Wiley
International Journal of Energy Research
Volume 2024, Article ID 6087208, 36 pages
https://doi.org/10.1155/2024/6087208

*Review Article*

# Exploring Insights in Biomass and Waste Gasification via Ensemble Machine Learning Models and Interpretability Techniques

**Ocident Bongomin**,[1,2] **Charles Nzila**,[1] **Josphat Igadwa Mwasiagi**,[1] **and Obadiah Maube**[3]

[1]*Department of Manufacturing, Textile and Industrial Engineering, School of Engineering, Moi University, P.O. Box 3900-30100, Eldoret, Kenya*
[2]*Africa Centre of Excellence II in Phytochemical, Textile and Renewable Energy (ACE II-PTRE), Moi University, P.O. Box 3900-30100, Eldoret, Kenya*
[3]*Department of Industrial and Mechatronics Engineering, School of Mechanical and Manufacturing Engineering, Technical University of Kenya, P.O. Box 52428-00200, Nairobi, Kenya*

Correspondence should be addressed to Ocident Bongomin; ocidentbongomin@gmail.com

Received 8 January 2024; Revised 2 May 2024; Accepted 13 May 2024

Academic Editor: Panagiotis Grammelis



This comprehensive review delves into the intersection of ensemble machine learning models and interpretability techniques for biomass and waste gasification, a field crucial for sustainable energy solutions. It tackles challenges like feedstock variability and temperature control, highlighting the need for deeper understanding to optimize gasification processes. The study focuses on advanced modeling techniques like random forests and gradient boosting, alongside interpretability methods like the Shapley additive explanations and partial dependence plots, emphasizing their importance for transparency and informed decision-making. Analyzing diverse case studies, the review explores successful applications while acknowledging challenges like overfitting and computational complexity, proposing strategies for practical and robust models. Notably, the review finds ensemble models consistently achieve high prediction accuracy (often exceeding $R^2$ scores of 0.9) for gas composition, yield, and heating value. These models (34% of reviewed papers) are the most applied method, followed by artificial neural networks (26%). Heating value (12%) was the most studied performance metric. However, interpretability is often neglected during model development due to the complexity of techniques like permutation and Gini importance. The paper calls for dedicated research on utilizing and interpreting ensemble models, especially for co-gasification scenarios, to unlock new insights into process synergy. Overall, this review serves as a valuable resource for researchers, practitioners, and policymakers, offering guidance for enhancing the efficiency and sustainability of biomass and waste gasification.

## 1. Introduction

Gasification, a transformative thermochemical process, holds a crucial role in converting a variety of carbon-based feedstocks, ranging from coal to biomass, plastics, sewage sludge, and municipal solid waste, into syngas [1]. This syngas, composed of methane ($CH_4$), hydrogen ($H_2$), carbon monoxide (CO), and carbon dioxide ($CO_2$), stands as a versatile resource for generating hydrogen gas, heat, or electricity through combustion. Biomass and coal, recognized as primary feedstocks for gasification, have significant implications for the energy landscape [2–4]. Despite its substantial benefits, challenges persist, including the assurance of syngas quality and addressing feedstock supply shortages for continuous operation [5–7].

In response to these challenges, co-gasification has emerged as a focal point of attention, presenting a promising solution to maintain a consistent feedstock supply and enhance gasification efficiency [7–9]. Co-gasification, particularly in the realm of coal-biomass co-gasification, offers

environmental advantages by minimizing tar formation and reducing pollutant emissions compared to conventional coal gasification processes [10–13]. However, the intricacies of gasification reactions, influencing syngas composition and heating value, remain persistent challenges for process control [14, 15]. Factors such as feedstock properties, reactor types, gasification agents, and operational conditions contribute to the complexity of syngas quality and composition [16–19].

The realm of biomass gasification is undergoing exciting advancements. Researchers are exploring methods to achieve ultralow carbon dioxide emissions, with Zhu et al. [20] proposing a system capturing and converting nearly 100% of $CO_2$. A key area of focus is the production of hydrogen-rich syngas, a versatile fuel precursor, as evidenced in the review by Makwana et al. [21]. Beyond conventional methods, innovative approaches like catalytic gasification [22, 23], chemical looping gasification [24–26], supercritical water gasification [27–29], and microwave-assisted gasification [30, 31] are being investigated. Economic viability is also a key focus, with studies comparing production methods for valuable products like methanol [32, 33] and exploring solar-assisted techniques for hydrogen and chemical recovery [34]. Furthermore, research is ongoing to optimize the gasification of organic solid waste (D. [35]) and delve deeper into reactor design, including evaluations of supercritical water gasification systems [36] and the integration of new reactors like plasma [37], multistage [38], and water-gas shift units [39–41]. Life cycle assessments are emphasizing the environmental and economic benefits of these advancements, with Fang et al. [42] analyzing concentrated solar thermal gasification for sustainable electricity generation. This focus on innovation and efficiency underscores the potential of gasification to become a cornerstone of a sustainable biorefinery industry.

Optimizing gasification processes hinges on a suite of powerful modeling techniques. Thermodynamic equilibrium models, while simple to use, offer a potentially limited view by assuming the system reaches a perfect state of balance [43]. For a more nuanced understanding, kinetic models incorporate reaction rates to predict how gasification unfolds over time [44]. Computational fluid dynamics (CFD) or numerical models delve even deeper, using complex simulations to analyze gas flow, temperature distribution, and reaction behavior within the gasifier itself [45–47]. Process simulation models provide a comprehensive perspective by integrating aspects of thermodynamics, kinetics, and reactor design to simulate the entire gasification system [48, 49]. Among these methods, machine learning (ML) models are rapidly gaining favor. Their ability to analyze vast datasets and identify hidden patterns makes them adept at optimizing gasification processes [50–52]. ML is a subfield of artificial intelligence (AI) that equips machines with the ability to learn from data without explicit programming, enabling data-driven predictive analytics across various industrial sectors [53–56]. ML models have the ability to continuously improve their performance on a specific task as they are exposed to more data. In the context of gasification, this translates to tasks like predicting optimal operating conditions, product yields, and even gasifier control strategies [57, 58].

Ensemble models, which combine the strengths of multiple ML algorithms, are particularly attractive due to their enhanced accuracy and robustness [59]. However, a key challenge remains: interpretability. The complex decision-making processes within ensemble models can be difficult to decipher, hindering transparency for users like managers and policymakers who need to understand the model's rationale [60–62]. This is a crucial point, as advancements in areas like hydrogen production [34] and waste management [35] hinge on clear communication and trust in the underlying models. Despite this hurdle, the versatility and power of modeling techniques, particularly in the rapidly evolving field of ML, offer a compelling path towards optimizing gasification processes for a more efficient and sustainable future. As research into interpretability techniques progresses, we can expect ML models, as a powerful subset of AI, to play an even greater role in unlocking the full potential of biomass gasification.

The latest research highlights the growing importance of ML in optimizing and understanding biomass thermochemical conversion processes [63]. Studies are applying ML models to predict a variety of factors, including biochar yield and surface area [64], total hydrogen production cost [65], and bio-oil production from pyrolysis [66]. ML is also being utilized to optimize processes themselves. For instance, researchers have developed a neural network model to control fluidized bed gasification [67] and optimize concentrated solar thermal gasification [68]. Additionally, ML can be used to analyze data from sensors for tasks like fuel feeding rate estimation [69] and determine optimal operating parameters for gasification [70]. The field encompasses a broad range of applications, including the design and optimization of catalysts [71], and even predicting solar-to-liquid fuel production [72]. This integration of ML across the entire conversion process chain, from feedstock analysis to product prediction, signifies its potential to revolutionize biomass conversion for a more efficient and sustainable biorefinery industry [73]. For further exploration, refer to related works on ML in hydrogen production, gasification control, and bioethanol production [74–76].

This paper tackles a critical gap in biomass and waste gasification research. Existing reviews, as shown in Table 1, have explored various facets like thermodynamics, kinetics, and general modeling approaches. However, these reviews often lack a focused exploration of two powerful tools: ensemble machine learning (ML) models and interpretability techniques. The emphasis on interpretability is crucial for gasification. These processes are inherently complex, with numerous variables impacting factors like syngas yield, tar formation, and overall efficiency. Traditional "black box" models, while effective in prediction, often lack transparency, making it difficult to understand the rationale behind their outputs. This is a significant hurdle for gasification optimization, as understanding the underlying decision-making processes within the model is vital for targeted improvements. Interpretable models, on the other hand, provide valuable insights into how these variables interact and influence the model's predictions. This allows researchers and engineers to not only predict outcomes but

Table 1: Comparison of the present review and published reviews on ML-aided thermochemical conversion of biomass and wastes.

| Study | Overview of biomass and waste gasification | Detailed discussion of ensemble models | Machine learning applications | Detailed exploration of interpretability analysis | Challenges and future directions |
|---|---|---|---|---|---|
| Liao and Yao [77] | ✗ | ✗ | ✓ | ✗ | ✓ |
| Zhang et al. [78] | ✗ | ✗ | ✓ | ✗ | ✓ |
| Umenweke et al. [79] | ✓ | ✗ | ✓ | ✗ | ✓ |
| Sedej et al. [80] | ✓ | ✗ | ✓ | ✗ | ✗ |
| Kushwah et al. [51] | ✓ | ✗ | ✓ | ✗ | ✓ |
| Ascher et al. [81–83] | ✓ | ✗ | ✓ | ✗ | ✓ |
| Zhang et al. [84] | ✓ | ✗ | ✓ | ✗ | ✓ |
| Khan et al. [85] | ✓ | ✗ | ✓ | ✗ | ✓ |
| Sakheta et al. [86] | ✗ | ✗ | ✓ | ✗ | ✓ |
| Alfarra et al. [87] | ✓ | ✗ | ✓ | ✗ | ✓ |
| Bin Abu Sofian et al. [88] | ✓ | ✗ | ✓ | ✗ | ✗ |
| This review | ✓ | ✓ | ✓ | ✓ | ✓ |

also to identify key factors that can be manipulated to optimize gasification processes for sustainability.

This review highlights the potential of ensemble ML models, such as random forests, to achieve this goal. Ensemble models combine the strengths of multiple individual models, often leading to more robust and accurate predictions compared to single models. By incorporating interpretability techniques like the Shapley additive explanations (SHAP) and PDP (partial dependence plots), the review explores how researchers can gain a deeper understanding of the decision-making processes within these ensemble models. This allows for targeted improvements in gasification technology by focusing on the most influential factors identified by the interpretability techniques. By showcasing successful applications of ensemble ML models and interpretability techniques in various gasification case studies, the paper provides a balanced perspective. It acknowledges the limitations of these methods, such as the computational complexity of some interpretability techniques and the potential challenges in generalizing models across diverse feedstocks and gasification setups. However, the review argues that the benefits of interpretability outweigh these limitations, particularly in the context of complex and dynamic processes like gasification.

Furthermore, the paper outlines future research directions, offering a roadmap for researchers and policymakers seeking to leverage these methods for heightened efficiency and sustainability in biomass and waste gasification processes. It delves into the critical aspect of interpretation, emphasizing the need for a profound understanding of decision-making processes within gasification models. Most notably, it highlights the potential oversight in previous studies, where interpretability has been neglected in model development goals, possibly due to the challenge of understanding and implementing interpretability approaches. Generally, this paper's innovation lies in its methodical focus on ensemble ML methods and interpretability within the broader spectrum of gasification studies. It is not merely a compilation of existing methodologies but an intricate exploration that bridges gaps in the current understanding of gasification processes. By providing in-depth insights into the role of ensemble models and interpretability methods, the paper contributes a valuable guide for researchers, practitioners, and policymakers aiming to enhance the efficiency and sustainability of gasification technologies.

## 2. Overview of Biomass and Waste Gasification

Gasification, a thermochemical process, converts carbon-containing materials into syngas, mainly consisting of CO, $H_2$, $CO_2$, and $CH_4$. Involving heating biomass or waste in a high-temperature reactor with controlled oxygen or steam, gasification prevents complete combustion. Diverse feedstock options, including agricultural residues and plastics, undergo preparation [89]. Gasifiers, complex systems with various types, are classified based on physical structure [90]. These include fixed bed (updraft and downdraft), fluidized bed (bubbling, circulating, and dual), entrained flow, plasma, and rotary kiln gasifiers [91, 92]. The gasification agent, such as air, carbondioxide, oxygen, or steam, varies in each type, influencing chemical reactions [93]. Molino et al. [94] extensively discuss the advantages and disadvantages of different reactors, including rotary kiln and plasma.

Gasification agents offer flexibility for outcomes, emphasizing hydrogen production or minimizing carbon dioxide emissions [95]. Syngas cleanup is crucial for environmental sustainability, ensuring suitability for power generation, chemical synthesis, and fuel production. Gasification proves promising for sustainable and efficient energy solutions, contributing to the circular economy with minimal environmental impact [96, 97]. Co-gasification, an innovative approach, involves simultaneously gasifying multiple feedstocks in a shared reactor [98]. Combining materials like agricultural residues and plastics exploits synergies in their structures and compositions [99]. The cross-interaction enhances the conversion of tarry compounds into valuable gaseous products [100, 101]. Co-gasification, whether in gasifiers or thermogravimetric analyzers, optimizes resource

utilization, improves efficiency, and contributes to sustainable energy solutions while addressing environmental challenges associated with organic waste.

*2.1. Innovative Gasification Concepts.* Innovative gasification concepts are transforming biomass and waste conversion. Chemical looping gasification utilizes metal oxide carriers for oxygen separation, enhancing efficiency and reducing emissions [102–104]. Catalytic gasification employs catalysts to expedite reactions, boosting efficiency and operational flexibility. Supercritical water gasification (SCWG) also known as hydrothermal gasification, operating at elevated temperatures and pressures, efficiently converts biomass into syngas without external gasifying agents [28, 105–107]. It is ideal for biomass and waste with high moisture content, considered an environmentally friendly and cost-effective approach for hydrogen production [108–110]. Distinctions between conventional and hydrothermal gasification are explored in detail by Umenweke et al. [79].

Gasification processes have evolved with two and three-stage implementations. Two-stage gasification involves an initial phase of pyrolysis or partial oxidation, producing intermediates that undergo a second stage for further conversion into syngas [111–113]. Three-stage gasification adds an additional step, providing enhanced control over reaction pathways and product composition [114, 115]. These advancements showcase the dynamic nature of gasification technologies, offering diversified approaches to meet specific energy and environmental goals. Ongoing research in these concepts holds promise for more efficient and sustainable energy solutions through biomass and waste conversion [69, 116]. Makwana et al. [21] comprehensively reviewed waste/biomass gasification for hydrogen-rich syngas production, providing an overview of gasification technologies.

*2.2. Applications and Significance.* Biomass and waste gasification, a versatile thermochemical process, is a forefront solution for sustainable energy and waste management, converting carbon-containing materials into syngas [117]. The produced syngas serves as a valuable fuel source for various power generation applications (such as gas engines, turbines, and combined heat and power systems), contributing to diversified energy sources and reducing reliance on conventional fossil fuels [118, 119]. Beyond power generation, applications extend to direct heating in industrial, district, and residential processes.

Gasification integrates with biochemical processes, enabling the production of biofuels, biochemicals, biochar, and other high-value products, supporting the sustainable evolution of the energy landscape [119–121]. The versatility of syngas extends to its refinement for synthetic fuels like synthetic natural gas (SNG), hydrogen, and liquid biofuels, offering cleaner burning options in transportation and industry [117]. Additionally, gasification plays a crucial role in waste treatment, reducing the volume of diverse waste streams, from municipal solid waste to agricultural residues, offering an ecofriendlier disposal method compared to traditional means.

The significance of biomass and waste gasification extends beyond immediate applications, encompassing environmental sustainability, reduced greenhouse gas emissions, and economic opportunities. By contributing to renewable energy sources, gasification helps diminish dependence on finite fossil fuels. Waste valorization provides dual benefits of efficient waste management and resource recovery [29, 122]. The reduced environmental impact of gasification, compared to traditional waste disposal methods, aligns with sustainability principles, offering a cleaner and more responsible approach to energy production and waste treatment [123]. Acknowledging these advantages, ongoing research and development efforts are crucial to overcoming challenges and optimizing the adoption of biomass and waste gasification technologies on a larger scale.

*2.3. Challenges and Opportunities.* Biomass and waste gasification represent cutting-edge technologies that hold great promise for converting organic materials into valuable energy products. This process involves the transformation of biomass and waste into a gaseous mixture known as syngas, comprising carbon monoxide, hydrogen, and methane [39]. While the potential benefits include the generation of clean energy and waste reduction, numerous challenges underscore the complexity of these gasification processes, ranging from feedstock variability to economic viability and environmental concerns [124].

*2.3.1. Process Challenges.* The gasification process is a complex interplay of thermochemical reactions occurring across solid, liquid, and gas phases, compounded by the inherent heterogeneity of biomass and waste feedstocks [125]. Variability in composition, moisture content, and physical properties introduces uncertainties affecting efficiency and reliability. Feedstocks may contain contaminants like sulfur, chlorine, and alkali metals, leading to corrosion, fouling, and undesirable by-products during gasification. Ash management poses operational challenges, with high ash content causing slagging and fouling, impacting overall system efficiency [124].

A pervasive challenge is tar formation, a complex mixture of organic compounds that can condense on equipment surfaces, causing operational issues. Effective tar removal or conversion processes are crucial for system integrity and efficiency. Incomplete gasification can result in char and other solid residues, diminishing overall process efficiency. Precise temperature control is critical as gasification reactions are highly temperature-dependent [121, 126]. A review by Ramos et al. [98] discussed the effect of different operation conditions on gasification performance, including temperature, fuel particle size, pressure, gasifying agent, and gasifier types.

Scaling up from laboratory to commercial-scale operations adds complexity. Factors like heat transfer, reactor design, and process optimization become intricate in larger systems. Syngas generated often requires cleanup processes to meet quality standards, removing impurities like particulates, tars, and contaminants. Operating gasification processes at elevated pressures for enhanced gas yields presents engineering challenges and requires robust, expensive

equipment [127]. Integrating gasification with downstream processes demands careful engineering and optimization for overall efficiency and economic viability [128, 129]. Economic challenges, encompassing both capital and operational expenses, are crucial for widespread adoption, especially amid competition from other renewable energy sources [130].

*2.3.2. Optimization Difficulties.* Efforts to optimize gasification face hurdles due to incomplete understanding of reaction mechanisms and limited data on specific feedstocks. Developing accurate models for predicting and controlling gasification behavior is challenging due to the process's complexity and diverse feedstocks, hindering the collection of reliable data, especially at the industrial scale [48, 131]. Environmental considerations, including emissions and ash disposal, add complexity. Compliance with regulatory standards requires careful monitoring and mitigation strategies. The lack of standardized testing protocols for different feedstocks and gasifier configurations further complicates system comparison and optimization [132].

Researchers address these challenges with innovative solutions. Machine learning algorithms analyze process data and identify optimal operating conditions. Data-driven modeling, combining experimental data with computational techniques, improves prediction accuracy. Multiobjective optimization strategies are developed to simultaneously optimize conflicting goals [71, 133]. Advanced gasifier designs and technoeconomic analysis contribute to enhancing the efficiency and sustainability of biomass and waste gasification technologies [134]. Despite challenges, the persistent focus on research and development signifies the potential of gasification as a transformative technology in the transition to a renewable energy future [135].

*2.3.3. Need for Advanced Modeling.* Advanced modeling in biomass and waste gasification involves sophisticated computational techniques to simulate and analyze complex thermochemical reactions. This encompasses various mathematical and computational models that simulate gasification systems under different conditions [125]. The primary goal is to gain a comprehensive understanding of interactions within the system for detailed predictions and optimizations [136]. These models consider factors like temperature, pressure, reaction time, and feedstock composition. Computational fluid dynamics (CFD) simulates fluid flow, heat transfer, and chemical reactions within gasifiers. This capability allows optimization of gasifier performance by predicting temperature profiles, reaction time, and species concentrations [137]. Kinetic modeling focuses on reaction rates during gasification, predicting pathways, and optimizing conditions [138].

Data-driven modeling, leveraging machine learning and statistical techniques, analyzes extensive datasets to enhance predictive capabilities and offer insights into optimal operating conditions [75, 139, 140]. These models complement traditional approaches like CFD and kinetic modeling, providing a holistic understanding of biomass and waste gasification processes. Process optimization is a primary application, fine-tuning operating parameters for maximum efficiency and identifying optimal conditions for desired outputs like syngas while minimizing undesired by-products. Additionally, they contribute to efficiency improvement by highlighting factors influencing gasification processes [14, 71, 141], aiding engineers in optimizing conditions, enhancing resource efficiency, and reducing the environmental footprint.

In resource utilization, advanced modeling assesses biomass and waste feedstocks' suitability, guiding researchers in identifying optimal compositions for specific gasification systems. Models like CFD and kinetic modeling play pivotal roles in designing and scaling up gasification systems, predicting larger-scale performance based on lab-scale experiments, and minimizing risks associated with deploying new technologies [51]. A detailed discussion on different modeling approaches for the gasification process has been reported by Ramos et al. [142]. Moreover, advanced modeling supports emissions prediction and control, assisting in designing systems compliant with environmental regulations. Life cycle assessment, technoeconomic analysis, economic evaluation, cost-benefit analysis, and risk mitigation are additional benefits, showcasing the indispensable role of advanced modeling in advancing sustainable and economically viable gasification technologies [42, 65, 68, 143].

## 3. Ensemble Models in Biomass and Waste Gasification

Ensemble models, a powerful machine learning technique, combine predictions from multiple models to achieve superior overall performance and generalization [144]. By leveraging the diversity of different models, ensembles overcome weaknesses in individual models and amplify their strengths. This approach, prevalent in classification, regression, and anomaly detection, consistently demonstrates improved accuracy, robustness, and generalization, making ensembles a valuable tool for machine learning tasks [145]. However, it is crucial to note that ensembles come with certain trade-offs. One significant consideration is the challenge of interpretability. As ensembles aggregate predictions from multiple models, understanding the decision-making process becomes more complex. The combined effects of diverse models can make it challenging to interpret and explain the rationale behind specific predictions. This lack of interpretability may be a limitation in applications where understanding the model's reasoning is essential for user trust or regulatory compliance [146, 147].

Ensemble techniques, by aggregating predictions from diverse models, offer advantages such as enhanced accuracy, robustness against overfitting, and improved generalization to new data. Applicable to various machine learning models, successful ensembling relies on ensuring diversity among base models, with each contributing unique insights or specializing in specific aspects of the data. This powerful technique significantly contributes to the machine learning toolbox, providing stable and reliable predictions [148, 149]. Ensemble ML integrates multiple conventional ML models or base learners into a single predictive model with enhanced accuracy [59, 150]. However, it is not mandatory for an ensemble model to achieve better accuracy than the most successful base learners [151]. For the case of

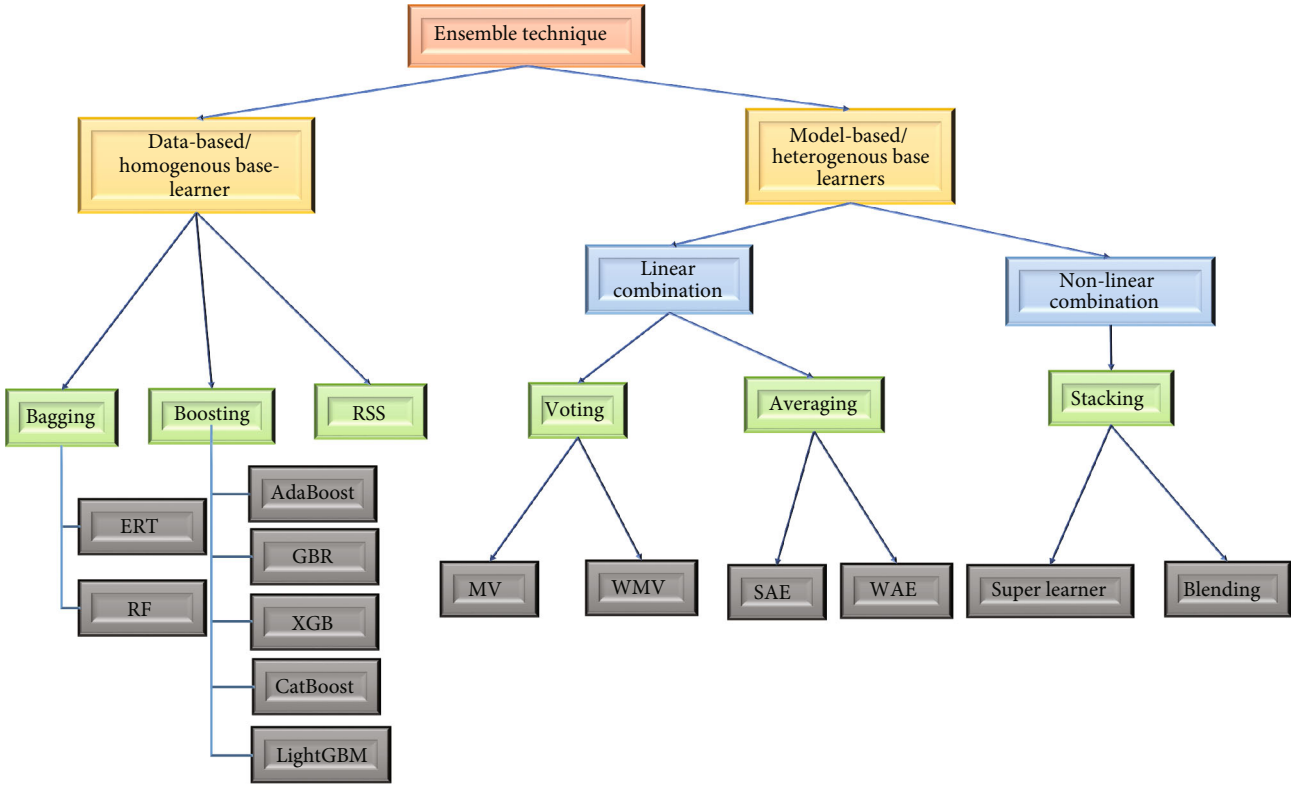


Figure 1: Classification of ensemble machine learning methods. Legend: MV: majority voting; SAE: simple averaging ensemble; WAE: weighted average ensemble; RSS: random subspace; WMV: weighted majority voting; RF: random forest; ERT: extremely randomized trees; GBR: gradient boosting regression.

regression problem, the accuracy of an ensemble model is typically an average of all base learners' performances [151]. In real-world problems, the same generalization performances may yield different predictions [152]. Therefore, combining the outputs from base learners can improve prediction accuracy, even when trained with small datasets (cross-validation can be used) [153, 154].

*3.1. Specific Ensemble Techniques.* Ensemble methods can be broadly classified into several categories, as depicted in Figure 1 [151, 155, 156]. These include bagging, boosting, voting, averaging, and stacking. Additional classifications exist, such as cascade generalization, cascading, delegating, arbitrating, and metadecision trees [157]; stochastic gradient boosting [158]; and Bayes Optimal Classifier [159]. Tree-based ensemble techniques like bagging and boosting are well-developed, but other approaches like voting, averaging, and stacking have also seen significant development. However, understanding and distinguishing between these techniques are essential [160]. To provide a comprehensive understanding of ensemble techniques, this section offers a detailed discussion on bagging, boosting, voting, averaging, and stacking. The summary of the advantages and limitations of each technique is presented in Table 2.

*3.1.1. Bagging Ensemble.* Bagging (bootstrap aggregating) leverages the bootstrap distribution to create subsets of data for training diverse base learners [155], with the ensemble's output aggregated through averaging (for regression) or voting (for classification) to mitigate variance errors [161]. Decision trees are often chosen as base models for bagging due to their capacity to capture intricate data interactions [161]. This method, exemplified by random forest (RF) and extremely randomized trees (ExtraTrees), is valued for its simplicity, effectiveness, and robustness (especially in scenarios where ensemble members exhibit different local minima) [59, 162].

Random forests (RFs) constitute an ensemble of tree predictors, where each tree's predictions rely on values sampled independently from a random vector with a uniform distribution across all trees in the forest [163]. A supervised learning meta-algorithm, RF, employs bootstrap aggregation to construct independent decision trees (DTs) [151, 164], combining the simplicity of DTs with enhanced flexibility for improved accuracy in both classification and regression problems [165, 166]. Hyperparameters such as the number of randomly selected predictors, splitting rules, and minimal node size can be optimized using techniques like $k$-fold cross-validation grid search or trial-and-error testing [165, 167].

Despite the proven success of RF in various domains, its exploration in gasification remains limited, with a handful of studies available [1]. Notably, Kardani et al. [150] achieved high accuracy in predicting gasification performance indicators using an RF model, while Ascher et al. [81] applied RF to predict multiple gasification characteristics with notable accuracy. These instances underscore the suitability of RF as a valuable tool for assessing and predicting gasification processes.

Table 2: Advantages and limitations of ensemble methods.

| Ensemble method | Advantages | Limitations |
|---|---|---|
| Bagging ensemble | (i) Reduces variance error<br>(ii) Works well when base learners have different local minima<br>(iii) Suitable for both regression and classification tasks | (i) May not significantly improve base model accuracy<br>(ii) Limited interpretability |
| Boosting ensemble | (i) Combines weak learners into a highly accurate model<br>(ii) Strong interpretability<br>(iii) Works well for both regression and classification | (i) Sensitive to noisy data<br>(ii) Computationally intensive<br>(iii) Require careful tuning of hyperparameters |
| Random subspace | (i) Decreases the likelihood of overfitting to specific features or noise in the data<br>(ii) Captures various aspects of the data, leading to better performance on unseen data<br>(iii) Each base learner focuses on different feature subsets, enhancing overall model diversity<br>(iv) Enables efficient exploration of the feature space, especially beneficial in high-dimensional datasets | (i) Interpreting the overall model becomes more complex due to diverse base learners<br>(ii) Managing multiple models and their predictions adds complexity to the modeling process<br>(iii) Noisy or irrelevant features in random subsets may degrade individual base learner performance<br>(iv) Training multiple models on different feature subsets can increase computational requirements |
| Voting ensemble | (i) Simple and easy to implement<br>(ii) Typically used for classification tasks<br>(iii) Reduce bias towards individual base learners | (i) Loss of information may occur in majority voting<br>(ii) Weighted majority voting may be challenging to implement |
| Averaging ensemble | (i) Provides an unweighted average of base learner outputs<br>(ii) Suitable for both regression and classification tasks<br>(iii) Weighted averaging can improve predictions | (i) Vulnerable to weaker base learners<br>(ii) Sensitive to overconfident base learners |
| Stacking ensemble | (i) Combines base learners in a hierarchical, nonlinear way<br>(ii) Maximizes prediction accuracy by leveraging the strengths of diverse models<br>(iii) Optimizes weights through cross-validation | (i) Requires a large number of base learners for optimal results<br>(ii) More complex to implement than other ensemble models<br>(iii) Overfitting can be a concern with smaller datasets |

*3.1.2. Boosting Ensemble.* Unlike bagging, boosting adopts a sequential approach to gradually build a collection of models with weak individual performance, progressively refining them into a highly accurate ensemble model [168, 169]. Boosting algorithms iteratively train and adjust the weights of numerous weak learners, producing a potent ensemble model. Examples of boosting algorithms encompass gradient boosting for regression (GBR), adaptive boosting (AdaBoost), gradient boosting machine (GBM), extreme gradient boosting (XGB), light gradient boosting machine (LightGBM), and categorical gradient boosting (CatGBM) [156, 170–172]. Gradient boosting models, particularly, have garnered extensive utilization in gasification processes, showcasing superiority over numerous other ensemble techniques [173]. Table 3 delineates the strengths and weaknesses of each boosting algorithm.

GBM and GBR serve as specific types of boosting technique for classification and regression problems, respectively [151]. GBR constructs a tree based on the errors (pseudoresiduals) made by the preceding tree, repeating this process until additional trees cease to enhance predictions [150]. GBR offers several advantages, including strong interpretability, robust generalization ability, fast convergence during training, and adaptability for datasets with multiple features [173]. Boosting ensemble models, particularly GBR and XGB, have been employed by various authors for gasification processes. For instance, Ascher et al. [81] and Wen et al. [174] utilized GBR to predict gasification performance metrics, achieving average prediction accuracies with $R^2$ scores of 0.9 and 0.87, respectively. Moreover, Ascher et al. [81] and Kardani et al. [150] applied XGB to model gasification processes, obtaining average prediction accuracies with $R^2$ scores of 0.84 and 0.95, respectively. AdaBoost was also employed by Ascher et al. [81] to predict 10 gasification characteristics, achieving an average prediction accuracy with an $R^2$ score of 0.85. While these examples demonstrate the effectiveness of boosting methods in gasification modeling, it is noteworthy that other boosting ensemble models like LightGBM and CatBoost have not yet been applied to gasification.

*3.1.3. Random Subspace Ensemble.* Random subspace method is an ensemble learning technique where multiple models are trained on different subsets of features randomly selected from the original feature set. Predictions from these models are then aggregated (usually by averaging) to improve overall model accuracy and robustness [175, 176]. This approach helps combat overfitting by reducing the influence of irrelevant or noisy features within the data. Unfortunately, there is no specific example of the random subspace algorithm being used directly in gasification modeling. Gasification models can be complex, involving numerous input features like feedstock properties, operating conditions, and desired products. The random subspace algorithm could be particularly advantages in such scenarios due to its ability to handle high dimensionality, improve

Table 3: Advantages and disadvantages of boosting ensemble algorithms.

| Boosting ensemble algorithms | Advantages | Disadvantages |
|---|---|---|
| Gradient boosting regression (GBR) | (i) Builds strong predictive models by iteratively improving weak models<br>(ii) Handles heterogeneous data types well<br>(iii) Provides flexibility in defining loss functions<br>(iv) Less prone to overfitting compared to AdaBoost | (i) Susceptible to overfitting if not properly tuned<br>(ii) More complex to implement compared to simpler algorithms<br>(iii) Requires careful tuning of hyperparameters |
| Adaptive boosting (AdaBoost) | (i) Can combine weak learners into a strong learner<br>(ii) Less susceptible to overfitting<br>(iii) Works well with a variety of base classifiers<br>(iv) Handles high-dimensional data effectively | (i) Sensitive to noisy data and outliers<br>(ii) Training can be time-consuming<br>(iii) May not perform well with complex datasets |
| Extreme gradient booting (XGB) | (i) Fast and scalable implementation<br>(ii) Handles missing data well<br>(iii) Provides regularization to prevent overfitting<br>(iv) Often achieves state-of-the-art performance | (i) Requires careful tuning of hyperparameters<br>(ii) Can be memory-intensive for large datasets<br>(iii) Prone to overfitting if not properly tuned |
| Light gradient boosting (LightGBM) | (i) Extremely fast training speed, making it suitable for large datasets<br>(ii) Efficient handling of high-dimensional data<br>(iii) Reduced memory usage compared to other gradient-boosting implementations<br>(iv) Supports categorical features without requiring one-hot encoding | (i) Prone to overfitting if not properly tuned<br>(ii) Requires careful tuning of hyperparameters<br>(iii) Less interpretability compared to simpler models |
| Categorical gradient boosting (CatBoost) | (i) Automatically handles categorical features without preprocessing<br>(ii) Robust to overfitting, thanks to built-in regularization techniques<br>(iii) Handles missing data internally<br>(iv) Provides strong performance with minimal hyperparameter tuning | (i) Slower training compared to some other gradient-boosting implementations<br>(ii) May require more memory for large datasets<br>(iii) Limited interpretability due to its complex nature |

prediction accuracy, and reduce overfitting. In general, the random subspace algorithm holds promise for gasification modeling, especially considering the complexity of these models. Further research could explore its effectiveness and potential benefits in this field.

*3.1.4. Voting Ensemble.* Voting is one of the simplest methods for aggregating prediction from multiple classifiers [177]. It involves deriving a final prediction by combining base learners [178]. Typically applied in two ways including major voting (MV) and weighted majority voting (WMV). Majority voting (MV) is the more prevalent approach for classification tasks [159, 179]. It employs several classification models to analyze each data point. The votes of these base learners are then counted, and the final label predictions are based on the most frequent vote [159, 180]. In contrast to MV, weighted majority voting (WMV) computes an unweighted average of the predicted class probability from base learners and selects the label with the highest average probability [180]. WMV introduces weights based on each classifier's performance on a validation set [181]. The final prediction considers the class with the highest weighted vote for each instance [182]. Unlike MV, where all models have equal weights, WMV assigns weights based on model performance [150]. However, determining appropriate voting weights for each class per classifier remains a challenge for WMV, limiting its effectiveness [183].

While MV is popular for classification tasks, it has not been applied in gasification, which is typically treated as a regression problem. However, WMV shows promise for gasification modeling. Kardani et al. [150] successfully applied WMV to merge the output of five regression-based ML models (decision tree regression (DTR), XGB, RF, multilayer perceptron (MLP), and SVR) to create an optimized ensemble model (OEM). This OEM achieved an impressive average prediction accuracy with an $R^2$ score of 0.98 for three gasification characteristics (LHV, LHVp, and gas yield), surpassing most base learners. This demonstrates the potential of WMV in constructing accurate models for gasification assessment, prompting further exploration of its suitability for regression problems.

*3.1.5. Averaging Ensemble.* Averaging, also referred to as voting regression, is an ensemble technique that combines predictions from multiple base learners through linear combination [184, 185]. There are two main averaging methods: simple averaging and weighted averaging. Simple averaging ensemble (SAE), also known as unweighted averaging or naïve averaging, is the most commonly averaging method, particularly for neural networks. SAE takes an unweighted average of the prediction from all the base learners and reports it as the final predicted score [159, 180, 185]. While similar to MV, SAE can be applied to both regression and classification problems [186]. A key

limitation of SAE is its susceptibility to the influence of weaker base learners and overconfident predictions [180].

Weighted average ensemble (WAE) extends SAE by assigning specific weights to each base learner's output based on their relative importance to the final prediction [186]. WAE typically offers improved prediction accuracy, calibration, and validation effectiveness compared to SAE [187]. Neither SAE nor WAE have been documented in predictive modeling for gasification processes. Further research is needed to explore their potential and suitability in this field.

*3.1.6. Stacking.* Stacking, also known as stacked generalization, is a unique ensemble learning framework categorized as a hierarchical or nonlinear combination [185, 188]. Unlike previously discussed techniques, stacking employs a metalearning model to combine base learners, aiming to maximize overall prediction accuracy [59, 189–191]. This method can leverage heterogeneous base learners (e.g., ANN, naive Bayesian, and logistic regression) to create potentially more powerful ensemble [192]. Stacking can be implemented with two or more layers [192, 193], and the metalearner can be any individual machine learning model, including generalized linear models [155]. There are two main approaches to implementing stacking: super learner and blending.

In the super learner approach, the metamodel is trained on the out-of-fold predictions made by base models using cross-validation [190, 191, 194, 195]. The super learner is advantageous for handling large numbers of base learners and selecting the optimal combination for a given dataset [196]. However, for very large datasets, this approach can become computationally expensive. Therefore, instead of optimizing the V-fold cross-validation [197], a single split cross-validation can be employed to optimize weights for an optimal combination [198]. Similar to the super learner, blending uses predictions from base learners to create a new dataset for the metamodel. However, blending utilizes a leave-out validation approach, setting aside a portion of the training data for validation [159, 191, 195]. While this makes the blending model simpler and faster to train, it may be prone to overfitting with smaller datasets [191].

Stacking ensemble models offer potential for gasification modeling, but limited research exists in this area. Ascher et al. [81] developed an ensemble using a super learner approach with six base learners (ANN, RF, GBR, XGB, AdaBoost, and SVM). However, details regarding the specific metalearning model used were not provided [82]. This ambiguity necessitates further studies to evaluate the efficacy of the super learner ensemble approach in gasification modeling and to determine if a true super learner or a simpler averaging method was actually implemented.

*3.2. Application of Machine Learning Models in Gasification.* Gasification performance is influenced by numerous factors, including feedstock properties and operational conditions [199–201]. Key performance indicators (KPIs) such as heating values, gas yield, CCE, gasification efficiency (GE), cold gas efficiency (CGE), and tar yield are used to assess the economic viability and quality of the gasification system [15, 202]. In this study, KPIs are categorized into four groups: gas composition, product yields and quality, heating values, and process efficiency. Machine learning (ML) has been applied to both regression and classification problems in gasification processes, with most research focusing on regression tasks [203]. Additionally, the majority of ML studies have examined single-feedstock gasification, with limited research on cogasification. To address this gap, training ML models with heterogeneous datasets encompassing various blend types and blending ratios is necessary. Table 4 provides examples of relevant studies on applying ML to gasification and cogasification processes.

In multiple input multiple output (MIMO) modeling, where there are multiple outputs (target variables), it is crucial to recognize that each target variable may have varying levels of prediction accuracy. Therefore, relying solely on the overall model accuracy can be misleading. Traditionally, researchers have often focused on the overall model performance, neglecting individual target variables. However, for effective application selection, assessing the prediction accuracy of each target variable is essential. This section analyzes the prediction accuracy of individual target variables to gain insights into the models' strengths and weaknesses for specific outputs. Despite variations in modeling approaches, comparisons between ML models are still possible [166, 215].

Artificial neural networks (ANNs) have been the most widely used technique for gasification modeling [23, 221, 222], as shown in Figure 2(a). Many other ML techniques have also been explored, including support vector SVR, DTR, linear regression (LR), polynomial regression (PR), Gaussian process regression (GPR), *K*-nearest neighbor (KNN), and ensemble methods (RF, GBR, XGB, AdaBoost, optimized ensemble model (OEM), and super learner). For a comprehensive analysis of these methods' advantages and limitations, refer to Umenweke et al. [79]. Heating value, including both lower heating value (LHV) and higher heating value (HHV), has been the primary target variable in gasification modeling studies (Figure 2(b)). Other relevant gasification performance metrics studied include gas composition (CO, $CO_2$, $H_2$, and $CH_4$), product yields and gas quality (total gas yield, $H_2$ yield, and tar yield), and process efficiency (CCE and CGE).

*3.2.1. Gas Composition Prediction.* Syngas composition is a critical indicator of gasification system performance [15, 223, 224]. The primary gas components in syngas typically include CO, $CO_2$, $H_2$, $CH_4$, and $H_2O$ [199, 225]. Additional components may include Tar ($C_{10}H_8$), $H_2S$, $N_2$, and unconverted carbon (char) [226]. In machine learning (ML) models, syngas composition can be modeled in two ways: multiple input single output (MISO) models (each gas species is predicted as a separate target variable) and multiple input multiple output (MIMO) models (all gas species are predicted collectively). Commonly predicted gases include CO, $CO_2$, $CH_4$, and $H_2$, with recent studies exploring $C_2H_4$ and $N_2$ prediction [81, 82].

Table 4: A summary of the literature on the machine learning modeling of gasification.

| Authors | Reactor type | Feedstock | Data collection methods | ML models | Data samples | Feature/input variables | Target/output variables | Key findings |
|---|---|---|---|---|---|---|---|---|
| Mutlu and Yucel [203] | Downdraft | Woody biomass | Experimentally using Arduino | Binary least squares SVM and multiclass RF classifiers | 5237 | TD ($T_0$, $T_1$, $T_2$, $T_3$, $T_4$, and $T_5$), ER, FR, MC, C, H, O, N, ash, and FC | GC ($CH_4$, $H_2$, CO, and $CO_2$) and HHV | $R^2$ exceed 96% for LS SVMR$^2$ and greater than 89% for the RF classifier |
| Elmaz et al. [204] | Downdraft | Woody biomass | Experimentally using Arduino | PR, SVR, DTR, and ANN-MLP | 4826 | TD ($T_0$, $T_1$, $T_2$, $T_3$, $T_4$, and $T_5$), ER, and FR | CO, $CO_2$, $CH_4$, $H_2$, and HHV | MLP and DTR outperformed other models with $R^2$ higher than 0.92 for $CH_4$, $H_2$, and HHV outputs, and $R^2$ exceed 0.85 for CO and $CO_2$ |
| Mutlu et al. [205] | Fluidized bed gasifier | High-ash coal and biomass | Literature and random resampling using a SNR value of 13Db | LR, PR, GPR,SVR, DTR, and ANN | 56 and 255 | T, CFR, BFR, FC, VM, ash, MC, AFR, SFR, C, H, O, and rate constant. | GY, CCE, HHV, and CGE | GPR outperformed other models for the both original and resampled datasets with $R^2$ exceeding 0.7 for all outputs |
| Ayub et al. [206] | Downdraft | 86 biomasses | Simulations using Aspen plus | ANN | 1032 | MC, VM, FC, ash, C, O, H, N, and S; required power (kW). | Temperature and AFR. | The best number of neurons was 60 with a test score achieved as an MSE of 1.497 and $R^2$ value of 0.9976. |
| Kartal and Özveren [207] | CFB | 56 biomasses | Simulation using Aspen plus | Deep learning model (MLP) | 1 million | C, H, O, T, FR, and SR | LHV | $R^2$ exceed 0.99 for all datasets |
| Serrano and Castelló [208] | Lab-scale BFB | Woody biomass | Literature | ANN (four different models) | 126 | C, H, O, MC, ash, T, ER, and TSN | Tar content | All ANN models have $R^2$ higher than 0.97 |
| Serrano et al. [209] | Lab-scale BFB | Woody biomass | Literature | ANN | 203 | C, H, O, MC, ash, ER, T, SFR, and BM. | GC ($CH_4$, $H_2$, CO, $CO_2$) and GY | $R^2$ exceed 0.94. |
| Ozonoh et al. [3] | Updraft, downdraft, fluidized bed, and entrained bed | Biomass, coal, and blends of biomass and coal. | Literature | ANN | 315 | C, H, N, S, O, ash, MC, VM, LHV, ER, and T | CGE, CCE, GY, GC ($CH_4$, $H_2$, CO, and $CO_2$), and LHV | $R^2$ greater than 0.7 achieved for all outputs |
| Yucel et al. [210] | Downdraft | Pinecone particles and wood pellet | Experimental | ANN and NARX neural network model | 3831 | TD ($T_0$-$T_5$), ER, AF, C, H, N, S, O, ash, MC, VM, and LHV | GC ($CH_4$, $H_2$, CO, and $CO_2$) and LHV | ANN models achieved $R^2$ greater than 0.99. |

Table 4: Continued.

| Authors | Reactor type | Feedstock | Data collection methods | ML models | Data samples | Feature/input variables | Target/output variables | Key findings |
|---|---|---|---|---|---|---|---|---|
| Shahbaz et al. [211] | Integrated fixed and circulating bed | Palm oil waste | Experimental | ANN | — | SFR, T, coal bottom ash, and CaO/biomass | GC ($CH_4$, $H_2$, CO, and $CO_2$) | $R^2$ exceeds 0.99) in almost all the cases |
| Wen et al. [174] | Updraft | Rice husks | Experimental | ANN and GBR | 74 | ER, T, SFR, C, H, N, S, O, ash, MC, VM, LHV, ER, and T | GC ($CH_4$, $H_2$, CO, and $CO_2$) | $R^2$ (0.80-0.89) for ANN model $R^2$ score (0.81-0.93) for GBR |
| Safarian et al. [212] | Downdraft | 50 various feedstocks | Simulation using Aspen | ANN | 1800 | M, VM, FC, ash, C, O, H, N, S, T, AFR, and SBR | SMFR of hydrogen | $R^2$ higher than 0.999 and RMSE less than 0.25. |
| Baruah et al. [213] | Downdraft | Various woody biomass | Experimental | ANN | 63 | C, H, O, ash, M, and T | GC ($CH_4$, $H_2$, CO, and $CO_2$) | $R^2$ exceed 0.99 for $CH_4$ and CO models and $R^2$ more than 0.98 for $CO_2$ and $H_2$ models. |
| Ozbas et al. [214] | Fixed bed | Olive pit | Experimental | KNN, LR, SVR, and DTR | 2036 | Time ($T$) | GC ($CH_4$, $H_2$, CO, and $CO_2$), HHV, and $H_2$ yield | In all models, $R^2$ exceed 0.9 for $H_2$ yield was obtained |
| Zhao et al. [215] | SCWG | Agricultural waste and municipal solid waste | Literature | GPR, ANN, SVR, and RF | 95 | C, H, O, and ash | $H_2$ yield | RF model outperformed GPR, ANN, and SVM models for predicting $H_2$ yield ($R^2$ value of 0.9782) |
| Kardani et al. [150] | Lab-scale fluidized bed gasifier. | Municipal solid waste | Literature | WMVE (OEM), XGB, RF, SVR, DTR, and MLP | 67 | MC, C, O, H, N, ash, S, ER, and T | LHV, LHVp, and GY | OEM performed better than others with a rank score of 36 |
| Ascher et al. [81] and Ascher et al. [83] | General gasifier types | Variety of feedstocks | Literature | GBR, XGB, AdaBoost, RF, ANN, SVR, and SL ensemble | 312 | C, H, S, ash, and MC, PS, T, gasifier operation mode, gasifier scale, ER, catalyst usage, gasifying agent, reactor type, and bed material. | GC ($CH_4$, $H_2$, CO, $CO_2$, $N_2$, and $C_2H_n$), LHV, Tar, GY, and char yield | GBR outperformed all model types with the highest performance for seven and eight out of ten outputs in terms of $R^2$ and RMSE |
| Li and Song [23] | Various gasifier types | Food waste, sludge, and mixed biomass waste | Literature | RF, SVR, ANN, and GBR | 342 | C, H, S, N, O, ash, SBR, ER, T | GC ($CH_4$, $H_2$, CO, $CO_2$), Tar, char, GY | GBR model showed good performance with test $R^2$ value range from 0.82 to 0.96 |
| Haq et al. [141] | SCWG | Sewage sludge | Literature | SVR, ANN, GPR, and ensembled tree | 125 | C, O, H, N, S, FC, VM, ash, T, pressure, and residence time | $H_2$ yield | GPR has the highest accuracy with $R^2$ value of 0.997. SVM has the lowest accuracy with $R^2$ value of 0.761 |
| Sison et al. [58] | CLG | Agricultural residues, energy crops, municipal waste, algae biomass, and poultry waste. | Aspen model | SVM, RF, and GBR | 236 | C, H, O, N, S, FC, VM, ash, T, OC, and SBR | $H_2$ yield and char yield | GBR has the highest $R^2$ score of 0.9453 for both $H_2$ and char yields. |

Table 4: Continued.

| Authors | Reactor type | Feedstock | Data collection methods | ML models | Data samples | Feature/input variables | Target/output variables | Key findings |
|---|---|---|---|---|---|---|---|---|
| Pandey et al. [216] | BFB | Variety of biomass feedstock | Literature | LR, GBR, LARS, RR, RF, BAG, and MLP | 558 | C, H, O, MC, ash, ER, T, BM, and SBR | GC ($H_2$, CO, $CO_2$, and $CH_4$) and GY | GBR has the highest prediction accuracy for all output with an average $R^2$ score of 0.847 |
| Aguado et al. [217] | Downdraft | exhausted olive pomace pellets | Experimental | MLP, LS-SVR, PR, ensemble (hybrid intelligent model), and *k*-means | 26,839 | CO, $CO_2$, $CH_4$, $C_nH_m$, $O_2$, and $N_2$ | LHV and $H_2$ | The hybrid intelligent model has prediction errors of only 0.134% |
| Kargbo et al. [135] | Two-stage gasifier | Waste wood | Literature | Ensemble (BANN) and ANN | 40 | DT, T, SBR, carrier gas flow rate, C, H, and O | GY, residue, tar, $H_2$, CO, $CO_2$, and $CH_4$ | BANN has the highest prediction accuracy of $R^2$ score of 0.999 the gas composition |
| Yang et al. [218] | | MSW | Literature | SVR, RF, and GBR | 317 | MSW properties and gasification parameters | GY, tar, char, GC, and LHV | GBR has the highest prediction accuracy with $R^2$ exceeding 0.926 |
| Devasahayam and Albijanic [57] | — | RSS and HDPE | Experimental | DTR, RF, GBR, and AdaBoost | 30 | T, RSS-PS, HDPE-PS, and HDPE % | $H_2$ yield | GBR provided the best score for both MSE and $R^2$ |
| Lahafdoozian et al. [40] | Water-gas shift reactor | Plastic (polystyrene and PE) | Simulation (Aspen model) | RF, ANN, GBR, CART, SVR, and KNN | 9000 | *T*, pressure | CGE, mass flow rate, and GC ($CH_4$, $H_2$, CO, and $CO_2$) | RF model demonstrates robust performance, achieving $R^2$ values exceeding 0.99. |
| Hasanzadeh and Azdast [219] | — | Polyethylene terephthalate | Experimental | PR | 121 | T and ER | CGE, $H_2$/CO, and LHV | $R^2$ exceeding 0.9 |
| Gil et al. [220] | Fluidized bed gasifier, | Biomass | Literature | GPR, XGB | — | T, SA, SR, SBR, C, H, O, ash, FC, VM, and HHV, | $H_2$, CO, $CO_2$ COMB, and $CH_4$, GY | Good predictive performance with $R^2$ values of 0.82 and 0.98 |

Note: PR: polynomial regression; SVM: support vector machine; SVR: support vector regression; KNN: *K*-nearest neighbor; DTR: decision tree regression; MLP: multilayer perceptron; GPR: Gaussian process regression; GBR: gradient boosting regression; XGB: extreme boosting regression; WMVE: weighted majority voting ensemble; VM: volatile matter; FC: fixed carbon; ER: equivalence ratio; FR: fuel flow/feed rate; T: temperature; AFR: air/fuel ratio; SFR: steam/fuel (biomass) ratio; SBR: steam biomass ratio; MC: moisture content; CFR: coal feed rate; BFR: biomass feed rate; COG: cogasification; PS: particle size; GC: gas composition; SR: steam feed rate; TSM: tar sampling method; BM: bed material; GY: gas yield; NARX: nonlinear autoregressive exogenous; SMFR: specific mass flow rate; TD: temperature distribution; CLG: chemical looping gasification; LR: multivariate linear regression; RR: ridge regression; LARS: least-angle regression; RF: random forest; BAG: bagging; BFB: bubbling fluidized bed; RS: rubber seed shell; HDPE: high density polyethylene; MSE: mean square error; COMB: total combustible gas; RSS: rubber seed shell; DT: devolatilization temperature; PE: polyethylene; SA: steam to air ratio; SR: stoichiometric ratio.

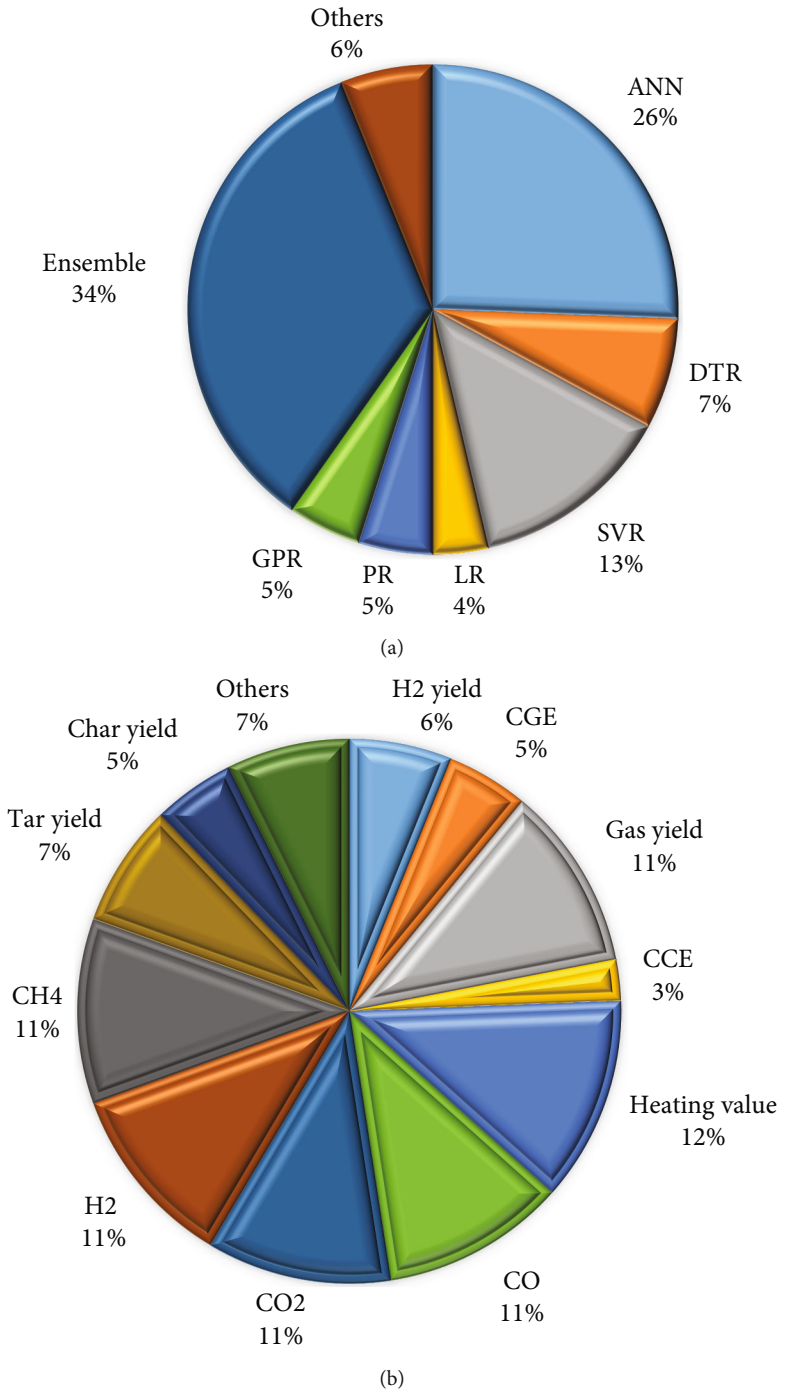


Figure 2: Comparative analysis of (a) machine learning methodologies and (b) gasification performance metrics within gasification research. The data used for constructing this figure can be accessed from the supplementary material (S1 and S2).

Studies have shown promising results using models like ANNs, support vector regression (SVR), decision tree regression (DTR), and gradient boosting regression (GBR) [82, 174, 204, 220]. These studies consistently demonstrated that ML models can effectively predict syngas composition, with $R^2$ values exceeding 0.7 for most gas components and even reaching as high as 0.92 in some cases (Supplementary material S3, S4, S5, and S6). Notably, ANNs consistently delivered good performance across various studies, highlighting their potential as a reliable tool for syngas composition prediction [82, 174, 204].

While some research focuses on comparing multiple models to identify the most effective approach [82, 204], others delve deeper into optimizing the performance of a single model, particularly ANNs [174]. Regardless of the chosen approach, ML offers a powerful and versatile tool for analyzing and predicting syngas composition. This allows for a deeper understanding of the intricate relationships between

process parameters and gas production, paving the way for optimizing gasification systems for improved efficiency and performance.

*3.2.2. Product Yields and Quality Prediction.* Optimizing gasification, especially when combining different feedstocks (Co-gasification), requires a deep understanding of its products [227]. This includes not just the total amount of gas produced (gas yield) but also the formation of tar, char, and the yields of specific gases like hydrogen. One crucial parameter is gas yield. This metric excludes inert gases like nitrogen and focuses on the usable gas per unit of dry feedstock, free of ash. It provides valuable insights into hydrogen-rich gas production, a key goal in many applications [200, 228]. Factors like steam addition and the speed of steam-related reactions significantly influence how much hydrogen yield ($m^3$/kg) increases as gasification progresses [229].

Machine learning (ML) offers a powerful tool for predicting gas yield (Supplementary material S7). Mutlu et al. [205] developed several ML models (SVR, DTR, PR, ANN, LR, and PLR) that all excelled at prediction, achieving $R^2$ values consistently above 0.85. Notably, GPR and SVR were particularly impressive with $R^2$ values of 0.9, followed by ANN (0.92), LR (0.91), PR (0.89), and DTR (0.88). Kardani et al. [150] built on this research by applying different ML models (SVR, DTR, ANN, XGB, RF, and OEM) to predict gas yield in a fluidized bed reactor. All models surpassed $R^2$ values of 0.85, with XGB leading the pack at 0.996, followed by OEM (0.99), RF (0.979), DTR (0.975), ANN (0.93), and SVR (0.886).

While research on overall gas yield is extensive, studies specifically focused on hydrogen yield ($H_2$) are less common. Ozbas et al. [214] proposed models (SVR, DTR, KNN, and LR) for $H_2$ yield prediction in a fixed bed reactor, all achieving $R^2$ values above 0.9. Similarly, Zhao et al. [215] applied different models (SVR, ANN, GPR, and RF) to an SCWG system, with all models exceeding $R^2$ values of 0.9. It is important to note that the limited input features used in these models by Ozbas et al. [214] and Zhao et al. [215] may restrict their applicability to broader scenarios. A recent study by Devasahayam and Albijanic [57] utilized various ML models (DTR, Adaboost, GBR, and RF) to predict hydrogen production from the cogasification of biomass and plastics. The achieved $R^2$ values were 0.76, 0.92, 0.99, and 0.68, respectively.

Another important gasification byproduct is tar, which consists of undesirable hydrocarbons, particularly aromatic compounds. Minimizing tar formation is crucial for achieving the desired gas quality [230]. It is important to remember that tar is a complex substance with five distinct classes. Focusing solely on total tar content or yield can be misleading [229]. Various factors, including feedstock composition and operating conditions, significantly influence tar production and its composition during biomass gasification [208, 231]. Tar yield, a key metric, is calculated by multiplying the total tar concentration in the gas product by the gas yield per unit mass of dry ash-free feedstock [208].

Despite its significance, predicting tar yield using ML models has received less attention, with only a few studies available [208]. Ascher et al. [81] addressed this gap by developing an ANN model for predicting tar yield in universal gasification systems, achieving high accuracy with an $R^2$ value exceeding 0.9. This aligns with the findings of Serrano and Castelló [208], who developed ANN models for lab-scale gasification, with the best model achieving an $R^2$ value of 0.977.

In contrast, Ascher et al. [83] conducted a broader study involving seven ML models (SVR, ANN, RF, GBR, XGB, AdaBoost, and super learner) for predicting tar yield in universal gasification systems. Most models exhibited lower accuracy in this broader study. SVR had the lowest $R^2$ value of 0.09. The XGB model outperformed others, achieving an $R^2$ value of 0.85, followed by GBR (0.76), super learner (0.75), ANN (0.62), and RF (0.59). Notably, the ANN model's performance in this study by Ascher et al. [82] was significantly lower compared to the results from their previous study and the study by Serrano and Castelló [208]. This highlights the potential influence of factors like the specific dataset and model configuration on ML prediction accuracy.

By providing accurate predictions of gas yield, hydrogen yield, and tar formation, ML models can significantly contribute to optimizing gasification processes for various applications. As research continues to explore and refine these models, we can expect even greater advancements in understanding and controlling gasification dynamics.

*3.2.3. Heating Value Prediction.* Heating value, also known as energy value or calorific value, is a crucial metric in gasification. It quantifies the energy released during the combustion of a fuel per unit volume or mass (MJ/$Nm^3$ or MJ/kg) relative to liquid and gaseous water [150, 222]. There are two main categories for gas heating values: lower heating value (LHV) and higher heating value (HHV). The distinction lies in how they account for water vapor in the combustion products. HHV considers the energy released when water vapor condenses back to liquid at a standard state, while LHV assumes the water vapor remains a gas. LHV is generally preferred for gasification assessments and is often derived from HHV [201].

Machine learning (ML) models have proven adept at predicting both HHV and LHV in gasification processes (supplementary material S8). Elmaz et al. [204] used four ML models (SVR, DTR, PR, and ANN) to predict HHV in downdraft gasification. ANN achieved the highest $R^2$ value (0.931), followed by DTR (0.921), SVR (0.886), and PR (0.858). Mutlu et al. [205] employed six ML models (SVR, DTR, PR, ANN, LR, and GPR) for downdraft gasification HHV prediction. GPR performed best ($R^2 = 0.78$). However, their lower $R^2$ values might be due to a smaller dataset (56 samples) compared to Elmaz et al. [204] who used 4826 samples. Kardani et al. [150] developed six ML models (SVR, DTR, ANN, XGB, RF, and OEM) for LHV prediction in lab-scale fluidized bed gasification. All models achieved $R^2$ values greater than 0.9, with OEM performing the best ($R^2 = 0.973$).

Additionally, Ascher et al. [83] used seven ML models (SVR, ANN, XGB, GBR, RF, AdaBoost, and super learner) for predicting LHV in general-type gasification systems. All models performed well except for ANN ($R^2 = 0.88$). The prediction accuracy of SVR, XGB, and RF was similar between their study and Kardani et al. [150]. Notably, differences in ANN model $R^2$ scores might be due to factors like a larger number of features considered by Ascher et al. [82] and potential overfitting in Kardani et al. [150] due to their smaller dataset. Hasanzadeh and Azdast [219] developed a polynomial regression (PR) model for LHV prediction, achieving impressive results with $R^2$ values exceeding 0.99. This highlights the potential of polynomial functions for accurate heating value prediction in gasification.

In general, various ML models have demonstrated strong performance for both HHV and LHV prediction. While ANNs have been promising, other approaches like PR and SVR have also shown effectiveness. As shown in the comparison between Elmaz et al. [204] and Mutlu et al. [205], the size and quality of the dataset can significantly impact the performance of ML models. Larger and more diverse datasets are generally preferred for robust model development. The choice of ML model and its hyperparameter tuning can influence prediction accuracy. While various models have been successful, exploring different options and optimizing their configurations could potentially lead to even better results. Selecting and engineering relevant features as input for the models can significantly enhance their effectiveness.

*3.2.4. Process Efficiency Prediction.* Carbon conversion efficiency (CCE), cold gas efficiency (CGE), and gasification efficiency (GE) are crucial metrics for evaluating the effectiveness of gasification processes. They provide insights into how well the feedstock is converted into usable gaseous products. CCE is calculated by dividing the moles of carbon in the product gas by the total moles of carbon in the feedstock [15]. Higher CCE values indicate a more significant conversion of carbon into valuable gaseous products like CO, $CO_2$, and $CH_4$ [232]. CGE compares the heating value of the produced gas mixture to the heating value of the feedstock [233]. It focuses solely on the gases exiting the gasification section and reflects how efficiently the potential energy in the feedstock is converted into usable energy in the gas [201]. GE, though less common, measures the overall conversion of feedstock into gaseous products. It is calculated as the total mass or moles of gas produced per unit of feedstock [224, 232].

While CCE and CGE are essential for efficiency evaluation, relatively few studies have explored using machine learning (ML) models for their direct prediction. Ascher et al. [82] suggested that CCE and CGE can be estimated based on other predicted parameters. However, directly evaluating ML model performance for CCE and CGE prediction remains valuable due to potential variations in results. Mutlu et al. [205] evaluated six ML models (SVR, ANN, LR, GPR, DTR, and PR) for CCE and CGE prediction. GPR achieved the best performance for CGE ($R^2 = 0.71$), while SVR led for CCE ($R^2 = 0.79$). However, it is important to consider limitations like dataset size, input selection, and hyperparameter tuning, as these factors can significantly impact prediction accuracy. Ozonoh et al. [3] used an ANN model and achieved higher $R^2$ values (0.9 for CGE and 0.84 for CCE) compared to Mutlu et al. [205]. This highlights the potential of ANNs for CCE and CGE prediction, but further validation is needed. Hasanzadeh and Azdast [219] developed a polynomial regression (PR) model for predicting CGE, achieving exceptional results with $R^2$ values exceeding 99%. This suggests the potential effectiveness of polynomial functions for CCE and CGE prediction.

Compared to gas yield and heating value prediction, research on ML for CCE and CGE prediction is scarce. More studies are needed to explore the effectiveness of various models and identify the best approaches for different gasification scenarios. The accuracy of ML models heavily relies on the quality and size of the training data. Limited access to high-quality datasets on CCE and CGE can hinder model development and generalizability. While some ML models can achieve high prediction accuracy, understanding the reasons behind their predictions can be challenging. This limits interpretability and the ability to refine gasification processes based on model insights.

## 4. Interpretability Techniques in the Context of Gasification

In the context of gasification, interpretability techniques play a vital role in understanding and explaining the complex processes and outcomes associated with thermochemical conversion. Gasification transforms carbonaceous materials into syngas, a versatile fuel used for electricity generation, chemical production, and other applications. Comprehensive understanding and interpretation of gasification system mechanisms are crucial for optimizing efficiency and minimizing environmental impact. Interpretability techniques contribute significantly to the development of advanced and sustainable gasification technologies by offering insights into these complex processes, facilitating optimization efforts, and ensuring reliable system operation [143].

Interpretability, broadly defined, refers to the ability to understand and explain a model or system's outputs or decisions in a clear and human-understandable way. This concept is particularly critical in machine learning and artificial intelligence (AI) to ensure transparency and trust in these decision-making models [234]. Interpretable models, like linear regression, allow for straightforward understanding of the relationship between input features and output. However, complex models, such as deep neural networks and ensemble models, can pose challenges due to their intricate architectures. Balancing model complexity with interpretability is an ongoing effort. Researchers are actively developing methods to enhance interpretability, providing valuable insights into the features and patterns influencing predictions.

Interpretability analysis, also known as feature importance assessment, variable importance assessment, or relative feature importance analysis, focuses on extracting

knowledge about the relationships learned by a model or the patterns within the underlying data [235, 236]. In simpler terms, interpretability refers to a human's ability to understand why a machine learning model makes a particular prediction. Highly interpretable models are easy to understand, and their results can be readily predicted by a human. In contrast, models with low interpretability are challenging to grasp [61, 82, 235].

*4.1. Importance of Understanding Models.* In the dynamic field of biomass and waste gasification, understanding machine learning models is paramount for all stakeholders, including operators, researchers, and policymakers. Clear model interpretation is essential, as it unlocks actionable insights that aid in process optimization and facilitate effective decision-making. This is particularly crucial in domains where model-based decisions have significant consequences, such as healthcare, finance, and law [237].

Several key reasons highlight the importance of understanding models across various domains, including data science and machine learning [238]. Firstly, interpretability and explainability are vital. They empower individuals to comprehend and explain the results and predictions of models, especially in high-stakes decision-making. Understanding the underlying model is crucial for troubleshooting and debugging issues. Furthermore, knowledge of the parameters and their impact allows practitioners to optimize and fine-tune models for better performance. Additionally, comprehension of model applicability and limitations guides practitioners in selecting the most suitable approach for a given problem and avoiding inappropriate model applications.

The importance of understanding models extends beyond technical aspects. It plays a pivotal role in feature engineering. Knowledge of a model's characteristics helps guide the selection of the most relevant data points to improve model accuracy. Effective communication of results to nontechnical audiences is also facilitated by understanding the model. If one grasps the inner workings of the model, they can translate complex concepts into clear explanations.

Understanding models is crucial for addressing biases and ensuring responsible model usage. By understanding how the model arrives at its conclusions, researchers and practitioners can identify and mitigate potential biases in the data or algorithms. Finally, as models are not static entities, continuous improvement relies on a deep understanding of models to inform decisions about updates or replacements based on evolving data or requirements. In conclusion, understanding models goes beyond technical considerations. It plays a critical role in informed decision-making, model reliability, and ethical considerations across diverse applications [239].

*4.2. Overview of Interpretability Techniques.* Interpretability techniques are crucial for unlocking the secrets of gasification processes, which are essential for achieving sustainable energy production and effective waste management. Broadly categorized into model-agnostic and model-specific methods, these techniques offer a range of benefits. They empower stakeholders to optimize gasification conditions, diagnose issues that may arise, and make informed decisions throughout the design and operation stages. This ultimately contributes to enhanced efficiency and sustainability in gasification processes.

Model-agnostic methods focus on understanding the overall behavior of the model and can be applied to various gasification models. Examples include sensitivity analysis, which quantifies the impact of individual input variables on the model's output. Feature importance analysis highlights the most significant features influencing the model's predictions. Partial dependence plots, on the other hand, provide visual insights into the relationships between input variables and the model's output [240].

For deeper insights specific to gasification models, model-specific methods are available. Rule extraction techniques, for instance, reveal the decision rules within rule-based models. Feature interaction analysis explores how input variables interact and influence the model's output. Additionally, counterfactual explanations help illustrate the cause-and-effect relationships of altering specific inputs [60, 241].

These interpretability techniques go beyond technical considerations. By empowering stakeholders with a deeper understanding of the model's inner workings, they enable process optimization, fault diagnosis, and informed decision-making. This multifaceted approach to interpretability, as illustrated in Figure 3, caters to different aspects of model interpretation. The figure categorizes the methods based on various characteristics, such as intrinsic vs. post hoc (built-in vs. applied afterwards), model-specific vs. model-agnostic, premodel (in-model) vs. postmodel (application stage), and global vs. local interpretability (providing insights into the entire model or specific aspects). A detailed analysis of the advantages and limitations of these methods within the gasification context is provided in Table 5 [82, 242]. This comprehensive approach ensures a holistic understanding of the intricate dynamics of gasification systems, paving the way for advancements in sustainable energy production and waste management.

The use of interpretability techniques to understand complex machine learning (ML) models in gasification systems is a nascent field. Many researchers have not prioritized interpretability as part of model development [243]. However, a growing body of research is exploring interpretability methods for complex models like ensembles and artificial neural networks (ANNs). Analysis of existing literature reveals five methods gaining traction for feature importance assessment: permutation importance, Gini importance, SHAP (Shapley additive explanations), sensitivity analysis, and partial derivative plots. Sensitivity analysis, particularly methods designed for ANNs, has been the most widely used (Figure 4(a)). However, its application has primarily focused on interpreting ANNs (Figure 4(b)).

Integrating interpretability techniques with ensemble models holds significant promise. While ensembles often outperform other models, their complexity can hinder understanding of their decision-making processes. Interpretability methods can bridge this gap, improving transparency and trust in these models. This is crucial in domains where understanding the rationale behind predictions is essential.

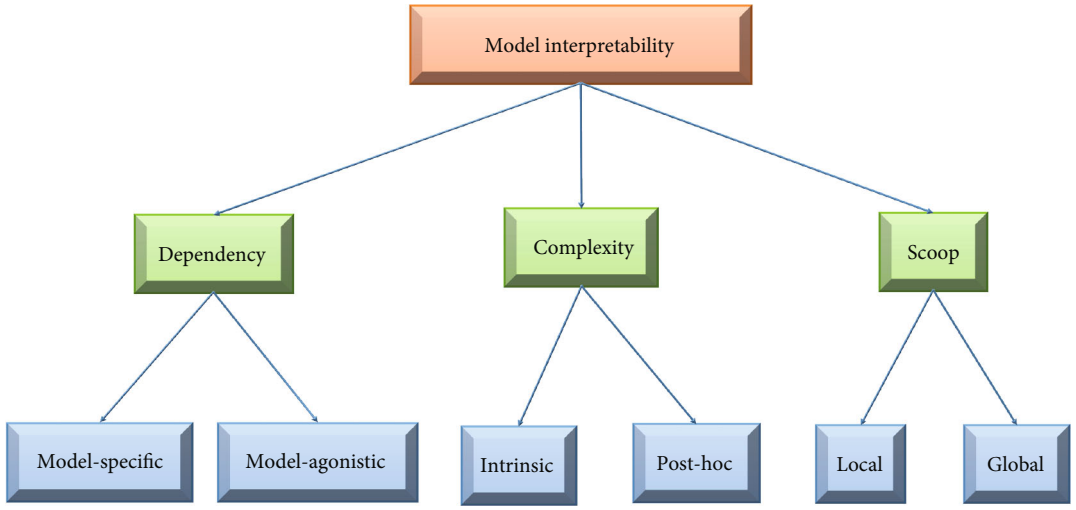


Figure 3: Classification of model interpretability analysis method.

Benefits include increased trust and accountability, improved model debugging, better feature selection, and more effective communication between developers and stakeholders [244].

However, challenges remain. These include developing efficient and scalable interpretability techniques, balancing accuracy with interpretability, standardizing methods, and adapting to new ensemble models [245]. Despite these challenges, the potential benefits of interpretability are substantial. Increased trust, more informed decision-making, and improved model performance all underscore the importance of further research in this area. Table 6 summarizes various studies involving model interpretability in gasification. It is worth noting some inconsistencies in the literature, with terms like "Gini importance" being misused as a type of sensitivity analysis [150]. To address this confusion, a critical evaluation of the five interpretability methods is provided elsewhere, offering a brief overview of their theory, strengths, weaknesses, and applications in feature importance analysis for gasification.

*4.2.1. Permutation Importance.* Permutation importance, also known by various terms including permutation feature importance [163, 246], permutation accuracy importance [164], or mean decrease accuracy [247], offers a valuable approach to understanding how features influence a model's predictions. It shares a similar algorithmic concept with random forests [163]. This method is specifically aimed at assessing the statistical significance of each feature's impact on the model's output [248]. It provides insight into feature importance by using the coefficients of linear models, which are often employed as interpretable models [249]. Permutation importance operates as a global interpretability technique. It works by modifying or permuting the order of input features. Subsequently, the impact of this manipulation on the model's performance is measured [215]. Essentially, this approach quantifies how much individual features contribute to the accuracy of a prediction model [250].

Within the context of permutation importance, a feature is considered crucial if permuting its values leads to a significant increase in prediction error. Therefore, the greater the increase in model error observed, the more important the feature is deemed to be [215]. Detailed explanations of interpreting feature importance results can be found in previous studies [251–253]. Permutation importance's primary strength lies in its ability to offer a global insight into the model's behavior and to automatically account for interaction effects among features [61, 254]. Additionally, it performs well with correlated variables and can be estimated using Equation (1) [240].

$$PFI_j\left(\hat{f}\right) = E\left(L\left(\hat{f}\left(X_{[j]}\right), Y\right)\right) - E\left(L\left(\hat{f}(X), Y\right)\right), \quad (1)$$

where $X_{[j]}$ is the $p$-dimensional random variable vector of features and the idea behind this method is to break the association between the $j$th feature and the target variable by permutating its feature values. If a feature is not useful for predicting an outcome, changing its values by permutation will not increase the expected loss.

Ascher et al. [83] compared the performance of permutation importance on four different tree-based ensemble models (RF, XGB, GBR, and AdaBoost). Their findings revealed that all models exhibited good agreement in ranking features, with temperature consistently emerging as the key feature. However, the overall importance of individual predictors could result from contributions of different submodels within the ensembles. The authors also noted slight variations between submodels for each output variable of each model type [82].

Another study by Zhao et al. [215] used permutation importance to interpret a random forest model for supercritical water gasification (SCWG) on $H_2$ yield. They separated the interpretation for biomass properties and SCWG process parameters. Their findings revealed that biomass concentration was the most influential feature among process parameters (temperature, residence time, and pressure). Hydrogen

Table 5: Advantages and limitations of interpretability methods.

| Technique | Dependency | Scoop | Advantages | Limitations |
|---|---|---|---|---|
| Permutation importance | Model-agnostic | Global | (i) Intuitive and highly compressed global interpretation<br>(ii) Considers interactions between features<br>(iii) Easy to implement for any model type | (i) Unclear whether training dataset or testing dataset can be used<br>(ii) Causes unrealistic data instances to be created during shuffling which limits interpretability |
| Partial dependence plot | Model-agnostic | Global | (i) Provides intuitive and clear interpretation<br>(ii) Very effective and straightforward for illustrating input/output relationship | (i) Makes assumption of independence between features<br>(ii) Becomes infeasible for MIMO models<br>(iii) Cannot capture confounding factors |
| SHAP | Model-agnostic | Local/ global | (i) Gives explanation based on strong game theory<br>(ii) Not only informs about the importance of features but also their relationship with the output<br>(iii) Predictions are fairly distributed among feature values | (i) High computational cost |
| Gini impurity | Model-specific | Global | (i) Provides a straightforward measure to interpret tree-based methods such as RF and GBR | (i) Not applicable to other model types<br>(ii) Biased towards inputs with more categories |
| Sensitivity analysis | Model-specific/ model-agnostic | Global/ local | (i) Effective and powerful algorithm to understand the stability of black box models | (i) High computational cost<br>(ii) Mainly used for ANN models |

Note: SHAP: Shapley additive explanations.

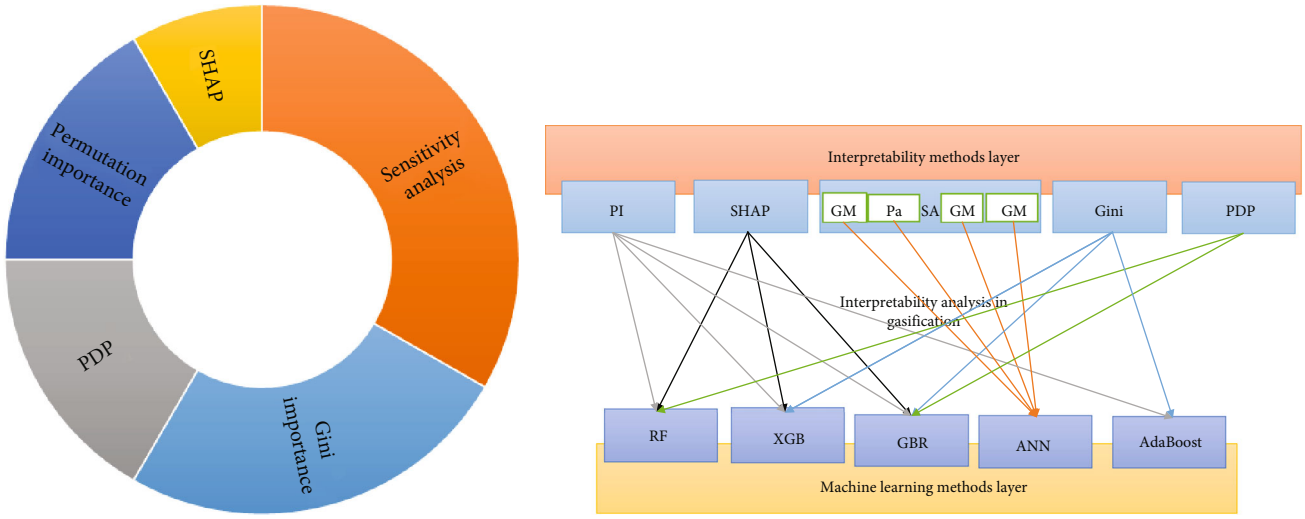


FIGURE 4: Interpretability methods in gasification: (a) extent of use among studies, (b) linking with ML models. Note. SHAP: Shapley additive explanations; PDP: partial dependence plot; GM: Garson method; SA: sensitivity analysis; PI: permutation importance; PaD: partial derivatives. Data used for creating (a) can be accessed in supplementary material S9.

and oxygen were identified as the most prominent features among the biomass properties considered (carbon and ash contents) [215]. It is important to note that SCWG is a novel gasification concept that differs from conventional gasification. As it utilizes feedstock with high moisture content (wet feedstock), process parameters like biomass concentration, temperature, and pressure play crucial roles in SCWG performance.

*4.2.2. Sensitivity Analysis.* Sensitivity analysis (SA), also known as uncertainty analysis, plays a crucial role in understanding complex models, often referred to as "black-box" models. SA assesses how the model's output reacts to changes in its input variables [3, 249]. This technique is particularly valuable for developing visual tools that help us inspect models with nonlinear or nonmonotonic behavior. By performing sensitivity analysis, we can gain valuable insights into the robustness and decision-making processes employed by machine learning models [255, 256]. There is a variety of SA methods available, categorized as local or global. Some common examples include one-at-a-time analysis, Morris methods, Sobol indices, perturb and profile methods, and techniques based on artificial neural networks [257–260].

When it comes to interpreting artificial neural network (ANN) models, two main categories of SA methods are commonly used: pure SA (like perturb and profile methods) and neuron network-based SA methods (like connection weights or Garson's algorithm) [261–263]. Within the neuron network-based methods, partial derivatives (Pad) and connection weight (Garson's algorithm) approaches have been applied to interpret ANN models [3]. Ozonoh et al. [3] argued that the Pad is a more reliable method compared to connection weights, step-wise methods, and perturb and profile methods. Using Pad to interpret their ANN model, they found that carbon was the most influential feature for all predicted outputs (CGE, gas yield, and LHV), followed by volatile matter and then temperature.

Several studies have employed SA techniques to understand how input features influence the outputs of ANN models. For instance, Serrano & Castelló [208] used the Garson method to analyze the influence of input features on tar yield in an ANN model. Their results revealed that the elemental composition (C, H, and O) of the biomass feedstock had a significant impact (over 45%) on tar yield, while the operating conditions (ER and temperature) of the gasification system contributed over 30%. Yucel et al. [210] applied the Garson method to interpret an ANN model with two different input groups. They found that temperature distributions were the most influential factors for all output variables in one case. However, another case showed different features (AF and AR) becoming more important for specific predictions. This highlights how the choice of model inputs can significantly impact the resulting gasification predictions. Baruah et al. [213] included reduction zone temperature alongside other variables as model inputs. The Garson method revealed that each input variable had a strong influence on the outputs (ranging from 8% to 31%), with reduction temperature being the most crucial factor for predicting CO and $H_2$. These case studies demonstrate how sensitivity analysis can be a powerful tool for understanding how different features influence the predictions of ANN models used in various applications.

*4.2.3. Partial Dependence Plot.* Partial dependence plots (PDPs) offer a valuable tool for visualizing and understanding the global relationship between a machine learning model's output and its individual input features within a reduced feature space [249]. However, it is important to be aware of their limitations. PDPs can produce misleading results when dealing with strongly correlated predictors. This is because they rely on extrapolating the model's

Table 6: A summary of literature on feature importance analysis for gasification.

| Authors | Method | Feedstock | ML model | Feature | Target | Key findings |
|---|---|---|---|---|---|---|
| Kardani et al. [150] | Gini importance or MDA | MSW | XGB | C, H, S, O, MC, Ash, ER, and T | LHV, LHVp, and gas yield | $T$ is the most influential for LHV, LHVp, and gas yield |
| Ascher et al. [81] | SHAP, Gini, and permutation importance | Biomass and waste | GBR, RF, XGB, and AdaBoost | 24 input variables including particle size, gasifying agents, moisture, reactor types, T, etc. | Gas compositions, tar, gas yield, LHV, and char yield | Feedstock's particle size and choice of gasifying agent are key for all the models, with particle size being the top for GBR, RF, and AdaBoost. |
| Zhao et al. [215] | Permutation importance and PDP | Agricultural waste and MSW | RF | C, H, O and ash, BC, T (P), and residence time ($t$) | $H_2$ yield | For biomass properties: H > O > C > ash order of importance; for process parameter: BC > T > P > t. Overall, BC is the key input for $H_2$ yield |
| Li and Song [23] | Gini importance and PDP | Food waste, sludge, and mixed biomass waste | GBR | C, H, S, N, O, ash, SBR, ER, and T | char, tar, gas, $H_2$, $CH_4$, $CO_2$, and CO | Ash, N, and temperature were the most important to $H_2$ yield |
| Haq et al. [141] | SHAP | Sewage sludge | GPR | H, N, O, C, MC, VM, Ash, T, P, and time | $H_2$ yield | Features im- pacts trends are T > MC > P > time |
| Sison et al. [58] | SHAP and PDP | Agricultural residues, energy crops, municipal waste, algae biomass, and poultry waste. | GBR | SBR, OC, C, H, N, S, O, T, FC, VM, Ash, and MC | $H_2$ yield, and char yield | T > SBR > OC, FC > C > Ash > VM > S > H > moisture (for H2 yield); ash > VM > S > C > H > FC (for char yield) |

Legend: ML: machine learning; T: temperature; P: pressure; GPR: process gaussian regression; GBR: gradient boosting regression; RF: random forest; ANN: artificial neural network; XGB: extreme gradient boosting; AdaBoost: adaptive boosting; MDA: mean decrease accuracy; MSW: municipal solid waste; LHVp: product gas lower heating value; LHV: lower heating value; C: carbon; H: hydrogen; N: nitrogen; S: sulphur; O: oxygen; PDP: partial dependence plots; ER: equivalence ratio; BC: biomass concentration; MC: moisture content; PaD: partial derivatives; $T_R$: reduction temperature; SBR: steam/biomass ratio; VM: volatile matter; TD ($T_0$-$T_5$): temperature distribution; SHAP: Shapley additive explanations; OC: oxygen carrier mass ratio to biomass.

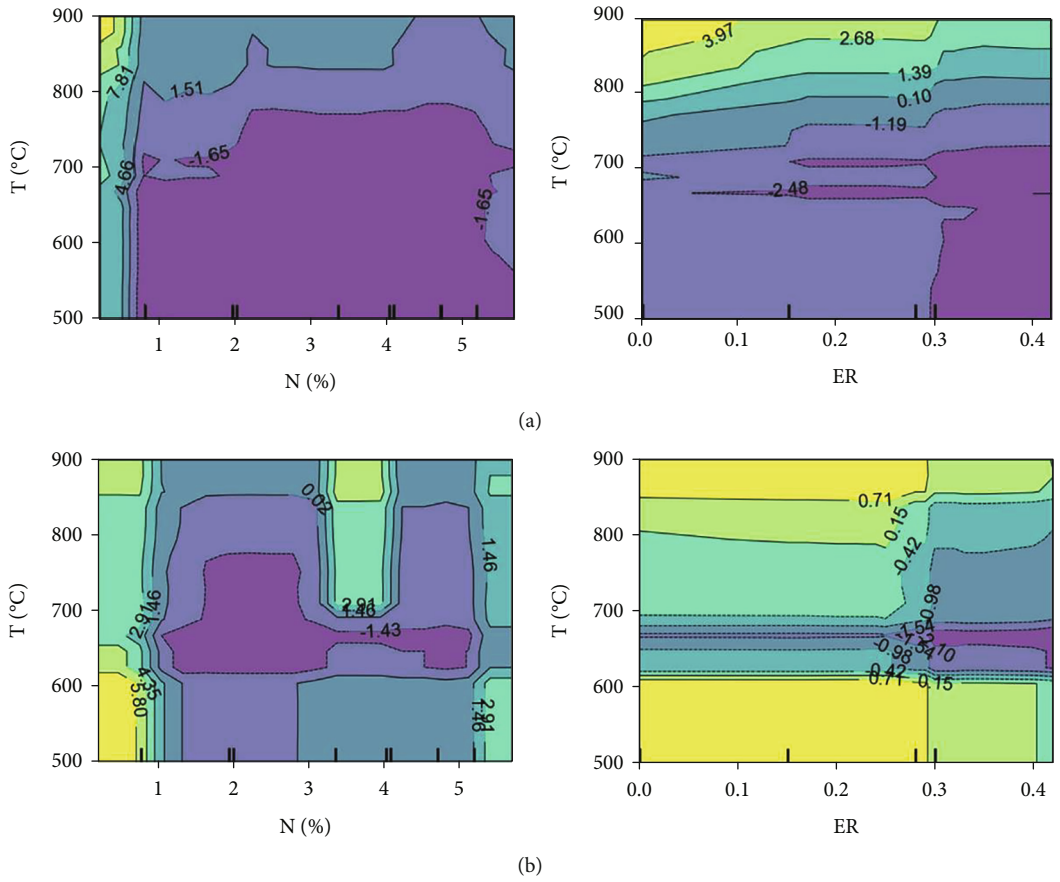


Figure 5: PDP plot on impacts of the interaction of temperature with other factors (a) on $H_2$ yield and (b) on $CO_2$ yield in syngas. Source: adapted from Li et al. [243]. ER: equivalent ratio; N: nitrogen atom. Copyright © Elsevier 2023.

response at predictor values that fall outside the range of values observed in the training data [264]. The calculation of PDPs can often be achieved using the following equation:

$$\widehat{f}_s(x_s) = \frac{1}{n}\sum_{i=1}^{n} f\left(x_s, x_c^{(i)}\right), \tag{2}$$

where $i$ is a generic observation, $x_s$ represent the features (s) of interest, $x_c^{(i)}$ are the other features included in the model, and $n$ is the number of instances which the dataset consists of.

Interpreting PDPs is relatively straightforward, as they share similarities with regression models. However, they have limitations. Firstly, PDPs might not be suitable for illustrating the input-output relationship in models with multiple inputs and multiple outputs (MIMO models). Secondly, PDPs may not capture the influence of confounding factors [82].

Despite these limitations, PDPs are often used alongside other interpretability methods. PDPs excel at visualizing how a specific input variable affects model predictions, while other methods, like permutation feature importance, help identify which input variables have the strongest influence on the model's output [253]. PDPs can be generated in both 1D and 2D formats. Two-dimensional PDPs are particularly useful for interpreting the interaction between two input variables and their combined impact on the predicted output [243].

There are examples of studies to showcase the strengths and weaknesses of PDPs as a tool for interpreting machine learning models. Zhao et al. [215] employed PDPs to illustrate the marginal effect of one or two features on the predictions made by a random forest (RF) model. They found that two-variable PDPs were effective in demonstrating how features like biomass concentration and temperature can have a synergistic impact on $H_2$ yield [215]. A more recent study by Li et al. [243] utilized PDPs to visualize the correlations between input variables and predicted targets using a gradient-boosting regression (GBR) model (Figure 5). PDPs were successful in illustrating the interactions between the most critical conditions (e.g., temperature) and the properties of the feedstock or other operational conditions on $H_2$ and $CO_2$ yield [243].

*4.2.4. Shapley Additive Explanations.* The Shapley additive explanations (SHAP) stand out as a powerful tool for interpreting machine learning models, offering a significant advancement over local interpretable model-agnostic explanations (LIME) [265]. Unlike LIME, which delves into local explanations for individual data points, SHAP focuses on the bigger picture. It is aimed at explaining the predictions of a

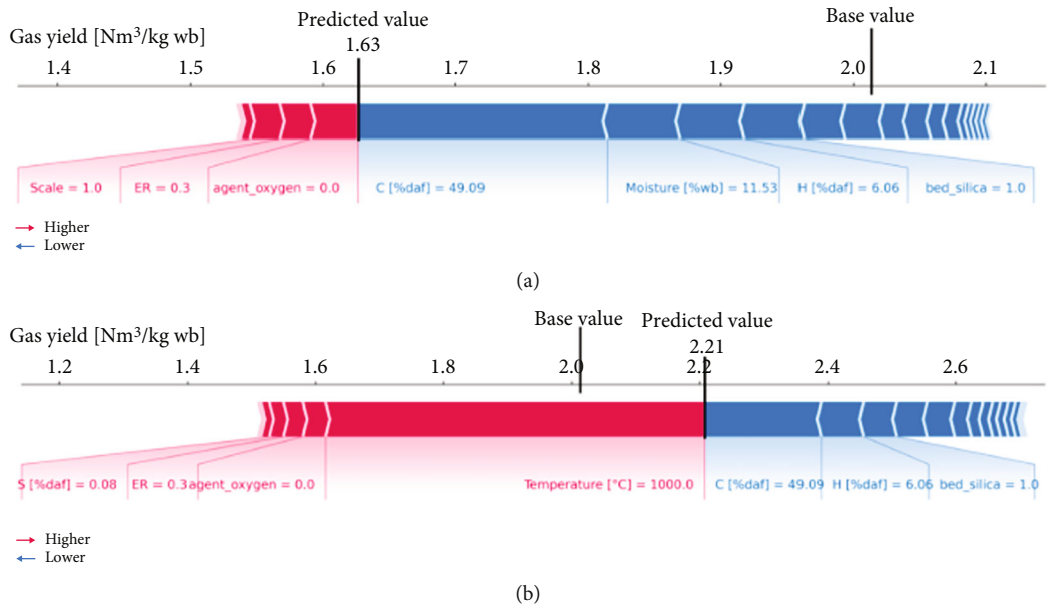


FIGURE 6: SHAP explanations of some selected individual predictions. The syngas yield from barley straw gasification is shown for (a) a base case at 800°C and (b) a high-temperature case at 1,000°C. Source: extracted from Ascher et al. [82]. Copyright © Elsevier 2022.

given data sample by quantifying the contribution of each feature to the model's overall prediction [266]. SHAP accomplishes this by leveraging coalitional game theory, a concept from economics. It calculates Shapley values, which essentially represent how much credit each feature deserves for the final prediction [249].

SHAP offers a significant advantage; it provides an intuitive visual representation of how various features influence a specific prediction. This visualization makes it easier to understand the reasoning behind the model's output. The Shapley value for a feature is calculated using a specific formula (Equation (3)), but the core concept is more important [267]. In simpler terms, the formula considers different combinations of features and calculates how much each feature contributes to the model's output within those combinations. The Shapley value represents the average of these contributions across all possible feature combinations. This approach provides a more comprehensive understanding of feature influence compared to methods that only look at individual features in isolation.

$$g\left(z'\right) = \varnothing_0 + \sum_{J=1}^{M} \varnothing_j z'_j, \tag{3}$$

where $g$ equals the model and $z'$ is the coalition vector. The $z'_j$ stands for a coalition vector where if a specific feature value is included, it is represented by 1 and 0 if it is not included. $M$ is the maximum size of the coalition, and $j$ is the feature attribution for a feature $j$, i.e., Shapley values. $\varnothing_0$ indicating a constant value when all inputs are missing. The generic $\varnothing_j$ is the weight against the feature contribution summation for the output of the model for overall feature combinations.

One of the key strengths of SHAP for global interpretation is that it goes beyond just feature importance. It also reveals the relationships between features and the model's output. Additionally, SHAP ensures that its feature attributions (contributions) are fairly distributed across all features, leading to more trustworthy results. These advantages have made SHAP a popular choice in various studies [61, 268]. SHAP can also be effectively combined with other interpretability methods, such as partial dependence plots (PDPs), to provide even more comprehensive interpretations [171].

For instance, Ascher et al. [83] applied SHAP to explain the predictions made by three different ensemble models across ten output variables. In the case of a model predicting gas yield, which used 24 input features, SHAP identified carbon, ER (equivalence ratio), and temperature as the key features. The authors also demonstrated how SHAP can be used to provide explanations for individual model predictions (Figure 6). They highlighted that such visualizations are valuable for communicating the prediction process of a machine learning model to stakeholders or policymakers [82].

*4.2.5. Gini Importance.* Gini importance, also known as impurity importance or mean decrease impurity, is a method introduced by Breiman [163] to assess the significance of variables in decision tree models [178, 243]. It works by quantifying the level of "impurity" during the process of building a decision tree. In simpler terms, impurity refers to how mixed the data samples are within a specific node of the tree. Features that are used more frequently to create separation points between these data samples are considered more significant by the model. Gini importance is typically used for classification problems, while regression tasks employ a similar approach with the sum of squares as the impurity measure [252, 269]. This can be estimated using Equation (4) [252].

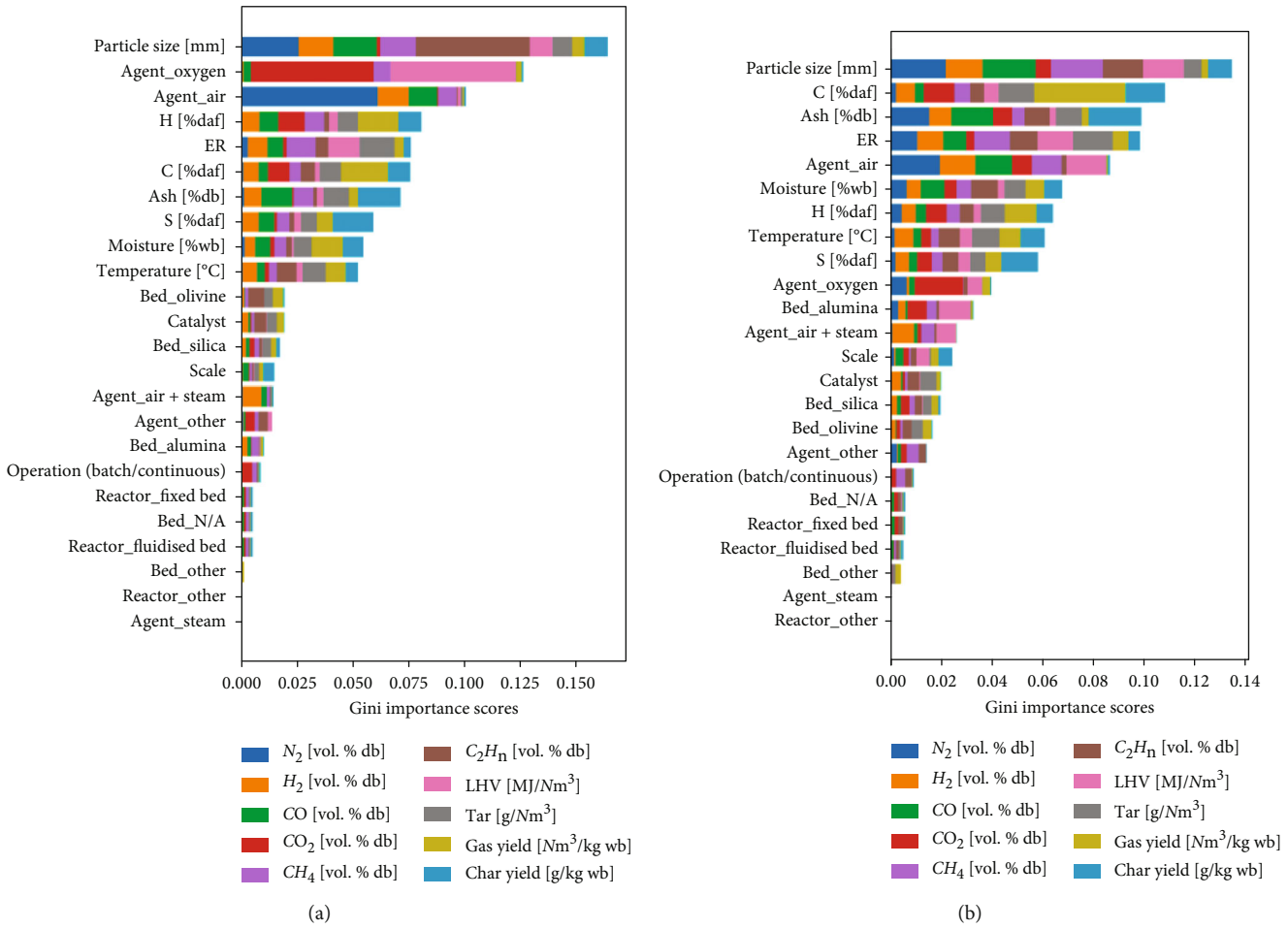


Figure 7: Gini feature importance of (a) gradient boosting models and (b) random forest models. Source: adapted from Ascher et al. [82]. Copyright © Elsevier 2022.

$$\widehat{\Gamma}(t) = \sum_{j=1}^{J} \widehat{\varphi}_j(t)\left(1 - \widehat{\varphi}_j(t)\right), \tag{4}$$

where $\widehat{\varphi}_j(t)$ is the class frequency for class $j$ in the node $t$.

While Gini importance offers a clear advantage in its simplicity and speed for interpreting tree-based models, it is not without limitations. One drawback is its tendency to favor categorical features with a high number of categories. Additionally, it can be biased towards features that are continuous and have little correlation with each other [163]. However, research suggests that Gini importance can sometimes identify important categorical features as well [82]. Another limitation is that this method is not applicable to other types of models, limiting its versatility [270]. Interestingly, Gini importance can yield similar interpretations of feature importance for GBR and RF ensemble models when compared to permutation importance [82].

Several studies have applied Gini importance to interpret machine learning models used in biomass gasification. For instance, Kardani et al. [150] used Gini importance to analyze an XGB model predicting target variables like lower heating value (LHV) and gas yield. Their findings revealed that temperature had the most significant influence on the outputs, with importance scores ranging from 35% to 40.8%. Ascher et al. [83] compared the performance of Gini importance on different ensemble models, including GBR, RF, XGB, and AdaBoost. They found that the particle size of the feedstock and the choice of gasifying agent were consistently important predictors across all models (Figure 7). Interestingly, they also observed that even when using the same interpretability method (Gini importance), feature importance can vary among different ensemble models. This highlights the importance of considering the specific model type when interpreting feature importance.

In a separate study, Li and Song [23] conducted an interpretation of a GBR model applied to gas composition, tar, char, and gas yields. Their analysis revealed that ash content, nitrogen content, and temperature were the most influential features when compared to $H_2$ yield in relation to the other input features, including hydrogen content, sulfur content, oxygen content, steam-to-biomass ratio, and equivalence ratio. Interestingly, temperature emerged as a key feature in both studies by Li and Song [23] and Kardani et al. [150]. However, it is worth noting that Kardani et al. [150] included proximate analysis of the feedstock, such as moisture content, as an input feature to their XGB model, which

was not considered in Li et al.'s study. Both studies employed GBR and XGB models, which belong to the boosting ensemble method, allowing for a meaningful comparison of their interpretations.

## 5. Challenges and Future Research Directions

Machine learning (ML) offers tremendous potential for optimizing and enhancing gasification processes. However, there are still challenges to address and exciting areas for further exploration.

*5.1. Challenges with Limited Datasets.* One major hurdle lies in the use of relatively small datasets, often collected from literature or generated through experiments. These limited datasets can restrict the prediction accuracy of ML models [203]. While there is no established minimum data size for optimal accuracy, it is generally understood that more data leads to better results. To overcome these limitations, researchers often employ resampling techniques like cross-validation during model development. However, limited exploration of alternative methods and the lack of consistent cross-validation implementation can hinder unbiased performance evaluations and scalability [205]. Therefore, focusing on improved feature selection, identification of crucial factors, and data preprocessing is vital for boosting the accuracy of ML models in gasification.

*5.2. Hyperparameter Optimization.* Hyperparameter optimization plays a key role in refining the performance of ML models. Traditionally, researchers have relied on the time-consuming trial-and-error method, particularly for tuning artificial neural network (ANN) models [271, 272]. Fortunately, alternative techniques are emerging to identify the most effective hyperparameter combinations. These include the grid search algorithm, particle swarm optimization, genetic algorithms, and Bayesian optimization methods [82, 141, 150, 215]. The impact of these optimization methods can vary depending on the ML model being used. Ensemble models (like random forest, XGBoost, gradient boosting regression, and AdaBoost) and tree-based methods (like decision tree regression) have shown less sensitivity to hyperparameter optimization [82]. However, a more comprehensive understanding of how these methods affect different models is still needed. Additionally, conducting comparative analyses of various optimization techniques would be valuable for advancing this field.

*5.3. Developing Generalizable ML Models.* While improving prediction accuracy and preventing overfitting have been major focuses, developing generalizable ML models that can handle complex real-world scenarios in gasification is equally important. Unfortunately, there is a limited number of studies exploring universal models capable of effectively handling high-dimensional or heterogeneous data, encompassing both continuous and categorical variables [273]. Furthermore, there is room for improvement in preprocessing categorical variables, with more exploration needed for alternative encoding methods. By emphasizing model generalizability, ML models can be tailored to address real-world challenges in gasification. For instance, cogasification introduces new factors like feedstock blend types and ratios. Predicting the synergistic behavior of these blends using ML models could be a valuable area of exploration. Depending on the specific feedstock combinations, different models might be required to effectively capture binary, ternary, and other combination types. Developing a generalized model that can adapt to new feedstock blends while considering their composition would be a significant advancement. Additionally, models capable of capturing novel gasification systems beyond conventional ones (e.g., two- and three-stage gasification/cogasification, chemical looping gasification/cogasification, catalytic gasification/cogasification) could significantly enhance generalizability and address real-world complexities.

*5.4. Interpreting Complex ML Models.* Most ML models are well-suited for harnessing heterogeneous data to create highly generalizable predictive models. However, interpreting models developed using high-dimensional and heterogeneous data remains a challenge, especially for those designed for predicting complex gasification systems. While various interpretable approaches exist, their suitability for specific models requires further investigation. For example, the SHAP method might not be compatible with AdaBoost models, even though it works well for certain ensemble models [82].

*5.5. Standardization and Frameworks.* Establishing frameworks for comparing different ML interpretation approaches and their corresponding results is crucial. This is because different interpretability methods used on the same model might reveal similar or distinct underlying logic and characteristics. Additionally, creating benchmarks for both existing and new interpretation methods specifically within the context of gasification is essential. High-quality benchmark datasets would be indispensable for achieving this. Exploring ways to combine the results obtained from multiple interpretability methods to provide a more comprehensive and coherent interpretation of the models is another valuable pursuit. Reproducibility is crucial in ML research. Standard practices and frameworks can ensure reproducibility across different studies.

*5.6. Deployment for Real-World Applications.* Finally, deploying ML models for real-world applications should be a core consideration during development [141]. This can be facilitated by creating graphical user interfaces (GUIs) and software applications that offer an interactive digital environment for predicting gasification performance based on user-provided input parameters, allowing engineers and technicians to leverage the power of ML for optimizing gasification processes in real time. This would bridge the gap between research and practical applications, accelerating the development and implementation of efficient and sustainable gasification technologies.

## 6. Limitations and Practical Implications

Despite these limitations, the review offers valuable insights with practical implications for researchers, practitioners, and policymakers working in the field of gasification.

Despite these limitations, the review offers valuable insights with practical implications for researchers, practitioners, and policymakers working in the field of gasification.

*6.1. Limitations of Current Studies.* The specific criteria used for selecting reviewed papers may have inadvertently excluded some relevant studies. Future reviews could benefit from expanding the search strategy or including a wider range of publication types (e.g., conference proceedings). Another limitation lies in the focus of this review. While ensemble models and interpretability methods are significant areas of research, other potentially relevant ML techniques (e.g., deep learning) or interpretability approaches might not have been covered in as much detail. Finally, the generalizability of the findings may be limited by the availability and quality of data used in the reviewed studies. Data scarcity and potential biases within datasets can affect the overall conclusions drawn about the performance of ML models in gasification.

*6.2. Practical Implications of This Review.* The emphasis on interpretable ensemble models highlights the need for researchers to not only achieve high prediction accuracy but also gain a deeper understanding of the factors influencing gasification processes. This knowledge can be crucial for optimizing these processes and ensuring their effectiveness. Standardized data collection and preprocessing protocols are essential for data quality and comparability. Establishing such protocols would facilitate collaboration among researchers and enable the development of more generalizable ML models. Developing user-friendly interfaces and software applications for the deployment of ML models is critical to bridge the gap between research and practical applications. Making these models accessible to engineers and technicians working on real-world gasification systems will accelerate the adoption and impact of this technology. The exploration of promising research directions, such as multimodal learning, physics-informed ML, and explainable AI for sustainability, provides a roadmap for future research efforts. By pursuing these directions, researchers can unlock the full potential of ML for advancing gasification technologies.

## 7. Conclusions and Recommendations

This comprehensive review delves into critical discussions surrounding emerging concepts in gasification, particularly focusing on ensemble machine learning (ML) modeling and interpretability analysis. Drawing insights from a diverse range of research findings, the study sheds light on the landscape of ML models employed in gasification. The review explores the use of both conventional ML models (support vector regression, Gaussian process regression, artificial neural networks, polynomial regression, linear regression, and decision tree regression) and ensemble ML models (random forest, gradient boosting regression, extreme gradient boosting, AdaBoost, super learners, and weighted majority voting). Interestingly, ensemble models were the most commonly employed model in roughly 34% of reviewed papers, followed by ANN models, representing 26% of the reviewed literature. The study also explores commonly applied interpretability methods like sensitivity analysis (Garson algorithm and partial derivative methods), Gini importance, SHAP (Shapley additive explanations), partial dependence plots, and permutation importance. This highlights the growing importance of understanding how these ML models arrive at their predictions.

The review emphasizes the substantial promise of ensemble ML models and their associated interpretability analyses for applications in gasification systems modeling. Significant advancements in this area are anticipated in the coming years. However, it is important to acknowledge that there is no single "best" ML method or interpretability approach. The optimal choice depends on the specific dataset, research question, and learning objectives. Consistency in results across multiple interpretability techniques can help build trust in the interpretation. A crucial consideration highlighted in this review is the inherent limitation of interpretability methods. The insights gleaned are ultimately limited by the content, quality, and quantity of the data used to construct the models. Therefore, careful selection of training data and features is essential to prevent the introduction of technical or scientific artifacts into both the models and their interpretations. This study offers significant contributions to the field of gasification research by highlighting the role of ensemble machine-learning modeling and model interpretability analysis. By examining a diverse array of ML models and interpretability methods, we provide insights into their applications and potential advancements in gasification systems modeling.

This paper underscores the pivotal contributions of ensemble models and interpretability techniques to the realm of biomass and waste gasification. Through a precise exploration of challenges and successes, the study emphasizes the potential of these methodologies in achieving cleaner and more sustainable energy production. The integration of advanced modeling not only showcases achievements but also addresses critical challenges like overfitting and computational complexities, paving the way for strategic approaches to mitigate these issues. Beyond the core focus on ensemble models and interpretability, the future of ML in gasification holds exciting prospects for further exploration:

(i) Explore incorporating sensor data, operational data, and even visual data from the gasification process into the ML models. This could lead to more comprehensive and accurate models that capture the intricate relationships between various factors and overall gasification performance

(ii) Integrate physics-based models with data-driven ML models. This approach would leverage the strengths of both approaches—physics-based models could provide a foundation for the ML models, while the data-driven approach could learn from experimental data to improve accuracy and efficiency

(iii) Utilize real-time data from sensor networks to develop ML models for closed-loop control of gasification processes. This would allow for dynamic

adjustments to operating conditions to optimize performance metrics like yield, heating value, and minimization of harmful emissions

(iv) Develop explainable AI (XAI) techniques specifically tailored to gasification, enabling researchers to not only understand how ML models arrive at predictions but also assess their environmental impact. This would be crucial for ensuring the sustainability and environmental benefits of gasification technologies

## Abbreviations

*Symbols/Nomenclature*

%: Percentage
>: Greater than
C: Carbon atom
$CH_4$: Methane
CO: Carbonmonoxide
$CO_2$: Carbondioxide
H: Hydrogen atom
$H_2$: Hydrogen molecule
N: Nitrogen atom
$N_2$: Nitrogen molecule
O: Oxygen atom
$P$: Pressure
$R^2$: Coefficient of determination
S: Sulphur atom
$T$: Gasification temperature
$t$: Residence time
$T_R$: Reduction temperature.

*Acronyms*

AdaBoost: Adaptive boosting
AFR: Air/fuel ratio
ANN: Artificial neural network
BC: Biomass concentration
BFB: Bubbling fluidized bed
BFR: Biomass feed rate
BM: Bed material
CCE: Carbon conversion efficiency
CGE: Cold gas efficiency
CFB: Circulating fluidized bed
CFR: Coal feed rate
daf: Dry and ash-free
db: Dry basis
DFB: Dual fluidized bed
DTR: Decision tree regression
ER: Equivalence ratio
FC: Fixed carbon
FR: Fuel flow/feed rate
GBM: Gradient boosting machine
GBR: Gradient boosting regression
GC: Gas composition
GPR: Process Gaussian regression
GY: Gas yield
KNN: $K$-nearest neighbor regression
LHV: Lower heating value
LHVp: Product gas lower heating value
MAD: Mean absolute deviation
MC: Moisture content
MDA: Mean decrease accuracy
ML: Machine learning
MLP: Multilayer perceptron
MSE: Mean square error
MSW: Municipal solid waste
MV: Majority voting
NARX: Nonlinear autoregressive exogenous
OEM: Optimized ensemble model
PaD: Partial derivatives
PDP: Partial dependence plots
PR: Polynomial regression
PS: Particle size
RF: Random forest
RMSE: Root mean square error
RSP: Random subspace
SAE: Simple averaging ensemble
SBR: Steam biomass ratio
SFR: Steam/fuel ratio
SHAP: Shapley additive explanations
SMFR: Specific mass flow rate
SR: Steam feed rate
SVM: Support vector machine
SVR: Support vector regression
TD($T_0$-$T_5$): Temperature distribution
TSM: Tar sampling method
VM: Volatile matter
WAE: Weighted average ensemble
WMV: Weighted majority voting
XGB: Extreme gradient boosting.

## Data Availability

We have shared additional data as a supplementary material.

## Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Authors' Contributions

Ocident Bongomin was responsible for the conceptualization (lead), writing—original draft preparation, writing—reviewing and editing, visualization, validation (equal), and formal analysis. Charles Nzila was responsible for the conceptualization (supporting), supervision (equal), resources (equal), and funding acquisition (equal). Josphat Igadwa Mwasiagi was responsible for the supervision (equal), resources (equal), validation (equal), and funding acquisition (equal). Obadiah Maube was responsible for the supervision (equal), validation (equal), and resources (equal).

## Acknowledgments

The authors gratefully acknowledge Africa Centre of Excellence II in Phytochemical, Textiles and Renewable Energy (ACE II-PTRE) at Moi University, for providing financial and technical support for this work and for awarding PhD scholarship to Mr. Ocident Bongomin.

## Supplementary Materials

The supplementary data associated with this article is attached as a separate file. The supplementary material contains quantitative data ($R^2$ score and frequency of appearance in previous studies) for comparison: (S1) gasification performance measures/target variables predicted in different studies, (S2) machine learning models used in gasification modeling, (S3) application of machine learning models for CO prediction, (S4) application of machine learning models for $CO_2$ Prediction, (S5) application of machine learning models for $CH_4$ prediction, (S6) application of machine learning models for $H_2$ prediction, (S7) application of machine learning for heating values (HHV/LHV) prediction, (S8) application of machine learning for hydrogen yield prediction, and (S9) interpretability methods used in gasification/cogasification modeling. *(Supplementary Materials)*

## References

[1] Y. Fang, L. Ma, Z. Yao, W. Li, and S. You, "Process optimization of biomass gasification with a Monte Carlo approach and random forest algorithm," *Energy Conversion and Management*, vol. 264, article 115734, 2022.

[2] T. M. Ismail, M. Shi, J. Xu, X. Chen, F. Wang, and M. A. el-Salam, "Assessment of coal gasification in a pressurized fixed bed gasifier using an ASPEN plus and Euler–Euler model," *International Journal of Coal Science & Technology*, vol. 7, no. 3, pp. 516–535, 2020.

[3] M. Ozonoh, B. O. Oboirien, A. Higginson, and M. O. Daramola, "Performance evaluation of gasification system efficiency using artificial neural network," *Renewable Energy*, vol. 145, pp. 2253–2270, 2020.

[4] M. Wang, Y. Wan, Q. Guo et al., "Brief review on petroleum coke and biomass/coal co-gasification: syngas production, reactivity characteristics, and synergy behavior," *Fuel*, vol. 304, article 121517, 2021.

[5] A. AlNouss, G. McKay, and T. Al-Ansari, "Production of syngas via gasification using optimum blends of biomass," *Journal of Cleaner Production*, vol. 242, article 118499, 2020.

[6] D. Baruah and D. C. Baruah, "Decision support system based planning of biomass gasification system for decentralised energy generation," *Renewable Energy Focus*, vol. 38, pp. 22–35, 2021.

[7] C. Ciliberti, A. Biundo, R. Albergo et al., "Syngas derived from lignocellulosic biomass gasification as an alternative resource for innovative bioprocesses," *Processes*, vol. 8, no. 12, pp. 1–38, 2020.

[8] M. Inayat and S. A. Sulaiman, "Effect of blending ratio on quality of producer gas from co-gasification of wood and coconut residual," *MATEC Web of Conferences*, vol. 225, article 05005, 2018.

[9] Y. Yang, R. K. Liew, A. Muthaliar et al., "Gasification of refuse - derived fuel from municipal solid waste for energy production: a review," *Environmental Chemistry Letters*, vol. 19, no. 3, pp. 2127–2140, 2021.

[10] A. Kumar, "Experimental investigation of a dual stage ignition biomass downdraft gasifier for deriving the engine quality gas," *Ain Shams Engineering Journal*, vol. 14, no. 3, article 101912, 2023.

[11] S. Mariyam, M. Shahbaz, T. Al-Ansari, H. R. Mackey, and G. McKay, "A critical review on co-gasification and co-pyrolysis for gas production," *Renewable and Sustainable Energy Reviews*, vol. 161, article 112349, 2022.

[12] S. Pati, S. De, and R. Chowdhury, "Process modelling and thermodynamic performance optimization of mixedIndian lignocellulosicwaste co-gasification," *International Journal of Energy Research*, vol. 45, no. 12, pp. 17175–17188, 2021.

[13] H. A. Umar and S. A. Sulaiman, "Parametric optimisation through the use of Box-Behnken design in the co-gasification of oil palm trunk and frond for syngas production," *Fuel*, vol. 310, article 122234, 2022.

[14] A. Bahadar, R. Kanthasamy, H. H. Sait et al., "Elucidating the effect of process parameters on the production of hydrogen-rich syngas by biomass and coal co-gasification techniques: a multi-criteria modeling approach," *Chemosphere*, vol. 287, Part 1, article 132052, 2022.

[15] M. Shahbaz, T. Al-ansari, M. Inayat, S. A. Sulaiman, P. Parthasarathy, and G. Mckay, "A critical review on the influence of process parameters in catalytic co-gasification: current performance and challenges for a future prospectus," *Renewable and Sustainable Energy Reviews*, vol. 134, article 110382, 2020.

[16] P. Kumar, P. M. V. Subbarao, L. D. Kala, and V. K. Vijay, "Real-time performance assessment of open-top downdraft biomass gasifier system," *Cleaner Engineering and Technology*, vol. 7, article 100448, 2022.

[17] T. A. Lestander, F. Weiland, A. Grimm, M. Rudolfsson, and H. Wiinikka, "Gasification of pure and mixed feedstock components: effect on syngas composition and gasification efficiency," *Journal of Cleaner Production*, vol. 369, article 133330, 2022.

[18] I. Rosyadi, S. Suyitno, A. X. Ilyas, A. Faishal, A. Budiono, and M. Yusuf, "Producing hydrogen-rich syngas via microwave heating and co-gasification: a systematic review," *Biofuel Research Journal*, vol. 33, pp. 1573–1591, 2022.

[19] S. M. Santos, A. C. Assis, L. Gomes, C. Nobre, and P. Brito, "Waste gasification technologies: a brief overview," *Waste*, vol. 1, no. 1, pp. 140–165, 2023.

[20] Y. Zhu, J. Miao, Y. Wang et al., "Ultra-low carbon dioxide emissions for biomass gasification using air with nearly 100% CO2 capture and conversion," *Chemical Engineering Journal*, vol. 484, article 149778, 2024.

[21] J. Makwana, A. D. Dhass, P. V. Ramana, D. Sapariya, and D. Patel, "An analysis of waste/biomass gasification producing hydrogen-rich syngas: a review," *International Journal of Thermofluids*, vol. 20, article 100492, 2023.

[22] A. C. Cruz, M. E. Ramalho, and R. M. Pilão, "Catalytic co-gasification of glycerol/fat mixtures: experimental vs thermodynamic equilibrium results," *Biomass Conversion and Biorefinery.*, vol. 12, no. 10, pp. 4575–4583, 2022.

[23] W. Li and Y. Song, "Artificial neural network model of catalytic coal gasification in fixed bed," *Journal of the Energy Institute*, vol. 105, pp. 176–183, 2022.

[24] A. Abdalazeez, L. Tianle, Y. Cao, W. Wang, S. Abuelgasim, and C. Liu, "Syngas production from chemical looping gasification of rice husk-derived biochar using BaFe2O4 as an oxygen carrier," *Journal of the Energy Institute*, vol. 105, pp. 376–387, 2022.

[25] P. C. Kuo, Z. Sun, F. Özdemir, M. Aziz, and W. Wu, "CO2 utilization in chemical looping gasification and CO-gasification of lignocellulosic biomass components over iron-based oxygen carriers: thermogravimetric behavior, synergistic effect, and reduction characteristics," *Journal of Environmental Chemical Engineering*, vol. 11, no. 3, article 109971, 2023.

[26] X. Liu, J. Hu, S. Zhao, W. Wang, Q. Zhang, and X. Yan, "Chemical looping co-gasification of wheat straw and lignite with calcium-enhanced iron-based oxygen carrier for syngas production," *Fuel Processing Technology*, vol. 227, article 107108, 2022.

[27] J. M. dos Santos Junior, Í. A. M. Zelioli, and A. P. Mariano, "Hybrid modeling of machine learning and phenomenological model for predicting the biomass gasification process in supercritical water for hydrogen production," *Eng*, vol. 4, no. 2, pp. 1495–1515, 2023.

[28] K. Heeley, R. L. Orozco, L. E. Macaskie, J. Love, and B. Al-Duri, "Supercritical water gasification of microalgal biomass for hydrogen production-a review," *International Journal of Hydrogen Energy*, vol. 49, pp. 310–336, 2024.

[29] A. Sadeghi Sheshdeh, M. R. Sabour, F. Mohammadi, J. Hui, M. Birkved, and B. Khoshnevisan, "Early environmental sustainability guidance on supercritical water gasification technologies for sugarcane bagasse management," *Sustainable Production and Consumption*, vol. 44, pp. 312–329, 2024.

[30] A. A. Arpia, T.-B. Nguyen, W.-H. Chen, C.-D. Dong, and Y. S. Ok, "Microwave-assisted gasification of biomass for sustainable and energy-efficient biohydrogen and biosyngas production: a state-of-the-art review," *Chemosphere*, vol. 287, Part 1, article 132014, 2022.

[31] Y. Zhang, C. Ke, W. Fu, Y. Cui, M. A. Rehan, and B. Li, "Simulation of microwave-assisted gasification of biomass: a review," *Renewable Energy*, vol. 154, pp. 488–496, 2020.

[32] S. Li, T. Laukkanen, D. Jiang, V. Vuorinen, and M. Järvinen, "AspenPlus-based techno-economic analysis of solar-assisted sorption-enhanced gasification for hydrogen and chemicals recovery from polyethylene terephthalate waste," *Energy Conversion and Management*, vol. 306, article 118318, 2024.

[33] F. Rajaee, G. Guandalini, M. C. Romano, and J. Ritvanen, "Techno-economic evaluation of biomass-to-methanol production via circulating fluidized bed gasifier and solid oxide electrolysis cells: a comparative study," *Energy Conversion and Management*, vol. 301, article 118009, 2024.

[34] F. Li, Y. Li, J. Cao et al., "Combination of $CO_2$ electrochemical reduction and biomass gasification for producing methanol: a techno-economic assessment," *Energy Conversion and Management*, vol. 307, article 118380, 2024.

[35] D. Peng, X. Xiang, Z. Deng, X. Zhou, B. Wang, and C. He, "Study on emission factor and reduction potential of organic solid waste gasification process," *Case Studies in Thermal Engineering*, vol. 53, article 103978, 2024.

[36] J. Dutzi, I. K. Stoll, N. Boukis, and J. Sauer, "Screening of ten different plants in the process of supercritical water gasification," *Sustainable Chemistry for the Environment*, vol. 5, article 100062, 2024.

[37] E. Sanjaya and A. Abbas, "Plasma gasification as an alternative energy-from-waste (EFW) technology for the circular economy: an environmental review," *Resources, Conservation and Recycling*, vol. 189, article 106730, 2023.

[38] S. Ghorbani, K. Atashkari, and M. Borji, "Three-stage model-based evaluation of a downdraft biomass gasifier," *Renewable Energy*, vol. 194, pp. 734–745, 2022.

[39] Y. Gao, M. Wang, A. Raheem et al., "Syngas production from biomass gasification: influences of feedstock properties, reactor type, and reaction parameters," *ACS Omega*, vol. 8, no. 35, pp. 31620–31631, 2023.

[40] M. Lahafdoozian, H. Khoshkroudmansouri, S. H. Zein, and A. A. Jalil, "Hydrogen production from plastic waste: a comprehensive simulation and machine learning study," *International Journal of Hydrogen Energy*, vol. 59, pp. 465–479, 2024.

[41] H. H. Shah, M. Amin, A. Iqbal et al., "A review on gasification and pyrolysis of waste plastics," *Frontiers in Chemistry*, vol. 10, pp. 1–38, 2023.

[42] Y. Fang, X. Li, S. Ascher et al., "Life cycle assessment and cost benefit analysis of concentrated solar thermal gasification of biomass for continuous electricity generation," *Energy*, vol. 284, article 128709, 2023.

[43] I. P. Silva, R. M. A. Lima, G. F. Silva, D. S. Ruzene, and D. P. Silva, "Thermodynamic equilibrium model based on stoichiometric method for biomass gasification: a review of model modifications," *Renewable and Sustainable Energy Reviews*, vol. 114, article 109305, 2019.

[44] Y. Zhang, S. A. Takyi, and Q. Chen, "Coal conversion to SNG: thermochemical insights through kinetic-free modeling in entrained flow gasifier," *Case Studies in Thermal Engineering*, vol. 56, article 104266, 2024.

[45] W. D. S. Fernando and J. Naser, "Numerical modelling of biomass gasification with steam injection for catalytic bed materials," *Results in Engineering*, vol. 21, article 101653, 2024.

[46] B. Pandey, Y. K. Prajapati, and P. N. Sheth, "CFD analysis of biomass gasification using downdraft gasifier," *Materials Today: Proceedings*, vol. 44, pp. 4107–4111, 2021.

[47] L. Stanger, A. Bartik, M. Hammerschmid et al., "Model predictive control of a dual fluidized bed gasification plant," *Applied Energy*, vol. 361, article 122917, 2024.

[48] Y. Ayub, J. Zhou, J. Ren, T. Shi, W. Shen, and C. He, "High-dimensional model representation-based surrogate model for optimization and prediction of biomass gasification process," *International Journal of Energy Research*, vol. 2023, Article ID 7787947, 14 pages, 2023.

[49] K. Rabea, S. Michailos, K. J. Hughes, D. Ingham, and M. Pourkashanian, "Comprehensive process simulation of a biomass-based hydrogen production system through gasification within the BECCS concept in a commercial two-stage fixed bed gasifier," *Energy Conversion and Management*, vol. 298, article 117812, 2023.

[50] M. Kuhn and K. Johnson, "Feature Engineering and Selection: A Practical Approach for Predictive Models," in *Feature Engineering and Selection*, CRC Press, Taylor & Francis Group, 2020.

[51] A. Kushwah, T. R. Reina, and M. Short, "Modelling approaches for biomass gasifiers: a comprehensive overview," *Science of the Total Environment*, vol. 834, article 155243, 2022.

[52] J. Li, X. Yao, and K. Xu, "A comprehensive model integrating BP neural network and RSM for the prediction and

optimization of syngas quality," *Biomass and Bioenergy*, vol. 155, article 106278, 2021.

[53] O. Bongomin, E. O. Nganyi, M. R. Abswaidi, E. Hitiyise, and G. Tumusiime, "Sustainable and dynamic competitiveness towards technological leadership of industry 4.0: implications for east African community," *Journal of Engineering*, vol. 2020, Article ID 8545281, pp. 1–22, 2020.

[54] O. Bongomin, G. G. Ocen, E. O. Nganyi et al., "Exponential disruptive technologies and the required skills of industry 4.0," *Journal of Engineering*, vol. 2020, Article ID 4280156, 17 pages, 2020.

[55] O. Bongomin, A. Yemane, B. Kembabazi et al., "Industry 4.0 disruption and its neologisms in major industrial sectors: a state of the art," *Journal of Engineering*, vol. 2020, Article ID 8090521, pp. 1–45, 2020.

[56] I. H. Sarker, "AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems," *SN Computer Science*, vol. 3, no. 2, p. 158, 2022.

[57] S. Devasahayam and B. Albijanic, "Predicting hydrogen production from co-gasification of biomass and plastics using tree based machine learning algorithms," *Renewable Energy*, vol. 222, article 119883, 2024.

[58] A. E. Sison, S. A. Etchieson, F. Güleç, E. I. Epelle, and J. A. Okolie, "Process modelling integrated with interpretable machine learning for predicting hydrogen and char yield during chemical looping gasification," *Journal of Cleaner Production*, vol. 414, article 137579, 2023.

[59] U. A. Dodo, E. C. Ashigwuike, J. N. Emechebe, and S. I. Abba, "Prediction of energy content of biomass based on hybrid machine learning ensemble algorithm," *Energy Nexus*, vol. 8, article 100157, 2022.

[60] B. M. Greenwell, B. C. Boehmke, and A. J. McCarthy, "A simple and effective model-based variable importance measure," https://arxiv.org/1805.04755. 2018.

[61] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, Cham: Springer International Publishing, 2019.

[62] V. Petsiuk, A. Das, and K. Saenko, "Rise: Randomized input sampling for explanation of black-box models," https://arxiv.org/1806.07421. 2018.

[63] R. F. Carvalhode, V. d. H. Pasolini, J. G. Fraga Breciani, A. B. Silva Costa, and R. C. d. Sousa, "Poultry manure combustion parameters to produce bioenergy: a thermogravimetric analysis by isoconventional models and machine learning," *Case Studies in Thermal Engineering*, vol. 53, article 103757, 2024.

[64] A. Hai, G. Bharath, M. F. Patah et al., "Machine learning models for the prediction of total yield and specific surface area of biochar derived from agricultural biomass by pyrolysis," *Environmental Technology and Innovation*, vol. 30, article 103071, 2023.

[65] S. Rodgers, A. Bowler, L. Wells et al., "A surrogate model for the economic evaluation of renewable hydrogen production from biomass feedstocks via supercritical water gasification," *International Journal of Hydrogen Energy*, vol. 49, pp. 277–294, 2024.

[66] H. Lee, I. H. Choi, and K. R. Hwang, "Machine learning prediction of bio-oil production from the pyrolysis of lignocellulosic biomass: recent advances and future perspectives," *Journal of Analytical and Applied Pyrolysis*, vol. 179, article 106486, 2024.

[67] I. Khurram Faridi, E. Tsotsas, W. Heineken, M. Koegler, and A. Kharaghani, "Development of a neural network model predictive controller for the fluidized bed biomass gasification process," *Chemical Engineering Science*, vol. 293, article 120000, 2024.

[68] Y. Fang, X. Li, X. Wang, L. Dai, R. Ruan, and S. You, "Machine learning-based multi-objective optimization of concentrated solar thermal gasification of biomass incorporating life cycle assessment and techno-economic analysis," *Energy Conversion and Management*, vol. 302, article 118137, 2024.

[69] Y. Ögren, A. Sepman, E. Fooladgar, F. Weiland, and H. Wiinikka, "Development and evaluation of a vision driven sensor for estimating fuel feeding rates in combustion and gasification processes," *Energy and AI*, vol. 15, article 100316, 2024.

[70] S. Tomasek, Á. Bárkányi, A. Egedy, and N. Miskolczi, "Model-based determination of optimal operating parameters for different solid waste gasification," *Chemical Engineering Journal Advances*, vol. 17, article 100586, 2024.

[71] N. Wang, H. He, Y. Wang et al., "Machine learning-driven optimization of Ni-based catalysts for catalytic steam reforming of biomass tar," *Energy Conversion and Management*, vol. 300, article 117879, 2024.

[72] M. W. Shahzad, V. H. Nguyen, B. B. Xu et al., "Machine learning assisted prediction of solar to liquid fuel production: a case study," *Process Safety and Environmental Protection*, vol. 184, pp. 1119–1130, 2024.

[73] D. A. Akinpelu, O. A. Adekoya, P. O. Oladoye, C. C. Ogbaga, and J. A. Okolie, "Machine learning applications in biomass pyrolysis: from biorefinery to end-of-life product management," *Digital Chemical Engineering*, vol. 8, article 100103, 2023.

[74] A. Coşgun, M. E. Günay, and R. Yıldırım, "A critical review of machine learning for lignocellulosic ethanol production via fermentation route," *Biofuel Research Journal*, vol. 10, no. 2, pp. 1859–1875, 2023.

[75] W. George Davies, S. Babamohammadi, Y. Yang, and S. Masoudi Soltani, "The rise of the machines: a state-of-the-art technical review on process modelling and machine learning within hydrogen production with carbon capture," *Gas Science and Engineering*, vol. 118, article 205104, 2023.

[76] H. Li, J. Chen, W. Zhang et al., "Machine-learning-aided thermochemical treatment of biomass: a review," *Biofuel Research Journal*, vol. 10, no. 1, pp. 1786–1809, 2023.

[77] M. Liao and Y. Yao, "Applications of artificial intelligence-based modeling for bioenergy systems: a review," *GCB Bioenergy*, vol. 13, no. 5, pp. 774–802, 2021.

[78] Y. Zhang, Y. Ji, and H. Qian, "Progress in thermodynamic simulation and system optimization of pyrolysis and gasification of biomass," *Green Chemical Engineering*, vol. 2, no. 3, pp. 266–283, 2021.

[79] G. C. Umenweke, I. C. Afolabi, E. I. Epelle, and J. A. Okolie, "Machine learning methods for modeling conventional and hydrothermal gasification of waste biomass: a review," *Bioresource Technology Reports*, vol. 17, article 100976, 2022.

[80] O. Sedej, E. Mbonimpa, T. Sleight, and J. Slagley, "Artificial neural networks and gradient boosted machines used for regression to evaluate gasification processes: a review," *Journal of Energy and Power Technology*, vol. 4, no. 3, p. 1, 2022.

[81] S. Ascher, W. Sloan, I. Watson, and S. You, "A comprehensive artificial neural network model for gasification process prediction," *Applied Energy*, vol. 320, article 119289, 2022.

[82] S. Ascher, X. Wang, I. Watson, W. Sloan, and S. You, "Interpretable machine learning to model biomass and waste gasification," *Bioresource Technology*, vol. 364, article 128062, 2022.

[83] S. Ascher, I. Watson, and S. You, "Machine learning methods for modelling the gasification and pyrolysis of biomass and waste," *Renewable and Sustainable Energy Reviews*, vol. 155, article 111902, 2022.

[84] W. Zhang, Q. Chen, J. Chen et al., "Machine learning for hydrothermal treatment of biomass: a review," *Bioresource Technology*, vol. 370, article 128547, 2023.

[85] M. Khan, S. Raza Naqvi, Z. Ullah et al., "Applications of machine learning in thermochemical conversion of biomass-a review," *Fuel*, vol. 332, article 126055, 2023.

[86] A. Sakheta, R. Nayak, I. O'Hara, and J. Ramirez, "A review on modelling of thermochemical processing of biomass for biofuels and prospects of artificial intelligence-enhanced approaches," *Bioresource Technology*, vol. 386, article 129490, 2023.

[87] F. Alfarra, H. K. Ozcan, P. Cihan, A. Ongen, S. Y. Guvenc, and M. N. Ciner, "Artificial intelligence methods for modeling gasification of waste biomass: a review," *Environmental Monitoring and Assessment*, vol. 196, no. 3, p. 309, 2024.

[88] A. D. A. Bin Abu Sofian, H. R. Lim, K. W. Chew et al., "Hydrogen production and pollution mitigation: enhanced gasification of plastic waste and biomass with machine learning & storage for a sustainable future," *Environmental Pollution*, vol. 342, article 123024, 2024.

[89] S. Pang, "Fuel flexible gas production: biomass, coal and biosolid wastes," in *Fuel Flexible Energy Generation*, pp. 241–269, Elsevier Ltd, 2020.

[90] N. Indrawan, B. Simkins, A. Kumar, and R. L. Huhnke, "Economics of distributed power generation via gasification of biomass and municipal solid waste," *Energies*, vol. 13, no. 3390, pp. 1–18, 2020.

[91] T. Kittivech and S. Fukuda, "Investigating agglomeration tendency of co-gasification between high alkali biomass and woody biomass in a bubbling fluidized bed system," *Energies*, vol. 13, no. 1, p. 56, 2020.

[92] D. Mallick, P. Mahanta, and V. S. Moholkar, "Co–gasification of coal/biomass blends in 50 kWe circulating fluidized bed gasifier," *Journal of the Energy Institute*, vol. 93, no. 1, pp. 99–111, 2020.

[93] V. S. Sikarwar, M. Hrabovský, G. Van Oost, M. Pohořelý, and M. Jeremiáš, "Progress in waste utilization via thermal plasma," *Progress in Energy and Combustion Science*, vol. 81, article 100873, 2020.

[94] A. Molino, S. Chianese, and D. Musmarra, "Biomass gasification technology: the state of the art overview," *Journal of Energy Chemistry*, vol. 25, no. 1, pp. 10–25, 2016.

[95] A. Gani, M. Zaki, R. Mamat et al., "Analysis of technological developments and potential of biomass gasification as a viable industrial process: a review," *Case Studies in Chemical and Environmental Engineering*, vol. 8, article 100439, 2023.

[96] O. Bongomin and P. Nziu, "A critical review on the development and utilization of energy systems in Uganda," *The Scientific World Journal*, vol. 2022, Article ID 2599467, 25 pages, 2022.

[97] R. Eghtedaei, S. A. Mirhosseini, M. J. Esfahani et al., "Co-gasification of biomass and municipal solid waste for hydrogen-rich gas production," *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, vol. 39, no. 14, pp. 1491–1496, 2017.

[98] A. Ramos, E. Monteiro, V. Silva, and A. Rouboa, "Co-gasification and recent developments on waste-to-energy conversion: a review," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 380–398, 2018.

[99] S. K. Thengane, A. Gupta, and S. M. Mahajani, "Co-gasification of high ash biomass and high ash coal in downdraft gasifier," *Bioresource Technology*, vol. 273, pp. 159–168, 2019.

[100] V. Esmaeili, J. Ajalli, A. Faramarzi, M. Abdi, and M. Gholizadeh, "Gasification of wastes: the impact of the feedstock type and co-gasification on the formation of volatiles and char," *International Journal of Energy Research*, vol. 44, no. 5, pp. 3587–3606, 2020.

[101] S. Salavati, C. Zhang, S. Zhang, Q. Liu, M. Gholizadeh, and X. Hu, "Cross-interaction during co-gasi fication of wood , weed, plastic, tire and carton," *Journal of Environmental Management*, vol. 250, article 109467, 2019.

[102] C. Lu, X. Zhang, Y. Gao et al., "Parametric study of catalytic co-gasification of cotton stalk and aqueous phase from wheat straw using hydrothermal carbonation," *Energy*, vol. 216, article 119266, 2021.

[103] W. X. Peng, S. B. Ge, A. G. Ebadi, H. Hisoriev, and M. J. Esfahani, "Syngas production by catalytic co-gasification of coal-biomass blends in a circulating fluidized bed gasifier," *Journal of Cleaner Production*, vol. 168, pp. 1513–1517, 2017.

[104] H. Song, H. Yang, C. Zhao et al., "Co-gasification of petroleum coke with coal at high temperature: effects of blending ratio and the catalyst," *Fuel*, vol. 307, article 121863, 2022.

[105] D. Hantoko, E. Kanchanatip, M. Yan, J. Lin, and Z. Weng, "Co-gasification of sewage sludge and lignite coal in supercritical water for H2 production: a thermodynamic modelling approach," *Energy Procedia*, vol. 152, pp. 1284–1289, 2018.

[106] A. Seçer, E. Fakı, Ş. Türker Üzden, and A. Hasanoğlu, "Hydrothermal co-gasification of sorghum biomass and çan lignite in mild conditions: an optimization study for high yield hydrogen production," *International Journal of Hydrogen Energy*, vol. 45, no. 4, pp. 2668–2680, 2020.

[107] Y. Xiao, S. Xu, Y. Liu, and C. Qiao, "Catalytic steam co-gasification of biomass and coal in a dual loop gasification system with olivine catalysts," *Journal of the Energy Institute*, vol. 93, no. 3, pp. 1074–1082, 2020.

[108] N. Ghavami, K. Özdenkçi, G. Salierno, M. Björklund-Sänkiaho, and C. De Blasio, "Analysis of operational issues in hydrothermal liquefaction and supercritical water gasification processes: a review," *Biomass Conversion and Biorefinery*, vol. 13, no. 14, pp. 12367–12394, 2023.

[109] J. Huang, C. Zhu, X. Lian et al., "Catalytic supercritical water gasification of glucose with in-situ generated nickel nanoparticles for hydrogen production," *International Journal of Hydrogen Energy*, vol. 44, no. 38, pp. 21020–21029, 2019.

[110] S. Nanda, J. A. Okolie, R. Patel et al., "Catalytic hydrothermal co-gasification of canola meal and low-density polyethylene using mixed metal oxides for hydrogen production," *International Journal of Hydrogen Energy*, vol. 47, no. 100, pp. 42084–42098, 2022.

[111] Y. S. Jeong, J. W. Kim, M. W. Seo, T. Y. Mun, and J. S. Kim, "Characteristics of two-stage air gasification of polystyrene

with active carbon as a tar removal agent," *Energy*, vol. 219, article 119681, 2021.

[112] H. O. Kargbo, J. Zhang, and A. N. Phan, "Optimisation of two-stage biomass gasification for hydrogen production via artificial neural network," *Applied Energy*, vol. 302, article 117567, 2021.

[113] M. A. Yazdanpanah Jahromi, K. Atashkari, and M. Kalteh, "Development of a high-temperature two-stage entrained flow gasifier model for the process of biomass gasification and syngas formation," *International Journal of Energy Research*, vol. 43, no. 11, pp. 5864–5878, 2019.

[114] B. Hu, Q. Huang, A. Buekens, Y. Chi, and J. Yan, "Co-gasification of municipal solid waste with high alkali coal char in a three-stage gasifier," *Energy Conversion and Management*, vol. 153, pp. 473–481, 2017.

[115] Y. Hu, K. Pang, L. Cai, and Z. Liu, "A multi-stage co-gasification system of biomass and municipal solid waste (MSW) for high quality syngas production," *Energy*, vol. 221, article 119639, 2021.

[116] H. Song, G. Yang, P. Xue et al., "Recent development of biomass gasification for H2 rich gas production," *Applications in Energy and Combustion Science*, vol. 10, article 100059, 2022.

[117] D. K. Singh, R. Raj, J. V. Tirkey et al., "Progress and utilization of biomass gasification for decentralized energy generation: an outlook & critical review," *Environmental Technology Reviews*, vol. 12, no. 1, pp. 1–36, 2023.

[118] L. Ding, M. Yang, K. Dong et al., "Mobile power generation system based on biomass gasification," *International Journal of Coal Science & Technology*, vol. 9, no. 1, 2022.

[119] X. He, C. H. Wang, and C. A. Shoemaker, "Multi-objective optimization of an integrated biomass waste fixed-bed gasification system for power and biochar co-production," *Computers and Chemical Engineering*, vol. 154, article 107457, 2021.

[120] D. K. Omol, O. Acaye, D. F. Okot, and O. Bongomin, "Production of fuel oil from municipal plastic wastes using thermal and catalytic pyrolysis," *Journal of Energy Research and Reviews*, vol. 4, no. 2, pp. 1–8, 2020.

[121] V. S. Sikarwar, M. Zhao, P. Clough et al., "An overview of advances in biomass gasification," *Energy and Environmental Science*, vol. 9, no. 10, pp. 2939–2977, 2016.

[122] P. R. Rout, D. S. Pandey, M. Haynes-Parry et al., "Sustainable valorisation of animal manures via thermochemical conversion technologies: an inclusive review on recent trends," *Waste and Biomass Valorization*, vol. 14, no. 2, pp. 553–582, 2023.

[123] M. G. Rasul and S. Begum, "Energy recovery from solid waste: application of gasification technology," in *Handbook of Ecomaterials (Vol. 2)*, CQUniversity, 2019.

[124] D. M. Filho, S. T. Coelho, and D. Perecin, "Opportunities and challenges of gasification of municipal solid waste (MSW) in Brazil," *Energies*, vol. 15, no. 8, p. 2735, 2022.

[125] M. Hussain, O. Ali, N. Raza, H. Zabiri, A. Ahmed, and I. Ali, "Recent advances in dynamic modeling and control studies of biomass gasification for production of hydrogen rich syngas," *RSC Advances*, vol. 13, no. 34, pp. 23796–23811, 2023.

[126] M. Sharma and R. Kaushal, "Advances and challenges in the generation of bio-based fuels using gasifiers: a comprehensive review," *International Journal of Ambient Energy*, vol. 41, no. 14, pp. 1645–1663, 2020.

[127] J. C. Schmid, F. Benedikt, J. Fuchs, A. M. Mauerhofer, S. Müller, and H. Hofbauer, "Syngas for biorefineries from thermochemical gasification of lignocellulosic fuels and residues—5 years' experience with an advanced dual fluidized bed gasifier design," *Biomass Conversion and Biorefinery*, vol. 11, no. 6, 2021.

[128] M. Pérez-Fortes, V. He, A. Nakajo, J. Schiffmann, F. Maréchal, and J. van Herle, "Techno-economic optimization of an integrated biomass waste gasifier–solid oxide fuel cell plant," *Frontiers in Energy Research*, vol. 9, pp. 1–20, 2021.

[129] K. Stasiak, M. Amiri, and D. Mikielewicz, "Negative carbon dioxide gas power plant integrated with gasification of sewage sludge," *Energy*, vol. 262, article 125496, 2023.

[130] A. I. M. C. Ramos, *Experimental Study of the Operating Conditions Impact on Fuel Blends Co-Gasification: Environmental and Socio-Economic Assessment*, University of Porto, 2020.

[131] A. A. Rabah, "Optimization of syngas quality for Fischer-Tropsch synthesis," *Journal of Energy*, vol. 2023, Article ID 1842187, 12 pages, 2023.

[132] N. Kamaruzaman, N. Abdul Manaf, D. Milani, and A. Abbas, "Assessing the current state of biomass gasification technology in advancing circular economies: a holistic analysis from techno-economic-policy perspective in Malaysia and beyond," *Chemical Engineering Research and Design*, vol. 199, pp. 593–619, 2023.

[133] M. Puig-Gamero, L. Sanchez-Silva, F. Dorado, and P. Sánchez, "Multi-criteria analysis for selecting the optimum blend in the co-gasification process," *Computers and Chemical Engineering*, vol. 141, article 106983, 2020.

[134] M. Akbari, A. O. Oyedun, and A. Kumar, "Ammonia production from black liquor gasification and co-gasification with pulp and waste sludges: a techno-economic assessment," *Energy*, vol. 151, pp. 133–143, 2018.

[135] H. O. Kargbo, J. Zhang, and A. N. Phan, "Robust modelling development for optimisation of hydrogen production from biomass gasification process using bootstrap aggregated neural network," *International Journal of Hydrogen Energy*, vol. 48, no. 29, pp. 10812–10828, 2023.

[136] M. Meena, H. Kumar, N. D. Chaturvedi et al., "Biomass gasification and applied intelligent retrieval in modeling," *Energies*, vol. 16, no. 18, p. 6524, 2023.

[137] H. Al-Qadasi and G. M. Ozkan, "CFD analysis of biomass steam gasification in fluidized bed gasifier: a parametric study by the assessment of drying stage," *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, vol. 43, no. 19, pp. 2369–2390, 2021.

[138] K. Rabea, S. Michailos, M. Akram, K. J. Hughes, D. Ingham, and M. Pourkashanian, "An improved kinetic modelling of woody biomass gasification in a downdraft reactor based on the pyrolysis gas evolution," *Energy Conversion and Management*, vol. 258, article 115495, 2022.

[139] S. K. Arumugasamy, A. Selvarajoo, and M. A. Tariq, "Artificial neural networks modelling: gasification behaviour of palm fibre biochar," *Materials Science for Energy Technologies*, vol. 3, pp. 868–878, 2020.

[140] D. Cerinski, J. Baleta, H. Mikulčić, R. Mikulandrić, and J. Wang, "Dynamic modelling of the biomass gasification process in a fixed bed reactor by using the artificial neural network," *Cleaner Engineering and Technology*, vol. 1, article 100029, 2020.

[141] Z. U. Haq, H. Ullah, M. N. A. Khan, S. R. Naqvi, and M. Ahsan, "Hydrogen production optimization from sewage

sludge supercritical gasification process using machine learning methods integrated with genetic algorithm," *Chemical Engineering Research and Design*, vol. 184, pp. 614–626, 2022.

[142] A. Ramos, E. Monteiro, and A. Rouboa, "Numerical approaches and comprehensive models for gasification process: a review," *Renewable and Sustainable Energy Reviews*, vol. 110, pp. 188–206, 2019.

[143] D. D. Sapariya, U. J. Patdiwala, H. Panchal, P. V. Ramana, J. Makwana, and K. K. Sadasivuni, "A review on thermochemical biomass gasification techniques for bioenergy production," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 6, no. 3, pp. 1–34, 2021.

[144] Y. Sun, Z. Li, X. Li, and J. Zhang, "Classifier selection and ensemble model for multi-class imbalance learning in education grants prediction," *Applied Artificial Intelligence*, vol. 35, no. 4, pp. 290–303, 2021.

[145] P. Mahajan, S. Uddin, F. Hajati, and M. A. Moni, "Ensemble learning for disease prediction: a review," *Healthcare*, vol. 11, no. 12, 2023.

[146] S. Ali, T. Abuhmed, S. El-Sappagh et al., "Explainable artificial intelligence (XAI): what we know and what is left to attain trustworthy artificial intelligence," *Information Fusion*, vol. 99, article 101805, 2023.

[147] V. Hassija, V. Chamola, A. Mahapatra et al., "Interpreting black-box models: a review on explainable artificial intelligence," *Cognitive Computation*, vol. 16, no. 1, pp. 45–74, 2024.

[148] W. Feng, J. Gou, Z. Fan, and X. Chen, "An ensemble machine learning approach for classification tasks using feature generation," *Connection Science*, vol. 35, no. 1, 2023.

[149] A. Zizaan and A. Idri, "Evaluating and comparing bagging and boosting of hybrid learning for breast cancer screening," *Scientific African*, vol. 23, article e01989, 2024.

[150] N. Kardani, A. Zhou, M. Nazem, and X. Lin, "Modelling of municipal solid waste gasification using an optimised ensemble soft computing model," *Fuel*, vol. 289, article 119903, 2021.

[151] A. Sharafati, S. B. H. S. Asadollah, and M. Hosseinzadeh, "The potential of new ensemble machine learning models for effluent quality parameters prediction and related uncertainty," *Process Safety and Environmental Protection*, vol. 140, pp. 68–78, 2020.

[152] M. Bagheri, A. Akbari, and S. A. Mirbagheri, "Advanced control of membrane fouling in filtration systems using artificial intelligence and machine learning techniques: a critical review," *Process Safety and Environmental Protection*, vol. 123, pp. 229–252, 2019.

[153] B. Minaei-Bidgoli, H. Parvin, H. Alinejad-Rokny, H. Alizadeh, and W. F. Punch, "Effects of resampling method and adaptation on clustering ensemble efficacy," *Artificial Intelligence Review*, vol. 41, no. 1, pp. 27–48, 2014.

[154] O. Sagi and L. Rokach, "Ensemble learning: a survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, pp. 1–18, 2018.

[155] S. Lee, N. Nguyen, A. Karamanli, J. Lee, and T. P. Vo, "Super learner machine-learning algorithms for compressive strength prediction of high performance concrete," *Structural Concrete*, vol. 24, no. 2, pp. 2208–2228, 2023.

[156] I. D. Mienye and Y. Sun, "A survey of ensemble learning: concepts, algorithms, applications, and prospects," *IEEE Access*, vol. 10, pp. 99129–99149, 2022.

[157] C. Giraud-Carrier, "Combining base-learners into ensembles," in *Metalearning: Applications to Automated Machine Learning and Data Mining 2022 Feb 22 (pp. 169-188)*Springer International Publishing, Cham.

[158] H. I. Erdal, "Two-level and hybrid ensembles of decision trees for high performance concrete compressive strength prediction," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 7, pp. 1689–1697, 2013.

[159] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: a review," *Engineering Applications of Artificial Intelligence*, vol. 115, article 105151, 2022.

[160] Y. Yang, "Ensemble learning," in *Temporal Data Mining Via Unsupervised Ensemble Learning*, pp. 35–56, Elsevier, 2017.

[161] D. De Clercq, Z. Wen, F. Fei, L. Caicedo, K. Yuan, and R. Shang, "Interpretable machine learning for predicting biomethane production in industrial-scale anaerobic co-digestion," *Science of the Total Environment*, vol. 712, article 134574, 2020.

[162] C. H. Chen, K. Tanaka, M. Kotera, and K. Funatsu, "Comparison and improvement of the predictability and interpretability with ensemble learning models in QSPR applications," *Journal of Cheminformatics*, vol. 12, no. 1, p. 19, 2020.

[163] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[164] J. Xing, H. Wang, K. Luo, S. Wang, Y. Bai, and J. Fan, "Predictive single-step kinetic model of biomass devolatilization for CFD applications: a comparison study of empirical correlations (EC), artificial neural networks (ANN) and random forest (RF)," *Renewable Energy*, vol. 136, pp. 104–114, 2019.

[165] J. Xing, K. Luo, H. Wang, and J. Fan, "Estimating biomass major chemical constituents from ultimate analysis using a random forest model," *Bioresource Technology*, vol. 288, article 121541, 2019.

[166] J. Xing, K. Luo, H. Wang, Z. Gao, and J. Fan, "A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches," *Energy*, vol. 188, article 116077, 2019.

[167] Q. Tang, Y. Chen, H. Yang et al., "Prediction of bio-oil yield and hydrogen contents based on machine learning method: effect of biomass compositions and pyrolysis conditions," *Energy and Fuels*, vol. 34, no. 9, pp. 11050–11060, 2020.

[168] M. H. D. M. Ribeiro and L. dos Santos Coelho, "Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series," *Applied Soft Computing*, vol. 86, article 105837, 2020.

[169] Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*, Taylor & Francis Group, 2012.

[170] N. Aslam, I. U. Khan, S. Mirza et al., "Interpretable machine learning models for malicious domains detection using explainable artificial intelligence (XAI)," *Sustainability*, vol. 14, no. 12, p. 7375, 2022.

[171] C. Cakiroglu, G. Bekdaş, S. Kim, and Z. W. Geem, "Explainable ensemble learning models for the rheological properties of self-compacting concrete," *Sustainability*, vol. 14, no. 21, article 14640, 2022.

[172] H. L. Zheng, S. Y. An, B. J. Qiao, P. Guan, D. S. Huang, and W. Wu, "A data-driven interpretable ensemble framework based on tree models for forecasting the occurrence of COVID-19 in the USA," *Environmental Science and Pollution Research*, vol. 30, no. 5, pp. 13648–13659, 2023.

[173] L. Leng, L. Yang, X. Lei et al., "Machine learning predicting and engineering the yield, N content, and specific surface area

of biochar derived from pyrolysis of biomass," *Biochar*, vol. 4, no. 1, 2022.

[174] H. T. Wen, J. H. Lu, and M. X. Phuc, "Applying artificial intelligence to predict the composition of syngas using rice husks: a comparison of artificial neural networks and gradient boosting regression," *Energies*, vol. 14, no. 10, p. 2932, 2021.

[175] T. Boot and D. Nibbering, "Forecasting using random subspace methods," *Journal of Econometrics*, vol. 209, no. 2, pp. 391–406, 2019.

[176] J. Shin, "Random subspace ensemble learning for functional near-infrared spectroscopy brain-computer interfaces," *Frontiers in Human Neuroscience*, vol. 14, pp. 1–9, 2020.

[177] D. Ballabio, R. Todeschini, and V. Consonni, "Recent advances in high-level fusion methods to classify multiple analytical chemical data," *Data Handling in Science and Technology*, vol. 31, pp. 129–155, 2019.

[178] M. A. Yaman, A. Subasi, and F. Rattay, "Comparison of random subspace and voting ensemble machine learning methods for face recognition," *Symmetry*, vol. 10, no. 11, p. 651, 2018.

[179] V. C. Osamor and A. F. Okezie, "Enhancing the weighted voting ensemble algorithm for tuberculosis predictive diagnosis," *Scientific Reports*, vol. 11, no. 1, article 14806, 2021.

[180] C. Ju, A. Bibaut, and M. van der Laan, "The relative performance of ensemble methods with deep convolutional neural networks for image classification," *Journal of Applied Statistics*, vol. 45, no. 15, pp. 2800–2818, 2018.

[181] S. K. Satapathy, A. K. Jagadev, and S. Dehuri, "Weighted majority voting based ensemble of classifiers using different machine learning techniques for classification of EEG signal to detect epileptic seizure," *Informatica*, vol. 41, no. 1, pp. 99–110, 2017.

[182] A. Dogan and D. Birant, "A weighted majority voting ensemble approach for classification," in *2019 4th International Conference on Computer Science and Engineering (UBMK)*, pp. 1–6, Samsun, Turkey, 2019.

[183] Y. Zhang, H. Zhang, J. Cai, and B. Yang, "A weighted voting classifier based on differential evolution," *Abstract and Applied Analysis*, vol. 2014, Article ID 376950, 6 pages, 2014.

[184] B. Erdebilli and B. Devrim-İçtenbaş, "Ensemble voting regression based on machine learning for predicting medical waste: a case from Turkey," *Mathematics*, vol. 10, no. 14, p. 2466, 2022.

[185] V. Nourani, G. Elkiran, and S. I. Abba, "Wastewater treatment plant performance analysis using artificial intelligence - an ensemble approach," *Water Science and Technology*, vol. 78, no. 10, pp. 2064–2076, 2018.

[186] I. K. Nti, A. F. Adekoya, and B. A. Weyori, "A comprehensive evaluation of ensemble learning for stock-market prediction," *Journal of Big Data*, vol. 7, no. 1, 2020.

[187] S. I. Abba, N. T. T. Linh, J. Abdullahi et al., "Hybrid machine learning ensemble techniques for modeling dissolved oxygen concentration," *IEEE Access*, vol. 8, pp. 157218–157237, 2020.

[188] M. Massaoudi, S. S. Refaat, I. Chihi, M. Trabelsi, F. S. Oueslati, and H. Abu-Rub, "A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for short-term load forecasting," *Energy*, vol. 214, article 118874, 2021.

[189] P. Kumari and D. Toshniwal, "Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance," *Journal of Cleaner Production*, vol. 279, article 123285, 2021.

[190] G. Valdes, Y. Interian, E. Gennatas, and M. Van der Laan, "The conditional super learner," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 10236–10243, 2022.

[191] T. Wu, W. Zhang, X. Jiao, W. Guo, and Y. Alhaj Hamoud, "Evaluation of stacking and blending ensemble learning methods for estimating daily reference evapotranspiration," *Computers and Electronics in Agriculture*, vol. 184, article 106039, 2021.

[192] S. Shafieian and M. Zulkernine, "Multi-layer stacking ensemble learners for low footprint network intrusion detection," *Complex and Intelligent Systems*, vol. 9, no. 4, pp. 3787–3799, 2023.

[193] N. Liu, H. Gao, Z. Zhao, Y. Hu, and L. Duan, "A stacked generalization ensemble model for optimization and prediction of the gas well rate of penetration: a case study in Xinjiang," *Journal of Petroleum Exploration and Production Technology*, vol. 12, no. 6, pp. 1595–1608, 2022.

[194] M. J. Van Der Laan, E. C. Polley, and A. E. Hubbard, "Super learner," *Statistical Applications in Genetics and Molecular Biology*, vol. 6, no. 1, 2007.

[195] J. Yao, X. Zhang, W. Luo, C. Liu, and L. Ren, "Applications of stacking/blending ensemble learning approaches for evaluating flash flood susceptibility," *International Journal of Applied Earth Observation and Geoinformation*, vol. 112, article 102932, 2022.

[196] R. Wyss, S. Schneeweiss, M. Van Der Laan, S. D. Lendle, C. Ju, and J. M. Franklin, "Using super learner prediction modeling to improve high-dimensional propensity score estimation," *Epidemiology*, vol. 29, no. 1, pp. 96–106, 2018.

[197] A. I. Naimi and L. B. Balzer, "Stacked generalization: an introduction to super learning," *European Journal of Epidemiology*, vol. 33, no. 5, pp. 459–464, 2018.

[198] C. Ju, M. Combs, S. D. Lendle et al., "Propensity score prediction for electronic healthcare databases using super learner and high-dimensional propensity score methods," *Journal of Applied Statistics*, vol. 46, no. 12, pp. 2216–2236, 2019.

[199] M. Dhrioua, K. Ghachem, W. Hassen, A. Ghazy, L. Kolsi, and M. N. Borjini, "Simulation of biomass air gasification in a bubbling fluidized bed using Aspen plus: a comprehensive model including tar production," *ACS Omega*, vol. 7, no. 37, pp. 33518–33529, 2022.

[200] M. A. Hamad, A. M. Radwan, D. A. Heggo, and T. Moustafa, "Hydrogen rich gas production from catalytic gasification of biomass," *Renewable Energy*, vol. 85, pp. 1290–1300, 2016.

[201] R. G. Jenkins, "Thermal gasification of biomass–a primer," in *Bioenergy: Biomass to Biofuels and Waste to Energy*, Elsevier, 2nd edition, 2020.

[202] N. Calderon Henao, O. J. Venturini, Y. Castillo Santiago et al., "Energy and economic assessment of a system integrated by a biomass downdraft gasifier and a gas microturbine," *Processes*, vol. 10, no. 11, p. 2377, 2022.

[203] A. Y. Mutlu and O. Yucel, "An artificial intelligence based approach to predicting syngas composition for downdraft biomass gasification," *Energy*, vol. 165, pp. 895–901, 2018.

[204] F. Elmaz, Ö. Yücel, and A. Y. Mutlu, "Predictive modeling of biomass gasification with machine learning-based regression methods," *Energy*, vol. 191, article 116541, 2020.

[205] A. Y. Mutlu, Ö. Yücel, and F. Elmaz, "Evaluating the effect of blending ratio on the co-gasification of high ash coal and

biomass in a fluidized bed gasifier using machine learning," *Mugla Journal of Science and Technology*, vol. 5, no. 1, pp. 1–12, 2019.

[206] H. M. U. Ayub, M. Rafiq, M. A. Qyyum, G. Rafiq, G. S. Choi, and M. Lee, "Prediction of process parameters for the integrated biomass gasification power plant using artificial neural network," *Frontiers in Energy Research*, vol. 10, pp. 1–9, 2022.

[207] F. Kartal and U. Özveren, "A deep learning approach for prediction of syngas lower heating value from CFB gasifier in Aspen plus®," *Energy*, vol. 209, article 118457, 2020.

[208] D. Serrano and D. Castelló, "Tar prediction in bubbling fluidized bed gasification through artificial neural networks," *Chemical Engineering Journal*, vol. 402, article 126229, 2020.

[209] D. Serrano, I. Golpour, and S. Sánchez-Delgado, "Predicting the effect of bed materials in bubbling fluidized bed gasification using artificial neural networks (ANNs) modeling approach," *Fuel*, vol. 266, article 117021, 2020.

[210] O. Yucel, E. S. Aydin, and H. Sadikoglu, "Comparison of the different artificial neural networks in prediction of biomass gasification products," *International Journal of Energy Research*, vol. 43, no. 11, pp. 5992–6003, 2019.

[211] M. Shahbaz, S. A. Taqvi, A. C. Minh Loy et al., "Artificial neural network approach for the steam gasification of palm oil waste using bottom ash and CaO," *Renewable Energy*, vol. 132, pp. 243–254, 2019.

[212] S. Safarian, S. M. Ebrahimi Saryazdi, R. Unnthorsson, and C. Richter, "Modeling of hydrogen production by applying biomass gasification: artificial neural network modeling approach," *Fermentation*, vol. 7, no. 2, p. 71, 2021.

[213] D. Baruah, D. C. Baruah, and M. K. Hazarika, "Artificial neural network based modeling of biomass gasification in fixed bed downdraft gasifiers," *Biomass and Bioenergy*, vol. 98, pp. 264–271, 2017.

[214] E. E. Ozbas, D. Aksu, A. Ongen, M. A. Aydin, and H. K. Ozcan, "Hydrogen production via biomass gasification, and modeling by supervised machine learning algorithms," *International Journal of Hydrogen Energy*, vol. 44, no. 32, pp. 17260–17268, 2019.

[215] S. Zhao, J. Li, C. Chen, B. Yan, J. Tao, and G. Chen, "Interpretable machine learning for predicting and evaluating hydrogen production via supercritical water gasification of biomass," *Journal of Cleaner Production*, vol. 316, article 128244, 2021.

[216] D. S. Pandey, H. Raza, and S. Bhattacharyya, "Development of explainable AI-based predictive models for bubbling fluidised bed gasification process," *Fuel*, vol. 351, article 128971, 2023.

[217] R. Aguado, J. L. Casteleiro-Roca, D. Vera, and J. L. Calvo-Rolle, "A hybrid intelligent model to predict the hydrogen concentration in the producer gas from a downdraft gasifier," *International Journal of Hydrogen Energy*, vol. 47, no. 48, pp. 20755–20770, 2022.

[218] Y. Yang, H. Shahbeik, A. Shafizadeh et al., "Predicting municipal solid waste gasification using machine learning: a step toward sustainable regional planning," *Energy*, vol. 278, article 127881, 2023.

[219] R. Hasanzadeh and T. Azdast, "Machine learning utilization on air gasification of polyethylene terephthalate waste," *Waste Management Bulletin*, vol. 2, no. 1, pp. 75–82, 2024.

[220] M. V. Gil, K. M. Jablonka, S. Garcia, C. Pevida, and B. Smit, "Biomass to energy: a machine learning model for optimum gasification pathways," *Digital Discovery*, vol. 2, no. 4, pp. 929–940, 2023.

[221] R. Mikulandrić, D. Lončar, D. Böhning, R. Böhme, and M. Beckmann, "Artificial neural network modelling approach for a biomass gasification process in fixed bed gasifiers," *Energy Conversion and Management*, vol. 87, pp. 1210–1223, 2014.

[222] D. S. Pandey, S. Das, I. Pan, J. J. Leahy, and W. Kwapinski, "Artificial neural network based modelling approach for municipal solid waste gasification in a fluidized bed reactor," *Waste Management*, vol. 58, pp. 202–213, 2016.

[223] A. Molino and G. Braccio, "Synthetic natural gas SNG production from biomass gasification - thermodynamics and processing aspects," *Fuel*, vol. 139, pp. 425–429, 2015.

[224] M. Shahbaz, S. Yusup, A. Inayat et al., "Syngas production from steam gasification of palm kernel shell with subsequent CO2capture using CaO sorbent: an Aspen plus modeling," *Energy and Fuels*, vol. 31, no. 11, pp. 12350–12357, 2017.

[225] S. Valin, S. Ravel, P. Pons de Vincent et al., "Fluidised bed gasification of diverse biomass feedstocks and blends—an overall performance study," *Energies*, vol. 13, no. 14, p. 3706, 2020.

[226] R. Khonde, S. Hedaoo, and S. Deshmukh, "Prediction of product gas composition from biomass gasification by the method of Gibbs free energy minimization," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 43, no. 3, pp. 371–380, 2021.

[227] J. Hu, J. Shao, H. Yang et al., "Co-gasification of coal and biomass: synergy, characterization and reactivity of the residual char," *Bioresource Technology*, vol. 244, Part 1, pp. 1–7, 2017.

[228] A. Anniwaer, N. Chaihad, A. C. A. Zahra et al., "Steam co-gasification of Japanese cedarwood and its commercial biochar for hydrogen-rich gas production," *International Journal of Hydrogen Energy*, vol. 46, no. 70, pp. 34587–34598, 2021.

[229] Z. Zhang and S. Pang, "Experimental investigation of biomass devolatilization in steam gasification in a dual fluidised bed gasifier," *Fuel*, vol. 188, pp. 628–635, 2017.

[230] M. Asadullah, "Barriers of commercial power generation using biomass gasification gas: a review," *Renewable and Sustainable Energy Reviews*, vol. 29, pp. 201–215, 2014.

[231] F. L. Chan and A. Tanksale, "Review of recent developments in Ni-based catalysts for biomass gasification," *Renewable and Sustainable Energy Reviews*, vol. 38, pp. 428–438, 2014.

[232] Z. Khan, S. Yusup, M. M. Ahmad, and N. A. Rashidi, "Integrated catalytic adsorption (ICA) steam gasification system for enhanced hydrogen production using palm kernel shell," *International Journal of Hydrogen Energy*, vol. 39, no. 7, pp. 3286–3293, 2014.

[233] A. Alamia, A. Larsson, C. Breitholtz, and H. Thunman, "Performance of large-scale biomass gasifiers in a biorefinery, a state-of-the-art reference," *International Journal of Energy Research*, vol. 41, no. 14, pp. 2001–2019, 2017.

[234] A. Erasmus, T. D. P. Brunet, and E. Fisher, "What is interpretability?," *Philosophy and Technology*, vol. 34, no. 4, pp. 833–862, 2021.

[235] F. Doshi-Velez and B. Kim, *Towards A rigorous science of interpretable machine learning*, Ml, 2017.

[236] Z. C. Lipton, "The mythos of model interpretability," *Communications of the ACM*, vol. 61, no. 10, pp. 36–43, 2018.

[237] B. J. M. Ale and D. H. Slater, "Complexity for complexity—-how advanced modeling may limit its applicability for decision-makers," *Risk Analysis*, pp. 1–11, 2023.

[238] T. Räz and C. Beisbart, "The importance of understanding deep learning," *Erkenntnis*, vol. 89, no. 5, pp. 1823–1840, 2024.

[239] A. Aldoseri, K. N. Al-Khalifa, and A. M. Hamouda, "Re-thinking data strategy and integration for artificial intelligence: concepts, opportunities, and challenges," *Applied Sciences*, vol. 13, no. 12, p. 7082, 2023.

[240] T. Laugel, M. J. Lesot, C. Marsala, X. Renard, and M. Detyniecki, "Comparison-Based Inverse Classification for Interpretability in Machine Learning," in *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations. IPMU 2018. Communications in Computer and Information Science*, J. Medina, M. Ojeda-Aciego, J. L. Verdegay, D. A. Pelta, I. P. Cabrera, B. Bouchon-Meunier, and R. R. Yager, Eds., vol. 853, Springer, Cham, 2018.

[241] A. Adadi and M. Berrada, "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018.

[242] R. Elshawi, M. H. Al-Mallah, and S. Sakr, "On the interpretability of machine learning-based model for predicting hypertension," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, 2019.

[243] J. Li, L. Li, Y. W. Tong, and X. Wang, "Understanding and optimizing the gasification of biomass waste with machine learning," *Green Chemical Engineering*, vol. 4, no. 1, pp. 123–133, 2023.

[244] T. Parr, J. Hamrick, and J. D. Wilson, "Nonparametric feature impact and importance," *Information Sciences*, vol. 653, no. 2, article 119563, 2024.

[245] O. Saadallah and Z. Rouissi, "Interpretable meta-active learning for regression ensemble learning," *CEUR Workshop Proceedings*, vol. 3470, pp. 89–103, 2023.

[246] M. Muschalik, F. Fumagalli, B. Hammer, and E. Hüllermeier, "Agnostic explanation of model change based on feature importance," *KI-Kunstliche Intelligenz*, vol. 36, no. 3, pp. 211–224, 2022.

[247] X. Yuan, M. Suvarna, S. Low et al., "Applied machine learning for prediction of CO2adsorption on biomass waste-derived porous carbons," *Environmental Science and Technology*, vol. 55, no. 17, pp. 11925–11936, 2021.

[248] B. Gregorutti, B. Michel, and P. Saint-Pierre, "Correlation and variable importance in random forests," *Statistics and Computing*, vol. 27, no. 3, pp. 659–678, 2017.

[249] K. Rasheed, A. Qayyum, M. Ghaly, A. Al-Fuqaha, A. Razi, and J. Qadir, "Explainable, trustworthy, and ethical machine learning for healthcare: a survey," *Computers in Biology and Medicine*, vol. 149, article 106043, 2022.

[250] A. Fisher, C. Rudin, and F. Dominici, "All models are wrong, but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously," *Journal of Machine Learning Research*, vol. 20, no. 177, pp. 1–81, 2019.

[251] J. Li, X. Zhu, Y. Li, Y. W. Tong, Y. S. Ok, and X. Wang, "Multi-task prediction and optimization of hydrochar properties from high-moisture municipal solid waste: application of machine learning on waste-to-resource," *Journal of Cleaner Production*, vol. 278, article 123928, 2021.

[252] S. Nembrini, I. R. König, and M. N. Wright, "The revival of the Gini importance?," *Bioinformatics*, vol. 34, no. 21, pp. 3711–3718, 2018.

[253] M. Rzychoń, A. Zogała, and L. Róg, "An interpretable extreme gradient boosting model to predict ash fusion temperatures," *Minerals*, vol. 10, no. 6, p. 487, 2020.

[254] R. Rodríguez-Pérez and J. Bajorath, "Interpretation of machine learning models using Shapley values: application to compound potency and multi-target activity predictions," *Journal of Computer-Aided Molecular Design*, vol. 34, no. 10, pp. 1013–1026, 2020.

[255] M. J. Ankenbrand, L. Shainberg, M. Hock, D. Lohr, and L. M. Schreiber, "Sensitivity analysis for interpretation of machine learning based segmentation models in cardiac MRI," *BMC Medical Imaging*, vol. 21, no. 1, p. 27, 2021.

[256] S. Ritsema, *A Feature Sensitivity and Dependency Analysis Approach for Model Explainability*, University of Twente, 2021.

[257] J. de Oña and C. Garrido, "Extracting the contribution of independent variables in neural network models: a new approach to handle instability," *Neural Computing and Applications*, vol. 25, no. 3–4, pp. 859–869, 2014.

[258] B. Iooss and P. Lemaître, "A review on global sensitivity analysis methods," *Operations Research/Computer Science Interfaces Series*, vol. 59, pp. 101–122, 2015.

[259] K. G. Link, M. T. Stobb, J. PaolaDi et al., "A local and global sensitivity analysis of a mathematical model of coagulation and platelet deposition under flow," *PLoS One*, vol. 13, no. 7, article e0200917, 2018.

[260] J. D. Olden and D. A. Jackson, "Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks," *Ecological Modelling*, vol. 154, no. 1–2, pp. 135–150, 2002.

[261] M. Cao, N. F. Alkayem, L. Pan, and D. Novák, "Advanced methods in neural networks-based sensitivity analysis with their applications in civil engineering," in *Artificial Neural Networks-Models and Applications (Vol. 25, issue 4, pp. e275–e281)*, InTech, 2016.

[262] M. Gevrey, I. Dimopoulos, and S. Lek, "Review and comparison of methods to study the contribution of variables in artificial neural network models," *Ecological Modelling*, vol. 160, no. 3, pp. 249–264, 2003.

[263] M. H. Shojaeefard, M. Akbari, M. Tahani, and F. Farhani, "Sensitivity analysis of the artificial neural network outputs in friction stir lap joining of aluminum to brass," *Advances in Materials Science and Engineering*, vol. 2013, Article ID 574914, 7 pages, 2013.

[264] D. W. Apley and J. Zhu, "Visualizing the effects of predictor variables in black box supervised learning models," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 82, no. 4, pp. 1059–1086, 2020.

[265] V. Vishwarupe, P. M. Joshi, N. Mathias, S. Maheshwari, S. Mhaisalkar, and V. Pawar, "Explainable AI and interpretable machine learning: a case study in perspective," *Procedia Computer Science*, vol. 204, no. 2021, pp. 869–876, 2022.

[266] R. Omobolaji Alabi, A. Almangush, M. Elmusrati, I. Leivo, and A. A. Mäkitie, "An interpretable machine learning prognostic system for risk stratification in oropharyngeal cancer," *International Journal of Medical Informatics*, vol. 168, article 104896, 2022.

[267] V. D'Amato, R. D'Ecclesia, and S. Levantesi, "Firms' profitability and ESG score: a machine learning approach," *Applied Stochastic Models in Business and Industry*, vol. 40, no. 2, pp. 243–261, 2024.

[268] J. Li, L. Pan, M. Suvarna, and X. Wang, "Machine learning aided supercritical water gasification for H2-rich syngas production with process optimization and catalyst screening," *Chemical Engineering Journal*, vol. 426, article 131285, 2021.

[269] H. Ishwaran, "The effect of splitting on random forests," *Machine Learning*, vol. 99, no. 1, pp. 75–118, 2015.

[270] P. Wei, Z. Lu, and J. Song, "Variable importance analysis: a comprehensive review," *Reliability Engineering and System Safety*, vol. 142, pp. 399–432, 2015.

[271] B. V. Ayodele, S. I. Mustapa, R. Kanthasamy, M. Zwawi, and C. K. Cheng, "Modeling the prediction of hydrogen production by co-gasification of plastic and rubber wastes using machine learning algorithms," *International Journal of Energy Research*, vol. 45, no. 6, pp. 9580–9594, 2021.

[272] H. S. M. Yahya, T. Abbas, and N. A. S. Amin, "Optimization of hydrogen production via toluene steam reforming over Ni–Co supported modified-activated carbon using ANN coupled GA and RSM," *International Journal of Hydrogen Energy*, vol. 46, no. 48, pp. 24632–24651, 2021.

[273] F. Elmaz and Ö. Yücel, "Data-driven identification and model predictive control of biomass gasification process for maximum energy production," *Energy*, vol. 195, pp. 1–14, 2020.