Research Week 2020

Integrating domain knowledge, interpretable deep learning, and uncertainty quantification for computational prediction of cell response to drug combinations

Nathaniel Evans, B.S., Aurora Blucher, Shannon McWeeney

OHSU

Keywords

Precision Oncology, Drug Response, Deep Learning, Graphical Models, Graph Convolutional Networks, Visible Neural Networks, Uncertainty Quantification, Bayesian Modeling

Abstract

A major challenge in precision oncology is accurately modeling the molecular machinery that governs cellular sensitivity or resistance to a given drug or drug combination. While predicting drug response is an active area, model performance and interpretability vary depending on the method and prediction task, and leaves room for improvement. Challenges for drug response prediction include high feature dimensionality, measurement variability, low volume of functionally annotated tumor genomes, and limited understanding of drug-target interactions. Moreover, while domain literature describes many of the molecular interactions critical to drug response, it can be difficult to systematically encode this information in a framework conducive to predictive modeling. One effective family of methods use graph-based representations of domain knowledge captured in databases such as Reactome [3], KEGG [4], STRING [2] and Pathway Commons [8]. Graph structures are well-suited to encode numerous feature types and narrow the model hypothesis space by limiting feature (nodes) and interactions (edges) according to supporting domain evidence. Recently, the deep learning field has proposed applications specific to graph structures, such as Graph Convolutional Networks (GCN) [7], that show promise. The flexibility and predictive value of deep learning is renowned across fields such as machine learning, signal processing, and image recognition; however, this often comes at a loss of interpretability since the latent (or underlying) representations rarely map to real world concepts. To address this issue, Visible Neural Networks (VNNs) were proposed [1,6] to constrain latent space representations to hierarchies based on real-world pathway domain knowledge. Further, the rise in popularity of Bayesian methods have made uncertainty quantification highly approachable, and are well suited for high-dimensional and data-poor problems such as drug response prediction. In this talk, I'll discuss how these methods may be combined in a biology-centric framework that provides robust drug-response prediction, an interpretable latent space, and uncertainty quantification.