

Screening chemických knihoven je důležitou součástí výzkumu léčiv. Existující chemické knihovny obsahují miliony sloučenin. Screening v takovém rozsahu je velice nákladný, z tohoto důvodu je často používáno virtuálního screeningu. Existuje několik variant virtuálního screeningu, ligand-based virtual screening je jednou z nich. Využívá podobnosti screenovaných sloučenin ke známým sloučeninám. Na výsledky virtuálního screeningu má vliv nejen použitá podobnostní metoda, ale také zvolená reprezentace chemických sloučenin.

V této práci prezentujeme reprezentaci chemických sloučenin založenou na fingerprintech. Naše reprezentace využívá fragmentů chemické sloučeniny k její reprezentaci jakožto celku. Každý fragment je reprezentován fyzikálně-chemickými vlastnostmi. Reprezentace je vysoce parametrizovatelná a to zejména v oblasti výběru fyzikálně-chemických vlastností a jejich aplikace. Pro otestování naší reprezentace jsme využili existujícího frameworku pro benchmark virtuálních screeningů. Výsledky ukázaly, že naše reprezentace je srovnatelná s nejlepšími existujícími, navíc na některých datasetech dosáhla nejlepších výsledků.