

# DeepPurpose: a Deep Learning Library for Drug-Target Interaction Prediction and Applications to Repurposing and Screening

Kexin Huang<sup>1</sup> Tianfan Fu<sup>2</sup> Lucas M. Glass<sup>3</sup> Marinka Zitnik<sup>1</sup> Cao Xiao<sup>3</sup> Jimeng Sun<sup>4</sup>

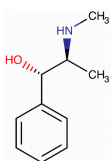
<sup>1</sup>Harvard <sup>2</sup>Georgia Tech <sup>3</sup>IQVIA <sup>4</sup>UIUC CS



## Drug-Target Interaction



## Input Data Format

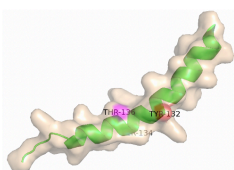


*Pseudoephedrine*

### SMILES

Simplified molecular-input line-entry system

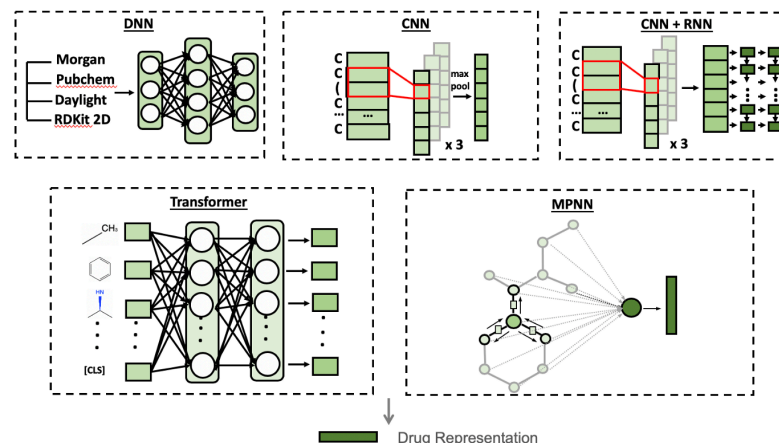
CC(C(C1=CC=CC=C1)O)NC



*Alpha-2A receptor*  
Amino Acid Sequence

MFHQEPLAEGSFAPMGLSPDAGNASWNGTEAPGGGARATPYSLVQLTLVCLAGLLMLTDFGNVLIIVFTSRALKAPONLFLVSLASDILVATLVIPFLANEVNGWVYFGKAWCEIYALDLVLFCTSSVHLCAISLDRYWSTQAIENLKRTPRRKAIITVWVSAVIFPLISIEKGGGGGQPPAEPKCEINDQKWYIISGIGSFFAPCLMILYVRYGLAKRRTVPPSRGPDVAVAAPGCTERRPGLGPERSGAGPGAEEPLPTQLNGAPGEPAPAPGRDLDLESSSDHAERPPGPRRPERGPRGKARASQVKGPDSPRRGPGATGIGTPAAGPGEERVGAAKASRWGRQNRKRFVLAIVIGVPCVWFFFFFTYLTAVGCSVPTLTKFFWFGYCNSSLNPVYITFNHDFRAFKKLCRGDRKRV

## Drug Encoders

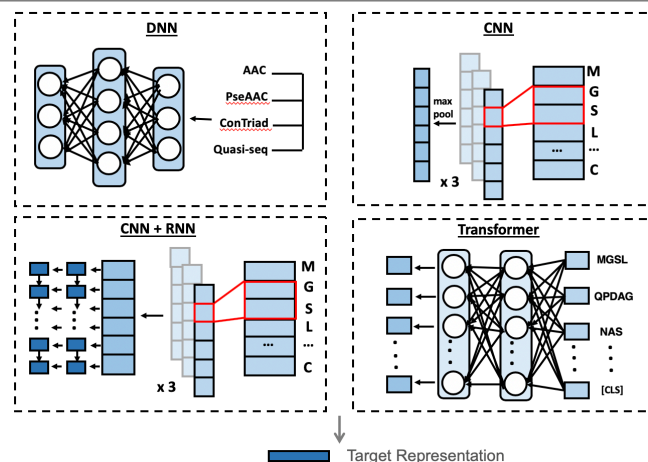


## DeepPurpose One-Line Mode

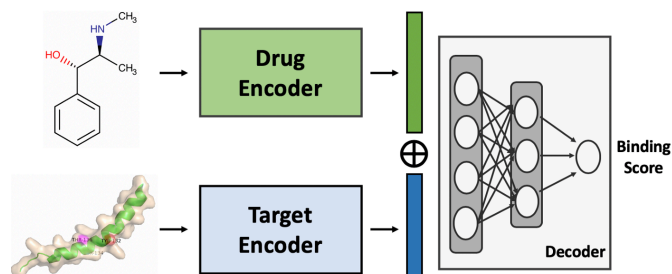
```
>>> from DeepPurpose import oneliner
>>> from DeepPurpose.dataset import *
>>>
>>> oneliner.repurpose(*read_file_target_sequence('target.txt'), \
                      *read_file_repurposing_library('repurpose.txt'))
```

Rank	Drug Name	Target Name	Binding Score
1	* Sofosbuvir	SARS-CoV2 3CL Protease	190.25
2	Daclatasvir	SARS-CoV2 3CL Protease	214.58
3	Vicriviroc	SARS-CoV2 3CL Protease	315.70
4	* Simeprevir	SARS-CoV2 3CL Protease	396.53
5	Etravirine	SARS-CoV2 3CL Protease	409.34
6	* Amantadine	SARS-CoV2 3CL Protease	419.76
7	Letermovir	SARS-CoV2 3CL Protease	460.28
8	Rilpivirine	SARS-CoV2 3CL Protease	470.79
9	+ Darunavir	SARS-CoV2 3CL Protease	472.24
10	+ Lopinavir	SARS-CoV2 3CL Protease	473.01
11	Maraviroc	SARS-CoV2 3CL Protease	474.86
12	Fosamprenavir	SARS-CoV2 3CL Protease	487.45
13	+ Ritonavir	SARS-CoV2 3CL Protease	492.19

## Target Encoders



## Encoder-Decoder Framework



## DeepPurpose DTI Framework

```
>>> from DeepPurpose import models
>>> from DeepPurpose.utils import *
>>> from DeepPurpose.dataset import *
>>>
>>> X_drug, X_target, y = load_process_DAVIS(SAVE_PATH, binary=False)
>>>
>>> drug_encoding, target_encoding = 'CNN', 'CNN'
>>> train, val, test = data_process(X_drug, X_target, y, drug_encoding, \
                                target_encoding, split_method='random', \
                                frac=[0.7, 0.1, 0.2], random_seed = 1)
>>>
>>> config = generate_config(drug_encoding, target_encoding, \
                            cls_hidden_dims = [1024, 1024, 512], \
                            train_epoch = 100, LR = 0.001, batch_size = 256, \
                            cnn_drug_filters = [32, 64, 96], \
                            cnn_drug_kernels = [4, 8, 12], \
                            cnn_target_filters = [32, 64, 96], \
                            cnn_target_kernels = [4, 8, 12])
>>>
>>> model = models.model_initialize(**config)
>>> model.train(train, val, test)
```