ENStaDTI: An Ensemble Graph-Learning and Multilayer Stacking-Based Deep Learning Pipeline to Predict Novel Drug-Target Interactions for Large-Scale Drug Repurposing and Drug Discovery

Amalapurapu, Shreya (School: Texas Academy of Mathematics and Science)

The inefficient treatment of thousands of diseases contributes to the need for a quick and cost-efficient alternative to current drug discovery methods. Drug development is extremely time-consuming and expensive; therefore, drug repurposing is the most efficient way to identify treatment options for these diseases. Drug Target Interactions (DTI) are vital in identifying promising drug-repurposing candidates since they indicate how drugs and proteins interact. ENStaDTI utilizes novel network-based techniques in combination with deep-graph learning on heterogeneous multiplex biological networks to identify strong drug-repurposing possibilities. The pipeline has three steps: creating a graph containing known DTIs and representing the similarity amongst drugs and targets, generating lower dimensional drug and target node embeddings using the network, and employing deep learning on the obtained embeddings for DTI predictions. Node embeddings are computed using a novel approach where five embedding algorithms are applied individually on the heterogeneous multiplex network. A novel late-integration-based stacking ensemble classifier is then developed on the embeddings, outputting strong DTI predictions, which are further validated using molecular docking studies. ENStaDTI was then run on all possible drug-target combinations, identifying thousands of new drug-target interactions, spanning more than 7,000 diseases, including orphan, antibiotic-resistant, and viral diseases. ENStaDTI was able to predict literature-supported and previously unknown DTIs, therefore, validating it as a tool that can be used for large-scale drug repurposing and discovery.

Awards Won: