工程科学学报
Chinese Journal of Engineering

机器学习在有机固体废物资源化的应用进展

马鸿志 刘忆婵 赵继华 费凡 高明 汪群慧

Advances in machine learning applications to resource technology for organic solid waste

MA Hongzhi, LIU Yichan, ZHAO Jihua, FEI Fan, GAO Ming, WANG Qunhui

引用本文:
马鸿志, 刘忆婵, 赵继华, 费凡, 高明, 汪群慧. 机器学习在有机固体废物资源化的应用进展[J]. 北科大：工程科学学报, 2025, 47(3): 550–561. doi: 10.13374/j.issn2095–9389.2024.07.10.001
MA Hongzhi, LIU Yichan, ZHAO Jihua, FEI Fan, GAO Ming, WANG Qunhui. Advances in machine learning applications to resource technology for organic solid waste[J]. *Chinese Journal of Engineering*, 2025, 47(3): 550–561. doi: 10.13374/j.issn2095–9389.2024.07.10.001

在线阅读 View online: https://doi.org/10.13374/j.issn2095–9389.2024.07.10.001

工程科学学报, 第 47 卷, 第 3 期: 550−561, 2025 年 3 月
Chinese Journal of Engineering, Vol. 47, No. 3: 550−561, March 2025
https://doi.org/10.13374/j.issn2095-9389.2024.07.10.001; http://cje.ustb.edu.cn

机器学习在有机固体废物资源化的应用进展

马鸿志[1,2,3)], 刘忆婵[1,2)], 赵继华[1,2)], 费　凡[1,2)✉], 高　明[1,2)], 汪群慧[1,2)]

1) 北京科技大学能源与环境工程学院, 北京 100083　2) 工业典型污染物资源化处理北京市重点实验室, 北京 100083　3) 伊犁师范大学资源与环境学院, 伊宁 835000

✉通信作者, E-mail: feifan@ustb.edu.cn

摘　要　机器学习(ML)方法, 以其卓越的数据解析和模式识别能力, 已在有机固体废物(OSW)处理领域展现出显著的应用潜力. 随着对 OSW 处理需求的日益增长及技术革新的推进, ML 在该领域的应用正迅速普及. 聚焦 ML 技术在 OSW 资源化处理中的应用, 首先界定了 OSW 的范畴, 针对 OSW 处理中存在的异质性和复杂性问题, 指出了传统处理技术在进行 OSW 产量预测和条件优化时的局限性. 通过对 2018—2023 年相关学术成果进行系统梳理和分析, 揭示了 ML 在 OSW 处理中的研究趋势和潜力. 特别是发现以人工神经网络(ANN)、支持向量机(SVM)、决策树(DT)、随机森林(RF)和极端梯度提升(XGBoost)为代表的常用模型结合遗传(GA)优化算法, 成为提高 OSW 处理效率和资源回收率的研究热点. 分析了这些模型在源头产生与分类、热化学转化处理、厌氧生物处理和好氧堆肥等具体应用中的现状及应用频率, 同时评估了它们的优缺点及适用性. 研究发现, ML 技术能够有效提高 OSW 处理的预测精度和工艺优化能力, 尤其是在废物特性预测和生物处理过程模拟方面展现出显著优势. 然而, 数据质量、模型的泛化能力以及算法选择仍然是 ML 技术应用中的关键挑战. 为此, 提出开发综合模型、加强跨学科技术融合等一系列解决策略, 以期为 OSW 资源化提供科学指导和技术支持.

关键词　有机固体废物; 机器学习; 资源化; 高效处理; 建模预测; 过程优化

分类号　TP181;X705

Advances in machine learning applications to resource technology for organic solid waste

MA Hongzhi[1,2,3)], *LIU Yichan*[1,2)], *ZHAO Jihua*[1,2)], *FEI Fan*[1,2)✉], *GAO Ming*[1,2)], *WANG Qunhui*[1,2)]

1) School of Energy and Environmental Engineering, University of Science and Technology Beijing, Beijing 100083, China

2) Beijing Key Laboratory of Resource-oriented Treatment of Industrial Pollutants, Beijing 100083, China

3) School of Resources and Environment, Yili Normal University, Yining 835000, China

✉Corresponding author, E-mail: feifan@ustb.edu.cn

ABSTRACT　Machine learning (ML) techniques, with their advanced data analysis and pattern recognition capabilities, are highly effective for addressing the complexities of organic solid waste (OSW) treatment and resource recovery. As global waste generation continues to increase, the need for efficient and sustainable OSW management solutions is growing. Traditional waste treatment technologies often face challenges in managing the heterogeneous and complex nature of OSW, which varies widely in composition. In contrast, ML can optimize treatment processes, improve resource recovery rates, and enhance decision-making. This study explores a range of commonly used ML models, including artificial neural network (ANN), support vector machine (SVM), decision tree, random forest, and extreme gradient boosting (XGBoost). These models have been used to predict waste characteristics, classify diverse types of

收稿日期: 2024−07−10

基金项目: 国家重点研发计划资助项目(2022YFE0105700); 鄂尔多斯市科技计划项目(ZD20232319); 国家自然科学基金青年科学基金资助项目(72004067); 国家重点实验室专项基金课题(22K02ESPCT); 新疆维吾尔自治区重大科研专项(2022B02021)

OSW, and optimize treatment parameters across various processes, such as thermochemical conversion, anaerobic digestion, and aerobic composting. A key focus of this work is the combination of ML models with optimization algorithms like Genetic Algorithm, which improves the performance of ML models by optimizing hyperparameters and enhancing prediction accuracy. This approach is particularly useful in complex processes such as biological treatment and resource recovery, where ML models can predict waste characteristics and optimize treatment conditions. This work also presents a comprehensive analysis of the application frequency of these ML models in various stages of OSW treatment, including source generation, classification, and treatment processes like pyrolysis, gasification, and composting. This analysis identifies the strengths and weaknesses of each model, highlighting the importance of selecting the most appropriate ML approach based on the specific characteristics of the OSW treatment task. ANN, for example, is particularly useful for complex, nonlinear relationships within biological treatment processes, while SVM is effective for small datasets and high-dimensional data. Despite the promise of ML in OSW management, there are key challenges that remain unresolved. These include issues related to data quality, such as missing or incomplete datasets, and the generalization ability of ML models across different treatment scenarios. Furthermore, selecting the right ML model for a specific task requires careful consideration of the data structure, the complexity of the problem, and the desired outcomes. The full potential of ML in OSW treatment may not be realized without addressing these challenges. This work proposes strategies for overcoming these challenges and improving the effectiveness of ML in OSW treatment. One strategy involves developing integrated models that combine multiple ML techniques to leverage their respective strengths. For example, the ensemble learning method, which integrates the outputs of multiple models, has been demonstrated to improve prediction accuracy and robustness. Another strategy is the use of reinforcement learning and transfer learning, which effectively address dynamic environments and small datasets, respectively. Finally, this work highlights the need for future research to focus on the integration of ML models with real-time process monitoring and control systems. By linking ML with data-driven control strategies, such as model predictive control, it may be possible to develop fully automated, intelligent OSW treatment systems that optimize resource recovery and minimize environmental impact. The work concludes by recommending that researchers continue exploring the combinations of ML with advanced control techniques to push the achievement boundaries in sustainable waste management.

KEY WORDS organic solid waste; machine learning; resource utilization; efficient treatment; modeling and simulation; process optimization

有机固体废物(OSW)是在生产和生活活动中产生的含有有机成分的固体废物,我国的OSW包含餐厨垃圾、城市污泥等生活源,农业秸秆、畜禽粪污等农业源,菌渣、酒糟等工业源废弃物,涉及面广,总量位居全球首位[1]. OSW具有污染物和资源的双重属性,若处理不当,将会造成严重的环境污染和资源浪费[2]. 目前,OSW的有效处理方法主要包括生化处理(如好氧堆肥、厌氧生物处理)和热转化处理(如焚烧、热解、气化)[3−5]. 两种处理方法各有优势,但由于OSW的来源多样性与成分异质性[6],处理效果波动较大,工艺条件参数设置显著影响处理效率,这无疑增加了处理技术的难度. 鉴于此,本文将探讨提高OSW处理效率和效果的策略. 首先,从废物管理的源头分析,揭示废物特性模式和趋势,为后续处理提供先验知识;其次,构建基于数据驱动的预测模型,估算不同来源OSW的产量,模拟处理过程,优化操作参数,实现产量和效率的最大化.

传统建模方法在处理OSW问题时存在局限性,特别是在成分异质性导致的实验结果差异和预测精度降低方面. 例如Patra和Lalhriatpuii[7]采用的线性回归模型在预测甲烷排放时对数据的高精度要求和计算量,在异常值面前缺乏鲁棒性等. 近年来,机器学习(ML)作为人工智能的分支,展示出强大的特征表示能力,可以克服传统方法缺乏全局理解的局限,推动OSW处理的智能化、精细化和高值化[8−10]. ML是一种能够自主获取和整合知识的系统,近年来在水气固废物处理等环境领域正逐渐展现出其巨大的潜力. ML的一般原理是通过归纳推理,学习经验数据中的输入和输出变量之间的关系,训练得出某种模型并以此在新情况下做出决策[11],在处理复杂非线性问题时具备节省时间、提高预测精度以及降低实验成本等优势.

目前,已有多种ML算法应用于OSW管理与处理,如神经网络用于废物特性评估、图像识别用于垃圾分类、回归算法用于厌氧发酵和生物质转化的模拟与优化. 尽管训练阶段需投入计算资源,但其长期效益显著超越初期成本. 因此,ML在OSW循环利用中具有巨大的应用前景,值得深入研究.

本研究旨在探讨 ML 技术在 OSW 资源化领域的研究现状和发展趋势, 分析不同 ML 模型在物理、化学和生物处理中的适用性, 评估其优势和局限性, 为 OSW 的机器学习辅助处理提供全面的技术路线图.

1　机器学习在有机固体废物领域的模型构建

模型构建是 OSW 资源化的关键环节, 通过模拟和分析废物处理过程, 为高效回收和可持续发展提供科学依据. 如图 1 所示, 一个完整的 ML 过程通常涉及四个关键阶段: 训练、交叉验证、测试和应用.

模型构建的基础在于数据的质量和结构, 需要建立一个详尽的数据库, 涵盖物理化学特性、微生物种群、生物降解参数、环境影响评估等, 并结合文献和实验数据进行特征选择. 数据预处理包括缺失值填补、异常值剔除、数据归一化和特征选择, 以确保数据一致性和高质量.

选择适合的 ML 模型进行训练是关键. 基础 ML 模型, 如决策树(DT)、支持向量机(SVM)和人工神经网络(ANN), 在不同处理环节中已有广泛应用. 例如, SVM 在分类任务中表现优异, 而随机森林(RF)在应对多变量和复杂性问题上优势明显. 近年来深度学习(DL)模型, 特别是卷积神经网络(CNN)能够自动提取图像特征, 大幅提升分类精度, 展现了在垃圾分类任务中的巨大潜力[12]. 此外, 集成学习方法, 如 RF 和极限梯度提升(XGBoost), 通过集成多个基本学习器来提高预测准确性和处理效率, 特别是在应对数据不平衡问题时表现尤为出色[13].

在模型的训练过程中, 超参数调优也是至关重要的环节. 超参数的选择直接影响模型的预测精度和泛化能力. 例如, 在 XGBoost 模型中, 优化基本学习器的超参数能够显著提高资源回收率的预测精度. 在 CNN 中, 通过调整网络层数和卷积核大小, 可以有效提升垃圾分类中的图像处理效果. 交叉验证通常用于优化超参数, 以防模型过拟合, 从而增强模型的泛化能力.

模型的评估一般通过测试集数据的决定系数(R^2)来衡量, R^2值越接近 1, 表明模型的预测准确性越高. 经过验证的模型可以应用于 OSW 处理的不同环节, 为工艺流程的优化提供决策支持. 最终, 借助特征重要性评估和局部解释模型等方法, 揭示模型的决策过程, 为 OSW 处理提供更深入的理解, 从而指导实际工艺流程的优化设计, 提高资源回收率和处理效率[14].

2　机器学习在有机固体废物领域的应用现状

为了解机器学习(ML)在有机固体废物(OSW)资源化领域的研究现状, 本研究在 Web of Science、Science Direct 等数据库中, 以“机器学习(machine learning)”、“有机固体废弃物(organic solid waste)”、“生物质(biomass)”、“预测(prediction)”、“算法(algorithm)”、“模型优化(model optimization)”和“处理(treatment)”等为关键词, 检索了 2018 至 2023 年间的相关学术成果. 通过筛选主题相关的文章, 统计了文献数量和研究领域的分布, 结果如图 2 所示.

图 2 展示了 OSW 处理相关文献数量在 2018 至 2023 年间的增长趋势. 从图 2(a) 可见, ML 在 OSW 资源化中的研究热度逐年增加, 热点集中在“废物产生与分类”和“热化学处理”领域, 尤其是厌氧消

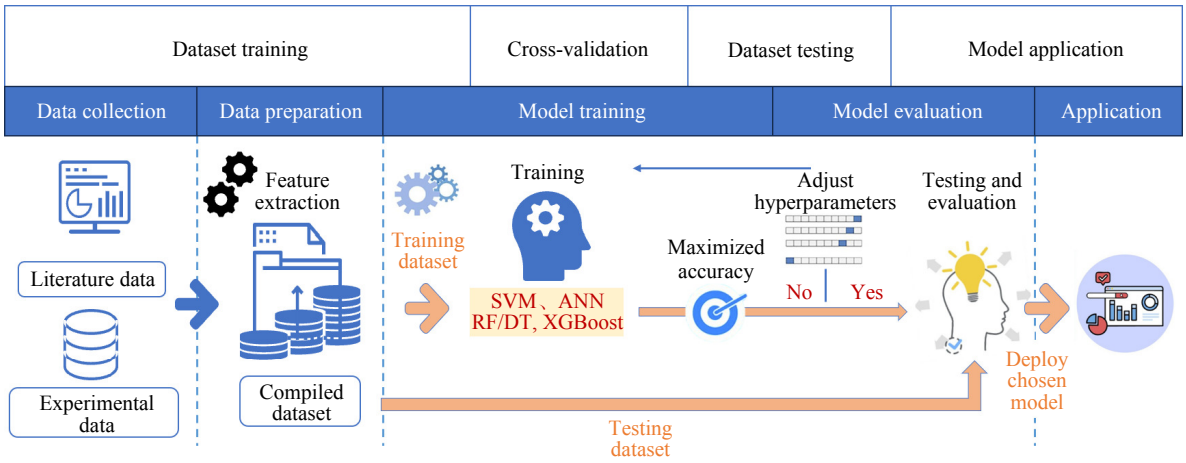

图 1　ML 在 OSW 中的模型构建与应用过程

Fig.1　Machine learning modeling and applications in the organic solid waste field

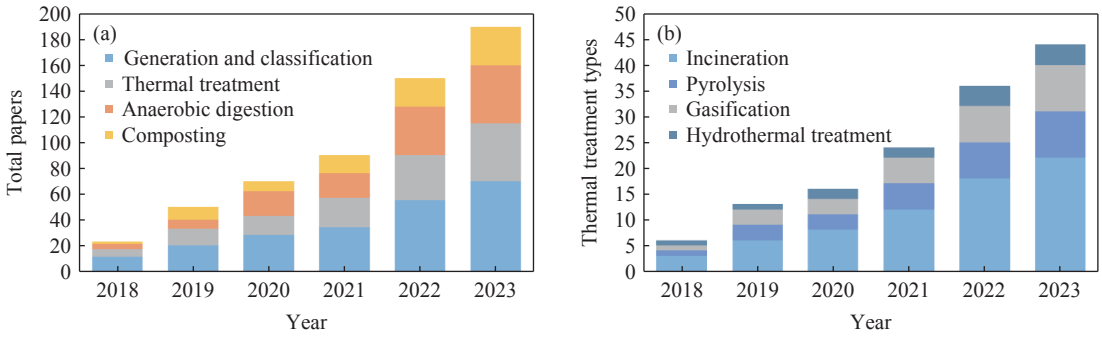

图 2 相关文章的分布与趋势统计. (a) OSW 处理; (b) 热处理技术

Fig.2 Distribution and trend statistics of related articles in organic solid waste (OSW) treatment: (a) OSW treatment; (b) heat treatment technology

化和好氧堆肥. 随着对微生物生态和生物转化过程的深入理解, ML 在优化厌氧发酵产物(如中长链脂肪酸)、热化学处理产物(如生物燃料)及堆肥过程中的生物炭应用等方面展现出重要作用[15−17]. ML 的介入不仅提升了废物处理效率, 还推动了高值化产品的开发. 图 2(b) 进一步细分了热化学处理技术的文献分布情况, 包括焚烧、热解、气化和水热处理. 数据显示, 焚烧相关研究占据热化学处理文献的最大比例, 显示其在 OSW 处理中的重要性. 然而, 气化技术的文献数量显著增加, 尤其是在能源回收方面逐渐成为热点. 热解技术在转化为高附加值产品方面显示出潜力, ML 模型为工艺优化和产物预测提供了可靠工具. 相比之下, 水热处理文献较少, 但因其在处理高水分废物中的优势, 未来仍具发展空间

综上, ML 技术在 OSW 资源化处理中已取得显著进展, 研究方向由基础处理技术向复杂废物转化过程转变, 着重提高资源回收率与环境效益.

2.1 产生与分类

我国城市生活垃圾(MSW)中有机成分约占总量的 60%, 高水分含量影响了垃圾的热值和分类回收效率. 针对 OSW 的管理, 需要根据成分和处理方式进行区分, 确保有效管理. 管理过程会涉及产生、收集、存储、分类等多个环节, 合理利用 ML 方法可以对相应的环节进行建模与优化. 近年来, ANN 模型在废物产生量预测中表现出优越性. Kulisz 和 Kujawska 开发了一个 ANN 模型来预测波兰废物产生量[18], 研究综合了社会、经济和人口因素, 预测 R^2 达到 0.989. 类似地, Elshaboury 等[19] 采用相同特征, 集成粒子群优化(PSO)算法提高了 ANN 的预测精度. 此外, Soni 等[20] 比较了六种不同的 ML 模型, 通过将 ANN 与遗传算法(GA)集成准确估计了印度新德里废物产量(R^2=0.87). 这些研究结果不仅证实了 ANN 模型在废物产量预测方面的有效性, 还展示了通过与其他算法集成可以进一步提升预测精度的潜力.

随着垃圾分类的细化, OSW 分类研究逐渐受到重视. 例如 Zheng 等[21] 使用 Fisher 判别模型通过高光谱成像对聚乙烯(PP)、聚氯乙烯(PVC)等有机塑料废弃物进行分类, 预测精度达 100%. Serranti 等[22] 则通过红外高光谱成像系统(NIR-HIS)完成了建筑垃圾中聚烯烃垃圾的分类. 还有研究聚焦在具备复杂成分的 OSW 的分类与理化特性预测上. Toğaçar 等[23] 在 2020 年首次提出了一种精确度高达 99.95% 的卷积神经网络(CNN), 对可回收垃圾以及有机垃圾进行分类. Tao 等[24] 在 2023 年提出了结合 ML 模型与高光谱成像技术的新概念设计, 实现了垃圾分类及快速表征. 在参数与界面优化后, ANN 模型对 MSW 中无机/有机组分的分类精确度接近 100%, 对有机组分的 C、H、O、N 元素以及热值预测的精准率达到 90% 以上. 这些成果为 OSW 的物理分拣系统设计提供了新思路, 并有助于热化学处理的过程管理以及生物处理的路线决策.

2.2 热化学处理

热化学处理涉及到 OSW 的污染排放、能量回收与化学转化, 是减量化、能源化、资源化的关键, 通过焚烧、热解、气化等方式将 OSW 转化为热能, 并输出电力或蒸汽等能源. 此外, 过程中还能产生生物炭、液体燃料等有价值的化学物质, 在化学品、肥料和能源生产中具有重要应用. ML 技术为 OSW 的高效热化学转化提供了工艺建模的新途径.

在减量化方面, 研究多集中于焚烧过程中污染物排放的预测与分析, 例如一氧化碳[25] 与二噁英[26]. Tavoosi 等[27] 利用非线性神经网络(RBFNN)模型对生物质焚烧炉排放进行预测, 有研究进一步提出了一种多输出生物质控制系统的模型预测

控制[28], 有效提高了生物质锅炉系统的控制性能. ML 技术还被用于预测垃圾焚烧飞灰中的重金属含量, 识别重金属的浸出模式[29], 评估重金属固化效果的影响因素[30], 为飞灰的合理处置和利用提供依据.

在能源化方面, 精确预测热解过程的固液气三相产物特性, 是实现资源化调控的基础. Li 等[31]开发的 ANN 综合模型成功预测了合成气的组成和效率. Zhao 等[32] 构建的氧化热解产物分布预测模型 R^2高达 0.992. Leng 等[33] 建立了三相产物分布和生物油热值预测模型, 并确定了 480 °C 为鼓泡流化床中快速热解制油的最佳温度. Shu 等[34] 建立了基于干基或湿基物理构成与元素分析的多层感知神经网络(MLP–ANN)模型, 用于城市生活垃圾的低温热值预测, 模型精度高达 98%. Taki 和 Rohani[35] 对城市固废的高位热值进行建模, 结果显示径向基函数神经网络(RBF–ANN)模型的预测准确性最高. 这些研究表明, ML 在生物质气化和热值预测方面具有重要应用潜力, 能够有效提升气相产物的品质.

在资源化方面, 生物质在热化学处理中产生的生物油和生物炭具有重要应用价值. 生物油可以进一步加工为生物燃料、生物化学品和生物材料等高价值产品[36]. Aghbashlo 等[37] 最近回顾了 ML 在生物油研究中的应用, 例如 Tang 等[38] 研究发现 RF 模型在预测生物油产量及其氢含量方面优于多元逻辑回归(MLR)模型. 除了生产过程方面的显著研究进展, 对生物油的转化和利用过程的深入研究同样至关重要. Ma 等[39] 探索了使用乳化技术在较低温度下生产乳化生物沥青的可能性, 这为生物油的可持续高值利用提供了一种新方法. 结合 ML 技术提出将体积分数为 10.37% 的生物油和质量分数为 3.53% 的乳化剂结合使用的一种最佳乳化工艺. 此外, 生物油的生产成本也是商业化的一个重大问题, Liu 等[40] 成功对藻油和甲醇制备生物油的过程进行了 ML 模拟, 确定最优工艺条件从而提高整体的效率. 未来的研究应侧重于使用 ML 技术对生物油系统进行实时过程监测和控制, 以提高产品产量. 生物炭作为一种具有高孔隙率和养分供应能力的稳定碳质物质, 近年来已被广泛用作 OSW 好氧堆肥的补充剂. 随着 ML 的引入也取得了许多突破性的研究成果. 例如 Zhu 等[41] 根据生物质特性和热解条件成功预测了木质纤维素生物质的生物炭产量和碳含量(C-char), 通过 ML 建模证明了热解温度对二者的主导作用, 为生物炭生产的优化提供了新思路. 未来的研究应进一步探索生物炭在土壤改良和温室气体减排中的潜力, 并通过 ML 模型细化热解条件对生物炭特性的影响, 实现高效生产.

2.3　厌氧生物处理

厌氧生物处理通过厌氧和兼性微生物在无氧条件下, 将有机大分子降解为小分子物质, 如挥发性脂肪酸(VFAs)、醇类、甲烷和氢气等, 这些产物被视为可再生化学品的重要来源[42–43]. 根据反应程度和产物类型可以将厌氧生物处理进一步分为厌氧消化(AD) 和厌氧发酵 (AF). AD 包括水解、酸化和产甲烷阶段, AF 仅涉及上述的前两个步骤, 通过选择性抑制产气, 可以回收各种高附加值的液体产品, 如乳酸、乙醇和短链脂肪酸[44], 由于传统机械模型在生物转化过程建模方面存在局限, 近年来, ML 技术被广泛用于开发和优化厌氧生物处理工艺[45], 显著提高了工艺的稳定性和效率.

在 AD 模拟中, 大多数 ML 应用集中于根据工艺参数和原料特性预测和优化沼气生产. 一些研究通过 ANN 模型实现沼气产生量的精准预测, R^2 位于 0.87 ~ 1 之间, GA 进一步优化了产气量[46–47]. 其他 ML 模型如 MLR、KNN、SVM、RF、XGBoost 等也被用于预测 AD 性能. Wang 等[48] 使用 RF 和逻辑回归多类 (GLMNET) 来预测甲烷产量, 并使用底物中的主要成分作为输入特征. 还有研究对 AD 过程当中的氮、碳趋向进行了预测, 如 Li 等[49] 建立 ANN 模型预测化学需氧量(COD)浓度、CH_4 产量、NH_4^+-N 浓度, 模型的 R^2 大于 0.7. Alejo 等[50] 通过 SVM 的运用, 对两级 AD 工艺中的总氨氮量进行了预测, 模型 R^2 为 0.9. 反向传播神经网络(BP–ANN)被用于 AD 碱度建模, 表现出优异的性能(R^2= 0.99)[48]. 总之, 这些 ML 应用显著提升了 OSW 在 AD 处理中的效率及稳定性.

对 AF 流程的理解和优化同样受益于 ML 模型的应用. Kazemi 等[51] 成功开发了一个 ML 框架, 评估了 AF 过程中 VFA 浓度与操作条件之间的关系, R^2高达 0.999, 且在故障检测(FD)的稳健性方面也表现出色. Xu 等[52] 开发的 XGBoost 模型在使用操作条件作为输入来预测基于零价铁(ZVI)的发酵性能上准确性最高. Vendruscolo 等[53] 运用 ML 识别了 AF 系统中与生物气产量相关的关键功能微生物, 如 *Candidatus Cloacomonas*、*Methanospirillum* 等, 揭示了微生物群落结构的季节性变化规律, 为微生物管理提供了新的视角.

此外, ML 技术还可以应用于更为复杂的厌氧

生物转化平台. 例如, 碳链延长(CE)技术作为一种新兴的碳资源回收生物技术, 是在 AF 的基础上首先通过厌氧混合菌将 OSW 转化为液体产品作为中间体, 再通过反向 β 氧化或脂肪酸生物合成(FAB)途径延伸为具有更高经济价值的中链脂肪酸(MCFA). CE 技术由于工艺的强可操作性和产品的高产值优势而逐渐受到重视, 但与 AD 和 AF 相比, MCFA 生产的动力学模型研究仍处于起步阶段. 早期, Coelho 等[54] 使用残余甘油和屠宰场废水作为底物, 通过 *C. kluyveri* 的生物动力学模型模拟 CE 过程[55]. 然而, 由于 CE 过程的复杂性, 更多学者开始尝试利用 ML 替代传统模型, 由于 CE 过程的复杂性, 越来越多的研究者尝试利用 ML 模型替代传统动力学模型, 为不同动态条件下的 MCFA 产量预测提供了新的解决方案, 展现出广阔的应用前景. 例如 Long 等[15] 结合 XGBoost、RF、GA 等算法证明了 ML 在促进 MCFA 生产方面的可行性(R^2=0.89), 同时优化了从食物垃圾中生产 MCFA 的条件, 生产率提高了 113%[56]. 未来可以继续探索如何将训练好的 ML 模型与其他厌氧生物处理控制策略相结合, 以实现更加全面和自动化的过程优化.

2.4 好氧堆肥

好氧堆肥是一种用于稳定 OSW 的生物化学方法, 主要处理易于微生物分解的有机物, 如蛋白质和纤维素, 最终产物可用于土壤改良[57]. 确保堆肥产品的质量对于其在土壤中的有效和安全应用至关重要. 然而, 堆肥过程中伴随着氨气和温室气体的排放, 产品中还可能残留重金属和抗生素[58]. 这对环境和健康构成潜在风险. ML 技术因其处理非线性关系的能力, 能够对复杂微生物参与的堆肥反应进行建模, 显著降低试错成本. 根据统计, 最近的研究主要集中在堆肥腐熟度、工艺参数和污染物排放的预测和优化上.

腐熟度是反映有机物降解和生物化学稳定度的指标. Sharma 等[59] 以 C/N、CO_2 释放率和 GI(种子发芽指数)为输入变量, 预测了堆肥成熟度, 结果表明 ANN 模型的表现优于响应面算法, R^2 为 0.94. Muthuveni 等[60] 还根据堆肥成熟度指数评估了 ANN 在预测采用不同基质组合的蚯蚓堆肥过程中的性能, 研究结果表明, 反向传播神经网络(BPNN)显示出极佳的预测准确性(R^2>0.99). 此外, 还有研究将 CNN 引入, 克服了传统 ANN 的特征提取限制, 对堆肥图像进行直接分析, 来对堆肥的腐熟度进行评估, 并将模型准确度进行优化, 且精确度超出了 99%[61−62].

ML 模型通常会结合优化算法对堆肥工艺进行优化. Soto-Paz 等[63] 通过 ANN 和粒子群优化(PSO)算法的集成, 首次分析了两个操作参数(混合比、翻堆频率)的同时变化对堆肥工艺和产品质量的影响. Bayındır 等[64] 利用级联正向神经网络(CFNN)对牛粪和城市固体废物共堆肥中的环境因素进行建模, 并结合 GA 算法优化堆肥时间和材料混合比.

ML 方法有助于对好氧堆肥中有害物质的控制. Alavi 等[65] 采用 ANN 模型对堆肥过程中四环素(TC)的降解进行了有效模拟(R^2=0.99). 通过敏感性分析得出时间在 TC 降解中的决定性影响(相对重要性为 80.43%). Yamawaki 等[66] 通过 RF 和 XGBoost 模型评估了聚乳酸基生物塑料的降解效率, RF 模型表现最佳(R^2=0.78). 其他研究成功地将 ANN、SVM 等 ML 模型与各种生物传感器相结合用于预测 OSW 堆肥中有机污染物的降解[67−68], 有效拓展了生物传感器的浓度检测范围和检测准确性.

3 机器学习在有机固体废物领域的模型选择

根据处理工艺的选择, OSW 的处理可以分为物理、化学和生物处理, 处理目标和技术要求都有各自的特点, 需要针对性地选择和应用适合的的 ML 方法. 例如, 在物理处理中, 通常更关注废物的分选与减量化, ML 模型则多用于分类与优化物流路径; 在化学处理过程中, 重点在于废物的化学转化与稳定化, ML 可帮助预测热化学处理的产物分布与能量回收效率; 生物处理则依赖微生物的代谢活动降解有机物, ML 可用于模拟微生物的行为和优化发酵条件. 不同的 ML 模型在这些领域中发挥着独特作用, 各自具有优势与局限. 通过对模型进行全面比较, 可以更清楚地理解其特性, 从而帮助研究人员根据 OSW 处理需求, 优化工艺流程、提升资源化效率并做出更科学的决策.

各种 ML 模型在 OSW 资源化管理与处理领域的应用频率统计结果如图 3 所示. ANN 是 OSW 领域中最常用的 ML 模型, 占应用总数的 40.89%, 其次是 SVM (16.44%) 和 RF/DT (19.52%). 其他模型如多元线性回归(MLR)、K-最近邻(KNN)、极端梯度提升 (XGBoost)、梯度推进回归树 (GBRT) 等占总模型的 23.15%. 此外, 值得注意的是, 统计得到遗传算法(GA)作为一种优化算法, 常与这些模型结合以优化其超参数、结构和全局性能. 在所有研究中的使用占比达 7.31%. GA 尤其适用于复杂

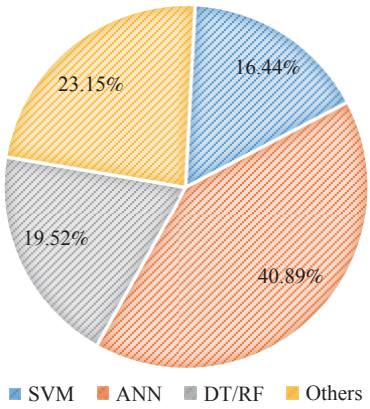

图 3　应用于 OSW 资源化的各种 ML 模型的频率统计

Fig.3　Frequency of various machine learning algorithms(including SVM, ANN,RF/DT and others) applied to the OSW field

的多维优化问题, 例如在 ANN 中优化神经网络的层数与权重[69], 在 SVM 中选择最优的核函数参数[70], 或者在集成学习中改进基本学习器的性能[71], 从而提升模型的预测能力与稳定性.

首先针对 ANN、SVM 和 DT/RF 这三类主要模型, 统计了其分别在 OSW 的物理、化学、生物处理领域的应用频率, 并相应地绘制了频率热图(图 4).

如图 4 所示, 在 OSW 处理领域中, ANN 的使用频率占据首位, 表明 ANN 可能是目前最流行的 ML 模型. 然而, ANN 模型在数据输入不足时容易遇到过拟合问题, 因此在样本量较少的情况下, SVM 算法的优势更为明显. 在化学与物理领域中, DT/RF、SVM 和 ANN 等模型已有广泛应用. 相较于 ANN 和 SVM, DT 和 RF 模型更容易解释和实现, 并且对缺失值具有较高的容忍度, 这是一个显著的优势, 因为数据缺失在 OSW 相关研究中非常常见. 此外, DT 和 RF 在处理包含不相关特征的数据集时表现更佳, 但在处理强相关特征的数据集时则表现平平. 目前, DT/RF 在城市生活垃圾生成预测、城市生活垃圾分类、热解性能估算等方面都表现出了较强的能力, 而在生化领域应用相对较少, 主要是因为它们在处理复杂非线性问题时存在局限. 相反, SVM 和 ANN 在生化过程建模中表现出色. Nair 等[72] 提出了有关 ANN 对厌氧消化中甲烷产量的预测能力的担忧, 因为模型仅达到了 R^2 为 0.73 的预测精度. 分析表明, 这可能是因为模型仅使用了化学或物理指标作为输入变量, 而未考虑生物指标. 类似地, Alejo 等[50] 使用 SVM 和 ANN 模型, 以化学需氧量、进水总氨氮和总挥发性固体量为输入数据, 对消化出水总氨氮浓度进行预测, 但预测精度仅为 0.81 和 0.77. 实际上, 由于相关参数难以明确, 目前涉及复杂生化过程的研究很少使用生物指示剂进行模型训练.

为了更好地分析不同机器学习模型在 OSW 处理领域的应用和发展趋势, 本文引入了时间轴分析. 从图 5 可以看出, 2018 年的研究主要以 DT/RF 和 SVM 为主, 反映了早期 OSW 处理领域对传统分类与回归模型的依赖. 随着时间推移, XGBoost 和 ANN 逐渐成为主流模型. XGBoost 的使用频率在 2020 年后大幅增长, 表明集成学习在处理复杂非线性问题和大数据集时的优势逐渐显现, 而 ANN

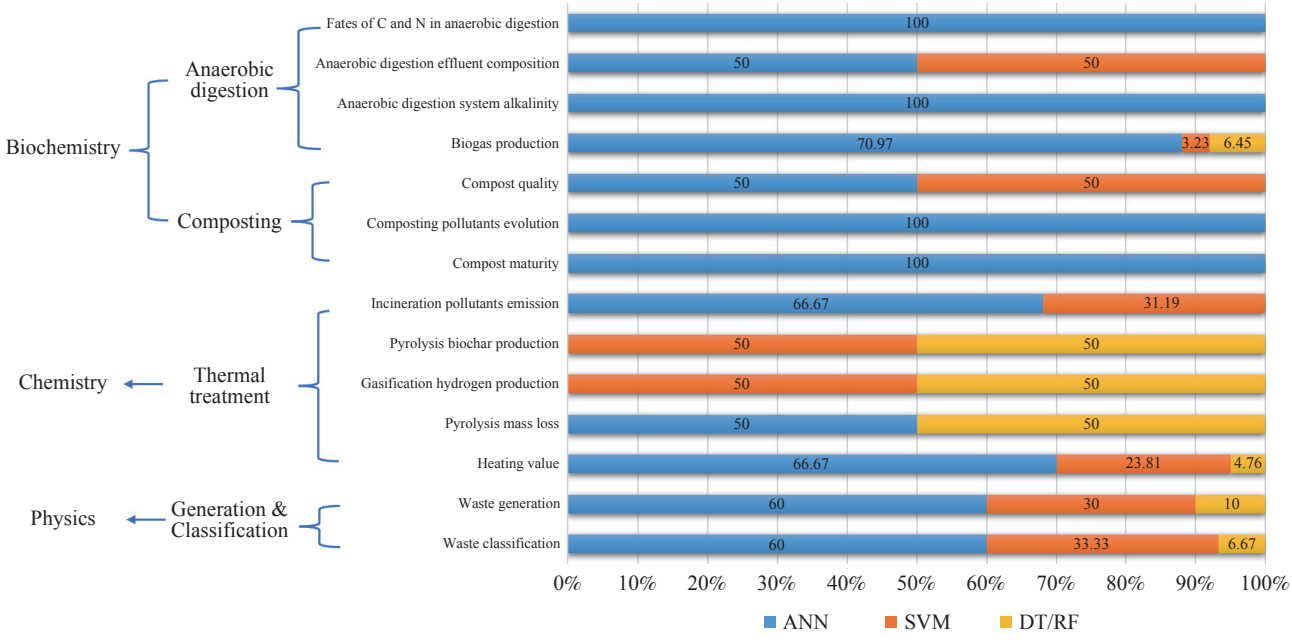

图 4　ANN、SVM、DT/RF 模型在 OSW 各领域的应用频率热图

Fig.4　Heatmap of the application frequencies of ANN, SVM, and DT/RF models in various OSW treatment methods

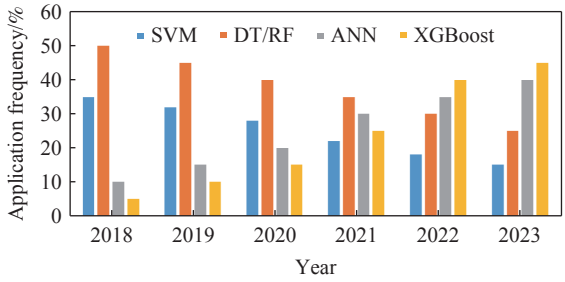

图 5 2018—2023 年 OSW 处理领域机器学习模型应用频率趋势

Fig.5 Usage trends of machine learning models in OSW treatment (2018—2023)

凭借其强大的特征表示能力, 尤其在处理非结构化数据时, 使用频率在后期显著增加. 相比之下, SVM 的使用频率逐年下降, 表明其在应对复杂数据时的局限性.

由此可见, 数据集的规模对模型选择至关重要. SVM 适合于数据量较小的情况; 当数据量适中时, 集成树模型(如 RF 和 XGBoost)通常优于 ANN, 因为它们对数据缺失具有较高的容忍度. 随着数据量的增加, 对模型能力的要求提高, 过拟合风险降低, ANN 的优势逐渐显现, 而集成学习的优势相对减少. 表 1 对近年的常用模型的适用性进行了总结和分析, 为未来研究开发和模型选择提供了有效参考, 并强调了不断改进模型评估工作的必要性.

总而言之, 在 OSW 处理领域, ML 的应用呈现出不同时间段的研究趋势和方法选择的多样性, 且预测目标各异. 研究者应首先基于数据的结构和规模, 结合具体的处理需求针对性地选择适宜的 ML 模型. 例如, 在数据量较小时, SVM 模型因其出色的高维数据处理能力常被采用; 而在大型非线性数据集上, ANN 与集成学习模型(如 XGBoost、RF)表现尤为优异. 此外, 输入变量的选择也应基于问题所涉及的基本过程. 对于复杂且机理不明确的生化过程, 输入变量的选择需要更加谨慎, 并且必须考虑微生物等内在作用. 模型建立之后, GA 算法的引入可以有效改进模型的超参数选择与优化过程, 显著提升模型的预测精度与稳定性, 特别是在 OSW 处理的动态复杂环境中.

4 机器学习的潜在挑战与未来展望

4.1 潜在挑战

尽管 ML 技术在 OSW 处理领域展现出巨大潜力, 但其应用仍面临一些挑战, 可能限制其进一步发展.

(1) 数据需求量大: ML 模型的训练和验证往往依赖于大量数据. 在数据稀缺, 尤其是高维数据集的情况下, 模型可能会陷入过拟合, 仅“记住”而非“学习”数据模式, 导致训练效率降低.

(2) 问题复杂性: 实际问题通常具有高度复杂性, 其形式化后的目标函数同样复杂. 目前, 尚未存在普遍有效的算法来寻找这类复杂目标函数的最优解, 这要求我们探索新的优化技术.

表 1 ANN、SVM、RF/DT、XGBoost 的优点、局限性及相应的适用领域

Table 1 Advantages, limitations, and corresponding application fields of ANN、SVM、RF/DT、GA in OSW

ML model	Advantages	Limitations	Suitable application fields
ANN	(1) Deep learning capability (2) Adaptability to various data types (3) Automatic feature extraction (4) Efficient graphics processing unit (GPU) training	(1) Requires large datasets (2) Complex tuning and architecture (3) Computationally intensive	(1) Predictive maintenance of waste-to-energy plants (2) Real-time waste monitoring (3) Dynamic waste routing
SVM	(1) Effective in high-dimensional spaces (2) Versatile, with various kernels (3) Regularization controls overfitting (4) Global optimization	(1) Memory intensive with large datasets (2) Sensitive to parameter tuning (3) Primarily for binary classification	(1) Landfill gas prediction (2) Odor-control optimization (3) Resource-recovery planning
DT	(1) Fast learning (2) Minimal data preprocessing	(1) Prone to overfitting (2) Limited nonlinear capabilities	(1) Composting optimization (2) Waste generation and classification (3) Landfill diversion strategies (4) Pyrolysis performance estimation (5) Bioprocess optimization
RF	(1) Robust ensemble method (2) Addresses nonlinearity (3) Less prone to overfitting than DT	(1) Slower training than DT (2) Complexity in interpretation (3) Not ideal for high dimensions	
XGBoost	(1) High efficiency and fast training on large datasets (2) Addresses missing data well (3) Regularization controls overfitting (4) Scalable for high dimensions	(1) Computationally intensive (2) Complex interpretation (3) Requires careful parameter tuning	(1) Resource-recovery efficiency prediction (2) Real-time process optimization (3) Multivariable process optimization (4) Pyrolysis and gasification modeling

(3) 数据分布不均: 在某些反应过程中, 由于数据点稀缺(如羟基自由基反应中反应速率过快或过慢的情况), 模型的预测精度可能受到影响. 此外, 不同数据集的实验条件和侧重点差异, 使得数据整合成为一个难题.

(4) 模型可解释性: 以 ANN 为代表的某些模型因其“黑箱”特性而受到批评. 这些模型的内部机制难以理解, 变量的重要性难以评估, 为模型的应用带来了不确定性.

尽管存在上述挑战, ML 在 OSW 处理和管理方面的应用潜力依然巨大, 对环境保护和资源利用具有重要价值.

4.2　未来展望

面对现有挑战, 可以考虑开展如下的策略:

(1) 集成模型的应用: 目前大多数研究集中于单一机器学习模型. 然而, 集成学习模型, 如 XGBoost 和 RF, 相比单一模型, 通常具有更低的过拟合风险、更精准的预测能力以及更强的稳健性. 这些集成模型在处理大规模数据、非线性问题以及多变量优化时表现出色, 已成为解决复杂问题的重要工具. GA 则作为一种优化工具, 常用于优化其他模型的参数设置, 提升其性能[73].

(2) 跨学科技术融合: 将机器学习与其他先进技术结合, 如将机器学习与可见近红外漫反射光谱技术结合, 可快速预估堆肥成熟度和酶活性[74]; 结合物联网技术, 可为提升废弃物管理智能化水平提供创新解决方案.

(3) 数据不足问题的解决策略: 除了数据提取、实验设计、特征选择和数据增强等传统的筛选和净化数据集的策略, 强化学习(RL)和迁移学习(TL)作为前沿的机器学习方法, 提供了有效的替代方案. TL 通过将已有模型应用到新的数据集上, 减少了对大规模数据的依赖, 例如将化学反应模型应用于 OSW 的热解或发酵过程优化. 而 RL 则通过不断试验和反馈机制, 优化处理流程. 特别是在 OSW 处理的动态环境中, 强化学习可以有效提升处理效率并确保系统的稳定性.

(4) “黑箱”问题的深入研究: 已有方法可以解释输入变量对模型的贡献, 部分实现“黑箱”内部的可视化[75]. 未来研究中应更频繁地应用这些解释方法, 并不断探索更简单、有效的新方法, 以进一步阐明“黑箱”问题.

5　结论

本研究考察了机器学习(ML)技术在有机固体废物(OSW)处理与资源化领域的应用现状与发展趋势, 分析了 ML 在 OSW 管理和处理中的多种应用场景. 结果表明当前应用主要聚焦在 OSW 的源头管理方面, 其次为热化学处理、厌氧生物处理、好氧堆肥. 此外, 发展趋势表明, ML 方法可以为实现废弃生物质高值化提供强有力的技术支持. 本研究进一步深入对比了人工神经网络(ANN)、支持向量机(SVM)、决策树(DT)/随机森林(RF)和极端梯度提升(XGBoost)等常用 ML 模型在 OSW 处理中的应用频率、原理、适用性和局限性, 同时讨论了遗传算法(GA)作为一种优化算法在领域中的应用情况. ANN 因其在处理大型非线性数据集方面的卓越性能而得到广泛应用, 而 GA 则因其解决复杂问题的优化能力常与 ML 模型结合使用以提高性能. ML 在废物管理优化、资源利用提升、污染风险降低等多个领域都有巨大的应用潜力, 不仅提高了 OSW 处理的效率和准确性, 还为环境工程领域提供了新的视角和解决方案. 同时, 本研究也指出了 ML 应用的潜在限制, 包括数据质量、模型泛化能力及算法选择问题, 并提出了相应的解决策略.

参　考　文　献

[1] Arun C, Sivashanmugam P. Study on optimization of process parameters for enhancing the multi-hydrolytic enzyme activity in garbage enzyme produced from preconsumer organic waste. *Bioresour Technol*, 2017, 226: 200

[2] Laurent A, Bakas I, Clavreul J, et al. Review of LCA studies of solid waste management systems: Part I: Lessons learned and perspectives. *Waste Manag*, 2014, 34(3): 573

[3] Kumar A, Samadder S R. Performance evaluation of anaerobic digestion technology for energy recovery from organic fraction of municipal solid waste: A review. *Energy*, 2020, 197: 117253

[4] Salman C A, Schwede S, Thorin E, et al. Predictive modelling and simulation of integrated pyrolysis and anaerobic digestion process. *Energy Procedia*, 2017, 105: 850

[5] Wang H X, Xu J L, Yu H X, et al. Study of the application and methods for the comprehensive treatment of municipal solid waste in Northeastern China. *Renew Sustain Energy Rev*, 2015, 52: 1881

[6] Nikku M, Deb A, Sermyagina E, et al. Reactivity characterization of municipal solid waste and biomass. *Fuel*, 2019, 254: 115690

[7] Patra A K, Lalhriatpuii M. Development of statistical models for prediction of enteric methane emission from goats using nutrient composition and intake variables. *Agric Ecosyst Environ*, 2016, 215: 89

[8] Andrade Cruz I, Chuenchart W, Long F, et al. Application of machine learning in anaerobic digestion: Perspectives and challenges. *Bioresour Technol*, 2022, 345: 126433

[9] Guo H N, Wu S B, Tian Y J, et al. Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: A review. *Bioresour Technol*, 2021, 319: 124114

[10] Portugal I, Alencar P, Cowan D. The use of machine learning algorithms in recommender systems: A systematic review. *Expert Syst Appl*, 2018, 97: 205

[11] Cipullo S, Snapir B, Prpich G, et al. Prediction of bioavailability and toxicity of complex chemical mixtures through machine learning models. *Chemosphere*, 2019, 215: 388

[12] Chen C, Liang R, Song M Y, et al. Noise-assisted data enhancement promoting image classification of municipal solid waste. *Resour Conserv Recycl*, 2024, 209: 107790

[13] Ahmad A, Yadav A K, Singh A, et al. A comprehensive machine learning-coupled response surface methodology approach for predictive modeling and optimization of biogas potential in anaerobic co-digestion of organic waste. *Biomass Bioenergy*, 2024, 180: 106995

[14] Wu Q L, Bao X, Guo W Q, et al. Medium chain carboxylic acids production from waste biomass: Current advances and perspectives. *Biotechnol Adv*, 2019, 37(5): 599

[15] Long F, Fan J, Xu W C, et al. Predicting the performance of medium-chain carboxylic acid (MCCA) production using machine learning algorithms and microbial community data. *J Clean Prod*, 2022, 377: 134223

[16] Agrawal P, R G, Dhawane S H. Prediction of biodiesel yield employing machine learning: Interpretability analysis *via* shapley additive explanations. *Fuel*, 2024, 359: 130516

[17] Ma J J, Zhang S, Liu X J, et al. Machine learning prediction of biochar yield based on biomass characteristics. *Bioresour Technol*, 2023, 389: 129820

[18] Kulisz M, Kujawska J. Prediction of municipal waste generation in Poland using neural network modeling. *Sustainability*, 2020, 12(23): 10088

[19] Elshaboury N, Mohammed Abdelkader E, Al-Sakkaf A, et al. Predictive analysis of municipal solid waste generation using an optimized neural network model. *Processes*, 2021, 9(11): 2045

[20] Soni U, Roy A, Verma A, et al. Forecasting municipal solid waste generation using artificial intelligence models-a case study in India. *SN Appl Sci*, 2019, 1(2): 162

[21] Zheng Y, Bai J R, Xu J N, et al. A discrimination model in waste plastics sorting using NIR hyperspectral imaging system. *Waste Manage*, 2018, 72: 87

[22] Serranti S, Gargiulo A, Bonifazi G. Classification of polyolefins from building and construction waste using NIR hyperspectral imaging system. *Resour Conserv Recycl*, 2012, 61: 52

[23] Toğaçar M, Ergen B, Cömert Z. Waste classification using AutoEncoder network with integrated feature selection method in convolutional neural network models. *Measurement*, 2020, 153: 107459

[24] Tao J Y, Gu Y D, Hao X L, et al. Combination of hyperspectral imaging and machine learning models for fast characterization and classification of municipal solid waste. *Resour Conserv Recycl*, 023, 188: 106731

[25] Bunsan S, Chen W Y, Chen H W, et al. Modeling the dioxin emission of a municipal solid waste incinerator using neural networks. *Chemosphere*, 2013, 92(3): 258

[26] Norhayati I, Rashid M. Adaptive neuro-fuzzy prediction of carbon monoxide emission from a clinical waste incineration plant. *Neural Comput Appl*, 2018, 30(10): 3049

[27] Tavoosi J, Mohammadzadeh A. A new recurrent radial basis function network-based model predictive control for a power plant boiler temperature control. *Int J Eng*, 2021, 34(3): 667

[28] Alitasb G K, Salau A O. Multiple-input multiple-output Radial Basis Function Neural Network modeling and model predictive control of a biomass boiler. *Energy Rep*, 2024, 11: 442

[29] Liu Z Y, Lu M, Zhang Y F, et al. Identification of heavy metal leaching patterns in municipal solid waste incineration fly ash based on an explainable machine learning approach. *J Environ Manag*, 2022, 317: 115387

[30] Guo L S, Xu X, Wang Q, et al. Machine learning-based prediction of heavy metal immobilization rate in the solidification/stabilization of municipal solid waste incineration fly ash (MSWIFA) by geopolymers. *J Hazard Mater*, 2024, 467: 133682

[31] Li J S, Yao X W, Xu K L. A comprehensive model integrating BP neural network and RSM for the prediction and optimization of syngas quality. *Biomass Bioenergy*, 2021, 155: 106278

[32] Zhao S H, Xu W J, Chen L H. The modeling and products prediction for biomass oxidative pyrolysis based on PSO-ANN method: An artificial intelligence algorithm approach. *Fuel*, 2022, 312: 122966

[33] Leng L J, Zhang W J, Liu T G, et al. Machine learning predicting wastewater properties of the aqueous phase derived from hydrothermal treatment of biomass. *Bioresour Technol*, 2022, 358: 127348

[34] Shu H Y, Lu H C, Fan H J, et al. Prediction for energy content of Taiwan municipal solid waste using multilayer perceptron neural networks. *J Air Waste Manag Assoc*, 2006, 56(6): 852

[35] Taki M, Rohani A. Machine learning models for prediction the Higher Heating Value (HHV) of Municipal Solid Waste (MSW) for waste-to-energy evaluation. *Case Stud Therm Eng*, 2022, 31: 101823

[36] Leng L J, Li H, Yuan X Z, et al. Bio-oil upgrading by emulsification/microemulsification: A review. *Energy*, 2018, 161: 214

[37] Aghbashlo M, Peng W X, Tabatabaei M, et al. Machine learning technology in biodiesel research: A review. *Prog Energy Combust Sci*, 2021, 85: 100904

[38] Tang Q H, Chen Y Q, Yang H P, et al. Prediction of bio-oil yield and hydrogen contents based on machine learning method: Effect of biomass compositions and pyrolysis conditions. *Energy Fuels*, 2020, 34(9): 11050

[39] Ma Z Y, Wang H N, Li Y L, et al. Optimized bio-oil emulsification for sustainable asphalt production: A step towards a low-carbon pavement. *Constr Build Mater*, 2024, 419: 135218

[40] Liu Y, Sayed B T, Sivaraman R, et al. Novel and robust machine learning model to optimize biodiesel production from algal oil using CaO and CaO/Al_2O_3 as catalyst: Sustainable green energy. *Environ Technol Innov*, 2023, 30: 103018

[41] Zhu X Z, Li Y N, Wang X N. Machine learning prediction of biochar yield and carbon contents in biochar based on biomass characteristics and pyrolysis conditions. *Bioresour Technol*, 2019, 288: 121527.

[42] Greses S, Tomás-Pejó E, González-Fernández C. Short-chain fatty acids and hydrogen production in one single anaerobic fermentation stage using carbohydrate-rich food waste. *J Clean Prod*, 2021, 284: 124727

[43] Dong W J, Yang Y L, Liu C, et al. Caproic acid production from anaerobic fermentation of organic waste - Pathways and microbial perspective. *Renew Sustain Energy Rev*, 2023, 175: 113181

[44] Perez-Esteban N, Vinardell S, Vidal-Antich C, et al. Potential of anaerobic co-fermentation in wastewater treatments plants: A review. *Sci Total Environ*, 2022, 813: 152498

[45] Iglesias-Iglesias R, Kennes C, Veiga M C. Valorization of sewage sludge in co-digestion with cheese whey to produce volatile fatty acids. *Waste Manage*, 2020, 118: 541

[46] Barik D, Murugan S. An artificial neural network and genetic algorithm optimized model for biogas production from co-digestion of seed cake of karanja and cattle dung. *Waste Biomass Valorization*, 2015, 6(6): 1015

[47] Abu Qdais H, Bani Hani K, Shatnawi N. Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm. *Resour Conserv Recycl*, 2010, 54(6): 359

[48] Wang X M, Bai X, Li Z F, et al. Evaluation of artificial neural network models for online monitoring of alkalinity in anaerobic co-digestion system. *Biocheml Eng J*, 2018, 140: 85

[49] Li H, Ke L T, Chen Z, et al. Estimating the fates of C and N in various anaerobic codigestions of manure and lignocellulosic biomass based on artificial neural networks. *Energy Fuels*, 2016, 30(11): 9490

[50] Alejo L, Atkinson J, Guzmán-Fierro V, et al. Effluent composition prediction of a two-stage anaerobic digestion process: Machine learning and stoichiometry techniques. *Environ Sci Pollut Res Int*, 2018, 25(21): 21149

[51] Kazemi P, Bengoa C, Steyer J P, et al. Data-driven techniques for fault detection in anaerobic digestion process. *Process Saf Environ Protect*, 2021, 146: 905

[52] Xu W C, Long F, Zhao H, et al. Performance prediction of ZVI-based anaerobic digestion reactor using machine learning algorithms. *Waste Manage*, 2021, 121: 59

[53] Vendruscolo E C G, Mesa D, Rissi D V, et al. Microbial communities network analysis of anaerobic reactors fed with bovine and swine slurry. *Sci Total Environ*, 2020, 742: 140314

[54] Coelho M M H, Morais N W S, Pereira E L, et al. Potential assessment and kinetic modeling of carboxylic acids production using dairy wastewater as substrate. *Biochem Eng J*, 2020, 156: 107502

[55] Candry P, Radić L, Favere J, et al. Mildly acidic pH selects for chain elongation to caproic acid over alternative pathways during lactic acid fermentation. *Water Res*, 2020, 186: 116396

[56] Long F, Fan J, Liu H. Prediction and optimization of medium-chain carboxylic acids production from food waste using machine learning models. *Bioresour Technol*, 2023, 370: 128533

[57] Guo X X, Liu H T, Wu S B. Humic substances developed during organic waste composting: Formation mechanisms, structural properties, and agronomic functions. *Sci Total Environ*, 2019, 662: 501

[58] Déportes I, Benoit-Guyod J L, Zmirou D. Hazard to man and the environment posed by the use of urban waste compost: A review. *Sci Total Environ*, 1995, 172(2-3): 197

[59] Sharma D, Pandey A K, Yadav K D, et al. Response surface methodology and artificial neural network modelling for enhancing maturity parameters during vermicomposting of floral waste. *Bioresour Technol*, 2021, 324: 124672

[60] Muthuveni M, Deebika S, Boopathy T, et al. I-optimal mixture design and artificial neural network for the sustainable production of vermicompost. *Biomass Convers Biorefin*, 2024, 14(9): 10147

[61] Xue W, Hu X J, Wei Z, et al. A fast and easy method for predicting agricultural waste compost maturity by image-based deep learning. *Bioresour Technol*, 2019, 290: 121761

[62] Kujawa S, Mazurkiewicz J, Czekała W. Using convolutional neural networks to classify the maturity of compost based on sewage sludge and rapeseed straw. *J Clean Prod*, 2020, 258: 120814

[63] Soto-Paz J, Alfonso-Morales W, Caicedo-Bravo E, et al. A new approach for the optimization of biowaste composting using artificial neural networks and particle swarm optimization. *Waste Biomass Valorization*, 2020, 11(8): 3937

[64] Bayındır Y, Cagcag Yolcu O, Aydın Temel F, et al. Evaluation of a cascade artificial neural network for modeling and optimization of process parameters in co-composting of cattle manure and municipal solid waste. *J Environ Manag*, 2022, 318: 115496

[65] Alavi N, Sarmadi K, Goudarzi G, et al. Attenuation of tetracyclines during chicken manure and bagasse co-composting: Degradation, kinetics, and artificial neural network modeling. *J Environ Manag*, 2019, 231: 1203

[66] Yamawaki R, Tei A, Ito K, et al. Decomposition factor analysis based on virtual experiments throughout Bayesian optimization for compost-degradable polymers. *Appl Sci*, 2021, 11(6): 2820

[67] Kim S, Lee M H, Wiwasuku T, et al. Human sensor-inspired supervised machine learning of smartphone-based paper microfluidic analysis for bacterial species classification. *Biosens Bioelectron*, 2021, 188: 113335

[68] Galan E A, Zhao H R, Wang X K, et al. Intelligent microfluidics: The convergence of machine learning and microfluidics in materials science and biomedicine. *Matter*, 2020, 3(6): 1893

[69] Li R, Xu A K, Zhao Y, et al. Genetic algorithm (GA) - Artificial neural network (ANN) modeling for the emission rates of toxic volatile organic compounds (VOCs) emitted from landfill working surface. *J Environ Manag*, 2022, 305: 114433

[70] Ramachandran S, Jayalal M L, Vasudevan M, et al. Combining machine learning techniques and genetic algorithm for predicting run times of high performance computing jobs. *Appl Soft Comput*, 2024, 165: 112053

[71] Wang Z N, Wu F X, Hao N, et al. The combined machine learning model SMOTER-GA-RF for methane yield prediction during anaerobic digestion of straw lignocellulose based on random forest regression. *J Clean Prod*, 2024, 466: 142909

[72] Nair V V, Dhar H, Kumar S, et al. Artificial neural network based modeling to evaluate methane yield from biogas in a laboratory-scale anaerobic bioreactor. *Bioresour Technol*, 2016, 217: 90

[73] Beltramo T, Hitzmann B. Evaluation of the linear and non-linear prediction models optimized with metaheuristics: Application to anaerobic digestion processes. *Eng Agric Environ Food*, 2019, 12(4): 397

[74] Chakraborty S, Das B S, Nasim Ali M, et al. Rapid estimation of compost enzymatic activity by spectral analysis method combined with machine learning. *Waste Manage*, 2014, 34(3): 623

[75] Olden J D, Jackson D A. Illuminating the "black box": A randomization approach for understanding variable contributions in artificial neural networks. *Ecol Model*, 2002, 154(1-2): 135