



# A Review of Approaches for Predicting Drug–Drug Interactions Based on Machine Learning

Ke Han<sup>1,2\*†</sup>, Peigang Cao<sup>3†</sup>, Yu Wang<sup>1</sup>, Fang Xie<sup>1</sup>, Jiaqi Ma<sup>1</sup>, Mengyao Yu<sup>1</sup>, Jianchun Wang<sup>1</sup>, Yaoqun Xu<sup>1</sup>, Yu Zhang<sup>1</sup> and Jie Wan<sup>4\*</sup>

<sup>1</sup>Heilongjiang Provincial Key Laboratory of Electronic Commerce and Information Processing, School of Computer and Information Engineering, Harbin University of Commerce, Harbin, China, <sup>2</sup>College of Pharmacy, Harbin University of Commerce, Harbin, China, <sup>3</sup>Beidahuang Industry Group General Hospital, Harbin, China, <sup>4</sup>Laboratory for Space Environment and Physical Sciences, Harbin Institute of Technology, Harbin, China

## OPEN ACCESS

### Edited by:

Xiujuan Lei,  
Shaanxi Normal University, China

### Reviewed by:

Leyi Wei,  
Shandong University, China  
Lei Deng,  
Central South University, China

### \*Correspondence:

Ke Han  
thruster@163.com  
Jie Wan  
wanjie@hit.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Experimental Pharmacology and Drug  
Discovery,  
a section of the journal  
Frontiers in Pharmacology

**Received:** 14 November 2021

**Accepted:** 20 December 2021

**Published:** 28 January 2022

### Citation:

Han K, Cao P, Wang Y, Xie F, Ma J,  
Yu M, Wang J, Xu Y, Zhang Y and  
Wan J (2022) A Review of Approaches  
for Predicting Drug–Drug Interactions  
Based on Machine Learning.  
*Front. Pharmacol.* 12:814858.  
doi: 10.3389/fphar.2021.814858

Drug–drug interactions play a vital role in drug research. However, they may also cause adverse reactions in patients, with serious consequences. Manual detection of drug–drug interactions is time-consuming and expensive, so it is urgent to use computer methods to solve the problem. There are two ways for computers to identify drug interactions: one is to identify known drug interactions, and the other is to predict unknown drug interactions. In this paper, we review the research progress of machine learning in predicting unknown drug interactions. Among these methods, the literature-based method is special because it combines the extraction method of DDI and the prediction method of DDI. We first introduce the common databases, then briefly describe each method, and summarize the advantages and disadvantages of some prediction models. Finally, we discuss the challenges and prospects of machine learning methods in predicting drug interactions. This review aims to provide useful guidance for interested researchers to further promote bioinformatics algorithms to predict DDI.

**Keywords:** machine learning, drug–drug interactions, similarity, network diffusion, prediction

## INTRODUCTION

Drug–drug interactions (DDI) can occur when two or more drugs are used in combination (Baxter and Preston, 2010). Such interactions may enhance or weaken the efficacy of drugs, cause adverse drug reactions (ADRs) that can even be life-threatening in severe cases (Classen et al., 1997; Agarwal et al., 2020), and cause a drug to be withdrawn from the market (Lazarou et al., 1998). According to the U.S. Centers for Disease Control and Prevention, more than 10% of people take five or more drugs at the same time. Even worse, 20% of older adults take at least 10 drugs (Hohl et al., 2001), which greatly increases the risk of ADR. With an increasing number of approved drugs, the possibility for interactions between drugs increases accordingly (Khorri et al., 2011). Therefore, predicting DDI in advance is both urgent and increasingly difficult in clinical practice.

*In vivo* and *in vitro* experiments can facilitate the identification of DDI, but cannot be performed in some cases due to laboratory limitations and/or high cost (Safdari et al., 2016). Thus, it is particularly important to develop computational methods to solve problems of identifying DDI. Current computational approaches to identify DDI can be divided into two categories: 1) extraction of DDI from literature, electronic medical records, and spontaneous reports; 2) use of known DDI to predict unknown DDI.

**TABLE 1** | DDI prediction methods based on machine learning.

Category	Method	Description
Traditional similarity	Vilar et al. (2012), Rus-Rao Zhang et al. (2009), Gottlieb et al. (2012)	Drug A and drug B interact to produce A specific effect, and it is likely that A drug similar to drug A (or drug B) interacts with drug B (or drug A) to produce the same effect
Traditional classification	Li et al. (2015), Jian-Yu et al. (2016), Kastrin et al. (2018)	The prediction task is simulated as a binary classification problem. Drugs interaction and non-interaction pairs were used to construct classification models
Network diffusion	Link prediction PPIN Cami et al. (2013), Yan et al. (2019), Zhang et al. (2015), Park et al. (2015), Sridhar et al. (2016)	Using drugs as nodes, and their extensive connections and interactions as edges, to predict unknown interactions. Lable propagation, recursive least squares (RSL), traversal of graph and other methods are also used for link prediction
	Graph embedding Decagon Wu et al., 2020, Feng et al. (2020), Ma et al. (2018), Liu et al. (2019), DeepDDI Ryu et al. (2018), Lee et al. (2019), Hou et al. (2019), Yifan et al. (2020)	Transform the graph into a low-dimensional space in which the information about the structure diagram is preserved. Automatically learn node representation in low dimensional space for prediction
Matrix factorization	IPF Vilar et al. (2013), MRMF Zhang et al. (2018), Shtar et al. (2019), ISCMF Rohani et al. (2020), DDINMF Yu et al. (2018), TMFUF Shi et al. (2018)	Matrix factorization decomposes the known DDI matrix into several potential matrices constrained by collective similarity, and then reconstructs the potential matrix to obtain a new interaction matrix
Ensemble-based approach	MLKNN Zhang et al. (2017)	Combine multiple methods to predict unknown DDI.
Based on literature	Tari et al. (2010), Tatonetti et al. (2012b), Kolchinsky et al. (2013)	Firstly, statistical or text mining methods are used to extract the reasonable relationship between drugs from unstructured data sources, and then machine learning methods are used to predict the unknown drug-drug interaction from the extracted drug-drug interaction information

## Extraction of DDI

A large number of DDI are contained in unstructured articles, but with the explosion of biomedical literature, it has become a huge challenge to identify useful information from the vast literature and synchronize it within drug databases (Rodríguez-Terol et al., 2009; Pathak et al., 2013). Extraction of DDI is achieved by one of two approaches: pattern-based approaches and characteristics-based machine learning. The current pattern-based approach is being phased out because it relies on domain knowledge to manually classify DDI. With the emergence of annotated corpus (Segura et al., 2013), the method of extracting DDI based on machine learning becomes more and more popular. Moreover, extracting DDI from unstructured text data does not provide an early warning or identify unknown DDI, while machine learning can effectively predict it in advance (Kanehisa et al., 2010; Chen et al., 2019; Song et al., 2021).

## Prediction of DDI

Only known DDI can be extracted from unstructured articles. However, if the relevant DDI can be predicted in advance before a drug is put onto the market, drugs that cannot be used in combination can be identified. These identified DDI's can prevent many medical errors. We first divide machine learning into traditional and non-traditional categories. In traditional machine learning methods, it is divided into similarity-based method and classification-based method. There are four broad categories of non-traditional machine learning. 1) Network propagation-based approach. The network propagation-based approach can be divided into link prediction and graph embedding according to different methods of network (graph) processing. The link prediction method takes biomedical entities

as nodes and their complex interactions as edges to predict unknown relationship interactions and identify false or missing interactions. The method of graph embedding is to transform the known network (graph) into a low-dimensional space through the embedding layer and retain the information of the network (graph). 2) Matrix factorization. The matrix factorization method is to decompose the known drug interaction matrix into N low-dimensional space matrices using different decomposition methods, and then recombine them to obtain the matrix predicting drug interaction. 3) Ensemble-based approaches. The ensemble-based approaches which combine various methods for predicting drug interactions with the goal of achieving better results. 4) Literature-based methods. This approach first uses NLP to extract drug interactions from unstructured data as data sets. The extracted data are then used to predict unknown drug interactions. In the following article, the first part will introduce the database frequently used in the experiment in detail. The second part introduces several methods for predicting DDI. As shown in **Table 1**. The third part summarizes the article and gives their own views.

## DATASETS

Predicting DDI requires the use of multiple characteristics of drugs and known DDI. The most commonly used databases are: DrugBank (Knox et al., 2010), SIDER (Kuhn et al., 2016), TWOSIDES (Tatonetti et al., 2012a), Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2017). Certain databases were common in the literature-based approach, such as

**TABLE 2** | Common database used to predict drug interactions.

Database	Entities	URL	Brief description
DrugBank	Drugs, Targets, Proteins	<a href="http://www.drugbank.ca/">http://www.drugbank.ca/</a>	Contains a lot of drug information and protein or drug target information
SIDER	Drugs, ADRs	<a href="http://sideeffects.embl.de/">http://sideeffects.embl.de/</a>	Adverse drug reactions of large drugs
TWOSIDES	Drugs, ADRs	<a href="http://tatonettilab.org/resources/tatonetti-stm.html">http://tatonettilab.org/resources/tatonetti-stm.html</a>	Adverse drug reactions of large drugs
PubChem	Structure	<a href="https://pubchem.ncbi.nlm.nih.gov/">https://pubchem.ncbi.nlm.nih.gov/</a>	An open repository of chemical structures and their biological test results
KEGG	DDIs, Proteins	<a href="http://www.kegg.jp/">http://www.kegg.jp/</a>	Metabolic pathways of hyperlinks between metabolites and protein/enzyme information
Medline	Abstract	<a href="https://en.wikipedia.org/wiki/MEDLINE/">https://en.wikipedia.org/wiki/MEDLINE/</a>	Contains abstracts of several biomedical articles

MedlinePlus and PubChem (Kim et al., 2019). The database is described in **Table 2**. DrugBank contains more than 4,100 drug entries, more than 800 FDA-approved small molecule and biotech drugs, and more than 3,200 experimental drugs. In addition, more than 14,000 protein or drug target sequences were associated with these drug entries. Each drug entry contains more than 80 data fields, with half of the information dedicated to drug/chemical data and the other half dedicated to drug target or protein data. SIDER newly released version of SIDER 4 incorporates data about drugs, targets and side effects into a more complete picture of drug mechanisms of action and how they cause adverse reactions. Included 1,430 drugs, 5,880 ADRs. PubChem is an open repository of chemical structures and their biological test results. Contains 247.3 million substance descriptions, 96.5 million unique chemical structures, provided by 629 data sources from 40 countries. It also contains 237 million bioactivity test results from 1.25 million bioassays covering more than 10,000 target protein sequences. A valid database of KEGG protein pathway information. The database is used to capture drug pathways by mapping drug targets. There are currently 75 protein pathway maps. KEGG drug database has 10,979 pieces of drug related information and 501,689 pieces of DDIs relationship. TWOSIDES is a database of drug-drug interactions with side effects. The database contains 868,221 significant associations between 59,220 drug pairs and 1,301 adverse events. MedLinePlus contains 233 abstracts of biomedical articles.

## MACHINE LEARNING-BASED APPROACH

### Similarity-Based Approach

The basic concept of traditional similarity-based approaches for prediction of DDI is as follows: if drug A and drug B interact with each other to produce a specific effect, then drugs like drug A (or drug B) are likely to produce the same effect with drug B (or drug A). With regard to drug similarity, interactions between new drugs are predicted through the fusion of similar characteristics of multiple drugs (Su et al., 2019a; Zeng et al., 2019; Zeng et al., 2020a; Fu et al., 2020; Mo et al., 2020; Zhuang et al., 2020; Shaker et al., 2021).

Vilar et al. (2012) proposed a large-scale approach based on identification of molecular similarities to analyze interactions of

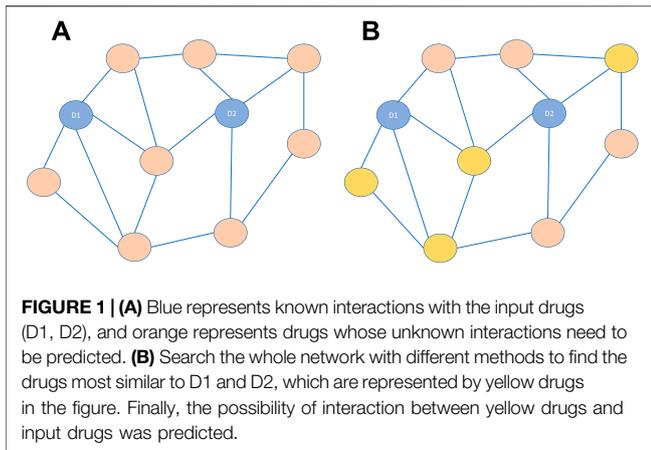
multiple types of drugs caused by inhibition of metabolic enzymes, transporters, and even pharmacological targets. To obtain molecular similarity, the authors first collected and processed drug molecules, then represented the resulting molecular structure as a bit vector that encoded the presence or absence of molecular features, where each feature was assigned a specific location. Finally, the calculation and data representation of similarity measurement are presented. Tanimoto coefficient (TC) was used to measure molecular fingerprints. 0 indicates the greatest dissimilarity, and 1 indicates the greatest similarity.

Ferdousi et al. (2017) calculated the similarity of drug pairs using the Rus-Rao approach based on similarity measurements of 12 binary vectors. The greater the similarity, the greater the likelihood of drugs interactions. Pharmacokinetic DDI (Zhang et al., 2009) describes the process by a drug affects the absorption, distribution, metabolism, or excretion of another drug; whereas, pharmacodynamic DDI (Imming et al., 2006) involves the process by which two or more drugs affect the same receptor to cause synergistic or harmful effects. Gottlieb et al. (2012) used a logistic classifier to infer interactions between pharmacodynamics and pharmacokinetics, as well as their severity, by integrating the similarity measurements of seven different drugs and building classification characteristics.

### Classification-Based Approach

The traditional classification-based approach involves simulating the DDI prediction task as a binary classification problem. DDI pairs and non-DDI pairs are used to build classification models. For binary classification, known interactions are used as inputs, and other drug pairs may have undetected or unobserved interactions that need to be predicted. In machine learning, similar problems are generally converted to semi-supervised learning tasks (Zhao et al., 2020; Hu et al., 2021a). In the classification task, a model is often built using classifiers such as logistic regression, Bayesian, k-nearest neighbor, random forest, and support vector machines (SVM) to predict DDI.

Li et al. (2015) designed a probability ensemble approach employing a Bayesian network model and similarity algorithm to predict drug pairs from molecular and pharmacological characteristics. Jian-Yu et al. (2016) proposed a new semi-supervised fusion algorithm based on a local classification model and Dempster–Shafer evidence theory. With this



approach, new DDI may be predicted based on structural and side-effect similarity (Zhao et al., 2019). Kastrin et al. (2018) treated the process of predicting DDI as a binary classification task by predicting unknown interactions of randomly selected drugs in five large DDI databases using a link-prediction technique, and enhanced the network topology characteristics using four semantic characteristics.

Similarity based on the traditional method and based on the traditional classification method are obtained to predict the unknown drug interactions between very good results, but in these methods, the characteristics of drugs and drug interactions cannot get a good integration between known, there would be no way to use the known information to fully predict drug interactions. So, we need to develop more efficient computational methods to predict unknown drug interactions.

## Network Propagation-Based Approach

The network propagation-based approach predicts unknown DDI using a network of drug structural information or a network formed on the basis of known DDI (Lotfi Shahreza et al., 2018). In biomedical research, it is common to use information available in the network (figure) to predict unknown information or interactions, such as drug–disease associations, DDI, and protein–protein interactions. To solve these problems, link-prediction and graph-embedding approaches are generally used as detailed below.

### Link Prediction

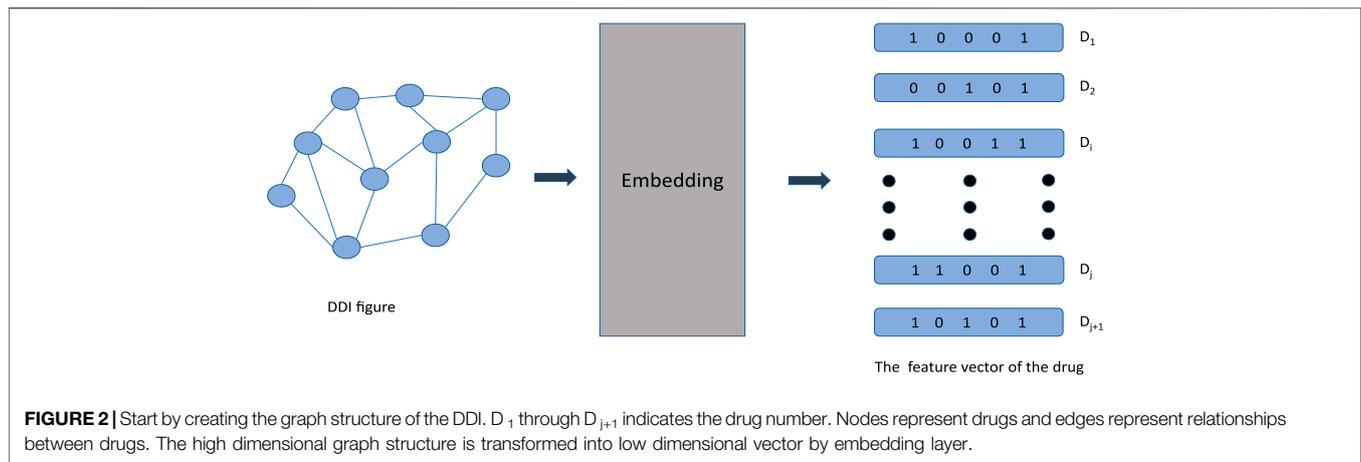
Link prediction uses biomedical entities as nodes, and their extensive connections and interactions as edges, to predict unknown interactions and identify false or missing interactions. In **Figure 1**. Initially, link-prediction approaches assessed the similarity between nodes based on local topological characteristics (Zhou et al., 2009). However, random walk algorithms have become more frequently used for link prediction thanks to the development of global network topology. In addition, many approaches have been employed for link prediction, such as label propagation (Zhang P et al., 2015), probability soft logic (PSL) model (Bach et al., 2015), and graph traversal algorithms (Hu et al., 2020; Hu et al., 2021b).

The similarity-based approaches mentioned above often ignore the structural information encoded in drug biological networks and their interactions. Both the network similarity-based approach and memory network reasoning approach can solve this problem. Specifically, the network similarity-based approach is presented with graphs or reasoned directly through graph structures. Yan et al. (2019) proposed a binary vector-based approach incorporating drug chemistry, biological, and phenotypic data. Briefly, comprehensive drug characteristic similarity was calculated by the isotope similarity approach, a node-based drug network diffusion approach was used to calculate the initial score of the relationship between new drugs, and new DDI were deduced using the recursive least squares (RLS) algorithm.

Cami et al. (2013) described that the DDI of unknown drugs may be predicted by building a network of known DDI. In such a network, drugs are expressed as nodes, while known interactions between known drugs are expressed as edges to predict unknown edges. The first is to integrate data from a number of different sources, including safety data, taxonomic data and data related to the intrinsic characteristics of the drug. The authors construct a network representation of all DDI, where each node represents a drug, and each connecting two nodes represents a known interaction between the two drugs. Next, the binary variables representing the presence or absence of interactions between drug I and drug J were modeled as Bernoulli random variables and three types of covariate functions. By fitting all possible univariate logistic regression models, the authors began to develop models to assess the univariate effect and significance of each covariable.

Most approaches use only first-order similarity, while label propagation considers higher-order similarity. Zhang W et al. (2015) proposed prediction of DDI by integrating clinical side effects extracted from prescription drug packages and United States Food and Drug Administration (FDA) Adverse Event Reporting System, as well as chemical structures extracted from PubChem. They also proposed a framework for comprehensive label propagation that considers similarities of higher-order interactions. First, the authors used the Jaccard index to calculate the similarity between all fingerprints. The authors then create a matrix so that the rows and columns represent the drug, and each cell represents the TC between the fingerprint and the drug. Therefore, drug information from chemical structure, prescription package inserts, and FAERS is transferred to a matrix of chemical similarity, label side effect similarity, and off-label side effect similarity. The authors then use a tag propagation algorithm. The label propagation algorithm solves the following problem: given a weighted network without direction, estimate the labels of the remaining unlabeled nodes. The author takes different drugs as nodes on the network and calculates the edge weights on the network using the drug similarity evaluated by the method in the last section. For each drug, all other drugs in the labeling network are positive, and unlabeled drugs will have DDI with this drug if they know there is a DDI associated with this drug, and use the label to spread the likelihood of this drug network estimate.

The random walk algorithm provides a powerful function to randomly select objects from adjacent nodes in the network, and



then repeats this process constantly to capture information in the network. Park et al. (2015) conceived the method of predicting pharmacokinetic DDI by comparing signal propagation and protein–protein interactions in the network. To achieve this, they applied random walk and restart algorithms to simulate signal propagation from drug targets and capture the possibility of distant interference. The probability of each protein was calculated by the random walk with restart (RWR) algorithm. Protein probabilities are used to represent the effect of drug targets on the protein–protein interaction (PPI) network. The RWR algorithm simulates random walkers until the probability of all proteins on the PPI network is saturated. Next, the authors calculated the protein fraction, which represents the overlapping effects of the two drugs on the same protein. In addition, DDIScore is calculated by summing up the protein fractions of all proteins. The authors used DDIScore as a measure of the likelihood of DDI occurring between drugs.

The advantages of PSL are highly extensible and easy to extend (Bach et al., 2015). Sridhar et al. (2016) inferred DDI from the similarities of multiple drugs and networks of their known interactions, using the joint probability under the PSL framework. Importantly, this approach can be easily extended by different informants and similarities for a variety of applications.

### Graph-Embedding Approach

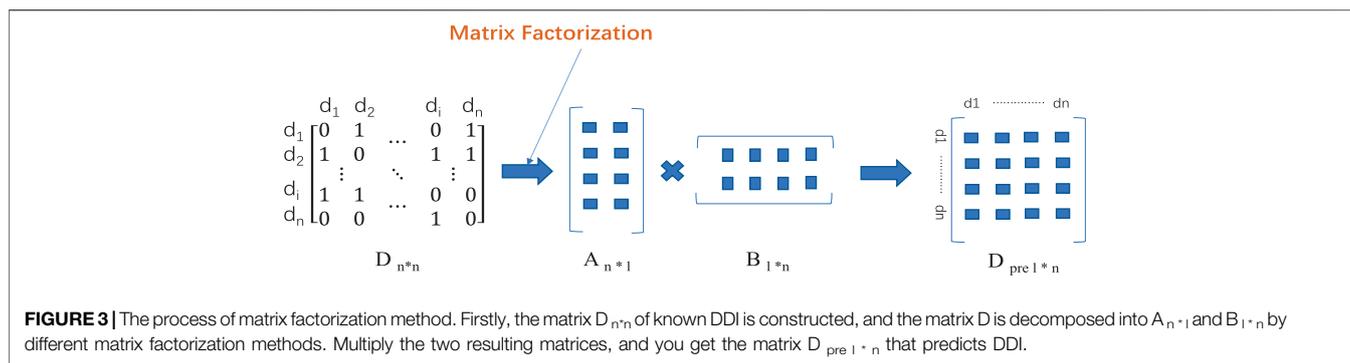
The purpose of graph-embedding approaches is to transform a graph into a low-dimensional space in which structural information about the diagram is preserved. Nodes in the low-dimensional space of automatic learning indicate that prediction has been conducted. Notably, in the field of biology graph-embedding approaches have proven to be more effective in the field of biology than traditional approaches (Tang et al., 2015). In **Figure 2**. Below, graph convolution network (GCN), automated encoders and deep neural networks (DNN) used to predict DDI are described.

The GCN summarizes all the characteristic vectors of adjacent nodes in the graph by considering the characteristics of data structure extraction, and then forms a summary in the spatial domain (Wu et al., 2020). Marinka et al. (2018) built a GCN

architecture (Decagon) for predicting DDI. Multi-model diagrams were constructed using protein–protein interactions, drug–protein interactions, and drug–drug interactions. By using Decagon to predict multi-relation link on multi-model graph, we can not only identify whether any two drugs interact with each other in the graph, but also determine the type of interaction. Ma et al. (2018) considered each type of drug characteristic as a view, calculated the similarity of each view, and then used a multi-view graphical automatic encoder to integrate drug similarity. Subsequently, an attention mechanism was employed to select the view, therefore improving the explanatory nature of the experiment. For modeling, each drug was a node in the drug association network, which was extended by the GCN to embed the characteristics and edges of the multi-view node.

Feng et al. (2020) Combination of GCN and DNN models has also been used to extract structural drugs from the DDI network, in order to predict DDI. DDI predictions can be solved in a three-step approach. First, the potential eigenvectors of each drug were obtained through the function F1. Then the potential vectors of the two drugs are aggregated into an eigenvector to represent the drug pair. Finally, the network is reconstructed through F2. The F1 function is called a feature extractor and the F2 function is called a predictor. Firstly, the GCN model was established to extract the latent features of low dimensional embedding of drugs from DDI network. The latent eigenvectors of the drug are then aggregated to represent drug pairs. Finally, the fused eigenvectors are fed into DNN to predict DDI.

An automatic encoder is an unsupervised neural network whose input and output errors can be minimized through an encoder and decoder. Liu et al. (2019) proposed a multi-pattern, deep autoencoder, drug expression-learning approach based on DDI prediction, which can simultaneously learn the uniform expression of the drug from its functional network. The authors integrate all four characteristics of the drug (chemical substructure, target, enzyme and pathway) to learn about drug characterization and then develop an effective model to predict drug interactions. Consider each drug data set as a view of the drug signature network, so there are five drug signature networks. We use deep neural networks to deal with representational learning. In the DDI-MDAE model, each drug signature



network is trained through a deep automatic coding channel and shares the representation of the drug in the hidden layer for simultaneous learning. A unified sharing representation captures the interrelationships between different networks. On this basis, the author employs four operators to represent drug pairs and train a random forest classifier to predict potential drug–drug interactions.

Deep learning can automatically extract the characteristics of drugs from a dataset and conduct autonomous learning through a multi-layer network to predict unknown DDI. As an artificial neural network with multiple processing layers, DNN can be used to learn highly abstract expressions (Cheng et al., 2018; Su et al., 2019b; Wang et al., 2020a; Cai et al., 2020; Jia et al., 2020; Li et al., 2020; Zhu et al., 2020; Cai et al., 2021; Jin et al., 2021; Liu J et al., 2021; Liu Q et al., 2021; Su et al., 2021; Zhao et al., 2021). To predict unknown DDI, DNN-based approaches often build a framework using a DNN generated from a variety of drug data. Ryu et al. (2018) proposed that a variety of DDI can be predicted by generating a structural similarity profile of drugs that can be used as the “DeepDDI” for the prediction characteristic vector. SMILES were used to generate a feature vector called structural similarity profile (SSP) for each drug in the drug pair. The SSP is designed to effectively capture the unique structural characteristics of a given drug and correlate that characteristic with a set of reported DDI types. In order to predict the DDI type of a given drug pair, two SSPs were generated for each drug pair and combined into a single vector after dimensionality reduction. The combined SSP was the feature vector of the drug pair. A combined SSP of all DDI in a DDI dataset was created and the entire set was used to develop a DNN for accurately predicting DDI types. The author uses cross entropy as loss function and Adam optimization method to train DNN by minimizing prediction error.

Lee et al. (2019) proposed the use of three automatic encoders and a deep feed-forward network to predict DDI. Structural similarity profiles (SSP), target gene similarity profiles (TSP) and Gene Ontology (GO) term similarity profiles (GSP) were measured by autoencoder for dimension reduction (Wang et al., 2020b; Zeng et al., 2020b; Wang Y et al., 2020). The three autoencoders are all isomorphic, the size of input layer and output layer are 3194 and 600 respectively, and the size of hidden layer are 1000, 200 and 1000 respectively. The deep feedforward network has an input layer of size 600, six hidden

layers of size 2000 and an output layer of size 106. Batch size is 256, the learning rate of autoencoder is 0.001, and the learning rate of feedforward network is 0.0001. The activation functions of autoencoder and feedforward network are sigmoid and ReLU. Sigmoid is used as the activation function of the output layer of the feedforward network. The number of epochs was 850, using Adam as the optimizer for the feedforward network and RMSprop as the Autoencoder optimizer. To avoid overfitting, The author uses apply dropout of 0.3 and batch normalization to feedforward networks and Autoencoders.

Hou et al. (2019) proposed the use of DNN to predict DDI, with drugs expressed as a characteristic generated by the SMILE code and entered into a DNN. Yifan et al. (2020) proposed a DDI multimodal deep-learning framework that predicts DDI event types by combining chemical substructures, targets, enzymes, and pathways with deep learning; four drug characterization vectors were calculated and put into the DNN network for training.

### Matrix Factorization-Based Approach

Matrix factoring provides a mathematical basis for various modeling of the biological information problem (Martin et al., 2016). The matrix-factoring approach breaks down the DDI matrix into several matrices, extracts potential characteristics therefrom, and rebuilds the matrix to identify new DDI. In **Figure 3**. Traditional matrix-factoring approaches involve single-value decomposition (SVD) (Sarwar et al., 2000), non-negative matrix factoring (NMF) (Lee and Seung, 1999), and probability matrix factorization (PMF) (Mnih and Salakhutdinov, 2008). However, newly developed matrix decomposition models based on neural networks have been improved in terms of DDI performance. Although the matrix factorization method can achieve good results in predicting drug interactions, the explicability of matrix factorization method is poor and further research is needed. In the matrix factorization method, attention should also be paid to how to better integrate the characteristics of drugs as constraints of matrix decomposition.

Vilar et al. (2013) developed a new approach based on drug interaction profile fingerprinting (IPF). The IPF matrix is used to measure the similarity of drugs, and the interactive probability matrix is calculated by multiplying the DDI matrix by the IPF matrix. New DDI are predicted from the resulting interactive probability matrix. The author first generated an established

database of drug interactions (matrix M1). The collected DDI set was converted into a binary matrix of 928\*928, with a value of 1 indicating the interaction between the two drugs and a value of 0 indicating no interaction. Then, the similarity matrix M2 of interactive profile is generated, which is divided into three steps: the first step is IPFs calculation, the second step is fingerprint similarity calculation, and the last step is M2 construction matrix. The matrix makes rows and columns represent drugs, and each number represents contour similarity based on TC interactions between corresponding drug pairs. Finally, you can predict the new DDIs (matrix M3). Multiply matrix M1 by matrix M2 to obtain M3, and then generate a new set of predicted DDI from M3, and capture the biological effects provided by the original DDI source in M1 and associate them with the new DDI.

Rohani et al. (2020) proposed an integrated similarity constraint matrix factorization (ISCMF) to predict DDI. It can be divided into two steps: first, the integrated similarity matrix is generated as the constraint matrix of matrix decomposition, and then the matrix decomposing the matrices of known drug interactions to obtain two matrices containing potential similarity. Multiply these two matrices to get the matrix that predicts DDI. In the generation of integration similarity, eight similarity matrices were generated based on eight drug characteristics, and k optimal subsets were selected by low entropy and redundancy. Finally, Similarity matrix of K subsets was integrated by similarity network fusion (SNF) method. Based on the known drug interaction matrix, the author decomposed this matrix as the input, integrated similarity as the constraint conditions of decomposition, and finally obtained two decomposition matrices. If you multiply these two matrices, you get the matrix that predicts DDI. Because in decomposition, the loss function has multiple regularization and similarity constraints. Therefore, it can be assumed that the new interaction is a combination of known interactions and similarity matrices.

Zhang et al. (2018) considered multiple characteristics of drugs to calculate similarity matrices. Next, they assumed that the matrices were multiples, projected the drugs onto the low-dimensional space of the interactive space, and introduced multiple normalizations. This group also proposed multi relational matrix factorization of DDI prediction, under which potential DDI are predicted by introducing multi relational matrix factorization of drug characteristics into the matrix factorization. Shtar et al. (2019) proposed prediction of DDI using adjacent matrix factorization (AMF) and adjacent matrix factorization propagation (AMFP). The authors used only known DDI as inputs to predict unknown DDI. AMF breaks down the matrix into an adjacent matrix of DDI, while AMFP (an extension of AMF) spreads potential factors from each drug to interacting drugs on an AMF basis.

Traditional approaches for predicting DDI can only predict their probability, not increases or decreases of drug efficacy during interaction. Yu et al. (2018) proposed a DDI-non-negative matrix factorization (DDINMF) approach to predict conventional and synthetic DDI based on semi-non-negative matrix factors. This approach can predict not only DDI, but

also whether such interactions will enhance or decrease drug efficacy. DDINMF consists of a training stage and prediction stage. The authors expressed interactions in the DDI dataset as a symmetric interaction matrix, which was divided into basic and potential matrices using NMF. Shi et al. (2018) designed the triple matrix factorization (TMF)-based unified framework approach, which uses TMF to connect the adjacent matrix of the DDI network with the characteristic matrix of the drugs.

## Ensemble-Based Approach

The ensemble-based approach combines multiple approaches to predict unknown DDI. Zhang P et al. (2015) proposed that the computational burden of multi-label cases may be reduced by selecting appropriate information dimensions based on the mutual characteristics and side effects of drugs. Combined use of genetic algorithms and the multi-label k-nearest neighbor algorithm can define the optimal characteristic size and enables development of prediction models. A novel multi-label K-nearest adjacency method based on function selection (FS-MLKNN) is proposed, which can simultaneously determine key feature sizes and construct high-precision multi-label prediction models. FS-MLKNN takes two steps to establish the relationship between characteristic vectors and side effects. Firstly, information dimensions are selected by mutual information between functional dimensions and side effects to reduce the computational burden of multi-label learning. Then, genetic algorithm (GA) and multi-label K-nearest neighbor point method (MLKNN) were combined to determine the optimal feature size and develop a prediction model.

Zhang et al. (2017) built a prediction model based on various characteristics of drugs and known data about DDI according to neighbor-recommendation, random walk, and matrix disturbance approaches, which use flexible and diverse frameworks to combine different models with different ensemble rules. Deepika and Geetha (2018) predicted DDI through positive-unlabeled (PU) learning (Elkan and Keith, 2008) and meta-learning (Lemke et al., 2015), and proposed a learning framework for semi-supervised classifiers based on SVM. The PU-based classifier was used to generate meta-knowledge from the network, and the meta-classifier was designed to predict the probability of DDI from the generated meta-knowledge.

## Literature-based Approach

Literature-based prediction of DDI consists of two steps: first, extraction of the reasonable relationship between drugs from unstructured data sources (Vilar et al., 2018) (literature, electronic medical records, spontaneous reports, etc.) with a statistical or text-mining method, followed by use of natural language processing technique; second, prediction of unknown DDI from extracted information about the interactions between drugs using machine learning.

Tari et al. (2010) predicted DDI by combining text mining and reasoning. The process involved two stages: natural language extraction and reasoning. The authors used a parsing tree to extract various interactions and applied logical rules to predict interactions based on extracted interactions between new and

existing drugs. Tatonetti et al. (2012b) divided FAERS into two sets of data: reports involving only one drug and reports involving two drugs, and constructed eight “clinically major” adverse event models. In each model, the drug information described was an introduction to the frequency of adverse events extracted from FAERS, and a logistic regression classifier was used to distinguish drugs that caused major clinical adverse events under study from those that did not; prediction was conducted based on the drug combination for each model. Kolchinsky et al. (2013) evaluated the performance of several classifiers, such as logistics regression, SVM, and discriminatory analysis, to distinguish relevant abstracts and PubMed articles containing evidence for pharmacodynamic DDI. Notably, their approach is also helpful to link causal mechanisms to potential DDI.

## CONCLUSION

The occurrence of DDI affects the treatment of patients and has become a serious problem for patient safety and drug management. The harm caused by DDI will be greatly reduced if machine learning can be used to efficiently predict DDI. To this end, it is urgent to develop better-performing machine learning approaches. This article describes existing machine learning-based approaches for predicting DDI. In the past 10 years, machine learning has been widely applied in bioinformatics

## REFERENCES

- Agarwal, S., Agarwal, V., Agarwal, M., and Singh, M. (2020). Exosomes: Structure, Biogenesis, Types and Application in Diagnosis and Gene and Drug Delivery. *Curr. Gene Ther.* 20, 195–206. doi:10.2174/1566523220999200731011702
- Bach, S. H., Broecheler, M., Huang, B., and Getoor, L. (2015). *Hinge-loss Markov Random fields and Probabilistic Soft Logic*. arXiv preprint arXiv:1505.04406.
- Baxter, K., and Preston, C. L. (2010). *Stockley's Drug Interactions*. Pharmaceutical Press London.
- Cai, L., Lu, C., Xu, J., Meng, Y., Wang, P., Fu, X., et al. (2021). Drug Repositioning Based on the Heterogeneous Information Fusion Graph Convolutional Network. *Brief. Bioinformatics* 22 (6), bbab319. doi:10.1093/bib/bbab319
- Cai, L., Wang, L., Fu, X., Xia, C., Zeng, X., and Zou, Q. (2020). ITP-pred: an Interpretable Method for Predicting, Therapeutic Peptides with Fused Features Low-Dimension Representation. *Brief. Bioinformatics* 22 (4), bbaa367. doi:10.1093/bib/bbaa367
- Cami, A., Manzi, S., Arnold, A., and Reis, B. Y. (2013). Pharmacointeraction Network Models Predict Unknown Drug-Drug Interactions. *PLoS One* 8, e61468. doi:10.1371/journal.pone.0061468
- Chen, X., Shi, W., and Deng, L. (2019). Prediction of Disease Comorbidity Using HeteSim Scores Based on Multiple Heterogeneous Networks. *Curr. Gene Ther.* 19, 232–241. doi:10.2174/1566523219666190917155959
- Cheng, L., Hu, Y., Sun, J., Zhou, M., and Jiang, Q. (2018). DincRNA: a Comprehensive Web-Based Bioinformatics Toolkit for Exploring Disease Associations and ncRNA Function. *Bioinformatics* 34, 1953–1956. doi:10.1093/bioinformatics/bty002
- Classen, D. C., PestotnikEvans, S. L. R. S., Evans, R. S., Lloyd, J. F., and Burke, J. P. (1997). Adverse Drug Events in Hospitalized Patients. Excess Length of Stay, Extra Costs, and Attributable Mortality. *Jama* 277, 301–306. doi:10.1001/jama.1997.03540280039031
- Deepika, S. S., and Geetha, T. V. (2018). A Meta-Learning Framework Using Representation Learning to Predict Drug-Drug Interaction. *J. Biomed. Inform.* 84, 136–147. doi:10.1016/j.jbi.2018.06.015

and achieved good results. Under most of the existing approaches, drug similarity is taken as the most fundamental starting point for better prediction of DDI, assisted by a variety of other means. However, most current DDI predictions are limited to the interactions between two drugs. In the future work, we should not only pursue the accuracy of predicting the probability of drug-drug interactions, but also pursue the ability to accurately predict the types of drug-drug interactions. However, because the use of multiple drugs has become a trend in clinical medicine, it is urgent to develop methods to predict interactions between multiple drugs. It is our opinion that a number of excellent ways to solve this problem will be available in the near future.

## AUTHOR CONTRIBUTIONS

Conception and design: KH, PC, and JW. Collected literatures: YW, FX, and JM. Figures and Tables: MY and JCW. Writing original draft and editing: YW, YX, YZ, and JW. Grammar and editing: KH and PC.

## FUNDING

The work was supported by the Natural Science Foundation of Heilongjiang Province under Grant LH2020F009.

- Deng, Y., Xu, X., Qiu, Y., Xia, J., Zhang, W., and Liu, S. (2020). A Multimodal Deep Learning Framework for Predicting Drug-Drug Interaction Events. *Bioinformatics* 36, 4316–4322. Oxford, England. doi:10.1093/bioinformatics/btaa501
- Elkan, C., and Keith, N. (2008). Learning Classifiers from Only Positive and Unlabeled Data. *Knowledge Discov. Data mining* 08. doi:10.1145/1401890.1401920
- Feng, Y. H., Zhang, S. W., and Shi, J. Y. (2020). DPDDI: a Deep Predictor for Drug-Drug Interactions. *BMC Bioinformatics* 21, 419–515. doi:10.1186/s12859-020-03724-x
- Ferdousi, R., Safdari, R., and Omid, Y. (2017). Computational Prediction of Drug-Drug Interactions Based on Drugs Functional Similarities. *J. Biomed. Inform.* 70, 54–64. doi:10.1016/j.jbi.2017.04.021
- Fu, X., Cai, L., Zeng, X., and Zou, Q. (2020). StackCPPred: a Stacking and Pairwise Energy Content-Based Prediction of Cell-Penetrating Peptides and Their Uptake Efficiency. *Bioinformatics* 36, 3028–3034. doi:10.1093/bioinformatics/btaa131
- Gottlieb, A., Stein, G. Y., Oron, Y., Rupp, E., and Sharan, R. (2012). INDI: a Computational Framework for Inferring Drug Interactions and Their Associated Recommendations. *Mol. Syst. Biol.* 8, 592. doi:10.1038/msb.2012.26
- Hohl, C. M., Dankoff, J., Colacone, A., and Afilalo, M. (2001). Polypharmacy, Adverse Drug-Related Events, and Potential Adverse Drug Interactions in Elderly Patients Presenting to an Emergency Department. *Ann. Emerg. Med.* 38, 666–671. doi:10.1067/mem.2001.119456
- Hou, X., You, J., and Hu, P. (2019). “Predicting Drug-Drug Interactions Using Deep Neural Network,” in Proceedings of the 2019 11th International Conference on Machine Learning and Computing (ICMLC). 19. doi:10.1145/3318299.3318323
- Hu, Y., Qiu, S., and Cheng, L. (2021b). Integration of Multiple-Omics Data to Analyze the Population-specific Differences for Coronary Artery Disease. *Comput. Math. Methods Med.* 2021, 7036592. doi:10.1155/2021/7036592
- Hu, Y., Zhang, H., Liu, B., Gao, S., Wang, T., Han, Z., et al. (2020). rs34331204 Regulates TSPAN13 Expression and Contributes to Alzheimer's Disease with Sex Differences. *Brain* 143, e95. doi:10.1093/brain/awaa302

- Hu, Y., Sun, J.-y., Zhang, Y., Zhang, H., Gao, S., Wang, T., et al. (2021a). rs1990622 Variant Associates with Alzheimer's Disease and Regulates TMEM106B Expression in Human Brain Tissues. *BMC Med.* 19, 11. doi:10.1186/s12916-020-01883-5
- Imming, P., Sinning, C., and Meyer, A. (2006). Drugs, Their Targets and the Nature and Number of Drug Targets. *Nat. Rev. Drug Discov.* 5, 821–834. doi:10.1038/nrd2132
- Jia, C., Bi, Y., Chen, J., Leier, A., Li, F., and Song, J. (2020). PASSION: an Ensemble Neural Network Approach for Identifying the Binding Sites of RBPs on circRNAs. *Bioinformatics* 36, 4276–4282. doi:10.1093/bioinformatics/btaa522
- Jian-Yu, S. G. K., Shang, X. Q., and Siu-Ming, Y. (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), 512–515. doi:10.1109/bibm.2016.7822571
- Jin, Q., Cui, H., Sun, C., Meng, Z., and Su, R. (2021). Free-form Tumor Synthesis in Computed Tomography Images via Richer Generative Adversarial Network. *Knowledge-Based Syst.* 218, 106753. doi:10.1016/j.knsys.2021.106753
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* 45, D353–D61. doi:10.1093/nar/gkw1092
- 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, D355–D360. doi:10.1093/nar/gkp896
- Kastrin, A., Ferik, P., and Leskošek, B. (2018). Predicting Potential Drug-Drug Interactions on Topological and Semantic Similarity Features Using Statistical Learning. *PLoS One* 13, e0196865. doi:10.1371/journal.pone.0196865
- Khori, V., Semnani, S. H., and Roshandel, G. H. (2011). Frequency Distribution of Drug Interactions and Some of Related Factors in Prescriptions. *Med. J. Tabriz Univ. Med. Sci.* 27, 29–32.
- Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2019). PubChem 2019 Update: Improved Access to Chemical Data. *Nucleic Acids Res.* 47, D1102–D09. doi:10.1093/nar/gky1033
- Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., et al. (2010). DrugBank 3.0: a Comprehensive Resource for 'omics' Research on Drugs. *Nucleic Acids Res.* 39, D1035–D1041. doi:10.1093/nar/gkq1126
- Kolchinsky, A., Lourenço, A., Li, L., and Rocha, L. M. (2013). "Evaluation of Linear Classifiers on Articles Containing Pharmacokinetic Evidence of Drug-Drug Interactions," in Pacific Symposium on Biocomputing.
- Kuhn, M., Letunic, I., Jensen, L. J., and Bork, P. (2016). The SIDER Database of Drugs and Side Effects. *Nucleic Acids Res.* 44, D1075–D1079. doi:10.1093/nar/gkv1075
- Lazarou, J., PomeranzPomeranz, B. H., and Corey, P. N. (1998). Incidence of Adverse Drug Reactions in Hospitalized Patients: a Meta-Analysis of Prospective Studies. *Jama* 279, 1200–1205. doi:10.1001/jama.279.15.1200
- Lee, D. D., and Seung, H. S. (1999). Learning the Parts of Objects by Non-negative Matrix Factorization. *Nature* 401, 788–791. doi:10.1038/44565
- Lee, G., Park, C., and Ahn, J. (2019). Novel Deep Learning Model for More Accurate Prediction of Drug-Drug Interaction Effects. *BMC Bioinformatics* 20, 415. doi:10.1186/s12859-019-3013-0
- Lemke, C., Budka, M., and Gabrys, B. (2015). Metalearning: a Survey of Trends and Technologies. *Artif. Intell. Rev.* 44, 117–130. doi:10.1007/s10462-013-9406-y
- Li, F., Chen, J., Leier, A., Marquez-Lago, T., Liu, Q., Wang, Y., et al. (2020). DeepCleave: a Deep Learning Predictor for Caspase and Matrix Metalloprotease Substrates and Cleavage Sites. *Bioinformatics* 36, 1057–1065. doi:10.1093/bioinformatics/btz721
- Li, P., Huang, C., Fu, Y., Wang, J., Wu, Z., Ru, J., et al. (2015). Large-scale Exploration and Analysis of Drug Combinations. *Bioinformatics* 31, 2007–2016. doi:10.1093/bioinformatics/btv080
- Liu, J., Su, R., Zhang, J., and Wei, L. (2021). Classification and Gene Selection of Triple-Negative Breast Cancer Subtype Embedding Gene Connectivity Matrix in Deep Neural Network. *Brief. Bioinform.* 22 (5), bbaa395. doi:10.1093/bib/bbaa395
- Liu, Q., Chen, J., Wang, Y., Li, S., Jia, C., Song, J., et al. (2021). DeepTorrent: a Deep Learning-Based Approach for Predicting DNA N4-Methylcytosine Sites. *Brief Bioinform.* 22. doi:10.1093/bib/bbaa124
- Liu, S., Huang, Z., Qiu, Y., Chen, Y. P., and Zhang, W. (2019). "Structural Network Embedding Using Multi-Modal Deep Auto-Encoders for Predicting Drug-Drug Interactions," in IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, November 18–21, 2019 (IEEE), 445–450.
- Lotfi Shahreza, M., Ghadiri, N., Mousavi, S. R., Varshosaz, J., and Green, J. R. (2018). A Review of Network-Based Approaches to Drug Repositioning. *Brief Bioinform.* 19, 878–892. doi:10.1093/bib/bbx017
- Ma, T., Xiao, C., Zhou, J., and Wang, F. (2018). *Drug Similarity Integration through Attentive Multi-View Graph Auto-Encoders*. arXiv preprint arXiv:1804.10850.
- Marinka, Z., Agrawal, M., and Jure, L. (2018). Modeling Polypharmacy Side Effects with Graph Convolutional Networks. *Bioinformatics* 34, i457–i466. Oxford, England. doi:10.1093/bioinformatics/bty294
- Martin, S., Marinka, Z., Blaž, Z., Jernej, U., and Tomaž, C. (2016). Orthogonal Matrix Factorization Enables Integrative Analysis of Multiple RNA Binding Proteins. *Bioinformatics* 32 (10), 1527–1535. Oxford, England. doi:10.1093/bioinformatics/btw003
- Mnih, A., and Salakhutdinov, R. R. (2008). "Probabilistic Matrix Factorization," in *Advances in Neural Information Processing Systems*, 1257–1264.
- Mo, F., Luo, Y., Fan, D., Zeng, H., Zhao, Y., Luo, M., et al. (2020). Integrated Analysis of mRNA-Seq and miRNA-Seq to Identify C-MYC, YAP1 and miR-3960 as Major Players in the Anticancer Effects of Caffeic Acid Phenethyl Ester in Human Small Cell Lung Cancer Cell Line. *Curr. Gene Ther.* 20, 15–24. doi:10.2174/1566523220666200523165159
- 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, e0140816. doi:10.1371/journal.pone.0140816
- Pathak, J., Kiefer, R. C., and Chute, C. G. (2013). Using Linked Data for Mining Drug-Drug Interactions in Electronic Health Records. *Stud. Health Technol. Inform.* 192, 682–686.
- Rodríguez-Terol, A., Caraballo, M. O., Palma, D., Santos-Ramos, B., Molina, T., Desongles, T., et al. (2009). Quality of Interaction Database Management Systems. *Farmacia Hospitalaria (English Edition)* 33, 134–146.
- Rohani, N., Eslahchi, C., and Ali, K. (2020). 'iscmf: Integrated Similarity-Constrained Matrix Factorization for Drug-Drug Interaction Prediction. *Netw. Model. Anal. Health Inform. Bioinformatics* 9, 1–8. doi:10.1007/s13721-019-0215-3
- Ryu, J. Y., Kim, H. U., and Lee, S. Y. (2018). Deep Learning Improves Prediction of Drug-Drug and Drug-Food Interactions. *Proc. Natl. Acad. Sci. U S A* 115, E4304–E11. doi:10.1073/pnas.1803294115
- Safdari, R., Ferdousi, R., Azizheris, K., Niakan-Kalhari, S. R., and Omid, Y. (2016). Computerized Techniques Pave the Way for Drug-Drug Interaction Prediction and Interpretation. *Bioimpacts* 6, 71–78. doi:10.15171/bi.2016.10
- Sarwar, B., George, K., Joseph, K., and Riedl, J. (2000). *Application of Dimensionality Reduction in Recommender System-A Case Study*. Minnesota Univ Minneapolis Dept of Computer Science.
- Segura, B., IsabelMartinez, P., and Zazo, M. H. (2013). "Semeval-2013 Task 9: Extraction of Drug-Drug Interactions from Biomedical Texts (Ddiextraction 2013)," in *Association for Computational Linguistics*.
- Shaker, B., TranMong, K. M., Jung, C., and Na, D. (2021). Introduction of Advanced Methods for Structure-Based Drug Discovery. *Cbio* 16, 351–363. doi:10.2174/1574893615999200703113200
- Shi, J. Y., Huang, H., Li, J. X., Lei, P., Zhang, Y. N., Dong, K., et al. (2018). TMFUF: a Triple Matrix Factorization-Based Unified Framework for Predicting Comprehensive Drug-Drug Interactions of New Drugs. *BMC Bioinformatics* 19, 411–437. doi:10.1186/s12859-018-2379-8
- Shtar, G., Rokach, L., and Shapira, B. (2019). Detecting Drug-Drug Interactions Using Artificial Neural Networks and Classic Graph Similarity Measures. *PLoS One* 14, e0219796. doi:10.1371/journal.pone.0219796
- Song, B., Li, F., Liu, Y., and Zeng, X. (2021). 'Deep Learning Methods for Biomedical Named Entity Recognition: a Survey and Qualitative Comparison. *Brief. Bioinformatics* 22 (6), bbab282. doi:10.1093/bib/bbab282
- Sridhar, D., Fakhraei, S., and Getoor, L. (2016). A Probabilistic Approach for Collective Similarity-Based Drug-Drug Interaction Prediction. *Bioinformatics* 32, 3175–3182. doi:10.1093/bioinformatics/btw342
- Su, R., Liu, X., Wei, L., and Zou, Q. (2019b). Deep-Resp-Forest: A Deep forest Model to Predict Anti-cancer Drug Response. *Methods* 166, 91–102. doi:10.1016/j.jymeth.2019.02.009

- Su, R., Wu, H., Xu, B., Liu, X., and Wei, L. (2019a). Developing a Multi-Dose Computational Model for Drug-Induced Hepatotoxicity Prediction Based on Toxicogenomics Data. *Ieee/acm Trans. Comput. Biol. Bioinform* 16, 1231–1239. doi:10.1109/TCBB.2018.2858756
- Su, R., Liu, X., Jin, Q., Liu, X., and Wei, L. (2021). Identification of Glioblastoma Molecular Subtype and Prognosis Based on Deep MRI Features. *Knowledge-Based Syst.* 232, 107490. doi:10.1016/j.knsys.2021.107490
- Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., and Mei, Q. (2015). “Line,” in Proceedings of the 24th International Conference on World Wide Web, 1067–1077. doi:10.1145/2736277.2741093
- Tari, L., Anwar, S., Liang, S., Cai, J., and Baral, C. (2010). Discovering Drug-Drug Interactions: a Text-Mining and Reasoning Approach Based on Properties of Drug Metabolism. *Bioinformatics* 26, i547–53. doi:10.1093/bioinformatics/btq382
- Tatonetti, N. P., Fernald, G. H., and Altman, R. B. (2012a). A Novel Signal Detection Algorithm for Identifying Hidden Drug-Drug Interactions in Adverse Event Reports. *J. Am. Med. Inform. Assoc.* 19, 79–85. doi:10.1136/amiajnl-2011-000214
- Tatonetti, N. P., Ye, P. P., Daneshjoui, R., and Altman, R. B. (2012b). Data-driven Prediction of Drug Effects and Interactions. *Sci. Transl. Med.* 4, 125ra31–25ra31. doi:10.1126/scitranslmed.3003377
- Vilar, S., Friedman, C., and Hripcsak, G. (2018). Detection of Drug-Drug Interactions through Data Mining Studies Using Clinical Sources, Scientific Literature and Social media. *Brief Bioinform* 19, 863–877. doi:10.1093/bib/bbx010
- Vilar, S., Harpaz, R., Uriarte, E., Santana, L., Rabadan, R., and Friedman, C. (2012). Drug-drug Interaction through Molecular Structure Similarity Analysis. *J. Am. Med. Inform. Assoc.* 19, 1066–1074. doi:10.1136/amiajnl-2012-000935
- Vilar, S., Uriarte, E., Santana, L., Tatonetti, N. P., and Friedman, C. (2013). Detection of Drug-Drug Interactions by Modeling Interaction Profile Fingerprints. *PLoS One* 8, e58321. doi:10.1371/journal.pone.0058321
- Wang, J., Shi, Y., Wang, X., and Chang, H. (2020b). A Drug Target Interaction Prediction Based on LINE-RF Learning. *Cbio* 15, 750–757. doi:10.2174/1574893615666191227092453
- Wang, J., Wang, H., Wang, X., and Chang, H. (2020a). Predicting Drug-Target Interactions via FM-DNN Learning. *Cbio* 15, 68–76. doi:10.2174/1574893614666190227160538
- Wang, Y., Li, F., Bharathwaj, M., Rosas, N. C., Leier, A., Akutsu, T., et al. (2020). DeepBL: a Deep Learning-Based Approach for In Silico Discovery of Beta-Lactamases. *Brief Bioinform* 22 (4), bbaa301. doi:10.1093/bib/bbaa301
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Yu, P. S. (2020). A Comprehensive Survey on Graph Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 4–24. doi:10.1109/TNNLS.2020.2978386
- Yan, C., Duan, G., Zhang, Y., Wu, F. X., Pan, Y., and Wang, J. (2019). “IDNDDI: An Integrated Drug Similarity Network Method for Predicting Drug-Drug Interactions,” in *Bioinformatics Research and Applications*. doi:10.1007/978-3-030-20242-2\_8
- Yu, H., Mao, K. T., Shi, J. Y., Huang, H., Chen, Z., Dong, K., et al. (2018). Predicting and Understanding Comprehensive Drug-Drug Interactions via Semipositive Matrix Factorization. *BMC Syst. Biol.* 12, 14–10. doi:10.1186/s12918-018-0532-7
- Zeng, X., Song, X., Ma, T., Pan, X., Zhou, Y., Hou, Y., et al. (2020b). Repurpose Open Data to Discover Therapeutics for COVID-19 Using Deep Learning. *J. Proteome Res.* 19, 4624–4636. doi:10.1021/acs.jproteome.0c00316
- Zeng, X., Zhu, S., Liu, X., Zhou, Y., Nussinov, R., and Cheng, F. (2019). deepDR: a Network-Based Deep Learning Approach to In Silico Drug Repositioning. *Bioinformatics* 35, 5191–5198. doi:10.1093/bioinformatics/btz418
- Zeng, X., Zhu, S., Lu, W., Liu, Z., Huang, J., Zhou, Y., et al. (2020a). Target Identification Among Known Drugs by Deep Learning from Heterogeneous Networks. *Chem. Sci.* 11, 1775–1797. doi:10.1039/c9sc04336e
- Zhang, L., Zhang, Y. D., Zhao, P., and Huang, S. M. (2009). Predicting Drug-Drug Interactions: an FDA Perspective. *AAPS J.* 11, 300–306. doi:10.1208/s12248-009-9106-3
- Zhang, P., Wang, F., Hu, J., and Sorrentino, R. (2015). Label Propagation Prediction of Drug-Drug Interactions Based on Clinical Side Effects. *Sci. Rep.* 5, 12339. doi:10.1038/srep12339
- Zhang, W., Chen, Y., Li, D., and Yue, X. (2018). Manifold Regularized Matrix Factorization for Drug-Drug Interaction Prediction. *J. Biomed. Inform.* 88, 90–97. doi:10.1016/j.jbi.2018.11.005
- 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, 18. doi:10.1186/s12859-016-1415-9
- Zhang, W., Liu, F., Luo, L., and Zhang, J. (2015). Predicting Drug Side Effects by Multi-Label Learning and Ensemble Learning. *BMC Bioinformatics* 16, 365. doi:10.1186/s12859-015-0774-y
- Zhao, T., Hu, Y., Peng, J., and Cheng, L. (2020). DeepLGP: a Novel Deep Learning Method for Prioritizing lncRNA Target Genes. *Bioinformatics* 36, 4466–4472. doi:10.1093/bioinformatics/btaa428
- Zhao, T., Hu, Y., and Cheng, L. (2021). Deep-DRM: a Computational Method for Identifying Disease-Related Metabolites Based on Graph Deep Learning Approaches. *Brief Bioinform* 22, bbaa212. doi:10.1093/bib/bbaa212
- Zhao, X., Chen, L., Guo, Z.-H., and Liu, T. (2019). Predicting Drug Side Effects with Compact Integration of Heterogeneous Networks. *Cbio* 14, 709–720. doi:10.2174/1574893614666190220114644
- Zhou, T., Lü, L., and Zhang, Y.-C. (2009). Predicting Missing Links via Local Information. *Eur. Phys. J. B* 71, 623–630. doi:10.1140/epjb/e2009-00335-8
- Zhu, Y., Li, F., Xiang, D., Akutsu, T., Song, J., and Jia, C. 2020. Computational Identification of Eukaryotic Promoters Based on Cascaded Deep Capsule Neural Networks, *Brief Bioinform*, 22. bbaa299doi:doi:10.1093/bib/bbaa299
- Zhuang, J., Dai, S., Zhang, L., Gao, P., Han, Y., Tian, G., et al. (2021). Identifying Breast Cancer-Induced Gene Perturbations and its Application in Guiding Drug Repurposing. *Cbio* 15, 1075–1089. doi:10.2174/1574893615666200203104214

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

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Han, Cao, Wang, Xie, Ma, Yu, Wang, Xu, Zhang and Wan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.