

Supporting Information

Hierarchical H-QSAR Modeling Approach for Integrating Binary, Multiclass, and Regression Models of Acute Oral Systemic Toxicity

Xinhao Li¹, Nicole C. Kleinstreuer ^{2,3}, and Denis Fourches^{1,*}

1 Department of Chemistry, Bioinformatics Research Center, North Carolina State University, Raleigh, NC 27695, United States.

2 Division of Intramural Research/Biostatistics and Computational Biology Branch, NIEHS, RTP, North Carolina 27709, United States.

3 National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, NIEHS, RTP, North Carolina 27709, United States.

* Whom correspondence should be sent. Email: dfourch@ncsu.edu

Table S1. Hyperparameter Search Spaces.

Algorithm	Parameters Space
kNN	<p>RDkit and Mordred: <code>{'n_neighbors': [5,9,15,19,25,35,45,55,71], 'weights': ['distance'],'p': [1,2]}</code></p> <p>ECFP6_Bits and MACCS: <code>{'n_neighbors': [5,9,15,19,25,35,45,55,71], 'weights': ['distance'], 'metric': ['jaccard', 'dice', 'rogerstanimoto']}</code></p> <p>ECFP6_Counts: <code>{'n_neighbors': [5,9,15,19,25,35,45,55,71], 'weights': ['distance'], 'metric': ['hamming', 'canberra', 'braycurtis']}</code></p>
SVM	<code>{'C': [0.01, 0.1, 1, 10, 100, 200, 400, 1000], 'kernel': ['linear']}</code> , <code>{'C': [0.01, 0.1, 1, 10, 100, 200, 400, 1000], 'gamma': [100,10,1,1e-1,1e-2, 1e-3], 'kernel': ['rbf']}</code>
RF	<code>{'bootstrap': [True, False], 'max_depth': [5, 20, 35, 50, 65, 80, None], 'max_features': ['log2', 'sqrt'], 'min_samples_leaf': [2, 4, 6], 'min_samples_split': [2, 5, 10], 'n_estimators': [500, 1500]}</code>
XGBoost	<code>{'learning_rate': [0.01, 0.1], 'max_depth': [3, 6, 10], 'min_child_weight': [1, 3, 5], 'gamma': [0, 1, 5], 'subsample': [0.6, 0.7, 0.8, 0.9, 1.0], 'colsample_bytree': [0.5, 0.6, 0.7, 0.8, 0.9, 1.0], 'n_estimators': [500, 1500]}</code>

Table S2. Hyperparameter Tuning of Regression Models

Algorithm	Feature	Best Parameter Setting
kNN	ECFP6_Bits	{'metric': 'dice', 'n_neighbors': 15, 'weights': 'distance'}
	ECFP6_Counts	{'metric': 'braycurtis', 'n_neighbors': 15, 'weights': 'distance'}
	MACCS	{'metric': 'rogerstanimoto', 'n_neighbors': 9, 'weights': 'distance'}
	RDKit	{'n_neighbors': 9, 'p': 1, 'weights': 'distance'}
	Mordred	{'n_neighbors': 9, 'p': 1, 'weights': 'distance'}
	H-Feature	{'n_neighbors': 45, 'p': 2, 'weights': 'distance'}
SVM	ECFP6_Bits	{'C': 1, 'gamma': 0.01, 'kernel': 'rbf'}
	ECFP6_Counts	{'C': 10, 'gamma': 0.0, 'kernel': 'rbf1'}
	MACCS	{'C': 10, 'gamma': 0.1, 'kernel': 'rbf'}
	RDKit	{'C': 1, 'gamma': 1, 'kernel': 'rbf'}
	Mordred	{'C': 10, 'gamma': 0.1, 'kernel': 'rbf'}
	H-Feature	{'C': 1, 'gamma': 0.01, 'kernel': 'rbf'}
RF	ECFP6_Bits	{'n_estimators': 1500, 'min_samples_split': 2, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': None, 'bootstrap': False}
	ECFP6_Counts	{'n_estimators': 500, 'min_samples_split': 2, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': None, 'bootstrap': False}
	MACCS	{'n_estimators': 500, 'min_samples_split': 5, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': 80, 'bootstrap': False}
	RDKit	{'n_estimators': 1500, 'min_samples_split': 2, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': 65, 'bootstrap': False}
	Mordred	{'n_estimators': 1500, 'min_samples_split': 10, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': 65, 'bootstrap': False}
	H-Feature	{'n_estimators': 1500, 'min_samples_split': 2, 'min_samples_leaf': 4, 'max_features': 'sqrt', 'max_depth': 20, 'bootstrap': False}
XGBoost	ECFP6_Bits	{'subsample': 0.9, 'n_estimators': 1500, 'min_child_weight': 3, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.7}
	ECFP6_Counts	{'subsample': 1.0, 'n_estimators': 500, 'min_child_weight': 5, 'max_depth': 10, 'learning_rate': 0.1, 'gamma': 0, 'colsample_bytree': 0.6}
	MACCS	{'subsample': 0.6, 'n_estimators': 1500, 'min_child_weight': 1, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.7}
	RDKit	{'subsample': 1.0, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.5}
	Mordred	{'subsample': 0.9, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.8};
	H-Feature	{'subsample': 0.6, 'n_estimators': 500, 'min_child_weight': 3, 'max_depth': 6, 'learning_rate': 0.01, 'gamma': 1, 'colsample_bytree': 0.6}

Table S3. Hyperparameter Tuning of Binary Models

Algorithm	Feature	Best Parameter Setting
kNN	ECFP6_Bits	{'metric': 'dice', 'n_neighbors': 19, 'weights': 'distance'}
	ECFP6_Counts	{'metric': 'braycurtis', 'n_neighbors': 19, 'weights': 'distance'}
	MACCS	{'metric': 'rogerstanimoto', 'n_neighbors': 23, 'weights': 'distance'}
	RDKit	{'n_neighbors': 15, 'p': 1, 'weights': 'distance'}
	Mordred	{'n_neighbors': 15, 'p': 1, 'weights': 'distance'}
	H-Feature	{'n_neighbors': 71, 'p': 1, 'weights': 'distance'}
SVM	ECFP6_Bits	{'C': 1, 'gamma': 0.1, 'kernel': 'rbf'}
	ECFP6_Counts	{'C': 1, 'gamma': 0.01, 'kernel': 'rbf'}
	MACCS	{'C': 1, 'gamma': 0.1, 'kernel': 'rbf'}
	RDKit	{'C': 1, 'gamma': 1, 'kernel': 'rbf'}
	Mordred	{'C': 10, 'gamma': 0.1, 'kernel': 'rbf'}
	H-Feature	{'C': 10, 'gamma': 0.001, 'kernel': 'rbf'}
RF	ECFP6_Bits	{'n_estimators': 1500, 'min_samples_split': 2, 'min_samples_leaf': 4, 'max_features': 'sqrt', 'max_depth': 80, 'bootstrap': False}
	ECFP6_Counts	{'n_estimators': 1500, 'min_samples_split': 5, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': 35, 'bootstrap': True}
	MACCS	{'n_estimators': 1500, 'min_samples_split': 2, 'min_samples_leaf': 2, 'max_features': 'log2', 'max_depth': 35, 'bootstrap': False}
	RDKit	{'n_estimators': 1500, 'min_samples_split': 2, 'min_samples_leaf': 2, 'max_features': 'log2', 'max_depth': 65, 'bootstrap': False}
	Mordred	{'n_estimators': 1500, 'min_samples_split': 5, 'min_samples_leaf': 4, 'max_features': 'sqrt', 'max_depth': 80, 'bootstrap': False}
	H-Feature	{'n_estimators': 1500, 'min_samples_split': 5, 'min_samples_leaf': 6, 'max_features': 'log2', 'max_depth': 35, 'bootstrap': False}
XGBoost	ECFP6_Bits	{'subsample': 0.6, 'n_estimators': 1500, 'min_child_weight': 5, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 1, 'colsample_bytree': 0.9}
	ECFP6_Counts	{'subsample': 0.6, 'n_estimators': 1500, 'min_child_weight': 1, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.7}
	MACCS	{'subsample': 0.6, 'n_estimators': 500, 'min_child_weight': 5, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 1, 'colsample_bytree': 0.8}
	RDKit	{'subsample': 0.8, 'n_estimators': 1500, 'min_child_weight': 1, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.6}
	Mordred	{'subsample': 1.0, 'n_estimators': 1500, 'min_child_weight': 3, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.7}
	H-Feature	{'subsample': 0.6, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 3, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.5}

Table S4. Hyperparameter Tuning of Multiclass Models

Algorithm	Feature	Best Parameter Setting
kNN	ECFP6_Bits	{'metric': 'dice', 'n_neighbors': 9, 'weights': 'distance'}
	ECFP6_Counts	{'metric': 'braycurtis', 'n_neighbors': 9, 'weights': 'distance'}
	MACCS	{'metric': 'rogerstanimoto', 'n_neighbors': 15, 'weights': 'distance'}
	RDKit	{'n_neighbors': 9, 'p': 1, 'weights': 'distance'}
	Mordred	{'n_neighbors': 9, 'p': 1, 'weights': 'distance'}
	H-Feature	{'n_neighbors': 35, 'p': 1, 'weights': 'distance'}
SVM	ECFP6_Bits	{'C': 10, 'gamma': 0.01, 'kernel': 'rbf'}
	ECFP6_Counts	{'C': 10, 'gamma': 0.01, 'kernel': 'rbf'}
	MACCS	{'C': 10, 'gamma': 0.1, 'kernel': 'rbf'}
	RDKit	{'C': 10, 'gamma': 1, 'kernel': 'rbf'}
	Mordred	{'C': 10, 'gamma': 1, 'kernel': 'rbf'}
	H-Feature	{'C': 0.1, 'kernel': 'linear'}
RF	ECFP6_Bits	{'n_estimators': 1500, 'min_samples_split': 2, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': 80, 'bootstrap': False}
	ECFP6_Counts	{'n_estimators': 1500, 'min_samples_split': 2, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': 65, 'bootstrap': False}
	MACCS	{'n_estimators': 500, 'min_samples_split': 2, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': 35, 'bootstrap': False}
	RDKit	{'n_estimators': 500, 'min_samples_split': 5, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': 80, 'bootstrap': False}
	Mordred	{'n_estimators': 500, 'min_samples_split': 5, 'min_samples_leaf': 2, 'max_features': 'sqrt', 'max_depth': None, 'bootstrap': False}
	H-Feature	{'n_estimators': 500, 'min_samples_split': 10, 'min_samples_leaf': 6, 'max_features': 'sqrt', 'max_depth': None, 'bootstrap': False}
XGBoost	ECFP6_Bits	{'subsample': 0.7, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 10, 'learning_rate': 0.1, 'gamma': 1, 'colsample_bytree': 0.9}
	ECFP6_Counts	{'subsample': 0.8, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 10, 'learning_rate': 0.1, 'gamma': 1, 'colsample_bytree': 1.0}
	MACCS	{'subsample': 0.6, 'n_estimators': 1500, 'min_child_weight': 3, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.5}
	RDKit	{'subsample': 0.6, 'n_estimators': 1500, 'min_child_weight': 1, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.8}
	Mordred	{'subsample': 0.8, 'n_estimators': 1500, 'min_child_weight': 3, 'max_depth': 6, 'learning_rate': 0.1, 'gamma': 5, 'colsample_bytree': 0.8}
	H-Feature	{'subsample': 0.6, 'n_estimators': 1500, 'min_child_weight': 3, 'max_depth': 10, 'learning_rate': 0.01, 'gamma': 5, 'colsample_bytree': 0.7}

Table S5. 10-Fold Cross-Validation Performances of Regression Models

Algorithm	Feature	RMSE	R ²	MAE	MSE
Base Models					
kNN	ECFP6_Bits	0.646±0.017	0.487±0.026	0.477±0.014	0.418±0.022
	ECFP6_Counts	0.647±0.017	0.486±0.027	0.477±0.012	0.419±0.023
	MACCS	0.633±0.028	0.508±0.043	0.462±0.016	0.402±0.035
	RDKit	0.588±0.026	0.576±0.032	0.43±0.019	0.346±0.031
	Mordred	0.609±0.023	0.545±0.031	0.443±0.015	0.371±0.028
SVM	ECFP6_Bits	0.668±0.02	0.454±0.024	0.49±0.016	0.446±0.026
	ECFP6_Counts	0.623±0.013	0.524±0.02	0.465±0.009	0.388±0.016
	MACCS	0.61±0.026	0.543±0.031	0.452±0.019	0.373±0.031
	RDKit	0.628±0.025	0.516±0.036	0.458±0.017	0.395±0.031
	Mordred	0.613±0.024	0.54±0.026	0.447±0.015	0.376±0.03
RF	ECFP6_Bits	0.628±0.022	0.517±0.026	0.466±0.016	0.394±0.028
	ECFP6_Counts	0.619±0.02	0.531±0.023	0.459±0.015	0.383±0.024
	MACCS	0.597±0.024	0.562±0.032	0.44±0.017	0.357±0.029
	RDKit	0.637±0.025	0.503±0.031	0.461±0.015	0.406±0.032
	Mordred	0.597±0.027	0.562±0.032	0.44±0.018	0.357±0.033
XGBoost	ECFP6_Bits	0.62±0.025	0.529±0.03	0.46±0.017	0.385±0.031
	ECFP6_Counts	0.621±0.019	0.527±0.021	0.461±0.014	0.386±0.024
	MACCS	0.595±0.027	0.565±0.029	0.435±0.018	0.355±0.032
	RDKit	0.594±0.025	0.567±0.028	0.435±0.018	0.353±0.03
	Mordred	0.584±0.028	0.582±0.033	0.428±0.018	0.341±0.033
Hierarchical Models					
kNN	H-Features	0.559±0.024	0.617±0.028	0.405±0.018	0.312±0.027
SVM	H-Features	0.551±0.024	0.628±0.026	0.395±0.018	0.304±0.026
RF	H-Features	0.549±0.025	0.631±0.028	0.396±0.019	0.301±0.027
XGBoost	H-Features	0.550±0.024	0.628±0.027	0.398±0.018	0.303±0.027

Table S6. 10-Fold Cross-Validation Performances of Binary Models

Algorithm	Feature	Accuracy	Balance Accuracy	F1-Score	MCC	AUROC
Base Models						
kNN	ECFP6 Bits	0.746±0.014	0.737±0.013	0.745±0.014	0.479±0.028	0.737±0.013
	ECFP6 Counts	0.749±0.012	0.74±0.011	0.748±0.012	0.486±0.023	0.74±0.011
	MACCS	0.758±0.011	0.747±0.01	0.756±0.011	0.503±0.022	0.747±0.01
	RDKit	0.763±0.013	0.754±0.013	0.762±0.013	0.514±0.027	0.754±0.013
	Mordred	0.77±0.016	0.763±0.016	0.769±0.016	0.529±0.032	0.763±0.016
SVM	ECFP6 Bits	0.7±0.009	0.665±0.007	0.675±0.01	0.393±0.016	0.665±0.007
	ECFP6 Counts	0.764±0.013	0.753±0.012	0.762±0.012	0.515±0.027	0.753±0.012
	MACCS	0.782±0.014	0.772±0.014	0.78±0.014	0.552±0.029	0.772±0.014
	RDKit	0.766±0.015	0.755±0.015	0.764±0.015	0.519±0.031	0.755±0.015
	Mordred	0.774±0.012	0.767±0.011	0.773±0.012	0.537±0.023	0.767±0.011
RF	ECFP6 Bits	0.758±0.011	0.743±0.012	0.754±0.011	0.502±0.025	0.743±0.012
	ECFP6 Counts	0.758±0.014	0.74±0.014	0.753±0.014	0.505±0.03	0.74±0.014
	MACCS	0.784±0.013	0.774±0.013	0.782±0.013	0.556±0.028	0.774±0.013
	RDKit	0.786±0.011	0.776±0.009	0.784±0.011	0.561±0.022	0.776±0.009
	Mordred	0.789±0.014	0.778±0.013	0.787±0.014	0.567±0.027	0.778±0.013
XGBoost	ECFP6 Bits	0.751±0.017	0.74±0.017	0.749±0.017	0.488±0.034	0.74±0.017
	ECFP6 Counts	0.766±0.013	0.756±0.014	0.764±0.013	0.519±0.028	0.756±0.014
	MACCS	0.772±0.015	0.762±0.015	0.77±0.015	0.532±0.03	0.762±0.015
	RDKit	0.788±0.01	0.78±0.009	0.787±0.01	0.566±0.02	0.78±0.009
	Mordred	0.792±0.014	0.783±0.013	0.791±0.014	0.573±0.028	0.783±0.013
Hierarchical Models						
kNN	H-Features	0.796±0.017	0.79±0.017	0.796±0.017	0.584±0.034	0.79±0.017
SVM	H-Features	0.799±0.015	0.792±0.015	0.798±0.015	0.588±0.03	0.792±0.015
RF	H-Features	0.805±0.015	0.799±0.014	0.804±0.015	0.601±0.029	0.799±0.014
XGBoost	H-Features	0.802±0.011	0.795±0.01	0.801±0.011	0.594±0.021	0.795±0.01

Table S7. 10-Fold Cross-Validation Performances of Multiclass Models

Algorithm	Feature	Accuracy	Balance Accuracy	MCC	F1-Score
Base Models					
kNN	ECFP6_Bits	0.605±0.022	0.537±0.023	0.373±0.034	0.594±0.023
	ECFP6_Counts	0.607±0.014	0.542±0.017	0.375±0.022	0.596±0.015
	MACCS	0.605±0.013	0.529±0.014	0.37±0.024	0.592±0.015
	RDKit	0.61±0.018	0.542±0.014	0.382±0.022	0.6±0.019
	Mordred	0.614±0.009	0.547±0.011	0.388±0.015	0.604±0.011
SVM	ECFP6_Bits	0.589±0.019	0.522±0.019	0.348±0.029	0.579±0.02
	ECFP6_Counts	0.606±0.018	0.542±0.017	0.374±0.027	0.597±0.018
	MACCS	0.604±0.01	0.545±0.013	0.377±0.012	0.598±0.011
	RDKit	0.609±0.016	0.548±0.015	0.384±0.019	0.602±0.018
	Mordred	0.606±0.01	0.545±0.016	0.376±0.021	0.598±0.012
RF	ECFP6_Bits	0.615±0.018	0.494±0.015	0.374±0.02	0.577±0.02
	ECFP6_Counts	0.619±0.015	0.494±0.012	0.38±0.017	0.58±0.018
	MACCS	0.627±0.015	0.531±0.018	0.398±0.029	0.608±0.017
	RDKit	0.636±0.011	0.535±0.008	0.412±0.012	0.616±0.012
	Mordred	0.638±0.011	0.541±0.013	0.416±0.022	0.619±0.012
XGBoost	ECFP6_Bits	0.603±0.019	0.521±0.02	0.361±0.025	0.587±0.02
	ECFP6_Counts	0.617±0.019	0.535±0.02	0.384±0.029	0.601±0.019
	MACCS	0.629±0.011	0.547±0.01	0.407±0.015	0.616±0.012
	RDKit	0.633±0.013	0.538±0.011	0.409±0.014	0.616±0.014
	Mordred	0.618±0.016	0.522±0.017	0.383±0.024	0.599±0.018
Hierarchical Models					
kNN	H-Features	0.654±0.018	0.576±0.013	0.447±0.023	0.64±0.02
SVM	H-Features	0.655±0.015	0.569±0.013	0.448±0.018	0.639±0.017
RF	H-Features	0.659±0.016	0.587±0.015	0.458±0.023	0.648±0.017
XGBoost	H-Features	0.658±0.015	0.588±0.012	0.456±0.023	0.647±0.017

Table S8. Test Set Performances of Regression Models

Algorithm	Feature	RMSE	R ²	MAE	MSE
Base Models					
kNN	ECFP6_Bits	0.616	0.528	0.444	0.380
	ECFP6_Counts	0.620	0.523	0.449	0.384
	MACCS	0.598	0.555	0.427	0.358
	RDKit	0.596	0.559	0.428	0.355
	Mordred	0.590	0.566	0.422	0.349
SVM	ECFP6_Bits	0.627	0.512	0.458	0.393
	ECFP6_Counts	0.598	0.556	0.440	0.357
	MACCS	0.576	0.589	0.418	0.331
	RDKit	0.616	0.529	0.4452	0.379
	Mordred	0.602	0.550	0.430	0.362
RF	ECFP6_Bits	0.604	0.548	0.445	0.365
	ECFP6_Counts	0.600	0.553	0.440	0.360
	MACCS	0.580	0.583	0.424	0.336
	RDKit	0.561	0.610	0.404	0.315
	Mordred	0.575	0.590	0.417	0.331
XGBoost	ECFP6_Bits	0.590	0.568	0.437	0.348
	ECFP6_Counts	0.588	0.571	0.430	0.346
	MACCS	0.566	0.602	0.404	0.321
	RDKit	0.563	0.606	0.412	0.318
	Mordred	0.567	0.600	0.411	0.322
Base Model Consensus		0.551	0.623	0.398	0.304
Hierarchical Models					
kNN	H-Features	0.534	0.647	0.381	0.285
SVM	H-Features	0.529	0.652	0.376	0.280
RF	H-Features	0.529	0.652	0.376	0.280
XGBoost	H-Features	0.531	0.649	0.377	0.283
Hierarchical Model Consensus		0.526	0.656	0.374	0.277

Table S9. Test Set Performances of Binary Models

Algorithm	Feature	Accuracy	Balance Accuracy	F1-Score	MCC	AUROC
Base Models						
kNN	ECFP6_Bits	0.748	0.735	0.745	0.478	0.735
	ECFP6_Counts	0.756	0.743	0.754	0.495	0.743
	MACCS	0.768	0.754	0.765	0.520	0.754
	RDKit	0.771	0.759	0.769	0.527	0.759
	Mordred	0.770	0.760	0.768	0.525	0.760
SVM	ECFP6_Bits	0.734	0.709	0.724	0.449	0.709
	ECFP6_Counts	0.757	0.745	0.755	0.498	0.745
	MACCS	0.780	0.767	0.777	0.545	0.767
	RDKit	0.763	0.751	0.761	0.510	0.751
	Mordred	0.775	0.765	0.774	0.536	0.765
RF	ECFP6_Bits	0.763	0.745	0.758	0.510	0.745
	ECFP6_Counts	0.756	0.735	0.750	0.497	0.735
	MACCS	0.783	0.770	0.780	0.551	0.770
	RDKit	0.779	0.766	0.777	0.544	0.766
	Mordred	0.782	0.770	0.780	0.550	0.770
XGBoost	ECFP6_Bits	0.746	0.733	0.744	0.475	0.733
	ECFP6_Counts	0.766	0.753	0.764	0.517	0.753
	MACCS	0.772	0.758	0.769	0.528	0.758
	RDKit	0.782	0.772	0.781	0.550	0.772
	Mordred	0.787	0.777	0.785	0.560	0.777
Base Model Consensus		0.790	0.777	0.788	0.567	0.777
Hierarchical Models						
kNN	H-Features	0.799	0.790	0.798	0.586	0.790
SVM	H-Features	0.796	0.788	0.795	0.581	0.788
RF	H-Features	0.801	0.793	0.800	0.590	0.793
XGBoost	H-Features	0.801	0.793	0.800	0.591	0.793
Hierarchical Consensus		0.800	0.792	0.799	0.589	0.792

Table S10. Test Set Performances of Multiclass Models

Algorithm	Feature	Accuracy	Balance Accuracy	F1-Score	MCC
Base Models					
kNN	ECFP6 Bits	0.625	0.555	0.614	0.407
	ECFP6 Counts	0.636	0.572	0.626	0.426
	MACCS	0.627	0.563	0.616	0.411
	RDKit	0.633	0.572	0.624	0.422
	Mordred	0.636	0.578	0.627	0.427
SVM	ECFP6 Bits	0.624	0.522	0.598	0.394
	ECFP6 Counts	0.634	0.545	0.614	0.413
	MACCS	0.631	0.542	0.614	0.411
	RDKit	0.633	0.548	0.617	0.413
	Mordred	0.635	0.548	0.618	0.416
RF	ECFP6 Bits	0.633	0.516	0.602	0.411
	ECFP6 Counts	0.641	0.524	0.610	0.426
	MACCS	0.652	0.566	0.637	0.447
	RDKit	0.659	0.569	0.642	0.456
	Mordred	0.663	0.575	0.648	0.464
XGBoost	ECFP6 Bits	0.631	0.550	0.618	0.412
	ECFP6 Counts	0.639	0.555	0.624	0.425
	MACCS	0.649	0.574	0.638	0.445
	RDKit	0.653	0.567	0.638	0.448
	Mordred	0.635	0.550	0.619	0.417
Base Model Consensus		0.669	0.578	0.651	0.475
Hierarchical Models					
kNN	H-Features	0.678	0.612	0.668	0.493
SVM	H-Features	0.678	0.611	0.668	0.492
RF	H-Features	0.678	0.615	0.667	0.492
XGBoost	H-Features	0.671	0.615	0.663	0.483
Hierarchical Model Consensus		0.681	0.618	0.670	0.497

Table S11. Performances of Hierarchical Consensus Models based on Prediction Zones.

Models	Metric	OutZone	InZone	HalfZone_B	HalfZone_M	Overall
Regression	RMSE	0.490	0.622	0.804	0.588	0.526
	R2	0.714	0.464	0.183	0.315	0.656
	MAR	0.349	0.558	0.575	0.375	0.374
	MSE	0.240	0.387	0.646	0.238	0.277
Binary	Accuracy	0.839	0.571	0.819	0.528	0.800
	BA	0.831	0.583	0.775	0.527	0.792
	F1-Score	0.838	0.571	0.814	0.529	0.799
	MCC	0.666	0.167	0.585	0.053	0.589
	AUROC	0.831	0.583	0.775	0.527	0.792
Multiclass	Accuracy	0.705	0.533	0.437	0.706	0.681
	BA	0.668	0.343	0.401	0.250	0.618
	F1-Score	0.698	0.444	0.433	0.585	0.670
	MCC	0.548	0.125	0.182	0.000	0.497
%AD		79%	1%	9%	11%	100%

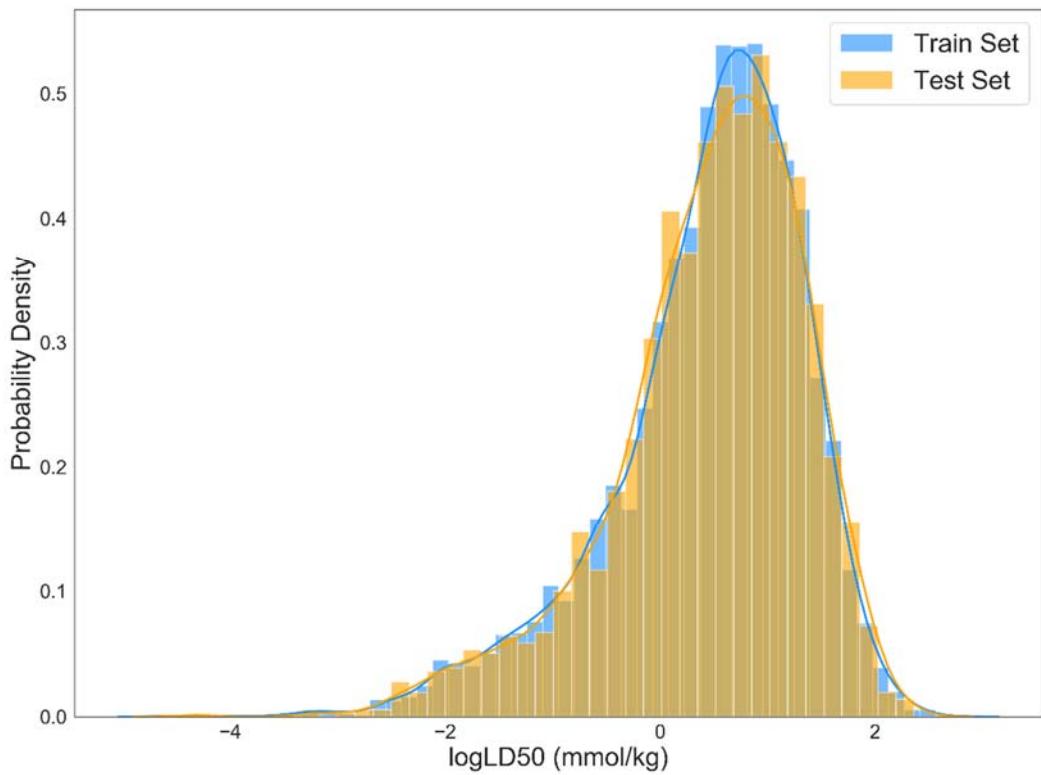


Figure S1. Distribution of $\log\text{LD}_{50}$ (mmol/kg) of Train and Test Sets.

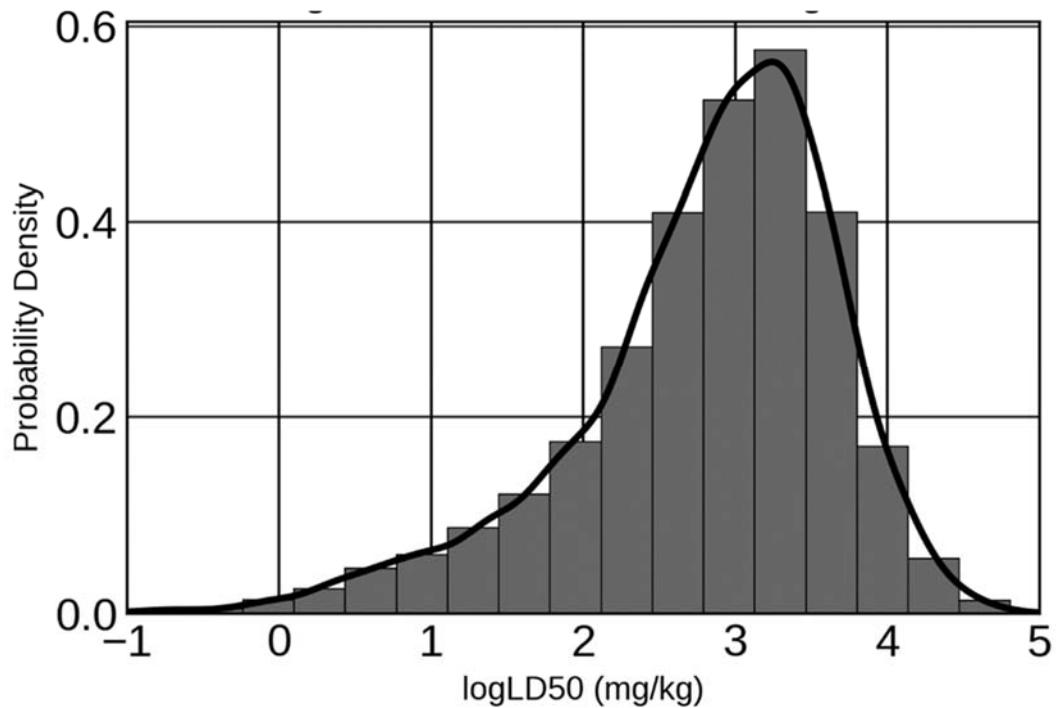


Figure S2. Distribution of logLD50 (mg/kg) of the Training Set.