An enhancement of bayesian inference network for ligand-based virtual screening using minifingerprints

Abstract

Selection and identification of a subset of compounds from libraries or databases, which are likely to possess a desired biological activity is the main target of ligand-based virtual screening approaches. The main challenge of such approaches is achieving of high recall of active molecules. To this end, different models of Bayesian network have been developed. In this study, we enhance the Bayesian Inference Network (BIN) using a subset of selected molecule's features. In this approach, a few features that represent the Minifingerprints (MFPs) were filtered from the molecular fingerprint features based on an analysis of distributions of molecular descriptors and structural fragments into large compound data set collections. Simulated virtual screening experiments with MDL Drug Data Report (MDDR) data sets showed that the proposed method provides simple ways of enhancing the cost effectiveness of ligand-based virtual screening searches, especially for higher diversity data set.