

机器学习在有机固废全链条处置中的应用进展

张子杭¹, 许丹¹, 胡艳军², 管文洁¹, 王树荣¹

(1. 浙江大学能源清洁利用国家重点实验室, 浙江 杭州 310027;

2. 浙江工业大学能源与动力工程研究所, 浙江 杭州 310014)

摘要:有机固废热转化过程会发生一系列复杂的热化学反应, 给深入理解其机理、优化转化过程技术参数及实现产物定向调控等带来挑战。基于数据驱动的机器学习建模方法可推动有机固废的智能化和精细化处置。综述了基于机器学习的智能建模方法在有机固废全链条处置中的应用, 总结了机器学习对有机固废上游产生量与理化特性、中游热转化过程及产物特性、下游产物利用与效益评估等的预测并对比了不同模型的适用场景及优缺点, 以期构建有机固废全链条智能化处置方案, 实现集无害化、减量化、资源化、高值化及智能化于一体的有机固废高效处置模式, 为实际固废的有效管理提供一定的指导意义。

关键词:有机固废; 热转化; 机器学习; 全链条处置

中图分类号: X705

文献标识码: A

Application progress of machine learning in the whole chain disposal of organic solid waste

ZHANG Zihang¹, XU Dan¹, HU Yanjun², GUAN Wenjie¹, WANG Shurong¹

(1. State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, China;

2. Institute of Energy and Power Engineering, Zhejiang University of Technology, Hangzhou 310014, China)

Abstract: A series of complex thermochemical reactions occur in the thermal conversion process of organic solid waste, which brings challenges to the in-depth understanding of its mechanism, the optimization of technical parameters for the conversion process and the directional regulation of products. The data-driven machine learning modeling method can promote the intelligence and refinement of organic solid waste disposal. In this review, the application of the intelligent modeling method based on machine learning in the whole chain disposal of organic solid waste was summarized. The prediction of machine learning on the upstream production and physical and chemical characteristics of organic solid waste, the thermal conversion process and product characteristics in the middle stream, and the utilization and benefit evaluation of downstream products were reviewed. The applicable scenarios, advantages and disadvantages of different models were compared. The purpose was to construct an intelligent disposal scheme for the whole chain of organic solid waste, and realize an efficient disposal mode of organic solid waste integrating harmlessness, reduction, resource utilization, high value and intelligence, and to provide some guides for the effective management of solid waste.

Keywords: Organic solid waste; Thermal conversion; Machine learning; Whole chain disposal

0 引言

平提高以及城市化进程加速, 城市固废产量也大幅增加。据估计, 到 2030 年和 2050 年, 全球城市

随着经济高速发展、人口快速增长、工业化水

收稿日期: 2022-12-10

基金项目: 浙江省“领雁”研发攻关计划资助(2022C03092); 国家自然科学基金区域创新联合基金(U21A20142)

第一作者: 张子杭(2000—), 男, 湖北孝感人, 博士生, 生物质和有机固废处置过程的智能化。E-mail: 12227022@zju.edu.cn

通讯作者: 王树荣(1972—), 男, 浙江嵊州人, 教授, 生物质和有机固废的热转化利用。E-mail: srwang@zju.edu.cn

固废产量将分别达到 25.9 亿吨和 34.0 亿吨^[1],其中约 40%~70% 为有机固废^[2],主要包括餐厨垃圾、塑料等^[3],此外,有机固废还包括农业废弃物(如作物秸秆、畜禽粪便)和工业固废(如工业污泥)等^[4],有机固废成分复杂,含大量有机物,若不能进行有效管理、处置、回收将会对生态环境以及人类健康造成严重威胁^[5],同时也会导致资源浪费。为此,如何清洁、高效地实现有机固废无害化、减量化与资源化是固废有效管理中亟需解决的问题。

目前,常见有机固废可通过生物化学(填埋、堆肥、厌氧消化)和热化学转化(焚烧、热解、气化、水热转化)等方法进行处置^[6-7],相比于生物化学处置方法,热化学转化方法在减少二次污染的同时可以更大程度地实现有机固废的减量化和资源化^[8-9]。然而,热转化过程中发生的一系列复杂反应为深入理解反应机理、高效转化工艺优化及反应产物定向调控带来挑战。此时,对处置过程进行先建模从而指导过程的反向设计是解决上述难题的有效方案。建模过程可分为基于内在反应机理的建模和基于数据驱动的建模^[10],机理建模高

度依赖于对反应内在机理的理解,而基于数据驱动的智能建模方法仅依赖于对历史经验数据的学习,就可挖掘并总结蕴藏在经验数据中的规律,最终运用这些规律实现建模后的精准决策。为此,基于数据驱动的机器学习建模方法可推动有机固废热转化处置的智能化、精细化和高值化。此外,有机固废管理与处置全环节还包括上游产生、分类、特性以及下游污染监测、产物利用及效益评估等过程^[11],上游环节为中游热转化处置的路线决策提供先验知识,下游环节完善中游处置的效益评估及产物的高值利用过程,运用机器学习建模方法对上游产生、下游应用及评估环节进行预测有助于实现有机固废全环节的高效管理与处置。

综上所述,本文重点对基于机器学习的智能建模方法在有机固废热转化处置中的应用进行了综述,同时对其上游产生、分类、特性预测与下游评估、应用中涉及的机器学习方法进行总结,以期构建机器学习辅助有机固废全链条高效处置的方案(见图 1),实现集无害化、减量化、资源化、高值化以及智能化于一体的有机固废高效处理模式,为实际固废的有效管理与处置提供一定的指导意义。



图 1 机器学习在有机固废全链条处置中的应用示意图

Fig.1 The diagram of machine learning application in the whole chain disposal of organic solid waste

1 数据驱动的机器学习建模

1.1 有机固废处置中机器学习建模及应用流程

基于数据驱动的机器学习算法已被用于建立一系列输入变量与输出变量间的关系并对输出结果进行精准预测^[12],其既可预测回归问题也可预测分类问题。当预测输出结果取值为一系列连续的数值时,该任务被定义为回归,回归任务也是机器学习辅助有机固废处置过程中最常见的,如对有机固废产生量、热转化失重曲线、产物产率的预测等。

有机固废处置中机器学习建模及应用流程见图 2,基于有机固废及其热转化特性建立数据库,对数据库中的数据进行预处理,主要包括缺失值填补、异常值剔除、数据归一化、特征选择等。随后选择模型进行训练,为提升模型预测性能,需对超参数进行调优,在超参数调优过程中采用交叉验证^[13]方式防止模型过拟合并提高模型的泛化能力。交叉验证过程如图 3 所示,首先将数据集划分为训练集和测试集,进一步地将训练集均匀划分为 K 份(K 折交叉验证),其中 K-1 份用于训练模型,剩下的一份用于验证,此步骤重复 K 次,

最终取 K 次验证集中结果的均值作为模型性能用于筛选最优超参数。最佳超参数确定后,训练集数据用于再次训练模型,最终模型性能由测试集进行评估。最后将模型投入实际应用,对具体案例进行预测,同时借助模型的可解释性分析反馈指导工艺的优化设计。

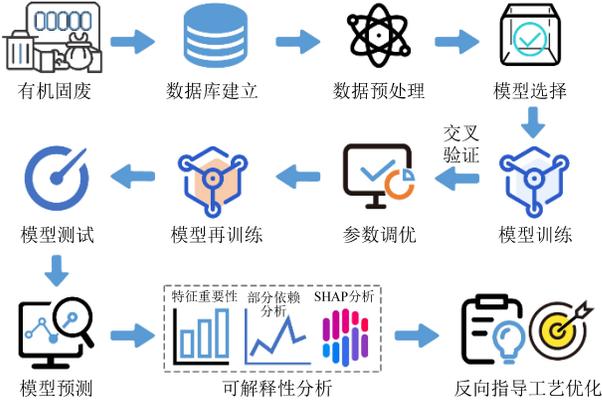


图2 有机固废处置中机器学习建模及应用流程示意图
Fig.2 The diagram of machine learning modeling and application process in disposal of organic solid waste

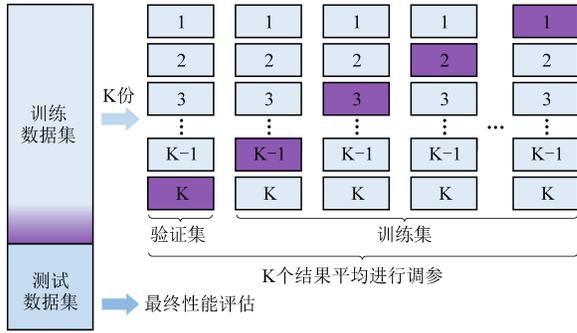


图3 K折交叉验证原理示意图

Fig.3 The principle diagram of K-fold cross validation

模型可解释性分析主要包括:特征重要性分析、部分依赖分析(Partial Dependence Analysis, PDA)和SHapley Additive exPlanations (SHAP)分析。特征重要性是分析各个输入特征对模型决策过程的影响程度,如置换特征重要性^[14],通过计算改变特征后导致模型预测误差的增量来评估该特征的重要性,误差增量越大则该特征越重要。部分依赖分析用于探究一个或两个特征对模型预测结果的边际效应^[13]。SHAP分析是基于博弈论的一种方法,用来将收益和成本同时按贡献分配给各个玩家,玩家指样本的特征,通过计算每个特征的SHAP值得到各特征对预测值的贡献,从而对模型预测结果进行解释^[15]。

1.2 模型性能评估指标

有机固废处置中常见的机器学习回归模型包

括经典模型:K近邻回归(K-Nearest Neighbors, KNN)、线性回归(Linear Regression, LR)、多项式回归(Polynomial Regression, PLR)、岭回归(Ridge Regression, RR)、支持向量机回归(Support Vector Machine, SVM)等;基于树结构的模型:决策树(Decision Tree, DT)、随机森林(Random Forest, RF)、梯度提升树(Gradient Boosting Decision Tree, GBDT)、极限梯度提升树(eXtreme Gradient Boosting, XGB)等;基于人工神经网络(Artificial Neural Network, ANN)的模型:多层感知机(Multi-layer Perceptron, MLP)、自适应神经模糊推理系统(Adaptive Network-based Fuzzy Inference System, ANFIS)、深度神经网络(Deep Neural Network, DNN)等;基于种群的模型:遗传算法(Genetic Algorithm, GA)、粒子群优化算法(Particle Swarm Optimization, PSO)等。其中经典模型在线性问题、小样本数据上表现出色,而树模型和神经网络模型通常用于处理复杂的非线性问题,如有机固废热转化产物特性预测、热转化失重曲线等;基于种群的模型主要作为优化算法使用。

为评估回归模型的拟合效果,通常引入决定系数(Coefficient of determination, R^2)、均方根误差(Root Mean Square Error, RMSE)和平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)对模型进行评估, R^2 是衡量模型预测能力的指标,取值范围为0~1;RMSE、MAPE评估了模型预测值与实际值之间的偏差情况,一般 R^2 越大、RMSE和MAPE越小,则模型的预测性能越好。三者的计算公式如下^[16-17]:

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

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

$$MAPE = \frac{100}{N} \sum_{i=1}^N \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (3)$$

式中, N 为测试集总样本数, \hat{y}_i 代表第*i*个样本的预测值, y_i 代表第*i*个样本的真实值, \bar{y} 为测试集总样本的平均值。

2 机器学习辅助有机固废智能化管理与处置

2.1 有机固废产生及其理化特性预测

机器学习在有机固废管理与处置的上游环节

主要用于预测有机固废的产生及其理化特性,对于产生量的预测有助于掌握有机固废产生量的动态变化规律,从而为其的产生控制、规划处置提供有效信息;对于有机固废理化特性的预测(如:水分含量、热值等)可为中游热转化处置的路线决策提供先验知识。

Nguyen 等^[18]以涵盖经济、人口、消费和固废产生特性的 10 个特征(城市人口数量、消费品零售总额、人均月收入、人均住房面积、人口密度、人均月消费支出等)作为输入,通过 LR、KNN、SVM、RF 等算法对城市固废产量进行预测,发现 RF 和 KNN 模型具有良好的预测能力,预测 R^2 均大于 0.96,所开发的机器学习将有助于规划、设计和实施城市固废综合管理计划。

Qi 等^[19]提出了一种基于衰减全反射-傅里叶变换红外光谱(Attenuated Total Refraction-Fourier Transform Infrared Spectrometer, ATR-FTIR)和不同机器学习算法:SVM、RF、偏最小二乘回归(Partial Least Squares Regression, PLSR)、连续投影算法(Successive Projections Algorithm, SPA)的无损测定多源固废(纺织品、纸张、皮革和木材)含水率的方法,建立了一种含水率预测的组合回归模型,组合模型能够高效预测多源固废含水率,水分预测的验证和测试集 R^2 分别达到 0.96、0.97。该方法可避免对物体的损坏,降低重复人工检测的成本,实现快速高精度的检测,同时对于进一步监测和管理固废处理处置过程、提高固废分类效率、减少二次污染具有重要意义。

Tao 等^[20]提出了一种通过红外光谱和机器学习模型对固废进行快速表征的新方法,该混合模型包括从原始红外光谱中提取核心信息的特征压缩部分、区分无机组分的分类部分和预测有机组分元素组成和热值的回归部分。利用特征压缩后的光谱信息作为输入进行建模,通过参数优化,该混合模型对固废的 C、H、O 含量和热值预测的准确率分别达到 95.54%、85.53%、92.40% 和 92.49%。最终进行稳健性分析验证了红外光谱包含足够的信息来表征固废的燃料特性,从而为固废处理和利用的上游分拣系统设计提供了思路,同时也助于中游处置路线的决策。

Taki 等^[21]利用四种机器学习方法:径向偏置函数人工神经网络(Radial Basis Function - Artificial Neural Network, RBF-ANN)、多层感知器人工神经网络(MLP-ANN)、SVM 和 ANFIS 对具

有六种不同输入(C、H、O、N、S、灰分、水分含量)的城市固废高位热值(Higher Heating Value, HHV)进行建模。结果表明,RBF-ANN 比其他模型能更准确地预测固废的 HHV。MLP-ANN、SVM 和 ANFIS 模型预测的 MAPE 分别为 7.3%、11.77% 和 23.76%,而具有最佳超参数 RBF-ANN 预测的 MAPE 为 0.45%,由此人工神经网络可作为一种具有高精度和可靠性的实用工具,用于垃圾焚烧厂的设计和管理。

最终,上游环节机器学习预测有机固废产生量及热转化特性可为中游热转化处置的路线决策提供先验知识。

2.2 热转化处置过程预测及工艺反向设计

热转化处置属于有机固废全链条处置的中游环节,也是减量化和能源化处理思路的技术核心,主要包括热解、气化、燃烧、水热转化等,研究表明利用机器学习建模可以对热转化过程进行精准建模预测,有助于实现有机固废高效热转化处置的智能化、精细化和高值化的决策分析。

2.2.1 热解

对热解产物生成特性等进行精准预测是对产物进行定向调控的前提。顾春晗等^[22]利用 BP 神经网络对 4 种固废(西瓜皮、玉米芯、市政污泥、桉木)的热解产物分布进行建模分析,以物料特性(固定碳、挥发分、灰分、C、H 含量)和反应工况(热解温度、停留时间、载气流量、升温速率)作为输入,三相产物产率作为输出,最终预测 R^2 为 0.97, RMSE 为 2.30,为不同种类固废的热解产物分布预测提供了一个可行的方案。为进一步探究输入特征与目标响应之间的关系,Gu 等^[23]基于 ANN 模型对固废热解的三相产物分布进行预测,同时借助敏感性分析探究不同输入变量的重要性以及单因素对输出结果的影响。Leng 等^[24]建立了基于 GBDT、RF、SVM 和 MLP 的三相产物分布和生物油热值预测模型,借助特征重要性及 PDA 分析得出 480 °C 是生物质在鼓泡流化床中快速热解制油的最佳温度。

除对热解产物进行预测外,对热解过程及行为的预测有助于深入理解热解反应,RF、多元线性回归(Multiple Linear Regression, MLR)、GBDT 和 MLP 算法被用于预测固废热解的 TG 曲线,以热解温度、升温速率、热解气氛为输入,发现 RF 模型预测精度最高, R^2 和 RMSE 分别为 0.999 7 和 0.42^[25]。其他复杂深度学习算法也用于 TG 曲线

的预测, Lin 等^[26]以垃圾分类产生的废渣为对象, 采用深度学习算法研究废渣的热解行为并验证和预测其 TG 数据, 模型输入为热解温度和来自于 TG-FTIR 的吸光度特征, 建立的一维卷积神经网络-长短期记忆神经网络(1D-Convolutional Neural Network-Long Short-Term Memory, 1D-CNN-LSTM)对废渣热解 TG 数据的拟合度 R^2 为 93.91%。此外, 机器学习模型也用于共热解特性的预测。Wen 等^[27]利用 LR、RF 和 DT 模型对纺织染色污泥和熏香棒的共热解特性进行了研究, 预测了共热解 TG 以及 DTG 数据, 分析了输入特征(混合比、升温速率、载气气氛、热解温度)的重要性及交互关系, 由此指导反应过程的优化。此外, 其他模型(SVM、XGB、ANFIS 等)也用于热解反应的预测^[12, 28, 29]。

2.2.2 气化

基于机器学习的气化过程建模主要用于对合成气产率进行预测同时针对目标组分进行调控以提高气化气品质。Pandey 等^[30]利用多层前馈神经网络模型对流化床反应器中的城市固废气化过程建模, 以气化温度、当量比以及原料碳、氢、氧、氮、硫、灰分、水分含量作为输入预测合成气产率, 结果表明, 模型的预测性能与实验数据集有较好的吻合, R^2 为 98%~99%。对于混合固废, Xiao 等^[31]以木材、纸张、厨余垃圾、聚乙烯、纺织品的百分比以及气化温度、当量比作为输入变量, 建立 ANN 模型成功预测了合成气产率及热值, 训练集和测试集上的相对误差分别约 15% 和 20%, 一定程度上可以指导工业应用。除神经网络模型外, DT、XGB、RF 和 SVM 模型也用于对城市固废的气化过程进行建模^[32], 利用 PSO 算法对五种模型进行优化并预测了合成气产率和热值, 发现机器学习算法与优化算法结合的混合模型都具有较高的预测精度。

针对目标气体组分进行调控及基于模型可解释性分析优化反应参数是机器学习建模的主要优势所在。Li 等^[33]利用 4 种数据驱动的机器学习方法(RF、GBDT、SVM、ANN)对固废气化过程进行建模, 旨在对合成气成分进行预测并对气化过程进行可解释性分析进而优化气化参数。结果表明, GBDT 模型在预测三相产物分布和合成气组成成分方面性能最佳($R^2 = 0.82 \sim 0.96$), 进一步进行可解释性分析, 结合特征重要性发现原料和气化条件(原料灰分、碳、氮、氧含量和气化温度)是影

响焦炭、焦油和合成气分布的重要因素。此外通过部分依赖分析对工艺过程及原料选择进行优化, 得出 800 °C 下气化当量比在 2.7~3.0 内, 碳含量较高(>48%)、氮含量低(<0.5%)和灰分(1%~5%)的原料可以实现较高的 H_2 产率, 此结果对有机固废气化生产富氢合成气的工业应用具有指导意义。

2.2.3 燃烧

在有机固废燃烧过程中机器学习建模方法主要用于加深对燃烧反应的理解以及对燃烧工况进行在线监控。Ni 等^[34]采用 TG-FTIR 方法研究污水污泥(SS)和咖啡残渣(CIR)的共燃特性, 同时借助主成分分析(Principal Components Analysis, PCA)和 ANN 分别确定不同 SS 和 CIR 比例下共燃过程发生的主要反应并对 TG 实验数据进行预测, 第一主成分(PC1)和第一主成分(PC2)的方差贡献率分别为 54.43% 和 45.06%, PC1 主要在有机挥发物的释放阶段起主导作用, PC2 在水分蒸发和大分子物质燃烧中起主导作用, 模型预测具有高的精准度($R^2 = 0.9998$), 预测结果有助于指导原料配比和反应器优化, 从而降低工业应用成本。You 等^[35]采用 ANN、ANFIS、SVM 和 RF 模型对循环流化床焚烧炉焚烧城市固废的热值进行在线等级划分预测, 最终得出 ANFIS 模型在预测精度方面优于其他三个模型, 此外高斯过程回归模型(Gaussian Process Regression, GPR)也用于燃烧热值的在线预测^[36]。除对燃烧特性进行在线监控外还可对炉膛运行工况进行监测, Yan 等^[37]提出一种预测城市固废燃烧过程炉温的异构特征集成建模方法, 通过基础训练特征、模型池、集成模型的构建对燃烧过程炉温进行建模, 结果表明, 集成模型具有较好的预测精度、泛化能力和稳定性, 训练效率较高, 该预测模型可以指导现场操作人员及时、准确地判断炉温变化趋势, 对城市生活垃圾焚烧过程中炉温的优化控制具有重要意义。

2.2.4 水热转化

水热转化又分为水热炭化(Hydrothermal Carbonization, HTC)、水热液化(Hydrothermal Liquefaction, HTL)和水热气化(Hydrothermal Gasification, HTG), 机器学习在水热转化中的应用主要体现在对产物产率以及特性的预测, 同时借助机器学习分析方法, 可对转化过程进行优化以提高目标产物的选择性。

基于机器学习的 HTC 过程主要用于研究水

热炭特性。Li 等^[38]基于水热炭数据集开发和优化了预测水热炭特性的机器学习模型(SVM 和 RF),以原料特性和反应条件为输入预测水热炭的产率、热值、能量回收率、能量密度,使用基于博弈论的模型可解释方法(SHAP 值)探索特征对目标的影响同时探索特征重要性和相关性,进而进行特征筛选以改进模型,发现具有最优超参数下的支持向量机算法性能更优,取得了较高的预测准确率($R^2 = 0.90$)。Li 等^[17]以原料特性(C、H、N、O、固定碳、灰分、挥发分含量)和反应条件(反应时间 t , 温度 T 和水含量)为输入对比了 RF、SVM 和 DNN 模型的 HTC 过程炭产率及特性的预测性能,发现具有 2 个隐藏层(每层 128 个神经元)、学习率为 0.001 和丢弃率为 0.1 的最优 DNN 模型在对水热炭的多任务预测(产率、能量回收率、碳回收率、热值、碳含量和 H/C、O/C、N/C 原子比)中表现出最高的准确度,测试集上 R^2 为 0.91。

对于 HTL 过程,Katongtung 等^[39]采用核岭回归(Kernel Ridge Regression, KRR)、RF、SVM、XGB 将原料特性和操作条件共 17 个变量作为特征进行输入,预测湿生物质 HTL 的原油产量和高位热值,其中 XGB 预测效果最佳,进行特征重要性分析发现温度是对油产率和热值预测影响最大的特征。同时,原料特性对 XGB 模型决策的贡献超过 55%,此外,部分依赖图(Partial Dependence Plot, PDP)分析揭示了重要特征对输出结果的影响及各特征之间的相互作用,有助于深入理解 HTL 反应过程。Shafizadeh 等^[40]使用 ANN、广义加性模型(Generalized Additive Model, GAM)、SVM 和 GPR 对固废水热液化过程进行建模,其中 GPR 模型预测精度最高($R^2 = 0.926$),进一步基于 SHAP 方法和 PDP 分析详细评估了输入变量对输出响应的重要性及影响趋势并优化反应工况,最大限度地减少生物油的杂原子(N、O)含量。此外机器学习模型还用于多任务回归预测,Li 等^[41]采用机器学习模型(DT、RF、GBDT)以原料特性、工艺参数、溶剂特性为输入,同时预测 HTL 过程生物油产量、氮含量和能量回收率,并利用 PSO 算法优化实验条件,最终通过实验验证了模型预测的可靠性。

对于 HTG 过程,机器学习模型主要用于预测 H_2 产率以及温度、压力、进料浓度和反应时间等参数对 H_2 产率和选择性的影响。Gopirajan 等^[42]提出了一种基于可调决策支持系统和可调推荐系统

的机器学习方法,以原料特性、操作条件、催化参数作为特征进行输入,用于预测 HTG 过程中 H_2 产率并优化反应条件(温度、压力、催化剂用量、反应时间和溶剂与原料的比例),最终预测结果与实际结果相近,预测准确率高于 94%。Li 等^[43]开发了一种耦合数据驱动和机理建模方法用于湿生物质的 HTG 过程,以原料特性和操作参数作为输入,将 GBDT 与 PSO 算法有机结合构建集成优化模型来预测湿生物质 HTG 合成气的产量并优化气体组成,然后通过 Aspen Plus 模拟验证和解释 GBDT 模型的预测结果。此外,通过 Aspen Plus 对基于机器学习筛选的最优工艺条件进行了模拟验证,误差约为 10% 和 20%,且 Aspen Plus 模拟解释反应温度和原料固含量是对 H_2 生成影响最大的因素,该研究提出了数据驱动模型和机理模型的集成建模方法并阐述二者如何相互补充与验证,有助于加速机器学习对 HTG 的实验设计。

2.3 机器学习指导处置产物利用及过程评估预测

热转化处置产物的利用及过程效益评估是有机固废全链条管理与处置中的下游环节,运用机器学习建模方法对下游应用及评估进行预测可完善中游处置的效益评估及产物的高值利用过程,有助于实现有机固废全环节智能化管理与处置。

有机固废处置后的下游环节,机器学习可辅助处置产物的高值利用并进行风险监测。Zhu 等^[44]通过机器学习模拟了热解炭对重金属的吸附过程,使用 ANN 和 RF 模型对生物炭吸附六种重金属(Pb、Cd、Ni、As、Cu 和 Zn)进行建模,根据生物炭特性(pH、比表面积、阳离子交换量、灰分含量、粒径、C 含量、(O+N)/C、O/C、H/C)、金属特性(电荷数、离子半径、电负性)、吸附条件(温度和溶液 pH)以及重金属与生物炭的初始浓度比共 15 个特征预测生物炭的吸附能力,发现 RF 模型($R^2 = 0.97$)预测性能优于 ANN 模型($R^2 = 0.95$)。特征重要性分析表明生物炭特性对吸附效果的影响最为显著(54%),其中阳离子交换量和 pH 占生物炭特性中的 66%,而生物炭比表面积仅占 2%。所开发的模型可根据现有条件直接预测生物炭对目标金属的去除效率,从而减少实验与试错成本,特征重要性分析为高效处理废水中重金属所需的生物炭特性提供重要的调控方向。

Liu 等^[45]广泛收集了包含城市生活垃圾燃烧飞灰特性和重金属浸出的数据集,进行了基于 XGB 模型和博弈分析的可解释机器学习方法来预

测 6 种重金属 (Cr、Cd、Cu、Ni、Pb 和 Zn) 的浸出潜力,发现稳定的飞灰中 Cd 和 Pb 的超标风险仍然很高。其中,可溶性氯含量是影响预测重金属浸出行为的最重要因素,结合 SHAP 分析量化了可溶性氯对 Cd、Cu、Pb 和 Zn 重金属浸出的影响,其每增加 0.5 mol/L 数量级将导致 Cd、Cu、Pb 和 Zn 的浸出浓度将分别增加 223.5%、215.2%、216.5% 和 222.6%。最后,提出了飞灰垃圾填埋场的快速判定标准,当垃圾燃烧后产生的飞灰中可溶性氯含量小于 0.412 mol/L 时是可以接受的,否则需采取进一步措施以稳定飞灰。该模型的提出对燃烧产生的飞灰进行了风险监测,完善了固废处置全环节的管理。此外,机器学习模型还用于对有机固废燃烧过程中二噁英的生成进行在线监测^[46],实现处置过程的实时管理。

在水热转化过程中会不可避免地产生大量的副产物废水(水相),如何处理和利用这种副产物是水热转化技术的瓶颈问题之一。Leng 等^[47]对比和优化了 GBDT 和 RF 模型,以元素组成(C、H、O、S、N 和灰分含量)、组分分布(蛋白质、脂质和碳水化合物含量)和操作条件(温度、反应时间、固体含量)作为输入,预测固废水热处理后所得水相的 pH、总氮、总有机碳和总磷含量,结果表明 GBDT 模型可以准确预测单目标和多目标的水相特性。基于特征重要性和部分依赖分析表明,原料 N 含量、固含量和温度是预测水相 pH、总氮和总磷含量的三个最关键的特征,而对于总有机碳含量的预测则是反应时间、脂质含量和温度三个因素最为重要。该研究提出的 ML 模型为水热转化水相的处置和利用提供了一定的指导意义。

在实际有机固废处置与管理过程中对整个处

置系统的能源、环境、经济效益进行评估具有重要意义。Cheng 等^[48]集成了一个机器学习建模、生命周期评估(Life Cycle Assessment, LCA)和经济分析框架,评估了慢速热解系统作为一种产能负碳排放技术的潜力。基于 RF 模型开发了用于预测生物炭产量和慢速热解系统的能源生产、碳排放和经济可行性,具体为能源投资回报率(Energy Return On Investment, EROI)、全球变暖潜能值(Global Warming Potential, GWP)、热解炭的最低产品售价(Minimum Product Selling Price, MPSP)。结果表明,RF 模型在基于历史经验数据集来预测实验室规模和中试规模的固废热解途径方面具有巨大潜力。LCA 结果表明,固废的缓慢热解提供了正的 EROI 和负 GWP 排放,高温有利于能源生产和负碳排放,但低的反应温度下经济性会更好。

3 不同机器学习模型适用场景及特性对比

基于上述机器学习建模方法在有机固废管理与处置的上、中、下游环节中的应用,表 1 为不同模型的适用场景及特性对比,经典模型(KNN、LR、PLR、RR、SVM)主要用于有机固废管理与处置上游环节的产生量、特性预测,树模型(DT、RF、GBDT、XGB)主要运用于中游热转化及下游应用环节的预测,神经网络模型则贯穿整个有机固废管理与处置环节,是运用最多的模型。各类模型的适用场景既与目标问题的复杂程度有关同时也与其自身特性有关(见表 1),基于树结构集成和神经网络的模型通常可以处理复杂的非线性问题,若想进一步提升预测精度及模型的可解释性,融合不同模型的优势,则可考虑使用混合模型。

表 1 不同机器学习模型适用场景及特性对比

Table 1 Comparison of applicable scenarios and characteristics of different machine learning models

机器学习模型	适用场景	优势	不足
经典模型 (KNN、LR、PLR、RR、SVM)	主要用于固废管理与处置上游环节固废的产生量、特性(含水率、HHV)预测 ^[18, 19, 21]	模型简单,适用于小样本数据,SVM 自身可抗过拟合	对异常值及缺失值敏感,KNN、SVM 对于大样本数据训练效率低
树模型 (DT、RF、GBDT、XGB)	用于中游处置、下游应用环节,热解行为及产物 ^[25, 44] 、气化气产率及组分 ^[32, 33] 、水热转化的预测 ^[39, 47]	易解释,预测性能好,模型训练速度快,自动处理缺失值	DT 对非线性问题拟合度较弱,集成模型容易过拟合
神经网络模型 (ANN、ANFIS、DNN)	贯穿于整个管理与处置环节,上游固废特性 ^[21] 、中游各类热转化过程 ^[17, 23, 26, 30, 34] 、下游应用预测 ^[44]	预测性能好,处理复杂非线性问题能力强	容易过拟合,模型可解释性差、训练速度较慢
混合模型 (PSO-RF、PSO-XGB、PSO-ANN)	主要用于中游环节复杂热化学转化过程的建模、产物分析及其特性预测 ^[32, 43]	混合模型预测精度高,可实现多模型的优势互补	模型训练速度相对较慢,存在过拟合风险

4 结 论

本文主要对基于机器学习的智能建模方法在有机固废热转化处置中的应用进行了总结,同时对上游固废产生、特性预测以及下游产物利用与风险评估中的机器学习应用进行了梳理,基于数据驱动的机器学习模型可以对固废管理与处置过程进行精准的建模预测,进而实现处置工艺的反向设计与优化、产物预测与调控(高值利用)、风险评估与监测,并用于构建智能化处置体系。此外,耦合机器学习的固废处置应朝着更精准、更高效、更智能的方向发展:(1)为提高模型的预测准确度与泛化能力,可以考虑混合模型的建立(相比于单个模型,可以实现多模型集成的优势互补),混合模型往往具有更高的预测准确性,更低的过拟合风险和更强的鲁棒性;(2)将基于数据驱动的机器学习模型与理论驱动的理论模型有机结合,理论驱动的理论模型更容易解释热转化激励并有很好的理论支撑,而数据驱动的理论模型更具灵活性和拓展性,二者可以相互补充以更好地剖析具体问题;(3)进一步增强机器学习模型的可解释性分析以突破传统的“黑匣子”困境,机器学习建模可以得到较为精确的结果,为进一步理解模型内在决策过程以更好指导实际生产,应开发更多类似特征重要性、部分依赖分析、SHAP 分析等可解释性分析方法;(4)需要基于更多人工智能的工具和框架来搭建高效的数据处理系统从而有效管理固废处置,形成高效、智能的管理与处置系统,优化固废全链条处置的各个环节,从而更好地服务实际生产应用。

参考文献 (References) :

[1] Kaza Silpa, Yao Lisa, Bhada-Tata Perinaz, et al. What a waste 2.0: A global snapshot of solid waste management to 2050 [M]. World Bank Publications, 2018.

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

[3] Ali Asim, Mahar Rasool Bux, Soomro Razium Ali, et al. Fe₃O₄ nanoparticles facilitated anaerobic digestion of organic fraction of municipal solid waste for enhancement of methane production [J]. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2017, 39(16): 1815-1822.

[4] Guo Haonan, Wu Shubiao, Tian Yingjie, et al. Application of machine learning methods for the prediction of organic solid waste treatment and recycling processes: A review [J]. Biore-source Technology, 2021, 319: 124114.

[5] Triassi Maria, Alfano Rossella, Illario Maddalena, et al. Environmental pollution from illegal waste disposal and health effects: A review on the “Triangle of Death” [J]. International Journal of Environmental Research and Public Health, 2015, 12(2): 1216-1236.

[6] Salman Chaudhary Awais, Schwede Sebastian, Thorin Eva, et al. Predictive modelling and simulation of integrated pyrolysis and anaerobic digestion process [J]. Energy Procedia, 2017, 105: 850-857.

[7] Wang Hanxi, Xu Jianling, Yu Haixia, et al. Study of the application and methods for the comprehensive treatment of municipal solid waste in northeastern China [J]. Renewable and Sustainable Energy Reviews, 2015, 52: 1881-1889.

[8] 苏毅, 朱惠春, 张金亮, 等. 城市垃圾热化学转化处理技术进展与应用 [J]. 工业锅炉, 2015(1): 7-14.

SU YI, ZHU Huichun, ZHANG Jinliang, et al. Progress and application of thermochemical treatment technology for municipal solid waste [J]. Industrial Boiler, 2015(1): 7-14.

[9] 孙浩胜, 秦俊, 王君. 有机固废热化学处置研究进展 [C]// 中国环境科学学会: 2021 年全国有机固废处理与资源化利用高峰论坛, 2021: 97-104.

[10] Wang Zhengxin, Peng Xinggan, Xia Ao, et al. The role of machine learning to boost the bioenergy and biofuels conversion [J]. Bioresource Technology, 2022, 343: 126099.

[11] Ihsanullah I, Alam Gulzar, Jamal Arshad, et al. Recent advances in applications of artificial intelligence in solid waste management: A review [J]. Chemosphere, 2022: 136631.

[12] Tang Qinghui, Chen Yingquan, Yang Haiping, et al. Machine learning prediction of pyrolytic gas yield and compositions with feature reduction methods: Effects of pyrolysis conditions and biomass characteristics [J]. Bioresource Technology, 2021, 339: 125581.

[13] Zhu Xinzhe, Li Yanan, Wang Xiaonan. Machine learning prediction of biochar yield and carbon contents in biochar based on biomass characteristics and pyrolysis conditions [J]. Biore-source Technology, 2019, 288: 121527.

[14] Onsree Thossaporn, Tippayawong Nakorn. Machine learning application to predict yields of solid products from biomass torrefaction [J]. Renewable Energy, 2021, 167: 425-432.

[15] Hart Sergiu. Shapley Value [M]//EATWELL JOHN, MILGATE MURRAY, NEWMAN PETER. Game Theory. London: Palgrave Macmillan UK, 1989: 210-216.

[16] Ullah Zahid, Khan Muzammil, Raza Naqvi Salman, et al. A comparative study of machine learning methods for bio-oil yield prediction - A genetic algorithm-based features selection [J]. Bioresource Technology, 2021, 335: 125292.

[17] Li Jie, Zhu Xinzhe, Li Yanan, et al. Multi-task prediction and optimization of hydrochar properties from high-moisture municipal solid waste: Application of machine learning on waste-to-resource [J]. Journal of Cleaner Production, 2021, 278: 123928.

[18] Nguyen X Cuong, Nguyen T Thanh Huyen, La D Duong, et al. Development of machine learning - based models to forecast

- solid waste generation in residential areas: A case study from Vietnam[J]. *Resources, Conservation and Recycling*, 2021, 167: 105381.
- [19] Qi Ya Ping, He Pin Jing, Lan Dong Ying, et al. Rapid determination of moisture content of multi-source solid waste using ATR-FTIR and multiple machine learning methods[J]. *Waste Management*, 2022, 153: 20-30.
- [20] Tao Junyu, Liang Rui, Li Jian, et al. Fast characterization of biomass and waste by infrared spectra and machine learning models [J]. *Journal of Hazardous materials*, 2020, 387: 121723.
- [21] Taki Morteza, Rohani Abbas. Machine learning models for prediction the Higher Heating Value (HHV) of Municipal Solid Waste (MSW) for waste-to-energy evaluation [J]. *Case Studies in Thermal Engineering*, 2022, 31: 101823.
- [22] 顾春晗,汪小愁,宋谦石,等.基于BP神经网络的固体废弃物热解产物分布规律预测[J].*燃烧科学与技术*,2022(2): 133-140.
- GU Chunhan, WANG Xiaohan, SONG Qianshi, et al. Prediction of distribution law of solid waste pyrolysis products based on bp neural network [J]. *Combustion Science and Technology*, 2022(2): 133-140.
- [23] Gu Chunhan, Wang Xiaohan, Song Qianshi, et al. Prediction of gas-liquid-solid product distribution after solid waste pyrolysis process based on artificial neural network model[J]. *International Journal of Energy Research*, 2021, 45(9): 13786-13800.
- [24] Leng Erwei, He Ben, Chen Jingwei, et al. Prediction of three-phase product distribution and bio-oil heating value of biomass fast pyrolysis based on machine learning[J]. *Energy*, 2021, 236: 121401.
- [25] Zhang Junhui, Liu Jingyong, Evrendilek Fatih, et al. TG-FTIR and Py-GC/MS analyses of pyrolysis behaviors and products of cattle manure in CO₂ and N₂ atmospheres: Kinetic, thermodynamic, and machine-learning models[J]. *Energy Conversion and Management*, 2019, 195: 346-359.
- [26] Lin Kunsen, Tian Lu, Zhao Youcai, et al. Pyrolytic characteristics of fine materials from municipal solid waste using TG-FTIR, Py-GC/MS, and deep learning approach: Kinetics, thermodynamics, and gaseous products distribution[J]. *Chemosphere*, 2022, 293: 133533.
- [27] Wen Shaoting, Buyukada Musa, Evrendilek Fatih, et al. Uncertainty and sensitivity analyses of co-combustion/pyrolysis of textile dyeing sludge and incense sticks: Regression and machine-learning models [J]. *Renewable Energy*, 2020, 151: 463-474.
- [28] Pathy Abhijeet, Meher Saswat, P Balasubramanian. Predicting algal biochar yield using eXtreme Gradient Boosting (XGB) algorithm of machine learning methods [J]. *Algal Research*, 2020, 50: 102006.
- [29] Li Yize, Gupta Rohit, You Siming. Machine learning assisted prediction of biochar yield and composition via pyrolysis of biomass[J]. *Bioresource Technology*, 2022, 359: 127511.
- [30] Pandey Daya Shankar, Das Saptarshi, Pan Indranil, et al. Artificial neural network based modelling approach for municipal solid waste gasification in a fluidized bed reactor [J]. *Waste Management*, 2016, 58: 202-213.
- [31] Xiao Gang, Ni Ming-jiang, Chi Yong, et al. Gasification characteristics of MSW and an ANN prediction model [J]. *Waste Management*, 2009, 29(1): 240-244.
- [32] Kardani Navid, Zhou Annan, Nazem Majidreza, et al. Modeling of municipal solid waste gasification using an optimised ensemble soft computing model[J]. *Fuel*, 2021, 289: 119903.
- [33] Li Jie, Li Lanyu, Tong Yen Wah, et al. Understanding and optimizing the gasification of biomass waste with machine learning[J/OL]. *Green Chemical Engineering*, 2022: 1-12 [2022-12-01]. <https://doi.org/10.1016/j.gce.2022.05.006>.
- [34] Ni Zhanshi, Bi Haobo, Jiang Chunlong, et al. Investigation of co-combustion of sewage sludge and coffee industry residue by TG-FTIR and machine learning methods [J]. *Fuel*, 2022, 309: 122082.
- [35] You Haihui, Ma Zengyi, Tang Yijun, et al. Comparison of ANN (MLP), ANFIS, SVM, and RF models for the online classification of heating value of burning municipal solid waste in circulating fluidized bed incinerators [J]. *Waste Management*, 2017, 68: 186-197.
- [36] Birgen Cansu, Magnanelli Elisa, Carlsson Per, et al. Machine learning based modelling for lower heating value prediction of municipal solid waste[J]. *Fuel*, 2021, 283: 118906.
- [37] Yan Aijun, Guo Jingcheng, Wang Dianhui. Heterogeneous feature ensemble modeling with stochastic configuration networks for predicting furnace temperature of a municipal solid waste incineration process[J]. *Neural Computing and Applications*, 2022, 34(18): 15807-15819.
- [38] Li Jie, Pan Lanjia, Suvama Manu, et al. Fuel properties of hydrochar and pyrochar: Prediction and exploration with machine learning[J]. *Applied Energy*, 2020, 269: 115166.
- [39] Katontung Tossapon, Onsree Thossaporn, Tippayawong Nakorn. Machine learning prediction of biocrude yields and higher heating values from hydrothermal liquefaction of wet biomass and wastes[J]. *Bioresource Technology*, 2022, 344: 126278.
- [40] Shafizadeh Alireza, Shahbeig Hossein, Nadian Mohammad Hossein, et al. Machine learning predicts and optimizes hydrothermal liquefaction of biomass [J]. *Chemical Engineering Journal*, 2022, 445: 136579.
- [41] Li Jie, Zhang Weijin, Liu Tonggui, et al. Machine learning aided bio-oil production with high energy recovery and low nitrogen content from hydrothermal liquefaction of biomass with experiment verification [J]. *Chemical Engineering Journal*, 2021, 425: 130649.
- [42] Gopirajan P V, Gopinath K P, Sivaranjani G, et al. Optimization of hydrothermal gasification process through machine learning approach: Experimental conditions, product yield and pollution [J]. *Journal of Cleaner Production*, 2021, 306:172302

- [43] Li Jie, Suvarna Manu, Pan Lanjia, et al. A hybrid data-driven and mechanistic modelling approach for hydrothermal gasification[J]. *Applied Energy*, 2021, 304: 117674.
- [44] Zhu Xinzhe, Wang Xiaonan, Ok Yong Sik. The application of machine learning methods for prediction of metal sorption onto biochars [J]. *Journal of Hazardous materials*, 2019, 378: 120727.
- [45] Liu Zeyuan, Lu Min, Zhang Yufeng, et al. Identification of heavy metal leaching patterns in municipal solid waste incineration fly ash based on an explainable machine learning approach[J]. *Journal of Environmental Management*, 2022, 317: 115387.
- [46] Xia Wanjun, Jiang Yanping, Chen Xiaohong, et al. Application of machine learning algorithms in municipal solid waste management: A mini review [J]. *Waste Management & Research*, 2021, 40(6) : 609–624.
- [47] Leng Lijian, Zhang Weijin, Liu Tonggui, et al. Machine learning predicting wastewater properties of the aqueous phase derived from hydrothermal treatment of biomass [J]. *Bioresource Technology*, 2022, 358: 127348.
- [48] Cheng Fangwei, Luo Hongxi, Colosi Lisa M. Slow pyrolysis as a platform for negative emissions technology: An integration of machine learning models, life cycle assessment, and economic analysis [J]. *Energy Conversion and Management*, 2020, 223: 113258.