

DOI: 10.19333/j.mfkj.2017080320707

基于需求预测的服装 O2O 货品管理探析

李玉灵¹ 单丹微¹ 杨以雄^{1,2} 王晶晶¹

(1. 东华大学 服装与艺术设计学院, 上海 200051; 2. 东华大学 现代服装设计与技术教育部重点实验室, 上海 200051)

摘要: 针对服装企业 O2O(Online to Offline, 线上到线下或线下到线上) 渠道背景下, 货品管理中存在的需求预测问题进行探析。利用决策树算法和自动回归技术, 分别建立按 O2O 销售渠道和实体店铺地域划分的货品需求预测模型。根据案例企业 ERP 系统中线上及线下的实际历史销售数据, 由模型计算得到未来第 1 季度的预测值, 并与线上及线下的实际销售值进行对比分析, 验证模型的准确度。结合案例企业不同渠道实际配货补货需求及事例解析, 得出需求预测模型可帮助企业对货品进行合理分配, 减少库存成本及缺货风险。

关键词: 需求预测; 服装企业; O2O; 货品管理

中图分类号: TS 941.1 文献标志码: A

Study on demand forecasting in apparel O2O products management

LI Yuling¹, SHAN Danwei¹, YANG Yixiong^{1,2}, WANG Jingjing¹

(1. College of Fashion and Design, Donghua University, Shanghai 200051, China; 2. Key Laboratory of Clothing Design and Technology, Ministry of Education, Donghua University, Shanghai 200051, China)

Abstract: The problem of demand forecasting in the management of product under the background of the clothing enterprise O2O channels was studied in this paper. According to the O2O sales channels and the location of the store, the products demand forecasting models were established by using decision tree algorithm and auto regression techniques. On the basis of the actual sales data online and offline in enterprise ERP system, the predicted values of the first quarter were calculated by the model. Compared to the actual sales values online and offline, the accuracy of the model was verified. With the actual distribution end replenishment demand of different channels, the model of products demand forecasting can help enterprises to make reasonable allocation of goods, and to reduce inventory cost and shortage risk.

Keywords: demand forecasting; apparel enterprises; Online to Offline; products management

服装企业货品管理包括订货、配货、补货、货品推广与分析、库存处理等, 前期的货品预算直接影响企业的库存控制, 而货品需求的准确预测将有助于企业对货品管理进行有效决策^[1]。近年来, 随着电商经营渠道的加入, 传统线下实体店铺的货品管理方式已不再适用于服装企业 O2O(Online to Offline, 线上到线下或线下到线上) 渠道运营。线上渠道的

加入使得货品需求难以准确把握, 双渠道冲突导致线上及线下货品资源分配难度增加、货品供应无法及时跟上消费需求变化等问题^[2]。如何对 O2O 渠道模式下的货品需求做出准确预测, 实现精准营销成为服装企业货品管理的重要经营课题。

目前, 国内外针对服装货品需求预测的研究, 主要是围绕单一的线上或者线下实体店铺的货品销售预测, 而针对 O2O 双渠道货品需求预测方案的研究则尚处于起步阶段^[3-5]。由于服装货品需求受渠道、地域、气候等多重因素影响, 其销售数据往往呈现复杂的非线性特征^[6]。为此, 迄今货品需求预测的研究, 比较有效的方法是利用决策树算法和自动回归技术建立预测模型, 通过误差分析不断改进预测方法, 以此提高预测的精准度^[7]。本文结合案例

收稿日期: 2017-08-31

基金项目: 上海高校知识服务平台项目(13S107024); 东华大学非线性科学研究所专项资金(INS-1401); 同济大学上海国际创新研究院项目(DB17016)

第一作者简介: 李玉灵, 硕士生, 研究方向为服装产业经济研究。通信作者: 杨以雄, E-mail: yyx@dhu.edu.cn。

企业实际经营特点,基于对数据挖掘技术中的决策树理论及自动回归技术,探索并建立适用于O2O渠道的货品需求预测模型,并根据预测数据为案例企业配货补货提供决策参考。

1 需求预测模型的选择

本文研究的案例对象为L品牌服装企业(以下简称L企业),是一家海外品牌服装集团在国内设立的本土化全资销售子公司,主要负责中国区的商品企划、货品采购、零售运营管理和市场拓展。O2O经营模式为企业直接掌控线上线下渠道的货品管理,运营方式为订货制,无生产补货,且共享库存,即线上货品存储在各地区门店中,线下门店同时满足2种(线上、线下)渠道的货品需求。货品需求预测将以此为对象进行模型选择与建立。

1.1 模型的选择

在数据挖掘技术中,时间序列是重要研究方向之一,其预测方法是基于大量时间序列数据,对未来走势做出有效客观分析^[8],从而实现对销售数据季节性、周期性和趋势性的有效模拟,以提高预测精度^[9]。决策树算法是一种典型的归纳分类算法^[10],以实际数据为基础,生成结果为树结构,树的节点表示属性测试条件,树的分枝表示测试输出^[11],其特点为速度快,计算量小,数据挖掘准确性高,且便于理解^[12]。

据此,本文选择决策树算法,结合处理时间序列的技术——自动回归技术,建立自动回归树(Auto-Regression Tree, ART),利用SQL Server 2012数据库系统实现数据挖掘预测模型^[13]。该模型优势表现为:①能够处理大量数据;②可处理非线性关系数据;③可对周期性数据建模;④分段式的线性预测使得结果便于理解;⑤容易学习且费用低^[14]。

1.2 模型构成

服装货品需求预测以大量历史销售数据为基础,目标是尽可能准确地预测店铺未来销售量和销售额,故可将预测需求分为以下2种组合:

① 销售渠道:对同品类产品的线上渠道和线下渠道的销售量(金额)进行预测。

② 店铺地域:对国内某一地区或某家店铺的销售量(金额)进行预测。

根据服装企业O2O渠道经营特点,结合ART模型,得到构建模型的步骤方法。货品需求预测模型基本步骤见图1。其中,源数据主要由企业ERP系统(Enterprise Resource Planning,企业资源计划)获取,对象产品以L企业C品牌连衣裙品类为例,对ART模型的构建与验证过程进行解析。

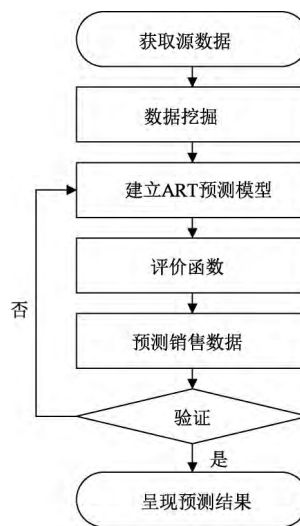


图1 货品需求预测模型基本步骤

2 货品需求预测模型的构建与验证

通过SQL Server系统建立数据库,利用SSAS构建自动回归树(ART)模型,主要分为建立模型、验证模型、部署和更新模型3个部分^[13]。

2.1 建立模型

2.1.1 定义问题

截至2016年12月,C品牌在国内共有69家线下实体店铺(包含品牌集合店铺),分布在东北、华北、华东、华南、华中、西北及西南等7大区域。在此,以C品牌连衣裙(A品类)2015年1月至2016年12月期间(共106周)的每周销售数据为基础,按销售渠道和店铺地域2种分类划分方式,建立货品需求预测模型,对销售金额进行预测。

2.1.2 准备数据

从企业ERP系统提取C品牌连衣裙A品类2015年1月至2016年12月的历史销售周数据,并按2种销售渠道和7个店铺地域进行数据划分。鉴于国内店铺数量每年均有一定增减,并且货品数量根据店铺等级不同而有所区别,故对系统内相关数据统一作均值处理。A品类按渠道划分的销售数据见表1。A品类按地域划分的销售数据见表2。

表1 A品类按渠道划分的销售数据(均值,部分)
元/(店·周)

渠道	周次					
	1	2	3	4	...	106
线下渠道	1 889.28	2 749.57	344.35	1 133.33	...	3 388.99
线上渠道	2.96	99.91	45.88	41.36	...	142.32
合计	1 892.24	2 849.48	390.23	1 174.69	...	3 531.31

注: C品牌销售数据二次整理。

表 2 A 品类按地域划分的销售数据(均值 部分)

地区	元/(店·周)					
	周次					
	1	2	3	4	...	106
东北	1 416.00	2 120.00	216.00	1 120.00	...	2 300.00
华北	2 273.33	3 275.32	562.08	1 115.56	...	2 980.00
华东	2 028.36	2 226.84	179.05	345.37	...	3 398.18
华南	998.33	2 023.00	593.98	1 836.67	...	5 371.67
华中	2 664.00	3 844.93	648.00	2 545.56	...	3 998.00
西北	1 440.00	4 193.33	720.00	1 280.00	...	1 180.00
西南	1 932.50	4 030.00	135.00	810.00	...	2 825.00
合计	12 752.52	21 713.42	3 054.11	9 053.16	...	22 052.85

注: C 品牌销售数据二次整理。

2.1.3 浏览和统计数据

运用 SQL Server 系统的导入导出工具将源数据(A 品类),即销售数据采用 EXCEL 文件导入创建的数据库。通过新建数据源视图进行选择、组织以及操作数据源的数据,查看数据分布并进行统计分析。以按地域划分的销售数据为例,通过软件计算得出 C 品牌 7 个地区 2015—2016 年的周销售数据统计值。A 品类 2015—2016 年不同地域销售数据统计值见表 3,其中标准偏差反映一组数据的离散程度,即相比平均值的误差上下波动的幅度。服装销售数据由于受地方气候因素影响较大,其标准偏差偏大属于可接受范围。

表 3 A 品类 2015—2016 年不同地域销售数据统计值

地区	元/(店·周)			
	最大值	最小值	平均值	标准偏差
东北地区	14 407.06	216.00	5 773.64	3 185.13
华北地区	14 835.56	33.33	4 862.73	2 344.38
华东地区	18 405.77	179.05	6 761.23	3 278.32
华南地区	12 185.48	593.98	6 089.40	2 319.02
华中地区	22 712.60	648.00	7 282.52	3 728.75
西北地区	11 206.67	393.33	3 555.23	2 217.25
西南地区	20 670.00	135.00	6 548.18	3 634.01

2.1.4 生成模型

根据数据挖掘算法,挖掘结构用于挖掘问题的数据列以及这些列的信息,以定义问题的域;挖掘模型则是基于挖掘结构构建,用特定的挖掘算法对挖掘结构进行应用计算。应用 SQL Server 系统中的 Analysis Services 项目(SSAS 关系型数据分析服务)提供的回归算法生成挖掘模型(ART 模型)。通过设置特定参数控制决策树的生长、形状以及设置序列周期和处理缺失值等,优化模型预测精度^[15]。ART 模型的参数设置方法见表 4。

表 4 ART 模型的参数设置方法

参数	设置方法
Complexity_Penalty	浮点类型参数,范围在 0~1 之间,决定决策树拆分点,当输入属性少于 10 时,该值设为 0.5
Minimum_Support	指定决策树每一个节点至少包含的分支数,一般默认值为 10
Missing_Value_Substitution	处理数据缺失值的替代,通常采用过去值替代法、平均值替代法以及指定数字替代法
Periodicity_Hint	设置数据周期,可区分模型的优劣,可指定任意多个周期,本文以 1 年为周期,以周为参数设置为 {1,53}
Auto_Detect_Periodicity	控制算法在数据中找出自然周期的强烈程度

预测值由软件直接生成,预测结果由 SSAS 项目的时序查看器显示,包含图表和模型 2 部分:图表显示生成的预测折线图,可根据预测需求设定预测周数,并在可视化窗口查询预测数据;模型显示基于历史销售数据进行交叉计算,形成模型决策树的分布和挖掘图例,可得到相应的回归公式。

2.1.4.1 按渠道划分的预测

根据连衣裙 A 品类线上及线下的历史销售数据,运用时序算法进行交叉计算,得到线上货品需求预测模型。A 品类线上渠道历史销售及预测见图 2。A 品类线上渠道销售额预测模型截图见图 3。

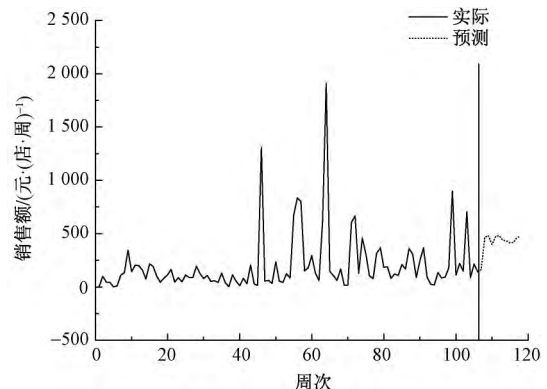


图 2 A 品类线上渠道历史销售及预测

图 3(a)中, w 为周数,周数 44.888 取整数 45, t 为预测周,线下 F_{t-2} 为线下渠道在预测周的前 2 周的历史销售额,元·/(店·周)。在图 3(b)中,事例总计 = 事例在数据库中的出现的次数;期限 = 预测回归方程中的各个变量;系数(列) = 各变量的具体回归系数;菱形标志的分布位置和形状分别表示该节点数据的平均差和标准差的大小:位置越居中,代表平均差越小,预测值越准确;宽度越窄,代表标准差越小,即预测误差越小。

决策树模型中的每个节点都对应一个回归公式。由图 3 整理可得连衣裙线上渠道销售额预测回归公式如下:

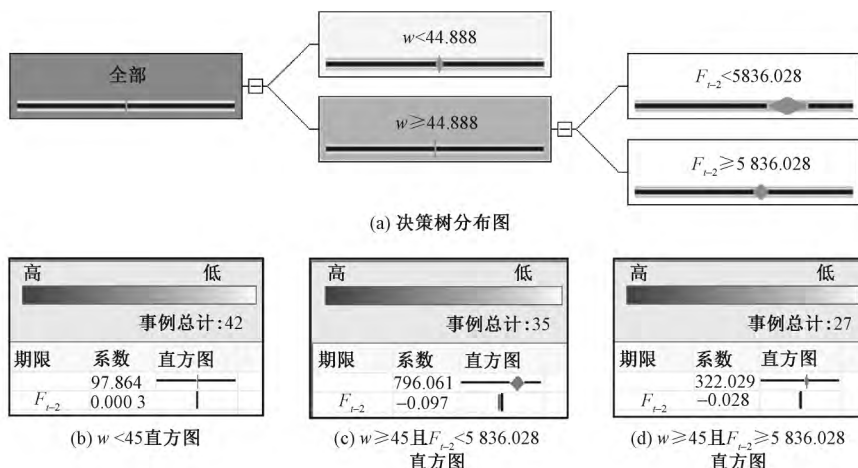


图3 A 品类线上渠道销售额预测模型

$$\begin{cases} O_t = 97.864 + 0.0003 \times F_{t-2} & w < 45 \\ O_t = 796.061 - 0.0970 \times F_{t-2} & w \geq 45 \text{ 且 } F_{t-2} < 5836.028 \\ O_t = 322.029 - 0.0280 \times F_{t-2} & w \geq 45 \text{ 且 } F_{t-2} \geq 5836.028 \end{cases} \quad (1)$$

式中: t 为预测周 $t-1$ 为 t 的前1周 $t-2$ 为 $t-1$ 的前1周; O_t 为线上连衣裙销售额预测值,元/(店·周); F_t 为线下连衣裙销售额预测值,元/(店·周); O_{t-1} 为线上连衣裙历史销售额,元/(店·周); F_{t-1} 为线下连衣裙历史销售额,元/(店·周); w 为周数。

同理,可得到相应的连衣裙线下渠道销售额预测值及回归树模型和回归公式。

2.1.4.2 按地域划分的预测

对连衣裙东北地区线下渠道与全国线下渠道周销售额进行交叉预测,挖掘地区与全国之间的关系。

A 品类东北地区线下渠道历史销售额及预测见图4。A 品类东北地区线下渠道销售额预测模型截图见图5。其中 N_{t-1} 为东北地区线下连衣裙历史销售额, F_{t-53} 为全国线下连衣裙在上年同期的历史销售额,元/(店·周)。其他同图3。

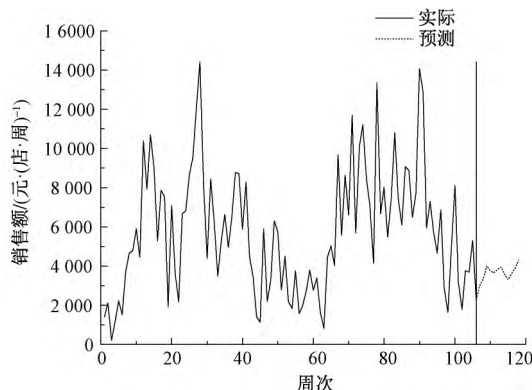


图4 A 品类东北地区线下渠道历史销售额及预测

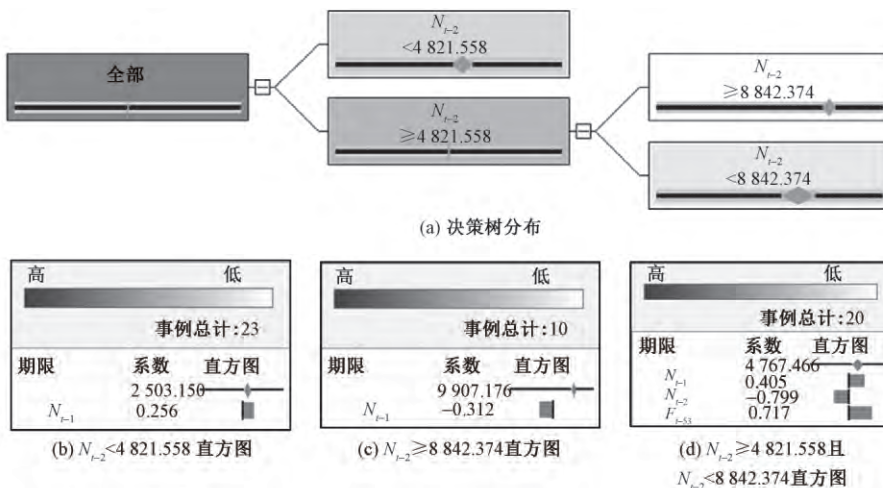


图5 A 品类东北地区线下渠道销售额预测模型

根据图5中2个拆分点可得东北地区线下渠道销售的预测回归公式:

$$\begin{cases} N_t = 2\,503.150 + 0.256 \times N_{t-1} & N_{t-2} < 4\,821.558 \\ N_t = 9\,907.176 - 0.312 \times N_{t-1} & N_{t-2} \geq 8\,842.374 \\ N_t = 4\,767.466 + 0.405 \times N_{t-1} - 0.799 \times N_{t-2} + \\ \quad 0.717 \times F_{t-53} & \\ N_{t-2} \geq 4\,821.558 \text{ 且 } N_{t-2} < 8\,842.374 \end{cases} \quad (2)$$

式中: N_t 为东北地区线下连衣裙销售额预测值, 元/(店·周); F_t 为全国线下连衣裙销售额预测值, 元/(店·周); N_{t-1} 为东北地区线下连衣裙历史销售额, 元/(店·周); F_{t-53} 为全国线下连衣裙历史销售额, 元/(店·周)。

依此类推, 可以得到其他地区线下渠道销售的预测值及回归树模型和回归公式。

2.2 验证模型

以渠道销售预测为例, 通过对 C 品牌连衣裙 2015—2016 年共 106 周销售数据的分析与建模, 得到 2017 年第 1 季度的货品需求预测值, 并使用实际值与预测值的均方误差 (Mean Squared Error, MSE) 和平均绝对误差 (Mean Absolute Error, MAE) 评价预测的准确度。以渠道划分为例, 2017 年 1~12 周的实际与预测销售额对比见图 6。

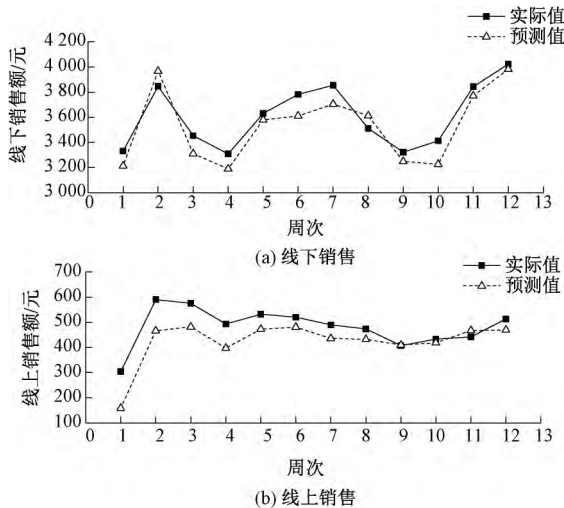


图 6 2017 年 1—12 周的实际与预测销售额对比

均方误差用于评价数据的变化程度, 平均绝对误差指预测值的绝对误差取平均值, 均方误差和平均绝对误差数值越小, 表示预测的精确度越高。在连衣裙线上渠道和线下渠道的预测中, 均方误差分别为 120.46 和 74.83, 平均绝对误差分别为 112.16 和 61.75。根据相关研究算法的误差分析^[16], 以上预测结果的均方误差以及平均绝对误差均在可接受范围, 且图 6 显示, 预测值与实际销售额基本吻合, 说明本文选择并建立的货品需求预测模型可用于服装企业 O2O 货品管理。

2.3 部署和更新模型

服装企业在实际货品需求预测的应用过程中, 由于 O2O 渠道易产生线上及线下的交互影响, 为降低诸多不确定因素, 保证较高的预测准确度, 可通过不断输入新的销售数据, 利用实时数据, 对模型进行检验与修正, 形成持续更新的动态需求预测结构。同时, 企业也可考虑以年、月或天为单位, 按照上述模型构建方式实施部署和更新, 对不同款式、颜色、尺码进行预测或管控。服装动态需求预测框架见图 7。

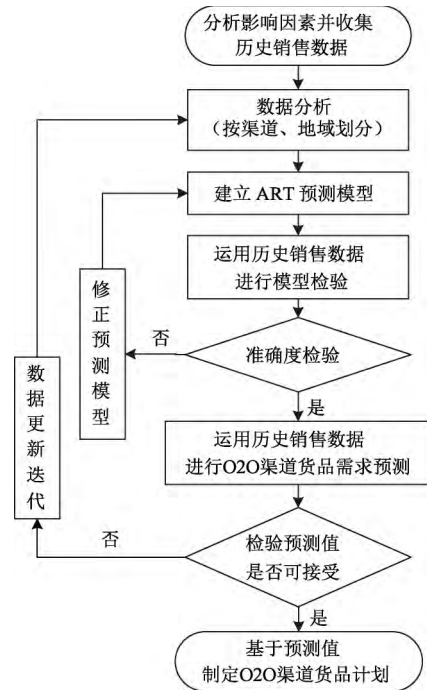


图 7 服装动态需求预测框架

3 基于需求预测的货品管理应用分析

从配货补货等经营要素对 C 品牌进行货品管理: 基于货品需求预测模型, 制定区域店铺业绩指标, 并结合企业设定库存周转率 (以下简称周转率), 决策各波段线上线渠道新品上市时的配货量或补货量。其中, 周转率计算公式如下:

$$\text{周转率} = \text{销售数量(金额)} / \text{平均库存} \times 100\% \quad (3)$$

3.1 配货补货需求分析

C 品牌的配货补货需考虑店铺及商品等级、周转率以及业绩指标等因素, 以确定下一波段货品投放金额 (或数量)。其中周转率是衡量单位时间内库存利用效率的有效指标, 是企业用于店铺销售业绩是否达标的主要考核指标之一。

本文考虑对店铺品类的配货补货, 基于前文建立的货品需求预测模型, 有效利用不同渠道和区域

的数据库信息,计算得出该渠道店铺某品类下一波段销售预测值,并综合去年同期和当期若干周的销售情况以及企业的整体发展规划要求,对预测值进行修正,作为店铺该品类波段业绩指标。结合店铺业绩指标和目标周转率,可求得满足店铺达成指标的平均库存。根据C品牌平均每周均有波段新品上市,且分到每家店铺货品数量较少,本文以周为时间单位,每家店铺的配货补货量用金额表示。

定义变量 k 为目标周库存周转率; Q 为配货补货金额,元/(店·周); K_f 为店铺业绩指标,元/(店·周); I 为店铺每周平均库存,元/(店·周); f 为店铺品类货品需求预测值,元/(店·周); p 为业绩指标制定影响浮动比率,一般取 $-10\% \sim 10\%$ 。

由ART预测模型可获知单店单品类货品需求预测值 f ,得到如下配货补货量计算公式:

$$K_{f(i)} = f_{(i)} \times (1 + p) \quad (4)$$

$$K_{f(i+1)} = f_{(i+1)} \times (1 + p) \quad (5)$$

$$I_{(i)} = \frac{K_{f(i)}}{k} \quad (6)$$

$$I_{(i+1)} = \frac{K_{f(i+1)}}{k} \quad (7)$$

则第 i 周需向该店铺该品类铺货补货金额为:

$$Q_{(i)} = I_{(i+1)} + I_{(i)} + K_{f(i)} \quad (8)$$

若 $I_{(i+1)} \leq I_{(i)} - K_{f(i)}$,表明上一周末库存量可以满足下一周预计的平均库存,不需要对该店铺进行补货。

3.2 案例解析

C品牌连衣裙品类年库存周转率为6%,以1年12个月,每月4周计算,则可设平均每周库存周转率为12.5%,即企业设定目标周库存周转率 $k = 12.5\%$ 。同时假设2017年前2周的业绩指标参照货品需求预测量,无影响浮动,即 $p = 0$ 。

基于前文按渠道和地域划分的货品需求预测模型,预测2017年第1周和第2周的线上及线下货品需求。2017年东北地区平均每家店铺连衣裙需求量见表5。

表5 2017年东北地区平均每家店铺连衣裙需求量
元/(店·周)

周次	线下	线上	合计
1	2 809.93	187.77	2 997.70
2	3 223.38	116.49	3 339.87

由于库存共享,将线上线下2渠道货品需求量相加,得出平均每家店铺的连衣裙第1周和第2周的预测需求量分别为 $f_{(1)} = 2 997.70$ 元/(店·周), $f_{(2)} = 3 339.87$ 元/(店·周)。代入式(8)求得需在

第1周向东北地区店铺平均配货补货的连衣裙金额为:

$$Q_{(1)} = I_{(2)} - I_{(1)} + K_{f(1)} = 5 735.06 \text{ 元/(店·周)}$$

而L企业实际在2017年第1周向东北地区平均每家店铺配货补货C品牌连衣裙19 712元/(店·周),远大于上述所求得理论建议值,且在第2周连衣裙实际共销售3 840元/(店·周),虽略大于预测需求量3 339.87元/(店·周),但还是造成大量库存存储在店铺中,增加了店铺的运营成本和区域间频繁调货的困境。

相对而言,本文基于需求预测模型计算所得的配货补货金额在基本满足区域店铺实际销售需求的情况下,一方面可以保证店铺货品的正常周转,降低运营成本,另一方面把大部分货品储存在总仓库,服务于全国店铺,更易于企业对货品的统筹规划,减少店铺库存不均,减少缺货或库存过剩风险的发生。

4 结束语

本文针对案例企业提出基于货品需求预测模型O2O货品管理应用方案,建立按不同渠道和地域划分的预测模型,由C品牌实际历史销售数据得到的短期预测值与实际需求基本吻合,由此验证了模型在O2O渠道应用中的可行性。根据预测值分析货品需求,以配货补货为例,结合店铺业绩指标和目标周库存周转率,得到店铺达成指标的货品需求量,与企业实际配货补货金额相比不仅能帮助企业实现货品正常周转,降低运营成本,而且可减少缺货及库存过剩风险,为同类型服装品牌企业的货品管理提供借鉴和参考。

参考文献:

- [1] 倪冬梅,赵秋红,李海滨. 需求预测综合模型及其与库存决策的集成研究[J]. 管理科学学报,2013,16(9):44-52.
- [2] 赵秋锦,周静. 品牌服装零售企业双渠道冲突管理策略研究[J]. 丝绸,2015(4):36-41.
- [3] 张细香,马卫民,刘建勋. 采用基因库构建的季节性服装需求预测[J]. 纺织学报,2014,35(5):142-148.
- [4] 孟志青,马珂,郑英. 基于核函数技术的时尚服装需求预测方法[J]. 计算机科学,2016,43(S2):455-460,465.
- [5] WEI D, GENG P, YING L, et al. A prediction study on e-commerce sales based on structure time series model and web search data [C]//The 26th Chinese Control and Decision Conference. Changsha: IEEE, 2014:5346-5351.

- [6] 孙晓静,高慧,陈云. 基于聚类分析和决策树算法的服装销售预测模型[J]. 中国管理信息化,2015(9): 64-67.
- [7] 高融,杨以雄,史晓云. 服装品牌产品销售预测[J]. 东华大学学报(自然科学版),2014,40(5): 651-657.
- [8] 袁亚丽. 时序算法在销售预测中的应用研究[J]. 微计算机信息,2009(15): 249-250.
- [9] 岑琴. 数据挖掘技术在饰品设计中的应用[J]. 商场现代化,2011(1):75-76.
- [10] DAI Q Y, ZHANG C P, WU H. Research of decision tree classification algorithm in data mining [J]. International Journal of Database Theory and Application,2016,9(5): 1-8.
- [11] 徐伟,王林章,李宣东. 基于分类树的随机测试用例生成[J]. 计算机科学,2009,36(1):263-266.
- [12] MEEK C, CHICKERING D M, HECKERMAN D. Autoregressive tree models for time-series analysis [C]// In Proceedings of the Second International SIAM Conference on Data Mining. Arlington: DBLP,2002:229-244.
- [13] MACLENNAN J, TANG Z H, CRIVAT B. 数据挖掘原理与应用[M]. 黄艳,程文俊,译. 北京:清华大学出版社,2010:209-230.
- [14] 于天池. 移动商务环境下服装行业动态销售预测技术研究与应用[D]. 上海:上海交通大学,2008.
- [15] 刘继清,黄金花. 基于改进决策树算法的设备故障智能诊断模型[J]. 制造业自动化,2011,33(7): 30-33.
- [16] 刘金平. 基于 Microsoft 时序算法在电力负荷预测中的应用[J]. 电脑知识与技术,2011,7(8): 1923-1926.