

协同过滤推荐系统综述

赵俊逸^{1,2}, 庄福振³, 敖翔^{1,2}, 何清^{1,2}, 蒋慧琴⁴, 马岭⁴

¹郑州大学河南先进技术研究院, 郑州 中国 450003

²中国科学院计算技术研究所智能信息处理重点实验室 北京 中国 100190

³北京航空航天大学人工智能研究院 北京 中国 100191

⁴郑州大学信息工程学院 郑州 中国 450003

摘要 随着互联网和信息计算的飞速发展, 衍生了海量数据, 我们已经进入信息爆炸的时代。网络中各种信息量的指数型增长导致用户想要从大量信息中找到自己需要的信息变得越来越困难, 信息过载问题日益突出。推荐系统在缓解信息过载问题中起着非常重要的作用, 该方法通过研究用户的兴趣偏好进行个性化计算, 由系统发现用户兴趣进而引导用户发现自己的信息需求。目前, 推荐系统已经成为产业界和学术界关注、研究的热点问题, 应用领域十分广泛。在电子商务、会话推荐、文章推荐、智慧医疗等多个领域都有所应用。传统的推荐算法主要包括基于内容的推荐、协同过滤推荐以及混合推荐。其中, 协同过滤推荐是推荐系统中应用最广泛最成功的技术之一。该方法利用用户或物品间的相似度以及历史行为数据对目标用户进行推荐, 因此存在用户冷启动和项目冷启动问题。此外, 随着信息量的急剧增长, 传统协同过滤推荐系统面对数据的快速增长会遇到严重的数据稀疏性问题以及可扩展性问题。为了缓解甚至解决这些问题, 推荐系统研究人员进行了大量的工作。近年来, 为了提高推荐效果、提升用户满意度, 学者们开始关注推荐系统的多样性问题以及可解释性问题。由于深度学习方法可以通过发现数据中用户和项目之间的非线性关系从而学习一个有效的特征表示, 因此越来越受到推荐系统研究人员的关注。目前的工作主要是利用评分数据、社交网络信息以及其他领域信息等辅助信息, 结合深度学习、数据挖掘等技术提高推荐效果、提升用户满意度。对此, 本文首先对推荐系统以及传统推荐算法进行概述, 然后重点介绍协同过滤推荐算法的相关工作。包括协同过滤推荐算法的任务、评价指标、常用数据集以及学者们在解决协同过滤算法存在的问题时所做的工作以及努力。最后提出未来的几个可研究方向。

关键词 推荐系统; 协同过滤; 稀疏性; 深度学习

中图分类号 TP391 DOI号 10.19363/J.cnki.cn10-1380/tn.2021.09.02

Survey of Collaborative Filtering Recommender Systems

ZHAO Junyi^{1,2}, ZHUANG Fuzhen³, AO Xiang^{1,2}, HE Qing^{1,2}, JIANG Huiqin⁴, MA Ling⁴

¹Henan Institute of Advanced Technology, Zhengzhou University, Zhengzhou 450003, China

²Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China

³Institute of Artificial Intelligence, Beihang University, Beijing 100191, China

⁴School of Information Engineering, Zhengzhou University, Zhengzhou 450003, China

Abstract With the rapid development of Internet and information computing, we have entered the era of information explosion. The exponential growth of all kinds of information in the network makes it more and more difficult for users to find the information they need from a large amount of information, and the problem of information overload is becoming increasingly prominent. Recommender system plays a very important role in alleviating the problem of information overload. This method studies the user's interest preference for personalized calculation, and the system finds the user's interest, and then guides the user to find their own information needs. At present, recommender system has become a hot issue in industry and academia. It has a wide range of applications, such as e-commerce, conversation recommendation, article recommendation, intelligent medical and so on. Traditional recommendation algorithms mainly include content-based recommendation, collaborative filtering recommendation and hybrid recommendation. Among them, collaborative filtering recommendation is one of the most widely used and successful technologies in recommender system. This method uses the similarity between users or items and historical behavior data to recommend target users, so there are problems of user cold start and project cold start. In addition, with the rapid growth of information, traditional collaborative filtering recommender system will encounter serious data sparsity and scalability problems in the face of the rapid growth of data. In order to alleviate or even solve these problems, recommender system researchers have done a lot of work. In recent years, in order to improve the recommendation effect and user satisfaction, scholars began to pay attention to the diversity and

通讯作者: 何清, 博士, 研究员, Email: heqing@ict.ac.cn

本文得到国家重点研发计划课题(No. 2017YFB1002104), 国家自然科学基金(No. 61976204, No. U1811461), 中科院青年创新促进会支持

收稿日期: 2021-04-30; 修改日期: 2021-08-08; 定稿日期: 2021-08-10

interpretability of recommender system. Because deep learning method can learn an effective feature representation by discovering the nonlinear relationship between users and items in the data, it has attracted more and more attention of recommender system researchers. The current work is mainly to use scoring data, social network information and other areas of information and other auxiliary information, combined with deep learning, data mining and other technologies to improve the recommendation effect and improve user satisfaction. In this regard, this paper first summarizes the recommender system and traditional recommendation algorithm, and then focuses on the related work of collaborative filtering recommendation algorithm. It includes the task of collaborative filtering recommendation algorithm, evaluation index, common data sets and the work and efforts of scholars in solving the problems of collaborative filtering algorithm. Finally, several possible research directions in the future are proposed.

Key words recommender system; collaborative filtering; sparsity; deep learning

1 引言

互联网和信息计算的快速发展, 衍生了海量数据, 我们已经进入信息爆炸的时代。然而这些信息并不完全是人们所关心的, 用户要从大量信息中寻找有用信息变得越来越困难。另一方面, 信息的生产方在绞尽脑汁把用户感兴趣的信息推送给用户, 而每个人的兴趣又不尽相同, 所以可以实现千人千面的推荐系统应运而生。简单来说, 推荐系统是根据用户的浏览习惯确定用户兴趣, 通过发掘用户的行为, 将合适的信息推荐给用户, 满足用户的个性化需求。推荐系统的结构通常由在线和离线两部分组成。在线部分即前端展示部分; 离线部分由后端的日志系统和推荐算法系统组成。离线部分通过学习用户资料和行为日志建立模型, 在新的上下文背景中计算相应的推荐内容, 将其呈现于在线页面中, 如图 1 所示。目前, 推荐系统已经成为产业界和学术界关注、研究的热点问题, 应用领域十分广泛, 在电子商务^[1]、会话推荐^[2]、文章推荐^[3]、智慧医疗^[4]等领域都有所应用。

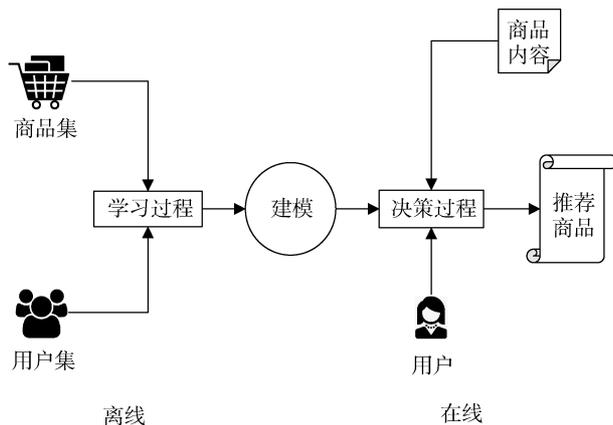


图 1 推荐系统的推荐过程

Figure 1 Recommendation process of recommender system

传统的推荐算法主要包括基于内容的推荐(Content-based Recommendation)^[3]、协同过滤推荐

(Collaborative Filtering-based Recommendation, CFR)^[5]和混合推荐(Hybrid Recommendation)^[6]。

(1) 基于内容的推荐。该推荐方法主要利用项目内容进行推荐, 是信息过滤技术的延续与发展。其通过机器学习方法从内容的特征描述中得到用户的偏好信息而不需要依据用户对项目的评价意见。

基于内容的推荐过程一般为: 首先抽取表示该项目的项目特征; 然后利用用户过去感兴趣以及不感兴趣的项目特征数据学习该用户的偏好特征; 最后通过比较用户的偏好特征和候选项目特征为用户推荐一组相关性最大的项目^[3]。

(2) 协同过滤推荐。该方法假设拥有相似兴趣的用户可能会喜欢相似的项目或者用户可能对相似的项目表现相似的偏好程度。其核心思想是基于近邻的推荐算法, 利用用户或物品间的相似度以及历史行为数据以对目标用户进行有效推荐^[5]。

(3) 混合推荐。该方法将多种推荐技术进行混合, 使相互弥补缺点获得更好的推荐效果。最常见的是将协同过滤推荐与内容推荐相结合^[6]。

表 1 为上述三类推荐方法的优缺点对比。可以发现, 协同过滤推荐与基于内容的推荐相比, 其优点在于: 首先, 可以帮助用户发现新的兴趣, 协同过滤可以发现与用户已知兴趣不同的潜在兴趣偏好; 其次, 通过共享他人经验, 可避免内容分析的不完全和不准确, 并且能够基于复杂概念(如个人品味)进行过滤; 最后, 能够通过使用其他相似用户的反馈信息加快个性化学习速度。因此, 协同过滤成为目前应用最广泛最成功的推荐技术之一。

然而随着时代的发展, 各类应用及服务出现井喷式增长, 传统协同过滤推荐系统面对数据的快速增长逐渐暴露出一些问题, 如冷启动、数据稀疏、可扩展性、多样性以及可解释性等问题。这些问题极大地影响了推荐系统的性能, 给推荐系统的发展带来挑战。因此, 如何缓解甚至解决这些问题以提高推荐系统的推荐效果, 对用户和企业来说都具有重要意义, 学者们为此进行了大量研究。

表 1 推荐算法优缺点对比

Table 1 Comparison of advantages and disadvantages of recommendation algorithms

推荐方法	优点	缺点
基于内容的推荐	不需要其他用户数据; 能推荐新项目及非流行项目; 推荐结果直观; 无需领域知识 新异兴趣发现;	稀疏性问题; 用户冷启动问题; 复杂属性不好处理; 需要足够数据构造分类器
协同过滤推荐	无需领域知识; 推荐个性化、自动化程度高; 能处理复杂的非结构化对象	稀疏性问题; 冷启动问题; 可扩展性问题; 质量取决于历史数据集;
混合推荐	弥补单一推荐技术的缺点	实现难度高; 执行效率低

本文首先对协同过滤推荐算法的相关技术及评价指标进行梳理与对比分析, 随后详细介绍学者们在解决传统协同过滤推荐算法存在的各种问题及挑战时做出的努力, 最后提出未来的几个发展方向。

2 基于协同过滤的推荐

协同过滤推荐系统通过计算系统中用户之间的相似度, 根据他们之间的相似模式来预测项目。该方法利用“用户-项目评分矩阵”进行推荐。如图 2 所示, 在推荐系统中根据用户对项目的评分可以得到“用户-项目评分矩阵”, 其中矩阵中的值表示用户对项目的评分, “-”表示用户未对项目进行评分。实际场景中通常只存在用户对少量项目的评分, 因此矩阵往往非常稀疏。

User/Item	1	2	3	4	5	...
1	5	-	4.5	-	-	...
2	-	4.5		-	5	...
3	4.2	-	-	4.8	-	...
4	-	-	4.5	5	-	...
5	-	3	4	-	-	...
...

图 2 用户-项目评分矩阵

Figure 2 User-Item rating matrix

2.1 协同过滤推荐过程

协同过滤主要根据相似用户之间的现有评分为用户推荐项目。如图 3 所示, 主要分为以下两步^[7]:

(1) 系统将整个空间表示为二维矩阵。在二维矩阵 R 中, U 和 I 分别对应用户和项目。 R_{ij} 是用户 U_i 对项目 I_j 的评分。如果用户未对项目评分则项目值设为 0。

(2) 使用矩阵 R 预测用户 i 对项目 j 的评分, 以

便为用户可能喜欢的 N 个项目做出推荐。

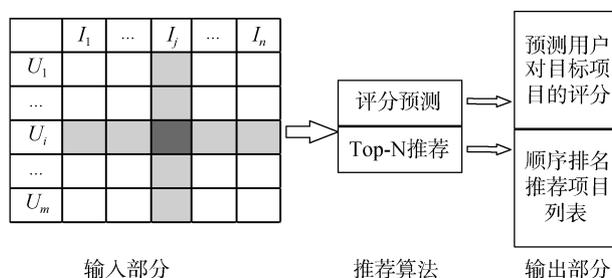


图 3 协同过滤推荐的一般过程

Figure 3 General process of collaborative filtering recommendation

2.2 协同过滤方法

有两种类型的协同过滤方法被广泛研究, 即基于内存的协同过滤(Memory-based CF)和基于模型的协同过滤(Model-based CF)。

2.2.1 基于内存的协同过滤

基于内存的协同过滤方法一般采用最近邻技术, 利用用户的历史喜好信息计算用户之间的距离, 然后利用目标用户的邻居用户对商品评价的加权值来预测目标用户对特定商品的喜好程度, 推荐系统根据喜好程度对目标用户进行推荐。

该方法包括基于用户的方法(User-based CF)和基于项目的方法(Item-based CF)。基于用户的方法根据用户之间的评分行为相似度预测用户的评分, 如图 4 所示; 基于项目的方法根据预测项目与用户实际选择项目之间的相似度预测用户评分, 如图 5 所示。由于这两种方法的原理不同, 因此在不同应用场景中的表现不同。基于用户的推荐更社会化, 反映用户所在的兴趣群体中物品的热门程度; 基于项目的推荐更个性化, 反映用户自己的兴趣传承。具体对比如表 2 所示。

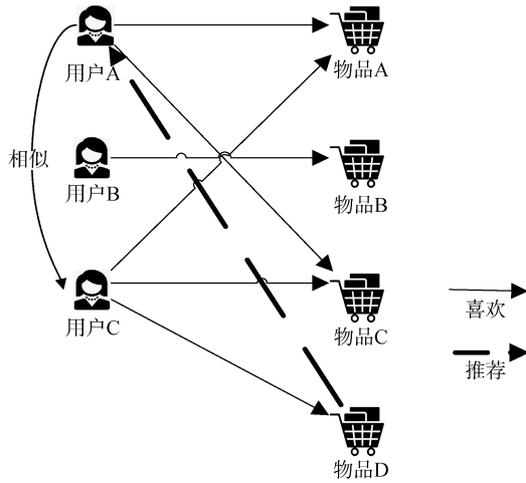


图 4 基于用户的协同过滤原理
Figure 4 User-based CF

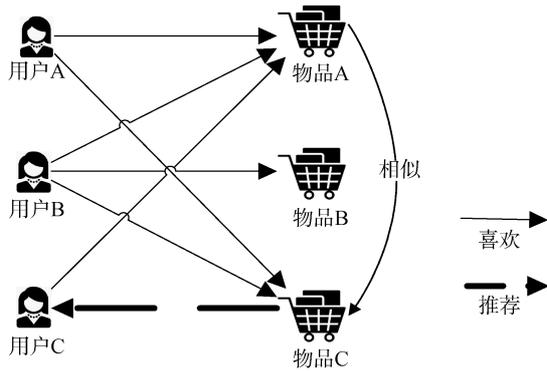


图 5 基于项目的协同过滤原理
Figure 5 Item-based CF

在基于内存的协同过滤技术中, 计算相似度的常用方法包括皮尔森(Pearson)相关系数、余弦相似度、杰卡德(Jaccard)相似系数以及欧氏距离。这些相似度计算方法没有绝对的优劣之分, 适用场景也并非绝对。因此应根据实际的应用场景、数据特点灵活选择或进行组合。

(1) 皮尔森相关系数。皮尔森相关系数用于度量任意两变量间线性相关的程度, 系数值越大表明两者相关性越强, 系数的输出范围为 $[-1, 1]$ 。0 代表两者无相关性, 负值为负相关, 正值为正相关。两者的相似性可表示为:

$$\text{sim}(i, j) = \frac{\sum_{x \in I_{ij}} (R_{i,x} - \bar{R}_i)(R_{j,x} - \bar{R}_j)}{\sqrt{\sum_{x \in I_{ij}} (R_{i,x} - \bar{R}_i)^2 \sum_{x \in I_{ij}} (R_{j,x} - \bar{R}_j)^2}} \quad (1)$$

其中, I_{ij} 表示用户 i 和用户 j 共同评分过的项目集合, $R_{i,x}$ 和 $R_{j,x}$ 分别表示用户 i 和用户 j 对项目 x 的评分, \bar{R}_i 和 \bar{R}_j 分别表示用户 i 和用户 j 对所有项目评分的平均值。

(2) 余弦相似度。余弦相似度通过计算两个向量的夹角余弦值来评估它们的相似度。输出范围和皮尔森相关性系数一致, 含义也相似。公式为:

$$\text{sim}(x, y) = \frac{\sum_{i=1}^n R_{x,i} R_{y,i}}{\sqrt{\sum_{i=1}^n R_{x,i}^2} \sqrt{\sum_{i=1}^n R_{y,i}^2}} \quad (2)$$

表 2 基于内存的协同过滤方法对比

Table 2 Comparison of Memory-based CF

	基于用户的协同过滤	基于项目的协同过滤
性能	适用于用户较少的场景	适用于项目数远小于用户数的场景
实时性	用户有新行为时不一定造成推荐结果的立即变化	用户有新行为时一定会导致推荐结果的实时变化
冷启动	存在用户冷启动问题	存在项目冷启动问题
适用领域	实时性较强, 用户个性化兴趣不太明显的领域。 例如: 新闻网站	长尾物品丰富, 用户个性化需求强烈的领域。 例如: 电子商务、电影网站

(3) 修正的余弦相似度。修正的余弦相似度利用减去用户对项目的平均评分的方式改善余弦相似度中未考虑不同用户对项目的评分标准存在差异的问题。公式为:

$$\text{sim}(x, y) = \frac{\sum_{i \in I_{xy}} (R_{x,i} - \bar{R}_X)(R_{y,i} - \bar{R}_Y)}{\sqrt{\sum_{i \in I_{xy}} (R_{x,i} - \bar{R}_X)^2} \sqrt{\sum_{i \in I_{xy}} (R_{y,i} - \bar{R}_Y)^2}} \quad (3)$$

其中, I_{xy} 表示用户 x 和用户 y 共同评分过的项目集合, $R_{x,i}$ 和 $R_{y,i}$ 分别表示用户 x 和用户 y 对项目 i 的评

分, \bar{R}_X 和 \bar{R}_Y 分别表示用户 x 和用户 y 对所有项目评分的平均值。

(4) 杰卡德相似系数。杰卡德相似系数是衡量两个集合之间相似度的一种指标。用户 u 和用户 v 共同评分过的项目集合的数量与他们评分过的总项目集合的数量之间的比例称为杰卡德相似系数。由于该方法无法反映用户的具体评分偏好信息, 因此通常用于评估用户是否会对项目进行评分而不是预测用户对项目的具体评分。公式为:

$$\text{sim}(u, v) = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|} \quad (4)$$

其中 $N(u)$, $N(v)$ 分别表示用户 u 和用户 v 评分过的项目集合。

(5) 欧氏距离。对于用户 u_i 和 u_j , 欧氏距离即计算这两个用户的评分向量在 n 维向量空间中的绝对距离, 公式如下:

$$d(u_i, u_j) = \sqrt{\sum (\bar{x}_i - \bar{x}_j)^2} \in [0, \infty) \quad (5)$$

在实际使用中, 通常还需要将欧氏距离进行归一化处理, 可以得出两者之间的相似性为:

$$\text{sim}(u_i, u_j) = \frac{1}{1 + d(u_i, u_j)} \in (0, 1] \quad (6)$$

2.2.2 基于模型的协同过滤

基于模型的协同过滤方法通过建模的方式模拟用户对项目的评分行为。其使用机器学习与数据挖掘技术, 从训练数据中确定模型并将模型用于预测未知商品评分。常见模型包括聚类模型^[8-10]、贝叶斯模型^[11]、矩阵分解^[12]等。

(1) 聚类。聚类算法基于预定义模型构造数据, 是一种无监督机器学习算法。按照聚类对象的不同可以分为用户聚类模型、项目聚类模型以及用户-项目联合聚类模型。用户聚类模型^[13]即将兴趣相近的用户集聚成一簇; 项目聚类模型^[14-15]即将相似项目进行聚类; 还可以同时考虑用户和项目的聚类结果进行综合分析。聚类结果应该具有较高的簇内相似性和较低的簇间相似性。

通常使用闵可夫斯基距离(Minkowski Distance)和皮尔森相关性计算数据对象 $X=(x_1, x_2, x_3, \dots, x_n)$ 和 $Y=(y_1, y_2, y_3, \dots, y_n)$ 之间的相似度。闵可夫斯基距离定义如下:

$$d(X, Y) = \sqrt[q]{\sum_{i=1}^n |x_i - y_i|^q} \quad (7)$$

其中 n 是对象的维数, x_i 和 y_i 分别是对象 x 和对象 y 第 i 维度上的值。 q 是一个正整数。当 $q=1$ 时该距离称为曼哈顿距离(Manhattan Distance); 当 $q=2$ 该距离称为欧氏距离。

最常用的聚类算法是 K-means。该算法实现起来相对简单且具有很高的推荐精度。

(2) 贝叶斯模型。由于推荐问题可以转化为分类问题, 因此可以通过机器学习领域中的分类算法来解决推荐问题。贝叶斯网络基于条件概率和贝叶斯定理。该网络是最流行的概率框架, 能够推导出用户或项目之间的概率依赖关系^[11]。其使用决策树表示概率表, 对每个推荐项都构造一个单独的树。树中节

点的分支选择取决于用户对特定项目的评分。用户对预测项目评分的概率向量存储在树的节点中。

(3) 关联规则挖掘。关联规则挖掘通常基于用户购买商品之间的关联关系预测用户偏好。具体来说, 关联规则基于物品维度构思考虑, 分析出数据库中所有包含属性等特征信息模块间的隐藏关联, 进而找出满足给定项集支持度和项集置信度的多个模块间存在的所有依赖关系^[16]。根据关联规则推荐原理产生的推荐算法步骤如图 6 所示。

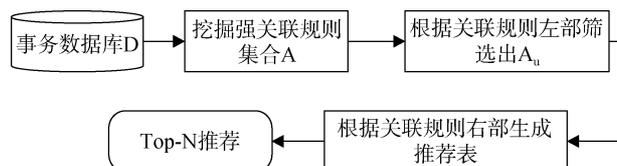


图 6 关联规则推荐步骤

Figure 6 Recommendation steps of association rules

(4) 矩阵分解^[12]。矩阵分解是推荐系统中最常用的协同过滤模型之一。该模型利用用户-项目评分矩阵预测用户对项目的评分, 通过学习用户潜在向量 U 和项目的潜在向量 V , 使 U 和 V 内积近似于用户真实评分 X , 得到预测的评分 R 可以写成:

$$r_{ij} = u_i^T v_j \quad (8)$$

其中 i 是用户索引, j 是项目索引, r_{ij} 表示用户 i 对项目 j 的预测评分, u 是用户的潜在向量, v 是项目的潜在向量。然后计算其损失函数为:

$$L_{MF} = \|X - R\|^2 + \lambda(\|U\|_F^2 + \|V\|_F^2) \quad (9)$$

其中 $\|\cdot\|_F^2$ 是第二范数。 λ 是正则化项。

(5) 马尔可夫决策过程(Markov decision processes, MDP)。基于 MDP 的协同过滤算法将推荐转化为一种序列优化问题, 使用 MDP 生成推荐。MDP 的关键优势在于它们考虑每个推荐的长期影响和每个推荐的算术平均值。MDP 可以被定义为一个四元组: $\langle S, A, R, P_r \rangle$, 其中 S 表示状态集合, A 表示行为集合; R 表示每个“状态/行为”对的奖励函数, P_r 表示给定任一行行为的每一状态对之间的转移概率^[17]。

表 3 为基于内存的协同过滤方法与基于模型的协同过滤方法的优缺点对比。通过比较可以看出, 基于模型的协同过滤方法能更好的解决稀疏性、可扩展性等问题。

2.3 协同过滤任务

基于协同过滤的推荐系统可以根据任务的不同分为评分预测和 Top-N 推荐。

表3 协同过滤方法的优缺点比较

Table 3 Comparison of advantages and disadvantages of collaborative filtering methods

协同过滤方法	优点	缺点
基于内存的协同过滤	容易实施; 容易添加数据; 无需考虑项目内容	依赖评分数据; 数据稀疏时性能下降; 冷启动问题; 可扩展性有限
基于模型的协同过滤	易于解决稀疏性、可扩展性和其他问题; 提高预测性能; 推荐结果直观	建立模型的成本高; 需要在预测性能和可扩展性之间权衡; 信息降维导致有用信息丢失

(1) 评分预测。该任务的主要应用场景为评价网站,如豆瓣网站。通过建模用户对项目的历史评分记录得到用户的兴趣模型,进而预测用户对未评分项目的评分。最常用的技术是矩阵分解技术。

(2) Top-N 推荐。该任务的主要应用场景为购物网站或一般拿不到显式评分信息的网站,如亚马逊网站、淘宝网站。通过用户的隐式反馈信息为其推荐可能感兴趣的前 N 个商品。最常用的技术为贝叶斯个性化排名(Bayesian Personalized Ranking, BPR)。

2.4 评价指标

不同的推荐任务使用的评价指标往往也不同,评分预测任务通常需要预测准确度,Top-N 推荐任务通常需计算查准率、查全率等准确度指标以及覆盖度、多样性等非准确度指标。

2.4.1 评分准确度

该方法用于衡量算法预测的评分与用户的实际评分之间的贴近程度。令 r_{ua} 表示用户 u 对项目 α 的真实评分, r'_{ua} 表示用户 u 对项目 α 的预测评分, E^P 表示测试集。主要的评价指标为:

(1) 平均绝对误差(Mean Absolute Error, MAE)

该方法因计算简单、通俗易懂被广泛应用。其通过计算预测评分与真实评分之间的绝对误差得到,公式为:

$$MAE = \frac{1}{|E^P|} \sum_{(u,\alpha) \in E^P} |r_{ua} - r'_{ua}| \quad (10)$$

(2) 均方根误差(Root Mean Square Error, RMSE)

该方法加大了对用户项目评分的错误预测惩罚,因而对系统的评测更加苛刻。公式为:

$$RMSE = \sqrt{\frac{1}{|E^P|} \sum_{(u,\alpha) \in E^P} (r_{ua} - r'_{ua})^2} \quad (11)$$

2.4.2 排名准确度

令 $R(u)$ 表示模型根据用户在训练集上的行为计算出的推荐项目列表, $T(u)$ 表示用户在测试集上的真实喜爱列表,主要评测方式为:

(1) 查准率。该方法可得到分类正确的正样本个数占分类器判定为正样本的样本个数的比例。用于体现所预测的推荐列表中有多少物品是用户感兴趣的。公式如下:

$$Precision = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|} \quad (12)$$

(2) 召回率。该方法用于体现用户真实喜爱列表中的物品有多少是被推荐算法预测出的。公式如下:

$$Recall = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |T(u)|} \quad (13)$$

(3) F1 分数。该方法能够兼顾精确率和召回率,可以看作是模型精确率和召回率的一种加权平均。公式如下:

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (14)$$

(4) 折损累计增益(Discounted Cumulative Gain, DCG)。目的在于让排名越靠前的结果越能影响最后的结果。公式为:

$$DCG_p = \sum_{i=1}^p \frac{2^{rel_i} - 1}{\log_2(i+1)} \quad (15)$$

其中, rel_i 表示 i 位置上的相关度。

(5) 归一化折损累计增益(Normalized Discounted cumulative gain, NDCG)。由于搜索结果随着检索词的不同,返回的数量是不一致的,而 DCG 是一个累加的值,无法对两个不同的搜索结果进行比较,因此需要归一化处理。公式如下:

$$NDCG_p = \frac{DCG_p}{IDCG_p} \quad (16)$$

IDCG_p 为理想情况下的最大 DCG 值,即:

$$IDCG_p = \sum_{i=1}^{|\text{REL}|} \frac{2^{rel_i} - 1}{\log_2(i+1)} \quad (17)$$

其中, |REL| 表示结果按照相关性从大到小排序后,前 p 个结果组成的集合。

2.4.3 覆盖度

覆盖度标识推荐系统对物品长尾的发掘能力,用以衡量系统是否能推荐到所有物品。设现有用户集合为 U , 系统为每个用户 u 提供 Top-N 推荐列表为 $R(u)$, 则覆盖度为:

$$\text{Coverage} = \frac{|U_{u \in U} R(u)|}{|I|} \quad (18)$$

2.4.4 多样性

为提高用户满意度,推荐应覆盖用户的所有兴趣点,甚至包括用户未发觉的潜在兴趣,体现推荐多样性。设 $\text{sim}(i,j) \in [0,1]$ 为物品 i,j 的相似性。则用户 u 的推荐列表 $R(u)$ 的多样性可定义为:

$$\text{Diversity}(R(u)) = \frac{\sum_{i,j \in R(u), i \neq j} (1 - \text{sim}(i,j))}{\frac{1}{2} |R(u)| (|R(u)| - 1)} \quad (19)$$

2.4.5 新颖度

新颖度指的是向用户推荐他并不知道物品。平均流行度是评估新颖度的最简便方法。该方法为推荐列表中所有物品的度的总和平均值。推荐列表中物品的平均流行度越小,新颖度越高。

2.4.6 健壮性

健壮性指的是系统抵御作弊的能力,学者在研究推荐算法时可以利用算法对用户生成一个推荐列

表,然后往数据集中注入一定的噪声数据,然后再次生成推荐列表,最后对比两次推荐列表的相似度,以此评估算法的健壮性。

2.5 常用数据集

推荐系统中常用的数据集主要有 MovieLens^①、Jester^②、Book-Crossings^③、Last.fm^④、Wikipedia^⑤。具体内容如表 4 所示。

其中,MovieLens 是电影评分的集合,有 3 种大小的数据集;Jester 包含 150 个笑话,大约 600 万评分;Book-Crossings 是关于图书评分的数据集;Last.fm 提供音乐推荐的数据集,包含用户最喜爱的艺术家列表及其播放次数;Wikipedia 是其用户撰写的百科全书,被广泛用于社交网络分析、图像和数据库实现测试以及用户行为研究。

3 国内外研究现状

随着数据的指数级增长,传统协同过滤推荐系统的问题越来越明显,主要包括:冷启动与稀疏性问题、可扩展性问题、多样性问题以及可解释性问题。为了应对这些问题与挑战,研究人员们做出了很多努力。本节将详细介绍学者们在解决协同过滤推荐算法相关问题与挑战时所做的工作。表 5 为相关工作汇总。

表 4 常用数据集介绍

Table 4 An introduction to common datasets

数据集	用户数	项目数	评分数	密集程度(%)
MovieLens 1M	6040	3883	1000209	4.26
MovieLens 10M	69878	10681	10000054	1.33
MovieLens 20M	138493	27278	20000263	0.52
Jester	124113	150	5865235	31.5
Book-Crossing	92107	271379	1031175	0.0041
Last.fm	1892	17632	92834	0.28
Wikipedia	5583724	4936761	417996366	0.0015

3.1 冷启动与稀疏性问题

在协同过滤方法中,项目的推荐基于用户过去的偏好,因此当出现新用户或新项目的时候无法做出推荐。该问题被称为冷启动问题。此外,由于电子

商务规模的扩大,用户数据和项目数据急剧增加,而用户评价过的项目或用户间重叠的项目数量过少,使得用户-项目矩阵出现极端稀疏性,导致推荐效果不理想。

① <https://grouplens.org/datasets/movielens/>

② <http://eigentaste.berkeley.edu/dataset/>

③ <http://www2.informatik.uni-freiburg.de/~ctiegl/BX/>

④ <https://grouplens.org/datasets/hetrec-2011/>

⑤ https://en.wikipedia.org/wiki/Wikipedia:Database_download#English-language_Wikipedia

表5 协同过滤算法工作梳理
Table 5 Review of collaborative filtering algorithm

协同过滤挑战	相关文章
评分数据	文献[1]、文献[18]、文献[19]、文献[20]、文献[21]、文献[22]、文献[23]、文献[24]、文献[25]、文献[26]、文献[27]、文献[28]、文献[29]
隐语义模型	文献[8]、文献[13]、文献[30]、文献[31]、文献[32]、文献[33]、文献[34]
神经网络	文献[35]、文献[36]、文献[37]、文献[38]、文献[39]、文献[40]、文献[41]、文献[42]
冷启动与稀疏性	辅助信息 注意力机制 文献[50]、文献[51]、文献[52] 社交网络 文献[44]、文献[45]、文献[46]、文献[47]、文献[48]、文献[49] 异质信息网络 文献[53]、文献[54]、文献[55]、文献[56]、文献[57]、文献[58]、文献[59]
迁移协同过滤	单域迁移 迁移知识 文献[60]、文献[61]、文献[62]、文献[63] 聚合知识 文献[64]、文献[65]、文献[66]、文献[67]、文献[68]、文献[69]、文献[70]、文献[75] 多域迁移 文献[71]、文献[72]、文献[73]、文献[74]
主成分分析	文献[5]、文献[76]
可扩展性	用户聚类 文献[9]、文献[10]、文献[77]、文献[78]、文献[79]、文献[80] 聚类 项目聚类 文献[4]、文献[14]、文献[15]、文献[81]
多样性	文献[82]、文献[83]、文献[84]、文献[85]、文献[86]、文献[87]、文献[88]
可解释性	文献[2]、文献[89]、文献[90]、文献[91]、文献[92]、文献[93]、文献[94]、文献[95]、文献[96]、文献[97]、文献[98]、文献[99]、文献[100]

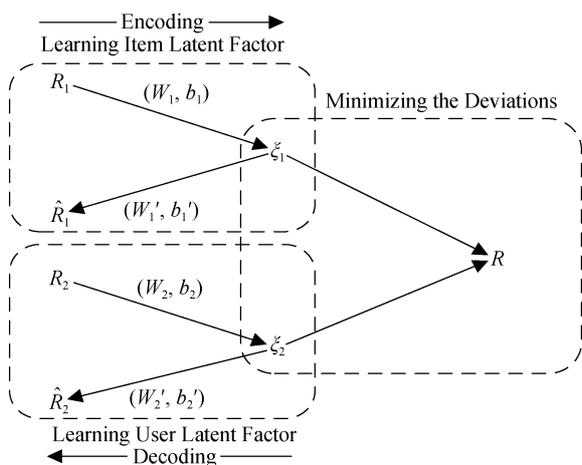
协同过滤方法可以根据它们使用的信息源分为三条通路。开创性的工作是仅利用评分数据或签到数据作为单一来源用于推荐; 为了提高推荐性能, 随后一些算法考虑了额外的信息(如用户的社交网络或商品的属性); 第三是迁移协同过滤(Transfer Collaborative Filtering, TCF), 它尝试从源域学习知识应用到目标域以实现更好推荐。

3.1.1 利用评分数据

矩阵分解技术是协同过滤最流行的模型, 其将一个用户-项目评分矩阵分解为两个低秩的用户特定矩阵和项目特定矩阵, 然后利用分解后的矩阵进行进一步的预测。非负矩阵分解(Non-negative Matrix Factorization, NMF)^[18]是一种有效的多元数据分解方法, 可以将评分矩阵分解为用户和项目信息。Paterek^[19]提出改进正则奇异值分解来预测用户对项目的偏好。Salakhutdinov 和 Mnij 进一步提出概率矩阵分解(Probabilistic Matrix Factorization, PMF)^[20]和贝叶斯概率矩阵分解(Bayesian Probabilistic Matrix Factorization, BPMF)^[21]。其中, PMF 使用一种带有高斯观测噪声的概率线性模型来表示用户和物品的潜在特征, 该模型在大且稀疏的评分数据上表现良好。BPMF 通过综合所有模型参数和超参数自动控制模型容量。Zhou 等^[22]将矩阵分解与决策树算法相结合, 提出功能矩阵分解模型(Functional Matrix Factorization, FMF), 在冷启动过程中为用户选择合适物品, 通过对物品打分子学习用户的偏好。Gu 等^[23]考虑了相

似用户分布与相似项目分布之间的关系, 即用户部分和项目部分对推荐结果的贡献取决于它们的相似度, 对此提出一种动态加权协同过滤方法(Dynamic-weighted Collaborative Filtering approach, DWCF)来解决稀疏性问题。

随着深度学习的发展, 学者们将深度学习技术用于评分矩阵中以提高推荐性能。Salakhutdinov 等人^[24]首次使用 RBM 对数据进行建模。Phung 等人^[25]探索并扩展了用于协同过滤任务的玻尔兹曼机, 其以原则性的方式集成了相似性和共现性。Zhuang 等^[26]提出一种基于自动编码器的表示学习框架(Recommendation via Dual-Autoencoder, ReDa)。如图 7 所示, 该框架将深度学习直接用于评分信息, 可以通过广泛利用评分信息获得更好的用户和项目的隐语义。具体来说, ReDa 使用堆叠自动编码器(Stacked Autoencoders, SAE)同时分别学习用户和项目的潜在表示, 并利用学习到的表示使训练数据的误差最小化。Zhuang 等^[1]认为推荐系统的主要目标是预测用户对一些他们之前没有注意到的商品的兴趣, 在训练已知数据用于预测未知数据时不仅要保持已知数据值的一致性, 还要保持值排序的一致性。同时考虑值和值排序可能比只关注一方更为合理, 也更可能具有高准确性。因此, 在 ReDa 基础上考虑了用户-项目对的成对排序损失, 提出一种基于成对约束表示学习的协同排序框架(Representation Learning with Pair-wise Constraints, REAP), 进一步提升推荐系统的性能。

图7 ReDa 框架(R 为评分矩阵)Figure 7 ReDa Framework (R is the rating matrix)

Lee 等^[27]考虑到以往采用负采样可能将实际为正但却未观察到的信息作为负样本这一局限性, 提出 BUIR 框架。该框架不需要负采样而是采用两个相互学习的不同编码网络来保证正相关的用户表示和项目表示彼此相似, 并将随机数据增强应用于编码器输入中以缓解数据稀疏问题。Zhu 等^[28]考虑了解决项目冷启动时存在的偏向问题, 提出一种可学习的后处理框架作为增强公平性的模型蓝图, 在此基础上提出联合学习生成模型和分数缩放模型, 使模型在保持推荐效用的同时提高公平性。Lin 等^[29]为解决用户冷启动问题提出一种新的元学习推荐器, 即任务自适应神经过程 (Task-adaptive Neural Process, TaNP)。该模型将观察到的每个用户的交互直接映射到预测分布中, 避免了基于梯度的元学习模型的一些训练问题。此外还引入任务自适应机制以平衡模型容量与自适应可靠性之间的平衡。

3.1.2 添加辅助信息

通过在协同过滤方法中加入辅助信息能够进一步缓解冷启动与稀疏性问题, 提高推荐性能。

Vandenoord 等人^[30]将主题模型引入到学习过程中, 从文本信息中学习隐语义; McAuley 和 Leskovec^[31]提出了一种利用主题模型来理解评论文本评分的方法。在文献^[32]中进一步利用评分和评论, 使主题与评级维度保持一致, 以提高预测准确性。Ren 等人^[33]提出了一个潜在变量模型, 该模型可以根据用户意见和社会关系生成解释并预测项目评分。Rodrigues^[8]提出一种将基于项目的协同过滤与用户人口统计信息结合到聚类加权机制中的框架, 能为新用户提供良好的推荐服务。Ji 等^[34]提出一种基于矩阵分解的可伸缩协同过滤算法, 利用用户类别和用户关键字这两个决策矩阵代替单一的用户-项目评分矩阵进行

预测。Yan 等^[13]利用用户的年龄、性别、职业等特征对用户进行聚类, 之后计算属于每个聚类的所有用户对每个项目的评分的平均值, 以此来解决新用户问题。

Kim 等^[35]将卷积神经网络 (Convolutional Neural Networks, CNN) 无缝地集成到 PMF 模型中, 提出卷积矩阵分解模型 (Convolutional Matrix Factorization, convMF), 该模型可以捕获上下文信息用于评分预测。Wang 等^[36]提出了一种深度知识感知网络 (Deep Knowledge-aware Network, DKN), 它将知识图表示融入到新闻推荐中。Wang 等^[37]将使用堆叠去噪自动编码器模型的内容信息与使用协同过滤的评分矩阵相结合。此外, 其在文献^[38]中还利用 CNN 从音乐信号中学习潜在表示用于音乐推荐。Wei 等^[39]结合时间感知 CF 和深度学习方法解决项目冷启动问题。Bansal 等^[40]使用循环神经网络 (Recurrent Neural Network, RNN) 将文本序列编码为隐语义, 然后使用顺序敏感编码器结合隐语义模型解决冷启动与稀疏性问题。文献^[41]表明协同过滤可以看作一个序列预测问题, 将 RNN 用于解决冷启动问题具有很强的竞争性。

Li 等^[42]将矩阵分解与边缘化去噪自动编码器耦合并结合来自用户和项目的辅助信息对矩阵分解进行了泛化。He 等^[43]提出一种贝叶斯二元神经网络框架 (Bayesian Dual Neural Network framework, BDNet), 不仅考虑用户和项目的辅助信息, 同时还考虑权重的不确定性问题。具体来说, 该框架由两个神经网络组成, 首先使用 MLP 学习用户和项目的 one-hot 编码的公共低维空间, 然后使用元素相乘的方法联结这两个隐藏表示到新向量中。同样通过 MLP 将用户和项目的辅助信息投影到另一个共享的潜在空间中, 并使用元素相乘的方法将用户和项目的辅助信息隐藏表示联结成一个新的辅助信息向量。面对辅助信息不同类型, 采用不同类型的神经网络建模 (如, 对文本使用循环神经网络, 对图片使用卷积神经网络)。最后将两个神经网络的输出合并以产生最终预测。此外, 将高斯先验添加到所有的权重和偏置中可得到校准的概率分布, 以做出更稳健的预测。

由于好友推荐可以增加推荐的信任度且可以解决冷启动问题, 学者们开始考虑利用用户的社交关系进行推荐。Lin 等^[44]利用社交信息解决项目冷启动问题。Ma 等^[45]同时利用用户的社交网络信息和评分记录, 解决了数据稀疏和预测精度差的问题。其在文献^[46]中提出了额外的社会正则化项, 确保两个具有相似口味的朋友的潜在特征向量之间的距离更近。

Yang 等人^[47]将现有的评分数据与朋友关系结合, 推断出特定类别的社会信任圈。Cantador 等人^[48]利用用户和项目信息, 根据社会标签的加权列表来定义 Top-N 推荐。此外, 他们还在文献^[49]中对社交系统中不同类型信息对项目推荐的影响进行了比较研究。

由于注意力机制在机器翻译中被证明是非常有效的, 基于注意力的推荐模型近年来得到了发展。Chen 等人^[50]将隐语义模型(latent factor model, LFM)与基于项目和成分级注意力的邻域模型相结合, 对多媒体推荐中的隐式反馈进行建模, 取得了较好的效果。Han 等^[51]提出一种考虑个体多样性的自适应深层隐语义模型(Adaptive Deep Latent Factor Model, ADLFM), 首先使用 CNN 和 LFM 生成项目的潜在表示, 然后使用余弦相似度计算项目之间相关度表示为注意力因素, 并将其作为项目权重, 最后将用户已评分项目的加权和作为用户表示。ADLFM 是第一个利用项目描述及注意力因素建模用户表示的工作。Chen 等^[52]提出了一个基于协同过滤的模型, 即注意力驱动因子模型(Attention-driven Factor Model, AFM), 根据用户对商品不同角度的关注程度进行推荐。模型中的门控注意单元(Gated Attention Unit, GAUs)利用评分和商品特征评估用户偏好, 如图 8 所示, 将用户 u_i 的评分表示与商品的不同特征表示 x_j 通过 GAU 筛选出一个最有价值的输出用于预测。该方法可以针对不同用户生成注意力分布, 根据用户的注意力分布, AFM 可以得到很高的预测精度。此外, 还可以通过向用户展示注意力分布图来帮助提高用户对推荐的信任。

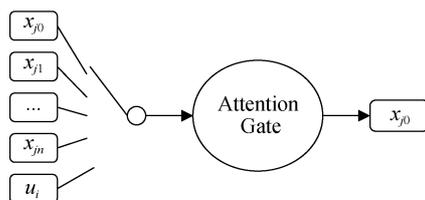


图 8 门控注意单元
Figure 8 GAUs

异质信息网络作为一种有效的信息建模方法, 也逐渐受到人们的关注。该种网络可以通过构建多种类型的实体以及实体之间的联系表示各种数据信息, 将不同数据信息应用到推荐任务中以带来更好的推荐效果。基于元路径进行相似度度量是最重要和最基础的一个方向。Sun 等^[53]提出 PathSim 算法, 通过抽取用户间对称的元路径度量用户之间的相似性。Shi 等^[54]提出基于元路径的双向随机游走算法

HeteSim, 用于度量网络中任意节点间的相关性。Wang 等^[55]提出 OptRank 算法, 通过利用社会标签系统中包含的异质信息来缓解冷启动问题。Luo 等^[56]提出一种基于异质关系的社会协同过滤算法。Shi 等^[57]提出加权异质信息网络的概念, 提出 SemRec 算法用于个性化推荐。随着网络表示学习的兴起, Shi 等^[58]提出 HERec 方法, 该方法基于元路径的随机游走生成节点序列以学习节点的嵌入表示, 并将其集成至矩阵分解框架用于推荐。Zhao 等^[59]提出一种基于异质网络嵌入的推荐方法 HetNERec。该方法从异质信息网络中提取多个共现关系来构建共现网络以解决基于元路径方法产生的稀疏性问题。

3.1.3 迁移协同过滤

迁移学习通过将将一个领域(即源领域)的知识迁移到另外一个领域(即目标领域)使目标领域能够取得更好的学习效果。跨域推荐系统(Cross Domain Recommendation System, CDRS)通过迁移学习, 利用其他域的评分信息解决稀疏问题以提高目标域的预测。现有的 CDRS 大致可分为迁移知识的方法和聚合知识的方法。

(1) 迁移知识。该方法将从源域学习的模式直接转移到目标域以提高目标域的推荐精度, 如图 9 所示。

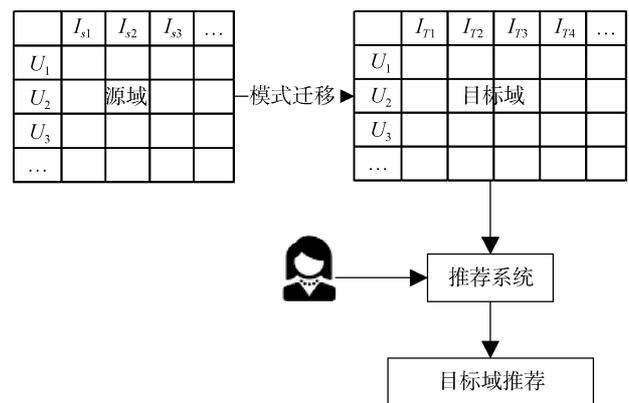


图 9 基于迁移知识的跨域推荐
Figure 9 CDRS based on transfer knowledge

Lee 和 Seung^[60]提出第一个用评分模式进行跨域推荐的方法。该方法根据域的相似度得分识别全局最近邻, 然后使用关联规则发现由一组邻居用户共同评价的项目, 最后使用基于用户的 CF 算法进行评分预测, 使用包含目标项目的用户规则进行增强。码本迁移(Codebook Transfer, CBT)^[61]将用户-项目评分模式从源域中的密集评分矩阵转移到目标域中的稀疏评分矩阵中, 该方法不需要用户或项目的重叠。Li 等^[62]将该思想扩展到概率框架的方法中, 提出评分

矩阵生成模型 (Rating-Matrix Generative Model, RMGM), 该模型不根据密集的源域数据来构建代码本, 而是提取所有评分矩阵的共享模式。此外, 用户或项目并不只属于单独的集群而是引入概率分布, 允许用户和项目属于多个集群, 具有不同的参与度。

(2) 聚合知识。该方法同时从源域和目标域中学习模型, 并期望这些域能够相互补充。在此类方法中, 源域和目标域应共享同一用户表示^[63], 如图 10 所示。

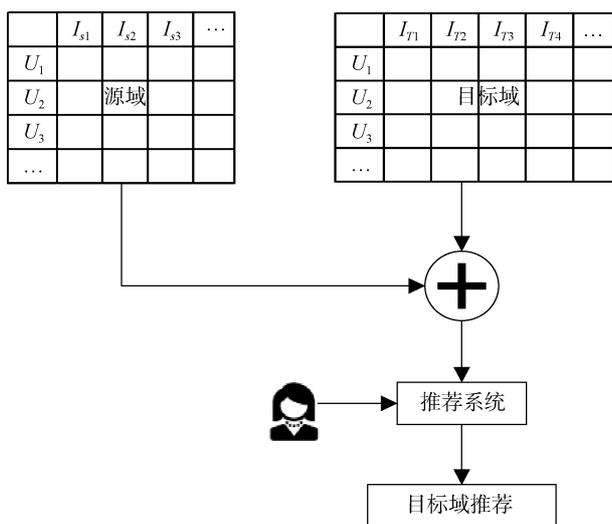


图 10 基于聚合知识的跨域推荐

Figure 10 CDRS based on aggregated knowledge

Shapira 等^[64]将 Facebook 点赞组成的数据集作为用户偏好证明了当可用用户偏好稀疏时, 使用聚合的方法能显著提高准确性。Shi 等^[65]利用用户贡献标签作为共同特征, 提出一种广义标签诱导的跨域协同过滤模型。文献[66]提出多视图深度神经网络 (Multi-View Deep Neural Network, MVDNN), 该模型是推荐系统中一种跨域的多视图深度学习, 其使用深度学习方法将用户和项目映射到潜在空间中, 在该空间中用户与他最喜欢的商品之间具有最大相似度并且商品特征从不同域中联合学习。Lian 等^[67]提出基于神经网络的跨领域推荐模型 (Content-boosted Collaborative Filtering Neural Network For Cross Domain Recommender Systems, CCCFNet), 该模型利用神经网络将项目的辅助信息引入协同过滤中。文献[68]考虑到 MVDNN 只使用辅助信息而没有考虑协同过滤, 没有利用用户和项目之间相关性这一优势, 因此将 BNet 框架中考虑了权重不确定性的协同过滤方法与 MVDNN 框架中的跨域方法相结合, 提出一种使用贝叶斯神经网络的通用跨域框架 (General Cross-domain framework via Bayesian Neural network, GCBAN)。Yu 等^[69]考虑到源域和目标域中

没有用户或项目重叠时直接共享两个域的嵌入可能会导致嵌入错位, 因此引入评论文本信息, 将两个域的嵌入分别与对应的文本语义空间对齐, 由此保证两个域的嵌入可以进行更准确的空间对齐。Zhang 等^[70]提出一种双迁移学习框架以解决数据带来的长尾问题, 该框架从模型级和项目级两个层面联合学习知识迁移, 通过合并两种迁移确保头部项目学习的知识可以很好地应用于长尾分布中的尾部项目表示学习。

由于现实场景中的许多应用往往涉及到不同领域, 学者们开始考虑利用多个源域的信息辅助目标域做出推荐。Moreno 等^[71]提出多域迁移学习方法 (Transfer Learning for Multiple Domains, TALMUD), 该方法首先从不同源域中生成多个输出, 然后分配权重将它们整合在一起。Zhuang 等^[72]提出一种基于共识正则化的多源迁移协同过滤框架 (Transfer Collaborative Filtering Framework from Multiple Sources via Consensus regularization, TRACER)。该框架基于 CBT 跨域协同过滤方法, 从多个源域迁移知识用于协同过滤, 并设计处理不同领域产生预期结果不一致问题的共识正则化。Yan 等^[73]通过在跨域模型参数上加入很小却很重要的干扰, 发现模型性能下降很多, 认为跨域模型的鲁棒性不如单域模型。因此提出一种对抗跨域网络 (Adversarial Cross-Domain Network, ACDN), 通过动态生成对抗样本来训练跨域推荐模型, 从而提高泛化性能。该对抗样本通过在输入样本中添加很小却有意义的干扰来生成, 使模型以高置信度给出错误输出。由于在输入层添加噪声可能会改变语义, 因此将对抗样本应用于底层推荐模型的参数中。ACDN 将 MLP++^[74]作为基础模型, 在 MLP++ 目标函数中加入对抗样本以使其最小化。ACDN 的目标函数为:

$$L_{ACDN}(\Theta) = L_{MLP++}(\theta_{emb}, \theta_f) + \lambda L_{MLP++}(\theta_{emb} + n_{adv}, \theta_f) \quad (20)$$

其中 n_{adv} 表示嵌入参数 θ_{emb} 的对抗干扰, $\theta_{emb} + n_{adv}$ 是对抗样本的表示。

文献[75]是跨域推荐在实际场景的一次应用, 利用线上网购行为与线下消费行为共同推动线下零售推荐。提出概率图模型, 即线上到线下主题模型 (Online-to-Offline Topic Modeling, O2OTM)。该模型使用隐含狄利克雷分布 (Latent Dirichlet Allocation, LDA) 主题模型对线上和线下活动进行主题建模并建立线上线下主题之间的关系。

3.2 可扩展性问题

随着互联网和信息计算的发展, 用户和项目的

数量也急剧增长, 传统协同过滤推荐算法的计算过程和运算量将呈指数型增长, 会使推荐系统遭受很严重的可扩展性问题; 此外, 系统需要准确、实时地为系统中的所有用户进行推荐, 而这需要推荐系统具有较高的可扩展性。在协同过滤技术中, 基于模型的方法更易解决系统的可扩展性问题。

Yu 等^[76]考虑了非参数矩阵分解方法, 令 SVD 和概率主成分分析(probabilistic Principal Component Analysis, pPCA)的隐语义由数据驱动, 维数随着数据规模的增加而增加, 使非参数矩阵分解模型在大规模问题上更加有效。李聪^[5]提出基于 PCA 和竞争网络(Self-Organizing Map, SOM)聚类的协同过滤模型, 该方法首先利用基于项目的协同过滤补全空值评分, 然后利用主成分分析数据补全之后的项目, 最后在主成分空间中对用户进行 SOM 聚类。

基于用户聚类的工作中, Schafer 等^[9]首先对用户聚类以生成若干聚类中心, 然后计算用户与聚类中心的相似度生成类别所属的程度矩阵, 最后在矩阵中搜寻目标用户的最近邻。Gao 等^[10]考虑到聚类算法容易将附近用户划分为不同集群, 而只为同一集群的用户进行推荐, 导致附近用户流失错过推荐。因此, 提出一种基于交叉聚类的方法, 首先定义关联度来寻找相邻的聚类, 然后合并用户聚类中的邻居来考虑近邻用户, 从而获得更多推荐机会。Koochi 等^[77]采用子空间聚类的方法寻找邻居用户来解决稀疏性问题。其通过将评分矩阵构造为三个子集来为用户的邻居创建了三层树, 实验证明在解决稀疏数据时是有效的。Xue 等^[78]利用 K-means 聚类算法, 根据用户的评分对用户进行聚类, 然后在聚类基础上进行平滑和邻域选择以填充用户-项目矩阵中的空白元素。Wang 等^[79]利用 K-means 算法对用户评分进行聚类, 解决协同过滤中用户和项目数量非线性增长时的稀疏性和扩展性问题。Kim 等^[80]将 K-means 聚类应用于网上购物系统, 根据用户的购买行为对其进行分类, 并利用遗传算法解决 K-means 聚类中的局部最优问题。

基于项目聚类的工作中, 文献[4]将项目聚类与注意力因素相结合, 将其应用于疾病诊断中。Verma 等^[14]提出一种使用协同过滤与模糊 C 均值(Fuzzy C-means, FCM)聚类算法的推荐系统。该系统利用 FCM 聚类方法进行项目聚类, 使用协同过滤聚类进行评分预测。文献[15]基于 Hadoop 进行项目聚类, 解决系统扩展性差、性能低的问题。文献[81]认为在找到用户邻居时应考虑目标项与每个共评项目之间的相似性。

3.3 多样性问题

该问题是指用户只能获得其用户画像中已知的或已定义的物品推荐, 阻碍用户发现新物品或其他选择。然而推荐多样性是一切推荐系统的目标。

Acker 和 McReynolds^[82]首次提出“新颖性”这一概念。随后, Rajus^[83]研究了在人格特征、人口统计学变量和消费者背景下的探索性行为。Zhang 等^[84]利用观察到的用户行为探索猎奇特征, 在单一领域对其建模, 如图 11 所示。其中 Z 表示猎奇水平; ϕ 表示选择效用分布, 即个体对每个选择的偏好; X 表示动作序列, 取值取决于此时的猎奇水平、选择效用分布 ϕ 和先前选择的动作。

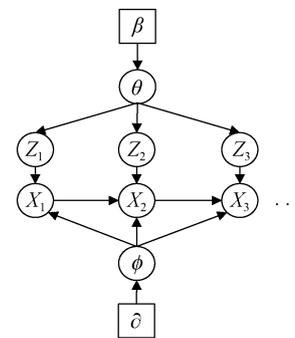


图 11 猎奇模型

Figure 11 Novel Seeking Model (NSM)

Zhuang 等^[85]受其启发, 在多个域中探索猎奇行为, 提出跨领域猎奇模型(Cross-domain Novelty Seeking Trait, CDNST), 该模型基于序列行为从辅助域迁移知识, 有效提高了推荐性能。Cheng 等^[86]将多样化推荐视为一项监督学习任务, 定义了两个耦合优化问题, 提出一种多样化协同过滤算法来解决这两个问题。Gogna 和 Majumdar^[87]利用矩阵补全来平衡准确性和多样性。Xu 等^[88]定义了一个具有高满意度和低初始兴趣度的新概念: Serendipity。基于此概念提出一个两阶段推荐问题, 包括 Serendipity 预测和个性化推荐。在此基础上, 开发了神经意外发现的推荐(Neural Serendipity Recommendation, NSR)方法。如图 12 所示, NSR 利用 MF 来保证推荐的准确性, 利用 MLP 来捕获新颖度, 通过合并 MLP 和 MF 去预测 Serendipity, 然后提出一种候选过滤方法, 它可以最大限度的从预测的 Serendipity 中获取知识, 缓解数据稀疏问题并给出个性化的推荐。

3.4 可解释性问题

该问题是由于没有解释和理由的推荐是缺乏说服力的。给出合理解释可以提升推荐系统的透明度, 提升用户对推荐系统的信任度和接受度, 进而提升

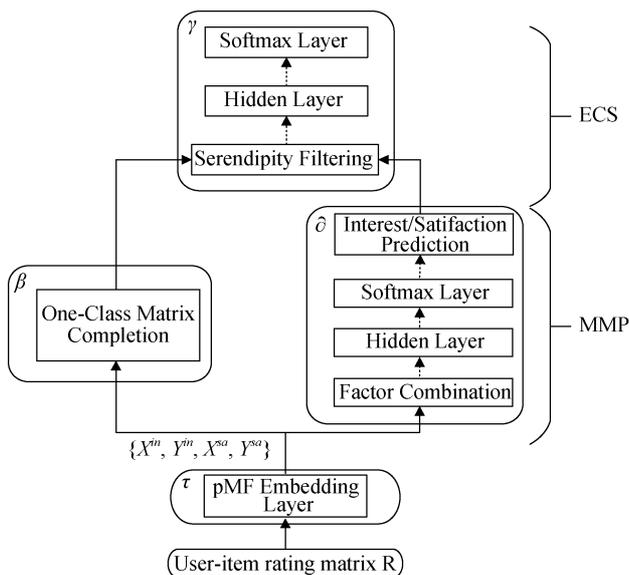


图 12 神经意外发现推荐模型
Figure 12 Neural serendipity recommendation model

用户对推荐产品的满意度。例如, 文献[2]在对话推荐中加入解释, 有效的提高推荐可信度。如图 13 所示。

Model: I recommend *Pulp Fiction*. This is a dark comedy with a great cast.
 User: I don't want to watch a comedy right now.
 Model: How about *Ice Age*? It is a very good anime with a lot of action adventure.
 User: I don't like anime, but action movie sounds goods.
 Model: I recommend *Mission Impossible*. This is by far the best of the action series.
 User: Sounds great. Thanks for the recommendation!

Predefined Template	Recommended Item	Generated Explanation
---------------------	------------------	-----------------------

图 13 可解释性会话推荐

Figure 13 Interpretable conversation recommendation

文献 [89] 提出一种可解释矩阵分解模型 (Explainable Matrix Factorization, EMF), 计算一个可解释的 Top-N 项目排名列表。此外, 引入新的解释质量指标, 分别是平均解释精度 (Mean Explainability Precision, MEP) 和平均解释召回 (Mean Explainability Recall, MER)。Zhang 等 [90] 提出一种基于模型的协同过滤方法, 即显因子模型 (Explicit Factor Model, EFM)。该模型除了使用评分数据以外, 还基于项目特征和用户评论的情感分析作为数据源来生成可解释推荐。随后, 在文献 [91] 中将词云用于推荐中提供解释。Chen 等 [92] 将 EFM 扩展为张量分解, 通过从文本评论中提取项目特征构造了用户-项目-特征立方体。基于这个立方体进行成对学习以预测用户对特征和项目的偏好并以此做出个性化推荐。Bauman 等 [93] 提出情感效用逻辑模型 (Sentiment Utility Logistic Model, SULM), 该模型通过提取特征及用户对这

些特征的情感, 将两者整合到矩阵分解模型中, 对未知情感和评分进行拟合生成推荐。该方法不仅能为用户提供推荐的项目还能为项目提供推荐特征, 特征即为推荐的解释。Qiu 等 [94] 和 Hou 等 [95] 通过整合项目评分和评论, 研究了基于方面级别的隐语义模型。

Su 等 [96] 提出多矩阵分解 (Multi-Matrix Factorization, MMF) 模型。MMF 方法中具有权重的属性评分能得到每个属性对推荐的贡献程度, 从而可以提供可解释性。Kouki 等 [97] 基于概率模型提出一个混合推荐模型, 利用混合模型分析用户偏好, 并通过对在线用户的调查评价了文本、可视化及图形化的解释。Shulman 和 Wolf [98] 建立了基于每个用户决策树的可解释推荐系统以及基于单个属性值的决策规则, 从而得到一个协同过滤解决方案, 该方案可以直接解释它提供的每个评分。Hada 等 [99] 提出可以提供推荐解释的端到端框架 ReXPlug, 其利用即插即用语言模型生成可解释性推荐, 并利用联合训练的交叉注意力网络来个性化推荐解释。Ghazimatin 等 [100] 提出人机协同框架 Elixir, 利用推荐和解释的反馈去学习用户特定的潜在偏好向量, 同时通过基于项目相似性的邻域标签传播可以克服稀疏性。

4 未来研究方向

随着信息技术和互联网行业的发展, 信息过载成为人们处理信息的挑战。推荐系统能根据用户的浏览习惯, 确定用户的兴趣, 通过发掘用户的行为, 将合适的信息推荐给用户, 满足用户的个性化需求。随着图神经网络、知识图谱等研究的发展, 将其与推荐系统相结合以做出更准确的推荐是必要的。

(1) 推荐系统与图神经网络

近年来, 由于图结构的强大表现力, 用机器学习方法分析图的研究越来越受到重视。图神经网络是一类基于深度学习的处理图域信息的方法。由于其较好的性能和可解释性, GNN 最近已成为一种广泛应用的图分析方法。推荐系统中的用户-项目交互关系可以看作一个图, 即用户和项目作为节点, 两者之间的交互作为边。同时还可以融入用户的社交网络、项目的属性信息等。

面对推荐中的用户与项目为不同类型的节点, 可以应用异质信息网络; 面对数据稀疏以及长尾问题, 可以引入邻居采样等技术; 面对图的大规模问题, 可以进行采样以及子图训练等优化方式。

(2) 推荐系统与知识图谱

在多数推荐场景中, 物品中可能包含丰富的知

识信息, 刻画这些知识的网络结构被称为知识图谱。一方面, 知识图谱中的结构化知识有助于缓解推荐中的冷启动等问题; 另一方面, 使用知识图谱能极大扩展物品信息, 强化物品之间的联系, 为推荐提供丰富的参考价值, 为推荐结果带来额外的多样性和解释性。由于现有的模型网络结构大多都是静态的, 而在真实场景中, 知识图谱具有一定的时效性。如何刻画这种时间演变的网络并在推荐时充分考虑时序信息是需要我们考虑的。

(3) 交互式推荐系统

目前的推荐结果大多由平台和算法主导, 用户比较被动, 只能接收推荐引擎输送过来的推荐结果。推荐过于单一或者推荐不感兴趣时用户没有调优与反馈的权利。可交互的推荐系统不仅可以通过数据的迭代来更新推荐算法, 还可以通过人机交互的方式帮推荐算法感知人们的情感色彩, 使推荐系统不仅能推断用户对物品的喜好, 还能感知用户对推荐算法本身的调优方向。

交互式推荐系统可分为两种思路: 一种是结合自然语言处理, 构建对话机器人, 通过与用户交谈来获得用户意图和偏好。其难点在于对话的理解与生成, 以及如何在对话过程中对推荐场景有所感知; 另一种是交互式的意图挖掘, 利用用户的少量交互行为, 快速得到用户偏好以完成推荐任务。其难点在于交互形式倾向于强化学习算法的应用。

(4) 推荐系统的可解释性

近年来, 学者们发现在推荐时给出推荐理由会增加用户对推荐的认可度和接受度, 没有解释和理由的推荐是缺乏说服力的。互联网上的虚拟物品推荐, 如果能够像线下推荐一样, 在给出推荐的同时提供推荐解释、说明推荐原因, 可以提升推荐系统的透明度, 还能提升用户对推荐系统的信任度和接受度, 进而提升用户对推荐产品的满意度。在可解释性研究中, 一方面, 可以通过知识图谱来增强解释能力; 另一方面, 可以建立模型无关的可解释推荐框架。此外, 还可以考虑结合生成模型进行对话式推荐。

5 结论

通过对协同过滤推荐算法进行回顾, 可以发现解决传统协同过滤算法存在的冷启动与稀疏性、可扩展性、多样性以及可解释性问题时, 目前的工作主要是利用评分数据、社交网络信息以及其他领域信息等辅助信息, 结合深度学习、数据挖掘等技术提高推荐效果、提升用户满意度。随着图神经网络、

知识图谱等研究的发展, 将其与推荐系统相结合以做出更准确的推荐是未来一个重要的发展方向。

参考文献

- [1] Zhuang F Z, Luo D, Yuan N J, et al. Representation Learning with Pair-Wise Constraints for Collaborative Ranking[C]. *The Tenth ACM International Conference on Web Search and Data Mining*, 2017: 567-575.
- [2] Chen Z X, Wang X T, Xie X, et al. Towards Explainable Conversational Recommendation[C]. *The Twenty-Ninth International Joint Conference on Artificial Intelligence*, 2020.
- [3] Nair A M, Benny O, George J. Content Based Scientific Article Recommendation System Using Deep Learning Technique[M]. *Inventive Systems and Control*. Singapore: Springer Singapore, 2021: 965-977.
- [4] Guzmán G, Torres-Ruiz M, Tambonero V, et al. A Collaborative Framework for Sensing Abnormal Heart Rate Based on a Recommender System: Semantic Recommender System for Healthcare[J]. *Journal of Medical and Biological Engineering*, 2018, 38(6): 1026-1045.
- [5] Li C. *Research on the Bottleneck Problems of Collaborative Filtering In E-Commerce Recommender Systems*[D]. Hefei: Hefei University of Technology, 2009.
(李聪. 电子商务推荐系统中协同过滤瓶颈问题研究[D]. 合肥: 合肥工业大学, 2009.)
- [6] Xu H L, Wu X, Li X D, et al. Comparison Study of Internet Recommendation System[J]. *Journal of Software*, 2009, 20(2): 350-362.
(许海玲, 吴潇, 李晓东, 等. 互联网推荐系统比较研究[J]. *软件学报*, 2009, 20(2): 350-362.)
- [7] Wang L. *Research on e-commerce personalized system based on Collaborative Filtering Technology*[D]. East China Normal University, 2008.
(王亮. 基于协同过滤技术的电子商务个性化系统研究[D]. 华东师范大学, 2008)
- [8] Rodrigues C M, Rathi S, Patil G. An Efficient System Using Item & User-Based CF Techniques to Improve Recommendation[C]. *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, 2016: 569-574.
- [9] Schafer J B, Konstan J A, Riedl J. *E-Commerce Recommendation Applications*[M]. *Applications of Data Mining to Electronic Commerce*. Boston, MA: Springer US, 2001: 115-153.
- [10] Gao M, Cao F Y, Huang J Z. A Cross Cluster-Based Collaborative Filtering Method for Recommendation[C]. *2013 IEEE International Conference on Information and Automation*, 2013: 447-452.
- [11] Kim E, Kim M, Ryu J. Collaborative Filtering Based on Neural Networks Using Similarity[M]. *Advances in Neural Networks – ISNN 2005*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005:

- 355-360.
- [12] Shani G, Brafman R I, Heckerman D. An MDP-Based Recommender System [J]. *Journal of Machine Learning Research*, 2005, 6: 1265-1295.
- [13] Ling Y X, Guo D K, Cai F, et al. User-Based Clustering with Top-N Recommendation on Cold-Start Problem[C]. *2013 Third International Conference on Intelligent System Design and Engineering Applications*, 2013: 1585-1589.
- [14] Verma S K, Mittal N, Agarwal B. Hybrid Recommender System Based on Fuzzy Clustering and Collaborative Filtering[C]. *2013 4th International Conference on Computer and Communication Technology*, 2013: 116-120.
- [15] Zhao W, Zhang H T. The Improved Item-Based Clustering Collaborative Filtering Algorithm Based on Hadoop[C]. *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference*, 2017: 2416-2419.
- [16] Wang Q. *Research and Application of Recommendation Algorithm Based on Association Rules and Collaborative Filtering*[D]. Zhenjiang, China: Jiangsu University of Science and Technology, 2020. (王琦. 基于关联规则和协同过滤的推荐算法研究与应用[D]. 镇江: 江苏科技大学, 2020.)
- [17] Rendle S, Freudenthaler C, Gantner Z, et al. BPR: Bayesian Personalized Ranking from Implicit Feedback [C]. *The Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. 2009: 452-461.
- [18] Lee D D, Seung H S. Learning the Parts of Objects by Non-Negative Matrix Factorization[J]. *Nature*, 1999, 401(6755): 788-791.
- [19] Paterek A. Improving Regularized Singular Value Decomposition for Collaborative Filtering[J]. *KDD Cup and Workshop*, 2007: 2-5.
- [20] Mnih A, Salakhutdinov R R. Probabilistic Matrix Factorization[J]. *Advances in Neural Information Processing Systems*, 2007, 20: 1257-1264.
- [21] 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.
- [22] Zhou K, Yang S H, Zha H Y. Functional Matrix Factorizations for Cold-Start Recommendation[C]. *The 34th international ACM SIGIR conference on Research and development in Information Retrieval*, 2011: 315-324.
- [23] Gu L, Yang P, Dong Y Q. An Dynamic-Weighted Collaborative Filtering Approach to Address Sparsity and Adaptivity Issues[C]. *2014 IEEE Congress on Evolutionary Computation*, 2014: 3044-3050.
- [24] Salakhutdinov R, Mnih A, Hinton G. Restricted Boltzmann Machines for Collaborative Filtering[C]. *The 24th international conference on Machine learning - ICML '07*, 2007: 791-798.
- [25] Truyen T T, Phung D Q, Venkatesh S. Ordinal Boltzmann Machines for Collaborative Filtering [C]. *The Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. 2009: 548-556.
- [26] Zhuang F Z, Zhang Z Q, Qian M D, et al. Representation Learning via Dual-Autoencoder for Recommendation[J]. *Neural Networks*, 2017, 90: 83-89.
- [27] Lee D H, Kang S, Ju H, et al. Bootstrapping User and Item Representations for One-Class Collaborative Filtering [J]. ArXiv preprint ArXiv:2105.06323, 2021.
- [28] Zhu Z W, Kim J, Nguyen T, et al. Fairness among New Items In Cold Start Recommender Systems[C]. *The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021: 767-776.
- [29] Lin X X, Wu J, Zhou C, et al. Task-Adaptive Neural Process for User Cold-Start Recommendation[C]. *The Web Conference 2021*, 2021: 1306-1316.
- [30] Van Den Oord A, Dieleman S, Schrauwen B. Deep Content-Based Music Recommendation[C]. *Neural Information Processing Systems Conference. Neural Information Processing Systems Foundation*, 2013, 26.
- [31] McAuley J, Leskovec J. Hidden Factors and Hidden Topics: Understanding Rating Dimensions with Review Text[C]. *The 7th ACM conference on Recommender systems*, 2013: 165-172.
- [32] Ling G, Lyu M R, King I. Ratings Meet Reviews, a Combined Approach to Recommend[C]. *The 8th ACM Conference on Recommender systems*, 2014: 105-112.
- [33] Ren Z C, Liang S S, Li P J, et al. Social Collaborative Viewpoint Regression with Explainable Recommendations[C]. *The Tenth ACM International Conference on Web Search and Data Mining*, 2017: 485-494.
- [34] Ji K, Shen H. Using Category and Keyword for Personalized Recommendation: A Scalable Collaborative Filtering Algorithm[C]. *2014 Sixth International Symposium on Parallel Architectures, Algorithms and Programming*, 2014: 197-202.
- [35] Kim D, Park C, Oh J, et al. Convolutional Matrix Factorization for Document Context-Aware Recommendation[C]. *The 10th ACM Conference on Recommender Systems*, 2016: 233-240.
- [36] 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 on World Wide Web - WWW'18*, 2018: 1835-1844.
- [37] Wang H, Wang N Y, Yeung D Y. Collaborative Deep Learning for Recommender Systems[C]. *The 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015: 1235-1244.
- [38] Wang X X, Wang Y. Improving Content-Based and Hybrid Music Recommendation Using Deep Learning[C]. *The 22nd ACM inter-*

- national conference on Multimedia*, 2014: 627-636.
- [39] Wei J, He J H, Chen K, et al. Collaborative Filtering and Deep Learning Based Recommendation System for Cold Start Items[J]. *Expert Systems With Applications*, 2017, 69: 29-39.
- [40] Bansal T, Belanger D, McCallum A. Ask the GRU: Multi-Task Learning for Deep Text Recommendations[C]. *The 10th ACM Conference on Recommender Systems*, 2016: 107-114.
- [41] Ko Y J, Maystre L, Grossglauser M. Collaborative Recurrent Neural Networks for Dynamic Recommender Systems[C]. *Asian Conference on Machine Learning. PMLR*, 2016: 366-381.
- [42] Li S, Kawale J, Fu Y. Deep Collaborative Filtering via Marginalized Denoising Auto-Encoder[C]. *The 24th ACM International on Conference on Information and Knowledge Management*, 2015: 811-820.
- [43] He J, Zhuang F Z, Liu Y C, et al. Bayesian Dual Neural Networks for Recommendation[J]. *Frontiers of Computer Science*, 2019, 13(6): 1255-1265.
- [44] Lin J, Sugiyama K, Kan M Y, et al. Addressing Cold-Start In App Recommendation: Latent User Models Constructed from Twitter Followers[C]. *The 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2013: 283-292.
- [45] Ma H, Yang H X, Lyu M R, et al. SoRec: Social Recommendation Using Probabilistic Matrix Factorization[C]. *The 17th ACM Conference on Information and Knowledge Management*, 2008: 931-940.
- [46] Ma H, Zhou D Y, Liu C, et al. Recommender Systems with Social Regularization[C]. *The Fourth ACM International Conference on Web Search and Data Mining*, 2011: 287-296.
- [47] Yang X W, Steck H, Liu Y. Circle-Based Recommendation In Online Social Networks[C]. *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012: 1267-1275.
- [48] Cantador I, Bellogin A, Vallet D. Content-Based Recommendation In Social Tagging Systems[C]. *The Fourth ACM Conference on Recommender Systems*, 2010: 237-240.
- [49] Bellogin A, Cantador I, Castells P. A Comparative Study of Heterogeneous Item Recommendations In Social Systems[J]. *Information Sciences*, 2013, 221: 142-169.
- [50] Chen J Y, Zhang H W, He X N, et al. Attentive Collaborative Filtering: Multimedia Recommendation with Item- and Component-Level Attention[C]. *The 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2017: 335-344.
- [51] Han J Y, Zheng L, Xu Y B, et al. Adaptive Deep Modeling of Users and Items Using Side Information for Recommendation[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(3): 737-748.
- [52] Chen J W, Zhuang F Z, Hong X, et al. Attention-Driven Factor Model for Explainable Personalized Recommendation[C]. *SIGIR '18: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018: 909-912.
- [53] Sun Y Z, Han J W, Yan X F, et al. Pathsim: Meta Path-Based Top-k Similarity Search In Heterogeneous Information Networks [J]. *The VLDB Endowment*, 2011, 4(11): 992-1003.
- [54] Shi C, Kong X N, Huang Y, et al. HeteSim: A General Framework for Relevance Measure In Heterogeneous Networks[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2014, 26(10): 2479-2492.
- [55] Feng W, Wang J Y. Incorporating Heterogeneous Information for Personalized Tag Recommendation In Social Tagging Systems[C]. *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012: 1276-1284.
- [56] Luo C, Pang W, Wang Z, et al. Hete-CF: Social-Based Collaborative Filtering Recommendation Using Heterogeneous Relations[C]. *2014 IEEE International Conference on Data Mining*, 2014: 917-922.
- [57] 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.
- [58] Shi C, Hu B B, Zhao W X, et al. Heterogeneous Information Network Embedding for Recommendation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2019, 31(2): 357-370.
- [59] Zhao Z Y, Zhang X J, Zhou H, et al. HetNERec: Heterogeneous Network Embedding Based Recommendation[J]. *Knowledge-Based Systems*, 2020, 204: 106218.
- [60] Févotte C, Idier J. Algorithms for Nonnegative Matrix Factorization with the B-Divergence[J]. *Neural Computation*, 2011, 23(9): 2421-2456.
- [61] Li B, Yang Q, Xue X. Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction[C]. *The 21st International Joint Conference on Artificial Intelligence*. 2009: 2052-2057.
- [62] Li B, Yang Q, Xue X Y. Transfer Learning for Collaborative Filtering via a Rating-Matrix Generative Model[C]. *The 26th Annual International Conference on Machine Learning - ICML '09*, 2009: 617-624.
- [63] Berkovsky S, Kuflik T, Ricci F. Cross-Domain Mediation In Collaborative Filtering[M]. *User Modeling 2007*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007: 355-359.
- [64] Shapira B, Rokach L, Freilikhman S. Facebook Single and Cross Domain Data for Recommendation Systems[J]. *User Modeling and User-Adapted Interaction*, 2013, 23(2-3): 211-247.

- [65] Shi Y, Larson M, Hanjalic A. Generalized Tag-Induced Cross-Domain Collaborative Filtering[EB/OL]. 2013: arXiv: 1302.4888[cs.IR]. <https://arxiv.org/abs/1302.4888>
- [66] Elkahky A M, Song Y, He X D. A Multi-View Deep Learning Approach for Cross Domain User Modeling In Recommendation Systems[C]. *The 24th International Conference on World Wide Web*, 2015: 278-288.
- [67] Lian J X, Zhang F Z, Xie X, et al. CCCFNet: A Content-Boosted Collaborative Filtering Neural Network for Cross Domain Recommender Systems[C]. *The 26th International Conference on World Wide Web Companion*, 2017: 817-818.
- [68] He J, Liu R, Zhuang F Z, et al. A General Cross-Domain Recommendation Framework via Bayesian Neural Network[C]. *2018 IEEE International Conference on Data Mining*, 2018: 1001-1006.
- [69] Yu W H, Lin X, Ge J F, et al. Semi-Supervised Collaborative Filtering by Text-Enhanced Domain Adaptation[C]. *The 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020: 2136-2144.
- [70] Zhang Y, Cheng D Z, Yao T S, et al. A Model of Two Tales: Dual Transfer Learning Framework for Improved Long-Tail Item Recommendation[C]. *The Web Conference 2021*, 2021: 2220-2231.
- [71] Moreno O, Shapira B, Rokach L, et al. TALMUD: Transfer Learning for Multiple Domains[C]. *The 21st ACM International Conference on Information and Knowledge Management*, 2012: 425-434.
- [72] Zhuang F Z, Zheng J, Chen J W, et al. Transfer Collaborative Filtering from Multiple Sources via Consensus Regularization[J]. *Neural Networks*, 2018, 108: 287-295.
- [73] Yan H R, Zhao P P, Zhuang F Z, et al. Cross-Domain Recommendation with Adversarial Examples [C]. *International Conference on Database Systems for Advanced Applications*, 2020: 573-589.
- [74] Hu G N, Zhang Y, Yang Q. CoNet: Collaborative Cross Networks for Cross-Domain Recommendation[C]. *The 27th ACM International Conference on Information and Knowledge Management*, 2018: 667-676.
- [75] Luo P, Yan S, Liu Z Q, et al. From Online Behaviors to Offline Retailing[C]. *The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016: 175-184.
- [76] Yu K, Zhu S H, Lafferty J, et al. Fast Nonparametric Matrix Factorization for Large-Scale Collaborative Filtering[C]. *The 32nd international ACM SIGIR conference on Research and development in information retrieval*, 2009: 211-218.
- [77] Koohi H, Kiani K. A New Method to Find Neighbor Users that Improves the Performance of Collaborative Filtering[J]. *Expert Systems With Applications*, 2017, 83: 30-39.
- [78] Xue G R, Lin C X, Yang Q, et al. Scalable Collaborative Filtering Using Cluster-Based Smoothing[C]. *The 28th annual international ACM SIGIR conference on Research and development in information retrieval*, 2005: 114-121.
- [79] Wang P. A Personalized Collaborative Recommendation Approach Based on Clustering of Customers[J]. *Physics Procedia*, 2012, 24: 812-816.
- [80] Kim K J, Ahn H. A Recommender System Using GA K-Means Clustering In an Online Shopping Market[J]. *Expert Systems With Applications*, 2008, 34(2): 1200-1209.
- [81] Choi K, Suh Y. A New Similarity Function for Selecting Neighbors for each Target Item In Collaborative Filtering[J]. *Knowledge-Based Systems*, 2013, 37: 146-153.
- [82] Acker M, McReynolds P. The "Need for Novelty": A Comparison of Six Instruments[J]. *The Psychological Record*, 1967, 17(2): 177-182.
- [83] Raju P S. Optimum Stimulation Level: Its Relationship to Personality, Demographics, and Exploratory Behavior[J]. *Journal of Consumer Research*, 1980, 7(3): 272-282.
- [84] Zhang F Z, Yuan N J, Lian D F, et al. Mining Novelty-Seeking Trait across Heterogeneous Domains[C]. *The 23rd international conference on World wide web*, 2014: 373-384.
- [85] Zhuang F Z, Zhou Y M, Ying H C, et al. Sequential Recommendation via Cross-Domain Novelty Seeking Trait Mining[J]. *Journal of Computer Science and Technology*, 2020, 35(2): 305-319.
- [86] Cheng P Z, Wang S Q, Ma J, et al. Learning to Recommend Accurate and Diverse Items[C]. *The 26th International Conference on World Wide Web*, 2017: 183-192.
- [87] Gogna A, Majumdar A. Balancing Accuracy and Diversity In Recommendations Using Matrix Completion Framework[J]. *Knowledge-Based Systems*, 2017, 125: 83-95.
- [88] Xu Y B, Yang Y J, Wang E, et al. Neural Serendipity Recommendation[J]. *ACM Transactions on Knowledge Discovery from Data*, 2020, 14(4): 1-25.
- [89] Abdollahi B, Nasraoui O. Explainable Matrix Factorization for Collaborative Filtering[C]. *The 25th International Conference Companion on World Wide Web*, 2016: 5-6.
- [90] Zhang Y F, Lai G K, Zhang M, et al. Explicit Factor Models for Explainable Recommendation Based on Phrase-Level Sentiment Analysis[C]. *The 37th international ACM SIGIR conference on Research & development in information retrieval*, 2014: 83-92.
- [91] Zhang Y F. Incorporating Phrase-Level Sentiment Analysis on Textual Reviews for Personalized Recommendation[C]. *The Eighth ACM International Conference on Web Search and Data Mining*, 2015: 435-440.
- [92] Chen X, Qin Z, Zhang Y F, et al. Learning to Rank Features for Recommendation over Multiple Categories[C]. *The 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, 2016: 305-314.

- [93] Bauman K, Liu B, Tuzhilin A. Aspect Based Recommendations: Recommending Items with the most Valuable Aspects Based on User Reviews[C]. *The 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017: 717-725.
- [94] Qiu L, Gao S, Cheng W L, et al. Aspect-Based Latent Factor Model by Integrating Ratings and Reviews for Recommender System[J]. *Knowledge-Based Systems*, 2016, 110: 233-243.
- [95] Hou Y F, Yang N, Wu Y, et al. Explainable Recommendation with Fusion of Aspect Information[J]. *World Wide Web*, 2019, 22(1): 221-240.
- [96] Su Y X, Erfani S M, Zhang R. MMF: Attribute Interpretable Collaborative Filtering[C]. *2019 International Joint Conference on Neural Networks*, 2019: 1-8.
- [97] Kouki P, Schaffer J, Pujara J, et al. User Preferences for Hybrid Explanations[C]. *The Eleventh ACM Conference on Recommender Systems*, 2017: 84-88.
- [98] Shulman E, Wolf L. Meta Decision Trees for Explainable Recommendation Systems[C]. *The AAAI/ACM Conference on AI, Ethics, and Society*, 2020: 365-371.
- [99] Hada D V, Vijaikumar M, Shevade S K. ReXPlug: Explainable Recommendation Using Plug-and-Play Language Model[C]. *The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021: 81-91.
- [100] Ghazimatin A, Pramanik S, Roy R S, et al. ELIXIR: Learning from User Feedback on Explanations to Improve Recommender Models[C]. *The Web Conference 2021*, 2021: 3850-3860.



赵俊逸 于 2018 年在烟台大学软件工程专业获得工学学士学位。现在郑州大学计算机技术专业攻读硕士学位。研究领域为机器学习与数据挖掘。研究兴趣包括: 医疗学者评价、异质图学习。Email: zhaojunyi9504@gs.zzu.edu.cn



庄福振 于 2011 年在中国科学院研究生院获得工学博士学位。现任北京航空航天大学人工智能研究院研究员, 博士生导师。研究领域为机器学习与数据挖掘, 研究兴趣包括: 迁移学习, 多任务学习, 推荐系统以及知识图谱。Email: zhuang-fuzhen@buaa.edu.cn



敖翔 于 2015 年在中国科学院计算技术研究所计算机软件与理论专业获得工学博士学位。现任中国科学院计算技术研究所副研究员、硕士生导师。研究领域为金融数据挖掘, 研究兴趣包括金融异常行为检测, 金融文本理解等。Email: aoxiang@ict.ac.cn



何清 于 2000 年在北京师范大学应用数学专业获得博士学位。现任中国科学院计算技术研究所, 智能信息处理重点实验室研究员、博士生导师。研究领域为机器学习与数据挖掘。研究兴趣包括: 数据挖掘、分布式计算。Email: heqing@ict.ac.cn



蒋慧琴 于 2004 年在日本千叶大学计算机与信息系统科学专业获得博士学位, 现任郑州大学教授。研究领域为医疗人工智能。研究兴趣包括: 医学影像处理与分析、深度学习及应用等。Email: iehqjiang@zzu.edu.cn



马岭 于 1997 年在北京航空航天大学航空宇航制造专业获得博士学位, 现任郑州大学教授。研究领域为机器视觉及应用。研究兴趣包括: 图像处理与分析、深度学习及应用等。Email: ielma@zzu.edu.cn