ISSN: 1648-715X/eISSN: 1648-9179

VILNIUS TECH
Vilnius Gediminas Technical University

INTERNATIONAL JOURNAL OF STRATEGIC PROPERTY MANAGEMENT

2026
Volume 30
Issue 1
Pages 16–31
https://doi.org/10.3846/ijspm.2026.25981

A HORIZON ON THE EVOLUTION OF MACHINE LEARNING APPLICATIONS IN REAL ESTATE

Ruilin WANG[1], Hongqin ZHANG[2*], Fanrong LI[1]

[1] *College of Economics, Shanxi University of Finance and Economics, 030012 Shanxi, China*
[2] *School of Public Finance and Economics, Shanxi University of Finance and Economics, 030006 Shanxi, China*

Article History:
- received 23 January 2025
- accepted 21 October 2025

Abstract. Machine learning (ML) in the real estate industry has transformed property assessment, administration, and promotion, tackling significant issues including market instability and pricing precision. Notwithstanding considerable progress in predictive, descriptive, prescriptive analytics, and automation, current research mostly emphasises technological and operational efficiencies, overlooking the integration of environmental, social, economic, and governance (ESEG) sustainability dimensions. This monitoring constrains the advancement of comprehensive and accountable real estate solutions corresponding to sustainable development objectives. This study aims to address these gaps by systematically analyzing publication trends, key contributors, and thematic clusters, incorporating sustainability principles via a combination of bibliometric and content analysis approaches. The study uncovers publication trends, key research themes, and their alignment with ESEG criteria. The results highlight significant research clusters in predictive and descriptive analytics while revealing a notable deficiency in sustainability-focused studies. Implications of this study underscore the necessity for incorporating ESEG dimensions into ML-driven real estate practices, promoting resilient, equitable, and environmentally responsible industry advancements. This study provides actionable insights for stakeholders to enhance sustainable ML adoption, fostering long-term viability and societal well-being in the real estate sector.

Keywords: machine learning, real estate, sustainability, bibliometric analysis, content analysis, ESEG framework.

*Corresponding author. E-mail: *zhanghq@sxufe.edu.cn*

1. Introduction

The incorporation of machine learning (ML) into the real estate industry has transformed the methods by which stakeholders analyse, forecast, and make informed decisions. Real estate, a major global economic sector, includes residential, commercial, and financial areas, each marked by intricate and fluctuating market circumstances. The utilisation of machine learning techniques in real estate improves predictive precision, optimises investment strategies, and enhances operational efficiencies, thereby tackling significant challenges such as market volatility, pricing accuracy, and resource allocation (Jafary et al., 2024; Vivekananda & Shidlyali, 2024). The increasing volume of data produced in real estate transactions, along with developments in machine learning algorithms, highlights the necessity and immediacy of utilising these technologies to sustain competitiveness and foster innovation in the business.

The current advancements in ML applications in real estate encompass predictive analytics for price forecasting, descriptive analytics for market segmentation, and prescriptive analytics for investment optimisation (Chen et al., 2024; Ram et al., 2024). Neural Networks (NNs), Random Forests (RF), and Deep Learning (DL) models are widely employed to analyse large datasets, reveal hidden patterns, and generate actionable insights (Tin et al., 2024; Yang, 2024). Despite advancements, notable gaps persist, especially in integrating sustainability dimensions into ML frameworks. Although economic and operational efficiencies are adequately addressed, ESEG factors are frequently under-represented, hindering the development of comprehensive and responsible real estate solutions. This oversight limits the sector's ability to effectively contribute to sustainable development goals and address wider societal impacts.

Although machine learning is increasingly utilised in real estate, the literature is still fragmented, overly technical, and inconsistent regarding geographic regions, property types, and sustainability objectives. Ongoing challenges include data quality, model overfitting, regulatory resistance, interpretability difficulties, and insufficient methodological standardization (Khandaskar et al., 2023; Kuppan et al., 2024; Mukherjee et al., 2024). Additionally, numerous ML models depend significantly on historical data, which restricts their flexibility in dynamic or emerging markets. The incorporation of ESEG factors is seldom explored thoroughly, with most studies neglecting the ethical, regional, and infrastructural limitations impacting ML implementation. These constraints underscore the necessity for a systematic, critical evaluation that both consolidates existing insights and assesses the methodological soundness and sustainability focus of ML applications in the real estate sector.

This study seeks to address gaps by analysing ML applications in real estate using bibliometric and content analysis methods. This study assesses the alignment of ML applications with the dimensions of the sustainability framework ESEG by mapping identified research themes to these criteria. The methodology includes identifying relevant keywords, selecting databases, collecting and preprocessing data, and using analytical tools to reveal trends and themes. This dual approach highlights the current state of ML in real estate and emphasises the necessity of incorporating sustainability considerations to promote a more resilient and responsible industry.

The main contributions of this study are:

a) Providing an in-depth overview of publication trends, key authors, and influential institutions in the intersection of ML and real estate.
b) Identifying and categorizing major research themes and their maturity levels within the real estate sector.
c) Mapping research themes to the ESEG sustainability framework to assess the alignment of ML applications with broader sustainability goals.
d) Highlighting underexplored areas, particularly the integration of ESEG dimensions, to guide future research and industry practices.
e) Offering actionable recommendations for stakeholders to enhance the adoption of sustainable ML practices in real estate.

Section 2 reviews literature on ML applications in real estate, highlighting key themes and methodological challenges. Section 3 details the research methodology, including data collection, preprocessing, and techniques for bibliometric and content analysis. Section 4 presents bibliometric results, identifying influential authors, institutions, and publication trends. Section 5 offers an in-depth content analysis of machine learning techniques, their applications, and alignment with ESEG sustainability criteria. Section 6 addresses ethical considerations, methodological limitations, and research gaps in the current ML landscape. Finally, Section 7 concludes with a summary of findings, implications, and recommendations for future research and practice.

2. Related work

Several studies have explored the intersection of ML and real estate, each contributing unique insights into the application and impact of these technologies. Some studies examined how AI and ML are transforming property management, investment decisions, and development processes, emphasizing the shift towards a more data-driven industry (Conway, 2018). Another study conducted a bibliometric analysis of real estate bubble research, highlighting a focus on housing price formation and identifying a lack of advanced ML techniques in this area (Vergara-Perucich, 2023). Wang (2023) reviewed the digital transformation of real estate, noting the emerging field of smart real estate and the central role of data-driven topics like ML and big data. Another study surveyed AI-based ML methods for urban real estate prediction, revealing a preference for simpler ML models and a need for larger datasets and model explainability demonstrated the superior performance of ML algorithms in property price prediction compared to traditional models, while also addressing ethical concerns related to privacy and fairness (Tekouabou et al., 2024). Lastly, a study compared recent publications using ML algorithms in real estate, emphasizing the economic significance of ML but overlooking its role in sustainable development (Breuer & Steininger, 2020).

Despite these comprehensive reviews, there remains a significant gap in integrating bibliometric and content analysis with sustainability frameworks within ML applications in real estate. Existing literature predominantly focuses on enhancing economic efficiencies and predictive accuracies without thoroughly addressing how these technological advancements align with ESEG criteria. Additionally, most studies either perform bibliometric analyses or content analyses in isolation, limiting the depth of insights into ML research's structural and thematic dimensions in real estate. This study addresses these gaps using a combined bibliometric and content analysis approach, mapping identified ML research themes to the sustainability framework dimensions ESEG. Doing so provides a holistic assessment of how ML applications in real estate can contribute to sustainable development goals, thereby offering a more integrated and comprehensive understanding that previous studies have overlooked.

3. Methodology

This study utilises a systematic methodology that integrates quantitative and qualitative approaches across four phases: identifying search keywords, selecting databases, data collection and preprocessing, and performing bibliometric and content analyses. This combination comprehensively analyzes real estate ML applications, highlighting measurable trends and detailed thematic insights. Figure 1 visually summarises the methodology framework, detailing the sequential steps taken to meet the research objectives and enhance understanding of the topic (Altarturi et al., 2020).

The initial phase emphasises identifying search keywords that denote the convergence of ML and real estate. Keywords like "machine learning", "neural networks", and "property valuation" were chosen according to the research objectives and recognised knowledge gaps.

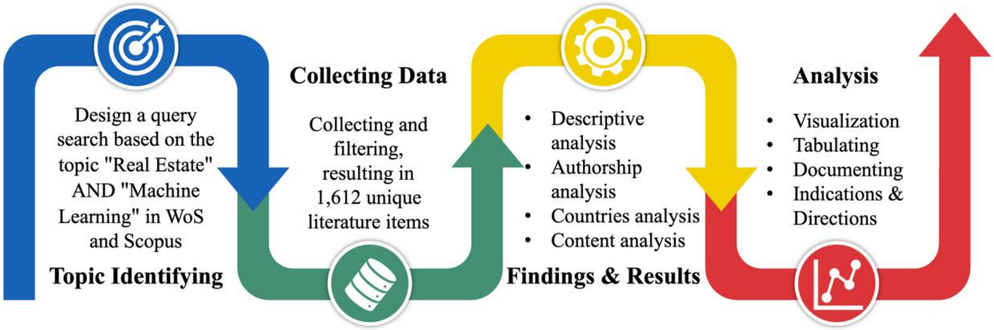

Figure 1. Methodology phases of the study

Boolean operators (AND, OR) were utilised to broaden the search's scope and accuracy by including secondary terms such as "prediction models", "random forests", and "deep learning". This strategic keyword formulation ensured the retrieval of studies that comprehensively represent ML methodologies in real estate while excluding unrelated results. This phase illustrates a qualitative approach to ensure data collection aligns with the study's thematic focus (Sahran et al., 2024).

The second and third phases entail database selection, data collection, and preprocessing, highlighting quantitative rigour. The Web of Science (WoS) and Scopus databases were chosen for their comprehensive coverage, quality citation data, and strong bibliometric capabilities (Altarturi et al., 2025). A systematic search retrieved articles published from 2011 to 2025 using the identified keywords. A dataset was screened to eliminate duplicates and exclude irrelevant studies according to established criteria. Articles specifically discussing ML applications in real estate were included, whereas those with peripheral or unrelated references were omitted; no filtering was performed on the basis of sustainability or ESEG attributes at this stage. The filtering process yielded a refined dataset of 1612 publications, ensuring data reliability and relevance for subsequent analyses.

The final phase combines quantitative bibliometric analysis with qualitative content analysis, highlighting the study's methodological distinctiveness. Bibliometric analysis utilised *R* (version 4.3.2) and the 'bibliometrix' package, revealing publication trends, citation counts, and collaboration networks. Content analysis employed advanced Natural Language Processing (NLP) techniques (Altarturi et al., 2023), such as trigram analysis and greedy clustering, to identify and categorise research themes. Themes aligned with the sustainability framework dimensions ESEG to evaluate the integration of the sustainability estate's ML research. This dual approach merges quantitative objectivity with qualitative depth, yielding a comprehensive understanding of the field and actionable recommendations for future research and practice. All tables in the findings section use the following standard abbreviations: NP (Number of Publications), TC (Total Citations), h-index (Hirsch index), PYS (Publication Year Started), SCP (Single Country Publication), and MCP (Multiple Country Publication).

4. Findings

4.1. Descriptive analysis

The Summary Information section provides key dataset statistics and real estate ML application trends. Table 1 shows key metrics from 1612 literature items published between 2011 and 2025 in 964 journals, conferences, and books. The dataset has 3832 authors, averaging 3.3 per document, indicating a collaborative but not fragmented research environment. The 12.03% international co-authorship rate suggests a gradual increase in global collaborations, which may increase field diversity and methodologies. A moderate academic impact is indicated by an average of 8.648 citations per document, and the extensive use of Keywords Plus (4697) and Author's Keywords (4126) indicates a broad and multifaceted research focus necessary for real estate ML applications.

Table 1. Summary information about the literature in the analyzed dataset

Description of the dataset	Value
Timespan	2011:2025
Sources (journals, books, conferences, etc.)	964
Documents	1612
Average citations per document	8.648
Keywords plus	4697
Author's keywords	4126
Authors	3832
Authors of single-authored documents	178
Single-authored documents	209
Co-authors per documents	3.3
International co-authorships %	12.03

Table 2 shows that 51.5% of the dataset is journal articles and 28.2% conference papers. This balance shows a strong foundation in peer-reviewed research and active conference discussions that facilitate timely knowledge exchange. The 13.0% proceedings paper rate emphasises the need to quickly share new findings. The low number of book chapters (3.7%) and reviews (1.9%) suggests the need for more comprehensive and integrative studies to advance theoretical frameworks and synthesise knowledge.

Table 2. Types of literature items in the analyzed dataset

Description of the dataset	Value	% of dataset
Article	830	51.5%
Conference paper	455	28.2%
Proceedings paper	209	13.0%
Book chapter	59	3.7%
Review	31	1.9%
Others (book, editorial, etc.)	28	1.7%

The low share of other formats (1.7%) suggests expanding publication avenues to create a more diverse and resilient real estate ML research landscape.

4.1.1. Annual growth of publications

The number of publications on ML applications in real estate increased from 47 in 2011 to 353 in 2024, indicating a significant upward trend. This trend commenced in 2017 and persists to the present (Figure 2). This rapid ascent illustrates the real estate sector's increased interest in and investment in ML, likely driven by technological advancements and the growing demand for data-driven insights. The significant rise from 2017 to 2024 suggests that machine learning is becoming increasingly vital in addressing complex real estate issues. The decline to fifteen articles in 2025 may indicate a temporary lull, reflecting market saturation or a shift towards niche research topics. Emphasising innovative applications and promoting interdisciplinary collaborations to address emerging real estate dynamics is essential for sustaining growth in the field.

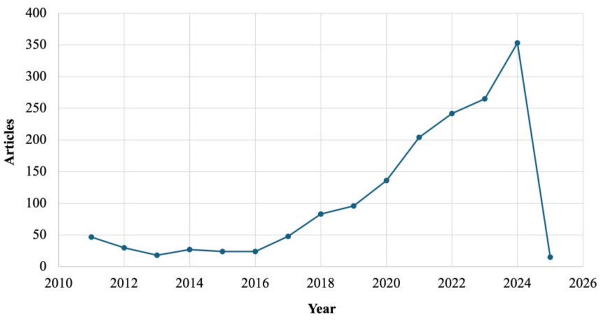

Figure 2. Publication annual growth in the field of ML for real estate

4.1.2. Top 15 impacting sources

The examination of the leading 15 sources in ML applications in real estate indicates a concentration of significant research within a limited number of journals, as illustrated in Table 3. Expert Systems with Applications has an h-index of 12 and significant total citations, highlighting its importance in publishing influential research that advances the field. Land Use Policy and Sustainability exhibit robust citation records, underscoring their essential roles in merging machine learning with sustainable real estate practices and land management strategies. Journals such as the Journal of Property Research and the International Journal of Housing Markets and Analysis exhibit considerable influence, reflecting a concentrated academic interest in the relationship between ML and property markets.

Despite the significance of these key sources, the data reveals a broader context in which most sources contribute little, with 800 of 964 sources publishing only once and 280 receiving no citations. This indicates a fragmented research landscape, with significant work concentrated in select journals, which may restrict diverse perspectives and impede the broad dissemination of innovative findings. The emergence of 130 new sources from 2023 onwards indicates growing diversification; however, the prevalence of single-contribution sources suggests challenges in achieving sustained engagement and recognition in the field. Enhancing the field's robustness requires increased collaboration and encouragement of publications in a broader range of reputable journals to promote balanced advancements in ML applications in real estate.

Table 3. The impact of the top journals in the field of ML in real estate

Source	NP	h-index	TC	PYS
Expert Systems with Applications	19	12	866	2011
Land Use Policy	15	10	465	2011
Sustainability	18	10	380	2017
Journal of Property Research	15	9	558	2013
International Journal of Housing Markets and Analysis	18	8	196	2011
ISPRS International Journal of Geo-Information	12	8	230	2017
Journal of Real Estate Finance and Economics	11	8	114	2020
Cities	12	7	187	2021
Land	15	7	155	2020
Lecture Notes in Computer Science	19	7	192	2011
International Journal of Strategic Property Management	13	6	194	2017
Electronics	5	5	61	2019
Journal of Property Investment & Finance	20	5	135	2018
Planning Malaysia	5	5	64	2019
Property Management	7	5	95	2016

4.1.3. Top impacting literature items

Citation metrics indicate the respect and influence of academic publications in real estate machine learning. Table 4 delineates the most-cited publications from 2015 to

Table 4. The top impactful articles in the field of ML in real estate

Item#	Title	Year	TC
PI1	Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data	2015	189
PI2	A novel machine learning model for estimation of sale prices of real estate units	2016	189
PI3	Novel machine-learning model for estimating construction costs considering economic variables and indexes	2018	173
PI4	Addressing disasters in smart cities through UAVs path planning and 5G communications: A systematic review	2021	145
PI5	Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics	2012	144
PI6	3D room layout estimation from a single RGB image	2020	143
PI7	Prediction accuracy in mass appraisal: A comparison of modern approaches	2013	135
PI8	Predicting property prices with machine learning algorithms	2021	129
PI9	The mass appraisal of the real estate by computational intelligence	2011	122
PI10	Grading buildings on energy performance using city benchmarking data	2019	114

2021 that exerted considerable influence. Influential studies such as PI1 and PI2 from 2015 and 2016, respectively, with 189 citations, have demonstrated how machine learning enhances property value precision. The elevated citation counts of these works indicate their methodological advancements and practical applicability, potentially impacting subsequent field studies and applications.

Of the 1,612 papers, 540 garnered no citations and 230 received a single citation, suggesting that a significant portion of the literature remains unacknowledged. This disparity may prevent many studies from reaching or influencing academic and professional groups, highlighting gaps in the distribution and acceptance of research findings. A select number of seminal works are emphasised among the 570 publications with citation counts exceeding one per year. The under-citation of these works indicates a necessity for promotion and increased engagement from individuals. We should promote collaboration, publish in prestigious journals, and employ interdisciplinary methodologies to enhance the impact of future research. This will bridge these gaps and enhance the integration of machine learning into real estate.

4.1.4. Top contributing affiliations

The distribution of publications in ML applications in real estate shows a notable concentration of research output among a limited number of institutions, as Figure 3 illustrated. The leading 15 affiliations, comprising only 1% of all contributing institutions, generate more than 18% of total publications. Universities like the University of Hong Kong, University of Regensburg, and Wroclaw University of Science and Technology significantly contribute articles, underscoring their essential roles in this research domain. This concentration highlights the prevalence of specific research hubs that likely influence innovation and establish research agendas in the field. The geographical diversity of these institutions across Asia, Europe, and Australia indicates a wide international interest and diverse regional strategies for integrating ML in real estate.

The data reveals a fragmented research landscape, with 745 of 1,390 affiliations contributing only one publication each. This disparity indicates restricted engagement and resource allocation among most institutions, which may limit diverse perspectives and impede the overall growth of the field. The emergence of 130 new affiliations from 2023 onwards indicates growing interest, although established

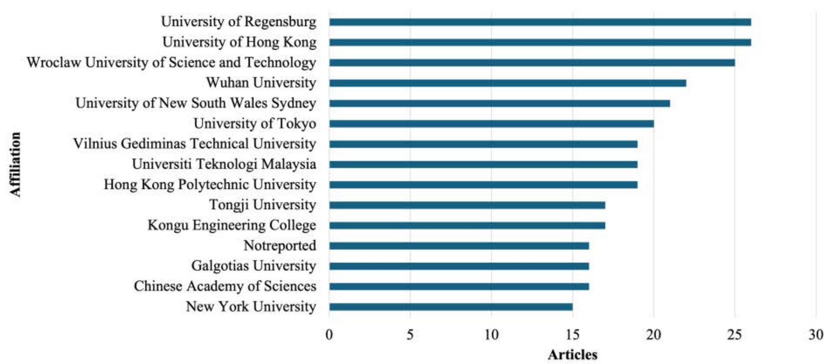

Figure 3. Top contributing affiliations in the field of ML in real estate

institutions continue to dominate. Efforts should be made to support and collaborate with a broader range of institutions to promote a more inclusive and diverse research environment. Improving cross-institutional partnerships and fostering contributions from under-represented regions may reduce current concentration, facilitating a more equitable and thorough development of ML applications in real estate.

4.2. Countries analysis

China and the USA dominate the landscape of ML applications in real estate, contributing a substantial share of publications and citations, as Table 5 shows. China's significant production and citation metrics demonstrate its robust investment and leadership in this research domain, whereas the USA retains a notable scholarly impact despite reduced international collaboration. Australia and Spain exhibit significant international collaboration, indicating that their research gains from varied perspectives and partnerships, potentially improving the quality and relevance of their studies.

In contrast, the majority of countries contribute little, as 70 of the 88 nations represent under 1% of contributions, while 35 lack any international collaborations. This fragmentation suggests insufficient engagement and resource allocation, which may hinder diversity and innovation in the field. The abundance of publications with minimal or no citations reveals challenges in research dissemination and recognition, emphasising the necessity for improved visibility and greater international collaboration to create a more inclusive and impactful research landscape.

This section examines global collaboration in ML applications within real estate, focussing on robust partnerships between China and the USA, as well as China and the UK. These alliances capitalise on each nation's technological capabilities and market knowledge, driving major progress and shaping research priorities. Collaborations among Australia, India, and Germany reflect a trend of integrating diverse regional perspectives, enhancing the robustness and applicability of research outcomes.

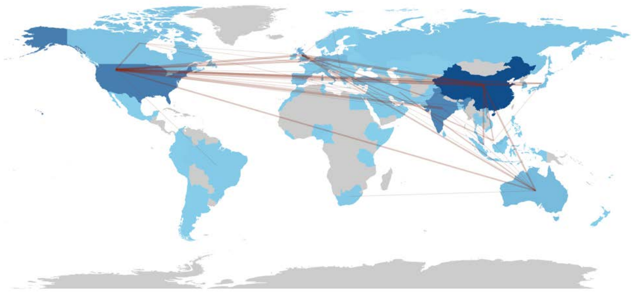

Figure 4. Countries' collaboration on the ML for real estate

Despite notable partnerships, international collaboration is generally limited, with most countries participating in few or rare joint studies, as Figure 4 shows. The concentration of ideas within a few nations may limit diversity and impede global inclusivity in the field. Broadening partnerships beyond dominant nations is crucial for integrating diverse cultural and methodological perspectives, enhancing the research environment, and more effectively tackling region-specific real estate issues.

4.3. Authorship analysis

The authorship landscape in ML applications within real estate is characterized by a significant concentration of contributions among a few key researchers, alongside widespread fragmentation. As Table 6 shows, Prominent authors such as Tadeusz Lasota, Zbigniew Telec, and Bogdan Trawiński from Polish institutions each have a substantial number of publications and demonstrate respectable h-

Table 5. Analysis of countries' impact and productivity in the field of ML in real estate

Country	No. of publication	Freq. appearance	SCP	MCP	MCP %	TC
China	310	587	276	34	11	2414
USA	154	315	126	28	18.2	2878
India	132	215	123	9	6.8	471
United Kingdom	52	112	41	11	21.2	531
Poland	49	91	46	3	6.1	244
South Korea	47	85	45	2	4.3	537
Germany	45	81	41	4	8.9	371
Italy	45	76	44	1	2.2	535
Australia	40	82	28	12	30	728
Spain	37	64	26	11	29.7	473
Malaysia	33	63	29	4	12.1	98
Turkey	25	51	25	0	0	286
Japan	24	53	20	4	16.7	147
Russia	22	32	20	2	9.1	199
Canada	18	42	17	1	5.6	140

Table 6. Top impactful authors in the field of ML for real estate

Author	NP	Fraction	h_index	TC	PYS
Tadeusz Lasota	16	3.49	7	141	2011
Zbigniew Telec	16	3.44	6	130	2011
Bogdan Trawiński	16	3.35	7	135	2011
Marcelo Cajias	10	5.17	3	42	2021
Rotimi Boluwatife Abidoye	9	3.65	9	213	2016
Toshihiko Yamasaki			4	36	2018
Laura Gabrielli	8	2.45	3	26	2020
Marek Walacik	7	3.17	4	37	2021
Aurora Greta Ruggeri	7	2.25	3	26	2020
Massimiliano Scarpa	7	2.25	3	26	2020

indices, indicating their pivotal roles in advancing the field. These authors lead in productivity and contribute to the foundational knowledge that underpins ongoing research efforts. However, despite their contributions, the overall citation impact remains modest, suggesting that even the most prolific authors have yet to achieve widespread recognition within the broader academic community.

In contrast, most authors in this field show minimal engagement, with more than 3,100 of 3,832 authors contributing only one publication. Furthermore, numerous authors obtain few citations, with 1,162 having none and 480 receiving only a single citation. This pattern indicates a fragmented research environment in which individual contributions are minimal, possibly reflecting a publish-or-perish culture that prioritises quantity over quality. Fragmentation can obstruct the formation of cohesive theoretical frameworks and disrupt ongoing scholarly dialogue. The emergence of 648 new authors from 2023 onwards indicates increasing interest; however, the prevalence of single-publication authors suggests limited long-term engagement in the field. To enhance the research community, it is crucial to promote ongoing collaboration and support systems that emphasise impactful, high-quality research rather than simply increasing publication volume.

4.4. Content analysis

4.4.1. Author keywords

Author keywords analysis using NLP methods for trends analysis approach summary identifies key patterns in ML for real estate research. The Timeline Bubble Plot visualization (Figure 5) shows relationships among the most common author keywords in the analyzed dataset and indicates the changing areas of emphasis in a research field. The rise of deep learning and AI terms in the early 2020s signifies a transition in the industry towards advanced technologies aimed at enhancing model accuracy and explainability. Advanced methods have not sufficiently

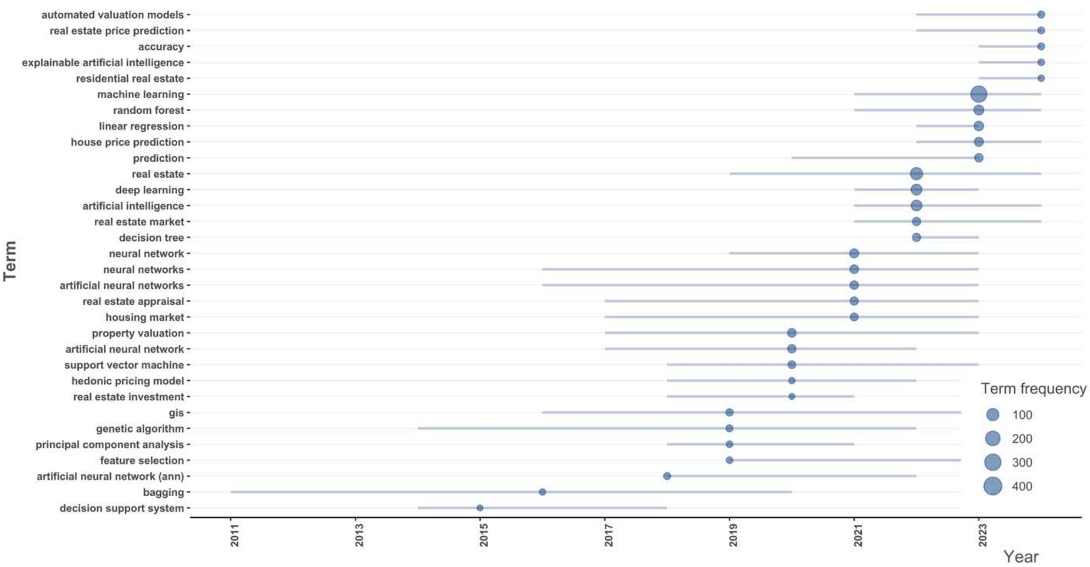

Figure 5. Topic trends in the field of ML for real estate

represented sustainability concepts. Sustainable development and automated valuation models are insufficiently emphasized in real estate practices, indicating a deficiency in environmental and economic sustainability. Future research must enhance sustainability aspects to guarantee that ML applications maximize economic results while fostering environmental and social welfare.

The sector has prioritized property valuation, real estate appraisal, and housing market analysis since 2017, demonstrating its commitment to prediction accuracy and investment strategies. Deep learning and artificial intelligence keywords increased in the early 2020s, indicating the industry's shift toward cutting-edge technologies for more accurate and explainable models. Despite using advanced methods, sustainability keywords are underrepresented. Sustainable development and automated valuation models are underemphasized in real estate practices, suggesting a gap in environmental and economic sustainability. Future research should strengthen sustainability dimensions to ensure that ML applications optimize economic outcomes and promote environmental and social well-being.

The recent dominance of ML, RF, and prediction models highlights the field's commitment to data-driven decision-making and forecasting in real estate. Nevertheless, the fragmentation across various topics such as NNs and Decision Trees (DT) indicates a dispersed research focus that may hinder the development of cohesive theoretical frameworks. Integrating sustainability with these advanced methodologies could bridge existing gaps, fostering a more comprehensive and responsible approach to real estate management. Encouraging interdisciplinary research that combines ML with sustainability principles will enhance the sector's ability to address multifaceted challenges and drive sustainable growth.

4.4.2. Keywords plus

The thematic analysis of Keywords Plus through NLP and co-occurrence clustering identifies specific research trends in ML applications in the real estate sector. As Figure 6 illustrates, three main clusters represent different field aspects in the network visualization. The Red Cluster emphasizes regression analysis and neural networks to develop and improve real estate predictive models. The Blue Cluster emphasises advanced technologies such as artificial intelligence, big data, and decision support systems, highlighting the integration of tools to improve decision-making and data processing in real estate. The Green Cluster includes economic and market-oriented concepts like real estate appraisal, housing market, and real estate investment, highlighting the sector's emphasis on market analysis and investment strategies.

The network's centrality measures show how keywords link research themes. In the Red Cluster, "machine learning" demonstrates significant betweenness and PageRank, serving as a crucial link between technical methodologies and practical real estate applications. In the Blue and Green Clusters, "artificial intelligence" and "regression analysis" are central, highlighting their role in connecting research domains and promoting interdisciplinary methods. The analysis indicates a significant lack of sustainability-related terms, including "sustainable development" and "environmental impact", in the major clusters. This gap shows that technological advances are adequately represented but sustainability is not. Addressing this shortfall is crucial for optimising economic outcomes and promoting environmental and social sustainability in ML applications within real estate. To promote responsible real estate management, future research should include sustainability themes.

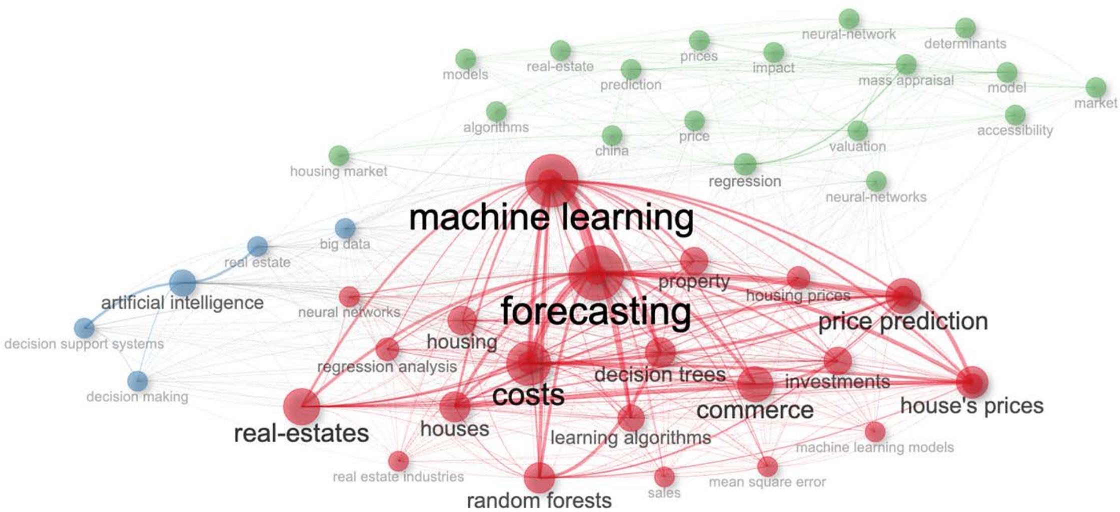

Figure 6. Network collaboration between Keywords Plus in the field of ML for real estate

4.4.3. Abstract and title analysis

Advanced NLP techniques for abstracts and titles reveal real estate ML applications' thematic structure and development trends. Thematic mapping (Figure 7) uses trigram analysis and greedy clustering techniques to rank research themes by centrality and density, indicating maturity and interconnections. Thematic mapping highlights key study fields such NNs models, ML algorithms, and hedonic price models. The primary and dense motor themes highlight their importance in real estate prediction and decision-making.

We conducted an expert-coding exercise to validate that the trigram-based greedy clustering captured coherent themes. Two senior scholars in real-estate analytics independently classified a random 10% subsample of abstracts and titles into the algorithm-defined clusters. Inter-rater reliability between the two experts was high, with a Cohen's $K = 0.89$, indicating strong agreement beyond chance. After resolving a few disagreements, mainly on boundary cases between some clusters, they produced a consensus coding. This consensus matched the algorithm's classification in 94% of the cases, confirming that the automated clustering aligned closely with expert judgment, with only minor manual adjustments needed to enhance interpretability.

The investigation shows inadequacies, especially in sustainability. There are a few sustainability phrases like "sustainable development", "environmental impact" or "green real estate". Technical approaches and market-oriented topics dominate. This shows that while ML is commonly employed to improve economic outcomes and market analysis, environmental and social sustainability are still neglected. Without sustainability-focused research, the sector may struggle to address resource efficiency, climate resilience, and social equality. Future research should incorporate sustainability to promote responsible research and improve economic optimization, environmental stewardship, and social well-being with ML applications.

Fragmented clusters like "floor plan images" and "principal component analysis" reflect a fragmented research emphasis that may hinder theoretical framework formation. Clusters like "deep neural networks" are advanced and interconnected, but the absence of interdisciplinary integration with sustainability themes requires more detailed study methods. Integrating sustainability into technical clusters helps overcome gaps and build a cohesive research environment. Integrating sustainability with modern ML methods might help future real estate research solve complicated problems, boosting technological innovation and sustainable growth.

5. Machine learning in real estate

ML has emerged as a transformative force in the real estate sector, offering tools that enable data-driven decision-making, operational efficiency, and innovative applications across various domains (Mandati et al., 2024; Vivekananda & Shidlyali, 2024). This section explores the practical implementation of ML in real estate, emphasizing its predictive, descriptive, prescriptive analytics capabilities, and the automation of operational processes.

To enhance understanding of ML applications, Figure 8 presents a structured taxonomy that outlines different ML approaches (supervised, unsupervised, reinforcement learning, and deep learning), commonly used ML algorithms, and their real-world applications in residential, commercial, and financial real estate. This visualization provides a clear overview, which aids in the following in-depth discussions within each analytics domain, helping stakeholders recognise how particular techniques tackle practical challenges and seize opportunities in real estate.

After this overview, subsections 5.1 to 5.1.4 delve into the key analytics methods- predictive, descriptive, and prescriptive- and explore how automation contributes to improving operational efficiency. Each section emphasises relevant ML techniques, real-world applications, implementations specific to various domains, and offers critical reflections on methodological limitations and their alignment with ESEG sustainability standards.

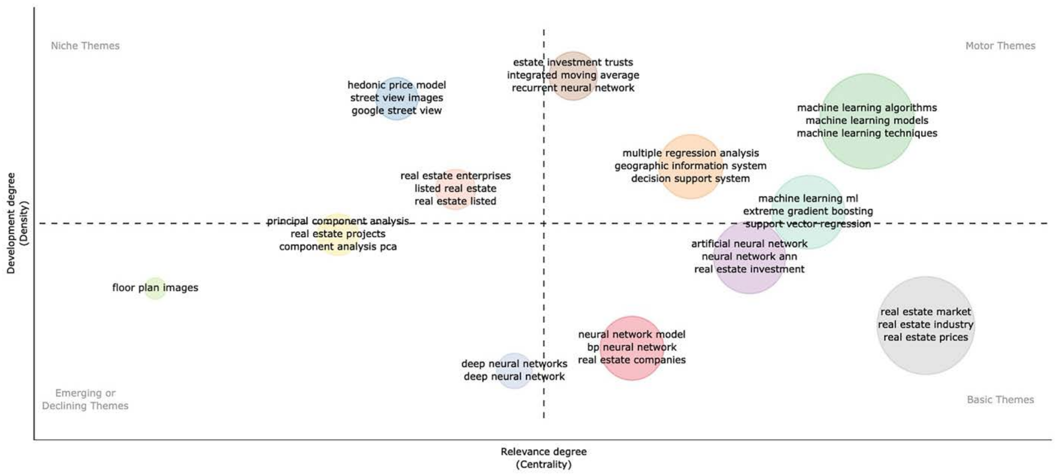

Figure 7. Thematic map of the ML for real estate research content

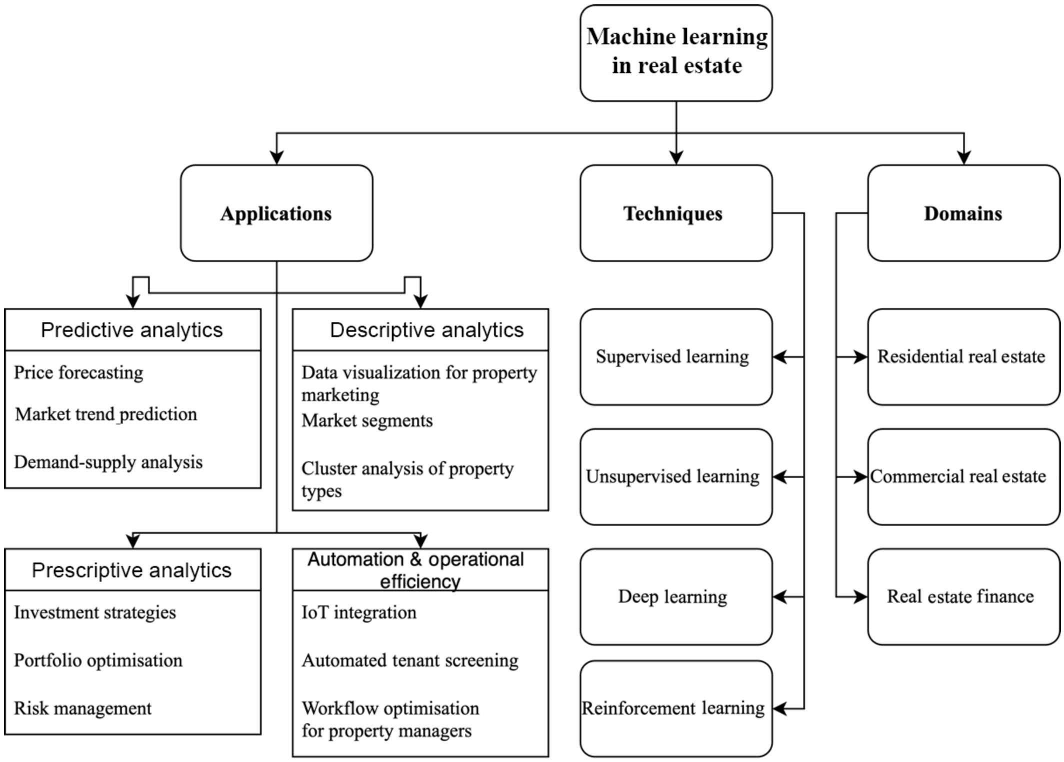

Figure 8. Taxonomy of the ML approaches and algorithms for the real estate field and its domains

5.1. Machine learning applications in real estate

This section explores the practical implementation of ML in real estate, emphasising its predictive, descriptive, and prescriptive analytics capabilities and automation for operational efficiency. In this integrated discussion, we will not only delve into how these analytics techniques revolutionise real estate practices but also examine their connections to specific ML approaches (such as *supervised*, *unsupervised*, *and reinforcement learning* (*RL*)) and their applications in diverse real estate domains, including residential, commercial, and financial sectors (Lubis et al., 2024; Panda et al., 2024). By understanding these interdependencies, stakeholders can better appreciate how ML bridges the gap between theoretical advancements and practical, real-world implementations, ensuring sustainable innovation and growth in the industry (Ganguly & Mukherjee, 2024; Rodriguez-Serrano, 2025).

The following subsections provide a detailed breakdown of these analytics approaches, showcasing their significance, associated ML techniques, and domain-specific applications.

5.1.1. Predictive analytics

Real estate predictive analytics combine real-time and historical data and advanced statistical and ML models to forecast future events. This strategy provides practical insights in commercial, residential, and financial sectors to help investors, developers, and politicians make better decisions. Predictive analytics analyzes demographic trends, economic data, and property features to assess demand-supply dynamics, market trends, and property value changes (Vivekananda & Shidlyali, 2024). These projections are enhanced utilizing regression, time series analysis, and advanced algorithms to aid strategic planning in response to changing market conditions.

Predictive analytics relies on supervised learning. RF, linear regression, and gradient boosting machines (GBMs) are widely used because they can handle structured data and make accurate predictions. LR and its advanced forms, like Ridge and Lasso regression, are extensively applied in predicting property prices by analysing variables like location, size, and neighbourhood amenities (Ganguly & Mukherjee, 2024). Ensemble methods, such as RF and GBMs, further enhance predictive power by capturing nonlinear relationships and interactions between variables (Purushotham et al., 2024). Moreover, Deep learning techniques, particularly NNs, improve the accuracy and scalability of predictive analytics when processing large, complex datasets. For example, Recurrent NNs (RNNs) are ideal for forecasting market trends by analysing temporal datasets like transaction histories and macroeconomic indicators (Tin et al., 2024). Convolutional NNs (CNNs) extract features from images, aiding market analysis and property assessment. RL has niche applications in predictive analytics, though it's less common than supervised learning. RL models optimize pricing strategies based on changing demand, competition, and market factors. RL algorithms have achieved a 5–10% revenue boost in dynamic pricing applications (Gadipudi & Kalpana Kalaimani, 2024).

Residential real estate uses predictive analytics for price prediction and rental rate optimization. Supervised learning models like LR, RF, and GBMs forecast property size, location, and amenities to help buyers and investors make data-driven decisions (Sobana et al., 2024). Commercial real estate predictive analytics aids strategic leasing, property valuation, and rental forecasting. GNNs predict foot traffic by analysing consumer behaviour, environmental factors, and location-based data, helping retailers optimise staffing, promotions, and inventory (Boz et al., 2023).

Although the integration of ESEG sustainability criteria remains underdeveloped, predictive analytics in real estate predominantly emphasizes economic factors, such as price forecasting and market predictions. Machine learning techniques like Random Forest and Decision Trees have been effectively applied to predict whole-life carbon emissions (WLCE) and optimize energy efficiency during the early stages of building design to align predictive models more closely with ESEG objectives (Elias & Issa, 2024; Zheng et al., 2024). These integrations can facilitate more sustainable investment decisions, like incorporating social equity and environmental responsibility into predictive frameworks. Expanding predictive analytics to include ESEG metrics systematically can enhance its strategic value, supporting resilient and sustainable real estate development (Walacik & Chmielewska, 2024).

5.1.2. Descriptive analytics

Descriptive analytics utilizes historical data to assist real estate professionals in identifying trends, patterns, and anomalies. This strategy enhances understanding in residential, commercial, and financial domains by consolidating extensive records. The effective conveyance of insights and facilitation of data-driven decision-making critically relies on tools such as dashboards, scatter plots, and visualisers. Descriptive analytics has demonstrated significant value in volatile markets by elucidating macroeconomic consequences and pricing dynamics (Bhattacharyya et al., 2018). Data visualisation transforms complex data sets into comprehensible formats, such as interactive dashboards, heat maps, and graphs. These instruments expose patterns in rental rates, transaction volumes, and price variations, thereby helping stakeholders to spot areas of development and track changes in the market. Many times, advanced geospatial analytics combines demographic data and sales records to produce logical visual outputs that enable quicker and more informed judgments.

Descriptive analytics frequently relies on unsupervised learning techniques to uncover hidden patterns and groupings within datasets:

a) Clustering: algorithms like k-means, DBSCAN, and hierarchical clustering identify groupings based on shared attributes such as property type, location, or price range (Chaudhry et al., 2023). For instance, clustering can differentiate high-end urban properties from suburban family-oriented developments, aiding pricing strategies and investment planning.
b) Dimensionality reduction: techniques such as Principal Component Analysis, t-SNE, and UMAP simplify high-dimensional datasets, improving interpretability without significant information loss. These methods support spatial data modelling and highlight factors like location and amenities for property pricing (Cheng et al., 2024).

Supervised learning is often linked with predictive analytics, but it can also be used to support descriptive analytics. Market segmentation is improved with the help of classification models like DT and random forests, which group properties into predefined categories using historical data. According to Skovajsa (2023), this method enhances strategic planning and refines segmentation.

Descriptive analytics improves property segmentation and enhances understanding of customers in the residential sector. Clustering techniques enable professionals to categorize properties into affordable rentals, high-end residences, or niche markets. Visualization tools such as heat maps illustrate price fluctuations and demographic changes, informing investment decisions and marketing strategies (Nur & Siregar, 2024). Descriptive analytics facilitates decision-making for commercial properties by summarising foot traffic data, tenant demographics, and rental trends. Geospatial visualisation methods combine transaction histories with consumer data to enhance retail space utilisation and improve lease agreements. Clustering algorithms assist in delineating submarkets and identifying high-demand locations (Boz et al., 2023; Ganguly & Mukherjee, 2024). Descriptive analytics in finance focuses on identifying trends in mortgage activity, loan default rates, and credit risk profiles. Historical financial data is summarised to reveal patterns through models like decision trees and clustering, which aid in portfolio optimisation and risk assessment. Heat maps and dashboards provide a visual representation of insights, facilitating improved management of financial risks by lenders (Akinjole et al., 2024; Hung & Lo, 2024).

Descriptive analytics plays a crucial role in comprehending historical trends in real estate markets, but its capability to visualise ESEG-related performance metrics is still not fully realised. Sophisticated visualisation tools like geospatial analy tics and BIM can effectively depict the environmental impact of buildings, energy efficiency, and demographic changes (Asif et al., 2024; Sridharan & Maddern, 2022). Incorporating ESEG performance assessment tools enhanced by AI models greatly improves descriptive analytics, allowing stakeholders to visualize governance anomalies, environmental risks, and social impacts easily. Incorporating these ESEG dimensions allows descriptive analytics to promote responsible market interventions, steering strategic decisions towards sustainable development goals (Filippi & Carbone, 2025).

5.1.3. Prescriptive analytics

Prescriptive analytics uses advanced ML models and optimisation to give real estate professionals actionable recommendations to evaluate multiple scenarios and make optimal decisions. Instead of forecasting future trends, prescriptive analytics uses predictive models and scenario-based evaluations to recommend the best course of action. It balances returns and uncertainty in investment strategies, portfolio optimisation, and risk management (Alkhaldi, 2023; Huang, 2024). Prescriptive analytics analyzes returns, market trends, and risk factors to help real estate investors strategize. NNs and RF capture complex dataset relationships better than statistical methods. Real estate investment trusts use these models to improve property acquisition and diversification. Iterative scenario simulations match investment decisions to risk tolerance and market conditions, ensuring robust returns in dynamic environments (Yang, 2024).

RL plays a crucial role in prescriptive analytics, particularly in portfolio optimization and dynamic pricing. RL agents master sequential decision-making to boost cumulative rewards through environmental interaction. RL models adjust real estate investment portfolios in real-time, factoring in time-sensitive data, risk levels, and market trends, enhancing performance in fluctuating markets (Ram et al., 2024). Multi-agent RL frameworks boost flexibility in intricate situations with various participants. Other supervised learning methods like RF, DT, and NNs enhance scenario evaluation by modeling the links between input variables (e.g., economic indicators) and desired outcomes (e.g., investment returns). Ensemble methods enhance predictions and provide strong guidance for property acquisition and diversification strategies (Huang, 2024; Piao, 2023). These clusters improve recommendation relevance by targeting specific market segments (Skovajsa, 2023). Ensemble methods such as XGBoost and stacking are frequently utilized to investigate diverse scenarios, improve predictions, and model the impacts of economic pressures. These methodologies evaluate market volatility, regulatory changes, and environmental hazards, enabling stakeholders to adopt proactive strategies (Leng, 2024; Rao & Ravi, 2018).

Current prescriptive analytics methods successfully enhance investment choices and risk management, but often fail to address essential ESEG sustainability factors. AI-powered scenario planning and reinforcement learning techniques incorporate environmental and social dimensions of sustainability, including energy efficiency and climate resilience, into prescriptive models (Darko et al., 2023; Rane et al., 2024). Furthermore, prescriptive frameworks can utilise AI-powered ESEG models, including NLP and CNNs, to conduct thorough scenario analyses that focus on governance compliance and transparency (Gupta & Agarwal, 2024). Emphasizing these sustainability factors within prescriptive analytics enables more comprehensive scenario evaluations, promoting robust and ethically responsible real estate investments.

5.1.4. Automation and operational efficiency

Real estate automation using ML, IoT, and workflow optimisation improves efficiency and service quality. Automation streamlines residential and commercial processes, lowering costs, improving decision-making, and increasing tenant satisfaction. Scalable, data-driven property management is enabled by predictive maintenance and automated tenant screening (Bechina & Arntzen, 2022). IoT sensors and ML-driven analytics optimise energy usage, security, and tenant comfort in intelligent property management systems. IoT integration into building management systems has cut HVAC energy consumption by 30% and operational costs by 20% (Ram et al., 2024). Predictive maintenance, occupancy sensors, and automated adjustments boost efficiency, but increased connectivity requires strong cybersecurity (Zaman et al., 2024).

Supervised learning plays a critical role in automation by training models to predict outcomes and automate decisions. For example:

a) Logistic regression and RF: these algorithms score tenant applications based on predefined risk parameters, automating background checks, credit evaluations, and rental history verifications. This reduces manual work and improves turnaround times (Chen et al., 2024).
b) Classification models: used for tenant profiling and risk assessment, these models categorise tenants into risk levels, ensuring compliance and improving community safety (Yadav et al., 2024).

Unsupervised techniques enhance workflow optimisation by uncovering hidden patterns in operational data. For instance:

a) Clustering algorithms: group properties based on energy consumption or maintenance needs, enabling targeted interventions that improve efficiency.
b) Dimensionality reduction: simplifies large IoT datasets while preserving critical information for actionable insights (Cheng et al., 2024).

RL is used for dynamic decision-making in real-time property management, such as optimising energy usage and predictive maintenance schedules. By learning from interactions with building systems, RL algorithms adjust HVAC operations and lighting systems to balance tenant comfort and energy efficiency (Ghasemi & Ebrahimi, 2024). Moreover, Deep learning supports automation through advanced image and video recognition. For instance, CNNs process surveillance footage to detect security breaches, while RNNs analyse occupancy patterns to predict maintenance needs (Mandati et al., 2024).

Automation powered by machine learning, Internet of Things (IoT), and workflow optimization focuses on boosting operational and economic efficiency. However, it is crucial to integrate more ESEG dimensions. Intelligent property management systems leveraging IoT and AI analytics can substantially lower energy usage and enhance resource optimization, closely aligning with environmental and governance standards (Asif et al., 2024). For example, smart technologies like IoT-enabled Building Management Systems have shown considerable energy savings and increased operational transparency, playing a major role in achieving sustainable development goals (Stroparo et al., 2023). Enhancing the integration of social sustainability elements, like automated impartial tenant screening and monitoring community safety, can strengthen ESEG objectives by fostering social equity and inclusiveness. This holistic method ensures that automation in real estate actively promotes sustainability and ethical management practices.

5.2. Challenges and ethical considerations

Real estate evaluation, management, and marketing have changed due to ML. As with any technological leap, these advancements present challenges and ethical issues that must be considered, as discussed below:

a) *Data quality issues*: incomplete or outdated datasets can reduce the accuracy of predictive models in property valuation and market forecasting (Vivekananda & Shidlyali, 2024).
b) *Algorithmic bias*: ML models trained on biased historical data can perpetuate inequalities, especially in tenant screening and market segmentation, necessitating careful data curation and regular model audits to ensure fairness (Jafary et al., 2024).
c) *Privacy and security concerns*: the use of IoT devices in smart property management systems raises significant privacy and security issues, requiring robust cybersecurity measures and transparent data usage policies to protect sensitive tenant information (Zaman et al., 2024).
d) *Opacity of Black-Box models*: advanced machine learning techniques, such as deep learning, often lack transparency, complicating stakeholder understanding and trust in outcomes. This hinders adoption and highlights the need for explainable models (Tin et al., 2024).
e) *Balancing automation with human interaction*: ML-driven automation improves efficiency but raises ethical issues regarding workforce displacement and diminished personal interaction, underscoring the necessity of integrating automation with human roles (Cholaraja et al., 2024).
f) *Accessibility divide*: smaller real estate enterprises face barriers to adopting ML innovations due to high costs and technical requirements, calling for democratized access to ML tools through open-source platforms or affordable solutions (Iyelolu et al., 2024).
g) *Environmental impact*: the computational demands of ML, particularly deep learning models, increase energy consumption and carbon footprint, emphasizing the need for energy-efficient algorithms and renewable energy use in data centers to align with sustainability goals (Sangarya et al., 2024).

In addition to the challenges previously mentioned, numerous methodological limitations of existing ML applications in real estate call for careful consideration. A major concern is the reliance on extensive, high-quality datasets to train sophisticated models such as neural networks. In markets with limited data or highly localized data, these models may overfit, which diminishes their generalizability and predictive precision (Galante et al., 2024). Conversely, simpler models such as multiple linear regression can occasionally outperform more complex ones when dealing with limited data. This highlights the importance of selecting models based on dataset quality. Additionally, machine learning-driven pricing models can unintentionally establish feedback loops, where valuations generated by algorithms affect future assessments, leading to compounded errors and skewing market dynamics (Malik & Manzoor, 2023). Such dynamics highlight the risk of over-reliance on automated predictions without sufficient human oversight.

6. Discussion and future directions

ML has transformed real estate practices by improving predictive capabilities, operational efficiency, and decision-making tools. This comprehensive study found that real estate ML applications are mostly predictive, descriptive, prescriptive, and automation and operational efficiency. These categories use ML methods like supervised learning, unsupervised learning, RL, and deep learning to solve residential, commercial, and financial problems.

The results show that innovative algorithms like NNs, random forests, and regression models are essential for predicting market trends, property prices, and demand-supply dynamics. Descriptive analytics improves market segmentation and trend identification by clustering and dimensionality reduction. Prescriptive analytics optimizes investment strategies and risk management using RL and ensemble methods. IoT and machine learning-driven systems boost operational efficiency and automation, simplifying property management and tenant interactions.

The study identified data quality, algorithmic bias, privacy and security, black-box model transparency, the need to balance automation with human interaction, accessibility gaps, and environmental impacts as major challenges. Given these challenges, ML must be used in a balanced way to meet its practical and ethical constraints.

Machine learning in real estate has endless potential for growth and change. Notable prospects include:

a) *Improved predictive models*: with the ongoing progress in ML algorithms, especially in deep learning and ensemble methods, predictive models will become even more accurate and resilient.
b) *Integration of sustainability frameworks*: future research should focus on embedding ESEG criteria into ML models, ensuring that real estate developments are not only economically viable but also environmentally sustainable and socially equitable.

c) *Interdisciplinary approaches*: combining ML with other emerging technologies, such as blockchain for transparent transactions and augmented reality for virtual property tours, can create more comprehensive and integrated real estate solutions.
d) *Make ML solutions scalable and accessible*: to make ML solutions more widely used in the real estate industry, we need to build scalable ML platforms that meet the needs of all stakeholders, including small and medium-sized businesses.
e) *Develop AI with ethics in mind*: to earn people's trust and make sure ML apps are good for society, developers must prioritize ethical AI practices like transparency, accountability, and fairness.
f) *Smart cities and IoT integration*: as urbanization accelerates, ML will play a pivotal role in the development of smart cities by optimizing infrastructure, enhancing sustainability, and improving the quality of life for residents through intelligent real estate management systems.

Future research should explicitly incorporate ESEG dimensions criteria to foster sustainable solutions. Improved predictive and prescriptive analytics must embed clear environmental effects metrics, including carbon emissions and energy efficiency, within real estate decision-making frameworks (Walacik & Chmielewska, 2024). Furthermore, greater emphasis must be placed on social aspects like inclusivity and equity within algorithmic practices and market interventions, guaranteeing that ML-driven solutions foster social justice and enhance community well-being (Filippi & Carbone, 2025). By utilizing blockchain and AI-based compliance monitoring technologies, governance can be improved by enhancing transparency, accountability, and traceability. A multidisciplinary strategy that merges machine learning with sustainability frameworks, urban planning, and ethical technology practices will be essential for overcoming existing challenges and constructing resilient urban areas that align with global sustainable development goals.

7. Conclusions

This research delves into the evolution of ML applications in the real estate sector, highlighting significant advancements while pinpointing critical gaps related to sustainability. Even though there has been notable progress in predictive, descriptive, and prescriptive analytics and operational automation, most existing applications focus on technological efficiency and economic outcomes, often neglecting the full integration with ESEG dimensions. This study identifies key themes, trends, and methodological enhancements, critically assessing their link to ESEG dimensions. The findings indicate that ML-driven predictive and descriptive analytics are well-established and effective in addressing economic challenges; however, their ability to incorporate broader sustainability metrics remains significantly under-arched. Similarly, prescriptive analytics and automation techniques show promising potential but necessitate a clear incorporation of sustainability elements to enhance their value for stakeholders and society. To effectively confront these shortcomings, future research and industry practices must emphasise intentionally integrating ESEG metrics into ML models, particularly by incorporating indicators of environmental performance (like carbon emissions and energy consumption), social outcomes (such as equity and community health), and governance transparency (including ethical data usage and regulatory compliance). Adopting interdisciplinary approaches and leveraging advancements like blockchain technology for improved governance, the IoT for enhanced environmental results, and ethical AI frameworks for equitable social impacts is essential. By consciously including sustainability standards, ML applications in real estate can transcend traditional efficiency-focused strategies and make meaningful contributions towards achieving global Sustainable Development Goals (SDGs). This integrated perspective will strengthen market resilience and operational effectiveness and foster ethical accountability, environmental stewardship, and social equity in real estate practices, ultimately aiding in developing more sustainable and equitable urban landscapes.

Acknowledgements

This work was supported by A Study on the Development Path of New-Quality Productive Forces Based on the Theory of Productive Forces in Das Kapital (Project No. 2024ZZ012).

References

Akinjole, A., Shobayo, O., Popoola, J., Okoyeigbo, O., & Ogunleye, B. (2024). Ensemble-based machine learning algorithm for loan default risk prediction. *Mathematics*, *12*(21), Article 3423. https://doi.org/10.3390/math12213423

Alkhaldi, F. (2023). Maximizing business potential: A framework for implementing prescriptive analytics. In S. G. Yaseen (Ed.), *Cutting-edge business technologies in the big data era* (pp. 251–259). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42455-7_23

Altarturi, H. H. M., Nor, A. R. M., Jaafar, N. I., & Anuar, N. B. (2025). A bibliometric and content analysis of technological advancement applications in agricultural e-commerce. *Electronic Commerce Research*, *25*, 805–848. https://doi.org/10.1007/s10660-023-09670-z

Altarturi, H. H. M., Saadoon, M., & Anuar, N. B. (2020). Cyber parental control: A bibliometric study. *Children and Youth Services Review*, *116*, Article 105134. https://doi.org/10.1016/j.childyouth.2020.105134

Altarturi, H. H. M., Saadoon, M., & Anuar, N. B. (2023). Web content topic modeling using LDA and HTML tags. *PeerJ Computer Science*, *9*, Article e1459. https://doi.org/10.7717/peerj-cs.1459

Asif, M., Naeem, G., & Khalid, M. (2024). Digitalization for sustainable buildings: Technologies, applications, potential, and challenges. *Journal of Cleaner Production*, *450*, Article 141814. https://doi.org/10.1016/j.jclepro.2024.141814

Bechina, K., & Arntzen, A. A. (2022). A system of systems approach to smart building management: An AI vision for facility management. In *2022 17th Annual System of Systems Engineering Conference* (*SOSE*) (pp. 452–456). IEEE. https://doi.org/10.1109/SOSE55472.2022.9812696

Bhattacharyya, S., Bhattacharyya, S., & Kaur, P. (2018). *System and method for performing descriptive analysis in data mining* (United States Patent No. US20180225390A1). https://patents.google.com/patent/US20180225390A1/en

Boz, H. A., Bahrami, M., Balcisoy, S., & Pentland, A. (2023). Footfall prediction using graph neural networks. In *2023 31st Signal Processing and Communications Applications Conference* (*SIU*) (pp. 1–3). IEEE. https://doi.org/10.1109/SIU59756.2023.10224021

Breuer, W., & Steininger, B. I. (2020). Recent trends in real estate research: A comparison of recent working papers and publications using machine learning algorithms. *Journal of Business Economics*, *90*(7), 963–974. https://doi.org/10.1007/s11573-020-01005-w

Chaudhry, M., Shafi, I., Mahnoor, M., Vargas, D. L. R., Thompson, E. B., & Ashraf, I. (2023). A systematic literature review on identifying patterns using unsupervised clustering algorithms: A data mining perspective. *Symmetry*, *15*(9), Article 1679. https://doi.org/10.3390/sym15091679

Chen, W., Farag, S., Butt, U., & Al-Khateeb, H. (2024). Leveraging machine learning for sophisticated rental value predictions: A case study from Munich, Germany. *Applied Sciences*, *14*(20), Article 9528. https://doi.org/10.3390/app14209528

Cheng, S., Blanco, M. N., Larson, T. V., Sheppard, L., Szpiro, A., & Shojaie, A. (2024). *Principal component analysis balancing prediction and approximation accuracy for spatial data* (No. arXiv:2408.01662). arXiv. https://doi.org/10.48550/arXiv.2408.01662

Cholaraja, K., M, V., B, S., Vishal S, R. M., & S, R. N. (2024). Justrent (a house rental management system). In *2024 10th International Conference on Advanced Computing and Communication Systems* (*ICACCS*) (pp. 2483–2487). IEEE. https://doi.org/10.1109/ICACCS60874.2024.10717280

Conway, J. (2018). *Artificial intelligence and machine learning*: *Current applications in real estate* [Master's thesis, Massachusetts Institute of Technology]. DSpace@MIT. https://dspace.mit.edu/handle/1721.1/120609

Darko, A., Glushakova, I., Boateng, E. B., & Chan, A. P. C. (2023). Using machine learning to improve cost and duration prediction accuracy in green building projects. *Journal of Construction Engineering and Management*, *149*(8), Article 04023061. https://doi.org/10.1061/JCEMD4.COENG-13101

Elias, R., & Issa, R. R. A. (2024). Machine learning-based generative design optimization of the energy efficiency of Florida single-family houses. In *Computing in Civil Engineering 2023* (pp. 640–648). ASCE. https://doi.org/10.1061/9780784485224.077

Filippi, F. D., & Carbone, C. (2025). Digital technologies for urban and social innovation in affordable, smart, and sustainable neighbourhoods. In *Recent advances and prospects in urban e-planning* (pp. 349–382). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-5956-3.ch014

Gadipudi, S. B., & Kalpana Kalaimani, R. (2024). Reinforcement learning for dynamic pricing under competition for perishable products. In *2024 28th International Conference on System Theory*, *Control and Computing* (*ICSTCC*) (pp. 297–302). IEEE. https://doi.org/10.1109/ICSTCC62912.2024.10744760

Galante, M., Giove, S., & Rosato, P. (2024). *Neural networks and linear models in real estate appraisal*: *The impact of sets selection procedures*. https://doi.org/10.48264/VVSIEV-20243505

Ganguly, P., & Mukherjee, I. (2024). Enhancing retail sales forecasting with optimized machine learning models. In *2024 4th International Conference on Sustainable Expert Systems* (*ICSES*) (pp. 884–889). IEEE. https://doi.org/10.1109/ICSES63445.2024.10762950

Ghasemi, M., & Ebrahimi, D. (2024). *Introduction to reinforcement learning* (No. arXiv:2408.07712). arXiv. https://doi.org/10.48550/arXiv.2408.07712

Gupta, S., & Agarwal, S. (2024). *Applications of machine learning and artificial intelligence in environmental*, *social and governance* (*ESG*) *sector* (SSRN Scholarly Paper No. 5086611). Social Science Research Network. https://doi.org/10.2139/ssrn.5086611

Huang, S. (2024). Advancing portfolio optimization: The convergence of machine learning and traditional financial models. *Applied and Computational Engineering*, *57*, 206–211. https://doi.org/10.54254/2755-2721/57/20241335

Hung, M.-T., & Lo, H.-C. (2024). Risk analysis of mortgage loan default for bank customers and AI machine learning. *Journal of Applied Finance* & *Banking*, *14*(6), 33–50. https://doi.org/10.47260/jafb/1463

Iyelolu, T. V., Agu, E. E., Idemudia, C., & Ijomah, T. I. (2024). Driving SME innovation with AI solutions: Overcoming adoption barriers and future growth opportunities. *International Journal of Science and Technology Research Archive*, *7*(1), 36–54. https://doi.org/10.53771/ijstra.2024.7.1.0055

Jafary, P., Shojaei, D., Rajabifard, A., & Ngo, T. (2024). Automating property valuation at the macro scale of suburban level: A multi-step method based on spatial imputation techniques, machine learning and deep learning. *Habitat International*, *148*, Article 103075. https://doi.org/10.1016/j.habitatint.2024.103075

Khandaskar, S., Panjwani, C., Patil, V., Fernandes, D., & Bajaj, P. (2023). House and rent price prediction system using regression. In *2023 International Conference on Sustainable Computing and Smart Systems* (*ICSCSS*) (pp. 1733–1739). IEEE. https://doi.org/10.1109/ICSCSS57650.2023.10169290

Kuppan, K., Bhaskar Acharya, D., & B, D. (2024). Foundational AI in Insurance and real estate: A survey of applications, challenges, and future directions. *IEEE Access*, *12*, 181282–181302. https://doi.org/10.1109/ACCESS.2024.3509918

Leng, R. (2024). Exploring AI's role in enhancing risk assessment models in financial quantitative trading. *Journal of Applied Economics and Policy Studies*, *12*, 1–5. https://doi.org/10.54254/2977-5701/12/2024110

Lubis, A. S., Zulfan, Chania, M. F., & Adha, I. M. (2024). Analysis of the use and application of mathematics in economics: Demand and supply functions. *Journal of Research in Mathematics Trends and Technology*, *6*(1), 16–23. https://doi.org/10.32734/jormtt.v6i1.17603

Malik, N., & Manzoor, E. (2023). *Does machine learning amplify pricing errors in the housing market? -- The economics of machine learning feedback loops* (No. arXiv:2302.09438). arXiv. https://doi.org/10.48550/arXiv.2302.09438

Mandati, R., Anderson, V., Chen, P., Agarwal, A., Dokic, T., Barnard, D., Finn, M., Cromer, J., Mccauley, A., Tutaj, C., Dave, N., Besharati, B., Barnett, J., & Krall, T. (2024). *Integrating artificial intelligence models and synthetic image data for enhanced asset inspection and defect identification* (No. arXiv:2410.11967). arXiv. https://doi.org/10.48550/arXiv.2410.11967

Mukherjee, A., Nambiar, R., & Choudhury, D. (2024). Machine learning based real estate price prediction. In *2024 8th International Conference on Inventive Systems and Control* (*ICISC*) (pp. 351–358). IEEE. https://doi.org/10.1109/ICISC62624.2024.00067

Nur, M. F., & Siregar, A. (2024). Exploring the use of cluster analysis in market segmentation for targeted advertising. *IAIC Transactions on Sustainable Digital Innovation* (*ITSDI*), *5*(2), 158–168. https://doi.org/10.34306/itsdi.v5i2.665

Panda, A. R., Rout, P., & Gautam, A. (2024). Optimizing marketing strategy by performing customer segmentation. In *2024 International Conference on Advances in Computing Research on Science Engineering and Technology* (*ACROSET*) (pp. 1–6). IEEE. https://doi.org/10.1109/ACROSET62108.2024.10743643

Piao, J. (2023). Portfolio optimization based on deep learning and factor constraints. *Advances in Economics*, *Management and Political Sciences*, *48*, 264–273. https://doi.org/10.54254/2754-1169/48/20230454

Purushotham, K., Bangarappa, Kodipalli, A., & Rao, T. (2024). Real-time house price predictions with regression analysis. In *2024 IEEE Recent Advances in Intelligent Computational Systems* (*RAICS*) (pp. 1–4). IEEE. https://doi.org/10.1109/RAICS61201.2024.10689962

Ram, K. S. R., Selvaganapathy, M., Nandhini, I., Tamilselvi, M., Kalaivani, M., & Natarajan, B. (2024). Enhanced investment decision making with a reinforcement learning-based multi-agent portfolio management system. In *2024 International Conference on Data Science and Network Security* (*ICDSNS*) (pp. 1–6). IEEE. https://doi.org/10.1109/ICDSNS62112.2024.10691210

Rane, N., Choudhary, S., & Rane, J. (2024). *Artificial intelligence driven approaches to strengthening environmental*, *social*, *and governance* (*ESG*) *criteria in sustainable business practices*: *A review* (SSRN Scholarly Paper No. 4843215). Social Science Research Network. https://doi.org/10.2139/ssrn.4843215

Rao, S. N., & Ravi, S. C. (2018). *System and method for identifying*, *monitoring and mitigating risks* (United States Patent No. US20180308174A1). https://patents.google.com/patent/US20180308174A1/en

Rodriguez-Serrano, J. A. (2025). Prototype-based learning for real estate valuation: A machine learning model that explains prices. *Annals of Operations Research*, *344*(1), 287–311. https://doi.org/10.1007/s10479-024-06273-1

Sahran, F., Altarturi, H. H. M., & Anuar, N. B. (2024). Exploring the landscape of AI-SDN: A comprehensive bibliometric analysis and future perspectives. *Electronics*, *13*(1), Article 26. https://doi.org/10.3390/electronics13010026

Sangarya, V., Bradford, R., & Kim, J.-E. (2024). Estimating environmental cost throughout model's adaptive life cycle. *Proceedings of the AAAI/ACM Conference on AI*, *Ethics*, *and Society*, *7*, 1281–1291. https://doi.org/10.1609/aies.v7i1.31723

Skovajsa, Š. (2023). Review of clustering methods used in data-driven housing market segmentation. *Real Estate Management and Valuation*, *31*(3), 67–74. https://doi.org/10.2478/remav-2023-0022

Sobana, P., Balakumaran, M., Bharathkumar, S., Boopathi, P., & Harish, J. (2024). House price prediction using machine learning. In *Challenges in information*, *communication and computing technology*. CRC Press. https://doi.org/10.1201/9781003559085-121

Sridharan, K., & Maddern, C. R. (2022). Methodology for realising ESG performance of corporate real estate assets through digital data architectures. *Corporate Real Estate Journal*, *11*(4), 347–362. https://doi.org/10.69554/ZBNP5967

Stroparo, T. R., Guerra, A. de L. e R., Cordeiro, E. da S., Bochniak, B., Bortolotti, M. A., & Júnior, O. da S. L. (2023). Objetivos de desenvolvimento sustentável (ODS), environmental, social and governance (ESG) e artificial intelligence (AI): Tríplice abordagem para a sustentabilidade corporativa. *AKRÓPOLIS - Revista de Ciências Humanas da UNIPAR*, *31*(2), Article 2. https://doi.org/10.25110/akropolis.v31i2.2023-11369

Tekouabou, S. C. K., Gherghina, Ş. C., Kameni, E. D., Filali, Y., & Idrissi Gartoumi, K. (2024). AI-based on machine learning methods for urban real estate prediction: A systematic survey. *Archives of Computational Methods in Engineering*, *31*(2), 1079–1095. https://doi.org/10.1007/s11831-023-10010-5

Tin, T. T., Wei, C. J., Min, O. T., Feng, B. Z., & Xian, T. C. (2024). Real estate price forecasting utilizing recurrent neural networks incorporating genetic algorithms. *International Journal of Innovative Research and Scientific Studies*, *7*(3), 1216–1226. https://doi.org/10.53894/ijirss.v7i3.3220

Vergara-Perucich, J.-F. (2023). A systematic bibliometric analysis of the real estate bubble phenomenon: A comprehensive review of the literature from 2007 to 2022. *International Journal of Financial Studies*, *11*(3), Article 106. https://doi.org/10.3390/ijfs11030106

Vivekananda, M. N., & Shidlyali, P. A. (2024). Enhancing housing price prediction using AI and machine learning: A stacked regression meta-modeling approach. In *2024 8th International Conference on I-SMAC* (*IoT in Social*, *Mobile*, *Analytics and Cloud*) (*I-SMAC*) (pp. 1954–1960). IEEE. https://doi.org/10.1109/I-SMAC61858.2024.10714734

Walacik, M., & Chmielewska, A. (2024). Real estate industry sustainable solution (environmental, social, and governance) significance assessment—AI-powered algorithm implementation. *Sustainability*, *16*(3), Article 1079. https://doi.org/10.3390/su16031079

Wang, F. (2023). The present and future of the digital transformation of real estate: A systematic review of smart real estate. *Business Informatics*, *17*(2), 85–97. https://doi.org/10.17323/2587-814X.2023.2.85.97

Yadav, V., Bhatnagar, R., & Kumar, U. (2024). Innovative computational techniques for DSSCs using machine learning: A review. *Discover Electronics*, *1*(1), Article 17. https://doi.org/10.1007/s44291-024-00022-1

Yang, J. (2024). Data-driven investment strategies in international real estate markets: A predictive analytics approach. *International Journal of Computer Science and Information Technology*, *3*(1), 247–258. https://doi.org/10.62051/ijcsit.v3n1.32

Zaman, M., Al Islam, M., Zohrabi, N., & Abdelwahed, S. (2024). A machine learning-based temperature control and security protection for smart buildings. In *2024 IEEE International Conference on Smart Computing* (*SMARTCOMP*) (pp. 290–295). IEEE. https://doi.org/10.1109/SMARTCOMP61445.2024.00070

Zheng, L., Mueller, M., Luo, C., & Yan, X. (2024). Predicting whole-life carbon emissions for buildings using different machine learning algorithms: A case study on typical residential properties in Cornwall, UK. *Applied Energy*, *357*, Article 122472. https://doi.org/10.1016/j.apenergy.2023.122472