Machine learning methods as virtual screening tools in computer-aided drug design

Sabina Smusz[1,2], Rafał Kurczab[2], Andrzej J. Bojarski[2]

[1]Faculty of Chemistry, Jagiellonian University,
R. Ingardena 3, 30-060 Kraków
[2] Institute of Pharmacology, Polish Academy of Sciences,
Smętna 12, 31-343 Kraków
sabina.smusz@gmail.com

Virtual screening (VS) techniques are widely used in the process of drug design as a tool for evaluating large libraries of chemical compounds and identifying the active ones. There are two approaches for VS methods: structure-based and ligand-based. The former selects active compounds through modeling and evaluating their interaction with a specific target based on 3D structure of the active site. The latter relies on knowledge of compounds that bind to a receptor for building a pharmacophore model that is then used to determine whether a candidate ligand is active or not.

Different machine learning methods have been tested in their potential application in ligand-based virtual screening. We analysed and compared their performance depending on the number of active compounds present in the training set, the type of fingerprints and the type of filter used for attributes selection in data preprocessing.

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.

References:

[1]. Plewczynski, D.; Spieser, S. A. H.; Koch, U. J. Chem. Inf. Model. **2006,** 46, 1098-1106. Geppert, H.; Vogt, M.; Bajorath, J. J. Chem. Inf. Model. **2010,** 50, 205–216.
[2]. Schwaighofer, A.; Schroeter, T.; Mika, S.; Blanchard, G. Combinatorial Chemistry & High Throughput Screening, **2009,** 12, 453-468.
[3]. Agarwal, S.; Dugar, D.; Sengupta, S. J. Chem. Inf. Model. **2010,** 50, 716–731.