P-57

Meta-Learning as an Improvement of Machine Learning Methods Performance in Virtual Screening.

Sabina Smusz [a,b], *Rafał Kurczab* [a], *Andrzej J. Bojarski* [a]

[a] *Department of Medicinal Chemistry, Institute of Pharmacology, Polish Academy of Sciences, Smętna 12t, 31-343 Kraków*
[b] *Department of Crystallochemistry of Drugs, Faculty of Chemistry, Jagiellonian University, 3 Ingardena Street, 30-060 Kraków*
e-mail: smusz@if-pan.krakow.pl

Computational methods are widely used in pharmaceutical industry - both ligand and structure-based approaches are applied to virtual screening tasks, where large libraries of chemical compounds undergo evaluation in order to select drug candidates [1].
Recently, many applications of machine learning methods in this process have been reported. Their major task is to assign objects (in our case: molecules) into classes (here: active or inactive), but they can also carry out numerical classification that might be helpful e.g. in predicting ADME properties [2]. In order to improve the performance of classification algorithms, meta-learning strategy was developed. It gains knowledge from analyzing a number of subtasks, increasing efficiency of base classifier by using additional methods such as bagging and boosting [3].
We took into account four meta-learning algorithms implemented in WEKA package [4] and three different base classifiers. Their performance was examined depending on the type of molecular fingerprints used for representing chemical structures and the number of active compounds present in the training sets. Time required for building predictive models was also measured. For the case study, we used the 5-HT_{1A} antagonists taken from MDDR database (actives), and a set of structures randomly selected from ZINC database [5] (inactives).

Acknowledgements
The study was partly supported by a grant PNRF-103-AI-1/07 from Norway through the Norwegian Financial Mechanism within the Polish-Norwegian Research Fund; http://www.cns-platform.eu/

[1] Sousa S. F., Cerqueira N. M. F. S. A., Fernandes P. A., Ramos M. J.: *Comb. Chem. High Throughput Screen.* 13 (2010), 442.
[2] Melville J. L., Burke E. K., Hirst J. D.: *Comb. Chem. High Throughput Screen.* 12 (2009), 332.
[3] Vilalta R., Drissi Y.: *Artif. Intell. Rev.* 18 (2002), 77.
[4] Hall M., Frank E., Holmes G., Pfahringer B., Reutemann P., Witten I. H.: *SIGKDD Explorations* 11 (2009), 10.
[5] Irwin J. J., Shoichet B. K.: *J. Chem. Inf. Model.* 45 (1), 177.