International Journal of Medical and Pharmaceutical Sciences
Open access peer-reviewed Journal
Volume 2, Issue 9, Page 335-344.
https://doi.org/10.5281/zenodo.22795105

Research Article

Machine Learning-Based Prediction of Drug-Induced Hepatotoxicity Using Molecular Descriptors: A QSAR Approach

Mukthiyar Ahamed*, Ramsha Thabassum, Inayath Jan N., Misba Aliya

Department of Pharmacology, Shantha College of Pharmacy, Peresandra, Chikkaballapur-562104, Karnataka, India

Background: Drug-induced hepatotoxicity is one of the leading causes of drug withdrawal and treatment discontinuation during drug development and clinical practice. Early prediction of hepatotoxicity using computational approaches can minimize adverse drug reactions and reduce the cost and time associated with experimental toxicity studies. Machine learning techniques have emerged as effective tools for predicting toxicity based on molecular descriptors. Objective: The present study aimed to develop and compare machine learning models for predicting hepatotoxic drugs using molecular descriptors generated from chemical structures and analyzed with Orange Data Mining software. Materials and Methods: A dataset consisting of 99 drugs, including hepatotoxic and non-hepatotoxic compounds, was collected from publicly available databases. Canonical SMILES were obtained from PubChem and molecular descriptors were generated using PaDEL-Descriptor software. Initially, 1,876 molecular descriptors were generated for each compound. Missing values were handled through data preprocessing, and feature selection was performed using the Rank widget in Orange Data Mining. The top 20 ranked descriptors were selected for model development. Four machine learning algorithms, namely Random Forest, Logistic Regression, Naive Bayes, and Decision Tree, were trained and evaluated using stratified five-fold cross-validation. Model performance was assessed using the area under the receiver operating characteristic curve (AUC), classification accuracy, precision, recall, F1-score, and Matthew's correlation coefficient (MCC). Results: Among the evaluated models, Random Forest demonstrated the best predictive performance with an AUC of 0.840, classification accuracy of 73.7%, F1-score of 0.737, precision of 0.738, recall of 0.737, and MCC of 0.475. Naive Bayes ranked second with an AUC of 0.800, followed by Decision Tree (0.764) and Logistic Regression (0.608). Confusion matrix analysis further confirmed that Random Forest correctly classified the highest number of hepatotoxic and non-hepatotoxic drugs. Conclusion: The findings suggest that Random Forest is an effective machine learning approach for predicting hepatotoxicity using molecular descriptors. This computational strategy may support early toxicity assessment during drug discovery and contribute to safer drug development.

Keywords: Machine Learning; Hepatotoxicity; Molecular Descriptors; Random Forest; Orange Data Mining; Drug-Induced Liver Injury; PaDEL-Descriptor.

INTRODUCTION

Drug-induced liver injury (DILI), commonly referred to as hepatotoxicity, is one of the major safety concerns encountered during drug development and post-marketing surveillance. The liver plays a central role in drug metabolism, detoxification, and elimination, making it particularly susceptible to toxic effects caused by pharmaceutical compounds. Hepatotoxicity can range from mild, transient elevations in liver enzymes to severe liver failure requiring transplantation or resulting in death. It is recognized as one of the leading causes of drug withdrawal from the pharmaceutical market and remains a significant challenge for regulatory authorities and healthcare professionals. [1,2]

Conventional methods for evaluating hepatotoxicity primarily rely on in vitro cell culture experiments, in vivo animal studies, and clinical trials. Although these methods provide valuable information regarding drug safety, they are often expensive, labor-intensive, and time-consuming. [3] Furthermore, animal models may not always accurately predict human liver toxicity because of interspecies differences in drug metabolism and physiological responses. Consequently, there is an increasing demand for computational approaches capable of predicting hepatotoxicity during the early stages of drug discovery. [1,4] Recent advances in artificial intelligence (AI) and machine learning (ML) have transformed pharmaceutical research by enabling rapid analysis of complex biological and chemical datasets. Machine learning algorithms can identify hidden relationships between molecular properties and biological activities, allowing accurate prediction of toxicity before experimental validation. These computational models have become valuable tools in cheminformatics, quantitative structure–activity relationship (QSAR) studies, virtual screening, adverse drug reaction prediction, and drug safety assessment. [5] Molecular descriptors are numerical values that represent the structural, physicochemical, topological, and electronic properties of chemical compounds. They provide comprehensive information about molecular architecture and are widely used as input variables in machine learning models. [2] Descriptor calculation software such as PaDEL-Descriptor can automatically generate thousands of molecular descriptors from the chemical structure of a compound represented by its Simplified Molecular Input Line Entry System (SMILES). Feature selection techniques are then employed to identify the most informative descriptors, thereby reducing dimensionality and improving prediction accuracy. [6] Several machine learning algorithms have been successfully applied for toxicity prediction, including Random Forest, Logistic Regression, Naive Bayes, Support Vector Machine, Artificial Neural Networks, and Decision Trees. Among these algorithms, Random Forest has gained considerable attention because of its robustness, ability to handle high-dimensional datasets, resistance to overfitting, and excellent predictive performance. Logistic Regression provides a simple and interpretable statistical approach for binary classification, whereas Naive Bayes offers efficient probabilistic classification based on Bayes' theorem. Decision Trees are widely appreciated for their easy interpretation and visualization of classification rules. [7,8] Orange Data Mining is an open-source visual programming platform that enables users to perform machine learning and data analysis without extensive programming knowledge. The software provides an intuitive workflow environment for data preprocessing, feature selection, model training, cross-validation, and performance evaluation. Its graphical interface makes it particularly suitable for pharmaceutical researchers seeking to develop predictive models efficiently while maintaining reproducibility. [9] In the present study, molecular descriptors were generated for a dataset of hepatotoxic and non-hepatotoxic drugs using PaDEL-Descriptor software. Following preprocessing and feature selection, four supervised machine learning algorithms, namely Random Forest, Logistic Regression, Naive Bayes, and Decision Tree, were developed and evaluated using Orange Data Mining software.[10,11] Model performance was assessed using five-fold stratified cross-validation and multiple evaluation metrics, including the area under the receiver operating characteristic curve (AUC), classification accuracy, precision, recall, F1-score, Matthews correlation coefficient (MCC), receiver operating characteristic (ROC) analysis, and confusion matrix. The primary objective was to identify the most effective machine learning algorithm for predicting hepatotoxicity and to demonstrate the potential of computational methods as reliable tools for early toxicity assessment in drug discovery. [12]

2. MATERIALS AND METHODS

2.1 Study Design

The present study employed a computational machine learning approach to predict hepatotoxicity using molecular descriptors derived from chemical structures. The overall workflow consisted of dataset collection, molecular descriptor generation, data preprocessing, feature selection, machine learning model development, and performance evaluation. [13] The workflow is illustrated in Figure 1.

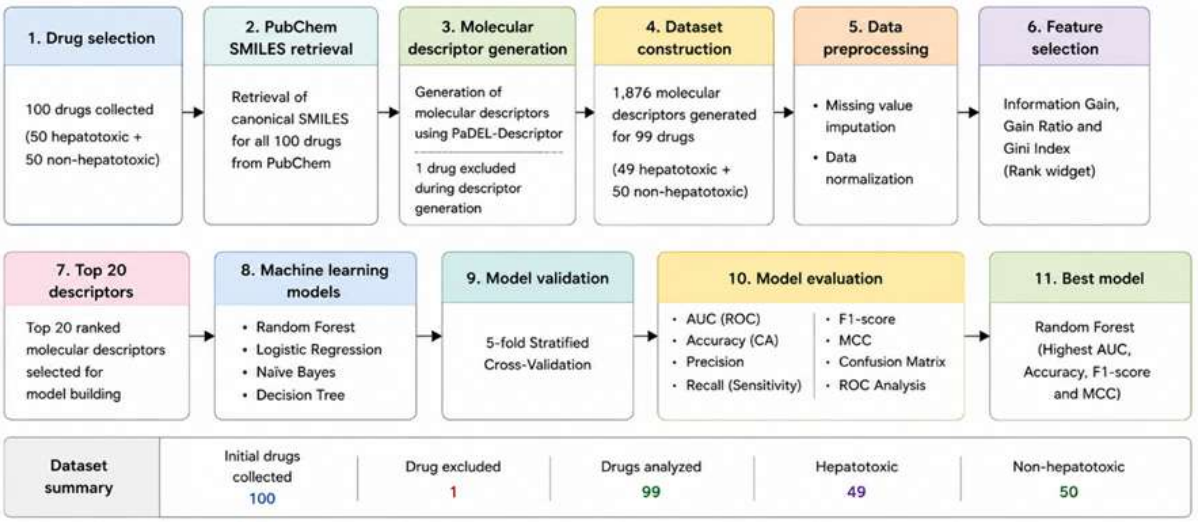

Figure 1. Overall workflow of the machine learning-based approach used for prediction of drug-induced hepatotoxicity

2.2 Dataset Collection

A total of 100 approved drugs were initially collected from publicly available databases. The dataset consisted of both hepatotoxic and non-hepatotoxic compounds. Drug names, PubChem Compound Identifiers (CID), canonical SMILES, and class labels were compiled into a structured dataset. During descriptor generation, one compound (Insulin) could not be processed because molecular descriptor calculation is primarily designed for small molecules. Consequently, the final dataset contained 99 drugs, including 50 hepatotoxic and 49 non-hepatotoxic compounds. [14]

The binary class labels were assigned as follows:

- 1 = Hepatotoxic drug
- 0 = Non-hepatotoxic drug

2.3 Collection of Molecular Structures

Canonical SMILES for each drug were retrieved from the PubChem database. The SMILES notation was used as the molecular representation for descriptor calculation because it provides a standardized and machine-readable description of chemical structures.

2.4 Molecular Descriptor Generation

Molecular descriptors were generated using PaDEL-Descriptor software. The canonical SMILES file containing all selected compounds was imported into the software, and molecular descriptor calculation was performed using the default descriptor settings. [15]

Initially,

- 1,876 molecular descriptors

were generated for each compound, representing various molecular properties including:

- Constitutional descriptors
- Topological descriptors
- Electronic descriptors
- Geometrical descriptors
- Hybrid descriptors
- Physicochemical descriptors

The generated descriptor dataset was exported in CSV format for subsequent machine learning analysis.

2.5 Data Preprocessing

The descriptor dataset was imported into Orange Data Mining software (Version 3.39) for preprocessing. [8]

The following preprocessing steps were performed:

- Missing values were handled using the Impute widget.
- Descriptor columns containing invalid values (Infinity) were corrected before model development.
- Drug names were assigned as Meta attributes.
- The hepatotoxicity class label was assigned as the Target variable.

These preprocessing steps ensured that the dataset was suitable for supervised machine learning.

2.6 Feature Selection

The descriptor dataset contained 1,876 molecular descriptors, many of which contributed little to classification performance. Feature selection was therefore performed using the Rank widget available in Orange Data Mining.

The descriptors were ranked according to:

- Information Gain
- Information Gain Ratio
- Gini Decrease

The top 20 ranked molecular descriptors were selected for model development. Feature selection reduced dataset dimensionality, minimized redundant information, and improved model interpretability while maintaining predictive performance.

2.7 Machine Learning Model Development

Four supervised machine learning algorithms were employed for hepatotoxicity prediction: [30]

- Random Forest
- Logistic Regression
- Naive Bayes
- Decision Tree

All models were implemented using the default settings available in Orange Data Mining. The selected descriptors were used as predictor variables, whereas the hepatotoxicity class label served as the response variable. [16,17]

2.8 Model Validation

Model performance was evaluated using 5-fold stratified cross-validation. During validation, the dataset was divided into five approximately equal subsets.

In each iteration,

- four subsets were used for model training,
- one subset was used for testing.

This process was repeated until every subset had served as the testing dataset. Cross-validation reduces bias and provides a reliable estimate of model performance.

2.9 Performance Evaluation

Model performance was assessed using the following statistical parameters:

- Area Under the Receiver Operating Characteristic Curve (AUC)
- Classification Accuracy (CA)
- Precision
- Recall (Sensitivity)
- F1-score
- Matthews Correlation Coefficient (MCC)

Additionally,

- Receiver Operating Characteristic (ROC) curves
- Confusion Matrices

were generated to visualize classification performance and compare the predictive ability of different machine learning algorithms. [30]

2.10 Software Used

Software and Purpose

PubChem- Retrieval of molecular structural and canonical SMILES

PaDEL- Descriptor Molecular descriptor generation

Microsoft Excel Dataset preparation and preprocessing

Orange Data Mining (Version 3.39)- Feature selection, model development, and performance evaluation. [8]

3. RESULTS

3.1 Dataset Preparation and Molecular Descriptor Generation

A total of 100 drugs were initially collected for hepatotoxicity prediction. Canonical SMILES were retrieved from the PubChem database and molecular descriptors were generated using PaDEL-Descriptor software. During descriptor calculation, one compound (Insulin) could not be processed because molecular descriptor generation is intended for small molecules. Consequently, the final dataset consisted of 99 drugs, including 50 hepatotoxic and 49 non-hepatotoxic compounds. Initially, 1,876 molecular descriptors were generated for each compound. The dataset was preprocessed by correcting invalid values and handling missing data before model development.

3.2 Feature Selection

Feature ranking was performed using the Rank widget in Orange Data Mining. Information Gain, Information Gain Ratio, and Gini Decrease were used to evaluate descriptor importance. The top 20 molecular descriptors were selected for model development to reduce dimensionality and eliminate redundant variables. The selected descriptors were used as input features for all machine learning classifiers.

		#	Info. gain	Gain ratio	Gini
1	RDF105s		0.302	0.151	0.178
2	RDF115s		0.272	0.136	0.162
3	RDF105p		0.272	0.136	0.162
4	RDF105v		0.272	0.136	0.162
5	RDF110v		0.266	0.133	0.162
6	RDF110m		0.266	0.133	0.162
7	RDF110s		0.255	0.128	0.155
8	RDF10u		0.254	0.127	0.150
9	RDF10i		0.230	0.115	0.137
10	RDF115p		0.225	0.112	0.140
11	RDF115v		0.225	0.112	0.140
12	RDF105m		0.225	0.112	0.140
13	nAtomLC		0.219	0.110	0.137
14	RDF20p		0.219	0.109	0.131
15	SP-3		0.218	0.109	0.135
16	RDF10e		0.215	0.107	0.134
17	RDF10m		0.215	0.107	0.134
18	RDF120s		0.213	0.107	0.133
19	RDF20i		0.213	0.106	0.126
20	RDF140m		0.212	0.117	0.134

Figure 2. Ranking of the top 20 molecular descriptors selected using the Rank widget based on information gain, gain ratio, and Gini decrease.

3.3 Performance of Machine Learning Models

Four supervised machine learning algorithms were evaluated using 5-fold stratified cross-validation. Model performance was assessed using AUC, classification accuracy, precision, recall, F1-score, and Matthews Correlation Coefficient (MCC).

Table 2. Performance comparison of machine learning models

Model	AUC	Accuracy	Precision	Recall	F1 score	MCC
Random Forest	0.840	73.7%	0.738	0.737	0.737	0.475
Naïve bayes	0.800	72.7%	0.728	0.727	0.727	0.456
Decision tree	0.764	68.7%	0.688	0.687	0.686	0.375
Logistic Regression	0.608	62.6%	0.627	0.626	0.625	0.253

Among the evaluated classifiers, Random Forest achieved the highest predictive performance with an AUC of 0.840, classification accuracy of 73.7%, precision of 0.738, recall of 0.737, F1-score of 0.737, and MCC of 0.475. Naive Bayes demonstrated the second-best performance, whereas Logistic Regression produced the lowest predictive accuracy.

3.4 Confusion Matrix Analysis

The confusion matrices provided additional insight into the classification performance of each algorithm.

Table 3. Confusion matrix summary

Model	TN	FP	FN	TP
Random Forest	35	14	12	38
Naïve Bayes	37	12	15	35
Decision Tree	31	18	13	37
Logistic Regression	28	21	16	34

The Random Forest classifier correctly classified 38 hepatotoxic drugs and 35 non-hepatotoxic drugs, with relatively fewer false predictions compared with the other models. Logistic Regression exhibited the highest number of misclassifications, while Naive Bayes showed balanced performance with fewer false positives but slightly more false negatives.

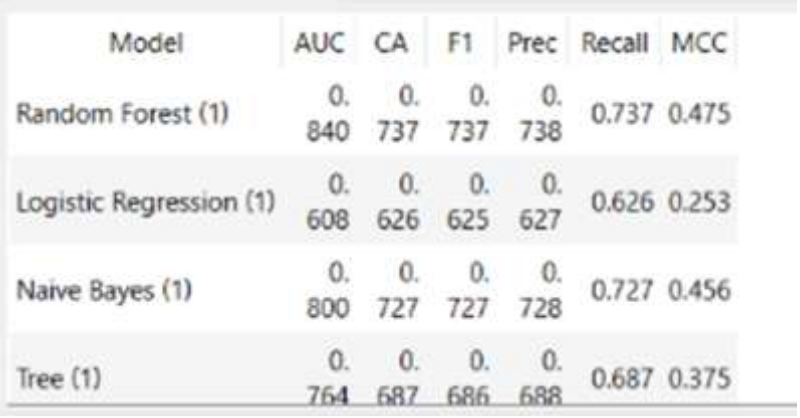

Figure 3. Confusion matrices of the four machine learning models.

3.5 Receiver Operating Characteristic (ROC) Analysis

Receiver Operating Characteristic (ROC) analysis was performed to compare the discrimination ability of the developed machine learning models.

The Random Forest classifier produced the highest ROC curve with an AUC of 0.840, indicating superior discrimination between hepatotoxic and non-hepatotoxic drugs. Naive Bayes also demonstrated good predictive capability with an AUC of 0.800, whereas Decision Tree and Logistic Regression showed comparatively lower performance.

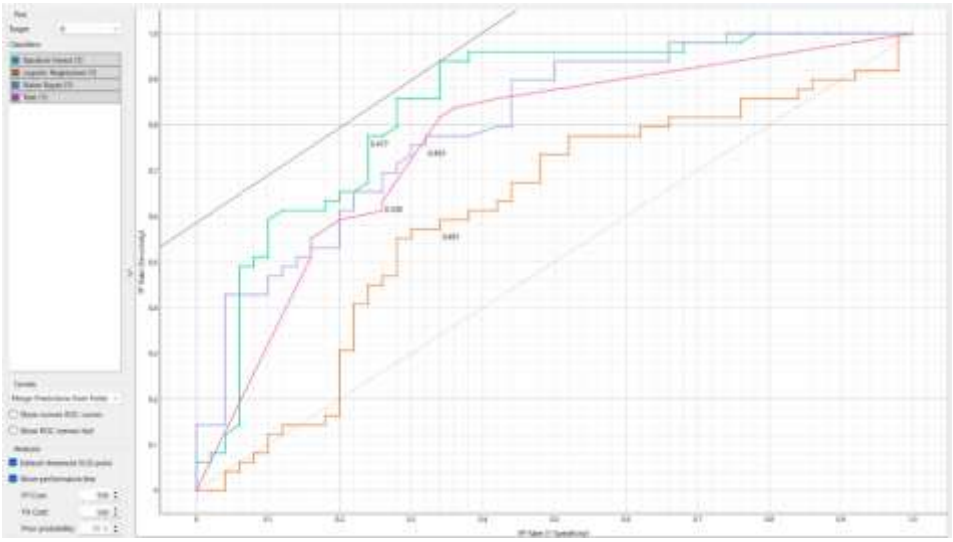

Figure 4. ROC curves of Random Forest, Naive Bayes, Decision Tree, and Logistic Regression models.

3.6 Overall Model Comparison

The comparative analysis demonstrated that Random Forest consistently outperformed the other classifiers across all evaluation metrics. Its higher AUC, classification accuracy, F1-score, and MCC indicate that the algorithm effectively captured the complex relationships among the selected molecular descriptors. Naive Bayes achieved performance comparable to Random Forest, suggesting that probabilistic classification also provides reliable hepatotoxicity prediction. Decision Tree exhibited moderate predictive ability, whereas Logistic Regression showed the lowest overall performance, indicating that linear classification may be less suitable for the present dataset. Overall, the findings indicate that Random Forest is the most effective machine learning algorithm for hepatotoxicity prediction using the selected molecular descriptors.

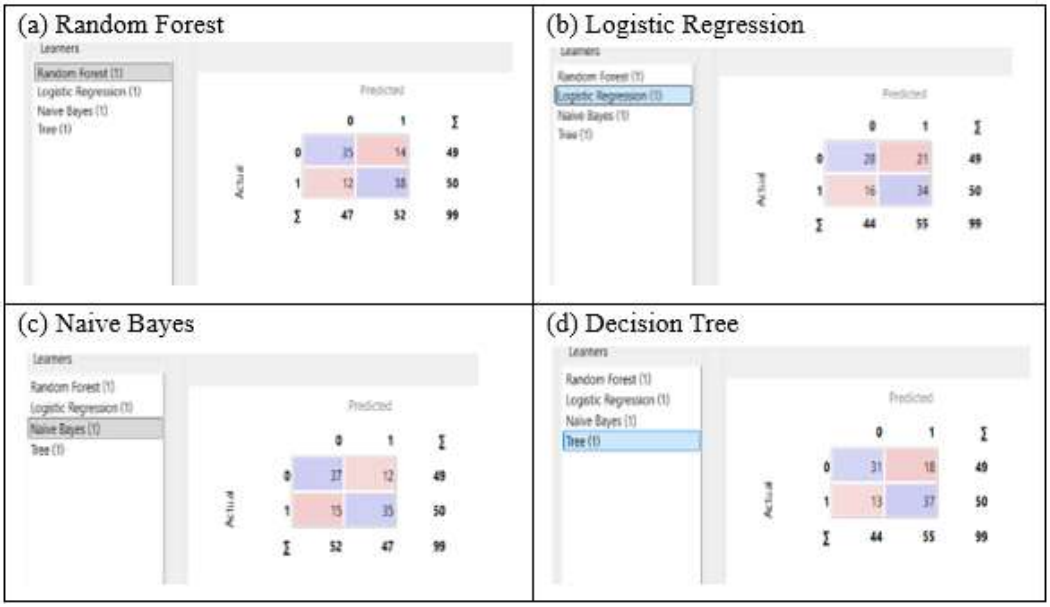

Figure 5. Confusion matrices of the machine learning classifiers developed using the top 20 selected molecular descriptors: (a) Random Forest, (b) Logistic Regression, (c) Naive Bayes, and (d) Decision Tree.

DISCUSSION

Drug-induced hepatotoxicity remains one of the most important causes of adverse drug reactions, frequently

leading to treatment discontinuation, regulatory restrictions, and withdrawal of drugs from the pharmaceutical market. Consequently, there is increasing interest in computational approaches that can identify potentially hepatotoxic compounds before expensive laboratory or clinical testing. In the present study, machine learning models were developed using molecular descriptors to predict hepatotoxicity, and their performance was systematically compared. [21,22] Among the four evaluated algorithms, Random Forest demonstrated the best overall performance, achieving an AUC of 0.840, classification accuracy of 73.7%, F1-score of 0.737, and MCC of 0.475. These findings suggest that Random Forest was more effective at distinguishing hepatotoxic from non-hepatotoxic compounds than Logistic Regression, Naive Bayes, and Decision Tree. The superior performance of Random Forest can be attributed to its ensemble learning strategy. Unlike single decision trees, Random Forest constructs multiple decision trees from randomly selected subsets of data and combines their predictions. This approach reduces overfitting, improves generalization, and enables the model to capture complex, non-linear relationships between molecular descriptors and hepatotoxicity. Because molecular descriptors often describe diverse structural and physicochemical characteristics, ensemble-based methods are well suited for modeling these interactions. Naive Bayes achieved the second-highest predictive performance with an AUC of 0.800 and an accuracy of 72.7%. Although Naive Bayes assumes that predictor variables are conditionally independent, it still produced competitive results. This suggests that the selected molecular descriptors retained useful information for distinguishing hepatotoxic compounds despite possible correlations among descriptors. Decision Tree produced moderate predictive performance with an AUC of 0.764 and an accuracy of 68.7%. Decision trees are easy to interpret and visualize; however, they are more susceptible to overfitting and instability when trained on relatively small datasets. These characteristics may explain their lower performance compared with Random Forest. Logistic Regression showed the lowest predictive performance, with an AUC of 0.608 and an accuracy of 62.6%. Logistic Regression models linear relationships between predictor variables and outcomes. Because hepatotoxicity is influenced by multiple interacting molecular features, a simple linear classifier may not adequately capture the complexity of these relationships. Feature selection played an important role in this study. Initially, 1,876 molecular descriptors were generated using PaDEL-Descriptor software. After ranking the descriptors using Information Gain, Information Gain Ratio, and Gini Decrease, the top 20 descriptors were selected for model development. Reducing the number of descriptors decreased redundancy and simplified the dataset while maintaining good predictive performance. Feature selection is particularly valuable in cheminformatics because high-dimensional datasets often contain descriptors that contribute little to classification. [23] The confusion matrix analysis provided further insight into model performance. Random Forest correctly identified 38 hepatotoxic and 35 non-hepatotoxic drugs while producing comparatively fewer false classifications than the other algorithms. These findings are consistent with the ROC analysis, in which Random Forest produced the largest area under the curve, indicating superior discrimination between the two classes. [25,26] The findings of this study are generally consistent with previous reports demonstrating that ensemble learning methods, particularly Random Forest, perform well in toxicity prediction and quantitative structure–activity relationship (QSAR) modeling. Random Forest has been widely applied in computational toxicology because of its robustness, ability to process high-dimensional molecular descriptor datasets, and resistance to overfitting. [28,29] Despite these encouraging results, several limitations should be acknowledged. First, the study included a relatively small dataset of 99 drugs, which may limit the generalizability of the developed models. Second, only four machine learning algorithms were evaluated. Additional algorithms such as Support Vector Machine, Extreme Gradient Boosting (XGBoost), LightGBM, Artificial Neural Networks, or Deep Learning models could be investigated in future studies. Third, the present work relied exclusively on molecular descriptors derived from chemical structures. Integrating additional information such as molecular fingerprints, biological targets, gene expression profiles, or pharmacokinetic properties may further improve predictive performance. Overall, the results demonstrate that machine learning combined with molecular

descriptors provides an effective computational strategy for hepatotoxicity prediction. Such approaches can support early-stage drug discovery by identifying potentially hepatotoxic compounds before extensive experimental evaluation, thereby reducing development costs and improving drug safety. [26,27]

CONCLUSION

The present study demonstrated the potential of machine learning techniques for predicting drug-induced hepatotoxicity using molecular descriptors. A dataset comprising 99 drugs was analyzed following molecular descriptor generation with PaDEL-Descriptor and feature selection using Orange Data Mining. Four machine learning algorithms-Random Forest, Logistic Regression, Naive Bayes, and Decision Tree-were evaluated using five-fold cross-validation. Among the evaluated models, Random Forest exhibited the best overall performance, achieving an AUC of 0.840, classification accuracy of 73.7%, F1-score of 0.737, and MCC of 0.475. These findings indicate that Random Forest was more effective than Logistic Regression, Naive Bayes, and Decision Tree in distinguishing hepatotoxic from non-hepatotoxic drugs within the present dataset. The study also demonstrated that feature selection effectively reduced the dimensionality of the molecular descriptor dataset while maintaining strong predictive performance. The selected descriptors captured important structural and physicochemical characteristics associated with hepatotoxicity and contributed to improved model development. Overall, the proposed computational workflow provides a practical approach for the early prediction of hepatotoxicity during drug discovery. Although further validation using larger and more diverse datasets is required, the findings suggest that machine learning combined with molecular descriptors may support toxicity assessment and facilitate the identification of safer drug candidates.

Declarations

Availability of Data

The dataset generated and analyzed during the present study is available from the corresponding author upon reasonable request.

Funding

The authors received no external funding for this study.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Limitations

- Relatively small dataset (99 drugs).
- Only four machine learning algorithms were evaluated.
- External validation using an independent dataset was not performed.

Descriptor-based models do not account for biological or clinical factors that may influence hepatotoxicity.

REFERENCES

1. Yap CW. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. J Comput Chem. 2011;32(7):1466–1474.
2. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32.
3. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random forest: A classification and regression tool for compound classification and QSAR modeling. J Chem Inf Comput Sci. 2003;43(6):1947–1958.
4. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273–297.
5. Friedman JH. Greedy function approximation: A gradient boosting machine. Ann Stat. 2001;29(5):1189–1232.
6. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. Proc 22nd ACM SIGKDD Int Conf Knowl Discov Data Min. 2016:785–794.
7. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: An update. SIGKDD Explor. 2009;11(1):10–18.
8. Demšar J, Curk T, Erjavec A, Gorup Č, Hočevar T, Milutinovič M, et al. Orange: Data mining

toolbox in Python. J Mach Learn Res. 2013; 14:2349–2353.
9. Chen M, Hong H, Fang H, Kelly R, Zhou G, Borlak J, et al. Quantitative structure–activity relationship models for predicting drug-induced liver injury based on FDA-approved drug labeling annotation. Toxicol Sci. 2013;136(1):242–252.
10. Zhu XW, Sedykh A, Zhu H, Tropsha A. Predicting drug-induced liver injury using QSAR approaches. Expert Opin Drug Metab Toxicol. 2014;10(10):1527–1543.
11. Low Y, Uehara T, Minowa Y, Yamada H, Ohno Y, Urushidani T, et al. Predicting drug-induced hepatotoxicity using QSAR and toxicogenomics approaches. Chem Res Toxicol. 2011;24(8):1251–1262.
12. He S, Ye T, Wang R, Zhang C, Sun G, Sun X. The development and application of in silico models for drug-induced liver injury. RSC Adv. 2018; 8:28970–28980.
13. Liu R, Schyman P, Wallqvist A. Computational models using multiple machine learning algorithms for predicting drug hepatotoxicity with the DILIrank dataset. Int J Mol Sci. 2020;21(6):2114.
14. Temple RJ, Himmel MH. Safety of newly approved drugs: Implications for prescribing. JAMA. 2002;287(17):2273–2275.
15. Reuben A, Koch DG, Lee WM. Drug-induced acute liver failure: Results of a U.S. multicenter prospective study. Hepatology. 2010;52(6):2065–2076.
16. Ye H, Nelson LJ, Gómez Del Moral M, Martínez-Naves E, Cubero FJ. Dissecting the molecular pathophysiology of drug-induced liver injury. World J Gastroenterol. 2018;24(13):1373–1385.
17. Moriwaki H, Tian YS, Kawashita N, Takagi T. Mordred: A molecular descriptor calculator. J Cheminform. 2018;10(1):4.
18. Papa E, Sangion A, Taboureau O, Gramatica P. Quantitative prediction of rat hepatotoxicity by molecular structure. Int J Quant Struct Prop Relatsh. 2018;3(2):49–60.
19. Netzeva TI, Worth AP, Aldenberg T, Benigni R, Cronin MTD, Gramatica P, et al. Current status of methods for defining the applicability domain of QSARs. Altern Lab Anim. 2005;33(2):155–173.
20. Sheridan RP, Feuston BP. QSAR models in drug discovery using Random Forest. J Chem Inf Comput Sci. 2003;43(6):1947–1958.
21. Iorga A, Dara L. Cell death in drug-induced liver injury. Adv Pharmacol. 2019; 85:31–74.
22. Albrecht W, Kappenberg F, Brecklinghaus T, Stoeber R, Marchan R, Zhang M, et al. Prediction of human drug-induced liver injury in relation to oral doses and blood concentrations. Arch Toxicol. 2019;93(6):1609–1637.
23. Pramanik S, Roy K. Modeling bioconcentration factor using mechanistically interpretable descriptors computed from PaDEL-Descriptor. Environ Sci Pollut Res Int. 2014;21(4):2955–2965.
24. Arab I, Barakat K. ToxTree: Descriptor-based machine learning models for cardiotoxicity liability predictions. arXiv. 2021.
25. Dahl GE, Jaitly N, Salakhutdinov R. Multi-task neural networks for QSAR predictions. arXiv. 2014.
26. Boulesteix AL, Janitza S, Kruppa J, König IR. Overview of Random Forest methodology and practical guidance with emphasis on computational biology. Wiley Interdiscip Rev Data Min Knowl Discov. 2012;2(6):493–507.
27. Kotsampasakou E, Ecker GF. Predictive models for drug-induced liver injury: Current status and future directions.
28. Tropsha A. Best practices for QSAR model development, validation, and exploitation.
29. OECD. Guidance document on the validation of quantitative structure–activity relationship (QSAR) models.
30. He S, Wang R, Zhang C, Sun X. Machine learning approaches for predicting drug-induced liver injury using molecular descriptors and fingerprints.

Cite: Mukthiyar Ahamed*, Ramsha Thabassum, Inayath Jan N., Misba Aliya, Machine Learning-Based Prediction of Drug-Induced Hepatotoxicity Using Molecular Descriptors: A QSAR Approach, Int. J. Med. Pharm. Sci., 2026, 2 (9), 335-344. https://doi.org/10.5281/zenodo.22795105