



Artificial Intelligence-Based Application to Explore Inhibitors of Neurodegenerative Diseases

Leping Deng^{1†}, Weihe Zhong^{1†}, Lu Zhao^{1,2†}, Xuedong He¹, Zongkai Lian¹, Shancheng Jiang^{1*} and Calvin Yu-Chian Chen^{1,3,4*}

¹ Artificial Intelligence Medical Center, School of Intelligent Systems Engineering, Sun Yat-sen University, Shenzhen, China, ² Department of Clinical Laboratory, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China, ³ Department of Medical Research, China Medical University Hospital, Taiwan, China, ⁴ Department of Bioinformatics and Medical Engineering, Asia University, Taiwan, China

OPEN ACCESS

Edited by:

Zijun Zhang,
City University of Hong Kong,
Hong Kong

Reviewed by:

Li Zhuang,
City University of Hong Kong,
Hong Kong
Chongjun Yan,
Dongbei University of Finance and
Economics, China

*Correspondence:

Shancheng Jiang
jiangshch3@mail.sysu.edu.cn
Calvin Yu-Chian Chen
chenyuchian@mail.sysu.edu.cn

[†]These authors have contributed
equally to this work

Received: 14 October 2020

Accepted: 30 November 2020

Published: 22 December 2020

Citation:

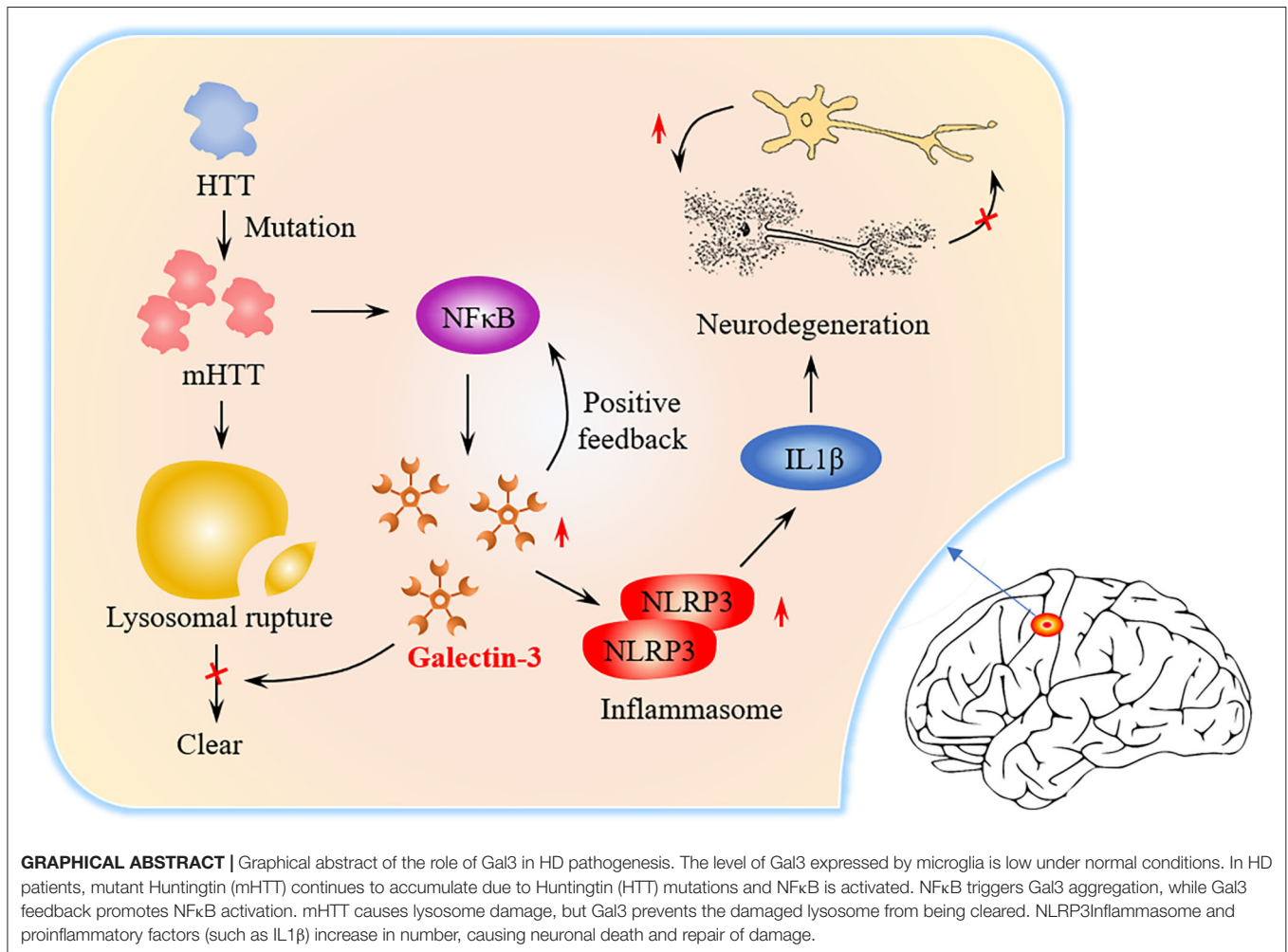
Deng L, Zhong W, Zhao L, He X, Lian Z, Jiang S and Chen C-YC (2020) Artificial Intelligence-Based Application to Explore Inhibitors of Neurodegenerative Diseases. *Front. Neurobot.* 14:617327. doi: 10.3389/fnbot.2020.617327

Neuroinflammation is a common factor in neurodegenerative diseases, and it has been demonstrated that galectin-3 activates microglia and astrocytes, leading to inflammation. This means that inhibition of galectin-3 may become a new strategy for the treatment of neurodegenerative diseases. Based on this motivation, the objective of this study is to explore an integrated new approach for finding lead compounds that inhibit galectin-3, by combining universal artificial intelligence algorithms with traditional drug screening methods. Based on molecular docking method, potential compounds with high binding affinity were screened out from Chinese medicine database. Manifold artificial intelligence algorithms were performed to validate the docking results and further screen compounds. Among all involved predictive methods, the deep learning-based algorithm made 500 modeling attempts, and the square correlation coefficient of the best trained model on the test sets was 0.9. The XGBoost model reached a square correlation coefficient of 0.97 and a mean square error of only 0.01. We switched to the ZINC database and performed the same experiment, the results showed that the compounds in the former database showed stronger affinity. Finally, we further verified through molecular dynamics simulation that the complex composed of the candidate ligand and the target protein showed stable binding within 100 ns of simulation time. In summary, combined with the application based on artificial intelligence algorithms, we unearthed the active ingredients 1,2-Dimethylbenzene and Typhic acid contained in *Crataegus pinnatifida* and *Typha angustata* might be the effective inhibitors of neurodegenerative diseases. The high prediction accuracy of the models shows that it has practical application value on small sample data sets such as drug screening.

Keywords: artificial intelligence, deep belief network, molecular dynamic simulation, galectin-3, neurodegenerative disease

INTRODUCTION

Neurodegenerative diseases (ND) cause the progressive death of central neurons, leading to brain dysfunction and the development of diseases, such as Huntington's disease (HD) (Macdonald et al., 1993), Alzheimer's disease (AD) (McKhann et al., 1984) and Parkinson's disease. ND often result from the aberrant deposition of aggregated host proteins (Voet et al., 2019).



At present, the mechanism of ND is not clear, but inflammation is considered to be a common factor (Saijo et al., 2010). Galectin-3 (Gal3) is an important member of the galectin family (Romero and Gabius, 2019). Gal3 is a key molecule linking inflammation and decreased insulin sensitivity (Li et al., 2016). Recently, more and more studies have shown that Gal3 is closely related to ND. Gal3 plays an important role in regulating inflammation (Henderson and Sethi, 2009). Extensive research on Gal3 in the central nervous system has shown that Gal3 promotes inflammation (Shin, 2013). Inhibition of Gal3 can help reduce inflammation in ND (Ramirez Hernandez et al., 2020). In the brains of AD patients, Gal3 promotes the activation of microglia (Ramirez et al., 2019), and inhibition of Gal3 may be a potential pharmacological method for the treatment of AD (Boza-Serrano et al., 2019). The latest research found that the brain Gal-3 content of patients and mice with HD is higher than normal. Inflammation can be controlled and the accumulation of mutant Huntingtin is reduced by inhibiting Gal3 (Siew et al., 2019). From the signal

transmission process (**Graphical Abstract**), it can be identified that inhibition of Gal3 may become a new drug target for HD treatment.

Understanding the basic laws of target protein-drug interactions is the basis of molecular targeted drug design, which plays a vital role in drug discovery and drug design (Rahman et al., 2020). Gal3 is one of the most potential target proteins for treating ND. Molecular docking method and active ingredient screening techniques are used to screen out drug molecules that have inhibitory effects on the target protein from drug database (Abdolmaleki et al., 2017). Traditional Chinese medicine is a medicine with great modern potential (Wen et al., 2019). Greater than 85 percent of patients diagnosed with COVID-19 in China have received Chinese medicine for adjuvant treatment (Yang et al., 2020). Therefore, discovering and designing the Chinese medicine prescription inhibitors of Gal3 is expected to have curative effects ND treatment. With the continuous improvement of computer performance, artificial intelligence (AI)-based methods are increasingly applied to

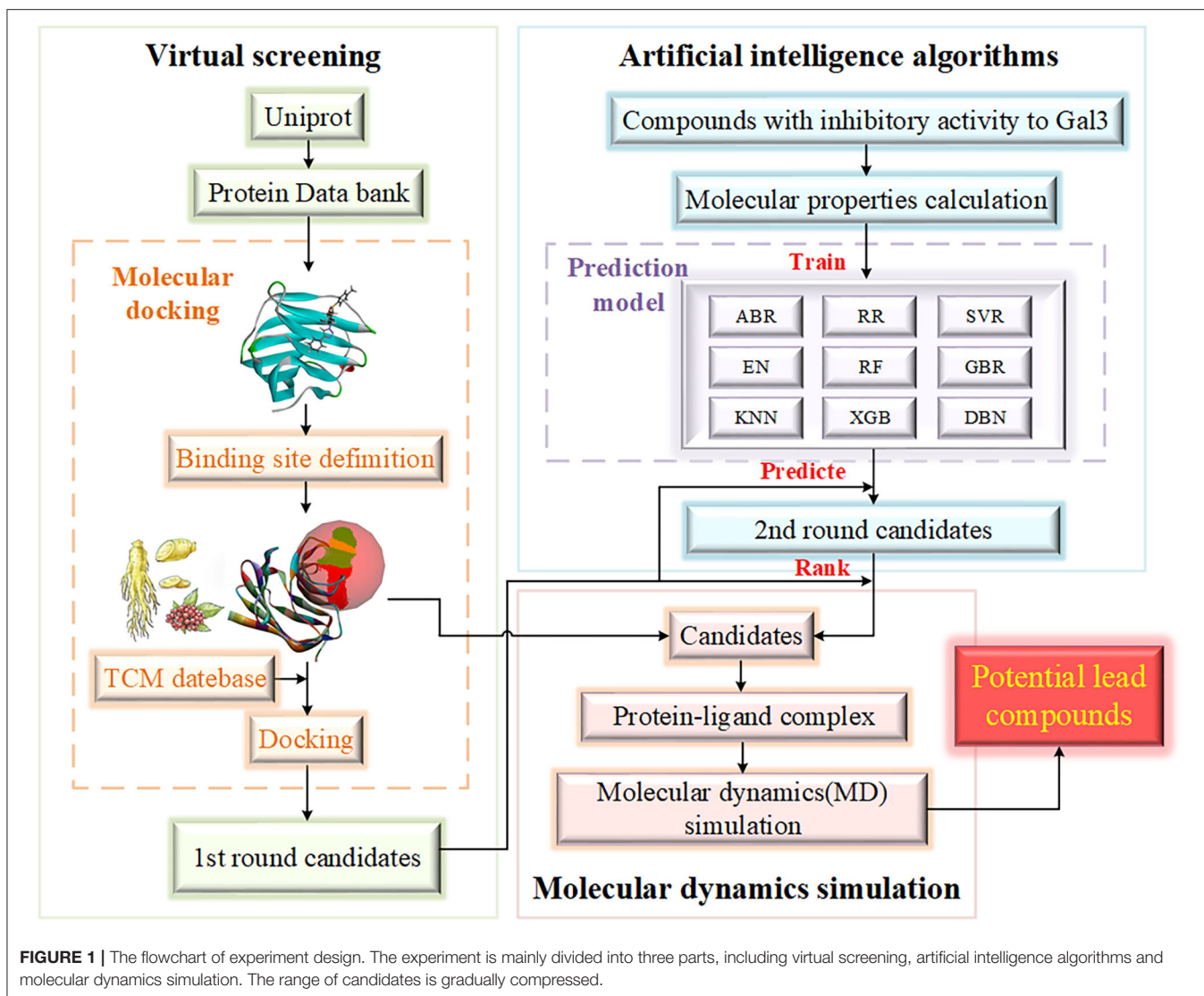


FIGURE 1 | The flowchart of experiment design. The experiment is mainly divided into three parts, including virtual screening, artificial intelligence algorithms and molecular dynamics simulation. The range of candidates is gradually compressed.

various stages of drug discovery (Chen et al., 2018a; Schneider et al., 2020; Senior et al., 2020). Machine learning methods are used to predict biologically active properties (Kaiser et al., 2018; Correia et al., 2020). SVM and other methods were used to establish four quantitative prediction models of the inhibitory activity value of HIV-1 integrase inhibitors (Xuan et al., 2013). The DeepTox algorithm shows good accuracy in predicting the toxicity of compounds (Mayr et al., 2016). Drug-drug interactions prediction (Zhang et al., 2019), biomolecular properties prediction (Hessler and Baringhaus, 2018) and quantum mechanical property prediction are combined with AI that is used in pharmacodynamic research of new synthetic drug candidates, which can greatly save costs. The combination of AI and traditional Chinese medicine may be a new development trend of modern Chinese medicine in the future (Liu et al., 2017).

Based on the discovery of Gal3 as a key target protein in a new pathogenesis closely related to HD, the purpose

of this study to screen potential compounds that inhibit Gal3 has been determined. The contribution of our study is 2-fold. Methodologically, we have added AI algorithms to establish a compound activity value prediction model based on traditional drug screening methods, including molecular docking and molecular dynamics (MD) simulation (Kumar et al., 2020), accelerating the process of new drug discovery. Screening drug molecules from traditional Chinese medicine (TCM) database is an innovation and beneficial supplement to screening from the general database. Practically, the underlying relationship between compound activity values and input molecular properties can be acquired through algorithm models. Target variable is an important evaluation index of drugs, which can provide researchers with reference, and the high accuracy of the model improves reliability. **Figure 1** provides the flowchart of experiment design.

MATERIALS AND METHODS

Molecular Docking Screening

Molecular docking is a virtual screening tool and can identify lead compounds from large small molecule databases, which reducing the number of experimental screening compounds and thus shortening the research cycle (Elmezayen et al., 2020; Pant et al., 2020). Molecular docking can be used to investigate the possible weak interactions between small molecule ligands and large molecule receptors and to calculate their affinity (Liu et al., 2019). The sequence of Gal3 was obtained from UniProt knowledgebase [(Identifier: P17931) (Bateman et al., 2017), and the crystal structure was obtained from RCSB Protein Data Bank (PDB ID: 6QLR) (Burley et al., 2019; Kumar et al., 2019), with a resolution of 0.97 Å. Through preprocessing operation, including removing crystal water molecules in the composite crystals, replenishing missing hydrogen atoms and optimizing energy by using the CHARMM27 force field (Brooks et al., 2009), a receptor protein with high confidence binding site was presented. The binding site was defined with the pro-ligand. A total of 18,776 molecule compounds obtained from the TCM database (TCM Database @Taiwan) (Chen, 2011) and 148,120 molecule compounds from ZINC database (<http://zinc.docking.org/>) (Irwin and Shoichet, 2005) were used as ligands for molecular docking, respectively. The original ligand in the complex was used as the control ligand, which was used as a reference for docking results. All involved experiments were implemented on LigandFit module in Discovery Studio Client v17.2.0.16349 (DS). Ligandfit has the functions of automatic search and confirmation of the active site of the receptor molecule, conformationally flexible multi-ligand docking, and evaluation of interaction scores based on force fields (Venkatachalam et al., 2003).

Artificial Intelligence-Based Prediction Models

Data Collection and Processing

Relevant information (structural formula and IC₅₀) of the small molecules reported that have inhibitory effect on Gal3 were collected from literatures, invention patents, and drug generation companies. Chemdraw was used to draw the structural formula of molecules (Mills, 2006). In addition, from open source small molecule databases such as PubChem (Kim et al., 2019), ChEMBL (Gaulton et al., 2017) and ZINC, we have downloaded the corresponding three-dimensional structure containing Gal3 inhibitors. All collected compounds were saved as Mol format files and converted to SDF format through Chem3D. The sample with clear IC₅₀ value could be kept, and 56 molecular samples were included. Chem3D software was used to minimize the molecular posture energy of all molecule samples. All qualified molecular samples were also subject to molecular attitude energy minimization processing. The IC₅₀ value was changed to pIC₅₀ as target variable by equation (1). All collected and sorted sample molecules were imported into DS to calculate molecular properties of 204 types that used as the input feature set.

$$pIC_{50} = 6 - \log_{10}(IC_{50}) \quad (1)$$

Algorithm 1: AdaBoost

1. Initialize weights distribution of training samples:
 $D_1 = (\omega_{11}, \omega_{12}, \dots, \omega_{1i}, \dots, \omega_{1N}), \omega_{1i} = 1/N, i = 1, 2, \dots, N$
2. For $m = 1, 2, \dots, M$ multiple iterations:
 - (1) Training the weighted D_m sample set to obtain the base learner $G_m(x)$
 - (2) Calculating the maximum error of the training set: $E_m = \max |y_i - G_m(x_i)|$
 - (3) Calculating the relative error of each sample:

$$e_{mi} = \frac{|y_i - G_m(x_i)|}{E_m}$$
 - (4) Calculating the regression error rate:

$$e_m = \sum_{i=1}^N \omega_{mi} e_{mi}$$
 - (5) Calculating the weight coefficients of weak learners:

$$\alpha_m = \frac{e_m}{1 - e_m}$$
 - (6) Updating the weight distribution of the sample set:

$$\omega_{m+1,i} = \frac{\omega_{mi}}{Z_m} \alpha_m^{1 - e_{mi}}$$

$$Z_m = \sum_{i=1}^N \omega_{mi} \alpha_m^{1 - e_{mi}}$$

$$D_{m+1,i} = (\omega_{m+1,1}, \omega_{m+1,2}, \dots, \omega_{m+1,i}, \dots, \omega_{m+1,N})$$
3. Output the ultimate strong learner

$$f(x) = \sum_{m=1}^M (\ln \frac{1}{\alpha_m}) \alpha_m G_m(x)$$

end

We performed AdaBoost, random forest, XGBoost, deep belief network and other models to build models for predicting pIC₅₀ values. These models are widely used AI-based machine learning models, with fast convergence speed and advantages in processing small sample data, which is very suitable for our small-scale data experiments.

AdaBoost Model

The kernel of adaptive boosting algorithm (AdaBoost) is to build a strong learner by connecting multiple weak learners (Huang et al., 2020). AdaBoost algorithm adjusts the weight of samples in the training set of each round by increasing the weight of the samples that were incorrectly predicted in the previous round (Ratsch et al., 2001). The iteration continues until the predetermined error rate is reached or the specified maximum number of iterations is reached (Figure 2). The ultimate strong learner is combined by linearly weighting and summing all base learners, and the base learner with small error rate has a larger weight coefficient. Details of the AdaBoost algorithm is presented in Algorithm 1. In this study the optimal number of decision trees was set as 15 based on some initial trials.

Ridge Regression Model

Ridge regression (RR) works well on condition that the number of independent variables is more than the sample size. As shown in equation 2, the RR model adds a penalty term of the L2 norm to the objective function of the ordinary linear regression mode (Yang and Wen, 2018), which contributed to the biased estimation of the regression coefficient β . Generally, RR is a regression method that solves the ill-conditioned matrix problem at the cost of giving up unbiasedness and reducing accuracy. The

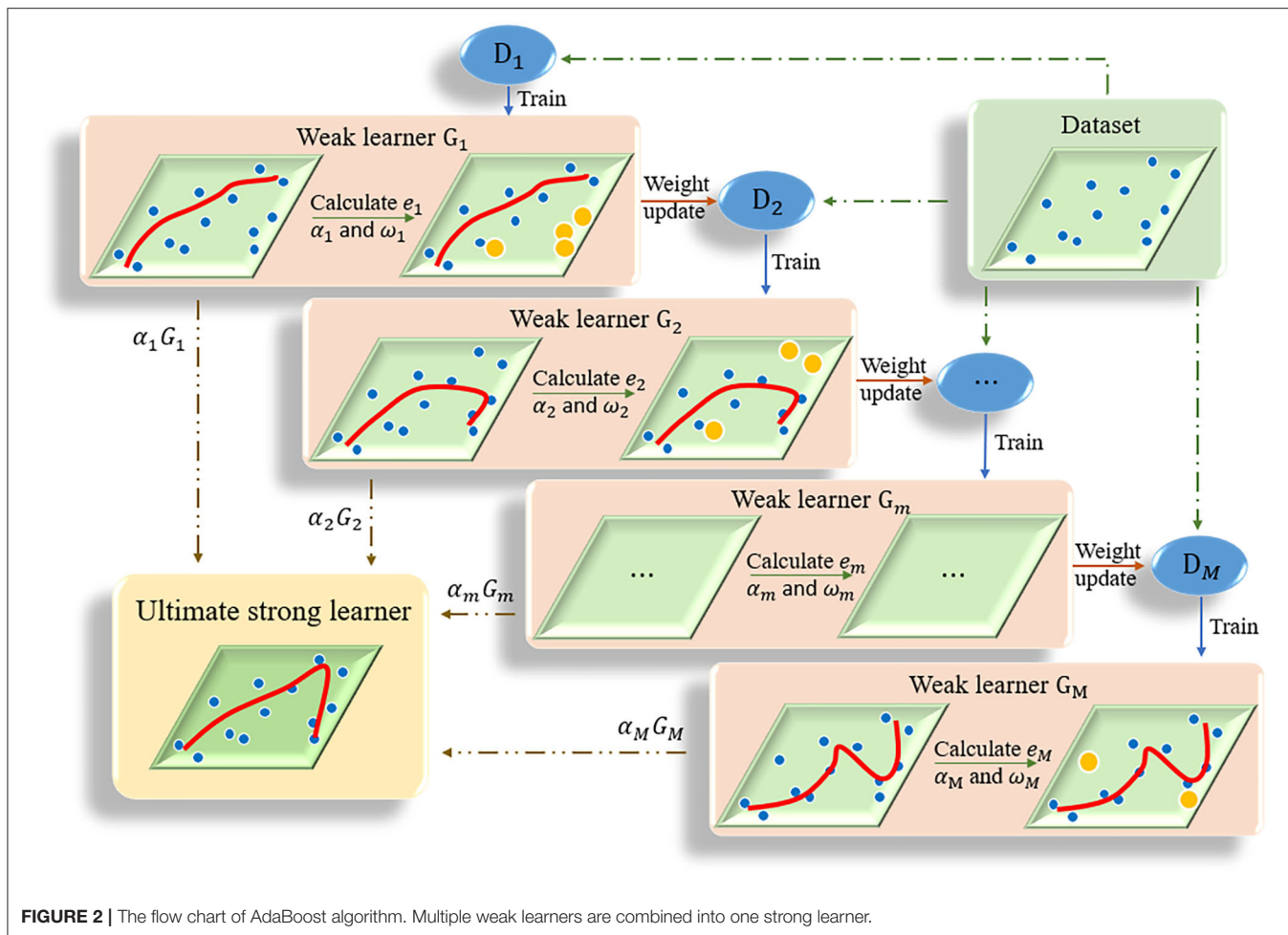


FIGURE 2 | The flow chart of AdaBoost algorithm. Multiple weak learners are combined into one strong learner.

alpha parameter was set to 0.05 in the case study.

$$\begin{aligned}
 J(\beta) &= (y - X\beta)^T + \lambda\beta^T\beta \\
 \Rightarrow \beta &= (X^T X + \lambda I)^{-1} X^T y
 \end{aligned}
 \tag{2}$$

SVM Model

Support vector machine (SVM) can be divided into support vector classification (SVC) and support vector regression (SVR) in practical applications (Chang and Lin, 2011). SVR is designed to fit each training sample and retain all the main features that characterize the algorithm to minimize errors. The kernel function is used to replace the linear term in the linear equation to make the original linear algorithm non-linear, which is used to achieve non-linear regression (Figure 3). SVR has high accuracy and strong generalization ability to solve small sample data. It has better applicability for the diversity of drug molecular characteristics and less sample data. In this study, the SVR algorithm was used to quantitatively predict the inhibitory activity of Gal3 inhibitors. All data were plotted in 28-dimensional space, and the error tolerance parameter was set to 0.39.

Elastic Net Model

Elastic net (EN) is a linear regression model trained using L1, L2 norms as prior regular terms (Zou and Hastie, 2005). EN is very effective in cases where multiple features are interconnected. The cost function of elastic network is as shown in Equation 3. When r is equal to 0, it is RR, when r is equal to 1, it is Lasso regression (LR). EN is compromised in RR and LR.

$$J(\theta) = MSE(\theta) + \gamma\alpha \sum_{i=1}^n |\theta_i| + \frac{1-\gamma}{2}\alpha \sum_{i=1}^n \theta_i^2
 \tag{3}$$

Random Forest Model

Random forest (RF) algorithm uses bootstrap method to generate training set (Breiman, 2001). Through random row and column (samples and features) sampling, seven decision trees were constructed to form a decision tree forest. The final prediction result was obtained by weighted average or voting. RF can achieve parallel learning and has a good filtering effect on noise and abnormal data (Athey et al., 2019). The model training process is given in Algorithm 2.

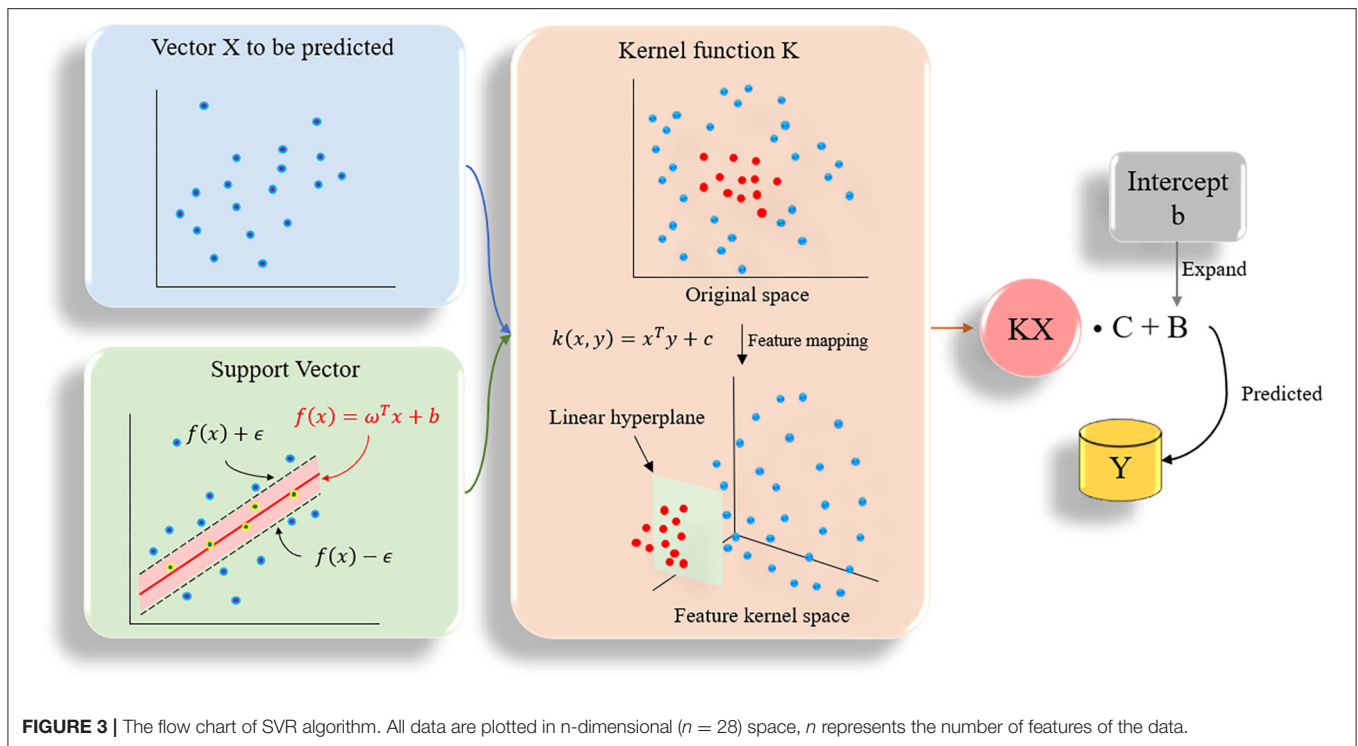


FIGURE 3 | The flow chart of SVR algorithm. All data are plotted in n-dimensional ($n = 28$) space, n represents the number of features of the data.

Algorithm 2 Random forest

1. Training set $D_{feature} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$
2. For $t = 1, 2, \dots, T$ do
 - (1) Random sampling m sample points, constructing a training set D_t
 - (2) Using D_t to train a decision tree
 - (3) Least squares regression tree as an example

Choosing the optimal segmentation variable i and the Segmentation Point s :

$$\min_{i,s} \left[\min_{c_1} \sum_{x_j \in R_1(i,s)} (y_j - c_1)^2 + \min_{c_2} \sum_{x_j \in R_2(i,s)} (y_j - c_2)^2 \right]$$

The selected pairs (i, s) were used to divide the region and determine the corresponding output values: $R_1(i, s) = \{x | x^i \leq s\}$, $R_2(i, s) = \{x | x^i > s\}$

$$c_\kappa = \frac{1}{N_\kappa} \sum_{x_j \in R_\kappa(i,s)} y_j, x \in R_\kappa, \kappa = 1, 2$$

Continue to iterate steps a and b for satisfying the condition

The input space was divided into K regions:

$$f(x) = \sum_{\kappa=1}^K c_\kappa I(x \in R_\kappa)$$

end

Gradient Boosting Regression Model

Gradient boosting regression (GBR) is an algorithm that improves by learning from mistakes. Each calculation of it is to reduce the residual error of the previous time and establish a

Algorithm 3 Gradient boosting regression

1. Initialize $F_0(x) = \arg \min_\gamma \sum_{i=1}^n L(y_i, \gamma)$
2. For $m = 1, 2, \dots, M$:
 - (1) For $i = 1, 2, \dots, N$ compute pseudo-residuals: $r_{im} = - \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{m-1}(x)}$
 - (2) Fit a base regression tree to the targets r_{im} giving terminal regions $R_{jm}, j = 1, 2, \dots, J_m$
 - (3) compute multiplier γ_{jm} : $= \arg \min_\gamma \sum_{x_i \in R_{jm}} L(y_i, F_{m-1}(x_i) + \gamma)$
 - (4) Update: $F_m(x) = F_{m-1}(x) + \sum_{j=1}^{J_m} \gamma_{jm} I(X \in R_{jm})$
3. Output $\hat{f}(x) = f_M(x)$
end

new model in the direction of the negative gradient (Chen et al., 2018b). The least square error was used as the loss function of regression prediction, and the number of decisions was set to 7. The sub-models are integrated as the final predictive model. The algorithm flow is given in Algorithm 3.

K-Nearest Neighbor Model

K-nearest neighbor (KNN) randomly divides the matrix into a training subset and a test subset, and returns the divided samples and labels. The Euclidean distance between the samples was calculated and sorted according to the distance (Abdel-Basset

Algorithm 4: XGBoost

1. Training set $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$
2. Loss function MSE (mean square error):

$$L(\theta) = \sum_i \left(y_i - \bar{y}_i \right)^2$$

where, y_i is the real value, \bar{y}_i is the predicted value.

3. For $k = 1, 2, \dots, K$ do

(1) The definition of the predicted model:

$$\bar{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F$$

K is the number of trees, F includes all possible trees, f_k is a specific tree,

$f_k(x_i)$ represents the predicted value of x_i on the k th tree.

(2) Minimizing objective function $L_{obj}(\theta)$ to acquire f_k .

$$L_{obj}(\theta) = \sum_i^n L(\theta) + \sum_{k=1}^K \Omega(f_k)$$

$\Omega(f_k)$ is the complexity of tree f_k .

4. Output f_k

e n-dimensional d

et al., 2020). The average value of the k ($k = 4$) samples closest to the target to be predicted was selected as the regression prediction value of the new sample.

XGBoost Model

XGBoost (XGB) algorithm improves the accuracy of the algorithm by adding the number of decision trees (Lai et al., 2020). The detailed process is shown as Algorithm 4. We used the `xgb.DMatrix` function. The representation method of the data in `libsvm` is a sparse matrix, which is very suitable for a large number of features and sparse. When there are missing values in the sample, XGB can automatically learn the split direction.

Deep Belief Network

Compared with traditional artificial neural networks, deep learning-based frameworks with huge numbers of multiple hidden layers maintain two-way fidelity of information transferring between different levels of abstraction during model learning (LeCun et al., 2015). Deep learning algorithms are widely used in drug design research such as protein structure prediction and disease diagnosis (Li et al., 2020; Senior et al., 2020). Deep belief networks (DBN) is the foundation of deep learning. In this study, we trained a quantitative prediction model of the inhibitory activity value of Gal3 inhibitors using a simple fully connected neural network with three hidden layers (Figure 4). The activation function was configured as rectified linear unit (ReLU) (Agarap, 2018). Furthermore, we introduce the Dropout method was to reduce the amount of calculation and increase the robustness of the model, by randomly pruning some neural units in the hidden layer with a predetermined probability (Zhang et al., 2018).

Molecular Dynamics Simulation

Molecular docking methods can narrow the scope of searches from large compound databases. Combining the molecular docking score and the target value predicted by artificial intelligence-based models, four optimal compounds (6318, 5372, 22157, 7649) ligands were finally screened as candidates for MD simulation. MD simulation can simulate and analyze the ligand-receptor movement process, and obtain some key information from the process to verify their stability (Song et al., 2019). SwissParam was chosen as the tool for generating topology files (Zoete et al., 2011), the mol2 file of candidates were submitted to obtain topology files with parameters such as atom type, charge and bonding conditions. The four candidate small molecules were combined with the processed Gal3 receptor protein and divided into four groups A, B, C, and D for MD simulation. CHARMM27 force field was used to describe the receptor protein (Sapay and Tieleman, 2011). Using TIP3P water molecule model, adding water as the solvent of the complex system, while adding NaCl and sodium ions to maintain the electrical neutrality of the system. The steepest descent method was used to optimize the energy of 5,000 steps to make the system reach near the lowest point of energy. After the optimization, it entered the equilibrium stage, and the positions of proteins and ligands needed to be restricted. First, balance for 10 ns under the constant number of atoms, constant volume and constant temperature (NVT) ensemble, and the system temperature rose from 0 to 310 K. Then it was equilibrated for 10 ns under the constant number of atoms, constant pressure and constant temperature (NPT) ensemble, and the temperature was kept at 310 K, which was used to simulate the physiological environment in the human body. The V-rescale method is used for temperature coupling, and the Parrinello-Rahman method is used for pressure coupling. After the equilibrium stage was over, the restriction was released, and each combination was subjected to 100 ns MD simulation under the NPT ensemble. In the simulation process, periodic boundary conditions (PBC) were used in all directions to eliminate possible boundary effects. The time step was set to 2fs, and the coordinate file and the energy file were recorded every 2ps.

RESULTS AND DISCUSSION**Molecular Docking**

Compounds from the TCM database were used for high-throughput virtual screening against Gal3, and the top 10 compounds were shown in Table 1. According to the docking scores of screenings, 6318, 5372, 7649, and 22157 (Table 1) were selected for further analysis. 2D diagram of these molecules in docking results showed potential interactions between key residues and ligands, including hydrogen bonds, van der Waals, salt bridge, Pi-Pi stacked, etc (Figure 5). These four TCM compounds have a common property, a carboxylate group in their chemical structures. Based on the docking poses of Gal3 in Figure 6, the α , β -unsaturated carbonyl groups of 6318 formed two hydrogen bonds (2.4 and 2.5 Å) with LYS176 and ARG144 of Gal3, respectively. The methoxy of 6318 formed a hydrogen bond (2.2 Å) with ARG162. The carboxylate group of 5372 formed a 2.1 Å hydrogen bond with LYS176. 7649's benzoic acid groups interacted with LYS176 and ARG144 through three hydrogen

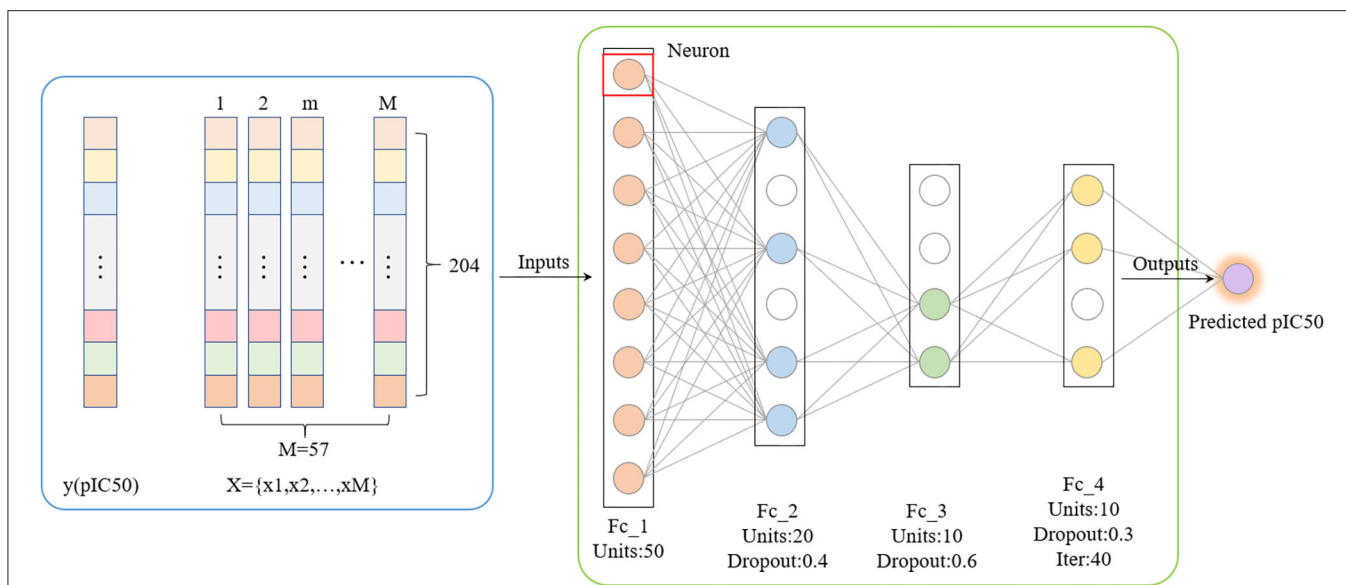


FIGURE 4 | Deep Belief Network generated by applying dropout. The mean square error is used as the loss function. The neural network weights are updated by an Adam optimizer with a learning rate of 0.0006. Rectified Linear Units and Dropout methods are utilized.

TABLE 1 | Docking score, predicted activity value of top ten TCM candidates, top two ZINC candidates and control ligand.

Index	Name	Predicted value (pIC50)									Docking Score
		ABR	RR	SVR	EN	RF	GBR	KNN	XGB	DBN	
6318	Chinese Hawthorn	7.119	6.830	6.791	6.851	6.665	6.839	6.693	7.093	6.162	125.276
22157	Longbract Cattail Pollen	7.155	7.519	7.532	7.440	6.696	6.937	6.554	7.093	5.743	116.132
5372	Carnation	6.951	7.023	7.148	7.128	7.061	7.037	6.887	6.380	5.591	119.682
7649	Staphyleaceae	6.650	7.956	8.281	8.099	6.886	6.779	6.942	6.510	6.629	116.693
2246	Amur Adonis	6.951	6.873	6.967	7.012	6.816	7.002	6.938	6.380	5.537	117.959
14992	White Mulberry Fruit	5.658	4.901	4.750	4.933	5.768	6.088	5.838	5.929	2.528	114.298
2670	Java Brucea	5.793	4.861	4.774	4.926	5.849	6.444	6.257	5.598	3.829	106.878
8713	Whiteflower Leadword	7.143	7.048	7.057	7.083	6.726	6.937	6.552	7.211	2.044	110.309
22676	Common Threewingnut	6.076	8.094	7.947	7.853	5.767	5.723	6.177	5.866	3.541	110.031
210	Fresh Common Ginger	6.509	5.143	5.171	5.100	6.347	6.011	6.020	6.769	2.963	106.732
ZINC000019363537	Tetraethylenepentamine	5.301	4.622	4.630	4.538	5.319	5.133	5.742	5.438	0.611	102.451
ZINC000019364225	Trientine	5.818	5.455	5.276	5.313	5.577	5.836	5.759	5.598	0.529	90.915
Control	J4N	6.180	6.838	6.902	6.899	6.542	6.034	6.979	6.510	4.962	56.381

bonds (1.7, 1.6, and 1.7 Å). The carboxylate group of 22157 had H-bonding interactions with ARG144 and ASN160 (1.7 and 2.4 Å), and the β -ketone carboxyl group engaged with ASP148 and LYS176 through two H-bonds. The hydrogen bond interaction of the compounds with LYS176 and ARG144 indicated that these were two key residues. The same experiment had been done with ZINC database and the results showed that the binding affinity was weaker than that from the TCM database.

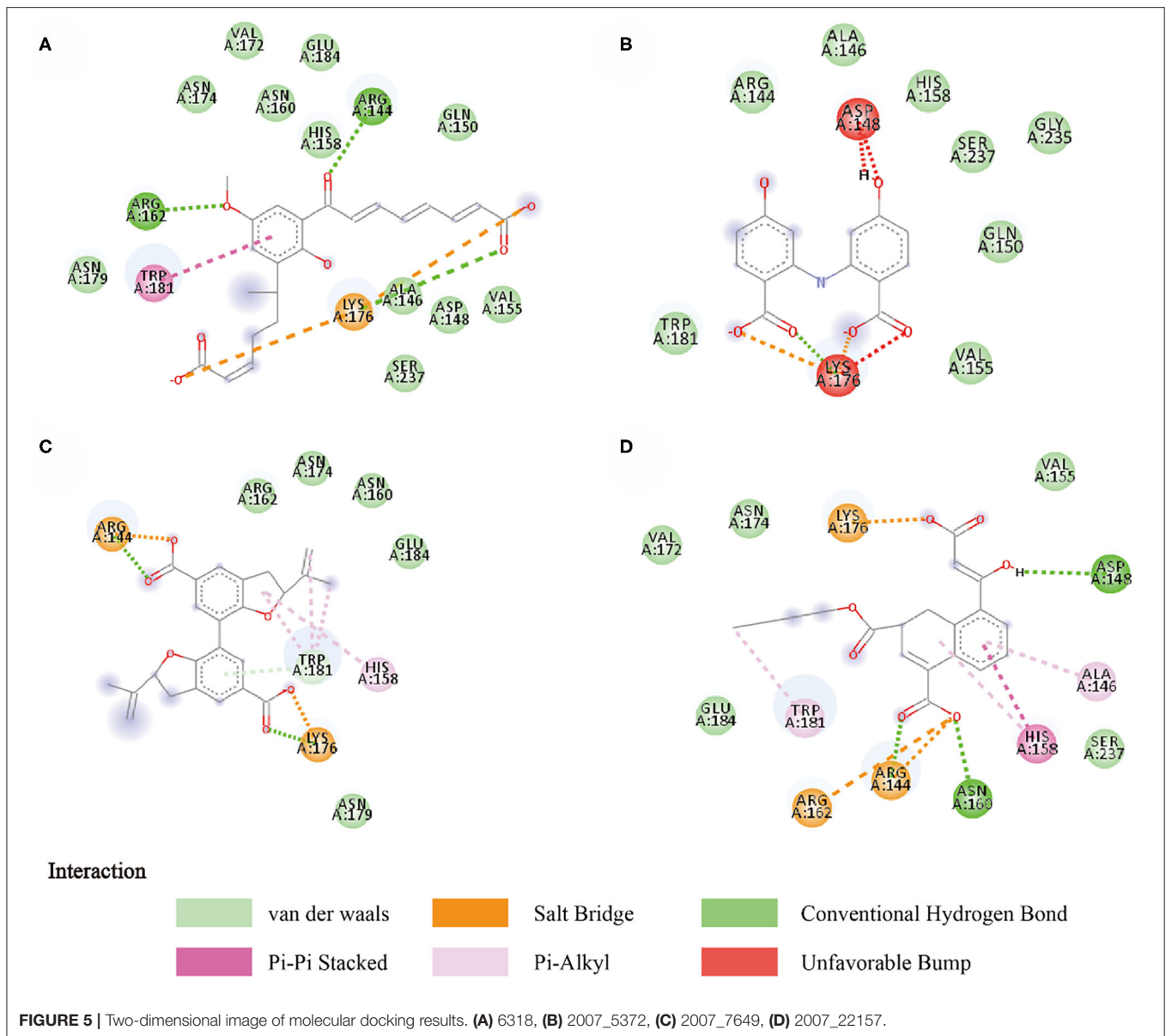
Prediction Results of Models

During the whole validation process, the prediction accuracy on testing sets is used to quantify the performance of different prediction algorithms. Both Mean Square Error (MSE) and square correlation coefficient (R-square) are set as the metric for prediction accuracy. MSE is the expectation of the square of the

difference between the predicted value and the true value, which is used to evaluate the predicted result. The smaller the MSE, the stronger the model's ability to fit the experimental data. R-square represents the quality of a fit through changes in data. The normal value range is 0 to 1. The closer to 1 indicates that the variable of the equation (input feature) has a stronger ability to explain Y (pIC50). **Table 2** lists the R square and MSE values of all prediction models.

Feature Selection

In this study, we obtained only 56 samples, but the feature dimension is as high as 204. Theoretically, this can easily lead machine learning models to get over-fitting. We used the following methods to perform feature dimensionality reduction to search for representative features. Firstly, Pearson correlation

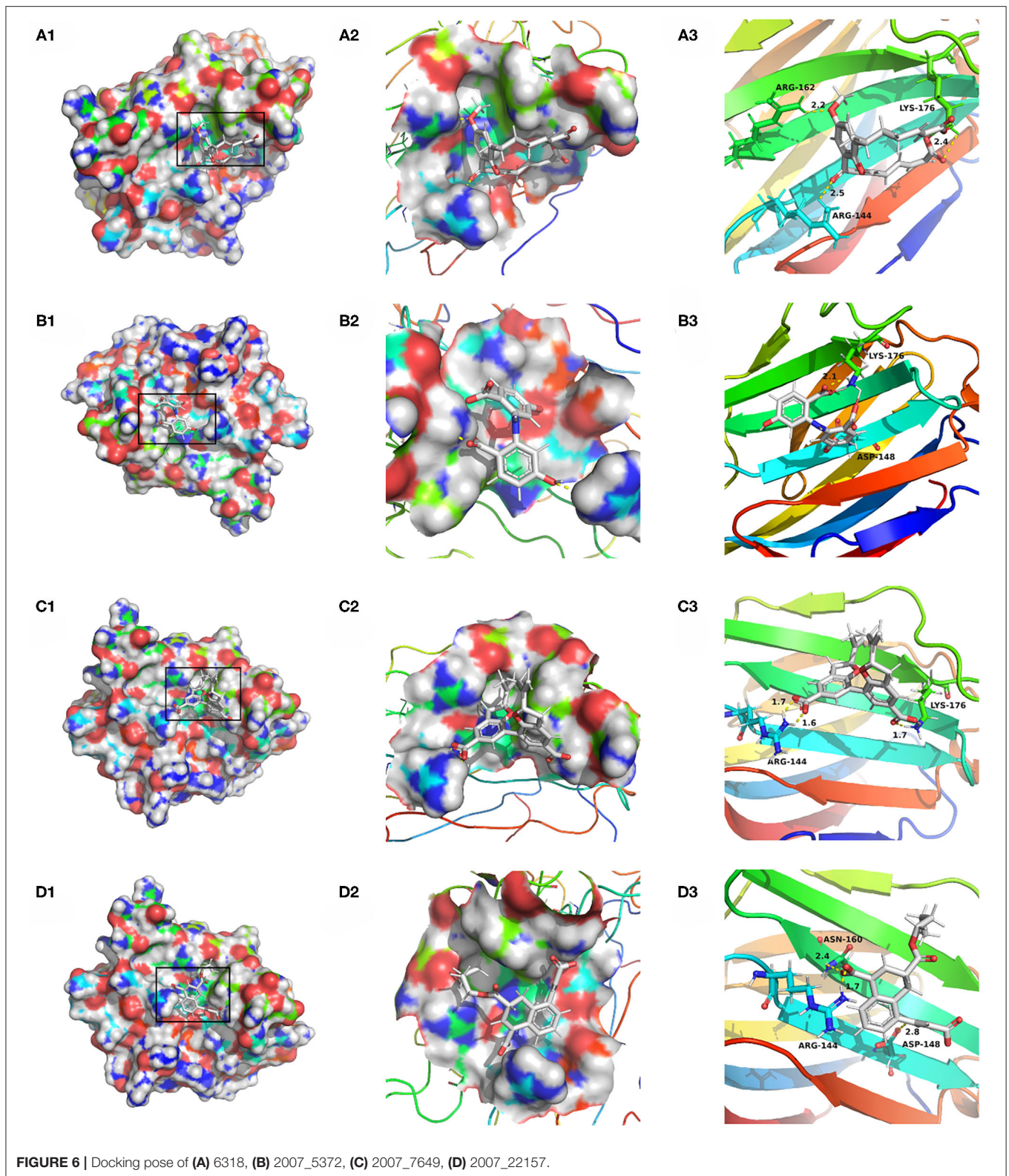


coefficient has very good applicability in the characterization of correlation. **Figure 7A** is a heat map of Pearson correlation coefficients between all features, the deeper the red, the stronger the correlation. It can usually be considered to have a strong correlation when the correlation is >0.9 , and parameters would be eliminated. Then, principal component analysis (2D PCA and 3D PCA) achieved dimensionality reduction by integrating the original single variable to obtain a new set of comprehensive variables (**Figures 8A,B**). What's more, the elements with sample feature variance <0.05 were eliminated by calling the Variance Threshold library function, and the Lasso function was used to select the variables of the sample data based on the penalty method. Finally, the original coefficients were compressed, and the insignificant variables were directly discarded. **Figure 7B** is a heat map of Pearson correlation coefficients between the remaining 28 features after eigenvalue preprocessing. These 28 indicators were set as input features for all involved machine

learning models. Through a 5-fold cross-validation method, the predictive ability of the sample model is evaluated. Correlation between predicted values and actual values (pIC50) of ABR, RR, SVR, EN, RF, GBR, KNN, XGB models were shown in **Figure 9**.

RR and EN Model

For ridge regression, elastic net and lasso regression models, model regularization is introduced to reduce the over-fitting, but the method of restraining weight is different. The R-square of the RR model and the EN model on the training sets were 0.87 and 0.85, respectively, but the R-square of the EN model on the test sets was 0.84, which was significantly better than 0.79 of the RR model (**Figures 9B,D**). This might be because only a few key features are related to pIC50 in all features, and the RR model retains those irrelevant features, which increases the fitting error. In this study, the number of features was much larger than the



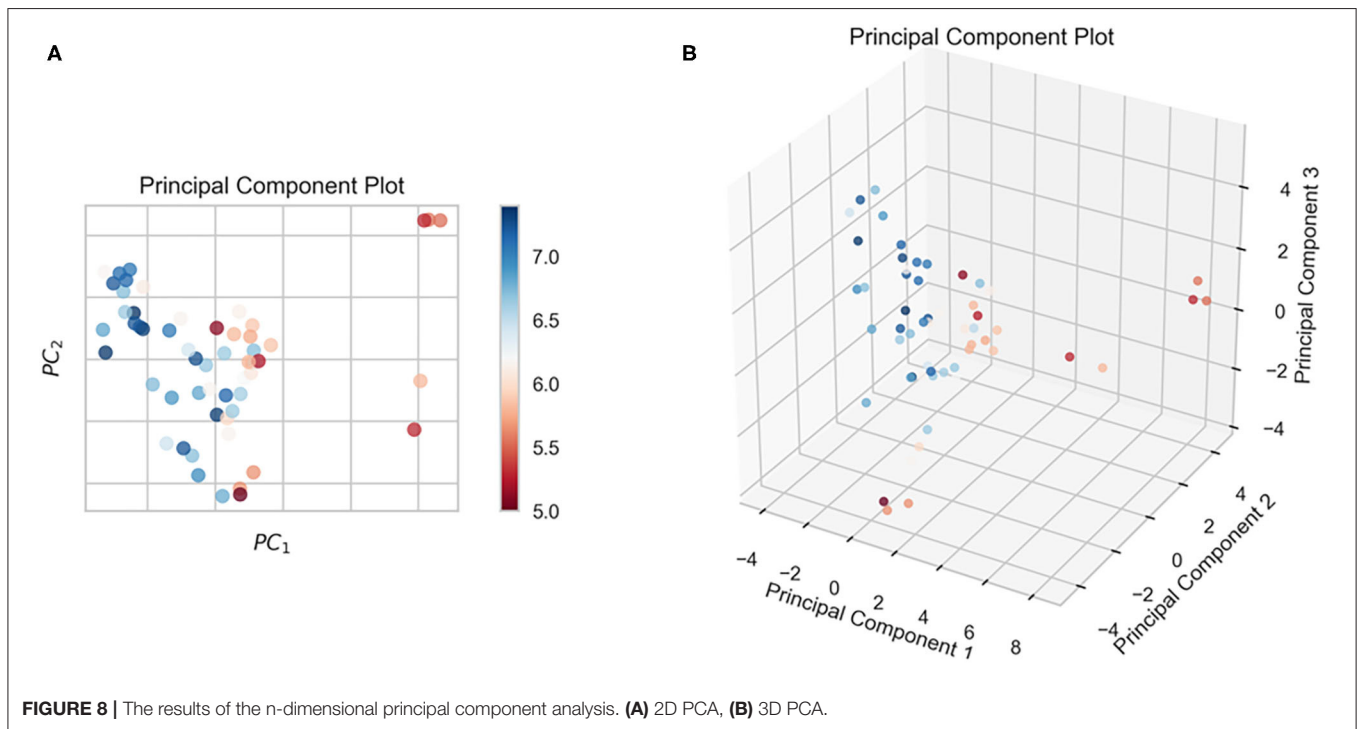
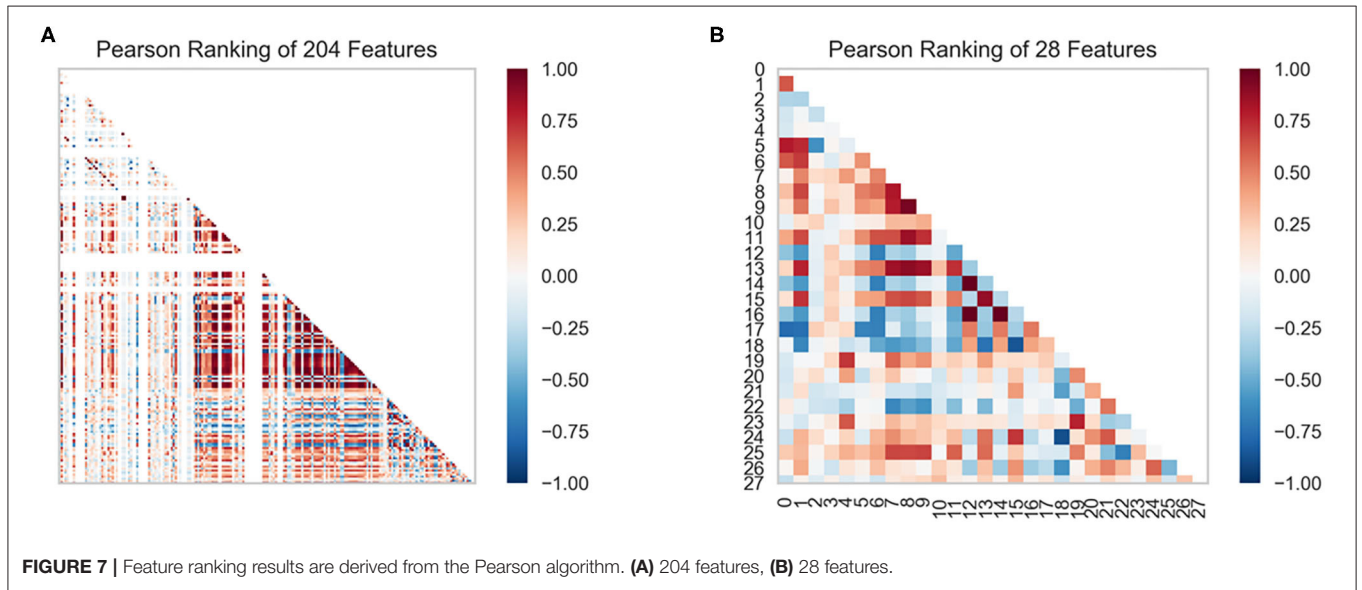
number of samples. The EN model could reduce the weight of non-key features to zero, and only retained a few key features, so it showed better fitting results.

ABR, RF, GBR, and XGB Model

The R-square of the ABR, RF, GBR, and XGB models on the training sets were all higher than 0.9 (Table 2), which was better

TABLE 2 | R-square and mean squared error (MSE) values of all trained models.

	ABR	RR	SVR	EN	RF	GBR	KNN	XGB	DBN
R2_train	0.990	0.869	0.850	0.851	0.920	0.991	1.000	0.970	0.935
R2_test	0.923	0.795	0.848	0.837	0.848	0.859	0.770	0.901	0.900
MSE_train	0.002	0.056	0.064	0.064	0.036	0.001	0.000	0.014	0.025
MSE_test	0.027	0.087	0.064	0.069	0.052	0.049	0.080	0.034	0.045



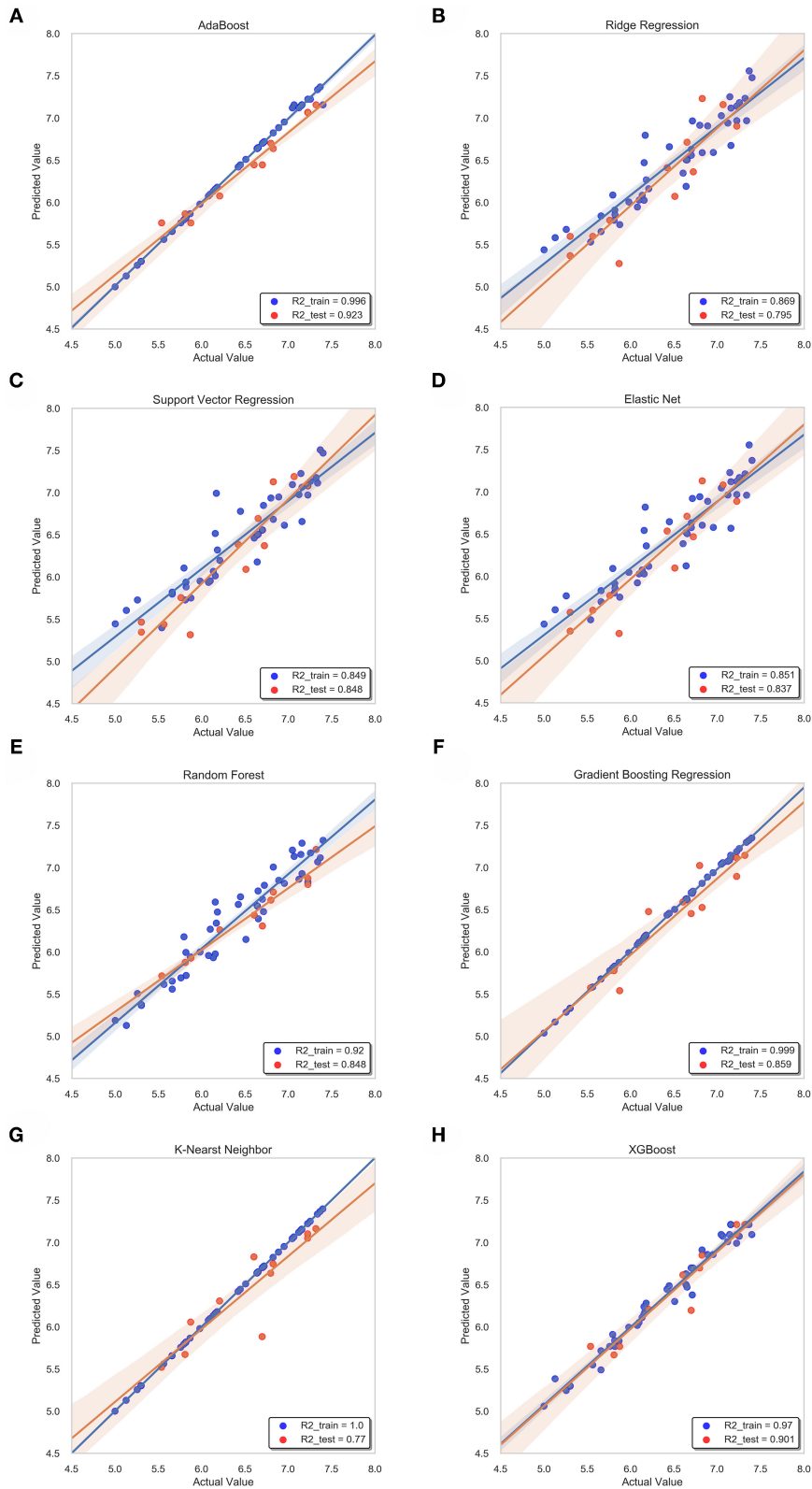
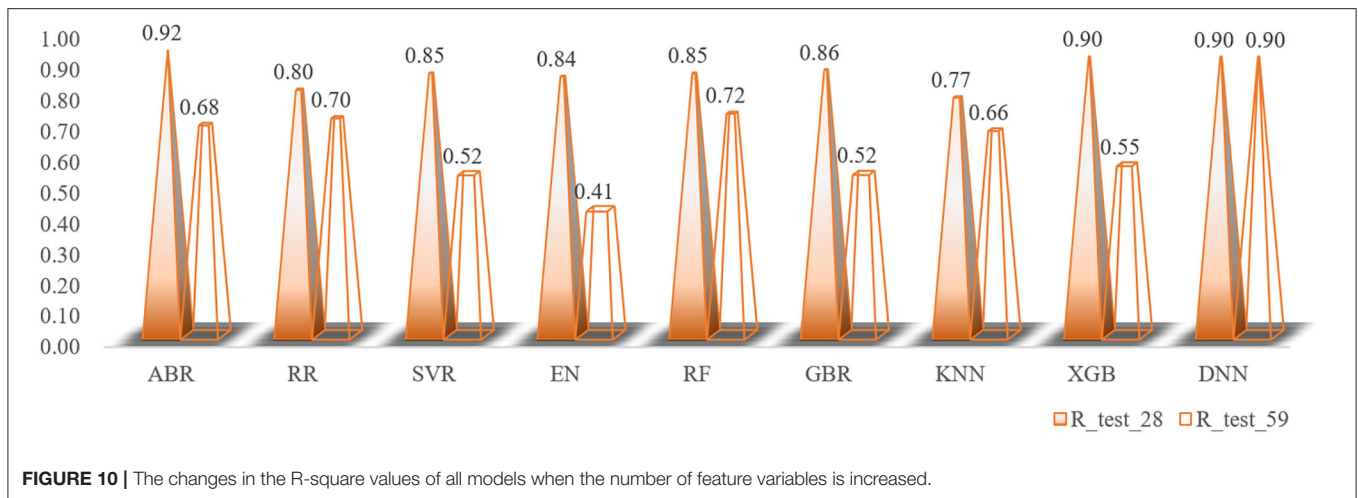


FIGURE 9 | Correlation between predicted value and actual value of machine learning model. Training set (blue dots) and testing set (red dots) are shown. The confidence interval is 95%. **(A)** ABR, **(B)** RR, **(C)** SVR, **(D)** EN, **(E)** RF **(F)** GBR, **(G)** KNN, **(H)** XGB.



than the RR and EN models. This was because they all used integrated learning methods and used different strategies to combine individual learners into a committee, which improved the generalization ability of the model. Based on bagging integration method, the prediction result of RF was obtained by parallel calculation of all decision trees. It made the model susceptible to large deviations due to the influence of individual wrong decision trees, so the R-square on the training sets was only 0.92 (Figure 9E). The R-square of the ABR and GBR models on the training sets were both close to 1 (Figures 9A,F). This was most likely because they used the Boosting method to upgrade the weak learner to a strong learner and the sample size was relatively small. ABR continuously optimized the sample weights of each round of training, the trained model with the testing sets mean square error (MSE) of 0.027 and R-square of 0.92 (Table 2), which showed the best performance. XGB added a regular term to the cost function that reduced the variance of the model and avoided overfitting. With the help of the variable importance indicator `feature_importances_`, we obtained the weight of the importance of the XGB model variable. It could be seen that `ALogP_MR` and `ES_Count_aaC` were the top two eigenvalues. The distribution of predicted values around actual values of the XGB model were plotted in Figure 9H, and the R-square on the training sets and the test sets reached 0.97 and 0.9, respectively.

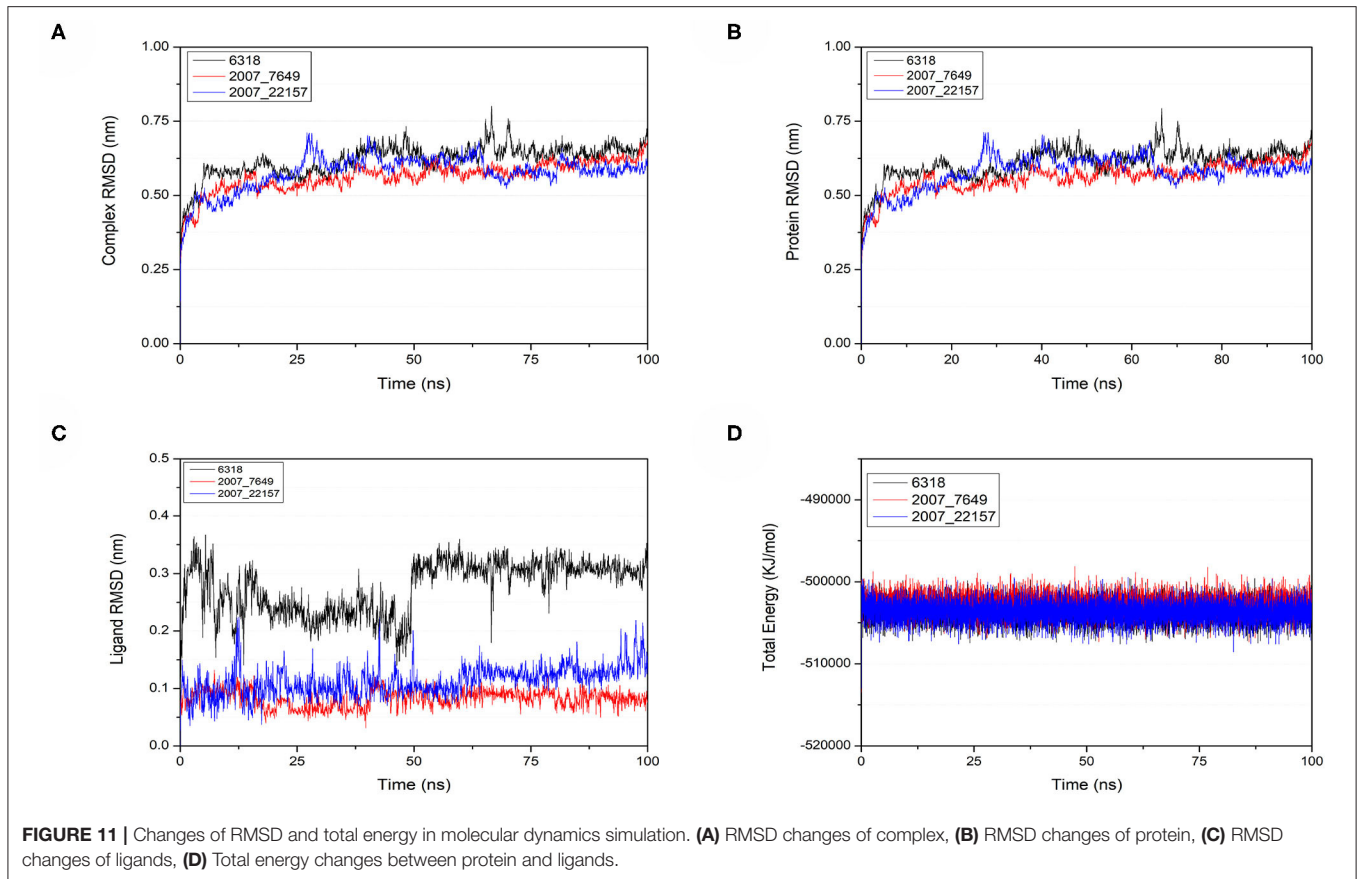
SVR and KNN Model

Before ensemble learning and neural network algorithms showed superior performance, SVM algorithm basically occupied a dominant position, especially in the field of classification. For Gal3 inhibitor studies, there were actually fewer samples available for reference. Even so, since SVR basically did not involve probability measurement and the law of large numbers, the model still showed high prediction accuracy, and the R-square on the test sets was close to 0.85 (Figure 9C). However, the shortcoming of SVR was also exposed. When the feature dimension was much larger than the number of samples, the SVR model tended to ignore the correlation of mutual characteristics, so the R-square on the training sets was only 0.85. After

repeated calculation and verification, the linear kernel was finally selected, and the error tolerance was set to 0.39 to ensure that the model had sufficient generalization ability to avoid overfitting. The fitting curve of the KNN model was perfect, and the R-square on the training sets had reached 1, which was suspected of overfitting. The number *k* of favorable features was filtered through the built-in function of Scikit-learn, and it was verified that the model performs best when *k* was equal to 4. When the independent variable dimension was small, the KNN model could significantly reduce the error. But as the variable gradually increased, the mean value of the dependent variable corresponding to the closest value to the target might deviate from the actual value exponentially. The above reasons caused the model accuracy to be significantly reduced. The R-square on the test sets was only 0.77, which was the worst performance among all models (Table 2).

Deep Belief Network

The total number of sample data we obtained was only 56, far less than the number of 204 feature dimensions. In view of this situation, traditional machine learning models usually removed most of the features during the preprocessing process. Although a model with higher prediction accuracy could be obtained, this actually consumed the credibility of the model. DBN could avoid this problem as much as possible, allowing more features to participate in training. Although it took thousands of times the training time, the trained model had better prediction reliability and more convincing. Using the Dropout method, the dropout rates of the first, second and third layers of the neural network were 0.4, 0.6, and 0.3, respectively, and a total of 500 modeling attempts were made. The R-square of the best trained model on the training sets and test sets were 0.94 and 0.9, respectively. We had tried in the previous machine learning model. When the parameter threshold was set to 0.01 and alpha was set to 0.001, 59 features are retained. However, the accuracy of all models was significantly reduced. The R-square on some model test sets was even lower than 0.5 (Figure 10). Through comparison, it could

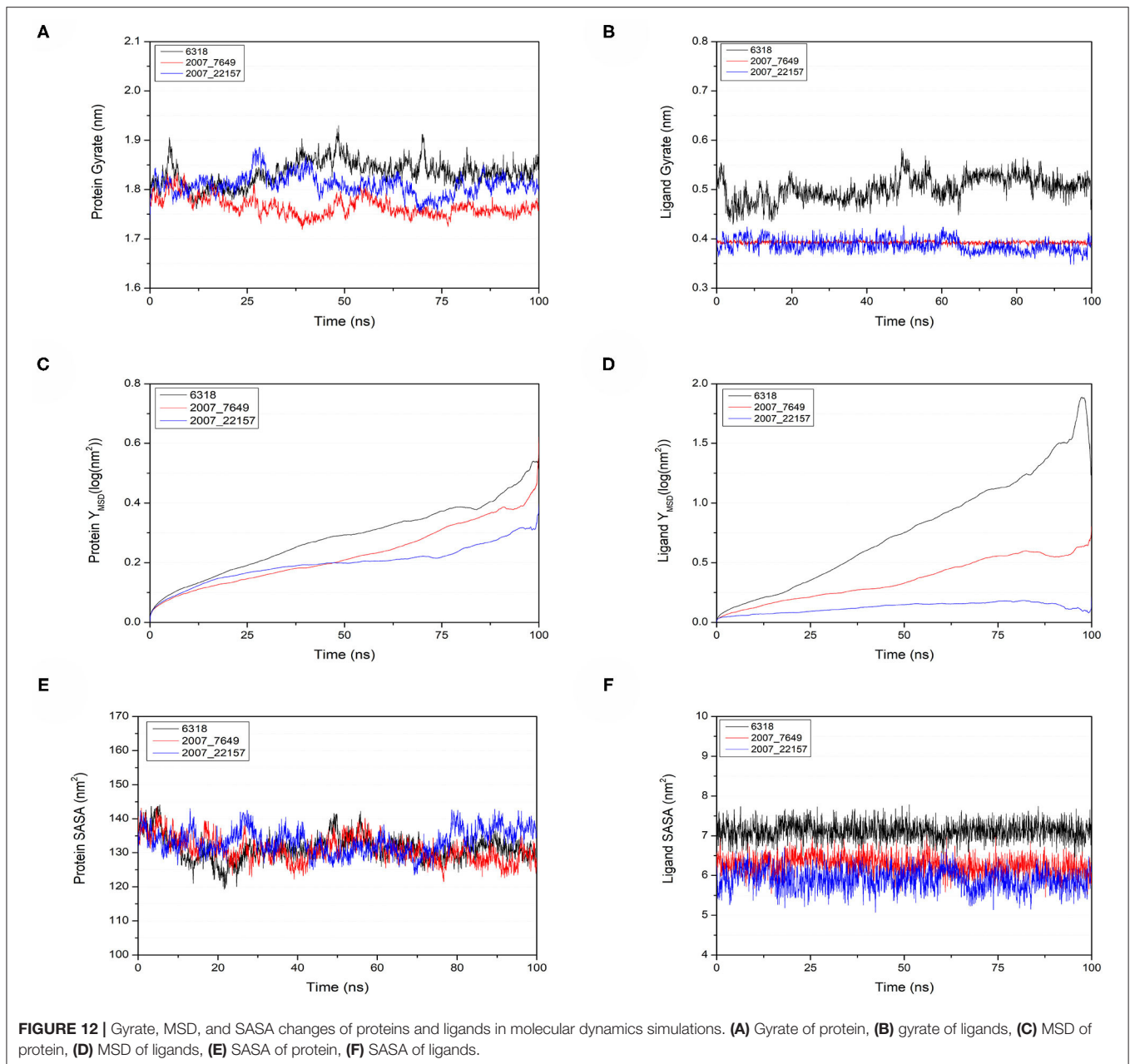


be found that the DBN model still showed high accuracy under low sample set and high feature dimension.

Molecular Dynamics Simulation

To verify the binding stability of the receptor-ligand complexes, we performed molecular dynamics simulations in 100 ns with Gromacs 2018 software (Kutzner et al., 2019). Unfortunately, 5372 was detached from the binding site during the molecular dynamic simulation, although it showed good results in the molecular docking and AI-based models prediction process. The escape of 5372 indicated that its binding stability to Gal3 was poor. The RMSD was calculated to evaluate the deviation of the structure from the original starting structure over the course of the simulations. In the RMSD results (**Figures 11A–C**), the RMSD values of the three candidates and Gal3 complexes shown an upward trend at initial 10 ns, then tended to stabilize with a relatively flat curve and maintained around 0.5–0.6 nm. The protein RMSD change curve is similar to the complex RMSD change curve. Based on ligand RMSD, 6318 had higher fluctuation rate than other candidates, which may explain high ligand gyration value. The RMSD results suggested that candidates 7649 and 22157 have higher binding stability to Gal3 protein. Besides, the total energy of simulation systems in the 100 ns process was calculated to analyze the energy changes in the complexes. And the results shown that the energy of protein-ligand complexes was stable and it had been maintained between

about $-510\ 000$ to $-500\ 000$ kJ/mol (**Figure 11D**). The radius of gyration could give a measure of the compactness of the structures, and can also give a measure of the atomic mass relative the molecular center mass. As shown in **Figures 12A,B**, protein gyrate and ligands gyrate were stable in general during the MD simulation process, and the gyrate of 6318 was higher than other two ligands (include target protein system), which was consistent with the RMSD results. MSD revealed the movement of atoms from the initial position to the final stage of MD simulation, indicating the movement trend of each ligand or protein. The low and stable MSD value of the ligand shows the stability of the binding, while the decrease of the MSD value indicates that the ligand may be close to the binding pocket. The extremely high MSD value and the increasing MSD value mean the ligand has a tendency to escape. In the process of the whole simulation (**Figures 12C,D**), although the MSD values of Gal3 and three ligands both increased, the changes of all measured MSD values maintained in a low range. It was worth noting that 6318 obtained much higher ligand MSD value compared to other two ligands, which indicates the ligand 6318 have a trend of escaping from the binding pocket. From the SASA calculated results, we can analyze the hydrophilicity and hydrophobicity of the simulation system. The solvent accessible area of Gal3 decreased significantly from 0 to 15 ns, and then remained to a relatively stable area (**Figure 12E**). Meanwhile, as shown in **Figure 12F**, the SASA values of all ligands were very stable in the simulation process.

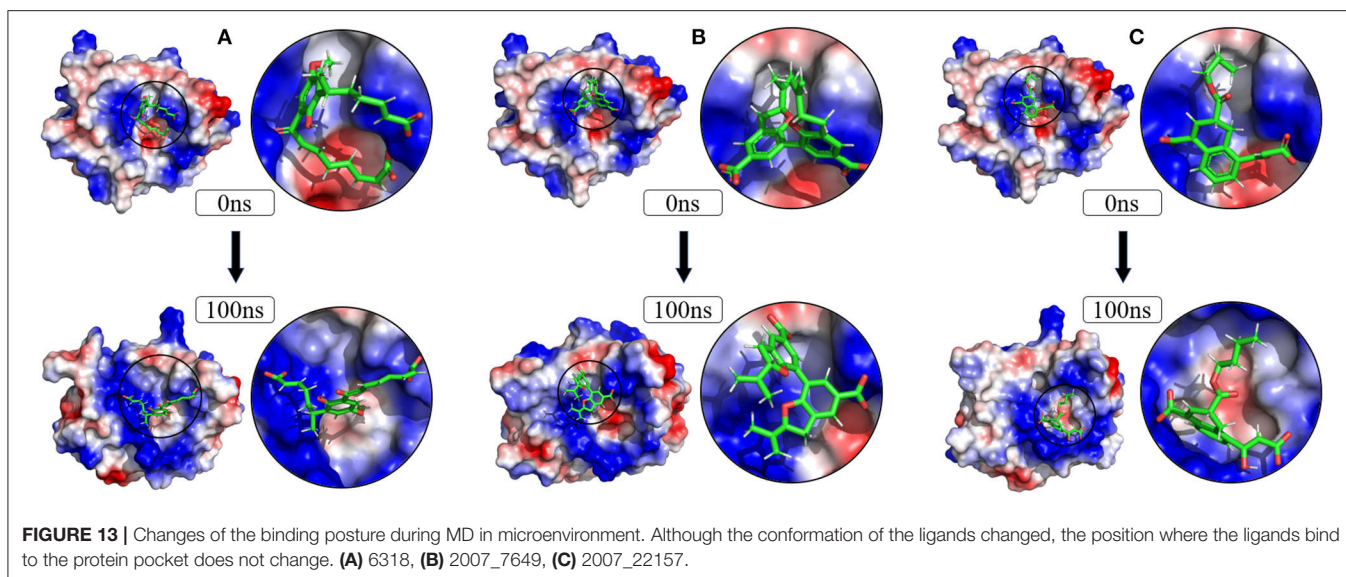


This can further indicate the stability of the complex systems. Finally, the binding postures of three ligands and target protein in the initial and final conformations were displayed in **Figure 13**. The ligands rotated in the same pocket but not detached from the target protein.

CONCLUSION

The new discovery of Gal3 in the pathogenesis of HD provides a new target and new method for ND treatment. In order to find potential inhibitors of Gal3, we have completed the following work and obtained satisfactory results. Using molecular docking methods, we initially screened a batch of small molecules with

relatively stable docking from a large drug molecule library. Multiple artificial intelligence-based models were constructed, and known Gal3 inhibitors were used as sample sets to train the models. From the performance parameters of the model, all models achieved high overall accuracy sensitivity. The R-square of XGBoost model on the test sets was higher than other algorithms, and there was no overfitting on the training sets. We not only screened from the TCM database, but also used the ZINC database to do the same. The results showed that the molecules from the TCM database performed better than the ZINC database in terms of binding stability and pIC50 value predicted by AI models. Comparing the prediction results of all models, we completed further screening and



narrowed the candidate range. Finally, through MD simulation, we further verified the stability of the complexes, the final candidate ligand and the target protein complexes showed stable binding throughout the simulation time. Combining all experimental results, the active ingredients 1,2-Dimethylbenzene and Typhic acid contained in *Crataegus pinnatifida* and *Typha angustata* may become the new drug formulation for ND treatment. We provide a new strategy with applying AI-based methods to the drug screening process, which can greatly reduce the cost of new drug development. Screening drug molecules from the TCM database is an innovation and beneficial supplement to screening from the general database. In summary, this study has explored a highly accurate integrated architecture to reduce the drug screening process. With the application of artificial intelligence, medical practitioners exclude candidates with low probability based on prediction results, which reduces the risk of downstream decision-making for better resource planning and allocation. The proposed integration method shows high accuracy under different algorithm models, indicating that artificial intelligence-based drug development has application prospects. Artificial intelligence-based application is an improvement and supplement to the existing traditional drug screening based on molecular interaction relationships.

REFERENCES

- Abdel-Basset, M., El-Shahat, D., El-henawy, I., de Albuquerque, V. H. C., and Mirjalili, S. (2020). A new fusion of grey wolf optimizer algorithm with a two-phase mutation for feature selection. *Expert Syst. Appl.* 139:112824. doi: 10.1016/j.eswa.2019.112824
- Abdolmaleki, A., Ghasemi, J. B., and Ghasemi, F. (2017). Computer aided drug design for multi-target drug design: SAR/QSAR, molecular docking and pharmacophore methods. *Curr. Drug Targets* 18, 556–575. doi: 10.2174/1389450117666160101120822
- Agarap, A. F. (2018). Deep learning using rectified linear units (ReLU) arXiv. arXiv 7.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SJ and CC designed research. LD, WZ, LZ, XH, and ZL worked together to complete the experiment and analyze the data. SJ, CC, LD, and WZ wrote the manuscript together. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the National Nature Science and Foundation of China Grand No. 71801031, in part by the Guangdong Basic and Applied Basic Research Foundation project, China, No. 2019A1515011962, in part by the Fundamental Research Funds for the Central Universities No. 20lgpy183, the Science, Technology & Innovation Commission of Shenzhen Municipality [JCYL 20170818165305521], and in part by the Guangzhou science and technology fund [Grant No. 201803010072].

- Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized random forests. *Ann. Stat.* 47, 1148–1178. doi: 10.1214/18-AOS1709
- Bateman, A., Martin, M. J., O'Donovan, C., Magrane, M., Alpi, E., Antunes, R., et al. (2017). UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 45, D158–D169. doi: 10.1093/nar/gkw1099
- Boza-Serrano, A., Ruiz, R., Sanchez-Varo, R., Garcia-Revilla, J., Yang, Y., Jimenez-Ferrer, I., et al. (2019). Galectin-3, a novel endogenous TREM2 ligand, detrimentally regulates inflammatory response in Alzheimer's disease. *Acta Neuropathol.* 138, 251–273. doi: 10.1007/s00401-019-02013-z
- Breiman, L. (2001). Random forests. *Machine Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

- Brooks, B. R., Brooks, C. L. III, Mackerell, A. D. Jr., Nilsson, L., Petrella, R. J., Roux, B., et al. (2009). CHARMM: the biomolecular simulation program. *J. Comput. Chem.* 30, 1545–1614. doi: 10.1002/jcc.21287
- Burley, S. K., Berman, H. M., Bhikadiya, C., Bi, C., Chen, L., Di Costanzo, L., et al. (2019). RCSB protein data bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.* 47, D464–D474. doi: 10.1093/nar/gky1004
- Chang, C.-C., and Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2:27. doi: 10.1145/1961189.1961199
- Chen, C. Y. C. (2011). TCM database@Taiwan: the world's largest traditional Chinese medicine database for drug screening *in silico*. *PLoS ONE* 6:5. doi: 10.1371/journal.pone.0015939
- Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., and Blaschke, T. (2018a). The rise of deep learning in drug discovery. *Drug Discov. Today* 23, 1241–1250. doi: 10.1016/j.drudis.2018.01.039
- Chen, X., Huang, L., Xie, D., and Zhao, Q. (2018b). EGBMMDA: extreme gradient boosting machine for MiRNA-disease association prediction. *Cell Death Dis.* 9:3. doi: 10.1038/s41419-017-0003-x
- Correia, J., Resende, T., Baptista, D., and Rocha, M. (2020). “Artificial intelligence in biological activity prediction,” in *Practical Applications of Computational Biology and Bioinformatics*, eds F. FdezRiverola, M. Rocha, M. S. Mohamad, N. Zaki, and J. A. CastellanosGarzon, (Cham: Springer), 164–172. doi: 10.1007/978-3-030-23873-5_20
- Elmezayen, A. D., Al-Obaidi, A., Sahin, A. T., and Yeleki, K. (2020). Drug repurposing for coronavirus (COVID-19): *in silico* screening of known drugs against coronavirus 3CL hydrolase and protease enzymes. *J. Biomol. Struct. Dyn.* doi: 10.1080/07391102.2020.1758791. [Epub ahead of print].
- Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., et al. (2017). The ChEMBL database in 2017. *Nucleic Acids Res.* 45, D945–D954. doi: 10.1093/nar/gkw1074
- Henderson, N. C., and Sethi, T. (2009). The regulation of inflammation by galectin-3. *Immunol. Rev.* 230, 160–171. doi: 10.1111/j.1600-065X.2009.00794.x
- Hessler, G., and Baringhaus, K. H. (2018). Artificial intelligence in drug design. *Molecules* 23:2520. doi: 10.3390/molecules23102520
- Huang, Q., Chen, Y., Liu, L., Tao, D., and Li, X. (2020). On combining biclustering mining and AdaBoost for breast tumor classification. *IEEE Trans. Knowl. Data Eng.* 32, 728–738. doi: 10.1109/TKDE.2019.2891622
- Irwin, J. J., and Shoichet, B. K. (2005). ZINC—a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* 45, 177–182. doi: 10.1021/ci049714+
- Kaiser, T. M., Burger, P. B., Butch, C. J., Pelly, S. C., and Liotta, D. C. (2018). A machine learning approach for predicting HIV reverse transcriptase mutation susceptibility of biologically active compounds. *J. Chem. Inf. Model.* 58, 1544–1552. doi: 10.1021/acs.jcim.7b00475
- Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2019). PubChem 2019 update: improved access to chemical data. *Nucleic Acids Res.* 47, D1102–D1109. doi: 10.1093/nar/gky1033
- Kumar, D., Kumari, K., Jayaraj, A., Kumar, V., Kumar, R. V., Dass, S. K., et al. (2020). Understanding the binding affinity of noscapines with protease of SARS-CoV-2 for COVID-19 using MD simulations at different temperatures. *J. Biomol. Struct. Dyn.* 1–14. doi: 10.1080/07391102.2020.1752310
- Kumar, R., Ignjatovic, M. M., Peterson, K., Olsson, M., Leffler, H., Ryde, U., et al. (2019). Structure and energetics of ligand-fluorine interactions with galectin-3 backbone and side-chain amides: insight into solvation effects and multipolar interactions. *Chemmedchem* 14, 1528–1536. doi: 10.1002/cmde.2019.00293
- Kutzner, C., Pall, S., Fechner, M., Eszternmann, A., de Groot, B. L., and Grubmueller, H. (2019). More bang for your buck: improved use of GPU nodes for GROMACS 2018. *J. Comput. Chem.* 40, 2418–2431. doi: 10.1002/jcc.26011
- Lai, T., Chen, R., Yang, C., Li, Q., Fujita, H., Sadri, A., et al. (2020). Efficient robust model fitting for multistructure data using global greedy search. *IEEE Trans. Cybernetics* 50, 3294–3306. doi: 10.1109/TCYB.2019.2900096
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539
- Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., et al. (2020). Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. *Radiology* 296, E65–E71. doi: 10.1148/radiol.2020200905
- Li, P., Liu, S., Lu, M., Bandyopadhyay, G., Oh, D., Imamura, T., et al. (2016). Hematopoietic-derived galectin-3 causes cellular and systemic insulin resistance. *Cell* 167, 973–984.e12. doi: 10.1016/j.cell.2016.10.025
- Liu, R., Chen, P., Li, X., Wu, Z., Gao, X., Chen, X., et al. (2017). Artificial intelligence sense technology: new technology in pharmaceutical sciences. *Chin. J. Pharm. Anal.* 37, 559–567. doi: 10.16155/j.0254-1793.2017.04.01
- Liu, X.-Y., Lv, X., Wang, P., Ai, C.-Z., Zhou, Q.-H., Finel, M., et al. (2019). Inhibition of UGT1A1 by natural and synthetic flavonoids. *Int. J. Biol. Macromol.* 126, 653–661. doi: 10.1016/j.ijbiomac.2018.12.171
- Macdonald, M. E., Ambrose, C. M., Duyao, M. P., Myers, R. H., Lin, C., Srinidhi, L., et al. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on huntington's disease chromosomes. The huntington's disease collaborative research group. *Cell* 72, 971–983. doi: 10.1016/0092-8674(93)90585-E
- Mayr, A., Klambauer, G., Unterthiner, T., and Hochreiter, S. (2016). DeepTox: toxicity prediction using deep learning. *Front. Environ. Sci.* 3:80. doi: 10.3389/fenvs.2015.00080
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on alzheimer's disease. *Neurology* 34, 939–944. doi: 10.1212/WNL.34.7.939
- Mills, N. (2006). ChemDraw ultra 10.0. *J. Am. Chem. Soc.* 128, 13649–13650. doi: 10.1021/ja0697875
- Pant, S., Singh, M., Ravichandiran, V., Murty, U. S. N., and Srivastava, H. K. (2020). Peptide-like and small-molecule inhibitors against Covid-19. *J. Biomol. Struct. Dyn.* 1–10. doi: 10.1080/07391102.2020.1757510
- Rahman, M. R., Islam, T., Zaman, T., Shahjaman, M., Karim, M. R., Huq, F., et al. (2020). Identification of molecular signatures and pathways to identify novel therapeutic targets in Alzheimer's disease: insights from a systems biomedicine perspective. *Genomics* 112, 1290–1299. doi: 10.1016/j.ygeno.2019.07.018
- Ramirez Hernandez, E., Sanchez-Maldonado, C., Mayoral Chavez, M. A., Hernandez-Zimbron, L. F., Patricio Martinez, A., Zenteno, E., et al. (2020). The therapeutic potential of galectin-1 and galectin-3 in the treatment of neurodegenerative diseases. *Expert Rev. Neurother.* 20, 439–448. doi: 10.1080/14737175.2020.1750955
- Ramirez, E., Sanchez-Maldonado, C., Mayoral, M. A., Mendieta, L., Alariste, V., Patricio-Martinez, A., et al. (2019). Neuroinflammation induced by the peptide amyloid-beta (25-35) increase the presence of galectin-3 in astrocytes and microglia and impairs spatial memory. *Neuropeptides* 74, 11–23. doi: 10.1016/j.npep.2019.02.001
- Ratsch, G., Onoda, T., and Muller, K. R. (2001). Soft margins for AdaBoost. *Machine Learn.* 42, 287–320. doi: 10.1023/A:1007618119488
- Romero, A., and Gabius, H. J. (2019). Galectin-3: is this member of a large family of multifunctional lectins (already) a therapeutic target? *Expert Opin. Ther. Targets* 23, 819–828. doi: 10.1080/14728222.2019.1675638
- Saijo, K., Crotti, A., and Glass, C. K. (2010). Nuclear receptors, inflammation, and neurodegenerative diseases. *Adv. Immunol.* 106, 21–59. doi: 10.1016/S0065-2776(10)06002-5
- Sapay, N., and Tieleman, D. P. (2011). Combination of the CHARMM27 force field with united-atom lipid force fields. *J. Comput. Chem.* 32, 1400–1410. doi: 10.1002/jcc.21726
- Schneider, P., Walters, W. P., Plowright, A. T., Sieroka, N., Listgarten, J., Goodnow, R. A., et al. (2020). Rethinking drug design in the artificial intelligence era. *Nat. Rev. Drug Discov.* 19, 353–364. doi: 10.1038/s41573-019-0050-3
- Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., et al. (2020). Improved protein structure prediction using potentials from deep learning. *Nature* 577, 706–710. doi: 10.1038/s41586-019-1923-7
- Shin, T. (2013). The pleiotropic effects of galectin-3 in neuroinflammation: a review. *Acta Histochem.* 115, 407–411. doi: 10.1016/j.acthis.2012.11.010
- Siew, J. J., Chen, H.-M., Chen, H.-Y., Chen, H.-L., Chen, C.-M., Soong, B.-W., et al. (2019). Galectin-3 is required for the microglia-mediated brain inflammation in a model of huntington's disease. *Nat. Commun.* 10:3473. doi: 10.1038/s41467-019-11441-0
- Song, W., Yen, H.-Y., Robinson, C. V., and Sansom, M. S. P. (2019). State-dependent lipid interactions with the A2a receptor revealed by MD simulations using *in vivo*-mimetic membranes. *Structure* 27, 392–403.e3. doi: 10.1016/j.str.2018.10.024

- Venkatachalam, C. M., Jiang, X., Oldfield, T., and Waldman, M. (2003). LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. *J. Mol. Graph. Model.* 21, 289–307. doi: 10.1016/S1093-3263(02)00164-X
- Voet, S., Srinivasan, S., Lamkanfi, M., and van Loo, G. (2019). Inflammasomes in neuroinflammatory and neurodegenerative diseases. *Embo Mol. Med.* 11:e10248. doi: 10.15252/emmm.201810248
- Wen, P., Hu, T.-G., Linhardt, R. J., Liao, S.-T., Wu, H., and Zou, Y.-X. (2019). Mulberry: a review of bioactive compounds and advanced processing technology. *Trends Food Sci. Technol.* 83, 138–158. doi: 10.1016/j.tifs.2018.11.017
- Xuan, S., Wu, Y., Chen, X., Liu, J., and Yan, A. (2013). Prediction of bioactivity of HIV-1 integrase ST inhibitors by multilinear regression analysis and support vector machine. *Bioorg. Med. Chem. Lett.* 23, 1648–1655. doi: 10.1016/j.bmcl.2013.01.081
- Yang, X., and Wen, W. (2018). Ridge and lasso regression models for cross-version defect prediction. *IEEE Trans. Reliability*, 67, 885–896. doi: 10.1109/TR.2018.2847353
- Yang, Y., Islam, M. S., Wang, J., Li, Y., and Chen, X. (2020). Traditional Chinese medicine in the treatment of patients infected with 2019-new coronavirus (SARS-CoV-2): a review and perspective. *Int. J. Biol. Sci.* 16, 1708–1717. doi: 10.7150/ijbs.45538
- Zhang, W., Jing, K., Huang, F., Chen, Y., Li, B., Li, J., et al. (2019). SFLN: a sparse feature learning ensemble method with linear neighborhood regularization for predicting drug-drug interactions. *Inf. Sci.* 497, 189–201. doi: 10.1016/j.ins.2019.05.017
- Zhang, Y.-D., Pan, C., Sun, J., and Tang, C. (2018). Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. *J. Comput. Sci.* 28, 1–10. doi: 10.1016/j.jocs.2018.07.003
- Zoete, V., Cuendet, M. A., Grosdidier, A., and Michielin, O. (2011). SwissParam: a fast force field generation tool for small organic molecules. *J. Comput. Chem.* 32, 2359–2368. doi: 10.1002/jcc.21816
- Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Series B Stat. Methodol.* 67, 301–320. doi: 10.1111/j.1467-9868.2005.00503.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Deng, Zhong, Zhao, He, Lian, Jiang and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.