[BIOINFO 2025]

Substructure-driven Junction-Tree for Interpretability in ADMET Tasks

Myungjin Lee[1], Juhyeon Mo[1], Yeeun Lee[2], and Sangsoo Lim[1,*]

[1]*Department of Computer Science and Artificial Intelligence, Dongguk University*
[2]*Department of AI Convergence, Dongguk University*
**Corresponding author: sslim@dgu.ac.kr*

Accurate prediction of molecular properties such as absorption, distribution, metabolism, excretion, and toxicity (ADMET) is crucial in the drug discovery process. However, ensuring the interpretability of prediction models remains a significant challenge. Although existing Graph Neural Network (GNN)-based models demonstrate high predictive accuracy, they often suffer from the 'black box' limitation of failing to explain which molecular substructures contribute to specific properties. In this paper, we propose SJoINT: a novel dual-encoder model that explicitly utilizes molecular substructure information to achieve accuracy and interpretability. SJoINT takes an atom-level molecular graph and a corresponding junction tree, representing the molecular scaffold and functional groups, as dual inputs. It learns the relationship between these two representations through an iterative cross-attention mechanism. This process enables the model to identify the key substructures associated with particular ADMET properties, thereby improving its interpretability. Experiments conducted on the MoleculeNet benchmark show that SJoINT achieves state-of-the-art performance on the SIDER task for predicting toxic side effects. This demonstrates the effectiveness of our approach to integrating micro-level (atomic) and macro-level (substructure) information, particularly for predicting complex drug toxicity profiles. In conclusion, our work suggests that SJoINT is a promising approach that provides accurate predictions and offers interpretable insights by elucidating the relationship between molecular substructures and drug properties.