

知识驱动的推荐系统：现状与展望

阳德青^{1,3}, 夏西¹, 叶琳¹, 薛吕欣¹, 肖仰华^{2,3}

¹ 复旦大学大数据学院 上海 中国 200433

² 复旦大学计算机科学技术学院 上海 中国 200433

³ 复旦爱数认知智能联合研究中心 上海 中国 200433

摘要 个性化推荐系统能够根据用户的个性化偏好和需要, 自动、快速、精准地为用户提供其所需的互联网资源, 已成为当今大数据时代应用最广泛的信息检索系统, 具有巨大的商业应用价值。近年来, 随着互联网海量数据的激增, 人工智能技术的快速发展与普及, 以知识图谱为代表的大数据知识工程日益受到学界和业界的高度关注, 也有力地推动推荐系统和个性化推荐技术也迈入到知识驱动与赋能的发展阶段。将知识图谱中蕴含的丰富知识作为有用的辅助信息引入推荐系统, 不仅能够有效应对数据稀疏、语义失配等传统推荐系统难以避免的问题, 还能帮助推荐系统产生多样化、可解释的推荐结果, 并更好地完成跨领域推荐、序列化推荐等具有挑战性的推荐任务, 从而提升各类实际推荐场景中的用户满意度。本文将现有融入知识图谱的各种推荐模型按其采用的推荐算法与面向的推荐场景不同进行分类, 构建科学、合理的分类体系。其中, 按照推荐方法的不同, 划分出基于特征表示的和基于图结构的两大类推荐模型; 按推荐场景划分, 特别关注多样化推荐、可解释推荐、序列化推荐与跨领域推荐。然后, 我们在各类推荐模型中分别选取代表性的研究工作介绍, 还简要对比了各个模型的特点与优劣。此外, 本文还结合当下人工智能技术和应用的发展趋势, 展望了认知智能推荐系统的发展前景, 具体包括融合多模态知识的推荐系统, 具有常识理解能力的推荐系统, 以及解说式、劝说式、抗辩式推荐系统。本文的综述内容和展望可作为推荐系统未来研究方向的有益参考。

关键词 推荐系统; 知识图谱; 嵌入表示; 深度学习; 认知智能

中图法分类号 TP391.3 DOI 号 10.19363/J.cnki.cn10-1380/tn.2021.09.03

Knowledge-enhanced Recommender Systems: A Survey and Prospect

YANG Deqing^{1,3}, XIA Xi¹, YE Lin¹, XUE Lyuxin¹, XIAO Yanghua^{2,3}

¹ School of Data Science, Fudan University, Shanghai 200433, China

² School of Computer Science, Fudan University, Shanghai 200433, China

³ FUDAN-AISHU Recognitive Intelligence Research Center, Shanghai 200433, China

Abstract Personalized recommender systems can provide users with their expected internet resources automatically, rapidly and precisely, based on users' preferences and demands. As the broadly applied information retrieval systems in the big data era, personalized recommender systems have shown great commercial values in various applications. In recent years, the knowledge engineering of big data which takes knowledge graph (KGs) as representative, has become the hot spot in the communities of both industry and academic, thanks to the sharp increasement of the big data on Web, as well as the rapid development and popularization of artificial intelligence technologies. It also promotes recommender systems to step into the phase of being driven and enabled by knowledge. With the rich knowledge in KGs used as valuable side information, the inevitable problems harming traditional recommender systems, such as data sparsity and semantic mismatch, can be addressed well. Moreover, the knowledge is helpful to generate diverse and explainable recommendation results, as well as achieve cross-domain and sequential recommendation. As a result, user experiences in various real recommendation scenarios are enhanced. In this paper, we classify the existing recommendation models incorporating KGs at first, according to their algorithm design and applied recommendation scenarios (tasks), to construct a scientific and reasonable taxonomy. Specifically, according to recommendation algorithm design, we divide the KG-based recommendation models into the class based on feature representation and the class based on graph topology. For recommendation scenarios, we specially focus on diversified recommendation, explainable recommendation, sequential recommendation and cross-domain recommendation. Next, we introduce some representative works respectively for each category, and briefly compare their characteristics including advantages and disadvantages. Furthermore, inspired by the current trend of artificial intelligence technologies and various applications, we also try to forecast the development trend of recommender sys-

通讯作者: 阳德青, 副教授, Email: yangdeqing@fudan.edu.cn

本课题得到上海市“科技创新行动计划”人工智能科技支撑专项项目(No. 21511100400, No. 19511120400)资助。

收稿日期: 2021-05-07; 修改日期: 2021-08-09; 定稿日期: 2021-08-10

tems with cognitive intelligence. To be specific, these future recommender systems include the systems built with multi-modal knowledge and common sense, as well as the explainable, persuasive and polemical recommender systems. The survey and prospect in this paper can provide constructive preferences for future work in the research field of recommender systems.

Key words recommender system; knowledge graph; embedding; deep learning; cognitive intelligence

1 引言

大数据时代互联网上的海量数据在为网络用户提供丰富信息的同时,也给用户便捷地获取自己所需的信息带来了挑战。而个性化推荐因为能够自动、精准地为用户提供满足其个性化需求的各类信息,目前已经成为绝大多数应用程序与网站标配的信息检索功能。大量商家与平台的成功经验已经充分证明个性化推荐系统能够带来巨大的商机与价值,并且成为人工智能学界和业界共同关注的热点。

1.1 推荐系统的基本原理

个性化推荐系统的任务可以简要描述为:对于给定的目标用户 u , 将最匹配 u (即 u 最可能喜好) 的物品 (也称为项目) 推荐出来。其中, 物品与用户的匹配程度 (即用户对物品的偏好程度) 是推荐算法 (模型) 计算的核心目标, 可用给定 u 后能发现候选物品 i 的条件概率 $P(i|u)$ 来量化。因此, 推荐系统的任务目标可用以下公式来表示^[1]:

$$i_0 = \arg \max_{i \in I} P(i/u) \quad (1)$$

其中的 I 是所有候选物品的集合。在很多推荐模型中, $P(i|u)$ 常用 \hat{y}_{ui} 表示, 说明该条件概率度量的是 u 与 i 的匹配程度 (分值)。公式 1 表明系统要推荐给用户的是匹配概率值最大的物品 i_0 。如果系统要为用户推荐一组物品 (也称为列表式推荐), 则可对所有候选物品按概率值排序, 将排序前 n 位的物品作为输出。传统的推荐算法中, 如基于协同过滤的推荐^[2-3], $P(i|u)$ 的计算依据往往来自用户与物品的历史交互记录 (如评分、点击或购买记录)。用户-物品交互记录往往又可分为显式反馈 (explicit feedback) 和隐式反馈 (implicit feedback) 两类。显式反馈包括用户对物品的具体评分 (评级), 因为能从中明确看出用户对物品的偏好程度。不少推荐算法是对没有观察到交互记录的 “用户-物品” 对预测其可能的匹配分值 r (即用户对物品的评分), 包括很多基于矩阵分解的推荐算法^[3-4], 这种任务往往也称为评分预测。这类推荐算法的目标可视为学习下式中的用户-物品交互函数 f :

$$R = f(U \times I) \quad (2)$$

其中, U 和 I 分别代表系统中所有用户与物品的集合,

R 则是所有用户-物品对的匹配分值集合。

隐式反馈则一般指用户的点击^[4]、浏览等行为。之所以称为隐式反馈是因为, 无法确定用户一定喜欢通过这类行为交互过的物品。例如有可能用户浏览了一个商品页面后并不满意, 不会下单购买。反之, 对于用户没有交互过的物品, 也无法确定用户肯定不喜欢该物品, 因为有可能用户根本就不知道该物品的存在。很多电商网站或者广告系统中关于用户点击率 (CTR) 的预测^[4]即可视为隐式反馈的预测。关于隐式反馈的建模, 一般做法是将观测到的用户 u 与物品 i 的交互记录视为正样本, 即 $y_{ui} = 1$; 而 u 与 i 之间没有观测交互记录则视为负样本, 即 $y_{ui} = 0$ 。一般而言, 对于任何一个用户, 其负样本要远远多于其正样本, 因此很多模型样本集合的构建都采取负采样的手段。即, 按照设定的正负样本数比例, 从用户未交互过的物品集合中随机选取相应数量的物品作为负样本。

1.2 推荐系统面临的挑战

随着个性化推荐的实际场景愈发多样复杂, 用户的要求也日渐提高, 推荐系统面临的挑战越来越多, 其中包括:

1) 数据稀疏性问题。很多推荐系统在数据稀疏的场景下往往难以取得令人满意的推荐效果。例如, 用户与物品的历史交互数据不充足使得诸如协同过滤这类推荐算法^[2-3]无法基于相关交互记录来准确地计算用户与物品的相关性 (匹配值), 因而很难产生精准的推荐结果。这一问题对于新注册的用户或者新上架的商品尤其突出, 关于完全没有任何交互记录的用户/物品的推荐常被称为冷启动 (cold-start) 推荐问题。

2) 语义失配问题。发掘用户与物品之间、用户与用户 (物品与物品) 之间的关联是推荐系统完成推荐的关键步骤。而传统推荐算法由于无法准确捕获实体之间或特征 (如画像标签) 之间的语义匹配或关联使得这一关键步骤难以实现。例如, 基于关联规则的推荐模型^[5]不会为买过 “口红” 的顾客推荐 “高跟鞋”, 因为它很难通过历史的交易记录来发现两者之间的相关性, 造成潜在的商机流失。还有同义词、一词多义等问题造成的语义失配也会影响推荐效

果。例如, 标签“十面埋伏”可以是成语、电影名、小说名, 甚至游戏的名字, 不充分掌握与其相关的语义知识, 则无法准确识别特定场景下该标签的真实指代。

3) 同质化的推荐结果。选择与用户之前交互物品的相似物品来推荐给用户是推荐算法的一个基本原则。虽然这种合理的推荐原则很大程度上能保证推荐的准确性(因为符合用户的偏好), 但是难免产生同质化的推荐结果, 即所谓“信息茧房”效应。例如, 推荐模型会判断一位购买过华为 Mate 40 的用户是高端智能手机的潜在客户, 于是向其推荐 iPhone 12。这一推荐看似准确捕获了客户的偏好, 但他短期内一般不会再买一台新手机, 甚至其有可能实华为的忠实粉丝而坚决不会去买竞争品牌的手机。所以对这位用户而言, 该推荐结果缺乏商业价值, 甚至可能引起用户的反感。

4) 推荐结果缺乏可解释性。协同过滤作为最经典的推荐算法, 其核心思想是基于群体智慧, 但群体的行为模式并不一定适用于个体的行为选择。而基于矩阵分解的隐因子模型^[6-7]中产生的隐因子本身就缺乏可解释性, 每个因子都难以确定是对应于用户偏好或者物品属性中哪一个具体的特征。最近几年出现的深度学习推荐模型尽管效果明显, 但众所周知, 其计算过程缺乏透明性和可解释性是其固有的一大缺陷。

1.3 知识图谱对推荐系统的作用

针对数据稀疏性问题, 不少人士将交互记录以外的辅助信息(side information)引入推荐系统, 来对用户和物品进行更精准、丰富的画像, 从而使得推荐系统对于冷启动用户/物品也能实现精准的推荐。利用到的辅助信息包括用户的注册信息、社交关系^[8-9]、标签^[9-10], 以及商品目录、品牌等属性^[11]。辅助信息尽管能帮助推荐系统, 但要充分地获取往往需要不菲的代价。

近几年, 人们发现知识图谱(knowledge graph, KG)中蕴含的丰富数据, 尤其是关于物品的各类知识, 可作为一类重要的辅助信息引入推荐系统, 以实现物品特征更好地刻画, 从而提升推荐效果。本文将这类结合知识图谱的推荐系统称为**知识驱动的推荐系统**或**知识化推荐系统**。

除了帮助克服数据稀疏性问题, 知识图谱改善推荐系统的作用还体现在以下几方面:

1) 实现精准的语义匹配。基于知识图谱中丰富的实体与关系信息, 或经由简单的知识推理^[12]就能发现很多实体或特征之间的语义匹配和关联, 从而

支持推荐模型产生更加准确的结果。如前述的“口红”与“高跟鞋”的语义关联问题, 可以通过概念图谱(存有概念-实体的 isA 关系)发现“口红 $\xrightarrow{\text{isA}}$ 女性用品 $\xleftarrow{\text{isA}}$ 高跟鞋”这一语义链, 从而将两个商品联系起来。

2) 完成多样化推荐。推荐模型可以从知识图谱中发现物品间的异质关联, 从而推荐出不同类型的物品, 产生多样化的推荐结果。例如, 如果一个用户看过电影《盗梦空间》, 推荐系统根据知识图谱中电影间导演、演员和体裁这 3 种不同属性的关联, 向该用户推荐《敦刻尔克》(都由克里斯托弗·诺兰导演)、《泰坦尼克号》(都由莱昂纳多·迪卡普里奥主演)和《终结者》(都是科幻体裁电影)三部完全不同风格的电影。

3) 提供可解释依据。知识图谱中存有的不同物品间的关联即可用作推荐结果的可解释依据。还以上例中看过电影《盗梦空间》的用户为例, 将模型推荐出的三部电影与看过的这部电影之间的关联(即共同的导演、演员和体裁)直接展现给用户, 可让用户直观地了解系统如此推荐的理由, 从而更加愿意接受推荐的结果。

1.4 本文贡献与组织结构

本文的贡献可简单归纳为:

1) 对现有的知识化推荐系统进行科学合理的分类, 系统、深入地综述当前代表性推荐模型的基本原理和技术创新等, 并针对性地展开比较。

2) 介绍相关的推荐模型如何有效地结合知识图谱, 以应对当前实际推荐场景中的各种任务挑战。

3) 结合人工智能技术与应用场景的发展趋势, 展望推荐系统的未来发展方向, 尤其是认知智能推荐系统的发展前景, 为相关研究提供参考意见。

下文的组织结构为: 第 2 章介绍知识化推荐系统的相关背景知识; 第 3 章根据推荐算法的特点, 对现有知识化推荐模型进行分类介绍; 第 4 章从实际推荐场景的角度, 介绍知识图谱对个性化推荐系统带来的帮助; 第 5 章展望知识驱动的推荐系统向认知智能推荐系统发展的前景; 第 6 章为本文结论。

相比已有的知识化推荐系统综述^[13], 本文除了从各推荐模型的算法特点上进行分类介绍, 还特别针对多样化推荐、可解释推荐、序列化推荐和跨领域推荐等当前具有挑战性的现实推荐场景, 介绍相关的推荐模型如何完成这些挑战性任务, 使读者更能从实际应用的角度理解相关模型的特点和优势。

图 1 列出了本文所提及的知识驱动推荐模型的参考文献, 其分类体系即按照推荐(模型)方法和推荐场景这两种分类方式。

此外, 本文不仅关注推荐模型和算法本身的创

新方向, 还紧密结合当下人工智能学界和业界的发展趋势展望知识化推荐模型的未来发展方向, 即预见推荐系统将向认知智能方向发展, 为从事人工智能和推荐系统研究的人士提供了有益的参考。

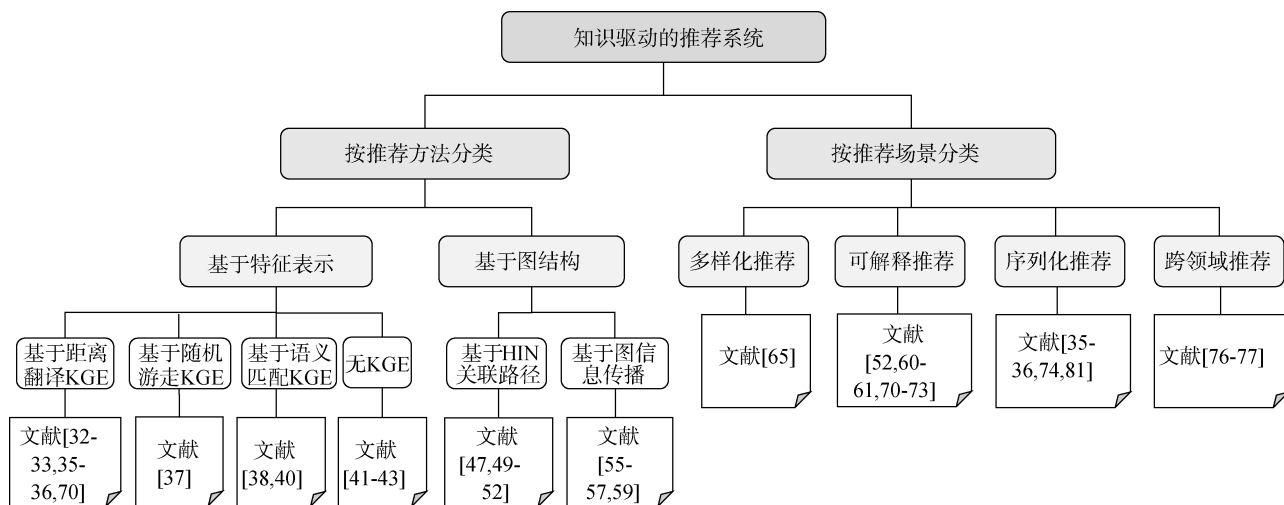


图 1 本文介绍的知识化推荐模型分类体系

Figure 1 The taxonomy of knowledge-driven recommender systems

2 背景知识

为了对知识驱动的推荐系统中各种算法和模型有更清晰的认识, 需要掌握一定的背景知识。本章就当中最重要的相关内容做简要介绍。

2.1 传统的推荐系统

传统意义上, 人们习惯将推荐系统大致分为基于协同过滤, 基于内容和混合式推荐系统三类。

协同过滤算法又可细分为基于用户^[2]和基于物品的协同过滤^[3]。基于用户的协同过滤是假设历史交互记录相似(有很多共同的交互物品)的用户具有相同或相近的偏好, 因此先从用户-物品交互矩阵中学习用户之间的相似性, 找到与目标用户相似的用户集合, 根据集合中用户对候选物品的交互情况(包括评分)来推断目标用户对候选物品的偏好。而基于物品的协同过滤则是对称的做法, 也是从历史交互记录中发现与候选物品相似的物品集合, 基于目标用户对该集合中相似物品的交互情况来推断其对候选物品的偏好。很多基于矩阵分解的推荐模型^[6-7]也可划归为基于协同过滤的推荐。这类推荐方法受用户-物品历史交互记录的制约, 常面临数据稀疏性甚至冷启动问题。

基于内容的推荐则是基于用户与物品特征(也称为内容或画像)之间的相关性来计算两者的匹配程度。特征来源包括标签^[9-10]、用户社交关系^[8-9]、

商品属性^[11]等, 一般通过特征工程将其转换成特征向量, 再基于特征向量间的相似性度量来计算用户与物品的匹配程度。因此, 相似性度量在该类推荐方法中很重要, 常用度量方法包括余弦距离。该类方法的挑战性往往在于特征提取的难度, 对于冷启动的用户和物品同样难以抽取足够的特征进行准确的画像。

混合式的推荐系统一定程度上解决了上述两类推荐系统面临的困难。它集成了前两类方法的特点, 将用户与物品的内容特征作为辅助信息融入协同过滤框架之中^[14], 往往可以达到更好的推荐效果。根据不同的结合方法又可以细分为四种: 分别实现协同过滤与基于内容的方法, 再聚合两者的预测结果; 将内容特征融入协同过滤算法之; 在基于内容的方法之中加入协同过滤的特征; 构造一个融合协同过滤与基于内容相统一的模型。本文关注的知识驱动的推荐系统可以归属为混合式的推荐。

2.2 知识图谱与异构信息网络

谷歌公司在 2012 年公布在其搜索引擎中引入知识图谱后显著提升了搜索效果, 从而引起世人对知识图谱的关注。知识图谱可理解为一种大规模的语义网络, 也是现今知识库的主流形式。现有的知名英文知识图谱包括 FreeBase^[15]、YAGO^[16]、DBpedia^[17]和 Probase^[18]等, 代表性的中文知识图谱则有 CN-DBpedia^[19]和 Zhishi.me^[20]等。

知识图谱中的知识主要以资源描述框架(RDF)形式进行存储,具体表现为一条知识(事实)由一个知识三元组 $\langle e_h, r, e_t \rangle$ 描述。其中, e_h 和 e_t 分别是头实体(或概念)和尾实体(或概念),在图谱中用节点表示; r 是两者间的关系,在图谱中用边来表示。例如, \langle 盗梦空间,导演,克里斯托弗·若兰 \rangle 这个三元组就描述了“电影《盗梦空间》的导演是克里斯

托弗·诺兰”这条知识。知识图谱中的大量实体往往对应了推荐系统中的很多物品(如电影、音乐、各种商品等),其中存储的与物品实体相关的知识就描述了物品各种属性(特征)。如图2所示的有向图就是一个电影知识图谱片段,其中包含了几部电影实体及其各种属性(主演、导演、体裁)所关联的其他实体。

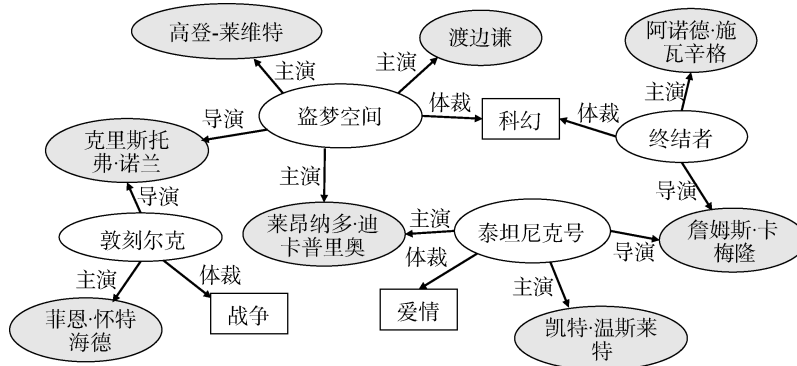


图2 电影知识图谱片段
Figure 2 A toy example of movie KG

一个至少包含两种节点类型或/且至少两种边类型的图被称为异构信息网络(heterogeneous information network, 简称 HIN)。知识图谱中的实体(节点)与关系(边)类型多样,因此也可视为异构信息网络(图)。HIN 中连接两个节点的路径称为元路径(meta path),可用于抽取 HIN 中的关联特征,从而为相关任务的建模服务^[21-22]。图2所示的电影知识图谱即一个异构信息网络,当中的元路径“盗梦空间 $\xrightarrow{\text{导演}}$ 克里斯托弗·诺兰 $\xleftarrow{\text{导演}}$ 敦刻尔克”连接了两个电影类型节点,当中经过了一个人物类型节点,而经过的两条边则都是“导演”关系。该路径描述了其两端电影的语义相关性。

2.3 知识图谱表示学习

受近几年深度学习模型成功应用于各类信息处理任务的激励,新一代推荐系统中也广泛采用深度学习模型。很多知识驱动的推荐系统利用知识的方式都是先将实体、关系表示成低维的稠密向量,常称作嵌入表示(embedding),本文将这种分布式表示向量统称为知识表示向量。然后,知识表示向量输入深度学习模型,作为计算最终推荐分数(即目标用户与候选物品的匹配概率值)的依据。知识表示向量的学习过程是以拟合知识图谱的拓扑结构特征为目标,一般称为知识图谱表示学习(knowledge graph embedding, 简称 KGE)。知识图谱的拓扑结构特征反映的即是实体(包括概念)间的各种关联,这正是知识图

谱能帮助到推荐系统的关键所在。因此,在很多知识驱动的推荐模型中,KGE 是必不可少的步骤。

知识图谱表示学习研究领域,已经提出了很多 KGE 算法,其中最具代表性的一类是基于(表示向量)距离翻译的模型。该类模型的最早代表是 TransE^[23],它将知识三元组 $\langle e_h, r, e_t \rangle$ 中的尾实体 e_t 视作头实体 e_h 利用关系 r 进行翻译得到。基于这一思想设计出如下损失函数:

$$f_r(h, t) = \|h + r - t\|_2^2 \quad (3)$$

在TransE基础上,又有学者提出了一些改进型的翻译模型,例如TransD^[24]、TransR^[25]、TransH^[26]等,以弥补TransE模型在一些复杂场景下知识表示学习能力的不足。

除了基于距离翻译的模型,还有基于图中随机游走的 KGE 算法,其学出的节点表示向量一样能够充分反映知识图谱(即异构信息网络)的拓扑结构特征。这类方法最早的代表是 DeepWalk^[27],它借鉴了词嵌入(Word Embedding)^[28]表示学习中的 Skip-gram^[28]模型思想,将图中随机游走经过的一系列节点视作高度相关的节点(类似词嵌入表示学习中共同出现于上下文中的单词),模型为这些相关节点习得的表示向量会更加接近。在 DeepWalk 之后,又有学者先后提出了 Node2Vec^[29]、Metapath2Vec^[30]和 HIN2Vec^[31]等改进模型。

显而易见, KGE 的学习目标有别于推荐任务的目标(拟合用户-物品交互记录), 因此使用 KGE 的推荐模型大都采用多任务学习框架。在多任务学习框架中, 相关模型的参数训练方式又可分为依次训练(预训练)、交替训练和联合训练三类^[13]。

依次训练先通过KGE习得知识图谱中的实体和关系向量, 然后将这些低维向量引入推荐系统以获得用户和物品的向量表示, 再基于推荐任务目标对向量参数进行调优。联合训练则是将KGE与推荐任务的目标函数组合成如下的联合目标函数, 在一个统一的端到端(end to end)框架下共同训练各自任务中的参数。

$$\mathcal{L} = \mathcal{L}_R + \lambda \mathcal{L}_K \quad (4)$$

其中, \mathcal{L}_R 和 \mathcal{L}_K 分别是推荐和KGE的目标函数, λ 是取值(0,1)的控制参数。交替训练将KGE与推荐看作独立但又相关的任务, 在不同模型训练轮次中交替训练各自目标函数中的参数。

3 基于方法分类的知识化推荐

根据推荐方法的设计原理, 现有的知识化推荐系统模型可大致分为基于特征表示的和基于图结构的两大类。本章按照这种分类方式分别选近几年的代表性研究成果进行介绍。本文提出的分类体系仅是作者的一家之言, 其他综述^[13]中提出的分类方式同样值得参考。最后, 为了便于读者对比本文介绍的各种知识化推荐模型, 我们在第4章的表1中简要列出了相关模型所采用的技术、所适用的推荐任务和场景, 并概述其特点(包括优缺点)。

3.1 基于特征表示的方法

基于特征表示的知识化推荐的一般做法是: 通过对知识图谱进行 KGE 先习得知识表示向量, 它们都是与物品和用户相关的特征表示, 这一过程可视作特征工程。然后, 推荐模型基于知识特征表示产生用户和物品的表示, 再将其输入特别设计的推荐模块以产生满意的推荐结果。

推荐系统中的物品一般都能对应于知识图谱中的很多实体节点, 因此直接对图谱应用 KGE 模型就能够获得物品的表示向量。而通过用户与物品间的交互关系, 用户也能作为一类新的节点加入知识图谱。例如, 图2中看过电影《盗梦空间》的用户可以作为一个节点, 并附带一条边指向“盗梦空间”实体节点的交互关系边加入该图谱。对于包含用户节点的知识图谱, KGE 模型也能直接学出用户的表示向量。倘若图谱中没有用户节点, 因为用户交互过的物

品能代表用户的偏好特征, 则可以通过某种方式(或算法)将历史交互物品的嵌入表示综合成用户的表示向量。此外, 也有不少知识化推荐系统采用其他更为复杂的方式来获得用户和物品的表示向量。

假定 u 和 i 分别表示目标用户 u 和候选物品 i 的表示向量, 推荐模型大都基于 u 、 i 来计算两者的匹配分数, 即

$$P(i|u) = f(u, i) \quad (5)$$

当中的函数 f 即推荐计算模块, 可直接用向量内积操作或者是复杂的深度神经网络计算(如多层感知机)。还有一些模型在公式5中还将关系表示 r 作为 f 函数的输入。

现有基于特征表示的知识化推荐系统中, 我们根据其知识特征表示方法的不同, 将其又细分为下面几种类型。

3.1.1 基于距离翻译特征表示学习的模型

该类模型中的 KGE 是前文提及的基于(知识表示向量)距离翻译的模型。Zhang 等人^[32]提出的 CKE 模型是本类知识化模型的早期代表。该模型中, 来源于知识图谱的是结构知识, 先通过 TransR 模型^[25]习得其结构嵌入表示(即图谱中的实体表示)。物品图像(如电影海报)和文字(如评论)则由自编码器习得其嵌入表示, 与结构的嵌入表示一起组成物品的综合表示。物品表示与用户表示再通过内积计算得到针对隐式反馈的推荐分数。CKE 中的推荐任务与 KGE 任务采用联合训练方式。

引入知识图谱的新闻推荐系统 DKN^[33]的 KGE 则采用 TransD^[24], 以生成新闻标题中出现的实体嵌入表示, 还利用词嵌入模型^[28]获得标题中单词的嵌入表示。接下来使用 KCNN 模型^[34]将所有特征表示有效综合成新闻的表示。在获得新闻表示的基础上, 通过注意力网络来产生用户的表示, 而推荐计算模块则使用多层感知器网络来算出最终的用户与待推荐新闻匹配分值。该模型中推荐任务与 KGE 任务采用依次训练方式。

推荐任务与 KGE 任务采用交替训练方式的知识化推荐模型则有 MKM-SR^[35]。该模型针对的是会话推荐(session-based recommendation)任务, 属于序列化推荐(sequential recommendation)。MKM-SR 通过引入用户-物品交互行为的具体类型(如加入购物车、收藏、下单等)来增强会话(实际对应了用户)的初始表示, 然后利用软注意力机制来产生会话的最终表示。同时, 考虑到 TransH^[26]能更好地处理一对多、多对多的实体间关系, MKM-SR 用它来学习知识图谱中的物品实体表示, 从而引入了知识的信息。最终的推荐

分数是将会话和物品的表示输入多层感知机计算得到。针对序列化推荐的知识化推荐模型还包括 Chorus^[36], 当中采用 TransE^[23]来学习知识表示, 下一章再单独介绍针对这类推荐任务的模型。

3.1.2 基于随机游走特征表示学习的模型

除了利用基于距离翻译的 KGE 以外, 也有不少知识化推荐系统采用基于随机游走的 KGE。例如, KTGAN 模型^[37]中的 KGE 直接采用 Metapath2Vec^[30], 它是 DeepWalk^[27]的改进模型。除了知识向量, KTGAN 模型还利用词嵌入模型习得用户与物品的标签嵌入表示, 一起组成用户/物品优化的表示向量。然后, 将其输入生成-对抗网络^[32](GAN)来计算最终的推荐分数。

3.1.3 基于语义匹配特征表示学习的模型

另有一些知识化推荐模型中的 KGE 则是基于语义匹配的。例如, RCF^[38]先用基于关系语义匹配的 KGE 模型——DistMult^[39]来对物品之间的关系进行建模, 以获得物品的表示向量。其表示学习目标函数中的得分函数如下:

$$f(i, r, j) = \mathbf{q}_i^T \cdot \text{diag}(\mathbf{r}) \cdot \mathbf{q}_j \quad (6)$$

其中, i 和 j 代表两个物品, \mathbf{q} 是物品的嵌入表示, $\text{diag}(\mathbf{r})$ 是关系的对角矩阵。然后, 物品表示与用户表示也像其他模型一样输入集成了注意力机制与多层感知机的推荐模块, 产生最终的推荐分数。RCF 中向量的表示学习与推荐任务的学习是采用联合练的方式。而 MKR 模型^[40]则采用了交替训练方式, 它并没有直接使用现成的 KGE 方法, 而是采用模型作者自创的一个基于语义匹配的特征表示学习模型。

3.1.4 无 KGE 特征表示的模型

以上介绍的基于特征表示的知识化推荐模型, 如2.3节所介绍的, 都先通过KGE来获得特征(知识)表示向量, 而KGE的优化目标往往有别于推荐任务本身的目标(但高度相关)。还有一些基于特征表示的知识化推荐模型, 其中的特征表示并不需要经过一个学习目标有别于推荐任务目标的KGE, 而是直接使用特征的随机初始化表示向量, 直接由推荐任务的目标来监督其参数训练。

这类模型的代表包括 DIN 模型^[41], 当中每个物品的表示直接由特征 ID(如商品 ID、点评 ID、类别 ID 等)所映射的嵌入表示向量组成。用户历史上交互过的物品与待推荐物品表示一起经过注意力激活单元产生更优的表示后, 再通过求和池化(sum pooling)操作得到用户的表示。然后, 用户与待推荐物品的表示再经过多层感知机求出推荐分数。DIN 中虽然没

有提及知识图谱, 但其物品的特征其实就是物品知识图谱中的各自属性(值), 因此将其也划归知识化推荐模型。此外, AKR 模型^[42]和 ACAM 模型^[43]也属于该类模型, 其实现知识化推荐的方案和 DIN 的思路基本一致。前者还使用了卷积操作来压缩用户的特征表示维度, 而后者则用更具特色的协同注意力来代替注意力。

这类模型中, 特征初始表示向量中的参数因为没有经过 KGE 目标的训练, 其本身的内容无法反映知识上的关联。但是, 对于具有共同特征的两个物品, 其表示向量却有一定程度的相似性(因为包含相同的特征表示内容), 正好能反映两个物品因共享某些特征而具有的知识关联。即便表示向量的参数是基于后续推荐任务目标来训练, 仍能保持物品表示向量间的相关性, 这就是该类模型仍具有优良推荐性能的原因。

除了以上介绍的这些研究工作, 文献[44-46]中提出的推荐模型也都属于基于特征表示的知识化推荐模型。

3.1.5 基于特征表示的方法小结

以上 3.1.1 至 3.1.3 小节介绍的基于特征表示的知识化推荐模型, 其最终推荐效果尽管与其推荐计算模块的设计有关, 但在很大程度上会受其采用的 KGE 方法的性能影响。而且因为 KGE 的学习目标与推荐任务的目标有所区别, 所以基本都涉及多任务学习的框架, 而采用哪种多任务参数的训练方式(预训练、联合训练还是交替训练)也会影响最终的推荐效果。因此, 当中很多模型的作者在其研究中大都基于对比实验的结果来选择最优的 KGE 模型和多任务参数训练方式。

相较而言, 3.1.4 小节中介绍的没有采用 KGE 的模型, 其推荐效果则主要看其推荐主模块设计的质量, 包括其推荐任务目标函数的设计。还要特别指出的是, 无论哪种基于特征表示的推荐模型, 其性能都会受其使用的知识图谱质量影响, 即从中抽出的知识特征的多少与相关性(significance)都会影响最终的推荐效果。正如很多基于特征工程的机器学习建模, 特征的质量往往是决定模型任务效果的关键因素, 甚至其对效果的重要性超过模型本身。

3.2 基于图结构的方法

第二大类知识驱动的推荐系统采用基于图结构的方法来融入外部知识, 也可以大致再细分为两类。第一类是将用户节点和交互关系补入物品知识图谱, 将这样的大图视为一个大规模的 HIN, 然后利用当中的关联路径信息来捕获用户与物品间的(语义)

关联特征, 以此来推断用户对物品的偏好。这类方法也被称为基于路径的方法^[17]。第二类是最近几年新提出的基于图信息传播的模型, 以图的拓扑结构来监督知识特征的表示学习, 从而改善推荐效果。下面分别对这两个细类的模型做介绍。

3.2.1 基于 HIN 关联路径的模型

该类模型的早期代表包括 Yu 等人^[47]提出的 HeteRec。当中, 先从 HIN 中抽取 L 条不同类型的关联用户和物品的元路径, 再基于 PathSim^[48]来计算用户与物品的相关度(相似度), 以优化用户-物品的交互矩阵。后来, 因为意识到不同元路径对不同用户应有不同的权重, Yu 等人^[49]又提出了 HeteRec-p, 在实现推荐的过程中对用户进行了聚类, 最终的推荐分数要将目标用户与其所在用户群组的相似度考虑在内, 而相似度仍是由基于元路径的关联特征计算而得。

上述基于元路径相似度的模型只关注用户交互过的物品(即喜欢的物品), Shi 等人^[50]提出的 SemRec 模型则不仅考虑了用户不喜欢的物品, 还利用图谱中每条边上的具体关系(属性值)产生带权重的元路径。先将每条带权元路径分解成多条原子元路径(atomic meta path), 其相关度的计算仍采用已有的方法。所有原子元路径的相关度加权和才作为该带权元路径算出的相关度。

为了弥补元路径表示能力的不足, Zhao 等人^[51]设计的 FMG 模型中, 先将 HIN 分解成多个元图(meta graph)。然后, 在每个原图中应用矩阵分解产生用户和物品的表示向量。再通过因子分解机(FM)来融合不同元图中用户/物品的表示特征, 以算出最终的推荐分数。

HIN 中连接目标用户和待推荐物品的元路径不仅可以用作计算两者相关度的强弱, 因其直观易懂, 还可用作推荐的可解释依据。如 1.3 节中提及的例子, 向看过《盗梦空间》的用户推荐《敦刻尔克》的理由能显式地从元路径“盗梦空间 $\xrightarrow{\text{导演}}$ 克里斯托弗·诺兰 $\xleftarrow{\text{导演}}$ 敦刻尔克”中获知。Xian 等人^[52]意识到元路径的这点优势, 首次将强化学习与知识图谱结合, 提出了可解释推荐模型 PGPR。作者设计了一个基于多跳评分函数的软奖励策略来帮助智能体(agent)从 HIN 中寻找最优的连接目标用户与候选物品的元路径。这些路径正好可用作推荐结果的可解释理由。

3.2.2 基于图信息传播的模型

最近几年, 随着图神经网络^[53-54](GNN)的流行,

越来越多的知识化推荐模型也采用 GNN 的架构来学习实体和关系表示, 从获得更优的推荐结果。GNN 的基本思想可简单概括为基于邻居信息聚合的嵌入表示传播。

较早采用这种思想来引入知识的推荐模型是 RippleNet^[55], 该模型的图中所传播的表示向量嵌入的就是用户对物品的偏好信息。其基本设计思想基于: 用户的偏好不仅能用其交互过的物品来表征, 知识图谱中与这些物品关联的其他实体也能在一定程度上表征用户的偏好。模型首先将用户交互过的物品视作用户的原始兴趣, 然后通过迭代计算得到用户原始兴趣的 K 阶邻居(即图中通过 K 步跳转可达)的表示向量。该过程可视作用户的兴趣传播, 距离越远的邻居能表征用户兴趣的程度越弱, 如同涟漪散开的波纹, 模型的名称由此而来。最终, 一个用户 u 的偏好表示向量 \mathbf{u} (即 u 的画像)按照如下公式计算:

$$\mathbf{u} = \mathbf{o}_u^1 + \mathbf{o}_u^2 + \dots + \mathbf{o}_u^K \quad (7)$$

上式中的 \mathbf{o}_u^i 是用户 u 的 i 阶邻居向量, 其在实际计算中包含了范围在 i 跳以内的所有邻居信息。

RippleNet 的作者是基于传播思想自创了用户偏好传播的模型, 而 KGCM^[56]和 KGAT^[57]则分别直接借鉴了经典的 GNN 模型——图卷积网络^[53](GCN)和图注意力网络^[54](GAT)。KGCM 中的迭代信息传播和 RippleNet 正好相反, 即先从距离候选物品节点 H 跳的邻居开始, 是向着候选物品节点逐渐缩小范围的向心传播。KGAT 中, 则是利用 TransR^[25]先习得物品实体的知识表示作为物品的初始表示, 其信息传播的方向则是以候选物品节点为起点向外传播。

后来, 有学者发现基于 GNN 的推荐模型会随着传播深度的增加出现过度平滑问题^[58]。为此, Chen 等人^[59]设计出一种新的异构传播方法, 提出了 HKIPN 模型。该模型用到的 HIN 也是包含物品知识图谱和用户-物品交互关系的统一图。在 HKIPN 的传播过程中, 用一种层次化方法来显式、并发地聚合节点的高阶邻居。共分为三层, 第一层是用 TransR 做 KGE, 第二层是层次化和注意力传播, 第三层则是集合所有层次的表示并计算推荐分数。

基于图信息传播的模型中同样有不少是针对可解释推荐的任务^[60], 我们将在下一章再另行论述。本节提到的代表性模型在表 2 中列出, 包括它们各自采用的路径相关度计算方法和图信息的传播模型。

除了以上介绍的这些研究工作, 文献[61-64]中提出的推荐模型也都属于基于图结构的知识化推荐模型。

3.2.3 基于图结构的方法小结

3.2.1 小节介绍的基于 HIN 关联路径的模型大多没有涉及深度学习模型, 模型参数较少, 所以计算效率较高, 但准确率等推荐效果不及深度学习模型。

3.2.1 小节介绍的基于图信息传播的模型, 基本都使用了 GNN 模型, 虽然其推荐效果提升明显, 但因为模型参数太多, 导致模型的迭代训练收敛慢, 耗费时间过多(尤其是当 GNN 层数较多时)。从最基本的原理来看, 基于图结构的模型与基于特征表示的模型并无本质的区别。这是因为 KGE 对特征表示向量的参数训练也是基于知识图谱的图结构而设计的优化目标。现有的知识化推荐模型, 基本上都是将知识图谱的拓扑结构特征(反映的是实体间的知识关联)作为最核心、最重要的信息加以利用。

此外, 本章中介绍的很多模型还是针对下文第四章提到的几类具有挑战性的实际任务场景而设计的。例如, 前述基于 HIN 路径关联的知识化推荐模型 PGPR^[52], 其挖掘出的目标用户与候选物品的关联路径能够直观地表达出两者间的语义联系, 因而可直接用于对推荐结果的解释说明。

4 基于场景分类的知识化推荐

随着互联网生态与业态的发展, 推荐系统要面对越来越多挑战性的实际场景, 用户对推荐效果的要求也越来越高。本章再分别介绍一些推荐模型如何有效地引入知识以应对以下几个具有代表性和挑战性的推荐场景。

4.1 多样化推荐

准确率并不是体现推荐效果的唯一指标, 仅以准确率为目标设计推荐算法往往会导致推荐结果的高度同质性, 形成所谓“信息茧房”效应。例如, 模型会判断一位购买过华为 Mate 40 的用户是高端智能手机的潜在客户, 于是向其推荐 iPhone 12。这一推荐看似准确捕获了客户的偏好, 但他短期内一般不会再买一台新手机, 所以这样的推荐是无意义的。近几年, 多样化推荐因为可以避免同质化推荐的不良效果, 受到越来越多的关注。

知识图谱中存有物品实体间多样的语义关联, 因而有助于推荐模型产生多样化的推荐结果。Lu 等人^[65]提出的 DivKG 模型就利用了知识的辅助信息来让推荐模型在准确率和多样性上达到平衡。该模型工作过程基本分两步, 第一步还是用 TransE^[23]做 KGE, 习得的用户和物品向量可以用来计算用户与物品, 以及物品与物品之间的相关性; 这些相关性用于第二步的行列式点过程(DPP)^[66]中的核矩阵

(kernel matrix), 以计算每个候选推荐列表的多样性得分。实现多样化的推荐结果。DPP 已被不少面向多样化推荐的模型采用^[67], 被证明是一种有效解决多样化推荐的算法。

4.2 可解释推荐

除了前文介绍的 PGPR 模型^[52], 还有不少其他引入知识的推荐模型也有效地解决了可解释推荐的问题。例如, Xiang 等人^[60]提出的 KPRN 基于知识图谱中的路径推理来为推荐提供可解释依据。KPRN 以目标用户和待推荐物品作为输入, 先在知识图谱中搜索用户和物品之间的路径, 基于这些路径的推理来推断用户对物品的偏好程度。模型分为三部分: 第一部分是实体、实体类型和关系映射到表示向量, 基于这些向量构成的路径表示就包含了更多的语义信息; 第二部分是用长短期记忆网络^[68](LSTM)对路径进行学习和编码表示, 以捕捉路径序列中的依赖信息; 第三部分是将多条路径通过池化操作综合成一个表示, 再计算出最终用户会与待推荐物品交互的概率。KPRN 模型中找出的每一条目标用户到候选物品间的路径都对应了一条解释理由, 而模型为这些路径算出的分数则可作为每条理由被用来解释向该目标用户推荐该候选物品的可能性。

另外, 有些模型并不为目标用户计算值得推荐的物品, 而仅为已经产生的推荐结果从知识图谱中寻找可解释依据, 这类任务称为事后(post-hoc)可解释推荐^[69]。如文献[70]提出的算法, 先将用户、商品、商品品牌、评论关键词等组成一个知识图谱, 在图中用广度优先算法找出所有关联目标用户和被推荐商品的路径。然后, 基于路径上每个实体节点的向量计算每条路径的分值, 最后将分值最高的路径作为该推荐结果的可解释依据。如图 3 所示的商品知识图谱中, 发现 Bob 以往的商品评论中提到过“iOS”, 而其他用户对 iPad 的评论中也提到了“iOS”, 并且 iPad 与 Bob 曾经购买过的某个商品 i 都属于 Apple

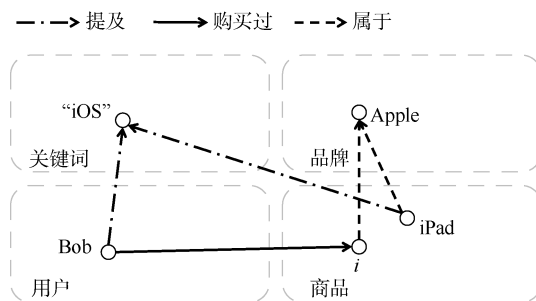


图 3 商品知识图谱中用户与商品间的可解释路径
Figure 3 The explainable paths between users and products in product KG

品牌。该图中可发现两条 Bob 到 iPad 的关联路径, 可作为 Bob 购买 iPad 的可解释依据候选。

除上述模型外, 文献[60-61,70-73]也是利用知识图谱完成可解释推荐的代表性研究工作。

4.3 序列化推荐

序列化推荐任务不同于一般的评分预测和列表式推荐, 它是基于用户历史上(尤其是最近)的交互记录来预测用户下一个会交互的物品。由于用户交互序列中的物品之间往往反映出前后的依赖关系, 因此序列化推荐系统都会考虑时间维度特征, 大都基于序列化的模型来设计。序列化推荐中还有一个子类称为会话(session)推荐, 即基于用户最近一个会话期间的交互记录来预测下一交互物品。

前文介绍过的 MKM-SR^[35]和 Chorus^[36]都是引入知识图谱来改善序列化推荐效果的模型。知识的引入更有利于模型获知序列中物品之间的语义关联, 从而能更准确地捕获当中的相互依赖, 这是提升序列化模型性能的关键所在。在此, 我们补充介绍另一知识化序列推荐模型的代表工作 KSR 模型^[74]。该模型中, 用户的表示分为两部分: 一部分是序列化的用户偏好, 另一部分是物品属性级的用户偏好。用户的序列化偏好通过门控循环单元^[68](GRU)计算而得; 物品属性级的用户偏好则通过键值对记忆网络^[75](Key-Value Memory Network)产生。物品属性知识的引入增强了模型中物品的语义表示。另外, 基于关系的表示向量, KSR 还可获取显式的用户对物品属

性级的兴趣, 提高了序列化推荐中的可解释性。

4.4 跨领域推荐

跨领域推荐在当前实际应用场景中经常涉及的任务, 例如向电影观众推荐音乐作品, 或向微博用户推荐淘宝商品。不同领域的异构性(用户不同、物品不同或特征异构)给这种推荐任务带来了巨大的挑战, 而知识图谱提供的丰富知识则有助于关联不同领域的用户、物品和特征。

在知识化跨领域推荐模型中, 我们先介绍基于实体关联的跨领域推荐方法^[76]。该方法的作者首先利用英文知识图谱 DBpedia^[17]构建一个丰富的语义网络, 当中包含用户的 POI、音乐家等实体, 及其类别与属性等。然后, 作者设计了一种基于图的权重迭代传播算法, 计算出用户所在的 POI 与音乐家基于关联路径的相关度, 从而实现为 POI 上的用户推荐音乐家的跨领域推荐。以图 4 为例, 图中展现了一个将维也纳国家歌剧院关联到奥地利籍作曲家 Gustav Mahler 的语义网络(KG 片段), 当中包括了建筑的类型、地址、建造时间等信息, 以及音乐家的出生年代、逝世地点、作曲风格等知识。该网络中从最左边的源头实体(歌剧院)到最右边的目标实体(音乐家)的连通路径正是两者存在关联的证据, 也是跨领域推荐的重要依据。在该语义网络中, 为了找出最匹配 POI 的音乐家, 需要计算所有目标节点(待推荐音乐家)与源头节点(用户 POI)的关联程度, 并推荐关联最紧密的目标节点。

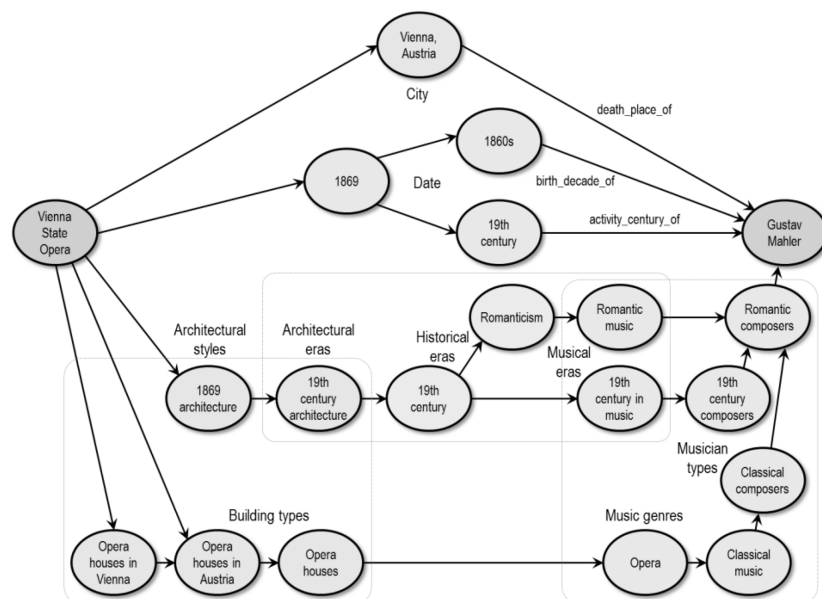


图 4 POI 到音乐家的跨领域图谱关联路径^[76]

Figure 4 The KG paths between POIs to musicians^[76]

另外, Yang 等人^[77]还利用知识来发掘特征间的语义关联, 以完成跨领域推荐任务。在其推荐场景中, 用户和物品都用标签进行画像, 但是两个领域的异构性使得用户与物品的特征(即标签)差异性交大, 无法以特征的重合程度来判断用户与物品的相关性强弱。为有效计算处于一个领域的目标用户与另一领域中候选物品的相关性, 作者引入百科知识图谱, 其实体间的关系对应于百科各页面之间的超链接。先借用显式语义分析模型^[78](ESA)来产生实体(能对应标签特征)的概念向量, 而概念向量间的距离正好度量了特征间的语义强弱。有了特征间的语义关联, 再基于流行排序^[79]来计算目标用户与候选物品间的匹配度, 从而实现跨领域推荐。

4.5 小结

本章针对几类颇具挑战性的实际推荐场景, 又分别介绍了一些代表性的知识化推荐模型。可以看

出, 针对特定的推荐任务并不一定需要针对性地设计特别的算法。很多模型的算法基本框架仍可对应到第三章的方法分类体系中(参见表 1、表 2)。在此, 我们简单再小结一下被引入的知识帮助推荐系统更好地胜任这些推荐场景的关键所在: 多样化推荐需要利用好物品类型、体裁这类(属性)知识, 因为它们才是“多样”的具体表现; 可解释推荐是从目标用户(或用户交互过的物品)节点到候选物品节点的通路上抽取两者间的语义关联; 序列化推荐是利用好知识图谱发掘交互序列中相邻(或相近)物品之间的序列知识关联; 跨领域推荐则是通过知识图谱来解决不同领域中实体或特征间的语义匹配和关联。

5 认知智能推荐系统的发展前景

知识驱动的推荐系统发展日趋成熟, 现在已经产

表 1 基于特征表示方法的各种知识驱动的推荐模型对比

Table 1 Some representative KG-based models

| 模型名 | 知识特征表示学习方法 | 推荐计算模块 | 适用任务与场景 | 特点概述 |
|------------------------|---|-------------|------------|-------------------------|
| CKE ^[32] | TransR ^[25] 、自编码器 | 向量内积 | 列表式推荐 | |
| DKN ^[33] | TransD ^[24] | 注意力+多层感知机 | 列表式推荐 | 能充分发挥知识特征的作用, 但推荐效果易 |
| MKR ^[40] | 自创的语义匹配 | 多层感知机 | 点击预测和列表式推荐 | 受其采用的 KGE 模型性能和多任务参数训练 |
| KTGAN ^[37] | Metapath2Vec ^[30] | 生成-对抗网络 | 列表式推荐 | 方式的影响 |
| RCF ^[38] | DistMult ^[39] | 注意力+多层感知机 | 列表式推荐 | |
| KSR ^[74] | TransE ^[23] | 记忆网络+多层感知机 | 序列化推荐 | 能实时跟踪用户的最新偏好变化, 但记忆网 |
| MKM-SR ^[35] | TransH ^[26] | 软注意力+多层感知机 | 序列化推荐 | 络的操作频繁且复杂 |
| Chorus ^[36] | TransE ^[23] | 向量内积或多层感知机 | 序列化推荐 | 除了 KGE, 用户-物品历史交互序列中的会话 |
| DIN ^[41] | 无 | 注意力+多层感知机 | 点击预测 | 划分也会影响最终推荐效果 |
| AKR ^[42] | 无 | 注意力+多层感知机 | 列表式推荐 | 省去了 KGE 模型的设计与多任务框架, 训练 |
| ACAM ^[43] | 无 | 协同注意力+多层感知机 | 列表式推荐 | 方式直接, 但更易受引入的知识特征质量影 |
| DivKG ^[65] | TransE ^[23] 或 TransH ^[26] | 行列式点过程 | 列表式多样化推荐 | 响 |
| | | | | 更看重推荐多样化的性能指标 |

表 2 基于图结构方法的各种知识驱动的推荐模型对比

| 模型名 | 路径相关度计算方法 | 图信息传播模型 | 适用任务与场景 | 特点概述 |
|---------------------------|----------------------------|---------------------|------------|-------------------------|
| HeteRec ^[47] | PathSim ^[48] | | 列表式推荐 | 虽然推荐效果比不上深度学习模型, 但所采 |
| HeteRec-p ^[49] | PathSim ^[48] | | 列表式推荐 | 用的路径相关度计算方法参数少, 计算效率 |
| SemRec ^[50] | PathSim ^[48] | | 评分预测 | 较高 |
| FMG ^[51] | 矩阵分解 | | 评分预测 | 利用了矩阵分解与因子分解机的优势 |
| PGPR ^[52] | 强化学习 | | 列表式可解释推荐 | 推荐准确率和可解释性都非常高, 但模型设 |
| RippleNet ^[55] | | 自创 | 点击预测与列表式推荐 | 计与训练都更复杂 |
| KGCN ^[56] | | GCN ^[53] | 点击预测与列表式推荐 | 推荐效果与 GNN 模型本身的优劣有关, 一般 |
| KGAT ^[57] | | GAT ^[54] | 列表式推荐 | 而言, 较多的图传播层数(跳转数)有更好的性 |
| HKIPN ^[59] | | 自创 | 点击预测与列表式推荐 | 能, 但更消耗训练时间, 且容易产生过渡平滑 |
| KPRN ^[60] | 分布式表示+LSTM ^[68] | | 列表式可解释推荐 | 问题 |
| | | | | 计算结果既是推荐物品产生的依据也是推荐 |
| | | | | 理由选择的依据 |

生众多具有技术特点又有实际应用效果的范例。个性化推荐作为人工智能最具代表性的应用,其未来发展方向也和人工智能的发展趋势密切相关。鉴于让机器像人类一样具备高阶认知能力已成为人工智能领域发展的必然趋势,利用知识研发具有认知智能的推荐系统也极具发展前景。

5.1 在推荐系统中实现认知智能的挑战

当前的人工智能技术和应用发展虽然如火如荼,但大多还是停留在图像识别、语音识别等感知智能的阶段,机器还无法像人类那样具备高阶认知的能力,还距离“强人工智能”甚远。认知智能是以知识的发现和应用为核心内容的高阶智能形式,是让机器具备高阶认知能力的关键。人类的高阶认知包括演绎、归纳、推理、理解和解释等能力,其中理解和解释能力对机器而言尤为重要。具备高阶认知能力的机器应能理解数据、理解自然语言进而理解现实世界,具体表现在机器具备解释数据、过程的能力,进而能够解释现象,甚至像人类一样可以进行推理、规划等。然而,让机器具备高阶认知能力并非易事,以下几方面的问题也对建成具有认知智能的推荐系统提出了很大挑战。

1) 特定推荐场景或单模态下的数据稀疏问题。虽然像 BERT^[80]这类预训练语言模型已在很多自然语言处理和推荐任务^[81]中展现出优异的性能,甚至体现出一定的认知能力,但是其训练要基于大规模的语料库。在一些特定的推荐场景中,不仅无法获得充足的用户-物品交互记录,而且与推荐任务紧密相关的语料数据也十分缺乏。如果推荐领域的语料(如商品评论)的语义、语法等特征明显有别于其他领域,则无法使用通用语料来训练语言模型。另外,很多推荐领域中,尽管都存在文本以外的多模态数据(图像、音频、视频等)能为模型提供更丰富的信息,但多模态的特征抽取、信息融合等要比单模态的相关任务复杂,本身就具有很大挑战性。

2) 缺乏常识的支持。常识是人所共知的基本知识,因此很少显式地存储于现有知识库中。同样因为很少显式地在源数据语料中提及,难以用现有的知识抽取模型抽取出来。但是,常识的缺乏却很大程度上限制了机器对一些普通现象的正确理解,也使得推荐模型难以发现基于常识的实体关联。例如,没有关联交易的数据支持,推荐模型无法捕获“口红”与“高跟鞋”的关联,但这种关联却能从“女性用品”的常识中获得。

3) 模型在可解释性、高阶推理、因果推断等方面能力的缺陷。虽然深度学习模型在分类、内容生

成等任务上效果显著,但其黑盒属性使其工作原理严重缺乏可解释性。另外,现有大多数模型在高阶推理、因果推断上的能力还是非常欠缺的。这些认知能力的缺乏使得推荐模型无法产生更多准确的推荐结果,其产生的推荐结果也难以让人理解其推荐的缘由。

5.2 具有认知智能的全新个性化推荐

综上,知识图谱和相关技术的利用应该以帮助智能系统克服上述短板为出发点。个性化推荐系统应当向具有认知智能的推荐系统方向发展。据此,对于推荐系统领域的研究工作,我们尝试建议以下几个具有研究意义和应用价值的具体方向。

5.2.1 多模态知识驱动的推荐

相比于单模态知识,多模态的知识可以提供更加丰富的辅助信息,会进一步提升推荐系统的效果。推荐系统引入多模态知识的一个关键是多模态知识的表示学习。文献[82]指出,多模态知识的表示学习方法可分为基于特征和基于实体两类。基于特征的做法是把实体的各模态信息(如电影的海报图片、剧情简介等)视作实体额外的辅助信息,分别通过某种表示学习模型习得相应的模态特征表示后,再聚合成实体的统一增强表示,前述的 CKE^[32]即采用此种方法。而最新提出的 MKGAT 模型^[82]则采用基于实体的方法,即把实体不同的模态信息都视为多模态知识图谱中的一个实体,和原有的物品等实体都建立起关联。然后,分别用 ResNet^[83]和词嵌入模型^[28]得到图片与文本实体的嵌入表示,再通过 KGE 来获得物品实体的嵌入表示。如此获得的物品表示实际上是被多模态知识(实体)表示增强过的,因而有助于推荐模型获得更优的推荐效果。

5.2.2 具备常识理解的推荐

具备常识是让推荐系统进一步提升效果的关键。以电商网站上的推销为例,向购买了“口红”的用户推荐购买“高跟鞋”而不是“皮鞋”,或向购买了“西装”的用户推荐购买“领带”而不是“围巾”,这样的精准推荐需要诸如“使用口红和高跟鞋的是女士”、“穿西装、打领带的是男士”这样的常识支持。尤其是在历史购买记录稀疏,不足以挖掘出准确的物品关联规则情况下,常识对推荐系统的帮助更为显著。常识的获取、理解与推理,常识知识图谱的构建等问题的研究在近几年方兴未艾,但也因为常识自身的特点使相关研究工作面临不小的挑战。将常识融入推荐系统以提升个性化推荐效果的工作同样面临全新的挑战,相关研究已受到阿里巴巴等具有实际业务需求的业界公司高度关注,也将很快成

为推荐系统研究领域的新热点。

5.2.3 解说式、劝说式与抗辩式推荐

具有认知智能的推荐系统实际上可被视为一种具有高阶认知能力的机器人, 其与用户的交互方式可参照对话机器人。在此, 我们预计为用户提供解说式、劝说式与抗辩式推荐即将成为流行的推荐方式。

1) 解说式推荐即推荐系统向用户推荐物品时还能够解释说明推荐的原因, 实际上对应了4.2节介绍的可解释推荐。知识图谱中蕴含的实体间关联能够有效弥补黑盒推荐模型在可解释性上的短板, 是让推荐系统甚至包括其他智能系统具备可解释能力的关键。

2) 劝说式推荐是指在向用户推荐一个主要物品同时, 或已知用户历史交互行为的前提下, 能够劝说用户同时接受其他相关物品的推荐。该种推荐方式能充分反映新一代推荐系统在认知能力上的提升, 要求系统能够充分认知到不同物品间的潜在关联。例如, 劝说一位刚购买过“登山杖”的用户再购买“墨镜、头巾、水壶”等其他商品是合理的建议。这样的相关商品推荐可以基于丰富的关联购物记录发掘出来, 但交易数据稀疏时就要求推荐系统能够认知如下的知识(常识)推理链才能达成:

“登山杖→爬山用→在野外→阳光刺眼→需要墨镜
...→风大→需要头巾
...→气温高→口渴→喝水→需要水壶”

因为墨镜、头巾、水壶三件商品在很多知识图谱中不会和登山杖一起归属于“户外商品”这个概念, 之前的知识化推荐模型很难发现它们之间的关联。机器要认知这些推理链, 必须像人类一样具备常识和高阶推理能力。

3) 抗辩式推荐是指系统能充分了解用户真正所需, 因此当用户表达出的需求与系统的理解有偏差时, 能够提供看似截然不同但其实是真正准确的推荐。仍以向一个户外运动爱好者推荐商品为例, 假设该用户先购买了“帐篷”, 然后在其搜索“雨披”时, 系统却建议用户优先购买“防潮垫、睡袋”。这是因为基于推荐系统的认知, 相比雨披, 防潮垫和睡袋才是与帐篷更紧密关联的物品(没有防潮垫帐篷无法使用, 没有睡袋没法晚上在帐篷里睡觉)。即便该用户已经买过防潮垫和睡袋, 但面对这种善意的“提醒”也是可以理解和接受的。

6 结论

本文作者充分调研了现有知识驱动化推荐系统的相关研究工作, 在对其进行合理分类的基础上,

系统性地综述了代表性的模型和算法基本思想, 从而为相关研究提供基础。相比其他知识化推荐系统综述, 本文还梳理了几类具有挑战性的实际推荐场景, 针对性地介绍了相关知识化推荐模型如何有效应对这些场景, 可使读者对现有相关工作有更加深入的理解。此外, 本文还基于人工智能领域向认知智能发展的趋势, 建议了几个颇具前景的认知智能推荐系统发展方向, 希望能为相关研究提供有价值的参考意见。

参考文献

- [1] XIAO Yanghua. Knowledge Graph – Concepts and Techniques[M]. BEIJING: Publishing House of Electronics Industry, 2019.
(肖仰华. 知识图谱-概念与技术[M]. 北京: 电子工业出版社, 2019.)
- [2] Jin R, Chai J Y, Si L. An Automatic Weighting Scheme for Collaborative Filtering[C]. *The 27th annual international ACM SIGIR conference on Research and development in information retrieval*, 2004: 337-344.
- [3] Linden G, Smith B, York J. Amazon.com Recommendations: Item-to-Item Collaborative Filtering[J]. *IEEE Internet Computing*, 2003, 7(1): 76-80.
- [4] Zhou G R, Zhu X Q, Song C R, et al. Deep Interest Network for Click-through Rate Prediction[C]. *The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018: 1059-1068.
- [5] Li J Y, Tang B, Cercone N. Applying Association Rules for Interesting Recommendations Using Rule Templates[C]. *Advances in Knowledge Discovery and Data Mining, 8th Pacific-Asia Conference*, 2004: 166-170.
- [6] Salakhutdinov R, Mnih A. Bayesian Probabilistic Matrix Factorization Using Markov Chain Monte Carlo[C]. *The 25th international conference on Machine learning - ICML '08*, 2008: 880-887.
- [7] Koren Y, Bell R, Volinsky C. Matrix Factorization Techniques for Recommender Systems[J]. *Computer*, 2009, 42(8): 30-37.
- [8] Massa P, Avesani P. Trust-Aware Recommender Systems[C]. *The 2007 ACM conference on Recommender systems*, 2007: 17-24.
- [9] Yang D Q, Chen L H, Liang J Q, et al. Social Tag Embedding for the Recommendation with Sparse User-Item Interactions[C]. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2018: 127-134.
- [10] Sen S, Vig J, Riedl J. Tagommenders: Connecting Users to Items through Tags[C]. *The 18th International Conference on World wide Web*, 2009: 671-680.
- [11] Ziegler C N, Lausen G, Schmidt-Thieme L. Taxonomy-Driven Computation of Product Recommendations[C]. *The thirteenth*

- ACM International Conference on Information and Knowledge Management*, 2004: 406-415.
- [12] Guan S P, Jin X L, Jia Y T, et al. Knowledge Reasoning over Knowledge Graph: A Survey[J]. *Journal of Software*, 2018, 29(10): 2966-2994.
(官赛萍, 靳小龙, 贾岩涛, 等. 面向知识图谱的知识推理研究进展[J]. *软件学报*, 2018, 29(10): 2966-2994.)
- [13] Zhang Q, Zhang L, Qin C, et al. A Survey on Knowledge Graph-Based Recommender Systems[J]. *Scientia Sinica Informationis*, 2020, 50(7): 937-956.
- [14] Schein A I, Popescul A, Ungar L H, et al. Methods and Metrics for Cold-Start Recommendations[C]. *The 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2002: 253-260.
- [15] Bollacker K, Evans C, Paritosh P, et al. Freebase: A Collaboratively Created Graph Database for Structuring Human Knowledge[C]. *The 2008 ACM SIGMOD International Conference on Management of Data*, 2008: 1247-1250.
- [16] Suchanek F M, Kasneci G, Weikum G. Yago: A Core of Semantic Knowledge[C]. *The 16th International Conference on World Wide Web*, 2007: 697-706.
- [17] Lehmann J, Isele R, Jakob M, et al. DBpedia – a Large-Scale, Multilingual Knowledge Base Extracted from Wikipedia[J]. *Semantic Web*, 2015, 6(2): 167-195.
- [18] Wu W T, Li H S, Wang H X, et al. Probase: A Probabilistic Taxonomy for Text Understanding[C]. *The 2012 ACM SIGMOD International Conference on Management of Data*, 2012: 481-492.
- [19] Xu B, Xu Y, Liang J Q, et al. CN-DBpedia: A Never-Ending Chinese Knowledge Extraction System[M]. *Advances in Artificial Intelligence: From Theory to Practice*. Cham: Springer International Publishing, 2017: 428-438.
- [20] Niu X, Sun X R, Wang H F, et al. Zhishi.me - Weaving Chinese Linking Open Data[M]. *The Semantic Web – ISWC 2011*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011: 205-220.
- [21] Yu X, Ren X, Sun Y Z, et al. Recommendation In Heterogeneous Information Networks with Implicit User Feedback[C]. *The 7th ACM Conference on Recommender Systems*, 2013: 347-350.
- [22] Yu X, Ren X, Sun Y Z, et al. Personalized Entity Recommendation: A Heterogeneous Information Network Approach[C]. *The 7th ACM International Conference on Web Search and Data Mining*, 2014: 283-292.
- [23] A. Bordes, N. Usunier, A. Garcia-Duran, et al. Translating Embeddings for Modeling Multi-Relational Data[C]. *Neural Information Processing Systems*, 2013: 1-9.
- [24] Ji G L, He S Z, Xu L H, et al. Knowledge Graph Embedding via Dynamic Mapping Matrix[C]. *The 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2015: 687-696.
- [25] Che F H, Zhang D W, Tao J H, et al. ParamE: Regarding Neural Network Parameters As Relation Embeddings for Knowledge Graph Completion[J]. *The AAAI Conference on Artificial Intelligence*, 2020, 34(3): 2774-2781.
- [26] Z. Wang, J. Zhang, J. Feng, et al. Knowledge Graph Embedding by Translating on Hyperplanes[C]. *The 28th AAAI Conference on Artificial Intelligence*, 2014, 28(1).
- [27] Perozzi B, Al-Rfou R, Skiena S. DeepWalk: Online Learning of Social Representations[C]. *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2014: 701-710.
- [28] Mikolov T, Chen K, Corrado G, et al. Efficient Estimation of Word Representations In Vector Space[EB/OL]. 2013
- [29] Grover A, Leskovec J. Node2vec: Scalable Feature Learning for Networks[C]. *The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016: 855-864.
- [30] Dong Y X, Chawla N V, Swami A. Metapath2vec: Scalable Representation Learning for Heterogeneous Networks[C]. *The 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017: 135-144.
- [31] Fu T Y, Lee W C, Lei Z. HIN2Vec: Explore Meta-Paths In Heterogeneous Information Networks for Representation Learning[C]. *The 2017 ACM on Conference on Information and Knowledge Management*, 2017: 1797-1806.
- [32] Zhang F Z, Yuan N J, Lian D F, et al. Collaborative Knowledge Base Embedding for Recommender Systems[C]. *The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016: 353-362.
- [33] Wang H W, Zhang F Z, Xie X, et al. DKN: Deep Knowledge-Aware Network for News Recommendation[C]. *The 2018 World Wide Web Conference*, 2018: 1835-1844.
- [34] Kim Y. Convolutional Neural Networks for Sentence Classification[C]. *The 2014 Conference on Empirical Methods in Natural Language Processing*, 2014: 1746-1751.
- [35] Meng W J, Yang D Q, Xiao Y H. Incorporating User Micro-Behaviors and Item Knowledge into Multi-Task Learning for Session-Based Recommendation[C]. *The 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020: 1091-1100.
- [36] Wang C Y, Zhang M, Ma W Z, et al. Make it a Chorus: Knowledge- and Time-Aware Item Modeling for Sequential Recommendation[C]. *The 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020: 109-118.
- [37] Yang D Q, Guo Z K, Wang Z Y, et al. A Knowledge-Enhanced Deep Recommendation Framework Incorporating GAN-Based

- Models[C]. *2018 IEEE International Conference on Data Mining*, 2018: 1368-1373.
- [38] Xin X, He X N, Zhang Y F, et al. Relational Collaborative Filtering: Modeling Multiple Item Relations for Recommendation[C]. *The 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2019: 125-134.
- [39] Yang B, Yih W, He X, et al. Embedding Entities and Relations for Learning and Inference in Knowledge Bases[EB/OL]. 2014: ArXiv preprint ArXiv:1412.6575.
- [40] Wang H W, Zhang F Z, Zhao M, et al. Multi-Task Feature Learning for Knowledge Graph Enhanced Recommendation[C]. *WWW '19: The World Wide Web Conference*, 2019: 2000-2010.
- [41] Zhou G R, Zhu X Q, Song C R, et al. Deep Interest Network for Click-through Rate Prediction[C]. *The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018: 1059-1068.
- [42] Shen C L, Yang D Q, Xiao Y H. A Deep Recommendation Model Incorporating Adaptive Knowledge-Based Representations[C]. *Database Systems for Advanced Applications*, 2019: 481-486.
- [43] Yang D Q, Song Z C, Xue L X, et al. A Knowledge-Enhanced Recommendation Model with Attribute-Level Co-Attention[C]. *The 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020: 1909-1912.
- [44] Guo Z K, Yang D Q, Liu B C, et al. Co-Refining User and Item Representations with Feature-Level Self-Attention for Enhanced Recommendation[C]. *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2020: 494-501.
- [45] Yang D Q, Wang Z Y, Jiang J Y, et al. Knowledge Embedding towards the Recommendation with Sparse User-Item Interactions[C]. *The 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2019: 325-332.
- [46] Cao Y X, Wang X, He X N, et al. Unifying Knowledge Graph Learning and Recommendation: Towards a Better Understanding of User Preferences[C]. *WWW '19: The World Wide Web Conference*, 2019: 151-161.
- [47] Yu X, Ren X, Sun Y Z, et al. Recommendation In Heterogeneous Information Networks with Implicit User Feedback[C]. *The 7th ACM conference on Recommender systems*, 2013: 347-350.
- [48] Sun Y Z, Han J W, Yan X F, et al. Pathsirn: Meta Path-Based Top-k Similarity Search In Heterogeneous Information Networks [J]. *The VLDB Endowment*, 2011, 4(11): 992-1003.
- [49] Yu X, Ren X, Sun Y Z, et al. Personalized Entity Recommendation: A Heterogeneous Information Network Approach[C]. *The 7th ACM International Conference on Web Search and Data Mining*, 2014: 283-292.
- [50] Shi C, Zhang Z Q, Luo P, et al. Semantic Path Based Personalized Recommendation on Weighted Heterogeneous Information Networks[C]. *The 24th ACM International on Conference on Information and Knowledge Management*, 2015: 453-462.
- [51] Zhao H, Yao Q M, Li J D, et al. Meta-Graph Based Recommendation Fusion over Heterogeneous Information Networks[C]. *The 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017: 635-644.
- [52] Xian Y K, Fu Z H, Muthukrishnan S, et al. Reinforcement Knowledge Graph Reasoning for Explainable Recommendation[C]. *The 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2019: 285-294.
- [53] Defferrard M, Bresson X, Vandergheynst P. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering [C]. *Annual Conference on Neural Information Processing Systems*, 2016: 3837-3845.
- [54] Veličković P, Cucurull G, Casanova A, et al. Graph Attention Networks[EB/OL]. 2017
- [55] Wang H W, Zhang F Z, Wang J L, et al. RippleNet: Propagating User Preferences on the Knowledge Graph for Recommender Systems[C]. *The 27th ACM International Conference on Information and Knowledge Management*, 2018: 417-426.
- [56] Wang H W, Zhao M, Xie X, et al. Knowledge Graph Convolutional Networks for Recommender Systems[C]. *WWW '19: The World Wide Web Conference*, 2019: 3307-3313.
- [57] Wang X, He X N, Cao Y X, et al. KGAT: Knowledge Graph Attention Network for Recommendation[C]. *The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019: 950-958.
- [58] Li Q M, Han Z C, Wu X M. Deeper Insights into Graph Convolutional Networks for Semi-Supervised Learning[EB/OL]. 2018
- [59] Chen Q H, Tan H B, Lin G Y, et al. A Hierarchical Knowledge and Interest Propagation Network for Recommender Systems[C]. *2020 International Conference on Data Mining Workshops*, 2020: 119-126.
- [60] Wang X, Wang D X, Xu C R, et al. Explainable Reasoning over Knowledge Graphs for Recommendation[J]. *The AAAI Conference on Artificial Intelligence*, 2019, 33: 5329-5336.
- [61] Joseph K, Jiang H. Content Based News Recommendation via Shortest Entity Distance over Knowledge Graphs[C]. *WWW '19: Companion of the World Wide Web Conference*, 2019: 690-699.
- [62] Sun Z, Yang J, Zhang J, et al. Recurrent Knowledge Graph Embedding for Effective Recommendation[C]. *The 12th ACM Conference on Recommender Systems*, 2018: 297-305.
- [63] Sha X, Sun Z, 0002 J Z. Attentive Knowledge Graph Embedding for Personalized Recommendation [EB/OL]. 2019: ArXiv Preprint ArXiv:1910.08288.
- [64] Zhao J, Zhou Z, Guan Z Y, et al. IntentGC: A Scalable Graph

- Convolution Framework Fusing Heterogeneous Information for Recommendation[C]. *The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019: 2347-2357.
- [65] Gan L, Nurbakova D, Laporte L, et al. Enhancing Recommendation Diversity Using Determinantal Point Processes on Knowledge Graphs[C]. *The 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020: 2001-2004.
- [66] Alex Kulesza, Ben Taskar. k-DPPs: Fixed-size Determinantal Point Processes[C]. *The 28th International Conference on Machine Learning*, 2011: 1193-1200.
- [67] Chen L M, Zhang G X, Zhou H N. Fast Greedy MAP Inference for Determinantal Point Process to Improve Recommendation Diversity[EB/OL]. 2017: arXiv: 1709.05135[cs.IR]. <https://arxiv.org/abs/1709.05135>
- [68] Dive into Deep Learning. <https://d2l.ai/index.html>. 2020.
- [69] Zhang Y F, Chen X. Explainable Recommendation: A Survey and New Perspectives [EB/OL]. 2018: ArXiv Preprint ArXiv:1804.11192v4.
- [70] Ai Q Y, Azizi V, Chen X, et al. Learning Heterogeneous Knowledge Base Embeddings for Explainable Recommendation[J]. *Algorithms*, 2018, 11(9): 137.
- [71] Ma W Z, Zhang M, Cao Y, et al. Jointly Learning Explainable Rules for Recommendation with Knowledge Graph[C]. *WWW'19: The World Wide Web Conference*, 2019: 1210-1221.
- [72] Zhang Y, Xu X R, Zhou H N, et al. Distilling Structured Knowledge into Embeddings for Explainable and Accurate Recommendation[C]. *The 13th International Conference on Web Search and Data Mining*, 2020: 735-743.
- [73] Huang X W, Fang Q, Qian S S, et al. Explainable Interaction-Driven User Modeling over Knowledge Graph for Sequential Recommendation[C]. *The 27th ACM International Conference on Multimedia*, 2019: 548-556.
- [74] Huang J, Zhao W X, Dou H J, et al. Improving Sequential Recommendation with Knowledge-Enhanced Memory Networks[C]. *SIGIR '18: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018: 505-514.
- [75] Miller A, Fisch A, Dodge J, et al. Key-Value Memory Networks for Directly Reading Documents[C]. *The 2016 Conference on Empirical Methods in Natural Language Processing*, 2016: 1400-1409.
- [76] Kaminskas M, Fernández-Tobías I, Cantador I, et al. Ontology-Based Identification of Music for Places[M]. *Information and Communication Technologies in Tourism 2013*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013: 436-447.
- [77] Yang D Q, He J R, Qin H Z, et al. A Graph-Based Recommendation across Heterogeneous Domains[C]. *The 24th ACM International Conference on Information and Knowledge Management*, 2015: 463-472.
- [78] Gabrilovich E, Markovitch S. Wikipedia-Based Semantic Interpretation for Natural Language Processing[J]. *Journal of Artificial Intelligence Research*, 2009, 34: 443-498.
- [79] Dengyong Z, Jason W, Arthur G, et al. Ranking on Data Manifolds[C]. *Neural Information Processing Systems*, 2003: 169-176.
- [80] Jacob D, Ming-Wei C, Kenton L, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[C]. *The 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019: 4171-4186.
- [81] Sun F, Liu J, Wu J, et al. BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer[C]. *The 28th ACM International Conference on Information and Knowledge Management*, 2019: 1441-1450.
- [82] Sun R, Cao X Z, Zhao Y, et al. Multi-Modal Knowledge Graphs for Recommender Systems[C]. *The 29th ACM International Conference on Information & Knowledge Management*, 2020: 1405-1414.
- [83] He K M, Zhang X Y, Ren S Q, et al. Deep Residual Learning for Image Recognition[C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770-778.



阳德青 于 2013 年在复旦大学计算机软件与理论专业获得博士学位。现任复旦大学大数据学院单位副教授。研究领域为人工智能、推荐系统、知识图谱等。研究兴趣包括: 基于知识图谱的推荐、知识抽取等。Email: yangdeqing@fudan.edu.cn



夏西 于 2019 年在上海交通大学园林专业获得学士学位, 现在复旦大学大数据学院应用统计专业攻读专业硕士学位。研究领域为推荐系统。Email: 20210980067@fudan.edu.cn



叶琳 于 2020 年在西安电子科技大学计算机科学与技术专业获得工学学士学位。现在复旦大学大数据学院应用统计专业攻读专业硕士学位。研究领域为推荐系统。
Email: yelin_smile@163.com



薛吕欣 于 2019 年在复旦大学信息与计算科学专业获得理学学士学位。现在复旦大学大数据学院统计学专业攻读硕士学位。研究领域为推荐系统。研究兴趣包括可解释推荐, 图机器学习。Email: lxxue19@fudan.edu.cn



肖仰华 于 2009 年在复旦大学计算机软件与理论专业获得博士学位。现任复旦大学计算机科学技术学院教授, 知识工场实验室主任。研究领域为人工智能、自然语言处理、知识图谱等。研究兴趣包括: 大图数据管理、知识抽取、常识理解等。
Email: shawyh@fudan.edu.cn