Front. Comput. Sci., 2025, 19(5): 195902
https://doi.org/10.1007/s11704-024-31063-0

REVIEW ARTICLE

Application of machine learning in drug side effect prediction: databases, methods, and challenges

Haochen ZHAO[1,2], Jian ZHONG[1,2], Xiao LIANG[1,2], Chenliang XIE[1,2], Shaokai WANG (✉)[3]

1 School of Computer Science and Engineering, Central South University, Changsha 410083, China
2 Hunan Provincial Key Lab on Bioinformatics, Central South University, Changsha 410083, China
3 David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON N2L 3G1, Canada

Abstract Drug side effects have become paramount concerns in drug safety research, ranking as the fourth leading cause of mortality following cardiovascular diseases, cancer, and infectious diseases. Simultaneously, the widespread use of multiple prescription and over-the-counter medications by many patients in their daily lives has heightened the occurrence of side effects resulting from Drug-Drug Interactions (DDIs). Traditionally, assessments of drug side effects relied on resource-intensive and time-consuming laboratory experiments. However, recent advancements in bioinformatics and the rapid evolution of artificial intelligence technology have led to the accumulation of extensive biomedical data. Based on this foundation, researchers have developed diverse machine learning methods for discovering and detecting drug side effects. This paper provides a comprehensive overview of recent advancements in predicting drug side effects, encompassing the entire spectrum from biological data acquisition to the development of sophisticated machine learning models. The review commences by elucidating widely recognized datasets and Web servers relevant to the field of drug side effect prediction. Subsequently, The study delves into machine learning methods customized for binary, multi-class, and multi-label classification tasks associated with drug side effects. These methods are applied to a variety of representative computational models designed for identifying side effects induced by single drugs and DDIs. Finally, the review outlines the challenges encountered in predicting drug side effects using machine learning approaches and concludes by illuminating important future research directions in this dynamic field.

Keywords machine learning, drug side effects, computational models, databases, Web servers

1 Introduction

The drug design and development process has seen acceleration thanks to advancements in biomedical technology [1]. However, it remains a lengthy, expensive, and high-risk endeavor. Statistics reveal that approximately 90% of new drugs fail following their initial human trials, and a predominant reason for this failure is the occurrence of drug side effects, contributing to 35% of the overall failure rate [2]. Drug side effects, often referred to as adverse drug reactions, are officially defined by China's State Food and Drug Administration as "harmful reactions that occur under normal usage and dosage of qualified drugs and have nothing to do with the purpose of medication" [3]. Notably, drug side effects have become the fourth leading cause of death, ranking just behind cardiovascular disease, cancer, and infectious diseases [4]. In the mid-20th century, a lack of rigorous management in the drug research and development process resulted in numerous catastrophic "drug harm" incidents worldwide, leading to the deaths of nearly 20,000 people due to drug side effects and countless others left injured [5]. These incidents not only impose an economic burden on patients but also strain healthcare systems, resulting in increased morbidity and mortality rates [6]. Moreover, combination therapy, the simultaneous use of multiple medications, often provides superior efficacy compared to single-drug therapy and is a preferred approach [7]. In reality, many patients, particularly the elderly, need to concurrently consume various prescription and over-the-counter drugs to manage distinct yet frequently unrelated medical conditions. However, adverse drug events resulting from Drug-Drug Interactions (DDIs) account for 30% of all adverse drug events, and these events are clinically relevant for up to 80% of elderly cancer patients. [8]. Hence, the study of drug side effects of is paramount importance. It can enhance patient safety by identifying and monitoring potential adverse events, consequently reducing the risks associated with treatment [9]. It also facilitates the improvement of treatment protocols, empowering healthcare professionals to make more informed decisions and provide better patient management, all while enhancing patient autonomy and knowledge regarding their treatments [10]. Furthermore, pharmaceutical companies benefit from such

Received December 28, 2023; accepted April 2, 2024

E-mail: shaokai.wang@uwaterloo.ca

Special Issue—Application of computational techniques in drug discovery

research as it aids in refining the design and development of drugs, thereby enhancing their safety and efficacy [11].

In recent years, there has been a global emphasis on the safety of widely used pharmaceuticals [12]. The focus on drug safety extends throughout the entire lifecycle of drug research and development and carries substantial implications for public health and well-being [13]. For example, during the initial drug design phase, researchers meticulously select specific biomolecules or cellular structures as targets for drug intervention or regulation to influence particular physiological processes [14]. This careful selection helps minimize the risk of adverse side effects, as drugs can act precisely on their intended targets without disturbing normal physiological processes. In the drug molecule design process, medicinal chemists work diligently to optimize molecular structures, enhancing selectivity and affinity for the intended target [15]. Such high selectivity reduces the chances of adverse effects on non-target areas, consequently lowering adverse side effects. As drug development progresses, researchers conduct extensive studies on drug metabolism and toxicity to understand the drug's metabolic pathways in the human body and identify potential harmful effects [16]. This knowledge is pivotal for predicting and mitigating the risk of adverse side effects. However, this stage is inherently complex, involving intricate analyses of structure and Structure-Activity Relationships (SARs), often requiring costly and time-consuming safety pharmacology or risk assessment procedures [17]. During the clinical trial phase, drugs are tested on patients to evaluate their efficacy and safety [18]. If adverse side effects are discovered, researchers can further optimize the drug's dosage or treatment duration or decide whether to continue developing the drug. However, despite comprehensive screening and testing before FDA approval, determining drug side effects is subject to various limitations such as the human body's complex genetic landscape, metabolic variations, and limitations in clinical trial sample sizes [19]. Post-marketing surveillance has, therefore, emerged as a core concern for drug safety [20]. However, observations during the post-marketing phase heavily rely on public databases containing reports submitted by various healthcare professionals. Due to the voluntary nature of reporting, some adverse reactions may go unreported, thereby diminishing the comprehensive assessment of the overall safety profile of the drug. Moreover, The mechanistic pathways of drug side effects are highly intricate. On the one hand, the similarity in drug targets may lead to unintended biological responses by affecting proteins similar to the therapeutic target. On the other hand, drug metabolism within the body can generate metabolites with varying activities or toxicity, and individual sensitivity to these metabolites contributes to differences in side effects. Factors such as immune system reactions, drug interactions, genetic elements, drug concentrations, and individual variations also play roles in shaping the process of drug side effect formation. Therefore, the development of reliable computational models for drug side effect prediction has become an important research focus.

Recently, the application of machine learning methods in scientific research has gained remarkable popularity. These methods offer the capability not only to efficiently process vast datasets but also to distill and elucidate patterns and trends within the data, and even predict unsolved problems. This versatility finds applications across various domains, encompassing medicine, biology, engineering, finance, and more [21]. Using machine learning algorithms to explore drug side effects offers numerous advantages. Firstly, it enables the processing of extensive medical data, amalgamating information from diverse sources to yield a comprehensive understanding of adverse effects associated with various drugs, even those that are rare or emerging. Secondly, machine learning automatically detects patterns and correlations in the data, thereby aiding in the early identification of adverse effects that might otherwise go unnoticed. Furthermore, it can predict potential adverse effects and identify patients who may be more susceptible to such effects, allowing for proactive measures to reduce patient risks. Recently, the expanding literature on drug side effects, coupled with the proliferation of websites and databases containing comprehensive information on drugs, side effects, and associated biological entities, offers researchers ample opportunities to gather data and advance the development of machine learning methods for predicting drug side effects. While numerous review articles have addressed drug side effect prediction, most have concentrated on methods for predicting drug side effect associations. Compared with previous review work, we sort out machine learning-based prediction methods for side effects caused by single drugs and DDIs. Moreover, the study of prediction of the frequency and severity of side effects for drugs is highlighted. The methods are systematically categorized, with each falling into one of four distinct groups based on specific prediction tasks. The subsequent sections of this article are structured as follows: Section 2 provides insights into various datasets and Web servers applicable to drug side effect prediction. Sections 3 and 4 delve into the details of various machine learning methods designed for this purpose. Section 5 claims the current challenges and future research directions in drug side effect prediction. Finally, Section 6 concludes the paper, presenting potential future trends and promising research directions aimed at advancing drug side effect prediction.

2 Data sources

Machine learning-based drug side effect prediction relies extensively on various crucial datasets, which offer information about drug structures, protein targets, and known drug side effects. These datasets are instrumental in the inference and identification of potential drug side effects and protein target properties. Table 1 enumerates the vital datasets used or potentially used in drug side effect prediction. They encompass information about drug side effects, DDIs and the features of drugs, side effects and targets. These datasets assume a pivotal role in the research of drug side effect prediction, serving as the foundation for training and validating machine learning models. Moreover, the network servers dedicated to drug side effect prediction are invaluable tools in the realms of drug research and medicine. These Web servers offer convenient online resources and tools to both

Table 1 Databases for detecting drug side effects

Databases	Latest update	Introduction	URL
SIDER [22]	2015/10 V4.1	A database containing information on marketed medicines and their recorded adverse drug reactions.	See sideeffects.embl.de/ website
STITCH [23]	2016/01 V5	A database of known and predicted interactions between chemicals and proteins.	See stitch.embl.de/ website
BIOSNAP [24]	2018	A biological network collecting various types of interactions between FDA-approved drugs.	See snap.stanford.edu/biodata/ website
TWOSIDES [25]	2019/11	A database of DDIs safety signals mined from the FDA's Adverse Event Reporting System.	See tatonettilab.org/resources/nsides/ website
Offsides [25]	2019/11	A database of individual drug side effect signals mined from the FDA's Adverse Event Reporting System.	See tatonettilab.org/resources/nsides/ website
DDInter [26]	2020/09 V1.0	A comprehensive and open-access database providing abundant annotations for each DDI association.	See ddinter.scbdd.com/ website
DailyMed [27]	2021/10	A database containing labeling submitted to the Food and Drug Administration (FDA) by companies.	See dailymed.nlm.nih.gov/dailymed/ website
MecDDI [28]	2023/02	A database offering the mechanism underlying 110,000 DDIs by explicit description and graphic illustration.	See mecddi.idrblab.net/ website
ADReCS [29]	2023/03 V3.2	A comprehensive drug side effect ontology database providing both standardization and hierarchical classification of drug side effect terms.	See bioinf.xmu.edu.cn/ADReCS/index.jsp website
ChEMBL [30]	2023/05 V33	A manually curated database of bioactive molecules with drug-like properties, including chemical, bioactivity and genomic data.	See ebi.ac.uk/chembl/ website
TTD [31]	2023/07	A database providing information about the therapeutic targets and corresponding drugs information.	See db.idrblab.net/ttd/ website
KEGG [32]	2023/09 V108	A database resource for understanding high-level functions and utilities of the biological system.	See genome.jp/kegg/drug/ website
PubChem [33]	Quarterly update	A database containing information physical properties and biological activities of the compounds.	See pubchem.ncbi.nlm.nih.gov/ website
PharmGKB [34]	Quarterly update	A comprehensive resource that curates knowledge about the impact of genetic variation on drug response.	See pharmgkb.org/ website
Drugs [35]	Quarterly update	A database providing accurate and independent information on more than 24,000 prescription drugs, over-the-counter medicines and natural products.	See drugs.com/ website
UniProt [36]	Quarterly update	A database providing protein sequence and functional information.	See uniprot.org/ website
FAERS [37]	Quarterly update	A database containing information on adverse event reports, medication error reports, and product quality complaints resulting in adverse events that were submitted to the FDA.	See open.fda.gov/data/faers/ website

researchers and medical professionals. They empower users with machine learning models for predicting various facets of drug side effects. This includes drug-side effect associations, drug-protein interactions, DDIs, and more. Additionally, these servers incorporate data visualization tools, aiding users in comprehending and interpreting prediction outcomes. Regular updates and database maintenance ensure that users have access to the most current information. Table 2 highlights some online servers for drug side effect prediction, which equip researchers with powerful instruments to holistically analyze and apply the available data and online resources, enhancing comprehension of the interplay between drugs and side effects. This, in turn, contributes to improved prediction accuracy and positively impacts drug safety and medical research.

3 Predicting side effects caused by single drugs

The machine learning method for single-drug side effect prediction can be categorized based on its task type into the binary prediction of drug-side effect associations and multi-label prediction of side effects for a drug. Moreover, recognizing the limited application of these methods in drug risk-benefit assessment, some studies have shifted their focus towards analyzing and predicting the occurrence frequency or severity of individual drug side effects. This focus is crucial

Table 2 Web servers for detecting drug side effects

Servers	Year	Introduction	URL
DDI-CPI [38]	2014	A Web server can make real-time for predicting side effects caused by DDIs based only on molecular structure.	See cpi.bio-x.cn/ddi/ website
NetInfer [39]	2020	A Web server for predicting targets, therapeutic and adverse effects of drugs via network-based inference methods.	See lmmd.ecust.edu.cn/netinfer/ website
DDI-Preditor [40]	2021	A Web server dedicated to quantitative prediction of the impact on drug exposure of DDIs mediated by cytochromes P450 3A4, 2D6, 2C9, 2C19, and 1A2.	See www.ddi-predictor.org/ website
BioChemDDI [41]	2022	A Web server for predicting DDI by fusing biochemical information and structural information through self-attention mechanism.	See 120.77.11.78/BioChemDDI/ website
DeepLGF [42]	2022	A Web server for predicting DDI by using local-global information calculated based on biological knowledge graph.	See 120.77.11.78/DeepLGF/ website
DDI-GCN [43]	2023	A Web server for predicting DDIs by utilizes graph convolutional networks (GCN) based on chemical structures.	See wengzq-lab.cn/ddi/ website
MecDDI [28]	2023	A Web server for clarifying the mechanisms underlying >1,78,000 DDIs by explicit descriptions and graphic illustrations and providing a systematic classification for all collected DDIs based on their clarified mechanisms.	See mecddi.idrblab.net/ website

for patient care in clinical practice and equally essential for pharmaceutical companies, as it mitigates the risk of drug withdrawal from the market.

3.1 Drug-side effect association prediction

The drug side effect association prediction involves binary classification, aiming to determine whether a given drug is associated with a specific side effect. Machine learning models for drug side effect association prediction typically take drug side effect pairs as input samples, incorporating features related to both drugs and side effects. The output for each sample is a binary vector. The machine learning-based process for predicting these associations can be outlined as follows: Initially, feature extraction is performed for both drugs and side effects, followed by the fusion of these features into a comprehensive feature vector. Next, a trainable model is constructed to classify and predict the likelihood of a drug causing a specific side effect based on this feature vector. The output of the model typically provides a probability score, $P(y \mid x)$, indicating the presence (1) or absence (0) of a side effect for the drug, where y represents a binary label (see Fig. 1). Machine learning methods for drug-side effect association prediction are broadly categorized into traditional machine learning-based methods, matrix factorization-based methods, KG-based methods, network-based methods, and deep learning-based methods (see Table 3).

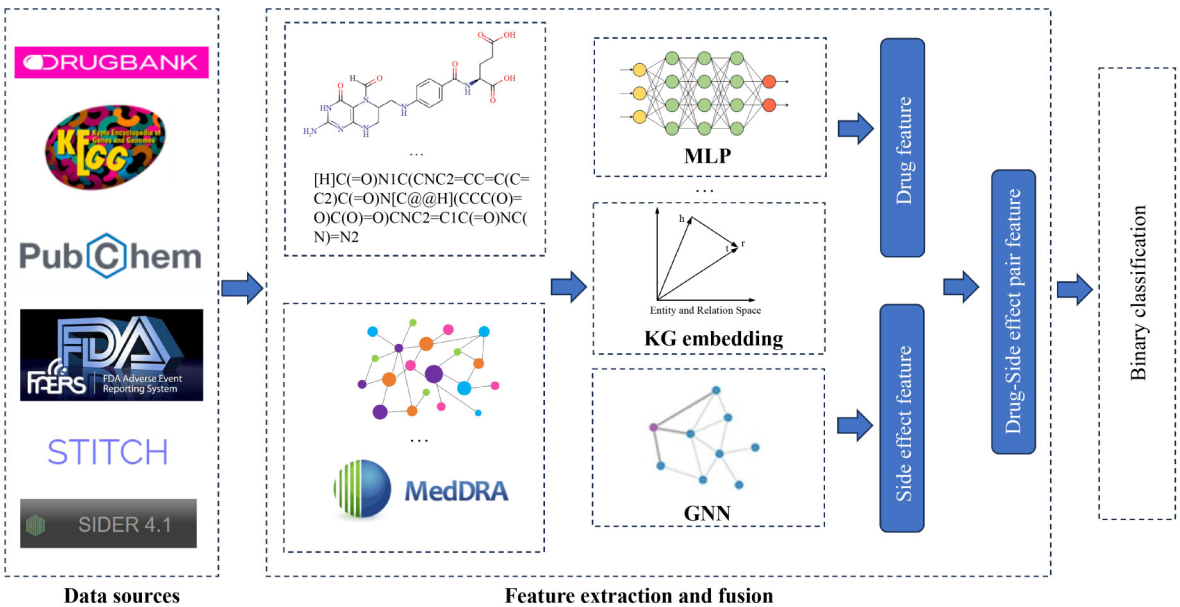

Fig. 1 Overview of drug–side effect association prediction based on machine learning algorithms

Table 3 Methods for predicting drug-side effect associations with binary classification

Method	Year	Remarks	Validation	Dataset	Code or software	F1	AUC	Precision	AUPR	ACC	Recall
Zhao et al. [44]	2018	Traditional machine learning-based method	10-fold cross-validation	SIDER	No	/	0.849	/	/	/	/
FGRMF [45]	2018	Matrix factorization-based method	5-fold cross-validation	SIDER	Yes	0.428	0.949	0.433	0.413	0.988	0.424
IC-PNM [46]	2019	Network-based method	training and test sets	FAERS	No	/	0.910	/	/	/	/
Liang et al. [47]	2020	Traditional machine learning-based method	10-fold cross-validation	SIDER	No	0.959	/	0.997	/	0.975	0.923
Ietswaart et al. [48]	2020	Traditional machine learning-based method	5-fold cross-validation	FAERS	Yes	/	/	/	/	/	/
TMF Guo et al. [49]	2020	Matrix factorization-based method	5-fold cross-validation	SIDER	Yes	/	0.943	/	0.680	/	/
Dasgupta et al. [50]	2021	Knowledge graph-based method	5-fold cross-validation	OMOP	No	0.904	0.945	/	/	/	/
Lee et al. [51]	2021	Deep learning-based method	training and test sets	SIDER	No	0.664	0.846	0.925	/	0.514	0.518
Wu et al. [52]	2022	Traditional machine learning-based method	10-fold cross-validation	SIDER	No	/	0.969	/	0.977	/	/
Yang et al. [53]	2022	Traditional machine learning-based method	5-fold cross-validation	SIDER	No	/	0.875	/	/	/	/
Joshi et al. [54]	2022	Knowledge graph-based method	5-fold cross-validation	SIDER	No	0.837	0.912	0.821	/	0.832	0.857
MPGNN-DSA [55]	2022	Network-based method	10-fold cross-validation	Luo's data	No	/	0.936	/	/	/	/
GCRS [56]	2022	Deep learning-based method	5-fold cross-validation	SIDER	No	/	0.957	/	0.272	/	/
CLMF-NTK [57]	2023	Matrix factorization-based nethod	5-fold cross-validation	SIDER	No	/	0.949	/	0.678	/	/
MSDSE [58]	2023	Deep learning-based method	10-fold cross-validation	SIDER	Yes	0.609	/	/	0.676	/	/
TCSD [59]	2023	Deep learning-based method	5-fold cross-validation	SIDER	No	/	0.977	/	0.351	/	/
GGSC [60]	2023	Deep learning-based method	5-fold cross-validation	SIDER	No	/	0.969	/	0.340	/	/

* The experimental results of all methods in the table are from the corresponding references.

3.1.1 Traditional machine learning-based methods

Traditional machine learning methods for predicting drug side effect associations typically involve combining drug and side effect features as inputs to binary classifiers. In this approach, drug and side effect features are manually extracted before input into the predictive model, requiring human effort for identification and extraction. Several models based on traditional machine learning algorithms have been proposed for predicting drug-side effect associations. For instance, Zhao et al. [44] developed a similarity-based model to predict the association between drugs and side effects based on the dataset from the SIDER [61] database. Their approach involved calculating drug similarities using five features for each drug and employing an RF model for binary classification. The model demonstrates effective performance on the benchmark dataset, with an AUC of 84.9%. Liang et al. [47] introduced a refined negative sample selection strategy using a Random Walk with Restart (RWR) algorithm on a drug network constructed according to the chemical-chemical interactions. This approach distinguishes drugs less likely to have certain side effects as negative samples, and machine learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Network (ANN) are employed to construct classification models. Ietswaart et al. [48] discretized the AC50 values of drugs into one-hot encoded features for predicting drug side effects, training separate RF models for each distinct side effect. Wu et al. [52] proposed a multiple-feature sampling approach for predicting drug side effects, incorporating diverse drug features, including fingerprints, molecular structures, and interactions with target proteins. This method, using multi-type features, achieves accurate predictions for drug side effects using RF classifiers. Similarly, Yang et al. [53] developed a model to understand the phenotypic effects of drugs by examining associations between molecular substructures and phenotypes. This method establishes quantitative associations using L1 Regularized Logistic Regression (L1LOG) and L1 Regularization Support Vector Machine (L1SVM), and predicts potential drug-ATC code and drug-side effect associations, providing deeper insights into the intricate relationships between drug substructures and their phenotypic effects.

3.1.2 Matrix factorization-based methods

The fundamental idea behind matrix factorization methods is to decompose the input matrix into two lower-dimensional matrices while ensuring that the product of these two matrices approximates the original input matrix. In the original matrix, rows represent different drugs, columns represent potential side effects, and each cell contains a value indicating whether a specific drug causes a particular side effect. Since not all drug-side effect relationships are known, this matrix is often sparse and incomplete. Recently, researchers have accurately predicted drug side effects based on common techniques in matrix factorization. For example, Zhang et al. [45] developed a feature-derived graph regularized matrix factorization method, called FGRMF, for predicting unobserved side effects in approved drugs. FGRMF transforms the association between drugs and side effects into a low-dimensional representation, revealing hidden characteristics of both drugs and side effects. Moreover, it builds a graph using drug features and integrates graph regularization to maintain the structural integrity of the drug graph. Guo et al. [49] developed the Triple Matrix Factorization (TMF) model to predict drug-side effect associations. TMF integrates multiple data sources and fuses drug and side effect kernel matrices using Kernel Target Alignment-based Multiple Kernel Learning. It employs low-rank approximation and bipartite projection, using two latent feature matrices to capture drug and side effect properties. On the Pauwels [62], Liu [63], and Mizutani [64] datasets, the TMF shows superior performance compared to other existing methods, obtaining AUPRs of 67.7%, 68.5%, and 68.0%, respectively. Ding et al. [57] proposed CLMF-NTK, a new method for predicting drug-side effect associations using matrix factorization. It utilizes the Neural Tangent Kernel to learn drug and side effect similarity matrices. The model maximizes the correntropy criterion in its objective function to mitigate noise effects and uses an alternating optimization algorithm to derive latent subspace matrices for drugs and side effects.

3.1.3 Knowledge graph-based methods

Knowledge Graph (KG)-based methods integrate varied data and utilize structural features for drug-side effect association prediction through embedding and path analysis. For example, Dasgupta et al. [50] proposed a method to predict drug side effects using literature-derived KGs. The method first extracts entities (drugs, side effects, etc.) and relations from medical literature using natural language processing. Then, it learns representations for each entity on the constructed KG by using DeepWalk [65] and TransE [66] methods. To reduce the impact of extraction errors, the method uses the confidence scores obtained during extraction for weighting. Finally, the concatenated embeddings of drugs and side effects are used to train a machine-learning classifier for drug-side effect associations. Joshi et al. [54] proposed a novel method for predicting drug side effects based on extracted KG embeddings. The method introduces drug-pathway and drug-gene relationships to build a drug KG with six types of nodes and five types of relations. Then the node2vec [67] algorithm is applied to learn embeddings for drug and side effect nodes. These embeddings capture the non-linear relationships between drugs and side effects. Finally, the method uses a customized deep neural network, named KGDNN, to predict and classify the joint embedding features. The method is evaluated on real datasets and shows high prediction accuracy.

3.1.4 Network-based methods

Network-based methods analyze connections within heterogeneous biological networks, using topology and clustering to identify the relationships between drugs and side effects. For example, Yao et al. [55] proposed the MPGNN-DSA method for predicting drug-side effect associations, which operates through a cohesive two-step process. Initially, the method involves creating a Heterogeneous Information Network (HIN) that amalgamates various biomedical datasets, providing a comprehensive backdrop. Concurrently, it

employs a Meta-path-based feature learning module within this network. This module is key in capturing the intricate relationships and characteristics, focusing on accurately representing both drugs and side effects. Utilizing these in-depth features, the method then applies a fully connected neural network for drug-side effect association predictions. In Ji et al.'s study [46], they introduced a novel Information Component Guided Pharmacological Network Model (IC-PNM) for predicting drug-side effect associations, utilizing a combination of network pharmacology and Bayesian statistics. The model is trained on 33,947 drug-side effect pairs from FAERS [37] database with 2010 version and validated on 21,065 new pairs from FAERS during 2011-2015. IC-PNM enhances prediction accuracy by taking into account the frequency of drug-side effect associations. It penalizes the drug-ADE signals when their sample sizes are small since these small sample drug-ADE signals are prone to accidental findings. The experimental results demonstrate an improvement in prediction accuracy for IC-PNM after excluding small samples, with the AUC increasing from 82% to 91%.

3.1.5 Deep learning-based methods

Deep learning methods are more suited for handling complex nonlinear relationships and large-scale, high-dimensional data. In this approach, manual feature extraction is not required as the network automatically identifies relevant features through the learning process, making the analysis and processing more efficient and accurate. Lee et al. [51] developed a novel hybrid deep learning model that integrates a Graph Convolutional Neural Network (GCNN) with Inception modules and a Bidirectional Long Short-Term Memory (BiLSTM) network for the descriptive prediction of drug side effects. The GCNN component effectively learns from the molecular structures of drugs, while the BiLSTM handles textual descriptions related to these side effects. This unique combination of two deep neural networks enables the model to predict drug side effects with high accuracy and interpretability. Trained on a relatively small dataset, the model still achieves impressive performance metrics, including an AUC score of 84.6% and a precision score of 92.5%. Yu et al. [58] proposed a multi-structural deep learning framework, named MSDSE, for predicting drug side effects by integrating multi-scale features of drugs. MSDSE extracts three different types of drug features: SMILES sequences, chemical substructures, and molecular structures. It includes a Convolutional Neural Network (CNN) module to encode local information about the multi-scale features of drugs and a multi-head self-attention module to merge various drug features. Experiments demonstrate that MSDSE outperforms the current state-of-the-art models, showing its robustness and effectiveness. Xuan et al. [56] proposed GCRS, a novel model for predicting drug-side effect associations by integrating specific and common topologies, as well as pairwise attributes from multiple drug-side effect heterogeneous graphs. The model uses Graph Convolutional Autoencoders (GCA) to encode the unique topology of each graph and a shared-parameter GCA for common topologies. A representation-level attention mechanism in GCRS is established to adaptively fuse different topology representations, and an attribute-level attention mechanism is introduced to discern the importance of various attributes in drug-side effect pair embeddings. Then, Xuan et al. [59] proposed a novel method for predicting drug side effect associations named TCSD, which is based on a heterogeneous graph transformer and capsule network approach. This method constructs a complex heterogeneous graph between drugs and side effects, integrating various types of information and attention mechanisms. Additionally, it employs a capsule network strategy to enhance the learning of positional information, thereby more effectively predicting the relationship between drugs and their potential side effects. The model demonstrates a strong predictive capability, achieving an AUC of 97.7% in five-fold cross-validation. Recently, Xuan et al. [60] developed a novel method, named GGSC, for predicting drug-side effect associations by deeply integrating diverse topologies and attributes from multiple heterogeneous graphs. GGSC utilizes graph convolutional autoencoders and multilayer perceptions for feature extraction, alongside representation-level attention and self-calibrated convolutional neural networks for enhanced attribute learning. The results from comparative experiments indicate that GGSC outperformed several of the contemporary advanced prediction methods in terms of predictive performance. Additionally, case studies conducted on five different drugs highlighted GGSC's proficiency in identifying potential candidates for drug-related side effects.

3.2 Multi-label drug side effect prediction

The prediction of drug side effects can be conceptualized as a multi-label prediction task. In this context, a set of drugs is represented as $\{Drug_1, Drug_2,..., Drug_n\}$, where n is the total number of drugs. Each drug is associated with a set of drug descriptors denoted as $\{Descriptor_1, Descriptor_2,..., Descriptor_p\}$, where p represents the number of drug descriptors or input features. Additionally, there exists a label set of side effects, $\{SE_1, SE_2,..., SE_q\}$, where q denotes the total number of potential drug side effects. The drug descriptor sets function as features for machine learning, while the set of side effects serves as the labels. Both sets are utilized for training and prediction within the model. The output of the model is a label vector of length q, indicating the presence or absence of each corresponding side effect for a given drug. This multi-label approach to side effect prediction is visually represented in Fig. 2, where 0 denotes the absence of a side effect, and 1 denotes the presence of the side effect. Machine learning methods for multi-label drug side effect prediction are broadly categorized into traditional machine learning-based methods, kernel-based methods, KG-based methods, Network-based methods and deep learning-based methods (see Table 4).

3.2.1 Traditional machine learning-based methods

In multi-label drug-side effect prediction, traditional machine learning algorithms like SVM, Decision Trees (DT), and RF play the crucial roles and are frequently utilized in various methods. These methods convert drugs into computable

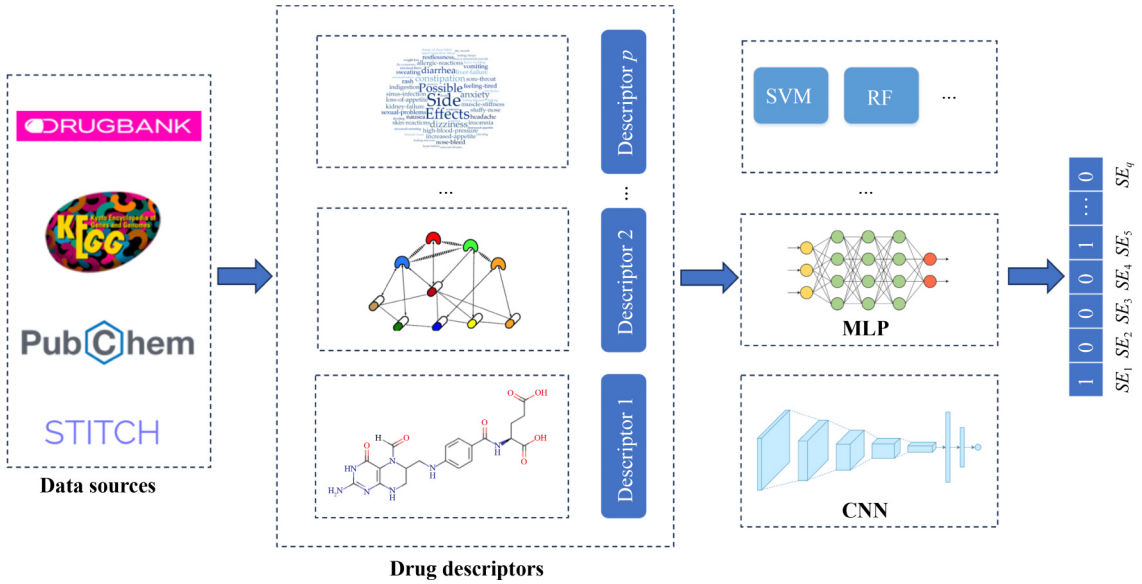

Fig. 2 Overview of multi-label drug side effect prediction based on machine learning algorithms

Table 4 Methods for predicting drug side effects with multi-label classification

Method	Year	Remarks	Validation	Dataset	Code or software	F1	AUC	Precision	AUPR	ACC	Recall
Liu et al. [63]	2012	Traditional machine learning-based method	5-fold cross-validation	SIDER	No	/	0.952	0.662	/	0.967	0.631
Yamanishi et al. [68]	2012	Kernel based method	5-fold cross-validation	SIDER	No	/	/	/	0.209	/	/
Huang et al. [69]	2013	Traditional machine learning-based method	10-fold cross-validation	SIDER	No	/	0.700	/	/	/	/
Cheng et al. [70]	2013	Network-based method	10-fold cross-validation	CTD, SIDER, OFFSIDES	No	/	0.902	/	/	/	/
FS-MLKNN [71]	2015	Traditional machine learning-based method	5-fold cross-validation	SIDER	Yes	/	0.875	/	0.477	/	/
Ngufor et al. [72]	2015	Traditional machine learning-based method	training and test sets	SIDER, FAERS	No	/	0.800	/	/	/	/
DSEP [73]	2015	Deep learning-based method	5-fold cross-validation	SIDER	No	/	0.893	/	0.414	/	/
Rahmani et al. [74]	2016	Network-based method	LOOCV	SIDER	No	/	/	/	/	/	/
Raja et al. [75]	2017	Traditional machine learning-based method	10-fold cross-validation	Biomedical literature	No	0.900	/	/	/	/	/
Bean et al. [76]	2017	Network-based method	10-fold cross-validation	SIDER	Yes	/	0.920	/	/	/	/
Dey et al. [77]	2018	Deep learning-based method	10-fold cross-validation	SIDER	No	/	/	/	/	0.977	/
SDHINE [78]	2018	Deep learning-based method	10-fold cross-validation	SIDER, OFFSIDES	No	/	0.841	/	/	/	/
Ding et al. [79]	2018	Kernel based method	5-fold cross-validation	SIDER	Yes	/	0.951	/	0.655	/	/
DeepSide [80]	2019	Deep learning-based method	3-fold cross-validation	SIDER	Yes	/	0.809	/	/	/	/
Wang et al. [81]	2019	Deep learning-based method	5-fold cross-validation	SIDER	No	/	0.844	/	/	/	/
Ding et al. [82]	2019	Kernel based method	5-fold cross-validation	SIDER	No	/	0.949	/	0.672	/	/
Bongini et al. [83]	2022	Network-based method	training, validation and test sets in a 8:1:1 ratio	SIDER	Yes	/	/	/	/	0.863	/

* The experimental results of all methods in the table are from the corresponding references.

feature descriptors, which are subsequently used to identify their potential side effects. The study by Liu et al. [84] focused on predicting drug side effects using three types of drug descriptors: phenotypic, chemical, and biological. They collect a dataset with 832 drugs and 1,385 side effects and test various machine learning classifiers, including SVM, K-Nearest Neighbor (KNN), Logistic Regression (LR), RF, and Naive Bayes, on different combinations of these descriptors. The study reveals that among all the classifiers tested, the SVM classifier achieves the highest accuracy, with a rate of 96.8%. Similarly, Huang et al. [69] findings reveal that integrating multiple data types, including chemical substructures of drugs, protein-protein interactions, and drug target information, improved the accuracy of drug side effect prediction compared to models using individual data sources. Zhang et al. [71] developed FS-MLKNN, a method combining feature selection and multi-label KNN, to predict drug side effects using drug-related targets and chemical properties. They collect a dataset of 1,080 drugs and 2,260 side effects from SIDER and compare FS-MLKNN with methods from Mizutani [64], Liu [63], and Pauwels [62]. FS-MLKNN performs well, achieving AUPR scores of 42.9% on Pauwel's, 40.0% on Mizutani's, and 48.0% on Liu's datasets. Ngufor et al. [72] enhance the model's capabilities by introducing a novel variational Bayesian ensemble method that integrates four base models (SVM, KNN, LR, extreme LR) for

predicting drug side effects. Their dataset comprised 888 drugs and 1,450 side effects, incorporating biological and chemical properties as well as demographic data for making predictions. The ensemble technique, which assigns optimal weights to each base model, achieved a notable AUC of 80%, surpassing other ensemble methods like majority voting, sum rule, and stacking in drug side effect prediction. Raja et al. [75] presented a novel method for improving the prediction of drug side effects using drug-gene interaction (DGI) data. Their study focuses on integrating DGIs with existing DDIs to enhance the predictive performance of the model. The approach uses a variety of classifiers, achieving an F1-score of 90% when both DDI and DGI features are used, compared to 87% with DDI features alone.

3.2.2 Kernel based methods

Kernel learning for drug side effect prediction is a computational technique that uses kernel functions to map drug and side effect features into a higher-dimensional space. This approach identifies complex, non-linear patterns, enabling the prediction of unknown drug side effects. It enhances the drug discovery process by revealing intricate relationships between drugs and potential side effects in a more discernible feature space. Yamanishi et al. [68] developed a method utilizing multiple kernel regression models to concurrently predict 969 side effects for 658 approved drugs. This method integrates the chemical and biological data of these drugs, combining the chemical space of drug structures with the biological space of drug target proteins in a cohesive framework. Additionally, they perform an extensive prediction of side effects for numerous uncharacterized drug molecules housed in DrugBank [85]. To validate these predictions, they confirm them using independently sourced information. This suggests that their method could be effectively used to gain insights into the side-effect profiles of drugs that have not yet been characterized. In the study conducted by Ding et al. [82], a novel computational method was developed for predicting drug side effects. This method utilizes a multiple kernel learning algorithm, integrating both drug and side-effect data for enhanced predictive accuracy. This method employs Centered Kernel Alignment-based Multiple Kernel Learning (CKA-MKL) and Kronecker Regularized Least Squares (Kronecker RLS) to enhance prediction accuracy. Also, Ding et al. [79] developed a semi-supervised model using a Multiple Kernel Learning (MKL) algorithm to predict drug side effects, integrating multiple kernels from drug and side-effect spaces. By employing graph-based semi-supervised learning, the model effectively combines drug and side effect information, enhancing the accuracy of prediction.

3.2.3 Network-based methods

In the field of multi-label drug side effect prediction, network-based and KG methodologies are frequently employed. For example, Cheng et al. [70] developed MetaADEDB, an extensive database for adverse drug events, which compiles over 520,000 drug-ADE associations from 3,059 unique compounds, including 1,330 drugs, and annotates them using Medical Subject Headings (MeSH). Then, the authors developed the phenotypic network inference model (PNIM) to predict the potential side effects of drugs. Similarly, Rahmani et al. [74] constructed a new Human Drug Network (HDN) by incorporating side effect entities in previous drug networks composed of drug and disease entities. The Restart Walk Random (RWR) method is proposed in this research for predicting the side effects of drugs by propagating information on the HDN. Bean et al. [76] developed the drug-related KG by integrating public datasets and proposed a machine learning algorithm based on enrichment analysis to predict drug side effects. To address the lack of validation in previous drug side effect prediction studies, they validated the results of the prediction model using electronic health records (EHRs). Bongini et al. [83] developed a new approach by first creating a detailed network dataset from various data sources, including drug-gene, drug-drug, and gene-gene interactions, sourced from multiple public databases. Their innovative use of Graph Neural Networks (GNNs) in a multi-label classification setting allows for effective processing of this complex network data, enhancing the prediction of multiple side effects for drugs.

3.2.4 Deep learning-based methods

Deep learning-based methods for multi-label drug-side effect prediction rely on neural network architectures like CNN or Multi-Layer Perceptron (MLP). These networks are capable of autonomously learning complex representations directly from drug features. For example, Niu et al. [73] developed the DSEP neural network, a deep learning model based on MLP, for predicting drug side effects. The model is trained on a dataset comprising 996 drugs and 4,192 side effects from the SIDER database. It utilizes input drug features including chemical descriptors, drug targets, and chemical substructures. The DSEP model utilizes an MLP as its classifier, achieving a high AUC of 89.3%. Dey et al. [77] developed a deep learning model using a CNN module to predict drug side effects without the need for pre-defined molecular substructures. Their approach involves using 2D or 3D graphical representations of drugs to extract detailed chemical features like atom types, bond numbers and types, and atomic valence. These features are then converted into neural fingerprints for analysis. The ability of the model to identify key molecular substructures associated with specific side effect labels through feature analysis and statistical evaluation represents an advancement in understanding molecular-level risk factors. Hu et al. [78] developed SDHINE, a deep learning approach for predicting drug side effects, incorporating protein-protein interaction (PPI) data with drug information from databases like DrugBank and SIDER. This model stands out for its use of semi-supervised stacked denoising auto-encoders (SDAE) to create detailed drug representations. The innovation of SDHINE lies in its integration of PPI data, enhancing accuracy beyond traditional methods focused on direct drug similarities. Wang et al. [81] developed a Deep Neural Network (DNN) model that utilizes the chemical and biological properties of drugs, along with Drug2Vec [86] semantic features extracted from biomedical literature, to detect and predict drug side effects. Uner et al. [80] developed

DeepSide, a deep learning framework that leverages gene expression profiles, chemical structures, and metadata to accurately predict drug side effects. This framework compares five sophisticated deep learning architectures and demonstrates that drug chemical structures offer superior predictive insights compared to gene expression data, with a CNN module based on SMILES strings delivering the most effective results.

3.3 Frequency and severity prediction of drug side effects

Estimating the frequency and severity of drug side effects is crucial for drug risk-benefit assessments, essential in clinical practice for patient care, and of importance to pharmaceutical companies as it aids in preventing the withdrawal of drugs from the market. Our investigation reveals that there are currently five research works (see Table 5) and one research work dedicated to predicting frequency and severity of drug side effects based on machine learning algorithms, respectively. In 2020, Galeano et al. [87] conducted pioneering research on the prediction of drug side effect frequency. Inspired by the concept of movie recommendation systems [88–90], they developed a system akin to movie recommendations but specifically designed for suggesting the frequency of drug side effects. Galeano et al.'s model [87] categorize drug side effects into five frequency classes and harnessed a non-negative matrix factorization model to extract latent feature vectors representing drugs and side effects. These vectors encoded the intricate relationships between drugs and side effects. Leveraging biological interaction data, the model then supplemented and predicted the frequency of unknown drug side effects by fitting side effect frequency values to feature vectors. The model's predictions on side effect frequency offer valuable supplementary hypotheses for risk-benefit assessments during later stages of clinical trials. While Galeano et al.'s model [87] demonstrates good performance, it has limitations, such as a lack of integration with useful features like drug similarity and side effect similarity information. Addressing these limitations, Zhao et al. [91] developed a graph attention model, named MGPred, for predicting drug side effect frequencies using multi-view data. This model aggregates various data sources, including chemical structure similarity, side effect semantic similarity, and known drug-side effect frequency information, into distinct views. Through a graph attention mechanism, these view vectors are automatically reorganized to create a unified predictive embedding for predicting side effect frequencies. MGPred exhibits improved predictive performance compared to Galeano et al.'s work and takes into account multiple aspects of drugs and side effects. Moreover, recognizing that existing methods for predicting drug side effect frequencies rely on known data, making it challenging to predict frequencies for new drugs lacking side effect information, Zhao et al. [92] developed a multi-task model, named SDPred, for drug-side effect associations and corresponding frequencies. SDPred integrates rich features by fusing various similarity matrices with deep convolutional neural networks, even for new drugs lacking prior drug-side effect associations or frequency data. In 2022, Xu et al. [93] proposed a deep learning model (DSGAT) based on the molecular structures of drugs. Using a widely adopted encoder-decoder framework, DSGAT can predicts the drug-side effect associations frequency of drug side effects simultaneously. DSGAT treats atoms in drug molecules as nodes and bonds as edges in a molecular structure graph, employing multi-layer graph attention networks to learn the characteristics of individual atoms. representations for side effects by extracting features from side effect cosine similarity graphs through multi-layer graph attention networks. Finally, a matrix decomposition method serves as the decoder. While this model can predict the frequency of side effects for known molecular structures, it treats drug molecules as independent entities and does not consider potential correlations between drugs and lacks interpretability. While Galeano et al.'s model [87] demonstrating good predictive performance and interpretability, fails to fully utilize multi-view data for drugs and side effects. Addressing this limitation, Wang et al. [94] proposed NRFSE, a neighborhood-regularization method for predicting drug side effect frequencies based on multi-view data. NRFSE employs category-weighted Non-negative Matrix Factorization (NMF) [95] to decompose the drug side effect frequency matrix. It uses a Gaussian distribution to represent unknown drugs, assigning prediction scores for drug-side effect pairs. By considering various views, including drug side effect frequency, chemical structure, Gene Ontology (GO) annotations of drug targets, and Medical Dictionary of Regulatory Activity (MDRA) terminology, NRFSE ensures that the most similar drugs and side effects have similar embeddings and adaptively determines the optimal weight for each view.

Some computational methods have developed to predict potential drug side effects from heterogeneous drug databases. However, the majority of these methods do not address a key concern raised by pharmacologists and pharmaceutical companies - how to assess the severity of drug side effects. To tackle this issue, Zhao et al. [96] established a framework named GCAP, which not only identifies the severity of a drug-

Table 5 Machine learning methods for predicting frequencies and severity of drug side effects based on SIDER

Method	Year	Remarks	Code or software	AUC	AUPR	RMSE	MAE	PCC
Galeano's model [87]	2020	Matrix factorization-based method	Yes	0.907	0.216	1.298	0.953	0.478
MGPred [91]	2021	Deep learning-based method	Yes	0.762	0.120	0.643	0.486	0.734
SDPred [92]	2022	Deep learning-based method	Yes	0.919	0.230	0.593	0.433	0.781
DSGAT [93]	2022	Deep learning-based method	Yes	0.917	0.243	1.031	0.754	0.557
NRFSE [94]	2023	Matrix factorization-based method	Yes	0.926	0.272	1.008	0.767	0.580
GCAP [96]	2023	Deep learning-based method	Yes	0.956	0.946	/	/	/

* The experimental results of all methods in the table are from NRFSE [91] and GCAP [92].

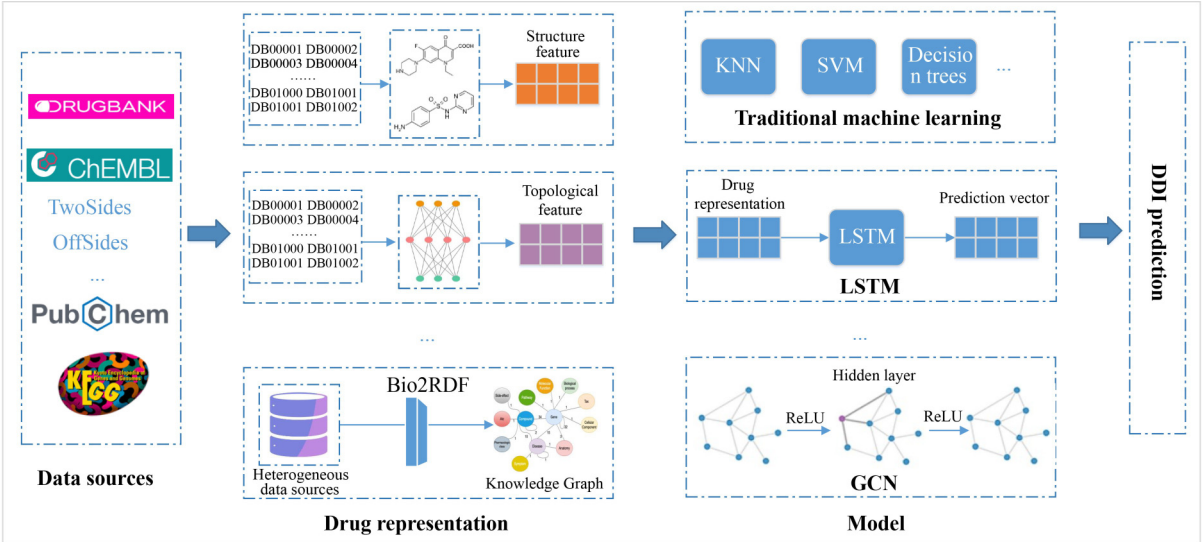

Fig. 3 Overview of side effect prediction resulting from DDIs using machine learning algorithms

side effect pair but also categorizes it into one of seven severity levels. GCAP constructs drug molecular graphs and encodes SMILES sequences into numerical matrices. For side effects, GCAP constructs a Directed Acyclic Graph (DAG) containing semantic descriptors related to each side effect. These descriptors are used to obtain feature vectors for the side effects. The GNN module processes the drug molecules, while the CNN module handles SMILES sequences. By stacking the representations of drugs and side effects and their associations, GCAP uses a multi-head attention mechanism to derive representations for drug-side effect pairs. By deploying different downstream classifiers for each task, GCAP achieves accurate predictions of potential serious side effects and categorizes side effects based on severity.

4 Predicting side effects caused by DDIs

Combination therapy, often employed in practical treatment, becomes a common choice, especially in dealing with complex diseases or symptoms. To address the issue of drug side effects, it is essential to investigate adverse effects caused not only by individual drugs but also by DDIs. Due to the various types of DDIs, such as absorption interaction, distribution interaction, metabolism interaction, excretion interaction, synergy interaction, antagonism interaction, and additive effect interaction, there is a close relationship between these different types of interactions and adverse reactions. This paper focuses on DDIs that may lead to adverse drug reactions. Recently, researchers have begun to design some multi-label and multi-class classifiers for predicting the side effects caused by DDIs, providing more meaningful insights into the potential mechanisms of combined drug use and adverse reactions (see Fig. 3). In multi-class classifiers for predicting side effects caused by DDIs, the objective is to categorize DDIs into specific pharmacological effects and side effects, offering a more detailed description of interaction types. This approach can identify specific interaction types, providing a richer understanding of how DDIs differently. In multi-label classifiers for predicting side effects caused by DDIs, each instance of DDIs can be associated with multiple side effect labels, indicating one or more potential adverse reactions or events corresponding to the DDIs. Moreover, since the majority of studies on predicting side effects caused by DDIs are based on deep learning methods, we categorize existing approaches based on the source data used to represent drug or DDI pairs (see Table 6). Specifically, we differentiate methods that derive representations from molecular structures and drug-related biological entities for a comprehensive discussion.

4.1 Molecular structure-based methods

In recent years, several molecular structure-based methods have successfully predicted the side effects of DDIs. These methods typically use chemical structures, biological activity, and other molecular features to represent drugs and primarily evaluate their performance using data from the DrugBank database. For instance, Ryu et al. [97] developed the DeepDDI framework, utilizing the similarity of drug chemical structures as drug features to predict pharmacological effects and side effects for drug-drug and drug-food component pairs. Based on 192,284 DDIs from DrugBank encompassing 86 pharmacological effect labels and side effect labels, DeepDDI achieves an average accuracy of 92.4%. Considering comprehensive features may enhance the predictive performance of the model, Deng et al. [22] proposed DDIMDL, a multi-modal deep learning framework. DDIMDL integrates various drug features using deep learning frameworks, constructing sub-models based on DNNs for chemical substructures, targets, enzymes, and pathways of drugs. The combined model achieved an accuracy of 88.5% and an AUPR of 92.1%. Although DDIMDL achieves effective performance, it faces underfitting issues due to a scarcity of some rare DDI events. Deng et al. [103] developed the META-DDIE method to address this, using meta-knowledge learned from common DDI events to improve predictive performance for pharmacological effects and side effects of rare DDIs. META-DDIE demonstrated superiority in predicting rare events and contributed to identifying key

Table 6 Machine learning methods for predicting side effects caused by DDIs

Method	Year	Task type	Remarks	Validation	Dataset	Code or software	Macro-F1	AUROC	Macro-precision	AUPR	ACC	Macro-Recall
DeepDDI [97]	2018	Multi-class	Structure-based method	training, validation and test sets	Drugbank	Yes	/	/	/	/	0.924	/
DDIMDL [22]	2019	Multi-class	Structure-based method	5-fold cross-validation	Drugbank	Yes	0.759	0.998	0.847	0.921	0.885	0.7182
Hou et al. [98]	2019	Multi-class	Structure-based method	training, validation and test sets in a 6:2:2 ratio	Drugbank	No	/	0.942	/	/	0.932	/
SSI-DDI [99]	2021	Multi-class	Structure-based method	training, validation and test sets	Drugbank	Yes	/	0.984	/	0.981	0.945	/
MDNN [100]	2021	Multi-class	KG-based method	5-fold cross-validation	Drugbank	No	0.830	0.998	0.862	0.967	0.918	0.820
MUFFIN [101]	2021	Multi-class	KG-based method	5-fold cross-validation	Drugbank	Yes	0.950	/	0.957	/	0.965	0.948
SumGNN [102]	2021	Multi-class	KG-based method	training, validation and test sets in a 7:1:2 ratio.	Drugbank	Yes	0.869	/	/	/	0.927	/
MUFFIN [101]	2021	Multi-label	KG-based method	6-fold cross-validation	TwoSides	Yes	/	0.916	/	0.703	/	/
SumGNN [102]	2021	Multi-label	KG-based method	training, validation and test sets in a 7:1:2 ratio.	TwoSides	Yes	/	0.949	/	0.934	/	/
META-DDIE [103]	2022	Multi-class	Structure-based method	5-fold cross-validation	Drugbank	Yes	/	/	0.816	/	/	/
MDDI-SCL [104]	2022	Multi-class	Structure-based method	5-fold cross-validation	Drugbank	Yes	0.876	0.998	0.880	0.978	0.938	0.877
TBPM-DDIE [105]	2022	Multi-class	Structure-based method	6-fold cross-validation	Drugbank	No	0.843	0.999	0.870	0.970	0.920	0.830
DM-DDI [106]	2022	Multi-class	Network-based method	5-fold cross-validation	Drugbank	Yes	0.852	0.999	0.879	0.964	0.908	0.839
MDF-SA-DDI [107]	2022	Multi-class	Network-based method	6-fold cross-validation	Drugbank	Yes	0.888	0.999	0.909	0.974	0.930	0.876
STNN-DDI [108]	2022	Multi-class	Structure-based method	5-fold cross-validation	Drugbank	Yes	/	/	/	/	/	/
GMPNN-CS [84]	2022	Multi-class	Structure-based method	3-fold cross-validation	Drugbank	Yes	/	0.985	0.936	0.9794	0.953	0.972
LaGAT [109]	2022	Multi-class	KG-based method	4-fold cross-validation	Drugbank	Yes	0.929	/	/	/	0.960	/
DeepMDDI [110]	2022	Multi-class	Network-based method	5-fold cross-validation	Drugbank	Yes	/	0.983	/	0.792	0.967	/
STNN-DDI [108]	2022	Multi-label	Structure-based method	10-fold cross-validation	TwoSides	Yes	/	0.955	0.852	0.921	0.897	/
GMPNN-CS [84]	2022	Multi-label	Structure-based method	3-fold cross-validation	TwoSides	Yes	/	0.901	0.784	0.872	0.828	0.837
R2-DDI [111]	2023	Multi-class	Structure-based method	training, validation and test sets in a 3:1:1 ratio	Drugbank	Yes	0.982	0.997	/	0.996	0.982	/
DSN-DDI [112]	2023	Multi-class	Network-based method	4-fold cross-validation	Drugbank	Yes	0.969	0.995	/	0.994	0.969	/
ACDGNN [113]	2023	Multi-class	Network-based method	training, validation and test sets in a 6:2:2 ratio	Drugbank	Yes	0.941	0.988	0.956	0.984	0.967	0.937
MCFF-MTDDI [114]	2023	Multi-class	KG-based method	5-fold cross-validation	Drugbank	Yes	0.955	0.998	0.972	0.976	0.977	0.946
R2-DDI [111]	2023	Multi-label	Structure-based method	training, validation and test sets in a 3:1:1 ratio	TwoSides	Yes	0.873	0.915	/	0.878	0.862	/
DSN-DDI [112]	2023	Multi-label	Network-based method	3-fold cross-validation	TwoSides	Yes	0.988	0.999	/	0.999	0.988	/
MCFF-MTDDI [114]	2023	Multi-label	KG-based method	4-fold cross-validation	TwoSides	Yes	/	0.933	/	0.717	/	/

* The experimental results of all methods in the table are from the corresponding references.

factors leading to DDI events. Lin et al. [104] proposed MDDI-SCL, a supervised contrastive learning method for predicting side effects caused by DDIs. Comprising three modules, MDDI-SCL demonstrates effectiveness in various tasks, outperforming previous state-of-the-art methods. Shao et al. [105] developed TBPM-DDIE, a transformer-based pre-training method for predicting the pharmacological effects and side effects of DDIs. The authors use the ChEMBL dataset as the TBPM-DDIE's pre-training dataset, which contains 1,722,297 unlabeled compound SMILES sequences. The F1-Score and AUPR of TBPM-DDIE reach 84.3% and 96.9%, respectively, showing its ability to make effective use of unlabeled data in predicting side effects caused by DDIs.

Moreover, some methods decompose the prediction of side effects caused by DDIs into identifying their respective substructure interactions. For example, Yu et al. [108] developed a novel Substructure-aware Tensor Neural Network for predicting side effects caused by DDIs (STNN-DDI). This model constructs a Substructure-Substructure Interaction (SSI) space by learning a three-dimensional tensor of (substructure, substructure, interaction type) triples. Mapping drugs into the SSI space enables STNN-DDI to predict side effects caused

by DDIs in a unified and interpretable way for both transduction and induction scenarios. Performance evaluation experiments on the dataset collected by Zhu et al. [115] show that STNN-DDI improves AUC, AUPR, Accuracy, and Precision compared to the best-performing baseline. STNN-DDI assumes that the interaction between two given drugs is influenced by their local chemical substructures, and the DDI type is determined by the connections between different substructures. However, it overlooks the importance of each substructure for the final prediction. Addressing this concern, Nyamabo et al. [99] introduced a deep learning framework, named SSI–DDI, that directly learns substructure representations from drug molecular graphs. It utilizes a common attention mechanism [116] to discern the importance of each substructure for the final prediction. On a dataset consisting of 1,704 drugs and 191,400 DDIs related to 86 DDI pharmacological effect and side effect labels, collected from DrugBank, SSI–DDI achieves ACC and AUC values of 94.5% and 98.4%, respectively. To facilitate communication between substructures of drugs, Nyamabo et al. [84] developed a gated message-passing neural network (GMPNN-CS). This network learns representations of various sizes and shapes of chemical substructures from drug molecular graphs for predicting side effects caused by DDIs. GMPNN-CS treats edges as gates controlling the message passing flow, partitioning substructures in a learnable way. The final prediction is based on the interaction information learned from their substructures. Lin et al. [111] proposed the R2-DDI framework, encoding the drug and relationship embeddings and constructing relationship-aware fine features for improved accuracy of predicting side effects caused by DDIs. The model extends with a multi-branch structure and incorporates consistency training across different branches to enhance generalization ability. Performance evaluation experiments on the dataset from Nyamabo et al. [84] show that R2-DDI outperformed previous baseline methods.

4.2 Biological entity-based methods

The associations and interactions between drugs and biological entities are crucial for understanding drug functions. By considering the relationships among entities such as drugs, genes, proteins, and more, algorithms can make predictions in a richer biological context, providing more comprehensive information. Deep learning methods for predicting side effects caused by DDIs based on drug-related biological entities can be categorized into Network Topology Structure (NTS)-based methods and KG-based methods.

4.2.1 Network topology structure-based methods

Methods based on network topology play a crucial role in predicting side effects caused by DDIs. These approaches explore the connectivity patterns, degree distributions, and other topological features among nodes in networks constructed based on the associations between drugs and biological entities. They can capture potential interaction relationships, providing valuable insights into the complexity of drug interactions. This network topology-based approach contributes robust tools for predicting side effects resulting from DDIs. For example, Feng et al. [110] developed DeepMDDI, employing a deep Relational Graph Convolutional Network (R-GCN) encoder to capture the topological features of the DDI network. Compared with other multi-class classifiers, DeepMDDI achieves an AUC of 98.3% and AUPR of 79.2%. However, DeepMDDI solely focuses on drug-related features, overlooking potential information in drug-related biological entities like targets and genes. Yu et al. [113] proposed an Attention-based Cross-Domain Graph Neural Network (ACDGNN) for predicting side effects caused by DDIs, incorporating different entities related to drugs and propagating information through cross-domain operations. ACDGNN effectively predicted the potential pharmacological effect labels and side effect labels of DDIs and surpassed existing models, considering abundant information in drug-related biomedical entities in heterogeneous biological networks. To achieve deep integration of drug features and topological relationships, Kang et al. [106] developed a method, named DM-DDI, based on the deep fusion of drug features and topological relationships. DM-DDI adopts a deep fusion strategy to combine drug features and topological structure to learn representative drug embeddings for predicting side effects caused by DDIs, achieving a prediction accuracy of 90.8% and an AUPR of 96.4%, respectively. Lin et al. [107] developed MDF-SA-DDI, which predicts DDI pharmacological effect labels and side effect labels based on multi-source feature fusion and transformer self-attention mechanism. MDF-SA-DDI combines the features of drugs in four different ways and inputs the combined drug features to four different drug fusion networks (conjoined network, convolutional neural network, and two autoencoders) to obtain the features of the drug-drug pairs. The performance evaluation dataset, collected by Deng et al. [22], contains 1,258 drugs, 323,539 DDI, and 100 pharmacological effect labels and side effect labels. The AUPR and F1-Score of MDF-SA-DDI reach 97.4% and 88.8%, respectively. Recently, Li et al. [112] proposed DSN-DDI, a dual-view drug representation learning network for predicting side effects caused by DDIs. DSN-DDI iteratively employs local and global representation learning modules to simultaneously learn drug substructures from single drugs (intra-view) and drug-drug pairs (inter-view). The authors collected 1,706 drugs, 191,808 DDI, and 86 pharmacological effect labels and side effect labels from DrugBank to evaluate the model. In the multi-class classification, the accuracy of DSN-DDI on the transduction setting reaches 96.9%. In the multi-label classification, the accuracy of DSN-DDI on the transduction setting reaches 98.8%.

4.2.2 Knowledge graph-based methods

Compared to (NTS)-based methods, KG-based methods aim to express semantic relationships between entities. These relationships are often challenging to capture using traditional structured data representation methods. By embedding relationships into a graph, KGs can provide richer semantic information. For example, Hong et al. [109] developed the Link-aware Graph ATtention (LaGAT) method for predicting side effects caused by DDIs. LaGAT utilizes the embedding of one drug as a query vector for calculating attention weights,

considering topological neighbor nodes to obtain semantic information about the other drug. LaGAT outperforms baseline methods in binary and multi-class classification based on the datasets collected by Lin et al. [117], and Yu et al. [102], respectively. Lyu et al. [100] proposed a Multi-modal Deep Neural Network (MDNN) for predicting side effects caused by DDIs. MDNN incorporates a dual-path framework, including a drug KG-based path and a heterogeneous feature-based path. MDNN explores the complementarity between drug multi-modal representations through a multi-modal fusion neural layer. Chen et al. [101] developed a multi-scale feature fusion deep learning model, named MUFFIN, for joint learning of drug representations based on drug structures and KG. MUFFIN uses a two-layer cross strategy to fuse multi-modal features effectively. It achieves high accuracy in different datasets, demonstrating its effectiveness in both binary and multi-class classification. Yu et al. [102] proposed the knowledge Summary Graph Neural Network (SumGNN) for predicting side effects caused by DDIs. SumGNN employs a subgraph extraction module, a self-attention-based subgraph summarization scheme, and a multi-channel knowledge and data integration module. Han et al. [114] presented a multi-channel feature fusion model, named MCFF-MTDDI, for predicting side effects caused by DDIs. MCFF-MTDDI addresses feature redundancy and KG noise issues by extracting chemical structure features, drug-drug pairs' additional label features, and KG features. Through a multi-channel feature fusion module, it effectively fuses these features and achieves high accuracy in predicting side effects caused by DDIs.

5 Challenges and future direction

The emergence of drug side effects poses a serious threat to the safety of patient treatment and medicine management. During the past decade, machine learning has been widely applied in the field of bioinformatics, providing a powerful tool for reducing the potential harm caused by drug side effects. However, despite the progress made in many studies dedicated to predicting drug side effects, this field still faces multiple challenges.

5.1 Imbalanced dataset

In real-world scenarios, the occurrence of drug side effects naturally varies, with common side effects like headaches or nausea being more prevalent than rarer ones, such as severe allergies. The long-tail effect in drug side effects can lead to an imbalanced distribution of samples in datasets [118]. In the context of imbalanced data, predictive models tend to excel at identifying and predicting the majority class (common side effects) while exhibiting reduced performance in predicting rarer ones. Moreover, since the number of drug pairs with DDI identified from biomedical databases is lower than those without DDI, the imbalance issue in datasets also presents a crucial and intricate challenge in the prediction of side effects caused by DDIs. Adverse interaction events between different drugs in DDI datasets often exhibit an uneven distribution, particularly affecting the model's recognition ability for the minority class, especially when dealing with low-frequency DDI events during training.

To address the aforementioned issues, various strategies can be employed. For example, one approach involves balancing the sample sizes of different classes in the dataset through methods such as oversampling or undersampling. Alternatively, the Synthetic Minority Oversampling TEchnique (SMOTE) can be applied to generate synthetic samples, thereby increasing the sample count of rare events. Employing ensemble learning methods, such as voting or boosting, to combine results from multiple classifiers can enhance the model's predictive performance for minority classes. Adjusting the weights of rare side effects or drug pairs is another potential solution.

5.2 Data quality and fusion

Data quality and fusion play pivotal roles in the accurate prediction of drug side effects. The scarcity of detailed information on long-term and rare side effects for newly introduced drugs, coupled with reporting mechanism imperfections and limitations in clinical trials, contributes to potential omissions and data loss. This scenario can result in mislabeling positive instances as negative samples, undermining predictive model performance. Individual variations in drug responses, influenced by genetic factors, age, gender, lifestyle, and other health considerations, introduce complexity. Achieving accurate predictions requires comprehensive health records, including genetic backgrounds and medical histories, which pose challenges due to resource costs and privacy concerns. Moreover, to enhance prediction accuracy, integrating multiple drug descriptors is essential, covering aspects like drug chemistry, physical and chemical properties, biological activity, and pharmacokinetic properties. However, challenges arise from the diversity and complexity of descriptors, requiring specific processing methods, expertise, and sophisticated data processing and fusion techniques. Effectively integrating data from diverse sources must minimize information overlap and redundancy. Acquiring high-quality data is challenging but crucial, as prediction accuracy depends on comprehensive and accurate datasets. While the combination of descriptors theoretically enhances accuracy, addressing challenges in data processing, integration, computation, and quality is imperative in practical implementation. Future work should focus on understanding individual differences and developing personalized drug side effect prediction models.

5.3 Predicting side effects caused by drug combinations

Combination therapy can enhance the overall therapeutic efficacy by leveraging different mechanisms, pathways, or target interactions. However, it may also introduce additional challenges such as drug interactions and cumulative side effects. Compared to predicting the side effects of individual drugs, predicting the side effects of combination therapy poses more difficulties and avenues for exploration. Firstly, drug combinations introduce a more intricate network of interactions, where each drug may influence the metabolism and effects of others. Understanding these complex network structures and the dynamics of interactions presents a challenge. Secondly, combination therapy may lead to the

stacking of side effects, where the adverse effect of one drug may interact with specific reactions of another. Accurately capturing such effects requires an in-depth exploration of potential correlations between different drugs. Considering the diversity of drug combinations, the dimensionality of datasets also may grow exponentially. Effectively handling and analyzing this vast amount of data poses severe challenges to algorithms and computational capabilities. Moreover, existing research on the side effects of combination therapy primarily focuses on interactions between two drugs. However, patients often take an average of 3 to 5 drugs, and especially in the case of chronic patients developing acute illnesses, the number of drugs may exceed 10. Future research needs to expand to explore the adverse effects of interactions among more than two drugs.

5.4 Predicting side effects of traditional Chinese medicines

Predicting side effects of Traditional Chinese Medicines (TCMs) is gaining increasing attention, given the rich history of traditional Chinese medicine spanning thousands of years and its proven clinical efficacy. Natural products such as artemisinin and paclitaxel, derived from TCMs, have played a crucial role in saving millions of lives worldwide. The integration of artificial intelligence (AI) in TCMs is on the rise. However, challenges arise due to the relatively limited clinical data for TCMs, and the potential imbalance in the occurrence frequency of adverse reactions among different herbal medicines, making it challenging for machine learning models during the learning and prediction processes. Furthermore, traditional Chinese medicine often consists of complex formulations composed of multiple herbs, introducing a diverse range of components and potential interactions that increase the complexity of predictive models. Moving forward, researchers can enhance the construction of large-scale, high-quality datasets related to adverse reactions in TCMs. This involves integrating data from various medical institutions and research centers to encompass a broader patient population and diverse types of herbal medicines. In the development of machine learning algorithms, it is essential to consider the pharmacological characteristics of each component in TCMs and the potential interactions among multiple components. Lastly, a multidisciplinary research approach that integrates knowledge from traditional Chinese medicine, pharmacology, and computer science is crucial to collectively advance the field of predicting adverse reactions in traditional Chinese medicine.

5.5 Model interpretability

Interpretability stands out as a common challenge encountered in drug side effect prediction models within machine learning applications. These models are frequently perceived as “black boxes”, particularly due to their intricate structures and a multitude of parameters [119]. While the model can capture intricate high-order interactions involving biological properties, chemical structures, and multiple drug interactions, conveying these relationships in an intuitive manner remains elusive, impeding a profound comprehension of the prediction process. Moreover, elucidating individual differences between models across different patients presents a significant hurdle. Even when providing an overarching explanation of the process, the complex and diverse biological characteristics, genomic information, and clinical histories of each patient introduce complexity. Consequently, articulating how the model generates specific predictions at the individual level becomes a complex and challenging task. Future work will focus on developing more transparent and interpretable algorithms, such as feature importance analysis and data visualization tools. This will not only help enhance the trust of medical professionals in these advanced models but also make the model’s prediction results easier to understand and accept.

5.6 Generative large model

Generative large models have made significant progress in many fields recently, such as ChatGPT (See openai.com/blog/chatgpt website). These models have demonstrated excellent performance in various fields with their powerful language generation and comprehension abilities. It is expected that generative large models will play a key role in drug side effect prediction in the future. Firstly, the integration of generative large models with biomedical KGs is expected to enhance the model’s ability to understand potential interactions and capture the diversity of biomedical information more comprehensively by learning the complex relationships between entities such as drugs, proteins, and genes in the KG. Secondly, generative large models can integrate information from various data sources, including text, graphs, and experimental data, to more accurately capture the complexity of DDIs and improve the accuracy of predictions. In addition, further developments include generative models for individual patient differences, which fully consider information such as the patient’s genome and clinical history to achieve personalized drug side effect prediction, providing more targeted support for precision medicine. This will make medical decisions more consistent with the individual characteristics of patients, thereby improving treatment outcomes. By combining generative large models with advanced machine learning techniques, future drug side effect prediction models will be even more powerful.

6 Conclusion

Understanding and preventing drug side effects holds a profound influence on drug development and utilization, profoundly impacting patients’ physical and mental well-being. Traditional artificial drug experimentation methods are not only expensive but also time-consuming, rendering comprehensive testing a challenging task. However, with the advent of advanced technologies, particularly in the realm of machine learning and the availability of extensive biochemical medical data, combining these two has emerged as a pivotal approach for predicting drug side effects. This amalgamation offers notable advantages in terms of speed, efficiency, and cost-effectiveness.

This article provides a comprehensive review of computational methods that employ machine learning to identify side effects stemming from individual drugs as well as DDIs. It outlines the fundamental principles behind

establishing predictive models and introduces commonly utilized databases and Web servers employed in the detection of drug side effects. Finally, the article deliberates on the current challenges and future avenues in machine learning-based methods for discerning drug side effects.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China (Nos. 62072473, U22A2041), the National Funded Postdoctoral Program of China (No. G2C20233162), and the Natural Science Foundation of Hunan Province (No. 2022JJ30750). This work was carried out in part using computing resources at the High-Performance Computing Center of Central South University.

Competing interests The authors declare that they have no competing interests or financial conflicts to disclose.

References

1. Berdigaliyev N, Aljofan M. An overview of drug discovery and development. Future Medicinal Chemistry, 2020, 12(10): 939−947
2. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? Acta Pharmaceutica Sinica B, 2022, 12(7): 3049−3062
3. Aronson J K. Adverse drug reactions: history, terminology, classification, causality, frequency, preventability. In: Talbot J, Aronson J K, eds. Stephens' Detection and Evaluation of Adverse Drug Reactions: Principles and Practice. 6th ed. Chichester: John Wiley & Sons, Ltd, 2011, 1−119
4. Church D L. Major factors affecting the emergence and re-emergence of infectious diseases. Clinics in Laboratory Medicine, 2004, 24(3): 559−586
5. Zullo A, Large M, Amoros E, Martin J L. Estimated number of seriously injured road users admitted to hospital in France between 2010 and 2017, based on medico-administrative data. BMC Public Health, 2021, 21: 649
6. De Kinderen R J A, Evers S M A A, Rinkens R, Postulart D, Vader C I, Majoie M H J M, Aldenkamp A P. Side-effects of antiepileptic drugs: the economic burden. Seizure, 2014, 23(3): 184−190
7. Mokhtari R B, Homayouni T S, Baluch N, Morgatskaya E, Kumar S, Das B, Yeger H. Combination therapy in combating cancer. Oncotarget, 2017, 8(23): 38022−38043
8. Rao A, Cohen H J. Symptom management in the elderly cancer patient: fatigue, pain, and depression. JNCI Monographs, 2004, 2004(32): 150−157
9. Forster A J, Worthington J R, Hawken S, Bourke M, Rubens F, Shojania K, Van Walraven C. Using prospective clinical surveillance to identify adverse events in hospital. BMJ Quality & Safety, 2011, 20(9): 756−763
10. Papaioannou D, Cooper C, Mooney C, Glover R, Coates E. Adverse event recording failed to reflect potential harms: a review of trial protocols of behavioral, lifestyle and psychological therapy interventions. Journal of Clinical Epidemiology, 2021, 136: 64−76
11. Bender R, Beckmann L, Lange S. Biometrical issues in the analysis of adverse events within the benefit assessment of drugs. Pharmaceutical Statistics, 2016, 15(4): 292−296
12. Rácz A, Bajusz D, Miranda-Quintana R A, Héberger K. Machine learning models for classification tasks related to drug safety. Molecular Diversity, 2021, 25(3): 1409−1424
13. Call K T, Riedel A A, Hein K, McLoyd V, Petersen A, Kipke M. Adolescent health and well-being in the twenty-first century: a global perspective. Journal of Research on Adolescence, 2002, 12(1): 69−98
14. Eisenhauer E A, O'Dwyer P J, Christian M, Humphrey J S. Phase I clinical trial design in cancer drug development. Journal of Clinical Oncology, 2000, 18(3): 684−684
15. Fliri A F, Loging W T, Thadeio P F, Volkmann R A. Analysis of drug-induced effect patterns to link structure and side effects of medicines. Nature Chemical Biology, 2005, 1(7): 389−397
16. Dimidjian S, Hollon S D. How would we know if psychotherapy were harmful? American Psychologist, 2010, 65(1): 21−33
17. Bugrim A, Nikolskaya T, Nikolsky Y. Early prediction of drug metabolism and toxicity: systems biology approach and modeling. Drug Discovery Today, 2004, 9(3): 127−135
18. Flynn K E, Kramer J M, Dombeck C B, Weinfurt K P. Participants' perspectives on safety monitoring in clinical trials. Clinical Trials, 2013, 10(4): 552−559
19. Naranjo C A, Busto U, Sellers E M. Difficulties in assessing adverse drug reactions in clinical trials. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 1982, 6(4−6): 651−657
20. Gomes K L G, Da Silva R E, Da Silva Junior J B, Bosio C G P, Novaes M R C G. Post-marketing authorisation safety and efficacy surveillance of advanced therapy medicinal products in brazil, the European union, the united states and japan. Cytotherapy, 2023, 25(10): 1113−1123
21. Sarker I H. Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2021, 2(3): 160
22. Deng Y, Xu X, Qiu Y, Xia J, Zhang W, Liu S. A multimodal deep learning framework for predicting drug-drug interaction events. Bioinformatics, 2020, 36(15): 4316−4322
23. Szklarczyk D, Santos A, Von Mering C, Jensen L J, Bork P, Kuhn M. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. Nucleic Acids Research, 2016, 44(D1): D380−D384
24. Huang K, Xiao C, Glass L M, Sun J. MolTrans: molecular interaction transformer for drug-target interaction prediction. Bioinformatics, 2021, 37(6): 830−836
25. Tatonetti N P, Fernald G H, Altman R B. A novel signal detection algorithm for identifying hidden drug-drug interactions in adverse event reports. Journal of the American Medical Informatics Association, 2012, 19(1): 79−85
26. Xiong G, Yang Z, Yi J, Wang N, Wang L, Zhu H, Wu C, Lu A, Chen X, Liu S, Hou T, Cao D. DDInter: an online drug-drug interaction database towards improving clinical decision-making and patient safety. Nucleic Acids Research, 2022, 50(D1): D1200−D1207
27. Shingjergji K, Celebi R, Scholtes J, Dumontier M. Relation extraction from dailymed structured product labels by optimally combining crowd, experts and machines. Journal of Biomedical Informatics, 2021, 122: 103902
28. Hu W, Zhang W, Zhou Y, Luo Y, Sun X, Xu H, Shi S, Li T, Xu Y, Yang Q, Qiu Y, Zhu F, Dai H. MecDDI: clarified drug-drug interaction mechanism facilitating rational drug use and potential drug-drug interaction prediction. Journal of Chemical Information and

Modeling, 2023, 63(5): 1626−1636
29. Cai M C, Xu Q, Pan Y J, Pan W, Ji N, Li Y B, Jin H J, Liu K, Ji Z L. ADReCS: an ontology database for aiding standardization and hierarchical classification of adverse drug reaction terms. Nucleic Acids Research, 2015, 43(D1): D907−D913
30. Mendez D, Gaulton A, Bento A P, Chambers J, De Veij M, Félix E, Magariños M P, Mosquera J F, Mutowo P, Nowotka M, Gordillo-Marañón M, Hunter F, Junco L, Mugumbate G, Rodriguez-Lopez M, Atkinson F, Bosc N, Radoux C J, Segura-Cabrera A, Hersey A, Leach A R. ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Research, 2019, 47(D1): D930−D940
31. Chen X, Ji Z L, Chen Y Z. TTD: therapeutic target database. Nucleic Acids Research, 2002, 30(1): 412−415
32. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Research, 2000, 28(1): 27−30
33. Kim S, Thiessen P A, Bolton E E, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker B A, Wang J, Yu B, Zhang J, Bryant S H. PubChem substance and compound databases. Nucleic Acids Research, 2016, 44(D1): D1202−D1213
34. Whirl-Carrillo M, McDonagh E M, Hebert J M, Gong L, Sangkuhl K, Thorn C, Altman R B, Klein T E. Pharmacogenomics knowledge for personalized medicine. Clinical Pharmacology & Therapeutics, 2012, 92(4): 414−417
35. Pinkoh R, Rodsiri R, Wainipitapong S. Retrospective cohort observation on psychotropic drug-drug interaction and identification utility from 3 databases: drugs. com®, lexicomp®, and epocrates®. PLoS One, 2023, 18(6): e0287575
36. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Research, 2019, 47(D1): D506−D515
37. Delfino J G, Krainak D M, Flesher S A, Miller D L. MRI-related FDA adverse event reports: A 10-yr review. Medical Physics, 2019, 46(12): 5562−5571
38. Luo H, Zhang P, Huang H, Huang J, Kao E, Shi L, He L, Yang L. *DDI-CPI*, a server that predicts drug-drug interactions through implementing the chemical-protein interactome. Nucleic Acids Research, 2014, 42(W1): W46−W52
39. Wu Z, Peng Y, Yu Z, Li W, Liu G, Tang Y. NetInfer: a Web server for prediction of targets and therapeutic and adverse effects via network-based inference methods. Journal of Chemical Information and Modeling, 2020, 60(8): 3687−3691
40. Moreau F, Simon N, Walther J, Dambrine M, Kosmalski G, Genay S, Perez M, Lecoutre D, Belaiche S, Rousselière C, Tod M, Décaudin B, Odou P. Does DDI-predictor help pharmacists to detect drug-drug interactions and resolve medication issues more effectively? Metabolites, 2021, 11(3): 173
41. Ren Z H, Yu C Q, Li L P, You Z H, Pan J, Guan Y J, Guo L X. BioChemDDI: Predicting drug-drug interactions by fusing biochemical and structural information through a self-attention mechanism. Biology, 2022, 11(5): 758
42. Ren Z H, You Z H, Yu C Q, Li L P, Guan Y J, Guo L X, Pan J. A biomedical knowledge graph-based method for drug-drug interactions prediction through combining local and global features with deep neural networks. Briefings in Bioinformatics, 2022, 23(5): bbac363
43. Zhong Y, Zheng H, Chen X, Zhao Y, Gao T, Dong H, Luo H, Weng Z. DDI-GCN: Drug-drug interaction prediction via explainable graph convolutional networks. Artificial Intelligence in Medicine, 2023, 144: 102640
44. Zhao X, Chen L, Lu J. A similarity-based method for prediction of drug side effects with heterogeneous information. Mathematical Biosciences, 2018, 306: 136−144
45. Zhang W, Liu X, Chen Y, Wu W, Wang W, Li X. Feature-derived graph regularized matrix factorization for predicting drug side effects. Neurocomputing, 2018, 287: 154−162
46. Ji X, Wang L, Hua L, Wang X, Zhang P, Shendre A, Feng W, Li J, Li L. Improved adverse drug event prediction through information component guided pharmacological network model (IC-PNM). IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2021, 18(3): 1113−1121
47. Liang H, Chen L, Zhao X, Zhang X. Prediction of drug side effects with a refined negative sample selection strategy. Computational and Mathematical Methods in Medicine, 2020, 2020: 1573543
48. Ietswaart R, Arat S, Chen A X, Farahmand S, Kim B, DuMouchel W, Armstrong D, Fekete A, Sutherland J J, Urban L. Machine learning guided association of adverse drug reactions with *in vitro* target-based pharmacology. EBioMedicine, 2020, 57: 102837
49. Guo X, Zhou W, Yu Y, Ding Y, Tang J, Guo F. A novel triple matrix factorization method for detecting drug-side effect association based on kernel target alignment. BioMed Research International, 2020, 2020: 4675395
50. Dasgupta S, Jayagopal A, Hong A L J, Mariappan R, Rajan V. Adverse drug event prediction using noisy literature-derived knowledge graphs: algorithm development and validation. JMIR Medical Informatics, 2021, 9(10): e32730
51. Lee C Y, Chen Y P P. Descriptive prediction of drug side-effects using a hybrid deep learning model. International Journal of Intelligent Systems, 2021, 36(6): 2491−2510
52. Wu Z, Chen L. Similarity-based method with multiple-feature sampling for predicting drug side effects. Computational and Mathematical Methods in Medicine, 2022, 2022: 9547317
53. Yang J, Zhang D, Cai Y, Yu K, Li M, Liu L, Chen X. Computational prediction of drug phenotypic effects based on substructure-phenotype associations. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2023, 20(1): 256−265
54. Joshi P, Masilamani V, Mukherjee A. A knowledge graph embedding based approach to predict the adverse drug reactions using a deep neural network. Journal of Biomedical Informatics, 2022, 132: 104122
55. Yao W, Zhao W, Jiang X, Shen X, He T. MPGNN-DSA: a meta-path-based graph neural network for drug-side effect association prediction. In: Proceedings of 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2022, 627−632
56. Xuan P, Wang M, Liu Y, Wang D, Zhang T, Nakaguchi T. Integrating specific and common topologies of heterogeneous graphs and pairwise attributes for drug-related side effect prediction. Briefings in Bioinformatics, 2022, 23(3): bbac126
57. Ding Y, Zhou H, Zou Q, Yuan L. Identification of drug-side effect association via correntropy-loss based matrix factorization with neural tangent kernel. Methods, 2023, 219: 73−81
58. Yu L, Xu Z, Qiu W, Xiao X. MSDSE: predicting drug-side effects based on multi-scale features and deep multi-structure neural network. Computers in Biology and Medicine, 2024, 169: 107812
59. Xuan P, Li P, Cui H, Wang M, Nakaguchi T, Zhang T. Learning multi-types of neighbor node attributes and semantics by heterogeneous graph transformer and multi-view attention for drug-related side-effect prediction. Molecules, 2023, 28(18): 6544
60. Xuan P, Xu K, Cui H, Nakaguchi T, Zhang T. Graph generative and adversarial strategy-enhanced node feature learning and self-calibrated pairwise attribute encoding for prediction of drug-related side effects. Frontiers in Pharmacology, 2023, 14: 1257842
61. Kuhn M, Letunic I, Jensen L J, Bork P. The SIDER database of drugs and side effects. Nucleic Acids Research, 2016, 44(D1): D1075−D1079
62. Pauwels E, Stoven V, Yamanishi Y. Predicting drug side-effect profiles: a chemical fragment-based approach. BMC Bioinformatics, 2011, 12(1): 169

63. Liu M, Wu Y, Chen Y, Sun J, Zhao Z, Chen X W, Matheny M E, Xu H. Large-scale prediction of adverse drug reactions using chemical, biological, and phenotypic properties of drugs. Journal of the American Medical Informatics Association, 2012, 19(e1): e28−e35
64. Mizutani S, Pauwels E, Stoven V, Goto S, Yamanishi Y. Relating drug-protein interaction network with drug side effects. Bioinformatics, 2012, 28(18): i522−i528
65. Perozzi B, Al-Rfou R, Skiena S. DeepWalk: Online learning of social representations. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2014, 701−710
66. Bordes A, Usunier N, Garcia-Durán A, Weston J, Yakhnenko O. Translating embeddings for modeling multi-relational data. In: Proceedings of the 26th International Conference on Neural Information Processing Systems. 2013, 2787−2795
67. Grover A, Leskovec J. node2vec: Scalable feature learning for networks. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016, 855−864
68. Yamanishi Y, Pauwels E, Kotera M. Drug side-effect prediction based on the integration of chemical and biological spaces. Journal of Chemical Information and Modeling, 2012, 52(12): 3284−3292
69. Huang L C, Wu X, Chen J Y. Predicting adverse drug reaction profiles by integrating protein interaction networks with drug structures. Proteomics, 2013, 13(2): 313−324
70. Cheng F, Li W, Wang X, Zhou Y, Wu Z, Shen J, Tang Y. Adverse drug events: database construction and in silico prediction. Journal of Chemical Information and Modeling, 2013, 53(4): 744−752
71. Zhang W, Liu F, Luo L, Zhang J. Predicting drug side effects by multi-label learning and ensemble learning. BMC Bioinformatics, 2015, 16(1): 365
72. Ngufor C, Wojtusiak J, Pathak J. A systematic prediction of adverse drug reactions using pre-clinical drug characteristics and spontaneous reports. In: Proceedings of 2015 International Conference on Healthcare Informatics. 2015, 76−81
73. Niu S Y, Xin M Y, Luo J, Liu M Y, Jiang Z R. DSEP: A tool implementing novel method to predict side effects of drugs. Journal of Computational Biology, 2015, 22(12): 1108−1117
74. Rahmani H, Weiss G, Méndez-Lucio O, Bender A. ARWAR: a network approach for predicting adverse drug reactions. Computers in Biology and Medicine, 2016, 68: 101−108
75. Raja K, Patrick M, Elder J T, Tsoi L C. Machine learning workflow to enhance predictions of adverse drug reactions (ADRs) through drug-gene interactions: application to drugs for cutaneous diseases. Scientific Reports, 2017, 7(1): 3690
76. Bean D M, Wu H, Iqbal E, Dzahini O, Ibrahim Z M, Broadbent M, Stewart R, Dobson R J B. Knowledge graph prediction of unknown adverse drug reactions and validation in electronic health records. Scientific Reports, 2017, 7(1): 16416
77. Dey S, Luo H, Fokoue A, Hu J, Zhang P. Predicting adverse drug reactions through interpretable deep learning framework. BMC Bioinformatics, 2018, 19(21): 476
78. Hu B, Wang H, Wang L, Yuan W. Adverse drug reaction predictions using stacking deep heterogeneous information network embedding approach. Molecules, 2018, 23(12): 3193
79. Ding Y, Tang J, Guo F. Identification of drug-side effect association via semisupervised model and multiple kernel learning. IEEE Journal of Biomedical and Health Informatics, 2019, 23(6): 2619−2632
80. Uner O C, Cinbis R G, Tastan O, Cicek A E. DeepSide: a deep learning framework for drug side effect prediction. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2022, 20(1): 330−339.
81. Wang C S, Lin P J, Cheng C L, Tai S H, Yang Y H K, Chiang J H. Detecting potential adverse drug reactions using a deep neural network model. Journal of Medical Internet Research, 2019, 21(2): e11016
82. Ding Y, Tang J, Guo F. Identification of drug-side effect association via multiple information integration with centered kernel alignment. Neurocomputing, 2019, 325: 211−224
83. Bongini P, Scarselli F, Bianchini M, Dimitri G M, Pancino N, Lió P. Modular multi-source prediction of drug side-effects with DruGNN. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2023, 20(2): 1211−1220
84. Nyamabo A K, Yu H, Liu Z, Shi J Y. Drug-drug interaction prediction with learnable size-adaptive molecular substructures. Briefings in Bioinformatics, 2022, 23(1): bbab441
85. Wishart D S, Feunang Y D, Guo A C, Lo E J, Marcu A, Grant J R, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. DrugBank 5. 0: a major update to the DrugBank database for 2018. Nucleic Acids Research, 2018, 46(D1): D1074−D1082
86. Habib M, Faris M, Alomari A, Faris H. AltibbiVec: a word embedding model for medical and health applications in the Arabic language. IEEE Access, 2021, 9: 133875−133888
87. Galeano D, Li S, Gerstein M, Paccanaro A. Predicting the frequencies of drug side effects. Nature Communications, 2020, 11(1): 4575
88. Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems. Computer, 2009, 42(8): 30−37
89. Ekstrand M D, Riedl J T, Konstan J A. Collaborative filtering recommender systems. Foundations and Trends® in Human-Computer Interaction, 2011, 4(2): 81−173
90. Ma H, Zhou D, Liu C, Lyu M R, King I. Recommender systems with social regularization. In: Proceedings of the 4th ACM International Conference on Web Search and Data Mining. 2011, 287−296
91. Zhao H, Zheng K, Li Y, Wang J. A novel graph attention model for predicting frequencies of drug-side effects from multi-view data. Briefings in Bioinformatics, 2021, 22(6): bbab239
92. Zhao H, Wang S, Zheng K, Zhao Q, Zhu F, Wang J. A similarity-based deep learning approach for determining the frequencies of drug side effects. Briefings in Bioinformatics, 2022, 23(1): bbab449
93. Xu X, Yue L, Li B, Liu Y, Wang Y, Zhang W, Wang L. DSGAT: predicting frequencies of drug side effects by graph attention networks. Briefings in Bioinformatics, 2022, 23(2): bbab586
94. Wang L, Sun C, Xu X, Li J, Zhang W. A neighborhood-regularization method leveraging multiview data for predicting the frequency of drug-side effects. Bioinformatics, 2023, 39(9): btad532
95. Lee D D, Seung H S. Algorithms for non-negative matrix factorization. In: Proceedings of the 13th International Conference on Neural Information Processing Systems, 2000, 535–541
96. Zhao H, Ni P, Zhao Q, Liang X, Ai D, Erhardt S, Wang J, Li Y, Wang J. Identifying the serious clinical outcomes of adverse reactions to drugs by a multi-task deep learning framework. Communications Biology, 2023, 6(1): 870
97. Ryu J Y, Kim H U, Lee S Y. Deep learning improves prediction of drug-drug and drug-food interactions. Proceedings of the National Academy of Sciences of the United States of America, 2018, 115(18): E4304−E4311
98. Hou X, You J, Hu P. Predicting drug-drug interactions using deep neural network. In: Proceedings of the 11th International Conference on Machine Learning and Computing. 2019, 168−172
99. Nyamabo A K, Yu H, Shi J Y. SSI-DDI: substructure-substructure interactions for drug-drug interaction prediction. Briefings in Bioinformatics, 2021, 22(6): bbab133
100. Lyu T, Gao J, Tian L, Li Z, Zhang P, Zhang J. MDNN: A multimodal

deep neural network for predicting drug-drug interaction events. In: Proceedings of the 30th International Joint Conference on Artificial Intelligence. 2021, 3536−3542

101. Chen Y, Ma T, Yang X, Wang J, Song B, Zeng X. MUFFIN: multi-scale feature fusion for drug-drug interaction prediction. Bioinformatics, 2021, 37(17): 2651−2658
102. Yu Y, Huang K, Zhang C, Glass L M, Sun J, Xiao C. SumGNN: multi-typed drug interaction prediction via efficient knowledge graph summarization. Bioinformatics, 2021, 37(18): 2988−2995
103. Deng Y, Qiu Y, Xu X, Liu S, Zhang Z, Zhu S, Zhang W. META-DDIE: predicting drug-drug interaction events with few-shot learning. Briefings in Bioinformatics, 2022, 23(1): bbab514
104. Lin S, Chen W, Chen G, Zhou S, Wei D Q, Xiong Y. MDDI-SCL: predicting multi-type drug-drug interactions via supervised contrastive learning. Journal of Cheminformatics, 2022, 14(1): 81
105. Shao Z, Qian Y, Dou L. TBPM-DDIE: transformer based pretrained method for predicting drug-drug interactions events. In: Proceedings of the 46th IEEE Annual Computers, Software, and Applications Conference (COMPSAC). 2022, 229−234
106. Kang L P, Lin K B, Lu P, Yang F, Chen J P. Multitype drug interaction prediction based on the deep fusion of drug features and topological relationships. PLoS One, 2022, 17(8): e0273764
107. Lin S, Wang Y, Zhang L, Chu Y, Liu Y, Fang Y, Jiang M, Wang Q, Zhao B, Xiong Y, Wei D Q. MDF-SA-DDI: predicting drug-drug interaction events based on multi-source drug fusion, multi-source feature fusion and transformer self-attention mechanism. Briefings in Bioinformatics, 2022, 23(1): bbab421
108. Yu H, Zhao S, Shi J. STNN-DDI: a substructure-aware tensor neural network to predict drug-drug interactions. Briefings in Bioinformatics, 2022, 23(4): bbac209
109. Hong Y, Luo P, Jin S, Liu X. LaGAT: link-aware graph attention network for drug-drug interaction prediction. Bioinformatics, 2022, 38(24): 5406−5412
110. Feng Y H, Zhang S W, Zhang Q Q, Zhang C H, Shi J Y. deepMDDI: A deep graph convolutional network framework for multi-label prediction of drug-drug interactions. Analytical Biochemistry, 2022, 646: 114631
111. Lin J, Wu L, Zhu J, Liang X, Xia Y, Xie S, Qin T, Liu T Y. R^2-DDI: relation-aware feature refinement for drug-drug interaction prediction. Briefings in Bioinformatics, 2023, 24(1): bbac576
112. Li Z, Zhu S, Shao B, Zeng X, Wang T, Liu T Y. DSN-DDI: an accurate and generalized framework for drug-drug interaction prediction by dual-view representation learning. Briefings in Bioinformatics, 2023, 24(1): bbac597
113. Yu H, Li K, Dong W, Song S, Gao C, Shi J. Attention-based cross domain graph neural network for prediction of drug-drug interactions. Briefings in Bioinformatics, 2023, 24(4): bbad155
114. Han C D, Wang C C, Huang L, Chen X. MCFF-MTDDI: multi-channel feature fusion for multi-typed drug-drug interaction prediction. Briefings in Bioinformatics, 2023, 24(4): bbad215
115. Zhu J, Liu Y, Wen C. MTMA: Multi-task multi-attribute learning for the prediction of adverse drug-drug interaction. Knowledge-Based Systems, 2020, 199: 105978
116. Lu J, Yang J, Batra D, Parikh D. Hierarchical question-image co-attention for visual question answering. In: Proceedings of the 30th International Conference on Neural Information Processing Systems. 2016, 289−297
117. Lin X, Quan Z, Wang Z J, Ma T, Zeng X. KGNN: Knowledge graph neural network for drug-drug interaction prediction. In: Proceedings of the 29th International Joint Conference on Artificial Intelligence. 2021, 2739−2745
118. Tang K, Huang J, Zhang H. Long-tailed classification by keeping the good and removing the bad momentum causal effect. In: Proceedings of the 34th International Conference on Neural Information Processing Systems. 2020, 1513−1524
119. Benítez J M, Castro J L, Requena I. Are artificial neural networks black boxes? IEEE Transactions on Neural Networks, 1997, 8(5): 1156−1164

Haochen Zhao received his PhD degree in computer science from Central South University, China. Now He is a Postdoctoral fellow in Computer Science at Central South University, China. His current research interests include bioinformatics and machine learning.

Jian Zhong received his MS degree in computer science from Central South University, China. Now He is a PhD candidate in Computer Science at Central South University, China. His current research interests include bioinformatics and machine learning.

Xiao Liang is a PhD candidate in School of Computer Science and Engineering at Central South University, China. His current research interests include machine learning and its applications in bioinformatics.

Chenliang Xie is an MS student in School of Computer Science and Engineering at Central South University, China. Her current research interests include bioinformatics and machine learning.

Shaokai Wang is a PhD candidate in Cheriton School of Vomputer Science, University of Waterloo, Canada. His current research interests include bioinformatics, proteomics and machine learning.