# Supplementary

Supplementary Table 1. Detail information of the datasets.

| Dataset | Compounds | Task type | Number of targets | Data split | Evaluation |
|---|---|---|---|---|---|
| Tox21 | 8014 | Classification | 12 | Random | ROC-AUC |
| SIDER | 1427 | Classification | 27 | Random | ROC-AUC |
| HIV | 41127 | Classification | 1 | Scaffold | ROC-AUC |
| BBBP | 2050 | Classification | 1 | Scaffold | ROC-AUC |
| ESOL | 1128 | Regression | 1 | Random | RMSE |
| FreeSolv | 642 | Regression | 1 | Random | RMSE |
| Lipophilicity | 4200 | Regression | 1 | Random | RMSE |
| Malaria | 9998 | Regression | 1 | Random | RMSE |
| CEP | 29978 | Regression | 1 | Random | RMSE |

Supplementary Table 2. The number of compounds for each target in the Tox21 dataset.

| Target | NR-AR | NR-AR-LBD | NR-Ahr | NR-Aromatase | NR-ER | NR-ER-LBD |
|---|---|---|---|---|---|---|
| # of data | 7439 | 6902 | 6691 | 5940 | 6316 | 7112 |
| # of active | 310 | 238 | 785 | 307 | 798 | 356 |
| % of active compound | 24.00 | 29.00 | 8.52 | 19.35 | 7.91 | 19.98 |
| Target | NR-PPAR-gamma | SR-ARE | SR-ATAD5 | SR-HSE | SR-MMP | SR-p53 |
| # of data | 6583 | 5935 | 7232 | 6594 | 5920 | 6909 |
| # of active | 189 | 961 | 267 | 379 | 936 | 431 |
| % of active compound | 34.83 | 6.18 | 27.09 | 17.40 | 6.32 | 16.03 |

Supplementary Table 3. Hyperparameters of QSAR model for each task.

| Task | Type of task | Embedding size | # of feature maps | Learning rate | Window size of filter |
|---|---|---|---|---|---|
| **Tox21-single** | Classification | 200 | 512 | 1e-4 | 5 |
| **Tox21-multi** | Classification | 200 | 1024 | 1e-4 | 5 |
| **SIDER-single** | Classification | 100 | 512 | 1e-3 | 5 |
| **SIDER-multi** | Classification | 100 | 512 | 1e-4 | 5 |
| **HIV** | Classification | 200 | 1024 | 1e-5 | 5 |
| **BBBP** | Classification | 200 | 1024 | 1e-5 | 5 |
| **ESOL** | Regression | 100 | 2048 | 5e-4 | 1 |
| **FreeSolv** | Regression | 200 | 2048 | 1e-4 | 1 |
| **Lipophilicity** | Regression | 100 | 2048 | 5e-4 | 1 |
| **Malaria** | Regression | 100 | 2048 | 1e-4 | 1 |
| **CEP** | Regression | 100 | 2048 | 5e-4 | 1 |

Supplementary Table 4. The ROC-AUC scores of our model in each trial on the classification tasks (Top: validation set, bottom: test set).

| Validation set | Tox21 | | SIDER | | HIV | BBBP |
|---|---|---|---|---|---|---|
| | Multi-task | Single task | Multi-task | Single task | | |
| Trial 1 | 0.878 | 0.839 | 0.864 | 0.746 | 0.791 | 0.965 |
| Trial 2 | 0.873 | 0.821 | 0.833 | 0.723 | 0.812 | 0.960 |
| Trial 3 | 0.887 | 0.827 | 0.865 | 0.729 | 0.791 | 0.956 |
| Trial 4 | 0.878 | 0.851 | 0.850 | 0.732 | 0.805 | 0.957 |
| Trial 5 | 0.875 | 0.843 | 0.831 | 0.725 | 0.800 | 0.964 |
| Average | **0.878** | **0.836** | **0.848** | **0.731** | **0.800** | **0.960** |
| STDEV | 0.005 | 0.012 | 0.016 | 0.009 | 0.009 | 0.004 |

| Test set | Tox21 | | SIDER | | HIV | BBBP |
|---|---|---|---|---|---|---|
| | Multi-task | Single task | Multi-task | Single task | | |
| Trial 1 | 0.874 | 0.841 | 0.845 | 0.732 | 0.750 | 0.707 |
| Trial 2 | 0.870 | 0.818 | 0.845 | 0.734 | 0.756 | 0.719 |
| Trial 3 | 0.874 | 0.812 | 0.834 | 0.734 | 0.746 | 0.711 |
| Trial 4 | 0.876 | 0.828 | 0.826 | 0.742 | 0.757 | 0.720 |
| Trial 5 | 0.886 | 0.826 | 0.829 | 0.717 | 0.759 | 0.709 |
| Average | **0.876** | **0.825** | **0.836** | **0.732** | **0.754** | **0.713** |
| STDEV | 0.006 | 0.011 | 0.009 | 0.009 | 0.005 | 0.006 |

Supplementary Table 5. The ROC-AUC score on the Tox21 dataset against each target.

| Target | | NR-AR | NR-AR-LBD | NR-Ahr | NR-Aromatase | NR-ER | NR-ER-LBD |
|---|---|---|---|---|---|---|---|
| Multi-task | Average | 0.851 | 0.952 | 0.877 | 0.832 | 0.757 | 0.917 |
| | STDEV | 0.043 | 0.041 | 0.018 | 0.060 | 0.024 | 0.019 |
| Single task | Average | 0.810 | 0.898 | 0.896 | 0.847 | 0.725 | 0.812 |
| | STDEV | 0.027 | 0.040 | 0.014 | 0.042 | 0.037 | 0.025 |

| Target | | NR-PPAR-gamma | SR-ARE | SR-ATAD5 | SR-HSE | SR-MMP | SR-p53 |
|---|---|---|---|---|---|---|---|
| **Multi-task** | Average | 0.933 | 0.840 | 0.905 | 0.826 | 0.913 | 0.911 |
| | STDEV | 0.011 | 0.009 | 0.038 | 0.040 | 0.012 | 0.013 |
| **Single task** | Average | 0.788 | 0.803 | 0.844 | 0.765 | 0.873 | 0.839 |
| | STD EV | 0.081 | 0.033 | 0.033 | 0.028 | 0.013 | 0.019 |

Supplementary Table 6. The ROC-AUC scores on the SIDER dataset against each target*.

| | | Target 1 | Target 2 | Target 3 | Target 4 | Target 5 | Target 6 | Target 7 | Target 8 | Target 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Multi-task** | Average | 0.873 | 0.848 | 0.463 | 0.843 | 0.881 | 0.862 | 0.953 | 0.870 | 0.801 |
| | STDEV | 0.029 | 0.026 | 0.121 | 0.023 | 0.048 | 0.013 | 0.018 | 0.042 | 0.040 |
| **Single task** | Average | 0.782 | 0.715 | 0.449 | 0.765 | 0.816 | 0.744 | 0.834 | 0.677 | 0.670 |
| | STDEV | 0.008 | 0.045 | 0.161 | 0.013 | 0.072 | 0.068 | 0.055 | 0.039 | 0.038 |
| | | Target 10 | Target 11 | Target 12 | Target 13 | Target 14 | Target 15 | Target 16 | Target 17 | Target 18 |
| **Multi-task** | Average | 0.853 | 0.814 | 0.853 | 0.818 | 0.756 | 0.886 | 0.866 | 0.876 | 0.755 |
| | STDEV | 0.012 | 0.050 | 0.042 | 0.066 | 0.035 | 0.022 | 0.011 | 0.056 | 0.013 |
| **Single task** | Average | 0.809 | 0.714 | 0.784 | 0.684 | 0.622 | 0.743 | 0.789 | 0.708 | 0.642 |
| | STDEV | 0.034 | 0.058 | 0.084 | 0.030 | 0.052 | 0.044 | 0.037 | 0.057 | 0.076 |
| | | Target 19 | Target 20 | Target 21 | Target 22 | Target 23 | Target 24 | Target 25 | Target 26 | Target 27 |
| **Multi-task** | Average | 0.819 | 0.867 | 0.882 | 0.875 | 0.780 | 0.867 | 0.879 | 0.901 | 0.828 |
| | STDEV | 0.041 | 0.021 | 0.026 | 0.016 | 0.067 | 0.018 | 0.009 | 0.049 | 0.025 |
| **Single task** | average | 0.749 | 0.735 | 0.765 | 0.787 | 0.673 | 0.748 | 0.819 | 0.820 | 0.716 |
| | STDEV | 0.079 | 0.067 | 0.046 | 0.059 | 0.059 | 0.023 | 0.044 | 0.083 | 0.021 |

(* Target 1 Hepatobiliary disorders; Target 2 Metabolism and nutrition disorders; Target 3 Product issues; Target 4 Eye disorders; Target 5 Investigations; Target 6 Musculoskeletal and connective tissue disorders; Target 7 Gastrointestinal disorders; Target 8 Social circumstances; Target 9 Immune system disorders; Target 10 Reproductive system and breast disorders; Target 11 Neoplasms benign, malignant and unspecified (incl cysts and polyps); Target 12 General disorders and administration site conditions; Target 13 Endocrine disorders; Target 14 Surgical and medical procedures; Target 15 Vascular disorders; Target 16 Blood and lymphatic system disorders; Target 17 Skin and subcutaneous tissue disorders; Target 18 Congenital, familial and genetic disorders; Target 19 Infections and infestations; Target 20 Respiratory, thoracic and mediastinal disorders; Target 21 Psychiatric disorders; Target 22 Renal and urinary disorders; Target 23 Pregnancy, puerperium and perinatal conditions; Target 24 Ear and labyrinth disorders; Target 25 Cardiac disorders; Target 26 Nervous system disorders; Target 27 Injury, poisoning and procedural complications)

Supplementary Table 7. The RMSE scores of our model in each trial on the regression tasks. (Top: validation set, bottom: test set)

| Validation set | ESOL | FreeSolv | Lipophilicity | Malaria | CEP |
|---|---|---|---|---|---|
| **Trial 1** | 1.00 | 2.65 | 0.86 | 1.04 | 1.23 |
| **Trial 2** | 1.09 | 1.78 | 0.81 | 1.02 | 1.27 |
| **Trial 3** | 1.12 | 2.28 | 0.86 | 1.06 | 1.31 |

| | | | | | |
|---|---|---|---|---|---|
| **Trial 4** | 1.10 | 1.93 | 0.83 | 0.96 | 1.30 |
| **Trial 5** | 1.12 | 2.00 | 0.83 | 1.04 | 1.28 |
| **Average** | **1.09** | **2.13** | **0.84** | **1.02** | **1.28** |
| **STDEV** | 0.05 | 0.34 | 0.02 | 0.04 | 0.03 |

| Test set | ESOL | FreeSolv | Lipophilicity | Malaria | CEP |
|---|---|---|---|---|---|
| **Trial 1** | 1.15 | 1.33 | 0.86 | 1.00 | 1.30 |
| **Trial 2** | 1.00 | 1.67 | 0.81 | 1.01 | 1.35 |
| **Trial 3** | 0.92 | 1.88 | 0.84 | 0.99 | 1.31 |
| **Trial 4** | 1.13 | 1.44 | 0.83 | 1.04 | 1.31 |
| **Trial 5** | 1.07 | 1.48 | 0.85 | 1.01 | 1.40 |
| **Average** | **1.06** | **1.56** | **0.84** | **1.01** | **1.34** |
| **STDEV** | 0.10 | 0.22 | 0.02 | 0.02 | 0.04 |

Supplementary Table 8. The ROC-AUC scores of our QSAR model using the circular fingerprint featurizer on Tox21 dataset. (Top: Validation set, bottom: test set)

| Validation set | Window size 5 | Window size 11 |
|---|---|---|
| **Trial 1** | 0.562 | 0.622 |
| **Trial 2** | 0.618 | 0.618 |
| **Trial 3** | 0.620 | 0.629 |

| | | |
|---|---|---|
| Trial 4 | 0.602 | 0.609 |
| Trial 5 | 0.629 | 0.637 |
| Average | **0.606** | **0.623** |
| STDEV | 0.027 | 0.011 |

| Test set | Window size 5 | Window size 11 |
|---|---|---|
| Trial 1 | 0.543 | 0.612 |
| Trial 2 | 0.587 | 0.633 |
| Trial 3 | 0.587 | 0.626 |
| Trial 4 | 0.620 | 0.629 |
| Trial 5 | 0.591 | 0.622 |
| Average | **0.586** | **0.624** |
| STDEV | 0.027 | 0.008 |