations for intent and conformity, which are refined through graph propagation and dynamically fused via an attention mechanism. We evaluate MV-LLMRec on the Amazon-Movie and Amazon-Book datasets and show that it significantly outperforms state-of-the-art baselines, validating our approach.

**Code** — <https://github.com/Smingaeng/MV-LLMRec>

## Introduction

Recent advances in recommender systems increasingly integrate Large Language Models (LLMs) for their rich semantic understanding of users and items. This line of work helps compensate for sparse interaction signals and capture latent preferences. However, most LLM-augmented recommenders still interpret user-item interactions through a monolithic, single-perspective. This approach fails to fully disentangle the composite motivations behind a user’s choice. In this context, disentangled representation learning emerges as a promising direction for modeling the inherently multifaceted nature of interactions.

Disentanglement promises gains in both interpretability and accuracy. However, while previous work has modeled user intent, the conformity factor - how much a user follows popularity (Zhao et al. 2023) - has been largely treated with structural signals alone and is rarely captured with the semantic expressiveness of LLMs. Moreover, effective methods to learn from these multiple views and to fuse them dynamically into a single, comprehensive preference signal for recommendation still remain notably elusive.

To address these gaps, we propose **Multi-View Representation Learning with Large Language Models for Recommendation (MV-LLMRec)**. MV-LLMRec is a novel framework that models each interaction through three independent views—structural, intent, and conformity.

\*Corresponding Author

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

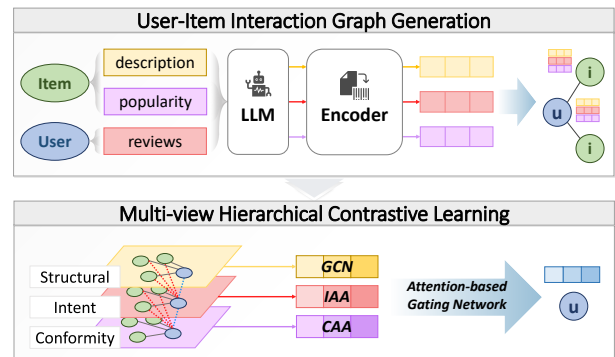


Figure 1: Framework of MV-LLMRec.

These view-specific representations are then dynamically fused to produce a holistic preference model.

## Methodology

As illustrated in Figure 1, MV-LLMRec performs multi-view representation learning, and for effective training we also apply multi-task optimization. Building on IRL-Rec (Wang et al. 2025), our framework adopts its LLM prompting and embedding scheme while extending them with three view-specific semantics and a view-aware aggregation. We prompt an LLM to generate view-specific semantics, propagate them on the user-item graph to learn per-view representations, and fuse the three views via an attention-based gating network to obtain the final user and item embeddings. We jointly optimize complementary objectives, including hierarchical contrastive learning.

## Multi-View Representation Learning

**LLM-based Semantic Information Generation.** To inject semantics that are difficult to infer from graph structure alone, we use LLM to generate three view-specific texts: a profile that encapsulates a user’s preferences and summarizes the types of items that appeal to them; an intent that captures user needs in a specific context and item’s context; and a conformity signal indicating the propensity to follow popularity of a user and the social influence of an item. These texts are encoded by a pretrained encoder to obtain initial view-level embeddings.

Dataset	Model	Metrics			
		R@5	R@10	N@5	N@10
Amazon Movie	LightGCN	0.0250	0.0430	0.0147	0.0204
	RLMRec-Con	<u>0.0262</u>	0.0425	<u>0.0155</u>	<u>0.0208</u>
	RLMRec-Gen	0.0250	<u>0.0435</u>	0.0146	0.0205
	IRLLRec	0.0064	0.0101	0.0042	0.0053
	<b>Ours</b>	<b>0.0288</b>	<b>0.0490</b>	<b>0.0169</b>	<b>0.0235</b>
	$\Delta$ Improv. (%)	+ 9.92	+ 12.64	+ 9.03	+ 12.98
Amazon Book	LightGCN	0.0561	0.0902	0.0578	0.0694
	RLMRec-Con	<u>0.0633</u>	0.0981	<u>0.0621</u>	0.0742
	RLMRec-Gen	0.0601	0.0937	0.0615	0.0727
	IRLLRec	0.0620	<u>0.0991</u>	0.0617	<u>0.0746</u>
	<b>Ours</b>	<b>0.0661</b>	<b>0.1037</b>	<b>0.0653</b>	<b>0.0784</b>
	$\Delta$ Improv. (%)	+ 4.42	+ 4.64	+ 5.15	+ 5.09

Table 1: Recommendation performance. The best is highlighted in **bold**, the second best is underlined.

**Multi-View Embedding Propagation.** We propagate information on the user–item graph and learn embeddings while respecting the characteristics of each view.

- **Structural View.** Adopts standard LightGCN (He et al. 2020) message passing to capture pure collaborative filtering—the structural relationships between users and items. This view yields *GCN embeddings*.
- **Intent View.** Uses an LLM-derived adaptive mask to reweight neighbor messages and yields *intent-aware aggregation (IAA)* embeddings:  $\mathbf{E}_{IAA}^{(l)} = \mathbf{G}_{int}^{(l)} \mathbf{E}^{(l)}$ , where  $\mathbf{G}_{int}^{(l)} = (\mathbf{D}_{int}^{(l)})^{-1} \tilde{\mathbf{A}}^{(l)}$ ; the final  $\mathbf{E}_{IAA} = \sum_{l=0}^{L-1} \mathbf{E}_{IAA}^{(l)}$ .
- **Conformity View.** Captures popularity and conformity propensity. Analogous to the Intent View, we aggregate with the conformity-weighted operator  $\mathbf{G}_{conf}^{(l)}$  to obtain *conformity-aware aggregation (CAA)* embeddings.

**Adaptive Fusion.** To combine the resulting view-specific embeddings, we employ an attention-based gating network that computes specific importance over the three views and outputs a weighted sum as the final embedding.

## Multi-Task Optimization

We design multi-objective losses so that each component learns its target behavior effectively.

**Main Recommendation Task.** We optimize a Bayesian Personalized Ranking (BPR) loss by training the predicted score of an observed positive sample to exceed that of an unobserved negative sample.

**Knowledge Distillation.** View-specific graph embeddings learned via propagation are aligned with their corresponding LLM-based semantic embeddings using an InfoNCE loss.

**Hierarchical Contrastive Learning.** We encourage each view to learn complementary, task-relevant information while enforcing consistent representations for the same node across views:

$$\mathcal{L}_{HCL} = -\frac{\lambda}{2|B_u|} \sum_{i \in B_u} \sum_{\substack{v, w \in \mathcal{V} \\ v \neq w}} \log \frac{\exp(\langle \hat{\mathbf{e}}_i^v, \hat{\mathbf{e}}_i^w \rangle / \tau)}{\sum_{j \in B_u} \exp(\langle \hat{\mathbf{e}}_i^v, \hat{\mathbf{e}}_j^w \rangle / \tau)}.$$

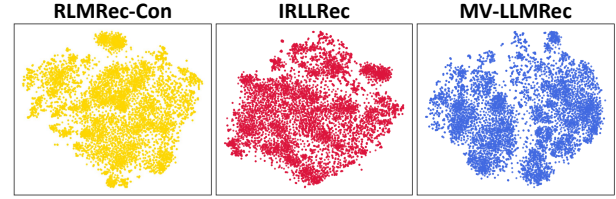


Figure 2: Item Embeddings on Amazon-Movie.

where  $B_u$  denotes the user mini-batch,  $\mathcal{V}$  represents the set of views, and both  $\tau$  and  $\lambda$  are positive.

## Experiments

We evaluate the proposed MV-LLMRec and examine its effectiveness with the following research questions:

- **RQ1:** Does MV-LLMRec improve recommendation accuracy over state-of-the-art methods?
- **RQ2:** Does MV-LLMRec learn higher-quality item embeddings than existing methods?

**Setup.** We evaluate on Amazon-Movie and Amazon-Book and compare our framework against LightGCN-based RLMRec (Ren et al. 2024) and IRLLRec. The Amazon-Book split from RLMRec is used, and Amazon-Movie is constructed under the same protocol. We prompt Qwen2.5-14B-Instruct to generate view-specific semantics and embed them using OpenAI embedding models.

**Accuracy (RQ1).** As shown in Table 1, the proposed MV-LLMRec consistently outperforms all compared methods on both Amazon-Movie and Amazon-Book. In particular, averaged over the four evaluation metrics, MV-LLMRec achieves relative gains of 11.1% on Amazon-Movie and 4.8% on Amazon-Book. These results suggest that disentangling embeddings with an LLM and then fusing multi-view information can improve recommendation accuracy.

**Representation (RQ2).** For a qualitative assessment of the learned item embeddings, we visualize them in a two-dimensional space using t-SNE, as shown in Figure 2. MV-LLMRec exhibits overall clearer inter-cluster separation and intra-cluster compactness than the baselines. This structure is desirable in that semantically similar items form dense groups while distinct groups are well separated, suggesting that our model learns less noisy, more discriminative representations. This qualitative evidence aligns closely with the performance gains in the quantitative results.

## Conclusions

We extend LLM-augmented recommenders by disentangling structural, intent, and conformity signals. MV-LLMRec prompts an LLM for view-specific semantics and propagates each view on the user–item graph under a hierarchical contrastive objective for cross-view consistency. It consistently outperforms strong baselines and yields cleaner embeddings, confirming that disentangling multi-faceted signals provides a more faithful preference representation.

## Acknowledgments

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2025-25435830).

## References

- He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; and Wang, M. 2020. LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation. In *SIGIR*, 639–648.
- Ren, X.; Wei, W.; Xia, L.; Su, L.; Cheng, S.; Wang, J.; Yin, D.; and Huang, C. 2024. Representation Learning with Large Language Models for Recommendation. In *WWW*, 3464–3475.
- Wang, Y.; Sang, L.; Zhang, Y.; and Zhang, Y. 2025. Intent Representation Learning with Large Language Model for Recommendation. In *SIGIR*, 1870–1879.
- Zhao, W.; Tang, D.; Chen, X.; Lv, D.; Ou, D.; Li, B.; Jiang, P.; and Gai, K. 2023. Disentangled Causal Embedding with Contrastive Learning for Recommender System. In *WWW*, 406–410.