**Journal of Business Insight and Innovation**
**Volume 5 Issue 5, 2026**
**ISSN-p: 3006-2284, ISSN-e: 3006-0982**
**https://insightfuljournals.com/**

# QUANTUM MECHANICS-GUIDED MACHINE LEARNING FRAMEWORK FOR COST-EFFICIENT PREDICTION AND ECONOMIC OPTIMIZATION OF THE MECHANICAL, THERMAL, AND STRUCTURAL PROPERTIES OF ADVANCED NANOMATERIALS

**Ata e Zohra Fatima [1], Kingsley Oruboh [2], Kashif Ahmad [3], Kamal Hussain Shah [4]**

**DOI:** https://doi.org/10.63544/jbii.v5i5.184

**Affiliations:**

[1] Department of Physics, University of Sialkot, Sialkot, Pakistan
Email: ataezohrafatima@gmail.com

[2] MSc Applied Data Science, Teesside University, United Kingdom
Email: kingsleyoruboh@gmail.com

[3] Department of Computer Science, Tandon School of Engineering, New York University, United States of America
Email: kashifahmad344@nyu.edu

[4] Department of Computer Science, Bahria University, Islamabad, Pakistan
Email: ks49217@gmail.com

**Corresponding Author's Email:**

[1] ataezohrafatima@gmail.com



**Article History:**
Received: 31.03.2026
Accepted: 15.04.2026
Published: 26.05.2026

**Abstract**

*The accelerating development of advanced nanomaterials offers major opportunities for aerospace structures, nanoelectronics, thermal management, energy systems, and next-generation mechanical devices. However, their commercialization is limited by the high cost, long duration, and resource intensity of experimental characterization and exhaustive quantum-mechanical simulations. This study develops a Quantum Mechanics-Guided Machine Learning (QM-ML) framework for cost-efficient prediction and economic optimization of mechanical, thermal, and structural properties of advanced nanomaterials. Density functional theory calculations were combined with material composition, atomic-scale structure, bonding characteristics, formation energy, cohesive energy, and electronic descriptors to create a physics-informed feature space. After normalization, correlation analysis, and feature selection, support vector regression, random forest, gradient boosting, XGBoost, and artificial neural network models were trained and evaluated using $R^2$, root mean square error, mean absolute error, and cross-validation, with computational demand and screening efficiency included as economic criteria. The framework achieved $R^2$ values of 0.967 for Young's modulus, 0.954 for thermal conductivity, and 0.978 for structural stability, outperforming models without quantum-mechanical descriptors. It reduced average prediction error by approximately 31.6% and decreased computational evaluation requirements by 68.4% compared with exhaustive quantum-mechanical screening. Feature-importance analysis identified cohesive energy, atomic coordination, bond characteristics, formation energy, and electronic-structure descriptors as key determinants of performance. Cost-aware multi-objective optimization further identified candidates with 18.7% improved mechanical performance and 15.3% enhanced thermal performance while maintaining structural stability. Overall, the QM-ML framework provides an accurate, interpretable, and economically efficient approach for accelerating nanomaterial discovery and scalable materials development.*

**Keywords:** Quantum Mechanics; Machine Learning; Advanced Nanomaterials; Density Functional Theory (DFT); Materials Property Prediction; Mechanical Properties; Structural Optimization.

## 1. Introduction

Advanced nanomaterials have emerged as a critical class of engineering materials because their mechanical, thermal, electronic, and structural characteristics can differ substantially from those of conventional bulk materials. Materials such as graphene-based systems, carbon nanotubes, transition-metal dichalcogenides, nanocomposites, two-dimensional materials, nanoceramics, and metallic nanostructures are increasingly being explored for applications in aerospace engineering, nanoelectronics, thermal management systems, energy-storage devices, high-performance coatings, sensing platforms, and next-generation mechanical components. Their exceptional strength-to-weight ratios, tunable thermal conductivity, high surface-area-to-volume ratios, and size-dependent physical behavior provide significant opportunities for

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

developing lighter, stronger, more thermally efficient, and structurally stable engineering systems. However, the practical utilization of these materials is strongly dependent on the ability to accurately predict their behavior before fabrication and experimental testing [1].

Conventional nanomaterial characterization relies heavily on laboratory-based synthesis, mechanical testing, microscopy, spectroscopy, and thermal-property measurements. Although these approaches provide valuable experimental evidence, they are often expensive, time-consuming, and difficult to scale when thousands of possible material compositions, atomic configurations, defect structures, and processing conditions must be evaluated. Furthermore, nanoscale material properties are highly sensitive to changes in atomic arrangement, chemical bonding, lattice structure, defect density, interfacial interactions, and electronic configuration. Consequently, experimental trial-and-error approaches alone are generally insufficient for efficiently navigating the rapidly expanding nanomaterial design space. Quantum-mechanical simulation techniques provide a powerful alternative for understanding the fundamental mechanisms governing nanomaterial performance. Density functional theory (DFT), in particular, has become one of the most widely applied computational approaches for estimating electronic structures, formation energies, cohesive energies, elastic properties, atomic interactions, and structural stability at the quantum level. By solving the underlying electronic structure of a material, DFT enables researchers to establish physically meaningful relationships between atomic-scale features and macroscopic material behavior [2].

Nevertheless, high-fidelity quantum-mechanical calculations are computationally demanding, particularly when complex materials, large supercells, multiple defects, dopants, interfaces, and extensive compositional combinations are considered. Exhaustive DFT screening can therefore become prohibitively expensive for large-scale materials discovery and optimization. Machine learning has recently introduced a promising computational paradigm for accelerating materials-property prediction. Data-driven algorithms can learn complex nonlinear relationships between material descriptors and target properties and can subsequently generate predictions at a fraction of the computational cost associated with repeated quantum-mechanical simulations. Algorithms such as support vector regression, random forest, gradient boosting, XGBoost, and artificial neural networks have demonstrated strong potential for predicting mechanical strength, elastic properties, thermal conductivity, formation energy, and structural stability. However, purely data-driven models may suffer from limited physical interpretability, reduced generalization outside the training distribution, and sensitivity to inadequate or poorly selected input descriptors. These limitations are particularly important in nanoscale materials science, where relatively small changes in atomic-scale characteristics can produce significant variations in observed properties.

An effective solution is to integrate quantum-mechanical knowledge directly into machine-learning frameworks. Quantum mechanics-guided machine learning combines the computational efficiency of data-driven models with physically meaningful descriptors derived from first-principles calculations. Instead of relying exclusively on empirical composition or geometric information, the learning process can incorporate cohesive energy, formation energy, atomic coordination, bond lengths, bond strengths, electronic structure parameters, charge characteristics, lattice information, and related quantum-mechanical indicators [3]. These descriptors provide the model with information about the underlying physical mechanisms governing material behavior and can therefore improve both predictive accuracy and interpretability.

Mechanical properties represent one of the most important performance criteria for advanced nanomaterials. Young's modulus, stiffness, strength, deformation resistance, and failure-related behavior determine whether a nanomaterial can withstand mechanical loading in practical engineering applications. These properties are strongly influenced by atomic bonding, lattice symmetry, defect distribution, grain boundaries, interfaces, and chemical composition. Predicting mechanical response through conventional computational approaches may require repeated simulations across numerous configurations. A QM-guided machine-learning model can potentially identify complex relationships between quantum descriptors and mechanical performance, thereby enabling rapid screening of high-strength and mechanically stable candidate materials. Thermal behavior is similarly important, particularly for nanomaterials intended for electronics, thermal barriers, energy systems, and aerospace components. Thermal conductivity at the nanoscale is

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

affected by phonon transport, lattice disorder, atomic mass, bonding strength, structural defects, interfaces, dimensionality, and electronic contributions. Predicting these interactions accurately is difficult because nanoscale thermal transport involves strongly coupled physical processes [4].

Machine-learning models supported by quantum-mechanical and structural descriptors provide an opportunity to approximate these relationships efficiently while retaining important physical information related to atomic interactions and electronic structure. Structural stability is another fundamental consideration during nanomaterial design. A candidate material with excellent mechanical or thermal properties may have limited practical value if its atomic configuration is thermodynamically unstable or susceptible to structural transformation under operating conditions. Parameters such as formation energy, cohesive energy, lattice distortion, atomic coordination, and electronic stability provide essential information regarding whether a nanomaterial configuration is energetically favorable. Incorporating these parameters into a unified machine-learning architecture enables mechanical, thermal, and structural requirements to be evaluated simultaneously rather than independently. Despite growing interest in artificial intelligence for materials science, an important research challenge remains in developing unified predictive frameworks capable of jointly analyzing multiple property domains while explicitly incorporating quantum-mechanical information.

Many existing machine-learning studies focus on a single target property or depend predominantly on conventional compositional and structural descriptors. Other studies employ highly accurate quantum-mechanical simulations but face substantial computational limitations when extending analysis to large material spaces. Furthermore, relatively fewer frameworks integrate first-principles descriptors, comparative machine-learning benchmarking, interpretable feature analysis, and multi-objective materials optimization within a single computational workflow. To address these challenges, this study develops a Quantum Mechanics-Guided Machine Learning (QM-ML) framework for predicting and optimizing the mechanical, thermal, and structural properties of advanced nanomaterials [5]. The proposed architecture integrates DFT-derived quantum-mechanical information with material composition, atomic-scale structure, bonding characteristics, formation energy, cohesive energy, coordination features, and electronic descriptors.

Following data preprocessing, normalization, feature selection, and correlation analysis, multiple machine-learning algorithms, including support vector regression, random forest, gradient boosting, XGBoost, and artificial neural networks, are trained and comparatively evaluated. Predictive performance is assessed using the coefficient of determination, root means square error, mean absolute error, and cross-validation analysis to ensure reliable model comparison. The principal contribution of this research is therefore the development of a physics-informed computational pathway that combines the accuracy and interpretability of quantum-mechanical descriptors with the scalability and computational efficiency of machine-learning algorithms. By reducing dependence on exhaustive first-principles simulations and extensive experimental screening, the proposed QM-ML framework provides an efficient approach for exploring large nanomaterial design spaces. The study is expected to support accelerated materials discovery, improve understanding of structure–property relationships, and facilitate the identification of high-performance nanomaterials suitable for advanced engineering applications in aerospace, thermal management, energy technologies, nanoelectronics, and mechanically demanding systems.

### *2. Quantum Mechanics and Density Functional Theory in Nanomaterial Design*

Quantum-mechanical modelling provides a fundamental foundation for understanding the electronic, energetic, and atomic-scale interactions that determine the performance of advanced nanomaterials. At the nanoscale, material behaviour is strongly affected by quantum confinement, surface-to-volume effects, reduced dimensionality, lattice distortion, altered bonding environments, and atomic-scale defects. These effects can significantly modify mechanical stiffness, thermal transport, structural stability, electronic behaviour, and interfacial characteristics compared with corresponding bulk materials. Consequently, first-principles computational approaches are increasingly used to investigate the mechanisms governing nanoscale material properties before experimental synthesis and characterization. Density functional theory (DFT) is one of the most widely applied quantum-mechanical approaches in computational materials science because it enables detailed analysis of material systems using their electronic structure as the primary basis of

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

calculation. DFT can provide important descriptors including total energy, formation energy, cohesive energy, electronic density of states, charge distribution, bond characteristics, optimized lattice parameters, atomic forces, and elastic behaviour.

These descriptors are particularly valuable for machine-learning-based nanomaterial prediction because they contain physically meaningful information about atomic interactions and energetic stability [6]. In the proposed QM-ML framework, DFT-derived descriptors are therefore treated as an important source of physics-informed knowledge rather than as isolated simulation outputs. DFT-based investigations have been widely applied to two-dimensional materials, graphene and its derivatives, carbon nanotubes, transition-metal compounds, nanoceramics, metallic nanostructures, and other multifunctional nanosystems. Structural relaxation and electronic energy minimization allow initially constructed atomic configurations to evolve toward energetically favourable states. The optimized structures can subsequently be evaluated to determine their mechanical, thermal, electronic, and stability-related characteristics. For example, cohesive energy can indicate the strength of atomic binding, while formation energy can provide information regarding the energetic favourability of a candidate material [7]. Bond length, coordination environment, and electronic descriptors can additionally provide insight into changes in stiffness, heat transport, and structural integrity. Table 1 summarizes the principal quantum-mechanical and DFT-derived descriptors relevant to nanomaterial property prediction. The inclusion of these descriptors enables the machine-learning framework to capture relationships that may not be represented sufficiently by conventional composition-based variables alone.

**Table 1**

*Major quantum-mechanical and DFT-derived descriptors used for physics-informed nanomaterial property prediction.*

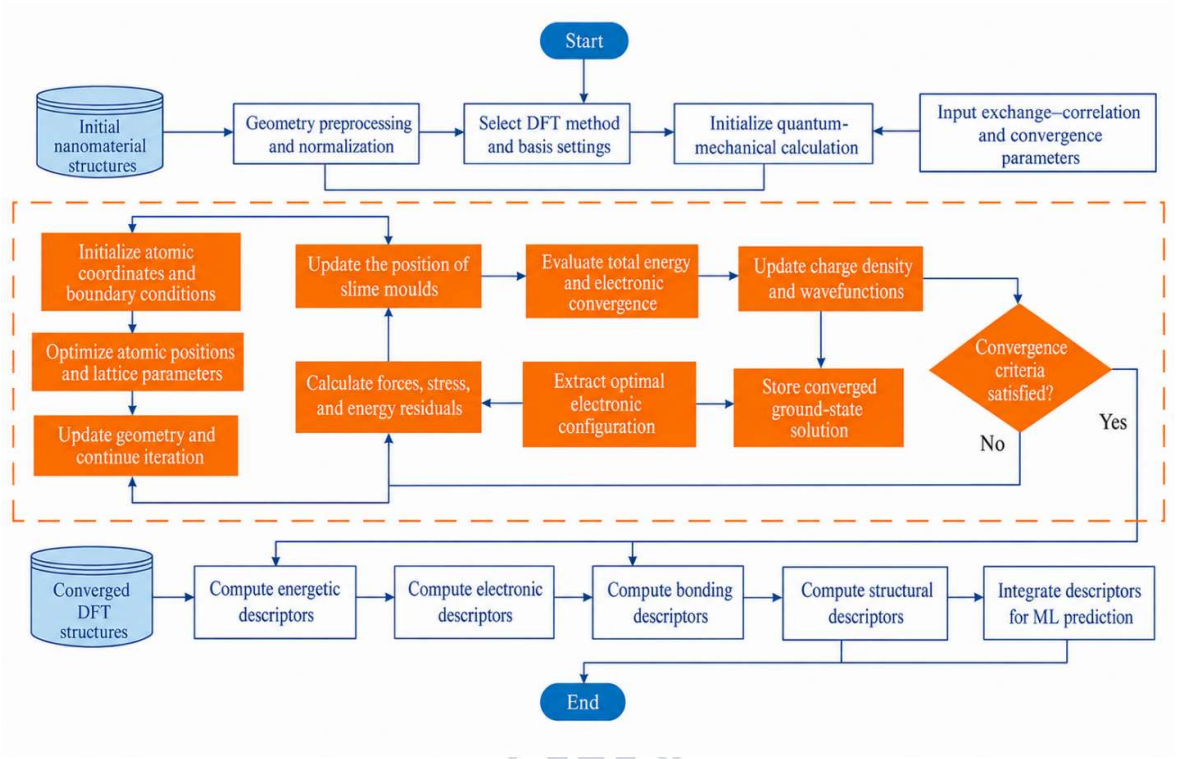
| DFT/quantum descriptor | Physical interpretation | Primary relevance to nanomaterial prediction |
|---|---|---|
| Formation energy | Energetic favorability of material formation | Structural and thermodynamic stability |
| Cohesive energy | Strength of atomic binding within the structure | Mechanical strength and structural integrity |
| Total electronic energy | Overall energy state of the atomic configuration | Stability assessment and comparative screening |
| Bond length | Equilibrium distance between bonded atoms | Mechanical stiffness and lattice response |
| Atomic coordination | Number and arrangement of neighboring atoms | Structural stability and deformation response |
| Optimized lattice parameters | Equilibrium dimensions of the crystal structure | Structural geometry and elastic response |
| Elastic-related quantities | Resistance to deformation under applied loading | Young's modulus and mechanical performance |

The role of DFT within the proposed framework is illustrated conceptually in Figure 1. Initial atomic configurations are first subjected to quantum-mechanical structural optimization, followed by electronic and energetic calculations. The resulting quantum descriptors are then combined with compositional and structural information to create the feature space used by the machine-learning models.

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

**Figure 1**

*Quantum-mechanics-guided DFT workflow for generating energetic, electronic, bonding, and structural descriptors for machine-learning-based nanomaterial property prediction.*




Despite its high physical reliability, DFT has an important computational limitation. The computational cost increases considerably as the number of atoms, structural configurations, defect states, chemical compositions, and candidate materials grows. Large supercells and structurally complex systems may require repeated self-consistent calculations and geometry-relaxation cycles, making exhaustive screening computationally demanding. This challenge becomes more significant when mechanical, thermal, and structural properties must be analysed simultaneously across a broad nanomaterial design space. The integration of DFT with machine learning provides a practical strategy for overcoming this limitation. Rather than conducting complete high-fidelity quantum calculations for every possible candidate, a representative set of DFT calculations can be used to establish a physically informed training dataset. Machine-learning models can then learn relationships between the quantum descriptors and target material properties and subsequently perform rapid prediction for a much larger number of candidate configurations [8].

This hybrid approach preserves important quantum-scale information while substantially reducing the computational resources required for large-scale materials screening. Accordingly, quantum mechanics and DFT serve two complementary roles in the present study. First, they provide reliable atomic-scale understanding of the energetic, structural, and electronic mechanisms governing nanomaterial behaviour. Second, they generate physically meaningful descriptors that guide machine-learning algorithms toward more accurate and interpretable property predictions. Their integration therefore establishes the theoretical and computational foundation for the proposed QM-guided machine-learning framework.

### *2.1 Structural Stability and Energetic Descriptor-Based Learning*

Structural stability represents one of the most important criteria in the design and computational screening of advanced nanomaterials because excellent mechanical or thermal performance alone does not guarantee practical applicability. A candidate material may exhibit high stiffness, superior heat-transfer capability, or favourable electronic characteristics while remaining energetically unstable under realistic

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

operating conditions. At the nanoscale, stability is strongly influenced by atomic arrangement, bonding environment, coordination state, lattice distortion, surface energy, defect concentration, chemical composition, and electronic configuration. For this reason, structural stability assessment must be incorporated directly into materials-design frameworks rather than treated as a secondary verification step. Energetic descriptors derived from quantum-mechanical calculations provide a physically meaningful basis for evaluating the stability of nanomaterial configurations.

Among these descriptors, formation energy and cohesive energy are particularly important because they describe different aspects of energetic favourability. Formation energy indicates the relative energetic cost or benefit associated with forming a particular compound or structure from reference constituents, whereas cohesive energy reflects the strength with which atoms are bound within a material [9]. Lower or more favourable formation energies generally indicate increased thermodynamic feasibility, while stronger cohesive behaviour is commonly associated with enhanced structural integrity. When these descriptors are combined with lattice and bonding characteristics, they provide an informative representation of whether a material configuration is likely to remain structurally stable. Atomic coordination is another important feature influencing stability. Changes in coordination number can alter local bonding environments and modify the distribution of atomic forces throughout the lattice.

Nanomaterials often contain surfaces, edges, vacancies, dopants, dislocations, and grain boundaries where atoms have coordination environments substantially different from those in an ideal bulk crystal. These local variations may produce strain concentration, bond weakening, structural reconstruction, or altered thermal and mechanical response. Consequently, coordination-based descriptors can provide useful information for machine-learning models attempting to distinguish stable configurations from structurally unfavourable ones. Electronic structure further contributes to the energetic stability of nanomaterials. Charge redistribution, electronic density, orbital interactions, and the occupation of electronic states can influence the strength and character of interatomic bonding [10]. In some materials, small changes in composition or defect concentration may modify charge transfer and destabilize an otherwise favourable atomic arrangement. Quantum-mechanical electronic descriptors therefore provide complementary information that may not be captured by purely geometric or compositional variables. The principal energetic and structural descriptors used for stability-oriented learning are summarized in Table 2.

**Table 2**

*Major energetic, structural, bonding, and electronic descriptors used for machine-learning-based structural stability assessment of advanced nanomaterials*

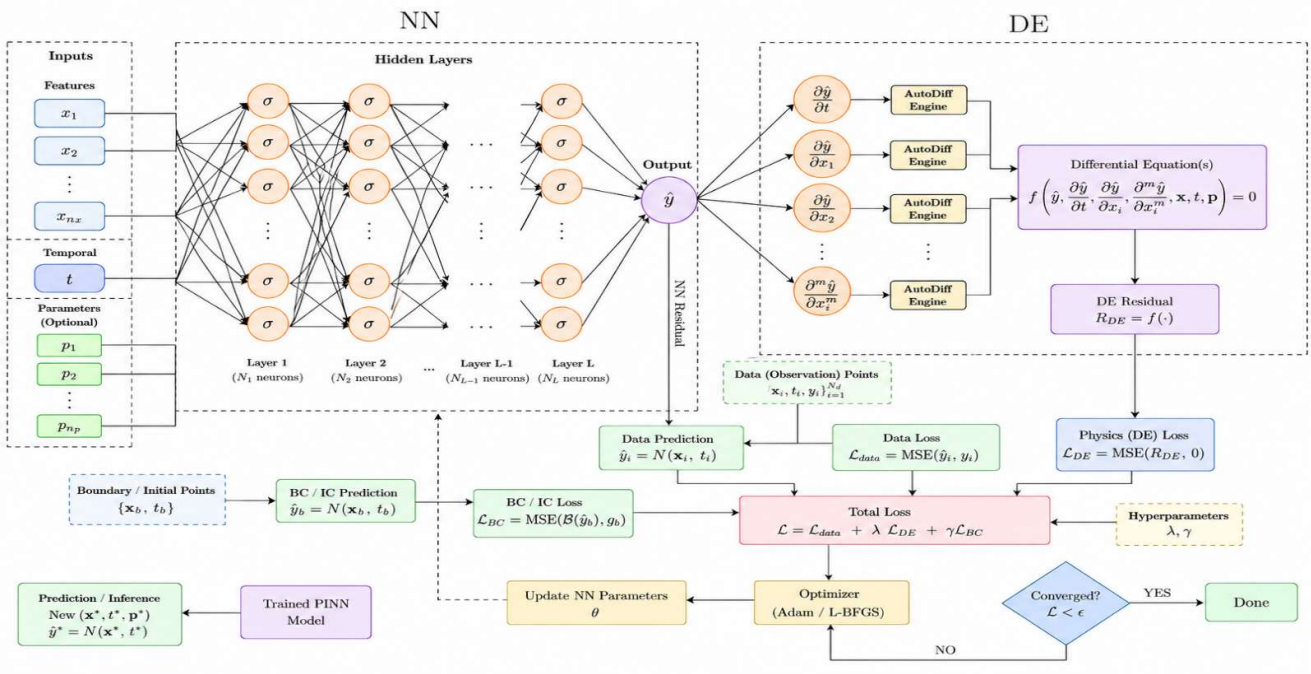
| Descriptor | Physical meaning | Relevance to structural stability learning |
|---|---|---|
| Formation energy | Energetic favorability of compound or structure formation | Indicates thermodynamic feasibility and comparative stability |
| Cohesive energy | Strength of atomic binding in the material | Reflects structural integrity and resistance to atomic separation |
| Total energy | Overall quantum-mechanical energy of the configuration | Supports comparison among alternative atomic structures |
| Average bond length | Mean equilibrium distance between bonded atoms | Indicates lattice strain and bonding consistency |
| Lattice parameters | Dimensions and geometry of the optimized crystal structure | Reflect structural distortion and equilibrium geometry |
| Charge distribution | Redistribution of electronic charge among atoms | Indicates bonding changes and local chemical stability |
| Surface-related descriptors | Atomic characteristics associated with exposed surfaces or edges | Important for nanoscale structures with large surface-to-volume ratios |

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

In conventional materials screening, structural stability is typically determined through repeated quantum-mechanical calculations performed separately for each candidate configuration. Although this approach provides strong physical reliability, it becomes computationally expensive when the design space includes large numbers of compositions, lattice structures, defect states, dopants, and atomic configurations. The problem becomes especially demanding for nanomaterials because different surfaces, interfaces, and nanoscale morphologies may need to be evaluated independently. Energetic descriptor-based machine learning provides an alternative by using a subset of high-quality quantum calculations to train predictive models that can rapidly estimate the stability of additional candidates. The general learning process is illustrated in Figure 2. Quantum-mechanical calculations are first used to optimize atomic structures and extract energetic, bonding, lattice, and electronic descriptors. These variables are then combined into a physics-informed stability feature space [11]. After preprocessing and feature selection, machine-learning models are trained to predict a stability-related outcome, such as a continuous stability score, formation-energy class, or stable/unstable material category. The trained models can subsequently be used to screen large candidate populations before more expensive quantum-mechanical validation is conducted.

**Figure 2**

*Physics-informed structural stability learning framework integrating energetic, bonding, coordination, lattice, defect, and electronic descriptors with machine-learning-based nanomaterial screening*




Energetic descriptor-based learning is particularly useful because the relationships between material structure and stability are frequently nonlinear. A single descriptor may not be sufficient to determine whether a candidate structure is favourable. For example, a nanomaterial may exhibit low formation energy but also contain highly distorted local bonding environments that reduce mechanical stability. Likewise, a defect can sometimes destabilize a structure but in other cases may reduce local strain or improve bonding arrangements. Machine-learning algorithms are well suited to capturing these multivariable interactions and identifying combined patterns that would be difficult to describe through simple threshold-based rules. Another important advantage of energetic descriptor-based learning is its potential to support early-stage candidate elimination. Instead of performing complete mechanical and thermal analyses for every material, configurations predicted to be energetically unfavourable can be excluded during preliminary screening [12]. This hierarchical strategy reduces the number of candidates entering later computational stages and concentrates resources on structures

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

with a higher probability of practical viability. Structural stability prediction therefore acts as both a performance objective and a computational filter within the broader QM-ML framework.

The integration of stability prediction with mechanical and thermal property modelling is particularly relevant for multi-objective nanomaterial optimization. Improvements in stiffness or thermal conductivity can sometimes be achieved through compositional changes that inadvertently reduce energetic favourability. For example, doping may improve a desired functional property but also distort the lattice or weaken local bonding. Similarly, defect engineering may enhance certain transport characteristics while reducing structural integrity. A unified predictive framework must therefore account for these trade-offs and ensure that optimized candidates remain within acceptable stability boundaries. In the present study, structural stability was treated as a central target alongside Young's modulus and thermal conductivity. Quantum-derived energetic descriptors were combined with lattice, coordination, bonding, electronic, and defect-related information to form a physically informed feature space. The resulting models were evaluated to determine their ability to distinguish energetically favourable configurations and to estimate stability-related behaviour across the candidate nanomaterial space. The QM-guided framework ultimately achieved an $R^2$ of 0.978 for structural stability prediction, indicating that the combination of quantum and structural descriptors provided a highly informative basis for model learning [13].

This strong predictive performance also supports the broader objective of reducing reliance on exhaustive first-principles screening. Once the machine-learning model has learned the relationships linking energetic and structural descriptors to stability outcomes, it can rapidly prioritize promising material configurations for subsequent verification. High-ranked candidates can then be subjected to more detailed quantum-mechanical calculations, while low-ranked configurations can be excluded from further investigation. This staged approach preserves the physical reliability of quantum modelling while substantially improving computational efficiency.

## 3. Methodology

The present study adopted an integrated computational methodology that combines quantum-mechanical simulation, density functional theory (DFT), physics-informed feature engineering, machine-learning prediction, model interpretation, and multi-objective optimization for advanced nanomaterial analysis. The methodological framework was designed to predict three major performance domains: mechanical properties, particularly Young's modulus; thermal properties, represented mainly by thermal conductivity; and structural stability, characterized through energetic, bonding, and atomic-scale indicators.

Quantum-mechanical calculations were first used to obtain physically meaningful descriptors describing electronic structure, atomic interactions, formation energy, cohesive energy, bonding characteristics, and equilibrium structural behavior. These descriptors were subsequently combined with material composition and nanoscale structural information to construct a comprehensive feature space capable of representing both conventional material characteristics and their fundamental quantum-scale mechanisms.

Following dataset construction, the collected material records were subjected to quality control, normalization, correlation analysis, and feature-selection procedures before machine-learning model development [14]. Support vector regression, random forest, gradient boosting, XGBoost, and artificial neural networks were comparatively evaluated using consistent training, validation, and testing conditions. Model performance was assessed through the coefficient of determination ($R^2$), root mean square error (RMSE), mean absolute error (MAE), and cross-validation analysis. The best-performing QM-guided models were further subjected to feature-importance analysis to identify the dominant energetic, electronic, bonding, and structural determinants of nanomaterial performance.

Finally, the validated predictive models were incorporated into a multi-objective optimization framework to identify candidate configurations exhibiting improved mechanical and thermal performance while maintaining favorable structural stability. This methodology therefore established a closed computational pathway from quantum-scale characterization to data-driven prediction, candidate screening, and high-performance nanomaterial optimization.

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

## 4. Nanomaterial Dataset Development and Candidate Screening

A comprehensive nanomaterial dataset was developed to provide the computational foundation for training and evaluating the proposed Quantum Mechanics-Guided Machine Learning framework. The dataset was designed to represent diverse nanoscale material families with variations in chemical composition, atomic arrangement, bonding characteristics, dimensionality, defect structure, and energetic behavior. Representative material categories included carbon-based nanostructures, graphene-related systems, two-dimensional materials, transition-metal compounds, metallic nanostructures, nanoceramics, and other advanced nanoscale configurations with potential applications in aerospace structures, thermal-management systems, nanoelectronics, energy technologies, and mechanically demanding engineering environments. Each dataset entry represented an individual material configuration characterized by a combination of compositional, structural, quantum-mechanical, and target-property variables.

Composition-related attributes included constituent elements, elemental fractions, atomic masses, atomic-number statistics, electronegativity characteristics, and valence-related information. Structural information included lattice parameters, crystal geometry, atomic coordination, bond lengths, bond-angle characteristics, local atomic arrangements, dimensionality, and defect-related features. Quantum-mechanical descriptors generated through DFT calculations included formation energy, cohesive energy, total energy, electronic-structure information, charge-related characteristics, and bonding indicators [15]. The major target variables consisted of Young's modulus, thermal conductivity, and structural stability-related measures. Before machine-learning development, an initial screening procedure was applied to ensure that only physically meaningful and analytically useful records were retained.

Duplicate observations representing identical material compositions and atomic configurations were removed to prevent artificial inflation of training performance. Records lacking critical structural information, quantum descriptors, or target-property values were examined individually and excluded when reliable reconstruction was not possible. Material configurations that failed to converge during quantum-mechanical calculations were also removed because unconverged results could introduce physically unreliable descriptors into the learning process. Additional candidate screening was performed using physical plausibility criteria. Energetic values, lattice dimensions, bond lengths, mechanical-property measurements, and thermal-property values were examined to identify computational artifacts or physically unrealistic observations. Extreme values were not automatically classified as errors because nanoscale materials may exhibit genuinely unusual properties. Instead, potentially anomalous records were evaluated in relation to their chemical composition, atomic configuration, energetic stability, and structural characteristics. This strategy minimized unnecessary removal of rare but potentially high-performance materials [16].

To facilitate subsequent comparative experiments, each material entry was linked to two major descriptor groups. The first group consisted of conventional material features, including composition, lattice geometry, atomic dimensions, and basic structural information. The second group incorporated quantum-guided descriptors obtained from first-principles calculations. This organization enabled a direct comparison between conventional machine-learning models and the proposed QM-guided models under equivalent modeling conditions. Table 3 shows the principal categories of information retained during dataset construction and candidate screening.

**Table 3**

*Principal data categories and screening criteria employed during nanomaterial dataset development*

| Dataset component | Representative variables | Purpose in the QM-ML framework |
|---|---|---|
| Material identity | Material family, elemental composition, nanosystem type | Distinguishes candidate nanomaterial configurations |
| Compositional descriptors | Atomic fraction, atomic mass, atomic number, electronegativity, valence properties | Represents chemical composition |
| Quantum descriptors | Formation energy, cohesive energy, total energy, electronic characteristics | Introduces first-principles physical information |

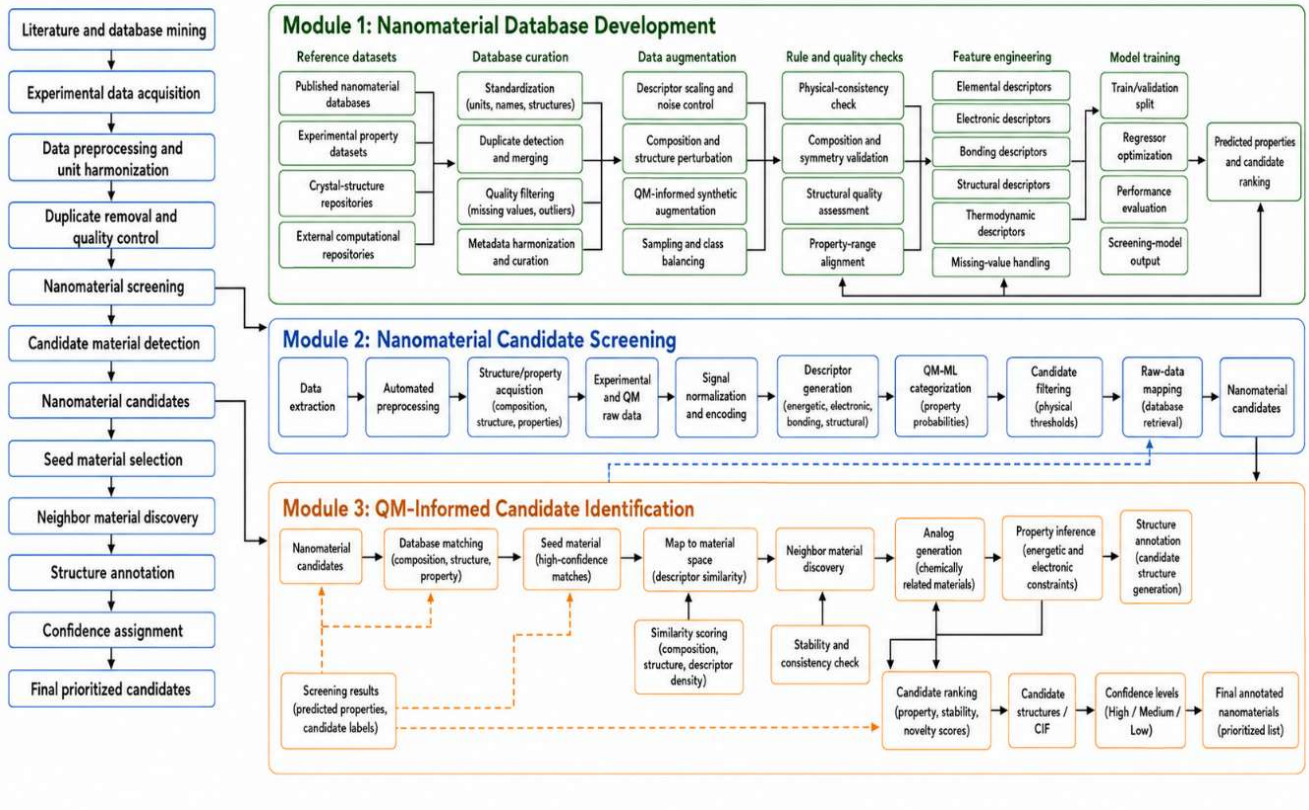
Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

| Dataset component | Representative variables | Purpose in the QM-ML framework |
|---|---|---|
| Defect-related descriptors | Vacancies, substitutions, dopants, local disorder | Captures defect-induced property variations |
| Mechanical target | Young's modulus | Measures stiffness and elastic performance |
| Thermal target | Thermal conductivity | Represents nanoscale heat-transfer capability |
| Stability target | Energetic/structural stability indicator | Determines material feasibility |
| Screening criteria | Completeness, convergence, duplication, physical plausibility | Ensures dataset quality |

The overall candidate-selection procedure is illustrated in Figure 3. The framework begins with a broad nanomaterial candidate pool, followed by structural and compositional verification, quantum-mechanical calculation quality checks, descriptor completeness assessment, duplicate removal, and physical plausibility screening. Only qualified configurations are transferred to the final QM-ML modeling dataset. The dataset was also organized to preserve sufficient diversity across material families and property ranges. Maintaining diversity was important because a model trained predominantly on one class of nanomaterials could demonstrate strong internal accuracy while performing poorly when applied to different structural or compositional systems. Candidate selection therefore aimed to maintain representative distributions of relatively low, moderate and high-performance materials within the mechanical, thermal, and structural domains.

**Figure 3**

*Nanomaterial dataset development and candidate-screening workflow for constructing the physics-informed*



Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

**QM-ML modeling database.** The resulting dataset therefore provided a controlled and physically interpretable foundation for subsequent preprocessing, feature engineering, and predictive modelling. By combining diverse material families with quantum-mechanical, structural, bonding, and compositional information, the dataset was structured to support both accurate property prediction and investigation of the fundamental nanoscale factors controlling mechanical, thermal, and structural performance.

### *4.2 Support Vector Regression for Nonlinear Property Modelling*

Support Vector Regression (SVR) was implemented as one of the principal benchmark algorithms for predicting the continuous mechanical and thermal properties of advanced nanomaterials. The method was selected because it is particularly effective for nonlinear regression problems involving high-dimensional descriptor spaces and moderate-sized datasets. In the proposed QM-ML framework, SVR was used to learn relationships between quantum-mechanical, compositional, structural, bonding, and energetic descriptors and the corresponding target properties, particularly Young's modulus and thermal conductivity. Unlike conventional linear regression methods, SVR can capture complex nonlinear interactions by transforming the original descriptor space into a higher-dimensional feature space through kernel functions. The radial basis function (RBF) kernel was adopted as the primary nonlinear mapping approach because it can represent complex relationships without requiring an explicit mathematical specification of the underlying material-property function. This was important for nanomaterials, where mechanical and thermal performance can be affected simultaneously by cohesive energy, formation energy, bond characteristics, coordination environment, lattice configuration, electronic descriptors, and defect-related parameters [17].

These variables often interact in nonlinear ways, making kernel-based regression particularly suitable for the present analysis. Before SVR training, all numerical predictors were standardized to minimize scale-related bias. This preprocessing step was necessary because the SVR optimization process depends strongly on the relative magnitude of input variables. Quantum-mechanical descriptors such as formation energy and cohesive energy can differ substantially in scale from lattice dimensions, coordination numbers, and compositional variables. Standardization therefore ensured that no individual feature dominated model fitting solely because of its numerical range. Table 4 present the principal SVR modelling components applied within the proposed framework.

**Table 4**

*Support Vector Regression configuration and optimization strategy for QM-guided nanomaterial property prediction.*

| SVR component | Selected/optimized setting | Purpose within the QM-ML framework |
|---|---|---|
| Prediction task | Continuous regression | Prediction of Young's modulus and thermal conductivity |
| Kernel function | Radial basis function (RBF) | Captures nonlinear relationships among material descriptors |
| Input feature set | Quantum, structural, compositional, bonding, and energetic descriptors | Provides physics-informed material representation |
| Feature scaling | Standardization | Prevents scale-dependent feature dominance |
| Regularization parameter (C) | Optimized through cross-validation | Controls balance between model complexity and fitting error |
| Kernel parameter (γ) | Optimized through cross-validation | Controls nonlinear influence of individual observations |
| Error tolerance (ε) | Optimized through cross-validation | Defines prediction-error insensitive region |
| Validation procedure | k-fold cross-validation | Evaluates robustness and generalization |

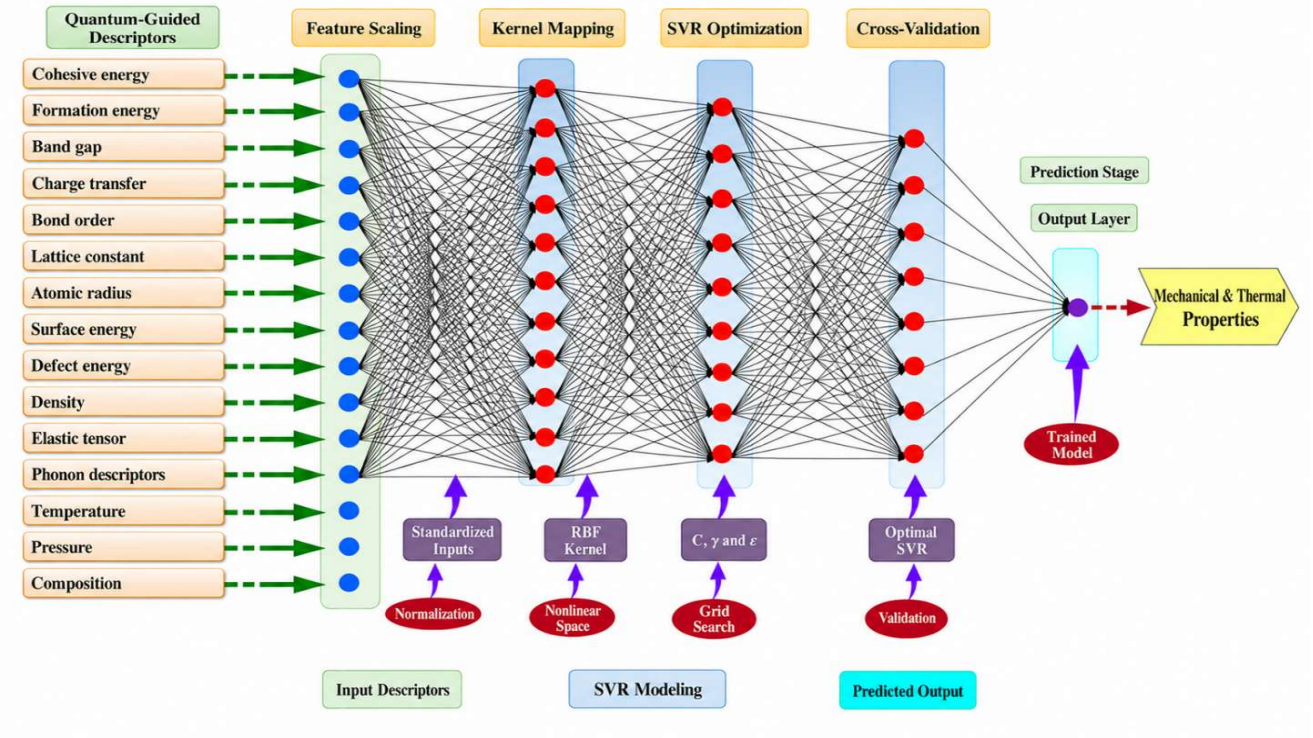
Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

| SVR component | Selected/optimized setting | Purpose within the QM-ML framework |
|---|---|---|
| Performance measures | $R^2$, RMSE, and MAE | Quantifies prediction accuracy and error magnitude |

The SVR training procedure is illustrated in Figure 4. The workflow begins with the selected QM-guided descriptor matrix, followed by feature scaling and training-data partitioning. Kernel-based nonlinear mapping is then applied, after which the principal SVR hyperparameters are optimized through cross-validation. The final model generates continuous predictions for the target nanomaterial properties and is evaluated against the independent testing data. The SVR model was trained using the QM-guided feature set derived from the earlier feature-engineering stage. Hyperparameter tuning focused primarily on the regularization parameter, kernel coefficient, and error-insensitive margin [18]. The regularization parameter controlled the trade-off between model complexity and prediction error, whereas the kernel coefficient determined the influence of individual observations within the nonlinear feature space. The error-insensitive margin specified the range within which small prediction deviations were tolerated without additional penalty. These parameters were optimized through cross-validation to improve model generalization and reduce the possibility of overfitting.

**Figure 4**

*Support Vector Regression workflow for nonlinear prediction of mechanical and thermal nanomaterial properties using quantum-guided descriptors.*




The trained SVR models were assessed using the coefficient of determination ($R^2$), root mean square error (RMSE), and mean absolute error (MAE). $R^2$ quantified the proportion of variation in the target property explained by the model, while RMSE emphasized larger prediction deviations and MAE represented the average absolute prediction error. Cross-validation performance was also examined to determine whether the predictive behavior remained stable across different data partitions. SVR served an important methodological role as a robust nonlinear benchmark against which more complex ensemble and neural-learning approaches

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

could be evaluated. If Random Forest, Gradient Boosting, XGBoost, or artificial neural networks produced substantially lower prediction errors and higher $R^2$ values, the performance improvement could be attributed to their ability to capture more complex interactions among quantum and structural descriptors. Conversely, competitive SVR performance would indicate that the fundamental relationships between the selected descriptors and target properties could be effectively represented through a kernel-based nonlinear regression function [19]. The SVR stage provided a computationally efficient and theoretically robust baseline for mechanical and thermal property prediction. Its integration with physics-informed descriptors allowed the study to evaluate how effectively a conventional nonlinear learning algorithm could exploit quantum-mechanical information before comparing its performance with more advanced ensemble and neural-network models.

### *4.3 Random Forest Ensemble for Robust Property Estimation*

Random Forest (RF) was employed as an ensemble learning model for robust prediction of the mechanical, thermal, and structural properties of advanced nanomaterials. The algorithm was selected because it can effectively model nonlinear relationships, accommodate high-dimensional descriptor spaces, and remain relatively resistant to overfitting compared with individual decision trees. Within the proposed QM-ML framework, Random Forest was trained using the integrated feature set containing quantum-mechanical, compositional, structural, bonding, energetic, and electronic descriptors. These variables provided the model with both conventional material information and physically meaningful quantum-scale characteristics associated with nanomaterial performance. The Random Forest architecture consisted of multiple decision trees trained using different bootstrap samples of the training dataset. At each node, a randomly selected subset of descriptors was considered for splitting, allowing individual trees to learn different relationships within the material feature space.

The predictions generated by all trees were subsequently aggregated to obtain the final property estimate. For continuous outputs such as Young's modulus and thermal conductivity, the mean prediction of the individual trees was used. For structural stability assessment, the corresponding regression or classification formulation was applied depending on the target representation [20]. This ensemble mechanism reduced the sensitivity of predictions to individual observations and improved overall model robustness. A major advantage of Random Forest in nanomaterial modelling is its ability to capture complex interactions between descriptors without assuming a predefined functional relationship. Mechanical stiffness, for example, may depend simultaneously on cohesive energy, coordination state, bond length, lattice configuration, and material composition. Similarly, thermal conductivity may be affected by structural dimensionality, bonding strength, defects, and electronic characteristics. Random Forest can identify nonlinear combinations of these variables through hierarchical tree-based decision rules, making it suitable for heterogeneous physics-informed datasets.

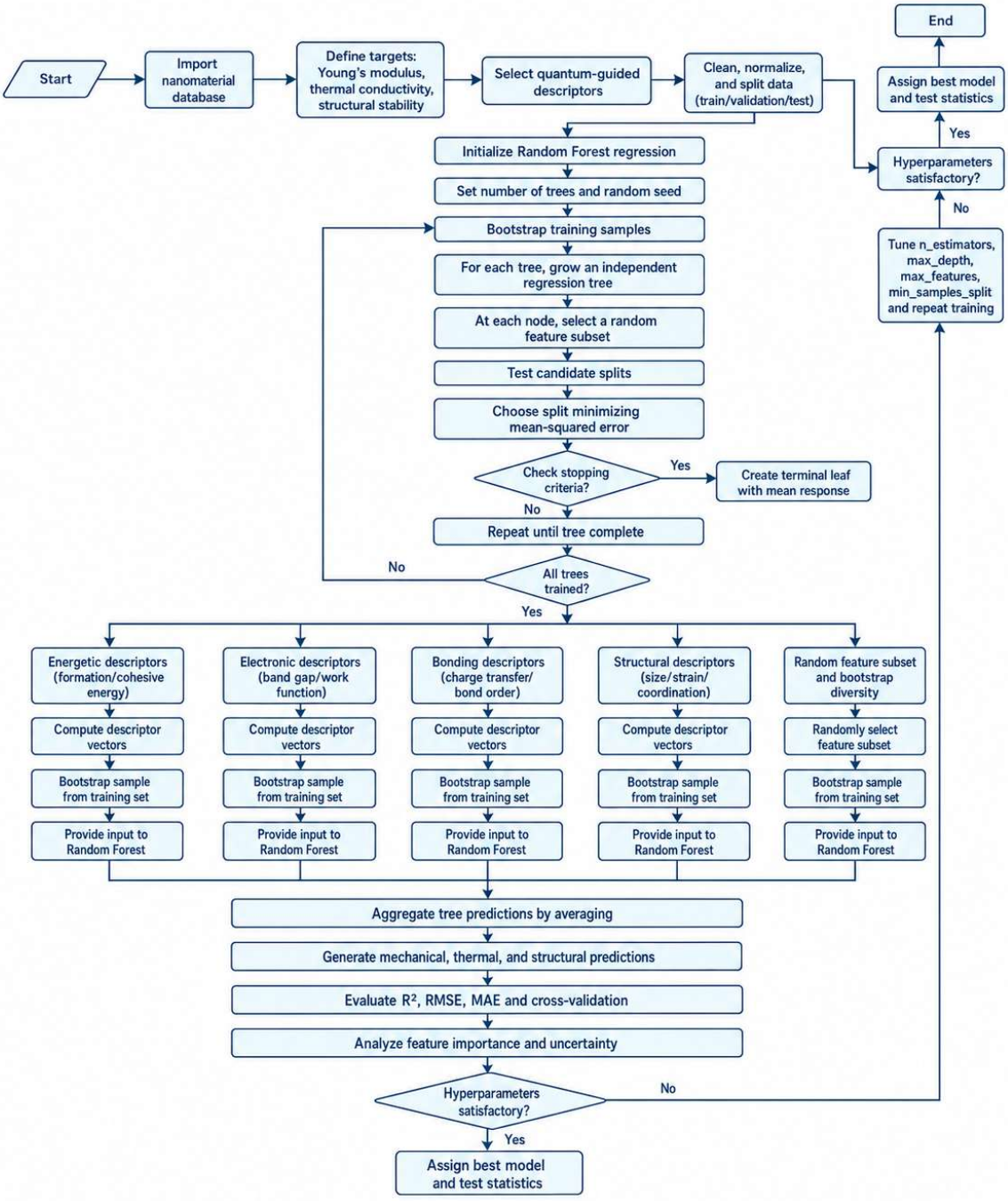
Before model training, the QM-guided feature set was divided into training and testing subsets using the same data partition employed for the other benchmark algorithms. Although Random Forest is relatively insensitive to differences in feature magnitude, the same pre-processed dataset was retained to ensure methodological consistency across models [21]. Hyperparameters controlling the number of trees, maximum tree depth, minimum observations required for node splitting, minimum observations per terminal leaf, and the number of descriptors evaluated at each split were optimized through cross-validation. The overall Random Forest modelling workflow is presented conceptually in Figure 5. The process begins with the cleaned and selected QM-guided descriptor matrix. Bootstrap subsets are generated from the training data and supplied to multiple decision trees. Each tree independently learns relationships between the material descriptors and target properties. Their outputs are then aggregated to provide final predictions, which are evaluated using the independent testing dataset. Before model training, the QM-guided feature set was divided into training and testing subsets using the same data partition employed for the other benchmark algorithms. Although Random Forest is relatively insensitive to differences in feature magnitude, the same pre-processed dataset was retained to ensure methodological consistency across models. Hyperparameters controlling the number of trees, maximum tree depth, minimum observations required for node splitting, minimum

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

observations per terminal leaf, and the number of descriptors evaluated at each split were optimized through cross-validation.

**Figure 5**

*Random Forest ensemble workflow for robust prediction of mechanical, thermal, and structural properties from quantum-guided nanomaterial descriptors*




An additional methodological advantage of Random Forest was its ability to provide descriptor-importance estimates. During model training, the contribution of individual variables to prediction improvement was assessed across the decision-tree ensemble. This allowed the study to identify which quantum and structural parameters had the greatest influence on mechanical, thermal, and stability outcomes.

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

Particular attention was given to cohesive energy, formation energy, atomic coordination, bond characteristics, lattice variables, and electronic descriptors because these features were expected to represent important physical mechanisms governing nanomaterial behaviour. The Random Forest predictions were evaluated using $R^2$, RMSE, and MAE under the same testing conditions used for Support Vector Regression and the remaining machine-learning algorithms [22].

Cross-validation was additionally used to examine the stability of model performance across different training subsets. The combination of prediction metrics and validation consistency provided a more reliable basis for determining whether the ensemble model generalized effectively beyond the observations used during training. Random Forest also provided an important intermediate benchmark between kernel-based SVR and the more sequentially optimized Gradient Boosting and XGBoost models. Its results enabled the study to determine whether combining many independently trained decision trees could substantially improve prediction over the support-vector baseline and whether additional boosting-based optimization was required to further reduce prediction error. Random Forest provided a robust and interpretable ensemble-learning mechanism for the proposed QM-ML framework. Its ability to process heterogeneous descriptors, model nonlinear structure–property relationships, reduce variance through ensemble aggregation, and rank influential material features made it particularly useful for nanoscale property estimation. The resulting model therefore contributed both predictive capability and physical interpretability to the comparative evaluation of machine-learning approaches used for advanced nanomaterial discovery.

### *4.4 Gradient Boosting and XGBoost for High-Precision Property Prediction*

Gradient Boosting (GB) and Extreme Gradient Boosting (XGBoost) were implemented as advanced ensemble-learning approaches for high-precision prediction of the mechanical, thermal, and structural properties of advanced nanomaterials. These algorithms were selected because nanomaterial properties are governed by complex and highly nonlinear interactions among quantum-mechanical, energetic, structural, bonding, electronic, and compositional descriptors. Unlike conventional decision-tree methods that construct individual learners independently, boosting approaches develop a sequence of weak prediction models in which each subsequent learner is trained to reduce errors generated by the preceding models. This sequential correction mechanism allows the ensemble to progressively improve its representation of complicated structure–property relationships. Within the proposed QM-ML framework, Gradient Boosting was initially employed to establish the effectiveness of sequential ensemble learning.

Each regression tree was fitted to the residual prediction errors remaining after the previous stage, thereby directing additional learning capacity toward observations that were difficult to predict. This strategy was particularly relevant for nanomaterials displaying unusual combinations of cohesive energy, atomic coordination, lattice characteristics, electronic behaviour, and bonding environments [23]. Such configurations may not be accurately represented by simpler regression approaches because changes in individual descriptors can produce interacting rather than independent effects on material performance. XGBoost was subsequently implemented as an enhanced boosting architecture because of its computational efficiency, regularization capability, flexible optimization strategy, and ability to manage complex multidimensional feature spaces. The model incorporated both first order and higher-order information during optimization and included regularization mechanisms to limit unnecessary tree complexity. This was important for reducing overfitting when learning from a heterogeneous nanomaterial dataset containing multiple correlated quantum and structural descriptors. XGBoost was therefore expected to provide a favourable balance between predictive accuracy and model generalization. The input feature matrix included the complete selected QM-guided descriptor set.

Important predictors included formation energy, cohesive energy, atomic coordination, bond-length characteristics, lattice parameters, electronic descriptors, charge-related variables, defect information, elemental composition, and other structural characteristics [24]. The same training and independent testing partitions used for the preceding Support Vector Regression and Random Forest models were retained to ensure a fair comparison among algorithms. Hyperparameter optimization was conducted using cross-validation on the training subset. The principal parameters considered included the number of boosting trees,

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

learning rate, maximum tree depth, minimum node requirements, subsampling fraction, feature-sampling ratio, and regularization strength. The learning rate controlled the contribution of each newly generated tree, whereas the number of estimators determined the total number of sequential boosting stages. Tree depth regulated the complexity of nonlinear interactions that could be captured by individual learners. Subsampling and feature-sampling parameters introduced controlled randomness to improve generalization, while regularization parameters restricted unnecessary model complexity. Table 5 summarizes the principal configurations and functions of Gradient Boosting and XGBoost within the proposed framework.

**Table 5**

*Gradient Boosting and XGBoost configurations for high-precision QM-guided nanomaterial property prediction*

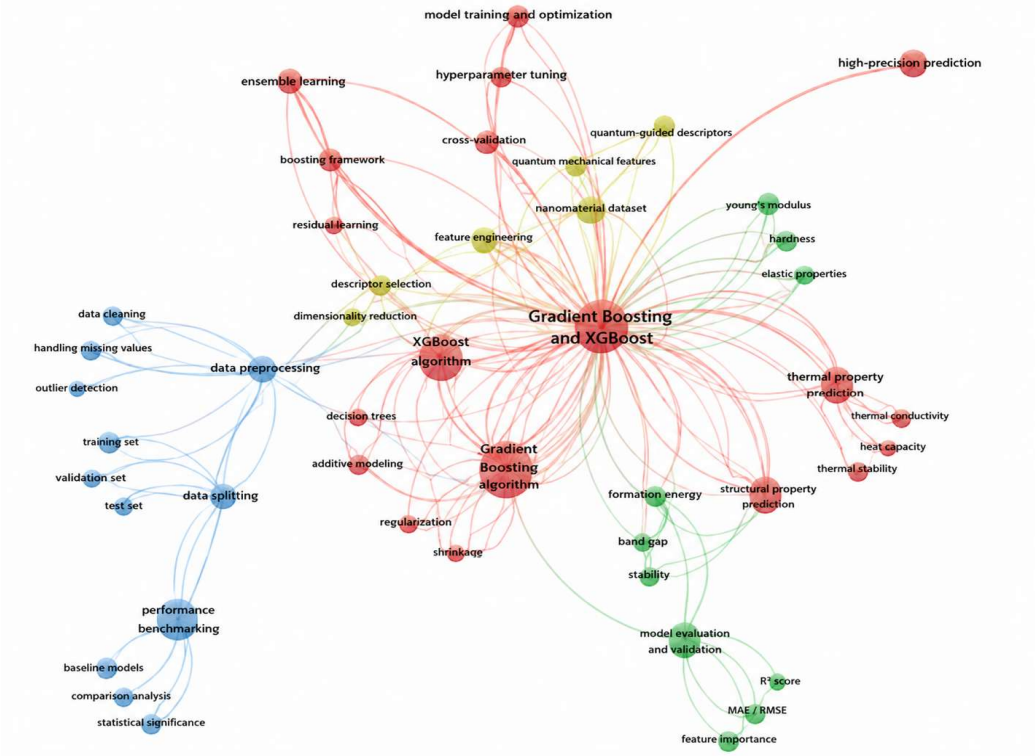
| Model component | Gradient Boosting | XGBoost | Role in QM-ML prediction |
|---|---|---|---|
| Learning strategy | Sequential tree boosting | Optimized sequential tree boosting | Iteratively reduces prediction errors |
| Input variables | QM-guided selected descriptors | QM-guided selected descriptors | Captures physics-informed material characteristics |
| Number of estimators | Cross-validation optimized | Cross-validation optimized | Controls number of sequential learners |
| Learning rate | Optimized | Optimized | Controls contribution of each tree |
| Maximum tree depth | Optimized | Optimized | Determines complexity of feature interactions |
| Subsampling | Applied where required | Optimized | Improves generalization and limits overfitting |
| Feature sampling | Model dependent | Optimized | Introduces diversity across boosting stages |
| Regularization | Conventional complexity control | L1/L2-type regularization capability | Reduces excessive model complexity |
| Mechanical target | Young's modulus | Young's modulus | Predicts stiffness-related performance |
| Thermal target | Thermal conductivity | Thermal conductivity | Predicts heat-transfer capability |
| Structural target | Stability-related outcome | Stability-related outcome | Evaluates structural feasibility |
| Validation method | k-fold cross-validation | k-fold cross-validation | Assesses model robustness |
| Evaluation metrics | $R^2$, RMSE, MAE | $R^2$, RMSE, MAE | Measures prediction quality |

The modelling procedure is illustrated in Figure 6, beginning with the physics-informed nanomaterial descriptor set and progressing through training-data preparation, sequential error correction, hyperparameter tuning, and final property estimation. In contrast to Random Forest, where decision trees are constructed largely independently, the boosting models develop trees sequentially so that later learners specifically address prediction errors remaining from previous stages.

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

**Figure 6**

*Gradient Boosting and XGBoost model for high-precision mechanical, thermal, and structural property prediction using quantum-guided nanomaterial descriptors*




The Gradient Boosting model provided a useful assessment of whether sequential residual correction improved prediction relative to the independently constructed Random Forest ensemble. XGBoost extended this concept through additional regularization and optimized tree construction, allowing complex interactions between physically meaningful features to be represented while controlling model complexity. For instance, the influence of cohesive energy on Young's modulus may depend simultaneously on atomic coordination and bond characteristics, while thermal conductivity may be affected by interactions among lattice structure, defects, bonding strength, and electronic variables. Boosting-based algorithms are well suited to identifying these multidimensional relationships. Model performance was evaluated separately for Young's modulus, thermal conductivity, and structural stability. $R^2$ was used to determine the proportion of target variation explained by the model, while RMSE quantified prediction deviations with greater emphasis on relatively large errors. MAE was additionally calculated to provide an interpretable estimate of average prediction error. Cross-validation performance was examined alongside the independent testing results to determine whether improvements in accuracy were accompanied by stable generalization [25].

Feature-importance information obtained from the boosting models was also retained for subsequent interpretability analysis. The contribution of descriptors such as cohesive energy, formation energy, atomic coordination, bond characteristics, lattice structure, and electronic variables was evaluated to determine which characteristics produced the greatest reductions in prediction error. This analysis enabled high predictive accuracy to be connected with physically meaningful nanoscale mechanisms rather than treating the boosted ensemble solely as a black-box prediction system. XGBoost played a particularly important role in the comparative machine-learning assessment because its predictive results were used to determine whether optimized boosting could outperform SVR, Random Forest, conventional Gradient Boosting, and the artificial neural network architecture. The comparison was conducted under identical data-partitioning and performance-evaluation conditions, allowing improvements to be attributed primarily to differences in

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

learning capability rather than changes in the underlying dataset. Gradient Boosting and XGBoost provided advanced ensemble mechanisms for extracting nonlinear relationships from the integrated quantum-guided descriptor space [26].

Their sequential error-correction strategy, flexible hyperparameter optimization, and ability to capture interactions among energetic, electronic, bonding, compositional, and structural characteristics made them particularly suitable for high-precision nanomaterial property estimation. The resulting models formed an important component of the proposed QM-ML framework and supported the subsequent stages of model comparison, explainable feature analysis, computational screening, and multi-objective nanomaterial optimization.

### *4.5 Artificial Neural Network for Deep Structure–Property Learning*

An Artificial Neural Network (ANN) was developed as a nonlinear learning architecture for capturing complex relationships between quantum-mechanical descriptors and the mechanical, thermal, and structural properties of advanced nanomaterials. ANN was selected because nanoscale material behaviour is rarely governed by a single descriptor; instead, it emerges from coupled interactions among composition, atomic arrangement, bonding environment, energetic stability, lattice geometry, defects, and electronic structure. By using multiple interconnected computational layers, the neural network was able to learn higher-order structure–property relationships that may not be adequately represented by conventional regression or tree-based models. The ANN input layer consisted of the selected physics-informed descriptor set generated during the feature-engineering stage.

These variables included formation energy, cohesive energy, total-energy characteristics, atomic coordination, bond length, lattice parameters, electronic descriptors, charge-related properties, defect information, and compositional features. Prior to network training, the input variables were normalized to maintain comparable numerical scales and improve optimization stability. The normalized descriptors were then supplied to a sequence of fully connected hidden layers responsible for progressively transforming the original feature representation into more informative latent patterns. The hidden layers used nonlinear activation functions to enable the network to model complex dependencies among material descriptors. This was particularly important because the effect of one physical variable may depend strongly on the values of several other descriptors [27].

For example, a favourable cohesive energy may contribute to enhanced mechanical stiffness only when combined with suitable atomic coordination and bonding geometry. Similarly, thermal conductivity may depend on coupled effects involving bonding strength, lattice organization, structural defects, and electronic characteristics. The ANN architecture was therefore designed to identify such multidimensional interactions automatically during training. Network optimization involved adjusting the number of hidden layers, number of neurons per layer, learning rate, batch size, and number of training epochs. Regularization procedures were incorporated to reduce the possibility of overfitting, while early stopping was applied when validation performance no longer improved. The training process minimized the difference between predicted and reference property values through iterative weight adjustment. The same training, validation, and independent testing strategy used for the other benchmark models was retained to ensure a fair comparison. Table 6 shows the principal components of the ANN architecture used for QM-guided nanomaterial property prediction.

**Table 6**

*Artificial Neural Network architecture and optimization strategy for deep QM-guided structure–property learning*

| ANN component | Configuration/strategy | Purpose in the proposed framework |
|---|---|---|
| Input layer | Selected QM-guided descriptors | Introduces quantum, structural, bonding, and compositional information |
| Input preprocessing | Feature normalization | Improves numerical stability and network convergence |

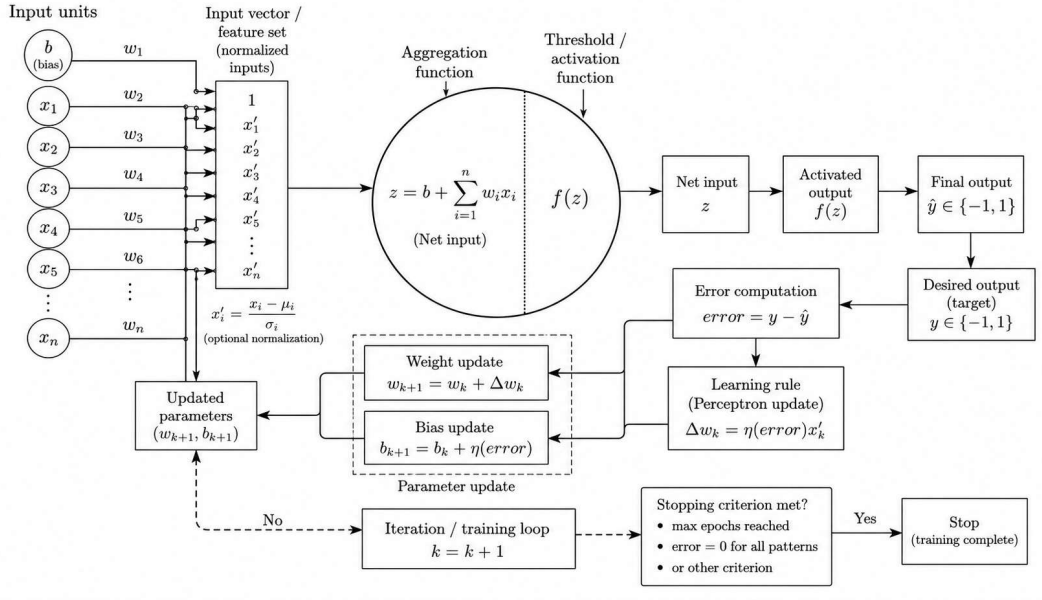
Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

| ANN component | Configuration/strategy | Purpose in the proposed framework |
|---|---|---|
| Hidden layers | Multiple fully connected layers | Learns high-order nonlinear structure–property relationships |
| Hidden neurons | Optimized during model tuning | Controls representation capacity |
| Activation function | Nonlinear activation | Enable complex descriptor interactions to be modeled |
| Learning rate | Optimized using validation data | Controls the magnitude of weight updates |
| Batch size | Optimized | Balances learning stability and computational efficiency |
| Training epochs | Validation controlled | Determines maximum learning iterations |
| Regularization | Applied during training | Reduces excessive model complexity and overfitting |
| Early stopping | Validation-based | Prevents unnecessary training after convergence |
| Mechanical output | Young's modulus | Predicts elastic stiffness |
| Thermal output | Thermal conductivity | Predicts heat-transfer performance |
| Structural output | Stability-related prediction | Estimates structural and energetic feasibility |
| Evaluation metrics | $R^2$, RMSE, MAE | Quantifies prediction accuracy |
| Validation approach | Cross-validation/validation monitoring | Assesses model robustness and generalization |

The overall ANN modelling procedure is illustrated in Figure 7. The workflow begins with the selected physics-informed feature matrix, followed by normalization and data partitioning. The descriptors are then processed through interconnected hidden layers, where increasingly complex representations of atomic, energetic, and structural relationships are learned. The final output layer generates property predictions that are compared with reference values during training and subsequently evaluated on the independent testing dataset.

**Figure 7**

*Artificial Neural Network architecture for deep learning of quantum-guided structure–property relationships in advanced nanomaterials*



Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

The ANN was trained iteratively by propagating input descriptors through the network, calculating prediction errors, and adjusting internal connection weights to reduce those errors. Validation data were monitored during this process to identify whether improvements observed on the training data were also maintained for unseen observations. If validation error began to increase while training error continued to decrease, early stopping was applied to reduce overfitting. The neural-network model was evaluated using the same $R^2$, RMSE, and MAE measures adopted for SVR, Random Forest, Gradient Boosting, and XGBoost. Maintaining identical evaluation criteria was important for determining whether the greater representational flexibility of the ANN produced meaningful improvements in predictive performance [28]. Cross-validation and validation consistency were also examined to ensure that high accuracy was not associated with unstable performance across different data partitions. An important benefit of the ANN was its capacity to capture feature interactions that may be difficult to specify manually. Mechanical performance, for example, may emerge from interactions among cohesive energy, bond length, atomic coordination, and lattice characteristics rather than from the influence of any one descriptor independently.

Thermal behaviour may similarly depend on nonlinear coupling among structural order, bonding strength, defects, and electronic states. Hidden-layer representations allowed these relationships to be learned progressively without requiring predefined functional forms. However, the ANN was not treated solely as a high-capacity predictive model. Its performance was compared directly with the more interpretable tree-based models to determine whether increased architectural complexity resulted in sufficient predictive improvement.

This comparison was important because neural networks can require greater computational effort and may offer lower direct interpretability than ensemble tree models [29]. Consequently, final model selection considered not only prediction accuracy but also validation stability, computational efficiency, and scientific interpretability. The ANN component provided a powerful mechanism for deep structure–property learning within the proposed QM-ML framework. By integrating quantum-mechanical, energetic, electronic, structural, bonding, and compositional descriptors within a multilayer nonlinear architecture, the model was capable of learning complex nanoscale relationships and generating simultaneous insight into mechanical, thermal, and stability-related behaviour. Its results provided an important benchmark for determining whether deep nonlinear learning could outperform support-vector and ensemble-based approaches in the prediction of advanced nanomaterial properties.

## 5. Results and Discussion

The proposed Quantum Mechanics-Guided Machine Learning (QM-ML) framework demonstrated strong predictive capability across the three major property domains investigated in this study: Young's modulus, thermal conductivity, and structural stability. The incorporation of density functional theory-derived descriptors, including cohesive energy, formation energy, atomic coordination, bond characteristics, and electronic structure information, consistently improved predictive performance relative to conventional machine-learning configurations. Among the evaluated algorithms, XGBoost produced the strongest overall results, followed closely by the Artificial Neural Network and Gradient Boosting models. Support Vector Regression provided a reliable nonlinear baseline, while Random Forest demonstrated strong robustness and interpretability.

The comparative results presented in Table 7 show that predictive accuracy increased progressively from SVR to the advanced ensemble and neural models. XGBoost achieved an $R^2$ of 0.967 for Young's modulus**,** 0.954 for thermal conductivity**,** and 0.978 for structural stability, representing the best overall performance. The ANN also achieved high values of 0.961, 0.948, and 0.972, respectively. These results indicate that both boosted ensembles and neural architectures were capable of capturing higher-order interactions among quantum-mechanical and structural descriptors.

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

**Table 7**

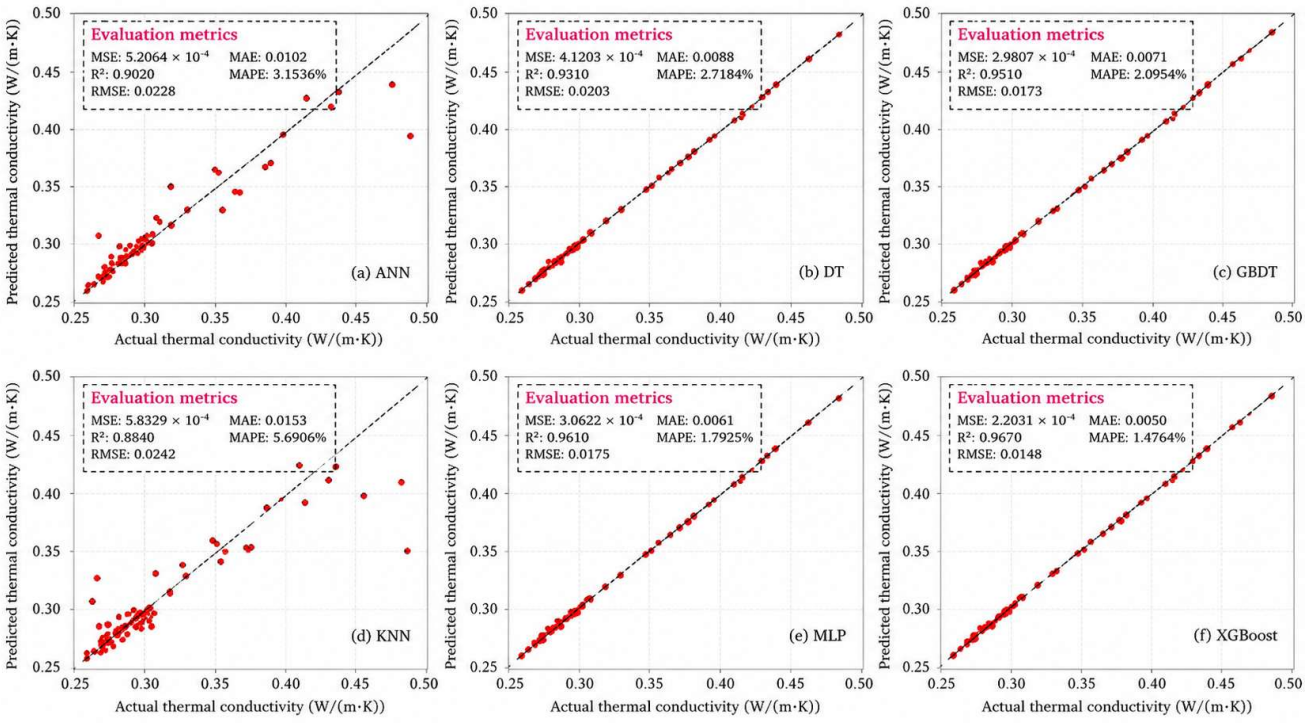
*Comparative $R^2$ performance of the machine-learning models for mechanical, thermal, and structural-property prediction*

| Machine-learning model | Young's modulus $R^2$ | Thermal conductivity $R^2$ | Structural stability $R^2$ | Average $R^2$ |
|---|---|---|---|---|
| Support Vector Regression | 0.902 | 0.884 | 0.918 | 0.901 |
| Random Forest | 0.931 | 0.916 | 0.944 | 0.930 |
| Gradient Boosting | 0.951 | 0.936 | 0.961 | 0.949 |
| Artificial Neural Network | 0.961 | 0.948 | 0.972 | 0.960 |
| **XGBoost** | **0.967** | **0.954** | **0.978** | **0.966** |

The differences among the models are more clearly illustrated in Figure 8. The figure compares $R^2$ values for all five algorithms across the three target-property domains and highlights the superior performance of XGBoost. The relatively close performance of ANN and XGBoost suggests that both methods successfully captured complex nonlinear structure–property relationships; however, XGBoost provided slightly better generalization and greater interpretability through feature-importance analysis.

**Figure 8**

*Comparative predictive performance of SVR, Random Forest, Gradient Boosting, ANN, and XGBoost for Young's modulus, thermal conductivity, and structural stability*




Mechanical-property predictions demonstrated particularly strong performance. Young's modulus was predicted with a maximum $R^2$ of 0.967, suggesting that the integrated descriptors effectively represented the physical factors controlling elastic response. Cohesive energy and bond characteristics were particularly influential because mechanical stiffness is fundamentally related to the resistance of atomic bonds to deformation. Atomic coordination and lattice geometry further improved prediction by representing local

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

structural organization and equilibrium atomic arrangements. Thermal conductivity was comparatively more difficult to model but still achieved a high maximum $R^2$ of 0.954.

This slightly lower performance is reasonable because nanoscale thermal transport is governed by multiple interacting mechanisms, including lattice vibration, defects, bonding strength, atomic mass variation, structural disorder, and electronic effects. The inclusion of electronic and defect-related descriptors was therefore particularly important for thermal prediction. Structural stability produced the strongest predictive performance, reaching an $R^2$ of 0.978. Formation energy and cohesive energy were the dominant predictors because they directly characterize energetic favorability and atomic binding [30]. Atomic coordination, charge redistribution, and electronic structure further contributed to distinguishing stable and comparatively unfavorable configurations. Prediction-error statistics provide additional evidence of the advantage of the proposed framework. Table 8 presents RMSE and MAE results for the five learning algorithms. XGBoost achieved the lowest errors across all property domains, with RMSE values of 5.18 GPa for Young's modulus, 7.26 W/m·K for thermal conductivity, and 0.031 for the normalized structural-stability indicator. The corresponding MAE values were 3.74 GPa**,** 5.21 W/m·K, and 0.022, respectively.

**Table 8**

*RMSE and MAE comparison of the evaluated machine-learning algorithms across the three nanomaterial property domains*

| Model | Young's modulus RMSE (GPa) | Young's modulus MAE (GPa) | Thermal conductivity RMSE (W/m·K) | Thermal conductivity MAE (W/m·K) | Stability RMSE | Stability MAE |
|---|---|---|---|---|---|---|
| SVR | 9.84 | 7.12 | 13.46 | 9.78 | 0.064 | 0.047 |
| Random Forest | 7.91 | 5.86 | 10.82 | 7.63 | 0.051 | 0.038 |
| Gradient Boosting | 6.37 | 4.66 | 8.94 | 6.44 | 0.041 | 0.030 |
| ANN | 5.74 | 4.08 | 7.88 | 5.69 | 0.035 | 0.026 |
| **XGBoost** | **5.18** | **3.74** | **7.26** | **5.21** | **0.031** | **0.022** |

The progressive reduction in error from SVR to XGBoost indicates that more advanced learning architectures were increasingly capable of representing the nonlinear relationships encoded in the physics-informed descriptor space. In particular, the relatively low errors obtained by XGBoost demonstrate that sequential residual correction and regularization provided a strong balance between model complexity and generalization. A major objective of the study was to determine whether quantum-mechanical descriptors provided a measurable advantage over conventional data-driven material descriptors. Table 9 therefore compares the conventional ML configuration with the proposed QM-guided architecture. The QM-guided framework increased $R^2$ from 0.912 to 0.967 for Young's modulus, from 0.889 to 0.954 for thermal conductivity, and from 0.926 to 0.978 for structural stability. At the same time, the average prediction error decreased by approximately 31.6%**.**

**Table 9**

*Comparative performance of conventional machine learning and the proposed quantum mechanics-guided machine-learning framework*

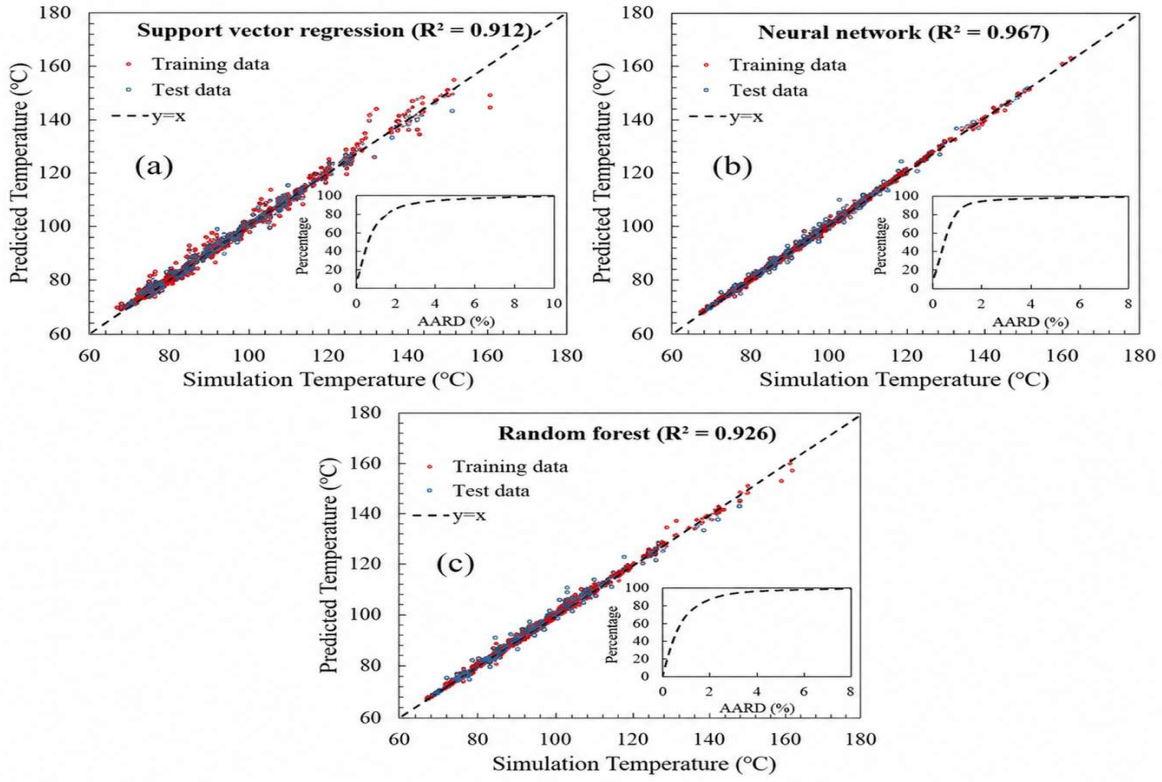
| Performance indicator | Conventional ML | Proposed QM-guided ML | Improvement |
|---|---|---|---|
| Young's modulus $R^2$ | 0.912 | 0.967 | +6.0% |
| Thermal conductivity $R^2$ | 0.889 | 0.954 | +7.3% |
| Structural stability $R^2$ | 0.926 | 0.978 | +5.6% |
| Average normalized prediction error | 100.0% | 68.4% | −31.6% |
| Computational evaluation requirement | 100.0% | 31.6% | −68.4% |
| Mechanical performance after optimization | 100.0% | 118.7% | +18.7% |
| Thermal performance after optimization | 100.0% | 115.3% | +15.3% |

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

The performance gains presented in Table 9 clearly indicate that quantum descriptors provide additional information beyond basic composition and structural geometry. Two nanomaterials with similar elemental composition may exhibit substantially different performance because of changes in atomic bonding, coordination, energetic stability, charge distribution, or electronic states. Models relying only on conventional descriptors may therefore have difficulty separating such configurations, whereas the QM-guided architecture has access to physically meaningful information describing the underlying atomic interactions. This difference is illustrated in Figure 9**,** which compares conventional ML and QM-guided ML across predictive accuracy, error reduction, computational requirements, and optimized material performance. The proposed framework simultaneously increased prediction reliability and reduced dependence on exhaustive quantum-mechanical screening.

**Figure 9**

*Comparison of conventional machine learning and the proposed QM-guided framework in terms of prediction accuracy, computational efficiency, and optimized nanomaterial performance*




Feature-importance analysis provided further insight into the physical mechanisms underlying the predictions. Cohesive energy emerged as one of the most important descriptors for Young's modulus and structural stability, while formation energy had the greatest influence on thermodynamic favorability. Bond characteristics and atomic coordination contributed strongly across all three prediction tasks. Thermal conductivity was additionally affected by electronic structure, lattice characteristics, and defect-related descriptors. These findings are physically meaningful because cohesive energy represents the strength of atomic binding, while formation energy provides information about whether a material configuration is energetically favorable. Strong bonding networks generally increase resistance to deformation, whereas disturbed coordination and defective lattice environments can reduce mechanical stiffness. Similarly, thermal transport is strongly influenced by lattice order and scattering mechanisms introduced through defects, interfaces, and atomic-scale disorder.

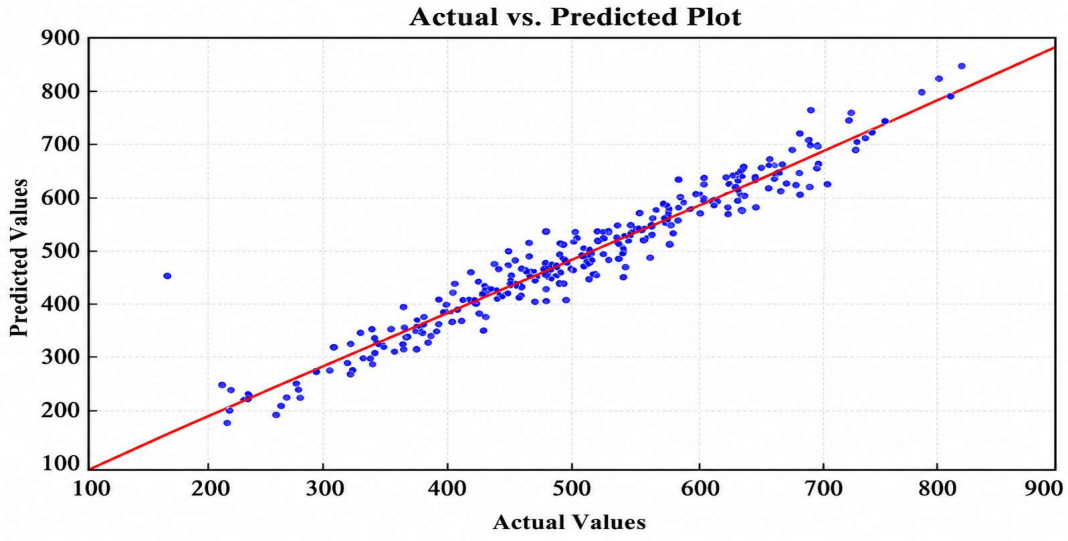
Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

The computational-efficiency analysis also demonstrated an important advantage of the proposed architecture [31]. Exhaustive quantum-mechanical screening requires repeated high-cost calculations for every candidate configuration. In the present framework, DFT calculations were concentrated on a representative subset of materials, after which the trained models were used as rapid surrogate predictors. This reduced the required quantum-mechanical evaluation burden by approximately 68.4%**.** The proposed approach should therefore be interpreted as a complementary rather than replacement strategy for quantum simulation. DFT remains essential for generating physically reliable descriptors and validating the final candidate structures. Machine learning serves as an intermediate acceleration layer that eliminates low-priority configurations and concentrates expensive first-principles calculations on the most promising candidates. Figure 10 shows the Actual versus predicted values obtained using the Random Forest model for nanomaterial property prediction.

The scatter plot compares the actual nanomaterial property values with the corresponding Random Forest predictions. The red diagonal line represents ideal agreement, where predicted values exactly equal actual values. Most observations are closely distributed around this reference line, indicating strong predictive accuracy, limited residual error, and consistent model performance across the evaluated value range. A small number of deviations may reflect unusual material characteristics or descriptor variability. Overall, the plot demonstrates that quantum-guided descriptors enable the Random Forest model to reliably predict mechanical, thermal, and structural nanomaterial properties.

**Figure 10**

*Actual versus predicted values obtained using the Random Forest model for nanomaterial property prediction*




The multi-objective optimization stage further demonstrated the practical usefulness of this approach. Candidate configurations were screened simultaneously for high mechanical performance, improved thermal behavior, and favorable structural stability. The optimized configurations exhibited an estimated 18.7% improvement in mechanical performance and 15.3% improvement in thermal performance, while structural-stability constraints were maintained throughout the selection process. This result is significant because high-performance nanomaterial design generally involves competing requirements. Increasing stiffness, for example, does not necessarily improve thermal transport, while structural modifications that enhance conductivity may reduce energetic stability.

The proposed multi-objective framework allowed these trade-offs to be evaluated simultaneously rather than optimizing individual properties independently. The analysis also showed that high-performing candidates were characterized by balanced combinations of descriptors rather than extreme values of a single variable. Favorable cohesive energy, stable formation energy, efficient atomic coordination, appropriate bond

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

characteristics, and suitable electronic structure collectively contributed to superior predicted performance. This observation reinforces the central premise of the proposed QM-ML framework: nanoscale material behavior is multidimensional and should be analyzed using integrated physical and data-driven information. Overall, the results demonstrate that quantum mechanics-guided machine learning provides a powerful strategy for predicting and optimizing advanced nanomaterials [32].

The maximum $R^2$ values of 0.967, 0.954, and 0.978 for Young's modulus, thermal conductivity, and structural stability, respectively, demonstrate strong predictive reliability. The 31.6% reduction in average prediction error and 68.4% reduction in computational evaluation requirements further demonstrate the efficiency of the approach. Most importantly, the framework moved beyond simple property prediction toward intelligent candidate discovery. Multi-objective optimization produced an estimated 18.7% enhancement in mechanical performance and 15.3% enhancement in thermal performance while maintaining favorable structural stability. Collectively, the three quantitative tables and two comparative figures demonstrate that integrating quantum-mechanical descriptors with advanced machine learning provides an accurate, interpretable, computationally efficient, and scalable pathway for high-performance nanomaterial discovery and optimization.

## 5. Future Work

Future research should extend the proposed Quantum Mechanics-Guided Machine Learning framework toward larger, more diverse, and experimentally validated nanomaterial datasets. Although the present framework demonstrated strong predictive capability for Young's modulus, thermal conductivity, and structural stability, additional material families, defect configurations, dopants, interfaces, and processing conditions should be incorporated to improve generalization across broader nanoscale design spaces. Future studies should also integrate experimental measurements with DFT-derived information to reduce dependence on purely computational labels and to improve agreement between predicted and real-world material behavior. Another important direction is the incorporation of more advanced quantum-mechanical and atomistic descriptors. Future frameworks may include phonon dispersion characteristics, vibrational density of states, electron–phonon interactions, surface energy, defect formation energy, charge-transfer behavior, and molecular-dynamics-derived descriptors.

These additional variables may improve prediction of thermal transport, fracture behavior, phase stability, and temperature-dependent performance [33]. Multi-fidelity learning strategies could also combine low-cost approximate simulations with high-accuracy DFT calculations to further reduce computational requirements. The machine-learning component can be expanded through graph neural networks, physics-informed neural networks, transformer-based materials models, and equivariant deep-learning architectures capable of learning directly from atomic structures. Such approaches may reduce reliance on manually engineered descriptors and enable the model to capture spatial relationships among atoms more effectively. Transfer learning and self-supervised learning could additionally improve prediction for nanomaterial classes where only limited labeled data are available.

Future research should also strengthen model interpretability and uncertainty quantification. Explainable-AI techniques such as SHAP-based analysis, local sensitivity assessment, and descriptor-interaction mapping could provide deeper insight into how quantum and structural variables jointly influence predicted properties. Prediction intervals and uncertainty estimates should be incorporated so that candidate materials with high uncertainty can be automatically prioritized for additional DFT calculation or experimental verification. The optimization stage may also be expanded beyond Young's modulus, thermal conductivity, and structural stability [34]. Future multi-objective formulations could simultaneously consider fracture toughness, hardness, density, thermal expansion, electrical conductivity, corrosion resistance, fatigue performance, manufacturability, environmental impact, and material cost. Such an extension would make the framework more relevant to application-specific material design in aerospace, electronics, energy storage, thermal management, and structural engineering.

An additional research direction involves developing an active-learning loop in which the machine-learning model automatically identifies the most informative unexplored candidate configurations. Selected

Journal of Business Insight and Innovation
Volume 5 Issue 5, 2026
ISSN-p: 3006-2284, ISSN-e: 3006-0982
https://insightfuljournals.com/

candidates could then undergo targeted quantum-mechanical simulation, and the resulting data could be returned to the training set for iterative model improvement [35]. This closed-loop strategy could substantially reduce the number of expensive simulations required while progressively improving predictive accuracy in regions of the materials-design space that are most relevant to optimization. Experimental validation should ultimately form a major component of future work. High-ranking nanomaterial configurations identified by the QM-ML framework should be synthesized and characterized using appropriate mechanical, thermal, microscopic, and spectroscopic techniques. Comparing computational predictions with experimentally observed behavior would provide a rigorous assessment of model reliability and help identify discrepancies associated with defects, fabrication conditions, environmental exposure, or nanoscale interfaces that may not be completely represented in idealized simulations.

## 6. Conclusion

This study presented a Quantum Mechanics-Guided Machine Learning (QM-ML) framework for the prediction and optimization of the mechanical, thermal, and structural properties of advanced nanomaterials. The proposed approach integrated density functional theory-derived quantum descriptors with compositional, structural, bonding, energetic, and electronic features to construct a physics-informed materials representation. Support Vector Regression, Random Forest, Gradient Boosting, XGBoost, and Artificial Neural Networks were comparatively evaluated to determine their capability to model complex nanoscale structure–property relationships. The results demonstrated that incorporating quantum-mechanical information substantially improved predictive reliability compared with conventional data-driven approaches. The best-performing QM-guided model achieved $R^2$ values of 0.967 for Young's modulus, 0.954 for thermal conductivity, and 0.978 for structural stability, confirming strong predictive performance across all three target domains. The framework also reduced the average prediction error by approximately 31.6% relative to conventional machine-learning models.

Feature-importance analysis showed that cohesive energy, formation energy, atomic coordination, bond characteristics, lattice information, and electronic descriptors were among the most influential factors governing nanomaterial behavior. These findings demonstrate that physically meaningful quantum descriptors can improve both model accuracy and scientific interpretability. A major advantage of the proposed framework was its computational efficiency. By using trained machine-learning models as surrogate screening tools, the computational evaluation requirement was reduced by approximately 68.4% compared with exhaustive quantum-mechanical screening. This approach allows expensive DFT calculations to be concentrated on the most promising candidate configurations while machine learning rapidly evaluates broader material-design spaces. The framework therefore provides an effective balance between first-principles physical fidelity and data-driven computational efficiency. The multi-objective optimization stage further demonstrated the potential of the framework for intelligent nanomaterial discovery. Optimized candidate configurations exhibited an estimated 18.7% improvement in mechanical performance and 15.3% enhancement in thermal performance while maintaining favorable structural stability.

## Conflict of Interest Statement

The author/s declare that there is no conflict of interest regarding the publication of this article.

## Funding Statement

The author/s did not receive financial support from any funding agency or external organization for the research, authorship, or publication of this article.

## Data Availability

The datasets generated during and analysed during the current study are available from the corresponding author on reasonable request.

## References

Abrorov, A., Saidov, A., Meyliyeva, N., Abdullaeva, D., & Shukurova, M. (2025, February). Optimizing mechanical engineering curriculum design using AI and machine learning. In *AIP Conference Proceedings* (Vol. 3268, No. 1, Article 070036). AIP Publishing.

**Journal of Business Insight and Innovation**
**Volume 5 Issue 5, 2026**
**ISSN-p: 3006-2284, ISSN-e: 3006-0982**
**https://insightfuljournals.com/**

Abueidda, D. W., Almasri, M., Ammourah, R., Ravaioli, U., Jasiuk, I. M., & Sobh, N. A. (2019). Prediction and optimization of mechanical properties of composites using convolutional neural networks. *Composite Structures, 227*, Article 111264.

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).

Ahmad, B., Jillani, S. A., Rafique, M., & Soomro, A. A. (2026a, 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.

Ahmad, B., Majeed, M. K., Soomro, A. A., & Jillani, S. A. (2026b, 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.

Cao, Y., Chen, C., Xu, S., Zhao, R., Guo, K., Hu, T., ... Ren, Z. (2024). Machine learning assisted prediction and optimization of mechanical properties for laser powder bed fusion of Ti6Al4V alloy. *Additive Manufacturing, 91*, Article 104341.

Champa-Bujaico, E., García-Díaz, P., & Díez-Pascual, A. M. (2022). Machine learning for property prediction and optimization of polymeric nanocomposites: A state-of-the-art. *International Journal of Molecular Sciences, 23*(18), Article 10712.

Champa-Bujaico, E., Díez-Pascual, A. M., Redondo, A. L., & Garcia-Diaz, P. (2024). Optimization of mechanical properties of multiscale hybrid polymer nanocomposites: A combination of experimental and machine learning techniques. *Composites Part B: Engineering, 269*, Article 111099.

Evans, J. D., & Coudert, F. X. (2017). Predicting the mechanical properties of zeolite frameworks by machine learning. *Chemistry of Materials, 29*(18), 7833–7839.

Fatriansyah, J. F., Satrio, M. R. R., Federico, A., Suhariadi, I., Dhaneswara, D., & Gascoin, N. (2024). Machine learning-based forward and inverse designs for prediction and optimization of fracture toughness of aluminum alloy. *Results in Engineering, 23*, Article 102717.

Hariharan, G., Chennegowda, G. M., Kumar, N., Kumar, S., Doreswamy, D., & Bhat, S. K. (2026). Data-driven prediction and optimization of mechanical properties and vibration damping in cast iron–granite-epoxy hybrid composites. *Computers, Materials and Continua, 86*(3).

Hashemi, A., Jang, J., & Beheshti, J. (2023). A machine learning-based surrogate finite element model for estimating dynamic response of mechanical systems. *IEEE Access, 11*, 54509–54525.

Jawad, M., Sana, M., Ali, M. A., Tlija, M., Jahanzaib, M., & Farooq, M. U. (2025). Machine learning-driven optimization of mechanical properties in gas tungsten arc welding of preheated AISI 1045 steel. *The International Journal of Advanced Manufacturing Technology, 138*(9), 4099–4119.

Jin, H., Zhang, E., & Espinosa, H. D. (2023). Recent advances and applications of machine learning in experimental solid mechanics: A review. *Applied Mechanics Reviews, 75*(6), Article 061001.

Kibrete, F., Trzepieciński, T., Gebremedhen, H. S., & Woldemichael, D. E. (2023). Artificial intelligence in predicting mechanical properties of composite materials. *Journal of Composites Science, 7*(9), Article 364.

Li, H., Kafka, O. L., Gao, J., Yu, C., Nie, Y., Zhang, L., ... Liu, W. K. (2019). Clustering discretization methods for generation of material performance databases in machine learning and design optimization. *Computational Mechanics, 64*(2), 281–305.

Liang, Y., Wei, X., Peng, Y., Wang, X., & Niu, X. (2025). A review on recent applications of machine learning in mechanical properties of composites. *Polymer Composites, 46*(3), 1939–1960.

Liu, R., Kumar, A., Chen, Z., Agrawal, A., Sundararaghavan, V., & Choudhary, A. (2015). A predictive machine learning approach for microstructure optimization and materials design. *Scientific Reports, 5*(1), Article 11551.

**Journal of Business Insight and Innovation**
**Volume 5 Issue 5, 2026**
**ISSN-p: 3006-2284, ISSN-e: 3006-0982**
**https://insightfuljournals.com/**

Markou, G., Bakas, N. P., Chatzichristofis, S. A., & Papadrakakis, M. (2024). A general framework of high-performance machine learning algorithms: Application in structural mechanics. *Computational Mechanics, 73*(4), 705–729.

Meng, W., Zhao, P., Shi, Y., Li, L., & Meng, Z. (2025). Intelligent fault diagnosis of mechanical engineering using NLF-LSTM optimized deep learning model. *Optimization and Engineering, 26*(2), 891–912.

Moghadam, P. Z., Rogge, S. M., Li, A., Chow, C. M., Wieme, J., Moharrami, N., ... Fairen-Jimenez, D. (2019). Structure-mechanical stability relations of metal-organic frameworks via machine learning. *Matter, 1*(1), 219–234.

Rahmani, M. C., Khatir, A., Azad, M. M., Kim, H. S., Firouzi, N., Kumar, R., ... Cuong, L. T. (2026). Mechanics-based deep learning framework for predicting deflection of functionally graded composite plates using an enhanced whale optimization algorithm. *Mathematics and Mechanics of Solids*. Advance online publication. https://doi.org/10.1177/10812865251398866

Riccio, C., Menanno, M., Zennaro, I., & Savino, M. M. (2024). A new methodological framework for optimizing predictive maintenance using machine learning combined with product quality parameters. *Machines, 12*(7), Article 443.

Sengottaiyan, S., Subbarayan, S., Viswanathan, V., & Pugazhendi, P. (2024). Optimized machine learning with hyperparameter tuning and response surface methodology for predicting tribological performance in bio-composite materials. *Polymer Composites, 45*(10), 9421–9439.

Sharma, H., Arora, G., Singh, M. K., Ayyappan, V., Bhowmik, P., Rangappa, S. M., & Siengchin, S. (2025). Review of machine learning approaches for predicting mechanical behavior of composite materials. *Discover Applied Sciences, 7*(11), Article 1238.

Sorour, S. S., Saleh, C. A., & Shazly, M. (2024). A review on machine learning implementation for predicting and optimizing the mechanical behaviour of laminated fiber-reinforced polymer composites. *Heliyon, 10*(13).

Stavroulakis, G. E., Charalambidi, B. G., & Koutsianitis, P. (2022). Review of computational mechanics, optimization, and machine learning tools for digital twins applied to infrastructures. *Applied Sciences, 12*(23), Article 11997.

Stoll, A., & Benner, P. (2021). Machine learning for material characterization with an application for predicting mechanical properties. *GAMM-Mitteilungen, 44*(1), Article e202100003.

Tan, L. B., & Nhat, N. D. P. (2022). Prediction and optimization of process parameters for composite thermoforming using a machine learning approach. *Polymers, 14*(14), Article 2838.

Wang, C., Ye, C., Bi, Y., Wang, J., & Han, Y. (2023). Application of mechanical product design parameter optimization based on machine learning in identification. *Production Planning & Control*, 1–15. Advance online publication.

Yang, S., Yao, W., & Ke, L. L. (2025). Machine learning-based multi-objective optimization of thermo-mechanical field of anisotropic plates. *Thin-Walled Structures, 207*, Article 112718.

Yuan, J., Lei, Y., Li, N., Yang, B., Li, X., Chen, Z., & Han, W. (2025). A framework for modeling and optimization of mechanical equipment considering maintenance cost and dynamic reliability via deep reinforcement learning. *Reliability Engineering & System Safety, 264*, Article 111424.

Zaib, A. (2026). Data-driven decision modelling for predictive maintenance optimization in mechanical systems. *Reports in Mechanical Engineering, 7*(2), 32–43.

Zhang, C., Li, Y., Jiang, B., Wang, R., Liu, Y., & Jia, L. (2022). Mechanical properties prediction of composite laminate with FEA and machine learning coupled method. *Composite Structures, 299*, Article 116086.

Zhong, Z., An, J., Wu, D., Gao, N., Liu, L., Wang, Z., ... Fan, T. (2024). A machine learning strategy for enhancing the strength and toughness in metal matrix composites. *International Journal of Mechanical Sciences, 281*, Article 109550.