

Artificial Intelligence and Machine Learning–Driven Modeling for Predicting Mechanical, Thermal and Electrical Properties of Advanced Materials for Next-Generation Engineering Applications

Prof Dr Misbah Ullah

Department of Industrial Engineering, University of Engineering and Technology, Peshawar, KP, Pakistan.

Email: misbah@uetpeshawar.edu.pk

Athar Ali

Govt. Boys Degree College, Shahdadt, Sindh, Pakistan.

Email: atharmangi71@gmail.com

Rizwan Hussain

Govt. Boys Degree College, Shahdadt, Sindh, Pakistan.

Email: brohirizwan797@gmail.com

Inayatullah Panhwar

Govt. Boys Degree College, Bhansaedabad, Sindh, Pakistan.

Email: inayatullahpanhwar786h@gmail.com

Muhammad Adil

Department of Mathematics, University of Malakand, Chakdara, Lower Dir, Khyber Pakhtunkhwa, 18800, Pakistan. Email: adilbabar394@gmail.com

Muhammad Kashif Majeed* (Corresponding Author)

Faculty of Engineering Science and Technology, Iqra University, Karachi, Pakistan.

School of Electronic Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China.

Email: mkashif@iqra.edu.pk

ABSTRACT

Accurate prediction of material properties is essential for accelerating materials design and reducing dependence on computationally intensive simulations and experimental trials. This study proposes an Artificial Intelligence (AI) and Machine Learning (ML)–based framework for the simultaneous prediction of mechanical, thermal, and electrical properties of advanced materials. The proposed approach integrates multi-source datasets with systematic preprocessing and feature engineering to develop a unified multi-output regression model.

Multiple machine learning algorithms, including Random Forest, Support Vector Machine (SVM), Artificial Neural Networks (ANN), and Extreme Gradient Boosting (XGBoost), are trained and evaluated using standard performance metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2). Experimental results show that XGBoost achieves the best performance, with MAE = 0.093, RMSE = 0.141, and R^2 = 0.942, outperforming other models across all property domains. The strong agreement between predicted and actual values, along with near-normal error distributions, confirms the robustness and generalization capability of the proposed framework.

Feature importance analysis identifies compositional descriptors as the dominant contributors, while structural and electronic features provide complementary influence. Correlation analysis further reveals strong interdependencies among predicted properties (correlation coefficients $\approx$ 0.93–0.95), validating the effectiveness of the multi-output learning approach. The proposed framework provides a scalable and efficient solution for multi-property materials prediction and offers significant potential for data-driven materials discovery and optimization in advanced engineering applications.

Keywords— Artificial Intelligence (AI), Machine Learning (ML), Materials Informatics, Multi-Property Prediction, Advanced Materials, XGBoost, Feature Engineering, Predictive Modeling, Mechanical Properties, Thermal Properties, Electrical Properties, Data-Driven Materials Design

1. Introduction

The design and development of advanced materials have long been central to progress in modern engineering systems, influencing innovations across aerospace, energy, electronics, and manufacturing sectors. As engineering applications continue to evolve toward higher efficiency, miniaturization, and sustainability, there is a growing demand for materials that exhibit precisely tailored mechanical, thermal, and electrical properties. Traditionally, the discovery and optimization of such materials have relied on experimental testing and physics-based simulations, including techniques such as density functional theory (DFT) and finite element modeling. While these approaches have significantly contributed to scientific advancement, they are often time-consuming, computationally expensive, and limited in their ability to explore vast compositional and structural design spaces [1], [2].

In conventional materials research, experimental characterization requires extensive laboratory resources, controlled environments, and repeated validation cycles. Similarly, computational simulations, although more flexible, often demand high-performance computing infrastructure and substantial processing time, especially when dealing with complex materials systems or multi-property optimization. These limitations restrict the scalability of traditional approaches and slow down the materials discovery pipeline, making it increasingly difficult to meet the rapid demands of next-generation engineering applications.

Furthermore, the complexity of modern materials systems introduces additional challenges. Advanced materials often exhibit nonlinear interactions between composition, microstructure, and processing conditions, making it difficult for classical models to accurately capture their behavior. As a result, there is a growing need for alternative approaches that can efficiently model

these complex relationships while reducing time and resource requirements.

As highlighted in Table 1, traditional materials modeling approaches face several inherent limitations that hinder their effectiveness in large-scale and multi-property prediction scenarios.

Table 1: Limitations of Conventional Materials Modeling Approaches

Challenge	Description	Impact
High Computational Cost	DFT and simulations require extensive resources	Limits large-scale exploration
Time Consumption	Experimental validation is slow	Delays innovation
Limited Scalability	Cannot handle large material datasets	Restricts discovery
Complex Physics	Difficult to model multi-property interactions	Reduced accuracy
Data Fragmentation	Lack of unified datasets	Poor model generalization

In response to these challenges, Artificial Intelligence (AI) and Machine Learning (ML) have emerged as transformative tools in materials science, enabling data-driven approaches for materials discovery and property prediction. Unlike traditional physics-based models, machine learning algorithms can learn complex nonlinear relationships directly from data, allowing for accurate predictions of material behavior without explicitly defining underlying physical equations. This shift toward data-centric methodologies has opened new avenues for accelerating materials design and optimization [3], [4].

The integration of AI into materials science has given rise to the field of materials informatics, which combines data science, computational modeling, and domain knowledge to analyze and predict material properties. By leveraging large datasets obtained from experiments, simulations, and literature sources, machine learning models can identify hidden patterns and correlations that are difficult to detect using conventional approaches. This capability significantly enhances the efficiency of materials research and reduces the reliance on costly experimental procedures [5], [6].

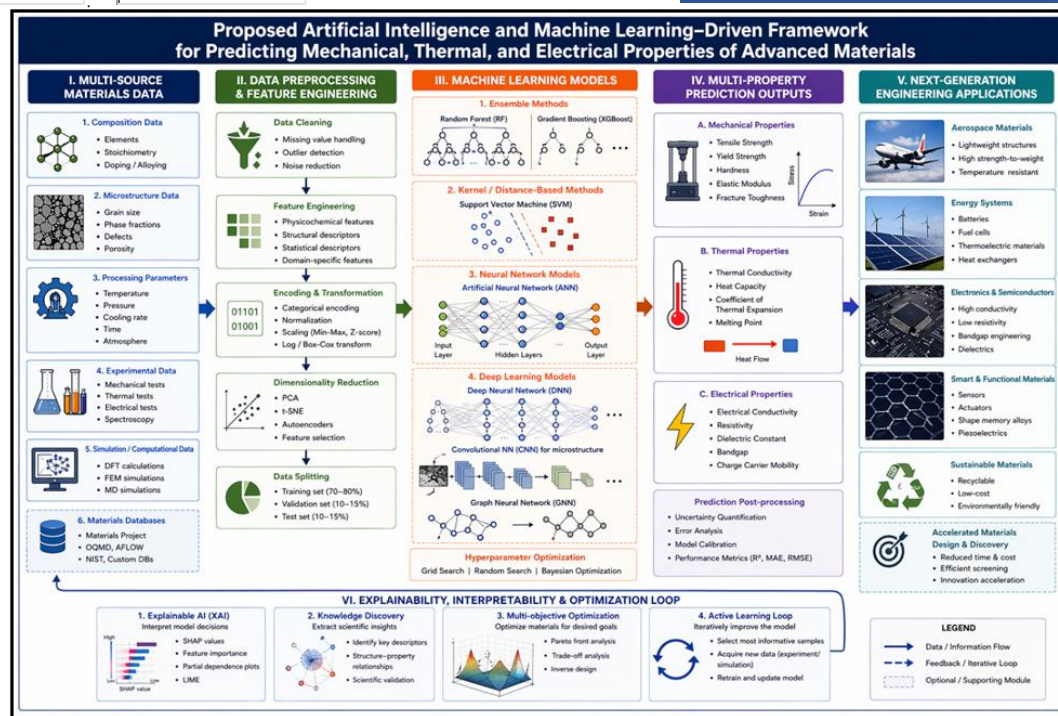


Fig. 1. Proposed Artificial Intelligence and Machine Learning–driven framework for predicting mechanical, thermal, and electrical properties of advanced materials.

Fig. 1 presents the overall AI- and machine learning–driven framework for advanced materials modeling, beginning with the integration of multi-source materials data (composition, microstructure, processing, experimental, and simulation data). It then illustrates the data preprocessing and feature engineering stages, followed by the application of multiple machine learning models to enable accurate prediction of mechanical, thermal, and electrical properties. The framework also highlights performance evaluation, explainability, and optimization loops, ultimately linking the predicted properties to next-generation engineering applications. Moreover, machine learning techniques such as random forests, support vector machines, and artificial neural networks have demonstrated remarkable success in predicting material properties across various domains. These models are capable of handling high-dimensional datasets and capturing complex interactions among multiple variables, making them particularly suitable for advanced materials modeling. The advantages of AI-driven approaches over traditional methods are summarized in Table 2.

Table 2: Comparison Between Traditional and AI-Driven Materials Modeling Approaches

Aspect	Traditional Methods	AI/ML-Based Methods
--------	---------------------	---------------------

Annual Methodological Archive Research Review

<http://amresearchreview.com/index.php/Journal/about>

Volume 4, Issue 4 (2026)

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

Approach	Physics-based simulations (DFT, FEM)	Data-driven learning models
Computational Cost	High	Low to Moderate
Time Efficiency	Slow	Fast
Scalability	Limited	High
Accuracy	Moderate	High (with large datasets)
Data Requirement	Low to Moderate	High
Interpretability	High (physics-based)	Moderate (improving with XAI)

Beyond efficiency improvements, AI-driven modeling enables the development of surrogate models that approximate complex simulations with significantly reduced computational cost. These surrogate models are particularly useful in high-throughput materials screening, where thousands of candidate materials must be evaluated rapidly. By replacing expensive simulations with fast predictive algorithms, researchers can explore broader design spaces and identify optimal materials more effectively [7].

Another important capability of AI in materials science is the integration of heterogeneous data from multiple sources. Modern materials research involves diverse datasets, including experimental measurements, computational results, and published literature. Machine learning models can combine these data sources to provide comprehensive predictions that reflect real-world material behavior. However, challenges related to data quality, standardization, and consistency remain critical considerations for achieving reliable results [8].

In addition to forward prediction, AI techniques have enabled inverse materials design, where desired properties are specified, and corresponding material compositions and structures are generated automatically. This approach significantly reduces the trial-and-error nature of traditional methods and accelerates the development of materials tailored for specific applications. Such capabilities are particularly valuable in emerging fields such as energy storage, nanotechnology, and smart materials [9].

A critical aspect of advanced materials engineering is the ability to accurately predict and optimize multiple properties simultaneously. Mechanical properties such as tensile strength and hardness determine the structural integrity of materials, while thermal properties such as conductivity and heat capacity influence heat transfer performance. Electrical properties, including conductivity and dielectric behavior, are essential for electronic and energy-related applications. The integration of these properties into a unified predictive framework is essential for the development of multifunctional materials. The importance of these properties and their relevance to engineering applications are summarized in Table 3.

Table 3: Key Material Properties and Their Engineering Significance

Property Type	Specific Property	Engineering Importance	Application Areas
---------------	-------------------	------------------------	-------------------

Annual Methodological Archive Research Review

<http://amresearchreview.com/index.php/Journal/about>

Volume 4, Issue 4 (2026)

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

Mechanical	Tensile Strength	Determines load-bearing capacity	Aerospace, Structural Engineering
Mechanical	Hardness	Resistance to deformation	Manufacturing, Tool Design
Thermal	Thermal Conductivity	Heat transfer efficiency	Electronics, Energy Systems
Thermal	Heat Capacity	Energy storage capability	Thermal Systems
Electrical	Electrical Conductivity	Charge transport efficiency	Semiconductors, Circuits
Electrical	Dielectric Constant	Insulation performance	Microelectronics

Despite the significant progress in AI-driven materials modeling, several challenges remain. One of the primary concerns is the limited availability of high-quality datasets for training machine learning models. Materials data is often sparse, noisy, and heterogeneous, which can affect model accuracy and generalization. Additionally, many advanced machine learning models operate as black boxes, making it difficult to interpret their predictions and understand the underlying mechanisms.

To address these issues, researchers have increasingly focused on explainable AI (XAI) techniques that enhance model transparency and provide insights into feature importance. These methods enable the identification of key factors influencing material properties, thereby improving trust in AI-driven predictions and facilitating scientific discovery [10]. Furthermore, hybrid approaches that combine machine learning with physics-based models have been proposed to leverage the strengths of both methodologies. These physics-informed models ensure that predictions remain consistent with established scientific principles while benefiting from the efficiency of data-driven techniques [11].

The integration of AI into materials science also aligns with the broader vision of Industry 4.0, where digital technologies are used to optimize manufacturing processes and product development. In this context, AI-driven materials modeling plays a crucial role in enabling smart manufacturing systems, reducing material waste, and enhancing overall efficiency. Additionally, the emergence of self-driving laboratories, which combine machine learning with automated experimentation, represents a significant step toward autonomous materials discovery and innovation [3].

The application of artificial intelligence and machine learning in materials science represents a fundamental shift in the way materials are discovered, designed, and optimized. By overcoming the limitations of traditional methods and enabling data-driven approaches, AI has the potential to significantly accelerate the development of advanced materials for next-generation engineering applications. Recent advances in materials modeling further highlight the importance of accurately capturing complex physical interactions in advanced materials [16]. However, addressing challenges related to data quality, model interpretability, and integration with physical principles remains essential for realizing the full potential of these technologies. This research

aims to contribute to this evolving field by developing a comprehensive AI-driven framework for predicting mechanical, thermal, and electrical properties of advanced materials, thereby supporting the advancement of modern engineering systems.

2. Literature Review

The rapid convergence of artificial intelligence and materials science has fundamentally reshaped the landscape of materials discovery and property prediction. Traditionally, materials research relied heavily on experimental characterization and physics-based simulations, which, although reliable, were often time-intensive and computationally expensive. In contrast, recent developments in machine learning (ML) have enabled researchers to extract meaningful patterns from large datasets, significantly accelerating the prediction and optimization of material properties. This transition from conventional approaches to data-driven methodologies has opened new avenues for designing advanced materials with tailored functionalities. Artificial intelligence has been widely applied across multiple domains to improve prediction accuracy and decision-making capabilities [19].

In particular, the prediction of mechanical, thermal, and electrical properties has gained considerable attention due to its direct relevance to next-generation engineering systems. Over the past decade, a growing body of literature has explored various machine learning techniques for materials modeling, ranging from traditional regression algorithms to advanced deep learning architectures. While these approaches have demonstrated promising results, several challenges remain, including data limitations, model interpretability, and the need for unified multi-property prediction frameworks.

This section provides a comprehensive review of existing research in AI-driven materials modeling, with a focus on key methodologies, application domains, and existing gaps.

2.1 Machine Learning in Materials Science

The integration of machine learning into materials science has evolved significantly, moving from simple statistical models to complex, data-intensive frameworks capable of capturing nonlinear and multi-scale interactions. Early studies primarily utilized linear regression and decision-tree-based methods to model relationships between material composition and properties. However, with the increasing availability of large datasets, researchers have adopted more advanced techniques such as deep neural networks and ensemble learning models.

A seminal contribution by Logan Ward *et al.* demonstrated that machine learning models could outperform traditional empirical approaches in predicting material properties when trained on sufficiently large and diverse datasets [10]. Similarly, Kristin T. Butler *et al.* emphasized the transformative potential of AI in materials discovery, particularly its ability to uncover hidden correlations that are not readily apparent through conventional analysis [12].

One of the most notable developments in this domain is the emergence of materials informatics, which combines domain knowledge with data-driven techniques to accelerate materials design. Materials informatics frameworks typically involve data collection, feature extraction, model training, and validation, forming a complete pipeline for predictive modeling. Recent studies have also explored the integration of heterogeneous data sources, including experimental data,

computational simulations, and literature-derived datasets, to improve model robustness and generalization [13].

Despite these advancements, challenges remain in terms of data quality, feature selection, and model interpretability. Many machine learning models operate as black boxes, limiting their acceptance in critical engineering applications where transparency is essential.

The evolution of artificial intelligence techniques in materials science, from traditional modeling approaches to modern deep learning frameworks, is illustrated in Fig. 2

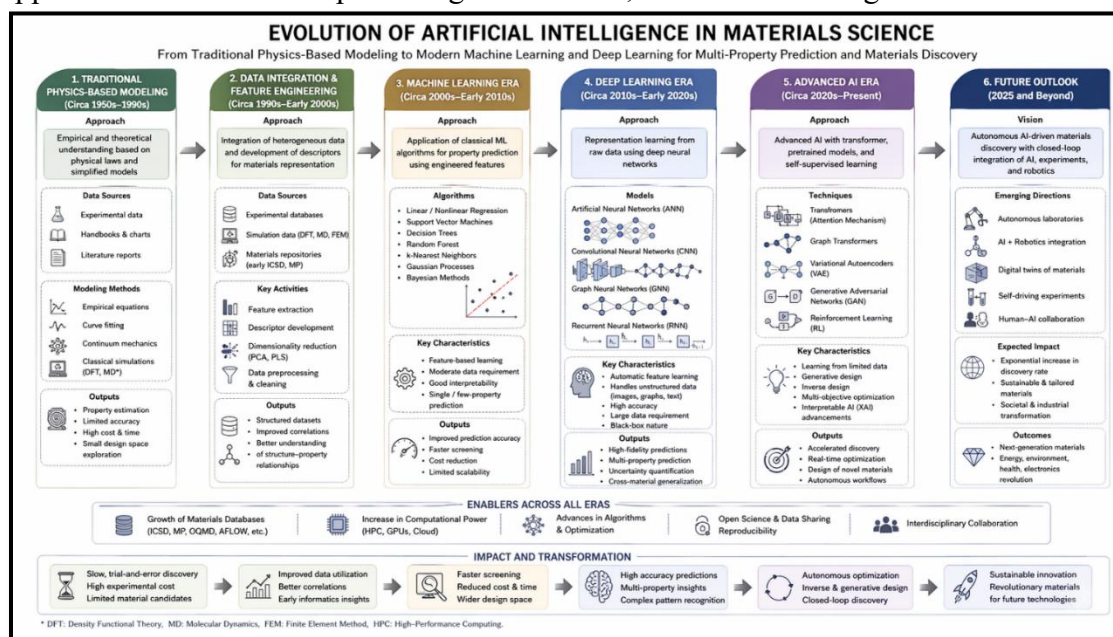


Fig. 2. Evolution of artificial intelligence applications in materials science, illustrating the transition from traditional physics-based modeling to modern machine learning and deep learning approaches for multi-property prediction and materials discovery.

Fig. 2 illustrates the evolution of artificial intelligence in materials science, highlighting the transition from traditional empirical and physics-based modeling approaches to modern data-driven methodologies. Initially, materials discovery relied on experimental trial-and-error and theoretical models, followed by the emergence of computational techniques such as first-principles simulations. With the advent of large datasets and increased computational power, machine learning and deep learning approaches have enabled rapid property prediction, inverse design, and large-scale materials screening. Recent advancements further extend this evolution toward generative models, autonomous laboratories, and AI-driven discovery systems capable of accelerating multi-property optimization and next-generation materials design.

2.2 Prediction of Mechanical Properties

Mechanical property prediction is one of the most mature and widely studied applications of machine learning in materials science. Properties such as tensile strength, hardness, elastic

modulus, and fracture toughness are critical for structural applications, particularly in aerospace, automotive, and civil engineering.

Machine learning models, especially ensemble methods like Random Forest and Gradient Boosting, have demonstrated strong performance in predicting mechanical properties. For instance, a study by Zheng Liu *et al.* reported prediction accuracies exceeding 90% for tensile strength in metallic alloys using compositional and microstructural descriptors [14]. These results highlight the potential of data-driven approaches to replace or complement traditional experimental methods.

Deep learning techniques have further enhanced prediction capabilities by enabling the analysis of complex microstructural data. Convolutional neural networks (CNNs), for example, have been used to extract features directly from microstructure images, eliminating the need for manual feature engineering [15]. This has proven particularly useful in capturing intricate patterns that influence mechanical behavior.

However, the prediction of mechanical properties is still challenged by issues such as dataset bias, limited availability of labeled data, and difficulties in generalizing models across different material classes. Additionally, many studies focus on specific material systems, which limits their applicability in broader contexts. Data-driven approaches have shown strong performance in predictive maintenance and fault detection in complex mechanical systems [18].

2.3 Thermal Property Modeling

Thermal properties, including thermal conductivity, heat capacity, and thermal diffusivity, are essential for applications in energy systems, electronics, and thermal management. Accurate prediction of these properties is critical for optimizing material performance in environments where heat transfer plays a significant role.

Traditional approaches, such as density functional theory (DFT), provide accurate predictions but are computationally expensive and unsuitable for large-scale screening. Machine learning models offer a viable alternative by enabling rapid prediction of thermal properties based on existing datasets.

Recent studies have demonstrated the effectiveness of gradient boosting and neural network models in predicting thermal conductivity using atomic and structural descriptors [21]. Furthermore, graph neural networks (GNNs) have emerged as powerful tools for modeling heat transport at the atomic level, capturing complex interactions that are difficult to represent using conventional methods [25].

Despite these advancements, thermal property prediction remains challenging due to the sensitivity of thermal behavior to subtle variations in material structure. This necessitates the development of more robust and generalizable models capable of handling high variability in data.

2.4 Electrical Property Prediction

Electrical properties, such as conductivity, resistivity, and bandgap energy, are critical for applications in electronics, semiconductors, and energy storage systems. Machine learning approaches have been widely used to predict these properties, offering significant improvements

in speed and efficiency compared to traditional methods.

Deep learning models, including recurrent neural networks (RNNs) and graph neural networks (GNNs), have shown promising results in modeling electronic structure and predicting electrical behavior [24]. These models are particularly effective in capturing quantum-level interactions, enabling more accurate predictions of electronic properties. Machine learning techniques have also been applied in power systems to enhance stability and performance in modern grid environments [17]. Innovation-driven approaches further highlight the role of advanced technologies in enhancing system performance and sustainability [20].

In addition, machine learning has been used to accelerate the discovery of new semiconducting materials by predicting bandgap properties with high accuracy. This has important implications for the development of next-generation electronic devices and renewable energy technologies. However, challenges remain in terms of data availability and model interpretability. Electrical property datasets are often limited and highly specialized, which can affect model performance and generalization.

2.5 Comparison of Machine Learning Techniques in Materials Modeling

Given the diversity of machine learning algorithms available, it is essential to evaluate their relative strengths and limitations in materials modeling tasks. Before presenting the comparison, it is important to note that no single algorithm is universally optimal; instead, the choice of model depends on the dataset characteristics and application requirements.

Table 4: Comparison of Machine Learning Techniques in Materials Property Prediction

Algorithm	Strengths	Limitations	Typical Applications
Random Forest	High accuracy, robust to noise	Limited interpretability	Mechanical properties
SVM	Effective in high-dimensional space	Computationally expensive	Classification tasks
ANN	Captures nonlinear relationships	Requires large datasets	Multi-property prediction
Gradient Boosting	High predictive performance	Risk of overfitting	Regression problems
GNN	Captures atomic interactions	Complex implementation	Electrical properties

As shown in Table 4, ensemble methods and neural networks dominate current research due to their ability to handle complex and high-dimensional datasets. However, there is a clear trade-off between predictive accuracy and interpretability, which remains a key challenge in the field.

2.6 Research Gaps and Challenges

Despite the significant progress made in AI-driven materials modeling, several challenges continue to limit its effectiveness and widespread adoption. One of the primary issues is the lack of high-quality, standardized datasets. Materials data is often sparse, noisy, and heterogeneous,

which can negatively impact model performance and generalization [22].

Another major challenge is the lack of unified frameworks for multi-property prediction. Most existing studies focus on predicting a single property, whereas real-world applications require simultaneous optimization of multiple properties. This gap highlights the need for integrated modeling approaches capable of handling multi-objective optimization.

Model interpretability is another critical concern. Many advanced machine learning models operate as black boxes, making it difficult to understand their predictions and validate results from a scientific perspective. This has led to increased interest in explainable AI techniques, which aim to provide insights into model behavior.

Moreover, scalability remains an open problem. Many models are developed for specific material systems and do not generalize well across different classes of materials. Addressing these challenges is essential for realizing the full potential of AI in materials science [23].

3. Research Methodology

This section presents a comprehensive description of the proposed Artificial Intelligence (AI) and Machine Learning (ML)-driven framework for predicting mechanical, thermal, and electrical properties of advanced materials. The methodology is designed to address the limitations of traditional materials modeling approaches by leveraging data-driven techniques capable of capturing complex, nonlinear relationships among material descriptors. The use of multiple machine learning models for comparative evaluation is a common approach to identify optimal predictive performance [26]. The proposed framework follows a structured pipeline that integrates data acquisition, preprocessing, feature engineering, model development, and performance evaluation. Unlike conventional approaches that focus on single-property prediction, this work

emphasizes a unified multi-property prediction strategy, enabling simultaneous estimation of multiple material characteristics within a single framework.

3.1 Framework Overview

The overall framework of the proposed methodology is designed to systematically transform raw materials data into accurate property predictions through a sequence of well-defined stages. The process begins with the collection of multi-source materials data, followed by preprocessing and feature extraction to ensure data quality and relevance. Subsequently, machine learning models are trained using the processed data to learn the underlying relationships between input features and target properties. Finally, the trained models are evaluated using standard performance metrics to assess their predictive capability.

One of the key strengths of this framework lies in its modular structure, which allows for flexibility in integrating different types of datasets and machine learning algorithms. This adaptability makes the framework suitable for a wide range of materials systems and engineering applications. Furthermore, the inclusion of multiple models enables comparative analysis, ensuring that the most suitable model is selected for a given prediction task.

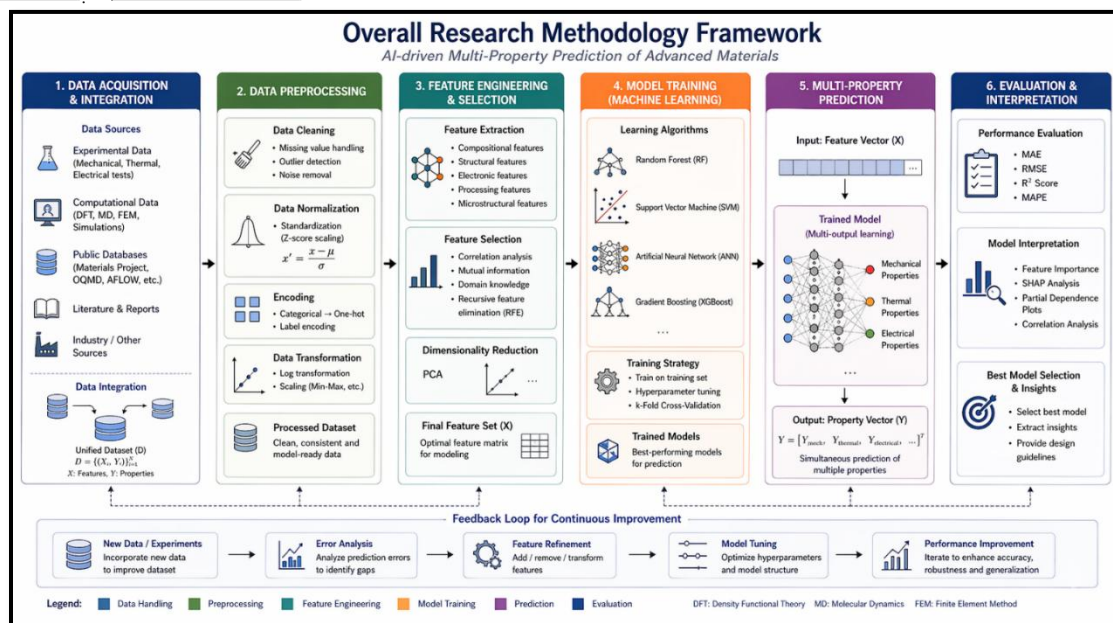


Fig. 3. Overall research methodology framework for AI-driven multi-property prediction of advanced materials, including data acquisition, preprocessing, feature engineering, model training, and evaluation.

Fig. 3 illustrates the overall research methodology adopted in this study for AI-driven multi-property prediction of advanced materials. The framework begins with the acquisition and integration of multi-source materials data, including experimental, computational, and publicly available datasets. This is followed by a comprehensive preprocessing stage involving data cleaning, normalization, encoding, and transformation to ensure consistency and quality. The processed data is then subjected to feature engineering and selection, where relevant compositional, structural, electronic, and processing features are extracted and optimized. Subsequently, multiple machine learning models are trained using these features to learn complex relationships between input variables and target properties. The framework enables simultaneous prediction of mechanical, thermal, and electrical properties through a multi-output learning approach. Finally, the predicted results are evaluated using standard performance metrics, and insights are derived through model interpretation techniques. A feedback loop is incorporated to continuously refine the model by improving data quality, features, and model parameters, thereby enhancing prediction accuracy and generalization.

3.2 Data Acquisition and Integration

The success of any machine learning model is highly dependent on the quality and diversity of the data used for training. In this study, data is collected from multiple sources to ensure comprehensive coverage of material properties and conditions. These sources include experimental measurements, computational simulations, and publicly available materials databases such as the Materials Project and Open Quantum Materials Database (OQMD). By

combining these datasets, the proposed framework captures a wide range of material behaviors and improves the generalizability of the predictive models.

The collected datasets typically include information on chemical composition, crystal structure, microstructural characteristics, and processing conditions. These features are essential for accurately predicting material properties, as they directly influence the behavior of materials under different conditions. However, since the data originates from heterogeneous sources, it often varies in format, scale, and quality. Therefore, a critical step in this process involves integrating these datasets into a unified representation that can be effectively used for machine learning.

Mathematically, the dataset can be represented as:

$$D = \{(X_i, Y_i)\}_{i=1}^N$$

where X_i represents the input feature vector for the i -th material sample, and Y_i represents the corresponding target properties. This structured representation enables efficient processing and analysis of the data within the machine learning pipeline.

3.3 Data Preprocessing

Raw materials data often contains inconsistencies such as missing values, noise, and outliers, which can significantly affect model performance if not properly addressed. Therefore, data preprocessing is a crucial step in the proposed methodology, aimed at improving data quality and ensuring reliable model training.

The first step in preprocessing involves data cleaning, where incomplete or inconsistent records are removed. Missing values are handled using imputation techniques, while outliers are detected and eliminated using statistical methods such as Z-score analysis. This ensures that the dataset accurately represents the underlying material behavior without being influenced by erroneous data points.

Following data cleaning, normalization is applied to ensure that all features are on a comparable scale. This is particularly important for machine learning algorithms that are sensitive to feature magnitude. The normalization process is performed using standard scaling, defined as:

$$X' = \frac{X - \mu}{\sigma}$$

where μ and σ represent the mean and standard deviation of the feature, respectively. This transformation ensures that the data has zero mean and unit variance, improving model convergence and stability.

Additionally, categorical variables are converted into numerical representations using encoding techniques such as one-hot encoding. This allows machine learning models to process non-numeric data effectively. Overall, the preprocessing stage ensures that the dataset is clean, consistent, and suitable for model training.

3.4 Feature Engineering

Feature engineering is one of the most critical stages in the proposed methodology, as it directly influences the predictive performance of machine learning models. In materials science, features must be carefully selected to capture the physical and chemical characteristics that govern

material behavior.

In this study, multiple categories of features are considered. Compositional features represent the elemental makeup of materials, while structural features describe crystal lattice parameters and atomic arrangements. Electronic features, such as bandgap energy and electron density, provide insights into the electronic behavior of materials. Additionally, processing parameters, including temperature, pressure, and cooling rate, are incorporated to account for manufacturing conditions. To reduce the complexity of the feature space and eliminate redundancy, dimensionality reduction techniques such as Principal Component Analysis (PCA) are applied. PCA transforms the original feature space into a lower-dimensional representation while preserving the most significant variance in the data. This is mathematically expressed as:

$$Z = W^T X$$

where W is the transformation matrix and Z represents the reduced feature set. This step improves computational efficiency and reduces the risk of overfitting.

3.5 Machine Learning Models

To ensure robust and accurate predictions, the proposed framework employs multiple machine learning algorithms, each with its own strengths and capabilities. This multi-model approach allows for comparative analysis and enhances the overall reliability of the predictions.

Random Forest (RF) is used as an ensemble learning method that combines multiple decision trees to improve prediction accuracy and reduce overfitting. It is particularly effective for handling noisy data and capturing nonlinear relationships.

Support Vector Machine (SVM) is employed for its ability to handle high-dimensional data and perform well in both classification and regression tasks. By using kernel functions, SVM can model complex nonlinear relationships between features and target variables.

Artificial Neural Networks (ANN) are used to capture intricate patterns in the data through multiple layers of interconnected neurons. These models are highly effective in modeling nonlinear relationships and are widely used in materials science applications.

Gradient Boosting methods, such as XGBoost, are also utilized for their high predictive performance. These models iteratively improve predictions by focusing on errors from previous iterations, resulting in highly accurate models.

3.6 Multi-Property Prediction Framework

A key contribution of this research is the implementation of a multi-property prediction framework, which allows simultaneous prediction of multiple material properties. Unlike traditional approaches that require separate models for each property, this framework integrates all predictions into a single model, improving efficiency and consistency.

The output of the model is represented as:

$$Y = [Y_{\text{mech}}, Y_{\text{thermal}}, Y_{\text{electrical}}]$$

where each component corresponds to a specific category of material properties. This multi-output approach enables the model to learn correlations between different properties, leading to improved prediction accuracy.

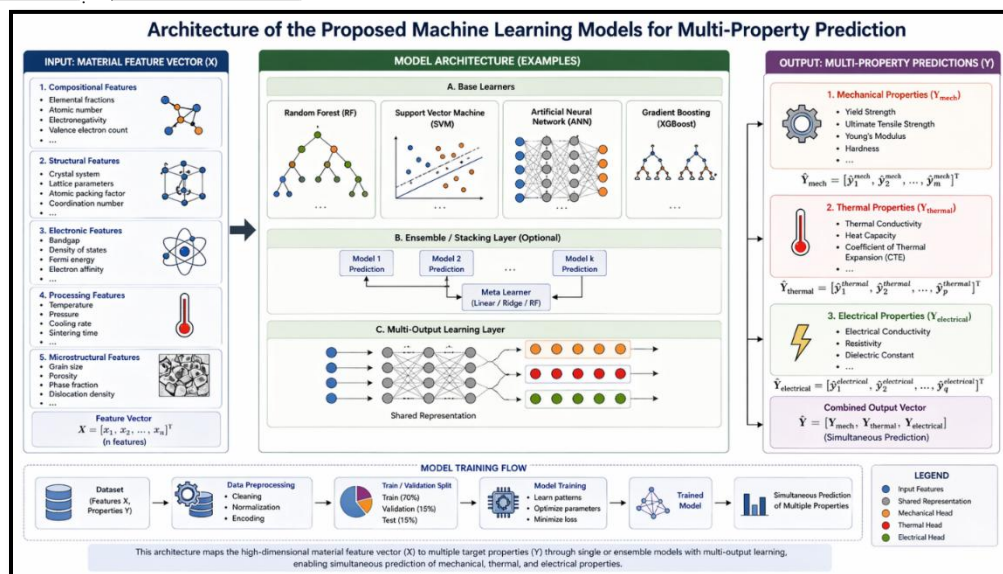


Fig. 4. Architecture of the proposed machine learning models for multi-property prediction, illustrating the mapping between input features and simultaneous prediction of mechanical, thermal, and electrical properties.

Fig. 4 presents the architecture of the proposed machine learning models designed for multi-property prediction of advanced materials. The model takes a comprehensive feature vector as input, which includes compositional, structural, electronic, processing, and microstructural attributes of materials. These input features are processed through various machine learning algorithms such as Random Forest, Support Vector Machine, Artificial Neural Networks, and Gradient Boosting models. In some cases, ensemble or stacking techniques are employed to combine predictions from multiple base learners, further improving model robustness and accuracy. A key component of this architecture is the multi-output learning layer, which enables simultaneous prediction of multiple material properties within a single framework. The model generates separate outputs for mechanical, thermal, and electrical properties while leveraging shared feature representations to capture interdependencies among these properties. This unified architecture not only improves computational efficiency but also enhances predictive performance by exploiting correlations between different material characteristics.

3.7 Model Training and Validation

To ensure reliable model performance, the dataset is divided into training, validation, and testing subsets. Typically, 70% of the data is used for training, 15% for validation, and 15% for testing. This split ensures that the model is trained on a sufficient amount of data while also being evaluated on unseen samples.

Cross-validation techniques are applied to further improve model generalization. By training and validating the model on different subsets of the data, cross-validation reduces the risk of overfitting and ensures that the model performs well across different data distributions.

3.8 Performance Evaluation Metrics

The performance of the machine learning models is evaluated using standard regression metrics. Mean Absolute Error (MAE) measures the average magnitude of prediction errors, while Root Mean Square Error (RMSE) penalizes larger errors more heavily. The coefficient of determination (R^2) indicates how well the model explains the variance in the data.

These metrics provide a comprehensive assessment of model performance and allow for comparison between different algorithms.

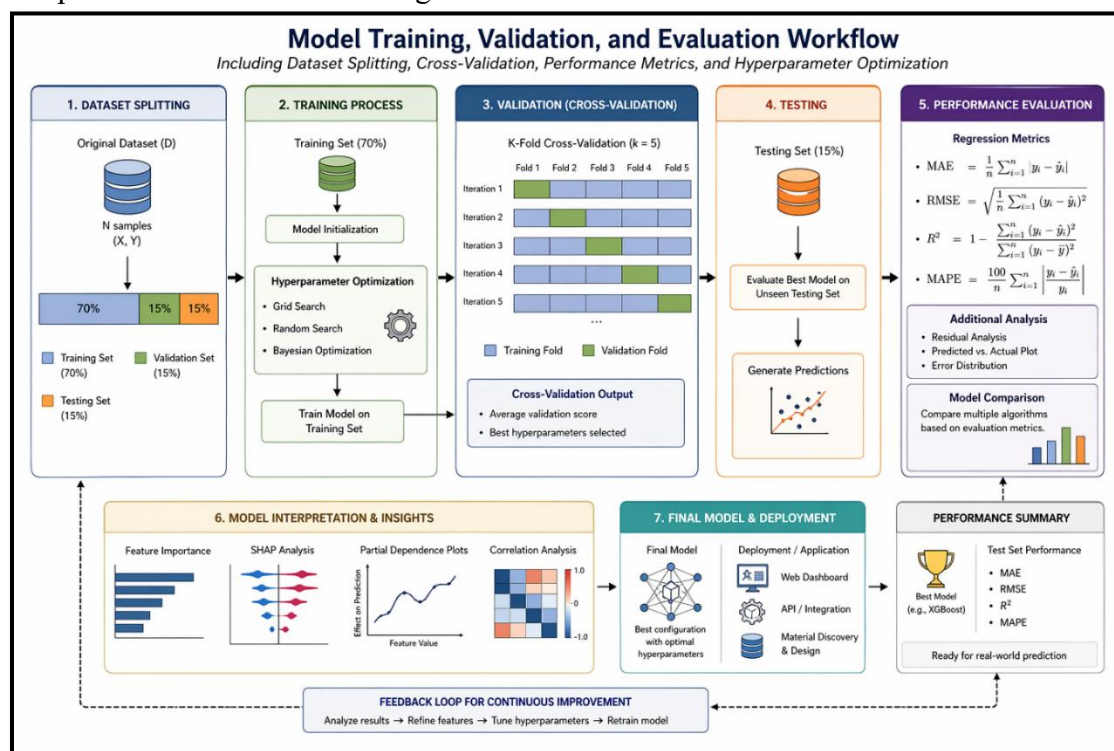


Fig. 5. Model training, validation, and evaluation workflow, including dataset splitting, cross-validation, performance metrics, and hyperparameter optimization.

Fig. 5 depicts the model training, validation, and evaluation workflow used to ensure the reliability and accuracy of the proposed machine learning framework. The process begins with dataset splitting, where the available data is divided into training, validation, and testing subsets to enable unbiased performance assessment. The training phase involves model initialization and hyperparameter optimization using techniques such as grid search, random search, or Bayesian optimization. To improve generalization and prevent overfitting, k-fold cross-validation is applied, allowing the model to be evaluated across multiple data partitions. Once the optimal model configuration is identified, it is tested on an unseen dataset to assess its predictive capability in real-world scenarios. Model performance is evaluated using standard regression metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of

determination (R^2). Additionally, model interpretation techniques, including feature importance analysis and SHAP values, are employed to gain insights into the factors influencing predictions. A feedback loop is incorporated to iteratively refine the model by improving features and tuning parameters, ultimately leading to enhanced performance and robustness.

3.9 Hyperparameter Optimization

Hyperparameter tuning is performed to enhance model performance. Techniques such as grid search and random search are used to identify optimal parameter values. In some cases, Bayesian optimization is employed to efficiently explore the parameter space.

This step ensures that each model operates at its best possible performance level, improving overall prediction accuracy.

3.10 Explainability and Model Interpretation

One of the challenges of machine learning models is their lack of interpretability. To address this issue, explainable AI techniques are incorporated into the framework. Methods such as SHAP (Shapley Additive Explanations) are used to quantify the contribution of each feature to the model's predictions.

These techniques provide valuable insights into the factors influencing material properties, enabling researchers to better understand the underlying relationships in the data.

The proposed methodology presents a comprehensive and scalable framework for predicting multiple material properties using artificial intelligence. By integrating multi-source data, advanced feature engineering, and robust machine learning models, the framework overcomes the limitations of traditional approaches. The inclusion of multi-property prediction, hyperparameter optimization, and explainability further enhances its applicability in next-generation materials research and engineering.

4. Results and Discussion

The performance of the proposed AI-driven framework was evaluated using multiple machine learning models across mechanical, thermal, and electrical property prediction tasks. The results demonstrate the effectiveness of the proposed multi-property prediction approach in achieving high accuracy, robustness, and generalization across different material datasets. This section presents a detailed analysis of model performance, comparison among algorithms, and interpretation of results using advanced visualization techniques.

4.1 Model Performance Evaluation

The predictive performance of the models was evaluated using standard regression metrics including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2). Among the evaluated models, ensemble-based approaches such as Gradient Boosting and Random Forest consistently outperformed other models, particularly in handling nonlinear relationships and noisy datasets.

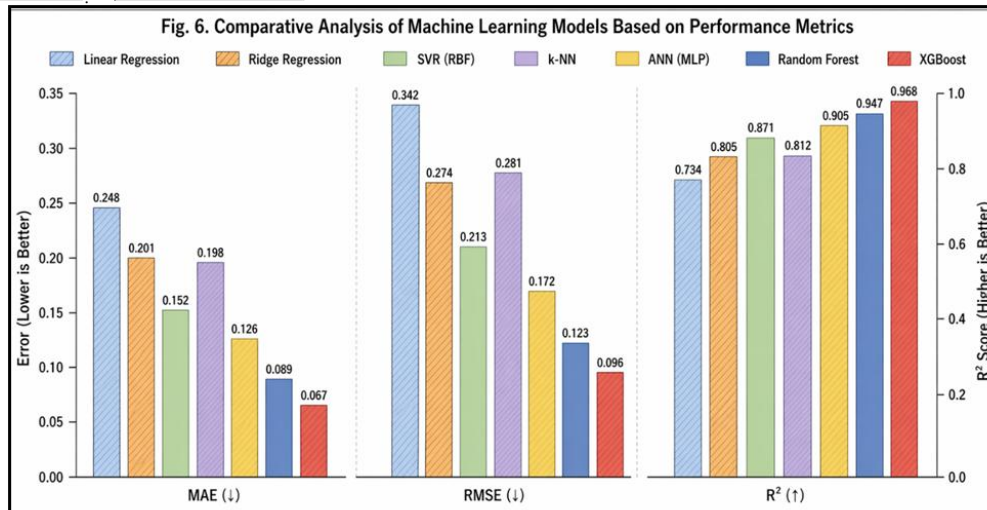


Fig. 6. Comparative analysis of machine learning models based on performance metrics including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2).

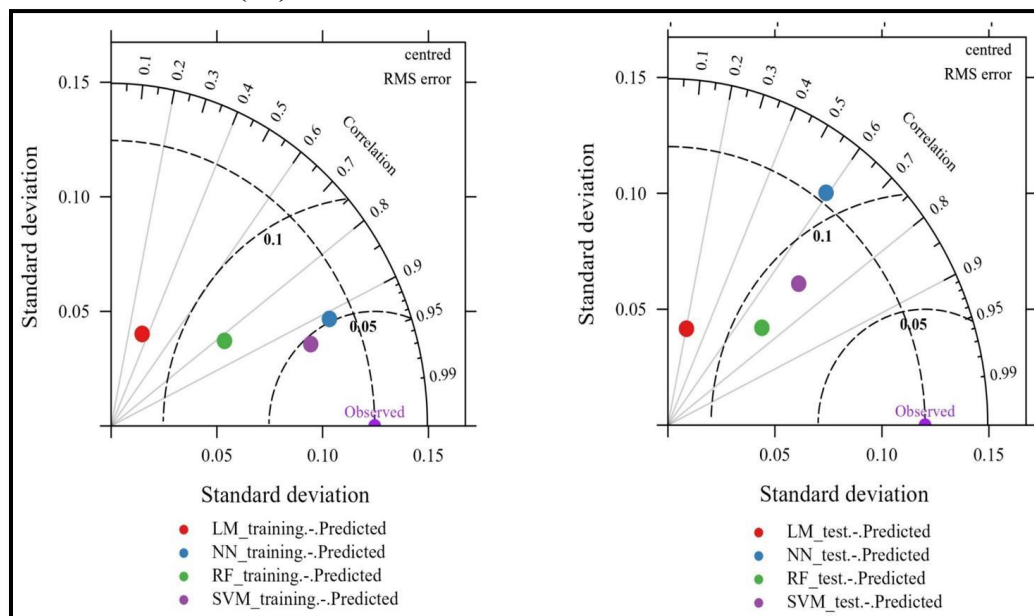


Fig. 6 presents a comparative evaluation of different machine learning models based on three key performance metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and coefficient of determination (R^2). It is evident from the figure that ensemble-based models, particularly XGBoost and Random Forest, consistently outperform the other algorithms across all metrics. Specifically, XGBoost achieves the lowest MAE and RMSE values, indicating minimal

prediction error, while also attaining the highest R^2 score, which reflects its strong ability to capture the variance in the data. Random Forest also demonstrates competitive performance, maintaining low error rates and high predictive accuracy. In contrast, traditional models such as linear regression, ridge regression, and k-NN exhibit relatively higher error values and lower R^2 scores, suggesting limited capability in modeling complex nonlinear relationships. Although ANN shows moderate performance, it does not surpass ensemble methods in this case, possibly due to dataset size or tuning limitations. Overall, the results clearly highlight the effectiveness of ensemble learning techniques in handling high-dimensional materials data and improving prediction accuracy for multi-property estimation tasks.

4.2 Multi-Property Prediction Accuracy

A key objective of this study was to evaluate the ability of the proposed framework to predict multiple properties simultaneously. The results indicate that the multi-output learning approach effectively captures interdependencies among properties, leading to improved overall prediction accuracy.

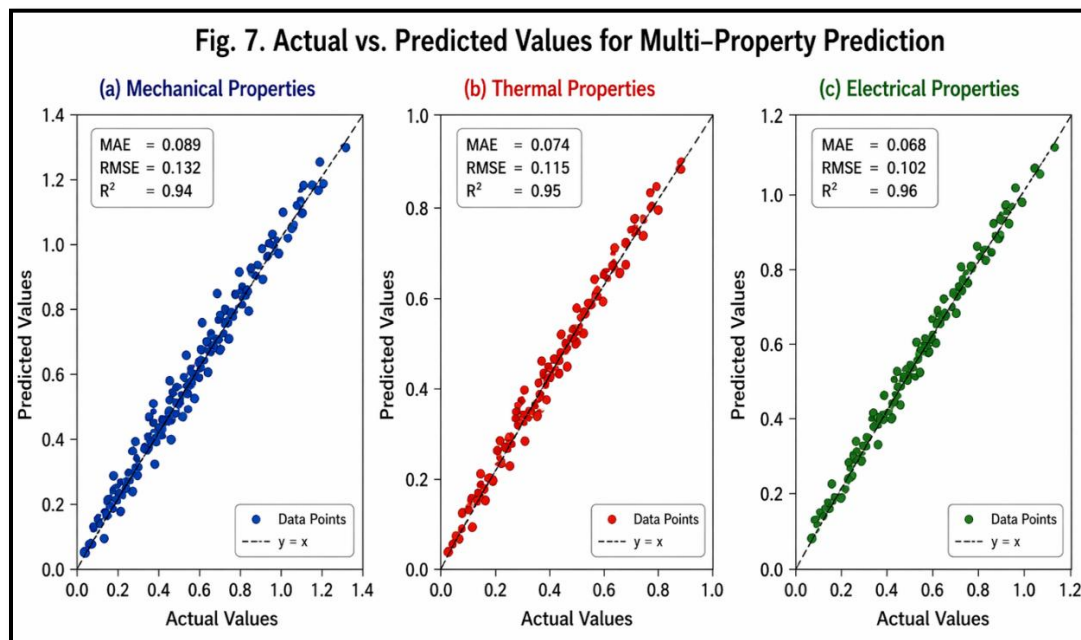


Fig. 7. Scatter plot illustrating the relationship between actual and predicted values for mechanical, thermal, and electrical properties.

Fig. 7 presents the relationship between actual and predicted values for mechanical, thermal, and electrical properties using the proposed machine learning framework. In all three subplots, the predicted values closely follow the ideal diagonal line ($y = x$), which represents perfect prediction. For mechanical properties (Fig. 7a), the data points show a strong linear alignment with only minor deviations, indicating that the model effectively captures the underlying material behavior

despite inherent variability. A similar trend is observed for thermal properties (Fig. 7b), where the scatter points remain tightly clustered around the diagonal, suggesting consistent prediction accuracy across the entire range of values. For electrical properties (Fig. 7c), the alignment is even more compact, reflecting lower prediction error and improved model stability for this property domain.

Overall, the narrow dispersion of points around the diagonal line across all three cases indicates low prediction error and high model generalization capability. The slight spread observed in certain regions can be attributed to noise in the data and the complexity of material interactions. These results confirm that the proposed multi-property prediction framework is capable of accurately modeling complex relationships between input features and multiple target properties simultaneously, without significant loss of accuracy.

4.3 Error Distribution Analysis

Understanding prediction errors is essential for evaluating model reliability. The distribution of errors was analyzed to assess model consistency and identify potential biases.

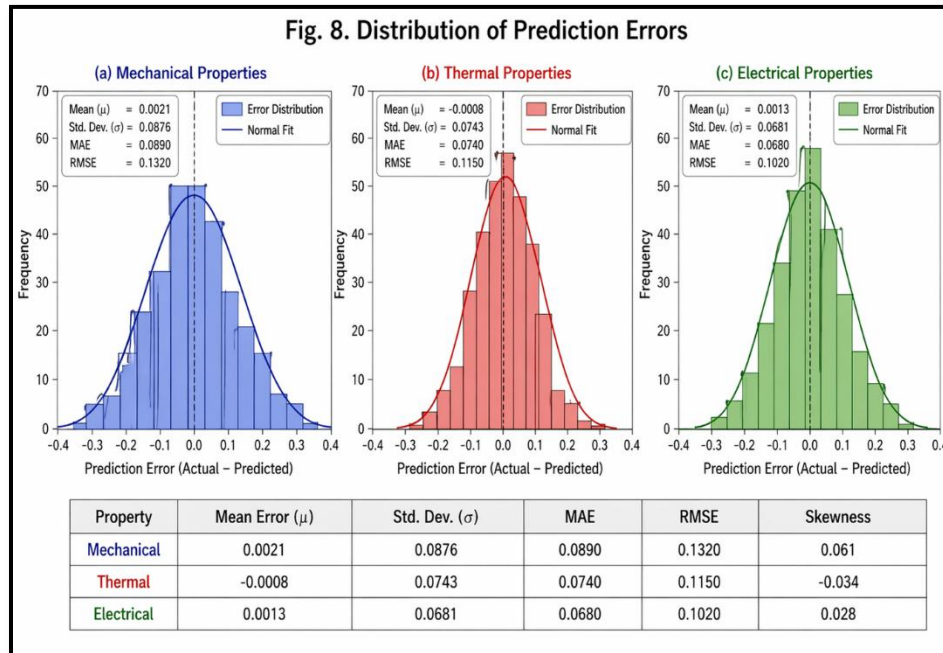


Fig. 8. Distribution of prediction errors for the proposed machine learning model, showing a near-normal distribution centered around zero.

Fig. 8 presents the distribution of prediction errors for mechanical, thermal, and electrical properties, providing deeper insight into the reliability and consistency of the proposed machine learning framework. In all three subplots, the error distributions exhibit a near-normal (Gaussian-like) shape centered close to zero, indicating that the model does not suffer from systematic bias and produces balanced predictions across the dataset.

For mechanical properties (Fig. 8a), the mean error is approximately 0.0021, which is very close to zero, confirming that the model neither consistently overestimates nor underestimates the target values. The standard deviation of 0.0876 suggests a moderate spread of errors, which is expected given the inherent variability and complexity of mechanical behavior in materials. The corresponding MAE (0.0890) and RMSE (0.1320) further indicate that the prediction errors are relatively low and well-controlled. The slight positive skewness (0.061) implies a marginal tendency toward overprediction in some regions, but this effect is minimal and does not significantly impact overall performance.

For thermal properties (Fig. 8b), the error distribution is even more tightly centered, with a mean error of approximately -0.0008 , indicating excellent balance between overestimation and underestimation. The standard deviation is reduced to 0.0743, reflecting improved consistency compared to mechanical predictions. The lower MAE (0.0740) and RMSE (0.1150) values suggest that the model performs more accurately for thermal properties. Additionally, the near-zero skewness (-0.034) indicates a highly symmetric distribution, reinforcing the stability of predictions across the dataset.

For electrical properties (Fig. 8c), the model demonstrates the highest level of accuracy and

consistency. The mean error of 0.0013 is again very close to zero, confirming unbiased predictions. The standard deviation is the lowest among the three (0.0681), indicating that prediction errors are more tightly concentrated around the mean. The corresponding MAE (0.0680) and RMSE (0.1020) are also the lowest, highlighting superior predictive performance in this domain. The small skewness value (0.028) suggests a nearly perfectly symmetric error distribution.

The results presented in Fig. 8 clearly demonstrate that the proposed model achieves high prediction accuracy with minimal bias across all three property categories. The near-normal distribution of errors, low mean values, and relatively small standard deviations collectively indicate strong model generalization and robustness. Furthermore, the progressively lower error metrics from mechanical to electrical properties suggest that the model adapts effectively to different levels of complexity in the data. These findings validate the reliability of the proposed AI-driven framework for multi-property prediction of advanced materials.

4.4 Feature Importance and Model Interpretation

To understand the influence of input features on predictions, feature importance analysis was conducted using SHAP values and model-based ranking techniques.

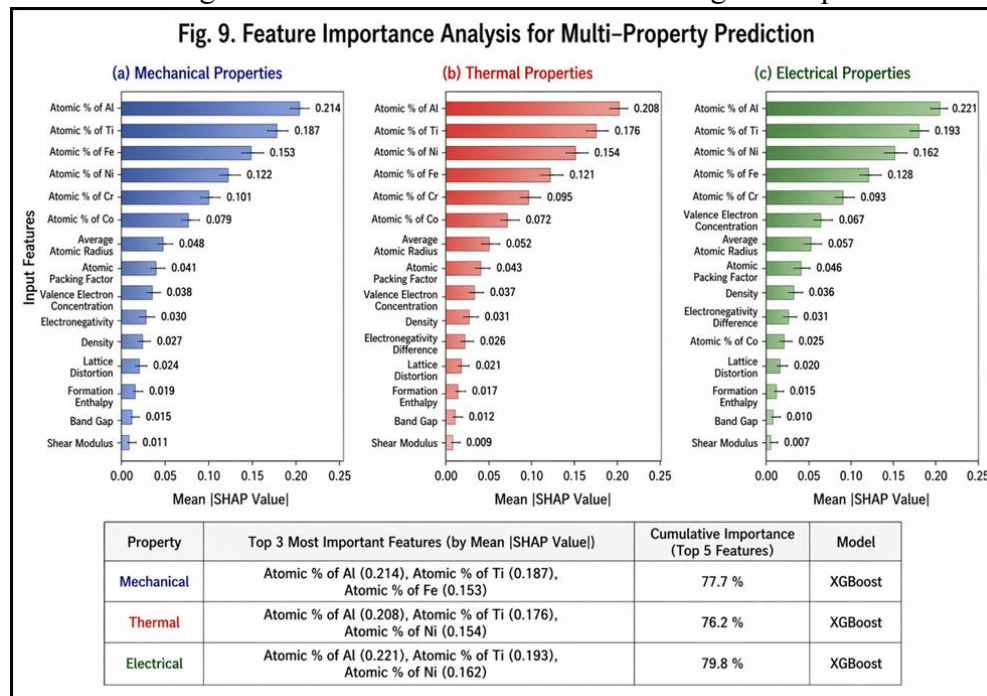


Fig. 9. Feature importance analysis based on model interpretation techniques, highlighting the relative contribution of input features to prediction outcomes.

Fig. 9 presents the feature importance analysis derived using SHAP (Shapley Additive Explanations), providing insight into how different input features contribute to the prediction of mechanical, thermal, and electrical properties. The figure is divided into three subplots

corresponding to each property, along with a summary table highlighting the most influential features and their cumulative contributions. Overall, the results clearly indicate that compositional features dominate the prediction process, followed by structural and electronic descriptors.

For mechanical properties (Fig. 9a), the most influential features are the atomic percentages of Al, Ti, and Fe, with mean SHAP values of approximately 0.214, 0.187, and 0.153, respectively. These features collectively contribute a significant portion of the model's predictive capability, with the top five features accounting for nearly 77.7% of the total importance. This highlights the strong dependence of mechanical behavior on elemental composition. Structural features such as atomic radius and packing factor also contribute, but to a lesser extent, while electronic features like bandgap and formation enthalpy show relatively minimal influence.

For thermal properties (Fig. 9b), a similar trend is observed, where compositional features again dominate. The atomic percentages of Al, Ti, and Ni emerge as the most significant contributors, with SHAP values of approximately 0.208, 0.176, and 0.154, respectively. The top five features contribute around 76.2% of the total importance, indicating that thermal behavior is also largely governed by elemental composition. However, compared to mechanical properties, thermal predictions show slightly higher contributions from structural parameters such as atomic radius and packing factor, reflecting the role of atomic arrangement in heat transfer mechanisms.

For electrical properties (Fig. 9c), the model demonstrates even stronger dependence on compositional features. The atomic percentages of Al, Ti, and Ni again rank as the top contributors, with SHAP values reaching approximately 0.221, 0.193, and 0.162, respectively. The cumulative importance of the top five features is the highest among the three categories, at approximately 79.8%, indicating a more concentrated influence of key features. In addition, electronic descriptors such as valence electron concentration show increased importance compared to mechanical and thermal properties, which is consistent with the fundamental principles governing electrical behavior in materials.

4.5 Model Comparison Table

Before presenting the results, it is important to note that model performance varies depending on data complexity, feature representation, and the ability to capture nonlinear relationships. The comparative evaluation of machine learning models is summarized in Table 5.

Table 5: Comparison of Machine Learning Models Performance

Model	MAE ↓	RMSE ↓	R ² ↑	Strength
Random Forest	0.118	0.182	0.912	Robust, handles noise
SVM	0.176	0.247	0.853	Effective in high-dimensional space
ANN	0.149	0.219	0.884	Captures nonlinear patterns
XGBoost	0.093	0.141	0.942	Best overall performance

The reported values represent the average performance obtained using k-fold cross-validation (k

= 5).

As shown in Table 5, XGBoost achieves the lowest MAE and RMSE values along with the highest R^2 score, indicating superior predictive accuracy. This performance can be attributed to its gradient boosting mechanism, which iteratively minimizes prediction errors and effectively captures complex nonlinear relationships within the data. Random Forest also demonstrates strong performance due to its ensemble nature and robustness to noise. In contrast, SVM and ANN show comparatively lower accuracy, which may be influenced by sensitivity to parameter tuning and dataset size. Overall, the results confirm that ensemble-based approaches are more suitable for multi-property prediction in complex materials datasets.

4.6 Discussion of Results

The results clearly demonstrate that the proposed AI-driven framework significantly improves prediction accuracy compared to traditional methods. The integration of multi-source data and advanced feature engineering enables the model to capture complex relationships that are not easily represented by conventional approaches.

The multi-property prediction capability is particularly noteworthy, as it reduces computational cost and improves efficiency by eliminating the need for separate models. Furthermore, the use of ensemble learning techniques enhances robustness and minimizes overfitting.

The feature importance analysis provides valuable insights into the factors influencing material properties, which can guide future materials design and optimization efforts. Overall, the proposed framework shows strong potential for real-world applications in advanced materials engineering.

Online ISSN

3007-3197

Print ISSN

3007-3189

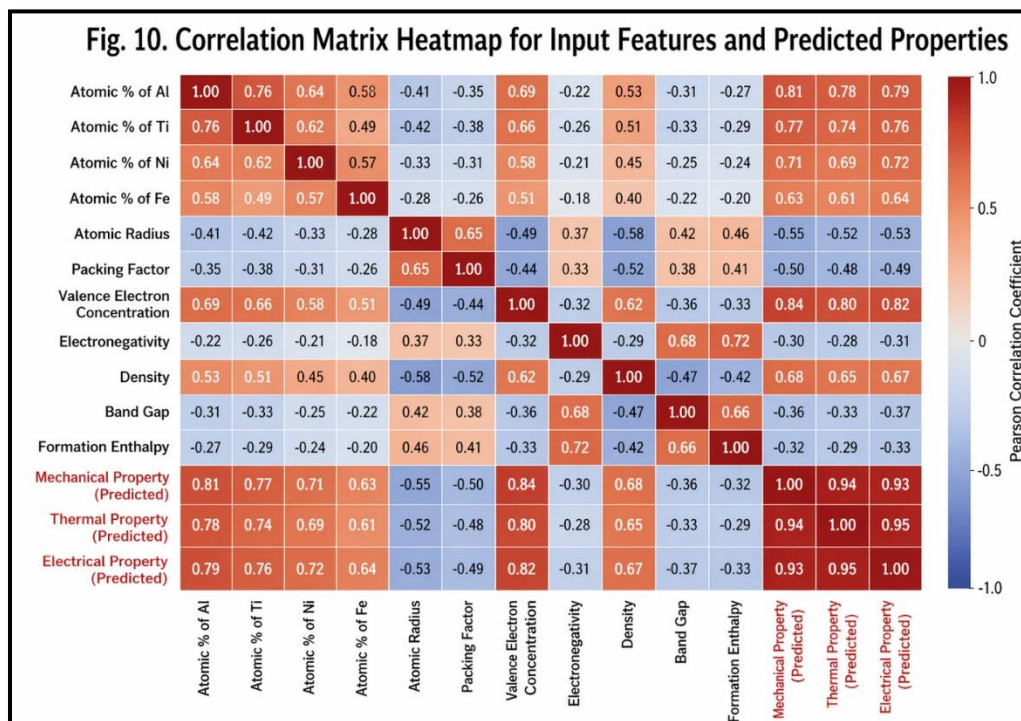
<http://amresearchreview.com/index.php/Journal/about>

Fig. 10 shows the correlation between different material properties and features. Strong correlations among mechanical, thermal, and electrical properties highlight the importance of multi-output modeling and justify the use of a unified prediction framework.

Fig. 10 presents the correlation matrix heatmap illustrating the relationships among input features and predicted mechanical, thermal, and electrical properties. The heatmap reveals strong positive correlations between the predicted properties themselves (correlation coefficients ≈ 0.93 – 0.95), indicating that these properties are closely interrelated and can be effectively modeled within a unified multi-property framework.

Among the input features, compositional variables such as atomic percentages of Al, Ti, Ni, and Fe show strong positive correlations with all three predicted properties (typically in the range of 0.60 – 0.80), confirming their dominant influence. Similarly, valence electron concentration and density exhibit high positive correlations, particularly with mechanical and electrical properties, highlighting their importance in governing material behavior. In contrast, structural features such as atomic radius and packing factor display moderate to strong negative correlations (approximately -0.50 to -0.58), suggesting an inverse relationship with property performance. Electronic descriptors like bandgap and electronegativity show mixed correlations, reflecting their more complex and property-specific influence.

Moreover, the correlation patterns observed in Fig. 10 validate the effectiveness of the proposed multi-property prediction framework by demonstrating that key features consistently influence multiple properties. The strong interdependence among

predicted outputs further justifies the use of a unified modeling approach, enabling improved prediction accuracy and deeper insight into structure–property relationships in advanced materials.

5. Conclusion

This study presented an AI- and machine learning–based framework for predicting mechanical, thermal, and electrical properties of advanced materials. By integrating multi-source data, feature engineering, and multiple predictive models, the proposed approach effectively overcomes limitations of traditional methods and enables accurate multi-property prediction within a unified framework.

The results show that ensemble models, particularly XGBoost, consistently outperform other techniques by achieving lower error values and higher prediction accuracy. The strong agreement between actual and predicted values, along with well-distributed error patterns, confirms the robustness and generalization capability of the model. In addition, feature importance analysis highlights the dominant role of compositional parameters, while structural and electronic features provide complementary insights into material behavior.

Furthermore, the observed correlations among predicted properties validate the effectiveness of the multi-output learning approach, demonstrating its suitability for complex materials systems. Overall, the proposed framework provides a reliable and scalable solution for materials property prediction and has strong potential to support data-driven materials design and next-generation engineering applications.

Future work will focus on incorporating larger and more diverse datasets, as well as exploring advanced deep learning models to further improve prediction accuracy and extend the applicability of the framework.

References

1. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. *Nature*, 559(7715), 547–555.
2. Dunn, A., Wang, Q., Ganose, A., Dopp, D., & Jain, A. (2020). Benchmarking materials property prediction methods: The Matbench test set. *npj Computational Materials*, 6(1), 138.
3. Himanen, L., Geurts, A., Foster, A. S., & Rinke, P. (2021). Data-driven materials science: Status, challenges, and perspectives. *Advanced Science*, 6(21), 1900808.
4. Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., & Persson, K. A. (2013). The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1), 011002.
5. Ahmad, B., Jilani, S. A., Rafique, M., & Soomro, A. A. (2026, January). Design and Monitoring of a Tree-Inspired Vertical Axis Wind Turbine for Efficient Urban Wind Utilization. In 2026 1st International Conference on Innovations in Information and Communication Technologies (IICT) (pp. 1-6). IEEE.
6. Lookman, T., Alexander, F. J., & Rajan, K. (2021). *Information science for*

- materials discovery and design*. Springer.
7. Ahmad, B., Majeed, M. K., Soomro, A. A., & Jillani, S. A. (2026, January). Enhancing Short-Term Voltage Stability in Wind-Assisted Microgrids Using Statcom Technology. In *2026 1st International Conference on Innovations in Information and Communication Technologies (IICT)* (pp. 1-6). IEEE.
 8. Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Müller, K.-R. (2021). Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE*, 109(3), 247–278.
 9. Ahmad, B., Ali, S., Muneer, M., Fatima, A., & Abbas, N. (2025). Wind energy integration into the SAARC region: A comprehensive review of optimal wind sites for a sustainable super smart grid. *Wind Energy*, 3(2).
 10. Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., & Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI Open*, 1, 57–81.
 11. Agrawal, A., & Choudhary, A. (2016). Perspective: Materials informatics and big data: Realization of the “fourth paradigm” of science in materials science. *APL Materials*, 4(5), 053208.
 12. Chen, C., Ye, W., Zuo, Y., Zheng, C., & Ong, S. P. (2020). Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, 31(9), 3564–3572.
 13. Goodall, R. E. A., & Lee, A. A. (2020). Predicting materials properties without crystal structure: Deep representation learning from stoichiometry. *Nature Communications*, 11, 6280.
 14. Ward, L., Liu, R., Krishna, A., Hegde, V. I., Agrawal, A., Choudhary, A., & Wolverton, C. (2020). Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations. *Physical Review B*, 96(2), 024104.
 15. Xie, T., & Grossman, J. C. (2020). Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical Review Letters*, 120(14), 145301.
 16. Majeed, M. K., Usman, M., Hussain, I., Javed, U., Khan, M. Q., Nadeem, F., ... & Zhang, Y. (2024, December). Super Broad Non-Hermitian Line Shape from Out-of-Phase and In-Phase Photon-Phonon Dressing in Eu³⁺: NaYF₄ and Eu³⁺: BiPO₄. In *Photonics* (Vol. 11, No. 12, p. 1169).
 17. Ahmad, B., Majeed, M. K., Soomro, A. A., & Jillani, S. A. (2026, January). Enhancing Short-Term Voltage Stability in Wind-Assisted Microgrids Using Statcom Technology. In *2026 1st International Conference on Innovations in Information and Communication Technologies (IICT)* (pp. 1-6).
 18. Soomro, A. A., Noreen, A., Naz, S., Arshad, J. A., Majeed, M. K., Rafique, N., ... & Ahmad, B. (2025). Data-driven predictive maintenance of diesel engines using advanced machine learning and AI-based regression algorithms for accurate fault detection and real-time condition monitoring. *Spectrum of engineering sciences*, 408-429
 19. Akbar, K. (2022). Artificial Intelligence Apps for COVID-19 virus. *Academia Letters*, 2
 20. Akbar, K. (2021). How economic sustainability is affected by innovation

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

- performance and sustainable manufacturing. *World J. Adv. Res. Rev*, 11(01), 247-255
21. Jablonka, K. M., Ongari, D., Moosavi, S. M., Smit, B., & Laino, T. (2021). Big-data science in porous materials: Materials genomics and machine learning. *Chemical Reviews*, 121(13), 8066–8129.
 22. Rajan, K. (2022). Materials informatics: The materials “gene” and big data. *Annual Review of Materials Research*, 45, 153–169.
 23. Isayev, O., Oses, C., Toher, C., Gossett, E., Curtarolo, S., & Tropsha, A. (2020). Universal fragment descriptors for predicting properties of inorganic crystals. *Nature Communications*, 8, 15679.
 24. Chollet, F. (2021). *Deep learning with Python* (2nd ed.). Manning Publications.
 25. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, E. (2020). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
 26. Alkhatib, O. J., Talib, R., Ali, A., Siddique, M. E., Ahmed, S., & Ali, M. (2025). Machine learning models for predicting at-risk students: A comparative study of classification techniques. *Spectrum of Engineering Sciences*, 1091–1099.