HYPER-GRAPH ENHANCED CAUSAL RECOMMENDATION WITH GRAPH CONTRASTIVE NETWORKS FROM HYPERBOLIC SPACE IN MOOCS

By

LUO HAO

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

August 2024

FSKTM 2024 21

© COPYRIGHT UPM

All material contained within the thesis, including without limitation text, logos, icons, photographs, and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

HYPER-GRAPH ENHANCED CAUSAL RECOMMENDATION WITH GRAPH CONTRASTIVE NETWORKS FROM HYPERBOLIC SPACE IN MOOCS

By

LUO HAO

August 2024

Chairman : Nor Azura binti Husin, PhD
Faculty : Computer Science and Information Technology

Massive Open Online Courses (MOOCs) offer vast academic resources like online course but also introduce challenges such as information overload, making it difficult for users to find relevant content. Recommendation Systems (RS) aim to address this issue by predicting user preferences and suggesting suitable courses. However, traditional RS methods, such as collaborative filtering, content-based, and hybrid approaches, face limitations including data sparsity, cold-start problems, and an inability to effectively leverage graph information within MOOCs.

Graph-based models have been proposed to resolve these limitations, but they often struggle with sparsity issues, leading to biased recommendations, or generate graph noise due to improper contrasting pairs. Moreover, existing methods typically overlook the role of causal inference in user performance, which is crucial for providing personalized and meaningful recommendations in MOOCs.

© COPYRIGHT UPM

To address these challenges, we propose the Hypergraph enhanced Causal Recommendation framework with Hyperbolic Graph Contrastive Networks (HGCR-HGCN), which leverages hyperbolic graph contrastive networks to model user-user and concept-concept relationships more effectively. Our approach introduces hypergraph representations of user-concept interactions, employs both hyperbolic and Euclidean space views of graph structures, and optimizes recommendation accuracy through contrastive learning. Additionally, causal inference is incorporated to identify key factors influencing user performance.

The framework is evaluated against state-of-the-art baselines using a large-scale MOOC dataset, demonstrating improvements in recommendation accuracy, diversity, and novelty. Our findings suggest that HGCR-HGCN offers a robust solution to the problems of data sparsity, biased recommendations, and user performance analysis in MOOCs.

In summary, this research presents an innovative approach that combines hypergraph modeling, contrastive learning, and causal inference to deliver more accurate and personalized recommendations in MOOCs.

Keywords: Recommendation Systems, Graph Contrastive Networks, Hyper-Graph Modeling, Causal Inference

SDG: GOAL 4: Quality Education

© COPYRIGHT UPM

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

HYPER-GRAPH DIPERKASAKAN UNTUK CADANGAN SEBAB AKIBAT DENGAN RANGKAIAN KONTRASTIF GRAF DARI RUANG HIPERBOLIK DALAM MOOCS

Oleh

LUO HAO

Ogos 2024

Pengerusi : Nor Azura binti Husin, PhD
Fakulti : Sains Komputer dan Teknologi Maklumat

Kursus Dalam Talian Terbuka Besar-besaran (MOOC) menawarkan sumber akademik yang luas seperti kursus dalam talian tetapi turut memperkenalkan cabaran seperti lambakan maklumat, menyukarkan pengguna untuk mencari kandungan yang relevan. Sistem Cadangan (RS) bertujuan untuk menangani isu ini dengan meramalkan keutamaan pengguna dan mencadangkan kursus yang sesuai. Walau bagaimanapun, kaedah RS tradisional, seperti penapisan kolaboratif, berasaskan kandungan, dan pendekatan hibrid, menghadapi had seperti kekurangan data, masalah permulaan sejuk, dan ketidakupayaan untuk memanfaatkan maklumat graf dengan berkesan dalam MOOCs.

Model berasaskan graf telah dicadangkan untuk menyelesaikan had ini, tetapi sering menghadapi masalah kekurangan data, yang membawa kepada cadangan berat sebelah, atau menghasilkan bunyi graf akibat pasangan kontras yang tidak sesuai. Selain itu, kaedah sedia ada biasanya mengabaikan peranan inferens sebab-akibat dalam prestasi

© COPYRIGHT UPM

pengguna, yang penting untuk menyediakan cadangan yang diperibadikan dan bermakna dalam MOOCs.

Untuk menangani cabaran ini, kami mencadangkan rangka kerja Cadangan Sebaban Diperkasakan Hypergraf dengan Rangkaian Kontrastif Graf Hiperbolik (HGCR-HGCN), yang memanfaatkan rangkaian kontrastif graf hiperbolik untuk memodelkan hubungan pengguna-pengguna dan konsep-konsep dengan lebih berkesan. Pendekatan kami memperkenalkan perwakilan hypergraf interaksi pengguna-konsep, menggunakan pandangan ruang hiperbolik dan Euclidean bagi struktur graf, dan mengoptimumkan ketepatan cadangan melalui pembelajaran kontrastif. Selain itu, inferens sebab-akibat digabungkan untuk mengenal pasti faktor utama yang mempengaruhi prestasi pengguna.

Rangka kerja ini dinilai berbanding dengan asas terkini menggunakan set data MOOC berskala besar, menunjukkan peningkatan dalam ketepatan, kepelbagaian, dan kebaharuan cadangan. Penemuan kami mencadangkan bahawa HGCR-HGCN menawarkan penyelesaian kukuh kepada masalah kekurangan data, cadangan berat sebelah, dan analisis prestasi pengguna dalam MOOCs.

Kesimpulannya, penyelidikan ini mempersembahkan pendekatan inovatif yang menggabungkan pemodelan hypergraf, pembelajaran kontrastif, dan inferens sebab-akibat untuk memberikan cadangan yang lebih tepat dan diperibadikan dalam MOOCs.

Kata Kunci: Sistem Cadangan, Rangkaian Kontrastif Graf, Pemodelan Graf Hiper, Inferens Sebab

SDG: MATLAMAT 4: Kualiti Pendidikan

© COPYRIGHT UPM

ACKNOWLEDGEMENTS

First and foremost, I am extremely grateful to my supervisors, Dr. Nor Azura Husin for her invaluable continuous support, encouragement and patience in all the time of my academic research and daily life. I would like to also express my profound gratitude to my supervisory committee members, Assoc. Prof. Dr. Teh Noranis Mohd Aris, Dr. Mohd Yunus Sharum and Dr. Maslina Zolkepli for the guidance and support given along the research conduct. Their immense knowledge and plentiful experience have encouraged me all the time. Without their tremendous understanding and encouragement in the past four years, it would be difficult for me to complete my study.

I would also like to thank all the officers and staff in Faculty of Computer Science and Information Technology, UPM. It is their kind help and support that have made my study a wonderful time. Furthermore, a word of thank to all of my friends in UPM especially Sina Abdipoor; thank you for being great companions whilst this research is being carried out.

Finally, I would also like to express my deepest gratitude to my family and parents who have supported me emotionally and gave me encouragement along my study.

© COPYRIGHT UPM

This thesis was submitted to the Senate of the Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Nor Azura binti Husin, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Teh Noranis binti Mohd Aris, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Maslina binti Zolkepli, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Mohd Yunus bin Sharum, PhD
Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 13 March 2025

© COPYRIGHT UPM

TABLE OF CONTENTS

© COPYRIGHT UPM

© COPYRIGHT UPM

LIST OF TABLES

© COPYRIGHT UPM

LIST OF FIGURES

© COPYRIGHT UPM

LIST OF ABBREVIATIONS

MOOCs	Massive Open Online Courses
PR	Personalized Recommendations
RS	Recommendation system
CBF	Content-based Filtering
CF	Collaborative Filtering
CRS	Causal inference-based Recommendation System
SM	Statistical Methods
NN	Neural Network
CBN	Causal Bayesian Network
GNNs	Graph-based Neural Networks
RL	Reinforcement Learning
HGCN	Hyperbolic Graph Contrastive Networks
HGCR	Hypergraph enhance Causal Recommendation
HGCR-HGCN	Hypergraph Enhanced Causal Recommendation with Hyperbolic Graph Contrastive Networks
HR	Hit Ratio
NDCG	Normalized Discounted Cumulative Gain
MRR	Mean Reciprocal Rank
PR	Personalized Recommendations
CL	Contrastive Learning
CNN	Convolutional Neural Network
RNNs	Recurrent Neural Networks
GANs	Generative Adversarial Networks
LSTM	Long Short-Term Memory
DNN	Deep Neural Network

© COPYRIGHT UPM

HGAT	Hypergraph Attention Networks
SEM	Structural Equation Modeling
HDNNs	Hyperbolic Deep Neural Networks
BPR	Bayesian Personalized Ranking

© COPYRIGHT UPM

CHAPTER 1

INTRODUCTION

1.1 Background

MOOC (Massive Open Online Course) refers to an online learning platform that provides access to a wide range of courses and educational resources to a large number of users from all over the world. MOOCs offer users the opportunity to acquire new skills and knowledge, often at no cost, and at their own pace and convenience (Uddin et al., 2021). MOOCs data is a valuable resource for researchers and educators who are interested in improving the quality and effectiveness of online learning (Zhao et al., 2021). However, with the vast number of available courses and resources, users often face the challenge of finding the most appropriate and relevant ones that meet their individual needs and interests. Undoubtedly, the complex nature of MOOCs could lead to information processing overload for users, which could become a major bottleneck for online education (Tarus et al., 2017; Fan et al., 2019; Khalid et al., 2021).

To this end, previous research has often focused on utilizing recommender methods to reduce the dropout rates or procrastination behaviors in Massive Open Online Courses (MOOCs), while also recommending content such as courses and learning paths that users might find interesting (Gardner et al., 2019; Yao et al., 2020; Prenkaj et al., 2020). MOOCs, as a collective term, can be seen as a series of instructional videos, with each video containing various knowledge concepts. For example, in an Artificial Intelligence MOOC course, the videos may cover multiple concepts like "Machine Learning", "Deep Learning", "Neural Networks" and so on. Recently,

© COPYRIGHT UPM

scholars like (Gong et al., 2020) argue that directly recommending courses or videos to users often overlooks their specific interests in certain knowledge points. Piao (2021) also contends that the recommendation of courses and videos does not adequately consider user interests or learning needs, as a substantial number of knowledge concepts covered in courses and videos are not given due attention. In this study, the emphasis is on predicting and suggesting knowledge concepts that may capture the interest of users on MOOC platforms.

In addition, MOOCs data is heterogeneous and rich in graph-structured data (Zhu et al., 2020). The different types of entities and their relationships, such as users, courses, videos, and knowledge concepts, are expressed as a graph model shown in Figure 1.1. Each dashed circle in Figure 2.1 represents an identity of a course with many knowledge concepts, and each rectangle represents an online video. Student A is shown as learning knowledge concepts that are of interest to them, either because they are going to learn them or have already learnt them in previous times. Finally, effectively using heterogeneous graph-structured data to provide Personalized Recommendations (PR) remains a significant challenge (Wang et al., 2022).

© COPYRIGHT UPM

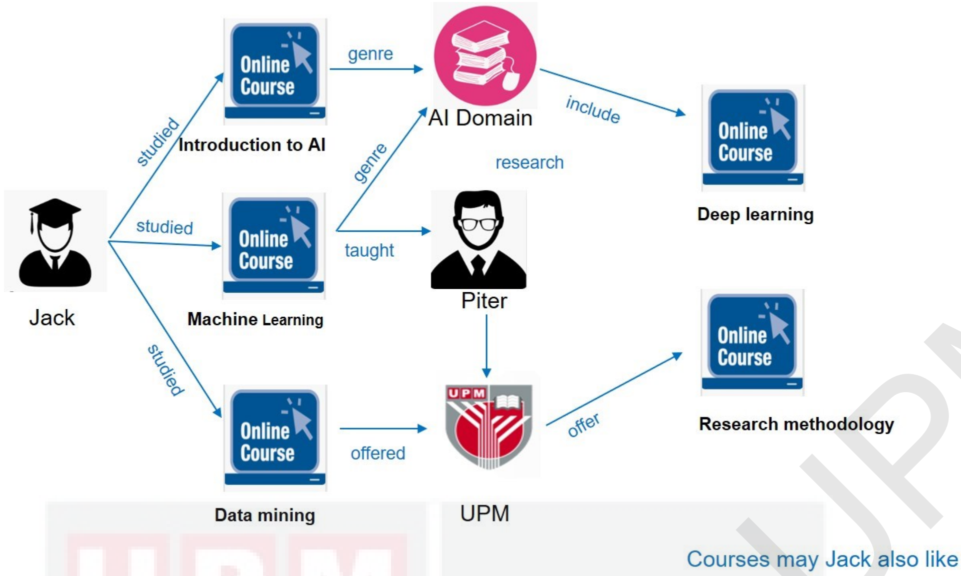

Figure 1.1 : The Heterogeneity Graph-Structure RS in MOOCs

A recommendation system (RS), on the other hand, refers to a technique used to provide personalized recommendations to users based on their individual preferences and behavior. In the context of MOOCs, recommendation systems (Zhang et al., 2019) play a crucial role in providing users with customized recommendations for courses, learning resources, and activities.

Recommendation systems are important for MOOCs for several reasons. Firstly, MOOCs offer a vast number of educational resources, and a personalized recommendation system can help users navigate this content more effectively by providing them with relevant and customized recommendations. This can help users save time and effort and enhance their overall learning experience (Ye et al., 2021). Secondly, personalized recommendation systems (Gulzar et al., 2018) can help improve the engagement and retention of users in MOOCs. By providing users with recommendations that are tailored to their interests and learning progress, users are more likely to remain engaged and motivated to continue learning. Finally, recommendation systems in MOOCs can also help improve the overall effectiveness of MOOCs by providing users with recommendations that are more aligned with their learning goals and preferences. This can help users achieve better learning outcomes and improve their overall satisfaction with the MOOC (Tian & Liu, 2021).Thus, recommendation systems play a critical role in enhancing the learning experience and outcomes of users in MOOCs, and are a key component of the MOOC ecosystem (Chanaa & El Faddouli, 2019).

Without a doubt, Recommendation System（RS）is the most widely used technology for providing personalized recommendations to users (Fu et al., 2016).RS are information filtering systems (Ma et al., 2017) that utilizes user's historical behavior and preference data to predict items or content that users may like (as shown in Figure 1.2). Arguably, RS are essential for online education, as they can help students discover relevant learning resources, personalize their learning experience, and improve their learning outcomes (Herath & Jayaratne, 2018). According to a study by Chen & Lee (2022), the adoption of recommendation systems significantly improves the user experience and engagement in MOOC platforms. In addition, recommendation systems have been shown to increase the retention rate of users and enhance their learning outcomes by providing relevant and personalized course recommendations (Assami et al., 2018).

© COPYRIGHT UPM

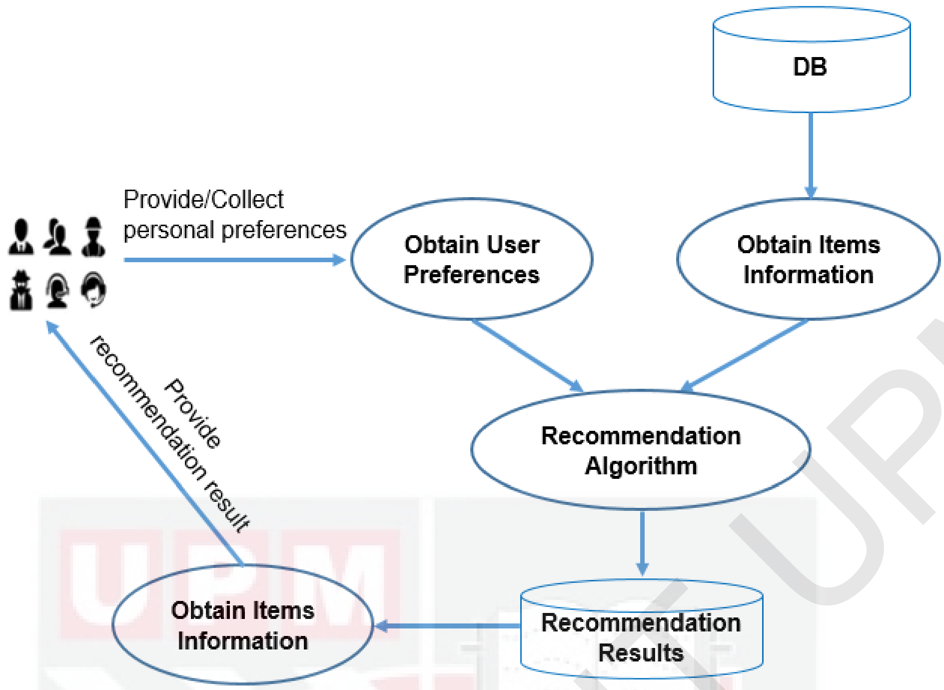

Figure 1.2 : Recommendation System General Model

Traditional RS approaches mainly include content-based filtering (CBF) (Son & Kim, 2017; Huang & Lu, 2018; Neamah & El-Ameer, 2018) , collaborative filtering (CF) (Zhang et al., 2016; Bozyi et al., 2018; Obeidat et al., 2019), and hybrid methods (Geetha et al., 2018; Cai et al., 2020; Riyahi & Sohrabi, 2020). More generally, these algorithms have shown advantages in capture users' preferences and provide personalized learning guidance based on user and item's explicit or implicit feedback, such as enrollment and viewing, to improve their learning performance (Xian et al., 2017). In addition, many traditional methods focus on use correlation in historical data to make predictions. For example, collaborative filtering (CF) methods use features and similarities between features to predict users' favorite items (Obeidat et al., 2019). However, there are many causal relationships in the real world, such as the relationship between the learning sequence of certain knowledge, that cannot be measured by similarity. Therefore, many researchers have proposed causal inference based recommendation systems (CRS), which extract complex causal relationships within data through causal reasoning to improve the performance of personalize recommendation (Xu et al., 2023；Zhu et al., 2023).

Most of the existing approaches for CRS are based on statistical methods (SM) and neural network (NN). SM for causal inference rely on regression analysis and experimental design to estimate the causal effect of an intervention on an outcome variable, by controlling for other factors that may influence the outcome variable. Another common technique of NN applied is causal Bayesian network (CBN) which represent causal relationships between variables as a directed acyclic graph, and use this graph to infer the causal effects of interventions. However, some unpreventable limitations from SM method such as cold-start issue (which happens when the user or item is new join the RS, SM for causal inference cannot establish causal relationships beyond the history data they are trained on) and data bias from NN technique (which lead to neural networks learning spurious causal relationships, resulting in inaccurate causal inferences) has called on academics to find alternative solutions.

Recently, the ability of graph-based neural networks (GNNs) to handle structured input and explore high-order information has made them the new state-of-the-art method for solving traditional reinforcement learning (RL) problems (Gao et al., 2022). GNNs can learn to represent users and items in a latent space that captures their causal relationships, which can then be used to make accurate recommendations even for new users and items, and to mitigate the effects of data bias (Norton et al., 2022). Besides, MOOCs' full with a vast of heterogeneous graph-structured data (Ye et al., 2021), GNNs can take advantage of graph-structured data to learn representations of

© COPYRIGHT UPM

users and items in hypergraph and hyperbolic spaces, leading to better recommendation results. In contrast, traditional neural networks use structured data but cannot handle graph-structured data, which contains rich information that can be leveraged for improved recommendations (Ming Chen & Zhou, 2020). Many advantages of the GNNs approach encourage this research to propose a novel approach for addressing the problem of processing heterogeneous graph-structured data in graph-based recommendation systems that incorporates causal relationships. Based on a thorough search of the relevant literature, it found no prior research on applying hypergraph learning to enhance causal recommendation based on hyperbolic graph contrastive networks (HGCN) in the MOOCs domain. Therefore, this study demonstrates the capability of the GNN approaches based on HGCN algorithm to learn the representations of items and users which is more effective in heterogeneous graph-structured data in CRS. Both the noise generation problem and the sparsity problem of user-concept interactions are addressed by the proposed HGCN method, which implements the graph representation on hyperbolic space. Furthermore, due to its ability to gather adequate semantic information from contrasting pairings in various spaces, the HGCN technique has the potential to significantly increase the quality of recommendations (Euclidean spaces and hyperbolic spaces). At last, it leverages hypergraph modeling to capture the complex causal relationships between users, courses, and other contextual factors to enhance the causal recommendation.

This research works focus on the development of <u>Hypergraph</u> enhance <u>C</u>ausal <u>R</u>ecommendation (HGCR) based on HGCN approach and introduce hypergraph learning and causal inference representation into the graph-based recommendation system. To explore the impact of hypergraph learning and causal inference on graph-

© COPYRIGHT UPM

based recommendation, it propose **HGCR-HGCN**, which enhances HGCN with hypergraph-based causal inference to identify the most relevant factors influencing user behavior and performance in MOOCs. Furthermore, in order to prove the importance of different components in the model and the impact of different hyperparameters on the model, it also conducted ablation experiments and sensitivity analysis. Lastly, comparison the proposed HGCR-HGCN approach to benchmark work by using the metrics of Hit Ratio (HR), Normalized Discounted Cumulative Gain (NDCG), and Mean Reciprocal Rank (MRR). The results show that the proposed approaches are effetely to recommend personized item to user.

1.2 Problem Statement

The popularity of Massive Open Online Courses (MOOCs) has grown significantly in recent years, providing users with access to a vast of educational resources. However, the large scale and diverse nature of MOOCs pose significant challenges for personalized recommendation systems, particularly in terms of accurately modeling the complex relationships between users, courses, and other relevant factors. Effectively using graph-structured heterogeneous MOOCs data to provide Personalized Recommendations (PR) remains a significant challenge. Furthermore, a lot of MOOC recommendation systems overlook user interests and learning requirements for particular knowledge concepts in favor of promoting courses and videos that cover a broad variety of knowledge concepts (Gong et al., 2020; Piao, 2021 ; Jiang et al., 2023).

The core research problem in improving MOOC recommendation systems involves enhancing the capability to provide accurate, diverse, and contextually appropriate

© COPYRIGHT UPM

course recommendations that align with individual user needs and learning behaviors. The challenge arises from several interconnected issues: (1) **Complex Relationships**. The dynamic and diverse nature of MOOC platforms results in complex interactions between users, courses, and various other factors that are difficult to model accurately with traditional methods. Data Sparsity and Cold-Start Problem: New courses and users continually entering the system exacerbate data sparsity and cold-start issues, leading to recommendations that may not effectively reflect user preferences or course relevance. (2) **Inadequacy of Current Graph-Based Methods**. While Graph Neural Networks (GNNs) offer potential solutions, they often assume homogeneous graph structures that fail to capture the intricate and heterogeneous relationships prevalent in MOOCs. (3) **Overfitting in Advanced Models.** Techniques like Contrastive Learning (CL), intended to address data sparsity, often introduce noise that can lead to overfitting, thus degrading the model's ability to generalize and accurately capture meaningful relationships. (4) **Neglect of Causality.** Most existing models focus on correlation rather than causation, ignoring the underlying causal relationships that influence user behaviors and preferences in educational contexts. The majority of conventional recommendation techniques rely on basic user- concept interaction data, like enrollment, clicking, watching, and bookmarking, to forecast or suggest the user's ideas, such as CF (Rabahallah et al., 2018), CB (Huang & Lu, 2019), or hybrid techniques (Geetha et al., 2018). However, these methods often stuck with personalized predict users' preferences and learning behaviors, making it challenging to provide accuracy and variety recommendations in light of the sparsity, cold start, and the problem of cannot effectively deal with graph (Medio et al., 2020a).

© COPYRIGHT UPM

The introduction of Graph Neural Networks (GNNs) has improved the performance of recommendation systems by learning more expressive user and concept representations. However, GNNs still suffer from limitations, particularly in addressing the data sparsity problem. In MOOCs, with an ever-expanding collection of courses and concepts, the limited interaction data available for new users or newly introduced courses leads to suboptimal recommendations. Additionally, conventional GNNs typically assume a homogeneous graph structure, failing to account for the complex, heterogeneous relationships in MOOCs, which can reduce the accuracy of recommendations.

Moreover, while contrastive learning (CL) has emerged as a powerful tool to combat data sparsity by leveraging self-supervised signals, it introduces new challenges. Methods like the InfoNCE model, which add noise to the embedding space to generate contrastive views, risk overfitting due to the introduction of uniform noise. This overfitting can degrade the model's ability to generalize and capture meaningful user-concept relationships, ultimately impacting recommendation performance.

Lastly, ignore the causal relationship between users and knowledge concepts. Current recommender systems deduce user preferences by analyzing data correlations, encompassing behavioral correlation in collaborative filtering, feature-feature relationships, and feature-behavior correlation in click-through rate prediction(C. Gao, Zheng, et al., 2022). Unfortunately, the real world is governed by causality rather than mere correlation, and it's crucial to recognize that correlation does not the same as causation. Moreover, most existing graph-based methods typically focus on the relationships between nodes in a graph, without explicitly considering the causal

© COPYRIGHT UPM

relationships between them, and also fail to reveal their mutual influence and function when making recommendation, resulting in the lack of interpretability of recommendation.

Hypothesis: We hypothesize that incorporating hypergraph modeling, causal inference, graph contrastive learning, and hyperbolic space representations will enhance recommendation accuracy and diversity by overcoming key issues like data sparsity, graph noise, and the lack of causal reasoning in current recommender systems.

To address these challenges, this paper proposes a novel approach, called Hypergraph Enhanced Causal Recommendation with Hyperbolic Graph Contrastive Networks (HGCR-HGCN), which combines hypergraph modeling, causal inference, graph contrastive learning, and hyperbolic representation space techniques to provide more accurate and effective recommendations for MOOC users.

1.3 Research Objectives

The primary objective of this study is to develop a highly personalized recommendation system for MOOCs that adeptly accounts for the individual preferences and diverse educational needs of users. Given the complex dynamics of user-course interactions and the influence of various contextual factors—such as user engagement patterns, course popularity, and peer effects—this research aims to intricately model these relationships to deliver tailored educational recommendations.

© COPYRIGHT UPM

To effectively address these challenges, this study proposes a structured approach centered around the following refined objectives:

1. Development of a Hypergraph-Based Recommendation Framework. To construct and assess a novel recommendation framework that utilizes hypergraph modeling combined with hyperbolic space representation. This framework is designed to capture the multi-dimensional and complex relationships between users and courses, thereby enhancing the accuracy and personalization of course recommendations.

2. Causal Inference in Recommendation Systems. To introduce and validate a causal inference approach that elucidates the complex causal relationships between users and educational concepts. This objective is pivotal and directly addresses the core research focus on causality, as emphasized in the revised problem statement. It seeks to provide a more scientifically grounded basis for generating precise recommendations by understanding not just correlations but the causative factors that influence user learning and course selection.

3. Graph Contrastive Learning for Enhanced Recommendations. To develop and implement a graph contrastive networks method that leverages different views of graph representations within embedding spaces. This method aims to enhance the robustness and quality of the recommendations by improving the model's ability to generalize across diverse user scenarios and to mitigate issues related to data sparsity and overfitting.

Each of these objectives contributes to the overarching goal of this research by sequentially addressing the identification of relevant factors, the development of an advanced theoretical framework, and the empirical testing and validation of the proposed models. This structured approach ensures that the research progresses logically from foundational concepts to practical applications, leading to innovations in the field of personalized educational recommendations. To accomplish the goal, this

© COPYRIGHT UPM

study takes the following steps:

i. To develop and evaluate a new knowledge concept recommendation framework that leverages hypergraph modeling and hyperbolic space representation to improve recommendation performance.

ii. To propose a new causal inference representation approach for capture the complex causal relationships among the user and concept to improve more precise recommendations.

iii. To propose a new Graph contrastive networks method as different views of graph representations embedding spaces to enhance the quality of recommendations.

1.4 Research Scope

This research focuses on the development and evaluation of a novel recommendation framework for Massive Open Online Courses (MOOCs). The proposed framework HGCR-HGCN involves the exploration of hypergraph modeling, causal inference, graph contrastive learning, and hyperbolic space representation techniques, with a focus on their application within MOOCs. The study aims to address challenges in existing recommendation systems, including data sparsity, cold-start problems, and the inability to utilize heterogeneous graph-structured information effectively. The goal is to enhance the personalization and effectiveness of course recommendations, thereby advancing online education by leveraging hyperbolic graph structures and causal relationships in recommendation models.

The research explicitly aims to develop and evaluate the HGCR-HGCN Framework: To provide a robust and comprehensive solution to the challenges faced in MOOC recommendation systems by leveraging hypergraph modeling for user-concept

© COPYRIGHT UPM

interactions, causal inference to establish meaningful relationships, and graph contrastive learning to enhance generalization and recommendation quality.

The scope of the thesis has been structured as follows:

- Recommendation Systems in MOOCs: Investigating the landscape of recommendation systems for MOOCs, highlighting their significance in enhancing the learning experience for users.
- Challenges in Current Recommendation Methods: Identifying limitations in traditional recommendation methods, including data sparsity, cold-start problems, and challenges with using heterogeneous graph-structured data.
- Hypergraph Modeling: Developing novel hypergraph modeling techniques to represent user-concept interactions within MOOCs, including constructing hypergraphs, feature extraction, and hyper-edge weighting.
- Causal Inference in Recommendation Systems: Focusing on the integration of causal inference techniques to establish meaningful causal relationships between users and learning concepts, ultimately leading to more targeted and precise recommendations.
- Graph Contrastive Learning: Introducing a graph contrastive learning framework to learn from positive and negative samples effectively. This approach will involve developing contrastive loss functions and training algorithms to optimize the graph neural network for better personalization.
- Utilizing Hyperbolic Space: Leveraging hyperbolic space representation for enhanced flexibility and effectiveness, focusing on its advantages over Euclidean spaces in representing hierarchical and complex user-course relationships.
- Evaluation and Validation: Conducting experimental evaluations on real-world MOOC datasets to validate the performance and effectiveness of the proposed HGCR-HGCN framework.

© COPYRIGHT UPM

- Contributions to Online Education: Demonstrating how the proposed framework and findings contribute to the field of recommendation systems in online education, with the potential to improve learning experiences and outcomes for MOOC participants.

1.5 Research Contribution

This research has made a positive contribution to the field of MOOCs recommendation. In particular, the collection, processing, and utilization of the heterogeneous graph structure data in MOOCs can help recommendation systems better mine users' personalized needs. The following are the contributions of this study, which can help improve the current practice of knowledge concept recommendation.

- A novel knowledge concept recommendation Framework that combines hypergraph modeling, causal inference, graph contrastive learning, and hyperbolic space representation techniques to provide more accurate and effective recommendations for MOOC users. Furthermore, the meta-path construction and hypergraph construction were firstly proposed to express the heterogeneous graph-structure data of MOOCs for create different types of relationship. This framework is not only potentially a major breakthrough in the development of personalized recommendation systems for MOOCs, but also has important implications for educational technology and online learning.

- A new causal relationship representation approach based on hypergraph learning. The proposed research leverages hypergraph based causal inference techniques to identify the most relevant factors influencing user behavior and performance in MOOCs. This can provide valuable insights into the factors that impact learning outcomes in MOOCs and inform the development of more effective recommendation algorithms.

© COPYRIGHT UPM

- A new Graph contrastive multi-view method as different views of graph representations from hyperbolic space and Euclidean space to enhance the quality of recommendations. This study firstly employs two embedding spaces: one hyperbolic and the other Euclidean. By utilizing their increased expressiveness, it produces a variety of perspectives of the input graph. These techniques can effectively capture the complex, multi-modal nature of MOOC data and provide a more accurate representation of the relationships between users, courses, and contextual factors.

Arguably, the combination of the new knowledge concept recommendation Framework, the new causal relationship representation, and graph contrastive multi-view model can further improve personalized recommendation in MOOCs in the future.

1.6 Organization of the Thesis

There are six chapters in the thesis, and each chapter has several sections and subsections. Chapter 1 serves as an introduction, providing an overview of the research background, outlining the primary issues within existing recommendation systems, particularly those related to MOOCs, and presenting the research objectives. The specific scope of the research and the significance of the contributions of this work are discussed in separate sections in Chapter 1.

Chapter 2 provides a comprehensive review of the related literature in the field of personalized recommendation systems for MOOCs. It discusses the existing recommendation techniques, including collaborative filtering, content-based filtering, and hybrid approaches. It also reviews the recent advances in hypergraph modeling, causal inference, graph contrastive learning, hyperbolic representation space, and their

© COPYRIGHT UPM

potential applications in enhancing personalized recommendation in MOOCs.

Chapter 3 presented an overview of the research frameworks, from research planning to the proposed method. Firstly, the research steps are briefly introduced. The following section detailed introduces the research framework and there five parts. Finally, this chapter also discusses the relevant software and hardware used in the MOOCs recommendation system.

Chapter 4 presents the experiments conducted on the proposed HGCR-HGCN approaches, including the framework of the algorithms and the experimental settings. This chapter focuses on the detailed description of the research framework, the method of HGCR-HGCN, and the setting of the experiment, such as the setting of hyperparameters. The summary of experimental setup is presented at the end of the chapter.

In Chapter 5, it presents a comprehensive overview of the experimental procedures for the HGCR-HGCN approaches with some important questions. A total of four distinct experiments were conducted, encompassing hyperparameter optimization, ablation studies, sensitivity assessments, and comparative analyses, all of which are elaborated upon in this chapter. The chapter culminates in a summary of the key findings derived from these experiments.

Chapter 6 summarizes the research objectives, contributions, and findings, and highlights the limitations and future directions of the proposed framework. It discusses the potential applications of the proposed framework in other educational contexts beyond MOOCs, and suggests possible directions for future research.

© COPYRIGHT UPM

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter provides an overview of key methodologies in recommendation systems and their limitations. Section 2.2 discusses classical approaches, including collaborative filtering and content-based methods, which are foundational but face challenges such as data sparsity and inability to model complex relationships. Section 2.3 explores hypergraph learning, highlighting its capacity to capture high-order interactions, though scalability and interpretability remain issues. Section 2.4 reviews graph contrastive learning, which enhances representations through self-supervision but depends on effective augmentations and hyperparameter tuning. Section 2.5 introduces causal inference methods aimed at mitigating biases but notes the complexity of defining causal structures. Section 2.6 examines hyperbolic representation spaces for modeling hierarchical data, with challenges in computational efficiency and integration. Finally, Section 2.7 summarizes the strengths and weaknesses of these approaches, underscoring the need for integrated solutions.

2.2 Classical Recommendation Approaches

Existing roadmap of recommendation system can be categorized into two main parts, one is the traditional methods like collaborative filtering, content-based filtering, and hybrid approach. Another one is the deep learning-based methods.

Traditional recommender systems have played a crucial role in personalizing content and product recommendations across various domains. Two of the most widely used

© COPYRIGHT UPM

techniques in this field are Collaborative Filtering (CF) and Content-Based Filtering (CBF). Despite their successes, these approaches face several limitations, prompting the development of hybrid systems and, more recently, deep learning-based models. In this section, we provide an overview of these traditional recommendation methods and discuss their associated challenges, as shown in Table 2.1.

A. Collaborative Filtering

Collaborative Filtering (CF) (Fan et al., 2019) is one of the most established methods in recommendation systems. It operates on the premise that users who have similar preferences in the past will also have similar preferences in the future (Onah & Sinclair, 2015). CF can be categorized into two main types: user-based (Zheng et al., 2018) and item-based filtering (Li et al., 2020). User-based CF recommends items by identifying users with similar tastes and preferences, while item-based CF uses item similarity to predict what the user might like (Patel & Patel, 2020).

Obeidat et al. (2019) proposed a collaborative recommendation system designed to suggest online courses based on similarities in students' course history. The system employs clustering techniques to group students with similar course selections, enhancing the quality of association rules generated. Specifically, it compared the Apriori algorithm's (Yabing, 2013) performance using the entire dataset versus its performance on clusters, finding that clusters improved the coverage of association rules. Additionally, they applied the algorithm to analyze course sequences, confirming the findings from Apriori regarding the importance of course dependency in recommendations. While the study demonstrates the benefits of clustering in generating more relevant recommendations, potential limitations include its reliance

© COPYRIGHT UPM

on pre-existing course history data, which may not address the cold start problem for new users effectively.

A hybrid evolutionary approach introduced by Mohammadpour et al. (2019) to enhance clustering in collaborative filtering recommender systems. Their method combines Genetic Algorithm (GA) and Gravitational Emulation Local Search (GELS) to iteratively optimize item clustering for offline collaborative filtering. The approach begins with a population of randomized solutions, which are refined using GA and further improved with GELS. Simulations on standard datasets reveal that while the hybrid method requires significant runtime, it improves key performance metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coverage. However, the increased computational demand highlights a trade-off between accuracy and efficiency that may limit its practicality in real-time systems.

Moreover, Xian et al. (2017) addressed privacy concerns in collaborative filtering by integrating differential privacy into the SVD++ model, a matrix factorization approach that enhances recommendation accuracy by incorporating implicit feedback. Recognizing that users' heightened awareness of data privacy impacts the quality and availability of data, the authors proposed three privacy-preserving methods: gradient perturbation, objective-function perturbation, and output perturbation. Their findings demonstrate that these approaches effectively protect user privacy while maintaining predictive accuracy. Additionally, they provide theoretical and empirical insights into balancing privacy protection strength and recommendation performance. However, the study may face challenges in scaling to large datasets or optimizing parameter selection for varying privacy requirements.

© COPYRIGHT UPM

A popular approach, proposed by Rosa et al. (2020), is the Cluster-based Local Similarity (CBLS) model, which aims to enhance the accuracy of neighborhood-based collaborative filtering by addressing the limitations of global similarity measures. Their method incorporates clustering algorithms (affinity propagation and K-means) to group similar items and build localized similarity models for each cluster, leveraging techniques such as rating normalization and resource allocation. Experiments conducted on benchmark datasets demonstrated that CBLS outperforms traditional and state-of-the-art collaborative filtering models in terms of accuracy metrics, such as mean absolute error (MAE) and root-mean-square error (RMSE). However, the study may face challenges in computational efficiency when applied to large-scale systems due to clustering overhead.

Despite its widespread use, CF faces several challenges. One significant issue is the cold start problem, which occurs when there is insufficient data for new users or items. Without past ratings or interactions, it becomes difficult to recommend items accurately (Çano & Morisio, 2017). Furthermore, CF systems often struggle with sparsity, as not all users rate all items, leading to incomplete user-item interaction matrices. This limits the system's ability to make accurate predictions, especially for niche items (Idrissi et al., 2019). Finally, CF is computationally expensive for large-scale systems, making it challenging to scale as the number of users and items increases (Wang et al., 2019).

B. Content-Based Filtering

Content-Based Filtering (CBF) approaches recommend items based on the attributes of the items and the user's previous interactions. The fundamental assumption in CBF

© COPYRIGHT UPM

is that users will likely prefer items that are similar to those they have liked in the past (Badriyah et al., 2018). CBF systems build a user profile by tracking the features of items a user has interacted with, such as categories or keywords, and then recommend items with similar characteristics (Shah et al., 2016).

Patra et al. (2020) proposed a content-based recommendation system aimed at improving the reusability of datasets by connecting datasets from Gene Expression Omnibus (GEO) with relevant biomedical literature. The study addresses the weak linkage between datasets and related literature, which hampers scientific productivity. Using an information retrieval approach, the authors generated semantic features from MEDLINE article titles and abstracts and evaluated methods such as TF-IDF, BM25, Latent Semantic Analysis, Latent Dirichlet Allocation, word2vec, and doc2vec for similarity scoring. The best-performing model, BM25, achieved high precision metrics in recommending articles. Novel re-ranking and normalization techniques further improved results. Despite its effectiveness, the study acknowledges that this work is an initial step and highlights the potential for enhancing dataset reuse through refined recommendations.

In addition, Son and Kim (2017) introduced a novel content-based filtering (CBF) approach utilizing multiattribute networks to improve recommendation system performance. Their method calculates item similarities by considering multiple attributes and analyzing both direct and indirect connections in a network. The approach incorporates centrality and clustering techniques to capture mutual item relationships and structural patterns, ensuring diverse recommendations and mitigating common issues like sparsity and over-specialization. Experiments using

© COPYRIGHT UPM

MovieLens data demonstrated that the proposed method outperformed traditional approaches in accuracy and robustness. Additionally, it effectively addresses the cold start problem by relying solely on user-specific historical ratings, making it less dependent on external data. However, future work may explore extending its applicability to more dynamic datasets or hybrid systems.

Furthermore, Huang and Lu (2018) addressed the issue of information overload on MOOC platforms and the limitations of basic course search and classification functions. They proposed a content-based recommendation model for MOOCs, using the "iCourse" platform as a case study. Their method emphasizes the use of detailed course descriptions to enhance recommendation precision. Experimental results demonstrated that the model significantly outperformed random recommendations, with accuracy improving as the comprehensiveness of course descriptions increased. While effective, the study highlights the need for further exploration into improving descriptive data quality to enhance recommendation performance.

While CBF can overcome the cold start problem for items by recommending new items based on their attributes, it faces challenges related to overspecialization (Rahman et al., 2022). This occurs when the system recommends only items that are similar to those the user has already interacted with, potentially leading to a lack of diversity in recommendations (Salehi & Kmalabadi, 2012). Additionally, content-based systems rely heavily on accurate item descriptions, which may not always be available or complete for all items (Kumar & Jadon, 2024).

C. Hybrid Approach

© COPYRIGHT UPM

To overcome the limitations of CF and CBF, hybrid recommender systems combine the strengths of both approaches. Hybrid approach can alleviate the cold start problem associated with CF by incorporating item attributes from CBF, while they mitigate the overspecialization issue of CBF by incorporating user interaction data from CF (Geetha et al., 2018).

Yadav et al. (2024) investigated the effectiveness of recommendation systems (RS) across various industries, focusing on content-based, collaborative, and hybrid approaches. Their study involved both qualitative and quantitative analyses to evaluate these methods, emphasizing the unique importance of qualitative metrics for movie recommendation systems. Results showed that hybrid systems consistently outperformed standalone methods in terms of recommendation accuracy and relevance. However, the paper does not deeply explore the computational trade-offs or implementation challenges of hybrid approaches, which could provide additional insights for practical application.

The research by Cai et al. (2020) introduces a hybrid recommendation model that incorporates many-objective optimization to enhance recommendation systems (RS) by addressing multiple user-centric factors such as accuracy, diversity, novelty, and coverage. By combining three different recommendation techniques, the model improves robustness and user satisfaction. The authors employed a many-objective evolutionary algorithm (MaOEA) to solve the model and conducted extensive experiments to evaluate its performance. The results demonstrate that the proposed approach effectively balances accuracy with diversity and novelty, providing more comprehensive recommendations. However, the study does not deeply address

© COPYRIGHT UPM

potential computational challenges in implementing MaOEA for real-time applications.

In addition, Rawaa and Rojalina (2024) proposed a hybrid e-learning recommender system called CFISAR, which integrates collaborative filtering models (CFMs) with fine-grained sentiment analysis (FSA) to enhance recommendations on educational platforms. The system leverages user reviews and ratings to build personalized learner profiles, using CFMs to capture latent factors and FSA to refine content rankings based on sentiment. A novel mechanism for introducing new items into the system is also presented, where initial ratings are predicted using FSA. Experimental results highlight the system's effectiveness, with Asymmetric SVD achieving a low MSE of 0.699 and the Peephole algorithm demonstrating superior accuracy of 92.64%. Despite these promising outcomes, the system's reliance on deep learning and word embeddings may pose computational challenges for large-scale real-time applications.

Hybrid recommendation systems offer several advantages and challenges. They enhance accuracy by combining multiple techniques like collaborative and content-based filtering, addressing the limitations of individual methods (Çano & Morisio, 2017). They improve diversity and novelty by balancing these aspects with accuracy, providing a broader range of recommendations. Additionally, they effectively address data sparsity and cold-start problems, ensuring better performance for new users or sparsely rated items. Hybrid systems are adaptable, integrating dynamic features like fine-grained sentiment analysis to ensure continuous relevance. They also optimize recommendations for domain-specific needs, providing highly contextual suggestions.

© COPYRIGHT UPM

However, hybrid methods face challenges such as increased complexity in implementation, requiring the integration of various algorithms and substantial computational resources (Phalle et al., 2024). The computational overhead from training complex models can be significant. Potential overfitting may arise without proper validation, particularly in limited datasets. These systems are data-dependent, necessitating high-quality and diverse data, which might not always be available. Scalability also poses a challenge, as maintaining real-time performance with increasing data volume is demanding. Despite these challenges, thoughtful design and optimization are crucial for realizing the full potential of hybrid systems.

D. Deep Learning Approach

In recent years, deep learning has become one of the most influential methodologies in the field of artificial intelligence and machine learning. It functions on the premise of learning hierarchical feature representations from data through neural networks with multiple layers (Rane et al., 2024). Deep learning algorithms are broadly categorized into three main types based on their architecture and application areas. Convolutional Neural Networks (CNNs) are primarily used for image and video processing, where they leverage convolutional layers to automatically learn spatial hierarchies of features, making them highly effective for tasks like image classification and object detection (Lecun et al., 2023). Recurrent Neural Networks (RNNs) are designed for sequential data, excelling in time series analysis, natural language processing, and speech recognition by retaining information over time. Long Short-Term Memory (LSTM) networks, a specialized RNN, handle long-range dependencies more effectively (Hochreiter & Schmidhuber, 1997). Generative Adversarial Networks (GANs) consist of two networks, a generator and a

© COPYRIGHT UPM

discriminator, which are trained in opposition, allowing GANs to generate realistic images, videos, and other media (Mahdizadehaghdam et al., 2019). These three types of deep learning algorithms cater to different problem domains: CNNs for spatial data, RNNs for sequential data, and GANs for data generation.

Wang et al. (2020) address the challenge of effectively recommending suitable courses to learners in the context of Massive Open Online Courses (MOOCs). The proposed method utilizes an attention-based convolutional neural network (CNN) to personalize course recommendations. The authors first represent learner behaviors and learning histories as feature vectors. Then, an attention mechanism enhances the relevance estimation by minimizing discrepancies between predicted and actual user ratings during the neural network training phase. The final trained model is used to recommend the top-n courses to learners. The framework of the recommendation system is also introduced. Experimental results highlight the method's potential to improve recommendation accuracy and user satisfaction. However, the paper does not provide a detailed evaluation or comparison with other state-of-the-art recommendation approaches, leaving room for further research on the system's scalability and applicability in diverse MOOC contexts.

Moreover, Zhang et al. (2018) proposed a recommendation model that combines collaborative filtering (CF) algorithms with deep learning techniques to address challenges such as data sparsity and limited scalability in traditional CF methods. The model consists of two key components: a feature representation approach based on a quadric polynomial regression model, which improves latent feature extraction by enhancing matrix factorization, and a deep neural network (DNN) model that predicts

© COPYRIGHT UPM

user rating scores using these latent features as input. The model was evaluated on three public datasets and demonstrated improved recommendation performance compared to existing algorithms. However, the study does not delve into the computational efficiency or potential limitations of scalability when applying the model to larger datasets, leaving room for further investigation.

Another research proposed by Chen et al. (2024) investigates the use of deep neural networks (DNNs) to automatically classify the semantic content of MOOC course reviews, addressing the challenges posed by the traditional manual coding of unstructured textual data. Utilizing a dataset of 102,184 reviews from 401 MOOCs, the study develops DNN models, with a focus on recurrent convolutional neural networks (RCNNs), to categorize semantic features of the reviews. The findings demonstrate that RCNNs perform effectively in identifying and learning semantic categories, significantly reducing the manual workload and improving the efficiency of feedback analysis. This approach allows for timely insights into learner experiences, enabling course providers to implement targeted interventions to enhance instructional design. However, the study does not explore potential scalability issues or comparative analysis with alternative classification methods, suggesting directions for future research.

In addition, Ahmadian Yazdi et al. (2022) propose a dynamic educational resource recommender system that integrates advanced deep learning models, including Multi-Layer Perceptron (MLP), BiLSTM, and LSTM, enhanced by an attention mechanism. The study addresses the challenge of guiding students to suitable learning resources within web-based educational systems, which often struggle to match the effectiveness

© COPYRIGHT UPM

of traditional classrooms. Unlike previous approaches relying on DBN networks, which focus on users' recent interests, the proposed system considers both current and long-term user preferences. The results demonstrate superior performance, achieving an accuracy of 0.96 and a loss of 0.0822, thereby outperforming similar methods in delivering personalized recommendations. However, the study does not elaborate on computational efficiency or scalability, highlighting areas for future exploration.

In summary, traditional recommendation techniques such as collaborative filtering and content-based filtering have been foundational in the development of recommender systems. However, they each face significant challenges, including cold start issues, sparsity, scalability, and overspecialization. Hybrid approaches have attempted to address some of these limitations, but these systems also introduce complexity and scalability concerns. With the advent of deep learning, new models have been developed that offer promising solutions, particularly in terms of capturing complex patterns in user behavior. Nonetheless, the challenges associated with data sparsity, cold start, and scalability remain critical obstacles for the effective deployment of recommendation systems in large-scale applications. Future research should focus on overcoming these challenges, particularly in the context of dynamic, real-time recommendation systems.

© COPYRIGHT UPM

Table 2.1 : Summarizes the Previous Research Work in Classical Recommendation Approaches

Articles	Algorithm	Contributions	Limitations
(Obeidat et al., 2019)	CF	Enhanced course recommendations by clustering students with similar course histories, improving the coverage of association rules; analyzed course sequences to highlight dependency in recommendations.	Relies on existing course history data, making it ineffective for new users (cold start problem).
(Mohammadpour et al., 2019)	CF	Optimized item clustering for collaborative filtering, improving accuracy metrics such as MAE and RMSE while enhancing coverage; demonstrated the effectiveness of combining GA and GELS for iterative optimization.	High computational demand limits practical application in real-time systems.
(Rosa et al., 2020)	CF	Addressed global similarity limitations in collaborative filtering by developing localized similarity models using clustering (affinity propagation, K-means); improved accuracy metrics (MAE, RMSE) over state-of-the-art methods.	Clustering overhead could impact computational efficiency in large-scale systems.
(Patra et al., 2020)	CBF	Linked GEO datasets with biomedical literature to improve dataset reusability; novel re-ranking techniques enhanced results.	Initial step with limited exploration; potential for refinement to maximize dataset reuse impact.
(Son & Kim, 2017)	CBF	Calculated similarities using multiattribute networks; integrated centrality and clustering for diverse, accurate recommendations; tackled sparsity and cold start problems effectively.	Limited applicability to dynamic datasets; potential for extending to hybrid recommendation systems.
(Huang & Lu, 2019)	CBF	Improved recommendation precision on the "iCourse" platform using detailed course descriptions; demonstrated significant accuracy improvements with more comprehensive data.	Dependent on the quality of descriptive data; further work needed to enhance data comprehensiveness.

© COPYRIGHT UPM

Table 2.1 : Continued

(Yadav et al., 2024)	Hybrid Method	Investigated RS across industries; demonstrated that hybrid methods outperform standalone approaches in accuracy and relevance.	Lacked in-depth analysis of computational trade-offs and implementation challenges of hybrid methods.
(Cai et al., 2020)	Hybrid Method	Addressed user-centric factors (accuracy, diversity, novelty, coverage); utilized MaOEA for balancing trade-offs; improved robustness and user satisfaction.	Did not thoroughly explore computational challenges in applying MaOEA in real-time systems.
(Rawaa Alatrash, Rojalina Priyadarshini, 2024)	Hybrid Method	Integrated CFMs and FSA for enhanced e-learning recommendations; proposed a novel method for new item introduction; achieved high accuracy and low error rates.	Computational demands due to deep learning and word embeddings limit scalability for real-time use.
(Wang et al., 2020)	CNN	Personalized MOOC course recommendations using learner behavior and history vectors; enhanced relevance estimation with attention mechanism; improved accuracy and user satisfaction.	Lacked detailed comparison with state-of-the-art methods; scalability and applicability issues in diverse MOOC contexts not addressed.
(Zhang et al., 2018)	DNN	Addresses data sparsity and scalability challenges in CF. Demonstrated improved recommendation performance on three public datasets compared to existing methods.	Lacks detailed discussion on computational efficiency. Scalability issues with larger datasets not fully addressed.
(Chen et al., 2024)	RNN	Automated semantic content classification of MOOC course reviews; reduced manual workload and enabled timely insights into learner experiences for targeted instructional improvements.	Scalability issues and comparative analysis with other classification methods not explored.
(Ahmadian et al., 2022)	LSTM	Dynamic educational resource recommendations; incorporated both current and long-term preferences; achieved high accuracy (0.96) and low loss (0.0822).	Computational efficiency and scalability not elaborated, limiting real-world application insights.

© COPYRIGHT UPM

2.3 Hypergraph Learning

Hypergraph learning, as an extension of traditional graph structures, has emerged as a powerful paradigm for representing and analyzing complex, high-dimensional relationships in data. Unlike conventional graphs, which are limited to pairwise interactions modeled through edges, hypergraphs introduce hyperedges, allowing connections among multiple nodes simultaneously (Kim et al., 2024). This flexibility enables the modeling of higher-order interactions, where relationships among entities extend beyond simple pairwise connections, capturing intricate data dependencies that traditional graphs fail to represent effectively.

The motivation behind hypergraph learning lies in its ability to represent overlapping, hierarchical, and multi-way relationships in diverse contexts. For instance, in social networks, hypergraphs can model group interactions rather than just individual relationships, offering a richer representation of social dynamics (Sun et al., 2023). Similarly, in biological systems, hypergraphs effectively capture the multifaceted interactions among genes or proteins, which often involve groups rather than pairs. These capabilities make hypergraphs particularly suitable for domains where high-dimensional and complex relational structures are inherent (Antelmi et al., 2023).

The adaptability of hypergraph frameworks extends to hierarchical representations, allowing nodes and hyperedges to encapsulate data at different levels of granularity (Zhang et al., 2023). This makes them particularly advantageous for tasks requiring nuanced representations, such as in natural language processing for document and sentence structures or in computer vision for multi-object interactions within images (Liao et al., 2022). The unique ability of hypergraphs to bridge local and global relationships within data highlights their pivotal role in advancing the understanding and analysis of complex systems.

In deep learning, hypergraphs serve as a powerful framework to model intricate relationships in diverse and complex datasets. Unlike traditional graphs, hypergraphs can represent higher-order interactions by allowing hyperedges to connect multiple nodes simultaneously, thus capturing relationships that are overlapping, hierarchical, or multi-faceted. This versatility makes them particularly valuable for domains such as social networks, where group dynamics and collective behaviors are key, biological systems that demand the modeling of multi-protein or gene interactions, and natural language processing tasks, where word and sentence structures often exhibit overlapping relationships (Lawson et al., 2024). Additionally, hypergraphs offer a robust mechanism to encapsulate both local and global structures within data, enabling more comprehensive and nuanced analysis compared to traditional graph-based approaches. Their ability to integrate various levels of data abstraction enhances their applicability across a wide range of machine learning tasks (Kim et al., 2024).

Several methodologies have been developed to harness the power of hypergraphs in deep learning, focusing on both local and global structural learning. Hypergraph Convolutional Networks (HGCN) proposed by Yadati et al. (2019) employs convolutional filters designed specifically for hypergraphs, effectively capturing local structural features and interactions around each node. This approach models localized relationships and enhances node representations in hypergraph-structured data. Another method, HGAT (Hypergraph Attention Networks) utilizes attention mechanisms to prioritize the most relevant nodes and hyperedges, focusing on

© COPYRIGHT UPM

capturing global relationships within the hypergraph. By learning weighted connections, HGAT provides flexibility in analyzing varying interaction strengths (Chen et al., 2020). The next method HGNN (Hypergraph Neural Networks) combines the strengths of HGCN and HGAT by integrating both local neighborhood information and global structural relationships. This dual approach enables a more comprehensive representation of node and edge features, making HGNN suitable for a wide range of applications, from social network analysis to biological data modeling (Feng et al., 2019).

Hypergraphs offer significant advantages in modeling higher-order relationships, overlapping structures, and hierarchical interactions, which traditional graph methods struggle to capture. Their adaptability makes them particularly valuable for tasks requiring sophisticated and nuanced data representations, such as social network analysis and biological data modeling (H. Tian & Zafarani, 2024). However, hypergraph learning faces notable challenges. Designing convolution filters that manage multi-node hyperedges without escalating computational complexity is a key hurdle (Arya et al., 2024). Scalability poses another obstacle, as the high computational cost limits applications to large datasets (Raghavendra, 2024). Additionally, the effective selection of hyperedges, especially in scenarios with an overwhelming number of potential connections, remains an unresolved challenge critical to improving model efficiency and accuracy. These challenges underline the need for further advancements in hypergraph learning methodologies.

To address the challenges of hypergraph learning, researchers have developed several innovative strategies. Higher-order polynomial filters and learnable aggregations,

© COPYRIGHT UPM

such as Long Short-Term Memory (LSTM) networks, enable efficient feature integration by capturing intricate relationships across multi-node hyperedges (Gao et al., 2022). To improve scalability, techniques like hypergraph sparsification reduce computational complexity while preserving structural integrity, and distributed training frameworks facilitate processing on large-scale datasets (Zhang et al., 2022). For dynamic hypergraphs, incremental learning and meta-learning schemes provide adaptability, allowing models to efficiently update with new data and robustly respond to changes in hypergraph structures. These solutions collectively enhance the efficiency, scalability, and robustness of hypergraph learning.

Hypergraph learning represents a powerful and evolving field in deep learning, providing a robust framework for analyzing complex data relationships. By overcoming existing challenges and refining methodologies, hypergraphs hold the potential to drive innovation across domains, including natural language processing, computer vision, and social network analysis. This chapter presents a comprehensive overview of hypergraph learning, summarizing its definition, methods, applications, advantages, challenges, and potential solutions.

2.4 Graph Contrastive Learning

Recent advancements in Graph Contrastive Learning (GCL) (Zhong et al., 2021, Liu et al., 2022, Cai et al., 2023) have introduced a variety of methods aimed at improving the learning of graph representations through self-supervised learning (SSL) (Liu et al., 2023, Liu et al., 2023) techniques. Traditional GCL methods, such as Deep Graph Infomax (DGI) (Veličković et al., 2019), use graph convolutional networks (GCNs) to maximize mutual information between subgraph patch representations and high-

© COPYRIGHT UPM

level graph summaries. DGI has shown competitive performance in node classification tasks but remains limited by its generalization to more complex graph tasks and large graphs. In contrast, InfoGraph (Sun et al., 2020) extends GCL by maximizing mutual information between graph-level representations and substructure representations at various scales, outperforming state-of-the-art methods in graph classification and molecular property prediction. The introduction of InfoGraph further enhances its semi-supervised learning capabilities, offering competitive performance in semi-supervised tasks.

To capture higher-order information, H-GCL (Zhu et al., 2023) utilizes hypergraphs, contrasting representations from both the original graph and its hypergraph view, showing improved performance in node and graph classification. This approach also uses a diffusion model to bridge the gap between these two views, although scalability remains a challenge. Similarly, TP-GCL (Li et al., 2024) transforms graphs into hypergraphs using clique expansion and high-order adjacency tensors, significantly improving generalization and performance on sparse labeled data. However, further exploration is needed to assess its scalability to larger datasets.

Other recent methods have focused on the robustness and flexibility of GCL. For example, GROC (Jovanović et al., 2021) investigates adversarial robustness in GCL by introducing adversarial transformations, including edge insertion and removal, which increases robustness against attacks. On the other hand, GraphCL (Zhu et al., 2023)and SimGRACE (Xia et al., 2022) aim to simplify and improve the efficiency of GCL. GraphCL employs various graph augmentations to achieve superior generalizability and robustness, while SimGRACE eliminates the need for traditional

© COPYRIGHT UPM

data augmentations, instead using two encoders for contrasting perturbed views of the graph, making it more efficient.

Moreover, GraphMAE (Hou et al., 2022) introduces a masked graph autoencoder that focuses on feature reconstruction rather than graph structure, addressing training robustness issues in generative SSL for graphs. Extensive experiments show that GraphMAE outperforms both contrastive and generative baselines across multiple datasets. However, its scalability and broader applicability require further research.

In summary, while significant progress has been made in GCL, the methods vary in their approaches to capturing higher-order structural information, robustness, and scalability. Techniques like hypergraph-based GCL (H-GCL and TP-GCL) and multi-view contrastive learning (GraphCL and SimGRACE) enhance the expressiveness of graph representations, while others like DGI, InfoGraph, and GraphMAE focus on optimizing mutual information and improving the training process, which is shown in Table 2.2. Despite these advancements, challenges remain in scaling these methods to larger datasets and improving their generalization to a wider range of graph-related tasks.

Table 2.2 : Summarizes the Previous Research Work in Graph Contrastive Learning

Articles	Algorithm	Contribution	Limitation
(Wang et al., 2021)	Hypergraph Contrastive Learning (**H-GCL**)	Introduces hypergraphs to capture higher-order structural information, improving performance in node and graph classification tasks. Uses diffusion models for better alignment.	Scalability and effectiveness in other graph-related tasks are yet to be explored.
(Veličković et al., 2019)	Deep Graph Infomax (**DGI**)	Maximizes mutual information between subgraph patch representations and high-level summaries, showing competitive performance in node classification without random walks.	Generalization to complex tasks and scalability to very large graphs need further research.
(Monti et al., 2017)	**GROC**	Enhances robustness through adversarial transformations, including edge removal and insertion, improving robustness against adversarial attacks.	Needs more validation across diverse datasets and applications.
(Berg et al., 2017)	**InfoGraph**	Maximizes mutual information between graph-level representations and substructures at different scales. InfoGraph extends to semi-supervised settings. Outperforms baselines.	Scalability and generalization to other graph tasks require further investigation.
(Ying et al., 2018)	Multi-View Graph Representation Learning (**MVGRL**)	Contrasts structural views of graphs, showing that contrasting first-order neighbors and graph diffusion outperforms multi-scale encodings. Sets new state-of-the-art results.	Performance may not benefit from additional views or multi-scale encodings, suggesting diminishing returns.
(Chen et al., 2023.)	Tensorized Hypergraph Contrastive Learning (**TP-GCL**)	Transforms regular graphs into hypergraphs with high-order adjacency tensors, improving generalization and handling sparse labeled data.	Further exploration needed on scalability and performance across different graph tasks.
(He et al., 2020)	Graph Contrastive Learning (**GraphCL**)	Uses four types of graph augmentations, achieving competitive generalizability, transferability, and robustness across multiple settings.	Full potential may require refining augmentation parameters and incorporating advanced GNN architectures.

© COPYRIGHT UPM

Table 2.2 : Continued

(Xiao Wang et al., 2019)	Simple Graph Contrastive Learning (**SimGRACE**)	Uses two encoders to generate correlated views without traditional data augmentations, offering flexibility and efficiency.	Applicability to complex graph tasks and scalability to larger datasets need further investigation.
(Gonget al., 2020a)	Graphical Mutual Information (**GMI**)	Measures mutual information between input graphs and their high-level representations. Invariant to isomorphic transformations and can be efficiently estimated.	Further exploration needed for scalability and broader application to graph tasks.
(Hu & Rangwala, 2019)	Graph Masked Autoencoder (**GraphMAE**)	Focuses on feature reconstruction rather than graph structure, improving robustness during training and outperforming both contrastive and generative baselines.	Scalability and application to a broader range of graph tasks need further research.

© COPYRIGHT UPM

2.5 Causal Inference

Causal inference has emerged as a pivotal area of study within graph-based deep learning, offering novel approaches to uncover causal relationships within complex datasets (Wei et al., 2022). Graph-structured data, prevalent in domains such as social networks, biological systems, and recommendation systems, necessitates advanced methods to model and understand the underlying causal mechanisms that drive interactions and dependencies within these systems (Sha et al., 2021). By incorporating causal reasoning into graph-based models, researchers can move beyond static data patterns and gain deeper insights into the dynamic forces that shape real-world phenomena. This enhanced understanding of causality enables more accurate predictions and facilitates more informed decision-making across diverse applications.

Several key approaches are employed in graph-based causal inference to address the challenges of representing and reasoning about causal relationships. Structural Equation Modeling (SEM) defines equations to capture causal interactions within the graph, estimating parameters to predict node effects while controlling for confounding variables (Arif et al., 2024). Farnell & Mirra (2023) proposed Causal Bayesian Networks utilize directed acyclic graphs (DAGs) to infer conditional probabilities of nodes based on their parent relationships, providing a framework for elucidating the causal structure of the system. Finally, Counterfactual reasoning techniques model hypothetical outcomes of interventions, providing valuable insights into potential changes within the graph's structure and the resulting consequences (Mu & Li, 2024).These methodologies collectively enable robust analysis and interpretation of graph-structured data, addressing biases and confounding factors to reveal the true causal dependencies within the system.

© COPYRIGHT UPM

Causal graph learning finds valuable applications across various domains. In recommendation systems, it facilitates the prediction of user behavior and preferences by analyzing historical data and leveraging graph structures to represent users, items, and their interactions (Vuong, 2023). By understanding the causal relationships between user preferences, item characteristics, and contextual factors, these systems can generate more personalized and effective recommendations, improving user engagement and satisfaction (Y. Ma & Tresp, 2021). Beyond recommendation systems, causal inference in graphs finds applications in understanding complex biological networks, modeling disease progression, and simulating the potential outcomes of interventions in various fields (Liao et al., 2022).

Despite its significant promise, causal inference in graph-based deep learning faces several challenges. Scalability remains a critical concern, as large datasets and complex relationships can significantly increase computational demands, necessitating the development of more efficient and scalable models (Gupta et al., 2020). Furthermore, effective negative sample selection within large datasets, particularly in recommendation systems, poses a significant bottleneck (Zhu et al., 2023). Addressing these challenges requires innovative solutions, such as advanced sampling strategies, including synthetic data generation, causal transfer learning, and multitask learning, which can enhance model robustness and generalization (Harianja & Lumbantoruan, 2019). Integrating self-supervised learning with causal inference can also yield more powerful graph models, further advancing the field (Xu et al., 2023).

© COPYRIGHT UPM

Graph-based deep learning for causal inference is a rapidly evolving field with transformative potential across diverse disciplines. By addressing the current challenges and refining existing methodologies, researchers can unlock deeper insights into causal mechanisms, driving innovation in recommendation systems, complex system modeling, and beyond. Future research should focus on improving scalability, enhancing causal representation methods, and exploring interdisciplinary applications to maximize the impact of this promising approach.

2.6 Hyperbolic Representation Space

Hyperbolic space, a non-Euclidean geometry with constant negative curvature, has emerged as a promising alternative to Euclidean space in deep learning (Chamberlain et al., 2017). Its inherent flexibility and ability to effectively model hierarchical or tree-like data structures have driven significant advancements across various domains, including natural language processing (NLP), representation learning, and graph-based modeling (Liang et al., 2022). Traditional deep learning approaches heavily rely on Euclidean geometry, which, while simple and intuitive, often struggles to represent complex, hierarchical relationships within data (Chami et al., 2019). This limitation hinders the effective modeling of intricate data interactions observed in domains requiring nuanced representations, such as social networks or hierarchical ontologies. In contrast, hyperbolic space offers a more suitable framework for these scenarios, enabling the representation of data with inherent hierarchy and complex relationships more efficiently (Qi et al., 2019; Guo et al., 2021). Its negative curvature allows for more compact representations, facilitating efficient data modeling and potentially improving the performance of deep learning models.

© COPYRIGHT UPM

Hyperbolic deep neural networks (HDNNs) have gained popularity recently because they can give high fidelity embeddings with few dimensions from deep representations in the hyperbolic space, particularly for data with hierarchical structure (Peng et al., 2022). A hyperbolic neural architecture of this kind is rapidly applied to various scientific domains, including as computer vision, graph embedding, and natural language processing. The encouraging outcomes show how much better it is than its counterpart in Euclidean space in terms of capability, significant compactness of the model, and physical interpretability.

Hyperbolic neural networks involve using hyperbolic space as the underlying geometry for deep neural networks, allowing for more flexible representations of complex data (Yang et al., 2022). Hyperbolic embeddings involve mapping high-dimensional data into hyperbolic space, allowing for more efficient and accurate representation learning. Hyperbolic attention networks involve using attention mechanisms in hyperbolic space, allowing for more accurate and efficient modeling of complex relationships and interactions between entities.

Hyperbolic space has several advantages over Euclidean space in the context of deep learning (Qi et al., 2019). One of the key advantages is its ability to represent hierarchical and tree-like structures more accurately, which is especially important in natural language processing and other fields where data has a hierarchical or tree-like structure. Another advantage is its ability to represent complex data with fewer dimensions, reducing the computational complexity of deep learning models. Hyperbolic space also has the ability to capture complex relationships and interactions between entities that are not easily captured by Euclidean space (Guo et al., 2021). For

© COPYRIGHT UPM

example, hyperbolic space can represent groups of entities that interact with each other in complex ways, such as in social networks or biological systems.

While hyperbolic spaces show promise for modeling graph-structured data more effectively than traditional deep learning methods, there remain several challenges that warrant further research attention. Addressing these outstanding issues will help realize the full potential of hyperbolic deep learning.

One key challenge is evaluation methodology (Chami et al., 2019). Unlike Euclidean domains, there exists no agreed-upon benchmark datasets and tasks to rigorously and consistently assess the performance of hyperbolic models. The development of standardized evaluation protocols is needed to objectively measure progress in the field. Secondly, optimization in non-Euclidean geometries like hyperbolic spaces is more complex than in flat vector spaces (Guo et al., 2021). Existing optimization techniques from Euclidean deep learning do not always translate readily. More work is required to derive optimization methods tailored for hyperbolic embeddings that stabilize training and converge efficiently. A third issue relates to scalability (Xu et al., 2022). As graphs modeling real-world systems grow exponentially in size, current hyperbolic neural networks do not scale gracefully to massive graph diameters. approaches capable of accommodating vast graphs with millions of nodes are urgently needed. At last challenge in hyperbolic space-based deep learning is the selection of hyperparameter (Dai et al., 2021). In other words, how to select the curvature and dimensionality of the hyperbolic space that will provide the most accurate and efficient representation of the data. This issue is particularly important for applications such as natural language processing, where the structure of the data is complex and difficult

© COPYRIGHT UPM

to capture with traditional Euclidean space-based models.

To address this challenge, researchers are exploring different techniques for hyperparameter optimization, such as hypergraph neural networks (Li et al., 2021), hyperbolic contrastive learning (Yue et al., 2023) and hyperbolic meta-path learning(Luo et al., 2023). These techniques aim to find the optimal hyperparameters for the hyperbolic space-based deep learning model, while minimizing the computational cost and time required for optimization.

Hyperbolic space provides a powerful framework for learning representations of complex relationships and interactions in high-dimensional data (Qi et al., 2019). By developing new methods and algorithms that can address the challenges and bottlenecks, researchers can unlock the full potential of hyperbolic space in deep learning. This will enable new discoveries and insights into the complex relationships and interactions that underlie many important phenomena, and will open up new avenues for research and innovation in fields ranging from natural language processing to social network analysis (Li et al., 2021).

In conclusion, hyperbolic space presents a transformative paradigm in deep learning, offering substantial advantages over Euclidean geometry for modeling hierarchical and complex relationships. By addressing the existing challenges, researchers can unlock its full potential, driving innovation across various domains, including NLP, graph analysis, and social network modeling. Continued advancements in optimization techniques, scalable architectures, and hyperparameter tuning will be crucial for the broader adoption and impact of hyperbolic deep learning. Future

© COPYRIGHT UPM

research should also focus on exploring interdisciplinary applications, leveraging the compactness and interpretability of hyperbolic space to address novel challenges in diverse fields.

2.7 Summary

Traditional methods, such as collaborative filtering and content-based filtering, struggle to effectively handle the heterogeneous and high-dimensional nature of MOOC data. CF suffers from cold-start issues, data sparsity, and scalability challenges, while CBF often results in overspecialization, limiting recommendation diversity. These limitations necessitate exploring advanced methods to better model complex relationships and user preferences.

Graph-based approaches have made significant strides in capturing complex interactions in recommendation systems. However, these methods often rely heavily on labeled data, limiting their ability to generalize to unlabeled or sparse datasets. While graph contrastive networks (GCNs) offer potential solutions to label scarcity and robustness issues, they face challenges such as overfitting and the need for effective negative sample selection. Additionally, hypergraph-based learning, which can capture higher-order relationships, is constrained by computational inefficiencies and difficulties in hyperedge selection and convolution operations, which limit its broader application in large-scale educational platforms.

Emerging techniques such as causal inference and hyperbolic graph contrastive networks provide promising avenues for addressing these gaps. Causal inference can effectively model cause-and-effect relationships, alleviating cold-start and sparsity

© COPYRIGHT UPM

issues by analyzing user behavior at a deeper level. Meanwhile, hyperbolic space representations allow for more nuanced modeling of hierarchical or tree-like structures, offering improved scalability and performance in sparse data scenarios. However, challenges remain in seamlessly integrating these advanced methods into a unified framework. Future research should focus on developing scalable architectures, optimizing computational efficiency, and validating these approaches on diverse MOOC datasets to bridge the identified gaps and push the boundaries of online course recommendation systems.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter proposes a technique for enhancing causal recommendation in MOOCs using hypergraphs and graph contrastive networks from hyperbolic space. The aim of this study is to improve the internal relation representation of complex graph structure in MOOC data. The concept of hypergraph modeling is introduced to capture the complex relationships among various entities in MOOCs, representing these relationships as hypergraphs, a better understanding of the causal dependencies between different elements can be achieved, leading to more accurate recommendations. Section 3.2 provides an elaborate explanation of the research steps, while Section 3.3 offers a succinct overview of the hardware and software utilized, followed by a chapter summary.

3.2 Research Steps

This section provides an overview of the research steps involves five main phases: pre-processing, feature extraction, hypergraph construction, the proposed method, classification and evaluation. The overview of the research steps are represented as in Figure 3.1.

© COPYRIGHT UPM

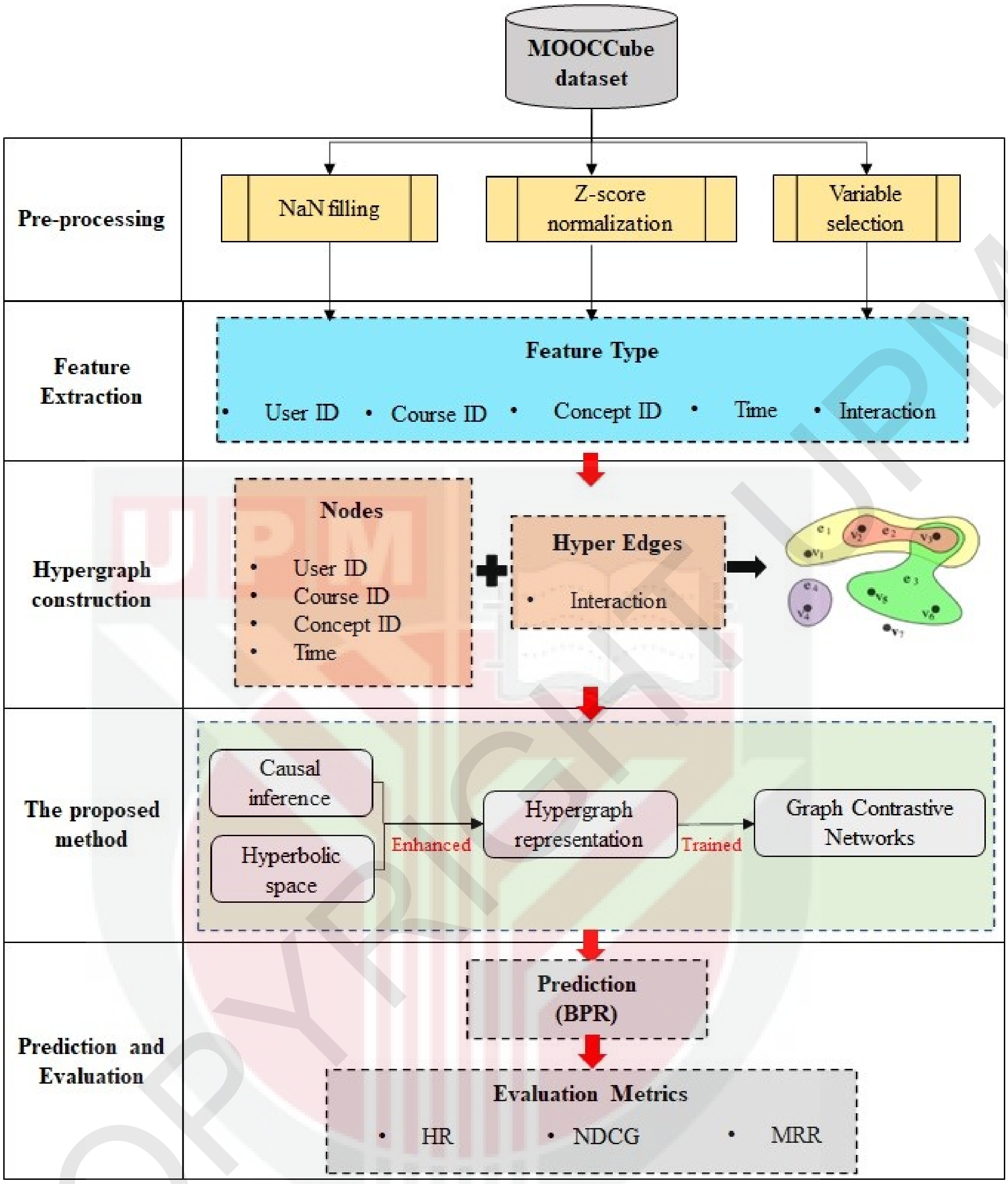

Figure 3.1 : Research Steps. This is the MOOC Recommendation System Workflow. The ground truth refers to the actual items a user has interacted with or chosen, hypergraph refers to how users, courses, and concepts are linked based on various types of relationships. Besides, embeddings are generated as part of the "Hypergraph representation".

3.2.1 MOOCCube Dataset

MOOCs (Massive Open Online Courses) have emerged as a popular form of online learning (Suresh & Srinivasan, 2020), providing users with access to a vast number of courses and educational resources from around the world. As the popularity of MOOCs has grown, so has the need for large-scale data warehouses to store and analyze the vast amount of data generated by MOOC users. One such data warehouse is MOOCCube (Yu et al., 2020) an open-source platform designed to serve research related to MOOCs. The data repository is now available at http://moocdata.cn/data/MOOCCube.

According to Yu et al., (2020), MOOCCube is one of the biggest and most varied collections of educational materials. It includes specific data about student behavior, like the length of time spent learning, the frequency of learning, and the breaks between learning videos. Nearly 5 million student video viewing records, including information from almost 200,000 pupils, are included in the database. 2,005 individuals, 21,037 concepts, 600 courses, 22,403 videos, 137 schools, 138 teachers, and their relationships make up the dataset, which offers a rich and varied source of information for study and analysis. Of particular interest to researchers and educators is the ability to analyze user behavior and make personalized recommendations for learning. With over 930,553 interactions between users and concepts, MOOCCube provides a wealth of data for modeling user behavior and developing effective recommendation systems. By selecting specific courses and videos for processing, researchers can analyze user behavior and provide personalized recommendations based on individual preferences and learning progress as shown in Table 3.1.

Overall, MOOCCube plays a critical role in advancing research and understanding in the field of MOOCs, providing researchers and educators with a powerful tool for analyzing user behavior and developing personalized recommendation systems to

© COPYRIGHT UPM

enhance the learning experience and outcomes of MOOC users.

Table 3.1 : Details of the MOOCCube Dataset for Our Research

Entities	Statistics	Relations	Statistics
users	2,005	**user-concept**	930,553
concepts	21,037	**user-course**	13,696
courses	600	**course-video**	42,117
videos	22,403	**teacher-course**	1875
schools	137	**video-concept**	295,475
teachers	138	**course-concept**	150,811

Researchers in course recommendation heavily rely on the MOOCCube dataset, a comprehensive open-source resource developed by Yu et al., (2020), This large-scale data warehouse goes beyond just courses; it offers valuable insights into student behavior, learning concepts, relationships between entities, and even external resources. This wealth of information allows researchers to develop more sophisticated and effective course recommendation algorithms (T. Liu et al., 2022). In contrast to analogous current educational resource databases, it functions on a grander scale and offers a greater variety and depth of data. The database includes information on 2,005 users, 21,037 concepts, 600 courses, 22,403 videos, 137 schools, 138 teachers, and their interrelationships. The wealth of information contained within the datasets, makes it an invaluable resource for researchers and educators alike. The dataset includes 858,072 interactions in the training set and 72,481 in the test set, for a total of 930,553 interactions between users and concepts. Table 3.2 provides a comprehensive overview of the dataset's statistics, which clearly demonstrate the breadth and complexity of the information that it has collected.

© COPYRIGHT UPM

Table 3.2 : Summary of Dataset with Number of Each Important Features for Training and Testing

Datasets	Statistics	Datasets	Statistics
users	2,005	**concepts**	21,037
courses	600	**videos**	22,403
schools	137	**teachers**	**138**
	Total interactions:	**930,553**	
Training set	**858,072**	Testing set	**72,481**

3.2.2 Pre-Processing

The primary goal of the first step is to prepare the dataset for use in the recommendation model by cleaning, transforming, and organizing. It involves several methods which include NaN filling(Stepanek, 2020), Z-score normalization(Fei et al., 2021), and Variable selection(Andersen & Bro, 2010). NaN filling is a data imputation technique used to replace missing values in a dataset with a suitable substitute. Z-score normalization is a data standardization technique that transforms each data point into a dimensionless score representing its distance from the mean in units of standard deviation. Variable selection is a statistical method that identifies a subset of relevant variables from a larger set of variables. It is used to improve the performance of statistical models.

3.2.3 Feature Extraction

In this phase, the features were selected and extracted. The extracting the relevant features from the MOOCCube dataset for constructing a hypergraph based on the interactions between users, courses, and concepts. Specifically, it extracts the following features from the dataset: user ID, course ID, concept ID, time stamp, and interaction type (e.g., view, comment, like). The feature extraction has ability to provide detailed information about user behavior. This contains data on how long

© COPYRIGHT UPM

users spend studying, how often they access course materials, and how long it takes them to complete each learning task. Such data provides valuable insights into how users engage with course material and can be used to inform the development of new and more effective learning strategies.

3.2.4 Hypergraph Construction

The next phase is to construct a hypergraph based on the feature extraction of interactions between users, concepts, and courses in the MOOC dataset. The hypergraph is constructed with users, courses, and concepts as nodes, and interactions as hyper-edges that connect the nodes. Construct a hypergraph by identifying nodes representing MOOC elements, creating edges linking related nodes, assigning weights reflecting relationship strength, and visualizing the graph to uncover patterns and relationships. The Hypergraph construction details will be further discussed in Chapter 4.

3.2.5 The Propose Method

This phase proposes a technique for enhancing causal recommendation in MOOCs using hypergraphs and graph contrastive networks from hyperbolic space. The proposed method involves two main steps: Firstly, the techniques of causal inference and hyperbolic space are used to enhance hypergraph representation. Secondly, the hypergraph representation embedding is used to training graph contrastive networks (GCNs). The propose method details will be further discussed in Chapter 4 subsection 4.3.

© COPYRIGHT UPM

3.2.6 Prediction and Evaluation

The final phase reported the result of the proposed methods. In this phase, the prediction process was run by using Bayesian Personalized Ranking (BPR). BPR (Rendle et al., 2009), a technique derived from Bayesian analysis. The core idea behind BPR is to model user preferences based on pairwise comparisons of items. In the training phase, the algorithm learns latent factors for users and items by optimizing a pairwise ranking objective function. Once trained, BPR can be used for predictions by estimating the preference score for a user-item pair. This is typically achieved by computing the dot product of the latent factor vectors associated with the user and the item. The higher the dot product, the higher the predicted preference score, allowing for the ranking of items to make personalized recommendations. At last, the results were evaluated in terms of HR, NDCG, and MRR. The results evaluation details will be further discussed below.

In addition to the commonly employed evaluation metrics that traditional RS used are discussed earlier in Chapter 2, various metrics have been introduced to assess the performance of prevalent models, including Hit Ratio (HR) (Gong et al., 2021), Normalized Discounted Cumulative Gain (NDCG) (Pradhan & Pal, 2019), and Mean Reciprocal Rank (MRR) (L. Yang et al., 2018). The subsequent sections will provide detailed explanations of these three widely utilized evaluation metrics, serving as the evaluation models for this study.

As stated in Equation 2.1, HR@k is the percentage of pertinent concepts in the test set that fall inside the top-k concepts of the recommendations (Gong et al., 2021).

© COPYRIGHT UPM

$$\text{HR@k} = \frac{1}{N}\sum_{u}\mathbb{I}(|R_u \cap T_u|) \quad \text{(Equation 2.1)}$$

To evaluate the performance of the recommendation model, it utilizes the Normalized Discounted Cumulative Gain (NDCG) metric at rank k (NDCG@k). NDCG@k considers the ranking positions of relevant concepts within a test dataset of size N. Here, I(x) represent an indicator function, assigning a value of 1 if x is greater than 0, and 0 otherwise (W. Wang, 2022). The specific calculation of NDCG@k is detailed in Equation 2.2.

$$\text{NDCG@k} = \frac{1}{Z}\sum_{j=1}^{k}\frac{2^{\mathbb{I}(|R_u \cap T_u|)}-1}{\log_2(j+1)} \quad \text{(Equation 2.2)}$$

where Z is a normalization factor that represents the score attained by an ideal top-k ranking. When positive notions are stated in reciprocal rankings, their mean reciprocal rank (MRR) is calculated (Z. Li et al., 2021), which can be expressed as in Equation 2.3.

$$MRR = \frac{1}{N}\sum_{1}^{N}\frac{1}{rank_i} \quad \text{(Equation 2.3)}$$

3.3 Hardware and Software Requirement

To evaluate the performance of our personalized MOOC recommendation system, we conducted a series of experiments using Python software. The experiments ran on a 64-bit Linux Professional machine equipped with an NVIDIA A100 GPU and an Intel Xeon processor clocked at 2.20 GHz, with 16.00 GB of RAM. It's important to note that while the experiments leveraged Python, diacritic recognition itself relied on a separate, established algorithm.

© COPYRIGHT UPM

3.4 Summary

This chapter outlines the research methodology employed for feature extraction and hypergraph construction of MOOCs datasets. Chapters 4 will delve into the detailed methods for hypergraph enhanced causal recommendation (HGCR) and hyperbolic graph contrastive networks (HGCNs). Additionally, this chapter discusses the data utilized in this research and the hardware and software requirements for conducting the study. The chapter concludes with a discussion of the performance measures that are used to assess the precision and quality of the suggested techniques.

© COPYRIGHT UPM

CHAPTER 4

HYPERGRAPH ENHANCED CAUSAL RECOMMENDATION FRAMEWORK WITH HYPERBOLIC GRAPH CONTRASTIVE NETWORKs (HGCR-HGCN)

4.1 Introduction

The proposed HGCNs approach for Hypergraph enhanced Causal Recommendation named as HGCR-HGCN will be presented in this chapter. The algorithm is basically analogous to the GCN in (Cai et al., 2023), while GCNs adopts multi-view method instead of the traditional Euclidean single-view method. Meta-path structure and hypergraph structure will be used as the input of GNN to implement graph embedding by learning the representation of the relationships between user-user, concept-concept, and user-concept. Subsequently, by leveraging hyperbolic and Euclidean spaces as embedding spaces, this multi-view framework creates many views of the input graph by utilizing their enhanced expressiveness. The causal inference is used to enhanced the hypergraph representation which will then generate the recommendation item for each user by using GCNs. The effectiveness of the recommendation rank list is evaluated using a variety of metrics.

This chapter introduces the HGCR-HGCN framework and its methodologies. Section 4.2 outlines the framework and its algorithm, detailing how hypergraph convolution and causal reasoning are integrated. Section 4.3 describes the step-by-step workflow of the HGCR-HGCN method, emphasizing its approach to capturing complex interactions. Section 4.4 elaborates on the hyperparameter optimization objective, focusing on balancing performance and computational efficiency. Section 4.5 presents the experimental design, including dataset selection, Pre-processing step, Feature

© COPYRIGHT UPM

extraction, prediction, and performance measure. Finally, Section 4.6 summarizes the chapter, highlighting the methodological contributions and experimental rigor of HGCR-HGCN.

4.2 Framework of HGCR-HGCN

Hypergraph Enhanced Causal Recommendation with Graph Contrastive Networks from Hyperbolic Space in MOOCs is a research study that proposes a new methodology for improving recommendation systems in MOOCs by using hypergraphs and deep learning techniques. The study aims to address the limitations of conventional recommendation systems by introducing hypergraphs, causal inference, and hyperbolic space, which allow for a more flexible and expressive representation of relationships between entities.

To tackle the sparsity issue in user-concept interactions and address noise generation concerns, the proposal suggests implementing graph representation in hyperbolic space, building on recent advancements in geometric graph mining within hyperbolic spaces. This approach utilizes graph contrastive targets to project the graph representations into an angular space, significantly boosting their ability to distinguish between similar and dissimilar graphs. The framework functions as a multi-view system, employing both hyperbolic and Euclidean spaces as embedding spaces. This combination leverages the enhanced expressiveness of both spaces to generate diverse perspectives of the input graph.

In addition, to better learn the rich interaction information in heterogeneous graphs, it constructed a hypergraph structure and causal inference to learn graph representations, to address the problems of graph structure data not being well utilized and causal relationships being ignored.

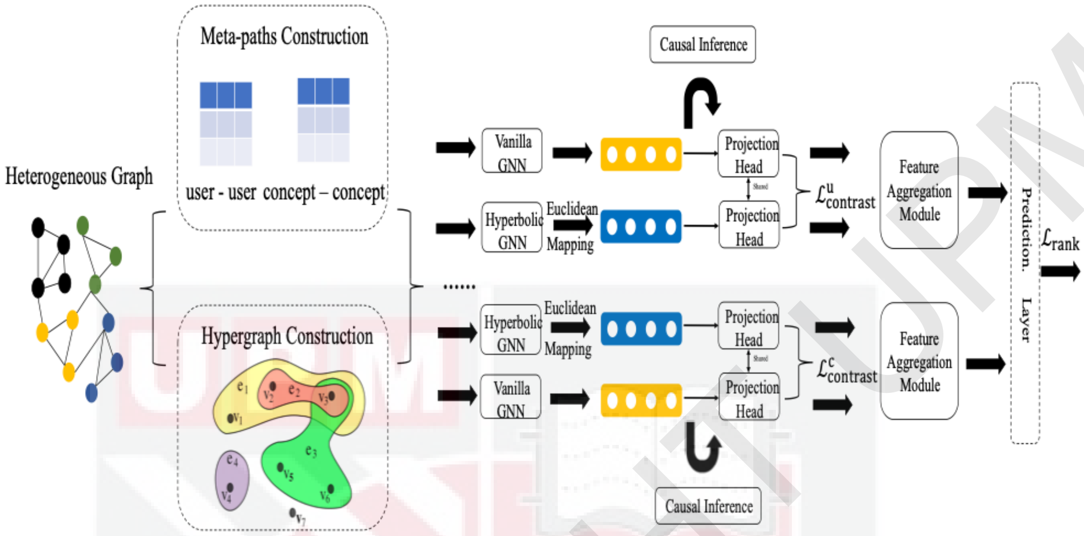

Figure 4.1 : The framework of HGCR-HGCN

The HGCR-HGCN framework involves several main structures, including Meta-path construction, Hypergraph construction, and Causal inference as shown in Figure 4.1. The system utilizes a heterogeneous graph that connects users A and courses X in MOOC. Meta-path structure will be used as the input of GNN to implement graph embedding by learning the representation of the relationships between user-user, concept-concept. The hypergraphs are constructed to represent relationships between user-concept, and employs various types of GNN, including Vanilla GNN, Hyperbolic GNN, and Euclidean Mapping to learn embeddings. These different types of methods are along with shared projection heads and feature aggregation modules. The prediction layer uses a causal inference approach to make recommendations based on the learned embeddings from the hypergraph by using GNN. Finally, the framework also includes Graph Contrastive Networks (GCNs), which is used to further improve the performance of the recommendation system. This technique involves the comparison of different views of the graph structure to learn more robust and discriminative representations of entities. By leveraging contrastive learning, the framework can improve its ability to distinguish between similar entities and provide more accurate recommendations. The algorithm of proposed HGCR-HGCN is shown in Algo. 4.1.

Algorithm 1 Algorithm for HGCR-HGCN

Require: MOOCs data as a heterogeneous graph G
Ensure: MOOCs prediction output y_{pred}
Initialize $\textsc{MetaPaths}()$ as an empty list.
$\{\text{MetaPaths} \leftarrow \textsc{ConstructMetaPaths}(G)\}$
for nodes $u \in V$ of type t_u, nodes $v \in V$ of type t_v **do**
  if $\textsc{IsConnected}(u, v, G)$ **then**
    Find a path $p \leftarrow \textsc{FindPath}(u, v, G)$
    if $p \neq$ null **then**
      Add p to MetaPaths list
end for
Hypergraph $\leftarrow \textsc{ConstructHypergraph}(G)$
for path $\in$ MetaPaths **do**
  if path.type = 'Vanilla' **then**
    $X_{\text{vanilla}} \leftarrow \textsc{VanillaGNN}(\text{path}, G)$
    $Z_{\text{vanilla}} \leftarrow \textsc{ProjectionHead}(X_{\text{vanilla}})$
  else if path.type = 'Hyperbolic' **then**
    $X_{\text{hyperbolic}} \leftarrow \textsc{HyperbolicGNN}(\text{path}, G)$
    $X_{\text{euclidean}} \leftarrow \textsc{EuclideanMapping}(X_{\text{hyperbolic}})$
    $Z_{\text{hyperbolic}} \leftarrow \textsc{ProjectionHead}(X_{\text{euclidean}})$
  $Z \leftarrow Z \cup \{Z_{\text{vanilla}}, Z_{\text{hyperbolic}}\}$
end for
function $\textsc{ContrastiveLearning}(Z_{\text{vanilla}}, Z_{\text{hyperbolic}})$
  $n \leftarrow$ number of elements in Z_{vanilla} and $Z_{\text{hyperbolic}}$
  Initialize Loss $L \leftarrow 0$
  for $i \leftarrow 1$ to $nandj \leftarrow 1$ to n **do**
    if $i \neq j$ **then**
      $s_{\text{pos}} \leftarrow \textsc{Similarity}(Z_{\text{vanilla}}[i], Z_{\text{hyperbolic}}[j])$
      $s_{\text{neg}} \leftarrow \textsc{Similarity}(Z_{\text{vanilla}}[i], Z_{\text{vanilla}}[j])$
      $L \leftarrow L + \textsc{Loss}(s_{\text{pos}}, s_{\text{neg}})$

$L \leftarrow \frac{L}{n(n-1)}$
$Z_{\text{agg}} \leftarrow \textsc{FeatureAggregationModule}(Z)$
$\{Z_{\text{causal}} \leftarrow \textsc{CausalInference}(Z_{\text{agg}})\}$
Identify the set of confounders C
Adjust for C using a statistics (e.g., regression)
Estimate the causal effect of Z_{agg} on the label Y
Perform sensitivity analysis to assess the robustness
Return the causal effect estimate Z_{causal}
$y_{\text{pred}} \leftarrow \textsc{PredictionLayer}(Z_{\text{causal}})$
return y_{pred}

The first step of the algorithm involves preparing the MOOC data by cleaning and transforming it into a format compatible with the HGCR-HGCN framework. This

© COPYRIGHT UPM

involves representing the MOOC data as a heterogeneous graph, which serves as the input, and generating a ranked list of MOOCs that the algorithm predicts the learner will be most interested in. Before constructing metapaths, initialize the list of metapaths to be empty (Line 1). Then, construct metapaths using the heterogeneous graph as input (Line 2). The construction of metapaths occurs from Line 3 to Line 4. Meta-paths are used to model complex relationships between nodes in a graph. If a path can be found, it will be added to the metapaths list.

Secondly, hypergraphs are constructed using the meta-paths to capture the rich relationships between entities in the MOOCs dataset (Line 5 to Line 6). If the path type is Vanilla, the Vanilla GNN is used to generate the projection head (Line 6). Hyperbolic space, as a non-Euclidean geometry, is used to represent complex relationships between entities. The HGCR-HGCN algorithm utilizes both hyperbolic and Euclidean space representations of the hypergraphs to learn more comprehensive representations of users and concepts. If the path type is hyperbolic, the hyperbolic GNN employs Euclidean mapping to generate the projection head (Line 6 to Line 8).

Thirdly, graph contrastive learning is a technique used to train GNNs by learning from positive and negative examples based on similarity (Line 8 to Line 9). The HGCR-HGCN algorithm uses graph contrastive learning to enhance the recommendations by learning from negative examples. Feature aggregation is influenced by causal inference (Line 10 to Line 11). Causal inference is used to identify the causal relationships between confounders (Line 12). Adjustment for confounders is performed using statistical methods (e.g., regression) to estimate the causal effect within the feature aggregation module (Line 13 to Line 14).

© COPYRIGHT UPM

Sensitivity analysis is then conducted to assess robustness, and the causal effect estimate is returned to the prediction layer (Line 15 to Line 16). Finally, this information is used to generate the prediction result (Line 17 to Line 18).

4.3 Method of HGCR-HGCN

This chapter dives into the details of our proposed HGCR-HGCN methods. As shown in Figure 4.6, the framework consists of five components, meta-path selection, graph representation, causal inference, graph contrastive learning, and recommendation prediction. It will discuss each part in more detail in the following paragraphs.

4.3.1 Meta-path Selection

According to the previous research(Shi et al., 2019; Piao, 2021), meta-paths have demonstrated their capability in establishing entity-entity relationships. This study focuses on examining user-user and concept-concept relationships utilizing various meta-paths, similar to those employed in the earlier research. The MOOCCube dataset incorporates six meta-paths, with four designated for users and two for concepts, as detailed in Table 4.1.

Table 4.1 : Different Meta-Paths Used in MOOCCube Dataset

Type	Meta-Path
User	$user \rightarrow concept \rightarrow user$
	$user \rightarrow course \rightarrow user$
	$user \rightarrow video \rightarrow user$
	$user \rightarrow course \rightarrow teacher \rightarrow course \rightarrow user$
Concepts	$concept \rightarrow user \rightarrow concept$
	$concept \rightarrow course \rightarrow concept$

© COPYRIGHT UPM

The novelty of our proposed meta-path selection approach lies in its dynamic and context-aware mechanism that adapts to individual user behaviors and the diverse learning paths in MOOCs. Unlike traditional methods that use static meta-paths or rely solely on predefined relationships between users and courses, this approach incorporates temporal dynamics, such as the sequence in which courses are taken, and contextual influences, such as peer interactions and user engagement patterns. By leveraging these additional factors, the meta-path selection is capable of capturing the evolving learning preferences of users and adapting recommendations accordingly. This dynamic selection is further enhanced by utilizing a hypergraph-based representation, allowing for the modeling of higher-order relationships that connect users, concepts, and content in a multi-faceted way. Such a rich representation ensures that recommendations are not only personalized but also responsive to the changing educational needs of users over time, thereby significantly improving the effectiveness of the recommendation system in MOOCs.

In order to more clearly show the meta-path selection, it is also necessary to define it. Based on a student's viewing history, it utilizes their learned concepts, course information, videos, and other pertinent contextual details to predict and recommend concepts that align with their interests. Given a set of n users, $\mathcal{U} = \{u_1, ..., u_n\}$, and m concepts, C = $\{c_1,, ..., c_m\}$, it construct an implicit feedback matrix R $\in R^{\mathrm{n}\times\mathrm{m}}$, where each entry $r_{u,c}$ = 1 if user u has learned concept c and $r_{u,c}$ = 0 otherwise, as described in (Piao, 2021). To effectively address this problem, it can represent it using a heterogeneous graph denoted as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, comprising an object set $\mathcal{V}$ and a link set $\mathcal{E}$. Additionally, the heterogeneous graph is accompanied by an object type mapping function ϕ: $\mathcal{V} \rightarrow O$ and a link type mapping function ψ: $\mathcal{E} \rightarrow \mathcal{R}$. As specified in [31], $\mathcal{R}$ represents predefined object and link types, where $|O|+|\mathcal{R}|>2$.

The MOOC database in this study is represented as a heterogeneous graph, which effectively captures the complex relationships between various entities within the MOOC ecosystem. In this heterogeneous graph, the six main categories of entities are users, concepts, videos, courses, schools, and teachers. These entities are interconnected through various links that represent their diverse relationships, forming a comprehensive network of knowledge and interactions. A network schema provides a comprehensive description of a network's meta-structure, complementing its heterogeneous graph representation. This schema, denoted as $\mathcal{S} = (\mathcal{R}; \mathcal{O})$, facilitates the extraction of semantically meaningful meta-paths between entities. Formally, a meta-path is a path within the network schema represented as $O_1 \xrightarrow{R_1} O_2 \xrightarrow{R_2} \dots \xrightarrow{R_l} O_{l+1}$. It encapsulates a composite relation R = $R_1 * R_2 * \dots * R_l$ between object O_1 and object O_{l+1}, where the composition operator for relations is shown by $*$.

4.3.2 Hypergraph Construction

The hypergraph construction will involve four main phases as shown in Figure 4.2, which including hypergraph construction, hypergraph representation, hypergraph training algorithms, and recommendation generation.

© COPYRIGHT UPM

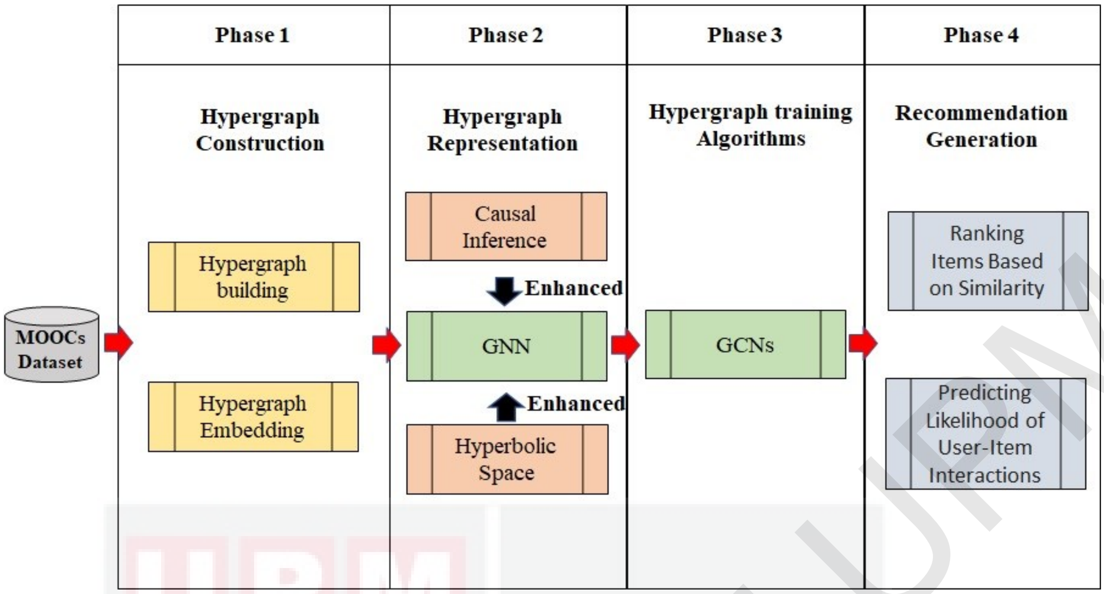

Figure 4.2 : The Phases of Hypergraph Construction

The hypergraph construction process typically involves the following steps: The first step in hypergraph construct is to building the hypergraph itself, based on the hyperedges and node representations(Yu Zhu et al., 2016). Hypergraph construction is a fundamental task in deep learning that involves the representation and analysis of complex relationships between entities in a hypergraph. Hypergraphs are a generalization of graphs that allow for more flexible and expressive representations of relationships between entities. In hypergraphs(Saghafi et al., 2023), an edge or hyperedge can connect more than two nodes, allowing for the representation of higher-order relationships between entities.

Once the hypergraph has been built, the next step is to learn the hypergraph embeddings(Yu Zhu et al., 2016). This can be done using various techniques such as deep learning models, graph convolutional networks, or hypergraph neural networks(C. Liu et al., 2023). The resulting embeddings must accurately capture the relationships between hyperedges and nodes while preserving the higher-order structure of the hypergraph. The graph convolutional network, which are used to process the graph structure and extract meaningful information. Graph convolutional network have been shown to be effective in capturing the complex relations between entities in a graph and are particularly useful in addressing the sparsity and heterogeneity of MOOCs data.

The second phase is to enhance the hypergraph with causal relationships between the entities. The causal relationships are identified using a causal inference algorithm, which takes into account the temporal order of the interactions and the potential confounding factors. Specifically, it uses the do-calculus algorithm(Hyttinen et al., 2015) to identify the causal relationships between the nodes in the hypergraph, based on the temporal order of the interactions and the observed variables. After that add the causal relationships as edges in the hypergraph, which allows for more accurate representation of the causal relationships between different entities. The methodology also involves the use of hyperbolic space for efficient representation of the hypergraph (Y. Li et al., 2021)and the use of causal inference algorithm for identifying causal relationships(Geiger et al., 2021). The hypergraph representation is enhanced using hyperbolic embedding, which allows for efficient representation of the complex relationships between different entities. The data sources used in the study include the MOOC dataset, which contains information on user interactions with courses and concepts, and the hyperbolic embedding algorithm and the causal inference algorithm.

Thirdly, the enhanced hypergraph is then used to train a graph contrastive networks (GCNs) for recommendation(L. Liu et al., 2022). The GCNs is a deep learning model (Mengru Chen et al., 2023)that learns the representations of the entities and the relationships between them, and uses them to make personalized recommendations for the users. The GCNs is trained using a contrastive loss function, which encourages the model to learn the differences between positive and negative examples. Specifically, it randomly samples pairs of nodes from the hypergraph, and train the GCNs to maximize the similarity between positive pairs and minimize the similarity between negative pairs.

At last, after learning the representations of nodes and hyperedges, HGRS employ them to generate recommendations for each user. This process involves various approaches, including ranking items based on their similarity to the user's representation (J. Wang et al., 2021)and predicting the likelihood of user-item interactions(H. Dai et al., 2016). Ranking items based on similarity involves recommending items with higher similarity scores to the user, while predicting user-item interactions involves recommending items based on predicted interaction probabilities.

The network design of a network-based recommendation algorithm is a crucial step that significantly influences the algorithm's performance. While traditional networks are limited to representing pairwise, point-to-point relationships, hypergraphs offer the advantage of capturing higher-order relationships involving three or more nodes. The process of constructing the hypergraphs involves two main steps: hyper-edge definition and hypergraph building. Hyper-edge definition involves identifying the entities that make up the hyper-edges and defining the hyper-edges themselves, which are constructed using various techniques such as clustering, community detection, or spectral partitioning. The resulting hyper-edges are then transformed into a matrix or

© COPYRIGHT UPM

tensor representation for easy processing by deep learning algorithms. Hypergraph building involves constructing the hypergraph itself, based on the hyper-edges and node representations. A hypergraph is referred to be "homogeneous" when all of the nodes connected by hyperedges are of the same kind; otherwise, it is called "heterogeneous." In this research, it employs a heterogeneous hypergraph to represent the relationships between users and items, allowing the algorithm to capture the intricate interactions between these entities.

Just as it defined meta-path in Section A, the users set $\mathcal{U} = \{u_1, ..., u_n\}$, and concepts set C = $\{c_1,, ..., c_m\}$, m, n ∈ ℕ denoting the number of users and concepts, respectively. In this setup, hyperedges connecting concepts and users are represented as the sets of the form $e \subseteq V$, where V denotes the combined set of users U and concepts C. We define a user-concept hypergraph G as the tuple $H_g = (V, E, \boldsymbol{w})$, where $E = \{e_1, e_2, \ldots, e_m\}$ is a collection of hyperedges connecting nodes in V, and $\boldsymbol{w} = \{w_1, w_2, \ldots, w_m\}$ is a corresponding set of weights associated with these hyperedges. The number of nodes in hyperedge e_i is termed the hyperedge's degree, calculated as d $(e_i) = \sum_{j=1}^{M+N} \mathrm{h}(V_j, e_i)$, where h (V_j, e_i) is an element of the incidence matrix H between nodes and hyperedges.

4.3.3 Graph Representation

For the requirements of the model for graph representation, two methods are adopted here and directly implemented using neural network models, where A is the derived adjacency matrix and X is the matrix of node feature vectors. The model incorporates two graph encoders. One, denoted as $\boldsymbol{g_E}$, operates within the standard Euclidean space

© COPYRIGHT UPM

commonly used in deep learning. The other, $\boldsymbol{g_H}$, leverages the more recently explored hyperbolic space for encoding. In the graph encoder, the common graph convolutional network (GCN) (Kipf & Welling, 2017) is adopted, which can learn the node representation of the graph by checking the neighbors of the graph. To learn the user (concept) representation related to the meta path, it applied the linear formula of the hierarchical propagation rule as shown in Equation 4.1:

$$Z_i^{l+1} = W_i^{l+1} \cdot Z_i^l + b_i^{l+1} \qquad \text{(Equation 4.1)}$$

Where Z_i^l denotes the representation of node i in layer l, W_i^{l+1} denotes the weight matrix of node i in layer l +1, and b_i^{l+1} denotes the bias vector of node i in layer l +1. To learn the representations of users and concepts in the hypergraph, it adopts the non-linear formula of the hierarchical propagation rule. The non-linear formula can be implemented by adding a non-linear activation function, such as ReLU(Agarap, 2018) activation function, to capture more complex relationships as shown in Equation 4.2:

$$Z_i^{l+1} = \sigma(W_i^{l+1} \cdot Z_i^l + b_i^{l+1}) \qquad \text{(Equation 4.2)}$$

In the equation, Z_i^l denotes the representation of node i in layer l, W_i^{l+1} denotes the weight matrix of node i in layer l+1, and b_i^{l+1} denotes the bias vector of node i in layer l+1. Subsequently, the representation in Euclidean space of a user u for a meta-path MP_i will be $e_u^{MP_i}$ and for hypergraph HG_i will be $e_u^{HG_i}$. The embeddings for users in Euclidean space and hyperbolic space can be represented as follows:

$$e_u = e_u^{MP_i} \odot e_u^{HG_i} = \boldsymbol{g_E}(X, A) \odot \boldsymbol{g_E} f(X, A) \qquad \text{(Equation 4.3)}$$

$$h_u = e_h^{MP_i} \odot e_h^{HG_i} = exp_x^c\left(\boldsymbol{g_E}(X, A) \odot \boldsymbol{g_E} f(X, A)\right) \qquad \text{(Equation 4.4)}$$

Here, the method of combining representations is to cross-embed them, where $\odot$ is the Hadamard product and f is the representation function. For example, it can use graph convolutional neural networks (GCN) to learn hypergraph representations. In GCN, the representation function f is a graph convolutional layer. Graph convolutional layers can capture the neighbor relations in hypergraphs through the adjacency matrix A. And the embeddings in Euclidean space and hyperbolic space for concept is derived as e_c and h_c , respectively.

4.3.4 Causal Inference

Graph embeddings are essential to the entire system, as they have strong representation power. To improve graph representation embeddings, causal inference is another method used to enhance hypergraph representation. This approach is predicated on the idea of a causal graph, which illustrates the causal connections between different variables. Variables are shown as nodes in a causal graph, and directed edges represent the causal links between the variables. The direction of the edge indicates the causal relationship.

For example, if there is a directed edge from variable A to variable B, then this means that variable A effect variable B. The formula is shown in Equation 4.5.

$$P(Y|do(X)) = \sum_z P(Y|X, z)P(z) \qquad \text{(Equation 4.5)}$$

Where P(Y|do(X)) is the probability of Y given an intervention on X, $P(Y|X, z)$ is the probability of Y given X and z, $P(z)$ is the probability of z. This formula is known as the backdoor adjustment formula. It can be used to estimate the effect of an intervention on X on Y, even if there are confounding variables that affect both X and Y.

4.3.5 Graph Contrastive Learning

Over Euclidean embeddings, hyperbolic representations perform better in taxonomic, hierarchical, and entailment data. This is because disconnected subtrees from the latent hierarchical structure detangle and cluster in the embedding space due to the negative curvature of the embedding space. By contrasting pairs of embeddings from these distinct spaces, it can extract rich semantic information. To leverage this advantage, it proposes a contrastive learning approach that constructs pairs from both Euclidean and hyperbolic spaces. The goal is to use graph contrastive learning of Euclidean and hyperbolic embeddings to extract informative semantics from the opposing viewpoints of the input graph. It is not possible to compare embeddings directly from one space to another without mapping one space to the other, when a different embedding is present. To compare Euclidean and hyperbolic embeddings, it first translates the Euclidean graph embedding into hyperbolic space using an exponential function:

$$e\rightarrow (h_u^{MP_i} \odot h_u^{HG_i}) = exp_x^c\,(h_u^{MP_i} \odot h_u^{HG_i}) \quad \text{(Equation 4.6)}$$

$$e\rightarrow (h_c^{MP_i} \odot h_c^{HG_i}) = exp_x^c\,(h_c^{MP_i} \odot h_c^{HG_i}) \quad \text{(Equation 4.7)}$$

Following that, the representations are run through a shared projection head f, a multilayer perceptron MLP with two hidden layers (Pinkus, 1999) and parametric rectified linear unit PReLU (M. Zhao et al., 2021) nonlinearity, to obtain the final graph representations $h_u^{MP_i} \odot h_u^{HG_i}$ and $e\rightarrow (h_u^{MP_i} \odot h_u^{HG_i})$ for user, the $h_c^{MP_i} \odot h_c^{HG_i}$ and $e\rightarrow (h_c^{MP_i} \odot h_c^{HG_i})$ for concept. To bolster the discriminative and resilient

© COPYRIGHT UPM

qualities of representations within generalized spaces like hyperbolic spaces, we suggest integrating a contrastive loss mechanism with residual components in the angular space, thereby enhancing the discriminative power of the loss function. Angular θ_{ij} will be represented as:

$$\theta_{ij}^{u} = \arccos(\frac{(h_u^{MP_i} \odot h_u^{HG_i})\ \mathrm{e}\!\rightarrow (h_u^{MP_i} \odot h_u^{HG_i})}{\| (h_u^{MP_i} \odot h_u^{HG_i}) \| \cdot \| \mathrm{e}\!\rightarrow (h_u^{MP_i} \odot h_u^{HG_i}) \|} \qquad \text{(Equation 4.8)}$$

Insufficient decision margins can result in erroneous decisions when minor disturbances arise around the decision boundary. To address this issue, it introduces a novel training objective for graph representation learning of positive pairs, incorporating an additive angular residual m between contrastive pairs as the formulated shown in Equation 4.9 and 4.10.

$$L_{contrast}^{u} = -\log \frac{e^{\cos(\theta_{ii*}^{u}+m)/\tau}}{e^{\cos(\theta_{ii*}^{u}+m)/\tau} + \sum_{j \neq i} e^{\cos(\theta_{ij}^{u})/\tau}} \qquad \text{(Equation 4.9)}$$

$$L_{contrast}^{c} = -\log \frac{e^{\cos(\theta_{ii*}^{c}+m)/\tau}}{e^{\cos(\theta_{ii*}^{c}+m)/\tau} + \sum_{j \neq i} e^{\cos(\theta_{ij}^{c})/\tau}} \qquad \text{(Equation 4.10)}$$

This is the formula for the final contrastive loss:

$$L_{contrast} = L_{contrast}^{u} + L_{contrast}^{c} \qquad \text{(Equation 4.11)}$$

4.3.6 Recommendation Prediction

At last, the user features and the concepts features are aggregated in the Euclidean space of each meta-path and hypergraph construction for the final prediction as follows: $e_u = \sum_{j \in ||MP \odot HG||} e_u^{MP_j \odot HG_j}$, and $e_c = \sum_{j \in ||MP \odot HG||} e_c^{MP_j \odot HG_j}$. As the same way, it aggregates the features in hyperbolic space, $h_u = \sum_{j \in ||MP \odot HG||} h_u^{MP_j \odot HG_j}$, and

© COPYRIGHT UPM

$h_c = \sum_{j\in||MP\odot HG||} h_c^{MP_j \odot HG_j}$. Drawing from the user and concept representations and embeddings that have been obtained e_u, e_c, and h_u, h_c. Based on the acquired users and concepts representations or embeddings, the following calculation formula can be used to obtain the preference score of user u for concept c using the matrix factorization framework:

$$\hat{y}_{u,c} = e_u^T \Theta_e e_c + \alpha \cdot h_u^T \Theta_h h_c \qquad \text{(Equation 4.12)}$$

The preference score for user u and concept c, denoted by $\hat{y}_{u,c}$, is determined using two trainable matrices, Θ_e and Θ_h, which project user and concept representations into a shared space. The trainable parameter α is used to balances the prediction scores from these two different views. It uses Bayesian individualized Ranking (BPR) for individualized ranking (Rendle et al., 2009), a technique derived from Bayesian analysis. In a collection of new concepts, BPR recommends that a learned concept be placed higher than a random one. This can be expressed mathematically as follows:

$$L_{rank} = \sum_{u,i,j,i\neq j} -\ln\left(\sigma(w(u,i)\hat{y}_{u,i} - w(u,j)\hat{y}_{u,j})\right) + \gamma\|\Theta\|^2 \qquad \text{(Equation 4.13)}$$

The notation u, i, j represents a triplet relationship involving a user u, an interacted concept i, and an unknown concept j for that user. Term $\hat{y}_{u,i} - \hat{y}_{u,j}$ signifies the preference discrepancy between interacted and unknown concepts. The loss function, modulated by the sigmoid activation function σ, prioritizes interacted concepts in its first term, promoting their higher ranking. The causal weight of user u on item i is w (u, i).and causal weight of user u on item j is w (u, j). The parameter γ acts as the regularization parameter for the *L2* norm, controlling the severity of the regularization penalty. Θ represents the collection of parameters that need to be learned. Our proposed HGCR-HGCN's ultimate loss function can be expressed as follows:

© COPYRIGHT UPM

$$L = L_{contrast} + \beta L_{rank} \quad \text{(Equation 4.14)}$$

where the multitask learning coefficient is represented by β.

4.4 Hyperparameter Optimization Objective

In the realm of deep learning, the challenge of hyperparameter optimization, or tuning, revolves around selecting an optimal set of hyperparameters for a given learning algorithm. Hyperparameters are parameters utilized to govern the learning process, influencing the algorithm's behavior and performance. The objective is to strategically choose values for these hyperparameters to enhance the efficiency and effectiveness of the learning algorithm.

In this study, hyperparameter optimization objective is to obtain the optimal values of the parameters **α** and **β** in Eq.12 and Eq.14, angular residual **m** in Eq.9 and Eq.10 respectively. The reason for hyperparameter optimization is that the parameters **α**, **β**, and **m** have a great impact on the performance of the model. Sections 5.3.1 and 5.4.1 provide specific instructions for adjusting the hyper-parameters.

4.5 Experimental Setup

The specific experimental setup of the proposed method in this study are shown in Table 4.2. First, the proposed method is run on the real MOOC dataset MOOCCube. In the second step, the preprocessing stage is performed, which mainly includes processing of null values and outliers. Then, the heterogeneous graph feature extraction stage mainly adopts aggregation and transformation methods. Among them, two methods are mainly used for graph construction representation, namely, the meta-

© COPYRIGHT UPM

path construction method and the hypergraph construction method. The meta-path construction corresponds to the aggregation method, and the hypergraph construction uses the hyperbolic space feature mapping method, which corresponds to the transformation-based method. Then, GCN is used for contrastive learning, and causal inference are added when clustering features in the last layer of contrastive learning. Finally, the BPR method is used to obtain the predicted ranking results, and the obtained results are measured by HR, NDCG, and MRR.

Table 4.2 : The Experimental Setup of HGCR-HGCN

SETUP	METHOD
Dataset Used	MOOCCube
Pre-processing Step	• NaN filling • Z-score normalization • Variable selection
Heterogeneous Graph Feature Extraction	• aggregation method (meta-path construction) • transformation method (hypergraph construction ← hyperbolic space)
Contrastive Learning	GCN
Prediction	BPR
Performance Measure	• HR • NDCG • MRR

4.6 Summary

In this chapter, the proposed method HGCR-HGCN involves several advanced techniques, including Meta-paths Construction, Causal Inference, and Heterogeneous Graph. The system utilizes hypergraphs to represent relationships between entities and employs Vanilla GNN, Hyperbolic GNN, and Euclidean Mapping to learn embeddings. To predict MOOC recommendations, a shared Projection Head and Feature Aggregation Module are utilized. Additionally, the system includes a Causal Inference layer to improve the recommendation accuracy. Overall, the system aims to

© COPYRIGHT UPM

provide more accurate and effective recommendations for MOOC users, thereby improving their learning experiences. In addition to the introduction of the above methods, this section also proposes several hyperparameters and experimental Setup that need to be optimized, so as to prepare for the experiment and the analysis of experimental results in the next chapter.

© COPYRIGHT UPM

CHAPTER 5

EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Introduction

This section starts by outlining the research questions. It then describes the baseline methods and deeply discussion the experimental results of this proposed method. The datasets and evaluation metrics were already shown detail in chapter 3.2.

5.2 Baseline Methods

To demonstrate the effectiveness of our method, it compares with several state-of-the-art methods including:

- **CMF** (Z. Zhao et al., 2015) tackles the challenge of analyzing multiple behavior types by decomposing their data matrices simultaneously.
- **Metapath2vec** (Dong et al., 2017) utilizes meta-path-based random walks alongside heterogeneous skip-gram models to adeptly capture the diverse neighborhood of a node.
- **NMTR** (C. Gao et al., 2019) employs a joint optimization technique rooted in multitask learning to effectively integrate the strengths of NCF modeling.
- **EHCF** (Chong Chen et al., 2020) stands as a model that adeptly captures fine-grained user-item interactions and efficiently optimizes its parameters without relying on negative sampling.
- **ACKRec** (Gong et al., 2020) tackles the knowledge concept recommendation problem by framing it as a rating prediction task, where a concept's rating is determined by the frequency of interactions between a user and that concept.
- **LightGCN** (X. He et al., 2020) streamlines the Graph Convolutional

© COPYRIGHT UPM

Network (GCN) architecture to enhance its comprehensibility and suitability for recommendation tasks.

- **MOOCIR** (Piao, 2021) utilizes learned user and concept representations to predict a user's preference for selecting specific concepts.
- **ROME** (Luo et al., 2023)is a graph contrastive multi-view framework for MOOCs recommendation, utilizing hyperbolic angular space to effectively capture the intricate relationships between users and concepts.

5.3 Experimental Results

5.3.1 Hyperparameter Optimization

To initialize various parameters, including user and concept nodes, the meta-path list, and the loss function, the Gaussian distribution (mean value = 0, standard deviation = 0.1) is employed without any prior pretraining (Qiu et al., 2019). To effectively optimize the model and prevent overfitting, we implement Adam (Kingma & Ba, 2015) as the optimizer and incorporate the $L2$ norm as the regularization method. The model will be trained for 500 global steps, with a learning rate of 0.01 that doesn't decay over time. Each batch will contain 1024 samples and there will be 30 batches per epoch. The maximum argument of the hyperbolic tangent function is set as 15 and the small epsilon value used for numerical stability is set as 1e-5 in Eq.4.7. Lastly, Eq. 4.13 sets the dimension of latent features for user (concept) embeddings to 10 and 100, respectively, and the regularization parameter λ to 1e-8. The details of all the important parameters for our research are summarized in Table 5.1.

© COPYRIGHT UPM

Table 5.1 : Summarize All the Important Parameters for Our Research

Parameter	**Value**
Gaussian Distribution	Mean = 0, Std. Deviation = 0.1
Optimizer	Adam (Kingma & Ba, 2015)
Regularization Method	L2 Norm
Training Steps	500 Global Steps
Learning Rate	0.01 (No decay)
Batch Size	1024 Samples
Batches per Epoch	30
Max Argument for tanh Function	15
Epsilon for Numerical Stability	1e^-5
Latent Feature Dimensions (Users)	10
Latent Feature Dimensions (Concepts)	100
Regularization Parameter (λ)	1e^-8
Validation Set Construction	Last interacted concept per user
Epochs for Training	500
Model Selection Metric	Highest MRR (Mean Reciprocal Rank)
Baseline - CMF Embedding Size	64
Baseline - Metapath2vec Embedding Size	128
Metapath2vec Walk Length	40
Baseline - NMRT Embedding Size	64
NMRT L2 Regularization Weight	0.001
Baseline - EHCF Embedding Size	128 (Two Layers)
EHCF Dropout Rate	20%
Baseline - ACKRec Embedding Size	128 (Two Layers)
ACKRec Dropout Rate	20%
LightGCN Layers	2
LightGCN Hidden Unit Embedding Size	64
LightGCN Dropout Rate	20%
ROME and MOOCIR Regularization Parameter	1e^-8
ROME and MOOCIR Learning Rate	0.01
Latent Feature Dimensions (Matrix Factorization)	30
Latent Feature Dimensions (Users/Concepts)	100

To prevent overfitting, the model constructs a validation set using each user's last interacted concept. It randomly pairs each known concept with 99 unknown ones and runs for 500 epochs, monitoring performance on the validation set (details in Section 5.3.2). Ultimately, it chooses the model with the highest MRR (Mean Reciprocal Rank) on the validation set. MRR measures how high the ground truth concept ranks among 100 candidates. Other metrics can also be used based on preference.

© COPYRIGHT UPM

To optimize performance, the research meticulously tuned hyper-parameters identified as crucial in existing research for each baseline method, leaving others at their default settings. In CMF, it uses linear regression with stochastic gradient descent to learn the parameters and set the embedding size as 64. For Metapath2vec, both user and item embedding dimensions were set to 128 and the walk length as 40. In NMRT, the embedding size between user and item set of 64 and L2 regularization weight choose 0.001. For EHCF and ACKRec, it first both used two layers and 128 as embedding size, the second one is dropout rate set as 20%. LightGCN's best performance occurred with two layers and a hidden unit embedding size of 64, for dropout rate sets of 20%. Finally, ROME and MOOCIR first set the regularization parameter as 1e-8 and the learning rate is set as 0.01, second the dimensionality of latent features for matrix factorization (MF) was set to 30, while user (concept) embeddings were 100-dimensional.

5.3.2 Comparison with Benchmark Work

To evaluate the performance of the proposed method, the MOOC dataset is split into training and test sets. The evaluation metrics used include HR@k, NDCG@k, and MRR. This research will compare with several state-of-the-art baselines on a real-world dataset (MOOCCube). The comparison results are reported in Table 5.2 and 5.3, from which we have the following observations:

© COPYRIGHT UPM

Table 5.2 : The Recommendation Accuracies of Method Measured by HR@K. Bold Denotes Indicate the Best Model Performance during All of the Methods, and Underline Indicates the Best Model Performance in This Type

Methods	HR@5	HR@10	HR@20	HR@50
CMF	0.620	0.741	0.828	0.859
NMTR	<u>0.651</u>	<u>0.772</u>	0.829	0.891
EHCF	0.649	0.768	<u>0.833</u>	<u>0.898</u>
ACKRec	0.659	0.764	0.842	0.881
LightGCN	0.688	0.759	0.851	0.894
MOOCIR	<u>0.704</u>	<u>0.836</u>	<u>0.872</u>	<u>0.919</u>
Metapaht2vec	0.642	0.773	0.873	0.906
ROME	<u>0.710</u>	<u>0.855</u>	<u>0.907</u>	<u>0.938</u>
HGCR-HGCN (ours)	**0.728**	**0.895**	**0.928**	**0.952**

Table 5.3 : The Recommendation Accuracies of Method Measured by NDGC@K and MRR. Bold Denotes Indicate the Best Model Performance during All of the Methods, and Underline Indicates the Best Model Performance in This Type

Methods	NDCG@5	NDCG @10	NDCG @20	NDCG @50	MRR
CMF	0.432	0.479	0.499	0.509	0.418
NMTR	0.470	0.516	0.542	0.591	0.449
EHCF	<u>0.475</u>	<u>0.521</u>	<u>0.547</u>	<u>0.593</u>	<u>0.462</u>
ACKRec	0.503	0.538	0.557	0.601	0.475
LightGCN	0.515	0.548	0.569	0.617	0.477
MOOCIR	<u>0.520</u>	<u>0.562</u>	<u>0.594</u>	<u>0.637</u>	<u>0.484</u>
Metapaht2vec	0.468	0.511	0.537	0.582	0.440
ROME	<u>0.539</u>	<u>0.587</u>	<u>0.612</u>	<u>0.655</u>	<u>0.497</u>
HGCR-HGCN (ours)	**0.569**	**0.625**	**0.651**	**0.667**	**0.518**

HGCR-HGCN, our proposed model, consistently demonstrates superior performance compared to all the baseline methods. For instance, it is noteworthy that HGCR-HGCN outperforms all baseline models by a substantial margin of 2.54% (0.710→0.728), 4.68% (0.855→0.895), 2.32% (0.907→0.928) and 1.49% (0.938→0.952) of HR metric for Top-k recommendation. As for the NDCG metric, AGNN achieves the obviously highest score when k = 5, 10, 15, 20 with the improvements of 5.57%, 6.47%, 6.37%, and 1.83%, respectively. This demonstrates the validity of our hyperbolic GCN-based solution and the benefits of introducing hypergraph representation and causal inference.

© COPYRIGHT UPM

A. Traditional similarity-based recommendation methods like CMF, NMTR, and EHCF, which these methods rely on various techniques to measure item or user similarities. It is evident that Method EHCF exhibits the best performance among three similarity-based recommendation methods and it also used as a comparison model for analysis in this research. In comparison to these methods, our approach demonstrates superior performance across key metrics, including HR@k, NDCG@k, and MRR. The hybrid nature of our method, incorporating graph neural network and graph contrastive learning, contributes to higher HR@k compare with the EHCF which improve as 10.62% (0.649→0.728), 20.18% (0.768→0.895), 17.65% (0.833→0.928), 6.01% (0.898→0.952) respectively in Table 5.2. This shows the indicating a better ability to place relevant items within the top-k recommendations. NDCG@k values showcase the enhanced ranking quality of our recommendations compared to similarity-based models. It achieves a better score compare with EHCF model when k=5, 10, 20, 50 with the improvements of 19.79%, 19.96%, 19.01%, and 12.48%, respectively in Table 5.3. Moreover, our approach demonstrates a notable improvement of 12.12% (0.462→0.518) in the MRR metric when compared to the EHCF model, which represents the best-performing among the three similarity-based models. signifying the effectiveness in ranking the first relevant item in Table 5.3. These results collectively suggest that our method excels in providing accurate, diverse, and well-ranked recommendations, outperforming traditional similarity-based recommending methods across multiple evaluation metrics.

B. Compared to similarity-based recommending methods (like CMF, NMTR and EHCF), graph-based recommending methods (such as ACKRec, LightGCN and MOOCIR) generally outperform similarity-based methods. Compare these two graph-based approaches, the result in Table 5.2 and Table 5.3 shows the model of MOOCIR get better performance during these methods. As a result, it infers that efficient exploitation of graph semantic information can result in discriminative abilities. Second, almost all of the evaluation metrics show that our proposed method performs significantly better than other baselines. Take the Hit Ratio for example, it shows

© COPYRIGHT UPM

comparatively high scores at 0.895 and 0.928 when k = 10, 20, which is 7.06%, and 6.42% over the second-best baseline. In regard to NDCG, it achieves the obviously highest score when k=5, 10, 20, 50 with the improvements of 9.42%, 11.21%, 9.60%, and 4.71%, respectively. As for MRR, our method still achieves the obviously highest score 0.518, which is 7.02% over the best baseline MOOCIR.

C. Meta-path enhanced graph-based methods (i.e., Metapath2vec, ROME, and HGCR-HGCN) have a significant improvement over traditional graph-based method (ie., ACKRec, LightGCN and MOOCIR) as shown in Table 5.2 and Table 5.3. This is because meta-paths allow for the definition of meaningful relationships, enabling the model to capture diverse and higher-order connections within the graph. This flexibility, absent in traditional methods, leads to more comprehensive representations that better reflect the intricacies of the graph structure. By considering task-specific meta-paths, these methods adapt to specific applications, providing personalized and accurate solutions. The ability to mitigate data sparsity, capture nuanced interactions, and enhance expressiveness makes meta-path enhanced graph-based methods more effective in modeling complex relationships within graphs, leading to superior performance across various tasks.

D. As the value of K increases, HGCR-HGCN consistently maintains its superior performance across the benchmarks. The results presented in Figure 5.1 demonstrate the effectiveness of HGCR-HGCN in enhancing the performance of the ranking recommendation task, specifically Top-K recommendation. Compare our method with three different types baseline in Figure 5.2, the K=10 shows the best preference of HR@K and NDCG@K.

© COPYRIGHT UPM

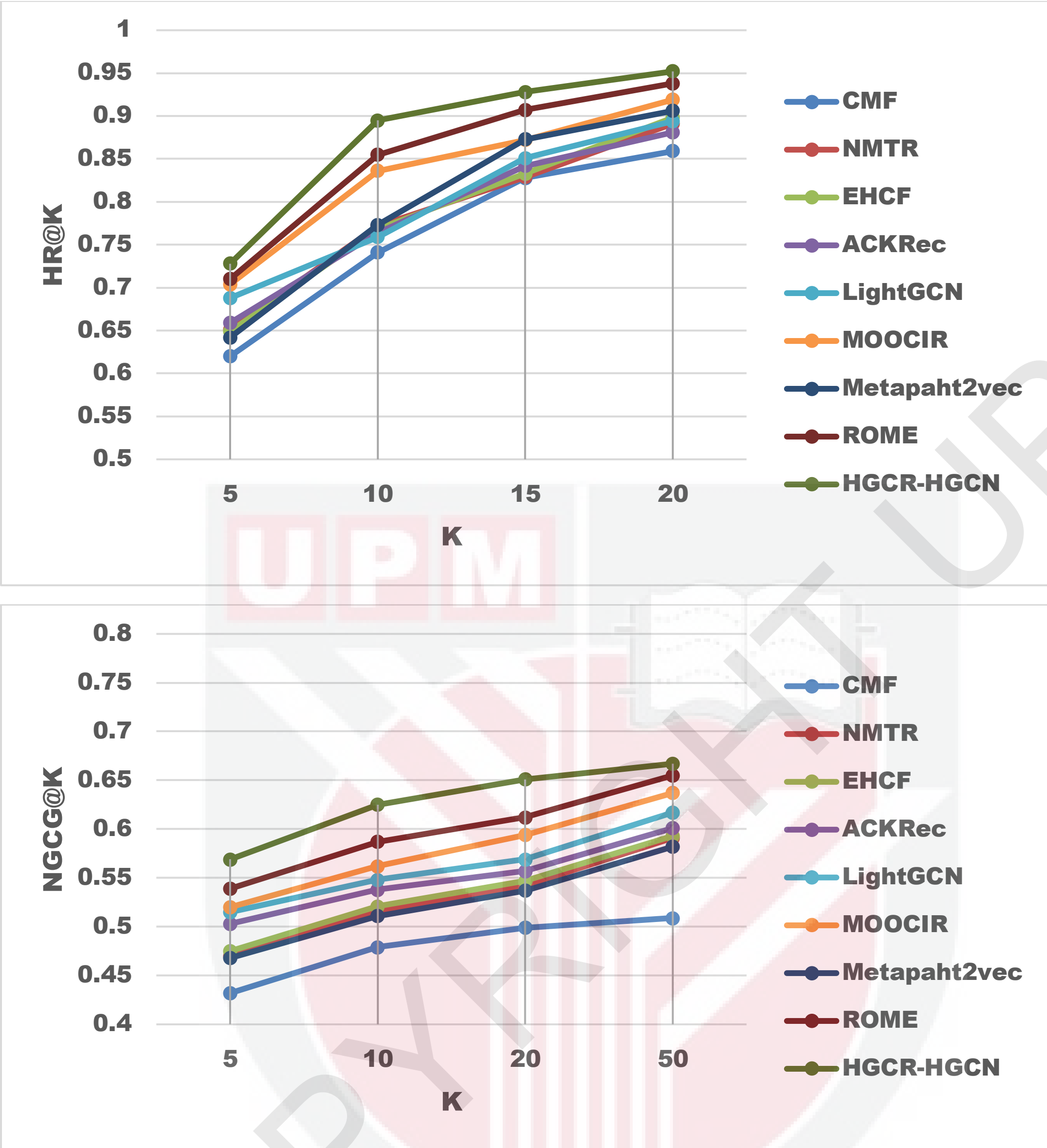

Figure 5.1 : Results of HR@K and NDCG@K in Top-K Recommendation

© COPYRIGHT UPM

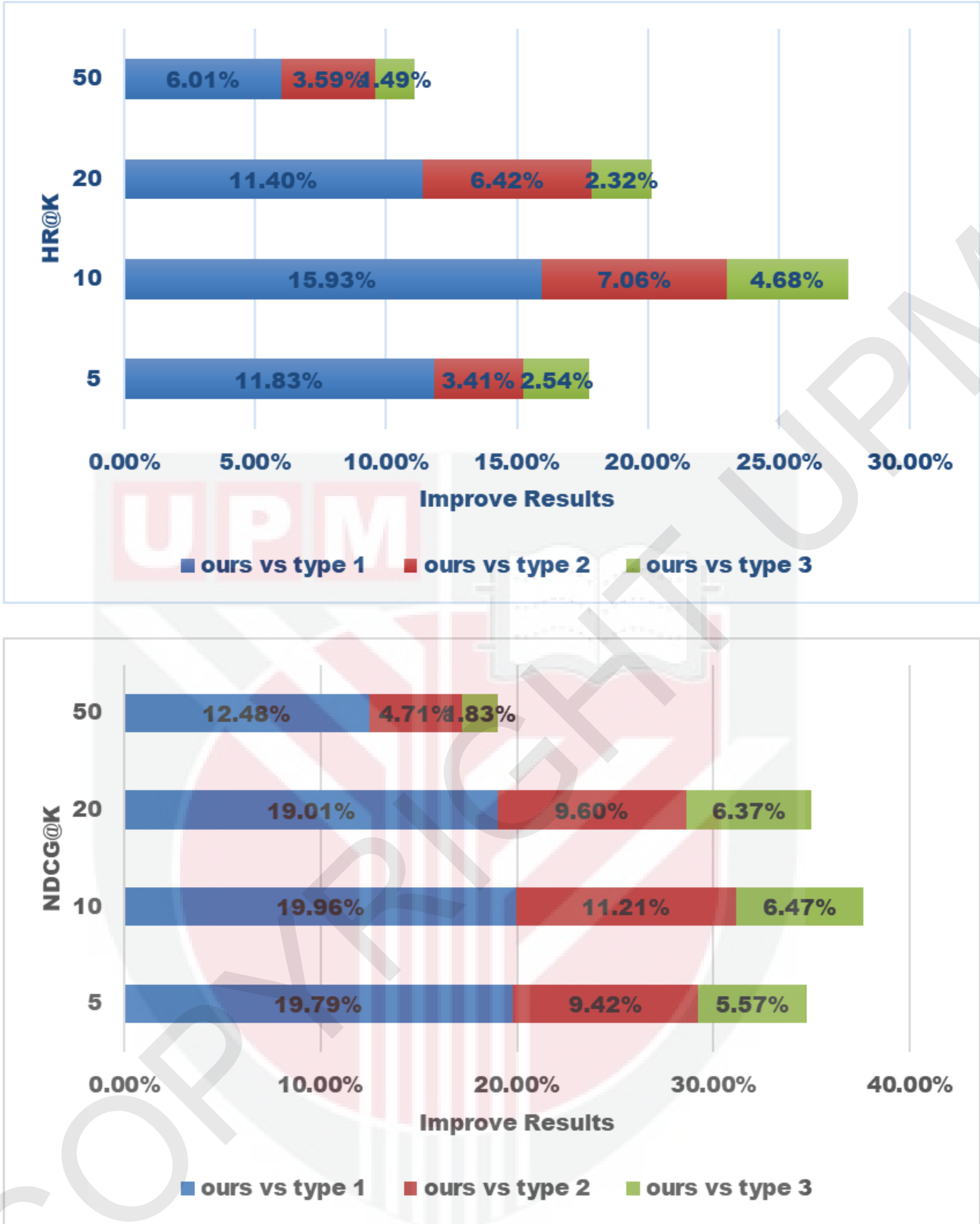

Figure 5.2 : Improve Results between Ours and Different Types

These results demonstrate our method is effective because of these three ways: (i) It leverages hypergraph embeddings that are able to capture richer structural information about nodes beyond the pairwise relationships in a traditional graph. By modeling higher-order connections between nodes, the hypergraph embeddings may better

© COPYRIGHT UPM

represent the complex relationships in MOOC data, including both direct and indirect influences. (ii) It performs causal recommendation in a hyperbolic space rather than Euclidean space. Embedding the graph in a hyperbolic geometric setting allows it to naturally model hierarchical properties that are inherent in MOOC datasets. (iii) It employs contrastive learning on the hyperbolic graph embeddings to induce clusters of similar nodes. This could result in embeddings that better discern semantic similarities and produce higher quality recommendations.

5.4 Experimental Analysis

In order to assess our HGCR-HGCN method's performance even further, this part undertakes a number of in-depth experiments, such as the ablation study and the impact of hyper-parameters.

5.4.1 Ablation Study

This approach involves using hypergraphs to represent the users and courses in the MOOCs, and using hyperbolic space to model the relationships and interactions between them. Hypergraph based causal inference will be applied to identify the most relevant factors influencing learner behavior and performance in MOOCs. The HGCN model uses contrastive learning to learn representations that capture the similarities and differences between users and courses, and uses a hyperbolic space-based distance metric to measure the similarity between the representations.

© COPYRIGHT UPM

Table 5.4 : Components Included in the Variants of HGCR-HGCN, Include with Checkmarks (√) or Blanks Indicating Not Include. Ignore This Part Checkmarks (⊗)

Methods	HG[a]	CI[b]	HB[c]	HGCI[d]	HGCN[e]
Ours\HG (no HG)		√	√	⊗	√
Ours\CI (no CI)	√		√	⊗	√
Ours\HB (no HB)	√	√		√	⊗
Ours\HGCI (no HGCI)	⊗	⊗	√		√
Ours\HGCN (no HGCN)	√	√	⊗	√	
HGCR-HGCN (Ours)	√	√	√	√	√

[a]The Hypergraph constructure (HG) component.
[b]The Causal Inference (CI) component.
[c]The Hyperbolic space (HB) component.
[d]The Hypergraph enhanced Causal Inference (HGCI) component.
[e]The Hyperbolic space Graph Contrastive Network (HGCN) component.

To understand the individual contributions of HGCR-HGCN's components, an ablation study was conducted. Each component was systematically removed or replaced, and its impact on MOOC recommendation performance was observed. Table 5.4 outlines the components included in each HGCR-HGCN variant, which are defined as follows:

- **Ours\HG (no HG):** This is to demonstrate the important of the hypergraph representation component, which is a variant of HGCR-HGCN via getting use and concept interaction from it. That is, it only uses meta-path constructure to learn user-user and concept-concept' representations.

- **Ours\CI (no CI):** This is to show the necessity of considering the Causal Inference component, that is, learning the hide relation removing causal relationships between users and concepts from HGCR-HGCN.

- **Ours\HB (no HB):** This is to evaluate the importance of the hyperbolic space capture the complex, multi-modal nature of MOOC data. That is a variant of HGCR-HGCN via substitutes the hyperbolic representations in the multi-view framework with vanilla Euclidean representations.

- **Ours\HGCN (no HGCN):** This method removes the multi-view framework by excluding hyperbolic representations and contrastive learning in HGCR-

© COPYRIGHT UPM

HGCN. The objective of this variant is to validate the effectiveness of hyperbolic graph contrastive networks in extracting meaningful semantic information from input graphs by employing graph contrastive learning of both Euclidean and hyperbolic embeddings.

- **Ours\HGCI (no HGCI):** This is to evaluate how important of the hypergraph construction and causal inference component in the model of HGCR-HGCN.

Table 5.5 : The Impact of the Components to the Recommendation Performance

Methods	HR @5	HR @10	HR @20	HR @50	NDCG @5	NDC G@10	NDC G@20	NDC G@50	MRR
Ours\HG	0.706	0.869	0.903	0.932	0.553	0.612	0.643	0.657	0.507
Ours\CI	0.710	0.868	0.901	0.939	0.556	0.613	0.639	0.655	0.501
Ours\HB	0.713	0.866	0.896	0.934	0.552	0.611	0.641	0.658	0.505
Ours\HGCN	0.701	0.863	0.887	0.924	0.548	0.601	0.632	0.647	0.491
Ours\HGCI	0.699	0.857	0.891	0.928	0.541	0.604	0.634	0.645	0.494
Ours	**0.728**	**0.895**	**0.928**	**0.952**	**0.569**	**0.625**	**0.651**	**0.667**	**0.518**

To verify the impact of each part of the model on the overall model, the following analysis can be performed based on the results shown in the Table 5.4:

i. HGCR-HGCN consistently outperforms its other variants, as illustrated in Table 5.5. Compared to the Ours\HG (no HG) variant, our model achieves improvements of 3.12%, 2.99%, 2.77%, and 2.15% for k values of 5, 10, 20, and 50, respectively. This demonstrates the crucial role of the HG component, as relying solely on user-user or concept-concept embeddings derived from the meta-path structure is insufficient for achieving optimal results.

ii. As evident from Table 5.5, Ours\CI (no CI) consistently underperforms compared to HGCR-HGCN, exhibiting declines of 2.54%, 3.11%, 2.99%, and 1.38% in HR@K, respectively. This performance gap underscores the importance of the CI component, as relying solely on learners' or concepts' similarity relationships without considering their latent causal connections can only yield suboptimal results.

© COPYRIGHT UPM

iii. HGCR-HGCN surpasses Ours\HB (no HB), demonstrating clear performance gains as evident from Table 5.5. On average, it achieves a 2.74% improvement in HR@K. These results highlight the significance of exploring the hyperbolic geometric perspective for multi-view representation learning, emphasizing that limiting representation learning to Euclidean space may lead to suboptimal outcomes. The multi-view representation learning enhances semantics and addresses recommendation challenges, as removing hyperbolic representations significantly reduces performance. This highlights importance of incorporating such representations.

iv. HGCR-HGCN outperforms Ours\HGCN (no HGCN), that is, when it replaces the hyperbolic graph contrastive network with the GCN method, the results drop by 3.80% on average for HR@K. It demonstrates the effectiveness of the HGCN model uses contrastive learning to learn representations that capture the similarities and differences between users and concept, and uses a hyperbolic space-based distance metric to measure the similarity between the representations.

v. To understand the impact of the two components, hypergraph representation and causal inference, on the HGCR-HGCN model, the following detailed analysis will be conducted below.

(1). From the Table 5.5, Ours\HGCI (no HGCI): achieves worse results than HGCR-HGCN on MOOCCube dataset, decreasing by 4.15%, 4.43%, 4.14%, and 2.59% on HR@K, respectively. HGCR-HGCN outperforms Ours\HGCI (no HGCI), demonstrating the necessity of the hypergraph representation and causal inference components.

(2). To further understand the impact of the two components on the model, we conducted the following experiments: First, we removed hypergraph representation alone, which is defined as Ours\HG (no HG) variant. The experimental results showed that the HR@K performance decreased by an average of 2.76%. Second, we removed causal inference alone, which is defined as Ours\CI (no CI) variant. The experimental results showed that the HR@K performance decreased by an average of 2.51%. Finally,

© COPYRIGHT UPM

we conducted an experiment by removing HG and CI simultaneously, which is defined as Ours\HGCI (no HGCI) variant. The experimental results showed that the HR@K performance decreased most significantly by 3.80%. As shown in Figure 5.3, the comparative experiments again proved that HG and CI have a significant impact on the model. When our HGCR-HGCN model includes these two components, it significantly improves the model performance.

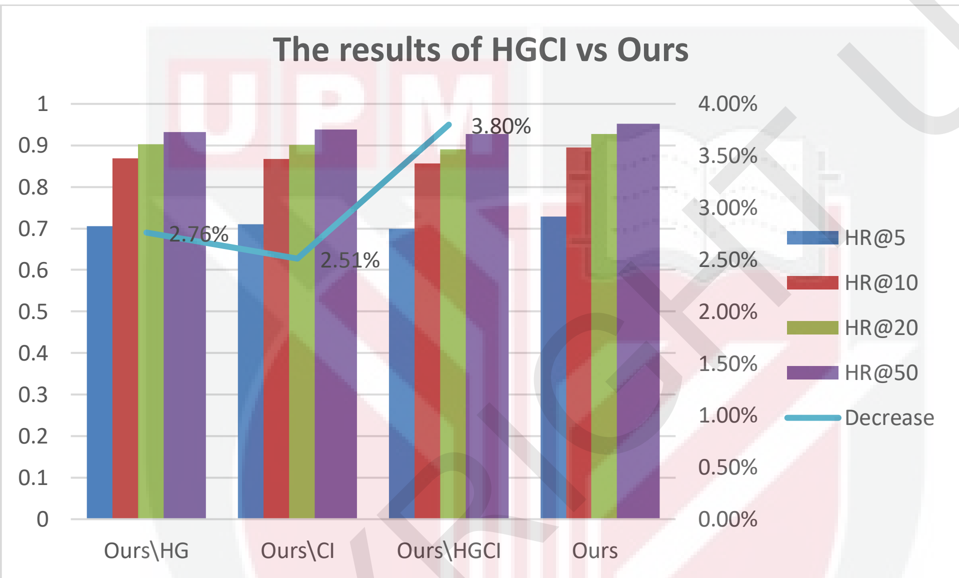

Figure 5.3 : The Results of HGCI vs Ours

(3). To further evaluate the impact of the two components on the model, we also introduced the calculation of weights. At the same time, we summarized the methods and functions of the two components, as shown in Table 5.6. The weight calculation here used sensitivity analysis, which can be used to calculate the sensitivity of the part to the model output. The greater the sensitivity, the greater the impact of the part on the model output. Using sensitivity analysis to calculate the weight of a part in the entire model, the following formula can be used: weight = sensitivity / total sensitivity. According to the data in Table 5.6, the weights of hypergraph representation and causal inference can be calculated to be 0.32 and 0.27, respectively.

© COPYRIGHT UPM

Table 5.6 : Hypergraph Construction and Causal Inference

Hypergraph construction in **MOOCs recommendation**	
Method One	Hypergraph representation
Describe	Constructing a hypergraph to model relationships among various elements in online MOOCs, such as users, courses, and resources.
Strength	Captures complex relationships and dependencies among different components in online MOOCs, providing a comprehensive and flexible representation for analyzing and understanding the structure and dynamics of the MOOC ecosystem.
Metric	0.32
Causal Inference in MOOCS recommendation	
Method One	Propensity Score Matching
Describe	Estimating causal effects by matching treated and control groups based on their propensity to receive a treatment.
Strength	Enables the identification of causal relationships between variables in online MOOCs, providing insights into the impact of interventions or treatments on user outcomes.
Metric	0.27

In summary, the results of our method and its variants provide valuable insights into the effectiveness of different components. The multi-view framework with hyperbolic representations and contrastive learning proves to be crucial in enhancing the recommendation system's performance. Furthermore, the additional mapping in the angular space and the utilization of hyperbolic representations contribute to further improvements. These findings emphasize the significance of incorporating multi-view graph representation learning, exploring geometric views, and leveraging contrastive learning in recommendation systems.

5.4.2 Hyperparameters Sensitivity Analysis

In this section, it investigates the effects of three hyperparameters on our model's performance. These hyperparameters include two balance coefficients, denoted as α and β, in Equations 12 and 14, and the margin of the angular space, represented as m, in Equations 9 and 10. To evaluate their impact, it conducts experiments using the HR metric. It began by varying the value of the hyperparameter α from 0.1 to 0.9 while

© COPYRIGHT UPM

holding other parameters fixed. α controls the influence of representations derived from the hyperbolic space on the overall model. As shown in Figure 5.4, it observes that higher values of α lead to improved recommendation performance, indicating that incorporating representations from hyperbolic space enhances the expressiveness of our HGCR-HGCN model. Notably, the model's results remained relatively stable when α was within the range of 0.6 to 0.7.

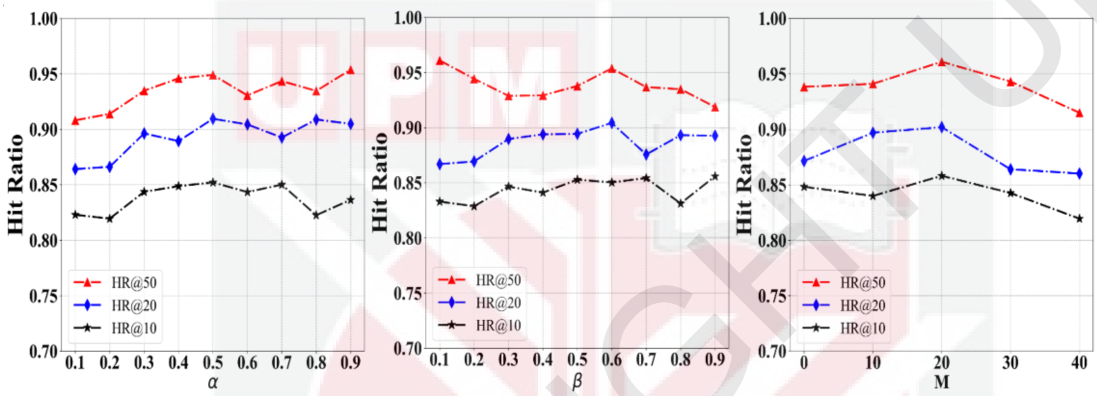

Figure 5.4 : Hyperparameters Sensitivity Analysis of α, β, and M

Next, it fixed all other hyperparameters and explored the impact of varying β from 0.1 to 0.9. β determines the balance between the two optimization objectives. Based on our sensitivity analysis, it found that the model's performance was not significantly impacted when β was close to 0.5.

Regarding the margin parameter M of the angular space, it incrementally varied it from 0 to 40 degrees in steps of 10 degrees. M directly affects the discriminative capability of the model. Our experiments revealed that performance remained stable as M ranged from 10 to 20 degrees.

© COPYRIGHT UPM

Informed by these sensitivity analyses, it set the hyperparameters α, β, and M to values of 0.8, 0.3, and 20 degrees respectively, as these settings yielded favorable results. α was set higher to emphasize representations from hyperbolic space, while β was set lower to slightly prioritize proximity over contrastive learning. M was allocated an intermediate value within the stable range noted. Together, these configurations facilitated effective recommendation performance as demonstrated through our empirical evaluations.

5.5 Discussion

In addition to the aforementioned points, the proposed hypergraph-enhanced causal recommendation framework with graph contrastive networks (GCN) for MOOCs also addresses the challenge of sparsity in recommendation systems. In MOOC platforms, the number of available courses and resources can be vast, and users' interactions with these items are often sparse. Traditional recommendation methods struggle to provide accurate and diverse recommendations in such sparse scenarios. However, by leveraging the hypergraph structure, the proposed framework can effectively capture high-order relations among items and users, even in sparse data situations. This enables the framework to make more informed and accurate recommendations.

Furthermore, the framework incorporates graph contrastive networks (GCNs) to learn representations of items and users. GCNs have shown great success in capturing complex relationships in graph data. By applying GCNs to the hypergraph structure, the framework can learn meaningful representations that capture the dependencies and interactions among items and users. The graph contrastive loss function used in training encourages the representations of positive samples (i.e., items that users have

© COPYRIGHT UPM

interacted with) to be closer to each other in the embedding space compared to negative samples (i.e., items that users have not interacted with). This learning mechanism helps improve the discriminative power of the representations and enhances the recommendation accuracy.

The evaluation of the proposed framework involves extensive experiments on a real-world MOOC dataset. This dataset likely contains information about users, courses, resources, and their interactions. The experiments compare the performance of the proposed framework against several state-of-the-art baselines, which may include traditional collaborative filtering methods, matrix factorization techniques, or graph-based recommendation models. The evaluation metrics used could include accuracy measures, such as precision, recall, or mean average precision, as well as diversity measures to assess the variety of recommended items.

The experimental results demonstrate the superiority of the proposed framework over the baseline methods in terms of recommendation accuracy and diversity. This indicates that the hypergraph structure, along with the GCN-based model and the graph contrastive training, effectively address the limitations of traditional recommendation methods in MOOC platforms. The framework's ability to leverage high-order relations, capture complex interactions, and learn meaningful representations contributes to its improved performance.

By analyzing the performance of different components, such as the hypergraph structure and the GCN model, the study provides insights into their individual contributions to the recommendation accuracy and diversity. This analysis helps

© COPYRIGHT UPM

validate the effectiveness of each component and provides a better understanding of their role in enhancing the recommendation performance. It also helps guide future improvements and optimizations of the framework.

Overall, the proposed hypergraph-enhanced causal recommendation framework with graph contrastive networks presents a novel approach to addressing the limitations of traditional recommendation methods in MOOC platforms. It leverages the hypergraph structure to capture high-order relations, employs graph contrastive networks for representation learning, and demonstrates promising results in terms of recommendation accuracy and diversity. The study contributes to the field of graph-based learning by showcasing the application of hypergraph structures and graph contrastive networks in the context of MOOC recommendation systems. Future research directions could involve exploring further enhancements to the framework, such as incorporating additional contextual information or evaluating its performance on larger and more diverse datasets.

5.6 Summary

The paper tackles the challenge of recommending online MOOCs (Massive Open Online Courses) by introducing a novel approach that leverages graph representation learning from two distinct spaces. This approach not only utilizes the standard Euclidean space representations learned by vanilla Graph Neural Networks (GNNs), but also incorporates the rich semantic information captured by GNNs operating in the hyperbolic space. To effectively combine these perspectives, we map both representations into a unified hyperbolic space and employ a contrastive learning framework with an angular margin to achieve more discriminative representation

© COPYRIGHT UPM

learning. Extensive experimentation on the widely used MOOCCube dataset demonstrates the superiority of our proposed method compared to several state-of-the-art recommendation algorithms. Looking towards the future, we envision extending our method to a broader range of relevant applications, including semi-supervised recommendation systems and other real-world scenarios with complex relational data.

© COPYRIGHT UPM

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Introduction

The thesis's main contributions are wrapped up in this chapter. It also describes the possible ways to expand or enhance the work that is offered in the thesis. In light of this, this thesis is a significant attempt to suggest a similarity check method for online MOOC recommendations. This research contributes to the field by developing the HGCR-HGCN framework, which integrates hypergraph construction, causal inference, and hyperbolic space representation to improve personalized recommendations in MOOCs. The use of contrastive learning on hyperbolic embeddings further enhances the quality of user and concept representations, leading to more accurate and context-aware recommendations. Additionally, the proposed method has been comprehensively evaluated against state-of-the-art baselines, demonstrating its effectiveness in advancing personalized learning in online education.

6.2 Conclusion

Massive Open Online Courses (MOOCs) have revolutionized the learning landscape, offering individuals convenient access to a wide range of educational resources and courses from prestigious institutions. However, the sheer abundance of available courses and resources presents a significant challenge for users in discovering the most relevant ones that align with their interests and goals. To tackle this challenge, MOOC platforms have developed recommendation systems that leverage users' preferences and behavior to provide personalized and pertinent course recommendations. These

© COPYRIGHT UPM

recommendation systems aim to enhance the learning experience and outcomes, ultimately assisting users in achieving their educational objectives.

Traditional recommendation methods employed in MOOC platforms, such as collaborative filtering and content-based filtering, possess certain limitations. These include issues like the cold start problem, where it is challenging to provide recommendations for new users or items with limited data, sparsity in the user-item interaction matrix, which hinders accurate recommendations, and scalability problems when dealing with large datasets. To overcome these limitations, recent research has delved into the utilization of graph-based models, specifically graph neural networks (GNNs), to bolster the performance of recommendation systems. These models leverage the inherent relationships among users, courses, and resources in a MOOC platform to offer more precise and effective recommendations.

In this paper, it proposes a novel hypergraph-enhanced causal recommendation framework, complemented by graph contrastive networks (GCN), tailored specifically for MOOCs. Our framework capitalizes on the hypergraph structure, enabling the capture of higher-order relationships among users and items. It employs a GCN-based model to learn representations of items and users, which are crucial for making accurate recommendations. Furthermore, the GCN model is trained using a graph contrastive loss, which encourages the representations of positive samples (users who liked a particular course) to be closer to each other than the representations of negative samples (users who did not show interest in the course). The proposed HGCR-GCNs framework has the potential to significantly advance the state-of-the-art in personalized recommendation systems for MOOCs, by leveraging the unique

© COPYRIGHT UPM

advantages of hypergraph modeling and graph contrastive learning. This research can contribute to the development of more effective and efficient recommendation systems for MOOCs, and ultimately promote better learning outcomes for MOOC users.

The task of this research is to develop and evaluate a new recommendation framework that leverages hypergraph modeling to capture the complex relationships between users, courses, and other contextual factors, and graph contrastive learning to enhance the quality of recommendations by learning from negative examples.

More specifically, the objectives of this research including developing a hypergraph-based model for representing MOOC user-concept interactions, which captures the complex, multi-modal nature of MOOC data. To investigate the use of hypergraph-based causal inference to identify the most relevant factors influencing user behavior and performance in MOOCs. To develop a graph contrastive learning framework that can effectively learn from negative examples and enhance recommendation quality. And to evaluate the proposed HGCR-GCNs framework against state-of-the-art recommendation algorithms on large-scale MOOC datasets, and demonstrate its effectiveness in improving recommendation accuracy, diversity, and novelty.

To evaluate the effectiveness of our proposed framework, it conducted comprehensive experiments on a real-world MOOC dataset. The results clearly demonstrate that our framework outperforms several state-of-the-art baselines in terms of recommendation accuracy and diversity. Additionally, it conducted ablation studies to further validate the effectiveness of each component within our proposed framework.

© COPYRIGHT UPM

Looking ahead, there are several promising avenues for future research in the field of MOOCs recommendation systems. One potential direction is exploring the integration of hypergraph-based models, which have shown promise in capturing higher-order relationships, with graph-based models to further enhance the recommendation performance. Additionally, incorporating additional data sources, such as user feedback and social network information, could contribute to improving personalization and relevance in recommendations. Evaluating recommendation systems using diverse metrics, including novelty, serendipity, and fairness, would provide a more holistic understanding of recommendation quality. Finally, ensuring scalability and efficiency in recommendation systems to handle larger datasets and cater to a growing user base is critical for real-time personalized recommendations.

In summary, our proposed framework has demonstrated its potential to enhance the recommendation performance in MOOC platforms. However, further research is needed to explore the integration of different graph-based models, incorporate additional data sources, evaluate using diverse metrics, and improve scalability and efficiency. Addressing these areas of research will contribute to the ongoing development and improvement of recommendation systems in MOOCs, ultimately benefiting users by enhancing their learning experience and helping them achieve their educational goals.

6.3 Future Works

Although our proposed framework showed promising results, there are several avenues for future research in the field of MOOCs recommendation systems. Firstly, further investigation can be conducted to explore the integration of other graph-based

© COPYRIGHT UPM

models, such as hypergraph-based models, to enhance the recommendation performance. Hypergraph based models have shown potential in capturing higher-order relationships among users, courses, and resources, and their combination with graph-based models could lead to improved recommendations. Possible paths include developing hybrid graph-hypergraph architectures that can dynamically switch between different types of relationships based on the context of the recommendation or experimenting with hypergraph attention mechanisms to better capture the relative importance of various connections.

Secondly, the incorporation of additional data sources, such as user feedback and social network information, can be explored to further enhance the personalization and relevance of the recommendations. User feedback, such as ratings and reviews, can provide valuable insights into users' preferences and help refine the recommendation algorithms. Social network information, on the other hand, can be leveraged to identify influential users or communities that can contribute to more accurate recommendations. Possible paths could involve developing multi-modal recommendation frameworks that combine textual reviews, ratings, and social interactions into a unified model or employing graph-based clustering techniques to identify community structures that can inform the recommendation process.

Furthermore, the evaluation of recommendation systems can be extended to include other metrics and dimensions, such as novelty, serendipity, and fairness. These metrics can provide a more comprehensive understanding of the quality and effectiveness of the recommendations and ensure that they are diverse, unexpected, and unbiased. Possible paths include conducting user studies to assess user satisfaction with novelty

© COPYRIGHT UPM

and serendipity in recommendations, or implementing fairness-aware algorithms that mitigate biases based on user demographics or other sensitive attributes.

Lastly, the scalability and efficiency of the recommendation systems can be further improved to handle larger datasets and serve a larger user base. As MOOC platforms continue to grow in popularity and attract more users, it is crucial to develop scalable algorithms and infrastructure to deliver personalized recommendations in real-time. Possible paths include exploring distributed computing frameworks such as Apache Spark to parallelize the recommendation process, or employing approximate nearest neighbor search techniques to enable real-time performance on massive datasets.

In conclusion, the proposed framework has demonstrated the potential to enhance the recommendation performance in MOOC platforms. However, there are still opportunities for future research to explore the integration of different graph-based models, incorporate additional data sources, evaluate using diverse metrics, and improve scalability and efficiency. Addressing these areas will contribute to the ongoing development and improvement of recommendation systems in MOOCs, ultimately benefiting the learning experience and outcomes of users.

© COPYRIGHT UPM

REFERENCES

Agarap, A. F. (2018). Deep Learning using Rectified Linear Units (ReLU). *ArXiv Preprint ArXiv:1803.08375.*, *1*, 2–8.

Ahmadian Yazdi, H., Seyyed Mahdavi Chabok, S. J., & Kheirabadi, M. (2022). Dynamic Educational Recommender System Based on Improved Recurrent Neural Networks Using Attention Technique. *Applied Artificial Intelligence*, *36*(1). https://doi.org/10.1080/08839514.2021.2005298

Al Arif, T. Z. Z., Kurniawan, D., Handayani, R., Hidayati, & Armiwati. (2024). EFL University Students' Acceptance and Readiness for e-Learning: A Structural Equation Modeling Approach. *Electronic Journal of E-Learning*, *22*(1), 1–16. https://doi.org/10.34190/ejel.22.1.3063

Andersen, C. M., & Bro, R. (2010). Variable selection in regression-a tutorial. *Journal of Chemometrics*, *24*(11–12), 728–737. https://doi.org/10.1002/cem.1360

Antelmi, A., Cordasco, G., Polato, M., Scarano, V., Spagnuolo, C., & Yang, D. (2023). A Survey on Hypergraph Representation Learning. *ACM Computing Surveys*, *56*(1), 1–38. https://doi.org/10.1145/3605776

Arya, D., Gupta, D. K., Rudinac, S., & Worring, M. (2024). Adaptive Neural Message Passing for Inductive Learning on Hypergraphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–18. https://doi.org/10.1109/TPAMI.2024.3434483

Assami, S., Daoudi, N., & Ajhoun, R. (2018). Personalization criteria for enhancing learner engagement in MOOC platforms. *IEEE Global Engineering Education Conference, EDUCON*, *2018-April*, 1265–1272. https://doi.org/10.1109/EDUCON.2018.8363375

Badriyah, T., Azvy, S., Yuwono, W., & Syarif, I. (2018). Recommendation system for property search using content based filtering method. *2018 International Conference on Information and Communications Technology, ICOIACT 2018*, *2018-Janua*, 25–29. https://doi.org/10.1109/ICOIACT.2018.8350801

Berg, R. van den, Kipf, T. N., & Welling, M. (2017). *Graph Convolutional Matrix Completion*.

Bozyi\ugit, A., Bozyi\ugit, F., Kilinç, D., & Nasibo\uglu, E. (2018). Collaborative filtering based course recommender using OWA operators. *2018 International Symposium on Computers in Education (SIIE)*, 1–5.

Cai, D., Qian, S., Fang, Q., Hu, J., Ding, W., & Xu, C. (2023). Heterogeneous Graph Contrastive Learning Network for Personalized Micro-Video Recommendation. *IEEE Transactions on Multimedia*, *25*, 2761–2773. https://doi.org/10.1109/TMM.2022.3151026

© COPYRIGHT UPM

Cai, Xingjuan, Hu, Z., Zhao, P., Zhang, W. S., & Chen, J. (2020). A hybrid recommendation system with many-objective evolutionary algorithm. *Expert Systems with Applications*, *159*, 113648. https://doi.org/10.1016/j.eswa.2020.113648

Cai, Xuheng, Huang, C., Xia, L., & Ren, X. (2023). LIGHTGCL: SIMPLE YET EFFECTIVE GRAPH CONTRASTIVE LEARNING FOR RECOMMENDATION. *11th International Conference on Learning Representations, ICLR 2023*, 1–16.

Çano, E., & Morisio, M. (2017). Hybrid recommender systems: A systematic literature review. In *Intelligent Data Analysis* (Vol. 21, Issue 6, pp. 1487–1524). https://doi.org/10.3233/IDA-163209

Chamberlain, B. P., Clough, J., & Deisenroth, M. P. (2017). *Neural Embeddings of Graphs in Hyperbolic Space*.

Chami, I., Ying, R., Ré, C., & Leskovec, J. (2019). Hyperbolic graph convolutional neural networks. *Advances in Neural Information Processing Systems*, *32*(NeurIPS).

Chanaa, A., & El Faddouli, N.-E. (2019). Context-aware factorization machine for recommendation in massive open online courses (MOOCs). *2019 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS)*, 1–6.

Chen, Chaofan, Cheng, Z., Li, Z., & Wang, M. (2020). Hypergraph attention networks. *Proceedings - 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2020*, *December 2020*, 1560–1565. https://doi.org/10.1109/TrustCom50675.2020.00215

Chen, Chong, Zhang, M., Zhang, Y., Ma, W., Liu, Y., & Ma, S. (2020). Efficient Heterogeneous Collaborative Filtering without negative sampling for recommendation. *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*. https://doi.org/10.1609/aaai.v34i01.5329

Chen, Mengru, Huang, C., Xia, L., Wei, W., Xu, Y., & Luo, R. (2023). Heterogeneous Graph Contrastive Learning for Recommendation. *WSDM 2023 - Proceedings of the 16th ACM International Conference on Web Search and Data Mining*, 544–552. https://doi.org/10.1145/3539597.3570484

Chen, Ming, & Zhou, X. (2020). DeepRank: Learning to rank with neural networks for recommendation. *Knowledge-Based Systems*, *209*, 106478. https://doi.org/10.1016/j.knosys.2020.106478

Chen, Y. H., Huang, N. F., Tzeng, J. W., Lee, C. A., Huang, Y. X., & Huang, H. H. (2022). A Personalized Learning Path Recommender System with LINE Bot in MOOCs Based on LSTM. *2022 11th International Conference on Educational and Information Technology, ICEIT 2022*, 40–45. https://doi.org/10.1109/ICEIT54416.2022.9690754

© COPYRIGHT UPM

Dai, H., Wang, Y., Trivedi, R., & Song, L. (2016). *Deep Coevolutionary Network: Embedding User and Item Features for Recommendation*. https://doi.org/10.475/123

Dai, J., Wu, Y., Gao, Z., & Jia, Y. (2021). A Hyperbolic-to-Hyperbolic Graph Convolutional Network. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 154–163. https://doi.org/10.1109/CVPR46437.2021.00022

Dong, Y., Chawla, N. V., & Swami, A. (2017). Metapath2vec: Scalable representation learning for heterogeneous networks. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *Part F1296*, 135–144. https://doi.org/10.1145/3097983.3098036

Fan, W., Wang, J., Ma, Y., Tang, J., Yin, D., & Li, Q. (2019). Deep social collaborative filtering. *RecSys 2019 - 13th ACM Conference on Recommender Systems*, 305–313. https://doi.org/10.1145/3298689.3347011

Farnell, D. J. J., & Mirra, R. M. (2023). Using Directed Acyclic Graphs (DAGs) to Represent the Data Generating Mechanisms of Disease and Healthcare Pathways: A Guide for Educators, Students, Practitioners and Researchers. In *Teaching Biostatistics in Medicine and Allied Health Sciences*. https://doi.org/10.1007/978-3-031-26010-0

Fei, N., Gao, Y., Lu, Z., & Xiang, T. (2021). Z-Score Normalization, Hubness, and Few-Shot Learning. *Proceedings of the IEEE International Conference on Computer Vision*, 142–151. https://doi.org/10.1109/ICCV48922.2021.00021

Feng, Y., You, H., Zhang, Z., Ji, R., & Gao, Y. (2019). Hypergraph Neural Networks. *The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19) Hypergraph*, 19.

Fu, D., Liu, Q., Zhang, S., & Wang, J. (2016). The Undergraduate-Oriented Framework of MOOCs Recommender System. *Proceedings - 2015 International Symposium on Educational Technology, ISET 2015*, 115–119. https://doi.org/10.1109/ISET.2015.31

Gao, C., He, X., Gan, D., Chen, X., Feng, F., Li, Y., Chua, T., & Jin, D. (2019). Neural multi-task recommenddation from multi-behaviour Data. *In 2019 IEEE 35th International Conference on Data Engineering (ICDE)*, 1–4.

Gao, C., Wang, X., He, X., & Li, Y. (2022). Graph neural networks for recommender system. *WSDM 2022 - Proceedings of the 15th ACM International Conference on Web Search and Data Mining*, 1623–1625. https://doi.org/10.1145/3488560.3501396

Gao, C., Zheng, Y., Wang, W., Feng, F., He, X., & Li, Y. (2022). Causal Inference in Recommender Systems: A Survey and Future Directions. *ArXiv Preprint ArXiv:2208.12397.*, *1*(1).

© COPYRIGHT UPM

Gao, Y., Zhang, Z., Lin, H., Zhao, X., Du, S., & Zou, C. (2022). Hypergraph Learning: Methods and Practices. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(5), 2548–2566. https://doi.org/10.1109/TPAMI.2020.3039374

Gardner, J., Yang, Y., Baker, R. S., & Brooks, C. (2019). Modeling and experimental design for MOOC dropout prediction: A replication perspective. *EDM 2019 - Proceedings of the 12th International Conference on Educational Data Mining*, *Edm*, 49–58.

Geetha, G., Safa, M., Fancy, C., & Saranya, D. (2018). A Hybrid Approach using Collaborative filtering and Content based Filtering for Recommender System. *Journal of Physics: Conference Series*, *1000*(1). https://doi.org/10.1088/1742-6596/1000/1/012101

Geiger, A., Lu, H., Icard, T., & Potts, C. (2021). Causal Abstractions of Neural Networks. *Advances in Neural Information Processing Systems*, *12*(2), 9574–9586.

Gong, J., Wang, C., Zhao, Z., & Zhang, X. (2021). Automatic Generation of Meta-Path Graph for Concept Recommendation in MOOCs. *Electronics*, *10*(14), 1671.

Gong, J., Wang, S., Wang, J., Feng, W., Peng, H., Tang, J., & Yu, P. S. (2020). Attentional Graph Convolutional Networks for Knowledge Concept Recommendation in MOOCs in a Heterogeneous View. *SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 79–88. https://doi.org/10.1145/3397271.3401057

Gulzar, Z., Leema, A. A., & Deepak, G. (2018). PCRS: Personalized Course Recommender System Based on Hybrid Approach. *Procedia Computer Science*, *125*, 518–524. https://doi.org/10.1016/j.procs.2017.12.067

Guo, H., Tang, J., Zeng, W., Zhao, X., & Liu, L. (2021). Multi-modal entity alignment in hyperbolic space. *Neurocomputing*, *461*, 598–607. https://doi.org/10.1016/j.neucom.2021.03.132

Gupta, R., Bhardwaj, K. K., & Sharma, D. K. (2020). Transfer Learning. *Machine Learning and Big Data: Concepts, Algorithms, Tools and Applications*, 337–360. https://doi.org/10.1002/9781119654834.ch13

Harianja, E., & Lumbantoruan, G. (2019). Integrating MLP with Algorithm with AHP Modification for Car Evaluation. *Journal of Physics: Conference Series*, *1361*(1). https://doi.org/10.1088/1742-6596/1361/1/012022

He, X., Deng, K., Wang, X., Li, Y., Zhang, Y. D., & Wang, M. (2020). LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation. *SIGIR 2020 - Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 639–648. https://doi.org/10.1145/3397271.3401063

© COPYRIGHT UPM

Herath, D., & Jayaratne, L. (2018). A personalized web content recommendation system for E-learners in E-learning environment. *2017 National Information Technology Conference, NITC 2017*, *2017-Septe*, 89–95. https://doi.org/10.1109/NITC.2017.8285650

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hou, Z., Liu, X., Cen, Y., Dong, Y., Yang, H., Wang, C., & Tang, J. (2022). GraphMAE: Self-Supervised Masked Graph Autoencoders. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 594–604. https://doi.org/10.1145/3534678.3539321

Hu, Q., & Rangwala, H. (2019). Academic performance estimation with attention-based graph convolutional networks. *EDM 2019 - Proceedings of the 12th International Conference on Educational Data Mining*, 69–78.

Huang, R., & Lu, R. (2019). Research on Content-based MOOC Recommender Model. *2018 5th International Conference on Systems and Informatics, ICSAI 2018*, *Icsai*, 676–681. https://doi.org/10.1109/ICSAI.2018.8599503

Huang, R., & Lu, R. (2018). Research on Content-based MOOC Recommender Model. *2018 5th International Conference on Systems and Informatics (ICSAI)*, 676–681.

Hyttinen, A., Eberhardt, F., & Järvisalo, M. (2015). Do-calculus when the true graph is unknown. *Uncertainty in Artificial Intelligence - Proceedings of the 31st Conference, UAI 2015*, 395–404.

Idrissi, N., Hourrane, O., Zellou, A., & Benlahmar, E. H. (2019). A Restricted Boltzmann Machine-based Recommender System for Alleviating Sparsity Issues. *ICSSD 2019 - International Conference on Smart Systems and Data Science*, 1–5. https://doi.org/10.1109/ICSSD47982.2019.9003149

Jiang, L., Liu, K., Wang, Y., Wang, D., Wang, P., Fu, Y., & Yin, M. (2023). Reinforced Explainable Knowledge Concept Recommendation in MOOCs. *ACM Transactions on Intelligent Systems and Technology*, *14*(3). https://doi.org/10.1145/3579991

Jovanović, N., Meng, Z., Faber, L., & Wattenhofer, R. (2021). Towards Robust Graph Contrastive Learning. *Proceedings of ACM Conference (Conference'17)*, *1*(1). http://arxiv.org/abs/2102.13085

Khalid, A., Lundqvist, K., Yates, A., & Ghzanfar, M. A. (2021). Novel online Recommendation algorithm for Massive Open Online Courses (NoR-MOOCs). *Plos One*, *16*(1), e0245485.

Kim, S., Lee, S. Y., Gao, Y., Antelmi, A., Polato, M., & Shin, K. (2024). A Survey on Hypergraph Neural Networks: An In-Depth and Step-By-Step Guide. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 6534–6544. https://doi.org/10.1145/3637528.

© COPYRIGHT UPM

3671457

Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–13.

Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, 1–14.

Kumar, S., & Jadon, P. (2024). Machine Learning Techniques and Advancement in Recommendation System. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4934448

Lawson, S., Donovan, D., & Lefevre, J. (2024). An application of node and edge nonlinear hypergraph centrality to a protein complex hypernetwork. *Plos One*, 1–20. https://doi.org/10.1371/journal.pone.0311433

Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (2023). *Gradient-based learning applied to document recognition To cite this version : HAL Id : hal-03926082 Gradient-Based Learning Applied to Document Recognition*. *86*(11), 2278–2324.

Li, M., Meng, L., Ye, Z., Yang, Y., Cao, S., Xiao, Y., & Zhao, H. (2024). TP-GCL: graph contrastive learning from the tensor perspective. *Frontiers in Neurorobotics*, *18*(May), 1–9. https://doi.org/10.3389/fnbot.2024.1381084

Li, X., Li, X., Tang, J., Wang, T., Zhang, Y., & Chen, H. (2020). Improving Deep Item-Based Collaborative Filtering with Bayesian Personalized Ranking for MOOC Course Recommendation. *International Conference on Knowledge Science, Engineering and Management*, 247–258.

Li, Y., Chen, H., Sun, X., Sun, Z., Li, L., Cui, L., Yu, P. S., & Xu, G. (2021). Hyperbolic Hypergraphs for Sequential Recommendation. *International Conference on Information and Knowledge Management, Proceedings*, *1*(1), 988–997. https://doi.org/10.1145/3459637.3482351

Li, Z., Liu, H., Zhang, Z., Liu, T., & Shu, J. (2021). Recalibration convolutional networks for learning interaction knowledge graph embedding. *Neurocomputing*, *427*, 118–130. https://doi.org/10.1016/j.neucom.2020.07.137

Liang, F., Qian, C., Yu, W., Griffith, D., & Golmie, N. (2022). Survey of Graph Neural Networks and Applications. *Wireless Communications and Mobile Computing*, *2022*. https://doi.org/10.1155/2022/9261537

Liao, W., Bak-Jensen, B., Pillai, J. R., Wang, Y., & Wang, Y. (2022). A Review of Graph Neural Networks and Their Applications in Power Systems. *Journal of Modern Power Systems and Clean Energy*, *10*(2), 345–360. https://doi.org/10.35833/MPCE.2021.000058

© COPYRIGHT UPM

Liu, C., He, T., Zhu, H., Li, Y., Xie, S., & Hosam, O. (2023). A Survey of Recommender Systems Based on Hypergraph Neural Networks. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 13828 LNCS*. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-28124-2_10

Liu, L., Kang, Z., Ruan, J., & He, X. (2022). Multilayer graph contrastive clustering network. *Information Sciences*, *613*, 256–267. https://doi.org/10.1016/j.ins.2022.09.042

Liu, T., Wu, Q., Chang, L., & Gu, T. (2022). A review of deep learning-based recommender system in e-learning environments. *Artificial Intelligence Review*, *55*(8), 5953–5980. https://doi.org/10.1007/s10462-022-10135-2

Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., & Tang, J. (2023). Self-Supervised Learning: Generative or Contrastive. *IEEE Transactions on Knowledge and Data Engineering*, *35*(1), 857–876. https://doi.org/10.1109/TKDE.2021.3090866

Liu, Y., Jin, M., Pan, S., Zhou, C., Zheng, Y., Xia, F., & Yu, P. S. (2023). Graph Self-Supervised Learning: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, *35*(6), 5879–5900. https://doi.org/10.1109/TKDE.2022.3172903

Luo, H., Husin, N. A., & Aris, T. N. M. (2023). ROME: A Graph Contrastive Multi-View Framework From Hyperbolic Angular Space for MOOCs Recommendation. *IEEE Access*, *11*. https://doi.org/10.1109/ACCESS.2022.3232552

Ma, H., Wang, X., Hou, J., & Lu, Y. (2017). Course recommendation based on semantic similarity analysis. *2017 3rd IEEE International Conference on Control Science and Systems Engineering (ICCSSE)*, 638–641.

Ma, Y., & Tresp, V. (2021). Causal Inference under Networked Interference and Intervention Policy Enhancement. *Proceedings of Machine Learning Research*, *130*, 3700–3708.

Mahdizadehaghdam, S., Panahi, A., & Krim, H. (2019). Sparse generative adversarial network. *Proceedings - 2019 International Conference on Computer Vision Workshop, ICCVW 2019*, 3063–3071. https://doi.org/10.1109/ICCVW.2019.00369

Mohammadpour, T., Bidgoli, A. M., Enayatifar, R., & Javadi, H. H. S. (2019). Efficient clustering in collaborative filtering recommender system: Hybrid method based on genetic algorithm and gravitational emulation local search algorithm. *Genomics*, *111*(6), 1902–1912. https://doi.org/10.1016/j.ygeno.2019.01.001

Monti, F., Bronstein, M. M., & Bresson, X. (2017). Geometric matrix completion with recurrent multi-graph neural networks. *Advances in Neural Information Processing Systems*, *2017-Decem*(Nips), 3698–3708.

© COPYRIGHT UPM

Mu, F., & Li, W. (2024). A Causal Approach for Counterfactual Reasoning in Narratives. *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, *1*, 6556–6569. https://aclanthology.org/2024.acl-long.354

Neamah, A. A., & El-Ameer, A. S. (2018). Design and Evaluation of a Course Recommender System Using Content-Based Approach. *2018 International Conference on Advanced Science and Engineering (ICOASE)*, 1–6.

Norton, V., Miller, M., Pierremeyerunistrahotmailcom, P. M., Ward, J., Agnew, G., Norton, V., Miller, M., Meyer, P., Ward, J., & Agnew, G. (2022). *Heterogeneous Graph Recommendation Model based on Graph Neural Network Heterogeneous Graph Recommendation Model based on Graph Neural Network*. 0–7.

Obeidat, R., Duwairi, R., & Al-Aiad, A. (2019). A Collaborative Recommendation System for Online Courses Recommendations. *Proceedings - 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications, Deep-ML 2019*, 49–54. https://doi.org/10.1109/Deep-ML.2019.00018

Onah, D., & Sinclair, J. (2015). Collaborative filtering recommendation system : a framework in massive open online courses. *INTED2015 Proceedings*, *March*, 1249–1257. https://doi.org/10.13140/RG.2.1.5023.4409

Patel, K., & Patel, H. B. (2020). A state-of-the-art survey on recommendation system and prospective extensions. *Computers and Electronics in Agriculture*, *178*(September), 105779. https://doi.org/10.1016/j.compag.2020.105779

Patra, B. G., Maroufy, V., Soltanalizadeh, B., Deng, N., Zheng, W. J., Roberts, K., & Wu, H. (2020). A content-based literature recommendation system for datasets to improve data reusability – A case study on Gene Expression Omnibus (GEO) datasets. *Journal of Biomedical Informatics*, *104*(May 2019), 103399. https://doi.org/10.1016/j.jbi.2020.103399

Peng, W., Varanka, T., Mostafa, A., Shi, H., & Zhao, G. (2022). Hyperbolic Deep Neural Networks: A Survey. In *IEEE Transactions on Pattern Analysis and Machine Intelligence* (Vol. 44, Issue 12, pp. 10023–10044). IEEE. https://doi.org/10.1109/TPAMI.2021.3136921

Phalle, Tejashri Sharad; Bhushan, S. (2024). Content Based Filtering And Collaborative Filtering: A Comparative Study. *Journal of Advanced Zoology*, *45*, 96.

Piao, G. (2021). Recommending Knowledge Concepts on MOOC Platforms with Meta-path-based Representation Learning. *International Educational Data Mining Society.*

Pinkus, A. (1999). Approximation theory of the MLP model in neural networks. *Acta Numerica*, *8*, 143–195. https://doi.org/10.1017/S0962492900002919

Prenkaj, B., Velardi, P., Distante, D., & Faralli, S. (2020). A Reproducibility Study of Deep and Surface Machine Learning Methods for Human-related Trajectory Prediction. *International Conference on Information and Knowledge Management, Proceedings*, 2169–2172. https://doi.org/10.1145/3340531.3412088

Qi, L., Maximilian, N., & Douwe, K. (2019). Hyperbolic Graph Neural Networks. *Advances in Neural Information Processing Systems*, *32*(NeurIPS).

Qiu, R., Huang, Z., Li, J., & Yin, H. (2019). Rethinking the item order in session-based recommendation with graph neural networks. *International Conference on Information and Knowledge Management, Proceedings*, 579–588. https://doi.org/10.1145/3357384.3358010

Rabahallah, K., Mahdaoui, L., & Azouaou, F. (2018). MOOCs recommender system using ontology and memory-based collaborative filtering. *ICEIS 2018 - Proceedings of the 20th International Conference on Enterprise Information Systems*, *1*(Iceis 2018), 635–641. https://doi.org/10.5220/0006786006350641

Raghavendra, S. (2024). Scalability of Data Science Algorithms : Empowering Big Data Analytics. *Artificial Intelligence and Soft Computing Techniques*, *2*, 101.

Rahman, M., Shama, I. A., Rahman, S., & Nabil, R. (2022). Hybrid Recommendation System To Solve Cold Start Problem. *Journal of Theoretical and Applied Information Technology*, *100*(11), 3562–3580.

Rane, N. L., Mallick, S. K., Kaya, Ö., & Rane, J. (2024). *Machine learning and deep learning architectures and trends : A review* (Issue 2024).

Rawaa Alatrash, Rojalina Priyadarshini, H. E. (2024). Collaborative filtering integrated fine-grained sentiment for hybrid recommender system. *The Journal of Supercomputing*, *80*(4), 4760–4807.

Rendle, S., Freudenthaler, C., Gantner, Z., & Schmidt-Thieme, L. (2009). BPR: Bayesian personalized ranking from implicit feedback. *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, UAI 2009*.

Riyahi, M., & Sohrabi, M. K. (2020). Providing effective recommendations in discussion groups using a new hybrid recommender system based on implicit ratings and semantic similarity. *Electronic Commerce Research and Applications*, *40*(December 2018), 100938. https://doi.org/10.1016/j.elerap.2020.100938

Saghafi, F., Mirzaei, S., & Fahim, A. (2023). Product Recommender System Using Hypergraph Neural Networks. *2023 9th International Conference on Web Research, ICWR 2023*, 189–195. https://doi.org/10.1109/ICWR57742.2023.10139136

Salehi, M., & Kmalabadi, I. N. (2012). A Hybrid Attribute–based Recommender System for E–learning Material Recommendation. *IERI Procedia*, *2*, 565–570. https://doi.org/10.1016/j.ieri.2012.06.135

© COPYRIGHT UPM

Sha, X., Sun, Z., & Zhang, J. (2021). Electronic Commerce Research and Applications Hierarchical attentive knowledge graph embedding for personalized recommendation. *Electronic Commerce Research and Applications*, *48*(June), 101071. https://doi.org/10.1016/j.elerap.2021.101071

Shah, L., Gaudani, H., & Balani, P. (2016). Survey on Recommendation System. *International Journal of Computer Applications*, *137*(7), 43–49. https://doi.org/10.5120/ijca2016908821

Shi, C., Hu, B., Zhao, W. X., & Yu, P. S. (2019). Heterogeneous information network embedding for recommendation. *IEEE Transactions on Knowledge and Data Engineering*, *31*(2), 357–370. https://doi.org/10.1109/TKDE.2018.2833443

Son, J., & Kim, S. B. (2017). Content-based filtering for recommendation systems using multiattribute networks. *Expert Systems with Applications*, *89*, 404–412. https://doi.org/10.1016/j.eswa.2017.08.008

Stepanek, H. (2020). Thinking in pandas: How to use the python data analysis library the right way. In *Thinking in Pandas: How to Use the Python Data Analysis Library the Right Way*. https://doi.org/10.1007/978-1-4842-5839-2

Sun, F. Y., Hoffmann, J., Verma, V., & Tang, J. (2020). INFOGRAPH: UNSUPERVISED AND SEMI-SUPERVISED GRAPH-LEVEL REPRESENTATION LEARNING VIA MUTUAL INFORMATION MAXIMIZATION. *8th International Conference on Learning Representations, ICLR 2020*, *2018*, 1–16.

Sun, X., Cheng, H., Liu, B., Li, J., Chen, H., Xu, G., & Yin, H. (2023). Self-Supervised Hypergraph Representation Learning for Sociological Analysis. *IEEE Transactions on Knowledge and Data Engineering*, *35*(11), 11860–11871. https://doi.org/10.1109/TKDE.2023.3235312

Suresh, K., & Srinivasan, P. (2020). Massive Open Online Courses – Anyone Can Access Anywhere at Anytime. *Shanlax International Journal of Education*, *8*(3), 96–101. https://doi.org/10.34293/education.v8i3.2458

Tarus, J. K., Niu, Z., & Yousif, A. (2017). A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. *Future Generation Computer Systems*, *72*, 37–48. https://doi.org/10.1016/j.future.2017.02.049

Tian, H., & Zafarani, R. (2024). Higher-Order Networks Representation and Learning: A Survey. *ACM SIGKDD Explorations Newsletter*, *26*(1), 1–18. https://doi.org/10.1145/3682112.3682114

Tian, X., & Liu, F. (2021). Capacity Tracing-Enhanced Course Recommendation in MOOCs. *IEEE Transactions on Learning Technologies*, *14*(3), 313–321. https://doi.org/10.1109/TLT.2021.3083180

Uddin, I., Imran, A. S., Muhammad, K., Fayyaz, N., & Sajjad, M. (2021). A Systematic Mapping Review on MOOC Recommender Systems. *IEEE Access*,

© COPYRIGHT UPM

9, 118379–118405. https://doi.org/10.1109/ACCESS.2021.3101039

Veličković, P., Fedus, W., Hamilton, W. L., Bengio, Y., Liò, P., & Devon Hjelm, R. (2019). Deep graph infomax. *7th International Conference on Learning Representations, ICLR 2019*, 1–17.

Veras De Sena Rosa, R. E., Guimarães, F. A. S., Mendonça, R. D. S., & Lucena, V. F. De. (2020). Improving Prediction Accuracy in Neighborhood-Based Collaborative Filtering by Using Local Similarity. *IEEE Access*, *8*, 142795–142809. https://doi.org/10.1109/ACCESS.2020.3013733

Vuong, M. (2023). *Background on Graph Neural Networks*. *October*, 3–5.

Wang, J., Xie, H., Au, O. T. S., Zou, D., & Wang, F. L. (2020). Attention-based CNN for personalized course recommendations for MOOC learners. *2020 International Symposium on Educational Technology (ISET)*, 180–184.

Wang, J., Xie, H., Wang, F. L., Lee, L. K., & Au, O. T. S. (2021). Top-N personalized recommendation with graph neural networks in MOOCs. *Computers and Education: Artificial Intelligence*, *2*(September 2020), 100010. https://doi.org/10.1016/j.caeai.2021.100010

Wang, Q., Nguyen, Q. V. H., Yin, H., Huang, Z., Wang, H., & Cui, L. (2019). Enhancing collaborative filtering with generative augmentation. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 548–556. https://doi.org/10.1145/3292500.3330873

Wang, W. (2022). Attentional Meta-path Contrastive Graph Convolutional Networks for Knowledge Concept Recommendation. *2022 Tenth International Conference on Advanced Cloud and Big Data (CBD)*, 200–205. https://doi.org/10.1109/CBD58033.2022.00043

Wang, Xiang, He, X., Wang, M., Feng, F., & Chua, T. S. (2019). Neural graph collaborative filtering. *SIGIR 2019 - Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 165–174. https://doi.org/10.1145/3331184.3331267

Wang, Xiao, Ji, H., Cui, P., Yu, P., Shi, C., Wang, B., & Ye, Y. (2019). Heterogeneous graph attention network. *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, *May*, 2022–2032. https://doi.org/10.1145/3308558.3313562

Wang, Xinhua, Jia, L., Guo, L., & Liu, F. (2022). *Multi-aspect heterogeneous information network for MOOC knowledge concept recommendation*.

Wei, Y., Wang, X., Nie, L., Li, S., Wang, D., & Chua, T. S. (2022). Causal Inference for Knowledge Graph based Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 1–11. https://doi.org/10.1109/TKDE.2022.3231352

© COPYRIGHT UPM

Xia, J., Wu, L., Chen, J., Hu, B., & Li, S. Z. (2022). SimGRACE: A Simple Framework for Graph Contrastive Learning without Data Augmentation. *WWW 2022 - Proceedings of the ACM Web Conference 2022*, *1*(1), 1070–1079. https://doi.org/10.1145/3485447.3512156

Xian, Z., Li, Q., Huang, X., & Li, L. (2017). New SVD-based collaborative filtering algorithms with differential privacy. *Journal of Intelligent and Fuzzy Systems*, *33*(4), 2133–2144. https://doi.org/10.3233/JIFS-162053

Xieling Chen, D. Z. C. ， H. X. (2024). Deep neural networks for the automatic understanding of the semantic content of online course reviews. *Education and Information Technologies*, *29*(4), 3953–3991.

Xu, S., Ji, J., Li, Y., Ge, Y., Tan, J., & Zhang, Y. (2023). Causal Inference for Recommendation: Foundations, Methods and Applications. *ArXiv Preprint ArXiv*, 1–28.

Xu, X., Pang, G., Wu, D., & Shang, M. (2022). *Joint hyperbolic and Euclidean geometry contrastive graph*. *609*, 799–815. https://doi.org/10.1016/j.ins.2022.07.060

Yabing, J. (2013). Research of an Improved Apriori Algorithm in Data Mining Association Rules. *International Journal of Computer and Communication Engineering*, *2*(1), 25–27. https://doi.org/10.7763/ijcce.2013.v2.128

Yadati, N., Yadav, P., Louis, A., Nimishakavi, M., Nitin, V., & Talukdar, P. (2019). HyperGCN: A new method of training graph convolutional networks on hypergraphs. *Advances in Neural Information Processing Systems*, *32*(NeurIPS 2019).

Yadav, A., Srivastava, G., & Kumar, S. (2024). A Hybrid Approach to Movie Recommendation System. *Journal of Management and Service Science (JMSS)*, *04*(062), 1–14. https://doi.org/10.54060/a2zjournals.jmss.62 http://creativecommons.org/licenses/by/4.0/

Yang, M., Zhou, M., Li, Z., Liu, J., Pan, L., Xiong, H., & King, I. (2022). Hyperbolic Graph Neural Networks: A Review of Methods and Applications. *ArXiv Preprint ArXiv:2202.13852.*

Yao, M., Sahebi, S., & Behnagh, R. F. (2020). Analyzing student procrastination in MOOCs: A multivariate Hawkes approach. *Proceedings of The 13th International Conference on Educational Data Mining (EDM)*, *Edm*, 280–291.

Ye, B., Mao, S., B, P. H., Chen, W., & Bai, C. (2021). *Recommendation in MOOCs* (Vol. 1). Springer International Publishing. https://doi.org/10.1007/978-3-030-82147-0

Ying, R., He, R., Chen, K., Eksombatchai, P., Hamilton, W. L., & Leskovec, J. (2018). Graph convolutional neural networks for web-scale recommender systems. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 974–983. https://doi.org/10.1145/3219819.

© COPYRIGHT UPM

3219890

Yu, J., Luo, G., Xiao, T., Zhong, Q., Wang, Y., Feng, W., Luo, J., Wang, C., Hou, L., Li, J., Liu, Z., & Tang, J. (2020). MOOCCube: A large-scale data repository for NLP applications in MOOCs. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 3135–3142. https://doi.org/10.18653/v1/2020.acl-main.285

Yue, Y., Lin, F., Yamada, K. D., & Zhang, Z. (2023). Hyperbolic Contrastive Learning. *ArXiv Preprint ArXiv:2302.01409*.

Zhang, F., Yuan, N. J., Lian, D., Xie, X., & Ma, W. Y. (2016). Collaborative knowledge base embedding for recommender systems. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13-17-Augu*, 353–362. https://doi.org/10.1145/2939672.2939673

Zhang, H., Huang, T., Lv, Z., Liu, S., & Yang, H. (2019). MOOCRC: A highly accurate resource recommendation model for use in MOOC environments. *Mobile Networks and Applications*, *24*(1), 34–46.

Zhang, J., Chen, Y., Xiao, X., Lu, R., & Xia, S. T. (2022). LEARNABLE HYPERGRAPH LAPLACIAN FOR HYPERGRAPH LEARNING. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *2022-May*, 4503–4507. https://doi.org/10.1109/ICASSP43922.2022.9747687

Zhang, L., Luo, T., Zhang, F., & Wu, Y. (2018). A Recommendation Model Based on Deep Neural Network. *IEEE Access*, *6*. https://doi.org/10.1109/ACCESS.2018.2789866

Zhang, T., Liu, Y., Shen, Z., Ma, X., Chen, X., Huang, X., Yin, J., & Jin, J. (2023). Learning from Heterogeneity: A Dynamic Learning Framework for Hypergraphs. *ArXiv Preprint ArXiv*, *00*(0), 1–15. http://arxiv.org/abs/2307.03411

Zhao, M., Zhong, S., Fu, X., Tang, B., Dong, S., & Pecht, M. (2021). Deep Residual Networks with Adaptively Parametric Rectifier Linear Units for Fault Diagnosis. *IEEE Transactions on Industrial Electronics*, *68*(3), 2587–2597. https://doi.org/10.1109/TIE.2020.2972458

Zhao, Y., Ma, W., Jiang, Y., & Zhan, J. (2021). A MOOCs Recommender System Based on User's Knowledge Background. *International Conference on Knowledge Science, Engineering and Management*, 140–153.

Zhao, Z., Cheng, Z., Hong, L., & Chi, E. H. (2015). Improving user topic interest profiles by behavior factorization. *WWW 2015 - Proceedings of the 24th International Conference on World Wide Web*, 1406–1416. https://doi.org/10.1145/2736277.2741656

Zheng, L., Lu, C. T., Jiang, F., Zhang, J., & Yu, P. S. (2018). Spectral collaborative filtering. *RecSys 2018 - 12th ACM Conference on Recommender Systems*, 311–

© COPYRIGHT UPM

319. https://doi.org/10.1145/3240323.3240343

Zhong, H., Wu, J., Chen, C., Huang, J., Deng, M., Nie, L., Lin, Z., & Hua, X. S. (2021). Graph Contrastive Clustering. *Proceedings of the IEEE International Conference on Computer Vision*, 9204–9213. https://doi.org/10.1109/ICCV48922.2021.00909

Zhu, J., Zeng, W., Zhang, J., Tang, J., & Zhao, X. (2023). Cross-view graph contrastive learning with hypergraph. *Information Fusion*, *101867*, 99.

Zhu, Yaochen, Ma, J., & Li, J. (2023). *Causal Inference in Recommender Systems: A Survey of Strategies for Bias Mitigation, Explanation, and Generalization*.

Zhu, Yifan, Lu, H., Qiu, P., Shi, K., Chambua, J., & Niu, Z. (2020). Heterogeneous teaching evaluation network based offline course recommendation with graph learning and tensor factorization. *Neurocomputing*, *415*, 84–95.

Zhu, Yu, Guan, Z., Tan, S., Liu, H., Cai, D., & He, X. (2016). Heterogeneous hypergraph embedding for document recommendation. *Neurocomputing*, *216*, 150–162. https://doi.org/10.1016/j.neucom.2016.07.030

© COPYRIGHT UPM

BIODATA OF STUDENT

Luo Hao was born in Yunnan, China on 16th February 1988. He received the master's degree from Yunnan University, Yunnan, China. He is currently pursuing the Ph.D. degree in intelligent systems with University Putra Malaysia. He not only has many years of scientific research and teaching experience, but also a lot of engineering practice experience, and he is good at machine learning and deep learning algorithms. His current research interest includes the application of recommendation system for the educational field.

© COPYRIGHT UPM

LIST OF PUBLICATIONS

Luo, H., Husin, N. A., & Aris, T. N. M. (2022). ROME: A graph contrastive multi-view framework from hyperbolic angular space for MOOCs recommendation. *IEEE Access*, 11, 9691-9700.

LUO, H., HUSIN, N. A., ABDIPOOR, S., ARIS, T. N. M., SHARUM, M. Y., & ZOLKEPLI, M. (2024). OBJECT-ORIENTED ONLINE COURSE RECOMMENDATION SYSTEMS BASED ON DEEP NEURAL NETWORKS. *Journal of Theoretical and Applied Information Technology*, 102(3).

© COPYRIGHT UPM