HeLU-DTI: Drug Target Interaction Prediction via Heterogeneous Graph Neural Networks and Large Language Models

Ye, Andy (School: Eastside Preparatory School)

Drug discovery, a complex and time-consuming process, heavily relies on determining affinities between drugs and targets. With the introduction of computational methods, researchers can now utilize vast datasets combined with deep learning to uncover hidden patterns in drug-target bindings, saving valuable time and resources. However, powerful existing models still have limitations. To address these shortcomings, we develop a novel end-to-end model, named HeLU-DTI. The algorithm involves using unsupervised pre-trained large language models to generate rich representations of drug and protein data, followed by using a knowledge graph to compute embeddings for diseases, their phenotypes, and drug effects. The multi-modal information is then aggregated into a heterogeneous graph neural network, allowing effective information propagation between interconnected nodes. Afterward, a multi-layer linear perceptron is employed to accurately predict corresponding affinities. The model's accuracy was evaluated on two datasets, BindingDB and BioSNAP. The results show that the model outperforms four baseline models, demonstrating its significance in precisely classifying unknown relationships between drugs and targets. To assess generalizability through unseen testing, the model was trained on an independent BindingDB dataset and tested on an independent NR dataset. The model achieved high accuracy alongside providing accurate predictions of unseen drug-target interactions. HeLU-DTI not only predicts the ground truth but also goes beyond it by identifying reasonable targets or corresponding drugs that were not previously known. The model has the potential to significantly improve drug development by providing a more efficient and accurate way of predicting drug-target interactions.

Awards Won:

Fourth Award of $500