

Supplement

Lisa-Marie Rolli^{1*}, Loulwah Arnaout¹ and Andrea Volkamer¹

¹Data Driven Drug Design, Center for Bioinformatics, Saarland Informatics Campus, Saarland University, PharmaScienceHub (PSH), Campus, Saarbrücken, 66123, Saarland, Germany.

*Corresponding author(s). E-mail(s): lisa-marie.rolli@uni-saarland.de;
Contributing authors: loar00001@stud.uni-saarland.de;
volkamer@cs.uni-saarland.de;

1 Hyperparameters

Model	Hyperparameters	Value
SVM [1]	probability	True
	random_state	42
	class_weight	'balanced'
RF [1]	random_state	42
	class_weight	'balanced'
CatBoost [2]	iterations	1000
MLP [1]	early_stopping	True
	random_state	42
	learning_rate	'adaptive'
	solver	'auto'
	batch_size	

Table 1 Fixed hyperparameters for SVM, RF, CatBoost, and MLP.

2 Package Versions

package name	version
tabPFN	2.2.1
scikit-learn	1.6.1
pandas	2.3.3
numpy	2.3.5
tabPFN-extensions	0.1.6
autogluon.tabular	1.4.0
catboost	1.2.8
scipy	1.16.2
statsmodels	0.14.5
matplotlib	3.10.7
seaborn	0.13.2
mrmr-selection	0.2.8
datsail	1.2.3
scikit-posthocs	0.12.0

Table 2 Open source Python libraries that we used, along with the respective package version.

3 Additional Results

3.1 DR Method Comparison For Physicochemical Properties

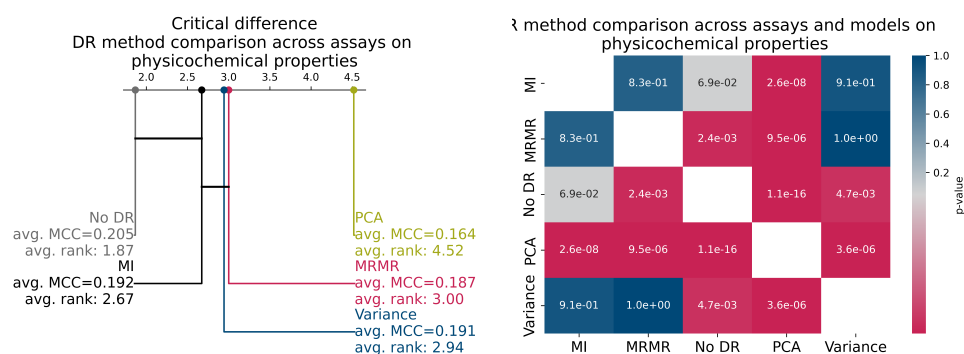


Fig. 1 Results of the Nemenyi test performed to compare DR methods across all combinations with physicochemical properties for all tested assays. The left side shows the critical difference plot; if there is no significant difference between the average rank of different DR methods, they will be connected by a vertical line. The right plot shows the p-value heat map; a red p-value indicates that the difference was significant with significance level of 0.05.

3.2 Dataset Characteristics

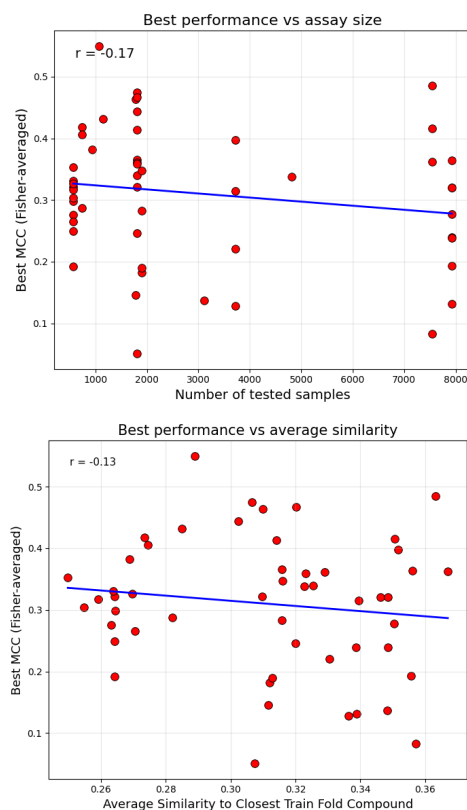


Fig. 2 Relationship between best achieved performance and assay size (left) or Tanimoto similarity between test fold and closest training compound (right).

References

- [1] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011)
- [2] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A.: Catboost: unbiased boosting with categorical features. In: *Advances in Neural Information Processing Systems*, vol. 31. Curran Associates, Inc., Red Hook, NY, USA (2018)