

Using surrogate models and variable importance to better understand Random Forests regression fitting.

Manlio Migliorati¹, Anna Simonetto¹

¹ University of Brescia, Department Civil, Environmental, Architectural Engineering and Mathematics (DICATAM), Brescia, via Branze 43, Italy (E-mail: anna.simonetto@unibs.it)

Interpretability mechanisms helping users in better understanding machine learning models are crucial for Artificial Intelligence acceptance. In this manuscript, our experience in interpretation of random forest regression via surrogate models, i.e. models trying to replicate in an interpretable framework an original fitting difficult to understand, is reported. It is shown how, beyond classical R² analysis, adequacy of surrogate models can be assessed via variable importance analysis.

Keywords: interpretability, surrogate models, variable importance