Predictive Modeling of Optimal Cancer Therapies

Xu, Joyce

Synergistic drug combinations for cancer have the potential to save millions of lives each year through lowering toxicity, limiting side effects, and drastically increasing the efficacy of the given treatment. Through predictive modeling of synergy, this project developed a comprehensive, clinically applicable tool to optimize cancer therapies. The baseline model utilized a gradient boosting machine, trained on 11 initial features regarding the monotherapy information of each drug, to predict the synergy of combinations. Additional properties of both the compounds and the cancer cell line were subsequently mined from various online databases and clustered to construct informative features for the model. A target interaction network was created from protein-protein interaction information downloaded from another database (BioGrid), and a genetic algorithm wrapper was developed to perform feature selection and optimize model performance. The final cross-validated model achieved a strong classification accuracy (0.83) and AUC (0.79) for identifying clinically significant instances of synergy, which is not only on par with the most recent research but outperforms many existing methods. In particular, the emphasis placed on cell line information in this model helped account for the differing disease and genomic profile of each patient, contributing to the rise of personalized medicine. Features selected for by the genetic algorithm can be additionally analyzed as biomarkers for synergy and used to further biological understanding of drug interactions.

Awards Won:

Third Award of $1,000

American Statistical Association: Fourth Award of $500