Applied Science Innovation (ASI)
Published by College of Physical and Applied Sciences, MOUAU

- *Received:* 04 June 2026
- *Revised*: 16 June 2026
- *Accepted:* 18 July 2026
- *Published:* 27 July 2026

PREDICTION OF THE CORROSION INHIBITION EFFICIENCY OF *NEWBOULDIA LAEVIS* EXTRACT ON AA7075-T7351 IN 1.0 M HCL ENVIRONMENT USING SUPERVISED MACHINE LEARNING MODELS

Solomon C. Udensi[1]*

[1]Department of Physics, Federal University of Technology, Owerri, Imo State, Nigeria.

***Corresponding author: solomon.udensi@futo.edu.ng**

ABSTRACT

This study investigates the application of supervised machine learning (ML) techniques in predicting the corrosion inhibitive efficiency (IE) of *Newbouldia laevis* (NBL) extract on aluminium alloy AA7075-T7351 in a 1.0 M HCl environment. A gravimetric dataset comprising 65 samples and 16 features - including dip time (DipTime), inhibitor mass (Mass), corrosion rate (CRmmYr), and weight loss (Loss) - was analysed. Preprocessing steps involved the removal of quasi-constant and redundant variables, alongside feature selection using Kendall correlation and gradient boosting decision tree methods. Four regression algorithms - Lasso, Ridge, Support Vector Regressor (SVR), and Extreme Gradient Boosting (XGBoost) - were implemented and evaluated using varying train-test sizes. Model performance was assessed using root mean square error (RMSE) and coefficient of determination (R^2). Results indicated that an 80:20 train-test split yielded optimal performance across models. Lasso regression achieved the highest predictive accuracy (R^2 = 97.45%) with the lowest RMSE, followed closely by Ridge and SVR, while XGBoost showed slightly lower test performance but higher training accuracy. Feature importance analysis revealed that inhibitor mass and dip time were consistently significant predictors of IE. The findings demonstrate that ML models, particularly Lasso regression, provide reliable and efficient alternatives for predicting corrosion inhibition, reducing experimental cost, time, and environmental impact associated with extensive use of natural resources.

Keywords: Machine learning, Corrosion Inhibition, Green extract, Aluminum alloy, Feature engineering.

1. Introduction

The advantages of trees to man cannot be over emphasized. Trees byproducts, which most times are bio-degradable, say, may be extracted from the leaves (Udensi *et al*., 2020; Desai *et al*., 2023), roots (Obot & Obi-Egbedi, 2009; Kaya *et al*., 2023), tree barks (Chaubey *et al*., 2017; Ali *et al*., 2019), fruits (Da Rocha *et al*., 2010; Sanaei *et al*., 2019) among others, have been reported to yield extracts which are excellent corrosion inhibitors in various media (Udensi *et al*., 2021; Hossain *et al.*, 2021). However, some of these

trees are felled down before these byproducts are acquired, mostly without replacement, thereby adding to the extinction process that is injurious to the planet. In addition, the ever-growing costs of corrosion (Komary, *et al.*, 2023) indicate that trees alone cannot solve all corrosion problems without a drastic imbalance in nature. The enormous amount of time, sometimes in years, required to complete research on simulated or on-field corrosion experiments means that a constant feed of these exhaustible extracts are needed. Worse still, these experiments may not be favorable, leading to colossal waste of our flora and time. These days, accelerated simulated experiments that last, but a few hours Diao *et al.*, 2021) have been devised to obtain a pattern for the inhibiting action of these extracts. Subsequently, use of algorithms and statistical models are made to simulate the desired property after learning the sequence of events.

Hitherto, these algorithms are reported to have been used to model the; initiation time of steel's corrosion embedded in concrete (Salami *et al*., 2020), aging defects of pipelines (Akhlaghi, *et al.*, 2023), corrosion rate in CO_2 (Abbas *et al*., 2018) and marine environments (Zhi *et al*., 2019). In addition, researchers used these machine learning models to inspect corrosion stages in materials (Sanchez *et al*., 2020), detect corrosion for automatic image analysis (Nash *et al*., 2022), characterize the spatial distribution of pitting corrosion (Abass *et al*., 2020), predict corrosion rate using electrochemical impedance spectroscopy (EIS) method (Aghaaminiha, *et al.*, 2021) among others. However, gravimetrically investigated inhibitive effectiveness depends jointly on the dip time of the specimen in the testing environment, mass of inhibitor, weight loss of the specimen, area of specimen exposed etc. These variables are used in machine learning models, as features in predicting the accuracy of a target outcome (Alvarsson, *et al.*, 2016). There is virtually no exhaustive gravimetric study, in which ML has been used to determine the predictive accuracy of an inhibitor.

Therefore, this study presents a comprehensive supervised machine learning, of a small corrosion dataset using non-linear regression algorithms such as extreme gradient boosting (XGBoost), support vector regressor (SVR) (Moein, *et al.*, 2023), and the related ridge regression (RR) lasso regression (LR) to predict the inhibitive efficiency of *Newbouldia laevis* (NBL) extract on AA7075-T7351 in 1.0 M HCl corrosive environment.

2. Materials and Method

2.1 Corrosion data

The dataset used in this study was one obtained by immersing aluminium alloy AA7075-T7351 in 1.0 M HCl, containing various masses of the inhibitor extracts (100 – 500 mg/L), at different dip times ranging from 60 to 780 minutes and tabulated. The shape of the data frame is 65 x 16 (65 rows by 16 columns). The features include the: dip time (DipTime); corrosion rate in millimeter per year (CRmmYr); mass of NBL inhibitor in grams (Mass); acidic environment concentration (Conc); specimen's initial weight (IniWt), final weight (FnlWt), weight loss (Loss), surface area (SurfArea), volume (Vol), density (Den) and inhibitive efficiency of the NBL extract (IE). Other features are the specimen's length (Length), breadth (Breadth); thickness (Thickness); diameter of the hole on the specimen for easy dip/withdrawal (Diam) and the surface coverage of the inhibiting action (SC). The target for this study is IE of the NBL extract.

2.2 Feature selection and importance

Kendall correlation analysis was used to weigh whether any pair of the input features are consistent or inconsistent with each other (1, 0, and -1 for perfectly positive, no and perfectly negative correlations

respectively), and the visibly paired is chosen (Wen, *et al.*, 2019). Kendall correlation analysis is normally applied for small datasets, and is given mathematically as in equation 1.

$$\tau = \frac{n_c - n_d}{\left(\frac{n(n-1)}{2}\right)} \quad \ldots \quad (1)$$

where

n_c = number of concordant pairs (*x, y*)
n_d = number of discordant pairs
n = data size
$\frac{n(n-1)}{2}$ = number of ways to choose any pair from the data size.

In addition, the Gradient boosting decision tree (GBDT) was also used as ensemble learning technique, based on a set of features in the study, to make decision that improves the overall model (Yan *et al.*, 2020), tree after tree. In this study, 100 trees were used. Hence, the XGBoost optimized objective function, Θ, comprising the mean squared error (MSE) loss and a regularization term, is defined in equation 2.

$$\Theta = \sum_{j=1}^{n} loss(y_j, \hat{y}_j) + \sum_{k=1}^{k} \Omega(f_k) \quad \ldots \quad (2)$$

Where n equals number of training samples; y_j equals actual label for the j-th sample; $\hat{y}_j$ equals predicted value for the j-th sample; $loss$ equals the function calculating the difference between the true and predicted values; k equals the number of leaves in all the trees and $\Omega(f_k)$ equals the regularization term for the k-th tree.

2.3. Experimental procedure

Four algorithms, namely: Lasso regression, Ridge regression, support vector machine and XGBoost, were tested on a small corrosion dataset, in the python environment, under default parameter conditions, using the scikit-learn library. In order to determine which of the algorithms predicted the corrosion inhibitive efficiency of NBL in 1.0 M HCl better. The dataset was divided into testing and training sets for, the algorithms' inhibition efficiencies optimization and the models' prediction accuracies classification respectively. This study considered different ratios of test set and train set respectively, specifically, 80:20, 70:30, 60:40, 50:50, 40:60, 30:70, and 20:80. The mathematical formula for RMSE metric (i.e. the square root of MSE) is given mathematically in equation 3. The lower the error value of RMSE the better the prediction would be.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i-1}^{n}\left(Z_i - \hat{Z}_\iota\right)^2} \quad \ldots \quad (3)$$

Another complimentary verification metric (use of RMSE alone is not adequate) considered is the coefficient of determination (R^2), which quantifies the goodness of fit the models. It is given mathematically by equation 4.

$$R^2 = 1 - \frac{\sum_{i-1}^{n}\left(Z_i - \hat{Z}_\iota\right)^2}{\sum_{i-1}^{n}\left(Z_i - \bar{\bar{Z}}_\iota\right)^2} \quad \ldots \quad (4)$$

Where Z_i, $\hat{Z}_\iota$ and $\bar{\bar{Z}}_\iota$ in equations 3 and 4 are respectively the actual, predicted and average values.

3. Results and Discussion

3.1 Empirical data analysis

Figure 1 shows the pair plots obtained from our data, via "Seaborn library" in Python, to visualize the relationships between the corrosion features in order to determine the inhibition trend. The plots helped to identify data points that are markedly different in value from other data points in the dataset. The features and target, from the pair plots showed the existence of little collinearities with each pair. The plots revealed two quasi-constant features in "conc" and "Diam", a reminder that the values in these columns were almost constant throughout the dataset, hence of minimal consequence. In addition, the plots revealed that the target "IE" was equal to a hundred "SC", verifiable from the axes, signifying perfect correlation.

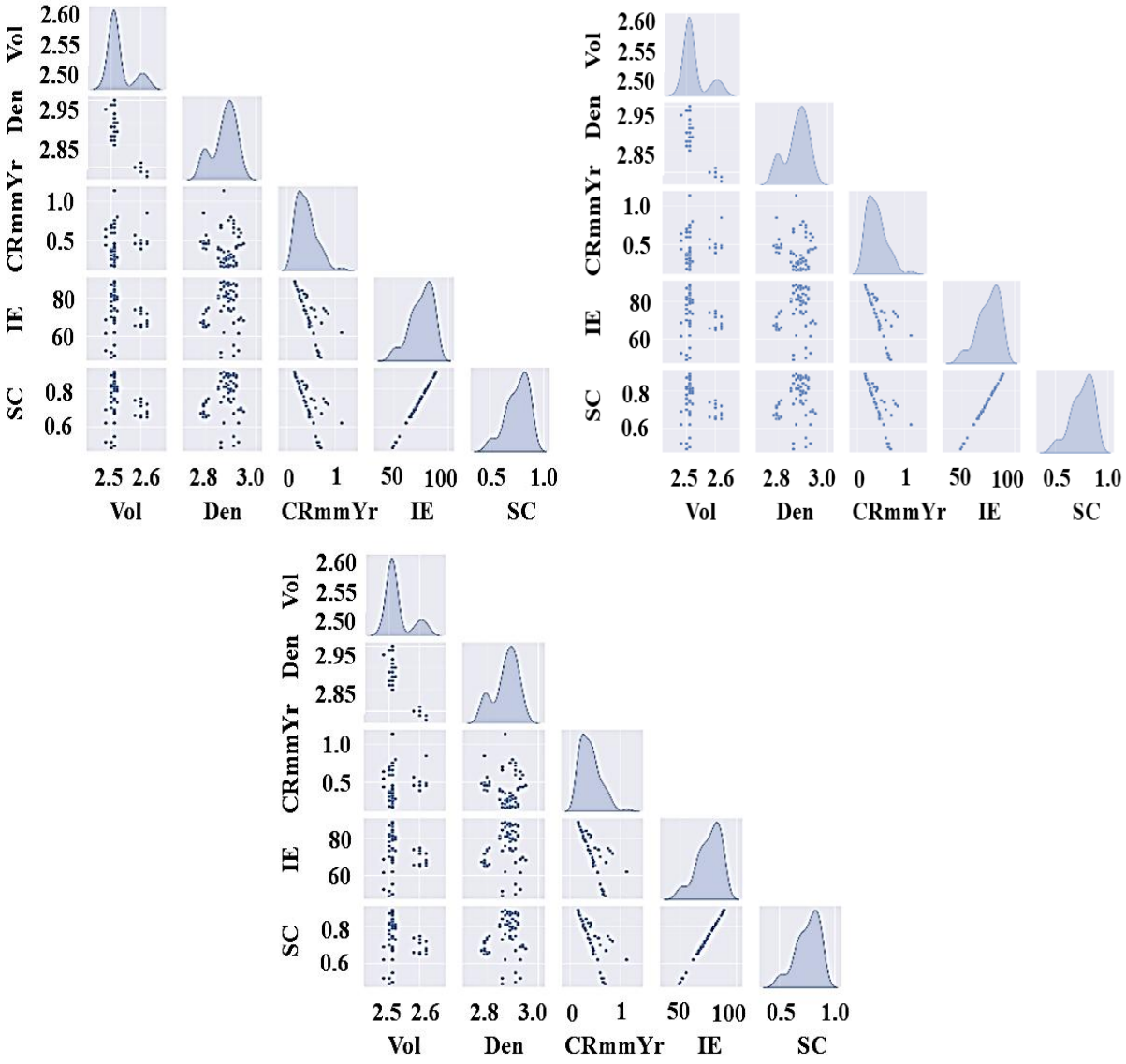


Figure 1: The pair plots depicting the pairwise relationships in the dataset

The three features; "conc", "Diam" and "SC" were discarded, and the Kendall correlation matrix plotted in Figure 2. The Kendall's tau correlation coefficients (values) in the map show the ranking of the variables, an indication of whether there is a perfect agreement or disagreement in the rank of variables. The value of 0.84 showed a positive correlation of "Mass" and "IE", meaning that an increase in the inhibitor "Mass" will necessarily bring about a corresponding increase in "IE". Also, the biggest negative correlation was seen between "CRmmYr" and "Mass" (i.e. -0.63), indicating, rightly, that an increase in "CRmmYr" would bring about corresponding decrease in "IE'. Hence, the ranking of the features with the target was "Mass", "Thickness", "Den", "FnlWt", "DipTime", "IniWt", "Vol", "length", "SurfArea", "breadth", "loss", "CRmmYr".

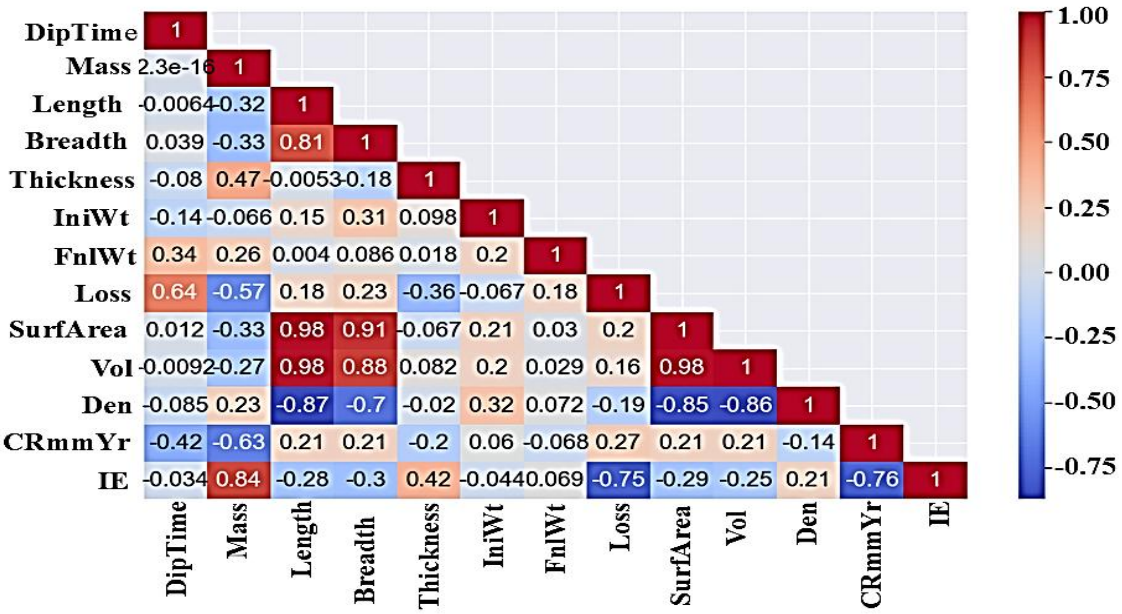


Figure 2: Kendall correlation map showing the pairwise relationships in the dataset

3.2 Feature selection for "IE" prediction

The model performance was improved by dropping the quasi – constant variables namely; "conc", "Diam" and "SC". Secondly, the dimensionality of the dataset was reduced without the three features. Thirdly, smaller dimensionality means reduced overfitting of the algorithms and improved protection from the noise in the data (Xiong *et al., 2024*). Figure 3 shows the graph of variance of features obtained after preprocessing of the dataset. "DipTime" showed the presence of extreme values because of the large variance of features score (i.e. 51187.5). This is corroborated from the dataset where the range of "DipTime" was 60 to 780 minutes. Features with very low variance of features contributed minimally to the predicting algorithms Kondo *et al.*, 2019).

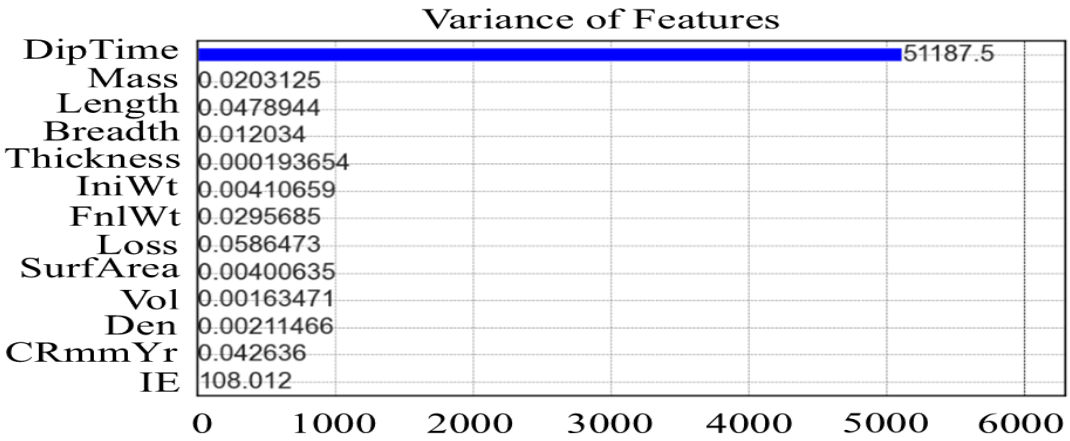


Figure 3: Variance of feature scores for NBL inhibitive action on AA7075-T7351 in 1.0 M HCl environment

3.3 Comparison of the algorithms' predictive performance

The 'percent test score' (i.e. R^2 scores) versus the 'percent train size' of the four algorithms used to model the small data set of the inhibitive efficiency of NBL on AA7075-T7351 in 1.0 M HCl corrosive environments are shown in Figure 4 (left). It would be seen that the best train size which gave rise to the best set of test scores, for all four models was 80 % (i.e. 20 % test size), and followed the order; Lasso (97.45 %), Ridge (97.36 %), SVR (97.00 %) and XGBoost (95.85 %). Lower train sizes saw XGBoost falling off rapidly from the pack, reaching a lowest test score of 79.08 % for 20% train size. The results were corroborated by the plot of 'RMSE of test size' versus 'percent train size', shown in Figure 4 (right), where Lasso, having the least RMSE score performed best and closely followed by Ridge algorithm. Three algorithms, namely, SVR, Lasso and Ridge could replace each other at train size of 30 %, because their test scores and RMSE values were approximately the same. Xgboost algorithm was thought to have 'underfit' the small dataset because the small training size might not have been representative of the whole data distribution, leading to the failure to generalize (i.e. learning to optimize the loss function in order to minimize predicted errors in the ensemble) unseen data (Wang,*et al.*, 2022).

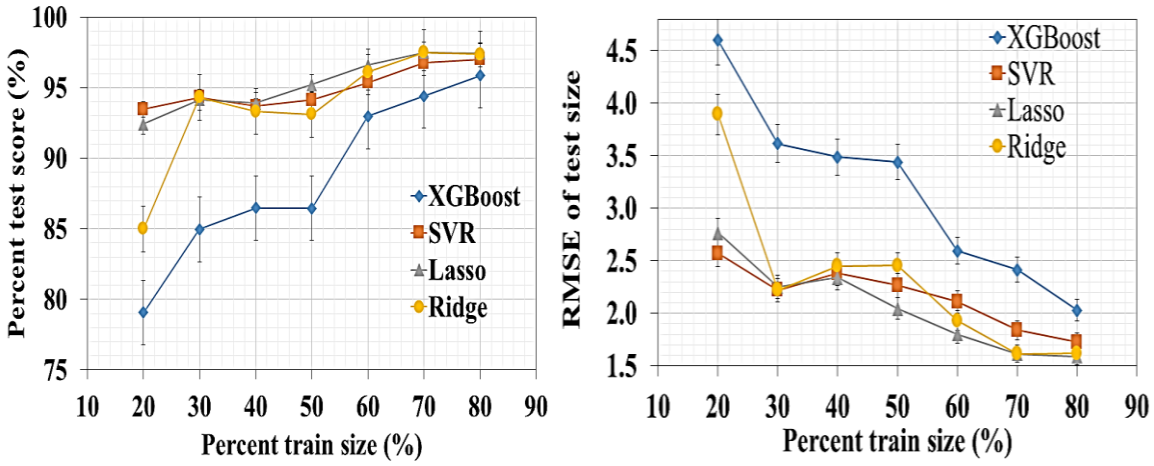


Figure 4: Algorithms percentage test score (left) and the predictive inhibitive performance of the algorithms as a function of percent train size

3.4 Feature Importance of the algorithms

The feature importance graphs of the four algorithms used in this work are independent of each other and are shown in Figure 5. The number of features considered by the four algorithms are 7 for XGBoost, 6 for Ridge, 4 for each of Lasso and SVR. However, in SVR algorithm there are two positively correlated and two negatively correlated features with the target variable IE.

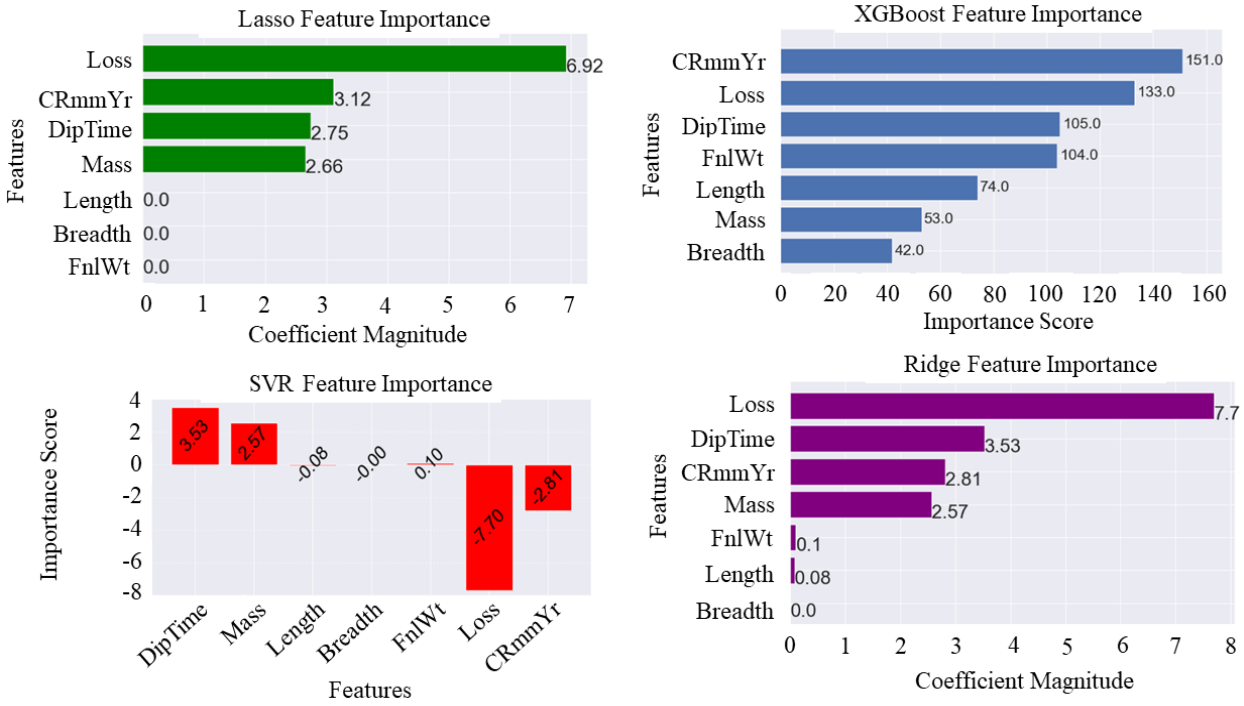


Figure 5: Algorithms choice of features importance for the inhibitive efficiency prediction of NBL extract on AA7075-T7351 in 1.0 M HCl environment. Lasso (top left), XGBoost (top right), SVR (bottom left) and Ridge (bottom right).

This is as a result of SVR's use of a hyperparameter that helped interpret importance in terms of weights given to each feature (Udensi & Ejelonu, 2024). These results show that the order of feature importance selection is XGBoost, Ridge, Lasso and SVR in determining the inhibitive efficiency of NBL. XGBoost considers CRmmYr as the most important variable, while Ridge, Lasso and SVR consider weight loss, and DipTime respectively. The features that recur in all the four algorithms DipTime and mass of NBL, meaning that they correlated positively with the target IE.

Figure 6 shows one of the XGBoost decision trees in which the most significant feature for IE prediction was CrmmYr, with the root node splitting at 0.4. Subsequent splits on DipTime reveal a conditional relationship, where exposure time modulates the effect of CRmmYr on IE (Xu *et al.*, 2025: Kang, Kim, & Lim, 2023). This highlights the importance of contact time for AA7075 in the corrosive environment, reflecting time-dependent surface processes that influence inhibition efficiency.

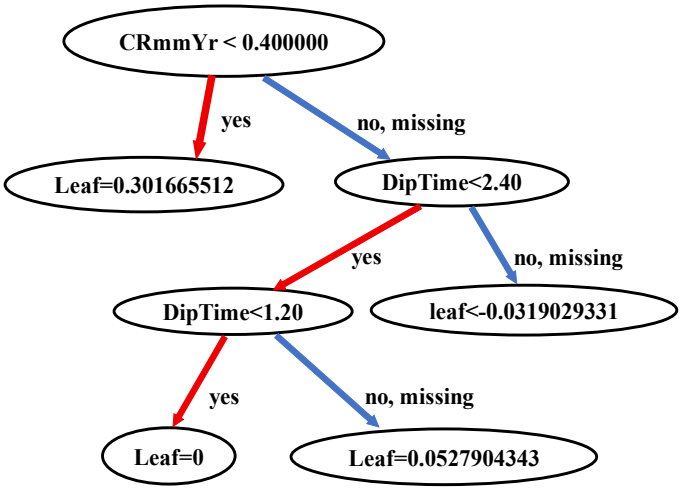


Figure 6: Algorithms choice of features importance for the inhibitive efficiency prediction of NBL extract on AA7075-T7351 in 1.0 M HCl environment. Lasso (top left), XGBoost (top right), SVR (bottom left) and Ridge (bottom right)

3.5 Coefficient of determination (R^2) of the algorithms

The R^2 scores, (the coefficient of determination), were used to predict the accuracy of the IE of the NBL extract for the four tested algorithms. The graphs of predicted values versus actual values were plotted and are shown in Figure 7. The XGBoost algorithm achieved the highest training score of 99.77%, likely because it utilized the most features during training.

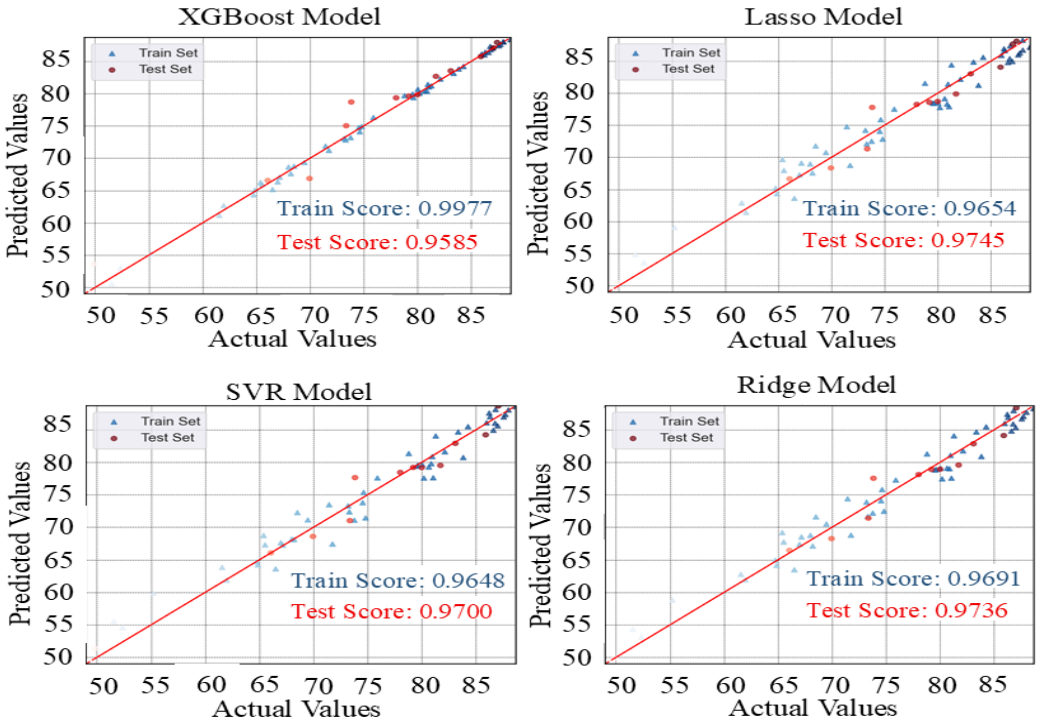


Figure 7: NBL Inhibitive Efficiency Prediction Accuracy in 1.0 M HCl: XGBoost, Lasso, SVR, and Ridge (20% Test Size)

Following closely behind were Ridge, Lasso, and SVR, with scores of 96.91%, 96.54%, and 96.48%, respectively. SVR performed similarly to the other algorithms because it used DipTime and NBL extract mass, which are positively correlated with IE, indicating they are key variables in reducing corrosion (Udensi & Ejelonu, 2024).

Table 1 shows the R^2 scores for the training and test sets of the four algorithms, along with the RMSE and MSE values for the NBL extract in a 1.0 M HCl environment. The test R^2 scores (accuracy) decrease in the order: Lasso (97.45%), Ridge (97.36%), SVR (97.00%), and XGBoost (95.85%).

Table 1: The coefficient of determination (R^2) for training set and test set, root-mean square error (RMSE), mean square error (MSE) for the accuracy prediction of the inhibitive efficiency of NBL extract in 1.0 M HCl using XGBoost, Lasso, SVR and Ridge algorithms respectively.

Algorithm	R^2_score		RMSE	MSE	R^2 Accuracy
	Training set	Test set			
XGBoost	0.9977	0.9585	2.0280	4.1129	0.9585
Lasso	0.9654	0.9745	1.5906	2.5300	0.9745
SVR	0.9648	0.9700	1.7233	2.9699	0.9700
Ridge	0.9691	0.9736	1.6168	2.6140	0.9736

This trend is consistent with the RMSE and MSE values, where lower values indicate better performance (Diao, Yan & Gao, 2021: Dong *et al*., 2024). Lasso performed best, with the lowest RMSE of 1.5906 because its use of L1 regularization focused on the most important features and avoided overfitting, allowing it to make accurate prediction on new datasets. SVR underperformed relative to other algorithms because it lacked the built-in feature selection mechanisms, rendering it less effective at isolating the most predictive features. Also, its sensitivity to hyperparameters and kernel choice can reduce generalization, particularly when the fundamental relationships are mostly linear (He, Wang & Jiang, 2008).

3.6 Models' capabilities

The comparison between experimental inhibitive efficiency values and predictions from four machine learning algorithms - XGBoost, SVR, Lasso, and Ridge – is shown in Figure 8. The bar charts reveals that all models exhibited strong predictive performance with close agreement to the actual data across the evaluated dataset. Among them, XGBoost provided the most accurate predictions, with a slight tendency to overestimate in certain instances. In contrast, SVR, Lasso, and Ridge presented nearly identical results, indicating comparable and stable performance. The algorithms achieved higher accuracy at higher (70% and above) inhibitive efficiency values, while minor deviations were more noticeable at lower ranges (50 – 60%). Overall, the findings highlighted the reliability of all four algorithms, with XGBoost showing a marginal advantage in predictive accuracy.

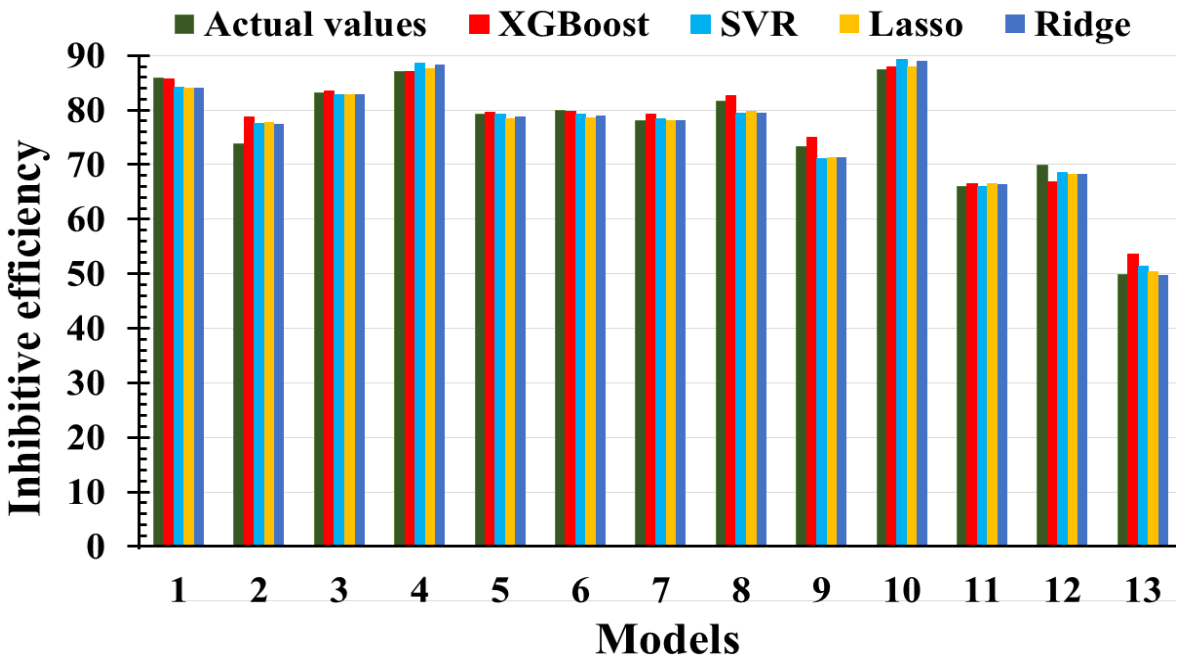


Figure 8: The predictive performance of XGBoost, Lasso, SVR and Ridge algorithms for the inhibitive efficiency of NBL in 1.0 M HCl corrosive environment.

4 Conclusion

This study successfully demonstrated the effectiveness of machine learning techniques in predicting the corrosion inhibitive efficiency of *Newbouldia laevis* extract on AA7075-T7351 aluminium alloy in an acidic medium. The results showed that careful preprocessing, including the removal of quasi-constant and redundant features, significantly improved model performance and reduced overfitting. Among the four algorithms evaluated, Lasso regression emerged as the best-performing algorithm, achieving the highest predictive accuracy and lowest root mean square error, owing to its ability to perform feature selection and mitigate overfitting through L1 regularization.

Ridge and SVR models also exhibited strong and consistent predictive capabilities, while XGBoost, although highly accurate during training, showed signs of underfitting on the small dataset. Feature importance analysis across all models consistently identified inhibitor mass and dip time as the most influential variables governing inhibition efficiency, highlighting their critical role in corrosion control.

Overall, the study confirms that machine learning provides a robust and efficient framework for modeling corrosion inhibition processes, especially when experimental datasets are limited. This approach not only accelerates prediction but also minimizes reliance on laborious laboratory experiments and the excessive use of plant-based materials, thereby supporting sustainable research practices. Future work may focus on expanding dataset size, optimizing hyperparameters, and integrating hybrid or deep learning approaches to further enhance predictive performance.

Conflict of Interest

No conflict of interest was declared by the author

References

Abass, A., Wada, K., Matsunaga, H., Remes, H., & Vuorio, T. (2020). Quantitative characterization of the spatial distribution of corrosion pits based on nearest neighbor analysis. *Corrosion*, *76*(9), 861-870.

Abbas, M. H., Norman, R., & Charles, A. (2018). Neural network modelling of high pressure CO2 corrosion in pipeline steels. *Process Safety and Environmental Protection*, *119*, 36-45.

Aghaaminiha, M., Mehrani, R., Colahan, M., Brown, B., Singer, M., Nesic, S., ... & Sharma, S. (2021). Machine learning modeling of time-dependent corrosion rates of carbon steel in presence of corrosion inhibitors. *Corrosion Science*, *193*, 109904.

Akhlaghi, B., Mesghali, H., Ehteshami, M., Mohammadpour, J., Salehi, F., & Abbassi, R. (2023). Predictive deep learning for pitting corrosion modeling in buried transmission pipelines. *Process Safety and Environmental Protection*, *174*, 320-327.

Ali, I. H., Idris, A. M., & Suliman, M. H. (2019). Evaluation of leaf and bark extracts of Acacia tortilis as corrosion inhibitors for mild steel in seawater: Experimental and studies. *International Journal of Electrochemical Science*, *14*(7), 6406-6419.

Alvarsson, J., Lampa, S., Schaal, W., Andersson, C., Wikberg, J. E., & Spjuth, O. (2016). Large-scale ligand-based predictive modelling using support vector machines. *Journal of cheminformatics*, *8*, 1-9.

Chaubey, N., Savita, Singh, V. K., & Quraishi, M. A. (2017). Corrosion inhibition performance of different bark extracts on aluminium in alkaline solution. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, *22*(1), 38-44.

Da Rocha, J. C., Gomes, J. A. D. C. P., & D'Elia, E. (2010). Corrosion inhibition of carbon steel in hydrochloric acid solution by fruit peel aqueous extracts. *Corrosion Science*, *52*(7), 2341-2348.

Desai, P. S., Desai, F., Patel, A., & Parmar, B. (2023). Anticorrosive properties of Eucalyptus (Nilgiris) leaves extract on 2S grade aluminium in acid solutions. *Applied Surface Science Advances*, *16*, 100414.

Diao, Y., Yan, L., & Gao, K. (2021). Improvement of the machine learning-based corrosion rate prediction model through the optimization of input features. *Materials & Design*, *198*, 109326.

Dong, Z., Ding, L., Meng, Z., Xu, K., Mao, Y., Chen, X., ... & Poursaee, A. (2024). Machine learning-based corrosion rate prediction of steel embedded in soil. *Scientific Reports*, *14*(1), 18194.

He, W., Wang, Z., & Jiang, H. (2008). Model optimizing and feature selecting for support vector regression in time series forecasting. *Neurocomputing*, *72*(1-3), 600-611.

Hossain, N., Chowdhury, M. A., Iqbal, A. P., Islam, M. S., Omar, N. Y. S., & Saifullah, A. Z. A. (2021). Paederia Foetida leaves extract as a green corrosion inhibitor for mild steel in hydrochloric acid solution. *Current Research in Green and Sustainable Chemistry*, *4*, 100191.

Kang, Y., Kim, M. G., & Lim, K. M. (2023). Machine-learning based prediction models for assessing skin irritation and corrosion potential of liquid chemicals using physicochemical properties by XGBoost. *Toxicological Research*, *39*(2), 295-305.

Kaya, F., Solmaz, R., & Gecibesler, I. H. (2023). Adsorption and corrosion inhibition capability of Rheum ribes root extract (Işgın) for mild steel protection in acidic medium: A comprehensive electrochemical, surface characterization, synergistic inhibition effect, and stability study. *Journal of Molecular Liquids*, *372*, 121219.

Komary, M., Komarizadehasl, S., Tošić, N., Segura, I., Lozano-Galant, J. A., & Turmo, J. (2023). Low-cost technologies used in corrosion monitoring. *Sensors*, *23*(3), 1309.

Kondo, M., Bezemer, C. P., Kamei, Y., Hassan, A. E., & Mizuno, O. (2019). The impact of feature reduction techniques on defect prediction models. *Empirical Software Engineering*, *24*(4), 1925-1963.

Moein, M. M., Saradar, A., Rahmati, K., Mousavinejad, S. H. G., Bristow, J., Aramali, V., & Karakouzian, M. (2023). Predictive models for concrete properties using machine learning and deep learning approaches: A review. *Journal of Building Engineering*, *63*, 105444.

Nash, W., Zheng, L., & Birbilis, N. (2022). Deep learning corrosion detection with confidence. *npj Materials degradation*, *6*(1), 26.

Obot, I. B., & Obi-Egbedi, N. O. (2009). Ginseng root: a new efficient and effective eco-friendly corrosion inhibitor for aluminium alloy of type AA 1060 in hydrochloric acid solution. *International Journal of Electrochemical Science*, *4*(9), 1277-1288.

Salami, B. A., Rahman, S. M., Oyehan, T. A., Maslehuddin, M., & Al Dulaijan, S. U. (2020). Ensemble machine learning model for corrosion initiation time estimation of embedded steel reinforced self-compacting concrete. *Measurement*, *165*, 108141.

Sanaei, Z., Ramezanzadeh, M., Bahlakeh, G., & Ramezanzadeh, B. (2019). Use of Rosa canina fruit extract as a green corrosion inhibitor for mild steel in 1 M HCl solution: A complementary experimental, molecular dynamics and quantum mechanics investigation. *Journal of industrial and engineering chemistry*, *69*, 18-31.

Sanchez, G., Aperador, W., & Cerón, A. (2020). Corrosion grade classification: a machine learning approach. *Indian Chemical Engineer*, *62*(3), 277-286.

Udensi, S. C. & Ejelonu, B. O. (2024). Gravimetric inhibition efficiency prediction model of AA7075-T7351 alloy using *Treculia Africana* extract in 1.0 M HCl through input feature optimization. *Surface Science and Technology* 2, 22.

Udensi, S. C., Ekpe, O. E., & Nnanna, L. A. (2020). *Newbouldia laevis* leaves extract as tenable eco-friendly corrosion inhibitor for aluminium alloy AA7075-T7351 in 1 M HCl corrosive environment: gravimetric, electrochemical and thermodynamic studies. *Chemistry Africa*, *3*(2), 303-316.

Udensi, S. C., Ekpe, O. E., & Nnanna, L. A. (2021). Corrosion inhibition performance of low cost and eco-friendly *Treculia africana* leaves extract on aluminium alloy AA7075-T7351 in 2.86% NaCl solutions. *Scientific African*, *12*, e00791.

Wang, Y., Su, F., Guo, Y., Yang, H., Ye, Z., & Wang, L. (2022). Predicting the microbiologically induced concrete corrosion in sewer based on XGBoost algorithm. *Case Studies in Construction Materials*, *17*, e01649.

Wen, C., Zhang, Y., Wang, C., Xue, D., Bai, Y., Antonov, S., ... & Su, Y. (2019). Machine learning assisted design of high entropy alloys with desired property. *Acta Materialia*, *170*, 109-117.

Xiong, W., Yu, G., Ma, J., & Liu, S. (2024). A novel robust adaptive subspace learning framework for dimensionality reduction: W. Xiong et al. *Applied Intelligence*, *54*(19), 8939-8967.

Xu, L., Wen, S., Huang, H., Tang, Y., Wang, Y., & Pan, C. (2025). Corrosion failure prediction in natural gas pipelines using an interpretable XGBoost model. *Insights and applications. Energy*, *325*, *136157*.

Yan, L., Diao, Y., Lang, Z., & Gao, K. (2020). Corrosion rate prediction and influencing factors evaluation of low-alloy steels in marine atmosphere using machine learning approach. *Science and technology of Advanced Materials*, *21*(1), 359-370.

Zhi, Y., Fu, D., Zhang, D., Yang, T., & Li, X. (2019). Prediction and knowledge mining of outdoor atmospheric corrosion rates of low alloy steels based on the random forests approach. *Metals*, *9*(3), 383.