Toronto Metropolitan University

Data Science & Analytics Graduate Studies

LLM-Augmented Knowledge Graph-Based Recommendation Systems

Kartikey Chauhan

Toronto Metropolitan University | Data Science & Analytics Graduate Studies

Domain under Study

- Recommender systems have been widely applied to address the issue of information overload in various internet services, exhibiting promising performance in scenarios ranging from search engines, E-commerce, to social media sites and news portals.
- In the general domain, the traditional knowledge recommendation method is *collaborative filtering (CF),* which usually suffers from the cold start problem and sparsity of user-item interactions.
- To address these limitations, incorporating *knowledge graphs (KG)* as side information to improve the recommendation performance has attracted attention to increasing researchers, and have achieved state of the art performance. (e.g. LightGCN)
- However, KGs are difficult to construct and evolve by nature, and existing methods often lack considering textual information. LLM-augmented KGs, that leverage *Large Language models* (LLM) for different KG tasks such as embedding, completion, construction can be a way to help overcome these challenges and lead to better recommendation systems.

List of the Research Papers

- Wang, Xiang, et al. *“KGAT: Knowledge Graph Attention Network for Recommendation.”*, 2019.
- Volokhin, et al. *“Augmenting graph convolutional networks with textual data for recommendations”*, 2023
- Pan, Shirui, et al. *“Unifying Large Language Models and Knowledge Graphs: A Roadmap.”*, 2024
- W. Wei, et al. *“LLMRec: Large Language Models with Graph Augmentation for Recommendation”*, 2024

KGAT: Knowledge Graph Attention Network for Recommendation

- **Background:** Side information, such as item attributes and user profiles can be used in recommendation systems to provide more accurate, diverse, and explainable recommendations. Traditional collaborative filtering (CF) methods focus on user-item interactions, while supervised learning (SL) methods like factorization machines (FM) fail to capture the relations among instances or items, which are crucial for distilling collaborative signals from collective user behaviors.
- **Aim:** The main aim of the paper is to develop a method that can effectively exploit high-order relations in collaborative knowledge graphs (CKGs) for recommendation in an efficient, explicit, and end-to-end manner.
- **Methodology:** The authors propose a new method called **Knowledge Graph Attention Network (KGAT),** which employs the following key components:
  - **Embedding layer:** Parameterizes each node (user, item, attribute) as a vector representation by preserving the structure of the CKG using knowledge graph embedding.
  - **Attentive embedding propagation layers:** Recursively propagate embeddings from a node's neighbors to update its representation, using a knowledge-aware attention mechanism to learn the importance of each neighbor during propagation.
  - **Prediction layer:** Aggregates the representations of a user and an item from all propagation layers to output the predicted matching score.
- **Results:**
  - Extensive experiments on three public benchmarks (Amazon-book, Last-FM, and Yelp2018) demonstrate the effectiveness of KGAT.
  - KGAT consistently outperforms state-of-the-art methods of that time, improving over the strongest baselines by up to 10.05% in terms of recall and NDCG.
- **Conclusion:** The results demonstrate the significance of capturing high-order relations and the potential of graph neural networks for knowledge-aware recommendation.

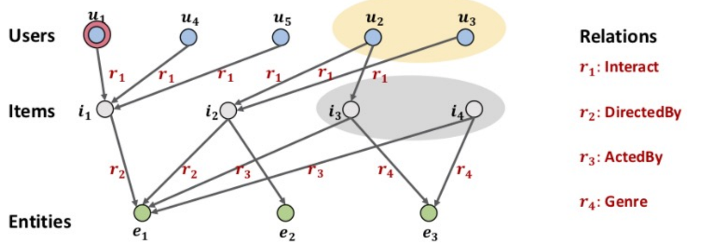


Figure 1: A toy example of collaborative knowledge graph.

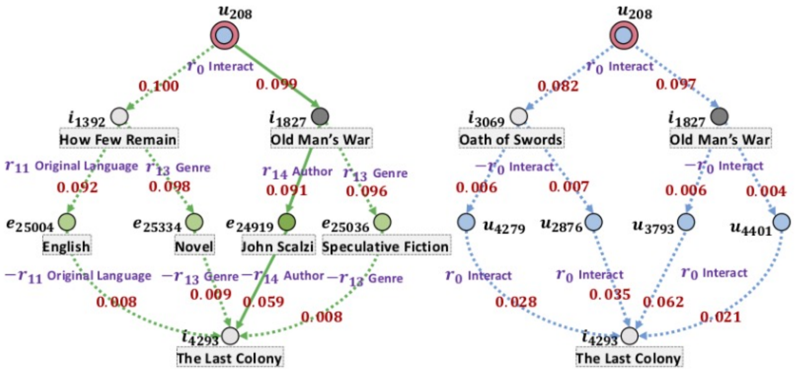
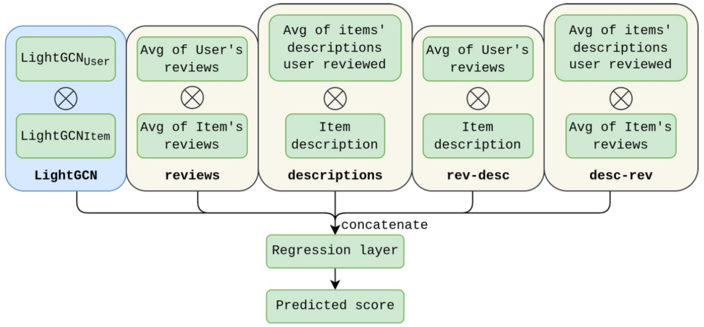


Figure 4: Real Example from Amazon-Book.

Augmenting Graph Convolutional Networks with Textual Data for Recommendations

- **Background:** Graph Convolutional Networks (GCNs) have shown state-of-the-art performance for collaborative filtering-based recommender systems, but most systems use only user-item interaction graphs, ignoring available textual information about users and items.
- **Aim:** The paper proposes an effective and general method, TextGCN, that utilizes rich textual information about graph nodes, specifically user reviews and item descriptions, using pre-trained text embeddings. The aim is to integrate this textual information into item recommendations to augment graph embeddings obtained using LightGCN, a state-of-the-art graph network.
- **Methodology:**
  - Use LightGCN to obtain node embeddings from the user-item interaction graph.
  - Create textual representations of users and items from reviews and descriptions using sentence embeddings.
  - Combine the graph and textual representations as features using dot products.
  - Train a regression layer on these features to predict user-item scores.
- **Results:**
  - TextGCN achieves statistically significant improvements of 7-23% over the LightGCN baseline across various evaluation metrics on several large-scale review datasets.
  - Incorporating textual data captures semantic signals not available from graph connections alone.
  - The most important feature is the similarity between a user's averaged item description vector and the candidate item vector.
- **Conclusion:** The paper establishes that unstructured textual data can be effectively exploited to improve recommendations using Graph Neural Networks, as it captures information not present in the user-item interaction graph. A simple approach of augmenting lightweight graph embeddings with textual representations leads to substantial performance gains.

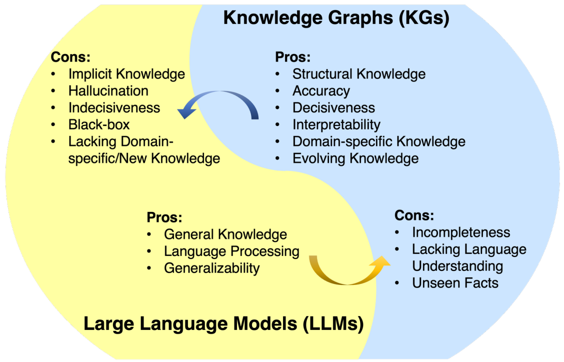


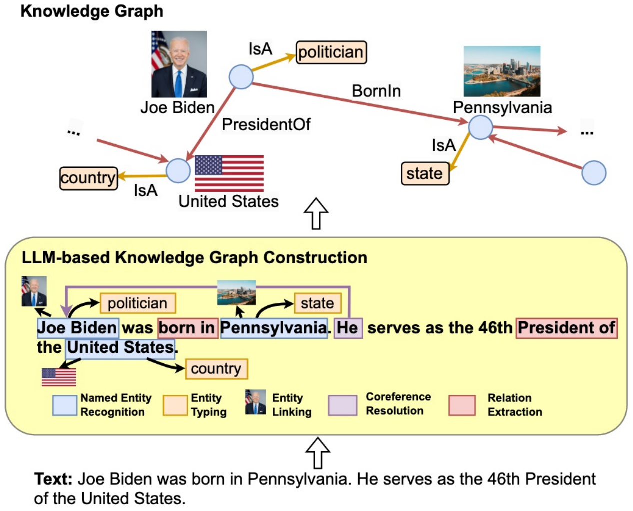
Model	Recall	Precision	Hit rate	nDCG
CF (BPR)	0.1422	0.0391	0.4662	0.1029
Graph Baselines:				
SAGE	0.0963	0.0263	0.3479	0.0674
GAT	0.1366	0.0359	0.4452	0.0984
GATv2	0.1384	0.0364	0.4503	0.0993
GCN	0.1419	0.0374	0.4584	0.1029
"Single"	0.1162	0.0318	0.4081	0.0833
LightGCN	0.1690	0.0455	0.5210	0.1244
LightGCN w DNS	0.1813	0.0490	0.5467	0.1353
TextGCN:				
XGBoost	0.1539	0.0372	0.4736	0.1075
GBDT	0.1749	0.0453	0.5308	0.1304
Linear	0.1833	0.0460	0.5308	0.1350
Linear w DNS	**0.1923**	**0.0485**	**0.5481**	**0.1428**

Toronto Metropolitan University
Data Science & Analytics Graduate Studies

Unifying Large Language Models and Knowledge Graphs.

- **Background:** Large language models (LLMs) have achieved remarkable success and generalizability in various applications. However, they often fall short of capturing and accessing factual knowledge. Knowledge graphs (KGs) are structured data models that explicitly store rich factual knowledge. Nevertheless, KGs are hard to construct and existing methods in KGs are inadequate in handling the incomplete and dynamically changing nature of real-world KG's.
- **Aim:** The paper aims to present a comprehensive roadmap for unifying LLMs and KGs to leverage the strengths of both technologies.
- **Methodology:** The methodology involves proposing three general frameworks for integrating LLMs and KGs:
  - **KG-enhanced LLMs:** Incorporating KGs into the pre-training and inference phases of LLMs to enhance their understanding and generation of knowledge.
  - **LLM-augmented KGs:** Utilizing LLMs for various KG tasks such as embedding, completion, construction, and question answering to improve KG performance and applicability.
  - **Synergized LLMs + KGs:** Developing a unified framework where LLMs and KGs work together in a mutually beneficial manner for bidirectional reasoning, leveraging both data and knowledge.
- **Results:**
  - The proposed frameworks aim to mitigate these issues by combining the strengths of LLMs and KGs, leading to enhanced performance in knowledge representation, reasoning, and various downstream tasks.
- **Conclusion:**
  - Unifying LLMs and KGs presents a forward-looking approach to overcoming the limitations of each technology while leveraging their respective advantages.





LLMRec: Large Language Models with Graph Augmentation for Recommendation

- **Background:** The challenge of data sparsity in recommendation systems has been a long-standing issue, with previous attempts to mitigate it by incorporating side information often introducing problems such as noise, availability issues, and low data quality.Recent advancements in large language models (LLMs) offer extensive knowledge bases and strong reasoning capabilities, presenting a new avenue for enhancing recommender systems.
- **Aim:** To propose a novel framework, **LLMRec**, that leverages LLMs for graph augmentation in recommender systems, aiming to address the challenges of sparse implicit feedback and low-quality side information. LLMRec seeks to enhance the performance of recommender systems by employing three LLM-based graph augmentation strategies and a denoised data robustification mechanism.
- **Methodology:**
  - **LLM-based Implicit Feedback Augmentation:** LLMs are used to sample positive and negative user-item pairs from a candidate pool, which are then added to the original training data.
  - **LLM-based Attribute Augmentation:** LLMs are used to generate user profiles and item attributes, which are then incorporated into the recommender model.
  - **Denoised Robustification:** Techniques like noise pruning and masked autoencoder-based feature enhancement are used to make the model robust to the augmented data.
- **Results:** The authors evaluate LLMRec on two real-world datasets and show that it outperforms state-of-the-art recommendation methods in terms of various metrics, such as Recall@K and NDCG@K. The results demonstrate the effectiveness of the proposed LLM-based augmentation and denoised robustification strategies.
- **Conclusion:** LLMRec is a promising approach that leverages the power of LLMs to enhance recommender systems by augmenting user-item interactions and item attributes. The novel framework and its components (graph augmentation strategies and robustification mechanism) offer a promising direction for future research in leveraging LLMs to improve recommender systems.

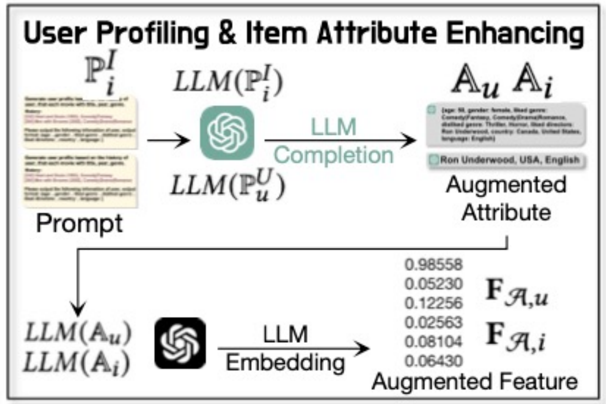


Table 1: Statistics of the Original and Augmented Datasets

Dataset		Netflix			MovieLens		
Graph	Ori.	# U	# I	# E	# U	# I	# E
		13187	17366	68933	12495	10322	57960
	Aug.	**# E:**	**26374**		**# E:**	**24990**	
Ori. Sparsity		99.970%			99.915%		
Att.	Ori.	U: None	I: year, title		U: None	I: title, year, genre	
	Aug.	U[1536]:	**age, gender, liked genre, disliked genre, liked directors, country, and language**				
		I[1536]:	**director, country, language**				
Modality		Textual[768], Visiual [512]			Textual [768], Visiual [512]		

* Att. represents attribute, Ori. represents original, and Aug. represents augmentation. Number in [] represents the feature dimensionality.

Toronto Metropolitan University | Data Science & Analytics Graduate Studies

References

- [1] X. Wang, X. He, Y. Cao, M. Liu, and T.-S. Chua, "KGAT: Knowledge Graph Attention Network for Recommendation," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Jul. 2019, pp. 950–958. doi: 10.1145/3292500.3330989.

- [2] S. Volokhin, M. D. Collins, O. Rokhlenko, and E. Agichtein, "Augmenting Graph Convolutional Networks with Textual Data for Recommendations," in *Advances in Information Retrieval*, vol. 13981, J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, and A. Caputo, Eds., in Lecture Notes in Computer Science, vol. 13981. , Cham: Springer Nature Switzerland, 2023, pp. 664–675. doi: 10.1007/978-3-031-28238-6_58.

- [3] S. Pan, L. Luo, Y. Wang, C. Chen, J. Wang, and X. Wu, "Unifying Large Language Models and Knowledge Graphs: A Roadmap," *IEEE Trans. Knowl. Data Eng.*, pp. 1–20, 2024, doi: 10.1109/TKDE.2024.3352100

- [4] W. Wei et al., "LLMRec: Large Language Models with Graph Augmentation for Recommendation." arXiv, Jan. 06, 2024. Accessed: Mar. 20, 2024. [Online]. Available: http://arxiv.org/abs/2311.00423