Predicting Structural Similarity Between Molecules Using Graph Neural Networks

Deng, Sichen (School: High School Affiliated to Shanghai Jiao Tong University)

Molecule properties and functions are highly influenced by their structures. Investigating the structural similarity between molecules is a fundamental task in chemistry-related fields, which is able to benefit a wide range of downstream tasks. Graph edit distance (GED) is a representative metric for measuring the structural similarity between molecules. However, exactly calculating the GED is an NP-hard problem. In this paper, we use graph neural networks to process a pair of molecules and output their representations, finally feeding the two representations into a regression model to predict their ground-truth GED. The experimental results show that our model significantly outperforms other molecule representation learning methods in GED prediction. Moreover, our model is shown to be significantly more time-efficient than the algorithm that calculates the exact GED. The proposed methodology can provide guidance for similar molecule retrieval and drug discovery.