

基于关联规则的电商产品组合销售热点研究

刘玉林

(安徽商贸职业技术学院经济贸易系,安徽芜湖 241002)

摘要:在对电商中产品组合销售的商业理解和数据理解基础上,选择三只松鼠碧根果礼盒为例,运用关联规则等数据挖掘方法,找出子产品组合热点,给出定性分析和组合建议等后处理。结果表明关联规则在电商产品组合销售热点中有着非常良好的分析效果。

关键词: 电商;产品组合销售;关联规则;热点

中图分类号: F272 **文献标志码:** A **文章编号:** 1673-3851 (2016) 06-0528-04 **引用页码:** 120103

产品组合销售,又被称为搭配销售,通常指将相关产品搭配在一起成套销售的方式。对于电商而言,组合销售的产品必须要是消费者需要并且愿意购买的产品,从而突出组合销售的优点,给顾客提供方便,并且提升产品销售量^[1]。

电商商家在考虑组合销售产品时,可以从目前现有市场上产品组合销售的特点分析,找到组合热点,从而作为决策的参考意见。因此,本文以市场上某款产品销售为例,基于关联规则技术分析组合热点^[2],从而为相关的研究和决策提供一些建议和思路。

一、研究流程设计

对产品组合销售热点的研究,符合数据挖掘的基本流程,包括商业理解、数据理解、数据准备、建模和后处理^[3],具体见图1。

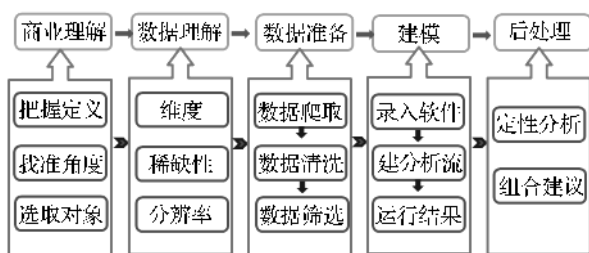


图1 产品组合热点数据挖掘流程设计

图1流程中,“准确把握产品组合销售的内涵,找准分析的角度,合理选择研究对象”是商业理解的核心;“合理确立数据的维度、对稀缺性的筛选、适度设置分辨率”是数据理解的关键;在数据准备阶段,利用爬虫软件爬取网络数据,对数据噪音和异常值清洗,做好数据筛选等初处理工作;在建模中利用关联规则找出频繁项集,将数据处理成布尔数据导入SPSS Modeler 软件,运行 Apriori 模块得出结果;最后进行定性分析和产品组合销售建议等后处理工作^[4]。

二、实证分析

本文实证分析选择电商“三只松鼠”品牌碧根果礼盒为例,分析该企业如何通过市场上碧根果礼盒产品组合热点来设计自有品牌礼盒中子产品组合,以此说明基于关联规则的数据挖掘在电商产品组合销售热点中的运用。

(一)商业理解

安徽三只松鼠电子商务有限公司“三只松鼠”品牌创立于2012年,短短三年发展销量突破20亿元,其品牌价值估值40亿,创造了电子商务干果行业发展奇迹。

在商业理解阶段,三只松鼠旗舰店在平时、节假日销售中往往采用产品组合销售模式进行,即以“礼

盒、套装”等形式提供产品。因为碧根果是其主打产品,同时在干果市场上占有较大份额,因此碧根果礼盒装往往成为三只松鼠旗舰店销售主流。在与其管理人员讨论中,如何组合产品才能更受市场欢迎,起到消费者满意和提升销售的作用,成为管理者首要面对的问题。当然,通过消费者问卷调查能够帮助企业完成一些市场探测性结论,或者产品组合实验也能够起到效果。在这里,本文以电子商务中所有网商的碧根果礼盒装产品为分析对象,利用数据挖掘中关联规则分析产品组合的频繁项集,以此来说明目前产品组合销售的热点现状,其结论可以作为三只松鼠旗舰店管理层决策的重要参考。

(二)热点分析数据理解

在淘宝网中宝贝栏目以“碧根果礼盒”为关键词进行搜索,在搜索页面中“合并同款宝贝”。结果显

示淘宝网(含天猫店)中所有提供碧根果礼盒的商家产品,产品信息包含名称、价格、付款人数、商家地址;详细页面中还提供商品详情,包含保质期、净含量、配料表等。

在数据理解阶段,首先要确定研究对象的维度,产品组合销售热点首要分析产品组合中的配料信息,维度选择产品组合名称、配料表;其次稀缺性中每款碧根果礼盒中都有对配料表的详细描述,能够保证产品组合销售热点分析;最后,分辨率中对配料表中以具体种类区别,去掉口味或者命名差异,例如“紫薯花生”和“多味花生”都视为花生,“山楂片”和“山楂球”都视为山楂。

(三)热点分析数据准备

对淘宝网(含天猫店)所有搜索到的碧根果礼盒进行软件爬取,共获得 538 款产品组合销售数据(同款已被合并)。该数据描述性分析见表 1。

表 1 碧根果礼盒爬取数据总体情况

总数据量	天猫店数据量	淘宝店数据量	配料表中子产品区间/种	价格区间/元	月销量区间/件
538	109	429	2~12	41.9~486	0~19631

表 1 显示其中天猫店共有 109 款、淘宝店共有 429 款组合销售产品;产品组合中最少有 2 种子产品,最多有 12 种子产品;价格最低 41.9 元,最高 486 元;销量最高为 19631 件/月,最低为 0 件/月。

将 538 款产品组合中产品名称和配料表进行提取,对配料表进行分列,并按照数据理解中的分辨率

进行归并,定义缺失值数据^[5],将产品名称以数字序号代替等,完成数据清洗工作。清洗完成的数据形式见表 2。通过表 2 发现虽然商家将产品命名为碧根果礼盒,但有些产品中并没有碧根果子产品,因此将该款产品组合从分析中删去,最后数据量为 471 款。

表 2 产品组合数据清洗后初步形式(部分)

序号	子产品								
	1	2	3	4	5	6	7	8	...
1	开心果	腰果	碧根果	榛子	扁桃仁	葡萄干	杏仁	(空)	...
2	猪肉脯	夏威夷果	炒米	炒米	碧根果	芒果干	开心果	花生	...
3	夏威夷果	碧根果	巴旦木	话梅	瓜子	腰果	花生	瓜子	...
...

将表 2 中子产品归并成 1 列后通过 excel 数据 透视图得出子产品种类表 3。

表 3 子产品次数表

子产品	次数	子产品	次数	子产品	次数	子产品	次数
碧根果	471	兰花豆	18	茶	4	枸杞	1
夏威夷果	406	棒子	15	甘草	4	怪味豆	1
巴旦木	249	豆瓣	15	甘栗	4	桂皮	1
腰果	205	香蕉干	15	曲奇	4	果蔬干	1
开心果	163	玉米片	14	香榧	4	黑豆	1
瓜子	149	芒果干	12	瓜蒌子	3	胡桃	1

表 3 续

子产品	次数	子产品	次数	子产品	次数	子产品	次数
花生	139	豌豆	11	果仁	3	烤椰片	1
杏仁	108	猪肉脯	10	菠萝蜜干	2	蓝莓	1
核桃	92	蔓越莓干	9	黄金豆	2	荔枝干	1
葡萄干	84	牛肉干	8	可可	2	木瓜干	1
蚕豆	79	西梅	7	鳕鱼丝	2	柠檬片	1
松子	78	杨梅	7	奶粉	2	牛轧糖	1
枣	78	黄焦干	6	葡萄糖	2	肉松饼	1
山核桃	63	煎豆	6	巧克力	2	素肉	1
青豆	58	葵花籽	6	圣女果干	2	椰蓉	1
话梅	43	无花果干	6	笋干	2	鱼豆腐	1
榛子	43	鲍鱼果	5	鱿鱼	2	鱼皮	1
炒米	41	锅巴	5	草莓干	1	子弹肠	1
桂圆	39	牛肉	5	蛋皮	1		
山楂	37	板栗	4	凤梨干	1		

在关联规则中对次数出现 1 次及其以下的,无法分析频繁项集,因此筛选除去子弹肠、鱼皮等子产品。在筛选结果上利用 excel 中的 if 函数构建布尔型数据表 4^[6]。至此,完成数据准备工作。

表 4 子产品布尔型数据表(部分)

巴旦木	板栗	榛子	鲍鱼果	菠萝蜜干	蚕豆	茶	炒米	...
0	0	0	0	0	1	0	1	...
0	0	0	0	0	1	0	1	...
1	0	0	0	0	0	0	1	...
1	0	0	0	0	1	0	1	...
1	0	0	0	0	1	0	1	...
1	0	0	0	0	1	0	0	...
1	0	0	0	0	1	0	0	...
0	0	0	0	0	1	0	0	...
0	0	1	0	0	0	0	0	...
0	0	0	0	0	0	0	1	...
...

(四)热点分析数据 SPSS Modeler 建模

在 SPSS Modeler 中建立“流”分析,“流”设计见图 2^[7]。其中源数据导入表 4 布尔型数据表,类型中测量选择“标志”和角色选择“两者”。在 Apriori 规则中设置支持度阈值为 10%,置信度阈值为 80%。

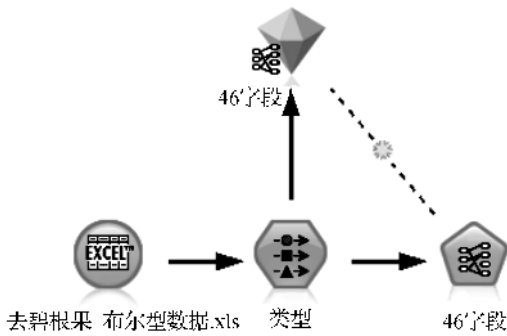


图 2 SPSS Modeler 中关联规则流分析

运行软件得到结果见表 5。

表 5 SPSS Modeler 中 Apriori 规则运行结果

后项	前项	支持度 / %	置信度 / %	规则支持 / %
夏威夷果	腰果	39.92	87.23	34.82
夏威夷果	花生	25.90	80.33	20.81
夏威夷果	腰果 and 巴旦木	18.90	89.89	16.99
夏威夷果	蚕豆	15.92	84.00	13.38
夏威夷果	松子	15.07	84.51	12.74
夏威夷果	蚕豆 and 花生	11.46	85.19	9.77
夏威夷果	花生 and 巴旦木	11.25	84.91	9.55
夏威夷果	开心果 and 腰果	11.04	82.69	9.13
夏威夷果	青豆	10.61	82.00	8.70
夏威夷果	花生 and 腰果	10.19	87.50	8.92

三、产品组合销售热点建议

在热点分析数据关联规则结果上通过定性分析,解读目前电子商务中碧根果礼盒产品组合销售热点,并以此给出网上组合中子产品建议。

(一)定性分析

表5显示,夏威夷果是网商碧根果礼盒搭配中最喜爱的子产品,所有的频繁项目都指向夏威夷果,这对于网商做出产品组合销售策略是非常有意义的。

同时,结果中支持度最高的为“腰果” \Rightarrow “夏威夷果”,支持度超过39.92%,是目前碧根果产品组合销售中关联规则最强的,可见目前电子商务中网商倾向于搭配“腰果—夏威夷果”;在频繁项集中“花生” \Rightarrow “夏威夷果”也是网商喜爱的组合,其他包括巴旦木、开心果、蚕豆、松子、青豆等也比较流行。当然,结果显示“腰果、巴旦木” \Rightarrow “夏威夷果”支持度也较高,也是网商采用的组合热点。同时将“蚕豆、花生”、“花生、巴旦木”、“开心果、腰果”、“花生、腰果”等与“夏威夷果”搭配也是目前碧根果礼盒中的热度较高的考虑^[8]。

(二)组合建议

经过子产品关联规则的挖掘,在三只松鼠考虑销售碧根果礼盒时,有以下建议参考^[9]。

1.“碧根果、腰果、夏威夷果”三果组合策略

从结果可以看出,“腰果” \Rightarrow “夏威夷果”支持度超过39.92%,碧根果搭配腰果之后,网商再次搭配夏威夷果的比例超过87.23%,可见网商对于“碧根果、腰果、夏威夷果”组合的看好。因此,建议三只松鼠销售碧根果礼盒首要考虑为“碧根果、腰果、夏威夷果”的三果组合。

2.“碧根果、花生、夏威夷果”多品类组合策略

在现有碧根果礼盒中“花生” \Rightarrow “夏威夷果”支持度达到25.9%(超过1/4),即如果碧根果搭配花生,再搭配夏威夷果是目前1/4网商的选择。同时碧根果搭配花生,再次搭配夏威夷果的置信度为80.33%,比例较高。该组合中花生是消费者喜爱并且是非果类的零食产品,因此建议三只松鼠丰富碧根果礼盒系列,考虑推出“碧根果、花生、夏威夷果”多品类组合产品。

3.“开心果、巴旦木”等多样性组合策略

在结果中巴旦木、开心果、蚕豆、松子、青豆等也是网商喜爱组合的子产品热点^[10]。当然,有些是三项搭配,比如“碧根果、蚕豆、夏威夷果”;有的是四项搭配,比如“碧根果、开心果、腰果、夏威夷果”。因此,建议三只松鼠可以进一步增加碧根果礼盒品质,将上述支持度较高的搭配进行合理考虑。

四、结 语

本文选择三只松鼠碧根果礼盒为研究对象通过电商中碧根果礼盒产品信息爬取、数据准备、建模等方法,利用SPSS Modeler中Apriori规则找出电商市场中产品组合热点,进行定性分析和组合建议等后处理,最终分析产品组合销售热点的现状。事实证明,关联规则在挖掘子产品搭配组合、探讨产品组合销售方面有非常好的效果。

参考文献:

- [1] 唐克冰. 茶叶销售与节日庆典搭配的营销思路分析[J]. 福建茶叶, 2016, 38(1): 63-64.
- [2] 赵静. 基于市场特征的数字出版物销售商的产品组合策略[J]. 编辑之友, 2014(5): 65-67.
- [3] 吴晓英, 明均仁. 基于数据挖掘的大数据管理模型研究[J]. 情报科学, 2015, 33(11): 131-134.
- [4] 吉顺权, 周毅. 产品用户评论在企业竞争情报中的应用: 基于产品特征的关联规则数据挖掘[J]. 现代情报, 2015, 35(6): 114-121.
- [5] 花琳琳, 施念, 杨永利, 等. 不同缺失值处理方法对随机缺失数据处理效果的比较[J]. 郑州大学学报(医学版), 2012, 47(3): 315-318.
- [6] 何伟全, 王天志, 张广斌, 等. 基于关联规则挖掘的高校学生意外伤害的时间特征研究: 以云南部分高校为例[J]. 云南师范大学学报(哲学社会科学版), 2015, 47(4): 111-119.
- [7] 李刚, 潘跃. 基于SPSS Modeler的超售方法在京沪高铁票额管理中的应用研究[J]. 铁路计算机应用, 2014, 23(8): 33-35.
- [8] 周怡, 黄昌富. 信息经济时代的产品组合决策[J]. 统计与决策, 2007(9): 151-153.
- [9] 杜建耀. 企业产品组合的优化[J]. 经济管理, 2003(4): 71-74.
- [10] 马骋远, 白思俊, 郭云海. 不确定环境战略导向的复杂产品研发项目组合选择[J]. 企业改革与管理, 2016(3): 20-21.

Research on Influential Factors of Terminal Distribution Satisfaction of Online Customers for Fresh Products

XIE Linshan^a, HAN Shuguang^b

(a. School of Economics and Management; b. School of Sciences, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: This paper studies influential factors of terminal distribution satisfaction of online customers for fresh products. Based on the theory of service quality satisfaction, TAM model and the consumer research, this paper chooses reliability, service care, service guarantee, service usability and service response as the research indexes of influential factors of consumer satisfaction. The regression analysis of the data shows the reliability and service guarantee mostly influence consumer satisfaction for terminal distribution of fresh products. The result suggests that online fresh product operation can improve consumer satisfaction through enhancing distribution time accuracy and freshness.

Key words: terminal distribution; fresh products; consumer satisfaction; regression analysis

(责任编辑:陈和榜)

(上接第 531 页)

The Hot Research of E-Business Product Portfolio Sales Based on Association Rules

LIU Yulin

(Department of Economics & Trade, Anhui Business College, Wuhu 241002, China)

Abstract: The article selects Three Squirrels' Pecans gift boxes as the example based on the business understanding and data understanding of product portfolio sales in e-commerce, applies the data mining methods such as association rules to look for the hot product portfolio, and then gives qualitative analysis and combination suggestions. The study shows that association rules have a very good analysis effect on the hot research of product portfolio sales.

Key words: E-business; product portfolio sales; association rules; hot research

(责任编辑:陈和榜)