

Investigasi Efisiensi Inhibisi Korosi Senyawa Turunan Benzimidazole Berbasis *Ensemble Learning*

Novianto Nur Hidayat¹, Cornellius Adryan Putra Sumarjono², Muhamad Akrom^{3*}, Gustina Alfa Trisnapradika⁴

^{1,2,3,4}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
Jl. Imam Bonjol No. 207, Semarang, Indonesia 50131

*email: m.akrom@dsn.dinus.ac.id

(Naskah masuk: 15 September 2023; diterima untuk diterbitkan: 19 Januari 2024)

ABSTRAK – Di sektor material, korosi merupakan masalah besar yang mengakibatkan kerugian finansial tahunan. Salah satu taktik utama untuk mencegah kerusakan material karena lingkungan yang korosif adalah penghambatan korosi. Dalam hal ini, bahan kimia yang berasal dari benzimidazol telah dikenal sebagai agen yang memungkinkan untuk mencegah korosi. Namun, pemahaman menyeluruh tentang sifat korosi (yang rumit dan sulit diprediksi secara langsung) pada material diperlukan untuk menilai kemanjuran penghambatan senyawa ini. Untuk mengatasi masalah tersebut, penelitian ini memprediksi seberapa baik bahan kimia turunan benzimidazol dalam menekan korosi pada permukaan logam menggunakan teknik mesin learning (ML). Karena pembelajaran mesin dapat memproses dan menganalisis data yang kompleks dan membuat model prediksi berdasarkan pola yang ditemukan, hal ini dianggap sebagai jawaban yang memungkinkan. Langkah-langkah dalam metode ML yang diusulkan adalah pengumpulan data, pengolahan data, pemilihan algoritma ML, evaluasi model, dan pemilihan model terbaik. Temuan penelitian ini menunjukkan bahwa Gradient Boosting Regressor (GBR), khususnya, dapat menjadi alat yang berguna untuk memperkirakan seberapa baik senyawa yang berasal dari benzimidazol akan berfungsi sebagai penghambat korosi. Kontribusi utama dari penelitian ini adalah untuk menawarkan sudut pandang baru mengenai penerapan pembelajaran mesin untuk mengantisipasi inhibitor korosi secara akurat. Hal ini pada akhirnya dapat menghasilkan terciptanya material dengan ketahanan korosi yang meningkat. Hal ini menunjukkan kemampuan pembelajaran mesin untuk mengatasi masalah rumit di bidang kimia material.

Kata Kunci – Machine learning; Inhibisi korosi; Benzimidazole.

Investigation of Corrosion Inhibition Efficiency of Benzimidazole Derivative Compounds Based on Ensemble Learning

ABSTRACT – In the materials sector, corrosion represents a significant challenge, leading to annual financial losses. One of the primary strategies employed to mitigate material damage in corrosive environments is corrosion inhibition. In this context, compounds derived from benzimidazole have been identified as potential agents for corrosion prevention. Nevertheless, a comprehensive comprehension of the intricate and challenging-to-predict corrosion properties of materials is imperative to assess the inhibitory efficacy of these compounds. Addressing this challenge, the present research endeavors to predict the corrosion suppression capabilities of benzimidazole derivative chemicals on metal surfaces through the application of machine learning (ML) techniques. Given its capacity to process and analyze intricate data, along with its ability to formulate predictive models based on identified patterns, machine learning emerges as a promising solution. The proposed ML methodology involves sequential steps, namely data collection, data processing, ML algorithm selection, model evaluation, and the identification of the most effective model. The outcomes of this study propose that the Gradient Boosting Regressor (GBR), in particular, may serve as a valuable tool for estimating the effectiveness of compounds derived from benzimidazoles as corrosion inhibitors. The primary contribution of this research lies in providing a novel perspective on the utilization of machine learning to accurately forecast corrosion inhibitors, with potential implications for the development of materials exhibiting heightened corrosion resistance. This underscores the capability of machine learning to address intricate challenges within the realm of materials chemistry.

Keywords – Machine learning; Corrosion inhibition; Benzimidazole.

1. PENDAHULUAN

Proses korosi merupakan penguraian bahan yang disebabkan oleh reaksi kimia logam pada lingkungan yang banyak mengandung zat korosif [1]. Dalam proses ini, bahan korosif seperti oksigen di udara mengoksidasi logam untuk menghasilkan oksida, hidroksida, atau garam logam sebagai produk sampingan dari proses korosi. Reaksi korosif ini mempunyai beberapa dampak negatif, seperti menurunkan kualitas dan kinerja material, memperpendek masa pakainya, dan menimbulkan kerugian finansial yang besar [2], [3]. Jenis logam yang terlibat, kondisi lingkungan korosif (seperti kelembaban, pH, suhu, dan konsentrasi zat korosif), dan aspek lain seperti tekanan mekanis atau gesekan merupakan beberapa faktor yang mempengaruhi kecepatan korosi [4]. Interaksi dua logam berbeda dalam elektrolit, yang dikenal sebagai galvanik, aktivitas mikroorganisme seperti bakteri, atau korosi yang didorong oleh tegangan semuanya dapat mempercepat proses korosi [5, 6]. Memahami penyebab yang mendasari, menciptakan teknik untuk manajemen korosi, dan menilai kinerja material dalam lingkungan korosif semuanya termasuk dalam penelitian korosi [7]. Banyak industri, termasuk sektor kimia, minyak dan gas, otomotif, dan konstruksi, dapat memperoleh manfaat besar dari pengendalian korosi [8].

Senyawa kimia yang dikenal sebagai senyawa benzimidazol ($C_7H_6N_2$) terdiri dari cincin heterosiklik yang memiliki struktur inti yang sama dengan benzena (C_6H_6) dan imidazol ($C_3H_3N_2$) [9]. Banyak turunan benzimidazol telah dipelajari dan ditemukan memiliki rentang aktivitas molekuler [10]-[12]. Karakteristik ini menjadikan bahan kimia berguna dalam berbagai industri, seperti kimia bahan, agrokimia, dan obat-obatan. Selain digunakan dalam sintesis kimia, senyawa benzimidazol juga digunakan sebagai katalis, pengatur laju korosi, dan pigmen organik [13]. Namun demikian, uji coba eksperimental benzimidazol sebagai inhibitor korosi menghabiskan banyak uang, waktu, dan sumber daya [14], [15]. Kemajuan teknis saat ini menggabungkan pendekatan mekanika kuantum dengan aplikasi algoritmik untuk mempercepat proses desain dan mencari material baru melalui model ML. Produksi model prediktif, klasifikasi, dan pengelompokan, termasuk yang terkait dengan korosi, menjadi lebih mudah dengan disiplin kecerdasan buatan. Karena adanya hubungan (juga dikenal sebagai hubungan properti-struktur kuantitatif/quantitative structure-property relationship (QSPR)) antara karakteristik senyawa dan struktur molekul, teknik ML telah mendapatkan popularitas baru-baru ini dalam pencarian bahan-bahan baru [16].

Li et al. [17] menguji senyawa turunan

benzimidazole sebagai penghambat korosi menggunakan pendekatan ML berbasis model QSPR dengan algoritma SVM. Hasilnya adalah model menghasilkan kinerja prediksi sebesar $R^2 = 0.96$ dan $RMSE = 6.79\%$. Berdasarkan metrik evaluasi tersebut, model yang dikembangkan masih dapat ditingkatkan akurasi sebagai model prediktif. Tnatngan utama dalam pengembangan ML adalah pengembangan model yang akurat. Dalam karya ini, kami mempelajari cara memprediksi efisiensi penghambatan korosi (CIE) bahan kimia yang berasal dari turunan benzimidazol menggunakan ML berbasis model QSPR. Temuan penelitian ini menjelaskan pemodelan ML untuk menciptakan kemungkinan bahan kimia yang menghambat korosi. Dengan melakukan hal ini, material mungkin lebih terlindung dari kerusakan korosi [18]. Ini merupakan langkah menuju identifikasi bahan kimia penghambat korosi yang efektif dengan menggunakan metode yang lebih ekonomis dan efisien, sehingga mengurangi kebutuhan akan uji coba yang memakan waktu dan mahal.

2. METODE PENELITIAN

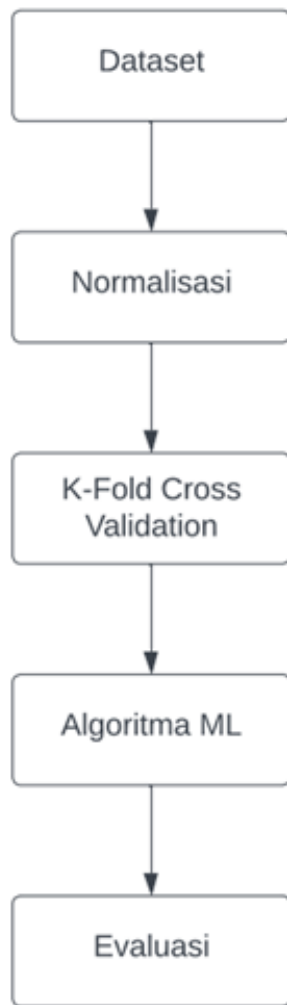
2.1. Dataset, fitur, dan target

Dataset yang digunakan dalam penelitian ini berasal dari literatur yang dipublikasikan dan mencakup rincian 20 senyawa inhibitor korosi berbasis senyawa turunan benzimidazole [17].

Dataset tersebut terdiri atas dua atribut utama, yaitu sebanyak 12 deskriptor kimia kuantum/*quantum chemical property* (QCP) dan 1 sifat penghambatan korosi/*corrosion inhibition efficiency* (CIE). QCP yang terdiri dari HOMO, LUMO, polarisasi (α), muatan total (Q), volume (V), potensial ionisasi (I), afinitas elektron (A), elektrofilitas (ω), fraksi elektron yang ditransfer (ΔN), benzena aromatik indeks ($\Delta NICS(1)B$), dan indeks aromatik imidazol ($\Delta NICS(1)I$) berfungsi sebagai fitur (variabel independen), dan CIE berfungsi sebagai target (variabel dependen) dalam pemodelan ML [19], [20]. Rangkuman fitur-fitur QCP dan target CIE yang digunakan pada [19] dan [20] dimuat di Tabel 1.

2.2. Model ML

Pada studi berbasis ML ini, kami menggunakan bahasa pemrograman Python yang dieksekusi menggunakan software Jupyter Notebook. Untuk mencari model optimum sebagai prediktor CIE senyawa turunan benzimidazol, kami menyelidiki berbagai algoritma ensemble dalam penelitian ini. *Bagging regressor* (BR), *extra trees* (ETR), *gradient boosting* (GBR), *random forest* (RF), *stacking* (SR), dan *voting* (VR) adalah beberapa algoritma ensemble yang diuji [21]-[26]. Langkah pemodelan diilustrasikan pada Gambar 1.



Gambar 1. Tahapan pemodelan ML

Tahap pra-pemrosesan data dilakukan untuk menghilangkan *noise* dan menskalakan data (normalisasi). Langkah ini untuk mengurangi sensitivitas terhadap fitur tertentu. Tahap berikutnya dilanjutkan dengan teknik *cross-validation* (CV) dengan model *k-fold* dengan menerapkan $k = 10$. Data dibagi menjadi k lipatan (*fold*), dimana satu *fold* digunakan sebagai set pengujian dan *fold* lainnya sebagai set pelatihan. Teknik ini untuk mengatasi bias dan varians pada data [27]–[31].

Metrik regresi seperti *mean squared error* (MSE), *root mean square error* (RMSE), *mean absolute error* (MAE), dan koefisien determinasi (R^2) digunakan untuk menilai efektivitas model prediksi yang dihasilkan pada tahap evaluasi model.

3. HASIL DAN PEMBAHASAN

Naskah Di antara model ML yang diuji yang disajikan pada Tabel 2, GBR menunjukkan performa prediksi terbaik bila diukur menggunakan metrik penilaian yang berbeda ($R^2 = 0.997$, RMSE = 0.549, MSE = 0.301, dan MAE = 0.427). Model dengan nilai

MSE, RMSE, dan MAE yang rendah dikombinasikan dengan nilai R^2 yang tinggi merupakan model yang kami anggap terbaik. Hal ini menunjukkan bahwa model tersebut cocok dengan data yang tersedia dan memiliki daya prediksi yang kuat. Sebaran titik data yang dihasilkan model GBR yang cenderung cukup dekat dengan garis prediksi yang digunakan (Gambar 2) mendukung kesimpulan tersebut.

Tabel 2. Performa model ML

Model	R^2	RMS E	MSE	MAE
GBR	0.997	0.549	0.301	0.427
RF	0.986	2.218	4.920	2.015
BR	0.971	2.262	5.118	2.178
ETR	0.969	2.282	5.209	2.231
SR	0.933	2.469	6.097	3.377
VR	0.926	2.475	6.124	3.406
SVR [20]	0.974	2.445	5.980	1.955

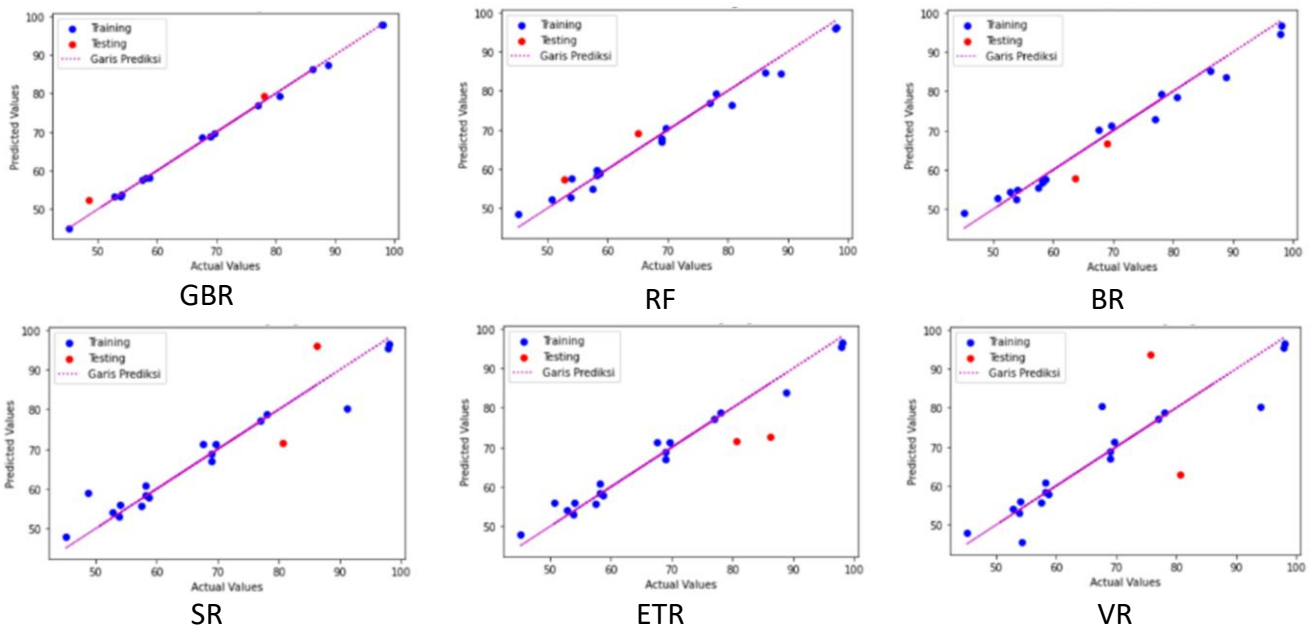
Secara umum, model GBR terbukti menjadi pilihan paling akurat untuk memperkirakan CIE senyawa yang berasal dari benzimidazol. Selain itu, terbukti bahwa hasil prediksi model GBR lebih baik dibandingkan penelitian sebelumnya pada dataset yang sama dengan menggunakan model Support Vector Regressor (SVR), yang memiliki RMSE sebesar 6,79 dan nilai R^2 sebesar 0,979 [17]. Hal ini menunjukkan bahwa untuk bahan kimia turunan benzimidazol, model GBR menghasilkan prediksi yang lebih tepat dan menyerupai nilai CIE sebenarnya.

Dari Gambar 3, terlihat bahwa GBR memiliki kinerja terbaik. Model-model lainnya memiliki performa yang lebih rendah dalam memprediksi data, dengan perbedaan kinerja yang cukup signifikan. Analisis ini memberikan pemahaman yang lebih baik tentang perbandingan relatif antara performa model-model tersebut berdasarkan nilai RMSE selama pelatihan untuk setiap *fold*-nya.

Deskriptor muatan total (Q) dan potensial ionisasi (I), masing-masing, merupakan faktor paling penting dalam menentukan hasil prediksi model, berdasarkan analisis fitur utama dapat dilihat Gambar 4.

Tabel 1. Fitur-fitur dan target yang dirujuk dari [19] dan [20]

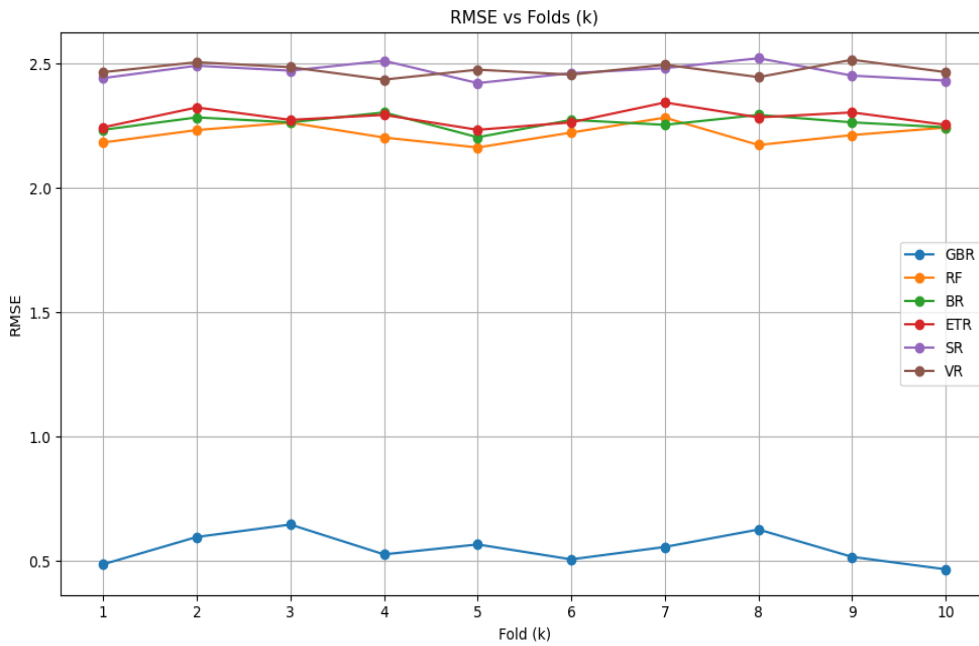
Inh#	EH	EL	α	Q	V	I	A	ω	ΔN	ΔN_{ICS} (1)B	ΔN_{ICS} (1)I	CIE
1	-7.56	-2.04	116.29	-1.74	63.83	7.38	2.40	4.79	-0.31	-11.53	-9.27	52.76
2	-8.99	-2.56	156.66	-2.16	81.80	7.39	4.10	10.03	-0.39	-11.22	-9.06	52.80
3	-7.27	-2.88	230.17	-2.74	106.97	7.04	3.22	6.90	-0.28	-11.21	-8.02	68.95
4	-7.33	-3.46	164.65	-3.11	82.53	7.10	3.82	9.11	-0.32	-10.35	-6.13	69.58
5	-6.86	-1.09	183.12	-3.74	96.68	5.53	2.13	4.31	0.06	-10.25	-5.89	97.99
6	-7.04	-1.05	188.23	-3.77	96.68	5.28	2.04	4.13	0.09	-9.82	-5.74	97.79
7	-8.28	-3.67	145.26	-2.20	75.72	7.94	4.02	9.12	-0.52	-10.97	-9.44	50.77
8	-7.64	-1.97	162.18	-2.34	84.60	7.09	3.10	6.51	-0.29	-11.41	-8.78	53.99
9	-8.29	-3.42	144.74	-2.25	75.72	7.88	3.79	8.32	-0.50	-11.22	-9.09	57.49
10	-7.40	-3.23	166.30	-3.15	82.53	7.16	3.61	8.17	-0.33	-10.62	-5.97	80.70
11	-9.07	-1.21	144.67	-2.98	77.45	6.58	1.52	3.25	0.00	-10.99	-6.08	88.75
12	-10.39	-2.18	137.58	-1.77	72.18	7.38	3.48	7.56	-0.37	-11.54	-8.44	45.06
13	-10.52	-1.53	122.71	-2.48	67.84	7.09	2.62	5.27	-0.25	-11.35	-7.11	58.19
14	-10.23	-1.79	146.76	-2.90	80.26	7.18	2.42	4.84	-0.26	-11.46	-8.67	68.97
15	-9.66	-1.25	148.44	-2.90	80.26	6.59	2.21	4.42	-0.11	-11.04	-6.60	78.03
16	-7.42	-1.82	188.12	-2.78	94.22	6.77	2.77	5.70	-0.20	-11.31	-7.89	67.55
17	-10.38	-1.26	131.91	-2.65	70.64	6.65	2.13	4.27	-0.11	-11.09	-6.25	76.93
18	-10.40	-1.93	152.21	-1.82	74.99	6.98	2.89	5.94	-0.25	-11.47	-7.91	58.13
19	-10.34	-1.90	140.56	-2.59	77.45	7.24	2.56	5.14	-0.29	-11.45	-8.69	58.76
20	-8.09	-1.91	169.20	-2.85	89.87	5.67	2.50	5.27	-0.01	-10.47	-9.14	86.18



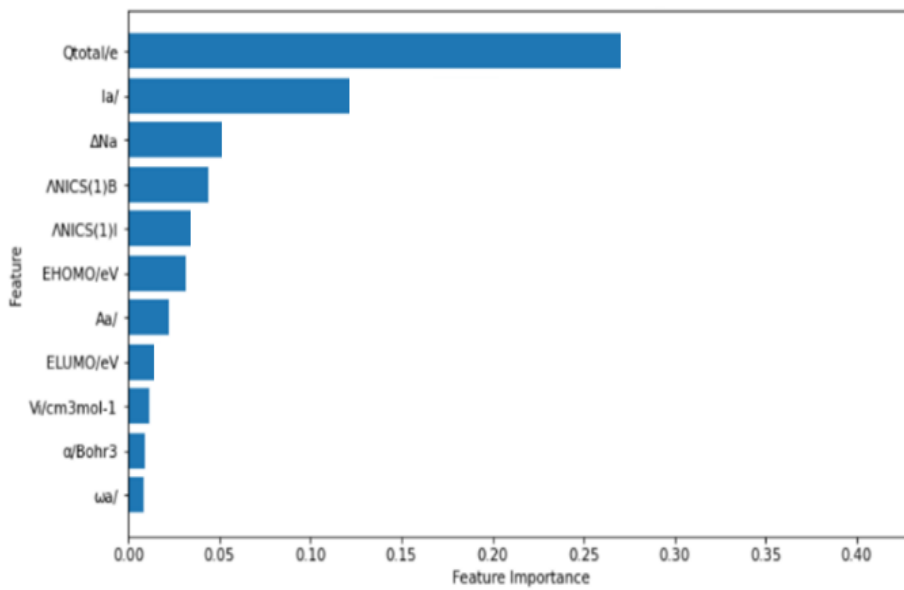
Gambar 2. Distribusi data poin

Kedua variabel ini mempunyai pengaruh yang kuat terhadap daya prediksi model, terbukti dengan korelasinya yang tinggi dengan tujuan prediksi CIE

[32]-[35]. Selain itu, terbukti bahwa fitur tambahan menunjukkan hubungan positif yang kuat dengan target CIE. Hal ini menunjukkan bahwa fitur-fitur ini dan nilai target yang akan diprediksi memiliki



Gambar 3. Plot RMSE dan k-fold untuk setiap model



Gambar 4. Analisis fitur penting

4. KESIMPULAN

hubungan yang kuat, yang membantu menjelaskan mengapa model GBR dapat memprediksi nilai dengan akurasi tinggi. Namun demikian, performa model ini untuk memprediksikan senyawa inhibitor korosi baru dalam kasus riil masih perlu diuji, yang bisa menjadi topik pembahasan dalam karya berikutnya.

Dengan membandingkan beberapa model ML, model optimal untuk menginvestigasi nilai CIE senyawa turunan benzimidazol menggunakan pendekatan ML telah diselidiki. Dengan nilai R^2 tertinggi = 0.997 dan nilai terendah RMSE = 0.549, MSE = 0.301, dan MAE = 0.427, maka ditetapkan bahwa model GBR merupakan model yang paling akurat jika dibandingkan dengan model lainnya yang diuji. Performa model ini untuk memprediksikan senyawa inhibitor korosi baru dalam kasus riil masih perlu diuji, yang bisa menjadi topik pembahasan dalam karya berikutnya. Selain

itu, dapat juga diusulkan efek penambahan fungsi polynomial terhadap peningkatan akurasi model. Namun demikian, hasil penelitian ini menawarkan wawasan penting dalam pengembangan teknik eksplorasi desain material penghambat korosi yang praktis dan efisien di industri.

DAFTAR PUSTAKA

- [1] M. Akrom, A. G. Saputro, A. L. Maulana, A. Ramelan, A. Nuruddin, S. Rustad, and H. K. Dipojono, "DFT and microkinetic investigation of oxygen reduction reaction on corrosion inhibition mechanism of iron surface by *Syzygium Aromaticum* extract," *Applied Surface Science*, vol. 615, p. 156319, 2023. [Online]. Available: <https://doi.org/10.1016/j.apsusc.2022.156319>.
- [2] M. Sugeng, F. M. Ismail, and J. P. Utomo, "Analysis of differences in corrosion rates from weight loss and polarization on pipe with ASTM G59 and ASTM G31 Standard Corrosion Testing," *Tera Journal*, vol. 2, no. 1, 2022.
- [3] S. Budi, M. Akrom, G. A. Trisnapradika, T. Sutojo, and W. A. E. Prabowo, "Optimization of Polynomial Functions on the NuSVR Algorithm Based on Machine Learning: Case Studies on Regression Datasets" *Scientific Journal of Informatics*, vol. 10, no. 2, 2023. [Online]. Available: <https://doi.org/10.15294/sji.v10i2.43929>.
- [4] M. Akrom, "DFT Investigation of *Syzygium Aromaticum* and *Nicotiana Tabacum* Extracts as Corrosion Inhibitor," *Science Tech: Jurnal Ilmu Pengetahuan dan Teknologi*, vol. 8, no. 1, p. 42, 2022. [Online]. Available: <https://doi.org/10.30738/st.vol8.no1.a11775>.
- [5] M. Akrom, "Investigation of Natural Extract as Green Corrosion Inhibitors In Steel Using Density Functional Theory," *Jurnal Teori dan Aplikasi Fisika*, vol. 10, no. 1, 2022.
- [6] A. A. Rosidah, V. A. Setyowati, and M. Choir, "Effect of Current and Coating Time on the Layer Thickness and Corrosion Rate of Electroplated AISI 1045," *SPECTA Journal of Technology*, vol. 5, no. 1, 2021.
- [7] R. Napitupulu, J. Daely, R. Manurung, and C. Manurung, "Effect of chrom electroplating time on low carbon steel on hardness, corrosion rate, and layer thickness," *Citra Science Technology*, vol. 1, no. 2, 2022. [Online]. Available: doi: 10.2421/cisat.v1i2.38.
- [8] A. Husodo, "Optimizing the painting process on ship decks to slow down corrosion," *JPB: Jurnal Patria Bahari*, vol. 3, no. 1, 2023. [Online]. Available: <https://doi.org/10.54017/jpb.v3i1.76>.
- [9] M. Akrom, " Experimental Investigation of Natural Plant Extracts as A Green Corrosion Inhibitor in Steel," *Journal of Renewable Energy and Mechanics*, vol. 5, no. 1, p 1-15, 2022. [Online]. Available: <https://doi.org/10.25299/rem.2022.vol5.no01.887>.
- [10] S. Maharani, N. Y. Haryono, and B.D. Mariana, "Analysis of the Effect of Concentration of Benomil Fungicide on Sterilization of Tawangmangu Citrus Plant Stem Tissue Culture," *Proceedings of Life and Applied Sciences*, vol. 1, 2022.
- [11] N. V. Putranto, M. Akrom, and G. A. Trisnapradika, "Implementasi Fungsi Polinomial pada Algoritma Gradient Boosting Regressor: Studi Regresi pada Dataset Obat-Obatan Kadaluarsa Sebagai Material Antikorosi," *Techno.com*, vol. 9, no. 2, p. 172, 2023.
- [12] I. Panduwiguna, I. Hardiana, S. A. Ogiuna, and M.S. Latief, "Gerd Disease Prescribing Patterns In The Inpatient Installation Of XYZ Hospital Jakarta," *KRYONAUT PHARMACY JOURNAL*, vol. 2, no. 1, 2023. [Online]. Available: <https://doi.org/10.59969/jfk.v2i1.21>.
- [13] M. Akrom, S. Rustad, and H. K. Dipojono, "Machine learning investigation to predict corrosion inhibition capacity of new amino acid compounds as corrosion inhibitors," *Results in Chemistry*, vol. 6, p. 101126, 2023. [Online]. Available: <https://doi.org/10.1016/j.rechem.2023.101126>.
- [14] M. A. Quraishi, D.S. Chauhan, and V.S. Saji, "Heterocyclic biomolecules as green corrosion inhibitors," *Journal of Molecular Liquids*, vol. 341, p. 117265, 2021. [Online]. Available: <https://doi.org/10.1016/j.molliq.2021.117265>.
- [15] D. Prasad, R. Singh, Z. Safi, N. Wazzan, and L. Guo, "De-scaling, experimental, DFT, and MD-simulation studies of unwanted growing plant as natural corrosion inhibitor for SS-410 in acid medium," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 649, p. 129333, 2022. [Online]. Available: <https://doi.org/10.1016/j.colsurfa.2022.129333>.
- [16] M. Akrom, U. Sudibyo, A. W. Kurniawan, N. A. Setiyanto, A. Pertiwi, A. N. Safitri, N. Hidayat, H. Al Azies, W. Herowati, "QSPR-Based Artificial Intelligence in the Study of Corrosion Inhibitors," *JoMMiT : Journal of Multi Media and IT*, vol. 7, no. 1, p. 015-020, 2023. [Online]. Available: <https://doi.org/10.46961/jommit.v7i1.721>.
- [17] L. Li, X. Zhang, S. Gong, H. Zhao, Y. Bai, Q. Li, and L. Ji, "The discussion of descriptors for the QSAR model and molecular dynamics

- simulation of benzimidazole derivatives as corrosion inhibitors," *Corrosion Science*, vol. 99, p. 76–88, 2015. [Online]. Available: <https://doi.org/10.1016/j.corsci.2015.06.003>.
- [18] M. Akrom and T. Sutojo, "Investigation of QSPR-Based Machine Learning Models in Pyrimidine Corrosion Inhibitors," *Eksergi*, vol. 20, no. 2, 2023, doi: 10.31315/e.v20i2.9864.
- [19] D. Leni, Y. P. Kusuma, R. Sumiati, Muchlisinalahuddin, and Adriansyah, "Comparison of Machine Learning Algorithms for Predicting Mechanical Properties of Low Alloy Steel," *Journal of Materials, Manufacturing, and Energy Engineering*, vol. 5, no. 2, 2022. [Online]. Available: <https://doi.org/10.30596/rmme.v5i2.11407>.
- [20] 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," *Computational and Theoretical Chemistry*, vol. 1229, p. 114307, 2023. [Online]. Available: <https://doi.org/10.1016/j.comptc.2023.114307>.
- [21] C. Beltran-Perez, A. A. A. Serrano, G. Solis-Rosas, A. Martinez-Jimenez, R. Orozco-Cruz, A. Esponzoza-Vazquez, and A. Miralrio, "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," *International Journal of Molecular Sciences*, vol. 23, no. 9, 2022. [Online]. Available: <https://doi.org/10.3390/ijms23095086>.
- [22] 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, 2022.
- [23] T. Sutojo, S. Rustad, M. Akrom, A. Syukur, GF Shidik, and HK Dipojono, "A machine learning approach for corrosion small datasets," *npj Materials Degradation*, vol. 7, no. 1, 2023. [Online]. Available: <https://doi.org/10.1038/s41529-023-00336-7>.
- [24] O. Tigga, J. Pal, and D. Mustafi, "A Comparative Study of Multiple Linear Regression and K Nearest Neighbours using Machine Learning," *Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT), IEEE*, 2023. [Online]. Available: <https://doi.org/10.1109/ICECCT56650.2023>.
- [25] D. Rahmawati, T. Kristanto, B.F.S. Pratama, and D.B. Abiansa, "Prediction of Foreign Travelers During the COVID-19 Pandemic Using Simple Linear Regression Methods," *Journal of Information Systems Research*, vol. 3, no. 3, 2022. [Online]. Available: <https://doi.org/10.47065/josh.v3i3.1507>.
- [26] H.D. Panduwinata, S. Suyitno, and M.N. Huda, "Weibull Regression Model on Classified Continuous Data," *Exponential*, vol. 13, no. 2, 2022.
- [27] T. Fushiki, "Estimation of prediction error by using K-fold cross-validation," *Statistics and Computing*, vol. 21, 2011.
- [28] S. Bates, T. Hastie, and R. Tibshirani, "Cross-Validation: What Does It Estimate and How Well Does It Do It?," *Journal of the American Statistical Association*, 2023. [Online]. Available: <https://doi.org/10.1080/01621459.2023.2197686>.
- [29] L. Leng, L. Yang, X. Lei, W. Zhang, Z. Ai, Z. Yang, H. Zhan, J. Yang, X. Yuan, H. Peng, and H. Li, "Machine learning predicting and engineering the yield, N content, and specific surface area of biochar derived from pyrolysis of biomass," *Biochar*, vol. 4, no. 1, p. 63, 2022. [Online]. Available: <https://doi.org/10.1007/s42773-022-00183-w>.
- [30] T. Zhu, S. Li, L. Li, and C. Tao, "A new perspective on predicting the reaction rate constants of hydrated electrons for organic contaminants: Exploring molecular structure characterization methods and ambient conditions," *Science of the Total Environment*, vol. 904, p. 166316, 2023. [Online]. Available: <https://doi.org/10.1016/j.scitotenv.2023.166316>.
- [31] 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," *Materials Today Communications*, vol. 35, p. 106402, 2023. [Online]. Available: <https://doi.org/10.1016/j.mtcomm.2023.106402>.
- [32] U. M. Khaire and R. Dhanalakshmi, "Stability of feature selection algorithms: A review," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 4, p. 1060–1073, 2022. [Online]. Available: <https://doi.org/10.1016/j.jksuci.2019.06.012>.
- [33] A. Bommert, X. Sun, B. Bischl, J. Ruhnenfurher, and M. Lang, "Benchmark of filter methods for feature selection in high-dimensional gene expression survival data," *Computational Statistics and Data Analysis*, 2020. [Online]. Available: <https://doi.org/10.1016/j.csda.2019.106839>.
- [34] F. M. Haikal, M. Akrom, and G. A. Trisnapradika, "Perbandingan Algoritma

- Multilinear Regression dan Decision Tree Regressor Dalam Memprediksi Efisiensi Penghambatan Korosi Piridazin," *Edumatic: Jurnal Pendidikan Informatika*, vol. 7, no. 2, 2023. [Online]. Available: <https://doi.org/10.29408/edumatic.v7i2.22127>.
- [35] M. Akrom, T. Sutojo, A. Pertiwi, S. Rustad, and H. K. Dipojono, " Investigation of Best QSPR-Based Machine Learning Model to Predict Corrosion Inhibition Performance of Pyridine-Quinoline Compounds," *Journal of Physics: Conference Series*, vol. 2673, no. 1, p. 012014, 2023. [Online]. Available: <https://doi.org/10.1088/1742-6596/2673/1/012014>.