

文章编号: 2095-2163(2023)07-0150-05

中图分类号: F272.1

文献标志码: A

基于 SVM 算法的生鲜农产品质量安全预警研究

李永飞^{1,2}, 贺桂英^{1,2}, 张金^{1,2}

(1 西安邮电大学 现代邮政学院, 西安 710061; 2 邮政行业技术研发中心, 西安 710061)

摘要: 为有效预警生鲜农产品质量突发事件,最大限度地加强生鲜农产品质量安全管理,提高农产品冷链效益。本文从供应链角度搭建生鲜农产品质量安全预警指标体系,构建基于支持向量机算法的预警模型,同时利用 Matlab 软件下 LIBSVM 工具箱优化支持向量机参数,最后通过实例分析验证生鲜农产品质量安全预警模型对农产品质量预警的可行性和有效性。

关键词: 预警; 质量安全; 生鲜农产品; 支持向量机

Research on quality and safety early warning of fresh agricultural products based on SVM algorithm

LI Yongfei^{1,2}, HE Guiying^{1,2}, ZHANG Jin^{1,2}

(1 School of Modern Posts, Xi'an University of Posts and Telecommunications, Xi'an 710061, China;

2 Technology Research and Development Center of Postal Industry, Xi'an 710061, China)

[Abstract] In order to effectively warn the quality emergencies of fresh agricultural products, the quality and safety management of fresh agricultural products could be strengthened to the greatest extent, and the cold chain benefits of agricultural products could be improved. From the perspective of supply chain, this paper establishes the quality and safety early warning index system of fresh agricultural products, constructs the early warning model based on support vector machine algorithm, and optimizes the parameters of support vector machine by using LIBSVM toolbox under Matlab software. Thereafter, an example is given to verify the feasibility and effectiveness of the quality and safety early warning model of fresh agricultural products for agricultural product quality early warning.

[Key words] early warning; quality safety; fresh agricultural products; Support Vector Machine

0 引言

随着盒马集市、多多买菜、美团优选等社区团购平台的兴起,城乡居民对于高品质且安全的生鲜农产品的消费需求也日趋旺盛,对生鲜农产品的质量安全关注度正不断提高。由于生鲜农产品具有时效性、易腐性、季节性等特点,就使得在质量控制阶段对生鲜农产品的质量安全预警成为极其重要的一个环节^[1-2]。因此,当前形势下适时有效地进行生鲜农产品质量安全预警,有利于降低生鲜农产品腐损率,提升大众的满意度和幸福感,同时对于实现生鲜农产品全程冷链环节的实时监控具有重要意义。

近年来,国内外学者对生鲜农产品质量安全预警的研究成果主要集中在以下几个方面。一是关于

生鲜农产品质量安全预警指标体系的构建。曾欣平等学者^[3]从供应链角度出发,构建乳制品质量安全风险评价指标体系,利用熵权法计算各评价指标的权重系数,然后运用可拓学理论建立乳制品质量安全预警模型。宋宝娥^[4]在供应链视角下,分析农产品供应链各环节,从而建立评价指标体系,构建了基于集值统计加速迭代法的农产品安全预警模型。山红梅等学者^[5]依据生鲜农产品冷链物流特点构建了以农产品加工、低温仓储服务、冷藏运输服务、冷链物流信息化和人员素质五方面为主的生鲜农产品冷链物流综合服务水平评估指标体系,提出了基于云模型-IAHP 的生鲜农产品冷链物流服务水平评估模型,改进了云模型中的云合并算法,对生鲜农产品冷链物流服务水平评估具有实用性。二是生鲜农

基金项目: 国家社会科学基金项目(18XGL001)。

作者简介: 李永飞(1977-),男,博士,副教授,主要研究方向:物流与供应链管理、交通运输规划与管理、质量管理等;贺桂英(1995-),女,硕士研究生,主要研究方向:物流与供应链管理、质量管理;张金(1994-),男,硕士研究生,主要研究方向:物流与供应链管理、质量管理。

收稿日期: 2022-07-20

产品质量安全预警系统的设计。张洋等学者^[6]针对近年来频发的农产品安全事件,依据数据挖掘技术,采用关联规则挖掘和 Dempsters 组合规则的方式,提出了一个农产品供应链风险预警系统,以确保供应链农产品质量。马长路等学者^[7]利用物联网技术,检测生鲜乳中相关指标,当检测到的指标或监控到的参数不在标准范围时,通过系统进行预警。三是生鲜农产品质量安全预警模型的构建。杨玮等学者^[8]分析果蔬冷链过程影响质量的因素,从而建立预警指标体系,利用 BP 神经网络建立果蔬安全预警模型,该模型可以有效预测果蔬在冷链物流中的风险。高亚男等学者^[9]对食品安全相关数据和以往应用的智能化方法进行归纳分析,根据数据特点和已存在方法的不足提出了先验风险概率与模糊层级划分相结合的风险值计算规则,并应用 LightGBM 模型结合专家经验干预策略进行风险值校正和预测。在预警研究中,其他领域应用的预测方法也值得借鉴,张中丹等学者^[10]为了有效预警电网自然灾害风险,研究了基于区块链(block chain, BC)技术的电网灾害预警决策系统。

上述研究成果为生鲜农产品质量安全预警的实践提供了良好的理论基础和方法依据。但是,国内外学者针对生鲜品这类特定农产品的质量安全的文献也较为罕见。因此,本文以生鲜农产品质量安全为研究对象,分析了生鲜农产品在冷链物流过程中的风险因素,进而建立质量安全预警指标体系,构建基于支持向量机算法的预警模型,最后结合数据进行验证分析。

1 构建生鲜农产品质量安全预警指标

从事生鲜农产品质量安全预警的基础性工作是从构建生鲜农产品质量安全预警指标。从供应链角度入手,分析生鲜农产品物流运作的生产、环境、物流、监管等各个环节,依据指标选取原则筛选影响生鲜农产品质量安全的因素,并依据国家相关标准构建预警指标体系。

本文结合相关文献,考虑到生鲜农产品不同于一般农产品的相关特征,构建出生鲜农产品质量安全预警指标体系(见表1)。其中,一级指标共4项,分别为生产因素、环境因素、物流信息和监管过程。二级指标包括保存周期、加工转换率、温度、湿度等8项。通过3位专家的评判打分,运用模糊层次分析法(FAHP)^[11-13]得到表1中的生鲜农产品质量安全预警指标权重。

表1 生鲜农产品质量安全预警指标体系

Tab. 1 Quality and safety early warning index system of fresh agricultural products

一级指标	权重	二级指标	权重	综合权重
生产因素	0.084	保存周期	0.667	0.056
		加工转换率	0.333	0.028
环境因素	0.164	温度	0.750	0.123
		湿度	0.250	0.041
物流信息	0.485	货物破损率	0.570	0.276
		运输时效性	0.430	0.209
监管过程	0.267	质量抽检合格率	0.670	0.179
		农药残留合格率	0.330	0.088

2 建立生鲜农产品质量安全预警模型

支持向量机(Support Vector Machine, SVM)^[14-15]是 Cortes 和 Vapnik 等学者 1995 年首次提出,从模型原理来看,在解决小样本、非线性及高维模式识别中表现出许多特有的优势,并能够推广应用到函数拟合等其他机器学习问题中。在风险预警的研究中,支持向量机算法已得到了广泛应用。具体过程如下:

(1)划分训练集和测试集。训练集用来训练预警模型,测试集用来测试预警模型的准确程度。预警模型选用前40组数据作为训练集,后10组数据作为测试集。

(2)数据标准化处理。标准化处理训练样本,对调整指标范围,减少预警误差率,杜绝选择不当数据具有重要意义。采用极值标准化法,即将属性数据按照比例缩放,使之落入一个小的特定区间,如 $[-1, +1]$ 、 $[0, 1]$ 等,以进一步分析数据的属性。

(3)核函数训练。支持向量机常用的核函数有线性核函数、多项式核函数、RBF核函数以及sigmoid核函数等。在划分训练集后,选择高斯核函数(Radial Basis Function, RBF)进行训练,文中选择高斯核函数是考虑其分类效率较高,适用范围更广。计算公式一般形式可表述为:

$$K(x_i, x_j) = \exp\left\{-\frac{\|x_i - x_j\|^2}{2g^2}\right\} \quad (1)$$

其中, $\|x_i - x_j\|$ 为样本点 x_i 与 x_j 之间的距离, g 是高斯核函数的内核参数。

(4)交叉验证。为了得到更好的安全预警拟合值,提高预警准确度,通过交叉验证(K-fold Cross Validation, K-CV)优化惩罚函数 C 和高斯核函数的参数 g 。

(5) 实际应用。将得到的预测值和真实值进行拟合对比,得到产品各环节回归预测的结果图。由拟合度分析验证预警模型的准确度,利用预警模型对剩余数据进行预警。

3 实例研究

3.1 采集与处理数据

通过对西安某冷链运输公司以及该公司合作的

食品产业园进行实地调查和专家咨询^[16]。针对预警指标评价体系,随机筛选 50 组原始数据(以下取 10 组样本演示),具体资料统计数据 and 专家打分部分结果见表 2。

将收集好的 10 组原始数据进行标准化处理,并结合层次分析法,计算各环节目标结果数据,见表 3。

表 2 苹果冷链物流各环节的原始数据

Tab. 2 Raw data of all links of apple cold chain logistics

样本号	生产因素		环境因素		物流信息		监管过程	
	保存周期/周	加工转换率/%	温度/%	湿度/%	货物破损率/%	运输时效性/%	质量抽检合格率/%	农药残留合格率/%
1	5	54	3.5	88	8	88	91	94
2	23	40	5.1	86	5	76	88	79
3	12	37	4.8	96	7	97	90	90
4	15	43	5.3	94	3	88	94	78
5	13	51	5.2	96	6	76	87	90
6	14	48	6.1	89	8	131	78	97
7	21	50	4.4	77	7	76	86	78
8	8	54	3.5	94	5	88	88	86
9	15	34	5.3	88	3	97	87	87
10	22	56	5.3	96	3	88	90	88

表 3 标准化处理结果

Tab. 3 Standarization results

样本	指标							
	1	2	3	4	5	6	7	8
1	0	0.909 1	0	0.578 9	1.000 0	0.218 2	0.812 5	0.842 1
2	1.000 0	0.272 7	0.615 4	0.473 7	0.400 0	0	0.625 0	0.052 6
3	0.388 9	0.136 4	0.500 0	1.000 0	0.800 0	0.381 8	0.750 0	0.631 6
4	0.555 6	0.409 1	0.692 3	0.894 7	0	0.218 2	1.000 0	0
5	0.444 4	0.772 7	0.653 8	1.000 0	0.600 0	0	0.562 5	0.631 6
6	0.500 0	0.636 4	1.000 0	0.631 6	1.000 0	1.000 0	0	1.000 0
7	0.888 9	0.727 3	0.346 2	0	0.800 0	0	0.500 0	0
8	0.166 7	0.909 1	0	0.894 7	0.400 0	0.218 2	0.625 0	0.421 1
9	0.555 6	0	0.692 3	0.578 9	0	0.381 8	0.562 5	0.473 7
10	0.944 4	1.000 0	0.692 3	1.000 0	0	0.218 2	0.750 0	0.526 3

3.2 LIBSVM 仿真预测

本文将收集的 50 组数据的输入和输出分别采用支持向量机的 LIBSVM 工具箱进行训练预测。选用 RBF 核函数对生鲜农产品各个环节和总体质量状况进行训练,同时采用交叉验证(K-fold Cross Validation, K-CV)优化惩罚函数 C 和高斯核函数的

参数 g, 以提高预测精度^[17-18]。表 4 为运用 K-CV 法得到的各个环节最优参数 *Bestc* 和 *Bestg*。

利用 Matlab 自带的 LIBSVM 工具箱构建 SVM 预测模型,将得到的预测值和真实值进行拟合对比,得到生鲜农产品各环节回归预测的结果,如图 1~图 5 所示。

表 4 K-CV 法获得的各环节最优参数 *Bestc*、*Bestg*

Tab. 4 Optimal parameters *Bestc* and *Bestg* of each link obtained by K-CV method

内容	K-CV 参数优化	
	<i>Bestc</i>	<i>Bestg</i>
生产因素	8	0.031 3
环境因素	8	0.031 3
物流信息	16	0.031 3
监管过程	32	0.031 3
总体预测	8	0.031 3

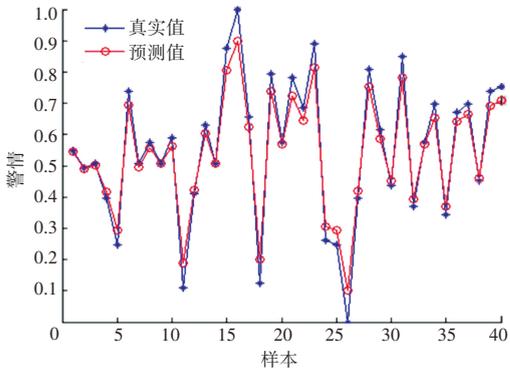


图 1 生产因素回归预测结果图

Fig. 1 Regression prediction results of production factors

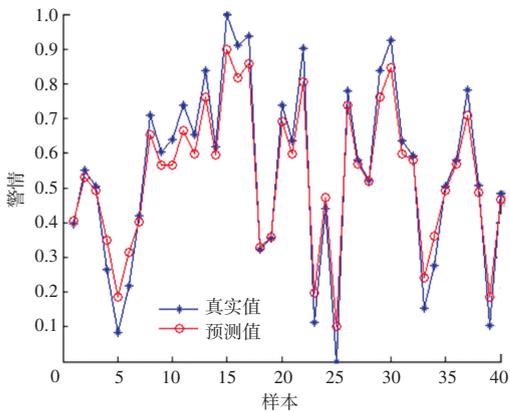


图 2 环境因素回归预测结果图

Fig. 2 Regression prediction results of environmental factors

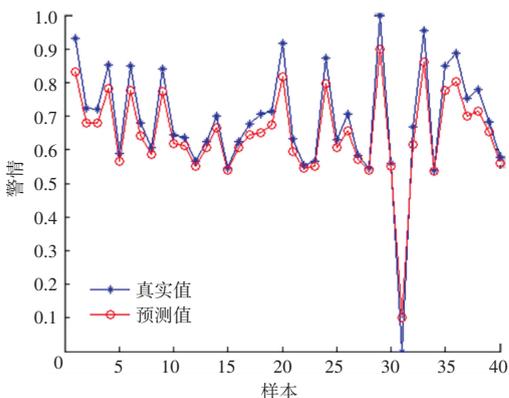


图 3 物流信息回归预测结果图

Fig. 3 Regression prediction results of logistics information

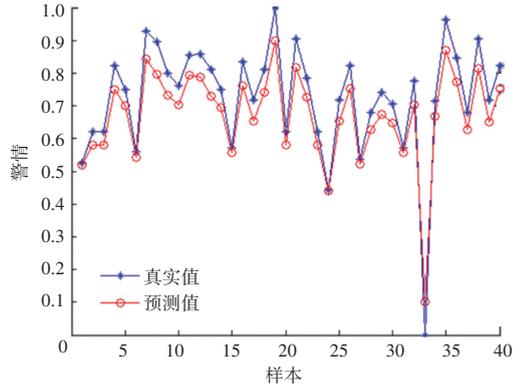


图 4 监管过程回归预测结果图

Fig. 4 Regression prediction results of supervision process

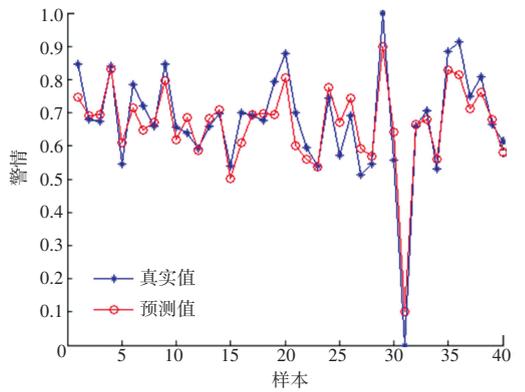


图 5 总体回归预测结果

Fig. 5 Overall regression prediction results

由图 1~图 5 分析可知,训练集的预测输出曲线和真实输出曲线均能较好地拟合,误差值小,结果具有较高的精准性,可运用该模型对剩余数据进行有效预测。

3.3 预警结果分析

为了将预测警情以直观方式显示,可将最终的预测结果值根据不同区间划分为不同的警情等级,结果见表 5。

表 5 警情等级划分表

Tab. 5 Classification of alarm level

预测范围	0~0.2	0.2~0.3	0.3~0.4	0.4~0.5	0.5~1.0
警情	重警	中警	轻警	微警	无警

由拟合度分析可知支持向量机用于生鲜农产品质量安全预警具有可靠性,因此选用该模型对剩余 10 组数据进行预测(取 5 组为演示样本),结果见表 6。

表6 警情预警结果

Tab. 6 Alarm results

环节样本	生产因素	环境因素	物流信息	监管过程
1	0.625 8	0.727 5	0.605 9	0.764 7
	无警	无警	无警	无警
2	0.635 6	0.345 4	0.929 2	0.631 7
	无警	轻警	无警	无警
3	0.363 3	0.672 8	0.688 0	0.576 0
	轻警	无警	无警	无警
4	0.460 9	0.432 2	0.559 2	0.599 8
	微警	微警	无警	无警
5	0.470 8	0.761 7	0.694 4	0.783 4
	微警	无警	无警	无警

输出结果在(0.5~1]范围内属于无警,即农产品处于安全状态下。由表6可以看出大部分样本处于无警范围内,个别样本处于微警范围内。由于物流过程实现的是生鲜农产品在空间和时间上的转移,需要始终保持生鲜农产品处于适当的温湿度范围内。因此对物流过程中运输工具和冷库的标准化、制冷设备的先进性要求较高。一般情况下,苹果最佳环境温度在0~5℃,环境相对湿度约为85%~90%。而所选取的苹果质量安全原始数据中温度和湿度有超出适宜范围的值,故会出现微预警,计算结果符合实际情况。

4 结束语

针对生鲜农产品在冷链物流过程中的质量安全问题,通过对整个冷链物流作业环节中关键因素进行筛选,确定监测目标和安全指标,构建基于支持向量机算法的预警模型。预警结果表明,支持向量机在处理生鲜农产品质量安全预警问题时相比传统预警方法效果更好。该预警方法为避免生鲜农产品在冷链物流过程中出现腐败、变质、过期等问题提供了方法依据,适用于果蔬、肉制品和乳制品等农产品在冷链物流过程中的质量安全预警。除此之外,所提出的预警模型可应用于生鲜农产品质量安全预警和追溯系统,针对业务主体的全程冷链要求,面向消费者和供应链相关企业,保证生鲜农产品在整个供应链周期内实现风险的提前警示和事后的信息追溯,具有广阔的应用前景。

参考文献

- [1] NAKANDALA D, LAU H C W, ZHAO Li, Development of a hybrid fresh food supply chain risk assessment model [J]. International Journal of Production Research, 2017, 55 (14): 4180-4195.
- [2] DENG Xiuquan, YANG Xinmiao, ZHANG Yue, et al. Risk propagation mechanisms and risk management strategies for a sustainable perishable products supply chain [J]. Computers & Industrial Engineering, 2019, 135: 1175-1187.
- [3] 曾欣平,吕伟,刘丹. 基于供应链和可拓物元模型的乳制品企业农产品质量安全风险预警研究[J]. 环境与安全工程, 2019, 26 (03): 145-151.
- [4] 宋宝娥. 基于集值统计和供应链的食品安全预警模型探析[J]. 统计与决策, 2014 (12): 56-58.
- [5] 山红梅,杨雪静. 基于云模型—IAHP的生鲜农产品冷链物流服务水平评估[J]. 公路交通科技, 2020, 37(10): 142-148.
- [6] 张洋,陈伟炯,付姗姗. 基于数据挖掘的农产品供应链风险预警系统研究[J]. 广西大学学报(自然科学版), 2018, 43(03): 1118-1125.
- [7] 马长路,曲爱玲,吕加平,等. 基于物联网的生鲜乳安全检测与预警系统研究[J]. 中国乳品工业, 2016, 44(02): 53-56.
- [8] 杨玮,曹薇. BP神经网络在果蔬冷链物流预警中的应用[J]. 计算机工程与科学, 2015 (09): 1707-1711.
- [9] 高亚男,王文倩,王建新. 集成模糊层级划分的LightGBM食品安全风险预警模型:以肉制品为例[J]. 农产品科学, 2021, 42 (01): 197-207.
- [10] 张中丹,杨德州,王洲,等. 基于区块链技术的电网灾害预警决策系统[J]. 现代电力, 2021, 38(03): 307-315.
- [11] 刘昶,吴君华,张凤娇,等. 基于ARIMA-BP组合模型的城市公交客运量预测[J]. 智能计算机与应用, 2022, 12(07): 80-83.
- [12] 杜明英,杨珊珊,高向辉,等. 基于模糊综合评价法的农产品安全网络舆情预警研究[J]. 商业经济研究, 2018(09): 188-192.
- [13] 李佳洁,徐然,李江华. 基于模糊层次分析法的农产品冷链物流发展影响因素分析——以北京市为例[J]. 农产品质量与安全, 2015(01): 35-40.
- [14] DONG Qingxing, COOPER O. An orders-of-magnitude AHP supply chain risk assessment framework [J]. International Journal of Production Economics, 2016, 182: 144-156.
- [15] 游清顺,王建新,张秀宇,等. 基于支持向量机的食品安全抽检数据分析方法[J]. 软件工程, 2019, 22(02): 29-31.
- [16] 张新亚,沈菊红,刘楷. 一种输入数据为模糊数的模糊支持向量机[J]. 计算机工程与应用, 2017, 53(20): 122-127.
- [17] SONG H, KIM J, KIM B, et al. Development of a food temperature prediction model for real time food quality assessment [J]. International Journal of Refrigeration-Revue Internationale du Froid, 2019, 98: 468-479.
- [18] ISMAIL A, ABDLERAZEK S, ELHENAWY I. Development of smart healthcare system based on speech recognition using Support Vector Machine and dynamic time warping [J]. Sustainability, 2020, 12(6): 2403.