Original Article

Check for updates

Available online at
www.heca-analitika.com/hjas

Heca Journal of Applied Sciences

Vol. 3, No. 1, 2025

Predicting AXL Tyrosine Kinase Inhibitor Potency Using Machine Learning with Interpretable Insights for Cancer Drug Discovery

Teuku Rizky Noviandy [1], Ghifari Maulana Idroes [2], Essy Harnelly [3], Irma Sari [4], Fazlin Mohd Fauzi [5] and Rinaldi Idroes [4,6,*]

[1] Department of Information Systems, Faculty of Engineering, Universitas Abulyatama, Aceh Besar 23372, Indonesia; rizky_si@abulyatama.ac.id (T.R.N.)
[2] Department of Nuclear Engineering and Engineering Physics, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia; ghifarimaulana145@gmail.com (G.M.I.)
[3] Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia; essy.harnelly@usk.ac.id (E.H.)
[4] Department of Pharmacy, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia; irmasari@usk.ac.id (I.S.); rinaldi.idroes@usk.ac.id (R.I.)
[5] Faculty of Pharmacy, Universiti Teknologi MARA Selangor, Puncak Alam Campus, 42 300 Bandar Puncak Alam, Selangor, Malaysia; fazlin5465@uitm.edu.my (F.M.F.)
[6] School of Mathematics and Applied Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia

* Correspondence: rinaldi.idroes@usk.ac.id

Article History

Received 7 January 2025
Revised 25 February 2025
Accepted 8 March 2025
Available Online 15 March 2025

Keywords:
AXL tyrosine kinase
Machine learning
Drug discovery
SHAP analysis
Cancer therapeutics

Abstract

AXL tyrosine kinase plays a critical role in cancer progression, metastasis, and therapy resistance, making it a promising target for therapeutic intervention. However, traditional drug discovery methods for developing AXL inhibitors are resource-intensive, time-consuming, and often fail to provide detailed insights into molecular determinants of potency. To address this gap, we applied machine learning techniques, including Random Forest, Gradient Boosting, Support Vector Regression, and Decision Tree models, to predict the potency (pIC_{50}) of AXL inhibitors using a dataset of 972 compounds with 550 molecular descriptors. Our results demonstrate that the Random Forest model outperformed others with an R^2 of 0.703, MAE of 0.553, RMSE of 0.720, and PCC of 0.841, showcasing strong predictive accuracy. SHAP analysis identified critical molecular features, such as RNCG and TopoPSA(NO), as key contributors to inhibitor potency, providing interpretable insights into structure-activity relationships. These findings highlight the potential of machine learning to accelerate the identification and optimization of AXL inhibitors, bridging the gap between computational predictions and rational drug design and paving the way for effective cancer therapeutics.

1. Introduction

AXL, a receptor tyrosine kinase, is critical in various physiological and pathological processes, including cell survival, proliferation, migration, and immune response [1]. Overexpression or dysregulation of AXL has been implicated in the progression of several cancers, including hepatocellular carcinoma [2], non-small cell lung cancer [3], and breast cancer [4]. It is critical in promoting metastasis, therapy resistance, and poor patient prognoses. As a result, AXL has emerged as a promising therapeutic target in cancer drug discovery, with inhibitors of AXL tyrosine kinase showing potential

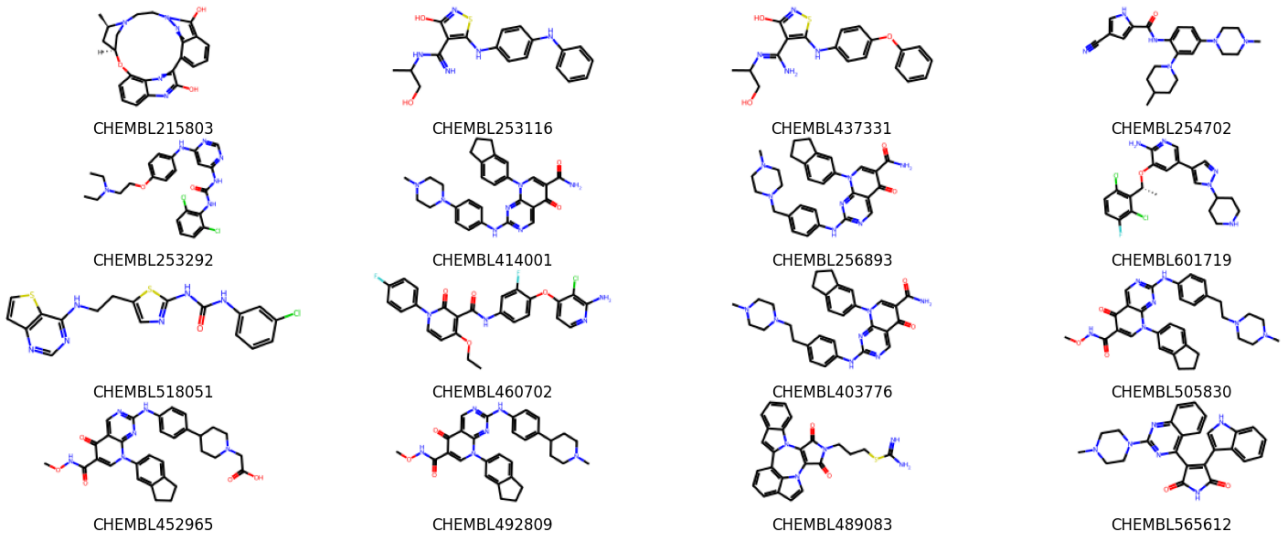

Figure 1. Representative chemical structures of AXL inhibitors from the ChEMBL database.

for mitigating disease progression and enhancing treatment efficacy [5].

Traditional approaches to drug discovery for targeting AXL rely on experimental high-throughput screening (HTS) and structure-based drug design [6, 7]. While these methods have contributed significantly to identifying potential AXL inhibitors, they are resource-intensive, time-consuming, and often yield limited insights into the underlying mechanisms determining inhibitor potency [8]. Moreover, these conventional methods fail to fully grasp the complex interactions between molecular features that affect drug effectiveness, making optimization and rational design more challenging [9].

In recent years, artificial intelligence (AI) and machine learning have revolutionized drug discovery by enabling data-driven approaches that can accelerate target identification, lead optimization, and drug design [10, 11]. Machine learning algorithms can uncover complex patterns and relationships within large datasets, offering predictive capabilities that surpass traditional methods [12]. This paradigm shift in computational biology has facilitated breakthroughs in identifying novel drug candidates with high precision and efficiency, even in challenging therapeutic targets like AXL.

Machine learning offers several advantages for drug discovery, including its ability to integrate diverse data sources, process high-dimensional datasets, and generalize predictive models for inhibitor potency [13, 14]. Moreover, the incorporation of interpretable machine learning techniques, such as SHAP (Shapley Additive Explanations), has mitigated the "black box" nature of many machine learning models [15, 16]. By providing explainable insights into the contribution of specific molecular features to model predictions, these techniques empower researchers to better understand the structural and chemical properties governing AXL inhibition, aiding rational drug design and discovery.

This study aims to use machine learning to predict the potency of AXL tyrosine kinase inhibitors while providing interpretable insights into the molecular determinants of their efficacy. With advanced machine learning techniques and interpretable frameworks, we seek to accelerate the identification and optimization of promising AXL inhibitors, ultimately contributing to developing effective cancer therapeutics.

2. Materials and Methods

2.1. Dataset Collection and Preparation

The dataset used in this study was obtained from the ChEMBL database [17], specifically targeting the tyrosine-protein kinase receptor (CHEMBL4895), commonly known as AXL. Initially, 1,113 compounds associated with AXL activity were retrieved. To ensure data quality and reliability, duplicate entries and compounds with missing values were removed, resulting in a final dataset of 972 compounds for analysis. A representative selection of these compounds is illustrated in Figure 1. The depicted compounds are known to interact with the AXL kinase and exhibit varying degrees of activity.

The dataset included two key attributes: SMILES and IC_{50} values. The SMILES (Simplified Molecular Input Line Entry System) representation is a textual format that encodes the structure of chemical compounds. It provides a compact and machine-readable description of molecules, enabling their use as inputs for computational modeling [18]. The IC_{50} (half-maximal inhibitory concentration) is a widely used metric in drug discovery that quantifies the

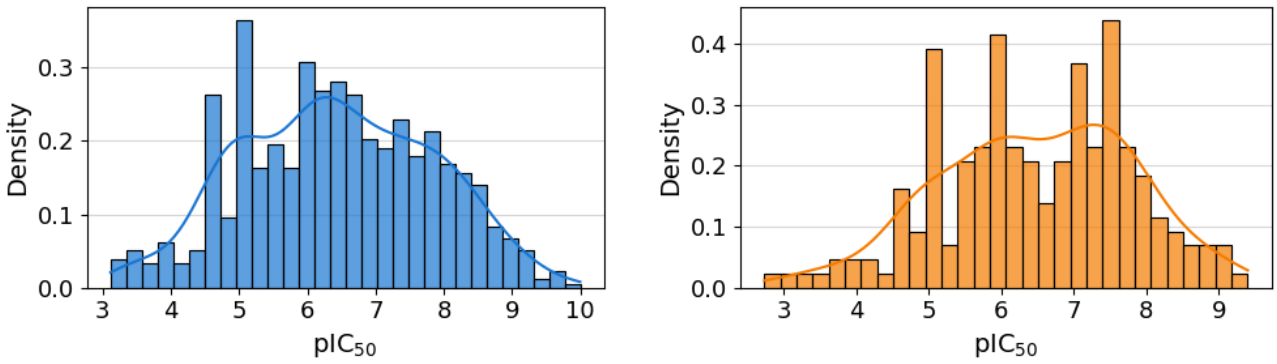

Figure 2. Distribution of pIC_{50} values for AXL inhibitors in (a) the training set and (b) the testing set, highlighting the range and variability in compound potency across both datasets.

concentration of a compound required to inhibit a specific biological or biochemical function by 50%. This value reflects the potency of a compound as an inhibitor.

To enhance the interpretability and stability of the IC_{50} data, the raw IC_{50} values (expressed in molar units) were transformed into pIC_{50} values [19]. This transformation converts the IC_{50} values into a logarithmic scale, where higher pIC_{50} values correspond to greater inhibitor potency. The resulting pIC_{50} values were used as the target variable in the machine-learning models to predict AXL inhibitor potency.

2.2. Molecular Descriptor Calculation

2D molecular descriptors were calculated to represent the compounds' chemical properties quantitatively. Molecular descriptors are numerical representations of molecules' physicochemical, structural, and topological characteristics [20]. They provide a way to capture complex molecular attributes such as size, shape, polarity, aromaticity, and the presence of functional groups, which are crucial for understanding and predicting biological activities like drug potency. These descriptors serve as features for machine learning models, enabling them to identify patterns and relationships between molecular structure and inhibitor activity.

In this study, molecular descriptors were computed using the Mordred software package, which is widely used for generating a comprehensive set of chemical features [21]. A total of 1,329 descriptors were initially calculated for each compound in the dataset. However, not all descriptors are equally informative, and some may introduce noise or redundancy into the model. A preprocessing pipeline was applied to refine the feature set to address this. Descriptors with zero variance across all compounds were removed, as they provided no distinguishing information [22]. Additionally, highly correlated descriptors with a correlation coefficient greater than 0.95 were excluded to reduce multicollinearity and improve the interpretability and generalizability of the models [23].

After preprocessing, a final set of 550 molecular descriptors was retained. These descriptors represent a diverse and meaningful selection of chemical features, providing the necessary inputs for machine-learning models to predict AXL inhibitor potency with accuracy and reliability.

2.3. Exploratory Data Analysis

The data was split into training and test sets to prepare the dataset for machine learning, with 80% of the compounds allocated for training and 20% for testing. This split ensures that the machine learning models are trained on most data while retaining an independent subset for evaluating model performance on unseen compounds [24]. By doing so, we can assess the generalization ability of the models and ensure their reliability in predicting AXL inhibitor potency.

The distribution of the pIC_{50} values in the dataset was visualized to understand the range and variability of compound potency (Figure 2). This step provides insight into the overall data distribution and helps identify any imbalances or patterns that may influence the model's performance.

A t-distributed Stochastic Neighbor Embedding (t-SNE) analysis was also performed on the training set. t-SNE is a dimensionality reduction technique to visualize high-dimensional data in a lower-dimensional [25]. The purpose of applying t-SNE was to explore the relationships and clustering of the molecular descriptors within the dataset. By projecting the high-dimensional descriptor space into a 2D plot, we can identify patterns, groupings, or outliers in the chemical space. This may provide further insights into the structure-activity relationships governing AXL inhibitor potency.

2.4. Machine Learning Models

To predict the potency of AXL tyrosine kinase inhibitors, we trained and evaluated four machine-learning regression models: Random Forest, Gradient Boosting, Support Vector Regression, and Decision Tree. These models were selected for their ability to handle diverse data structures and relationships and their complementary approaches to regression tasks. Random Forest and Gradient Boosting were chosen for their strong performance in handling high-dimensional data and capturing complex non-linear relationships. At the same time, Support Vector Regression was included for its ability to model intricate patterns using kernel functions. Decision Tree, though simpler, serves as a baseline for interpretability and comparison. The diversity of these methods provides a robust framework for identifying the best-performing model and understanding the predictive relationships between molecular descriptors and inhibitor potency. All models were implemented using Scikit-learn with default hyperparameters, and a fixed random state of 42 was used to ensure the reproducibility of the results.

Random Forest is an ensemble learning method that builds multiple decision trees during training and combines their outputs to make predictions [26]. It is particularly suited for high-dimensional datasets, such as those with molecular descriptors, due to its robustness to overfitting and ability to capture nonlinear patterns. By averaging predictions from multiple trees, Random Forest reduces variance and delivers stable results, making it a strong candidate for regression tasks [27].

Gradient Boosting takes a different approach by building models sequentially, where each new model focuses on correcting the errors of the previous ones. This iterative process minimizes a predefined loss function through gradient descent, allowing Gradient Boosting to effectively capture intricate patterns in the data. Its ability to model complex relationships makes it a popular choice for tasks involving structure-activity relationships, where subtle molecular features may strongly influence biological activity [28].

Support Vector Regression offers a mathematically rigorous approach by fitting a hyperplane in a high-dimensional feature space that best represents the data with some error tolerance. By leveraging kernel functions, such as the radial basis function (RBF), Support Vector Regression can model complex, nonlinear relationships. This flexibility makes it well-suited for capturing the nonlinearity often present in molecular structure-activity data, adding depth to the model comparison [29].

Decision Tree Regression, while simpler than the other models, provides interpretable predictions by partitioning the feature space into regions using decision rules. It captures nonlinearity through piecewise constant approximations and serves as a baseline for understanding the dataset. Although Decision Trees are prone to overfitting independently, they form the foundation for more sophisticated ensemble methods like Random Forest and Gradient Boosting, making them an essential component of this analysis [30].

By training these four models, we aimed to explore a variety of regression approaches, from interpretable and simple methods to more complex and flexible techniques. This comprehensive comparison allows us to identify the most effective model for predicting AXL inhibitor potency and to gain insights into the molecular features that govern activity.

2.5. Model Evaluation

To evaluate the performance of the regression models, we utilized four metrics: R^2 (coefficient of determination), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Pearson Correlation Coefficient (PCC). Each metric serves a specific purpose in assessing the predictive accuracy, error magnitude, and strength of the relationship between predicted and actual values, providing a comprehensive evaluation of the models.

The R^2 (coefficient of determination) measures the proportion of the variance in the target variable (pIC_{50}) that the model explains. It assesses how well the model captures the overall trends in the data. A higher R^2 value indicates better predictive performance, with 1 representing a perfect fit and 0 indicating no explanatory power [31].

The Mean Absolute Error (MAE) quantifies the average absolute difference between the predicted and actual values, making it a straightforward and interpretable measure of model precision. MAE is less sensitive to outliers than other metrics, providing a reliable sense of the average error magnitude [32].

The Root Mean Square Error (RMSE), on the other hand, calculates the square root of the average squared differences between predicted and actual values. By squaring the errors, RMSE penalizes larger deviations more heavily than MAE, making it especially useful when large errors are of greater concern. A lower RMSE reflects better model performance [33].

The Pearson Correlation Coefficient (PCC) evaluates the linear correlation between the predicted and actual values. It indicates the strength and direction of the

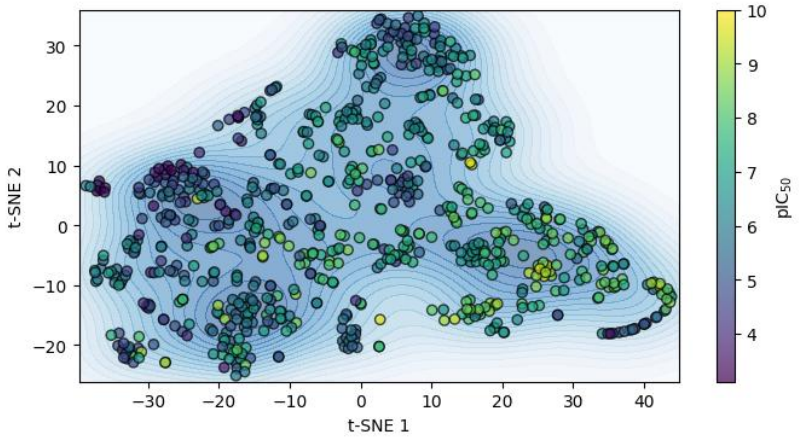

Figure 3. t-SNE visualization of molecular descriptors for AXL inhibitors, with points colored by pIC_{50} values. Higher pIC_{50} values (yellow) indicate greater potency, while lower values (purple) reflect weaker inhibition.

relationship, ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation), with 0 implying no linear correlation. PCC provides a complementary perspective to the error-based metrics by focusing on the relationship rather than the magnitude of errors [34].

2.6. Model Interpretation

To gain deeper insights into the predictions of the best-performing model, we employed SHAP (Shapley Additive Explanations), a widely used framework for interpreting machine learning models [35]. SHAP provides a unified approach to explain the contribution of each feature to the model's predictions by assigning Shapley values [36, 37]. These values are derived from cooperative game theory and quantify the impact of individual features on the predicted outcomes.

By using SHAP, we identified the most influential molecular descriptors driving the predictions of AXL inhibitor potency. This interpretability helps uncover structure-activity relationships, offering valuable guidance for rational drug design and optimization [38]. The feature importance analysis highlights key chemical properties and enables domain experts to validate the model's predictions and ensure alignment with prior knowledge in drug discovery.

Visualizations of SHAP values, such as summary and dependency plots, were used to interpret the model's behavior [39, 40]. These plots reveal how specific features influence the model's predictions across the dataset and allow for identifying non-linear or interaction effects. The interpretability provided by SHAP bridges the gap between model complexity and human understanding, making the machine-learning predictions more actionable for AXL tyrosine kinase inhibitor discovery.

3. Results and Discussion

3.1. t-SNE Visualization of Molecular Descriptor Space

We conducted a t-distributed Stochastic Neighbor Embedding (t-SNE) analysis on the molecular descriptor data to investigate the distribution and structure-activity relationships in the dataset. The t-SNE projection, illustrated in Figure 3, reduces the high-dimensional chemical feature space to two dimensions, clearly visualizing molecular clustering and distribution. Each point in the plot corresponds to a compound, with its color reflecting its pIC_{50} value (inhibitor potency), as indicated by the color bar.

The t-SNE visualization reveals distinct groupings of compounds based on their molecular features, suggesting potential structure-activity relationships governing AXL inhibitor potency. Compounds with higher pIC_{50} values (yellow) are relatively sparse but cluster together in specific regions of the chemical space. This clustering indicates that these potent compounds may share similar structural or physicochemical properties that contribute to their inhibitory activity. On the other hand, compounds with lower pIC_{50} values (purple) are distributed more broadly, suggesting greater structural diversity among less potent inhibitors.

The continuous gradient of colors across the t-SNE plot further highlights the nuanced relationships between molecular features and AXL inhibition potency. This observation suggests that certain molecular descriptors play a pivotal role in determining compound activity, with clear transitions in potency levels evident in the visualization. Outliers or isolated points in the plot may represent compounds with unique chemical properties or unexpected activity profiles, which could warrant further investigation to identify novel chemical scaffolds for AXL inhibition.

Table 1. Performance metrics for machine learning models on the test set, evaluating AXL inhibitor potency prediction.

Model	R2	MAE	RMSE	PCC
Random Forest	0.703	0.553	0.720	0.841
Gradient Boosting	0.687	0.585	0.740	0.829
Support Vector Regression	0.688	0.567	0.738	0.830
Decision Tree	0.357	0.757	1.060	0.685

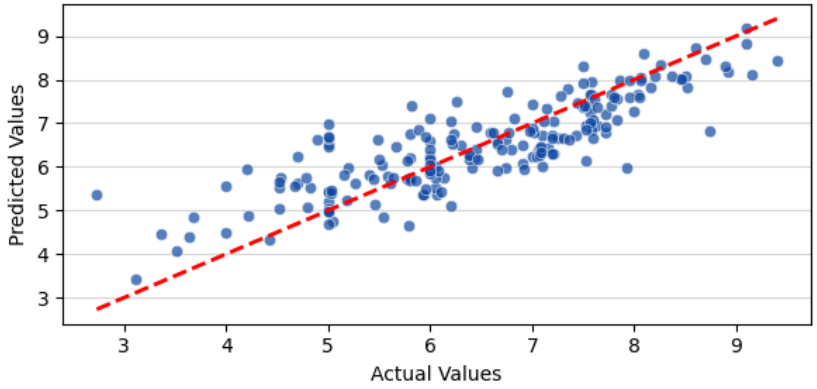

Figure 4. The Random Forest model's actual versus predicted pIC_{50} values.

This initial analysis underscores the potential of combining molecular descriptors with machine learning techniques to uncover meaningful patterns in the chemical space. The observed clustering and correlations provide a foundation for predictive modeling, enabling us to build robust machine-learning models for predicting the potency of AXL tyrosine kinase inhibitors.

3.2. Machine Learning Model Performance Evaluation

The predictive performance of four machine learning models was evaluated on the test set using four metrics: R^2, MAE, RMSE, and PCC. The results are summarized in Table 1.

The Random Forest model demonstrated the best performance, achieving an R^2 of 0.703, the lowest MAE (0.553), and the lowest RMSE (0.720). It also displayed a strong linear correlation between predicted and actual values, as reflected by its high PCC (0.841). These results highlight the model's ability to capture the complex relationships between molecular descriptors and inhibitor potency while maintaining generalizability on unseen data.

Gradient Boosting and Support Vector Regression showed comparable performance, with R^2 values of 0.687 and 0.688, respectively. Both models demonstrated similar MAE and RMSE scores, with Gradient Boosting slightly outperforming Support Vector Regression regarding RMSE (0.740 vs. 0.738). However, their PCC values (0.829 and 0.830) suggest a marginally weaker correlation compared to Random Forest. These models effectively captured non-linear patterns in the data. Still, their performance slightly lagged behind Random Forest, possibly due to overfitting or less optimal handling of the high-dimensional descriptor space.

The Decision Tree model, while interpretable, performed significantly worse than the other models. It achieved an R^2 of 0.357, the highest error metrics (MAE: 0.757, RMSE: 1.060), and the lowest PCC (0.685). This result underscores the limitations of simple tree-based models in handling the complexity of molecular descriptor data and highlights the necessity of ensemble or advanced techniques for this task.

The performance metrics indicate that ensemble methods, particularly Random Forest, are well-suited for predicting AXL inhibitor potency. The strong performance of these models demonstrates their ability to generalize across diverse molecular structures, making them valuable tools for drug discovery tasks. Incorporating Random Forest into the pipeline further supports its potential for identifying structure-activity relationships and aiding rational design.

As the Random Forest model demonstrated the best performance in predicting AXL inhibitor potency, its predictions were further analyzed to evaluate its reliability and interpretability. Figure 4 illustrates the actual versus predicted pIC_{50} values for the test set, showing a strong agreement between the predicted and actual pIC_{50} values, with points clustering closely around the red dashed line, representing the perfect correlation line (slope = 1). This indicates that the Random Forest model effectively captures the underlying relationships between molecular descriptors and AXL inhibitor potency. The spread of points is minimal, particularly in

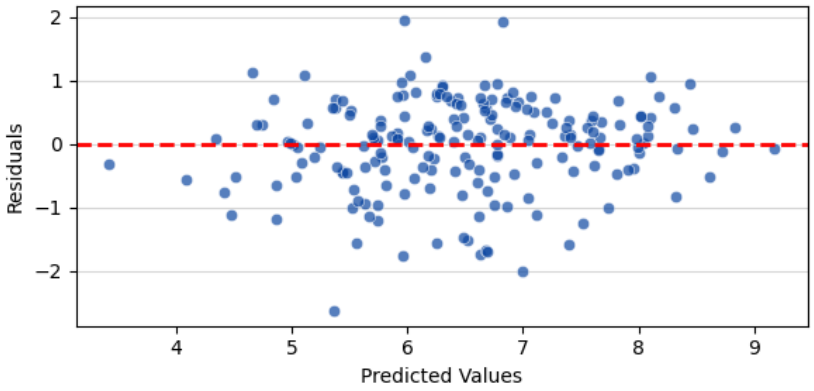

Figure 5. The residual plot for the Random Forest model shows evenly distributed residuals and minimal systematic bias across predicted pIC_{50} values.

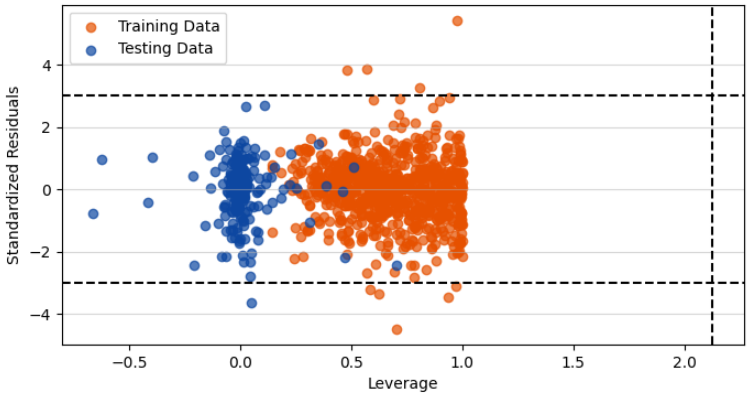

Figure 6. William's plot depicts leverage and standardized residuals for training and test datasets, confirming model robustness and identifying potential outliers.

the middle pIC_{50} range, suggesting robust model performance across most of the potency spectrum.

However, minor deviations are observed at the extreme ends of the pIC_{50} values, where the model exhibits slightly higher prediction errors. These deviations could result from the relatively lower representation of compounds in these dataset regions, highlighting a potential area for improvement. Nevertheless, the alignment demonstrates the model's capacity for generalizing to unseen data and supports its suitability for predictive tasks in AXL inhibitor discovery.

A residual plot (Figure 5) was generated to further assess the Random Forest model's performance. Residuals, defined as the difference between actual and predicted pIC_{50} values, provide insights into the model's prediction errors and potential biases. The residuals are evenly distributed around the red dashed line (residual = 0), indicating that the model does not exhibit systematic over- or under-prediction across the predicted range. This uniformity suggests that the Random Forest model captures the complex relationships in the dataset effectively and does not suffer from significant bias.

Most residuals fall within a narrow range, reflecting the model's accuracy. However, several points with larger residuals are observed, particularly at the lower and higher ends of the predicted pIC_{50} range. These outliers may be due to inherent noise in the data, less frequent representation of extreme potency compounds, or unique molecular properties not fully captured by the descriptors used.

The Williams plot (Figure 6) was used to evaluate the robustness and reliability of the Random Forest model by analyzing leverage and standardized residuals for both the training and test datasets. Leverage quantifies the influence of individual data points on the model, while standardized residuals measure the deviation of predicted values from actual values. The horizontal dashed lines on the plot represent the ±3 threshold for standardized residuals, beyond which points are considered outliers. The vertical dashed line marks the critical leverage value.

For the training data (orange points), most data points fall within the acceptable range for standardized residuals and leverage. This indicates that the model is not excessively biased or overfitting to specific compounds. A

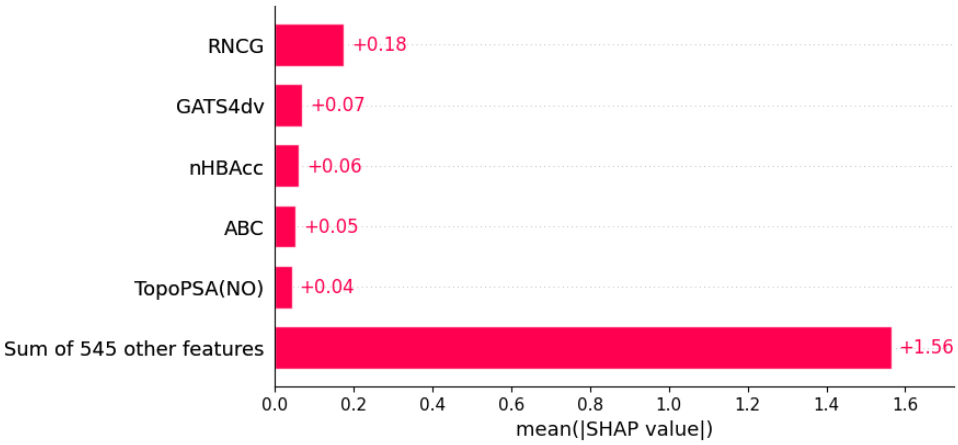

Figure 7. The SHAP feature importance bar plot highlights the top molecular descriptors contributing to AXL inhibitor potency predictions.

few training points exceed the leverage threshold, suggesting these are influential compounds with unique chemical features or extreme descriptor values. These compounds may significantly impact the model's predictions and are worth further investigation for their potential as outliers or sources of unique insights.

For the test data (blue points), most points are similarly confined within the acceptable limits, with very few points displaying high leverage or residual values. This highlights the generalization capability of the Random Forest model, as it performs consistently well on unseen data without significant deviations or over-reliance on specific data points. Overall, the Williams plot demonstrates that the model performs robustly across both datasets, with a few high-leverage or outlier compounds that could warrant additional exploration. These compounds might represent opportunities for discovering novel chemical scaffolds or improving the quality and diversity of the dataset.

3.3. SHAP Analysis for Interpretability and Feature Importance

To interpret the predictions of the Random Forest model and identify the key molecular descriptors influencing AXL inhibitor potency, SHAP (Shapley Additive Explanations) analysis was performed. The feature importance results are visualized in Figure 7, which highlights the top five descriptors contributing to the model's predictions.

Figure 7 shows the mean absolute SHAP values for the most important molecular descriptors, providing actionable insights for AXL inhibitor design. The descriptor RNCG has the highest impact on model predictions, with a mean SHAP value of +0.18, indicating its strong influence on potency. This suggests optimizing structural or physicochemical properties captured by RNCG could enhance inhibitor effectiveness. Other key descriptors, including GATS4dv, nHBAcc, ABC, and TopoPSA(NO), highlight the role of molecular topology, hydrogen bonding capacity, and polar surface area in drug-receptor interactions. For example, nHBAcc (hydrogen bond acceptor count) suggests that increasing hydrogen bond acceptors may improve binding affinity. At the same time, TopoPSA(NO) (topological polar surface area) indicates that modulating polarity could optimize drug permeability and target interaction. These insights provide a rational basis for structural modifications, guiding the design of more potent and selective AXL inhibitors by fine-tuning molecular properties that drive activity. The cumulative contribution of the remaining 545 features, with a mean SHAP value of +1.56, underscores the dataset's complexity, reinforcing the need for machine learning-driven feature selection to streamline drug optimization.

To complement the feature importance analysis presented in the bar plot (Figure 7), a SHAP beeswarm plot (Figure 8) was generated to provide a detailed overview of the variability and distribution of SHAP values for individual predictions across the dataset. This plot illustrates the SHAP values for the top five most important molecular descriptors and the remaining features' cumulative contribution. Each point represents a single compound, with the color indicating the feature value (red for high and blue for low). The horizontal spread of points reflects the range of SHAP values, showing the positive or negative impact of the feature on the model's prediction and highlighting the extent to which each molecular descriptor influences AXL inhibitor potency, offering valuable insights into the structural characteristics driving activity.

Unlike the bar plot, which summarizes the overall average importance of each feature, the beeswarm plot

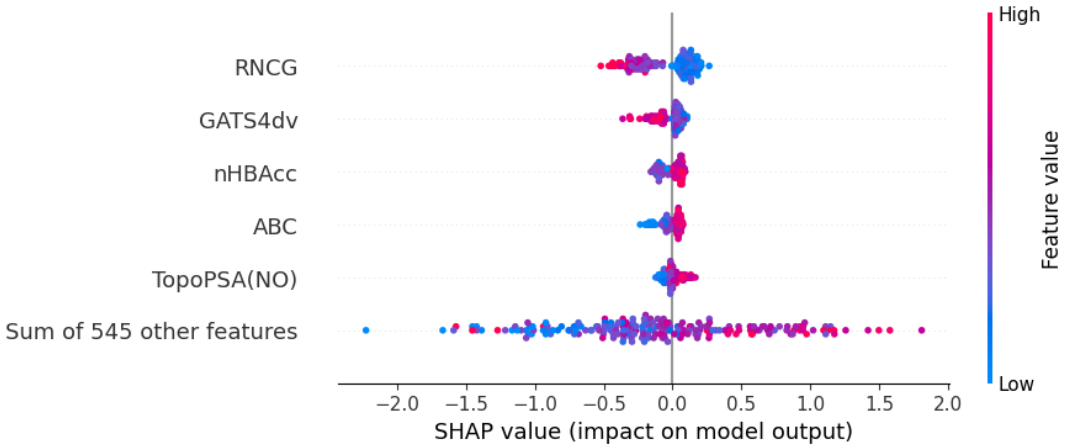

Figure 8. SHAP beeswarm plot displaying the variability and distribution of feature contributions for the top molecular descriptors across individual predictions.

captures the variability and interaction effects of descriptors at the level of individual predictions. For example, the descriptor RNCG consistently contributes positively to model predictions, but the extent of its impact varies depending on the compound. In contrast, TopoPSA(NO) shows positive and negative contributions, suggesting a more complex relationship with AXL inhibitor potency.

A SHAP dependence plot was generated to investigate the relationship between the most important. This plot provides insight into how variations in the feature value impact the predicted pIC_{50} values.

Figure 9a shows that as the value of RNCG increases, its SHAP value generally decreases, indicating a negative correlation between RNCG and the predicted potency of AXL inhibitors. Compounds with lower RNCG values tend to exhibit higher positive SHAP values, suggesting a stronger contribution to higher predicted pIC_{50} values. Conversely, compounds with higher RNCG values tend to negatively influence the model's predictions.

Figure 9b shows a distinct pattern in which compounds with lower GATS4dv values (ranging between -2 and 0) generally have positive SHAP values, indicating a favorable contribution to higher predicted pIC_{50} values. As GATS4dv values increase beyond 0, the SHAP values shift to negative, suggesting a diminishing contribution to inhibitor potency. Interestingly, the plot reveals a clear separation in the SHAP value distribution, with compounds clustering into two distinct regions. This indicates that GATS4dv may interact with other descriptors or reflect distinct chemical subgroups within the dataset, which warrants further investigation.

Figure 9c indicates a generally positive relationship between nHBAcc values and SHAP values. Compounds with lower nHBAcc values (less than -1) exhibit predominantly negative SHAP values, suggesting that fewer hydrogen bond acceptors diminish the predicted potency. As the nHBAcc values increase towards positive values, the SHAP values progressively shift upwards, indicating an increasing contribution to higher predicted pIC_{50} values. This trend aligns with the importance of hydrogen bonding in drug-receptor interactions, as compounds with more hydrogen bond acceptors are likely to exhibit stronger binding affinity.

The plot also reveals a plateau in SHAP values for nHBAcc values near or above 0, suggesting that additional hydrogen bond acceptors may no longer significantly enhance potency beyond a certain threshold. This insight is valuable for prioritizing compounds during optimization, as it highlights a diminishing return on increasing nHBAcc.

Figure 9d reveals a positive correlation between ABC values and SHAP values. The SHAP values are predominantly negative for lower ABC values (below -1), indicating a detrimental effect on the predicted pIC_{50} values. As the ABC values increase beyond 0, the SHAP values transition to positive, suggesting that higher ABC values contribute favorably to the predicted potency of AXL inhibitors.

This trend implies that ABC, potentially capturing specific structural or physicochemical properties, is an essential determinant of inhibitor activity. Notably, the plot demonstrates a smooth, nonlinear relationship, with the effect of ABC plateauing at higher values, indicating that further increases in ABC may have a limited impact on potency beyond a certain threshold.

Figure 9e reveals a positive correlation between TopoPSA(NO) values and their SHAP contributions to the model's output. The SHAP values are negative for lower

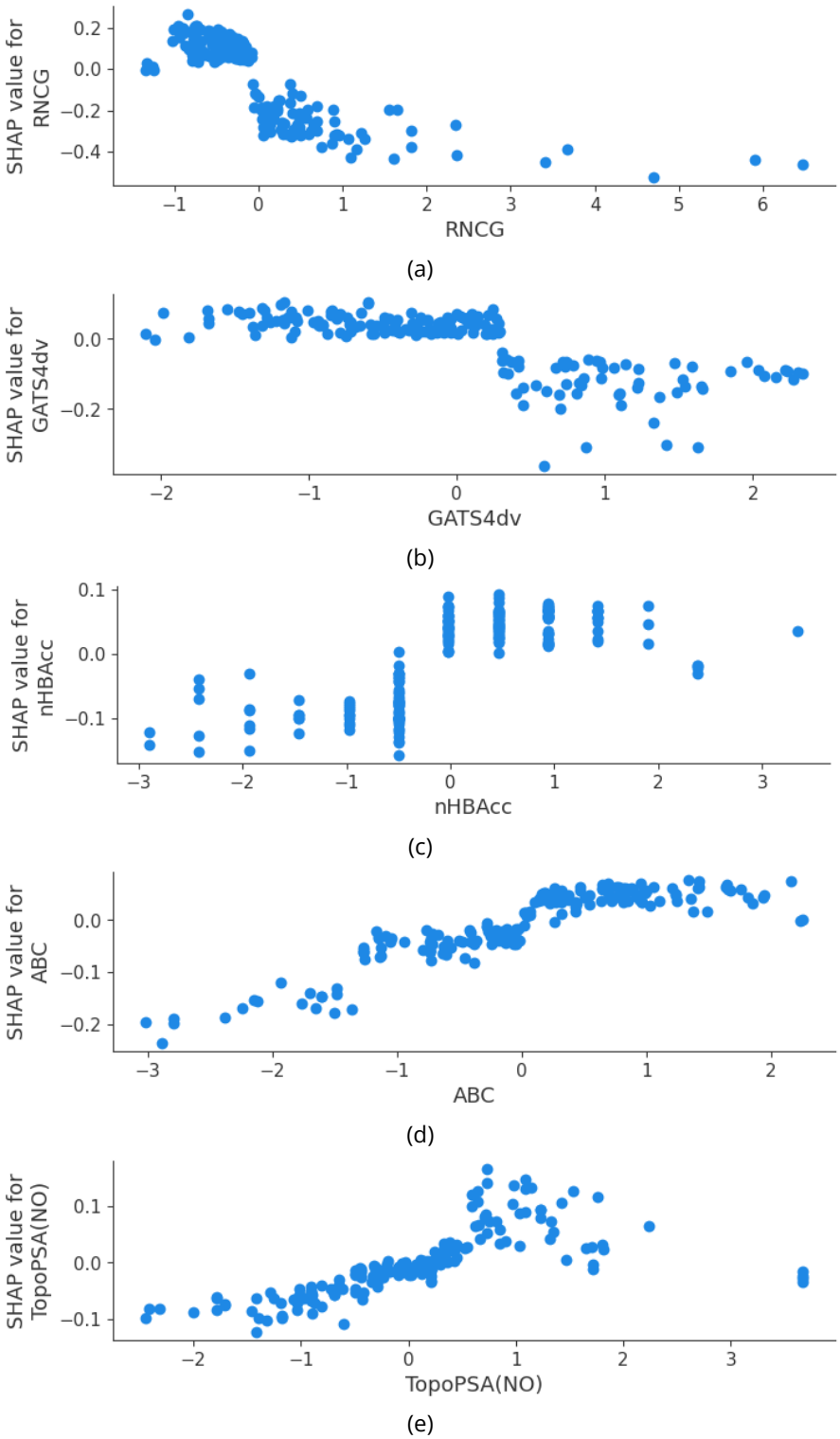

Figure 9. SHAP dependence plots showing the relationship between key molecular descriptors and their influence on predicted AXL inhibitor potency, (a) RNCG; (b) GATS4dv; (c) nHBAcc; (d) ABC; (e) TopoPSA(NO).

TopoPSA(NO) (below -1), indicating a reduced contribution to predicted potency. As TopoPSA(NO) values increase beyond 0, the SHAP values shift to positive, demonstrating an increasing influence on higher pIC_{50} predictions. This trend suggests that the topological polar surface area (linked to nitrogen and oxygen atoms) is a critical determinant of inhibitor activity.

The plot also highlights a nonlinear relationship, with TopoPSA(NO) values between 1 and 2 having the strongest positive impact, after which the effect appears to plateau. This observation suggests that optimizing polar surface area within a specific range may enhance AXL inhibitor potency, while further increases beyond this range may offer diminishing returns.

3.4. Limitations and Future Directions

While the study demonstrates the effectiveness of machine learning models, particularly Random Forest, in predicting AXL inhibitor potency and uncovering structure-activity relationships, several limitations should be addressed to enhance the robustness and applicability of these findings.

One significant limitation lies in the dataset itself. Despite preprocessing to ensure quality, the dataset contains inherent biases, such as an uneven distribution of pIC_{50} values and the underrepresenting of certain chemical scaffolds. This imbalance may lead to suboptimal performance when predicting the potency of compounds with rare or underrepresented structural features. Future work could focus on curating a more diverse dataset by incorporating additional high-quality data from experimental studies or public databases to better capture the chemical space relevant to AXL inhibition.

Another limitation involves the reliance on 2D molecular descriptors. Although these descriptors provide valuable insights into compounds' structural and physicochemical properties, they may fail to fully capture the 3D conformations and dynamic interactions critical to drug-receptor binding. Integrating 3D molecular descriptors or molecular docking scores into the modeling pipeline could significantly enhance predictive performance and biological relevance. Moreover, developing machine learning models that incorporate hybrid approaches, such as combining molecular dynamics simulations with experimental data, could provide deeper insights into the mechanisms of AXL inhibition.

Finally, while this study emphasizes prediction and interpretability, transitioning from computational predictions to practical drug discovery remains challenging. Future work should prioritize experimental validation of the predicted high-potency compounds and employ active learning techniques to iteratively refine models using experimental feedback. This iterative framework could accelerate the identification and optimization of promising AXL inhibitors.

4. Conclusions

In this study, we successfully leveraged machine learning techniques to predict the potency of AXL tyrosine kinase inhibitors, providing interpretable insights into the molecular features driving their activity. Among the models evaluated, the Random Forest model demonstrated the best predictive performance, with strong generalization and minimal bias across diverse chemical structures. SHAP analysis identified key molecular descriptors, such as RNCG and TopoPSA(NO), as critical determinants of AXL inhibitor potency, offering valuable guidance for rational drug design. While the results highlight the potential of machine learning in accelerating drug discovery, limitations such as dataset bias and the exclusion of 3D molecular interactions underline the need for future work incorporating diverse datasets, 3D structural features, and experimental validation. These findings advance AI-driven drug discovery, paving the way for developing effective cancer therapeutics targeting AXL.

Author Contributions: Conceptualization, T.R.N. and R.I.; methodology, T.R.N., E.H. F.M.F. and R.I.; software, T.R.N. and G.M.I.; validation, E.H. and R.I.; formal analysis, T.R.N., G.M.I. and I.S.; investigation, T.R.N. and G.M.I.; resources, E.H.; data curation, F.M.F. and R.I.; writing—original draft preparation, T.R.N., G.M.I., and E.H.; writing—review and editing, I.S., F.M.F. and R.I.; visualization, T.R.N. and G.M.I.; supervision, R.I.; project administration, R.I.; funding acquisition, R.I. All authors have read and agreed to the published version of the manuscript.

Funding: This study does not receive external funding.

Ethical Clearance: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study is available upon request from the corresponding author.

Conflicts of Interest: All the authors declare no conflicts of interest.

References

1. Zhu, C., Wei, Y., and Wei, X. (2019). AXL Receptor Tyrosine Kinase as a Promising Anti-cancer Approach: Functions, Molecular Mechanisms and Clinical Applications, *Molecular Cancer*, Vol. 18, No. 1, 153. doi:10.1186/s12943-019-1090-3.
2. Batur, T., Argundogan, A., Keles, U., Mutlu, Z., Alotaibi, H., Senturk, S., and Ozturk, M. (2021). AXL Knock-Out in SNU475 Hepatocellular Carcinoma Cells Provides Evidence for Lethal Effect Associated with G2 Arrest and Polyploidization, *International Journal of Molecular Sciences*, Vol. 22, No. 24, 13247. doi:10.3390/ijms222413247.
3. Yoshimura, A., Yamada, T., Serizawa, M., Uehara, H., Tanimura, K., Okuma, Y., Fukuda, A., Watanabe, S., Nishioka, N., Takeda, T., Chihara, Y., Takemoto, S., Harada, T., Hiranuma, O., Shirai, Y., Shukuya, T., Nishiyama, A., Goto, Y., Shiotsu, S., Kunimasa, K., Morimoto, K., Katayama, Y., Suda, K., Mitsudomi, T., Yano, S., Kenmotsu, H., Takahashi, T., and Takayama, K. (2023). High Levels of AXL Expression in Untreated EGFR-Mutated Non-Small Cell Lung Cancer Negatively Impacts the Use of Osimertinib, *Cancer Science*, Vol. 114, No. 2, 606–618. doi:10.1111/cas.15608.
4. Ozyurt, R., and Ozpolat, B. (2023). Therapeutic Landscape of AXL Receptor Kinase in Triple-Negative Breast Cancer, *Molecular Cancer Therapeutics*, Vol. 22, No. 7, 818–832. doi:10.1158/1535-7163.MCT-22-0617.
5. Tang, Y., Zang, H., Wen, Q., and Fan, S. (2023). AXL in Cancer: A Modulator of Drug Resistance and Therapeutic Target, *Journal of Experimental & Clinical Cancer Research*, Vol. 42, No. 1, 148. doi:10.1186/s13046-023-02726-w.
6. Phatak, S. S., Stephan, C. C., and Cavasotto, C. N. (2009). High-Throughput and In Silico Screenings in Drug Discovery, *Expert*

Opinion on Drug Discovery, Vol. 4, No. 9, 947–959. doi:10.1517/17460440903190961.

7. van Dongen, M., Weigelt, J., Uppenberg, J., Schultz, J., and Wikström, M. (2002). Structure-Based Screening and Design in Drug Discovery, *Drug Discovery Today*, Vol. 7, No. 8, 471–478. doi:10.1016/S1359-6446(02)02233-X.

8. Mukherjee, D., Sharma, T., Lunawat, A. K., Awasthi, A., Kurmi, B. Das, Kumar, M., Gupta, G. Das, and Thakur, S. (2024). Advancement of Artificial Intelligence in Drug Discovery: A Comprehensive Review, *Current Artificial Intelligence*, Vol. 03. doi:10.2174/0129503752322569241104114248.

9. Lynch, C., Sakamuru, S., Ooka, M., Huang, R., Klumpp-Thomas, C., Shinn, P., Gerhold, D., Rossoshek, A., Michael, S., Casey, W., Santillo, M. F., Fitzpatrick, S., Thomas, R. S., Simeonov, A., and Xia, M. (2024). High-Throughput Screening to Advance In Vitro Toxicology: Accomplishments, Challenges, and Future Directions, *Annual Review of Pharmacology and Toxicology*, Vol. 64, No. 1, 191–209. doi:10.1146/annurev-pharmtox-112122-104310.

10. Tiwari, P. C., Pal, R., Chaudhary, M. J., and Nath, R. (2023). Artificial intelligence revolutionizing drug development: Exploring opportunities and challenges, *Drug Development Research*, Vol. 84, No. 8, 1652–1663. doi:10.1002/ddr.22115.

11. Dara, S., Dhamercherla, S., Jadav, S. S., Babu, C. M., and Ahsan, M. J. (2022). Machine Learning in Drug Discovery: A Review, *Artificial Intelligence Review*, Vol. 55, No. 3, 1947–1999. doi:10.1007/s10462-021-10058-4.

12. Noviandy, T. R., Idroes, G. M., Mohd Fauzi, F., and Idroes, R. (2024). Application of Ensemble Machine Learning Methods for QSAR Classification of Leukotriene A4 Hydrolase Inhibitors in Drug Discovery, *Malacca Pharmaceutics*, Vol. 2, No. 2, 68–78. doi:10.60084/mp.v2i2.217.

13. Kumar, S. A., Ananda Kumar, T. D., Beeraka, N. M., Pujar, G. V., Singh, M., Narayana Akshatha, H. S., and Bhagyalalitha, M. (2022). Machine Learning and Deep Learning in Data-Driven Decision Making of Drug Discovery and Challenges in High-Quality Data Acquisition in the Pharmaceutical Industry, *Future Medicinal Chemistry*, Vol. 14, No. 4, 245–270. doi:10.4155/fmc-2021-0243.

14. Priya, S., Tripathi, G., Singh, D. B., Jain, P., and Kumar, A. (2022). Machine Learning Approaches and Their Applications in Drug Discovery and Design, *Chemical Biology & Drug Design*, Vol. 100, No. 1, 136–153. doi:10.1111/cbdd.14057.

15. Noviandy, T. R., Idroes, G. M., and Hardi, I. (2024). An Interpretable Machine Learning Strategy for Antimalarial Drug Discovery with LightGBM and SHAP, *Journal of Future Artificial Intelligence and Technologies*, Vol. 1, No. 2, 84–95. doi:10.62411/faith.2024-16.

16. Zhang, Y., Xu, F., Zou, J., Petrosian, O. L., and Krinkin, K. V. (2021). XAI Evaluation: Evaluating Black-Box Model Explanations for Prediction, *2021 II International Conference on Neural Networks and Neurotechnologies (NeuroNT)*, IEEE, 13–16. doi:10.1109/NeuroNT53022.2021.9472817.

17. Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., and Overington, J. P. (2012). ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery, *Nucleic Acids Research*, Vol. 40, No. D1, D1100–D1107. doi:10.1093/nar/gkr777.

18. Wigh, D. S., Goodman, J. M., and Lapkin, A. A. (2022). A Review of Molecular Representation in the Age of Machine Learning, *WIREs Computational Molecular Science*, Vol. 12, No. 5. doi:10.1002/wcms.1603.

19. Thakur, A., Kumar, A., Sharma, V. kumar, and Mehta, V. (2022). PIC50: An Open Source Tool for Interconversion of PIC50 Values and IC50 for Efficient Data Representation and Analysis, *BioRxiv*, 2010–2022.

20. Grisoni, F., Consonni, V., and Todeschini, R. (2018). Impact of Molecular Descriptors on Computational Models, 171–209. doi:10.1007/978-1-4939-8639-2_5.

21. Moriwaki, H., Tian, Y. S., Kawashita, N., and Takagi, T. (2018). Mordred: A Molecular Descriptor Calculator, *Journal of Cheminformatics*, Vol. 10, No. 1, 1–14. doi:10.1186/s13321-018-0258-y.

22. Ojha, P. K., and Roy, K. (2011). Comparative QSARs for Antimalarial Endochins: Importance of Descriptor-Thinning and Noise Reduction Prior to Feature Selection, *Chemometrics and Intelligent Laboratory Systems*, Vol. 109, No. 2, 146–161. doi:10.1016/j.chemolab.2011.08.007.

23. Yang, H., Du, Z., Lv, W.-J., Zhang, X.-Y., and Zhai, H.-L. (2019). In Silico Toxicity Evaluation of Dioxins Using Structure–Activity Relationship (SAR) and Two-Dimensional Quantitative Structure–Activity Relationship (2D-QSAR), *Archives of Toxicology*, Vol. 93, No. 11, 3207–3218. doi:10.1007/s00204-019-02580-w.

24. Noviandy, T. R., Maulana, A., Idroes, G. M., Suhendra, R., Afidh, R. P. F., and Idroes, R. (2024). An Explainable Multi-Model Stacked Classifier Approach for Predicting Hepatitis C Drug Candidates, *Sci*, Vol. 6, No. 4, 81. doi:10.3390/sci6040081.

25. Soni, J., Prabakar, N., and Upadhyay, H. (2020). Visualizing High-Dimensional Data Using t-Distributed Stochastic Neighbor Embedding Algorithm, 189–206. doi:10.1007/978-3-030-43981-1_9.

26. Noviandy, T. R., Maulana, A., Idroes, G. M., Emran, T. B., Tallei, T. E., Helwani, Z., and Idroes, R. (2023). Ensemble Machine Learning Approach for Quantitative Structure Activity Relationship Based Drug Discovery: A Review, *Infolitika Journal of Data Science*, Vol. 1, No. 1, 32–41. doi:10.60084/ijds.v1i1.91.

27. Noviandy, T. R., Idroes, G. M., Maulana, A., Afidh, R. P. F., and Idroes, R. (2024). Optimizing Hepatitis C Virus Inhibitor Identification with LightGBM and Tree-structured Parzen Estimator Sampling, *Engineering, Technology & Applied Science Research*, Vol. 14, No. 6, 18810–18817. doi:10.48084/etasr.8947.

28. Suhendra, R., Husdayanti, N., Suryadi, S., Juliwardi, I., Sanusi, S., Ridho, A., Ardiansyah, M., Murhaban, M., and Ikhsan, I. (2023). Cardiovascular Disease Prediction Using Gradient Boosting Classifier, *Infolitika Journal of Data Science*, Vol. 1, No. 2, 56–62. doi:10.60084/ijds.v1i2.131.

29. Noviandy, T. R., Idroes, G. M., Tallei, T. E., Handayani, D., and Idroes, R. (2024). QSAR Modeling for Predicting Beta-Secretase 1 Inhibitory Activity in Alzheimer's Disease with Support Vector Regression, *Malacca Pharmaceutics*, Vol. 2, No. 2, 79–85. doi:10.60084/mp.v2i2.226.

30. Sasmita, N. R., Ramadeska, S., Kesuma, Z. M., Noviandy, T. R., Maulana, A., Khairul, M., and Suhendra, R. (2024). Decision Tree versus k-NN: A Performance Comparison for Air Quality Classification in Indonesia, *Infolitika Journal of Data Science*, Vol. 2, No. 1, 9–16. doi:10.60084/ijds.v2i1.179.

31. Winarsih, A., Idroes, R., Zulfiani, U., Yusuf, M., Mahmudi, M., Saiful, S., and Rahman, S. A. (2023). Method Validation for Pesticide Residues on Rice Grain in Aceh Besar District, Indonesia Using Gas Chromatography-Electron Capture Detector (GC-ECD), *Leuser Journal of Environmental Studies*, Vol. 1, No. 1, 18–24. doi:10.60084/ljes.v1i1.37.

32. Maulana, A., Idroes, G. M., Kemala, P., Maulydia, N. B., Sasmita, N. R., Tallei, T. E., Sofyan, H., and Rusyana, A. (2023). Leveraging Artificial Intelligence to Predict Student Performance: A Comparative Machine Learning Approach, *Journal of Educational Management and Learning*, Vol. 1, No. 2, 64–70. doi:10.60084/jeml.v1i2.132.

33. Noviandy, T. R., Maulana, A., Irvanizam, I., Idroes, G. M., Maulydia, N. B., Tallei, T. E., Subianto, M., and Idroes, R. (2025). Interpretable Machine Learning Approach to Predict Hepatitis C Virus NS5B Inhibitor Activity Using Voting-Based LightGBM and SHAP, *Intelligent Systems with Applications*, Vol. 25, 200481. doi:10.1016/j.iswa.2025.200481.

34. Wu, Y., Huo, D., Chen, G., and Yan, A. (2021). SAR and QSAR Research on Tyrosinase Inhibitors Using Machine Learning

Methods, *SAR and QSAR in Environmental Research*, Vol. 32, No. 2, 85–110. doi:10.1080/1062936X.2020.1862297.

35. Lundberg, S. M., and Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions, *Advances in Neural Information Processing Systems*, Vol. 30.

37. Moncada-Torres, A., van Maaren, M. C., Hendriks, M. P., Siesling, S., and Geleijnse, G. (2021). Explainable Machine Learning Can Outperform Cox Regression Predictions and Provide Insights in Breast Cancer Survival, *Scientific Reports*, Vol. 11, No. 1, 6968. doi:10.1038/s41598-021-86327-7.

38. Iqbal, A. B., Masoodi, T. A., Bhat, A. A., Macha, M. A., Assad, A., and Shah, S. Z. A. (2025). Explainable AI-Driven Prediction of APE1 Inhibitors: Enhancing Cancer Therapy with Machine Learning Models and Feature Importance Analysis, *Molecular Diversity*. doi:10.1007/s11030-025-11133-6.

36. Noviandy, T. R., Idroes, G. M., Syukri, M., and Idroes, R. (2024). Interpretable Machine Learning for Chronic Kidney Disease Diagnosis: A Gaussian Processes Approach, *Indonesian Journal of Case Reports*, Vol. 2, No. 1, 24–32. doi:10.60084/ijcr.v2i1.204.

39. Noviandy, T. R., Idroes, G. M., and Hardi, I. (2025). Integrating Explainable Artificial Intelligence and Light Gradient Boosting Machine for Glioma Grading, *Informatics and Health*, Vol. 2, No. 1, 1–8. doi:10.1016/j.infoh.2024.12.001.

40. Prendin, F., Pavan, J., Cappon, G., Del Favero, S., Sparacino, G., and Facchinetti, A. (2023). The Importance of Interpreting Machine Learning Models for Blood Glucose Prediction in Diabetes: An Analysis Using SHAP, *Scientific Reports*, Vol. 13, No. 1, 16865. doi:10.1038/s41598-023-44155-x.