

基于个性化导购的商品智能动态推荐系统

An Intelligent Dynamic Recommendation System for Commodity Based on Personalized Guide System

姚剑^① YAO Jian; 余炎^② YU Yan; 黄诗盛^③ HUANG Shi-sheng; 赵培^① ZHAO Pei

(^①上海遥薇(集团)有限公司, 上海 201802; ^②上海遥薇智能科技工程有限公司, 上海 201802;

^③上海旭微物联网科技有限公司, 上海 201802)

(^①Shanghai Yaowei (Group) Co., Ltd., Shanghai 201802, China; ^②Shanghai Yaowei Intelligent Science and Technology Engineering Co., Ltd., Shanghai 201802, China; ^③Shanghai Xuwei Internet of Things Technology Co., Ltd., Shanghai 201802, China)

摘要: 为了节省顾客商品选购时间, 帮助顾客从海量商品中快速找寻意购物品, 本文尝试将关联规则与个性化导购相结合, 引入 FP-tree 关联规则算法, 利用商品推荐公式中的关联度, 同时综合利用商家利益最大化、热销商品等因素, 建立商品推荐综合评分公式。基于综合评分值对待推荐商品进行动态排序, 并将排序结果推送给客户, 实现个性化商品智能动态推荐服务。

Abstract: In order to save shopping time and to help customers find the favorite goods conveniently from numerous merchandises, this paper attempted to develop an intelligent recommendation system for individual commodity based on association rules. We established the comprehensive score formula of commodity selection by introducing the FP-tree association rules, using the correlation index of the recommended formula and synthesizing some factors such as the maximum of benefits, hot commodity, etc.. The merchandise recommendation service is based on the rank of comprehensive score, the sorted results are finally pushed to the customer.

关键词: 个性化商品导购; 数据挖掘; 关联规则; 智能推荐系统

Key words: guiding-purchase of individual commodity; data mining; association rules; intelligent recommendation system

中图分类号: TP311.1

文献标识码: A

文章编号: 1006-4311(2017)35-0199-03

DOI: 10.14018/j.cnki.cn13-1085/n.2017.35.086

0 引言

近年来, 随着信息技术的飞速发展。网上购物方式在人们的生活中比重已经越来越大。商品推荐系统主要是基于网站的用户购买记录与浏览行为, 通过算法对商品库进行匹配, 将当前热门的、用户收藏的或者风格类似的商品推荐给潜在购买顾客^[1]。

个性化智能推荐最终目标就是让任何一个访问电商平台的用户, 在进入平台页面时, 系统能够根据用户日常的行为偏好和习惯, 提前知道用户意向购买商品, 在还没有发生点击行为时, 系统能自动将特定商品推荐到用户访问的页面, 提升平台用户下单转化率。即使在用户没有访问平台时, 商家通过与用户日常浏览互联网行为轨迹的平台进行联盟合作, 在联盟平台推送用户希望购买的商品广告和商品链接, 刺激和引导用户点击购买。即使在用户没有打开电脑时, 能够通过信息和邮件的方式, 根据用户平常的购买频次和周期, 在特定的时间推送到用户手机和电脑^[2]。

1 工作流程及具体实现

1.1 推荐系统的工作流程

商品动态推荐系统的整个工作流程包括以下四个步

骤: 商品数据采集、商品数据预处理、商品推荐综合评分、动态智能推荐^[3]。

1.2 商品数据预处理

商品数据预处理主要是数据清洗与数据变换, 通过数据预处理后得到的商家主推修正系数、商品热销度、商品毛利率修正系数、商品关联度。数据预处理的具体步骤如下:

Step1: 商品热销度评分是根据商家最近一段时间(例如 1 个月内)商品销售记录, 汇总后排名经规范化后的计算得分。最高取值为 1, 最低取值为 0。

$$\gamma_1 = \frac{Q_i - Q_{\min}}{Q_{\max} - Q_{\min}} \quad (1)$$

式(1)中 γ_1 为商品热销度, 取值范围在 0-1 之间; Q_i 为商品销售量; Q_{\max} 为该商家最近 1 个月内有销售记录的商品最大销售量; Q_{\min} 为该商家最近 1 个月内有销售记录的商品最小销售量。

Step2: 商家主推系数 γ_2 是针对新商品做出排序得分, 最高为 1, 最低为 0.1, 系数越高表示商户越期望将此商品推荐给顾客。

Step3: 商品毛利率修正系数是根据商品单价和成本计算得到毛利率, 再经归一化后得到。

$$\gamma_3 = \frac{\rho_{\text{单价}} - \rho_{\text{成本}}}{\rho_{\text{单价}}} \quad (2)$$

式(2)中 γ_3 为特定商品的毛利率修正系数; $\rho_{\text{单价}}$ 为商品销售单价; $\rho_{\text{成本}}$ 为商品采购成本。

Step4: 商品关联度修正系数 γ_4 是根据历史订购数据, 采用 FP 关联规则算法得到商品关联度推荐评分值, 再经过归一化, 反映商品与顾客已采购商品的关联程度。

关联规则是形如 $X \rightarrow Y$ 的函数(X 项集和 Y 项集不相交)。本文引入支持度与置信度来衡量关联规则强度。支持度(support)用于确定规则可以对给定数据集的频繁程度; 置信度(confidence)用于确定项集 Y 在包含项集 X 的事务

基金项目: 上海市 2015 年度“科技创新行动计划”高新技术领域项目(15511102100) 本研究受上海市 2015 年度“科技创新行动计划”高新技术领域项目——基于大数据分析技术的社区便民服务平台和示范应用(项目编号: 15511102100)资助, 同时, 研究成果在上海市嘉定区智慧社区管理平台已经得到了实际应用。

作者简介: 姚剑(1979-), 男, 上海人, 本科, 研究方向为智慧社区与物联网; 余炎(1979-), 男, 上海人, 本科, 研究方向为智慧城市; 黄诗盛(1981-), 男, 上海人, 本科, 研究方向为计算机软件; 赵培(1961-), 男, 上海人, 本科, 研究方向为智慧学校。

中出现的频繁程度^[4]。

支持度(s)形式定义:

$$s(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{N} \quad (3)$$

置信度(c)形式定义:

$$c(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)} \quad (4)$$

支持度较低的规则,表示发现频率较低,可能是偶然发生。可以通过确定阈值来删除这些支持度较低的规则。本文的商品智能推荐系统采用 FP 增长算法,根据最小支持度找出频繁项集,再由频繁项集产生满足一定支持度和置信度条件的强关联规则。

1.3 动态推荐系统的具体实现

1.3.1 基于关联规则的推荐算法

经过数据预处理之后,基于关联规则的商品智能推荐算法,可以描述如下:

引进 FP-tree 算法后,该算法只需要进行两次商品数据库扫描,直接压缩数据库便可形成一个频繁模式树,同时生成相应的关联规则^[5]。

算法关键步骤:第一步利用商品数据库构造 FP-tree;第 2 步从 FP-tree 中挖掘频繁模式。

①构建 FP-tree;

②FP 增长算法产生频繁项集;

③产生强关联规则。

1.3.2 商品推荐综合评分

本文所述的商品推荐系统是对商品关联系数、商家主推、商品毛利率、商品热销度的综合考量,特定商品推荐综合评分计算公式如下:

$$S_{\text{综合评分}} = (E - Y)AY^T \quad (4)$$

$$\text{其中 } E = (1, 1, 1, 1), Y = (\gamma_1, \gamma_2, \gamma_3, \gamma_4), A = \begin{bmatrix} 0 & \alpha_2 & \alpha_3 & \alpha_4 \\ \alpha_1 & 0 & \alpha_3 & \alpha_4 \\ \alpha_1 & \alpha_2 & 0 & \alpha_4 \\ \alpha_1 & \alpha_2 & \alpha_3 & 0 \end{bmatrix}$$

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 10.$$

$S_{\text{综合评分}}$ 是特定商品推荐综合评分,取值范围 0~40; γ_1 为特定商品热销度的评分,取值范围 0~1, α_1 为其权值; γ_2 商家对特定商品的主推修正系数,取值范围 0.1~1.0, α_2 为其权值; γ_3 是特定商品毛利率修正系数,取值范围 0.1~1.0, α_3 为其权值; γ_4 是特定商品与已选商品关联度修正系数,取值范围 0.1~1.0, α_4 为其权值。(注:4 种评分因素分配的权值不同,则商品推荐综合评分值也会不同)

传统综合评分公式只能对每个指标加权求和得到综合评分,而该新公式在此基础上加入后缀影响因式,可以自行根据指标数值的情况提升重要指标(占权值较大的指标)排序时的优势地位,即权值大的指标值越大,在总体中排序更具优势。但综合评分排序并不完全按照权值排序,还会考虑其他指标。

2 智能推荐系统的应用实例

为验证推荐综合评分算法有效性,将此算法与一般综合评分算法进行验证和比较。

某顾客在某店铺已选购了智能手机,该店铺收集相关商品数据如表 1,包含 12 个商品组和 20 个商品。

表 1 菜品数据集

序号	商品集
1	智能手机,智能手环,玩具机器人,平板电脑,蓝牙音箱,充电宝
2	智能手机,旅行箱,自拍杆,无线鼠标,VR 眼镜
3	空气净化器,智能手环,旅行箱
4	智能手机,智能手环,平板电脑,旅行箱,路由器,无人机
5	智能手机,玩具机器人,平板电脑,行车记录仪,血压计
6	智能手机,智能手环,旅行箱,自拍杆,体重秤
7	智能手机,平板电脑,净水器,电饭煲,扫地机器人
8	自拍杆,折叠自行车
9	智能手机,智能手环,玩具机器人,平板电脑,自拍杆,路由器
10	智能手机,智能手环,平板电脑,自拍杆,充电宝
11	智能手机,智能手环,平板电脑,旅行箱,路由器
12	智能手机,智能手环,玩具机器人,平板电脑,旅行箱,摄像机

通过扫描表 1 商品数据集,统计得到每个商品的支持度计数如表 2 所示。

表 2 商品支持度计数

商品	计数
智能手机	10
平板电脑	8
智能手环	7
旅行箱	5
玩具机器人	4
自拍杆	4
路由器	3
充电宝	2
无线鼠标	1
VR 眼镜	1
蓝牙音箱	1
无人机	1
行车记录仪	1
血压计	1
体重秤	1
净水器	1
电饭煲	1
扫地机器人	1
折叠自行车	1
摄像机	1

设最小支持度为 2,丢弃非频繁项,得到频繁项,按照支持度递减的排序为:智能手机、平板电脑、智能手环、旅行箱、玩具机器人、自拍杆、路由器、充电宝。

接着按照上述算法构建 FP-tree 和 FP 增长算法,产生以智能手机结尾的频繁项集,并将其中归一化后达到最小置信度 0.40 的具有强关联规则的项集输出:{平板电脑、智能手机}、{智能手环、智能手机}、{旅行箱、智能手机}、{玩具机器人、智能手机}、{自拍杆、智能手机},它们的置信度分别为 0.80 0.70 0.50 0.40 0.40。

整理与智能手机具有强关联规则的 5 个商品的销售量、单价、成本的数据如表 3 示。

表 3 5 个商品的 3 种原始数据表

商品名称	月销售量/份	单价/元	成本/元	备注
平板电脑	116	16	9.6	最热销的商品 销售量=188; 最冷销的商品 销售量=9
智能手环	152	5	2.0	
旅行箱	143	20	10	
自拍杆	134	15	5.7	
玩具机器人	161	15	6.0	

根据原始数据及式(1)~式(4),计算出热销度评分、毛

浅谈高职院校优质校建设

On the Construction of High Quality Schools in Higher Vocational Colleges

杨洁 YANG Jie

(浙江经贸职业技术学院教务处 杭州 310018)

(Zhejiang Institute of Economics and Trade Academic Affairs Office Hangzhou 310018 China)

摘要:建设优质高职院校是示范高职院校建设的深化与拓展,也是高职院校加强内涵建设的必然要求,根本上是为了满足人们日益增长的对优质高职教育的需求。通过强化内涵建设,加强“双师型”队伍建设,推进现代学徒制改革试点,重视“诊改”工作等,促进高职院校优质校建设。

Abstract: The construction of high quality vocational colleges is to deepen and expand the construction of demonstration vocational colleges, and is also the inevitable request of strengthening the connotation construction of higher vocational colleges, which is to basically meet the growing demand for quality of higher vocational education. By strengthening the connotation construction, strengthening the construction of "double qualified" team, promoting the pilot reform of modern apprenticeship, attaching importance to the "diagnosis and reform" work, the construction of high quality schools in higher vocational colleges is promoted.

关键词: 高职院校; 优质校; 建设

Key words: higher vocational college; high quality school; construction

中图分类号: G718

文献标识码: A

文章编号: 1006-4311(2017)35-0201-02

DOI:10.14018/j.cnki.cn13-1085/n.2017.35.087

0 引言

近年来,随着我国高职教育从外延式扩展向内涵质量提升的转变,以及国家示范(骨干)高职院校的建设要求,高质量逐渐成为新一轮高职院校建设的奋斗目标。《高等职业教育创新发展行动计划(2015-2018)》(教职成[2015]9号)已明确提出,将要建成200所左右的优质专科高等职业院校。所谓优质高职院校,就是不断“优质化”的高职院校,主要体现在办学理念先进、办学定位准确、产教深度

作者简介: 杨洁(1968-),女,浙江临海人,研究方向为高等教育管理。

融合、人才培养质量高、社会服务能力强以及社会认可度高等方面。

1 优质校建设背景

1.1 我国已经建成世界上最大规模的职业教育体系。据统计,截止2016年,全国共有职业院校1.23万所,在校生2680.21万人,职业院校共开设近千个专业,近10万个专业点,基本覆盖国民经济各领域,是教育领域名副其实的“半壁江山”。职业教育虽然体量庞大、发展速度惊人,然而在整个社会领域的认可度仍然很低,面临着如何做大做强的挑战。职业教育只有走内涵式发展的道路,才能从

3 结束语

本文主要设计一种基于关联规则综合评分算法的销售服务智能推荐系统。根据已购商品关联度、商家利益最大化、热销商品等因素,以及因素重要度来创建商品推荐综合公式。基于综合评分对各商品进行推荐排序,排序结果最终以移动方式推送给顾客。商品动态推荐服务,能够帮助顾客快速发现感兴趣的商品,同时有助于商家发现商品间关联关系,并采取措施,同时提高客户满意度与营业利润,从而实现消费者和商家的双赢。

参考文献:

- [1] 张素智, 赵亚楠, 杨芮. 推荐系统研究[J]. 湖北民族学院学报(自然科学版), 2017(01): 1-6.
- [2] 丁雪. 基于数据挖掘的图书智能推荐系统研究[J]. 情报理论与实践, 2010(05): 107-110.
- [3] 任明枢. Web数据挖掘及其在电子商务中的应用[D]. 山东科技大学, 2004.
- [4] 胡迎松, 韩苹, 陈中新. 一个基于Agent的个性化推荐系统[J]. 计算机应用研究, 2006(04): 78-80.
- [5] 李煌, 汪晓岩, 庄镇泉. 基于关联规则挖掘的个性化智能推荐服务[J]. 计算机工程与应用, 2002(11): 200-204, 229.

表4 商品与其4种数据变量

商品名称	热销度评分	毛利率	主推度	关联度
平板电脑	0.60	0.40	0.80	0.80
智能手环	0.80	0.40	0.50	0.70
旅行箱	0.75	0.50	0.75	0.50
自拍杆	0.70	0.62	0.90	0.40
玩具机器人	0.85	0.60	0.70	0.40

根据经验设定权值,由传统综合评分公式和本文综合评分公式分别计算得到各商品的综合分数如表5。

表5 评分方法对比

评分方式	商品名称				
	智能手环	平板电脑	自拍杆	旅行箱	玩具机器人
传统评分	6.000	7.850	6.000	5.875	5.775
本文评分	7.420	7.460	6.236	6.531	6.203

对比分析两种方法,由传统方法计算出的智能手环和自拍杆的分数都是6,故当顾客选购了智能手机和行车记录仪,这两类商品处于同等推荐位置,但显然是不合理的,因为智能手环与智能手机和行车记录仪关联度是0.7,远远大于自拍杆与智能手机和行车记录仪的关联度0.4,智能手环应比自拍杆更具有推荐优势位置,且智能手环和自拍杆的分数差为1.184,也在合理范围内,不会太悬殊。

自拍杆和玩具机器人的关联度都是0.4,自拍杆热销