



考虑多层次潜在特征的个性化推荐模型

申情^{1,2}, 郭文宾¹, 楼俊钢^{1,3}, 余强国²

(1. 湖州师范学院信息工程学院, 浙江 湖州 313000; 2. 湖州学院理工学院, 浙江 湖州 313000;
3. 浙江省现代农业资源智慧管理与应用研究重点实验室, 浙江 湖州 313000)

摘要: 个性化推荐已成为解决信息过载的最有效手段之一, 也是海量数据挖掘研究领域的热点技术。然而传统推荐算法往往只使用用户对物品的评分信息, 而缺少对用户与物品潜在特征的综合考虑。基于因子分解机、宽神经网络、交叉网络和深度神经网络的融合, 提出一种新的考虑多层次潜在特征的模型, 可以提取用户与物品的浅层潜在特征、低阶非线性潜在特征、线性交叉潜在特征以及高阶非线性潜在特征。在 4 个常用的数据集上的实验结果表明, 考虑用户与物品多层次潜在特征可以有效提高个性化推荐的预测精度。最后, 研究了嵌入层维度以及神经元数量等因素对新模型预测性能的影响。

关键词: 个性化推荐; 层次化潜在特征; 深度学习

中图分类号: TP391.3

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2022035

Personalized recommendation model with multi-level latent features

SHEN Qing^{1,2}, GUO Wenbin¹, LOU Jungang^{1,3}, YU Qiangguo²

1. School of Information Engineering, Huzhou University, Huzhou 313000, China

2. School of Science and Engineering, Huzhou College, Huzhou 313000, China

3. Zhejiang Province Key Laboratory of Smart Management & Application of
Modern Agricultural Resources, Huzhou 313000, China

Abstract: Personalized recommendation has become one of the most effective means to solve information overload, and it is also a hot technology in the research field of massive data mining. However, traditional recommendation algorithms often only use the user's rating information on the item, and lack a comprehensive consideration of the potential characteristics of the user and the item. The factorization machine, wide neural network, crossover network and deep neural network were combined to extract the shallow latent features, low-order nonlinear latent features, linear cross latent features, and high-order nonlinear latent features of users and items. Thus, a new deep learning personalized recommendation model with multilevel latent features was established. The experimental results on four commonly used data sets show that considering the multi-level potential features of users and items can effectively improve the prediction accuracy of personalized recommendations. Finally, the influence of factors such as the dimensions of the embedding layer and the number of neurons on the prediction performance of the new model was studied.

Key words: personalized recommendation, hierarchical latent feature, deep learning

收稿日期: 2021-08-09; 修回日期: 2021-10-26

基金项目: 浙江省重点研发计划项目 (No.2020C01097)

Foundation Item: The Key Research and Development Program of Zhejiang Province (No.2020C01097)



0 引言

互联网的快速发展,为广大消费者带来了海量的信息,消费者难以轻松获取其需要的信息,由于信息过量而引发的信息过载已成为日常生活中面临的重要问题。如何让消费者在海量信息中找到需要的信息已成为重要的课题^[1]。推荐系统(recommendation system, RS)作为解决信息过载问题的有效手段之一受到了学术界与产业界的广泛关注。

传统的推荐算法主要分为两种:基于内容的推荐^[2]与基于协同过滤(collaborative filtering, CF)的推荐^[3]。传统的推荐算法只采用用户对物品的评分信息进行推荐,并未采用用户与物品的基本信息,导致传统推荐算法数据稀疏性强,推荐精度一般。为解决此问题,有学者提出在推荐模型中融合用户与物品的潜在信息,并引入深度学习(deep learning, DL)技术,建立考虑潜在信息的深度学习个性化推荐模型^[4-9],但目前大多数研究中,考虑潜在特征的层次较少,不能反映多种层次潜在特征之间的组合性质。为更全面挖掘用户和物品基本信息的多层次潜在特征,提出一种考虑多层次潜在特征的个性化推荐模型。在该模型中,用户与物品基本信息作为输入,采用因子分解机(factor machine, FM)提取低层次潜在特征,宽神经网络(wide neural network, WNN)提取低阶非线性潜在特征,交叉网络(cross network, CN)提取线性交叉层次潜在特征,深度神经网络(deep neural network, DNN)提取高阶非线性潜在特征。在获取4种层次的潜在特征基础上,将4种层次的潜在特征进行融合,获取层级融合特征,输入深度神经网络学习融合特征所携带的多层次潜在特征信息进行评分预测。通过4个公开的真实数据进行实验,实验结果表明,本文模型相较于其他相关模型,在预测精度上有一定提升。

1 相关工作

迄今为止,CF算法依然是学术界与产业界应用最广泛的推荐算法。但像CF算法这种传统的推荐算法中仅考虑用户与物品评分信息进行推荐,并未考虑用户与物品基本属性信息与其潜在特征,致使推荐性能难以进一步提升。为此,有学者将深度学习应用在推荐领域来提升推荐系统性能。例如,Guilherme等^[10]将深度学习技术引入传统协同过滤推荐算法中,对传统算法存在的缺陷进行改进。Mylavarapu等^[11]提出一种用于高级应用的协同过滤和人工神经网络的推荐系统,采用人工神经网络从项目中提取内容特征,支持向量机(support vector machine, SVM)专注于计算时间动态和项目特征,负责未知数据预测的过程,以缓解推荐系统中完全冷启动问题。Tang等^[12]提出了基于云模型的反向化神经网络推荐算法。首先,该算法使用云模型定性和定量变换方法处理用户额定值,为用户生成多个云预测值,这些值构成云层。然后,云层加入神经网络,提高评级预测的准确性。薛峰等^[13]提出基于深度神经网络和加权隐反馈的推荐算法(DeepNASVD++),采用深度神经网络建模用户与物品之间的关系,使用注意力机制计算历史交互物品在建模用户隐式反馈时的权重。杨洁等^[14]提出了基于混杂社会网络的Web服务推荐框架及算法。该网络加入了服务提供者这一元素,可提供更多的真实信息。根据提出的服务推荐框架设计了用户候选服务信任值预测算法和服务推荐算法,提升个性化推荐性能。Hui^[15]等将知识图谱视为异构网络添加辅助信息,提出了行为和知识特征统一嵌入的推荐模型,从用户历史行为和知识图谱中挖掘用户潜在偏好,为用户提供更准确和多样化的推荐。Zhen^[16]等提出了一种基于离线挖掘、实时挖掘和深度

学习技术相结合的个性化推荐系统架构,其核心思想是从多个电子商务平台实时收集和预处理数据,收集用户的个性化数据,为后续数据挖掘做准备,利用数据挖掘技术,满足用户的个性化需求。邵英玮等^[17]通过经济学的相关研究,提出了商品潜在互补性发现推荐模型,从经济学角度刻画商品间关系。基于经济学理论中的需求交叉弹性,提出互补性发现模型,联合商品价格与购买历史挖掘商品间的互补关系,并基于此提出了融合商品互补关系的双重注意力机制推荐模型。Fu等^[18]提出了一种新颖的深度学习的方法,该方法通过事先了解用户和项目模仿有效的智能推荐。在初始阶段,分别学习用户和项目的对应低维向量,从而嵌入反映用户—用户和项目—项目相关性的语义信息。在预测阶段,采用前馈神经网络模拟用户和物品之间的交互,将相应的预训练代表向量作为神经网络的输入。Yan等^[19]引入了一种新的基于深度学习的推荐模型解决稀疏性和冷启动推荐问题,通过数据预处理层、嵌入层、卷积网络层、共享层和自动编码器层深入挖掘潜在信息及其对用户功能和项目特征的相关性。之后,通过结合目标评分数据与处理后的用户和项目特征数据,在深度神经网络中进行最终评分预测。Tang等^[20]提出了一种卷积序列嵌入推荐模型应用在Top-N推荐领域。主要想法是将最近项目的序列嵌入到时间和潜在空间中的“图像”中,并使用卷积滤波器将顺序模式作为图像的局部特征。这种方法提供了一个统一且灵活的网络结构,用于捕获用户一般偏好和顺序模式。邓路佳等^[21]将因子分解机与高斯混合模型进行结合,在高斯聚类之后应用因子分解机对聚类结果进行学习,构建广告点击率预测模型,获得了较好的实验结果。Yang等^[22]提出一个基于神经网络的推荐构架,通过嵌入标签的分布式特点表征特征之间的相关性,充分揭

示用户与物品之间的隐含交互关系,利用该隐含交互关系对用户进行推荐。Yu等^[23]将深度因子分解机模型应用到社交广告的个性化推荐中,其中,因子分解机部分负责提取低阶特征,深度神经网络部分负责提取高阶特征,通过两个层次的特征结合提升广告推荐效果。Lian等^[24]提出了一种新颖的压缩交互网络,其目的是以显式方式在矢量方式上生成特征交互。将压缩交互网络和经典深度神经网络组合到一个统一的模型中,并将该新模型命名为深度因子分解机(eXtreme deep factorization machine, xDeepFM)。一方面, xDeepFM能够显式学习某些边界度特征交互;另一方面,它可以隐式学习任意的低阶和高阶特征交互。Maxim等^[25]提出一种将矩阵分解、因子分解机等与深度学习进行结合提升推荐效果的个性化推荐模型。该模型设计了一种特殊的并行化方案,该方案利用嵌入表上模型的并行性减轻内存限制,同时利用数据并行性从全连接层进行横向扩展计算。Song等^[26]提出了一种有效且高效的方法,称为AutoInt,可以自动学习输入特征之间的高阶特征交互,将数值和分类特征映射到相同的低维空间中。然后,提出了带有残差连接的多头自注意神经网络,以对低维空间中的特征相互作用进行显式建模。利用多头自注意神经网络的不同层,可以对输入特征的特征组合的不同顺序进行建模。

从现有的研究可以看出,已经出现了一些考虑多层次潜在特征的混合推荐模型,但考虑潜在特征层次较少。为了更全面地挖掘用户与物品的潜在特征组合,提升推荐模型性能,需要将多种层次潜在特征考虑在内。

2 考虑多层次潜在特征的个性化推荐模型

考虑多层次潜在特征的个性化推荐模型框架如图1所示。以用户与物品的基本属性信息



构建用户-物品特征作为模型输入,通过 FM、WNN、CN、DNN 分别提取用户与物品不同层次的潜在特征,将各层次潜在特征进行融合以作为层级融合特征,将层级融合特征输入一个全连接层中进行评分预测。具体来说,所提出的推荐模型主要由 4 部分构成。

在第 1 部分中,将用户与物品的基本属性信息向量化表示为用户-物品联立特征向量。算法超体框架如图 1 所示,利用独热(one-hot)编码与多热(multi-hot)编码对用户与物品的基本信息编码,编码后得到高维稀疏向量。独热编码操作如图 1(a)所示,物品 ID 经编码后得到代表每个 ID 的高维稀疏向量,每个向量内仅有一维包含信息,而多热编码与独热编码主要区别在于其每个向量内可能有二维或二维以上包含信息。之后使

用嵌入层(embedding layer, EL)网络将高维稀疏向量映射为低维稠密向量。嵌入层网络进行降维的过程示意图如图 1(b)所示,将高维稀疏向量输入低维神经网络中所得即降维后的低维稠密向量。再对低维稠密向量进行拼接操作,组合成为用户-物品联立特征向量。

在第 2 部分中,提取用户-物品多层次的潜在特征。利用多种特征提取算法充分挖掘用户-物品多种层次潜在特征。FM 提取用户-物品低层次潜在特征, WNN 提取用户-物品低阶非线性潜在特征, CN 提取用户-物品线性交叉层次潜在特征, DNN 提取用户-物品高阶非线性层次潜在特征。

其中,FM 通过特征间的线性结合与两两交互提取用户-物品低层次潜在特征。如式(1),普

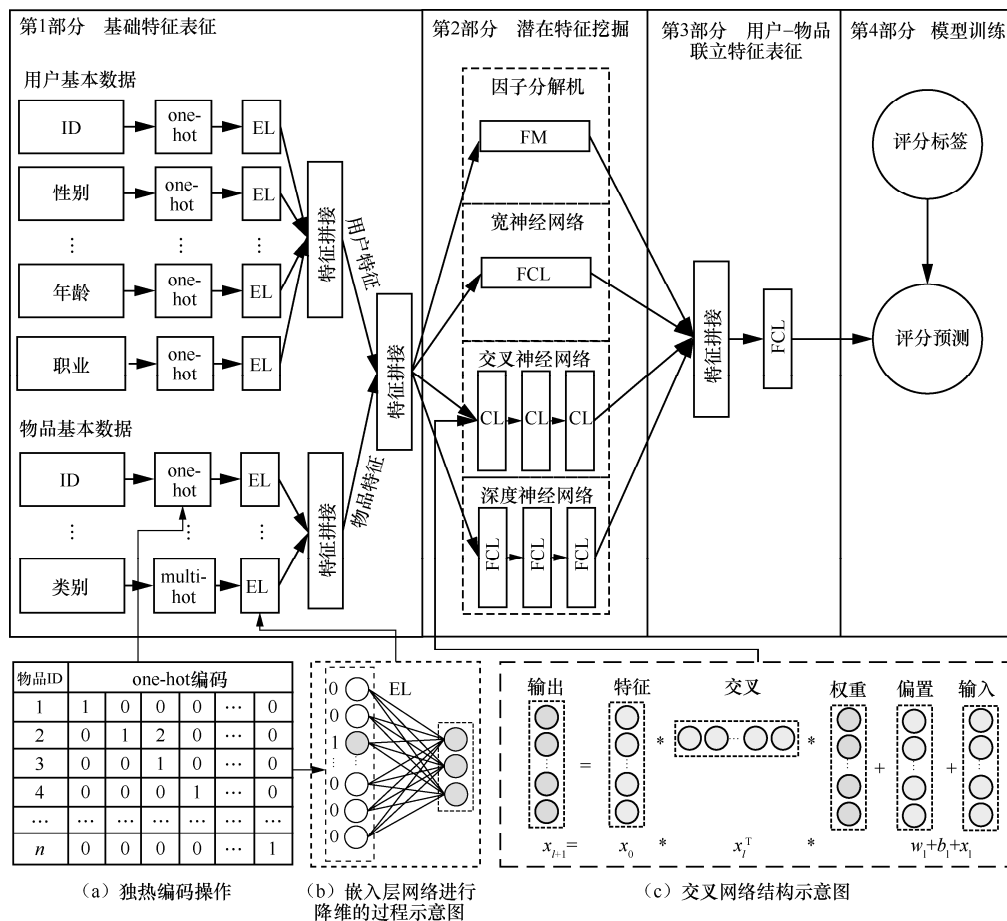


图1 个性化推荐模型框架

通的线性模型仅对各个特征进行加权求和, 忽略了特征和特征之间的潜在交互关系。FM 在线性结合的基础上考虑特征之间的关联计算, FM 算法的表达式如式 (2) 所示。

$$Y = w_0 + \sum_{i=1}^n w_i x_i \quad (1)$$

$$Y = w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n w_{ij} x_i x_j \quad (2)$$

其中, w_0 为常数偏置, n 为样本的特征个数。

由于直接计算 FM 算法中交叉项的时间复杂度与空间复杂度都极高, 故引入隐向量 $V_i = (v_{i1}, v_{i2}, \dots, v_{ik})$ 求解交叉项, 如式 (3) 所示。

$$V = \begin{pmatrix} v_{11} & v_{12} & v_{13} & \dots & v_{1k} \\ v_{21} & v_{22} & v_{23} & \dots & v_{2k} \\ v_{31} & v_{32} & v_{33} & \dots & v_{3k} \\ \dots & \dots & \dots & \dots & \dots \\ v_{n1} & v_{n2} & v_{n3} & \dots & v_{nk} \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \dots \\ v_4 \end{pmatrix} \quad (3)$$

用 $V_i V_j^T$ 对 w_{ij} 进行求解, 如式 (4) 所示。

$$\hat{W} = V V^T = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{pmatrix} \begin{pmatrix} v_1^T & v_2^T & \dots & v_n^T \end{pmatrix} \quad (4)$$

故 FM 算法二次项求解过程如式 (5) 所示。

$$\begin{aligned} & \sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j \\ &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \langle v_i, v_j \rangle x_i x_j - \frac{1}{2} \sum_{i=1}^n \langle v_i, v_i \rangle x_i x_i \\ &= \frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n \sum_{f=1}^k v_{i,f} v_{j,f} x_i x_j - \sum_{i=1}^n \sum_{f=1}^k v_{i,f} v_{i,f} x_i x_i \right) \quad (5) \\ &= \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{i,f} x_i \right) \left(\sum_{j=1}^n v_{j,f} x_j \right) - \sum_{i=1}^n v_{i,f}^2 x_i^2 \right) \\ &= \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{i,f} x_i \right)^2 - \sum_{i=1}^n v_{i,f}^2 x_i^2 \right) \end{aligned}$$

其中, $\langle v_i, v_j \rangle$ 为向量内积运算, 计算规则如式 (6) 所示。

$$\langle v_i, v_j \rangle = \sum_{f=1}^k v_{i,f} \cdot v_{j,f} \quad (6)$$

WNN 通过一个具有单个神经元的神经网络, 并在非线性激活函数的作用下提取用户—物品低阶非线性潜在特征。WNN 计算规则如式 (7) 所示。其中, Y_i 为网络输出, $f(\cdot)$ 为非线性激活函数, w_i 为网络权重, x_i 为输入, b_i 为偏置。

$$Y_i = f(w_i \cdot x_i + b_i) \quad (7)$$

CN 通过若干层交叉网络结构实现对输入特征进行交叉潜在特征提取。交叉网络结构示意图如图 1(c) 所示, 交叉网络层计算规则为原始输入与当前层网络输入交叉之后考虑权重, 再与当前层网络输入和偏置相加即为网络输出。特征交叉计算式如式 (8) 所示。其中, x_{l+1} 代表网络输出, x_0 代表原始输入, x_l 代表当前层网络输入, w_l 与 b_l 分别代表网络的权重与偏置。

$$x_{l+1} = x_0 x_l^T w_l + b_l + x_l = f(x_l, w_l, b_l) + x_l \quad (8)$$

DNN 由若干层全连接神经网络构成, 每层神经网络都经过非线性激活函数的作用, 多层全连接神经网络叠加, 以提取用户—物品高阶非线性层次潜在特征。每一层全连接神经网络的计算规则如式 (7) 所示。但 DNN 中每层全连接神经网络都具有多个神经元, 并非如 WNN 仅有单个神经元。

在第 3 部分中, 将用户—物品多种层次潜在特征进行融合, 并提取用户—物品层级融合潜在特征。在获取到用户—物品 4 种层次的潜在特征基础上, 将 4 种层次潜在特征进行拼接, 并送入一个全连接神经网络内, 以学习拼接向量中携带的用户—物品层级融合潜在特征。所应用全连接神经网络的计算规则同 DNN 内全连接神经网络计算规则。

在第 4 部分中, 进行模型训练与性能评估。将数据集划分为训练集与测试集, 利用训练集上



用户对物品的评分作为标签，以预测评分与实际评分之间的均方根误差作为损失函数，采用随机梯度下降的方法进行模型训练。在训练完成后使用测试集以预测评分与实际评分之间的标准差与均方根误差对模型性能进行评估。

3 实验分析与讨论

3.1 数据集和评价指标

为验证本文方法的有效性，采用 MovieLens-100k、MovieLens-1m、MovieLens- latest-small 和 FilmTrust 4 个常用的电影推荐数据集进行实验研究，实验数据集详情见表 1。这些数据集除包含评分信息，还包含用户与电影的基本属性信息，如用户 ID、用户年龄、用户性别、用户职业、电影 ID、电影类别等。

表 1 实验数据集详情

数据集	数据集大小	评分区间
MovieLens-100k	10 万余条评分，943 位用户，1 682 部电影	1~5
MovieLens-1m	100 万余条评分，6 040 位用户，3 952 部电影	1~5
MovieLens-latest-small	100 836 条评分，610 位用户，9 742 部电影	0.5~5
FilmTrust	35 497 条评分，1 508 位用户，2 071 部电影	0.5~4

本文中采用推荐系统中广泛应用的平均绝对误差（mean absolute error, MAE）与均方根误差（root mean square error, RMSE）衡量模型的性能，其计算式分别如式（9）、式（10）所示，其中， $f(x)_i$ 为预测评分， Y_i 为实际评分。MAE 与 RMSE 的值越小，说明算法性能越好。

$$MAE = \frac{1}{n} \sum_{i=1}^n |(f(x)_i - Y_i)| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f(x)_i - Y_i)^2} \quad (10)$$

另外采用数据稀疏度（sparsity rate, SR）指

标表示训练集稀疏性，训练集占比越小，SR 就越大，表明训练集数据就越稀疏，代表用户—物品评分矩阵中所含评分数据就越少，可用于训练的评分数据就越少，SR 计算式如式（11）所示。

$$SR = 1 - \frac{\text{评分总数}}{\text{用户数} \times \text{物品数}} \quad (11)$$

3.2 对比模型

本文选取对比模型包括：UBCF（user based collaborative filtering，基于用户的协同过滤）算法^[27]、DNN^[5]、FM（factorization machine，因子分解机）^[4]、WD（wide&deep，宽度&深度）^[7]、DeepFM^[8]和 DCN（deep & cross network，深度交叉网络）^[9]。具体介绍如下。

- UBCF：该算法属于传统推荐算法，仅采用用户对物品的评分信息对用户做出推荐，缺失对用户与物品基本信息的考虑。算法基本思想为根据用户对物品的历史评分构建用户画像，之后计算用户间相似度，发现与当前用户相似度较高的 K 个用户，并推荐这 K 个用户所偏好的物品。
- DNN：由若干层全连接神经网络构成，每层全连接神经网络都应用了非线性激活函数。
- FM：由 Rendle^[4]提出一种称为 FM 的特征交互提取模型。该模型结合了支持向量机的优势，与 SVM 相比，FM 使用分解参数对变量之间的两两交互进行建模，这样在数据稀疏的情况下仍然可以对特征之间的交互关系进行有效建模。
- WD：由 Cheng 等^[7]提出的一种名为 Wide & Deep 的推荐模型。在数据稀疏的情况，深度神经网络往往会忽略用户项目互动稀疏但相关性较高的情况。故该模型结合了宽度神经网络与深度神经网络的特点，对两种网络模型共享输入层进

行联合训练,以同时应用宽神经网络的强泛化性能与深度神经网络的强学习性能。

- **DeepFM**: 由 Guo 等^[8]提出的一种名为 Deep FM 的模型,该模型主要强调低阶特征和高阶特征交互的端到端学习。模型基本思路为,在新的神经网络体系结构中结合用于推荐的因子分解机和用于特征学习的深度神经网络,实现了低阶特征与高阶特征的交互。与 WD 模型相比,该模型将 wide 部分替换为 FM。
- **DCN**: 由 Wang 等^[9]提出的一种名为 DCN 的点击率预测模型。该模型提出了一种新颖的交叉网络结构,这种网络结构每一层的输出都由上一层网络与网络输入进行交叉得到,随着其层数的增加,特征之间的交叉越来越深。相对于 WD 与 Deep FM,该模型是深度神经网络与交叉网络的并行组合。

3.3 实验结果分析

为确定 SR 对模型性能的影响,首先将每个数据集都按不同比例进行训练集与测试集划分,在训练集占比不同的情况下对训练集的 SR 进行计算(见表 2),接着在不同数据集、不同的 SR 上分别计算对比模型与本文模型的预测性能。

表 2 不同数据集按比例划分训练集与测试集的 SR

训练集占比	80%	70%	50%	30%
MovieLens-100k	0.949 6	0.955 9	0.968 5	0.981 1
MovieLens-1m	0.964 3	0.968 7	0.977 7	0.986 6
MovieLens-latest-small	0.986 4	0.988 1	0.991 5	0.994 9
FilmTrust	0.990 9	0.992 0	0.994 3	0.996 5

表 3~表 10 分别给出了本文模型与所选对比模型在数据集 MovieLens-100k、MovieLens-1m、MovieLens-latest-small 和 FilmTrust 训练集占比不同(即不同数据稀疏度)情况的 MAE 与 RMSE 的实验结果(表中数据均为 20 次实验结果取平均值所得)。

表 3 数据集 MovieLens-100k 训练集占比不同(不同数据稀疏度)情况的 MAE

训练集占比	80%	70%	50%	30%
UB	0.825 4±0.022 7	0.838 5±0.013 2	0.842 9±0.017 9	0.882 3±0.027 2
FM	0.744 9±0.002 3	0.745 8±0.005 4	0.747 7±0.007 9	0.761 3±0.009 4
DNN	0.732 5±0.009 3	0.741 4±0.019 9	0.745 8±0.013 1	0.760 5±0.017 2
DeepFM	0.719 9±0.005 3	0.729 5±0.006 9	0.742 2±0.004 5	0.759 1±0.006 4
DCN	0.723 2±0.007 3	0.730 1±0.005 7	0.744 9±0.007 8	0.758 3±0.008 9
WD	0.719 4±0.005 4	0.728 9±0.004 9	0.741 1±0.009 5	0.759 3±0.010 1
本文模型	0.710 2±0.004 3	0.724 9±0.006 7	0.735 2±0.007 9	0.754 7±0.005 3

表 4 数据集 MovieLens-100k 训练集占比不同(不同数据稀疏度)情况的 RMSE

训练集占比	80%	70%	50%	30%
UB	1.028 3±0.034 1	1.043 3±0.017 3	1.054 2±0.021 5	1.094 3±0.029 5
FM	0.944 1±0.003 3	0.945 3±0.007 7	0.948 5±0.009 4	0.967 5±0.013 1
DNN	0.936 3±0.011 1	0.942 7±0.014 9	0.947 4±0.006 6	0.966 1±0.009 2
DeepFM	0.920 1±0.006 9	0.928 9±0.006 1	0.938 6±0.008 6	0.959 2±0.011 7
DCN	0.922 9±0.009 4	0.931 1±0.008 3	0.941 6±0.010 9	0.959 7±0.009 1
WD	0.919 4±0.006 8	0.927 3±0.009 1	0.937 9±0.014 2	0.958 1±0.007 4
本文模型	0.909 8±0.005 3	0.926 1±0.004 3	0.931 2±0.005 5	0.952 4±0.008 5



表 5 数据集 MovieLens-1m 训练集占比不同（不同数据稀疏度）情况的 MAE

训练集占比	80%	70%	50%	30%
UB	0.821 3±0.008 1	0.830 3±0.012 4	0.851 4±0.015 8	0.876 9±0.019 3
FM	0.674 5±0.018 2	0.675 6±0.019 6	0.706 8±0.021 4	0.734 2±0.025 1
DNN	0.692 6±0.019 7	0.701 8±0.020 3	0.719 4±0.024 6	0.738 5±0.026 8
DeepFM	0.672 8±0.007 8	0.672 7±0.009 1	0.688 4±0.012 6	0.706 8±0.013 4
DCN	0.673 6±0.014 1	0.673 3±0.015 7	0.689 5±0.017 3	0.714 5±0.016 2
WD	0.674 7±0.009 3	0.673 9±0.010 6	0.691 7±0.011 4	0.718 1±0.013 6
本文模型	0.663 5±0.011 7	0.668 1±0.012 9	0.685 2±0.013 5	0.702 3±0.015 1

表 6 数据集 MovieLens-1m 训练集占比不同（不同数据稀疏度）情况的 RMSE

训练集占比	80%	70%	50%	30%
UB	1.019 8±0.007 2	1.028 7±0.009 1	1.047 6±0.012 7	1.073 7±0.015 3
FM	0.862 3±0.017 8	0.864 7±0.019 4	0.885 4±0.021 3	0.915 8±0.023 8
DNN	0.877 4±0.015 4	0.886 4±0.018 3	0.897 2±0.020 6	0.923 8±0.026 8
DeepFM	0.856 5±0.008 7	0.862 3±0.009 5	0.873 4±0.010 7	0.899 1±0.011 3
DCN	0.857 2±0.005 7	0.863 7±0.008 6	0.875 9±0.007 4	0.903 4±0.009 1
WD	0.858 6±0.017 5	0.863 1±0.013 2	0.879 3±0.014 4	0.908 9±0.012 8
本文模型	0.849 3±0.019 7	0.857 2±0.014 5	0.872 6±0.015 9	0.891 2±0.013 7

表 7 数据集 MovieLens-latest-small 训练集占比不同（不同数据稀疏度）情况的 MAE

训练集占比	80%	70%	50%	30%
UB	0.924 7±0.053 3	0.935 0±0.046 2	0.951 7±0.041 7	0.968 4±0.038 9
FM	0.692 6±0.015 5	0.699 8±0.017 6	0.711 8±0.019 4	0.728 3±0.021 3
DNN	0.711 7±0.023 5	0.723 9±0.031 6	0.735 6±0.033 4	0.741 2±0.036 1
DeepFM	0.684 3±0.012 7	0.688 2±0.014 2	0.701 8±0.016 7	0.711 3±0.017 8
DCN	0.680 2±0.007 6	0.679 5±0.009 4	0.695 7±0.010 3	0.708 4±0.012 7
WD	0.682 2±0.011 8	0.684 7±0.013 7	0.697 3±0.015 6	0.709 8±0.017 5
本文模型	0.671 4±0.013 6	0.673 3±0.014 6	0.690 2±0.016 2	0.701 5±0.018 4

表 8 数据集 MovieLens-latest-small 训练集占比不同（不同数据稀疏度）情况的 RMSE

训练集占比	80%	70%	50%	30%
UB	1.016 0±0.068 3	1.022 6±0.053 9	1.036 4±0.058 6	1.057 6±0.049 3
FM	0.893 4±0.013 6	0.897 9±0.015 7	0.905 2±0.018 7	0.917 3±0.020 4
DNN	0.901 3±0.027 2	0.916 3±0.032 8	0.928 3±0.037 1	0.937 6±0.036 8
DeepFM	0.878 9±0.017 5	0.886 1±0.018 3	0.902 5±0.020 9	0.921 6±0.021 7
DCN	0.877 8±0.009 4	0.884 7±0.010 9	0.899 4±0.012 4	0.919 6±0.013 4
WD	0.875 9±0.013 3	0.881 4±0.014 7	0.901 6±0.016 2	0.920 2±0.017 9
本文模型	0.867 2±0.015 8	0.870 4±0.016 4	0.893 7±0.017 8	0.914 5±0.019 2

表 9 数据集 FilmTrust 训练集占比不同（不同数据稀疏度）情况的 MAE

训练集占比	80%	70%	50%	30%
UB	0.685 9±0.012 6	0.698 5±0.013 9	0.729 9±0.016 4	0.794 2±0.019 2
FM	0.639 1±0.013 4	0.645 9±0.014 7	0.660 8±0.017 3	0.689 3±0.018 6
DNN	0.651 2±0.022 7	0.659 2±0.025 4	0.678 4±0.026 2	0.691 5±0.029 3
DeepFM	0.630 9±0.011 9	0.634 5±0.013 5	0.661 3±0.014 9	0.674 1±0.016 1
DCN	0.633 8±0.012 5	0.637 3±0.013 9	0.665 9±0.015 7	0.678 6±0.017 3
WD	0.631 7±0.009 6	0.640 9±0.010 4	0.667 4±0.013 2	0.670 3±0.014 5
本文模型	0.621 4±0.011 3	0.634 8±0.012 6	0.655 0±0.014 8	0.665 3±0.015 8

表 10 数据集 FilmTrust 训练集占比不同（不同数据稀疏度）情况的 RMSE

训练集占比	80%	70%	50%	30%
UB	0.953 9±0.018 7	0.978 4±0.021 4	1.038 9±0.024 3	1.151 0±0.027 9
FM	0.831 6±0.017 3	0.846 2±0.019 8	0.857 7±0.020 7	0.889 4±0.023 2
DNN	0.839 3±0.023 1	0.843 8±0.025 5	0.866 8±0.027 1	0.898 7±0.029 6
DeepFM	0.810 3±0.011 7	0.823 1±0.012 8	0.838 9±0.013 5	0.846 7±0.015 2
DCN	0.811 7±0.010 7	0.825 6±0.011 6	0.841 9±0.012 9	0.854 6±0.014 7
WD	0.809 8±0.008 4	0.817 1±0.009 3	0.836 5±0.010 4	0.843 9±0.011 5
本文模型	0.800 2±0.010 8	0.812 4±0.011 2	0.827 4±0.012 6	0.837 6±0.013 9

由表 3~表 10 实验结果可知, 本文模型在预测评分与实际评分之间的 MAE 与 RMSE 上均有所提高。在训练集占比 80% 时, 模型在各个数据集上表现效果最佳。在数据集 MovieLens-100k 上, 相对于属于传统推荐算法的 UB, 在 MAE 上提升 11.52%, 在 RMSE 上提升 11.85%。相对于考虑单层次潜在特征的 FM 与 DNN, 在 MAE 上分别提升 3.58% 与 2.08%, 在 RMSE 上分别提升 3.45% 与 2.54%。相对于考虑两种层次潜在特征的 DeepFM、DCN 与 WD, 在 MAE 上分别提升 0.97%、1.3% 与 0.92%, 在 RMSE 上分别提升 1.03%、1.31% 与 0.96%。

在数据集 MovieLens-1m 上, 本文模型相对于属于传统推荐算法的 UB, 在 MAE 上提升 15.78%, 在 RMSE 上提升 17.05%。相对于考虑单层次潜在特征的 FM 与 DNN, 在 MAE 上分别提升 1.1% 与 2.91%, 在 RMSE 上分别提升

1.3% 与 2.81%。相对于考虑两种层次潜在特征的 DeepFM、DCN 与 WD, 在 MAE 上分别提升 0.93%、1.01% 与 1.12%, 在 RMSE 上分别提升 0.72%、0.79% 与 0.93%。

在数据集 MovieLens-latest-small 上, 本文模型相对于属于传统推荐算法的 UB, 在 MAE 上提升 25.33%, 在 RMSE 上提升 14.88%。相对于考虑单层次潜在特征的 FM 与 DNN, 在 MAE 上分别提升 2.12% 与 4.03%, 在 RMSE 上分别提升 2.62% 与 3.41%。相对于考虑两种层次潜在特征的 DeepFM、DCN 与 WD, 在 MAE 上分别提升 1.29%、0.88% 与 1.08%, 在 RMSE 上分别提升 1.17%、1.06% 与 0.87%。

在数据集 FilmTrust 上, 本文模型相对于属于传统推荐算法的 UB, 在 MAE 上提升 6.45%, 在 RMSE 上提升 15.37%。相对于考虑单层次潜在特征的 FM 与 DNN, 在 MAE 上分别提升

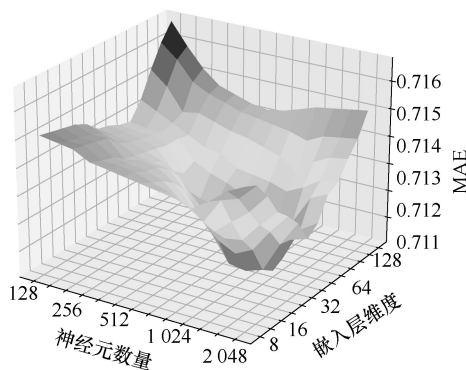


1.05%与2.98%，在RMSE上分别提升3.14%与3.91%。相对于考虑两种层次潜在特征的DeepFM、DCN与WD，在MAE上分别提升0.95%、1.24%与1.03%，在RMSE上分别提升1.01%、1.15%与0.96%。

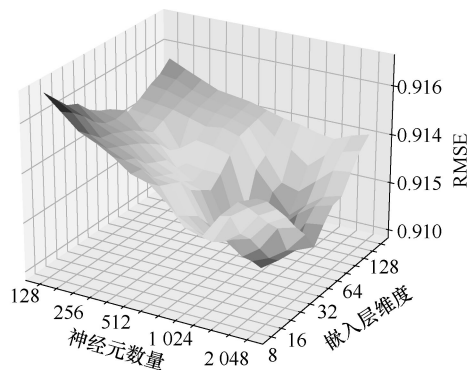
由表3~表10实验结果可以看出，WD模型的低阶非线性特征与高阶非线性特征的特征组合可以有效提升个性化推荐性能，本文模型在此基础上加入FM与交叉神经网络，进一步提取用户-物品之间低层次特征与线性交叉层次特征，将多种层次潜在特征进行拼接之后作为一个单层全连接神经网络的输入，使得个性化推荐性能在MAE与RMSE上均获得一定提升。在模型训练阶段，全连接神经网络中的权重会根据特征进行分配，重要特征的权重相对较高，因此不会造成重要特征被掩盖。

3.4 嵌入层维度与神经元数量对模型性能影响分析

此外，本文还分析了嵌入层维度与全连接层神经元数量对模型性能的影响。所提模型在不同数据集上选取不同嵌入层维度与不同神经元数量情况下MAE与RMSE的实验结果如图2、图3、图4、图5所示。由图中可以看出，随着嵌入层维度与图1第3部分中全连接层神经元数量的增加，模型效果有所改善，在嵌入层维度为32、神经元数量为1024时，模型效果最佳。但当二者过大时，模型效果变得较差。这表明在适当增大嵌入层维度与第3部分全连接层神经元的数量会使得模型更好地学习用户与物品原始编码中携带的信息与拼接向量中携带的用户-物品层级融合潜在特征，但是当二者过大时会导致模型训练效率低下并且训练结果较差。

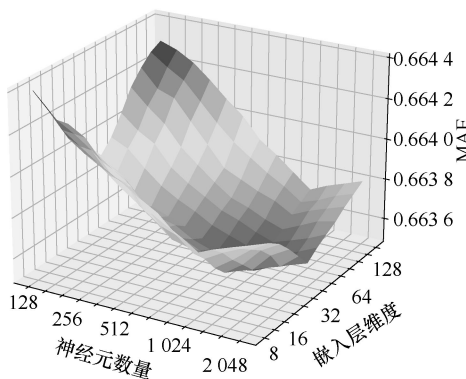


(a) MovieLens-100k MAE实验结果

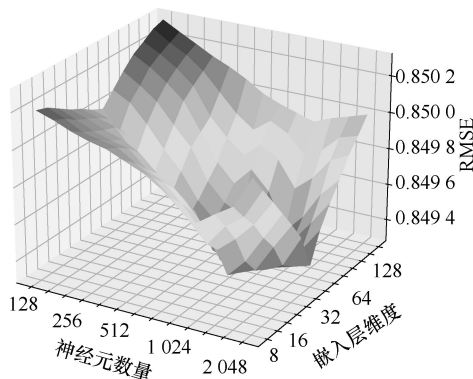


(b) MovieLens-100k RMSE实验结果

图2 MovieLens-100k 不同嵌入层维度与不同神经

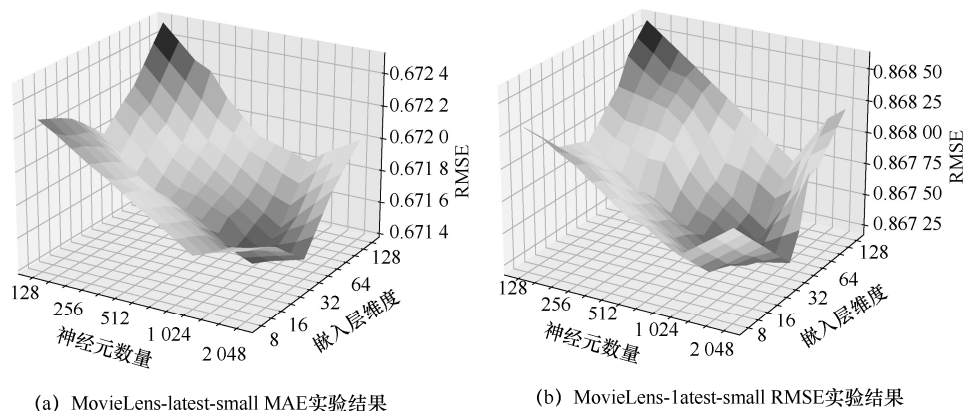


(a) MovieLens-1m MAE实验结果



(b) MovieLens-1m RMSE实验结果

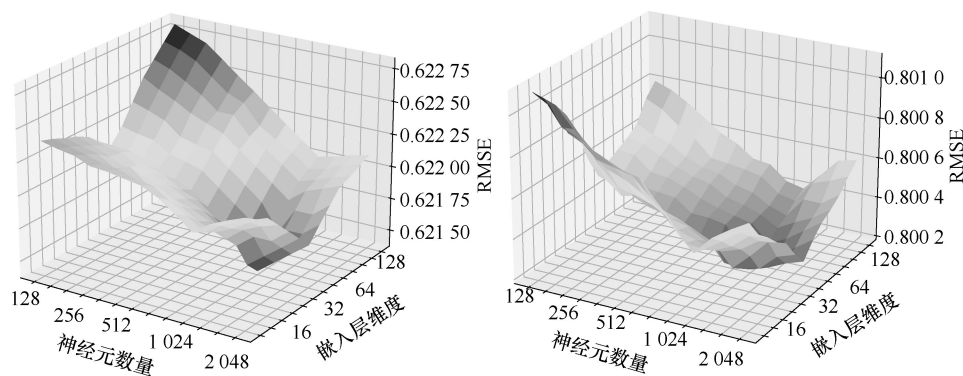
图3 MovieLens-1m 不同嵌入层维度与不同神经元数量下 MAE 与 RMSE 实验结果



(a) MovieLens-latest-small MAE实验结果

(b) MovieLens-latest-small RMSE实验结果

图4 MovieLens-latest-small 不同嵌入层维度与不同神经元数量下 MAE 与 RMSE 实验结果



(a) FilmTrust MAE实验结果

(b) FilmTrust RMSE实验结果

图5 FilmTrust 不同嵌入层维度与不同神经元数量下 MAE 与 RMSE 实验结果

本文模型与几种对比模型在不同数据集上单次实验所需时间的对比见表 11。从表 11 中可以看出, 本文模型所需时间略有提高, 其主要原因是本文模型考虑潜在特征的层次较多, 模型中特征提取算法数量较多, 但所需时间与其他模型差别不大, 基本上属于同一时间等级, 在模型性能提升较大的情况下, 时间消耗的增长在可接受范围内。

4 结束语

本文提出了一种考虑多层次潜在特征的个性化推荐模型。该模型结合 DeepFM、DCN、DM 三者提取不同层次的潜在特征的优点, 将多种层次的潜在特征考虑在内, 构建评分预测模型。为验证该模型的有效性, 将其与传统的推

表 11 模型与部分对比模型单次迭代所需时间对比 (单位: s)

数据集	MovieLens-100k	MovieLens-1m	MovieLens-latest-small	FilmTrust
FM	68.09	230.63	54.61	13.62
DNN	66.77	273.29	57.34	20.15
DeepFM	93.61	411.64	89.38	23.81
DCN	74.82	400.74	91.97	21.39
WD	65.68	381.83	87.65	20.34
本文模型	123.41	987.35	113.47	33.84



荐算法和考虑潜在特征的混合推荐模型进行对比,在4个公开数据集上对模型性能进行评估,实验结果证明该模型可以有效提升推荐性能。同时实验还评估了对本文模型性能有所影响的两个参数,不同的嵌入层维度与模型第三部分全连接层神经元数量。本文对模型的时间复杂度以模型单次迭代时间方式进行了评估,考虑了较多的潜在层次,导致模型时间复杂度较高。后续研究将对不同层次潜在特征的组合方式进行研究,使得层级融合特征向量内可以更好地包含多种层次潜在特征,且考虑更多层次的潜在特征提取,并尝试模型不同部分之间的参数共享,以降低模型复杂度。

参考文献:

- [1] 项亮. 推荐系统实践[M]. 北京: 人民邮电出版社, 2012.
XIANG L. The Development of recommendation system[M]. Beijing: Posts & Telecom Press, 2012.
- [2] BPRATTOL, CARTAS, FENUG, et al. Semantics-aware content-based recommender systems: Design and architecture guidelines[J]. Neurocomputing, 2017(254): 79-85.
- [3] YANG S, KORAYEM M, ALJADDA K, et al. Combining content-based and collaborative filtering for job recommendation system: a cost-sensitive statistical relational learning approach[J]. Knowledge-Based Systems, 2017(136): 37-45.
- [4] RENDLE S. Factorization machines[C]//Proceedings of 2010 IEEE International Conference on Data Mining. Piscataway: IEEE Press, 2010: 995-1000.
- [5] ZHANG L B, LUO T J, ZHANG F, et al. A recommendation model based on deep neural network[J]. IEEE Access, 2018, 6: 9454-9463.
- [6] BI J W, LIU Y, FAN Z P. A deep neural networks based recommendation algorithm using user and item basic data[J]. International Journal of Machine Learning and Cybernetics, 2020, 11(4): 763-777.
- [7] CHENG H T, KOC L, HARMSSEN J, et al. Wide & deep learning for recommender systems[C]//DLRS 2016: Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. [S.l.:s.n.], 2016: 7-10.
- [8] GUO H F, TANG R M, YE Y M, et al. DeepFM: a factorization-machine based neural network for CTR prediction[C]//Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. California: International Joint Conferences on Artificial Intelligence Organization, 2017.
- [9] WANG R X, FU B, FU G, et al. Deep & cross network for ad click predictions[C]//ADKDD'17: Proceedings of the ADKDD'17. [S.l.:s.n.], 2017: 1-7.
- [10] GUILHERME BRANDAO MARTINS, JOAO PAULO PAPA, HOJJAT ADELI. Deep learning techniques for recommender systems based on collaborative filtering[J]. Expert Systems, 2020, 37(6): 1-21.
- [11] MYLAVARAPU B K. Collaborative filtering and artificial neural network based recommendation system for advanced applications[J]. Journal of Computer and Communications, 2018, 06(12): 1-14.
- [12] TANG H, LEI M, GONG Q, et al. A BP neural network recommendation algorithm based on cloud model[J]. IEEE Access, 2019(7): 35898-35907.
- [13] 薛峰, 刘凯, 王东, 等. 基于深度神经网络和加权隐反馈的个性化推荐[J]. 模式识别与人工智能, 2020, 33(4): 295-302.
XUE F, LIU K, WANG D, et al. Personalized recommendation algorithm based on deep neural network and weighted implicit feedback[J]. Pattern Recognition and Artificial Intelligence, 2020, 33(4): 295-302.
- [14] 杨洁, 朱咸军, 周献中, 等. 基于混杂社会网络的个性化Web服务推荐方法[J]. 电子学报, 2020, 48(2): 341-349.
YANG J, ZHU X J, ZHOU X Z, et al. Personalized web service recommendation based on heterogeneous social network[J]. Acta Electronica Sinica, 2020, 48(2): 341-349.
- [15] HUI B, ZHANG L, ZHOU X, et al. Personalized recommendation system based on knowledge embedding and historical behavior[J]. Applied Intelligence, 2022(52): 954-966.
- [16] ZHEN W, ALLAM M, LIANG M B. Research on e-commerce personalized recommendation system based on big data technology[C]//Proceedings of 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA). Piscataway: IEEE Press 2021, 909-913.
- [17] 邵英玮, 张敏, 马为之, 等. 融合商品潜在互补性发现的个性化推荐方法[J]. 软件学报, 2020, 31(4): 1090-1100.
SHAO Y W, ZHANG M, MA W Z, et al. Integrating latent item-item complementarity with personalized recommendation systems[J]. Journal of Software, 2020, 31(4): 1090-1100.
- [18] FU M S, QU H, YI Z, et al. A novel deep learning-based collaborative filtering model for recommendation system[J]. IEEE Transactions on Cybernetics, 2019, 49(3): 1084-1096.
- [19] YAN W J, WANG D, CAO M J, et al. Deep auto encoder model with convolutional text networks for video recommendation[J].

- IEEE Access, 2019(7): 40333-40346.
- [20] TANG J X, WANG K. Personalized top-N sequential recommendation via convolutional sequence embedding[C]// Proceedings of WSDM '18: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. [S.l.:s.n.], 2018: 565-573.
- [21] 邓路佳, 刘平山. 基于 GMM-FMs 的广告点击率预测研究[J]. 计算机工程, 2019, 45(5): 122-126.
- DENG L J, LIU P S. Research on click-through rate prediction of advertisement based on GMM-FMs[J]. Computer Engineering, 2019, 45(5): 122-126.
- [22] YANG D Q, CHEN L H, LIANG J Q, et al. Social tag embedding for the recommendation with sparse user-item interactions[C]//Proceedings of 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). Piscataway: IEEE Press, 2018: 127-134.
- [23] 郁豹, 李振华, 张凯, 等. 基于 DeepFM 模型的广告推荐系统研究[J]. 计算机应用与软件, 2019, 36(7): 307-310, 316.
- YU B, LI Z H, ZHANG K, et al. Advertisement recommendation system based on deepfm model[J]. Computer Applications and Software, 2019, 36(7): 307-310, 316.
- [24] LIAN J X, ZHOU X H, ZHANG F Z, et al. xDeepFM: combining explicit and implicit feature interactions for recommender systems[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2018.
- [25] NAUMOV M, MUDIGERE D, SHI H J M, et al. Deep learning recommendation model for personalization and recommendation systems[EB]. 2019.
- [26] SONG W P, SHI C C, XIAO Z P, et al. AutoInt: automatic feature interaction learning via self-attentive neural networks[C]// Proceedings of CIKM '19: Proceedings of the 28th ACM International Conference on Information and Knowledge Management. [S.l.:s.n.], 2019: 1161-1170.
- [27] HONGTAO X, QINGSHENG Z, HONGCHUN Q, et al.

User-based collaborative recommendation filtering algorithm using extremely valued ratings[J]. International Journal of Digital Content Technology and Its Applications, 2011, 5(9): 47-54.

[作者简介]



申倩 (1982-), 女, 湖州学院副教授, 主要研究方向为个性化推荐、多目标优化、智能决策等。



郭文宾 (1996-), 男, 湖州师范学院信息工程学院硕士生, 主要研究方向为个性化推荐。



楼俊钢 (1982-), 男, 博士, 湖州师范学院信息工程学院教授, 主要研究方向为智能信息处理、个性化推荐。



余强国 (1977-), 男, 湖州学院理工学院高级工程师, 主要从事模式识别和智能控制等。