

POSTER PRESENTATION

Open Access

Potency-directed similarity searching using support vector machines

Kathrin Heikamp*, Anne M Wassermann, Jürgen Bajorath

From 7th German Conference on Chemoinformatics: 25 CIC-Workshop
Goslar, Germany. 6-8 November 2011

Searching for active compounds in screening databases is one of the main tasks in chemoinformatics [1,2]. For this purpose, different approaches have been developed, with an increasing interest in machine learning and data mining methods [3]. Among these, support vector machine (SVM) learning has proven to be a powerful search technique in many instances [3]. Several applications of SVMs have been reported that focus on compound ranking in similarity searching [4-6]. However, similarity search and machine learning methods that are commonly utilized for virtual screening generally do not take compound potency information into account. Regardless of the applied methods, one typically attempts to distinguish “active” from “inactive” compounds. With the exception of QSAR models adapted for compound screening, no approaches have thus far been introduced that incorporate potency information as a parameter and direct search calculations toward the recognition of potent hits. Here, an SVM approach for potency-directed virtual screening is introduced [7]. A newly designed structure-activity kernel and an SVM linear combination model take potency information of reference molecules into account. Applied to high-throughput screening data sets, this potency-directed SVM approach met or exceeded the recall performance of standard SVM ranking and led to a notable enrichment of highly potent hits in database selection sets.

Published: 1 May 2012

References

1. Stumpfe D, Bajorath J: **Applied virtual screening: strategies, recommendations, and caveats.** In *Methods and Principles in Medicinal Chemistry, Virtual Screening, Principles, Challenges, and Practical Guidelines*. Weinheim: Wiley-VCH;Sottriffer C. 2011:73-103.

Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universität Bonn, Dahlmannstr. 2, D-53113 Bonn, Germany

2. Eckert H, Bajorath J: **Molecular similarity analysis in virtual screening: foundations, limitations and novel approaches.** *Drug Discov Today* 2007, **12**:225-233.
3. Geppert H, Vogt M, Bajorath J: **Current trends in ligand-based virtual screening: molecular representations, data mining methods, new application areas, and performance evaluation.** *J Chem Inf Model* 2010, **50**:205-216.
4. Jorissen RN, Gilson MK: **Virtual screening of molecular databases using a support vector machine.** *J Chem Inf Model* 2005, **45**:549-561.
5. Wassermann AM, Geppert H, Bajorath J: **Searching for target-selective compounds using different combinations of multiclass support vector machine ranking methods, kernel functions, and fingerprint descriptors.** *J Chem Inf Model* 2009, **49**:582-592.
6. Agarwal S, Dugar D, Sengupta S: **Ranking chemical structures for drug discovery: a new machine learning approach.** *J Chem Inf Model* 2010, **50**:716-731.
7. Wassermann AM, Heikamp K, Bajorath J: **Potency-directed similarity searching using support vector machines.** *Chem Biol Drug Des* 2011, **77**:30-38.

doi:10.1186/1758-2946-4-S1-P12

Cite this article as: Heikamp et al.: Potency-directed similarity searching using support vector machines. *Journal of Cheminformatics* 2012 **4**(Suppl 1):P12.

Publish with **ChemistryCentral** and every scientist can read your work free of charge

“Open access provides opportunities to our colleagues in other parts of the globe, by allowing anyone to view the content free of charge.”

W. Jeffery Hurst, The Hershey Company.

- available free of charge to the entire scientific community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
<http://www.chemistrycentral.com/manuscript/>


ChemistryCentral