Finance Research Letters 58 (2023) 104306

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Exploring XAI techniques for enhancing model transparency and interpretability in real estate rent prediction: A comparative study

Ian Lenaers *, Lieven De Moor

Vrije Universiteit Brussel, Faculty of Social Sciences and Solvay Business School, Department of Business, Pleinlaan 2, B-1050, Brussels, Belgium

A R T I C L E I N F O

Keywords:
Explainable artificial intelligence techniques
CatBoost
Machine learning
Rent prediction
Residential real estate rent market

A B S T R A C T

Black-box artificial intelligence (AI) models are popular in real estate research, but their lack of interpretability raises concerns. To address this, explainable AI (XAI) techniques have been applied to shed light on these models. This paper presents a comparative study of six global XAI techniques on a CatBoost model for Belgian residential rent prediction. Results show that while some techniques offer substitute insights, others provide complementary perspectives on the model's behavior. Employing multiple XAI techniques is crucial to comprehensively understand rents drivers which contributes to transparency, interpretability, and model governance in the real estate industry, advancing the adoption of (X)AI.

1. Introduction

The increasing adoption of artificial intelligence (AI) in the financial industry has brought both significant opportunities and challenges. While AI models offer remarkable predictive capabilities, their opaque black-box nature raises concerns about transparency, interpretability, and model governance (Adadi and Berrada, 2018; Barredo Arrieta et al., 2020; Das and Rad, 2020). Addressing these concerns is crucial for fostering trust, compliance with regulations, and ensuring ethical use of AI (Martens, 2022)

To address this opaqueness, the use of explainable AI (XAI) has emerged in finance research (Giudici, 2018), aiming to enhance model transparency and interpretability. XAI techniques provide insights into the patterns learned in AI models, enabling stakeholders to understand the factors driving predictions, assess model biases, and comply with regulatory guidelines. As AI models have gained popularity for their ability to capture complex non-linear relationships and interaction effects, providing accurate predictions in real estate (Iban, 2022; Krämer et al., 2021; Lenaers et al., 2023; Lorenz et al., 2022), model transparency and interpretability are of utmost importance to understand and determine the relation of the drivers of real estate prices (McCluskey et al., 2013).

Our objective is twofold. First, we evaluate the complementarity and substitutability of different XAI techniques in uncovering rent drivers, enabling real estate stakeholders, including investors, policymakers, regulators, and real estate professionals, to gain a comprehensive understanding of the patterns learned in models and to make better-informed decisions based on transparent and interpretable information. Secondly, we aim to contribute to the broader discourse on model transparency and interpretability via XAI in finance, ultimately enhancing model governance and promoting responsible AI practices for advancing the adoption of (X)AI in finance.

To achieve these objectives, we conducted an analysis with XAI techniques on a CatBoost model – a tree-based model which is becoming the state-of-the-art successor of the XGBoost model – trained on a dataset comprising 78,788 properties in Belgium in 2022,

* Corresponding author.
E-mail address: Ian.Dave.J.Lenaers@vub.be (I. Lenaers).

https://doi.org/10.1016/j.frl.2023.104306
Received 10 May 2023; Received in revised form 20 July 2023; Accepted 6 August 2023
Available online 8 August 2023
1544-6123/

encompassing 24 features related to location factors and structural characteristics. In our study, we selected XAI techniques that were used in previous research but not compared to each other. Mora-Garcia et al. (2022) applied Permutation Feature Importance (PFI) and Partial Dependence Plots (PDPs), Kramer et al. (2021) applied PFI and Accumulated Local Effects (ALE) plots and Lorenz et al. (2022) applied PFI, ALE plots and PDPs. Iban (2022) and Lenaers et al. (2023) applied SHapley Additive exPlanations (SHAP) Feature Importance (FI) and SHAP Summary plots. Additionally, this study includes SHAP Dependence Plots (DPs) because of the shared mathematical basis with SHAP FI and Summary plots.

The remainder of this paper is structured as follows. Section 2 explains the methods applied. Section 3 presents the dataset. Section 4 presents and discusses the results. Section 5 concludes with a summary of the key findings.

2. Methodology

2.1. XAI techniques

2.1.1. PFI

PFI measures the importance of features in a model by calculating the increase in prediction error after permuting a feature (Fisher et al., 2019). Important features will increase model error more strongly, meaning that the prediction model heavily relies on that feature. Formally, the PFI_j for feature j is computed as follows (Molnar, 2022):

$$PFI_j = \frac{1}{n}\sum_{i=1}^{n}\left(L\left(y_i, \hat{y}_i^j\right) - L(y_i, \hat{y}_i)\right), \tag{1}$$

where n is the number of observations in the dataset, y_i the target variable for observation i, $\hat{y}_i$ the predicted target variable for observation i, $\hat{y}_i^j$ the predicted target variable for observation i after randomly permuting the feature j in the dataset and L the chosen loss function that quantifies the goodness-of-fit. Note that permuting data introduces randomness. Thus, the results of the procedure depend on the obtained configuration of permuted values. Hence, one performs the procedure several times such that the uncertainty associated with the calculated FIs will be assessed (Biecek and Burzykowski, 2021).

2.1.2. PDPs

The PDP shows the marginal effect that a feature has on the predicted outcome of a model (Biecek and Burzykowski, 2021). Thus, a PDP shows the relationship between the target variable and a feature (Masís, 2021). The value of a partial dependence function for model f and features x_S is defined as follows (Goldstein et al., 2014):

$$f_S(x_S) = E_{X_C}[f(x_S, X_C)] \tag{2}$$

where C is the complement set of S[1](i.e., x_S indicates the feature(s) for which the PD function should be plotted and X_C refers to the other features used in model f). Thus, it is the expected value of the model predictions when x_S is fixed over the (marginal) distribution of X_C, i.e., over the joint distribution of all features apart from x_S. However, the true distribution of x_S is unknown. As such, the partial dependence function $\hat{f}_S$ is estimated by calculating averages of actual feature values $x_C^{(i)}$ in the training data while keeping x_S constant:

$$\hat{f}_S(x_S) = \frac{1}{n}\sum_{i=1}^{n}\hat{f}\left(x_S, x_C^{(i)}\right). \tag{3}$$

However, disadvantages of PDPs are that they assume independence between features (Masís, 2021) and do not reflect whether some regions have fewer data points or not (Molnar, 2022). If this is the case, PDPs are unreliable and biased. These limitations are addressed in ALE plots. For an in-depth description, interested readers are referred to Goldstein et al. (2014).

2.1.3. ALE plots

Where PDPs yields biased results if the features are strongly correlated, ALE plots alleviate this issue by averaging the changes in the predictions for each feature and accumulating them over a grid. The ALE function $f_{S,ALE}$ for observation x of feature S, x_S, is defined as (Biecek and Burzykowski, 2021):

$$f_{S,ALE}(x_S) = \int_{z_{0,S}}^{x_S} E_{X_C|x_S}\left[\hat{f}^S(x_S, X_C)|x_S = z_S\right]dz_S - c, \tag{4}$$

where c is a constant and $E_{X_C|x_S}[\hat{f}^S(x_S, X_C)|x_S = z_S]$ the expected value of the model's predictions, given the input value x_S and all possible values of the other features, denoted by X_C. dz_S is the differential change in the feature S over which one integrates and $z_{0,S}$ the minimum value of the chosen feature S. Interested readers are directed to the work of Apley & Zhu (2019) for a detailed explanation of estimation. However, there are still disadvantages, e.g. grid selection may influence the plots and can mask heterogeneity in the data

[1] In this study we focus on the interpretation of one feature per plot, so S comprises only one feature.

Table 1
Summary statistics for the target variable and features.

Feature name	Data type	Nr. of observations	Min.	Max.	Mean	Median	St. Dev.	Skew.	Kurt.	Cor. with target var.
Monthly rent price (in EUR)	Continuous	78,788	370	4100	899.55	800	361.89	2.76	12.41	
Area (in m^2)	Continuous	59,644	62	5332	704.55	358	890.64	2.54	6.96	0.00
Build year	Continuous	42,167	1700	2023	1991.03	2006	35.26	−1.72	4.44	−0.11
Building area (in m^2)	Continuous	58,366	25	2637	237.60	128	317.55	3.68	16.49	0.02
Distance to bus stop (in m)	Continuous	69,273	1	5008	184.96	141	195.27	5.44	51.61	0.08
Distance to village center (in m)	Continuous	69,273	1	13,948	1487.24	1019	1501.01	2.34	7.05	0.00
Distance to school (in m)	Continuous	69,273	0	7857	353.34	246	365.17	3.72	25.77	0.06
Distance to train station (in m)	Continuous	69,273	0	13,876	1922.89	1210	1946.01	1.87	3.32	0.01
Double glass (0 = no, 1 = yes)	Binary	78,788	0	1	0.71	1	0.45	−0.92	−1.15	0.07
Energy consumption (in kWh/m^2 per year)	Continuous	49,200	8	1482	225.37	188	157.84	1.85	5.74	−0.07
Floor	Integer	65,384	0	38	1.44	1	1.85	3.37	24.75	−0.08
Garden (0 = no, 1 = yes)	Binary	78,788	0	1	0.65	1	0.48	−0.62	−1.62	0.08
Garden area (in m^2)	Continuous	45,881	24	3970	484.57	222	665.27	2.65	7.84	−0.02
Habitable area (in m^2)	Continuous	78,788	24	510	101.92	91	45.43	1.56	3.98	0.54
Housing type (0 = apartment, 1 = house)	Binary	78,788	0	1	0.26	0	0.44	1.12	−0.74	0.29
*Mobility score**	Continuous	78,788	0.25	1	0.80	0.82	0.12	−0.95	0.71	−0.02
New build (0 = no, 1 = yes)	Binary	78,788	0	1	0.13	0	0.34	2.18	2.77	0.03
Number of bathrooms	Integer	71,429	1	4	1.26	1	0.52	2.16	5.21	0.45
Number of bedrooms	Integer	78,788	1	6	2.00	2	0.89	0.87	0.94	0.52
Number of floors	Integer	44,678	0	43	2.97	3	2.57	4.50	43.77	0.09
Number of parking spaces	Integer	38,489	0	20	1.38	1	1.48	5.06	44.92	0.14
Number of sides (1, 2, 3, 4)	Ordinal	43,905	1	4	2.44	2	0.78	0.75	−0.19	0.17
Number of toilets	Integer	61,774	0	4	1.27	1	0.57	1.10	2.14	0.50
*Proneness to flooding***	Nominal	58,693	/	/	/	/	/	/	/	/
Solar panels (0 = no, 1 = yes)	Binary	78,788	0	1	0.08	0	0.27	3.11	7.65	0.03

* The mobility score indicates the environmental impact for travel from one's property. It shows how well facilities such as schools, stores, public transport, … are accessible by bike or on foot. ** Categories are yes, possible and no.

(Molnar, 2022). Furthermore, even though ALE plots are not biased in the case of strongly correlated features, interpretation remains difficult.

2.1.4. *Global SHAP interpretations*

SHAP are a model-agnostic technique that helps to understand the importance of features in making a prediction. SHAP are developed by Lundberg & Lee (2017) and extend from the concept of Shapley values in game theory to fairly divide player contributions when they collectively achieve a given outcome. The SHAP of feature j for observation x, $\phi_j(x)$, is defined as:

$$\phi_j(x) = \sum_{S\subseteq\{1,\ \dots,M\}\backslash\{j\}} \frac{|S|!(M-|S|-1)!}{M!}[f(x_S\cup\{j\})-f(x_S)], \tag{5}$$

where j is the feature evaluated, M the total number of features, S a subset of the full feature set $\{1, \dots, M\}$ that does not include the feature j, x_S a subset of features in S, and f the model's prediction function.

Aggregations of SHAP offer global interpretations, which help understand the overall behavior of a model (Molnar, 2022). First, the SHAP FI measures the importance of each feature in a model by computing the average absolute SHAP value for that feature across the dataset. The SHAP FI for feature j, $SHAPFI_j$, is defined as (Molnar, 2022):

$$SHAP\ FI_j = \frac{1}{n}\sum_{i=1}^{n}\left|\phi_j(x)\right|, \tag{6}$$

where n is the number of observations in the dataset, $\phi_i(x)$ the SHAP for feature j of observation x, and the absolute value is taken to ensure that the importance scores are positive. Second, the SHAP Summary plot provides an overall summary of the model's behavior by showing the distribution of SHAP across all observations in the dataset (Molnar, 2022). Third, the SHAP DP is a scatter plot that shows the effect a single feature has on the predictions made by the model, while accounting for the interactions and dependencies between the other features (Molnar, 2022).

3. Data and descriptive statistics

To illustrate our approach, we use a dataset that contains 24 features and the target variable monthly rent of 78,788 Belgian residential real estate properties from January 1st, 2022, until December 31st, 2022. To provide insight into the distribution of the features and their (linear) relationship with the target variable, summary statistics are provided in Table 1.

Table 2
Evaluation metrics for the CatBoost model.

	Model	MAE	MedAE	RMSE	MAPE	MedAPE
Train dataset	*Ridge*	181.06	131.95	260.01	20.03%	15.94%
	XGBoost	108.49	81.46	149.18	12.53%	9.73%
	CatBoost	130.54	94.70	183.84	14.73%	11.47%
Test dataset	*Ridge*	179.06	132.29	254.09	20.16%	15.99%
	XGBoost	143.47	100.42	212.63	16.06%	12.23%
	CatBoost	142.41	98.45	210.84	15.75%	11.98%

with MAE = mean average error, MedAE = median average error, RMSE = root mean square error, MAPE = mean average percentage error, and MedAPE = median average percentage error.

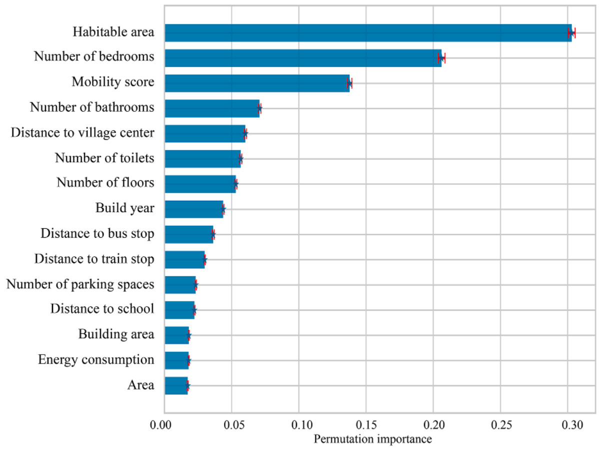

Fig. 1. PFI plot for the CatBoost model.

We sort the features by decreasing FI and plot them for both techniques. Remark that some features can be more important in PFI, while others can be more important in SHAP FI. Both techniques use different computational methods, leading to variations in feature importance and thus in the ranking of features. Thus, some features that are in the top 15 PFI do not appear in the top 15 SHAP FI and vice versa. To calculate the PFIs, MAPE was used as the loss function and the procedure was executed 100 times to create an idea of variability in the results due to randomness.

4. Results

The results are generated in python (version 3.8.15), using PyCaret (version 3.0.0) for model training, optuna (version 3.0.5) for hyperparameter tuning, shap (version 0.41.0), sklearn (version 1.1.3) and alibi (version 0.9.1) for XAI visualizations. Following Molnar (2022), the visualizations are created using the train dataset to gain an insight into the patterns learned.

4.1. Data pre-processing, model training and model evaluation

Before model training, we preprocess the data. We perform a classical train-test split, using 80% of the dataset for model training and 20% for model evaluation. To address missing values, we utilize a simple imputation method. This method involves filling in the missing values by the mean for continuous and integer features, and the mode for categorical and binary features. After imputing missing values, we process categorical features by creating dummy variables using one-hot encoding.

The CatBoost model is trained using the train dataset with hyperparameter tuning and 10-fold cross-validation to reduce the likelihood of overfitting. To improve the efficiency and effectiveness of the hyperparameter tuning process, we use the tree-structured parzen estimator, a Bayesian optimization approach, that has been shown to outperform traditional hyperparameter tuning methods (Yang and Shami, 2020). During the hyperparameter tuning process, we use the mean absolute percentage error (MAPE) as scoring metric.

While the focus of the study is not to demonstrate predictive performance of the CatBoost model, we nevertheless provide evaluation metrics regarding goodness-of-fit of the model for both train and test dataset in Table 2 to strengthen the credibility of both the model and the subsequent interpretations with XAI techniques. This table also reports metrics for two commonly used models in real estate, namely a ridge model (Hinrichs et al., 2021) and an XGBoost model (Iban, 2022; Krämer et al., 2021; Lenaers et al., 2023; Lorenz et al., 2022; Mora-Garcia et al., 2022) to put the CatBoost model in perspective. Relatively small differences are observed for the reported metrics of the CatBoost model between train and test dataset. The XGBoost model has similar metrics to the CatBoost model on the test dataset but note that the differences between train and test dataset are large, indicating overfitting. In addition, the ridge model seems not to overfit, but generally achieves poorer performances due to the lack of capturing non-linearities and interaction

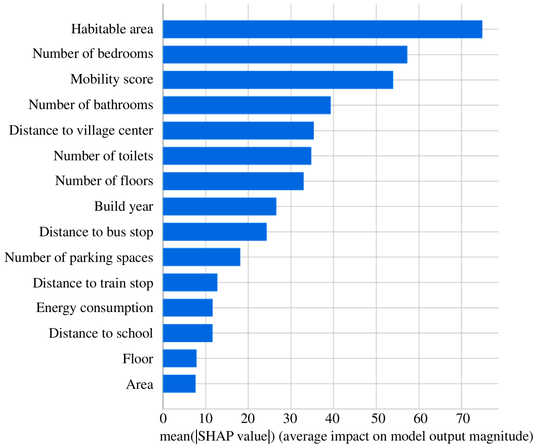

Fig. 2. SHAP FI plot for the CatBoost model.
We sort the features by decreasing FI and plot them for both techniques. Remark that some features can be more important in PFI, while others can be more important in SHAP FI. Both techniques use different computational methods, leading to variations in feature importance and thus in the ranking of features. Thus, some features that are in the top 15 PFI do not appear in the top 15 SHAP FI and vice versa. To calculate the PFIs, MAPE was used as the loss function and the procedure was executed 100 times to create an idea of variability in the results due to randomness. Note that the TreeExplainer was used for the SHAP visualizations for the CatBoost model as it generates computationally faster results than the KernelExplainer in shap (version 0.41.0).

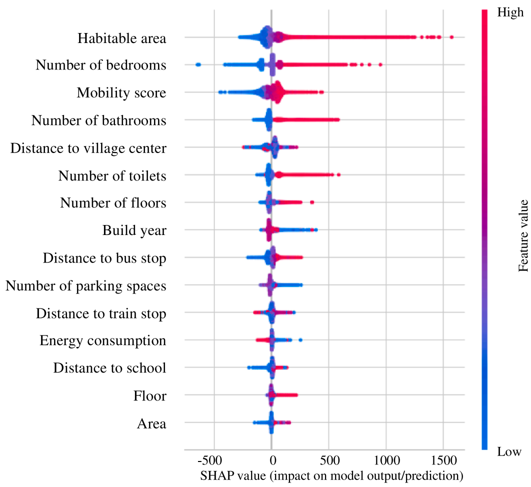

Fig. 3. SHAP Summary plot for the CatBoost model.
The SHAP Summary plot shows a horizontal bar chart for each feature, indicating the magnitude and direction of its impact on the model's output. The features are sorted according to their importance, with the most influential features at the top. The color of each bar represents the value of the corresponding feature, with red indicating high values and blue indicating low values. Overlapping points are jittered in y-axis direction, so we get a sense of the distribution of the SHAP per feature. Note that the TreeExplainer was used for the SHAP visualizations for the CatBoost model as it generates computationally faster results than the KernelExplainer in shap (version 0.41.0).

effects, as opposed to tree-based models.

4.2. Comparison of XAI techniques

Our analysis demonstrates complementarity and substitutability of XAI techniques for the CatBoost model, which allows for non-linearities and interaction effects. Figs. 1 and 2 show the results of feature importance. We find that habitable area, number of bedrooms and mobility score are the most important features in predicting rents, according to the PFI. This finding is confirmed by the SHAP FI, although this latter makes them seem less distinct. From this, it is argued that PFI and SHAP FI are substitutes, yet they impose slightly different nuances due to the underlying calculation methods. In addition, the SHAP Summary plot in Fig. 3 confirms the former plots and offers additional insights on the direction of relationships and indications of non-linearity.

To delve deeper into non-linearity and get a full understanding of the relationships, we use PDPs, ALE plots and SHAP DPs. In

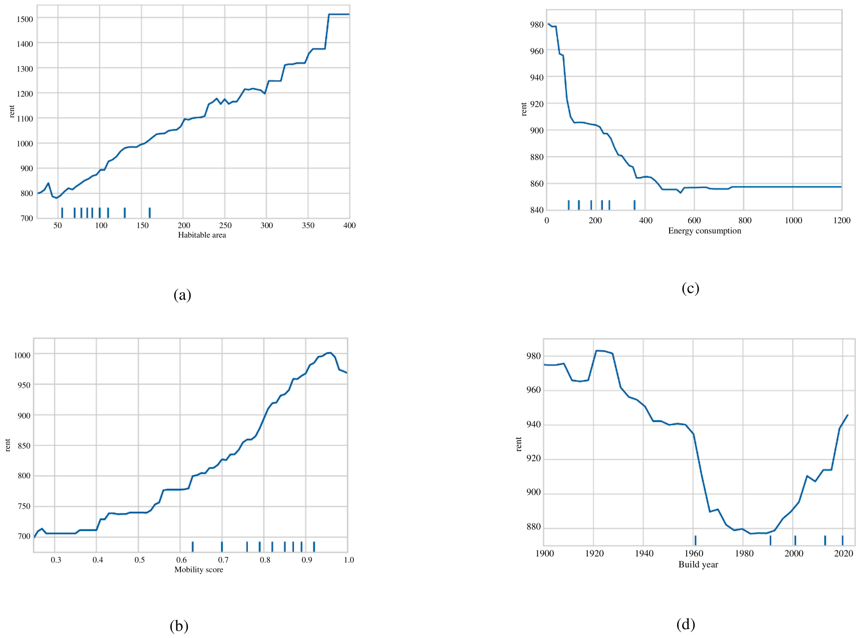

Fig. 4. PDPs for the CatBoost model for selected features.

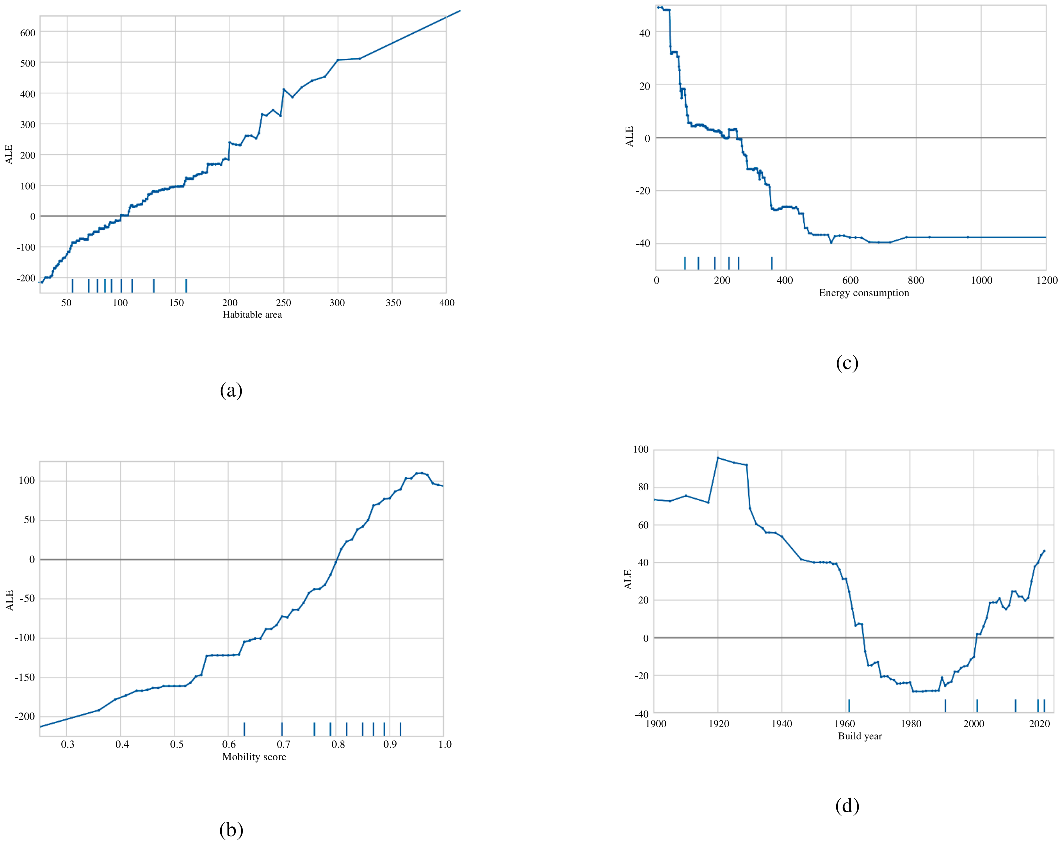

Fig. 5. ALE plots for the CatBoost model for selected features.

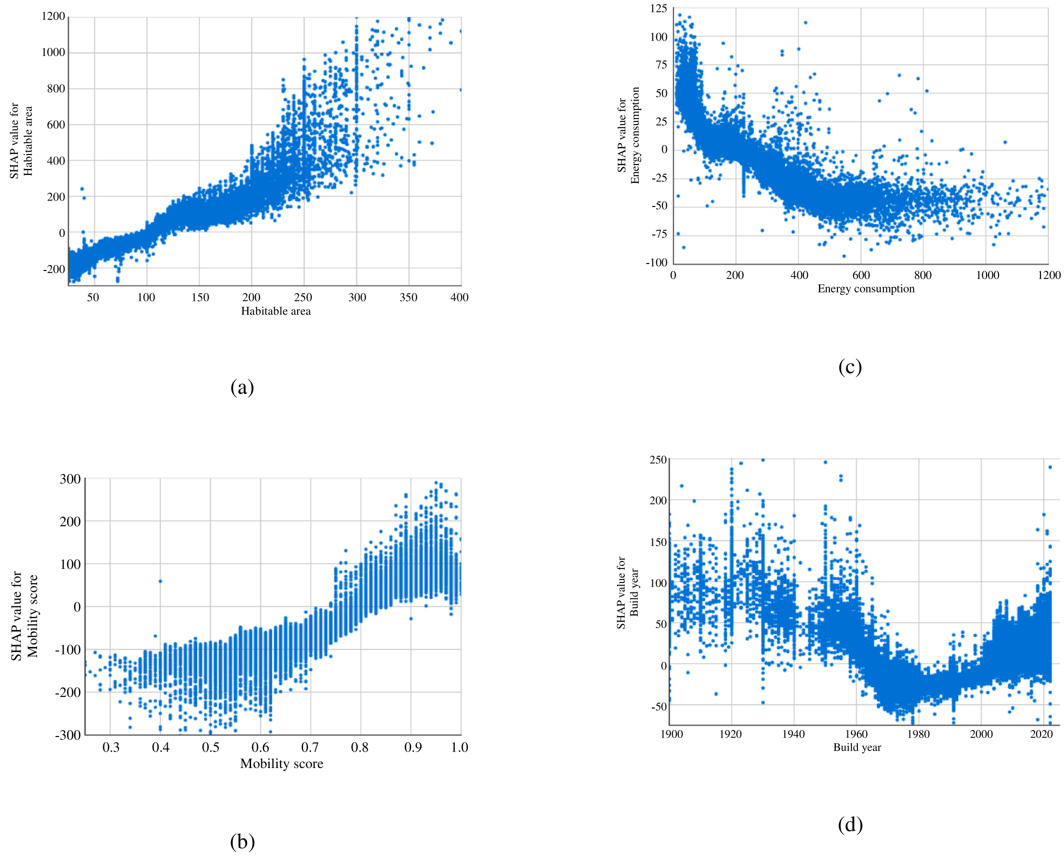

Fig. 6. SHAP DPs* for the CatBoost model for selected features.
* The SHAP DP shows the values of the feature on the x-axis and the corresponding SHAP on the y-axis, with each point representing a single observation in the train dataset. The SHAP indicate the contribution of the feature to the difference between the actual prediction and the average predicted rent. Note that the TreeExplainer was used for the SHAP visualizations for the CatBoost model as it generates computationally faster results than the KernelExplainer in shap (version 0.41.0).

Figs. 4, 5 and 6, these techniques consistently show both linear and non-linear relationships between selected features and rent. For example, habitable area shows a positive linear relationship with increasing variability for higher areas, indicating that larger properties command higher rents. Mobility score shows a positive relationship with predicted rent. Energy consumption shows a non-linear relationship, with a strong decreasing relation initially, followed by a stable decreasing relation and finally a constant negative relation. Build year shows a U-shaped relationship, suggesting that newer homes meeting modern standards and renovated older homes yield higher rents. This validates previous research by Iban (2022), Krämer et al. (2021), Lenaers et al. (2023), Lorenz et al. (2022) and Mora-Garcia et al. (2022), indicating the presence of non-linearities and the need to use non-linear models to capture relationships in real estate modeling. While PDPs and ALE plots show substitute views, SHAP DPs provide a more detailed perspective by illustrating the direction, magnitude, and variability of relationships. As such, PDPs, ALE plots and SHAP DPs are seen as complementary to PFI, SHAP FI and Summary plots. PDPs are considered a substitute for ALE plots. Finally, SHAP DPs are seen as complementary. Together, these XAI techniques provide a comprehensive understanding of the importance and the relationships of features, offering different levels of granularity to understand the complex dynamics between rent and its determinants in the model.

5. Conclusions

Our study employed six XAI techniques to enhance model transparency and interpretability in the real estate industry. Through an analysis of a CatBoost model trained with Belgian residential rental data, consisting of 78,788 properties and 24 features in 2022, we revealed valuable insights into the patterns learned in the model.

By comparing and examining various XAI techniques, we found that they both complemented and substituted each other, offering unique perspectives and insights into model. While PFI and SHAP FI are substitute measures of feature importance, the SHAP Summary plot provides complementary insights into the direction of the relationships between rents and its determinants. Similarly, PDPs and ALE plots offer substitute views of the relationships while SHAP DPs offer complementary perspectives by illustrating the variability of relationships.

Our findings highlight the importance of employing multiple XAI techniques to gain a comprehensive understanding of the factors driving rent. We advocate for the inclusion of an additional interpretation step in modeling frameworks to apply a combination of XAI techniques, as each technique offers unique insights and perspectives. This pluralistic interpretation will empower stakeholders in the real estate industry to make informed decisions, assess risks, and analyze market dynamics more effectively. By enhancing the transparency and interpretability of rent prediction models, we strengthen model governance practices within financial services organizations. This fosters trust, reduces potential biases, and promotes responsible use of AI technologies, which will advance the adoption of (X)AI in the real estate industry.

Our study contributes to the growing body of knowledge on transparent and responsible AI practices in the domain of finance. However, further validation of our methodology is encouraged by applying it to other real estate or financial-economic datasets and models to deepen our understanding of the substitutability, complementarity, benefits, and limitations of XAI techniques in finance.

Availability of data and material

The data was provided by Realo N.V. which is located at Poel 16, 9000 Ghent, Belgium.

CRediT authorship contribution statement

Ian Lenaers: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Lieven De Moor:** Resources, Supervision, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share the data, however data can be made available upon request.

References

Adadi, A., Berrada, M., 2018. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access 6, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052.

Apley, D.W., Zhu, J., 2019. *Visualizing the Effects of predictor variables in black box supervised learning models* (arXiv:1612.08468). arXiv. http://arxiv.org/abs/1612.08468.

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., 2020. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012.

Biecek, P., Burzykowski, T., 2021. Explanatory Model Analysis: Explore, Explain, and Examine Predictive Models, 1st edition. Chapman and Hall/CRC.

Das, A., Rad, P., 2020. *Opportunities and challenges in explainable artificial intelligence (XAI): a survey* (arXiv:2006.11371). arXiv. https://doi.org/10.48550/arXiv.2006.11371.

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* (arXiv:1801.01489). arXiv. https://doi.org/10.48550/arXiv.1801.01489.

Giudici, P., 2018. Fintech risk management: a research challenge for artificial intelligence in finance. Front. Artificial Intelligence 1. https://www.frontiersin.org/articles/10.3389/frai.2018.00001.

Goldstein, A., Kapelner, A., Bleich, J., Pitkin, E., 2014. *Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation* (arXiv:1309.6392). arXiv. https://doi.org/10.48550/arXiv.1309.6392.

Hinrichs, N., Kolbe, J., Werwatz, A., 2021. Using shrinkage for data-driven automated valuation model specification – a case study from Berlin. J. Property Res. 38 (2), 130–153. https://doi.org/10.1080/09599916.2021.1905690.

Iban, M.C., 2022. An explainable model for the mass appraisal of residences: the application of tree-based machine learning algorithms and interpretation of value determinants. Habitat Int. 128, 102660 https://doi.org/10.1016/j.habitatint.2022.102660.

Krämer, B., Stang, M., Nagl, C., & Schäfers, W. (2021). *Explainable AI in a real estate context—exploring the determinants of residential real estate values* (SSRN scholarly paper 3989721). 10.2139/ssrn.3989721.

Lenaers, I., Boudt, K., De Moor, L., 2023. Predictability of Belgian residential real estate rents using tree-based ML models and IML techniques. *Int. J. Housing Mark. Anal. Ahead-of-Print*(ahead-of-print). https://doi.org/10.1108/IJHMA-11-2022-0172.

Lorenz, F., Willwersch, J., Cajias, M., Fuerst, F., 2022. Interpretable machine learning for real estate market analysis. Real Estate Econ. 0 (0), 1–31. https://doi.org/10.1111/1540-6229.12397.

Lundberg, S., Lee, S.I., 2017. *A unified approach to interpreting model predictions* (arXiv:1705.07874). arXiv. https://doi.org/10.48550/arXiv.1705.07874.

Martens, D., 2022. Data Science Ethics: Concepts, Techniques, and Cautionary Tales. Oxford University Press.

Masís, S., 2021. Interpretable Machine Learning With Python: Learn to Build Interpretable High-Performance Models With Hands-On Real-World Examples. Packt Publishing.

McCluskey, W.J., McCord, M., Davis, P.T., Haran, M., McIlhatton, D., 2013. Prediction accuracy in mass appraisal: a comparison of modern approaches. J. Property Res. 30 (4), 239–265. https://doi.org/10.1080/09599916.2013.781204.

Molnar, C., 2022. Interpretable machine learning: a guide for making black box models explainable. Independently Published.
Mora-Garcia, R.T., Cespedes-Lopez, M.F., Perez-Sanchez, V.R., 2022. Housing price prediction using machine learning algorithms in COVID-19 times. Land (Basel) 11 (11), 11. https://doi.org/10.3390/land11112100. Article.
Yang, L., Shami, A., 2020. On hyperparameter optimization of machine learning algorithms: theory and practice. Neurocomputing 415, 295–316. https://doi.org/10.1016/j.neucom.2020.07.061.