Fusing Kernels using Geometric Mean of Kernel Matrices

Pooya Zakeri 1, 2, Yves Moreau 1, 2

1) KU Leuven, Department of Electrical Engineering-ESAT, SCD-SISTA, Leuven, Belgium
2) iMinds Future Health Department, Leuven, Belgium

E-mail: pooya.zakeri@esat.kuleuven.be

Kernel methods are being more interesting object when we want to integrate heterogeneous biological data. Various methods have been proposed to fuse multiple kernels instead of using a single one. Most techniques for multiple kernel learning (MKL) focus on learning convex linear combination of kernels. In addition to limitation of linear combination, going with such approaches could cause losing potentially useful data information.
We propose a new technique that combines kernels in taking Geometric mean instead of convex linear combination. Under the assumption of independence for the various data sources, information from multiple probabilistic biological networks could be combined by the Noisy-Or model. We show that geometric mean of positive definite kernels could be considered as the equivalent of the Noisy-Or model for them.
We evaluate our method for classification on protein fold recognition data set composed of 12 different representative models of protein samples. In order to compare and test the performance of our proposed approach and other MKL algorithms, we implemented ten MLK algorithms. The overall accuracy on the protein fold recognition test set obtained by our method is about 75%. This result indicates that our proposed approach is comparable to the results of best existing approach.

Plasticity models for Cytochrome P450 metabolite prediction: an integrative approach

D.P. Geerke 1, L. Peric-Hassler 1, R. Pool 1, N. Bon 1, D. Truan 1, C.R. Vosmeer 1, C. Oostenbrink 2, A.K. Malde 3, A.E. Mark 3, M. el-Kebir 4, S. Canzar 4, G.W. Klau 4

1) Division of Molecular and Computational Toxicology, VU University Amsterdam, the Netherlands
2) Institute of Molecular Modeling and Simulation, BOKU University, Vienna, Austria
3) Molecular Dynamics Group, University of Queensland, Brisbane, Australia
4)Algorithmic computational biology, CWI Life Sciences, Amsterdam, the Netherlands

E-mail: d.p.geerke@vu.nl

Background
Drug metabolism plays a major role in possible toxicity of drugs. Therefore it is highly important to be able to predict into which metabolites drugs are converted. Cytochrome P450s (CYPs) metabolize 75% of the marketed drugs. However, predicting CYP metabolite formation is difficult because of the large protein flexibility (plasticity) of CYPs.

Methods
Our efforts will be presented to combine molecular dynamics simulation techniques (MD), virtual-screening methods and graph-theory approaches, in order to develop an accurate, industrially attractive method for CYP metabolite prediction that efficiently includes plasticity effects.

Results
By carefully selecting protein conformations from MD, binary trees have been constructed for two of the most flexible CYPs, to improve virtual screening-based metabolite prediction towards industrially relevant accuracy. Subsequently, we could show that use of the selected protein conformations is necessary to efficiently and accurately predict CYP selectivity. An important step in making selectivity prediction applicable in industry is automated generation of topologies for use in MD. By using graph theory, the issue of charge-group partitioning could be solved.

Conclusion
By combining our MD, virtual screening and graph-theory approaches, an integrative method for efficient CYP metabolite prediction has come within reach, which will now be extended towards other CYPs.