Supporting Information

Transformer-based deep learning method for optimizing ADMET properties of lead compounds

Lijuan Yang[1,2,3,4], Chao Jin[1], Guanghui Yang[1,4], Zhitong Bing[1,4], Liang Huang[2], Yuzhen Niu[5*] and Lei Yang[1,4*]

1. Institute of modern physics, Chinese Academy of Science, Lanzhou 730000, China
2. School of Physics and Technology, Lanzhou University, Lanzhou 730000, China
3. School of Physics, University of Chinese Academy of Science, Beijing 100049, China
4. Advanced Energy Science and Technology Guangdong Laboratory, Huizhou 516000, China
5. Shandong Laboratory of Yantai Advanced Materials and Green, Yantai 264006, China

* Corresponding to: Yuzhen Niu, Email address: niuyzh12@lzu.edu.cn
* Corresponding to: Lei Yang, Email address: lyang_imp@outlook.com

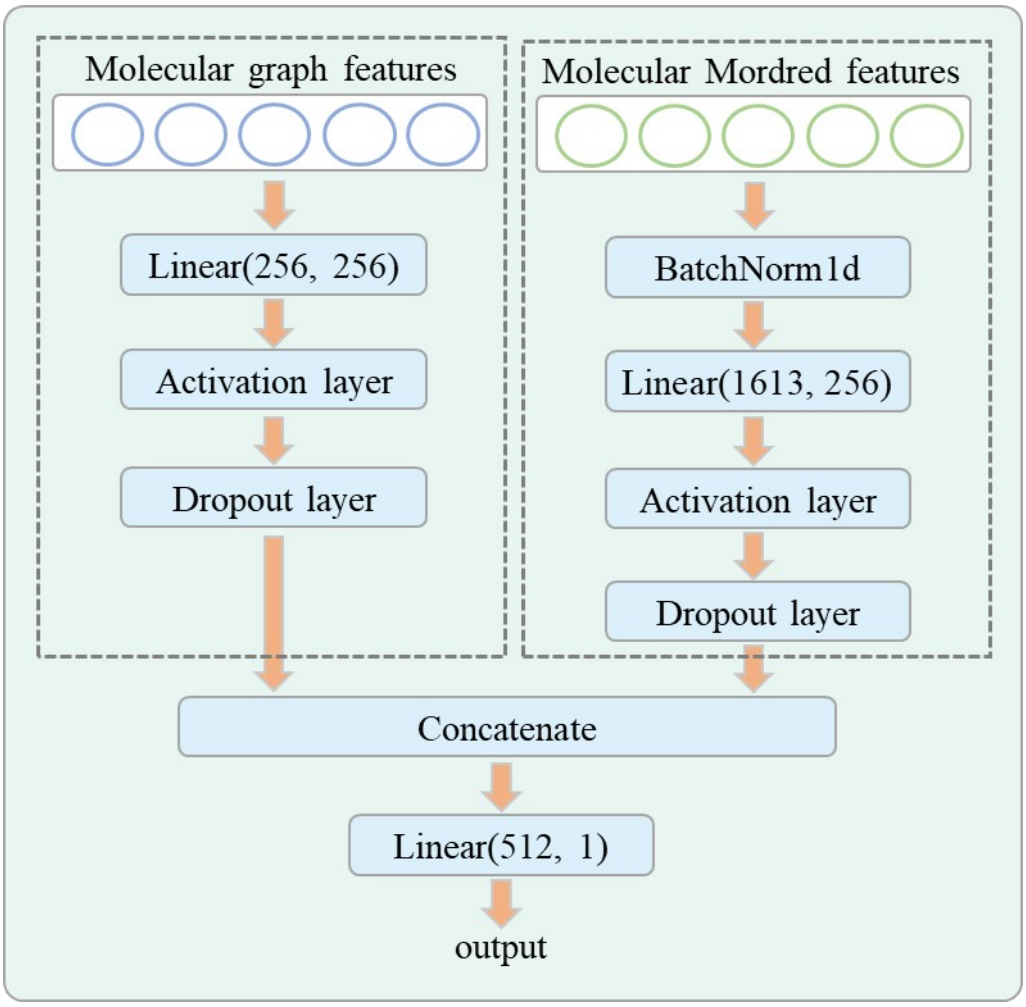

Figure S1. The detailed flow of the prediction layer.

Table S1. The hyperparameter settings for predictive and optimization models.

	ADMET prediction model	Constraints-Transformer
Parameters	Optimizer=Adam Loss Function=BCE/ SmoothL1 Lr=1e-5 Num_encoder_layer=3 Num_heads=8 D_model=256 D_ff = 1024 Dense_dropoutrate=0.5 Dense_activation =LeakyRELU(0.1) Epoch=200 Batch_size=24/32	Optimizer=Adam Loss Function=KLDivLoss Num_encoder_layer=6 Num_decoder_layer=6 Num_heads=8 D_model=256 D_ff = 2048 Epoch=100 Batch_size=256 Num_samples=10

Table S2. The label of each property.

Property	**Measured unit**	**Threshold**	**Threshold in *ln* scale**	**Discrete property labels**	**Changed property labels**
lnS	mol/L	Low: ≤ 0.05 High: > 0.05	Low: ≤ -3.0 High: > -3.0	lnS _ (inf, -7) lnS _-7 lnS _-6 ... lnS _2 lnS _(2, inf)	lnS _low-high lnS _high-low lnS _no-change
logD	--	--	--	logD_(inf, -5) logD_-5 logD_-4 ... logD_5 logD_(5, inf)	logD_(-inf, -5] logD_(-5, -4] logD_(-4, -3] ... logD_(4, 5] logD_(5, inf)
CL	μL/min/mg	Low: ≤ 20 High: > 20	Low: ≤ 3.0 High: > 3.0	CL_(inf, -5) CL_-5 CL_-4 ... CL_5 CL_(5, inf)	CL_low-high CL_high-low Cl_no-change

Table S3. Evaluation result on TDC ADMET benchmark group.

Dataset	Matrix	Current Top1	Bert & Mordred	Only Bert	Only Mordred	Rank
caco2	$MAE^{\downarrow}$	0.288±0.011	**0.277±0.002**	0.331±0.004	0.321±0.010	1st
lipophilicity		0.533±0.005	**0.524±0.005**	0.562±0.022	0.591±0.056	1st
solubility_aqsoldb		0.727±0.004	**0.727±0.004**	0.771±0.009	0.780±0.024	1st
ppbr		8.251±0.115	7.879±0.102	**7.722±0.165**	10.889±0.269	1st
ld50		0.588±0.005	**0.603±0.003**	0.607±0.165	0.656±0.009	3rd
hia	$AUROC^{\uparrow}$	0.988±0.002	**0.979±0.008**	0.936±0.021	0.933±0.020	3rd
pgp		0.946±0.001	**0.936±0.000**	0.889±0.010	0.931±0.004	2nd
bioav		0.748±0.033	0.685±0.010	0.668±0.037	**0.721±0.005**	3rd
bbb		0.905±0.001	**0.911±0.002**	0.868±0.007	0.899±0.004	1st
cyp3a4 substrate		0.680±0.005	**0.681±0.000**	0.652±0.008	0.669±0.012	1st
herg		0.874±0.014	**0.835±0.014**	0.789±0.017	0.801±0.013	3rd
Ames		0.865±0.002	**0.855±0.001**	0.804±0.006	0.844±0.002	3rd
dili		0.933±0.011	**0.933±0.005**	0.903±0.003	0.908±0.005	1st
cyp2c9 inhibition	$AUPRC^{\uparrow}$	0.794±0.004	**0.745±0.002**	0.709±0.006	0.737±0.005	5th
cyp2d6 inhibition		0.721±0.001	**0.624±0.006**	0.655±0.009	0.611±0.013	3rd
cyp3a4 inhibition		0.882±0.001	0.849±0.003	0.803±0.011	**0.852±0.003**	3rd
cyp2c9 substrate		0.433±0.017	**0.444±0.014**	0.396±0.015	0.416±0.022	1st
cyp2d6 substrate		0.677±0.047	**0.725±0.012**	0.579±0.006	0.678±0.030	1st
Vdss	$\rho^{\uparrow}$	0.612±0.018	**0.632±0.015**	0.585±0.014	0.583±0.026	1st
Half life		0.416±0.009	**0.485±0.020**	0.396±0.021	**0.492±0.029**	1st
Cl-hepa		0.491±0.006	**0.422±0.011**	0.364±0.031	0.405±0.008	3rd
Cl-micro		0.625±0.002	**0.583±0.005**	0.549±0.015	0.578±0.013	6th