PATRIZIA

Can a machine understand real estate pricing?

Evaluating machine learning approaches with big data

Dr. Marcelo Cajias

ERES 2019

The author especially thanks PATRIZIA AG for contributing the data set to conduct this study. All statements of opinion reflect the current estimations of the author and do not necessarily reflect the opinion of PATRIZIA AG or its associated companies

PATRIZIA | © 2019

A simple case study to understand Machine Learning

In order to identify the value add from an institutional research perspective

The dataset

- Munich residential market
- >65k observations
- Asking rents explained by dwellings' size

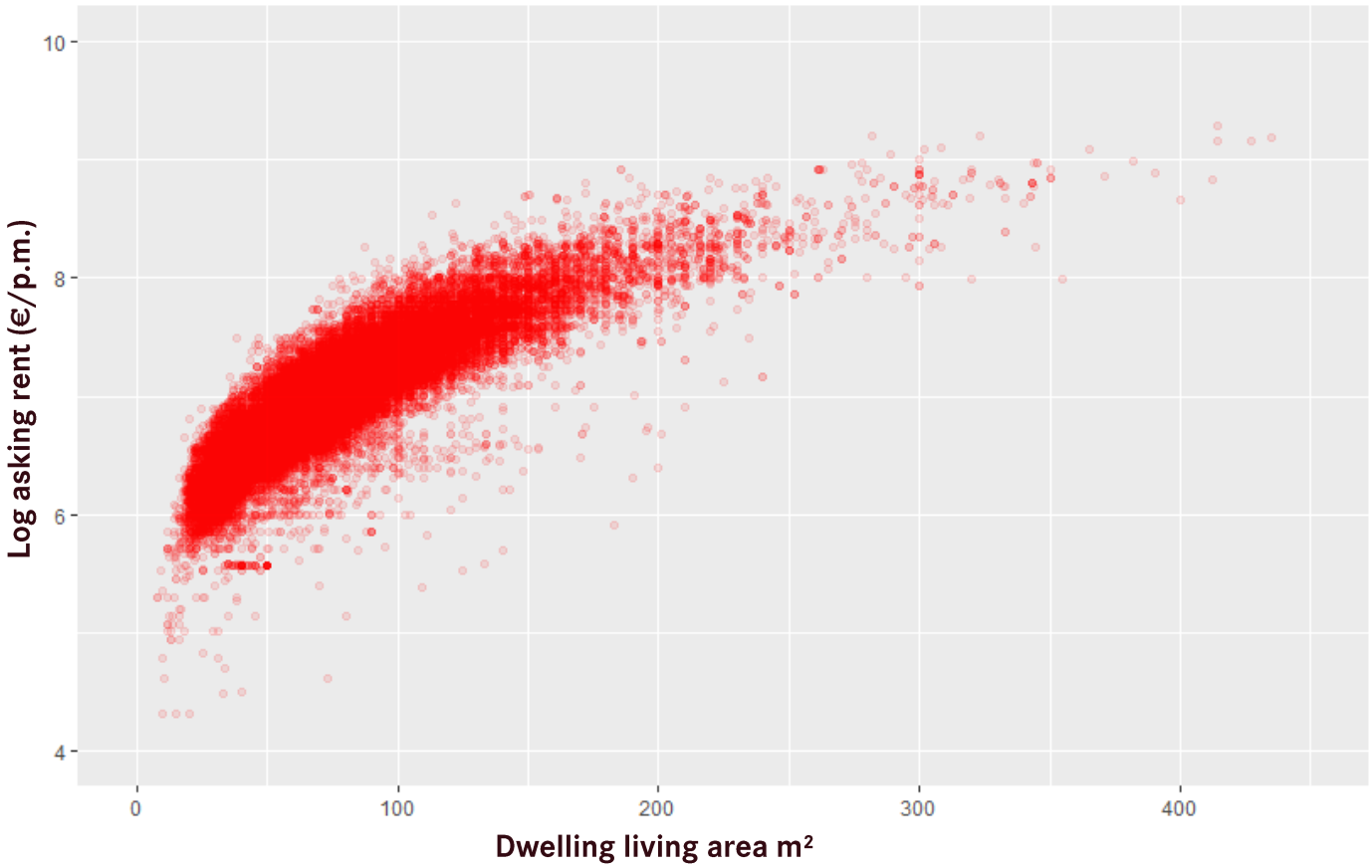
The target

Forecast asking rents with the lowest error

The methods

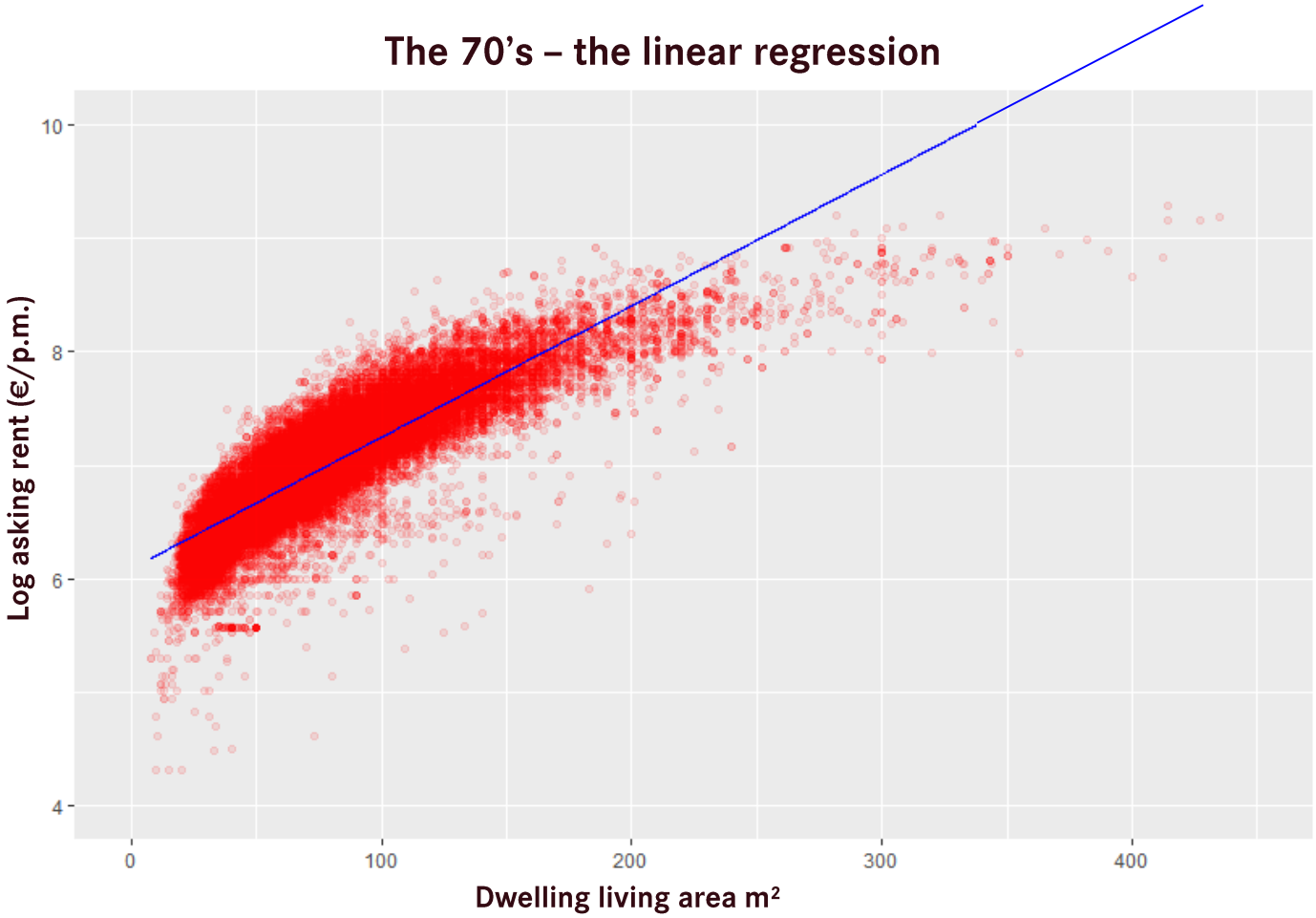
70's
90's
2000's
Nowadays = ML

The case study in a picture: log rents as a function of living size



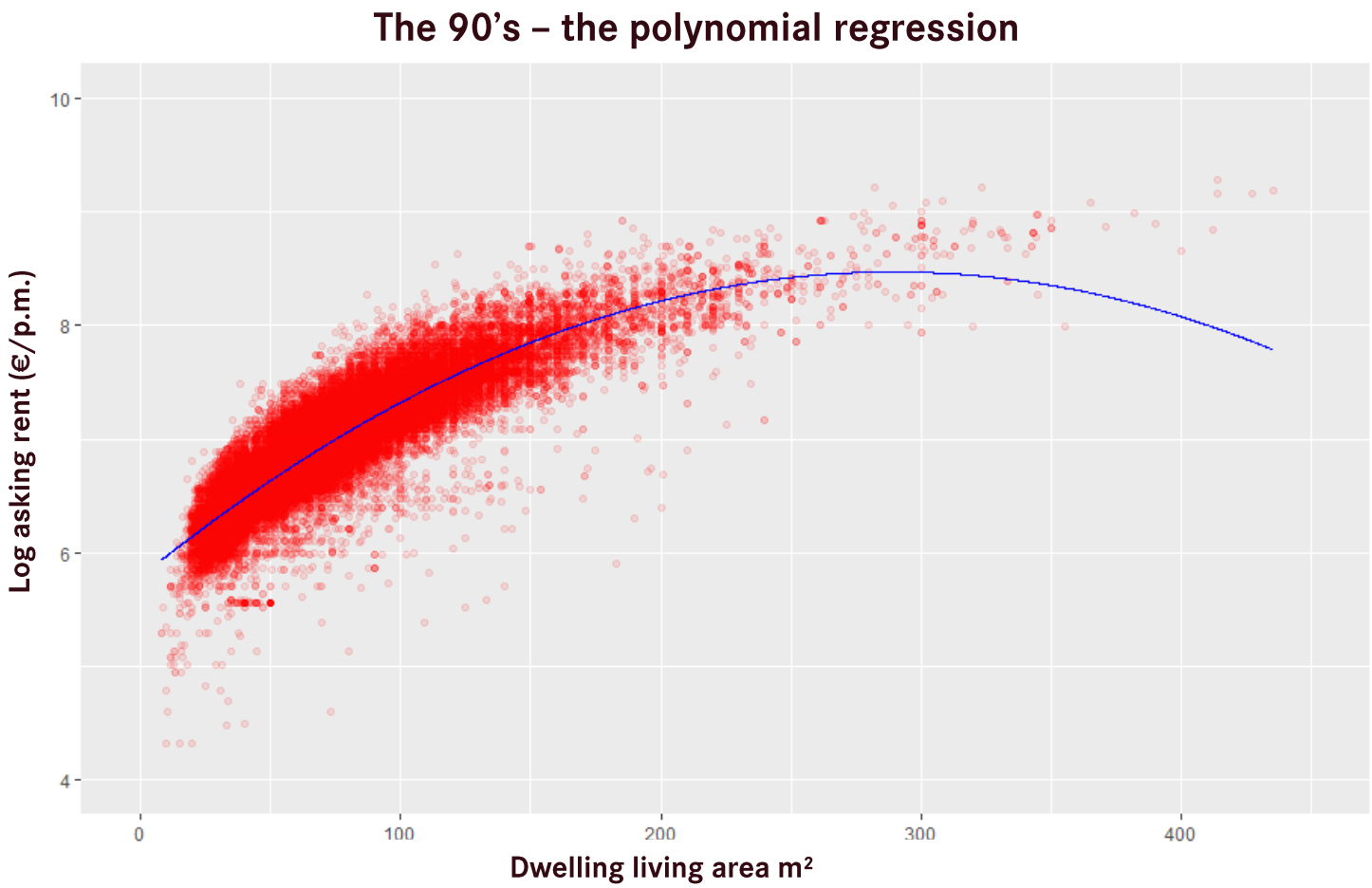
Source: PATRIZIA via Rstudio in EC2 on AWS.

The 70's – the linear regression



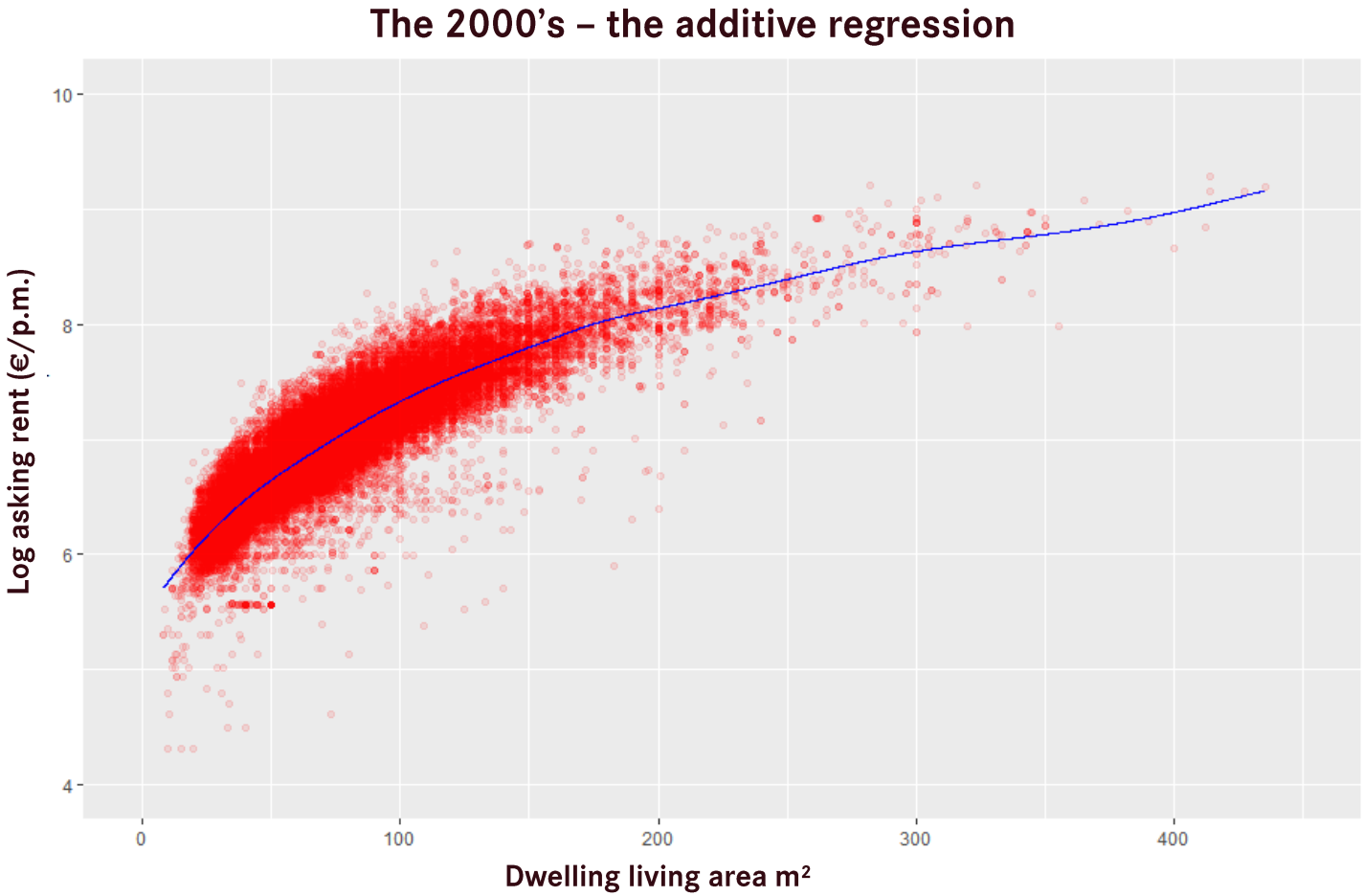
Source: PATRIZIA via Rstudio in EC2 on AWS.

The 90's – the polynomial regression



Source: PATRIZIA via Rstudio in EC2 on AWS.

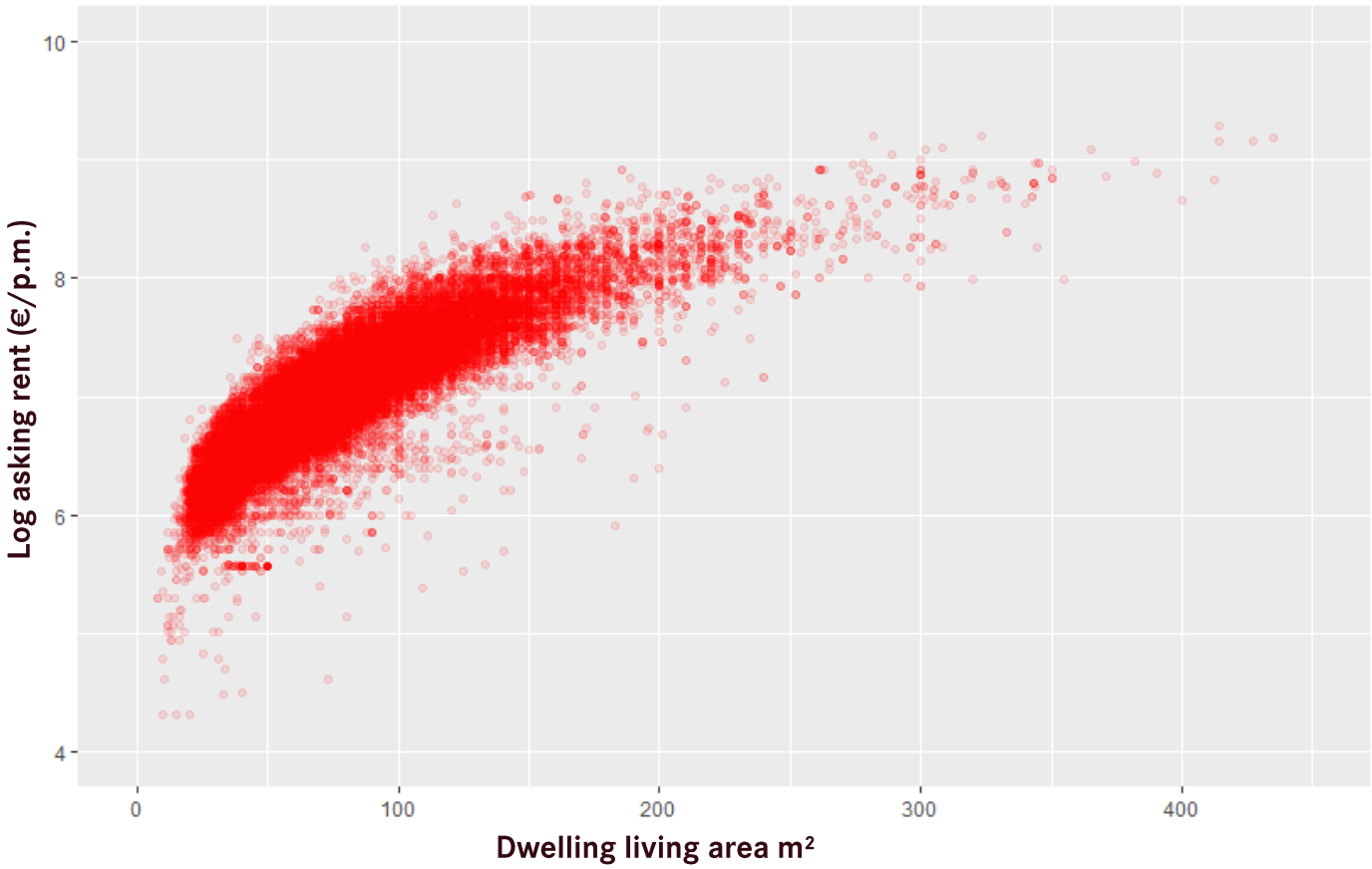
The 2000's – the additive regression



Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Learning process

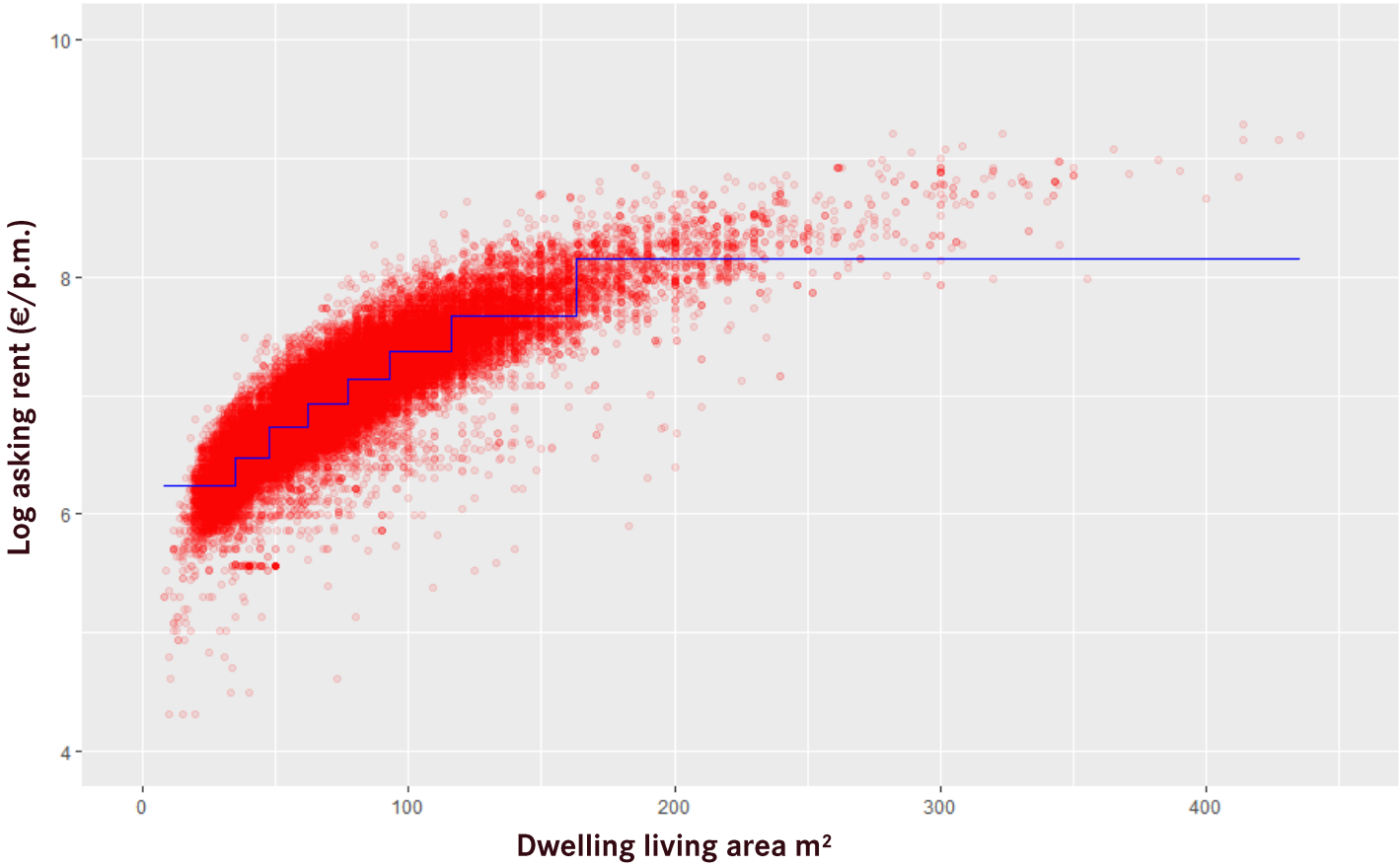


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process

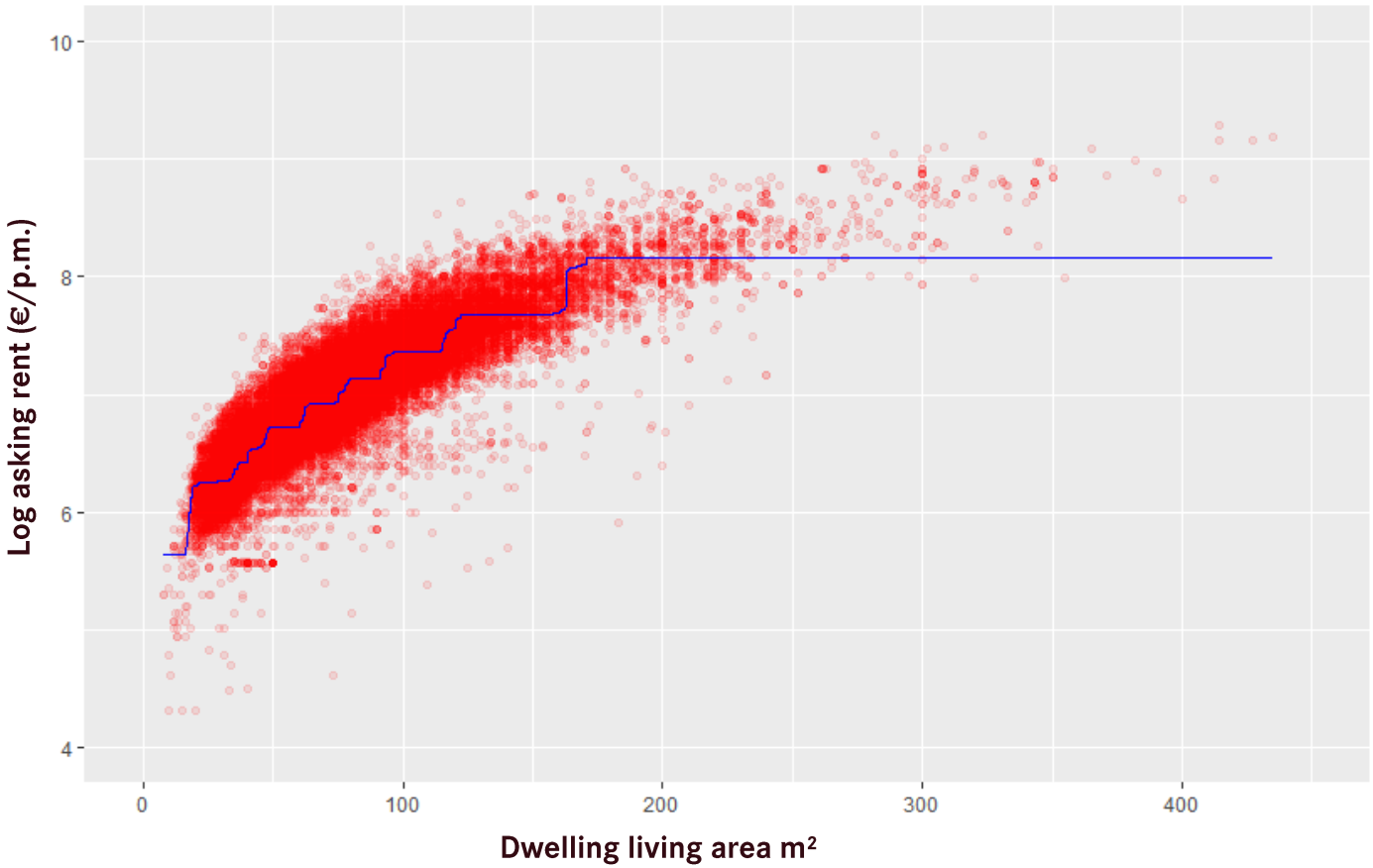


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process

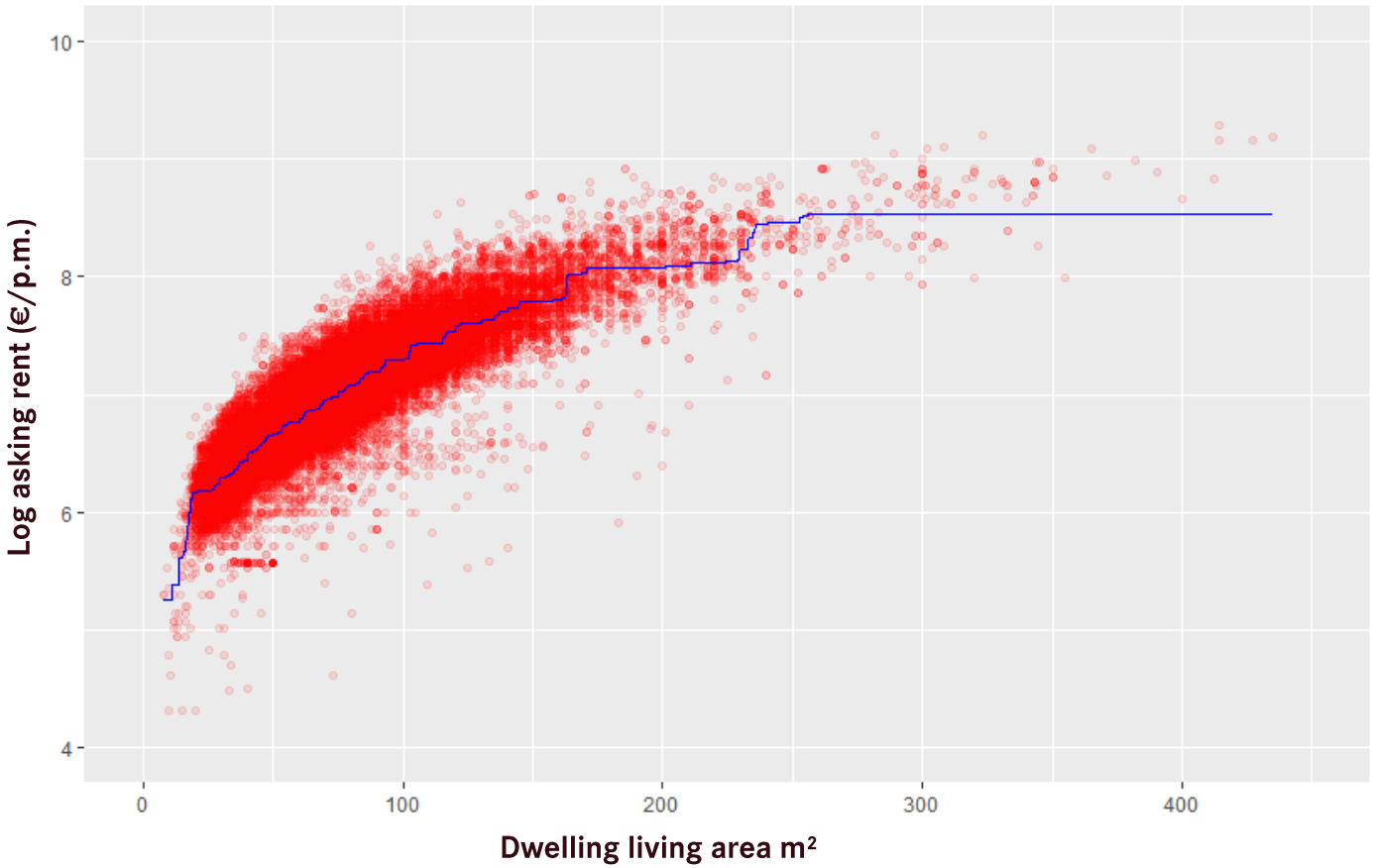


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process

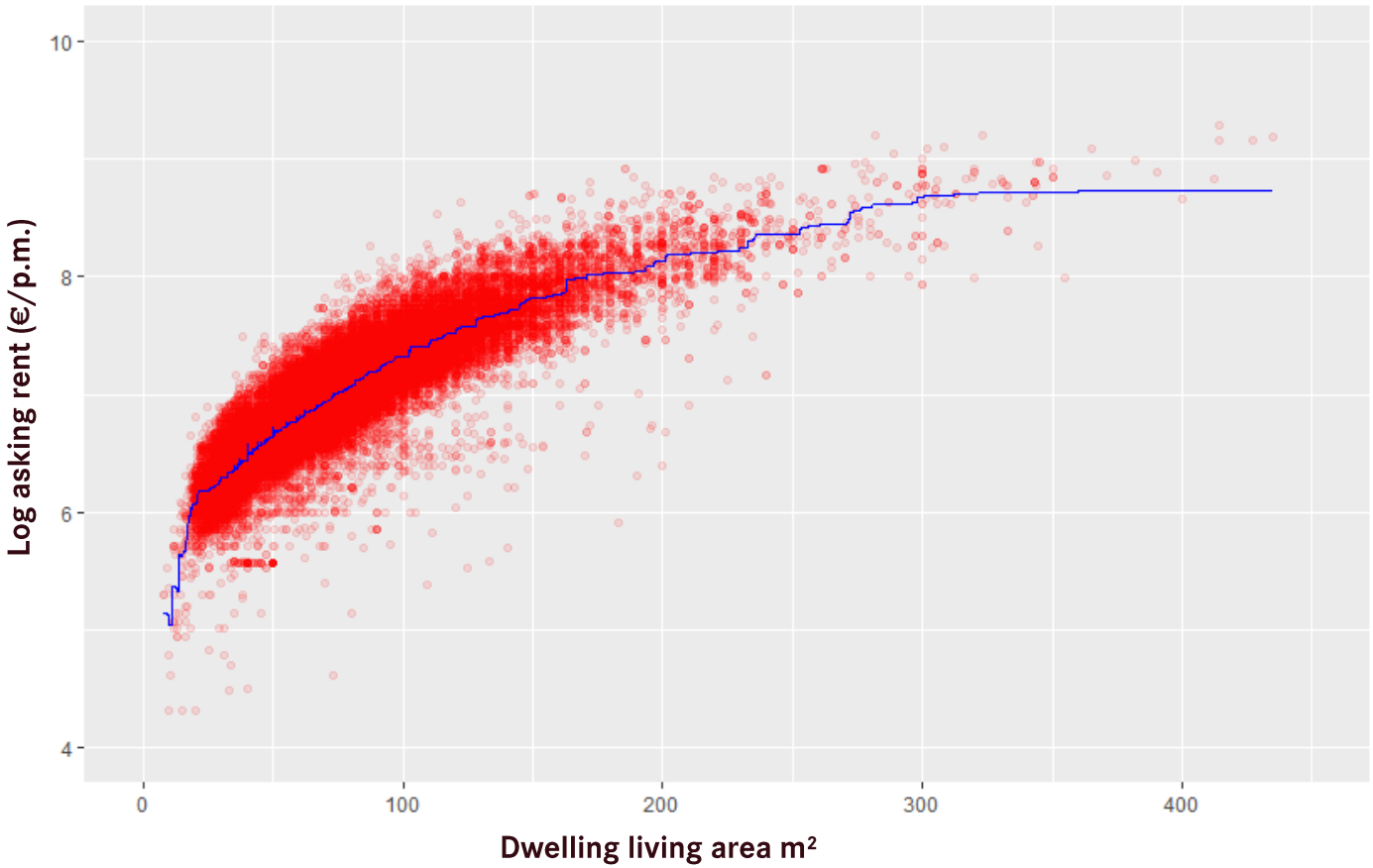


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process

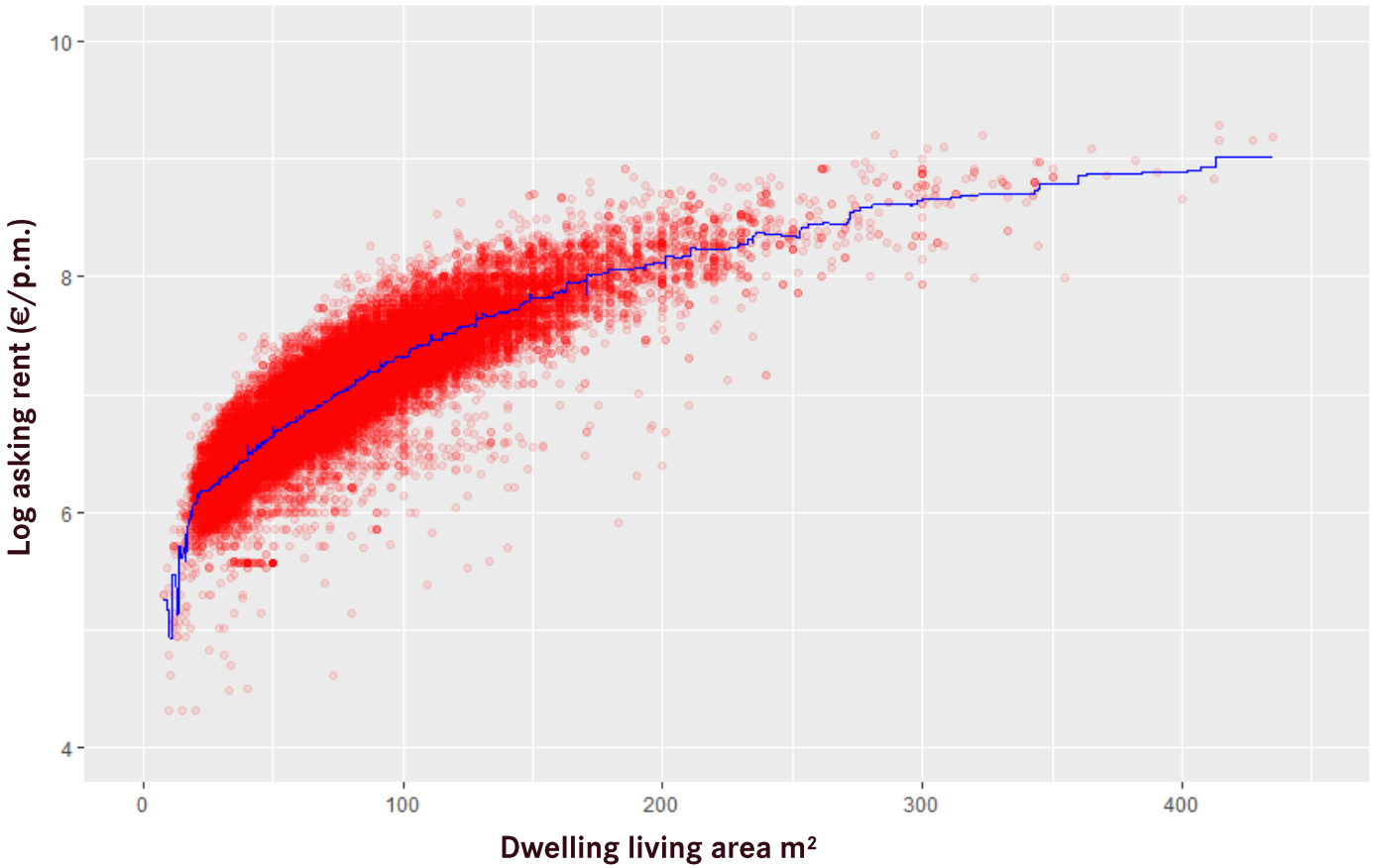


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process

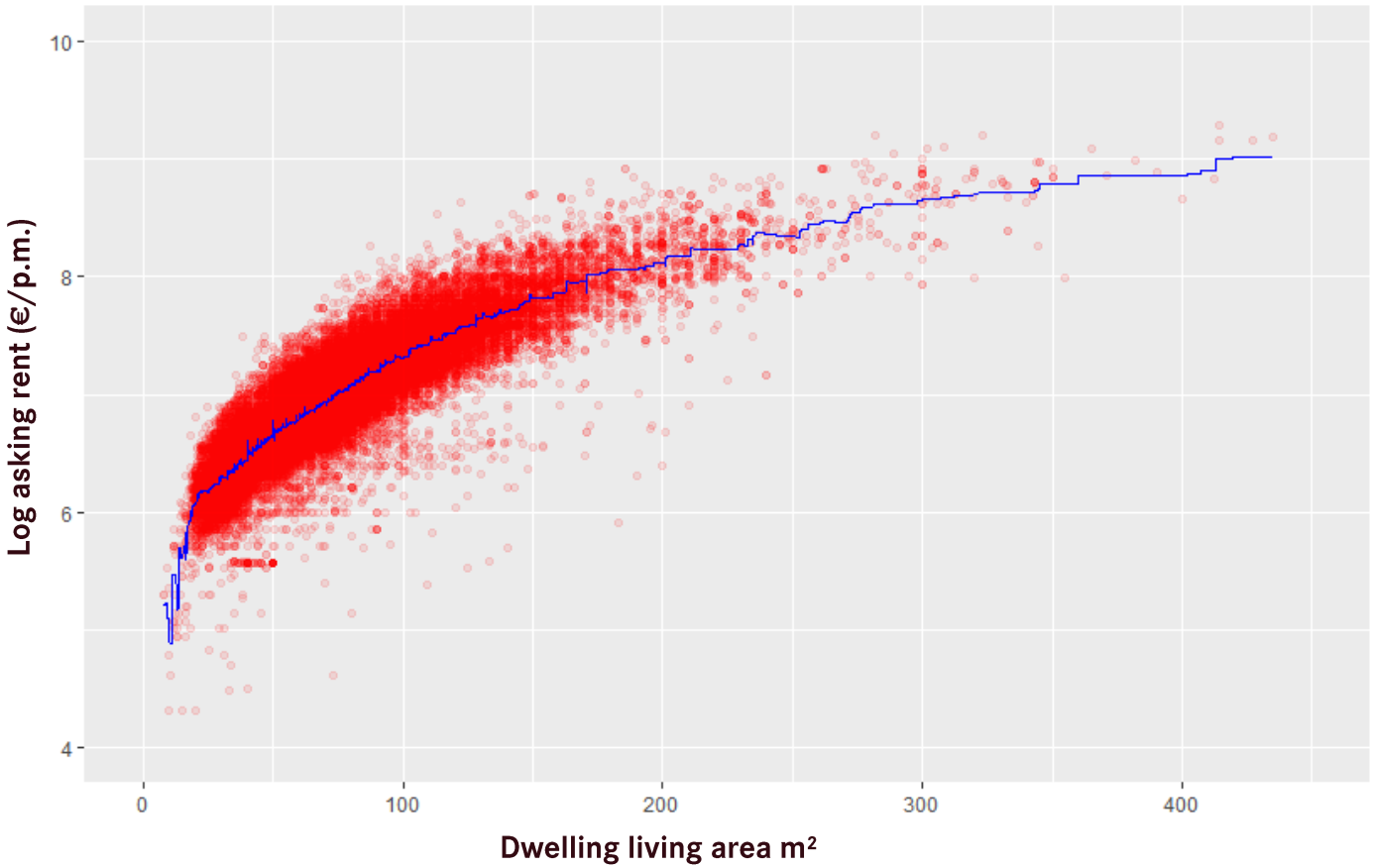


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process

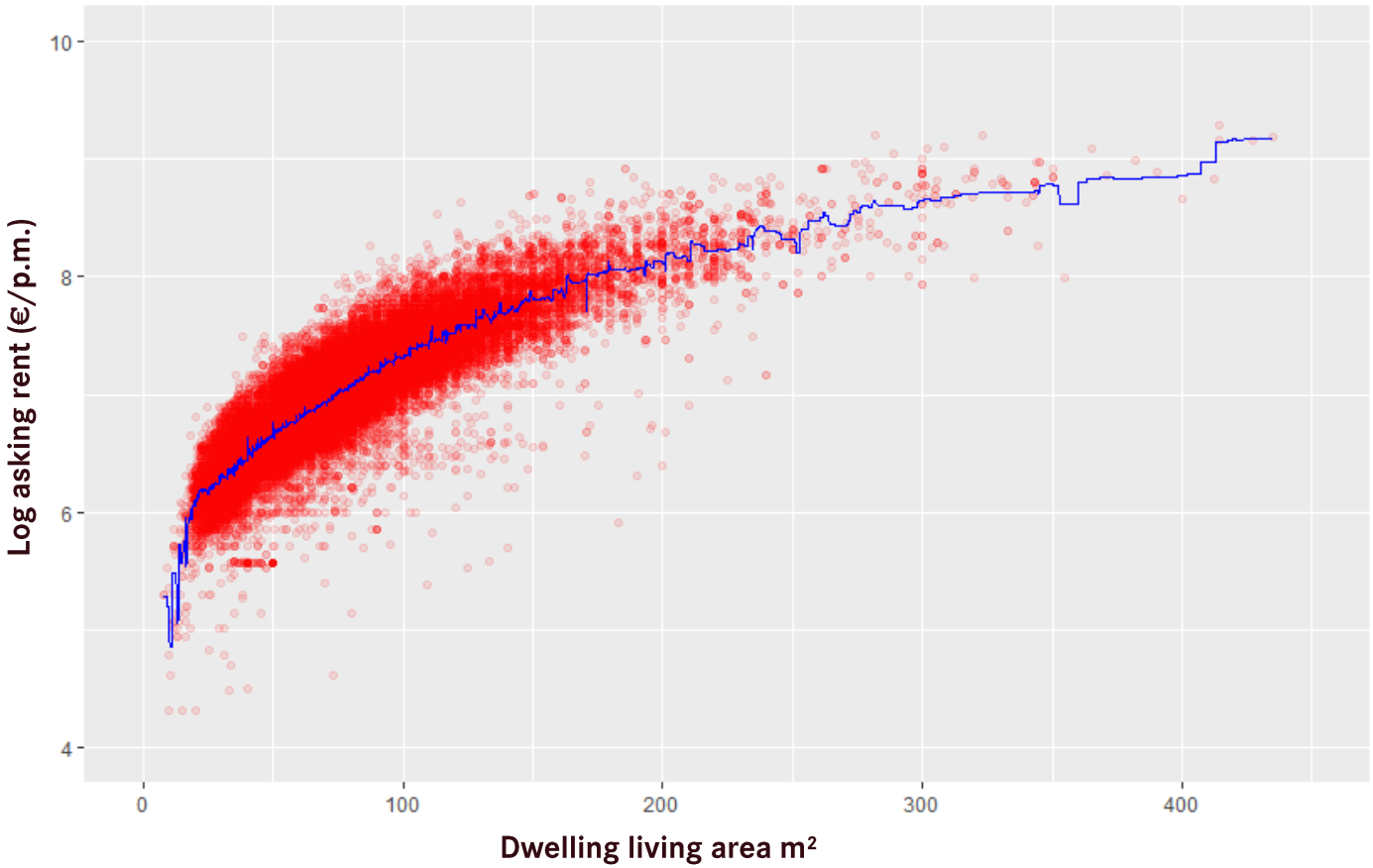


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process

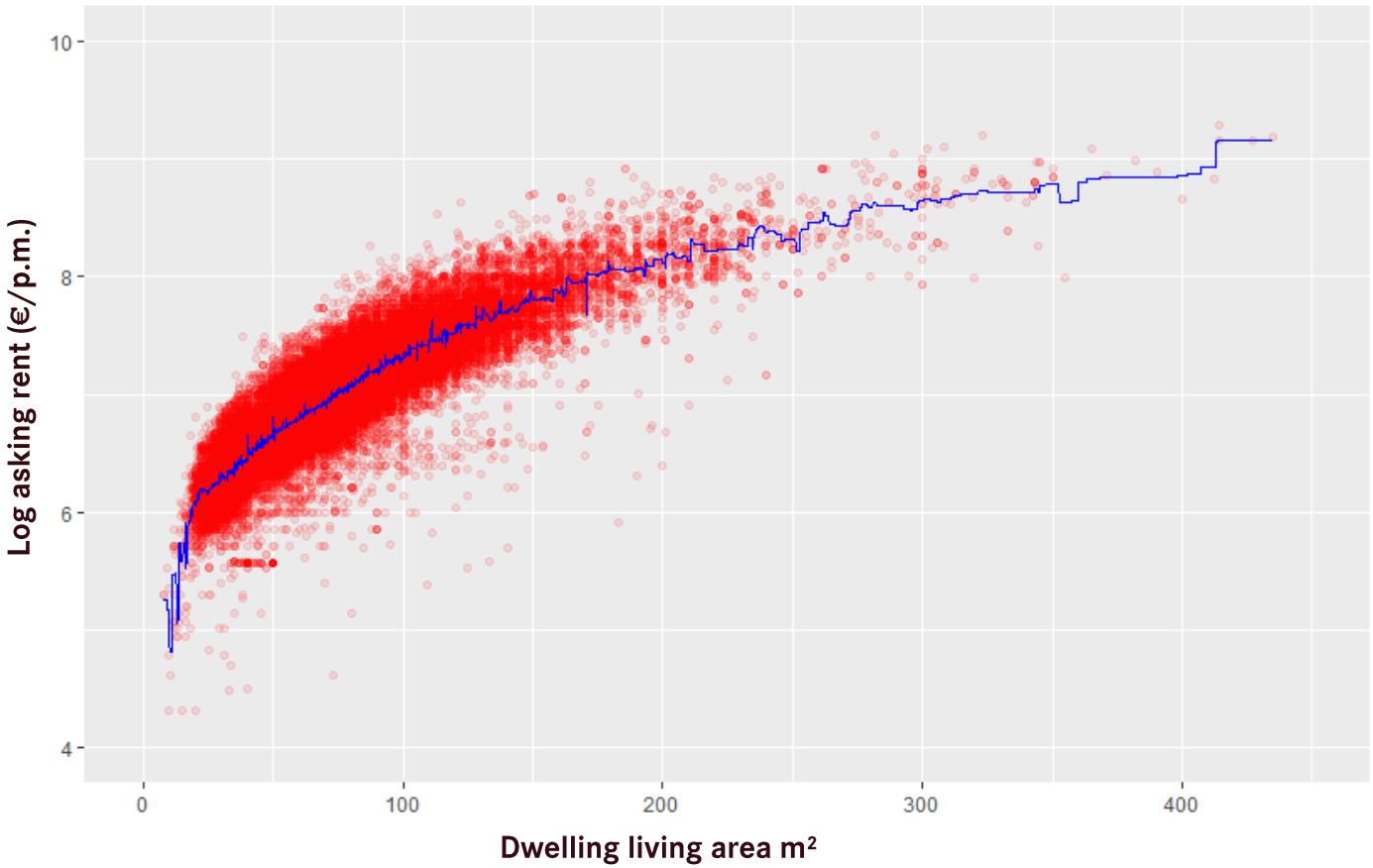


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process

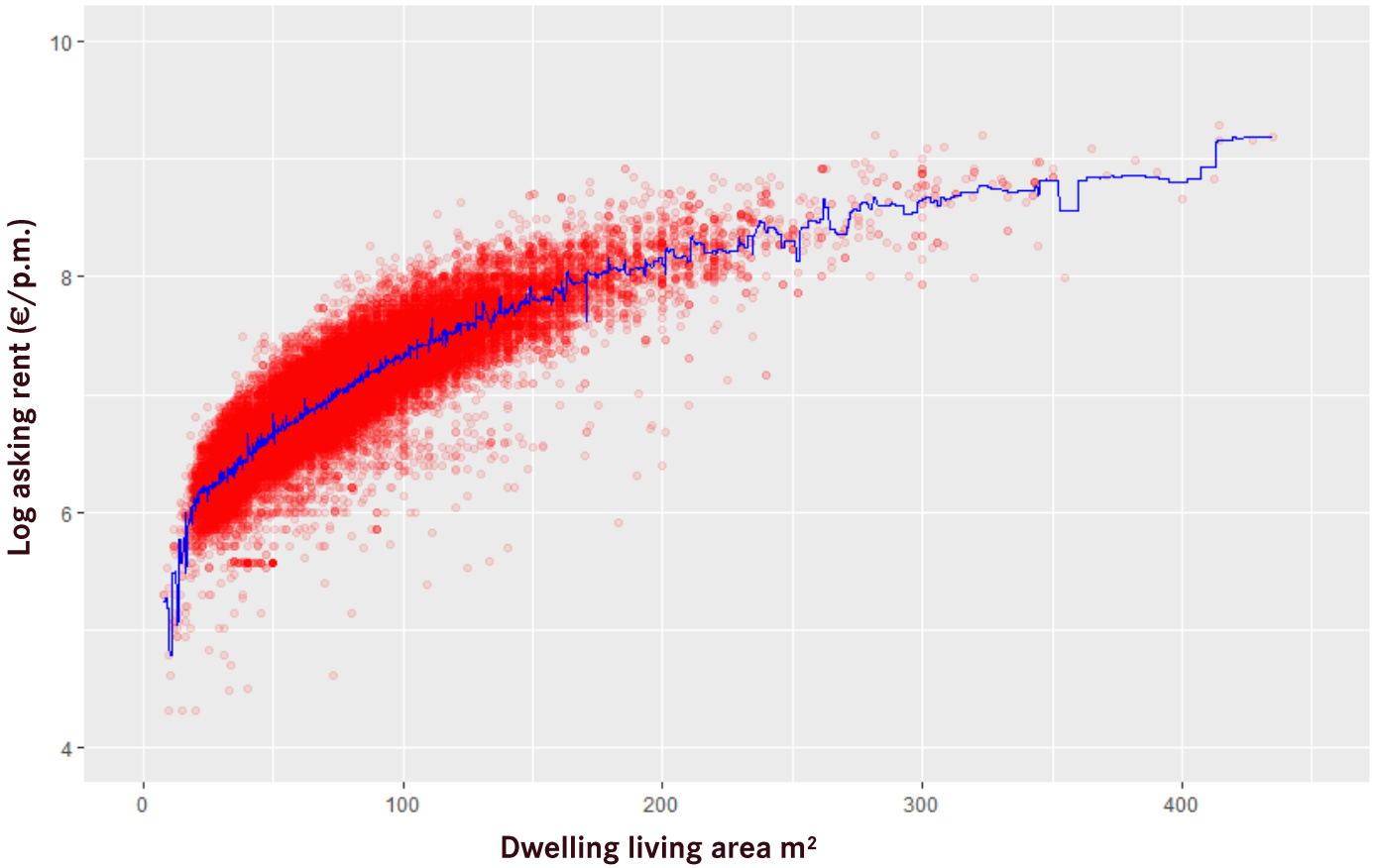


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process

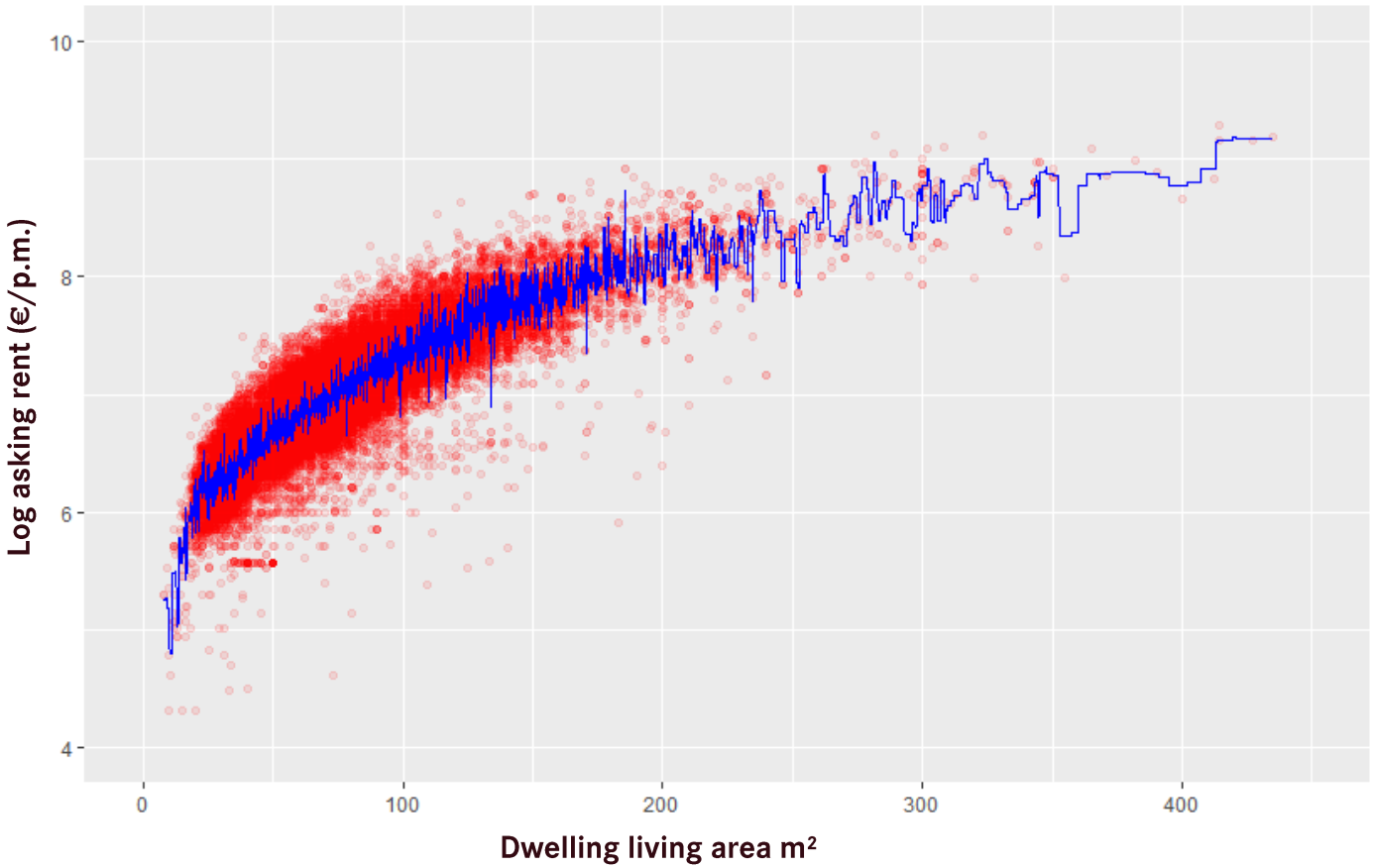


Source: PATRIZIA via Rstudio in EC2 on AWS.

Machine learning: nothing but a very very complex regression

Let the machine understand your problem

Learning process



Source: PATRIZIA via Rstudio in EC2 on AWS.

Institutional value add of machine learning for real estate

In this simplified case study

The target	The methods	The forecasting error
Forecast asking rents with the lowest error	70ies	
	90ies	836 €/p.m.
	2000ies	341 € /p.m.
	Nowadays = ML	326 € /p.m.
		291 €/p.m.

Source: PATRIZIA via Rstudio in EC2 on AWS.

Research questions and econometric approach

Research questions – qualitative

- What is ML?
- Is ML a black box?
- Are the BLUE assumptions dead?
- Where is the real estate industry along the ML-cycle?

Research questions – quantitative

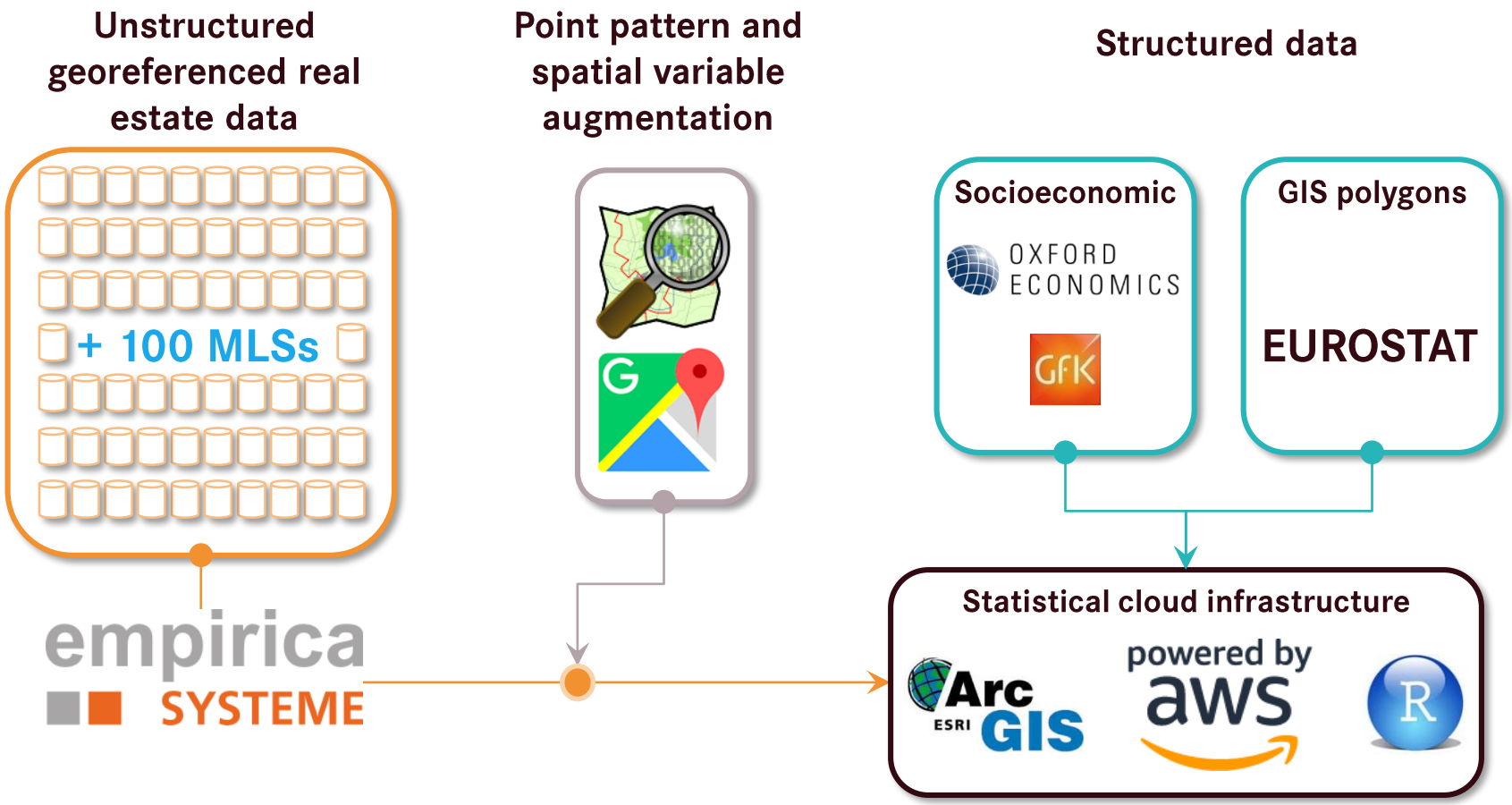
- What methods can we consider for real estate?
- How are these methods defined?
- How well do they perform in predicting asking prices?

Econometric methods

- GAMLSS
- XGboost
- Gradient Boosting Regression
- Support Vector Regression
- Random Forest Regression

Source: PATRIZIA via Rstudio in EC2 on AWS.

Big data approach: cloud infrastructure allows efficient merging and analysis of data sources



Source: PATRIZIA, based on GfK, Oxford Economics, Eurostat via Rstudio in EC2 on AWS.

Munich‘s residential market: 4 different sources of information

Dwelling specific variables

Living area, age, bathtub, built-in kitchen, parking slot, terrace, balcony, elevator

Spatial variables

Distance to 13 amenities like school, supermarket, rail station, etc. + spatial grid on latitude and longitude

Socio economic variables

Purchasing power per household and number of households on ZIP code level

Intertemporal variables

Listing date, i.e. month, year as time trends

Explanatory variable

Asking rents in log €/p.m.

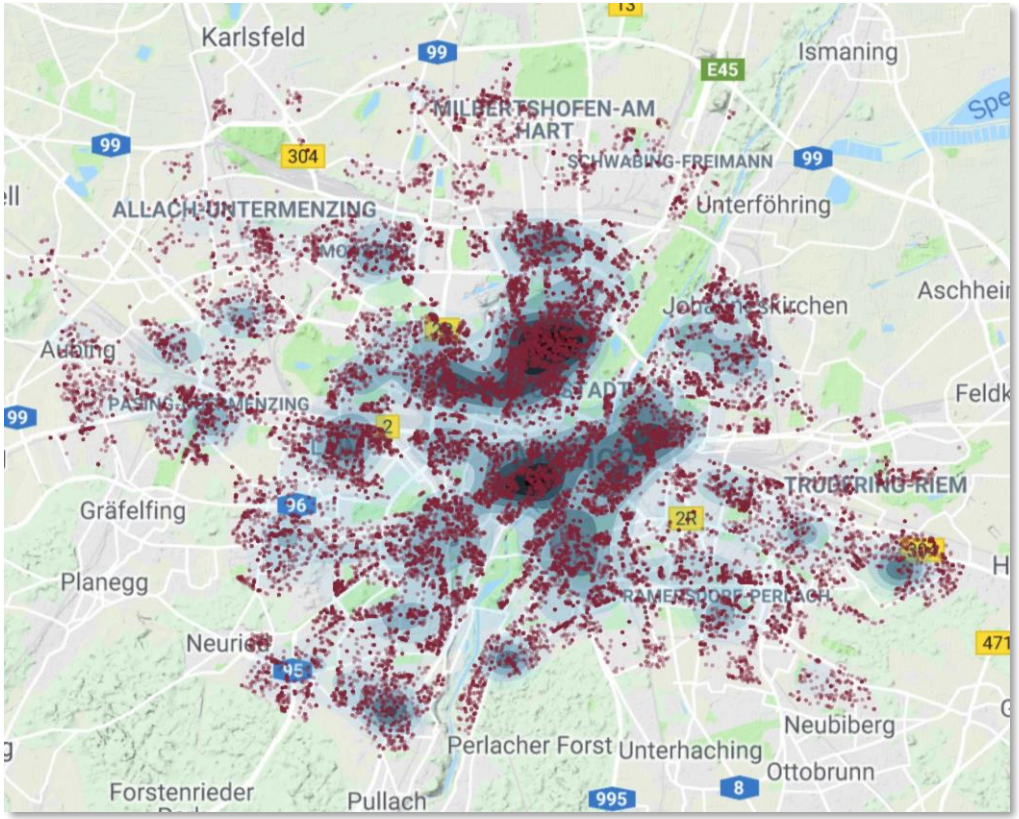
Source: PATRIZIA

Munich‘s residential market: 4 different sources of information

Explanatory variable

Asking rents in log €/p.m.

n=62,643 observations from 2013 until Q1 2019 in Munich DE

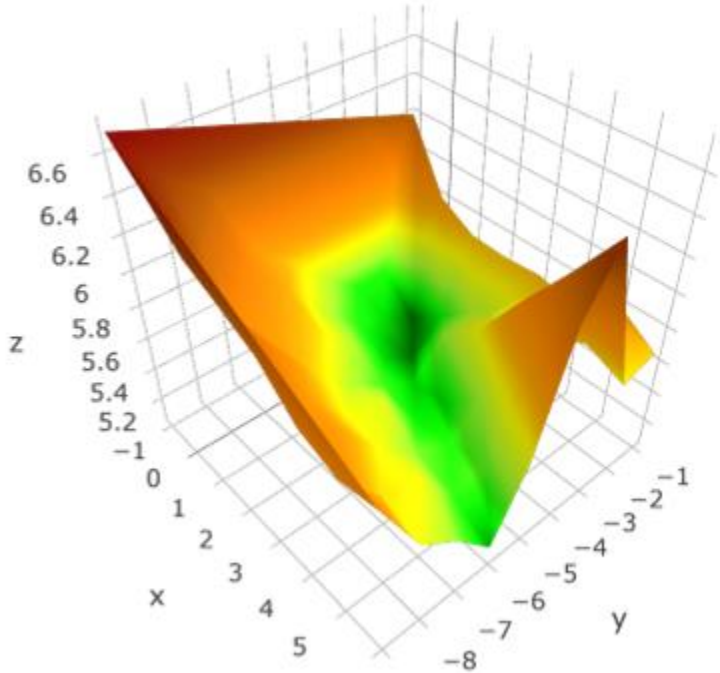


Source: PATRIZIA, Empirica, Google API, OSM, Rstudio via EC2 in AWS.

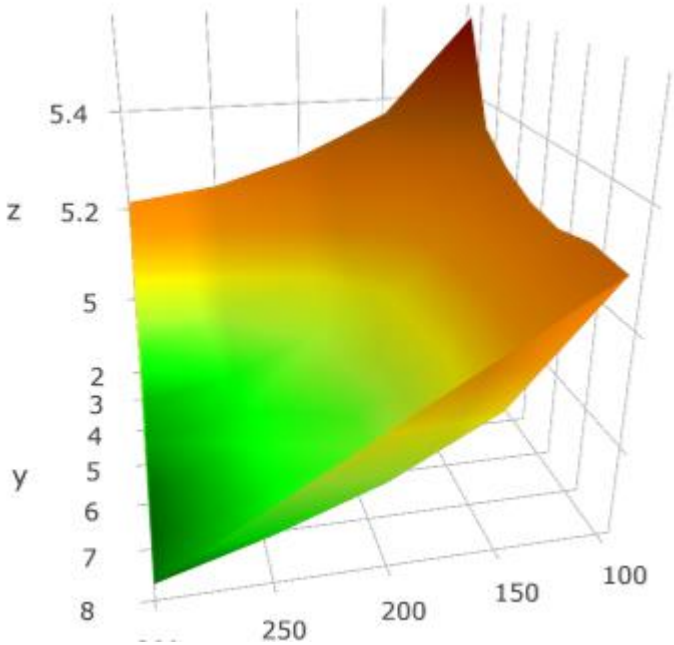
Grid search results – extract: SVR and RFR

The plots show MSE when searching the optimal hyperparameters

Support vector regression

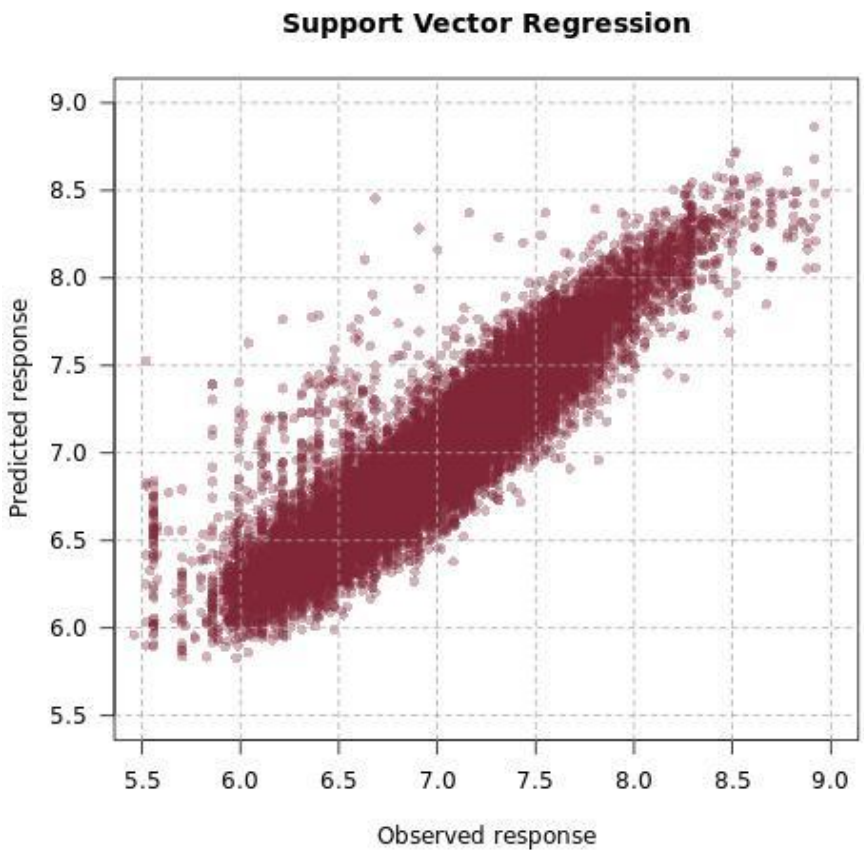


Random forest regression



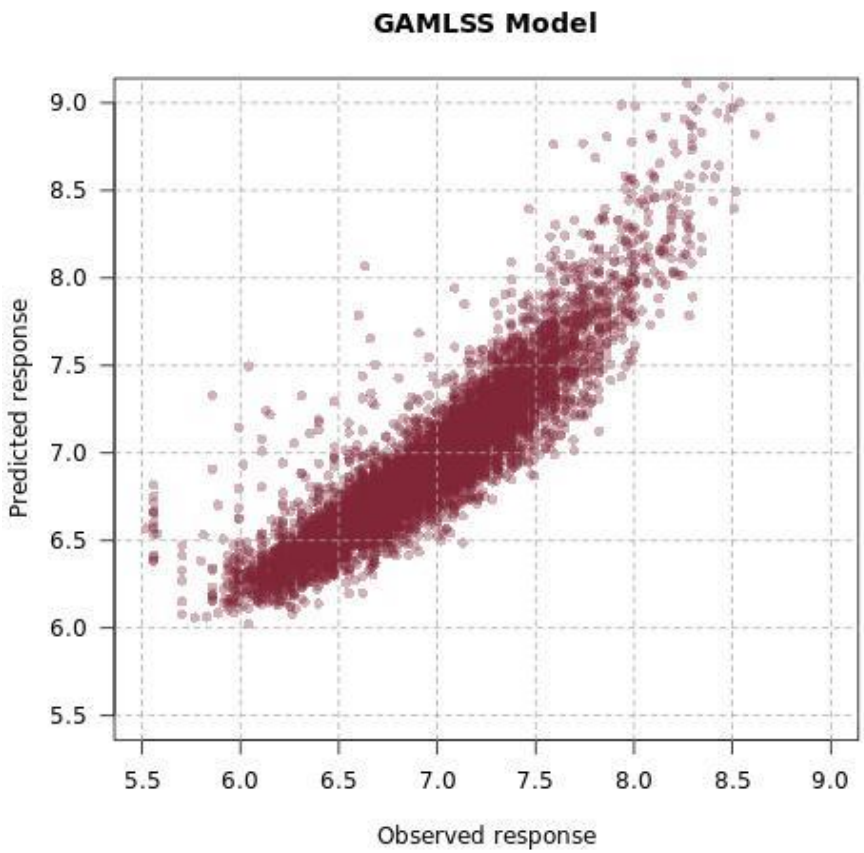
Source: PATRIZIA via Rstudio in EC2 on AWS.

Results: Support vector regression



RMSE=0.18
R^2= 85%

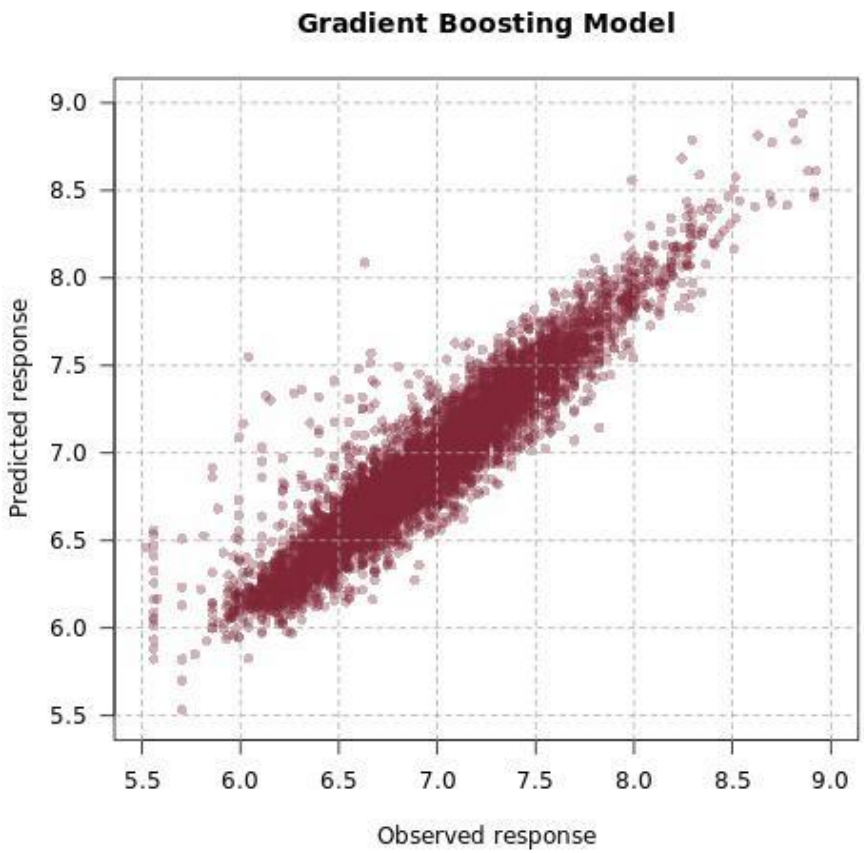
Results: GAMLSS



RMSE=0.169
R^2=87 %

Source: PATRIZIA via Rstudio in EC2 on AWS.

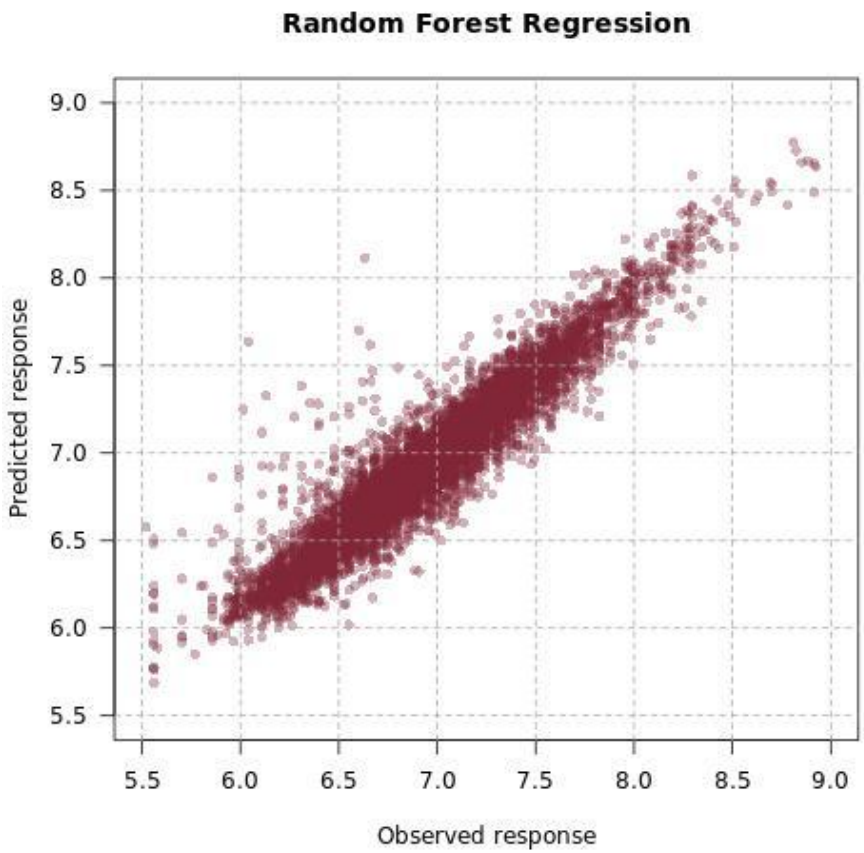
Results: Gradient boosting regression



RMSE=0.160
R^2=88 %

Source: PATRIZIA via Rstudio in EC2 on AWS.

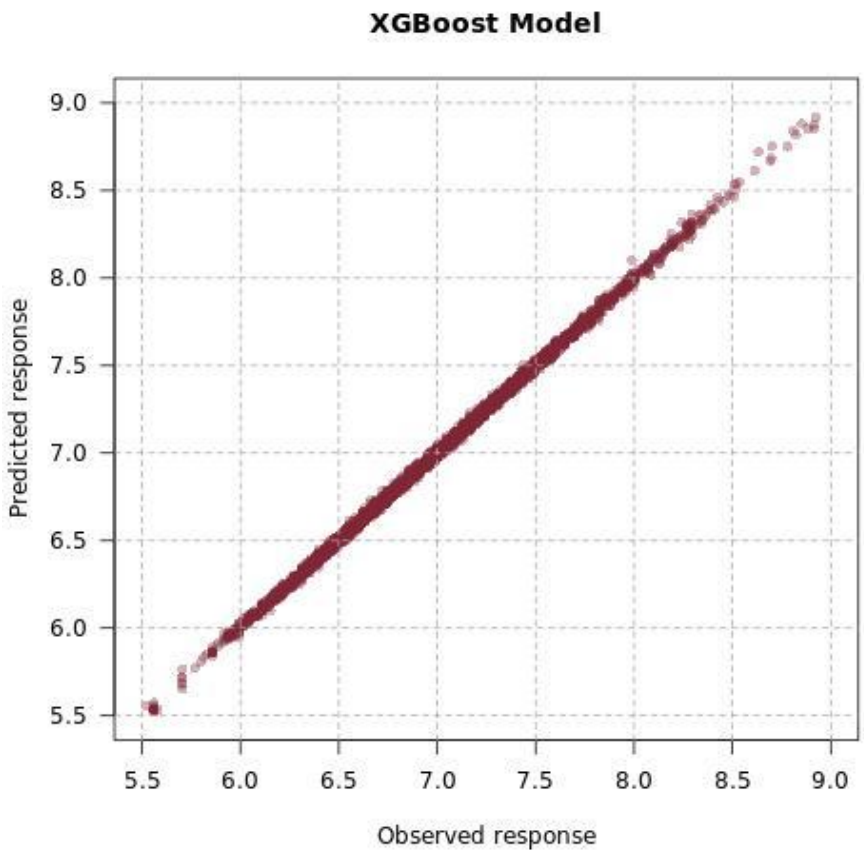
Results: Random forest regression



RMSE=0.152
R^2= 90%

Source: PATRIZIA via Rstudio in EC2 on AWS.

Results: XGBoost



RMSE=0.015
R^2= 99%

Source: PATRIZIA via Rstudio in EC2 on AWS.

Conclusion and value add for institutional investors

Machine learning methods are used for hedonic prediction and forecasting of residential rents.

The methods help investors benchmarking assets and portfolios for management decisions.

The accuracy of the methods vary but allows exact management decisions in view of the strong understanding of the data.

Model calibration and search for the optimal hyperparameter are time intensive.

Machine learning methods “came here to stay”

Source: PATRIZIA via Rstudio in EC2 on AWS.

Thanks!

By accepting this material, recipient acknowledges, understands and accepts the following:

The information contained herein is directed only at professional clients and intended solely for use by the recipient. No part of this document or the information herein may be distributed, copied or reproduced in any manner, in whole or in part, without our prior written consent. This document is for information and illustrative purposes only. It does not constitute advice, a recommendation or a solicitation of an offer to buy or sell shares or other interests, financial instruments or the underlying assets, nor does this document contain any commitment by PATRIZIA Immobilien AG or any of its affiliates. Whilst prepared in good faith, the information contained in this document does not purport to be comprehensive. PATRIZIA Immobilien AG and its affiliates provide no warranty or guarantee in relation to the information provided herein and accept no liability for any loss or damage of any kind whatsoever relating to this material. The information herein is subject to change without notice.