

BART Matching Method: Using Bayesian Additive Regression Tree for Data Matching

Authors : Gianna Zou

Abstract : Propensity score matching (PSM), introduced by Paul R. Rosenbaum and Donald Rubin in 1983, is a popular statistical matching technique which tries to estimate the treatment effects by taking into account covariates that could impact the efficacy of study medication in clinical trials. PSM can be used to reduce the bias due to confounding variables. However, PSM assumes that the response values are normally distributed. In some cases, this assumption may not be held. In this paper, a machine learning method - Bayesian Additive Regression Tree (BART), is used as a more robust method of matching. BART can work well when models are misspecified since it can be used to model heterogeneous treatment effects. Moreover, it has the capability to handle non-linear main effects and multiway interactions. In this research, a BART Matching Method (BMM) is proposed to provide a more reliable matching method over PSM. By comparing the analysis results from PSM and BMM, BMM can perform well and has better prediction capability when the response values are not normally distributed.

Keywords : BART, Bayesian, matching, regression

Conference Title : ICAISM 2022 : International Conference on Advances in Informatics, Statistics and Mathematics

Conference Location : New York, United States

Conference Dates : August 08-09, 2022