SUPPLEMENTARY DATA

ADMETLAB 3.0: AN UPDATED COMPREHENSIVE ONLINE ADMET PREDICTION PLATFORM ENHANCED WITH BROADER COVERAGE, IMPROVED PERFORMANCE, API FUNCTIONALITY, AND DECISION SUPPORT

AUTHORS

Li Fu[1,†], Shaohua Shi[2,†], Jiacai Yi[3,†], Ningning Wang[4], Yuanhang He[1], Zhenxing Wu[5], Jinfu Peng[1], Youchao Deng[1], Wenxuan Wang[1], Chenkun Wu[3], Aiping Lyu[2], Xiangxiang Zeng[6], Wentao Zhao[3], Tingjun Hou[5,*] and Dongsheng Cao[1,*]

[1] Xiangya School of Pharmaceutical Sciences, Central South University, Changsha, Hunan, 410013, P.R. China
[2] School of Chinese Medicine, Hong Kong Baptist University, Kowloon, Hong Kong SAR, 999077, P.R. China
[3] School of Computer Science, National University of Defense Technology, Changsha, Hunan, 410073, P.R. China
[4] Xiangya Hospital of Central South University, Changsha, Hunan, 410008, P.R. China
[5] College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang, 310058, P.R. China
[6] Department of Computer Science, Hunan University, Changsha, Hunan, 410082, P.R. China

* To whom correspondence should be addressed. Tel: +86-731-89824761; Email: oriental-cds@163.com
Correspondence may also be addressed to Tingjun Hou. Tel: +86-571-88208412; Email: tingjunhou@zju.edu.cn

† These authors contribute to this paper equally.

Table of Contents

DMPNN framework

Yang et al (1). introduced an open-source Python package called Chemprop designed for implementing DMPNN models. This package offers a robust and efficient solution tailored for molecular property prediction tasks and has garnered extensive utilization in fields such as drug discovery and materials science (2). Within the DMPNN framework, there exist two distinct stages: message passing and readout. Here, a graph G serves as an illustrative example, representing node (atom) features as x_v and edge (bond) features as $\mathrm{e_{vw}}$.Initially, the edge hidden state h_{vw}^0 is initialized using eq 1. This equation concatenate atom and bond features by passing them through the learned matrix W_I and applying the rectified linear unit (RELU) activation function. This initialization defines the edge hidden state, which undergoes subsequent updates during the message passing process. In eq 1, τ represent the RELU activation function, W_I denotes a learned matrix, and $cat(\,)$ signifies a basic concatenation operation.

$$h_{vw}^0 = \tau\big(W_i cat(x_v, e_{vw})\big) \quad (1)$$

The first phase of message passing computes interactions between atom v and atom w. This is achieved by summing the hidden states of all bonds connected to atom v while excluding the hidden state of bonds from atom w, as described by m_{vw}^{t+1} in eq 2. This step captures information about the neighboring atoms and their connections to individual atoms within vw. In eq 2, $h_{kv}^t \in R^D$ represents bond features at layer $t \in \{1,2,\dots,T\}$, $k \in \{N(v)\backslash w\}$ denotes the set of nodes connected to v excluding w.

$$m_{vw}^{t+1} = \sum_{k\in\{N(v)\backslash w\}} h_{kv}^t \quad (2)$$

Following this, a new hidden message at depth 1 is created by summing the product of the initial hidden state and the learned matrix $\mathrm{W_m}$ with the message. This resultant output undergoes further processing using the activation function τ, denoted as $\mathrm{h_{vw}^{t+1}}$ in eq 3.

$$h_{vw}^{t+1} = \tau(h_{vw}^0 + W_m m_{vw}^{t+1}) \quad (3)$$

In the final message passing layer (at t=T), the updated hidden states $\mathrm{h_{vw}^T}$ are summed to create the ultimate message for each atom, as described in eq 4. This action aggregates information about all neighboring atoms and their relationships into the final message for each atom.

$$m_v = \sum_{w\in N(v)} h_{vw}^T \quad (4)$$

The hidden state $\mathrm{h_v}$ for each atom is derived by concatenating the initial atom features with the message vector, as indicated in eq 5.

$$h_v = \tau(W_i cat(x_v, m_v)) \quad (5)$$

Finally, employing eq 6, the hidden states $\mathrm{h_v}$ of each atom are summed to generate a molecular feature vector. This step aggregate information from all atoms in the molecule into a unified molecular feature vector, facilitating property prediction. It encompasses both structural and attribute information of the entire molecule, offering a comprehensive representation for further property prediction.

$$h = \sum_{v \epsilon N(v)} h_v \quad (6)$$

In the DMPNN-Des model, preceding the readout phase, vector h is concatenated with descriptor vectors and collectively processed using a fully connected feedforward neural network to predict properties. The algorithm is implemented using the open-source Chemprop package (2). Regarding the raw datasets, we trained each dataset using Chemprop, employing random segmentation ratios [0.8, 0.1, 0.1] for training, testing, and evaluation. The batch process is iterated five times, and the average RMSE value and variance from these iterations are calculated to evaluate the robustness of the model.

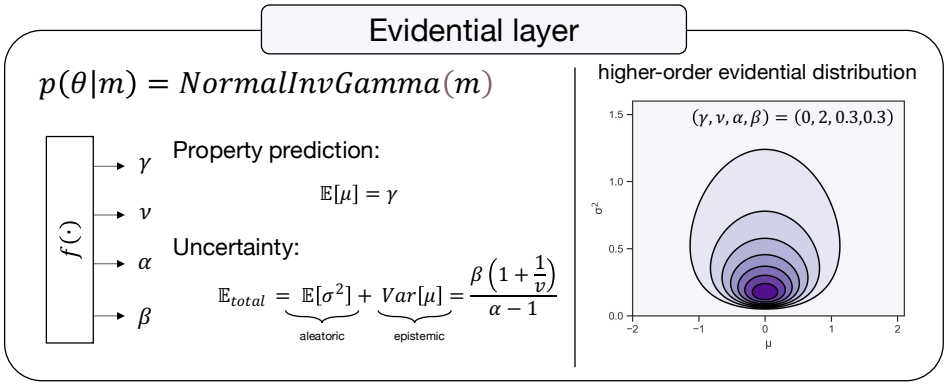
Uncertainty estimation method

In pharmaceutical research, especially in ADMET assessment, unreliable predictions might lead to misjudging drug efficacy, causing missed opportunities in drug development. Therefore, within AI-assisted drug development, quantifying predictive reliability is crucial for guiding subsequent research decisions by medicinal chemists (3).

Regression model

Within the regression model, ADMETlab 3.0 employs the evidence-based deep learning technique proposed by Amini et al (4). Evidential deep learning extends the concept of learning probability distribution parameters, further predicting higher-order distributions of the original likelihood parameters themselves. These higher-order parameters define the evidence distribution, thereby capturing the model's predictions and the degree of evidence associated with those predictions. In contrast to Bayesian neural networks that set priors on neural network weights, the evidence learning approach estimates uncertainty by directly learning the parameters defining this evidence distribution. It encompasses both epistemic and aleatoric uncertainties, eliminating the need for sampling and thus obviating the necessity of sampling procedures.

In a regression setting, the training samples comprise $\mathrm{D} = \{\mathrm{x_i}, \mathrm{y_i}\}_{\mathrm{i=1}}^{\mathrm{N}}$, where the target values $\mathrm{y_i} \in \mathbb{R}$ comprise an i.i.d. Gaussian distribution defined by mean and variance $\theta = \{\mu,\ \sigma^2\}$. Within the context of an evidential depth model, these parameters are presumed unknown and replaced by probabilistic estimates, achieved by placing a Gaussian distribution on the unknown mean μ and an Inverse-Gamma prior on the unknown variance σ^2. We obtain a higher-order distribution (also termed the evidential distribution), depicted as $\mathrm{p}(\theta|\mathrm{m})$, represented by a Normal Inverse-Gamma distribution. This evidential distribution is determined through four parameters, m $=\{\gamma, \nu, \alpha, \beta\}$. The model can capture predictive uncertainty by learning these parameters, as demonstrated in Figure 2. In this work, the uncertainty estimation of the regression model is achieved by setting the chemprop package(2) uncertainty_method to evidential_total.



Moreover, to better assist users in assessing the reliability of model predictions based on model uncertainty, furthermore, we furnish the range of model RMSE within different uncertainty intervals. This aids users in assessing the reliability of model predictions based on varying levels of uncertainty.

Classification model

In the classification model, uncertainty is estimated using the Monte Carlo dropout t approach (5). This method considers dropout in deep neural networks as an approximate Bayesian inference of deep Gaussian processes. Specifically, the dropout technique involves applying dropout before each layer during training and maintaining dropout activation during the inference process. This allows for the generation of prediction distributions using different random masks, approximating the posterior of deep Gaussian processes. The variance of this distribution serves as an estimation of predictive uncertainty (6-8).

During experiments with uncertainty quantification, an ensemble of models generates a prediction distribution for each molecule using a dropout-enabled network with a sample size of 10. The probability of 0.1 to use for Monte Carlo dropout uncertainty estimation. Let y_t represent the prediction from a single model within the ensemble, which contains T = 10 models. For a query sample x_i, the prediction $\hat{y}$ is represented as the means of all predictions and the uncertainty of this sample $U(x)$ can be provided by the variance σ_t^2 of the prediction distribution.

$$U(x) = \sigma^2 = \frac{1}{T-1}\sum_{t=1}^{T}(y_t - \hat{y})^2 \quad (7)$$

We employed the method proposed by Dolezal et al (9). to determine the optimal uncertainty threshold value, θ. This method establishes an uncertainty threshold, where predictions below this threshold are more likely to be correct than those with higher levels of uncertainty. To find the uncertainty threshold that optimally separates predictions into likely-correct (high-confidence) and likely-incorrect (low-confidence), we calculated the sensitivity and specificity for misprediction for all possible uncertainty thresholds. The corresponding Youden's index (J) for each uncertainty threshold θ_i is the calculated as

$$J_i = Se_i + Sp_i - 1 \quad (8)$$

The optimal uncertainty threshold θ is the defined as the threshold which maximized the Youden's index:

$$\theta = argmax J_i \quad (9)$$

The single threshold is then used for all predicted made by the model. We take a binary a binary approach to confidence using the uncertainty threshold, with confidence of the classification model defined as

$$C(x) = \begin{cases} high - confidence & \sigma^2(x) < \theta \\ low - confidence & \sigma^2(x) \geq \theta \end{cases} \quad (10)$$

In other words, prediction uncertainty exceeding this value designates the model's prediction as low confidence, while prediction uncertainty below this threshold indicates high confidence in the model's prediction. This threshold will be used to assess the reliability of prediction in classification tasks within the ADMET models.

Supplementary Table 1. Data information of 77 predictive models

Dataset	Total (Positive/Negative)	Training set (Positive/Negative)	Test set (Positive/Negative)	Validation set (Positive/Negative)
logS	4797	3837	480	480
logD7.4	19155	15324	1915	1916
logP	12682	10145	1268	1269
melting point	9384	7507	938	939
boiling point	5758	4606	576	576
pKa (acidic)	2750	2200	275	275
pKa (basic)	2992	2393	299	300
Caco-2 Permeability	6502	5201	650	651
MDCK Permeability	1140	912	114	114
Pgp inhibitor	2209 (1315/894)	1767 (1052/715)	220 (131/89)	222 (132/90)
Pgp substrate	1185 (586/599)	947 (468/479)	119 (59/60)	119 (59/60)
HIA	1160 (1022/138)	927 (817/110)	116 (102/14)	117 (103/14)
F20%	992 (753/239)	793 (602/191)	99 (75/24)	100 (76/24)
F30%	992 (666/326)	792 (532/260)	100 (67/33)	100 (67/33)
F50%	1474 (800/674)	1179 (640/539)	147 (80/67)	148 (80/68)
PAMPA	2035 (651/1384)	1627 (520/1107)	203 (65/138)	205 (66/139)
BCRP inhibitor	2799 (1098/1701)	2238 (878/1360)	280 (110/170)	281 (110/171)
MRP1 inhibitor	1015 (636/379)	811 (508/303)	102 (64/38)	102 (64/38)
BSEP inhibitor	763 (475/288)	610 (380/230)	76 (47/29)	77 (48/29)
OATP1B1 inhibitor	2372 (1343/1029)	1897 (1074/823)	237 (134/103)	238 (135/103)
OATP1B3 inhibitor	2228 (1540/688)	1782 (1232/550)	223 (154/69)	223 (154/69)
BBB	2865 (1649/1216)	2291 (1319/972)	287 (165/122)	287 (165/122)
PPB	4712	3769	471	472
VDss	2440	1952	244	244
Fu	2575	2060	257	258
CYP1A2 inhibitor	12614 (5872/6742)	10090 (4697/5393)	1261 (587/674)	1263 (588/675)
CYP1A2 substrate	366 (176/190)	292 (140/152)	37 (18/19)	37 (18/19)
CYP2C19 inhibitor	12611 (5770/6841)	10088 (4616/5472)	1261 (577/684)	1262 (577/685)
CYP2C19 substrate	256 (107/149)	204 (85/119)	26 (11/15)	26 (11/15)
CYP2C9 inhibitor	12089 (4014/8075)	9671 (3211/6460)	1208 (401/807)	1210 (402/808)
CYP2C9 substrate	811 (325/486)	648 (260/388)	81 (32/49)	82 (33/49)
CYP2D6 inhibitor	13073 (2535/10538)	10458 (2028/8430)	1307 (253/1054)	1308 (254/1054)
CYP2D6 substrate	877 (435/442)	701 (348/353)	87 (43/44)	89 (44/45)
CYP3A4 inhibitor	12339 (5092/7247)	9870 (4073/5797)	1234 (509/725)	1235 (510/725)
CYP3A4 substrate	979 (497/482)	782 (397/385)	98 (50/48)	99 (50/49)
CYP2B6 inhibitor	551 (197/354)	440 (157/283)	55 (20/35)	56 (20/36)
CYP2B6 substrate	2859 (260/2599)	2287 (208/2079)	286 (26/260)	286 (26/260)
CYP2C8 inhibitor	533 (312/221)	425 (249/176)	53 (31/22)	55 (32/23)
HLM Stability	15521 (6071/9450)	12416 (4856/7560)	1552 (607/945)	1553 (608/945)
CL-plasma	831	664	83	84
T1/2	3427	2741	343	343
hERG Blocker	13845 (6922/6923)	11075 (5537/5538)	1384 (692/692)	1386 (693/693)
Human Hepatotoxicity	2304 (1299/1005)	1843 (1039/804)	230 (130/100)	231 (130/101)
DILI	467 (235/232)	373 (188/185)	46 (23/23)	48 (24/24)
AMES Mutagenicity	7575 (4222/3353)	6059 (3377/2682)	757 (422/335)	759 (423/336)
ROA	7327 (2799/4528)	5861 (2239/3622)	733 (280/453)	733 (280/453)
FDAMDD	1197 (561/636)	956 (448/508)	120 (56/64)	121 (57/64)
Skin Sensitization	405 (274/131)	323 (219/104)	40 (27/13)	42 (28/14)
Carcinogenicity	1041 (516/525)	832 (412/420)	104 (52/52)	105 (52/53)
Eye Corrosion	2298 (886/1412)	1837 (708/1129)	230 (89/141)	231 (89/142)
Eye Irritation	5219 (3874/1345)	4175 (3099/1076)	521 (387/134)	523 (388/135)
Respiratory	1388 (835/553)	1110 (668/442)	138 (83/55)	140 (84/56)
A549 Cytotoxicity	3302 (277/3025)	2641 (221/2420)	330 (28/302)	331 (28/303)
hERG Blocker 10um	9876 (5090/4786)	7900 (4072/3828)	988 (509/479)	988 (509/479)
Hematotoxicity	2374 (752/1622)	1898 (601/1297)	237 (75/162)	239 (76/163)
Hek293 Cytotoxicity	6898 (1455/5443)	5518 (1164/4354)	689 (145/544)	691 (146/545)
Genotoxicity	641 (264/377)	512 (211/301)	64 (26/38)	65 (27/38)
Drug-induced Nephrotoxicity	565 (287/278)	451 (229/222)	57 (29/28)	57 (29/28)
Drug-induced Neurotoxicity	684 (329/355)	547 (263/284)	68 (33/35)	69 (33/36)
Ototoxicity	2808 (1103/1705)	2246 (882/1364)	280 (110/170)	282 (111/171)
RPMI-8226 Immunitoxicity	48811 (3877/44934)	39048 (3101/35947)	4881 (388/4493)	4882 (388/4494)

BCF	676	540	68	68
IGC50	1787	1429	179	179
LC50FM	816	652	82	82
LC50DM	347	277	35	35
NR-AR	7312 (266/7046)	5848 (212/5636)	732 (27/705)	732 (27/705)
NR-AR-LBD	6862 (233/6629)	5489 (186/5303)	686 (23/663)	687 (24/663)
NR-AhR	6603 (763/5840)	5282 (610/4672)	660 (76/584)	661 (77/584)
NR-Aromatase	5887 (256/5631)	4708 (204/4504)	589 (26/563)	590 (26/564)
NR-ER	6166 (669/5497)	4932 (535/4397)	617 (67/550)	617 (67/550)
NR-ER-LBD	7052 (342/6710)	5641 (273/5368)	705 (34/671)	706 (35/671)
NR-PPAR-gamma	6586 (197/6389)	5268 (157/5111)	659 (20/639)	659 (20/639)
SR-ARE	5652 (865/4787)	4521 (692/3829)	565 (86/479)	566 (87/479)
SR-ATAD5	7170 (249/6921)	5735 (199/5536)	717 (25/692)	718 (25/693)
SR-HSE	6319 (360/5959)	5055 (288/4767)	632 (36/596)	632 (36/596)
SR-MMP	5913 (892/5021)	4729 (713/4016)	591 (89/502)	593 (90/503)
SR-p53	6915 (456/6459)	5531 (364/5167)	692 (46/646)	692 (46/646)

Supplementary Table 2. Optimal hyperparameters for DMPNN models

Hyperparameter	Regression					Classification			
	Absorption	Distribution	Excretion	PCP	Toxicity	Absorption	Metabolism	Tox21	Toxicity
depth	2	5	4	3	4	2	4	4	6
dropout	0.1	0.1	0.1	0.15	0.05	0.4	0.3	0	0.35
ffn_hidden_size	1900	1300	1900	2200	1200	1900	1900	1300	2300
ffn_num_layers	3	3	3	3	3	3	3	3	1
hidden_size	900	1200	1100	900	1100	1900	900	600	1400
batch_size	128	128	128	128	128	128	128	128	128
epochs	200	200	200	200	200	200	200	200	200

Supplementary Table 3. Optimal hyperparameters for DMPNN-Des models

Hyperparameter	Regression					Cassification			
	Absorption	Distribution	Excretion	PCP	Toxicity	Absorption	Metabolism	Tox21	Toxicity
depth	5	5	4	4	4	4	5	6	4
dropout	0.25	0.3	0.3	0.1	0.15	0.05	0.15	0.25	0.3
ffn_hidden_size	800	800	300	2000	300	1400	400	400	1900
ffn_num_layers	1	3	1	3	3	1	3	3	2
hidden_size	2000	2000	1100	1000	2400	1200	600	1200	800
batch_size	128	128	128	128	128	128	128	128	128
epochs	200	200	200	200	200	200	200	200	200

Supplementary Table 4. Predictive performance of regression models

Dataset	Subset	DMPNN-Des			DMPNN			MGA		
		R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
logS	Test set	0.877±0.013	0.746±0.027	0.495±0.010	0.869±0.010	0.773±0.018	0.516±0.006	0.764±0.026	1.034±0.042	0.775±0.042
	Validation set	0.885±0.011	0.740±0.028	0.501±0.015	0.880±0.009	0.758±0.024	0.515±0.011	0.772±0.026	1.042±0.054	0.793±0.039
logD7.4	Test set	0.902±0.004	0.398±0.008	0.290±0.006	0.894±0.003	0.414±0.006	0.305±0.005	0.579±0.021	0.822±0.016	0.649±0.014
	Validation set	0.896±0.006	0.404±0.011	0.297±0.009	0.889±0.005	0.417±0.011	0.311±0.011	0.554±0.018	0.837±0.019	0.661±0.016
logP	Test set	0.958±0.003	0.378±0.017	0.253±0.009	0.956±0.004	0.385±0.018	0.263±0.011	0.780±0.016	0.861±0.026	0.662±0.019
	Validation set	0.959±0.005	0.375±0.018	0.257±0.007	0.956±0.005	0.388±0.016	0.266±0.004	0.783±0.014	0.861±0.020	0.665±0.018
melting point	Test set	0.862±0.009	37.423±1.466	26.646±0.925	0.852±0.009	38.741±1.441	27.709±0.832	0.859±0.028	37.658±3.978	27.45±2.896
	Validation set	0.861±0.012	36.868±1.748	26.236±0.968	0.847±0.010	38.692±1.356	27.884±0.989	0.854±0.033	37.540±4.725	27.421±3.41
boiling point	Test set	0.940±0.012	20.360±2.230	12.099±1.393	0.926±0.011	22.607±2.036	14.522±1.242	0.900±0.053	25.644±7.215	15.543±2.113
	Validation set	0.942±0.009	20.874±1.321	12.548±0.56	0.929±0.008	23.112±0.964	14.65±0.560	0.932±0.019	22.478±2.938	15.637±1.91
pKa (acidic)	Test set	0.880±0.013	1.164±0.068	0.737±0.065	0.880±0.008	1.163±0.040	0.737±0.027	0.649±0.036	1.990±0.080	1.437±0.071
	Validation set	0.892±0.013	1.102±0.068	0.698±0.019	0.886±0.012	1.131±0.079	0.721±0.062	0.645±0.036	1.999±0.143	1.431±0.094
pKa (basic)	Test set	0.872±0.026	1.142±0.118	0.707±0.055	0.868±0.013	1.164±0.073	0.749±0.031	0.689±0.037	1.785±0.129	1.328±0.050
	Validation set	0.870±0.024	1.163±0.138	0.715±0.067	0.870±0.030	1.158±0.147	0.744±0.072	0.702±0.023	1.764±0.146	1.294±0.102
Caco-2 Permeability	Test set	0.743±0.018	0.325±0.013	0.242±0.011	0.742±0.020	0.325±0.007	0.241±0.005	0.652±0.023	0.378±0.009	0.293±0.004
	Validation set	0.758±0.015	0.314±0.011	0.236±0.009	0.752±0.014	0.317±0.004	0.234±0.003	0.691±0.028	0.354±0.012	0.276±0.008
MDCK Permeability	Test set	0.700±0.062	0.293±0.034	0.205±0.010	0.695±0.058	0.296±0.035	0.203±0.008	0.629±0.074	0.327±0.043	0.237±0.019
	Validation set	0.693±0.070	0.297±0.050	0.196±0.022	0.677±0.074	0.304±0.053	0.197±0.023	0.580±0.101	0.346±0.054	0.245±0.020
PPB	Test set	0.824±0.031	11.382±0.648	5.976±0.429	0.788±0.056	12.439±1.182	6.594±0.668	0.705±0.054	14.771±0.715	9.041±0.538
	Validation set	0.820±0.035	11.508±0.880	6.151±0.489	0.804±0.029	12.049±0.774	6.523±0.488	0.746±0.045	13.684±0.664	8.662±0.405
VDss	Test set	0.760±0.045	0.301±0.024	0.162±0.010	0.742±0.050	0.312±0.026	0.173±0.016	0.581±0.032	0.399±0.009	0.298±0.008
	Validation set	0.813±0.040	0.260±0.029	0.143±0.018	0.811±0.042	0.261±0.028	0.153±0.020	0.594±0.030	0.385±0.014	0.295±0.013
Fu	Test set	0.894±0.024	0.229±0.027	0.135±0.012	0.879±0.027	0.244±0.029	0.149±0.020	0.695±0.041	0.390±0.018	0.289±0.017
	Validation set	0.885±0.017	0.250±0.024	0.148±0.008	0.885±0.015	0.250±0.017	0.153±0.010	0.731±0.015	0.383±0.017	0.287±0.012
CL-plasma	Test set	0.667±0.046	2.912±0.311	1.783±0.160	0.635±0.070	3.053±0.441	2.006±0.285	0.627±0.165	2.957±0.353	2.060±0.273
	Validation set	0.655±0.106	3.028±0.371	1.817±0.211	0.587±0.108	3.323±0.360	2.150±0.245	0.662±0.081	2.744±0.283	1.924±0.221
T1/2	Test set	0.653±0.070	0.877±0.111	0.420±0.026	0.629±0.110	0.900±0.141	0.449±0.037	0.638±0.147	1.013±0.200	0.535±0.064
	Validation set	0.748±0.052	0.965±0.143	0.455±0.012	0.740±0.071	0.983±0.171	0.464±0.027	0.745±0.149	0.919±0.152	0.522±0.044
Bioconcentration Factor	Test set	0.822±0.073	0.566±0.108	0.430±0.071	0.837±0.054	0.546±0.086	0.418±0.055	0.746±0.074	0.674±0.107	0.504±0.069
	Validation set	0.786±0.047	0.608±0.042	0.454±0.029	0.749±0.074	0.655±0.065	0.475±0.041	0.793±0.071	0.630±0.112	0.467±0.064
IGC50	Test set	0.862±0.016	0.382±0.018	0.273±0.013	0.856±0.020	0.391±0.032	0.281±0.023	0.792±0.029	0.468±0.029	0.348±0.018
	Validation set	0.890±0.026	0.350±0.041	0.249±0.027	0.884±0.018	0.363±0.033	0.262±0.017	0.832±0.033	0.439±0.017	0.334±0.014
LC50FM	Test set	0.686±0.053	0.909±0.125	0.669±0.070	0.634±0.068	0.986±0.174	0.727±0.116	0.587±0.201	0.998±0.133	0.685±0.052
	Validation set	0.702±0.120	0.968±0.186	0.665±0.094	0.628±0.169	1.075±0.214	0.730±0.129	0.747±0.099	0.738±0.081	0.543±0.040
LC50DM	Test set	0.727±0.073	0.747±0.048	0.503±0.041	0.695±0.123	0.778±0.098	0.526±0.046	0.699±0.064	0.805±0.086	0.580±0.030

	Validation set	0.751±0.061	0.742±0.083	0.497±0.047	0.748±0.066	0.747±0.089	0.506±0.056	0.692±0.083	0.804±0.120	0.591±0.082

Supplementary Table 5. Predictive performance of classification models

Dataset	Subset	DMPNN-Des			DMPNN			MGA		
		AUC	ACC	MCC	AUC	ACC	MCC	AUC	ACC	MCC
Pgp-inhibitor	Test set	0.915±0.011	0.850±0.010	0.687±0.021	0.896±0.015	0.833±0.007	0.652±0.014	0.915±0.014	0.838±0.016	0.662±0.033
	Validation set	0.908±0.007	0.841±0.029	0.670±0.063	0.895±0.008	0.827±0.019	0.641±0.041	0.928±0.015	0.867±0.017	0.724±0.036
Pgp-substrate	Test set	0.892±0.019	0.815±0.034	0.634±0.068	0.894±0.010	0.829±0.022	0.661±0.041	0.890±0.022	0.807±0.035	0.616±0.071
	Validation set	0.900±0.021	0.831±0.021	0.665±0.042	0.896±0.022	0.821±0.019	0.645±0.037	0.898±0.031	0.805±0.039	0.610±0.079
HIA	Test set	0.897±0.054	0.911±0.018	0.571±0.090	0.891±0.037	0.894±0.027	0.508±0.133	0.901±0.067	0.877±0.032	0.548±0.114
	Validation set	0.933±0.038	0.927±0.022	0.646±0.133	0.929±0.046	0.931±0.016	0.685±0.084	0.921±0.051	0.898±0.036	0.629±0.105
F20%	Test set	0.909±0.047	0.857±0.030	0.616±0.087	0.907±0.024	0.855±0.034	0.579±0.101	0.818±0.036	0.757±0.051	0.404±0.094
	Validation set	0.919±0.027	0.875±0.032	0.668±0.086	0.909±0.023	0.853±0.030	0.616±0.071	0.799±0.066	0.774±0.050	0.461±0.133
F30%	Test set	0.919±0.026	0.851±0.040	0.666±0.094	0.915±0.030	0.853±0.031	0.666±0.069	0.785±0.070	0.694±0.085	0.356±0.153
	Validation set	0.911±0.035	0.869±0.034	0.701±0.078	0.911±0.025	0.853±0.025	0.668±0.060	0.818±0.072	0.753±0.040	0.457±0.105
F50%	Test set	0.791±0.025	0.716±0.033	0.432±0.063	0.760±0.031	0.705±0.035	0.413±0.062	0.762±0.012	0.703±0.021	0.409±0.043
	Validation set	0.804±0.020	0.715±0.015	0.429±0.029	0.787±0.017	0.704±0.021	0.416±0.057	0.753±0.042	0.679±0.036	0.360±0.074
PAMPA	Test set	0.731±0.017	0.735±0.024	0.364±0.055	0.717±0.023	0.718±0.028	0.316±0.071	0.771±0.029	0.729±0.017	0.411±0.035
	Validation set	0.761±0.018	0.741±0.015	0.384±0.028	0.745±0.027	0.735±0.013	0.363±0.046	0.784±0.019	0.748±0.013	0.443±0.039
BCRP inhibitor	Test set	0.949±0.015	0.892±0.013	0.777±0.030	0.942±0.008	0.870±0.007	0.731±0.016	0.938±0.022	0.864±0.027	0.725±0.053
	Validation set	0.948±0.012	0.891±0.011	0.775±0.025	0.946±0.016	0.882±0.022	0.760±0.042	0.940±0.012	0.862±0.036	0.727±0.067
MRP1 inhibitor	Test set	0.956±0.023	0.928±0.019	0.846±0.042	0.953±0.021	0.924±0.014	0.839±0.027	0.946±0.021	0.876±0.015	0.750±0.025
	Validation set	0.971±0.008	0.938±0.017	0.869±0.036	0.974±0.009	0.944±0.007	0.883±0.016	0.950±0.021	0.844±0.021	0.692±0.044
OATP1B1inhibitor	Test set	0.826±0.034	0.760±0.044	0.517±0.083	0.812±0.034	0.755±0.030	0.500±0.062	0.720±0.024	0.660±0.025	0.318±0.046
	Validation set	0.850±0.024	0.783±0.016	0.561±0.037	0.829±0.012	0.770±0.019	0.530±0.040	0.706±0.042	0.664±0.040	0.327±0.080
OATP1B3 inhibitor	Test set	0.882±0.022	0.833±0.025	0.614±0.058	0.878±0.020	0.832±0.017	0.607±0.037	0.744±0.045	0.665±0.033	0.350±0.048
	Validation set	0.884±0.019	0.846±0.031	0.647±0.057	0.874±0.019	0.844±0.012	0.631±0.028	0.725±0.031	0.669±0.041	0.325±0.077
BSEP inhibitor	Test set	0.940±0.013	0.883±0.023	0.747±0.050	0.918±0.026	0.856±0.018	0.692±0.042	0.922±0.028	0.864±0.047	0.715±0.094
	Validation set	0.935±0.028	0.874±0.038	0.733±0.080	0.909±0.055	0.844±0.045	0.667±0.101	0.948±0.018	0.905±0.041	0.799±0.089
BBB Penetration	Test set	0.908±0.004	0.836±0.014	0.666±0.030	0.892±0.019	0.817±0.019	0.626±0.041	0.889±0.013	0.817±0.019	0.624±0.040
	Validation set	0.905±0.018	0.844±0.014	0.680±0.031	0.888±0.009	0.824±0.008	0.642±0.018	0.905±0.017	0.826±0.011	0.643±0.023
CYP1A2 inhibitor	Test set	0.942±0.003	0.874±0.003	0.748±0.006	0.926±0.005	0.858±0.006	0.715±0.012	0.931±0.005	0.847±0.008	0.695±0.016
	Validation set	0.935±0.003	0.864±0.005	0.729±0.010	0.917±0.002	0.851±0.009	0.701±0.018	0.929±0.005	0.851±0.006	0.702±0.013
CYP1A2 substrate	Test set	0.806±0.062	0.739±0.092	0.480±0.191	0.824±0.063	0.717±0.059	0.455±0.121	0.767±0.059	0.683±0.042	0.379±0.078
	Validation set	0.850±0.012	0.779±0.032	0.562±0.061	0.863±0.015	0.795±0.031	0.593±0.061	0.833±0.062	0.790±0.056	0.585±0.115
CYP2C19 inhibitor	Test set	0.915±0.005	0.843±0.009	0.689±0.016	0.902±0.005	0.831±0.008	0.660±0.018	0.899±0.006	0.824±0.017	0.655±0.028
	Validation set	0.921±0.011	0.846±0.017	0.698±0.032	0.906±0.008	0.839±0.010	0.677±0.019	0.910±0.008	0.831±0.011	0.669±0.020
CYP2C19 substrate	Test set	0.884±0.030	0.797±0.033	0.583±0.066	0.850±0.049	0.822±0.073	0.636±0.154	0.841±0.056	0.773±0.077	0.552±0.127
	Validation set	0.829±0.115	0.778±0.104	0.551±0.220	0.796±0.081	0.713±0.107	0.421±0.221	0.798±0.092	0.698±0.078	0.394±0.153
CYP2C9 inhibitor	Test set	0.917±0.009	0.854±0.009	0.682±0.020	0.902±0.009	0.844±0.012	0.652±0.023	0.902±0.008	0.816±0.014	0.632±0.023
	Validation set	0.921±0.008	0.853±0.009	0.685±0.016	0.907±0.008	0.843±0.012	0.651±0.025	0.899±0.006	0.808±0.008	0.621±0.017

CYP2C9 substrate	Test set	0.782±0.042	0.720±0.049	0.433±0.100	0.777±0.009	0.700±0.026	0.384±0.040	0.728±0.039	0.678±0.056	0.372±0.099
	Validation set	0.783±0.025	0.730±0.044	0.438±0.092	0.808±0.045	0.754±0.055	0.489±0.117	0.744±0.051	0.646±0.047	0.315±0.068
CYP2D6 inhibitor	Test set	0.886±0.010	0.869±0.012	0.575±0.033	0.870±0.009	0.868±0.010	0.572±0.024	0.891±0.005	0.796±0.023	0.512±0.035
	Validation set	0.898±0.010	0.880±0.006	0.608±0.016	0.875±0.013	0.865±0.007	0.564±0.019	0.889±0.005	0.801±0.021	0.521±0.035
CYP2D6 substrate	Test set	0.844±0.057	0.754±0.051	0.514±0.099	0.802±0.043	0.752±0.035	0.505±0.068	0.811±0.053	0.736±0.036	0.478±0.079
	Validation set	0.836±0.027	0.766±0.033	0.539±0.063	0.801±0.017	0.751±0.017	0.502±0.034	0.790±0.028	0.726±0.033	0.456±0.066
CYP3A4 inhibitor	Test set	0.916±0.005	0.834±0.009	0.667±0.020	0.887±0.007	0.813±0.008	0.615±0.018	0.908±0.013	0.821±0.020	0.647±0.037
	Validation set	0.913±0.007	0.835±0.011	0.666±0.020	0.892±0.007	0.816±0.006	0.622±0.012	0.905±0.006	0.814±0.010	0.632±0.014
CYP3A4 substrate	Test set	0.798±0.034	0.720±0.012	0.444±0.027	0.764±0.043	0.715±0.032	0.434±0.065	0.817±0.031	0.749±0.051	0.500±0.103
	Validation set	0.836±0.028	0.760±0.030	0.525±0.060	0.820±0.027	0.740±0.036	0.481±0.073	0.835±0.020	0.758±0.011	0.518±0.023
CYP2B6 inhibitor	Test set	0.808±0.061	0.770±0.030	0.505±0.063	0.821±0.024	0.741±0.029	0.434±0.055	0.806±0.109	0.733±0.118	0.462±0.219
	Validation set	0.859±0.038	0.782±0.039	0.540±0.100	0.865±0.055	0.786±0.042	0.544±0.093	0.839±0.047	0.768±0.046	0.530±0.094
CYP2B6 substrate	Test set	0.927±0.013	0.947±0.010	0.679±0.071	0.950±0.013	0.950±0.011	0.688±0.074	0.914±0.022	0.884±0.034	0.529±0.099
	Validation set	0.909±0.024	0.945±0.010	0.638±0.053	0.889±0.018	0.934±0.012	0.568±0.079	0.892±0.021	0.878±0.034	0.470±0.051
CYP2C8 inhibitor	Test set	0.918±0.032	0.823±0.064	0.637±0.132	0.918±0.026	0.838±0.039	0.666±0.081	0.888±0.034	0.811±0.044	0.617±0.088
	Validation set	0.932±0.032	0.862±0.034	0.717±0.072	0.911±0.038	0.851±0.042	0.697±0.086	0.931±0.034	0.833±0.039	0.667±0.080
HLM Stability	Test set	0.882±0.007	0.808±0.009	0.609±0.019	0.877±0.006	0.803±0.004	0.592±0.008	0.830±0.013	0.731±0.009	0.478±0.022
	Validation set	0.885±0.008	0.805±0.014	0.602±0.027	0.875±0.002	0.806±0.008	0.597±0.013	0.829±0.011	0.727±0.017	0.473±0.015
hERG Blockers	Test set	0.937±0.006	0.828±0.017	0.680±0.026	0.933±0.003	0.863±0.008	0.729±0.017	0.934±0.005	0.866±0.005	0.732±0.010
	Validation set	0.937±0.006	0.829±0.017	0.682±0.028	0.934±0.004	0.865±0.008	0.733±0.015	0.935±0.011	0.871±0.012	0.743±0.025
Human Hepatotoxicity	Test set	0.811±0.047	0.729±0.039	0.452±0.081	0.822±0.015	0.723±0.019	0.441±0.042	0.755±0.046	0.690±0.049	0.379±0.102
	Validation set	0.794±0.018	0.712±0.023	0.413±0.047	0.811±0.032	0.720±0.033	0.428±0.070	0.765±0.027	0.691±0.040	0.383±0.075
DILI	Test set	0.860±0.052	0.787±0.075	0.590±0.143	0.902±0.037	0.826±0.027	0.658±0.050	0.891±0.030	0.804±0.015	0.625±0.022
	Validation set	0.899±0.039	0.825±0.052	0.667±0.094	0.895±0.052	0.792±0.056	0.597±0.113	0.900±0.034	0.821±0.067	0.650±0.126
AMES Mutagenicity	Test set	0.882±0.007	0.785±0.015	0.576±0.028	0.874±0.009	0.778±0.002	0.563±0.009	0.888±0.011	0.812±0.015	0.620±0.031
	Validation set	0.893±0.010	0.796±0.015	0.595±0.028	0.879±0.012	0.780±0.008	0.566±0.018	0.888±0.014	0.807±0.016	0.611±0.033
ROA	Test set	0.865±0.011	0.761±0.022	0.534±0.036	0.846±0.010	0.761±0.011	0.515±0.024	0.852±0.014	0.777±0.023	0.543±0.045
	Validation set	0.842±0.014	0.746±0.025	0.508±0.037	0.836±0.011	0.755±0.011	0.505±0.021	0.855±0.013	0.779±0.022	0.541±0.041
FDAMDD	Test set	0.862±0.031	0.778±0.031	0.572±0.053	0.861±0.036	0.783±0.034	0.574±0.073	0.843±0.051	0.753±0.067	0.513±0.128
	Validation set	0.788±0.032	0.702±0.040	0.426±0.074	0.772±0.027	0.710±0.030	0.434±0.060	0.818±0.038	0.757±0.033	0.527±0.058
Skin Sensitization	Test set	0.787±0.043	0.780±0.062	0.491±0.152	0.768±0.072	0.745±0.024	0.365±0.072	0.759±0.066	0.725±0.040	0.419±0.082
	Validation set	0.831±0.067	0.781±0.035	0.497±0.083	0.801±0.120	0.762±0.096	0.427±0.258	0.764±0.074	0.681±0.090	0.340±0.184
Carcinogencity	Test set	0.715±0.029	0.668±0.040	0.349±0.078	0.729±0.044	0.633±0.037	0.296±0.072	0.702±0.038	0.637±0.022	0.277±0.047
	Validation set	0.745±0.029	0.662±0.028	0.342±0.059	0.752±0.029	0.647±0.049	0.324±0.103	0.742±0.048	0.670±0.052	0.342±0.104
Eye Corrosion	Test set	0.993±0.003	0.971±0.006	0.940±0.012	0.990±0.004	0.948±0.005	0.896±0.012	0.991±0.005	0.957±0.009	0.911±0.017
	Validation set	0.996±0.002	0.968±0.010	0.934±0.021	0.992±0.002	0.961±0.006	0.922±0.012	0.988±0.003	0.953±0.016	0.902±0.032
Eye Irritation	Test set	0.980±0.003	0.948±0.007	0.864±0.020	0.974±0.004	0.918±0.016	0.780±0.044	0.977±0.008	0.943±0.011	0.854±0.024
	Validation set	0.973±0.005	0.946±0.007	0.857±0.018	0.967±0.006	0.919±0.006	0.784±0.017	0.981±0.007	0.944±0.011	0.861±0.024
Respiratory	Test set	0.867±0.04	0.775±0.042	0.524±0.091	0.836±0.024	0.761±0.021	0.493±0.047	0.874±0.019	0.793±0.012	0.573±0.031
	Validation set	0.861±0.023	0.783±0.019	0.541±0.041	0.858±0.017	0.769±0.029	0.512±0.063	0.881±0.025	0.794±0.039	0.586±0.069

A549 Cytotoxicity	Test set	0.820±0.017	0.905±0.007	0.351±0.043	0.795±0.032	0.920±0.006	0.314±0.067	0.846±0.028	0.800±0.025	0.352±0.044
	Validation set	0.863±0.029	0.904±0.011	0.383±0.058	0.846±0.017	0.927±0.007	0.449±0.070	0.853±0.028	0.802±0.037	0.354±0.038
hERG Blocker (10um)	Test set	0.840±0.017	0.691±0.023	0.426±0.035	0.828±0.009	0.699±0.009	0.429±0.013	0.844±0.014	0.767±0.015	0.535±0.030
	Validation set	0.844±0.012	0.698±0.029	0.448±0.046	0.833±0.012	0.704±0.014	0.449±0.027	0.839±0.009	0.758±0.012	0.516±0.025
Hek293 Cytotoxicity	Test set	0.848±0.017	0.832±0.005	0.535±0.025	0.831±0.013	0.839±0.009	0.497±0.016	0.856±0.021	0.787±0.025	0.498±0.043
	Validation set	0.837±0.023	0.818±0.014	0.503±0.045	0.825±0.027	0.833±0.013	0.465±0.040	0.866±0.022	0.793±0.027	0.511±0.040
Hematotoxicity	Test set	0.819±0.022	0.754±0.024	0.463±0.046	0.824±0.032	0.755±0.048	0.468±0.088	0.771±0.021	0.692±0.034	0.355±0.038
	Validation set	0.825±0.015	0.749±0.016	0.445±0.031	0.828±0.028	0.763±0.020	0.474±0.047	0.794±0.045	0.724±0.049	0.412±0.088
Genotoxicity	Test set	0.941±0.033	0.908±0.043	0.814±0.085	0.933±0.027	0.889±0.039	0.781±0.070	0.925±0.039	0.841±0.049	0.676±0.109
	Validation set	0.961±0.027	0.888±0.037	0.774±0.078	0.945±0.027	0.876±0.029	0.748±0.067	0.953±0.029	0.870±0.038	0.741±0.076
Drug-induced Nephrotoxicity	Test set	0.850±0.034	0.735±0.027	0.481±0.045	0.813±0.056	0.756±0.077	0.519±0.153	0.752±0.055	0.680±0.056	0.365±0.111
	Validation set	0.785±0.046	0.729±0.030	0.471±0.065	0.788±0.077	0.719±0.057	0.445±0.111	0.807±0.047	0.736±0.039	0.478±0.080
Drug-induced Neurotoxicity	Test set	0.804±0.044	0.701±0.051	0.417±0.109	0.808±0.047	0.731±0.060	0.484±0.108	0.823±0.031	0.725±0.052	0.456±0.097
	Validation set	0.784±0.048	0.694±0.058	0.399±0.120	0.820±0.030	0.720±0.039	0.458±0.083	0.820±0.081	0.749±0.063	0.500±0.127
Ototoxicity	Test set	0.776±0.017	0.669±0.049	0.361±0.058	0.786±0.011	0.720±0.013	0.411±0.020	0.759±0.010	0.733±0.016	0.434±0.034
	Validation set	0.786±0.012	0.692±0.035	0.406±0.053	0.791±0.027	0.738±0.028	0.450±0.059	0.790±0.027	0.744±0.025	0.458±0.062
RPMI-8226 Immunitoxicity	Test set	0.878±0.008	0.927±0.005	0.478±0.025	0.866±0.006	0.929±0.002	0.401±0.006	0.880±0.004	0.827±0.017	0.405±0.019
	Validation set	0.881±0.005	0.930±0.003	0.500±0.009	0.874±0.008	0.930±0.003	0.419±0.028	0.888±0.009	0.824±0.020	0.415±0.022
NR-AR	Test set	0.883±0.016	0.956±0.016	0.534±0.076	0.839±0.021	0.962±0.008	0.520±0.078	0.840±0.026	0.880±0.022	0.314±0.024
	Validation set	0.904±0.026	0.955±0.012	0.531±0.057	0.886±0.032	0.959±0.010	0.548±0.076	0.893±0.068	0.885±0.028	0.372±0.082
NR-AR-LBD	Test set	0.959±0.026	0.973±0.005	0.667±0.049	0.933±0.027	0.968±0.009	0.621±0.081	0.912±0.021	0.929±0.017	0.452±0.048
	Validation set	0.969±0.025	0.972±0.006	0.666±0.083	0.959±0.019	0.972±0.005	0.676±0.050	0.950±0.040	0.931±0.020	0.508±0.090
NR-AhR	Test set	0.924±0.011	0.862±0.040	0.550±0.047	0.894±0.023	0.898±0.009	0.558±0.023	0.909±0.021	0.837±0.012	0.500±0.027
	Validation set	0.938±0.015	0.872±0.027	0.571±0.024	0.927±0.012	0.904±0.019	0.608±0.058	0.930±0.023	0.847±0.021	0.537±0.048
NR-Aromatase	Test set	0.895±0.019	0.902±0.030	0.347±0.044	0.831±0.030	0.934±0.009	0.358±0.035	0.861±0.040	0.832±0.012	0.288±0.024
	Validation set	0.893±0.014	0.903±0.021	0.362±0.023	0.848±0.028	0.937±0.017	0.401±0.062	0.904±0.053	0.842±0.019	0.346±0.073
NR-ER	Test set	0.812±0.034	0.797±0.081	0.373±0.076	0.784±0.036	0.853±0.019	0.392±0.023	0.770±0.058	0.800±0.016	0.327±0.031
	Validation set	0.806±0.033	0.805±0.046	0.363±0.039	0.773±0.021	0.841±0.034	0.355±0.042	0.817±0.065	0.822±0.023	0.388±0.068
NR-ER-LBD	Test set	0.904±0.019	0.930±0.019	0.482±0.058	0.896±0.025	0.942±0.011	0.492±0.041	0.853±0.052	0.862±0.021	0.343±0.061
	Validation set	0.919±0.013	0.933±0.020	0.503±0.073	0.911±0.019	0.951±0.012	0.548±0.064	0.894±0.069	0.868±0.025	0.397±0.083
NR-PPAR-gamma	Test set	0.913±0.035	0.940±0.019	0.424±0.064	0.874±0.039	0.957±0.003	0.422±0.062	0.873±0.054	0.891±0.013	0.301±0.054
	Validation set	0.903±0.024	0.947±0.024	0.473±0.081	0.887±0.037	0.951±0.008	0.391±0.083	0.907±0.066	0.894±0.013	0.354±0.082
SR-ARE	Test set	0.854±0.016	0.774±0.060	0.448±0.058	0.819±0.026	0.824±0.009	0.443±0.048	0.837±0.008	0.780±0.015	0.412±0.032
	Validation set	0.880±0.016	0.780±0.059	0.481±0.050	0.839±0.026	0.819±0.009	0.454±0.045	0.870±0.040	0.794±0.025	0.465±0.065
SR-ATAD5	Test set	0.940±0.030	0.942±0.019	0.494±0.085	0.926±0.030	0.963±0.009	0.525±0.096	0.901±0.040	0.869±0.032	0.324±0.092
	Validation set	0.947±0.009	0.943±0.008	0.506±0.037	0.929±0.035	0.962±0.005	0.535±0.053	0.929±0.047	0.870±0.026	0.368±0.077
SR-HSE	Test set	0.860±0.020	0.899±0.026	0.417±0.044	0.814±0.040	0.909±0.018	0.359±0.050	0.836±0.032	0.830±0.029	0.303±0.046
	Validation set	0.869±0.039	0.895±0.045	0.412±0.071	0.829±0.046	0.915±0.031	0.399±0.061	0.881±0.068	0.852±0.029	0.378±0.094
SR-MMP	Test set	0.941±0.016	0.857±0.048	0.604±0.062	0.911±0.017	0.881±0.021	0.610±0.045	0.915±0.013	0.850±0.023	0.572±0.046
	Validation set	0.935±0.008	0.859±0.043	0.615±0.063	0.919±0.008	0.880±0.020	0.613±0.031	0.933±0.022	0.863±0.022	0.606±0.052

SR-p53	Test set	0.890±0.027	0.882±0.034	0.415±0.036	0.870±0.027	0.914±0.010	0.445±0.022	0.854±0.016	0.813±0.032	0.339±0.031
	Validation set	0.905±0.021	0.885±0.044	0.489±0.056	0.884±0.012	0.917±0.012	0.486±0.042	0.898±0.049	0.823±0.031	0.392±0.074

Supplementary Table 6. Regression model error at various uncertainty threshold

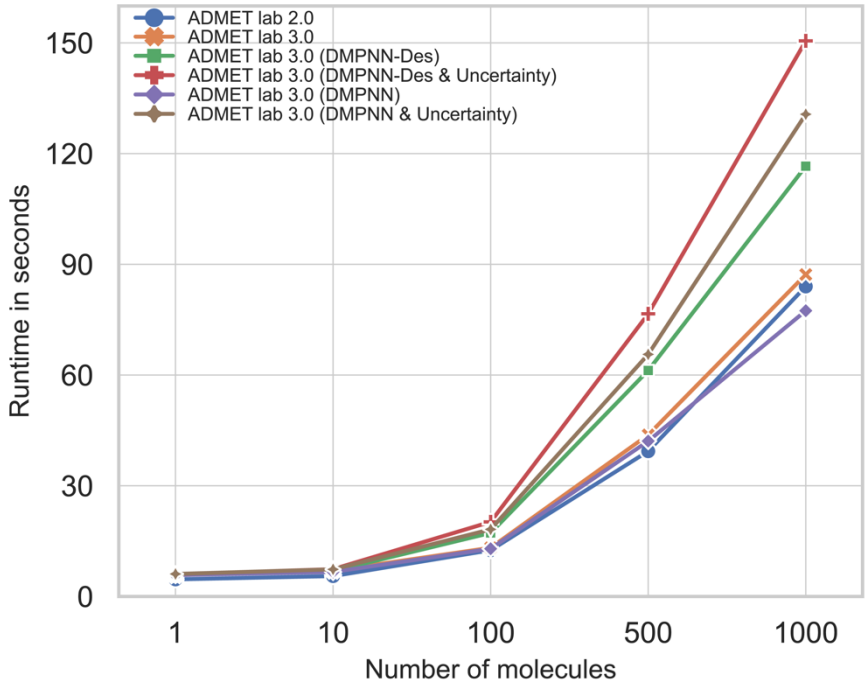
Dataset	DMPNN-Des		DMPNN	
	Uncertainty Interval	RMSE Interval	Uncertainty Interval	RMSE Interval
logS	[0,00000, 0.05255]	[0,000, 0.455]	[0,00000, 0.02043]	[0,000, 0.396]
	(0.05255, 0.07554]	(0.455, 0.502]	(0.02043, 0.02895]	(0.396, 0.475]
	(0.07554, 0.11779]	(0.502, 0.605]	(0.02895, 0.04577]	(0.475, 0.581]
logD7.4	[0,00000, 0.03084]	[0,000, 0.308]	[0,00000, 0.01615]	[0,000, 0.291]
	(0.03084, 0.03940]	(0.308, 0.326]	(0.01615, 0.02035]	(0.291, 0.323]
	(0.03940, 0.05330]	(0.326, 0.348]	(0.02035, 0.02746]	(0.323, 0.348]
logP	[0,00000, 0.01940]	[0,000, 0.235]	[0,00000, 0.00929]	[0,000, 0.209]
	(0.01940, 0.02572]	(0.235, 0.260]	(0.00929, 0.01244]	(0.209, 0.234]
	(0.02572, 0.03713]	(0.260, 0.282]	(0.01244, 0.01834]	(0.234, 0.261]
melting point	[0,00000, 172.21813]	[0,000, 27.034]	[0,00000, 73.31254]	[0,000, 27.048]
	(172.21813, 224.61369]	(27.034, 29.597]	(73.31254, 98.17039]	(27.048, 30.403]
	(224.61369, 315.17204]	(29.597, 32.751]	(98.17039, 140.63725]	(30.403, 31.989]
boiling point	[0,00000, 58.04890]	[0,000, 7.667]	[0,00000, 27.80054]	[0,000, 10.998]
	(58.04890, 76.09868]	(7.667, 10.598]	(27.80054, 36.23693]	(10.998, 11.454]
	(76.09868, 111.23996]	(10.598, 12.481]	(36.23693, 50.53419]	(11.454, 13.300]
pKa (acidic)	[0,00000, 0.08959]	[0,000, 0.394]	[0,00000, 0.02547]	[0,000, 0.230]
	(0.08959, 0.17127]	(0.394, 0.452]	(0.02547, 0.05581]	(0.230, 0.483]
	(0.17127, 0.34059]	(0.452, 0.773]	(0.05581, 0.11198]	(0.483, 0.658]
pKa (basic)	[0,00000, 0.12694]	[0,000, 0.427]	[0,00000, 0.04285]	[0,000, 0.362]
	(0.12694, 0.18051]	(0.427, 0.574]	(0.04285, 0.06346]	(0.362, 0.491]
	(0.18051, 0.29106]	(0.574, 0.686]	(0.06346, 0.10124]	(0.491, 0.680]
Caco-2 Permeability	[0,00000, 0.00506]	[0,000, 0.208]	[0,00000, 0.00147]	[0,000, 0.244]
	(0.00506, 0.00932]	(0.208, 0.243]	(0.00147, 0.00299]	(0.244, 0.269]
	(0.00932, 0.01785]	(0.243, 0.279]	(0.00299, 0.00732]	(0.269, 0.300]
MDCK Permeability	[0,00000, 0.00350]	[0,000, 0.187]	[0,00000, 0.00104]	[0,000, 0.201]
	(0.00350, 0.00561]	(0.187, 0.204]	(0.00104, 0.00190]	(0.201, 0.271]
	(0.00561, 0.01024]	(0.204, 0.251]	(0.00190, 0.00358]	(0.271, 0.292]
PPB	[0,00000, 0.12481]	[0,000, 2.336]	[0,00000, 0.15943]	[0,000, 5.906]
	(0.12481, 0.26645]	(2.336, 5.600]	(0.15943, 0.27067]	(5.906, 5.873]
	(0.26645, 1.52185]	(5.600, 7.412]	(0.27067, 1.05816]	(5.873, 7.309]
VDss	[0,00000, 0.00180]	[0,000, 0.164]	[0,00000, 0.00091]	[0,000, 0.172]
	(0.00180, 0.00269]	(0.164, 0.202]	(0.00091, 0.00130]	(0.172, 0.211]
	(0.00269, 0.00433]	(0.202, 0.243]	(0.00130, 0.00219]	(0.211, 0.253]
Fu	[0,00000, 0.00209]	[0,000, 0.201]	[0,00000, 0.00108]	[0,000, 0.177]
	(0.00209, 0.00314]	(0.201, 0.205]	(0.00108, 0.00173]	(0.177, 0.188]
	(0.00314, 0.00532]	(0.205, 0.236]	(0.00173, 0.00335]	(0.188, 0.238]
cl-plasma	[0,00000, 0.08178]	[0,000, 1.669]	[0,00000, 0.04806]	[0,000, 1.251]
	(0.08178, 0.16374]	(1.669, 1.928]	(0.04806, 0.08851]	(1.251, 1.617]
	(0.16374, 0.36927]	(1.928, 2.172]	(0.08851, 0.21065]	(1.617, 2.610]
t1/2	[0,00000, 0.00287]	[0,000, 0.201]	[0,00000, 0.00139]	[0,000, 0.180]
	(0.00287, 0.00518]	(0.201, 0.265]	(0.00139, 0.00234]	(0.180, 0.275]
	(0.00518, 0.01082]	(0.265, 0.368]	(0.00234, 0.00577]	(0.275, 0.342]
BCF	[0,00000, 0.00957]	[0,000, 0.438]	[0,00000, 0.01057]	[0,000, 0.745]
	(0.00957, 0.01488]	(0.438, 0.455]	(0.01057, 0.01911]	(0.745, 0.662]
	(0.01488, 0.02420]	(0.455, 0.570]	(0.01911, 0.04322]	(0.662, 0.691]
IGC50	[0,00000, 0.00397]	[0,000, 0.295]	[0,00000, 0.00366]	[0,000, 0.283]
	(0.00397, 0.00644]	(0.295, 0.351]	(0.00366, 0.00614]	(0.283, 0.285]
	(0.00644, 0.01099]	(0.351, 0.363]	(0.00614, 0.01303]	(0.285, 0.339]
LC50FM	[0,00000, 0.00689]	[0,000, 0.458]	[0,00000, 0.00734]	[0,000, 0.495]
	(0.00689, 0.01150]	(0.458, 0.500]	(0.00734, 0.01498]	(0.495, 0.650]
	(0.01150, 0.02091]	(0.500, 0.595]	(0.01498, 0.03465]	(0.650, 0.684]
LC50DM	[0,00000, 0.01234]	[0,000, 0.573]	[0,00000, 0.02166]	[0,000, 0.855]
	(0.01234, 0.01882]	(0.573, 0.512]	(0.02166, 0.03473]	(0.855, 0.920]
	(0.01882, 0.03202]	(0.512, 0.674]	(0.03473, 0.06622]	(0.920, 0.964]

Supplementary Table 7. The optimal uncertainty thresholds determined by maximum Yuden's index for various properties in classification model.

Dataset	DMPNN-Des		DMPNN	
	Uncertainty threshold	Max Youden's index	Uncertainty threshold	Max Youden's index
PAMPA	0.00116	0.244	0.00023	0.367
Pgp-inhibitor	0.00068	0.418	0.00049	0.518
Pgp-substrate	0.00059	0.343	0.00013	0.555
HIA	0.00050	0.602	0.00001	0.496
F20%	0.00075	0.485	0.00014	0.469
F30%	0.00069	0.412	0.00047	0.466
F50%	0.00150	0.296	0.00025	0.455
CYP1A2 inhibitor	0.00024	0.486	0.00060	0.556
CYP1A2 substrate	0.00091	0.380	0.00010	0.438
CYP2C19 inhibitor	0.00030	0.428	0.00058	0.473
CYP2C19 substrate	0.00089	0.288	0.00040	0.404
CYP2C9 inhibitor	0.00021	0.459	0.00031	0.461
CYP2C9 substrate	0.00567	0.344	0.00024	0.319
CYP2D6 inhibitor	0.00053	0.456	0.00057	0.525
CYP2D6 substrate	0.00263	0.282	0.00023	0.324
CYP3A4 inhibitor	0.00027	0.458	0.00044	0.505
CYP3A4 substrate	0.00033	0.291	0.00030	0.407
CYP2B6 inhibitor	0.00538	0.363	0.00023	0.400
CYP2B6 substrate	0.00018	0.461	0.00057	0.660
CYP2C8 inhibitor	0.00067	0.483	0.00031	0.552
HLM Stability	0.00176	0.403	0.00045	0.335
hERG Blockers	0.00027	0.521	0.00007	0.543
hERG Blocker (10um)	0.00156	0.517	0.00042	0.600
DILI	0.00008	0.669	0.00008	0.674
AMES Mutagenicity	0.00047	0.646	0.00011	0.640
ROA	0.00196	0.409	0.00024	0.316
FDAMDD	0.00025	0.662	0.00041	0.642
Skin Sensitization	0.00046	0.693	0.00019	0.712
Carcinogenicity	0.00063	0.367	0.00011	0.344
Eye Corrosion	0.00029	0.729	0.00019	0.696
Eye Irritation	0.00062	0.549	0.00011	0.575
Respiratory	0.00017	0.559	0.00005	0.470
Human Hepatotoxicity	0.00024	0.608	0.00006	0.612
Drug-induced Neurotoxicity	0.00029	0.460	0.00028	0.484
Ototoxicity	0.00104	0.322	0.00050	0.293
Hematotoxicity	0.00185	0.478	0.00049	0.430
Drug-induced Nephrotoxicity	0.00075	0.375	0.00036	0.366
Genotoxicity	0.00147	0.369	0.00033	0.338
RPMI-8226 Immunitoxicity	0.00044	0.327	0.00039	0.286
A549 Cytotoxicity	0.00359	0.328	0.00020	0.370
Hek293 Cytotoxicity	0.00080	0.206	0.00033	0.194
NR-AhR	0.00007	0.656	0.00016	0.691
NR-AR	0.00003	0.590	0.00029	0.653
NR-AR-LBD	0.00268	0.337	0.00031	0.396
NR-Aromatase	0.00255	0.262	0.00046	0.242
NR-ER	0.00402	0.288	0.00088	0.351
NR-ER-LBD	0.00200	0.300	0.00037	0.227
NR-PPAR-gamma	0.00115	0.376	0.00044	0.351
SR-ARE	0.00074	0.263	0.00034	0.225
SR-ATAD5	0.00005	0.538	0.00046	0.633
SR-HSE	0.00043	0.567	0.00056	0.589
SR-MMP	0.00080	0.542	0.00056	0.607
SR-p53	0.00135	0.417	0.00045	0.443

Supplementary Table 8. Runtime analysis in seconds for submissions of 1 to 1000 molecules for ADMETlab 2.0 and ADMETlab 3.0 with different modelling options.

Number of molecules	ADMETlab2.0	ADMETlab3.0 (web portal)	ADMETlab3.0 (DMPNN-Des)	ADMETlab3.0 (DMPNN-Des & Uncertainty)	ADMETlab3.0 (DMPNN)	ADMETlab3.0 (DMPNN & Uncertainty)
1	4.63	5.75	5.80	6.03	5.79	6.09
10	5.52	6.97	6.89	7.39	6.63	7.34
100	12.52	13.19	17.21	20.20	12.93	18.18
500	39.36	43.78	61.23	76.58	42.10	65.60
1000	84.00	87.23	116.59	150.54	77.42	130.66



Supplementary Figure 1. Runtime analysis in seconds for submissions of 1 to 1000 molecules for ADMETlab 2.0, ADMETlab 3.0 (DMPNN-Des), ADMETlab 3.0 (DMPNN-Des & Uncertainty), ADMETlab 3.0 (DMPNN), and ADMETlab 3.0 (DMPN & Uncertainty).

References

1. Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H., Guzman-Perez, A., Hopper, T., Kelley, B., Mathea, M. *et al.* (2019) Analyzing Learned Molecular Representations for Property Prediction. *J Chem Inf Model*, **59**, 3370-3388.
http://www.ncbi.nlm.nih.gov/pubmed/31361484
http://dx.doi.org/10.1021/acs.jcim.9b00237

2. Heid, E., Greenman, K.P., Chung, Y., Li, S.C., Graff, D.E., Vermeire, F.H., Wu, H., Green, W.H. and McGill, C.J. (2024) Chemprop: A Machine Learning Package for Chemical Property Prediction. *J Chem Inf Model*, **64**, 9-17.
http://www.ncbi.nlm.nih.gov/pubmed/38147829
http://dx.doi.org/10.1021/acs.jcim.3c01250

3. Begoli, E., Bhattacharya, T. and Kusnezov, D. (2019) The need for uncertainty quantification in machine-assisted medical decision making. *Nature Machine Intelligence*, **1**, 20-23.
http://dx.doi.org/10.1038/s42256-018-0004-1

4. Amini, A., Schwarting, W., Soleimany, A. and Rus, D. Deep Evidential Regression. *CoRR*, **abs/1910.02600**.

5. Gal, Y. and Ghahramani, Z. *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning*.

6. Wen, M.J. and Tadmor, E.B. (2020) Uncertainty quantification in molecular simulations with dropout neural network potentials. *Npj Comput Mater*, **6**.

7. Tong, X., Wang, D., Ding, X., Tan, X., Ren, Q., Chen, G., Rong, Y., Xu, T., Huang, J., Jiang, H. *et al.* (2022) Blood-brain barrier penetration prediction enhanced by uncertainty estimation. *J Cheminform*, **14**, 44.
http://www.ncbi.nlm.nih.gov/pubmed/35799215
http://dx.doi.org/10.1186/s13321-022-00619-2

8. Wang, S. and Manning, C. (2013) In Sanjoy, D. and David, M. (eds.), *Proceedings of the 30th International Conference on Machine Learning*. PMLR, Proceedings of Machine Learning Research, Vol. 28, pp. 118--126.

9. Dolezal, J.M., Srisuwananukorn, A., Karpeyev, D., Ramesh, S., Kochanny, S., Cody, B., Mansfield, A.S., Rakshit, S., Bansal, R., Bois, M.C. *et al.* (2022) Uncertainty-informed deep learning models enable high-confidence predictions for digital histopathology. *Nat Commun*, **13**, 6572.

http://www.ncbi.nlm.nih.gov/pubmed/36323656
http://dx.doi.org/10.1038/s41467-022-34025-x