Hnet-DTI: Incorporating heterogeneous information network for drug-target interaction prediction

Gökhan	Özsarı	gozsari@ceng.metu.edu.tr	Turkey	Middle East Technical University
Ahmet Süreyya	Rifaioğlu	ahmet.rifaioglu@uni-heidelberg.de	Germany,Heidelberg University	
Tunca	Dogan	tuncadogan@gmail.com	Turkey	Hacettepe University
Volkan	Atalay	vatalay@metu.edu.tr	Turkey	Middle East Technical University

Identifying drug-target interactions (DTIs) is crucial in drug discovery. In silico prediction of interactions between drugs and target proteins is needed to accelerate the drug discovery process. Although many DTI prediction methods have been proposed in recent years, there is still room for improvement. This study proposes a graph neural network-based method to exploit the heterogeneous information network for drug-target interaction prediction. The study includes two major contributions:

Heterogeneous information network: We first constructed a heterogeneous information network that includes two types of nodes as nodes and three types of edges. The network includes 15 291 protein nodes, 3 444 drug nodes, 475 850 protein-protein interaction edges, 2 296 964 drug-drug interaction edges and 6 018 drug-protein interaction edges.

DTI prediction method: We applied a three-layer graph convolution neural network (GCN) to learn the embeddings of drugs and proteins by utilizing neighborhood information in the network. GCNs were employed for each node type to aggregate neighborhood information. The final embeddings were a concatenation of embeddings from three layers of GCN. Finally, the embeddings of drugs and proteins were multiplied with a learnable weight matrix to predict the drug-target interactions.

We evaluated the performance of the proposed method by using a 10-fold cross-validation technique and compared the proposed method baseline classifiers: random forest (RF) and support vector machines (SVM). The results indicate that the proposed method outperforms the baseline classifiers by reaching a 0.898 AUROC score.