Research of Drug-Target Affinity Prediction based on Deep Learning

Anonymous submission

Abstract

The evaluation of drug-target interactions is of great significance in drug discovery and reuse, and affinity values reflect the strength of drug-target interactions. Currently, many deep learning methods are used for predicting drug-target affinities, but the vast majority of them focus on utilizing a single feature information. In response to the above issues, a drug-target affinity prediction model was designed. Firstly, the structural features of the drug/target were obtained through pre-trained model. Secondly, relationship graphs were generated based on similarity and affinity information, then extracting relationship features based on these graphs. Finally, some feature fusion technologies were used to fuse the structural and relationship features, and affinity prediction was performed based on these features. The experimental results on the Davis dataset and KIBA dataset indicate that fusing multiple feature information can effectively assist in affinity inference.

Introduction

In May 2022, the World Health Organization (WHO) released the 2022 World Health Statistics, which showed that as population growth and life expectancy increase, the total number of deaths from non communicable diseases is also increasing(Organization et al. 2022). This to some extent reflects that the supply and demand between drugs and patients have not yet reached a dynamic balance, and there is a certain gap between the two. In particular, under the influence of COVID-19, the public's demand for drugs is more urgent, and the pharmaceutical R&D industry is facing considerable pressure. However, from the perspective of the entire drug development process, a large number of experiments are required in the early stage to ensure the safety of the drug, which makes the cost of drug development high and the development speed slow. The reason for conducting large-scale experiments is to find and determine the target proteins that enable drugs to exert effective effects, which is the key to drug development and the main factor restricting the development process. Although high-throughput screening(Bajorath 2002), genomics, proteomics, and systems biology have been widely applied in drug discovery and drug reuse(Rudrapal, Khairnar, and Jadhav 2020), their assistance in large-scale biological experiments is still limited. Therefore, conducting research on drug target interaction prediction (DTI) is of great significance.

Based on the currently known drug target information, their quantity is limited and the vast majority are obtained through in vitro or biochemical experiments. PubChem(Kim et al. 2019) contains approximately 35 million compounds, with less than 7,000 containing target protein information; ChEMBL(Gaulton et al. 2017) covers over 1.9 million compounds, including over 10,000 drugs and over 12,000 targets; BindingDB(Gilson et al. 2016) contains over 7000 target data, 730,000 drug data, and over 1.65 million drug target interaction information. Although these data have important reference value, exploring unknown drug target information will have a greater driving effect on drug development. Obviously, conducting large-scale biological experiments to complete this task is undoubtedly expensive and time-consuming. Therefore, the prediction of drug target interactions based on computational methods has received widespread attention in the industry. It is worth mentioning that deep learning, as an efficient computational method, is commonly used to complete large-scale prediction or classification tasks. It extracts useful features from limited data and performs calculations; Especially with the support of strong computing power, this method can achieve its goals accurately and quickly, which greatly reduces time and technical costs.

More importantly, multiple fields in bioinformatics have been cross integrated with deep learning techniques. Because this technology can extract information from large amounts of biological data and effectively process high-dimensional and unstructured data, it can provide more accurate judgments in tasks such as gene expression analysis, protein structure prediction, disease diagnosis, etc. In fact, there have been cases of predicting drug target interactions through deep learning methods in previous literature(Thafar et al. 2019), which provides a reference for further research in this field. However, many studies view drug target interaction prediction as a binary problem, ignoring the value of drug target binding affinity. The binding affinity value is a continuous value that reflects the strength of the interaction between drug target pairs, so it cannot be ignored in studying drug target interactions.

Related Work

In this study, the focus was on investigating drug target prediction methods based on deep learning. Refer-

ence(Bagherian et al. 2021) divided them into six main branches, but overall they can be classified into three categories: feature-based methods, similarity based methods, and cross mixing methods.

The main design idea of feature based methods is to convert drugs and targets into feature vectors of a certain length through embedding techniques and input them into a prediction model. Ozturk, H. et al.(Öztürk, Özgür, and Ozkirimli 2018) used convolutional neural networks (CNNs) to embed representations of drugs and proteins, and then concatenated them and input them into deep neural networks (DNNs) for regression prediction; On this basis, Nguyen, T. et al.(Nguyen et al. 2021) used a multi type graph neural network (GNN) to embed drug molecular structure diagrams. Similarly, Jiang, M. et al.(Jiang et al. 2020) reduced the two-dimensional structure of proteins based on protein contact maps and sent them together with drug molecular structure maps to GNN for prediction. Li, M.(Li et al. 2022) designed a bidirectional attention mechanism to aggregate embedded representations of drugs and proteins from graph attention networks (GAT) and CNN, which can capture important regions of drugs and proteins. Yang et al.(Yang et al. 2022) introduced tight connections and constructed a multi-scale graph neural network based on them to obtain local and global information.

The similarity based method predicts based on the similarity between drug drug and target target. Zhang, X. et al.(Zhang et al. 2017) viewed the prediction model as an optimization problem, predicting unknown drug target interactions by identifying clusters in the drug target similarity network and maximizing cluster consistency. J. Shim et al. multiplied the similarity matrix between drugs and targets, and then input it into 2D-CNN to predict interactions. X. Ru et al.(Ru et al. 2022) extracted features using similarity and neighborhood relationships, and predicted the affinity value and priority order of drug targets through a ranking learning framework.

The cross mixing method is a combination of the above two methods, supplemented by deep learning methods or network-based methods to extract feature information and similarity information, thus requiring larger storage space during the training process. Thafar, M. A. et al.(Thafar et al. 2022) constructed a weighted isomerization graph based on drug drug similarity, target target similarity, and drug target affinity values, and used techniques such as graph mining to generate or extract features to predict the binding relationship between drugs and targets. Similarly, Shao, K. et al.(Shao et al. 2022) proposed an attention based heterogeneous graph model that utilizes GCN and GAT to obtain embedded representations of drugs and targets, and finally uses an inner product decoder for interaction prediction.

Proposed Solution

The research objective of this article is to predict the affinity between drugs and targets. The proposed deep learning based drug target affinity prediction model consists of three modules, and the framework is shown in Figure 1.

Module 1. feature extraction of drug/target structure. Select appropriate feature extraction models based on the structural characteristics of drugs and proteins; For the molecular structure of drugs, it can be viewed as a two-dimensional graph. For protein structure, it can be regarded as a one-dimensional amino acid sequence. Therefore, graph neural networks and natural language processing models can be used separately to assist feature extraction, and finally, a multi-layer feed-forward neural network (MLP) can be used to achieve linear mapping of feature dimensions.

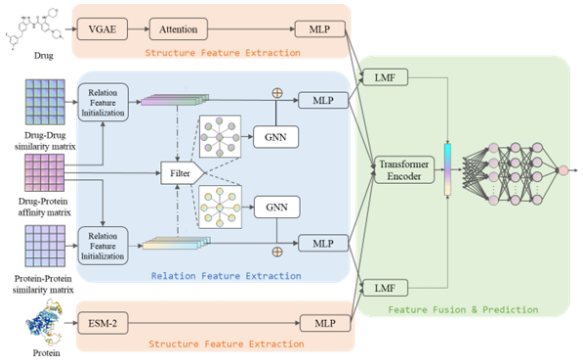


Figure 1: Structure diagram of drug target affinity prediction model based on deep learning

Module 2. Drug/Target Relationship Feature Extraction. The similarity matrix and affinity matrix of drugs/targets reflect the commonalities between drugs/targets, so relationship features can be initialized based on the similarity matrix and affinity matrix. On the other hand, similar drugs/-targets can be filtered and screened based on similarity, and a relationship graph can be constructed to extract common feature information of the drugs/targets to be predicted. This information can be concatenated as a feature onto the initial relationship feature of the drugs/targets. Similarly, it is necessary to perform linear mapping on the features for the next step of feature fusion.

Module 3. Feature Fusion and Prediction. After obtaining the structural and relational features, feature fusion technology is used to fuse the two. The design of the fusion module is divided into two schemes: the fusion of structural and relational features of drugs/targets, and the fusion of structural and relational features of drugs and targets. After the above fusion is completed, it is spliced and input into a DNN for affinity prediction.

Structural feature extraction module

Drug structure embedding. For drug structure embedding, we consider inputting the molecular structure diagram of the drug into a pre trained VGAE model for direct embedding, and then unifying the feature dimensions through a linear mapping.

The structural information of drug molecules is attached to their corresponding Simplified Molecular Linear Input Specification (SMILES). RDKit technology can convert SMILES strings into the form of molecular structure graphs, where each atom is treated as a node in the graph and the chemical bonds between atoms are treated as edges. In addition, for node features, we no longer use the unique hot

Table 1: Atomic node characteristics

id	features	dimension
1	The one-hot encoding of atoms	44
2	The one-hot encoding of Atomic Node Degree	11
3	The one-hot encoding of the total number of H bound to atoms	11
4	The one-hot encoding of implicit H-numbers bound to atoms	11
5	Does the atom have aromaticity	1
	sum	78

encoding of a single atom for initialization, but instead use the integration of other features related to the atomic structure as node features. Compared to this method, it can obtain more information, thus more comprehensively describing and representing the composition of molecules. Table 1 lists the components of atomic feature initialization. After completing the above operations, we initially obtained a drug molecular structure diagram that conforms to the graph data format.

Although the initialized structural features comprehensively aggregate information, the feature object only describes a single atom. In other words, the feature information only describes local features, and information on how nodes are connected and the distribution of molecular features has not been obtained. Obviously, for us, what is more needed is to use global features to describe the structural characteristics of the entire drug molecule. Therefore, we propose to use the VGAE model for embedding learning of drug molecular structure diagrams.

Protein Structure Embedding At present, there are various pre trained models available for embedding protein structures, and they are all trained in large databases. Therefore, in this study, ESM2 150M was selected as the protein structure pre training model, which has a total of 33 layers, a parameter quantity of 650M, and an embedded dimension of 640. Here, we use the outputs of layers 1, 32, and 33 of the model as protein structural features after average pooling. Similarly, perform a linear mapping on it at the end.

Relationship feature extraction module

The relationship feature extraction module is mainly aimed at mining valuable information based on the similarity matrix and affinity values. It can be divided into three steps. Firstly, the relationship features are initialized based on the similarity matrix. Secondly, filtering methods are used to generate similar relationship graphs. Finally, data mining is performed on the relationship graphs to obtain relationship features and concatenate them with structural features.

The initialization of relationship features is mainly based on the similarity matrix and affinity matrix. Table 2 shows the composition of relationship matrix initialization. Our main approach is to extract the statistics of similarity and affinity according to Table 2 to form new features, and then concatenate them as initial relational features.

Afterwards, in order to fully explore the common features between similar drugs/targets, an internal relationship graph of drugs/targets based on similarity and affinity was constructed. This method mainly relies on the following assumptions: (1) Similar drugs act on similar targets, and vice versa. (2) The affinity values generated by similar drugs acting on the same target are similar. (3) The affinity values generated by the action of the same drug on similar targets are similar.

Table 2: Relationship feature composition

id	features	dimension
1	mean value	1
2	25th quartile	1
3	50th quartile	1
4	75th quartile	1
5	85th quartile	1
6	95th quartile	1
7	Top 5 maximum values	5
8	Top 5 maximum values	5
	sum	16

Based on this, we introduce the method of selecting relationship nodes and GNN to assist in mining relationship features. The selection of relationship nodes is mainly based on the similarity matrix, and a star shaped relationship graph is constructed by selecting the 25 drugs/targets that are closest to the predicted drugs/targets. The star shaped relationship graph is a graph formed by connecting nodes around a single node to form edges. Here, we will use the predicted drug/target as the central node and neighboring nodes as surrounding nodes

Feature fusion and prediction module

The feature fusion module mainly uses LMF based methods and Transformer based methods. For the structural and relational features of drugs/targets, the LMF algorithm with fast calculation speed and fewer parameters is used. For the feature fusion of drugs and targets, the encoder part of the Transformer framework is used. Finally, all fused features are concatenated and sent to DNN for prediction. DNN is a three-layer neural network architecture, connected between each layer using the activation function PReLu.

Experiments

Evaluating indicator

The evaluation indicators used in the experimental section of this article are concordance index, mean square error, r_m^2, and area under the precision recall curve.

Concordance Index (CI) measures whether the predicted order of two random drug target pairs is the same as their true order, where the predicted order and true order are determined by comparing the predicted and true values of two drug target pairs, respectively (Steck et al. 2007). The higher the CI value, the higher the accuracy of the model. The calculation formula is as follows:

$$CI = \frac{1}{m} \sum_{y_i > y_j} \rho(\hat{y}_i - \hat{y}_j)$$

here,

$$\rho(x) = \begin{cases} 1, x > 0 \\ 0.5, x = 0 \\ 0, x < 0 \end{cases}$$

Where m is the standardized constant. y_i and y_j represent the true affinity values. $\hat{y}_i$ and $\hat{y}_j$ represent the predicted values of y_i and y_j, respectively.

Mean Square Error (MSE) reflects the degree of difference between the true and predicted values, with smaller MSE values indicating higher accuracy of the model. The calculation formula is as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Among them, y and $\hat{y}$ respectively contain the true and predicted values of data. y_i and $\hat{y}_i$ represent the true and predicted values of the th sample, respectively.

r_m^2 is a parameter used to evaluate the external predictive performance of QSAR models. When the r_m^2 of the test set is greater than 0.5, it indicates that the model is an acceptable, robust, and non chance obtained model (Pratim Roy et al. 2009). The higher the r_m^2, the more stable the model is.The calculation formula is as follows:

$$r_m^2 = r^2(1 - \sqrt{r^2 - r_0^2})$$

Here, r^2 represents the square of the correlation coefficient, and r_0^2 represents the square of the correlation coefficient with a intercept of zero

Area Under Precision Recall (AUPR) is mainly applied to binary prediction. Here, in order to measure the performance of the model's binary prediction, we convert quantitative datasets into binary datasets by setting affinity thresholds; Specifically, for the Davis dataset, it will be used $pK_d = 7$ as a threshold; The KIBA dataset will be set $pK_d = 12.1$ as a threshold (He et al. 2017).

Experimental Design

For the GNN module in relation feature extraction, a double-layer neural network structure of GCN+GAT is adopted; For the feature fusion module, use an 8-layer 6-head Transformer encoder and an LMF algorithm with rank=8.

The experiment used the Davis and KIBA datasets for model training, validation, and testing, and fine tuned the hyperparameters of the model using nested five fold cross validation. During the training phase, the model is iteratively trained for 1000 rounds with a batch size of 256, and the Adam algorithm is used to optimize the entire network. In the validation stage, the optimal hyperparameter combination is mainly selected based on the CI value of the model on the validation set, with a maximum iteration round of 500 set on the validation set.

The experiment in this section mainly explores the following two questions:

Firstly, how does our model perform compared to traditional algorithms. Compared with traditional drug target affinity prediction models, our model introduces structural and relational features, and completes the prediction task by organically integrating the two. This section of the experiment will compare the performance of our model with seven other affinity prediction models (KronRLS algorithm (Pahikkala et al. 2015), SimBoost algorithm (He et al. 2017), DeepDTA (Öztürk, Özgür, and Ozkirimli 2018), GraphDTA (Yang et al. 2022), FusionDTA (Yuan, Chen, and Chen 2022), MGraphDTA (Yang et al. 2022)) on different datasets.

Second, what is the contribution of each module in the model to the final performance. Our model consists of 5 modules, namely drug structure feature extraction module, target structure feature extraction module, drug relationship feature extraction module, target relationship feature extraction module, and feature fusion module. Different modules may have varying contributions to the model, and this part of the experiment will explore the impact of different modules on model performance through ablation experiments.

Results and Analysis

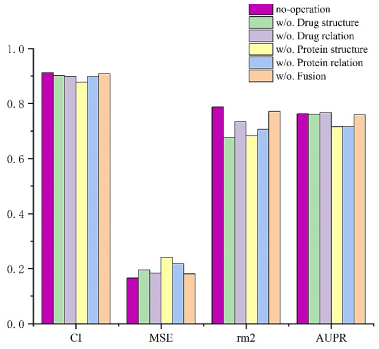
The experimental results of comparing the performance of the proposed model with other drug target affinity prediction algorithms are shown in Table 4-2. The experimental results are the average of the results obtained by each model in three experiments on each dataset, and the best performance on each dataset is highlighted in bold. The results demonstrate the following findings.

On two datasets, our model performed better than other models in predicting the affinity of unknown drugs/targets, especially on the Davis dataset. Our model achieved good improvement in both, and these indicators, and the values were close to the best performance in traditional algorithms. Specifically, in terms of, and indicators, our model performs as 0.166, 0.778, and 0.773. Compared to the second best experimental results, our model has improved by 19.8%, 6.1%, and 6.9% on these indicators, respectively; It is worth mentioning that the second best result only increased by 0.4%, 4.6%, and 0.7% compared to the third best result, and these improvements sometimes do not come from the same model. For the KIBA dataset with a large amount of data, our model performs similarly to current popular algorithms, and the overall improvement effect is not significant, with only a slight improvement in metrics. After comparing and analyzing with the Davis dataset, it is speculated that several factors may limit the performance improvement of the model on large datasets.

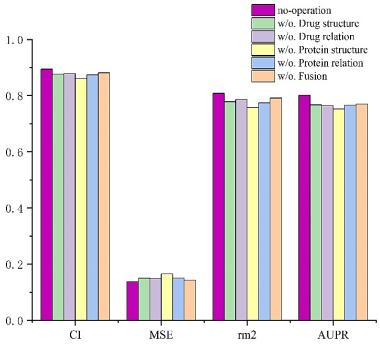
1.The protein protein similarity distribution in the KIBA dataset is clustered between 0.0 and 0.2, and the differences between them are not significant (the protein protein similarity distribution in the Davis dataset is between 0.3 and 0.7, and is approximately normal distribution). Therefore, it may affect the generated relationship features to some extent, and the model is difficult to capture these subtle differences, which may have a negative impact on the final prediction.

2. From the analysis of current popular model structures, it is found that learnable modules such as GNN, LSTM, CNN, etc. are used for extracting structural features of drugs and targets. These learnable models will focus on extracting useful feature information for affinity value prediction based

on the loss function, to some extent reducing the impact of noise on the model.



(a)



(b)

Figure 2: Comparison of ablation experimental results of modules on different datasets (a) Experimental results on Davis dataset (b) Experimental results on KIBA dataset

Table 3 shows the ablation experimental results of different modules on two datasets, and Figure 2 visualizes them separately. The experimental data shows that on the Davis dataset, protein structure embedding has the deepest impact on the model, resulting in a decrease of 3.8%, 45.2%, 15.4%, and 6.4% in the performance of various indicators, respectively. Therefore, this indicates that protein structure embedding plays a greater role in our model; In addition, from the perspective of indicators, the structural characteristics of drugs have the greatest impact on the stability of the model. The absence of structural characteristics of drugs reduces the performance of the model to 0.678, with a decrease of about 14%; It is worth mentioning that the absence of drug relationship features actually leads to a slight improvement in the model's performance on AUPR, but this improvement is not significant, only 0.004. Therefore, it can still be considered that the drug relationship does not have a significant effect on the model's performance on this indicator. From the comprehensive comparison of the four indicators, the fusion module has the smallest impact on the model. Compared with the non operational model, the non fusion module only reduces the performance of each indicator by 0.3%, 9.0%, 2.1%, and 0.53%, respectively. On the KIBA dataset, compared with other features, the loss of protein structure feature information has the greatest impact on the accuracy of model prediction. The results showed that the absence of protein structure feature increased the prediction error to 0.165, while the CI value decreased to 0.862. Similar to the Davis dataset, it was also found that the fusion module had the smallest impact on the model, with decreases of 1.4%, 3.6%, 2.1%, and 4.1%, respectively.

Based on the experimental results, we found that fusing multiple feature information can improve the inference performance of the model to a certain extent. Among them, protein structure feature information has the greatest impact on the final prediction performance, so it is of the highest importance; The feature fusion method has the least impact on the performance of the model, so it is possible to consider improving the structural design of this module in the future to achieve better enhancement effects.

Conclusion

Starting from the internal structure and external relationship characteristics of drugs/targets, this article studies feature extraction and fusion techniques to better assist in predicting drug target affinity. The main work is as follows:

Propose a drug target affinity prediction model based on deep learning methods. Firstly, the structure and relationship features of drugs/targets are introduced, and pre trained models and custom feature extraction methods are used to initialize the structure and relationship features. A relationship graph is constructed based on similarity and relationship features are mined. Finally, feature fusion technology is used to fuse the drug structure features, drug relationship features, target structure features, and target relationship features, And send it to the prediction module to complete the affinity value prediction. This model uses graph neural networks and feature extraction and fusion methods to embed drugs and targets, and based on this, achieves more accurate prediction of drug target affinity values.

References

Bagherian, M.; Sabeti, E.; Wang, K.; Sartor, M. A.; Nikolovska-Coleska, Z.; and Najarian, K. 2021. Machine learning approaches and databases for prediction of drug–target interaction: a survey paper. *Briefings in bioinformatics*, 22(1): 247–269.

Bajorath, J. 2002. Integration of virtual and high-throughput screening. *Nature Reviews Drug Discovery*, 1(11): 882–894.

Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A. P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L. J.; Cibrián-Uhalte, E.; et al. 2017. The ChEMBL database in 2017. *Nucleic acids research*, 45(D1): D945–D954.

Gilson, M. K.; Liu, T.; Baitaluk, M.; Nicola, G.; Hwang, L.; and Chong, J. 2016. BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic acids research*, 44(D1): D1045–D1053.

Table 3: Comparison of performance of various models on different datasets (note: the best result is in bold, the second best result is underlined)

Model	Davis				KIBA			
	CI	*MSE*	r_m^2	*AUPR*	*CI*	*MSE*	r_m^2	*AUPR*
KronRLS	0.871	0.379	0.407	0.661	0.782	0.411	0.342	0.635
SimBoost	0.872	0.282	0.644	0.709	0.836	0.222	0.629	0.760
DeepDTA	0.878	0.261	0.630	<u>0.714</u>	0.863	0.194	0.673	0.788
GraphDTA	0.893	0.229	—	—	0.891	0.139	—	—
FusionDTA	**0.913**	0.208	<u>0.743</u>	—	**0.906**	<u>0.130</u>	0.793	—
MGraphDTA	0.900	<u>0.207</u>	0.710	—	<u>0.902</u>	**0.128**	<u>0.801</u>	—
ours	<u>0.912</u>	**0.166**	**0.788**	**0.763**	0.894	0.138	**0.808**	**0.802**

He, T.; Heidemeyer, M.; Ban, F.; Cherkasov, A.; and Ester, M. 2017. SimBoost: a read-across approach for predicting drug–target binding affinities using gradient boosting machines. *Journal of cheminformatics*, 9(1): 1–14.

Jiang, M.; Li, Z.; Zhang, S.; Wang, S.; Wang, X.; Yuan, Q.; and Wei, Z. 2020. Drug–target affinity prediction using graph neural network and contact maps. *RSC advances*, 10(35): 20701–20712.

Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; et al. 2019. PubChem 2019 update: improved access to chemical data. *Nucleic acids research*, 47(D1): D1102–D1109.

Li, M.; Lu, Z.; Wu, Y.; and Li, Y. 2022. BACPI: a bi-directional attention neural network for compound–protein interaction and binding affinity prediction. *Bioinformatics*, 38(7): 1995–2002.

Nguyen, T.; Le, H.; Quinn, T. P.; Nguyen, T.; Le, T. D.; and Venkatesh, S. 2021. GraphDTA: Predicting drug–target binding affinity with graph neural networks. *Bioinformatics*, 37(8): 1140–1147.

Organization, W. H.; et al. 2022. World health statistics 2022.

Öztürk, H.; Özgür, A.; and Ozkirimli, E. 2018. DeepDTA: deep drug–target binding affinity prediction. *Bioinformatics*, 34(17): i821–i829.

Pahikkala, T.; Airola, A.; Pietilä, S.; Shakyawar, S.; Szwajda, A.; Tang, J.; and Aittokallio, T. 2015. Toward more realistic drug–target interaction predictions. *Briefings in bioinformatics*, 16(2): 325–337.

Pratim Roy, P.; Paul, S.; Mitra, I.; and Roy, K. 2009. On two novel parameters for validation of predictive QSAR models. *Molecules*, 14(5): 1660–1701.

Ru, X.; Ye, X.; Sakurai, T.; and Zou, Q. 2022. NerLTR-DTA: drug–target binding affinity prediction based on neighbor relationship and learning to rank. *Bioinformatics*, 38(7): 1964–1971.

Rudrapal, M.; Khairnar, S. J.; and Jadhav, A. G. 2020. Drug repurposing (DR): an emerging approach in drug discovery. *Drug repurposing-hypothesis, molecular aspects and therapeutic applications*, 10.

Shao, K.; Zhang, Y.; Wen, Y.; Zhang, Z.; He, S.; and Bo, X. 2022. DTI-HETA: prediction of drug–target interactions based on GCN and GAT on heterogeneous graph. *Briefings in Bioinformatics*, 23(3): bbac109.

Steck, H.; Krishnapuram, B.; Dehing-Oberije, C.; Lambin, P.; and Raykar, V. C. 2007. On ranking in survival analysis: Bounds on the concordance index. *Advances in neural information processing systems*, 20.

Thafar, M.; Raies, A. B.; Albaradei, S.; Essack, M.; and Bajic, V. B. 2019. Comparison study of computational prediction tools for drug-target binding affinities. *Frontiers in chemistry*, 7: 782.

Thafar, M. A.; Alshahrani, M.; Albaradei, S.; Gojobori, T.; Essack, M.; and Gao, X. 2022. Affinity2Vec: drug-target binding affinity prediction through representation learning, graph mining, and machine learning. *Scientific reports*, 12(1): 4751.

Yang, Z.; Zhong, W.; Zhao, L.; and Chen, C. Y.-C. 2022. MGraphDTA: deep multiscale graph neural network for explainable drug–target binding affinity prediction. *Chemical science*, 13(3): 816–833.

Yuan, W.; Chen, G.; and Chen, C. Y.-C. 2022. FusionDTA: attention-based feature polymerizer and knowledge distillation for drug-target binding affinity prediction. *Briefings in Bioinformatics*, 23(1): bbab506.

Zhang, X.; Li, L.; Ng, M. K.; and Zhang, S. 2017. Drug–target interaction prediction by integrating multiview network data. *Computational biology and chemistry*, 69: 185–193.