[P38] Generative Topographic Mapping approach as a Virtual Screening tool: a benchmarking study

Arkadii Lin[a,b], Dragos Horvath[a], Gilles Marcou[a], Alexandre Varnek[a], and Bernd Beck[b]

[a] *Laboratory of Chemoinformatics, Faculty of Chemistry, University of Strasbourg, 4, Blaise Pascal str., 67081 Strasbourg, France*
[b] *Department of Medicinal Chemistry, Boehringer Ingelheim Pharma GmbH & Co. KG, Birkendorferstrasse 65, 88397 Biberach an der Riss, Germany*

Generative Topographic Mapping (GTM) is a dimensionality reduction method that can be used for large chemical data visualization and analysis. [1] Recently it was tested as a tool for large chemical databases comparison (PubChem-17, ChEMBL-17, and FDB-17). [2] It was also tested as a machine learning method for Quantitative Structure-Activity Relation (QSAR) tasks. [3, 4]

Here, the GTM approach is tested as a ligand-based Virtual Screening (VS) tool and compared with some popular machine-learning methods such as Random Forest, Neural Networks, and Similarity searching with data fusion. For this purpose, four "universal" maps built for 1.5M ChEMBL compounds and efficiently separating actives and inactives for 618 biological targets were used in virtual screening. Activity landscapes for 9 targets prepared for these maps were used in virtual screening of Directory of Useful Decoys (DUD) compounds. Benchmarking results show that universal GTMs perform slightly worse than the "local" GTM built for the ligands of particular target whereas the predictive performance of the latter is similar to that of popular machine-learning approaches (see Figure 1). Notice that universal GTM supports multiple predictive landscapes on a single map, thus, representing an interesting tool for data visualization, analysis and virtual screening

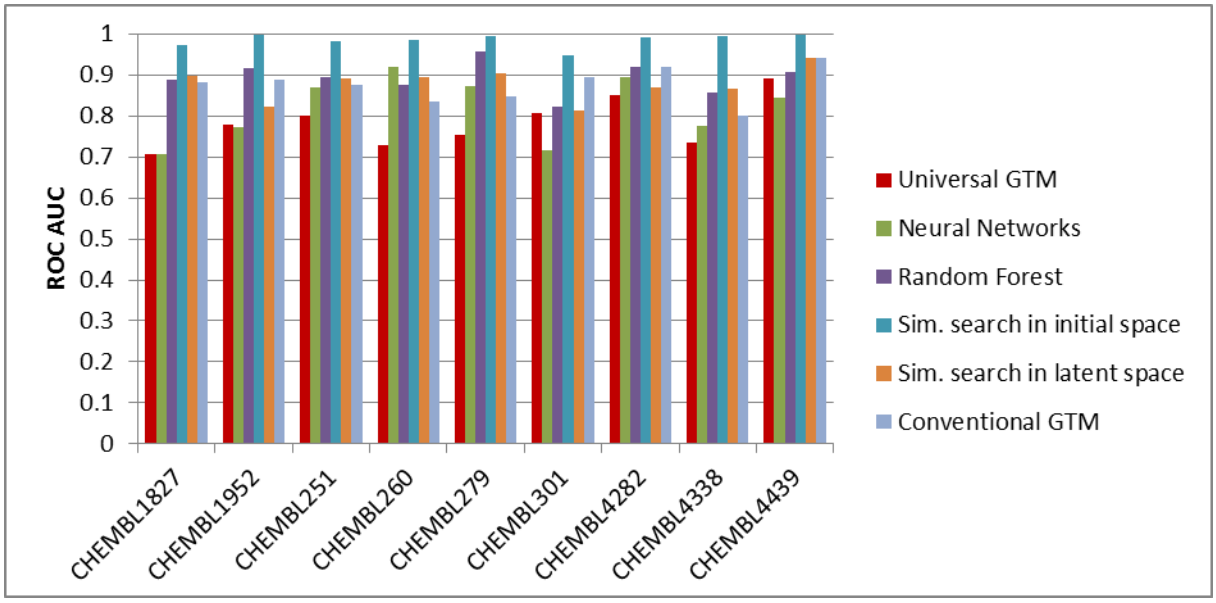

Figure 1. ROC AUC values obtained in virtual screening of 9 datasets from DUD using the models developed on ChEMBL compounds

Bibliography:

[1] Bishop CM, Svensén M, Williams CK. GTM: *Neural computation*. **1998**, 10(1), 215-34.
[2] Lin, A., Horvath, D., Afonina, V., Marcou, G., Reymond J.L. and Varnek, A. *ChemMedChem*. doi:10.1002/cmdc.201700561.
[3] Gaspar HA, Baskin II, Marcou G, Horvath D, Varnek A. *J. Chem.Inf.Mod.*. **2015**, 55(1), 84-94.
[4] Sidorov P, Gaspar H, Marcou G, Varnek A, Horvath D. *J. Comp.-Aided Mol. Design*. **2015**, 29(12), 1087-108.