Deep learning in drug discovery: an integrative review and future challenges

Heba Askr[1] · Enas Elgeldawi[2] · Heba Aboul Ella[4] · Yaseen A. M. M. Elshaier[5] · Mamdouh M. Gomaa[2] · Aboul Ella Hassanien[3]

Abstract

Recently, using artificial intelligence (AI) in drug discovery has received much attention since it significantly shortens the time and cost of developing new drugs. Deep learning (DL)-based approaches are increasingly being used in all stages of drug development as DL technology advances, and drug-related data grows. Therefore, this paper presents a systematic Literature review (SLR) that integrates the recent DL technologies and applications in drug discovery Including, drug–target interactions (DTIs), drug–drug similarity interactions (DDIs), drug sensitivity and responsiveness, and drug-side effect predictions. We present a review of more than 300 articles between 2000 and 2022. The benchmark data sets, the databases, and the evaluation measures are also presented. In addition, this paper provides an overview of how explainable AI (XAI) supports drug discovery problems. The drug dosing optimization and success stories are discussed as well. Finally, digital twining (DT) and open issues are suggested as future research challenges for drug discovery problems. Challenges to be addressed, future research directions are identified, and an extensive bibliography is also included.

Keywords Drug discovery · Artificial intelligence · Deep learning · Drug–target interactions · Drug–drug similarity · Drug side-effects · Drug sensitivity and response · Drug dosing optimization · Explainable artificial intelligence · Digital twining

1 Introduction

The examination of how various drugs interact with the body and how a medication needs to act on the body to have a therapeutic impact is known as drug discovery. Drug discovery strategy constitutes from different approaches as physiology-based and target based. This strategy is based on information about the ligand and the target. In this regard, our attention was directed in certain topics especially drug (ligand)–target interactions, drug sensitivity and response, drug–drug interaction, and drug–drug similarity. For certain diseases such as cancer or pandemic situations as COVID-19, more than one drug combination is required

✉ Aboul Ella Hassanien
aboitcairo@cu.edu.eg

Extended author information available on the last page of the article

Fig. 1 The main building blocks of the paper

to alleviate the prognosis and pathogenesis interactions. Despite all the recent advances in pharmaceuticals, medication development is still a labor-intensive and costly process. As a result, several computational algorithms are proposed to speed up the drug discovery process (Betsabeh and Mansoor 2021).

As DL models progress and the drug data size is getting bigger, a slew of new DL-based approaches is cropping up at every stage of the drug development process (Kim et al. 2021). In addition, we've seen large pharmaceutical corporations migrate toward AI in the wake of the development of DL approaches, eschewing outmoded, ineffective procedures to increase patient profit while also increasing their own (Nag et al. 2022). Despite the DL impressive performance, it remains a critical and challenging task, and there is a chance for researchers to develop several algorithms that improve drug discovery performance. Therefore, this paper presents a SLR that integrates the recent DL technologies and applications in drug discovery. This review study is the first one that incorporates the recent DL models and applications for the different categories of drug discovery problems such as DTIs, DDIs similarity, drug sensitivity and response, and drug-side effects predictions, as well as presenting new challenging topics such as XAI and DT and how they help the advancement of the drug discovery problems. In addition, the paper supports the researchers with the most frequently used datasets in the field.

The paper is developed based on six building blocks as shown in Fig. 1. More than 300 articles are presented in this paper, and they are divided across these building blocks. The papers are selected using the following criteria:

- The papers which published from 2000 to 2022.
- The papers which published in IEEE, ACM, Elsevier, and Springer have more priority.

The following analytical questions are discussed and completely being answered in the paper:

- AQ1: What DL algorithms have been used to predict the different categories of drug discovery problems?
- AQ2: Which deep learning methods are mostly used in drug dosing optimization?
- AQ3: Are there any success stories about drug discovery and DL?
- AQ4: What about the newest technologies such as XAI and DT in drug discovery?
- AQ5: What are the future and open works related to drug discovery and DL?

The remainder of this review paper is organized as: Sect. 2 presents a review of related studies; Sect. 3 covers the various DL techniques as an overview. Section 4 presents the organization of DL applications in drug discovery problems through explaining each drug discovery problem category and gives a literature review of the DL techniques used. Section 5 discusses the numerous benchmark data sets and databases that have been employed in the drug development process. Section 6 presents the evaluation metrics used for each drug discovery problem category. The drug dose optimization, successful stories, and XAI are introduced in Sect. 7, Sect. 8, and Sect. 9. DT and open problems are suggested as future research challenges in Sects. 10 and 11. Section 12 presents a discussion of the analytical questions. Finally, Sect. 13 concludes the paper.

2 Review of related studies

Although the drug discovery is a large field and has different research categories, there is a few review studies about this field and each related study has focused only on a one research category such as reviewing the DL applications for the DTIs. This section aims to review these related studies and a summary is presented in Table 1.

Kim et al. (2021) presented a survey of DL models in the prediction of drug–target interaction (DTI) and new medication development. They start by providing a thorough summary of many depictions of drugs and proteins, DL applications, and widely used exemplary data sets to test and train models. One good point for this study, they identify a few obstacles to the bright future of de novo drug creation and DL-based DTI prediction. However, the major drawback of this study was that it did not consider the latest technology in DL application for the DTIs such as XAI and DTs.

Rifaioglu et al. (2019) presented the recent ML applications in Virtual Screening (VS) with the techniques, instruments, databases, and materials utilized to create the model. They outline what VS is and how crucial it is to the process of finding new drugs. Good points for this study, they highlighted the DL technologies that are accessible as open access programming libraries and provided instances of VS investigations that resulted in the discovery of novel bioactive chemicals and medications, tool kits and frameworks, and can be employed for the foreseeable future's computational drug discovery (including DTI prediction). However, they did not consider the drug dose optimization in their literature review.

Sachdev and Gupta (2019) presented the various feature based chemogenomic methods for DTIs prediction. They offer a thorough review of the different methodologies, datasets, tools, and measurements. They give a current overview of the various feature-based methodologies. Additionally, it describes relevant datasets, methods for determining medication

Table 1 Related studies included DL for drug discovery

Reference	Main context	Year	Category
Kim et al. (2021)	Survey of DL models in drug–target interaction (DTI) and new medication development	2021	DTI
Rifaioglu et al. (2019)	A discussion about the most recent uses of ML techniques in VS (DTI application), including DL	2019	DTI
Sachdev et al. (2019)	Survey of the feature based chemogenomic methods for DTIs prediction	2019	DTI

or target properties, and evaluation measures. Although the study considered the initial integrated review which concentrate only on DTI feature-based techniques, they did not consider the latest technology in DL application for the DTIs such as XAI and DTs.

3 Deep learning (DL) techniques

Detecting spam, recommending videos, classifying images, and retrieving multimedia ideas are just a few of the techniques used are just a few of the applications where machine learning (ML) has lately gained favor in research. Deep learning (DL) is one of the most extensively utilized ML methods in these applications. The ongoing appearance of new DL studies is due to the unpredictability of data acquisition and the incredible progress made in hardware technologies. DL is based on conventional neural networks but outperforms them significantly. Furthermore, DL uses transformations and graph technology to build multi-layer learning models (Kim et al. 2021). With their groundbreaking invention, Machine Learning and Deep Learning have revolutionized the world's perspective. Deep learning approaches have revolutionized the way we tackle problems. Deep learning models come in various shapes and sizes, capable of effectively resolving problems that are too complex for standard approaches to tackle. We'll review the various deep learning models in this section (Sarker 2021).

3.1 Classic neural networks

As shown in Fig. 2, Multi-layer perceptron are frequently employed to recognize Fully Connected Neural Networks. It involves converting the algorithm into simple two-digit data inputs (Mukhamediev et al. 2021). This paradigm allows for both linear and nonlinear functions to be included. The linear function is a single line with a constant multiplier that multiplies its inputs. Sigmoid Curve, Hyperbolic Tangent, and Rectified Linear Unit are three representations for nonlinear functions. This model is best for categorization and regression issues with real-valued data and a flexible model of any kind.

3.2 Convolutional neural networks (CNN)

As shown in Fig. 3, The classic convolutional neural network (CNN) model is an advanced and high-potential variant ANN Which developed to manage escalating complexity levels, as well as data pretreatment and compilation. It is based on how an animal's visual cortex's neurons are arranged (Amashita et al. 2018). One of the most flexible algorithms for the

processing of data with and without images is CNNs. CNN can be processed through 4 phases:

- For analyzing basic visual data, such as picture pixels, it includes one input layer that is often the case a 2D array of neurons.
- Some CNNs analyze images on their inputs using a single-dimensional output layer of neurons coupled to distributed convolutional layers.
- Layer number 3, called as the sampling layer, is included in CNNs o restrict the number of neurons which It took part in the relevant network levels.
- The sampling and output layers are joined by one or more connected layers in CNNs.

This network concept can potentially aid in extracting relevant visual data in pieces or smaller units. In the CNN, the neurons are responsible for the group of neurons from the preceding layer.

After the input data has been included into the convolutional model, the CNN is constructed in four steps:

- Convolution: The method produces feature maps based on supplied data., which are then subjected to a purpose.
- Max-Pooling: It aids CNN in detecting an image based on supplied changes.
- Flattening: The data is flattened in this stage so that a CNN can analyze it.
- Full Connection: It's sometimes referred to as a "hidden layer" which creates the loss function for a model.

Image recognition, image analysis, image segmentation, video analysis, and natural language processing (NLP) (Chauhan et al. 2018; Tajbakhsh et al. May 2016; Mohamed et al. 2020; Zhang et al. 2018) are among the tasks that CNNs are capable of.

3.3 Recurrent neural networks (RNNs)

RNNs were first created to help in sequence prediction. These networks rely solely on data streams with different lengths as inputs. For the most recent forecast, the knowledge of its previous state is used as an input value by the RNN. As a result, it can help a network's short-term memory achievers (Tehseen et al. 2019). As shown in Fig. 4, The Long Short-Term Memory (LSTM) method, for example, is renowned for its adaptability.
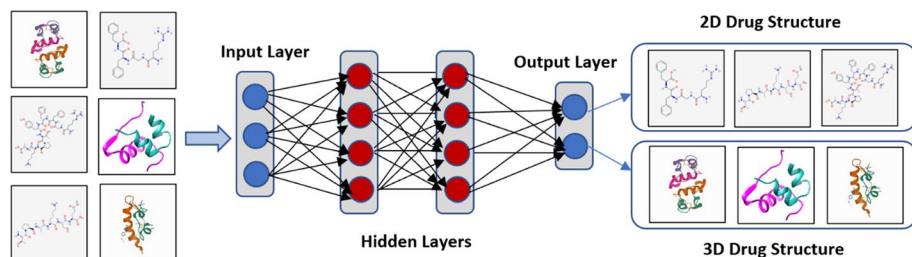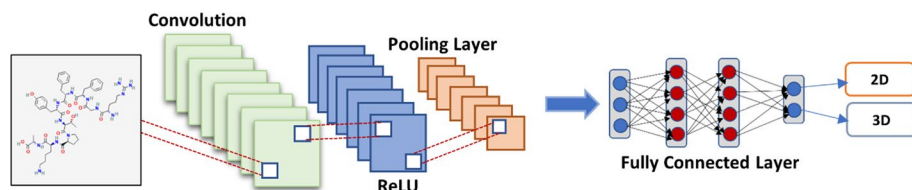

Fig. 2 Multilayer Perceptron or ANN

Fig. 3 Convolutional Neural Networks (CNN)

LSTMs, which are advantageous in predicting data in time sequences using memory, and LSTMs, which are useful in predicting data in time sequences using memory, are two forms of RNN designs that aid in the study of problems. The three gates are Input, Output, and Forget. Gated RNNs are particularly helpful for temporal sequence prediction using memory-based data. Both types of algorithms can be used to address a range of issues, including image classification (Chandra and Sharma 2017), sentiment analysis (Failed 2018), video classification (Abramovich et al. 2018), language translation (Hermanto et al. 2015), and more.

3.4 Generative adversarial networks: GAN

As shown in Fig. 5, It combines a Generator and a Discriminator DL neural network approach. The Discriminator helps to discriminate between real and fake data while the Generator Network creates bogus data (Alankrita et al. 2021).

Both networks compete with one another as The Discriminator still distinguishes between actual and fake data, and the Generator keeps making fake data look like real data. The Generator network will generate simulated data for the authentic photos if a picture library is necessary. Then, a deconvolution neural network would be created. Then, an Image Detector network would be utilized to discriminate between fictitious and real images. This competition would eventually help the network's performance. It can be employed in creating images and texts, enhancing the image and discovering new drugs.

3.5 Self-organizing maps (SOM)

As shown in Fig. 6, Self-Organizing Maps operate by leveraging unsupervised data to decrease a model's number of random variables (Kohonen 1990). Given that every synapse is linked to both its input and output nodes, the output dimension in this DL approach is set as a two-dimensional model. The competition between each data point and its model representation in the Self-Organizing Maps, the weight of the closest nodes or Best Matching Units is adjusted (BMUs). The value of the weights varies based on how close a BMU is. The value represents the node's position in the network because weights are a node attribute in and of themselves. It's great for evaluating dataset frameworks that don't have a Y-axis value or project explorations that don't have a Y-axis value.

3.6 Boltzmann machines

As shown in Fig. 7, the nodes are connected in a circular pattern because there is no set orientation in this network model. This deep learning technique is utilized to generate model
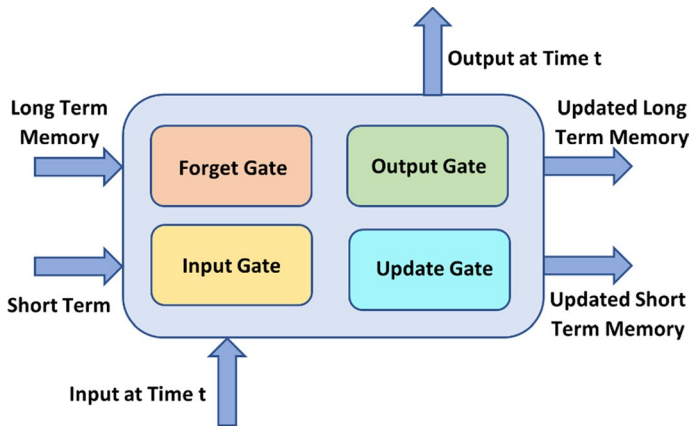
Fig. 4 LSTM Network

parameters because of its uniqueness. The Boltzmann Machines model is stochastic, unlike all preceding deterministic network models. It can monitor systems, create a binary recommendation platform, and analyze specific datasets (Hinton 2011).

The architecture of the Boltzmann Machine is a two-layer neural network. The visible or input layer is the first, while the hidden layer is the second. They are made up of several neuron-like nodes that carry out computations. These nodes are interconnected at different levels but are not linked across nodes in the same layer. As a result, there is no connectivity between layers, which is one of the Boltzmann machine's disadvantages. When data is supplied into these nodes, it is transformed into a graph, and they process it and learn all the parameters, motifs, and relations between them before deciding whether to transmit it. As a result, an Unsupervised DL model is often known as a Boltzmann Machine.
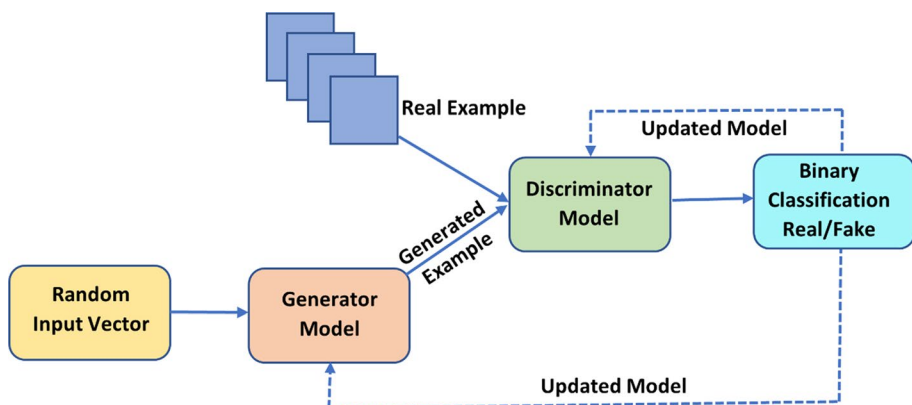

Fig. 5 GAN: *Generative Adversarial Networks*

3.7 Autoencoders

As shown in Fig. 8, This algorithm, one of the most popular deep learning algorithms, automatically based on its inputs, applies an activation function, and decodes the result at the end. Because of the backlog, there are fewer types of data produced, and the built-in data structures are used to their fullest extent (Zhai et al. 2018).

There are various types of autoencoders:

– Sparse: The generalization technique is used when the hidden layers outnumber the input layer to decrease the overfitting. It constrains the loss function and restricts the autoencoder from utilizing all its nodes simultaneously.
– Denoising: In this case, randomly, the inputs are adjusted and made to equal 0.
– Contractive: When the hidden layer outnumbers the input layer, to avoid overfitting and data duplication, a penalty factor is introduced to the loss function.
– Stacked: When another hidden layer is added to an autoencoder, it results in two stages of encoding and Initial stages of decoding.

Feature identification, establishing a strong recommendation model, and adding features to enormous datasets are some of the difficulties it can solve.

4 Organization of DL applications in drug discovery problems

The evolution of safe and effective treatments for human is the primary goal of drug discovery (Kim et al. 2021). Drug discovery is the problem of finding the suitable drugs to treat a disease (i.e., a target protein) which relies on several interactions. This paper divides the drug discovery problems into four main categories, as presented in Fig. 9. They are drug–target interactions, drug–drug similarity, drug combinations side effects, and drug sensitivity and response predictions. The following subsections provide a literature review of DL with these problems and some of the investigated literature articles related to each category are summarized in Table 2.
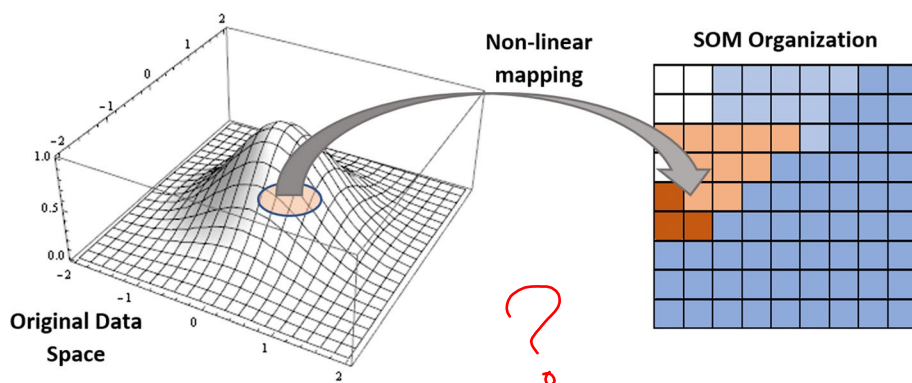

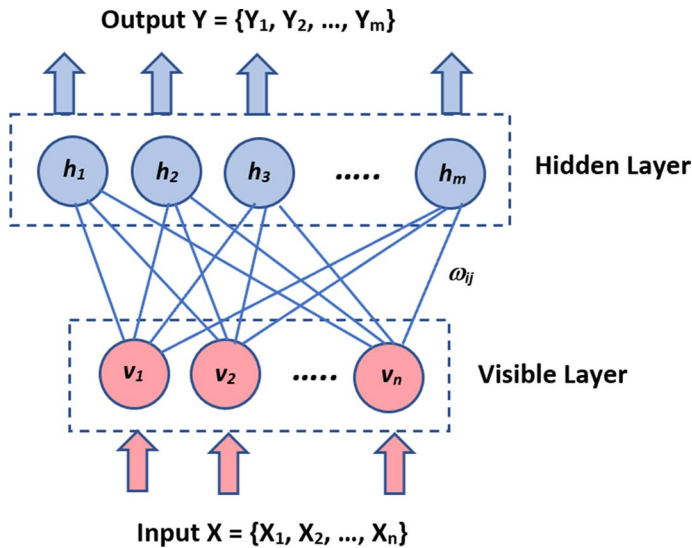**Fig. 6** Self-Organizing Maps (SOM)

Fig. 7 Boltzmann Machines

4.1 Drug–target interactions prediction using DL

Drug repurposing attempts to uncover new uses for drugs that are already on the market and have been approved. It has attracted much attention since it takes less time, costs less money, and has a greater success rate than traditional de novo drug development (Thafar et al. 2022). The discovery of drug–target interactions is the initial step in creating new medications, as well as one of the most crucial aspects of drug screening and drug-guided synthesis (Wang et al. 2020a). Exploring the link between possible medications and targets can aid researchers in better understanding the pathophysiology of targets at the drug level, which can help with the disease's early detection, treatment prognosis, and drug design. This is well known as drug–target interactions (DTIs) (Lian et al. 2021). Achieving success to the drug repositioning mechanism largely reliant on DTI's forecast because it reduces the number of potential medication candidates for specific targets. The approaches based on molecular docking and the approaches based on drugs are the two basic tactics used in traditional computational methods. When target proteins' 3D structures aren't available, the effectiveness of molecular docking is limited. When there are only a few known binding molecules for a target, drug-based techniques typically produce subpar prediction results.
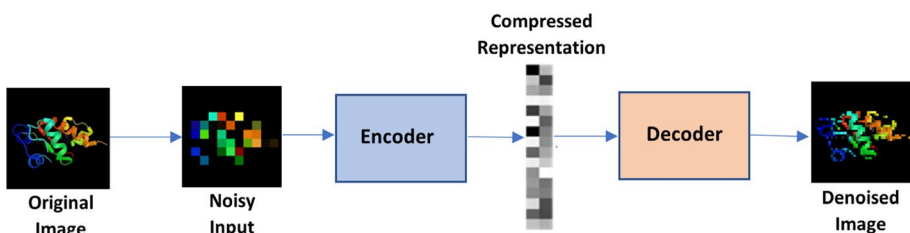

Fig. 8 Autoencoders

DL technologies overcome the restrictions of the high-dimensional structure of drug and target protein by using unstructured-based approaches which do not need 3D structural data or docking for DTI prediction. Therefore, this section provides a recent comprehensive review of DL-based DTIs prediction models (Chen et al. 2012).

As shown in Fig. 10, there are known interactions (solid lines) and unknown interactions (dashed lines) between diseases (proteins) and drugs. DTIs forecast unknown interactions or what diseases (or target proteins) a new drug might treat. According to their input features, we divided the latest DL models used to predict DTIs into three categories: drug-based models, structure (graph)-based models, and drug-protein(disease)-based models.

4.1.1 Drug-based models

Figure 10A shows drug-based models that assume a potential drug will be like known drugs for the target proteins. It calculates the DTI using the target's medication information. Similarity search strategies are used in these models, which postulate that structurally similar substances have similar biological functions (Thafar et al. 2019; Matsuzaka and Uesawa 2019). These methods have been used for decades to select compounds in vast compound libraries employing massive computer jobs or solve problems using human calculations. Deep neural network models gradually narrow the gap between in silico prediction and empirical study, and DL technology can shorten these time-consuming procedures and manual operations.
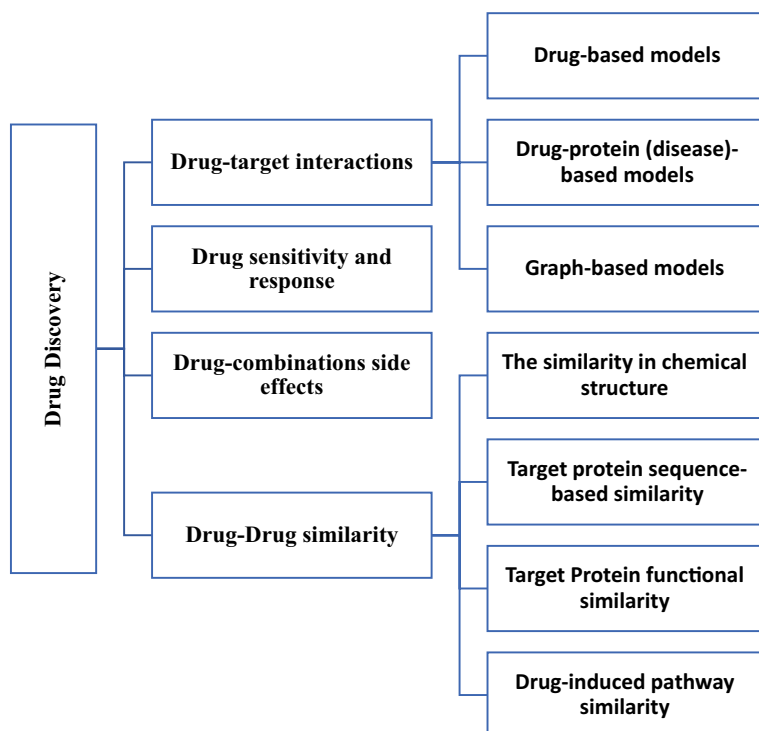

Fig. 9 Drug discovery problem categories

Table 2 Classification of articles related to drug discovery and DL

References	Publishing year	Method	Advantages	Drawbacks
Drug–target interactions (DTIs)				
Yang et al. (2019)	2019	A graph convolutional model that consistently equals or outperforms models with constant molecular descriptors and earlier graph neural designs on both open-source and closed-source data sets	The model performs equally well or better on 12 of the 19 open data sets	The proposed model is underperforming other models when either (1) the 3D information from the other models, (2) the data set is especially small, or (3) the classes are particularly imbalanced
Hirohara et al. (2018)	2020	A brand-new CNN for assessing data about chemical compounds. The CNN processes picture data similarly to traditional CNNs using a SMILES-based feature matrix	The standard fingerprint method used for the virtual screening of chemical compounds performed worse than the CNN based on SMILES string. Additionally, the use of motif detection with learnt filters allowed the detection of undiscovered functional group substructures in addition to significant recognised substructures like protein-binding sites	The model performed better than DNN using simply ECFP in the TOX 21 Challenge 2014 experiment, but somewhat worse than models utilising both ECFP and DeepTox features
Shin et al. (2019)	2019	A novel representation of a molecule employing the self-attention method that has been pre-trained using massive data on chemicals that is available to the public	Regarding four evaluation metrics, the model performs better than all other techniques now in use. Furthermore, the strategy is successful in including all the existing EGFR medications in the top-30 promising prospects, as demonstrated by the example study of discovering drug candidates targeting a cancer protein (EGFR)	The protein was ignored by the model. One explanation is because, on average, a protein sequence is 10 times longer than a molecule sequence, which requires a significant amount of computation time. Another justification is the requirement of a protein dataset with adequate data to pre-train the model
Lee et al. (2019)	2019	A new DTI prediction model by identifying regional residue trends from the complete sequences of the target protein via CNN	Other protein descriptors, such as CTD and SW scores, do not perform as well as the local properties of protein sequences that have been found. Compared to techniques requiring 3D structures, it can be more broadly utilized to predict DTIs	To attain enhanced performance, more detailed chemical characteristics are used for DTI forecasting. Chemical elaboration can be replaced by the consideration of 3D structure information

Table 2 (continued)

References	Publishing year	Method	Advantages	Drawbacks
Wang et al. (2020b)	2020	A novel method MDADTI to predict DTIs based on MDA	MDA lowered the size of the characteristic of medications and objectives, which accelerate the training of MDADTI	Did not consider predicting the binding affinity scores for drug–target pairs
Wan et al. (2019)	2019	DeepCPI is a new and scalable approach that predicts innovative CPIs by combining data-driven representation learning and DL (DTIs)	To effectively utilize the enormous amount of data on the interactions between compounds and proteins accessible from expansive databases, including PubChem and ChEMBL, the synthesis of the efficient feature embedding methodologies with the potent DL model is especially helpful	Better prediction results may be reached by combining other accessible data, like as gene expression and protein structures, into the suggested DL model
Vazquez et al. (2020)	2020	Model for defining hybrid LB (ligand-based) + SB (structure-based) computational schemes in VS (Virtual screening) studies	VS has been a powerful alternative to high-throughput screening assays. The performance of VS campaigns is significantly impacted by the inherent weaknesses of both LB and SB procedures while being the most popular way for exposing innovative hit compounds in the early phases of drug discovery	Although the progress is encouraging, it can be assumed that two primary elements will determine if these integrated tactics are adopted. Initially, a thorough benchmarking of the various combination tactics Secondly, convenient access to screening of focused chemical library searches should be made possible by the capacity to integrate the combined LB and SB techniques into automated modelling platforms
Xie et al. (2018)	2018	A framework that offers better prospects for inferencing and for DTI prediction based on L1000 database transcriptome data from pharmacological perturbation and gene knockdown experiments	The outcomes showed that our system can find DTIs that are more trustworthy than those discovered by previous techniques	investigating a possibility to discover an internal characteristic for assessing fresh DTI potential

Table 2 (continued)

References	Publishing year	Method	Advantages	Drawbacks
Drug sensitivity and response				
Dincer et al. (2018)	2018	A low-dimensional representation (LDR) is encoded using the DeepProfile framework, which learns a variational autoencoder (VAE) network using tens of thousands of publicly available gene expression samples. This network is then used to predict complicated disease symptoms	DeepProfile successfully disentangles data inconsistencies and discovers a useful LDR that precisely foretells complicated phenotypes of many malignancies	The work needs some future directions which include: (1) training DeepProfile using samples from various cancers and (2) expanding the application of Deep-Profile by utilizing multi-omics data to produce more insightful embeddings for cancer
Rampášek et al. (2019)	2019	'Dr. VAE', the first unified machine learning method for the semi-supervised learning for drug response prediction that successfully exploits the prior knowledge of drug-induced transcriptomic perturbations	First, the flexibility of This paradigm made it possible to incorporate the effects of transcriptional disruption into the framework for predicting medication reaction in a special way Second, neural networks with the ability to simulate complex non-linear interactions are employed to parametrize all conditional distributions that make up their model, in addition to variational posteriors employed by Dr. VAE to make a rough inference	The study could be limited in numerous ways. As it has repeatedly been demonstrated to offer the greatest predictive potential in several prior research on treatment response, they first just took the gene expression modality into consideration. Second, they modeled CMap-L1000v1 perturbations following a 6-h treatment period. Since many feedback regulatory processes take 6 h to appear, it may be argued that these studies alone do not give a complete picture of the transcriptome response

Table 2 (continued)

References	Publishing year	Method	Advantages	Drawbacks
Ahmed et al. (2020)	2020	Two neural network models based on graphs for predicting drug sensitivity as well as a network-based feature selection approach	To predict drug sensitivity, using a network to pick features. first identifies additional illustrative characteristics based on the network of gene co-expression. Second, Random Forest performs better than all the other established prediction techniques., Third, compared to DNN, the graph-based neural network models exhibit higher drug response prediction capability., Fourth, the performance of the prediction is reliant on the drug and could be related to the drug's mode of action (MOA)	In this study, there was little improvement for DNNs using graphs. Multi-omics data can be used for analysis to further improve drug prediction accuracy. In addition to the extensive biological features, data acquisition using multi-omics the genomic, epigenomic, and transcriptomic traits of each cell line in the cohort and offer more precise molecular fingerprints to forecast medication reaction than single-omics data alone
Ren et al. (2022)	2022	DeepGRMF, which demonstrated enhanced capability for predicting drug sensitivities	model can predict new responses to drugs or medications not previously seen, which It makes it possible to use already FDA approved medications in new ways to both cure cancer and maybe find new molecules that fight cancer	the DeepGRMF model only uses data from gene expression profiling, incorporating genetic changes further and epigenetic data will probably help the model perform even better
Drug–drug interactions (DDIs) side effect				
Jae et al. (2018)	2018	DeepDDI that accurately forecasts DDI types for medication combinations as well as drug-food component pairings using as inputs names and chemical structures	According to the seven common performance criteria, DeepDDI's feature vector of pharmaceuticals, which combines an improved DNN and SSP, demonstrated great accuracy (84.8–93.2%)	Only two medications are included in the best DDI database provided by Drug-Bank. The DNN can be updated by training DeepDDI for numerous medications and food ingredients once DDI data for multiple drugs and/or food ingredients becoming accessible

Table 2 (continued)

References	Publishing year	Method	Advantages	Drawbacks
Xiang et al. (2020)	2020	An overview of various graph embedding approaches is provided, along with an assessment of how well they perform on link prediction and node categorization, two significant biological tasks	Through comprehensive testing, they discover that current graph embedding techniques can compete favorably with state-of-the-art techniques or even out-perform them in a variety of biological prediction tasks	The idea of pre-training followed by fine-tuning becomes increasingly intriguing as there are more entities that may have pre-trained embeddings as biomedical data volume increases. The relationships between the freshly created network propagation (diffusion) methods and the graph embedding techniques may also have an impact on future research
Xifan et al. (2020)	2020	A multimodal deep learning system called DDIMDL uses deep learning and a variety of pharmacological characteristics to forecast DDI events	The experimental findings demonstrate that DDIMDL performs better than the approaches that were examined in terms of efficiency and accuracy	to enhance the forecast of DDI events. First, there needs to be consideration of unique strategies for dealing with the unbalanced dataset due to the extreme imbalance in the amount of DDIs for various occurrences. Second, certain events have insufficient numbers of interactions, making it easier for the DL approach to underfit, hence techniques to increase the size of the event dataset can be investigated
Karim et al. (2019)	2019	A new ML approach for forecasting DDIs using a variety of sources	The incorporation of drug data from many sources using knowledge graphs. This candidate will get comprehensive background information on medica-tions, illnesses, processes, proteins, enzymes, chemical structures, etc	The approach's inability to offer reasons for the projected DDIs is one of its limitations

Table 2 (continued)

References	Publishing year	Method	Advantages	Drawbacks
Bongini et al. (2022)	2022	Graph Neural Networks (GNNs) are exploited to predict DSEs	A GNN-based predictor aids in foreseeing the incidence of adverse effects. Additionally, its use with novel candidate medications would assist drug discovery studies save time and money while avoiding health problems for the volunteers taking part in clinical trials	An exciting future path is the creation of a prediction based on a GNN that It has the capacity to analyze molecular structural formulae, which are displayed as graphs. Features could be added to these molecular graphs to improve them by way of genes and links between genes and drugs
Iorio et al. (2010)	2010	A broad method to identify previously unknown uses for well-known medications as well as to forecast the molecular actions and MOA of new compounds	Using information buried in the gene expression profiles after drug treatment to identify pharmacological MoA similarities	The approach's primary drawback is the network's small number of chemicals. If a chemical is not like any of the medications in the network, no inferences about its MoA or its biological effects can be made because the approach is predicated on comparing how similar two drugs are
Wang et al. (2019)	2019	An approach using networks to find clinically effective drug combinations for disorders	Provides a network-level analysis of the comparative efficacy and harmful interactions of treatment combinations	The applicability of their findings to different disorders must be investigated in further research
Heba et al. (2021)	2021	The ML-based on similarity for drug interaction prediction (SMDIP) framework, which incorporates popular ML models utilizing internal similarity models based on products to choose the characteristics and benefit from more sparse feature space to enhance DDI prediction efficiency	creating a generalizable machine learning framework based on similarities to reveal innovative DDIs having excellent forecasting abilities; and (2) Showcasing superior predictor elements in comparison to the body of existing research	SMDIP exhibits some restrictions. SMDIP is first and foremost a binary classifier framework with DDI results. It cannot, however, predict how serious a DDI will be. Second, SMDIP cannot make recommendations for the best course of action to be done for the management of DDI since it cannot forecast the unfavorable effects brought on by the interaction

Researchers may now use deep neural networks to analyze medicines and predict drug-related features, including as bioactivities and physicochemical qualities, thanks to using benchmark packages like MoleculeNet (Wu et al. 2018) and DeepChem (). As a result, basic neural networks like MLP and CNN have been used in numerous drug-based DL approaches (Zeng et al. 2020; Yang et al. 2019; Liu et al. 2017). The representation power of molecular descriptors was often the focus of ADMET investigations rather than the model itself (Zhai et al. 2018; Liu et al. 2017; Kim et al. 2016; Tang et al. 2014). Hirohara et al. trained a CNN model with the SMILES string and then used learned attributes to discover motifs using significant structures for locations that bind proteins or unidentified functional groupings (Hirohara et al. 2018). Atom pairs and pharmacophoric donor–acceptor pairings have been employed by Wenzel et al. (2019) as adjectives in multi-task deep neural networks to predict microsomal metabolic liability. Gao et al. (2019) compared 6 different kinds of 2D fingerprints in the prediction of affinity between proteins and drugs using ML methods such as RF, single-task DNN, and multi-task DNN models. Matsuzaka and Uesawa (2019) used 2D pictures of 3D chemical compounds to train a CNN model to predict constitutive androstane receptor agonists. They optimized the greatest performance in snapshots of a 3D ball-and-stick model taken at various angles or coordinates. Therefore, the method outperformed seven common 3D chemical structure forecasts.

Since the GCN's development, drug related GCN models have created depictions of graphs which concerned with molecules that incorporate details on the chemical structures by adding up the adjacent atoms' properties (Gilmer et al. 2017).

GCNs have been employed as 3D descriptors instead of SMILES strings in a lot of research, and it's been discovered that these learned descriptors outperform standard descriptors in prediction tests and are easier to understand (Shin et al. 2019; Ozturk et al. 2018; Yu et al. 2019). Chemi-net employed GCN models to represent molecules and compared the performance of single-task and multi-task DNNs on their own QSAR datasets
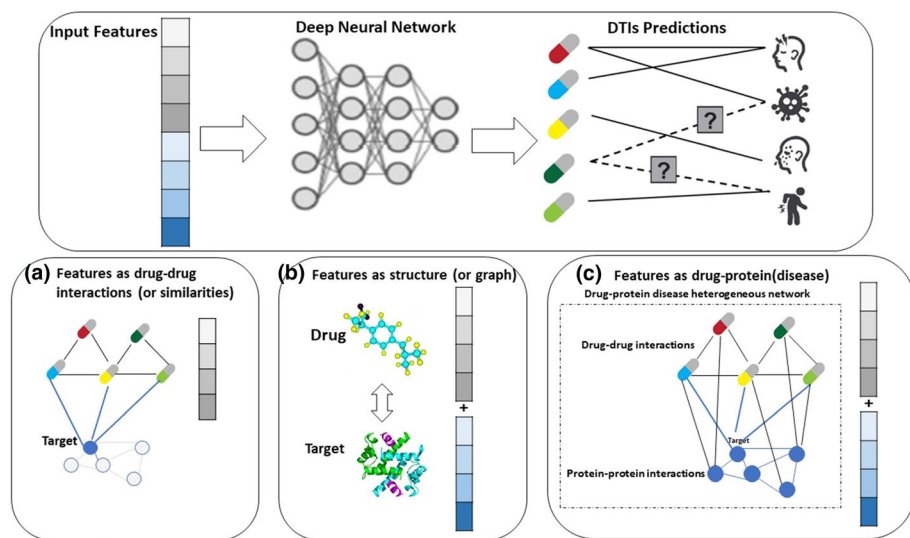

Fig. 10 DL models used for predicting the DTIs are grouped into three categories: **a** drug-based models, **b** structure (graph)-based models, and **c** drug-protein(disease)-based models

(Liu et al. 2019a). Yang et al. (2019) introduced the directed message passing neural network, which uses a directed message-passing paradigm, as a more advanced model (D-MPNN). They tested their approaches on 19 publicly available and 16 privately held datasets and discovered that in most situations, they were correct. The D-MPNN models outperformed the previous models. In two datasets, they underperformed and were not as resilient as typical 3D descriptors when the sample was small or unbalanced. The D-MPNN model was then employed by another research group to correctly forecast a kind of antibiotic named HALICIN, which demonstrated bactericide effects in models for mice (Stokes et al. 2020). This was the first incident that resulted in the finding of an antibiotic by using DL methods to explore a large-scale chemical space that current experimental methodologies cannot afford. The application of attention-based graph neural networks is another interesting contemporary method (Sun et al. 2020a). Edge weights and node features can be learned together since a molecule's graph representations can be altered by edge properties. As a result, Shang et al. suggested a multi-relational GCN with edge attention (Shang et al. 2018). For each edge, they created a reference guide on attention spans. Because it is used throughout the molecule, the approach can handle a wide range of input sizes.

In the Tox21 and HIV benchmark datasets, they found that this model performed better than the random forest model. As a result, the model may effectively learn pre-aligned features from the molecular graph's inherent qualities. Withnall et al. (2020) extended the MPNN model with AMPNN (attention MPNN), which is an attention technique that the message forwarding step employs weighted summation. Moreover, they termed the D-MPNN model the edge memory neural network because it was extended by the same attention mechanism as the AMPNN (EMNN). Although it is computationally more intensive than other models, this model fared better than others on the uniformly absent information from the maximal unbiased validation (MUV) reference.

4.1.2 Structure (graph)-based models

Unlike the drug- and structure-based models in Fig. 10b, protein targets and medication information should be included. Typical molecular docking simulation methods aim to predict the geometrically possible binding of known tertiary structure drugs and proteins. Atom sequences and amino acid residues can be used to express both the medicine as well as the target. Descriptors based on sequences were selected because DL approaches may be implemented right away with non-significant pre-processing of the entering data.

The Davis kinase binding affinity dataset (Davis et al. 2011) and the KIBA dataset (Sun et al. 2020a) were used in that study. DeepDTA, suggested by Ozturk et al. (2018), outperformed moderate ML approaches such as KronRLS (Nascimento et al. 2016) and SimBoosts (Tong et al. 2017) by applying solely information about the sequence of a CNN model based on the SMILES string and amino acid sequences. Wen et al. used ECFPs and protein sequence composition descriptors as examples of common and basic features and trained them using semi-supervised learning via a deep belief network (Wen et al. 2017). Another study, DeepConv-DTI, built a deep CNN model using only an RDKit Morgan fingerprint and protein sequences (Lee et al. 2019). They also used the pooled convolution findings to capture local residue patterns of target protein sequences, resulting in high values for critical protein areas like actual binding sites.

The scoring feature, which ranks the protein-drug interaction with 3D structures and makes the training data parametric to forecast values for binding affinities of targeted

proteins, is used to predict binding affinity values or binding pocket sites of the target proteins as a key metric for the structure-based regression model. The protein–drug complexes' 3D structural characteristics were included in the CNNs by AtomNet (Wallach et al. 2015). They placed 3D grids with set sizes (i.e., voxels) in comparison to protein–drug combinations, with every cell in the grid representing structural properties at that position. Several researchers have examined the situation since then, deep CNN models that use voxels to predict binding pocket location or binding affinity (Wang et al. 2020b; Ashburner et al. 2000; Zhao et al. 2019). In comparison to common docking approaches such as AutoDock Vina (Trott and Olson 2010) or Smina (Koes et al. 2013), these models have shown enhanced performance. This is since CNN models are relatively impervious even with large input sizes. It can be taught and is resilient to input data noise.

Many DTI investigations using GCNs based on structure-based approaches have been reported (Feng et al. 2018; Liu et al. 2016). Feng et al. (2018) used both ECFPs and GCNs as pharmacological characteristics. In the Davis et al. (2011), Metz et al. (2011), and KIBA Tang et al. (2014) benchmark datasets, their methods outperformed prior models such as KronRLS (Nascimento et al. 2016) and SimBoost (Tong et al. 2017). However, they did agree that their GCN model couldn't beat their ECFP model due to time and resource constraints in implementing the GCN. In a different DTI investigation study, Torng et al. employed a graph without supervision to become familiar with constant size depictions of protein binding sites (Torng and Altman 2019). The pre-trained GCN model was then trained using the newly created protein pocket GCN, the drug GCN model, on the other hand, used attributes to be trained and which were generated automatically. They concluded that without relying on target–drug complexes, their model effectively captured protein–drug binding interactions.

Because the models that implement the attention mechanism have key qualities that enable the model to be interpreted, attention-based DTI prediction approaches have evolved (Hirohara et al. 2018; Liu et al. 2016; Perozzi et al. 2014).

For protein sequences, Gao et al. (2017) employed compressed vectors with the LSTM RNNs and the GCN for drug structures. They concentrated on demonstrating their method's capacity to deliver biological insights into DTI predictions. To do so, Mechanisms for two-way attention were employed. to calculate the binding of drug–target pairs (DTPs), allowing for flexible interpretation of superior data from target proteins, such as GO keywords. Shin et al. (2019) introduced the Molecule transformer DTI (MT-DTI) approach for drug representations, which uses the self-attention mechanism. The MT-DTI model was tweaked to perfection and assessed using two Davis models Using pre-trained parameters from the 97 million chemicals PubChem (Davis et al. 2011) and (KIBA) (Tang et al. 2014) benchmark datasets, which are both publicly available. However, the attention mechanism was not used to depict the protein targets because it would take too long to calculate the target sequence in an acceptable amount of time. Pre-training is impossible due to a lack of target information.

On the other hand, attention DTA presented by Zhao et al. incorporates a CNN attention mechanism model to establish the weighted connections between drug and protein sequences (Zhao et al. 2019). They showed that these attention-based drug and protein representations have good MLP model affinity prediction task performance. DeepDTIs used external, experimental DTPs to infer the probability of interaction for any given DTP. Four of the top ten predicted DTIs have previously been identified, and one was discovered to have a poor glucocorticoid receptor binding affinity (Huang et al. 2018). DeepCPI was used to predict drug–target interactions. Small-molecule interactions with the glucagon-like

peptide one receptor, the glucagon receptor, and the vasoactive intestinal peptide receptor have been tested in experiments (Wan et al. 2019).

4.1.3 Drug–protein(disease)-based models

According to poly pharmacology, most medicines have multiple effects on both primary and secondary targets. The biological networks involved, as well as the drug's dose, influence these effects. As a result, the drug–protein(disease)-based models shown in Fig. 10c are particularly beneficial when evaluating protein promiscuity or drug selectivity (Cortes-Ciriano et al. 2015). Furthermore, Neural networks that can do multiple tasks are ideal for simultaneously learning the properties of many sorts of data (Camacho et al. 2018). Several DL model applications, such as drug-induced gene-expression patterns and DTI-related heterogeneous networks, leverage relational information for distinct views. A network-based strategy employs heterogeneous networks includes a variety of nodes and edges kinds (Luo et al. 2017; David et al. 2019). The nodes in these networks have a local similarity, which is a significant aspect of these models. One can anticipate DTIs using their connections and topological features when a network of similarity with medications as its nodes and drug–drug similarity values as a measure of the edges' weights is investigated. Machine to support vectors (Bleakley and Yamanishi 2009; Keum and Nam 2017), Machine learning techniques that use heterogeneous networks as prediction frameworks include the regularized least square model (RLS) (Liu et al. 2016; Xia et al. 2010; Hao et al. 2016) and random walk with the restart model Nascimento (Lian et al. 2021; Nascimento et al. 2016). DTI prediction research using networks have employed DL to enhance the methods used to forecast associations today for evaluating the comparable topological structures of drug and target networks that are bipartite and tripartite linked networks, owing to the increased interest in the usage of DL technologies (drug, target, and disease networks) (Hassan-Harrirou et al. 2020; Lamb et al. 2006; Korkmaz 2020; Townshend et al. 2012; Vazquez et al. 2020). Zong et al. (2017) used the DeepWalk approach to collect local latent data, compute topology-based similarity in tripartite networks, and demonstrate the technology's promise as a medication repurposing solution.

Relationship-based features collected by training the AE were used in some network-based DTI prediction studies. Zhao et al. (2020) developed a DTI-CNN prediction model that combined depth information that is low-dimensional but rich with a heterogeneous network that has been taught using the stacked AE technique. To construct the topological similarity matrix of drug and target, Wang et al. used a deep AE and mutually beneficial pointwise information in their analysis (Wang et al. 2020b). Peng et al. (2020) employed a denoising Autoencoder to pick network-based attributes and decrease the representation dimensions in another investigation.

By helping the self-encoder learn to denoise, the anti-aliasing effect (Autoencoder) enhances high-dimensional images with noise, input data that is noisy and incomplete, allowing the encoder to learn more reliably. These approaches, however, have a drawback in that it is challenging to foresee recent medications or targets, a problem. The problem of recommendation systems' "cold start" is known as the "cold start" problem (Bedi et al. 2015). The size and form of the network have a big impact on these models, so if the network isn't big enough, they will not be able to collect all the medications or targets that aren't in the network (Lamb et al. 2006).

Various investigations have also utilized Gene expression patterns as chemogenomic traits to predict DTIs. This research presumes that medications with similar expression patterns have similar effects on the same targets (Hizukuri et al. 2015; Sawada et al. 2018).

The revised version of CMAP, the LINCS-L1000 database, has been integrated into the DL DTI models in recent works (Subramanian et al. 2017; Thafar et al. 2020; Karpov et al. 2020; Arus-Pous et al. 2020). Based on the LINCS pharmacological perturbation and knockout gene data, using a deep neural network, Xie et al. developed a binary classification model (Xie et al. 2018).

On the other hand, Lee and Kim employed as a source of expression signature genes medication and target features. They used node2vec to train the rich data by examining three elements of protein function, including pathway-level memberships and PPI (Lee and Kim 2019). Saho and Zhang employed a GCN model to extract drug and target attributes from LINCS data and a CNN model to forecast DTPs by extracting latent features in DTIGCCN (Shao et al. 2020). The Gaussian kernel function was identified to aid in the production of high-quality graphs, and as a result, this hybrid model scored better on classification tests.

DeepDTnet employs a heterogeneous drug–gene-disease network to uncover known drug targets containing fifteen types of chemicals and genomic, phenotypic, and cellular network properties. DeepDTnet predicted and experimentally confirmed topotecan, a new direct inhibitor of the orphan receptor linked to the human retinoic acid receptor (Zeng et al. 2020).

4.2 Drug sensitivity and response prediction using DL

Drug response is the clinical outcome treated by the drug of interest (https://www.scien cedirect.com/topics/drug-response). This is due to the normally low ratio of samples to measurements each sample, which makes traditional feedforward neural networks unsuitable. The main idea of drug response prediction is shown in Fig. 11. The DL method takes the heterogenous network of drug and protein interactions as inputs and predicts the response scores. Although the widespread use of the deep neural network (DNN) approaches in various domains and sectors, including related topics like computational chemistry (Gómez-Bombarelli et al. 2018), DNNs have only lately made their way into drug response prediction. Overparameterization, overfitting, and poor generalization are common outcomes of recent simulation datasets. However, more public data has become available recently, and freshly built DNN models have shown promise. As a result, this section summarizes current DL computational problems and drug response prediction breakthroughs.

Since the 1990s, neural networks have been used to predict drug response (El-Deredy et al. 1997) revealed that data from tumor nuclear magnetic resonance (NMR) spectra might be used to train a neural network and can be utilized to predict drug response in gliomas and offer information on the metabolic pathways involved in drug response.

In 2018, The DRscan model was created by Chang et al. (2018), and it uses a CNN architecture that was trained on 1000 drug reaction studies per molecule. Compared to other traditional ML algorithms like RF and SVM, their model performed much better. CDRscan's ability to incorporate genomic data and molecular fingerprints is one of the reasons it outperformed these baseline models. Furthermore, its convolutional design has been demonstrated to be useful in various machine learning areas. A neural network called an autoencoder attempts to recreate the original data from the compressed form after

compressing its input. As proven by Way and Greene (2018), this is very useful for feature extraction, which condensed a gene expression profile with 5000 dimensions with a maximum of 100 dimensions, some of which revealed to significant characteristics such as the patient's sexual orientation or melanoma status. Using variational autoencoders, Dincer et al. (2018) created DeepProfile, a technique for learning a depiction of gene expression in AML patients in eight dimensions that is then fitted to a Lasso linear model for treatment response prediction with superior results to that of no extracting features.

Ding et al. (2018) proposed a deep autoencoder model for representation learning of cancer cells from input data consisting of gene expression, CNV, and somatic mutations.

In 2019, MOLI (Multi-omics Late Integration) (Sharifi-Noghabi et al. 2019) was a deep learning model that incorporates multi-omics data and somatic mutations to characterize a cell line. Three separate subnetworks of MOLI learn representations for each type of omics data. A final network identifies a cell's response as responder or non-responder based on concatenated attributes. Those methods share two characteristics: integrating multiple input data (multi-omics) and binary classification of the drug response. Although combining several forms of omics data can improve the learning of cell line status, it may limit the method's applicability for testing on different cell lines or patients because the model requires extra data beyond gene expression.

Furthermore, a certain threshold of the IC50 values should be set before binary classification of the drug response, which may vary depending on the experimental condition, such as drug or tumor types. Twin CNN for drugs in SMILES format (TCNNS) (Liu et al. 2019b) takes a one-hot encoded representation of drugs and feature vectors of cell lines as the inputs for two encoding subnetworks of a One-Dimensional (1D) CNN. One-hot encodings of drugs in TCNNS are Simplified Molecular Input Line Entry System (SMILES) strings which describe a drug compound's chemical composition. Binary feature vectors of cell lines represent 735 mutation states or CNVs of a cell. KekuleScope (Cortés-Ciriano and Bender 2019) adopts transfer learning, using a pre-trained CNN on ImageNet data. The pre-trained CNN is trained with images of drug compounds represented as Kekulé structures to predict the drug response.

Yuan et al. (2019) offer GNNDR, a GNN-based technique with a high learning capacity and allows drug response prediction by combining protein–protein interactions (PPI)
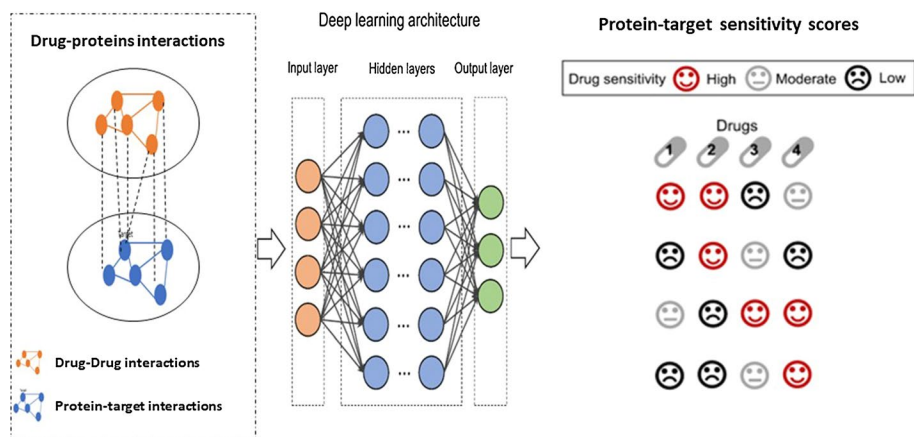

Fig. 11 Drug binding with proteins and drug sensitivity (response) scores prediction

information with genomic characteristics. The value of including protein information has been empirically proven. The proposed method offers a viable avenue for the discovery of anti-cancer medicines. Semi-supervised variational autoencoders for the prediction of monotherapy response were examined by the Rampášek et al. (2019). In contrast to many conventional ML methodologies, together developed a model for predicting medication reaction that took advantage of expression of genes before and after therapy in cell lines and demonstrated enhanced evaluation on a variety of FDA-approved pharmaceuticals. Chiu et al. (2019) trained a deep drug response predictor after pre-training autoencoders using mutation data and expression features from the TCGA dataset. The use of pretraining distinguishes their strategy from others. Compared to using only the labeled data, the pretraining process permits un-labelled data from outside sources, like TCGA, as opposed to just gene expression profiles obtained from drug reaction tests, resulting in a significant increase in the number of samples available and improved performance.

Chiu et al. (2019) and Li et al. (2019) used a combination of auto-encoders and predicted drug reactions in cell lines with deep neural networks and malignancies that had been gnomically characterized. To anticipate cell lines reactions to drug combinations, in https://string-db.org/cgi/download.pl?sessionId=uKr0odAK9hPs used deep neural encoders to link genetic characteristics with drug profiles.

In 2020, Wei et al. (2020) anticipate drug risk levels (ADRs) based on adverse drug reactions. They use SMOTE and machine learning techniques in their studies. The proposed framework was used to investigate the mechanism of ADRs to estimate degrees of drug risk and to assist with and direct decision-making during the changeover from prescription to over-the-counter medications. They demonstrated that the best combination, PRR-SMOTE-RF, was built using the above architecture and that the macro-ROC curve had a strong classification prediction effect. They suggested that this framework could be used by several drug regulatory organizations, including the FDA and CFDA, to provide a simple but dependable method for ADR signal detection and drug classification, as well as an auxiliary judgement basis for experts deciding on the status change of Rx drugs to OTC drugs. They propose that more ML or DL categorization algorithms be tested in the future and that computational complexity be factored into the comparison process. Kuenzi et al. (2020) built DrugCell, an interpretable DL algorithm of personal cancer cells based on the reactions of 1235 tumor cell lines to 684 drugs. Genotypes of cancer cause conditions in cellular systems combined with medication composition to forecast therapeutic outcome while also learning the molecular mechanisms underlying the response. Predictions made by DrugCell in cell lines are precise and help to categorize clinical outcomes. The study of DrugCell processes results in the development of medication combinations with synergistic effects, which we test using combinatorial CRISPR, in vitro drug–drug screening, and xenografts generated from patients. DrugCell is a step-by-step guide to building interpretable predictive medicine models.

Artificial Neural Networks (ANNs) that operate on graphs as inputs are known as Graph Neural Networks (GNNs). Deep GNNs were recently employed for learning representations of low-dimensional biomolecular networks (Hamilton 2020; Wu et al. 2020). Ahmed et al. (2020) used two separate GNN methods to develop a GNN using GE and a network of genes that are expressed together. This is a network that depicts the relationship between gene pairs' expression.

The CNN is one of the neural network models adopted for drug response prediction. The CNN has been actively used for image, video, text, and sound data due to its strong ability to preserve the local structure of data and learn hierarchies of features. In 2021, several methods had been developed for drug response prediction, each of which utilizes different input data for prediction (Baptista et al. 2021).

Nguyen et al. (2021) proposed a method to predict drug response called GraphDRP, which integrates two subnetworks for drug and cell line features, like CNN in Liu et al. (2019b) and Qiu et al. (2021). Gene expression data from cancer cell lines and medication response data, the author finds predictor genes for medications of interest and provides a reliable and accurate drug response prediction model. Using the Pearson correlation coefficient, they employed the ElasticNet regression model to predict drug response and fine-tune gene selection after pre-selecting genes. They ran a regression on each drug twice, once using the IC50 and once with the area under the curve (AUC), to obtain a more trustworthy collection of predictor genes (or activity area). The Pearson correlation coefficient for each of the 12 medicines they examined was greater than 0.6. With 17-AAG, IC50 has the highest Pearson correlation coefficient of 0.811.

In contrast, AUC has the highest Pearson correlation coefficient of 0.81. Even though the model developed in this study has excellent predictive performance for GDSC, it still has certain flaws. First, the cancer cell line's properties may differ significantly from those of in vivo malignancies, and it must be determined whether this will be advantageous in a clinical trial. Second, they primarily use gene expression data to predict drug response. While drug response is influenced by structural changes such as gene mutations, it is also influenced by gene expression levels. To improve the prediction capacity of the model, more research is needed to use such data and integrate it into the model.

In 2022, Ren et al. (2022) suggested a graph regularized matrix factorization based on deep learning (DeepGRMF), which uses a variety of information, including information on drug chemical composition, their effects on cell biology signaling mechanisms, and the conditions of cancer cells, to integrate neural networks, graph models, and matrix-factorization approaches to forecast cell response to medications. DeepGRMF trains drug embeddings so that drugs in the embedding space with similar structures and action mechanisms, (MOAs) are intimately linked. DeepGRMF learns the same representation embeddings for cells, allowing cells with similar biological states and pharmacological reactions to be linked. The Cancer Cell Line Encyclopedia (CCLE) and On the Genomics of Drug Sensitivity in Cancer (GDSC) datasets, DeepGRMF outperforms competing models in prediction performance. In the Cancer Genome Atlas (TCGA) dataset, the suggested model might anticipate the effectiveness of a treatment plan on lung cancer patients' outcomes. The limited expressiveness of our VAE-based chemical structure representation may explain why new cell line prediction outperforms innovative drug sensitivity prediction in terms of accuracy. A family of neural graph networks has recently been shown to depict better chemical structures that can be investigated in the future. Pouryahya et al. (2022) proposed a new network-based clustering approach for predicting medication response based on OMT theory. Gene-expression profiles and cheminformatic drug characteristics were used to cluster cell lines and medicines, and data networks were used to represent the data. Then, RF model was used regarding each pair of cell-line drug clusters. by comparison, prediction-clustered based models regarding the homogenous data are anticipated to enhance drug sensitivity and precise forecasting and biological interpretability.

4.3 Drug–drug interactions (DDIs) side effect prediction using DL

Drugs are chemical compounds consumed by people and interact with protein targets to create a change. The drugs may alter the human body positively or negatively. Drug side effects are the undesirable alterations medications cause in the human body. These adverse

effects might range from moderate headaches to life-threatening reactions like cardiac arrest, malignancy, and death. They differ depending on the person's age, gender, stage of sickness, and other factors (Kuijper et al. 2019). In the laboratory, to determine whether the medications have any unfavorable side effects, several tests are conducted on them. However, these examinations are both pricey and additionally lengthy. Recently, many computational algorithms for detecting medication adverse effects have been created. Computational methodologies are replacing laboratory experiments.

On the other hand, these methods do not provide adequate data to predict drug–drug interactions (DDIs). The phenomenon of DDIs is discussed in Fig. 12. The desired effects of a drug resulting from its interaction with the intended target and the unfavorable repercussions emerging from drug interactions with off targets make up a drug's entire reaction on the human body (undesirable effects). Even though A medication has a strong affinity for binding to one target, it binds to several proteins as well with varied affinities, which might cause adverse consequences (Liu et al. 2021). Predicting DDIs can assist in reducing the likelihood of adverse reactions and optimizing the medication development and post-market monitoring processes (Arshed et al. 2022). Side effects of DDIs are often regarded as the leading cause of drug failure in pharmacological development. When drugs have major side effects, the market is quickly removed from them. As a result, predicting side effects is a fundamental requirement in the drug discovery process to keep drug development costs and timelines in check and launch a beneficial drug in terms of patient health recovery.

Furthermore, the average drug research and development cost is $2.6 billion (Liu et al. 2019). As a result, determining the possibility of negative consequences is important for lowering the expense and risk of medication development. The researchers use various computer tools to speed up the process. In pharmacology and clinical application, DDI prediction is a difficult topic, and correctly detecting possible DDIs in clinical studies is crucial for patients and the public. Researchers have recently produced a series of successes utilizing deep learning as an AI technique to predict DDIs by using drug structural properties and graph theory (Han et al. 2022). AI successfully detected potential drug interactions, allowing doctors to make informed decisions before prescribing prescription combinations to patients with complex or numerous conditions (Fokoue et al. 2016).

Therefore, this section comprehensively reviews the researchers' most popular DL algorithms to predict DDIs.

In 2016, Tiresias is a framework proposed by Achille Fokoue et al. (2017) for discovering DDIs. The Tiresias framework uses a large amount of drug-related data as input to generate DDI predictions. The detection of the DDI approach begins using input data that has been semantically integrated, resulting in a knowledge network that represents drug properties and interactions using additional components like enzymes, chemical structures, and routes. Numerous similarity metrics between all pharmacological categories were determined using a knowledge graph in a scalable and distributed setting. To forecast the DDIs, a large-scale logistic regression prediction model employs calculated similarity metrics. According to the findings, the Tiresias framework was proven to help identify new interactions between currently available medications and freshly designed and existing drugs. The suggested Tiresias model's necessity for big, scaled medication information was negative, resulting in the developed model's high cost.

In 2017, Reza et al. (2017) developed a computational technique for predicting DDIs based on functional similarities among all medicines. Several major biological aspects were used to create the suggested model: carriers, enzymes, transporters, and targets (CETT). The suggested approach was implemented on 2189 approved medications, for
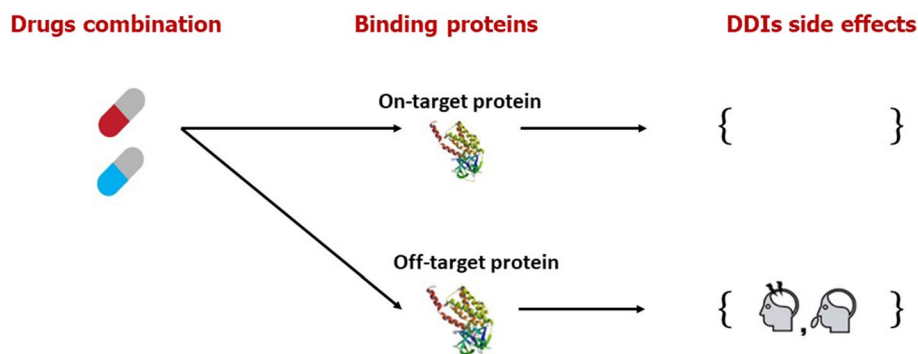
Fig. 12 Drug binding with proteins and DDI side effects

which the associated CETTs were obtained, and binary vectors to find the DDIs were created. Two million three hundred ninety-four thousand seven hundred sixty-seven potential drug–drug interactions were assessed, with over 250,000 unidentified possible DDIs discovered. Inner product-based similarity measures (IPSMs) offered good values predicted for detecting DDIs among the several similarity measures used. The lack of pharmacological data was a key flaw in this strategy, which resulted in the erroneous detection of all potential pairs of DDIs.

In 2018, Ryu et al. (2018) proposed a model that predicts more DDI kinds using the drug's chemical structures as inputs and applied multi-task learning to DDI type prediction in the same vein Decagon (Zitnik et al. 2018) models polypharmacy side effects using a relational GNN. To comprehend the representations of intricate nonlinear pharmacological interactions, Chu et al. (2018) utilized an auto-encoder for factoring. To predict DDIs, Liu et al. (2019c) presented the DDI-MDAE based on shared latent representation, a multimodal deep auto-encoder. Recently, interest in employing graph neural networks (GNNs) to forecast DDI has increased. Distinct aggregation algorithms lead to different versions of GNNs to efficiently assemble the vectors of its neighbors' feature vectors (Asada et al. 2018) uses a convolutional graph network (GCN) to encode the molecular structures to extract DDIs from text. Furthermore, Ma et al. (2018) has incorporated attentive Multiview graph auto-encoders into a coherent model.

Chen (2018) devised a model for predicting Adverse Drug Reactions (ADR). SVM, LR, RF, and GBT were all used in the predictive model. The DEMO dataset, which contains properties such as the patient's age, weight, and sex, and the DRUG dataset, which includes features such as the drug's name, role, and dosage, were employed in this model. Males make up 46% of the sample, while females make up 54%. The developed model had a fair forecasting accuracy for a representative sample set. Furthermore, the outputs revealed that the suggested model is only accurate for a significant number of datasets.

To anticipate the possible DDI, Kastrin et al. (2018) employed statistical learning approaches. The DDI was depicted as a complex network, with nodes representing medications and links representing their potential interactions. On networks of DDIs, the procedure for predicting links was represented as a binary classification job. A big DDI database was picked randomly to forecast. Several supervised and unsupervised ML approaches, such as SVM, classification tree, boosting, and RF, are applied for edge prediction in various DDIs. Compared to unsupervised techniques, the supervised link prediction strategy

generated encouraging results. To detect the link between the pharmaceuticals, The proposed method necessitates Unified Medical Language System (UMLS) filtering, which provided a dilemma for the scientists. Furthermore, the suggested system only considers fixed network snapshots, which is problematic for DDI's system because It's a fluid system.

In 2019, Lee et al. (2019) proposed a deep learning system for accurately forecasting the results of DDIs. To learn more about the pharmacological effects of a variety of DDIs, an assortment of auto-encoders and a deep feed-forward neural network was employed in the suggested method that were honed utilizing a mix of well-known techniques. The results revealed that using SSP alone improves GSP and TSP prediction accuracy, and the autoencoder is more powerful than PCA at reducing profile features. In addition, the model outperformed existing approaches and included numerous novel DDIs relevant to the current study Yue et al. (2020) combines numerous graphs embedding methods for the DDI job, while models DDI as link prediction with the help of a knowledge graph (Karim et al. 2019). There's also a system for co-attention (Andreea and Huang 2019), which presented a deep learning model based solely on side-effect data and molecular drug structure. CASTER in Huang et al. (2020) also based on drug chemical structures, develops a framework for dictionary learning to anticipate DDIs (Chu et al. 2019) and proposes using semi-supervised learning to extract meaningful information for DDI prediction in both labeled and unlabeled drug data. Shtar et al. (2019) used a mix of computational techniques to predict medication interactions, including artificial neural networks and graph node factor propagation methods such as adjacency matrix factorization (AMF) and adjacency matrix factorization with propagation (AMFP). The Drug-bank database was used to train the model, containing 1142 medications and 45,297 drug drugs. With 1442 drugs and 248,146 drug–drug interactions, the trained model was tested from the drug bank's most recent version. AMF and AMFP were also used to develop an ensemble-based classifier, and the outcomes were assessed using the receiver operating characteristic (ROC) curve. The findings revealed that the suggested a classifier that uses an ensemble delivers important drug development data and noisy data for drug prescription. In addition, drug embedding, which was developed during the training of models utilizing interaction networks, has been made available. To anticipate adverse drug events caused by DDIs, Hou et al. (2019) suggested a deep neural network architecture model. The suggested model is based on a database of 5000 medication codes obtained from Drug Bank. Using the computed features, it discovers 80 different types of DDIs. Tensor Flow-GPU was also used to create the model, which takes 4432 drug characteristics as input.

Medicines for inflammatory bowel disease (IBD) can predict how they will react; the trained model has an accuracy of 88 percent. The findings also revealed that the model performs best when many datasets are used. Detecting negative effects of drugs with a DNN Model was proposed by Wang et al. (2019). The model predicts ADRs by using synthetic, biological, and biomedical knowledge of drugs. Drug data from SIDER databases was also incorporated into the model. The proposed system's performance was improved by distributing. Using a word-embedding approach, determine the association between medications using the target drug representations in a vector space. The suggested system's fundamental flaw was that it only worked well with ordinary SIDER databases.

In 2020, numerous AI-based methods were developed for DDI event prediction, including evaluating chemical structural similarity using neural graph networks (Huang et al. 2020). Attempts to forecast DDI utilizing different data sources have also been made, such as leveraging similarity features to create pharmacological features for the DDI job predicting occurrences (Deng et al. 2020).

With the help of word embeddings, part-of-speech tags, and distance embeddings. Bai et al. (2020) suggested a deep learning technique that executes the DDI extraction task and supports the drug development cycle and drug repurposing. According to experimental data, the technique can better avoid instance misclassifications with minimal pre-processing. Moreover, the model employs an attention technique to emphasize the significance of each hidden state in the Bi-LSTM layers.

A tool for extracting features regarding a graph convolutional network (GCN) and a predictor based on a DNN. Feng et al. (2020) suggested DPDDI, an effective and robust approach for predicting potential DDIs by utilizing data from the DDI network lacking a thought of drug characteristics (i.e., drug chemical and biological properties). The proposed DPDDI is a useful tool for forecasting DDIs. It should benefit from other DDI-related circumstances, such as recognizing unanticipated side effects and guiding drug combinations. The disadvantage of this paradigm is that it ignores drug characteristics.

Zaikis and Vlahavas (2020), by developing a bi-level network with a more advanced level reflecting the network of biological entities' interactions, suggested a multi-level GNN framework for predicting biological entity links. Lower levels, however, reflect individual biological entities such as drugs and proteins, although the proposed model's accuracy needs to be enhanced.

In 2021, To overcome the DDI prediction, Lin et al. (2021) suggested an end-to-end system called Knowledge Graph Neural Network (KGNN). KGNN expands the use of spatial GNN algorithms to the knowledge graph by selectively various aggregators of neighborhood data, allowing it to learn the knowledge graph's topological structural information, semantic relations, and the neighborhood of drugs and drug-related entities. Medical risks are reduced when numerous medications are used correctly, and drug synergy advantages are maximized. For multi-typed DDI pharmacological effect prediction, Yue et al. (2021) used knowledge graph summarization. Lyu et al. (2021) also introduced a Multimodal Deep Neural Network (MDNN) framework for DDI event prediction. On the drug knowledge graph, a graph neural network was used, MDNN effectively utilizes topological information and semantic relations. MDNN additionally uses joint representation structure information, and heterogeneous traits are studied, which successfully investigates the multimodal data's complementarity across modes. Karim et al. (2019) built a knowledge graph that used CNN and LSTM models to extract local and global pharmacological properties across the network. DANN-DDI is a deep attention neural network framework proposed by Liu et al. (2021). To anticipate unknown DDIs, it carefully incorporates different pharmacological properties (Chun and Yi-Ping Phoebe 2021) and developed a deep hybrid learning (DL) model to provide a descriptive forecasting of pharmacological adverse reactions. It was one of the initial hybrid DL models through conception models that could be interpreted. The model includes a graph CNN through conception models to improve the learning efficiency of chemical drug properties and bidirectional long short-term memory (BiLSTM) recurrent neural networks to link drug structure to adverse effects. After concatenating the outputs of the two networks (GCNN and BiLSTM), a fully connected network is utilized to forecast pharmacological adverse reactions. Regardless of the classification threshold, the model obtains an AUC of 0.846. It has a 0.925 precision score. Even though a tiny drug data set was used for adverse drug response (ADR) prediction, the Bilingual Evaluation Understudy (BLEU) concluded results were 0.973, 0.938, 0.927, and 0.318, indicating considerable achievements. Furthermore, the model can correctly form words to explain pharmacological adverse reactions and link them to the drug's name and molecular structure. The projected drug structure and ADR relationship will guide safety pharmacology research at the preclinical stage and make ADR detection easier early in the drug development process. It can also

aid in the detection of unknown ADRs in existing medications. DDI extraction using a deep neural network model from medical literature was proposed by Mohsen and Hossein (). This model employs an innovative approach of attracting attention to improve the separation of essential words from other terms based on word similarity and location concerning candidate medications. Before recognizing the type of DDIs, this method calculates the results of a bi-directional long short-term memory (Bi-LSTM) model's attention weights in the deep network architecture. On the standard DDI Extraction 2013 dataset, the proposed approach was tested. According to the findings of the experiments, they were able to get an F1-Score of 78.30, which is comparable to the greatest outcomes for stated existing approaches.

In 2022, Pietro et al. (2022) introduced DruGNN, a GNN-based technique for predicting DDI side effects. Each DDI corresponds to a class in the prediction, a multi-class, multi-label node classification issue. To forecast the side effects of novel pharmaceuticals, they use a combination inductive-transudative learning system that takes advantage of drug and gene traits (induction path) and knowledge of known drug side effects (transduction path). The entire procedure is adaptable because the base for machine learning can still be used if the graph dataset is enlarged to include more node properties and associations. Zhang et al. (2022) proposed CNN-DDI, a new semi-supervised algorithm for predicting DDIs that uses a CNN architecture. They first extracted interaction features from pharmacological categories, targets, pathways, and enzymes as feature vectors. They then suggested a novel convolution neural network as a predictor of DDIs-related events based on feature representation. Five convolutional layers, two full-connected layers, and a CNN-based SoftMax layer make up the predictor. The results reveal that CNN-DDI superior to other cutting-edge techniques, but it takes longer to complete (Jing et al. 2022) presented DTSyn. This unique dual-transformer-based approach can select probable cancer medication combinations. It uses a multi-head attention technique to extract chemical substructure-gene, chemical-chemical, and chemical-cell-line connections. DTSyn is the initial model that incorporates two transformer blocks to extract linkages between interactions between genes, drugs, and cell lines, allowing a better understanding of drug action processes. Despite DTSyn's excellent performance, it was discovered that balanced accuracy on independent data sets is still limited. Collecting more training data is expected to solve the problem. Another issue is that the fine-granularity transformer was only trained on 978 signature genes, which could result in some chemical-target interactions being lost.

Furthermore, DTSyn used expression data as the only cell line attributes. To fully represent the cell line, additional omics data may be added going forward, including methylation and genetic data. He et al. (2022) proposed MFFGNN, a new end-to-end learning framework for DDI forecasting that can effectively combine information from molecular drug diagrams, SMILES sequences, and DDI graphs. The MFFGNN model used the molecular graph feature extraction module to extract global and local features from molecular graphs.

They run thorough tests on a variety of real-world datasets. The MFFGNN model routinely beats further cutting-edge models, according to the findings. Furthermore, the module for multi-type feature fusion configures the gating mechanism to limit the amount of neighborhood data provided to the node.

4.4 Drug–drug similarity prediction using DL

Drug similarity studies presume that medications with comparable pharmacological qualities have similar activation mechanisms, and side effects are used to treat problems like each other (Brown 2017; Zeng et al. 2019).

The drug-pharmacological similarity is critical for various purposes, including identifying drug targets, predicting side effects, predicting drug–drug interactions, and repositioning drugs. Features of the chemical structure (Lu et al. 2017; O'Boyle 2016), protein targets (Vilar 2016; Wang et al. 2014), side-effect profiles (Campillos et al. 2008; Tatonetti et al. 2012), and gene expression profiles (Iorio et al. 2010) provide a multi-perspective viewpoint for forecasting medications that are similar and can correct for data gaps in different data sources and offer fresh perspectives on drug repositioning and other uses. The main idea of drug–drug similarity is presented in Fig. 13. The vector represents the drug features, and the links reflect the similarity between the two drugs.

4.4.1 Drug similarity measures

The similarity estimations are calculated based on chemical structure, target protein sequence-based, target protein functional, and drug-induced pathway similarities.

4.4.1.1 The similarity in chemical structure DrugBank (2019) provides tiny molecule medicine chemical structures in SDF molecular format. Invalid SDFs can be recognized and eliminated, such as those with a NA value or fewer than three columns in atom or bond blocks. For valid compounds, atom pair descriptors can be computed, pairwise comparison of compounds, δ_c (*di*, *dj*), was evaluated using atom pairs using the Tanimoto coefficient, which is defined as the number of atom pairs in each fraction shared by two different compounds divided by their union (Eq. 1).

$$\delta_c(di, dj) = \left| AP_i \cap AP_j \right| \Big/ \left| AP_i \cup AP_j \right| \tag{1}$$

where AP_i and AP_j are atom pairs from pharmaceuticals d_i and dj, respectively, the numerator is the total number of atom pairs in both compounds, while the denominator is the number of common atom pairs in both compounds.

4.4.1.2 Target protein sequence-based similarity DrugBank provides all small molecule drugs have target sequences in FASTA format. The basic Needleman-Wunsch et al. (1970) dynamic programming approach for global alignment can be used to compare pairwise protein sequences. The proportion of pairwise sequence identity (Raghava 2006) can be represented as the corresponding sequence similarity. Equation 2 was used to calculate drug–drug similarity based on target sequence similarities:

$$\delta_t(d_i, d_j) = \frac{\left(\sum_{x \in T_i} \max_{\forall y \in T_j} \{S(x, y)\} + \sum_{x \in T_j} \max_{\forall x \in T_i} \{S(y, x)\} \right)}{\left(|T_i| * |T_j| \right)} \tag{2}$$

where δt (*di*, *dj*) denotes target-based similarity between medicines di and dj. Drugs di target a group of proteins known as Ti. Tj is a set of proteins that pharmaceuticals dj target and S(x,y) is a similarity metric based on symmetric sequences between two targeted proteins, x ∈ Ti and y ∈ Tj. Overall, Eq. 2 calculates the average of the best matches, wherein each first medicine's target is only connected to the second medicine's most comparable phrase, and vice versa.
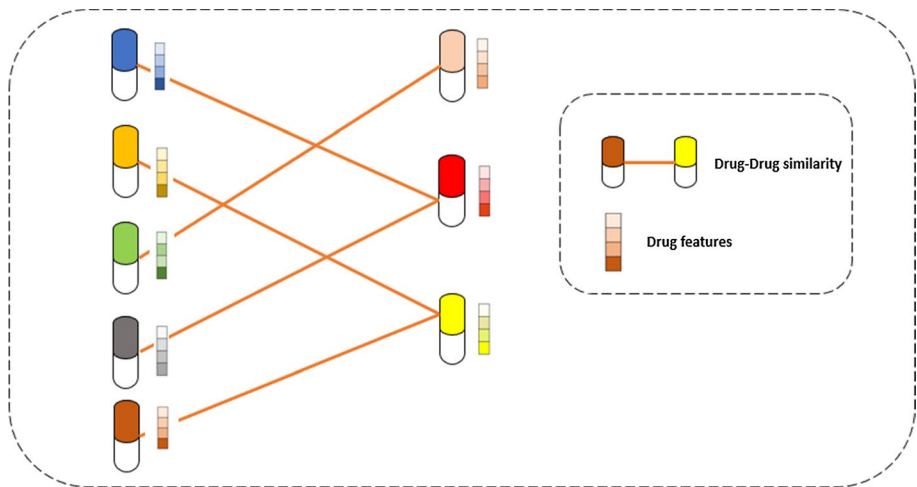
Fig. 13 Drug–drug similarity main idea

4.4.1.3 Target protein functional similarity Protein targets that are overrepresented by comparable biological functions and have similar sequences imply shared pharmacological mechanisms and downstream effects (Passi et al. 2018). As a result, each protein has a set of Gene Ontology (GO) concepts from all three categories associated with it, such as cellular components (CC), molecular functions (MF), and biological processes (BP). We filtered out GO keywords that were either very specialized (with 15 linked genes) or very general (with 100 genes). DrugBank (2019) provided the Human Protein–Protein Interaction (PPI) network. Wang et al. (2007) proposed leveraging the topology of the GO graph structure to determine the semantic similarity of their linked GO terms, which was used to determine how functionally comparable two drugs are, such as $\delta_f (d_i, d_j)$. Using a best-match average technique, any two GO keywords are compared for pairwise semantic similarity connected with di and d_j were aggregated into a single semantic similarity measure and presented into a final similarity matrix.

4.4.1.4 Drug-induced pathway similarity A medication pair that triggers similar pathways or overlaps shows that the drugs' mechanisms of action are similar, which is useful information for drug similarities and repositioning research (Zeng et al. 2015). Kanehisa and Goto (2000) was used to find the pathways activated by each small molecule medication. Using dice similarity, the similarity in pairs of any two options was calculated based on their constituent genes' closeness. After that, a pathway-based similarity score was calculated for each medication pair d_i and d_j, i.e., $\delta_p (d_i, d_j)$, was calculated using Eq. 3:

$$\delta_p\left(d_i, d_j\right) = \max_{\forall x \in P_i, \forall y \in P_j} \{DSC(x, y)\},\qquad(3)$$

where P_i and P_j are a group of drug-induced pathways d_i and d_j, respectively; x and y are two paths represented by a group of genes that make up their constituents, and $DSC(x, y) = 2|x \cap y|/(|x| + |y|)$ is the probability of a pair of dice matching, this determines how much the two trajectories overlap. When no gene is shared by any two pathways

produced by the comparing drug pair, the similarity is set to 0.0. Overall, Eq. 3 implies that if two medications stimulate one or more identical pathways, the maximum pathway-based similarity will be achieved (s).

4.4.2 DL for drug similarity prediction

Wang et al. (2019) introduced a gated recurrent units (GRUs) model that employs similarity to predict drug–disease interactions. In this approach, CDK turned the SMILES into 2D chemical fingerprints, and the Jaccard score of the 2D chemical fingerprints was used to compare the two medicines. This section comprehensively reviews the researchers' most popular DL algorithms to predict drug similarity.

Hirohara et al. (2018) employed a CNN to learn molecular representation. The network is given the molecule's SMILES notation as input to feed into the convolutional layers in this scenario. The TOX 21 dataset was used.

To conduct similarity analysis, Cheng et al. (2019) used the Anatomical Therapeutic Chemical (ATC) based on the drug ATC classification systems and code-based commonalities of drug pairs. The authors created interaction networks, performed drug pair similarity analyses, and developed a network-based methodology for identifying clinically effective treatment combinations for a specific condition.

Xin et al. (2016) presented a Ranking-based k-Nearest Neighbour (Re-KNN) technique for medication repositioning. The method's key feature combines the Ranking SVM (Support Vector Machine) algorithm and the traditional KNN algorithm. Chemical structural similarity, target-based similarity, side-effect similarity, and topological similarity are the types of similarity computation methodologies they used. The Tanimoto score was then used to determine the similarity between the two profiles.

Seo et al. (2020) proposed an approach that combined drug–drug interactions from DrugBank, network-based drug–drug interactions, polymorphisms in a single nucleotide, and anatomical hierarchy of side effects, as well as indications, targets, and chemical structures.

Zeng et al. (2019) developed an assessment of clinical drug–drug similarity derived from data from the clinic and used EHRs to analyse and establish drug–diagnosis connections. Using the Bonferroni adjusted hypergeometric P value, they created connections between drugs and diagnoses in an EMR dataset. The distances between medications were assessed using the Jaccard similarity coefficient to form drug clusters, and a k-means algorithm was devised.

Dai et al. (2020) reviewed, summarized representative methods, and discussed applications of patient similarity. The authors talked about the values and applications of patient similarity networks. Also, they discussed the ways to measure similarity or distance between each pair of patients and classified it into unsupervised, supervised, and semi-supervised.

Yan et al. (2019) created BiRWDDA, a new computational methodology for medication repositioning that combines bi-random walk and various similarity measures to uncover potential correlations between diseases and pharmaceuticals. First drug and disease–disease similarities are assessed to identify optimal drug and disease similarities. The information entropy is evaluated between the similarity of medicine and disease to determine the right similarities. Four drug–drug similarity metrics and three disease–disease similarity measurements were calculated depending on some drug- and

disease-related characteristics to create a heterogeneous network. The drug's protein sequence information, the extracted drug interaction from DrugBank then utilized the Jaccard score to determine this similarity, the chemical structure, derived canonical SMILES from DrugBank, and the side effect, respectively the four drug–drug similarities.

Yi et al. (2021) constructed the model of a deep gated recurrent unit to foresee drug–disease interactions that likely employ a wide range of similarity metrics and a kernel with a Gaussian interaction profile. Based on their chemical fingerprints, the similarity measure is utilized to detect a distinguishing trait in medications. Meanwhile, based on established disease–disease relationships, the Gaussian interactions profile kernel is used to derive efficient disease features. After that, a model with a deep gated recurrent cycle is created to anticipate drug-disease interactions that could occur. The outputs of the experiments showed that the suggested algorithm could be used to anticipate novel drug indications or disease treatments and speed up drug repositioning and associated drug research and discovery.

To forecast DDIs, Yan et al. (2022) suggested a semi-supervised learning technique (DDI-IS-SL). DDI-IS-SL uses the cosine similarity method to calculate drug feature similarity by combining chemical, biological, and phenotypic data. Drug chemical structures, drug–target interactions, drug enzymes, drug transporters, drug routes, drug indications, drug side effects, harmful effects of drug discontinuation, and DDIs that have been identified are all included in the integrated drug information.

Heba et al. (2021) used DrugBank to develop a machine learning framework based on similarities called "SMDIP" (Similarity-based ML for Drug Interaction Prediction), where they calculated drug–drug similarity utilizing a Russell–Rao metric for the biological and structural data that is currently accessible on DrugBank to represent the limited feature area. The DDI classification is carried out using logistic regression, emphasizing finding the main predictors of similarity. The DDI key features are subjected to six machine learning models (NB: naive Bayes; LR: logistic regression; KNN: k-nearest neighbours; ANN: neural network; RFC: random forest classifier; SVM: support vector machine).

For large-scale DDI prediction, Vilar et al. (2014) provided a procedure combining five similar drug fingerprints (Two-dimensional structural fingerprints, fingerprinting of interaction profiles, fingerprints of the target profile, Fingerprints of ADE profiles, and pharmacophoric techniques in three dimensions).

Song et al. (2022) used similarity theory and a convolutional neural network to create global structural similarity characteristics. They employed a transformer to extract and produce local chemical sub-structure semantic characteristics for drugs and proteins. To create drug and protein global structural similarity characteristics, The Tanimoto coefficient, Levenshtein distance, and CNN are all utilized in this study.

5 Benchmark datasets and databases

Drug development or discovery has been based on a range of direct and indirect data sources and has regularly demonstrated strong predictive capability in finding confirmed repositioning candidates and other applications for computer-aided drug design. This section reviews the most important and available benchmark datasets and databases used in the drug discovery problem and which the researchers may need according to each problem category. Thirty-five datasets are summarized in Table 3.

6 Evaluation metrics

Performance measures are required for evaluating machine learning models (Benedek et al. 2021). The measures serve as a tool for comparing different techniques. They aid in comparing many approaches to identify the best one for execution. This section describes the many metrics defined for the four categories of drug discovery difficulties below.

Table 4 shows the metrics employed in drug discovery problems—understanding the metrics aids in assessing the effectiveness of various prediction systems. True positives (TP) are drug side effects that have been recognized appropriately, False positives (FP) are adverse pharmacological effects that aren't present but were detected by the model, and True negatives (TN) are pharmacological side effects that do not exist but that the model failed to detect. False negatives (FN) are adverse pharmacological effects the model did not predict.

7 Drug dosing optimization

Drugs are vital to human health and choosing the proper treatment and dose for the right patient is a constant problem for clinicians. Even when taken as studied and prescribed, drugs have adverse impact profiles with varying response rates. As a result, all medications must be well-managed, especially those utilized in treating critical ailments or with a tight exposure window between efficacy and toxicity. Clinicians follow typical guidelines for the first dosage, which is not always optimal or secure for every patient, especially if the medicine no longer is evaluated in various dosages for various patient types. Precision dosage can revolutionize by increasing perks in health care while reducing drug therapy risks. While precise dosing will probably influence some pharmaceuticals significantly, perhaps not essential or practical to apply to all drugs or therapeutic classes. As a result, recognizing the characteristics that make medications suitable for precision dosage targets will aid in directing resources to where they'll have the most impact. Precision-dosing meds with a high priority and therapeutic classes could be crucial in achieving increased health care performance, safety, and cost-effectiveness (Tyson et al. 2020).

Due to standard, fixed dosing procedures or gaps in knowledge, imprecise drug dosing in specific subpopulations increases the risk of potentiating adverse effects due to supratherapeutic or subtherapeutic concentrations (Watanabe et al. 2018). Currently, the Food and Medicine Administration (FDA) simply requires a drug to be statistically better than a non-inferior to placebo of the existing treatment standard. This does not guarantee that the medicine will benefit most patients in clinical trials, especially if malignancies treatment can be tough, like diffuse intrinsic pontine glioma (DIPG) and unresectable meningioma, where rates of therapy response can be exceedingly low (Fleischhack et al. 2019).

There are essential aspects for dose optimization (https://friendsofcancerresearch.org/wpcontent/uploads/Optimizing_Dosing_in_Oncology_Drug_Development.pdf) that vary based on the product, the target population, and the available data to find the most effective dose, which varies based on the product, the target population, and the available data:

- Therapeutic properties: Drug features such as small molecule vs. large molecule and agonist vs. antagonist impact how drugs interact with the body regarding safety and efficacy. The therapeutic characteristics impact the first doses used in dose-finding studies and the procedures used to determine which doses should be used in registrational trials.

Table 4 The important metrics for drug discovery problems

Metrics	Description	Equations
Accuracy	It aids in determining the total number of side effects appropriately identified using the proposed technique	$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN)$
Sensitivity and specificity	The sensitivity metric aids in the identification of accurately identified positive samples. Specificity, on the other hand, detects the correctly established negative samples	$\text{Sensitivity} = TP/(TP + FN)$, $\text{Specificity} = TN/(TN + FP)$
Precision	The precision metric essentially identifies the relevant cases among all the expected events. As a result, it can be defined as true side effects resulting from all of a given technique's expected adverse effects	$\text{Precision} = TP/(TP + FP)$
Recall	The total number of real-life situations that have been discovered out of a total of accurate occurrences is referred to as recall	$\text{Recall} = TP/(TP + FN)$
F-measure	This statistic is used to assess a technique's accuracy. Precision and recall's harmonic mean can be defined as this	$\text{F-measure} = 2 \times ((\text{precision} \times \text{recall})/(\text{precision} + \text{recall}))$
ROC curve and AUC score	The receiving operator characteristic (ROC) curve is used to evaluate a model's performance across a variety of criterion values. It contrasts the ratio of true positives to false positives at various threshold values. The area under the ROC curve is measured by the AUC value. It aids in the model's proper evaluation because it is scale invariant and regardless of the value of the threshold	
PR curve and AUPR value	Although AUC cure has been used to evaluate the efficacy of several pharmacological Prediction models for side effects, most field-specific data sets are unbalanced, and the methodologies may exaggerate the AUC scores. As a result, the precision–recall (PR) curve has been employed to assess performance. At various threshold points, it compares the value of precision to recall. In addition, The AUPR value is a summary of the PR value over a wide range of thresholds	
Rate of predictions	The percentage of accurate projections and the frequency with which targets are missed have also been used to evaluate performance in some techniques	Rate of correct predictions = number of predictions/total number of experimentally known instances
Computational cost	The computational cost of a technology is defined as the overall time it takes to implement it on a given system. In most cases, it is measured in seconds	

Table 4 (continued)

Metrics	Description	Equations
Multilabel evaluation metrics	Some research regard side effect prediction to be a multilabel classification problem, and thus use multilabel evaluation measures to assess performance. Hammering loss, one error, coverage, and ranking loss are some of the most widely utilized measures	
Binding Affinity Score	Several assessment indicators can be used to examine it such as mean square error (MSE) and root mean square error (RMSE). These standards have been applied to rate the quality of the predictive QSAR model. The average squared difference between the predicted and actual scores for binding affinity is known as the MSE. The RMSE is the squaredMSE, as the name implies	

- Patient populations: Patient demographics vary depending on tumour kind, stage of disease, and comorbidities. Understanding how diverse factors influence the drug's efficacy may justify modifying the dose correspondingly, especially in the context of enlarged clinical trial populations.
- Supplemental versus original approval: Differences in disease features and patient demographics between tumour types and treatment settings, such as monotherapy versus combination therapy, must be considered when assessing whether additional dose exploration is required for a supplemental application. In cases when more dose exploration is required, the research design can include previous exposure-response knowledge from the initial approval.

8 Drug discovery and XAI

The topic of XAI addresses one of the most serious flaws in ML and DL algorithms: model interpretability and explain ability. Understanding how and why a prediction is formed becomes increasingly crucial as algorithms grow more sophisticated and can forecast with greater accuracy. It would be impossible to trust the forecasts of real-world AI applications without interpretability and explain ability. Human-comprehensible explanations will increase system safety while encouraging trust and sustained acceptance of machine learning technologies (). XAI has been studied to circumvent the limitations of AI technologies due to their black-box nature. In contrast to making decisions and model justifications which may be provided by AI approaches like DL and XAI (Zhang et al. 2022). Attention has been attracted to XAI approaches (Lipton 2018; Murdoch et al. 2019) to compensate for the lack of interpretability of some ML models as well as to aid human decision-making and reasoning (Goebel et al. 2018). The purpose of presenting relevant explanations alongside mathematical models is to help students understand them better by (1) Making the decision-making process more transparent (Doshi-Velez and Kim 2017), (2) correct predictions should not be made for the wrong motives (Lapuschkin et al. 2019), (3) avoid biases and discrimination that are unjust or unethical (Miller 2019), and (4) close the gap between ML and other scientific disciplines. Effective XAI can also help scientists in navigating the scientific process (Goebel et al. 2018), enabling people to fine-tune their understanding and opinions on the process under inquiry (Chander et al. 2018). We hope to provide an overview of recent XAI drug discovery research in this section.

XAI has a place in drug development. While the precise definition of XAI is still up for controversy (Guidotti et al. 2018), the following characteristics of XAI are unquestionably beneficial in applications of drug design (Lipton 2018):

- Transparency is accomplished by understanding how the system came to a specific result.
- The explanation of why the model's response is suitable serves as justification. It is instructive to provide new information to human decision-makers.
- Determining the reliability of a prediction to estimate uncertainty.

The molecular explanation of pharmacological activity is already possible with XAI (Xu et al. 2017; Ciallella and Zhu 2019), as well as drug safety and organic synthesis planning (Dey et al. 2018). If It's working overtime, XAI will be important in processing and interpreting increasingly complex chemical data, as well as creating new pharmaceutical

ideas, all while preventing human bias (Boobier et al. 2017). Application-specific XAI techniques are being developed to quickly reply to unique scientific issues relating to the Pathophysiology and biology of the human may be boosted by pressing drug discovery difficulties such as the coronavirus pandemic.

AI tools can increase their prediction performance by increasing model complexity. As a result, these models become opaque, with no clear grasp of how they operate. Because of this ambiguity, AI models are not generally utilized in important industries such as medical care. As a result, XAI focuses on understanding what goes into AI model prediction to meet the demand for transparency in AI tools. AI model interpretability approaches can be categorized depending on the algorithms used, a scale for interpreting, and the kind of information (Adadi and Mohammed 2018). Regarding the objectives of interpretability, approaches grouped as white-box model development, black-box model explanation, model fairness enhancement, and predictive sensitivity testing (Guidotti et al. 2018).

According to the gradient-based attribution technique (Simonyan et al. 2014), the network's input features are to blame for the forecast. Because this strategy is commonly employed when producing a DNN system's predictions, it may be a suitable solution for various black-box DNN models in DDI prediction (Quan, et al. 2016; Sun et al. 2018). In addition, DeepLIFT is a frequent strategy for implementing on top of DNN models that have been demonstrated to be superior to techniques based on gradients (Shrikumar et al. 2017). As opposed to that, the Guided Backpropagation model may be used to construct network architectures (Springenberg 2015). A convolutional layer with improved stride can be used instead of max pooling in CNN to deal with loss of precision. This method could be employed in CNN-based DDI prediction, as shown in Zeng et al. (2015).

Furthermore, in the Tao et al. (2016) was implemented neural networks that parse natural language. Using rationales, this method aimed to achieve the small pieces of input text. This method's design comprises two parts: a generator and an encoder that seek for text subsets that are closely connected to the predicted outcome. Because NLP-based models are used to extract DDIs (Quan et al. 2016), the above methods should be examined for usage in improving the model's clarity.

Aside from that, XAI has created methods for developing white-box models, including linear, decision tree, rule-based, and advanced but transparent models. However, these approaches are receiving less attention due to their weak ability to predict, particularly in the NLP-based sector, such as in the DDIs the job of extracting. Several ideas to address AI fairness have also been offered. Nonetheless, while extracting DDIs, only a small number of these scholarly studies looked at non-tabular data impartiality, such as text-based data. Many DDIs experiments used the word embedding method (Quan et al. 2016; Zhang 2020; Bolukbasi 2016). As a result, attempts to ensure fairness in DDI research should be considered more. To ensure the reliability of AI models, numerous methods also make an effort to examine the sensitivity of the models. Regarding their Adversarial Example-based Sensitivity Analysis, Zügner et al. (2018) used this model to explore graph-structured data. The technique looks at making changes to links between nodes or node properties to target node categorization models. Because graph-based methods are frequently utilized in DDIs research (Lin et al. 2021; Sun et al. 2020b), methods like those used in the previous study suggest that they might be used in a DDIs prediction model. In RNN, word embedding perturbations (Miyato et al. 1605) are also worth addressing. Significantly, the input reduction strategy utilized by Feng et al. (2018) to expose hypersensitivity in NLP models could be applied to DDI extraction studies. The DDIs study of Schwarz et al. (2021) attempted to provide model interpretability using Attention ratings derived at all levels of modeling in their DDIs study. The significance of similarity matrices to the vectors for

medication depiction is determined using these scores, and drug properties that contribute to improved encoding are identified using these scores. This method makes use of data that travels through all tiers of the network.

Graph neural networks (GNNs) and their explain ability are rapidly evolving in the field of graph data. GNNExplainer in Ying et al. (2019) uses mask optimization to learn soft masks for edge and node attributes to elaborate on the forecasts. Soft masks have been initiated at random and regarded as trainable variables. After that, the masks are then combined in comparison to the first graph using multiplications on a per-element basis by GNNExplainer. After that, by enhancing the exchange of information between the forecasts from the first graph and the recently acquired graph, the masks are maximized. Even when various regularization terms, such as element-by-element entropy, motivate optimal disguises for stealth, the resulting Masks remain supple.

In addition, because the masks are tuned for each input graph separately, it's possible that the explanations aren't comprehensive enough. To elaborate on the forecasts, PGExplainer (Luo et al. 2020) discovers approximated discrete edge masks. To forecast edge masks, it develops a mask predictor that is parameterized. It starts by concatenating node embeddings to get the embeddings for each edge in an input graph. The predictor then forecasts the chances of each edge being selected using the edge embeddings, that regarded as an evaluation of significance. The reparameterization approach is then used to sample the approximated discrete masks. Finally, the mutual information between the previous and new forecasts is optimized to train the mask predictor. GraphMask (Schlichtkrull et al. 2010) describes the relevance of edges in each GNN layer after the fact. It uses a classifier, like the PGExplainer, to forecast if an edge may be eliminated and does not impact the original predictions. A binary concrete distribution (Louizos et al. 1712) and a reparameterization method are used to roughly represent separate masks. The classifier is additionally trained by removing a term for a difference, which evaluates the difference between network predictions over the entire dataset. ZORRO (Thorben et al. 2021) employs discrete masks to pinpoint key input nodes and characteristics. A greedy method is used to choose nodes or node attributes from an input network. ZORRO chooses one node characteristic with the greatest fidelity score for each stage. The objective function, fidelity score, measures the degree of the recent forecasts resemble the model's original predictions by replacing the rest of the nodes/features with random noise values and repairing chosen nodes/features. The non-differentiable limitation of discrete masks is overcome because no training process is used.

Furthermore, ZORRO avoids the problem of "introduced evidence" by wearing protective masks. The greedy mask selection process, on the other hand, may result in optimal local explanations. Furthermore, because masks are generated for each graph separately, the explanations may lack a global understanding. Causal Screening (Xiang et al. 2021) investigates the attribution of causality to various edges in the input graph. It locates the explanatory subgraph's edge mask. The essential concept behind causal attribution is to look at how predictions change when an edge is added to the present explanatory subgraph, called the influence of causality. It examines the causal consequences of many edges at each step and selects one to include in the paragraph. It selects edges using the individual causal effect (ICE), which assesses the difference in information between parties after additional edges are introduced to the subgraph.

Causal Screening, like ZORRO, is a rapacious algorithm that generates undetectable masks without any prior training. As a result, it does not suffer due to the issue of the evidence presented. However, it is possible to lack worldwide comprehension and be caught in optimum local explanations. SubgraphX (Yuan et al. 2102) investigates deep graph

model subgraph-level explanations. It uses the Monte Carlo Tree Search (MCTS) method (Silver et al. 2017) to effectively investigate various subgraphs by trimming nodes and choose the most significant subgraph from the search tree's leaves as the explanation for the prediction.

Furthermore, the Shapley values can be used to update the mask generation algorithm's objective function. Its produced subgraphs are more understandable by humans and suited for graph data than previous perturbation-based approaches. However, the computational cost is higher because the MCTS algorithm explores distinct subgraphs.

9 Success stories about using DL in drug discovery

Big pharmaceutical companies have migrated toward AI as DL methodologies have advanced, abandoning conventional approaches to maximize patient and company profit. AstraZeneca is a multinational, science-driven, worldwide pharmaceutical company that has successfully used artificial intelligence in each stage of drug development, from virtual screening to clinical trials. They could comprehend current diseases better, identify new targets, plan clinical trials with higher quality, and speed up the entire process by incorporating AI into medical science. AstraZeneca's success is a shining illustration of how combining AI with medical science can yield incredible results. Their collaborations with other AI-based companies demonstrate their continual attempts to increase AI utilization. One such cooperation is with Ali Health, an Alibaba subsidiary that wants to provide AI-assisted screening and diagnosis systems in China (Nag et al. 2022).

SARS-CoV-2 virus outbreak placed many businesses under duress to develop the best medicine in the shortest amount of time feasible. These businesses have turned to employ AI in conjunction based on the data available to attain their goals. Below are some examples of firms that have been successful in identifying viable strategies to combat the COVID-19 virus because of their efforts.

Deargen, a South Korean startup, developed the MT-DTI (Molecule Transformer Drug Target Interaction Model), a DL-based drug-protein interaction prediction model. In this approach, the strength of an interaction between a drug and its target protein is predicted using simplified chemical sequences rather than 2D or 3D molecular structures. A critical protein on the COVID-19-causing virus SARS-CoV-2 is highly likely to bind to and inhibit the FDA-approved antiviral drug atazanavir, a therapy for HIV. It also discovered three more antivirals, as well as Remdesivir, a not-yet-approved medicine that is currently being studied in patients. Deagen's ability to uncover antivirals utilizing DL approaches is a significant step forward in pharmaceutical research, making it less time-consuming and more efficient. If such treatments are thoroughly evaluated, there is a good chance that we will be able to stop the epidemic in its tracks (Beck et al. 2020; Scudellari 2020).

Another example is Benevolent AI, a biotechnology company in London leverages medical information, AI, and machine learning to speed up health-related research. They've identified six medicines so far, one of which, Ruxolitinib, is claimed to be in clinical trials for COVID19 (Gatti et al. 2021). To find prospective medications that might impede the procedure for viral replication of SARS-CoV-2, The business has been utilizing a massive reservoir of information pertaining to medicine, together Utilizing data obtained from the scientific literature by their AI system and ML. They received FDA permission to use their planned Baricitinib medication in conjunction with Remdesivir, which resulted in a higher recovery rate for hospitalized COVID19 patients (Richardson et al. 2020).

Skin cancer is a form of cancer that is very frequent around the globe. As the rate at which skin cancer continues to rise, it is becoming increasingly crucial to diagnose it initially developed, research demonstrate that early identification and therapy improve the survival rate of skin cancer patients. With the advancement of medical research and AI, several skin cancer smartphone applications have been introduced to the market, allowing people with worrisome lesions to use a specialized technique to determine whether they should seek medical care. According to studies, over 235 dermatology smartphone apps were developed between 2014 and 2017 (Flaten et al. 2020). Previously, they worked by sending a snapshot of the lesion over the internet to a health care provider. Still, thanks to smartphones' internal AI algorithms, these applications can detect and classify images of lesions as high or low risk and Immediately assess the patient's risk and offer advice. Skin-Vison (Carvalho et al. 2019) is an example of a successful application.

10 Future challenges

10.1 Digital twinning in drug discovery

The development and implementation of Industry 4.0 emerging technologies allow for creation of digital twins (DTs), that promotes the modification of the industrial sector into a more agile and intelligent one. A DT is a digital depiction of a real entity that interacts in dynamic, two-way links with the original. Today, DTs are being used in a variety of industries. Even though the pharmaceutical sector has grown to accept digitization to embrace Industry 4.0, there is yet to be a comprehensive implementation of DT in pharmaceutical manufacture. As a result, it is vital to assess the pharmaceutical industry's success in applying DT solutions (Chen et al. 1088).

New digital technologies are essential in today's competitive marketplaces to promote innovation, increase efficiency, and increase profitability (Legner et al. 2017). AI (Venkatasubramanian 2019), Internet of Things (IoT) devices (Venkatasubramanian 2019; Oztemel and Gursev 2018), and DTs have all piqued the interest of governments, agencies, academic institutions, and corporations (Bao et al. 2018). Industry 4.0 is a concept offered by a professional community to increase the level of automation to boost productivity and efficiency in the workplace.

This section provides a quick look at the evolution of DT and its application in pharmaceutical and biopharmaceutical production. We begin with an overview of the technology's principles and a brief history, then present various examples of DTs in pharmacology and drug discovery. After then, there will be a discussion of the significant technical and other issues that arise in these kinds of applications.

10.1.1 History and main concepts of digital twin

The idea of making a "twin" of a process or a product returned to NASA's Apollo project in the late 1960s (Rosen et al. 2015; Mayani et al. 2018; Schleich et al. 2017), when it assembled two identical space spacecraft. In this scenario, the "twin" was employed to imitate the counterpart's action in real-time.

The DT, according to Guo et al. (2018), is a type of digital data structure that is generated as a separate entity and linked to the actual system. Michael Grieves presented

the original meaning of a DT in 2002 at the University of Michigan as part of an industry presentation on product lifecycle management (PLM) (Grieves 2014; Grieves and Vickers 2017; Stark et al. 2019). However, the first actual use of this notion, which gave origin to the current moniker, occurred in 2010, when NASA (the United States National Aeronautics and Space Administration) attempted to create virtual spaceship simulators for testing (Glaessgen and Stargel 2012).

A digital reproduction or representation of a physical thing, process, or service is what a DT is in theory. It's a computer simulation with unique features that dynamically connect the physical and digital worlds. The purpose of DTs is to model, evaluate, and improve a physical object in virtual space til it matches predicted performance, at which time it can be created or enhanced (if already built) in the real world (Kamel et al. 2021; Marr 2017).

Since then, DT technology has acquired popularity in both business and academia. Main components of DTs presently exist, as shown in Fig. 14. Still, the theoretical model comprises three parts: the real entity in the actual world, the digital entity in the virtual space, and the interconnection between them (Glaessgen and Stargel 2012).

In an ideal world, the digital component would have all the system's information that could be acquired from its physical counterpart (Kritzinger et al. 2018). When integrated with AI, IoT, and other recent intelligent systems, a DT can forecast how an object or process will perform.

10.1.2 Digital twin in pharmaceutical manufacturing

Developing a drug is lengthy and costly, requiring efforts in biology, chemistry, and manufacturing, and it has a low success rate. An estimated 50,000 hits (trial versions of compounds that are subsequently tweaked to develop a medication in the future) are evaluated to develop a successful drug. Only one in every 12 therapeutic compounds, clinical trials have been performed on humans, makes it to market successfully. Toxicity (A medication's capacity to offer a patient with respite and slow the progression of a disease) and lack of effectiveness contribute to more than 60% of all drug failures (Subramanian 2020).

Making the appropriate decisions about which targets, hits, leads, and compounds to pursue is important to a drug's successful market introduction. However, the decision is based on in vitro (Experimental system in a test tube or petri dish.) and in vivo (experiments in animals.) systems, both of which have a shaky correlation with clinical outcomes (Mak et al. 2014). Answers to the following inquiries would be provided by a perfect decision support system for drug discovery:

- What is the magnitude of any target's influence on the desired clinical result?
- Is the potential compound changing the target enough to change clinical outcomes?
- Is the chemical sufficiently selective and free of side effects or harmful consequences?
- Is the ineffectiveness attributable to the drug's failure to reach its target?
- Has the trial chosen the appropriate dose and dosing regimen?
- Are there any surrogate or biomarkers such as cholesterol that serves as a proxy for the illness's root cause that can forecast a drug's success or failure?
- Have the correct patients been chosen for the study?
- Is it possible to identify hyper- and hypo-responders before the study begins?

Therapeutic failures are prevalent and difficult to address, given the complex process of developing drugs based on the points above. This issue must be addressed by combining data and observations from many stages of the drug development process and developing a system that can forecast an experiment's outcome or a chemical modification's influence on a therapeutic molecule. This highlights the significance of DT in the field of drug discovery.

In the United States, funding organizations such as DARPA, NSF, and DOE have aggressively supported bioprocess modeling at the genomic and cellular levels, resulting in high-profile programs such as BioSPICE (Kumar and Feidler 2003). These groups have shown that smaller models built to answer specific issues can greatly influence drug development efficiency. This would make it possible to apply the prediction methodology to various stages of the drug discovery and research process, including confirmation of the target, enhancing leads, and choosing candidates, Recognition of biomarkers, fabrication of assays and screens, and the improvement of clinical trials.

The pharmaceutical business is embracing the overall digitization trend in tandem with the US FDA's ambition to establish an agile, adaptable pharmaceutical manufacturing sector that delivers high-quality pharmaceuticals without considerable regulatory scrutiny (O'Connor et al. 2016). Industries are beginning to implement Industry 4.0 and DT principles and use them for development and research (Barenji et al. 2019; Steinwandter et al. 2019; Lopes et al. 2019; Kumar et al. 2020; Reinhardt et al. 2020). Pharma 4.0 (Ierapetritou et al. 2016) is a digitalization initiative that integrates Industry 4.0 with International Council for Harmonisation (ICH) criteria to model a combined operational model and production control plan.
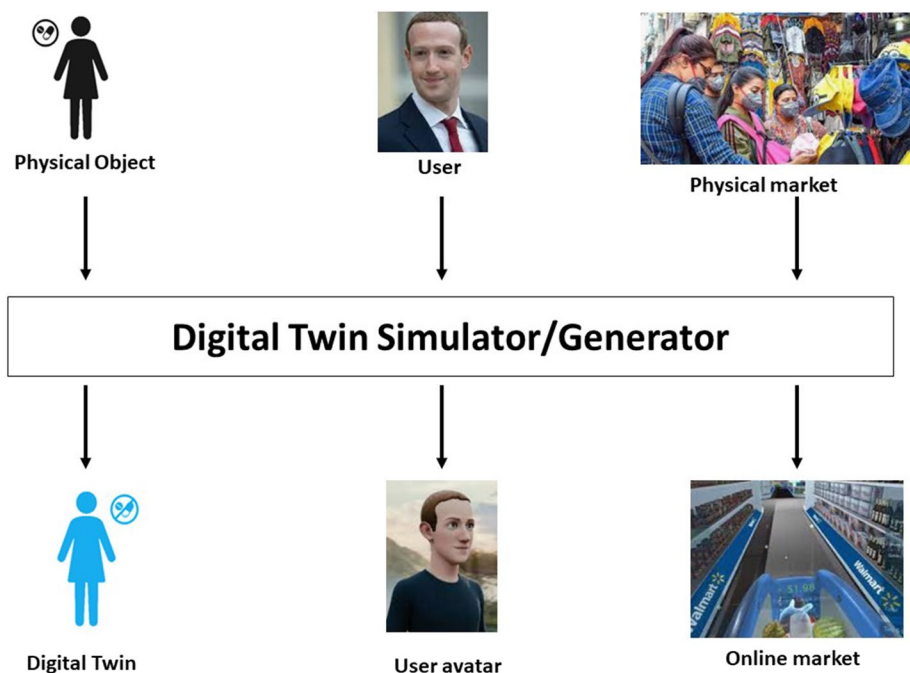

Fig. 14 Main components of DT

As shown in Fig. 15, live monitoring of the system `by the Process Analytical Technology (PAT), data collection from the machinery, the supplementary and finished goods, and a worldwide modelling and software for data analysis are some of the key requirements for achieving smart manufacturing with DT (Barenji et al. 2019). Quality-by-Design (QbD) and Continuous Manufacturing (CM) (Boukouvala et al. 2012), flowsheet modeling (Kamble et al. 2013), and PAT implementations (James et al. 2006) have all been used by the pharmaceutical industry to achieve this. Although some of the instruments have been thoroughly examined, DTs' entire integration and development is still a work in progress.

The pharmaceutical industry has used PAT in different programs across the steps involved in producing drugs (Nagy et al. 2013). Even though this has resulted in a rise in the use of PAT instruments, their implementations are limited to research and development rather than manufacturing on a large scale (Papadakis et al. 2018). They have been successful in decreasing production costs and enhancing product quality monitoring in the small number of examples where they have been used in manufacturing (Simon et al. 2019). The development of various PAT approaches, as well as their convincing implementation is a vital component of a scheme for surveillance and control (Boukouvala et al. 2012) and has given a foundation for obtaining essential data from the physical component.

Papadakis et al. (2018) recently provided a framework for identifying efficient reaction paths for pharmaceutical manufacture (Rantanen and Khinast 2015), which comprises modeling reaction route workflows discovery, analysis of reactions and separations, process simulation, assessment, optimization, and the use (Sajjia et al. 2017).

To develop models, data-driven modeling methods require the gathering and using of many substantial experiments, and the resulting models are solely reliant on the datasets provided. Artificial neural networks (ANN) (Pandey et al. 2006; Cao et al. 2018), multivariate statistical analysis, and in Monte Carlo Badr and Sugiyama (2020) are all commonly used in pharmaceutical manufacturing. These methods are less computationally costly, but the prediction outside the dataset space is frequently unsatisfactory due to the trained absence of underlying physics understanding in models. Using IoT devices in pharmaceutical manufacturing lines results in massive data collection volumes. The virtual component must receive this collection of process data and CQAs quickly and effectively. Additionally, for accurate prediction, several pharmaceutical process models need material properties. As a result, to provide virtual component access to all datasets, a central database site is necessary (Lin-Gibson and Srinivasan 2019).

10.1.3 Digital twin in biopharmaceutical manufacturing

The synthesis of big molecule-based entities in various combinations that has applications in the treatment of inflammatory, microbial, and cancer issues, is the focus of biopharmaceutical manufacturing (Glaessgen and Stargel 2012; Narayanan et al. 2020). The demand for biologic-based medications has risen in recent years, necessitating greater production efficiency and efficacy (Kamel et al. 2021). As a result, many businesses are switching from batch to continuous production and implementing intelligent manufacturing systems (Lin-Gibson and Srinivasan 2019). DT can aid in decision-making, risk analysis, product creation, and process prediction., which incorporates the physical plant, data collecting, data analysis, and system control (Tao et al. 2018).

biological products' components and structures are intimately connected to treatment effectiveness (Read et al. 2010) and are very sensitive to cell-line. Operating conditions thorough
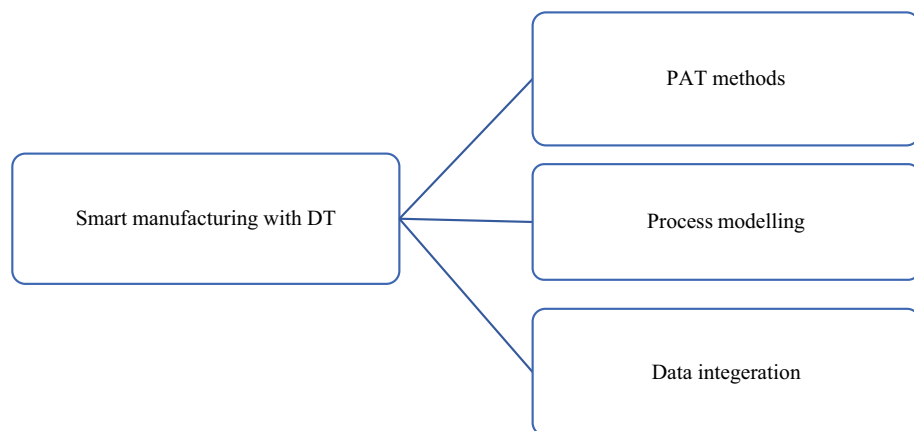
Fig. 15 Main categories of smart manufacturing with DT

actual plant's virtual description in a simulation environment is required to apply DT in biopharmaceutical manufacturing (Tao et al. 2018). This means that each unit activity inside an integrated model's simulation should accurately reflect the crucial process dynamics. Previous reviews Narayanan et al. (2020) Tang et al. (2020) Farzan et al. (2017) Baumann and Hubbuch (2017) Smiatek et al. (2020) and Olughu et al. (2019) focused on process modelling methodologies for both upstream and downstream operations.

Data from a biopharmaceutical monitoring system is typically diverse regarding data kinds and time scales. A considerable amount of data is collected during biopharmaceutical manufacture thanks to the deployment of real-time PAT sensors. As a result, data pre-processing is required to deal with missing data, visualize data, and reduce dimensions (Gangadharan et al. 2019). In batch biopharmaceutical production, Casola et al. (2019) presented data mining-based techniques for stemming, classifying, filtering, and clustering historical real-time data. Lee et al. (2012) combined different spectroscopic techniques and used data fusion to forecast the composition of raw materials.

10.2 AI-driven digital twins in today's pharmaceutical drug discovery

In the pharmaceutical industry, challenges are emerging from clinical studies that make drug development incomplete, sluggish, uncertain, and maybe dangerous. For example, It is not a true reflection of reality where clinical trials can take into account that in the real world, just a small portion of a big and diverse population is depicted among the many billions of humans on the planet where it is not possible to get a view of how each person based on how they will respond to a medicine. Clinical trials' rigorous requirements for physical and mental health in some cases also result in failure because of a lack of qualified participants. Pharmaceutical firms battle to provide the precise number and kind of participants needed to comply with the stringent requirements of clinical trial designs. Also, in most trials, the actual drug is replaced by a placebo as this helps contrast how sick individuals behave when they are not administered the experimental medication; This implies that at least some trial participants do not receive it. Here, These issues can be solved by using digital twins, which can imitate a range of patient features, giving a fair representation of how a medicine affects a larger population. AI-enabled digital twinning may reduce the trial's setup by revealing how susceptible a patient

is to various inclusion and exclusion criteria as a result, patients can be rapidly identified, and digital twins can predict a patient's reaction, and placebos won't be required. Therefore, the new treatment can be assured for every patient in the trial, and digital twins can reduce the dangerous impact of drugs in the early stages by decreasing the number of patients who need to be tested in the real world. Figure 16 illustrates a framework by running all possible combinations. All treatment protocols are tested on a digital twin of the patient to discover an appropriate treatment protocol for this patient. Doing this quickly and accurately can lead to providing the best quality treatment for the patient without experimenting with the patient, which saves effort, cost, and accuracy in determining an appropriate treatment protocol for patients.

11 Open problems

This section discusses important issues to consider regarding progression from preclinical to clinical and implementation in practice that necessitate new ML solutions to assist transparent, usable, and data-driven decision-making procedures to accelerate drug discovery and decrease the number of failures in clinical development phases.

– Complex disorders, such as viral infections and advanced malignancies frequently necessitate drug combinations (Julkunen et al. 2020; White et al. 2021). For example, kinase inhibitor combos or single compounds that block several kinases may improve therapeutic efficacy and duration while combating treatment resistance in cancer (Attwood et al. 2021). While several ML models have been created to predict response pairs of drug–dose combinations, higher-order combination effects can be predicted in a systematic way involving more than two medicines or targets is still a problem. In cancer cell lines, tensor learning methods have permitted reliable prediction of paired drug combination dose-response matrices (Smiatek et al. 2020). This computationally efficient learning approach could use extensive pharmacogenomic data, determine which drug combinations are most successful for additional in vitro or in vivo testing in many kinds of preclinical models, such as higher-order combinations among novel therapeutic compounds and doses.
– While possible toxicity and effectiveness that is targeted are important criteria for clinical development success, most existing ML models for predicting response to the therapy accentuate effectiveness as the primary result. As a result, careful examination, and harmful effects prediction of instances in simulated and preclinical settings is required to strike a balance between the effectiveness of the toxicity and therapy that is acceptable to accelerate the next stages of drug development (Narayanan et al. 2020). Applying single-cell data and ML algorithms to develop combinations of anticancer drugs has shown the potential to boost the likelihood of clinical success (Tao et al. 2018). Transfer of knowledge and deconvolution techniques for in silico cell set (Avila et al. 2020) may offer effective ways to reduce the requirement to generate a lot of single-cell data to predict combination therapy responders and impacts of toxicity, as well as the recommended dosage that optimizes both efficacy and safety.
– In addition, patient data and clinical profiles must be used to validate the in-silico therapy response forecasts. This real data for ML predictions is crucial for progress in medicine and establishing the practical value and providing clinical guidance in making decisions. A no-go decision was made early, for example, if the substance has harmful consequences. Many of the present issues encountered when using machine learning
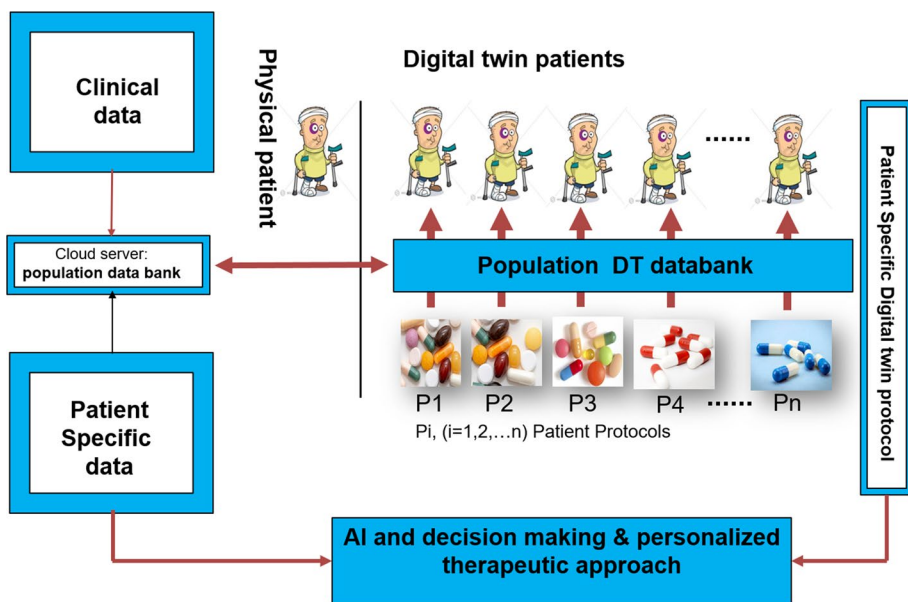
Fig. 16 AI-driven digital twins in today's pharmaceutical drug discovery

for drug discovery, particularly in clinical development, are since current AI algorithms do not meet the requirements for clinical research. As a result, ML model validation requires systematic and comprehensive high-quality clinical data sets. The discovery methods must be thoroughly evaluated for accuracy and reproducibility using community-agreed performance measures in various settings, not just a small collection of exemplary data sets. sharing and exploiting private patient information is possible with systems that isolate the code from the data or use the model to data method (Guinney and Saez-Rodriguez 2018), which It makes it possible for federated learning to utilise patient-level data for model construction and thorough assessment.

– Even if there are many applications for drug discovery, The majority of ML and particularly DL models remain "black boxes", and interpretation by a human specialist is sometimes tricky (Jiménez-Luna et al. 2020). Implementing mathematical models as online decision support tools must be understandable to users to obtain confidence. Comprehensible, accessible, and explainable models should clearly state the optimization goals, such as synergy, efficacy, and/or toxicity.

– DTI prediction is a notable example of fields of drug discovery research. It has been ongoing more than 10 years and aims to enhance the effectiveness of computational models using various technologies. The most recent computational approaches for predicting DTIs are DL technologies. These use unstructured-based approaches that don't need 3D structural data or docking to get over the drug and target protein's high-dimensional structure restrictions. Despite the DL's outstanding performance, regression inside the DTI prediction remains a critical and difficult issue, and researchers could develop several strategies to improve prediction accuracy. Furthermore, data scarcity and the lack of a standardized benchmark database are still considered current research gaps.

– While DL approaches show promise in detecting drug responses, especially when dealing with large amounts of data, drug response prediction research is in its first stages, and more efficient and relevant models are needed.

– While DL techniques have shown to be effective in detecting DDIs, especially when dealing with large amounts of data, more promising algorithms that focus on complex molecular reactions need to be developed.

– Only a few studies in the drug discovery field have investigated their models' explain ability, leaving much room for improvement. The explanations generated by XAI for human decision-making must be not insignificant, not artificial, and helpful to the scientific community. Until now, ensuring that XAI techniques achieve their goals and produce trustworthy responses would necessitate a combined effort amongst DL specialists, chemo informaticians and chemists, biologists, data scientists, and other subject matter experts. As a result, we believe that more developed methodologies to explain black-box models for drug discovery fields like DDIs, drug–target interactions, drug sensitivity, and drug side effects must be considered in the future to ensure model fairness or strict sensitivity evaluations of models. Further exploration of the capabilities and constraints of the existing chemical language for defining these models will be critical. The development of novel interpretable molecular representations for DL and the deployment of self-explanatory algorithms alongside sufficiently accurate predictions will be a critical area of research in the coming years. Because there are currently no methods that combine all the stated advantageous XAI characteristics (transparency, justification, informativeness, and uncertainty estimation), consensus techniques that draw on the advantages of many XAI approaches and boost model dependability will play a major role in the short and midterm. Currently, there is no open-community platform for exchanging and refining XAI software and model interpretations in drug discovery. As a result, we believe that future study into XAI in drug development has much potential.

12 Discussion

This section presents a brief about how the proposed analytical questions in Sect. 2 are being answered through the paper.

– AQ1: What DL algorithms have been used to predict the different categories of drug discovery problems?

   Several DL algorithms have been used to predict the different categories of drug discovery problems as deeply illustrated in Sect. 4 with respect to the main categories of drug discovery problems in Fig. 8. In addition, a summary of a sample of these algorithms, their methods, advantages and weaknesses are presented in Table 2.

– AQ2: Which deep learning methods are mostly used in drug dosing optimization?

   Recognizing the characteristics that make medications suitable for precision dosage targets will aid in directing resources to where they'll have the most impact. Employing DL in drug dosing optimization is a big challenge which increases the health care performance, safety, and cost-effectiveness as presented in Sect. 7.

– AQ3: Are there any success stories about drug discovery and DL?

   With the advancement of DL methods, we've seen big pharmaceutical businesses migrate toward AI, such as 'AstraZeneca' which is a global multinational pharmaceuti-

cal business that has successfully used AI in every stage of drug development. Several success stories have been presented in Sect. 9.

- AQ4: What about using the newest technologies such as XAI and DT in drug discovery?

  The topic of XAI addresses one of the most serious flaws in ML and DL algorithms: model interpretability and explain ability. It would be impossible to trust the forecasts of real-world AI applications without interpretability and explain ability. Section 8 presents the literature that address this issue. A digital twin (DT) is a virtual representation of a living thing that is connected to the real thing in dynamic, reciprocal ways. Today, DTs are being used in a variety of industries. Even though the pharmaceutical sector has grown to accept digitization to embrace Industry 4.0, there is yet to be a comprehensive implementation of DT in pharmaceutical manufacture. Success stories regarding employing DT into drug discovery is presented in Sect. 10.

- AQ5: What are the future and open works related to the drug discovery and DL?.

  Through the paper, we present how DL succeed in all aspects of drug discovery problems, However, it is still a very important challenge for future research. Section 11 covers these challenges.

Figure 17 presents the percentage of the different DL applications for each building block of our study. It is well observed that the most percentage segment is dedicated for the drug discovery and DL because it is the main core of our research.

13 Conclusion

Despite all the breakthroughs in pharmacology, developing new drugs still requires a lot of time and costs. As DL technology advances and the amount of drug-related data grows, a slew of new DL-based approaches is cropping up at every stage of the drug development process. In addition, we've seen large pharmaceutical corporations migrate toward AI in the wake of the development of DL approaches.

Although the drug discovery is a large field and has different research categories, there is a few review studies about this field and each related study has focused only on a one research category such as reviewing the DL applications for the DTIs. So, the main goal of our research is to present a systematic Literature review (SLR) which integrates the recent DL technologies and applications for the different categories of drug discovery problems Including, Drug–target interactions (DTIs), drug–drug similarity interactions (DDIs), drug sensitivity and responsiveness, and drug-side effect predictions. That is associated with the benchmark data sets and databases. Related topics such as XAI and DT and how they support the drug discovery problems are also discussed. In addition, the drug dosing optimization and success stories are presented as well. Finally, we suggest open problems as future research challenges.

Although the DL has proved its strength in drug discovery problems, it is still a promising open research area for the interested researchers. In this paper, they can find all they want to know about using DL in various drug discovery problems. In addition, they can find success stories and open areas for future research.

Given the recent success of DL approaches and their use by pharmaceuticals in identifying new medications, it seems clear that current DL techniques being highly regarded in the next generation of enormous data investigation and evaluation for drug discovery and development.
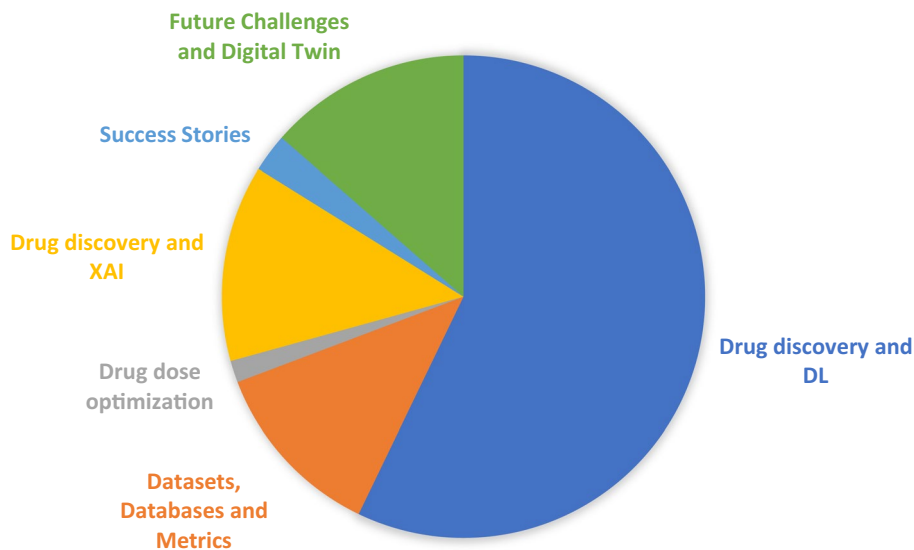
Fig. 17 Percentages of DL applications for each category

Declarations

Competing interests The authors declare no competing interests.

References

Abramovich I, Ben-Yehuda T, Cohen R (2018) Low-complexity video classification using recurrent neural networks. IEEE Int Conf Sci Electr Eng Israel (ICSEE) 2018:1–4. https://doi.org/10.1109/ICSEE.2018.8646076

Adadi A, Mohammed B (2018) Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access 6:2169–3536

Ahmed KT, Park S, Jiang Q et al (2020) Network-based drug sensitivity prediction. BMC Med Genomics 13:193

Alankrita A, Mamta M, Gopi B (2021) Generative adversarial network: an overview of theory and applications. Int J Inf Manag Data Insights 1(1):100004

Amashita R, Nishio M, Do RKG et al (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9:611–629. https://doi.org/10.1007/s13244-018-0639-9

Andreea D, Yu-Hsiang H, Petar V, Pietro L, Jian T (2019) Drug–drug adverse effect prediction with graph co-attention. https://arxiv.org/abs/1905.00534

Arshed MA, Mumtaz S, Riaz O, Sharif W, Abdullah S (2022) A deep learning framework for multi drug side effects prediction with drug chemical substructure. Int J Innovat Sci Technol 4(1):19–31

Arus-Pous J, Patronov A, Bjerrum EJ, Tyrchan C, Reymond JL, Chen H, Engkvist O (2020) SMILES-based deep generative scaffold decorator for de-novo drug design. J Cheminform 12:1–18

Asada M, Miwa M, Sasaki Y (2018) Enhancing drug–drug interaction extraction from texts by molecular structure information. In: proceedings of the 56th annual meeting of the association for computational linguistics. 2, pp 680–685, https://doi.org/10.18653/v1/P18-2108

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al (2000) Gene ontology: tool for the unification of biology. Nat Genet 25:25–29

Attwood MM, Fabbro D, Sokolov AV et al (2021) Trends in kinase drug discovery: targets, indications and inhibitor design. Nat Rev Drug Discov 20(11):839–861

Avila C, Alquicira-Hernandez J, Powell JE et al (2020) Benchmarking of cell type deconvolution pipelines for transcriptomics data. Nat Commun 11(1):5650

Azad AKM, Dinarvand M, Nematollahi A, Swift J, Lutze-Mann L, Vafaee F (2021) A comprehensive integrated drug similarity resource for in-silico drug repositioning and beyond. Brief Bioinform 22(3):bbaa126. https://doi.org/10.1093/bib/bbaa126

Badr S, Sugiyama H (2020) A PSE perspective for the efficient production of monoclonal antibodies: integration of process, cell, and product design aspects. Curr Opin Chem Eng 27:121–128

Bao J, Guo D, Li J, Zhang J (2018) The modelling and operations for the digital twin in the context of manufacturing. Enterp Inf Syst 13:534–556

Baptista D, Ferreira PG, Rocha M (2021) Deep learning for drug response prediction in cancer. Briefings Bioinform 22:360–379

Barenji RV, Akdag Y, Yet B, Oner L (2019) Cyber-physical-based PAT (CPbPAT) framework for Pharma 4.0. Int J Pharm 567:118445

Baumann P, Hubbuch J (2017) Downstream process development strategies for effective bioprocesses: Trends, progress, and combinatorial approaches. Eng Life Sci 17:1142–1158

Beck BR, Shin B, Choi Y, Park S, Kang K (2020) Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug–target interaction deep learning model. Comput Struct Biotechnol J 18:784–790

Bedi P, Sharma C, Vashisth P, Goel D, Dhanda M (2015) Handling cold start problem in Recommender Systems by using Interaction Based Social Proximity factor. In: Proceeding of the 2015 international conference on advances in computing, communications and informatics, Kerala, India, 10–13 August 2015; pp 1987–1993

Benedek R, Stephen B, Andriy N, Michael U, Sebastian N, Eliseo P (2021) A unified view of relational deep learning for drug pair scoring. coRR V. https://arxiv.org/abs/2111.02916.

Betsabeh T, Mansoor ZJ (2021) Using drug–drug and protein-protein similarities as feature vector for drug–target binding prediction. Chemom Intell Lab Syst 217:104405. https://doi.org/10.1016/j.chemolab.2021.104405

Bleakley K, Yamanishi Y (2009) Supervised prediction of drug–target interactions using bipartite local models. Bioinformatics 25:2397–2403

Bolukbasi T (2016) Man is to computer programmer as woman is to homemaker? debiasing word embeddings. Advances in neural information processing systems, 2016; 29. In Identifying gender and sexuality of data subjects. https://cis.pubpub.org/pub/debiasing-word-embeddings-2016.

Bongini P, Pancino N, Dimitri GM, Bianchini M, Scarselli F, Lio P (2022) Modular multi-source prediction of drug side-effects with DruGNN. http://arxiv.org/abs/2202.08147.

Boobier S, Osbourn A, Mitchell JB (2017) Can human experts predict solubility better than computers? J Cheminform 9:63

Boukouvala F, Niotis V, Ramachandran R, Muzzio FJ, Ierapetritou MG (2012) An integrated approach for dynamic flowsheet modeling and sensitivity analysis of a continuous tablet manufacturing process. Comput Chem Eng 42:30–47

Brown AS, Patel CJ (2017) MeSHDD: literature-based drug-drug similarity for drug repositioning. J Am Med Inf Assoc 24(3):614–618

Camacho DM, Collins KM, Powers RK, Costello JC, Collins JJ (2018) Next-generation machine learning for biological networks. Cell 173:1581–1592

Campillos M et al (2008) Drug target identification using side-effect similarity. Science 321(5886):263–666. https://doi.org/10.1126/science.1158140

Cao H, Mushnoori S, Higgins B, Kollipara C, Fermier A, Hausner D, Jha S, Singh R, Ierapetritou M, Ramachandran R (2018) A systematic framework for data management and integration in a continuous pharmaceutical manufacturing processing line. Processes 6:53

Casola G, Siegmund C, Mattern M, Sugiyama H (2019) Data mining algorithm for pre-processing biopharmaceutical drug product manufacturing records. Comput Chem Eng 124:253–269

Chabner BA (2016) NCI-60 cell line screening: a radical departure in its time. J Natl Cancer Inst. https://doi.org/10.1093/jnci/djv388

Chander A, Srinivasan R, Chelian S, Wang J, Uchino K (2018) Working with beliefs: AI transparency in the enterprise. In: Joint proceedings of the ACM IUI 2018 workshops co-located with the 23rd acm conference on intelligent user interfaces 2068 (eds Said, A. and Komatsu, T.) (CEUR-WS.org, 2018)

Chandra B, Sharma RK (2017) On improving recurrent neural network for image classification. Int Jt Conf Neural Netw (IJCNN) 2017:1904–1907. https://doi.org/10.1109/IJCNN.2017.7966083

Chang Y, Park H, Yang HJ, Lee S, Lee KY, Kim TS, Jung J, Shin JM (2018) Cancer drug response profile scan (CDRscan): a deep learning model that predicts drug effectiveness from cancer genomic signature. Sci Rep 8:1–11

Chauhan R, Ghanshala KK, Joshi RC (2018) Convolutional neural network (CNN) for image detection and recognition. First Int Conf Secure Cyber Comput Commun (ICSCCC) 2018:278–282. https://doi.org/10.1109/ICSCCC.2018.8703316

Chen AW (2018) Predicting adverse drug reaction outcomes with machine learning. Int J Commun Med Public Health 5(3):901–904

Chen JY, Mamidipalli S, Huan T (2009) Happi: an online database of comprehensive human annotated and predicted protein interactions. BMC Genomics 10(1):S16

Chen X, Liu M-X, Yan G-Y (2012) Drug–target interaction prediction by random walk on the heterogeneous network. Mol BioSyst 8:1970–1978. https://doi.org/10.1039/C2MB00002D

Chen Y, Yang O, Sampat C, Bhalode P, Ramachandran R, Ierapetritou M (2020) Digital twins in pharmaceutical and biopharmaceutical manufacturing: a literature review. Processes 8(9):1088. https://doi.org/10.3390/pr8091088

Cheng F, Kovács IA, Barabási AL (2019) Network-based prediction of drug combinations. Nat Commun 10(1):1–11

Chiu Y-C, Chen H-IH, Zhang T, Zhang S, Gorthi A, Wang L-J, Huang Y, Chen Y (2019) Predicting drug response of tumors from integrated genomic profiles by deep neural networks. BMC Med Genomics 12:119

Chu X, Lin Y, Gao J, Wang J, Wang Y, Wang L (2018) Multi-label robust factorization autoencoder and its applicationin predicting drug–drug interactions. arXiv:1811.00208.

Chu X, Lin Y, Wang Y, Wang L, Wang J, Mlrda JG (2019) A multitask semi-supervised learning framework for drug–drug interaction prediction. In: proceedings of the international joint conference on artificial intelligence, pp 4518– 4524

Ciallella HL, Zhu H (2019) Advancing computational toxicology in the big data era by artificial intelligence: data-driven and mechanism-driven modeling for chemical toxicity. Chem Res Toxicol 32:536–547

Cortes-Ciriano I, Ain QU, Subramanian V, Lenselink EB, Méndez-Lucio O, IJzerman AP, Wohlfahrt G, Prusis P, Malliavin TE, van Westen GJP et al (2015) Polypharmacology modelling using proteochemometrics (PCM): recent methodological developments, applications to target families, and future prospects. Medchemcomm 6:24–50

Cortés-Ciriano I, Bender A (2019) KekuleScope: prediction of cancer cell line sensitivity and compound potency using convolutional neural networks trained on compound images. J Cheminform 11:1–16

Dai L, Zhu H, Liu D (2020) Patient similarity: methods and applications. http://arxiv.org/abs/2012.01976

David L, Arús-Pous J, Karlsson J, Engkvist O, Bjerrum EJ, Kogej T, Kriegl JM, Beck B, Chen H (2019) Applications of deep-learning in exploiting large-scale and heterogeneous compound data in industrial pharmaceutical research. Front Pharmacol 10:1303

Davis MI, Hunt JP, Herrgard S, Ciceri P, Wodicka LM, Pallares G, Hocker M, Treiber DK, Zarrinkar PP (2011) Comprehensive analysis of kinase inhibitor selectivity. Nat Biotechnol 29:1046–1051

De Carvalho TM, Noels E, Wakkee M, Udrea A, Nijsten T (2019) Development of smartphone apps for skin cancer risk assessment: progress and promise. JMIR Dermatol 2(1):e13376

De Kuijper GM, Risselada A, van Dijken R (2019) Monitoring drug side-effects. Handbook of intellectual disabilities. Springer, Cham, pp 275–301

"deepchem/deepchem: Democratizing Deep-Learning for Drug Discovery"; Quantum Chemistry, Materials Science and Biology; Available online: https://github.com/deepchem/deepchem (accessed on 15 April 2022).

Dey S, Luo H, Fokoue A, Hu J, Zhang P (2018) Predicting adverse drug reactions through interpretable deep learning framework. BMC Bioinform 19:476

Dincer AB, Celik S, Hiranuma N, Lee S-I (2018) DeepProfile: deep learning of cancer molecular profiles for precision medicine. bioRxiv. https://doi.org/10.1101/278739

Ding MQ, Chen L, Cooper GF, Young JD, Lu X (2018) Precision oncology beyond targeted therapy: combining omics data with machine learning matches the majority of cancer cells to effective therapeutics. Mol Cancer Res 16:269–278

Doshi-Velez F, Kim B (2017) Towards a rigorous science of interpretable machine learning. https://arxiv.org/abs/1702.08608

DrugBank (2019) DrugBank Release Version 5.1.3, chemical structures. https://www.drugbank.com

Dua D, Graff C (2017) UCI machine learning repository. https://archive.ics.uci.edu/ml/index.php

El-Deredy W et al (1997) Pretreatment prediction of the chemotherapeutic response of human glioma cell cultures using nuclear magnetic resonance spectroscopy and artificial neural networks. Cancer Res 57:4196–4199

Farzan P, Mistry B, Ierapetritou MG (2017) Review of the important challenges and opportunities related to modeling of mammalian cell bioreactors. AIChE J 63:398–408

Fatehifar M, Karshenas H (2021) Drug–drug interaction extraction using a position and similarity fusion-based attention mechanism. J Biomed Inf 115:103707. https://doi.org/10.1016/j.jbi.2021.103707

Feng S, et al (2018) Pathologies of neural models make interpretations difficult. http://arxiv.org/abs/1804.07781

Feng Q, Dueva E, Cherkasov A, Ester M (2018) PADME: a deep learning-based framework for drug–target interaction prediction. arXiv 2018; arXiv:1807.09741

Feng YH, Zhang SW, Shi JY (2020) DPDDI: a deep predictor for drug–drug interactions. BMC Bioinform 21:419. https://doi.org/10.1186/s12859-020-03724-x

Ferdousi R, Safdari R, Omidi Y (2017) Computational prediction of drug–drug interactions based on drugs functional similarities. J Biomed Inform. https://doi.org/10.1016/j.jbi.2017.04.021

Finn RD et al (2013) Pfam: the protein families database. Nucleic Acids Res 42(D1):D222–D230

Flaten HK, St Claire C, Schlager E, Dunnick CA, Dellavalle RP (2020) Growth of mobile applications in dermatology. Dermatol Online J 24(2):13–16

Fleischhack G, Massimino M, Warmuth-Metz M, Khuhlaeva E, Janssen G, Graf N et al (2019) Nimotuzumab and radiotherapy for treatment of newly diagnosed diffuse intrinsic pontine glioma (DIPG): a phase III clinical study. J Neurooncol 143:107–113. https://doi.org/10.1007/s11060-019-03140-z

Fokoue A, Sadoghi M, Hassanzadeh O, Zhang P (2016) Predicting drug–drug interactions through large-scale similarity-based link prediction. In: European semantic web conference 2016 May 29; pp 774–789

Fushman D, Shooshan SE, Rodriguez L, Aronson AR, Lang F, Rogers W, Tonning J (2018) A dataset of 200 structured product labels annotated for adverse drug reactions. Sci Data 5:180001

Gangadharan N, Turner R, Field R, Oliver SG, Slater N, Dikicioglu D (2019) Metaheuristic approaches in biopharmaceutical process development data analysis. Bioprocess Biosyst Eng 42:1399–1408

Gao Z et al (2008) PDTD: a web-accessible protein database for drug target identification. BMC Bioinf 9(1):104

Gao KY, Fokoue A, Luo H, Iyengar A, Dey S, Zhang P (2017) Interpretable drug target prediction using deep neural representation. In: Proceedings of the international joint conference on artificial intelligence, Melbourne, Australia, 19–25 August 2017

Gao K, Duy Nguyen D, Sresht V, Mathiowetz AM, Tu M, Wei G-W (2019) Are 2D fingerprints still valuable for drug discovery? Phys Chem Chem Phys 22:8373–8390

Gatti M, Turrini E, Raschi E, Sestili P, Fimognari C (2021) Janus kinase inhibitors and coronavirus disease (COVID)-19: rationale, clinical evidence and safety issues. Pharmaceuticals 14(8):738

Gaulton A et al (2011) ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 40(D1):D1100–D1107

Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE (2017) Neural message passing for quantum chemistry. 34th Int Conf Mach Learn ICML 3:2053–2070

Glaessgen EH, Stargel DS (2012) The digital twin paradigm for future NASA and US Air Force vehicles. In: Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference, Honolulu, HI, USA. https://ntrs.nasa.gov/citations/20120008178

Goebel R et al (2018) Explainable AI: the new 42? In: Holzinger A, Kieseberg P, Tjoa A, Weippl E (eds) Machine learning and knowledge extraction. CD-MAKE Lecture Notes in Computer Science. Springer, New York

Gómez-Bombarelli R et al (2018) Automatic chemical design using a data-driven continuous representation of molecules. ACS Cent Sci 4:268–276

Grieves M (2014) Digital twin: manufacturing excellence through virtual factory replication. Glob J Eng Sci Res. https://doi.org/10.5281/zenodo.1493930

Grieves M, Vickers J (2017) Digital twin: mitigating unpredictable undesirable emergent behavior in complex systems. Springer, Cham, pp 85–113

Guidotti R et al (2018) A survey of methods for explaining black box models. ACM Comput Surv 51:93

Guinney J, Saez-Rodriguez J (2018) Alternative models for sharing confidential biomedical data. Nat Biotechnol 36(5):391–392

Gunther S et al (2007) SuperTarget and Matador: resources for exploring drug–target relationships. Nucleic Acids Res 36:D919–D922

Hamilton WL (2020) Graph representation learning. Synth Lect Artif Intell Mach Learn 14:1–159

Han X, Xie R, Li X, Li J (2022) SmileGNN: drug–drug interaction prediction based on the smiles and graph neural network. Life (basel). 12(2):319. https://doi.org/10.3390/life12020319

Hao M, Wang Y, Bryant SH (2016) Improved prediction of drug–target interactions using regularized least squares integrating with kernel fusion technique. Anal Chim Acta 909:41

Hassan-Harrirou H, Zhang C, Lemmin T (2020) RosENet: improving binding affinity prediction by leveraging molecular mechanics energies with an ensemble of 3D convolutional neural networks. J Chem Inf Model 60:2791–2802

He C, Liu Y, Li H, Zhang H, Mao Y, Qin X, Liu L, Zhang X (2022) Multi-type feature fusion based on graph neural network for drug-drug interaction prediction. BMC Bioinf 23(1):1–8

Hecker N et al (2011) SuperTarget goes quantitative: update on drug–target interactions. Nucleic Acids Res 40(D1):D1113–D1117

Hermanto A, Adji TB, Setiawan NA (2015) Recurrent neural network language model for English-Indonesian machine translation: experimental study. Int Conf Sci Inf Technol (ICSITech) 2015:132–136. https://doi.org/10.1109/ICSITech.2015.7407791

Hinton G (2011) Boltzmann machines. In: Sammut C, Webb GI (eds) Encyclopedia of machine learning. Springer, Boston

Hirohara M, Saito Y, Koda Y, Sato K, Sakakibara Y (2018) Convolutional neural network based on SMILES representation of compounds for detecting chemical motif. BMC Bioinform 19:83–94

Hizukuri Y, Sawada R, Yamanishi Y (2015) Predicting target proteins for drug candidate compounds based on drug-induced gene expression data in a chemical structure-independent manner. BMC Med Genomics 8:82

Hou X, You J, Hu P (2019) Predicting drug–drug interactions using deep neural network. In: proceedings of the 11th international conference on machine learning and computing, pp 168–172

http://zinc.docking.org

https://bioinf-applied.charite.de/supernatural_new/index.php.

https://friendsofcancerresearch.org/wpcontent/uploads/Optimizing_Dosing_in_Oncology_Drug_Development.pdf.

https://ncats.nih.gov/tox21

https://pharmacodb.pmgenomics.ca/datasets/4

https://sites.broadinstitute.org/ccle/

https://string-db.org/cgi/download.pl?sessionId=uKr0odAK9hPs

https://www.cancer.gov/about-nci/organization/ccct/ctrp

https://www.ebi.ac.uk/chebi/

https://www.sciencedirect.com/topics/drug-response

Hu J, Gao J, Fang X, Liu Z, Wang F, Huang W, Wu H, Zhao G (2022) DTSyn: a dual-transformer-based neural network to predict synergistic drug combinations. bioRxiv. https://doi.org/10.1101/2022.03.29.486200

Huang C-T et al (2018) A large-scale gene expression intensity-based similarity metric for drug repositioning. iScience 7:40–52

Huang K, Xiao C, Hoang TN, Glass LM, Sun J (2020) Caster: predicting drug interactions with chemical substructure representation. In: AAAI 2020 34th AAAI Conference on Artificial Intelligence, American Association for Artificial Intelligence (AAAI) Press, pp 702–709

Ibrahim H, El Kerdawy AM, Abdo A, Eldin AS (2021) Similarity-based machine learning framework for predicting safety signals of adverse drug–drug interactions. Inf Med Unlocked 26:100699

Ierapetritou M, Muzzio F, Reklaitis G (2016) Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. AIChE J 62:1846–1862

Iorio F et al (2010) Discovery of drug mode of action and drug repositioning from transcriptional responses. PNAS 107(33):14621–14626. https://doi.org/10.1073/pnas.1000138107

Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves E, Barthorpe S, Lightfoot H et al (2016) A landscape of pharmacogenomic interactions in cancer. Cell 166:740–754

James M, Stanfield CF, Bir G (2006) A review of process analytical technology (PAT) in the US pharmaceutical industry. Curr Pharm Anal 2:405–414

Ji ZL, Han LY, Yap CW, Sun LZ, Chen X, Chen YZ (2003) Drug adverse reaction target database (DART). Drug Saf 26(10):685–690

Jiménez-Luna J, Grisoni F, Schneider G (2020) Drug discovery with explainable artificial intelligence. Nat Mach Intell 2(10):573–584

Julkunen H, Cichonska A, Gautam P et al (2020) Leveraging multi-way interactions for systematic prediction of pre-clinical drug combination effects. Nat Commun 11(1):6136

Kamath U, Liu J (2021) Explainable artificial intelligence: an introduction to interpretable machine learning. Springer, Cham

Kamble R, Sharma S, Varghese V, Mahadik K (2013) Process analytical technology (PAT) in pharmaceutical development and its application. Int J Pharm Sci Rev Res 23:212–223

Kamel Boulos MN, Zhang P (2021) Digital twins: from personalised medicine to precision public health. J Person Med 11(8):745

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28(1):27–30

Karim MR, Cochez M, Jares JB, Uddin M, Beyan O, Decker S (2019) Drug–drug interaction prediction based on knowledge graph embeddings and convolutional-LSTM network. In: Proceedings of the 10th ACM international conference on bioinformatics, computational biology and health informatics, pp 113–123

Karim MR, Cochez M, Jares JB, Uddin M, Beyan O, Decker S (2019) Drug–drug interaction prediction based on knowledge graph embeddings and convolutional-LSTM network. In: Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics 2019, pp 113–123

Karpov P, Godin G, Tetko IV (2020) Transformer-CNN: Swiss knife for QSAR modeling and interpretation. J Cheminform 12:17

Kastrin A, Ferk P, Leskošek B (2018) Predicting potential drug–drug interactions on topological and semantic similarity features using statistical learning. PLoS ONE 13(5):e0196865

Keum J, Nam H (2017) SELF-BLM: prediction of drug–target interactions via self-training SVM. PLoS ONE 12:e0171839

Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA et al (2016) PubChem substance and compound databases. Nucleic Acids Res 44:D1202–D1213

Kim J, Park S, Min D, Kim W (2021) comprehensive survey of recent drug discovery using deep learning. Int J Mol Sci 22:9983. https://doi.org/10.3390/ijms22189983

Koes DR, Baumgartner MP, Camacho CJ (2013) Lessons learned in empirical scoring with smina from the CSAR 2011 benchmarking exercise. J Chem Inf Model 53:1893–1904

Kohonen T (1990) The self-organizing map. Proc IEEE 78(9):1464–1480

Korkmaz S (2020) Deep learning-based imbalanced data classification for drug discovery. J Chem Inf Model 60:4180–4190

Kritzinger W, Karner M, Traar G, Henjes J, Sihn W (2018) Digital Twin in manufacturing: a categorical literature review and classification. IFAC-PapersOnLine 51:1016–1022

Kuenzi BM et al (2020) Predicting drug response and synergy using a deep learning model of human cancer cells. J Elsevier Cancer Cell 38(5):1535–6108. https://doi.org/10.1016/j.ccell.2020.09.014

Kuhn M et al (2010) A side effect resource to capture phenotypic effects of drugs. Mol Syst Biol 6(1):343

Kuhn M et al (2013) STITCH 4: integration of protein–chemical interactions with user data. Nucleic Acids Res 42(D1):D401–D407

Kumar SP, Feidler JC (2003) BioSPICE: a computational infrastructure for integrative biology. OMICS J Integr Biol 7(3):225. https://doi.org/10.1089/153623103322452350

Kumar S, Talasila D, Gowrav M, Gangadharappa H (2020) Adaptations of pharma 4.0 from industry 4.0. Drug Invent Today 14:405–415

Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN et al (2006) The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. Science 313:1929–1935

Lapuschkin S et al (2019) Unmasking clever Hans predictors and assessing what machines really learn. Nat Commun 10:1096

Lee CY, Chen YP (2021) Descriptive prediction of drug side-effects using a hybrid deep learning model. Int J Intell Syst 36(6):2491–2510

Lee H, Kim W (2019) Comparison of target features for predicting drug–target interactions by deep neural network based on large-scale drug-induced transcriptome data. Pharmaceutics 11:377

Lee HW, Christie A, Xu J, Yoon S (2012) Data fusion-based assessment of raw materials in mammalian cell culture. Biotechnol Bioeng 109:2819–2828

Lee G, Park C, Ahn J (2019) Novel deep learning model for more accurate prediction of drug–drug interaction effects. BMC Bioinform 20(1):415

Lee I, Keum J, Nam H (2019) DeepConv-DTI: prediction of drug–target interactions via deep learning with convolution on protein sequences. PLoS Comput Biol 15:1–21

Legner C, Eymann T, Hess T, Matt C, Böhmann T, Drews P, Mädche A, Urbach N, Ahlemann F (2017) Digitalization: opportunity and challenge for the business and information systems engineering community. Bus Inf Syst Eng 59:301–308

Lei T, Barzilay R, Jaakkola T (2016) Rationalizing neural predictions. In: 2016 conference on empirical methods in natural language processing, 2016; Austin, Texas: Association for computational linguistics, pp 107—117. https://aclanthology.org/D16-1011

Li M, Wang Y, Zheng R, Shi X, Wu F, Wang J, et al. (2019) Deepdsc: a deep learning method to predict drug sensitivity of cancer cell lines. IEEE/ACM transactions on computational biology and bioinformatics

Lian M, Du W, Wang X, Yao Q (2021) Drug–target interaction prediction based on multi-similarity fusion and sparse dual-graph regularized matrix factorization. IEEE Access 9:99718–99730. https://doi.org/10.1109/ACCESS.2021.3096830

Lin X, Quan Z, Wang Z-J, Ma T, Zeng X (2021) KGNN: knowledge graph neural network for drug–drug interaction prediction. In: Proceedings of the twenty-ninth international joint conference on artificial intelligence, Jaban; IJCAI'20

Lin-Gibson S, Srinivasan V (2019) Recent industrial roadmaps to enable smart manufacturing of biopharmaceuticals. IEEE Trans Autom Sci Eng 2019:1–8

Lipton ZC (2018) The mythos of model interpretability. Queue 16:31–57

Liu Y, Wu M, Miao C, Zhao P, Li X-L (2016) Neighborhood regularized logistic matrix factorization for drug–target interaction prediction. PLoS Comput Biol 12:e1004760

Liu B, Ramsundar B, Kawthekar P, Shi J, Gomes J, Luu Nguyen Q, Ho S, Sloane J, Wender P, Pande V (2017) Retrosynthetic reaction prediction using neural sequence-to-sequence models. R ACS Cent Sci 3:1103–1113

Liu N, Chen CB, Kumara S (2019) Semi-supervised learning algorithm for identifying high-priority drug–drug interactions. IEEE J Biomedic Health Inform. https://doi.org/10.1109/JBHI.2019.2932740

Liu K, Sun X, Jia L, Ma J, Xing H, Wu J, Gao H, Sun Y, Boulnois F, Fan J (2019a) Chemi-net: a molecular graph convolutional network for accurate drug property prediction. Int J Mol Sci 20:3389

Liu P, Li H, Li S, Leung KS (2019b) Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network. BMC Bioinform 20:408

Liu S, Huang Z, Qiu Y, Chen Y-PP, Zhang W (2019c) Structural network embedding using multi-modal deep auto-encoders for predicting drug–drug interactions. IEEE Int Conf Bioinform Biomed 2019:445–450. https://doi.org/10.1109/BIBM47256.2019.8983337

Liu S, Zhang Y, Cui Y, Qiu Y, Deng Y, Zhang W, Zhang Z (2021) Enhancing drug–drug interaction prediction using deep attention neural networks. BioRxiv. https://doi.org/10.1101/2021.03.16.435553

Lopes MR, Costigliola A, Pinto R, Vieira S, Sousa JMC (2019) Pharmaceutical quality control laboratory digital twin—a novel governance model for resource planning and scheduling. Int J Prod Res 58:1–15

Louizos C, Welling M, Kingma DP (2017) Learning sparse neural networks through l 0 regularization. http://arxiv.org/abs/1712.01312.

Lu Y, Guo Y, Korhonen AJB (2017) Link prediction in drug–target interactions network using similarity indices. BMC Bioinf 18(1):39. https://doi.org/10.1186/s12859-017-1460-z

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

Luo D, Cheng W, Xu D, Yu W, Zong B, Chen H, Zhang X (2020) Parameterized explainer for graph neural network. Adv Neural Inf Process Syst 33:19620–19631

Lyu T, Gao J, Tian L, Li Z, Zhang P, Zhang J (2021) MDNN: a multimodal deep neural network for predicting drug–drug interaction events. In: Proceedings of the thirtieth international joint conference on artificial intelligence (IJCAI-21), pp 3536–3542. https://doi.org/10.24963/ijcai.2021/487

Ma T, Xiao C, Zhou J, Wang F (2018) Drug similarity integration through attentive Multiview graph auto-encoders. In: IJCAI 2018, proceedings of the 27th international joint conference on artificial intelligence, pp 3477–3483

Mahajan D, Kumar D (2018) Sentiment analysis using RNN and Google translator. In: 2018 8th international conference on cloud computing, data science & engineering (Confluence), pp 798–802. https://doi.org/10.1109/CONFLUENCE.2018.8442924

Mak IWY, Evaniew N, Ghert M (2014) Lost in translation: animal models and clinical trials in cancer treatment. Am J Transl Res 6:114–118

Marr B (2017) What is digital twin technology and why is it so important? Forbes. https://www.forbes.com/sites/bernardmarr/2017/03/06/what-is-digital-twin-technology-and-why-is-it-so-important

Matsuzaka Y, Uesawa Y (2019) Prediction model with high-performance constitutive androstane receptor (CAR) using DeepSnap-deep learning approach from the tox21 10K compound library. Int J Mol Sci 20:4855

Maul J-T, Djamei V, Kolios AG, Meier B, Czernielewskiand J, Jungo P (2016) Efficacy and survival of systemic psoriasis treatments: an analysis of the SWISS registry SDNTT. Dermatology 232(6):640–647

Mayani MG, Svendsen M, Oedegaard SI (2018) Drilling digital twin success stories the last 10 years. In: Proceedings of the SPE Norway one day seminar, Bergen, Norway. https://doi.org/10.2118/191336-MS

Metz JT, Johnson EF, Soni NB, Merta PJ, Kifle L, Hajduk PJ (2011) Navigating the kinome. Nat Chem Biol 7:200–202

Miller T (2019) Explanation in artificial intelligence: insights from the social sciences. Artif Intell 267:1–38

Miyato T, Dai AM, Goodfellow I (2016) Adversarial training methods for semisupervised text classification. http://arxiv.org/abs/1605.07725

Mohamed C, Nsiri B, Abdelmajid S, Abdelghani EM, Brahim B (2020) Deep convolutional networks for image segmentation: application to optic disc detection. Int Conf Electr Inf Technol (ICEIT) 2020:1–3. https://doi.org/10.1109/ICEIT48248.2020.9113204

Mukhamediev RI, Symagulov A, Kuchin Y, Yakunin K, Yelis M (2021) From classical machine learning to deep neural networks: a simplified scientometric review. Appl Sci 11:5541. https://doi.org/10.3390/app11125541

Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, Yu B (2019) Definitions, methods, and applications in interpretable machine learning. Proc Natl Acad Sci USA 116:22071–22080

Nag S, Baidya ATK, Mandal A et al (2022) Deep learning tools for advancing drug discovery and development. 3 Biotech 12:110. https://doi.org/10.1007/s13205-022-03165-8

Nagy ZK, Fevotte G, Kramer H, Simon LL (2013) Recent advances in the monitoring, modelling, and control of crystallization systems. Chem Eng Res Des 91:1903–1922

Narayanan H, Luna MF, von Stosch M, Cruz Bournazou MN, Polotti G, Morbidelli M, Butte A, Sokolov M (2020) Bioprocessing in the digital age: the role of process models. Biotechnol J 15:e1900172

Nascimento ACA, Prudêncio RBC, Costa IG (2016) A multiple kernel learning algorithm for drug–target interaction prediction. BMC Bioinforma 17:46

Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 48(3):443–453

Nguyen T, Nguyen TT, Nguyen T, Le DH (2021) Graph convolutional networks for drug response prediction. IEEE/ACM Trans Comput Biol Bioinform 19:146–154

O'Connor TF, Yu LX, Lee SL (2016) Emerging technology: a key enabler for modernizing pharmaceutical manufacturing and advancing product quality. Int J Pharm 509:492–498

Oboyle NM, Sayle RA (2016) Comparing structural fingerprints using a literature-based similarity benchmark. J Cheminform 8(1):1–14. https://doi.org/10.1186/s13321-016-0148-0

Olughu W, Deepika G, Hewitt C, Rielly C (2019) Insight into the large-scale upstream fermentation environment using scaled-down models. J Chem Technol Biotechnol 94:647–657

Oughtred R, Rust J, Chang C, Breitkreutz BJ, Stark C, Willems A, Boucher L, Leung G, Kolas N, Zhang F, Dolma S (2021) The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Sci 30(1):187–200

Oztemel E, Gursev S (2018) Literature review of Industry 4.0 and related technologies. J Intell Manuf 31:127–182

Ozturk H, Ozturk A, Ozkirimli E (2018) DeepDTA: Deep drug–target binding affinity prediction. Bioinformatics 34:i821–i829

Pandey P, Katakdaunde M, Turton R (2006) Modeling weight variability in a pan coating process using Monte Carlo simulations. AAPS Pharm Sci Tech 7:E2–E11

Papadakis E, Woodley JM, Gani R (2018) Perspective on PSE in pharmaceutical process development and innovation. In Process. Systems engineering for pharmaceutical manufacturing. Elsevier, Amsterdam pp 597–656

Passi A et al (2018) RepTB: a gene ontology-based drug repurposing approach for tuberculosis. J Cheminform 10(1):24. https://doi.org/10.1186/s13321-018-0276-9

Peng J, Li J, Shang X (2020) A learning-based method for drug–target interaction prediction based on feature representation learning and deep neural network. BMC Bioinform 21:1–13

Perozzi B, Al-Rfou R, Skiena S (2014) DeepWalk: online learning of social representations. In: Proceeding of the ACM SIGKDD international conference on knowledge discovery and data mining, New York, NY, USA, 24–27 August 2014, pp 701–710

Poluzzi E, Raschi E, Piccinni C, De Ponti F (2012) data mining techniques in pharmacovigilance: analysis of the publicly accessible FDA adverse event reporting system (AERS). In: Data mining applications in engineering and medicine. London, United Kingdom: IntechOpen. https://doi.org/10.5772/50095

Pouryahya M, Oh JH, Mathews JC, Belkhatir Z, Moosmüller C, Deasy JO, Tannenbaum AR (2022) Pan-cancer prediction of cell-line drug sensitivity using network-based methods. Int J Mol Sci 23:1074. https://doi.org/10.3390/ijms23031074

Qiu K, Lee J, Kim H, Yoon S, Kang K (2021) Machine learning based anti-cancer drug response prediction and search for predictor genes using cancer cell line gene expression. Genomics Inform. https://doi.org/10.5808/gi.20076

Quan C et al (2016) Multichannel convolutional neural network for biological relation extraction. BioMed Res Int. https://doi.org/10.1155/2016/1850404

Raghava GP, Barton GJ (2006) Quantification of the variation in percentage identity for protein sequence alignments. BMC Bioinf 7(1):415. https://doi.org/10.1186/1471-2105-7-415

Rampášek L et al (2019) Improving drug response prediction via modeling of drug perturbation effects. Bioinformatics. https://doi.org/10.1093/bioinformatics/btz158

Rantanen J, Khinast J (2015) The future of pharmaceutical manufacturing sciences. J Pharm Sci 104:3612–3638

Read EK, Park JT, Shah RB, Riley BS, Brorson KA, Rathore AS (2010) Process analytical technology (PAT) for biopharmaceutical products: Part I. Concepts and applications. Biotechnol Bioeng 105:276–284

Reinhardt IC, Oliveira DJC, Ring DDT (2020) Current perspectives on the development of industry 4.0 in the pharmaceutical sector. J Ind Inf Integr 18:100131

Ren S, Tao Y, Yu K et al (2022) De novo prediction of Cell-Drug sensitivities using deep learning-based graph regularized matrix factorization. Pacif Symp Biocomput. https://doi.org/10.7490/f1000research.1118807.1

Reza F, Reza S, Yadollah O (2017) Computational prediction of drug–drug interactions based on drugs functional similarities. J Biomed Inform 70:54–64

Richardson P, Grifn I, Tucker C, Smith D, Oechsle O, Phelan A, Rawling M, Savory E, Stebbing J (2020) Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. Lancet (london, England) 395(10223):e30

Rifaioglu AS, Atas H, Martin MJ, Cetin-Atalay R, Atalay V, Dogan T (2019) Recent applications of deep learning and machine intelligence on in silico drug discovery: methods, tools and databases. Brief Bioinform 20:1878–1912

Rosen R, von Wichert G, Lo G, Bettenhausen KD (2015) About the importance of autonomy and digital twins for the future of manufacturing. IFAC-PapersOnLine 48:567–572

Ryu JY, Kim HU, Lee SY (2018) Deep learning improves prediction of drug–drug and drug–food interactions. PNAS 115(18):E4304–E4311

Sachdev K, Gupta MK (2019) A comprehensive review of feature-based methods for drug–target interaction prediction. J Biomed Inform 93:103159

Sajjia M, Shirazian S, Kelly CB, Albadarin AB, Walker G (2017) ANN analysis of a roller compaction process; in the pharmaceutical industry. Chem Eng Technol 40:487–492

Sarker IH (2021) Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN Comput Sci 2:420. https://doi.org/10.1007/s42979-021-00815-1

Sawada R, Iwata M, Tabei Y, Yamato H, Yamanishi Y (2018) Predicting inhibitory and activatory drug targets by chemically and genetically perturbed transcriptome signatures. Sci Rep 8:156

Schleich B, Anwer N, Mathieu L, Wartzack S (2017) Shaping the digital twin for design and production engineering. CIRP Ann 66:141–144

Schlichtkrull MS, De Cao N, Titov I (2020) Interpreting graph neural networks for NLP with differentiable edge masking. http://arxiv.org/abs/2010.00577

Schwarz K (2021) AttentionDDI: Siamese attention-based deep learning method for drug–drug interaction predictions. BMC Bioinf 22(1):412

Scudellari M (2020) Five companies using AI to fight coronavirus. https://spectrum.ieee.org/the-human-os/artificial-intelligence/medical-ai/companies-ai-coronavirus

Seo S, Lee T, Kim MH, Yoon Y (2020) Prediction of side effects using comprehensive similarity measures. BioMed Res Int. https://doi.org/10.1155/2020/1357630

Shang C, Liu Q, Chen KS, Sun J, Lu J, Yi J, Bi J (2018) Edge attention-based multi-relational graph convolutional networks. arXiv 2018; arXiv:1802.04944.

Shao K, Zhang Z, He S, Bo X (2020) DTIGCCN: prediction of drug–target interactions based on GCN and CNN. In: Proceedings of the 2020 IEEE 2nd international conference on tools with artificial intelligence (ICTAI), Baltimore, MD, USA, 9–11 November 2020, pp 337–342

Sharifi-Noghabi H, Zolotareva O, Collins CC, Ester M (2019) MOLI: multi-omics late integration with deep neural networks for drug response prediction. Bioinformatics 35:i501–i509

Shin B, Park S, Kang K, Ho JC (2019) Self-attention based molecule representation for predicting drug–target interaction. Proc Mach Learn Res 106:1–18

Shoemaker RH (2006) The NCI60 human tumour cell line anticancer drug screen. Nat Rev Cancer 6:813–823

Shrikumar A, Greenside P, Kundaje A (2017) Learning important features through propagating activation differences. In: Proceedings of the 34th international conference on machine learning 2017; 70, JMLR.org: Sydney, NSW, Australia. pp 3145–3153

Shtar G, Rokach L, Shapira B (2019) Detecting drug–drug interactions using artificial neural networks and classic graph similarity measures. PLoS ONE 14(8):e0219796

Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A et al (2017) Mastering the game of go without human knowledge. Nature 550(7676):354–359

Simon LL, Kiss AA, Cornevin J, Gani R (2019) Process engineering advances in pharmaceutical and chemical industries: Digital process design, advanced rectification, and continuous filtration. Curr Opin Chem Eng 25:114–121

Simonyan K, Vedaldi A, Zisserman A (2014) Deep inside convolutional networks: visualising image classification models and saliency maps. In: 2nd international conference on learning representations, ICLR 2014, Banff, AB, Canada, April 14–16, 2014, Workshop Track Proceedings; http://arxiv.org/abs/1312.6034

Smiatek J, Jung A, Bluhmki E (2020) Towards a digital bioprocess. Replica: computational approaches in biopharmaceutical development and manufacturing. Trends Biotechnol 38(10):1141–1153. https://doi.org/10.1016/j.tibtech.2020.05.008

Song T, Zhang X, Ding M, Rodriguez-Paton A, Wang S, Wang G (2022) DeepFusion: a deep learning based multi-scale feature fusion method for predicting drug–target interactions. Methods 204:269–277

Springenberg JT (2015) Striving for simplicity: the all-convolutional Net. CoRR, http://arxiv.org/abs/1412.6806

Stark R, Fresemann C, Lindow K (2019) Development and operation of digital twins for technical systems and services. CIRP Ann 68:129–132

Steinwandter V, Borchert D, Herwig C (2019) Data science tools and applications on the way to Pharma 4.0. Drug Discov Today 24:1795–1805

Stokes JM, Yang K, Swanson K, Jin W, Cubillos-Ruiz A, Donghia NM, MacNair CR, French S, Carfrae LA, Bloom-Ackerman Z et al (2020) A deep learning approach to antibiotic discovery. Cell 180:688-702.e13

Subramanian K (2020) Digital twin for drug discovery and development—the virtual liver. J Indian Inst Sci 100:653–662. https://doi.org/10.1007/s41745-020-00185-2

Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, Gould J, Davis JF, Tubelli AA, Asiedu JK et al (2017) A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. Cell 171:1437-1452.e17

Sun X, Ma L, Du X, Feng J, Dong K (2018) Deep convolution neural networks for drug–drug interaction extraction. In: 2018 IEEE international conference on bioinformatics and biomedicine (BIBM), pp 1662–1668. https://doi.org/10.1109/BIBM.2018.8621405

Sun M, Zhao S, Gilvary C, Elemento O, Zhou J, Wang F (2020a) Graph convolutional networks for computational drug development and discovery. Brief Bioinform 21:919–935

Sun M, Wang F, Elemento O, Zhou J (2020b) Structure-based drug–drug interaction detection via expressive graph convolutional networks and deep sets. Proc AAAI Conf Artif Intell 34(10):13927–13928. https://doi.org/10.1609/aaai.v34i10.7236

System HSL (2006) Psychoactive Drug Screening Program. https://www.hsls.pitt.edu/obrc/index.php?page=URL1133202727

Tajbakhsh N et al (2016) Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans Med Imaging 35(5):1299–1312. https://doi.org/10.1109/TMI.2016.2535302

Tang J, Szwajda A, Shakyawar S, Xu T, Hintsanen P, Wennerberg K, Aittokallio T (2014) Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. J Chem Inf Model 54:735–743

Tang P, Xu J, Louey A, Tan Z, Yongky A, Liang S, Li ZJ, Weng Y, Liu S (2020) Kinetic modeling of Chinese hamster ovary cell culture: factors and principles. Crit Rev Biotechnol 40:265–281

Tao F, Cheng J, Qi Q, Zhang M, Zhang H, Sui F (2018) Digital twin-driven product design, manufacturing and service with big data. Int J Adv Manuf Technol 94:3563–3576

Tatonetti NP et al (2012) Data-driven prediction of drug effects and interactions. Sci Transl Med 4(125):12531. https://doi.org/10.1126/scitranslmed.3003377

Tatonetti NP, Patrick PY, Daneshjou R, Altman RB (2012) Data driven prediction of drug effects and interactions. Sci Transl Med 4(125):125ra31-125ra31

Tehseen Z, Usman Z (2019) Long short-term memory recurrent neural network architectures for Urdu acoustic modelling. Int J Speech Technol 22(1):21–30. https://doi.org/10.1007/s10772-018-09573-7

Thafar M, Raies AB, Albaradei S, Essack M, Bajic VB (2019) Comparison study of computational prediction tools for drug–target binding affinities. Front Chem 7:782. https://doi.org/10.3389/fchem.2019.00782

Thafar MA, Olayan RS, Olayan RS, Ashoor H, Ashoor H, Albaradei S, Albaradei S, Bajic VB, Gao X et al (2020) DTiGEMS: drug–target interaction prediction using graph embedding, graph mining, and similarity-based techniques. J Cheminform 12:1–17

Thafar MA, Alshahrani M, Albaradei S et al (2022) Affinity2Vec: drug–target binding affinity prediction through representation learning, graph mining, and machine learning. Sci Rep 12:4751. https://doi.org/10.1038/s41598-022-08787-9

Thorben F, Megha Kh, Avishek A (2021) Hard masking for explaining graph neural networks. In Submitted to international conference on learning representations https://openreview.net/forum?id=uDN8pRAdsoC

Tian X, Xin M, Luo J, Jiang Z (2016) Using the ranking-based KNN approach for drug repositioning based on multiple information. Springer, Cham, pp 317–327

Tong H, Heidemeyer M, Ban F, Cherkasov A, Ester M (2017) SimBoost: A read-across approach for predicting drug–target binding affinities using gradient boosting machines. J Cheminform 9:1–14

Torng W, Altman RB (2019) Graph convolutional neural networks for predicting drug–target interactions. J Chem Inf Model 59:4131–4149

Townshend RJL, Powers A, Eismann S, Derry A (2021) ATOM3D: tasks on molecules in three dimensions. arXiv 2021: arXiv:2012.04035

Trißl S, Rother K, Müller H et al (2005) Columba: an integrated database of proteins, structures, and annotations. BMC Bioinformatics 6:81. https://doi.org/10.1186/1471-2105-6-81

Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem 31:455

Tyson RJ, Park CC, Powell JR, Patterson JH, Weiner D, Watkins PB, Gonzalez D (2020) Precision dosing priority criteria: drug, disease, and patient population variables. J Front Pharmacol. https://doi.org/10.3389/fphar.2020.00420

U. Consortium (2014) UniProt: a hub for protein information. Nucleic Acids Res 43(D1):D204–D212

Vazquez J, Lopez M, Gibert E, Herrero E, Luque FJ (2020) Merging ligand-based and structure-based methods in drug discovery: an overview of combined virtual screening approaches. Molecules 25:4723

Venkatasubramanian V (2019) The promise of artificial intelligence in chemical engineering: is it here, finally? AIChE J 65:466–478

Vermeer NS, Straus SM, Mantel-Teeuwisse AK, Domergue F, Egberts TC, Leufkens HG, De Bruin ML (2013) Traceability of biopharmaceuticals in spontaneous reporting systems: a cross sectional study in the FDA adverse event reporting system (FAERS) and surveillance databases. Drug Saf 36(8):617–625

Vilar S, Hripcsak GJ (2016) Leveraging 3D chemical similarity, target and phenotypic data in the identification of drug-protein and drug-adverse effect associations. J Cheminform 8(1):35. https://doi.org/10.1186/s13321-016-0147-1

Vilar S, Uriarte E, Santana L, Lorberbaum T, Hripcsak G, Friedman C, Tatonetti NP (2014) Similarity-based modeling in large-scale prediction of drug–drug interactions. Nat Protoc 9(9):2147–2163. https://doi.org/10.1038/nprot.2014.151

Wallach I, Dzamba M, Heifets A (2015) AtomNet: a deep convolutional neural network for bioactivity prediction in structurebased drug discovery. arXiv 2015: arXiv:1510.02855.

Wan F et al (2019) DeepCPI: a deep learning-based framework for large-scale in silico drug screening. Genom Proteomics Bioinform 17:478–495

Wang JZ et al (2007) A new method to measure the semantic similarity of GO terms. Bioinformatics 23(10):1274–1281. https://doi.org/10.1093/bioinformatics/btm087

Wang W et al (2014) Drug repositioning by integrating target information through a heterogeneous network model. Bioinformatics 30(20):2923–2930. https://doi.org/10.1093/bioinformatics/btu403

Wang CS, Lin PJ, Cheng CL, Tai SH, Kao Yang YH, Chiang JH (2019) Detecting potential adverse drug reactions using a deep neural network model. J Med Internet Res 21(2):e11016

Wang T, Yi HC, You ZH, Li LP, Wang YB, Hu L, Wong L (2019) A gated recurrent unit model for drug repositioning by combining comprehensive similarity measures and Gaussian interaction profile kernel. In: International conference on intelligent computing. Springer, Cham. pp 344–353

Wang YB, You ZH, Yang S et al (2020a) A deep learning-based method for drug–target interaction prediction based on long short-term memory neural network. BMC Med Inform Decis Mak 20:49. https://doi.org/10.1186/s12911-020-1052-0

Wang H, Wang J, Dong C, Lian Y, Liu D, Yan Z (2020b) A novel approach for drug–target interactions prediction based on multimodal deep autoencoder. Front Pharmacol 10:1–19

Watanabe JH, McInnis T, Hirsch JD (2018) Cost of prescription drug-related morbidity and mortality. Ann Pharmacother 52:829–837. https://doi.org/10.1177/1060028018765159

Way GP, Greene CS (2018) Extracting a biologically relevant latent space from cancer transcriptomes with variational autoencoders. Pac Symp Biocomput 23:80–91

Wei J, Lu Z, Qiu K, Li P, Sun H (2020) Predicting drug risk level from adverse drug reactions using SMOTE and machine learning approaches. IEEE Access 8:185761–185775. https://doi.org/10.1109/ACCESS.2020.3029446

Weinstein JN (2004) Integromic analysis of the NCI-60 cancer cell lines. Breast Dis 19:11–22

Wen M, Zhang Z, Niu S, Sha H, Yang R, Yun Y, Lu H (2017) Deep-learning-based drug–target interaction prediction. J Proteome Res 16:1401–1409

Wenzel J, Matter H, Schmidt F (2019) Predictive multitask deep neural network models for adme-tox properties: learning from large data sets. J Chem Inf Model 59:1253–1268

White J, Schiffer JT, Bender R et al (2021) Drug combinations as a first line of defense against coronaviruses and other emerging viruses. Mbio 12(6):e0334721

Withnall M, Lindelöf E, Engkvist O, Chen H (2020) Building attention and edge message passing neural networks for bioactivity and physical-chemical property prediction. J Cheminform 12:1

Wu Z, Ramsundar B, Feinberg EN, Gomes J, Geniesse C, Pappu AS, Leswing K, Pande V (2018) MoleculeNet: a benchmark for molecular machine learning. Chem Sci 9:513–530

Wu Z, Pan S, Chen F et al (2020) A comprehensive survey on graph neural networks. IEEE Trans Neural Netw Learn Syst 32:4–24

Xia Z, Wu LY, Zhou X, Wong ST (2010) Semi-supervised drug-protein interaction prediction from heterogeneous biological spaces. BMC Syst Biol 4:S6

Xiang W, Yingxin W, An Z, Xiangnan H, Tat-seng C (2021) Causal screening to interpret graph neural networks. In Submitted to international conference on learning representations. https://www.openreview.net/forum?id=nzKv5vxZfge

Xie L, He S, Song X, Bo X, Zhang Z (2018) Deep learning-based transcriptome data classification for drug–target interaction prediction. BMC Genomics 19:13–16

Xie Y, Peng J, Zhou Y, et al (2019) Integrating protein-protein interaction information into drug response prediction by graph neural encoding. 16 December 2019, Available at Research Square https://doi.org/10.21203/rs.2.18936/v1.

Xu Y, Pei J, Lai L (2017) Deep learning-based regression and multiclass models for acute oral toxicity prediction with automatic chemical feature extraction. J Chem Inf Model 57:2672–2685

Yan CK, Wang WX, Zhang G et al (2019) BiRWDDA: a novel drug repositioning method based on multisimilarity fusion. J Comput Biol 26(11):1230–1242

Yan C, Duan G, Zhang Y, Wu F-X, Pan Y, Wang J (2022) Predicting drug–drug interactions based on integrated similarity and semi-supervised learning. IEEE/ACM Trans Comput Biol Bioinf 19(1):168–179. https://doi.org/10.1109/TCBB.2020.2988018

Yang K, Swanson K, Jin W, Coley C, Eiden P, Gao H, Guzman-Perez A, Hopper T, Kelley B, Mathea M et al (2019) Analyzing learned molecular representations for property prediction. J Chem Inf Model 59:3370–3388

Yi HC, You ZH, Wang L et al (2021) In silico drug repositioning using deep learning and comprehensive similarity measures. BMC Bioinf 22:293. https://doi.org/10.1186/s12859-020-03882-y

Yifan D, Xinran X, Yang Q, Jingbo X, Wen Z, Shichao L (2020) A multimodal deep learning framework for predicting drug–drug interaction events. Bioinformatics 36:4316–4322

Ying Z, Bourgeois D, You J, Zitnik M, Leskovec J (2019) Gnnexplainer: generating explanations for graph neural networks. Adv Neural Inf Process Syst 32:9244–9255

Yu Y, Si X, Hu C, Zhang J (2019) A review of recurrent neural networks: Lstm cells and network architectures. Neural Comput 31:1235–1270

Yu Y, Huang K, Zhang C, Glass LM, Sun J, Xiao C (2021) SumGNN: multi-typed drug interaction prediction via efficient knowledge graph summarization. Bioinformatics 37(18):2988–2995

Yuan H, Yu H, Wang J, Li K, Ji S (2021) On explain-ability of graph neural networks via subgraph explorations. http://arxiv.org/abs/2102.05152

Yue X, Wang Z, Huang J, Parthasarathy S, Moosavinasab S, Huang Y, Lin SM, Zhang W, Zhang P, Sun H (2020) Graph embedding on biomedical networks: methods, applications, and evaluations. Bioinformatics 36(4):1241–1251. https://doi.org/10.1093/bioinformatics/btz718

Yunsheng B, Ken G, Yizhou S, Wei W (2020) Bi-level graph neural networks for drug–drug interaction prediction. J Comput Eng arXiv:2006.14002

Zaikis D, Vlahavas I (2020) Drug–drug interaction classification using attention based neural networks. In: 11th Hellenic conference on artificial intelligence, pp 34–40. https://doi.org/10.1145/3411408.3411461

Zeng H, Qiu C, Cui QJD (2015) Drug-path: a database for drug-induced pathways. J Biol Databases Curation. https://doi.org/10.1093/database/bav061

Zeng T, Rongjian L, Ravi M, Jieping Y, Shuiwang J (2015) Deep convolutional neural networks for annotating gene expression patterns in the mouse brain. BMC Bioinformatics 16(1):147

Zeng X et al (2019) Measure clinical drug–drug similarity using electronic medical records. Int J Med Inf 124:97–103. https://doi.org/10.1016/j.ijmedinf.2019.02.003

Zeng X, Zhu S, Lu W, Liu Z, Huang J, Zhou Y, Fang J, Huang Y, Guo H, Li L et al (2020) Target identification among known drugs by deep learning from heterogeneous networks. Chem Sci 11:1775–1797

Zhai J, Zhang S, Chen J, He Q (2018) Autoencoder and its various variants. In: 2018 IEEE international conference on systems, man, and cybernetics (SMC), pp 415–419. https://doi.org/10.1109/SMC.2018.00080

Zhang Y (2020) Predicting drug–drug interactions using multi-modal deep autoencoders based network embedding and positive-unlabeled learning. Methods 179:37–46

Zhang M-L, Zhou Z-H (2007) Ml-knn: a lazy learning approach to multi-label learning. Pattern Recogn 40(7):2038–2048

Zhang H, Liu D, Xiong Z (2018) Convolutional neural network-based video super-resolution for action recognition. In: 2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018), pp 746–750. https://doi.org/10.1109/FG.2018.00117

Zhang Y, Weng Y, Lund J (2022) Applications of explainable artificial intelligence in diagnosis and surgery. Diagnostics 12:237. https://doi.org/10.3390/diagnostics12020237

Zhang C, Lu Y, Zang T (2022) CNN-DDI: a learning-based method for predicting drug–drug interactions using convolution neural networks. BMC Bioinf 23:88. https://doi.org/10.1186/s12859-022-04612-2

Zhao Y, Zheng K, Guan B, Guo M, Song L, Gao J, Qu H, Wang Y, Shi D, Zhang Y (2020) DLDTI: a learning-based framework for drug–target interaction identification using neural networks and network representation. J Transl Med 18:434

Zhao Q, Xiao F, Yang M, Li Y, Wang J (2019) AttentionDTA: prediction of drug–target binding affinity using attention model. In: Proceedings of the 2019 IEEE international conference on bioinformatics and biomedicine, San Diego, CA, USA, 18–21 November 2019, pp 64–69

Zhou Y, Zhang Y, Lian X, Li F, Wang C, Zhu F, Qiu Y, Chen Y (2022) Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents. Nucleic Acids Res 50:1398–1407

Zitnik M, Agrawal M, Leskovec J (2018) Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics 34(13):i457–i466

Zitnik SM, Sosic R, Leskovec J (2018) Biosnap datasets: Stanford biomedical network dataset collection. http://snap.stanford.edu/biodata

Zong N, Kim H, Ngo V, Harismendy O (2017) Deep mining heterogeneous networks of biomedical linked data to predict novel drug–target associations. Bioinformatics 33:2337–2344

Zügner D, Akbarnejad A, Günnemann S (2018) Adversarial attacks on neural networks for graph data. In: Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery and Data Mining. 2018, Association for Computing Machinery: London, United Kingdom. pp 2847–2856

Authors and Affiliations

Heba Askr[1] · Enas Elgeldawi[2] · Heba Aboul Ella[4] · Yaseen A. M. M. Elshaier[5] · Mamdouh M. Gomaa[2] · Aboul Ella Hassanien[3]

　Heba Askr
　heba.askr@fcai.usc.edu.eg

　Enas Elgeldawi
　enas.elgeldawi@mu.edu.eg

　Heba Aboul Ella
　haboelella@ecu.edu.eg

　Yaseen A. M. M. Elshaier
　Yaseen.Elshaier@fop.usc.edu.eg

　Mamdouh M. Gomaa
　mamdouh.gomaa@mu.edu.eg

[1]　Faculty of Computers and Artificial Intelligence, University of Sadat City, Sadat City, Egypt

[2]　Computer Science Department, Faculty of Science, Minia University, Minia, Egypt

[3]　Faculty of Computers and Artificial Intelligence, Cairo University, Cairo, Egypt

[4]　Faculty of Pharmacy and Drug Technology, Chinese University in Egypt (CUE), Cairo, Egypt

[5]　Faculty of Pharmacy, University of Sadat City, Sadat City, Menoufia, Egypt