ORIGINAL RESEARCH

Predicting Renal Toxicity of Compounds with Deep Learning and Machine Learning Methods

Bitopan Mazumdar[1,2] · Pankaj Kumar Deva Sarma[1] · Hridoy Jyoti Mahanta[2,3]

Abstract

Renal toxicity prediction plays a vital role in drug discovery and clinical practice, as it helps to identify potentially harmful compounds and mitigate adverse effects on the renal system. Compound with inherent renal-toxic potential is one of the major concerns for drug development as it leads to failure in drug discovery. Predicting nephrotoxic probabilities of a compound at an early stage can be effective for reducing the drug failure rate. Hence, it is crucial to develop a mechanism to analyze the renal toxicity of a drug-candidate optimally and quickly. To mitigate the risks associated with renal toxicity, predictive models leveraging machine learning and deep learning techniques have gained significant attention. In this study, 287 human renal-toxic drugs and 278 non-renal-toxic drugs were collected to develop a deep learning model and 27 machine learning models using 8 kinds of fingerprints and Rdkit descriptors. The deep neural network model shows better generalization scores on five-fold cross-validation and Extra-tree model shows better performance score on test data. Structural alerts, specific chemical substructures associated with renal toxicity, offer a valuable tool for early toxicity assessment. Therefore, the substructures of renal toxic compounds were studied by applying association rule mining technique based on frequent itemset patterns. A method has been proposed for generating structural alerts and 10 structural alerts have been generated using the method.

Keywords Renal toxicity · Deep neural network · Machine learning · Drugs · Association rule mining

Introduction

Kidney is one of the major and vital organs of human body and its main function is to filter blood, remove toxins and to maintain homeostasis of the body fluids. Renal toxicity (RT) is the deterioration of the kidney function due to toxic effect of medications and/or chemicals. Exposure to certain drugs, chemicals, or environmental factors can lead to renal

toxicity, causing severe damage to renal cells and impairing kidney function. Renal toxicity is a significant concern in drug development and clinical practice, as it can lead to adverse drug reactions, treatment failures, and even life-threatening conditions.

To prevent kidney failure due to adverse effect, early detection of kidney damage is necessary. However, toxicity analysis of lead compounds during drug discovery is an expensive and time-consuming process. In current time, the machine learning (ML) and deep learning (DL) methods have gained considerable attention in the field of drug discovery [1–11]. These advanced computational approaches offer the potential to improve the accuracy, efficiency, and scalability of predicting renal toxicity, aiding in early identification and prevention of nephrotoxicity-related complications. By leveraging large-scale data analysis, pattern recognition, and predictive modeling, machine learning and deep learning methods can provide valuable insights into the field of nephrotoxicity, thereby guiding drug development and clinical decision-making processes [12–24].

✉ Pankaj Kumar Deva Sarma
   pankajgr@rediffmail.com

1   Department of Computer Science, Assam University, Silchar, Assam 788011, India

2   Advanced Computation and Data Sciences Division, CSIR-North East Institute of Science and Technology, Jorhat, Assam 785006, India

3   Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, Uttar Pradesh 201002, India

Considering the various factors involved, predicting renal toxicity through machine learning model can be an alternative, effective, and a cost-efficient approach as compared to the traditional methods. Apart from renal-toxicity prediction, ML algorithms have also contributed as well as outperformed in different areas like virtual screening, antiviral predictions, drug repurposing, membrane permeability, other toxicity prediction [1, 7–11], etc. triggering an emerging space of research in computational drug discovery.

These algorithms can learn from historical data comprising chemical descriptors, structural information, and various molecular properties to create predictive models that can classify compounds as either nephrotoxic or non-nephrotoxic. By identifying the specific features or combinations of features that correlate with renal toxicity, these models can provide insights into the underlying mechanisms of nephrotoxicity [25–28]. With the ability to capture complex relationships and patterns in large and diverse datasets, deep learning models can provide highly accurate predictions of renal toxicity. These models can analyze diverse data types, including molecular structures, gene expression profiles, and clinical data, to uncover hidden associations and gain a comprehensive understanding of the factors contributing to nephrotoxicity [12–20].

Another approach to toxicity analysis involves the use of structural alerts, which are specific substructures or chemical moieties associated with known toxic effects. By identifying and evaluating these structural alerts, researchers can make informed decisions regarding the potential toxicity risks associated with specific compounds [25–29]. These features can include functional groups, chemical motifs, or specific arrangements of atoms within a molecule. By applying computational methods, researchers can uncover patterns and associations between these structural alerts and various toxicological endpoints, such as nephrotoxicity, hepatotoxicity, genotoxicity, or cardiotoxicity. Approaches employed for generating structural alerts for renal toxicity include expert knowledge-based approaches, rule-based systems, and data-driven methods. Expert knowledge-based approaches rely on the expertise of toxicologists and medicinal chemists to identify substructures linked to renal toxicity. Rule-based systems utilize predefined rules and criteria to classify compounds based on the presence of specific substructures. Data-driven methods employ computational algorithms and machine learning techniques to extract structural alerts from large toxicological databases.

Data sources related to kidney studies include patient registries and epidemiology data, electronic health records (EHR) and healthcare administrative data, clinical trials, mobile devices and wearable sensors, molecular data repositories (genomics, epigenomics, transcriptomics, proteomics, and metabolomics) [12]. However, labeled data for developing ML-based model to predict renal-toxicity of compounds are very less and all the relevant studies have been conducted using SIDER dataset [30].

In this study, new deep neural network (DNN) and ML model have been proposed for predicting renal-toxicity of compounds. Moreover, a method has been proposed for generating structural alerts. This is done by applying association rule mining technique based on frequent itemset patterns to generate structural alerts about renal toxic compounds. Association among the fingerprint-based substructures was also studied to identify substructures that maybe responsible for the renal toxicity of the drug molecule. In addition, models developed using fingerprints and descriptors have been compared in this study.

The paper is organized as follows: in "Introduction", the background knowledge about renal toxicity is introduced and motivation for design of DNN and machine learning models for detection of renal toxic drugs is emphasized. In "Related work", related works in this domain is discussed. In "Materials and methods", materials and methods concerning prediction of renal toxic drugs are discussed. This include description of the dataset, descriptor generation, feature selection, and various algorithms of machine learning possible to be applied in solving drug discovery problems along with application of association rule mining based on frequent pattern mining for developing techniques to analyze the molecular substructures of drug molecules for detection of possible causes of renal toxicity. In "Result and analysis", the experimental results are presented, compared with other such results, and analyzed. In "Pattern mining for structural alert generation", a method has been proposed for generating structural alerts about renal toxic compounds by applying association rule mining technique based on frequent itemset patterns. The paper is concluded with a discussion about the findings.

Related Work

The quick assessment of the renal toxic potential of compounds is crucial in reducing the failures of drug development. In vivo testing for drug-induced nephrotoxicity assessment is complex, expensive, time-consuming, and not suitable for screening large numbers of compounds, particularly virtual chemicals. Additionally, experimental results are susceptible to different factors such as model animals, technology, and environmental conditions. Computational toxicology methods offer significant advantages over biological experimental approaches for estimating nephrotoxicity of compounds: (1) rapid prediction for large compound sets, (2) prediction of compound toxicity based on structure alone, even without synthesis. Therefore, the development of fast and accurate computational tools for nephrotoxicity risk

estimation holds great importance. While numerous ML and DL models have been developed for different kinds of toxicity prediction over the years, but only a few have focused on renal toxicity.

Machine learning (ML) has been used for predicting kidney malfunction [13], drug-induced adverse effects [14], xenobiotic-induced renal-toxicity [15], chemical-induced nephrotoxicity [16], detection of nephrotoxicity of a specific chemical like Tacrolimus [17]. Again, machine learning has been used for evaluating chronic [18] and acute [19] kidney disease, predicting complicacy after kidney transplant [20, 21], assessment of renal cancer tumors [22], prediction of kidney injury after a medical treatment surgery [23, 24]. Mostly, linear regression, random forest, support vector machine, k-nearest-neighbor, naive Bayes, XGBoost, artificial neural network have been used to develop these models. However, most of the above models were not exactly relevant for predicting renal-toxicity of compounds since these models were predicting the risk of disease in patients.

Lee et al. [25] use molecular fingerprints and metabolites of 540 compounds to build an SVM-based ML model for predicting renal toxicity of compounds with an accuracy of 80%. In the study, structural alerts for nephrotoxic compounds were identified using information gain and substructure fragment analysis [25]. Since metabolites are used as features, first these metabolites need to be generated for predicting toxicity of new compounds.

Lei et al. [26] developed an SVM-based model for predicting urinary tract toxicity using 279 compounds. They have achieved the highest accuracy of 90.8%. In the study, structural alerts were generated using a software, SARpy [29]. Lei et al. conducted a comprehensive summary of the models that have been reported and are associated with urinary tract toxicity up until the year 2017. In the study, only 279 compounds were used to train and test the models.

Shi et al. [27], have used molecular descriptors of 565 chemicals and different ML and DL algorithms to develop models for predicting renal toxicity of compounds. The study achieved the highest accuracy of 75.9% in the training set. The study, performed five-fold cross-validation on the test set and the best model was selected based on the cross-validation scores. Again, they have generated the structural alerts by calculating f-score and positive rate.

Gong et al. [28] have developed ML-based models for predicting nephrotoxicity of chemical drugs and Chinese herbal medicines. In their study, they have used molecular fingerprints of 777 compounds. They have achieved the highest accuracy of 86% using light gradient boosting algorithm. However, their test data includes compounds that are present in their training set. Therefore, reliability of this model is questionable. Again, Gong et al. have identified structural alerts by SARpy.

Gardiner et al. [13] have developed models for prediction of animal drug toxicity using gene expression data and three ML-based algorithms. In their study, Gaussian process-based Bayesian model was used to measures the model uncertainty. However, to predict renal toxicity of a compound, first gene expression needs to be calculated since gene expressions were used as features to train the model. In this study no structural alerts were generated.

Above studies have not checked the relation between combinations of substructures and renal toxicity. Therefore, in this study, association rule mining [31] technique was used to generate the alerts by considering all the combinations of all the substructures. Again, some of the models were not evaluated properly and some models need wet lab experimental data before prediction. Therefore, it is required to deploy a properly-tested model that can predict nephrotoxicity of compounds without much overhead. The most relevant related works are summarized in Table 1.

Materials and Methods

The models and structural alerts were developed using python 3 environments in Ubuntu system (version 20.4) which was installed in a Dell PC consisting of i7 core and 16 GB RAM.

Model Development

Dataset Preparation

In this study, 565 marketed drugs with human nephrotoxic and non-nephrotoxic labels were collected from a published dataset [27]. Shi et al. extracted 287 human nephron-toxic compounds from Side Effect Resource (SIDER) database [30] and 278 non-nephrotoxic compounds from Zhang's work [32]. SIDER is a widely used repository of adverse drugs reactions on human. Since a compound can have different SMILES notion so it is not useful for duplicate removal. Therefore, a unique identifier, InchiKeys has been generated using OpenBabel [33] to check duplicate and ambiguous compounds in the dataset.

Descriptor Generation

Molecular descriptors and molecular fingerprints are widely used forms of a molecule representation. These representations were used in most of the machine learning studies. Fingerprints embed the presence or absence of a particular structural feature in a vector. On the other hand, molecular descriptors list the numeric values of different physio-chemical properties of a chemical. In this study, 2D molecular

Table 1 Related works in the field of predicting renal toxicity using ML and/or DL

Ref	Model	# Compounds	Structural alerts	Disadvantage/shortcoming/limitation
[25]	SVM	540	Information gain and substructure fragment analysis	Metabolites are used as features, so these metabolites need to be generated before predicting toxicity. So extra time and effort is needed. Did not consider combinations of substructures in generating the alerts
[26]	SVM	279	A software, SARpy was used to generate structural alerts	A small number of chemicals were used to develop the model. Did not consider combinations of substructures in generating the alerts
[13]	Gaussian process, LightGBM	429	No	Gene expressions were used as features to train the model. For predicting renal toxicity of unseen compounds first gene expressions needs to be calculated. Computation of gene-expression is time consuming and expensive. Structural alerts were not generated
[27]	Consensus	565	F score and positive rate were used to generate structural alerts	Performed five-fold cross-validation on test set. Did not evaluate combinations of substructures in generating the alerts
[28]	SVM	777	A software, SARpy was used to generate structural alerts	The presence of training set instances in test set. Did not consider combinations of substructures in generating the alerts

descriptors of 565 compounds were calculated using RDKIT software [34] and eight kinds of molecular fingerprints namely MACCS, Extended, Graph-only, Klekota-Roth, Pubchem, Atom-pair, Substructure, and Estate fingerprints were generated using Padel descriptor software [35].

Outliers Study

A widely used technique, principal component analysis (PCA) was used to plot the data in two-dimensional graph for visually detecting outliers in the dataset. Later, those datapoints whose first principal component (PC1) is more than or lower than interquartile range were deleted.

Feature Selection

Correlation among the features is a well-known method for reducing the number of features. In this study, result of the correlation was further filtered using forward feature selection method to overcome the curse of dimensionality issue. In the first step, Pearson's correlation coefficient was calculated to see the related features. The correlation value between pairs of features is used to keep only one feature when the correlation among the features is more than 90%. In the next step, sequential feature selection method [36] was used to extract those features which gave the best accuracy score of logistic regression model.

DL and ML Algorithm

Deep Neural Network: The artificial neural network that contains more than one hidden layer are termed as Deep Neural Network (DNN). DNN architecture has neurons in each hidden layer and these neurons execute some mathematical calculations and forward this calculated value to the succeeding layer based on the bias, weight, and activation function [37].

XGBoost: XGBoost [38] classifier uses boosting technique where N base-trees are built in such a manner, so that mth tree decreases the errors of the predecessor's tree. The mth tree learns from its predecessors and updates the residual errors of $(m - 1)$th tree. This model was highly used in developing ML predictive models.

Extremely Randomized Tree (Extra-tree) [39]: Given a $A \times B$ dimensional dataset (where $A =$ number of datapoints and $B =$ number of independent variables), an Extra-tree of E decision trees (estimators) is constructed by iteratively choosing n samples (where $n < A$) using random sampling with replacement technique. The Extra-tree model's final classification depends on the maximum voting of the E estimators. Extra-tree differs from Random Forest (RF) in the way of selecting the cut points for splitting its nodes. In RF it selects the best split whereas Extra-tree makes random selection for this purpose.

Training and Testing

The dataset obtained after removing the outliers was divided into training and testing set in the ratio of 8:2. In the training

phase, one bagging classifier (Extra-tree), one boosting classifier (XGBoost) and one deep learning (DL) based classifier (DNN) was fed with fingerprints and descriptors of compounds in the training set. The hyper-parameter space of Extra-tree and XGBoost classifiers were searched and best hyper-parameters were selected with the help Grid-Search CV [36]. Since, the hyper-parameter space of DNN is large, Bayesian optimization [40] was used to extract the optimal hyper-parameters. As the output of the Bayesian optimization method did not gave good result, hyper-parameters of DNN are tuned manually in the next phase.

The tunned parameter of Extra-tree classifiers includes number of estimators, max_features, max_depth, min_samples_leaf, min_samples_split, criterion. Similarly, for XGBoost classifier it includes number of estimators, max_depth, learning_rate, colsample_bytree, gamma, and regularization. The optimized hyper-parameters for deep neural network include number of hidden layers, number of neurons in hidden layers, optimizers, activation functions, learning rate, drop-out rate, l2-regulirazation, number of epochs, and batch-size.

To validate the models, five-fold cross-validation was used in the training set. Again, following metrics were calculated on the test-set to evaluate a model—ROC-AUC, accuracy, specificity, sensitivity, Matthew's correlation coefficient (MCC), and Cohen's Kappa coefficient (κ). The standard definitions of these metrics were used in this study.

Structural Alert Generation

(i) Pattern discovery through association rule mining:

An association rule implies that in a database of transactions whenever there are occurrences of set of attributes X in some transactions then another set of attributes Y also occurs along with X in the same transactions. Association rules are of the form $X \Rightarrow Y$ where X and Y are sets of items or attributes of a transaction database and the meaning of this rule is whenever an itemset X is present in transactions of the database the itemset Y is also co-present along with X such that $X \cap Y = \phi$ [41]. There is no common attribute in the sets of attributes X and Y. Thus $X \Rightarrow Y$ is an if-then rule which connects set of attributes X with set of attributes Y if X is already present. In other words, association rules provide occurrence patterns of attributes of objects in large databases of transactions as if-then rules. Using association rule, occurrence patterns of features or attributes can be discovered in the transaction databases.

The acceptability of an association rule depends on how good or interesting or strong the rule is. This is measured using different parameters of strength-like support, confidence, lift, etc. Support and confidence [31] are two widely used parameters to denote the merits of the rule. Support of a rule $X \Rightarrow Y$ is determined by the frequency of the combined occurrences of all the attributes in the attribute set X in the antecedent and in the attribute set Y in the consequent in the entire database of transactions. That is the frequency of occurrence of the set $X \cup Y$ in the entire database of transactions is the support of the rule $X \Rightarrow Y$ and it means how often X and Y occur together as percentage of total number of transactions.

For determination of effective association rules, a pre-specified minimum support threshold is provided as input. As a result, the sets of attributes which have support equal or more than the pre-specified minimum support are called frequent itemsets. All the frequent itemsets are computed in a database of transactions to discover the association rules. The frequent itemsets or attribute sets thus represent frequently occurring patterns of attributes and their relationships during occurrences are represented as association rules.

Confidence of an association rule is another parameter to measure its effectiveness. For a rule $X \Rightarrow Y$, its confidence c is defined as $c\%$ of the transactions in the database that support X also support Y. Confidence of a rule measures the extent a set of attributes (X) is dependent on the set of attribute Y. In other words, the confidence of the rule $X \Rightarrow Y$ is nothing but the conditional probability of occurrence of the set of attribute Y in the database of transactions when the set of attributes X has already occurred. In terms of support, the confidence c of the rule $X \Rightarrow Y$ is given as:

$$c = \text{support}(X \cup Y)/ \text{support}(X).$$

It is required to discover all such rules which are having support and confidence equal or more than a certain pre-specified minimum threshold on support and confidence.

There are two major computational tasks involved in the process of discovery of association rules from a database of transactions with respect to certain pre-specified minimum thresholds on support and confidence. First, to discover all the frequent set of attributes and second, to generate all the association rules from the frequent set of attributes discovered.

There are two broad approaches of discovery of frequent set of attributes—one, by generating candidate sets of attributes and the other is without generating candidate sets of attributes. A number of algorithms are designed for discovery of frequent itemsets and association rules based on these two approaches of discovery of frequent sets of attributes. Most prominent among these are the a priori algorithm [31] and the Frequent Pattern (FP) Tree Growth algorithm [41]. There are many other algorithms for discovery of frequent itemsets and association rules but most of these algorithms are designed based on the principles of either the a priori algorithm or the Frequent Pattern (FP) Tree Growth algorithm.

(ii) Why and how Frequent Itemset mining and Association Rule Discovery is relevant in analyzing drug molecules for Renal Toxicity:

For analyzing the renal toxicity of drug molecules, the molecular fingerprints of the drug molecules are generated. This allows the listing of the constituent types like MACCS FP, substructure FP, estate FP, etc. for each drug molecule as transactions. When this is done for all the collected drug molecules a database of transactions containing all the molecular fingerprints is generated in which every row corresponds to a particular drug molecule and every column represents a specific fingerprint or substructure. The presence of particular fingerprint for a molecule is marked as 1 and while its absence is marked as 0 in the corresponding row of each molecule. Thus, a complete binary transaction dataset of molecular fingerprints is prepared for entire set of molecules. Once this dataset is prepared in which one of the columns (attribute) is about the toxicity of the molecules and for those molecules which are renal toxic, this column has value 1 otherwise the value of this attribute is 0. Now frequent attribute set discovery and association rule mining algorithms can be applied on this dataset.

This is mining frequent patterns and association rules in database of molecular fingerprints for the purpose of discovery of frequent molecular fingerprints among the types like MACCS FP, substructure FP, estate FP, etc. in the renal toxic drug molecules. This will lead to identification of the most frequent occurrences of various molecular fingerprints and their associations with renal toxicity of the drug molecules. This, in turn is likely to help in finding the causes of toxicity in these molecules without performing wet laboratory experiments and may also help in focusing upon designing techniques for possible elimination of toxicity in these molecules. Further, this may lead to many other domain-based analysis on these substructures and other types of molecular fingerprints and the corresponding molecules.

Based on the survey works undertaken in this study it has been observed that frequent pattern mining in databases of molecular fingerprints is not undertaken in any of the reviewed literature. However, it is envisaged that based on frequent pattern mining and association rule mining vital information about the interdependencies among the molecular fingerprints can be discovered not only for renal toxic drug molecules but for other kinds of drug molecules as well. This is an unsupervised machine learning model of computing which is based on discovery driven approach. Therefore, a need is felt for frequent pattern mining in such molecular finger prints of drug molecules and accordingly by executing the FP tree growth algorithm many frequent patterns of molecular fingerprints and their correspondence consisting of MACCS FP, substructure FP, estate FP, etc. are discovered from the drug molecules which are highly

informative, previously unknown, novel and shall be useful from the point of view of design and discovery of drug molecules. The discovered patterns carry such useful information which cannot be predicted based on supervised learning techniques as such techniques evaluates the prospects and prediction of whether a drug molecule is renal toxic or not. When such a prediction is combined with the results of frequent pattern mining and the discovered associations among the fingerprints, then more concrete information is obtained and it can be prescribed for further experimentation in wet laboratory environment.

(iii) Structural alerts through pattern mining:

By applying association rule mining technique in the database of molecular fingerprints of a large number of compounds, association among their constituent substructures can be discovered. Based on such rules, structural alerts for nephrotoxic compounds can be generated. For generating structural alerts, first the renal toxic compounds were filtered from the original dataset. In the next step, three types of fingerprints namely MACCS FP, substructure FP, and estate FP were generated for all the renal toxic compounds which resulted in three datafiles. The MACCS FP, estate FP, and substructure FP based datafiles represents the presence or absence (1 or 0) of 166, 79, and 301 substructures in a compound, respectively. The MACCS FP datafile consists of 287 rows corresponding to the number of compounds, 166 columns corresponding to the number of substructures and 1 column for representing the renal toxicity. In the third step, these datafiles were used as input in the FP-tree algorithm to find the relationship between the powerset of substructures and renal toxicity. Complete methodology of this study is presented in Fig. 1.

Result and Analysis

Data Distribution

In the study, two principal components of 565 compounds (including 287 renal toxic drugs and 278 non-renal toxic drugs) are plotted in Fig. 2. From this figure, it can be observed that the range of the second principal component (PC2) is relatively wide and both training and test sets share a similar chemical space. However, some outliers can be detected based on the first principal component (PC1), so all the 32 outliers were removed from the dataset. After removing the outliers, 523 compounds are randomly divided into training and test set in the ratio of 8 is to 2. It is worth to mention that the number of toxic drugs is approximately equal to the number of non-toxic drugs, so this dataset does not suffer from class imbalance issue.
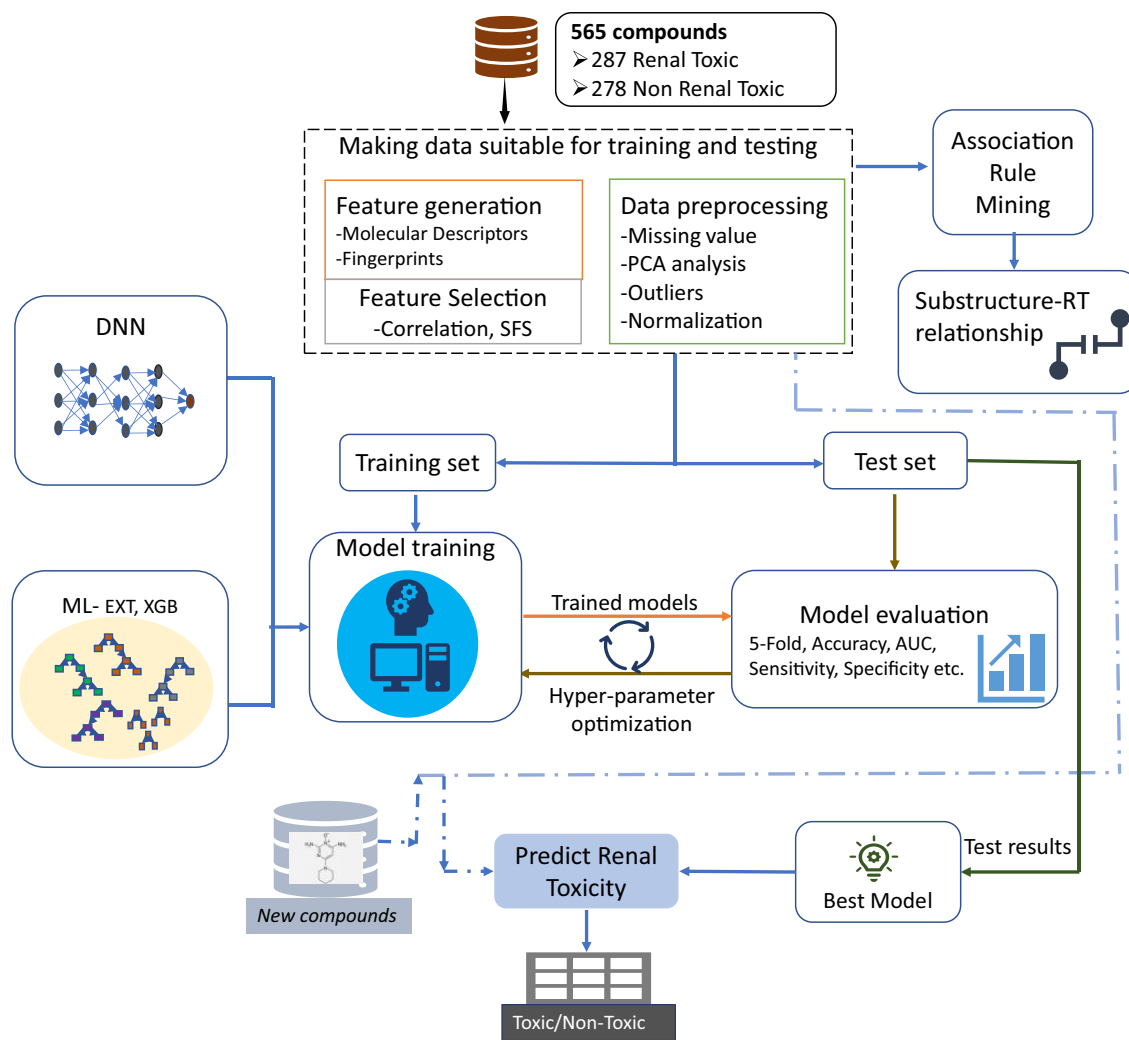
Fig. 1 Methodology for developing predictive models and structural alerts for determining renal toxicity of compounds
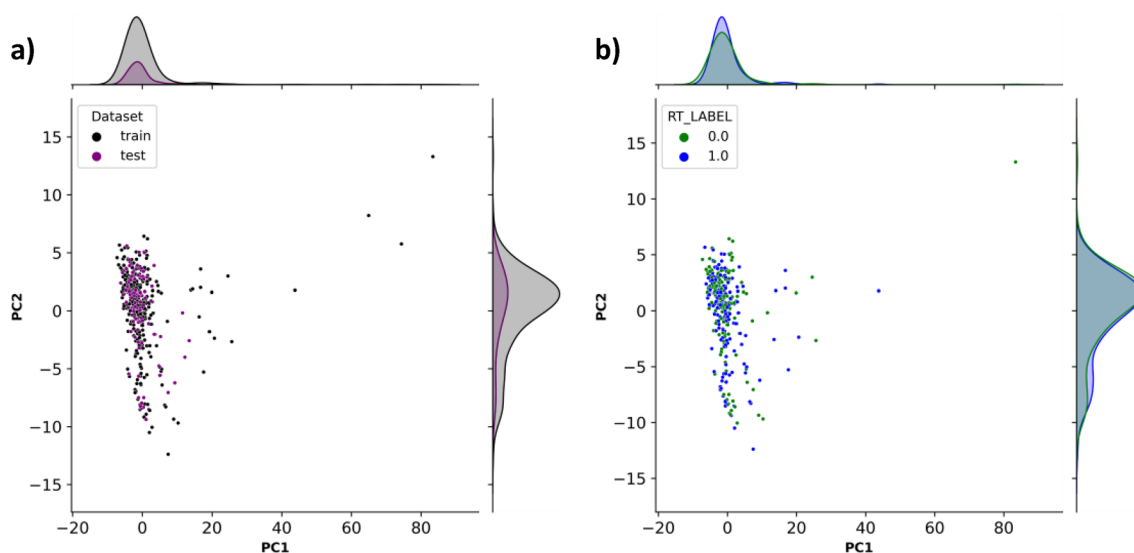

Fig. 2 Data distribution of compounds: **a** distribution of training and testing set data; **b** distribution of renal-toxic and non-renal-toxic data

Model Development

In this study, model development starts with missing value calculation, removal of outliers, selection of descriptors, and min–max normalization. It was observed that Rdkit was not able to generate values for 12 attributes of 1 compound. Since the percentage of missing value is less than 0.5% these values are imputed using the mean value of the attribute. Further, the 208 molecular descriptors were screened to select the important features. Initially, 67 features were removed based on the Pearson's correlation coefficient. In the second step, out of the 141 features, 138 features were selected based on the score obtained from the sequential feature selection method. The correlation heat-map and feature scores are shown in Figs. 3 and 4, respectively.

In the training phase, initially various ML algorithms like Logistic regression, SVM, KNN, Naïve Bayes, Random Forest, Extra-tree, XGBoost, and DL algorithms like Convolutional Neural Network, transfer learning, deep neural network are explored. However, accuracy of most of these models are below 70% on the test set so only three techniques have been considered for further exploration. In the second phase, hyper-parameters of one bagging classifier (Extra-tree), one boosting classifier (XGBoost), and one DL-based classifier (DNN) were tuned to enhance these models. In case of Extra-tree and XGBoost classifier, optimized parameters are obtained using GridSearchCV. In case of DNN, initially Bayesian optimization technique was used to find the optimized hyper-parameters. In the later stage, resultant values of the Bayesian method were used to reduce the hyper-parameter search space and this reduced space is manually explored to enhance the performance of the DNN model. The corresponding values of the tuned hyper-parameters are shown in Table 2.

In this study, model validation was started with five-fold cross-validation on training data and followed by evaluating different performance metrics of models on the test set.

The ROC-AUC of 19 models in the five-fold were studied to check the underfitting, overfitting, and for tunning hyper-parameters of the models. Table 3 depicts the performance of the datasets and models in terms of mean ROC-AUC scores in the five-fold cross-validation. It was seen that the models trained using Atom-pair, MACCS, Grap-only, and Rdkit datasets shows ROC-AUC of more than 0.80. The DNN model trained with Rdkit descriptors shows the highest scores of 0.88 on the cross-validation and the XGBoost model trained with the substructure fingerprints shows the least score of 0.61. It was observed that scores of 52 folds are greater than the average ROC-AUC (0.75) and scores of 42 models are below the average, hence it can be inferred that some models are not generalizing well. Therefore, all the models are not considered for further study.

The five-fold cross-validation ROC-AUC curves of three models are shown in Fig. 5. From this figure it can be observed that the performance of the DNN model is very consistent in all the folds. As the five-fold ROC-AUC score of DNN is between 0.85 and 0.88, it can be inferred that this model is free from overfitting and underfitting. The difference between ROC-AUC scores of Extra-tree and XG Boost models were 0.6 and 0.7, respectively, so it can be inferred that these models slightly overfit.

Table 4 displays the different performance scores of all the models on the test set. It was observed that the models trained using Rdkit molecular descriptors have performed better than the models that were trained using fingerprints. The Extra-tree classifier has performed much better than the XGBoost and DNN model on the test set with 0.87 ROC-AUC and 82% accuracy. The XGBoost model trained using estate, MACCS, and substructure fingerprints shows the least ROC-AUC score of 0.73. In terms of accuracy the Extra-tree model trained using extended fingerprints shows the least score of 64%. Since the same Extra-tree algorithm gave the highest and lowest scores depending on the dataset it can be concluded that there is an impact of data
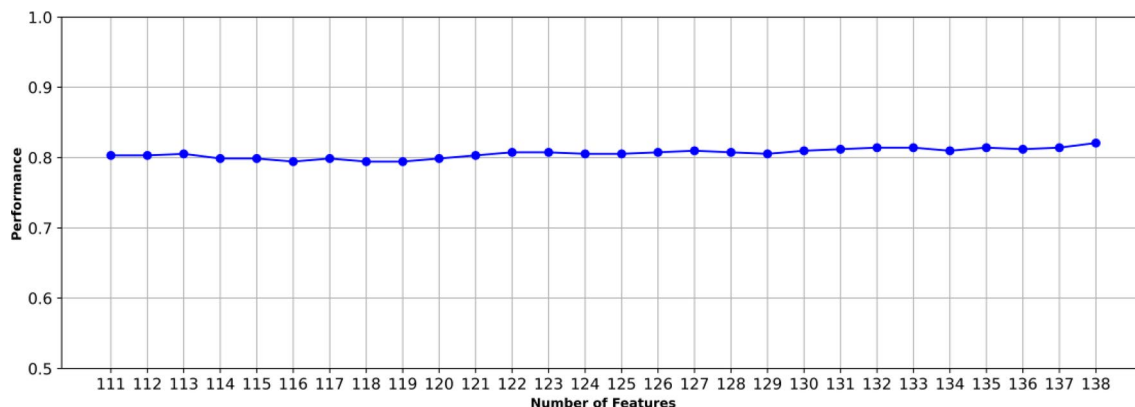
Fig. 3 Feature selection performance of logistic regression for different combinations of 111–138 features
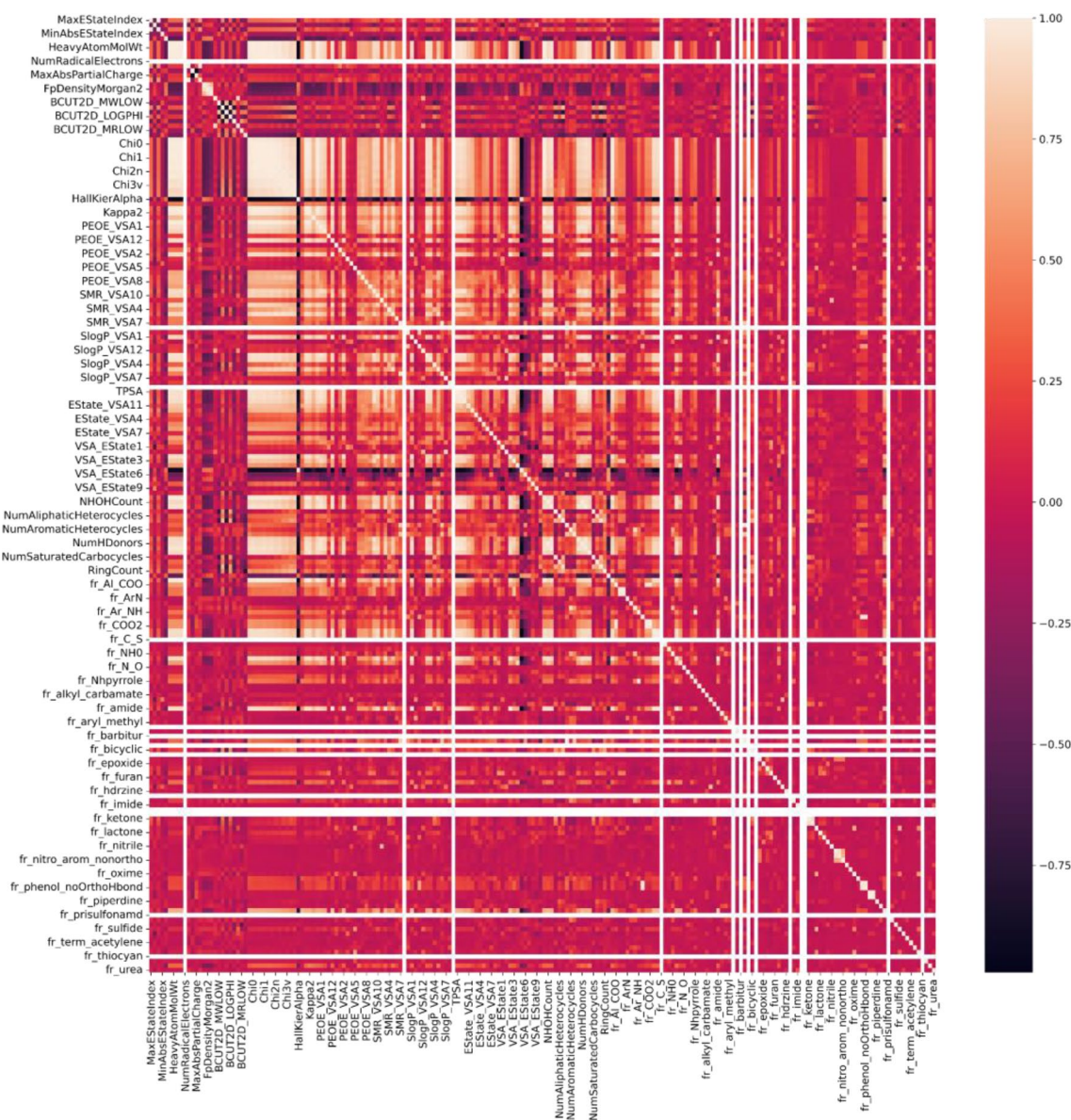
Fig. 4 Correlation between different features, correlation heat-map

representation (dataset) on the performance of the model. Again, seven models showed better scores than the average accuracy of 71% and 12 models' individual accuracy were lower than the average. Similarly, scores of ten models are above the average ROC-AUC (0.78) and ROC-AUC of 9 models are below the average. Examining all test-set scores, it can be inferred that both the algorithm and the data representation (molecular descriptors and fingerprints) plays an important role in developing renal toxicity prediction models.

It was found that the performance of DNN model on testing is lower than the training. Since all the important hyperparameters of DNN model have been explored exhaustively and the five-fold cross-validation scores were better than other models, the size of the test set may have impacted the DNN model performance and it is expected that the performance of DNN model will increase if the size of the dataset can be increased. It is worth to mention that the maximum accuracy of the DNN model was 0.79 on the test set while performing the *model.fit* operation.

Shi et al. [27] has developed models for predicting the nephrotoxicity of compounds using the same dataset. The prediction accuracy of their best model was 75.9%. The five-fold cross-validation ROC-AUC score of their best model (consensus) was 0.83. In the current study, the highest accuracy of 82.1% has been achieved by the Extra-tree model. Again, DNN model has achieved the highest ROC-AUC score of 0.86 on the five-fold cross-validation. Comparing

Table 2 Hyper-parameter search space and optimized values of DNN, Extra-tree, and XGBoost models

Model	Hyper-parameter	Search space	Bayesian selection	Manual selection
DNN	Optimizer	Rmsprop, Adadelta, Adam, SGD, Adagrad, Adamax, Nadam, Ftrl	Adam	Adam
	Activation Function	Softplus, Softsign, Relu, Sigmoid, Elu, Exponential, Leaky relu, Tanh, Selu	Elu	Relu
	Neurons	10–138	69	138
	LearningRate	0.0001–1	0.64	0.001
	BatchSize	16–256	217	128
	Epochs	20–1000	343	80
	Hidden layers	1–12	4	4
	Normalization	0–1	0.5635	
	DropoutRate	0–1	0.4849	0.5
	L2 regularization			0.01

Model	Hyper-parameter	Selected value	Model	Hyper-parameter	Selected value
Extra-tree	Bootstrap	FALSE	XGBoost	ColsampleBytree	0.2
	MaxDepth	None		Gamma	0.1
	MaxFeatures	0.5		LearningRate	0.05
	MinSamplesLeaf	1		MaxDepth	5
	MinSamplesSplit	2		Estimators	200
	Estimators	50		Reg. Alpha	0.5
				Reg. Lambda	2

Table 3 The mean ROC-AUC of five-fold cross-validation for 19 models on the training set

Model	F1	F2	F3	F4	F5	Mean	Max	Min	Range
Maccs_EXT	0.82	0.75	0.73	0.78	0.79	0.78	0.82	0.73	0.09
Maccs_XGB	0.79	0.77	0.73	0.79	0.77	0.77	0.79	0.73	0.06
extended_EXT	0.78	0.78	0.79	0.7	0.8	0.77	0.8	0.7	0.1
atompair_EXT	0.82	0.79	0.69	0.7	0.79	0.76	0.82	0.69	0.12
extended_XGB	0.77	0.76	0.77	0.69	0.77	0.75	0.77	0.69	0.08
pubchem_XGB	0.76	0.8	0.78	0.68	0.72	0.75	0.8	0.68	0.13
klekota_roth_EXT	0.8	0.77	0.71	0.71	0.75	0.75	0.8	0.71	0.08
pubchem_EXT	0.79	0.76	0.75	0.68	0.75	0.75	0.79	0.68	0.12
substructure_EXT	0.79	0.7	0.75	0.69	0.77	0.74	0.79	0.69	0.1
atompair_XGB	0.77	0.74	0.64	0.68	0.83	0.73	0.83	0.64	0.19
estate_EXT	0.75	0.75	0.72	0.7	0.7	0.73	0.75	0.7	0.06
graph_XGB	0.76	0.7	0.72	0.7	0.73	0.72	0.76	0.7	0.06
estate_XGB	0.7	0.79	0.68	0.7	0.68	0.71	0.79	0.68	0.11
graph_EXT	0.81	0.66	0.7	0.67	0.71	0.71	0.81	0.66	0.15
klekota_roth_XGB	0.78	0.73	0.63	0.67	0.73	0.71	0.78	0.63	0.15
substructure_XGB	0.74	0.72	0.7	0.61	0.75	0.7	0.75	0.61	0.14
Rdkit_XGB	0.75	0.74	0.8	0.74	0.81	0.77	0.81	0.74	0.07
Rdkit_EXT	0.78	0.75	0.8	0.77	0.81	0.78	0.81	0.75	0.06
Rdkit_DNN	0.86	0.86	0.88	0.85	0.85	0.86	0.88	0.85	0.03
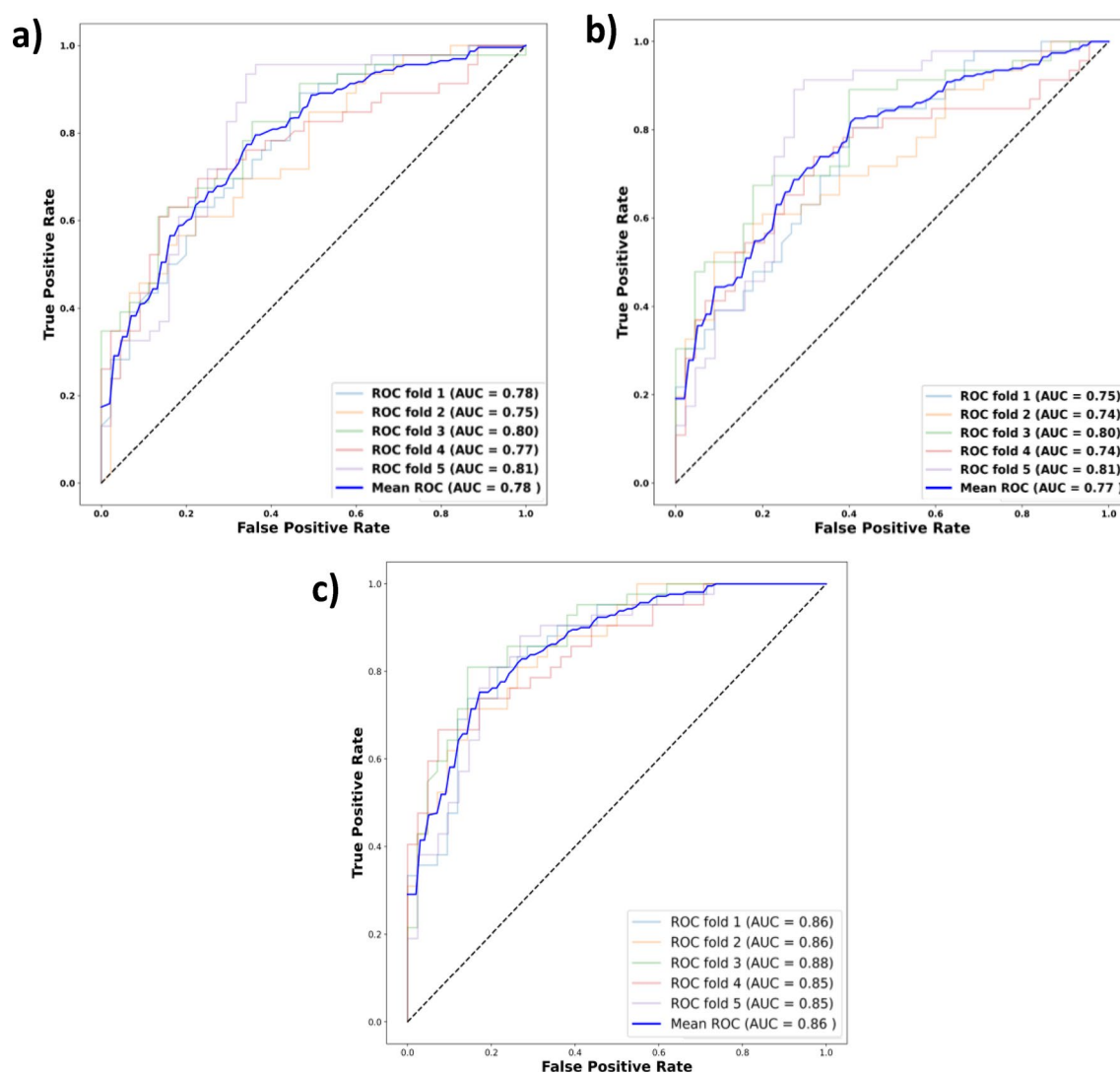
Fig. 5 Five-fold cross-validation ROC-AUC scores of the three classifiers trained on Rdkit descriptors: **a** Extra-tree, **b** XGBoost, and **c** DNN

the ROC-AUC value, models developed in this study is better than the Shi et al.

Pattern Mining for Structural Alert Generation

In this study, the association between substructure and renal-toxicity is studied using pattern mining for identifying substructures that are responsible for renal-toxicity. Co-occurrences of items are discovered in large transaction datasets to determine the influence of one set of items on another set of items when these occur together in large number of transactions. The idea here is to find possibility about any substructure that is responsible for a drug molecule exhibiting renal-toxicity through their associations with other functional groups of such molecules. The

result of the FP Tree growth algorithm [41] which is used for extracting the patterns is shown in Table 5.

From the pattern mining result, it was observed that 99% of the renal toxic compounds consists of substructure-based fingerprints—SubFP295, SubFP300, SubFP301, and SubFP307. 86–99% of nephron-toxic compounds consist of MACCS-based fingerprints—MACCSFP22, MAC-CSFP163, MACCSFP164, and MACCSFP165. Again, 88% toxic compounds consist of estate-based fingerprints—EStateFP9 and 80% toxic compounds consist of EstateFP35. This is based on support of an association rule which indicates the combined occurrences of the itemset on both the antecedent and consequent of an association rule together in the transactions of the database as percentage of the total number of the transactions in the whole database. Thus, for example consider the association rule appearing in the first row of Table 5, viz. Antecedent: SubFP295, SubFP307,

Table 4 Performance scores of various models trained on descriptors and fingerprints

Model	Accuracy	Sensitivity	Specificity	MCC	ROC-AUC	K
Rdkit_XGB	0.786	0.739	0.842	0.579	0.847	0.573
Rdkit_EXT	0.821	0.804	0.842	0.644	0.878	0.642
Klekota_Roth_XGB	0.743	0.754	0.732	0.487	0.782	0.487
Klekota_Roth_EXT	0.717	0.789	0.643	0.437	0.798	0.433
Pubchem_XGB	0.673	0.719	0.625	0.346	0.736	0.345
Pubchem_EXT	0.735	0.789	0.679	0.471	0.798	0.468
Estate_XGB	0.681	0.737	0.625	0.364	0.725	0.362
Estate_EXT	0.681	0.772	0.589	0.368	0.744	0.362
Extended_XGB	0.655	0.702	0.607	0.31	0.743	0.309
Extended_EXT	0.637	0.649	0.625	0.274	0.777	0.274
MACCS_XGB	0.69	0.754	0.625	0.383	0.734	0.38
MACCS_EXT	0.717	0.789	0.643	0.437	0.806	0.433
Substructure_XGB	0.673	0.684	0.661	0.345	0.734	0.345
Substructure_EXT	0.699	0.684	0.714	0.399	0.798	0.398
Graph_XGB	0.699	0.754	0.643	0.4	0.789	0.398
Graph_EXT	0.699	0.702	0.696	0.398	0.781	0.398
Atompair_XGB	0.699	0.702	0.696	0.398	0.761	0.398
Atompair_EXT	0.699	0.684	0.714	0.399	0.762	0.398
Rdkit_DNN	0.743	0.623	0.865	0.502	0.813	0.487

EXT extra-tree, *XGB* XGBoost, *DNN* deep neural network

Table 5 Substructure alerts based on the support and confidence of association rules

Antecedents	Consequent	Support	Confidence	#Antecedent
SubFP295, SubFP307, SubFP301	RT	0.99	1	3
SubFP307	RT	0.997	1	1
SubFP295	RT	0.997	1	1
SubFP295, SubFP307	RT	0.993	1	2
SubFP301	RT	0.993	1	1
SubFP295, SubFP301	RT	0.993	1	2
SubFP307, SubFP301	RT	0.99	1	2
SubFP295, SubFP307, SubFP301	RT	0.99	1	3
SubFP300	RT	0.993	1	1
SubFP300, SubFP301	RT	0.993	1	2
SubFP300, SubFP295	RT	0.993	1	2
MACCSFP22	RT	0.997	1	1
MACCSFP165	RT	0.934	1	1
MACCSFP22, MACCSFP165	RT	0.934	1	2
MACCSFP164	RT	0.92	1	1
MACCSFP164, MACCSFP22	RT	0.92	1	2
MACCSFP164, MACCSFP165	RT	0.868	1	2
MACCSFP164, MACCSFP22, MAC-CSFP165	RT	0.868	1	3
MACCSFP163	RT	0.916	1	1
MACCSFP163, MACCSFP22	RT	0.916	1	2
MACCSFP163, MACCSFP165	RT	0.909	1	2
EStateFP9	RT	0.889	1	1
EStateFP35	RT	0.805	1	1

SubFP301; Consequent: RT; Support: 0.99; and Confidence: 1. In formal notation, this rule is written as:

$$\{SubFP295, SubFP307, SubFP301\} -> \{RT\} \text{ with support}$$
$$= 0.99 \text{ and confidence } = 1.$$

In this rule the itemset {SubFP295, SubFP307, SubFP301, RT} is present or occurred in 99% of the transactions of the whole database. The confidence of the above rule is 1, i.e., 100%. The meaning of this rule is whenever the antecedent {SubFP295, SubFP307, SubFP301} of the above rule is found present in the transactions of the dataset (in this case it is present in 99% of the transactions) the itemset {RT} is also present in all these transactions as the confidence is 100%.

Now interpreting this rule in the context of the data set of compounds prepared for renal toxicity it can be inferred that 99% of the renal toxic compounds in the dataset contains the substructure-based fingerprints SubFP295, SubFP307, SubFP301 together in each compound. Thus, there may be a possibility of these substructures contributing to the toxic nature of the compounds. Further, from the frequent patterns and the association rules in Table 5, it is seen the substructures SubFP295 and SubFP307 are present individually in 99.7% of the renal toxic compounds and these two together are present in the 99.3% of the renal toxic compounds. Similarly, the substructure SubFP301 is present alone in 99.3% of the renal toxic compounds in the dataset and in combination with SubFP295 it is present in 99.3% and with SubFP307 it is present in 99%.

In this manner, in Table 5, the highly frequent fingerprint-based molecular substructures which are common to most of the renal toxic compounds in the dataset are discovered by the techniques of association rule mining. In this case the frequent itemset patterns and the rules are discovered with 60% minimum support threshold and maximum confidence of 100%. This was done to see both the minimum and the maximum occurrences of any substructure patterns which may lead to renal toxicity. However, rules with other values of minimum thresholds on support and confidence can also be discovered. In Table 5 most of the frequent occurrences are listed. As can be seen from Table 5 that certain fingerprint-based substructures occur in most of the compounds which exhibit renal toxicity. The possible influence of these substructures for the toxic behavior of these drug molecules may further be tested in wet lab conditions or based on the identity of the structures.

The FP Tree Growth algorithm [41] works by constructing a prefix tree to represent the itemsets present in the transaction database which contains the fingerprint-based substructures of the renal toxic compounds.

Conclusion

A dataset of only human nephron-toxic and non-nephron-toxic drugs is prepared in which eight types of fingerprints and one type of descriptor were used to represent the molecular structures to develop ML and DL based models for predicting renal toxicity of natural and artificial compounds. From the five-fold cross-validation result, the model with best prediction performance for renal toxicity was DNN. The best model for predicting renal toxicity of test-set drugs was found to be Extra-tree model. The performance of molecular descriptor-based models was better than the fingerprint-based models. At the same time, ten substructures of toxic drugs are identified by pattern mining. The presence of these substructures may indicate the likelihood of nephron-toxic potential of a compound. Structural alerts and models developed in this study can assist in assessing the risk of renal toxicity of compounds in drug discovery.

Declarations

Conflict of interest Authors have no conflict of interest.

Ethics approval NA.

Consent to participate NA.

Consent for publication NA.

References

1. Murugan NA, Priya GR, Sastry GN, Markidis S. Artificial intelligence in virtual screening: models versus experiments. Drug Discov Today. 2022. https://doi.org/10.1016/j.drudis.2022.05.013.
2. Xu Y, Wojtczak D. MC-NN: an end-to-end multi-channel neural network approach for predicting influenza A virus hosts and antigenic types. SN Comput Sci. 2023. https://doi.org/10.1007/s42979-023-01839-5.
3. Bozkurt C, Aşuroğlu T. Mortality prediction of various cancer patients via relevant feature analysis and machine learning. SN Comput Sci. 2023. https://doi.org/10.1007/s42979-023-01720-5.

4. Yadav S, Singh MK. Hybrid machine learning classifier and ensemble techniques to detect Parkinson's disease patients. SN Comput Sci. 2021. https://doi.org/10.1007/s42979-021-00587-8.

5. McGarry K, Nelson D, Ashton M. A method to explore the connectivity patterns of proteins and drugs for identifying disease communities. SN Comput Sci. 2020. https://doi.org/10.1007/s42979-020-00151-w.

6. Kang Y, Huang ST, Wu PH. Detection of drug–drug and drug–disease interactions inducing acute kidney injury using deep rule forests. SN Comput Sci. 2021. https://doi.org/10.1007/s42979-021-00670-0.

7. John L, Soujanya Y, Mahanta HJ, Narahari SG. Chemoinformatics and machine learning approaches for identifying antiviral compounds. Mol Inf. 2022. https://doi.org/10.1002/minf.202100190.

8. Madugula SS, John L, Nagamani S, Gaur AS, Poroikov VV, Sastry GN. Molecular descriptor analysis of approved drugs using unsupervised learning for drug repurposing. Comput Biol Med. 2021. https://doi.org/10.1016/j.compbiomed.2021.104856.

9. Nagamani S, Sastry GN. Mycobacterium tuberculosis cell wall permeability model generation using chemoinformatics and machine learning approaches. ACS Omega. 2021. https://doi.org/10.1021/acsomega.1c01865.

10. Filimonov DA, Lagunin AA, Gloriozova TA, Rudik AV, Druzhilovskii DS, Pogodin PV, Poroikov VV. Prediction of the biological activity spectra of organic compounds using the PASS online web resource. Chem Heterocycl Compd. 2014. https://doi.org/10.1371/journal.pone.0191838.

11. Druzhilovskiy DS, Rudik AV, Filimonov DA, Gloriozova TA, Lagunin AA, Dmitriev AV, Pogodin PV, Dubovskaya VI, Ivanov SM, Tarasova OA, Bezhentsev VM. Computational platform Way2Drug: from the prediction of biological activity to drug repurposing. Russ Chem Bull. 2017. https://doi.org/10.1007/s11172-017-1954-x.

12. Kaur N, Bhattacharya S, Butte AJ. Big data in nephrology. Nat Rev Nephrol. 2021. https://doi.org/10.1038/s41581-021-00439-x.

13. Gardiner LJ, Carrieri AP, Wilshaw J, Checkley S, Pyzer-Knapp EO, Krishna R. Using human in vitro transcriptome analysis to build trustworthy machine learning models for prediction of animal drug toxicity. Sci Rep. 2020. https://doi.org/10.1038/s41598-020-66481-0.

14. Wang Z, Clark NR, Maayan A. Drug-induced adverse events prediction with the LINCS L1000 data. Bioinformatics. 2016. https://doi.org/10.1093/bioinformatics/btw168.

15. Shi W, Bugrim A, Nikolsky Y, Nikolskya T, Brennan RJ. Characteristics of genomic signatures derived using univariate methods and mechanistically anchored functional descriptors for predicting drug- and xenobiotic-induced nephrotoxicity. Toxicol Mech Methods. 2008. https://doi.org/10.1080/15376510701857072.

16. Douali L, Villemin D, Cherqaoui D. Neural networks: accurate nonlinear QSAR model for HEPT derivatives. J Chem Inf Comput Sci. 2003. https://doi.org/10.1021/acs.molpharmaceut.7b00631.

17. Burghelea D, Moisoiu T, Ivan C, Elec A, Munteanu A, Iancu ȘD, Truta A, Kacso TP, Antal O, Socaciu C, Elec FI. The use of machine learning algorithms and the mass spectrometry lipidomic profile of serum for the evaluation of tacrolimus exposure and toxicity in kidney transplant recipients. Biomedicines. 2022. https://doi.org/10.3390/biomedicines10051157.

18. Tangri N, Ferguson TW. Role of artificial intelligence in the diagnosis and management of kidney disease: applications to chronic kidney disease and acute kidney injury. Curr Opin Nephrol Hypertens. 2022. https://doi.org/10.1097/MNH.0000000000000787.

19. Wei C, Zhang L, Feng Y, Ma A, Kang Y. Machine learning model for predicting acute kidney injury progression in critically ill patients. BMC Med Inform Decis Mak. 2022. https://doi.org/10.1186/s12911-021-01740-2.

20. Van Loon E, Zhang W, Coemans M, De Vos M, Emonds MP, Scheffner I, Gwinner W, Kuypers D, Senev A, Tinel C, Van Craenenbroeck AH. Forecasting of patient-specific kidney transplant function with a sequence-to-sequence deep learning model. JAMA Netw Open. 2021. https://doi.org/10.1001/jamanetworkopen.2021.41617.

21. Ramm S, Todorov P, Chandrasekaran V, Dohlman A, Monteiro MB, Pavkovic M, Muhlich J, Shankaran H, Chen WW, Mettetal JT, Vaidya VS. A systems toxicology approach for the prediction of kidney toxicity and its mechanisms in vitro. Toxicol Sci. 2019. https://doi.org/10.1093/toxsci/kfz021.

22. Shehata M, Alksas A, Abouelkheir RT, Elmahdy A, Shaffie A, Soliman A, Ghazal M, Abu Khalifeh H, Salim R, Abdel Razek AA, Alghamdi NS. A comprehensive computer-assisted diagnosis system for early assessment of renal cancer tumors. Sensors. 2021. https://doi.org/10.3390/s21144928.

23. Filiberto AC, Ozrazgat-Baslanti T, Loftus TJ, Peng YC, Datta S, Efron P, Upchurch GR Jr, Bihorac A, Cooper MA. Optimizing predictive strategies for acute kidney injury after major vascular surgery. Surgery. 2021. https://doi.org/10.1016/j.surg.2021.01.030.

24. Amiri S, Akbarabadi M, Abdolali F, Nikoofar A, Esfahani AJ, Cheraghi S. Radiomics analysis on CT images for prediction of radiation-induced kidney damage by machine learning models. Comput Biol Med. 2021. https://doi.org/10.1016/j.compbiomed.2021.104409.

25. Lee S, Kang YM, Park H, Dong MS, Shin JM, No KT. Human nephrotoxicity prediction models for three types of kidney injury based on data sets of pharmacological compounds and their metabolites. Chem Res Toxicol. 2013. https://doi.org/10.1021/tx400249t.

26. Lei T, Sun H, Kang Y, Zhu F, Liu H, Zhou W, Wang Z, Li D, Li Y, Hou T. ADMET evaluation in drug discovery. 18. Reliable prediction of chemical-induced urinary tract toxicity by boosting machine learning approaches. Mol Pharm. 2017. https://doi.org/10.1021/acs.molpharmaceut.7b00631.

27. Shi Y, Hua Y, Wang B, Zhang R, Li X. In silico prediction and insights into the structural basis of drug induced nephrotoxicity. Front Pharmacol. 2022. https://doi.org/10.3389/fphar.2021.793332.

28. Gong Y, Teng D, Wang Y, Gu Y, Wu Z, Li W, Tang Y, Liu G. In silico prediction of potential drug-induced nephrotoxicity with machine learning methods. J Appl Toxicol. 2022. https://doi.org/10.1002/jat.4331.

29. Ferrari T, Gini G, Bakhtyari NG, Benfenati E. Mining toxicity structural alerts from SMILES: a new way to derive structure activity relationships. In: 2011 IEEE symposium on computational intelligence and data mining (CIDM). 2011. https://doi.org/10.1109/CIDM.2011.5949444.

30. Kuhn M, Letunic I, Jensen LJ, Bork P. The SIDER database of drugs and side effects. Nucleic Acids Res. 2016. https://doi.org/10.1093/nar/gkv1075.

31. Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases. In: Proceedings of the 1993 ACM SIGMOD international conference on management of data. 1993. https://doi.org/10.1145/170035.170072.

32. Zhang H, Ren JX, Ma JX, Ding L. Development of an in silico prediction model for chemical-induced urinary tract toxicity by using naïve Bayes classifier. Mol Divers. 2019. https://doi.org/10.1007/s11030-018-9882-8.

33. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open babel: an open chemical toolbox. J Cheminform. 2011. https://doi.org/10.1186/1758-2946-3-33.

34. Landrum G, Tosco P, Kelley B, Ric, Cosgrove D, sriniker, gedeck, Vianello R, Schneider N, Kawashima E, N D, Jones G, Dalke A, Cole B, Swain M, Turk S, Savelyev A, Vaucher A, Wójcikowski M, Take I, Probst D, Ujihara K, Scalfani VF, godin G, Lehtivarjo J, Pahl A, Walker R, Berenger F, Biggs JD. strets123. rdkit/rdkit: 2023_03_2 (Q1 2023) Release. Zenedo. 2016. https://doi.org/10.5281/zenodo.591637.

35. Yap CW. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. J Comput Chem. 2011. https://doi.org/10.1002/jcc.21707.

36. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825–30.

37. Hinton OT, Hinton GE, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets. Neural Comput. 2006. https://doi.org/10.1162/neco.2006.18.7.1527.

38. Ramraj S, Uzir N, Sunil R, Banerjee S. Experimenting XGBoost algorithm for prediction and classification of different datasets. Int J Control Theory Appl. 2016. https://doi.org/10.1145/2939672.2939785.

39. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn. 2006. https://doi.org/10.1007/s10994-006-6226-1.

40. Nogueira F. Bayesian Optimization: Open source constrained global optimization tool for Python. Github. 2014. https://github.com/fmfn/BayesianOptimization

41. Han J, Pei J, Yin Y. Mining frequent patterns without candidate generation. ACM SIGMOD Rec. 2000. https://doi.org/10.1145/335191.335372.