Property Manag

Modelling the drivers of data science techniques for real estate professionals in the fourth industrial revolution era

Journal:	*Property Management*
Manuscript ID	PM-05-2022-0034.R1
Manuscript Type:	Research Paper
Keywords:	Asset valuation, Property management, Real estate, Investments

LIST OF TABLES

Table 1: 4IR potentials and disruptive technologies for real estate professionals

Authors	4IR potentials for real estate professionals	4IR disruptive technologies for real estate professionals
Gromova and Pupentsova (2020)	The technologies driven by the 4IR has the potentials of minimising the risk of real estate investment. Through using machine learning simulation, a real estate investor can pick a less risky investment.	Artificial intelligence using Machine learning
Kalyuzhnova (2018)	The blockchain technology assist in eliminating transaction cost associated with real estate investment. The transaction cost includes cost of information retrieval, negotiation cost and measurement cost.	Blockchain
Ottinger et al. (2021)	Data harvested from digital twin contributes to ensuring the maintenance and operational efficiency of real estate investment. The digital twin technology also supports the space optimisation of real estate investment	Digital twin
Baldominos et al. (2018)	The use of predictive models powered by machine learning assist real estate investors in identifying investment potentials in the real estate market. The model utilise machine learning to identify housing prices that are below the average market price. The model automatically synthesise various housing prices that are listed online and advise a prospective investor	Machine learning
Guan et al. (2014)	The use of Neuro fuzzy techniques driven by artificial intelligence assist in the prediction of rest estate prices based on vast accumulation of data.	Artificial intelligence using Neuro fuzzy
Yeh and Hsu (2018)	The case base reasoning is like the traditional comparative analysis method. But it utilises the principle of artificial intelligence in determining the comparative value of the property. The case base reasoning method assists in eliminating the bias associated with traditional comparative method.	Case based reasoning
Lee and Park (2020)	The rise of 4IR created new data that has the ability for ensuring accurate prediction of property prices. The study proposes the introduction of property visual information along with existing meta data like building size, age, number of bedrooms and others. The visual information was analysed using convolutional neural network.	Convolutional neural network
Law et al. (2019)	The study introduced the use of deep neural network model to automatically extract visual features from images aimed at estimating housing values in the United Kingdom. The deep learning driven by the 4IR provided the opportunity to accurately measure intangible factors that determine the value of real estate investment.	Deep learning

Table 2: Benefits of adopting modern data analysis techniques in the real estate practice

Benefits	Source	SEM coding
Reliability	Provost and Fawcett (2013)	EDA1
Accuracy	Lee and Park (2020)	EDA2
Comprehensiveness	Poursaeed et al. (2018) and Soman and Whyte (2020)	EDA3
Predictive capacity	Dhar (2013) and McCluskey et al. (2013)	EDA4
Inform planning policies	DeLisle et al. (2020)	EDA5
Enhance urban planning	Soman and Whyte (2020)	EDA6
Consistency	Zurada et al. (2011) and Provost and Fawcett (2013)	EDA7
Proper planning	Lee and Park (2020)	EDA8
Measure intangible assets	Soman and Whyte (2020)	EDA9

Table 3: Traditional and modern data analysis technique

Method	Technique	Source
Traditional	Investment/income analysis	Peterson and Flanagan (2009) Lin and Mohan (2011), Zurada et al. (2011), McCluskey et al. (2013) and Valier (2020)
	Discounted cash flow analysis	
	Sales comparison approach	
	Profit method	
	Residual/cost method	
	Multiple regression analysis	
Advanced technique originating from data science	Artificial neural network method	Guan et al. (2014), Niu and Niu (2019), Law et al. (2019), DeLisle et al. (2020), Lee and Park (2020), Valier (2020) and Khodabakhshian and Toosi (2021)
	Genetic algorithm	
	Fuzzy logic	
	Case based reasoning	
	Convolutional neural network	
	Spatial analysis method	
	Metaverse analysis based on non-fungible token (NFT)	Researcher (2021)

Table 4: Possible drivers for application of data science techniques for real estate professionals

Drivers	Literature sources
Use of blockchain for information sharing	Kalyuzhnova (2018) and Wouda and Opdenakker (2019)
Decentralised real estate structure	Kasprzak (2021)
Introduction of transdisciplinary skills in the real estate curriculum	Kim (2019) and Aliu and Aigbavboa (2020)
Real estate valuation exercise in classrooms	Kim (2019)
Data structure	DeLisle et al. (2020)
Virtual real estate market	Duan et al. (2021) and Ottinger et al. (2021)
Control of data	Sanquist et al. (2018)
Fragmented data	DeLisle et al. (2020)
Transparent real estate transaction history	Kasprzak (2021)
Autonomy in decision making	Ottinger et al. (2021) and Kasprzak (2021)
Utilisation of Proptech for real estate market decision	Poursaeed et al. (2018)
Training programs opportunities for students	Buglea and Şipoş (2015)
Inviting guest speakers from the industry	Aliu and Aigbavboa (2020)
Adoption of decentralised currency	Wouda and Opdenakker (2019)
Data synchronisation system	Provost and Fawcett (2013)
Data accessibility	DeLisle et al. (2020) and Molnar et al. (2019)
Introduction of modern valuation techniques into the educational system	Buglea and Şipoş (2015)
Restrictive culture	Sanquist et al. (2018)
Elimination of middlemen	(Buglea and Şipoş, 2015) and Eberhard et al. (2017)
Data quality	Peyré and Cuturi (2019) and DeLisle et al. (2020)

Table 5. Rotated component matrix

	Component			
	1	2	3	Variance explained
Use of blockchain for information sharing **(NRM1)**	0.853			**34.2%**
Virtual real estate market **(NRM2)**	0.849			
Adoption of decentralised currency **(NRM3)**	0.782			
Autonomy in decision making **(NRM4)**	0.763			
Utilisation of Proptech for real estate market decision **(NRM5)**	0.721			
Elimination of middlemen **(NRM6)**	0.684			
Decentralised real estate structure **(NRM7)**	0.654			

Transparent real estate transaction history **(NRM8)**	0.627			
Fragmented data **(NRM9)**	0.548			
Data synchronisation system **(DM1)**		0.781		
Data accessibility **(DM2)**		0.758		
Data structure **(DM3)**		0.693		**22.3%**
Data quality **(DM4)**		0.642		
Control of data **(DM5)**		0.584		
Restrictive culture **(DM6)**		0.510		
Introduction of modern valuation techniques into the educational system **(IES1)**			0.822	
Introduction of transdisciplinary skills in the real estate curriculum **(IES2)**			0.759	**11.7%**
Inviting guest speakers from the industry **(IES3)**			0.658	
Training programs opportunities for students **(IES4)**			0.552	
Real estate valuation exercise in classrooms **(IES5)**			0.510	

Table 6: CFA outer loadings and convergent analysis

Construct	Coding	Outer loadings	Composite reliability	AVE
Decentralised real estate market	DRM 1	0.810	0.946	0.685
	DRM 2	0.850		
	DRM 3	0.800		
	DRM 4	0.860		
	DRM 5	0.880		
	DRM 6	0.820		
	DRM 7	0.780		
	DRM 8	0.820		
Data management	DM 1	0.890	0.930	0.729
	DM 2	0.920		
	DM 3	0.880		
	DM 4	0.830		
	DM 5	0.730		

Inclusive educational system	IES 1	0.880	0.930	0.729
	IES 2	0.850		
	IES 3	0.900		
	IES 4	0.780		
Effective data analysis	EDA 1	0.880	0.947	0.782
	EDA 7	0.900		
	EDA 2	0.910		
	EDA 4	0.890		
	EDA 5	0.840		

Table 7: Discriminant validity analysis

	Effective data analysis	Decentralised real estate market	Data management	Inclusive educational system
Effective data analysis	**0.885**			
Decentralised real estate market	0.727	**0.828**		
Data management	0.773	0.779	**0.854**	
Inclusive educational system	0.811	0.762	0.802	**0.853**

Table 8. Model fit indices assessment

Fit indices	Recommended threshold	Final model
SRMR	0.05 to 0.07	0.056
TLI	0.95 to 1.00	0.962
CFI	0.93 to 1.00	0.944
RMSEA	0.05 to 0.08	0.062
NFI	0.60 to 1.00	0.930
IFI	0.93 to 1.00	0.945

Table 9: Structural model assessment and hypothesis testing

Impacts	Estimate	S.E.	C.R.	P	Remark

Enhanced data analysis << Decentralised real estate market	0.231	0.032	7.257	0.000	Significant effect
Enhanced data analysis << Data management	0.298	0.028	10.703	0.000	Significant effect
Enhanced data analysis << Inclusive educational system	0.471	0.034	13.749	0.000	Significant effect

LIST OF FIGURES

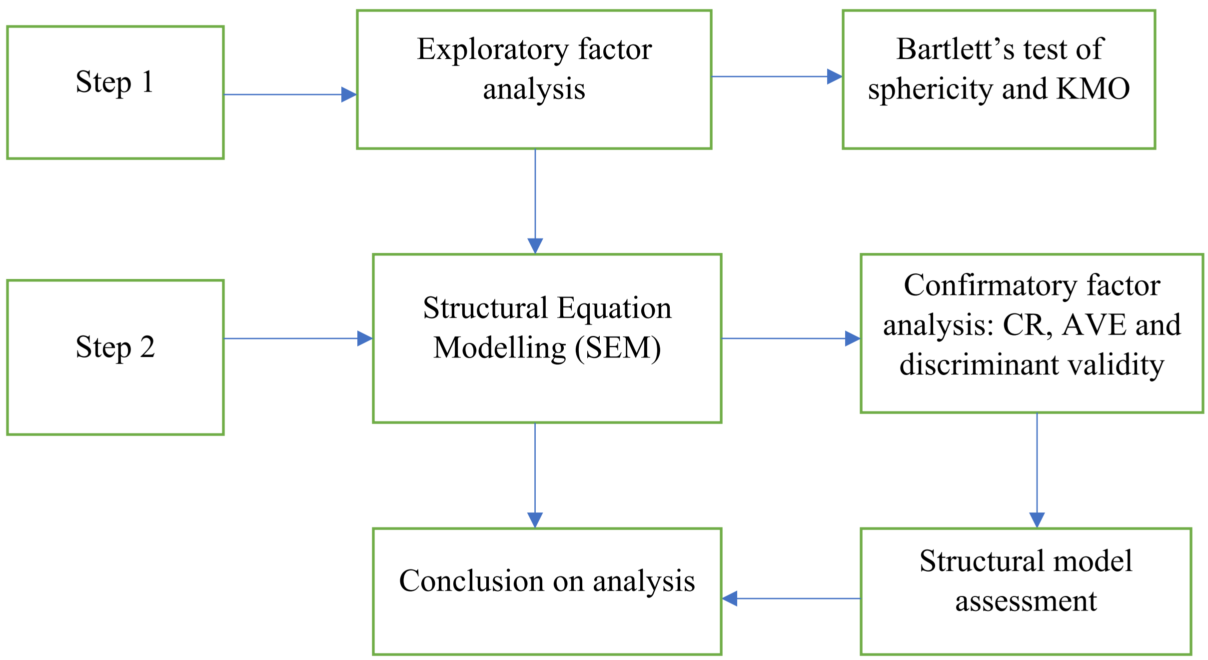

Figure 1: Research methodology framework

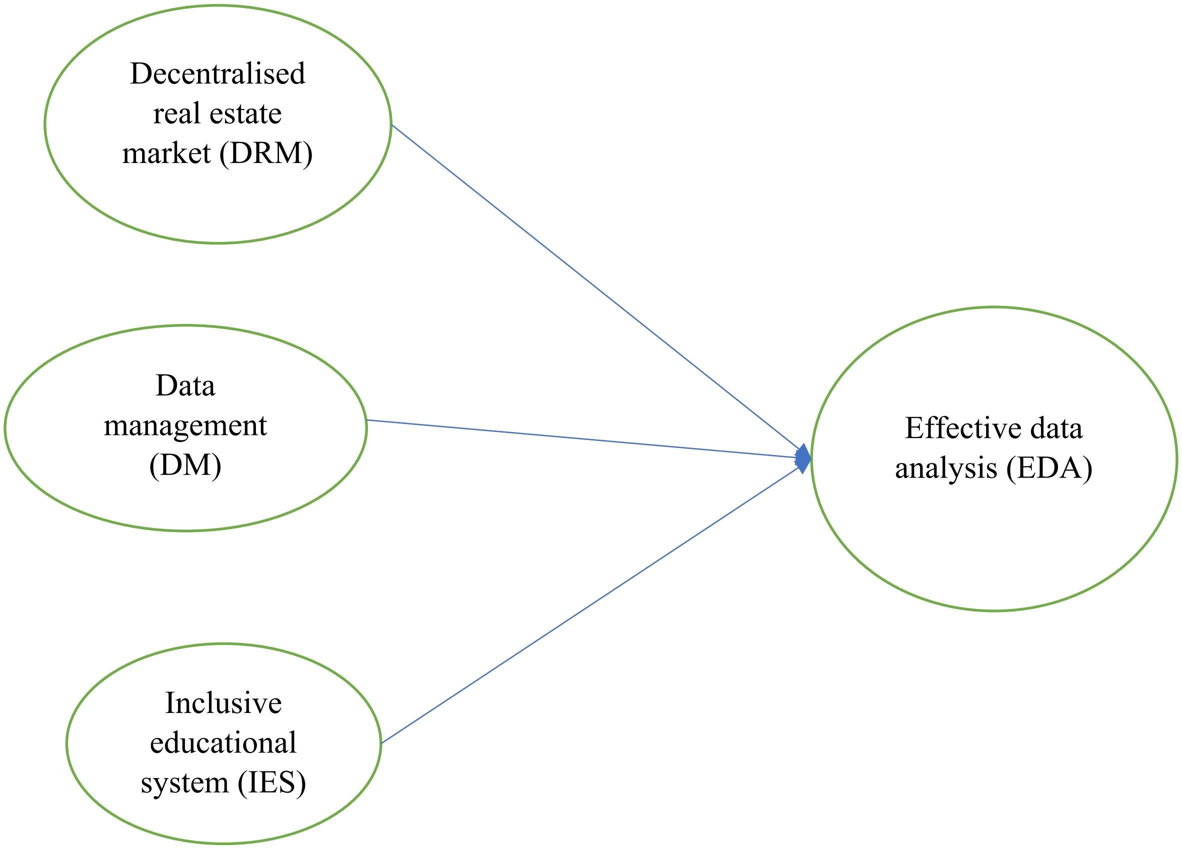

Figure 2: Conceptual framework of the study

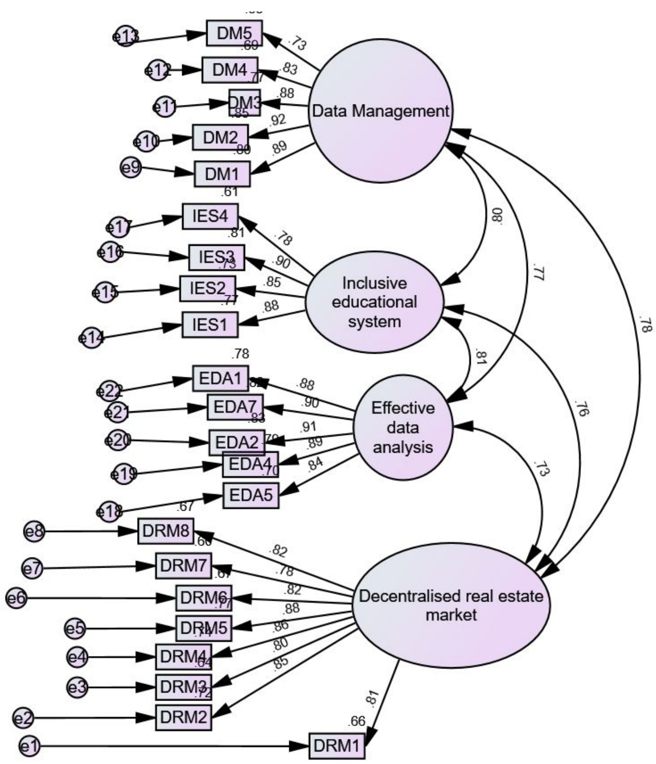

Figure 3: Confirmatory factor analysis

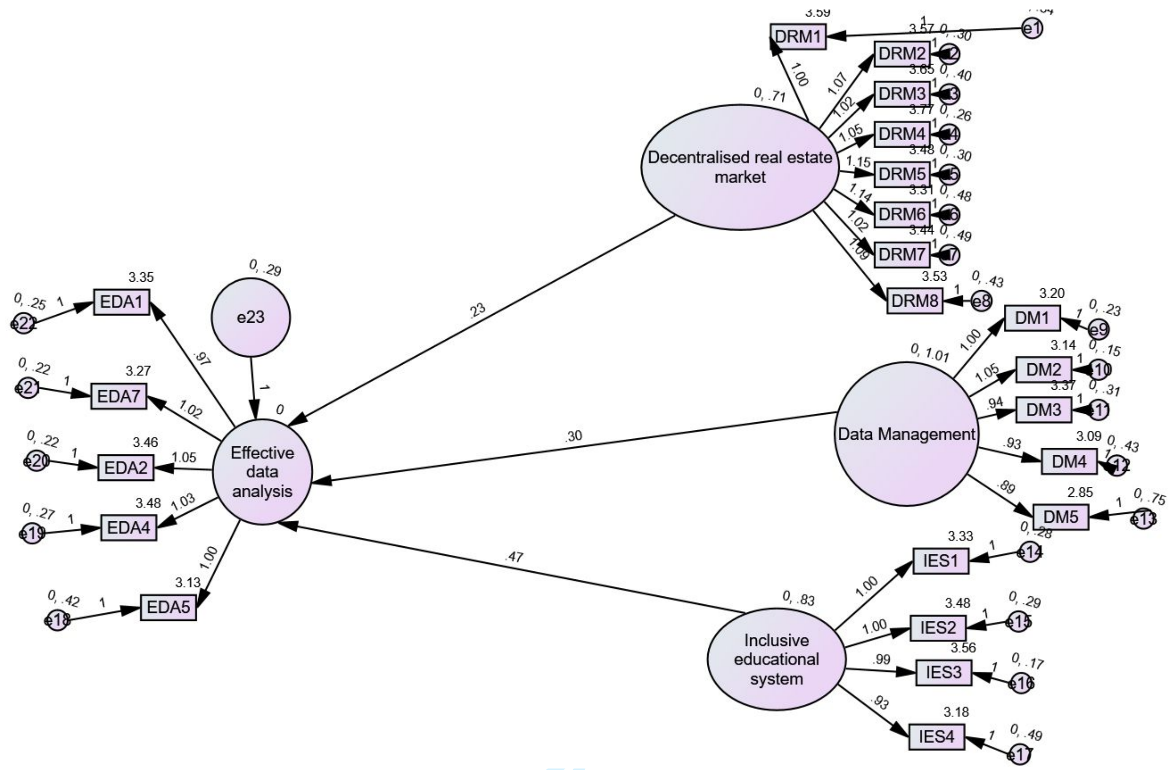

Figure 4: Structural model assessment

Modelling the drivers of data science techniques for real estate professionals in the fourth industrial revolution era

Abstract

Purpose- The study aims to develop a model that supports the application of data science techniques for real estate professionals in the fourth industrial revolution (4IR) era. The present 4IR era gave birth to big data sets and is beyond real estate professionals' analysis techniques. This has led to a situation where most real estate professionals rely on their intuition while neglecting a rigorous analysis for real estate investment appraisals. The heavy reliance on their intuition has been responsible for the underperformance of real estate investment, especially in Africa.

Design/methodology/approach – This study utilised a survey questionnaire to randomly source data from real estate professionals. The questionnaire was analysed using a combination of SPSS V24 and AMOS graphics V27 software. Exploratory factor analysis was employed to break down the variables (drivers) into meaningful dimensions helpful in developing the conceptual framework. The framework was validated using covariance-based structural equation modelling. The model was validated using fit indices like discriminant validity, SRMR, CFI, NFI, etc.

Findings- The model revealed that an inclusive educational system, decentralised real estate market, and data management system are the major drivers for applying data science techniques to real estate professionals. Also, real estate professionals' application of the drivers will guarantee an effective data analysis of real estate investments.

Originality/value- Numerous studies have clamoured for adopting data science techniques for real estate professionals. There is a lack of studies on the drivers that will guarantee the successful adoption of data science techniques. A modern form of data analysis for real estate professionals was also proposed in the study.

Keywords: Data science; Decentralised market; Fourth industrial revolution; Inclusive educational system; Real estate professional;

1. Introduction

The origin of the fourth industrial revolution (4IR) has been well contested across the literature (Sanquist et al., 2018, Barreto et al., 2017). However, the consensus is that it originated from the German manufacturing industry and was first coined by Klaus Schwab (Osunsanmi et al., 2018). Evidence from the extant literature and practice revealed that we are presently in the 4IR. Before arriving at the current revolution, the manufacturing industry has moved through three revolutions: mechanisation, electrification, and digitalization (Fragapane et al., 2020). Each of these revolutions was driven by unique principles and technologies. Eberhard et al. (2017) and Osunsanmi et al. (2021b) affirmed that a unique phenomenon of the 4IR era is the seamless connectivity and communication between machine to machine and machine to human. The seamless connectivity ensures the creation of large data sets that can effectively predict real estate investment (Choi et al., 2019).

Sanquist et al. (2018) argued that accurate prediction of real estate investment in the 4IR era is hinged on data availability. Kasprzak (2021) believed that a real estate investor needs more data in the 4IR period because the current revolution has disrupted the traditional data that affect property prices. The conventional data like location, number of bathrooms, neighbourhood crime, building age, size of the rooms, and others can not sufficiently predict the value of a real estate investment. Molnar et al. (2019) and Lee and Park (2020) revealed that current data like property picture, internet speed, and indoor environmental quality affect the choice of investing in real estate. Unfortunately, these recent data are challenging to analyse and sort in the first, second, and third industrial revolution era.

Fortunately, the analysis of data, sorting, and presentation has significantly improved in the 4IR era through the principles and technologies that originated from the domain of data science (Akwang and Chimah, 2021). This study describes data science as the discipline focused on using data for predicting an observation or event that will occur in the future. Provost and Fawcett (2013) describe data science as an interdisciplinary field utilising algorithms and scientific methods to extract knowledge from structured and unstructured data. The field, principles, and domain of data science boomed owing to the need to input heterogeneous and unstructured data like text, images, and videos for decision-making (Dhar, 2013). On the other hand, Peyré and Cuturi (2019) opined that the automatic generation of data by computers intended for usage by another computer created the field of professionals (data scientists) responsible for making meaningful information from the data. Data scientist adopts techniques like machine learning, deep learning, artificial intelligence, artificial neural networks, and others to create meaningful information from the large data set(Dhar, 2013)

Moreover, Niu and Niu (2019) affirmed that the rise of 4IR made it possible to utilise data science techniques like machine learning, artificial intelligence, and others for managing real estate investment. Machine learning techniques can accurately predict real estate for investment purposes (Baldominos et al., 2018, Lee and Park, 2020). Despite the potential of machine learning and other big data technologies originating from data science, application by real estate professionals is relatively scarce (Lin and Mohan, 2011). Also, Ajibola and Ogungbemi (2011) reported that most real estate stakeholders' decisions in Africa are not guided by data and thorough market analysis using modern techniques. Osunsanmi et al. (2021a) attributed the underperformance of real estate investments in Africa to the poor real estate market analysis using traditional multiple regression analysis.

Furthermore, Babawale (2013) indicated that multiple regression analysis is adopted for mass property tax appraisal in developing countries, especially in Africa. Unfortunately, the multiple regression analysis has been criticised for its weakness in modelling extensive real estate data features (Peterson and Flanagan, 2009, Guan et al., 2014). McCluskey et al. (2013) affirmed that multiple regression analysis's shortcomings are nonlinearity, multicollinearity, and heteroscedasticity. Yeh and Hsu (2018) opined that multiple regression analysis could not handle factors that are difficult to quantify. Antoniades (2021) related the use of multiple regression to the underperformance of real estate investments, which contributed to commercial bank failures. It affected commercial banks because real estate investors' decisions were not supported with effective predictive models to cater to disasters or economic recessions. It can be inferred that predictive models functions as a vaccine for the performance of real estate investment

Hence, to enhance the performance of real estate investment, studies such as (Gromova and Pupentsova, 2020, Baldominos et al., 2018, Guan et al., 2014) have proposed using data science techniques like machine learning, convolutional neural network, and others. Other studies, see (Valier, 2020, McCluskey et al., 2013, Peterson and Flanagan, 2009), compared the techniques originating from the domain of data science in the 4IR era and traditional multiple regression analysis. A common finding from the articles is that there is no enabling environment for adopting data science techniques for real estate professionals. Given the above and towards establishing an enabling environment and bridging the gap in the literature, this study focuses on modelling the drivers of data science for real estate professionals in the 4IR era using structural equation modelling. Our study is different from a similar study conducted by Soman and Whyte (2020) that examined data science challenges in construction. Their study focuses on the hindrance to adopting data science for construction professionals.

The findings from this study contribute significantly to academics and practice. To practice, this study provides the variable that supports the seamless application of data science techniques for real estate professionals. To knowledge, the findings from this study contribute to educating real estate professionals on the benefits of adopting data science techniques. The review of the extant literature contributes to understanding the critical technologies of the 4IR that are responsible for the disruptive potential of real estate investors. The subsequent section in this article presents the literature review related to the subject matter.

2. Real estate professionals practice in the fourth industrial revolution era

The fourth industrial revolution (4IR), also known as industry 4.0, originates in the manufacturing industry. Before the rise of the 4IR and its capacity to disrupt numerous industries, it has passed through three revolutions: mechanisation, electrification, and digitalisation. The first revolution (mechanization) replaced manual work with the invention of a steam engine. This revolution commenced around 1760, and the steam engine enabled the transition from farming and feudal society to a complex manufacturing process (Hirschi, 2018). The second revolution enabled the mass production of goods and services using electric energy (Osunsanmi et al., 2018). At the same time, the third revolution, built upon the second revolution supports the automation of activities within the manufacturing industry with information based on computers and the internet (Min et al., 2019, Maynard, 2015). The third revolution changed the pattern of making products that supported the screwing and welding of parts. We are presently in the 4IR era characterized by the seamless connection of the machine to machine and people to machine.

Morrar et al. (2017) affirmed that the industry 4.0 era is a world in which individuals move seamlessly within the digital domains and offline reality through the aid of connected technologies. Osunsanmi et al. (2020b) described 4IR as the apex industrial revolution responsible for its seamless connection of activities that disrupted the manufacturing industry's activities. Maynard (2015) believed that the ability of the 4IR era to integrate people fully and digitally controlled machines with the internet of things had provided an enormous advantage for the manufacturing industry. Likewise, Hirschi (2018) indicated that the ability of industry 4.0 to disrupt the production process in the manufacturing industry has contributed to the sector's growth. It can be inferred that the 4IR benefited the manufacturing industry due to its ability to introduce disruptive technologies.

Other sectors also benefited from the disruptive technologies of the 4IR. For instance, the financial sector activities have witnessed some disruption since the origin of the 4IR. Kosba et al. (2016) submitted that the 4IR disrupted the financial sector by introducing a blockchain technology known as the centralized ledger. The adoption of blockchain technology leads to the elimination of intermediaries. It supports the creation of a cryptographically secure ledger connected by computes that automatically verify a transaction before it is recorded. Recently, Osunsanmi et al. (2021b) revealed the benefit of the industry 4.0 technologies for managing the supply chain in the construction industry, which guarantees the resilience of the construction supply chain. As shown in Table 1, the literature review also revealed the benefits or the disruptive powers of 4IR technologies for real estate stakeholders and investments.

Insert Table 1 here

The studies presented in Table 1 were extracted from the Scopus database and focused on revealing the potential of 4IR for real estate professionals. The findings from these studies showed that the rise of the 4IR has disrupted the data needed for real estate professionals. Kalyuzhnova (2018) and Ottinger et al. (2021) discovered that 4IR technologies like blockchain and digital twin generate a new form of data that has the potential to determine the decisions made by real estate professionals. Ottinger et al. (2021) affirmed that data harvested from digital twin ensures real estate investment maintenance and operational efficiency. Unfortunately, analysing this new form of data created problems for real estate professionals as the traditional regression analysis can not sufficiently handle the modern form of data (Kasprzak, 2021). This implies that for an effective performance of real estate professionals in the 4IR era there is a need to collaborate with other professionals.

In support of the aforementioned Frey and Osborne (2017) recognised that the activities performed by real estate professionals are susceptible to changes or mutation, thereby creating a relationship between real estate professionals and other professionals. Similarly, Ajibola and Ogungbemi (2011) reported that a real estate professional could not perform effectively in isolation due to the numerous technologies that have made the word flat. Furthermore, the technologies driven by the 4IR have changed the world to a global village, creating the need for real estate to collaborate with other professionals (Kasprzak, 2021). This study calls for the collaboration of real estate professionals with data scientists focused on using techniques originating from their domain of making real estate informed decision

2.1 Adoption of data science techniques for real estate professionals

The 4IR has facilitated the creation of supercomputers that can collect and store a large amount of virtual data in an organised format (Akwang and Chimah, 2021). Unfortunately, deciding on a large amount of data remains challenging for an estate surveyor and valuer (Gromova and Pupentsova, 2020). Likewise, Kim (2019) asserted that estate surveyors find it difficult to create a helpful pattern from an extensive data set based on their curriculum. Similarly, Poursaeed et al. (2018) believed that professionals in the real estate practice find it difficult to make sense of big data without machine learning techniques. Lee and Park (2020) alleged that models based on machine learning proliferated the real estate practice owing to the boom of data scientists and the big data paradigm that was propelled by the 4IR. Guan et al. (2014) revealed that data mining and analysis approaches are driven by artificial intelligence that originated from data scientists has been tested as a preferable solution for real estate professionals for analysing big data.

McCluskey et al. (2013) and Gromova and Pupentsova (2020) opined that different modern data analysis and mining techniques are driven by artificial intelligence utilised by real estate professionals. The most popular artificial neural network (ANN) originated from the data scientist domain established to replicate human cognitive functioning. McCluskey et al. (2013) and Valier (2020) describes ANN as a digitised model that functions as a processing device supporting automatic pattern recognition, anomaly detection, and object classifications. Provost and Fawcett (2013) indicated that data scientists utilise ANN to cope with noisy data and train data to learn and compute from previously unknown relationships. Lin and Mohan (2011) reported that ANN supersedes traditional statistical techniques due to its ability to examine complex relationships and accurately predict prices. Provost and Fawcett (2013) proclaimed the ability of ANN to recognise a pattern in its data set to enhance its predictive capacity. The predictive capacity of ANN has recently been the subject of discussion of several papers within the real estate.

Moreover, some studies (see Guan et al., 2014, McCluskey et al., 2013, Lin and Mohan, 2011, Zurada et al., 2011, Peterson and Flanagan, 2009) have appraised the use of ANN in real estate. During this period (2009-2015) most of the research is focused on comparing the performance of ANN with the traditional multiple regression analysis used by real estate professionals. Zurada et al. (2011) discovered that artificial neural networks (ANN) perform better than conventional regression analysis. Likewise, McCluskey et al. (2013) compared ANN with the traditional hedonic linear regression using a sample size of 2694 residential properties. The scholar discovered that ANN reported better accuracy than MRA in predicting a residential property's value. Furthermore, Peterson and Flanagan (2009) found that ANN provides accurate predictions of property prices.

Recent literature (2015 till date) further confirmed the accuracy of using ANN for predicting property prices (Niu and Niu, 2019, Lee and Park, 2020). During the same period, an improvement on the ANN was witnessed. Lee and Park (2020) improved ANN and introduced the convolutional neural network (CNN). Accordingly, the authors discovered that the values of real estate investment could be effectively predicted by adding visual information to the existing metadata of the property. Law et al. (2019) and Poursaeed et al. (2018) also improved on ANN using CNN to determine the value of real estate investment. Another development in analysis during this period was the use of case-based reasoning. Case-based reasoning is the ability to solve a new problem using the solution of a past solved problem (Guan et al., 2014). Yeh and Hsu (2018) discovered that case base reasoning ensures consistency in the predicted value of real estate investment compared to the traditional comparative method. Aside from predicting property values, data scientists techniques were also used in conducting real estate market. Baldominos et al. (2018) adopted machine learning for identifying real estate investment opportunities. Table 2 presents the benefits of adopting the different data analysis techniques emanating from the domain of data scientists.

Insert Table 2 here

Despite the benefits of adopting the modern analysis method, as shown in Table 2, some African real estate professionals have not fully embraced this concept. Peterson and Flanagan (2009) reported that real estate professionals' favours multiple regression-based analysis over other forms of analysis. McCluskey et al. (2013) attributed the favouritism to the black box architecture of ANN and other modern data analysis methods originating from data scientists, limiting its application for real estate professionals. Kim (2019) and Khodabakhshian and Toosi (2021) opined that using ANN poses some challenges to real estate professionals due to the system's physical architecture and learning methods. Likewise, Zhang and Gao (2019) reported that the educational curriculum of real estate professionals had prevented the adoption of modern data analysis and construction management practices. To ensure that real estate professionals embrace the modern and advanced data analysis approach shown in Table 3, this study model the drivers that create an enabling environment for using the current data analysis.

Insert Table 3 here

The traditional and advanced data analysis methods available to real estate professionals are provided in Table 3. Peterson and Flanagan (2009) and Ajibola and Ogungbemi (2011) affirmed that real estate professionals' adoption of traditional analysis originated from their curriculum. Kalyuzhnova (2018) asserted that transformation from conventional analysis to advanced technique originating from data science became crucial owing to the growth of the 4IR. The 4IR ushered the possibility of using artificial intelligence and complex algorithm to analyse real estate data (Kasprzak, 2021). This study proposed a new data analysis technique called the metaverse analysis based on the blockchain network. The metaverse became popular in the early 2020s due to its merging virtual, augmented, and physical reality. The metaverse came with the opportunity to purchase property and other assets using the blockchain (Duan et al., 2021). The possibility of owning assets created the need for real estate professionals to value properties on the metaverse. Thus, this study models the drivers capable of ensuring real estate professionals value properties in the metaverse. The subsequent section examines the research methodology adopted in this study.

3. Research methods

The rise of the 4IR altered the real estate profession by introducing new data and analysis techniques from the data science domain (Gromova and Pupentsova, 2020, Kasprzak, 2021). This technique has been known for its accuracy in predicting real estate investment value. Accuracy in predicting the price of real estate investment is vital to all sectors of a nation's economy. The mortgage finance system and the property tax of a nation depend on an accurate prediction of a real estate investment (Babawale, 2013, Antoniades, 2021). Unfortunately, most real estate professionals are still using the traditional analysis method, which has a less predictive capacity (Peterson and Flanagan, 2009). Thus, this study is focused on modelling the drivers that enhance the utilisation of data science techniques for real estate professionals to enhance the predictive capacity of real estate investment.

The study commences with understanding the subject matter through an in-depth synthesis of the literature to understand the drivers supporting adopting data science techniques for real estate professionals. Table 4 presents the drivers extracted from literature and was used in preparing the questionnaire that was distributed to the respondents. The respondents are made up of real estate professionals comprising facility managers, real estate agents, and construction project managers in South Africa. The respondents (real estate professionals) were selected using a random sampling technique. Johnson and Onwuegbuzie (2004) and Dieronitou (2014) affirmed that random sampling provides enormous research advantages. The technique allows the calculation of sampling errors which supports the selection of reputable respondents and generalisation of the findings to the entire population.

Insert Table 4 here

A total of 228 respondents were selected randomly after adopting Cochran's formula with a 90% confidence level and a margin error of a 0.5% proportion of the population. The sample was selected from the list of registered real estate professionals in Gauteng Province. The population size was deemed accurate after reviewing similar studies (Aliu and Aigbavboa, 2020, Aghimien et al., 2021) that were conducted in South Africa. However, only 216 respondents to the questionnaire accurately with their response subjected to accurate and thorough analysis. The questionnaire was administered using an online platform (google forms) that took months and weeks (February to mid-March) to complete. In testing or ascertaining that the questions posed in the questionnaire are reliable, a Cronbach Alpha was calculated. The Cronbach Alpha gave a value of 0.845 above the recommended threshold of 0.70. According to Tavakol and Dennick (2011), a Cronbach Alpha with a value of 0.70 above is acceptable. It can also be deduced that the data emanating from the questionnaire is reliable.

The questionnaire used in this study was broken down into three sections. The first section of the questionnaire focused on the respondents' personal information like educational qualification, profession, affiliation, and working experience. The findings from the analysis revealed that all the respondents are educated with different qualifications. However, the majority (56%) of the respondents possess BSc/BTech qualification, while 32% have a master's degree as their highest qualification, and the remaining proportion (12%) have a PhD as their highest degree. More than half of the respondents (54%) are estate surveyors and valuers. This, therefore, makes their response crucial to the study owing to the aims and objective of the study. Other professions like construction project managers and architects account for the respondents' remaining proportion.

Regarding the affiliation of the respondents, most of them (56%) are affiliated with the estate agency affairs board (EAAB) and real estate business owners of South Africa (REBOSA). This agency is responsible for regulating the valuation and analysis of real estate investment in South Africa, making their response significant to this study. Other respondents are affiliated with the council for the built environment (CBE), the Engineering Council of South Africa (ECSA), and South African Association for the Quantity Surveyors (SAQSP).

The second section of the questionnaire evaluated the respondent's level of agreement with the benefits of using data analysis emanating from data science. While the last section focused on the variables that were used in modelling the drivers of data science techniques for real estate professionals. A structural equation modelling approach was used in modelling the drivers. Grotzinger et al. (2019) affirmed that the concept of SEM originated in the early 1918's from the field of medical science and used to study the complex relationship of bones. However, the review of the extant literature revealed that the application of SEM has moved from science to other domains like manufacturing, agriculture, education, and real estate. Recently, Aliu and Aigbavboa (2020) and Osunsanmi et al. (2021b) adopted SEM's concept to examine the complex relationships in the real estate-related industry.

Insert Figure 1

Figure 1 presents the research methodology framework that shows the steps in conducting the SEM. The figure showed that the first step entails breaking down the drivers or factors supporting adopting data science techniques into meaningful constructs. Exploratory factor analysis (EFA) using the varimax rotation method was used to determine the constructs supporting real estate professionals' data science applications. Jolliffe and Cadima (2016) recommended EFA as a useful tool for creating meaningful constructs from correlated variables. Statistical package for social science (SPSSV24) was used for conducting the EFA. Before the analysis, Kaiser Meyer Olkin (KMO) and Bartlett's test of sphericity was conducted to ascertain the validity of EFA data. The findings from the EFA were used to prepare a structural equation modelling. The covariance-based modelling (CB-SEM) structural equation modelling was an approach. According to Oke and Ogunsemi (2016) and Aliu and Aigbavboa (2020), CB-SEM is suitable for social science-based research, non-experimental research, and avoidance of excessive multicollinearity. Based on this reason, CB-SEM was deemed ideal for this study (social science-based research) that also aimed at avoiding multicollinearity for the model that has the potential of effective application of data science technique

Figure 1 showed that the next step was focused on conducting a confirmatory factor analysis (CFA). The CFA was performed to ensure that each construct's variables were accurate and adequate. Amos Graphics version 27 was used for the CFA and tests like discriminant validity, composite reliability (CR), and average variance explained (AVE). The confirmatory factor analysis was done following the principles of Schreiber et al. (2006) and Hair et al. (2017). The scholars recommended that each variable's loadings that make up a component should be between 0.70 and above in arriving at the best fit model. Also, the discriminant validity using the Fornell-Larcker criterion should be valid with a CR value between 0.70 to 0.95, and AVE should be above 0.5. The subsequent step was focused on assessing the structural model assessment as depicted in Figure 1. This study hypothesized that there is a significant impact of the drivers supporting the adoption of data science techniques on its successful application in the real estate field. The boot strapping in AMOS function was used in generating the outcome.

4. Results and Discussion

4.1 Exploratory factor analysis

This section presents the data analysis results related to the drivers of data science techniques for real estate professionals in the fourth industrial revolution era. The first step in the data analysis entails calculating the exploratory factor analysis. The exploratory factor analysis was conducted using principal component analysis (PCA). The PCA was performed to break the drivers of data science techniques for real estate professionals into reasonable dimensions. Osunsanmi et al. (2020a) and Jolliffe and Cadima (2016) also adopted the PCA for breaking variables extracted from literature into numerous dimensions. A total of twenty (20) variables were extracted from literatures regarding the drivers of data science techniques for real estate professionals.

The extracted variables were subjected to a PCA analysis using the varimax rotation method. The analysis was broken down into two stages, with the first stage focused on determining the validity of the data for PCA. The second stage involves using the varimax rotation method for breaking the variables into different dimensions. The validity of the data for PCA was determined using the Kaiser Meyer Olkin (KMO) and Bartlett's test of sphericity. The KMO provided a value of 0.623 which is above the stipulated threshold of 0.5, as recommended by (Ong and Puteh, 2017). The Bartlett test of sphericity was significant at 99% and 95% confidence levels. The test provided a chi-square value of 3435.180 with a degree of freedom of 1378. This confirms the data suitable for PCA according to Aliu and Aigbavboa (2020) recommendation and Ong and Puteh (2017).

Table 5 presents the rotated component matrix analysis showing the loadings and dimensions of the variables extracted from the literature. Before conducting the rotated component matrix, the Kaiser criterion analysis was conducted to determine the components extracted from the variable. According to Larsen and Warne (2010), the Kaiser criterion assists in identifying the extracted components or factors based on the eigenvalues. The analysis revealed that only three factors or components have eigenvalues above 1 and were thus submitted for the rotated component analysis using the varimax method. Table 5 shows the different loading for the variables that makes up the three-component including their variance. The variables in each component were arranged based on their loadings in a descending manner. The components were named following the variables, and the variables with the highest loading were given preferential consideration as recommended by Osunsanmi et al. (2020a) and Aghimien et al. (2021).

Insert table 5 here

Component 1: Decentralised real estate market

The first component in the exploratory factor analysis provided a variance of 34.2%. This implies that the component is responsible for 34% change in the drivers supporting data science techniques for real estate professionals. The component comprises nine variables, the topmost being the use of blockchain for information sharing, virtual real estate market, adoption of decentralised currency, autonomy in decision making, and utilisation of proptech for real estate market. It can be deduced from Table 6 that the variables making up the first components are closely related, with loading between 0.853 to 0.548. As Aghimien et al. (2021) recommended, the name given to a component depends on the variables with the highest loading within the component. Therefore, the first component was labelled the "Decentralised real estate market" due to the component's variables.

The decentralised real estate market is described as the utilisation of the 4IR technologies and principles like blockchain, virtual currency and others capable of supporting the direct exchange of information between real estate stakeholders. Kalyuzhnova (2018) also discovered that blockchain is one of the significant technologies supporting decentralised real estate. Likewise, Nieradka (2019) found that virtual reality and currency support real estate activities' decentralisation. Poursaeed et al. (2018) and DeLisle et al. (2020) provided the decentralised real estate market benefits. The authors submitted that a decentralised market supports transparency, eliminate intermediaries, and lowers transaction cost. This study contributed to the existing literature as it revealed that a decentralised real estate market will contribute to the utilisation of data science techniques for real estate professionals.

Component 2: Data management system

The second component accounted for a variance of 22.3%, as shown in Table 6. This implies that the second component is responsible for the 22% change or determinants in ensuring data science techniques for real estate professionals. The component is made up of six variables, and the variables with high loadings are data synchronisation system, data accessibility, data structure and data quality. Table 6 shows that the second component variables are related to loading between 0.781 and 0.510. However, two variables with the highest loadings are data synchronisation and accessibility. Therefore, this component was called "data management and was identified as the crucial factor that will foster the adoption of data science techniques for real estate professionals.

Lee and Park (2020) discovered that effective data management is crucial for adopting modern data analysis techniques for real estate professionals. Poursaeed et al. (2018) and Gromova and Pupentsova (2020) discovered that the data management technique adopted by a real estate professional is vital in the performance of modern data analysis like ANN, CNN and others. Furthermore, Kasprzak (2021) affirmed that effective data management is essential for the transformation of real estate practice in the fourth industrial revolution era. The author believed that data management practices like data structuring, quality, and accessibility are needed for transforming the practice of real estate professionals according to the 4IR. Likewise, Sanquist et al. (2018) discovered that data management functions as an effective tool for transforming corporate real estate in the 4IR era

Component 3: Inclusive educational system

The third component yielded a variance of 11.7%, as shown in Table 6, after conducting a varimax rotation method. It can be deduced that the third component accounted for approximately 12% change in real estate professionals' drivers of data science techniques. The third component comprises variables like the introduction of modern valuation techniques in the educational system, contrasting educational curriculum, transdisciplinary skills in the educational system and inviting guest speakers from the industry. The third component had loadings between 0.822 and 0.510 with high correlations. Based on the variables that make up the third component with preferential given to the variables with the highest loading, it was labelled the "inclusive educational system".

Buglea and Şipoş (2015) discovered that an inclusive education system is needed to re-orientate the real estate business environment in tandem with global practices. The findings from this study showed that higher education or university should establish an inclusive educational system that supports the application of modern data analysis techniques in making decisions for real estate. In comparison to the construction industry, Aliu and Aigbavboa (2020) recognised an inclusive educational system as a critical requirement for guaranteeing the employability of construction graduates. The scholar describes an inclusive educational system as collaboration between industry and the university. An inclusive educational system was described in this study as the introduction of modern data analysis principles and techniques into the real estate curriculum.

4.2 Structural equation modelling

The second objective was focused on determining the relationship or association between the drivers of data science techniques in real estate and their impact on enhancing data analysis. In meeting, this objective structural equation modelling was used in assessing the relationship, as shown in Figure 2. The findings from the exploratory factor analysis revealed that data management, decentralised real estate market, and inclusive educational system are the significant drivers of data science in the real estate industry. Figure 2 proposes a conceptual framework that hypothesise a substantial relationship between the drivers of data science in real estate in ensuring effective data analysis. According to the recommendation of Hair et al. (2010) and Mustafa et al. (2020), the confirmatory factor analysis is the first step in conducting SEM using AMOS.

Insert Figure 2 here

4.2.1 Confirmatory factor analysis

After conducting the exploratory factor analysis (EFA) to determine the drivers of data science principles for real estate professionals, a confirmatory factor analysis (CFA). The CFA was conducted using AMOS to test whether measures of the constructs are consistent with the nature of the constructs. This study adopted AMOS because it can determine the goodness of fit between data and hypothesized models. The requirement and recommendations of Oke and Ogunsemi (2016) and Mustafa et al. (2020) were used in developing the model. The authors recommended that the loadings for each variable should be above 0.6 and the average variance explained above 0.5.

Insert Table 6 here

Table 6 presents the findings from the confirmatory analysis showing the loading for each variable that makes up the constructs, including the composite reliability (CR) and average variance extracted (AVE). Towards achieving the best fit model, the analysis was conducted three times, and some variables that made up each construct were eliminated. For instance, variables that comprise the enhanced data analysis (EDA3, EDA6, EDA8, and EDA9) were eliminated from the model as they provided a loading below 0.6. Whereas variables like (EDA1= reliability, EDA7= enhance urban planning, EDA2 = accuracy, EDA4 = predictive capacity and EDA5 = inform planning policies) were retained. The retained variables represent the benefits of adopting a modern data analysis for real estate projects. The findings were also in tandem with the works of (Lee and Park, 2020, Lin and Mohan, 2011, Poursaeed et al., 2018), which discovered that modern data analysis enhanced the predictive capacity and accuracy of real estate value. It can be deduced from the confirmatory factor analysis that if the drivers of data science principles are implemented, it will ensure accurate prediction of real estate value, enhance urban planning and other positive benefits.

Insert table 7 here

Table 7 presents the discriminant validity analysis of the constructs. The test was conducted to determine whether measures of the constructs are consistent with the nature of the constructs. The discriminant validity test was conducted following the Fornell-Larcker criterion. Hair et al. (2019) stipulated that the discriminant test is deemed valid when all other correlations within the table are smaller when compared to the topmost correlation on each diagonal. On the other hand, Hair et al. (2010) declared that the discriminant test is valid if the square root of AVE for each construct is greater than the inter-construct co-relations. The topmost correlation in table 8 was bolded and arrived at by determining the square root of the AVE for each construct. A critical look at Table 8 revealed that the bolded topmost correlation is higher than other correlations. Thus, it can be deduced that the variables used to measure this study's four constructs are sufficient and discriminately valid.

Insert Figure 3 here

4.2.2 Structural model assessment

Towards assessing the structural model, numerous fit indices were examined as recommended by Oke and Ogunsemi (2016), Ong and Puteh (2017), and Aliu and Aigbavboa (2020). The fit indices are standardized root mean square (SRMR), comparative fit index (CFI), Tucker Lewis index (TLI), and other indices, as presented in Table 8. The table also shows the recommended threshold for each fit indices used to determine the validity of the model. A cursory look at table 8 revealed that all the fit indices are within the recommended threshold. According to Hair et al. (2010) and Rosenbaum and Spears (2009), the SRMR must be between 0.05 to 0.07. The value of SRMR in Table 8 confirmed that the model is fit for determining the drivers of data science in real estate. Further confirmation was done by checking other fit indices like TLI, CFI, RMSEA, NFI and IFI. As shown in Table 8, the value from the indices revealed that the model's structural component is valid.

Insert Table 8 here

4.2.3 Structural relationships among the constructs

This section assesses the model's structural relationship. This was done to test the hypothesis that was postulated in this study. This study postulated that data management would have a significant impact on ensuring effective data analysis. The same hypothesis was postulated for inclusive educational system and decentralised real estate market in ensuring effective data analysis. Figure

4 presents the graphic illustration of the hypothesis tested in this study. According to Mustafa et al. (2020), in AMOS the arrows represents the path coefficients between two constructs depicting their relationships. However, the construct with the arrows pointing towards it is referred to as the endogenous, explained or dependent variables. Thus, it can be deduced from figure 4 that the dependent variable is effective data analysis, and the independent variables are the three drivers of data science techniques (data management, inclusive educational system and decentralised real estate markets) for real estate professionals.

Insert Figure 4 here

The standardized values for the path co-efficient for each construct are presented in Figure 4. The values also represent their estimates, as shown in Table 9. It can be deduced from Table 9 and Figure 4 that there is a significant positive effect of the decentralised real estate market on ensuring effective data analysis in the 4IR era. The deduction was arrived at as the critical ratio (CR) value of the decentralised real estate market greater than 1.96. This study identified the decentralised real estate market as a forum that supports the direct transaction between sellers and buyers of real estate interest. The concept of a decentralised market became popular since the advent or discovery of blockchain technology. Wouda and Opdenakker (2019) and Kalyuzhnova (2018) affirmed that blockchain is the primary technology with the potential for decentralising the real estate market. The decentralisation of the real estate market as proposed in this study is based on three dimensions: financial, communication and value distribution. Communication decentralisation, also known as decentralised information management, is how all stakeholders involved in the real estate market have unrestricted communication access.

Insert Table 9 here

Table 9 shows that the critical ratio of data management yielded a value of 10.703 greater than 1.96. This confirms a significant impact of data management in ensuring effective data analysis in the 4IR era. It can also be inferred that proper data management will enhance data science practice by 11%. Kasprzak (2021) and DeLisle et al. (2020) discovered similar findings after recognising data management as the key to accurate data analysis for real estate professionals. This study describes data management as the ability to structure, collect, keep, and utilise data to ensure accurate prediction of real estate investment. Table 9 also confirms that an inclusive educational system significantly impacts applying data science principles for real estate practice. An inclusive educational system yielded a critical ratio value of 13.748. Buglea and Şipoş (2015) and Zhang and Gao (2019) were the significant supporters clamouring for inclusive educational systems for real estate professionals. They believed that an inclusive educational system supports introducing and implementing principles originating from other professions for real estate application. Figure 4 revealed that the drivers of data science techniques would ensure effective data analysis. Whereas the enhanced data analysis is measured by five variables (EDA1= reliability, EDA7= improve urban planning, EDA2 = accuracy, EDA4 = predictive capacity, and EDA5 = inform planning policies). Likewise, Yeh and Hsu (2018) and DeLisle et al. (2020) discovered that the shift into adopting modern analysis techniques provides reliable and high predictive results.

5. Conclusions and Recommendation

With the rise of the 4IR, the utilisation of analysis tools from other domains for analysing real estate investment moved from mission impossible or speculative to task critical. This is due to the new forms of data that emerged in the 4IR era. The 4IR supported the seamless connection between machine to machine and machine to human. As a result of the connection, large data sets or big data are created, which are often difficult to analyse using the traditional analysis methods available to real estate professionals. Data science techniques like genetic algorithm, artificial neural network, and case-based reasoning have been recommended to overcome this difficulty. This recommendation was made after discovering that the data science techniques outperformed real estate professionals' traditional analysis methods.

Despite the performance of the data science techniques method, its application in real estate is relatively scarce, especially in African countries. Most decisions made by real estate professionals in Africa are not backed by robust or sophisticated data analysis. They rely heavily on intuition and traditional analysis like multiple regression, comparison, and others. These conventional methods, especially regression analysis, have inherent shortcomings ranging from multicollinearity and difficulty in analysing unstructured data. Despite the shortcomings, the over-reliance on the traditional approach has been attributed to the absence of an enabling environment for adopting data science techniques. In creating an enabling environment, this study modelled the drivers or factors that support the application of data science techniques for real estate professionals in the 4IR era. The drivers were modelled using structural equation modelling (SEM) due to their ability to effectively show multiple relationships.

It was discovered that real estate professionals' enabling environment or drivers of data science techniques are based on a decentralised real estate market, data management, and inclusive educational system. The idea of an inclusive education system in this study is the introduction of modern data analysis principles and techniques into the real estate curriculum. To successfully adopt data science techniques, the real estate professional's curriculum should be altered. The alteration will support preparing the next generation of real estate professionals to adopt data science techniques. New practices and principles should be introduced to create an inclusive educational system, like modern valuation techniques and transdisciplinary skills. The transdisciplinary skills are described as tools and knowledge that real estate students use to adapt to data science, computing and other techniques. Also, the curriculum should include more exposure to current real estate practice through the invitation of guest speakers and mentors from the industry. Students undergoing mentoring from industry professionals tend to perform better and are exposed to modern analysis and valuation techniques.

This study discovered that another crucial driver for applying data science for real estate professionals is the decentralisation of the real estate market. The decentralised real estate market in this study is described as utilising the 4IR technologies and principles like blockchain, virtual currency and others capable of supporting the direct exchange of information between real estate stakeholders. Blockchain technology has been the primary tool for the decentralisation within real estate by creating smart contracts and digital tokens. The use of blockchain eliminates the middleman and supports the direct dealing among stakeholders instead of sharing information from a centralised system. The centralised system stifles communication and is responsible for poor data sharing among real estate professionals. Another crucial driver of data science technique for real estate professionals is a data management system. The model revealed that real estate data should be properly structured, synchronised, and accessible to all stakeholders to properly apply data science techniques.

The findings from the model were in tandem with the conceptual framework and the hypothesis that was postulated for this study. The model confirmed that a combination of the drivers would guarantee effective real estate investment data analysis. This will ensure real estate investment analysis's predictive capacity, reliability, and accuracy. The accurate prediction of prices has been the primary concern of real estate stakeholders. Thus, this study contributed to practice by providing a model or procedure that can guarantee accurate and reliable real estate investment analysis. It further contributes to practice as it allows for an enabling environment that supports the application of data science techniques for real estate professionals. It also contributes to practice by providing variables that support the application of data science techniques for professionals in the built environment. Applying data science techniques would, in return, enhance the market research of real estate professionals. The findings from the reviewed literature contribute to revealing the significant technologies of the 4IR that is responsible for the disruptive potential of real estate investors.

Finally, the activities performed by a real estate investor are susceptible to changes or mutation, thereby creating a marriage between real estate professionals and other professionals. Therefore, this study recommended that real estate professionals should not operate in silos but rather strive to combine the modern form of analysis with the traditional method. It is also recommended that adequate attention be given to real estate professionals' educational curriculum. The current curriculum should be altered to introduce transdisciplinary skills for real estate professionals. These skills will support real estate professionals in understanding the techniques, technologies, and principles from other domains. Due to the rise of the 4IR real estate professionals should concentrate on other data that will further enhance the accuracy of real estate investment. A further study can be conducted to add more variables to the model and use a larger sample.

References

Aghimien, L. M., Aigbavboa, C. O., Anumba, C. and Thwala, W. D. (2021), "A confirmatory factor analysis of the challenges of effective management of construction workforce in South Africa", *Journal of Engineering, Design and Technology,* Vol. 4 No. 2, pp. 283-295.

Ajibola, M. and Ogungbemi, A. (2011), "Importance of accessibility to reliable data for real estate practice", *Mediterranean Journal of Social Sciences,* Vol. 2 No. 2, pp. 223-223.

Akwang, N. E. and Chimah, J. N. (2021), "Research Data Management (RDM) in the Fourth Industrial Revolution (4IR) Era: The Case for Academic Libraries", *Handbook of Research on Information and Records Management in the Fourth Industrial Revolution*, IGI Global, pp. 17-37.

Aliu, J. and Aigbavboa, C. O. (2020), "Structural determinants of graduate employability: impact of university and industry collaborations", *Journal of Engineering, Design and Technology,* Vol. 19 No. 5, pp. 1080-1100.

Antoniades, A. (2021), "Commercial bank failures during the Great Recession: The real (estate) story", *ECB working paper,* Vol. 1 No. 1, pp. 1-30.

Babawale, G. (2013), "Designing appropriate valuation model for sustainable property tax system in developing countries", *International Journal of Law and Management,* Vol. 55 No. 3, pp. 226-246.

Baldominos, A., Blanco, I., Moreno, A. J., Iturrarte, R., Bernárdez, Ó. and Afonso, C. (2018), "Identifying real estate opportunities using machine learning", *Applied sciences,* Vol. 8 No. 11, p. 2321.

Barreto, L., Amaral, A. and Pereira, T. (2017), "Industry 4.0 implications in logistics: an overview", *Procedia Manufacturing,* Vol. 13 No. 1, pp. 1245-1252.

Buglea, R. and Şipoş, G. L. (2015), "Educational system orientation to requirements of real estate business environment. Accents and limits in intercultural communication", *Theoretical and Applied Economics,* Vol. 22 No. 3, pp. 35-44.

Choi, C., KIM, C. and KIM, C. (2019), "Towards sustainable environmental policy and management in the fourth industrial revolution: Evidence from big data analytics", *The Journal of Asian Finance, Economics, and Business,* Vol. 6 No. 3, pp. 185-192.

DeLisle, J. R., Never, B. and Grissom, T. V. (2020), "The big data regime shift in real estate", *Journal of Property Investment & Finance,* Vol. 38 No. 4, pp. 363-395.

Dhar, V. (2013), "Data science and prediction", *Communications of the ACM,* Vol. 56 No. 12, pp. 64-73.

Dieronitou, I. (2014), "The ontological and epistemological foundations of qualitative and quantitative approaches to research", *International journal of economics, commerce and management,* Vol. 2 No. 10, pp. 1-17.

Duan, H., Li, J., Fan, S., Lin, Z., Wu, X. and Cai, W. (2021), "Metaverse for social good: A university campus prototype", in *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 153-161.

Eberhard, B., Podio, M., Alonso, A. P., Radovica, E., Avotina, L., Peiseniece, L., Caamaño Sendon, M., Gonzales Lozano, A. and Solé-Pla, J. (2017), "Smart work: The transformation of the labour market due to the fourth industrial revolution (I4. 0)", *International Journal of Business & Economic Sciences Applied Research,* Vol. 10 No. 3, pp. 47-66.

Fragapane, G., Ivanov, D., Peron, M., Sgarbossa, F. and Strandhagen, J. O. (2020), "Increasing flexibility and productivity in Industry 4.0 production networks with autonomous mobile robots and smart intralogistics", *Annals of operations research,* Vol. 1 No. 1, pp. 1-19.

Frey, C. B. and Osborne, M. A. (2017), "The future of employment: How susceptible are jobs to computerisation?", *Technological forecasting and social change,* Vol. 114 No. 1, pp. 254-280.

Gromova, E. A. and Pupentsova, S. V. (2020), "Simulation modelling as a method of risk analysis in real estate valuation", in Vol. 898, pp. 45-52.

Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., Ip, H. F., Marioni, R. E., McIntosh, A. M. and Deary, I. J. (2019), "Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits", *Nature human behaviour,* Vol. 3 No. 5, pp. 513-525.

Guan, J., Shi, D., Zurada, J. M. and Levitan, A. S. (2014), "Analyzing massive data sets: an adaptive fuzzy neural approach for prediction, with a real estate illustration", *Journal of organizational computing and electronic commerce,* Vol. 24 No. 1, pp. 94-112.

Hair, J., Black, W., Babin, B. and Anderson, R. (2010), "Confirmatory factor analysis. Multivariate Data Analysis", *Pearson Education, Inc.: Upper Saddle River, NJ, USA,* Vol. 600, p. 638.
Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M. and Thiele, K. O. (2017), "Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods", *Journal of the Academy of Marketing Science,* Vol. 45 No. 5, pp. 616-632.
Hair, J. F., Risher, J. J., Sarstedt, M. and Ringle, C. M. (2019), "When to use and how to report the results of PLS-SEM", *European Business Review,* Vol. 33 No. 1, pp. 2-24.
Hirschi, A. (2018), "The fourth industrial revolution: Issues and implications for career research and practice", *The career development quarterly,* Vol. 66 No. 3, pp. 192-204.
Johnson, R. B. and Onwuegbuzie, A. J. (2004), "Mixed methods research: A research paradigm whose time has come", *Educational researcher,* Vol. 33 No. 7, pp. 14-26.
Jolliffe, I. T. and Cadima, J. (2016), "Principal component analysis: a review and recent developments", *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* Vol. 374 No. 2065, p. 1.
Kalyuzhnova, N. (2018), "Transformation of the real estate market on the basis of use of the blockchain technologies: Opportunities and problems", in Vol. 212, pp. 34-42.
Kasprzak, P. (2021), "The fourth industrial revolution and the real estate", *Recent trends in the real estate market and its analysis. SGH Warsaw,* Vol. 1 No. 1, pp. 1-20.
Khodabakhshian, A. and Toosi, H. (2021), "Residential Real Estate Valuation Framework Based on Life Cycle Cost by Building Information Modeling", *Journal of Architectural Engineering,* Vol. 27 No. 3, p. 04021020.
Kim, J. (2019), "Development of a real estate curriculum according to economic changes through job analysis: Focus on the real estate qualification system in South Korea", *Asia Life Sciences,* No. 1, pp. 117-124.
Kosba, A., Miller, A., Shi, E., Wen, Z. and Papamanthou, C. (2016), "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts", in *2016 IEEE symposium on security and privacy (SP)*, pp. 839-858.
Larsen, R. and Warne, R. T. (2010), "Estimating confidence intervals for eigenvalues in exploratory factor analysis", *Behavior research methods,* Vol. 42 No. 3, pp. 871-876.
Law, S., Paige, B. and Russell, C. (2019), "Take a look around: using street view and satellite images to estimate house prices", *ACM Transactions on Intelligent Systems and Technology (TIST),* Vol. 10 No. 5, pp. 1-19.
Lee, C. and Park, K.-H. (2020), "Using photographs and metadata to estimate house prices in South Korea", *Data Technologies and Applications,* Vol. 5 No. 2, pp. 280-292.
Lin, C. C. and Mohan, S. B. (2011), "Effectiveness comparison of the residential property mass appraisal methodologies in the USA", *International Journal of Housing Markets and Analysis,* Vol. 4 No. 224-243.
Maynard, A. D. (2015), "Navigating the fourth industrial revolution", *Nature nanotechnology,* Vol. 10 No. 12, pp. 1005-1006.
McCluskey, W. J., McCord, M., Davis, P., Haran, M. and McIlhatton, D. (2013), "Prediction accuracy in mass appraisal: a comparison of modern approaches", *Journal of Property Research,* Vol. 30 No. 4, pp. 239-265.
Min, J., Kim, Y., Lee, S., Jang, T.-W., Kim, I. and Song, J. (2019), "The fourth industrial revolution and its impact on occupational health and safety, worker's compensation and labor conditions", *Safety and health at work,* Vol. 10 No. 4, pp. 400-408.

Molnar, G., Savage, S. J. and Sicker, D. C. (2019), "High-speed Internet access and housing values", *Applied Economics,* Vol. 51 No. 55, pp. 5923-5936.

Morrar, R., Arman, H. and Mousa, S. (2017), "The fourth industrial revolution (Industry 4.0): A social innovation perspective", *Technology Innovation Management Review,* Vol. 7 No. 11, pp. 12-20.

Mustafa, M. B., Nordin, M. B. and Razzaq, A. B. A. (2020), "Structural Equation Modelling Using AMOS: Confirmatory Factor Analysis for Taskload of Special Education Integration Program Teachers", *Univers. J. Educ. Res,* Vol. 8 No. 1, pp. 127-133.

Nieradka, P. (2019), "Using virtual reality technologies in the real estate sector", *Annales Universitatis Mariae Curie-Skłodowska, Sectio H Oeconomia,* Vol. 53 No. 2, pp. 45-53.

Niu, J. and Niu, P. (2019), "An intelligent automatic valuation system for real estate based on machine learning", in *Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing*, pp. 1-6.

Oke, A. E. and Ogunsemi, D. R. (2016), "Structural equation modelling of construction bond administration", *Journal of Financial Management of Property and Construction,* Vol. 21 No. 3, pp. 192-211.

Ong, M. H. A. and Puteh, F. (2017), "Quantitative data analysis: Choosing between SPSS, PLS, and AMOS in social science research", *International Interdisciplinary Journal of Scientific Research,* Vol. 3 No. 1, pp. 14-25.

Osunsanmi, T., Aigbavboa, C. and Oke, A. (2018), "Construction 4.0: The Future of the Construction Industry in South Africa", *International Journal of Civil and Environmental Engineering,* Vol. 12 No. 3, pp. 206-212.

Osunsanmi, T., Aigbavboa, C. and Oke, A. (2021a), "Challenges of Conducting Market Research During Project Appraisals of Real Estate Investment", *Collaboration and Integration in Construction, Engineering, Management and Technology*, Springer, pp. 433-438.

Osunsanmi, T., Aigbavboa, C. O., Oke, A. and Onyia, M. E. (2020a), "Making a case for smart buildings in preventing corona-virus: focus on maintenance management challenges", *International Journal of Construction Management,* Vol. 2 No. 1, pp. 1-10.

Osunsanmi, T., Aigbavboa, C. O., Oke, A. E. and Liphadzi, M. (2020b), "Appraisal of stakeholders' willingness to adopt construction 4.0 technologies for construction projects", *Built Environment Project and Asset Management,* Vol. 3 No. 4, pp. 23-34.

Osunsanmi, T. O., Aigbavboa, C. O., Thwala, W. D. D. and Molusiwa, R. (2021b), "Modelling construction 4.0 as a vaccine for ensuring construction supply chain resilience amid COVID-19 pandemic", *Journal of Engineering, Design and Technology,* Vol. 5 No. 2, pp. 34-49.

Ottinger, N. B., Jordan Stein, E., Crandon, M. G. and Jain, A. (2021), "Digital twin: the Age of Aquarius in construction and real estate", *Journal of information technology,* Vol. 3 No. 2, pp. 20-34.

Peterson, S. and Flanagan, A. (2009), "Neural network hedonic pricing models in mass real estate appraisal", *Journal of real estate research,* Vol. 31 No. 2, pp. 147-164.

Peyré, G. and Cuturi, M. (2019), "Computational optimal transport: With applications to data science", *Foundations and Trends® in Machine Learning,* Vol. 11 No. 6, pp. 355-607.

Poursaeed, O., Matera, T. and Belongie, S. (2018), "Vision-based real estate price estimation", *Machine Vision and Applications,* Vol. 29 No. 4, pp. 667-676.

Provost, F. and Fawcett, T. (2013), "Data science and its relationship to big data and data-driven decision making", *Big data,* Vol. 1 No. 1, pp. 51-59.

Rosenbaum, M. S. and Spears, D. (2009), "Using group comparisons in AMOS to explore shopping as a travel driver", *International Journal of Culture, Tourism and Hospitality Research,* Vol. 3 No. 4, pp. 313-325.

Sanquist, N. J., Karpook, D. and Guelen, F. (2018), "Transforming corporate real estate for the Fourth Industrial Revolution", *Corporate Real Estate Journal,* Vol. 8 No. 2, pp. 165-179.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A. and King, J. (2006), "Reporting structural equation modeling and confirmatory factor analysis results: A review", *The Journal of educational research,* Vol. 99 No. 6, pp. 323-338.

Soman, R. K. and Whyte, J. K. (2020), "Codification challenges for data science in construction", *Journal of Construction Engineering and Management,* Vol. 146 No. 7, p. 04020072.

Tavakol, M. and Dennick, R. (2011), "Making sense of Cronbach's alpha", *International journal of medical education,* Vol. 2, p. 53.

Valier, A. (2020), "Who performs better? AVMs vs hedonic models", *Journal of Property Investment & Finance,* Vol. 38 No. 3, pp. 213-225.

Wouda, H. P. and Opdenakker, R. (2019), "Blockchain technology in commercial real estate transactions", *Journal of Property Investment & Finance,* Vol. 37 No. 6, pp. 570-579.

Yeh, I.-C. and Hsu, T.-K. (2018), "Building real estate valuation models with comparative approach through case-based reasoning", *Applied Soft Computing,* Vol. 65, pp. 260-271.

Zhang, S. and Gao, M. (2019), "Construction and Real Estate Education Curriculum Reform Based on BIM: Taking Legal System of Construction Contract Course as an Example", *ICCREM 2019: Innovative Construction Project Management and Construction Industrialization*, American Society of Civil Engineers Reston, VA, pp. 161-166.

Zurada, J., Levitan, A. and Guan, J. (2011), "A comparison of regression and artificial intelligence methods in a mass appraisal context", *Journal of Real Estate Research,* Vol. 33 No. 3, pp. 349-388.