Cambridge University Land Society
ALUMNI

CULS Working Paper Series

No. 01 (2021)

Interpretable Machine Learning for Real Estate Market Analysis

BY

FELIX LORENZ, UNIVERSITY OF REGENSBURG

JONAS WILLWERSCH, UNIVERSITY OF REGENSBURG

MARCELO CAJIAS, PATRIZIA AG

FRANZ FUERST, CULS FELLOW, UNIVERSITY OF CAMBRIDGE

Acknowledgments: *The authors thank PATRIZIA AG for supporting this study. All statements of opinions are those of the authors and do not necessarily reflect the opinion of PATRIZIA AG or its associated companies. Franz Fuerst is grateful to the Cambridge University Land Society (CULS) for continuous support of his research.*

Abstract

While Machine Learning (ML) excels at predictive tasks, its inferential capacity is limited due to its complex non-parametric structure. This paper aims to elucidate the analytical behavior of ML through Interpretable Machine Learning (IML) in a real estate context. Using a hedonic ML approach to predict unit-level residential rents for Frankfurt, Germany, we apply a set of model-agnostic interpretation methods to decompose the rental value drivers and plot their trajectories over time. Living area and building age are the strongest predictors of rent, followed by proximity to CBD and neighborhood amenities. Our approach is able to detect the critical distances to these centers beyond which rents tend to decline more rapidly. Conversely, close proximity to hospitality facilities as well as public transport is associated with rental discounts. Overall, our results suggest that IML methods provide insights into algorithmic decision-making by illustrating the relative importance of hedonic variables and their relationship with rental prices in a dynamic perspective.

Keywords: *Interpretable Machine Learning, Microeconomic Hedonic Pricing, Housing Markets, Rental Markets*

Introduction

Possible applications of Artificial Intelligence (AI) and Machine Learning (ML) are manifold and are rapidly gaining importance across a number of domains. While most members of the general public interact with ML algorithms on a daily basis (e.g. personalized web ads, mail spam filter, etc.), there is also a growing number of discoveries and implementations in research. Recently, Deepmind and its interdisciplinary research team solved one of the biggest challenges in biology with their AI-based system AlphaFold to predict how proteins fold – a problem that has been investigated for nearly 50 years (Senior et al., 2020). Further high stake domains include arrival planning in emergency department and cancer diagnosis in healthcare (Ahmad et al., 2018) or recidivism forecasting in criminal justice (Berk & Bleich, 2013).

But how is it that these methods are only gradually coming to the fore? The high predictive performance marks ML as a promising extension for existing regression as well as classification tasks due to their ability to incorporate complex patterns and deal with large datasets. However, because the methods are often perceived as opaque, their so-called ‘black box’ character is repeatedly criticized. Certain use cases such as an AI-based decision support of credit applications may improve and accelerate business operations of banks, however the sole decision of whether a credit may be granted or denied lacks accountability and does not represent a satisfactory outcome for neither the applicant nor the creditor. Consequently, explaining the inner working of an ML model is important to justify and validate how a certain decision is made as well as to discover new insights (Adadi & Berrada, 2018).

A similar picture can be seen for the application of AI in the real estate industry. Because real estate represents one of the largest asset classes worldwide (Kok et al., 2017), an adequate estimation of real estate prices and rents are of crucial importance for investors, landlords and tenants. By treating the property as the sum of its individual characteristics, the hedonic price regression has established itself as the main approach for price and rent estimation. ML models have proven to be helpful in real estate hedonic modelling especially for predictive purposes. Nevertheless, their inferential capabilities are limited, since the aforementioned missing transparency hides the inner logic and decision making process (Mullainathan & Spiess, 2017). But how to overcome this obvious weakness? One possibility is to design models in such a way that their complexity is kept low from the beginning to ensure interpretability. An example comes from Lechner et al. (2020), who have created a deep learning algorithm that manages to control a car based on only a few artificial neurons. As a result, the decisions made by the algorithm are easy to understand while maintaining robustness and functionality. Another possibility is to examine existing ML algorithms and their results with special analysis tools in order to establish interpretability. This is where this study picks up. The ML algorithm eXtreme

Gradient Boosting (XGB) is used for a hedonic estimation of rents in the city Frankfurt am Main, Germany, and forms the basis for the application of Interpretable Machine Learning (IML) methods. Different model-agnostic tools such as feature importance and feature effects are applied to illustrate how hedonic characteristics contribute to the final prediction of the applied ML model. To the best of the authors' knowledge, this is the first real estate related study to use ex-post IML methods to justify machine-based decision-making on the one hand, and on the other hand, to gain further insights into the individual value of certain hedonic characteristics of an apartment.

Literature review

For decades, hedonic models have formed the basis for empirically assessing prices and rents of properties based on their characteristics, such as amenities or location. A hedonic model estimates the effects of these characteristics by bundling them into a function and can thus determine the price of a property. The approach is commonly used because the concept offers many possible applications for a wide variety of problems.

According to Sirmans et al. (2005), origins of the hedonic model do not go back to just one founding father. Whereas Court (1939) first used a hedonic procedure to determine automobile prices, Lancaster (1966) and Rosen (1974) paved the way for the application in real estate. Since then, a large body of literature has emerged dealing with issues surrounding the relationship between the price or rent of a property and its characteristics. Essays by Sheppard (1999), Malpezzi (2002) and Sirmans et al. (2005) provide an overview of the diversity, but also the complexity of the questions that arise within hedonic research. However, the starting point is, as so often, the underlying data set or the available features of a property. Dubin (1988) argues that building characteristics that usually determine prices in a hedonic model can be grouped into three categories: Structural, location and neighbourhood variables. Can (1992) and Stamou et al. (2017) define them as follows: Structural variables describe the nature of an apartment, such as its size, the number of rooms or the age of the property. Location variables, on the other hand, such as distance to the central business district (CBD), define the geographic location. Neighbourhood variables tie in here and illustrate the socio-economic environment such as household income or the physical make-up of the closer environment. Often, the location and neighbourhood variables are considered together, as sometimes the distinction is not evident (Can, 1992, Haider & Miller, 2000, Des Rosiers et al., 2011, Stamou et al., 2017). In the recent past, much of the focus of studies has been on the effect of these locational or neighbourhood characteristics. Within this group, variables of interest come mainly from the environmental, infrastructure and social domains. With respect to features in the immediate environment of a property, Dumm et al. (2016), Rouwendal et al. (2017) and Jauregui et al. (2019) analyse the effect of proximity to water on price. Studies by Below et al. (2015) and Dumm et al. (2018) show the price impact of nearby subsurface conditions such as sinkholes or land erosion. Other issues such as the influence of distance to urban green spaces (Conway et al., 2010) or the presence of air pollution (Fernández-Avilés et al., 2012) also receive attention. Considering the group of neighbouring infrastructural facilities and their impact on properties, different studies emerged. Hoen et al. (2015), Hoen and Atkinson-Palombo (2016) and Wyman and Mothorpe (2018) study the effects of nearby electric facilities on property prices, such as wind turbines and power lines. Availability of transportation facilities such as of a highway and rail transit are investigated by Chernobai et al. (2011), Li (2020) and Chin et al. (2020). According to Theisen and Emblem (2018) and Zheng et al. (2016), the possibility of an easy access to early childhood education and training in the form of nearby kindergarten or schools is also a price-determining factor of residential properties. There are even more exotic themes such as the influence of strip clubs (Brooks et al., 2020) or the proximity to food trucks (Freybote et al., 2017). Nevertheless, factors in the immediate social environment can also play a role. For example, Goodwin et al. (2020) find that the presence of home ownership associations has price-determining effects. Seo (2018) shows that the neighbourhood condition is similarly price determining.

When it comes to the model design, the usual hedonic approach involves a parametric, semi- or non-parametric multiple regression analysis, which uses a pooled data set of properties and their individual features. Interestingly, the development of improved computational capabilities has recently allowed other methods such as ML to complement this estimation process. While the parametric hedonic price regression approach is largely applied for inferential purposes, its potential for predictive tasks is rather limited (Pérez-Rave et al., 2019). The scope of ML methods, however, is the other way around. While inference has hardly played a role so far due to the mostly opaque algorithms, the predictive qualities of these methods are much more pronounced. ML algorithms, like gradient tree boosting (GTB) (Friedman, 2001), random forest regression (RFR) (Breiman, 2001a) and support vector regression (SVR) (Smola & Schölkopf, 2004), are capable of artificially learning from the underlying data and continuously improving their predictive performance. Hence, these algorithms have shown remarkable accuracy. In the real estate literature, various studies demonstrate the performance of ML algorithms and parametric hedonic models, including Lam et al. (2009) and Kontrimas and Verikas (2011) for SVR, Yoo et al. (2012), Antipov and Pokryshevskaya (2012) and Yao et al. (2018) for RFR and van Wezel et al. (2005) and Kok et al. (2017) for boosting methods such as GTB. Furthermore, Zurada et al. (2011), Mayer et al. (2019) and Ho et al. (2021) document the performance of different ML methods.

However, these methods are viewed critically due to their black box character (McCluskey et al., 2013), since the final result often delivers the raw prediction without letting one know how it came to the respective conclusion. As Mayer et al. (2019) state, the predictive accuracy is only achieved by reduced comprehensibility of the ML models due to its ability to artificially capture highly complex pattern within the underlying data. In consequence, researchers are mostly faced with the trade-off between what is predicted (prediction) and why the prediction took place (inference).

In general, many ML methods, such as SVR, RFR and GTB, provide model transparency since there is an understanding of how the underlying algorithm works and the algorithm can be described mathematically without further knowledge of the data – although the structure of ML methods is increasingly complex. Nevertheless, model interpretability in terms of identifying and understanding what factors impact the final predictions seems to be the bottleneck for an overall acceptance and implementation of ML methods, because sole measures like predictive accuracy are “an incomplete description of most real-world tasks” (Doshi-Velez & Kim, 2017).

In the real estate literature, first approaches have been made to combine predictive and inferential purposes within a ML context. Pérez-Rave et al. (2019) propose a variable selection approach called “incremental sample with resampling” tested on two data sets of property prices. They apply random forests to varying subsamples to predict the final property prices. Variables are identified as important, if the feature is used in the final prediction rule of the RFRs for 95% of the subsamples. The final inferential interpretation is based on a parametric hedonic model using only the ML-selected variables. Moreover, Pace and Hayunga (2020) analyse the informational content of residuals from linear, spatial hedonic regression and ML models. After applying regression trees, they find that spatial information is still present in the residuals of ML models. Although single trees are easy to understand and their decision rule can be illustrated graphically, they show limited predictive performance and tend to be unstable due to high sensitivity to changes in the data or tuning parameter.

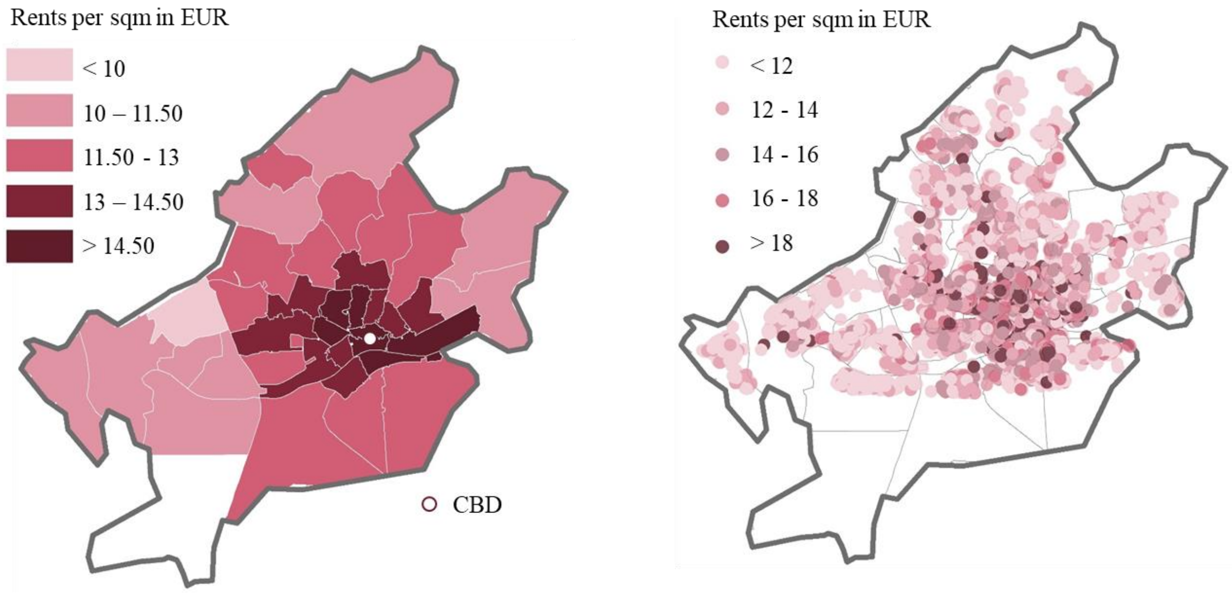
To conclude this section, this rather young field of research opens up the possibility to further engage with the interpretability of ML models and the impact of hedonic characteristics. In the following, we present the data set of our analysis and describe the methods we use to enable the interpretability of ML-based predictions. After that we discuss the results and summarize our findings in the conclusion.

Data

The sample for our analysis comprises 52,966 observations of residential rents in Frankfurt am Main, Germany. The country is the fourth largest economy worldwide and known as a safe haven for both domestic and cross-border real estate investments. With one of the lowest home ownership ratios of 51% being well below the European average, Germany is seen as a rental market rather than a homeowner market. Frankfurt represents the leading financial hub in continental Europe and is hosting the European Central Bank and the Frankfurt Stock Exchange amongst many important financial institutions. Its metropolitan region is home to more than 5.8 million inhabitants.

Rental data stems from Empirica Systeme, one of the largest German provider of real estate data, which comprises, amongst others, real estate listings of leading German Multiple Listing Systems (MLS). Data preparation and cleaning is performed to account for duplicates and erroneous data points. As the study focuses on the urban rental market in Frankfurt that is mainly determined by apartment rentals, we exclude single, semi-detached and terraced houses. We furthermore leave out student apartments, senior living accommodations, furnished co-living spaces, and short-stay apartments to control for highly specialized sub-markets that are expected to bias the overall rental market. Figure 1 provides two maps of the rental distribution in the data sample for Frankfurt. It highlights the average rent per sqm in every ZIP Code (left) and displays all observations gathered (right). Both maps indicate that the highest rents are found in the center, while lower rents tend to occur in the outskirts. There are no rental observations in the most southern part of Frankfurt due to highly forested areas and the airport of Frankfurt.

Figure 1: Distribution of rents and observations of the Frankfurt data sample



Notes: The left map shows average rents per sqm for each ZIP code. The right map depicts all observations. Both cover the Frankfurt city area from 2013 to 2019. The thin grey lines display the ZIP codes.

Besides the rent as target variable, the data contain information on structural characteristics in terms of living area, building age, floor and whether a kitchen, parking spot, balcony, terrace, bathtub and elevator is present or whether an apartment is refurbished. We add socio-economic data from Growth from Knowledge, Germany’s largest market research institute. Since all rental data points are georeferenced, we are able to add a spatial gravity layer based on data from Eurostat, the German statistical office and Open Street Map to account for spatial information and therefore add several location variables. We include the distance to the CBD as well as to numerous important amenities. Proximity to bus and railway station account for public transport and accessibility. Bakery, supermarket, convenience and department store distances comprice the local supply. Bar, beer garden and café represent the access to hospitality. While distances to school and park allow insights on public amenities, proximity to car wash and traffic signal incorporate adverse effects mainly due to noise emissions.

MLS are frequently used in German rental markets from professional as well as from private landlords. Moreover, since neither landlords nor tenants are obliged to disclose contract information in Germany, listing data is the main source of information for both researchers and practitioners.[1] In addition, it should be noted that rental price formation in major German cities is generally dominated by the offering party since residential vacancy rates in metropolitan areas are remarkably low.[2] A look at individual renting scenarios reveals that a landlord regularly receives inquiries in the double-digit range for an apartment that has been advertised. In consequence, the rental decision is not based on auction procedures but rather on timely application and best (personal and solvent) fit for the landlord. In the literature, Cajias and Freudenreich (2018) demonstrate that German residential markets are subject to low Time-on-Market and diminishing degrees of overpricing. As Gröbel (2019) suggests, asking data in Germany "reflect the currently prevailing overall market situation". Although we do not claim that rental listing precisely reflect the agreed contract rent, we expect the listing rents to be a useful framework for the ongoing analysis.

[1] See e.g. Gröbel and Thomschke (2018) using German rental listing prices in research as well as well-established applications of listing data e.g. F+B Residential Index or Empirica Real Estate Index in practice.

[2] According to CBRE, the vacancy rate for residential real estate in the city of Frankfurt am Main marks 0.4% of the stock. Moreover, Immobilienscout 24, the leading online listing platform for real estate in Germany, reports 198 clicks on average for an online apartment advertisement.

Table 1: Descriptive Statistics of the dataset for Frankfurt am Main (2013 – 2019)

	Unit	Mean	Median	Std.Dev
Rent	EUR/month	1,036.123	884	638.175
Living area	sqm	78.175	72	36.688
Floors	Integer	2.396	2	2.328
Age (relative to 2017)	Integer	49.377	48	39.701
Bathtub	Binary	0.564	1	0.496
Refurbished	Binary	0.242	0	0.428
Built-in kitchen	Binary	0.688	1	0.463
Balcony	Binary	0.633	1	0.482
Parking	Binary	0.487	0	0.500
Elevator	Binary	0.449	0	0.497
Terrace	Binary	0.136	0	0.342
Purchasing Power	EUR/HH/ZIP	50,390	49,993	5,798
CBD_distance	Km.	3.616	3.604	1.896
Bar_distance	Km.	0.722	0.511	0.636
Beergarden_distance	Km.	1.135	0.937	0.759
Cafe_distance	Km.	0.346	0.240	0.325
Bakery_distance	Km.	0.370	0.245	0.403
Convenience store_distance	Km.	0.849	0.589	0.748
Department store_distance	Km.	1.550	1.306	0.997
Supermarket_distance	Km.	0.252	0.223	0.167
Bus station_distance	Km.	3.062	2.667	1.566
Railway station_distance	Km.	0.835	0.581	0.685
Traffic signals_distance	Km.	0.186	0.157	0.135
Car wash_distance	Km.	1.266	1.234	0.584
Park_distance	Km.	0.266	0.236	0.158
School_distance	Km.	0.302	0.278	0.167

Notes: The table reports the summary statistics comprising data as of January 2013 to December 2019. Age is calculated as the difference of the building age to the year 2017. All distance variables are calculated as the distance to the specific dwelling in kilometers. Binary variables report whether the dwelling includes a certain characteristic (1) or not (0). Rent is presented as euro per month. Information on households (HH) is reported on ZIP level. SD: standard deviation, Min: minimum value, Max: maximum value.

Table 1 shows the descriptive statistics. We find a mean asking rent of 1,036.12 EUR p.m. (euros per month). An average apartment is 78.175 sqm located on the 2nd floor in a property that was built in 1968. The apartment contains a bathtub, a built-in-kitchen, a balcony, but neither a parking slot nor an elevator. On average, it is 3.62 km away from the CBD, 350 meters to the nearest café and 250 meters to the closest supermarket. The bus and railways station are 3 km and 0.84 km away, whereas the nearest school is located 300 meters nearby. The mean household purchasing power amounts to 50,390 EUR p.m. [3]

[3] In Appendix 3, we provide a full set of correlation coefficients for all variables.

Methodology

ML has proven its predictive power in the literature and is commonly used by real estate professionals to inform their decision making (RICS, 2017). We apply a tree-based approach to build the foundation for further analysis. As Pace and Hayunga (2020) state, a regression tree (RT) is easy-to-understand while still being capable of identifying complex pattern. That is because trees can capture non-linear relationships as well as interactions. In its core, a RT can be understood as nested if-else conditions. Tree-based models divide the data in distinct subsets and make a prediction for every subset (which usually is the average outcome of all observations in the specific subset). The division is made by several splitting steps, in which iteratively a feature variable is chosen and its feature space is split in a way that a certain criterion is affected most (e.g. the prediction error is reduced most) until a stopping point is reached.

Since single trees are prone to misspecification, ensembles are used to aggregate and combine the prediction rule of multiple trees. We choose XGB as an ensemble boosting method, which has shown to be capable of accurately predicting property prices and rents and at the same time yield robust estimation results.[4] Developed by Chen and Guestrin (2016), it is a promising approach for regression, as well as for classification, as it contains specific features that won it several Kaggle[5] competitions in the recent past. In its basic concept, boosting fits an initial tree, calculates the residuals of the initial prediction, and fits another tree on the residuals to stepwise reduce the prediction error and incrementally enhance the final prediction rule. To prevent overfitting cross-validation is applied.

Because the internal logic and consequently the rationale behind the individual predictions is rather hidden, the use of ML often lacks transparency. In consequence, a growing body of

[4] In general, tree-based ensemble algorithms are based on two different approaches, namely boosting and bagging. See e.g. Hastie et al. (2009) for a more detailed introduction to the fundamentals of ML models.

[5] Kaggle is one of the leading online platforms for the data science community and regularly hosts data competitions. For further information see https://www.kaggle.com

literature on IML[6] has evolved in recent years to further ‘improve trust’ in algorithmic decisions (See e.g. Adadi & Berrada, 2018; Carvalho et al., 2019; Arrieta et al., 2020 or Linardatos et al., 2021). In general, tree-based ML methods show some sort of algorithmic transparency, since their underlying concept and theory is comprehensible and mathematically described (James et al., 2013). Nevertheless, it is not evident, which feature[7] and to what extent it contributes to the prediction.

One possibility to understand how predictions are achieved in this context is to use **interpretable ML models.**[8] Like in parametric models, specific restrictions limit the complexity of the model and therefore allow inferential insights. RTs are a well-known example of interpretable ML models if e.g. the depth of the tree is limited. As Molnar (2020) states, short trees with a depth up to three splits are interpretable in a comprehensive way, since a maximum combination of three if-else-conditions as the decision rule is enough to explain how the model yield a certain prediction.

Limiting the models complexity often results in depriving ML much of its effect, since their flexible structure enables a strong predictive performance (Breiman, 2001b)[9]. Consequently, (post-hoc) model-agnostic **interpretation methods** have been developed, which separate the explanatory framework and the ML model, thus preserving its predictive capabilities. In contrast to interpretable models, the ML model remains a black box, with the separated interpretation methods aiming at extracting interpretable information post-hoc. Model-agnostic tools benefit from their flexibility because they do not depend on a specific ML method and can be applied to various learners (Ribeiro et al., 2016).

[6] In the context of IML, the term Explainable Artificial Intelligence (XAI) is often used synonymously.
[7] To describe the covariates, hedonic literature mainly refers to them as variables or characteristics, while research on IML generally uses the term features.
[8] Interpretable ML models are also referred to as transparent models, since they are considered to be understandable by itself.
[9] See e.g. Shmueli (2010) for further discussion on the trade-off between model accuracy and interpretability.

Interpretation methods differ on whether their focus is on feature importance or feature effects. The first one aims at evaluating which feature contributes the most to the prediction, whereas the second one sheds light on how a single feature contributes to the prediction. The methods are perceived as typical and useful tools to show the impact of features in ML models and explain the inner working on a global level (Hastie et al., 2009). We use the FeatureEffect and FeatureImp functions both implemented in the iml package in R (R Core Team, 2020).

Feature importance (FI) measures the relevance of a single feature for the prediction. The importance of a feature is calculated by permutation of the observed feature values and its effect on the prediction error, keeping all other features constant. Based on the concept of Breiman (2001a) for random forests, Fisher et al. (2019) provides a model-agnostic framework for measuring the covariates contribution to the accuracy of an ML model called 'model reliance'.

Let X be the feature matrix, Y the dependent variable and f the ML model, with the prediction error e being measured by a loss function $L(Y, f(X))$. The feature importance is defined as the ratio of the model error after permutation to the original model error before switching features.

$$FI(f) = \frac{e_{perm}(f)}{e_{orig}(f)} \tag{1}$$

The permutated error is thereby calculated as the expected error of the ML model based on the permuted feature matrix X_{perm}.

$$e_{perm}(f) = \mathrm{E}L(Y, f(X_{perm})) \tag{2}$$

To visualize the most important features, every variable is ranked and plotted according to their FI. Alternatively, the FI score can also be calculated as the difference of both errors, although the ratio provides the advantage of higher comparability. We use the Mean Absolute Error (MAE) as loss function. By switching the feature values of all observations (e.g. an observation with 1 for a kitchen being present is switched to 0), FI calculates how much this change leads to an observable decrease in prediction accuracy. It can consequently identify whether the specific feature contributes to the overall prediction or whether its change does not perceptibly affect the outcome. Lastly, we average the importance measures over 100 repeated permutations. As Fisher et al. (2019) states, FI is a helpful tool to identify influential features and increase the transparency of black box models.

In addition to the individual importance, **feature effects** show how a single feature influences the predicted outcome of an ML model. After the training process, a ML model has learned a specific relationship between the covariates and the target variable that can be analysed. Partial Dependence (PD) plots visualize the marginal effects of features on the model's prediction (Friedman, 2001). The plots are based on partial dependence functions which highlight the effect of one feature on the target variable when the average effects of all other features are accounted for. PD plots reveal useful information e.g. whether the relationship can be explained linearly or in a more complex manner.

Let once again X_j be the vector of the *j* variables and *n* be the number of observations. The PD is the effect of features of a subset X_S by marginalizing over all other features in the complement subset X_C (Zhao & Hastie, 2021). Given the ML model f, the partial function f_{x_S} is defined as:

$$f_{x_S}(x_S) = E_{x_C}[f(x_S, x_C)] = \int f(x_S, x_C)\, d\mathbb{P}(x_C) \tag{3}$$

With $d\mathbb{P}(x)$ being the marginal distribution of X_C. Marginalizing over all other features leads to a function that is solely dependent on the features X_S to be analyzed. The partial function f_{x_S} is estimated using the Monte Carlo method to average over actual features values $x_C^{(i)}$ while keeping X_S constant:

$$f_{x_S}(x_S) = \frac{1}{n}\sum\nolimits_{i=1}^{n} f\left(x_S, x_C^{(i)}\right) \tag{4}$$

As shown in Greenwell (2017), all values of feature x_S (e.g. living area) are in a first step replaced with the particular feature value (e.g. of the first observations). The ML model predicts expected output values for the newly created dataset (where all observations have the same constant feature value x_S). Averaging over these predictions calculates the marginal effect at the particular feature value. This step is repeated n times to obtain a marginal effect for all observed feature values. Finally, the single feature values are plotted against the resulting f_{x_S}. For a linear hedonic model, e.g. based on ordinary least squares (OLS), a PD plot would show a straight line representing the specific estimated coefficient. As Zhao and Hastie (2021) state, PD plots are a valuable visualization tool to interpret how the prediction of ML models depend on specific features.

Econometric results

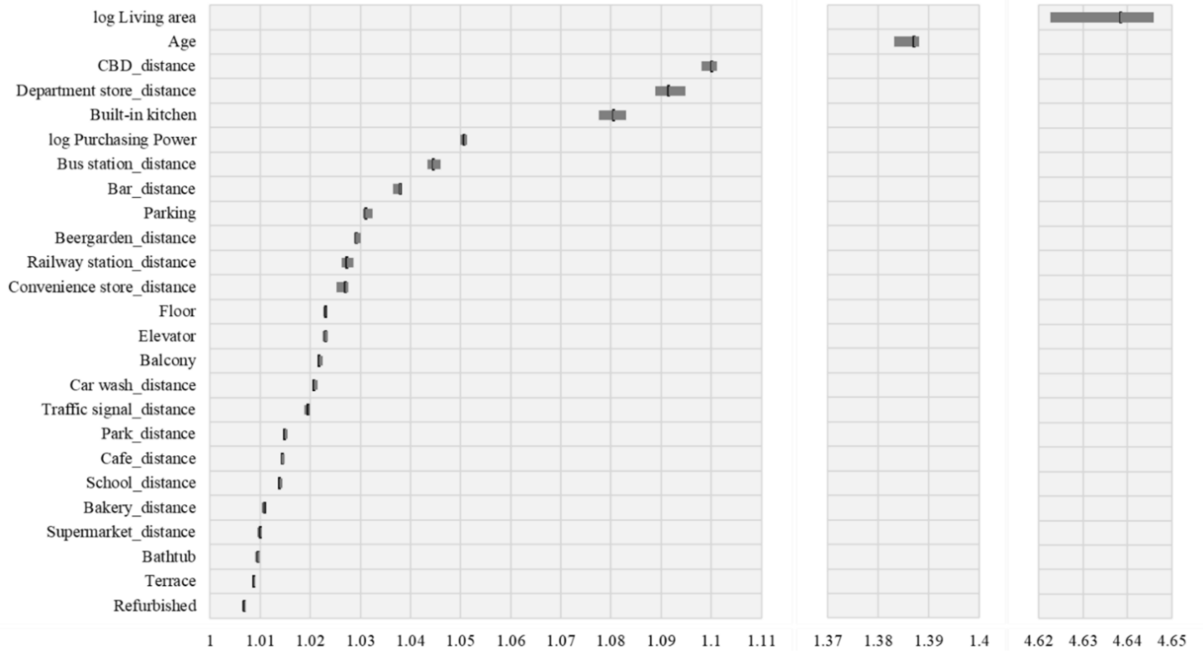
To set up a functional ML framework, we first train the XGB algorithm on our dataset of rental prices described in the data section. We apply random cross-validation with five folds and five repetitions. The tuning process takes 16 hours with 72 central processing units (CPUs) running simultaneously. The final XGB model is trained with $\eta = 0.243$, $\gamma = 0.0431$, $\lambda = 28.99$ and $\alpha = 22.64$. The out of sample rental prediction with XGB yields to a R^2 of 92.50%. The mean absolute percentage error marks 11.13%. Moreover, 57.96% of all predictions deviate less than 10% from the observed values. The tuned XGB algorithm subsequently allows a post-hoc analysis with a set of model-agnostic interpretation tools to identify feature importance and feature effects.[10]

Feature importance of the hedonic characteristics

Figure 2 provides the relevance of all characteristics for the ML prediction based on FI. The features are individually ranked on the y-axis from most important at the top to least important at the bottom. The x-axis provides information of how much prediction accuracy changes when the feature values are permutated. Median values are plotted with the bar denoting the 5% and 95% quantiles. Feature importance ratios exceeding 1 indicate an observable impact on the overall prediction. Ratios that tend towards 1 imply a rather negligible influence of the features.

[10] To ensure basic hedonic functionality of a hedonic rent estimation, we apply linear, spatial and non-linear methods in advance. The corresponding methodology and the results are presented and discussed in Appendix 1 (methodology) and 2 (results). All variables show expected signs and do not contradict findings from related literature.

Figure 2: Feature importance of the hedonic characteristics



Note: The figure displays the median values of the relative feature importance obtained with XGB. MAE is chosen as loss function. Variables are ranked based on their FI score. The bar denotes the 5% and 95% quantiles of the distribution of FI scores after 100 repetitions. A break in the horizontal axis is conducted to ease readability.

It is not surprising, that living area and age are seen to have by far the biggest impact on rental prediction. Their median values highlight that randomly permuting living area and age individually 100 times, increases the model error by a factor of 4.64 and 1.39, while keeping all other variables constant. Furthermore, distance to the CBD and to a department store are of high importance and associated with an increase in MAE of 1.10 and 1.09. We expect both variables to be a suitable proxy for a good location.[11] Moreover, the presence of a built in kitchen is also heavy influential. The purchasing power per household is followed by the distances to the bus station and the nearest bar and beer garden.[12] The existence of a parking spot complements the ten most influential variables. We will not discuss the remaining variables in detail since their contribution seems rather marginal. The small distribution of FI for all variables demonstrated by the 5% and 95% quantile indicates that the results are stable over all repetitions. To summarize, feature importance ranks how relevant a variable is for the predictive

[11] In major German cities, department stores are usually located either close to the city center or in highly frequented and therefore good shopping locations.

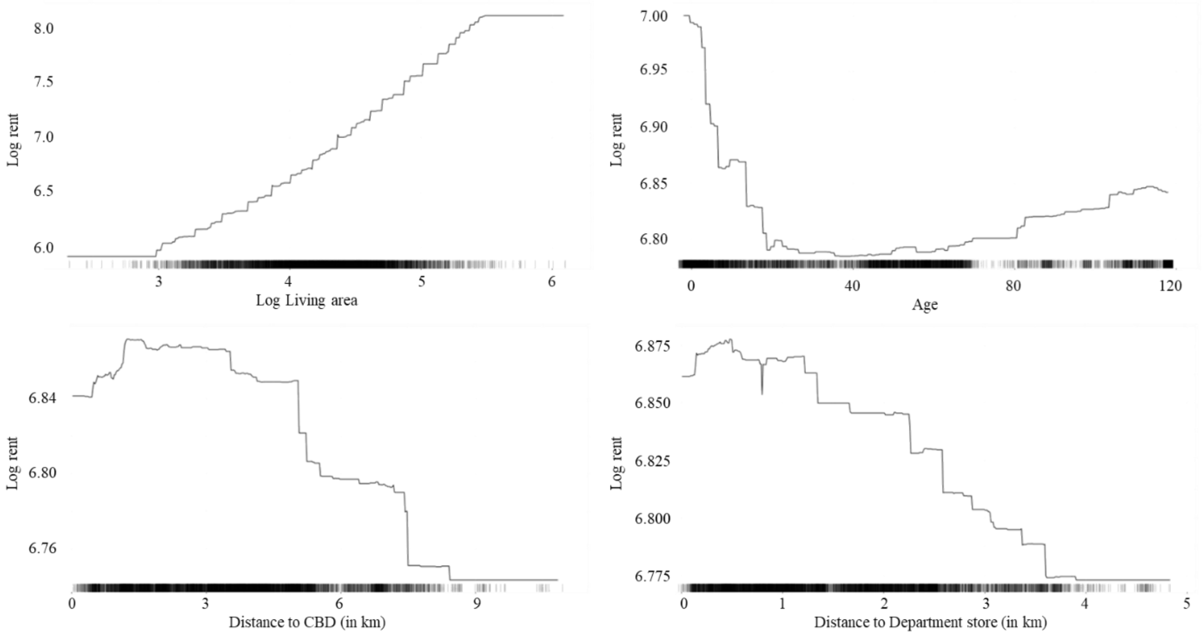
[12] Beer gardens are perceived as important hospitality institutions in Germany and thus the result is not surprising.

task as it provides which variables are more or less influential for an ML model. One can thus obtain a first impression whether an algorithmic hedonic model delivers reliable results that are based on a plausible understanding of the economic context. However, FI does not provide any information about the sign. To clarify e.g. whether a small or large distance is decisive, we investigate feature effects in a next step.

Feature effects of the hedonic characteristics

PD plots enable an analysis of how a certain feature influences the rental prediction and which relationships between residential rents and property characteristics has been traced by the algorithm. While the X-Axis provides information on the independent variable with the stacked black lines indicating the amount of observations, the Y-Axis shows the respective rent level. Since marginal effects are calculated and averaged for every feature value, PD plots require high computational power. Thus, we plot the partial dependence for the year 2019, whose generation took eight hours of computing time.

Figure 3: PD Plots for living area, age, distances to CBD and department store in 2019



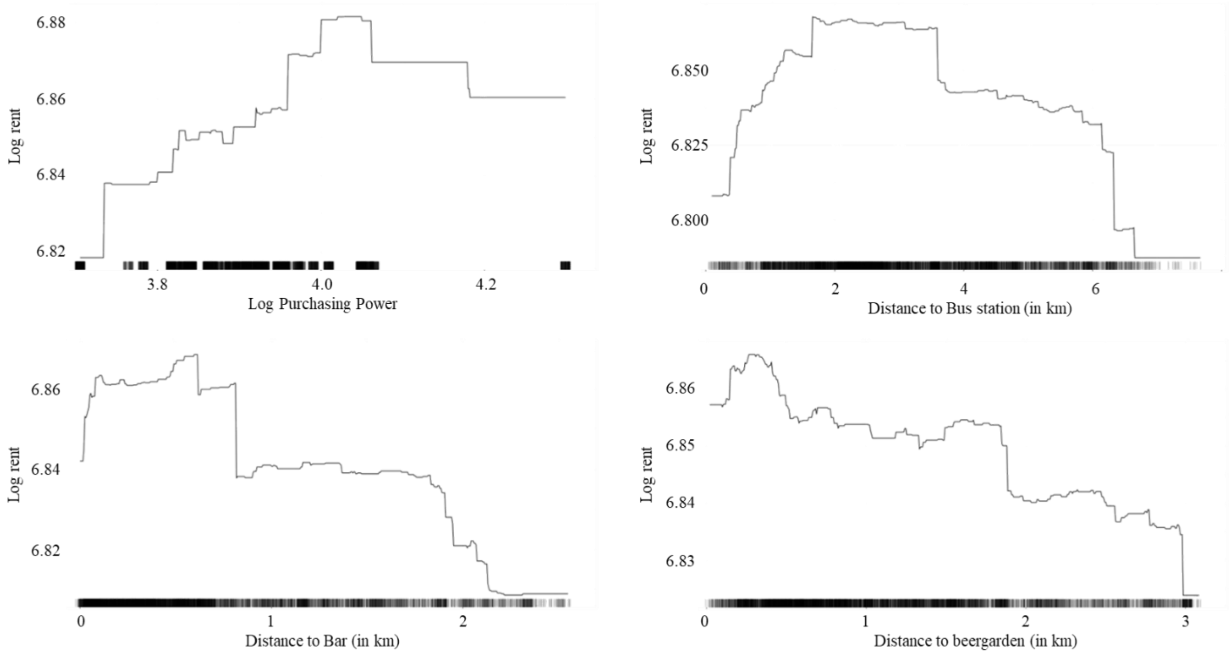
Note: The figure displays the partial dependence of the most important feature regarding two structural characteristics and distance to CBD and department store. The vertical axis denotes the feature values of log rent level while the horizontal axis represents the covariates feature values. Stacked black lines display the number of observations. *Source:* Own depiction.

Figure 3 demonstrates how rental prices are associated with the four most influential characteristics living area, age and distance to CBD and department store. We start with the most important feature living area, which is incorporated as the natural logarithm. Since the PD plot highlights a linear relationship, the commonly applied log-log transformation can be confirmed as a good approximation of the positive relationship between living area and rent. Recent hedonic literature on property prices provides similar findings for the positive relationship (e.g. Dumm et al., 2016, Dumm et al., 2018 or Stamou et al., 2017). Age is perceived to be more complex, though intuitive. We find rental values to decrease with greater age until a building year of 1990-2000. While newly build apartments obtain highest rents, depreciation, changes in living preference as well as increasing requirements on energy-efficient construction most likely result in a steep decline in rental values. This is followed by an indifference of rental values up to 1940th. Frankfurt was heavily bombed in World War II, with emergence constructions of social housing provided by the government in the following decades. Therefore, historical pre-war buildings face higher rents. Consequently, building age displays a u-shaped relationship, as e.g. incorporated in Mayer et al. (2019).

Distance to CBD is perceived to be highly influential. In general, we find rental prices to decline with greater distance to the city center. Hedonic literature suggests similar conclusions since authors such as Osland (2010) or Zheng et al. (2016) also find a negative relationship between property prices and distance to the city center. However, the opposite effect is visible for close proximity. We expect tenants to appreciate separation from very urban areas. A graphical turning point can be found at about 1.5 km, followed by moderate decline in rental prices. Interestingly, apartments close to the CBD face comparable rental values than the ones in 5 km distance. A steep decrease in rent levels can be seen beyond 5 and 7.5 km.

Regarding local supply, department stores are rather linearly and negatively associated with rental values. The proximity to shopping facilities results in increasing rents. We do not find an equivalent distance variable in the hedonic literature, however, Dubé and Legros (2016) show a positive price effect for properties not more than 1 km away from a shopping center. Interesting to note, the distance to department store drops sharply at about 1.5 and 2.5 km. This could indicate a critical distance for consumer goods. However, FI identifies supermarket as the least important distance variable. We assume that a high density of supermarkets in urban areas ensure local supply for everyday goods and therefore result in a negligible influence on rental values. In contrast, we assume different circumstances in rural communities. With minor influence due to the limited appearance of department stores, we expect the importance of supermarket to be more pronounced in non-metropolitan areas. Furthermore, FI ranks the presence of a built-in kitchen as important. Gröbel and Thomschke (2018) find a significant positive relationship between built-in kitchens and rents in Berlin (Germany). However, due to its binary nature, the visualization with PD plots is limited.

Figure 4: PD Plots for purchasing power, distances to bus station, bar and beer garden

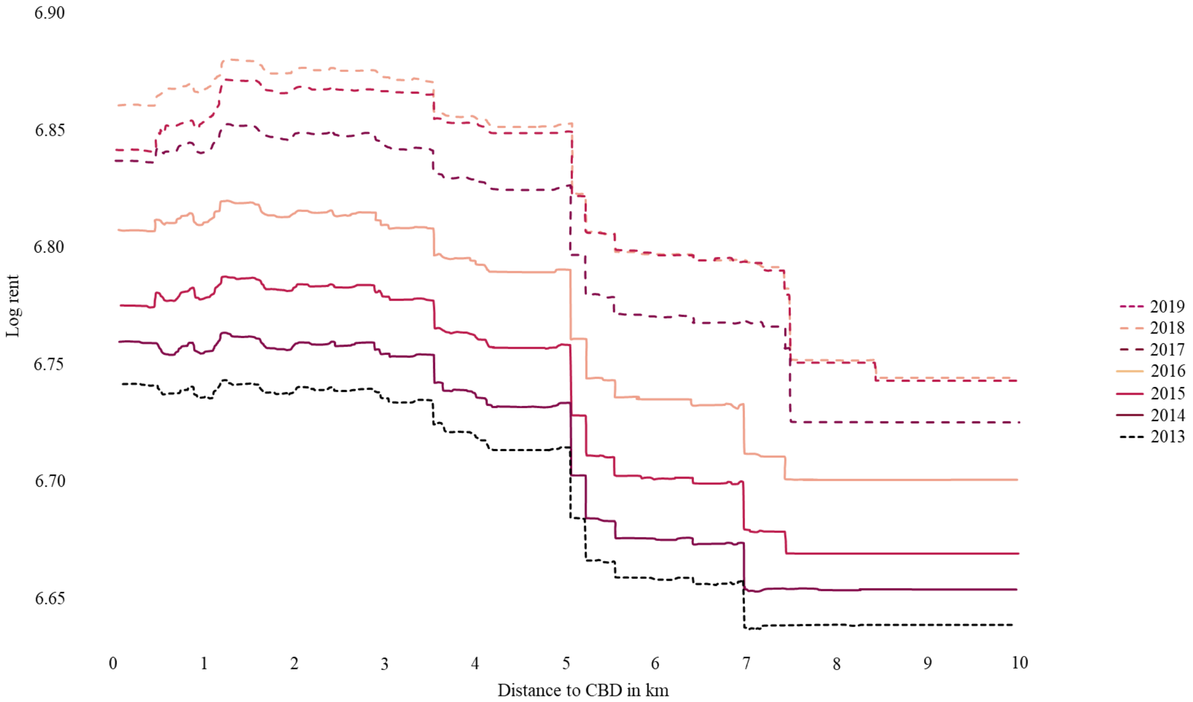


Note: The figure displays the partial dependence of the most important feature regarding two structural characteristics and distance to CBD and department store. The vertical axis denotes the feature values of log rent level while the horizontal axis represents the covariates feature values. Stacked black lines display the number of observations. *Source:* Own depiction.

The next most important characteristics displayed in Figure 4 are, according to FI, purchasing power and distance to bus station, bar and beer garden. We find socio-demographic information to show a rather linear relationship. Neighborhoods with high purchasing power are associated with more expensive apartments and thus the variable is perceived as a characteristic of a good residential area. A steep increase in rental values for high wealth districts could reflect the segment of high-rise apartments in residential towers. While the construction of high-rise buildings is restricted in most German cities, Frankfurt has early incorporated tower buildings in urban planning. These do not only represent the highest price segment in the residential market of Frankfurt but have shown to be driver of residential prices and rents in the last years.

Interesting to note, the distance to bar, beer garden and bus station have shown to affect the overall prediction the most out of all hospitality and public transport features. All three variables show a non-linear relationship with residential rents. We find the distance to a bar to be positively associated with rental values up to approx. 700 meters. While a bar in close proximity would result in lower rents, the access to hospitality leads to an increase in rental values only from a certain distance. We expect tenants to face a trade-off between accessibility and negative externalities such as noise. The same relationship holds for the variable bus station. A location further away from a central bus hub is linked to higher rental values up to approx. 1.7 km. Since central hubs are related to mostly high urban density and traffic, we assume that tenants appreciate locational separation. The plot reveals the relationship to be quite constant until 3.5 km, followed by declining rental prices. The accessibility to central hubs through different means of transport seems to overlay negative effect of a larger distance. However, after 3.5 km, we find this effect to become visible and apartments that are poorly located in terms of transport face discounts for low accessibility. The presence of a parking spot complements the ten most influential variables, yet as a binary variable it is not displayed as a PD plot.

Figure 5: PD plot for rent and distance to CBD for the years 2013 to 2019



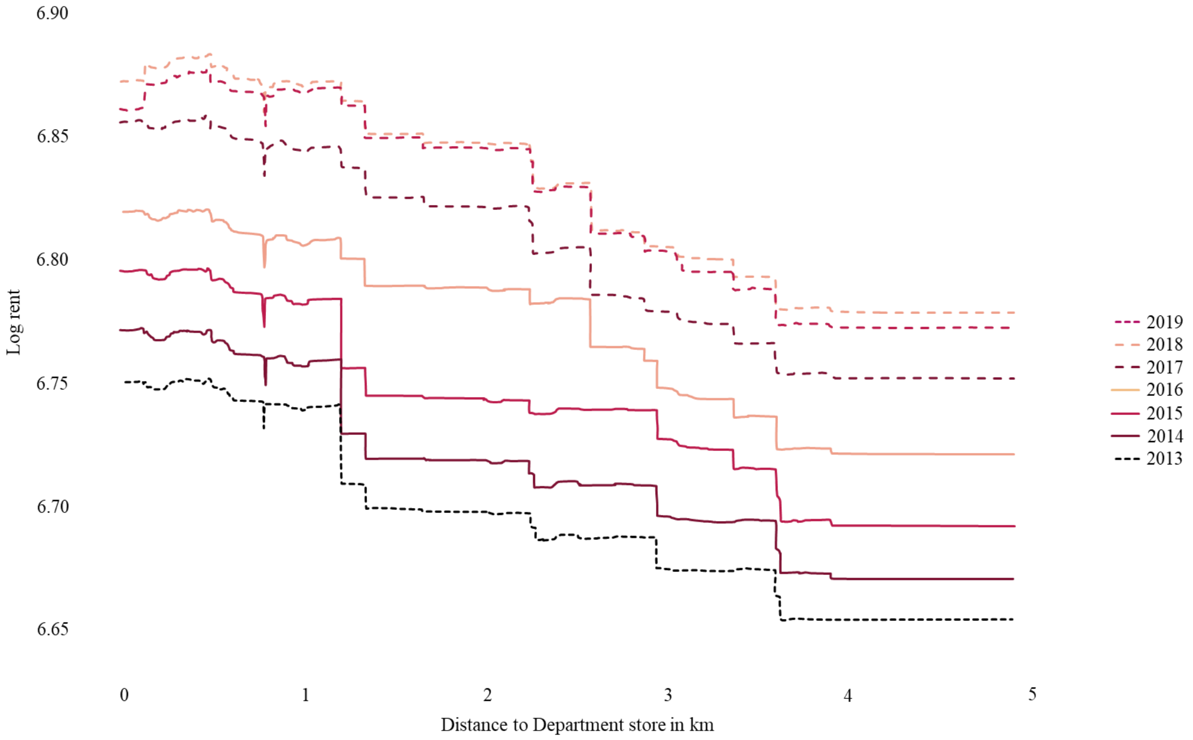
Note: The figure displays the partial dependence of important variables over different periods. The vertical axis denotes the feature values of the log rent level while the horizontal axis denotes the covariates feature values. *Source:* Own depiction.

Adding a temporal dimension to our analysis by displaying feature effects on a yearly basis enables us in a last step to illustrate temporal dynamics of the effects of hedonic characteristics. We demonstrate the latter by analyzing the distance to the CBD (Figure 5) and the distance to a department store (Figure 6).

At first, Figure 5 shows a negative relationship between rents and the distance to CBD across time. A continuous upwards shift for all feature values indicates increasing rent levels during the observed period. Only the graph of the year 2019 behaves differently, since it moves below 2018 for closer proximity and analogous from 5 km distance onwards. This development could be attributed to a declining preference for downtown locations in combination with overall stable rent levels in recent years. Although the course of all lines is quite similar, we find some differences. First, a drop in rental prices at a distance of 5 km is less pronounced for 2017, 2018 and 2019 than for previous years. This possibly indicates that residential locations further away from the center experienced rent increases due to a growing preference for sub-urban areas during the last years. Second, another major decline can be recognized at 7 km for 2013 to 2016.

In the following years 2017 to 2019, however, this is only noticeable at a distance of approx. 7.5 km, but the downturn is considerably stronger. Both changes indicate that residential locations in medium distance to the center (5 to 7.5 km) experienced stronger rent increases compared to central as well as periphery location. We would assume that high demand in central locations results in a preference shift towards apartments further away from the CBD.

Figure 6: PD plot for rent and distance to department store for the years 2013 to 2019



Note: The figure displays the partial dependence of important variables over different periods. The vertical axis denotes the feature values of the log rent level while the horizontal axis denotes the covariates feature values. *Source:* Own depiction.

In Figure 6, a negative relationship between rents and the distance to a department store is displayed, yet a similar pattern for the graphs can be seen in terms of comparable upwards shift of rents throughout all periods and 2019 being slightly below 2018. A first major decline is visible at approximately 1.2 km, with the years 2013, 2014 and 2015 experiencing a stronger decrease. From 2.6 km distance, the picture is the other way around. Whereas rents fell rapidly from 2016 to 2019, the downturn was not as strong as in previous years. The findings indicate that while locations between 1.2 km and 2.8 km gained popularity, locations in close proximity as well as further away remained more or less stable. Appendix 4 provides additional and centered PD plots for the features Distance to Distance to department store. Centered PD plots aid and underpin the interpretation of the differences in PDs throughout the years.

Ultimately, the feature effects technique yields greater transparency of how the different inputs contribute to the final estimation of the ML model. By visualizing the individual relations between the variables and the rent to be estimated, this method demonstrates which (economic) rational the algorithm has learned from the data and accordingly integrated into its internal calculations.

Conclusion

This paper sheds light on how Machine Learning (ML) based decision making in hedonic modelling can be made more transparent. We visualize and investigate the relationship between residential rents and a set of hedonic variables which was learned by a ML model. Based on a residential dataset of more than 52k apartments in Frankfurt am Main, Germany, we apply the eXtreme Gradient Boosting algorithm (XGB) for rental prediction. Model-agnostic Interpretable Machine Learning (IML) methods are subsequently used to examine feature importance and feature effects. Feature importance (FI) reveals that living area, age and the distance to CBD and a department store influence the overall rental prediction the most. In contrast, the least important features are several structural dummy variables and the distance to a supermarket and a bakery, albeit in an urban setting with presumably excellent coverage with every-day shopping facilities.

We plot the partial dependences (PD) for the influential variables that were detected in the preceding analysis to highlight feature effects. Although the relationship of rental values and the distance to CBD and department store is mainly linear, major declines at specific proximity values indicate that critical distances to the center as well as to local supply exist. Furthermore, there seems to be a difference in rent level to the wealthiest neighborhoods. Interestingly, we find that close proximity to hospitality and public transport is associated with rental discounts. In addition, the inspection of PD plots on a yearly basis reveals that especially apartments in a medium distance to the city center face considerable higher rent increases over the years. We assume both an increasing preference for less urban areas as well as peaking rent in the center to be possible reasons.

To conclude, interpretation methods can reveal the rationale behind the ML models estimation by demonstrating what relationship the algorithm detects in the underlying data. Peeking inside the black box enables researchers to reenact how a ML model arrived at its prediction and will help to gain new insights, ease practical applications and enhance reliability in algorithmic decisions.

The insights gained by these methods are relevant not only for research but also for practice in the private as well as public sector. Since real estate professionals commonly use ML to inform their decision making (RICS, 2017), model-agnostic methods provide a useful framework to effectively handle AI-based results. Whereas the advantages of these methods have already been discussed in detail, difficulties and limitations must also be pointed out. First of all, there are challenges in terms of computing power. Whereas parametric or semi-parametric methods are usually able to estimate hedonic models within seconds, ML-based methods such as XGB take considerably longer. This also applies to the application of IML. Furthermore, it should be noted that data availability is of course essential for hedonic models. Even with ML-based models, an omitted variable bias can drastically reduce the informative value and thus the applicability. Admittedly, the data set of this study is quite extensive, but there are of course other additional apartment features imaginable that could influence the meaning of the results.

IML is a rapidly evolving field with new methods and applications being continuously proposed. Although this research area has achieved a degree of stability (Molnar et al., 2020), it is still in its infancy and faces several challenges to overcome. On the one hand, there is a need to define what interpretability means to then evaluate how black box models can be made more interpretable. On the other hand, the sensitivity of interpretation methods is of high importance, since not only these methods, but also the ML techniques are dynamically developing. To further improve trust in algorithmic decisions, ongoing research is necessary. We expect IML methods to be a valuable addition to the hedonic practice, both because it contributes to the transparency of ML models and because it provides insights on potentially unknown relationships in real estate hedonic modelling.

Appendix

Appendix 1

We apply different hedonic methods that have been used regularly in the literature. First, we deploy a hedonic OLS modelling approach to estimate the effects of property characteristics on rental prices. Linear hedonic regression represents the standard approach in modelling real estate prices and rents and is frequently used in housing studies (Mayer et al., 2019). The hedonic regression describes the rent Y as the sum of the predicted values of its characteristics X_j:

$$Y = \beta_0 + \sum_{j=1}^{J} X_j \beta_j + \varepsilon \tag{5}$$

In accordance to the real estate literature, a semi-log functional form with log-transformation of the dependent variable is conducted. Property characteristics include structural, socio-economic neighborhood and locational features. Proximity variables account for the spatial distance to public amenities and transport. Further spatial effects are modelled via spatial expansion by incorporating the coordinates in terms of longitude and latitude (Bitter et al., 2007, Chrostek & Kopczewska, 2013, Pace & Hayunga, 2020). Furthermore, temporal dummies are included for the specific month and year.

Many authors argue that property prices and rents may contain two key figures, namely spatial autocorrelation and spatial heterogeneity, that can require the spatial extension of hedonic models (LeSage, 1999). Since the occurrence of spatial effects can lead to misspecifications and biased results in the OLS framework (Anselin, 1988), we additionally apply a spatial autoregressive regression (SAR) with the following functional form:

$$Y = X\beta + \rho WY + \varepsilon \tag{6}$$

ρWY denotes a spatial lag of the target variable Y, with W being the spatial weight matrix that specifies the spatial structure, and ρ representing the spatial lag parameter.

However, linear models are subject to various restrictions due to their functional parametric form that can yield to misspecifications (Mason & Quigley, 1996; Pace, 1998). Because relationships in housing markets appear often to be non-linear, hedonic modelling can require the incorporation of more flexible functional forms to account for nonlinearity (Bontemps et al., 2008; Brunauer et al., 2013). Hence, a semi-parametric generalized additive model (GAM) is further considered.

$$Y \ = \ \beta_o + \sum_{j=1}^{J} X_j \beta_j + \sum_{p=1}^{P} f_p(X_p) + \varepsilon \qquad (7)$$

GAM relaxes the linearity assumption by replacing the parametric linear relationship with non-parametric smoothers (e.g. splines, near neighbor and kernel smoothers). The linear equation is expanded by p smooth functions f_p in order to identify latent non-linear effects.

The results of the aforementioned methods are presented in Appendix 2. The coefficients provide expected signs and confirm a good model fit by showing acceptable R^2.

Appendix 2: Results of the OLS, GAM and SAR estimation

Dependent variable: log Rent per month

	OLS			GAM			SAR		
log Living area	0.939	***	(0.002)	0.900	***		0.928	***	(0,008)
Floors	0.002	***	(0.0004)	0.003	***	(0.0003)	0.003	***	(0,002)
Age (relative to 2017)	-0.0002	***	(0.00003)	s 8.000	***		-0.000	***	(0,0001)
Bathtub	-0.032	***	(0.002)	-0.016	***	(0.001)	-0.032	***	(0,006)
Refurbished	-0.015	***	(0.002)	0.005	***	(0.002)	-0.013	***	(0,007)
Built-in kitchen	0.084	***	(0.002)	0.077	***	(0.002)	0.077	***	(0,007)
Balcony	0.011	***	(0.002)	0.025	***	(0.002)	0.012	***	(0,007)
Parking	0.053	***	(0.002)	0.032	***	(0.002)	0.048	***	(0,008)
Elevator	0.053	***	(0.002)	0.020	***	(0.002)	0.048	***	(0,009)
Terrace	0.041	***	(0.002)	0.020	***	(0.002)	0.041	***	(0,009)
log Purchasing Power	0.406	***	(0.011)	0.069	***	(0.002)	0.313	***	(0,040)
CBD_distance	-0.019	***	(0.001)	s 8.692	***		-0.014	***	(0,002)
Bar_distance	-0.031	***	(0.002)	s 8.579	***		-0.024	***	(0,008)
Beergarden_distance	-0.020	***	(0.002)	s 8.631	***		-0.015	***	(0,005)
Cafe_distance	-0.014	***	(0.003)	s 8.700	***		-0.011	***	(0,010)
Bakery_distance	-0.011	***	(0.003)	s 8.842	***		-0.016	***	(0,009)
Convenience store_distance	-0.036	***	(0.002)	s 8.144	***		-0.035	***	(0,007)
Department store_distance	-0.006	***	(0.001)	s 8.580	***		-0.008	***	(0,005)
Supermarket_distance	-0.018	***	(0.006)	s 6.487	***		-0.029	***	(0,020)
Bus station_distance	-0.028	***	(0.001)	s 8.794	***		-0.017	***	(0,004)
Railway station_distance	-0.020	***	(0.002)	s 8.757	***		-0.020	***	(0,007)
Traffic signals_distance	0.086	***	(0.007)	s 8.243	***		0.075	***	(0,024)
Car wash_distance	0.012	***	(0.002)	s 8.763	***		0.007	***	(0.006)
Park_distance	-0.024	***	(0.006)	s 8.343	***		-0.019	***	(0.020)
School_distance	-0.003	***	(0.005)	s 8.412	***		0.008	***	(0,006)
Constant	-34.043	***	(3.087)	2.405	***	(0.100)	-22.860	***	(11,359)
rho							0.131	***	
time controls	Yes			Yes			Yes		
locational controls	Yes			Yes			Yes		
observations	52,966			52,966			52,966		
R^2	0.880						0.885		
adjusted R^2	0.880			0.898					
UBRE				0.028					

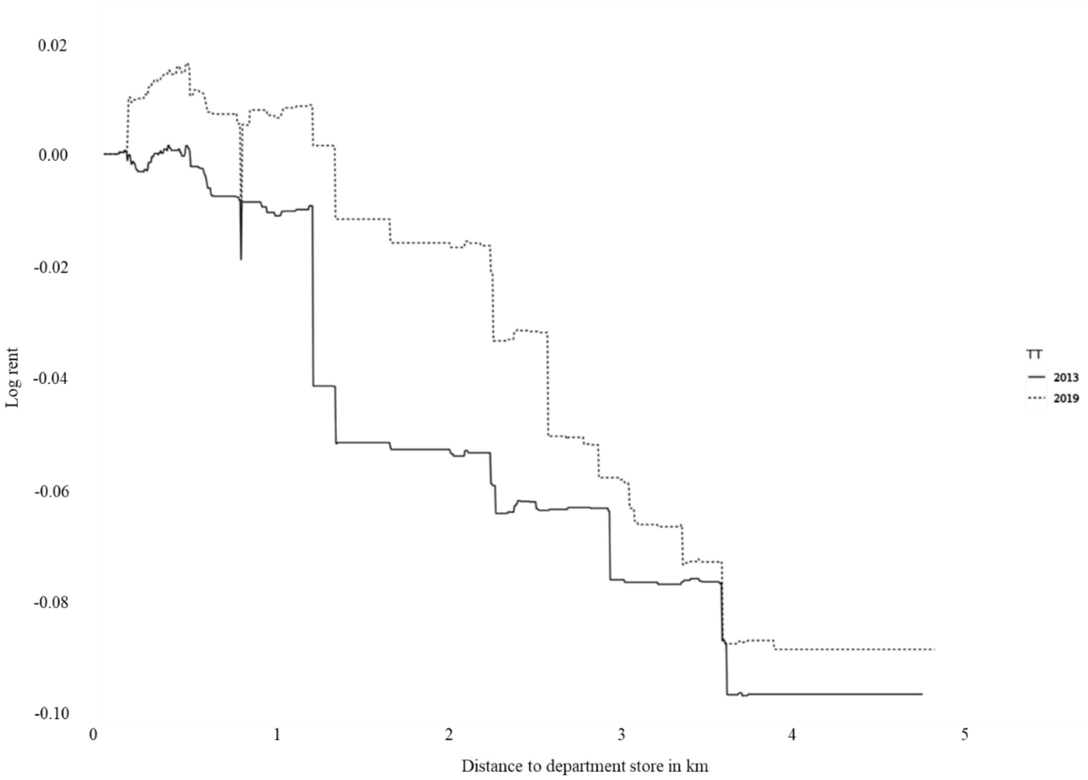
Note: *p<0.1; **p<0.05; ***p<0.01, standard errors are displayed in parentheses. The GAM column reports the estimated degrees of freedom of the smooth terms (s) as well as their joint significance. Time controls (year and month) as well as location controls (apartment coordinates) are included in all models.

Appendix 3: Correlation matrix

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.
1.	Log rent p.m.	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.	Log living area	0,89	1	0	0	0	0	0	0	0	0	0	0	0	0	0,01	0	0,90	0	0	0	0	0,21	0,92	0	0	0
3.	Floor	0,07	0,02	1	0	0,05	0	0	0	0	0	0	0	0	0	0,23	0	0	0	0	0	0	0	0	0	0	0
4.	Age	-0,14	-0,10	-0,09	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5.	Bathtub	0,12	0,18	-0,01	-0,06	1	0,74	0	0	0	0	0	0,22	0	0,02	0	0	0,77	0	0	0,01	0,51	0	0,01	0,62	0	0,05
6.	Refurbished	-0,09	-0,08	-0,04	0,23	0,00	1	1,00	0	0	0	0	0	0	0,91	0	0	0	0	0	0	0	0	0	0	0,46	0
7.	Bulit-in-kitchen	0,32	0,21	0,05	-0,13	0,02	0,00	1	0	0	0	0	0,01	0	0	0,01	0	0	0	0	0	0	0	0	0	0	0
8.	Balcony	0,19	0,19	0,10	-0,27	0,12	-0,06	0,06	1	0	0	0	0,85	0	0	0	0	0	0	0	0	0,09	0	0,02	0	0	0
9.	Parking	0,37	0,33	0,05	-0,57	0,08	-0,12	0,24	0,21	1	0	0	0	0	0,24	0	0	0	0	0	0	0,04	0	0	0	0	0
10.	Elevator	0,24	0,12	0,28	-0,56	0,03	-0,17	0,18	0,23	0,45	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11.	Terrace	0,21	0,21	-0,17	-0,20	0,05	-0,06	0,09	-0,09	0,21	0,11	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0,07	0
12.	Purchasing power	0,10	0,11	-0,10	0,04	0,01	0,05	0,01	0,00	0,01	-0,06	0,05	1	0	0	0	0	0	0	0	0	0	0	0	0	0,01	0
13.	CBD_distance	-0,22	-0,07	-0,11	-0,04	0,01	0,02	-0,12	-0,03	-0,02	-0,18	0,03	0,42	1	0	0	0	0	0	0	0	0	0	0	0	0	0
14.	Bar_distance	-0,20	-0,04	-0,13	-0,08	0,01	0,00	-0,16	0,02	-0,01	-0,19	0,04	0,19	0,42	1	0	0	0	0	0	0	0	0	0	0	0	0
15.	Biergarten_distance	-0,06	0,01	-0,01	-0,28	0,03	-0,11	-0,01	0,09	0,18	0,13	0,06	-0,06	0,07	0,05	1	0	0	0	0	0	0	0	0	0	0	0
16.	Cafe_distance	-0,15	-0,03	-0,10	-0,14	0,03	-0,03	-0,15	0,07	0,04	-0,10	0,06	0,18	0,35	0,53	0,20	1	0	0	0	0	0	0	0	0	0	0
17.	Bakery_distance	-0,09	0,00	-0,07	-0,12	0,00	-0,02	-0,08	0,04	0,06	-0,06	0,06	0,22	0,33	0,37	0,11	0,31	1	0	0	0	0	0	0	0	0	0
18.	Convenience store_distance	-0,11	0,03	-0,11	-0,22	0,02	-0,05	-0,07	0,06	0,13	-0,02	0,09	0,45	0,54	0,48	0,38	0,42	0,42	1	0	0	0	0	0	0	0	0
19.	Department store_distance	-0,18	-0,02	-0,09	-0,21	0,05	-0,05	-0,11	0,07	0,11	-0,07	0,06	0,17	0,43	0,47	0,46	0,43	0,22	0,58	1	0	0	0	0	0	0	0
20.	Supermarket_distance	-0,05	0,03	-0,10	-0,08	0,01	-0,03	-0,07	0,03	0,03	-0,11	0,06	0,27	0,28	0,37	0,16	0,39	0,31	0,34	0,22	1	0	0	0	0	0	0
21.	Bus station_distance	-0,25	-0,09	-0,15	-0,08	0,00	0,02	-0,15	0,01	-0,01	-0,18	0,03	0,29	0,62	0,57	0,07	0,40	0,48	0,59	0,48	0,29	1	0	0	0	0	0
22.	Railway station_distance	-0,14	0,01	-0,12	-0,21	0,03	-0,05	-0,09	0,07	0,11	-0,06	0,07	0,41	0,59	0,40	0,45	0,43	0,35	0,65	0,64	0,34	0,52	1	0	0	0	0
23.	Traffic signals_distance	-0,08	0,00	-0,09	-0,03	0,01	0,02	-0,08	0,01	-0,01	-0,14	0,04	0,20	0,38	0,45	0,10	0,37	0,26	0,42	0,30	0,37	0,33	0,35	1	0	0	0
24.	Car wash_distance	0,08	0,08	-0,01	-0,05	0,00	-0,02	0,03	0,03	0,04	0,06	0,03	0,16	-0,13	0,31	0,16	0,09	0,03	0,13	-0,08	-0,03	-0,05	0,02	0,09	1	0	0
25.	Park_distance	-0,11	-0,03	-0,04	-0,06	0,03	0,00	-0,08	0,03	0,01	-0,08	0,01	0,01	0,24	0,32	-0,02	0,23	0,35	0,25	0,23	0,14	0,30	0,18	0,24	-0,08	1	0
26.	School_distance	-0,06	-0,02	0,03	-0,11	0,01	-0,03	-0,04	0,05	0,04	0,02	0,02	-0,08	0,05	0,19	0,16	0,24	0,17	0,20	0,24	0,30	0,10	0,16	0,25	0,01	0,15	1

Notes: Pearson correlation coefficients are displayed below the diagonal and p-values above

Appendix 4: Centred PD plot for Distance to department store



Note: The figure displays the partial dependence centered at lowest feature value. The vertical axis denotes the feature values of the log rent level while the horizontal axis denotes the covariates feature values. *Source:* Own depiction.

References

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access*, *6*, 52138–52160.

Ahmad, M. A., Eckert, C., & Teredesai, A. (2018). Interpretable machine learning in healthcare. *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 559–560.

Anselin, L. (1988). *Spatial econometrics: methods and models*. Springer Science & Business Media.

Antipov, E. A., & Pokryshevskaya, E. B. (2012). Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. *Expert Systems with Applications*, *39*(2), 1772–1778.

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., & Benjamins, R. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115.

Below, S., Beracha, E., & Skiba, H. (2015). Land erosion and coastal home values. *Journal of Real Estate Research*, *37*(4), 499–536.

Berk, R. A., & Bleich, J. (2013). Statistical procedures for forecasting criminal behavior: A comparative assessment. *Criminology & Public Policy*, *12*, 513–544.

Bitter, C., Mulligan, G. F., & Dall’erba, S. (2007). Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems*, *9*(1), 7–27.

Bontemps, C., Simioni, M., & Surry, Y. (2008). Semiparametric hedonic price models: assessing the effects of agricultural nonpoint source pollution. *Journal of Applied Econometrics*, *23*(6), 825–842.

Breiman, L. (2001a). Random forests. *Machine Learning*, *45*(1), 5–32.

Breiman, L. (2001b). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, *16*(3), 199–231.

Brooks, T. J., Humphreys, B. R., & Nowak, A. (2020). Strip Clubs, “Secondary Effects” and Residential Property Prices. *Real Estate Economics*, *48*(3), 850–885.

Brunauer, W., Lang, S., & Umlauf, N. (2013). Modelling house prices using multilevel structured additive regression. *Statistical Modelling*, *13*(2), 95–123.

Cajias, M., & Freudenreich, P. (2018). Exploring the determinants of liquidity with big data – market heterogeneity in German markets. *Journal of Property Investment & Finance*, *36*(1), 3–18.

Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, *22*(3), 453–474.

Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics*, *8*(8), 832.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 785–794.

Chernobai, E., Reibel, M., & Carney, M. (2011). Nonlinear Spatial and Temporal Effects of Highway Construction on House Prices. *The Journal of Real Estate Finance and Economics*, *42*(3), 348–370.

Chin, S., Kahn, M. E., & Moon, H. R. (2020). Estimating the Gains from New Rail Transit Investment: A Machine Learning Tree Approach. *Real Estate Economics*, *48*(3), 886–914.

Chrostek, K., & Kopczewska, K. (2013). Spatial prediction models for real estate market analysis. *Ekonomia*, *35*(0).

Conway, D., Li, C. Q., Wolch, J., Kahle, C., & Jerrett, M. (2010). A spatial autocorrelation approach for examining the effects of urban greenspace on residential property values. *The Journal of Real Estate Finance and Economics*, *41*(2), 150–169.

Court, A. T. (1939). Hedonic price indexes with automotive examples. In *The Dynamics of Automobile Demand* (pp. 99–117).

Des Rosiers, F., Dubé, J., & Thériault, M. (2011). Do peer effects shape property values? *Journal of Property Investment & Finance*, *29*(4/5), 510–528.

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *Working Paper. ArXiv:1702.08608*.

Dubé, J., & Legros, D. (2016). A Spatiotemporal Solution for the Simultaneous Sale Price and Time-on-the-Market Problem. *Real Estate Economics*, *44*(4), 846–877.

Dubin, R. A. (1988). Estimation of regression coefficients in the presence of spatially autocorrelated error terms. *Review of Economics and Statistics*, 466–474.

Dumm, R. E., Sirmans, G. S., & Smersh, G. T. (2016). Price variation in waterfront properties over the economic cycle. *Journal of Real Estate Research*, *38*(1), 1–26.

Dumm, R. E., Sirmans, G. S., & Smersh, G. T. (2018). Sinkholes and Residential Property Prices: Presence, Proximity, and Density. *Journal of Real Estate Research*, *40*(1), 41–68.

Fernández-Avilés, G., Minguez, R., & Montero, J.-M. (2012). Geostatistical air pollution indexes in spatial hedonic models: the case of Madrid, Spain. *Journal of Real Estate Research*, *34*(2), 243–274.

Fisher, A., Rudin, C., & Dominici, F. (2019). All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. *Journal of Machine Learning Research*, *20*(177), 1–81.

Freybote, J., Fang, Y., & Gebhardt, M. (2017). The impact of temporary uses on property prices: the example of food trucks. *Journal of Property Research*, *34*(1), 19–35.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, *29*(5), 1189–1232.

Goodwin, K. R., La Roche, C. R., & Waller, B. D. (2020). Restrictions versus amenities: the differential impact of home owners associations on property marketability. *Journal of Property Research*, *37*(3), 238–253.

Greenwell, B. M. (2017). pdp: An R Package for Constructing Partial Dependence Plots. *R Journal*, *9*(1), 421–436.

Gröbel, S. (2019). Analysis of spatial variance clustering in the hedonic modeling of housing prices. *Journal of Property Research*, *36*(1), 1–26.

Gröbel, S., & Thomschke, L. (2018). Hedonic pricing and the spatial structure of housing data–an application to Berlin. *Journal of Property Research*, *35*(3), 185–208.

Haider, M., & Miller, E. J. (2000). Effects of transportation infrastructure and location on residential real estate values: application of spatial autoregressive techniques. *Transportation Research Record*, *1722*(1), 1–8.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction* (2nd ed.). Berlin: Springer.

Ho, W. K., Tang, B.-S., & Wong, S. W. (2021). Predicting property prices with machine learning algorithms. *Journal of Property Research*, *38*(1), 48–70.

Hoen, B., & Atkinson-Palombo, C. (2016). Wind Turbines, Amenities and Disamenitites: Astudy of Home Value Impacts in Densely Populated Massachusetts. *Journal of Real Estate Research*, *38*(4), 473–504.

Hoen, B., Brown, J. P., Jackson, T., Thayer, M. A., Wiser, R., & Cappers, P. (2015). Spatial hedonic analysis of the effects of US wind energy facilities on surrounding property values. *The Journal of Real Estate Finance and Economics*, *51*(1), 22–51.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (6th ed.). New York: Springer.

Jauregui, A., Allen, M. T., & Weeks, H. S. (2019). A spatial analysis of the impact of float distance on the values of canal-front houses. *Journal of Real Estate Research*, *41*(2), 285–318.

Kok, N., Koponen, E.-L., & Martínez-Barbosa, C. A. (2017). Big Data in Real Estate? From Manual Appraisal to Automated Valuation. *The Journal of Portfolio Management*, *43*(6), 202–211.

Kontrimas, V., & Verikas, A. (2011). The mass appraisal of the real estate by computational intelligence. *Applied Soft Computing*, *11*(1), 443–448.

Lam, K. C., Yu, C. Y., & Lam, C. K. (2009). Support vector machine and entropy based decision support system for property valuation. *Journal of Property Research*, *26*(3), 213–233.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(2), 132–157.

Lechner, M., Hasani, R., Amini, A., Henzinger, T. A., Rus, D., & Grosu, R. (2020). Neural circuit policies enabling auditable autonomy. *Nature Machine Intelligence*, *2*(10), 642–652.

LeSage, J. (1999). The theory and practice of spatial econometrics. *University of Toledo. Toledo, Ohio*, *28*(11).

Li, T. (2020). The Value of Access to Rail Transit in a Congested City: Evidence from Housing Prices in Beijing. *Real Estate Economics*, *48*(2), 556–598.

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2021). Explainable AI: A Review of Machine Learning Interpretability Methods. *Entropy*, *23*(1), 18.

Malpezzi, S. (2002). Hedonic Pricing Models: A Selective and Applied Review. In T. O'Sullivan & K. Gibb (Eds.), *Housing Economics and Public Policy* (pp. 67–89). Blackwell Science Ltd.

Mason, C., & Quigley, J. M. (1996). Non-parametric hedonic housing prices. *Housing Studies*, *11*(3), 373–385.

Mayer, M., Bourassa, S. C., Hoesli, M., & Scognamiglio, D. (2019). Estimation and updating methods for hedonic valuation. *Journal of European Real Estate Research*, *12*(1), 134–150.

McCluskey, W. J., McCord, M., Davis, P. T., Haran, M., & McIlhatton, D. (2013). Prediction accuracy in mass appraisal: a comparison of modern approaches. *Journal of Property Research*, *30*(4), 239–265.

Molnar, C. (2020). *Interpretable machine learning.*

Molnar, C., Casalicchio, G., & Bischl, B. (2020). Interpretable Machine Learning--A Brief History, State-of-the-Art and Challenges. *Working Paper ArXiv:2010.09337*.

Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106.

Osland, L. (2010). An application of spatial econometrics in relation to hedonic house price modeling. *Journal of Real Estate Research*, *32*(3), 289–320.

Pace, R. K. (1998). Appraisal using generalized additive models. *Journal of Real Estate Research*, *15*(1), 77–99.

Pace, R. K., & Hayunga, D. (2020). Examining the Information Content of Residuals from Hedonic and Spatial Models Using Trees and Forests. *The Journal of Real Estate Finance and Economics*, *60*(1-2), 170–180.

Pérez-Rave, J. I., Correa-Morales, J. C., & González-Echavarría, F. (2019). A machine learning approach to big data regression analysis of real estate prices for inferential and predictive purposes. *Journal of Property Research*, *36*(1), 59–96.

R Core Team. (2020). *R: A language and environment for statistical computing.*

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Model-Agnostic Interpretability of Machine Learning. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*.

RICS. (2017). *The Future of Valuations*. Royal Institute of Chartered Surveyor.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, *82*(1), 34–55.

Rouwendal, J., Levkovich, O., & van Marwijk, R. (2017). Estimating the Value of Proximity to Water, When Ceteris Really Is Paribus. *Real Estate Economics*, *45*(4), 829–860.

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Žídek, A., Nelson, A. W. R., Bridgland, A., Penedones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D. T., Silver, D., Kavukcuoglu, K., & Hassabis, D. (2020). Improved protein structure prediction using potentials from deep learning. *Nature*, *577*(7792), 706–710.

Seo, W. (2018). Does neighborhood condition create a discount effect on house list prices? Evidence from physical disorder. *Journal of Real Estate Research*, *40*(1), 69–88.

Sheppard, S. (1999). Hedonic analysis of housing markets. In P. Cheshire & E. S. Mills (Eds.), *Applied Urban Economics: Vol. 3. Handbook of Regional and Urban Economics* (pp. 1595–1635).

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*(3), 289–310.

Sirmans, S., Macpherson, D., & Zietz, E. (2005). The Composition of Hedonic Pricing Models. *Journal of Real Estate Literature*, *13*(1), 1–44.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199–222.

Stamou, M., Mimis, A., & Rovolis, A. (2017). House price determinants in Athens: a spatial econometric approach. *Journal of Property Research*, *34*(4), 269–284.

Theisen, T., & Emblem, A. W. (2018). House prices and proximity to kindergarten – costs of distance and external effects? *Journal of Property Research*, *35*(4), 321–343.

van Wezel, M., Kagie, M. M., & Potharst, R. R. (2005). *Boosting the accuracy of hedonic pricing models.*

Wyman, D., & Mothorpe, C. (2018). The pricing of power lines: A geospatial approach to measuring residential property values. *Journal of Real Estate Research*, *40*(1), 121–154.

Yao, Y., Zhang, J., Hong, Y., Liang, H., & He, J. (2018). Mapping fine-scale urban housing prices by fusing remotely sensed imagery and social media data. *Transactions in GIS*, *22*(2), 561–581.

Yoo, S., Im, J., & Wagner, J. E. (2012). Variable selection for hedonic model using machine learning approaches: A case study in Onondaga County, NY. *Landscape and Urban Planning*, *107*(3), 293–306.

Zhao, Q., & Hastie, T. (2021). Causal Interpretations of Black-Box Models. *Journal of Business & Economic Statistics*, *39*(1), 272–281.

Zheng, S., Hu, W., & Wang, R. (2016). How much is a good school worth in Beijing? Identifying price premium with paired resale and rental data. *The Journal of Real Estate Finance and Economics*, *53*(2), 184–199.

Zurada, J., Levitan, A., & Guan, J. (2011). A comparison of regression and artificial intelligence methods in a mass appraisal context. *Journal of Real Estate Research*, *33*(3), 349–387.