Original Article

Check for updates

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

Malacca Pharmaceutics

Vol. 4, No. 1, 2026

QSAR Modeling of Beta-2 Adrenergic Receptor Ligands Using Molecular Descriptor–Based Machine Learning

Teuku Rizky Noviandy [1], Mohsina Patwekar [2], and Rinaldi Idroes [3,*]

[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 Pharmacology, Luqman College of Pharmacy, Karnataka 585102, India; mohsina.patwekar@gmail.com (M.P.)
[3] School of Mathematics and Applied Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia; rinaldi.idroes@usk.ac.id (R.I.)

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

Article History

Received 12 January 2026
Revised 15 March 2026
Accepted 23 March 2026
Available Online 31 March 2026

Keywords:
ADRB2
QSAR
Drug discovery
Ensemble methods

Abstract

The Beta-2 Adrenergic Receptor (ADRB2) is a well-characterized G protein–coupled receptor widely studied in pharmacology and drug discovery. In this study, quantitative structure–activity relationship (QSAR) models were developed using molecular descriptor–based machine learning approaches to predict the activity of ADRB2 ligands. A curated dataset of 745 ADRB2 ligands with experimentally determined IC_{50} values was obtained from the ChEMBL database, where IC_{50} values were used to define ligand activity. Two-dimensional molecular descriptors were calculated and preprocessed to remove low-variance and highly correlated features, resulting in a refined feature set for model development. The dataset was categorized into active and inactive compounds and divided into training and testing subsets. Four machine learning algorithms, Logistic Regression, Support Vector Machine, Gradient Boosting, and Random Forest, were implemented and evaluated using accuracy, precision, recall, F1-score, and ROC-AUC metrics. Among the models, Random Forest achieved the best performance, with an accuracy of 89.26%, an F1-score of 89.87%, and an AUC of 0.926, followed by Gradient Boosting with an accuracy of 87.92% and an AUC of 0.922. Analysis of physicochemical descriptors indicated that hydrogen-bond donor capacity (nHD) is statistically significantly associated with variation in compound activity toward ADRB2. In contrast, lipophilicity (LogP) and hydrogen-bond acceptor count (nHA) do not exhibit statistically significant differences between activity classes. Overall, the results demonstrate that molecular descriptor–based machine learning models, particularly ensemble methods, provide an effective framework for predicting ADRB2-related compound activity and for prioritizing candidate molecules in computational drug discovery.

1. Introduction

The Beta-2 Adrenergic Receptor (ADRB2) is a member of the G protein–coupled receptor (GPCR) superfamily that plays a central role in regulating physiological processes such as smooth muscle relaxation, bronchodilation, and metabolic activity [1, 2]. Due to its well-characterized structure, signaling mechanisms, and pharmacological relevance, ADRB2 has been extensively studied as a model receptor in drug discovery and molecular pharmacology [3]. A wide range of ligands, including agonists, antagonists, and inverse agonists, have been developed to modulate ADRB2 activity, making it a valuable system for investigating ligand–receptor interactions [4].

While ADRB2 agonists are widely used as bronchodilators in the treatment of asthma and chronic obstructive pulmonary disease (COPD), antagonists of ADRB2 are primarily of interest in broader pharmacological and biochemical contexts, including receptor selectivity, signaling modulation, and off-target interaction profiling [5]. In addition, ADRB2 is a well-curated target in public databases, where numerous compounds with experimentally determined bioactivity values are available [6]. This makes ADRB2 an attractive model system for developing and evaluating computational approaches in quantitative structure–activity relationship (QSAR) modeling.

Despite extensive experimental and computational studies, predicting the biological activity of ADRB2-targeting compounds remains challenging due to the complex and non-linear relationships between molecular structure and receptor binding [7]. Traditional drug discovery approaches are often time-consuming and resource-intensive, relying heavily on experimental screening and iterative optimization [8, 9]. These limitations highlight the importance of computational methods that can efficiently model structure–activity relationships and prioritize candidate molecules.

In recent years, machine learning techniques have significantly advanced QSAR modeling by enabling the analysis of large chemical datasets and capturing complex, non-linear patterns in molecular descriptor space [10]. Algorithms such as support vector machines, random forests, and gradient boosting have demonstrated strong predictive performance in various cheminformatics applications [11, 12]. By transforming molecular structures into numerical descriptors, machine learning models can learn the underlying relationships between chemical features and biological activity, thereby enabling rapid, cost-effective prediction of compound behavior.

Although numerous QSAR studies have applied machine learning to drug discovery, many are limited to single-model approaches, small datasets, or lack systematic comparison across different algorithms using standardized preprocessing pipelines. In addition, relatively few studies have focused specifically on ADRB2 antagonists using curated large-scale datasets and comprehensive descriptor filtering to reduce redundancy and improve model robustness.

In this study, we address these gaps by developing a molecular descriptor–based QSAR framework for predicting the activity of ADRB2 ligands using multiple machine learning algorithms. The main contributions of this work include the construction of a curated dataset of ADRB2 ligands with consistent IC_{50}-based classification from the ChEMBL database, the application of systematic descriptor preprocessing, including variance and correlation filtering, to improve data quality, and a comparative evaluation of both linear and non-linear machine learning models. Overall, this study provides a robust and reproducible computational approach for modeling ADRB2 antagonist activity and contributes to the broader application of machine learning in QSAR-based drug discovery.

2. Materials and Methods

2.1. Data Collection

The dataset used in this study was obtained from the ChEMBL database, a comprehensive repository of curated bioactivity data for drug discovery [13]. Compounds associated with the Beta-2 Adrenergic Receptor (ADRB2; ChEMBL ID: CHEMBL210) were retrieved based on their reported bioactivity data, with IC_{50} values representing ligand potency measured in inhibition assays.

A total of 745 compounds with experimentally determined half-maximal inhibitory concentration (IC_{50}) values were collected. In this study, IC_{50} values were used as a quantitative measure of compound potency to characterize functional inhibition of receptor activity in biochemical or cell-based assays. Only records with clearly defined IC_{50} values and corresponding molecular structures were included to ensure data quality and consistency.

Data curation was performed to improve reliability and remove potential sources of bias [14]. Duplicate entries, incomplete records, and compounds lacking valid structural information were excluded. When multiple measurements were reported for the same compound, inconsistent or ambiguous entries were removed to maintain dataset integrity. The resulting curated dataset provides a reliable foundation for QSAR modeling and machine learning analysis. Representative chemical structures of selected compounds from the dataset are shown in Figure 1.

2.2. Molecular Descriptors Calculation

Molecular descriptors are quantitative representations of chemical structures that encode physicochemical, topological, and structural properties of molecules into numerical values [15]. These descriptors serve as essential input features in QSAR modeling, enabling the translation of molecular information into a format suitable for computational analysis and machine learning

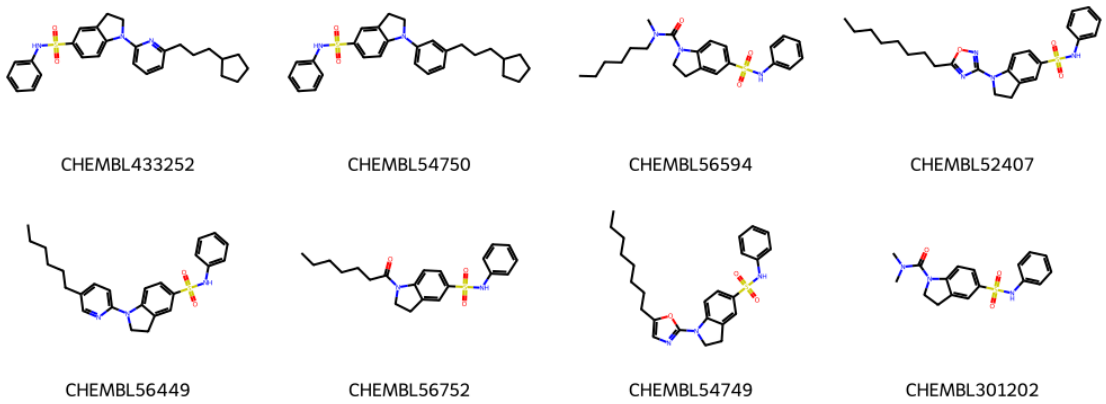

Figure 1. Representative chemical structures of selected compounds targeting ADRB2 obtained from the ChEMBL database.

algorithms [16]. In this study, two-dimensional (2D) molecular descriptors were calculated for all compounds using the Mordred software package. Mordred is a widely used cheminformatics tool capable of generating a comprehensive set of descriptors from molecular structures, covering various categories such as constitutional, topological, electronic, and geometrical properties [17].

2.3. Data Preprocessing

For the classification task, the biological activity of compounds targeting ADRB2 was categorized based on their reported IC_{50} values. Compounds with $IC_{50} \geq 1000$ nM were labeled as inactive, whereas those with $IC_{50} < 1000$ nM were classified as active [10]. In this context, IC_{50} values serve as a quantitative measure of compound potency in modulating receptor activity under experimental conditions [18].

Before model development, molecular descriptors underwent several preprocessing steps to improve data quality and reduce redundancy [19]. First, descriptors with low variance were removed using a variance threshold of 0.1, as such features contribute minimal discriminatory information [20]. Next, highly correlated (colinear) descriptors were eliminated using a correlation threshold of 0.80 to minimize multicollinearity and prevent model overfitting [12]. After preprocessing, the final dataset comprised 745 compounds and 155 molecular descriptors. The dataset included 385 inactive and 360 active compounds, indicating an approximately balanced dataset class distribution suitable for classification modeling.

Finally, the dataset was divided into training and testing subsets using an 80:20 ratio. The training set was used for model development and optimization, while the independent test set was reserved for evaluating the predictive performance and generalizability of the developed models [21].

2.4. Exploratory Data Analysis

To better understand the dataset's characteristics and evaluate potential distinctions between active and inactive compounds, an exploratory data analysis (EDA) was conducted. The key physicochemical properties associated with drug-likeness were analyzed based on Lipinski's Rule of Five [22]. The evaluated descriptors included molecular weight (MW), the logarithm of the partition coefficient (logP), the number of hydrogen-bond donors (nHD), and the number of hydrogen-bond acceptors (nHA). These properties are widely recognized as important indicators of molecular behavior, influencing solubility, permeability, and overall bioavailability.

To quantitatively compare the distributions of these descriptors between active and inactive compounds, descriptive statistics, including the mean, median, standard deviation, and interquartile range, were calculated. Furthermore, the Mann–Whitney U test was employed to determine whether statistically significant differences exist between the two classes, as it does not assume a normal distribution of the data.

Finally, boxplot visualizations were generated for each selected descriptor to provide a clear comparison of their distributions across activity classes. These visual analyses helped identify potential trends and distinguishing physicochemical features associated with differences in compound activity toward ADRB2.

2.5. Machine Learning Models

In this study, several supervised machine learning algorithms were employed to develop classification models to predict compound activity against ADRB2. The selected algorithms included Logistic Regression,

Table 1. Descriptive statistics and Mann-Whitney U test results for Lipinski's Rule of Five descriptors.

	MW		LogP		nHD		nHA	
	Inactive	**Active**	**Inactive**	**Active**	**Inactive**	**Active**	**Inactive**	**Active**
p-value	5.09×10^{-1}		1.45×10^{-1}		1.06×10^{-21}		6.71×10^{-1}	
Min	185.23	183.21	-0.34	-0.19	0	1	2	2
Max	770.97	1312.68	10.44	8.03	8	8	13	18
Median	477.05	490.63	4.07	3.75	3	4	6	6
Mean	492.27	505.08	4.09	3.91	3.03	3.92	6.34	6.76
Skew	-0.22	0.95	0.33	0.1	1.04	-0.08	-0.07	1.23
Kurtosis	-0.48	3.39	0.61	-0.43	3.82	-0.46	-0.16	1.96

Note: The p-value represents the result of the Mann-Whitney U test.

Support Vector Machine, Gradient Boosting, and Random Forest, representing both linear and non-linear modeling approaches commonly used in QSAR studies [23–25].

Logistic Regression was utilized as a baseline linear model due to its simplicity and interpretability in binary classification tasks. The Support Vector Machine algorithm was applied to capture complex decision boundaries by maximizing the margin between classes. Ensemble learning methods, including Random Forest and Gradient Boosting, were also implemented to improve predictive performance by combining multiple decision trees and reducing variance and bias.

All models were implemented using the scikit-learn library. Default hyperparameter settings provided by the library were used for each algorithm to ensure a fair and unbiased comparison of model performance without extensive parameter tuning. The models were trained on the training dataset and subsequently evaluated on the independent test set to assess their generalization ability [21]. This approach allows for a consistent comparison of different algorithms in predicting compound activity based on molecular descriptors.

2.6. Performance Evaluation

The performance of the developed classification models was assessed using several standard evaluation metrics on the independent test set [26]. These metrics included accuracy, precision, recall, and F1-score, which collectively provide a comprehensive evaluation of the model's performance in distinguishing active and inactive compounds targeting ADRB2.

Accuracy measures the overall proportion of correctly classified instances, while precision evaluates the proportion of true positive predictions among all predicted positives. Recall (sensitivity) reflects the model's ability to correctly identify active compounds, and the F1-score, defined as the harmonic mean of precision and recall, provides a balanced measure of classification performance, particularly in cases of class imbalance.

In addition to these metrics, the models' discriminative ability was further evaluated using the receiver operating characteristic (ROC) curve. The ROC curve illustrates the trade-off between the true positive rate and false positive rate across different classification thresholds. The area under the ROC curve (AUC) was calculated as a quantitative measure of model performance; higher AUC values indicate better classification performance. ROC curves were plotted for each model to visually compare predictive performance and assess effectiveness in distinguishing active from inactive compounds.

3. Results and Discussion

3.1. Results of Exploratory Data Analysis

Table 1 summarizes the descriptive statistics and results of the Mann–Whitney U test for key physicochemical descriptors based on Lipinski's Rule of Five, including molecular weight, lipophilicity, number of hydrogen bond acceptors, and number of hydrogen bond donors. These properties are widely associated with drug-likeness and can influence compound interactions with ADRB2. It should be noted that these descriptors were used solely for exploratory analysis and do not represent the complete set of features employed in model development.

The MW distribution showed similar characteristics for inactive and active compounds. The mean MW values were 492.27 and 505.08 for inactive and active compounds, respectively, with medians of 477.05 and 490.63. Although active compounds tended to have slightly higher molecular weights, the difference was not statistically significant, as indicated by the Mann–Whitney U test ($p = 5.09 \times 10^{-1}$). The skewness values suggest that inactive compounds exhibited a slightly left-skewed distribution (−0.22), whereas active compounds showed a right-skewed distribution (0.95), indicating the presence of higher-MW outliers. Additionally, the higher kurtosis value for active compounds (3.39) indicates a more peaked distribution with heavier tails than that of inactive compounds.

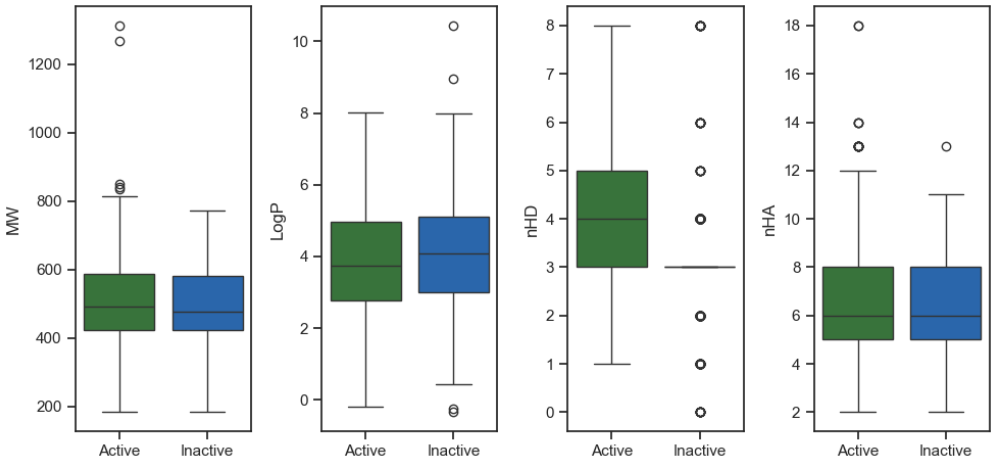

Figure 2. Boxplots of MW, LogP, nHD, and nHA for active vs. inactive ADRB2 compounds.

In contrast, LogP did not differ significantly between the two groups ($p = 1.45 \times 10^{-1}$). Inactive compounds showed a slightly higher mean LogP (4.09) compared to active compounds (3.91), with medians of 4.07 and 3.75, respectively. While these differences suggest that active compounds tend to be moderately less lipophilic, the lack of statistical significance indicates that lipophilicity alone may not be a strong differentiating factor for compound activity toward ADRB2. The relatively similar skewness values (0.33 for inactive and 0.10 for active) indicate near-symmetric distributions, although inactive compounds displayed slightly greater variability, as reflected by a higher kurtosis (0.61).

The number of hydrogen-bond donors (nHD) exhibited a statistically significant difference between active and inactive compounds ($p = 1.06 \times 10^{-21}$). Active compounds had a higher mean nHD value (3.92) compared to inactive compounds (3.03), with median values of 4 and 3, respectively. The distribution of nHD for inactive compounds was positively skewed (1.04). In contrast, active compounds showed a slightly negative skewness (−0.08), indicating a more balanced distribution with fewer high-value outliers.

In contrast, the number of hydrogen-bond acceptors (nHA) did not show a statistically significant difference between active and inactive compounds ($p = 6.71 \times 10^{-1}$). Active compounds exhibited slightly higher mean values (6.76 vs. 6.34), with identical median values (6). The distribution of nHA for active compounds was more positively skewed (1.23), indicating the presence of compounds with higher acceptor counts, whereas inactive compounds showed an approximately symmetric distribution (−0.07). Despite these differences, the substantial overlap between distributions suggests that nHA is not a strong distinguishing factor for compound activity.

Figure 2 illustrates the boxplot distributions of key Lipinski's Rule of Five descriptors: MW, LogP, nHD, and nHA for active and inactive compounds. These visualizations provide further insight into the differences and overlaps between the two classes. These descriptors were analyzed for exploratory purposes and do not represent the full feature set used in model development.

The MW boxplots show a substantial overlap between active and inactive compounds, confirming the lack of a statistically significant difference observed earlier. Both classes exhibit similar interquartile ranges and median values, indicating comparable central tendencies. However, active compounds display a wider spread with several high-value outliers exceeding 1200 Da, suggesting that while most compounds fall within a similar MW range, a subset of active compounds extends toward higher molecular weights. This observation aligns with the positive skewness reported for active compounds and indicates that MW alone is not a strong differentiating factor for compound activity toward ADRB2.

For LogP, the distributions of active and inactive compounds are relatively close, with only subtle differences observed. Inactive compounds tend to have slightly higher median and upper-range values, while active compounds show a somewhat narrower spread. The presence of extreme values in the inactive group contributes to greater variability. However, consistent with the statistical analysis, these differences are not statistically significant, suggesting that lipophilicity alone may not strongly distinguish compound activity.

Table 2. Performance comparison of machine learning models for predicting ADRB2 compound activity.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Logistic Regression	84.56	83.33	88.61	85.89
Gradient Boosting	87.92	88.61	88.61	88.61
Random Forest	89.26	89.87	89.87	89.87
Support Vector Machine	85.91	87.18	86.08	86.62

The nHD distributions reveal a clearer distinction between the two groups. Active compounds exhibit a broader interquartile range and a slightly higher median compared to inactive compounds, indicating, on average, a greater number of hydrogen-bond donors. Inactive compounds appear more concentrated around lower nHD values, with several outliers extending beyond the central range. This pattern is consistent with the statistically significant difference observed in Table 1. It suggests that hydrogen-bond donor capacity is associated with variations in compound activity toward ADRB2, although overlap between groups indicates it is not the sole determining factor.

In contrast, the nHA boxplots show considerable overlap between active and inactive compounds. Although active compounds tend to have slightly higher median values, the overall distributions are similar, and the presence of outliers in both groups indicates substantial variability. These observations are consistent with the statistical results, which show no significant difference in nHA, suggesting that the hydrogen-bond acceptor count alone is not a strong differentiating feature of compound activity.

3.2. Performance of Machine Learning Models

Table 2 summarizes the performance of the four machine learning models evaluated for predicting the activity of compounds targeting ADRB2, including Logistic Regression, Gradient Boosting, Random Forest, and Support Vector Machine. The models were assessed using accuracy, precision, recall, and F1-score to provide a comprehensive evaluation of classification performance.

Among the tested models, Random Forest achieved the highest overall performance, with an accuracy of 89.26% and balanced precision, recall, and F1-score values of 89.87%. This indicates that RF was the most effective model at distinguishing active from inactive compounds while maintaining consistency across all evaluation metrics. The strong performance of RF can be attributed to its ensemble nature, which combines multiple decision trees to capture complex, non-linear relationships within the molecular descriptor space while reducing overfitting.

Gradient Boosting also demonstrated strong predictive capability, achieving an accuracy of 87.92% and identical precision, recall, and F1-score values of 88.61%. The consistent performance across these metrics suggests that the model maintains a good balance between sensitivity and specificity. As another ensemble method, GB iteratively improves model performance by focusing on previously misclassified instances, which likely contributes to its high accuracy and robustness.

The Support Vector Machine model showed competitive performance, with an accuracy of 85.91% and an F1-score of 86.62%. It achieved relatively high precision (87.18%) and recall (86.08%), indicating a balanced ability to identify both active and inactive compounds. However, its performance was slightly lower than that of the ensemble methods, suggesting that while SVM is effective at handling complex decision boundaries, it may be less capable of capturing the dataset's full variability than tree-based ensembles.

Logistic Regression, used as a baseline linear model, achieved the lowest performance among the four models, with an accuracy of 84.56% and an F1-score of 85.89%. Although its recall was relatively high (88.61%), its lower precision (83.33%) indicates a higher rate of false positive predictions. This suggests that linear models may not fully capture the non-linear relationships between molecular descriptors and biological activity, which are better handled by more complex algorithms.

Figure 3 presents the confusion matrices for the four machine learning models, providing a detailed view of their classification performance in terms of true positives, true negatives, false positives, and false negatives. The Random Forest model shows the best overall performance, correctly classifying 71 active and 62 inactive compounds, with only 8 false positives and 8 false negatives. This balanced error distribution indicates strong reliability in distinguishing between the two classes. Gradient Boosting demonstrates similarly strong performance, with 70 true positives and 61 true negatives, but it has slightly higher misclassification (9 false positives and 9 false negatives) than Random Forest. The Support Vector Machine model exhibits a higher number of false negatives (14), indicating that some active compounds were misclassified as inactive. This

Figure 3. Confusion matrices of (a) Random Forest, (b) Gradient Boosting, (c) Support Vector Machine, and (d) Logistic Regression models on the test set.

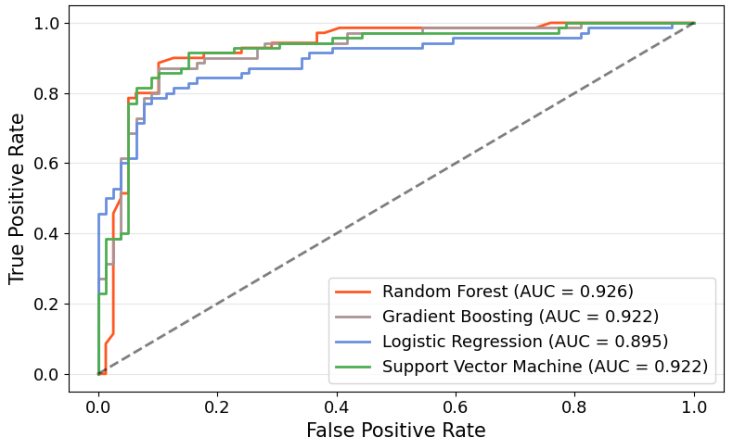

Figure 4. ROC curves of all models showing classification performance with corresponding AUC values.

suggests slightly reduced sensitivity despite reasonable overall performance. Logistic Regression has the highest number of false positives (11), indicating that more inactive compounds were incorrectly predicted as active. While it maintains moderate performance, it is less balanced compared to the other models.

Figure 4 presents the receiver operating characteristic (ROC) curves for all evaluated models, illustrating their ability to discriminate between active and inactive compounds across different classification thresholds. Among the models, Random Forest achieved the highest performance with an AUC of 0.926, indicating excellent discriminative capability. Gradient Boosting and Support Vector Machine showed comparable performance, with AUCs of 0.922, suggesting strong and consistent classification ability. Logistic Regression, while still performing well, had the lowest AUC of 0.895, reflecting relatively weaker discrimination compared to the other models.

Figure 5 shows the precision–recall (PR) curves for all models, providing additional insight into their performance, particularly regarding the balance between precision and recall. Random Forest maintains consistently high precision across a wide range of recall values, indicating strong performance in correctly identifying active compounds while minimizing false positives. Gradient Boosting and Support Vector Machine exhibit similar trends, with relatively stable precision at moderate to high recall levels. In contrast, Logistic Regression shows a more noticeable decline in precision as recall increases, suggesting reduced reliability when attempting to capture a higher proportion of active compounds.

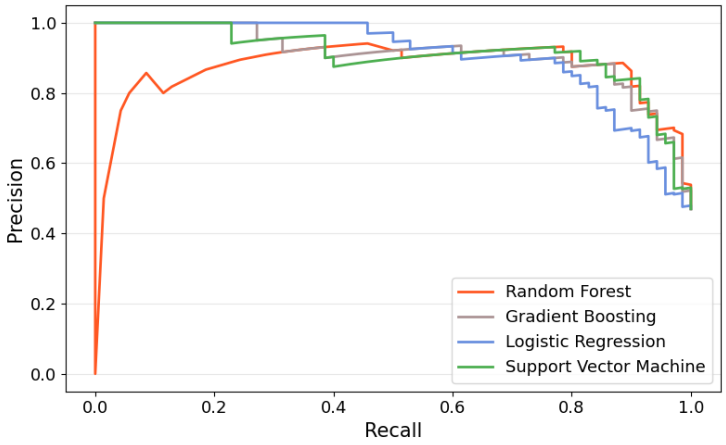

Figure 5. Precision–recall curves of all models illustrating the trade-off between precision and recall.

The results demonstrate that machine learning–based QSAR models can effectively predict the activity of ADRB2 ligands. Ensemble methods, particularly Random Forest and Gradient Boosting, consistently outperformed Logistic Regression and Support Vector Machine across multiple evaluation metrics. This highlights the importance of capturing non-linear relationships between molecular descriptors and compound activity, which are often present in complex pharmacological systems. Additionally, the analysis of molecular descriptors indicates that nHD is statistically significantly associated with variation in activity. In contrast, other descriptors, such as LogP and nHA, exhibit weaker or non-significant contributions.

Despite these promising results, several limitations should be considered. First, the study relies solely on 2D molecular descriptors, which may not fully capture the three-dimensional conformational and electronic properties of molecules that are critical for receptor binding. Second, default hyperparameters were used for all models, which may limit the optimization of model performance. Third, although relatively balanced, the dataset is limited to compounds available in the ChEMBL database and may not fully represent the chemical space of ADRB2-targeting compounds. Additionally, potential experimental variability in IC_{50} values could introduce noise into the dataset, affecting model accuracy.

Future work should focus on enhancing model performance and generalizability. Incorporating 3D descriptors, molecular fingerprints, or graph-based representations could provide more comprehensive structural information. Hyperparameter optimization and the application of advanced techniques such as deep learning or ensemble stacking may further improve predictive accuracy. Expanding the dataset with additional high-quality experimental data and performing external validation on independent datasets would also strengthen model robustness. Furthermore, integrating molecular docking or molecular dynamics simulations could provide mechanistic insights and support the interpretation of QSAR findings, ultimately aiding in the rational design of more potent compounds targeting ADRB2.

4. Conclusions

In this study, machine learning–based QSAR models were successfully developed to predict the activity of ADRB2 ligands using molecular descriptor data. Among the evaluated models, Random Forest achieved the highest performance, with an accuracy of 89.26%, an F1-score of 89.87%, and an AUC of 0.926, followed closely by Gradient Boosting with an accuracy of 87.92% and an AUC of 0.922. These results demonstrate the strong capability of ensemble methods to capture complex structure–activity relationships within the dataset. The analysis of physicochemical descriptors indicated that nHD is statistically significantly associated with variation in compound activity toward ADRB2. In contrast, other descriptors, such as LogP and nHA, exhibit weaker or non-significant contributions. Overall, the findings demonstrate that molecular descriptor–based machine learning models, particularly ensemble approaches, provide an effective framework for predicting ADRB2-related compound activity and for prioritizing candidate molecules in computational drug discovery.

Author Contributions: Conceptualization, T.R.N. and R.I.; methodology, M.P. and R.I.; software, T.R.N.; validation, R.I.; formal analysis, T.R.N. and M.P.; investigation, R.I.; resources, R.I.; data curation, M.P. and R.I.; writing—original draft preparation, T.R.N. and R.I.; writing—review and editing, R.I.; visualization, T.R.N.; 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 that support the findings of this study are available from the corresponding author upon reasonable request.

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

References

1. Yang, A., Yu, G., Wu, Y., and Wang, H. (2021). Role of B2-Adrenergic Receptors in Chronic Obstructive Pulmonary Disease, *Life Sciences*, Vol. 265, 118864. doi:10.1016/j.lfs.2020.118864.
2. Ali, D. C., Naveed, M., Gordon, A., Majeed, F., Saeed, M., Ogbuke, M. I., Atif, M., Zubair, H. M., and Changxing, L. (2020). β-Adrenergic Receptor, an Essential Target in Cardiovascular Diseases, *Heart Failure Reviews*, Vol. 25, No. 2, 343–354. doi:10.1007/s10741-019-09825-x.
3. Cho, Y. Y., Kim, S., Kim, P., Jo, M. J., Park, S.-E., Choi, Y., Jung, S. M., and Kang, H. J. (2025). G-Protein-Coupled Receptor (GPCR) Signaling and Pharmacology in Metabolism: Physiology, Mechanisms, and Therapeutic Potential, *Biomolecules*, Vol. 15, No. 2, 291. doi:10.3390/biom15020291.
4. Chandrasekaran, B., Samarneh, S., Jaber, A. M. Y., Kassab, G., and Agrawal, N. (2019). Therapeutic Potentials of A2B Adenosine Receptor Ligands: Current Status and Perspectives, *Current Pharmaceutical Design*, Vol. 25, No. 25, 2741–2771. doi:10.2174/1381612825666190717105834.
5. Xing, G., Woo, A. Y.-H., Pan, L., Lin, B., and Cheng, M.-S. (2020). Recent Advances in β 2 -Agonists for Treatment of Chronic Respiratory Diseases and Heart Failure, *Journal of Medicinal Chemistry*, Vol. 63, No. 24, 15218–15242. doi:10.1021/acs.jmedchem.0c01195.
6. de Azevedo, D. Q., Campioni, B. M., Pedroz Lima, F. A., Medina-Franco, J. L., Castilho, R. O., and Maltarollo, V. G. (2024). A Critical Assessment of Bioactive Compounds Databases, *Future Medicinal Chemistry*, Vol. 16, No. 10, 1029–1051. doi:10.1080/17568919.2024.2342203.
7. Khan, S., Sarfraz, A., Prakash, O., and Khan, F. (2024). Machine Learning-Based QSAR Modeling, Molecular Docking, Dynamics Simulation Studies for Cytotoxicity Prediction in MDA-MB231 Triple-Negative Breast Cancer Cell Line, *Journal of Molecular Structure*, Vol. 1315, 138807. doi:10.1016/j.molstruc.2024.138807.
8. Kalyankar, D. S., Bhagat, C. G., Kadu, A. D., Tambade, A. P., and Dhoran, K. S. (2024). AI-Driven Insights: Paving the Path to Next-Generation Therapeutics, *International Journal of Advanced Research in Science, Communication and Technology*, 372–378. doi:10.48175/IJARSCT-22854.
9. Dhudum, R., Ganeshpurkar, A., and Pawar, A. (2024). Revolutionizing Drug Discovery: A Comprehensive Review of AI Applications, *Drugs and Drug Candidates*, Vol. 3, No. 1, 148–171. doi:10.3390/ddc3010009.
10. 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.
11. Noviandy, T. R., Setiawan, R., Sufri, R., and Anisah, A. (2025). Optimizing Machine Learning Models for Cardiovascular Drug Candidate Screening with XGBoost and Successive Halving, *Open International Journal of Informatics*, Vol. 13, No. 2, 41–50. doi:10.11113/oiji2025.13n2.339.
12. Noviandy, T. R., Idroes, G. M., Patwekar, M., and Idroes, R. (2025). Fine-Tuning ChemBERTa for Predicting Activity of AXL Kinase Inhibitors in Oncogenic Target Modeling, *Grimsa Journal of Science Engineering and Technology*, Vol. 3, No. 2, 73–84. doi:10.61975/gjset.v3i2.98.
13. 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.
14. Amato, A., and Di Lecce, V. (2023). Data Preprocessing Impact on Machine Learning Algorithm Performance, *Open Computer Science*, Vol. 13, No. 1. doi:10.1515/comp-2022-0278.
15. Mauri, A., Consonni, V., and Todeschini, R. (2017). Molecular Descriptors, *Handbook of Computational Chemistry*, Springer International Publishing, Cham, 2065–2093. doi:10.1007/978-3-319-27282-5_51.
16. Dehmer, M., Varmuza, K., and Bonchev, D. (Eds.). (2012). *Statistical Modelling of Molecular Descriptors in QSAR/QSPR*, Wiley. doi:10.1002/9783527645121.
17. 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.
18. Noviandy, T. R., Maulana, A., Emran, T. B., Idroes, G. M., and Idroes, R. (2023). QSAR Classification of Beta-Secretase 1 Inhibitor Activity in Alzheimer's Disease Using Ensemble Machine Learning Algorithms, *Heca Journal of Applied Sciences*, Vol. 1, No. 1, 1–7. doi:10.60084/hjas.v1i1.12.
19. Goodarzi, M., Dejaegher, B., and Heyden, Y. Vander. (2012). Feature Selection Methods in QSAR Studies, *Journal of AOAC INTERNATIONAL*, Vol. 95, No. 3, 636–651. doi:10.5740/jaoacint.SGE_Goodarzi.
20. Simeon, S., Anuwongcharoen, N., Shoombuatong, W., Malik, A. A., Prachayasittikul, V., Wikberg, J. E. S., and Nantasenamat, C. (2016). Probing the Origins of Human Acetylcholinesterase Inhibition via QSAR Modeling and Molecular Docking, *PeerJ*, Vol. 4, e2322.
21. Joseph, V. R. (2022). Optimal Ratio for Data Splitting, *Statistical Analysis and Data Mining: The ASA Data Science Journal*, Vol. 15, No. 4, 531–538. doi:10.1002/sam.11583.
22. Chen, X., Li, H., Tian, L., Li, Q., Luo, J., and Zhang, Y. (2020). Analysis of the Physicochemical Properties of Acaricides Based on Lipinski's Rule of Five, *Journal of Computational Biology*, Vol. 27, No. 9, 1397–1406. doi:10.1089/cmb.2019.0323.
23. 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.
24. Suhendra, R., Suryadi, S., Husdayanti, N., Maulana, A., Noviandy, T. R., Sasmita, N. R., Subianto, M., Earlia, N., Niode, N. J., and Idroes, R. (2023). Evaluation of Gradient Boosted Classifier in Atopic Dermatitis Severity Score Classification, *Heca Journal of Applied Sciences*, Vol. 1, No. 2, 54–61. doi:10.60084/hjas.v1i2.85.
25. Wu, Z., Zhu, M., Kang, Y., Leung, E. L.-H., Lei, T., Shen, C., Jiang, D., Wang, Z., Cao, D., and Hou, T. (2021). Do We Need Different Machine Learning Algorithms for QSAR Modeling? A Comprehensive Assessment of 16 Machine Learning Algorithms on 14 QSAR Data Sets, *Briefings in Bioinformatics*, Vol. 22, No. 4, bbaa321. doi:10.1093/bib/bbaa321.
26. Roy, K., and Kar, S. (2015). How to Judge Predictive Quality of Classification and Regression Based QSAR Models?, Z. Ul-Haq; J. D. B. T.-F. in C. C. Madura (Eds.), , Bentham Science Publishers, 71–120. doi:10.1016/B978-1-60805-979-9.50003-2.