



Contribution ID: 47

Type: **not specified**

## Why Order Matters for Ordinal Prediction with Random Forests

*Thursday, 16 May 2024 10:20 (20 minutes)*

Traditionally, ordinal response data have been modeled through parametric models such as the proportional odds model. More recently, popular machine learning methods such as random forest (RF) have been extended for ordinal prediction. As RF does not inherently support ordinal response data, a common approach is assigning numeric scores to the ordinal response categories and learning a regression RF model on the numeric scores instead. However, this requires the pre-specification of said numeric scores. While some approaches simply use an integer representation of the  $k$  ordinal response categories (i.e.,  $1, 2, \dots, k$ ), other methods such as Ordinal Forest (OF; Hornung, 2019) and the Ordinal Score Optimization Algorithm (OSOA; Buczak et al., 2024) have been proposed which both internally optimize the numeric scores w.r.t. the predictive performance achieved when using them. For predicting unseen observations, both OF and OSOA rely on a Transform-First-Aggregate-After (TFAA) procedure, where for each new observation numeric score predictions are generated at the tree level and transformed back into the ordinal response category. In a second step, an aggregated prediction is then obtained via majority voting. In this work, we propose a novel prediction approach, where the numeric score predictions are first aggregated into a single, combined numeric score prediction which in turn is transformed back into a categorical prediction (i.e., Aggregate-First-Transform-After; AFTA). We show that AFTA prediction can notably enhance the predictive performance of OF and OSOA. Further, we propose Border Ranger (BR), a novel RF method for ordinal prediction that reaches similar predictive performance as the AFTA prediction enhanced OF while avoiding the computationally intensive optimization procedure. We evaluate all methods on simulation and real data.

### Type of presentation

Contributed Talk

**Primary author:** BUCZAK, Philip (TU Dortmund University)**Presenter:** BUCZAK, Philip (TU Dortmund University)**Session Classification:** Contributed session**Track Classification:** Spring Meeting