

基于自监督的预训练在推荐系统中的研究综述

杨纪元¹ 马沐阳¹ 任鹏杰¹ 陈竹敏¹ 任昭春¹ 辛鑫¹ 蔡飞² 马军¹

(1. 山东大学 计算机科学与技术学院, 山东省 青岛市 266237;

2. 国防科技大学 系统工程学院 湖南省 长沙市 410015)

摘要: 预训练模型旨在通过自监督学习从大规模无标签的数据中获得与具体任务无关的表征。伴随着深度学习的迅猛发展以及强有力的算力支持, 预训练技术在自然语言处理(NLP)和计算机视觉(CV)领域中众多下游任务上取得了显著的提升。因此, 如何设计能够融合语言、图像、文档和视频等各种信息的预训练模型以更好的服务下游任务成为了当前的研究热点。在推荐系统中用户交互记录的稀疏性问题和冷启动问题由来已久, 同时用户和物品的数量巨大, 如何得到良好的用户表征和物品表征也成为了一大挑战。为了缓解这些问题, 近来涌现了一批研究工作探讨如何将预训练技术应用在推荐场景下并构造预训练任务, 以此提升最终的推荐性能。本文重点综述现有的基于预训练的推荐模型研究进展, 对不同的预训练方法进行分类和比较, 并在三个推荐系统基准数据集上对一些代表性模型进行实验和分析, 相关的数据集和代码已开源, 希望能对相关领域的研究人员带来便利, 最后对预训练的推荐模型的未来发展趋势进行总结和展望。

关键词: 推荐系统; 综述; 预训练模型; 自监督学习

中图分类号: TP391

文献标识码: A

Self-supervised Pretraining for Recommender Systems: A Literature Survey

Jiyuan Yang¹, Muyang Ma¹, Pengjie Ren¹, Zhumin Chen¹, Zhaochun Ren¹, Xin Xin¹, Cai Fei², Jun Ma¹

(1. Department of Computer Science and Technology, Shandong University, Qingdao, Shandong Province, 266237, China;

2. School of systems engineering, National University of Defense Technology, Changsha, Hunan Province, 410015, China)

Abstract : Pretrained models aim to extract task-agnostic representations from large-scale unlabeled data through self-supervised learning. With the rapid development of deep learning and potent computational support, pretraining techniques have achieved remarkable enhancements in various downstream tasks in the fields of Natural Language Processing (NLP) and Computer Vision (CV). Consequently, current research hotspots include devising pretraining models that can effectively integrate various types of information, such as language, images, documents, and videos, to better serve downstream tasks. In recommendation systems, longstanding issues such as the sparsity of user interaction records and challenges related to cold start problems persist. Simultaneously, the enormous numbers of users and items add complexity with the pressing issue of obtaining representative user and item embeddings. To mitigate these problems, a batch of recent studies explores the application of pretraining techniques within the context of recommendation scenarios and the design of pretraining tasks in order to enhance the overall performance of recommendation. This paper extensively reviews the progress in research of recommendation models based on pre-training, classifies and compares different pre-training methods, and conducts extensive experiments and analyses on some representative models using three benchmark datasets for recommendation systems. The datasets and codes have been made open source to facilitate researchers in the field of information retrieval. Finally, it summarizes and prospects the future development trend of recommendation models based on pre-training.

收稿日期: 2023-08-10; **定稿日期:** 20XX-XX-XX

基金项目: 国家自然科学基金 (61902219, 61972234, 61672324, 61672322, 62072279, 62102234, 62202271, 62272274), 山东省自然科学基金 (ZR2021QF129, ZR2022QF004), 国家重点研发项目 (2020YFB1406704, 2022YFC3303004), 山东省科技创新重点项目 (2019JZZY010129), 腾讯犀牛鸟专项研究计划 (JR-WXG2021411)

Key words: Recommendation systems; Survey; Pre-training models; Self-supervised learning

0 引言

随着深度学习的迅猛发展,传统的特征工程已经逐渐被表征学习所替代^[0],如多层感知机(MultiLayer Perceptron, MLP)、卷积神经网络(Convolutional Neural Network, CNN)、循环神经网络(Recurrent Neural Network, RNN)、图卷积神经网络(Graph Convolutional Network, GCN)、注意力机制(Attention)、转换器(Transformer)等,这些神经网络模型能够通过有监督学习有效地捕获数据中所蕴含的特点和规律。近年来这些神经网络模型也被逐渐运用在推荐系统中用以编码用户交互的物品记录^{[2]-[4]},以此获得良好的用户表征和物品表征^{[5]-[7]}。比起传统的基于矩阵分解的机器学习方法,其推荐性能有所提高^{[8]-[9]}。但现实生活中带有标签的数据集非常少且人工标注成本很高,这导致了深度学习模型容易过拟合,并且在特定任务下训练得到的模型不具有很好的泛化性,这些因素都限制了其进一步的发展^{[10]-[11]}。研究者们开始思考如何利用大规模未标注数据进行训练进而得到通用的知识表征,然后在下游任务中利用特定领域数据进行微调,于是预训练模型应运而生^{[12]-[14]}。

以自然语言处理(Natural Language Processing, NLP)领域为例,预训练主要是通过自监督学习从大规模文本数据中获得与具体任务无关的模型,以此获得通用的自然语言表征。这其中蕴含着丰富的隐性语言学规则和人类的常识性知识^[15]。在实际应用中,使用预训练语言模型可以分成两个过程,首先在大量的文本上训练出一个预训练语言模型,然后在下游的任务上,根据实际情况对得到的预训练语言模型进行微调^[16]。用于训练的数据除了文本外,还可以融合图像、视频等多种模态^{[17]-[18]}。此类模型经过微调之后,可用于支持分类、序列标记、结构预测和序列生成等各种任务,以支持构建摘要生成、机器翻译、图片检索、视频注释等各种应用^{[19]-[21]}。预训练模型能够提供良好的参数初始化,不仅可以使下游任务的性能得到提高,还可以降低下游任务所依赖的标注语料的规模,有助于提高泛化性和收敛速度^[10]。

对推荐系统而言,当只涉及用户的历史交互记录时,推荐系统可以被建模成求解指定的物品在特定用户交互记录下可能被交互的联合概率分

布,而自然语言处理中的语言模型也是在建模给定一个序列去预测下一个词出现的概率,因此推荐问题与自然语言处理问题的抽象形式很相似。同时推荐系统中关于用户属性或者物品描述等侧信息(side information)中包含大量文本^{[22]-[23]},因此自然语言处理领域技术的进步也推动了推荐系统的发展。最近也有一些工作开始研究如何将预训练技术与推荐任务相结合^{[24]-[28]}。如Zeng等人^[29]将现有工作分为两类:一类是对预训练模型的迁移应用^{[30]-[32]},另一类是重新构建适合推荐场景的自监督任务来进行预训练^{[33]-[35]}。第一类工作更倾向于借助预训练好的语言或者视觉等模型,将推荐系统中的侧信息进行编码,再以知识迁移的方式把外部知识引入推荐系统,从而获取更好的用户和物品表征来服务主要的推荐模型。虽然这些丰富的侧信息(如知识图谱信息^[36]、用户信息(如年龄、职业等)^[37]、物品信息(如类别、价格等)^[38]以及评论信息^[39]等)通过预训练模型编码可以作为初步的用户和物品表征,但是这些表征中仅仅包含用户自身信息或者物品自身信息,而推荐中最重要的用户和物品之间的交互信息被忽略了,因此这并不是为推荐系统量身打造的预训练技术。

相比于第一类工作,第二类工作则更加关注用户与物品的交互记录,主要针对推荐场景构建预训练辅助任务。虽然推荐场景下的用户交互记录与自然语言处理场景下的文本在表征学习方面都可以抽象成利用最大似然估计建模联合条件概率的形式,但比起自然语言处理领域,用户交互记录的顺序关系中不存在像自然语言中天然的语法规则,同时在推荐系统中,用户数量和物品数量往往巨大,数据稀疏性问题^{[40]-[41]}尤为突出,因此如何发掘推荐数据上的自监督信号存在巨大挑战并成为当下研究的热点^[42]。近期许多工作^{[43]-[46]}开始研究如何根据大量的用户历史交互记录构造不同的自监督辅助任务来进行预训练,以获得海量用户和物品的表示,从而能够一定程度上缓解推荐数据上的稀疏问题^[47]。

本文重点探究第二类工作的进展,按照自监督学习的不同^[48],将现有基于预训练的推荐模型分为两类:生成式自监督推荐预训练和对比式自监督推荐预训练。其中,生成式自监督推荐预训练旨在重建部分或整个原始的用户物品交互序列,代表性做法是自回归(Auto-Regressive, AR)和自动编码器(Auto-Encoder, AE)模型,他们旨在建立用户物品交互序列的联合概率分布,或随机

掩码一些用户/物品(或者相关属性信息)并根据剩余的序列信息去预测这些掩码的用户/物品(或者相关属性信息)。对比式自监督推荐预训练的目标是设计一个编码器以及合理的负采样策略,通过噪声对比估计(Noise Contrastive Estimation, NCE)来指导模型学习出正例和负例之间的差别,或者是通过互信息最大化(Mutual Information Maximization, MIM)来建模输入交互信息中局部和整体之间的关系。为了展示当前预训练推荐模型的进展和性能,本文选取了八个代表性模型(SASRec^[49]、BERT4Rec^[50]、S3-Rec^[42]、ASReP^[51]、CL4SRec^[52]、MrTransformer (PE)^[25]、SGL^[53]和 DHCN^[54]) 在三个不同场景下的推荐数据集上(Amazon-Beauty¹⁾、MovieLens-1M²⁾和 Yelp³⁾)进行了实验。实验结果表明利用预训练技术能够融合丰富的侧信息,有助于提升推荐系统性能。预训练过程中不同的数据增广方式对性能的提升程度不同,还可能会引入对模型有益的偏差。

近期关于推荐系统预训练的学术研究引起了科研人员的关注,相继取得了一系列重要的研究成果,并且已经有一些相关的综述性工作被发表。比如 Zeng 等人^[168]回顾了借助推荐系统预训练的知识迁移,它主要关注的是推荐系统通过利用知识图在不同领域之间传递知识的工作,但是其只涵盖了少量推荐系统预训练的早期工作。Yu 等人^[169]对现有的自监督推荐系统进行了定义,并将现有方法分类为对比式、生成式、预测式和混合式四种方式,并对这四种方式进行了介绍和总结。本文与先前工作的不同主要体现在以下三点:(1) 本文涵盖了基于对比式和生成式的预训练推荐系统的近期研究,并首次对这两种方式进行了详尽的对比分析;(2) 本文首次对现有的基于自监督的预训练推荐系统的代表性方法在三个公共的标准推荐数据集上进行重现和分析,从实验角度更细粒度的分析不同的方法,得出了一些有价值的结论。(3) 实验用到的数据和代码已整理并开源,希望能促进推荐系统预训练方面的研究。

本文的篇章结构如下:第1部分“背景知识”简要介绍了自监督学习和推荐系统任务等背景知识。第2部分“研究现状”总结、对比和分析了三类基于预训练的推荐模型算法。第3部分“实验分析”在三个公共的标准推荐数据集上对八个代表性模型进行实验和分析。第4部分“总结与未来工作展望”对本文进行了简要总结并展望了预训练模型在推荐系统的未来研究趋势。

1 背景知识

本章将介绍预训练推荐模型的相关背景知识,首先介绍推荐系统的三个相关任务,然后介绍自监督学习相关技术。

1.1 推荐系统

为了缓解数据爆炸性增长导致的信息过载问题,推荐系统(Recommendation Systems, RSs)通过发现用户的个性化偏好来预测用户可能感兴趣的物品^{[55][56]}为用户提供了极大的便利。一方面随着电子商务的蓬勃发展,推荐系统在互联网中的重要地位越来越明显^[57]。国内比较著名的电子商务网站如淘宝、京东、苏宁易购等逐渐改变着人们的消费方式。在电商平台上,商品数量和用户规模巨大,用户往往很难从中挑出令自己满意的商品,推荐系统能够根据用户的历史购买记录分析用户兴趣,从而给用户提供一些购买建议。这不仅提升了用户的购买体验,也提高了网站的销售额^[58]。另一方面,智能手机等移动终端的迅猛发展也极大地促进了推荐系统的进步^[59]。用户可以对现实生活的各类信息结合自身体验进行及时的评价和打分,比如景点打分、餐馆打分、电影打分、音乐评论、好友关注等。因此,推荐系统应用广泛,与当前人们的生活息息相关,密不可分。

根据输入和输出的不同,推荐系统中定义了多种推荐任务。下面将具体描述每种推荐任务以及各自的主流研究进展。

1.1.1 评分预测

评分预测任务的主要是探究用户对物品的评分机制,涉及到的数据集主要是由用户评分构成,每一条评分记录可以看作一个三元组 (u, i, r) ,这表示了用户 u 对物品 i 给予了评分 r 。评分预测的目标是通过分析用户的历史评分记录以此来预测用户对其他物品的评分,如图1。

用户-物品 评分矩阵	物品1	物品2	物品3	物品4
用户1	3.0		4.0	5.0
用户2	1.0	2.0	?	3.0
用户3		3.0	4.0	

图1 评分预测推荐任务,根据用户之前对物品的评分预测用户对未交互物品的评分

1) <http://jmcauley.ucsd.edu/data/amazon/>

2) <https://grouplens.org/datasets/movielens/>

3) <https://www.yelp.com/dataset>

一般采用均方误差 (Mean Squared Error, MSE) 或平均绝对误差 (Mean Absolute Error, MAE) 作为损失函数来优化模型的学习。传统的做法可以分为三类, 分别是协同过滤方法、基于内容的方法以及混合推荐的方法。协同过滤主要是通过用户显式反馈 (例如, 用户的评分) 或隐式反馈 (例如, 用户的浏览记录) 来学习用户和物品的表示, 通过分析用户或者物品之间的相似性, 来预测用户可能感兴趣的物品。传统的主流做法大都是基于用户物品评分矩阵进行矩阵分解来进行^[60]。基于内容的方法是根据计算候选物品与用户过去喜欢的物品的相似性来进行推荐, 这其中还可以考虑到文本、图像以及知识图谱等各种辅助信息。混合推荐的方法结合了以上二者的优点, 结合了多种推荐技术。通常做法有加权融合不同推荐模型的推荐结果, 或组合不同数据源的特征或者复杂的多模型级联等^{[61][65]}。随着深度学习的发展, 多层感知机、循环神经网络、卷积神经网络、图卷积神经网络、注意力机制以及自编码等模型和机制被逐渐用于建模不同的信息源 (如物品评论、物品图片、用户的信息) 来学习更加精准的用户和物品表征。其中, 循环神经网络天然适合处理一些序列数据, 如用户的交互记录, 模型在提取用户的偏好时能够结合用户的浏览历史和动作序列, 逐渐丰富了序列推荐的任务^{[49][66]}。在处理长序列的时候, 使用循环神经网络容易遗失掉长程信息。而注意力机制可以通过提取序列或者文本中的重点来克服这一问题。通常用户的交互记录中每一个交互过的物品并不一定反映了平等的用户兴趣, 注意力机制通过给交互物品分配不同权重来刻画更加准确的用户表征^[67]。卷积神经网络适合建模一些文本信息, 如用户对特定物品的评论和被许多用户交互的物品的评论^[68]。基于评论的推荐假设用户对物品的评论能够体现出用户的偏好以及商品自身的特征信息, 这是一种语义信息丰富的隐式特征, 相比于单纯的利用显式评分特征, 文本信息一方面可以弥补评分稀疏性的问题, 另一方面也能够提升推荐系统的可解释性^{[69][70]}。图神经网络擅长捕捉节点间的连接和图数据的表征学习, 而推荐系统中用户和物品之间的交互记录又可以被天然的建模为二部图结构, 因此图神经网络也很适合用来建模用户和物品表征^[71]。同时, 图中还可包含一些用户属性、物品属性以及与物品相关的知识图谱, 知识图谱中丰富的语义关联有助于挖掘潜在联系, 关系的多样性能够帮助扩展用户兴趣提高推荐多样性^{[72][73]}, 通过图神经网络引入边的信息还可以缓

解推荐系统中的数据稀疏性 (data sparsity) 和冷启动 (cold-start) 的问题, 同时利用图技术在知识图谱上做用户兴趣的多跳推理也可以提升推荐系统的可解释性^{[36][74][75]}。

1.1.2 社交推荐

推荐系统是为了推荐用户可能感兴趣的产品, 为用户提供个性化服务逐渐成了很多平台的重要任务。但是很多平台只能获取到用户的隐式反馈信息, 而且每一个用户只在少量的商品上产生行为, 导致平台获取到的数据集非常稀疏, 这对于推荐算法极为不利。而事实上在网络社区中, 往往容易获取到用户的社交关系。这就使得用户的偏好除了可以从自身的行为中获取以外, 还可以从他们好友的行为中获取。基于用户社交关系的推荐假设具有相似兴趣的朋友对彼此有更强的影响, 朋友之间也更容易产生相同的行为^[76]。因此社交推荐引入了用户的社交网络信息, 通过探索用户的朋友的兴趣爱好来提升用户的个性化推荐, 如图2。

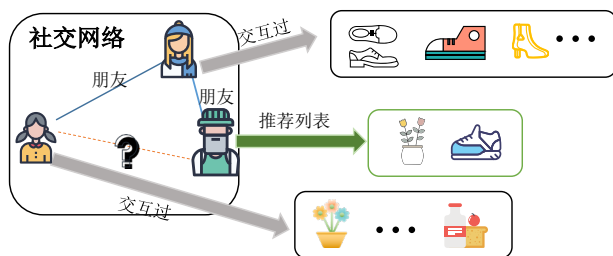


图2 社交推荐任务, 利用用户社交关系进行物品推荐

通过引入社交网络信息, 一方面可以缓解推荐系统冷启动问题^{[60][77][78]}, 另一方面还可以增加推荐的可解释性和用户对推荐的信任度^{[79]-[81]}。对于社交推荐任务存在两种数据, 一种是对物品的交互行为, 一种是对用户之间的社交网络关系。用户的社交网络可以表示为社交网络图, 其中节点代表不同的用户, 边代表用户之间的朋友关系, 用户对物品的行为可以表示为用户物品二部图, 节点为用户和物品, 边代表用户对物品的交互行为。这两种图还可以结合成整个图, 推荐算法通过建模该异构图来获得用户和物品表征^{[32][82]}。图卷积神经网络天然适合处理图数据, 大多数研究通常促使模型分别在社交空间和物品空间分别学习用户表征向量, 然后通过一些融合技术得到最终用户的表征^{[83][84]}。同时还有部分研究提出不同朋友的影响程度不同, 注意力机制被用于区分不同关系的重要性, 从而能够决定信息如何传播

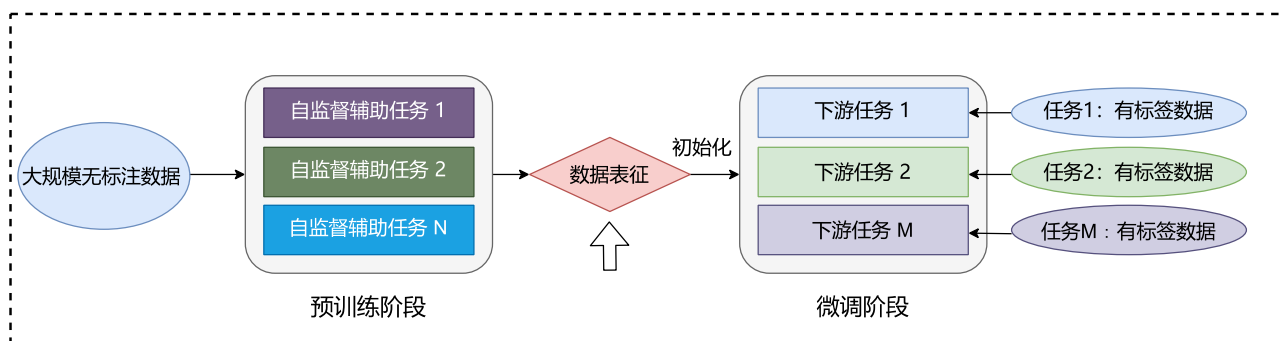


图4 自监督学习流程，在预训练阶段由若干自监督辅助任务学习到的数据表征，用于微调阶段下游任务的初始化

[85][86]。

1.1.3 序列推荐

传统的推荐任务大多没有考虑到用户交互记录中的时间信息，但实际上用户是按照时间顺序来进行交互的，因此该记录可以被建模为一个序列，这其中隐藏了用户在一段时间下隐式的兴趣转移。序列推荐旨在根据用户在某一时间段内的浏览或者购买记录来捕捉用户在短期或者长期下的兴趣爱好，从而很好地预测用户下一个时间点将会交互的物品。该任务假设一段时间内的记录反映了用户相似的兴趣和目的，如图3。例如一段时间内用户的音乐播放列表或者观影序列，往往反映了用户特定的喜好。随着时间的推移，用户的兴趣是动态且多样的，序列推荐的难点就是

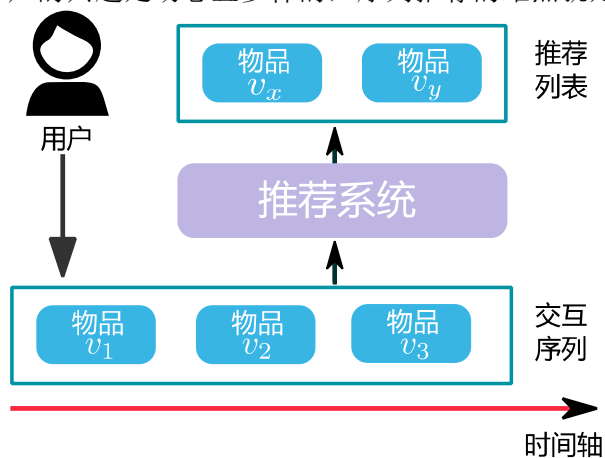


图3 序列推荐任务，根据用户的历史物品交互记录预测用户下次可能会点击的物品

如何准确地捕捉用户的兴趣爱好并以此为基础给出推荐列表^[87]。传统的序列推荐方法一般都是给定上一时刻的交互记录，根据马尔科夫链和马尔科夫决策过程来预测下一个交互^{[88][89]}。这对于建模动态序列数据确实有很大的进步，但是最重要的问题是它们只能建模比较短的序列，当考虑那些较长的序列时，就会导致状态空间迅速增大而变得不可控制^[67]。近年来，深度学习方法日益成熟，神经网络对处理变长的序列数据展示了

无穷的潜力。比起传统的方法，循环神经网络用于序列推荐的建模取得了显著的提升^[90]。同时上下文信息被证明对于行为建模是非常重要的，有研究在序列推荐中融入上下文信息并且提出了基于上下文的循环神经网络模型^{[91][92]}。同时，商品自身蕴含的侧信息也开始被考虑在内，有研究针对如何在序列信息中加入物品属性信息提出了一系列的平行循环神经网络架构^{[93][94]}。用户在物品上的停留时间长短也可能反映了用户不同的兴趣爱好，因此时间信息也被重点建模^[95]。由于用户当前序列中展示的兴趣可能包含长期主要兴趣和当前短期兴趣，有研究提出探索了带有注意力机制的混合编码器来建模用户的长短期爱好^{[67][96]}。这些主流的序列推荐模型都取得了不错的进步，它们的训练方式大多采用下一个预测的物品作为监督信号来指导模型学习。还有一部分研究将用户的交互序列建模成一张序列图，图中的节点表示交互的物品，边代表天然的交互顺序，图神经网络又展示了在图上捕捉序列化知识的能力^[91]。

推荐系统的目的是通过研究分析用户与物品的交互记录来获取用户的偏好从而为用户推荐他们有可能感兴趣的物品。但在现实场景中用户交互过的物品往往非常少，这导致了数据非常稀疏。而且不断会有新的物品和新的用户加入推荐系统造成严重的冷启动问题，这些挑战都会影响推荐算法的效果和性能。比较有效的做法是从其他领域或任务迁移知识或集成异质外部知识辅助推荐模型。而预训练模型通常的做法就是基于大规模无监督数据获取通用知识表征，然后基于下游任务进行微调以实现知识迁移。因此，预训练任务可以被很好地用于地挖掘用户-物品的交互数据以捕捉用户偏好，同时也可以将多种不同任务和数据的信息整合到一个通用的用户/物品表征空间，以适用于不同的推荐场景，从而可以帮助缓解推荐系统的这两大问题。本文重点关注预训练技术在推荐模型中的应用进展。

1.2 自监督学习

有监督学习 (supervised learning) 和无监督学习 (unsupervised learning) 是机器学习两种最基本的学习范式。监督学习旨在利用大量的有标注数据来训练模型, 通过不断计算模型的预测值和数据的真实标签之间的损失来进行反向传播, 从而使模型可以获得识别新样本的能力⁰。但由于其严重依赖于昂贵的人工标签, 因此模型在训练过程中仅仅关注于特定任务, 而忽视了数据本身隐藏的信息, 学习到的表征难以迁移, 同时依赖标签得到的模型也往往易受攻击, 因此监督学习的研究陷入了瓶颈期^[48]。而无监督学习不依赖任何标签值, 其通过对数据内在丰富的特征进行挖掘, 以找到隐藏的特定的数据模式。主要的建模方法有两种, 分别为聚类和降维, 他们旨在寻找数据本身的规律, 但是由于缺乏有标记的输入数据, 输入和输出均是未知的, 因此学习起来更加困难。

自监督学习能够使得神经网络用更少的标签、更少的样本学习到更好的表征^[97]。作为一种很有前途的可替代的学习范式, 自监督学习因其对下游任务有良好的泛化迁移能力而引起了广泛的关注^[98], 许多最先进的模型一直在遵循这一模式^{[12][18]}。例如, Chen 等人^[99]运用了自监督学习中的对比学习策略得到了图片的视觉特征表示, 这些表示能够很好的用于图片分类等任务。自监督学习结合了监督学习和无监督学习各自的优点, 主要是设计不同的自监督辅助任务从大规模的无监督数据中挖掘出数据自身的监督信号^[100], 通过这种方式构造的监督信号对网络进行有监督的训练, 从而可以学习到对下游任务有价值的表征^[16], 主要学习的过程如图4所示。自监督学习包含预训练和微调两个阶段, 在预训练阶段, 模型旨在构建若干个自监督辅助任务从大规模无标签数据中获得这些数据的表征, 在微调阶段, 这些数据表征能够作为良好的参数初始化加速不同下游任务的训练, 这些下游任务均对应各自有标签标注的数据集, 通常以有监督的方式进行训练。自监督学习的辅助任务通常是恢复经过转换、扭曲的部分输入数据或者挖掘与当前输入信息的不同模式、不同部分以及不同形式之间的关系。Liu 等人^[48]将现有的自监督学习大致分为了三类, 分别是生成式、对比式以及对抗式。但整体而言, 生成对抗的思想在推荐系统中的应用属于知识借鉴的范畴, 但针对特定的推荐场景, 核心问题还是如何建模用户表示和物品表示, 因此本文关于预训练技术在推荐系统中的近期研究主要集中在生成式

和对比式两方面, 下面将详细介绍这两种方法。

1.2.1 生成式自监督学习

生成式自监督学习旨在通过训练一个编码器 (encoder) 和解码器 (decoder) 来重建部分或整体的原始数据。训练过程首先是将输入数据编码成显式向量, 再基于该向量重构原始的输入数据。一类典型方法是自回归模型 (Auto-regressive Model), 目标是建立输入数据的联合概率分布, 通过链式法则, 这可以被转化成条件概率分布的乘积。以自然语言处理领域为例, 语言模型 (Language Model, LM) 的目标通常是最大化正向自回归因子分解的似然。类似的, 在推荐系统中, SASRec 模型^[49]将用户的交互记录建模为一种序列, 每次预测下一个时间步被交互的物品。这种模型能够很好的建模上下文信息, 训练的标签即为下一个时间步被交互的物品, 训练数据和标签都源于该条序列数据本身, 是自回归生成式自监督学习在推荐场景下的成功应用。但在这种建模方式下, 信息流一般是单向的。还有一类典型方法是自编码模型, 目标是从损坏或转换的输入中重构部分输入。首先编码器对输入进行编码得到隐含表征, 再通过解码器重构输入, 目标是使得输入和重构数据尽可能的相近, 模型通常通过最小化重构损失进行优化。以自然语言处理领域为例, 掩码语言模型 (Masked Language Model, MLM) 假设预测的单词是相互独立的, 首先将序列中的单词随机替换为掩码标记 (mask), 并根据剩余的上下文信息去预测它们。类似的, 在推荐系统中, BERT4Rec 模型^[50]将用户的交互序列中的物品进行随机掩码操作, 然后让模型根据上下文交互记录去预测被掩码的物品。与自回归模型相比, 自编码模型可以访问双向的上下文信息, 训练的标签源于被替换的部分输入^[101]。

生成式自监督学习模型架构可以看作是由编码器和解码器构成, 其能够在没有下游任务假设的情况下恢复原始数据分布, 这使得生成式自监督学习天然适合分类、生成等任务。但由于其大多数情况下通过最大似然估计来进行优化, 罕见样本将会造成损失值无穷大, 这就导致了其对于罕见样本非常敏感。同时该学习过程通常采用点到点 (point-wise) 的目标损失函数去建模, 因此得到的特征往往是样本级别或者像素级别等低层次的, 这导致模型无法捕捉到更高层次的语义信息, 限制了生成式模型的发展。

1.2.2 对比式自监督学习

大多数表征学习任务都希望对样本之间的关

系进行建模，因此在过去生成式模型占据着表征学习的主流地位。随着计算机视觉领域中 MoCo^[102] 和 SimCLR^[99] 等模型的提出，对比学习取得了突破性进展。对比式自监督学习旨在让模型在对比中学习不同样本的潜在特征的不同，其训练编码器将一组输入编码为显式向量，并在同一表示空间内度量他们之间的相似性（例如，噪声对比估计、互信息最大化）。一类是基于实例-实例的对比方法，这类做法通常根据一定的规则将输入数据进行数据增广（比如将图片旋转、缩放、裁剪等），从而获得同一条数据的多个视角（view），然后通过合理的采取负例来让模型判别这些输入的正样本和负样本数据的不同。不同编码器的设计能够影响这些样本投射到空间的表示，不同负例的选择会影响暴露给模型的信息，这些因素都会影响学习到的数据特征表示。

另一类是基于上下文-实例的对比方法，其旨在建模样本局部特征与上下文表征或者全局表征之间的归属关系。由于许多数据蕴含丰富的时空以及时序顺序关系，许多模型都把识别其各部分之间的相对位置作为辅助任务，如预测视频帧的前后关系、图片中不同区域的相对位置^{[103]-[105]}、或者句子之间的上下文关系等。以自然语言处理领域为例，训练 BERT 的下句预测（Next Sentence Prediction, NSP）任务中要求模型判断两个句子是否具有上下文关系^[17]。除此之外，还有模型更加侧重于学习局部信息和全局信息之间的明确归属关系，而忽视局部之间的相对位置信息。模型通常通过噪声对比估计来最大化互信息的下界来进行优化。以自然语言处理为例，InfoWord 模型^[106] 提出最大化句子的整体表征和句子中 n-grams 之间的互信息。类似的，在推荐系统中，S3-Rec 模型^[42] 提出最大化用户整条交互序列表征与个别物品的描述信息表征之间的互信息来学习用户和物品表征。

比起生成式自监督学习，对比式自监督学习的模型架构中只保留了编码器，这要求模型必须重构原始输入，而是强调在对正例和负例的对比中学习输入数据在特征空间中的差异，这适用于分类等下游任务，而且度量学习的目标可以帮助模型捕捉到更高层次的语义信息。但这个过程首先要保证负采样策略和数据增广的方法是有效的，这两点通常有很大的挑战性，也是当前预训练推荐模型的热点。

2 研究现状

本文重点研究构建适合推荐场景的预训练工

作，并按照自监督学习不同^[48]，将现有工作主要分为两类：生成式自监督推荐和对比式自监督推荐。下面首先汇总当前的预训练推荐模型，从几个不同的角度对现有模型做出大致的分类，然后详细展开介绍几种代表性模型。

2.1 基于预训练的推荐模型汇总

本节将从自监督方法分类、推荐任务、编码器、自监督辅助任务、是否用到侧信息等方面对当前基于预训练大部分推荐模型进行汇总整理，结果如表 1 所示。

从表 1 中可以看到，前 14 行属于生成式自监督推荐模型，后 10 行属于对比式自监督推荐模型。具体地，第 1-5 行属于生成式基于自回归方式的推荐模型，第 6-14 行属于生成式基于自编码方式的推荐模型，第 15-19 行属于基于实例-实例的正负样例对比的推荐模型，第 20-24 行属于基于上下文-实例的局部整体关系对比的推荐模型。这其中涉及到的推荐任务主要是序列推荐、评分预测以及与用户属性相关的预测任务。与此对应地，大多数模型采用的编码器都是 BERT 和 Transformer 等序列编码模型。在预训练过程中大多数模型采用各种预测任务作为自监督辅助任务，例如下一个物品预测、用户/物品属性预测、图重构、评论句子/单词预测等。利用到的信息除了最基本的用户-物品交互信息之外，部分模型还融入了物品相关的属性信息以及用户社交关系信息。为了获得数据的正例，其中对数据的增广方式可以分为三类，分别为针对交互记录级别、物品级别、属性级别，例如对用户的交互序列的片段进行掩码操作属于在交互记录级别的增广，对用户交互序列中的物品进行掩码和乱序操作属于在物品级别的增广，对用户/物品的属性进行掩码和随机删除属于在属性级别的增广。

2.2 生成式自监督推荐预训练

生成式自监督推荐旨在训练一个编码器来编码用户-物品的交互信息，并且训练解码器来重构原始输入的部分或整体信息。

一类典型方法是采用自回归的方式来建模用户的交互序列。如图 5 所示，对于用户的一条交互序列 $v = \{v_1, v_2, \dots, v_{t+1}\}$ ，基于自回归方式的模型将时刻 $1 \sim t$ 的子序列作为输入并将时刻 $2 \sim t+1$ 的子序列作为输出，用以建模用户的交互序列信息。自回归语言模型的目标通常是最大化正向自回归因子分解的似然，即建立当前输入序列的联合分布，这可以被分解为条件概

率的乘积, 每个变量概率依赖于之前的变量:

$$\max_{\theta} P_{\theta}(x) = \sum_{t=1}^T \log P_{\theta}(x_t | x_{1:t-1}) \quad (1)$$

其中, x_t 表示用户在 t 时刻交互物品 x , $x_{1:t-1}$ 表示用户在时刻 1 到时刻 $t-1$ 期间交互的物品序列, T 是整个交互序列的长度, θ 表示要更新的模型参数, 条件概率表示给定用户先前的交互物品序列为条件, 计算用户 t 时刻交互物品 x_t 的概率。序列推荐中下一项物品预测的标签来源于该序列数据本身, 因此该项任务可以被定义为自监督学习, 近几年来对序列推荐任务的建模受到了大多数学者的关注。

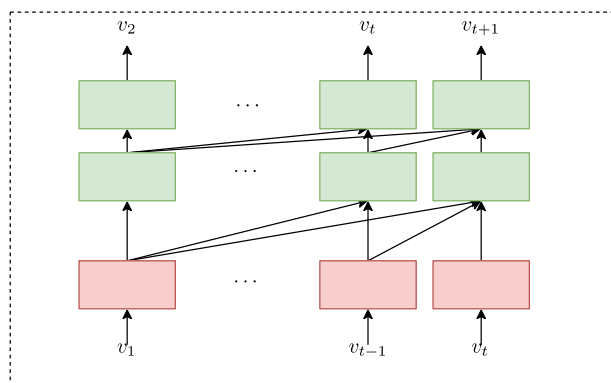


图5 自回归式模型 (以 SASRec^[49] 为例)

Kang 等人^[49]试图从用户的交互序列中识别哪些物品是“相关”, 并通过应用自注意力 (Self-Attention) 机制来预测下一个用户最有可能交互的物品。Liu 等人^[51]指出 Transformer 结构针对短序列建模效果不好, 他们首先采用逆序的 Transformer 结构进行预训练, 以此生成该条序列的伪标签来增长原始的输入序列, 然后采用正序的 Transformer 结构预测下一个物品。例如, 用户只交互了三个物品 ($A \rightarrow B \rightarrow C$), 模型首先通过将该序列逆序输入 Transformer ($C \rightarrow B \rightarrow A$), 然后生成该序列的伪标签以增长该短序列 ($C \rightarrow B \rightarrow A \rightarrow D \rightarrow E$), 然后再将该序列正序输入模型以预测下一个会被用户点击的物品 ($E \rightarrow D \rightarrow A \rightarrow B \rightarrow C$)。Xu 等人^[91]首先将用户交互的序列建模成图结构, 其中每个物品当作节点, 物品之间的先后顺序当作边, 通过图神经网络从而得到每一个物品的表征。接着利用 Transformer 结构对交互序列进行从左到右的建模, 以此预测用户下一个有可能交互的物品。自回归模型的优点是能够很好地建模上下文依赖关系, 但大多数情况下只能建模单向的信息流, 无法利用左右两侧的信息。

另一类典型的方法是采用自编码的方式来建模, 它的目标是从修改之后的输入序列中重构局部或原始的输入^{[135]~[138]}。如图 6 所示, 对于用户的一条交互序列 $v = \{v_1, v_2, \dots, v_t\}$, 自编码方式的模型将输入中的部分物品进行掩码 (mask) 操作, 在对应的输出位置重现该物品, 用以建模用户的交互序列。自编码模型主要是针对输入的用户交互序列 s 构建一个编码器, 通过编码得到序列的隐含表征 h , 再通过解码器中得到重构的输入序列 \hat{s} :

$$\begin{aligned} h &= f_{\text{encoder}}(s), \\ \hat{s} &= f_{\text{decoder}}(h). \end{aligned} \quad (2)$$

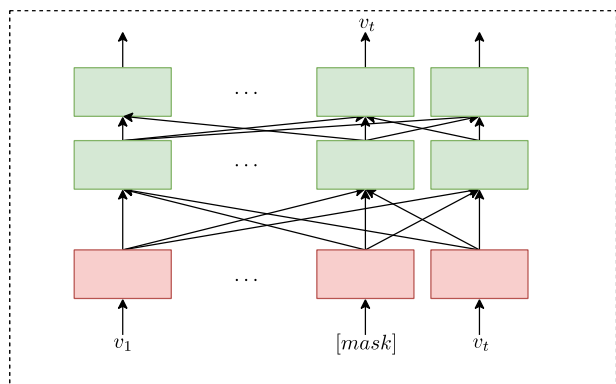


图6 自编码式模型 (以 BERT4Rec^[50] 为例)

模型优化目标使得原始输入 s 和重构输入 \hat{s} 尽可能地接近。针对推荐任务, 目前的研究大多是采用掩码物品或者其他信息的方式来还原被掩码的物品或其他信息^{[139]~[146]}。Sun 等人^[50]采用类似 BERT 的预训练方案, 在预训练过程中, 首先随机挑选用户的交互序列中的一些物品进行掩码操作 ($A \rightarrow [MASK] \rightarrow C$), 再根据剩余的物品 (A 和 C) 去预测序列中的掩码物品 ($[MASK] \rightarrow B$)。在微调阶段, 为了匹配序列推荐的任务, 只对用户交互序列的最后一个物品进行掩码操作, 以此来预测最后一个物品。Yuan 等人^[131]提出采用预训练-微调的策略来在数量较多的源域进行预训练, 得到用户的表征后迁移至数据稀疏的目标领域以此来服务下游任务; 在预训练阶段, 他们随机屏蔽序列中一定比例的物品, 然后预测被屏蔽的物品以获得用户偏好表征。在此基础上, 他们进一步指出尽管这取得了一些积极的知识迁移, 但它所服务的下游任务都是单独训练的。如果涉及到相同的用户, 这些任务可能有实质性的相似之处, 但是顺序的对下游任务挨个训练可能会导致灾难性的遗忘^[126]。因此他们提出基于人工神经网络的用户表征模型^[126], 它不仅提供通用的用户表征, 而

表 1 基于预训练的推荐模型汇总

行号	自监督	推荐任务	模型名称	编码器	自监督辅助任务	侧信息
1	生成式	序列推荐	SASRec ^[49]	Transformer	下个物品预测	×
2	生成式	序列推荐	AsRep ^[51]	Transformer	逆序序列生成	×
3	生成式	序列推荐	GC-SAN ^[91]	GCN & Self-attention layer	下个物品预测	×
4	生成式	序列推荐	SQN-SAC ^[41]	不限	下个物品预测	×
5	生成式	用户属性预测	Conure ^[126]	TCN	物品推荐和用户属性分类	√
6	生成式	序列推荐	BERT4Rec ^[50]	Transformer	掩码物品预测和下一个物品预测	×
7	生成式	评分预测	PMGT ^[127]	Transformer	图结构重构和掩码节点特征预测	√
8	生成式	序列推荐	UPRec ^[128]	BERT	掩码物品预测、用户属性预测和社交关系检测	√
9	生成式	评分预测	U-Bert ^[129]	BERT	掩码评论预测和观点打分预测	√
10	生成式	序列推荐	IERT ^[130]	BERT	下个单词/句子预测	√
11	生成式	序列推荐&用户属性预测	PeterRec ^[131]	NextITNet	下个物品预测	×
12	生成式	用户表征学习	ShopperBERT ^[28]	BERT	下个物品预测	√
13	生成式	物品推荐	CHEST ^[33]	Heterogeneous Subgraph Transformer	掩码节点/边预测和元图类型预测	√
14	生成式	序列推荐	MrTransformer (PE) ^[25]	Transformer	兴趣分离和重构	×
15	对比式	物品推荐	MSSL ^[132]	DNN	剪裁/掩码物品特征	√
16	对比式	物品推荐	CLRec ^[133]	不限	×	×
17	对比式	序列推荐	CL4SRec ^[52]	Transformer	剪裁/掩码/重新排序物品特征	×
18	对比式	序列推荐	DuoRec ^[43]	Transformer	采样具有相同目标物品的序列作为正例	×
19	对比式	物品推荐	SGL ^[53]	GCN	边/节点随机删除、随机游走	×
20	对比式	社交推荐	SEPT ^[82]	GCN	构建不同视图的图	√
21	对比式	序列推荐	S3-Rec ^[42]	Transformer	学习物品特征、物品、子序列以及序列之间的关系	√
22	对比式	序列推荐	DHCN ^[54]	GCN	构建多视图超图	×
23	对比式	序列推荐	Disentangled ^[134]	Transformer	学习序列中前半部分和后半部分之间的联系	×
24	对比式	社交推荐	S2-MHCN ^[83]	GCN	构建多视图超图	√

且在整个生命周期中具有连续学习能力，能够基于先前获得的知识快速学习新能力，避免遗忘旧知识。Liu 等人^[127]考虑了物品的多模态信息及其关系来学习物品表征，首先他们将物品之间的关系建模成一张同质图结构，其中节点是物品，边是物品之间的关联关系（如共同购买关系）。然后利用物品的多模态信息（文本信息，图片信息等）对图进行初始化，以获得物品的表征。在预训练阶段，对随机图游走获得的序列通过 Transformer 结构学习到更高阶的表征后，完成图结构预测以及掩码物品预测等任务。在微调阶

段，该模型可用于物品推荐，物品分类以及点击率预测任务。Xiao 等人^[128]提到之前用于预训练的工作忽视了异构的用户信息。因此他们提出使用用户信息以及社交网络来在预训练过程中构建自监督信号来帮助模型学习用户表征。整体网络采用 BERT 结构，在预训练阶段，他们提出三种任务，分别是随机掩码物品预测、用户属性预测以及社交网络监测。例如通过 BERT 得到用户交互的序列表征作为用户表征，以此预测用户的某个属性，如年龄，性别等。在学习到良好的用户表征和物品表征之后，在微调阶段，模型可用于物品

推荐。Qiu 等人^[129]指出用户的评论中反映了用户真实的兴趣。他们采用 BERT 的模型结构,在预训练阶段,模型关注信息充足的领域对用户表征进行建模。文中提出两种预训练策略,一种是掩码评论单词预测,即随机将用户评论中的单词进行掩码标记(The food [MASK] delicious),然后预测被遮掩的单词([MASK] \rightarrow is),另一种是评分预测,即根据用户给该物品的评论去预测用户的打分情况。在学习到用户表征后,他们将表征迁移到用户信息较少的领域。在微调阶段,他们只关注评分预测这一种预训练任务。Yang 等人^[130]指出现有深度学习大多使用静态的物品表征,这没有考虑到不同交互信息的上下文信息。他们提出一种基于上下文的序列推荐模型,在线下采用类似 BERT 结构训练物品表征,然后线上根据不同的任务再去进行微调。输出的上下文物品表征用于捕获用户的顺序行为以进行推荐。Shin 等人^[28]指出对于电商领域学习通用的用户表征将促进广泛的下流任务,如用户分析、目标定位和推荐任务。他们提出一个基于 Transformer 的针对电商场景的大型预训练模型,预训练任务是分别预测被掩码物品的四种特有类别属性和五类交互时间(年、月、日、周、时)。文中指出经过预训练的用户表征的在下流任务中的表现普遍比从头开始训练的效果更好,并且可以一定程度上缓解冷启动问题。比起自回归模型,自编码模型可以充分利用左右两侧的信息流对序列进行建模。

2.3 对比式自监督推荐预训练

对比式自监督推荐目标是设计一个编码器以及合理的负采样策略,通过噪声对比估计来指导模型学习出正例和负例之间的差别,或者是通过互信息最大化来建模输入交互信息之间的局部和整体之间的关系。

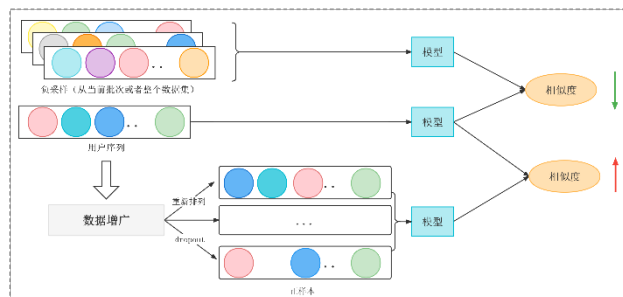


图7 基于不同用户交互记录关系的对比,将用户交互记录进行乱序、删除等操作后的序列作为正例,采样其他用户的交互序列作为负例,采用噪声对比估计指导模型学习

一类典型方法是对用户交互的物品之间的关

系进行建模,这类方法往往注重不同的数据增广方法,以此构造正负样本,然后采用噪声对比估计来指导模型学习^{[147]-[152]},如图7。

$$L = E_{x, x^+, x^k} \left[-\log \left(\frac{e^{f(x)f(x^+)}}{e^{f(x)f(x^+)} + \sum_{k=1}^K e^{f(x)f(x^k)}} \right) \right] \quad (3)$$

其中, x^+ 属于样本 x 增广后的正例, x^k 属于样本 x 的负例, K 为采样负例的个数。Xie 等人^[52]对用户的交互行为提出三种数据增广的方法,分别是随机以一定比例将交互的物品删除,或者随机对交互的物品进行掩码操作以及打乱交互序列。例如,用户交互的序列为 $(A \rightarrow B \rightarrow C)$,通过随机删除操作可得序列 $(A \rightarrow C)$,通过随机掩码操作可得序列 $([MASK] \rightarrow B \rightarrow C)$,通过随机打乱操作可得序列 $(C \rightarrow A \rightarrow B)$ 。模型假定对同一组特征的不同数据增广方式后得到的新特征互为正例,与其他物品特征互为负例。模型采用 SASRec^[49]的基本结构,在预训练阶段,随机选取三种数据增广方式中的两种进行编码得到相应的序列表征,然后优化同一条序列的不同增广表征更接近。在微调阶段,采用物品推荐的任务来训练模型。除了对物品级别的数据增广操作,Yao 等人^[132]提出对输入物品的稀疏分类特征进行掩码或者删除操作,以此来学习两组分类特征之间的内部关系。例如,用户交互的物品有特征 (X, Y, Z) ,通过随机掩码操作可得特征 $(X, [MASK], Z)$,通过随机删除操作可得特征 (Y, Z) 。模型假定对同一组特征的不同数据增广方式后得到的新特征互为正例,与其他物品特征互为负例。Qiu 等人^[43]提出一种内存增强的多实例对比预测编码模型,其采用对比预测编码框架,并根据推荐场景设计了基于多个实例的噪声对比估计损失,即在预测下个物品任务中,引入多个正样例和负样例。Wu 等人^[53]将用户的交互记录建模为用户-物品二部图的结构,他们提出三种在图上进行数据增广的方法。分别是对节点进行掩码,对边进行掩码以及图随机游走,以此生成图的三种不同增广数据。在预训练阶段,首先随机选取两种增广数据,通过图神经网络进行编码可得到每个节点的不同表征,然后优化同一张图的不同表征更接近。在微调阶段,采用图表征学习的任务来优化节点表征。Yu 等人^[82]考虑了社交网络中同质性的理论,对于同一个用户而言,他同时出现在用户-物品组成的二部图中以及社交网络中。文中提出将社交网络的用户表征进行数据增强,与二部图中的表征进行对齐,

表2 本文使用的数据集统计特性

数据集	用户数	物品数	属性数	交互数	最小长度	最大长度	平均长度	平均属性	稠密度
Amazon-Beauty	22,364	12,102	2,230	194,687	5	50	8.7	3.93	0.07%
MovieLens-1M	6,040	3,352	18	269,721	17	50	44.6	1.70	1.33%
Yelp	22,845	16,552	1,158	237,004	5	50	10.3	4.92	0.06%

随后三个编码器生成的三个用户表征均被认为是同一个用户的表征，以此来构建对比信号。除了关注于对输入的样例数据进行增广操作，还有一些研究关注于负例的选取。Zhou 等人^[133]指出在线的推荐系统面临着严重的暴露偏差，因此设计合理的采样策略对于提升推荐系统性能至关重要，并且他们从理论上证明了优化对比损失等同于优化减小暴露偏差。传统的负采样经常采用 in-batch 策略，即采用同一个 batch 内部其他样例当作负例。文中提出一种高效的负采样机制以缓解暴露偏差，即维护一个固定大小的队列，用来存放训练过的样例，从而增多负例的选择。为了提高效率减小运算复杂度，他们提出在队列中存储对应物品的表征。

另一类典型方法是通过重点建模交互序列的局部特征与全局上下文表征之间的归属关系，其目标是使局部表征与全局表征更相近^{[153]~[159]}，如图8。Zhou 等人^[42]指出当前针对序列推荐的现有工作中大多过于关注最后的推荐性能而只采用下个物品推荐的任务，没有充分挖掘上下文交互之间的关联关系，这很容易受到数据稀疏等问题的影响。文中提出一种基于互信息最大化的自监督的预训练方法，它能够在物品属性、物品、子序列以及序列之间捕捉局部和全局的相似性。Ma 等人^[134]指出已有的序列到物品的预测方法容易降低推荐系统的多样性。文中提出在子序列之间挖掘额外的自监督信号，首先将用户交互序列划分成两个子序列，然后对每一个子序列分别分离出用户的意图表征，再通过减小用户意图相近表征之间的距离来学习物品表征。

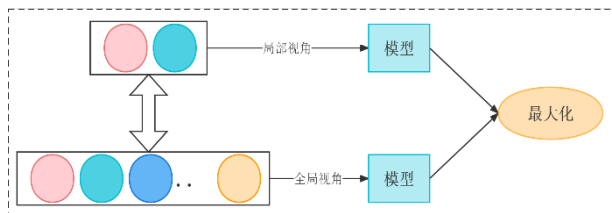


图8 基于同一用户交互记录局部与整体关系的对比，最大化两种视角下用户交互记录的互信息

除此之外，还有一些研究将用户交互记录建模为超图（hypergraph）结构，通过在超图上捕捉局部和整体关系来优化用户表征。Xia 等人^[54]提出一种双通道超图卷积神经网络，首先将用户

交互序列建模为两种超图，分别用来捕捉不同物品之间的关系和不同序列之间物品的共现频率，然后通过各自对应的超图卷积网络编码出不同的序列表征。在预训练阶段，其假定同一条用户交互序列构造的两种超图互为正样本，而和其他交互序列互为负样本，其通过最大化同一条序列在不同超图中学习到的序列表征之间的互信息来更新序列表示和物品表示。Yu 等人^[83]指出推荐系统由于数据稀疏性，当前研究通常利用社交关系来提升推荐性能。但用户交互复杂多样且用户之间的关系可能是高阶的，他们提出利用超图来建模社交推荐。文中提出了一种多通道的超图卷积网络，网络中的每个通道都对超图进行编码，并通过超图卷积得到各自的表征，然后聚合多通道学习到的嵌入信息，以此获得全面的用户表征形式来生成推荐结果。为了弥补聚合操作的损失，文中提出将自监督学习集成到超图卷积网络的训练中，通过最大化超图分层互信息来获得高阶用户关系表征。

3 实验分析

为了探究当前预训练技术在不同推荐任务以及信息利用方式的性能表现，本节挑选了八个代表性预训练推荐模型在三个不同场景下的公共推荐数据集上进行了实验，重点展示与对比不同预训练模型的推荐性能。

3.1 数据集

本工作采用推荐常用的三个数据集，情况如下：

- Amazon-Beauty：亚马逊相关数据集是从亚马逊公司的产品评论中抓取获得的数据集^[165]。该数据集记录了海量用户从1996年5月至2014年7月对亚马逊网站商品的1.428亿条评论以及商品的元数据，涵盖了众多领域。本文从中挑选了美容领域（Beauty），其中包含22,364个用户和12,102个物品，涉及到194,687个交互记录。
- MovieLens-1M：该数据集是电影领域的基准推荐数据集。本文采用MovieLens-1M版本，其中

包含了由 2000 年加入 MovieLens 的 6,040 名 MovieLens 用户对 3,352 部电影产生的 269,721 个匿名评分。

• Yelp: 该数据集来自美国著名商户点评网站, 囊括了海量用户对各地餐馆、购物中心、酒店等商户的打分、评论等, 包含约 16 万商户、863 万条评论和 20 万张图片数据, 是商业推荐基准数据集。由于该数据集非常庞大, 本文采取与工作^[42]一致的做法, 只使用了 2019 年度的用户记录。其中包含 22,845 个用户对 16,552 个商户的 237,004 条评论。

我们遵循与工作^{[42][49][50]}一致的数据预处理操作。如果用户对某物品进行了评分, 则将该评分转化为隐式的正反馈。首先将所有交互记录按照用户进行分组, 并根据时间对它们进行排序。由于本文不针对冷启动推荐, 本文与工作^{[23][166][167]}保持一致, 过滤掉了交互数量小于 5 的用户以及被交互次数小于 5 的物品, 并且限制用户交互序列长度的最大值为 50。数据集的详细统计学特性如表 2 所示。

3.2 评价指标及实验环境

本文采用留一法 (leave-one-out)^[50]来评估不同推荐算法的性能。首先按照时间顺序对用户交互的序列进行排列, 并将最后一项和倒数第二项作为测试集 (testing set) 和验证集

(validation set), 其余作为训练集 (training set)。在验证阶段, 输入训练集序列, 预测验证集的物品, 在测试阶段, 输入训练集和验证集的序列, 预测测试集的物品。与工作^{[23][165]}保持一致, 按照物品的流行度 (popularity) 为每一个真实的测试集中的物品随机采样 100 个用户未曾交互过的物品作为负例。本文使用指标 $Recall@k$, $MRR@k$ 和 $NDCG@k$ 并控制 $k = 5, 10, 20$ 。假设用户 u 预测的标签物品按照得分排名为 r_u , 评测指标计算如下:

• $Recall@k$: 该指标计算的是真实的用户交互物品在排名前 k 个候选推荐物品列表中的比例, 即如果 $r_u \leq k$, $Recall_u@k = 1$, 否则

$Recall_u@k = 0$ 。

• $MRR@k$: 该指标计算的是真实的用户交互物品在候选推荐物品列表中排名的倒数的平均

值, 即如果 $r_u \leq k$, $MRR_u@k = \frac{1}{r_u}$, 否则 $MRR_u@k = 0$ 。

• $NDCG@k$: 该指标也用于衡量真实的用户交互物品在候选推荐物品列表中的排名, 当排名越高, 值就越大, 即 $r_u \leq k$,

$$NDCG_u@k = \frac{1}{\log_2(1+r_u)}, \text{ 否则}$$

$NDCG_u@k = 0$ 。计算出用户的每个指标后, 单独对所有用户的每个指标进行平均得到最终的对应指标, 例如 Recall 指标, 具体计算方式如下:

$$Recall@k = \frac{1}{|U|} \sum_{u=1}^{|U|} Recall_u@k$$

其中, $|U|$ 表示用户数量。

在一台配置 2 个 Intel Xeon CPU (Gold 6226R @ 2.90GHz) 和 128GB RAM 的 Linux (Centos 7.9) 机器上使用 1 块 GeForce RTX 3090 GPU 进行所有的实验。所有实验均进行 3 次并记录平均值, 作为最终实验结果。优化器、正则化方式、学习率、衰减方式等实验参数均遵循论文给出的指导参数设置。

3.3 对比模型

为了展示当前主流的预训练推荐模型的性能和进展, 本文从现有工作中选取了八个模型进行实验, 具体模型如下:

• SASRec^[49]: 该模型使用一个从左到右的基于自注意的模型来捕捉用户的顺序行为。

• BERT4Rec^[50]: 该模型使用双向 Transformer 网络对用户偏好进行建模。在训练过程中, 它随机掩码用户交互序列中的部分物品, 并根据上下文信息进行预测; 在测试过程中, 它只掩码序列中最后一个物品。

• S3-Rec^[42]: 该模型采用 SASRec 为基础模型, 设计了四个辅助自监督目标。利用互信息最大化技术融合属性、物品、子序列和序列四种信息进行预训练。

• ASReP^[51]: 该模型采用逆序的 Transformer 结构进行预训练, 以此生成序列的伪标签来增长原始的输入序列, 然后采用正序的 Transformer 结构进行预测。

• CL4SRec^[52]: 该模型提出了三种数据增强方法 (随机删除、随机掩码和重排序), 然后通过最大化同一序列的不同增广方法下得到的隐含表示的一致性来进行预训练。

• MrTransformer (PE)^[25]: 该模型通过从不同的用户交互序列中学习出共同的和特有的用户兴趣

表示以此来分离出交互序列背后隐藏的多个用户兴趣表示。其自监督信号是从序列对中进行用户兴趣表示编辑操作从而重构各自序列的原始序列用户兴趣表示。

- SGL^[53]：该模型将用户的交互记录建模为用户-物品二部图的结构，并在图上提出对节点/边进行掩码以及图随机游走的三种数据增广方法。在预训练阶段随机选取两种增广数据通过图神经网络进行编码来得到每个节点的不同表示，通过最大化同一张图的不同表示的距离来学习节点表示。
- DHCN^[54]：该模型将用户交互序列建模为两种超图用来分别捕捉不同物品之间的关系和跨序列

Beauty 和 MovieLens-1M 这两个数据集上均优于传统的自回归式模型 SASRec 和自编码式模型 BERT4Rec。而且从整体上来说，模型 BERT4Rec 要优于模型 SASRec，这与部分工作实验结论一致^{[25][34][134]}。具体地，模型 S3-Rec 比起模型 BERT4Rec，在 Amazon-Beauty 数据集上，指标 Recall@10、NDCG@10 和 MRR@5 分别增长了 15.81%、13.52%和 16.61%；在 Yelp 数据集上，指标 Recall@20、NDCG@20 和 MRR@5 分别增长了 3.11%、1.97%和 6.83%。但是在 MovieLens-1M 数据集上，指标 Recall@5、NDCG@5 和 MRR@5 分别下降了 3.27%、8.13%和 10.38%。本文认为这是不同数据集本身的特性

表 3 在 Amazon-Beauty 数据集上的实验结果，每列中最好的结果用粗体标出，次优结果用下划线标出

Models	Re-call@5	Re-call@10	Re-call@20	NDCG@5	NDCG@10	NDCG@20	MRR@5	MRR@10	MRR@20
SASRec	36.58	46.54	58.67	27.91	31.12	34.17	25.79	26.36	27.19
BERT4Rec	39.85	49.20	61.02	30.49	33.50	36.48	27.39	28.62	29.44
S3-Rec	46.15	56.98	68.76	34.53	38.03	41.00	31.94	<u>32.13</u>	<u>32.94</u>
ASRep	37.71	46.97	58.42	29.05	32.04	34.92	26.18	27.41	28.19
CL4SRec	41.45	51.10	62.52	32.34	35.45	38.32	28.48	28.70	30.49
MrTransformer (PE)	42.28	52.10	63.86	32.68	35.85	38.82	29.50	30.81	31.62
(MrTransformer)	40.15	50.19	62.25	30.80	34.04	37.08	27.70	29.04	29.86
SGL	38.14	49.31	60.63	27.34	30.97	33.82	23.79	25.29	26.07
(SGL-p)	36.10	47.00	59.09	25.99	29.52	32.57	22.66	24.12	24.95
DHCN	<u>43.62</u>	<u>52.83</u>	<u>63.33</u>	<u>34.26</u>	<u>37.24</u>	<u>39.88</u>	<u>31.17</u>	32.29	33.12
(DHCN-p)	43.45	52.53	62.82	34.06	37.02	39.59	30.98	32.20	32.90

之间的联系。在预训练阶段，通过最大化不同超图编码得到的序列表示之间的互信息来学习不同节点表示，其负例由对序列的矩阵表示进行行列变换得到。

对于上述八个模型，本文使用原文中提供的代码，并使用原始论文中建议的超参数。为了便于在本文中获得的結果的可复现性，我们整理并共享了本文中使用的数据集、代码和参数文件，具体参见：
<https://github.com/nancheng58/Self-supervised-learning-for-Sequential-Recommender-Systems>。

3.4 实验结果分析

本节将展示上述八个预训练推荐模型在三个标准推荐数据集上的性能，实验结果如表 3-5 所示，从中可以得到如下结论。

首先，利用预训练技术融合丰富的侧信息有利于提升推荐系统的性能。模型 S3-Rec 在预训练期间融合了来自物品属性、物品、子序列以及序列这四个级别的侧信息，其性能在 Amazon-

造成的影响，从表 2 可以看到，数据集 Amazon-Beauty 和 Yelp 中每个物品的平均属性个数分别为 3.93 和 4.92，而 MovieLens-1M 中每个物品的平均属性个数只有 1.70。同时数据集 Amazon-Beauty 和 Yelp 的物品属性类别数分别为 2,320 和 1,158，而 MovieLens-1M 的物品属性类别数为 18，远低于前两个数据集。这说明当数据集中的属性信息非常稀疏时，模型难以从稀疏的信息中捕捉到用户的兴趣偏好，融合这些信息在预训练阶段反而会给模型引入更多的噪声，导致了模型性能的下降。但比起其他融合额外信息的预训练模型，模型 S3-Rec 在 MovieLens-1M 数据集上的效果次于 ASRep 和 MrTransformer (PE)，这说明在数据稀疏的场景下设计合理的预训练信息融合方式值得进一步探究。

第二，预训练阶段对用户交互序列进行数据增广对推荐性能整体上是有所提升的，并且不同的数据增广方式对推荐性能的影响不同。回顾上述五个利用到数据增广的对比模型，其中 ASRep 是

通过逆序生成序列伪标签进行序列长度增广，CL4SRec 是对序列在物品级别上进行增广，MrTransformer (PE) 是通过挖掘不同序列对之间

图/超图上进行增广也有利于提升性能。比起 BERT4Rec，在数据集 Amazon-Beauty 上，模型 CL4SRec 在指标 *Recall@5* 和 *MRR@5* 上分别增长了

表 4 在 MovieLen-1M 数据集上的实验结果，每列中最好的结果用粗体标出，次优结果用下划线标出

Models	Re-call@5	Re-call@10	Re-call@20	NDCG@5	NDCG@10	NDCG@20	MRR@5	MRR@10	MRR@20
SASRec	77.14	87.24	<u>93.26</u>	59.50	62.81	64.36	54.86	55.00	55.45
BERT4Rec	77.83	<u>87.31</u>	93.05	<u>61.12</u>	<u>64.23</u>	<u>65.68</u>	<u>55.55</u>	<u>56.85</u>	<u>57.25</u>
S3-Rec	75.28	86.21	93.03	56.15	59.73	61.46	49.78	51.29	51.76
ASRep	<u>77.98</u>	87.53	93.36	60.87	64.00	64.48	55.16	56.47	56.89
CL4SRec	73.84	84.16	91.54	56.61	59.98	61.86	50.87	52.28	52.80
MrTransformer (PE)	78.34	87.22	93.03	62.05	64.94	66.42	56.62	57.83	58.24
(MrTransformer)	77.72	85.98	92.28	61.67	64.50	66.19	56.60	57.78	58.26
SGL	61.61	76.82	87.62	43.01	47.92	50.68	36.86	38.89	39.66
(SGL-p)	59.06	76.29	88.21	40.60	46.19	49.23	34.52	36.83	37.68
DHCN	72.89	84.10	91.69	55.61	59.18	61.13	49.95	51.45	51.99
(DHCN-p)	72.39	83.37	91.70	55.30	58.97	60.98	49.63	51.16	51.72

表 5 在 Yelp 数据集上的实验结果，每列中最好的结果用粗体标出，次优结果用下划线标出

Models	Re-call@5	Re-call@10	Re-call@20	NDCG@5	NDCG@10	NDCG@20	MRR@5	MRR@10	MRR@20
SASRec	57.66	78.59	94.12	40.57	47.34	51.33	35.19	37.74	38.88
BERT4Rec	63.32	81.81	93.68	45.65	51.64	54.69	39.82	42.30	43.17
S3-Rec	64.24	83.51	96.60	46.15	52.40	55.77	42.54	42.78	43.74
ASRep	63.55	81.17	91.28	45.05	50.77	53.37	38.95	41.33	42.06
CL4SRec	63.78	81.14	92.22	46.44	52.08	54.93	40.71	43.06	43.86
MrTransformer (PE)	64.80	81.23	92.21	47.74	53.07	55.78	42.11	44.31	45.07
(MrTransformer)	63.97	82.18	93.90	45.99	51.90	54.92	40.05	42.51	43.36
SGL	66.17	<u>82.60</u>	92.23	48.67	54.01	56.49	42.89	45.11	45.82
(SGL-p)	62.57	80.39	92.20	45.64	51.43	54.46	42.89	45.11	45.82
DHCN	64.04	80.00	94.86	48.15	53.69	<u>57.20</u>	<u>42.93</u>	<u>45.21</u>	<u>46.20</u>
(DHCN-p)	64.28	82.35	<u>94.68</u>	<u>48.48</u>	<u>54.01</u>	57.42	43.29	45.57	46.53

的关系来进行增广，SGL 是对序列在图上进行增广，DHCN 是通过生成不同的序列超图来进行增广。从三个数据集的整体表现来看，这些模型得了最优或者次优的性能，尤其在 Amazon-Beauty 和 Yelp 数据集上，这两个数据集的平均序列较短，分别为 8.7 和 10.3，而数据集 MovieLens-1M 的平均序列长度为 44.6。该结论符合此类模型提出的假设，即对于平均序列长度更短的推荐场景更为有效。

由于 MovieLens-1M 平均序列长度足够长，传统的自回归 (SASRec) 和自编码 (BERT4Rec) 模型已经足够捕捉到用户的兴趣偏好，此时挖掘不同序列对之间的关系能够帮助更好的学习用户兴趣表示，模型 MrTransformer (PE) 取得了最优的效果。同时，对序列进行物品级别的增广或者在

4.01% 和 3.98%，模型 DHCN 在指标 *Recall@5* 和指标 *NDCG@5* 上分别增长了 9.46% 和 12.36%；在数据集 Yelp 上，模型 DHCN 在指标 *Recall@20* 和 *MRR@5* 上分别增长了 1.25% 和 7.81%，模型 SGL 在指标 *Recall@5* 和 *MRR@5* 上分别增长了 4.50% 和 7.70%。但在数据集 MovieLens-1M 上，模型 CL4SRec，SGL 和 DHCN 的性能均有所下降，本文认为这是该数据集的平均序列长度已经足够长，模型能够从中学出准确的用户偏好，因此简单的数据增广对性能的提升没有很大帮助。同时图神经网络通常是针对每个节点从其邻域收集特征，以更新其周围的局部图结构的表示。针对长序列构建的序列图，相邻节点之间连接的边增加，可能会导致图神经网络对节点的更新过于频繁，从而出现节点表示过平滑的现象，因此利用图结构建模序列的性能

比起 Transformer 结构性能会有所下降。模型 SGL 的性能在数据集 Amazon-Beauty 和 Yelp 上表现正常，但在数据集 MovieLens-1M 上下降最明显，通过实验分析，该模型可能更适合建模短序列场景，具体分析见章节 3.7。

第三，针对三个实验数据集上的整体结果而言，基于对比式的预训练模型效果整体上稍优于基于生成式的预训练模型。回顾上述模型，SASRec、BERT4Rec、ASReP 和 MrTransformer (PE) 是基于生成式的预训练模型；S3-Rec、CL4SRec、SGL 和 DHCN 是基于对比式的预训练模型。在数据集 Amazon-Beauty 上，S3-Rec 表现最优，DHCN 表现为次优；在数据集 MovieLens-1M 上，优；MrTransformer (PE) 表现最优，BERT4Rec 表现次在数据集 Yelp 上，S3-Rec、SGL 和 DHCN 表现较好。从中可看出对比式预训练模型的推荐性能较优，但这与具体的数据集特性有密切关系，如 Amazon-Beauty 和 Yelp 数据集的平均序列长度都较短，生成式模型（如 ASReP）更依赖序列自身信息，但是序列长度较短的序列存在的噪声更大。对于序列长度的探究将在后面章节 3.6 进行具体分析。同时由于对比式模型需要设计不同的数据增广方式和利用额外信息，数据集的稠密程度、平均序列长度等因素也会影响到推荐的性能，因此对比式预训练推荐模型也还有很大的提升空间，值得进一步研究。

3.5 预训练技术的影响

为了验证预训练技术对于推荐性能的影响，本文去掉了上述模型相应的预训练过程或者去掉了联合学习中相应的自监督损失函数，其结果如表 3-5 所示。值得注意的是，模型 S3-Rec、ASReP 和 CL4SRec 在去掉相应的自监督损失函数后，模型均简化成了 SASRec。

模型 MrTransformer (PE) 去掉用户兴趣编辑的预训练过程后，相当于在模型 BERT4Rec 的基础上增加了对用户兴趣的分离表示学习（即 MrTransformer）。模型 SGL 去掉图上的自监督学习后，相当于利用图卷积神经网络对用户-物品二部图上的节点和边进行编码（即 SGL-p）。模型 DHCN 去掉超图上的自监督学习之后，相当于利用卷积神经网络对序列超图上的节点和边进行编码（即 DHCN-p）。经过对比可以发现，当去除掉预训练过程后，大部分模型的推荐性能都下降了。具体来说，在数据集 Amazon-Beauty 上，模型 S3-Rec 在指标 $Recall@10$ 和 $NDCG@20$ 分别下降了 22.43% 和 19.98%；在数据集 MovieLens-1M 上，模型 SGL 在 $Recall@5$ 和 $MRR@10$ 分别下降了 4.31% 和 5.59%。

这说明预训练技术有助于帮助模型融合额外的信息，同时能够捕捉到数据增广操作对于用户兴趣表示的建模，但也能观察到在数据集 Yelp 上，模型 DHCN 在去掉预训练技术后，指标比之前提升，在 $Recall@5$ 上提升了 0.37%，这可能是由于该场景下引入的有关物品的额外信息反而干扰了模型对于用户兴趣的判断，因此研究在特定场景下的预训练推荐技术是有意义的并且值得进一步探究。

表 6 不同数据集中不同长度的序列占比统计

Length	Amazon-Beauty	MovieLens-1M	Yelp
(0, 20]	21,228 (94.92%)	177 (2.93%)	20,744 (90.80%)
(20, 30]	655 (2.92%)	684 (11.32%)	1,094 (4.78%)
(30, 40]	231 (1.03%)	543 (8.99%)	511 (2.23%)
(40, 50]	250 (1.11%)	4,636 (76.75%)	496 (2.17%)

3.6 不同序列长度的影响

由于上述模型在不同特性的数据集上表现不完全一致，并且这些数据集的序列长度有很大差异，因此本节将探究不同序列长度对于不同模型推荐性能的影响。回顾数据预处理阶段，本文过滤掉了交互数量小于 5 的用户并限制用户交互序列长度最大为 50。因此将用户交互序列按照长度划分为 4 段，分别为不超过 20, 20-30, 30-40, 40-50。所有数据集上不同长度的序列个数以及各自占比统计情况如表 6 所示。接着本文分别在各个数据集上评测了不同长度的序列对指标 $Recall@10$ 和 $NDCG@10$ 的影响，实验结果如图 9-14 所示，从中可以得到以下结论：

首先，总体上对比式模型的性能优于生成式模型，但是随着序列长度的逐渐增长，总体上两类模型的性能差距在变小。Amazon-Beauty 和 Yelp 两个数据集上的平均序列长度较短，而且其中长序列占比很少，两个数据集序列长度和长短序列的占比是相似的，但是模型在两个数据集的各个长度区间上的表现不同。具体来说，在 Amazon-Beauty 数据集上，随着序列长度的增长，大多数模型的性能获得了不同程度的上升，而且生成式模型对比式模型的性能提升更多。这说明虽然对比式模型引入了额外信息或者进行了不同程度的数据增广，提升了性能，但是同时会引入更多的噪声，因此在对长序列进行预测时性能和生成式模型相当，这也揭示了生成式模型的潜力。其次，不同数据集上预训练模型的表现也是不同的。在 Amazon-Beauty 数据集上，大多数模型的表现都是随着序列长度的增加而逐渐上升，但在 MovieLens-1M 和 Yelp 数据集上，大多数模

型在不同长度序列上表现更加稳定。本文认为这可能是由于不同数据集特性造成的, MovieLens-1M和Yelp数据集的序列平均长度均大于Amazon-Beauty的序列平均长度, 最简单的自回归模型在训练过程中已经能够捕捉到良好的序列表示, 因此在Yelp数据集上各个模型之间差异不大, 在测

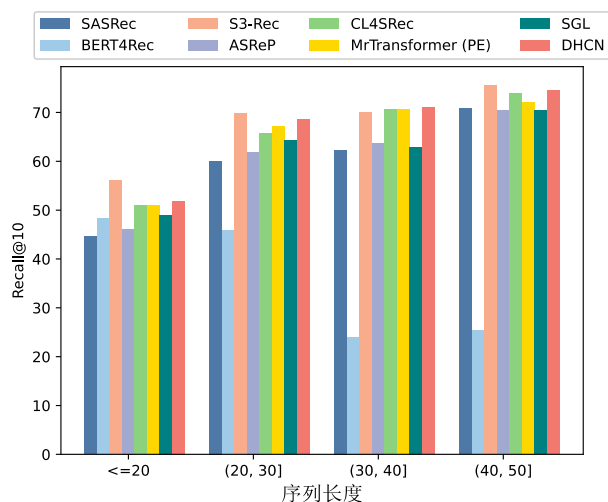


图9 数据集 Amazon-Beauty 上不同序列长度在 $Recall@10$ 上的性能

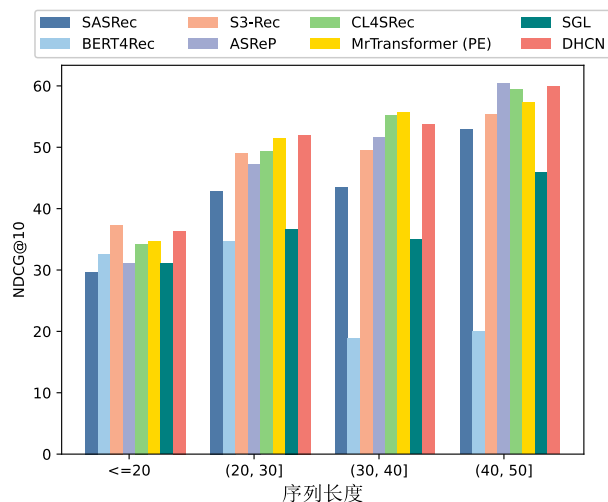


图10 数据集 Amazon-Beauty 上不同序列长度在 $NDCG@10$ 上的性能

试过程中, 序列长度的变化对模型本身的影响也不大。对于BERT4Rec模型而言, 在Amazon-Beauty和Yelp数据集上均呈现出针对长序列性能锐减的情况, 而在MovieLens-1M数据集上性能却随着长度增加而增加。本文认为在训练阶段当数据集的平均序列长度较小且大部分序列都为短序列时, 测试阶段中预测长序列就会较为困难。序列越长, 需要预测的掩码物品就会增多, 因此对于模型预测难度增大。

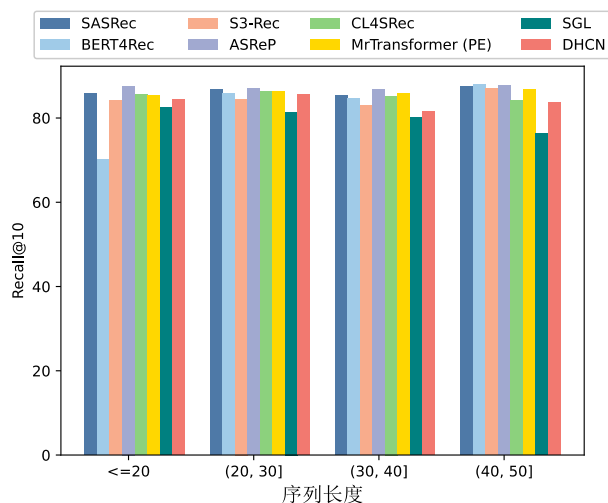


图11 数据集 MovieLens-1M 上不同序列长度在 $Recall@10$ 上的性能

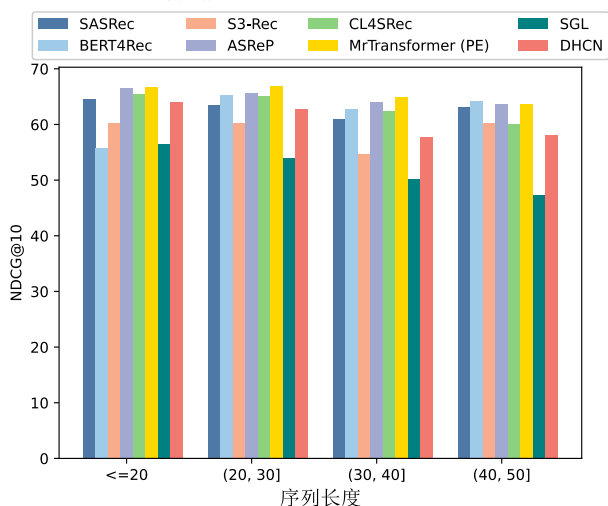


图12 数据集 MovieLens-1M 上不同序列长度在 $NDCG@10$ 上的性能

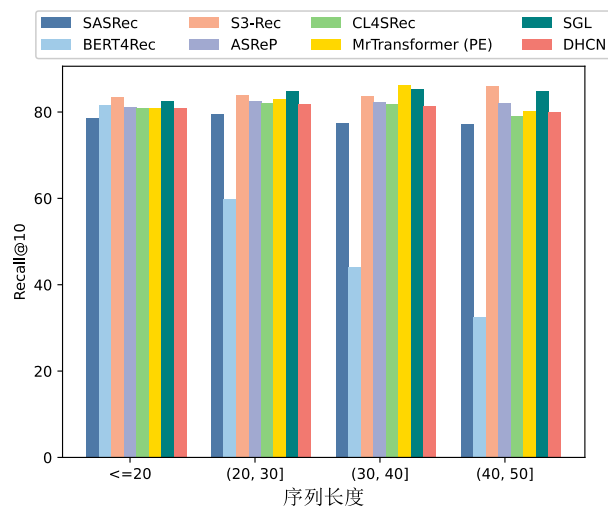


图13 数据集 Yelp 上不同序列长度在 $Recall@10$ 上的性能

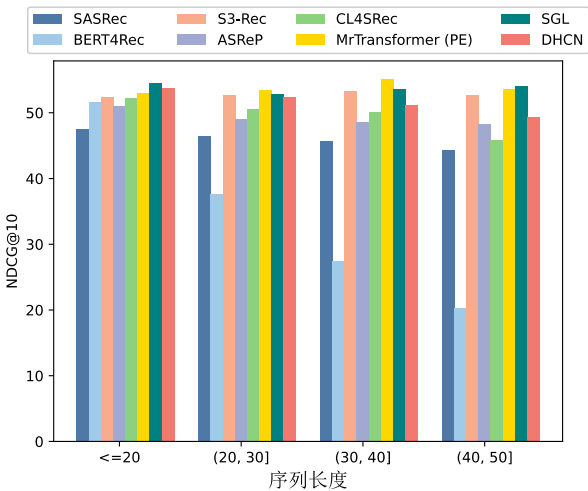


图 14 数据集 Yelp 上不同序列长度在 $NDCG@10$ 上的性能

例无关。另外，模型 SGL 没有显式建模时序信息，所以模型 SGL 可能更适合建模 Amazon-Beauty 和 Yelp 这种数据集上的短序列推荐场景。

4 总结与未来工作展望

预训练技术作为一种辅助手段可以从大规模无标签的用户交互记录中充分的学习良好的用户表示和物品表示，从而帮助解决冷启动和数据稀疏问题，以此提升推荐系统的准确性和多样性。本文对近期基于预训练技术的推荐模型进行了调研，并将现有工作按照自监督学习分为三类，分别为：生成式自监督推荐、对比式自监督推荐和对抗式自监督推荐。其中生成式自监督推荐模型

表 7 在 MovieLens-1M 数据集上对 SGL 模型的调参分析

Models	Re-call@5	Re-call@10	Re-call@20	NDCG@5	NDCG@10	NDCG@20	MRR@5	MRR@10	MRR@20
lr=0.001, layers=3, ssl_weight=0.02	61.51	76.87	88.28	39.66	47.92	50.83	36.77	38.85	39.66
lr=0.001, layers=4, ssl_weight=0.02	62.05	76.89	88.03	43.61	48.44	51.27	37.52	39.53	40.31
lr=0.001, layers=5, ssl_weight=0.02	61.71	77.45	87.81	43.17	48.31	50.94	37.06	39.21	39.47
lr=0.001, layers=4, ssl_weight=0.03	61.77	77.15	88.36	43.11	48.12	50.98	36.94	39.03	39.83
lr=0.001, layers=4, ssl_weight=0.04	61.23	77.05	87.67	43.08	48.23	50.95	37.07	39.22	39.98
lr=0.001, layers=4, ssl_weight=0.05	61.18	76.71	87.24	43.19	48.25	50.93	37.24	39.34	40.09
lr=0.002, layers=4, ssl_weight=0.02	61.29	76.77	87.37	43.22	48.27	50.97	37.23	39.34	40.10

3.7 SGL 模型性能分析

从表 3-5 中可以看出，模型 SGL 在数据集 MovieLens-1M 上性能出现了下降。为了进一步分析模型 SGL 下降的原因，本文对该模型设计了分析实验。对于模型 SGL 而言，其性能在数据集 Amazon-Beauty 和 Yelp 上表现正常，但在数据集 MovieLens-1M 上效果下降明显，为此本文进行了若干组参数分析，具体结果见表 7。本文尝试了学习率（learning rate）、层数（layers）和自监督学习任务比例（ssl_weight）等不同参数。我们可以看到，表 7 中针对模型 SGL 的调参结果和表 3 中 SGL 的结果表现持平，同时该模型在数据集 Amazon-Beauty 和 Yelp 上性能优于模型 SASRec 和 BERT4Rec，这说明该模型在 MovieLens-1M 上的表现与学习率、层数和自监督学习任务比

天然适合生成、分类等任务，但捕捉到的语义信息是低级的且对未出现过的样本较为敏感；对比式自监督推荐模型适合分类等任务，并且能够捕捉到更高层次的语义信息，但设计合适的负采样技术和编码器结构尤为重要，值得进一步探索；对抗式自监督推荐能够在生成对抗网络中引入用户交互的上下文信息，但目前在推荐系统中仍属于知识借鉴的范畴，因此近期工作主要集中在生成式和对比式两方面。为了探究当前模型的性能，本文挑选了八个代表性模型在三个不同应用场景下的标准数据集上进行了实验，结果表明利用预训练技术能够融合丰富的侧信息，有利于提升推荐系统的性能。在预训练阶段对用户交互序列进行数据增广对推荐性能整体上是有所提升的，并且不同的数据增广方式对推荐性能的影响不同，还可能会引入对模型有益的偏差。当交互序列长度小于 20 时，生成式模型占主要优势，随着交互序

列增长,对比式模型逐渐优于生成式模型。最后本文给出关于预训练技术在推荐系统领域可能的未来发展趋势:

1. 设计适合推荐场景的预训练自监督辅助任务。近几年推荐系统的发展很大程度上依托于自然语言处理技术的进步。用户与物品的交互记录存在天然的时序关系,这与自然语言处理中对文本的抽象形式很类似,因此很多任务(如序列推荐)可以直接利用自然语言处理技术。近期预训练技术与推荐系统的结合大多也只是将预训练技术在文本场景或者视觉领域下的尝试套用到推荐场景下,如掩码物品预测,或者预测交互物品的相对位置等。虽然取得了不错的效果,但是建模过程中没有充分考虑到推荐场景的特点,因为推荐系统的核心还是通过分析用户行为来进行用户观点挖掘、用户兴趣演化建模,同时大多数推荐场景均存在数据稀疏性问题,现有工作中这些基于启发式的对用户的交互序列进行数据增广,反而有可能会引入噪声,因此如何设计适合推荐场景的预训练自监督辅助任务从而建模良好的用户兴趣表示值得进一步探索。

2. 构建信息丰富的推荐数据集。当前学术界和工业界面对推荐系统存在一定差距,他们旨在提升的指标不同,例如学术界重点在于模型设计以此来提升推荐的召回率,工业界更加看重用户的点击率以及转化率。同时使用的数据集也存在一定差异,学术界常用的推荐数据集包含噪声少,但同时包含信息也很少,工业界包含信息众多,但由于涉及用户隐私,很多有价值的真实场景下的推荐数据并没有被很好地加以利用。为了更好地建模用户画像以及克服冷启动问题,构建尽可能内容丰富的推荐数据集能够起到很大的作用。同时,使用尽可能一致的数据集能够一定程度上减少差距,这能够使得预训练技术更快更方便的落地使用,一方面可以帮助用户画像建模,另一方面能够促进预训练技术的发展。

3. 训练推荐领域中的大模型。近期大多预训练技术在推荐系统中的尝试都是围绕一个或者几个推荐任务在单个数据集上进行预训练,这限制了模型的发展和对用户/物品表示的建模。与自然语言领域相比,推荐系统中至今还没有一个能够落地的大型预训练模型在多个数据集上用以同时训练多种任务以此获得通用的用户表示与物品表示。当然,在推荐领域训练大模型存在着很多挑战:首先,不同预训练任务可能有不同的优化目标,如何平衡不同的推荐任务是一个难点;其次,自然语言处理中的语料库虽然巨大,但词表大小

有限,因此学出来的通用语言表示能够合理的用于下游任务中。但与之不同的是,推荐的数据包含多个领域下的海量用户和海量物品,而用户在不同领域下的兴趣分布存在天然的差异,交互数据形式多样且非常稀疏,因此如何对这些数据进行预训练是一个难点;第三,如何合理地在预训练过程中融入丰富的侧信息(如用户/物品的多模态的属性信息、知识图谱、社交关系等)也是一个难点;第四,在对有限的数据集进行预训练之后,需要将该模型合理的用于下游任务。但因为数据集是不可枚举的,而且由于涉及的领域不同,其中包含的特征域也是无穷尽的,因此如何将已经训练好的预训练模型迁移到包含新的特征域的数据集上进行微调也是一大挑战;第五,大规模预训练模型会产生海量参数,这会占据大量内存和计算量,如何减少在低资源设备上进行调整带来的参数量也值得探究。以上这些挑战对于大规模预训练推荐模型的发展值得我们深思和研究。

参考文献

- [1] Liu Jian-Wei, Liu Yuan, Luo Xiong-Lin. Research progress of deep learning. Applications Research of Computers, 2014, 31(7): 1921-1930 (in Chinese) (刘建伟, 刘媛, 罗雄麟. 深度学习研究进展. 计算机应用研究, 2014, 31(7):1921-1930)
- [2] Fang Hui, Zhang Dan-Ning, Shu Yi-Heng, Guo Gui-Bing. 2020. Deep learning for sequential recommendation: Algorithms, influential factors, and evaluations. In ACM Transactions on Information Systems, 2020, 39(1): 1-42.
- [3] Wang Shou-Jin, Cao Long-Bing, Wang Yan, Sheng Quan Z, Orgun Mehmet, Lian De-Fu. A survey on session-based recommender systems. In ACM Computing Surveys, 2021, 54(7): 1-38.
- [4] Zheng Lin, Guo Nai-Cheng, Chen Wei-Hao, Yu Jin, Jiang Da-Zhi. Sentiment-guided sequential recommendation.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Xi'an, China, 2020: 1957-1960.
- [5] Cheng Derek Zhi-Yuan, Chi Ed, Hong Li-Chan, Yao Tian-Sheng, Chen Ting, Kang Wang-Cheng, Yi Xin-Yang. Learning to embed categorical features without embedding tables for recommendation.// Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining, Online, 2021: 840-850.
- [6] Guo Hui-Feng, Chen Bo, Tang Rui-Ming, Zhang Wei-Nan, Li Zhen-Guo, He Xiu-Qiang. An embedding

- learning framework for numerical features in ctr prediction.// Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining, Online, 2021: 2910-2918.
- [7] Wang Peng-Fei, Fan Yu, Xia Long, Zhao Wayne Xin, Niu Shao-Zhang, Huang Jimmy. KERL: A knowledge-guided reinforcement learning model for sequential recommendation.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Xi'an, China, 2020: 209-218.
- [8] Cho Sung Min, E. Park, S. Yoo. MEANTIME: Mixture of attention mechanisms with multi-temporal embeddings for sequential recommendation.// Proceedings of the ACM Conference on Recommender Systems, Online, 2020: 515-520.
- [9] Yuan Fa-Jie, He Xiang-Nan, Jiang Hao-Chuan, Guo Gui-Bing, Xiong Jian, Xu Zhe-Zhao, Xiong Yi-Lin. Future data helps training: Modeling future contexts for session-based recommendation.// Proceedings of the Web Conference, Online, 2020: 303-313.
- [10] Liu Rui-Qian, Ye Xia, Yue Zeng-Ying. A survey of pre-training models for natural language processing tasks. Journal of Computer Applications, 2020 (in Chinese) (刘睿珩, 叶霞, 岳增营. 面向自然语言处理任务的预训练模型综述. 计算机应用. 2020)
- [11] Wang Meng, Fu Wei-Jie, He Xiang-Nan, Hao Shi-Jie, Wu Xin-Dong. A survey on large-scale machine learning. In IEEE Transactions on Knowledge and Data Engineering, 2020, PP(99):1.
- [12] Qiu Xi-Peng, Sun Tian-Xiang, Xu Yi-Ge, Shao Yun-Fan, Dai Ning, Huang Xuan-Jing. Pre-trained models for natural language processing: A survey. Science China Technological Sciences, 2020: 1-26.
- [13] Lan Zhen-Zhong, Chen Ming-Da, Goodman Sebastian, Gimpel Kevin, Sharma Piyush, Soricut Radu. Albert: A lite bert for self-supervised learning of language representations.// Proceedings of the International Conference on Learning Representations, 2020.
- [14] Lewis Mike, Liu Yin-Han, Goyal Na-Man, Ghazvininejad Marjan, Mohamed Abdelrahman, Levy Omer, Stoyanov Ves, Zettlemoyer Luke. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension.// Proceedings of the Annual Meeting of the Association for Computational Linguistics, Online, 2020: 7871-7880.
- [15] Yue Zeng-Ying, Ye Xia, Liu Rui-Qian. A survey of pre-training technology based on language model. Journal of Chinese Information Processing, 2021, 35(9): 15-29 (in Chinese) (岳增营, 叶霞, 刘睿珩. 基于语言模型的预训练技术研究综述. 中文信息学报, 2021, 35(9): 15-29)
- [16] Han Xu, Zhang Zheng-Yan, Ding Ning, Gu Yu-Xian, Liu Xiao, Huo Yu-Qi, Qiu Jie-Zhong, Zhang Liang, Han Wen-Tao, Huang Min-Lie and others. Pre-trained models: Past, present and future. AI Open, 2021.
- [17] Devlin Jacob, Chang Ming-Wei, Lee Kenton, Toutanova Kristina. Bert: Pre-training of deep bidirectional transformers for language understanding.// Proceedings of the North American Chapter of the Association for Computational Linguistics, New Orleans, US, 2018: 4171-4186.
- [18] Gillioz Anthony, Casas Jacky, Mugellini Elena, Abou Khaled Omar. Overview of the transformer-based models for nlp tasks.// Proceedings of the Federated Conference on Computer Science and Information Systems, Sofia, Bulgaria, 2020: 179-183.
- [19] Zhang Jing-Qing, Zhao Yao, Saleh Mohammad, Liu Peter. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization.// Proceedings of the International Conference on Machine Learning, Online, 2020: 11328-11339.
- [20] Liu Yin-Han, Gu Jia-Tao, Goyal Naman, Li Xian, Edunov Sergey, Ghazvininejad Marjan, Lewis Mike, Zettlemoyer Luke. Multilingual denoising pre-training for neural machine translation. In Transactions of the Association for Computational Linguistics, 2020(8): 726-742.
- [21] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need.// Proceedings of the Neural Information Processing Systems, California, US, 2017: 5998-6008.
- [22] Sun Pei-Jie, Wu Le, Zhang Kun, Fu Yan-Jie, Wang Meng. Dual learning for explainable recommendation: Towards unifying user preference prediction and review generation.// Proceedings of the Web Conference, Online, 2020: 837-847.
- [23] Li Chen-Liang, Niu Xi-Chuan, Luo Xiang-Yang, Chen Zhen-Zhong, Quan Cong. A review-driven neural model for sequential recommendation.// Proceedings of the International Joint Conference on Artificial Intelligence, Macao, China, 2019: 2866-2872.
- [24] Chen Xu-Song, Liu Dong, Lei Chen-Yi, Li Rui, Zha Zheng-Jun, Xiong Zhi-Wei. Bert4sessrec: Content-based video relevance prediction with bidirectional encoder representations from transformer.// Proceedings of the Conference of Multimedia, Nice, France, 2019: 2597-2601.
- [25] Ma Mu-Yang, Ren Peng-Jie, Chen Zhu-Min, Ren Zhao-Chun, Liang Hua-Sheng, Ma Jun, de Rijke Maarten. Improving transformer-based sequential recommenders through preference editing. arXiv preprint

- arXiv:2106.12120, 2021.
- [26] Guo Qing-Yu, Zhuang Fu-Zhen, Qin Chuan, Zhu Heng-Shu, Xie Xing, Xiong Hui, He Qing. A survey on knowledge graph-based recommender systems. In *Scientia Sinica Informationis*, 2020, 50(7):937.
- [27] Lake Thom, Williamson Sinead A, Hawk, Alexander T, Johnson Christopher C, Wing Benjamin P. Large-scale collaborative filtering with product embeddings. arXiv preprint arXiv:1901.04321, 2019.
- [28] Shin Kyuyong, Kwak Hanock, Kim Kyung-Min, Kim Minkyu, Park Young-Jin, Jeong Jisu, Jung Seungjae. One4all user representation for recommender systems in e-commerce. arXiv preprint arXiv:2106.00573, 2021.
- [29] Zeng Zhe-Ni, Xiao Chao-Jun, Yao Yuan, Xie Ruo-Bing, Liu Zhi-Yuan, Lin Fen, Lin Le-Yu, Sun Mao-Song. Knowledge transfer via pre-training for recommendation: A review and prospect. *Frontiers in big Data*, 2021(4).
- [30] de Souza Pereira Moreira Gabriel, Rabhi Sara, Lee Jeong Min, Ak Ronay, Oldridge Even. Transformers4Rec: Bridging the gap between nlp and sequential/session-based recommendation.// *Proceedings of the ACM Conference on Recommender Systems*, Online, 2021: 143-153.
- [31] Qin Chuan, Zhu Heng-Shu, Zhuang Fu-Zhen, Guo Qing-Yu, Zhang Qi, Zhang Le, Wang Chao, Chen En-Hong, Xiong Hui. A survey on knowledge graph-based recommender systems. *Scientia Sinica Informationis*, 2020, 50(7): 937-956.
- [32] Guo Lei, Wen Yu-Fei, Wang Xin-Hua. Exploiting pre-trained network embeddings for recommendations in social networks. *Journal of Computer Science and Technology*, 2018, 33(4): 682-696.
- [33] Wang Hui, Zhou Kun, Zhao Wayne Xin, Wang Jing-Yuan, Wen Ji-Rong. Curriculum pre-training heterogeneous subgraph transformer for top-N recommendation. arXiv preprint arXiv:2106.06722, 2021.
- [34] Yuan Xu, Chen Hong-Shen, Song Yong-Hao, Zhao Xiao-Fang, Ding Zhuo-Ye, He Zhen, Long Bo. Improving sequential recommendation consistency with self-supervised imitation.// *Proceedings of the International Joint Conference on Artificial Intelligence*, Online, 2021: 3321-3327.
- [35] Zhou Xin, Sun Ai-Xin, Liu Yong, Zhang Jie, Miao Chun-Yan. SelfCF: A simple framework for self-supervised collaborative filtering. arXiv preprint arXiv:2107.03019, 2021.
- [36] Huang Jin, Zhao Wayne Xin, Dou Hong-Jian, Wen Ji-Rong, Chang Edward Y. Improving sequential recommendation with knowledge-enhanced memory networks.// *Proceedings of the ACM International Conference on Research and Development in Information Retrieval*. Ann Arbor Michigan, U.S, 2018: 505-514.
- [37] Wang Hong-Wei, Zhang Fu-Zheng, Hou Min, Xie Xing, Guo Min-Yi, Liu Qi. Shine: Signed heterogeneous information network embedding for sentiment link prediction.// *Proceedings of the ACM International Conference on Web Search and Data Mining*, California, US, 2018: 592-600.
- [38] Cao Yi-Xin, Wang Xiang, He Xiang-Nan, Hu Zi-Kun, Chua Tat-Seng. Unifying knowledge graph learning and recommendation: Towards a better understanding of user preferences.// *Proceedings of the Web Conference*, San Francisco, US, 2019: 151-161.
- [39] Zheng Lei, Noroozi Vahid, Yu Philip S. Joint deep modeling of users and items using reviews for recommendation.// *Proceedings of the ACM International Conference on Web Search and Data Mining*, Cambridge, UK, 2017: 425-434.
- [40] Natarajan Senthilselvan, Vairavasundaram Subramaniyaswamy, Natarajan Sivaramakrishnan, Gandomi Amir H. Resolving data sparsity and cold start problem in collaborative filtering recommender system using linked open data. *Expert Systems with Applications*, 2020, 149: 113248.
- [41] Xin Xin, Karatzoglou Alexandros, Arapakis Ioannis et al. Self-supervised reinforcement learning for recommender systems.// *Proceedings of the ACM International Conference on Research and Development in Information Retrieval*. Xi'an, China, 2020: 931-940.
- [42] Zhou Kun, Wang Hui, Zhao Wayne Xin, Zhu Yu-Tao, Wang Si-Rui, Zhang Fu-Zheng, Wang Zhong-Yuan, Wen Ji-Rong. S3-rec: Self-supervised learning for sequential recommendation with mutual information maximization.// *Proceedings of the ACM International Conference on Information and Knowledge Management*, Online, 2020: 1893-1902.
- [43] Qiu Rui-Hong, Huang Zi, Yin Hong-Zhi, Wang Zi-Jian. Contrastive learning for representation degeneration problem in sequential recommendation.// *Proceedings of the ACM International Conference on Web Search and Data Mining*, Online, 2022: 813-823.
- [44] Liu Zhuang, Ma Yun-Pu, Ouyang Yuan-Xin, Xiong Zhang. Contrastive learning for recommender system. arXiv preprint arXiv:2101.01317, 2021.
- [45] Liu Zhi-Wei, Chen Yong-Jun, Li Jia, Yu Philip S, McAuley Julian, Xiong Caiming. Contrastive self-

- supervised sequential recommendation with robust augmentation. arXiv preprint arXiv:2108.06479, 2021.
- [46] Zhang Jun-Wei, Gao Min, Yu Jun-Liang, Guo Lei, Li Jun-Dong, Yin Hong-Zhi. Double-scale self-supervised hypergraph learning for group recommendation.// Proceedings of the ACM International Conference on Information and Knowledge Management, Online, 2021: 2557-2567.
- [47] Wei Yin-Wei, Wang Xiang, Li Qi, Nie Li-Qiang, Li Yan, Li Xuan-Ping, Chua Tat-Seng. Contrastive learning for cold-start recommendation.// Proceedings of the Conference of Multimedia, Online, 2021: 5382-5390.
- [48] Liu Xiao, Zhang Fan-Jin, Hou Zhen-Yu, Li Mian, Wang Zhao-Yu, Zhang Jing, Tang Jie. Self-supervised learning: Generative or contrastive. In IEEE Transactions on Knowledge and Data Engineering, 2021.
- [49] Kang Wang-Cheng, McAuley Julian. Self-attentive sequential recommendation.// Proceedings of the IEEE International Conference on Data Mining, Singapore, 2018: 197-206.
- [50] Sun Fei, Liu Jun, Wu Jian, Pei Chang-Hua, Lin Xiao, Ou Wen-Wu, Jiang Peng. Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer.// Proceedings of the ACM International Conference on Information and Knowledge Management, Beijing, China, 2019: 1441-1450.
- [51] Liu Zhi-Wei, Fan Zi-Wei, Wang Yu, Yu Philip S. Augmenting sequential recommendation with pseudo-prior items via reversely pre-training transformer.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval, Online, 2021: 1608-1612.
- [52] Xie Xu, Sun Fei, Liu Zhao-Yang, Wu Shi-Wen, Gao Jin-Yang, Ding Bo-Lin, Cui Bin. Contrastive learning for sequential recommendation. arXiv preprint arXiv:2010.14395, 2020.
- [53] Wu Jian-Can, Wang Xiang, Feng Fu-Li, He Xiang-Nan, Chen Liang, Lian Jian-Xun, Xie Xing. Self-supervised graph learning for recommendation.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Xi'an, China, Online, 2021: 726-735.
- [54] Xia Xin, Yin Hong-Zhi, Yu Jun-Liang, Wang Qin-Yong, Cui Li-Zhen, Zhang Xiang-Liang. Self-supervised hypergraph convolutional networks for session-based recommendation.// Proceedings of the National Conference on Artificial Intelligence, New York, US, 2020: 4503-4511.
- [55] Wang Guo-Xia, Liu He-Ping. Survey of personalized recommendation system. Computer Engineering and Applications, 2012, 48(7) (in Chinese)
(王国霞, 刘贺平. 个性化推荐系统综述. 计算机工程与应用, 2012, 48(7))
- [56] Liu Jian-Guo, Zhou Tao, Wang Bing-Hong. Survey of personalized recommendation system. Progress in Natural Science, 2009, 19(1):15 (in Chinese)
(刘建国, 周涛, 汪秉宏. 个性化推荐系统的研究进展. 自然科学 进, 2009, 19(1):15)
- [57] Li Xing-Xing, Huang Xiao-Qin, Zhu Qing-Sheng. Research on e-commerce recommendation system. Computer Engineering and Science, 2004, 26(5): 7-10 (in Chinese)
(黎星星, 黄小琴, 朱庆生. 电子商务推荐系统研究. 计算机工程 与科学, 2004, 26(5): 7-10)
- [58] Bock Joel R, Maewal Akhilesh. Adversarial learning for product recommendation. In AI, 2020, 1(3): 376-388.
- [59] Meng Xiang-Wu, Hu Xun, Wang Li-Cai, Zhang Yu-Jie. Mobile recommender systems and their applications. Journal of Software, 2013, 24(1):18 (in Chinese)
(孟祥武, 胡勋, 王立才, 张玉洁. 移动推荐系统及其应用. 软件 学报, 2013, 24(1):18)
- [60] Xiang Liang. Recommended system practice, The People's Posts and Telecommunications Press, 2012 (in Chinese)
(项亮. 推荐系统实战, 人民邮电出版社, 2012)
- [61] Koren Y. The bellkor solution to the netflix grand prize. Netflix prize documentation, 2009, 81(2009): 1-10.
- [62] Töschler, Andreas, Michael Jahrer, Robert M. Bell. The bigchaos solution to the netflix grand prize. Netflix prize documentation (2009): 1-52.
- [63] Potté Martin, Martin Chabbert. The pragmatic theory solution to the netflix grand prize. Netflix prize documentation (2009).
- [64] Çano E, Morisio M. Hybrid recommender systems: A systematic literature review. Intelligent Data Analysis, 2017, 21(6): 1487-1524.
- [65] Burke R. Hybrid recommender systems: Survey and experiments. User modeling and user-adapted interaction, 2002, 12(4): 331-370.
- [66] Tang Jia-Xi, Wang Ke. Personalized top-n sequential recommendation via convolutional sequence embedding.// Proceedings of the ACM International Conference on Web Search and Data Mining, California, US, 2018: 565-573.
- [67] Li Jing, Ren Peng-Jie, Chen Zhu-Min, Ren Zhao-Chun, Lian Tao, Ma Jun. Neural attentive session-based recommendation.// Proceedings of the ACM International Conference on Information and Knowledge Management, Singapore, 2017: 1419-1428.

- [68] Seo Sungyong, Huang Jing, Yang Hao, Liu Yan. Interpretable convolutional neural networks with dual local and global attention for review rating prediction.// Proceedings of the ACM Conference on Recommender Systems, Como, Italy, 2017: 297-305.
- [69] Huang Chun-Li, Jiang Wen-Jun, Wu Jie, Wang Guo-Jun. Personalized review recommendation based on users' aspect sentiment. In ACM Transactions on Internet Technology, 2020(20): 1-26.
- [70] Liu Dong-Hua, Li Jing, Du Bo, Chang Jun, Gao Rong. Daml: Dual attention mutual learning between ratings and reviews for item recommendation.// Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining, Anchorage, AK, US, 2019: 344-352.
- [71] Wu Shu, Tang Yu-Yuan, Zhu Yan-Qiao, Wang Liang, Xie Xing, Tan Tie-Niu. Session-based recommendation with graph neural networks.// Proceedings of the National Conference on Artificial Intelligence, Hawaii, US, 2019: 346-353.
- [72] Wang Hong-Wei, Zhang Fu-Zheng, Wang Jia-Lin, Zhao Miao, Li Wen-Jie, Xie Xing, Guo Min-Yi. Ripplenet: Propagating user preferences on the knowledge graph for recommender systems.// Proceedings of the ACM International Conference on Information and Knowledge Management, Turin, Italy, 2018: 417-426.
- [73] Wang Hong-Wei, Zhao Miao, Xie Xing, Li Wen-Jie, Guo Min-Yi. Knowledge graph convolutional networks for recommender systems.// Proceedings of the Web Conference, San Francisco, US, 2019: 3307-3313.
- [74] Huang Jin, Ren Zhao-Chun, Zhao Wayne Xin, He Gao-Le, Wen Ji-Rong, Dong Da-Xiang. Taxonomy-aware multi-hop reasoning networks for sequential recommendation.// Proceedings of the ACM International Conference on Web Search and Data Mining, Melbourne, Australia, 2019: 573-581.
- [75] Hao Bo-Wen, Zhang Jing, Yin Hong-Zhi, Li Cui-Ping, Chen Hong. Pre-training graph neural networks for cold-start users and items representation.// Proceedings of the ACM International Conference on Web Search and Data Mining, Online, 2021: 265-273.
- [76] Fabian Abel, Eelco Herder, Geert-Jan Houben, Nicola Henze, and Daniel Krause. Cross-system user modeling and personalization on the social web.// Proceedings of the Conference on Uncertainty in Artificial Intelligence, Washington, US, 2013: 169-209.
- [77] Herlocker J L, Konstan J A, and Riedl J. Explaining collaborative filtering recommendations.// Proceedings of the ACM conference on Computer supported cooperative work, 2000: 241-250.
- [78] Lara Quijano-Sanchez, Christian Sauer, Juan A. Recio-Garcia and Belen Diaz-Agudo. Make it personal: A social explanation system applied to group recommendations. Expert Systems with Applications, 2017: 36-48.
- [79] Ma Hao, King Irwin, Lyu Michael R. Learning to recommend with social trust ensemble.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Paris, France, 2019: 203-210.
- [80] Ma Hao, King Irwin, Lyu Michael R. Learning to recommend with trust and distrust relationships.// Proceedings of the ACM Conference on Recommender Systems, New York, US, 2009: 189-196.
- [81] Zhang Yong-Feng. Research on the interpretability of personalized recommendation, Tsinghua University Press, 2019 (in Chinese)
(张永锋. 个性化推荐的可解释性研究. 清华大学出版社, 2019.)
- [82] Yu Jun-Liang, Yin Hong-Zhi, Gao Min, Xia Xin, Zhang Xiang-Liang, Hung Nguyen Quoc Viet. Socially-aware self-supervised tri-training for recommendation.// Proceedings of the ACM Conference on Knowledge Discovery and Data Mining, Online, 2021: 2084-2092.
- [83] Yu Jun-Liang, Yin Hong-Zhi, Li Jun-Dong, Wang Qin-Yong, Hung Nguyen Quoc Viet, Zhang Xiang-Liang. Self-supervised multi-channel hypergraph convolutional network for social recommendation.// Proceedings of the Web Conference, Online, 2021: 413-424.
- [84] Wang Xiang, He Xiang-Nan, Nie Li-Qiang, Chua Tat-Seng. Item silk road: Recommending items from information domains to social users.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Tokyo, Japan, 2017: 185-194.
- [85] Song Wei-Ping, Xiao Zhi-Ping, Wang Yi-Fan, Charlin Laurent, Zhang Ming, Tang Jian. Session-based social recommendation via dynamic graph attention networks.// Proceedings of the ACM International Conference on Web Search and Data Mining, Melbourne, Australia, 2019: 555-563.
- [86] Chen Tian-Wen, Wong Raymond Chi-Wing. An efficient and effective framework for session-based social recommendation.// Proceedings of the ACM International Conference on Web Search and Data Mining, Online, 2021: 400-408.
- [87] Pan Zhi-Qiang, Cai Fei, Ling Yan-Xiang, de Rijke Maarten. Rethinking item importance in session-based recommendation.// Proceedings of ACM International

- Conference on Research and Development in Information Retrieval. Xi'an, China, 2020: 1837-1840.
- [88] Andrew Zimdars, David Maxwell Chickering, and Christopher Meek. Using temporal data for making recommendations.// *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, Washington, US, 2001: 580-588.
- [89] Guy Shani, David Heckerman, and Ronen I. Brafman. An mdp-based recommender system. *Journal of Machine Learning Research*, 2005(6): 1265-1295.
- [90] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, Domonkos Tikk. Session-based recommendations with recurrent neural networks.// *Proceedings of the International Conference on Learning Representations*, 2015.
- [91] Xu Cheng-Feng, Zhao Peng-Peng, Liu Yan-Chi, Sheng Victor S, Xu Jia-Jie, Zhuang Fu-Zhen, Fang Jun-Hua, Zhou Xiao-Fang. Graph contextualized self-attention network for session-based recommendation.// *Proceedings of the International Joint Conference on Artificial Intelligence*, Macao, China, 2019: 3940-3946.
- [92] Liu Qiang, Wu Shu, Wang Di-Yi, Li Zhao-Kang, Wang Liang. Context-aware sequential recommendation.// *Proceedings of the IEEE International Conference on Data Mining*, Barcelona, Spain, 2016: 1053-1058.
- [93] Massimo Quadrana, Alexandros Karatzoglou, Balzs Hidasi, and Paolo Cremonesi. Personalizing session-based recommendations with hierarchical recurrent neural networks.// *Proceedings of the ACM Conference on Recommender Systems*, Como, Italy, 2017: 130-137.
- [94] Balázs Hidasi, Massimo Quadrana, Alexandros Karatzoglou, and Domonkos Tikk. Parallel recurrent neural network architectures for feature-rich session-based recommendations.// *Proceedings of the ACM Conference on Recommender Systems*, Boston, US, 2016: 241-248.
- [95] Veronika Bogina, Tsvi Kuflik. Incorporating dwell time in session-based recommendations with recurrent Neural networks.// *Proceedings of the ACM Conference on Recommender Systems*, Como, Italy, 2017: 57-9.
- [96] Sun Shi-Ming, Tang Yuan-He, Dai Ze-Mei, Zhou Fu. Self-attention network for session-based recommendation with streaming data input. In *IEEE Access*, 2019(7): 110499-110509.
- [97] Zhai Xiao-Hua, Oliver Avital, Kolesnikov Alexander, Beyer Lucas. S4I: Self-supervised semi-supervised learning.// *Proceedings of the IEEE International Conference on Computer Vision*, Seoul, South Kerean, 2019: 1476-1485.
- [98] Jing Long-Long, Tian Ying-Li. Self-supervised visual feature learning with deep neural networks: A survey. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 43(11): 4037-4058.
- [99] Chen Ting, Kornblith Simon, Norouzi Mohammad, Hinton Geoffrey. A simple framework for contrastive learning of visual representations.// *Proceedings of the Machine Learning Research*, Online, 2020: 1597-1607.
- [100] Jaiswal Ashish, Babu Ashwin Ramesh, Zadeh Mohammad Zaki, Banerjee Debapriya, Makedon Filia. A survey on contrastive self-supervised learning. In *Technologies*, 2021, 9(1): 2.
- [101] Ma Jian-Xin, Zhou Chang, Cui Peng, Yang Hong-Xia, Zhu Wen-Wu. Learning disentangled representations for recommendation.// *Proceedings of the Neural Information Processing Systems*, Vancouver, Canada, 2019: 32.
- [102] Chen Xin-Lei, Fan Hao-Qi, Girshick Ross, He Kai-Ming. Improved baselines with momentum contrastive learning. *arXiv preprint arXiv:2003.04297*.
- [103] Oord Aaron van den, Li Ya-Zhe, Vinyals Oriol. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*.
- [104] Doersch Carl, Gupta Abhinav, Efros Alexei A. Unsupervised visual representation learning by context prediction.// *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 2015: 1422-1430.
- [105] Wei Chen, Xie Ling-Xi, Ren Xu-Tong, Xia Ying-Da, Su Chi, Liu Jia-Ying, Tian Qi, Yuille Alan L. Iterative reorganization with weak spatial constraints: Solving arbitrary jigsaw puzzles for unsupervised representation learning.// *Proceedings of the Conference on Computer Vision and Pattern Recognition*, Long Beach, US, 2019: 1910-1919.
- [106] Kong Ling-Peng, d'Autume Cyprien de Masson, Ling Wang, Yu Lei, Dai Zi-Hang, Yogatama Dani. A mutual information maximization perspective of language representation learning.// *Proceedings of the International Conference on Learning Representations*, 2020.
- [107] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets.// *Proceedings of the Neural Information Processing Systems*, Montreal, Canada, 2014: 2672-2680.
- [108] Wang Hongwei, Wang Jia, Wang Jia-Lin, Zhao Miao, Zhang Wei-Nan, Zhang Fu-Zheng, Xie Xing, Guo Min-Yi. GraphGAN: Graph representation learning with generative adversarial nets.// *Proceedings of the National Conference on Artificial Intelligence*, New

- Orleans, US, 2018: 32(1).
- [109] Wang Jun, Yu Lan-Tao, Zhang Wei-Nan, Gong Yu, Xu Ying-Hui, Wang Ben-You, Zhang Peng, Zhang Dell. IRGAN: A minimax game for unifying generative and discriminative information retrieval models.// *Proceedings of the ACM International Conference on Research and Development in Information Retrieval*. Tokyo, Japan. 2017: 515–524.
- [110] Sun Zhong-Chuan, Wu Bin, Wu Yun-Peng, Ye Yang-Dong. APL: Adversarial pairwise learning for recommender systems. *Expert Systems with Applications*, 2019(118): 573–584.
- [111] Liu Ming-Yu, Oncel Tuzel. Coupled generative adversarial networks.// *Proceedings of the Neural Information Processing Systems*, Barcelona, Spain, 2016: 29.
- [112] Karras Tero, Aila Timo, Laine Samuli, Lehtinen Jaakko. Progressive growing of gans for improved quality, stability, and variation.// *Proceedings of the International Conference on Learning Representations*, 2018.
- [113] Zhang Han, Goodfellow Ian, Metaxas Dimitris, Odena Augustus. Self-attention generative adversarial networks.// *Proceedings of the Machine Learning Research*, Nagoya, Japan, 2019: 7354–7363.
- [114] Zhu Jun-Yan, Park Taesung, Isola Phillip, Efros Alexei A. Unpaired image-to-image translation using cycle-consistent adversarial networks.// *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 2017: 2223–2232.
- [115] Brock A, Donahue J, Simonyan K. Large scale gan training for high fidelity natural image synthesis.// *Proceedings of the International Conference on Learning Representations*, 2019.
- [116] Ma Li-Qian, Jia Xu, Sun Qian-Ru, Schiele Bernt, Tuytelaars Tinne, Gool Luc Van. Pose guided person image generation.// *Proceedings of the Neural Information Processing Systems*, California, US, 2017: 30.
- [117] Zhang Han, Xu Tao, Li Hong-Shen, Zhang Shao-Ting, Wang Xiao-Gang, Huang Xiao-Bei, Metaxas Dimitris. StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks.// *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 2017: 5907–5915.
- [118] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks.// *Proceedings of the International Conference on Learning Representations*, 2016.
- [119] Mirza M, Osindero S. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.
- [120] Makhzani A, Shlens J, Jaitly N, et al. Adversarial auto-encoders. arXiv preprint arXiv:1511.05644, 2015.
- [121] T.Karras, S. Laine, and T. Aila. A style-based generator architecture for generative adversarial networks.// *Proceedings of the Conference on Computer Vision and Pattern Recognition*, Long Beach, US, 2019: 4401–4410.
- [122] K. Clark, M.-T. Luong, Q. V. Le, and C. D. Manning. Electra: Pre-training text encoders as discriminators rather than generators.// *Proceedings of the International Conference on Learning Representations*, 2020.
- [123] Chae D K, Kang J S, Kim S W, et al. Cfgan: A generic collaborative filtering framework based on generative adversarial networks.// *Proceedings of the ACM International Conference on Information and Knowledge Management*, Turin, Italy, 2018: 137–146.
- [124] Yin Rui-Ping, Li Kan, Lu Jie, Zhang Guang-Quan. Rsy-GAN: Generative adversarial network for recommender systems.// *Proceedings of the International Joint Conference on Neural Networks*, Budapest, Hungary, 2019: 1–7.
- [125] Wang Cheng, Niepert Mathias, Li Hui. Recsys-dan: discriminative adversarial networks for cross-domain recommender systems. In *IEEE transactions on neural networks and learning systems*, 2019, 31(8): 2731–2740.
- [126] Yuan Fa-Jie, Zhang Guo-Xiao, Karatzoglou Alexandros, Jose Joemon, Kong Bei-Bei, Li Yu-Dong. One person, one model, one world: Learning continual user representation without forgetting.// *Proceedings of the ACM International Conference on Research and Development in Information Retrieval*. Online, 2021: 696–705.
- [127] Liu Yong, Yang Su-Sen, Lei Chen-Yi, Wang Guo-Xin, Tang Hai-Hong, Zhang Ju-Yong, Sun Ai-Xin, Miao Chun-Yan. Pre-training graph transformer with multi-modal side information for recommendation.// *Proceedings of the Conference of Multimedia*, Online, 2021: 2853–2861.
- [128] Xiao Chao-Jun, Xie Ruo-Bing, Yao Yuan, Liu Zhi-Yuan, Sun Mao-Song, Zhang Xu, Lin Le-Yu. UPRec: User-aware pre-training for recommender systems. arXiv preprint arXiv:2102.10989, 2021.
- [129] Qiu Zhao-Peng, Wu Xian, Gao Jing-Yue, Fan Wei. U-BERT: Pre-training user representations for improved recommendation.// *Proceedings of the National Conference on Artificial Intelligence*, Online, 2021: 4320–4327.
- [130] Yang Jing-Xuan, Xu Jun, Tong Jian-Zhuo, Gao Sheng,

- Guo Jun, Wen Ji-Rong. Pre-training of context-aware item representation for next basket recommendation. arXiv preprint arXiv:1904.12604, 2019.
- [131] Yuan Fa-Jie, He Xiang-Nan, Karatzoglou Alexandros, Zhang Li-Guang. Parameter-efficient transfer from sequential behaviors for user modeling and recommendation. // Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Xi'an, China, 2020: 1469-1478.
- [132] Yao Tian-Sheng, Yi Xin-Yang, Cheng Derek Zhi-Yuan, Yu Felix, Chen Ting, Menon Aditya, Hong Li-Chan, Chi Ed H, Tjoa Steve, Kang Jie-Qi and others. Self-supervised learning for large-scale item recommendations.// Proceedings of the ACM International Conference on Information and Knowledge Management, Online, 2020: 4321-4330.
- [133] Zhou Chang, Ma Jian-Xin, Zhang Jian-Wei, Zhou Jing-Ren, Yang Hong-Xia. Contrastive learning for debiased candidate generation in large-scale recommender systems.// Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining, Online, 2021: 3985-3995.
- [134] Ma Jian-Xin, Zhou Chang, Yang Hong-Xia, Cui Peng, Wang Xin, Zhu Wen-Wu. Disentangled self-supervision in sequential recommenders.// Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining, Online, 2020: 183-491.
- [135] Zhang Qi, Li Jing-Jie, Jia Qing-Lin, Wang Chu-Yuan, Zhu Jie-Ming, Wang Zhao-Wei, He Xiu-Qiang. Unbert: User-news matching bert for news recommendation.// Proceedings of the International Joint Conference on Artificial Intelligence, Online, 2021: 3356-3362.
- [136] Shang Jun-Yuan, Ma Teng-Fei, Xiao Cao, Sun Ji-Meng. Pre-training of graph augmented transformers for medication recommendation.// Proceedings of the ACM International Joint Conference on Artificial Intelligence, Macao, China, 2019: 5953-5959.
- [137] Jiang Ju-Yong, Luo Ying-Tao, Kim Jae Boum, Zhang Kai, Kim Sunghun. Sequential recommendation with bidirectional chronological augmentation of transformer. arXiv preprint arXiv:2112.06460, 2021.
- [138] Wu Chu-Han, Wu Fang-Zhao, Qi Tao, Lian Jian-Xun, Huang Yong-Feng, Xie Xing. PTUM: Pre-training user model from unlabeled user behaviors via self-supervision.// Proceedings of the Conference on Empirical Methods in Natural Language Processing, Online, 2020.
- [139] Wang Hui, Zhou Kun, Zhao Wayne Xin, Wang Jing-Yuan, Wen Ji-Rong. Curriculum pre-training heterogeneous subgraph transformer for top-n recommendation. arXiv preprint arXiv:2106.06722, 2021.
- [140] Lee Dongha, Kang SeongKu, Ju Hyunjun, Park Chanyoung, Yu Hwanjo. Bootstrapping user and item representations for one-class collaborative filtering.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval, Online, 2021: 1513-1522.
- [141] Cheng Ming-Yue, Yuan Fa-Jie, Liu Qi, Xin Xin, Chen En-Hong. Learning transferable user representations with sequential behaviors via contrastive pre-training.// Proceedings of the IEEE International Conference on Data Mining, Online, 2021: 51-60.
- [142] Liu Hao-Chen, Tang Da, Yang Ji, Zhao Xiang-Yu, Tang Ji-Liang, Cheng You-Long. Self-supervised learning for alleviating selection bias in recommendation systems. 2021.
- [143] Tao Ying-Hui, Gao Min, Yu Jun-Liang, Wang Zong-Wei, Xiong Qing-Yu, Wang Xu. Predictive and contrastive: Dual-auxiliary learning for recommendation. arXiv preprint arXiv:2203.03982, 2022.
- [144] Bian Shu-Qing, Zhao Wayne Xin, Zhou Kun, Cai Jing, He Yan-Cheng, Yin Cun-Xiang, Wen Ji-Rong. Contrastive curriculum learning for sequential user behavior modeling via data augmentation.// Proceedings of the ACM International Conference on Information and Knowledge Management, Online, 2021: 3737-3746.
- [145] Xia Xin, Yin Hong-Zhi, Yu Jun-Liang, Shao Ying-Xia, Cui Li-Zhen. Self-supervised graph co-training for session-based recommendation.// Proceedings of the ACM International Conference on Information and Knowledge Management, Online, 2021: 2180-2190.
- [146] Hao Bo-Wen, Yin Hong-Zhi, Zhang Jing, Li Cui-Ping, Chen Hong. A multi-strategy based pre-training method for cold-start recommendation. arXiv preprint arXiv:2112.02275, 2021.
- [147] Yang Yong-Hui, Wu Le, Hong Ri-Chang, Zhang Kun, Wang Meng. Enhanced graph learning for collaborative filtering via mutual information maximization.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Online, 2021: 71-80.
- [148] Li Yi-Cong, Chen Hong-Xu, Sun Xiang-Guo, Sun Zhen-Chao, Li Lin, Cui Li-Zhen, Yu Philip S, Xu Guan-Dong. Hyperbolic hypergraphs for sequential recommendation.// Proceedings of the ACM International Conference on Information and Knowledge Management, Online, 2021: 988-997.
- [149] C. Wang, W. Ma, and C. Chen. Sequential recommendation with multiple contrast signals. In ACM Transactions on Information Systems, 2022.

- [150]Cao Jiang-Xia, Lin Xi-Xun, Guo Shu, Liu Lu-Chen, Liu Ting-Wen, Wang Bin. Bipartite graph embedding via mutual information maximization.// Proceedings of the ACM International Conference on Web Search and Data Mining, Online, 2021: 635-643.
- [151]Cai De-Sheng, Qian Sheng-Sheng, Fang Quan, Xu Chang-Sheng. Heterogeneous graph contrastive learning network for personalized micro-video recommendation. In IEEE Transactions on Multimedia, 2022.
- [152]Long Xiao-Ling, Huang Chao, Xu Yong, Xu Huan-Ce, Dai Peng, Xia Liang-Hao, Bo Lie-Feng. Social recommendation with self-supervised metagraph informax network.// Proceedings of the ACM International Conference on Information and Knowledge Management, Online, 2021: 1160-1169.
- [153]Xie Ruo-Bing, Liu Qi, Wang Liang-Dong, Liu Shu-Kai, Zhang Bo, Lin Le-Yu. Contrastive cross-domain recommendation in matching. arXiv reprint arXiv:2112.00999, 2021.
- [154]Wang Chen, Liang Yue-Qing, Liu Zhi-Wei, Zhang Tao, Yu Philip S. Pre-training graph neural network for cross domain recommendation.// Proceedings of the IEEE International Conference on Cognitive Machine Intelligence, Atlanta, US, 2021: 140-145.
- [155]Lin Zi-Han, Tian Chang-Xin, Hou Yu-Peng, Zhao Wayne Xin. Improving graph collaborative filtering with neighborhood-enriched contrastive learning.// Proceedings of the Web Conference, 2022.
- [156]Chen Yong-Jun, Liu Zhi-Wei, Li Jia, McAuley Julian, Xiong Cai-Ming. Intent contrastive learning for sequential recommendation. arXiv preprint arXiv:2202.02519, 2022.
- [157]Guo Wei, Zhang Can, He Zhi-Cheng, Qin Jia-Rui, Guo Hui-Feng, Chen Bo, Tang Rui-Ming, He Xiu-Qiang, Zhang Rui. Miss: Multi-interest self-supervised learning framework for click-through rate prediction.// Proceedings of the IEEE International Conference on Data Engineering, 2022.
- [158]Yu Jun-Liang, Yin Hong-Zhi, Xia Xin, Chen Tong, Cui Li-Zhen, Nguyen Quoc Viet Hung. Graph augmentation-free contrastive learning for recommendation. arXiv preprint arXiv:2112.08679, 2021.
- [159]Liu Zhi-Wei, Chen Yong-Jun, Li Jia, Luo Man, Yu Philip S. Yu, Xiong Cai-Ming. Improving contrastive learning with model augmentation. arXiv preprint arXiv: 2203.15508, 2021.
- [160]Zhao Peng-Yu, Shui Tian-Xiao, Zhang Yuan-Xing, Xiao Ke-Cheng, Bian Kai-Gui. Adversarial oracular seq2seq learning for sequential recommendation.// Proceedings of the International Joint Conference on Artificial Intelligence, Online, 2021: 1905-1911.
- [161]Ren Rui-Yang, Liu Zhao-Yang, Li Ya-Liang, Zhao Wayne Xin, Wang Hui, Ding Bo-Lin, Wen Ji-Rong. Sequential recommendation with self-attentive multi-adversarial network.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Xi'an, China, 2020: 89-98.
- [162]Wu Qiong, Liu Yong, Miao Chun-Yan, Zhao Bin-Qiang, Zhao Yin, Guan Lu. PD-GAN: Adversarial learning for personalized diversity-promoting recommendation.// Proceedings of the International Joint Conference on Artificial Intelligence, Macao, China, 2019: 3870-3876.
- [163]Chen Xin-Shi, Li Shuang, Li Hui, Jiang Shao-Hua, Qi Yuan, Song Le. Generative adversarial user model for reinforcement learning based recommendation system.// Proceedings of the International Conference on Machine Learning, California, US, 2019: 1052-1061.
- [164]Bharadhwaj, Homanga and Park, Homin and Lim, Brian Y. RecGAN: Recurrent generative adversarial networks for recommendation systems.// Proceedings of the ACM Conference on Recommender Systems, Vancouver, Canada, 2018: 372-376.
- [165]McAuley Julian, Targett Christopher, Shi Qin-Feng, Hengel Anton van den. Image-based recommendations on styles and substitutes.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Santiago, Chile, 2015: 43-52.
- [166]Zhang Ting-Ting, Zhao Peng-Peng, Liu Yan-Chi, Sheng Victor S., Zhou Xiao-Fang. Feature-level deeper self-attention network for sequential recommendation.// Proceedings of the International Joint Conference on Artificial Intelligence, Macao, China, 2019: 4320-4326.
- [167]Meng Wen-Jing, Yang De-Qing, Xiao Yang-Hua. Incorporating user micro-behaviors and item knowledge into multi-task learning for session-based recommendation.// Proceedings of the ACM International Conference on Research and Development in Information Retrieval. Xi'an, China, 2020: 1091-1100.
- [168]Zeng Zhe-ni, Xiao Chao-jun, Yao Yuan, Xie Ruo-bing, Liu Zhi-yuan, Lin Fen, Lin Le-yu, Sun Mao-song. Knowledge transfer via pre-training for recommendation: A review and prospect. Frontiers in big Data, vol. 4, 2021.
- [169]Yu Jun-liang; Yin, Hong-zhi; Xia Xin; Chen Tong; Li Jun-dong; Huang Zi. Self-Supervised Learning for Recommender Systems: A Survey[J]. arXiv preprint

arXiv:2203.15876, 2022.



杨纪元(1999—), 硕士研究生, 主要研究领域为推荐系统。
E-mail: jiyuan.yang@mail.sdu.edu.cn



马沐阳(1995—), 共同第一作者, 博士, 主要研究领域为推荐系统。
E-mail: muyang0331@gmail.com



任鹏杰(1990—), 通信作者, 博士, 教授, 博士生导师, 主要研究领域为自然语言处理和信息检索。
E-mail: renpengjie@sdu.edu.cn