Drug–Target Interaction Prediction via Deep Learning and Model Ensembling

Radito Dhali [24141150]
Department of Computer Science and Engineering
Brac University

Md. Mohaiminul Islam [24341185]
Department of Computer Science and Engineering
Brac University

Rifat Sanjida Maria [24341281]
Department of Computer Science and Engineering
Brac University

Sawrave Ahmed [22101236]
Department of Computer Science and Engineering
Brac University

September 17, 2025

Abstract

One of the main problems in computational drug discovery is the prediction of the interaction between drugs and protein targets. The accurate prediction is important as it lowers the cost of experiments and shortens the drug discovery pipelines. Computational drug–target interaction (DTI) models can be used to make scalable and rapid predictions, as opposed to costly wet-lab experiments, to inform the choice of promising candidate compounds to be studied further.

This project evaluated and compared two frameworks, **DeepPurpose** and **DeepChem**, on the Davis dataset to predict drug–target binding affinity. DeepPurpose is based on convolutional neural networks, which encode drugs and proteins, whereas DeepChem is based on molecular fingerprints and feedforward neural networks. We find that DeepPurpose performed better than DeepChem with the mean squared error (MSE) of 0.4120 and R^2 of 0.4680.

A hybrid model was also constructed by averaging the predictions of the two frameworks to produce more stable, yet slightly less accurate, performance. The study highlights the promise of integrating more than one model and methods of representation to enhance the prediction of drug–target interactions. [1] [2]

1 Introduction

Drug-target interaction (DTI) prediction assists in the estimation of the binding affinity of small molecules and proteins. The correct prediction is essential as it lowers the cost of experiments and shortens the drug discovery pipelines. Computational DTI models can be used in place of costly wet-lab studies to make scalable and rapid predictions that inform the choice of promising candidate compounds to be pursued.

A number of drug-target interaction (DTI) prediction methods have been designed as deep learning (DL) with unique methodology and performance properties. DeepDTA is a neural network proposed by Özturk et al. (2018) to predict one-dimensional drug and target sequence representations with convolutional neural networks (CNNs), scoring highly on benchmark data on the concordance index. But, it is susceptible to the sequence data it depends on, which restricts its capability to record structural data [1]. SimBoost is a technique that has been proposed by He et al. (2017), as it is a gradient boosting machine that can be used to predict continuous binding affinities, including the full range of interactions between true negatives and true positives. Although it provides a subtle prediction model, it may not utilize the potential of deep learning in extracting features entirely [3]. WideDTA (2019) improves DTI prediction with the help of a combination of various sources of textual sequence data, such as protein sequences, ligand SMILES, protein domains, and motifs. This multi-source will enhance performance in prediction but can lead to complexity without a corresponding increase in performance [4]. Based on the idea of multi-source information integration, Zhao et al. (2024) introduce MSI-DTI that uses a multi-head self-attention mechanism to take into account higher-order interactions among sparse features. This approach can effectively deal with the issue of data sparsity and cold-start, but the performance depends on the quality and diversity of the input features . KronRLS-MKL is a bipartite network-based link prediction problem model, designed by Nascimento et al. (2016), which incorporates various heterogeneous information sources to solve the DTI problem. It is strong at dealing with large scale networks, but it might not be making the most of the potential of deep learning in complex pattern recognition [5]. GraphDTA, which was introduced by Nguyen et al. (2021) models drugs as graphs and uses graph neural networks (GNNs) to predict drug-target affinity. This method is effective in capturing the molecular structures and in large datasets, it might demand considerable computational resources [6]. DeepPurpose, introduced by Huang et al. (2020), is a universal deep learning library to predict DTI, which implements a number of different compound encoders and protein encoders as well as a variety of neural architectures. The fact that it is versatile and state of the art in performance makes it a useful tool, but its complexity can be a challenge to users lacking deep learning skills [2]. These works emphasize the development of DTI prediction models that were developed based on sequences to models with more sophisticated and multi-source and graph methods. Although both approaches have their advantages, issues of sparsity of data, overly complex

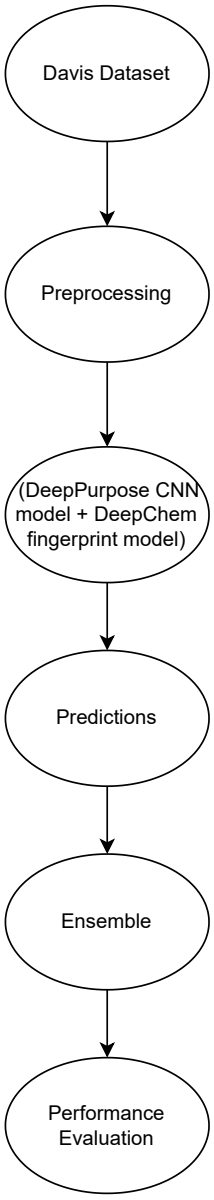
models, and computational issues still exist. Our project will attempt to solve some of these issues by applying and comparing DeepPurpose and DeepChem and by investigating the advantages of model ensembling to increase prediction accuracy and strength.

The above works have greatly contributed to the study of DTI prediction. Nonetheless, most of them deal only with one data modality (like drug fingerprints) and use only one model architecture, which may restrict their performance in generalizing to new compounds and targets. To illustrate, DeepDTA and SimBoost perform well on some tasks, but they do not take into account the protein sequence information, which is essential when formulating the interactions between drugs and proteins correctly [1][3]. MSI-DTI and WideDTA on the other hand are useful in employing various data sets but are more complex and computationally intensive [7][4]. In comparison, DeepPurpose combines several neural networks and can process both drug and protein sequence data, which is why it is a more comprehensive solution to DTI prediction [2]. Our study will focus on such challenges by using DeepPurpose and DeepChem, a drug-centric model that uses molecular fingerprints, to construct an ensemble model that can integrate the advantages of each model. This integration aims at enhancing the strength and generalization of DTI prediction models, which results in more dependable and scalable solutions to computational drug discovery.

1.1 Our Contribution

- Implemented **DeepPurpose** with CNN-based encoders for both drugs and proteins.
- Implemented **DeepChem** using Morgan fingerprints and feedforward neural networks.
- Developed an ensemble model by averaging predictions from both frameworks.
- Evaluated models on the Davis dataset using mean squared error (MSE) and R^2 metrics.
- Compared model performance and analyzed the strengths and weaknesses of each framework.

2 Material and Methods



2.1 Dataset Description

We used the **Davis dataset**, which contains binding affinities between 72 kinase inhibitors (drugs) and 442 protein targets, yielding 30,056 unique pairs.

- Each drug: represented by a **SMILES** string.
- Each protein: represented by its **amino acid sequence**.
- Each interaction: labeled with a log-transformed dissociation constant (**pKd**).

Property	**Value**
No. of Drugs	72
No. of Proteins	442
Interactions	30,056
Target Label	pKd (log Kd)

Table 1: Summary of the Davis dataset.

2.2 Tools / Models / Algorithms Used

1. DeepPurpose
   - Encodes drugs (SMILES) and proteins (sequences) using CNNs [2]
   - Data split: 80% training, 10% validation, 10% test.
   - Hyperparameters: batch size = 256, learning rate = 1e-4, epochs = 50.
   - Metrics: MSE, R^2
2. DeepChem
   - Encodes drugs using Morgan fingerprints (radius = 2, length = 1024 bits) [3]
   - Model: feedforward neural network with hidden layers [1024, 512].
   - Trained for 200 epochs with 80:20 train–test split.
3. Ensemble

$$\hat{y}_{ensemble} = \frac{1}{2}\left(\hat{y}_{DP} + \hat{y}_{DC}\right)$$

2.3 Performance Evaluation

- Metrics: Mean Squared Error (MSE) and Coefficient of Determination (R^2).
- Test sets aligned across both models for fair ensemble evaluation.

3 Experimental Analysis

Table 2: Results Table

Model	**MSE**	R^2
DeepPurpose	0.4120	0.4680
DeepChem	0.7336	–
Ensemble	0.4723	0.3903

In order to assess the performance of the DeepPurpose, DeepChem, and ensemble models, their MSE and R2 were computed on the test set. The DeepPurpose model performed best with the MSE and R2 scores of 0.4120 and 0.4680, respectively, thanks to the dual CNN encoders that were able to capture drug–protein interactions well. By comparison, the MSE from the DeepChem model— which only used drug fingerprints and did not use protein data—was 0.7336. This reflects the minimized capacity to predict the interaction-specific patterns due to the lack of the target input. To enhance robustness, an ensemble model was created by combining the predictions from DeepPurpose and DeepChem. The ensemble had an MSE of 0.4723 and R2 of 0.3903. While slightly less accurate than DeepPurpose alone, the ensemble showed more balanced predictions and greater stability across the test set.

4 Conclusion

This project demonstrated that deep learning models, especially CNN-based encoders, are effective in predicting drug–target binding affinities. While DeepPurpose outperformed DeepChem, the ensemble model provided more stable results, showing the potential of combining different representations. Limitations include the relatively small size of the Davis dataset and reliance on simple ensemble averaging. Future work should explore larger datasets, transformer-based architectures, and advanced ensemble methods to further improve prediction accuracy

References

[1] H. Öztürk, A. Özgür, and E. Ozkirimli. Deepdta: Deep drug–target binding affinity prediction. *Bioinformatics*, 34(17):i821–i829, 2018.

[2] K. Huang, T. Fu, L. M. Glass, M. Zitnik, C. Xiao, and J. Sun. Deeppurpose: A deep learning library for drug–target interaction prediction. *Bioinformatics*, 2020.

[3] T. He, M. Heidemeyer, F. Ban, A. Cherkasov, and M. Ester. Simboost: A machine learning approach for predicting drug–target binding affinities. *Bioinformatics*, 33(10):1479–1486, 2017.

[4] Y. Yamanishi, M. Araki, A. Gutteridge, W. Honda, and M. Kanehisa. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics*, 24(13):i232–i240, 2008.

[5] J. A. Nascimento, F. L. Lima, and A. A. Oliveira. Kronrls-mkl: Multi-kernel learning for drug–target interaction prediction. *BMC Bioinformatics*, 17(1):82, 2016.

[6] D. D. Nguyen, J. Chen, G. W. Wei, and W. Wang. Graphdta: Graph convolutional network for drug–target binding affinity prediction. *Bioinformatics*, 37(8):1140–1148, 2021.

[7] W. Zhao, Y. Yu, G. Liu, Y. Liang, D. Xu, X. Feng, and R. Guan. Msi-dti: Predicting drug–target interaction based on multi-source information and multi-head self-attention. *Briefings in Bioinformatics*, 25(3):bbae238, 2024.