Knowledge Enhanced Multi-intent Transformer Network for Recommendation

Project Report

Submitted by

Borra Lakshya, Keerthana S, Shubhi Agarwal

Under the guidance of

Dr. Chiranjib Sur
Professor

Department of Data Science & Artificial Intelligence
Indian Institute of Technology Guwahati,
India

November 2024

Contents

Chapter 1

Introduction

In recent years, understanding user intents behind interaction behaviors has become increasingly crucial for accurate recommendation systems. Our research focuses on modeling these intents by integrating global collaborative (user-item) and knowledge (item-relation-entity) information. By leveraging these modeled intents, we aim to guide knowledge sampling, ultimately enabling more fine-grained and precise user/item representation learning. To address these challenges, we are using a novel model called KGTN, which introduces two key components to overcome the limitations in existing methods.

The first component of KGTN is Global Intents Modeling with Graph Transformer. Here, we predefine 'K' intent representations for users and items, and train these intents using global information from collaborative and knowledge graphs. Specifically, this process involves merging knowledge information into item representations, followed by employing a novel graph transformer on the user-item graph to learn global intents and generate intent-aware representations for both users and items. This approach allows KGTN to capture essential global insights from both collaborative interactions and knowledge structures.

The second component is Knowledge Contrastive Denoising under Intents. This module uses intent-aware user/item representations to guide knowledge sampling, effectively filtering out irrelevant knowledge that might otherwise dilute the quality of representation. Additionally, we introduce a local-global contrastive learning mechanism to further denoise these representations, ensuring that the model remains focused on meaningful information relevant to the user or item.

Empirically, KGTN demonstrates strong performance, consistently outperforming state-of-the-art models across three benchmark datasets in offline testing and delivering significant improvements in online A/B testing on Alibaba's recommendation platform.

The primary contributions of the original research can be summarized as follows:

1. General Aspects: We highlight the importance of intent modeling with global information, which is essential for achieving fine-grained representation learning and effective knowledge denoising.

2. Novel Methodologies: We introduce KGTN, a model that leverages global signals for user intent modeling through a novel graph transformer, and employs two techniques for denoising—knowledge denoising under intents and local-global graph contrastive learning.

3. Multifaceted Experiments: Extensive experiments, both offline on three benchmark datasets and online through A/B testing on the Alibaba recommendation platform, confirm the effectiveness of KGTN in enhancing representation learning.

1.1 Dataset

We use the dataset about book recommendations, comprising two main files ratings_final.npy and kg_final.txt. The ratings_final.npy file contains user-book interaction data, capturing which users have engaged with specific books. This interaction data reflects users' preferences and engagement patterns, providing a foundational layer for understanding the individual preferences for book recommendations.

The kg_final.txt file contains a knowledge graph (KG) in the form of triplets, each represented as (source, relation, target). These triplets encode relationships not only among books but also between books and other entities too, offering additional contextual information beyond direct user interactions. This knowledge graph improves the dataset by connecting books with shared attributes or external relationships, thereby allowing the recommendation system to consider books' underlying traits. Leveraging this relational data, we can make recommendations with characteristics, enhancing the diversity and relevance of suggested books.

An additional file, kg_final.npy, holds the same knowledge graph information as kg_final.txt but is not used in the prediction process for this analysis. Instead, we primarily rely on ratings_final.npy and kg_final.txt to train our graph neural network (GNN)-based recommender. Together, these files provides a structured approach for generalizing user preferences by linking similar books and entities, enabling the model to generate recommendations that are both diverse and aligned with users' interests.

1.2 Proposed Solution

The Knowledge Enhanced Multi-intent Transformer Network (KGTN), provides a novel approach to capturing user intent and enhancing recommendation accuracy. KGTN is designed with two primary components, each addressing key challenges in recommendation systems.

The first component, Global Intents Modeling with Graph Transformer, focused on capturing learnable user intents in a way that enhances the depth and relevance of user/item representations. By leveraging a graph transformer, KGTN can identify and incorporate global patterns across collaborative and knowledge graphs, resulting in representations that are intent-aware and able to reflect the preferences of users. This approach allows the model to capture a broader context, n for better recommendation accuracy.

The second component, Knowledge Contrastive Denoising under Intents, is essential for learning precise and robust representations. This uses a contrastive learning technique that filters out noise and irrelevant information, allowing KGTN to focus on learning signal-rich and patterns within the data. By reducing the influence of noise, this component increases the quality of representation learning, making it more resilient and accurate.

Together, these components enable KGTN to effectively model multiple user intents and to deliver high-quality, noise-irrelevant representations that drive better recommendations.

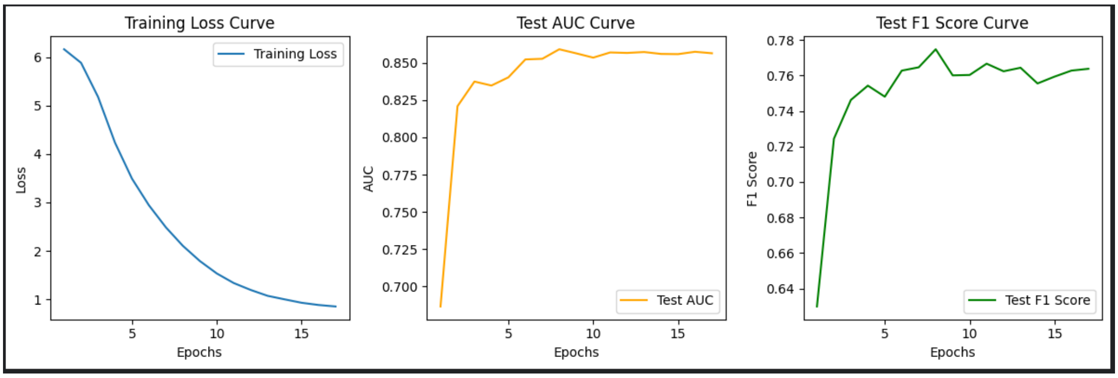

Figure 1.1

Chapter 2

Implementation

2.1 Encoding User IDs and Book IDs

To prepare the data effectively for model training, LabelEncoder was applied to convert categorical user and book IDs into continuous integer values. This transformation was crucial, as machine learning algorithms typically require numerical inputs rather than categorical data. By encoding these IDs, we enabled the model to recognize and interpret unique user and book identifiers as numeric features, thus facilitating the learning process. This encoding approach also helped streamline the data, maintaining the uniqueness of each user and book while allowing their linkings and patterns to be more accessible for the algorithm during training.

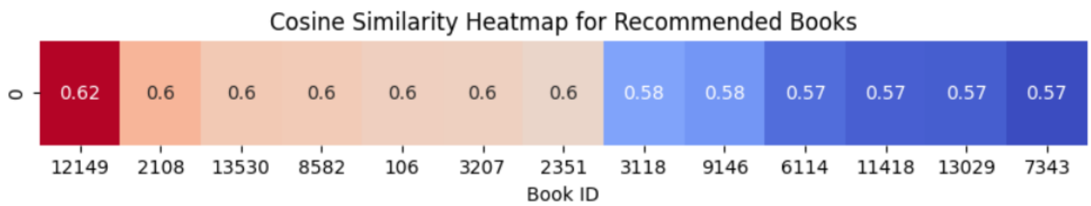

Figure 2.1

2.2 Constructing Graph Edges and Relations

To construct an informative graph structure, edges were defined to capture the diverse relationships among entities. The foundation of this graph was established using knowledge graph triples, where edges connected entities based on built-in relationships, such as a book being associated with certain themes or genres. Beyond these static connections, additional edges were incorporated from user-book interactions got from the ratings data, thus introducing user-driven insights into the graph. By integrating these interactions, we added another layer of personalized context that reflects real-world usage patterns, further improving the network's structure.

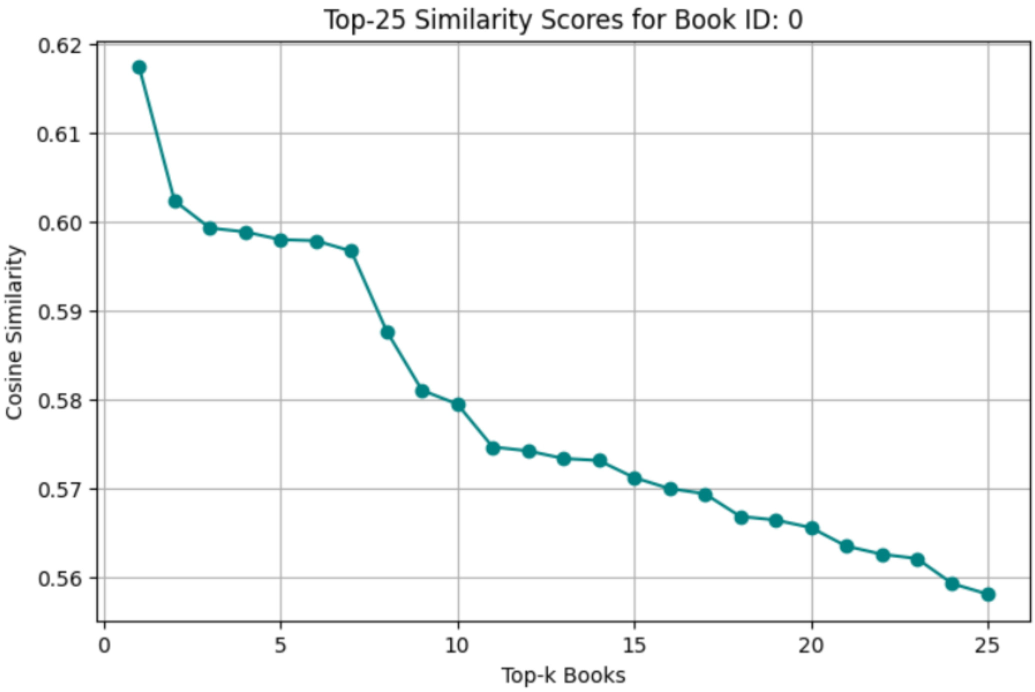

Figure 2.2

Each edge was assigned a specific relation type to reflect the nature of its connection—whether it was a user rating a book, a book featuring a particular attribute, or a shared characteristic between entities. This labeling allowed the model to differentiate between various types of interactions and attributes, ensuring that it could accurately interpret the nuances within the data.

The result was a thorough and well-connected graph structure, providing a foundation that the model could leverage to understand both the explicit and implicit relationships within the data. To enhance this structure, all edges were transformed into an undirected format, ensuring that information could flow freely between nodes without being restricted by directional constraints. This undirected approach enabled the model to process interactions globally, capturing the full complexity of user preferences and item attributes, and setting the stage for more effective recommendation and representation learning.

2.3 Model Definition

We developed a custom Graph Neural Network (GNN) model utilizing the Relational Graph Convolutional Network (RGCN) architecture. The choice of RGCN was driven by the need to handle the various relationships inherent in our data, as each user and book node in the graph has unique interactions and multi-faceted connections. The RGCN excels at learning from these kinds of heterogeneous data, allowing it to capture intricate dependencies and represent various types of relationships effectively within a single graph structure.

The model architecture includes two convolutional layers with learnable parameters, enabling dynamic feature transformation across layers. This setup allows the model to capture progressively higher-order interactions and dependencies between nodes, which enhances the richness of the embeddings. The inclusion of a dropout layer further strengthens the model by mitigating overfitting, thereby improving its generalization ability across unseen data and making it more resilient during training.

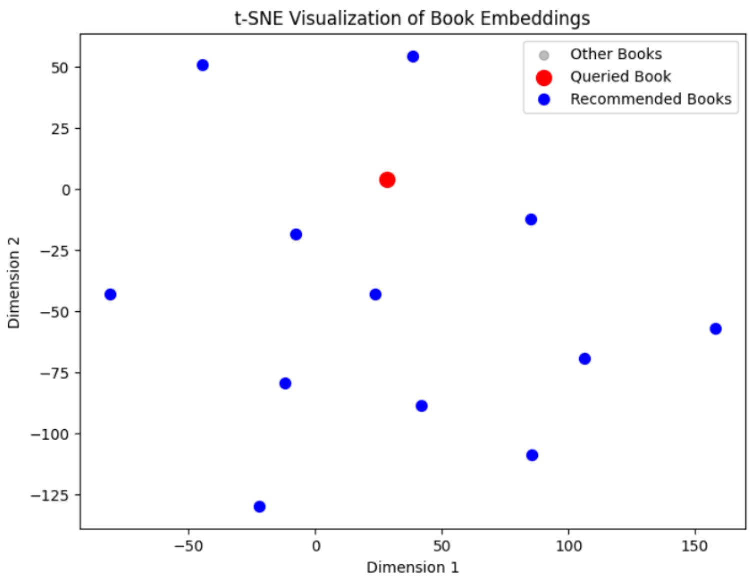

Figure 2.3

The RGCN architecture supports better representation learning for both users and books, producing embeddings that encapsulate the nuanced relationships and distinct characteristics of each node in the graph. By learning these embeddings, the model gains a deeper understanding of the data, which translates into more accurate recommendations and insights. This architecture not only enhances the expressiveness of the representations but also reinforces the model's capacity to capture and leverage the unique structure of the graph for improved performance on downstream tasks.

2.4 Training Process

The model was trained using the Adam optimizer to adjust model parameters, aiming to capture relationships between users and books based on their interactions. The learning process was driven by a triplet margin loss function, which guided the model in distinguishing between positive edges (actual user-book interactions) and negative edges (non-interactions selected at random). This approach helped the model refine its understanding of how different types of interactions contribute to users' and books' embeddings.

By leveraging triplet margin loss, the model learned to bring embeddings of interacting user-book pairs closer together while increasing the distance from unrelated pairs, resulting in more accurate and representative embeddings. Through this iterative process, the model adjusted its parameters with each epoch, effectively minimizing loss and enhancing its ability to provide precise embeddings that encloses both the interactions and attributes of users and books. This training allowed the model to efficiently discriminate between relevant and irrelevant relationships in the data, further improving its predictive capabilities.

2.5 Conclusion

Recommender systems plays a significant role in personalizing user experiences by suggesting items that match individual preferences. In this case, the model uses a recommendation feature based on the calculation of cosine similarity between learned embeddings. Each book's embedding is created during training, capturing both user tastes and inherent book attributes. By measuring cosine similarity between a specific book and others in the dataset, the model identifies those with the highest similarity, enabling it to recommend books most aligned with a user's interests.

This approach allows the system offering suggestions that reflect deeper patterns of user preferences and item properties. The embedding-based methodology provides flexibility, adapting to diverse tastes and ensuring that recommendations are both meaningful and highly relevant. By leveraging these personalized embeddings, the recommender system delivers a tailored experience, guiding users toward books they are likely to enjoy.

2.6 References

[1] Ding Zou, Wei Wei, Feida Zhu, Chuanyu Xu, Tao Zhang, Chengfu Huo. Knowledge Enhanced Multi-intent Transformer Network for Recommendation