

sich sehr homogene Gruppen bezüglich der Bioaktivität in diesem Netz bilden. Das zeigt, dass die gewählte Molekülrepräsentation auf Basis des Interaktionsmusters sehr gut geeignet ist.

Kern des Verständnisses der Variation der Bioaktivität verschiedener Moleküle an einer bestimmten biologischen Zielstruktur ist das Wissen darüber, wie die chemische Struktur die biologische Aktivität beeinflusst. Mit Techniken der Chemieinformatik lässt sich diese Information wie oben beschrieben vielfach sehr effizient extrahieren.

Literatur

- [1] a) BAUMANN, K. 1999: Uniform-length molecular descriptors for Quantitative Structure-Property Relationships (QSPR), Quantitative Structure-Activity Relationships (QSAR), classification studies, and similarity searching. *TrAC* **18**: 36–46. b) STIEFL, N. & K. BAUMANN 2003: Mapping Property distributions of molecular surfaces (MaP): Algorithm and evaluation of a novel 3D Quantitative Structure-Activity Relationship technique. *J. Med. Chem.* **46**: 1390–1407.
- [2] BREIMAN, L. 2001: Random Forests. *Mach. Learning* **45**: 5–32.
- [3] BAUMANN, K. 2003: Cross-validation as the objective function for variable selection techniques. *TrAC* **22**: 395–406.
- [4] POLITIS, D.N. 2013: Model-free model-fitting and predictive distributions. *Test* **22**: 183–221.
- [5] EFRON, B. 2014: Estimation and accuracy after model selection. *J. Amer. Statist. Assoc.* **109**: 991–1007.
- [6] WOLLENHAUPT, S. & K. BAUMANN 2014: InSARa: Intuitive and interactive SAR interpretation by reduced graphs and hierarchical MCS-based network navigation. *J. Chem. Inf. Model.* **54**: 1578–1595.