

Title: Exploratory and Predictive Analytics for Precision Medicine

Abstract:

Before personalized healthcare, treatment plans varied in effectiveness across patients with the same cancer type, where researchers have found differences in genetics and lifestyles that explain the varying outcomes. Precision Medicine is an emerging field in Biomedical Sciences that seeks to target a personalized treatment plan using an individual's information (e.g., biomarkers, environment, family history, personal habits, and dietary patterns). In this project, we utilize women's health data from the Women's Healthy Eating & Living (WHEL) study to measure relationships between metrics and assess the predictive power of patients' data regarding the recurrence likelihood of breast cancer events. In our exploratory analysis, we generated in-depth multivariate visualizations to identify correlations, particularly patterns involving variables of interest, such as quality of life, tumor type, recurrence status, and demographic variables that may reveal biases in data collection due to systemic socioeconomic factors. We found that (1) patients with more role limitations due to physical/emotional problems reported answers reflecting higher quality of life and mental health; (2) larger increases in alpha-carotene are associated with stable quality of life over time; and (3) the patient group with missing data points at study conclusion has markedly lower numbers of individuals with at least a college education. We then select learning features for our Machine Learning models using a binary breast cancer recurrence status as the target variable. We achieved 89% recall, 87% accuracy and precision, with resampling being the best approach to our imbalanced dataset.