

基于注意力机制的协同卷积动态推荐网络

汤文兵¹ 任正云¹ 韩芳¹

摘要 一直以来, 各种推荐系统关注于如何挖掘用户与物品特征间的潜在关联, 特征信息的充分利用有利于用户到物品的精准匹配. 基于矩阵分解和分解机的推荐算法是该领域的主流, 前者学习用户历史行为而后者分析对象特征关系, 但都难以兼顾用户行为与个体特征. 而近年来, 深度神经网络凭借其强大的特征学习能力和灵活可变的结构被应用到了推荐系统领域. 鉴于此, 本文提出了一种基于注意力机制的协同卷积动态推荐网络 (Attentionbased collaborative convolutional dynamic network, ACCDN), 它通过注意力机制实现用户历史行为、用户画像与物品属性的多重交互, 再通过卷积网络逐层捕捉更高阶的特征交互. 网络同时接受不同组块输出的低阶至高阶信息, 最后给出用户对指定物品青睐评分概率的预估. 而且本文还提出了一种基于无参时间衰减的用户兴趣标签来量化用户关注的变化. 通过比较若干先进模型在两个现实数据集的表现, 本文设计的动态推荐模型不但能够缓解推荐时滞性, 还能明显提高推荐质量, 为用户带来更好的个性化服务体验.

关键词 注意力机制, 2D 卷积网络, 动态推荐, 多重特征交互

引用格式 汤文兵, 任正云, 韩芳. 基于注意力机制的协同卷积动态推荐网络. 自动化学报, 2021, 47(10): 2438–2448

DOI 10.16383/j.aas.c190820

Attention-based Collaborative Convolutional Dynamic Network for Recommendation

TANG Wen-Bing¹ REN Zheng-Yun¹ HAN Fang¹

Abstract A variety of recommender systems focus on connecting users' and item' s features, and the full use of the potential information presented by features contributes to the accurate user-item matching. Recommendation algorithms based on matrix factorization and factorization machine have risen to the backbone in this field, with the former learning the users' behavior in history and the latter analyzing the individual features, but neither of them are taken into account by algorithms. Recently, deep neural networks are widely applied to recommendation because of their powerful representation learning and flexible structures. Thus, this paper proposes a novel attention-based collaborative convolutional dynamic network (ACCDN) for recommendation, which realizes the multiple interactions of the user behavior, user profiles and item attributes through attention mechanism and models higher-order interactions by convolution layers. And it absorbs all the low-to-high-order interaction features to predict the user preferable rating probability given an item. In addition, this paper puts forward the unique user interest label based on non-parametric time decay, as an auxiliary tool, to quantify the change of user focus. We conduct extensive experiments on two real-world datasets, which justifies that our proposed ACCDN not only alleviates the recommendation lag, but also improves the quality of recommendation significantly, bringing a better personalized service experience.

Key words Attention mechanism, 2D convolution network, dynamic recommendation, multiple feature interactions

Citation Tang Wen-Bing, Ren Zheng-Yun, Han Fang. Attention-based collaborative convolutional dynamic network for recommendation. *Acta Automatica Sinica*, 2021, 47(10): 2438–2448

信息时代下的数据量快速膨胀, 高渗透、高过载海量信息的冲击过渡稀释了用户注意力, 导致用户们在琳琅满目的产品面前难以抉择. 因此推荐系

统作用日趋凸显, 它准确而高效地跨接关系紧密的用户群和物品集, 放大长尾效应来维持供需平衡, 增大经济效益的同时提供智能化、专属化的用户体验, 实现从企业到客户的双赢.

特征的选择和处理方式在整个推荐过程中扮演着极其重要的角色. 根据特征处理方式的不同, 可分为基于内容、协同过滤和混合模型推荐系统^[1-2]. 狭义的特征一般指能够勾勒出用户或物品个体的定量描述, 是客观存在的, 多见于基于分解机^[3-4]的推荐模型; 而本文认为特征在广义上还应包含用户主

收稿日期 2019-12-03 录用日期 2020-03-25

Manuscript received December 3, 2019; accepted March 25, 2020

国家自然科学基金 (11572084, 11972115) 资助
Supported by National Natural Science Foundation of China (11572084, 11972115)

本文责任编辑 张民

Recommended by Associate Editor ZHANG Min

1. 东华大学信息科学与技术学院 上海 201620

1. College of Information Science and Technology, Donghua University, Shanghai 201620

观对物品产生的历史交互信息, 这也是矩阵分解类模型^[5-6]的主要处理对象. 推荐系统在掌握这些特征规律后, 预测用户行为, 提前替他们在众多候选物中筛选出“心仪”的产品. 因而若能够提取用户-物品对的特征关系, 并实时跟踪用户行为, 丰富信息的结合无疑会使推荐结果更加准确、更加个性化.

1 相关工作

早期多是基于内容的推荐系统, 它根据用户频繁浏览的物品类型给出推荐, 推荐方式简便直接但推荐内容单调乏味, 常见于广告和新闻平台^[7]. 而热门至今的协同过滤算法思想大为不同. 以基于用户推荐为例^[8], 先圈定行为近似的用户群, 再将他们关注过的物品推荐给彼此. 不少协同过滤模型通过用户显式反馈, 如收藏、转发和评价等, 来决定物品是否值得推荐. 而后更多平台提供了评分系统, 便于用户量化表达自己的喜好程度, 推荐过程因而变成了回归问题, 根据用户历史反馈预测评分, 并将可能得到高分的物品推荐出去. 于是矩阵近似盛行起来, 将高度稀疏的用户-物品评分矩阵分解为若干低阶矩阵的联合表示, 实现基于少数观测行为的评分预测. 其中以 Koren 等^[5-6]先后提出的 SVD++ (Singular value decomposition ++)、矩阵分解, Salakhutdinov 等^[9]基于极大似然估计的 PMF (Probabilistic matrix factorization) 和 Lee 等^[10]的局部低秩矩阵近似 (Local low-rank matrix approximation, LLORMA) 最为经典. 后来多用于图像处理的自编码器也能实现近似矩阵分解的功能, 如 AutoRec^[11]通过编码层获得低维向量, 解码层借此还原出原有评分的同时“填补”未知评分. Strub 等^[12]和 Wu 等^[13]对输入添加掩膜、高斯等噪声, 得到了系统鲁棒性更好的降噪自编码器. 这是自受限玻尔兹曼机^[14-15]后多层感知器 (Multi-layer perceptron, MLP) 在该领域的又一应用突破^[16].

以矩阵近似为代表的回归类推荐模型虽能在较准确还原观测数据的同时实现评分预测, 但往往因为缺乏利用甚至摒弃用户和物品特征而缺乏解释性, 且该类模型面对高稀疏性数据时通常表现不佳. 而 Rendle 提出的分解机 (Factorization machine, FM)^[3]着力分析不同特征间的关联度, 为指定用户推荐更贴切的物品, 同时还能缓解用户冷启动. 近年来, 深度神经网络 (Deep neural networks, DNNs) 在自然语言处理和计算机视觉的广泛应用也推动了推荐模型向深层结构的方向发展, 结合了上述方法与 DNNs 的混合模型在性能上表现更好. 先有 Strub 等于各层注入辅助特征优化 CDAE (Collaborative denoising autoencoders) 网

络^[12, 16-17], 后有携程团队结合用户和物品双自编码器支线实现矩阵分解^[18-19]. He 等认为低维向量级联的效果比做内积更好, 并让网络学习不同维度的隐因子向量组合, 提出了 NCF^[20] (Neural collaborative filtering) 混合模型. 特征处理方面, Google 团队认为 DNNs 特征学习能力很强但是特征记忆力较差, 提出了综合能力更优的深宽网络^[21]; 近年出现的 DeepFM (Deep factorization machine) 能比传统分解机学习更高阶的特征交互^[3, 22]. 而随着用户和物品特征字段的增多, DNNs 会学习更多无意义的特征交互信息, Lian 等就该问题提出了 xDeepFM^[23].

此外卷积神经网络 (Convolutional neural networks, CNNs) 和循环神经网络 (Recurrent neural networks, RNNs) 等深层架构被移用到了各自擅长的推荐应用中. Kim 等将 CNNs 嵌入 PMF 中实现基于评论文档分析的推荐^[9, 24-25], Covington 等将 DNNs 用于 Youtube 视频推荐, 成功接受千万级特征和千亿级训练样本^[26]. 而擅长处理时序任务的 RNNs 多用于音乐、新闻等推荐情景^[27-29]. 还不断涌现出特征处理巧妙、性能优良的模型变体, 比如协同过滤生成式对抗网络, 3D 卷积分解机 (Convolutional factorization machines, CFM), 协同记忆网络和处理特征不确定性的卷积网络等^[30-37].

然而相关研究工作中鲜有模型同时关注用户个体特征、行为习惯和物品属性特点三方面实现推荐工作. 因此本文提出了一种模型解释性更强、信息综合更多的基于注意力机制的协同卷积动态推荐网络, 具体工作和优势概括如下:

- 1) 提出了一种无参时间衰减的用户兴趣标签来量化用户关注角度随时间的变化, 给推荐候选物加入与推荐时间节点相关的奖惩机制, 加强推荐动态性;
- 2) 通过注意力机制实现用户行为、用户画像和物品属性的多重交互, 分析不同特征组间相关性, 更有效地捕捉有价值的高阶交互信息;
- 3) 引入卷积网络同时处理低阶原始特征信息和高阶特征交互信息, 让模型在尽可能多保留低阶信息的同时学习更高阶特征, 加速训练收敛;
- 4) 使用两个现实数据集来测试本文模型的推荐效果, 并与其他性能优良的推荐系统做横向比较, 结果显示本文模型在推荐质量和个性化上表现得更好.

2 模型准备

2.1 基于无参时间衰减的用户兴趣标签

物品类型越符合用户的口味, 用户就越容易对

其产生兴趣. 而物品往往会被贴上若干标签, 一类是本文关注的物品分类标签; 另一类是用户们评价留下的标签, 并衍生出了基于标签的推荐系统^[38-39]. 如今电商、视频、新闻资讯等平台都建议用户在注册时选择偏好类型以供推荐参考, 这能有效地解决用户冷启动.

但值得注意的是, 用户的兴趣会转移、关注的内容会变化, 或快或慢, 取决于情感、心理等主观因素, 也与时事热点、亲朋好友等客观环境有关. 于是结合用户注册的偏好类型和他们历史浏览的物品类型, 本文提出了一种可行的时间衰减算法来跟踪用户兴趣的变化. 首先, 规定用户兴趣和物品类别共享标签, 圈定有限的标签集合记 $C = \{c_1, c_2, \dots, c_n\}$. 标签分类可借助聚类、关联度分析等机器学习算法, 而有的数据集直接给出了分类标签.

假定用户 u 关于物品集的第一条历史操作 (评价、转发、收藏等) 的时间记为 t_{u1} , 最新的一条历史操作时间记为 t_{u2} , 则他的活跃时间跨度为 $\Delta T_u = t_{u2} - t_{u1}$. 推荐节点按月计整数更新, 更新速率可按照具体推荐情景调整. 接下来, 指定用户 u 遍历物品集 \mathcal{S} , $\kappa(u)$ 记录用户 u 留下历史痕迹的物品索引, $\bar{\kappa}(u)$ 则记录用户 u 从未接触过的物品, 有 $\kappa(u) \cup \bar{\kappa}(u) = \mathcal{S}$. 而后记用户兴趣标签向量为 $\mathbf{p}_u = [p_{u,1}, p_{u,2}, \dots, p_{u,n}]$, 而物品的分类标签向量记为 $\mathbf{c}_i = [c_1, c_2, \dots, c_n]$, 且 $\mathbf{p}_u, \mathbf{c}_i \in \mathbf{R}^n$. 则在用户历史行为遍历过程中 \mathbf{p}_u 更新方法如下:

$$\mathbf{p}_u = \mathbf{p}_{u,0} + \sum_{i \in \kappa(u)} \mathbf{c}_i \cdot e^{-\frac{t_{u2} - t_{u1}}{\Delta T_u}}, \quad t_{u,i} \in [t_{u1}, t_{u2}], \quad (1)$$

$$\text{Norm}(\mathbf{p}_u) = \frac{\mathbf{p}_u}{\|\mathbf{p}_u\|_1} \quad (2)$$

其中, $t_{u,i}$ 表示用户 u 对物品 i 产生操作的时间. 若用户于注册时提供了偏好信息则 $\mathbf{p}_{u,0}$ 为非零向量, 为初期推荐提供了引导和基准, 缓解了用户冷启动问题; 反之, 若用户不提供任何先验兴趣信息时 $\mathbf{p}_{u,0} = \mathbf{0}$, 则 \mathbf{p}_u 的更新将完全取决于注册后的用户行为. 最后, 在每个推荐节点按式 (12) 对 \mathbf{p}_u 元素归一化到 $[0,1]$ 区间, 如此每个元素值就量化表示用户对各类物品喜爱的热度和概率, 作为用户画像的一部分输入网络.

图 1 为例, 用雷达图可视化某月某位用户的兴趣向量, 通过 $\text{Norm}(\mathbf{p}_u) \cdot \mathbf{c}_i, i = 1, 2, 3, 4$ 计算用户与 4 部电影 的向量内积值, 依次为 0.37, 0.45, 0.58, 0.33, 显然电影 3 可能更符合用户的口味.

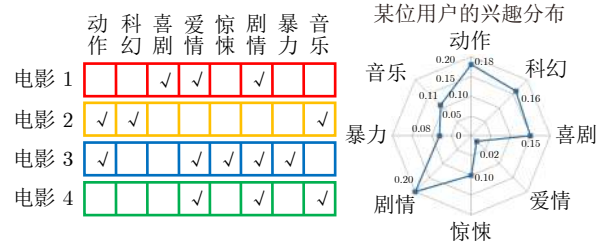


图 1 用户兴趣和物品 (如电影) 分类向量空间共享示意图
Fig.1 The demonstration of shared vector space between user's interests and item's types (e.g. movies)

本节提出的用户兴趣标签更新法节省了传统时间衰减法需要的超参调试, 让推荐系统考虑用户兴趣动态变化的同时缓解冷启动: 若用户不对新类别产生行为, 则衰减不起效, 维持原有爱好热度; 若用户关注了不同以往的新类别, 则增加新兴趣的热度, 此增彼减. 式 (1) 中的累加使 \mathbf{p}_u 记忆用户兴趣史, 基于自然常数的平滑衰减使得推荐更具平稳的动态性.

2.2 注意力机制: 实现特征嵌入向量的交互

分解机类的推荐模型^[3-4, 19, 22-23, 30-31] 专注于处理用户和物品特征的交互: 这些特征以 one-hot 或 multi-hot 字段的形式输入模型 (示例见图 2), 并通过嵌入层索引到各自对应的稠密嵌入向量, 它们才是分解机处理的真实对象. 典型的分解机^[3, 22] 逐对处理输入特征, 分析特征对 x_i 和 x_j 的关系, 得到推荐依据:

$$y_{FM} = \omega_0 + \sum_{i=1}^n \omega_1 x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n x_i x_j \cdot \text{dot}(\mathbf{e}_i, \mathbf{e}_j) \quad (3)$$

其中, $\langle \mathbf{e}_i, \mathbf{e}_j \rangle$ 是特征对 $\langle i, j \rangle$ 对应的嵌入向量对, 而 ω_0 和 ω_1 分别是全局偏置和一阶权重.

本文中注意力机制实现特征嵌入向量间的交互, 注意力机制是机器翻译的主要工具^[40-41], 它通过分析当前单词与上下文内容的关联程度, 将当前单词的注意力以不同的权重分配到上下文词句. Christakopoulou 等在^[32] 中提及同一件物品对于不同用户意义不同, 不同背景 (职业、收入、学历等)、人口统计信息 (年龄、性别等) 和心理 (爱好、性格

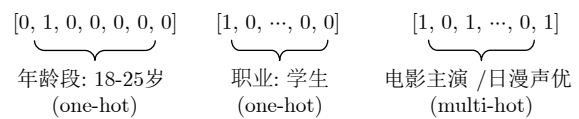


图 2 稀疏特征字段示例
Fig.2 The example of sparse feature fields

等). 因此, 本文认为注意力机制的思想同样适用于分析物品任一属性与用户个体的关系.

图 3 假设用户 u 对应 M 个字段的特征嵌入向量, 而物品 i 对应 N 个, 可构建用户画像矩阵 $E_M^u = [e_1^u; e_2^u; \dots; e_M^u]^T$ 和物品属性矩阵 $E_N^i = [e_1^i; e_2^i; \dots; e_N^i]^T$, 其中 $E_M^u \in \mathbf{R}^{M \times K}$, $E_N^i \in \mathbf{R}^{N \times K}$, K 是特征嵌入向量的维度. 则嵌入向量对 $\langle e_m^u, e_n^i \rangle$ 的注意力系数计算如下:

$$\alpha_{m,n}^{u,i} = e_m^u \cdot e_n^i, \quad m \in [1, M], n \in [1, N], m, n \in \mathbf{N}^* \quad (4)$$

而物品第 n 个属性 e_n^i 关于用户 u 的注意力向量为:

$$a_n^{i,u} = \sum_{m=1}^M \frac{e^{\alpha_{m,n}^{u,i}}}{\sum_{m=1}^M e^{\alpha_{m,n}^{u,i}}} \cdot e_m^u, \quad a_n^{i,u} \in \mathbf{R}^K \quad (5)$$

3 模型介绍: 基于注意力机制的协同卷积动态推荐网络 (ACCDN)

介绍完本文模型所需的动态用户兴趣标签更新法和注意力机制特征交互后, 下文将按照网络输入、注意力模块、卷积层和用户推荐等模块顺序介绍本文 ACCDN 模型框架.

3.1 网络输入

模型输入由用户、物品特征和用户行为值 (历史评分) 三部分组成. 不同数据集支持的评分选项不统一, 以本文实验所用数据集为例, MovieLens 评分 $r_{u,i} \in [1, 2, 3, 4, 5]$, 而 Niconico 评分项多达 10 个, 更有推荐数据集支持小数评分, 这不利于模

型学习用户在不同评分系统中的习惯. 因此评分需按式 (6) 做均值归一化, 作为用户 u 对物品 i 历史行为的量化值 $h_{u,i}^* \in [-1, 1]$.

$$h_{u,i}^* = \frac{r_{u,i} - \text{mean}(r_u)}{\text{std}(r_u)} \quad (6)$$

之所以这么做是考虑了不同用户给分的严苛度是不一样的, 以两种给分极端的人为例 (评分范围 1 ~ 5 分): 第一种人给分严苛且平均评分为 2.3, 那么预测得到 4 分的物品很可能被列入推荐列表内; 第二种人给分很宽松且平均评分为 4.2, 而这时 4 分的物品就不建议纳入推荐范畴, 模型需要寻找更多能够得到 5 分的物品作为推荐. 而这正是那些依赖于特征做预测的分解机类推荐系统所忽视的重要的个性化问题, 毕竟起初符合口味的物品在用户体验了解后也不一定能获得好评.

3.2 注意力机制模块

本文设计的注意力层重在实现用户历史行为 (如评分)、用户画像和物品属性的交互融合. 由于 one-hot 特征与嵌入向量是一一对应的关系, 而 multi-hot 特征与嵌入向量是一对多的关系, 因而对于不确定的一对多关系按式 (7) 加权融合为一个向量, 这种情况多出现于影视作品的主演、配音和风格等字段的特征.

$$e_n = \omega_{n,1} e_{n,1} + \omega_{n,2} e_{n,2} + \dots + \omega_{n,k} e_{n,k} \quad (7)$$

这样能保证 N 个特征字段在通过嵌入层后仍对应到 N 个嵌入向量, 式 (7) 中的权重 $\omega_{n,k}$ 将由网络学习. $h_{u,i}^*$ 会在特征交互中起到辅助作用, 因 $h_{u,i}^* \in [-1, 1]$, 当 $h_{u,i}^*$ 为正值时将增大用户、物品特征对之间的注意力系数, 反之将反向削弱, 于是

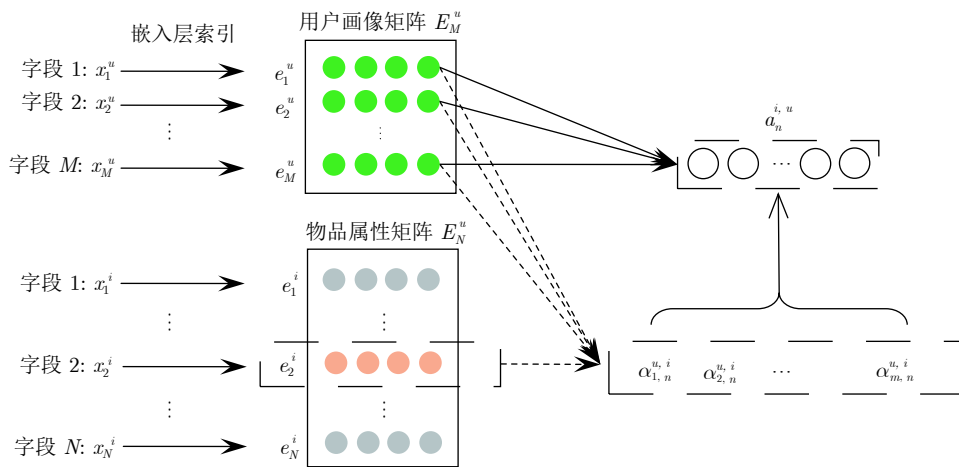


图 3 基于注意力机制的用户-物品属性交互

Fig. 3 Attentive interactions between the user's profile and item's attribute

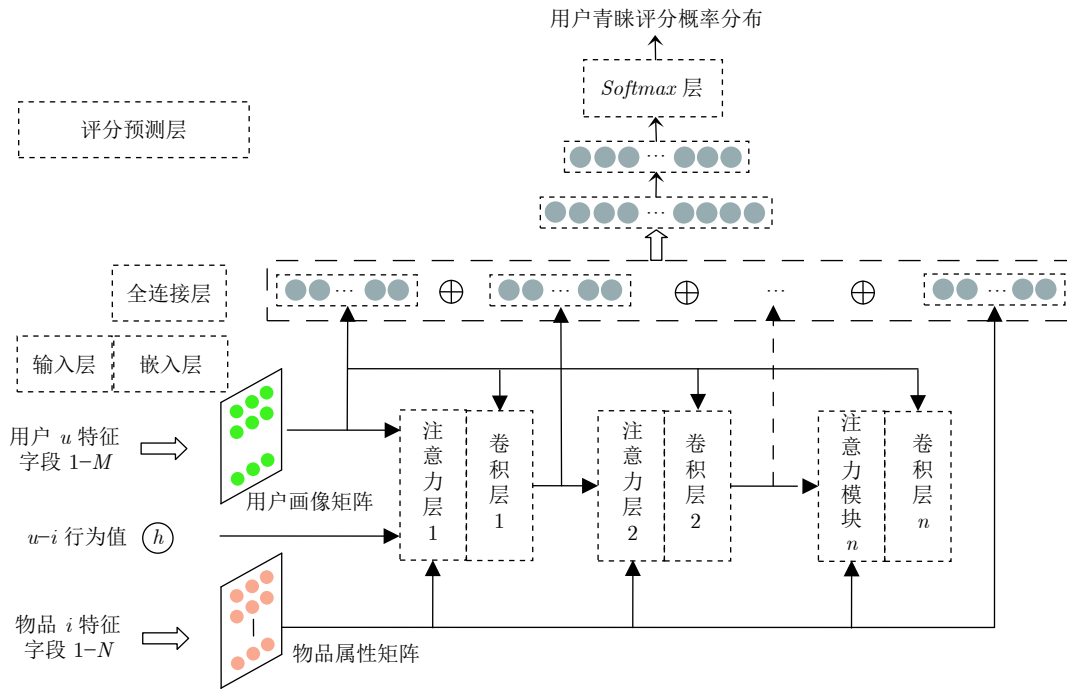


图 4 基于注意力机制的协同卷积动态推荐网络 (ACCDN)

Fig.4 Attention-based collaborative convolutional dynamic network (ACCDN)

式 (4) 和式 (5) 改写如下:

$$\mathbf{a}_n^{i,u} = \sum_{m=1}^M \frac{e^{\text{dot}(h_{u,i}^*, e_m^u, e_n^i)}}{\sum_{m=1}^M e^{\text{dot}(h_{u,i}^*, e_m^u, e_n^i)}} \cdot e_m^u, \quad \mathbf{a}_n^{i,u} \in \mathbf{R}^K \quad (8)$$

$$A^{u,i} = [\mathbf{a}_1^{i,u}; \mathbf{a}_2^{i,u}; \dots; \mathbf{a}_N^{i,u}]^T, \quad A^{u,i} \in \mathbf{R}^{N \times K} \quad (9)$$

其中, $\mathbf{a}_n^{i,u}$ 和 $A^{u,i}$ 分别是第 n 个物品属性对应的注意力向量和用户-物品注意力交互矩阵. 注意力机制为用户个体分配了推荐候选物品各个属性的关注度.

3.3 卷积层

通过第 3.2 节式 (9) 得到的注意力交互矩阵记录了用户物品特征组间关系的高阶信息, 同时也是卷积层的输入. 考虑到式 (8) 是对用户特征向量分配注意力系数后融合的结果, 在一定程度上会丢失关于物品属性的原始信息, 因此最直接有效的做法就是将物品属性矩阵也作为卷积层输入的一部分. 这能让卷积提取高阶特征的同时考虑低阶信息.

相对高阶的特征矩阵 $A^{u,i}$ 和低阶特征 E_N^i 因维度相同, 可叠成卷积层的输入交互块, 而二维卷积层的输入和输出可抽象为:

$$I^{\text{CNN}} = \text{stack}(A^{u,i}, E_N^i), \quad I^{\text{CNN}} \in \mathbf{R}^{N \times K \times 2} \quad (10)$$

$$O^{\text{CNN}} = \text{ReLU}(2\text{D_CNN}(I^{\text{CNN}})), \quad I^{\text{CNN}} \in \mathbf{R}^{N \times K} \quad (11)$$

本文构造了类似于传统卷积层处理的图像特征图, 但图 5 中使用的是 $1 \times K \times 2$ 的行卷积核, 这不同于感知野中常见的矩形块状卷积核, 行卷积适理解类似文本形式的特征并提取精简且有效的信息, 文献 [9, 25] 也提及了相关性质的分析. 行卷积核将低阶和高阶信息联系起来并减少冗余信息, 卷积层的输出将进入下一个注意力模块与用户画像交互, 输出的副本直接级联到全连接层, 避免相对低阶的信息随着网络的加深而丢失.

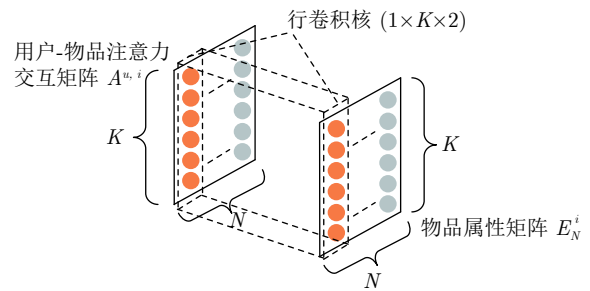


图 5 卷积层示意图

Fig.5 The demonstration of the convolutional layer

3.4 用户青睐评分概率的预测

本文模型最后一个模块是由多层全连接层组成

的 MLP, 它的输入层是一个由原始用户画像、物品属性和每个卷积层输出平铺级联而成的全连接层^[22-23,25,30-31]. 这借鉴了 NCF^[19] 中向量级联比内积效果更好, 将不同嵌入向量间的高阶关系交予多层隐藏神经元来处理. 而且将最低阶的特征嵌入向量直接跨接至最后网络部分, 有利于加快卷积层数增多时整个网络的训练收敛速度.

模型通过 Softmax 层输出用户对指定物品给出不同评分的概率, 相当于多分类问题, 但不同于直接预测评分的回归模型和推荐与否的二分类模型. 假设评分选项为 $r_{u,i} \in \{r_1, r_2, \dots, r_n\}$, 模型预测概率结果为 $\mathbf{y}_{u,i} = [y_1, y_2, \dots, y_n]$, 且满足 $\sum_{i=1}^n y_i = 1$. 值最大的 y_i 对应的 r_i 项就是用户最有可能给出的评分. 在最后的推荐表排名中, 可按如下计算用户 u 对于物品 i 的评分期望值:

$$\hat{r}_{u,i} = \mathbf{r}_{u,i} \otimes \mathbf{y}_{u,i} = \sum_{i=1}^n y_i r_i \quad (12)$$

4 实验

4.1 实验细节

按照以下规定评测本文模型的推荐准确率和质量: 1) 遍历每一位用户的所有历史行为按照 8:1:1 随机抽样添加到训练集、验证集和测试集中; 2) 在训练和测试模型时使用均值归一化后的评分作为 $u-i$ 行为值, 而在推荐阶段会涉及到用户从未接触过的物品故不存在对应评分, 使用 $h_{u,i}^* = \text{Norm}(\mathbf{p}_u \cdot \mathbf{c}_i)$ 作为 $u-i$ 行为值, 依旧满足 $h_{u,i}^* \in [-1, 1]$ 的要求; 3) 对于 multi-hot 字段特征按照 30% 比例随机掩膜去特征, 加入噪声来提高训练后推荐系统的鲁棒性; 4) 实验使用了两种数据集 MovieLens-1M 和 Niconico, 前者学习率设定为 0.01, MLP 各层隐藏神经元数为 [128, 64, 32, 5], 后者学习率设定为 0.1, MLP 各层隐藏神经元数为 [256, 128, 64, 10]; 5) 整个模型基于深度学习框架 Tensorflow 实现, 使用 Adagrad 优化器, MLP 使用 tanh 激活函数, 交叉熵式 (13) 作为损失函数.

$$J(\mathbf{r}_{u,i}, \mathbf{y}_{u,i} | \theta) = - \sum_{j=1}^n y_{u,i}^j \log(\hat{r}_{u,i}^j) \quad (13)$$

4.2 数据集

本文使用两个现实数据集来评估模型, 主要信息概述见表 1.

1) MovieLens-1M: 这是推荐系统领域非常经典的一个开源电影评分数据集, 现已由推荐系统概

表 1 数据集概述表

数据集	用户	物品	字段	样本	稀疏度
MovieLens-1M	6040	3883	5	1000 K	95.74%
Niconico	20566	13195	7	1045 K	99.62%

念的提出团队 GroupLens 汇总整理. 它提供了有关用户的人口统计数据 and 电影名称、风格等信息, 其中有 177 部电影无人评价.

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

2) Niconico: 这是在日本一个热门的日漫网络通过爬虫后整理得到的数据集. 它让推荐种类扩展到电影、电视剧、动漫连载等, 有着更为充足的用户样本和丰富的物品列表, 以及更高的稀疏度, 共 3784 部作品无人评价. 由于许多用户观看过后未留下评价信息, 本文过滤掉数据损失率达 97% 以上的用户, 剩余用户的缺失值用随机方法填补, 作为适量的数据噪声. 链接: <https://sp.nicovideo.jp/>

4.3 基准模型

基于矩阵分解的模型:

CFN^[17]: 为降噪自编码器网络的每一个隐藏层注入用户或物品的特征辅助信息.

基于分解机的模型:

FM^[3]: 最经典的分解机推荐模型, 分析特征组间的关系.

DeepFM^[22]: 提取多阶特征信息, 用更深的网络学习特征组对间的高阶关系.

混合模型:

NCF^[19]: 基于 DNNs 实现的协同过滤模型, 通过多组不同纬度的用户 - 物品隐因子向量丰富信息融合的方式.

xDeepFM^[23]: 独特的压缩交互网络结构来学习显隐式输入的高阶关系, 并能有效剔除冗余信息.

CFM^[30]: 通过 3D 卷积处理特征交互块, 以外积形式提取向量二阶信息.

4.4 评价标准

评分预测准确率可分析注意力-卷积层层数 L 、特征嵌入向量维度 K 和每层卷积核个数 T 三个超参对本文模型性能的影响. 实验着重于比较 Top-N 推荐列表的质量, 以 HR 和 NDCG^[42] 作为评价标准来比较 ACCDN 与上述模型的系统召回率和推荐列表正则化排名累积增益. 用每个用户推荐列表分析后的均值来分析模型, HR 和 NDCG 的值越高说明推荐质量越好, 此处 N 取 5, 10, 20, 30.

4.5 模型调参

本文的 ACCDN 推荐模型涉及三个关键的超参: 注意力-卷积层层数 L 、特征嵌入向量维度 K 和每层卷积核个数 T 。其中 K 是第 4.3 节中每个基准模型所需的超参, 而 T 也是 CFM 的参数。由于三个超参值的组合很多, 为了更好地体现单个超参对模型推荐的影响, 本文在固定两个超参为最佳值的情况下调节另一个参数, 以 Niconico 数据集为例见表 2。 L 的增加能够让模型提取用户和物品特征交互的更高阶信息; K 的增大能够将特征映射到更高维的稠密向量来丰富信息传输, 同时提高全连接层的宽度; 而 T 的增大可以让卷积层提炼更多模式的信息。权衡模型性能和训练代价后选择 $L = 3, K = 16, T = 32$ 作为 Niconico 推荐时的最佳参数组合。

4.6 实验结果与分析

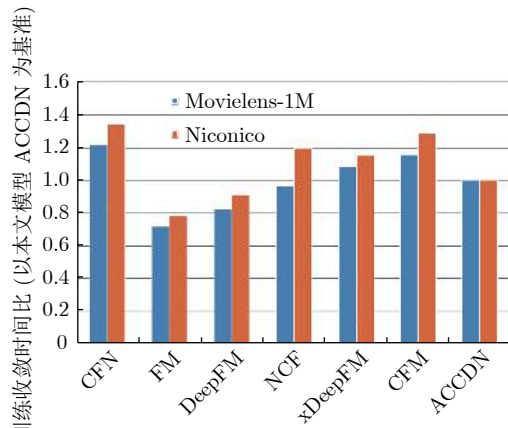
使用 GPU 训练所有模型, 考虑到模型在不同

表 2 超参对 ACCDN Top -30 推荐效果的影响 (Niconico 数据集)

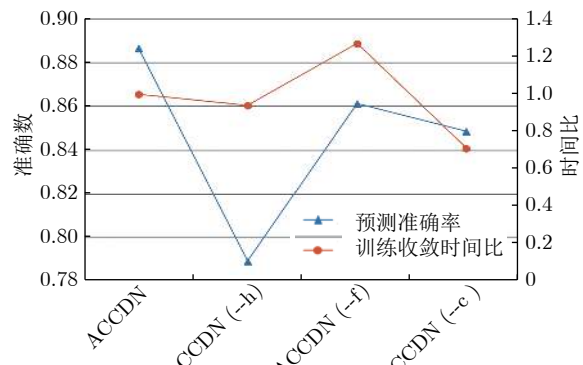
Table 2 ACCDN's hyper-parameters' influence on the Top-30 recommendation for Niconico

固定参数	调节参数	预测准确率 (%)	HR@30	NDCG@30
$K=16, T=32$	$L=1$	71.85	0.1659	0.0685
	$L=2$	82.54	0.2545	0.0872
	$L=3$	88.67	0.3001	0.0971
$L=3, T=32$	$K=8$	84.91	0.2683	0.0905
	$K=16$	88.67	0.3001	0.0971
	$K=24$	89.02	0.3089	0.0972
$L=3, K=16$	$T=8$	80.22	0.2244	0.0868
	$T=16$	84.58	0.2696	0.0919
	$T=32$	88.67	0.3001	0.0971

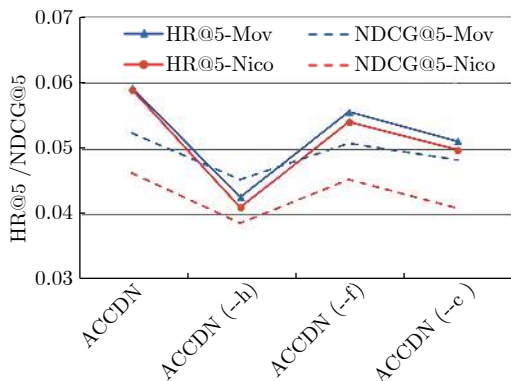
硬件条件下具体的收敛时间存在明显变化, 但是它们收敛的先后顺序是基本不变的, 故给出图 6 (a) 作为训练时间参考, 其中以 ACCDN 收敛时间作为



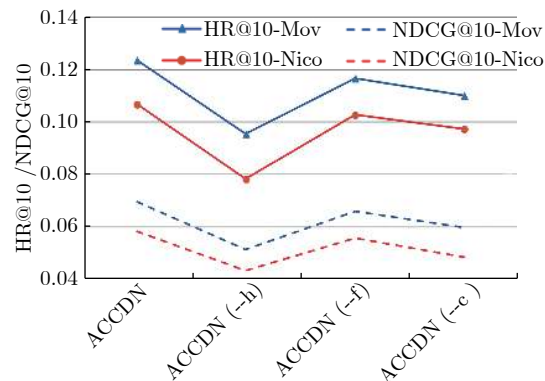
(a) 基准模型与 ACCDN 的训练收敛时间比
(a) The comparison of training convergence among models



(b) ACCDN 各模块对准确率和训练时间的影响
(b) The impacts of modules in ACCDN on accuracy and convergence



(c) ACCDN 各模块对 Top-5 推荐质量的影响
(c) The impacts of modules in ACCDN on Top-5 recommendation



(d) ACCDN 各模块对 Top-10 推荐质量的影响
(d) The impacts of modules in ACCDN on Top-10 recommendation

图 6 模型训练时间比较以及 ACCDN 各模块对推荐的影响

Fig.6 The comparison of the training costs among models and the impacts of the modules in ACCDN

基准值 1.0. CFN 因其隐藏层众多的神经单元明显收敛得最慢, FM 因只分析低阶特征收敛得最快, 而 DeepFM 则次之. ACCDN 的收敛速度与 NCF、xDeepFM 相当, 处于中等水平. 基于 3D 卷积网络的 CFM 运算过程最为复杂, 它在两个数据集上的训练代价是最高的, 所需时间比 ACCDN 多了约 19%.

关于图 6 (b)~(d) 中模型标记的说明: ACCDN (--h) 是在 ACCDN 的基础上不再输入用户历史行为值 $h_{u,i}^*$; ACCDN (--f) 是指不在全连接层额外输入原始的用户/物品特征向量信息; ACCDN (--c) 则是直接去除卷积层的模型. 通过分析图 5 (b)~(d) 可知, $h_{u,i}^*$ 记录了用户最主观的评分习惯, 屏蔽它对削弱推荐质量的影响最大, 准确率损失了约 11%, $h_{u,i}^*$ 有利于提高推荐的个性化和准确性. 而移除卷积层虽可以简化模型结构, 让模型训练收敛大幅加速, 但是卷积层是提炼压缩和提取高阶信息的关键模块, 缺少它会使推荐效果 (质量以及准确率等) 出现显著下降. 虽然不为全连接层输入原始特征信息不会明显影响模型的推荐性能, 但是显然没有原始数据的跨接不利于网络训练的收敛, 导致所需时间提高了 27%. 所以各个模块的协同工作才能综合优化模型的运作性能.

表 3 展示了 ACCDN 与基准模型在两个实验数据集上推荐质量的表现. 其中 CFN 在样本量极大且高度稀疏的 Niconico 数据集面前表现较差, 即

使它额外接受了特征辅助信息. FM、DeepFM 和 NCF 的推荐表现处于相近的中等水平, 而 xDeepFM、CFM 和本文模型 ACCDN 的推荐质量更好. 当 $N \in [5, 10]$ 较小时, xDeepFM 和 CFM 在 Top-N 推荐列表上的整体召回率 HR@N 比 ACCDN 略高, 而推荐排名质量 NDCG@N 旗鼓相当. 但随着推荐列表内容的丰富, 本文模型在指标上与其他基准模型的差距逐渐扩大, 意味着注意力机制与卷积层的组合能够高效处理特征关系, 即使在高稀疏度数据集下依旧能提供质量较好的推荐. 而就推荐排名质量而言, ACCDN 在 Top-N 推荐中是全体表现最佳的.

此外, ACCDN 由于基于时间用户兴趣标签的输入而善于跟踪用户兴趣的变化, 能在少量用户行为后快速响应出动态推荐, 以 3 次用户行为前后的 Top-5 推荐列表变化为例, 与同样性能优良的 xDeepFM 和 CFM 模型作比较的结果见表 4.

5 结语

本文提出了一种新式的推荐模型, 它通过注意力机制将用户行为值评分、用户画像和物品属性在特征向量空间中交互融合, 并借助 2D 卷积信息提炼和特征提取的能力逐层挖掘交互信息与特征向量的更高阶信息. 此外为了提高推荐动态性, 本文还提出了一种无参时间衰减用户兴趣标签赋予模型动

表 3 推荐列表评析结果

Table 3 The evaluation results of recommendation lists between baselines and ACCDN

HR	MovieLens-1 M				Niconico			
	HR@5	HR@10	HR@20	HR@30	HR@5	HR@10	HR@20	HR@30
CFN	0.0395	0.0788	0.1480	0.1923	0.0358	0.0676	0.1205	0.1747
FM	0.0498	0.0953	0.1899	0.2785	0.0471	0.0902	0.1659	0.2580
DeepFM	0.0577	0.1168	0.2101	0.3043	0.0543	0.1013	0.1928	0.2773
NCF	0.0543	0.1175	0.2081	0.2976	0.0485	0.0964	0.1799	0.2618
xDeepFM	0.0584	0.1250	0.2113	0.3109	0.0594	0.1032	0.2038	0.2843
CFM	0.0612	0.1233	0.2198	0.3177	0.0589	0.1054	0.2077	0.2917
ACCDN	0.0593	0.1237	0.2254	0.3253	0.0590	0.1069	0.2136	0.3001
NDCG	MovieLens-1 MM				Niconico			
	NG@5	NG@10	NG@20	NG@30	NG@5	NG@10	NG@20	NG@30
CFN	0.0326	0.0475	0.0610	0.0738	0.0308	0.0425	0.0535	0.0694
FM	0.0382	0.0504	0.0658	0.0790	0.0339	0.0492	0.0617	0.0760
DeepFM	0.0415	0.0549	0.0720	0.0853	0.0403	0.0544	0.0689	0.0832
NCF	0.0444	0.0612	0.0779	0.0901	0.0420	0.0535	0.0712	0.0859
xDeepFM	0.0493	0.0684	0.0852	0.0940	0.0448	0.0565	0.0723	0.0901
CFM	0.0470	0.0649	0.0815	0.0921	0.0487	0.0566	0.0728	0.0914
ACCDN	0.0524	0.0697	0.0862	0.1027	0.0463	0.0583	0.0750	0.0971

注: ACCDN(--h) 表示本文模型在注意力机制模块不加入用户行为值进行交互.

表 4 Top-5 动态推荐对比示例
Table 4 The examples of the dynamic Top-5 recommendation

用户1 喜欢类型 1) 动作 2) 冒险 3) 超凡 4) 运动			
原Top-5推荐			
xDeepFM	1. <i>Tengen Toppa Gurren Lagann</i> 动作/冒险/机甲 2. <i>Mononoke Hime</i> 动作/冒险/魔幻 3. <i>Fate/Zero2</i> 动作/超凡/魔幻 4. <i>Fullmetal Alchemis</i> 动作/冒险/魔幻 5. <i>Hunter x Hunter</i> 动作/冒险/超凡		
CFM	1. <i>One Piece</i> 动作/冒险/超凡/喜剧 2. <i>Fate/Zero2</i> 动作/超凡/魔幻 3. <i>JoJo no Kimyou na Bouken</i> 动作/超凡/冒险/青春 4. <i>Kizumonogatari II</i> 动作/悬疑/超凡 5. <i>Hellsing Ultimate</i> 动作/惊悚/超凡		
本文模型	1. <i>Hunter x Hunter</i> 动作/冒险/超凡 2. <i>JoJo no Kimyou na Bouken</i> 动作/超凡/冒险/青春 3. <i>Fate/Zero</i> 动作/超凡/魔幻 4. <i>Fate/Zero2</i> 动作/超凡/魔幻 5. <i>One Piece</i> 动作/冒险/超凡/喜剧		
增加三次用户行为: 1. <i>Tonari no Totoro</i> 动作/喜剧/校园/运动/青春 3. <i>Redline</i> 动作/赛车/科幻/运动			
新Top-5推荐			
xDeepFM	无变化	CFM	无变化
本文模型	1. <i>JoJo no Kimyou na Bouken</i> 动作/超凡/冒险/青春 2. <i>One Piece</i> 动作/冒险/超凡/喜剧 3. <i>Fate/Zero</i> 动作/超凡/魔幻 4. <i>Fairy Tail</i> 动作/冒险/喜剧/青春 5. <i>Hunter x Hunter</i> 动作/冒险/超凡		
用户2 喜欢类型 1) 惊悚 2) 超凡 3) 动作			
原Top-5推荐			
xDeepFM	1. <i>Hellsing Ultimate</i> 动作/惊悚/超凡 2. <i>Akira</i> 冒险/惊悚/超凡/科幻 3. <i>Paprika</i> 魔幻/惊悚/神秘 4. <i>Vampire Hunter D</i> 动作/魔幻/惊悚 5. <i>Another</i> 惊悚/恐怖/超凡		
CFM	1. <i>Hellsing Ultimate</i> 动作/惊悚/超凡 2. <i>Ajin Part 1: Shoudou</i> 动作/惊悚/神秘 3. <i>Ajin</i> 动作/惊悚/神秘 4. <i>Change!! Getter Robo</i> 动作/冒险/惊悚/科幻 5. <i>Memories</i> 惊悚/科幻		
本文模型	1. <i>Higurashi no Naku Koro ni</i> 惊悚/神秘/恐怖 2. <i>Tokyo Ghoul</i> 惊悚/超凡/动作/青春 3. <i>Change!! Getter Robo</i> 动作/冒险/惊悚/科幻 4. <i>Jigoku Shoujo</i> 惊悚/神秘/超凡 5. <i>Gakkou no Kaidan</i> 惊悚/超凡		
增加三次用户行为: 1. <i>Ano Natsu de Matteru</i> 喜剧/生活 2. <i>Ling Qi</i> 动作/喜剧/超凡 3. <i>One Piece</i> 动作/冒险/超凡/喜剧			
新Top-5推荐			
xDeepFM	无变化	CFM	无变化
本文模型	1. <i>Tokyo Ghoul</i> 惊悚/超凡/动作/青春 2. <i>Change!! Getter Robo</i> 动作/冒险/惊悚/科幻 3. <i>One Piece</i> 动作/冒险/超凡/喜剧 4. <i>Sankarea</i> 喜剧/超凡/生活 5. <i>Kemonozume</i> 动作/惊悚/超凡		

态跟踪用户兴趣变化的能力. ACCDN 还将初始特征、低阶到高阶交互信息级联在一起实现用户青睐评分的预测, 丰富模型输入信息的同时有效降低模型训练成本. 最后通过两个现实数据集将它与性能先进且具有代表性的模型做比较实验, 本文的 ACCDN 表现出更好的 Top-N 推荐质量. 之后我们还将继续研究学习先进的推荐算法和网络结构, 进一步提高本文推荐模型的性能和推荐的可解释性.

References

- Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 2015, 17(6): 734-749
- Zhang S, Yao L, Sun A, Tay Y. Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys*, 2017.
- Rendle S. Factorization machines. In: Proceedings of the IEEE International Conference on Data Mining. Sydney, Australia: IEEE, 2011. 995-1000
- Blondel M, Fujino A, Ueda N, Ishihata M. Higher-order factorization machines. In: Proceedings of Advances in Neural Information Processing Systems. New York, USA: Curran Associates Inc, 2016. 3359-3367
- Koren Y. Factorization meets the neighborhood: A multifaceted collaborative filtering model. In: Proceedings of 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Las Vegas Nevada, USA: SIGKDD, 2008. 426-434
- Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems. *Computer*, 2009, 42(8): 30-37
- Bharat K, Kamba T, Albers M. Personalized, interactive news on the Web. *Multimedia Systems*, 1998, 6(5): 349-358
- Herlocker J L, Konstan J A, Borchers A, Riedl J. An algorithmic framework for performing collaborative filtering. In: Proceedings of the 22nd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA: ACM, 1999. 230-237
- Salakhutdinov R, Mnih A. Probabilistic matrix factorization. In: Proceedings of Advances in Neural Information Processing Systems. New York, USA: NIPS, 2008. 1257-1264
- Lee J, Kim S, Lebanon G, Singer Y. Local low-rank matrix: approximation. In: Proceedings of 30th International Conference on Machine Learning. Atlanta, USA: ICML, 2013. 741-749
- Sedhain S, Menon A K, Samner S, Xie L. Autorec: autoencoders meet collaborative filtering. In: Proceedings of the 24th International Conference on World Wide Web. Florence, Italy: WWW,

2015. 111–112
- 12 Strub F, Mary J. Collaborative filtering with stacked denoising autoencoders and sparse inputs. *NIPS Workshop*, 2015.
- 13 Wu Y, DuBois C, Zheng A X, Ester M. Collaborative denoising auto-encoders for top-n recommender systems. In: Proceedings of the 9th ACM International Conference on Web Search and Data Mining. San Francisco, USA: WSDM, 2016. 153–162
- 14 Salakhutdinov R, Mnih A, Hinton G. Restricted Boltzmann machines for collaborative filtering. In: Proceedings of 24th International Conference on Machine Learning. Corvallis, USA: ICML, 2007. 791–798
- 15 Jia X, Li X, Li K, Gopalakrishnan V, et al. Collaborative restricted Boltzmann machine for social event recommendation. In: Proceedings of International Conference on Advances in Social Networks Analysis and Mining. San Francisco, USA: IEEE, 2016. 402–405
- 16 Li Jin-Zhong, Liu Guan-Jun, Yan Chun-Gang, Jiang Chang-Jun. Research advances and prospects of learning to rank. *Acta Automatica Sinica*, 2018, **44**(8): 1345–1369
(李金忠, 刘关俊, 闫春钢, 蒋昌俊. 排序学习研究进展与展望. 自动化学报, 2018, **44**(8): 1345–1369)
- 17 Strub F, Gaudel R, Mary J. Hybrid recommender system based on autoencoders. In: Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. Boston, USA: ICPS, 2016. 11–16
- 18 Dong X, Yu L, Wu Z, et al. A hybrid Collaborative filtering model with deep structure for recommender systems. In: Proceeding of 31st AAAI Conference on Artificial Intelligence. San Francisco, USA: AAAI, 2017. 1309–1315
- 19 Liang D, Krishnan R G, Hoffinan M D, Jebara T. Variational autoencoders for collaborative filtering. In: Proceedings of the 2018 Web Conference. Lyon, France: IEEE, 2018.
- 20 He X, Liao L, Zhang H, Nie L, Hu X, Chua T. Neural collaborative filtering. In: Proceedings of the 26th International Conference on World Wide Web. Perth, Australia: ACM, 2017. 173–182
- 21 Cheng H T, Koc L, Harmsen J, et al. Wide & deep learning for recommender systems. In: Proceedings of the 1st Workshop on the Deep Learning for Recommender Systems. Boston, USA: ICPS, 2016. 7–10
- 22 Guo H, Tang R, Ye Y, Li Z, He X. A factorization-machine based neural network for CTR prediction. arXiv preprint, arXiv: 1703.04247, 2017.
- 23 Lian J, Zhou X, Zhang F, Chen Z, Xie X, Sun G. xDeepFM: Combining explicit and implicit feature interactions for recommender systems. In: Proceedings of 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. London, UK: SIGKDD, 2018.
- 24 Zheng L, Noroozi V, Yu P S. Joint deep modeling of users and items using reviews for recommendation. In: Proceedings of the 10th ACM International Conference on Web Search and Data Mining. Cambridge, UK: ACM WSDM, 2017. 425–434
- 25 Kim D, Park C, Oh J, Lee S, Yu H. Convolutional matrix factorization for document context-aware recommendation. In: Proceedings of the 10th ACM Conference on Recommender Systems. Boston, USA: RecSys, 2016. 233–240
- 26 Covington P, Adams J, Sargin E. Deep neural networks for YouTube recommendations. In: Proceedings of the 10th ACM Conference on Recommender Systems. Boston, USA: RecSys, 2016. 191–198
- 27 Soh H, Sanner S, White M, Jamieson G. Deep sequential recommendation for personalized adaptive user interfaces. In: Proceedings of the 22nd International Conference on Intelligent User Interfaces. Limassol, Cyprus: ACM, 2017. 589–593
- 28 Suglia A, Greco C, Musto C, Gemmis M, Lops P, Semeraro G. A deep architecture for content-based recommendations exploiting recurrent neural networks. In: Proceedings of the 25th Conference on User Modeling Adaptation and Personalization. Bratislava, Slovakia: UMAP, 2017. 202–211
- 29 Li Z, Zhao H, Liu Q, Huang Z, Mei T, Chen E. Learning from history and present: Next-item recommendation via discriminatively exploiting user behaviors. arXiv, preprint, arXiv: 1808.01075, 2018.
- 30 Xin X, Chen B, He X, Wang D, Ding Y, Jose H M. CFM: Convolutional factorization machines for context-aware recommendation. In: Proceedings of the International Joint Conference on Artificial Intelligence. Macao, China: IJCAI, 2019. 3926–3932
- 31 Jiang J, Yang D, Xiao Y, Shen C. Convolutional Gaussian embeddings for personalized recommendation with uncertainty. In: Proceedings of the International Joint Conference on Artificial Intelligence. Macao, China: IJCAI, 2019. 2642–2648
- 32 Christakopoulou E, Karypis G. Local latent space models for top- n recommendation. In: Proceedings of 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. London, UK: SIGKDD, 2018. 1235–1243
- 33 Zhou X, Liu D, Lian J, Xie X. Collaborative metric learning with memory network for multi-relational recommender systems. In: Proceedings of the International Joint Conference on Artificial Intelligence. Macao, China: IJCAI, 2019.
- 34 Chae D, Kang J, Kim S, Lee J. CFGAN: A generic collaborative filtering framework based on generative adversarial networks. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management. Torino, Italy: CIKM, 2018. 137–146
- 35 Li Hui, Ma Xiao-Ping, Shi Jun, Li Cun-Hua, Zhong Zhao-Man, Cai Hong. A recommendation model by means of trust trans-

ition in complex network environment. *Acta Automatica Sinica*, 2018, **44**(2): 363–376

(李慧, 马小平, 施珺, 李存华, 仲兆满, 蔡虹. 复杂网络环境下基于信任传递的推荐模型研究. *自动化学报*, 2018, **44**(2): 363–376)

- 36 Chen J, Zhang H, He X, Nie L, Liu W, Chua T. Attentive collaborative filtering: Multimedia recommendation with item- and component-level attention. In: Proceedings of the 40th International ACM SIGIR Conference. Tokyo, Japan: SIGIR, 2017. 335–344
- 37 Feng Yong, Chen Yi-Gang, Qiang Bao-Hua. Social and comment text CNN model based automobile recommendation. *Acta Automatica Sinica*, 2019, **45**(3): 518–529
(冯永, 陈以刚, 强保华. 融合社交因素和评论文本卷积网络模型的汽车推荐研究. *自动化学报*, 2019, **45**(3): 518–529)
- 38 Xu Z, Chen C, Lukasiewicz O, Miao Y, Meng X. Tag-aware personalized recommendation using a deep-semantic similarity model with negative sampling. In: Proceedings of the 25th ACM International Conference on Information and Knowledge Management. Indianapolis, USA: CIKM, 2016. 1921–1924
- 39 Rawat Y S, Kankanhalli M S. ConTagNet: Exploiting user context for image tag recommendation. In: Proceedings of the 2016 ACM on Multimedia Conference. Amsterdam, Netherland: ACM, 2016. 1102–1106
- 40 Luong M, Pham H, Manning C D. Effective approaches to attention-based neural machine translation. arXiv preprint, arXiv: 1508.04025v5, 2015.
- 41 Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need. arXiv preprint, arXiv: 1706.03762v5, 2017.
- 42 Jarvelin K, Kekalainen J. Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems*, 2002, **20**(4): 422–446



汤文兵 东华大学信息科学与技术学院硕士研究生. 主要研究方向为深度学习, 数据挖掘与分析.

E-mail: wenbing_tang@hotmail.com
(**TANG Wen-Bing** Master student at the College of Information Science and Technology, Donghua University. His research interest covers deep learning, data mining and analysis.)



任正云 东华大学信息科学与技术学院教授. 主要研究方向为流程工业先进控制及应用, 系统建模与优化. 本文通信作者.

E-mail: renzhengyun@dhu.edu.cn
(**REN Zheng-Yun** Professor at the College of Information Science and Technology, Donghua University. His research interest covers advanced control and application of process, system modeling and optimization. Corresponding author of this paper.)



韩芳 东华大学信息科学与技术学院教授. 主要研究方向为神经动力学和智能系统.

E-mail: yadiahhan@dhu.edu.cn
(**HAN Fang** Professor at the College of Information Science and Technology, Donghua University. Her research interest covers neurodynamics and intelligent system.)