[BIOINFO 2025]

SubDyve: Subgraph-Driven Dynamic Propagation for Virtual Screening Enhancement Controlling False Positive

Jungseob Yi[1], Seoyoung Choi[2], Sun Kim[1,2,3,4] and Sangseon Lee[5,*]

[1]*Interdisciplinary Program in Artificial Intelligence, Seoul National University*
[2]*Department of Computer Science and Engineering, Seoul National University*
[3]*Interdisciplinary Program in Bioinformatics, Seoul National University*
[4]*AIGENDRUG Co., Ltd., Seoul*
[5]*Department of Artificial Intelligence, Inha University*
**Corresponding author:* ss.lee@inha.ac.kr

Virtual screening (VS) aims to identify bioactive compounds from vast chemical libraries, but remains difficult in low-label regimes where only a few actives are known. Existing methods largely rely on general-purpose molecular fingerprints and overlook class-discriminative substructures critical to bioactivity. Moreover, they consider molecules independently, limiting effectiveness in low-label regimes. We introduce SubDyve, a network-based VS framework that constructs a subgraph-aware similarity network and propagates activity signals from a small known actives. When few active compounds are available, SubDyve performs iterative seed refinement, incrementally promoting new candidates based on local false discovery rate. This strategy expands the seed set with promising candidates while controlling false positives from topological bias and overexpansion. We evaluate SubDyve on ten DUD-E targets under zero-shot conditions and on the CDK7 target with a 10-million-compound ZINC dataset. SubDyve consistently outperforms existing fingerprint or embedding-based approaches, achieving margins of up to +34.0 on the BEDROC and +24.6 on the $EF_{1\%}$ metric.