Neurocomputing 434 (2021) 80–89

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Prediction of drug-target interactions based on multi-layer network representation learning

Yifan Shang [a], Lin Gao [a], Quan Zou [b], Liang Yu [a,*]

[a] School of Computer Science and Technology, Xidian University, Xi'an 710071, Shaanxi, China
[b] Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China

ARTICLE INFO

Article history:
Received 6 September 2020
Revised 29 October 2020
Accepted 12 December 2020
Available online 28 December 2020
Communicated by Steven Hoi

Keywords:
Drug-target interaction
Similarity network
Multi-layer network representation learning
Neural network
Deep learning

ABSTRACT

The prediction of drug-target interactions aims to identify potential targets for the treatment of new and rare diseases. The large number of unknown combinations between drugs and targets makes them difficult to verify with experimental methods. There are computational methods that predict drug-target interactions; however, these methods are insufficient in integrating multiple types of data and managing network noise, which affects the accuracy of the prediction. We report a multilayer network representation learning method for drug-target interaction prediction that can integrate useful information from different networks, reduce noise in the multilayer network, and learn the feature vectors of drugs and targets. The feature vectors of the drug and the target are put into the drug-target space to predict the potential drug-target interactions. This work improves the method of multilayer network representation learning and prediction accuracy by increasing the parameter regularization constraints.

1. Introduction

With the advancement of medical technology, drug-target interactions (DTIs) have been discovered. Although the number of known drug-target interactions have increased, the number of proteins approved as drug targets is less than 10% of all human proteins [1]. According to statistics [2], most drugs have only a few targets, with an average of two to three targets for each drug. Therefore, for all drug-target pairs, the known drug-target interactions are very limited [3], and new interactions need to be predicted in order to discover more potential effects of available drugs. DTIs prediction is an important part of drug reposition that can reduce the cost and cycle of drug development. The cost and time required to determine all drug-target interactions using experimental methods cannot be estimated; therefore, computational methods are needed to predict drug-target interactions to narrow the scope, reduce costs, and save experimental time. The traditional calculation method mainly relies on two strategies based on molecular docking [4,5] and ligand [6,7]. The molecular docking method requires three-dimensional structural data of the predicted target protein, while the ligand method requires a large number of binding ligands; therefore these two strategies are not suitable for managing large-scale data. Current computational methods are based on the assumption that similar drugs may have the same target and vice versa. These methods are mainly divided into two categories according to the data used: single-type data and multi-type data integration.

Among the methods based on single-type data, there was a recent study [8] that used chemical structure data of drugs in term frequency and term frequency-inverse document frequency weighting methods. These methods proposed a Simplified Molecular Input Line Entry Specification (SMILES) kernel based on cosine similarity to predict drug-target interactions in multiple data sets by comparing multiple SMILES-based similarity methods, which were shown to be superior to other methods [8]. The similarity of phenotypic side effects is also used to infer whether drugs have common targets [9]. Zhu, el al. [10] conducted a groundbreaking study in which they developed a probabilistic model, called the Mixture Aspect Model (MAM), based on text mining of drug and target co-occurrence, and found hidden drug-target interactions. Recently, some articles continue to study on the basis of probabilistic models [11,12]. These analyses were based on one type of data, using different methods to predict drug-target interactions.

With the development of biotechnology, data types are becoming more and more abundant. Compared with a single type of data, there is information supplemented between multiple types of data, and the integration of different types of data can be used to mine for additional hidden information.

* Corresponding author.
E-mail address: lyu@xidian.edu.cn (L. Yu).

https://doi.org/10.1016/j.neucom.2020.12.068
0925-2312/

There are studies [13–15] that used the similarity network fusion (SNF) method [16] to nonlinearly fuse multiple similarity networks and provide information supplementation for different networks [17,18]. There are also two-way random walks in heterogeneous bipartite networks to infer potential drug-target interactions [19,20]. A study projected multiple types of data into a common feature space, integrated multiple networks into a single network, and then used collaborative matrix decomposition to make predictions [21]. There is also a method based on Laplacian regularized sparse subspace learning (LRSSL), which integrates a variety of drug characteristics such as drug chemical structure information, drug target information, and target labeling features projected into the common subspace, and assigns Laplace regularization terms to satisfy the smoothness of the subspace [22]. There are also methods such as a support vector machine (Support Vector Machine, SVM) and feature selection (Feature Selection, FS) that establish a predictive drug target interaction model [23].

However these methods have deficiencies in the integration of data. First, through network fusion, which directly uses the diffusion state as a feature or prediction score, the noise in different networks can affect results. Second, integrating multiple networks into one network or projecting into a common subspace may result in the loss of specific information from the different networks because the information from multiple data sources is mixed and cannot be distinguished from one another. These factors affect the accuracy of predicting drug-target interactions.

The Multi-layer network representation learning to predict drug target interactions (MEDTI) model is based on deep neural networks [24–27] and integrates multiple layers of similarity networks constructed using multiple types of drugs and targets, learns the drug and target compact feature vectors, respectively, and then puts the drug and target feature vectors into the drug target space. The known drug-target interactions are used as supervising information and the drug and target vectors proximity in the drug-target space is used to predict the new drug target interaction. The MEDTI not only captures the common topology of multilayer drug or target similarity networks and the unique network topology features of each layer of the network, but also uses the characteristics of deep neural networks to capture deep nonlinear feature transformations. The MEDTI's innovation is the application of deep learning ideas to drug-target interaction prediction and improving the original multilayer network representation learning method. By adding regularization constraints, the accuracy of prediction is improved. Experiments on the same data set show that the performance of MEDTI is better than other prediction methods. In addition, the integration of multiple types of data can significantly improve the prediction accuracy. We also verified the top eight drug-target pairs in the prediction results and demonstrated the reliability of the prediction results from different methods. These results show that MEDTI is a useful tool for integrating multiple types of data to predict unknown DTIs that may provide new insight into drug reposition and understanding drug action mechanisms.

2. Materials and methods

2.1. Datasets

The data sets used in this article are all from public databases. Drug chemical structure information and drug interaction data were collected from the DrugBank database [28]. Drug-disease relationship data were extracted from the Comparative Toxicogenomics Database (CTD) [29], and the side-effect data of the drugs were collected in the SIDER database [30]. Extract target sequence information and protein interaction data are from the Human Protein Reference Database (HPRD) [31], and protein-disease relationship data from the CTD database. The known drug-target interaction data were extracted from the DrugBank database. Tables 1 and 2 present the data statistics.

2.2. Construction of the similarity network

By collecting data from multiple public databases, multiple networks have been constructed to predict unknown DTIs. Based on the four types of data of drug chemical structure, drug-disease association, drug-drug interaction, and drug side effects, four drug similarity networks are constructed. Three types of protein similarity networks are constructed based on three types of data: protein sequence, protein-disease association, and protein–protein interaction. Enhance the heterogeneity of the similarity network through multiple types of data, which can provide multi-angle information to predict DTIs. See the supplementary materials for the specific steps of similarity network construction.

2.3. Problem description

We used $D = \{d_i, i = 1, \ldots, n\}$ for the drugs node set, $T = \{t_j, j = 1, \ldots, m\}$ for the proteins node set, $A_D \in \Re^{n \times n}$ for the drug similarity matrix, $A_T \in \Re^{m \times m}$ for the protein similarity matrix, n for the number of drugs, and m for the number of proteins. The element values in different similarity matrices represent the similarity between drugs or target proteins based on different measures, and the values of all elements in each similarity matrix are in the range of $[0, 1]$. The interaction between the drugs D and the proteins T is represented by the matrix $P \in \Re^{n \times m}$. If there is an interaction between the drug d_i and the protein t_j, then $P_{ij} = 1$; otherwise $P_{ij} = 0$. The drug and protein similarity matrices are A_D and A_T, and the drug target interaction matrix is P. Our goal was to predict the new (i.e. unknown) drug-target interactions in P.

2.4. The framework of the MEDTI method

The MEDTI method uses a deep neural network to integrate a multilayer similarity network and denoise each layer of the network. Through multilayer network representation learning, a feature vector representation of the drug and target was obtained and then the known drug target interaction was used as supervision information to predict unknown drug-target interactions. In addition, this method improved the learning method of multilayer network representation. Fig. 1 illustrates the MEDTI process overview.

2.5. Learning compact feature vectors

In order to maintain the common topology of the multilayer similarity network and the unique characteristics of each layer of the network, we used the Multimodal Deep Autoencoder (MDA) [32] method and improved upon it to learn the compact feature vectors of drugs and targets. Fig. 2 is a flowchart of learning feature vectors.

First, the restarted random walk algorithm (RWR) [33] was used on the drug and target similarity network (A_D, A_T) to capture the topology information of the similarity network and obtain the feature vectors of the drug and target. Compared with the vector representation of network nodes obtained by DeepWalk [34], node2vec [35] and other methods, the RWR method required fewer hyperparameters and had lower computational complexity. The RWR was calculated as follows:

Table 1
The number of nodes in heterogeneous networks.

Types of node	Drug	Target	Disease	Drug side effect
Number of nodes	882	1449	6902	5439

Table 2
The number of edges in heterogeneous networks.

Type of edge	Number of edges
Drug-Target	3185
Drug-Drug	173,585
Drug-Disease	369,072
Drug-Side Effect	119,382
Protein-Protein	3193
Protein-Disease	2,173,297

$$p_i^{(t)} = \alpha p_i^{(t-1)} A + (1-\alpha) p_i^{(0)} \tag{1}$$

where $p_i^{(t)}$ is the row vector of the network node i after t steps, $p_i^{(0)}$ is the initial one-hot encoding vector, the i-th position of the vector is 1, the rest is 0, A is the transition probability matrix after the rows of the similar networks matrix are normalized, and α is the relative probability of whether the restarting walk is biased towards the local topology information or global topology information of the network.

In order to obtain the higher-order structure information of the network, we use the method of previously described [36] to add $p_i^{(t)}$ in t steps in RWR:

$$r_i = \sum_{t=1}^{T} p_i^{(t)} \tag{2}$$

where T represents the total number of steps in the RWR algorithm. Each node in a similar network can obtain a vector r_i that contains high-order structure information; therefore the r_i of all nodes form a co-occurrence probability matrix R.

After obtaining the co-occurrence probability matrix R, the Positive Pointwise Mutual Information (PPMI) [37] calculation is used to obtain a high-quality vector representation of network nodes. The PPMI is used to calculate the association probability between nodes and contains rich network context information. The PPMI is calculated as:

$$x_{lm} = \max\left\{0, \frac{R_{lm} \sum l \sum m R_{lm}}{\sum l R_{lm} \sum m R_{lm}}\right\} \tag{3}$$

where X_{lm} represents the probability co-present node values l and m in a similar network.

Each layer of the similarity network between the drug and the target can obtain the feature vectors of the drug and the target using the above calculation, but the feature vectors obtained based on RWR will have a significant impact because of the lack of edges in the network and false similarities [38]. Therefore, when integrating multilayer networks, high-dimensional noise needs to be further processed to reduce the impact of noise. In this paper, the multilayer network representation learning method based on deep neural network proposed by MDA [32] was used; therefore, the characteristics of noise influence can be removed by the automatic encoder system [39] and the vector representation of drugs and targets is learned in an unsupervised manner. We also improved the MDA method to make it more suitable for predicting drug-target interactions that integrated multiple types of data. When integrating multilayer networks, use homogeneous networks for integration. The drug similarity networks are integrated to obtain the drug feature vector, and then the target feature vector is obtained in the same way. The integration steps were as follows:

(i) Encoding

Using the PPMI matrix $X^{(j)} \in \Re^{k \times k}, j \in \{1, \ldots, N\}$ of all similar networks calculated in the previous step, we calculated the nonlinear representation and performed denoising:

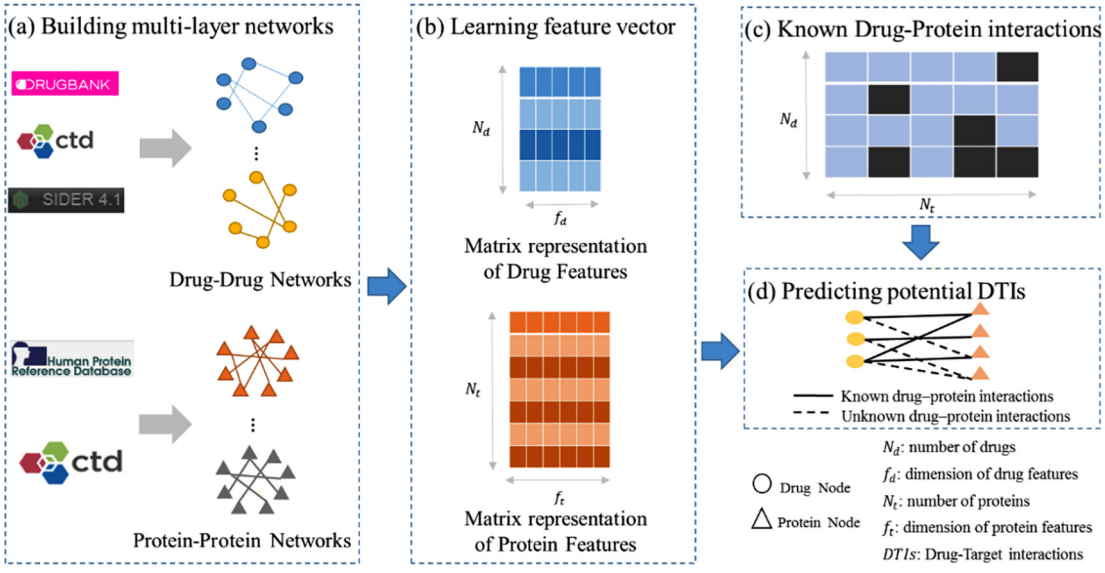


Fig. 1. A flowchart of the MEDTI method. (a) The multilayer drug and target similarity network is constructed from the data of the drug and the target, respectively. (b) According to the improved multilayer network representation method, the feature vectors of drugs and targets are learned separately. (c) Known drug-target interactions as supervisory information. (d) Map feature vectors of drugs and targets in the drug-target space and how supervisory information is used to predict unknown DTIs.

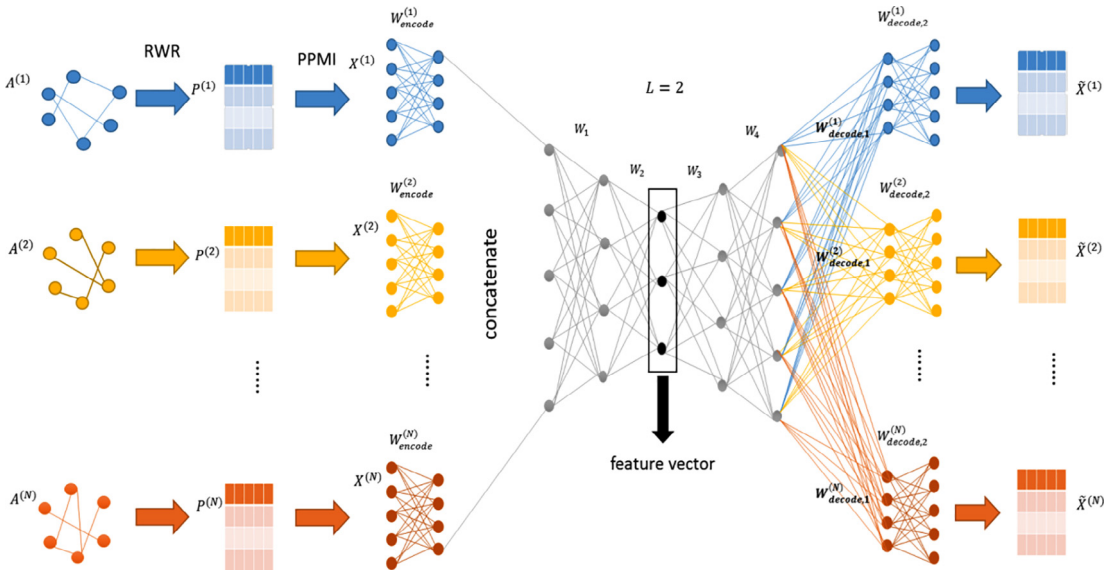


Fig. 2. Multilayer network representation learning method overview. First, multilayer networks were converted into vectors according to the restarted random walk algorithm (RWR) method. Then calculate the PPMI matrix of each layer network vector to get the high-quality vector. Finally, the multilayer network is combined to learn the feature vectors of the network nodes using MDA, which is used for the subsequent prediction work.

$$H_{encode}^{(j)} = \sigma\left(W_{encode}^{(j)} X^{(j)} + B_{encode}^{(j)}\right) \tag{4}$$

where $W_{encode}^{(j)} \in \Re^{d_j \times k}$ is the weight matrix, $B_{encode}^{(j)} \in \Re^{d_j \times k}$ is the bias matrix, and $\sigma(x) = \frac{1}{1+e^{-x}}$ is the sigmoid activation function.

Then, the characteristics of each layer of the network were concatenated, and the multilayer nonlinear activation function was used to calculate the common feature representation of the integrated network. The first layer was converted to:

$$H_{c,1} = \sigma\left(W_1\left[H^{(1)}, \ldots, H^{(N)}\right] + B_1\right) \tag{5}$$

where $\left[H^{(1)}, \ldots, H^{(N)}\right]$ was a concatenated activation function and the L layer was expressed as:

$$H_{c,l+1} = \sigma\left(W_l H_{c,l} + B_l\right) \tag{6}$$

where $l \in \{1, \ldots, L\}$ is the number of consecutive conversion layers.

(ii) Decoding

After obtaining the coding layer features $H_{c,l+1}$, the decoding operation was performed with the same number of layers as the coding layer:

$$H_{c,l+1} = \sigma\left(W_l H_{c,l} + B_l\right) \tag{7}$$

where $l \in \{L+1, \ldots, 2L\}$ is the number of decoding and conversion layers. Then the decoded common feature $H_{c,2L}$ was used to calculate the decoded representation $H_{decode}^{(j)}$ of each layer of the network:

$$H_{decode}^{(j)} = \sigma\left(W_{decode,1}^{(j)} H_{c,2L} + B_{decode,1}^{(j)}\right) \tag{8}$$

The decoding feature $H_{decode}^{(j)}$ of each layer network was used to restore the input PPMI matrix output $\hat{X}^{(j)}, j \in \{1, \ldots, N\}$:

$$\hat{X}^{(j)} = \sigma\left(W_{decode,2}^{(j)} H_{decode}^{(j)} + B_{decode,2}^{(j)}\right) \tag{9}$$

In order to minimize the gap between the original PPMI matrix $X^{(j)}$ and the reduction matrix $\hat{X}^{(j)}$, the objective function used was:

$$\theta = \arg\min_{\theta} L(\theta) = \arg\min_{\theta} \sum_{j=1}^{N} l\left(X^{(j)}, \hat{X}^{(j)}\right) + Z(\theta) \tag{10}$$

where $l(*)$ is the sample-wise binary cross-entropy function, and $Z(\theta) = Z\left\{W_{encode}^{(j)}, B_{encode}^{(j)}, W_{decode}^{(j)}, B_{decode}^{(j)}, W_l, B_l\right\}$ is the regularization constraint of all parameters when encoding and decoding the model.

In the model, the reverse standard propagation algorithm was used to optimize the loss function. The model was trained using a small batch stochastic gradient descent. After optimizing the objective function, the final intermediate layer $H_{c,L+1} \in \Re^{d \times k}$ was the required compact feature vector, where d is the set vector dimension, and k represents the number of drug or target nodes, the feature vectors of the drug and the target were obtained respectively.

2.6. Predicting drug-target interactions

The feature vectors of the drug and target that were learned in the previous step were then placed into the drug target space. If the feature vectors of the two were geometrically close in the drug target space, there was a potential for a drug-target interaction. The known drug-target interactions are then used as supervisory information to predict unknown interactions. We used the method described by Luo et al. [40] to learn the projection matrix of drugs and proteins, and then multiplied the projected drug and target feature vectors to obtain the drug-target interaction prediction matrix P (See supplementary materials for projection matrix solution).

3. Results

This study analyzed the performance of an MEDTI algorithm designed for the efficient prediction drug-target interactions. These experiments included two parts: (i) experiments on the data set collected by the MEDTI algorithm using different algorithms that included unimproved multilayer network representation learning algorithms to compare the accuracy of the prediction results with the advantages of the MEDTI algorithm. (ii) By comparing the integration of different networks, the impact on the prediction results changes, indicating that integration of multilayer networks

can help to improve the accuracy of drug-target interactions prediction models.

3.1. Performance comparison of different methods

The first task was to compare the available interaction prediction algorithms. We used the DDR [14] and LRSSL [22] algorithms for the comparison of multilayer network representation learning methods. The MNE [41] and original MDA [32] methods were used to obtain the feature vectors of drugs and targets, and then based on these feature vectors, the fourth step of the MEDTI algorithm was used to obtain the prediction results. In calculating the prediction results, a 10-fold cross-validation strategy was adopted. The area under the AUROC curve and the area under the accurate recall curve AUPR were used to evaluate the performance of each method [42,43]. We then verified the prediction results for the novel drug-target interactions.

First, we compared the MEDTI algorithm with the interaction prediction algorithm. The results in Fig. 3 showed that compared with LRSSL, the AUROC value was 1.93% higher and the AUPR value was 78.56% higher with MEDTI, while compared with DDR, the AUROC value was 3.14% higher, and the AUPR value was 21.06% higher using the MEDTI method. Because the method of network representation learning to capture more topological information in heterogeneous networks, the results are better than LRSSL and DDR. Second, compared with the multilayer network embedding learning method MNE algorithm, both the AUROC (8.74%) AUPR (6.01%) values were higher with MEDTI, which may have been because the MNE adopts the random walk method to embed the multilayer network which due to the missing and false information in each layer of the network, the random walk can lead to increased noise when integrating a multilayer network. Thus the quality of the feature vector is low, which affects the accuracy of the final prediction result. Finally, due to our improvement of the MDA method, we were also able to compare the prediction results of the MDA algorithm with those of the MEDTI algorithm. The AUROC value was 9.38% higher and the AUPR value was 6.54% higher when we used the MEDTI compared with the MDA model because the MDA algorithm uses a deep automatic encoder based on a multilayer neural network. In the representation of multilayer networks, as the number of layers of the neural network deepens, the parameters that need to be trained increase exponentially. In the case, the MEDTI algorithm is based on deep neural networks, there were too many parameters in the model training. When optimizing the objective function, it was easy to cause over-fitting of the parameters, resulting in overt accuracy in the results of the training data; however, the accuracy of the test data, cannot be applied to actual data. Therefore, this paper improved on the original MDA model to add the regularization constraint term $Z(\theta)$ of parameters in the objective function (10) to prevent the training model from overfitting. The experiment showed that after adding the regularization constraints of the parameters, the prediction accuracy was better than the results of the original MDA model.

3.2. Performance comparison between single and differentially integrated networks

Integrating multiple types of data helps to improve the accuracy of predicting drug-target interactions. Therefore, we conducted experiments to test whether the integrated network has an effect on the accuracy of drug target interaction prediction results.

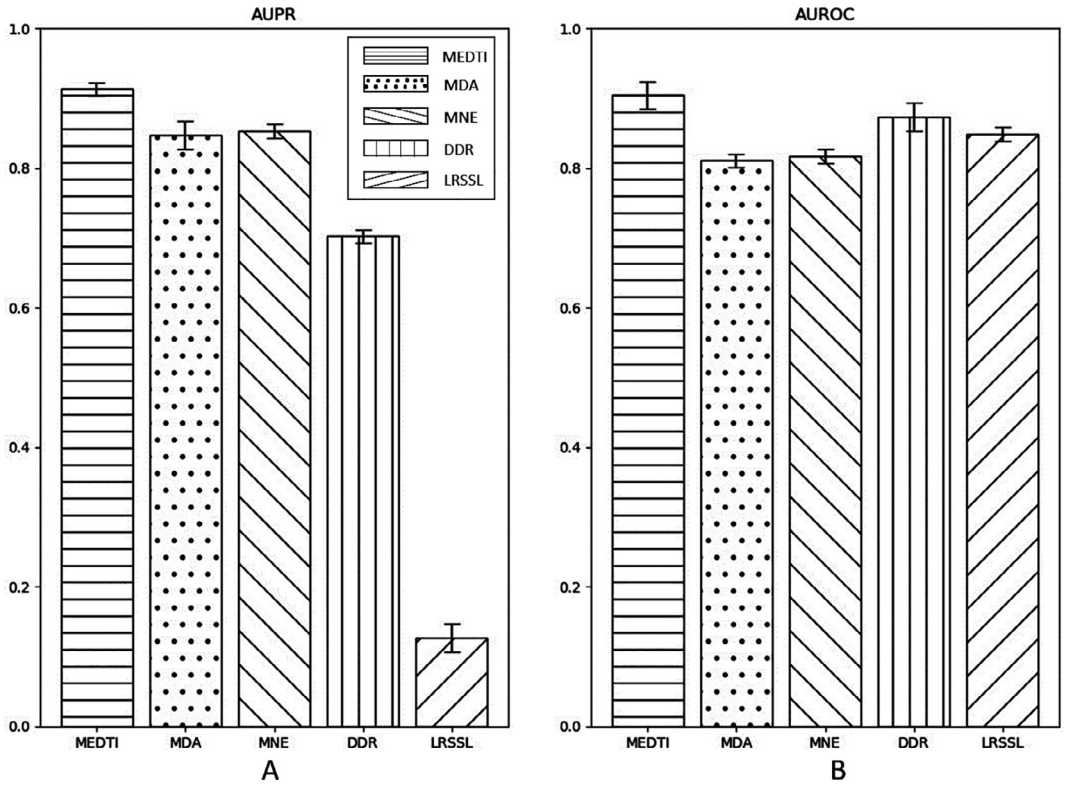


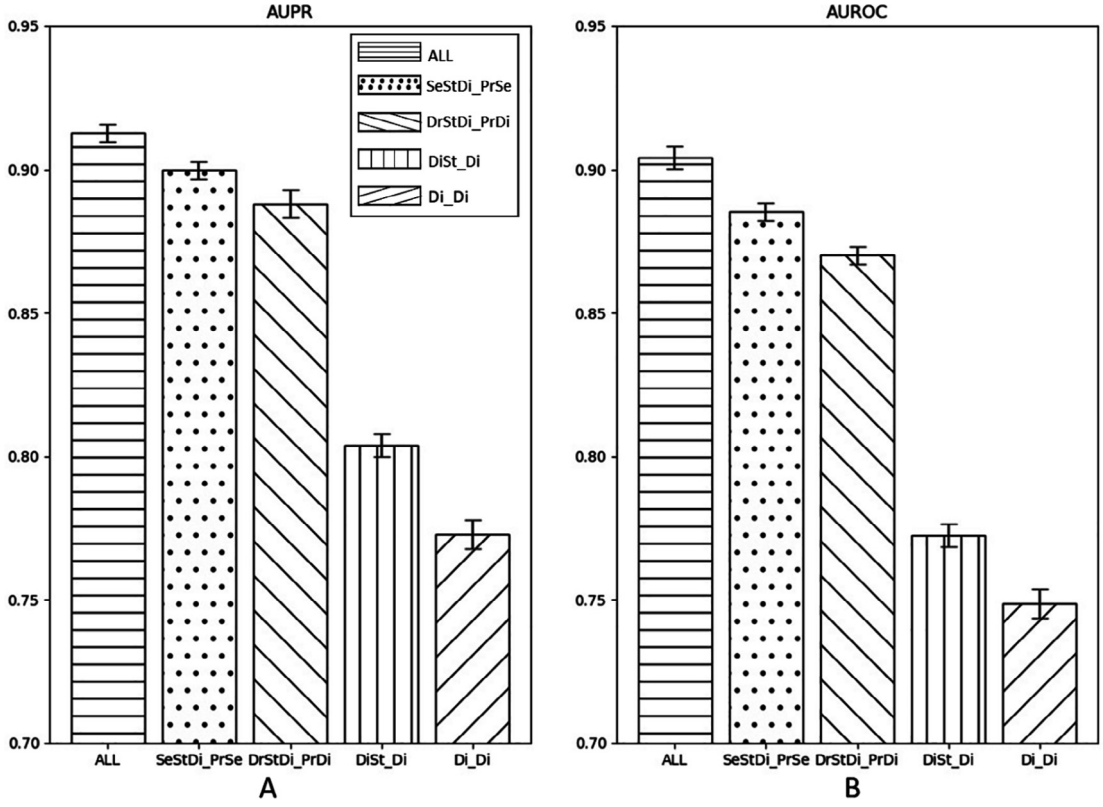
Fig. 3. Comparison of AUROC and AUPR values of different prediction algorithms for drug-target interactions. The A and B graphs compare the AUPR and AUROC values, respectively, of the different algorithms tested. The MEDTI was the algorithm proposed in this paper while MDA was the original multilayer network identification learning algorithm used to obtain the feature vectors of the drugs and targets. MNE is a multilayer network embedding method, and DDR and LRSSL are interaction prediction algorithms. All results were summarized over 10 trials and expressed as mean ± SD.

The MEDTI algorithm presented in this paper predicts drug-target interactions by integrating multiple types of data. As the Fig. 4 showed, in the result of ALL, the AUROC values were found to be 1.88% and 3.4% higher, and the AUPR values were 1.3% and 2.46% higher, than SeStDi_PrSe and DrStDi_PrDi, respectively; the AUROC value was 13.17% higher and the AUPR value is 10.89% higher than DiSt_Di; the AUROC value is 15.6% higher and the AUPR value is 13.99% higher than Di_Di. Because different types of data describe drugs or proteins from different perspectives, after building a multilayer network, the topology of the network has its own characteristics. These different networks have complementary information between each other, so the more data types integrated, the more accurate the prediction. In addition, when the network has the same number of integration layers but different types, the accuracy is also different between the results of the analysis, which is because different types of data can contribute different information to the networks.

3.3. Verification of prediction results

Table 3 lists 8 pairs of predicted drug target interactions. Excluding the known drug-target interaction scores, these 8 pairs of drug-target interaction prediction scores are the highest. Because the MEDTI prediction model uses known drug-target interactions as supervision information, and among the known interactions, the number of drugs acting on protein P35348 was the largest, there were several records of target P35348 in the prediction results. Therefore, we wanted to verify the prediction results in different ways, including literature verification, disease association verification and enrichment analysis verification.

The drugs and targets that corresponded to DB00321 and P14416 were Amitriptyline and the D(2) dopamine receptor (DRD2), respectively. A study investigated the effect of psychotropic drugs on the transcription of genes related to the risk



	ALL	SeStDi_PrSe	DrStDi_PrDi	DiSt_Di	Di_Di
Drug Side effects	✓	✓			
Drug Chemical structure	✓	✓	✓	✓	
Drug-Disease	✓	✓	✓		✓
Drug interactions	✓		✓	✓	
Protein intertactions	✓	✓	✓		
Protein Sequence	✓	✓			
Protein-Disease	✓		✓	✓	✓

C

Fig. 4. The effect of integrating different networks for drug-target interaction prediction. The A and B graphs illustrate the AUPR and AUROC values comparison between algorithms, respectively. The C is the data composition of the integrated network in the experiment. For example, ALL represents all similarity networks for integrating drugs and targets. All results were summarized over 10 trials and expressed as mean ± SD.

Table 3
The 8 predicted drug-target interactions in the top-50 list.

Rank	DrugBank ID	Drug Name	UniProt ID	Target Name
31	DB00321	Amitriptyline	P14416	D(2) dopamine receptor
32	DB01618	Molindone	P35348	Alpha-1A adrenergic receptor
38	DB00434	Cyproheptadine	P35348	Alpha-1A adrenergic receptor
41	DB00589	Lisuride	P35348	Alpha-1A adrenergic receptor
42	DB06216	Asenapine	P35368	Alpha-1B adrenergic receptor
43	DB00186	Lorazepam	P35348	Alpha-1A adrenergic receptor
46	DB05271	Rotigotine	P35348	Alpha-1A adrenergic receptor
48	DB08815	Lurasidone	P35348	Alpha-1A adrenergic receptor

of Parkinson 's disease development [44]. The effect of psychotropic drugs on the mRNA expression of 70 genes related to Parkinson's disease was found to be due to the drug Amitriptyline, which can affect and up-regulate the expression of *DRD2*. The corresponding drugs and targets of DB00589 and P35348 are Lisuride and Alpha-1A adrenergic receptor (ADRA1A), respectively. Newman Tancredi et al. [45] analyzed the binding patterns of anti-Parkinson agents and revealed a comparative pattern of affinity for different classes of monoaminergic receptors. They identified the effects of some human receptors, including ADRA1A, and found that the drug Lisuride has moderate effects on these receptors. Another study [46] reported that Lisuride had both agonist and antagonist effects on ADRA1A and suggested that its nano-level affinity for ADRA1A may alter its efficacy as a Parkinson's disease drug. Williams et al. [47] proposed their report that ADRA1A is a G protein-coupled receptor (GPCR) that can regulate the contraction of peripheral smooth muscle and the neuronal output of the central nervous system (CNS), making it an emerging treatment of neurodegenerative diseases. The central nervous system targeting drug, Lorazepam, is considered to be a positive allosteric modulator (PAMs) of ADRA1A [47].

In the drug-target interaction networks, a disease association helps to support the hypothesis that a drug may potentially affect the predicted target. For example, Lisuride is used to treat a variety of diseases, including dyskinesias [48], and seizures [49]. ADRA1A is also associated with diseases, such as Liver Cirrhosis [50], and Seizures [51]. If the diseases associated with Lisuride and ADRA1A intersect, the two diseases are related and Lisuride may have a potential effect on ADRA1A (Fig. 5).

Enrichment analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) [52] pathways can also be used to illustrate the interaction between drugs and targets. For example, pathway enrichment analysis of the drug Amitriptyline had a P-value < 0.01 and indicated pathway KEGG: hsa04080, which indicates a neuroactive ligand-receptor interaction and suggests, this pathway is involved in stimulating nerve tissue (Fig. 6). By querying the gene collection in this pathway (which includes DRD2), further suggesting that DRD2 participates in the biological process of Amitriptyline through their interaction.

Finally, all the 8 drug-target interactions were verified from three aspects (literature verification, disease association, enrichment analysis) and the results are shown in Table 4. We found that, the drug-target interactions ranked 31,41,43 were the most reliable of the predicted results.

4. Discussion

This paper collected current data about drugs and their targets from the public databases DrugBank, CTD, HPRD, and SIDER to develop multilayer network representation learning and deep learning for drug-target interaction prediction for improved prediction accuracy. We used the improved MDA algorithm by adding regularization constraint parameters to prevent the model from overfitting and affecting the accuracy of the final prediction results. In this paper, a ten-fold cross-validation was used to calculate the final prediction accuracy and prediction results from other methods were compared with the same data set. The comparison of AUROC and AUPR values showed that the accuracy of the MEDTI framework proposed in this paper were higher than those from other methods, which indicated that the multilayer network representation learning method was more effective and ultimately improved prediction accuracy. In addition, the results show that the prediction accuracy rate of integrating multi-type data was higher than that of single-type data; thus, it is feasible to improve model prediction accuracy rates through the integration of multi-type data. In the validation of the MEDTI model, three different verification methods were used, including literature verification, disease association verification, and the KEGG pathway enrichment analysis. These results showed that the reliability of the MEDTI predicted drug-target interactions could verified from different analytical processes.

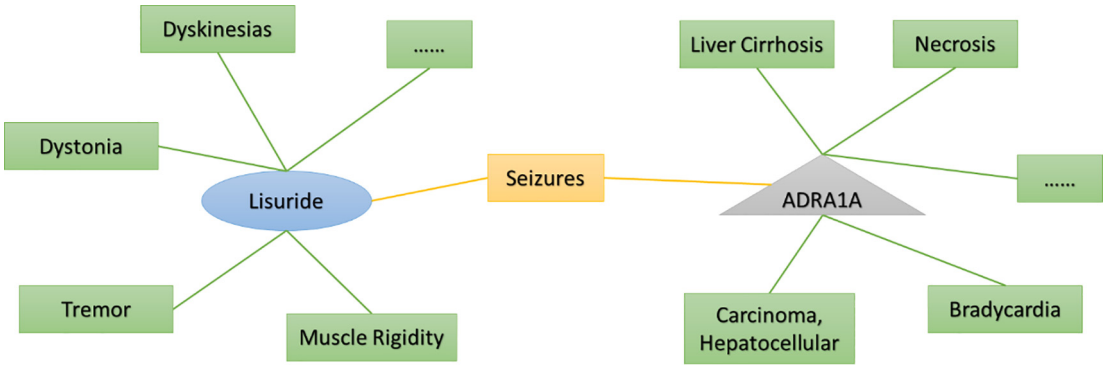


Fig. 5. The disease correlation between the drug Lisuride and the ADRA1A protein. The blue circle represents Lisuride, the gray triangle represents ADRA1A, and the green squares represent the diseases associated with the drug and protein independently. The yellow box indicates the disease that is associated with both Lisuride and ADRA1A.

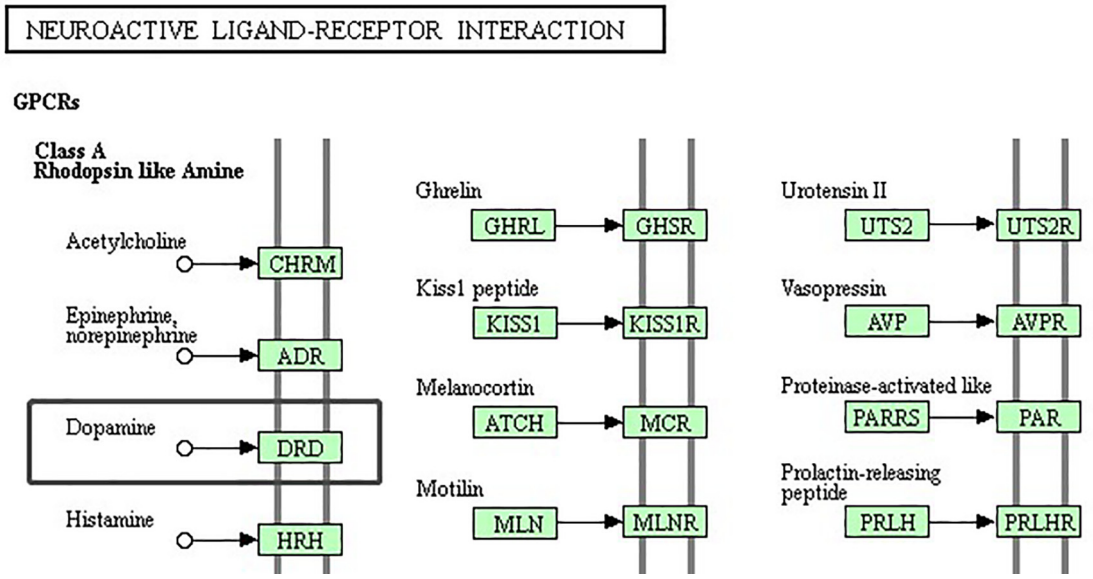


Fig. 6. KEGG hsa04080 partial enrichment pathway diagram. The figure shows part of the hsa04080 pathway in KEGG. The green is the gene on the pathway, and the black box is the location of the gene DRD2.

Table 4
The verification of eight predicted drug-target pairs from three aspects (literature verification, disease association, enrichment analysis).

Rank	Drug Name	Target Name	Literature verification	Disease Relationship	Pathway Enrichment Analysis
31	Amitriptyline	D(2) dopamine receptor	√	√	√
32	Molindone	Alpha-1A adrenergic receptor		√	√
38	Cyproheptadine	Alpha-1A adrenergic receptor		√	√
41	Lisuride	Alpha-1A adrenergic receptor	√	√	√
42	Asenapine	Alpha-1B adrenergic receptor		√	
43	Lorazepam	Alpha-1A adrenergic receptor	√	√	√
46	Rotigotine	Alpha-1A adrenergic receptor		√	√
48	Lurasidone	Alpha-1A adrenergic receptor		√	√

The MEDTI framework proposed here provides new insight into the integration of multiple types of data to predict drug-target interactions and can be used in other link prediction tasks with good scalability. Multilayer network representation learning integrates different types of data, reduces the impact of noise in the multilayer network, and captures the deep topology of the network. These can play a role in promoting link prediction tasks in other fields.

The MEDTI model has some limitations. When integrating multiple types of data, if there is a heterogeneous network with more than two types of nodes, the vector representation of the network features cannot be extracted. At present, the multilayer similarity network integrated by MEDTI is still a multilayer network with the same node type. However, heterogeneous networks contain richer network information and how to learn to extract feature vectors through heterogeneous networks to fully capture more complex network topology is still a challenge. At the same time, the feature vectors of drugs and targets have low interpretability, including the meaning represented by each dimension in the vector. This is a challenge that exists in current network representation learning methods and although the accuracy of the prediction results is high, how to interpret the feature vectors requires further study.

Funding

This work has been supported by the National Key Research and Development Program of China (grant no. 2018YFC0910403), and National Natural Science Foundation of China (grant nos. 62072353 and 61672406 and 61532014).

CRediT authorship contribution statement

Yifan Shang: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Lin Gao:** Resources, Writing - review & editing, Funding acquisition. **Quan Zou:** Writing - review & editing. **Liang Yu:** Conceptualization, Methodology, Formal analysis, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Thanks to all those who maintain excellent databases and to all experimentalists who enabled this work by making their data publicly available.

References

[1] P. Csermely, T. Korcsamaros, H.J.M. Kiss, G. London, R. Nussinov, Structure and dynamics of molecular networks: a novel paradigm of drug discovery. A comprehensive review, Pharmacol. Ther. 138 (2013) 333–408.

[2] T. Cheng, M. Hao, T. Takeda, S.H. Bryant, Y. Wang, Large-scale prediction of drug-target interaction: a data-centric review, AAPS J. 19 (2017) 1264–1275.

[3] B.R. Stockwell, Chemical genetics: ligand-based discovery of gene function, Nat. Rev. Genet. 1 (2000) 116–125.

[4] B.R. Donald, Algorithms in Structural Molecular Biology, MIT Press, Cambridge, MA, 2011.

[5] G.M. Morris, R. Huey, W. Lindtrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 30 (2009) 2785–2791.

[6] A.J. Fathima, G. Murugaboopathi, P. Selvam, pharmacophore mapping of ligand based virtual screening, molecular docking and molecular dynamic simulation studies for finding potent NS2B/NS3 protease inhibitors as potential anti-dengue drug compounds, Curr. Bioinform. 13 (2018) 606–616.

[7] M.J. Keiser, B.L. Roth, B.N. Armbruster, P. Ernsberger, J.J. Irwin, B.K. Schoichet, Relating protein pharmacology by ligand chemistry, Nat. Biotechnol. 25 (2007) 197–206.

[8] H. Ozturk, E. Ozkirimli, A. Ozgur, A comparative study of SMILES-based compound similarity functions for drug-target interaction prediction, BMC Bioinf. 17 (2016) 128.

[9] M. Campillos, M. Kuhn, A.C. Gavin, L.J. Jensen, P. Bork, Drug target identification using side-effect similarity, Science 321 (2008) 263–266.

[10] S. Zhu, Y. Okuno, G. Tsujimoto, H. Mamitsuka, A probabilistic model for mining implicit 'chemical compound–gene' relations from literature, Bioinformatics 21 (2005) 245–251.

[11] A. Munir, S.I. Malik, K.A. Malik, Proteome mining for the identification of putative drug targets for human pathogen *Clostridium tetani*, Curr. Bioinform. 14 (2019) 532–540.

[12] J. Wang, H. Wang, X. Wang, H. Chang, Predicting drug-target interactions via FM-DNN learning, Curr. Bioinform. 15 (2020) 68–76.

[13] W. Wang, S. Yang, X. Zhang, L. Jing, Drug repositioning by integrating target information through a heterogeneous network model, Bioinformatics 30 (2014) 2923–2930.

[14] R.S. Olayan, H. Ashoor, V.B. Bajic, DDR: efficient computational method to predict drug-target interactions using graph mining and machine learning approaches, Bioinformatics 34 (2018) 3779.

[15] N. Srivastava, B.N. Mishra, P. Srivastava, In-silico identification of drug lead molecule against pesticide exposed-neurodevelopmental disorders through network-based computational model approach, Curr. Bioinform. 14 (2019) 460–467.

[16] B.o. Wang, A.M. Mezlini, F. Demir, M. Fiume, Z. Tu, M. Brudno, B. Haibe-Kains, A. Goldenberg, Similarity network fusion for aggregating data types on a genomic scale, Nat. Methods 11 (3) (2014) 333–337, https://doi.org/10.1038/nmeth.2810.

[17] X. Liu, Z. Hong, J. Liu, Y. Lin, A. Rodriguez-Paton, Q. Zou, X. Zeng, Computational methods for identifying the critical nodes in biological networks, Brief Bioinform. 21 (2020) 486–497.

[18] X. Zhao, L. Chen, Z.-H. Guo, T. Liu, Predicting drug side effects with compact integration of heterogeneous networks, Curr. Bioinform. 14 (2019) 709–720.

[19] H. Luo, J. Wang, M. Li, J. Luo, X. Peng, F.X. Wu, Y. Pan, Drug repositioning based on comprehensive similarity measures and Bi-Random walk algorithm, Bioinformatics 17 (2016) 2664–2671.

[20] X.Y. Yan, S.W. Zhang, C.R. He, Prediction of drug-target interaction by integrating diverse heterogeneous information source with multiple kernel learning and clustering methods, Comput. Biol. Chem. 78 (2019) 460–467.

[21] I. S. Dhillon, Y. Koren, R. Ghani, T.E. Senator, R. Uthurusamy, Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013.

[22] X. Liang, P. Zhang, L. Yan, Y. Fu, F. Peng, L. Qu, M. Shao, Y. Chen, Z. Chen, LRSSL: predict and interpret drug–disease associations based on data integration using sparse subspace learning, Bioinformatics 33 (2017) 1187–1196.

[23] Y. Ding, J. Tang, F. Guo, Identification of drug-target interactions via multiple information integration, Inf. Sci. 418 (2017) 546–560.

[24] X. Zeng, Y. Zhong, W. Lin, Q. Zou, Predicting disease-associated circular rnas using deep forests combined with positive-unlabeled learning methods, Brief Bioinform. (2020), https://doi.org/10.1093/bib/bbz1080.

[25] Z. Lv, C. Ao, Q. Zou, Protein function prediction: from traditional classifier to deep learning, Proteomics 19 (2019) 1900119.

[26] B. Wu, H. Zhang, L. Lin, H. Wang, Y. Gao, L. Zhao, Y.P. Chen, R. Chen, L. Gu, A similarity searching system for biological phenotype images using deep convolutional encoder-decoder architecture, Curr. Bioinform. 14 (2019) 628–639.

[27] R. Su, X. Liu, L. Wei, Q. Zou, Deep-Resp-Forest: a deep forest model to predict anti-cancer drug response, Methods 166 (2019) 91–102.

[28] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, et al., DrugBank 5.0: a major update to the DrugBank database for 2018, Nucleic Acids Res. 46 (2018) D1074–1082.

[29] A.P. Davis, C.G. Murphy, R. Johnson, J.M. Lay, K. Lennon-Hopkins, C. Saraceni-Richards, D. Sciaky, B.L. King, M.C. Rosenstein, T.C. Wiegers, et al., The comparative toxicogenomics database: update 2013, Nucleic Acids Res. 41 (2013) D1104–D1114.

[30] M. Kuhn, I. Letunic, L.J. Jensen, P. Bork, The SIDER database of drugs and side effects, Nucleic Acids Res. 44 (D1) (2016) D1075–D1079, https://doi.org/10.1093/nar/gkv1075.

[31] T.S. Keshava Prasad, R. Goel, K. Kandasamy, S. Keerthikumar, S. Kumar, S. Mathivanan, D. Telikicherla, R. Raju, B. Shafreen, A. Venugopal, L. Balakrishnan, A. Marimuthu, S. Banerjee, D.S. Somanathan, A. Sebastian, S. Rani, S. Ray, C.J. Harrys Kishore, S. Kanth, M. Ahmed, M.K. Kashyap, R. Mohmood, Y.L. Ramachandra, V. Krishna, B.A. Rahiman, S. Mohan, P. Ranganathan, S. Ramabadran, R. Chaerkady, A. Pandey, Human Protein Reference Database–2009 update, Nucleic Acids Res. 37 (Database) (2009) D767–D772, https://doi.org/10.1093/nar/gkn892.

[32] V. Gilgorijevic, M. Barot, R. Bonneau, deepNF: deep network fusion for protein function prediction, Bioinformatics 43 (2018) 3873–3881.

[33] H. Tong, C. Faloutsos, J. Pan, Fast random walk with restart and its applications, Int. Conf. Data Min. (2006) 613–622.

[34] B. Perozzi, R. Alrfou, S. Skiena, DeepWalk: online learning of social representations, Knowl. Discov. Data Min. (2014) 701–710.

[35] A. Grover, J. Leskovec, node2vec: scalable feature learning for networks, KDD 2016 (2016) 855–864.

[36] S. Cao, W. Lu, Q. Xu, Deep neural networks for learning graph representations, in: Thirtieth AAAI Conference on Artificial Intelligence, 2016, pp. 1145–1152.

[37] O. Levy, Y. Goldberg, Neural word embedding as implicit matrix factorization, in: Advances in Neural Information Processing Systems, 2014, pp. 2177–2185.

[38] M. Kim, J. Leskovec, The network completion problem: inferring missing nodes and edges in networks, Proceedings of the Eleventh SIAM International Conference on Data Mining, 2011.

[39] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.A. Manzagol, Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion, J. Mach. Lean. Res. 11 (2010) 3371–3408.

[40] Y. Luo, X. Zhao, J. Zhou, J. Yang, Y. Zhang, W. Kuang, J. Peng, L. Chen, J. Zeng, A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information, Nat. Commun. 8 (2017) 573.

[41] H. Zhang, L. Qiu, L. Yi, Y. Song, Scalable multiplex network embedding, IJCAI (2018) 3082–3088.

[42] Z.M. Zhang, J.X. Tan, F. Wang, F.Y. Dao, Z.Y. Zhang, H. Lin, Early diagnosis of hepatocellular carcinoma using machine learning method, Front. Bioeng. Biotechnol. 8 (2020) 254.

[43] H. Lv, F.Y. Dao, D. Zhang, Z.X. Guan, H. Yang, W. Su, M.L. Liu, H. Ding, W. Chen, H. Lin, iDNA-MS: an integrated computational tool for detecting DNA modification sites in multiple genomes, iScience 23 (2020), 100991.

[44] E.C. Lauterbach, Psychotropic drug effects on gene transcriptomics relevant to Parkinson's disease, Prog. Neuro-Psychopharmacol. Biol. Psychiatry 38 (2012) 107–115.

[45] A. Newman-Tancredi, D. Cussac, V. Audinot, J.P. Nicolas, F. De Ceuninck, J.A. Boutin, M.J. Millan, Differential actions of antiparkinson agents at multiple classes of monoaminergic receptor. II. Agonist and antagonist properties at subtypes of dopamine D(2)-like receptor and alpha(1)/alpha(2)-adrenoceptor, J. Pharmacol. Exp. Ther. 303 (2002) 805–814.

[46] T. Gornemann, S. Jahnichen, B. Schurad, K.P. Latte, R. Horowski, J. Tack, M. Flieger, H.H. Pertz, Pharmacological properties of a wide array of ergolines at functional alpha(1)-adrenoceptor subtypes, Naunyn. Schmiedebergs. Arch. Pharmacol. 376 (2008) 321–330.

[47] L.M. Williams, X. He, T.M. Vaid, A. Abdul-Ridha, A.R. Whitehead, P.R. Gooley, R. A.D. Bathgate, S.J. Williams, D.J. Scott, Diazepam is not a direct allosteric modulator of α 1 -adrenoceptors, but modulates receptor signaling by inhibiting phosphodiesterase-4, Pharmacol. Res. Perspect. 7 (1) (2019) e00455, https://doi.org/10.1002/prp2.455.

[48] J.A. Strong, A. Dalvi, F.J. Revilla, A. Sahay, F.J. Samaha, J.A. Welge, J. Gong, M. Gartner, X. Yue, L. Yu, Genotype and smoking history affect risk of levodopa-induced dyskinesias in Parkinson's disease, Mov. Disord. 21 (5) (2006) 654–659, https://doi.org/10.1002/mds.20785.

[49] G. Al-Tajir, M.S. Starr, D-2 agonists protect rodents against pilocarpine-induced convulsions by stimulating D-2 receptors in the striatum, but not in the substantia nigra, Pharmacol. Biochem. Behav. 39 (1) (1991) 109–113, https://doi.org/10.1016/0091-3057(91)90405-Q.

[50] P. Sancho-Bru, R. Bataller, J. Colmenero, X. Gasull, M. Moreno, V. Arroyo, D.A. Brenner, P. Ginès, Norepinephrine induces calcium spikes and proinflammatory actions in human hepatic stellate cells, Am. J. Physiol. - Gastrointestinal Liver Physiol. 291 (5) (2006) G877–G884, https://doi.org/10.1152/ajpgi.00537.2005.

[51] M. Bhowmik, R. Khanam, N. Saini, D. Vohora, Activation of AKT/GSK3β pathway by TDZD-8 attenuates kainic acid induced neurodegeneration but not

seizures in mice, NeuroToxicology 46 (2015) 44–52, https://doi.org/10.1016/j.neuro.2014.11.008.
[52] M. Kanehisa, S. Goto, KEGG: Kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (2000) 27–30.

Yifan Shang received his MS degree in Computer Science and Technology from Xidian University, in 2020. His research interests include machine learning, data mining and network representation learning.

Lin Gao is Professor of Computer Science at the Xidian University, China. She was a visiting scholar at University of Guelph, Canada from 2004 to 2005. In 2016, as a senior visiting fellow, she visited University of Pennsylvania. Her research interests include Computational bioinformatics, data mining, graph theory and optimization. At present, she focused on computational model and algorithm in omics-data analysis, especially its application in cancer. Funded by the National Natural Science Foundation of China, National Key Research and Development Program of China the Foundation of the Ministry of Education of China and others.

Quan Zou is a professor of Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China. He received his PH.D. from Harbin Institute of Technology, P.R.China in 2009, and worked at Xiamen University and Tianjin University from 2009 to 2018. His research is in the areas of bioinformatics, machine learning and parallel computing. Now he is putting the focus on protein classification, genome assembly, annotation and functional analysis from the next generation sequencing data with parallel computing methods. Several related works have been published by Briefings in Bioinformatics, Bioinformatics, PLOS Computational Biology and IEEE/ACM Transactions on Computational Biology and Bioinformatics. Google scholar showed that his more than 100 papers have been cited more than 4000 times. He is also a reviewer for many impacted journals and NSFC(National Natural Science Foundation of China).

Liang Yu received the BSc, MSc and the PhD degrees in computer science from Xidian University, China, in 2003, 2006 and 2011, respectively. She worked in Xidian University from 2006 to present as an assistant professor, associate professor and professor. She was a visiting scholar at Harvard Medical School, USA from 2017 to 2018. Her research is in the areas of bioinformatics, network medicine, and machine learning. Several related works have been published by Proteomics, International Journal of Biological Sciences, IEEE/ACM Transactions on Computational Biology and Bioinformatics, etc. Funded by the National Natural Science Foundation of China, National Key Research and Development Program of China the Foundation of the Ministry of Education of China and others.