Results in Chemistry 6 (2023) 101126

Contents lists available at ScienceDirect

Results in Chemistry

journal homepage: www.sciencedirect.com/journal/results-in-chemistry

# Machine learning investigation to predict corrosion inhibition capacity of new amino acid compounds as corrosion inhibitors

Muhamad Akrom [a,b,*], Supriadi Rustad [b,*], Hermawan Kresno Dipojono [c,*]

[a] Study Program in Informatics Engineering, Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang 50131, Indonesia
[b] Research Center for Materials Informatics, Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang 50131, Indonesia
[c] Advanced Functional Materials Research Group, Institut Teknologi Bandung, Bandung 40132, Indonesia

ARTICLE INFO

*Keywords:*
Machine learning
QSPR
Corrosion inhibitor
Amino acid

ABSTRACT

This scientific paper aims to investigate the best machine learning (ML) for predicting the corrosion inhibition efficiency (CIE) value of amino acid compounds. The study applied a quantitative structure–property relationship (QSPR) model based on an ML approach to predict the CIE values of three new amino acid compounds, namely L-asparagine (LA), L-isoleucine (LI), and L-proline (LP). The result is that the Gradient Boosting Regressor (GBR) model is proven to be the best predictive model based on the coefficient of determination (R2) and root mean square error (RMSE) metrics used. The study found that the three amino acid compounds LA, LI, and LP tested had high CIE values, ranging from 90.49% to 93.67%. These results are also relevant to the CIE values resulting from experimental studies and show a trend that is by the adsorption energy trend. This engineering breakthrough can be used to predict the corrosion inhibition properties of new compounds before experimental synthesis.

## Introduction

Corrosion inhibitors based on organic compounds have been extensively investigated to protect against corrosion attacks on metals. They have received a lot of attention due to their high performance of corrosion inhibition, non-toxicity, environmental friendliness, affordability in cost, and ease of manufacture [1–3]. Corrosion-inhibiting properties are often found in organic compounds with heteroatoms (e.g., N, O, S, and P) and aromatic rings [4–6]. Amino acids are organic compounds that in their structure contain amino groups and carboxyl groups, and are environmentally friendly, therefore they have the potential to be promising corrosion inhibitors. Various types of amino acids have been tested experimentally for the corrosion resistance of steel. They showed an increase in the corrosion inhibition efficiency (CIE) as the inhibitor concentration increased [7–10]. In another study, Pour-Ali et al. [11] investigated 3 amino acid compounds, namely L-Asparagine (LA), L-Isoleucine (LI), and L-Proline (LP) as corrosion inhibitors using experimental methods. The results obtained were CIE values for the three compounds of 95, 93, and 91% respectively. But experimental testing is expensive, time-consuming, and resource-intensive [12–14].

To address this gap, recently, materials informatics has become popular because of the rapid development of technology and data science through machine learning (ML), especially in the design and development of new materials [15–17]. The ML approach based on the quantitative structure–property relationship model (QSPR) is often used to evaluate the performance of compounds because molecular attributes can be measured and related to the chemical structure of compounds [18–21]. QSPR is a fast, reliable, and affordable method in the field of materials informatics [22,23].

Khaled et al. [24] implemented a quantitative structure–activity relationship (QSAR) model to predict the performance of 28 amino acid compounds as corrosion inhibitors using quantum chemical properties (QCP) such as Hammett constants, dipole moment, HOMO energy, LUMO energy, energy gap, molecular area and volume as features input. His research shows that the artificial neural network (ANN) model predicts CIE values with a model performance metric of coefficient of determination ($R^2$) = 0.999. Using 19 amino acid compounds as samples, Zhao et al. [25] implemented the QASR model to correlate CIE with QCP such as HOMO energy, LUMO energy, energy gap, and dipole moment. The results show that the support vector machine (SVM) model used produces a prediction performance of $R^2$ = 0.97 and root mean square error (RMSE) = 1.48. Since both of the above studies used the same compounds, it is therefore hypothesized that they should be

* Corresponding authors.
*E-mail addresses:* m.akrom@dsn.dinus.ac.id (M. Akrom), srustad@dsn.dinus.ac.id (S. Rustad), dipojono@itb.ac.id (H. Kresno Dipojono).

https://doi.org/10.1016/j.rechem.2023.101126
Received 25 August 2023; Accepted 16 September 2023
Available online 18 September 2023

**Table 1**
Amino acid datasets.

| Dataset | Code | Number of compounds |
|---|---|---|
| Amino acid_1 [24] | AA1 | 28 |
| Amino acid_2 [25] | AA2 | 19 |
| Amino acid_C (dataset constructed) | AAC | 47 |

**Table 2**
Structure of novel amino acid molecules by RDKit.

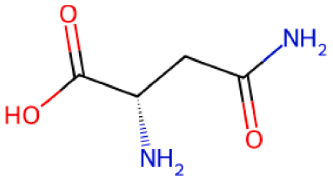
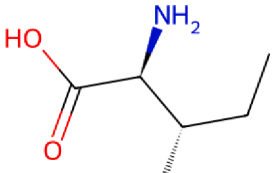
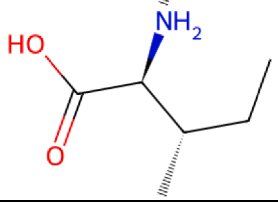
| Inhibitor | Molecule | Structure |
|---|---|---|
| LA | L-Asparagine ($C_4H_8N_2O_3$) |  |
| LI | L-Isoleucine ($C_6H_{13}NO_2$) |  |
| LP | L-Proline ($C_5H_9NO_2$) |  |

**Table 3**
QCP of novel amino acid molecules by DFT and Koopmans calculations by Pour-Ali et al. [11].

| Inhibitor | LUMO | HOMO | ΔE | χ | η | μ | ΔN |
|---|---|---|---|---|---|---|---|
| LA | −1.21 | −5.61 | 4.40 | 3.49 | 1.93 | 8.31 | 0.92 |
| LI | −1.14 | −5.83 | 4.68 | 3.41 | 2.20 | 2.84 | 0.82 |
| LP | −0.04 | −5.91 | 5.86 | 2.98 | 2.93 | 6.13 | 0.69 |

associated with the same best QSPR model fit for amino acids. Finding the best model with the highest level of prediction accuracy for amino acids is very important, it is also one of the challenges in developing ML approaches, especially when evaluating the possibility of new organic compounds acting as corrosion inhibitors.

In this paper, we construct a new dataset consisting of 47 amino acid compounds and focus on investigating the best QSPR model-based ML approach for predicting the CIE of amino acid compounds. The best model was used to predict the CIE of 3 new amino acid compounds, namely L-asparagine (LA), L-isoleucine (LI), and L-proline (LP). This work is critical in simplifying the manufacture and assessment of new anticorrosive materials prior to experimental studies.

## Method

### *Dataset of amino acid compounds*

From the available literature [24,25], we constructed a dataset containing 47 amino acid compounds. QCP such as HOMO and LUMO energies, energy gap (ΔE), ionization potential (I), electron affinity (A), global hardness (η), global softness (σ), electronegativity (χ), dipole moment (μ), and a fraction of electrons transferred (ΔN) were used as features and CIE as targets. QCP is usually calculated using the density functional theory (DFT) approach and Koopman theory [26]. QCP affects the corrosion inhibition capacities of anticorrosive agents [27,28]. Detailed information about the dataset used can be seen in Table 1. The structure of the 3 new amino acid compounds (LA, LP, and LI) are presented in Table 2. For these three molecules, we used QCP from the literature published by Pour-Ali et al. (Table 3) [11].

### *Machine learning model development*

We used two published datasets (AA1 and AA2) to assess the performance of the QSPR-based gradient boosting regressor (GBR), K-nearest neighbors (KNN), and support vector regression (SVR) models in predicting CIE values of amino acid compounds. In addition, validation was also carried out on the newly constructed dataset (AAC). The best model was then used to evaluate the corrosion inhibition capacity of the 3 new amino acid compounds (LA, LI, and LP).

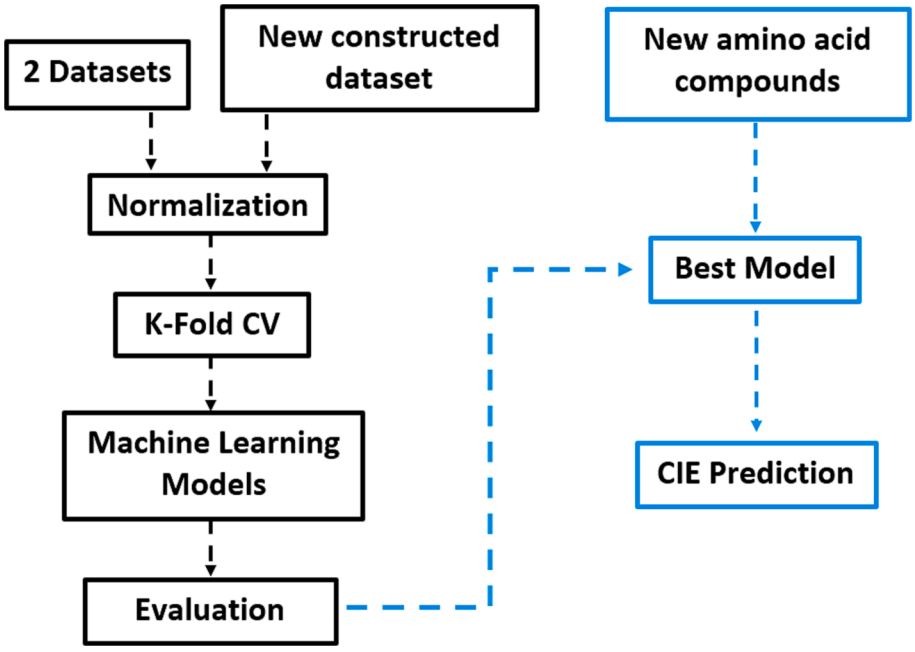



**Fig. 1.** Model development of QSPR-based ML.

**Table 4**
ML model performances for AA1 and AA2 datasets.

| Dataset | Model | $R^2$ | RMSE | MAE | MPE | Chi-square |
|---|---|---|---|---|---|---|
| AA1 | GBR | 0.999 | 1.50 | 0.51 | 1.63 | 0.96 |
| | KNN | 0.354 | 14.97 | 8.77 | 15.41 | 12.16 |
| | SVR | 0.132 | 17.36 | 13.35 | 18.68 | 15.43 |
| | Other work (ANN) [24] | 0.999 | na | na | na | na |
| AA2 | GBR | 0.999 | 0.24 | 0.39 | 1.08 | 0.89 |
| | KNN | 0.142 | 14.65 | 10.01 | 12.37 | 10.32 |
| | SVR | 0.011 | 15.74 | 11.79 | 13.91 | 10.18 |
| | Other work (SVM) [25] | 0.970 | 1.48 | na | na | na |

To avoid model sensitivity issues when working with oversized or inadequately sized data, a robust scaling approach is used as a data normalization tool. To reduce prediction errors, this preprocessing step is very important [29]. By using the k-fold cross-validation (k-fold CV) method, the stability of the ML method model is assessed by repeated training until the minimum statistical error value is reached [30]. To address variance and bias issues in ML, the data is divided into 5 sections, with one portion acting as the test pool and the other as the training pool in each training iteration. Fold k = 5 or 10 is often used, depending on the data set size [31]. An illustration of the development of the ML model can be seen in Fig. 1.

Model performance is measured using reliable metrics such as the coefficient of determination ($R^2$), root mean square error (RMSE), mean absolute error (MAE), model predictive error (MPE), and chi-square ($\chi^2$). The feasibility of the model is indicated by the value of $R^2$ close to 1. The difference between actual and predicted values is calculated using RMSE, MAE, MPE, and $\chi^2$. As these values decrease, the prediction error also decreases. A reduced statistical error indicates a better predictive model, so this statistic is used to evaluate the accuracy of the model [32–35].

$$R^2 = \frac{\sum_{i=1}^{n}(Y_i' - \overline{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \overline{Y}_i)^2} \quad (1)$$

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

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i' - Y_i| \quad (3)$$

$$MPE = \frac{100}{n}\sum_{i=1}^{n}\frac{Y_i - Y_i'}{Y_i'} \quad (4)$$

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

where $Y_i$, $\overline{Y}_i$, and $Y_i'$ represent the actual, mean of actual, and predicted values, respectively.

## Result and discussion

Table 4, and Figs. 2 and 3 show the performance of the prediction models (GBR, KNN, and SVR). In addition, Fig. 4 shows the CIE value

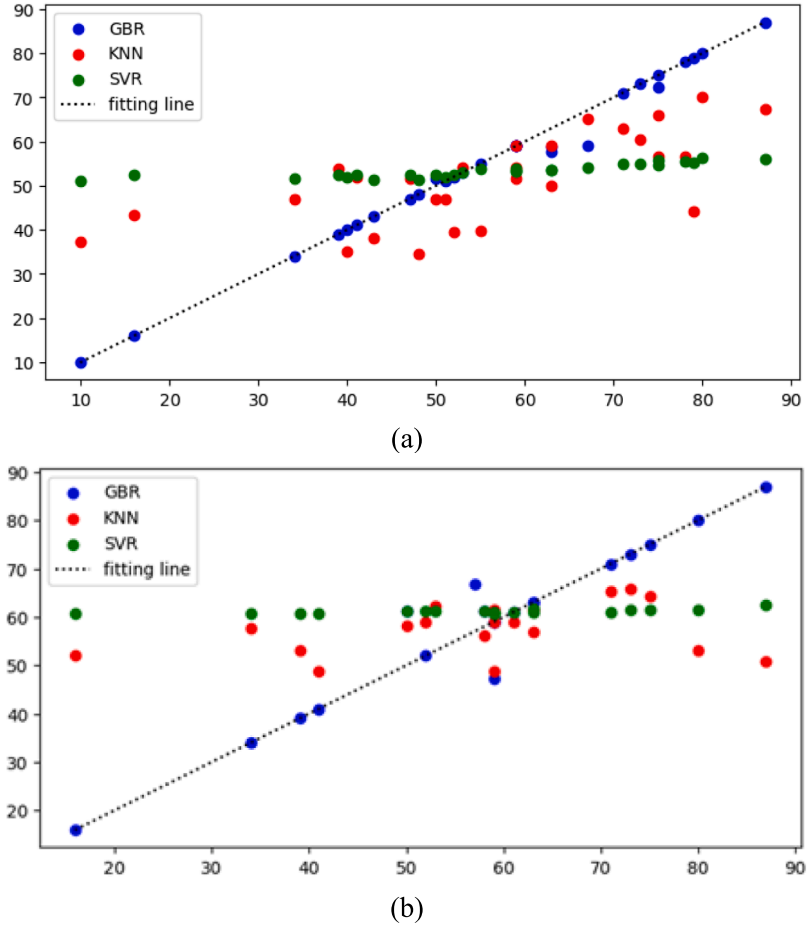



**Fig. 2.** Scatter plot of prediction data by models for (a) AA1 and (a) AA2 datasets.

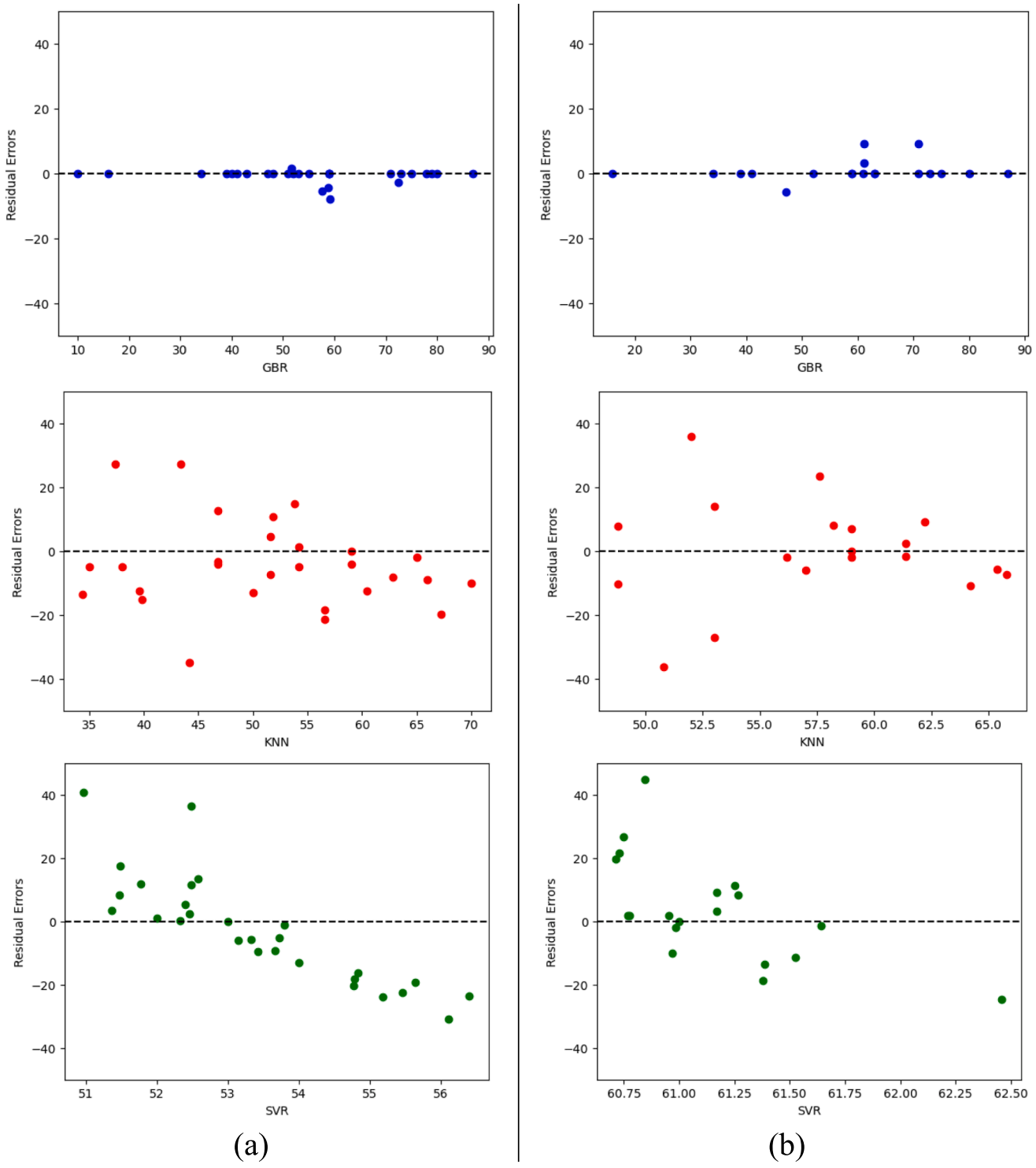



**Fig. 3.** Residual error plot of model predictions for (a) AA1 and (b) AA2 datasets.

predicted by the models against the experimental CIE value on the AA1 and AA2 datasets. As shown in Table 4, GBR shows the highest $R^2$ value and lowest RMSE, MAE, MPE, and Chi-square compared to the other two models on the AA1 and AA2 datasets. In addition, GBR is also superior to models from other works. These results are supported by Fig. 2 which shows that predictions at data points from GBR tend to be closest to the prediction line (fitting line). In addition, in the residual error plot in Fig. 3, the GBR residual data tends to be the smallest (close to the zero line). A model can accurately estimate the true value of the data tested if the residual error is low. These results show that GBR has the best predictive performance.

If GBR is the best model for each of the unaggregated datasets, it will be interesting to see whether it also shows predictive superiority over the reconstructed dataset (AAC). Table 5 compares the predictive performance of the three models on the AAC dataset. We find that GBR also shows its superiority as the best predictive model based on $R^2$ and RMSE values, also confirmed by the distribution of the predicted data points that are closest to the fitting line (Fig. 5(a)) and the CIE value pattern that is closest to the experimental value (Fig. 5(b)).

Since it proved to be superior, the GBR model was then used to assess

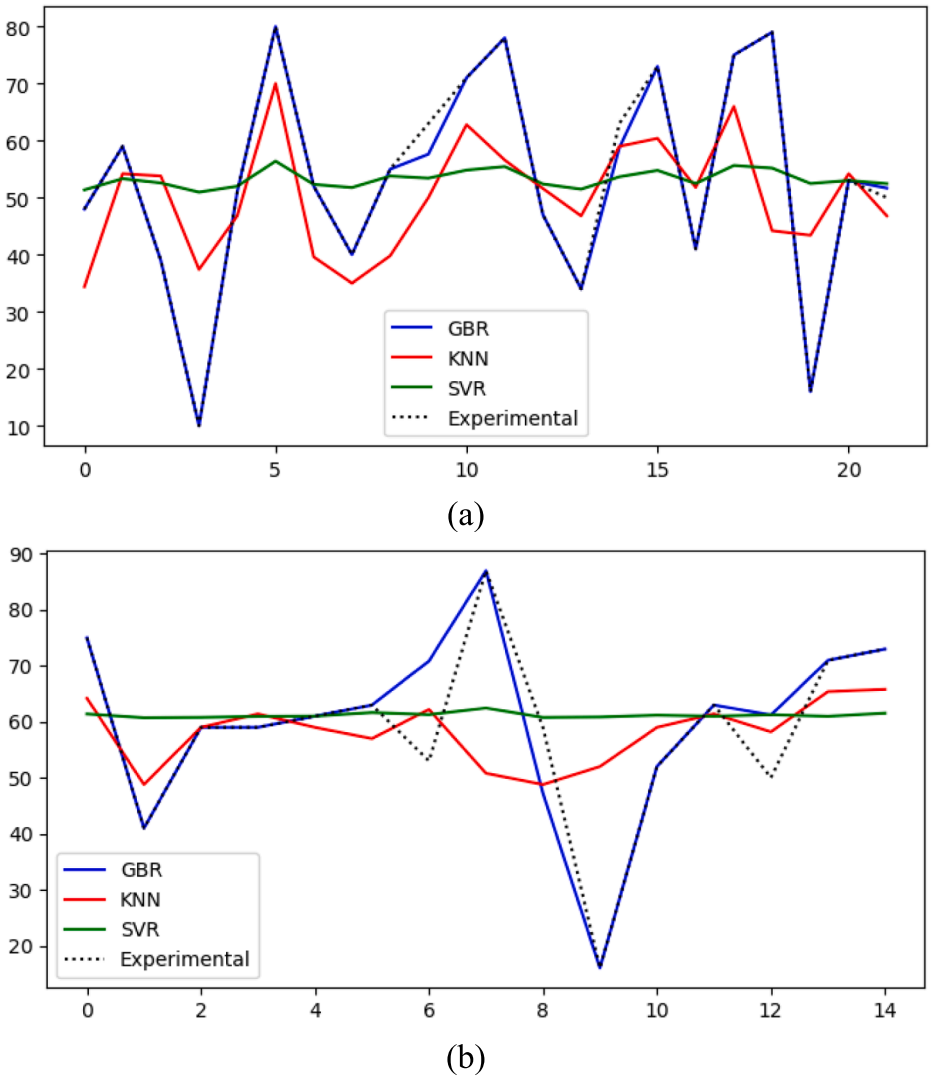



**Fig. 4.** The plot of CIE between model prediction and experimental values for (a) AA1 and (b) AA2 datasets.

**Table 5**
ML model performances for AAC dataset.

| Model | $R^2$ | RMSE |
|---|---|---|
| GBR | 0.999 | 0.59 |
| KNN | 0.437 | 12.32 |
| SVR | 0.051 | 15.82 |

three new amino acid compounds as potential corrosion inhibitor candidates. Table 6 shows that the estimated CIE of the new inhibitor compounds varies from 90.49 to 93.67% by GBR. The results indicated that this chemical would, in theory, function as a superior corrosion inhibitor. These results agree with experimental investigations by Pour-Ali et al. [11].

Since corrosion inhibition is related to the interaction of the inhibitor

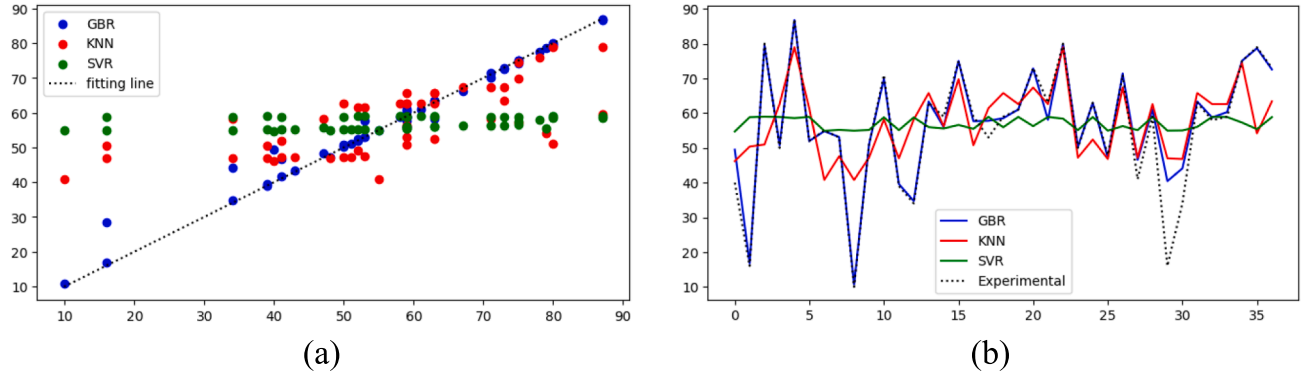



**Fig. 5.** (a) Scatter plot of data point prediction and (b) comparison of CIE between model prediction and experimental values.

**Table 6**
The CIE of novel amino acid molecules by GBR.

| Inhibitor | CIE (%) by GBR | CIE (%) by Experimental [11] | Eads (eV) by DFT [11] |
|---|---|---|---|
| LA | 93.67 | 95 | −2.65 |
| LI | 91.85 | 93 | −1.69 |
| LP | 90.49 | 91 | −1.30 |

with the metal surface, DFT calculations are used to assess the adsorption energy of amino acid inhibitor molecules on the iron surface. The behavior of inhibitor molecules adsorbed on metal surfaces influences how strongly they can prevent corrosion. The adsorption behavior of inhibitors is related to their ability to inhibit corrosion on metal surfaces. The adsorption energies (Eads) for the three molecules have also been calculated by Pour-Ali et al. [11] with molecular dynamics simulations to assess the interactions between molecules and the iron surface. The inhibitor molecules have significant adsorption energies ranging from −2.65 to −1.30 eV as seen in Table 6. It is important to note that the CIE trend and the adsorption energy trend are correlated, confirming the notion that the ability of the inhibitor molecule to interact with the metal surface through the formation of an adsorbed layer influences corrosion inhibition. Because the corrosion reaction on the metal surface can be inhibited in the presence of inhibitor compounds, its bonding behavior on the metal surface can protect the metal from corrosive attack [36,37]. To limit charge and mass transfer and protect the metal from corrosive conditions, molecular inhibitors must be able to form an adsorbed layer on the metal surface [38,39]. The presence of N and O heteroatoms and aromatic rings in the structure of the 3 molecules (Table 2) can also support the interaction between the molecule and the metal surface through the electron donor–acceptor process. The lone electron pairs of the heteroatom groups of the molecule and the empty d orbitals of the metal surface atoms, as well as the interactions between the electrons of the aromatic molecular ring and the empty d-orbitals of the metal surface atoms, can act as donors and acceptors, to facilitate the formation of adsorbed layers [40,41].

## Conclusion

In this work, a QSPR-based ML prediction model was developed to investigate the corrosion inhibition ability of amino acid compounds. The GBR model is the best predictor based on the $R^2$ and RMSE metrics. The accuracy of the GBR model was also seen when predicting the CIE values of the 3 new amino acid compounds (L-asparagine (LA), L-proline (LP), and L-isoleucine (LI)) tested as corrosion inhibitors, where the values agreed with experimental values. Another important finding is the trend of corrosion inhibition of the tested molecules in accordance with the trend of their adsorption energy. Future research will benefit from the complementary methodology of this study, especially when updating data on the ability of amino acid compounds to suppress corrosion. DFT simulations can help update the data set when new compounds are discovered by ML. Apart from amino acid derivatives, this strategy has the potential to be used for other types of molecules in exploring new corrosion inhibitors effectively and efficiently.

## CRediT authorship contribution statement

**Muhamad Akrom:** Conceptualization, Investigation, Software, Visualization, Writing - original draft, Writing - review & editing. **Supriadi Rustad:** Supervision, Conceptualization, Methodology. **Hermawan Kresno Dipojono:** Supervision, Methodology.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgments

All calculations were performed using the Computation Facility at the Research Center for Materials Informatics, Universitas Dian Nuswantoro.

## References

[1] Y. Cui, T. Zhang, F. Wang, New understanding on the mechanism of organic inhibitors for magnesium alloy, Corros. Sci. 198 (2022), 110118, https://doi.org/10.1016/J.CORSCI.2022.110118.
[2] H. Jin, D.J. Blackwood, Y. Wang, M.-F. Ng, T.L. Tan, First-principles study of surface orientation dependent corrosion of BCC iron, Corros. Sci. 196 (2022) 110029.
[3] T.K. Sarkar, V. Saraswat, R.K. Mitra, I.B. Obot, M. Yadav, Mitigation of corrosion in petroleum oil well/tubing steel using pyrimidines as efficient corrosion inhibitor: Experimental and theoretical investigation, Mater. Today Commun. 26 (2021), 101862, https://doi.org/10.1016/J.MTCOMM.2020.101862.
[4] C. Verma, E.E. Ebenso, M.A. Quraishi, Alkaloids as green and environmental benign corrosion inhibitors: An overview, Internat. J. Corr. Scale Inhibit. 8 (3) (2019) 512–528, https://doi.org/10.17675/2305-6894-2019-8-3-3.
[5] D. Daouda, T. Douadi, D. Ghobrini, N. Lahouel, H. Hamani, Investigation of some phenolic-type antioxidants compounds extracted from biodiesel as green natural corrosion inhibitors; DFT and molecular dynamic simulation, comparative study. AIP Conference Proceedings, American Institute of Physics Inc., 2019 doi: 10.1063/1.5138584.
[6] D.K. Kozlica, A. Kokalj, 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. 182 (2021), 109082, https://doi.org/10.1016/J.CORSCI.2020.109082.
[7] B. El Ibrahimi, A. Jmiai, L. Bazzi, S. El Issami, Amino acids and their derivatives as corrosion inhibitors for metals and alloys, Arab. J. Chem. 13 (1) (2020) 740–771.
[8] G.L.F. Mendonça, S.N. Costa, V.N. Freire, P.N.S. Casciano, A.N. Correia, P. de Lima-Neto, Understanding the corrosion inhibition of carbon steel and copper in sulphuric acid medium by amino acids using electrochemical techniques allied to molecular modelling methods, Corros. Sci. 115 (2017) 41–55, https://doi.org/10.1016/j.corsci.2016.11.012.
[9] A. Imjjad, K. Abbiche, M.D. Mellaoui, A. Jmiai, N. El Baraka, A. Ait Taleb, I. Bazzi, S. El Issami, M. Hilali, R.B. Said, M. Hochlaf, Corrosion inhibition of mild steel by aminobenzoic acid isomers in hydrochloric acid solution: Efficiency and adsorption mechanisms, Appl. Surf. Sci. 576 (2022) 151780.
[10] M. Fawzi Nassar, T. Zedan Taban, R. Fadhel Obaid, M. Hatem Shadhar, H. Abdulkareem Almashhadani, M.M. Kadhim, P. Liu, Study to amino acid-based inhibitors as an effective anti-corrosion material, J. Mol. Liq. 360 (2022) 119449.
[11] S. Pour-Ali, R. Tavangar, S. Hejazi, Comprehensive assessment of some L-amino acids as eco-friendly corrosion inhibitors for mild steel in HCl: Insights from experimental and theoretical studies, J. Phys. Chem. Solid 181 (2023) 111550.
[12] M. Akrom, S. Rustad, A.G. Saputro, H.K. Dipojono, Data-driven investigation to model the corrosion inhibition efficiency of Pyrimidine-Pyrazole hybrid corrosion inhibitors, Comput. Theor. Chem. 1229 (Nov. 2023), 114307, https://doi.org/10.1016/J.COMPTC.2023.114307.
[13] M.E.A. Ben Seghier, D. Höche, M. Zheludkevich, Prediction of the internal corrosion rate for oil and gas pipeline: Implementation of ensemble learning techniques, J. Nat. Gas Sci. Eng. 99 (2022) Mar, https://doi.org/10.1016/j.jngse.2022.104425.
[14] A.H. Alamri, N. Alhazmi, Development of data driven machine learning models for the prediction and design of pyrimidine corrosion inhibitors, J. Saudi Chem. Soc. 26 (6) (2022) Nov, https://doi.org/10.1016/j.jscs.2022.101536.
[15] R.I.D. Putra, A.L. Maulana, A.G. Saputro, Study on building machine learning model to predict biodegradable-ready materials, in: AIP Conference Proceedings, American Institute of Physics Inc., 2019, https://doi.org/10.1063/1.5095351.
[16] A. Agrawal, A. Choudhary, Deep materials informatics: Applications of deep learning in materials science, MRS Commun. 9 (3) (2019) 779–792.
[17] S. Lim, S. Chi, Xgboost application on bridge management systems for proactive damage estimation, Adv. Eng. Inf. 41 (2019) 100922.
[18] 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. 23 (9) (2022), https://doi.org/10.3390/ijms23095086.
[19] T.W. Quadri, L.O. Olasunkanmi, O.E. Fayemi, E.D. Akpan, H.-S. Lee, H. Lgaz, C. Verma, L. Guo, S. Kaya, E.E. Ebenso, Multilayer perceptron neural network-based QSAR models for the assessment and prediction of corrosion inhibition performances of ionic liquids, Comput. Mater. Sci 214 (2022) 111753.
[20] D. Kumar, V. Jain, 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. 195 (2022) 109960.
[21] E.H. El Assiri, M. Driouch, J. Lazrak, Z. Bensouda, A. Elhaloui, M. Sfaira, T. Saffaj, M. Taleb, Development and validation of QSPR models for corrosion inhibition of

carbon steel by some pyridazine derivatives in acidic medium, Heliyon 6 (10) (2020) e05067.
[22] A.A. Toropov, A.P. Toropova, QSPR/QSAR: State-of-art,weirdness, the future, Molecules 25 (6) (2020), https://doi.org/10.3390/molecules25061292. MDPI AG.
[23] M.E. Belghiti, S. Echihi, A. Dafali, Y. Karzazi, M. Bakasse, H. Elalaoui-Elabdallaoui, L.O. Olasunkanmi, E.E. Ebenso, M. Tabyaoui, Computational simulation and statistical analysis on the relationship between corrosion inhibition efficiency and molecular structure of some hydrazine derivatives in phosphoric acid on mild steel surface, Appl. Surf. Sci. 491 (2019) 707–722.
[24] K. F. Khaled, W. E. Org, and N. A. Al-Mobarak, "A Predictive Model for Corrosion Inhibition of Mild Steel by Thiophene and Its Derivatives Using Artificial Neural Network Characterization and Corrosion Protection Properties of Imidazole Derivatives on Mild Steel in 1.0 M HCl View project ELECTROCHEMICAL SCIENCE A Predictive Model for Corrosion Inhibition of Mild Steel by Thiophene and Its Derivatives Using Artificial Neural Network," 2012. [Online]. Available: https://www.researchgate.net/publication/236141845.
[25] H. Zhao, X. Zhang, L. Ji, H. Hu, Q. Li, Quantitative structure–activity relationship model for amino acids as corrosion inhibitors based on the support vector machine and molecular design, Corros. Sci. 83 (2014) 261–271, https://doi.org/10.1016/J.CORSCI.2014.02.023.
[26] C.T. Ser, P. Žuvela, M.W. Wong, Prediction of corrosion inhibition efficiency of pyridines and quinolines on an iron surface using machine learning-powered quantitative structure-property relationships, Appl. Surf. Sci. 512 (2020) 145612.
[27] T. Sutojo, S. Rustad, M. Akrom, A. Syukur, G.F. Shidik, H.K. Dipojono, A machine learning approach for corrosion small datasets, npj Mater. Degrad. 7 (1) (2023) Dec, https://doi.org/10.1038/s41529-023-00336-7.
[28] R.L. Camacho-Mendoza, L. Feria, L.Á. Zárate-Hernández, J.G. Alvarado-Rodríguez, J. Cruz-Borbolla, New QSPR model for prediction of corrosion inhibition using conceptual density functional theory, J. Mol. Model. 28 (8) (2022) Aug, https://doi.org/10.1007/s00894-022-05240-6.
[29] M. Ahsan, M. Mahmud, P. Saha, K. Gupta, Z. Siddique, Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance, Technologies (Basel) 9 (3) (2021) 52, https://doi.org/10.3390/technologies9030052.
[30] A. Botchkarev, A new typology design of performance metrics to measure errors in machine learning regression algorithms, Interdiscip. J. Inf. Knowl. Manag. 14 (2019) 45–76, https://doi.org/10.28945/4184.
[31] X. Yuan, Z. Ge, Z. Song, Soft sensor model development in multiphase/multimode processes based on Gaussian mixture regression, Chemom. Intel. Lab. Syst. 138 (Nov. 2014) 97–109, https://doi.org/10.1016/j.chemolab.2014.07.013.
[32] V.C. Anadebe, V.I. Chukwuike, S. Ramanathan, R.C. Barik, Cerium-based metal organic framework (Ce-MOF) as corrosion inhibitor for API 5L X65 steel in CO2-saturated brine solution: XPS, DFT/MD-simulation, and machine learning model prediction, Process Saf. Environ. Prot. 168 (2022) 499–512, https://doi.org/10.1016/J.PSEP.2022.10.016.
[33] V.C. Anadebe, P.C. Nnaji, O.D. Onukwuli, N.A. Okafor, F.E. Abeng, V. I. Chukwuike, C.C. Okoye, I.I. Udoh, M.A. Chidiebere, L. Guo, R.C. Barik, Multidimensional insight into the corrosion inhibition of salbutamol drug molecule on mild steel in oilfield acidizing fluid: Experimental and computer aided modeling approach, J. Mol. Liq. 349 (2022) 118482.
[34] S. Bafandeh, I. And, and M. Bolandraftar, "Application of K-nearest neighbor (KNN) approach for predicting economic events theoretical background Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background." [Online]. Available: www.ijera.com.
[35] P.D. Pately, M.R. Pately, N. Kaushik-Basu, T.T. Talele, 3D QSAR and molecular docking studies of benzimidazole derivatives as hepatitis C virus NS5B polymerase inhibitors, J. Chem. Inf. Model. 48 (1) (2008) 42–55, https://doi.org/10.1021/ci700266z.
[36] M. Akrom, S. Rustad, A.G. Saputro, A. Ramelan, F. Fathurrahman, 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. 35 (2023), 106402, https://doi.org/10.1016/J.MTCOMM.2023.106402.
[37] M. Akrom, A.G. Saputro, A.L. Maulana, A. Ramelan, A. Nuruddin, S. Rustad, H. K. Dipojono, DFT and microkinetic investigation of oxygen reduction reaction on corrosion inhibition mechanism of iron surface by Syzygium Aromaticum extract, Appl. Surf. Sci. 615 (2023) 156319.
[38] A. Dehghani, A.H. Mostafatabar, G. Bahlakeh, B. Ramezanzadeh, A detailed study on the synergistic corrosion inhibition impact of the Quercetin molecules and trivalent europium salt on mild steel; electrochemical/surface studies, DFT modeling, and MC/MD computer simulation, J. Mol. Liq. 316 (2020) 113914.
[39] A. Thakur, S. Kaya, A.S. Abousalem, A. Kumar, Experimental, DFT and MC simulation analysis of Vicia Sativa weed aerial extract as sustainable and eco-benign corrosion inhibitor for mild steel in acidic environment, Sustain. Chem. Pharm. 29 (2022) 100785.
[40] S. Kamal, M. Khalid, M. Shahnawaz Khan, M. Shahid, M.o. Ashafaq, M. I., M. Shahwaz Ahmad, M. Ahmad, M. Faizan, S. Ahmad, Synthesis, characterization and DFT studies of water stable Cd(II) metal–organic clusters with better adsorption property towards the organic pollutant in waste water, Inorganica Chim Acta 512 (2020) 119872.
[41] E. Ech-chihbi, et al., Computational, MD simulation, SEM/EDX and experimental studies for understanding adsorption of benzimidazole derivatives as corrosion inhibitors in 1.0 M HCl solution, J. Alloy. Compd. 844 (Dec. 2020), https://doi.org/10.1016/j.jallcom.2020.155842.