

Knowledge-Enhanced Explainable Hypergraph Convolution Network for Medication Recommendation

Zihan Zhang¹, Hongzhi Liu^{1*}, Xiaoshuang Guo¹, Tianqi Sun¹, Zhonghai Wu^{2†}

¹School of Software and Microelectronics, Peking University, Beijing, China

²National Engineering Center of Software Engineering, Peking University, Beijing, China

lusia@stu.pku.edu.cn, liuhz@pku.edu.cn, gxs@stu.pku.edu.cn, suntq@stu.pku.edu.cn, wuzh@pku.edu.cn

Abstract

Medication recommendation systems aim to provide personalized and safe medication options based on individual patient records. However, existing approaches often face challenges related to inadequate modeling of complex relationships within Electronic Health Records (EHRs), data sparsity, and a lack of explainability for recommendations. In this paper, we present a Knowledge-enhanced Explainable HyperGraph Convolution Network (KEHGCN) that constructs a hierarchical hypergraph structure to capture the multi-level relationships within EHR data. By incorporating external knowledge graphs, our approach introduces additional positive relations that help alleviate the impact of data sparsity on model learning. Furthermore, by performing generalized metapath construction and selection on the knowledge graph, our approach achieves effective knowledge filtering and extracts semantically meaningful metapaths, thereby further enhancing the explainability of the recommendation results. We also explicitly introduce negative relations present in the domain knowledge to improve the safety of medication recommendation. Extensive experiments on different hospital departments of MIMIC-III and MIMIC-IV datasets demonstrate that KEHGCN outperforms other state-of-the-art baselines.

Introduction

Medication recommendation aims to provide accurate and safe medications to a patient at the current visit. Instead of relying on uniform treatment protocols, these systems are designed to consider the distinct medical histories and current conditions of individual patients, thereby enabling more personalized and effective medication recommendations (Liu et al. 2024a; Kim et al. 2025).

However, unlike conventional recommendation tasks where user-item interactions are frequently updated (e.g., movies, shopping) (Sun et al. 2024), a key challenge in medication recommendation is the sparsity of Electronic Health Records (EHRs) (Ma et al. 2022). For instance, once a patient is prescribed a certain medication and achieves a satisfactory outcome, further interactions are rarely observed. In addition, most existing methods fail to fully exploit the

potential of EHRs themselves. Specifically, conventional approaches often model only simple, pairwise associations between clinical entities and ignore the complex, higher-order dependencies inherent in EHRs. These limitations reduce the ability to accurately model patient and medication preferences.

To alleviate the data sparsity issue, previous works have leveraged a variety of auxiliary information, including medication molecular structures (Yang et al. 2023; Chen et al. 2023; Liang et al. 2025), ontology coding systems (Yao et al. 2023), and even knowledge graphs (Liu et al. 2024b; Mishra and Shridevi 2024). While these approaches help mitigate sparse interactions by introducing heterogeneous side information, they typically focus on positive associations and rarely exploit negative relations. Furthermore, most of these methods offer limited explainability. Therefore, there is a pressing need for frameworks that not only achieve robust predictive performance, but also provide transparent and explainable medication recommendations.

In this paper, we propose a novel Knowledge-enhanced Explainable HyperGraph Convolution Network (KEHGCN) for medication recommendation. Specifically, we incorporate relations from external knowledge graphs and construct generalized metapaths, which not only alleviate the challenges of model learning caused by data sparsity in EHRs but also provide explainability for the recommendation results. To more effectively learn the EHRs themselves, we design a hierarchical hypergraph structure to capture the multi-level relationships among medical entities. We also incorporate negative relations into the loss function design to enhance safety. Extensive experiments verify the effectiveness of KEHGCN over state-of-the-art baselines.

The contributions of our work are summarized as follows:

- We propose a novel **K**nowledge-**E**xplainable **H**yper**G**raph **C**onvolution **N**etwork (KEHGCN) for medication recommendation. By leveraging external knowledge graphs to alleviate data sparsity and further modeling knowledge through generalized metapaths, our approach enhances the explainability of medication recommendation.
- We design a hierarchical hypergraph structure and knowledge-driven hypergraph convolution framework to better capture complex relations and enhance representa-

*Corresponding author: liuhz@pku.edu.cn

†Corresponding author: wuzh@pku.edu.cn

tion learning in EHR data, leading to more accurate medication recommendation.

- Our approach proposes utilizing negative relationships from knowledge graphs, such as contraindications between diseases and medications, to explicitly guide and improve the safety of medication recommendation.
- Extensive experiments on departmental partitioned MIMIC-III and MIMIC-IV datasets demonstrate that KEHGCN consistently outperforms state-of-the-art baselines on medication recommendation tasks, achieving superior accuracy, enhanced explainability, and improved safety.

Related Work

Medication Recommendation

Recent approaches have witnessed rapid progress in medication recommendation systems, evolving from early instance-based (Zhang et al. 2017) and multi-label classification (Choi et al. 2016) approaches focused on single-visit data to sequential models that exploit patient history for personalized recommendations (Yang et al. 2021b). Furthermore, leveraging Graph Neural Networks (GNNs) significantly enhances the ability to model structural relationships and medication interactions within complex medical data (Gao et al. 2022; Shang et al. 2019). Despite these advances, existing efforts mainly rely on conventional graph structures that are limited in representing higher-order dependencies (Jin et al. 2020). Moreover, these approaches typically focus on solely modeling the EHR data, overlooking the potential of integrating external knowledge to enhance representation learning.

Knowledge-Enhanced Recommendation

Knowledge graphs offer valuable information by encoding entities and their interactions, which can be utilized to address challenges like data sparsity and limited semantic understanding in recommendation systems (Wang et al. 2021; Liu, Zhu, and Wu 2024). In the field of medication recommendation, however, existing knowledge-enhanced approaches are mainly divided into two categories. Embedding-based methods utilize entity and relation embeddings to enrich representations (Gong et al. 2021). GNN-based methods leverage the rich structural information in knowledge graphs through graph neural networks (Saadat et al. 2024). While these strategies effectively improve predictive performance, they generally offer limited explainability, which is a crucial requirement in high-stakes scenarios such as medication recommendation.

Moreover, most current methods primarily focus on positive relations in knowledge graphs, while negative relations are less systematically incorporated. Although some studies include Drug-Drug Interactions (DDIs) (Ren et al. 2022; Kim et al. 2024), other clinically relevant negative relations, such as the contraindications between diseases and medications, are often overlooked. In addition, though the utilization of metapaths for explainability has been explored in other recommendation domains (Lai et al. 2024; Yang et al. 2025), such explicit modeling and explanation have not been

widely adopted in the domain of medication recommendation systems.

Preliminaries and Notations

To ensure effective and safe medication recommendations, it is essential to integrate information from EHR data with external domain knowledge.

Electronic Health Records (EHRs). EHRs are represented as a sequence of visits $V = \{v_1, v_2, \dots, v_N\}$, where each visit $v_t = \{d_t, p_t, m_t\}$. Here, $d_t \in \{0, 1\}^{|\mathcal{D}|}$, $p_t \in \{0, 1\}^{|\mathcal{P}|}$, and $m_t \in \{0, 1\}^{|\mathcal{M}|}$ are binary vectors indicating the presence or absence of each disease, procedure, and medication in the corresponding sets \mathcal{D} , \mathcal{P} , and \mathcal{M} at visit t . This formulation provides a unified, structured representation of the patient’s clinical history.

Medical Domain Knowledge. We extract a knowledge graph $G_u = (E, R)$, where E is the union of heterogeneous biomedical entities and R is the relationship between those entities. Each edge in the knowledge graph is represented as a triplet (h, r, t) , where $h, t \in E$ and $r \in R$. To encode disease-medication safety constraints, we extract a contraindication matrix $\mathbf{A}_{contra} \in \mathbb{R}^{|\mathcal{D}| \times |\mathcal{M}|}$ from the knowledge graph, where $\mathbf{A}_{contra}[i, j] = 1$ indicates that medication $m_j \in \mathcal{M}$ is contraindicated for disease $d_i \in \mathcal{D}$, and $\mathbf{A}_{contra}[i, j] = 0$ otherwise. Moreover, in order to capture potential DDI interactions, we incorporate the adjacency matrix $\mathbf{A}_{ddi} \in \mathbb{R}^{|\mathcal{M}| \times |\mathcal{M}|}$, where $\mathbf{A}_{ddi}[i, j] = 1$ if medications $m_i, m_j \in \mathcal{M}$ are known to have an adverse interaction.

Problem Definition. The model integrates the EHR data as well as external domain knowledge, including the knowledge graph G_u and DDI relationships as input. The goal of the medication recommendation task is to learn a function $f(\cdot)$, which can predict the medication set $\hat{y}_{u_i}^t = f(d_t^{u_i}, p_t^{u_i}, \dots, V_{t-1}^{u_i})$ for each patient u_i at the t -th visit, where $d_t^{u_i}$ denotes the current diagnoses, $p_t^{u_i}$ denotes the current procedures, $V_{t-1}^{u_i}$ is the patient’s prior visit records, and $\hat{y}_{u_i}^t$ is a multi-label output where each medication takes a value of 0 or 1. For clarity, we omit the patient index u_i in the subsequent discussion.

Method

As shown in Figure 1, the proposed KEHGCN model is composed of three main modules, including hierarchical hypergraph construction and generalized metapath extraction, knowledge-driven hypergraph convolution, and multi-view medication prediction. The hierarchical hypergraph structure and the extracted generalized metapaths are constructed specifically to capture the multi-level and semantic relationships within the EHR data and knowledge graphs. The knowledge-driven hypergraph convolution module aims to enhance entity representation learning by capturing both the co-occurrence and positive semantics. The multi-view medication prediction module integrates both metapath-based and instance-based perspectives to generate the final recommendations.

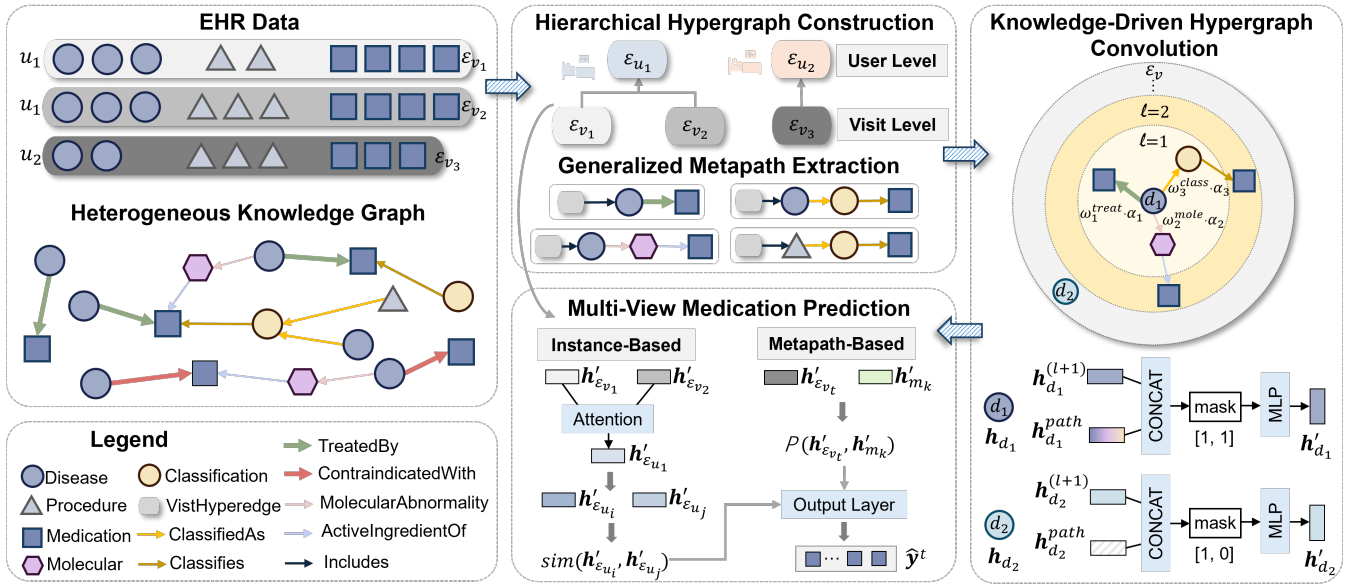


Figure 1: An overview of the proposed KEHGCN method.

Hierarchical Hypergraph Construction and Generalized Metapath Extraction

To comprehensively capture both the complex interplay of medical entities within individual clinical visits and the longitudinal care trajectory of each patient, we design a hierarchical hypergraph structure. For each patient, we construct visit-level hyperedges by unifying all diseases, procedures, and medications occurring during a single visit, and a patient-level hyperedge by aggregating all visit hyperedges across the patient’s medical history. Formally,

$$\mathcal{E}_{v_i} = \mathcal{D}_{v_i} \cup \mathcal{P}_{v_i} \cup \mathcal{M}_{v_i}, \quad \mathcal{E}_{u_j} = \{\mathcal{E}_{v_m}, \mathcal{E}_{v_{m+1}}, \dots, \mathcal{E}_{v_n}\}, \quad (1)$$

where \mathcal{E}_{v_i} denotes the visit hyperedge for the i -th visit (with disease set \mathcal{D}_{v_i} , procedure set \mathcal{P}_{v_i} , and medication set \mathcal{M}_{v_i}), and \mathcal{E}_{u_j} denotes the patient hyperedge that aggregates the visit hyperedges from the first (m) to the last (n) visit of patient u_j . Note that for the current query visit, the visit hyperedge \mathcal{E}_{v_t} is constructed by combining the current t -th diseases and procedures, and medications from the previous visit.

To filter knowledge and eliminate noise, we construct and select clinically positive metapaths from the external knowledge graph G_u provided by the Unified Medical Language System (UMLS) (Bodenreider 2004). This process leverages rich domain knowledge and enables our model to focus on semantic relations. Specifically, we define metapath Φ as a sequence of alternating entities and relations. For instance, metapath (*Disease, TreatedBy, Medication*) captures the direct treat relationship. Metapath (*Disease, MolecularAbnormality, Molecular, ActiveIngredientOf, Medication*) (Le 2025) links a disease to a medication through a shared molecular target. Metapath (*Disease, ClassifiedAs, Classification, Classifies, Medication*) (Maya et al. 2017) connects diseases and medications via clinical classifications.

In addition to the above conventional metapaths connecting diseases or procedures with medications through clinical relationships, we introduce a generalized metapath that includes visit hyperedge as a medical entity, e.g., (*VisitHyperedge, Includes, Disease, TreatedBy, Medication*). The augmented generalized metapath enhances the expressive power of conventional metapath by including visit-level co-occurrences and relationships, so that enables finer-grained medical entity modeling within the same visit.

We initialize each type of entity node with a d -dimensional embedding.

$$h_{e_i} = W_e e_i \in \mathbb{R}^d, \quad e \in \{d, p, m, o, c\}, \quad (2)$$

where each e_i represents the trainable vector for corresponding node, including disease node (d_i), procedure node (p_i), medication node (m_i), molecular structure node (o_i), and classification node (c_i), and $W_e \in \mathbb{R}^{|E_i| \times d}$ is the embedding matrix to learn.

Knowledge-Driven Hypergraph Convolution

To leverage both co-occurrence relationships and semantic connections among medical entities, we design and implement a knowledge-driven hypergraph convolution method. It involves updating the representations of entity nodes through hypergraph convolution to capture co-occurrence patterns, and applying metapath convolution to integrate semantic information from various metapaths. By concatenating the updated embeddings with a masking mechanism, our method creates comprehensive representations for each entity and visit hyperedge.

Hypergraph Convolution for Co-occurrence Relations.

In order to capture the co-occurrence relationships within the same visit, we utilize hypergraph convolution to update

the representations of the entities.

$$\mathbf{h}_{e_i}^{(l+1)} = \sigma \left(\frac{1}{|\mathcal{N}(\mathcal{E}_{v_j})|} \sum_{e_j \in \mathcal{N}(\mathcal{E}_{v_j})} \mathbf{h}_{e_j}^{(l)} \mathbf{W}_{e'} \right), \quad e \in \{d, p, m\}, \quad (3)$$

where $\mathbf{h}_{e_i}^{(l+1)}$ denotes the updated embedding of node e_i for at layer $l + 1$. $\mathcal{N}(\mathcal{E}_{v_j})$ denotes the set of nodes connected to the j -th visit hyperedge \mathcal{E}_{v_j} . $\mathbf{W}_{e'}$ represents the learnable weight matrix shared among different node types, and σ is the activation function sigmoid.

To effectively represent each visit hyperedge, we update the visit hyperedge by aggregating the nodes connected to it. This allows us to capture the co-occurrence relationships among the disease, procedure, and medication nodes for a specific visit. The updated visit hyperedge representation is derived as follows:

$$\mathbf{h}_{\mathcal{E}_{v_j}}^{(l+1)} = \sum_{e_j \in \mathcal{N}(\mathcal{E}_{v_j})} \alpha_{e_j} \mathbf{h}_{e_j}^{(l+1)}, \quad (4)$$

where $\mathbf{h}_{\mathcal{E}_{v_j}}^{(l+1)}$ denotes the embedding of visit hyperedge \mathcal{E}_{v_j} at layer $l + 1$, and α_{e_j} represents the self-attention weight assigned to node e_j .

By leveraging the connections among diseases, procedures, and medications within each visit, our approach can enhance the representation of individual entities and capture the visit-level prescribing logic to each clinical visit.

Metapath Convolution for Semantic Knowledge. As shown in Figure 2, the metapath convolution framework is designed to capture the complex semantic knowledge within metapaths. By employing both intra-metapath aggregation and inter-metapath fusion, we can obtain comprehensive representations of nodes by considering not only their local information within a specific metapath, but also the multiple semantic roles a node can acquire in various metapath contexts.

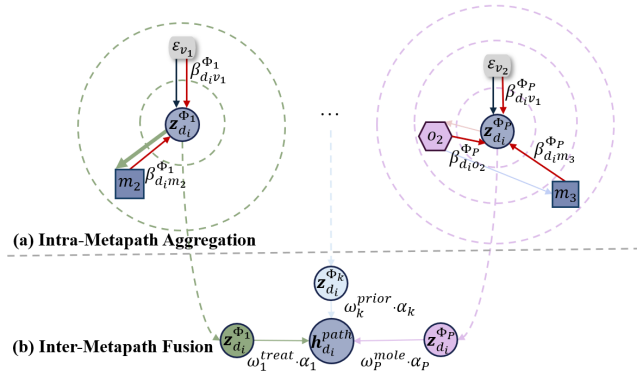


Figure 2: Overview of the integration process in both intra-metapath level and inter-metapath level.

(a) Intra-Metapath Aggregation. In order to effectively capture the semantic representations of nodes along a specified metapath Φ_k , we perform intra-metapath aggregation to

integrate information from all connections within that metapath.

$$\mathbf{z}_{e_i}^{\Phi_k} = \sigma \left(\sum_{e_j \in \mathcal{N}_{e_i}^{\Phi_k}} \beta_{e_i e_j}^{\Phi_k} \mathbf{h}_{e_j} \mathbf{W}_k \right), \quad e \in \{d, p, m, o, c, \mathcal{E}_v\}, \quad (5)$$

where $\mathbf{z}_{e_i}^{\Phi_k}$ denotes the aggregated representation of the target node e_i following metapath Φ_k . $\mathcal{N}_{e_i}^{\Phi_k}$ denotes the set of neighbors of node e_i within the specific metapath, including one hop, and multiple hops. The attention weight $\beta_{e_i e_j}^{\Phi_k}$ quantifies the significance of neighboring node e_j with respect to target node e_i in metapath Φ_k . \mathbf{W}_k is a learnable weight matrix, and σ denotes the activation function.

(b) Inter-Metapath Fusion. The primary goal of the inter-metapath fusion process is to not only integrate the embedding of a specific node obtained from various metapaths, but also to introduce an explainable layer through the design of prior weights. These weights allow us to allocate different levels of significance to distinct metapaths, enabling a more detailed understanding of relationships within the data. Mathematically, the fusion process can be represented as follows:

$$\mathbf{h}_{e_i}^{path} = \sum_{k=1}^P \left(\mathbf{z}_{e_i}^{\Phi_k} \cdot w_t^{prior} \cdot \alpha_k \right), \quad e \in \{d, p, m, \mathcal{E}_v\}, \quad (6)$$

where $\mathbf{z}_{e_i}^{\Phi_k}$ represents the embeddings of node e_i derived from the intra-metapath aggregation. w_t^{prior} denotes pre-defined prior weights assigned to different metapath types, with $w_{treat}^{prior} = 1$ (representing explicit treat relationships, which are deemed most significant), $w_{mole}^{prior} = 0.9$, and $w_{class}^{prior} = 0.8$. The variable α_k serves as a learnable weight that scales the prior weights, thus providing the flexibility to adjust the importance of each metapath in the overall representation dynamically. The combination of prior weights and learnable parameters allows the model to assign different scores to each metapath, which provides explainable recommendations by reflecting the contribution of each metapath on the final decision.

After obtaining the hypergraph convolution embedding $\mathbf{h}_{e_i}^{(l+1)}$ and metapath convolution embedding $\mathbf{h}_{e_i}^{path}$ for each entity node, we concatenate these two vectors together and then processed through a mask layer, followed by MLP to produce the updated node embedding representation \mathbf{h}'_{e_i} :

$$\mathbf{h}'_{e_i} = \text{MLP} \left(\text{mask} \odot \left[\mathbf{h}_{e_i}^{(l+1)}, \mathbf{h}_{e_i}^{path} \right] \right), \quad e \in \{d, p, m, \mathcal{E}_v\}, \quad (7)$$

where the mask layer is designed to selectively retain information based on the types of convolutions available for each node. For nodes with available metapath connections, the mask is $[1, 1]$ as node d_1 shown in Figure 1. Conversely, for nodes like d_2 that only possess hypergraph convolution without associated metapaths, the mask would be $[1, 0]$. This ensures the MLP processes only the relevant portions of the concatenated vector, effectively mitigating noise and enhancing the quality of the updated representation.

Multi-View Medication Prediction

In this section, we present a multi-view medication prediction that leverages both metapath-based and instance-based methods, which aim to enhance medication recommendation accuracy by integrating rich semantic knowledge and historical treatment information from patient visits.

Metapath-Based Prediction. In order to leverage the rich semantic knowledge captured in the metapaths, we compute the matching probability between the current visit query hyperedge $\mathbf{h}'_{\mathcal{E}_{v_t}}$ and each medication \mathbf{h}'_{m_k} .

$$P(\mathbf{h}'_{\mathcal{E}_{v_t}}, \mathbf{h}'_{m_k}) = \text{softmax} \left(\frac{(\mathbf{h}'_{\mathcal{E}_{v_t}})^T \cdot \mathbf{h}'_{m_k}}{\|\mathbf{h}'_{\mathcal{E}_{v_t}}\| \cdot \|\mathbf{h}'_{m_k}\|} \right), \quad (8)$$

where $P(\mathbf{h}'_{\mathcal{E}_{v_t}}, \mathbf{h}'_{m_k})$ quantifies how likely the current visit query will be prescribed with medication m_k .

Instance-Based Prediction. To enhance personalized medication recommendations by capturing the longitudinal medical trajectory of each patient, we apply attention aggregation over the visit hyperedges associated with each patient, thereby obtaining the patient hyperedge representation $\mathbf{h}'_{\mathcal{E}_{u_i}}$.

We then compute the cosine similarity $\text{sim}(\mathbf{h}'_{\mathcal{E}_{u_i}}, \mathbf{h}'_{\mathcal{E}_{u_j}})$ between the current target patient and others to identify similar medical histories. This approach facilitates recommendations for medications proven effective in comparable cases. The similarity for the current patient with himself or herself is defined as 1, which ensures the system to recommend medications that has been previously utilized. The final prediction score is obtained by multiplying the metapath-based score with the instance-based score:

$$\hat{y}_{m_k}^t = \text{sim}(\mathbf{h}'_{\mathcal{E}_{u_i}}, \mathbf{h}'_{\mathcal{E}_{u_j}}) \cdot P(\mathbf{h}'_{\mathcal{E}_{v_t}}, \mathbf{h}'_{m_k}). \quad (9)$$

By leveraging both the metapath-based similarity and the instance-based matching score, our approach ultimately improves the robustness and accuracy of medication recommendations.

Combined Loss for Parameter Learning

To optimize the learnable parameters, we introduce a combined loss to balance accuracy and safety. For accuracy, we employ two loss functions: binary cross-entropy loss L_{bce} , which measures the discrepancy between the predicted probabilities and the actual binary labels for each medication, and multi-label margin loss L_{multi} , which assesses the margin between the scores of relevant and irrelevant medications.

$$L_{bce} = - \sum_{i=1}^{|\mathcal{M}|} [y_{m_i}^t \log(\hat{y}_{m_i}^t) + (1 - y_{m_i}^t) \log(1 - \hat{y}_{m_i}^t)],$$

$$L_{multi} = \sum_{i,j: y_{m_i}^t=1, y_{m_j}^t=0} \frac{1}{|\mathcal{M}|} \max \left(0, 1 - (\hat{y}_{m_i}^t - \hat{y}_{m_j}^t) \right), \quad (10)$$

Attribute	MIMIC-III	MIMIC-IV	MIMIC-IV
	MED	GU	OMED
# patients	17,778	4,251	8,709
# diagnoses	5,205	3,617	9,192
# procedures	1,416	1,036	2,380
# medications	3,848	2,612	4,376
avg.# of visits	1.34	1.28	2.89
avg.# of diagnoses	16.34	8.61	29.13
avg.# of procedures	4.90	2.59	4.42
avg.# of medications	47.65	17.94	55.25

Table 1: Characteristics of Datasets

where $y_{m_i}^t$ represents the ground truth label for the i -th medication, $\hat{y}_{m_i}^t$ denotes the predicted probability of the i -th medication at time t .

For safety, we design two loss functions: the adverse drug-drug interaction loss L_{ddi} , based on the DDI adjacency matrix, and the contraindication loss L_{contra} , which encodes disease-medication safety constraints. The DDI loss captures potential harmful interactions between medications, and the contraindication loss utilizes the contraindication matrix \mathbf{A}_{contra} extracted from the knowledge graph.

$$L_{ddi} = \sum_{i=1}^{|\mathcal{M}|} \sum_{j=1}^{|\mathcal{M}|} \mathbf{A}_{ddi}[i, j] \cdot \hat{y}_{m_i}^t \cdot \hat{y}_{m_j}^t, \quad (11)$$

$$L_{contra} = \sum_{k=1}^{|\mathcal{D}|} \sum_{j=1}^{|\mathcal{M}|} \mathbf{A}_{contra}[k, j] \cdot \hat{y}_{m_j}^t \cdot y_{d_k}^t,$$

where matrix $\mathbf{A}_{contra}[k, j]$ indicates the contraindication relationship between disease d_k and medication m_j , with $y_{d_k}^t$ indicating whether disease d_k is present in the sample.

Finally, to balance between accuracy and safety during model training, we integrate all four losses into a comprehensive loss function.

$$L = \alpha(\beta L_{bce} + (1 - \beta)L_{multi}) + (1 - \alpha)(\lambda_1 L_{ddi} + \lambda_2 L_{contra}),$$

$$\alpha = \begin{cases} 1, & \text{if } \text{rate}_{harm} \leq \gamma \\ \max \left\{ 0, 1 - \frac{\text{rate}_{harm} - \gamma}{kp} \right\}, & \text{if } \text{rate}_{harm} > \gamma, \end{cases} \quad (12)$$

where rate_{harm} is the observed rate of harmful interactions, γ is the harm acceptance threshold for clinical use, and kp is a correction factor that adjusts the influence based on proportionality. The parameter β controls regularization strength, and the selection of λ_1 and λ_2 will be discussed in hyperparameter analysis section.

This combined approach ensures that the model not only delivers accurate medication recommendations but also prioritizes patient safety through the consideration of negative relations.

Experiments

Experimental Settings

Datasets and Pre-processing. Our method utilizes the publicly available MIMIC-III (Johnson et al. 2016) and MIMIC-IV (Johnson et al. 2023) datasets for training and testing.

EK	Methods	MIMIC-III-MED				MIMIC-IV-GU				MIMIC-IV-OMED			
		Jaccard \uparrow	DDI \downarrow	PRAUC \uparrow	F1 \uparrow	Jaccard \uparrow	DDI \downarrow	PRAUC \uparrow	F1 \uparrow	Jaccard \uparrow	DDI \downarrow	PRAUC \uparrow	F1 \uparrow
-	LR	0.3254	0.0959	0.6055	0.4646	0.2722	0.2720	0.5573	0.3555	0.2982	0.0689	0.5811	0.3893
-	ECC	0.2801	0.0916	0.6014	0.4027	0.2929	0.2523	0.5536	0.3778	0.3170	0.0835	0.5951	0.4119
-	RETAIN	0.3180	0.0956	0.5863	0.4696	0.2516	0.2481	0.4849	0.3542	0.3292	0.1492	0.5980	0.4188
-	DMNC	0.2847	0.0903	0.5526	0.4124	0.3171	0.2540	0.4134	0.4219	0.3404	0.0802	0.5219	0.4412
-	MICRON	0.3481	<u>0.0666</u>	0.6203	0.4928	0.2335	0.3146	0.5227	0.3107	0.3072	<u>0.0555</u>	0.5768	0.4022
-	LAMRec	<u>0.3929</u>	0.0714	0.6232	<u>0.5394</u>	0.3402	<u>0.1861</u>	0.5668	<u>0.4790</u>	0.3484	0.0855	0.5855	0.4782
\checkmark	GAMENet	0.3737	0.1130	0.6198	0.5250	0.3264	0.2790	0.5566	0.4480	<u>0.3515</u>	0.0595	<u>0.5989</u>	<u>0.4762</u>
\checkmark	COGNet	0.3765	0.1169	0.6028	0.5274	0.3193	0.1925	0.4993	0.4183	0.3191	0.0804	0.5164	0.4017
\checkmark	MoleRec	0.3843	0.0737	0.6205	0.5355	<u>0.3433</u>	0.2870	<u>0.5675</u>	0.4569	0.3415	0.0591	0.5924	0.4644
\checkmark	RASNet	0.3846	0.0673	<u>0.6259</u>	0.5361	0.2594	0.2127	0.4597	0.3516	0.2882	0.0985	0.5154	0.3812
\checkmark	KEHGCN	0.3999	0.0593	0.6320	0.5414	0.3554	0.1312	0.5779	0.4806	0.3688	0.0517	0.6199	0.5001

Table 2: Results of different methods. The best and the runner-up results are highlighted in **bold** and underline respectively.

The two datasets differ in terms of temporal coverage, data granularity, schema complexity, etc. To better simulate real-world hospital scenarios, we do not use the MIMIC datasets directly; instead, we organize the data based on departmental divisions using the services table from the Hosp module. Table 1 provides details on the datasets from three departments. For the UMLS knowledge graph, we extract the UMLS Metathesaurus full subset sources. We utilize the MRREL.RRF file to extract relationships and metapaths, and employ the MRCONSO.RRF file to establish code mapping relationships with the MIMIC datasets.

Compared Methods and Metrics. We compare our method with the representative state-of-the-art baselines. Methods like LR, ECC (Read et al. 2011), RETAIN (Choi et al. 2016), DMNC (Le, Tran, and Venkatesh 2018), MICRON (Yang et al. 2021a), and LAMRec (Tang et al. 2024) do not utilize external knowledge. In contrast, GAMENet (Shang et al. 2019), COGNet (Wu et al. 2022), MoleRec (Yang et al. 2023), and RASNet (Zhu et al. 2024) incorporate external knowledge into their frameworks. To assess the accuracy and safety of our proposed model, we follow the previous works (Tang et al. 2024) utilizing evaluation metrics including Jaccard Similarity Score, DDI rate, Precision-Recall AUC (PRAUC), and Average F1 Score for reference.

Implementation Details. We partition the dataset into training, validation, and test sets, allocating the first 2/3 for training, the middle 1/6 for testing, and the last 1/6 for validation. We set the loss function parameters $\beta = 0.95$, $kp = 0.05$, and acceptance rate $\gamma = 0.06$. The model is trained for 50 epochs using the Adam optimizer with an initial learning rate of $1e^{-4}$. Each baseline method is configured according to the default settings from the original papers or our fine-tuned parameters.

Performance Comparison

Table 2 presents the comparative performance of different methods. It is evident that our model KEHGCN outperforms all other baselines in all evaluated metrics, with statistically significant improvements (t-test, $p < 0.01$). Compared with methods that do not apply External Knowledge, our approach demonstrates superior results by integrating KG domain knowledge. As for methods involving External

Knowledge, our method also achieves notable performance by constructing hierarchical hypergraphs and extracting generalized metapaths.

In summary, our method not only enhances performance through the incorporation of knowledge graphs but also fully explores the relationships to support the learning of EHR data. By designing the generalized metapaths and hierarchical hypergraphs, our method is able to learn from multiple perspectives, thereby providing more accurate and reliable medication recommendations for patients.

Ablation Study

To verify the effectiveness of our proposed components, we conduct ablation studies by removing each of the key designs in KEHGCN. Specifically, we consider the following three variants of KEHGCN: 1) **w/o HGCN** removes the hypergraph convolution, which captures co-occurrence relations; 2) **w/o MCN** excludes the metapath convolution including intra-metapath aggregation and inter-metapath fusion, which learns the semantic knowledge inherent in the knowledge graph; 3) **w/GCN HH** substitutes hypergraph convolution with standard graph convolution; 4) **w/o Priors** excludes prior weights assigned to metapaths. The results in Table 3 demonstrate that removing any of the above components leads to a decline in performance, which indicates that each element of KEHGCN plays a crucial role in improving its overall performance.

Hyperparameter Analysis

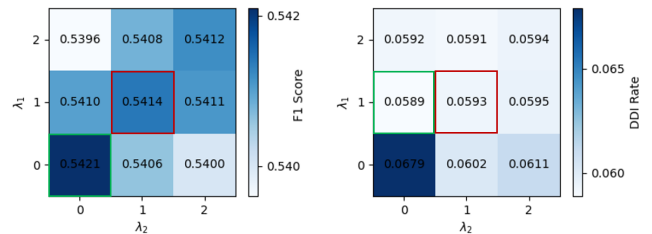


Figure 3: Impact of λ_1 and λ_2 on F1 score (accuracy) and DDI rate (safety) in MIMIC-III-MED Dataset.

Methods	MIMIC-III-MED				MIMIC-IV-GU				MIMIC-IV-OMED			
	Jaccard \uparrow	DDI \downarrow	PRAUC \uparrow	F1 \uparrow	Jaccard \uparrow	DDI \downarrow	PRAUC \uparrow	F1 \uparrow	Jaccard \uparrow	DDI \downarrow	PRAUC \uparrow	F1 \uparrow
w/o HGCN	0.3810	0.0675	0.6046	0.5271	0.3385	0.2720	0.5611	0.4522	0.3527	0.0559	0.5809	0.4666
w/o MCN	0.3940	0.0613	0.6289	0.5400	0.3548	0.2059	0.5764	0.4682	0.3618	0.0523	0.6091	0.4802
w/GCN HH	0.3845	0.0599	0.6114	0.5335	0.3401	0.2661	0.5627	0.4615	0.3578	0.0550	0.5871	0.4695
w/o Priors	0.3950	0.0597	0.6305	0.5410	0.3498	0.1607	0.5715	0.4733	0.3564	0.0591	0.6002	0.4768
KEHGCN	0.3999	0.0593	0.6320	0.5414	0.3554	0.1312	0.5779	0.4806	0.3688	0.0517	0.6199	0.5001

Table 3: Results of ablation study.

The impact of hyperparameters λ_1 and λ_2 on F1 score and DDI rate on MIMIC-III-MED dataset is illustrated in Figure 3, where the green box denotes the optimal parameter values for each metric, and the red box indicates the scenario when $\lambda_1 = 1, \lambda_2 = 1$ yields favorable outcomes for both indicators. These parameters are critical for balancing the losses associated with negative relations, where λ_1 governs the DDI relationships and λ_2 regulates the contraindications between diseases and medications. Experimental results reveal that when both parameters are set to zero, there is an increase in accuracy (F1). However, the increase is moderately limited because the recommendations involve negative interactions. When λ_1 is set to 1 while λ_2 remains 0, the DDI rate achieves its lowest value. However, the optimal balance between accuracy and safety is achieved at the parameter values [1, 1], which indicates that these specific settings effectively enhance performance while minimizing risks associated with safety concerns. By selecting these settings, we aim to ensure robust performance that adheres to both accuracy standards and safety protocols.

Case Study

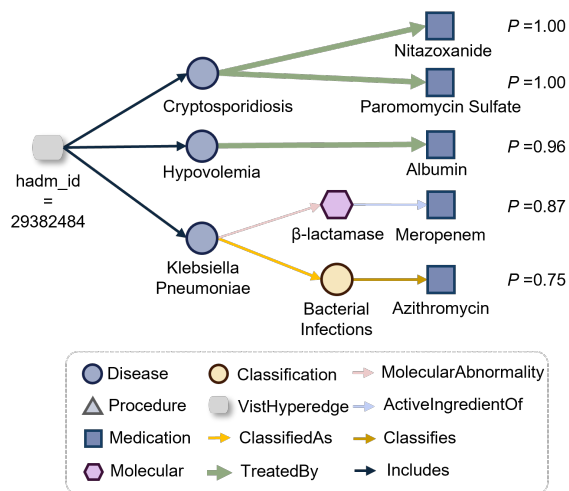


Figure 4: An example of top-5 metapaths for a specific visit.

One of the key innovations in our work is the provision of explainability in medication recommendations. In the meta-path convolution process, we assign prior weights to different types of metapaths. For metapaths that exhibit a direct treatment relationship, we set the maximum weight to

1. In addition, we introduce learnable weights to fine-tune these assignments. During each recommendation, we filter and present the top-5 metapaths with the highest weights to patients and clinicians.

Figure 4 demonstrates a case study from the MIMIC-IV-OMED dataset, which provides five highest-weighted metapaths for illustration. For direct treatment relationships, the metapaths retain their maximum weight even after the application of adjustable weights. This allows us to directly recommend a medication that corresponds to a specific disease. Such metapaths enhance explainability significantly, as they enable the users to quickly grasp the rationale behind the recommended medication. Additionally, when a disease results from an abnormal molecular structure, and the corresponding medication effectively addresses this anomaly, the explainability offered by such metapaths is remarkably beneficial. Finally, we also observe that metapaths related to medications and diseases within the same classification can provide valuable explanations.

To summarize, our approach not only enhances the accuracy and safety of medication recommendations but also prioritizes explainability, facilitating better understanding of healthcare decisions.

Conclusion

In this paper, we propose a novel Knowledge-enhanced Explainable HyperGraph Convolution Network (KEHGCN) that integrates knowledge graph and provides explainability in medication recommendation. To be specific, our method constructs a hierarchical hypergraph structure and applies a knowledge-driven hypergraph convolution network to better capture the multi-level relationships within EHR data. Moreover, the contraindication relations between diseases and medications derived from knowledge graph help regularize the loss function, which is evaluated in our hyperparameter analysis. For future work, we plan to incorporate other types of knowledge sources and multi-modal medical data, such as medical images and genomic information. Additionally, we will explore the integration of other types of negative relations to further enhance the safety of medication recommendations.

Acknowledgments

This work was supported by the S&T Program of Hebei (Grant No.252W7713D) and the National Key Research and Development Program of China (Grant No. 2022YFB2703301).

References

- Bodenreider, O. 2004. The Unified Medical Language System (UMLS): Integrating Biomedical Terminology. In *Nucleic acids research*, 32: D267–D270.
- Chen, Q.; Li, X.; Geng, K.; and Wang, M. 2023. Context-Aware Safe Medication Recommendations with Molecular Graph and DDI Graph Embedding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 7053–7060.
- Choi, E.; Bahadori, M. T.; Sun, J.; Kulas, J.; Schuetz, A.; and Stewart, W. 2016. RETAIN: An Interpretable Predictive Model for Healthcare using Reverse Time Attention Mechanism. In *Advances in Neural Information Processing Systems*, 3504–3512.
- Gao, C.; Wang, X.; He, X.; and Li, Y. 2022. Graph Neural Networks for Recommender System. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, 1623–1625.
- Gong, F.; Wang, M.; Wang, H.; Wang, S.; and Liu, M. 2021. SMR: Medical Knowledge Graph Embedding for Safe Medicine Recommendation. In *Big Data Research*, 23: 100174.
- Jin, Y.; Zhang, W.; He, X.; Wang, X.; and Wang, X. 2020. Syndrome-aware Herb Recommendation with Multi-Graph Convolution Network. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, 145–156.
- Johnson, A. E.; Pollard, T. J.; Shen, L.; Lehman, L.-w. H.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.; Celi, L. A.; and Mark, R. G. 2016. MIMIC-III, A Freely Accessible Critical Care Database. In *Scientific Data*, 3(1): 1–9.
- Johnson, A. E. W.; Bulgarelli, L.; Shen, L.; Gayles, A.; Shammout, A.; Horng, S.; Pollard, T. J.; Hao, S.; Moody, B.; Gow, B.; et al. 2023. MIMIC-IV, A Freely Accessible Electronic Health Record Dataset. In *Scientific Data*, 10(1): 1.
- Kim, T.; Heo, J.; Kim, H.; and Kim, S.-W. 2025. HI-DR: Exploiting Health Status-Aware Attention and an EHR Graph+ for Effective Medication Recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 11950–11958.
- Kim, T.; Heo, J.; Kim, H.; Shin, K.; and Kim, S.-W. 2024. VITA: 'Carefully Chosen and Weighted Less' Is Better in Medication Recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 8600–8607.
- Lai, K.-H.; Yang, Z.-R.; Lai, P.-Y.; Wang, C.-D.; Guizani, M.; and Chen, M. 2024. Knowledge-Aware Explainable Reciprocal Recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 8636–8644.
- Le, D.-H. 2025. Improving Computational Drug Repositioning Through Multi-Source Disease Similarity Networks. *Scientific Reports*, 15(30773).
- Le, H.; Tran, T.; and Venkatesh, S. 2018. Dual Memory Neural Computer for Asynchronous Two-view Sequential Learning. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1637–1645.
- Liang, S.; Li, X.; Mu, S.; Li, C.; Lei, Y.; Hou, Y.; and Tengfei, M. 2025. CIDGMed: Causal Inference-Driven Medication Recommendation with Enhanced Dual-Granularity Learning. In *Knowledge-Based Systems*, 309: 112685.
- Liu, H.; Zhao, N.; Zhou, J.; Qu, W.; and Ji, Z. 2024a. Chiral Molecular Graph Encoder for Medication Recommendation. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 1036–1041.
- Liu, H.; Zhu, Y.; and Wu, Z. 2024. Knowledge Graph-Based Behavior Denoising and Preference Learning for Sequential Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 36(6): 2490–2503.
- Liu, S.; Wang, X.; Zhao, X.; and Chen, H. 2024b. DKINet: Medication Recommendation via Domain Knowledge Informed Deep Learning. In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 3521–3526.
- Ma, X.; Wang, Y.; Chu, X.; Ma, L.; Tang, W.; Zhao, J.; Yuan, Y.; and Wang, G. 2022. Patient Health Representation Learning via Correlational Sparse Prior of Medical Features. In *IEEE Transactions on Knowledge and Data Engineering*, 35(17): 11769–11783.
- Maya, R.; Yoni, H.; Abdulhakim, T.; Steven, H.; and David, S. 2017. Learning a Health Knowledge Graph from Electronic Medical Records. *Scientific Reports*, 7(5994).
- Mishra, R.; and Shridevi, S. 2024. Knowledge Graph Driven Medicine Recommendation System Using Graph Neural Networks on Longitudinal Medical Records. In *Scientific Reports*, 14(1): 25449.
- Read, J.; Pfahringer, B.; Holmes, G.; and Frank, E. 2011. Classifier Chains for Multi-Label Classification. In *Machine learning*, 85(3): 333–359.
- Ren, Y.; Shi, Y.; Zhang, K.; Wang, X.; Chen, Z.; and Li, H. 2022. A Drug Recommendation Model Based on Message Propagation and DDI Gating Mechanism. In *IEEE Journal of Biomedical and Health Informatics*, 26(7): 3478–3485.
- Saadat, H.; Shah, B.; Halim, Z.; and Anwar, S. 2024. Knowledge Graph-Based Convolutional Network Coupled With Sentiment Analysis Towards Enhanced Drug Recommendation. In *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 21(4): 983–994.
- Shang, J.; Xiao, C.; Ma, T.; Li, H.; and Sun, J. 2019. GAMENet: Graph Augmented MEMory Networks for Recommending Medication Combination. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 1126–1133.
- Sun, T.; Guo, H.; Zhang, Z.; Liu, H.; and Wu, Z. 2024. Exploiting Multifaceted Nature of Items and Users for Session-Based Recommendation. In *Proceedings of the 2024 SIAM International Conference on Data Mining*, 580–588.
- Tang, Y.; Liu, N.; Yuan, H.; Yan, Y.; Liu, L.; Tan, W.; and Cui, L. 2024. LAMRec: Label-aware Multi-view Drug Recommendation. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, 2230–2239.

Wang, X.; Huang, T.; Wang, D.; Yuan, Y.; Liu, Z.; He, X.; and Chua, T.-S. 2021. Learning Intents behind Interactions with Knowledge Graph for Recommendation. In *Proceedings of the Web Conference 2021*, 878–887.

Wu, R.; Qiu, Z.; Jiang, J.; Qi, G.; and Wu, X. 2022. Conditional Generation Net for Medication Recommendation. In *Proceedings of the ACM Web Conference 2022*, 935–945.

Yang, C.; Xiao, C.; Glass, L.; and Sun, J. 2021a. Change Matters: Medication Change Prediction with Recurrent Residual Networks. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, 3728–3734.

Yang, C.; Xiao, C.; Ma, F.; Glass, L.; and Sun, J. 2021b. SafeDrug: Dual Molecular Graph Encoders for Recommending Effective and Safe Drug Combinations. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, 3735–3741.

Yang, N.; Zeng, K.; Wu, Q.; and Yan, J. 2023. MoleRec: Combinatorial Drug Recommendation with Substructure-Aware Molecular Representation Learning. In *Proceedings of the ACM Web Conference 2023*, 4075–4085.

Yang, T.; Ren, B.; Gu, C.; Ma, B.; He, T.; and Konomi, S. 2025. Towards Better Course Recommendations: Integrating Multi-Perspective Meta-Paths and Knowledge Graphs. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference*, 137–147.

Yao, Z.; Liu, B.; Wang, F.; Sow, D.; and Li, Y. 2023. Ontology-Aware Prescription Recommendation in Treatment Pathways Using Multi-Evidence Healthcare Data. In *ACM Transactions on Information Systems*, 41(4): 1–29.

Zhang, Y.; Chen, R.; Tang, J.; Stewart, W. F.; and Sun, J. 2017. LEAP: Learning to Prescribe Effective and Safe Treatment Combinations for Multimorbidity. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1315–1324.

Zhu, Q.; Han, F.; Yang, H.; Liu, J.; Hu, X.; and Wang, B. 2024. RASNet: Recurrent Aggregation Neural Network for Safe and Efficient Drug Recommendation. In *Knowledge-Based Systems*, 299: 112055.