Advance Sustainable Science, Engineering and Technology (ASSET)
Vol. 8, No.2, April 2026, pp. 02602028-01 ~ 02602028-09
ISSN: 2715-4211 DOI: https://doi.org/10.26877/asset.v8i2.2733

Comparative Study of Classical and Quantum Machine Learning Models: Insights into Quantum Advantage in Materials Informatics

Aris Tri Joko Harjanto*, Hindriyanto Dwi Purnomo, Hendry

Faculty of Information Technology, Universitas Kristen Satya Wacana, Jl. Dr. O. Notohamidjodjo, Salatiga 50715, Central Java, Indonesia

*982023024@student.uksw.edu

Abstract. Quantum Machine Learning (QML) has emerged as a promising paradigm for addressing increasing computational and representational demands in materials informatics. While classical models such as Support Vector Machines (SVM) achieve strong predictive performance, they often struggle to capture complex, highly correlated interactions in high-dimensional materials data. QML addresses this challenge by leveraging quantum-mechanical principles to construct expressive feature embeddings, where prospective quantum advantage lies in generating feature spaces that are difficult to approximate classically. In this study, 1,000 crystalline compounds from the Open Quantum Materials Database (OQMD) are evaluated in a binary classification task based on formation-energy stability. The dataset is normalized, reduced to four dimensions via Principal Component Analysis (PCA), and encoded into quantum circuits. Three QML models—QSVM, VQC, and QNN—are benchmarked against a classical SVM using repeated stratified evaluation. Results show that the classical SVM achieves the highest accuracy (91.8% ± 0.012), followed by QSVM (60.8% ± 0.035), while VQC and QNN perform significantly worse. This gap is driven by limited qubit capacity, encoding inefficiencies, restricted circuit expressivity, and optimization challenges. Nevertheless, QSVM demonstrates stable performance, suggesting that potential quantum advantage may emerge from improved feature encoding and kernel design rather than deeper variational circuits.

Keywords: Quantum machine learning, QSVM, VQC, QNN, materials informatics, quantum computing

(Received 2025-07-15, Revised 2025-09-05, Accepted 2026-03-02, Available Online by 2026-04-30)

1. Introduction

Materials informatics has become a central pillar in computational materials science, where data-driven approaches are increasingly used to predict properties, discover novel compounds, and accelerate the design of functional materials [1]. The rapid growth of publicly available databases—such as the Open Quantum Materials Database (OQMD), Materials Project, and NOMAD—combined with high-throughput Density Functional Theory (DFT) simulations, has led to large and complex datasets requiring scalable learning frameworks [2]. Classical machine learning methods, particularly kernel-based models such as Support Vector Machines (SVM), have demonstrated strong performance in this domain. However, as material descriptors become increasingly high-dimensional and correlated, classical models face challenges in efficiently capturing complex, non-linear interactions while maintaining scalability.

Quantum Machine Learning (QML) has been proposed as a promising alternative by leveraging quantum-mechanical principles—such as superposition, entanglement, and high-dimensional Hilbert space embeddings—to construct richer feature representations [3], [7]. In particular, quantum feature maps and variational circuits are hypothesized to enable enhanced representational capacity [4], [8], potentially allowing better separation of complex and highly correlated material features [4]. From this perspective, the prospective quantum advantage lies in the ability to generate expressive feature spaces and kernel representations that may be difficult to approximate classically.

Despite these theoretical motivations, empirical evidence for such advantages in practical, near-term settings remains limited. Current QML models often fail to consistently match classical baselines in realistic materials-informatics tasks. This gap can be attributed to several factors, including limited qubit availability [1], [4], information loss during classical-to-quantum encoding, restricted circuit expressivity, and optimization challenges such as barren plateaus and noisy gradient estimation [11], [12]. As a result, the conditions under which quantum models may offer practical advantages remain unclear [21], [22]. To address this gap, this work provides a systematic benchmark of three representative QML models: Quantum Support Vector Machine (QSVM), Variational Quantum Classifier (VQC), and Quantum Neural Network (QNN). These models are evaluated on a subset of 1,000 crystalline materials from OQMD, formulated as a binary classification task based on formation-energy stability. A classical SVM with optimized hyperparameters is used as the baseline.

Accordingly, this study investigates the following research questions:

- RQ1: Can near-term QML models achieve competitive predictive performance compared to a classical SVM under four-qubit encoding constraints?
- RQ2: How do QSVM, VQC, and QNN differ in terms of accuracy variability, convergence behavior, and computational cost?
- RQ3: Which methodological factors—such as encoding strategy, circuit design, and optimization protocol—most significantly influence QML performance?

By providing a transparent, statistically robust, and reproducible comparison, this work not only evaluates current QML capabilities but also identifies the practical limitations and potential pathways toward achieving quantum advantage in materials informatics.

2. Methods

2.1 Dataset Preparation and Feature Engineering

The dataset comprises 1,000 samples from the Open Quantum Materials Database (OQMD), with numerical descriptors of structural and thermodynamic properties. A binary label indicates material stability (1: stable, 0: unstable), with class balance preserved via stratified splitting.

Features were normalized using min–max scaling, and Principal Component Analysis (PCA) reduced the feature space to four dimensions to match qubit constraints. PCA was fitted on the training set and applied to the test set to avoid data leakage. The selected components retain most of the dataset variance while enabling efficient quantum encoding.

2.2 Quantum Computing Fundamentals

Quantum computing operates on qubits, which form the theoretical basis of QML [2], [3] and extend classical binary states into continuous complex superpositions. A qubit state can be written as $\alpha|0\rangle+\beta|1\rangle$, where the squared amplitudes correspond to measurement probabilities. This property enables efficient encoding of high-dimensional information within a single quantum system. The computational advantage of quantum systems further arises from entanglement, a non-classical correlation between qubits that allows joint representations beyond independent classical variables. Through unitary operations and multi-qubit gates (e.g., CNOT, CZ), quantum circuits can embed data into high-dimensional Hilbert spaces. These mechanisms—superposition, entanglement, and unitary transformations—form the foundation of Quantum Machine Learning (QML), enabling the construction of expressive feature maps and variational models.

2.3 Quantum Machine Learning Models

Quantum Machine Learning (QML) utilizes quantum circuits for high-dimensional feature embedding and variational learning. This study considers three approaches: QSVM, VQC, and QNN [17], [9], [14]. QSVM constructs a quantum kernel from feature-map circuits using state fidelities, offering stable performance without trainable parameters [8], [17], [10]. VQC employs parameterized circuits optimized via classical routines [9], [13], with performance dependent on circuit design and optimization dynamics. QNN extends this framework into deeper architectures with higher expressivity but increased computational cost and susceptibility to barren plateaus [14], [19], [23], [24]. These models represent kernel-based and variational paradigms for evaluating QML performance in materials informatics.

2.4 Challenges and Opportunities

Although QML provides theoretically powerful representational capabilities, its practical performance is constrained by NISQ-era limitations. Quantum circuits are susceptible to noise, and available qubit counts restrict the dimensionality of feasible encodings. Variational models often suffer from gradient vanishing effects [11], [12], hindering optimization. Nonetheless, QML remains promising for high-dimensional scientific domains—such as materials informatics—where quantum feature embeddings and scalable kernel constructions may offer long-term advantages as hardware fidelity and circuit design methodologies improve [25].

2.5 Quantum Feature Encoding

Quantum feature encoding is implemented using the ZZFeatureMap, which maps classical inputs into quantum states via single-qubit rotations and CZ-based entanglement to capture pairwise correlations. Two entangling repetitions are used to balance expressivity and NISQ constraints.

Let $x\in R^4$ be a sample after PCA. The feature map prepares a quantum state

$$|\phi(x)\rangle=U_{ZZFeatureMap}(x)|0\ldots0\rangle$$

where the circuit encodes both local features and cross-feature interactions. This embedding is shared across QSVM, VQC, and QNN, ensuring consistent comparison among models.

2.6 QSVM Configuration

The Quantum Support Vector Machine (QSVM) uses the Fidelity Quantum Kernel computed from the feature map described above [8], [17]. Kernel evaluation is performed by measuring the overlap between encoded states:

$$K(x_i,x_j)=|\langle\phi(x_i)|\phi(x_j)\rangle|^2.$$

The resulting kernel matrix is used as input to a classical SVC classifier with a precomputed kernel. No trainable parameters exist in the quantum circuit, making the QSVM relatively stable under noisy conditions. The kernel was evaluated using the Qiskit Estimator primitive with 1,024 shots and fixed random seeds for reproducibility. A classical SVC with regularization parameter C=1.0 was used for final classification.

2.7 VQC Architecture and Training

The Variational Quantum Classifier (VQC) consists of two components:

(1) the feature map for data encoding, and

(2) a trainable ansatz implemented using a hardware-efficient TwoLocal circuit [13].

The ansatz contains alternating layers of *RY* and *RZ* rotations, followed by CZ entanglement arranged in a fully connected pattern. Three repetitions were used, yielding a circuit with sufficient expressivity while maintaining tractable optimization. Trainable parameters were optimized using the SPSA optimizer [12] with a maximum of 50 iterations, learning rate scheduling, and gradient approximation based on stochastic perturbations.

The model was trained using cross-entropy loss, and convergence behaviour was monitored for signs of vanishing gradients, a known challenge in variational quantum learning. Classical–quantum hybrid training was executed using Qiskit EstimatorQNN with deterministic random seeds.

2.8 Quantum Neural Network (QNN) Setup

The Quantum Neural Network (QNN) used in this study extends the VQC structure by combining multiple feature-encoding and trainable layers into a unified network. Similar to VQC, the QNN leverages the ZZFeatureMap and a Three-Layer TwoLocal ansatz, but treats the entire circuit as a differentiable quantum network optimized via SPSA.

The QNN was implemented using the EstimatorQNN framework [14], where the output expectation values of the parameterized circuit form the logits for binary classification. Optimization followed the same training protocol as VQC but required more iterations per epoch due to the higher dimensionality of the parameter space. Although theoretically more expressive, the QNN is more sensitive to barren plateaus [11] and shot noise, making convergence slower compared to QSVM and VQC.

2.9 Classical Baseline

For baseline comparison, a classical Support Vector Machine (SVM) with an RBF kernel was trained on the same PCA-reduced data. Hyperparameters C=1.0 and γ=scale were selected based on standard grid-search ranges for materials classification tasks. The SVM was chosen as a strong benchmark due to its well-documented performance on moderate-dimensional structured datasets and its similarity to QSVM from a methodological perspective.

2.10 Computational Environment

All experiments were performed using Python 3.10, Qiskit 0.46, and scikit-learn 1.3. Quantum simulations used the Qiskit Aer Estimator backend running on a machine with an Intel Core i7 processor, 32 GB RAM, and no GPU acceleration. Random seeds were fixed across all runs to ensure reproducibility. Each QML model was executed five times using repeated stratified sampling to compute mean performance and standard deviations.

3. **Results and Discussion**

3.1 Overall Performance Comparison

The comparative evaluation of classical and quantum models is summarized in Table 1. Across five repeated runs using stratified sampling, the classical SVM with an RBF kernel achieved the highest performance with an average accuracy of 0.918 ± 0.012, confirming its strong capacity to model nonlinear relationships in the PCA-reduced feature space. Among the quantum models, QSVM obtained the highest average accuracy (0.608 ± 0.035), followed by VQC (0.363 ± 0.048) and QNN (0.298 ± 0.055). Table 1 highlights the performance gap between quantum and classical approaches under the current four-qubit encoding constraint.

Table 1. Classification accuracy of classical and quantum models

Model	Accuracy
Classical SVM (RBF)	0.918 ± 0.012
QSVM	0.608 ± 0.035
VQC	0.363 ± 0.048
QNN	0.298 ± 0.055

These results show a consistent trend across all runs: the classical SVM significantly outperforms the quantum models. However, the variance patterns indicate distinct computational behaviours. SVM displays high stability, while VQC and QNN show larger fluctuations, reflecting the susceptibility of parameterized quantum circuits to optimization noise and barren plateau effects.

3.2 Classical Baseline: SVM Performance

The classical SVM demonstrates strong discriminative capability on the four-dimensional PCA space. Its confusion matrix (Figure 3a) shows balanced true positive and true negative rates, with minimal misclassification near the decision boundary. The low variance across repeated runs further confirms its robustness for structured materials descriptors. From a computational perspective, SVM exhibits the lowest runtime and requires no quantum resources, making it an appropriate baseline for benchmarking emerging QML models.

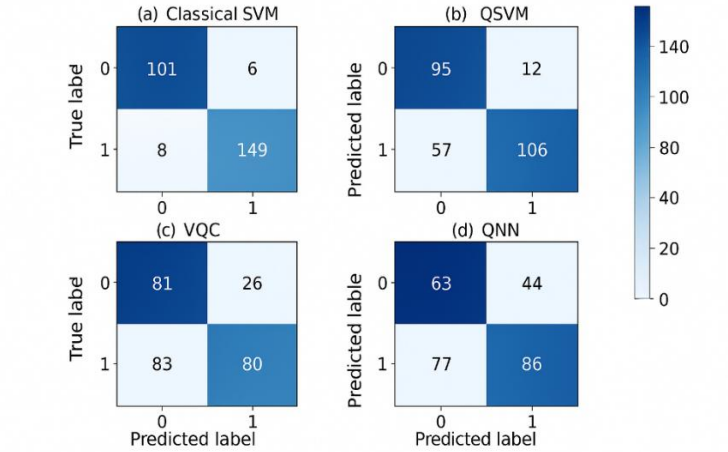

Figure 3. Confusion matrices of (a) Classical SVM, (b) QSVM, (c) VQC and (d) QNN

3.3 Quantum Support Vector Machine (QSVM)

QSVM achieves the highest quantum performance, with an accuracy of 0.608 ± 0.035 across repeated runs. The confusion matrix (Figure 3b) indicates moderate intra-class clustering but lower separability compared to the classical RBF kernel. This can be attributed to limited feature-map expressivity and dimensionality reduction: the ZZFeatureMap (two repetitions) captures pairwise interactions but may miss higher-order correlations, while projecting OQMD features into four PCA components introduces an information bottleneck.

Despite these limitations, QSVM demonstrates stable performance across runs, suggesting lower sensitivity to initialization and the absence of barren plateau issues common in variational models. This supports prior observations that quantum kernel methods can outperform variational approaches on small- to medium-scale datasets [8], [18].

3.4 Variational Quantum Classifier (VQC)

The Variational Quantum Classifier (VQC) achieved an average accuracy of 0.363 ± 0.048, with substantial variability across repeated runs. As shown in Figure 4(b), the distribution of accuracies indicates unstable convergence, a behavior commonly observed in variational circuits under NISQ constraints [12], particularly due to gradient vanishing effects [11]. The performance degradation is further influenced by parameter initialization sensitivity and stochastic optimization. These factors collectively hinder the VQC's ability to exploit the expressive capacity of the parameterized quantum circuit.

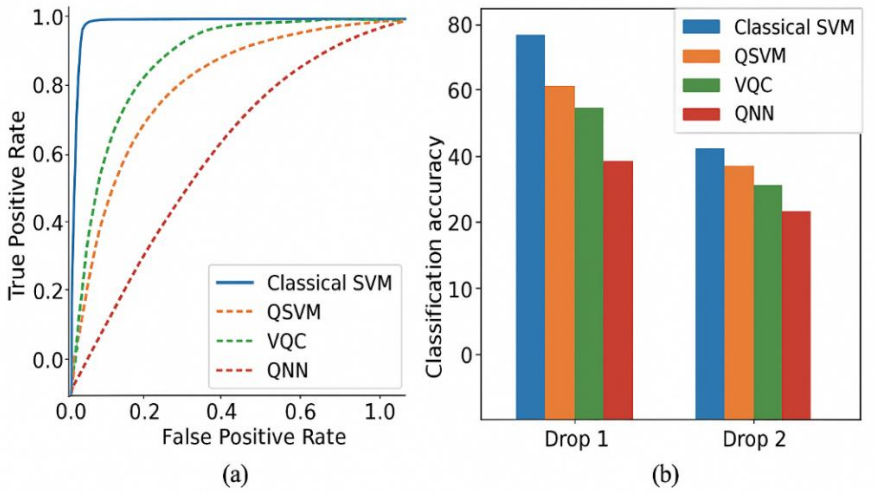

Figure 4. (a) ROC curves for Classical SVM, QSVM, VQC, and QNN.
(b) Classification accuracy distributions across five repeated runs for all models

Despite lower accuracy, VQC demonstrates partial learning signals in its early optimization steps, indicating that improved initializations, adaptive optimizers, or deeper but structured ansatz families may enhance performance.

3.5 Quantum Neural Network (QNN)

The QNN achieves the lowest accuracy (0.298 ± 0.055) and exhibits higher variance than VQC, indicating sensitivity to initialization and circuit depth. As shown in Figure 4(b), performance variability reflects unstable behavior across runs, consistent with barren plateau effects in deep variational circuits [12], [14]. These limitations arise from the large parameter space, noisy gradient estimates, and restricted expressivity under a four-qubit encoding. Such findings align with prior QML benchmarks, where QNNs often struggle to converge reliably under NISQ constraints without advanced training strategies.

3.6 Confusion Matrix and Class-wise Behavior

Figure 3 presents the confusion matrices for all four models. The classical SVM displays strong class balance, while QSVM exhibits mild bias toward the majority class. VQC and QNN show significantly weaker separation, with high false-negative rates, indicating underfitting.

This class-wise analysis highlights that although quantum models can capture partial structure within the PCA-reduced space, they are not yet competitive with classical methods for stability prediction tasks.

3.7 Resource Cost and Computational Efficiency

Table 2 summarizes the estimated computational resources for each model. SVM requires negligible overhead, while QSVM incurs cost during kernel construction. VQC and QNN have significantly higher runtime due to repeated circuit evaluations during parameterized optimization, highlighting known limitations of variational QML approaches [4], [12].

Table 2. Approximate computational cost per model

Model	Runtime	Circuit evaluations	Notes
SVM	Low	–	Fast classical optimization
QSVM	Medium	(O(n^2)) kernel evaluations	Depends on dataset size
VQC	High	~5,000–10,000	SPSA iterations dominate cost
QNN	Very High	>10,000	Deeper circuits + more parameters

This comparison reinforces that current NISQ hardware and simulators impose practical limits on variational QML approaches.

3.8 Summary of Findings

Overall, the classical SVM remains the superior model under the four-qubit encoding constraint, indicating that current QML approaches are still limited in practical materials-informatics settings. This performance gap can be attributed to restricted feature dimensionality, information loss during PCA-based encoding, and limited circuit expressivity, particularly for variational models. In addition, optimization challenges—such as barren plateaus and noisy gradient estimation—further degrade the performance of VQC and QNN.

Despite these limitations, QSVM demonstrates comparatively stable performance, suggesting that quantum kernel methods retain a key advantage: the ability to construct non-classical feature spaces without relying on complex parameter optimization. This indicates that potential quantum advantage may emerge from improved feature encoding and kernel design rather than increased circuit depth alone. These findings are consistent with recent QML benchmarking studies and highlight several directions for improvement, including increasing qubit capacity, refining encoding strategies, and developing structured or data-adaptive ansatz architectures [21], [22], [26]

4. Conclusion

This study benchmarked QSVM, VQC, and QNN against a classical SVM on a 1,000-sample OQMD subset with four-dimensional PCA encoding. The classical SVM achieved the highest accuracy (0.918 ± 0.012), followed by QSVM (0.608 ± 0.035), while VQC and QNN showed significantly lower performance (0.363 ± 0.048 and 0.298 ± 0.055). These results confirm that classical kernel methods remain superior under current NISQ and low-qubit constraints.

Among quantum models, QSVM demonstrated greater stability and consistently outperformed variational approaches, indicating that quantum kernel methods are more robust in limited-resource settings. In contrast, VQC and QNN exhibited high variance and unstable convergence, reflecting fundamental challenges such as barren plateaus, sensitivity to initialization, and noisy hybrid optimization. The observed performance gap is primarily driven by restricted qubit capacity, information loss during encoding, and limited circuit expressivity. Nevertheless, the results suggest that potential quantum advantage may arise not from increased variational depth, but from improved feature encoding and kernel design, particularly in low-parameter regimes [8], [18].

Future work will explore alternative feature maps, higher-dimensional encodings, and structured variational circuits, as well as larger datasets and real quantum hardware implementations. These directions are essential for clarifying the conditions under which QML can provide practical advantages in materials informatics.

Acknowledgements

The authors thank the open-source developers of Qiskit and scikit-learn, and the maintainers of OQMD for enabling reproducible research.

References

[1] Preskill J. Quantum computing in the NISQ era and beyond. Quantum. 2018;2:79. https://doi.org/10.22331/q-2018-08-06-79

[2] Schuld M, Petruccione F. Supervised Learning with Quantum Computers. Cham: Springer; 2018. https://doi.org/10.1007/978-3-319-96424-9

[3] Dunjko V, Briegel HJ. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. Rep Prog Phys. 2018;81(7):074001. https://doi.org/10.1088/1361-6633/aab406

[4] Bharti K, Cervera-Lierta A, Kyaw TH, et al. Noisy intermediate-scale quantum algorithms. Rev Mod Phys. 2022;94:015004. https://doi.org/10.1103/RevModPhys.94.015004

[5] Kirklin S, Saal JE, Meredig B, et al. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. npj Comput Mater. 2015;1:15010. https://doi.org/10.1038/npjcompumats.2015.10

[6] Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20:273–297. https://doi.org/10.1007/BF00994018

[7] Schuld M, Killoran N. Quantum machine learning in feature Hilbert spaces. Phys Rev Lett. 2019;122:040504. https://doi.org/10.1103/PhysRevLett.122.040504

[8] Havlíček V, Córcoles AD, Temme K, et al. Supervised learning with quantum-enhanced feature spaces. Nature. 2019;567:209–212. https://doi.org/10.1038/s41586-019-0980-2

[9] Peruzzo A, McClean J, Shadbolt P, et al. A variational eigenvalue solver on a photonic quantum processor. Nat Commun. 2014;5:4213. https://doi.org/10.1038/ncomms5213

[10] Fujii H, Yamamoto K. Quantum kernel method for machine learning. Phys Rev A. 2020;101(3):032302. https://doi.org/10.1103/PhysRevA.101.032302

[11] Grant E, Benedetti M, Cao S, et al. An initialization strategy for addressing barren plateaus in parametrized quantum circuits. Quantum. 2019;3:214. https://doi.org/10.22331/q-2019-12-09-214

[12] Cerezo M, Arrasmith A, Babbush R, et al. Variational quantum algorithms. Nat Rev Phys. 2021;3:625–644. https://doi.org/10.1038/s42254-021-00348-9

[13] Benedetti M, Lloyd E, Sack S, Fiorentini M. Parameterized quantum circuits as machine learning models. Quantum Sci Technol. 2019;4(4):043001. https://doi.org/10.1088/2058-9565/ab4eb5

[14] Abbas A, Sutter D, Zoufal C, et al. The power of quantum neural networks. Nat Comput Sci. 2021;1:403–409. https://doi.org/10.1038/s43588-021-00084-1

[15] Bausch T. Recurrent quantum neural networks. Quantum Inf Process. 2022;21:1. https://doi.org/10.1007/s11128-021-03344-3

[16] IBM Quantum. IBM Quantum services and systems. 2025. Available from: https://quantum-computing.ibm.com

[17] Rebentrost P, Mohseni M, Lloyd S. Quantum support vector machine for big data classification. Phys Rev Lett. 2014;113:130503. https://doi.org/10.1103/PhysRevLett.113.130503

[18] Schuld M. Quantum machine learning models are kernel methods. arXiv. 2021;arXiv:2101.11020. https://doi.org/10.48550/arXiv.2101.11020

[19] Tacchino F, Macchiavello C, Gerace D, Bajoni D. An artificial neuron implemented on an actual quantum processor. npj Quantum Inf. 2019;5:26. https://doi.org/10.1038/s41534-019-0140-4

[20] Jiang Z, Wang H, Wu K, Zhang S. Quantum algorithms for materials simulation. npj Comput Mater. 2023;9:100. https://doi.org/10.1038/s41524-023-01000-9

[21] Liu Y, et al. Benchmarking quantum machine learning algorithms across datasets. npj Quantum Inf. 2024.

[22] Jerbi S, et al. Challenges and opportunities in scaling QML. Nat Rev Phys. 2023.

[23] Lokeshwaran K, Komal Kumar N, Senthil Murugan J, Elanangai V, Sathya S. Benchmarking

Transformer Models Against Classical Approaches for Fake Review Detection on the Deceptive Opinion Spam Corpus. Int. J. Environ. Eng. Educ. [Internet]. 2025 Dec. 6 [cited 2026 May 2];7(3):182-95. Available from: https://ijeedu.com/index.php/ijeedu/article/view/334

[24] Raj J JS, Gopalan A. Compact Bi-slot Patch Antenna with Tapered Edges for Ka-Band Applications Featuring Machine Learning-Assisted Performance Prediction. Int. J. Environ. Eng. Educ. [Internet]. 2025 Dec. 6 [cited 2026 May 2];7(3):196-207. Available from: https://ijeedu.com/index.php/ijeedu/article/view/326

[25] Primawati P, Qalbina F, Mulyanti M, Yanuar F, Devianto D, Lapisa R, Rozi F. Predictive Maintenance of Old Grinding Machines Using Machine Learning Techniques . JAETS [Internet]. 2025Jun.8. 6(2):874-88. Available from: https://journal.yrpipku.com/index.php/jaets/article/view/6417

[26] Ilori AO, Ajanaku O. Characteristics of Synthesized Copper Oxide (CuO) Nanoparticles using Maize Husks: A Green Chemistry Approach. Int. J. Environ. Eng. Educ. [Internet]. 2022 Apr. 10 [cited 2026 May 2];4(1):1-4. Available from: https://ijeedu.com/index.php/ijeedu/article/view/68