

Abstrakt

Metody pro měření podobnosti hmotnostních spekter a struktur malých molekul jsou důležité pro pokroky v lékařské chemii, farmakologii a metabolomice. Mezi běžně využívané metody pro porovnání hmotnostních spekter molekul patří kosinová podobnost. Jedná se o míru podobnosti mezi dvěma nenulovými vektory, která měří kosinus úhlu mezi nimi. Porovnání hmotnostních spekter molekul umožňuje vyhledávání v databázích molekul, klastrování spekter a vyhledávání ve spektrálních knihovnách. Strukturální podobnost se pak měří na základě nejrozličnějších molekulárních fingerprintů, jako jsou například Daylight fingerprint, RDKit fingerprint, Atom-Pair Fingerprint, Topological Torsion Fingerprint, Extended-Connectivity Fingerprint a další. Tyto fingerprinty jsou pak porovnávány pomocí koeficientů podobnosti. Zmíněné metody pro porovnávání struktur a hmotnostních spekter molekul lze aplikovat pomocí bioinformatických knihoven RDKit a CDK pro generování a analýzu strukturálních fingerprintů a knihovny matchms pro porovnání hmotnostních spekter. Práce poskytuje teoretický přehled jak molekulárních deskriptorů, zahrnující rozmanité typy molekulárních fingerprintů a techniky pro měření strukturální podobnosti, tak principu hmotnostní spektrometrie a přístupu k porovnání hmotnostních spekter. Praktická část práce je zaměřena na analýzu strukturálních a spektrálních dat, získaných z databáze MoNA (MassBank of North America), pomocí nástrojů RDKit, CDK a MatchMS a na následné porovnání získaných strukturálních a spektrálních podobností.