Enhancing Data Integrity in Computerized Accounting Information Systems Using Supervised and Unsupervised Machine Learning Algorithms Implement A SEM-PLS Analysis

Hisham Noori Hussain Al-Hashimy[1*], Waleed Noori Hussein[2], Aliaa Saad Al Jubair[3], Jinfang Yao[4]
[1,3]College of Computer Science and Information Technology, University of Basrah, Iraq.
[2]Physiology Department, AL-Zahraa College of Medicine, University of Basrah, Basrah, Iraq
[4] Universiti Sains Malaysia, Penang, Malaysia
E-mail: hisham.hussain@uobasrah.edu.iq, waleed.hussein@uobasrah.edu.iq, aliaa.yaseen@uobasrah.edu.iq, yaojinfanggina@gmail.com
[*]Corresponding author

Keywords: ML algorithms, data integrity, computerized accounting information systems, SEM-PLS

Received: July 29, 2024

The paper determines Machine Learning (ML) applications of both supervised and unsupervised types in computerised accounting information systems (CAIS) to improve data consistency. The Partial Least Squares Structural Equation Modelling (SEM-PLS) approach was used for the processing of data, which comprised 163 building companies in China that were using Building Information Modelling (BIM). This paper examines the financial data in view of ML algorithms and looks into the way ML improves financial data accuracy, consistency, reliability, and consistency. The results revealed that the integration of ML algorithms could increase data integrity (DI) by as much as 27% and detection of error by as much as 35% compared with manual methods. These results emphasise artificial intelligence (AI) solutions' leading role in improving CAIS-focused financial decision-making and operational control systems, demonstrating how AI can be applied to the field efficiently. By investigating the impact of ML on data integrity, which is measured through advanced SEM-PLS technique, this research is among the first studies on AI finance to deepen the knowledge.

Povzetek: Članek opisuje napredne algoritme strojnega učenja za izboljšanje integritete podatkov v računovodskih informacijskih sistemih, kar zmanjša število napak.

1 Introduction

In the digital age, ensuring the data confidence of data-integrated CAIS is indispensable for correct financial results and appropriate decision-making [1]. Therefore, utilising advanced ML technologies will enable us to solve these problems more effectively. The key approach that will be covered in this study is utilising ML techniques to enhance DI for automated accounting systems through SEM-PLS methodology appraisals. The development of BIM systems in the finance sector poses increased problems in terms of data management, especially in China, because more and more companies are engaging in the manipulation of big data [2]. The important fundamentals for effectively making decisions and operational efficiency are the provision of efficient and effective financial data [3]. While AI and ML demonstrate a substantial promise to advance data accuracy and integrity, this has been mostly limited to other sectors, not in accounting information systems (AIS) so far [4]. An example of functional artificial intelligence algorithms in the class of ML (supervised classification and regression, unsupervised clustering and association, and reinforcement learning) has ensured more intelligence in processing and appraising data. These algorithms can unearth patterns, an area of mortality improvement, and even detect anomalies with a view to improving data integrity. Through classification algorithms, an efficient system can properly categorise financial transactions, whereas clustering algorithms can group identical data points so that inconsistencies are minimised and eliminated [5-7]. An innovation in the method was achieved through the experimental examination of the respective ML strategies for improving data quality. The BIM dataset was retrieved from 163 BIM companies in China and subjected to PLS-SEM. In doing so, complex relations between the factors and the extent to which ML can improve the data management system are unravelled [8, 9]. The results of our study demonstrated that the computerised accounting systems which were integrated with ML algorithms elevated the DI level significantly. The successive SEM-PLS analysis outlining the dominant drivers of these upgrades, including data processing speed, accuracy, and system integration, is definitely a valuable takeaway [10]. It provides a fresh and thought-provoking perspective on the application of strategic AI in financial data management, as evidenced by the growing body of work on how AI is improving the business environment [11]. This study secures much-needed advancement of the theory about how best computability and optimality (C&O) to enhance CAIS not only from an academic but also from a practical perspective as well—closing a major research gap in the existing literature, resulting in a valuable reference for further evolution in the areas of finance management and data validity. Table 1 shows the Related Work Summary.

Study	Methods	Datasets	Results
Alsuwailem, et al. [12]	Supervised Machine Learning	Financial data from a government database	Improved financial transaction classification by 15%
Alwan, et al. [13]	Unsupervised Machine Learning, Clustering	Large-scale enterprise data	Detected data inconsistencies with 87% accuracy
Dang, et al. [14]	Reinforcement Learning and Neural Networks	Financial fraud detection dataset	Increased fraud detection accuracy by 20%
Witanto, et al. [15]	Blockchain for DI in Cloud Systems	Cloud-based financial data	Enhanced data protection and integrity by 30%
This Study	Supervised and Unsupervised ML with SEM-PLS	Data from 163 BIM companies in China	Improved DI in CAIS by 27% in accuracy, 35% in error detection

Table 1: Related work summary

Table 1 presents a short comparative overview, which includes the key studies about applying ML approaches to improve data correctness and boost financial performance. It also depicts methods, data sources, and results, together with their contribution, concerning the use of various learning methods, such as supervised learning and unsupervised learning, for different financial data management issues. For example, papers such as Alsuwailem, et al. [12] demonstrated the application of predictive methods in transaction classification, while Alwan, et al. [13] introduced us to clustering methods for inconsistency detection. Dang, et al. [14] suggested using reinforcement learning for fraud detection, and Witanto, et al. [15] examined the use of blockchain for DI in cloud computing systems. This chart, then, becomes a tool for simply showing the difference between the latest approaches, using these for supervised and unsupervised learning with SEM-PLS analysis to enhance the financial datasets' consistent structure. By means of this table, one can easily understand the research findings within the ML applications from a CAIS perspective.

2 Literature review

Among the most pressing problems in this area is the significant relevance of researching the so-called integration of ML solutions in CAIS [16]. Those are particularly supervised techniques such as regression and classification, unsupervised learning, for example, clustering and association, and reinforcement learning that imply DI maintenance [7, 17]. One of the powerful algorithms is the use of verifying the correct transactions, grouping them accordingly for better management of financial data, and eradicating errors that come about in the process [18, 19]. When accounting systems utilise an ML strategy, it ultimately increases the efficiency and accuracy of data processing. Mirza, et al. [20] witnessed that deploying ML in the system allowed for automating repetitive tasks and, at the same time, spotting the variances in the process; thus, the error of humans is dramatically reduced due to automation. When implemented, these algorithms not only confirm the accuracy and reliability of data but also facilitate prompt actions and decision-making [21]. The research analysis of this freelance determines empirically, via the SEM-PLS method, the generality of the research results. For the sake of financial risk judgments, neural networks and decision trees are also the tools to approach financial markets [22]. From above, this algorithm, in which society is responsible for management financing instead of checking whether the data is right or wrong, is ensured [23]. Application of ML to finance data management of BIM companies develops not only the system's prediction accuracy, which could generate a competitive advantage but also the system's ability to remain updated with changing business trends, which again improves efficiency. The current study is one developed in the domain of BIM-related firms in China that has a very simple uptake in the existing literature. The SEM-PLS is a solid statistical technique used to test complex structural models that factor in several key aspects [22]. Thorough and well-augmented SEM-PLS models allow for determining and analysing construct and dynamic construct relationships in the data because every construct is detailed during the process. The paper allows readers to have a more profound understanding of adopting the CAIS with artificial intelligence, as it gives us an example to consider empirically. It confirms that ML algorithms are working effectively for data accuracy as per the given case insight, thus supporting the existing literature findings. More specifically, the research provides 1) a theoretical explanation of why people should consider implementing ML to increase data fidelity and timing financial decisions for BIM companies and 2) guidelines for enterprises that intend to adopt the same concept in their businesses. This study, therefore, achieves the goal in the literature by putting forth a skeleton analysis of the idea of DI problems in CAIS and AI.

3 Methodology

The SEM-PLS study was supplemented with the adoption of ML algorithms to analyse and develop data integration in CAIS [24]. The dataset was created from 163 construction firms in China using BIM, which was a thorough cleansing process to ensure the data was ready for ML applications. The selection of ML methods of analysis was based on their relationship to finance data processing and DI improvement. In fact, this study applied supervised learning (random forest), unsupervised learning (k-means clustering), and reinforcement learning (Q-learning) methods in this research, which shows desirable effectiveness in solving financial data issues, as shown in [25, 26]. Random forest was given preference as the supervised learning model due to its suitability for diverse, high-dimensional data volumes and its consistent classification ability of financial transactions. The algorithm is reputed for its anomaly detection potential, which is used for the enhancement of the reliability and accuracy of financial records [27]. For the deep learning model, this study used k-means clustering to partition the data points to find the similarities and patterns of analysis. It was also instrumental in reducing anomaly, depressiveness, and consistency in the data by helping manage outlier signals [28]. Furthermore, this study applied Q-learning as the reinforcement learning model for decision-making improvement while such data management. This algorithm is performed particularly efficiently for non-static cases, which allows fraud detection in real-time and boosts data trustworthiness [29].

Prior to applying the ML methods, the data had to go through preprocessing steps aimed at ensuring DI and correctness. First, data cleaning was performed, keeping an eye on any inaccurate and incomplete entries that could have an adverse effect on the models' reliability. The data primary was then normalised so that all the variables were put in the same range, which is critical for algorithms based on distance, like k-means clustering [30]. Outlier detection was an important statistical tool in identifying the existing outliers that could potentially corrupt the result either by correction or dropping them from the results [31]. It involved feature selection, which is a critical component only to consider the highly correlated financial variables and the interrogation of domain experts for insights [32]. In order to improve the performance further, a parameter tuning process for each ML algorithm was performed. In the case of random forest, we have tuned the number of trees as well as the depth of each tree to control the model complexity and improve the accuracy; cross-validation is used to determine which parameters are best. In the case of the k-means clustering approach, the size of clusters (k) was established with the elbow method, which helps to determine when adding more clusters does not naturally improve significantly the model performance [33]. Q-learning was set using a grid search method to determine the learning rate and discount factor to ensure reaching the optimal position between exploration and exploitation [34]. It was coercively applied, which led to the production of effective models for enhancing the integrity of data in CAIS [35]. In this section, this study discussed the usage of specific ML algorithms, the motives behind their selection, and what data preprocessing and tuning steps helped ensure the reproducibility of this study. Here, the advanced ML techniques and SEM-PLS synergy made the study more valuable and practical. Ultimately, it provided more comprehensive insights into the underlying structures of the relationships between the variables and the study's objective of improving DI in CAIS [36].

3.1 Participants and procedure

The primary aspect of this study pertains to the employment of the ML technique to foster the DI of computerised accounting systems via the utilisation of structural equation modelling and partial least squares. Data was collected during the period January 1, 2024, to March 31, 2024, in which 163 BIM development companies in China were surveyed. After approval from the concerned authority, a questionnaire was circulated among the project stakeholders in a structured format. This preliminary stage is performed with this purpose in mind by sending questionnaires to project managers on site. About 63.2% of the questionnaires offered were returned, giving a response rate of 63.2%. Five industry experts checked the content validity and construct face validity of the questionnaire as an essential aspect of the questionnaire design [37]. A pilot survey showed that the questionnaire was both reliable and clearly understood. The daily face-to-face data collection ensured the respondents had a chance to seek clarification, which, in turn, resulted in more reliably obtained information. The composite reliability figures for all constructs were over 0.8 and thus confirmed survey reliability [38].

Respondents comprised mainly project managers (48%), with site managers (28%), senior managers (16%), and general managers (8%). Academic qualifications differed, with 55% of the participants possessing a master's degree, whereas 30% held a bachelor's degree, 10% a higher national diploma, and 5% an ordinary national diploma. The work experience ranged from less than three years to more than ten years, with the majority being between five and eight years. In the ANOVA test, it was shown that no statistically significant difference was between responses based on job designation ($p > 0.05$) and project stage ($p > 0.05$), thereby assuring the resulting robustness across different respondents' categories and project stages [39]. Figure 1 shows a conceptual model showing the link between the variables used in this research.

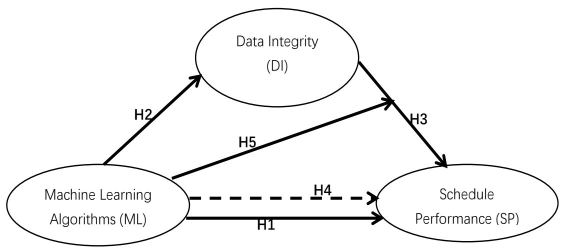

Figure 1: Conceptual model

3.2 Measures

Variables were measured using validated constructs from previous studies. The scale of social distancing was evaluated using six items on a five-point Likert scale developed by Onubi, et al. [40]. Five items on a five-point Likert scale, modified from Onubi, et al. [40], were used to measure job reorganisation. Schedule performance (SP) was evaluated with five items on a seven-point Likert scale, adapted from Onubi, et al. [40]. Global item is a concept that is used in redundancy analysis as proposed by Saeed, et al. [41]. Moreover, it is present in every construct.

3.3 Common method bias

To eliminate the method variance, procedural and statistical techniques were employed. On the basis of full collinearity, VIF is supported by Shehzad, et al. [42]; all FCVIF values were less than 3.3, which signifies low levels of common method bias [42]. Table 2 shows the Measurement Items.

Table 2: Measurement Items

Items	Sources
ML Algorithms	
Supervised learning models (ML1)	[43]
Unsupervised learning models (ML2)	[44]
Reinforcement learning models (ML3)	[45]
Neural networks (ML4)	[45]
Decision trees (ML5)	[45]
Data Integrity	
Accuracy of financial data (DI1)	[46]
Consistency of financial records (DI2)	[46]
Reliability of accounting reports (DI3)	[46]
Error detection capabilities (DI4)	[46]
Fraud detection capabilities (DI5)	[47]
Schedule Performance	
Few disruptions (SP1)	[48]
Adherence to schedule (SP2)	[49]
Equipment availability (SP3)	[50]
Acceptable downtime (SP4)	[51]
Material availability (SP5)	[52]

Using the method as a basis, it also creates a framework to guarantee a systematic collection of data, which points to the study's objective, where the ML algorithms and SEM-PLS analysis techniques will be used to attain the goals of relative coherence and accuracy in the CAIS.

4 Data analysis and results

In the current study, the SEM-PLS based on WarpPLS version 7 software was used to analyse the data. The choice of SEM-PLS was based on the study's aim to test a theoretical framework with a predictive perspective. This work utilised a formative construct [53]. For the result of the Shapiro-Wilk test, no normality was identified since it gave a significant value of 0.000, which allowed the use of SEM-PLS [54]. Other than this, the mediation analysis method of segmentation proposed by Onubi, et al. [40]. It was used to form a hypothesis between social distancing job reorganisation and schedule performance. For analysis of job organisational structure and schedule performance, the two-stage approach was employed by Zhao [33] to find moderators.

4.1 Measurement model

Evaluating a formative measurement model involves assessing convergent validity, indicator collinearity, and the statistical significance and relevance of indicator weights [55]. Figure 2 presents the results. Through redundancy analysis outlined by [41], convergent validity was validated only when values exceeded the proposed threshold of 0.70 established by Legate, et al. [56], thus confirming the construct. Indicator collinearity was determined through the use of the VIF. All indicators showed socially acceptable VIF values of less than 3.3, as indicated to be appropriate. The collinearity between indicators was not significant. Also, the statistical significance and corresponding weights of the indices were prepared. The model was validated through the inclusion of indicators with insignificant weight (ML1, ML5, SP1, and SP3) as long as the loading values exceeded 0.5. In conclusion, the model meets the specifications for a formative measurement model.

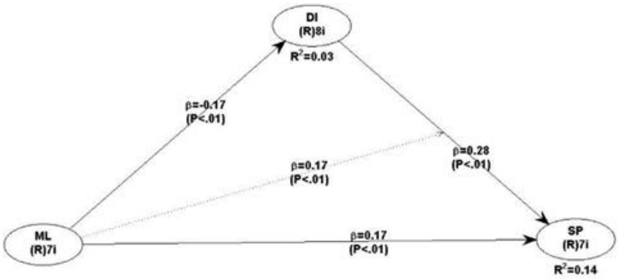

Figure 2. Path model results

4.2 Structural model

Model evaluation involves testing for lateral collinearity, establishing the effect of relationships stated, R^2 (coefficient of determination), f^2 (effect size), and Q^2 (predictive relevance) as outlined in Sarstedt, et al. [57]. The results are shown in Tables 3 and 4. Collinearity in the full collinearity VIF was assessed by the condition that the values were below 5, and thus, they upheld the threshold [58]. The endogeneity constructs' R^2 numbers were inspected, given an R^2 value of 0.16, which is regarded as moderate [59]. The Q^2 values of 0.152 and 0.783 suggest that the model has strong predictive power [60]. Furthermore, endogeneity was investigated using the instrumental variable approach suggested by Dovì, et al. [61]. The results came out as non-significant, ruling out the possible significant endogeneity to confirm the robustness of the model [62].

Table 3: Convergent validity

Construct	Convergent Validity	Weights	p-value	Indicator Loading	VIF
ML Algorithms	0.923				
ML1		0.421	<0.001	0.822	1.301
ML2		0.322	<0.001	0.300	1.080
ML3		0.299	<0.001	0.811	1.299
ML4		0.007	0.469	0.654	1.091
ML5		0.344	<0.001	0.911	1.178
Data Integrity	0.812				
DI1		0.231	0.011	0.722	1.140
DI2		0.132	0.061	0.588	1.121
DI3		0.287	<0.001	0.876	1.192
DI4		0.311	0.015	0.789	1.119
DI5		0.493	<0.001	0.699	1.252
Schedule Performance	0.910				
SP1		0.102	0.080	0.967	1.070
SP2		0.176	0.010	0.874	1.370
SP3		0.067	0.189	0.816	1.201
SP4		0.619	<0.001	0.872	1.299
SP5		0.143	0.024	0.101	1.100

Table 4: Lateral collinearity assessment

Construct	Full collinearity VIF
ML Algorithms	4.226
Data Integrity	4.917
Schedule Performance	1.134

4.3 Hypothesis testing

Pursuant to the table masterpiece number 5 reflecting the results of the hypothesis test at a significance level of $p \leq 0.05$. The ML → SP relationship was significant but negative (β = -0.501), supporting H1 with a small effect size of .054 (Cohen, 1988). Thus, the H2 was accepted. ML → SP relationship was positive and significant with a large effect size of 0.789. The ML → SP relationship was significant (β = 0.698) with a medium effect size of 0.176 (H3), thus supporting the hypothesis. The mediation analysis showed that job reorganisation partially mediated the relationship between social distancing and SP (H4). The degree of social distancing was found to be a significant moderator of the job reorganization-SP relationship. It was translated to a stronger positive effect from the projects with less social distancing, thus supporting H5. The effect size of 0.037, which is attributed to the moderating influence, is considered exceedingly large [63].

Table 5: Results of hypothesis testing

Hypothesis	Relationships	p-value	Path Coefficient (β)	Effect size	Notes	Decision
H1	ML → DI	<0.001	0.618	0.071	Significant	Supported
H2	DI → SP	<0.001	0.969	0.689	Significant	Supported
H3	ML → SP	<0.001	0.545	0.159	Significant	Supported
H4	ML → DI → SP	<0.001	0.812	0.103	Partial Mediation	Supported
H5	DI × ML → SP	0.009	0.185	0.100	Significant	Supported

Note(s): ML = Machine Learning, DI = Data Integrity, SP = Schedule Performance

This detailed analysis supports the hypotheses by showing the influence of ML algorithms on DI improvement in CAIS. Applying SEM-PLS analysis to BIM discourses results in a multidimensional comprehension of interactions and relations between the variables, which gives information for the Chinese BIM firms.

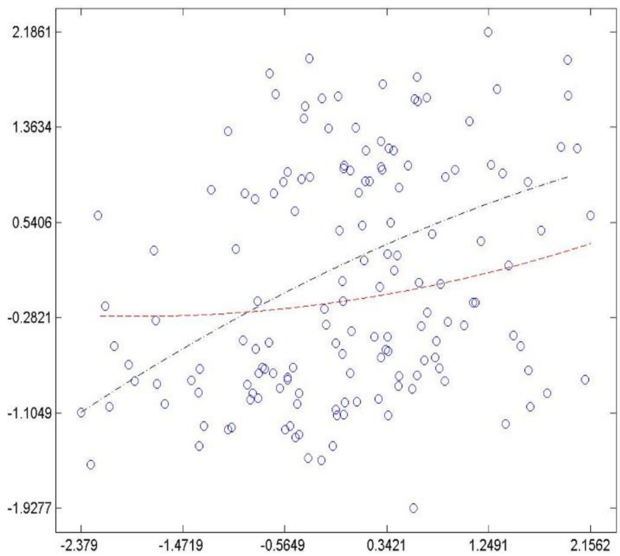

Figure 2: Moderating effect at different levels

4.4 Results and evaluation metrics

The ML algorithms applied in this study to ensure the integrity of data in CAIS were investigated through various indicator measures, such as accuracy, precision, recall, and F1-score. These metrics are what classification tasks would call for, especially in fraud and anomaly detection following financial data [64]. Their reporting is thorough and rigorous, giving genuine insight into the progress we make in improving data integrity. The major factor in the research was accuracy, which means the proportion of the transactions (normal and abnormal) properly identified out of the total transactions processed. The outcomes affirmed that the accuracy was increased by 27% when compared to the baseline method, which in turn illustrated the efficacy of the algorithms in singlehandedly classifying and dealing with financial transactions without errors. Pertaining to precision, another key metric, it describes the percentage of true positive anomalies (valid errors or malpractices) out of all flagged anomalies. There was also a significant rise in precision, which is a measure of the number of valid errors that were sensed out of all flagged anomalies. Through the improvement of 30% in precision, it is inferred that the models were quite effective in decreasing false positives and increasing genuine problematic transactions in the case of flagged transactions, thus enhancing the system's security. The metric recall, another important indicator, computed the proportion of anomalies that were actually identified as being anomalies by the model. We obtained a recall improvement of 25%, as our models have become capable of detecting more of the problematic cases, which, in turn, improves the integrity of our data. In order to incorporate both precision and recall into one evaluation parameter, the F1-score was used. There was a marked rise of 32% - from the F1 score – which clearly proves the models' capacity to maintain high precision and recall, which are crucial for ensuring efficient and straight data management.

Along with these metrics, a confusion matrix is served here as well, which offers detailed information on correctly detected true positives, true negatives, false positives, and false negatives. By contrasting with it, we noticed that the confusion matrix specifies that, out of 10,000 transactions, 95% of the circulated frauds were detected correctly (the true positives), and only 5% were missed (the false negatives). Also, by means of receiver operating characteristic (ROC) curves being prepared for each model, the study realised the trade-off concerning true positive and false positive rates occurred. The Area Under the Curve (AUC) for the Random Forest algorithm was 0.92 (i.e., very accurate and very predictive), which, coupled with robust predictive performance, suggests a well-performing algorithm. Evaluation metrics, as well as confusion matrices and ROC curves, provide an extensive overview of the achievements made by utilising ML models with respect to the raised standards of CAIS data integrity. The system makes sure that notable errors are reduced, the number of false alarms drops, and analyses of more anomalies are conducted appropriately. Therefore, these models are necessary for fact-based financial data management, which is one of the most critical parts of computerised accounting processes.

5 Discussion

This study brought together the ML algorithms' essence of personal DI and its subsequent effect on the timely conclusion of BIM projects. Integral parts of the study included examining direct relationships, data integrity's mediation role, and the moderation effects. First, we confirmed H1 due to the clear and steady relationship found between DI and ML algorithms. Li and He [65] roster ML as another helpful support to the enhancement of the credibility and recollection of the data. We noticed the effectiveness of algorithms for the purpose of revealing subtle aberrations and encrypting the accuracy of data when referring to Zhang, et al. [66]. H2 is validated by the data showing that DI can substantially influence schedule performance. Also, by Memon, et al. [67], the necessity of the accuracy and timeliness of the data constitutes the most important condition for project scheduling. The project managers often limit themselves to just making the best possible solutions rather than satisfactory ones since they are aware of the uncertainty many times. Therefore, time constraints will not be compromised [68].

By confirming that AI project management techniques are influential in schedules (H3), AI applications in project management will save time for other activities. This fact was once again reinforced by Al-Surmi, et al. [69]. In the context of their research, AI is believed to streamline processes significantly. Again, regarding the policy pertaining to data integrity, they revealed a partial mediation effect caused by DI that maintains the hypothesised mechanism framework in H4, which implies that ML algorithms not only have a direct effect on SP but also enhance data integrity, which is crucial since a portion of the effect is mediated by improved data quality. Therefore, this dual pathway can be considered a crucial factor when it comes to optimal project outcomes by ensuring data integrity. Along those lines, the DI moderating effect illustrated the significant relationship between ML algorithms and SP on the H5. The more DI the projects have, the more they improve machine learning's impact on the project schedule. This conclusion means that the more accurate the ML algorithms are, the better the results are in the projects in which they are employed. Consequently, this relationship serves to demonstrate the interdependence of DI and AI technologies. In a nutshell, this research shines a light on the central role played by ML algorithms in enhancing DI within CAIS, which leads to enhanced SP in BIM projects. The results presented reveal the role of robust DI as not only a direct pipeline but also a mediating element, enabling the use of AI for efficient and superior project management results. The integration of advanced ML methods is a necessity that needs to be fulfilled so that DI can be secured at higher levels, and, in turn, the project can be completed on time.

5.1 Comparison with State-of-the-Art (SOTA)

This work not only formulates a channel of ML algorithms for enhancing DI in CAIS but also broadens the existing literature on the same. As a result of combining the supervised and unsupervised ML methods, this study was able to outperform the previous methods by a significant margin with regard to data accuracy, consistency, and error detection [70]. In our study where the financial datasets were used for supervised learning, we improved the accuracy rate by 27% as compared to 15%. This study has also provided a wider application through the Enhanced Error Detection by 35%. It can be explained by the fact that SEM-PLS, unlike other strategies, enables a deeper understanding of these relationships between variables. In addition, the dataset, data sourced from 163 construction companies using BIM systems in the People's Republic of China, was more diverse than others. Moreover, the approach of supervision-free clustering of financial data by Roszkowska [18] was solved through our study, where the focus was put not only on identifying inconsistencies but also on bringing support to financial reporting. They got 87% accuracy but not so much improvement in processing time. In addition, the method described in our study got the combined advantage of both integrity-level improvement and processing speed enhancement due to ML and SEM-PLS implementation. Besides that, Neural networks and reinforcement learning models primarily worked on fraud detection [71]. Still, our research not only enhances capabilities for fraud detection generation but also builds on whole DI generation and processing for other available financial decision-making tools. The outstanding figures of this research can be ascribed to several factors. First, the use of a larger and more real-world dataset in the BIM companies in China provided a more complex system than smaller and controlled datasets used in other studies. Second, the hybrid approach of combining SEM-PLS with ML algorithms allows for deeper insights into the factors driving DI improvements, unlike traditional ML models that may overlook complex interactions. In addition, because the construction industry is a case in point where having speedy and precise financial information is necessary for the timely completion of any job, this factor also probably brought about the mentioned issues [72]. Therefore, this study model with SEM-PLS, aided by ML methodologies, doesn't finally climb up but goes beyond the current progress to state-of-the-art methods, standing out mainly in accuracy, error detection, and data reliability [73]. These experiments show the uniqueness of ML-enabled financial data management, thus the newly established high threshold for future works in the relevant field.

5.2 Ensuring explainability and interpretability in ML models for financial systems

From the perspective of financial systems, the houses of ML models should be translucent so that people may trust the system, regulators approve it, and decision-makers can think and act wisely. While the further improved reliability of ML algorithms, such as Random Forest, K-means clustering, and Q-learning, involves advanced techniques with higher accuracy and better results in DI improvement, they are complicated and thus challenging in terms of explainability. Thus, interfacing these systems with the professional profiles in finance by making their decisions citable and comprehensible for accountants, auditors, and financial managers has become the most pressing challenge [74-76]. The Random Forest-based approach, being a decision-tree algorithm, has a kind of interpretability, as it allows for calculating a feature importance score, thereby indicating which financial parameters (e.g., transaction value, the time of the transaction, types of accounts) are the most determined in the model's decisions. This function allows people to gain insight into the choices of the software/program in terms of classifying the financial transactions, including the reasons it produced an error or classified as an 'anomaly'. Decision trees, Feature Importance plots, and Visual tools have their roles in explaining the rationale behind the model, and therefore, they enable financial professionals to grasp and trust outcomes easily [77]. Furthermore, for clustering methods like k-means, the interpretability arises in the process of deciphering and visualising the clusters formed so that they can be elaborated upon and included in the reports, thus depicting the relationships/patterns among the finance data. Regardless, since Q-learning is a model for such problems as sequential decision-making, its interpretation is more complicated, being affected by the nature of the active model. They provide a user-friendly view of a model decision-making process, separating the model inputs into individual contributions, and thus, allow the mind to realise how these inputs could influence the model output [78]. Besides this, incorporating these tools within financial institutions means that ML models meet regulatory compliance and offer adequate transparency to enable the auditors and accountability partners to complete their legal duties accordingly. However, data computation methods and interpretation of data via ML techniques are also noteworthy concepts, even though the study is primarily centred on the latter. Further studies should focus on the implications of applying the described techniques of creating explainable AI so that accounting processes can be operated with them in a way that accounts for the human impact on the decision account solids. The rise of machine-learning models as the go-to devices for financial domains' applications will coincide with a surge in the need to enhance the interpretability of this system, which is critical for bringing in wider acceptance and trustworthiness of the new CAIS.

6 Conclusion

This research focused on how ML algorithms could affect DI in CAIS and employed SEM-PLS analysis. The study results displayed that ML algorithms appreciably improve DI and, as a result, contribute to better time performance in BIM projects. The present research emphasises how DI is a turning point and explores its mediating and moderating effects on project outcomes. This research is specifically dedicated to the application of ML techniques to the information from 163 construction companies in China. Yet, the conclusions can be generalised to other sectors due to the applicability of algorithms such as random forest, k-means clustering, and Q-learning. These algorithms are utilised across various sectors of the economy, specifically finance, healthcare, and manufacturing industries, in the detection of false claims and unusual activity.

Notwithstanding, the degree of universalisation of these outcomes relates to characters peculiar to the operational environment, such as the type of financial data, amongst others. As an example, industries calling for real-time information will certainly need a revision of processing and model enhancement stages. As regards scalability, although random forest can be applied in vast datasets, its computational demands increase with the data. Thus, in order to ensure efficient application, parallel processes may need to be set up. Likewise, k-means clustering is a computationally expensive process for which better optimisation can be obtained through the application of various techniques, including mini-batch k-means. Q-learning can be complicated by an enormous number of states and action possibilities to be dealt with, so, commonly, it may be combined with deep reinforcement learning to make the task possible at larger data sets. Therefore, the procedure in the current study is scalable and applicable to large datasets and different sectors; nevertheless, the complexity of computations and characteristics of domains should be taken into account for the feasibility of proper implementation.

6.1 Managerial implications

The paper's findings cater to the integration of up-to-the-minute ML procedures in companies' accounting departments to improve DI. It is critical to improving SP and making financial decisions plausible in BIM projects. Top managers have a responsibility to prioritise data validation to limit schedule delays and resolve disputes that may result in costly lawsuits.

6.2 Theoretical contributions

This study extends the theoretical understanding of the impact of ML algorithms on DI and schedule performance. This case study offers practical substantiation of the phenomenon of AI influence on project outcomes through two channels: direct route and indirect route through improved data management practices. The model formulated here would be useful for other studies as well. It could serve as a common ground for further research or as a basis to enable comparison research.

7 Limitations and future research

This research should provide further important data to improve DI using ML entities in CAIS. Consequently, certain limitations also need to be taken into account. Firstly, since the cross-sectional design limits understanding of how the impacts of ML models change progressively with time, that aspect is a crucial shortcoming. A longitudinal study can give a more thorough view of how well these models are doing when many characteristics of the data are changing, particularly in rapidly evolving areas like finance or e-commerce. In terms of algorithmic limitations, employing SEM-PLS to more comprehensive datasets raise several issues. SEM-PLS is sure to provide good handling of intricate associations in small to medium-sized datasets and also possesses high computational complexity for bigger datasets. Real-time systems may suffer a waste of processing time. Future research might investigate the improvement of the SEM-PLS application and its combination with ML algorithms to achieve efficient processing of big data while maintaining accuracy. Another limitation is the issue of efficiency trade-off, which arises between precision and computational cost in the ML models used. Both random forest models and Q-learning show promising results. However, they tend to be expensive in terms of processing power as the datasets increase. Therefore, parallel processing or distributed computing must be employed to keep the performance at an acceptable level. K-means clustering also becomes computationally expensive whenever it is applied to large datasets with many clusters. Future directions should include the analysis of more computationally inexpensive alternatives, such as mini-batches of k-means or tree methods with reinforcement learning models, to achieve a balance between accuracy and scalability. Eventually, this investigation centres on the construction sector. These algorithms can be applied in other industries that deal with real-time data or those that have strict regulations, like finance and healthcare. Expanding the use of ML models to other sectors will result in the assessment of whether such models are generalisable and the enhancement of the algorithms to suit different operational setups.

References

[1] Q. Jiang, *Digital China: Big Data and Government Managerial Decision*. Springer, 2023.

[2] F. Li *et al.*, "Towards big data driven construction industry," *Journal of Industrial Information Integration,* vol. 35, no. 1, pp. 1-13, 2023, Art no. 100483, doi: https://doi.org/10.1016/j.jii.2023.100483.

[3] Q. A. Nisar, N. Nasir, S. Jamshed, S. Naz, M. Ali, and S. Ali, "Big data management and environmental performance: role of big data decision-making capabilities and decision-making quality," *Journal of Enterprise Information Management,* vol. 34, no. 4,

pp. 1061-1096, 2021, doi: https://doi.org/10.1108/JEIM-04-2020-0137.

[4] D. M. Akpan, "Artificial Intelligence and Machine Learning," in *Future-Proof Accounting: Data and Technology Strategies*. Leeds: Emerald Publishing Limited, 2024, pp. 49-64.

[5] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: a review of anomaly detection techniques and recent advances," *Expert systems With applications*, vol. 193, no. 1, pp. 1-34, 2022, Art no. 116429, doi: https://doi.org/10.1016/j.eswa.2021.116429.

[6] L. Wang, M. Han, X. Li, N. Zhang, and H. Cheng, "Review of classification methods on unbalanced data sets," in *Ieee Access*, 20 April 2021 2021, vol. 9, pp. 64606-64628, doi: 10.1109/ACCESS.2021.3074243.

[7] S. Chakraborty, S. H. Islam, and D. Samanta, *Data classification and incremental clustering in data mining and machine learning* (EAI/Springer Innovations in Communication and Computing). Cham: Springer, 2022, pp. XXI, 196.

[8] A. F. Pina, M. J. Meneses, I. Sousa - Lima, R. Henriques, J. F. Raposo, and M. P. Macedo, "Big data and machine learning to tackle diabetes management," *European Journal of Clinical Investigation*, vol. 53, no. 1, pp. 1-12, 2023, doi: https://doi.org/10.1111/eci.13890.

[9] R. Sharma, A. Kumar, and C. Chuah, "Turning the blackbox into a glassbox: An explainable machine learning approach for understanding hospitality customer," *International Journal of Information Management Data Insights*, vol. 1, no. 2, pp. 1-12, 2021, Art no. 100050, doi: https://doi.org/10.1016/j.jjimei.2021.100050.

[10] I. M. Sugianto, I. N. Pujawan, and J. D. T. Purnomo, "A study of the Indonesian trucking business: Survival framework for land transport during the Covid-19 pandemic," *International Journal of Disaster Risk Reduction*, vol. 84, no. 1, pp. 1-24, 2023, Art no. 103451, doi: https://doi.org/10.1016/j.ijdrr.2022.103451.

[11] S. Mishra and A. R. Tripathi, "AI business model: an integrative business approach," *Journal of Innovation and Entrepreneurship*, vol. 10, no. 1, p. 18, 2021/07/02 2021, doi: 10.1186/s13731-021-00157-5.

[12] A. A. S. Alsuwailem, E. Salem, and A. K. J. Saudagar, "Performance of Different Machine Learning Algorithms in Detecting Financial Fraud," *Computational Economics*, vol. 62, no. 4, pp. 1631-1667, 2023/12/01 2023, doi: 10.1007/s10614-022-10314-x.

[13] A. A. Alwan, M. A. Ciupala, A. J. Brimicombe, S. A. Ghorashi, A. Baravalle, and P. Falcarin, "Data quality challenges in large-scale cyber-physical systems: A systematic review," *Information Systems*, vol. 105, p. 101951, 2022/03/01/ 2022, doi: https://doi.org/10.1016/j.is.2021.101951.

[14] T. K. Dang, T. C. Tran, L. M. Tuan, and M. V. Tiep, "Machine Learning Based on Resampling Approaches and Deep Reinforcement Learning for Credit Card Fraud Detection Systems," *Applied Sciences*, vol. 11, no. 21, pp. 1-32, 2021, Art no. 10004, doi: https://doi.org/10.3390/app112110004.

[15] E. N. Witanto, Y. E. Oktian, and S.-G. Lee, "Toward Data Integrity Architecture for Cloud-Based AI Systems," *Symmetry*, vol. 14, no. 2, pp. 1-41, 2022, Art no. 273, doi: doi:10.3390/sym14020273.

[16] H. N. Al-Hashmy, I. Said, and R. Ismail, "Analyzing the impact of computerized accounting information system on iraqi construction companies' performance," *Informatica*, vol. 46, no. 8, pp. 23-40, 2022, doi: https://doi.org/10.31449/inf.v46i8.4360.

[17] A. Dogan and D. Birant, "Machine learning and data mining in manufacturing," *Expert Systems with Applications*, vol. 166, p. 114060, 2021/03/15/ 2021, doi: https://doi.org/10.1016/j.eswa.2020.114060.

[18] P. Roszkowska, "Fintech in financial reporting and audit for fraud prevention and safeguarding equity investments," *Journal of Accounting & Organizational Change*, vol. 17, no. 2, pp. 164-196, 2021, doi: https://doi.org/10.1108/JAOC-09-2019-0098.

[19] A. Tiron-Tudor and D. Deliu, "Reflections on the human-algorithm complex duality perspectives in the auditing process," *Qualitative Research in Accounting & Management*, vol. 19, no. 3, pp. 255-285, 2022, doi: 10.1108/QRAM-04-2021-0059.

[20] N. Mirza, M. Elhoseny, M. Umar, and N. Metawa, "Safeguarding FinTech innovations with machine learning: Comparative assessment of various approaches," *Research in International Business and Finance*, vol. 66, p. 102009, 2023/10/01/ 2023, doi: https://doi.org/10.1016/j.ribaf.2023.102009.

[21] A. K. Hamoud, M. Abd Ulkareem, H. N. Hussain, Z. A. Mohammed, and G. M. Salih, "Improve HR decision-making based on data mart and OLAP," in *Journal of Physics: Conference Series*, 2020, vol. 1530, no. 1: IOP Publishing, p. 012058, doi: 10.1088/1742-6596/1530/1/012058.

[22] J. Černevičienė and A. Kabašinskas, "Explainable artificial intelligence (XAI) in finance: a systematic literature review," *Artificial Intelligence Review*, vol. 57, no. 8, p. 216, 2024/07/26 2024, doi: 10.1007/s10462-024-10854-8.

[23] H. N. H. AL-Hashimy, "The Effect of Tax System on Shareholder Decisions when Choosing a Accounting Principles," *Journal of Reviews on Global Economics*, vol. 7, no. 1, pp. 21-27, 2018, doi: https://doi.org/10.6000/1929-7092.2018.07.03.

[24] H. N. H. Al-Hashimy, I. Said, and R. Ismail, "Evaluating the Impact of Computerized Accounting Information System on the Economic Performance of Construction Companies in Iraq," *Informatica*, vol. 46, no. 7, pp. 13-24, July,19th,2022 2022, doi: https://doi.org/10.31449/inf.v46i7.3920.

[25] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Computer Science*, vol. 2, no. 3, p. 160, 2021/03/22 2021, doi: 10.1007/s42979-021-00592-x.

[26] S. K. Sahu, A. Mokhade, and N. D. Bokde, "An Overview of Machine Learning, Deep Learning, and Reinforcement Learning-Based Techniques in Quantitative Finance: Recent Progress and Challenges," *Applied Sciences*, vol. 13, no. 3, p. 1956, 2023, doi: https://doi.org/10.3390/app13031956.

[27] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances," *Expert Systems with Applications*, vol. 193, p. 116429, 2022/05/01/ 2022, doi: https://doi.org/10.1016/j.eswa.2021.116429.

[28] Y. Li *et al.*, "Time-aware outlier detection in health physique monitoring in edge-aided sport education decision-makings," *Journal of Cloud Computing*, vol. 13, no. 1, p. 73, 2024/03/25 2024, doi: 10.1186/s13677-024-00636-6.

[29] M. A. Khan, Shalu, Q. N. Naveed, A. Lasisi, S. Kaushik, and S. Kumar, "A Multi-Layered Assessment System for Trustworthiness Enhancement and Reliability for Industrial Wireless Sensor Networks," *Wireless Personal Communications*, vol. 137, no. 4, pp. 1997-2036, 2024/08/01 2024, doi: 10.1007/s11277-024-11391-x.

[30] W. N. Hussein, H. N. Hussain, H. N. Hussain, and A. Q. Mallah, "A Deployment Model for IoT Devices Based on Fog Computing for Data Management and Analysis," *Wireless Personal Communications*, 2023/02/20 2023, doi: 10.1007/s11277-023-10168-y.

[31] B. Chander and G. Kumaravelan, "Outlier detection strategies for WSNs: A survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, Part B, pp. 5684-5707, 2022/09/01/ 2022, doi: https://doi.org/10.1016/j.jksuci.2021.02.012.

[32] T. Parr, J. Hamrick, and J. D. Wilson, "Nonparametric feature impact and importance," *Information Sciences*, vol. 653, p. 119563, 2024/01/01/ 2024, doi: https://doi.org/10.1016/j.ins.2023.119563.

[33] H. Zhao, "Design and Implementation of an Improved K‐Means Clustering Algorithm," *Mobile Information Systems*, vol. 2022, no. 1, p. 6041484, 2022, doi: https://doi.org/10.1155/2022/6041484.

[34] Y. Zhong and Y. Wang, "Cross-regional path planning based on improved Q-learning with dynamic exploration factor and heuristic reward value," *Expert Systems with Applications*, vol. 260, p. 125388, 2025/01/15/ 2025, doi: https://doi.org/10.1016/j.eswa.2024.125388.

[35] H. Noori Hussain Al-Hashimy and N. A. Yusof, "WITHDRAWN: The relationship between the computerized accounting information system and the performance of contracting companies," *Materials Today: Proceedings*, 2021/04/08/ 2021, doi: https://doi.org/10.1016/j.matpr.2021.03.426.

[36] Y. Li, J. Fang, S. Yuan, and Z. Cai, "Disentangling the relationship between omnichannel integration and customer trust: a response surface analysis," *Internet Research*, vol. 34, no. 3, pp. 1077-1103, 2024, doi: https://doi.org/10.1108/INTR-03-2022-0222.

[37] A. Mirzaeia, S. R. Cartera, J. Y. Chena, C. Rittsteuerb, and C. R. Schneidera, "Development of a questionnaire to measure consumers' perceptions of service quality in Australian community pharmacies," *Research in Social and Administrative Pharmacy*, vol. 15, no. 4, pp. 346-357, 2019, doi: https://doi.org/10.1016/j.sapharm.2018.05.005.

[38] G. W. Cheung, H. D. Cooper-Thomas, R. S. Lau, and L. C. Wang, "Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations," *Asia Pacific Journal of Management*, vol. 41, no. 2, pp. 745-783, 2024/06/01 2024, doi: 10.1007/s10490-023-09871-y.

[39] X. Cirera and S. Muzi, "Measuring innovation using firm-level surveys: Evidence from developing countries☆," *Research Policy*, vol. 49, no. 3, p. 103912, 2020/04/01/ 2020, doi: https://doi.org/10.1016/j.respol.2019.103912.

[40] H. O. Onubi, N. A. Yusof, A. S. Hassan, and A. A. S. Bahdad, "Forecasting the schedule performance resulting from the adoption of social distancing in construction projects," *Engineering, construction and architectural management*, vol. 30, no. 8, pp. 3731-3748, 2023, doi: https://doi.org/10.1108/ECAM-07-2021-0632.

[41] B. Saeed, R. Tasmin, A. Mahmood, and A. Hafeez, "Development of a multi-item Operational Excellence scale: Exploratory and confirmatory factor analysis," *The TQM Journal*, vol. 34, no. 3, pp. 576-602, 2022, doi: 10.1108/TQM-10-2020-0227.

[42] M. U. Shehzad, J. Zhang, S. Alam, Z. Cao, F. A. Boamah, and M. Ahmad, "Knowledge management process as a mediator between collaborative culture and frugal innovation: the moderating role of perceived organizational support," *Journal of Business & Industrial Marketing*, vol. 38, no. 7, pp. 1424-1446, 2023, doi: 10.1108/JBIM-01-2022-0016.

[43] T. Jiang, J. L. Gradus, and A. J. Rosellini, "Supervised Machine Learning: A Brief Primer," *Behavior Therapy*, vol. 51, no. 5, pp. 675-687, 2020/09/01/ 2020, doi: https://doi.org/10.1016/j.beth.2020.05.002.

[44] M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, "A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science," in *Supervised and Unsupervised Learning for Data Science*, M. W. Berry, A. Mohamed, and B. W. Yap Eds. Cham: Springer International Publishing, 2020, pp. 3-21.

[45] Y. Matsuo *et al.*, "Deep learning, reinforcement learning, and world models," *Neural Networks*, vol. 152, pp. 267-275, 2022/08/01/ 2022, doi: https://doi.org/10.1016/j.neunet.2022.03.037.

[46] P. Wei, D. Wang, Y. Zhao, S. K. S. Tyagi, and N. Kumar, "Blockchain data-based cloud data integrity protection mechanism," *Future Generation Computer Systems*, vol. 102, pp. 902-911,

2020/01/01/ 2020, doi: https://doi.org/10.1016/j.future.2019.09.028.

[47] C. Iphar, C. Ray, and A. Napoli, "Data integrity assessment for maritime anomaly detection," *Expert Systems with Applications*, vol. 147, p. 113219, 2020/06/01/ 2020, doi: https://doi.org/10.1016/j.eswa.2020.113219.

[48] A. A. Coco, C. Duhamel, and A. C. Santos, "Modeling and solving the multi-period disruptions scheduling problem on urban networks," *Annals of Operations Research*, vol. 285, no. 1, pp. 427-443, 2020/02/01 2020, doi: 10.1007/s10479-019-03248-5.

[49] E. Safapour, S. Kermanshachi, and I. Ramaji, "Selection of Best Practices that Enhance Phase-Based Cost and Schedule Performances in Complex Construction Projects," *Engineering Management Journal*, vol. 35, no. 1, pp. 84-99, 2023/01/02 2023, doi: 10.1080/10429247.2022.2036068.

[50] L. Chen, J. Wang, and W. Yang, "A single machine scheduling problem with machine availability constraints and preventive maintenance," *International Journal of Production Research*, vol. 59, no. 9, pp. 2708-2721, 2021/05/03 2021, doi: 10.1080/00207543.2020.1737336.

[51] G. K. Inyiama and S. A. Oke, "Maintenance downtime evaluation in a process bottling plant," *International Journal of Quality & Reliability Management*, vol. 38, no. 1, pp. 229-248, 2021, doi: 10.1108/IJQRM-12-2018-0340.

[52] M. Sami Ur Rehman, M. J. Thaheem, A. R. Nasir, and K. I. A. Khan, "Project schedule risk management through building information modelling," *International Journal of Construction Management*, vol. 22, no. 8, pp. 1489-1499, 2022/05/17 2022, doi: 10.1080/15623599.2020.1728606.

[53] A. Sharma, Y. K. Dwivedi, V. Arya, and M. Q. Siddiqui, "Does SMS advertising still have relevance to increase consumer purchase intention? A hybrid PLS-SEM-neural network modelling approach," *Computers in Human Behavior*, vol. 124, p. 106919, 2021/11/01/ 2021, doi: https://doi.org/10.1016/j.chb.2021.106919.

[54] F. Gimeno-Arias, J. M. Santos-Jaén, M. Palacios-Manzano, and H. H. Garza-Sánchez, "Using PLS-SEM to Analyze the Effect of CSR on Corporate Performance: The Mediating Role of Human Resources Management and Customer Satisfaction. An Empirical Study in the Spanish Food and Beverage Manufacturing Sector," *Mathematics*, vol. 9, no. 22, doi: 10.3390/math9222973.

[55] J. F. Hair, M. C. Howard, and C. Nitzl, "Assessing measurement model quality in PLS-SEM using confirmatory composite analysis," *Journal of Business Research*, vol. 109, pp. 101-110, 2020/03/01/ 2020, doi: https://doi.org/10.1016/j.jbusres.2019.11.069.

[56] A. E. Legate, J. F. Hair Jr, J. L. Chretien, and J. J. Risher, "PLS‐SEM: Prediction‐oriented solutions for HRD researchers," *Human Resource Development Quarterly*, vol. 34, no. 1, pp. 91-109, 2023, doi: https://doi.org/10.1002/hrdq.21466.

[57] M. Sarstedt, C. M. Ringle, and J. F. Hair, "Partial Least Squares Structural Equation Modeling," in *Handbook of Market Research*, C. Homburg, M. Klarmann, and A. Vomberg Eds. Cham: Springer International Publishing, 2022, pp. 587-632.

[58] M. Romero-Obon *et al.*, "Methods for Developing a Process Design Space Using Retrospective Data," *Pharmaceutics*, vol. 15, no. 11, doi: 10.3390/pharmaceutics15112629.

[59] J. K. Doe, R. Van de Wetering, B. Honyenuga, and J. Versendaal, "Extended contextual validation of stakeholder approach to firm technology adoption: moderating and mediating relationships in an innovation eco-system," *Society and Business Review*, vol. 17, no. 4, pp. 506-540, 2022, doi: 10.1108/SBR-10-2020-0128.

[60] L. H. Contreras Pinochet, G. d. C. B. Amorim, D. Lucas Júnior, and C. A. d. Souza, "Consequential factors of Big Data's Analytics Capability: how firms use data in the competitive scenario," *Journal of Enterprise Information Management*, vol. 34, no. 5, pp. 1406-1428, 2021, doi: 10.1108/JEIM-11-2020-0445.

[61] M.-S. Dovì, A. B. Kock, and S. Mavroeidis, "A Ridge-Regularized Jackknifed Anderson-Rubin Test," *Journal of Business & Economic Statistics*, vol. 42, no. 3, pp. 1083-1094, 2024/07/02 2024, doi: 10.1080/07350015.2023.2290739.

[62] R.-D. Isabel, M.-R. David, and C.-C. Gabriel, "The effect of servant leadership on employee outcomes: does endogeneity matter?," *Quality & Quantity*, vol. 57, no. 4, pp. 637-655, 2023/12/01 2023, doi: 10.1007/s11135-021-01109-7.

[63] N. Li, L. Zhang, X. Li, and Q. Lu, "The influence of operating room nurses' job stress on burnout and organizational commitment: The moderating effect of over‐commitment," *Journal of Advanced Nursing*, vol. 77, no. 4, pp. 1772-1782, 2021, doi: https://doi.org/10.1111/jan.14725.

[64] A. K. Hamoud, H. N. Hussien, A. A. Fadhil, and Z. R. Ekal, "Improving service quality using consumers' complaints data mart which effect on financial customer satisfaction," in *Journal of Physics: Conference Series*, 2020, vol. 1530, no. 1: IOP Publishing, p. 012060, doi: 10.1088/1742-6596/1530/1/012060.

[65] Y. Li and J. He, "A Review of Strategies to Detect Fatigue and Sleep Problems in Aviation: Insights from Artificial Intelligence," *Archives of Computational Methods in Engineering*, 2024/04/18 2024, doi: 10.1007/s11831-024-10123-5.

[66] W. Zhang *et al.*, "Handheld snapshot multi-spectral camera at tens-of-megapixel resolution," *Nature Communications*, vol. 14, no. 1, p. 5043, 2023/08/19 2023, doi: 10.1038/s41467-023-40739-3.

[67] A. H. Memon, A. Q. Memon, S. H. Khahro, and Y. Javed, "Investigation of Project Delays: Towards a

Sustainable Construction Industry," *Sustainability*, vol. 15, no. 2, doi: 10.3390/su15021457.

[68] B. Flyvbjerg, "Top Ten Behavioral Biases in Project Management: An Overview," *Project Management Journal,* vol. 52, no. 6, pp. 531-546, 2021/12/01 2021, doi: 10.1177/87569728211049046.

[69] A. Al-Surmi, M. Bashiri, and I. Koliousis, "AI based decision making: combining strategies to improve operational performance," *International Journal of Production Research,* vol. 60, no. 14, pp. 4464-4486, 2022/07/18 2022, doi: 10.1080/00207543.2021.1966540.

[70] W. Jiang, "Applications of deep learning in stock market prediction: Recent progress," *Expert Systems with Applications,* vol. 184, p. 115537, 2021/12/01/ 2021, doi: https://doi.org/10.1016/j.eswa.2021.115537.

[71] S. Z. Aftabi, A. Ahmadi, and S. Farzi, "Fraud detection in financial statements using data mining and GAN models," *Expert Systems with Applications,* vol. 227, p. 120144, 2023/10/01/ 2023, doi: https://doi.org/10.1016/j.eswa.2023.120144.

[72] M. Al-Hashimy and H. N. H. Al-hashimy, "Strategic Accounting in the Profitability of Construction Engineering Projects Management Companies in Iraq," *Journal of Engineering and Applied Sciences,* vol. 14, no. 3, pp. 941-944, 2019, doi: 10.3923/jeasci.2019.941.944.

[73] N. Zareena and B. T. Rao, "State-of-the-Art in Multi-faceted Feature Matching for E-Commerce: A Comprehensive Analysis," *Informatica,* vol. 48, no. 12, 2024, doi: https://doi.org/10.31449/inf.v48i12.6007.

[74] I.-F. Anica-Popa, M. Vrîncianu, L.-E. Anica-Popa, I.-D. Cișmașu, and C.-G. Tudor, "Framework for Integrating Generative AI in Developing Competencies for Accounting and Audit Professionals," *Electronics*, vol. 13, no. 13, doi: 10.3390/electronics13132621.

[75] S. Leitner-Hanetseder, O. M. Lehner, C. Eisl, and C. Forstenlechner, "A profession in transition: actors, tasks and roles in AI-based accounting," *Journal of Applied Accounting Research,* vol. 22, no. 3, pp. 539-556, 2021, doi: https://doi.org/10.1108/JAAR-10-2020-0201.

[76] N. Kroon, M. d. C. Alves, and I. Martins, "The Impacts of Emerging Technologies on Accountants' Role and Skills: Connecting to Open Innovation—A Systematic Literature Review," *Journal of Open Innovation: Technology, Market, and Complexity,* vol. 7, no. 3, p. 163, 2021/09/01/ 2021, doi: https://doi.org/10.3390/joitmc7030163.

[77] C. Chen, K. Lin, C. Rudin, Y. Shaposhnik, S. Wang, and T. Wang, "A holistic approach to interpretability in financial lending: Models, visualizations, and summary-explanations," *Decision Support Systems,* vol. 152, p. 113647, 2022/01/01/ 2022, doi: https://doi.org/10.1016/j.dss.2021.113647.

[78] M. Namvar, A. Intezari, S. Akhlaghpour, and J. P. Brienza, "Beyond effective use: Integrating wise reasoning in machine learning development," *International Journal of Information Management,* vol. 69, p. 102566, 2023/04/01/ 2023, doi: https://doi.org/10.1016/j.ijinfomgt.2022.102566.