

TITLE: A Semi-Parametric BART Model for Causal Inference**AUTHORS:** Yusufhan Balci, Mark van de Wiel and Jeroen Hoogland – Amsterdam UMC

ABSTRACT: Estimating and inferring subgroup-specific treatment effects is a classic challenge in personalized medicine and causal inference. The field has evolved rapidly, with most research focusing on estimating the Conditional Average Treatment Effect (CATE) by leveraging Machine Learners, which assume sample sizes and signal-to-noise ratio's which are hard to come by in medical settings, given that treatment effect modifiers are often very small and hard to detect. We propose a Semi-Parametric Bayesian Additive regression Trees (BART) model combining a BART prior for the prognostic function, and a regularized linear function with interactions with an Horseshoe prior to capture treatment effect heterogeneity. BART has been shown to outperform other Causal Machine learning methods. The semi-parametric approach allows us to efficiently detect treatment effect heterogeneity, while also providing robust counterfactual predictions, which are disregarded by methods that solely focus on predicting the CATE. These counterfactual predictions are especially important, as they provide inference about their prognosis under different treatment paths. We further employ Shapley values to quantify the contribution of each covariate and its interactions to the final CATE estimate. Our method outperforms BCF when the data generating process for the treatment is linear with interaction terms, while remaining robust to misspecification of the prognostic. Even under misspecification, when the data generating process is monotonic, our method is able provide satisfactory ranking of the treatment effect. When applied to clinical trial data, our method is competitive with Bayesian Causal Forests in root mean squared error of the CATE predictions in the limited sample size settings typical in medical applications, while providing improved calibration and interpretation. We also provide diagnostic methods to compare different causal inference methods, in order to make the best practical choice between models.

PRESENTING AUTHOR BIO: Yusufhan Balci is affiliated with Amsterdam UMC and VU Amsterdam. His research focuses on Bayesian causal inference and treatment effect heterogeneity, combining flexible modelling with interpretable effect modifiers in clinical trials and observational studies.