

Lab 5: Prediction Challenge

Machine Learning for Economists, University of Cologne

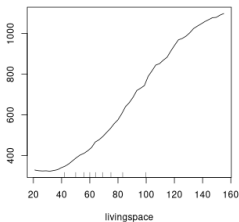
January 8, 2021

Model Approach

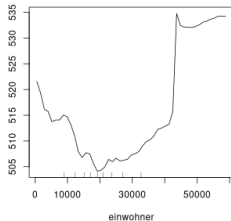
- additional variables used:
 - German federal states to differentiate between rents in East vs. West Germany
 - number of inhabitants to differentiate between rents in rural vs. urban areas
- method: Random Forest with hyperparameters determined by grid search
- procedure for hyperparameter tuning:
 - do a k-fold Cross-Validation for each hyperparameter combination
 - calculate the average RMSE
 - pick combination that minimizes the average RMSE

Estimated Model - Partial Dependences

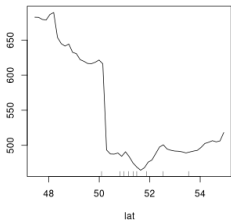
Partial Dependence on livingspace



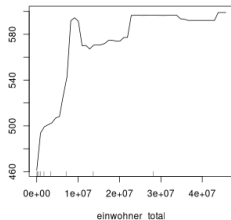
Partial Dependence on einwohner



Partial Dependence on lat



Partial Dependence on einwohner_total



Estimated Model - Variable Importance

