

---

Systems biology

## A multimodal deep learning framework for predicting drug-drug interaction events

**Yifan Deng<sup>1,2</sup>, Xinran Xu<sup>1</sup>, Yang Qiu<sup>1</sup>, Jingbo Xia<sup>1</sup>, Wen Zhang<sup>1,\*</sup> and Shichao Liu<sup>1,\*</sup>**

<sup>1</sup>College of Informatics, Huazhong Agricultural University, Wuhan 430070, China

<sup>2</sup>Electronic Information School, Wuhan University, Wuhan 430072, China

\*To whom correspondence should be addressed.

Associate Editor: XXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

### Abstract

**Motivation:** Drug-drug interactions (DDIs) are one of the major concerns in pharmaceutical research. Many machine learning based methods have been proposed for the DDI prediction, but most of them predict whether two drugs interact or not. The studies revealed that DDIs could cause different subsequent events, and predicting drug-drug interaction-associated events is more useful for investigating the mechanism hidden behind the combined drug usage or adverse reactions.

**Results:** In this paper, we collect DDIs from DrugBank database, and extract 65 categories of DDI events by dependency analysis and events trimming. We propose a multimodal deep learning framework named DDIMDL that combines diverse drug features with deep learning to build a model for predicting drug-drug interaction-associated events. DDIMDL first constructs deep neural network-based sub-models by respectively using four types of drug features: chemical substructures, targets, enzymes and pathways, and then adopts a joint DNN framework to combine the sub-models to learn cross-modality representations of drug-drug pairs and predict DDI events. In computational experiments, DDIMDL produces high-accuracy performances and has high efficiency. Moreover, DDIMDL outperforms state-of-the-art DDI event prediction methods and baseline methods. Among all the features of drugs, the chemical substructures seem to be the most informative. With the combination of substructures, targets and enzymes, DDIMDL achieves an accuracy of 0.8852 and an area under the precision-recall curve of 0.9208.

**Availability:** The source code and data are available at <https://github.com/YifanDengWHU/DDIMDL>

**Contact:** zhangwen@mail.hzau.edu.cn and scliu@mail.hzau.edu.cn

**Supplementary information:** Supplementary data are available at *Bioinformatics* online.

---

### 1 Introduction

The prevalence of polypharmacy is increasing over recent years, especially among the elders who suffer from multiple diseases. The recent study (Kantor *et al.*, 2015; Qato *et al.*, 2016) showed that 67% of elderly Americans took five or more medications in 2010–2011, including prescription drugs, over-the-counter drugs and dietary supplements. However, it has been revealed that drugs may interact with others when they are taken together, and unexpected drug-drug interactions (DDIs) may lead to unexpected adverse drug events (Edwards and Aronson, 2000; Pirmohamed *et al.*, 2004; Baxter and Preston, 2010). In this light,

the more DDIs we know, the better we can take effective measures to prevent such events. The determination of DDIs through wet experiments is labor-intensive and time-consuming. In recent years, researchers collected drug data from the literature, reports, etc., and constructed databases that facilitate the development of computational prediction methods. Therefore, machine learning methods can be applied for predicting DDIs to reduce time and cost (Wishart *et al.*, 2008, 2006; Wang *et al.*, 2009; Kanehisa *et al.*, 2010; Kuhn *et al.*, 2010; Li *et al.*, 2010; Knox *et al.*, 2011; Law *et al.*, 2014).

Many machine learning-based DDI prediction methods have been proposed, and are roughly classified into four categories: similarity-based methods, network-based methods, matrix factorization-based methods and ensemble learning-based methods. The similarity-based methods are one

1

major category among these methods, and assume that similar drugs may have interactions with the same drug. Vilar *et al.* (2012, 2014) predicted novel DDIs based on structural similarity and interaction profile fingerprint similarity. Gottlieb *et al.* (2012) obtained feature vectors based on seven types of drug-drug similarities to describe drug-drug pairs, and then built the logistic regression-based models to predict DDIs. Cheng and Zhao (2014) combined several drug-drug similarities to describe drug-drug pairs and adopted four classifiers to build prediction models. Shi *et al.* (2016) proposed a local classification-based method based on structure similarity and side effect similarity. Ferdousi *et al.* (2017) predicted DDIs based on drug functional similarities. Kastrin *et al.* (2018) proposed a statistical learning method to predict DDIs based on topological and semantic similarity features. The network-based methods infer novel DDIs through the constructed network. Zhang *et al.* (2015) used a label propagation approach on a high-order similarity-based network to predict DDIs and integrated multiple similarities for better solutions. Park *et al.* (2015) applied the random walk with restart on the protein-protein network to predict DDIs. Sridhar *et al.* (2016) proposed a probabilistic approach to infer DDIs from the network, which was constructed based on multiple drug-drug similarities and known interactions. Liu *et al.* (2019) proposed a multi-modal deep auto-encoder based drug representation learning method for the DDI prediction, which can learn unified representations of drugs simultaneously from drug feature networks. The matrix factorization-based methods decompose the adjacency matrix of DDIs into several matrices, and reconstruct the adjacency matrix to identify novel DDIs. Shi *et al.* (2016) presented a triple matrix factorization-based method named TMFUF. Yu *et al.* (2018) developed a novel method DDINMF based on the semi-nonnegative matrix factorization. Zhang *et al.* (2018) proposed the manifold regularized matrix factorization method for DDI prediction. The ensemble learning-based methods reasonably combine several models to achieve better performances than individual models. Deepika and Geetha (2018) adopted a semi-supervised learning framework with network representation learning and meta-learning from four drug datasets to predict DDIs. Zhang *et al.* (2017) applied ensemble methods with chemical, biological, phenotypic and network data to predict potential DDIs. Zhang *et al.* (2019) proposed the sparse feature learning ensemble method with linear neighborhood regularization(SFLNN) for the DDI prediction.

Generally, existing methods are designed to predict whether two drugs interact or not, and have greatly contributed to better understanding of DDIs. However, DDIs can lead to different biological consequences or events. Predicting DDI-associated events is a meaningful and challenging task, and has received some attention. Herrero-Zazo *et al.* (2013) built a manually annotated corpus for DDIs in biomedical texts. They collected DDIs from DrugBank and MedLine, and annotated the DDI relationships into four types: mechanism (pharmacokinetic mechanism), effect (pharmacodynamic mechanism), advice (not taking these two drugs together) and int (without further information). Ryu *et al.* (2018) classified the biological events collected from DrugBank into 86 types and built the deep learning-based model based on drug chemical substructures to predict DDI events. Later, Lee *et al.* (2019) directly merged three features as input of a DNN to build a prediction model.

Although the above works have made crucial efforts on the event prediction, there still exists space for improvement. First, DrugBank is a reliable data source and has unified syntax in describing DDIs, so we can extract DDI events through the standard descriptions and conduct further study. Second, there are a variety of drug features in DrugBank. Ryu *et al.* (2018) only made use of the drug chemical substructures. Considering more features is necessary for comprehensive study and better performances. Third, owing to redundancy between features, how to effectively combine different features is a challenge. Directly merging different feature vectors is a commonly used approach, but we need more

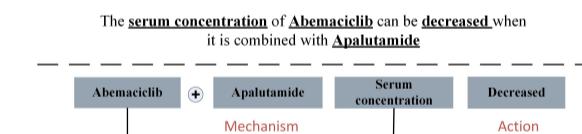
powerful and useful ways of combining different features. To address the above issues, we present a computational method named DDIMDL that combines diverse drug features with deep learning to predict DDI events. We define a standard protocol to analyze the DDI events in DrugBank and select 65 types of major events for analysis. Moreover, we collect four features of drugs: chemical substructures, targets, enzymes and pathways for modeling. In DDIMDL, four sub-models are constructed by using each drug feature and a joint DNN framework is used to combine the sub-models to learn cross-modality representations of drug-drug pairs. At last, we predict DDI-associated events with the learned cross-modality representations. To summarize, the main contributions of this paper are described as below:

- We focus on the fine-grained descriptions of known DDIs in DrugBank and build an interaction event dataset by applying semantic analysis to these descriptions.
- We present a multimodal deep learning framework named DDIMDL, which takes advantage of deep learning and diverse features of drugs to predict DDI events.
- The experimental results show that DDIMDL has high efficiency and produces high accuracy, and outperforms the compared methods.

## 2 Materials

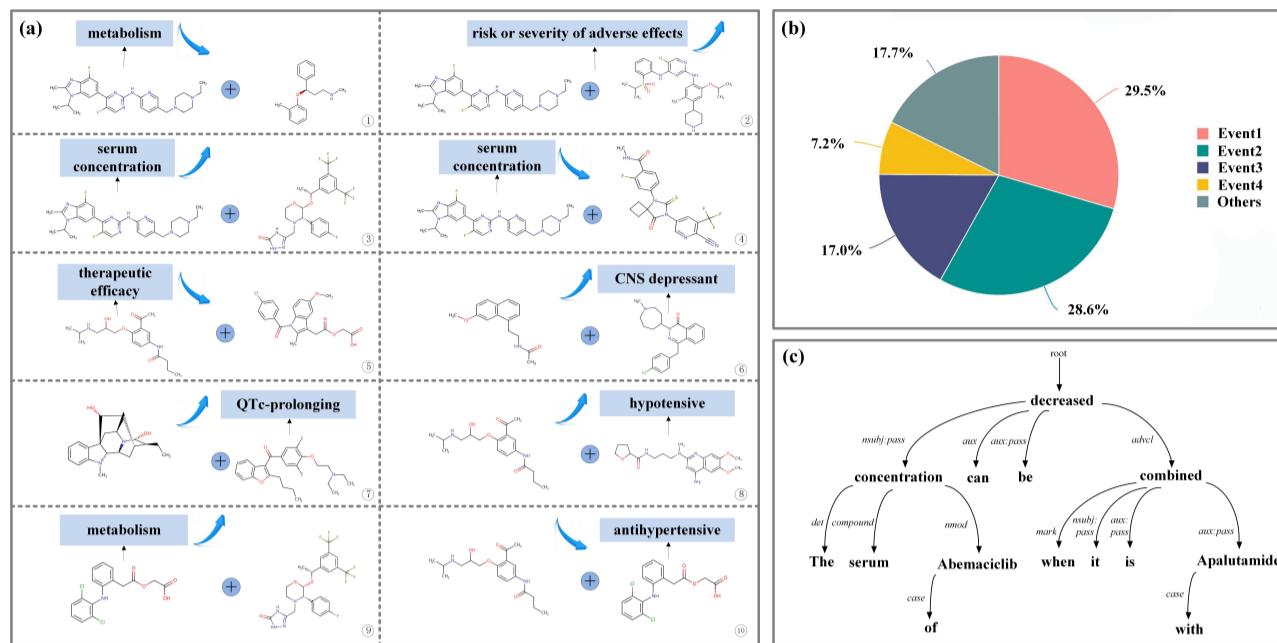
### 2.1 Datasets

DrugBank (Knox *et al.*, 2011) is a resource that provides comprehensive information about 12,151 drugs, including 3,844 FDA approved drugs and 5,867 experimental drugs. In the study, we collect DDIs as well as four features of drugs: chemical substructures, targets, pathways, and enzymes from DrugBank. After we obtain the pathways (KEGG ID) of a drug from DrugBank, we can use KEGG database to transform the KEGG ID into corresponding drug-pathway information. Targets and e of drugs are directly obtained from databases. We select drugs which have interactions with others and have all four features, and then obtain 641 drugs with 105,824 pairwise DDIs.



**Fig. 1.** An example about the sentence describing a DDI event

Further, we pay attention to the DDIs and their descriptions in the DrugBank, and observe that the interactions are described by several sentence-types. An example of "the serum concentration of Abemaciclib can be decreased when it is combined with Apalutamide" is shown in Figure 1. The sentence describes a DDI event between Abemaciclib and Apalutamide. For better understanding of the DDI events, we define the representation of DDI event as a four-tuple structure: **(drug A, drug B, mechanism, action)**, where the "mechanism" means the effect of drugs in terms of the metabolism, the serum concentration, the therapeutic efficacy and so on. The "action" represents the increase or decrease after lemmatization. Since descriptions about DDIs in DrugBank have fixed syntax, we apply StanfordNLP tool (Qi *et al.*, 2018) to obtain the dependency relationship. By using the tool, we can obtain a tuple for each word. We selected *index* (index of the word in the sentence), *text* (the word's content), *governor* (the index of the word



**Fig. 2.** (a): Top 10 frequent events numbered from #1 to #10. Event #1: the metabolism of drug A can be decreased when combined with drug B (19,620 DDIs). Event #2: the risk or severity of adverse effects can be increased when drug A is combined with drug B (18,992). Event #3: the serum concentration of drug A can be increased when it is combined with drug B (11,292). Event #4: the serum concentration of drug A can be decreased when it is combined with drug B (4,772). Event #5: the therapeutic efficacy of drug A can be decreased when used in combination with drug B (2,624). Event #6: drug B may increase the central nervous system depressant (CNS depressant) activities of drug A (2,264). Event #7: drug B may increase the QTc-prolonging activities of drug A (2,204). Event #8: drug B may increase the hypotensive activities of drug A (2,172). Event #9: the metabolism of drug A can be increased when combined with drug B (1,390). Event #10: drug B may decrease the antihypertensive activities of drug A (1,102). (b): Distribution of DDI events. (c): Dependency relationship tree of an event

which governs it) and *relation* (relationship with its governor) from the original tuple and construct our word expression profile  $W$ . Then we construct the dependency relationship tree as shown in Figure 2(c). We use abbreviations of the NLP terminology here, their full names can be found in Supplementary Table S2. Since the pre-trained model isn't trained on biomedical texts, it can't arrange proper parts of speech for some of the professional drugs' names. To solve the deficiency in the model, we build a list of drugs' names and perform named entity recognition to capture the drugs. The *action* of events is always the root of the dependency tree, thus we start from the *action* and search for the *mechanism*'s root which has a special dependency relationship (i.e. object, passive nominal subject) with the action, then perform traversal to its subtree and obtain the whole *mechanism*. We name the drugs whose efficiency is influenced *drugA* and the others *drugB*. In this way, our four-tuple structure  $\{drugA, drugB, mechanism, action\}$  is constructed. The detailed procedure is shown in Algorithm 1. In this way, we obtain a total of 110 types of events to describe DDIs. More than 99% of the interactions are associated with only one event, so we remove DDIs associated with more than one event. For the sake of analysis, we remove the rare events and select the events which have more than ten DDIs.

Thus, we obtain 572 drugs with 74,528 pairwise DDIs, which are associated with 65 types of events. We number these events from #1 to #65 according to the descending order of their occurrence frequency. The events numbered from #1 to #10 are given in Figure 2(a) as examples, and the percentage of all events is shown in Figure 2(b). Details of all events are provided in Supplementary Table S1.

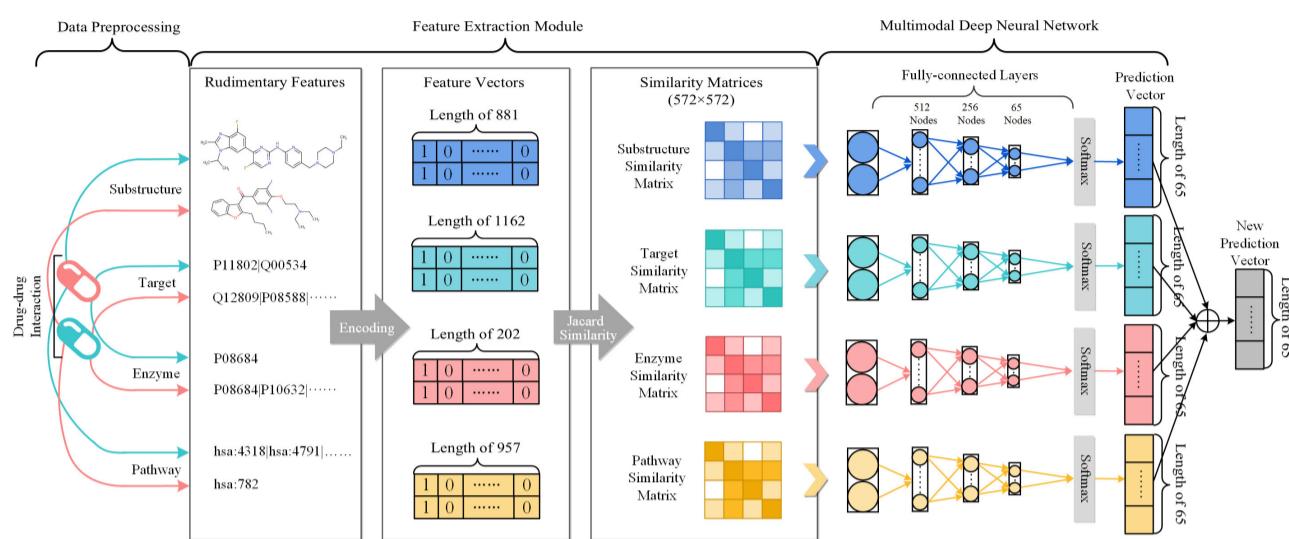
#### Algorithm 1 Procedure of extracting tuples from dependency relationship

**Input:** Word expression profile  $W = \{index, text, governor, relation\}$ ,  
A list  $D$  consisting of drugs' names.  
**Output:** Tuple  $\{drugA, drugB, mechanism, action\}$

- 1: Scan through the words to construct  $sons$  and find  $root$  of the tree.
- 2:  $action=W[root].text$
- 3:  $drugA, drugB = \text{named entity recognition according to } D$
- 4: Find the *mechanismRoot* which has a relationship of  
"obj" or "nsubj:pass" in  $sons[root]$ . Put it in *quene*.
- 5:  $\text{addMechanism}(node)$ :
- 6: **if**  $sons[node]=\text{null}$  **then**
- 7:   return
- 8: **else**
- 9:   **for**  $i$  in  $sons[node]$  **do**
- 10:     *quene.append*( $i$ )
- 11:     *addMechanism*( $i$ )
- 12:   **end for**
- 13: **end if**
- 14:  $\text{addMechanism}(mechanismRoot)$
- 15:  $\text{sort}(quene)$
- 16: **for**  $i$  in *quene* **do**
- 17:     *mechanism+=W[i].text*
- 18: **end for**
- 19: **return**  $\{drugA, drugB, mechanism, action\}$

## 2.2 Overview of DDIMDL method

In this study, we design a multimodal deep neural network named DDIMDL that combines diverse features of drugs to predict DDI-associated events.



**Fig. 3.** The pipeline of the proposed DDIMDL method

The pipeline of the DDIMDL method is shown in Figure 3. First, we use four types of drug features to calculate drug-drug similarities, and use them as the representations of drugs. The representations of drugs are respectively fed into the sub-models based on multiple-layer neural networks. Then, we combine the sub-models to learn cross-modality representations of drug-drug pairs and predict DDI events with the learned cross-modality representations. The key components of DDIMDL are introduced below.

### 2.2.1 Feature extraction module

The multimodal deep learning prediction method takes diverse features of drugs as modals and combines them to make predictions. The feature extraction and representations are critical for model construction.

As illustrated in Section 2.1, we have four features of drugs: chemical substructures, targets, enzymes and pathways, which bring diverse information about drugs. Each feature corresponds to a set of descriptors, and thus a drug can be represented by a binary feature vector, whose values (1 or 0) indicate the presence or absence of the corresponding descriptors. Here, we take the chemical substructures as an example. Pubchem defines 881 types of substructures, namely molecular fingerprints. According to the chemical substructures, the representation of a drug is an 881-dimensional bit vector and the value 1 or 0 denotes the presence or absence of a certain substructure type. Similarly, there are 1162 types of targets associated with drugs, and the target feature can encode a drug  $i$  into a 1162-dimensional bit vector. In this way, a drug can be represented as four types of vectors by using four features.

These features vectors have high dimensions and values of most dimensions are 0, thus we compress features and reduce the sparsity. Instead of using the bit vectors as input, we calculate the pairwise drug-drug similarity from bit vectors using Jaccard similarity measure. Jaccard similarity is calculated by Equation (1).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (1)$$

where  $A$  and  $B$  are in the form of sets for the bit vectors of two drugs;  $|A \cap B|$  is the intersection of  $A$  and  $B$ ;  $|A \cup B|$  is the union. We can obtain four  $572 \times 572$  pairwise drug-drug similarity matrices: chemical substructure similarity matrix  $S^{CS}$ , target similarity matrix  $S^T$ , enzyme similarity matrix  $S^E$  and pathway similarity matrix  $S^P$ . Based on a drug-drug

similarity matrix, each drug is then represented as a corresponding 572-dimensional row vector. Let  $V_i^{CS}, V_i^T, V_i^E, V_i^P$  denote different feature vectors of drug  $i$ . The representation of drug pair  $(i, j)$  is the concatenation of feature vectors of the two drugs, and the representations of drug pairs in four modals are denoted as  $(V_i^{CS}, V_j^{CS}), (V_i^T, V_j^T), (V_i^E, V_j^E)$  and  $(V_i^P, V_j^P)$ .

As shown in Figure 3, the feature extraction module generates different types of feature vectors, and use them as inputs to sub-models.

### 2.2.2 Construction of DDIMDL and model optimization

Since we have several features, we construct the sub-models based on each feature using the deep neural network (DNN).

A deep neural network (DNN) is an artificial neural network (ANN) with multiple layers between the input and output layers. The DNN can find the correct mathematical manipulation to turn the input into the output, whether it is a linear relationship or a non-linear relationship. The design of networks for sub-models is inspired by the bottleneck-like neural network VGG16 (Simonyan and Zisserman, 2014), which was originally proposed for image classification and won the silver medals in ILSVRC 2014. Bottleneck features are generated from a multi-layer perceptron in which one of the internal layers has a small number of hidden units, relative to the size of the other layers (Yu and Seltzer, 2011). This small layer creates a constriction in the network that forces the information pertinent to classification into a low dimensional representation. By using bottleneck features, we can reduce the number of parameters that need to be trained.

The forward propagation is calculated by the following formulas:

$$\begin{aligned} h^l &= W^l a^l + b^l \\ a^{l+1} &= \text{ReLU}(h^l) \end{aligned} \quad (2)$$

Where  $W^l$  implies the weights of neurons between layer  $l$  and layer  $l+1$ ,  $l = 1, 2, \dots, n$  and  $b^l$  is the bias.  $a^l$  and  $a^{l+1}$  are the input and output of the forward propagation. We use Rectified Linear Unit (ReLU) (Nair and Hinton, 2010) as the activation function. The sub-models produce the prediction after a softmax layer. Between layers, we add batch normalization layers (Ioffe and Szegedy, 2015) to accelerate the convergence, and add dropout layers (Srivastava *et al.*, 2014) to avoid over-fitting and enhance generalization ability.

The combination of different sub-models is important for DDIMDL. Here, the outputs from sub-models are combined by the average operator

to produce the final prediction. As shown in Figure 3, we construct sub-models from different features and then combine them to build the prediction model.

We adopt cross-entropy as the loss function, and empirically train and optimize the DDIMDL model. We use the early-stopping strategy (Prechelt, 1998) which automatically stops the training if no improvement is observed in 10 epochs. The strategy could prevent over-fitting while greatly accelerate the training procedure. We use a batch size of 256 and use Adam (Kingma and Ba, 2014) with default parameters as the optimization algorithm to train the networks.

### 3 Experiments and results

#### 3.1 Evaluation metrics

Researchers usually pay attention to three tasks in DDI prediction. The first is predicting unobserved interactions between known drugs, the second is predicting the interactions between known drugs and new drugs, the third is predicting the interactions between new drugs. We describe how to evaluate the performances of DDIMDL in implementing the three tasks.

For task 1, we apply 5-fold cross-validation (5-CV) to DDIs and split all DDIs into five subsets. We train models based on DDIs in the training set, and then make predictions for DDIs in the test set. For task 2 and task 3, we apply 5-CV to drugs instead of DDI pairs. We randomly split drugs into five folds, and used four of them as training drugs and the remaining as test drugs. For task 2, prediction models are constructed on the DDIs between training drugs, and then make predictions for DDIs between training drugs and test drugs. For task 3, prediction models are constructed on the DDIs between training drugs, and then make predictions for DDIs between test drugs.

Our task is the multi-class classification work. For evaluation, accuracy (ACC), area under the precision-recall-curve (AUPR), area under the ROC curve (AUC), F1 score and Precision are adopted as the evaluation metrics. We use micro metrics for AUPR and AUC, while macro metrics for others (micro-Precision, micro-Recall and micro-F1 are equal to accuracy in the multi-class problem).

#### 3.2 Parameter setting

The parameters could influence the performances of DDIMDL models. Sub-models are key components of DDIMDL, and the structures of sub-models are critical, and we also add the dropout layers. Therefore, we discuss the number of layers and the dropout rate.

First, we discuss the number of layers for structures of sub-models. We set the rule that the number of neurons for a layer is half of that for its former layer, and we fix the number of neurons of the last hidden layer to 256. We empirically consider two hidden layers, three hidden layers, four hidden layers and five hidden layers. The metric scores and training time under different configurations are shown in Table 1. Consequently, we adopt the network structure with three layers, because it achieves the best performances. Then, we discuss the influence of the dropout rate. We set the dropout rate from 0 to 0.5 with 0.1 as the step, and obtain the highest ACC and AUPR when the dropout rate equals 0.3.

Table 1. Metrics and training time of DDIMDL with different numbers of layers

layers	ACC	AUPR	AUC	F1	time(s)
2	0.7938	0.8346	0.9959	0.6228	1733.2590
3	<b>0.8852</b>	<b>0.9208</b>	<b>0.9979</b>	<b>0.7419</b>	<b>607.0318</b>
4	0.8836	0.9200	0.9977	0.7630	662.2027
5	0.8820	0.9184	0.9977	0.8820	873.2874

Therefore, we adopt the network with three layers, numbers of whose nodes are 512, 256 and 65, to build the sub-models in the following experiments, and set the dropout rate to 0.3.

#### 3.3 Feature evaluation

In this section, we evaluated the influence of different drug features on the DDI event prediction. We use each feature or feature combinations to build DDIMDL models, and adopt the metric scores of the models as the indicators of the usefulness of corresponding features or feature combinations. The results of all prediction models are shown in Table 2.

Among all drug features, the chemical substructure appears to be most informative and achieves an accuracy of 0.8623. The model based on the target feature produces an accuracy of 0.8338, and the model based on the pathway feature produces an accuracy of 0.8182. The enzyme feature leads to the model with an accuracy of 0.6687. The combination of features provides the significant improvement compared with individual features. The combination of substructures and targets produces the best performance among all combinations of two features; The combination of substructures, targets and enzymes performs best in all combinations of three features. Furthermore, we observe that using all features doesn't lead to better results than the combination of substructures, targets and enzymes. Therefore, the DDIMDL models in the following experiments are constructed by combining three features: substructures, targets and enzymes.

Moreover, the precision-recall curves of all DDIMDL models are shown in Figure 4(a). They also reveal intuitively that the application of multimodal learning enhances the performances of DDIMDL in the DDI event prediction.

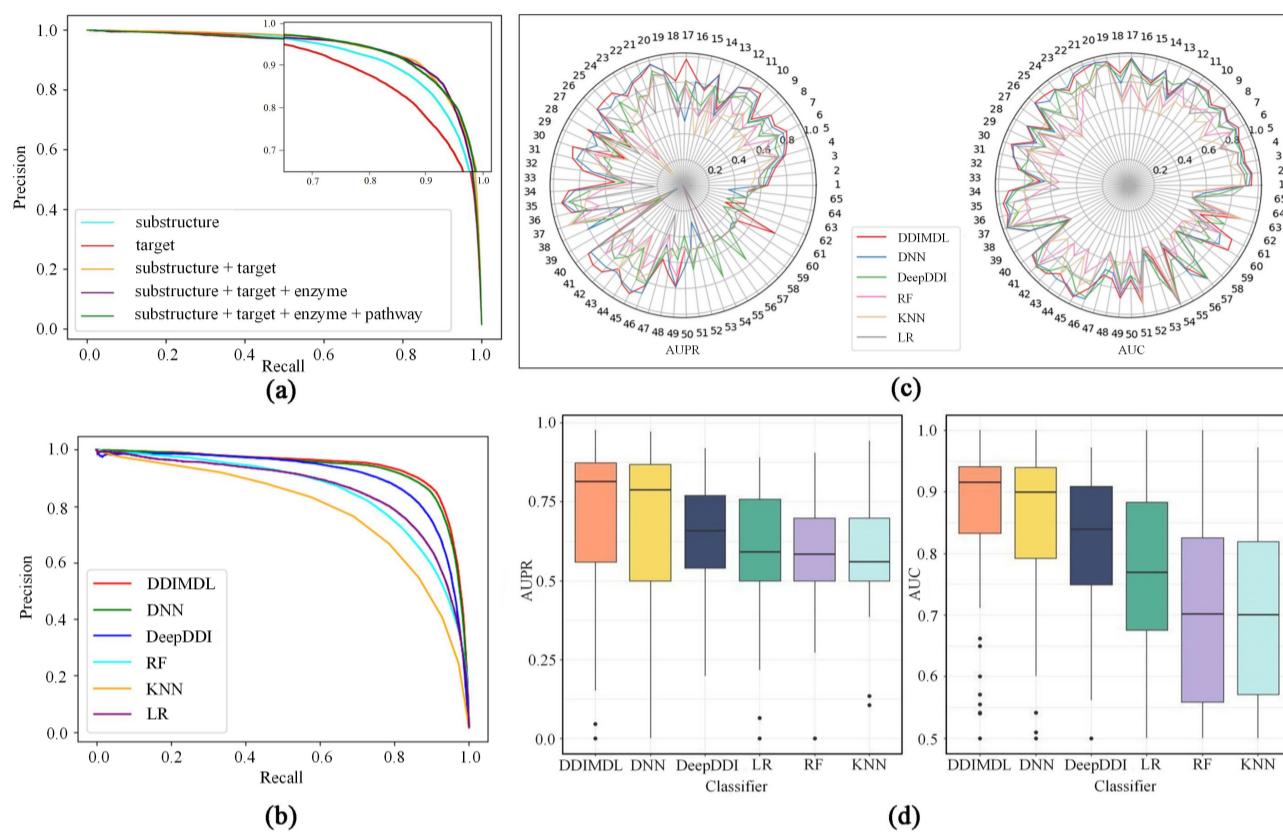
Table 2. The performance of DDIMDL with different feature combinations.

Feature set	ACC	AUPR	AUC	F1	Precision	Recall
S	0.8623	0.9136	0.9975	0.7324	0.7831	0.7006
T	0.8338	0.8979	0.9969	0.7084	0.7579	0.6788
P	0.8182	0.8876	0.9972	0.6875	0.7611	0.6495
E	0.6687	0.7384	0.9913	0.4105	0.4943	0.3714
S+T	0.8806	0.9192	0.9981	0.7625	0.8231	0.7283
S+P	0.8786	0.9188	0.9981	0.7611	0.8326	0.7223
S+E	0.8655	0.8939	0.9970	0.7263	0.8324	0.6821
T+P	0.8344	0.9004	0.9976	0.7012	0.7781	0.6660
T+E	0.8506	0.8860	0.9970	0.6974	0.7770	0.6564
P+E	0.8423	0.8809	0.9968	0.6664	0.7344	0.6279
S+T+P	0.8625	0.9202	<b>0.9982</b>	0.7330	0.7941	0.6950
S+T+E	<b>0.8852</b>	<b>0.9208</b>	0.9979	<b>0.7585</b>	<b>0.8471</b>	<b>0.7182</b>
S+P+E	0.8778	0.9153	0.9978	0.7321	0.8134	0.6905
T+P+E	0.8488	0.8956	0.9974	0.6967	0.7608	0.6591
S+T+P+E	0.8725	0.9178	0.9979	0.7361	0.8348	0.6938

<sup>1</sup> S:substructure, T:target, P:pathway and E:enzyme

#### 3.4 Method Comparison

We compare DDIMDL with one state-of-the-art event prediction method DeepDDI (Ryu *et al.*, 2018). We also consider several popular classification approaches, i.e. random forest (RF), k-nearest neighbor (KNN) and logistic regression (LR), and build sub-models as DDIMDL does, and combine them to make predictions. We compare DDIMDL with these models to demonstrate the advantages of sub-models based on networks. Moreover, we implement a deep neural network (DNN) which has the same structure as DDIMDL's sub-models, but the DNN



**Fig. 4.** Performances of all prediction models. (a) The precision-recall curves of DDIMDL models with different features or feature combinations. (b) AUPR and AUC of compared methods for each event. (c) The precision-recall curves of compared methods. (d) Boxplots displaying the AUPR and AUC of compared methods for each event.

Table 3. The performance of different methods

	Method	ACC	AUPR	AUC	F1	Precision	Recall
Task 1	DDIMDL	<b>0.8852</b>	<b>0.9208</b>	<b>0.9976</b>	<b>0.7585</b>	<b>0.8471</b>	<b>0.7182</b>
	DeepDDI	0.8371	0.8899	0.9961	0.6848	0.7275	0.6611
	DNN	0.8797	0.9134	0.9963	0.7223	0.8047	0.7027
	RF	0.7775	0.8349	0.9956	0.5936	0.7893	0.5161
	KNN	0.7214	0.7716	0.9813	0.4831	0.7174	0.4081
	LR	0.7920	0.8400	0.9960	0.5948	0.7437	0.5236
Task 2	DDIMDL	<b>0.6415</b>	<b>0.6558</b>	<b>0.9799</b>	<b>0.4460</b>	<b>0.5607</b>	<b>0.4319</b>
	DeepDDI	0.5774	0.5594	0.9575	0.3416	0.3630	0.3890
Task 3	DNN	0.6239	0.6361	0.9796	0.2997	0.4237	0.2840
	DDIMDL	0.4075	0.3635	0.9512	<b>0.1590</b>	<b>0.2408</b>	<b>0.1452</b>
	DeepDDI	0.3602	0.2781	0.9059	0.1373	0.1586	0.1450
	DNN	<b>0.4087</b>	<b>0.3776</b>	<b>0.9550</b>	0.1152	0.1836	0.1093

concatenates all features straightly as its input. The same structure is adopted by (Lee *et al.*, 2019). We compare DDIMDL with it to demonstrate the usefulness of multimodal deep learning. Thus, the models based on RF, KNN, LR and DNN are used as the baseline methods.

We implement DeepDDI according to the descriptions in (Ryu *et al.*, 2018). The network of DeepDDI has eight hidden layers, and each layer has 2048 nodes. We adjust the output nodes of DeepDDI to 65, because our problem has 65 types of events. We also optimize parameters of baseline methods. We set 100 decision trees for RF and set the neighbor number of KNN to 4.

Task 1 is important in drug-drug interaction prediction. We evaluate performances of all prediction methods for task 1, which predicts DDI-associated events between known drugs. The evaluation scores of all prediction models are shown in Table 3, and the results demonstrate that DDIMDL produces better performances than other methods in terms of all metrics. The precision-recall curves of all models are also shown in Figure 4(b), and DDIMDL performs best among all methods. Further, we investigate the performances of DDIMDL for each event and calculate the metric scores for events independently by using predicted scores and real labels. The AUPR scores and AUC scores of all prediction models for each event are shown in Figure 4(c). The original AUPR scores and AUC scores are listed in Supplementary Tables S3&S4. It is likely that the events with higher frequency can gain better performances. DDIMDL produces AUPR scores greater than 0.5 for the events numbered from #1 to #46 except the event #39. It can be observed there are 49 DDIs associated with event #39 and 38 of them are wrongly classified to event #1. The event #1 is described as "the metabolism of drug A can be decreased when combined with drug B"; the event #39 is described as "the serum concentration of the active metabolites of drug A can be reduced when the drug is used in combination with drug B resulting in a loss in efficacy". Therefore, both events are related to drug metabolism, and that is the possible reason why most instances of event #39 were classified into event #1. In general, Figure 4(c) demonstrates that DDIMDL produces greater AUPR scores and AUC scores than other methods in most types of events. To further analyze the performances of prediction models, we use Figure 4(d) to display AUPR scores and AUC scores of different methods for 65 types of events. These boxplots clearly show that DDIMDL produces statistically better performances for these events than compared methods

and it is worth mentioning that the DDIMDL based on multimodal deep learning significantly improves the performance of DNN.

Moreover, we evaluate the performances of prediction models for task 2 and task 3. Here, we mainly compare DDIMDL with DeepDDI and DNN, because the results for task 1 show that DeepDDI and DNN are more competitive. The performances of all prediction methods are shown in Table 3. It could be concluded that without prior knowledge about the new drugs, the performances of all models for task 2 and task 3 decrease, especially for task 3. The experimental results also demonstrate that DDIMDL outperforms DeepDDI for task 2 and task 3, which corroborates the efficiency of multimodal deep learning again. Although DDIMDL has similar ACC, AUC and AUPR as DNN for task 3, DDIMDL outperforms DNN in terms of macro F1, Precision and Recall.

The studies show that deep learning and diverse features are critical for the DDI event prediction, and our multimodal deep learning framework outperforms the traditional classifiers and deep network structures in previous studies (Ryu *et al.*, 2018; Lee *et al.*, 2019).

### 3.5 Case study

In this section, we conduct case studies to validate the usefulness of DDIMDL in practice.

We use all DDIs and their events in our dataset which were originally from DrugBank to train the prediction model, and then make predictions for other drug-drug pairs. We pay attention to five events with the highest frequencies numbered from #1 to #5, and check up on the top 20 predictions related to each event. We used the Interactions Checker tool provided by drugs.com to validate these predictions.

Table 4. Confirmed DDIs and their associated events

Event type	DrugBank IDs	Drug names
#1	DB00390;DB00153	Digoxin and Ergocalciferol
#2	DB01576;DB00574	Dextroamphetamine and Fenfluramine
#3	DB00978;DB00261	Lomefloxacin and Anagrelide
#4	DB00988;DB01142	Dopamine and Doxepin
#5	DB00515;DB01137	Cisplatin and Levofloxacin

Five drug-drug interaction events can be confirmed, and they are shown in Table 4. For example, the interaction between Dextroamphetamine and Fenfluramine is predicted to cause the event #2, and means the risk or severity of adverse effects can be increased when Dextroamphetamine is combined with Fenfluramine. According to drugs.com, the evidence shows that Fenfluramine may increase the effects of Dextroamphetamine, and side effects such as jitteriness, nervousness, anxiety, restlessness, and racing thoughts have been reported. The literature (Prior *et al.*, 2002) also supports this finding. More evidence about confirmed drug-drug interaction events is provided in the Supplementary Table S5.

### 4 Conclusion

This study obtains the DDI data from DrugBank, and applies NLP techniques to classify DDI-associated events into 65 types according to their descriptions' syntax, and compiles a dataset of 572 drugs, 74,528 interactions and 65 types of DDI-associated events. A multimodal deep learning framework named DDIMDL that combines diverse drug features with deep learning is presented for the DDI event prediction. Evaluated by using five-fold cross-validation, DDIMDL outperforms the existing DDI event prediction method and baseline methods. The case studies are also performed to identify the DDI events not included in our dataset, and several DDI-associated events, such as the event caused by the interaction

between Dextroamphetamine and Fenfluramine, are successfully found out. In conclusion, the usage of semantic analysis enables us to conspicuously classify the events from DrugBank, and multimodal learning provides a mighty way to integrate diverse features and cost reasonable training time. The multimodal deep learning framework is a promising tool for the DDI event prediction.

There exist issues we will address in the future work to improve the DDI event prediction. First, the numbers of DDIs are extremely imbalanced for different events, and we will consider novel techniques of dealing with the imbalanced dataset to relieve it. Second, the numbers of interactions are inadequate for some events, thus deep learning method is easy to go underfitting, and we will consider data augment techniques to enlarge the event dataset.

### 5 Acknowledgements

None

### 6 Funding

This work has been supported by the National Natural Science Foundation of China (61772381, 61572368), the Fundamental Research Funds for the Central Universities (2042017kf0219, 2042018kf0249, 2662019QD011), National Key Research and Development Program (2018YFC0407904), Huazhong Agricultural University Scientific & Technological Self-innovation Foundation. The funders have no role in study design, data collection, data analysis, data interpretation, or writing of the manuscript.

### References

- Baxter, K. and Preston, C. L. (2010). *Stockley's drug interactions*, volume 495. Pharmaceutical Press London.
- Cheng, F. and Zhao, Z. (2014). Machine learning-based prediction of drug-drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *J Am Med Inform Assoc*, **21**(e2), e278–86.
- Deepika, S. and Geetha, T. (2018). A meta-learning framework using representation learning to predict drug-drug interaction. *Journal of biomedical informatics*, **84**, 136–147.
- Edwards, I. R. and Aronson, J. K. (2000). Adverse drug reactions: definitions, diagnosis, and management. *The lancet*, **356**(9237), 1255–1259.
- Ferdousi, R., Safdari, R., and Omidi, Y. (2017). Computational prediction of drug-drug interactions based on drugs functional similarities. *Journal of biomedical informatics*, **70**, 54–64.
- Gottlieb, A., Stein, G. Y., Oron, Y., Ruppin, E., and Sharan, R. (2012). Indi: a computational framework for inferring drug interactions and their associated recommendations. *Molecular Systems Biology*, **8**(1), 592.
- Herrero-Zazo, M., Segura-Bedmar, I., Martínez, P., and Declerck, T. (2013). The ddi corpus: An annotated corpus with pharmacological substances and drug-drug interactions. *Journal of biomedical informatics*, **46**(5), 914–920.
- Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.
- Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., and Hirakawa, M. (2010). Kegg for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res*, **38**(Database issue), D355–60.
- Kantor, E. D., Rehm, C. D., Haas, J. S., Chan, A. T., and Giovannucci, E. L. (2015). Trends in prescription drug use among adults in the united states from 1999–2012prescription drug use in us adults 1999–2012prescription drug use in us adults 1999–2012. *JAMA*, **314**(17), 1818–1830.
- Kastrin, A., Ferk, P., and LeskoÅ¡ek, B. (2018). Predicting potential drug-drug interactions on topological and semantic similarity features using statistical learning. *PLoS one*, **13**(5), e0196865.
- Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C., and Wishart, D. S. (2011). Drugbank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res*, **39**(Database issue), D1035–41.
- Kuhn, M., Campillos, M., Letunic, I., Jensen, L. J., and Bork, P. (2010). A side effect resource to capture phenotypic effects of drugs. *Molecular systems biology*, **6**(1).

- Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A. C., Liu, Y., Maciejewski, A., Arndt, D., Wilson, M., Neveu, V., Tang, A., Gabriel, G., Ly, C., Adamjee, S., Dame, Z. T., Han, B., Zhou, Y., and Wishart, D. S. (2014). Drugbank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.*, **42**(Database issue), D1091–7.
- Lee, G., Park, C., and Ahn, J. (2019). Novel deep learning model for more accurate prediction of drug-drug interaction effects. *BMC bioinformatics*, **20**(1), 415.
- Li, Q., Cheng, T., Wang, Y., and Bryant, S. H. (2010). Pubchem as a public resource for drug discovery. *Drug Discov Today*, **15**(23-24), 1052–7.
- Liu, S., Huang, Z., Qiu, Y., Chen, Y.-P. P., and Zhang, W. (2019). Structural network embedding using multi-modal deep auto-encoders for predicting drug-drug interactions. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 445–450. IEEE.
- Nair, V. and Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814.
- Park, K., Kim, D., Ha, S., and Lee, D. (2015). Predicting pharmacodynamic drug-drug interactions through signaling propagation interference on protein-protein interaction networks. *PLOS ONE*, **10**(10), e0140816.
- Pirmohamed, M., James, S., Meakin, S., Green, C., Scott, A. K., Walley, T. J., Farrar, K., Park, B. K., and Breckenridge, A. M. (2004). Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *Bmj*, **329**(7456), 15–19.
- Prechelt, L. (1998). Early stopping-but when? In *Neural Networks: Tricks of the trade*, pages 55–69. Springer.
- Prior, F. H., Isbister, G. K., Dawson, A. H., and Whyte, I. M. (2002). Serotonin toxicity with therapeutic doses of dexamphetamine and venlafaxine. *The Medical journal of Australia*, **176**(5), 240–241.
- Qato, D. M., Wilder, J., Schumm, L. P., Gillet, V., and Alexander, G. C. (2016). Changes in prescription and over-the-counter medication and dietary supplement use among older adults in the united states, 2005 vs 2011. *JAMA internal medicine*, **176**(4), 473–482.
- Qi, P., Dozat, T., Zhang, Y., and Manning, C. D. (2018). Universal dependency parsing from scratch. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 160–170, Brussels, Belgium. Association for Computational Linguistics.
- Ryu, J. Y., Kim, H. U., and Lee, S. Y. (2018). Deep learning improves prediction of drug-drug and drug-food interactions. *Proceedings of the National Academy of Sciences*, **115**(18), E4304–E4311.
- Shi, J.-Y., Gao, K., Shang, X.-Q., and Yiu, S.-M. (2016). Lcm-ds: a novel approach of predicting drug-drug interactions for new drugs via dempster-shafer theory of evidence. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 512–515. IEEE.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sridhar, D., Fakhraei, S., and Getoor, L. (2016). A probabilistic approach for collective similarity-based drug-drug interaction prediction. *Bioinformatics*, **32**(20), 3175–3182.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, **15**(1), 1929–1958.
- Vilar, S., Harpz, R., Uriarte, E., Santana, L., Rabadian, R., and Friedman, C. (2012). Drug-drug interaction through molecular structure similarity analysis. *J Am Med Inform Assoc*, **19**(6), 1066–74.
- Vilar, S., Uriarte, E., Santana, L., Lorberbaum, T., Hripesak, G., Friedman, C., and Tatonetti, N. P. (2014). Similarity-based modeling in large-scale prediction of drug-drug interactions. *Nat Protoc*, **9**(9), 2147–63.
- Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., and Bryant, S. H. (2009). Pubchem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.*, **37**(Web Server issue), W623–33.
- Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., and Woolsey, J. (2006). Drugbank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.*, **34**(Database issue), D668–72.
- Wishart, D. S., Knox, C., Guo, A. C., Cheng, D., Shrivastava, S., Tzur, D., Gautam, B., and Hassanali, M. (2008). Drugbank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res.*, **36**(Database issue), D901–6.
- Yu, D. and Seltzer, M. L. (2011). Improved bottleneck features using pretrained deep neural networks. In *Twelfth annual conference of the international speech communication association*.
- Yu, H., Mao, K.-T., Shi, J.-Y., Huang, H., Chen, Z., Dong, K., and Yiu, S.-M. (2018). Predicting and understanding comprehensive drug-drug interactions via semi-nonnegative matrix factorization. *BMC systems biology*, **12**(1), 14.
- Zhang, P., Wang, F., Hu, J., and Sorrentino, R. (2015). Label propagation prediction of drug-drug interactions based on clinical side effects. *Scientific reports*, **5**, 12339.
- Zhang, W., Chen, Y., Liu, F., Luo, F., Tian, G., and Li, X. (2017). Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data. *BMC bioinformatics*, **18**(1), 18.
- Zhang, W., Chen, Y., Li, D., and Yue, X. (2018). Manifold regularized matrix factorization for drug-drug interaction prediction. *Journal of biomedical informatics*, **88**, 90–97.
- Zhang, W., Jing, K., Huang, F., Chen, Y., Li, B., Li, J., and Gong, J. (2019). Sfln: a sparse feature learning ensemble method with linear neighborhood regularization for predicting drug-drug interactions. *Information Sciences*, **497**, 189–201.