Contribution ID: **64** Type: **not specified**

# Wrapper-Based Variable Selection for Interacting Variables in Manufacturing

In manufacturing, identifying variables that influence process outcomes is essential for control and optimization. While wrapper-based variable selection methods, such as Conditional Boruta by Rotari and Kulahci(2025), has been proposed, they remain susceptible to rejecting variables that only influence the outcome through interactions with other variables. Therefore, the purpose of the suggested method in this paper, is to extend the application to work for cases where variables are only important through interactions, we will call these *"interacting variables"*.

This paper proposes *WarmStart Conditional Boruta*, an extension of the Conditional Boruta method introduced by Rotari and Kulahci(2025). The Conditional Boruta method is a wrapper-based variable selection method that iteratively selects a subset of variables and evalutes their importance, to identify relevant variables. However, it can fail to identify *interacting variables*, as variables have a risk of being rejected in early iterations if they are not included in the same subset as their interacting partners, leading to a failure to observe the interaction effect.

The WarmStart Conditional Boruta method introduces a memorization loop in which variables accepted during one iteration are carried over as mandatory conditioning variables in subsequent iterations. This increases the chance of including *interacting variables* with their interacting partners within subsets, thereby improving the likelihood that *interacting variables* are correctly identified and retained.

The method is evaluated on a synthetic dataset with *interacting variables*, as well as on a real-world high-dimensional manufacturing dataset where conventional variable selection has been insufficient. Results demonstrate improved selection of *interacting variables* compared to Conditional Boruta and other baseline methods. The method is computationally intensive by nature, making it best suited for manufacturing settings where process understanding is crucial but experimentation is impossible due to time, cost or process constraints.

## Special/ Invited session

## Classification

Both methodology and application

## Keywords

Variable Selection, Interactions, Manufacturing

**Primary author:** ENGSIG, Marcus (Technical University of Denmark)

**Co-author:** KULAHCI, murat (DTU)

**Presenter:** ENGSIG, Marcus (Technical University of Denmark)

**Track Classification:** AI: Machine Learning and Predictive Analytics