

**Monthly Progress Report  
February 2025**

**Submitted by  
Nithish Kumar V  
Reg. No: CS24D0002**

**Under the Guidance of  
Prof. B. Sivaselvan**



**Department of Computer Science and Engineering  
Indian Institute of Information Technology, Design and Manufacturing,  
Kancheepuram, Chennai 600127**

# Counterfactual Learning on Graphs: A Survey

[Paper Link](#)

## Abstract:

Graph-structured data are pervasive in the real-world such as social networks, molecular graphs and transaction networks. Graph neural networks (GNNs) have achieved great success in representation learning on graphs, facilitating various downstream tasks. However, GNNs have several drawbacks such as lacking interpretability, can easily inherit the bias of data and cannot model casual relations. Recently, counterfactual learning on graphs has shown promising results in alleviating these drawbacks. Various approaches have been proposed for counterfactual fairness, explainability, link prediction and other applications on graphs. To facilitate the development of this promising direction, in this survey, we categorize and comprehensively review papers on graph counterfactual learning. We divide existing methods into four categories based on problems studied. For each category, we provide background and motivating examples, a general framework summarizing existing works and a detailed review of these works. We point out promising future research directions at the intersection of graph-structured data, counterfactual learning, and real-world applications. To offer a comprehensive view of resources for future studies, we compile a collection of open-source implementations, public datasets, and commonly-used evaluation metrics. This survey aims to serve as a “one-stop-shop” for building a unified understanding of graph counterfactual learning categories and current resources.

## Conclusion:

In this survey, we present a comprehensive review of counterfactual learning on graphs from the problems of counterfactual fairness, counterfactual explanation, counterfactual link prediction and the real-world applications of graph counterfactual learning. This is the first survey for counterfactual learning on graphs. In particular, we first introduce the basic concept of counterfactual learning, then introduce a framework to give a unified understanding of the problems. We also summarise the datasets and metrics used in each category. Then we go beyond the counterfactual learning on graphs to its applications in many areas, such as physical systems, medical, etc. Finally, we also discuss future directions and encourage domain experts to contribute to essential and urgent topics in this area. We believe this survey can give starters the fundamental knowledge and inspire the domain experts to solve the urgent challenges in this area.

**Insights from the Paper:**

# Causal Inference for Knowledge Graph Based Recommendation

[Paper Link](#)

## Abstract:

Knowledge Graph (KG), as a side-information, tends to be utilized to supplement the collaborative filtering (CF) based recommendation model. By mapping items with the entities in KGs, prior studies mostly extract the knowledge information from the KGs and inject it into the representations of users and items. Despite their remarkable performance, they fail to model the user preference on attribute in the KG, since they ignore that (1) the structure information of KG may hinder the user preference learning, and (2) the user’s interacted attributes will result in the bias issue on the similarity scores. With the help of causality tools, we construct the causal-effect relation between the variables in KG-based recommendation and identify the reasons causing the mentioned challenges. Accordingly, we develop a new framework, termed Knowledge Graph-based Causal Recommendation (KGCR), which implements the deconfounded user preference learning and adopts counterfactual inference to eliminate bias in the similarity scoring. Ultimately, we evaluate our proposed model on three datasets, including Amazon-book, LastFM, and Yelp2018 datasets. By conducting extensive experiments on the datasets, we demonstrate that KGCR outperforms several state-of-the-art baselines, such as KGNN-LS (Wang et al., 2019), KGAT (Wang et al., 2019) and KGIN (Wang et al., 2021).

## Conclusion & Future Scope:

In this work, we propose to model the fine-grained user preference on attribute to improve the recommender system. Towards this end, we construct the causal graph to resolve the challenges in a causal view, which is the first attempt in the KG-based recommendation to the best of our knowledge. Analyzing the constructed causal graph, we attribute the challenges into the spurious relation negatively affecting the prediction and develop a Knowledge Graph based Causal Recommendation model (KGCR) to address the problem. Specifically, we design the deconfounded user preference learning to model the user preference on attribute by removing the confounder between the user preference and her/his interacted attributes. Furthermore, we leverage the counterfactual inference to eliminate the bias misleading the attribute-based similarity scores of user-item pairs. Despite the promising performance, there remains some problem that should be explored in future work. For instance, the distribution of attributes over all items may cause popularity bias in the interaction prediction. And, how to distinguish the

causal-effect relation [49], [50] between entities in the KG and leverage it in the recommendation is also a challenging problem.

## **Insights from the Paper:**

# Data Augmented Sequential Recommendation Based on Counterfactual Thinking

[Paper Link](#)

## Abstract:

Sequential recommendation has recently attracted increasing attention from the industry and academic communities. While previous models have achieved remarkable successes, an important problem may still hinder their performances, that is, the sparsity of the real-world data. In this paper, we propose a novel counterfactual data augmentation framework to alleviate the problem of data sparsity. In specific, our framework contains a sampler model and an anchor model. The sampler model aims to generate high-quality user behavior sequences, while the anchor model is trained based on the original and new generated samples, and leveraged to provide the final recommendation list. To implement the sampler model, we first design four types of heuristic methods based on either random or frequency-based strategies. And then, to improve the quality of the generated sequences, we propose two learning-based samplers by discovering the decision boundaries or increasing the sample informativeness. At last, we build an RL based model to automatically determine where to edit the history behaviors and how many items should be replaced. Considering that the sampler model can be imperfect, we, at last, analyze the influence of the noisy information contained in the generated sequences on the anchor model in theory, and design a simple but effective method to better serve the anchor model. We conduct extensive experiments to demonstrate the effectiveness of our model.

## Conclusion & Future Scope:

In this paper, we propose to improve sequential recommendation by enriching the user behavior sequences based on the idea of counterfactual thinking. To achieve this goal, we design seven models including both heuristic and learning-based methods. We also analyze the sample complexity of the designed framework, and propose a simple but effective method to control the noisy information. In the experiments, we evaluate our framework based on nine real-world datasets to demonstrate its effectiveness and generality. In the future, there are still many potential directions to improve this work. To begin with, it is interesting to explore other methods for relaxing the intractable optimization targets. Then, our framework is mainly based on the user/item ID information, however, in modern recommender systems, much more side

information, such as user reviews and item images can be leveraged to profile the users and items. Thus, it will be valuable to extend our framework to incorporate more heterogeneous knowledge.

## **Insights from the Paper:**

# Counterfactual Explainable Conversational Recommendation

[Paper Link](#)

## Abstract:

Conversational Recommender Systems (CRSs) fundamentally differ from traditional recommender systems by interacting with users in a conversational session to accurately predict their current preferences and provide personalized recommendations. Although current CRSs have achieved favorable recommendation performance, the explainability is still in its infancy stage. Most of the CRSs tend to provide coarse explanations and fail to explore the impact of minimal alterations on the recommendation decisions on items. In this paper, we are the first to incorporate the counterfactual techniques into CRS and propose a Counterfactual Explainable Conversational Recommender (CECR) to enhance the recommendation model from a counterfactual perspective. Counterfactual explanations can offer fine-grained reasons to explain users' real-time intentions, meanwhile generating counterfactual samples for augmenting the training dataset to enhance recommendation performance. Specifically, CECR adaptively learns users' preferences based on the conversation context and effectively responds to users' real-time feedback during multiple rounds of conversation. Furthermore, CECR actively generates counterfactual samples to augment the training set and thus leading to a constant improvement in recommendation performance. Empirical experiments carried out on three benchmark datasets show that our CECR outperforms state-of-the-art CRSs in terms of recommendation performance and explainability.

## Conclusion & Future Scop:

In conclusion, we introduce a Counterfactual Explainable Conversational Recommender named CECR to enhance CRS using counterfactual techniques. CECR adapts to user preferences during conversations based on online feedback, improving recommendation performance. By generating counterfactual samples, CECR continuously improves by augmenting the training set, leading to further performance improvement. Extensive experiments validate its effectiveness and explainability. Future work can explore integrating large language models like ChatGPT with multimedia processing capabilities in CRSs. This integration enables CRSs to handle different types of data, such as text, images, and videos, resulting in more diverse and engaging recommendations across various media formats.



## Insights from the Paper: