

机器学习在食品贮藏品质预测中的应用

代帅帅¹, 吴伟杰^{1,2*}, 牛犇¹, 房祥军¹, 陈慧芝¹, 陈杭君¹, 郜海燕^{1*}

(¹浙江省农业科学院食品科学研究所 农业农村部果品采后处理重点实验室 农业农村部蔬菜采后保鲜与加工重点实验室(部省共建) 浙江省果蔬保鲜与加工技术研究重点实验室 中国轻工业果蔬保鲜与加工重点实验室 杭州 310021

²省部共建农产品质量安全危害因子与风险防控国家重点实验室 杭州 310021)

摘要 食品贮藏和流通过程中会出现不同程度的品质劣变现象。随着人们对食品品质和安全重视程度的不断提高,开展食品贮运过程中的品质预测研究,对食品品质调控具有重要意义。本文综述机器学习在食品贮藏品质预测中的研究进展,包括常规的品质预测方法及其局限性。重点介绍近年发展快、应用广的集成学习和人工神经网络算法以及预测性能评估方法,展望机器学习在食品领域的未来发展趋势,为开展食品科学的交叉研究提供参考。

关键词 食品品质; 预测; 机器学习; 微生物; 模型评估

文章编号 1009-7848(2023)12-0337-12 **DOI:** 10.16429/j.1009-7848.2023.12.034

随着生活品质的提升与全民健康意识的增强,人们对食品需求量持续增大,对其品质要求不断提升^[1-2]。然而,食品在生产、贮藏、流通过程中会受到多种因素的影响而导致其品质劣变,如常见的微生物腐败变质^[3-4]、机械振动损伤^[5-6]、蒸腾作用^[7]、失水萎蔫^[8]等。此外,我国仅果蔬在运输环节的年损失量就高达 700 亿元^[9],存在着巨大的食品浪费。因此,通过预测的手段并及时调整储运条件,从而精准调控食品产后品质具有重要意义^[8]。然而,多因素、弱关联、高动态的品质信息特性阻碍了品质预测的研究发展。

随着计算机辅助技术的发展,研究人员开发了科学应用软件和信息系统,使得科研人员得以摆脱繁重的计算工作,也使决策得以实时实践^[9],相关的科研产出得到快速提升。然而,如今高通量、高并发、非结构化的数据特点又为实时处理提出新的时代课题。以机器学习为代表的人工智能技术经过几十年迭代已得到充分发展,在各个学科领域展现出强劲的适应能力^[10]。以人工智能为载体,相对新颖的现代预测研究是集计算机、数学、统计学等于一体的研究领域,也是近年来发展

比较快的一个交叉研究领域^[9],逐渐引起广大科研工作者的重视。

本文综述常规预测方法及局限性、机器学习及预测性能的评估方法等,重点介绍有关机器学习在贮藏品质预测方面的应用及泛化流程,为相关研究者提供参考。

1 食品品质预测研究的发展及现状

以食品品质为代表的数据具有多源耦合性、时变性、非对称性的特征,这导致食品品质感知和安全无法实现连续状态的可维持目标^[11]。早在 1825 年,Gompertz 等^[12]就关注到微生物生长与食品腐败的关联及对人体健康的影响,并试图建立微生物的生长模型。1985 年,Zamora 等^[13]基于数学模型建立了冷藏牛肉的微生物生长模型,测定了需氧菌总数、厌氧菌数、假单胞菌、热气梭菌、乳酸杆菌、大肠杆菌和酵母菌等随时间的变化。1996 年,Vankerschaver 等^[14]以莴苣为研究对象,分析温度、二氧化碳浓度对微生物生长及食品品质的影响,得到部分动态条件下的微生物生长模型。上述研究者仅建立了影响因子和品质劣变之间或各影响因子间的定性关系,并没有建立系统的定量关系,且不同研究人员依据的品质评价指标也不尽相同^[15],对食品品质调控不具备实际指导意义,由此也反映出现实环境中食品品质影响因子的多样性与复杂性。

收稿日期: 2022-12-01

基金项目: “十四五”国家重点研发计划项目
(2021YFD2100505)

第一作者: 代帅帅,男,硕士生

通信作者: 吴伟杰 E-mail: wuweijie87@163.com
郜海燕 E-mail: spsghy@163.com

1.1 基于微生物生长动力学的食品品质预测和调控

有学者指出,预测食品微生物生长,是食品品质安全管理的有效途径,它为 QMRA(Quantitative microal risk assessment)和 HACCP(Hazard Analysis Critical Control Point)提供了科学依据^[6]。以计算机为基础的现代微生物学预测,通过建立如外部环境因素与微生物之间关系的数学模型^[17],进而对食品品质和安全实现高效评估和有效控制。第十届食品预测建模国际会议(ICPMF10)的顺利召开,标志着微生物预测学领域走向成熟^[18]。

微生物预测的系统研究最早源于上世纪 20 年代对个别微生物热致死时间的计算,建立如温度、pH 值、二氧化碳水平、盐酸浓度、水分等影响因素和微生物行为的模型^[19-20]。1993 年,Whiting 和 Bchanan^[21]提出预测微生物的三级分类数学模型,自此基于数学建模法的微生物生长预测得到

系统发展,也促进了食品品质调控研究。具体来说,一级模型描述微生物随时间变化的基础关系,二级模型是在一级模型的基础上建立微生物随温度、pH 值、水分等环境变量变化的模型,三级模型是基于计算机辅助的二级模型,也被称为专家系统。表 1 是对基于三级预测模型统计的相关研究案例。

上述研究表明,三级预测模型在一定程度上可以预测食品微生物的生长趋势,为食品品质控制提供了生物调控途径,进一步拓展了保障食品安全的方法。基于此方法需要用科学合理的数学模型方法描述环境因素,而环境因素是实时变化与波动的,具有一定的偶然性和随机性,由此带来的预测结果是不稳定的^[22],存在着一定的局限性。此外,由于二级模型中不同环境因素变化的多样性,模型建立过程中需要经过繁杂的数学计算,因此不利于预测微生物学的发展^[23]。

表 1 三级预测模型的研究案例

Table 1 Research cases of three-level prediction models

研究团队	研究内容	研究结论
翁佩芳等 ^[24]	通过建立二级预测模型对榨菜腌制过程中植物乳杆菌的生长进行预测	建立了植物乳杆菌 L34 生长速率和环境因子之间的预测模型,得到自身偏差因子和准确因子
李苗云等 ^[25]	基于三级建模的肉品微生物生长预测模型的研究	综合多种预测方法得到更精确的微生物的生长预测模型
柳鑫等 ^[26]	基于二级数学模型建立的湿米粉中的微生物生长预测模型	利用 Gompertz 和平方根模型建立了湿米粉在不同时期的微生物生长预测模型
赵建兰 ^[27]	基于二级平方根模型建立的冷鲜鸡肉产品腐败微生物的生长预测模型	初始菌落数的变化对生长预测模型具有较大的影响
刘亚兵等 ^[22]	基于三级模型研究食品微生物生长的预测模型	微生物是影响食品品质劣变的主因之一
Park 等 ^[28]	利用多项式方程建立生菜的嗜水气单胞菌生长在温度(15~35℃)和相对湿度(RH,60%~80%)组合下的预测模型	温度 25℃和相对湿度 80%是嗜水气单胞菌的最佳生长条件,可为生菜的微生物风险定量评估提供可靠且实用的信息
Juneja 等 ^[29]	基于修正的 Ratkowsky 二级模型,建立意大利面冷却过程中芽孢中蜡样芽孢杆菌生长的模型	建立的芽孢杆菌生长模型有利于熟的面食在不同时间和温度下的品质调控
Tarlak 等 ^[30]	不同温度下鲜切生菜的沙门氏菌和内源乳酸菌的生长和存活模型	冷藏控制是有效的商业控菌手段

1.2 基于辅助软件的预测方法

辅助预测方法的本质是实现第三级数学模型的过程。进入 21 世纪,随着计算机学科的迅猛发展,基于数学方法建立的预测模型所面对的海量计算难题得到有效缓解,如于祝祝等^[32]提出的第三级分类数学预测模型,研究者只需在交互界面

中输入相关影响因子,无需了解建模技术或构建一级和二级模型即可自动得到分析结果。基于数学模型构建的辅助软件预测方法,已成为一种在食品工业和其它应用中更容易获取和使用的工具。

张宇婷等^[33]在 2008 年基于 Matlab 平台构建

了一个用于食品品质评价的微生物仿真系统,在形成仿真函数库的基础上,结合 VC6.0 编程软件开发出独立于 Matlab 的微生物预测软件。2020 年,于祝祝等^[32]基于三级数学模型,使用 Visual Basic 语言开发了大西洋鲑中微生物生长预测平台,在不同贮藏温度下,可以对微生物生长数量及曲线做出快速预测。2021 年,姜兰芳等^[34]基于 SPSS 主成分分析法,开发了面筋聚集仪预测小麦的品质,为掌握小麦品质及开展品质调控提供了有力依据。美国农业部微生物食品安全研究所开发的病原体模型程序 PMP (Pathogen Modeling Program),可根据预设的如温度、pH 值、水分等环境条件预测致病菌的生长或失活,而 PMP 缺乏在动态变化条件下的预测能力^[35]。基于软件的预测方法,若想建立动态和深层次的预测模型会受到软件内设参数的限制。针对上述困境,以机器学习为代表的人工智能技术正逐步改善这种局面。虽存在如建模、数据预处理等繁琐过程,但开源化、成本低、效果优的技术特点,得到科研工作者的广泛认可。

2 机器学习在食品品质预测中的应用

2.1 机器学习对食品品质调控的作用

面对高通量、高并发、非结构化的数据特点,

人脑从海量数据中提取有效信息的能力有限,因此需要借助计算机在一定约束或无约束条件下筛选有效信息并建立联系,揭示隐含数据中的本质客观规律、内在联系、发展趋势。

机器学习应用于食品贮藏品质调控中,主要有 3 个方面作用:1) 映射当前的食品品质或理化状态;2) 有效开展食品贮藏中的品质调控;3) 实现连续化动态控制。相较于 HACCP 国际体系及国标等固定标准而言,基于机器学习的品质预测,能够反映当前或未来一定时间范围内食品品质的变化情况,可做连续的动态标准,进而实现贮运过程中的品质动态调控。

2.2 机器学习算法的应用分类

机器学习算法可归纳为对数据进行分类、回归、聚类、降维等深层次处理的方法。对于输入端而言,可以分为有标签数据、无标签数据、混合标签数据以及决策流程和机理系统,分别对应机器学习算法中的监督学习、无监督学习、半监督学习和强化学习^[36-38]。输出端则可以根据结果的连续性分为两大类,输出数据若是离散值即是分类问题,连续值则称为回归问题。期望降低输出数据维度的处理方式称之为降维,用于发现数据潜在结构和联系并进行区分的输出方式是聚类。其具体分类见表 2。

表 2 机器学习相关算法分类表

Table 2 Classification tables of machine learning

分类	输入	输出	目的	应用情景
监督学习	有标签	分类、回归(有反馈)	预测结果	品质检测 病害预测
无监督学习	无标签	聚类(无反馈)	发现数据潜在结构及联系	自动分类 品质识别
半监督学习	部分有	降维(有反馈)	降低数据 标记难度	图像处理 天气预报
强化学习	决策流程或激励系统	一系列行动	长期利益最大化/优化	产线优化 药物设计

2.3 食品贮藏品质预测模型的构建流程

计算机硬件的发展与运算能力的提升,为多元数据分析与建模提供了基础。通常,将机器学习抽象为“输入”“函数”“输出”3 个核心过程,即将一些经过预处理的数据输入训练过的机器学习模型,进行无限制循环拟合,发现数据间具有某种定

量关系的函数即可做出预测^[39]。因此,借助机器学习解决实际问题可看作拟合函数的过程,即“给数据,做处理,出结果”的过程。本文通过综述相关内容,归纳总结基于机器学习构建食品贮藏品质预测模型的 4 个基本步骤(图 1):数据收集、数据预处理、训练模型、模型评估。

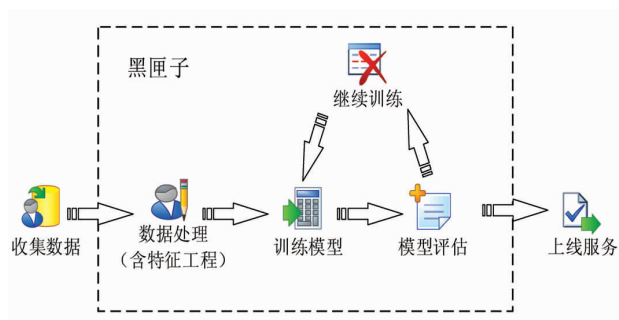


图1 泛化工作流程图

Fig.1 Generalization workflow

应用机器学习算法的首要条件是收集数据,一般是基于个人工实验内容产生的数据,也可以是相关数据库的数据。原始数据一般需要预处理,其目的是对有错误、缺失的数据进行修正、增补或对高维数据进行降维处理,主要包括缺失值处理、特征提取、数据降维、特征变换等^[38]。在泛化工作流程中,数据的预处理是必不可少的,将原始数据经特征工程处理转化为特征数据并作为输入数据使用,从本质上讲是高质量数据的展现过程。因此,机器学习训练效果的上限取决于数据和特征值的质量,而算法的选择及优化则是尽可能接近上限^[40]。其次是训练算法,将处理好的数据输入模型中进行训练,根据目标函数来检验模型的训练效果。上述步骤完成后,对程序输入相关数据集进行预测试验并得到预测结果^[41]。这一环节是整个机器学习中最困难的部分,需要不断评估训练模型并进行持续迭代优化,以达到预期效果。综上,机器学习的核心可归纳为“任务、特征、模型”

三大元素^[42]。

2.4 食品贮藏品质预测的典型算法

不同的数据集类型,应选择不同的算法。对于建立食品贮藏品质预测模型,其输出结果是关于数个影响因子的一条回归线,因此,食品贮藏品质预测可归纳为有监督的回归问题。

回归分析可选择的算法有线性回归、非线性回归、多项式回归、集成学习、人工神经网络等。有研究表明,数据之间多为非线性的关系,导致很多场合利用线性回归模型无法很好地拟合目标数据曲线^[43]。为了克服这种情况,引入非线性回归和多项式回归,其本质是线性回归的延伸。而集成学习和人工神经网络是为了克服单一预测和数据低层次关联的算法,该算法的数据输出效果好,也是目前研究较为热门、应用较广的算法类型。结合实际问题,重点介绍关于集成学习和人工神经网络相关算法。

2.4.1 集成学习 集成学习是通过多个独立模型学习再组合的预测算法,即该模型预测效果要高于其任一单一模型^[44]。在集成学习基础算法中主要有两类(如图2所示),一类是为了解决欠拟合问题而产生的boosting算法,另一类即是为了克服过拟合的bagging算法。此外,还有梯度提升树(GDBT)和极端梯度提升树(XGBoost),其中极端梯度提升树(XGBoost)在解决多数回归问题上具有十分优异的效果,如Hiura等^[45]在2021年利用XGBoost算法并基于ComBase数据库中的8个变量对食品中的单增李斯特菌进行预测,得到该菌

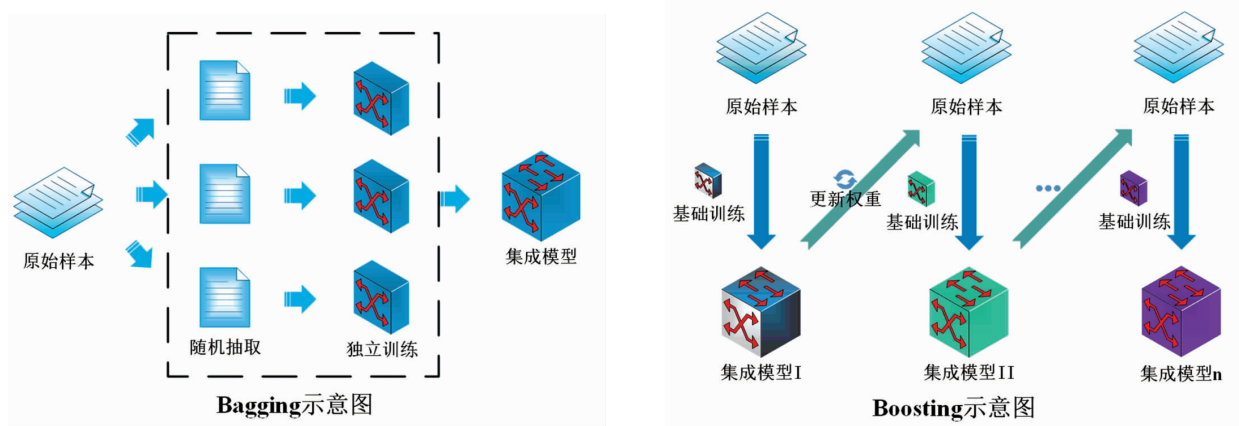


图2 集成学习示意图

Fig.2 Schematic diagram of integrated learning

准确的生长行为。上述研究表明,可以通过时间、温度、pH 值、水分活度等数据来预测食品中微生物种群行为,进一步提升食品安全相关的技术手段。

2.4.2 神经网络 如图 3 所示,神经网络也称为人工神经网络 (Artificial Neural Network, ANN),由一组相互关联的人工神经元(类神经元)构成的神经网络数学模型或基于计算机模拟生物神经网络的算法,通过模拟真实生物神经系统的交互作用处理高通量数据的过程^[36]。

姜鹏飞等^[43]综述了人工神经网络在食品工业中的应用,指出人工神经网络在食品微生物学及品质调控等领域具有广泛的发展空间。陈炎^[46]建立了基于 BP 人工神经网络的烟熏香肠多元品质预测模型,对烟熏香肠的加工及品质调控提供了理论依据。马惠玲等^[47]基于 GAN-BP-ANN 算法建立了苹果品质的预测模型,其预测准确度相较于 BP-ANN 算法提高了 4.7%,为苹果货架期及品质调控提供了一个有效方法。此外,人工神经网络不仅可以进行食品品质及微生物的预测,还可以对人工合成化合物的生物活性进行预测。2020 年, Badura 等^[48]为确定季铵盐对大肠杆菌菌株的抑制能力,设计了人工神经网络的分类和回归预测方法,得到人工神经网络可用于寻找潜在抗菌制剂的结论。通过以上研究分析得出,基于集成学习可以解决较为单一的传统预测模型所带来的预测精度低的问题,通过神经网络建立的预测模型可以解决数据之间关联性不强的问题,在一定程度上可以发现数据之间更多的潜在联系。

虽然已有很多类型的回归预测算法,但是不同算法对数据的处理具有一定倾向性。Tarazona 等^[49]采用 3 种算法对不同水分活度(0.96 和 0.99)和不同温度(20 ℃和 28 ℃)条件下食品中的真菌生长和毒素产生进行预测,结果表明,随机森林算法相较于神经网络和极端梯度提升树,预测效果更好。综上所述,对不同数据集进行预测,很难确定某个算法对某种数据类型的处理是最优的,只有根据实际数据集选择合适的训练模型,才能得到较好的预测效果。

2.5 集成技术

食品自身具有独特的光谱、介电常数、电磁、气味等特性,随着集成电子技术及物联网的飞速

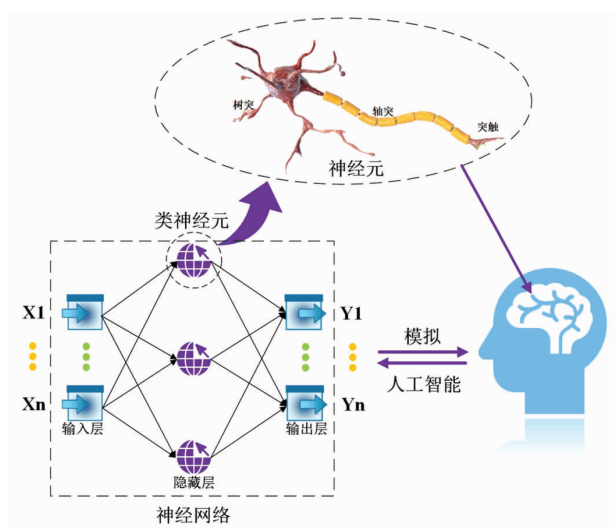


图 3 神经网络示意图

Fig.3 Neural network diagram

发展,借助相关传感器、集成设备、检测辅助技术等进行食品品质预测与调控的研究日趋火热。例如:

1) 高光谱成像。高光谱成像技术利用食品自身的光谱特性对其表面缺陷及外在品质数据化、可视化,结合其它算法进行品质预测。如张珏等^[50]基于高光谱成像技术和机器学习算法进行羊肉的新鲜度预测,预测集决定系数和均方根误差分别为 0.86 和 3.36。王凯丽^[51]通过高光谱成像技术结合降维对冬枣货架期进行预测,经 1%壳聚糖涂膜和 4 ℃条件下贮藏处理,可有效延缓贮藏 6 d 产生的褐变及劣变现象。通过选择最优贮藏条件,为冬枣采后贮藏和品质调控提供了理论依据和技术支持。

2) 近红外光谱。近红外光谱技术利用物体热效应及特征光谱可预测食品内在品质,如周娇娇等^[52]通过近红外光谱和化学计量学的方法建立了 pH 值、TVB-N、TBA 和 K 值的新鲜度定量预测模型,得到的相关系数分别为 0.961, 0.881, 0.955, 0.946 及交叉验证均方根误差分别为 0.049, 1.659, 0.047, 2.588, 表明所建模型具有较好的预测能力,为生鲜调理制品的品质调控提供了理论依据。

3) 拉曼光谱。结合相关算法可用于检测内在腐败菌及代谢产物,如眭亚南等^[53]通过拉曼光谱及前向逐步回归算法建立了明虾在不同贮藏条件

下的品质变化预测模型,其评价指标相关系数 r 值高达85%,实现了虾品质的快速检测及品质调控。

4) 介电特性。通过等效电容法测定食品的介电参数结合相关算法进行品质预测。如沈静波等^[54]通过介电参数和BP神经网络预测灵武长枣的新鲜度,其准确率为81.67%,通过构建此类模型可为果品的新鲜度预测提供新思路和方法。

5) 环境气体信息。通过电子鼻等气敏传感器搜集对其所处的环境气体信息进行预测。如张永超等^[55]建立了基于环境气体信息的BP神经网络苹果贮藏品质预测,得到的预测值与实际检测值间的相对误差不超5%。基于此可有效解决果蔬预测精度低的问题。

3 预测性能评估

构建基于机器学习的食品贮藏品质预测模型,建模的合理性决定了预测性能的高、低。因此,有效的评估方法判断模型是否具备实际运用能力是关键,也是建模中的最后一步。评估方法通常需要根据训练模型的不同而有所改变,对模型评估的分类及指标、模型评估方法及拟合情况进行分析。

3.1 评估分类及指标

评价模型的指标众多,不同指标所占权重不同,且作用不一。如果不能合理地运用评估指标,不仅无法发现食品品质预测模型的问题,而且有可能得到错误的结论。因此,评估性能指标的选择对整个模型的训练效果起着十分重要的作用。通常,输出结果一般为分类和回归2种类型。对于分类模型来说,评估指标有准确率、错误率、精确率、AUG指标、召回率、F1-score、操作者操作特征曲线等,可以用混淆矩阵对分类模型的准确性进行评估,它是最基本、最直观、计算最简单的方法^[37-38,56]。对于回归类型而言,衡量模型预测情况的指标是决定系数(R^2)、误差平方和(SSE)、均方根误差(RMSE)、相对平方误差(RSE)、相对绝对误差(RAE)、平均绝对误差(MAE)、平均绝对百分比误差(RAPE)等。通常,品质预测的输出值一般为连续值,因此重点介绍有关“回归”的相关指标。以下指标中, y_i 表示实际值, \hat{y}_i 表示预测值, \bar{y}_i 表示平均

值。

1) 决定系数(R^2),衡量多元回归方程中拟合程度,也是最基本的回归评估指标,通常越高越好。

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (1)$$

2) 误差平方和(SSE),表示结果数据和原始数据之间的误差平方和。

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

3) 均方根误差(RMSE),表示预测数据的离散程度,通常用同一单位衡量回归模型衡量预测效果。

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (3)$$

4) 相对平方误差(RSE),相较于RMSE而言,表示不同单位衡量模型预测效果。

$$RSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (4)$$

5) 相对绝对误差(RAE),区别于RSE而言,用来比较不同单位的模型误差。

$$RAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{\sum_{i=1}^n |y_i - \bar{y}_i|} \quad (5)$$

6) 平均绝对误差(MAE),反应预测值误差的真实情况。

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

7) 平均绝对百分比误差(MAPE),表示预测模型准确性的衡量指标。

$$MAPE = \frac{100}{i} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (7)$$

3.2 评估方法

通常经过预处理的数据集可利用重抽样的方法将其分为训练集和测试集(也称验证集)^[38]。训练集主要用于训练模型,而对模型进行测试的数据称为测试集。测试集是独立于训练集之外的数据,且完全不参与训练,用于验证模型的泛化性能(generalizability),即可用于其它相似数据的能力,也就是用于最终模型的评估验证。

然而,现实中的数据质量参差不齐、异构严

重,用于训练或测试的数据分布不一,会对预测的结果产生相应的影响。因此,在数据预处理阶段应使其平衡分布并均匀抽样。有关数据平衡处理方法可参考 Fuller 等^[57]在测定食品中蛋白质和氨基酸对肠道微生物合成及其消化率影响的研究方法。

3.3 拟合情况

模型评估用于评价训练好的模型表现效果。模型训练的结果可能会出现两种情况,一是欠拟合情况,产生欠拟合现象的原因是训练数据集中的特征数据过少,没有准确建立其样本数据特征关系,导致学习的标准粗糙,并不能准确预测数据。另一种情况是在训练样本中表现的过于优越,导致在使用测试集样本中的数据时表现不佳,出现这种情况的多数原因就是过拟合。

这两者之间的相同点是训练效果都很好,在测试数据上都会表现失真,不符合预测精度;区别是两者在训练集中的数据量,如果在训练数据较少的情况下出现训练效果较好而验证效果不好即是欠拟合,反之即是过拟合^[58]。

对于上述两种问题,有不同的解决办法。解决欠拟合问题所采用的方法是增加训练集的样本容量;而解决过拟合问题一般做法是抽取一部分训练样本中的数据作为测试数据,基于此可避免过拟合情况的发生以优化模型的训练效果^[57]。

4 总结及展望

品质预测已经逐渐发展成为评价和控制食品安全的重要手段之一。借助品质预测的方法,有利于对食品在加工、贮藏、流通过程中的品质做出科学合理的判断及安全控制。此外,通过整合在品质预测建模中积累的数据,有利于产品生产工艺控制、产品配方的制定与改良^[11],还有利于确定食品在加工、贮藏、流通过程中腐败的关键因素^[59]。

然而,对于食品品质预测还有很长的路要走。基于数学方法建立的单个预测模型,仅仅靠数个变量并不能准确预测出食品贮藏品质的情况,这是由于食品的组成成分复杂且相关因素易受环境变化等动态影响,许多未列入变量的影响因子在构建预测模型时却有着较大的权重。比如,Verheyen 等^[19]指出食品本身的微观结构对微生物生

长动力学有相关的影响,未来应综合食品微观结构层面的宏观二级模型和微观尺度的半机械模型,以更好地建立一个长久适用的预测模型。Koseki 等^[60]指出,科研工作者在建立传统的微生物随时间生长/死亡行为预测模型时,没有考虑到诸如个体细胞异质性等因素的拟合不确定性和变异性问题。以上研究有利食品加工和配送环境实际应用中品质的控制。因此,想要得到一个较为准确的预测模型,需持续不断地提供新的支撑数据并进行深入研究。

为了提高预测精度,相关学者结合了集成技术的预测方法,如非线性混合模型和机器学习结合的方法^[61]。然而,目前基于机器学习的预测手段也存在一些缺陷,比如预测结果具有不可解释性^[62]和不稳定性,尤其是这些关系往往是非线性的^[63]。此外,预测模型是对精确程度的估计,这取决于对数据进行特征工程的处理方法和输入数据的质量等因素^[64],不能实现高度的自动化处理,这在一定程度上阻碍了机器学习应用的广度和深度。

大数据和人工智能的结合为常规预测方法提供了新的选择,整合这些技术,并建立下一代预测技术将更加精准、便捷^[65]。未来可结合物联网、人工智能、区块链等新兴信息技术,围绕食品品质感知、品质信息的处理与传输、可信溯源 3 个方面展开研究,最终实现食品在贮运中品质调控的智能化、自动化、系统化,为食品全产业链实施实时监控、感知及调控,促进食品产业的信息化发展。

参 考 文 献

- [1] 曹锦萍, 陈焯芝, 孙翠, 等. 我国果蔬产地商品化技术支撑体系发展现状[J]. 浙江大学学报(农业与生命科学版), 2020, 46(1): 1-7.
CAO J P, CHEN Y Z, SUN C, et al. Development status of the technology supporting system for local commoditization of fruits and vegetables in chain[J]. Journal of Zhejiang University (Agriculture and Life Sciences), 2020, 46(1): 1-7.
- [2] 郭志明, 王郡艺, 宋焯, 等. 果蔬品质劣变传感检测与监测技术研究进展[J]. 智慧农业(中英文), 2021, 3(4): 14-28.

- GUO Z M, WANG J Y, SONG Y, et al. Research progress of sensing detection and monitoring technology for fruit and vegetable quality control [J]. *Smart Agriculture (Chinese and English)*, 2021, 3(4): 14–28.
- [3] 陈雯钰. 基于特征微生物生长动力学的食品包装保质期预测[D]. 无锡: 江南大学, 2013.
CHEN W Y. Shelf-life prediction of packaged food based on growth of specific spoilage organisms [D]. Wuxi: Jiangnan University, 2013.
- [4] 李海洲, 周小刚, 唐衍军. 区块链技术赋能果品供应链质量安全研究[J]. *中国果树*, 2021(5): 79–82.
LI H Z, ZHOU X G, TANG Y J. Research on quality and safety management of fruit supply chain enabled by blockchain technology [J]. *China Fruits*, 2021(5): 79–82.
- [5] 张壹钦, 郜海燕, 韩延超, 等. 振动胁迫下 pva 基微孔减振包装对草莓品质的影响[J]. *中国食品学报*, 2019, 19(11): 215–222.
ZHANG Y Q, GAO H Y, HAN Y C, et al. Effect of PVA-based microporous vibration-damping packaging on strawberry quality under vibration stress [J]. *Journal of Chinese Institute of Food Science and Technology*, 2019, 19(11): 215–222.
- [6] 周强, 韩延超, 陈杭君, 等. 电商物流贮藏过程杨梅品质的变化及货架寿命预测[J]. *中国食品学报*, 2021, 21(11): 112–119.
ZHOU Q, HAN Y C, CHEN H J, et al. Quality change and shelf-life prediction of myrica during e-commerce logistics storage [J]. *Journal of Chinese Institute of Food Science and Technology*, 2021, 21(11): 112–119.
- [7] 刘娟, 吴伟杰, 郜海燕, 等. 贮藏温度对鲜切火龙果品质及微生物的影响[J]. *中国食品学报*, 2017, 17(10): 168–175.
LIU J, WU W J, GAO H Y, et al. Effects of different storage temperatures on quality and microorganism of fresh-cut pitaya [J]. *Journal of Chinese Institute of Food Science and Technology*, 2017, 17(10): 168–175.
- [8] 罗海波, 何雄, 包永华, 等. 鲜切果蔬品质劣变影响因素及其可能机理[J]. *食品科学*, 2012, 33(15): 324–330.
LUO H B, HE X, BAO Y H, et al. Affecting factors and possible mechanisms of quality deterioration in fresh-cut fruits and vegetables [J]. *Food Science*, 2012, 33(15): 324–330.
- [9] MCMEEKIN T A, BARANYI J, BOWMAN J, et al. Information systems in food safety management [J]. *International Journal of Food Microbiology*, 2006, 112(3): 181–194.
- [10] 崔晓晖, 李伟, 顾诚淳. 食品科学大数据与人工智能技术[J]. *中国食品学报*, 2021, 21(2): 1–8.
CUI X H, LI W, GU C C. Big data of food science and artificial intelligence technology [J]. *Journal of Chinese Institute of Food Science and Technology*, 2021, 21(2): 1–8.
- [11] ADAMS M. Disciplines associated with food safety: Food microbiology [J]. *Encyclopedia of Food Safety*, 2014, 1: 28–32.
- [12] GOMPERTZ B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c. By Benjamin Gompertz, Esq. FR S[C]. In *Abstracts of the Papers Printed in the Philosophical Transactions of the Royal Society of London*, 1833(2): 252–253.
- [13] ZAMORA M C, ZARITZKY N E. Modeling of microbial growth in refrigerated packaged beef [J]. *Journal of Food Science*, 1985, 50(4): 1003–1006.
- [14] VANKERSCHAUVER K, WILLOCX F, SMOUT C, et al. The influence of temperature and gas mixtures on the growth of the intrinsic micro-organisms on cut endive: predictive versus actual growth [J]. *Food microbiology*, 1996, 13(6): 427–440.
- [15] 曹颖, 郜海燕, 陈杭君, 等. 荔枝加工品质评价研究进展[J]. *中国食品学报*, 2011, 11(8): 126–132.
CAO Y, GAO H Y, CHEN H J, et al. Progress in research on measurements of processing quality in Litchi [J]. *Journal of Chinese Institute of Food Science and Technology*, 2011, 11(8): 126–132.
- [16] 周康, 刘寿春, 李平兰, 等. 食品微生物生长预测模型研究新进展[J]. *微生物学通报*, 2008, 35(4): 589–594.
ZHOU K, LIU S C, LI P L, et al. New advances in predictive food microbial growth model [J]. *Microbiology China*, 2008, 35(4): 589–594.
- [17] 唐坚. 生菜的冰温保鲜及微生物预测模型的初步建立[D]. 上海: 上海师范大学, 2015.
TANG J. Freezing-point storage of lettuce and the

- preliminary establishment of prediction model of microorganisms[D]. Shanghai: Shanghai Normal University, 2015.
- [18] FERNANDO PEREZ-RODRIGUEZ E C A K, ANDERSON S, SANT'ANA V P V A. Special issue on 10th international conference of predictive modelling in foods towards a new paradigm in predictive microbiology[J]. International Journal of Food Microbiology, 2018, 291: 65-66.
- [19] VERHEYEN D, VAN IMPE J F M. The inclusion of the food microstructural influence in predictive microbiology: state-of-the-art[J]. Foods, 2021, 10(9): 2119.
- [20] STAVROPOULOU E, BEZIRTZOGLU E. Predictive modeling of microbial behavior in food[J]. Foods, 2019, 8(12): 654.
- [21] WHITING R C, BUCHANAN R L. A classification of models for predictive microbiology[J]. Food Microbiology, 1993, 10: 175-177.
- [22] 刘亚兵, 何腊平, 高泽鑫, 等. 食品微生物生长预测模型的研究[J]. 食品工业, 2016, 37(11): 159-164.
- LIU Y B, HE L P, GAO Z X, et al. Research on predictive model of food microorganism growth[J]. The Food Industry, 2016, 37(11): 159-164.
- [23] SKANDAMIS P N, JEANSON S. Colonial vs. Planktonic type of growth: mathematical modeling of microbial dynamics on surfaces and in liquid, semi-liquid and solid foods[J]. Frontiers in Microbiology, 2015, 6: 1178.
- [24] 翁佩芳, 雷兰兰, 吴祖芳. 榨菜腌制特征性微生物植物乳杆菌生长预测模型的研究[J]. 中国食品学报, 2012, 12(10): 67-75.
- WENG P F, LEI L L, WU Z F, et al. Growth prediction models of characteristic microbe *Lactobacillus plantarum* in pickled vegetables[J]. Journal of Chinese Institute of Food Science and Technology, 2012, 12(10): 67-75.
- [25] 李苗云, 田璐, 赵改名, 等. 肉品微生物生长预测模型研究进展[J]. 肉类研究, 2012, 26(12): 20-24.
- LI M Y, TIAN L, ZHAO G M, et al. Research progress in predictive modeling of microbial growth in meat products[J]. Meat Research, 2012, 26(12): 20-24.
- [26] 柳鑫, 文丽, 李莎, 等. 湿米粉中菌相分析与微生物生长预测模型的建立[J]. 中国酿造, 2013, 32(1): 65-70.
- LIU X, WEN L, LI S, et al. Analysis of main microflora and development of predictive models of microbe in wet rice noodle[J]. China Brewing, 2013, 32(1): 65-70.
- [27] 赵建兰. 气调包装冷鲜鸭肉产品腐败微生物及其生长预测模型研究[D]. 武汉: 武汉轻工大学, 2015.
- ZHAO J L. Research on spoilage microorganisms and growth prediction model in chilled duck with modified atmosphere packaging[D]. Wuhan: Wuhan Polytechnic University, 2015.
- [28] PARK S Y, CHOI S, HA S. Predictive modeling for the growth of *aeromonas hydrophila* on lettuce as a function of combined storage temperature and relative humidity[J]. Foodborne Pathogens and Disease, 2019, 16(6): 376-383.
- [29] JUNEJA V K, GOLDEN C E, MISHRA A, et al. Predictive model for growth of *bacillus cereus* at temperatures applicable to cooling of cooked pasta[J]. Journal of Food Science, 2019, 84(3): 590-598.
- [30] TARLAK F, JOHANNESSEN G, BASCÓN VILLEGAS I, et al. Modelling of the behaviour of *Salmonella enterica* serovar reading on commercial fresh-cut iceberg lettuce stored at different temperatures[J]. Foods, 2020, 9(7): 946.
- [31] KOWALIK J, LOBACZ A, ZULEWSKA J, et al. Analysis and mathematical modelling of the behaviour of *Escherichia coli* in the mascarpone cheese during cold storage[J]. International Journal of Food Science & Technology, 2018, 53(6): 1541-1548.
- [32] 于祝祝, 林洪, 王静雪. 大西洋鲑贮藏过程中微生物生长预测系统的构建[J]. 食品科学, 2020, 41(9): 139-144.
- YU Z Z, LIN H, WANG J X. Predictive system of microbial growth on atlantic salmon during storage[J]. Food Science, 2020, 41(9): 139-144.
- [33] 张宇婷, 孟雅娟, 闫国婷. Matlab在微生物生长预测模型中的应用[J]. 河北化工, 2008(1): 20-22.
- ZHANG Y T, MENG Y J, YAN G T. The application of matlab in microbiology growth prediction model[J]. Hebei Chemical Industry, 2008(1): 20-22.
- [34] 姜兰芳, 李晓丽, 曹勇, 等. 基于主成分分析的面筋聚集仪预测小麦品质[J]. 食品科学, 2022, 43(14): 85-92.
- JIANG L F, LI X L, CAO Y, et al. Prediction of

- wheat quality using glutopex method based on principal component analysis[J]. *Food Science*, 2022, 43(14): 85–92.
- [35] 方兰兰. 真空包装莲藕褐变特性和微生物生长预测模型研究[D]. 武汉: 武汉工业学院, 2012.
- FANG L L. Study on browning characteristics and microbial growth forecast model of vacuum-packaging lotus root[D]. Wuhan: Wuhan Polytechnic University, 2012.
- [36] 杜伟. 机器学习及数据挖掘在生物信息学中的应用研究[D]. 长春: 吉林大学, 2011.
- DU W. Research on application of machine learning and data mining in bioinformatics [D]. Changchun: Jilin University, 2011.
- [37] 李强, 衣杨, 吴忠道, 等. 基于机器学习的肠道菌群数据建模与分析研究综述[J]. *微生物学通报*, 2021, 48(1): 180–196.
- LI Q, YI Y, WU Z D, et al. Review of gut microbiome analysis prediction models and algorithms [J]. *Microbiology China*, 2021, 48(1): 180–196.
- [38] 李高磊, 黄玮, 孙浩, 等. 机器学习在微生物组宿主表型预测中的应用[J]. *微生物学报*, 2021, 61(9): 2581–2593.
- LI G L, HUANG W, SUN H, et al. Applications of machine learning in predicting host phenotype based on microbiome[J]. *Acta Microbiologica Sinica*, 2021, 61(9): 2581–2593.
- [39] 田有文, 吴伟, 卢时铅, 等. 深度学习在水果品质检测与分级分类中的应用[J]. *食品科学*, 2021, 42(19): 260–270.
- TIAN Y W, WU W, LU S Q, et al. Application of deep learning in fruit quality detection and grading [J]. *Food Science*, 2021, 42(19): 260–270.
- [40] 诸葛越. 百面机器学习: 算法工程师带你去面试[M]. 北京: 人民邮电出版社, 2018.
- ZHU G Y. The quest for machine learning: 100+ interview questions for algorithm engineer [M]. Beijing: Posts and Telecommunications Press, 2018.
- [41] 付宇. 基于 xgboost 模型的贵州省野外真菌菌种预测系统及可视化展示[D]. 北京: 北京林业大学, 2020.
- FU Y. Prediction and visualization of wild fungus species in guizhou province based in XGboost model [D]. Beijing: Beijing Forestry University, 2020.
- [42] 王聪, 姜舒文, 黄坤, 等. 机器学习在农产品供应链关键环节中的应用进展研究综述[J]. *中国农机化学报*, 2021, 42(5): 182–190.
- WANG C, JIANG S W, HUANG K, et al. A review on machine learning for applications in key links of agricultural products supply chain[J]. *Journal of Chinese Agricultural Mechanization*, 2021, 42(5): 182–190.
- [43] 姜鹏飞, 于文静, 孙娜, 等. 人工神经网络在食品工业中的应用[J]. *食品研究与开发*, 2021, 42(13): 188–196.
- JIANG P F, YU W J, SUN N, et al. Application of artificial neural network in food industry[J]. *Food Research and Development*, 2021, 42(13): 188–196.
- [44] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 560.
- ZHOU Z H. Machine learning[M]. Beijing: Tsinghua University Press, 2016: 560.
- [45] HIURA S, KOSEKI S, KOYAMA K. Prediction of population behavior of listeria monocytogenes in food using machine learning and a microbial growth and survival database [J]. *Scientific Reports*, 2021, 11(1): 10613.
- [46] 陈炎. 基于 bp 人工神经网络的烟熏香肠多元品质预测研究[D]. 合肥: 合肥工业大学, 2017.
- CHEN Y. Prediction of smoked sausage qualities based on BP artificial neural network [D]. Hefei: Hefei University of Technology, 2017.
- [47] 马惠玲, 曹梦柯, 王栋, 等. 苹果货架期 gan-bp-ann 预测模型研究[J]. *农业机械学报*, 2021, 52(11): 367–375.
- MA H L, CAO M K, WANG D, et al. Study on shelf-life prediction of apple with GAN-BP-ANN model[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(11): 367–375.
- [48] BADURA A, KRYSIŃSKI J, NOWACZYK A, et al. Application of artificial neural networks to prediction of new substances with antimicrobial activity against *Escherichia coli* [J]. *Journal of Applied Microbiology*, 2021, 130(1): 40–49.
- [49] TARAZONA A, MATEO E M, GÓMEZ J V, et al. Machine learning approach for predicting *Fusarium culmorum* and *F. Proliferatum* growth and mycotoxin production in treatments with ethylene-vinyl alcohol copolymer films containing pure components of essential oils[J]. *International Journal of Food Microbiology*, 2021, 338: 109012.

- [50] 张珏, 田海清, 王轲, 等. 基于高光谱成像技术的羊肉新鲜度预测[J]. 中国农业大学学报, 2020, 25(5): 94-103.
ZHANG Y, TIAN H Q, WANG K, et al. Non-destructive detection of lamb freshness based on hyperspectral imaging technology[J]. Journal of China Agricultural University, 2020, 25(5): 94-103.
- [51] 王凯丽. 基于高光谱成像技术的冬枣货架期预测研究[D]. 泰安: 山东农业大学, 2021.
WANG K L. Shelf-life prediction of winter jujube based on hyperspectral imaging[D]. Taian: Shandong Agricultural University, 2021.
- [52] 周娇娇, 吴潇扬, 陈周, 等. 近红外光谱技术快速预测团头鲂新鲜度[J]. 华中农业大学学报, 2019, 38(4): 120-126.
ZHOU J J, WU X Y, CHEN Z, et al. Using near-infrared reflectance spectroscopy to quickly predict the freshness of *Megalobrama amblycephala* [J]. Journal of Huazhong Agricultural University, 2019, 38(4): 120-126.
- [53] 眭亚南, 张雷蕾, 卢诗扬, 等. 基于拉曼光谱技术的不同贮藏条件下明虾品质变化预测模型的研究[J]. 光谱学与光谱分析, 2020, 40(5): 1607-1613.
SUI Y N, ZHANG L L, LU S Y, et al. Research on the shrimp quality of different storage conditions based on raman spectroscopy and prediction model [J]. Spectroscopy and Spectral Analysis, 2020, 40(5): 1607-1613.
- [54] 沈静波, 张海红, 马雪莲, 等. 基于介电特性的灵武长枣新鲜度预测[J]. 食品与机械, 2016, 32(1): 117-120.
SHEN J B, ZHANG H H, MA X L, et al. Prediction on freshness degree of lingwu long jujube on dielectric properties[J]. Food & Machinery, 2016, 32(1): 117-120.
- [55] 张永超, 赵录怀, 王昊, 等. 基于环境气体信息的bp神经网络苹果贮藏品质预测[J]. 江苏农业学报, 2020, 36(1): 194-198.
ZHANG Y C, ZHAO L H, WANG H, et al. Prediction of apple storage quality using BP neural network based on environmental gas information [J]. Jiangsu Journal of Agricultural Sciences, 2020, 36(1): 194-198.
- [56] 刘宏生, 冯华炜, 张力, 等. 机器学习在 maldi-tof ms 鉴定微生物中的应用 [J]. 微生物学报, 2020, 60(5): 841-855.
LIU H S, FENG H W, ZHANG L, et al. Application of machine learning in MALDI-TOF MS identification of microorganisms [J]. Acta Microbiologica Sinica, 2020, 60(5): 841-855.
- [57] FULLER M. Determination of protein and amino acid digestibility in foods including implications of gut microbial amino acid synthesis[J]. British Journal of Nutrition, 2012, 108(S2): S238-S246.
- [58] 吴杰, 吴雨濛. 机器学习在数据包络分析模型中的应用: 一种用于指标选择的智能机制[J]. 中国科学技术大学学报, 2022, 52(12): 41-49, 70, 72.
WU J, WU Y M. Machine learning in data envelopment analysis: A smart mechanism for indicator selection [J]. Journal of University of Science and Technology of China, 2022, 52(12): 41-49, 70, 72.
- [59] 赵瑞兰, 李远钊, 张培正. 预测食品微生物学概述及应用[J]. 中国食品卫生杂志, 2006, 18(5): 464-468.
ZHAO R L, LI Y Z, ZHANG P Z, et al. Predictive food microbiology brief introduction of theory and application[J]. Chinese Journal of Food Hygiene, 2006, 18(5): 464-468.
- [60] KOSEKI S, KOYAMA K, ABE H. Recent advances in predictive microbiology: theory and application of conversion from population dynamics to individual cell heterogeneity during inactivation process[J]. Current Opinion in Food Science, 2021, 39: 60-67.
- [61] QIAN W, VIENNET E, GLASS K, et al. Epidemiological models for predicting ross river virus in Australia: a systematic review [J]. PLOS Neglected Tropical Diseases, 2020, 14(9): e8621.
- [62] YEZHU Y, YANG W, YANG Y, et al. Interpretable machine learning methods for *in vitro* pharmaceutical formulation development [J]. Food Frontiers, 2021, 2(2): 195-207.
- [63] LOPATKIN A J, COLLINS J J. Predictive biology: modelling, understanding and harnessing microbial complexity [J]. Nature Reviews Microbiology, 2020, 18(9): 507-520.
- [64] TESSON V, FEDERIGHI M, CUMMINS E, et al. A systematic review of beef meat quantitative microbial risk assessment models [J]. International Journal of Environmental Research and Public Health, 2020, 17(3): 688.

Application of Machine Learning on Food Storage Quality Prediction

Dai Shuaishuai¹, Wu Weijie^{1,2*}, Niu Ben¹, Fang Xiangjun¹, Chen Huizhi¹, Chen Hangjun¹, Gao Haiyan^{*}

(¹Institute of Food Science, Zhejiang Academy of Agricultural Sciences, Key Laboratory of Post-Harvest Fruit Processing, Key Laboratory of Post-Harvest Vegetable Preservation and Processing, Ministry of Agriculture and Rural Affairs, Key Laboratory of Fruit and Vegetable Preservation and Processing Technology of Zhejiang Province, Key Laboratory of Light Industry Fruit and Vegetable Preservation and Processing, Hangzhou 310021

²State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-products, Hangzhou 310021)

Abstract During the process of food storage and circulation, there will be different degrees of quality deterioration. With the improvement of people's attention to food quality and safety, it is of great significance to carry out quality prediction research in the process of food storage and transportation for quality control. This paper reviews the research progress of machine learning in food storage quality prediction, including conventional quality prediction methods and limitations, and then focuses on the rapid development and wide application of integrated learning and artificial neural network algorithms, and prediction performance evaluation methods in recent years. Finally, it summarizes and looks forward to the future development trend of machine learning in the food field, and provides relevant references for the development of food science cross research.

Keywords food quality; prediction; machine learning; microorganisms; model evaluation