

Title: Utilization of latent semantic analysis in virtual screening

Author: Jiří Kolář

Department: Department of Software Engineering

Supervisor: RNDr. David Hoksza, Ph.D., Department of Software Engineering

Abstract: Aim of this thesis is to investigate utilisation of latent semantic indexing in Virtual screening. We have examined existing VS method called latent semantic structural indexing (LaSSI) and compared performance of different structural fingerprints. Additionally, we have developed a new model that compare fragments of molecules by usage of latent semantic indexing. Fragments are characterized by formula based counts and descriptors describing the physicochemical properties. Results of our methods are compared to VS techniques using directly standard fingerprints.

Keywords: virtual screening cheminformatics ligand-based fingerprints ECFP TT latent semantic analysis LaSSI