

# SCIENTIFIC REPORTS



OPEN

## Improved Classification of Blood-Brain-Barrier Drugs Using Deep Learning

Rui Miao<sup>1</sup>, Liang-Yong Xia<sup>1</sup>, Hao-Heng Chen<sup>1</sup>, Hai-Hui Huang<sup>2</sup> & Yong Liang<sup>3</sup>

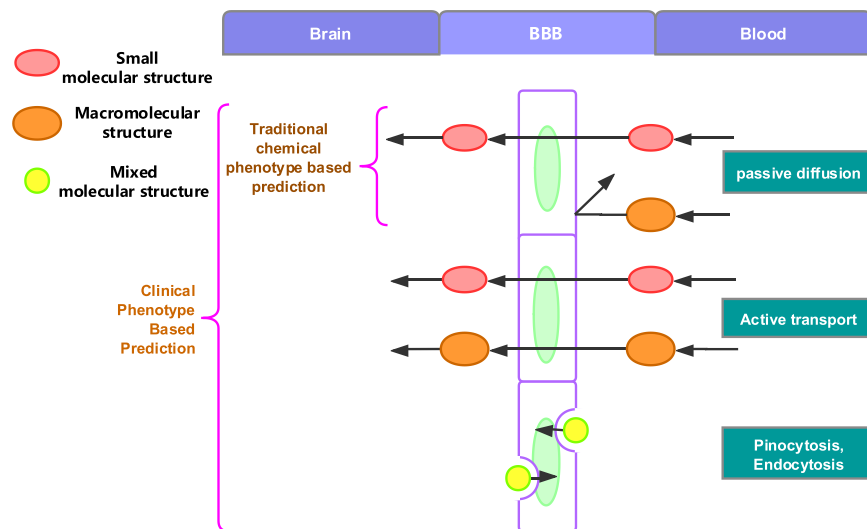
Blood-Brain-Barrier (BBB) is a strict permeability barrier for maintaining the Central Nervous System (CNS) homeostasis. One of the most important conditions to judge a CNS drug is to figure out whether it has BBB permeability or not. In the past 20 years, the existing prediction approaches are usually based on the data of the physical characteristics and chemical structure of drugs. However, these methods are usually only applicable to small molecule compounds based on passive diffusion through BBB. To deal this problem, one of the most famous methods is multi-core SVM method, which is based on clinical phenotypes about Drug Side Effects and Drug Indications to predict drug penetration of BBB. This paper proposed a Deep Learning method to predict the Blood-Brain-Barrier permeability based on the clinical phenotypes data. The validation result on three datasets proved that Deep Learning method achieves better performance than the other existing methods. The average accuracy of our method reaches 0.97, AUC reaches 0.98, and the F1 score is 0.92. The results proved that Deep Learning methods can significantly improve the prediction accuracy of drug BBB permeability and it can help researchers to reduce clinical trials and find new CNS drugs.

Currently, neurological diseases account for 28% of people with disabilities of all ages<sup>1</sup>. Despite the high prevalence associated with Central Nervous System (CNS) disease, effective medicines for these diseases are in scarcity. Researchers have done a lot of works on drug discovery. However, many tested compounds had failed due to lack of the ability to penetrate Blood-Brain-Barrier (BBB) rather than lack of potency, which made BBB get stuck in CNS drug discovery<sup>2-7</sup>. BBB is a special selective border with semi-permeability. This border can prevent certain substances (mostly harmful) from entering the brain tissues. BBB limit the passage of most of the external compounds (98%) to maintain CNS at the steady state<sup>8</sup>. Therefore, to determine a drug whether has BBB permeability is a pre-requirement of discovering CNS drugs<sup>9-12</sup>. Although the clinical experiment is the most accurate method of measuring BBB permeability<sup>13</sup>, it is difficult to do so due to the limitation of various types of drugs. Therefore, BBB permeability needs to be forecasted by the computer to save time and cost.

At present, the most widely used predictive methods are physical and chemical approaches, which mainly include topological polar surface area, hydrogen bond donors and acceptors, acidic and basic atomic number, ionization potential, silico methods and so on<sup>14-19</sup>.

Besides the physical and chemical methods, there are various supervised learning approaches, such as Support Vector Machine (SVM)<sup>20-23</sup>, Decision Tree (DT)<sup>24</sup> and K-Nearest Neighbor (KNN)<sup>25</sup> proposed for BBB drug prediction. In 2018, Wang *et al.* proposed a Silico prediction method which combines with Machine learning and resampling methods that can avoid imbalanced dataset and its accuracy of prediction reached 0.966<sup>26</sup>. All the methods mentioned above adopted physical or chemical features to train prediction models. In general, these methods only can be applied in small-molecule compounds that penetrate the BBB with passive diffusion. However, there are many molecules, such as glucose<sup>27,28</sup>, pass through BBB with more complex mechanism than passive diffusion which cannot be predicted (Fig. 1, right part). Therefore, to solve this problem, Gao *et al.* proposed a drug prediction method which based on drug side effects and drug indications<sup>29</sup>. This method basically solves the problem of drug entered brain with multi-mechanism and presents a new research direction of drug development for researchers (Fig. 1, left part). However, Gao only adopted multi-core SVM method and without

<sup>1</sup>Faculty of Information Technology, Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau, China. <sup>2</sup>School of Information Science and Engineering, Shaoguan University, No. 288, University Road, Zhenjiang District, Shaoguan City, Guangdong Province, China. <sup>3</sup>State Key Laboratory of Quality Research in Chinese Medicines, Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau, China. Correspondence and requests for materials should be addressed to Y.L. (email: [yliang@must.edu.mo](mailto:yliang@must.edu.mo))



**Figure 1.** Mechanisms of drugs passing BBB and the applicable scope of prediction methods<sup>29</sup>. The right part presents the blood vessel, which shows the mechanisms for drug passing BBB, and the left part is the brain, which shows the scope of clinical drug phenotype based and chemical feature based BBB permeability prediction methods<sup>29</sup>.

comparing the experiment results with other methods. What's more, the accuracy of multi-core SVM method only reached 0.76, AUC was 0.739 and F1 score was 0.76, which need to be improved urgently. For thousands of possible drugs, every 1% increase in accuracy can save a lot of drug clinical testing time. The 0.76-accuracy of existing SVM-based methods is far away to satisfy the realistic requirement.

This paper proposes a Deep Learning method in predicting the drug permeability of BBB which is based on clinical features. At present, Deep Learning method is widely used in the fields of image, sound, and text recognition, which have been achieved a majority of remarkable results. In recent years, some researchers have proposed many Deep Learning methods in the field of drug prediction and achieved excellent results<sup>30–32</sup>. However, the application of Deep Learning methods is still rare for the prediction of BBB permeability of CNS drugs. Therefore, our paper also tries to verify the Deep Learning method whether is effective in predicting the drug's BBB penetration based on clinical features. Compared with the existing methods, our method has the following advantages: (i) The average prediction accuracy of experiments with three datasets already achieved 0.97, the average AUC is 0.98, F1 score is 0.91. It significantly performed better than the multi-core SVM method, Decision Tree and the KNN method, which can help researchers save experiment time and discover new drugs. (ii) The accuracy, AUC and F1 scores of SVM methods with different datasets are fluctuated greatly, but the accuracy of the Deep Learning method, which proposed in this paper, is very stable and adaptable. (iii) The Deep Learning method can be applied in both simple diffusions of small molecule compounds and other compounds that diffuse through complex pathways. In summary, this paper proposes a Deep Learning method in drug prediction of BBB permeability which is based on the clinical features and our results are better than the previous researchers' results like multi-core SVM methods. In the future, we will experiment with more types of drug data and hope our method can be applied in different disease.

The remaining sections organized Section III introduces the datasets and how we established them. Section IV is talking about the Deep Learning methods, which is design for predicting BBB permeability. Section V is the performance analyses which compared the Deep Learning method with multi-core SVM, KNN and Decision Tree (DT) on the three datasets. In section VI, it is the discussion of the advantages of the Deep Learning method proposed in this paper. Section VII concludes and describes future work.

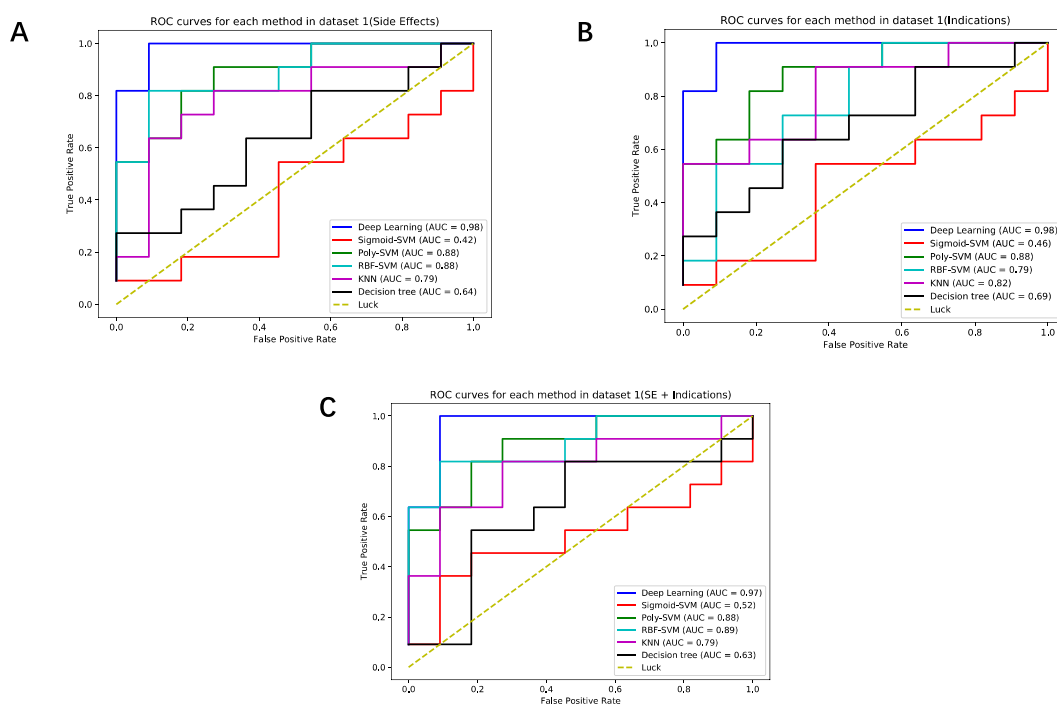
## Results

The experiments compare the Deep Learning method with Sigmoid-Support Vector Machine (Sigmoid-SVM), POLY-Support Vector Machine (POLY-SVM), Radial Basis Function-Support Vector Machine (RBF-SVM), K-Nearest Neighbor (KNN) and Decision Tree (DT). We tested the three datasets independently. Each test randomly assigned 1000 samples into mutually exclusive training sets (70%) and validation sets (30%). We also adopt 5-fold cross validation of the training datasets and validation datasets.

We adopt several evaluation methods to ensure the precision of the results. First, we calculate the accuracy on the training and validation datasets to evaluate the learning methods. However, the accuracy is not always valid for evaluating the learning performance in different situations, especially when the true and false samples of the dataset have large difference. Then we calculate the F1 score which is an indicator used in statistics to measure the accuracy of binary classification models, and we also consider the models' accuracy and the recall rate. Finally, in order to judge the performance of the learning models intuitively, we draw the ROC curve (Receiver Operating Characteristic curve) and calculate the AUC of the ROC curve (Area under the Curve of ROC). We also calculated all the indicators for the training and prediction datasets. Because the results analysis only requires the

Prediction performance in dataset 1	Method	Sigmoid-SVM			Poly-SVM			RBF-SVM		
		Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications
	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	
AUC	0.42 ± 0.08	0.79 ± 0.12	0.5252 ± 0.13	0.792 ± 0.12	0.81 ± 0.18	0.84 ± 0.09	0.88 ± 0.57	0.798 ± 0.21	0.84 ± 0.11	
Accuracy	0.495 ± 0.15	0.51 ± 0.11	0.52 ± 0.17	0.72 ± 0.1	0.63 ± 0.14	0.73 ± 0.08	0.89 ± 0.46	0.77 ± 0.11	0.74 ± 0.11	
F1	0.481 ± 0.17	0.5514 ± 0.12	0.607 ± 0.08	0.584 ± 0.14	0.37 ± 0.21	0.58 ± 0.11	0.76 ± 0.31	0.641 ± 0.15	0.73 ± 0.11	
Method	KNN			Decision Tree (DT)			Deep Learning			
	Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications	
AUC	0.795 ± 0.11	0.82 ± 0.08	0.791 ± 0.13	0.64 ± 0.15	0.69 ± 0.16	0.633 ± 0.09	<b>0.98 ± 0.02</b>	<b>0.98 ± 0.01</b>	<b>0.979 ± 0.02</b>	
Accuracy	0.806 ± 0.08	0.71 ± 0.13	0.74 ± 0.11	0.574 ± 0.12	0.69 ± 0.14	0.58 ± 0.12	<b>0.96 ± 0.02</b>	<b>0.965 ± 0.03</b>	<b>0.96 ± 0.02</b>	
F1	0.712 ± 0.07	0.68 ± 0.09	0.717 ± 0.09	0.567 ± 0.12	0.661 ± 0.09	0.568 ± 0.12	<b>0.902 ± 0.06</b>	<b>0.891 ± 0.09</b>	<b>0.904 ± 0.04</b>	

**Table 1.** Predictive performance comparisons with different learning methods in Dataset 1. Each test data field shows average ± std of 1000 random splits of training and test data.



**Figure 2.** (A) Drug-side-effects ROC curves with different methods in the validation part of Dataset 1. (B) Indication ROC curves with different methods in the validation part of Dataset 1. (C) Drug-side-effects (SE) + indications ROC curves with different methods in the validation part of Dataset 1.

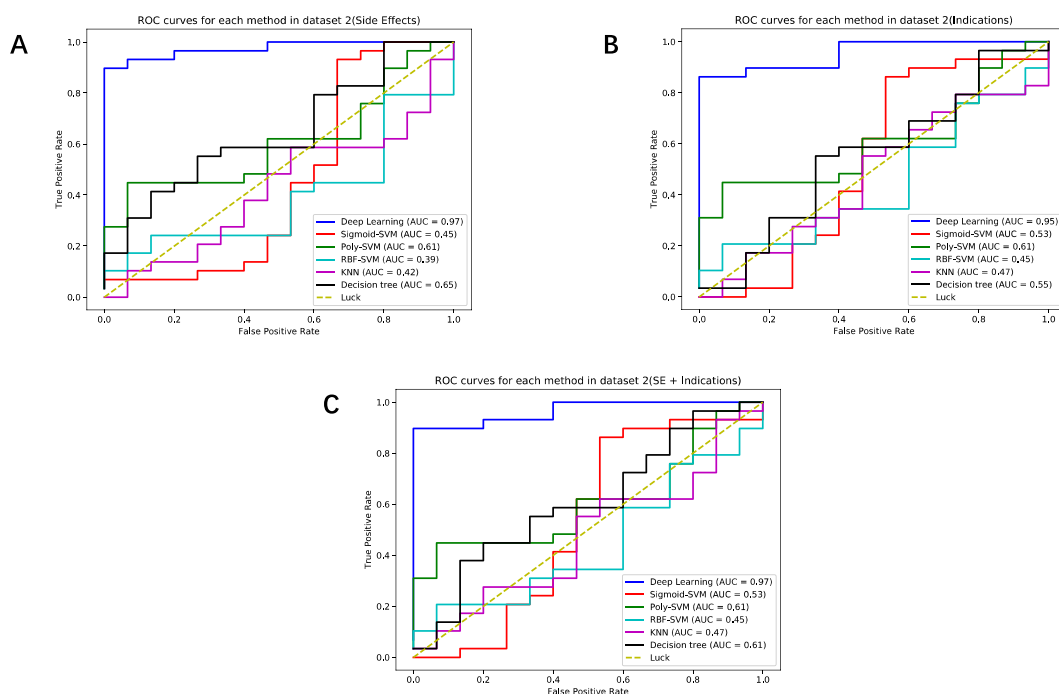
results of the predicted dataset, we did not list the results of the training dataset in the manuscript. The detailed results of dataset 1–3 are shown in the Supplementary Table S1

**Predictive performance of different methods with Dataset 1 and Dataset 2.** In this section, first of all, we established Datasets 1 and 2 and validated the performance of different learning methods with them, which based on drug's side effects, drug's indications and drug's side effects (SE) + indications. Then, we collected and analyzed the results of each individual test. Table 1 and Fig. 2 are the experiment outputs of Dataset 1 with different methods. According to Table 1, besides the Deep Learning method, the RBF-SVM method achieved the best results and its accuracy is 0.84, the AUC is 0.84 and the F1 score is 0.73. However, the performance of Deep Learning method is the best, the AUC increases by 13.9%, the accuracy increases by 12% and the F1 score increases by 17.4%. Therefore, the results show that Deep Learning method has better performance than the other methods on the experiments with Dataset 1.

In order to further verify the performance of the learning models, we do the experiments with Dataset 2, which has a larger sample number. The predictive performance of Dataset 2 is shown in Table 2. The drug-side-effects, indications' and drug-side-effects + indications' ROC curves of Dataset 2 are shown in Fig. 3. The experimental

Prediction performance in dataset 2	Method	Sigmoid-SVM			Poly-SVM			RBF-SVM		
		Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications
		Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction
AUC		0.452 ± 0.08	0.53 ± 0.09	0.535 ± 0.13	0.6111 ± 0.08	0.61 ± 0.07	0.66 ± 0.09	0.398 ± 0.13	0.45 ± 0.09	0.45 ± 0.08
Accuracy		0.51 ± 0.11	0.691 ± 0.13	0.52 ± 0.11	0.602 ± 0.13	0.591 ± 0.11	0.59 ± 0.24	0.41 ± 0.12	0.432 ± 0.13	0.45 ± 0.11
F1		0.474 ± 0.13	0.51 ± 0.1	0.5221 ± 0.08	0.5201 ± 0.14	0.51 ± 0.12	0.47 ± 0.23	0.531 ± 0.14	0.512 ± 0.08	0.41 ± 0.13
Method	Method	KNN			Decision Tree (DT)			Deep Learning		
		Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications
		Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction
AUC		0.421 ± 0.15	0.47 ± 0.18	0.472 ± 0.15	0.56 ± 0.15	0.51 ± 0.12	0.51 ± 0.12	<b>0.97 ± 0.02</b>	<b>0.9523 ± 0.03</b>	<b>0.971 ± 0.02</b>
Accuracy		0.415 ± 0.16	0.445 ± 0.14	0.51 ± 0.15	0.57 ± 0.11	0.52 ± 0.15	0.52 ± 0.15	<b>0.9621 ± 0.02</b>	<b>0.9235 ± 0.06</b>	<b>0.968 ± 0.03</b>
F1		0.541 ± 0.09	0.535 ± 0.17	0.53 ± 0.21	0.54 ± 0.15	0.56 ± 0.16	0.45 ± 0.25	<b>0.9008 ± 0.06</b>	<b>0.889 ± 0.08</b>	<b>0.911 ± 0.05</b>

**Table 2.** Predictive performance comparisons with different learning methods in Dataset 2. Each test data field shows average ± std of 1000 random splits of training and test data.



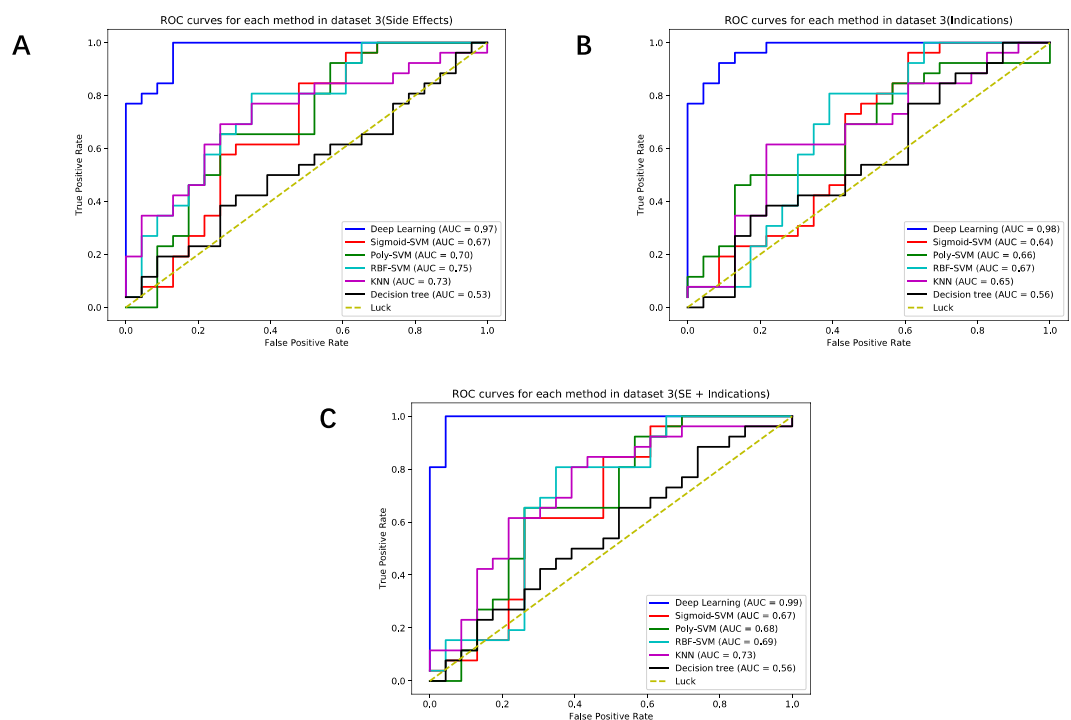
**Figure 3.** (A) Drug-side-effects ROC curves with different methods in the validation part of Dataset 2. (B) Indication ROC curves with different methods in the validation part of Dataset 2. (C) Drug-side-effects (SE) + indications ROC curves with different methods in the validation part of Dataset 2.

results show that the larger the sample number of a dataset, the larger difference between the results, then the advantage of the Deep Learning method is more clear. Compared Deep Learning method with the POLY-SVM method which is performing best among the other method, the AUC increases by 31%, the accuracy increases by 44.8%, and the F1 score increases by 44.1%. More experimental details of Datasets 1 and 2 are shown in the Supplementary Tables S2 and S3.

**Deep Learning method achieved higher performance in the Independent Dataset.** The third experiment is using the Independent Dataset (Dataset 3), because its result can produce more accurate and more objective performance assessment. The results of different methods on the Independent Dataset are shown in Table 3. The ROC curves of Drug-side-effects, Indication, and Drug-side-effects (SE) + indications on the Independent Dataset are shown in Fig. 4 respectively. According to Table 3 and Fig. 4, we knew that the Deep Learning method still has the best performance. Compared Deep Learning method with the best performing KNN method among the other methods, the AUC increases by 25.6%, the accuracy increases by 24%, and the F1 increases 22.6%. More experimental details of the Independent Dataset (Dataset 3) are shown in the Supplementary Table S4.

Prediction performance in dataset 3	Method	Sigmoid-SVM			Poly-SVM			RBF-SVM		
		Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications
	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	
AUC		0.675 ± 0.15	0.72 ± 0.18	0.672 ± 0.16	0.7058 ± 0.19	0.71 ± 0.15	0.685 ± 0.12	0.71 ± 0.12	0.71 ± 0.16	0.69 ± 0.15
Accuracy		0.51 ± 0.05	0.64 ± 0.18	0.64 ± 0.02	0.671 ± 0.14	0.53 ± 0.13	0.54 ± 0.11	0.735 ± 0.18	0.63 ± 0.12	0.69 ± 0.11
F1		0.42 ± 0.06	0.542 ± 0.21	0.584 ± 0.09	0.521 ± 0.16	0.51 ± 0.2	0.41 ± 0.06	0.601 ± 0.21	0.62 ± 0.15	0.561 ± 0.16
Method	KNN	Decision Tree (DT)			Deep Learning			Side Effects(SE)	Indications	SE + Indications
		Side Effects(SE)	Indications	SE + Indications	Side Effects(SE)	Indications	SE + Indications			
	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	Prediction	
AUC		0.73 ± 0.03	0.71 ± 0.11	0.734 ± 0.16	0.568 ± 0.16	0.67 ± 0.18	0.568 ± 0.12	<b>0.978 ± 0.02</b>	<b>0.98 ± 0.02</b>	<b>0.99 ± 0.01</b>
Accuracy		0.7 ± 0.11	0.72 ± 0.15	0.74 ± 0.19	0.584 ± 0.2	0.66 ± 0.18	0.59 ± 0.18	<b>0.978 ± 0.02</b>	<b>0.964 ± 0.03</b>	<b>0.98 ± 0.02</b>
F1		0.73 ± 0.07	0.71 ± 0.11	0.692 ± 0.11	0.516 ± 0.13	0.64 ± 0.16	0.519 ± 0.15	<b>0.91 ± 0.09</b>	<b>0.896 ± 0.1</b>	<b>0.918 ± 0.08</b>

**Table 3.** Predictive performance comparisons with different learning methods in Independent Dataset (Dataset 3). Each test data field shows average ± std of 1000 random splits of training and test data.



**Figure 4.** (A) Drug-side-effects ROC curves with different methods in the validation part of Independent Dataset (Dataset 3). (B) Indication ROC curves with different methods in the validation part of Independent Dataset (Dataset 3). (C) Drug-side-effects (SE) + indications ROC curves with different methods in the validation part of Independent Dataset (Dataset 3).

**Inter dataset validation.** To verify the versatility of the Deep Learning method, we also performed inter dataset validation. We used Dataset 2 as the training dataset and Dataset 1 as the validation dataset. The inter dataset validation results are shown in Table 4.

The results show that the Deep Learning method proposed in this paper achieves the ideal effect. The optimal accuracy is 0.97, the AUC is 0.98, and the F1 score is 0.92. It proves that the Deep Learning method has the best versatility among different datasets.

There is a brief summary of the experiments that the prediction accuracy of the Deep Learning method is very stable and always between 0.96 to 0.98. In addition, the best-processing method which besides the Deep Learning method is different in each dataset. At the same time, the fluctuation of accuracy is obvious which is influenced by the difference between the number of samples and the datasets and the number of positive and negative samples. What's more, the AUC and F1 score of the Deep Learning method also remain at a relatively high level. We also performed inter dataset validation to demonstrate the versatility of the Deep Learning method. Therefore, under the conditions described in this paper, we think the performance of the Deep Learning method is better than other existing methods.

Side Effects (SE)		Indications		SE + Indications	
Training	Prediction	Training	Prediction	Training	Prediction
<b>Deep Learning</b>					
0.975 ± 0.02	0.940 ± 0.04	0.987 ± 0.01	0.980 ± 0.02	0.971 ± 0.02	0.930 ± 0.06
0.940 ± 0.05	0.901 ± 0.07	0.982 ± 0.01	0.971 ± 0.25	0.965 ± 0.03	0.910 ± 0.08
0.930 ± 0.04	0.890 ± 0.09	0.941 ± 0.09	0.920 ± 0.06	0.931 ± 0.06	0.900 ± 0.09

**Table 4.** Predictive performance comparisons with different learning methods in Inter dataset validation. Each test data field shows average ± std of 1000 random splits of training and test data.

## Discussion

Research on neurological diseases has a long history. These kinds of researches can cure neurological diseases. At present, most researchers are still using various data mining algorithms based on different chemical characteristics to predict drugs' BBB permeability<sup>17,18</sup>. To further improve the performance of drug prediction models, researchers are still experimenting with many new physical and chemical features such as 2D molecular descriptors and molecular fingerprints, and machine learning methods like Gaussian process, Synthetic Minority Oversampling Technique (SMOTE) and SMOTE + edited nearest neighbor<sup>19,33–36</sup>. In fact, to improve the predictive methods, scientists have tried more than 1,000 chemical descriptors, many of which rely on esoteric quantum chemical calculations, and it is difficult to obtain accurate data using existing techniques<sup>37</sup>. In addition to the reason of computational complexity, there are some situations that chemical features are not available, such as some drugs/biologic with no precisely defined structures and most of the nutrients, nutrients analogs and certain physiologically important macro-molecules which pass through BBB must with more complex biological active mechanisms<sup>27,28,38</sup>. According to the cases mentioned above, if a model is trained with passive diffusion of BBB agents, the accuracy of BBB penetration prediction will be low. On the other hand, scientists can neither predict the mechanism by which a drug penetrates the BBB, nor predict the applicability of the model without the support of elaborate *in vivo* experiments. In order to solve this problem, the researchers have also made many attempts, such as: trying to establish an *in vitro* model. This method will clarify the mechanism of BBB development and help researchers predict the BBB permeability of drugs<sup>39,40</sup>. However, these methods still cannot completely solve the problem that small molecule drugs cannot be predicted. In this case, the researchers considered using drug side effects and drug indication information to predict BBB penetrate which the advantage is that most drugs have undergone an extensive clinical application and accumulated a wealth of information. These kinds of methods can greatly broaden the prediction range of CNS drugs.

For a long time, researchers often overlooked the relation between the clinical phenotype and efficacy of CNS drugs. In order to cross this barrier, Gao *et al.* have proved that data mining methods can effectively connect these two features<sup>29</sup>. However, there still has a problem of prediction with data mining methods which is the accuracy relatively low which means that clinical researchers still need to spend more time and effort to verify the effectiveness of the drug.

We think that due to the difference of features based on physics and chemistry, the relation between drug side effects and adaptability is more abstract and deeper. That means traditional machine learning methods might not find the relation between data and results very efficiently, and that is the reason why the classification result is not ideal. However, basically, the characteristic of Deep Learning method is suitable for handling the data with abstract relation. To solve the problem of the small number of drugs clinical data, we try several Deep Learning Network with different depth. The results prove that these kinds of datasets are not suitable for very deep network and it requires us to build a moderate-size Deep Learning model. Therefore, the purpose of our research is trying to find out a novel classification method that can more effectively predict the drug BBB permeability based on the clinical phenotype. The experiment result validates our thought that we can get an effective relation between clinical performance and efficacy of drugs with an appropriate size and depth Deep Learning model. Because these relations are on a deep level, the results of general machine learning models are not ideal which can have better performance with Deep Learning model. The performance of Deep Learning method proposed in this paper has been proved by the experiment results that we can greatly improve the final classification results. We think the method proposed in this paper is very helpful for CNS drug calculation and saving time and cost of clinical trials.

Despite the Deep Learning method proposed in this paper has lots of advantages, it is worth noting that this method still cannot predict how the drug penetrates BBB. This is of great significance to biology. Because in this case, we cannot distinguish between the side effects and secondary effects caused by the penetration of the compound into the BBB. Therefore, in the future, we consider combining drug clinical phenotypic effects and drug chemical structure characteristics, determining the general route of drug penetration into BBB. For example, if a drug appears to be permeable in a clinical phenotype-based model and not permeable in a physical and chemical-based model, the drug may enter the body indirectly through other means.

## Conclusion

This paper proposes the Deep Learning method to predict the permeability of Blood-Brain-Barrier based on clinical phenotype. There are three datasets with independent testing and the experimental results show that the Deep Learning method performs better than multi-core SVMs, KNNs and Decision Trees. What's more, the prediction accuracy of CNS drugs with our Deep Learning method increases more than 15%. The Deep Learning method

	Sample number	BBB permeability (true)	BBB permeability (false)	Data sources
Dataset 1	91	38	53	SIDER; <sup>20</sup>
Dataset 2	210	136	74	SIDER; <sup>37,42–44,46</sup>
Independent Dataset	161	76	85	SIDER; <sup>29,20,37,42–44</sup>

**Table 5.** The number of samples, BBB permeability true or false and data sources of the three datasets.

proposed in this paper adopted the clinical phenotypic approach, which means that our method has wider applicable scope and can reduce the workload of many clinical trials of drugs.

## Materials and Methods

**Datasets of clinical drug phenotypes.** According to the existing literature, this paper collected the drug names and SIDER datasets which have been proved that have BBB permeability true or false in the clinic.

The SIDER (<http://sideeffects.embl.de/>) dataset is a public dataset which contains a large number of drug side effects and drug indications<sup>41</sup>. We extracted the characteristics of the drug from this dataset. There is no existing complete BBB-permeable dataset on the Internet currently, so we refer to the literature which published in 2016<sup>29</sup> that collected experimental datasets from other six academic papers<sup>20,37,42–45</sup>. Based on this drug dataset, we classify the drugs into two categories, one is BBB permeability true and the other one is BBB permeability false.

The clinical drug phenotypes (side effects and indications) in the SIDER database were formatted according to the Medical Dictionary for Regulatory Activities (MedDRA, <http://www.meddra.org/>). MedDRA divides the clinical phenotype into 5 levels: Lowest Level Term (LLT), Preferred Term (PT), High-Level Term (HLT), High-Level Group Terms (HLGT) and System Organ Classes (SOC). PT is a special descriptor and it includes the information about symptoms, therapeutic adaptability diagnosis and so on. According to the High-Level Group Terms for neurological diseases (HLGT), we selected 43 terms as clinical phenotypic characteristics of drugs and the details were listed in Supplementary Table S5. Each HLGT also contained specific side effects and indications (PT). Then, took each drug's number of matching times under each specific HLGT group as training features<sup>29</sup>. More details are listed in Supplementary Table S6.

In a brief summary, we had established three datasets. The first dataset was referring to Doniger *et al.* paper which published in<sup>20</sup> and this dataset had 91 samples in total, of which 38 samples were BBB permeability true and 53 samples were BBB permeability false. The second dataset was referring to the papers published from<sup>29,37,42–44,46</sup> and this dataset had 210 samples in total, of which 136 samples were BBB permeability true and 74 samples were BBB permeability false. However, there was an imbalance in the sample distribution of Dataset 1 and Dataset 2. To solve the lopsidedness of the sample number of these datasets, we established the third Independent Dataset. The third dataset had 161 samples totally, of which 76 samples were BBB permeability true and 85 samples were BBB permeability false. The basic information of these datasets was shown in Table 5. The details of these datasets were given in Supplementary Tables S7–S9. The drug Side Effects and Indication based on SIDER dataset were listed in Supplementary Tables S10 and S11.

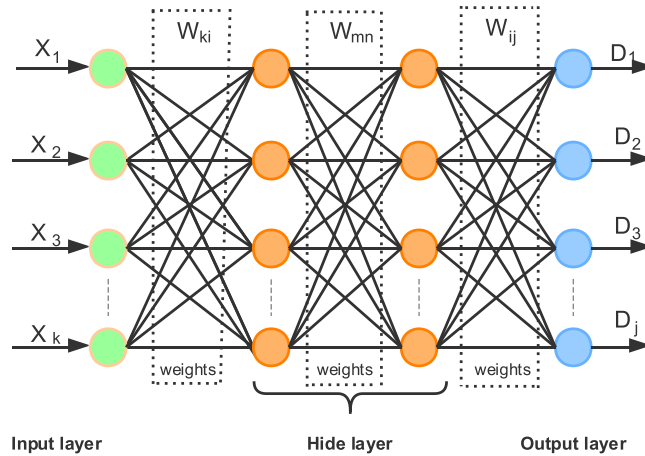
**System model of Deep Learning method.** Deep Learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These methods have already dramatically improved the state-of-the-art speech recognition, visual object recognition, object detection and many other domains such as drug discovery and genomics. Deep Learning is a model that can discover the more complicated structure of datasets by using the back-propagation algorithm. According to the discovered structure, the mode can change the internal parameters. The internal parameter of each layer is the result of the previous layer<sup>47</sup>. For different complex datasets, the number of layers required for Deep Learning is varied. We think that although the relation between clinical side effects and adaptability of drugs may be not so strong, there may have deeper relation between clinical expressiveness and final efficacy. That means clinical expressiveness will affect final efficacy. This relation is suitable for the main idea of Deep Learning which is trying to discover the deeper relation between the data through Multi-layer Network and back Propagation algorithm. Therefore, we try to establish a Deep Learning model to verify our thought. Based on the number of samples and dimensions of the drug datasets processed in this paper, we propose the four-layer Deep Learning model to deal with these datasets. The Deep Learning model which proposed in this paper is shown in Fig. 5.

**Hidden layer selection.** The number of nodes in the input layer and the output layer of the Deep Learning network. Here we calculated the number of nodes in the hidden layer using the following equation:

$$h = \sqrt{m + n} + \alpha \quad (1)$$

Where  $h$  is the number of hidden layer nodes,  $m$  is the number of input layer nodes,  $n$  is the number of output layer nodes,  $\alpha$  is an adjustment constant between 1 to 10, and generally,  $\alpha = 1$ .

**Forward pass subprocess.** Setting the weight between node  $i$  and node  $j$  is  $w_{ij}$ , the threshold of the node  $j$  is  $b_j$ , and the output value of each node is  $x_j$ . The output value of each node in the current layer is changed with the



**Figure 5.** The four-layer Deep Learning model constructed in this paper,  $x$  represents the data of each input node,  $D$  represents the data of each output node.  $W_{ki}$  is the weight between the input layer and the hidden layer,  $w_{mn}$  is the weight between the first hidden layer and second hidden layer and  $w_{ij}$  is the weight between the hidden layer and the output layer.

output value of all nodes in the previous layer. The weights and the thresholds of the nodes are implemented by an active function. The equations are as follows:

$$s_j = \sum_{i=0}^{m-1} w_{ij} + b_j \tag{2}$$

$$x_j = f(s_j) \tag{3}$$

where  $f$  is the active function represented by the sigmoid function, and its equation as following:

$$f(x) = \frac{A}{1 + e^{-\frac{x}{\beta}}} \tag{4}$$

The computation procedure is from top to bottom and then from left to right, and it needs to be observed strictly to finish the entire forward process.

*Reverse transfer subprocess.* After finishing the forward pass process, we need to construct the reverse transfer process. The most important thing in the reverse transfer process is the adjustment of the weights and thresholds between each adjacent layer. The specific adjustment steps are as follows:

Step 1. Assume that all results of the output layer are  $d_j$  and the equation of error function is as follows:

$$E(w, b) = \frac{1}{2} \sum_{j=0}^{n-1} (d_j - y_j)^2 \tag{5}$$

Step 2. According to the gradient descent method, the weights and thresholds of the functions are modified in several times in order to minimize the error function. The gradient of  $E(w, b)$  is divided by the correction of the weight vector at the current position. For the output node  $j$ :

$$\Delta w(i, j) = -\eta \frac{\partial E(w, b)}{\partial w(i, j)} \tag{6}$$

Step 3. In order to calculate the weights and thresholds between the hidden layer and the output layer, we derive the active function which represents by equation (4), then through equations (7) and (8) for  $w_{ij}$ , finally  $\delta_{ij}$  and  $b_j$  are calculated by the equations (9) and (10):

$$\begin{aligned} f'(x) &= \frac{Ae^{-\frac{x}{\beta}}}{B(1 + e^{-\frac{x}{\beta}})^2} \\ &= \frac{f(x)[A - f(x)]}{AB} \end{aligned} \tag{7}$$



$$\begin{aligned} \frac{\partial E(w, b)}{\partial w_{ij}} &= \frac{1}{\partial w_{ij}} \times \frac{1}{2} \sum_{j=0}^{n-1} (d_j - y_j)^2 \\ &= (d_j - y_j) \times f'(S_j) \times \frac{\partial S_j}{\partial w_{ij}} \\ &= (d_j - y_j) \times \frac{f(S_j)[A - f(S_j)]}{AB} \times \frac{\partial S_j}{\partial w_{ij}} \\ &= (d_j - y_j) \times \frac{f(S_j)[A - f(S_j)]}{AB} \times x_i \\ &= \delta_{ij} \times x_i \end{aligned} \tag{8}$$

$$\delta_{ij} = (d_j - y_j) \times \frac{f(S_j)[A - f(S_j)]}{AB} \tag{9}$$

$$\frac{\partial E(w, b)}{\partial b_j} = \delta_{ij} \tag{10}$$

Step 4. Calculate the thresholds between two hidden layers and between the input and hidden layers. In equations (11) and (12), we suppose that  $w_{mn}$  is the weight between the node  $m$  belongs to the first hidden layer and the node  $n$  belongs to the second hidden layer. The  $w_{ki}$  is the weight between the node  $K$  belongs to the input layer and the node  $i$  belongs to the hidden layer. The thresholds  $\delta_{ki}$  and  $\delta_{mn}$  are calculated by the equations (13) and (14):

$$\frac{\partial E(w, b)}{\partial w_{ki}} = \frac{1}{\partial w_{ki}} \times \frac{1}{2} \sum_{j=0}^{n-1} (d_n - y_n)^2 = \delta_{mn} \times x_m \tag{11}$$

$$\frac{\partial E(w, b)}{\partial w_{ki}} = \frac{1}{\partial w_{ki}} \times \frac{1}{2} \sum_{j=0}^{n-1} (d_i - y_i)^2 = \delta_{ki} \times x_k \tag{12}$$

$$\delta_{ki} = \sum_{j=0}^{n-1} \delta_{ki} \times w_{ki} \times \frac{f(S_k)[A - f(S_k)]}{AB} \tag{13}$$

$$\delta_{mn} = \sum_{j=0}^{n-1} \delta_{mn} \times w_{mn} \times \frac{f(S_m)[A - f(S_m)]}{AB} \tag{14}$$

Step 5. According to the gradient descent method and the formulas, which mentioned above, equations (15) and (16) are used to adjust the weights and thresholds between the hidden layer and the output layer. The equations (17) and (18) are used to adjust the weights and thresholds between two hidden layers. The equations (19) and (20) are used to adjust the weights and thresholds between the input layer and the hidden layer:

$$w_{ij} = w_{ij} - \eta \times \frac{\partial E(w, b)}{\partial w_{ij}} = w_{ij} - \eta_1 \times \delta_{ij} \times x_i \tag{15}$$

$$b_j = b_j - \eta_2 \times \delta_{ij} \tag{16}$$

$$w_{mn} = w_{mn} - \eta_1 \times \delta_{mn} \times x_{mn} \tag{17}$$

$$b_n = b_n - \eta_2 \times \delta_{mn} \tag{18}$$

$$w_{ki} = w_{ki} - \eta_1 \times \delta_{ki} \times x_k \tag{19}$$

$$b_i = b_i - \eta_2 \times \delta_{ki} \tag{20}$$

There is the whole procedure of the reverse transfer process in the Deep Learning method which is proposed in this paper. To complete the learning process of the entire Deep Learning network, the continuous adjustments of weights and thresholds are necessary. We can set an error threshold or a maximal number of cycles as a stop criterion to break off the entire learning process.

**Related methods for evaluation.** Nowadays, there are many usual methods in predicting the drug permeability of BBB, such as multi-core SVM, KNN, DT and so on. Therefore, we select several methods to compare with the Deep Learning method, which proposed in this paper in order to evaluate the performance of our method.

**Multi-core SVM method.** Multi-core SVM method is one of the most common methods in the published BBB permeability papers. For example, Gao *et al.* adopted POLY-SVM, RBF-SVM and normalized POLY-SVM methods in predicting the drug permeability of BBB<sup>29</sup>.

The SVM method assumes the hyperplane equation is  $w^T + b = 0$ . Let  $x$  be a vector of  $N$  dimensional input space. Let  $\phi(x) = (\varphi_1(x), \varphi_2(x), \dots, \varphi_M(x))$  denote the nonlinear transformation from the input space to the  $M$ -dimensional feature space. A superclass plane can be constructed in this feature space and the equation is<sup>48</sup>:

$$\sum_{j=1}^M w_j \phi_j(x) + b = 0 \quad (21)$$

where  $w_j$  is the weight that connects the feature space to the output space, and  $b$  is the offset.

If the data is not linearly separable, the kernel function will be used. The common kernel functions include Linear, Poly, RBF, Sigmoid and so on. Gao *et al.* paper proposed to use POLY-SVM, RBF-SVM and normalized POLY-SVM method in predicting the drug permeability of BBB which is based on clinical features<sup>29</sup>. However, in normalized POLY-SVM, the normalization only uses to preprocess the data and its influence on the results is slight. Therefore, we use another high performing method named Sigmoid-SVM method instead of normalized POLY-SVM in comparison.

**Drug prediction with KNN method.** KNN method is a kind of the classical data mining methods and it also has been used to predict drug penetration of BBB in many years.

KNN method is measuring the distance between different feature values. Its main idea is that if a sample in the feature space, most similar samples of  $K$ , which means the nearest neighbors in the feature space, belong to a certain category, then the sample also belongs to this category, where  $K$  is usually not greater than an integer of 20<sup>25</sup>.

**Drug prediction with Decision Tree (DT).** Decision Tree (DT) looks like the tree structure, which can be a binary tree or a non-binary tree. Each non-leaf node represents a feature attribute, each branch represents the output of the feature attribute in a range of values, and each leaf node stores a category<sup>27</sup>.

DT begins at the root node, then judge the corresponding feature in the item to be classified and selects the output branch according to its value until it reaches the leaf node. Finally, DT saved the category at the leaf node as the result of the decision<sup>49</sup>.

## References

1. Menken, M., Munsat, T. L. & Toole, J. F. The global burden of disease study: Implications for neurology. *Archives of Neurology* **57**, 418–420 (2000).
2. Pardridge, W. M. & Mietus, L. J. Transport of Steroid Hormones through the Rat Blood-Brain Barrier: PRIMARY ROLE OF ALBUMIN-BOUND HORMONE. *The Journal of Clinical Investigation* **64**, 145–154 (1979).
3. Harnish, P. P., Krutch, A. & Mukherji, M. Intravascular contrast media and the blood-brain barrier. Testing the new nonionic agent ioxilan. *Invest Radiol* **24**, 34–36 (1989).
4. Dieterich, H.-J. R., Reutershan, J. R., Felbinger, T. W. & Eltzhig, H. K. Penetration of Intravenous Hydroxyethyl Starch into the Cerebrospinal Fluid in Patients with Impaired Blood-Brain Barrier Function. *Anesthesia & Analgesia* **96**, 1150–1154 (2003).
5. Pardridge, W. M. The blood-brain barrier: Bottleneck in brain drug development. *NeuroRX* **2**, 3–14 (2005).
6. Saunders, N. R. *et al.* The rights and wrongs of blood-brain barrier permeability studies: a walk through 100 years of history. *Frontiers in Neuroscience* **8** (2014).
7. Hendricks, B., Cohen-Gadol, A. & C Miller, J. Novel delivery methods bypassing the blood-brain and blood-tumor barriers. *Neurosurgical Focus* **38**, E10 (2015).
8. Pardridge, W. M. Why is the global CNS pharmaceutical market so under-penetrated? *Drug Discovery Today* **7**, 5–7 (2002).
9. Davson, H. In *Implications of the Blood-Brain Barrier and Its Manipulation: Volume 1 Basic Science Aspects* (ed Edward A. Neuwelt) 27–52 (Springer US, 1989).
10. Esposito, P. *et al.* Corticotropin-Releasing Hormone and Brain Mast Cells Regulate Blood-Brain-Barrier Permeability Induced by Acute Stress. *Journal of Pharmacology and Experimental Therapeutics* **303**, 1061–1066 (2002).
11. Daneman, R. & Prat, A. The blood-brain barrier. *Cold Spring Harb Perspect Biol* **7**, a020412 (2015).
12. van Tellingen, O. *et al.* Overcoming the blood-brain tumor barrier for effective glioblastoma treatment. *Drug Resist Updat* **19**, 1–12 (2015).
13. Bickel, U. How to measure drug transport across the blood-brain barrier. *NeuroRX* **2**, 15–26 (2005).
14. Liu, X., Tu, M., Kelly, R. S., Chen, C. & Smith, B. J. Development of a computational approach to predict blood-brain barrier permeability. *Drug Metabolism and Disposition* **32**, 132–139 (2004).
15. Rautio, J., Laine, K., Gynther, M. & Savolainen, J. Prodrug Approaches for CNS Delivery. *The AAPS Journal* **10**, 92–102 (2008).
16. Cheng, F. *et al.* admetSAR: A Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties. *Journal of Chemical Information and Modeling* **52**, 3099–3105 (2012).
17. Kumar, R., Sharma, A. & Tiwari, R. K. Can we predict blood brain barrier permeability of ligands using computational approaches? *Interdisciplinary Sciences: Computational Life Sciences* **5**, 95–101 (2013).
18. Carpenter, T. S. *et al.* A method to predict blood-brain barrier permeability of drug-like compounds using molecular dynamics simulations. *Biophysical Journal* **107**, 630–641 (2014).
19. Vilar, S., Sobarzo-Sanchez, E., Santana, L. & Uriarte, E. Ligand and Structure-based Modeling of Passive Diffusion through the Blood-Brain Barrier. *Current Medicinal Chemistry* **25**, 1073–1089 (2018).
20. Doniger, S., Hofmann, T. & Yeh, J. Predicting CNS Permeability of Drug Molecules: Comparison of Neural Network and Support Vector Machine Algorithms. *Journal of Computational Biology* **9**, 849–864 (2002).
21. Pajouhesh, H. & Lenz, G. R. Medicinal chemical properties of successful central nervous system drugs. *NeuroRX* **2**, 541–553 (2005).
22. Zhang, D. *et al.* A Genetic Algorithm Based Support Vector Machine Model for Blood-Brain Barrier Penetration Prediction. *BioMed Research International* **2015**, 292683 (2015).

23. Jiang, L., Chen, J., He, Y., Zhang, Y. & Li, G. A method to predict different mechanisms for blood–brain barrier permeability of CNS activity compounds in Chinese herbs using support vector machine. *Journal of Bioinformatics and Computational Biology* **14**, 1650005 (2016).
24. Andres, C. & Hutter, M. C. CNS Permeability of Drugs Predicted by a Decision Tree. *QSAR & Combinatorial Science* **25**, 305–309 (2006).
25. Zhang, L., Zhu, H., Oprea, T. I., Golbraikh, A. & Tropsha, A. QSAR Modeling of the Blood–Brain Barrier Permeability for Diverse Organic Compounds. *Pharmaceutical Research* **25**, 1902 (2008).
26. Wang, Z. *et al.* In Silico Prediction of Blood–Brain Barrier Permeability of Compounds by Machine Learning and Resampling Methods. *ChemMedChem* **13**, 2189–2201 (2018).
27. Crone, C. Facilitated transfer of glucose from blood into brain tissue. *The Journal of physiology* **181**, 103–113 (1965).
28. Banks, W. A. The source of cerebral insulin. *European Journal of Pharmacology* **490**, 5–12 (2004).
29. Gao, Z., Chen, Y., Cai, X. & Xu, R. Predict drug permeability to blood–brain-barrier from clinical phenotypes: drug side effects and drug indications. *Bioinformatics* **33**, 901–908 (2017).
30. Gawehn, E., Hiss, J. A. & Schneider, G. Deep Learning in Drug Discovery. *Molecular Informatics* **35**, 3–14, <https://doi.org/10.1002/minf.201501008> (2016).
31. Zhang, L., Tan, J., Han, D. & Zhu, H. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug Discovery Today* **22**, 1680–1685, <https://doi.org/10.1016/j.drudis.2017.08.010> (2017).
32. Unterthiner, T. *et al.* In *Proceedings of the deep learning workshop at NIPS*. 1–9.
33. Wang, Z. *et al.* In Silico Prediction of Blood–Brain Barrier Permeability of Compounds by Machine Learning and Resampling. *Methods*. **13**, 2189–2201, <https://doi.org/10.1002/cmdc.201800533> (2018).
34. Raevsky, O. A., Grigorev, V. Y., Polianczyk, D. E., Raevskaja, O. E. & Dearden, J. C. Contribution assessment of multiparameter optimization descriptors in CNS penetration. *SAR and QSAR in Environmental Research* **29**, 785–800, <https://doi.org/10.1080/1062936x.2018.1514652> (2018).
35. Yuan, Y., Zheng, F. & Zhan, C.-G. J. T. A. J. Improved Prediction of Blood–Brain Barrier Permeability Through Machine Learning with Combined Use of Molecular Property-Based Descriptors and Fingerprints. **20**, 54, <https://doi.org/10.1208/s12248-018-0215-8> (2018).
36. Shityakov, S. & Förster, C. Y. J. H. *Biology, C. Computational simulation and modeling of the blood–brain barrier pathology*. **149**, 451–459, <https://doi.org/10.1007/s00418-018-1665-x> (2018).
37. Winkler, D. A. & Burden, F. R. Modelling blood–brain barrier partitioning using Bayesian neural nets. *Journal of Molecular Graphics and Modelling* **22**, 499–505 (2004).
38. Dimitrov, D. S. & Marks, J. D. In *Therapeutic Antibodies: Methods and Protocols* (ed Antony S. Dimitrov) 1–27 (Humana Press, 2009).
39. Sato, K. Consideration for future *in vitro* BBB models - technical development to investigate the drug delivery to the CNS. *Nihon Yakurigaku Zasshi* **152**, 287–294, <https://doi.org/10.1254/fpj.152.287> (2018).
40. Sharma, B., Luhach, K. & Kulkarni, G. T. In *Brain Targeted Drug Delivery System* (eds Huile Gao & Xiaoling Gao) 53–101 (Academic Press, 2019).
41. Kuhn, M., Letunic, I., Jensen, L. J. & Bork, P. The SIDER database of drugs and side effects. *Nucleic Acids Research* **44**, D1075–D1079 (2016).
42. Subramanian, G. & Kitchen, D. B. J. o. C.-A. M. D. Computational models to predict blood–brain barrier permeation and CNS activity. *Journal of Computer-Aided Molecular Design* **17**, 643–664 (2003).
43. Li, H. *et al.* Effect of Selection of Molecular Descriptors on the Prediction of Blood–Brain Barrier Penetrating and Nonpenetrating Agents by Statistical Learning Methods. *Journal of Chemical Information and Modeling* **45**, 1376–1384 (2005).
44. Abraham, M. H., Ibrahim, A., Zhao, Y. & Acree, W. E. Jr. A data base for partition of volatile organic compounds and drugs from blood/plasma/serum to brain, and an LFER analysis of the data. *Journal of Pharmaceutical Sciences* **95**, 2091–2100 (2006).
45. Law, V. *et al.* DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Research* **42**, D1091–D1097 (2014).
46. Wang, W., Kim, M. T., Sedykh, A. & Zhu, H. J. P. R. Developing Enhanced Blood–Brain Barrier Permeability Models: Integrating External Bio-Assay Data in QSAR Modeling. *Pharmaceutical Research* **32**, 3055–3065 (2015).
47. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
48. Scholkopf, B. & Smola, A. J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. (MIT Press, 2001).
49. Suenderhauf, C., Hamann, F. & Huwyler, J. Computational Prediction of Blood–Brain Barrier Permeability Using Decision Tree Induction. *Molecules* **17**, 10429–10445 (2012).

## Acknowledgements

This work is supported by the Macau Science and Technology Development Funds Grand No. 003/2016/AFJ and No. 0055/2018/A2 from the Macau Special Administrative Region of the People's Republic of China.

## Author Contributions

L.Y.X. and R.M. conceived the conception. R.M. also designed and developed the method, acquired and analyzed the data and result. H.H.C. and H.H.H. wrote, reviewed and revised the manuscript. Y.L. is the correspondence author. All authors have read and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-44773-4>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019