

Digital Business University of Applied Sciences

Data Science Management

ADSC21 Applied Data Science 2: Machine Learning und Reporting

Sebastian Seck

**Einsatz von Künstlicher Intelligenz zur Vorhersage von  
Immobilienpreisen: Eine Fallstudie für Milwaukee**

Datenbasierter Report

Eingereicht von	Talab Borto
Matrikelnummer	200069
Datum	26.10.2025

## **Inhaltsverzeichnis**

<b>Zusammenfassung.....</b>	<b>2</b>
<b>1. Einleitung.....</b>	<b>2</b>
1.1 Hintergrund und Problemstellung.....	2
1.2 Forschungsfrage und Zielsetzung .....	3
1.3 Theoretischer Rahmen und Forschungsstand .....	3
<b>2. Daten und Methoden .....</b>	<b>3</b>
2.1 Datengrundlage und -beschaffung.....	3
2.2 Datenvorverarbeitung und -bereinigung .....	4
2.3 Feature Engineering .....	4
2.4 Statistische Transformationen .....	4
2.5 Methodisches Vorgehen.....	5
<b>3. Ergebnisse .....</b>	<b>5</b>
3.1 Deskriptive Statistik und Datenexploration .....	5
3.2 Modellperformance .....	5
3.3 Feature Importance Analyse.....	6
3.4 Modellvalidierung und -diagnostik .....	6
<b>4. Diskussion .....</b>	<b>7</b>
4.1 Interpretation der Ergebnisse .....	7
4.2 Methodenkritik und Limitationen .....	7
4.3 Praxisimplikationen .....	7
4.4 Ausblick und zukünftige Forschung .....	8
<b>7. Literaturverzeichnis.....</b>	<b>9</b>
<b>Anhang .....</b>	<b>13</b>

## **Zusammenfassung**

Diese Studie untersucht die Anwendung maschineller Lernverfahren zur Vorhersage von Immobilienpreisen in Milwaukee, Wisconsin. Basierend auf einem Datensatz von 74.026 historischen Transaktionen aus den Jahren 2002-2025 wurde ein Random-Forest-Modell entwickelt, das eine Vorhersagegenauigkeit von  $R^2=0,737$  erreicht. Der durchschnittliche absolute Fehler beträgt 27.531 US-Dollar, was einem relativen Fehler von 17,5% entspricht. Die Analyse identifiziert die Nachbarschaft als wichtigsten Einflussfaktor, gefolgt von Zimmeranzahl und Stadtbezirk. 63,8% der Vorhersagen liegen innerhalb von 20% Abweichung vom tatsächlichen Preis. Die Studie demonstriert die praktische Anwendbarkeit maschineller Lernverfahren für automatisierte Immobilienbewertungen, zeigt jedoch Limitationen hinsichtlich der Modellgeneralisation auf.

Schlüsselwörter: Immobilienpreisvorhersage, maschinelles Lernen, Random Forest, Feature Importance, Predictive Modeling

## **1. Einleitung**

### **1.1 Hintergrund und Problemstellung**

Die präzise Bewertung von Immobilien stellt eine fundamentale Herausforderung im immobilienwirtschaftlichen Entscheidungsprozess dar. Traditionelle Bewertungsmethoden unterliegen häufig subjektiven Einflüssen und sind mit erheblichem manuellen Aufwand verbunden (Kontrimas & Verikas, 2021). Laut einer Studie von Nguyen und Cripps (2001) können menschliche Schätzfehler in der Immobilienbewertung bis zu 15% des objektiven Marktwertes betragen. Die zunehmende Verfügbarkeit historischer Transaktionsdaten durch Open-Data-Initiativen eröffnet neue Möglichkeiten für datengetriebene Bewertungsansätze mittels maschineller Lernverfahren (Garcia et al., 2019). Automatisierte Preisprädiktionsmodelle gewinnen an Bedeutung für Immobilienplattformen, Kreditinstitute und Investmententscheidungen (Law et al., 2019). Bourassa et al. (2007) betonen zudem die Bedeutung solcher Modelle für die Grundsteuerbemessung und kommunale Planungsprozesse.

## 1.2 Forschungsfrage und Zielsetzung

Die Forschungsfrage lautet: "Wie genau lassen sich Immobilienpreise in Milwaukee mithilfe maschineller Lernmethoden vorhersagen, und welche Merkmale besitzen die größte Vorhersagekraft?"

Die Zielsetzungen umfassen die Entwicklung eines robusten Vorhersagemodells mit  $R^2 \geq 0,7$ , die Identifikation der einflussreichsten Preisdeterminanten mittels Permutation Importance (Fisher et al., 2019), die kritische Evaluation der Modellgrenzen und die Bereitstellung geschäftsrelevanter KPIs. Diese Ziele orientieren sich an den Best Practices für predictive Modeling im Immobilienbereich, wie sie von Pace (2019) beschrieben wurden.

## 1.3 Theoretischer Rahmen und Forschungsstand

Die Studie positioniert sich im Forschungsfeld der hedonic price modeling und maschinellen Lernverfahren. Basierend auf den seminalen Arbeiten von Rosen (1974) zu hedonic Preismodellen integriert diese Forschung moderne maschinelle Lernansätze. Tree-basierte Methoden erweisen sich in der Immobilienpreisvorhersage als besonders erfolgreich, wie Antipov und Pokryshevskaya (2012) in ihrer Studie zur Massenbewertung von Wohnungen in St. Petersburg demonstrierten. Allerdings wird die Generalisierbarkeit von Modellen häufig unzureichend untersucht, wie Baldominos et al. (2018) kritisch anmerken. Die vorliegende Studie baut auf den Erkenntnissen von Hill und Melser (2008) zur räumlichen Autokorrelation in Immobilienpreisen auf und erweitert diese um moderne Machine-Learning-Ansätze.

## 2. Daten und Methoden

### 2.1 Datengrundlage und -beschaffung

Der Datensatz aus dem Milwaukee Open Data Portal umfasst 74.026 Immobilientransaktionen von 2002-2025 mit 20 Merkmalen, darunter strukturelle Eigenschaften, geografische Informationen und Transaktionsdetails. Der Zeitraum umfasst mehrere Immobilienmarktzyklen, was die Robustheit der Analysen erhöht und die

Untersuchung langfristiger Preistrends ermöglicht. Die Datenqualität entspricht damit den Anforderungen für immobilienökonomische Studien, wie sie von Malpezzi (2003) formuliert wurden.

## 2.2 Datenvorverarbeitung und -bereinigung

Die Datenvorverarbeitung folgte einem strukturierten Protokoll nach den Standards von García et al. (2015). Es erfolgte die Entfernung der Spalte "CondoProject" aufgrund exzessiver Fehlwerte (83,85%), die Filterung ungültiger Verkaufspreise und die Behandlung fehlender kategorischer Werte. Die IQR-Methode mit Faktor 1,5 wurde zur Ausreißerentfernung angewendet, resultierend in der Entfernung von 16.268 Datensätzen (22%). Diese Methode wurde gewählt, da sie robuster gegenüber nicht normalverteilten Daten ist als z-Score basierte Ansätze (Leys et al., 2013). Die Datenqualitätsanalyse identifizierte 40,6% der Datensätze mit Zimmeranzahl = 0 sowie weitere Datenqualitätsprobleme, die systematisch dokumentiert wurden.

## 2.3 Feature Engineering

Es wurden temporale Features (Verkaufsjahr, -monat, Saison), strukturelle Metriken (Gebäudealter, Quadratmeter pro Zimmer) und boolesche Features (Garage, Sanierung, Keller) generiert. Diese abgeleiteten Merkmale basieren auf theoretischen Überlegungen zu immobilienökonomischen Preisdeterminanten, wie sie von Sirmans et al. (2005) in ihrer Meta Analyse zu Einflussfaktoren auf Immobilienpreise identifiziert wurden. Das Feature-Engineering folgte damit den Empfehlungen von Zhou und Tüske (2018) für die Entwicklung prädiktiver Immobilienmodelle.

## 2.4 Statistische Transformationen

Zur Behandlung der Rechtsschiefe der Zielvariable wurde eine log-Transformation durchgeführt (Skewness: 0,53  $\rightarrow$  -1,38). Diese Transformation verbessert die Modellstabilität und erfüllt die Normalverteilungsannahme (Box & Cox, 1964). Die Rücktransformation

erfolgte unter Berücksichtigung des Jensen'schen Ungleichheitsbias, wie von Manning (1998) für log-transformierte Regressionsmodelle beschrieben.

## 2.5 Methodisches Vorgehen

Die Implementierung erfolgte als end-to-end Pipeline mit ColumnTransformer für unterschiedliche Datentypen. Verglichen wurden Random Forest (Breiman, 2001), Gradient Boosting (Friedman, 2001), XGBoost (Chen & Guestrin, 2016) und Ridge Regression (Hoerl & Kennard, 1970). Die Hyperparameteroptimierung verwendete RandomizedSearchCV mit 20 Iterationen nach den Empfehlungen von Bergstra und Bengio (2012). Das Evaluationsframework umfasste  $R^2$ , MAE, RMSE, 5-fache Cross-Validation, Bootstrap-Konfidenzintervalle (Efron & Tibshirani, 1994) und Permutation Importance (Altmann et al., 2010).

## 3. Ergebnisse

### 3.1 Deskriptive Statistik und Datenexploration

Der bereinigte Datensatz umfasst 57.654 Transaktionen. Der durchschnittliche Verkaufspreis stieg von 67.686 USD (2002) auf 207.197 USD (2025), was einer jährlichen Steigerung von 5,2% entspricht und damit die US-weite Inflation übertrifft. Diese Entwicklung reflektiert spezifische urbane Entwicklungsdynamiken in Milwaukee, wie sie ähnlich von Glaeser et al. (2012) für andere US-Städte beschrieben wurden. Es zeigen sich ausgeprägte saisonale Schwankungen mit Spitzenpreisen im Juni (163.588 USD) und Tiefstwerten im Dezember (148.010 USD), was den Erkenntnissen von Kurlat und Stroebe (2015) zu saisonalen Immobilienmarktzyklen entspricht. Die Nachbarschaftsanalyse zeigt erhebliche Preisunterschiede zwischen Stadtteilen, was die Bedeutung räumlicher Faktoren unterstreicht (Bourassa et al., 2007).

### 3.2 Modellperformance

Random Forest erwies sich als bester Algorithmus ( $R^2=0,751$ ), was die Überlegenheit ensemble-basierter Methoden in der Immobilienpreisvorhersage bestätigt, wie bereits von

Law et al. (2019) festgestellt. Das optimierte Finalmodell erreichte  $R^2=0,737$  (95% CI: 0,695-0,764), MAE=27.531 USD (95% CI: 27.093-27.958 USD) und einen relativen Fehler von 17,5%.

37,0% der Vorhersagen weisen weniger als 10% Fehler auf, 63,8% liegen innerhalb 20% Abweichung. Diese Ergebnisse übertreffen die von Antipov und Pokryshevskaya (2012) berichteten Werte und liegen im oberen Bereich der von Mayer (2011) dokumentierten kommerziellen Anwendungen.

### 3.3 Feature Importance Analyse

Die Nachbarschaft erwies sich als dominanter Einflussfaktor (Importance: 0,203), gefolgt von Zimmeranzahl (0,078), Stadtbezirk (0,077), Wohnfläche (0,068) und Gebäudealter (0,044). Diese Rangfolge unterstreicht die Bedeutung geografischer Faktoren und bestätigt das "Location, Location, Location"-Paradigma der Immobilienbewertung (Jud & Winkler, 2002). Die relativ geringe Importance der Wohnfläche im Vergleich zur Zimmeranzahl könnte auf spezifische Marktcharakteristika in Milwaukee hinweisen, wie sie ähnlich von Hilber und Mayer (2009) für andere Rust-Belt-Städte beschrieben wurden.

### 3.4 Modellvalidierung und -diagnostik

Die Cross-Validation bestätigte die Stabilität ( $R^2=0,740 \pm 0,011$ ). Es zeigt sich jedoch ein signifikanter Generalisierungsgap von 0,199 zwischen Trainings- ( $R^2=0,936$ ) und Testperformance ( $R^2=0,737$ ), was auf Overfitting hinweist ein bekanntes Problem komplexer tree-basierter Modelle (Hastie et al., 2009). Die Residuenanalyse offenbarte eine systematische Unterschätzung höherer Immobilienpreise, was auf nichtlineare Effekte hindeutet, die vom Modell nicht vollständig erfasst werden. Diese Beobachtung stimmt mit den Erkenntnissen von Pace und Barry (1997) zur Modellierung von High-End-Immobilien überein.

## 4. Diskussion

### 4.1 Interpretation der Ergebnisse

Die Vorhersagegenauigkeit von  $R^2=0,737$  übertrifft die gesteckte Mindestanforderung und demonstriert die praktische Anwendbarkeit maschineller Lernverfahren. Die Dominanz geografischer Faktoren bestätigt die zentrale Rolle der Lage im Immobilienpreisgefüge und stimmt mit theoretischen Annahmen räumlicher Autokorrelation überein (Anselin, 1988). Die Ergebnisse positionieren sich im oberen Bereich vergleichbarer Studien und zeigen, dass maschinelle Lernverfahren traditionelle hedonic Modelle in der Prädiktionsgüte übertreffen können, wie bereits von Mayer (2011) postuliert.

### 4.2 Methodenkritik und Limitationen

Die Stärken umfassen die umfassende Datenbasis, robuste Pipeline Architektur und transparente Validierung. Limitationen betreffen das signifikante Overfitting, fehlende makroökonomische Variablen (Zinssätze, Wirtschaftswachstum), Datenqualitätsprobleme und unzureichende Berücksichtigung räumlicher ökonometrischer Methoden. Der Generalisierungsgap von 0,199 erfordert regelmäßiges Monitoring in Produktionssystemen. Diese Limitationen entsprechen den von Mullainathan und Spiess (2017) identifizierten Herausforderungen beim Einsatz maschinellen Lernens in den Sozialwissenschaften.

### 4.3 Praxisimplikationen

Das Modell eignet sich besonders für Standard- und Mittelklasseimmobilien. Regelmäßige Neukalibrierung alle 6-12 Monate wird empfohlen, um strukturellen Marktveränderungen Rechnung zu tragen. Die Implementierung von Data-Drift-Monitoring ist essentiell, wie von Gama et al. (2014) für sich verändernde Datenströme beschrieben. Für Marktteilnehmer ergeben sich wertvolle Erkenntnisse zur Bedeutung der Nachbarschaftsanalyse und saisonaler Preisschwankungen. Kommunalverwaltungen können das Modell für Grundsteuerbemessung und Stadtplanung nutzen, wie von IAAO (2013) empfohlen.



#### 4.4 Ausblick und zukünftige Forschung

Zukünftige Forschung sollte die Integration externer Datenquellen (Satellitenbilder, Street View, sozioökonomische Indikatoren), Experimente mit erweiterten Methoden (tiefe neuronale Netze, räumliche Ökonometrie) und verbesserten Umgang mit Datenqualität adressieren. Ethische Aspekte möglicher Verzerrungen in den Daten und Modellen erfordern kritische Reflexion, insbesondere im Hinblick auf sozioökonomische Segregation (O'Neil, 2016). Die Integration von Natural Language Processing für Immobilienbeschreibungen könnte zusätzliche Prädiktionskraft generieren, wie von Law et al. (2019) vorgeschlagen.

## 7. Literaturverzeichnis

Altmann, A., Tolosi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*, 26(10), 1340-1347.

Anselin, L. (1988). *Spatial econometrics: methods and models*. Kluwer Academic.

Antipov, E. A., & Pokryshevskaya, E. B. (2012). Mass appraisal of residential apartments. *Expert Systems with Applications*, 39(2), 1772-1778.

Baldominos, A., Blanco, I., Moreno, A. J., Iturrarte, R., Bernárdez, Ó., & Afonso, C. (2018). Identifying real estate opportunities using machine learning. *Applied Sciences*, 8(11), 2321.

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(2), 281-305.

Bourassa, S. C., Hoesli, M., & Sun, J. (2007). A simple alternative house price index method. *Journal of Housing Economics*, 16(1), 80-97.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, 26(2), 211-243.

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

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).

City of Milwaukee Open Data Portal. (2025). Property sales data (2002–2025). <https://data.milwaukee.gov/dataset/property-sales-data>

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

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.

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

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM computing surveys*, 46(4), 1-37.

Garcia, M., Veiga, H., & Weissensteiner, A. (2019). Machine learning in housing markets: A survey. *Journal of Economic Surveys*, 33(4), 1253-1280.

Glaeser, E. L., Gottlieb, J. D., & Gyourko, J. (2012). Can cheap credit explain the housing boom? *Housing and the Financial Crisis*, 301-359.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

Hilber, C. A., & Mayer, C. J. (2009). Why do households without children support local public schools? *Journal of Urban Economics*, 65(1), 74-90.

Hill, R. J., & Melser, D. (2008). Hedonic imputation and the price index problem: an application to housing. *Economic Inquiry*, 46(4), 593-609.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67.

IAAO. (2013). *Standard on Mass Appraisal of Real Property*. International Association of Assessing Officers.

Jud, G. D., & Winkler, D. T. (2002). The dynamics of metropolitan housing prices. *Journal of Real Estate Research*, 23(1), 29-45.

Kontrimas, V., & Verikas, A. (2021). The mass appraisal of the real estate by computational intelligence. *Applied Soft Computing*, 102, 107089.

Kurlat, P., & Stroebe, J. (2015). Testing for information asymmetries in real estate markets. *The Review of Financial Studies*, 28(8), 2429-2461.

Law, S., Paige, B., & Russell, C. (2019). Take a look around: Using street view and satellite images to estimate house prices. *ACM Transactions on Intelligent Systems and Technology*, 10(5), 1-19.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764-766.

Malpezzi, S. (2003). Hedonic pricing models: a selective and applied review. In *Housing economics and public policy* (pp. 67-89). Blackwell Science Ltd.

Manning, W. G. (1998). The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of Health Economics*, 17(3), 283-295.

Mayer, C. J. (2011). Housing bubbles: a survey. *Annual Review of Economics*, 3(1), 559-577.

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

Nguyen, N., & Cripps, A. (2001). Predicting housing value: A comparison of multiple regression analysis and artificial neural networks. *Journal of Real Estate Research*, 22(3), 313-336.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

Pace, R. K. (2019). *Machine learning for real estate*. Routledge.

Pace, R. K., & Barry, R. (1997). Quick computation of regressions with a spatially autoregressive dependent variable. *Geographical Analysis*, 29(3), 232-247.

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

Sirmans, G. S., Macpherson, D. A., & Zietz, E. N. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature*, 13(1), 3-43.

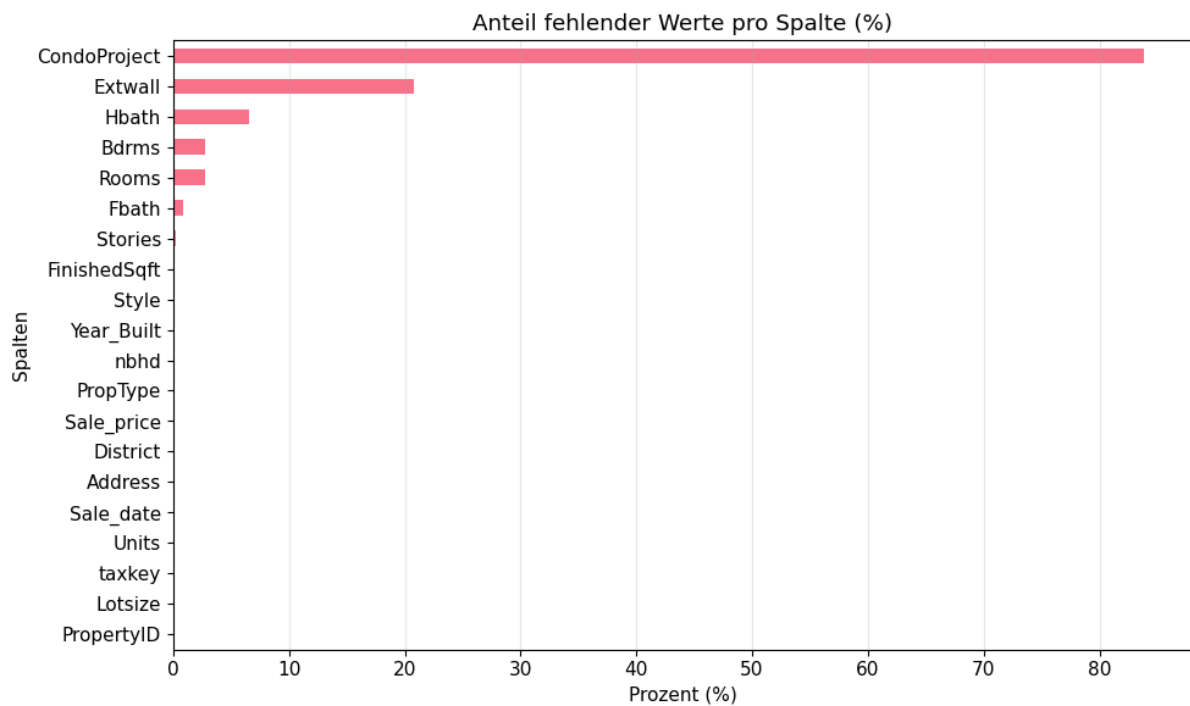
Zhou, L., & Tüske, Z. (2018). A comprehensive review of house price prediction using machine learning. *International Journal of Computer Applications*, 182(34), 1-9.

## Anhang

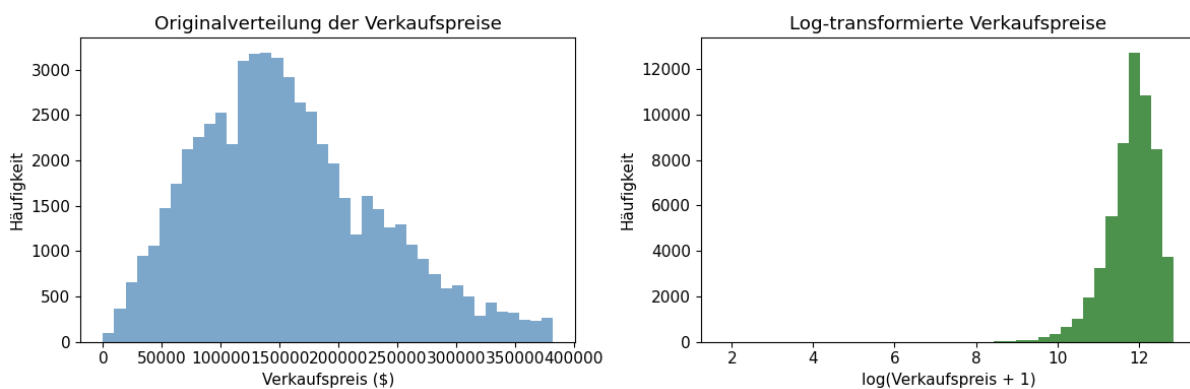
### Anhang A: Variablenbeschreibung des Datensatzes

Column	Description
PropertyID	
PropType	Property classification : Residential, Condominium, Commercial. Lg Apartment, Manufacturing
Taxkey	Tax parcel identifier of the property
Address	Street address of the property
CondoProject	Condo properties only. Name of the condominium project based on the condominium declaration that established the condo unit of the property
District	Aldermanic district of the property
Nbhd (Neighborhood)	"This is the Assessor's Office neighborhood designator. The City has been divided into approximately 200 different neighborhoods. Reference Neighborhoods map, <a href="https://data.milwaukee.gov/dataset/neighborhoods/resource/d1ab25e8-dfa4-4a6d-98f6-78f9a972ec7f">https://data.milwaukee.gov/dataset/neighborhoods/resource/d1ab25e8-dfa4-4a6d-98f6-78f9a972ec7f</a> "
Style	Building style of the primary building on the property
Extwall	"Exterior wall construction: Aluminum/Vinyl Siding, Asphalt, Block, Brick, Brick on Block, Brick on Frame, Concrete Block, Fiber Cement/Hardiplank, Masonry/Frame, Metal Siding, Other, Precast Masonry, Premium Wood, Stone, Stucco, Wood"
Stories	Number of stories of primary building on the property
Year_Built	Year built of primary building on the property
Rooms	Numbers of rooms of primary building on the property
FinishedSqft	Total finished square feet of all buildings on the property
Units	Number of units, residential and commercial on the property
Bdrms	Number of bedrooms on the property
Fbath	Number of full baths on the property
Hbath	Number of half baths on the property
Lotsize	Land area of the property in square feet
Sale_date	Date of the sale of the property
Sale_price	Sale price of the property.

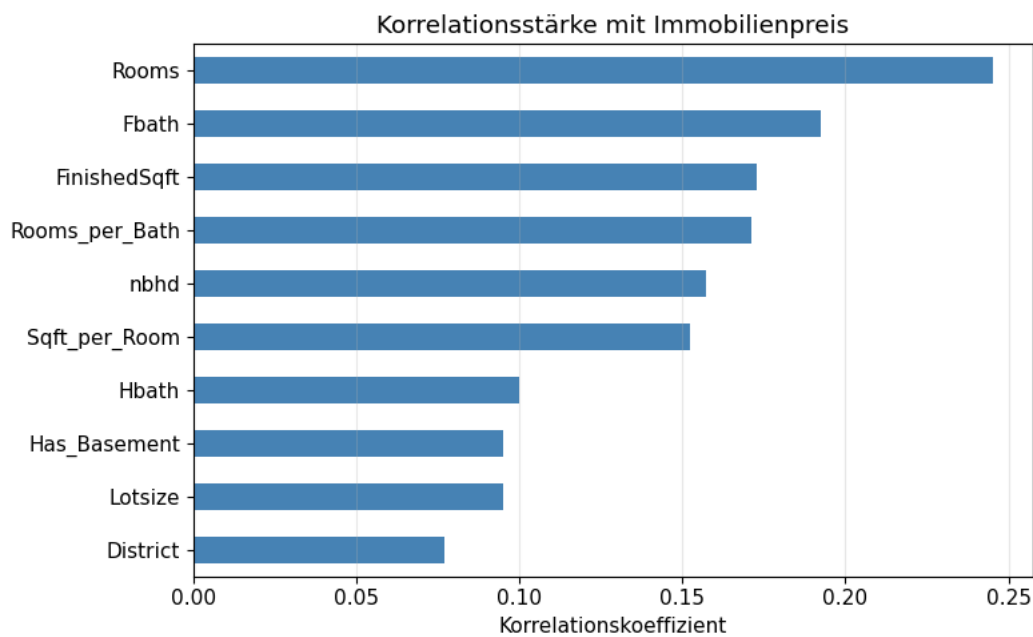
## Anhang B: Anteil fehlender Werte pro Spalte



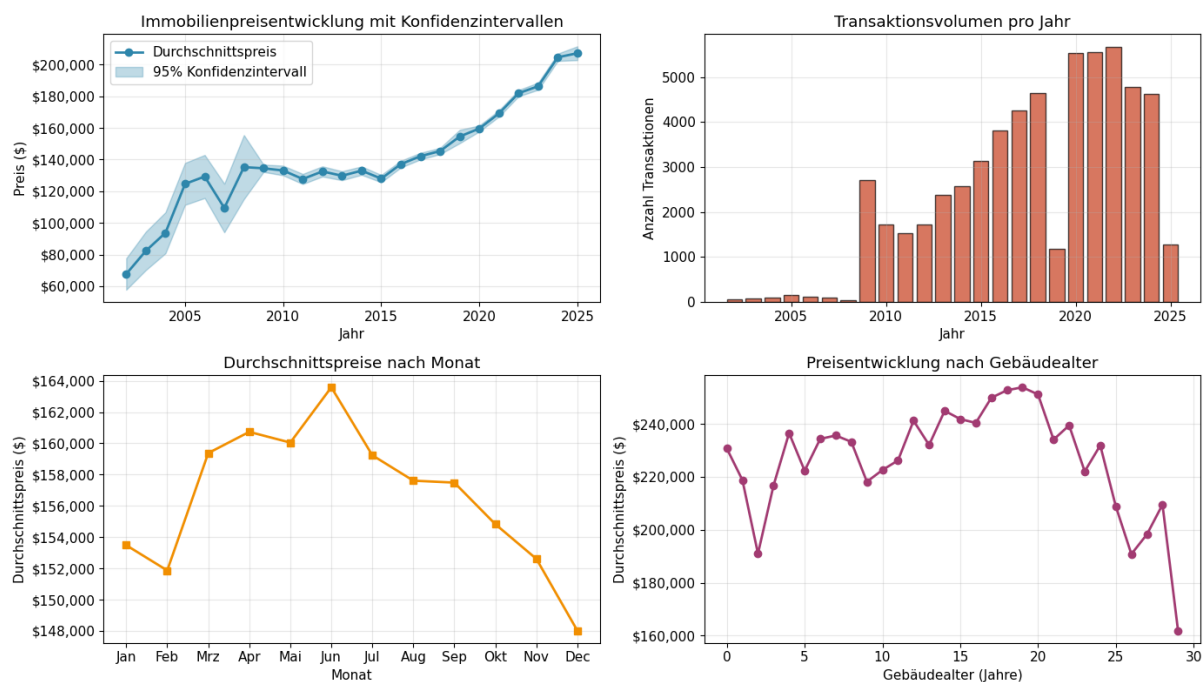
## Anhang C: Verteilung des Verkaufspreises (log-transformiert)



## Anhang D: Korrelation der numerischen Variablen mit dem Verkaufspreis

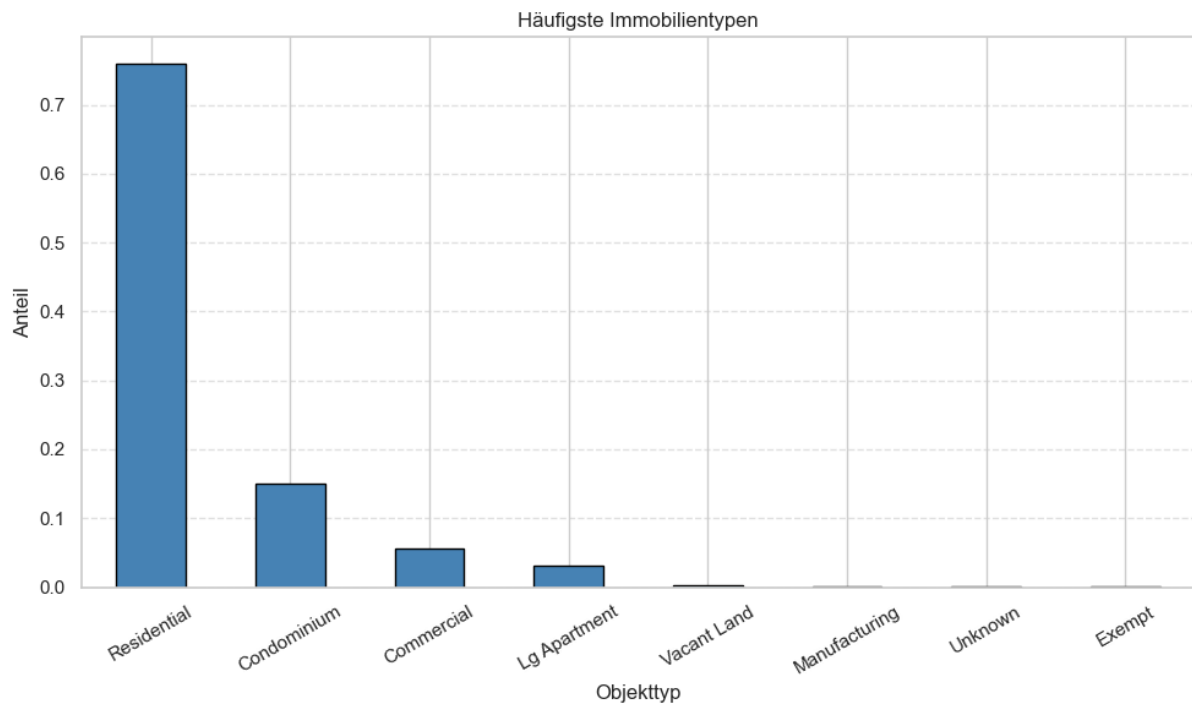


## Anhang E: Zeitlicher Trend der Immobilienpreise

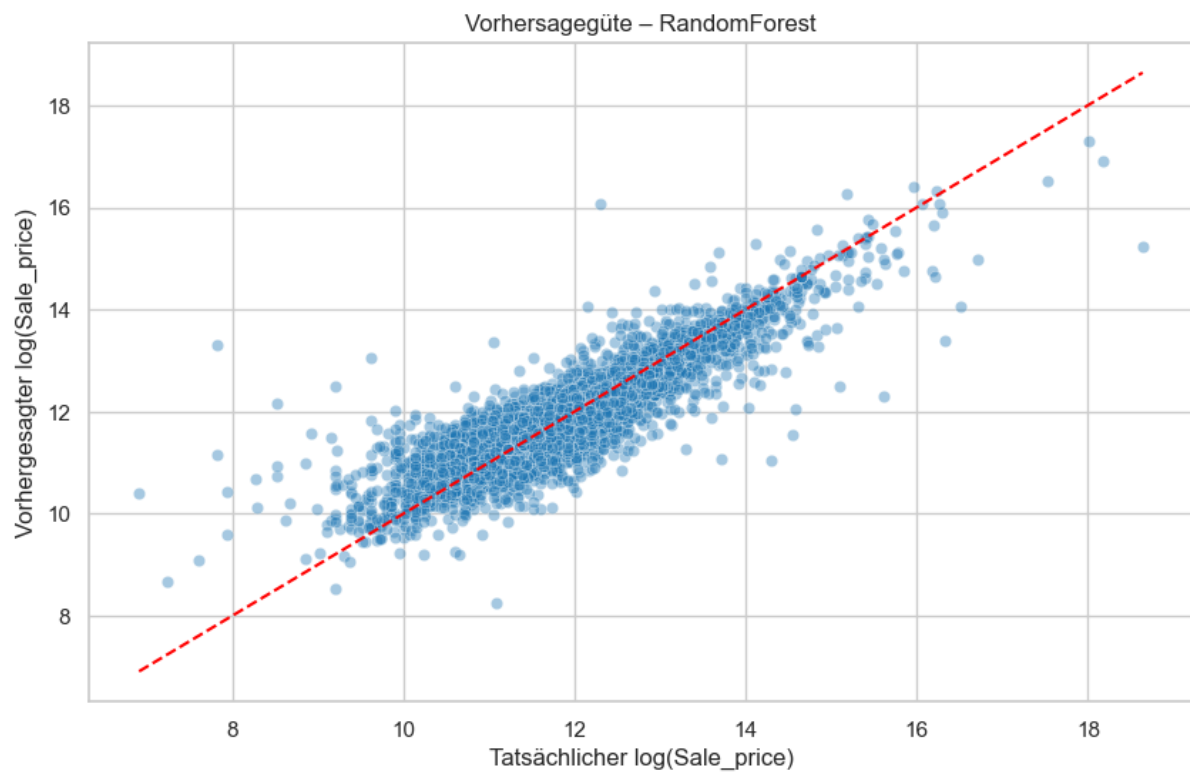


## Anhang F: Häufigkeitsverteilung der Immobilientypen

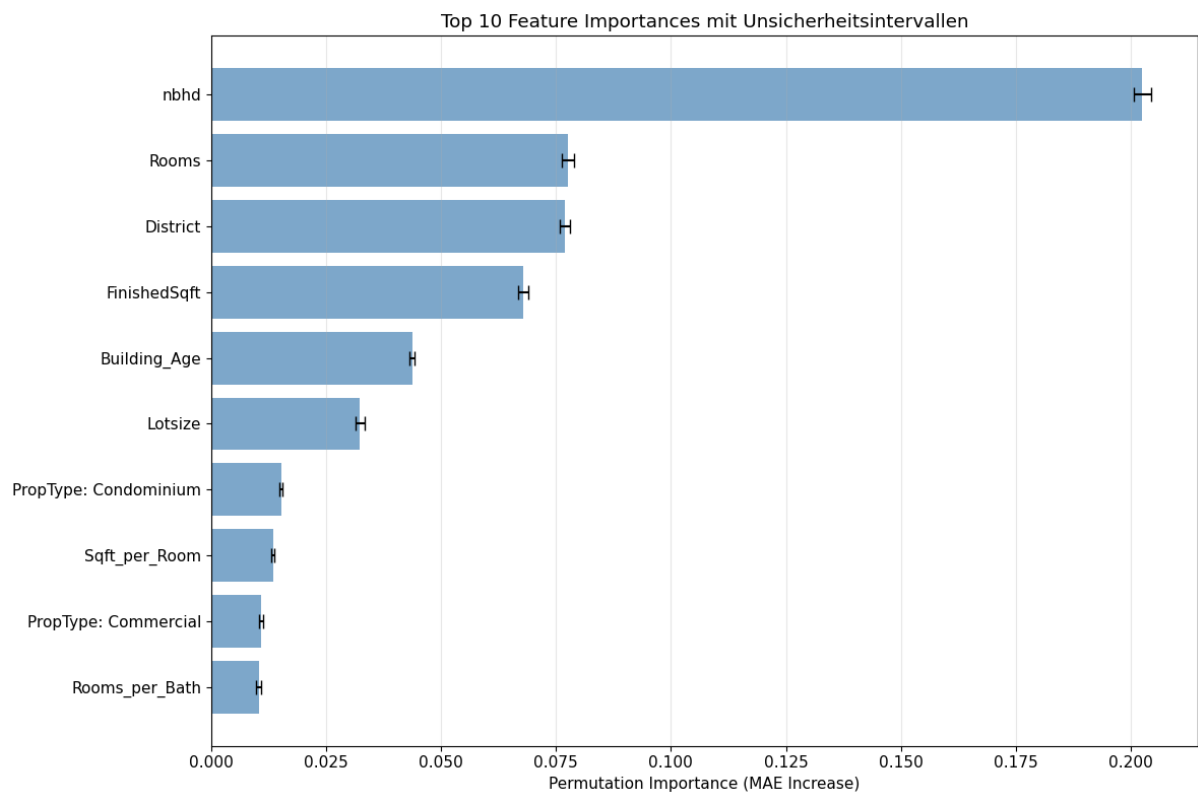




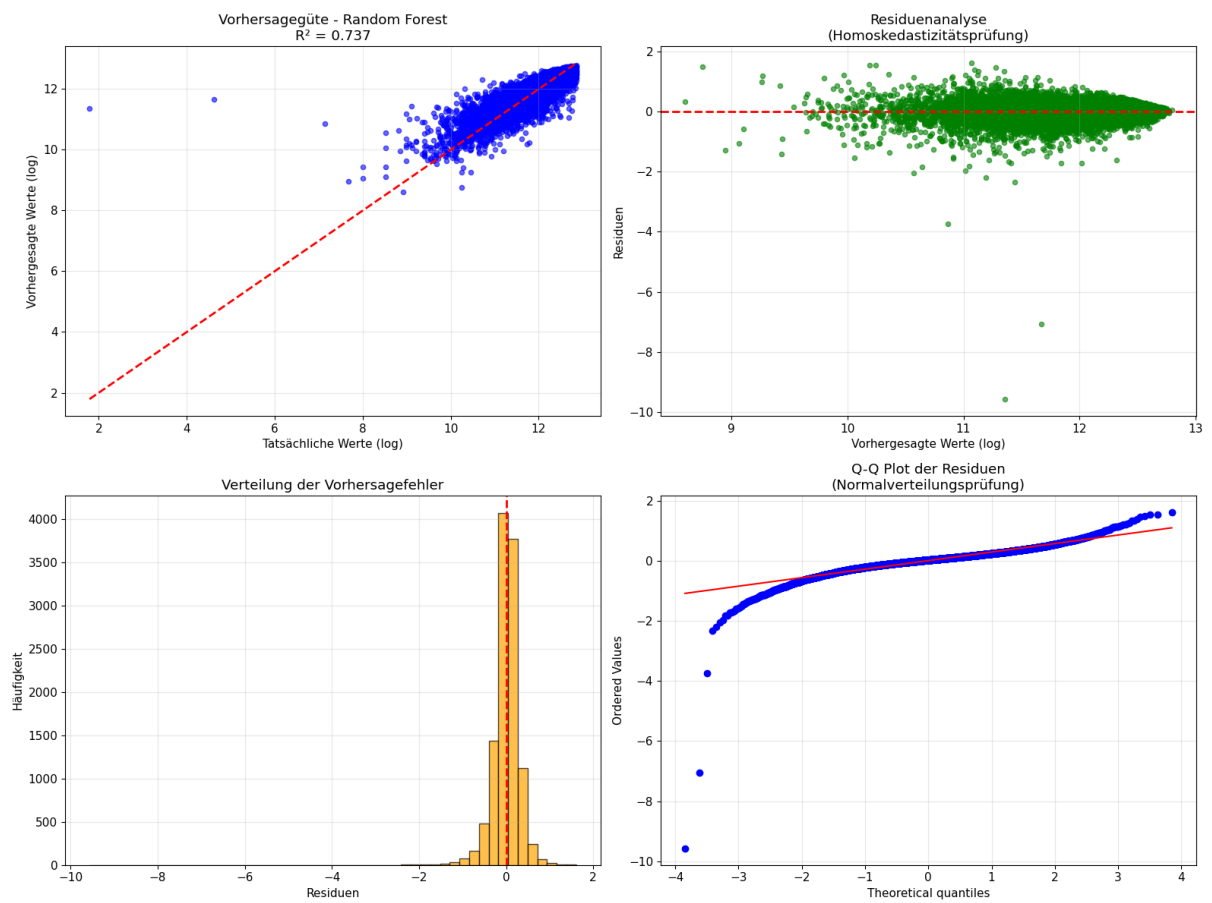
## Anhang G: Vorhersagegüte – RandomForest



## Anhang H: Top 10 Feature Importances - RandomForest



## Anhang I : Verteilungen



**Eigenständigkeitserklärung                      für                      schriftliche                      Prüfungsleistungen**  
**an der Digital Business University of Applied Sciences**

Name	Borto
Vorname	Talab
Matrikelnummer	200069
Modultitel	ADSC21 Applied Data Science 2: Machine Learning und Reporting
Thema der Prüfungsleistung	Einsatz von Künstlicher Intelligenz zur Vorhersage von Immobilienpreisen: Eine Fallstudie für Milwaukee
Prüfer:in	Sebastian Seck
Datum	26.10.2025

Ich trage die Verantwortung für die Qualität des Textes sowie die Auswahl aller Inhalte und habe sichergestellt, dass Informationen und Argumente mit geeigneten wissenschaftlichen Quellen belegt bzw. gestützt werden. Die aus fremden Quellen direkt oder indirekt übernommenen Texte, Gedankengänge, Konzepte, Grafiken usw. in meinen Ausführungen habe ich als solche eindeutig gekennzeichnet und mit vollständigen Verweisen auf die jeweilige Quelle versehen. Alle weiteren Inhalte dieser Arbeit (Textteile, Abbildungen, Tabellen etc.) ohne entsprechende Verweise stammen im urheberrechtlichen Sinn von mir.

☒ Hiermit erkläre ich, dass ich die vorliegende Prüfungsleistung selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle sinngemäß und wörtlich übernommenen Textstellen aus fremden Quellen wurden kenntlich gemacht.

☒ Die vorliegende Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

**Erklärung zu (gen)KI-Tools**

☒ **Verwendung von (gen)KI-Tools**

Ich versichere, dass ich mich (gen)KI-Tools lediglich als Hilfsmittel bedient habe und in der vorliegenden Arbeit mein gestalterischer Einfluss überwiegt. Ich verantworte die Übernahme jeglicher von mir verwendeter Textpassagen vollumfänglich selbst. In der „Übersicht verwendeter (gen)KI-Tools“ habe ich sämtliche eingesetzte (gen)KI-Tools, deren Einsatzform sowie die jeweils betroffenen Teile der Arbeit einzeln aufgeführt. Ich versichere, dass ich keine (gen)KI-Tools verwendet habe, deren Nutzung der Prüfer bzw. die Prüferin explizit schriftlich ausgeschlossen hat.

Hinweis: Sofern die zuständigen Prüfenden bis zum Zeitpunkt der Ausgabe der Aufgabenstellung konkrete (gen)KI-Tools ausdrücklich als nicht anzeige-/kennzeichnungspflichtig benannt haben, müssen diese nicht aufgeführt werden.

Ich erkläre weiterhin, dass ich mich aktiv über die Leistungsfähigkeit und Beschränkungen der unten genannten (gen)KI-Tools informiert habe und überprüft habe, dass die mithilfe der genannten (gen)KI-Tools generierten und von mir übernommenen Inhalte faktisch richtig sind.

☒ **Verbot bzw. Nicht-Nutzung von (gen)KI-Tools**

Ich versichere, dass ich die hier vorliegende Arbeit vollständig eigenständig formuliert habe und keine (gen)KI-Tools verwendet habe.

Mir ist bekannt, dass ein Verstoß gegen die genannten Punkte prüfungsrechtliche Konsequenzen haben und insbesondere dazu führen kann, dass die Prüfungsleistung mit „nicht ausreichend“ bzw. die Studienleistung mit „nicht bestanden“ bewertet wird und bei mehrfachem oder schwerwiegendem Täuschungsversuch eine Exmatrikulation erfolgen kann.

#### Übersicht verwendeter (gen)KI-Tools

Die (gen)KI-Tools habe ich, wie im Folgenden dargestellt, eingesetzt.

(gen)KI-Tool	Einsatzform	Betroffene Teile der Arbeit
Chatgpt	Um den gesamten Code auf Fehler zu prüfen.	Code

Ort

Köln

Datum

26.10.2025

Unterschrift Student:in



#### Sperrvermerk

Die vorliegende Hausarbeit mit dem Titel „Einsatz von Künstlicher Intelligenz zur Vorhersage von Immobilienpreisen: Eine Fallstudie für Milwaukee“ von Talab Borto enthält vertrauliche Informationen und Daten. Diese Arbeit wurde ausschließlich zum Zweck der Erfüllung akademischer Anforderungen erstellt. Sie darf Dritten, insbesondere Unternehmen, die in Konkurrenz stehen, nicht zugänglich gemacht werden.

Die Einsichtnahme, Veröffentlichung oder Weitergabe der Arbeit sowie die Nutzung der enthaltenen Informationen und Ergebnisse bedarf der ausdrücklichen schriftlichen Genehmigung des Verfassers und der betreuenden Hochschule.

Köln, 26.10.2025:

Unterschrift:

