

# 个性化特征的电子商务智能推荐系统

李元吉

(河南省驻马店财经学校, 河南 驻马店 463000)

**摘 要:** 以准确向用户推荐商品,提升电子商务网站销售量为目标,设计基于个性化特征的电子商务智能推荐系统。系统以个性化推荐引擎为核心,采集交易事务、商品特征、用户评价等数据,利用基于个性化特征的协同过滤推荐算法计算商品间相似度,确定新商品的近邻,根据近邻用户对新商品的评价结果选择商品进行推荐。测试结果表明,该系统的电子商务商品推荐误差小,有利于提升电子商务网站交易率,而且电子商务商品推荐性能明显优于其他推荐系统。

**关键词:** 个性化特征; 电子商务; 智能推荐; 数据挖掘; 商品特征

**中图分类号:** TP393 **文献标识码:** A **文章编号:** 1009-2552(2021)03-0131-05

**DOI:** 10.13274/j.cnki.hdzj.2021.03.024

## Personalized e-commerce intelligent recommendation system

LI Yuan-ji

(Henan Zhumadian Finance&Economics School, Zhumadian 463000, Henan Province, China)

**Abstract:** In order to accurately recommend goods to users and improve the sales volume of e-commerce websites, an intelligent e-commerce recommendation system based on personalized features is designed. The system takes the personalized recommendation engine as the core, collects the transactions, commodity characteristics, user evaluations and other datas, uses collaborative filtering recommendation algorithm based on personalized features to calculate the similarity between products, determines the neighbors of new products, predicts the user's evaluation of new products according to the evaluation results, and selects products for recommendation according to the evaluation results. The test results show that the e-commerce product recommendation error of the system is small, which is beneficial to improve the transaction rate of e-commerce website, and the e-commerce commodity recommendation performance is obviously better than other recommendation systems.

**Key words:** personalization feature; e-commerce; intelligent recommendation; data mining; commodity features

## 0 引 言

电子商务网站数量迅速上升,将电子商务网站浏览者变为实际消费者,满足消费者需求,通过不同形式提升消费者忠诚度是各电子商务网站面临的首要问题<sup>[1-3]</sup>。在此背景下研究具有个性化特征的电子商务推荐系统具有重要意义。

**作者简介:** 李元吉(1986-),男,硕士,讲师,研究方向为电子商务与网络营销。

以往普遍使用的电子商务推荐系统主要有,基于面向集成系统的推荐系统<sup>[4]</sup>和基于多分类支持向量机的推荐系统<sup>[5]</sup>,前者在生成推荐信息时忽略了新到商品推荐,后者则无法为新用户提供个性化推荐服务。针对当前复杂的电子商务网站结构以及用户与商品种类数量持续上升的趋势<sup>[6]</sup>,设计了基于个性化特征的电子商务智能推荐系统,利用基于个性化特征的协同过滤推荐算

法,以协同过滤技术为基础,弥补推荐过程中冷启动与数据稀疏等缺陷。

### 1 个性化特征的电子商务智能推荐系统

#### 1.1 系统逻辑架构

基于个性化特征的电子商务智能推荐系统采用三层逻辑架构<sup>[7]</sup>:数据层、业务逻辑层和表现层,如图1所示,其中实线箭头表示请求处理,虚线箭头表示个性化处理。

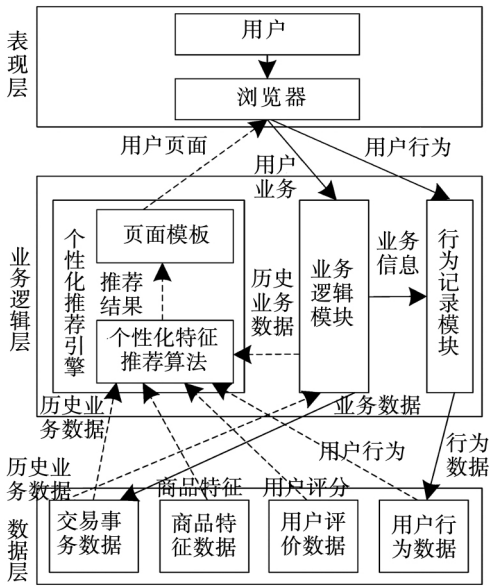


图1 系统整体逻辑架构

①数据层中主要利用 Web 挖掘技术在电子商务网站用户使用记录中挖掘电子商务网站用户的交易事务数据、商品特征数据、用户评价数据以及用户行为数据,该层利用分布式结构与 LDAP 技术为基础,负责电子商务网站实际数据存储与搜索,为业务逻辑层提供数据存储与管理服务<sup>[8]</sup>。

②业务逻辑层是系统的核心,包括行为记录模块、业务逻辑模块、个性化推荐引擎组成。行为记录模块的主要功能是记录用户的浏览信息与商品评价信息;业务逻辑模块连接商品购买流程<sup>[9]</sup>;个性化推荐引擎是业务逻辑层的核心,其中包含个性化特征推荐算法和页面模板两部分,前者负责根据数据挖掘结果提供推荐信息,后者基于 XML 技术实现推荐信息的个性化页面呈现。

③表现层负责用户与电子商务网站之间的交互<sup>[10]</sup>,基于 HTTP(S) 通信协议与网络浏览器完

成人机交互与信息流程的功能,反馈推荐信息。

#### 1.2 用户使用记录的挖掘

基于个性化特征的电子商务信息推荐以电子商务网站内用户使用记录信息为基础,因此需要利用 Web 挖掘技术对电子商务网站中的用户使用记录信息进行挖掘<sup>[11]</sup>。用户使用记录挖掘模型主要由数据预处理、模式识别与模式分析等部分组成,如图2所示。通过挖掘电子商务网站内用户使用记录信息,能够统计出用户频繁访问页面、用户交易量、商品特征量、单位时间访问量、用户评价等信息,基于这些信息,采用基于个性化特征的协同过滤推荐算法为用户提供个性化信息推荐服务。

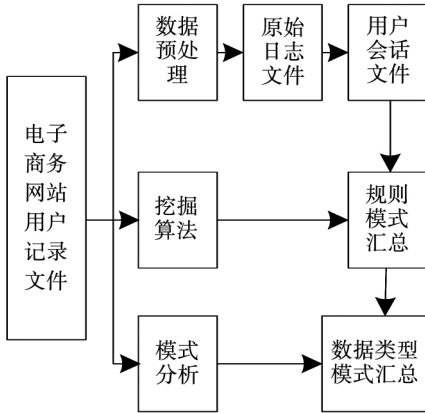


图2 用户使用记录挖掘模型

#### 1.3 基于个性化特征的协同过滤推荐算法

基于个性化特征的协同过滤推荐算法,基于商品的属性分析,确定新商品的相似商品;根据用户对相似商品的评价判断新商品的评分;利用协同过滤推荐技术确定目标用户最近邻居,基于最近邻居对于商品的评价向目标用户推荐新商品。

##### 1.3.1 算法的输入与输出

基于个性化特征的协同过滤推荐算法的输入和输出分别如①~④和⑤所示:

①商品特征量表、 $M \times L$  维矩阵、商品数量与特征量分别为  $M$  和  $L$  个,用  $PA_i (1 \leq i \leq L)$  表示商品特征属性。

②用户评价表、 $n \times m$  维矩阵、商品数量与用户数量为  $M$  和  $N$  个,用  $C_{(n,m)}$  表示商品评价值,其中  $1 \leq n \leq N, 1 \leq m \leq M$ 。

③ 输入阈值包括  $s$  和  $w$  表示的商品相似度阈值和邻居用户相似度阈值。

④ 分别输入  $P$ 、 $C$  和  $K$  表示的相似商品数量、相似用户数量和推荐商品数量。

⑤ 任意用户  $i$  最感兴趣的若干商品。

### 1.3.2 算法实现过程

基于个性化特征的协同过滤推荐算法实现过程中共分为三个环节,分别是:①基于商品的特征量表确定新商品的最近邻居集;②以用户对邻居商品评价为基础,判断用户对目标商品的评价;③基于协同过滤推荐算法生成推荐信息。

新商品最近邻居集确定:在商品特征量表内产生新信息的条件下(电子商务网站内出现新商品),确定新商品的最近邻居集(离线状态下)。对全部的商品集合  $j \in I$  确定商品  $j$  邻居项目,也就是同商品  $j$  相似度最高的商品,用  $SI_j$  表示。用  $Sim(j, p)$  表示商品同其他任意商品  $q \in 1, 2, \dots, j-1, j+1, \dots, I$  之间的相似度。逐个对比目标商品的特征量与其他任意商品的特征量,若  $PA_{j,i}$  与  $PA_{q,i}$  一致,则  $Sim(j, p) = Sim(j, p) + 1$ ; 若  $PA_{j,i}$  与  $PA_{q,i}$  不一致,则  $Sim(j, p)$  不变<sup>[12]</sup>。依照  $Sim(j, p)$  的大小进行由大到小的排序,对全部  $Sim(j, p), p \in 1, 2, \dots, j-1, j+1, \dots, I$  取前  $P$  个  $Sim(j, p)$  值最大的商品,将这  $P$  个商品作为目标商品的邻居商品  $SI_j$ 。用户对目标商品评价的判断:依照用户对于邻居商品的评价,利用式(1)判断用户  $a$  对目标商品  $j$  的评价  $PE_{a,j}$ :

$$PE_{a,j} = \frac{\sum_{n \in SI_j} Sim(j, n) * CE_{a,n}}{\sum_{n \in SI_j} Sim(j, n)} \quad (1)$$

式中,  $CE_{a,n}$  表示用户  $a$  对邻居商品集内商品  $n$  的评价。

通过对新商品的评价,用户评价表内前部的商品均产生评分<sup>[13]</sup>。利用式(2)计算任意用户  $i$  对商品  $j$  的评价:

$$CE_{i,j} = \begin{cases} CE_{i,j} & \text{if user } i \text{ rated item } j \\ PE_{i,j} & \text{if user } i \text{ not rated item } j \end{cases} \quad (2)$$

式中,  $PE$  表示系统判断评价,用户实际评价商品过程中采用用户评价  $CE$  取代  $PE$ ,由此可提升系

统推荐精度。

推荐信息的生成:基于新商品最近邻居集与用户对目标商品评价的判断结果,采用协同过滤推荐算法生成用户推荐商品。协同过滤推荐算法推荐过程包含查询最近邻居与生成推荐结果两部分。查询最近邻居的主要目的是针对任意用户  $i$ , 在全部用户空间内确定用户集合  $C = \{C_1, C_2, \dots, C_n\}, i \notin C$ , 令任意用户  $i$  与  $C_1$  之间的相似度  $Sim(i, C_1)$  值  $> Sim(i, C_2)$  值  $> \dots > Sim(i, C_n)$  值。在此过程中,利用 Pearson 相关系数评估相似度:

$$Sim(i, a) = \frac{\sum_{C \in I_{i,a}} (R_{C,i} - \bar{R}_i) (R_{C,a} - \bar{R}_a)}{\sqrt{\sum_{C \in I_{i,a}} (R_{C,i} - \bar{R}_i)^2} \sqrt{\sum_{C \in I_{i,a}} (R_{C,a} - \bar{R}_a)^2}} \quad (3)$$

式中  $R_{C,i}$ 、 $\bar{R}_i$  和  $\bar{R}_a$  分别为用户  $i$  对  $C$  的评价、用户  $i$  对商品的评价均值和用户  $a$  对商品的评价均值。

依照  $Sim(i, k)$  由大至小的排序,取  $Sim(i, k), k \in 1, 2, \dots, i-1, i+1, \dots, n$  内前  $C$  个最大的  $Sim(i, k)$  值对应的用户为目标用户的邻居用户集合  $Neighbor_{i,k}$ 。

基于上述的相似度评估能够确定目标用户的最近邻居集合,基于最近邻居集合最近邻居对商品的评价,利用式(4)能够确定目标用户  $i$  对商品的评价:

$$P_{i,j} = \bar{R}_i + \frac{\sum_{a \in Neighbor} Sim(i, a) * (R_{a,j} - \bar{R}_a)}{\sum_{a \in Neighbor} (|Sim(i, a)|)} \quad (4)$$

式中,  $Sim(i, a)$  和  $R_{a,j}$  分别表示用户  $i$  与用户  $a$  之间的相似性和用户  $a$  对商品  $j$  的评价。

基于以上过程能够判断用户对全部未评价商品的评价,基于判断结果的高低,选取最高的前  $K$  个商品作为推荐结果,利用页面模板进行规划,通过表示层向用户进行个性化推荐。

## 2 系统的仿真测试与分析

实验为验证本文设计的基于个性化特征的电子商务智能推荐系统的应用性能,以某服装类电

子商务网站为研究对象,将本文系统应用于研究对象商品推荐中。

2.1 系统性能研究

本文系统中存在一个重要参数,即相似用户数量。基于相似用户对商品的评价判断目标用户对商品的评价,向目标用户进行商品推荐。以准确率、召回率、覆盖率和流行度为本文系统性能的评价指标,各评价指标计算公式如下。

准确率 =  $\frac{\text{用户喜欢的商品}}{\text{全部推荐商品}} \times 100\%$  (5)

召回率 =  $\frac{\text{推荐商品}}{\text{用户喜欢的全部商品}} \times 100\%$  (6)

覆盖率 =  $\frac{\text{推荐商品}}{\text{全部商品}} \times 100\%$  (7)

流行度 =  $\frac{\text{同推荐商品产生关系的用户}}{\text{全部用户}} \times 100\%$  (8)

图3 所示为不同相似用户数量条件下,本文

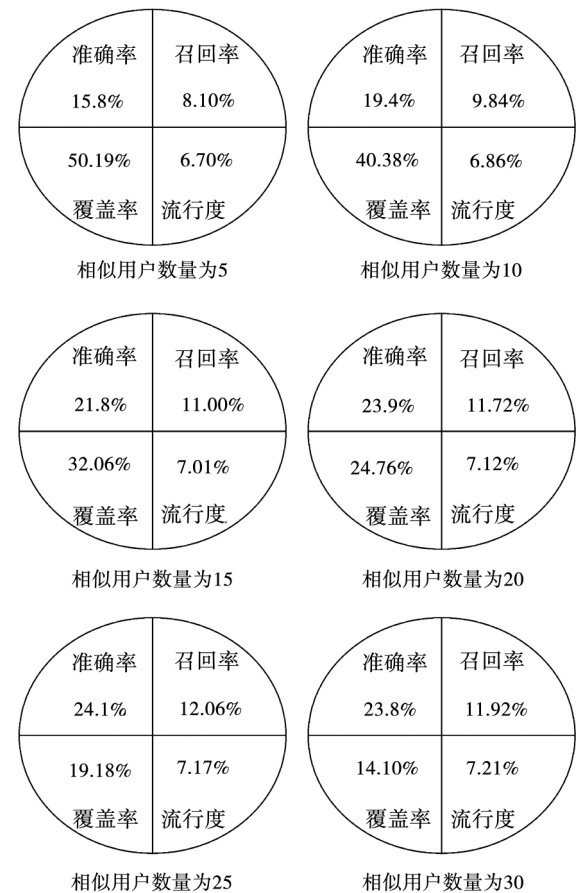


图3 本文系统性能评价结果

系统性能评价结果。由图3 能够得到,相似用户数量的波动对于本文系统性能产生明显影响。

①准确率与召回率: 本文系统的准确率与召回率同相似用户数量之间不存在线性相关性,当相似用户数量为25 时,系统的准确率与召回率相对较高。同时系统的准确率与召回率对相似用户数量之间的相关性并不高,因此相似用户数量取值在相应范围内,系统的准确率与召回率差异并不显著。

②覆盖率: 本文系统的覆盖率指标同相似用户数量之间为反比例关系,即相似用户数量越大,系统覆盖率评价越低。

③流行度: 本文系统的流行度指标同相似用户数量之间存在线性相关性,即相似用户数量越大,流行度评价结果越高,这主要是由于相似用户数量决定系统推荐过程中参考多少与目标用户相似的其它用户对商品的评价,参考的用户越多,系统推荐结果越趋近于热门商品。

综合图3 结果可知,本文系统中相似用户数量为20 时,可得到最优推荐结果。

2.2 推荐结果的对比

以文献[4]中面向集成系统的推荐系统与文献[5]中基于多分类支持向量机的推荐系统为对比系统,以平均绝对偏差(用户对商品评价的判断结果与实际用户评价间的偏差)为推荐结果质量对比指标,对比本文系统与对比推荐结果的质量,结果如图4 所示。分析图4 得到,采用三个推荐系统针对研究对象内60 个用户进行商品推荐,

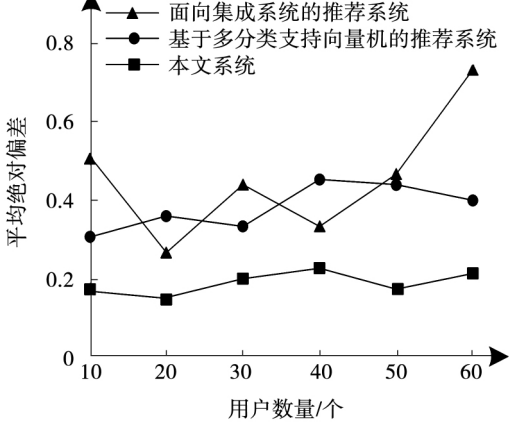


图4 推荐结果质量对比

本文系统推荐结果的平均绝对偏差低于 0.2,与两个对比系统相比明显下降,由此说明本文系统推荐结果具有较高精度。

图 5 所示为客观环境一致条件下,本文系统与两个对比系统进行商品推荐过程花费的时间对比结果。

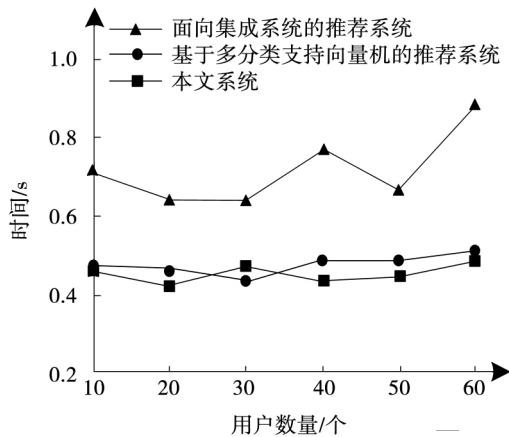


图 5 商品推荐时间对比结果

分析图 5 得到,本文系统进行商品推荐所用时间与基于多分类支持向量机的推荐系统所用时间大致相同,面向集成系统的推荐系统所用时间显著高于另两个系统。由此说明本文系统的推荐效率上也存在一定优势。

### 2.3 应用分析

为测试本文系统在实际应用中的应用效果,采用本文系统挖掘研究对象过去三年不同月份的用户行为数据,以用户的浏览量与交易率为分析指标,将过去三年研究对象中用户的浏览量与交易率均值作为对比标准,对比研究对象采用本文系统后用户的浏览量与交易率变化,结果如图 6 所示。由图 6 得到,每年的 6 月、11 月和 12 月是电子商务网站交易的高峰月。研究对象采用本文系统后,各月浏览量平均上升 12% 左右,以 2 月、6 月、11 月和 12 月上升比例最高;各月交易率平均上升 4% 左右,6 月、11 月和 12 月上升比例近 6%。以上结果说明本文系统能够显著提升研究对象用户浏览量与交易率。

### 3 结束语

随着电子商务网站数量迅速上升,各大电子商务网站为吸引消费者,提升市场占有率,各显

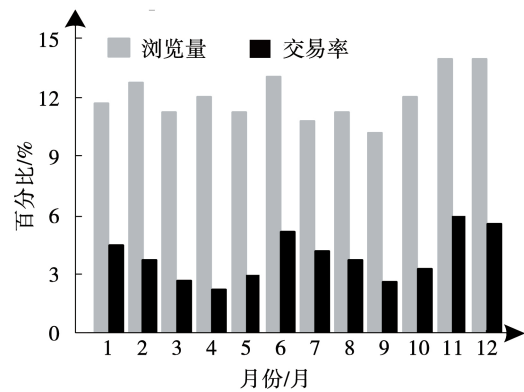


图 6 浏览量与交易率变化

所长。本文设计基于个性化特征的电子商务智能推荐系统,基于电子商务网站内用户行为数据,为用户准确推荐感兴趣的商品,以此吸引消费者,提升交易率。

### 参考文献:

- [1] 李桃迎,李墨,李鹏辉. 基于加权 Slope One 的协同过滤个性化推荐算法[J]. 计算机应用研究, 2017, 34(8): 2264-2268.
- [2] 田保军,胡培培,杜晓娟,等. Hadoop 下基于聚类协同过滤推荐算法优化的研究[J]. 计算机工程与科学, 2016, 38(8): 1615-1624.
- [3] 肖会敏,张锬,崔春生. 基于协同过滤的移动电子商务推荐算法[J]. 系统科学与数学, 2016, 36(8): 1265-1274.
- [4] 吴文李,范小朋,周庚申,等. 面向企业生产集成系统的物料推荐系统[J]. 中国机械工程, 2019, 30(15): 1856-1865.
- [5] 崔建双,车梦然. 基于多分类支持向量机的优化算法智能推荐系统与实证分析[J]. 计算机工程与科学, 2019, 41(1): 153-160.
- [6] Tomohiro M, Bijoy C, Nattapong K, et al. Designing a hadoop system based on computational resources and network delay for wide area networks[J]. Telecommunication Systems, 2019, 70(1): 13-25.
- [7] 李宇琦,陈维政,闫宏飞,等. 基于网络表示学习的个性化商品推荐[J]. 计算机学报, 2019, 42(8): 1767-1778.
- [8] 黄逸珺,杜梦甜,傅玉婷. 基于用户感知的个性化推荐系统效果研究——以淘宝电商平台为例[J]. 北京邮电大学学报: 社会科学版, 2018, 20(5): 24-31.

(下转第 142 页)

型,利用 FCM 算法进行日志数据分类,通过最小二乘法拟合直线方程计算数据运维容量,实现了电力信息通信数据智能运维技术分析。

通过实验得出,本文方法的故障检测耗时较短,电力信息数据运维智能性较好。解决了传统方法的资源耗比高问题,提升了数据运维的计算容量。但是由于技术限制,在运维容量上仍有提升空间。为此,需要通过提升云端处理器的计算性能,从而提升电力信息通信数据智能运维容量。

#### 参 考 文 献:

- [1] 孙保华,陈蕾,夏栋,等. 基于大数据平台的配电网智能化运维管控平台设计及应用[J]. 电气自动化, 2018, 40(6): 81-84.
- [2] 孙梦晨,丛伟,余江,等. 电网运维大数据背景下的继电保护通信系统故障定位方法[J]. 电力自动化设备, 2019, 39(4): 141-147.
- [3] 曾瑛,林斌,张正峰. 电力通信系统中运维数据采集技术的研究与分析[J]. 数据挖掘, 2018, 8(3): 112-120.
- [4] 王亚静,任瑞杰,张建臣,等. 基于场景的通信资源流程化管控机制分析与应用[J]. 电力信息与通信技术, 2018, 16(5): 68-73.
- [5] 周正钦,杜振波,王文瑞,等. 基于分层分布的变电站带电运维智能化技术及应用[J]. 电力系统保护与控制, 2019, 47(1): 150-157.
- [6] 张志强,黄佳胤,罗琨. 电网智能运维管理系统[J]. 云南电力技术, 2018, 46(4): 19-22.

- [7] 宋浒,杨波,郝悍勇,等. 基于流量预测的电力通信网流量均衡管理系统[J]. 电力信息与通信技术, 2018, 16(11): 53-57.
- [8] 陈远军,周哲,赵浩标. 电力变电站直流系统运行维护研究[J]. 电子设计工程, 2019, 27(20): 101-105.
- [9] 党芳芳,梅林,姬发家,等. 基于云计算的电网企业信息运维模式研究[J]. 电力大数据, 2019, 22(4): 50-54.
- [10] 徐砚,谷鹏. 电力监控网大数据分析平台研究与设计[J]. 通信技术, 2018, 51(8): 1908-1913.
- [11] 张辉,金燊,段寒硕,等. 基于动态推送时间优先级的电力通信现场运维信息推送方法[J]. 计算技术与自动化, 2019, 38(2): 58-63.
- [12] 林密,于祝芳,陈龙,等. 考虑通信影响的电力系统运维可穿戴设备可靠性评估方法[J]. 电子设计工程, 2020, 28(9): 107-111.
- [13] 蒋海明,夏丹妮,刘媛,等. 基于物联网技术的电力通信接入网的移动运维系统[J]. 电信科学, 2018, 34(10): 170-180.
- [14] 崔玉,吴奕,张志,等. 基于电力无线虚拟专网的继电保护智能移动运维系统设计及实现[J]. 电力系统保护与控制, 2018, 46(23): 175-181.
- [15] 荣雪琴,刘勇,刘昊,等. 基于时序数据库的电力运维系统关键技术研究[J]. 电测与仪表, 2018, 55(9): 48-51 83.

(责任编辑:丁一元)

(上接第 135 页)

- [9] 丁昭巧. 多 Agent 策略下的农业电子商务个性化推荐模型和接纳度研究[J]. 江苏农业科学, 2018, 46(1): 305-309.
- [10] 张付志,孙双侠,伊华伟. 基于非线性特征和 Cauchy 加权 M-估计量的鲁棒推荐算法[J]. 计算机学报, 2017(6): 1453-1469.
- [11] 何明,杨芃,要凯升,等. TEFRCF: 标签熵特征表示的协同过滤个性化推荐算法[J]. 计算机科学, 2018,

45(S1): 465-470 486.

- [12] 刘凯,张立民,周立军. 深度学习在信息推荐系统的应用综述[J]. 小型微型计算机系统, 2019, 40(4): 738-743.
- [13] 字云飞,李业丽,孙华艳. 基于深度神经网络的个性化推荐系统研究[J]. 电子技术应用, 2019, 45(1): 14-18 22.

(责任编辑:丁玥)