Hyperparameters	Classification	Regression
Learning rate	0.0001~0.001	0.0001~0.001
Learning rate decay	0.98~0.995	0.98~0.995
Batch Size	8,16	16,32
Dropout ratio	0~0.5	0~0.5
Graph pooling	set2set, collection	set2set, mean
Max Epochs	100	100
Learning Rate Decay	Linear	Linear

Table S31 (a): Hyperparameter optimization for MPG.

Hyperparameters	Fine-tuning
Max learning rate	0.0001~0.001
Max Epochs	100
Number of final learning	2~5
Batch Size	32
Dropout ratio	0~0.2
Self-attention layer	4~8
Dist coefficient	0.05~0.15
Dropout bond	0~0.4
Number of layers in FFN after MPN encoding	3~5

Table S31 (b): Hyperparameter optimization for GROVER.

Hyperparameters	Fine-tuning
Batch Size	8, 16, 32
learning rate	0.00001~0.01

Table S31 (c): Hyperparameter optimization for X-MOL.

Hyperparameters	Fine-tuning
Batch Size	8, 16, 32
Initial learning rate	1e-6~0.005
Initial base learning rate	1e-6~0.001

Table S31 (d): Hyperparameter optimization for MolCLR.

Hyperparameters	Fine-tuning
Batch Size	8, 16, 32
Initial learning rate	0.000001~0.01
Max Epochs	100

Table S31 (e): Hyperparameter optimization for CHEM-BERT.

Hyperparameters	Linear regression
Feature extraction	Polynomial Features
Degree	1-10
Hyperparameters	**Logistic regression**
Penalty	L1, L2
Intercept C	1~10
Solver	liblinear
Maximum number of iterations	100~10000
Hyperparameters	**MLP**
Activation	RELU
Solver	Adam
Learning rate	0.0001~0.005
Batch size	8, 16, 32
Maximum number of iterations	500-10000

Hyperparameters	**Random forest**
Number of estimators	100~10000
Criterion	Gini, Entropy (classification) MAE, MSE (regression)

Table S31 (f): Hyperparameter optimization for SMILES_Transformer.

Hyperparameters	**Fine-tuning**
Batch Size	8, 16, 32, 64, 128
Initial learning rate	0.0005~0.1

Table S31 (g): Hyperparameter optimization for Chemception, ADMET-CNN and QSAR-CNN.