Nathaniel Polley (CRCT, Toulouse)

"Follow the Trees: Novel XGBoost Forest Algorithm Elucidates Unforeseen Mechanisms of Action in Cancer Survival from Diagnostic Metabolic Flux Data"

Abstract:

Modern in vivo methods involving patient-derived xenograph models have facilitated countless advancements in experimental design due to the ability to apply multiple treatment combinations in parallel to existing human tissue. Nonetheless, the preliminary usage of in vivo screens in search of new biological targets in cancer research remains expensive, time-consuming, and rarely produces a sample size large enough for statistical significance to be inferred. As a consequence, the demand persists for an alternative method to accurately discover therapeutic mechanisms of action based upon genotypic and demographic variation directly from patient samples. Fortunately, the incorporation of metabolic flux data with modern machine learning methods has led to recent breakthroughs in survival prediction among AML patients. PolyMORPHOS (Metabolic-Oriented Regression Predicting Human Overall Survival) derives metabolism flux balances using bulk patient RNAseq data at diagnosis in order to calibrate a progressive XGBoost model, therefore providing a diverse metabolic landscape to predict the fate of patients at early, middle, and late stages of survival. Per each survival iteration, polyMORPHOS calculates the metabolic reactions most pertinent to death prediction, thus enabling researchers and clinicians to design individual therapies to best reverse any unfavorable prognosis with needlepoint precision. Additionally, we announce SOURIS (Simulation Of Universal Reactions In Silico), the most recent installment to our predictive analytics suite. Using the previous polyMORPHOS model, chronological metabolic survival conditions are simulated hundreds of times by random-sampling the initial patient cohort. Within each simulation, the cohort is stratified based upon high, baseline, or low expression of a variable chosen by the researcher e.g. gene, protein, age, mutation status, etc. The resulting metrics provide insight as to which metabolic reaction states result in morbidity or favorable outcome per each stratified patient characteristic. Hence, the precise elucidation of relevant mechanisms of action at early stages of the experimental process will empower researchers to strategically dedicate more resources to effective therapeutic design, thus further bridging the gap between theoretical and tangible results.