JOIN (Jurnal Online Informatika)
p-ISSN: 2528-1682, e-ISSN: 2527-9165
Volume 10 Number 1 | June 2025: 1-11
DOI: 10.15575/join.v10i1.1333

Variational Quantum Circuit-Based Quantum Machine Learning Approach for Predicting Corrosion Inhibition Efficiency of Expired Pharmaceuticals

Muhamad Akrom[1], Muhammad Reesa Rosyid[2], Lubna Mawaddah[2], Akbar Priyo Santosa[2]
[1]Research Center for Quantum Computing and Materials Informatics, Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia
[2]Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia

Article Info

Article history:

Received April 23, 2024
Revised November 15, 2024
Accepted November 15, 2024
Published April 01, 2025

Keywords:

Ansatz Designs
Corrosion Inhibitors
Drug Compounds
Quantum Machine Learning
Variational Quantum Circuit

ABSTRACT

This study examines the potential of quantum machine learning (QML) to predict the corrosion inhibition capacity of expired pharmaceutical compounds. The investigation employs a QSPR model, using features generated from density functional theory (DFT) calculations as input. At the same time, corrosion inhibition efficiency (CIE) values obtained from experimental data serve as the target output. The VQC model demonstrates varied performance across evaluation metrics, especially with encoding and ansatz design. The model achieves fine scores in evaluation metrics, with root mean square error (RMSE) of 6.15, mean absolute error (MAE) of 5.63, and mean absolute deviation (MAD) of 5.50. The research underscores the significance of larger datasets for enhancing predictive accuracy and points to QML's potential in exploring anti-corrosion materials. Although there are some limitations, this study provides a foundational framework for using QML to predict anti-corrosive properties.

Corresponding Author:

Muhamad Akrom,
Research Center for Quantum Computing and Materials Informatics, Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia
Email: m.akrom@dsn.dinus.ac.id

1. INTRODUCTION

Corrosion is a process that causes damage to metals due to the electrochemical interaction between the metal surface and its corrosive environment [1], [2]. Gradually, corrosion reduces the lifespan of materials, potentially shortening their expected service life. The global damage caused by corrosion reached at least $2.5 trillion, equivalent to 3.4% of the worldwide GDP in 2013, making mitigating corrosion effects a critical priority. It is estimated that 15-35% of these losses could be prevented by applying corrosion inhibitors [3]. Utilizing corrosion inhibition is highly effective in preventing corrosion. Corrosion inhibitors typically consist of organic compounds containing heteroatoms such as nitrogen (N), phosphorus (P), sulfur (S), arsenic (As), or oxygen (O) within their molecular structure because these compounds possess free electrons or π electrons on aromatic rings or double bonds, allowing for strong interactions between metal atoms and organic molecules to form a protective layer on the metal surface. This layer is absorbed at the interface of the corrosive solution due to structural similarities, leading to numerous compounds being tested as corrosion inhibitors [4], [5], [6].

Currently, materials informatics is gaining popularity due to the widespread development of technologies such as Machine Learning (ML), particularly in the design and development of new

p-ISSN: 2528-1682
e-ISSN: 2527-9165

materials [7], [8], [9]. The ML approach based on the Quantitative Structure-Property Relationship (QSPR) model is frequently employed to evaluate the performance of compounds because molecular attributes can be measured and related to the chemical structure of these compounds [3], [10], [11]. This approach replaces costly, time-consuming experimental procedures that require significant resources [12], [13], [14]. The ML approach to QSPR targets the Corrosion Inhibition Efficiency (CIE) derived from experimental results and utilizes Quantum Chemical Properties (QCP) obtained from Density Functional Theory (DFT) calculations as input [15].

Research such as that by Quadri et al. [16] utilized Artificial Neural Networks (ANN) and Multi Linear Regression (MLR) to design corrosion inhibitors based on pyridazine compounds, revealing that ANN produced optimal results for predicting Corrosion Inhibition Efficiency (CIE). Similarly, Akrom et al. [17] achieved accurate predictions for the CIE of diazine compounds as corrosion inhibitors, with predictions ranging from 85.2% to 94.99%. Almari et al. [18] also reported promising results in predicting corrosion inhibition for pyrimidine compounds using Partial Least Squares (PLS) and Random Forest (RF), with RF yielding more accurate predictions compared to PLS. Pham et al. [19] concluded a parallel investigation employing Gradient Boosting (GB) focused on expired drugs, where the model generated a Root Mean Square Error (RMSE) of 6.40. A similar focus was carried out by Perez et al. [3], who proposed a model with Autoregressive Exogenous Input (ARX), utilizing CML methods to identify corrosion inhibition compounds in expired drugs.

Although CML has provided significant insights into structure-property correlations, there is a strong case for exploring Quantum Machine Learning (QML) approaches as a complementary pathway. Features extracted from DFT, which describe the quantum properties of molecules in Hilbert space, can enhance our understanding of molecular structures, quantum characteristics, and atomic interactions, thereby strengthening this initiative [20], [21]. The application of QML in generative and discriminative models has shown promising potential. Examples include unsupervised learning in finance, kernel techniques for support vector machines, and binary classification problems in image recognition [22], [23]. Akrom et al. [24] developed a Quantum Neural Network (QNN) model to investigate organic pyrimidine compounds as corrosion inhibitors in the context of corrosion inhibitor design. The results indicated that the proposed model delivered good predictive performance. In another study, Akrom et al. [25] also proposed a Variational Quantum Circuit (VQC) model to predict the corrosion inhibition efficiency of hybrid pyridine-quinoline compounds. Both models leveraged features derived from quantum phenomena obtained through DFT calculations in these modeling efforts.

Therefore, this article's main objective is to introduce the QML paradigm for estimating the Corrosion Inhibition Efficiency (CIE) of expired drugs. It also aims to highlight the novel applications of the VQC method in this field and explain how this method can be utilized to design corrosion inhibition compounds.

2. METHODS

In this study, we developed and assessed a comparison of QML models utilizing Variational Quantum Circuits (VQC) to predict the corrosion inhibition capabilities of expired pharmaceutical compounds. Figure 1 depicts the framework of the constructed QML model, with a comprehensive explanation provided in Sections 2.1 to 2.8.

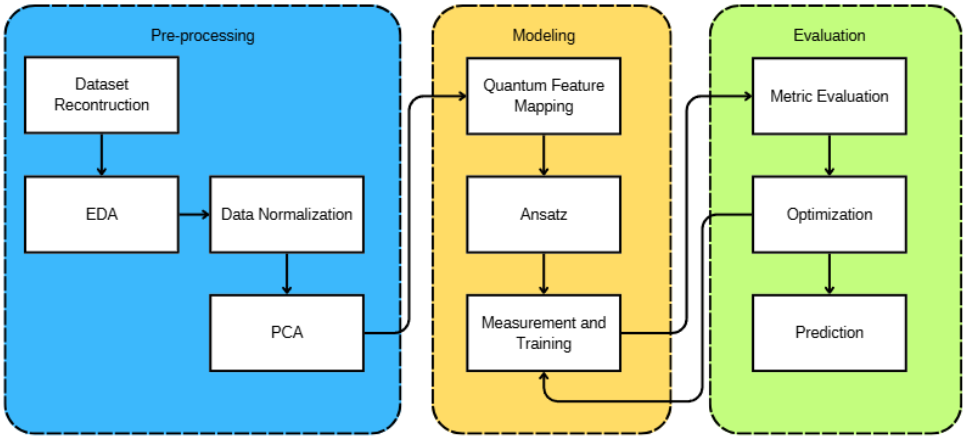

Figure 1. Research Flow Frameworks

2.1. Dataset Reconstruction

In this experiment, we utilized and restructured the dataset from the published study [3], [19]. The dataset includes 260 expired drugs identified as corrosion inhibitors. It contains two attributes: chemical property features and the Corrosion Inhibition Efficiency (CIE) value. The chemical property features encompass HOMO energy (eV), LUMO energy (eV), ionization potential (I) (eV), electron affinity (A) (eV), electronegativity (χ) (eV), electrophilicity (ω), global hardness (η) (eV), global softness (s) (eV), molecular weight (MW) (g/mol), polar surface area (PSA) (Å^2), polarizability (Å^3), acid dissociation constant (pKa), octanol-water partition coefficient (log P), water solubility (log S), the fraction of shared electrons (ΔN) and the CIE value is provided as the output [26], [27], [28] dataset depicted on Figure 2.

	HOMO (eV)	LUMO (eV)	I (eV)	A (eV)	χ (eV)	η (eV)	σ (eV)	ω (eV)	MW (g/mol)	PSA (Å2)	Polarizability (Å3)	pKa	Log P	Log S	ΔN	CIE (%)
0	-5.926	-2.103	5.926	2.103	4.0145	1.9115	0.523149	1.003625	214.25	122.06	20.38	2.80	-0.52	-2.40	0.780931	67.70
1	-6.088	-2.670	6.088	2.670	4.3790	1.7090	0.585138	1.094750	137.14	68.01	13.21	1.82	-0.70	0.01	0.766823	75.40
2	-5.914	-2.024	5.914	2.024	3.9690	1.9450	0.514139	0.992250	347.40	112.73	32.52	3.60	0.60	-3.10	0.779177	76.90
3	-6.419	-2.478	6.419	2.478	4.4485	1.9705	0.507485	1.112125	249.29	85.08	24.97	8.40	0.35	-2.70	0.647425	76.90
4	-6.587	-2.796	6.587	2.796	4.6915	1.8955	0.527565	1.172875	180.16	63.60	17.1	3.50	1.80	-2.10	0.608942	77.91

Figure 2. The model uses fifteen chemical property features as inputs and one target output, the CIE (%) value.

2.2. Exploratory Data Analysis (EDA)

The pre-processing phase, which includes Exploratory Data Analysis (EDA), is the initial and essential step in data-driven analysis methodologies. This phase addresses various challenges within the dataset, such as managing missing values, detecting anomalies, evaluating features with high correlation or low variance, identifying duplicate entries, recognizing skewed data distributions, establishing correlations between features and targets, and performing descriptive statistical analyses. These actions enhance understanding of data distribution, variability, and central tendencies. Insights gained from EDA clarify the key assumptions that inform model development and offer valuable guidance for improving subsequent data processing stages [29], [30]. Ultimately, EDA is a vital tool for navigating the complexities of datasets and supporting informed decision-making throughout the analytical process. The results of the EDA indicate a significant presence of missing values in the dataset, especially concerning the target variable. Only 80 clean data points were retained from the original 260 to address this issue. The remaining data points will be used in the modeling phase.

2.3. Normalization

Feature scaling methods are crucial for normalizing data, especially in large or heterogeneously scaled datasets, as they can significantly enhance model accuracy by reducing prediction errors [31], [32], [33]. The application of Robust Scaler in data preparation is significant because it can effectively

manage outliers in the dataset, which are data points that markedly differ from the majority. This scaling technique scales the data using the interquartile range (Q3 - Q1). Its robustness against outliers stems from normalizing the data using the middle 50% of the distribution before scaling. This approach minimizes the sensitivity of the scaling process by emphasizing the more representative central data distribution and diminishing the influence of extreme values.

$$X_{scaled} = \frac{X - Q_1}{Q_3 - Q_1} \quad (1)$$

In Equation 1, the initial value of a ata point is denote by the leter X.oneThe fitheisubscriptQbase1) of the fethreeture data coresponds to the 25th percentile, while the third quartile (Q3) represents the 75th percentile.

2.4. PCA

In this research, PCA is utilized to simplify the dimensions of both features and quantum circuits, enabling the characterization of datasets in a more concise form while retaining a substantial amount of the critical information from the original features [24], [34]. This technique enables the representation of datasets in a simplified form with reduced dimensions while retaining much of the key information from the original features. After applying PCA, the dataset was reduced to just 6 features from the original 15.

2.5. Basic Quantum Computing

Quantum computing is an innovative computing paradigm rooted in the principles of quantum physics, primarily utilizing entanglement and superposition. In quantum computation, information is represented similarly to how elementary particles and photons are modeled in traditional quantum mechanics, displaying both wave and particle properties. The fundamental unit of quantum information, the quantum bit (qubit), enables computations that are exceedingly challenging, if not nearly impossible, to achieve on traditional computers [35], [36], [37]. Quantum computers differ from classical computers in using qubits instead of bits. Unlike classical bits, which encode information strictly as 0 or 1 at any given time, qubits can represent these classical states (0 and 1) or exist in a superposition, simultaneously holding a combination of both states. This unique characteristic allows quantum computers to perform complex computations and explore multiple possibilities at once. The use of qubits necessitates a shift in data encoding and storage methods due to this superposition property, is defined neatly in Equation 2.

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (2)$$

In a two-dimensional complex Hilbert space with basis vectors |0⟩and |1⟩, the state of a qubit is represented as |ψ⟩ The complex coefficients α\alphaα and β indicate the degree to which the qubit is in a superposition of the states |0⟩and |1⟩. The likelihood of the qubit being in each state is given by the squared magnitudes $|\alpha|^2$and $|\beta|^2$, respectively. To ensure the conservation of total probability, a crucial requirement called normalization mandates that $|\alpha|^2+|\beta|^2=1$. This principle guarantees that the total probability of all possible qubit states sums to one, a foundational concept in quantum mechanics. Because α and β are complex, the qubit can exist in a complex superposition of states, allowing it to represent multiple values simultaneously through the principle of superposition. Leveraging qubit dependency, quantum computers can process large datasets more efficiently and tackle specific problems faster than classical systems [38], [39]. A fundamental concept in quantum computing is entanglement, which describes the interconnected states of multiple qubits, irrespective of the physical distance between them. This effect creates a dependency between the states of two qubits, highlighting their interconnected behavior. Utilizing this relationship allows quantum computers to manage large datasets more efficiently and solve specific problems faster than classical computers [38], [40].

Quantum circuits are composed of quantum gates that facilitate qubit manipulation during computations. The Hadamard (H) gate is essential, as it places qubits into a superposition state. When applied to qubits |0⟩ or |1⟩, the H gate transforms their states to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, respectively. For an n-qubit system, the combined state is represented by the tensor product of n single-qubit states, creating a superposition of 2^n basis states that includes every possible classical combination from |00 … 00⟩ to 11 … 11⟩ [41], [42]. Additionally, controlled-NOT (CNOT or CX) gates are frequently

used to generate entanglement between qubits, using one qubit as a control to determine if a NOT operation should be applied to the target qubit [42], [43]. Quantum circuits, formed by combining various quantum gates, are essential for quantum computing and information processing [44], [45].

2.6. QML Architecture and VQC Model

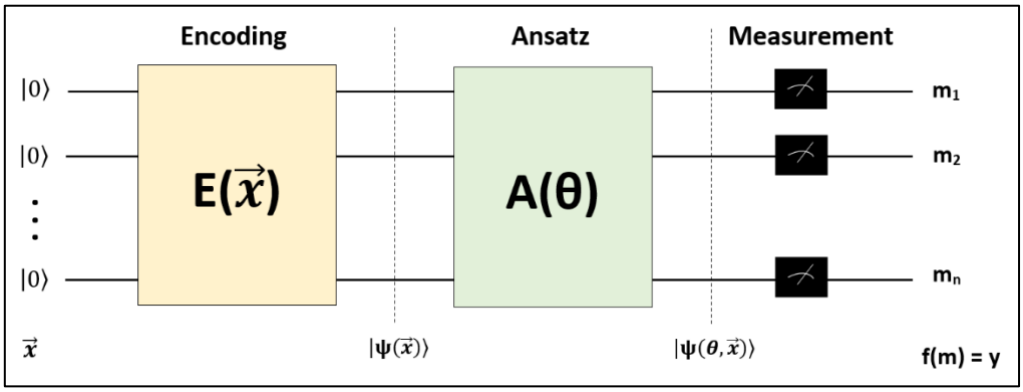

Figure 3. General architecture of QML

The primary distinction between quantum and classical computing lies in quantum systems' ability to leverage unique properties like superposition and qubit entanglement during computations [46], [47], [48]. Figure 3 presents the three core components of the QML architecture. The initial component, quantum feature mapping, functions as an input encoding technique that transforms classical data into quantum states within the Hilbert space, placing qubits into superposition states [49], [50]. As shown in Equation 3., the encoding $\mathbf{E}(\vec{\boldsymbol{x}})$is applied to the ground state $|\mathbf{0}\rangle$ of each qubit, allowing the classical input vector $\vec{\boldsymbol{x}}$ to be encoded as a quantum state vector $|\boldsymbol{\psi}(\vec{\boldsymbol{x}})\rangle$ for each observation. Each classical feature essentially corresponds to a mapping on a single qubit.

Equation 4. illustrates that the second component, known as the ansatz, $\mathbf{A}(\boldsymbol{\theta})$ utilizes flexible parameters (θ) to dynamically evolve the quantum states $|\boldsymbol{\psi}(\boldsymbol{\theta}, \vec{\boldsymbol{x}})\rangle$within the system, which can be expanded to n layers of circuits, typically incorporating CX gates and rotation gates at each layer. This parametrized circuit layout is particularly advantageous for tackling the problem, as its circuit depth increases linearly with the number of qubits. The ansatz serves as a foundational circuit framework where gates operate on designated subsystems, acting as an initial hypothesis for a training structure [51]. It offers a basic structure or format with adjustable variables, such as gate parameters (often angles), which can be fine-tuned through an optimization process [52]. The ansatz can be categorized into two primary types: the first aims to simplify the circuit to minimize quantum hardware computation. In contrast, the second seeks to identify the optimal ansatz for maximum performance on a specific task without reducing computation [25], [53].

The final component of the QML architecture involves determining the outcome, signifying the completion of circuit operations. The developed quantum states are measured to produce measurement results, denoted by the letter **m**. Subsequently, a classical computer performs additional post-processing on these results by decoding and extracting the output y as the function **f(m)**. This decoding process is the concluding step in the QML framework, finalizing the regression task and providing the ultimate output [25]. A loss function is computed by comparing the predicted values from the model with the target data, assessing how closely each sample's target values match the predictive model. Once the loss function is established, the next step in the model optimization process is to minimize this loss [54], [55].

$$\mathrm{E}(\vec{x})|0\rangle = |\psi(\vec{x})\rangle \quad (3)$$

$$|\psi(\theta, \vec{x})\rangle = \mathrm{A}(\theta)\mathrm{E}(\vec{x})|0\rangle \quad (4)$$

p-ISSN: 2528-1682
e-ISSN: 2527-9165

2.7. *QML Architecture and VQC Model*

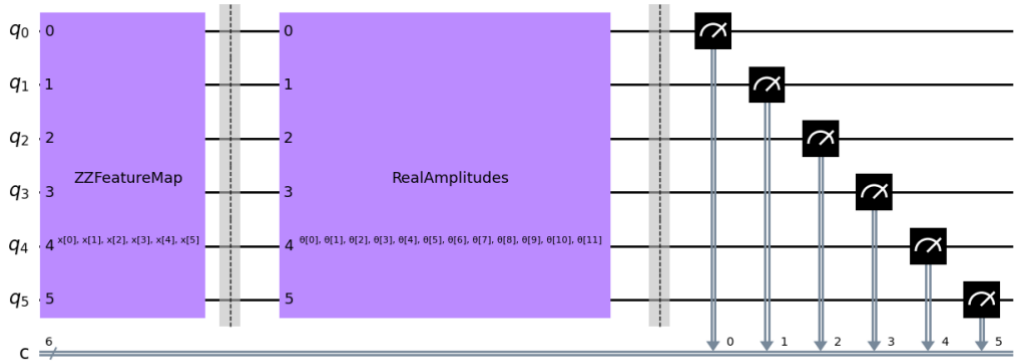

Figure 4. The structure of proposed VQC models VQC-1 (ZZFeatureMap as input encoding and RealAMplitudes as ansatz)

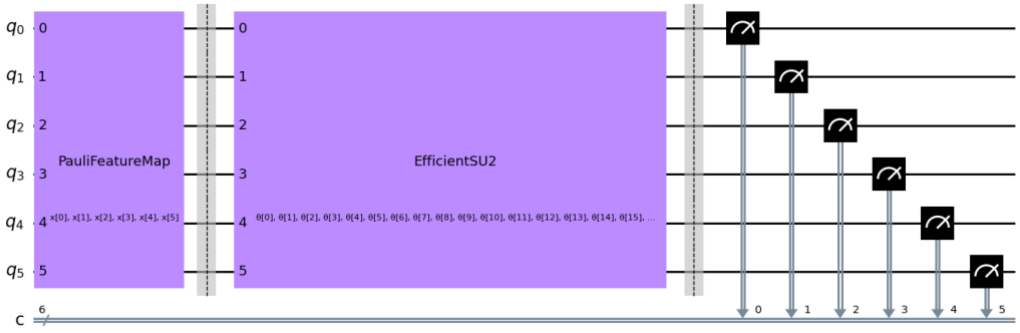

Figure 5. The structure of proposed VQC models (VQC-2 (PauliFeatureMap as input encoding and EfficientSU2 as ansatz)

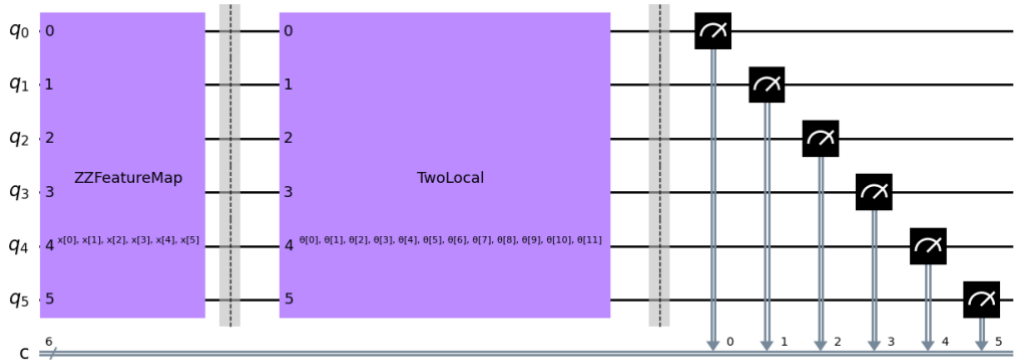

Figure 6. The structure of proposed VQC models VQC-3 (ZZFeatureMap as input encoding and TwoLocal as ansat

Hyperparameter tuning is employed to identify the best parameters for each model, enhancing its overall performance [56], [57]. Parameter optimization is important to improve model accuracy [58], [59]. As previously noted, this study developed VQC models. The effectiveness of all three models was evaluated in predicting the Chemical Information Encoding (CIE) of expired pharmaceutical compounds. the experiment involved 6 VQC models resulting from input encoding and ansatz. The input encoding techniques used are PauliFeatureMap and ZZFeatureMap, while the ansatz are RealAmplitudes, EfficientSU2, and TwoLocal, as depicted in Figure 3-5. These feature mapping parameters effectively convert classical data into more informative quantum representations, facilitating the search for optimal data encoding. In VQC classification tasks, quantum circuit-based variational parameters like RealAmplitudes, EfficientSU2, and TwoLocal are employed to determine the most suitable ansatz

structure for analyzing incoming data. By evaluating the similarity between data pairs in the quantum feature space, optimizing the encoding parameters and ansatz effectively addresses data classification challenges. It enhances the model's ability to learn and categorize data. Additionally, post-quantum parameterization involves optimization techniques aimed at minimizing the loss function, utilizing various methods such as analytic quantum gradient descent (AQGD), constrained optimization by linear approximations (COBYLA), and adaptive moment estimation (ADAM). Quantum models can continuously refine their performance in classification tasks by calculating and updating the gradient of the loss function, leading to more informative quantum feature mappings and improved ansatz for data processing.

2.8. QML Architecture and VQC Model

Key metrics such as RMSE, mean absolute error (MAE), and mean absolute deviation (MAD) evaluate a model's effectiveness. These metrics provide insights into how well the model's predictions align with the characteristics of the tested data, where lower values indicate fewer prediction errors and suggest greater accuracy of the model [9] [24], [31].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i' - Y_i)^2} \quad (5)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i' - Y_i| \quad (6)$$

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|Y_i' - \bar{Y}_i| \quad (7)$$

The parameters used in these evaluation formulas include in, which represents the number of samples in the dataset; Y_i, the actual observed values from the dataset; $\bar{Y}$, the mean value of the dataset, and Y_i', the predicted values generated by the model. Incorporating these characteristics into the evaluation criteria allows for a more comprehensive assessment of the model's performance and the predictive accuracy of the desired outcomes.

3. RESULT AND DISCUSSION

Table 1. Performance of diverse proposed VQC models

Model	Encoding	Ansatz	RMSE	MAE	MAD
VQC-1	ZZFeatureMap	RealAmlitude	6.27	5.51	5.21
VQC-2	ZZFeatureMap	EfficientSU2	6.30	6.12	6.09
VQC-3	ZZFeatureMap	TwoLocal	6.44	6.17	6.18
VQC-4	PauliFeatureMap	RealAmlitude	6.15	5.63	5.50
VQC-5	PauliFeatureMap	EfficientSU2	6.32	6.16	6.25
VQC-6	PauliFeatureMap	TwoLocal	6.50	6.34	6.35

The performance of the VQC in predicting the CIE values of expired pharmaceutical compounds as corrosion inhibitors is shown in Table 1. For RMSE, a lower score indicates better accuracy, as this metric calculates the average error between predictions and actual values. VQC-1 is more accurate than models VQC-2 and VQC-3, achieving the lowest RMSE score of 6.27 among models using ZZFeatureMap encoding. This result highlights that combining ZZFeatureMap encoding with the RealAmplitude ansatz effectively predicts data patterns.

Conversely, among the PauliFeatureMap models, VQC-6—which uses the TwoLocal ansatz with PauliFeatureMap encoding—shows the highest RMSE (6.50), MAE (6.34), and MAD (6.35) values, indicating more significant variability and reduced prediction accuracy. This suggests that this configuration may not adequately capture the dataset's complexity compared to other models. Notably, VQC-4 achieves competitive performance metrics (RMSE = 6.15, MAE = 5.63, MAD = 5.50) using PauliFeatureMap encoding with the RealAmplitude ansatz, effectively capturing data patterns comparable to those in ZZFeatureMap models. This highlights the importance of ansatz selection and encoding in optimizing VQC performance.

p-ISSN: 2528-1682
e-ISSN: 2527-9165

Overall, the investigation underscores the significant impact of encoding and ansatz settings on VQC model performance. Certain models, like VQC-6, illustrate the potential limitations of specific combinations, while others, such as VQC-1 and VQC-4, demonstrate the effectiveness of particular configurations in enhancing prediction accuracy. This comprehensive analysis highlights the need for tailored setups to optimize VQC performance for specific datasets and applications, providing insight into how different quantum circuit architectures influence prediction outcomes.

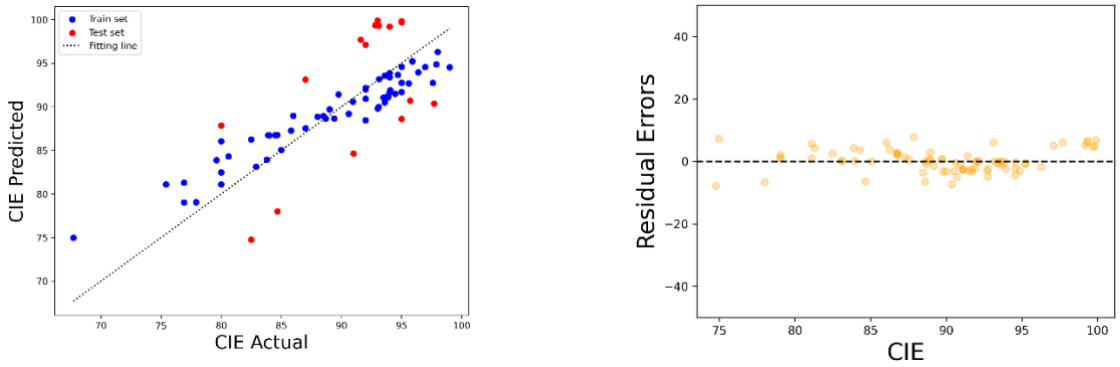

Figure 7. Scatter plots (left) and residual error plots (right) illustrate how various models predict individual data points.

The analysis presented in Figure 6, which includes scatter and residual error plots, demonstrates the effectiveness of the VQC model. The scatter plot shows a spread of data points around the diagonal line across all models, indicating variation in predicted values relative to actual values. In the scatter plot (left side), there is a strong correlation between actual and predicted values for both training data (blue points) and test data (red points), with the points closely aligning with the ideal line (y = x). This suggests that the model generally accurately predicts CIE values, particularly for the training data. In the residual plot (right side), residuals (the differences between predicted and actual values) are plotted against actual CIE values. The random distribution of residuals around the horizontal zero line indicates no systematic bias in the model's predictions. The even spread of residuals above and below the zero line suggests that prediction errors are not concentrated within a specific value range.

Table 2. Comparison with other works

	Model	RMSE	MAE	MAD	Ref.
QML:	VQC-1	6.27	5.51	5.21	This work
	VQC-4	6.15	5.63	5.50	This work
CML:	ARX	7.03	-	-	[3]
	GB	6.40	4.80	-	[19]

Table 2. compares QML models using the VQC method and CML models from previous studies in predicting the corrosion inhibition efficiency (CIE) of expired pharmaceutical compounds. The quantum models (VQC-1 and VQC-4) demonstrate competitive performance with classical models, especially regarding RMSE, where both models outperform ARX and GB. The comparison indicates that quantum models (QML), specifically VQC-1 and VQC-4, achieve prediction accuracy on par with classical models (CML) for CIE values. VQC-4 slightly outperforms VQC-1 in accuracy, with a lower RMSE (6.15 vs. 6.27), although it has slightly higher MAE and MAD values. The classical ARX model has the highest RMSE (7.03), indicating a more significant prediction error than the quantum and other classical models. In contrast, the GB model performs relatively well with an RMSE of 6.40 and a lower MAE (4.80), though MAD data for this classical model is unavailable.

In summary, VQC-4 surpasses VQC-1 by achieving lower prediction error, improved accuracy, and a more balanced distribution of errors centered around zero. The selection of encoding and ansatz configurations plays a crucial role in determining the effectiveness of QML models, with certain setups delivering enhanced predictive accuracy. Each QML model exhibits unique strengths and limitations

based on its encoding and ansatz choices. QML holds promise as a groundbreaking approach, transcending the constraints of traditional modeling techniques by uncovering new correlations and providing deep insights into corrosion inhibitors through molecular structure analysis. This innovative method has the potential to transform understanding and prediction in corrosion inhibitor research by revealing quantum-level details previously unexplored by conventional models.

Furthermore, research is ongoing to achieve reduced prediction errors. Increasing the quantity of samples in the dataset is one of our plans. Enlarging the dataset will aid the QML model's improved training and generalization by decreasing overfitting and increasing accuracy. (2) Construct the QML model using several circuit topologies. We can find more effective and efficient architectures for this task by experimenting with alternative quantum circuit designs, such as changing the amounts of qubits, types of gates, and depth of the circuits. (3) Put sophisticated feature engineering into practice. Improving the process of extracting and selecting features to better capture the underlying patterns pertinent to the quantum model in the data. (4) Examine strategies for mitigating quantum errors and using techniques to lessen the influence of mistakes and noise present in quantum computing to increase the accuracy of the findings. Our goal in developing these approaches is to improve the applicability, accuracy, and performance of QML models across various domains.

Moreover, research is ongoing to minimize prediction errors. One of our strategies is to increase the number of samples in the dataset. Expanding the dataset will help reduce overfitting and enhance accuracy, leading to better training and generalization of the QML model. (2) We plan to construct the QML model using various circuit topologies. By experimenting with different quantum circuit designs—such as adjusting the number of qubits, types of gates, and circuit depth—we aim to identify more effective and efficient architectures for this task. (3) We will implement advanced feature engineering to refine the features' extraction and selection process, allowing for better capture of the underlying patterns relevant to the quantum model in the data. (4) We will explore methods to mitigate quantum errors by employing techniques to reduce the impact of errors and noise inherent in quantum computing, thereby enhancing the accuracy of the results. Our objective in developing these approaches is to improve the applicability, accuracy, and performance of QML models across various domains.

4. CONCLUSION

This research investigates the potential of QML in predicting the corrosion inhibition capacity of expired pharmaceutical compounds. The proposed QML models show varying performance based on evaluation metrics, particularly in encoding and ansatz design, with the VQC method demonstrating competitive results against other CML methods. The study highlights the need for larger datasets to improve predictability while emphasizing QML's potential in exploring anti-corrosion materials. Despite its limitations, this research lays the groundwork for QML's ability to predict anti-corrosive properties.

ACKNOWLEDGEMENTS

This research was supported by Research Center for Quantum Computing and Materials Informatics Universitas Dian Nuswantoro. We thank our colleagues from Universitas Dian Nuswantoro, who provided insight and expertise that greatly assisted the research. However, they may not agree with all of the interpretations/conclusions of this paper.

REFERENCES

[1] Y. Cui, T. Zhang, and F. Wang, "New understanding on the mechanism of organic inhibitors for magnesium alloy," *Corros Sci*, vol. 198, p. 110118, Apr. 2022, doi: 10.1016/j.corsci.2022.110118.

[2] T. K. Sarkar, V. Saraswat, R. K. Mitra, I. B. Obot, and M. Yadav, "Mitigation of corrosion in petroleum oil well/tubing steel using pyrimidines as efficient corrosion inhibitor: Experimental and theoretical investigation," *Mater Today Commun*, vol. 26, p. 101862, Mar. 2021, doi: 10.1016/j.mtcomm.2020.101862.

[3] C. Beltran-Perez *et al.*, "A General Use QSAR-ARX Model to Predict the Corrosion Inhibition Efficiency of Drugs in Terms of Quantum Mechanical Descriptors and Experimental Comparison for Lidocaine," *Int J Mol Sci*, vol. 23, no. 9, p. 5086, May 2022, doi: 10.3390/ijms23095086.

[4] D. K. Kozlica, A. Kokalj, and I. Milošev, "Synergistic effect of 2-mercaptobenzimidazole and octylphosphonic acid as corrosion inhibitors for copper and aluminium – An electrochemical, XPS, FTIR and DFT study," *Corros Sci*, vol. 182, p. 109082, Apr. 2021, doi: 10.1016/j.corsci.2020.109082.

[5] M. Akrom, Investigation of natural extracts as green corrosion inhibitors in steel using density functional theory, Jurnal Teori Dan Aplikasi Fisika, 10(1), 89–102 (2022).

[6] C. Verma, E. E. Ebenso, and M. A. Quraishi, "Alkaloids as green and environmental benign corrosion inhibitors: An overview," *International Journal of Corrosion and Scale Inhibition*, vol. 8, no. 3, pp. 512–528, 2019, doi: 10.17675/2305-6894-2019-8-3-3.
[7] S. Lim and S. Chi, "Xgboost application on bridge management systems for proactive damage estimation," *Advanced Engineering Informatics*, vol. 41, p. 100922, Aug. 2019, doi: 10.1016/j.aei.2019.100922.
[8] M. Akrom, Green corrosion inhibitors for iron alloys: a comprehensive review of integrating data-driven forecasting, density functional theory simulations, and experimental investigation, Journal of Multiscale Materials Informatics, 1(1), 22-37 (2022), DOI: https://doi.org/10.62411/jimat.v1i1.10495.
[9] R. I. D. Putra, A. L. Maulana, and A. G. Saputro, "Study on building machine learning model to predict biodegradable-ready materials," 2019, p. 060003. doi: 10.1063/1.5095351.
[10] D. Kumar, V. Jain, and B. Rai, "Capturing the synergistic effects between corrosion inhibitor molecules using density functional theory and ReaxFF simulations - A case for benzyl azide and butyn-1-ol on Cu surface," *Corros Sci*, vol. 195, p. 109960, Feb. 2022, doi: 10.1016/j.corsci.2021.109960.
[11] T. W. Quadri *et al.*, "Multilayer perceptron neural network-based QSAR models for the assessment and prediction of corrosion inhibition performances of ionic liquids," *Comput Mater Sci*, vol. 214, p. 111753, Nov. 2022, doi: 10.1016/j.commatsci.2022.111753.
[12] A. H. Alamri and N. Alhazmi, "Development of data driven machine learning models for the prediction and design of pyrimidine corrosion inhibitors," *Journal of Saudi Chemical Society*, vol. 26, no. 6, p. 101536, Nov. 2022, doi: 10.1016/j.jscs.2022.101536.
[13] M. E. A. Ben Seghier, D. Höche, and M. Zheludkevich, "Prediction of the internal corrosion rate for oil and gas pipeline: Implementation of ensemble learning techniques," *J Nat Gas Sci Eng*, vol. 99, p. 104425, Mar. 2022, doi: 10.1016/j.jngse.2022.104425.
[14] M. Akrom, S. Rustad, A. G. Saputro, and H. K. Dipojono, "Data-driven investigation to model the corrosion inhibition efficiency of Pyrimidine-Pyrazole hybrid corrosion inhibitors," *Comput Theor Chem*, vol. 1229, p. 114307, Nov. 2023, doi: 10.1016/j.comptc.2023.114307.
[15] M. R. Rosyid, L. Mawaddah, A. P. Santosa, M. Akrom, S. Rustad, and H. K. Dipojono, "Implementation of quantum machine learning in predicting corrosion inhibition efficiency of expired drugs," *Mater Today Commun*, vol. 40, Aug. 2024, doi: 10.1016/j.mtcomm.2024.109830.
[16] T. W. Quadri *et al.*, "Development of QSAR-based (MLR/ANN) predictive models for effective design of pyridazine corrosion inhibitors," *Mater Today Commun*, vol. 30, p. 103163, Mar. 2022, doi: 10.1016/j.mtcomm.2022.103163.
[17] M. Akrom, S. Rustad, A. G. Saputro, A. Ramelan, F. Fathurrahman, and H. K. Dipojono, "A combination of machine learning model and density functional theory method to predict corrosion inhibition performance of new diazine derivative compounds," *Mater Today Commun*, vol. 35, p. 106402, Jun. 2023, doi: 10.1016/j.mtcomm.2023.106402.
[18] A. H. Alamri and N. Alhazmi, "Development of data driven machine learning models for the prediction and design of pyrimidine corrosion inhibitors," *Journal of Saudi Chemical Society*, vol. 26, no. 6, p. 101536, Nov. 2022, doi: 10.1016/j.jscs.2022.101536.
[19] T. H. Pham, P. K. Le, and D. N. Son, "A data-driven QSPR model for screening organic corrosion inhibitors for carbon steel using machine learning techniques," *RSC Adv*, vol. 14, no. 16, pp. 11157–11168, 2024, doi: 10.1039/D4RA02159B.
[20] S. Butmaratthaya, N. Buesamae, and U. Taetragool, "MNIST quantum classification models implementation and benchmarking," 2023, p. 070002. doi: 10.1063/5.0178776.
[21] N. Buesamae and U. Taetragool, "Tuning variational quantum classifier with automated design," 2023, p. 070003. doi: 10.1063/5.0179294.
[22] J. Alcazar, V. Leyton-Ortega, and A. Perdomo-Ortiz, "Classical versus quantum models in machine learning: insights from a finance application," *Mach Learn Sci Technol*, vol. 1, no. 3, p. 035003, Jul. 2020, doi: 10.1088/2632-2153/ab9009.
[23] M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, "Circuit-centric quantum classifiers," *Phys Rev A (Coll Park)*, vol. 101, no. 3, p. 032308, Mar. 2020, doi: 10.1103/PhysRevA.101.032308.
[24] M. Akrom, S. Rustad, and H. K. Dipojono, "Development of quantum machine learning to evaluate the corrosion inhibition capability of pyrimidine compounds," *Mater Today Commun*, vol. 39, p. 108758, Jun. 2024, doi: 10.1016/j.mtcomm.2024.108758.
[25] M. Akrom, S. Rustad, and H. K. Dipojono, "Variational quantum circuit-based quantum machine learning approach for predicting corrosion inhibition efficiency of pyridine-quinoline compounds," *Materials Today Quantum*, vol. 2, p. 100007, Jun. 2024, doi: 10.1016/j.mtquan.2024.100007.
[26] L. B. V. de Amorim, G. D. C. Cavalcanti, and R. M. O. Cruz, "The choice of scaling technique matters for classification performance," *Appl Soft Comput*, vol. 133, p. 109924, Jan. 2023, doi: 10.1016/j.asoc.2022.109924.
[27] I. Lukovits, E. Kálmán, and F. Zucchi, "Corrosion Inhibitors—Correlation between Electronic Structure and Efficiency," *CORROSION*, vol. 57, no. 1, pp. 3–8, Jan. 2001, doi: 10.5006/1.3290328.
[28] Y. Liu *et al.*, "A Machine Learning-Based QSAR Model for Benzimidazole Derivatives as Corrosion Inhibitors by Incorporating Comprehensive Feature Selection," *Interdiscip Sci*, vol. 11, no. 4, pp. 738–747, Dec. 2019, doi: 10.1007/s12539-019-00346-7.
[29] G. Ibarra-Vazquez, M. S. Ramírez-Montoya, and J. Miranda, "Data Analysis in Factors of Social Entrepreneurship Tools in Complex Thinking: An exploratory study," *Think Skills Creat*, vol. 49, p. 101381, Sep. 2023, doi: 10.1016/j.tsc.2023.101381.
[30] J. Linden and R. Marquis, "The influence of time on dynamic signature: An exploratory data analysis," *Forensic Sci Int*, vol. 348, p. 111577, Jul. 2023, doi: 10.1016/j.forsciint.2023.111577.
[31] M. Akrom, T. Sutojo, A. Pertiwi, S. Rustad, and H. Kresno Dipojono, "Investigation of Best QSPR-Based Machine Learning Model to Predict Corrosion Inhibition Performance of Pyridine-Quinoline Compounds," *J Phys Conf Ser*, vol. 2673, no. 1, p. 012014, Dec. 2023, doi: 10.1088/1742-6596/2673/1/012014.

[32] M. Ahsan, M. Mahmud, P. Saha, K. Gupta, and Z. Siddique, "Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance," *Technologies (Basel)*, vol. 9, no. 3, p. 52, Jul. 2021, doi: 10.3390/technologies9030052.
[33] M. Akrom, S. Rustad, A. G. Saputro, A. Ramelan, F. Fathurrahman, and H. K. Dipojono, "A combination of machine learning model and density functional theory method to predict corrosion inhibition performance of new diazine derivative compounds," *Mater Today Commun*, vol. 35, p. 106402, Jun. 2023, doi: 10.1016/j.mtcomm.2023.106402.
[34] T. Kumar, D. Kumar, and G. Singh, "Brain Tumour Classification Using Quantum Support Vector Machine Learning Algorithm," *IETE J Res*, vol. 70, no. 5, pp. 4815–4828, May 2024, doi: 10.1080/03772063.2023.2245350.
[35] A. Pyrkov *et al.*, "Quantum computing for near-term applications in generative chemistry and drug discovery," *Drug Discov Today*, vol. 28, no. 8, p. 103675, Aug. 2023, doi: 10.1016/j.drudis.2023.103675.
[36] Z. Deng, X. Wang, and B. Dong, "Quantum computing for future real-time building HVAC controls," *Appl Energy*, vol. 334, p. 120621, Mar. 2023, doi: 10.1016/j.apenergy.2022.120621.
[37] B. A. Alhayani, O. A. AlKawak, H. B. Mahajan, H. Ilhan, and R. M. Qasem, "Design of Quantum Communication Protocols in Quantum Cryptography," *Wirel Pers Commun*, Jul. 2023, doi: 10.1007/s11277-023-10587-x.
[38] P. Brown and H. Zhuang, "Quantum machine-learning phase prediction of high-entropy alloys," *Materials Today*, vol. 63, pp. 18–31, Mar. 2023, doi: 10.1016/j.mattod.2023.02.014.
[39] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, Sep. 2017, doi: 10.1038/nature23474.
[40] H. Gupta, H. Varshney, T. K. Sharma, N. Pachauri, and O. P. Verma, "Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction," *Complex & Intelligent Systems*, vol. 8, no. 4, pp. 3073–3087, Aug. 2022, doi: 10.1007/s40747-021-00398-7.
[41] R. Xia and S. Kais, "Hybrid Quantum-Classical Neural Network for Calculating Ground State Energies of Molecules," *Entropy*, vol. 22, no. 8, p. 828, Jul. 2020, doi: 10.3390/e22080828.
[42] Y. Kwak, W. J. Yun, S. Jung, and J. Kim, "Quantum Neural Networks: Concepts, Applications, and Challenges," Aug. 2021, [Online]. Available: http://arxiv.org/abs/2108.01468
[43] S. Aishwarya, V. Abeer, B. B. Sathish, and K. N. Subramanya, "Quantum Computational Techniques for Prediction of Cognitive State of Human Mind from EEG Signals," *Journal of Quantum Computing*, vol. 2, no. 4, pp. 157–170, 2020, doi: 10.32604/jqc.2020.015018.
[44] A. Sagingalieva, M. Kordzanganeh, N. Kenbayev, D. Kosichkina, T. Tomashuk, and A. Melnikov, "Hybrid Quantum Neural Network for Drug Response Prediction," *Cancers (Basel)*, vol. 15, no. 10, p. 2705, May 2023, doi: 10.3390/cancers15102705.
[45] N. Mishra *et al.*, "Quantum Machine Learning: A Review and Current Status," 2021, pp. 101–145. doi: 10.1007/978-981-15-5619-7_8.
[46] Z. Ozpolat and M. Karabatak, "Performance Evaluation of Quantum-Based Machine Learning Algorithms for Cardiac Arrhythmia Classification," *Diagnostics*, vol. 13, no. 6, p. 1099, Mar. 2023, doi: 10.3390/diagnostics13061099.
[47] M. J. Kholili, R. Muslim, and A. R. T. Nugraha, "A classical algorithm inspired by quantum neural network for solving a Bose-Hubbard-like system in phase-space representation," 2023, p. 070007. doi: 10.1063/5.0178381.
[48] T. Imanothai and U. Taetragool, "The effects of training quantum support vector machines with different samples from the same dataset," 2023, p. 070006. doi: 10.1063/5.0178310.
[49] G. Abdulsalam, S. Meshoul, and H. Shaiba, "Explainable Heart Disease Prediction Using Ensemble-Quantum Machine Learning Approach," *Intelligent Automation & Soft Computing*, vol. 36, no. 1, pp. 761–779, 2023, doi: 10.32604/iasc.2023.032262.
[50] E. I. Elsedimy, S. M. M. AboHashish, and F. Algarni, "New cardiovascular disease prediction approach using support vector machine and quantum-behaved particle swarm optimization," *Multimed Tools Appl*, vol. 83, no. 8, pp. 23901–23928, Aug. 2023, doi: 10.1007/s11042-023-16194-z.
[51] J. Qin, "Review of ansatz designing techniques for variational quantum algorithms," *J Phys Conf Ser*, vol. 2634, no. 1, p. 012043, Nov. 2023, doi: 10.1088/1742-6596/2634/1/012043.
[52] W. Aboumrad and D. Widdows, "Mod2VQLS: A Variational Quantum Algorithm for Solving Systems of Linear Equations Modulo 2," *Applied Sciences*, vol. 14, no. 2, p. 792, Jan. 2024, doi: 10.3390/app14020792.
[53] N. Nguyen and K.-C. Chen, "Quantum Embedding Search for Quantum Machine Learning," *IEEE Access*, vol. 10, pp. 41444–41456, 2022, doi: 10.1109/ACCESS.2022.3167398.
[54] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, Sep. 2017, doi: 10.1038/nature23474.
[55] M. Wieder, J. Fass, and J. D. Chodera, "Fitting quantum machine learning potentials to experimental free energy data: predicting tautomer ratios in solution," *Chem Sci*, vol. 12, no. 34, pp. 11364–11381, 2021, doi: 10.1039/D1SC01185E.
[56] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, Nov. 2020, doi: 10.1016/j.neucom.2020.07.061.
[57] Z. M. Alhakeem, Y. M. Jebur, S. N. Henedy, H. Imran, L. F. A. Bernardo, and H. M. Hussein, "Prediction of Ecofriendly Concrete Compressive Strength Using Gradient Boosting Regression Tree Combined with GridSearchCV Hyperparameter-Optimization Techniques," *Materials*, vol. 15, no. 21, p. 7432, Oct. 2022, doi: 10.3390/ma15217432.
[58] W. Herowati *et al.*, "Prediction of Corrosion Inhibition Efficiency Based on Machine Learning for Pyrimidine Compounds: A Comparative Study of Linear and Non-linear Algorithms," *KnE Engineering*, Mar. 2024, doi: 10.18502/keg.v6i1.15350.
[59] S. Budi *et al.*, "Implementation of Polynomial Functions to Improve the Accuracy of Machine Learning Models in Predicting the Corrosion Inhibition Efficiency of Pyridine-Quinoline Compounds as Corrosion Inhibitors," *KnE Engineering*, Mar. 2024, doi: 10.18502/keg.v6i1.15351.