

Interpretable Machine Learning - An Application Study Using the Munich Rent Index

Julia Brosig

Data Scientist at Sqooba Deutschland GmbH, Germany.

Abstract

Interpretable machine learning (IML) helps to understand decisions of black box models and thus improves confidence in machine learning models. To use interpretable machine learning methods, a black box model is fitted first, and on top of this model-agnostic interpretable machine learning methods are applied.

This paper analyses model-agnostic tools with regard to their global and local explainability. The methods are validated using a practical example of the estimation of the Munich rent index 2017.

In order to explain global decisions of the machine learning model, the Morris method and average marginal effects are compared. Comparison criteria are performance, available R packages or easy interpretability of results. Local methods concern one specific observation. LIME and Shapley values have been selected as local methods for analysis in this paper. The winning global and local method were then implemented and visualized in a dashboard, which can be found at <https://juliafried.shinyapps.io/MunichRentIndex/>.

In addition, the IML approach is compared with the model of the "original" Munich rent index 2017, which is based on simpler interpretable methods. This study shows that, model-agnostic methods provide explanations for machine learning models and the Munich rent index can be estimated with the IML approach. Model-agnostic interpretable machine learning offers enormous advantages because the underlying models are interchangeable and complex patterns in data can be explained globally and locally.

Keywords: *Interpretable Machine Learning; Black Box Models; Munich Rent Index; Shapley value.*
