

doi:10.19920/j.cnki.jmsc.2022.12.005

# 基于特征融合的生鲜商品短期销量组合预测<sup>①</sup>

徐小峰<sup>1</sup>, 余乐安<sup>2\*</sup>, 林姿汝<sup>1</sup>, 孙玉萍<sup>1</sup>

(1. 中国石油大学(华东)经济管理学院, 青岛 266580;

2. 四川大学商学院, 成都 610065)

**摘要:** 生鲜产品由于保质期短、易腐易损等特点,对短期销量预测的准确度和可靠性要求极高.为此,本文综合时间、定价、竞价、新鲜度等多种微观层面因素,开展了特征工程分析,并在此基础上提出了生鲜商品销量的ARIMA-NARX组合预测模型.该组合模型首先利用ARIMA模型描述销量时间序列中的线性规律,然后借助衍生降维处理后的特征矩阵,采用NARX捕捉ARIMA残差中的非线性关系,并利用NARX残差预测结果修正ARIMA预测值.最后,将该组合模型的预测结果与ARIMA、NARX、ARIMA-NAR、SVM及回归决策树等模型预测结果及真实观测值进行对比分析,通过MSE/MAPE预测误差评价和DM检验,验证了该组合预测模型的预测能力合理性和有效性,并能较大幅度提高生鲜产品短期销量的预测精度.

**关键词:** 生鲜商品; ARIMA-NARX; 组合预测; 特征工程

**中图分类号:** F205      **文献标识码:** A      **文章编号:** 1007-9807(2022)12-0102-22

## 0 引言

销量预测,尤其是来自供应链下游零售商的销量预测偏差,是决定整个链条或网络牛鞭效应大小的关键因素之一.精准而及时的销售预测对于提高整个供应链的运营效益至关重要.若销量预测不准确,可能会出现货品短缺或过剩,将直接影响企业利润水平.与此同时,销量预测准确性也将直接影响客户服务质量,例如缺货情况可能导致客户选择其他零售商,长此以往易导致客源流失.特别对于生鲜商品而言,其易腐、易损、保质期短等特点对销量预测的精准度提出了更高的要求.因此,生鲜电商把握顾客需求、实现销量精准预测,对于保证商品质量、提高顾客满意度、提升门店经营效益具有重要意义.

近年来,我国生鲜电商行业迅速兴起,除了中小型生鲜电商之外,阿里、京东等巨头纷纷入局,

不断加码冷链物流和生鲜供应链投资,同时也带来了线上线下融合的新零售模式.阿里旗下“盒马鲜生”构建了数据交互平台,实时传输并处理包括消费、供应等在内的数据,并设计了完整的编码体系,这一系统不仅为门店销量预测提供了技术支持,更利于整条生鲜供应链的管理<sup>[1]</sup>.总体而言,虽然我国是生鲜商品消费大国,但大多数生鲜电商在其供应链体系方面仍与发达国家存在差距,其中生鲜商品的零售销量预测就是关键环节之一.

生鲜商品属于季节性的易变质品,对客户订单响应速度、销量预估的要求更高.目前,对于生鲜商品销量预测的多为以月或年为周期的中长期预测,其预测精度低、效果不够理想.此外,预测方法多以单一预测模型为主,少数的组合预测也通常难以实现高精度的预测效果.基于以上两点,本研究将构建模型功能互补、能够实现高精度预测

① 收稿日期: 2019-10-10; 修订日期: 2021-06-23.

基金项目: 国家自然科学基金资助项目(71871222).

通讯作者: 余乐安(1976—),男,湖南常德人,博士,教授,博士生导师. Email: yulean@amss.ac.cn

的组合预测模型,对生鲜商品进行短期销量预测。

综上所述,本研究将综合考虑生鲜商品销量的多种影响因素,从微观层面对其进行短期组合预测。主要贡献如下:1) 构建了生鲜商品销量的 ARIMA-NARX (autoregressive integrated moving average-non-linear autoregressive with exogenous inputs) 组合预测模型,实现以单个零售店的单个存货单位 (stock keeping unit, SKU) 日销量为基本预测单元的短期预测;2) 基于生鲜商品的多维度微观因素数据,综合平稳化、特征转化、特征衍生、特征筛选和特征降维等技术,为实现细小粒度的微观精准预测提供高质量数据基础;3) 应用现实案例数据,对模型进行检验,并与其他单一或组合预测模型进行对比,结果显示本研究提出的组合预测模型误差率平均降低超过 50%,大大提高了预测精确度。

## 1 文献综述

生鲜商品具有保质期短、运输损耗率高、上市周期有限等特性,准确预测销量规模,快速响应市场需求显得尤为重要<sup>[2,3]</sup>。从预测范围看,目前关于易变质品的销量预测研究多为面向产业层面的长周期预测<sup>[4-7]</sup>。由于这种宏观预测可获取的历史统计数据规模较大、可信度较高,传统计量模型即可取得较理想的预测效果。同时,宏观层面的预测通常以年、月为预测周期,以大类商品综合销量为预测单位,预测粒度较粗导致误差抵消机会增多,若缩短预测周期,准确度就会降低<sup>[8]</sup>。

针对易变质商品,短周期销售预测可以更大程度地降低库存和产品时效方面的风险,同时保证销售效益<sup>[9]</sup>。关于短周期销量预测,目前研究在影响因素和预测方法两个方面进行了探索。首先,合理的影响因子矩阵能辅助模型理解数据,提高模型对非线性关系的捕捉能力,因此除了历史销量序列,不少研究对生鲜商品的销量影响因素进行了分析。例如, Arunraj 等<sup>[10]</sup>从气候、季节、价格、产品属性及顾客数量等方面构建香蕉零售销量影响因素体系,并对指标做出了进一步的统计分析;继而, Arunraj 等<sup>[11]</sup>将价格、季节性、节假日等外部因素视为不确定变量,构建带有外部变量的季节模型,解释了上述不确定性因素对需求造

成的影响;刘晶等<sup>[12]</sup>以线上平台交易数据为基础,从产品销售属性、买家个人因素及推广渠道等方面进行特征提取。已有文献侧重创新影响因素体系的指标宽度,在将相关指标进行量化的基础上,通过大量分析研究了影响因素作用于需求与销量的方式,建立起回归模型来描述指标对目标的作用。

在预测方法方面,已有研究基于自回归集成移动平均 (ARIMA)、神经网络 (artificial neural network, ANN)、支持向量机 (support vector machine, SVM) 等方法进行了一些改进。例如, Dellino 等<sup>[13]</sup>利用经 SPO (sequential parameter optimization) 改进的 ARIMA、ARIMAX (ARIMA model with exogenous variables) 及传递函数模型对不同货架期的生鲜食品销量进行了预测,发现同一模型在不同预测目标上预测效果有所不同;Huber 等<sup>[14]</sup>通过聚类分析筛选销售特征相近商品并有针对性地构建 ARIMAX 模型,以此保障模型的适应性;Wang 等<sup>[15]</sup>提出了一种基于 ABC (artificial bee colony) 的方法,通过与交互去噪技术和 ARIMA、SVR 等预测方法的结合,提高预测准确性和鲁棒性。Lee 等<sup>[16]</sup>对比逻辑回归 (LR)、移动平均 (MA) 以及反向传播神经网络 (BPNN) 在每日生鲜销量预测上的表现,认为 MA 可操作性强,LR 擅长揭示二进制数据蕴含的规律,而 BPNN 的预测效果则需要大量数据做支撑;徐琪等<sup>[17]</sup>引入 Bass 改进模型处理小样本数据,并进一步构建 SVM 销量预测模型,以此实现更优的拟合效果。此外,对于多种时间序列预测方法,杨海民<sup>[18]</sup>详细介绍了各类预测方法的模型及特点,例如 ARIMA 模型一般针对均值的单序列拟合预测,ANN 被用来获取时间序列中长期的趋势,而 SVM 则在解决小样本、高维数据和非线性问题方面展现出特有的优势。通过上述研究可发现由于每种预测方法对数据结构、适用条件都有特定要求,只能在具有特定特征的场景、时刻才达到预期的预测精度。

生鲜商品要实现短周期销量预测,需要收集供应链末端的基础数据,数据量大且噪声多,单一的线性或非线性预测模型难以全面揭示数据蕴含的规律。相较而言,组合预测方法能够综合两种及两种以上单一预测方法的优良性能,达到提高预

测精度的效果<sup>[19]</sup>. ARIMA 模型是目前较为成熟的时间序列预测方法,但本质上只能描述线性关系,对非线性关系的捕捉能力不足.为此,许多文献<sup>[20-24]</sup>引入带外部输入的非线性自回归神经网络模型(NARX)来处理时间序列的非线性特征. NARX 模型是循环神经网络(recurrent neural networks, RNN)的一种,它不仅具有静态神经网络的非线性映射能力,同时 NARX 内部设置了时延层(time-delay layer, TDL),能够将输出向量反馈给输入端,使得模型具有了动态特性,从而能更好地应用于时间序列多步预测.国内外不少文献将 NARX 应用在非线性系统建模工作中<sup>[20-24]</sup>,其容噪能力、收敛速度及泛化能力也得到了验证.可以发现,ARIMA 和 NARX 模型在解释线性及非线性关系方面存在一定互补性,将二者结合具有提升预测准确性潜力.

目前组合预测的组合途径主要包括分解输入端时间序列变量<sup>[25]</sup>、增加并筛选输入端影响因子<sup>[10, 26]</sup>、优化预测模型过程参数<sup>[13, 27]</sup>、通过残差预测修正初始预测结果<sup>[28-30]</sup>,以及将多种单一预测结果加权平均<sup>[8]</sup>等方式.考虑 ARIMA 及 NARX 的特点,研究将基于“残差补偿”思想<sup>[28-30]</sup>提出 ARIMA-NARX 组合预测模型,在理论意义上,实现了模型在预测线性及非线性关系上的优势互补,

构造出预测精度更高的组合预测模型,为预测领域的研究提供了新的思路和方法.同时,在现实意义上,通过对历史销售数据、时间、商品属性、营销策略等生鲜商品销量影响因素进行特征工程处理,充分挖掘销量数据蕴含的信息,构建粒度更小的商品销量预测模型对生鲜商品日销量进行预测,以满足企业按日管理库存的需求,为企业供应链末端的库存决策提供支持<sup>[31]</sup>.

## 2 模型构建

### 2.1 问题描述

由于生鲜商品的新鲜度随时间衰减明显,为保证稳定的商品品质,门店一般采取高频次的进货方式.但由此易带来较高的库存水平,进而使门店面临较高的商品过期损耗成本.反之,若降低库存水平则易发生缺货,而生鲜商品属于日常刚需,顾客等待意愿普遍较低,缺货易直接造成订单损失.可见,相比一般商品,生鲜商品对滞销和缺货的容忍程度更低,因此其对门店每日库存水平和实际需求的匹配程度提出了更高要求.研究将以生鲜商品  $i$  第  $t$  天在门店  $j$  的日销量为预测目标,根据历史销量及相关特征构建预测模型,预测结果将为门店  $j$  的进货行为提供基本进货决策依据.

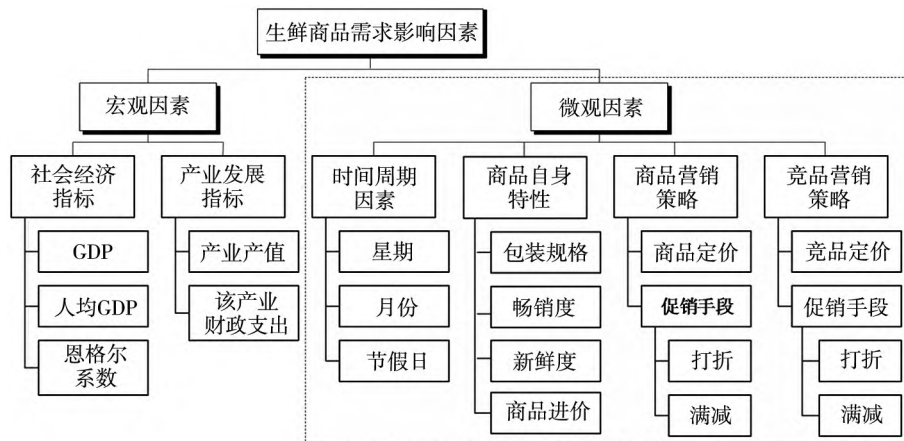


Fig.1 The classification system of influencing factors of fresh commodity demand

通常,顾客购买行为与多种因素相关,这些因素直接或间接地驱动着市场需求.根据相关文献<sup>[10,31]</sup>及零售商的销售经验,对生鲜商品的需求影响因素进行总结(如图1).从宏观和微观两个层面来看,宏观层面主要与社会总体经济情况和

产业发展趋势相关;微观层面可概括为时间周期性因素、商品自身特性、商品自身的营销策略、可替代品的营销策略等四类影响因素.其中,宏观因素主要对长期需求产生影响,短期内保持稳定,因此研究将主要从微观层面进行特征提取.



2.2 数据预处理

数据质量对模型表现至关重要,数据的量纲不一、缺失、噪声等问题决定了后续模型能否有效工作.为使数据能够适应模型需求,应首先对原始数据进行检测,纠正或删除异常数据记录,这个过程被称为数据预处理.本研究数据预处理过程主要涉及无量纲化、缺失值补充以及分类特征转化.

2.2.1 数据无量纲化

无量纲化是指将不同规格的数据转换到同一规格或将不同分布的数据转换到特定分布的过程,其不仅能够提升模型运算速度,同时能够避免某一个取值范围特别大的特征影响模型精度.由于销量样本数据存在较多噪音,选用对异常值相对不敏感的标准化(standardization)来进行缩放处理.

$$x^* = \frac{x - \mu}{\sigma}$$

(1)

式中  $x$  为待转换序列,  $\mu$  和  $\sigma$  分别为  $x$  的均值和标准差,将  $x$  按  $\mu$  进行中心化处理,并按  $\sigma$  进行缩放,即得到序列  $x^* \sim N(0,1)$ .

2.2.2 缺失数据处理

从实际业务场景中获取的原始数据不可避免存在很多缺失字段,若全部剔除缺失字段所在记录则可能会严重影响数据规模,因此需要采取合理的手段补全缺失值.将依据具体特征类型和内涵,选择临近值、均值、中值、众数或者特定值来填

补缺失值.

2.2.3 分类特征转化

为使数据适应编程及算法要求,需要将文字型分类特征转换为数值型.同时为尽可能传递给算法准确信息,引入虚拟变量来处理时间戳等名义变量.

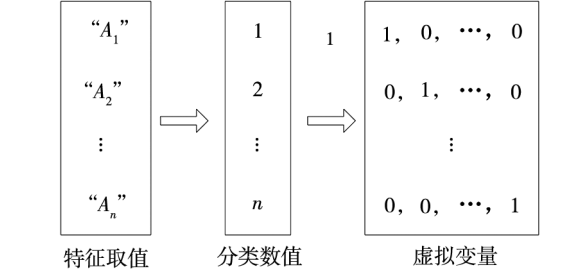


图 2 分类特征转化

Fig. 2 Transformation of classification feature

2.3 特征工程

“数据和特征很大程度上决定了机器学习的上限,模型和算法只是逼近这个上限而已<sup>[32]</sup>”特征工程是将原始数据转换为更能代表数据内涵的特征矩阵过程,利用特征矩阵训练模型,能够降低模型计算成本、提升预测潜力.特征工程包括特征提取、特征衍生、特征选择及特征降维等步骤(如图 3),研究将基于前文梳理的生鲜商品微观层面的需求影响因素,开展特征工程分析.对于特征提取,基于两个原则:1)特征对目标变量的解释程度高;2)特征间的相关程度低.基于此原则进行初始特征的选择.

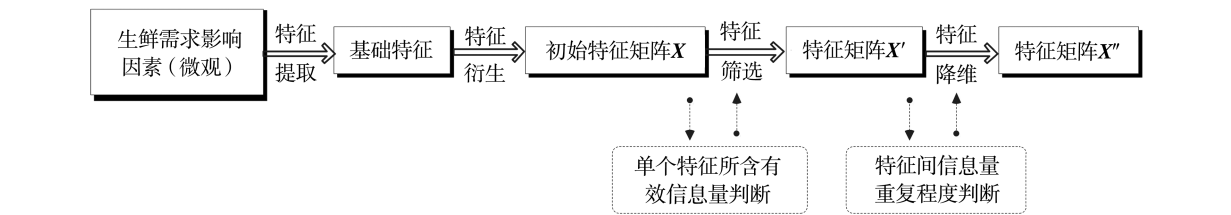


图 3 特征工程处理过程

Fig. 3 Feature engineering process

2.3.1 特征衍生

考虑到数据的可获取性、指标刻画的难易程度及准确性,在特征衍生之前,首先对图 1 中的指标调整作如下说明.其中,商品的包装规格可由单笔购买量体现.例如,客户购买多个同款商品说明包装规格过小,反之若多数客户购买单个同款商品,则其规格适中,满足单笔购买需求.

在销量时间序列预测问题中,由于数据本身特征数量十分有限,因此需要构造大量的衍生特征来增强模型的拟合能力.结合对业务场景的理解,基于图 4 的指标说明,基于历史数据,通过滑动窗口构建一系列统计特征、比值特征以及差分特征等(如表 1),采用滚动预测法,可以依靠当天的输入特征变量,来预测未来的生鲜销量.

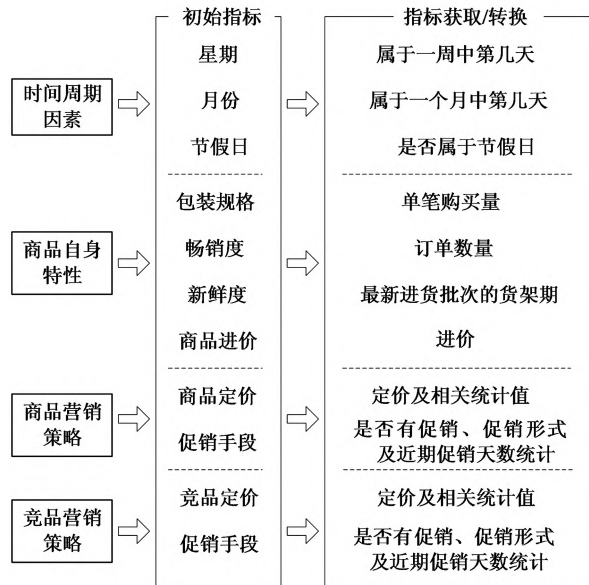


图 4 影响因素指标获取/转换说明

Fig.4 Instructions for indicator acquisition and conversion of influencing factors

表 1 生鲜销量预测特征衍生

Table 1 Feature creation for fresh product sales forecast

因素	基本特征及相关衍生特征
销量序列	销量 $Y_t$ ; [ $t - T$ , $t$ ] 时间段中销量的均值、中位数、标准差、最值、总和; $Y_t - Y_{t-1}$ 以及 [ $t - T$ , $t$ ] 期间该差分序列的均值、中位数、标准差、总和、最值.
时间周期因素	属于一周中第几天; 属于一个月中第几天; 是否属于节假日.
商品自身特性	单笔购买量; 订单数量; 最近进货批次的货架期; 商品进价 $p_t$ .
商品/竞品营销策略	定价 $P_t$ 、促销手段; [ $t - T$ , $t$ ] 时间段中定价的均值、中位数、标准差、最值; ( $P_t - P_{t-1}$ ) 以及 [ $t - T$ , $t$ ] 期间该差分序列的均值、中位数、标准差、最值; 商品 ( $P_t / p_t$ ) 以及 ( $P_t / p_t$ ) - ( $P_{t-1} / p_{t-1}$ ); [ $t - T$ , $t$ ] 时间段中促销天数之和.

### 2.3.2 特征筛选

通过特征提取和特征衍生得到的全部特征未必全部有效,而无效信息输入到预测模型中将降低求解效率,甚至影响预测精度.因此,需通过一定手段来筛选最有效的特征子集,作为模型训练的依据.特征的选择需要从特征的分散性及特征与目标的相关性两方面进行考虑.方差是衡量特

征是否发散的关键指标,而互信息指数能够评价定性自变量对因变量的相关性,两者均在特征筛选中表现出良好的性能<sup>[33-36]</sup>,研究将采用方差过滤和互信息法进行特征选择.

1) 方差过滤:方差是代表数据离散程度的统计值,对于单个特征来说,方差越小,意味着所有样本在这个特征上的取值差异越小,则该特征对样本区分的贡献度不高.反之,如果一个特征的方差越发散,意味着该特征在所有样本的取值差异越大,对区分样本的作用越大.因此,需根据样本实际情况设置一定阈值,应用方差过滤法来移除方差不满足此阈值的特征.默认情况下,它将会移除所有的零方差特征,即那些在所有的样本上的取值均不变的特征<sup>[33-36]</sup>.

2) 互信息法:互信息 (mutual information) 是度量两个事件集合间相关性的指标,表示两个变量间共同拥有的信息的含量.设两个随机变量  $X$  和  $Y$  的联合概率分布为  $p(x, y)$ , 边缘分布分别为  $p(x)$  和  $p(y)$ , 则互信息  $I(X; Y)$  是联合分布  $p(x, y)$  与边缘分布  $p(x)p(y)$  的相对熵,  $I(X; Y)$  定义如下<sup>[35]</sup>

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \ln \frac{p(x, y)}{p(x)p(y)} \quad (2)$$

互信息法能够捕捉特征和预测目标间任意的线性及非线性关系,通过互信息量的估计值判断特征与预测目标间相关关系的强弱,若  $I = 0$ ,则表示两变量相互独立,不存在重叠的信息,应舍弃该特征;反之,两者的相互依赖程度越高,  $I$  越大,共同信息越多.

### 2.3.3 特征降维

特征工程中的降维是指降低特征数量,目的是合并高维数据中的重复信息,并删除无效信息,用尽量少的特征高效地描述原特征矩阵带有的有效信息.若经筛选的特征矩阵中特征间显示有较强的相关性,研究将利用主成分分析法 (principal component analysis, PCA) 实现特征降维.

## 2.4 模型构建

### 2.4.1 ARIMA 模型介绍

1) 基本原理:ARIMA ( $p, d, q$ ) 模型是一种常用的时间序列预测的模型,其中参数  $p, d, q$  分别代表时序数据本身的滞后阶数、差分次数以及预测误差的滞后阶数.模型的基本原理是借助内生

变量,从自身历史数据中学习序列的变化规律,并用这种模式做预测。

2) 主要步骤: ARIMA 模型通过描述时间序列变量本身的变化规律,利用外推机制预测时间序列的未来走势. ARIMA 建模的步骤如图 5 所示。

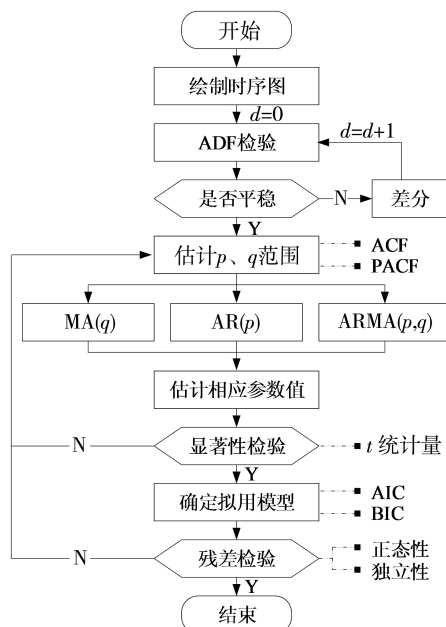


图 5 ARIMA 建模流程

Fig. 5 Modelling procedure of ARIMA

3) 适用范围: ARIMA 模型要求时序数据是稳定的,或者是通过差分化后的稳定时序数据;且在本质上只能捕捉线性关系。

#### 2.4.2 NARX 模型介绍

1) 原理: NARX 是非线性动态系统中应用最广泛的一种神经网络. 在静态多层神经元感知器基础上引入了 TDL (tapped delay lines) 延迟线模型以及输出反馈功能,从而通过反馈与记忆功能提高模型对历史数据的敏感性,实现网络动态性. NARX 是一个离散非线性系统,具体可表示为

$$y(t) = f[y(t-1), y(t-2), \dots, y(t-d_y), X(t-1), X(t-2), \dots, X(t-d_x)] \quad (3)$$

式中  $f$  是表示系统输入输出关系的非线性函数,  $y(t)$  和  $X(t)$  分别表示  $t$  时刻的输出序列和多输入矩阵,  $y(t)$  准确度与输入层和反馈层的延迟阶数  $d_y$  和  $d_x$  密切相关。

2) 主要步骤: NARX 主要由输入层、隐含层、输出层、输出反馈层及输入和输出延时构成,主要建模步骤<sup>[37]</sup>如图 6 所示。

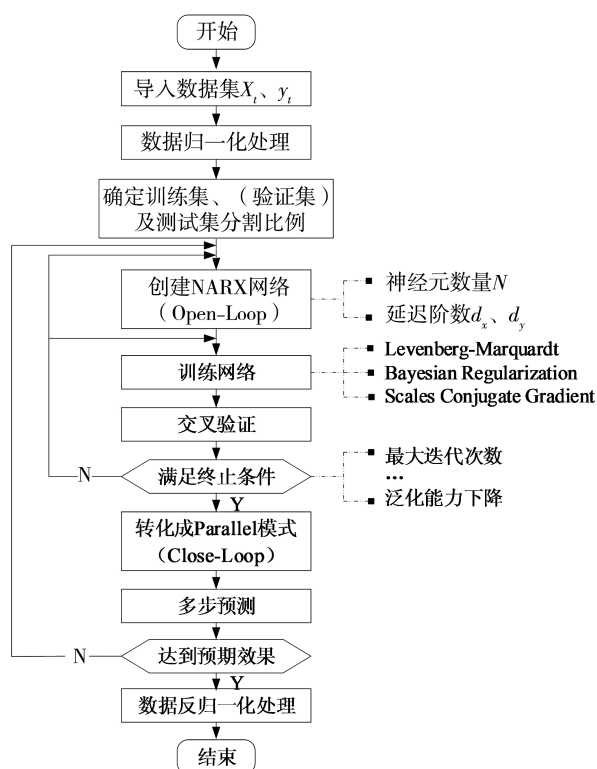


图 6 NARX 建模流程

Fig. 6 Modelling procedure of NARX

在标准的 NARX 网络结构中输出值被反馈到的输入端,如图 7(A). 但由于在 NARX 网络训练过程中训练目标已知,因此可以建立图 7(B)所示的 Open-Loop. Open-Loop 比 Close-Loop 训练更有效,因为其向输入端提供的是正确的历史输出值,有利于优化训练效果,同时 Open-Loop 是单纯的前馈神经网络,可以直接使用静态神经网络的建模函数. Open-Loop 只能向前预测一步,因此为进行多步预测,网络训练完成后将转换成 Close-Loop 模式进行迭代预测<sup>[23]</sup>。

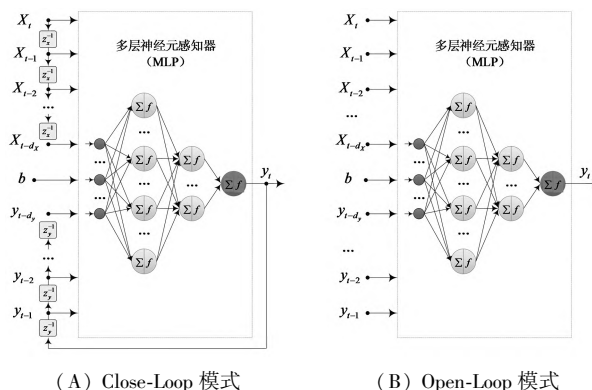


图 7 NARX 网络结构示意图

Fig. 7 Diagram of NARX network structure



3)适用范围:在非线性问题的时间序列预测方面有着优秀的记忆能力,对于中长期时效的非线性问题具有良好的预报、预测性能<sup>[38]</sup>。

#### 2.4.3 ARIMA-NARX 模型构建

ARIMA 和 NARX 模型在预测线性和非线性关系上优势互补,ARIMA 模型本质上只能描述线性关系,而不能捕捉数据中隐藏的非线性关系,NARX 则恰好相反.因此,研究在预测模型方面进行了创新,利用“残差补偿”思想构建 ARIMA-NARX 组合预测模型.首先,利用 ARIMA 捕捉历史销量数据序列中的线性关系,然后借

助特征矩阵利用 NARX 解释 ARIMA 预测所得残差中的非线性关系.模型的相比其它单一预测模型或组合预测模型而言,研究提出的 ARIMA-NARX 组合预测模型在理论上能够提高预测的精度.假设销量时间序列  $Y_t$  可转化为线性序列  $L_t$  与非线性残差序列  $R_t$  的加和,即  $Y_t = L_t + R_t$ ,以生鲜销量影响因素——时间周期性因素  $w_t$ 、商品自身营销策略  $\theta_t$ 、竞品营销策略  $\gamma_t$  及商品自身特性  $\eta_t$  等四类影响因子为基础进行特征处理,ARIMA-NARX 组合模型构建步骤如图 8 所示.

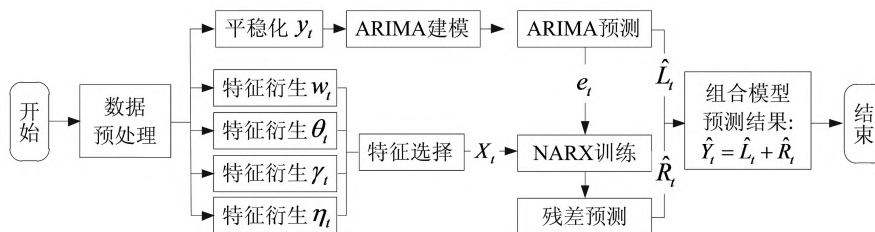


图 8 ARIMA-NARX 组合模型预测步骤

Fig. 8 Prediction steps of combined model ARIMA-NARX

**步骤 1** 利用 ARIMA 对历史数据进行线性部分拟合,将一非平稳的随机时间序列  $y_t$  经过  $d$  次差分后转化为一个平稳序列.经过 ARIMA 建模及预测后,预测结果记为  $\hat{L}_t$ ,残差序列  $e_t = Y_t - \hat{L}_t$ , $e_t$  中包含了 ARIMA 模型无法解释的非线性关系;

**步骤 2** 以  $e_t$  和  $w_t$ 、 $\theta_t$ 、 $\gamma_t$  及  $\eta_t$  等生鲜销量影响因素为基础开展特征工程,得到特征集合  $X_t$  作为 NARX 网络的输入输出数据,利用 NARX 建模逼近  $e_t$  的非线性规律,得到预测结果  $\hat{R}_t$ ;

**步骤 3** 利用残差预测值  $\hat{R}_t$  对 ARIMA 预测结果  $\hat{L}_t$  进行修正,得到组合预测结果  $\hat{Y}_t = \hat{L}_t + \hat{R}_t$ .

## 2.5 模型评估

### 2.5.1 评估方法

交叉验证是用来调节机器学习中的超参数,并对模型性能进行评价的方法.传统交叉验证方法将数据划分为  $k$  份,每次测试取其中一份作为测试集,其他  $k-1$  份作为训练集,从训练集中均匀随机抽样一部分样本作为验证集,综合  $k$  次训练误差的均值作为模型效果的度量值.对于时间序列预测问题,随机选择测试集的方法会打乱数据的时间顺序,导致数据泄露.因此应确保测试集

在时间顺序上处于训练集之后,采用如图 9 所示的嵌套交叉验证方法<sup>[39]</sup>,在保证训练集数据规模的前提下,采用滚动窗口按时间顺序多次取测试集,将测试集之前的数据作为训练集,以此得到多组不同的训练集/测试集.对每次测试得到的误差求平均,即得到模型精度的鲁棒估计.

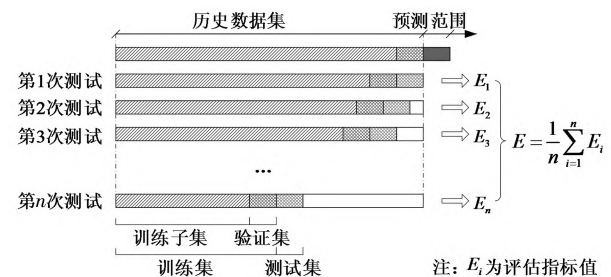


图 9 时间序列交叉验证过程

Fig. 9 Time series cross-validation process

### 2.5.2 评估指标

在交叉验证及评估预测精度过程中,销量预测模型均可采用均方误差 (mean square error, MSE) 和平均绝对百分比误差 (mean absolute percentage error, MAPE) 等指标作为评价标准<sup>[40-42]</sup>,前者反映预测值和观测值间的绝对误差,后者反映相对误差.

$$MSE = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n} \quad (4)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left( \frac{|Y_i - \hat{Y}_i|}{Y_i} \right) \times 100\% \quad (5)$$

式中  $Y_i$  为观测值,  $\hat{Y}_i$  代表预测值,  $n$  代表样本数量。

### 2.5.3 Diebold-Mariano 检验

考虑到上述评估指标仅能够作为比较模型预测精度的标准,却无法检验比较结果在统计上是否具有显著性。为此,Diebold 和 Mariano<sup>[43]</sup>提出一种渐进正态分布检验法,简称 DM 检验。DM 检验用于模型间的两两比较,其基本原理如下。

假设模型  $i$  预测误差序列为  $e_i$ ,  $g(e_i)$  代表上述评估指标,两模型相对损失函数值为  $d = g(e_1) - g(e_2)$ 。若两模型的预测能力无差别,则接受原假设

$$H_0: E(d) = E(g(e_1) - g(e_2)) = 0 \quad (6)$$

若  $E(d) > 0$  则说明模型一的预测能力劣于模型二,反之则说明预测能力优于对方。DM 统计量公式为

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi\hat{f}_d(0)}{T}}} \sim N(0,1) \quad (7)$$

其中  $\bar{d}$  为样本损失差均值,  $\hat{f}_d(0)$  是损失差频率为 0 时的谱线密度,  $\hat{f}_d(0)$  是  $f_d(0)$  的一致估计,经标准化处理 DM 统计量服从  $N(0,1)$  标准正态分布。

## 3 模型应用

四海绿仓公司(以下简称“S 公司”)是一家专注优质农产品销售的 FAF(farm and family)生鲜电商公司,致力于将生鲜电商由最后 1 km 推进到最后 100 m。针对传统农副产品流通链条冗长带来的一系列成本与安全等问题,S 公司对流通体系进行升级改造,建立起直接连接生产基地与社区居民的流通平台(图 10)。相比传统流通形式,S 公司作为供应链集成服务商具有较强的控制力,但同时也需要承担整个供应链的盈亏风险性。为此,S 公司计划通过精准销量预测实现对库存精细化管理。

### 3.1 数据获取及分析

#### 3.1.1 数据获取

以 S 公司香蕉品类在山东青岛黄岛区线上(由 S 公司总仓负责配送,记为 1 家门店)及线下

6 家门店的销量数据作为研究样本,数据选取范围从 2017 年 10 月 23 日到 2018 年 5 月 20 日(约 6 个月,28 周),包括 45 760 条门店 POS(point of sale)端数据以及 7 486 条进货数据,经统计整理得到 2 744 条样本数据。在此基础上,以单个门店单个 SKU 的每日销量为基本预测对象,并逐一一对每个 SKU 样本在未来一周(2018 年 5 月 21 日 ~ 2018 年 5 月 27 日)的销量趋势进行预测。需要注意的是,S 公司的线上门店不进行线下销售,但所有区域内的线上订单需由其配送;线下门店则负责社区的实物零售、送货上门及自提等业务。两者销售规律受时间周期、天气等因素影响程度不同,可用于验证预测模型的适用性。

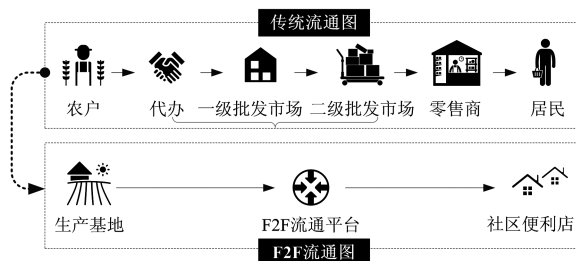


图 10 S 公司 FAF 生鲜商品流通体系

Fig. 10 FAF circulation system of Company S

需说明的是,S 公司香蕉品类中有销售记录的 SKU 主要包括广西国产香蕉、云南国产香蕉、都乐香蕉、菲律宾皇帝香蕉以及南美进口香蕉等多种类型。由于公司会根据市场行情不定期调整进货 SKU,导致任意 SKU 的销售数据连续覆盖时间都较短。并且,通过观察香蕉品类的销售数据,可以发现顾客对同类 SKU 的需求具有相似性,同时期在售 SKU 种类主要可按价格划分为精品(高价)和普通(低价)两类。因此为便于预测分析并达到样本数量要求,对基本 SKU 进行了合并处理,其中  $sku\_id = 1$ (普通)包括广西国产香蕉和云南国产香蕉,  $sku\_id = 2$ (精品)则包括都乐香蕉、菲律宾皇帝香蕉以及南美进口香蕉。根据样本数据统计两个 SKU 的相关特征如表 2 所示,可发现消费者对普通等级香蕉的购买需求是远大于精品类,说明消费者对生鲜商品的价格是非常敏感的。

表 2 各香蕉 SKU 销量统计特征

Table 2 Sales statistical characteristics for each SKU

id	Mean	std	Min	25%	50%	75%	Max
1	29.7	23.5	0.0	13.4	24.7	40.0	173.9
2	10.5	10.1	0.0	2.5	7.9	15.2	76.3



结合统计数据各 SKU 在公司各个门店的销量趋势如图 11 所示,由于 2018 年 2 月 16 日是春节,绝大多数门店不营业,因此该段数据对于预测意义不大,可以作删除处理. 为保持数据在一周的周期性,删除了 2018 年 2 月 12 日~2018 年 2 月 25 日两周的数据,后文将在此基础上进行预测及库存调拨研究.

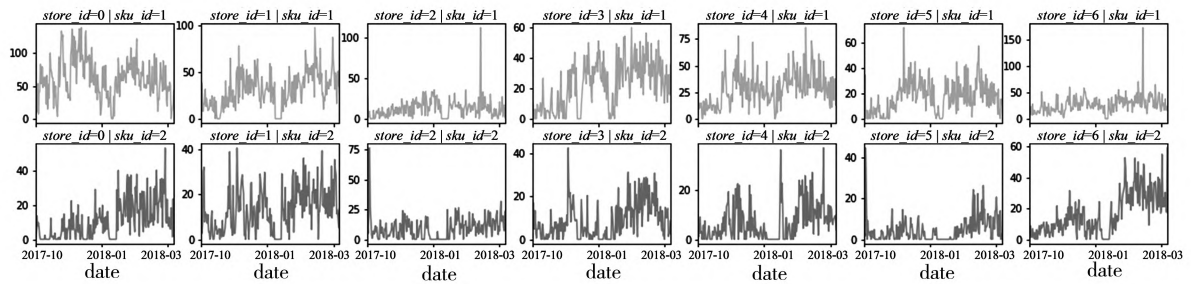


图 11 香蕉 SKU 在各门店的销量走势图

Fig. 11 Sales trend for banana SKU in each store

表 3 各门店香蕉品类销量统计特征

Table 3 Statistical characteristics of banana sales in each store

store_id	class	Mean	std	Min	25/%	50/%	75/%	Max
0	-	37.16	34.29	0.0	7.83	27.94	61.15	139.22
1	H	23.36	17.34	0.0	10.62	19.37	32.85	97.42
2	L	11.49	10.40	0.0	4.06	10.16	16.52	112.40
3	M	16.31	14.45	0.0	3.91	12.90	26.76	60.148
4	M	17.97	16.37	0.0	4.74	13.68	26.19	85.75
5	L	11.51	11.90	0.0	1.85	7.45	18.38	72.35
6	H	22.74	16.21	0.0	10.60	21.08	32.21	173.88

注: store\_id=0 代表总仓负责的线上销售额

## 2) 香蕉 SKU 以星期为周期的销量规律

为寻找周度销量规律,以样本数据为基础对每个 SKU 在每个门店的销量进行统计,如图 12 所示. 观察香蕉销量在一周内的分布,发现星期六

## 3.1.2 数据探索

### 1) 香蕉品类销量数据整体统计特征

观察样本数据在各个门店的历史销量特征(表 3),发现香蕉品类在不同门店之间销售规模及规律存在差异,若依据销量规模将门店划分为 Low(L)、Medium(M)、High(H),H 类门店的日均销量可以达到 L 类的两倍以上.

和星期天的销量普遍高于工作日,紧邻的星期一销量低迷,但星期二销量又出现了高峰. 可见,人们的购买意愿在工作日、非工作日及非工作日前后会有明显差别,有必要在特征工程中加以重视.

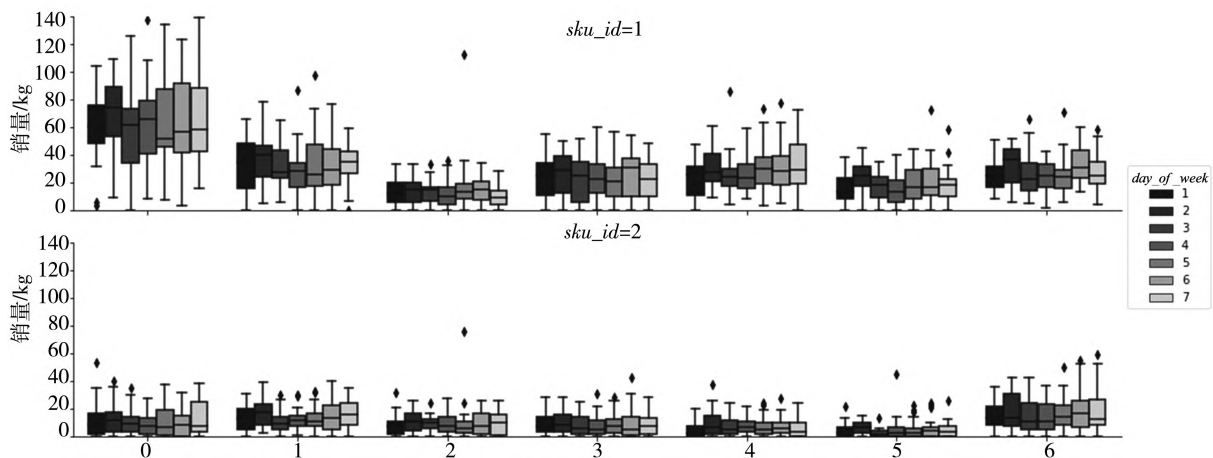


图 12 香蕉销量以周为周期的分布情况

Fig. 12 Distribution of banana sales on a weekly basis

3.1.3 特征工程

为保证特征处理含义具有一般性,以下将以全部 2 744 条样本数据呈现特征工程处理过程,实际预测过程需针对 14 个具体样本分别进行特征处理.

1) 特征衍生

依据生鲜商品的特征衍生思路,从销量序列

自身、时间周期因素、商品自身特性、商品自身及竞品的营销策略等方面提取特征共 59 个,并初步构建特征集合如下表 4 所示. 图 13 为各个特征间 Pearson 相关系数的热力图,可以直观地观察到部分同类特征存在相关性,为使特征信息能够高效表达,应对特征矩阵进一步筛选和降维.

表 4 特征集合初步构建

Table 4 Preliminary construction of feature set

因素	变量名	说明
销量序列	<i>sales</i>	销量序列(预测目标)
	<i>sales_3d_mean</i>	过去 3 天销量的均值
	<i>sales_7d_mean</i>	过去 7 天销量的均值
	<i>sales_7d_median</i>	过去 7 天销量的中位数
	<i>sales_7d_std</i>	过去 7 天销量的标准差
	<i>sales_7d_max</i>	过去 7 天销量的最大值
	<i>sales_7d_min</i>	过去 7 天销量的最小值
	<i>sales_14d_mean</i>	过去 14 天销量的均值
	<i>sales_14d_median</i>	过去 14 天销量的中位数
	<i>sales_14d_std</i>	过去 14 天销量的标准差
	<i>sales_14d_max</i>	过去 14 天销量的最大值
	<i>sales_14d_min</i>	过去 14 天销量的最小值
	<i>sales_diff</i>	第 $t$ 天与 $t-1$ 天销量的差值,记为 $\Delta Sales$
	<i>sales_diff_3d_mean</i>	过去 3 天 $\Delta Sales$ 的均值
	<i>sales_diff_7d_mean</i>	过去 7 天 $\Delta Sales$ 的均值
	<i>sales_diff_7d_median</i>	过去 7 天 $\Delta Sales$ 的中位数
	<i>sales_diff_7d_std</i>	过去 7 天 $\Delta Sales$ 的标准差
	<i>sales_diff_7d_max</i>	过去 7 天 $\Delta Sales$ 的最大值
	<i>sales_diff_7d_min</i>	过去 7 天 $\Delta Sales$ 的最小值
	<i>sales_diff_14d_mean</i>	过去 14 天 $\Delta Sales$ 的均值
	<i>sales_diff_14d_median</i>	过去 14 天 $\Delta Sales$ 的中位数
	<i>sales_diff_14d_std</i>	过去 14 天 $\Delta Sales$ 的标准差
时间周期因素	<i>day_of_week</i>	第 $t$ 天属于一周中的第几天(转换为哑变量),包括 <i>day_of_week_1-7</i>
	<i>day_of_month</i>	第 $t$ 天属于一个月中的第几天
	<i>official_holiday</i>	第 $t$ 天是否是法定节假日
	<i>official_holiday_before</i>	第 $t$ 天是否是节假日前一天
	<i>official_holiday_after</i>	第 $t$ 天是否是节假日后一天
商品自身特性	<i>purchase_mean</i>	第 $t$ 天顾客的平均单笔购买量
	<i>purchase_frequency</i>	第 $t$ 天订单数量
	<i>shelf_life</i>	商品累计货架期,即进货天数
	<i>cost</i>	第 $t$ 天在售批次的进价

续表 4  
Table 4 Continues

因素	变量名	说明
商品自身营销策略	<i>price</i>	第 $t$ 天商品定价
	<i>profit</i>	第 $t$ 天定价与进价的比率
	<i>profit_diff</i>	第 $t$ 天与 $t-1$ 天定价与进价的比率的差值
	<i>price_7d_mean</i>	过去 7 天定价的均值
	<i>price_14d_mean</i>	过去 14 天定价的均值
	<i>price_diff</i>	第 $t$ 天与 $t-1$ 天定价的差值, 记为 $\Delta Price$
	<i>price_diff_7d_mean</i>	过去 7 天 $\Delta Price$ 的均值
	<i>price_diff_14d_mean</i>	过去 14 天 $\Delta Price$ 的均值
	<i>promotion_1</i>	第 $t$ 天是否门店有满减促销
	<i>promotion_2</i>	第 $t$ 天是否商品自身有打折促销
	<i>promotion</i>	第 $t$ 天是否有促销(任意形式)
竞品营销策略	<i>promotion_7d_sum</i>	过去 7 天各类促销天数和
	<i>compete</i>	是否有同类商品在售
	<i>compete_sales</i>	第 $t$ 天竞品的销量
	<i>compete_sales_7d_mean</i>	过去 7 天竞品销量的均值
	<i>compete_sales_diff</i>	第 $t$ 天与 $t-1$ 天竞品销量差值, 记为 $\Delta Sales2$
	<i>compete_sales_diff_7d_mean</i>	过去 7 天 $\Delta Sales2$ 的均值
	<i>compete_price</i>	竞品定价
	<i>compete_price_diff</i>	第 $t$ 天与 $t-1$ 天竞品定价差值, 记为 $\Delta Price2$
	<i>compete_price_diff_7d_mean</i>	过去 7 天 $\Delta Price2$ 的均值

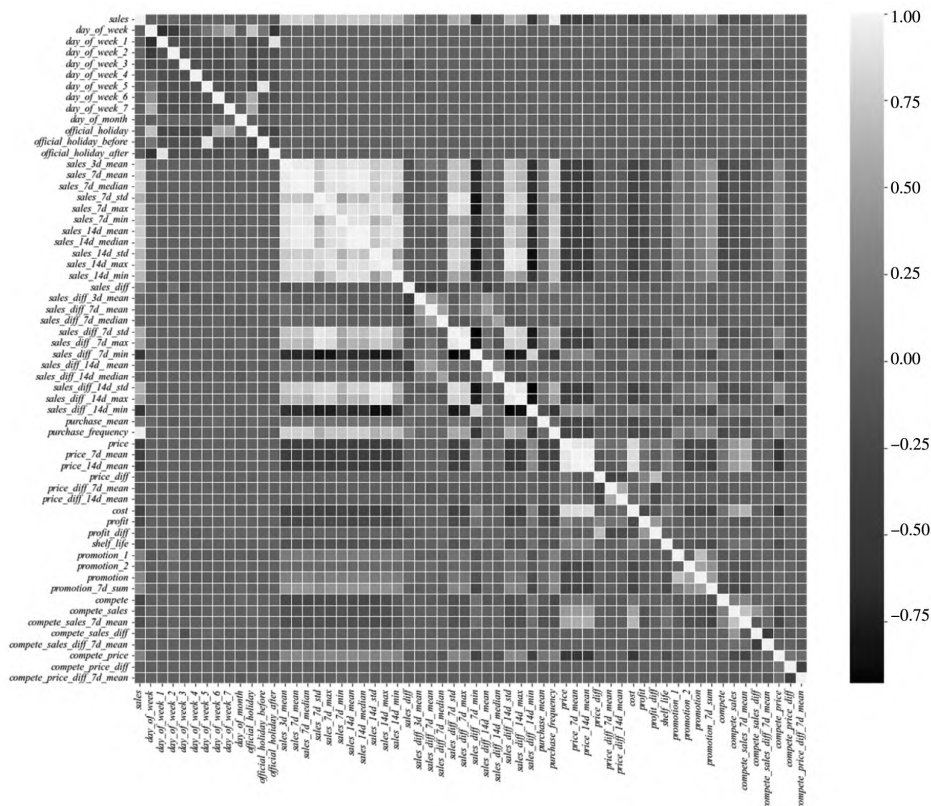


图 13 初始特征集合相关系数矩阵

Fig. 13 Initial correlation coefficient matrix of feature set



## 2) 特征筛选

利用 Python 中 Sklearn 算法模块对特征集合进行数据预处理,并依据特征自身方差(Var)以及特征与标签(实际观察值序列)间的互信息估计量(MIR)进行特征过滤,权衡模型表现和求解速度,分别设置  $\text{Var} = 0$  和  $\text{MIR} = 0$  为阈值,经过滤得到特征子集包含 54 个有效特征.依据互信息量对特征进行排序后发现购买频率以及单笔购买量与标签之间关系最紧密,次相关的是过去 7 天、14 天销量的均值、中位数、极值、标准差等一系列统计特征,进价、竞品定价、货架期及促销手段等特征与销量间也存在一定的相关关系.而  $\text{day\_of\_week\_1}$ 、 $\text{day\_of\_week\_3}$ 、 $\text{day\_of\_week\_5}$ 、 $\text{day\_of\_week\_6}$ 、 $\text{day\_of\_month}$  等 5 个特征的互信息估计量为 0,因此在后续预测中不予考虑.

## 3) 特征降维

进一步通过 KMO 和 Bartlett 检验发现,KMO 值为 0.827 且 Bartlett 球形度检验的  $p$  值小于 0.01.综合两个指标,说明变量间存在相关性,适合进行因子分析.利用 PCA 对特征矩阵进行主成分提取后,再次观察特征间的相关系数矩阵(图 14),可直观地发现各个特征间的相关性很小,且对目标变量的解释程度较高,说明经特征筛选和降维得到的特征矩阵有效地表达了初始特征集合所蕴含的信息,以上特征处理过程合理有效,后续将针对每个预测样本按上述步骤进行特征处理.

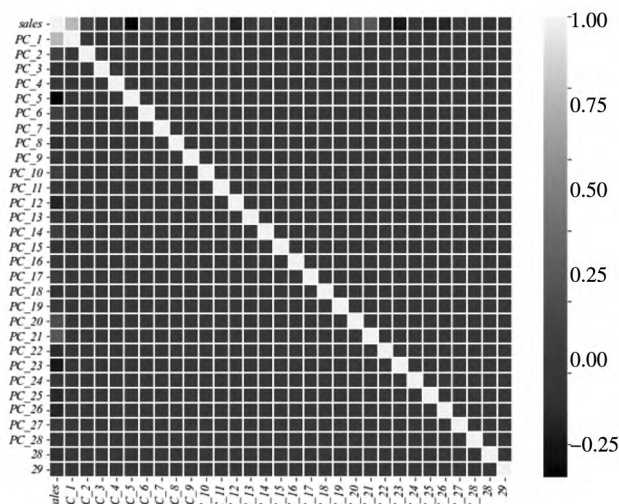


图 14 经过滤及降维处理的特征集合相关系数矩阵

Fig. 14 Correlation coefficient matrix of the feature set after filtering and dimensional reduction

## 3.2 ARIMA-NARX 销量组合预测模型实现

考虑到本研究的多个门店之间在预测环节无相关性,因此不失一般性,以下将以  $\text{store\_id} = 0$ 、 $\text{sku\_id} = 1$  为例,呈现 ARIMA-NARX 组合预测模型在 MATLAB 2018b 中的实现过程.

### 3.2.1 预测过程

1) 平稳性检验:观察预测目标商品销量时序图(图 15),发现数据没有明显的趋势性波动,可初步判断该时间序列具有平稳性;进一步,经单位根 ADF 检验,得到  $p$  值为 0 且  $t$  统计量为  $-2.9351$ ,小于显著性为 0.05% 的临界值  $-1.9423$ ,表明至少在 99.5% 的置信水平下可以拒绝原假设(序列存在单位根),因此认为该时间序列通过平稳性检验,无需差分处理,即  $d = 0$ .

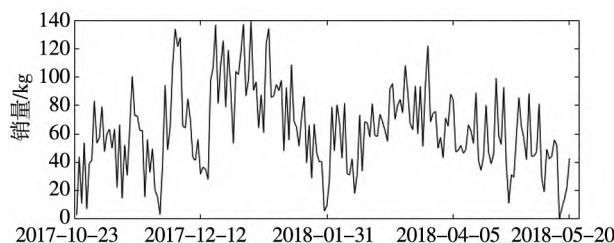


图 15 香蕉销量时间序列图 ( $\text{store\_id} = 0$ ,  $\text{sku\_id} = 1$ )

Fig. 15 Time series of banana sales ( $\text{store\_id} = 0$ ,  $\text{sku\_id} = 1$ )

2) ARIMA 模型识别:对序列相关系数图进行分析,从图 16 可以看出,ACF 和 PACF 均为拖尾,应建立  $\text{ARMA}(p, q)$  模型.根据 ACF 可基本确定  $q$  取值范围为  $[1, 5]$ ,由 PACF 确定  $p$  的取值范围为  $[1, 2]$ .

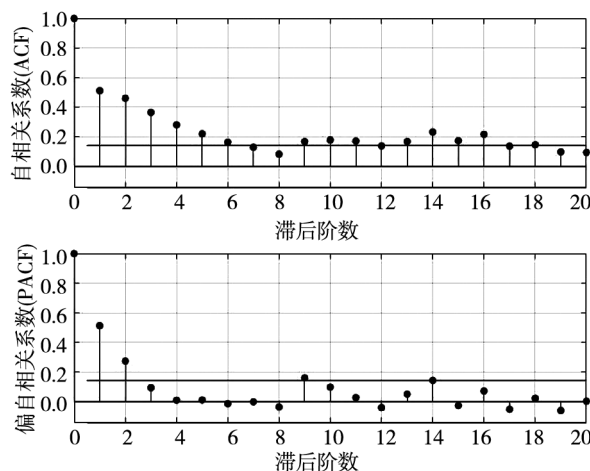


图 16 自相关系数与偏自相关系数

Fig. 16 Autocorrelation coefficient and partial correlation coefficient

3) ARIMA 模型选择: 根据确定的  $p, q$  范围, 采用循环遍历的方法定阶, 对比不同阶数 BIC 值 (表 5), 发现当  $p=1, q=1$  时 BIC 达到最小, 因此确定最佳模型形式为 ARMA (1, 1), 即  $(1 - \varphi_1 L)y_t = c + (1 + \theta_1 L)u_t$ ; 进一步估计模型参数并进行显著性检验 (表 6),  $p$  值均小于 0.05, 说明所得系数解释程度较高。

表 5 模型选择——BIC 值 ( $\times 10^3$ )Table 5 Model selection basis: BIC value ( $\times 10^3$ )

	$q=1$	$q=2$	$q=3$	$q=4$	$q=5$
$p=1$	1.821 2	1.826 2	1.831 1	1.832 6	1.837 1
$p=2$	1.825 8	1.830 9	1.836 2	1.834 1	1.839 3

表 6 ARIMA 模型参数显著性检验

Table 6 Parameter significance test of ARIMA model

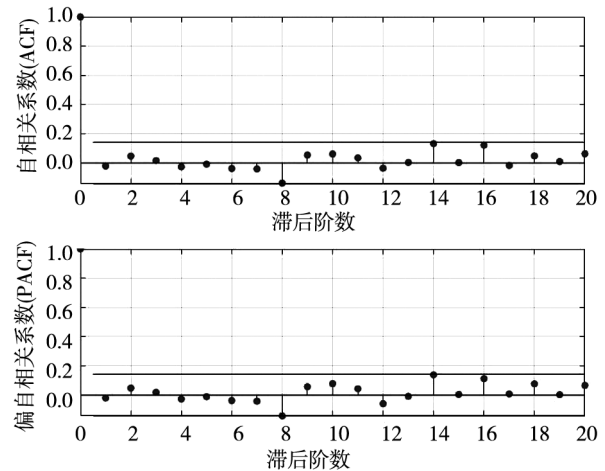
参数	值	标准误差	$t$ 统计量	$p$ 值
$c$	9.745 1	3.522 8	2.766 3	0.005
AR(1)	0.847 2	0.054 8	15.462 2	0.000
MA(1)	-0.483 8	0.095 5	-5.063 9	0.000

4) ARIMA 残差检验及输出: 参数估计之后, 需要对模型拟合残差序列进行独立性和正态分布检验. 图 17(A) 中, 残差 ACF 和 PACF 值均落在置信区间范围内, 独立性检验通过; 图 17(B) 中, 残差序列基本与标准正态分布一致, 符合正态分布. 利用通过显著性检验和残差检验的 ARMA (1, 1) 模型进行预测, 得到残差序列如图 17(C) 所示。

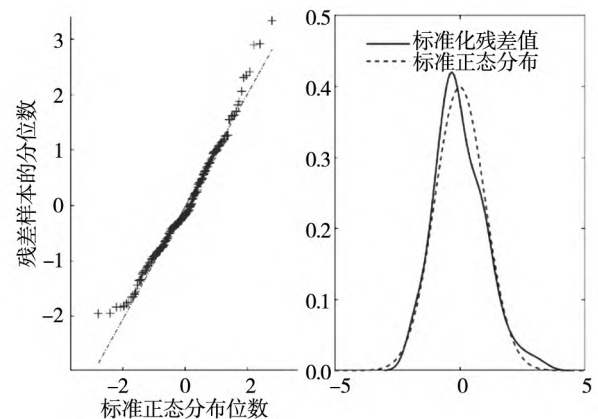
5) NARX 网络创建: 以残差序列作为输出数据, 经特征工程处理的特征矩阵作为输入数据, 经反复调参测试, 设置神经元个数为 20、延迟阶数为 7, 构建 Open-Loop NARX 网络结构如图 18 所示。

6) NARX 网络训练: 通过对比拟合效果和稳定性, 选择 Levenberg-Marquardt 算法进行网络训练, 训练过程交叉验证结果如图 19. 在训练 18 次之后, 验证集误差水平不再降低, 证明训练可以结束. 此时验证集 MSE 为 0.161 68. 拟合效果如图

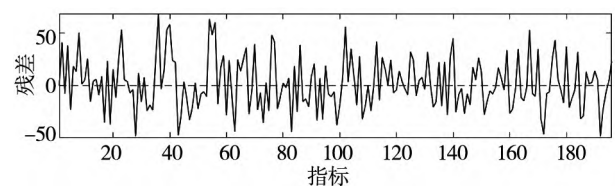
20 所示, 训练集拟合度为 0.954 97, 训练集、验证集及测试集综合拟合度为 0.861 63.



(A) 独立性检验



(B) 正态分布检验



(C) ARIMA 预测残差

图 17 ARIMA 预测残差检验及残差序列

Fig. 17 Prediction residual test and residual sequence of ARIMA model

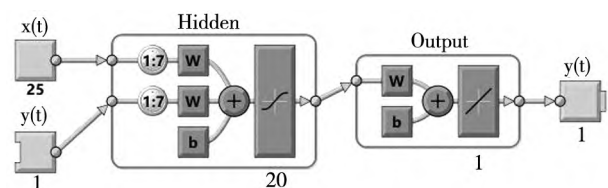


图 18 NARX 神经网络结构 (Open-Loop)

Fig. 18 Structure of NARX neural network (Open-Loop)

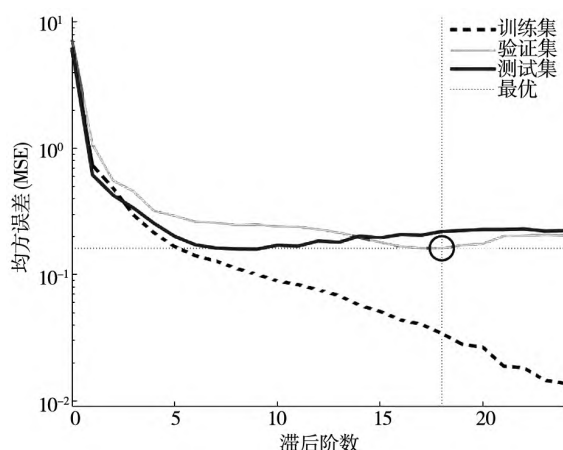


图 19 NARX 神经网络训练过程

Fig. 19 Training process of NARX neural network

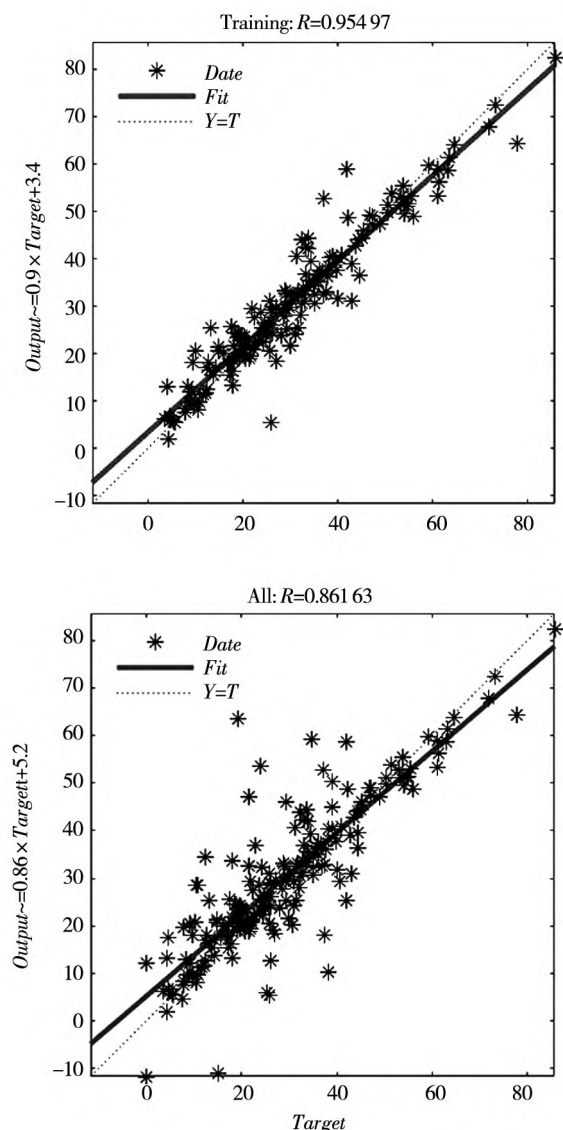


图 20 NARX 神经网络残差训练结果拟合度

Fig. 20 Fitting degree of residual training by NARX

7) NARX 残差预测: 由于在 NARX 网络训练过程中训练目标 (Target) 已知, 可以向输入端提供的是正确的历史观测值, 因此可以建立 Open-Loop 模式, Open-Loop 是单纯的前馈神经网络, 可以直接使用静态神经网络的建模函数. 但 Open-Loop 只能向前预测一步, 为进行多步预测, 网络训练完成后将转换成 Close-Loop 模式进行预测 (图 21), 对残差序列进行预测. 用残差预测结果修正 ARIMA 预测结果, 得到 ARIMA 与 ARIMA-NARX 组合预测模型拟合对比曲线 (图 22). 在  $store\_id = 0, sku\_id = 1$  的预测中, ARIMA-NARX 预测结果整体 MSE 为 2.89, MAPE 为 8.65%, 而 ARIMA 两项指标分别为 5.34 和 14.57%, 与 ARIMA 对比, 可以发现组合模型能够较好地捕捉销量变化.

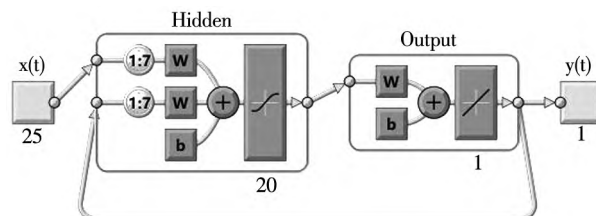


图 21 NARX 神经网络结构 (Closed-Loop)

Fig. 21 Structure of NARX neural network (Closed-Loop)

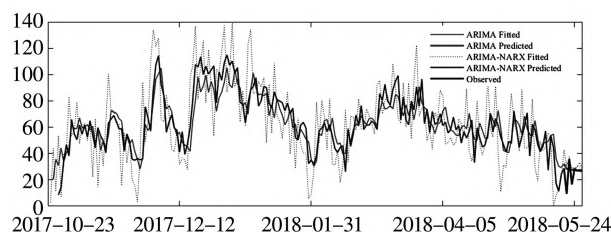


图 22 ARIMA 及 ARIMA-NARX 模型预测效果对比

(store\_id = 0, sku\_id = 1)

Fig. 22 Comparison of the prediction effect of ARIMA and

ARIMA-NARX model

### 3.2.2 预测结果分析

若在 NARX 预测过程中不考虑外部因素, 即仅依据序列自身进行预测, 则称为 NAR 神经网络, 按照残差补偿思想可构建 ARIMA-NAR 组合模型, 将 ARIMA-NARX 预测结果与 ARIMA-NAR 对比, 可直接反映特征矩阵的作用. 同时, 为说明 ARIMA-NARX 组合模型的预测水平, 采用 SVM 和回归决策树 (regression tree, RT) 进行对比, 二者在生鲜商品销量预测或时间序列预测方面有不少应用, 其中前者在小样本训练集上具有较好的



泛化能力,但核函数求解困难,后者预测结果可解释性高,但容易造成过拟合.图 23 呈现了 ARIMA-NARX 与 ARIMA、NARX、ARIMA-NAR、SVM

以及 RT 等预测方法的预测结果,预测目标包括香蕉品类( $sku\_id = 1 \sim 2$ )未来 7 天在各个门店( $store\_id = 0 \sim 6$ )的每日销量.

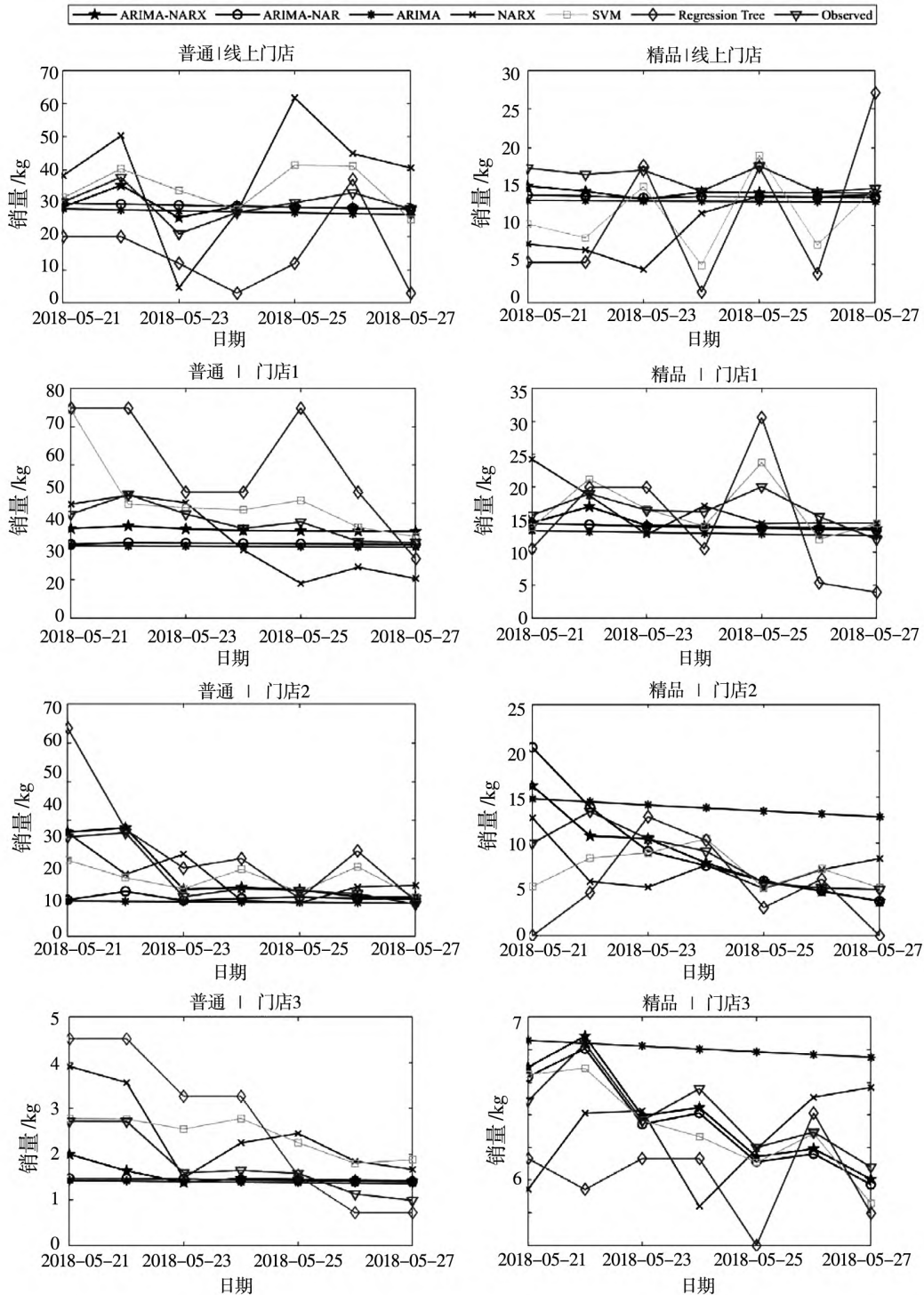
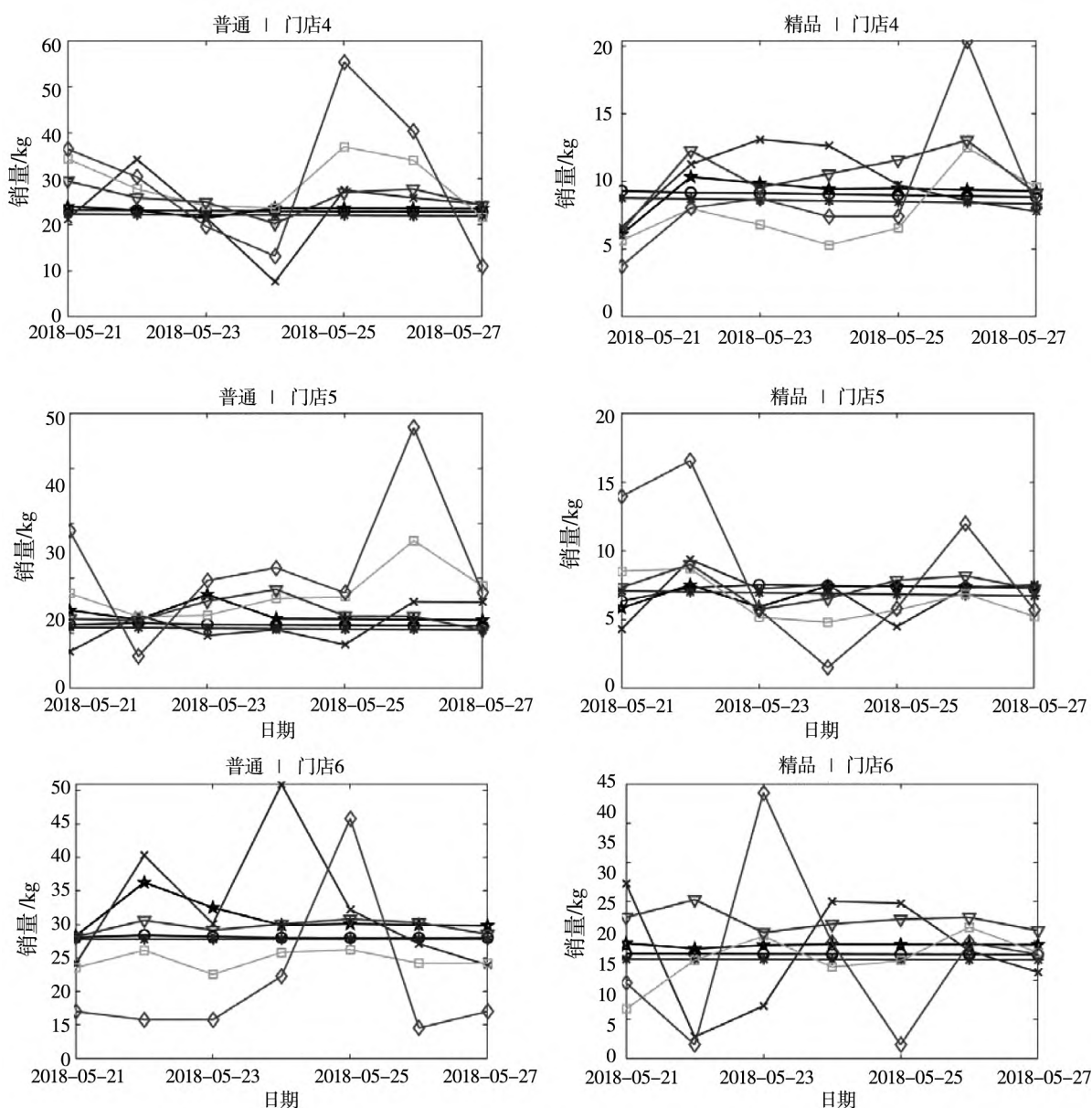


图 23 香蕉各 SKU 在各门店的销量预测结果对比

Fig. 23 Comparison of prediction result of sales trend for each banana SKU in each store



续图 23

Fig. 23 Continues

## 1) 不同预测方法预测效果对比

利用 MSE 和 MAPE 两个指标对模型预测误差进行评价(表 7),并基于上述两个指标对 ARIMA-NARX 与其他模型两两进行 DM 检验(表 8),发现检验结果均显著拒绝原假设,即 ARIMA-NARX 预测能力优于其他五种模型.按照 MSE 平均值对 ARIMA-NARX、ARIMA-NAR、ARIMA、SVM 及 RT 等五种模型预测效果从优到劣进行排序,得到 ARIMA-NARX > ARIMA-NAR > ARIMA > SVM > NARX > RT;按照 MAPE 平均值排序,得到

ARIMA-NARX > ARIMA-NAR > SVM > ARIMA > NARX > RT. 相比其他方法,ARIMA-NARX 模型在 MSE 和 MAPE 的表现均有较大优势.其中,对比 ARIMA 及 NARX 模型预测结果,ARIMA-NARX 预测精度得到大幅提升,说明 NARX 神经网络能够较好地解释销量数据序列中的非线性关系,NARX 与 ARIMA 模型存在互补性.对比 ARIMA-NAR,ARIMA-NARX 平均 MAPE 降低了 4.51 个百分点,表明特征矩阵的加入对于提高 NARX 的预测精度具有明显的作用.

表 7 模型预测误差评价  
Table 7 Evaluation model prediction error

sku	store	评价指标	预测方法					
			ARIMA-NARX	ARIMA	NARX	ARIMA-NAR	SVM	RT
1	0	MSE	2.89	5.34	15.90	4.87	7.38	17.27
		MAPE/%	8.65	14.57	45.88	12.78	21.39	53.70
	1	MSE	4.10	7.42	7.93	6.83	10.87	18.83
		MAPE/%	7.45	13.29	14.50	11.67	14.73	35.13
	2	MSE	1.26	9.40	6.41	8.53	6.01	12.19
		MAPE/%	8.99	31.94	40.37	27.80	31.01	54.02
	3	MSE	5.43	7.28	7.82	7.04	7.48	13.24
		MAPE/%	14.85	17.72	26.25	17.19	26.52	37.01
	4	MSE	3.64	4.47	6.69	3.83	5.13	13.57
		MAPE/%	13.06	15.37	20.82	13.20	16.22	43.21
	5	MSE	2.27	3.59	5.20	3.09	6.51	14.98
		MAPE/%	10.91	18.70	34.75	15.38	39.32	83.43
	6	MSE	2.55	2.11	9.09	1.82	5.05	13.03
		MAPE/%	5.57	6.23	21.39	5.16	16.79	42.98
2	0	MSE	2.27	3.21	7.36	2.70	6.13	10.09
		MAPE/%	10.43	17.59	34.42	13.93	32.19	55.81
	1	MSE	2.92	4.17	4.27	3.41	2.53	7.06
		MAPE/%	14.51	20.75	20.00	16.98	14.18	39.86
	2	MSE	2.64	5.91	3.96	4.03	2.82	5.57
		MAPE/%	18.67	84.02	37.95	24.54	23.69	52.24
	3	MSE	0.53	2.26	2.14	0.54	0.83	2.40
		MAPE/%	11.81	62.04	44.75	13.38	17.74	51.84
	4	MSE	1.69	2.80	2.49	2.38	3.38	3.92
		MAPE/%	10.94	23.81	19.37	19.44	25.32	31.05
	5	MSE	0.94	1.13	1.89	1.13	1.47	4.58
		MAPE/%	10.17	12.85	20.74	13.60	18.16	49.23
	6	MSE	3.48	5.19	8.21	4.52	6.15	12.04
		MAPE/%	11.28	18.01	23.22	15.51	17.64	35.59
	Min	MSE	0.53	1.13	0.15	0.54	0.83	2.40
		MAPE/%	5.57	6.23	14.50	5.16	14.18	31.05
	Max	MSE	5.43	9.40	15.90	8.53	10.87	18.83
		MAPE/%	18.67	84.02	44.75	27.80	39.32	83.43
	Average	MSE	2.62	4.59	6.38	3.91	5.12	10.63
		MAPE/%	11.24	25.49	28.89	15.75	22.49	47.51

表 8 ARIMA-NARX 与其他 5 种模型的 DM 检验 (DM 统计量)

Table 8 DM test of ARIMA-NARX with the other 5 models (DM statistics)

	ARIMA	NARX	ARIMA-NAR	SVM	RT
MSE	2.895 4 ***	2.702 7 ***	2.283 2 **	3.410 9 **	4.627 **
MAPE	1.840 9 **	4.595 2 **	2.920 7 ***	3.978 7 **	5.743 2 *

注: \*\*\* 为 1% 显著, \*\* 为 5% 显著, \* 为 10% 显著。

## 2) 细分预测目标预测效果对比

综合 6 种预测模型共 588 个预测点, 进一步

分析不同预测目标下的预测效果。

①在不同 SKU 的预测效果: 如图 24 所示, 观



察两种香蕉 SKU 的预测结果,发现模型对普通等级( $sku\_id = 1$ )的预测效果普遍优于精品等级( $sku\_id = 2$ ),结合前文数据探索可知普通等级的销量规模显著大于精品等级,规律性更稳定可能是使得其预测效果更优的原因。

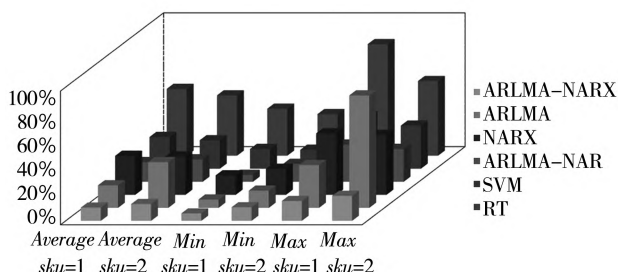


图 24 不同 SKU 预测效果对比 (MAPE)

Fig. 24 Comparison of prediction effects of different SKUs (MAPE)

②在不同门店的预测效果:如图 25 所示,观察不同门店的预测结果,发现模型对  $store\_1$  和  $store\_6$  的预测效果普遍更优,而对  $store\_2$  和  $store\_5$  的预测效果普遍欠佳。结合前文数据探索可知,  $store\_1$  和  $store\_6$  销量规模属于 H 类,而  $store\_2$  和  $store\_5$  属于 L 类, H 类顾客流大,总体销量更稳定,可能是使得其预测效果更优的原因。而  $store\_0$  属于线上销量,尽管销量规模较大,但面向人群随机性更大,购买行为不确定性较线下大,使得其销量预测误差相对较大。

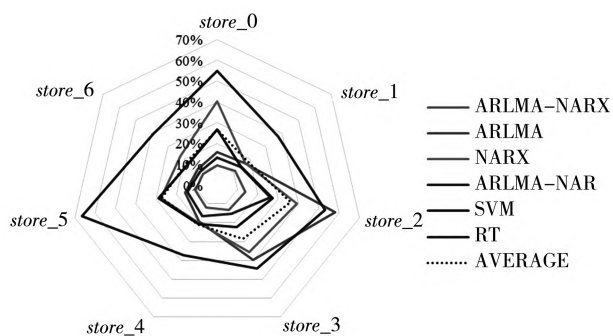


图 25 不同门店的预测效果 (MAPE)

Fig. 25 Comparison of prediction effects at different stores (MAPE)

③预测周期内不同日期的预测效果:如图 26 所示,观察模型对一个周期(2018 年 5 月 21 日 ~ 2018 年 5 月 27 日)的预测效果,发现所有模型在该周期星期三(2018 年 5 月 23 日)的平均预测精确度偏高,而对该周期内星期六(2018 年 5 月 26 日)的平均预测精度偏低,结合前文数据探索可知(图 12),该预测结果与香蕉销量的分布在历史

周期内当天的集中或分散程度相关。例如,历史星期三的香蕉销量相对稳定集中,则预测周期内周三的预测精度较高,历史数据中星期六的销量相对随机分散,则该预测周期内的周六预测精度较低。该结果说明历史销售数据中蕴含的规律参考性更高的日期,使得其未来的预测效果也更优。

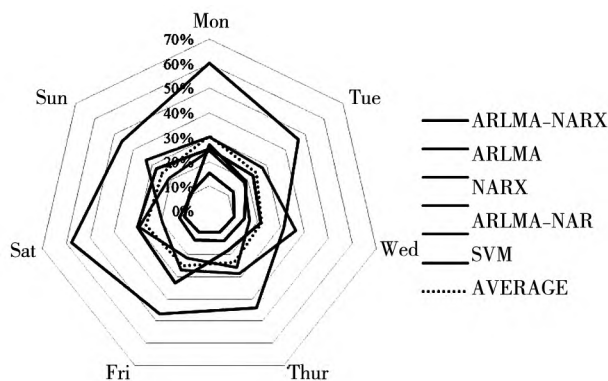


图 26 以星期为周期的预测效果对比 (MAPE)

Fig. 26 Comparison of prediction effect based on the week cycle (MAPE)

## 4 结束语

由于生鲜农产品具有易腐性、易损性等特性,且呈现出小批量、高频次的销售特点,使得销量趋势伴随更多不确定性。本研究以影响因素挖掘和预测模型改进为切入点,提出生鲜商品短期销量的预测方案。主要结论如下。

第一,基于时间周期性因素、商品自身的营销策略、可替代品的营销策略及商品自身特性等方面提取影响因子,通过特征衍生、筛选和降维处理,能够获得有效的特征矩阵,帮助模型理解数据关系,提高模型预测效果。

第二,ARIMA 和 NARX 模型在预测线性和非线性关系上优势互补,DM 检验结果显示,与 ARIMA、NARX、ARIMA-NAR、SVM 及 RT 等相比,通过残差补偿思想构建的 ARIMA-NARX 组合模型在生鲜商品短期销量预测中表现出更优的预测精度。

第三,生鲜商品短期销量预测的精准度与销售数据的规模、离散程度有关。比如客流量较大的门店、需求较大的商品以及一周中销量较为稳定

的时间,顾客群整体的购买行为较平稳,反映在销量上异常值较少,预测的准确度也相应较高。

当然,本研究分析仍存在一些不足的地方,比

如在新鲜度等部分影响因子指标刻画上不够精细,模型在实际应用中的可操作性需进一步提高,希望在下一步研究中继续关注改进。

## 参 考 文 献:

- [1]邓 涛,陈 玲,陈晗曦,等. 以“盒马鲜生”为例基于 GS1 的生鲜产品冷链供应链研究[J]. 中国自动识别技术, 2018, (4): 69–73.  
Deng Tao, Chen Ling, Chen Hanxi, et al. Research on cold chain supply chain of fresh products based on GS1 (taking “Freshhema” as an example)[J]. China Auto-ID, 2018, (4): 69–73. (in Chinese)
- [2]陈 军,但 斌,曹群辉,等. 短保质期变质产品的两次订货策略研究[J]. 管理科学学报, 2009, 12(3): 83–91.  
Chen Jun, Dan Bin, Cao Qunhui, et al. Ordering policy with two ordering opportunities for deteriorating items with short shelf life[J]. Journal of Management Sciences in China, 2009, 12(3): 83–91. (in Chinese)
- [3]郑宇婷,李建斌,陈植元,等. 不确定需求下的冷链分销商最优决策[J]. 管理科学学报, 2019, 22(1): 99–111.  
Zheng Yuting, Li Jianbin, Chen Zhiyuan, et al. Optimal decisions of cold chain distributor under uncertain in demand[J]. Journal of Management Sciences in China, 2019, 22(1): 99–111. (in Chinese)
- [4]王 帅,汤 铃,余乐安. 基于单变量分解集成的牛奶消费需求预测研究[J]. 系统科学与数学, 2013, 33(1): 11–19.  
Wang Shuai, Tang Ling, Yu Lean. Univariate decompose-ensemble method based milk demand forecasting[J]. Journal of System Science and Mathematical Science, 2013, 33(1): 11–19. (in Chinese)
- [5]Sagaert Y R, Aghezzaf E, Kourentzes N, et al. Tactical sales forecasting using a very large set of macroeconomic indicators [J]. European Journal of Operational Research, 2018, 264(2): 558–569.
- [6]Loureiro A L D, Miguéis V L, Da Silva L F M. Exploring the use of deep neural networks for sales forecasting in fashion retail[J]. Decision Support Systems, 2018, (114): 81–93.
- [7]Puchalsky W, Ribeiro G T, Da Veiga C P, et al. Agribusiness time series forecasting using wavelet neural networks and metaheuristic optimization: An analysis of the soybean sack price and perishable products demand[J]. International Journal of Production Economics, 2018, (203): 174–189.
- [8]张 钠,殷 哲,吕 飞,等. 面向连锁零售业的时间序列预测方法研究[J]. 工业工程与管理, 2014, 19(3): 60–65, 71.  
Zhang Na, Yin Zhe, Lü Fei, et al. Research on time series forecasting for the retail chain[J]. Industrial Engineering and Management, 2014, 19(3): 60–65, 71. (in Chinese)
- [9]Grigorios T. A survey of machine learning techniques for food sales prediction[J]. Artificial Intelligence Review, 2019, 52(1): 441–447.
- [10]Arunraj N S, Ahrens D. A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting[J]. International Journal of Production Economics, 2015, (170): 321–335.
- [11]Arunraj N S, Ahrens D, Fernandes M. Application of SARIMAX model to forecast daily sales in food retail industry[J]. International Journal of Operations Research and Information Systems, 2016, 7(2): 1–21.
- [12]刘 晶,和述群,朱清香,等. 基于深度学习的线上农产品销量预测模型研究[J]. 计算机应用研究, 2017, 34(8): 2291–2293, 2344.  
Liu Jing, He Shuqun, Zhu Qingxiang, et al. Online agricultural product sales prediction model research based on deep learning[J]. Application Research of Computers, 2017, 34(8): 2291–2293, 2344. (in Chinese)
- [13]Dellino G, Laudadio T, Mari R, et al. Microforecasting methods for fresh food supply chain management: A computational

- study[J]. *Mathematics and Computers in Simulation*, 2018, (147): 100–120.
- [14] Huber J, Gossmann A, Stuckenschmidt H. Cluster-based hierarchical demand forecasting for perishable goods[J]. *Expert Systems with Applications*, 2017, (76): 140–151.
- [15] Wang J, Wang Z, Li X, et al. Artificial bee colony-based combination approach to forecasting agricultural commodity prices[J]. *International Journal of Forecasting*, 2022, 38(1): 21–34.
- [16] Lee W, Chen C W, Chen K H, et al. A comparative study on the forecast of fresh food sales using logistic regression, moving average and BPnn methods[J]. *Journal of Marine Science and Technology*, 2012, 20(2): 142–152.
- [17] 徐 琪, 刘 峥. 基于 SVM 的短生命周期产品供应链双渠道需求预测模型[J]. *系统管理学报*, 2014, 23(2): 255–262, 270.
- Xu Qi, Liu Zheng. SVM-based dual-channel demand forecast model of perishables in the supply chain[J]. *Journal of Systems & Management*, 2014, 23(2): 255–262, 270. (in Chinese)
- [18] 杨海民, 潘志松, 白 玮. 时间序列预测方法综述[J]. *计算机科学*, 2019, 46(1): 21–28.
- Yang Haimin, Pan Zhisong, Bai Wei. Review of time series prediction methods[J]. *Computer Science*, 2019, 46(1): 21–28. (in Chinese)
- [19] 胡 浩, 闫 伟, 李泓明. 基于组合预测方法的城市道路短时交通流预测[J]. *工业工程与管理*, 2019, 24(3): 107–115.
- Hu Hao, Yan Wei, Li Hongming. Short-term traffic flow prediction of urban road based on combination forecasting method[J]. *Industrial Engineering and Management*, 2019, 24(3): 107–115. (in Chinese)
- [20] 耿 宏, 任道先, 杜 鹏. 基于 NARX 神经网络航空发动机参数动态辨识模型[J]. *计算机工程与应用*, 2017, 53(12): 241–248.
- Geng Hong, Ren Daoxian, Du Peng. Dynamic parameter identification model of aircraft engine based on NARX neural network[J]. *Computer Engineering and Application*, 2017, 53(12): 241–248. (in Chinese)
- [21] 谢吉洋, 闫 冬, 谢 垚, 等. 基于 NARX 神经网络的热负荷预测中关键影响因素分析[J]. *计算机应用*, 2018, 38(11): 3180–3187.
- Xie Jiyang, Yan Dong, Xie Yao, et al. Analysis of key factors in heat demand prediction based on NARX neural network[J]. *Journal of Computer Applications*, 2018, 38(11): 3180–3187. (in Chinese)
- [22] Koschwitz D, Frisch J, Treeck C. Data-driven heating and cooling load predictions for non-residential buildings based on support vector machine regression and NARX Recurrent Neural Network: A comparative study on district scale[J]. *Energy*, 2018, (165): 134–142.
- [23] Wunsch A, Liesch T, Broda S. Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX)[J]. *Journal of Hydrology*, 2018, (567): 743–758.
- [24] Lee W J, Na J, Kim K, et al. NARX modeling for real-time optimization of air and gas compression systems in chemical processes[J]. *Computers and Chemical Engineering*, 2018, (115): 262–274.
- [25] He K, Yu L, Tang L. Electricity price forecasting with a BED (Bivariate EMD Denoising) methodology[J]. *Energy*, 2015, (91): 601–609.
- [26] Tang L, Yu L, He K. A novel data-characteristic-driven modeling methodology for nuclear energy consumption forecasting[J]. *Applied Energy*, 2014, (128): 1–14.
- [27] 章杰宽. 智能组合预测方法及其应用[J]. *中国管理科学*, 2014, 22(3): 26–33.
- Zhang Jiekuan. Intelligent integration forecasting method and its application[J]. *Chinese Journal of Management Science*, 2014, 22(3): 26–33. (in Chinese)
- [28] 孙轶轩, 邵春福, 计 寻, 等. 基于 ARIMA 与信息粒化 SVR 组合模型的交通事故时序预测[J]. *清华大学学报(自然科学版)*, 2014, 54(3): 348–353, 359.



- Sun Yixuan, Shao Chunfu, Ji Xun, et al. Urban traffic accident time series prediction model based on combination of ARI-MA and information granulation SVR. *Journal of Tsinghua University (Science & Technology)*, 2014, 54(3): 348–353, 359. (in Chinese)
- [29] 张立杰, 寇纪淞, 李敏强, 等. 基于自回归移动平均及支持向量机的中国棉花价格预测[J]. *统计与决策*, 2013, 378(6): 30–33.
- Zhang Lijie, Kou Jisong, Li Minqiang, et al. Forecast of cotton price in China based on autoregressive moving average and support vector machine[J]. *Statistics & Decision*, 2013, 378(6): 30–33. (in Chinese)
- [30] 唐小我, 曾 勇. 组合预测误差校正模型的应用分析[J]. *管理科学学报*, 2002, 5(6): 53–64.
- Tang Xiaowo, Zeng Yong. Application of error correction models of combination forecasting[J]. *Journal of Management Sciences in China*, 2002, 5(6): 53–64. (in Chinese)
- [31] 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, 74.
- Ni Dongmei, Zhao Qiuhong, Li Haibin. Synthesized model for demand forecasting and its integration with inventory decision [J]. *Journal of Management Sciences in China*, 2013, 16(9): 44–52, 74. (in Chinese)
- [32] 毛 勇, 周晓波, 夏 铮, 等. 特征选择算法研究综述[J]. *模式识别与人工智能*, 2007, 20(2): 211–218.
- Mao Yong, Zhou Xiaobo, Xia Zheng, et al. A survey for study of feature selection algorithms[J]. *PR & AI*, 2007, 20(2): 211–218. (in Chinese)
- [33] 李郅琴, 杜建强, 聂 斌, 等. 特征选择方法综述[J]. *计算机工程与应用*, 2019, 55(24): 10–19.
- Li Zhiqin, Du Jianqiang, Nie Bin, et al. Summary of feature selection methods[J]. *Computer Engineering and Applications*, 2019, 55(24): 10–19. (in Chinese)
- [34] 袁 轶, 王新房. 一种基于方差的文本特征选择算法[J]. *计算机工程*, 2012, 38(12): 155–157.
- Yuan Yi, Wang Xinfang. Text feature selection algorithm based on variance[J]. *Computer Engineering*, 2012, 38(12): 155–157. (in Chinese)
- [35] 范雪莉, 冯海泓, 原 猛. 基于互信息的主成分分析特征选择算法[J]. *控制与决策*, 2013, 28(6): 915–919.
- Fan Xueli, Feng Haihong, Yuan Meng. PCA based on mutual information for feature selection[J]. *Control and Decision*, 2013, 28(6): 915–919. (in Chinese)
- [36] 张 尧. 基于互信息的特征选择方法研究[D]. 西安: 西安理工大学, 2019.
- Zhang Yao. Study on Feature Selection Based on Mutual Information[D]. Xi'an: Xi'an University of Technology, 2019. (in Chinese)
- [37] Ezzeldin R, Hatata A. Application of NARX neural network model for discharge prediction through lateral orifices[J]. *Alexandria Engineering Journal*, 2018, 57(4): 2991–2998.
- [38] 王 楠, 侯铁珊. 基于 NARX 网络的人民币汇率预测研究[J]. *东北大学学报(社会科学版)*, 2015, 17(1): 32–37.
- Wang Nan, Hou Tieshan. Research of RMB exchange rate forecasting based on the NARX model[J]. *Journal of Northeastern University (Social Science)*, 2015, 17(1): 32–37. (in Chinese)
- [39] Bergmeir C, Benítez J M. On the use of cross-validation for time series predictor evaluation[J]. *Information Sciences*, 2012, (191): 192–213.
- [40] 徐贤浩, 宋奇志. 短生命周期产品的预测方法[J]. *统计与决策*, 2006, (24): 161–163.
- Xu Xianhao, Song Qizhi. Shortlife cycle product forecasting method[J]. *Statistics & Decision*, 2006, (24): 161–163. (in Chinese)
- [41] 赵学斌, 李大学, 谢名亮. 短生命周期产品的销量预测模型研究[J]. *计算机工程与设计*, 2010, (11): 115–118.
- Zhao Xuebin, Li Daxue, Xie Mingliang. Study on sales forecasting models of short lifecycle products[J]. *Computer Engineering and Design*, 2010, (11): 115–118.

- neering and Design, 2010, (11): 115 – 118. (in Chinese)
- [42] Chung C, Niu S C, Sriskandarajah C. A sales forecast model for short-life-cycle products: New releases at blockbuster[J]. Operations Research, 2013, 53(5–6): 521 – 522.
- [43] Diebold F X, Mariano R S. Comparing predictive accuracy comparing predictive accuracy[J]. Journal of Business and Economic Statistics, 1995, 3(13): 253 – 263.

## Combination forecasting of short-term sales for fresh products based on feature fusion

*XU Xiao-feng*<sup>1</sup>, *YU Le-an*<sup>2\*</sup>, *LIN Zi-ru*<sup>1</sup>, *SUN Yu-ping*<sup>1</sup>

1. School of Economics and Management, China University of Petroleum, Qingdao 266580, China;
2. School of Business, Sichuan University, Chengdu 610065, China

**Abstract:** Due to the short shelf life and perishable characteristics, short-term fresh products sales forecast requires high accuracy and reliability. This paper carries out a feature engineering analysis for fresh products involving time, pricing, competitive product pricing, freshness and other micro-level factors. On this basis, an ARIMA-NARX combination forecasting model of fresh product sales is proposed. This combination forecasting model takes advantage of the ARIMA to capture the linear rule in the sales time series, and adopts the NARX to describe the nonlinear relationship in the ARIMA residual with the feature matrix processed by feature creation and dimension reduction. Then, the NARX residual prediction result is used to correct the predicted sales of ARIMA. Finally, the prediction results of the combined model are compared with the real observation values and prediction results of ARIMA, NARX, ARIMA-NAR, SVM and RT models. The MSE/MAPE value and DM test verify the rationality and effectiveness of the ARIMA-NARX model, which could improve the prediction accuracy of short-term fresh product sales significantly.

**Key words:** fresh product; ARIMA-NARX; combination forecasting; feature engineering