



Enhancing drug–drug interaction prediction by three-way decision and knowledge graph embedding

Xinkun Hao^{1,3} · Qingfeng Chen^{1,2,3} · Haiming Pan^{1,3} · Jie Qiu^{1,3} · Yuxiao Zhang^{1,3} · Qian Yu^{1,3} · Zongzhao Han^{1,3} · Xiaojing Du^{1,3}

Received: 7 December 2021 / Accepted: 15 February 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract

Drug–Drug interaction (DDI) prediction is essential in pharmaceutical research and clinical application. Existing computational methods mainly extract data from multiple resources and treat it as binary classification. However, this cannot unambiguously tell the boundary between positive and negative samples owing to the incompleteness and uncertainty of derived data. A granular computing method called three-way decision is proved to be effective in making uncertain decision, but it relies on supplementary information to make delay decision. Recently, biomedical knowledge graph has been regarded as an important source to obtain abundant supplementary information about drugs. This paper proposes a three-way decision-based method called 3WDDI, in combination with knowledge graph embedding as supplementary features to enhance DDI prediction. The drug pairs are divided into positive, negative and boundary regions by Convolutional Neural Network (CNN) according to drug chemical structure feature. Further, delay decision is made for objects in the boundary region by integrating knowledge graph embedding feature to promote the accuracy of decision-making. The empirical results show that 3WDDI yields up to 0.8922, 0.9614, 0.9582, 0.8930 for Accuracy, AUPR, AUC and F1-score, respectively, and outperforms several baseline models.

Keywords Drug–drug interactions · Three-way decision · Knowledge graph · Convolutional neural network

1 Introduction

Poly-pharmacy is increasingly common in real-world clinical settings, especially for aged and cancer patients who suffer from multiple diseases and need complex drug treatments. However, unexpected drug–drug interactions

(DDI) frequently occur and lead to adverse drug reactions, while several drugs are taken together. Unexpected DDI may threaten the patient's health, and give rise to re-hospitalization. This may directly or indirectly result in high burden on medical system. Therefore, it is of vital importance to identify potential drug–drug interactions before putting them on the market for clinical use.

✉ Qingfeng Chen
qingfeng@gxu.edu.cn

Xinkun Hao
haoxinkun@st.gxu.edu.cn

Haiming Pan
2574958437@qq.com

Jie Qiu
jgxyqj@126.com

Yuxiao Zhang
2636213416@qq.com

Qian Yu
823627043@qq.com

Zongzhao Han
1051360288@qq.com

Xiaojing Du
1779245288@qq.com

¹ School of Computer, Electronics and Information, Guangxi University, Nanning 530004, Guangxi, China

² Department of Computer Science and Information Technology, La Trobe University, Melbourne, VIA 3086, Australia

³ School of Computer Science and Engineering, Yulin Normal University, Yulin 537000, Guangxi, China

Traditional methods for DDI prediction mainly depend on *in vivo* and *in vitro* experiments, which are usually labor-intensive and limited in experimental scale. In recent years, with the development of computer technology and establishment of kinds of drug-related databases, a number of data-driven methods for DDI prediction have been proposed and yielded considerable results. Machine learning and deep learning methods are impressive due to their good performance. They tend to extract multiple drug-related features from open-source databases, such as DrugBank (Law et al. 2013), KEGG-DRUG (Kanehisa et al. 2015). And the derived drug-related features usually consist of drug molecular fingerprint, drug target information, drug phenotype information and so on. However, it is not only labor-intensive to obtain comprehensive drug-related information from diverse data sources, but also difficult to guarantee the quality of drug information.

Recently, knowledge graph has been widely studied and used in various research fields (Ji et al. 2021) for knowledge representation and knowledge integration. And several large-scale domain knowledge graphs have also been established in bioinformatics, such as Bio2RDF (Callahan et al. 2013) and DRKG (Ioannidis et al. 2020). Existing methods mainly utilize well-established knowledge graphs or integrate knowledge from multiple data sources to construct knowledge graphs (Abdelaziz et al. 2017; Karim et al. 2019; Dai et al. 2020). Knowledge graph embedding model (or knowledge representation learning model) is trained on the knowledge graph to generate embedding vectors for entities and relations as external knowledge, which is able to enhance traditional machine learning and deep learning methods for DDI prediction. And many studies have proven knowledge graph embedding vector is a powerful feature.

No matter obtaining drug features from multiple data sources or using knowledge graph embedding as feature ideally assume that we can obtain detailed and complete information about the drug all at once, and then make an immediate decision for the existence of DDI. However, only chemical structure information of a drug is easy to obtain, while others, such as target, pathway, need further experiments to discover. As a result, it is not easy to decide whether there is an interaction between two drugs based on incomplete information. And modeling the problem as a binary classification directly may hinder the accuracy of DDI prediction due to the uncertainty and incompleteness of drug information. In recent years, a granular computing method called three-way decision aiming to solve uncertain decision-making is proposed and widely applied in Artificial Intelligence field, such as image recognition (Huaxiong et al. 2016), recommendation system (Zhang et al. 2017a). The main idea of three-way decision is to divide the objects into three regions instead of two regions in binary

classification, and adopt delay decision for an additional region called boundary region until more information is obtained. To model the uncertain decision boundary and prompt the accuracy of DDI prediction, we thus introduce the three-way decision for DDI prediction. Meanwhile, delay decision of three-way decision tends to rely on a supplementary feature. Herein, we treat powerful knowledge graph embedding vector as the additional feature which contains global and abundant knowledge of drug.

In this paper, we treat the knowledge graph as auxiliary information and combine three-way decision to enhance drug–drug interaction prediction. The chemical structure information of drug is collected to generate drug structure feature for each drug pair in training set. A CNN model with strong feature extraction ability is adopted to divide training data into positive, negative and boundary region based on the drug structure feature. Unlike binary classification, immediate decision is not made for the samples in the boundary region. To achieve more accurate decision on drug–drug interaction in boundary region, knowledge graph embedding is trained to represent a drug pair, and further decision will be made by an enhancing model. Knowledge graph embedding feature shows strong representation ability and is significantly different from drug structure feature. The experimental results demonstrate that delay decision based on knowledge graph embedding feature produce more confidence prediction for drug pairs in the boundary region.

To summarize, this paper makes the following contributions:

- We introduce three-way decision to solve the uncertain decision of DDI prediction and utilize the knowledge graph embedding as supplementary feature to enhance DDI prediction.
- We propose a novel method called 3WDDI to predict the existence of drug–drug interaction.
- We compare our method with several classical models and state-of-the-art works on a popular dataset. The results show the proposed 3WDDI outperforms the baselines.

The remaining content of this paper is organized as follows. In Sect. 2, we introduce the background and related works briefly. In Sect. 3, the proposed method 3WDDI is introduced. And we report the experiment results and analysis in Sect. 4. The conclusions are offered in Sect. 5.

2 Related work

This section introduces the research background and relevant works, including DDI prediction, knowledge graph embedding and three-way decision.

2.1 DDI prediction

In the past few years, many works extract drug-related information from existing databases to construct drug pair representation and train machine learning models to predict drug–drug interactions. They mainly model the drug–drug prediction problem as binary classification. Given a representation of a drug pair, the machine learning models calculate the probability that the drug interaction exists. We roughly classify current methods into two categories: traditional machine learning-based method and deep learning-based method. The former focuses on effective drug pair representation. Gottlieb et al. (2012) constructed drug feature vectors based on seven types of drug–drug similarities to describe drug–drug pairs, and then applies logistic regression model to predict DDI. A heterogeneous network-assisted inference framework (Cheng and Zhao 2014) is proposed to assist the prediction of DDI, which integrates multiple similarity features and applies five classification models. Zhang et al. (2015) built a high-order similarity weight network and used a semi-supervised label propagation algorithm to predict drug–drug interactions. Drug–drug interaction (Ferdousi et al. 2017) is predicted by the functional similarity features of drug, including carriers, transporters, enzymes and targets (CTET). Zhang et al. (2017b) adopted ensemble learning model to predict DDI and evaluated the contribution of different features extracted from multiple data sources. Kastrin et al. (2018) introduced several statistical machine learning models to predict DDI in terms of semantic similarity and topology similarity.

Traditional machine learning-based methods provide a comprehensive guideline for information extraction and feature construction of drug. However, feature engineering is often time-consuming and labor-intensive. Recently, deep learning models have been widely studied in many fields due to its powerful ability of modeling complicated relations and extracting high-order features. Ryu et al. (2018) proposed a deep learning model called DeepDDI. They treat chemical drug substructure fingerprint (SSP) as input feature to predict 86 types of DDI interactions. In the same way, Lee et al. (2019) integrated three types of drug similarity fingerprints SSP, GSP and TSP. And they introduced an autoencoder to solve the curse of dimensionality aroused by additional fingerprints. Rohani and Eslahchi (2019) constructed several drug similarity matrices based on substructure, target, side effect, off-label side effect, pathway, transporter, and indication data. The subset of those matrices is then selected to integrate into an ensemble similarity matrix by a heuristic method. Finally, the drug representation denoted by the ensemble similarity matrix is fed into deep neural network for predicting DDI.

A multi-modality drug representation learning method based on autoencoder DDI-MDAT is proposed by Zhang et al. (2020), and positive-unlabeled (PU) learning setting is adopted for higher accuracy. Deng et al. (2020) designed a multi-modal deep learning framework and generated drug molecular structure, target, pathway and enzyme fingerprints to describe a drug pair.

The aforementioned methods have achieved remarkable results in DDI prediction. However, they usually predicted DDI by binary classification models, which cannot tell the clear boundary of the decision due to the uncertainty and incompleteness of drug information. Further, the process of feature construction is labor-intensive and the information extracted from different data sources lack of consistency and globality. In contrast, we consider uncertainty of drug information in real word by introducing three-way decision classification framework to model uncertain decision boundary for DDI prediction, and utilize knowledge graph embedding as the input feature of delay decision. In the following Sects. 2.2 and 2.3, we will give a brief introduction to knowledge graph embedding in DDI prediction and application of three-way decision.

2.2 Knowledge graph embedding

Knowledge graphs, a form of structured human knowledge, have aroused widespread interest from both the academia and the industry (Ji et al. 2021). Knowledge graph is composed of fact triple (h, r, t) , where h and t are the head entity and tail entity, respectively, and r denotes the relation between h and t . From the perspective of graph theory, knowledge graph belongs to heterogeneous graph, consisting of different types of nodes and edges corresponding to entities and relations in the knowledge graph. To utilize rich knowledge contained in knowledge graph, we need knowledge graph representation learning, namely knowledge graph embedding, aiming to map entities and relations into low-dimensional vectors while capturing their semantic information. Score function is a key component within knowledge graph embedding models. Current models can be roughly classified into two categories according to the score function: translation distance-based model and semantic similarity-based model.

Translation distance-based models, such as TransE (Bordes et al. 2013), TransH (Wang et al. 2014) and TransR (Lin et al. 2015) measure the plausibility of fact triples by calculating the distance between entities, where additive translation with relations as $h + r \approx t$ is widely used. Semantic similarity-based models measure the plausibility of fact triples by semantic matching. They usually adopt a multiplicative formulation, i.e., $h^\top M_r \approx t^\top$, to transform head entity near the tail in the representation

space. Typical models, such as DistMult (Yang et al. 2014) and ComplEx (Trouillon et al. 2016), can be divided into this class. Knowledge graph has been studied and applied in many fields, such as recommendation systems and question answering systems. There are also some applications of knowledge graph in bioinformatics (Lan et al. 2021b, 2022). And several works have also introduced knowledge graphs embedding to DDIs prediction. Abdelaziz et al. (2017) constructed knowledge graph based on multiple data source and proposed framework named Tiresias that generated drug pair representation according to cosine similarity of knowledge graph embedding vectors. Karim et al. (2019) applied knowledge graph as a tool for data integration to overcome data skewness, and conducted ComplEx over knowledge graph to obtain embedding vectors. DDIs are predicted by convolutional-LSTM model, where knowledge graph embedding is the only input. Instead of applying knowledge graph in feature generation, Dai et al. (2020) modeled multiple DDIs data as a knowledge graph and designed a knowledge graph embedding model based on Wasserstein Adversarial Autoencoder. DDI prediction is treated as link prediction over the knowledge graph. The above works have proved that knowledge graph is a promising tool for many tasks. Knowledge graphs related to drug are established based on the information derived from several different databases. They contain more detailed information about drug and the interactions with other entities, such as gene and protein, which enhances the performance of DDI prediction. The knowledge graph embedding feature can provide more consistent and global information for drugs.

2.3 Three-way decision

Three-way decision, as a new interpretation of rules in rough set theory, proposed by Yao (2010, 2021), aims to deal with complex and uncertain decision-making. It has been widely applied in several domains of Artificial Intelligence. Zhou et al. (2014) introduced three-way decision for e-mail spam filtering. Huaxiong et al. (2016) applied it for face recognition and designed cost-sensitive sequential three-way decision framework. Zhang et al. (2017a) designed a regression-based three-way recommender system to minimize the average cost by adjusting the thresholds for different behaviors. Li et al. (2017) developed a three-way decision model for handling the uncertain boundary to improve the binary text classification performance based on the rough set techniques and centroid solution. Zhang et al. (2019) introduces a three-way enhanced convolutional neural network model named 3W-CNN for sentence-level sentiment classification. Yu et al. (2020) proposed an active three-way clustering method to model uncertainty relations between objects and improve

the accuracy of high-dimensional multi-view clustering. Li and Huang (2020) integrated three-way decision into a fuzzy condition decision information system for credit card evaluation. And three-way decision was introduced into the traditional k -means clustering to combine knowledge of set-pair information granule (Zhang et al. 2021). The above works enrich the theoretic foundation of three-way decisions, and indicate that three-way decisions are capable of handling many practical decision problems. This paper introduces three-way decision to model uncertainty decision boundary of DDI prediction. Further, we treat the knowledge graph embedding feature as the delay decision feature of three-way decision for more accurate prediction.

3 Drug–drug Prediction using three-way decision

In this section, we first provide an overview of proposed 3WDDI framework in Sect. 3.1. The construction of the drug chemical structure feature is presented in Sect. 3.2. The acquisition of knowledge graph embedding features is later illustrated in Sect. 3.3. Finally, the components of three-way decision including boundary division and delay decision are offered in Sects. 3.4 and 3.5, respectively.

3.1 Overview

The procedure of our proposed method is showed in Figure 1. It takes drug chemical structure feature as the primary feature and knowledge graph embedding of drugs in KG as the supplementary feature to predict interaction value for drug–drug pair using three-way decision classification model. Therefore, we design our method as the following steps for DDI prediction:

1. drug chemical structure construction;
2. knowledge acquisition;
3. three-way decision and boundary division;
4. enhancing module.

In **step 1**, we collect the SMILES (Simplified Molecular Input Line Entry Specification) of drug from dataset and compute the drug chemical structure feature. In **step 2**, we train a knowledge graph embedding model on large-scale biomedical knowledge graph to generate embedding for candidate drugs. Then, based on the drug chemical structures of drug pairs, we divide the training samples into three regions including positive, negative and boundary in **step 3**. The boundary region contains samples that cannot be decided immediately into positive or negative. To further handle samples in the boundary region, we feed them to an enhancing module and treat knowledge graph embedding as input feature for the final result in **step 4**.

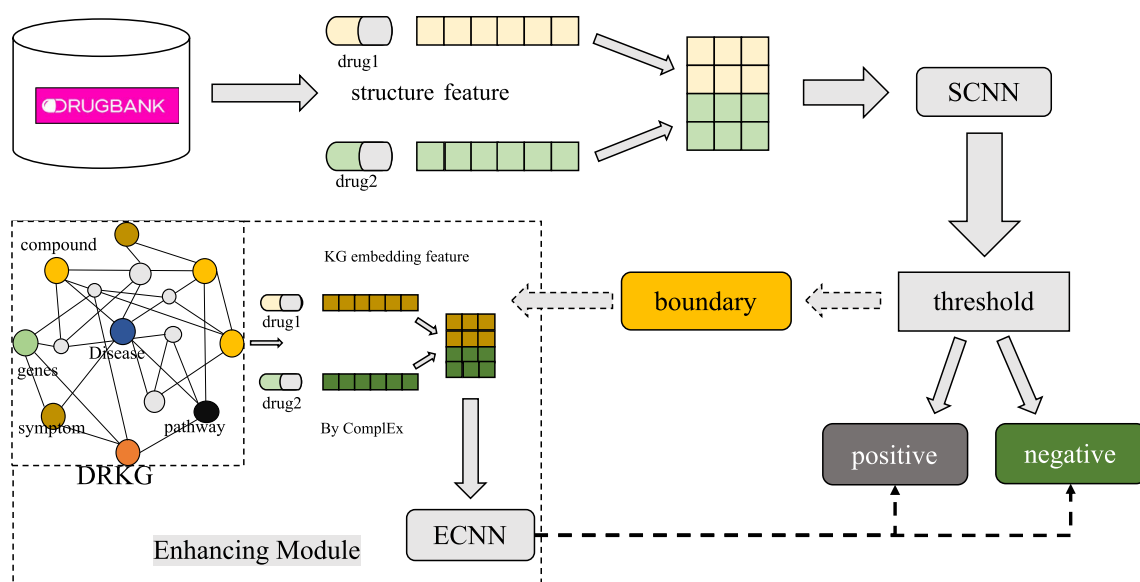


Fig. 1 Overview of 3WDDI

The final score of classification will be calculated in terms of the results from **step 3** and **step 4**. Next, we will describe our proposed method in more details.

3.2 Drug chemical structure feature construction

To train a deep learning model, we need to represent a drug pair as a continuous vector. Drug chemical structure information is critical and easy to obtain. Drug feature based on the chemical structure has already been successfully developed in many DDI prediction studies (Cheng and Zhao 2014; Zhang et al. 2015; Ryu et al. 2018). According to Zhang et al. (2017b), we construct the drug feature based on 881 types of chemical substructures, as known as chemical fingerprints defined in PubChem (Li et al. 2010). Each drug can be represented as an 881-dimensional bit vector where the value 1 or 0 denotes the presence or absence of the corresponding chemical substructure respectively, such as $-NH_2$, $-CH_3$.

We can obtain a high-dimensional and sparse feature vector for each candidate drug. To get a low-dimensional and dense drug feature vector, we calculate the pairwise drug–drug similarity from bit vectors using Jaccard similarity measure. Given drug d_i and d_j in candidate drug set D , and their feature vectors V_i and V_j , the Jaccard similarity can be defined as:

$$S(V_i, V_j) = \frac{M_{11}}{M_{01} + M_{10} + M_{11}}, \tag{1}$$

where M_{11} is the number of chemical substructures shared by d_i and d_j , or the number of dimensions where V_i and V_j both have the value of 1; M_{01} is the number of

substructures that d_j has but d_i does not; and M_{10} is the number of substructures that d_i has but d_j does not.

According to the obtained initial Jaccard similarity, we can obtain a $|D| \times |D|$ pair-wise drug–drug similarity matrix S where the element S_{ij} means the similarity between d_i and d_j . Then, each drug d_i can be represented as a $|D|$ -dimensional dense row vector called F_i^s in matrix S . As a result, for each drug pair (d_i, d_j) in dataset, we are able to concatenate F_i^s and F_j^s to get drug pair representation $[F_i^s; F_j^s]$. This drug chemical structure feature will be used as the input of decision function in Sect. 3.4 below.

3.3 Knowledge acquisition

Compared with the drug structure feature in Sect. 3.2, knowledge graph embedding feature is a distinctive feature containing global information of drugs. This section introduces how to generate the knowledge graph embedding feature for each drug in dataset.

We download a large-scale biomedical knowledge graph DRKG (Ioannidis et al. 2020), which covers all drugs in our dataset and other biology entities, such as gene, disease, pathway and so on, and relationships among them. A popular knowledge graph embedding model ComplEx (Trouillon et al. 2016) is used to generate embedding vectors for all entities and relations in DRKG. We denote the knowledge graph as $G = (N_e, N_r)$, where N_e and N_r is the set of entities and the set of relations, respectively. G is composed of entity–relation–entity triples $T_i = (h_i, r_i, t_i)$ where $h_i, t_i \in N_e$, $r_i \in N_r$. For each entity $e \in N_e$ and $r \in N_r$, knowledge graph embedding model aims to generate embedding vectors $e_e \in \mathbb{R}^{d_e}$ and $e_r \in \mathbb{R}^{d_r}$ where d_e

and d_r mean the dimension of e_e and e_r , respectively (e_x denotes the embedding vector of object x). Each embedding model has a scoring function $f : N_e \times N_r \times N_e \rightarrow \mathbb{R}$ to assign a score $f(h_i, r_i, t_i)$ for a possible triple (h_i, r_i, t_i) , which indicates its plausibility. Models are trained in a way such that for every fact triple $(h_i, r_i, t_i) \in T$ and fake triple $(h'_i, r'_i, t'_i) \notin T$. The models assign scores that satisfy $f(h_i, r_i, t_i) > 0$ and $f(h'_i, r'_i, t'_i) < 0$. A scoring function is generally a function of (e_h, e_r, e_t) .

In this paper, a popular semantic similarity-based model named ComplEx is applied to train embedding vectors for entities and relations. ComplEx represents entities and relations in complex space for modeling multi-type complicated relations. Given $h, t \in N_e$ and $r \in N_r$, it generates embedding vector $e_h, e_r, e_t \in \mathbb{R}^d$ and defines scoring function as:

$$f(h, r, t) = \text{Re}(\langle e_h, e_r, \bar{e}_t \rangle) = \text{Re} \left(\sum_{k=1}^d e_h^{(k)} e_r^{(k)} \bar{e}_t^{(k)} \right), \tag{2}$$

where $f(h, r, t) > 0$ for fact triples and $f(h, r, t) < 0$ for fake triples. Re denotes the real part of a complex number. \bar{e}_t is the conjugate vector of complex vector e_t .

According to complEx model, we can get an embedding vector for each entity and relation in the biomedical knowledge. For each drug d_i in dataset, we can find a corresponding compound entity e_i in the knowledge graph G . Here, we use embedding vector e_i of entity e_i as the representation of drug d_i . According to Eq. 2, for each drug pair (d_i, d_j) , we concatenate e_i and e_j to yield drug pair representation $[e_i; e_j]$, which is regarded as a new feature for boundary samples and fed to enhancing model for further decision. Considering the consistency of formula, we denote the drug pair knowledge graph embedding feature $[e_i; e_j]$ as $[F_i^e; F_j^e]$,

3.4 Three-way decision and boundary division

This section describes the theory of Three-way Decision and discusses how to divide decision boundary based on the drug chemical structural feature of Sect. 3.2.

Three-way decision is originally derived from rough set theory. In rough set theory, object x belongs to either set C or $\neg C$, where $C \cup \neg C = U$, and U is a finite non-empty set called the universe. For each object x , there are three decisions to choose, including dividing x to positive region POS (predicting $x \in C$), negative region NEG (predicting $x \notin C$) or boundary region BND (unpredicting). According to the differences between real label and decisional label of the object x , there will be six decision actions and corresponding six kinds of costs (Table 1): λ_{PP} ,

λ_{BP} , λ_{NP} and λ_{NN} , λ_{BN} , λ_{NN} . λ_{PP} , λ_{BP} , and λ_{NP} denote the costs of dividing x into POS, NEG and BND, respectively in case of $x \in C$. λ_{NN} , λ_{BN} , λ_{NN} mean the costs of dividing x into POS, NEG and BND, respectively while $x \in \neg C$. In the practical application environment, the values of costs are given by domain experts.

Generally, the inequations $\lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}$ and $\lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}$ should be satisfied, indicating that the cost of classifying an object x that belongs to C into the positive region POS(C) is less than or equal to the cost of classifying x into the boundary region BND(C), and both of these losses are strictly less than the cost of classifying x into the negative region NEG(C). The goal of three-way decision is to make a proper decision in a minimum cost. Based on the two inequations, Bayesian decision procedure suggests the following minimum-risk decision rules:

$$\text{decide } x \in \begin{cases} \text{POS}(C), \text{ If } P(C | x) \geq \alpha \\ \text{BND}(C), \text{ If } \alpha \leq P(C | x) \leq \beta, \\ \text{NEG}(C), \text{ If } P(C | x) \leq \beta \end{cases} \tag{3}$$

where α and β are a pair of threshold parameters used to divide decision boundary, and $\alpha = \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}$, $\beta = \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}$, $0 \leq \beta \leq \alpha < 1$; $P(C | x) \in (0, 1)$, called decision status value, is the predicted probability of x belonging to C , and is calculated by a decision function. In the application of three-way decision for machine learning, the decision function is generally a certain machine learning model. For example, Naive Bayes model is adapted as the decision function in the three-way spam filtering system based on Naive Bayes Zhou et al. (2014). Our method applies a CNN model as the decision function.

In our task, object x is a pair of drugs (d_i, d_j) . C denotes the set of drug pairs in which an interaction exists between two drugs, while $\neg C$ not. In other words, $x \in C$ indicates that d_i interacts with d_j , while $x \notin C$ indicates that d_i does not interact with d_j in contrast. $P(C | x)$ is the probability that d_i may interact with d_j . We introduce a CNN model as the decision function to predict the probability of $x \in C$. The drug substructure feature representation $[F_i^s; F_j^s]$ of drug pair (d_i, d_j) will be treated as the input feature of the CNN model called SCNN. The forward propagation of SCNN is calculated as:

Table 1 Costs matrix of three-way decision

	POS(C)	NEG(C)	BND(C)
C	λ_{PP}	λ_{BP}	λ_{NP}
$\neg C$	λ_{NN}	λ_{BN}	λ_{NN}

$$f_{scnn}([F_i^s; F_j^s]) = MLP(\text{vec}(\sigma([F_i^s; F_j^s] * \omega))). \quad (4)$$

In the above formula, $*$ is the convolution operator and ω is the convolution kernel. σ denotes the activation function and $\text{vec}()$ means reshaping a tensor to a vector. MLP is a multi-layer perceptron and its output layer contains only one neuron activated by a sigmoid function. Therefore, the output value of SCNN, represented as $f_{scnn}[F_i^s; F_j^s]$, ranging from 0 to 1, is regarded as the decision status value of three-way decision framework by which we can classify drug pairs (d_i, d_j) into POS(C), BND(C), and NEG(C) with two threshold parameters α and β . For drug pairs divided into BND(C), delay decision will be executed in Sect. 3.5.

3.5 Enhancing Module

Enhancing Module is an important component of three-way decision classification model, which makes it outperform binary classification models. According to the work of Zhang et al. (2019), we choose knowledge graph embedding of each drug from DRKG as supplementary features and the same CNN classifier in Sect. 3.4. We use the same model but different features in boundary division and enhancing module, while Zhang et al. (2019) uses the same feature of object but totally different classification models to guarantee the complementary property between boundary division and enhancing module. Global information contained in knowledge graph embedding focus on the drug interactions with other entities, while structure feature only considers local structure characteristics. In this way, different drug features between boundary division and enhancing module can maintain the complementary property between them. Meanwhile, treating the same deep learning model as decision models of boundary division and enhancing module can maintain strong ability of learning and feature extraction in both periods.

For drug pair (d_i, d_j) in BND(C), we have obtained the knowledge graph embedding representation of (d_i, d_j) , denoted as $[F_i^e; F_j^e]$ in Sect. 3.4. Then, following the Eq. 4, the second CNN that has the same network structure with SCNN, called ECNN will utilize the feature $[F_i^e; F_j^e]$ to produce a new prediction probability $f_{ecnn}[F_i^e; F_j^e]$ for drug pairs in BND(C). Finally, each drug pair in BND(C) will be classified into POS(C) or BND(C) according to the final prediction probability scored by ECNN. Given a drug pair (d_i, d_j) , the pseudo code of predicting it as positive or negative is showed in Algorithm 1 below.

Algorithm 1 The procedure of dividing drug pairs into positive or negative

Input: drug pair (d_i, d_j) ; drug chemical substructure feature $[F_i^s; F_j^s]$; drug knowledge embedding feature $[F_i^e; F_j^e]$; threshold α, β ; decision function $f_{scnn}()$; enhancing model $f_{ecnn}()$.

Output: divide (d_i, d_j) to POS(C) or NEG(C)

```

1: if  $\alpha < f_{scnn}([F_i^s; F_j^s]) < \beta$  then
2:   if  $f_{scnn}([F_i^s; F_j^s]) \geq 0.5$  then
3:     divide  $(d_i, d_j)$  to POS(C)
4:   else divide  $(d_i, d_j)$  to NEG(C)
5:   end if
6: else divide  $(d_i, d_j)$  to BND(C)
7:   if  $f_{ecnn}([F_i^e; F_j^e]) \geq 0.5$  then
8:     divide  $(d_i, d_j)$  to POS(C)
9:   else divide  $(d_i, d_j)$  to NEG(C)
10:  end if
11: end if

```

4 Experiment results

In this section, we introduce the experiment and analyze the result. We first introduce the used datasets in Sect. 4.1. Then, the baseline and parameter settings are detailed in Sect. 4.2. Finally, Sect. 4.3 discusses the results.

4.1 Datasets

DDIs label data and large-scale biology knowledge graph data are needed for our model training and evaluation. For DDIs data, we introduce a popular DDI data zhangDDI (Zhang et al. 2017b) as experiment data, which contain 548 drugs and 48,548 pairwise known DDIs. However, since zhangDDI was collected 4 years ago, some newly found DDIs by subsequent studies might not be included. Thus, instead of directly treating drug pairs without known interactions as negative samples, we validate them based on the latest DrugBank (Law et al. 2013) which covers millions of newly discovered DDIs. We remove negative DDIs that appear in DrugBank to update the dataset zhangDDI. Finally, we randomly select 48,548 negative sample from the validated negative samples set.

For the knowledge graph, we choose DRKG constructed for drug repurposing of COVID-19 (Ioannidis et al. 2020) as our knowledge graph, which contains several drug-related biological entities and relations. DRKG is a comprehensive biological knowledge graph related to genes, compounds, diseases, biological processes, side effects and symptoms that are extracted from six datasets including DrugBank, Hetionet, GNBR, String, IntAct and DGIdb.

Statistical information about DRKG is shown in Table 2. Compound entities in DRKG have already covered all 548 candidate drugs of zhangDDI. Drug (compounds in DRKG)-related entities and relations information will be maintained in embedding vectors generated by ComplEx to promote the prediction of unexpected DDIs.

As mentioned in Sect. 3, the problem of DDI prediction is modeled as binary classification. Therefore, we apply several evaluation metrics of classification to measure the performance of prediction model including accuracy (ACC), the area under the precision–recall curve (AUPR), the area under ROC (AUC), and F1 score. And we take ACC as our primary evaluation metrics.

4.2 Baselines and parameter settings

We compare our model against a variety of baselines which covers two state-of-the-art DDI prediction methods DeepDDI (Ryu et al. 2018), KGDDI (Karim et al. 2019) and three classical machine learning classification models including logistic regression (LR), random forest (RF), and k-nearest neighbor (KNN). DeepDDI is a deep learning method that predicts multi-type DDIs with dimension-reduction drug structure similarity feature. KGDDI applies drug embedding from knowledge graph and predicts DDIs with conv-LSTM. Both DeepDDI and KGDDI are similar to our model. In addition, to validate the effectiveness of delaying decision from three-way decision, we remove the delay decision of proposed model (3WDDI-delay) and test its performance.

We implement DeepDDI and KGDDI according to the original study (Ryu et al. 2018; Karim et al. 2019) and adjust the input and output layers for our experiment data. Chemical structure feature representation of a drug pair $[F_i^s; F_j^s]$ is treated as the input feature while 0, 1 labels are the expected prediction value to train DeepDDI. And knowledge graph embedding representation $[F_i^e; F_j^e]$ is used to train KGDDI. We carry out the other baseline methods with scikit-learn python package and adopt grid search to optimize parameters. We finally set regularization coefficient of LR to 100, size of decision trees for RF to 100, and the neighbor number of KNN to 95. Unlike the training of DeepDDI and KGDDI, we concatenate both the structure feature and the knowledge graph embedding as the final

Table 2 Statistics of DRKG

	Entities	Entity types	Relation types	Triples
DRKG	97,238	13	107	5,874,261

representation of a drug pair $[F_i^s; F_j^s; F_i^e; F_j^e]$ to train the other baselines.

In terms of our proposed method, both SCNN and ECNN have the same network structure except for the input layer. We design the CNN with one convolution layer and three full connection layers where convolution is conducted on kernel with 3×3 and 64 filters. The numbers of full connection neurons are 128, 64, 32, respectively. As described in Eq. 4, the output layer is set as one neuron and activated by a sigmoid function. The other hyperparameter settings are provided in Table 3. In addition, two thresholds of boundary division are the only parameters of the three-way decision. After evaluating on different threshold values, we find the best threshold parameters α as 0.1 and β as 0.9.

4.3 Results and discussion

In this section, we compare the performance of our proposed model with the baselines and discuss the effectiveness of delayed decision. Table 4 reports the ACC, AUPR and AUC of the proposed model and baselines, where all of the scores are average obtained from 5 runs and the bold indicates the best performance of each metric. As shown in Table 4, the proposed model 3WDDI significantly outperforms baselines and the join of three-way decision can achieve obvious promotion in several metrics. Specifically,

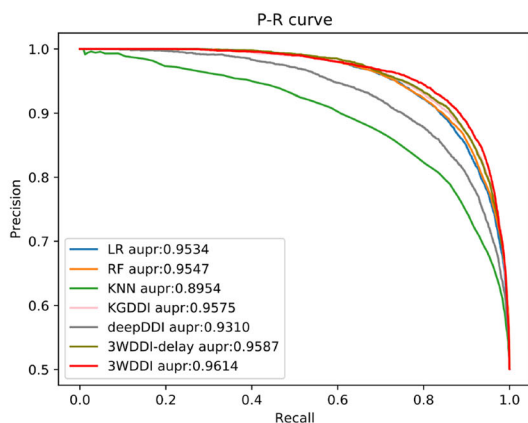
Table 3 Hyperparameter settings of CNN

Hyperparameter	Setting
Dropout	0.4
Batch normalization	Yes
Learning rate	0.001
Optimizer	Adam
Batch size	128
Activation function	ReLU

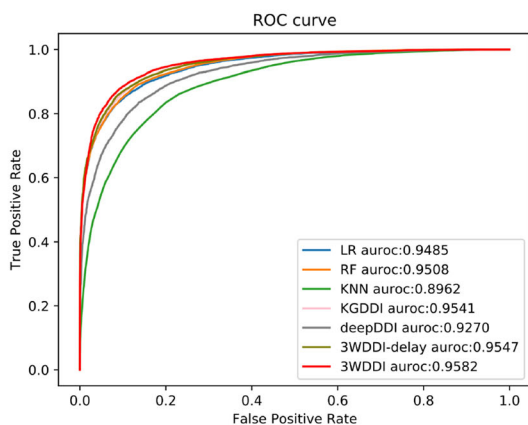
Table 4 Performance of proposed model comparative approaches

Method	ACC	AUPR	AUC	F1 score
DeepDDI (Ryu et al. 2018)	0.8472	0.9310	0.9270	0.8482
KGDDI (Karim et al. 2019)	0.8827	0.9575	0.9541	0.8827
LR	0.8720	0.9534	0.9485	0.8732
RF	0.8778	0.9547	0.9508	0.8780
KNN	0.8155	0.8954	0.8962	0.8236
3WDDI-delay	0.8850	0.9587	0.9547	0.8835
3WDDI	0.8922	0.9614	0.9582	0.8930

compared with structure-based method DeepDDI, 3WDDI achieves about 4.5% improvement on ACC, 3.0% on AUPR and 3.1% on AUC. In addition, our method is also slightly better than embedding-based method KGDDI, which are 1.0%, 0.4% and 0.4% on ACC, AUPR and AUC. RF model achieves the best performance among three classical classification methods. The possible reason is that the performance of ensemble learning model tends to be better than a single model. The precision–recall curves and ROC of 3WDDI against compared methods are shown in Fig. 2. To evaluate the performance of three-way decision, we remove the enhancing module of three-way decision (3WDDI-delay). It is observed that 3WDDI-delay drops 0.7%, 0.3%, and 0.4% on ACC, AUPR and AUC, suggesting that delayed decision with supplementary knowledge can improve the accuracy of DDIs prediction. The reasons of achieving better performance is that: (i) compared to DeepDDI that only uses drug structure feature, we introduce knowledge graph embedding as auxiliary feature to prediction samples in boundary region more accurately;



(a) The precision–recall curves of 3WDDI and compared



(b) The AUC curves of 3WDDI and compared methods

Fig. 2 Performances of all models

(ii) compared to KGDDI that only uses embedding feature, we add structure information as another feature to DDIs prediction; (iii) compared with classical binary decision, we adopt three-way decision classification to predict DDIs by delaying decision for uncertain DDIs.

5 Conclusion

This article proposes a novel method 3WDDI for drug–drug interaction prediction by combining three-way decision and application of knowledge graph. It applies a simple CNN structure as the decision function and the delay decision model. We prepare drug sub-structure similarity feature and drug embedding from knowledge graph DRKG based on ComplEx for 3WDDI. Then the drug sub-structure similarity feature is used to divide boundary region and provide classification result for the rest regions. The knowledge graph embedding of a drug pair is treated as new supplementary features by 3WDDI to carry out delay decision. We implement the proposed method and conduct comparison experiment on a widely used datasets. The experimental results show that 3WDDI outperforms DDI prediction models of baselines.

There is some work we may address in the near future to improve the DDI prediction. We will consider a more effective strategy to divide sample data into different regions. Further, diverse drug features and multi-omics data (Lan et al. 2020, 2021a; Chen et al. 2019, 2020) can be extracted to promote the performance of DDI prediction.

Acknowledgements The work reported in this paper was partially supported by a National Natural Science Foundation of China project 61963004 and 62072124, a key project of Natural Science Foundation of Guangxi 2017GXNSFDA198033, and a key research and development plan of Guangxi AB17195055.

Data availability Data sharing not applicable to this article as no datasets were generated.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- Abdelaziz I, Fokoue A, Hassanzadeh O et al (2017) Large-scale structural and textual similarity-based mining of knowledge graph to predict drug–drug interactions. *J Web Semant* 44:104–117
- Bordes A, Usunier N, García-Durán A et al (2013) Translating embeddings for modeling multi-relational data. *Adv Neural Inf Process Syst* 26:2787–2795
- Callahan A, Cruz-Toledo J, Ansell P et al (2013) Bio2rdf release 2: improved coverage, interoperability and provenance of life

- science linked data. Extended semantic web conference. Springer, New York, pp 200–212
- Chen Q, Lai D, Lan W et al (2019) ILDMSF: inferring associations between long non-coding rna and disease based on multi-similarity fusion. *IEEE/ACM Trans Comput Biol Bioinform* 18(3):1106–1112
- Chen Q, Qiao Y, Hu F et al (2020) Community detection in complex network based on APT method. *Pattern Recogn Lett* 138:193–200
- Cheng F, Zhao Z (2014) Machine learning-based prediction of drug-drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *J Am Med Inf Assoc* 21(e2):e278–e286
- Dai Y, Guo C, Guo W et al (2020) Drug-drug interaction prediction with Wasserstein adversarial autoencoder-based knowledge graph embeddings. *Brief Bioinform* 22(4):bbaa256
- Deng Y, Xu X, Qiu Y et al (2020) A multimodal deep learning framework for predicting drug-drug interaction events. *Bioinformatics* 36(15):4316–4322
- Ferdousi R, Safdari R, Omidi Y (2017) Computational prediction of drug-drug interactions based on drugs functional similarities. *J Biomed Inform* 70:54–64
- Gottlieb A, Stein GY, Oron Y et al (2012) INDI: a computational framework for inferring drug interactions and their associated recommendations. *Mol Syst Biol* 8(1):592
- Huaxiong L, Libo Z, Bing H et al (2016) Sequential three-way decision and granulation for cost-sensitive face recognition. *Knowl-Based Syst* 91:241–251
- Ioannidis VN, Song X, Manchanda S et al. (2020) DRKG- drug repurposing knowledge graph for covid-19. <https://github.com/gnn4dr/DRKG/>
- Ji S, Pan S, Cambria E et al (2022) A survey on knowledge graphs: representation, acquisition, and applications. *IEEE Trans Neural Netw Learn Syst* 33(2):494–514
- Kanehisa M, Sato Y, Kawashima M et al (2015) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 44(D1):D457–D462
- Karim MR, Cochez M, Jares JB et al. (2019) Drug-drug interaction prediction based on knowledge graph embeddings and convolutional-lstm network. In: International conference on bioinformatics, computational biology and health informatics, pp 113–123
- Kastrin A, Ferik P, Leskošek B (2018) Predicting potential drug-drug interactions on topological and semantic similarity features using statistical learning. *PLoS One* 13(5):1–23
- Lan W, Lai D, Chen Q et al (2020) LDICDL: Lncrna-disease association identification based on collaborative deep learning. *IEEE/ACM Trans Comput Biol Bioinform*. <https://doi.org/10.1109/TCBB.2020.3034910>
- Lan W, Dong Y, Chen Q et al (2021a) IGNSCDA: predicting circrna-disease associations based on improved graph convolutional network and negative sampling. *IEEE/ACM Trans Comput Biol Bioinform*. <https://doi.org/10.1109/TCBB.2021.3111607>
- Lan W, Dong Y, Chen Q et al (2021b) KGANCDA: predicting circRNA-disease associations based on knowledge graph attention network. *Brief Bioinform* 23(1):bbab494
- Lan W, Wu X, Chen Q et al (2022) GANLDA: Graph attention network for lncrna-disease associations prediction. *Neurocomputing* 469:384–393
- Law V, Knox C, Djombou Y et al (2013) Drugbank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* 42(D1):D1091–D1097
- Lee G, Park C, Ahn J (2019) Novel deep learning model for more accurate prediction of drug-drug interaction effects. *BMC Bioinform* 20(1):1–8
- Li Z, Huang D (2020) A three-way decision method in a fuzzy condition decision information system and its application in credit card evaluation. *Granul Comput* 5(4):513–526
- Li Q, Cheng T, Wang Y et al (2010) Pubchem as a public resource for drug discovery. *Drug Discov Today* 15(23):1052–1057
- Li Y, Zhang L, Xu Y et al (2017) Enhancing binary classification by modeling uncertain boundary in three-way decisions. *IEEE Trans Knowl Data Eng* 29(7):1438–1451
- Lin Y, Liu Z, Sun M et al. (2015) Learning entity and relation embeddings for knowledge graph completion. In: Twenty-ninth AAAI conference on artificial intelligence, pp 2181–2187
- Rohani N, Eslahchi C (2019) Drug-drug interaction predicting by neural network using integrated similarity. *Sci Rep* 9(1):1–11
- Ryu JY, Kim HU, Lee SY (2018) Deep learning improves prediction of drug–drug and drug–food interactions. *Proc Natl Acad Sci USA* 115(18):E4304–E4311
- Trouillon T, Welbl J, Riedel S et al. (2016) Complex embeddings for simple link prediction. In: International conference on machine learning, pp 2071–2080
- Wang Z, Zhang J, Feng J, et al. (2014) Knowledge graph embedding by translating on hyperplanes. In: Twenty-Eighth AAAI conference on artificial intelligence, pp 1112–1119
- Yang B, Yih W, He X, et al. (2014) Embedding entities and relations for learning and inference in knowledge bases. *arXiv preprint arXiv:1412.6575*
- Yao Y (2010) Three-way decisions with probabilistic rough sets. *Inf Sci* 180(3):341–353
- Yao Y (2021) Set-theoretic models of three-way decision. *Granul Comput* 6(1):133–148
- Yu H, Wang X, Wang G et al (2020) An active three-way clustering method via low-rank matrices for multi-view data. *Inf Sci* 507:823–839
- Zhang P, Wang F, Hu J et al (2015) Label propagation prediction of drug-drug interactions based on clinical side effects. *Sci Rep* 5(1):1–10
- Zhang HR, Min F, Shi B (2017a) Regression-based three-way recommendation. *Inf Sci* 378:444–461
- Zhang W, Chen Y, Liu F et al (2017b) Predicting potential drug-drug interactions by integrating chemical, biological, phenotypic and network data. *BMC Bioinform* 18(1):1–12
- Zhang Y, Zhang Z, Miao D et al (2019) Three-way enhanced convolutional neural networks for sentence-level sentiment classification. *Inf Sci* 477:55–64
- Zhang Y, Qiu Y, Cui Y et al (2020) Predicting drug-drug interactions using multi-modal deep auto-encoders based network embedding and positive-unlabeled learning. *Methods* 179:37–46
- Zhang C, Gao R, Qin H et al (2021) Three-way clustering method for incomplete information system based on set-pair analysis. *Granul Comput* 6(2):389–398
- Zhou B, Yao Y, Luo J (2014) Cost-sensitive three-way email spam filtering. *J Intell Inf Syst* 42(1):19–45

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