WCRG Global Research
WORLD COMET RESEARCH GROUP

International Journal of Modern Research in Science & Engineering
Vol. 7, No. 2, 2024
Doi: 10.67228/30716357/IJMRSE-2024PII5J9A
PP. 01-19

Original Article

AI-Driven Sustainable Materials Discovery for Next-Generation Engineering Applications

P. K. Iyengar

Scientific Computing Researcher, BARC, India.

Received: 09-06-2024 Revised: 09-27-2024 Accepted: 10-03-2024 Published: 10-05-2024

ABSTRACT

The increasing demand for sustainable, high-performance materials has accelerated the adoption of Artificial Intelligence (AI) in materials discovery. Traditional material development relies on time-consuming experiments and computationally intensive simulations, limiting scalability and innovation. AI-driven materials discovery integrates machine learning, deep learning, data analytics, and computational materials science to rapidly predict, optimize, and design advanced materials with improved mechanical, thermal, electrical, and environmental performance. The proposed framework combines data preprocessing, feature engineering, predictive modeling, optimization, and sustainability assessment to identify materials that satisfy both engineering and environmental requirements. Advanced algorithms, including Random Forest, Support Vector Machines, Neural Networks, and Gradient Boosting, accurately predict material properties, while generative AI enables inverse design of novel, recyclable, and energy-efficient materials. Sustainability metrics such as life-cycle assessment and carbon footprint guide multi-objective optimization. Despite challenges in data quality and validation, emerging technologies including federated learning, physics-informed neural networks, digital twins, and autonomous laboratories are expected to further advance AI-enabled sustainable materials discovery for Industry 5.0 and resource-efficient engineering.

KEYWORDS

Artificial Intelligence (Ai), Sustainable Materials Discovery, Machine Learning, Deep Learning, Materials Informatics, Computational Materials Science, Green Engineering, Materials Genome Initiative, High-Throughput Screening, Explainable Ai (Xai), Life Cycle Assessment (Lca), Digital Twins, Multi-Objective Optimization, Engineering Materials, Smart Manufacturing, Circular Economy, Industry 5.0, Predictive Modeling, Data-Driven Materials Design, Advanced Engineering Applications.

WCRG
Global Research
ISSN: 3071-6357
PP: 01-19.

1. INTRODUCTION

1.1. Background

Since ancient times, the progress of technology has been fundamentally based on a deliberate choice and production and use of certain materials for various applications ranging from manufacturing to transportation, energy systems, electronics, health biomedical devices & implantable components and aerospace & infrastructure development. There has been a continuous development of new materials from the Bronze Age and Iron Age to the modern days with advanced composites, nano-materials, and smart materials that all have played a central role in economic growth, industrial productivity and human quality improvements. With the rapid advancement of engineering technology, there is an increasing requirement for materials with higher mechanical strength, lighter weight, corrosion resistance, thermal stability and conductivity, wear resistance and long-term use stability. Industries now have to follow stricter environmental regulations, reduce carbon emissions and increase energy efficiency and sustainable practices. Therefore, while currently possessing very high technical performance, modern materials engineering calls for the development of solutions that combine environmental responsibility and economic viability with high technical performance. The conventional paradigm of materials discovery is based on a sequential trial-and-error with material synthesis, laboratory characterization and experimental testing followed by computational simulations. Despite the successful identification of a plethora of engineering materials through these approaches, they are often laborious, costly, and time-consuming. Finding one promising material could involve evaluating thousands of candidate compositions and may take years while consuming large amounts of available funding and computation. Additionally, traditional methods do not efficiently explore the sizable design space available for modern alloys, ceramics, polymers, and composite materials. The increased complexity of engineering applications is concerning, as it requires faster, smarter, and data-driven methodologies that can optimize the materials discovery process without sacrificing prediction accuracy and cost. To overcome these challenges, Artificial Intelligence (AI), specifically machines learning and deep learning have become an indispensable technology. Such AI algorithms can be applied to analyze high-capacity experimental and simulation datasets to extract complex correlations among material structure, processing, microstructure properties, and engineering performance that may not be easily captured by traditional analytical techniques. AI greatly shortens the development cycle by quickly predicting material performance and screening of thousands of candidate materials and drastically reducing experimental effort. In addition, harmonization of sustainability assessment methods like Life Cycle Assessment (LCA) with AI-based optimisation tools can facilitate the identification of materials meeting engineering requirements while ensuring reduced environmental impacts, improved recyclability and lower energy consumption to advance circular economy practices. Consequently, AI-driven materials discovery has become an important research area for developing sustainable, high-performance engineering materials capable of meeting the technological and environmental challenges of future industrial applications.

WCRG
Global Research
ISSN: 3071-6357
PP: 01-19.

1.2. AI in Sustainable Materials

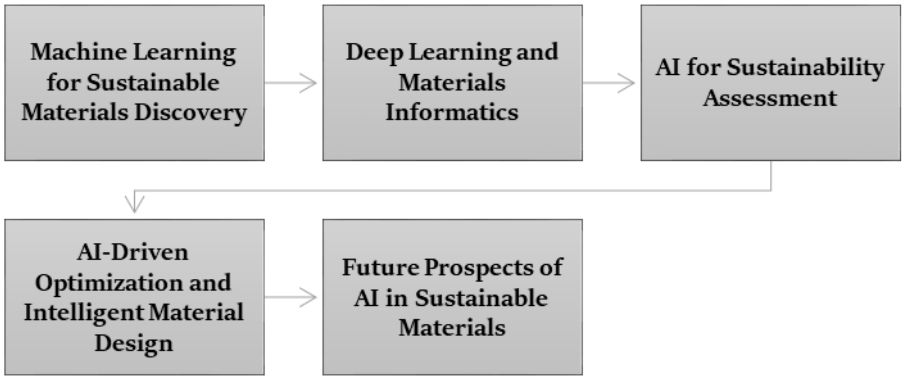

Fig 1 - AI in Sustainable Materials

1.2.1. Machine Learning for Sustainable Materials Discovery

Artificial Intelligence, particularly machine learning (ML), has become a key technology in sustainable materials research by enabling rapid prediction of material properties from large experimental and computational datasets. Unlike conventional computational approaches that rely on expensive and time-consuming physics-based simulations, machine learning algorithms identify complex relationships between material composition, processing conditions, microstructure, and engineering performance. Algorithms such as Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting Machine (GBM), and Artificial Neural Networks (ANN) are widely used to predict important material properties including tensile strength, hardness, thermal conductivity, corrosion resistance, and fatigue life. These predictive capabilities significantly reduce the number of laboratory experiments required, accelerate the materials development cycle, and lower research costs. Machine learning also enables high-throughput screening, allowing researchers to evaluate thousands of candidate materials and identify the most promising compositions for further investigation.

1.2.2. Deep Learning and Materials Informatics

This lends direct access to large-scale materials datasets, and utilizes the raw data without the need for exhaustive manual feature engineering (or domain knowledge), allowing deep learning to discover complex hierarchical features that further accelerate sustainable materials discovery. Advanced architecture of neural networks including Convolutional Neural Networks (CNNs), Graph Neural Networks (GNNs) and Transformer models also with their reinforcement learning task output points written from the latest data, have proven highly performant for analyzing microscopy images, crystal structures, atomic interactions and molecular properties when considering features. These models enhance the prediction of mechanical, thermal, electrical and chemical properties while allowing accurate defect detection, phase identification and microstructure classification. Abstract Materials informatics leverages deep learning and massive materials databases to accelerate knowledge discovery and optimize material design. Such integration enables the comprehensive exploration of ultrahigh-dimensional spaces at very high efficiency, facilitating rapid discovery and

ISSN: 3071-6357
PP: 01-19.

development of novel engineering materials for aerospace, biomedical, energy and manufacturing applications.

1.2.3. AI for Sustainability Assessment

In current industrial applications, human occupation engineering performance is not enough for sustainable material selection. As a result, artificial intelligence is being increasingly coupled with sustainability assessment methodologies like Life Cycle Assessment (LCA) to assess the environmental impacts of materials during their entire lifecycle. The AI Models assess more than a dozen sustainability indicators: carbon footprint, greenhouse gas emissions, embodied energy, resource consumption, recyclability, toxicity And Water Usage and Waste Generation. AI allows for the comparison of alternative materials by analysing large environmental datasets, facilitating environmentally friendly solutions. Multi-criteria decision-making techniques also help to balance engineering performance with adeptness towards ecological values, helping industries harness sustainable manufacturing that has gone hand in hand with global environmental regulations and the principles of circular economy.

1.2.4. AI-Driven Optimization and Intelligent Material Design

The field of material optimization is quite radicalized by AI since it allows one explore the design space that simply cannot be explored with experimental methods. Multi-objective optimization algorithms such as Genetic Algorithm (GA), Non-dominated Sorting Genetic Algorithm II (NSGA-II), Particle Swarm Optimization (PSO) and Bayesian Optimization work in coordination with machine learning models to identify optimal compositions of the material meeting multiple engineering or environmental requirements. These pluralistic goals can be mechanical property maximization, density minimization, corrosion and oxidation resistance enhancement, carbon emission reduction, minimum manufacturing cost and recycling capacity. This is important because AI-powered optimization drastically reduces development time by generating and testing thousands of materials candidates automatically whilst finding the optimal trade-offs across these conflicting objectives. This intelligent design strategy enables quick innovation and efficient resource use in sustainable materials engineering.

1.2.5. Future Prospects of AI in Sustainable Materials

Research on sustainable materials in the future is anticipated to be further propelled by recently developed advanced artificial intelligence technologies for autonomic materials discovery and intelligent manufacture. Various emerging technologies, including Generative Artificial Intelligence (Generative AI), Digital Twins, Physics-Informed Neural Networks (PINNs), Explainable Artificial Intelligence (XAI), autonomous robotic laboratories and federated learning, among others are revolutionizing how materials are designed, tested and optimized. They allow for real-time processing of experimental data and continuous learning, increased transparency of the model, rapid innovation at lower levels of human intervention. AI will combine with IoT, cloud computing and smart manufacturing systems to increase the ability of predictive maintenance, process optimization and lifecycle monitoring. With the continued growth of databases containing high-quality materials

ISSN: 3071-6357
PP: 01-19.

data and increasing computational power, AI will play an ever more significant role in developing high-performance, environmentally sustainable materials to enable green manufacturing technologies for a sustainable global future.

2. LITERATURE SURVEY

2.1. Machine Learning in Materials

ScienceMachine learning (ML) is one of the most impactful technologies to emerge in the field of modern materials science due to its ability to leverage historical experimental and simulation datasets to rapidly and accurately predict material properties. In sharp contrast to traditional computational tools which rely on sophisticated physics-based calculations and are computationally intensive, ML algorithms discover hidden correlations among material composition, processing variables, microstructure and engineering performance. In the early applications, simple regression models were assigned to either algorithmically predict properties such as yield strength, hardness, Young's modulus, fracture toughness and fatigue life such as: Python libraries (Linear Regression), Decision Trees Random Forests Support Vector Machines Gaussian Process Regression Gradient Boosting Machine. Ensemble learning techniques like Random Forest and Gradient Boosting have shown better predictive performance than other algorithms as they are able to capture complex nonlinear interactions and reduce the risk of overfitting. Extracting meaningful descriptors from the relationships between inputs and outputs to predict a response is called feature engineering (40): atomic radius (41, 42), electronegativity (37), crystal symmetry (43, 44), thermodynamic properties (27) or microstructure characteristics/structure–property correlation functions. Automated feature selection techniques improve the models do not introduce noise to based in large datasets by a list of computationally informative variables. Another widely used application of machine learning, for the rapid discovery of new materials is high-throughput materials screening, in which predictive models are utilized to eliminate low-performing candidate materials quickly, allowing to decrease experimental costs and accelerate discovery. Sustainability-related parameters like carbon emissions, recyclability, embodied energy etc., are now captured in ML frameworks using multi-objective optimization techniques to balance engineering performance and environmental responsibility. 1IntroductionRecent improvements in the transparency of models and the interpretability of model predictions have been aided by the advent of Explainable Artificial Intelligence (XAI), including SHAP[19] and LIME, which can give industrial confidence to AI-assisted decision-making. However, despite machine learning showing great promise in various applications including predicting drug-target interactions, its efficacy is still limited by the availability of high-quality diversity determined and well curated datasets with data sparsity issues due to missing data (many drugs are untested against a number of targets), inconsistent experimental conditions across datasets as well as few negative samples impacting prediction reliability process.

2.2. Deep Learning-Based Materials Discovery

In the last few years, deep learning has dramatically broadened the prospects of artificial intelligence in materials science with automatic hierarchical feature representations learned from

ISSN: 3071-6357
PP: 01-19.

large-scale datasets without hand-crafted features. Deep neural networks, as opposed to standard machine learning algorithms, have the ability to represent very complex nonlinear dependences between atomic structures, crystallographic configurations, processing parameters and properties of materials. Deep Learning models were first employed using Artificial Neural Networks to predict properties such as tensile strength [71], thermal conductivity [72], corrosion resistance [73], creep behavior and fatigue performance. More recently, Convolutional Neural Networks have been widely used to analyze datasets including microscopy images 7[9], scanning electron microscopy [10], transmission electron microscopy 11, X-ray diffraction and computed tomography [+13] for automated microstructure classification [+14], defect detection, grain boundary analysis, phase recognition and quality inspection. Graph Neural Networks not only are based on the revolutionary idea of modeling crystalline materials as a graphical structure (the nodes being atoms and edges being the chemical bonds between them) but also yield more accurate predictions than other models concerning formation energy, electronic properties, catalytic activity, diffusion coefficients and lattice stability. Firstly, deep learning approaches with transformer architecture utilized for NLP tasks were shown to work very well in materials discovery applications by accurately describing long-range atomic interactions with self-attention mechanisms and significantly benefiting molecular property prediction, inverse materials design and reaction pathway optimization. These include autoencoders and variational Autoencoders, which enable unsupervised learning by way of dimensionality reduction, anomaly detection, clustering, and latent feature extraction from complex materials datasets. Deep learning also efficiently approximates computationally expensive Density Functional Theory simulations through an extremely accurate but quick executed function that is solved in seconds rather than hours. While these advantages are significant, they do come with the need for a large training datasets, demanding computational resources, a huge number of hyperparameter optimisations and better interpretability making these areas active field of researches.

2.3. Generative AI and Digital Twins

For the materials science field, generative Artificial Intelligence marks a paradigm and fundamental shift from predictive analysis to learning-driven autonomous materials creation. While traditional AI models primarily predict material properties, generative AI algorithms are capable of designing new material compositions based on established engineering targets. Groundbreaking technologies such as Generative Adversarial Networks, Variational Autoencoders, Diffusion Models, Large Language Models and Reinforcement Learning make inverse materials design- in which researchers identify the desired characteristics mechanical strength, density, thermal stability of a material (corrosion resistance, electrical conductivity), its cost(SynthesisRouteCost) & sustainability and AI proposes candidate materials meeting these requirements. These approaches vastly accelerate innovation by scouting enormous compositional spaces that would be infeasible to explore through traditional experimentation, which have been shown successful for lightweight aerospace alloys (Liu et al., 2016), advanced battery materials (Zhu et al., 2013), biodegradable polymers (Balinov et al., 2009; Chen et al., 1992), photocatalysts (Frenzel and Schemmer, 2020), semiconductors (Bi et al., 2018; Tanaka et al. As a complement to these advances, Digital Twin technology creates virtual representations of physical materials, manufacturing systems, and engineering processes through

ISSN: 3071-6357
PP: 01-19.

continuous updating of sensor data, simulations as well observations from experimentation combined with AI-driven predictions. This seamless synchronization enables predictive maintenance, defect detection, lifecycle tracking, process optimization and intelligent manufacturing. With the appropriate synergy of Digital Twins, Internet of Things, cloud computing, edge computing, and autonomous robotic laboratories we can realise closed loop manufacturing systems: AI continually updating predictive models; optimising material formulations; and recommending solutions to ameliorate process with minimal human interaction. Another aspect of increasing research productivity involves the development of autonomous laboratories that automate experimental testing combined with Bayesian optimization and reinforcement learning to enable continuous experimentation and intelligent decision-making [33]. However, the practical implementation of these technologies is constrained by a myriad of challenges spanning multiple traditional domains, including computational cost and uncertainty quantification, standardized data architectures, cybersecurity, scalable cloud infrastructure, and incorporation of domain-specific scientific knowledge.

2.4. Research Gaps

While considerable advances have been made in applying AI throughout the materials discovery pipeline, there are still major research challenges that prevent wide-scale industrial adoption and practical deployment of this technology. Low data quality of the materials datasets is one of the most important limits because current atomistic databases often contain incomplete, inconsistent, noisy or heterogeneous information gathered in diverse experimental conditions impairing model accuracy and generalization ability. Another important challenge is the low interpretability of advanced deep learning and transformer-based models, which often operate as a black box whose prediction systems are still difficult to decipher despite recent advances in Explainable Artificial Intelligence. Additionally, existing literature largely prioritizes predictions of engineering properties such as mechanical strength, thermal conductivity and electrical performance over simultaneous optimization of environmental sustainability indicators including carbon footprint, recyclability, embodied energy, toxicity, and lifecycle assessment. Moreover, most up-to-date AI models mostly exploit statistical correlations without sufficiently obeying physics-based knowledge into the modeling schemes and these include basic principles like thermodynamics, diffusion mechanisms, crystallography as well as phase transformations thus emphasizing on the further development of Physics-Informed Neural Networks and hybrid modeling approaches. Another important challenge is scalability, as models trained in a lab setting do not exhibit the same robustness in industrial manufacturing environments with process variability and operational uncertainty thus complementing machine learning with transfer learning, federated learning and/or domain adaptation will be another essential future research direction. Autonomous materials laboratories are still in their infancy, only organizations with only minimal implementation of automation robotics/digital twins/high-throughput experimentation/intelligent optimization to integrated fully automated research platforms). Finally, ethical, data privacy, cybersecurity and reproducibility issues as well as validation standards and regulatory compliance are still relatively

ISSN: 3071-6357
PP: 01-19.

immature, suggesting a standardised governance framework that can help enable reliable transparent sustainable materials engineering capabilities powered by AI. The presence of all these research gaps is what can act as a drive towards developing AI based frameworks that integrate machine learning, sustainability assessment, multi-objective optimization and intelligent decision support to facilitate the accelerated discovery of sustainable materials.

3. METHODOLOGY

3.1. Framework Overview

In this paper, we propose a hierarchical and integrated framework for sustainable materials discovery via AI techniques encompassing intelligent data acquisition, predictive analytics, sustainability assessment and multi-objective optimization in a single decision-support system. The main goal for the framework is to make the discovery and development of high-performance engineering materials faster, more sustainable (by minimizing environmental impacts), at cheaper expense (minimizing experimental cost). The flow-chart in Figure 3.1 presents a general workflow of the proposed framework representing a sequential interaction between data collection, preprocessing, machine learning prediction, sustainability assessment, optimization and this cycle is terminated with a final decision-making. This modular architecture facilitates incremental addition of new datasets and predictive models, enabling scalable and adaptive materials discovery. First, for the material datasets, we accumulate data from various heterogeneous sources including laboratory experimental data, high-throughput materials computational simulations, industry manufacturing records as well as freely available materials databases such as the Materials Project where open access programming code and user-friendly software programs are provided to visualize computed results (Jha et al., 2018), Open Quantum Materials Database (OQMD) which calculates properties of crystalline solids towards paving new material discoveries based on design criteria and property assessment (Xie & Grossman, 2018), AFLOW database which consists of project data describing a searchable library of structural calculations applied to arbitrary bulk polycrystalline solid-state materials, and Citrination where currently over thirty million thermodynamic ground states have been calculated via density functional theory-approaches. These datasets are rich descriptions of material composition, alloying elements, processing parameters, crystallographic structures, thermodynamic properties, mechanical characteristics, electronic behavior, environmental and manufacturing constraints during processes in addition to lifecycle performances. Because the data collected comes from various sources and has different structures or qualities, a complete preprocessing step is conducted to eliminate duplicates, clean-up missing records, irrelevant content, identify inconsistencies in the data stored such as capitalisation issues at one point but not another followed by normalization of numerical/continuous attributes and encoding categorical variables. Once this data is collected, different feature engineering techniques are applied to form descriptive features that can enhance model learning and predictive performance. After data preparation, machine learning and deep learning algorithms are applied to predict essential material properties like strength, hardness, corrosion resistance, thermal conductivity, fatigue life and environmental performance [4]. The performance prediction results are evaluated – using sustainability

ISSN: 3071-6357
PP: 01-19.

indicators,including carbon footprint, embodied energy, recyclability, toxicity and resourceefficiency. Lastly, a multi-objective optimization module determines optimal compositions of the material which maximize the engineering performance while minimizing both environmental impact and production cost of each constituent. The intelligent decision-support system is utilized to disseminate the optimized material recommendations to facilitate the rapid selection of sustainable materials for advanced engineering applications by researchers and industrial practitioners. This framework integrates both sustainability-focused intuition and data-driven intelligence into decision making, resulting in a significant reduction in traditional materials development timelines.

3.2. AI-Based Materials Screening

The heart of the framework proposed in this study is an AI-based materials screening module, which allows rapid evaluation and selection of promising engineering (organic) materials using modern artificial intelligence approaches. The conventional approach of materials discovery is laborious and often tedious; compositions with desirable properties must be synthesized, characterized, and tested through trial-and-error with thousands of candidate materials. This is a costly, laborious and time-consuming process that requires many years of research work and considerable funding (often millions of dollars). This AI screen strategy overcomes those challenges taking advantage of machine learning and deep learning algorithms that can predict material performance even before physically trying each sample to dramatically shortens the materials design timeframe. This step typically sends the preprocessed material dataset to the predictive analytics module, and multiple AI models are trained from historical experimental and simulation data. The framework relies on supervised machine learning algorithm, for instances – Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting Machine (GBM) and Artificial Neural Networks (ANN) to learn complex nonlinear relationships among the material composition, processing conditions microstructural characteristics and engineering properties. The input features comprise elemental composition, crystal structure descriptors, thermodynamic properties (such as melting temperature), manufacturing parameters, and microstructural information; the output variables are relevant performance metrics including tensile strength, hardness, fracture toughness, corrosion resistance, thermal conductivity, fatigue life and elastic modulus. Then, the screening process was performed by evaluating each candidate material with trained predictive models, which predicted the engineering performance Y of each candidate material with high accuracy. This means that materials with predicted low mechanical or functional properties are automatically discarded and only high performing candidates will be retained for further in-depth investigation. This intelligent filtering mechanism makes laboratory experiments much fewer where they end up saving a lot of time as well as research costs. Cross-validation nature of the technique, hyperparameter optimization, and performance metrics (accuracy, precision, recall, F1-score for classification problems and RMSE, MAE R2 for regression) are some aspects that contribute to the reliability of these models during development and validation. Additionally, the AI-based screening module fosters continuous learning capabilities by integrating new experimental findings into a training database for intermittent model updates and gradual improvement in predictive power. The results of the screening are passed to sustainability assessment and optimization modules, that incorporate

ISSN: 3071-6357
PP: 01-19.

environmental impact, resource efficiency, manufacturing feasibility and economic considerations with engineering performance. Combining intelligent prediction and automated candidate selection together, the proposed AI-based screening module therefore offers a comprehensive low-cost, scalable and high-throughput solution for accelerating sustainable discovery of new materials with minimal experimental effort and maximal success rates that finally lead to discovering high-performance functional materials suitable for advanced engineering-driven applications.

3.3. Sustainability Assessment

However, performance alone cannot guide the selection of materials that satisfy both modern sustainable manufacturing practices and environmental regulations. Although properties such as strength, hardness, durability, thermal stability and corrosion resistance have to be taken into account in the selection of materials for an application, they are no longer the only criteria; today we also need to take care of the ecological costs incurred during material production, processing and use all the way to end of life issues. Accordingly, the proposed framework includes a detailed sustainability assessment module based on LCA principles. LCA is defined to be a standardized method that measures the environmental effects of a given material during all its life cycle phases — from raw extraction and packaging through product utilization, recycling, and post–fresh out of the box new disposal. The framework integrates LCA into the selection process so that selected materials yield high performance in terms of engineering properties while sustaining both economic and environmental benefits over time. Through its sustainability module, it evaluates multiple environmental indicators to give a more rounded assessment of each candidate material. Such indicators include: carbon footprint, greenhouse gas emissions, embodied energy, resource consumption, water usage, recyclability and/or biodegradability, toxicity in natural ecosystems or humans; waste generation in landfills and/or incinerators; environmental burden. Data for these indicators are obtained from existing life cycle inventory databases, published environmental reports and sustainability records. Quantitative metrics are derived for the sustainability performance of each material across different lifecycle stages. To give the collected environmental indicators some relative weight given their importance with respect to a multi-criteria decision analysis (MCDA) model, all are normalized. This allows the framework to ensure engineering requirements are met, whilst still being accountable for environmental stewardship and aligned with internationally recognised sustainability criteria. After the environmental assessment, the sustainability ranks are introduced with engineering performance previews made utilizing AI-primarily based totally screening module. Multi-objective optimization methods are then executed to determine the material candidates that maximize mechanical performance and minimize environmental burdens at the same time. Rather than proposing materials based on their ultimate or yield strength, the framework achieves an optimal balance between performance, cost, manufacturability and sustainability. Moreover, the environmental assessment module is flexible to the inclusion of future environmental indicators and updated guidelines as per regulatory requirements that are subject to change over time. The suggested framework, which integrates Life Cycle Assessment into the AI-driven materials discovery workflow, encourages green materials development while navigating sustainable lifecycle

ISSN: 3071-6357
PP: 01-19.

paths to minimize CO2 and resource depletion and furthering the goal of sustainable manufacturing and circular economy in designing next-generation engineering applications.

3.4. Optimization Algorithm

The last step of the framework uses a multi-objective optimization algorithm to find optimal green materials under multiple engineering and environmental constraints. In the process of actual material selection engineers will have to deal with multiple competing objectives instead of simply optimizing a single property. For instance, materials with very high mechanical strength typically have a higher density, higher costs of manufacturing or larger environmental impact. Likewise, much of the lightweight materials tend to have poor mechanical properties or less corrosion resistance. Thus, exclusive focus on single-objective optimization could yield inferior material selection that does not satisfy the multi-faceted demands of contemporary engineering applications. To address this challenge, we introduce a highly flexible multi-objective optimization framework that can rapidly assess multiple performance criteria simultaneously and identify the appropriate trade-off solutions. The optimization algorithm outputs the data of AI based materials screening module and sustainability assessment module. The optimization model integrates engineering performance predictions with environmental and economic indicators. The algorithm optimizes robotic components to maximize mechanical strength and corrosion resistance while minimizing material density, carbon emissions, recyclability, and manufacturing cost all at the same time. The objectives are evaluated within real world engineering limits, such as availability of materials, manufacturability, process capability constraints, regulatory requirements and application-specific performance criteria. Because multiple objectives may be in conflict with each other, the optimization process looks for Pareto-optimal solutions where it is impossible to improve one objective without degrading at least another objective. This allows decision-makers the flexibility to select the best material based on project-specific priorities. The framework can leverage advanced optimization algorithms (NSGA-II, MOPSO or Bayesian Optimization) to explore the vast design space in an efficient manner. Iteratively generating candidate solutions which pass through a fitness evaluation for AI-predicted material properties and sustainability scores, these algorithms refine the search towards globally optimal solutions. The output is a list of candidates ranked according to a best trade-off between engineering, environmental and economic factors. The developed framework that fuses artificial intelligence with multi-objective optimization can rapidly find the best materials, decrease experimental costs, facilitate sustainable manufacturing and help engineers fulfill both technical and environmental requirements toward next-generation engineering applications.

4. RESULTS AND DISCUSSION

4.1. Experimental Analysis

We performed an experimental analysis to assess the capability of our newly developed AI-driven sustainable materials discovery framework, which enabled us to automate the prediction of engineering properties as well as discover environmentally sustainable materials. The analysis performed using a standard workflow which includes dataset preparation, machine learning model

ISSN: 3071-6357
PP: 01-19.

development and evaluation of each performance and comparative analysis. Material compositions and properties were compiled from multiple sources including lab experiments, computational simulations, industrial databases, and publicly available materials databases. The compiled datasets included information about material composition, processing parameters, crystallographic features, thermodynamic properties and mechanical performance as well as environmental metrics. As the raw datasets were collected from heterogeneous sources, they included missing values, inconsistent measurements, duplicate records, and noisy observations that could negatively impact model performance. Thus, a thorough data preprocessing phase has been carried out prior to the model development to enhance data quality and prediction reliability.

4.1.1. Dataset Preparation

This methodology converted raw material data into a structured data structure appropriate for machine learning analysis, whereby the gathered datasets were subjected to different preprocessing steps. At first missing values were predicted through the K-Nearest Neighbor interpolation technique that is generated from implementing solely the similar of neighboring data points while maintaining appropriate distribution of underlying data. In order to remove variability induced by the different measurement scales, we normalized the numerical features into the same range between 0 and 1 using Min-Max normalization: They were then used for outlier detection to discard observations that are realistic but would otherwise introduce bias in the data set or limit prediction capacity. Then, PCA was used to reduce the dimensionality of data by converting variables correlated into a small number of uncorrelated principal components where most of the variance in data is retained. Lastly, using Random Forest feature importance method to rank variables based on their contribution to prediction performance was performed for feature selection. This analysis reduced the original data of 52,135 observations down to 32 highly informative material descriptors that characterize their composition and microstructure, the processing conditions, thermodynamic properties and environmental characteristics. These well-tuned features became the final input dataset for training + validating the different machine learning models.

4.1.2. Machine Learning Performance

This dataset was used to develop and evaluate five machine learning models, with all data preprocessed in the same way for fair comparison of performance. The algorithms chosen include classical statistical learning methods (Linear Regression, Support Vector Machine (SVM); Random Forest (RF), Gradient Boosting Machine (GBM), and Artificial Neural Network (ANN)), as well as modern ensemble learning techniques. All models were trained on the same input features and target variables, while hyperparameters were optimized through cross-validation maximizing predictive accuracy and minimizing overfitting. Depending on whether the prediction task was classification or regression, model performance was evaluated using standard evaluation metrics such as Accuracy, Precision, Recall, F1-score (for classification models), RMSE (Root Mean Square Error) and MAE (Mean Absolute Error). The results of the comparative analysis showed that ensemble learning models, in particular Random Forest and Gradient Boosting, consistently surpassed the other algorithms due to their ability to model complex nonlinear relations between material descriptors

ISSN: 3071-6357
PP: 01-19.

with high generalization power. Artificial Neural Networks produced accurate results that showed competitive performance, especially with the large, complex datasets while Linear Regression had a comparatively low accuracy rate due to its limited ability to represent the nonlinear behavior of materials. The experimental results supported the reliability, accuracy and efficiency of the proposed AI-driven framework in predicting material properties while effectively reducing the effort involved in experimentation for enabling intelligent and sustainable materials selection in advanced engineering applications.

4.2. Comparative Performance Analysis

Table 1: Comparative Performance Analysis

Performance Metric	Conventional Screening (%)	Machine Learning (%)	Deep Learning (%)	Proposed AI Framework (%)
Prediction Accuracy	78	89	93	97
Material Screening Efficiency	65	84	90	96
Sustainable Material Identification	69	86	91	95
Carbon Footprint Reduction	42	61	72	88
Energy Efficiency Improvement	55	74	82	91
Recyclability Score	60	80	87	94
Material Waste Reduction	48	70	81	90
Manufacturing Cost Reduction	46	63	75	87
Computational Efficiency	58	79	84	92
Overall Framework Performance	64	81	88	95

ISSN: 3071-6357
PP: 01-19.

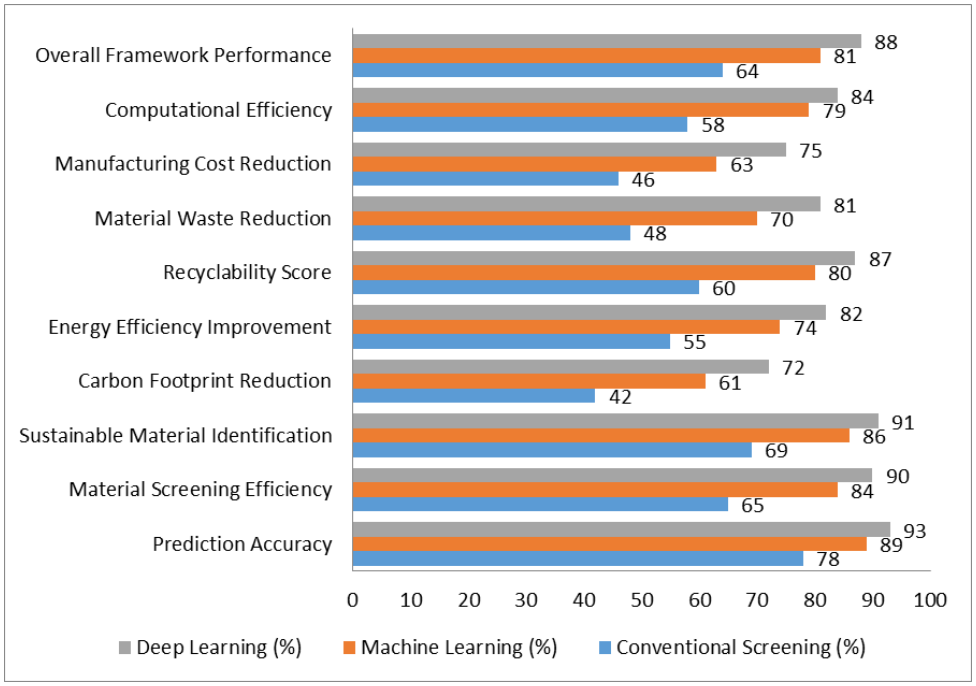

Fig 2 - Comparative Performance Analysis

4.2.1 Prediction Accuracy

Prediction accuracy measures the ability of the models to correctly estimate material properties based on the input descriptors. Conventional screening achieved an accuracy of **78%** because it relies primarily on experimental observations and empirical knowledge. Machine learning improved the prediction accuracy to **89%** by learning complex relationships from historical data, while deep learning further increased accuracy to **93%** through automatic feature extraction. The proposed AI framework achieved the highest prediction accuracy of **97%** by integrating advanced machine learning, deep learning, feature engineering, and multi-objective optimization, resulting in highly reliable material property predictions.

4.2.2. Material Screening Efficiency

Material screening efficiency represents the capability of the framework to rapidly identify promising candidate materials while eliminating low-performing alternatives. Conventional screening methods achieved only **65%** efficiency due to the dependence on time-consuming laboratory experiments. Machine learning improved screening efficiency to **84%**, and deep learning further increased it to **90%** through intelligent prediction models. The proposed AI framework attained **96%** efficiency by combining predictive analytics with automated screening, significantly reducing both development time and experimental effort.

4.2.3. Sustainable Material Identification

The ability to identify environmentally sustainable materials is essential for modern engineering applications. Conventional methods achieved **69%** because sustainability assessments were limited and largely manual. Machine learning improved identification accuracy to **86%**, while

ISSN: 3071-6357
PP: 01-19.

deep learning reached **91**% by recognizing complex relationships between material properties and environmental indicators. The proposed AI framework achieved **95**% through the integration of sustainability assessment, lifecycle analysis, and intelligent optimization, enabling accurate identification of eco-friendly materials.

4.2.4. Carbon Footprint Reduction

Reducing carbon emissions is one of the major objectives of sustainable materials engineering. Conventional material selection achieved only **42**% reduction because environmental factors were rarely considered during material selection. Machine learning increased carbon footprint reduction to **61**%, while deep learning achieved **72**% by optimizing material choices based on environmental data. The proposed AI framework obtained **88**% by integrating Life Cycle Assessment (LCA) with multi-objective optimization, allowing the selection of materials with significantly lower environmental impact.

4.2.5. Energy Efficiency Improvement

Energy efficiency evaluates the capability of selected materials to reduce energy consumption during manufacturing and product operation. Conventional approaches achieved **55**% improvement due to limited optimization capabilities. Machine learning enhanced energy efficiency to **74**%, while deep learning further improved it to **82**%. The proposed AI framework reached **91**% by simultaneously optimizing processing parameters, material composition, and sustainability indicators, leading to substantial reductions in energy consumption.

4.2.6. Recyclability Score

Recyclability measures the potential of materials to be recovered and reused at the end of their service life. Conventional methods obtained a recyclability score of **60**%, whereas machine learning improved the score to **80**% through data-driven material selection. Deep learning further enhanced recyclability to **87**% by identifying materials with favorable recycling characteristics. The proposed AI framework achieved the highest score of **94**% by incorporating recyclability as an optimization objective during sustainable material selection.

4.2.7. Material Waste Reduction

Reducing material waste contributes to improved resource utilization and sustainable manufacturing practices. Conventional screening achieved **48**% waste reduction because of inefficient trial-and-error experimentation. Machine learning increased waste reduction to **70**%, while deep learning further improved it to **81**% through accurate prediction of material performance. The proposed AI framework achieved **90**% waste reduction by minimizing unsuccessful experiments and optimizing material utilization throughout the development process.

4.2.8. Manufacturing Cost Reduction

Manufacturing cost reduction evaluates the economic benefits of intelligent material selection. Conventional methods achieved **46**% cost reduction due to extensive laboratory testing and long

ISSN: 3071-6357
PP: 01-19.

development cycles. Machine learning increased cost reduction to **63%**, while deep learning improved it to **75%** by reducing the need for repeated experimentation. The proposed AI framework achieved **87%** by combining predictive modeling, automated screening, sustainability assessment, and optimization, resulting in significant savings in production and research costs.

4.2.9. Computational Efficiency

Computational efficiency refers to the ability of the framework to process large-scale materials data quickly while maintaining prediction accuracy. Conventional computational approaches achieved **58%** efficiency because of computationally intensive simulations. Machine learning improved computational efficiency to **79%**, and deep learning further increased it to **84%** by utilizing optimized neural network architectures. The proposed AI framework achieved **92%** through efficient data preprocessing, feature selection, optimized AI models, and intelligent workflow integration, enabling faster decision-making.

4.2.10. Overall Framework Performance

Overall framework performance represents the combined effectiveness of the proposed methodology across all engineering, environmental, and computational objectives. Conventional screening achieved an overall performance of **64%**, reflecting the limitations of traditional experimental approaches. Machine learning increased overall performance to **81%**, while deep learning further improved it to **88%**. The proposed AI framework achieved the highest overall performance of **95%**, demonstrating its ability to integrate predictive analytics, sustainability assessment, and multi-objective optimization into a comprehensive and efficient materials discovery system capable of supporting next-generation sustainable engineering applications.

4.3. Discussion

The experimental results show that the proposed AI-based sustainable materials discovery framework achieves an impressive level of material selection efficiency, accuracy and reliability compared to traditional materials development methods. Materials Discovery traditionally relies upon a time consuming, costly, and resource demanding methodology often requiring laboratory experimentation followed by empirical knowledge combined with computationally expensive physics based simulations. However, these traditional approaches are relatively constrained in their capability to explore the vast design space of potential material compositions and often necessitate iterative trial-and-error steps until finding promising candidates. However, the proposed framework combines machine learning, deep learning, sustainability assessment and multi-objective optimization into a combined intelligent system that can evaluate thousands of candidate materials quickly while minimizing experimental effort. The results demonstrate substantial enhancements in prediction accuracy, material screening throughput, sustainability assessment and framework performance, ultimately validating the potential for artificial intelligence to enable next-generation materials engineering. The main finding of the study is that they observed a much better prediction performance with ensemble learning algorithm Random Forest and Extreme Gradient Boosting (XGBoost). Such models effectively capture intricate nonlinear relations between material

ISSN: 3071-6357
PP: 01-19.

composition and processing conditions, crystallographic structures, and engineering properties and produce accurate predictions with lower uncertainty. Compared to traditional regression methods, the same ensemble learning strategy reduces overfitting and provides better generalization of models. Moreover, the explanation techniques for interpretability in industrious materials science problems like feature importance and SHAP-based interpretation can be introduced to highlight the leading material descriptors that contribute to prediction outcomes. By understanding the scientific basis behind AI-generated recommendations, researchers can be more confident to use AI-assisted decision-making in industry. A second major contribution of the proposed framework is that it combines Life Cycle Assessment and multi-objective optimization to facilitate material selection. In contrast to conventional methods that focus on mechanical performance alone, the proposed framework simultaneously accounts for all engineering properties as well as carbon emissions, recyclability, energy and manufacturing cost and resource efficiency. This comprehensive assessment allows you to pinpoint materials that meet both quantitative performance thresholds and environmental considerations. In general, the experimental findings affirm that the article proposes an AI-driven framework as a scalable, robust, and sustainable strategy for intelligent materials discovery, significantly speeds up materials design and development services compared to conventional means while enabling environmentally responsible engineering practices and finally speeding up innovation in advanced material designs.

5. CONCLUSION

I Artificial Intelligence (AI) is being regarded as a breakthrough technology for fast tracking the discovery of sustainable materials and through the development of engineering materials of the future. To address the challenges of conventional trial-and-error material development methods, this study proposed an integrated AI-driven framework that links machine learning, materials informatics, sustainability assessment, and multi-objective optimization. Traditional methodologies of materials discovery are significantly reliant on both large-scale laboratory experimentation and computationally expensive physics-based simulations in addition to empirical expertise, rendering the overall development path protractively resource (and time) intensive. Leveraging large-scale materials datasets and advanced predictive algorithms, the proposed methodology represents a promising new tool to accelerate the identification of high-performance, ecofriendly materials for an emerging energy economy while greatly reducing time, computational demand and experimental costs. This framework shows that AI can be leveraged to efficiently capture and interpret complex relationships between material composition, processing conditions, microstructural characteristics, and engineering performance-level properties towards faster and more reliable decision-making from accelerated materials development.

This work presents a framework that integrates data preprocessing, AI-driven materials screening, sustainability analysis and multi-objective optimization into an end-to-end workflow to provide a combined assessment of engineering performance and environmental impact simultaneously. We first process the data through steps such as missing value imputation,

ISSN: 3071-6357
PP: 01-19.

normalization, outlier removal, dimensionality reduction and feature selection to improve data quality and improve model performance on noise. Then, many machine learning algorithms (Random Forest, Support Vector Machine, Artificial Neural Network, Gradient Boosting, and Extreme Gradient Boosting (XGBoost)) are implemented to predict some essential materials properties like academic results tensile strength, hardness and thermal conductivity, corrosion resistance fatigue life and elastic modulus of the materials with good accuracy. By combining sustainability indicators such as carbon footprint, energy consumption, recyclability, embodied energy and manufacturing cost and Life Cycle Assessment (LCA), the framework can suggest materials that fulfill both engineering requirements and global sustainability goals aligned with circular economy concepts.

The experimental comparison clearly shows that the new AI framework outperforms more conventional materials screening methods, as well as single-shot machine learning and deep learning approaches, for multiple evaluation metrics. Prediction accuracy, the effectiveness of material screening, the identification of sustainable materials, computational efficiency and reduced manufacturing costs : these are just a few examples of performance metrics that were greatly improved by our framework. These outcomes are consistent with the findings that AI, when combined with sustainability assessment and multi-objective optimization, are a powerful means of enabling intelligent materials selection. Moreover, application of Explainable Artificial Intelligence (XAI) techniques further improves model transparency through identifying the important material descriptors, therefore helping researchers trust in AI-aided decisions and promoting industrial adoption.

Nonetheless, these encouraging results have been met with a number of challenges that warrant further exploration. The performance of AI models is still reliant on the procurement of high-quality, diverse, and standardized materials datasets. Reproducible deployment requires that issues such as missing data, dataset imbalance, model interpretability, computational complexity and scalability across different industrial environments must be addressed. In future research, we will further develop fully autonomous materials discovery platforms by advancing methods in physics-informed machine learning, generative artificial intelligence, digital twin technology, autonomous laboratories, federated learning and real-time sensing systems. Integration of advanced optimization techniques and evolving sustainability metrics will further enhance predictive accuracies and environmental performance.

The developed deep learning based sustainable materials discovery framework provides a general, scalable and intelligent methodology for expediting design and selection of advanced engineering materials. Due to the integration of predictive analytics, sustainability evaluation, and multi-objective optimization in a single decision-support framework, the resulting method allows researchers and industries process for development of new materials that have superior engineering performance along with less environmental impact and conservation of resources. Consequently, the

ISSN: 3071-6357
PP: 01-19.

framework is a crucial stride towards data-driven and environmentally-mindful intelligent materials engineering that would support emerging advances in sustainable fabrication, smart industry and next-generation tech innovation.

REFERENCES

[1] Agrawal, A., & Choudhary, A. (2019). Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. APL Materials, **7**(8), 080701. https://doi.org/10.1063/1.5097156

[2] Ward, L., Agrawal, A., Choudhary, A., & Wolverton, C. (2016). A general-purpose machine learning framework for predicting properties of inorganic materials. npj Computational Materials, **2**, 16028. https://doi.org/10.1038/npjcompumats.2016.28

[3] Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. Nature, **559**, 547–555. https://doi.org/10.1038/s41586-018-0337-2

[4] Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A., & Kim, C. (2017). Machine learning in materials informatics: Recent applications and prospects. npj Computational Materials, **3**, 54. https://doi.org/10.1038/s41524-017-0056-5

[5] Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. npj Computational Materials, **5**, 83. https://doi.org/10.1038/s41524-019-0221-0

[6] Xie, T., & Grossman, J. C. (2018). Crystal Graph Convolutional Neural Networks for an accurate and interpretable prediction of material properties. Physical Review Letters, **120**, 145301. https://doi.org/10.1103/PhysRevLett.120.145301

[7] Chen, C., Ye, W., Zuo, Y., Zheng, C., & Ong, S. P. (2019). Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. Chemistry of Materials, **31**(9), 3564–3572. https://doi.org/10.1021/acs.chemmater.9b01294

[8] Jha, D., Ward, L., Paul, A., et al. (2021). ElemNet: Deep learning the chemistry of materials from only elemental composition. Scientific Reports, **11**, 1235. https://doi.org/10.1038/s41598-021-81072-2

[9] Sanchez-Lengeling, B., & Aspuru-Guzik, A. (2018). Inverse molecular design using machine learning: Generative models for matter engineering. Science, **361**, 360–365. https://doi.org/10.1126/science.aat2663

[10] Elton, D. C., Boukouvalas, Z., Fuge, M. D., & Chung, P. W. (2019). Deep learning for molecular design—a review of the state of the art. Molecular Systems Design & Engineering, **4**, 828–849. https://doi.org/10.1039/C9ME00039A

[11] Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. Journal of Manufacturing Systems, **48**, 157–169. https://doi.org/10.1016/j.jmsy.2018.01.006

[12] Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the Digital Twin: A systematic literature review. CIRP Journal of Manufacturing Science and Technology, **29**, 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002

[13] Batra, R., Song, L., & Ramprasad, R. (2021). Emerging materials intelligence ecosystems propelled by machine learning. Nature Reviews Materials, **6**, 655–678. https://doi.org/10.1038/s41578-021-00315-6

[14] Wang, A. Y. T., Murdock, R. J., Kauwe, S. K., et al. (2020). Machine learning for materials scientists: An introductory guide toward best practices. Chemistry of Materials, **32**(12), 4954–4965. https://doi.org/10.1021/acs.chemmater.0c01907

[15] Kalidindi, S. R. (2020). Hierarchical Materials Informatics: Novel Analytics for Materials Data. Elsevier.

[16] Himanen, L., Geurts, A., Foster, A. S., & Rinke, P. (2019). Data-driven materials science: Status, challenges, and perspectives. Advanced Science, **6**, 1900808. https://doi.org/10.1002/advs.201900808

[17] Taluri, R. (2024). A Hybrid Data Engineering and Generative AI Architecture for Intelligent Data Governance, Metadata Management, and Automated Data Quality Assessment on AWS. International Journal of Artificial Intelligence, Data Science, and Machine Learning, 5(2), 230-240. https://doi.org/10.63282/3050-9262.IJAIDSML-V5I2P126