Conference Name: Annual Conference on Business, Economics Research and AI, 18-20 November 2026, Bangkok
Conference Dates: 18-Nov-2026 to 20-Nov-2026
Conference Venue: Oakwood Suites Tiwanon Bangkok, 229 Thanon Krungthep-Nonthaburi, Bang Khen, Bangkok 11000, Thailand
Appears in: PEOPLE: International Journal of Social Sciences (ISSN 2454-5899)
Publication year: 2026

Chaichana & Lawson, 2026
Volume 2026, pp. 812-813
DOI- https://doi.org/10.20319/icssh.2026.812813
This paper can be cited as Chaichana, J. & Lawson, N. (2026). Weighting Methods for Nonresponse in Sample Survey using Machine Learning. Annual Conference on Business, Economics Research and AI, 18-20 November 2026, Bangkok. Proceedings of Social Science and Humanities Research Association (SSHRA), 2026, 812-813

WEIGHTING METHODS FOR NONRESPONSE IN SAMPLE SURVEY USING MACHINE LEARNING

Jittima Chaichana
Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, 1518 Pracharat 1 Road, Wongsawang, Bangsue, Bangkok 10800, Thailand
jittima542545@gmail.com

Nuanpan Lawson
Department of Applied Statistics, Faculty of Applied Science, King Mongkut's University of Technology North Bangkok, 1518 Pracharat 1 Road, Wongsawang, Bangsue, Bangkok 10800, Thailand
nuanpan.n@sci.kmutnb.ac.th

Abstract

Survey nonresponse produces a challenge for estimating population mean, as it can lead to bias in survey estimates. Weighting adjustment using auxiliary variables is a common approach to reducing nonresponse bias. However, auxiliary information can itself be partially missing, influencing the performance of weighting adjustment methods. This study aims to compare three weighting methods: random forest, XGBoost, and BART. A simulation study is conducted using R program consisting of a study variable and two auxiliary variables, one fully observed and the

other partially missing. The performance of each method is evaluated using bias and mean square error.

Keywords:
Nonresponse, Weighting Adjustment, Missing Auxiliary Data, Machine Learning