

DOI: 10.19812/j.cnki.jfsq11-5956/ts.20240728002

机器学习在白酒产地溯源研究中的应用进展

程铁轅¹, 肖宇¹, 薛康¹, 赵宇舒², 刘俊¹, 郭云霞³, 胡江涛^{4*}

(1. 成都海关技术中心, 成都 610041; 2. 四川省宜宾环球神州包装科技有限公司, 宜宾 644000;
3. 宜宾职业技术学院, 宜宾 644000; 4. 四川国际旅行卫生保健中心, 成都 610042)

摘要: 机器学习对于白酒产地溯源技术的应用效果具有重要价值。白酒产地溯源主要基于不同产地白酒在理化等方面的差异, 采用色谱、光谱、质谱等技术进行检测分析, 其一般技术路线: 对国内不同产地的白酒样品进行检测分析, 为寻找不同产地白酒之间的差异, 建立模型对获取的相关数据进行分析, 以便于更好地分类、预测。在这一过程中, 机器学习往往起着非常重要的作用, 既能够弥补一些现有数据分析方法的不足, 也能够很好地发挥相关检测技术的优势, 机器学习拓展了相关检测技术应用的深度和广度。本文综述了机器学习在白酒等食品、农产品、中药材产地溯源研究中的应用情况, 主要是与检测技术的联用情况, 总结分析了机器学习在白酒产地溯源研究中的应用效果, 并对其未来发展应用进行了展望, 可为白酒产地掺假行政执法提供重要技术支撑。

关键词: 机器学习; 白酒; 产地溯源; 算法; 模型

Progress of machine learning in the research of Baijiu origin traceability

CHENG Tie-Yuan¹, XIAO Yu¹, XUE Kang¹, ZHAO Yu-Shu², LIU Jun¹,
GUO Yun-Xia³, HU Jiang-Tao^{4*}

(1. Technology Center of Chengdu Customs, Chengdu 610041, China; 2. Sichuan Yibin Global Shenzhou Packaging Technology Co., Ltd., Yibin 644000, China; 3. Yibin Vocational and Technical College, Yibin 644000, China;
4. Sichuan International Travel Health Care Center, Chengdu 610042, China)

ABSTRACT: Machine learning is of great value to the application effect of Baijiu origin traceability technology. Baijiu origin traceability is mainly based on the differences of Baijiu from different origins in physical and chemical aspects, and detection and analysis are carried out by chromatography, spectrum, mass spectrometry and other technologies, its general technical route: To detect and analyze Baijiu samples from different places in China, in order to find the differences between Baijiu from different places, establish a model to analyze the relevant data obtained, so as to facilitate better classification and prediction. In this process, machine learning often plays a very important role, which can not only make up for the shortcomings of some existing data analysis methods, but also effectively leverage the advantages of related detection technologies. Machine learning has expanded the depth and breadth of the application of related detection technologies. This paper reviewed the application of machine learning in the

基金项目: 四川省重点研发项目(2023YFS0447)、成都海关科研项目(2023CK003)

Fund: Supported by the Key Research and Development Project of Sichuan Province (2023YFS0447), and the Chengdu Customs Research Project (2023CK003)

*通信作者: 胡江涛, 正高级工程师, 主要研究方向为食品安全检测。E-mail: 11166589@qq.com

*Corresponding author: HU Jiang-Tao, Professor, Sichuan International Travel Health Care Center, No.1, Tongzilin North Road, Wuhou District, Chengdu 610042, China. E-mail: 11166589@qq.com

research of origin traceability of Baijiu and other food, agricultural products, and traditional Chinese medicine, mainly in combination with detection technology, discussed the application effect of machine learning in the research of Baijiu origin traceability, and prospected its future development and application, which can provide important technical support for administrative law enforcement of Baijiu origin adulteration.

KEY WORDS: machine learning; Baijiu; origin traceability; algorithm; model

0 引言

白酒产地(产区)对于白酒产品经济价值具有重要影响。国内消费市场上一直存在“川酒云烟”的说法,四川等白酒优质产区奠定了特定产区白酒产品的高品质基础,赋予了产区内白酒产品更高的经济价值,从而获得更高的市场溢价。白酒优质产区具有得天独厚的自然条件,是出产高品质白酒的关键影响因素之一,是白酒产业高质量发展的重要基础,白酒产区保护受到了高度重视。受高额经济利益驱使以及信息不对称等因素的影响,当前国内消费市场上有关于白酒以次充好的问题仍较为普遍,尤其冒用优质产区知名品牌的违法行为,严重扰乱了市场秩序,损害了消费者权益。为加强我国白酒监管,采用先进科学技术实现白酒产地溯源尤为必要,而利用机器学习对产品和产地海量数据的分析与挖掘,对于白酒产地溯源技术的应用效果具有重要影响。

机器学习作为一门从数据中研究算法的学科,在数据分析领域扮演着重要角色,对很多研究领域产生了深刻影响,其与色谱、光谱、质谱等检测技术相结合,在食品、农产品等领域产地溯源研究中应用效果很好。从白酒产地溯源研究进展看,目前主要采用色谱、光谱、质谱等技术联用,再利用机器学习等进行建模分析以分类预测。当前白酒产地溯源研究仍处于起步阶段,亟需新技术、新方法的引入和应用。国内白酒产地众多,涉及海量数据,而机器学习具有自动处理大规模数据、泛化能力强等优势,其与相关溯源技术的高效联用,可为当前有关白酒产地掺假行政执法提供重要技术支撑,对于白酒产区保护具有重要意义。

1 机器学习概念、算法与模型

1.1 机器学习概念

机器学习概念最早由美国计算机科学家 SAMUEL 于 1959 年提出,首先将其定义为一种研究领域^[1],后来,美国计算机科学家 MITCHELL 提出了广泛认可的定义:机器学习领域涉及如何让计算机程序通过经验而自动改进的一类问题,同时书中还给出了一个较为正式的定义:对于某类任务 T (task)和性能度量 P (performance measure),计算机程序从经验 E (experience)中学习,然后他在任务 T 中的性能 P 随着经验 E 的提高而提高^[2]。总体上,机器学习

是一种基于数据的人工智能技术,通过让计算机从数据中学习规律和模式,从而实现了对未知数据的预测和分类,机器学习不需要编写规则和指令,而是通过输入数据信息进行模型建立和训练,优化算法的参数,从而实现自我学习和提高预测精度^[3]。

1.2 机器学习算法

KNUTH 对算法的定义:算法就是有穷规则的集合,其中的规则规定了解决某特定类型问题的运算序列^[4],在数学等学科中,算法是如何解决一类问题的明确规范。机器学习算法通过训练计算机系统来识别数据中的模式和规律并进行分类、预测,用于解决机器学习过程中的损失函数^[5]求解,包括最小二乘法、随机梯度下降法、牛顿法等。根据训练方式不同,机器学习算法主要分为监督学习、无监督学习、半监督学习、强化学习。比如,在监督学习中,算法从标记的训练数据中学习,然后应用所学知识于新的、未见过的数据^[6]。

1.3 机器学习模型

模型是对实际对象、系统或问题的简化描述和抽象表达,在不同领域具有不同的定义和表达方式,普遍认为模型是对现实世界的简化和抽象。机器学习模型本质上是一种接受数据作为输入并生成输出的函数。机器学习模型包括回归模型、分类模型、聚类模型、降维模型、集成模型、深度学习模型等。模型是指基于算法所建立的数学模型,与算法的主要联系和区别:通过算法可以获得模型,且同一算法可以适用于不同的模型。

2 机器学习在白酒等产地溯源研究中的应用情况

白酒产地溯源的主要研究思路是基于不同产地白酒所具有的特征进行分类、预测。针对不同产地白酒中的微量元素、乙醇等物质的碳同位素,以及白酒特征香味成分、感官等方面可能存在的差异,通过电感耦合等离子体质谱、稳定同位素质谱等仪器获取检测数据,并采用机器学习等对相关数据进行预处理、分类和预测。机器学习是人工智能和数据科学的核心之一,具有分析数据、图像识别、挖掘规律的强大功能,应用于多个领域且行之有效^[7]。机器学习在数据分析中具有效率高、准确性高等优势,在分类、预测等方面发挥着越来越重要的作用,可与检测技术

相结合应用于白酒产地溯源研究。

2.1 机器学习与质谱技术的联用

机器学习与质谱技术结合在食品、农产品真实性和产地溯源研究中应用广泛, 采用的机器学习算法主要有偏最小二乘法判别分析等, 部分案例见表 1。偏最小二乘法判别分析适用于处理分类和判别问题, 是用偏最小二乘回归方法, 对数据降维的同时建立回归模型, 并对回归结果进行判别分析, 常用于两组及以上的分类比较。质谱用于鉴定分子结构及定量分析, 为深入分析复杂化合物和大分子物质提供有效的分析方法^[17], 常与色谱等技术联用。白酒的很多产地信息可以通过质谱技术获取, 比如同一产区相同工艺的白酒往往在原料来源、配比等方面具有高度相似性, 导致白酒中乙醇等物质的稳定碳同位素或挥发性物质种类含量等可能具有产地差异, 即可通过稳定碳同位素技术及高分辨质谱技术等获取数据, 并结合偏最小二乘法判别分析等机器学习算法建模以分类、预测白酒产地。这其中, 同位素质谱技术应用效果尤为突出, 白酒中某些物质的碳同位素值可能与产地真实性有关, 稳定同位素反映的是分子水平的信息, 人为因素难以改变^[18]。通过质谱技术获取的数据, 除个别情况外(比如某些情况下可通过比较样品中某类或某几类物质的稳定碳同位素数据范围分类), 但多数情况下难以仅通过量值范围进行比较分类等, 这就需要构建机器学习模型, 借助监督学习算法等训练数据, 以达到更好的分类、预测效果。

2.2 机器学习与色谱技术的联用

机器学习与色谱技术结合在白酒产地溯源研究中应用潜力较大, 尤其在茶叶、中草药等产地溯源研究中应用

较多。在食品、农产品真实性和产地溯源研究中的部分案例见表 2, 采用的机器学习算法主要有支持向量机等。支持向量机适用于处理分类、回归等问题, 其目标是最大化决策边界边缘, 从而获得更好的泛化性。色谱技术是一种物理化学分析方法, 常与其他技术联用, 比如气相色谱与离子迁移谱联用法等^[26-28]。就白酒产地溯源研究而言, 针对白酒中的挥发性物质, 常用色谱技术有气相色谱法、液相色谱法和离子色谱法等, 通常基于不同产地白酒(工艺相同)香味成分可能存在种类含量等方面的差异, 其技术路线主要有: (1)对不同产地的白酒香味成分等进行种类含量分析, 结合主成分分析、线性判别分析等机器学习算法, 通过比较不同产地的白酒香味成分的种类及含量差异进行分类、预测; (2)对不同产地的白酒香味成分建立指纹图谱, 并结合机器学习算法进行分类、预测。此外, 白酒香味成分指纹图谱等与其产地之间很可能存在某种联系, 图谱作为一种图像, 可以结合机器学习中的深度学习技术等进行分类、预测。

2.3 机器学习与光谱技术的联用

机器学习与光谱技术结合的应用领域广泛, 在食品、农产品产地溯源研究中的部分案例见表 3, 采用的机器学习算法主要有 K 最邻近法等。K 邻近算法是一种基本的分类与回归算法, K 值的选择对于 K 最邻近法的分类结果具有重大影响。光谱技术是一种间接的分析检测手段, 无法直接从光谱信号中得出相关化合物的定性或定量信息, 其检测结果需结合有效的建模方法来体现^[39], 当前多用于白酒真实性鉴别。通过红外光谱、拉曼光谱等光谱技术能够获得大量的光谱数据, 光谱数据包含了噪声等干扰, 通常

表 1 机器学习与质谱技术联用的部分研究案例

Table 1 Partial research cases on the combination of machine learning and mass spectrometry technology

研究对象及目的	研究变量	机器学习	质谱技术	参考文献
陈皮产地溯源	Li 等 40 种元素	主成分分析、层次聚类分析、正交偏最小二乘法判别分析	电感耦合等离子体质谱法、电感耦合等离子体发射光谱法	[8]
番茄产地溯源	化合物	主成分分析、正交偏最小二乘法判别分析、双向正交偏最小二乘法判别分析	高效液相色谱-高分辨质谱法	[9]
赤霞珠葡萄酒产地溯源	矿物质元素	Fisher 判别分析	电感耦合等离子体质谱法	[10]
酱香型白酒产地溯源	Al、Fe、Zn、Rb、Sr 5 种元素	主成分分析、偏最小二乘法判别分析	电感耦合等离子体质谱法	[11]
法国红酒产地溯源	硼等微量元素、乙醇的 $\delta^{13}\text{C}$ 等	神经网络	电感耦合等离子体质谱法、稳定同位素质谱法	[12]
法国白兰地产地溯源	挥发性有机化合物指纹图谱	主成分分析、偏最小二乘法判别分析	质子转移反应质谱法	[13]
白酒产地溯源	异戊酸乙酯等挥发性成分	偏最小二乘法判别分析	气相色谱质谱法	[14]
崇明老白酒产地溯源	$\delta^2\text{H}$ 、 $\delta^{18}\text{O}$ 等 8 个变量	偏最小二乘法判别分析	稳定同位素质谱法、电感耦合等离子体发射光谱法	[15]
桃产地溯源	元素、碳同位素等	逐步判别法	电感耦合等离子体质谱法、稳定同位素质谱法	[16]

表 2 机器学习与色谱技术联用的部分研究案例

Table 2 Partial research cases on the combination of machine learning and chromatography technology

研究对象及目的	研究变量	机器学习	色谱技术	参考文献
黑茶产地识别	没食子酸等化学成分含量	支持向量机	高效液相色谱法	[19]
不同产地、不同香型白酒中活性羰基化合物的比较	7 种活性羰基化合物	聚类分析	气相色谱法	[20]
酒类真实性鉴别	香气成分	聚类分析、主成分分析	高效液相色谱法、气相色谱法	[21]
清香型白酒产地溯源	乙酸乙酯等 12 种主体香味成分	主成分分析	气相色谱法	[22]
不同产地酱香型白酒中挥发性含氮化合物的差异	挥发性含氮化合物	正交偏最小二乘法判别分析	气相色谱法	[23]
辣椒粉产地掺假	酚酸和多酚类物质	偏最小二乘法判别分析	高效液相色谱法	[24]
咖啡豆溯源	有机酸等	主成分分析、偏最小二乘判别分析	超高效液相色谱法等	[25]

表 3 机器学习与光谱技术联用的部分研究案例

Table 3 Partial research cases on the combination of machine learning and spectroscopic technology

研究对象及目的	研究变量	机器学习	光谱技术	参考文献
薏苡仁产地溯源	三维荧光光谱数据	改进的随机森林	激发-发射矩阵荧光光谱法	[29]
牦牛奶粉掺假检测与产地识别	近红外光谱数据	K 最邻近法、偏最小二乘回归法	近红外光谱法	[30]
茉莉花产地溯源	近红外漫反射光谱数据	主成分分析、线性判别分析、K 最邻近法	近红外漫反射光谱法	[31]
浓香型白酒产地溯源	比色信号	线性判别分析	紫外可见光谱法	[32]
阳山桃产地溯源	光谱指纹图谱	稀疏表示分类器	高光谱成像技术	[33]
黄芩产地溯源	光谱数据	主成分分析、卷积神经网络、支持向量机	高光谱成像技术	[34]
辣椒粉产地溯源	光谱数据	主成分分析、支持向量机等	近红外光谱技术	[35]
花生产地溯源	光谱信息等	DFFA—Net 网络结构	高光谱成像技术	[36]
蜂蜜产地溯源	光谱数据	主成分分析、数据驱动的簇类独立软模式法等	紫外—可见分光光度法	[37]
山药饮片产地溯源	光谱数据	支持向量机等	激光诱导击穿光谱法	[38]

先进行数据预处理,后借助机器学习算法对特征信息进行提取分类、预测。此外,光谱技术常与成像技术相结合,比如高光谱图像技术等,其在果实、种子等方面的应用较多,其图像信息的处理分析往往需要借助机器学习中的深度学习模型,深度学习具有自主学习、适应复杂任务、可扩展性强、泛化能力强等优点^[40]。在白酒产地溯源研究中,光谱技术与色谱技术类似,主要基于不同产地白酒(工艺相同)在光谱数据方面可能存在的差异。因光谱技术具有无损、高效等优势,与机器学习相结合时,可能在成品酒分析方面更具优势。

2.4 机器学习与智能感官技术的联用

机器学习与智能感官技术结合广泛应用于食品、农产品、中药材等领域,部分案例见表 4,采用的机器学习算法主要有卷积神经网络等。卷积神经网络是用于图像

等处理的重要深度学习模型,包括但不限于图像分类等。智能感官技术包括电子舌、电子鼻等,常与其他质谱技术等联用,其中,电子舌是 20 世纪 90 年代初期提出的一种识别,检测未知液体样品的智能检测技术,可被认为是多种溶液定量和定性分析的一种有前途的装置^[48]。感官评价是鉴别白酒品质的一个重要手段,对于白酒产品而言,电子舌技术可能更具优势,电子舌能够模拟人类味觉系统分析液体样品中的不同滋味,量化味觉信息,与人工品评相比,其优势是稳定性好、更客观且将评价结果数值化,更易于结合机器学习算法等深入分析。因此,可基于不同产地白酒在感官评价上存在的细微差异进行产地分类,该分类模型的建立,就需要选择合适的机器学习算法,两者联用在白酒产地溯源研究中具有一定的应用价值。

表 4 机器学习与智能感官技术联用的部分研究案例

研究对象及目的	研究变量	机器学习	智能感官技术	参考文献
白酒掺假鉴别	滋味(酸味等 8 个指标)	K 最邻近法、决策树	电子舌技术	[41]
三七产地鉴别	气味和滋味	主成分分析、偏最小二乘法和正交偏最小二乘法	电子鼻技术、电子舌技术	[42]
花椒产地快速溯源	电子舌及电子鼻信号	卷积神经网络、长短期记忆网络	电子鼻技术、电子舌技术	[43]
茯苓产地分类	电子鼻数据	卷积神经网络-长短期记忆网络等	电子鼻技术	[44]
石斛分类	电子鼻数据	深度卷积神经网络	电子鼻技术	[45]
六种中国白酒识别	电子鼻数据	Fisher 判别分析	电子鼻技术	[46]
春茶产地溯源	电子鼻信号	主成分分析、典则判别分析	电子鼻技术	[47]

2.5 机器学习与其他技术的联用

除上述检测技术外,机器学习还与显微镜技术等联用,多应用于医学、材料等领域。在白酒产地溯源研究中,机器学习结合显微镜技术的应用研究仍然很少,主要集中于白酒微观形态观测领域^[49],相关研究采取了原子力显微镜、扫描电子显微镜、冷冻电镜等技术。目前,制约其应用的原因主要在于:(1)显微镜技术主要观测微观形态,应用范围较窄;(2)用于观测液体样品时,因无法直接观测,均需对样品进行预处理,处理后的样品到底多大程度地反映原始信息仍然有待深入研究;(3)对于液体样品中的目标物而言,不同物质的最佳观测倍数通常并不完全一致,也会影响最终观测效果。未来,如能突破以上瓶颈,机器学习与显微镜技术有望广泛应用于白酒产地溯源研究中。

3 机器学习在白酒产地溯源研究中的应用效果

3.1 机器学习与单一检测技术

在白酒产地溯源研究中,机器学习主要是对采用检测技术获取的数据,通过建立模型进行分类、预测。不同的机器学习算法在处理数据时各具优势,这与数据点的数量和特征等密切相关。同一数据选择不同的机器学习算法时,所建模型通常在分类、预测效果方面存在差异,某一模型通常在特定检测数据上更具应用潜力,而检测数据又取决于所采用的检测技术。就白酒产品而言,基于白酒中不同指标(香味成分、某元素的稳定同位素等)区分产地时,通常需要不同的检测仪器,比如基于白酒中乙醇等物质的碳稳定同位素的产地差异,需采用同位素质谱仪获取乙醇等物质的碳稳定同位素值;若基于白酒中香味成分的产地差异,则需采用气质联用仪等。因此,基于不同的目标而采取对应的检测技术,通常获取的检测数据并不一致,此情形下,需要选择适合的机器学习算法。具体应用时,对获取的白酒相关数据,一般选择多种机器学习算法比较分析,确定适合该检测技术所获数据的最优算法,建立分类

或者预测准确率更高的机器学习模型。

3.2 机器学习与多种检测技术

在白酒产地溯源研究中,当采用单一检测技术无法应对或者应对效果不佳时,则考虑选择多种检测技术联用,既有串联也有并联。当数据为多源数据时,此时不同检测技术产生的数据在维度等方面存在差异,则数据处理相比于单一检测技术产生的数据可能更为复杂。此情形下,除与采用单一检测技术类似,涉及多种机器学习算法外,比如,在使用算法分类、预测之前,通常先对数据进行降维处理,还可能涉及数据融合技术^[50]。数据融合是集成多个数据源以产生比任何单个数据源提供的信息更一致、更准确和更有用的信息的过程^[51],数据科学与机器学习之间的联系十分紧密,数据和算法是人工智能三大支柱之二。以多种检测技术获取的数据,分析前可能需要采用数据融合技术进行预处理,此时,机器学习算法分析数据的实际效果,在一定程度上也取决于与数据融合技术的联用情况,即根据融合后的数据情况选择合适的机器学习算法。

3.3 机器学习不同算法模型

机器学习与白酒产地溯源技术的联用效果,还与多个算法的选择及联用情况密切相关。在食品等的产地溯源、质量评价等研究中,经常采用两个及两个以上的机器学习算法,其中较为常见的是主成分分析与其他算法的联用,比如主成分分析与偏最小二乘法、主成分分析与线性判别分析等联用。在溯源研究中,当出现样本量少、数据维度高的情况时,通常采用特征提取和特征选择以减少特征数量、降低维度,两者既可以单独使用也可以联用。其中,特征提取的主要方法有主成分分析等,主成分分析是一种降维统计方法,目的是减少数据维度、冗余特征等,缓解高维数据中存在的多重共线性问题,以提高分析效率,为应用其他算法建立分类模型等打下良好基础。特征选择的主要方法有过滤法等。

除上述情况外,机器学习与白酒产地溯源技术联用

效果还受真实数据量、酿造工艺、原酒贸易等影响。从真实数据量看,真实数据量与数据库构建紧密相关,可能在白酒生产企业层面建立数据库更能发挥相关检测技术的优势,更有利于相关研究应用落地;从酿造工艺看,除酿造原辅料尤其原料具有产地多源性外,比如原料之一的高粱来自东北、内蒙古,以及澳大利亚、美国等。相比于葡萄酒等发酵酒,白酒作为蒸馏酒,因其蒸馏工艺导致很多产地信息损失;从原酒贸易看,四川、贵州等白酒省份的浓、酱等香型原酒(或基酒)跨区域流通至全国各地。这些因素客观上制约了两者联用效果的更好发挥。

4 结束语

机器学习应用于白酒产地溯源研究以监督学习算法为主,实际应用过程中,为更好地分析相关数据,机器学习算法还常与遗传算法等联用。当然,机器学习尚存在一些不足,比如深度学习在图像分类识别中具有暗箱效应,不能观察整个学习过程等,可解释性差;另外机器学习模型训练过程中因数据不够等出现过拟合,以及训练效果不佳出现欠拟合等问题。鉴于白酒作为蒸馏酒的产品特性以及检测技术的发展水平,当前白酒产地溯源研究中应用的检测技术各有优劣。而机器学习在分类、预测等方面功能强大,其与检测技术的联用,可以更好地发挥各检测技术的优势,提升白酒产地溯源研究成效,具有很好的应用前景。

参考文献

- [1] SAMUEL AL. Some studies in machine learning using the game of checkers [J]. IBM J Res Dev, 1959, 3(3): 210-229.
- [2] MITCHELL T. Machine learning [Z]. 1997.
- [3] BI QF, GOODMAN KE, KAMINSKY J, *et al.* What is machine learning? A primer for the epidemiologist [J]. Am J Epidemiol, 2019, 188(12): 2222-2239.
- [4] KNUTH DE. The art of computer programming [Z]. 1997.
- [5] 李航. 统计学习方法[M]. 北京: 清华大学出版社, 2012.
LI H. Statistical learning method [M]. Beijing: Tsinghua University Press, 2012.
- [6] 张嘉晖. 蛋白质计算中的机器学习[J]. 物理学报, 2024, 73(6): 389-401.
ZHANG JH. Machine learning for in silico protein research [J]. Acta Phys Sin, 2024, 73(6): 389-401.
- [7] 仇皓雷, 王海燕. 机器学习在土壤性质预测研究中的应用进展[J/OL]. 生态学杂志, 1-15. [2024-05-11]. <http://kns.cnki.net/kcms/detail/21.1148.Q.20231120.1134.002.html>
QIU HL, WANG HY. Application of machine learning to the prediction of soil properties [J/OL]. Chin J Ecol, 1-15. [2024-05-11]. <http://kns.cnki.net/kcms/detail/21.1148.Q.20231120.1134.002.html>
- [8] 霍霏霖, 梁志森, 陈玉珍, 等. 多元素含量分析结合化学计量学对陈皮进行产地溯源[J]. 食品安全质量检测学报, 2023, 14(23): 227-233.
HUO PL, LIANG ZS, CHEN YZ, *et al.* Origin traceability of *Citri reticulatae* Pericarpium based on multi-elements content analysis combined with chemometrics [J]. J Food Saf Qual, 2023, 14(23): 227-233.
- [9] 王世成, 明旸, 张炜, 等. 基于高效液相色谱-高分辨质谱法的非靶标代谢组学方法分析鉴别不同产地番茄[J]. 食品安全质量检测学报, 2023, 14(20): 26-32.
WANG SC, MING Y, ZHANG W, *et al.* Analysis and identifying geographic origins of *Solanum lycopersicum* by untargeted metabolomics approach based on high performance liquid chromatography-high resolution mass spectrometry [J]. J Food Saf Qual, 2023, 14(20): 26-32.
- [10] 吕真真, 谢辉, 庞荣丽, 等. 基于矿质元素的赤霞珠葡萄酒产地溯源[J]. 食品研究与开发, 2023, 44(19): 159-164.
LV ZZ, XIE H, PANG RL, *et al.* Origin tracing of cabernet sauvignon wine based on mineral elements [J]. Food Res Dev, 2023, 44(19): 159-164.
- [11] 王丽, 卢君, 山其木格, 等. 赤水河流域不同地区酱香型白酒矿质元素特征分析及产地辨识初探[J]. 中国酿造, 2021, 40(9): 58-62.
WANG L, LU J, SHAN QMG, *et al.* Analysis of mineral elements and origin identification of sauce-flavor Baijiu in different regions of Chishui-river basin [J]. China Brew, 2021, 40(9): 58-62.
- [12] WU H, LIN GH, TIAN L, *et al.* Origin verification of French red wines using isotope and elemental analyses coupled with chemometrics [J]. Food Chem, 2020, 339: 127760.
- [13] MALFONDET N, BRUNERIE P, QUERE JL. Discrimination of French wine brandy origin by PTR-MS headspace analysis using ethanol ionization and sensory assessment [J]. Anal Bioanal Chem, 2021, 413: 3349-3368.
- [14] 李爱兰, 乔文君, 汪兴财, 等. 吹扫捕集-气质联用结合化学计量学用于不同产地和品牌浓香型白酒鉴别[J]. 化学试剂, 2021, 43(5): 576-583.
LI AIL, QIAO WJ, WANG XC, *et al.* Purge and trap-GC/MS combined with chemometrics for identification of strong aroma Baijiu from different origins and brands [J]. Chem Reag, 2021, 43(5): 576-583.
- [15] 刘星, 张其才, 姚春霞, 等. 崇明老白酒产地溯源判别[J]. 食品与机械, 2020, 36(11): 77-82, 96.
LIU X, ZHANG QC, YAO CX, *et al.* Geographical origin traceability discrimination of Chongming rice wine [J]. Food Mach, 2020, 36(11): 77-82, 96.
- [16] 习佳林, 郭阳, 李安, 等. 基于多元素和稳定同位素技术的桃产地溯源[J]. 食品安全质量检测学报, 2024, 15(9): 62-68.
XI JL, GUO Y, LI AN, *et al.* Traceability of the origin of *Prunus persica* based on multi-element and stable isotope techniques [J]. J Food Saf Qual, 2024, 15(9): 62-68.
- [17] 岳如玉, 慕宇, 范慧艺, 等. 质谱技术在乳蛋白糖化分析中的应用研究进展[J/OL]. 中国乳品工业: 1-8. [2024-05-12]. <http://kns.cnki.net/kcms/detail/23.1177.TS.20240423.1514.002.html>
YUE RY, MU Y, FAN HY, *et al.* Research progress on application of mass spectrometry in glycation analysis of milk protein [J/OL]. Chin Dairy Ind: 1-8. [2024-05-12]. <http://kns.cnki.net/kcms/detail/23.1177.TS.20240423.1514.002.html>
- [18] ROSSMANN A. Determination of stable isotope ratios in food analysis [J]. Food Rev Int, 2007, 17(3): 347-381.
- [19] 王贞红, 韩沅汐, 张立友, 等. 基于高效液相色谱指纹图谱结合化学计量学及机器学习的黑茶产地识别[J]. 食品安全质量检测学报, 2023,

- 14(18): 50–58.
WANG ZH, HAN YX, ZHANG LY, *et al.* Identification of dark tea origin based on high performance liquid chromatography fingerprint combined with chemometrics and machine learning [J]. *J Food Saf Qual*, 2023, 14(18): 50–58.
- [20] 司波, 袁雯雯, 贾梦玮, 等. 气相色谱-电子捕获检测器法分析不同产地、不同香型白酒中活性羰基化合物[J]. *食品与发酵工业*, 2021, 47(22): 241–248.
SI B, YUAN WW, JIA MW, *et al.* Reactive carbonyl species in Chinese Baijiu from different origins and flavor types by gas chromatography-electron capture detector [J]. *Food Ferment Ind*, 2021, 47(22): 241–248.
- [21] 陈焱洁, 陈少敏, 刘鸿钢, 等. 市售酒香精成分分析在酒类真实性鉴别中的应用[J]. *食品安全质量检测学报*, 2020, 11(16): 5440–5447.
CHEN DJ, CHEN SM, LIU HG, *et al.* Application of flavoring component analysis in liquor authenticity identification [J]. *J Food Saf Qual*, 2020, 11(16): 5440–5447.
- [22] 李艳敏, 张立严, 狄红梅. 主成分和判别分析在清香型白酒产地溯源中的应用[J]. *中国酿造*, 2018, 37(1): 145–148.
LI YM, ZHANG LY, DI HM. Application of principal component analysis and discriminant analysis in origin traceability of light-flavor Baijiu [J]. *China Brew*, 2018, 37(1): 145–148.
- [23] 吴李玲, 裴荣红, 李行, 等. 不同产地酱香型白酒中挥发性含氮化合物的差异分析[J]. *食品科学技术学报*, 2024, 42(3): 35–48.
WU LL, PEI RH, LI X, *et al.* Difference analysis of volatile nitro-containing compounds in soy sauce flavor-type Baijiu from different regions in China [J]. *J Food Sci Technol*, 2024, 42(3): 35–48.
- [24] CAMPMAJO G, RODRIGUEZ-JAVIER LR, SAURINA J, *et al.* Assessment of paprika geographical origin fraud by high-performance liquid chromatography with fluorescence detection (HPLC-FLD) fingerprinting [J]. *Food Chem*, 2021, 352: 129397.
- [25] 翟慧楠, 于菲, 唐玉妹, 等. 基于主要滋味成分指纹图谱和化学计量学鉴别海南不同地区咖啡豆[J]. *热带作物学报*, 2024, 45(5): 1007–1020.
ZHAI HN, YU F, TANG YM, *et al.* Identification of Hainan robusta coffee origins based on main taste components fingerprint and chemometrics [J]. *Chin J Trop Crop*, 2024, 45(5): 1007–1020.
- [26] 庞珂靖, 万国超, 刘振平, 等. 气相色谱-离子迁移谱结合化学计量学分析对新会陈皮的鉴别[J]. *食品科学*, 2024, 45(13): 275–281.
PANG KJ, WAN GC, LIU ZP, *et al.* Identification of Xinhui Chenpi (*Citri Reticulatae Pericarpium*) by gas chromatography-ion mobility spectrometry combined with chemometric analysis [J]. *Food Sci*, 2024, 45(13): 275–281.
- [27] 何婉琳, 施露, 林梦桦, 等. 基于气相色谱-离子迁移谱分析不同产地秋月梨品质差异[J]. *食品科学*, 2024, 45(5): 118–125.
HE WL, SHI L, LIN MH, *et al.* Quality analysis of ‘Akizuki’ pear fruit (*Pyrus pyrifolia*) from different geographical origins by gas chromatography-ion mobility spectrometry [J]. *Food Sci*, 2024, 45(5): 118–125.
- [28] 张碧莹, 隋雨萌, 张鑫, 等. 气相离子迁移谱技术在食品真实性鉴别中的应用[J/OL]. *食品科学*, 1-14. [2024-07-31]. <http://kns.cnki.net/kcms/detail/11.2206.TS.20240415.1129.002.html>
ZHANG BY, SUI YM, ZHANG X, *et al.* Application of gas chromatography-ion mobility spectrometry in food authentication [J/OL]. *Food Sci*, 1-14. [2024-07-31]. <http://kns.cnki.net/kcms/detail/11.2206.TS.20240415.1129.002.html>
- [29] 赵汉卿, 王斌, 陈瑶, 等. 基于改进随机森林算法的薏苡仁产地溯源研究[J]. *轻工学报*, 2023, 38(6): 70–77.
ZHAO HQ, WANG B, CHEN Y, *et al.* Research on geographical origin traceability of coix seed based on a modified random forest algorithm [J]. *J Light Ind*, 2023, 38(6): 70–77.
- [30] 彭海洋, 巫忠东, 林涛, 等. 基于近红外光谱技术的牦牛奶粉掺假检测与产地识别研究[J/OL]. *食品工业科技*, 1-14. [2024-06-27]. <https://doi.org/10.13386/j.issn1002-0306.2024020297>
PENG HY, WU ZD, LIN T, *et al.* Study on adulteration detection and origin identification of yak milk powder based on near-Infrared spectroscopy technology [J/OL]. *Food Ind Sci Technol*, 1-14. [2024-06-27]. <https://doi.org/10.13386/j.issn1002-0306.2024020297>
- [31] 张淑芳, 雷蕾, 雷顺新, 等. 近红外漫反射光谱的茉莉花产地溯源[J]. *光谱学与光谱分析*, 2023, 43(11): 3389–3395.
ZHANG SF, LEI L, LEI SX, *et al.* Traceability of geographical origin of jasmine based on near infrared diffuse reflectance spectroscopy [J]. *Spectrosc Spect Anal*, 2023, 43(11): 3389–3395.
- [32] 贾俊杰, 张宿义, 许涛, 等. 基于银纳米三角片蚀刻机制区分不同产地浓香型白酒[J]. *食品工业科技*, 2024, 45(10): 254–262.
JIA JJ, ZHANG SY, XU T, *et al.* Discrimination of strong aroma-type Baijiu of different origins based on Ag nanoprisms etching mechanism [J]. *Food Ind Sci Technol*, 2024, 45(10): 254–262.
- [33] SUN Y, LI YH, PAN LQ, *et al.* Authentication of the geographic origin of Yangshan region peaches based on hyperspectral imaging [J]. *Postharvest Biol Technol*, 2021, 171: 111320.
- [34] XIAO QL, BAI XL, GAO P, *et al.* Application of convolutional neural network-based feature extraction and data fusion for geographical origin identification of *Radix astragal* by visible/short-wave near-infrared and near infrared hyperspectral imaging [J]. *Sensors*, 2020, 20(17): 4940.
- [35] MEENA D, CHAKRABORTY S, MITRA J. Geographical origin identification of red chili powder using NIR spectroscopy combined with SIMCA and machine learning algorithms [J]. *Food Anal Methods*, 2024, 17(7): 1005–1023.
- [36] 石岩, 任宇琪, 王思远, 等. 自适应融合气体-光谱双模态信息花生产地溯源方法[J]. *农业机械学报*, 2024, 55(4): 176–183, 203.
SHI Y, REN YQ, WANG SY, *et al.* Adaptive fusion of gas spectral bimodal information for gas peanut origin traceability [J]. *Trans Chin Soc Agric Mach*, 2024, 55(4): 176–183, 203.
- [37] DIMAKOPOULOU-PAPAZOGLU D, PLOSKAS N, KOUTSOUMANIS K, *et al.* Identification of geographical and botanical origin of Mediterranean honeys using UV-vis spectroscopy and multivariate statistical analysis [J]. *J Food Meas Charact*, 2024, 18(5): 3923–3934.
- [38] 蔡羽, 赵志方, 郭连波, 等. 基于 LIBS 的山药饮片产地溯源研究[J]. *光谱学与光谱分析*, 2023, 43(1): 138–144.
CAI Y, ZHAO ZF, GUO LB, *et al.* Research on origin traceability of *Rhizoma dioscoreae* based on LIBS [J]. *Spectrosc Spect Anal*, 2023, 43(1): 138–144.
- [39] 翟双, 虞先国, 张贵宇, 等. 基于 FT-NIR 光谱技术结合 KPCA-MD-SVM 对白酒基酒的快速判别[J]. *现代食品科技*, 2022, 38(4): 248–253.
ZHAI S, TUO XG, ZHANG GY, *et al.* Rapid discrimination of base liquor

- for Baijiu based on FT-NIR spectroscopy and KPCA-MD-SVM [J]. *Mod Food Sci Technol*, 2022, 38(4): 248–253.
- [40] SU JM, ZHANG FP, YU CX, *et al.* Machine learning: Next promising trend for microplastics study [J]. *J Environ Manage*, 2023, 344: 118756.
- [41] 程铁轅, 夏于林, 张莹. 基于机器学习和电子舌技术的白酒掺假鉴别[J]. *食品工业*, 2021, 42(5): 288–291.
CHENG TY, XIA YL, ZHANG Y. Identification of adulteration of Chinese liquor based on machine learning and electronic tongue technology [J]. *Food Ind*, 2021, 42(5): 288–291.
- [42] 任玲慧, 张诗焉, 郭宜欣, 等. 基于电子鼻和电子舌融合技术的三七产地鉴别[J]. *现代中药研究与实践*, 2024, 38(1): 1–6.
REN LH, ZHANG SY, GUO YX, *et al.* Origin identification of *Panax notoginseng* based on fusion technology of electronic nose and tongue [J]. *Chin Med J Res Prac*, 2024, 38(1): 1–6.
- [43] 张擎, 杨晓婧, 金鑫宁, 等. 基于电子鼻和电子舌与1D-CNN-LSTM模型的花椒产地快速溯源检测[J]. *传感技术学报*, 2024, 37(5): 904–912.
ZHANG Q, YANG XJ, JIN XN, *et al.* Rapid traceability detection of *Zanthoxylum bungeanum* origin based on electronic nose and electronic tongue with 1D-CNN-LSTM model [J]. *Chin J Sens Actuator*, 2024, 37(5): 904–912.
- [44] 徐昊, 章检明, 王中鹏, 等. 基于电子鼻的深度卷积神经网络茯苓产地分类方法[J]. *传感器与微系统*, 2023, 42(12): 138–141.
XU H, ZHANG JM, WANG ZP, *et al.* DCNN for *Poria cocos* origin classification method based on E-nose [J]. *Transduc Microsyst Technol*, 2023, 42(12): 138–141.
- [45] WANG Y, DIAO J, WANG Z, *et al.* An optimized deep convolutional neural network for dendrobium classification based on electronic nose [J]. *Sens Actuator A*, 2020, 307(1): 111874.
- [46] YU HC, YIN Y, YUAN YX, *et al.* A KECA identification method based on GA for E-nose data of six kinds of Chinese spirits [J]. *Sens Actuators B*, 2021, 333(2): 129518.
- [47] 黄汇惠, 闫莎莎, 靳冬武, 等. 基于电子鼻技术的湄潭春茶产地溯源研究[J]. *传感技术学报*, 2023, 36(5): 825–832.
HUANG HH, YAN SS, JIN DW, *et al.* Study on origin tracing of Meitan spring tea based on electronic nose technology [J]. *Chin J Sens Actuator*, 2023, 36(5): 825–832.
- [48] WU X, TAHARA Y, YATABE R, *et al.* Taste sensor: Electronic tongue with lipid membranes [J]. *Anal Sci*, 2020, 36(2): 147–159.
- [49] 程铁轅. 冷冻电镜技术在白酒微观形态观测中的应用[J]. *酿酒科技*, 2021, (2): 49–54, 58.
CHENG TY. Application of cryo-electron microscopy in the observation of micromorphology of Baijiu [J]. *Liquor-Making Sci Technol*, 2021, (2): 49–54, 58.
- [50] 邓焯文, 陈喆, 付家顺, 等. 数据融合策略在食品产地溯源中的应用进展[J]. *分析化学*, 2023, 51(1): 11–21.
DENG ZW, CHEN Z, FU JS, *et al.* Application progress of data fusion strategy in food origin traceability [J]. *Chin J Anal Chem*, 2023, 51(1): 11–21.
- [51] HAGHIGHAT M, ABDEL-MOTTALEB M, ALHALABI W. Discriminant correlation analysis: Real-time feature level fusion for multimodal biometric recognition [J]. *IEEE T Inf Fore Sec*, 2016, 11(9): 1984–1996.

(责任编辑: 于梦娇 韩晓红)

作者简介

程铁轅, 正高级工程师, 主要研究方向为食品安全检测。
E-mail: chengtieyuan@yeah.net

胡江涛, 正高级工程师, 主要研究方向为食品安全检测。
E-mail: 11166589@qq.com