nature portfolio

Peer Review File

DeepDTAGen: A Multitask Deep Learning Framework for Drug-Target Affinity Prediction and Target-aware Drugs Generation

Corresponding Author: Professor Min Li

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Journal: Nature Communication
Manuscript#: NCOMMS-24-54172, Corresponding Author: Min Li
Title: DeepDTAGen: Multitask deep learning framework for Predicting Drug-Target Affinity and Generating Target-Specific Drugs

The authors have presented a promising computational multitask deep-learning framework called DeepDTAGen for predicting drug-target binding affinity as well as for generating new target-aware drug variants. Moreover, they introduced a novel algorithm, FetterGrad, which addresses the optimization challenges associated with multitask learning. The authors evaluated their method using three datasets in terms of several evaluation metrics (The Mean Squared Error, Concordance Index, R squared, and Area Under the Curve 197) to demonstrate their method's efficiency and effectiveness in predicting drug-target affinity. They also evaluated the novel drugs by evaluating the Chemical Properties of the Generated Drugs. They further assessed the novel drugs by analyzing their chemical properties. All evaluation processes highlighted the method’s effectiveness in both tasks, thereby contributing to advancements in drug discovery.
This work addresses a significant problem and proposes a sound methodology. The introduction is well-written, providing strong motivation and a clear description of the research. It also effectively covers relevant literature and related works. The experiments are well-conducted, but certain sections of the manuscript require further detail and improvements. As a result, revisions are necessary before the manuscript can be considered for publication in Nature Communications.

Comments 1.
The prediction of drug-target interactions (DTIs) can significantly accelerate drug development by facilitating computational drug repositioning. The authors should explain drug repositioning in the introduction and motivation.

Comments 2.
The authors have discussed several works on predicting drug-target interactions (DTIs), focusing on binary classification methods. They have also mentioned regression-based approaches such as DeepDTA, GraphDTA, and SimBoost. However, more recent works predicting drug-target binding affinity, such as "Affinity2Vec: drug-target binding affinity prediction through representation learning, graph mining, and machine learning." Scientific reports 12.1 (2022): 4751.", which utilizes the KIBA, DAVIS, and PDBind datasets, should also be referenced in the related work section or included in comparisons with state-of-the-art methods.

Comment 3.
In section, 1.2.1 Protein representation, the authors mentioned “Furthermore, we have extracted a 128*1000 dimensional matrix against each numerical sequence using the word embedding method,”.
It is crucial to specify which 'word embedding' methods were used, as various approaches capture different types of features and patterns. Clarifying these details is essential for both understanding the methodology and ensuring reproducibility. Additionally, the rationale behind selecting a 128-dimensional embedding needs to be explained. Did the authors perform any empirical analysis or hyperparameter tuning to determine this specific dimension?

Comment 4.
In section, 1.4 FetterGard model, the authors mentioned “To tackle the gradient conflicting issue in DTI prediction and Drug generation tasks. we developed …”, It is better to explain the gradient conflict issue in multi-task learning briefly and then explain how the authors propose their solution.

Comment 5.
In the results section, The performance comparison would be stronger if the variance were reported.

Comment 6.
In the results section, Y-randomization can be used as a validation tool to compare the performance of a model with pseudo-random models trained on permuted datasets, thereby demonstrating the significance of the results. The authors can consider employing this test or any other statistical test to validate the significance of their results.. (see Rucker, C.; Rucker, G.; Meringer, M. y-Randomization and its variants in QSPR/QSAR. Journal of chemical information and modeling 2007, 47, 2345–2357)

Comment 7.
I believe it is crucial to consider using the AUPR (Area Under the Precision-Recall Curve) as an evaluation metric for drug-target interactions, as it provides a more realistic assessment compared to AUC, particularly in the context of highly imbalanced data. This is especially important if the problem is converted from a regression task to a binary classification problem in the evaluation. The authors are encouraged to include AUPR in their evaluation to ensure a more accurate reflection of model performance in imbalanced settings.

Comment 8.
To better demonstrate the proposed methods' robustness to results, the authors should consider using different splitting methods for the prediction of drug-target binding affinity, such as a new drug or target setting. These methods better mimic real-world scenarios with few known interactions for new drugs or targets.

Comment 9.
To evaluate the generated drugs, although the authors provided an explanation of the Validity, Novelty, and Uniqueness measurements in the supplementary materials, I believe it would improve the flow of the text and enhance the overall comprehension if these explanations were incorporated into the main manuscript.

Comment 10.
The conclusion would be more robust and impactful if the authors included a discussion of potential future research directions and areas for further exploration.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
The article entitled “DeepDTAGen: Multitask deep learning framework for Predicting Drug-Target Affinity and Generating Target-Specific Drugs” was submitted. This paper introduces a novel multitask learning framework for drug discovery that predicts drug-target binding affinities and generates new drugs, using a unified approach with the FetterGrad algorithm to handle optimization challenges.

1. I could not understand what is innovative about this work. Drug-Target Affinity prediction methods are abundant, and new ones are reported daily. The same goes for molecular generation methods. Specifically, what is new about your approach? Without clarifying this, I believe it does not merit publication.
2. While the paper mentions multitask learning, it is unclear what exactly is being learned. It appears that Drug-Target Affinity prediction and molecular generation are simply executed sequentially.
3. In Figure 3, I noticed that the distribution of logP seems to vary significantly between panels (A), (B), and (C). However, upon closer inspection, I realized that the scales on the x-axes are different. Could you align the scales to make the graphs consistent?
4. Regarding molecular generation, I understand that metrics like QED, logP, and SAS are not used as input criteria for generating molecules. As a result, it seems that the generated molecules may be biased toward the input affinity, potentially neglecting considerations such as QED, logP, and SAS. What are your thoughts on this?
5. In Table 3, I find it questionable that the top five generated molecules each come from different input compounds. Could you provide details on how the input compounds were selected, how many molecules were generated, how many were synthesized, and what the distribution of activity values was after assay? I am weary of molecular generation papers that only report favorable results.
6. The molecules in Table 3 show biological activity for three different targets. However, which target were the molecules designed to address during generation? A critical issue in drug discovery is polypharmacology, meaning that "many drugs do not act on only a single target biomolecule but on multiple targets, affecting efficacy and toxicity phenotypes." Would it not be preferable to generate molecules with higher selectivity?
7. Is it possible to display the structures of the input drugs in Tables 3 and 4? I would like to compare the structural similarity

between the input and generated molecules.
8. How were the molecules in Table 4 selected?

(Remarks on code availability)
I apologize for the delay in providing this report. I attempted to set up the environment on my usual machine, but after three days, the process did not complete successfully. I then switched to a different machine, where I managed to set up the environment without issues.

The code explanations are very detailed, and I found the instructions for environment setup and DEMO execution easy to follow, with no ambiguity in understanding the commands.

With the environment successfully set up, I proceeded to run the DEMO scripts. However, I encountered errors that prevented successful execution of the scripts. Below are the details:

- **DEMO_(Affinity).py**
When executing `python DEMO_(Affinity).py`, I encountered the following error:
```
Traceback (most recent call last):
File "DEMO_(Affinity).py", line 7, in <module>
from required_files_for_demo.demo_utils import *
File "/home/???/DeepDTAGen/DEMO/required_files_for_demo/demo_utils.py", line 8, in <module>
from utils import *
ModuleNotFoundError: No module named 'utils'
```
I attempted to resolve this by installing the package using `pip install utils`, but a subsequent error occurred:
```
python DEMO_(Affinity).py
Traceback (most recent call last):
File "DEMO_(Affinity).py", line 9, in <module>
from required_files_for_demo.model_aff import DeepDTAGen
File "/home/???/DeepDTAGen/DEMO/required_files_for_demo/model_aff.py", line 10, in <module>
from utils import Tokenizer
ImportError: cannot import name 'Tokenizer' from 'utils' (/home/???/miniforge3/envs/DeepDTAGen/lib/python3.8/site-packages/utils/__init__.py)
```

- **DEMO_(Generation).py**
Similarly, when running `python DEMO_(Generation).py`, the following error was encountered:
```
Traceback (most recent call last):
File "DEMO_(Generation).py", line 9, in <module>
from required_files_for_demo.model_gen import DeepDTAGen
File "/home/???/DeepDTAGen/DEMO/required_files_for_demo/model_gen.py", line 10, in <module>
from utils import Tokenizer
ImportError: cannot import name 'Tokenizer' from 'utils' (/home/???/miniforge3/envs/DeepDTAGen/lib/python3.8/site-packages/utils/__init__.py)
```

As mentioned above, I attempted to execute both scripts, but errors were encountered, leading to failure in execution. Unfortunately, due to time constraints, I am unable to further investigate the root cause of these issues. Thus, I am reporting this as the result of my code review.

Reviewer #3

(Remarks to the Author)
1) General comments
This paper presents a novel multitask learning framework for predicting drug-target binding affinity (DTA) and generating novel drugs which is entitled DeepDTAGen. The approach leverages Graph Convolutional Neural Networks (GCNs) and Gated Convolutional Neural Network (Gated-CNN) to extract features from drugs and proteins, which are used for both tasks. Particularly, the authors developed algorithm FetterGard to tackle the optimization issues of two different tasks. The experimental results show that DeepDTAGen performs better than other methods both in prediction and generation tasks. Overall, the work is interesting and useful.

Minors:
1. SMILES refers to the Simplified Molecular Input Line Entry Specification, which is a notation that uses ASCII strings to explicitly describe molecular structures. In the text, "SMILE" has been used multiple times to refer to the singular form of "SMILES," which should be corrected.

2. What are the loss functions for the prediction and generation tasks? How are the losses of the two tasks integrated?

3. In the FetterGard model section 1.4, what method is used to determine which gradient is the conflicting one among different gradients? For example, how is it determined whether gradient A should be projected onto B or B onto A?

4. In Table 3, the fifth column, why does each molecule correspond to so many targets? Is this task being performed in a multi-target generation scenario?

5. There is no Discussion section, the authors should add this section and discuss the implications of your findings, potential limitations, and suggestions for future research directions.

6. What database does the Input Drug ID in Table 4 refer to? Why were these Drug IDs chosen as inputs?

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
After carefully reviewing the revised manuscript, I am fully satisfied with how the authors have addressed all of my comments. The revisions have significantly enhanced the clarity and rigor of the work. I appreciate the authors' excellent efforts and, based on these improvements, I recommend this research paper for publication in Nature Communications.

(Remarks on code availability)
The GitHub repository is well-structured and organized, featuring comprehensive comments and thorough documentation. Navigation is straightforward thanks to the logical arrangement of files and folders. The authors have provided detailed environment settings and included a demo to facilitate reproducing the paper's results. The README file is clear and provides step-by-step instructions for installing and running the application. Overall, the code is a highly usable resource for the community and significantly enhances the reproducibility of the research findings.

Reviewer #2

(Remarks to the Author)
Thank you for addressing my comments. Apart from your response to the first question below, I am satisfied with your answers and have no further questions.
1. I could not understand what is innovative about this work. Drug-Target Affinity prediction methods are abundant, and new ones are reported daily. The same goes for molecular generation methods. Specifically, what is new about your approach? Without clarifying this, I believe it does not merit publication.
I sincerely appreciate the authors' thoughtful and detailed responses. While it may be due to my own lack of understanding, I still do not believe that this work presents results innovative enough to warrant publication in Nature Communications. It does not demonstrate a significant improvement in accuracy, nor does it introduce a novel method that leads to new findings. I believe it would be more suitable for publication in a specialized journal.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
Thank you for author's response. I don't have further questions.

(Remarks on code availability)