

基于用户兴趣模型的推荐算法^①

于波, 杨红立, 冷淼

(中国科学院大学, 北京 100049)

(中国科学院 沈阳计算技术研究所, 沈阳 110168)

通讯作者: 杨红立, E-mail: yang201107219@163.com

摘要: 传统的协同过滤算法虽然可以很容易地挖掘出用户的兴趣爱好, 但存在数据冷启动和稀疏性问题. 针对这些问题, 提出一种基于用户兴趣模型的推荐算法. 首先通过 LDA 主题模型训练数据集得到物品-主题概率分布矩阵, 利用物品-主题概率分布矩阵得到用户历史兴趣模型, 然后结合用户历史行为信息和物品内容信息得到用户兴趣模型, 最后计算用户与候选集之间的相似度, 进行 TOP-N 推荐. 在豆瓣电影数据集上的实验结果表明, 改进后的推荐算法能够更好地处理稀疏数据和冷启动问题, 并且明显提高了推荐质量.

关键词: 协同过滤; 用户兴趣模型; LDA 主题模型; 推荐算法

引用格式: 于波, 杨红立, 冷淼. 基于用户兴趣模型的推荐算法. 计算机系统应用, 2018, 27(9): 182-187. <http://www.c-s-a.org.cn/1003-3254/6496.html>

Recommendation Algorithm Based on User Interest Model

YU Bo, YANG Hong-Li, LENG Miao

(University of Chinese Academy of Sciences, Beijing 100049, China)

(Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China)

Abstract: Although traditional collaborative filtering recommendation algorithm can easily find potential users' interests, it remains cold-start problem and sparsity problem. In order to solve these problems, a new hybrid recommendation algorithm is proposed. Firstly, this study builds topic distribution matrix through the LDA topic model, and user interest matrix is created using topic distribution matrix. Secondly, the user interest model is obtained by combining user's historical behavior information and user's content information. Finally, the TOP-N recommendation list is output after calculating the similarity of user and candidate movies. Experiments on the Douban Movies dataset reveals that the results obtained from improved recommendation algorithm are obviously better than that from traditional recommendation algorithm, and it can better deal with sparse data and cold-start problems.

Key words: collaborative filtering; user interest model; LDA topic model; recommendation algorithm

随着互联网特别是移动互联网的飞速发展, 网络上的资源呈指数级增长, 网上资源严重过载^[1,2]. 传统的搜索引擎技术已经不能满足人们的实际需要, 而推荐系统可以利用用户行为信息和物品内容信息挖掘人们可能感兴趣的物品, 从而进行个性化的信息推荐^[1,3]. 因此推荐系统利用本身独特的优势开始逐步融入到日常

生活中, 减短人们获取信息的时间, 它被大量地应用在网络商城、信息检索和互联网广告等领域^[1,2]. 好的推荐系统不仅仅可以给用户推荐个性化信息, 而且可以增加用户对网站的依赖性. 推荐系统同时也面临着众多的问题: 数据稀疏性问题^[4]、冷启动问题^[5]、计算量大^[6]、扩展性问题^[7]等. 目前国内外的众多学者已经针

① 收稿时间: 2018-01-04; 修改时间: 2018-01-23; 采用时间: 2018-02-01; csa 在线出版时间: 2018-08-16

对这些问题进行了一系列的研究. 文献[8]首先介绍了协同过滤推荐技术所面临的问题, 接着介绍了解决这些问题的方案, 最后指明了协同过滤推荐技术未来的研究重点. 文献[9]提出了一种基于共同评分和相似性权重的协同过滤推荐算法. 文献[10]提出将时间因素融入用户项目评分矩阵中, 以解决兴趣衰减的问题. 文献[11]提出利用基于用户协同过滤算法计算出邻居集, 再结合加权 Slope One 算法, 预测项目评分, 实现对用户个性化信息推荐. 文献[12]提出了为了降低用户推荐的计算时间, 提出了一种改进的层次聚类协同过滤用户推荐算法. 对用户聚类, 按照分类计算出每个用户的推荐列表. 上述这些研究着重于利用某种机制改进用户相似度的计算或者加速用户相似度的计算, 忽略了物品内容信息对用户偏好的影响. 在传统协同过滤机制中直接利用评分矩阵来挖掘出目标用户的偏好, 但并不能反映出目标用户真实的兴趣爱好. 例如电影《泰坦尼克号》的属性为爱情、灾难、莱昂纳多、卡梅隆等, 用户选择该电影是因为他是主角莱昂纳多的影迷, 但通过评分矩阵无法体现该用户的兴趣所在^[13]. 因此本文提出通过结合物品内容信息和用户行为信息来构建用户兴趣模型, 利用用户兴趣模型进行推荐, 这充分发挥了物品内容对用户兴趣爱好的影响. 推荐结果在准确率和召回率等方面相比传统的基于用户的协同过滤有着明显的提高.

1 相关技术

1.1 LDA 主题模型

主题模型^[14-17]是对文字隐含主题的一种建模方式, 它可以挖掘文本中潜在的主题. LDA^[18]是主题模型中最经典的一种算法. 它也是一种生成模型^[14], 认为一篇文档中每个单词都是经过“以一定概率选择某个主题, 并在这个主题中以一定概率选择某个单词”这一过程得到的. 根据 LDA 主题模型的生成过程^[16,17]的定义, 那么每篇文档中词语出现的概率如式(1)所示.

$$p(\text{word}|\text{document}) = \sum_{\text{topic}} p(\text{word}|\text{topic}) * p(\text{topic}|\text{document}) \quad (1)$$

LDA 的概率图模型如图 1 所示. 其中, M 是文档数量; K 是主题个数; V 是词袋长度; N_m 是第 m 篇文档中单词的总数; α 和 β 是先验参数; θ 是一个 $M \times K$ 的矩阵, θ_m 表示第 m 篇文档的主题分布. α 到 θ 到 Z 的过程

表示在生成第 m 篇文档时, 先确定第 m 篇文档的主题分布, 然后确定第 m 篇文档中的第 n 个词语的主题. φ 是一个 $K \times V$ 的矩阵, φ_k 表示第 k 个主题的词分布; β 到 φ 到 W 的过程表示在 K 个主题中, 选出编号为 $Z_{m,n}$ 的主题, 然后生成第 m 篇文档中的第 n 个词语 $W_{m,n}$. LDA 算法的输入是大规模的文档集合、两个超参数和主题个数, 经过 LDA 主题模型的训练之后得到两个分布, 分别为文档-主题概率分布和主题-单词概率分布.

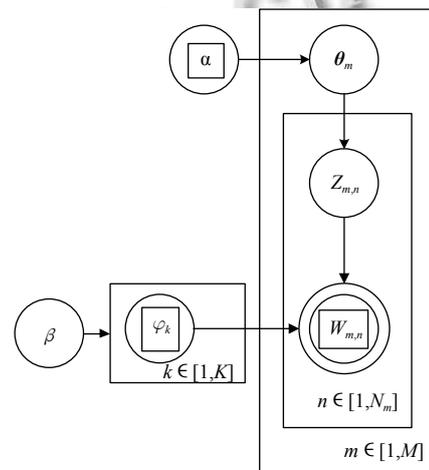


图 1 LDA 的概率图模型

1.2 基于用户的协同过滤

基于用户的协同过滤算法^[3,8]认为, 一个用户会喜欢和他有相似兴趣爱好的最近邻所喜欢的东西^[14], 其主要利用行为的相似度计算兴趣的相似度.

用户相似度的计算是基于共同评分的项目集进行的, 通常利用余弦相似度^[19,20]进行计算, 如式(2)所示:

$$\text{sim}(u, v) = \frac{\sum_{i \in N(u) \cap N(v)} \frac{1}{\ln(1+D(i))}}{\sqrt{|N(u)|} \sqrt{|N(v)|}} \quad (2)$$

其中, $D(i)$ 表示对物品 i 有过行为的用户集, $N(u)$ 表示用户 u 有过行为的物品集合.

基于用户的协同过滤算法如算法 1.

算法 1. 基于用户的协同过滤算法

输入: 评分矩阵 R , 物品集 D , 用户集 N , 目标用户 u

输出: 目标用户 u 的推荐列表

Begin:

$U_{\text{sim}} = \{\}, \text{sim} = \{\}, \text{rank} = \{\}$

for v in N :

for d in $N(u)$ and d in $N(v)$:

$\text{sim}(u, v) += 1 / (\ln(1 + \text{len}(D(d))))$;

```

end for
sim(u,v) /= (len(N(u)) * len(N(v)));
end for
Usim = sorted(sim(u,v))[0:N];
for v in Usim:
    for i in N(v) and i not in N(u):
        rank(i) += sim(u,v)
    end for
end for
return sorted(rank)[0:N]
    
```

2 用户兴趣模型的建立

虽然基于用户的协同过滤对物品冷启动问题不敏感,但需要解决第一推动力问题,即第一个用户如何发现新物品.如果将物品随机展示给用户,显然不太个性化,因此可以考虑利用物品的内容信息,将新物品投放给曾经喜欢过和它内容相似的物品的用户^[20].

通过分析用户的历史评分记录来创建用户历史兴趣模型,进而为用户推荐一组物品.用户的历史记录是有限的,因此存在数据稀疏性问题.针对此问题,提出了以用户行为和物品内容为基础来构建用户兴趣模型,从而为用户进行推荐.

2.1 LDA 主题模型建模

首先通过标题、导演、编剧、主演、类型和简介等等来给电影划分属性,生成电影-属性分布文件.然后在电影-属性分布上利用 LDA 主题模型进行建模,得到电影-主题概率分布,利用这个分布计算相似性.

给定电影集 $M = \{m_1, m_2, \dots, m_n\}$, 将每个电影看成一个单独的文档,对于文档中内容信息诸如像导演、主演等实体的话可以直接将这些实体作为电影属性,但诸如像简介等需要对文本内容进行分词,从字流变为词流,从词流中提取出命名实体,将这些命名实体也作为电影属性,形成电影-属性分布.在电影-属性分布上利用 LDA 算法进行建模,设置主题个数为 K ,从而得到主题特征序列^[21] $F = (f_1, f_2, \dots, f_k)$ 和电影-主题概率分布矩阵 Θ , 如式 (3) 所示:

$$\Theta = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1k} \\ w_{21} & w_{22} & \dots & w_{2k} \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nk} \end{bmatrix} \quad (3)$$

2.2 构建用户历史兴趣模型 UHIM

对于任意用户,利用评论过的电影和电影-主题概

率分布矩阵 Θ 进行数学运算,得到与 F 对应的权值向量称为用户历史兴趣模型 (User History Interest Model, UHIM),其数学式为 $UHIM = (w_{11}, w_{12}, \dots, w_{1b}, \dots, w_{1k})$, 其中, w_{1i} 表示主题词 f_i 在 UHIM 中的权值.用户 u 的 UHIM 中主题词 f_i 的权值计算如式 (4) 所示,该值代表了用户现有情况下的兴趣分布情况,较好地反映了用户历史的兴趣爱好.

$$w_{1ui} = \frac{\sum_{m \in M_u} w_{mi}}{|M_u|} \quad (4)$$

其中, M_u 是用户 u 评论的电影集合.

2.3 构建用户行为兴趣模型 UAIM

对于任意用户,利用评论过的电影计算用户行为之间的相似度,通过协同过滤将相似用户群的历史模型推荐给该用户,得到与 F 对应的权值向量称为用户行为兴趣模型 (User Actor Interest Model, UAIM),其数学式为 $UAIM = (w_{21}, w_{22}, \dots, w_{2b}, \dots, w_{2k})$, 其中, w_{2i} 表示主题词 f_i 在 UAIM 中的权值.在选择相似用户群时选择相似度最大的前 h 个用户.

用户 u 和用户 v 的行为相似度计算如式 (2) 所示.

用户 u 的 UAIM 中主题词 f_i 权值计算如式 (5) 所示:

$$w_{2ui} = \sum_{var \in U_{act}} \frac{sim_{act}(u, var)}{\sum_{vas \in U_{act}} sim_{act}(u, vas)} w_{1vasi} \quad (5)$$

其中, U_{act} 为用户 u 的行为相似用户群.

2.4 构建用户内容兴趣模型 UCIM

对于任意用户,结合电影的内容信息计算用户内容之间的相似度,通过协同过滤将相似用户群的历史兴趣模型推荐给该用户,得到与 F 对应的权值向量称为用户内容兴趣模型 (User Content Interest Model, UCIM),其数学式为 $UCIM = (w_{31}, w_{32}, \dots, w_{3b}, \dots, w_{3k})$, 其中, w_{3i} 表示主题词 f_i 在 UCIM 中的权值.在选择相似用户群时选择相似度最大的前 h 个用户.

设用户 u 评论的电影集 $M_u = \{m_{u1}, m_{u2}, \dots, m_{us}\}$, 历史模型 $UHIM_u = (w_{1u1}, w_{1u2}, \dots, w_{1uk})$, 用户 v 评论的电影集 $M_v = \{m_{v1}, m_{v2}, \dots, m_{vt}\}$, $UHIM_v = (w_{1v1}, w_{1v2}, \dots, w_{1vk})$.

用户 u 和用户 v 的内容相似度计算如式 (6) 所示.

$$sim_{con}(u, v) = \frac{UHIM_u * UHIM_v}{|UHIM_u||UHIM_v|} \quad (6)$$

用户 u 的 UCIM 中主题词 f_i 权值计算如式 (7)

所示:

$$w3_{ci} = \sum_{v_{cr} \in U_{con}} \frac{sim_{con}(u, v_{cr})}{\sum_{v_{cs} \in U_{con}} sim_{con}(u, v_{cs})} w1_{v_{cr}i} \quad (7)$$

其中, U_{con} 为用户 u 的内容相似用户群。

2.5 构建用户兴趣模型 UIM

得到目标用户的 UHIM、UAIM 和 UCIM 后, 将三个兴趣模型的主题词权值进行合并, 可得到用户兴趣模型 (User Interest Model, UIM), 其数学表达形式为 $UIM=(w4_1, w4_2, \dots, w4_i, \dots, w4_k)$, 其中, $w4_i$ 表示主题词 f_i 在 UIM 中的权值。

设目标用户 u 的 $UHIM_u=(w1_{u1}, w1_{u2}, \dots, w1_{uk})$ 、 $UAIM_u=(w2_{u1}, w2_{u2}, \dots, w2_{uk})$ 、 $UCIM_u=(w3_{u1}, w3_{u2}, \dots, w3_{uk})$, $UIM_u=(w_{u1}, w_{u2}, \dots, w_{uk})$, 则用户 u 的 UIM 中主题词 f_i 的权值计算如式 (8) 所示:

$$w4_{ui} = (1 - \alpha - \beta)w1_{ui} + \alpha w2_{ui} + \beta w3_{ui} \quad (8)$$

2.6 生成推荐列表

将候选电影集同目标用户 UIM 进行相似度计算, 进行 TOP-N 推荐。

设候选电影 $M_{cs}=\{m_1, m_2, \dots, m_h\}$. 目标用户 u 和候选电影之间的相似度计算如式 (9) 所示:

$$sim(UIM_u, m_i) = \frac{UIM_u * m_i}{|UIM_u| |m_i|} \quad (9)$$

构建用户兴趣模型算法的描述如算法 2。

首先将电影内容向量化, 利用 LDA 模型建模, 生成电影-主题概率分布矩阵 Θ ; 其次分析用户记录, 构建用户历史兴趣模型 UHIM; 再次利用用户记录计算用户行为相似度, 从而构建用户行为兴趣模型 UAIM; 最后结合物品内容信息计算用户内容相似度, 进而构建用户内容兴趣模型 UCIM, 融合 UHIM、UAIM 和 UCIM 生成用户兴趣模型 UIM。

算法 2. 构建用户兴趣模型的算法

输入: 电影-主题概率分布矩阵 Θ , 用户集 N , 目标用户 u

输出: 目标用户 u 的兴趣模型

Begin:

$UHIM_u=(w1_{u1}, \dots, w1_{u1}, \dots, w1_{uk}), UAIM_u=(w2_{u1}, \dots, w2_{u1}, \dots, w2_{uk}),$

$UCIM_u=(w3_{u1}, \dots, w3_{u1}, \dots, w3_{uk}), UIM_u=(w4_{u1}, \dots, w4_{u1}, \dots, w4_{uk}),$

$U_{con}=\{\}, U_{act}=\{\}, sim_{con}=\{\}, sim_{act}=\{\}$

// calculate $UHIM_u$

for d in $N(u)$:

$w1_{ui} += w_{di};$

end for

$w1_{ui} /= \text{len}(N(u));$

// calculate sim_{act}

for v in N :

for d in $N(u)$ and d in $N(v)$:

$sim_{act}(u, v) += 1 / (\ln(1 + \text{len}(D(d))));$

end for

$sim_{act}(u, v) /= (\text{len}(N(u)) * \text{len}(N(v)));$

$U_{act} = \text{sorted}(sim_{act})[0:h];$

end for

// calculate $UAIM_u$

for v in U_{act} : $sum_{act} += sim_{act}(u, v)$

end for

for v in U_{act} :

$w2_{ui} += sim_{act}(u, v) * w1_{vi} / sum_{act};$

end for

// calculate sim_{con}

$i = 0;$

while $i <= k$:

$sum1 += w1_{ui} * w1_{vi};$

$sumu += w1_{ui} * w1_{ui};$

$sumv += w1_{vi} * w1_{vi};$

end while

$sim_{con}(u, v) = sum1 / (\sqrt{sumu} * \sqrt{sumv});$

$U_{con} = \text{sorted}(sim_{con})[0:h];$

// calculate $UCIM_u$

for v in U_{con} :

$sum_{con} += sim_{con}(u, v);$

end for

for v in U_{con} :

$w3_{ui} += sim_{con}(u, v) * w1_{vi} / sum_{con};$

end for

// calculate UIM_u

$w4_{ui} = (1 - \alpha - \beta)w1_{ui} + \alpha w2_{ui} + \beta w3_{ui};$

return $UIM_u;$

3 实验结果及其分析

3.1 实验数据与实验评价指标

实验数据采用豆瓣电影网的 1337 部电影, 1535 个用户, 109 398 条评分记录作为实验数据集。

本文使用离线实验方法进行评测^[20]. 评价指标选用准确率/召回率和 F 值来评测推荐算法的精度. 召回率^[17,20]描述有多少真正产生过行为的物品包含在最终的推荐列表中; 准确率^[3,20]描述最终的推荐列表中包含多少真正产生过行为的物品; F 值^[19,21]是准确率和召回率之间的调和平均值. 对用户 u 推荐 N 个物品, 记为 $R(u)$; 用户 u 在测试集合上产生过行为的物品记为

$T(u)$. 召回率计算如式 (10) 所示.

$$Recall_u = \frac{\sum_u |R(u) \cap T(u)|}{\sum_u |T(u)|} \quad (10)$$

准确率计算如式 (11) 所示.

$$Precision_u = \frac{\sum_u |R(u) \cap T(u)|}{\sum_u |R(u)|} \quad (11)$$

F 值计算如式 (12) 所示.

$$F = \frac{2 * P * R}{P + R} \quad (12)$$

3.2 实验结果

(1) 主题个数 K 的确定

利用 LDA 主题模型进行建模时, 需要设定主题个数 K 的大小, 表 1 显示了当给每个用户推荐电影数目为 20 时, 不同的 K 值对召回率、准确率和 F 值的影响, 可以看出 K 为 20 的时候是最佳值.

表 1 不同的主题个数对应的召回率、准确率和 F 值

主题个数	10	15	20	25	30
召回率	0.165	0.187	0.206	0.185	0.176
准确率	0.289	0.305	0.327	0.3	0.291
F 值	0.2100	0.2318	0.2527	0.2288	0.2193

(2) 最近邻 h 的确定

构建 UAIM 和 UCIM 时需要设置最近邻的个数, 表 2 显示了当主题个数为 20 并且给每个用户推荐电影数为 20 的时候, 不同 h 值对召回率、准确率和 F 值的影响, 可以看出相似用户数为 30 的时候是最佳值.

表 2 不同的 h 值对应的召回率、准确率和 F 值

相似用户数	10	20	30	40	50
召回率	0.193	0.202	0.214	0.209	0.198
准确率	0.309	0.323	0.336	0.33	0.324
F 值	0.237	0.248	0.2614	0.2559	0.2457

(3) 不同推荐方法下召回率、准确率和 F 值.

图 2 显示了主题个数为 20, 最近邻为 30 时, 不同的推荐电影数目对三种推荐算法在召回率方面的影响. 随着推荐电影数量的增加, 召回率均随之增加.

图 3 显示了主题个数为 20, 最近邻为 30 时, 不同的推荐电影数目对三种推荐算法在 F 值方面的影响. 随着推荐电影数量的增加, 准确率随之降低.

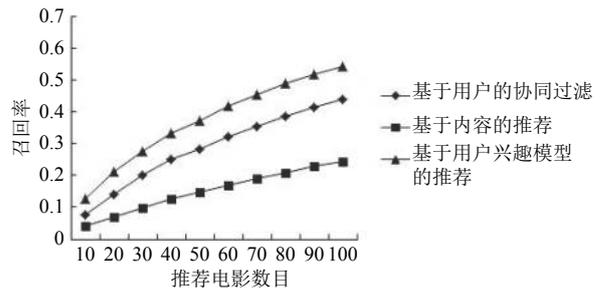


图 2 召回率对比

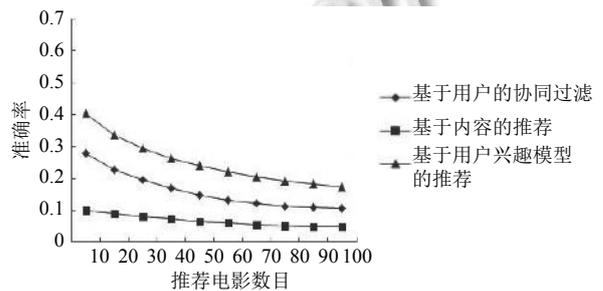


图 3 准确率对比

图 4 显示了主题个数为 20, 最近邻为 30 时, 不同的推荐电影数目对三种推荐算法在召回率方面的影响. 推荐电影数量为 40, F 值最佳.

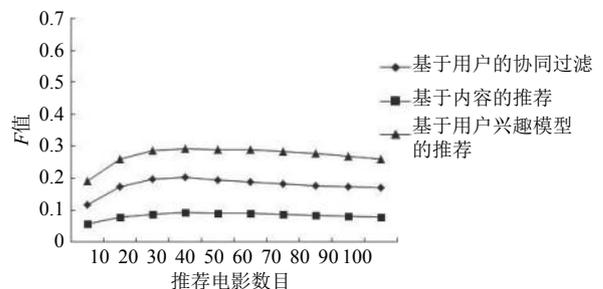


图 4 F 值对比

通过上述的对比可以看出, 本文采用的方法在推荐质量上要优于传统的推荐算法, 从而证明了该模型的可行性.

4 结语

本文针对传统协同过滤算法中面临的数据冷启动和稀疏性的问题, 提出了一种基于用户兴趣模型的电影推荐算法, 该算法首先利用用户记录和物品信息构建用户历史兴趣模型, 然后使用协同过滤算法挖掘出用户行为兴趣模型和用户内容兴趣模型, 最后将三种模型进行融合, 再与候选电影集进行相似度计算. 最后

通过实验验证,表明本文算法在一定程度上改善了数据稀疏性问题,提高了算法的推荐质量.面对日益增多的用户,数据量的急剧增加,算法的扩展性问题成为制约推荐系统实施的重要因素,尤其当用户量达到一定水平时,用户相似度的计算量是巨大的,传统的推荐算法会遇到严重的瓶颈问题,该问题解决不好,将会直接影响推荐系统的质量,那么下一步的重点是解决算法的扩展性问题.

参考文献

- 1 胡勋,孟祥武,张玉洁,等.一种融合项目特征和移动用户信任关系的推荐算法.软件学报,2014,25(8):1817-1830. [doi: 10.13328/j.cnki.jos.004491]
- 2 曹一鸣.协同过滤推荐瓶颈问题综述.软件,2012,33(12):315-321. [doi: 10.3969/j.issn.1000-386x.2012.12.088]
- 3 陶俊,张宁.基于用户兴趣分类的协同过滤推荐算法.计算机系统应用,2011,20(5):55-59. [doi: 10.3969/j.issn.1003-3254.2011.05.013]
- 4 Kim HN, Ji AT, Ha I. Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. Electronic Commerce Reserch and Applications, 2010, 9(1): 73-83. [doi: 10.1016/j.elerap.2009.08.004]
- 5 Schein A, Popescul A, Ungar LH, *et al.* Methods and metrics for cold-start recommendation. Reserch and Development in Information Retrieval, 2012, 22(1): 253-260.
- 6 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 & Data Engineering, 2015, 17(6): 734-749. [doi: 10.1109/TKDE.2005.99]
- 7 Koren Y, Bell R. Advances in collaborative filtering. US: Recommender Systems Handbook, 2011.
- 8 冷亚军,陆青,梁昌勇.协同过滤推荐技术综述.模式识别与人工智能,2014,27(8):720-734. [doi: 10.3969/j.issn.1003-6059.2014.08.007]
- 9 汪静,印鉴,郑利荣,等.基于共同评分和相似性权重的协同过滤推荐算法.计算机科学,2010,37(2):99-104. [doi: 10.3969/j.issn.1002-137X.2010.02.022]
- 10 董立岩,王越群,贺嘉楠,等.基于时间衰减的协同过滤推荐算法.吉林大学学报(工学版),2017,47(4):1268-1272.
- 11 刘毓,何锐.一种改进 SlopeOne 的学习推荐算法.西安邮电大学学报,2017,22(4):105-108.
- 12 张峻玮,杨洲.一种基于改进的层次聚类的协同过滤用户推荐算法研究.计算机科学,2014,41(12):176-178. [doi: 10.11896/j.issn.1002-137X.2014.12.038]
- 13 姚平平,邹东升,牛宝君.基于用户偏好和项目属性的协同过滤推荐算法.计算机系统应用,2015,24(7):15-21. [doi: 10.3969/j.issn.1003-3254.2015.07.003]
- 14 高娜,杨明.嵌入 LDA 主题模型的协同过滤推荐算法.计算机科学,2016,43(3):57-61.
- 15 黄璐,林川杰,何军,等.融合主题模型和协同过滤的多样化移动应用推荐.软件学报,2017,28(3):708-720. [doi: 10.13328/j.cnki.jos.005163]
- 16 唐晓波,谢力.基于主题的用户兴趣模型的构建及动态更新.情报理论与实践,2016,39(2):116-123.
- 17 卢露,朱福喜,高榕,等.基于用户-内容主题模型的兴趣点联合推荐算法.计算机工程与应用,2018,(4):154-159.
- 18 廉涛,马军,王帅强,等.LDA-CF:一种混合协同过滤方法.中文信息学报,2014,28(2):129-135. [doi: 10.3969/j.issn.1003-0077.2014.02.019]
- 19 吴毅涛,张兴明,王兴茂,等.基于用户模糊相似度的协同过滤算法.通信学报,2016,37(1):198-206.
- 20 项亮.推荐系统实践.北京:人民邮电出版社,2012.
- 21 杨武,唐瑞,卢玲.基于内容的推荐与协同过滤融合的新闻推荐方法.计算机应用,2016,36(2):414-418.