Abstract: Boosting Competing Risks

Alexis Bellot and Mihaela van der Schaar

Background and Problem - The co-occurrence of multiple diseases among the general population is an important problem as those patients have more risk of complications and represent a large share of health care expenditure. Predicting future time-to-event probabilities for these patients, such as for instance death from cancer or from cardiovascular disease (CVD), is a challenging problem because the risks of events are correlated (there are competing risks) while disease patterns may vary wildly even in narrow phenotypes as patients are highly heterogeneous.

Solution - We introduce a survival model with the flexibility to leverage a common representation of related events. Our model learns fully nonparametric survival estimates which give a highly flexible representation of a patients disease risk and associated survival probabilities without any assumptions on the data generating process.

Methods - Our method proceeds sequentially by jointly learning outcome-specific nonparametric survival estimators, that is a separate survival distribution for each potential disease. In every iteration we learn to improve predictions of previously misdiagnosed patients - a process called *boosting* - by drawing from the relationships learned from patients at risk of similar diseases. Final survival estimates result from a combination of predictions in every iteration In addition, our approach yields a measure of the relative covariate importance that accurately identifies relevant covariates for each disease, thereby enabling further medical understanding of disease interactions.

Results - We evaluated predictive performance with the cause-specific C-index on $72,000$ patients extracted from the Surveillance, Epidemiology, and End Results data suffering from Caner and simultaneously at risk of CVD. Our results give performance improvements ($2.5\% - 20\%$) of our method over existing techniques and provide novel insights through our variable importance measure.