

融合内容与协同过滤的混合推荐算法应用研究

李雪婷¹, 杨 抒², 赛亚热·迪力夏提¹, 赵昀杰¹

(1. 新疆农业大学 计算机与信息工程学院, 新疆 乌鲁木齐 830000;

2. 成都大学 计算机学院, 四川 成都 610000)

摘 要:随着大数据时代的到来,海量数据出现在人们眼前。从冗杂的信息中快速获取满足人们个性化需求的数据成为了一个棘手的问题,推荐算法是解决此类问题的有力工具。针对移动电商平台中信息量较大,人们难以快速获得所需信息的问题,提出了一种融合内容与协同过滤的混合推荐算法。该算法先利用基于 LFM 的协同过滤算法产生推荐结果;当面临向新用户或新物品进行推荐时,再利用基于内容的推荐算法产生推荐结果。采用级联的方式将两种推荐算法进行混合,在一定程度上缓解了用户冷启动问题、物品冷启动问题及数据稀疏问题,克服了单一算法的局限性。实验证明,相较于传统的基于用户和基于物品的推荐算法,该算法能够有效提高推荐结果的准确率、召回率及覆盖率,从而提升推荐质量。

关键词:混合推荐;协同过滤;基于内容;冷启动问题;推荐质量

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2021)10-0024-06

doi: 10.3969/j.issn.1673-629X.2021.10.005

Research on Application of Hybrid Recommendation Algorithm of Content Fusion and Collaborative Filtering

LI Xue-ting¹, YANG Shu², Saiyare DILIXIATI¹, ZHAO Yun-jie¹

(1. School of Computer and Information Engineering, Xinjiang Agricultural University, Urumqi 830000, China;

2. School of Computer, Chengdu University, Chengdu 610000, China)

Abstract: With the advent of the era of big data, massive data appears before people's eyes. It has become a thorny problem to quickly obtain data from miscellaneous information to meet people's personalized needs. Recommendation algorithm is a powerful tool to solve such problems. Aiming at the problem that it is difficult for people to get the required information quickly due to the large amount of information in mobile e-commerce platforms, a hybrid recommendation algorithm combining content fusion and collaborative filtering is proposed. Firstly, the collaborative filtering algorithm based on LFM is used to generate the recommendation results. When it comes to making recommendations to new users or new items, content-based recommendation algorithms are used to generate recommendation results. The two recommendation algorithms are mixed in a cascading way, which alleviate the cold start of users, cold start of items and data sparsity to some extent, and overcome the limitation of single algorithm. Experiment shows that compared with traditional user-based and item-based recommendation algorithms, the proposed algorithm can effectively improve the accuracy rate, recall rate and coverage rate of recommendation results, thus improving the quality of recommendation.

Key words: hybrid recommendation; collaborative filtering; content-based; cold start; recommendation quality

0 引 言

互联网数据量的激增导致信息过载成了大数据时代的显著问题。数据的增长率已远大于人们的处理能力,现有的技术已不能很好地处理日益加速增长的数据^[1]。从冗杂的数据中,人们难以快速准确地获取有价值的信息。根据热门信息的展示很难满足用户的个性化需求,为此推荐系统应运而生^[2]。近年来,推荐系

统在电商平台、短视频 App、社交网站等各个领域广泛应用,使其快速发展并具有一定商用价值。比如在电商平台中用于挖掘用户行为数据里的隐藏商业价值。

据相关部门统计,推荐系统对网上商品销售的贡献率为 20%~30%,已经成为电商平台中至关重要的工具之一^[3]。

收稿日期: 2020-11-27

修回日期: 2021-03-29

基金项目: 自治区“十三五”重点研发项目(2017B01002-4)

作者简介: 李雪婷(1995-),女,硕士研究生,研究方向为农业信息化;通讯作者: 杨 抒,副教授,研究方向为软件工程。

1 相关研究

推荐系统根据实现方式的不同可将其分为基于内容、基于协同过滤、基于深度学习和混合的推荐系统^[1]。其中基于内容的推荐系统能够较好地为用户提供个性化推荐服务,适应冷启动问题,但准确率不高。基于协同过滤的推荐系统应用最广泛^[4],能够较为准确地向用户进行推荐。该算法并不依赖于物品的相关内容信息,但需要有用户行为数据做支撑。当数据库中数据稀疏或者没有数据时,没有足够的信息使该算法进行用户行为分析,导致推荐质量不高,推荐结果的有效性较差^[5]。混合推荐算法是解决上述问题较为成熟的方法之一,该算法将多种推荐算法进行混合,在实际应用中能够使各个推荐算法发挥其各自的优势,并且在一定程度上避免了其各自的劣势,进行高质量推荐^[6-7]。目前,有众多学者将其在各个领域中应用。例如,何锴琦等人^[8]将两种传统基于内存的协同过滤方法进行混合,应用在 GitHub 项目个性化推荐中。冉从敬等人^[9]将基于内容和基于协同过滤的推荐算法进行混合,应用在高校专利个性化推荐中。王末等人^[10]将动态加权的混合过滤推荐方法应用在地理数据共享领域中。叶小榕等人^[11]将基于物品的协同过滤推荐算法和基于热点的推荐算法进行混合,应用在图书选取中,都取得了较好的效果。文中采用级联的方式将基于内容与协同过滤的推荐算法进行混合,将其应用在移动电商应用领域中,解决用户难以从大量信息中获取满足个性化需求数据的问题。

2 理论基础

2.1 基于内容的推荐算法

基于内容的推荐(content-based recommendation)算法最早用于信息检索^[12]。该算法将用户和物品的特征都用显式标签进行表示,利用得到的用户标签与物品标签构建用户画像与物品画像,并进行相似度度量,得出相似度较高的 TOP-N 项完成推荐^[13],其结构如图 1 所示。

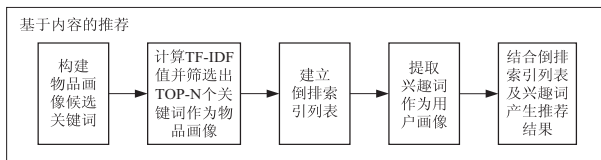


图 1 基于内容推荐算法结构

通过物品的分类、属性特征等信息构建物品画像的候选关键词。利用 TF-IDF (term frequency-inverse document frequency) 计算出候选关键词的权重值:

$$W(d, t) = \frac{n_{i,j}}{\sum_k n_{k,j}} * \lg\left(\frac{N}{n_i + 1}\right) \quad (1)$$

其中, $W(d, t)$ 为计算出的候选关键词 t 占文本 d 的比例; $n_{i,j}$ 为该词在文本中出现的次数; $\sum_k n_{k,j}$ 为文本中所有词出现的次数总和; N 为文本语料库中文本的总数; n_i 为该语料库中含候选关键词 t 的文本个数^[14]。

利用 TF-IDF 计算出的权重值选取 TOP-N 个候选关键词作为物品画像,建立倒排索引列表。根据用户在登录系统时的反馈信息收集用户兴趣词,并结合其属性特征构建用户画像。运用用户画像及物品画像进行相似度计算,其计算方法有很多种,如余弦相似度^[15]、Jaccard 相似度^[16]及 Pearson 相似度^[17]等。其中余弦相似度在计算过程中不考虑向量长度。Jaccard 相似度计算的是两个集合的交集元素个数占并集的比例,适合用于布尔向量表示。Pearson 相似度在计算过程中不仅考虑向量夹角余弦值还考虑向量长度。因此,选取 Pearson 相似度作为相似度计算方法:

$$P(A, B) = \frac{\sum_{i=1}^n (A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^n (A_i - \bar{A})^2} \sqrt{\sum_{i=1}^n (B_i - \bar{B})^2}} \quad (2)$$

最后,将相似度最高的 TOP-N 项作为推荐结果进行输出。

2.2 基于协同过滤的推荐算法

在基于协同过滤的推荐算法中,根据计算方式的不同可将其分为基于邻域和基于模型的推荐算法。文中采用的是 Yehuda Koren 等^[18]提出的基于隐语义模型(latent factor model, LFM)的推荐算法,其结构如图 2 所示。

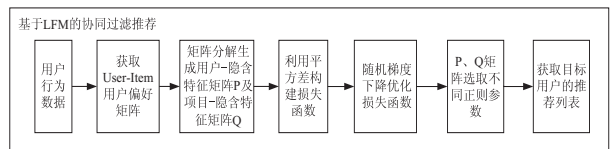


图 2 基于 LFM 的推荐算法结构

LFM 算法是一种基于矩阵分解(matrix factorization, MF)的协同过滤算法^[19-20],其原理如图 3 所示。

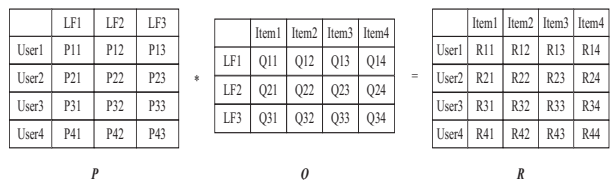


图 3 LFM 算法原理

其中, P 矩阵是 User-LF 矩阵,即用户-隐含因子矩阵; Q 矩阵是 LF-Item 矩阵,即隐含因子-物品矩阵; R 矩阵是 User-Item 矩阵,即用户-物品偏好矩阵。其中矩阵值 P_{11} 为用户 1 在隐含因子 1 上的权重值;矩阵值 Q_{11} 为隐含因子 1 在物品 1 上的权重值;矩阵值

R_{11} 为预测用户 1 对物品 1 的喜爱程度。在该算法的实现过程中,运用矩阵分解技术将 R 矩阵进行降维,使其分解成 P 矩阵与 Q 矩阵,通过隐含因子将用户与物品进行关联:

$$R_{ui} = PQ \quad (3)$$

再利用 P 矩阵与 Q 矩阵的乘积还原出 R 矩阵,将原有的稀疏矩阵进行填充,预测用户对物品的喜爱程度, F 为隐含因子数:

$$\hat{R}_{ui} = \sum_{f=1}^F P_{u,f} Q_{i,f} \quad (4)$$

如图 3,有 4 个用户,3 个物品及 3 个隐含因子。以用户 1 为例,用户 1 对隐含因子 LF_1 、 LF_2 、 LF_3 的关联程度分别为 P_{11} 、 P_{12} 、 P_{13} 。物品 1 与隐含因子 LF_1 、 LF_2 、 LF_3 的关联程度分别为 Q_{11} 、 Q_{21} 、 Q_{31} 。因此,用户 1 对物品 1 的感兴趣程度为:

$$R_{11} = P_{11} Q_{11} + P_{12} Q_{21} + P_{13} Q_{31} \quad (5)$$

利用平方差构建损失函数:

$$\text{Cost} = \sum_{u,i \in R} (R_{ui} - \hat{R}_{ui})^2 = \sum_{u,i \in R} (R_{ui} - \sum_{f=1}^F P_{u,f} Q_{i,f})^2 \quad (6)$$

为防止过拟合,加入正则化参数:

$$\text{Cost} = \sum_{u,i \in R} (R_{ui} - \sum_{f=1}^F P_{u,f} Q_{i,f})^2 + \lambda (\sum_u P_{u,f}^2 + \sum_i Q_{i,f}^2) \quad (7)$$

对损失函数求偏导:

$$\frac{\partial}{\partial P_{u,f}} \text{Cost} = -2 \sum_{u,i \in R} (R_{ui} - \sum_{f=1}^F P_{u,f} Q_{i,f}) (-Q_{i,f}) + 2\lambda P_{u,f} \quad (8)$$

$$\frac{\partial}{\partial Q_{i,f}} \text{Cost} = -2 \sum_{u,i \in R} (R_{ui} - \sum_{f=1}^F P_{u,f} Q_{i,f}) (-P_{u,f}) + 2\lambda Q_{i,f} \quad (9)$$

运用随机梯度下降法迭代计算,更新矩阵 P 和 Q ,

从而找到最优的 P 矩阵与 Q 矩阵:

$$P_{u,f} = P_{u,f} + a [(R_{ui} - \sum_{f=1}^F P_{u,f} Q_{i,f}) Q_{i,f} - \lambda_1 P_{u,f}] \quad (10)$$

$$Q_{i,f} = Q_{i,f} + a [(R_{ui} - \sum_{f=1}^F P_{u,f} Q_{i,f}) P_{u,f} - \lambda_2 Q_{i,f}] \quad (11)$$

其中, a 为学习率。

3 实验方法

3.1 实验数据

本实验采用的数据从马匹竞拍 App 的数据库中获取,数据集由两部分组成。第一部分是用户到马匹的评分数据集,收集 80 个用户在马匹竞拍 App 上的行为数据共 22 423 条。用户行为数据主要包括是否浏览,是否收藏,是否参与竞拍,用户浏览的马匹所具备的品种、价格、毛色、地区、性别、年龄等。并对其分别设置权重,计算用户对马匹的喜爱程度得出用户到马匹的评分数据集。其评分范围为 0~5,其结构如表 1 所示。

表 1 用户评分

用户 ID	马匹 ID	评分	评分日期
1	1	3	20200729
1	2	4	20200729
1	3	4	20200729
1	4	5	20200730
1	5	2	20200804
...

第二部分是马匹竞拍 App 中 925 条马匹数据。每条马匹数据都含有马匹的 RFID、品种、价格、毛色、地区、性别、年龄等属性,其结构如表 2 所示。

表 2 马匹信息

马匹 ID	马匹 RFID	品种	毛色	性别	地区	价格	年龄
1	900038000367247	焉耆马	锈黑毛	母马	巴音郭楞蒙古自治州	1w-5w	成驹
2	900038000338929	汉诺威	黑骝毛	母马	昌吉回族自治州	1w-5w	成驹
3	220170577100534	伊犁马	铁青毛	母马	伊犁哈萨克自治州	1w 以下	幼驹
4	900038000352178	哈萨克马	黄骝毛	母马	巴音郭楞蒙古自治州	1w 以下	幼驹
5	201907167000033	英纯血	黑棕毛	公马	喀什地区	1w-5w	成驹
6	900038000339461	哈萨克马	红骝毛	母马	塔城地区	1w-5w	成驹
...

在实验中,将数据集进行进一步划分,使其 80% 作为训练集,20% 作为测试集。

3.2 混合推荐算法

文中采用基于内容和协同过滤的混合推荐算法来实现竞拍 App 中马匹的推荐。将参与竞拍的马匹向

用户进行个性化推荐,产生较为可靠的推荐结果。使得用户能够快速找到心仪马匹,提升用户参与竞拍的积极性,增加参与同场竞拍人数,降低流拍比例,优化马匹竞拍 App 首页商品展示界面。其流程如图 4 所示。

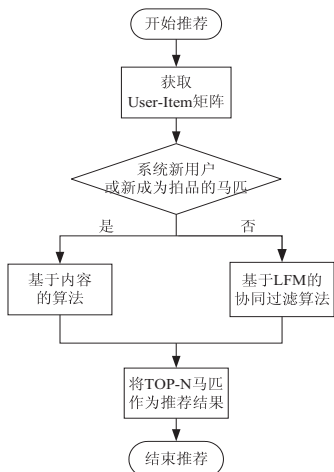


图 4 混合推荐算法流程

收集用户在马匹竞拍 App 上的行为数据,获取 User-Item 矩阵。判断当前用户是否为马匹竞拍 App 新注册用户或马匹竞拍 App 中是否有新成为拍品的马匹。如果既不是新用户,也没有新成为拍品的马匹,那么通过矩阵分解将 User-Item 矩阵做降维处理,使其分解成用户-隐含因子矩阵 P 及马匹-隐含因子矩阵 Q ,用隐含因子做连接用户与马匹的桥梁。使用平方差构建损失函数,再运用随机梯度下降法迭代计算,不断更新矩阵 P 和 Q ,并加入避免过拟合的正则参数,从而找到最优的 P 和 Q 矩阵。建立 LFM 模型,对用户未评分马匹的兴趣度进行预测,将 TOP-N 项作为推荐结果进行输出。以马匹竞拍 App 中 ID 为 10 的用户为例,当 $N = 10$ 时,推荐结果如表 3 所示。

表 3 基于协同过滤推荐结果

用户 ID	马匹 RFID	马匹属性
10	900038000367248	[焉耆马,黄骠毛,母马,巴音郭楞蒙古自治州,1w-5w,成驹]
10	900038000367249	[焉耆马,黄骠毛,母马,巴音郭楞蒙古自治州,1w-5w,成驹]
10	900038000367250	[焉耆马,红骠毛,母马,巴音郭楞蒙古自治州,1w-5w,成驹]
10	900038000367252	[焉耆马,朽栗毛,母马,巴音郭楞蒙古自治州,1w-5w,成驹]
10	900038000367255	[焉耆马,红骠毛,母马,巴音郭楞蒙古自治州,1w,幼驹]
10	900038000367256	[焉耆马,淡黑毛,母马,巴音郭楞蒙古自治州,1w-5w,成驹]
10	900038000367257	[焉耆马,朽栗毛,母马,巴音郭楞蒙古自治州,1w-5w,成驹]
10	900038000367258	[焉耆马,红骠毛,母马,巴音郭楞蒙古自治州,1w,幼驹]
10	900038000367259	[焉耆马,红骠毛,母马,巴音郭楞蒙古自治州,1w,幼驹]
10	900038000367260	[焉耆马,朽栗毛,母马,巴音郭楞蒙古自治州,1w-5w,成驹]

当有新注册用户使用马匹竞拍 App 或者有新马匹成为拍品时,数据库新增的数据没有任何历史行为,就会面临冷启动问题。此时利用基于内容的方式进行推荐,根据马匹的属性标签构建物品画像的候选关键

词。利用 TF-IDF 算出关键词的权重值,从中选取 TOP-N 个作为马匹画像。此处选取部分马匹画像进行展示,如表 4 所示。

表 4 马匹画像

马匹 ID	马匹属性	TF-IDF 值
325	[喀什地区,褐骠毛,伊犁马,1w]	[0.924616759492828,0.295859606931123,0.14202030893327,0.126633122325426]
326	[褐骠毛,伊犁哈萨克自治州,伊犁马,1w-5w]	[0.748545477766145,0.397294025923399,0.359321304809627,0.29934805408898]
327	[栗毛,伊犁哈萨克自治州,伊犁马,1w]	[0.701815824254909,0.370852462671326,0.335406983453149,0.29906732271914]
328	[褐骠毛,伊犁哈萨克自治州,伊犁马,1w]	[0.700243112741337,0.371657318959921,0.336134912878037,0.29971638464981]
329	[黑棕,伊犁哈萨克自治,伊犁马,公马]	[0.712051072757356,0.396149634233756,0.358286292228756,0.29937174352628]
330	[红骠毛,伊犁哈萨克自治州,伊犁马,1w-5w]	[0.654646760767797,0.452927048122322,0.409637012629694,0.34126571670582]
331	[黑棕,伊犁哈萨克自治,伊犁马,1w]	[0.683256258428058,0.3801296384767,0.343797461783619,0.3065487348972766]
332	[青毛,伊犁哈萨克自治,伊犁马,1w]	[0.77160519223435,0.331153369149135,0.299502265150378,0.267052700169445]
333	[黑棕,伊犁哈萨克自治,伊犁马,公马]	[0.712051072757356,0.396149634233756,0.358286292228756,0.29937174352628]
334	[红骠毛,伊犁哈萨克自治州,伊犁马,1w]	[0.69711243522337,0.37324897345389,0.337574439607013,0.3009999458934856]

根据用户在登录时对马匹进行的反馈信息收集用户兴趣词,结合用户的属性特征构建用户画像。以马

匹竞拍 App 中新注册 ID 为 144 的用户为例,构建的用户画像如表 5 所示。

表 5 用户画像

用户 ID	马匹属性	TF-IDF 值
144	‘伊犁哈萨克自治州’	1.0
144	‘伊犁马’	0.841 6
144	‘公马’	0.831 7
144	‘黑棕’	0.811 9
144	‘成驹’	0.693 1
144	‘1w-5w’	0.495
144	‘1w’	0.475 2

运用用户画像从物品画像中寻找最匹配的 TOP-N 项作为推荐结果进行输出。当 $N = 10$ 时,推荐结果如表 6 所示。

表 6 基于内容推荐结果

用户 ID	马匹 ID	相似度
144	810	0.971 271 634 101
144	90	0.971 130 132 751
144	18	0.970 082 938 671
144	450	0.970 041 392 337
144	14	0.969 721 317 291
144	399	0.969 713 622 928
144	120	0.969 686 388 969
144	368	0.969 326 225 775
144	237	0.969 166 576 862
144	367	0.969 144 940 376

4 实验结果及分析

4.1 评估准则

在本实验中选取的评估准则为以下三种:

(1) 准确率 (Precision): 在产生的 TOP-N 个推荐

结果中用户喜欢的物品个数和所有被推荐物品个数 N 的比值,用于衡量推荐系统对用户偏好的预测能力。其计算公式如下:

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

其中, U 为用户总量, $R(u)$ 为对每个用户产生的推荐结果集, $T(u)$ 为用户实际的行为结果集。

(2) 召回率 (Recall): 在产生的 TOP-N 个推荐结果中用户喜欢的物品个数和测试集中用户喜欢物品个数的比值,用于衡量推荐系统的实际推荐效果。其计算公式如下:

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

(3) 覆盖率 (Coverage): 产生的 TOP-N 个推荐结果个数与物品总数的比值,用于描述推荐系统对长尾物品的发掘能力。其计算公式如下:

$$\text{Coverage} = \frac{|U_{u \in U} R(u)|}{I} \quad (14)$$

其中, I 为物品总数。

4.2 实验结果分析

为证明本实验提出的方法推荐质量更佳,使用提出的基于内容和基于协同过滤的混合推荐算法与基于用户的协同过滤推荐算法和基于物品的协同过滤推荐算法进行对比实验。分别计算在推荐长度为 5、10、15、20、25 和 30 时,三种方法的准确率、召回率和覆盖率,其结果如表 7 ~ 表 9 所示。并绘制出推荐长度为 5、10、15、20、25 和 30 时的变化趋势,如图 5 ~ 图 7 所示。

表 7 准确率

算法	Top-5	Top-10	Top-15	Top-20	Top-25	Top-30
基于用户	47.00	39.00	31.90	29.90	24.18	23.18
基于物品	40.00	31.90	28.50	25.79	22.80	21.80
混合推荐	63.33	52.60	51.00	42.85	38.00	37.00

观察表 7 和图 5 可知,文中采用的融合内容与协同过滤的混合推荐方法在指定推荐马匹个数的情况下

准确率优于基于用户和基于物品的推荐方法。当推荐长度大于或等于 25 时,准确率逐渐趋于稳定状态。

表 8 召回率

算法	Top-5	Top-10	Top-15	Top-20	Top-25	Top-30
基于用户	19.50	21.00	31.25	31.50	35.70	36.10
基于物品	28.00	31.25	35.50	37.50	37.60	38.50
混合推荐	23.50	26.00	38.50	40.00	46.26	46.96

观察表 8 和图 6 可知,随着推荐长度的增加,混合推荐方法的召回率也逐渐优于基于用户和基于物品的

推荐方法。当推荐长度等于 15 时,混合推荐方法的召回率超过其基于用户和基于物品的推荐方法。

表 9 覆盖率

算法	Top-5	Top-10	Top-15	Top-20	Top-25	Top-30
基于用户	6.45	8.74	10.89	12.07	15.20	15.96
基于物品	3.98	5.93	7.18	9.11	13.89	14.21
混合推荐	7.18	12.09	18.54	20.26	21.56	22.00

观察表 9 和图 7 可知,混合推荐方法的覆盖率也 覆盖率逐渐趋于稳定状态。
 优于其他两种推荐方法。当推荐长度大于等于 25 时,

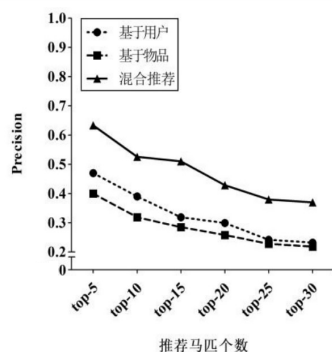


图 5 准确率对比

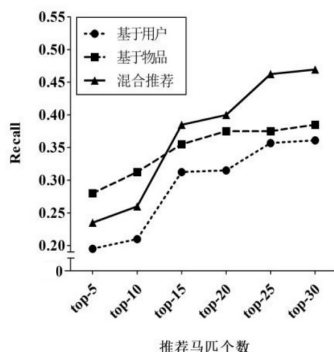


图 6 召回率对比

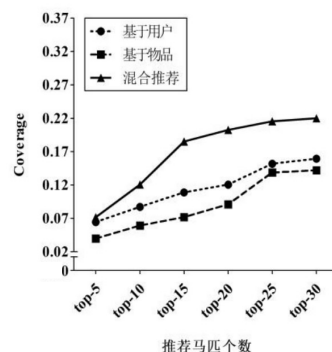


图 7 覆盖率对比

综上所述,混合推荐算法不仅能够解决冷启动问题,还具有更高的推荐质量,可以更好地为用户提供推荐服务。

5 结束语

随着电子设备智能化的发展,推荐系统在移动电商平台中得到了广泛应用。将统计分析方法与机器学习相结合,提出的融合内容与协同过滤的混合推荐算法能够规避单一算法的弊端,融合其各自优势。敏锐地捕捉用户需求,使用户能够从移动电商平台的大量数据中快速获取所需信息,产生有针对性的高质量推荐结果,减少用户时间成本,具有一定的应用价值。在真实的数据集下,将提出的混合推荐算法同传统的推荐算法进行对比实验,验证了该算法在准确率、召回率和覆盖率上具有一定的优越性。

参考文献:

- [1] 樊艳清,梁宏宇,纪佳琪. 协同过滤算法中相似度计算问题研究[J]. 计算机技术与发展,2020,30(8):91-96.
- [2] 张应辉,司彩霞. 基于用户偏好和项目特征的协同过滤推荐算法[J]. 计算机技术与发展,2017,27(1):16-19.
- [3] 刘宇,朱文浩. 基于内容和标签权重的混合推荐算法[J]. 计算机与数字工程,2020,48(4):773-777.
- [4] GHAZANFAR M A, PRUGEL-BENNETT A. A scalable, accurate hybrid recommender system[C]//Third international conference on knowledge discovery and data mining. Phuket: IEEE, 2010: 94-98.
- [5] 乔雨,李玲娟. 推荐系统冷启动问题解决策略研究[J]. 计算机技术与发展,2018,28(2):83-87.

- [6] 李忠俊,周启海,帅青红. 一种基于内容和协同过滤同构化整合的推荐系统模型[J]. 计算机科学,2009,36(12):142-145.
- [7] KASSAK O, KOMPAN M, BIELIKOVA M. Personalized hybrid recommendation for group of users: top-N multimedia recommender[J]. Information Processing and Management, 2016, 52(3): 459-477.
- [8] 何锴琦,马宇骁,张炎,等. 一种基于数据的 GitHub 项目个性化混合推荐方法[J]. 吉林大学学报:理学版,2020, 58(6):1399-1406.
- [9] 冉从敬,宋凯. 基于混合方法的高校专利个性化推荐模型构建[J]. 情报理论与实践,2020,43(10):93-98.
- [10] 王末,郑晓欢,王卷乐,等. 基于混合过滤的地学数据个性化推荐方法设计与实现[J]. 地理研究,2018,37(4):814-824.
- [11] 叶小榕,邵晴. 结合物联网和室内定位的手机图书馆推荐系统[J]. 科技导报,2016,34(23):127-136.
- [12] 杨凯,王利,周志平,等. 基于内容和协同过滤的科技文献个性化推荐[J]. 信息技术,2019,43(12):11-14.
- [13] 王光,张杰民,董帅含,等. 基于内容的加权粒度序列推荐算法[J]. 计算机工程与科学,2018,40(3):564-570.
- [14] 许甜华,吴明礼. 一种基于 TF-IDF 的朴素贝叶斯算法改进[J]. 计算机技术与发展,2020,30(2):75-79.
- [15] LIANG Changyong, LENG Yajun. Collaborative filtering based on information-theoretic co-clustering[J]. International Journal of Systems Science, 2014, 45(3): 589-597.
- [16] REINA D G, TORAL S L, JOHNSON P, et al. Improving discovery phase of reactive ad hoc routing protocols using Jaccard distance[J]. The Journal of Supercomputing, 2014, 67(1):131-152.

(下转第 37 页)