Computers & Industrial Engineering 220 (2026) 112253

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

A Blockchain 4.0-Digital Twin paradigm for prediction and reasoning in distributed construction project management

Rui Zhao [a], Lingming Kong [b], Meng Sun [b], Fan Xue [b,c,*]

[a] School of Computer Science and Engineering, Macau University of Science and Technology, Taipa, Macao SAR, China
[b] Department of Real Estate and Construction, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region of China
[c] National Center of Technology Innovation for Digital Construction Hong Kong Branch, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region of China

ARTICLE INFO

Keywords:
Blockchain 4.0
Digital Twin
Construction project management
Modular construction
Machine learning

ABSTRACT

Digital twin (DT) promises advanced analysis, prediction, and reasoning to address challenges in industrial scenarios. However, existing DT studies in literature predominantly focus on creating 3D digital replicas rather than on prediction and reasoning. This paper proposes a novel Blockchain 4.0-DT (BC4DT) paradigm to leverage blockchain-based project 'transactions' with automated and interpretable machine learning to realize prediction and reasoning in distributed Construction Project Management (CPM). This paper validates the BC4DT paradigm using a pioneering cross-border modular construction project, including manufacturing, transportation, and on-site assembly phases. As a result, BC4DT achieved a satisfactory prediction of a modular construction project's total production cycle ($R^2 = 0.972$). Furthermore, industrial experts verified BC4DT's reasoning regarding interpretable, significant project barriers; thereafter, steering recommendations were made for industrial practitioners based on BC4DT's reasoning and findings. The contribution of this paper is twofold. First, BC4DT is an innovative Blockchain 4.0 paradigm that fulfills DT's promises of prediction and reasoning for industries. Secondly, the findings in the distributed CPM case support industrial practitioners in learning from pioneering projects to improve training and production and to adopt novel modular construction on a large scale.

1. Introduction

Construction project management (CPM) is inherently distributed and complex (Bryde et al., 2013). The interweaving of complex and redundant CPM information often results in various problems. Examples of such problems include conflicting interests among stakeholders (Taylan et al., 2014), diverse roles of participants (Atlam et al., 2020; Vrchota et al., 2021), uncertainty in dynamic environments (Stanitsas et al., 2021), and outdated models and schedules (Xue et al., 2018). Wong et al. (2024) pointed out that between 50 and 80 percent of all problems on construction sites are caused by delays or a lack of accurate information. As a result, concerns have been raised about the inefficiencies of CPM; this is especially worrisome given productivity on construction projects has shown minimal gains over the past 50 years (Daboun et al., 2023). Currently, construction projects are becoming increasingly complex (Wang et al., 2020), while the composition of stakeholders involved in CPM is also growing more diverse (Li et al., 2024). Current CPM faces significant challenges in coordinating management and supporting data-driven decision-making (Owusu et al., 2019).

In recent years, the Digital twin (DT) has shown significant potential in solving the problem of collaboration in CPM (Qi & Tao 2018). In the practice of modern CPM, data is often limited to isolated workflow (or paper documents) (Zhang et al., 2025). But DT can better connect the data in the whole life cycle of construction projects (Mutesi & Kyakula 2011; Zhao et al., 2023). DT also facilitates optimization decisions by simulating the behavior of real systems (Boje et al., 2020). At present, the application of DT has been extended to many fields. These areas include carbon emission prediction, equipment safety maintenance and fire risk assessment (Atlam et al., 2020).

However, existing DT research has primarily focused on digital replicas, while the analytical, predictive, and reasoning capabilities essential for CPM remain underdeveloped. Such limitations severely constrain DT's potential to address the complexities of distributed CPM

* Corresponding author at: Room 512, Department of Real Estate and Construction, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region of China.

E-mail addresses: zhaorui@must.edu.mo (R. Zhao), u3009509@connect.hku.hk (L. Kong), sunmhku01@connect.hku.hk (M. Sun), xuef@hku.hk (F. Xue).

https://doi.org/10.1016/j.cie.2026.112253
Received 4 October 2025; Received in revised form 7 July 2026; Accepted 16 July 2026
Available online 17 July 2026

(Qi & Tao 2018). Specifically, numerous technical bottlenecks persist in practice regarding the integration of cross-organizational data flows, the implementation of automated processes, and the effective handling of unstructured information (Grieves & Vickers 2017).

Currently, blockchain technology offers a solution for coordinating distributed CPM and centralized DT (Bodkhe et al., 2020). As an emerging technology, blockchain is gaining increasing adoption in the construction industry due to its characteristics of decentralization, immutability, and transparent traceability (Wu et al., 2022). By providing a decentralized, tamper-proof ledger, all participants have identical copies of the data (Wu et al., 2022; Sadeghi et al., 2023; Li et al., 2022). Currently, the leading blockchains used for CPM are Blockchain 1.0 and 2.0, as well as web-based DApps. Blockchain 1.0 introduced digital currencies like Bitcoin and emphasized the programmability of currency. Blockchain 2.0 is built on smart contracts, which allow for the automated execution of agreement terms (Li et al., 2021; Perera et al., 2020). However, Blockchain 1.0, 2.0, and web-based DApps only partially meet DT's CPM needs. Applying blockchain to CPM DT faces two key limitations: higher latency and reliance on a stable internet connection (Wu et al., 2022; Mahmudnia et al., 2022). At the same time, new Blockchain 3.0 paradigms are emerging that reduce synchronization latency and maintain synchronization capabilities in network-less environments (Zhao et al., 2023). Blockchain 3.0 primarily functions as a high-performance ledger, while addressing scalability and interoperability challenges. Blockchain 4.0 signifies the convergence of blockchain with artificial intelligence technologies, aiming to enable automated prediction and reasoning within distributed industrial environments (Soori et al., 2023). However, the existing blockchain technology has not yet solved the problem of processing and analyzing many redundant data in CPM.

Therefore, this paper aims to propose a new Blockchain 4.0-Digital Twin (BC4DT) paradigm. BC4DT can realize evidence-based analysis, prediction and reasoning, which is also the function promised by DT to distribute CPM. To improve the data analysis, prediction and reasoning ability of DT, BC4DT paradigm integrates automated and interpretable ML into the DT of CPM based on blockchain. In addition, the BC4DT paradigm has been verified by a cross-border modular construction project. The main contribution of this paper is twofold: First, the proposed BC4DT paradigm has the core capabilities of data analysis, prediction and reasoning which goes beyond the DT method commonly used in CPM. Second, BC4DT has important practical application value. CPM based on BC4DT can help construction practitioners achieve more efficient and transparent project cycle management through interpretable ML algorithms. In addition, the function has also been tested by actual cases and qualitatively explored with industry experts.

Section 2 reviews related work, setting the stage for the proposed BC4DT framework detailed in Section 3. Feasibility is then examined through a specific case study in Section 4, while Section 5 offers a critical analysis of our experimental results. Section 6 concludes the study.

2. Literature review

2.1. Centralized DT for CPM

The UK National Infrastructure Commission (2017) points out that digital twins (DT) are essentially a form of data representation. This data representation can embed real-time data and other relevant information into infrastructure management processes. Although Grieves and Vickers (2017) defined DT as a digital replica for real-time monitoring and predictive analytics. The applications of DT in the construction industry remain constrained by interoperability issues. Alsakka et al. (2024) observed that early DT primarily focused on geometric collision detection based on Building Information Modeling (BIM). Consequently, these early DT implementations failed to meet the on-site management and monitoring capabilities required by CPM. Researchers are integrating DT with Internet of Things (IoT) technology. After integration, construction projects can achieve accurate real-time tracking of materials and equipment on site (Sreedharan et al., 2025).

Despite DT's immense potential in the construction sector, several bottlenecks hinder DT's application in CPM scenarios. For instance, Dangana et al. (2022) point out that significant delays in updating field data often compromise the reliability of DT-based BIM models. Furthermore, existing commercial platforms lack unified BIM standards. The absence of standardized protocols often forces practitioners to rely on manual data entry, which is prone to errors. The resulting semantic interoperability gaps further limit the widespread applicability of DT within CPM (Kong et al., 2025). At the same time, current technologies heavily rely on cloud architecture, which places extremely high demands on network bandwidth stability within CPM (Knebel et al., 2023).

Despite researchers' attempts to address the above issues, the results have not yet met expectations. While hybrid architecture integrating blockchain and DT can enhance data integrity (Yaqoob et al., 2020), computational inefficiency and latency issues remain persistent. Furthermore, current DT applications only replicate data, failing to deliver on DT's promised data-prediction capabilities (Zhong et al., 2023). The "black box" nature of ML algorithms often fails to meet CPM requirements for project implementation transparency and result credibility (Chen et al., 2023; Chen et al., 2025).

2.2. Distributed CPM

The goal of CPM is to manage the entire life cycle of construction projects (Sears et al., 2015). Project stakeholders want to implement CPM to ensure that deliverables meet time, budget and quality parameters. As Mostofi et al. (2025) emphasized, the key determinant of the successful implementation of construction projects is the establishment of a cohesive collaborative framework.

In order to solve the long-standing problems of lack of trust and low efficiency in CPM, new technologies such as building information model (BIM), Internet of things (IOT) and blockchain are widely used. Among them, in the design phase, BIM was used to support stakeholders to flexibly adjust material plans, and spatial conflicts were effectively resolved (Lu et al., 2019). The IoT collects key environmental and equipment data through sensors, and real-time monitoring of the construction site is realized (Oesterreich & Teuteberg, 2016). In addition, studies have confirmed that the application of the IoT in material traceability and quality inspection has significantly improved the overall safety and delivery efficiency of the project (Teizer et al., 2017; Li et al., 2022; Liu et al., 2020).

Although these new technologies have been widely adopted by CPM practitioners, the development of modern CPM still faces many obstacles (Oesterreich & Teuteberg, 2016; Alshboul, 2024). Existing literature shows that in the past three decades; six key challenges have been continuously considered in relevant research (Harris et al., 2021). The first challenge is system compatibility and institutional resistance to new technologies. The second challenge is the lack of capacity of construction practitioners and the changing regulatory policies, both of which constitute major obstacles. The third challenge is the project budget constraints and the inherent complexity of the project itself, which cannot be ignored (Galloway, 2006; Azhar et al., 2012; Kerzner, 1998).

2.3. Blockchain in CPM

Traditional centralized databases inherently suffer from a single point of failure (Liu et al., 2024). Blockchain technology is dedicated to solving the single truth problem. By using the distributed ledger and password authentication mechanism, the blockchain ensures that every node in the network maintains a complete copy of data (Li et al., 2018). In addition, the chain structure formed by the hash algorithm ensures the tamper proof nature of the blockchain (Luntovskyy & Guetter, 2019). Besides, the consensus mechanism within the blockchain (such as proof of work and proof of interest) shows its application potential in engineering collaboration and resource management scenarios (Wang et al., 2019).

In the current CPM, blockchain technology is driving many changes. The smart contract of blockchain significantly optimizes the project payment process by automatically executing terms. Research shows that smart contracts effectively eliminate the time delay caused by traditional manual review (Natanelov et al., 2022). Secondly, in the field of supply chain management, blockchain can realize end-to-end tracking of building materials from production to installation (Wu et al., 2022). It is worth noting that the integration of blockchain and Internet of things has been successfully implemented in large-scale construction projects (Lu et al., 2023). In addition, blockchain shows unique value in risk management. The traceability of blockchain provides reliable data support for solving engineering disputes (Sadeghi et al., 2023).

Although blockchain technology has broad application prospects, its application in construction projects still faces multiple challenges. At the technical level, the data processing capability of the existing blockchain system is difficult to meet the real-time interaction requirements of large-scale engineering projects. Interoperability between different blockchain platforms remains unresolved (Natanelov et al., 2022). In terms of law and regulation, there are significant regional differences in the recognition of the legal effect of smart contracts (Ma, 2020). In terms of industry awareness, most construction company managers have limited understanding of blockchain technology. Differences in knowledge severely limit the adoption of technological innovation in the construction sector (Olawumi et al., 2022). In addition, the initial deployment cost of blockchain system is still a major obstacle for small and medium-sized companies to adopt blockchain system (Perera et al., 2021).

2.4. Research gap

CPM is essentially distributed. Integrating centralized DT with CPM usually raises concerns about data trust and latency. Although blockchain provides a decentralized trust solution, the existing blockchain implementation in the construction industry (blockchain 1.0-3.0) is limited to distributed information storage and heavily relies on a stable network environment. In addition, the existing blockchain DT integration solutions lack data prediction and reasoning capabilities. Therefore, there is an urgent need for a new CPM paradigm that integrates blockchain and DT. The new paradigm needs to ensure data security and realize the intelligent data analysis capability promised by DT in CPM.

3. BC4DT framework for CPM

Fig. 1 outlines the research methodology of this paper. First, the problem identification and gap analysis phase identified challenges arising from the gap between centralized DT and distributed CPM. Subsequently, the system architecture design phase focused on developing a schema for CPM transaction data, establishing the framework for the paradigm, and defining its operational workflow. The developed paradigm was implemented in a modular construction case study during the BC4DT paradigm implementation phase. Finally, the evaluation and communication phase involved a thorough performance analysis and the communication of results from applying the BC4DT paradigm.

3.1. A scheme of CPM transaction data

Table 1 presents a structured record of tasks in a construction project, with each entry containing specific details necessary for task tracking and management. The table columns include the following key components.

- ID: A unique identifier for each record.
- Task ID: The identifier for the specific task associated with the record, which helps distinguish various activities.
- User ID: The ID of the user responsible for performing the task, linking the task to a particular individual.
- Org ID: The unique identifier of the organization responsible for executing the task.

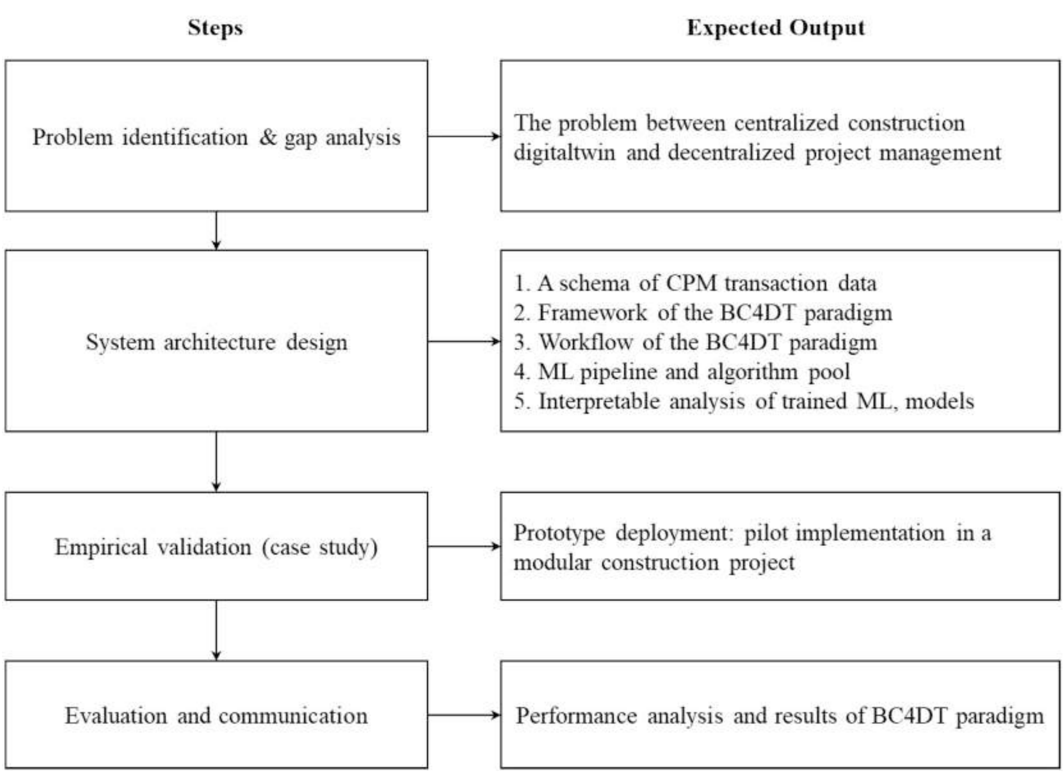

Fig. 1. Research methodology in this paper.

Table 1
Example transaction data of a construction project.

ID	Task ID	User ID	Org ID	Module ID	Timestamp	Status	General data
01	12	12	01	A-10F-01-M-M1	2021/8/2 9:47:11	1	{"inspectionTask": "...", "imgUrl": "...", "PDFUrl": "..."}
02	13	12	02	A-10F-01-M-M1	2021/8/2 9:49:33	0	{"inspectionTask": "...", "imgUrl": "...", "PDFUrl": "..."}

- Module ID: The code representing the specific module or subtask to which the task belongs, facilitating traceability.
- Timestamp: The exact time the task was recorded or updated, used for progress tracking.
- Status: The status identifier of the task; 1 represents "Active" or "Completed," 0 represents "Inactive" or "Pending".
- General Data: Contains supplementary information related to the task, such as inspection details, image links, or .pdf file links.

The general data fields primarily contain supplementary information about the project, such as the uploader's IP address, account details, signature hash, or verification information. All information in the general data fields is stored in JSON format. BC4DT employs a hybrid storage model, combining on-chain hash storage with off-chain physical storage. General data fields stored directly on the blockchain contain only a cryptographic index and an InterPlanetary File System (IPFS) content identifier. Raw sensitive data (such as IP addresses and user information) and large files (such as images and PDFs) are stored on IPFS. Data stored on IPFS is linked to the blockchain ledger via hash values.

3.2. The BC4DT framework

Fig. 2 illustrates the architecture of BC4DT. The application layer comprises two core functional modules: the Digital Twin Module, designed as a centralized digital twin unit, generates and maintains the digital mapping of the building entity; and the Application Module, built as an interactive unit for various building stakeholders, provides progress tracking, mobile and web applications to support project management needs in different scenarios.

The blockchain network support layer is composed of multiple complementary modules. Different modules jointly ensure the reliable operation and intelligent analysis ability of the system. The Cached Twin Module directly interfaces with the Digital Twin Module, which can realize the functions of fast data synchronization and offline query. The Artificial Intelligence Module is responsible for project cycle prediction, AutoML deployment, interpretable ML analysis and other tasks. The Artificial Intelligence Module also interacts with the blockchain network to realize the trusted storage of data. The core functions of the blockchain network are implemented collaboratively by multiple submodules. The User Management Module performs identity governance functions such as user registration, identity certificate issuance, and identity card management. As the core execution unit of the blockchain network, the Smart Contract Module can realize smart contract deployment, automatic transaction triggering and other operations. The Data Transaction Module provides basic functions such as consensus management, distributed ledger maintenance, peer-to-peer protocol operation, and ledger storage. The Additional Storage Module interfaces with the InterPlanetary File System (IPFS) to expand storage capacity.

3.3. The BC4DT workflow

Fig. 3 illustrates the workflow of BC4DT. The BC4DT workflow begins when users participating in CPM submit inspection data through the

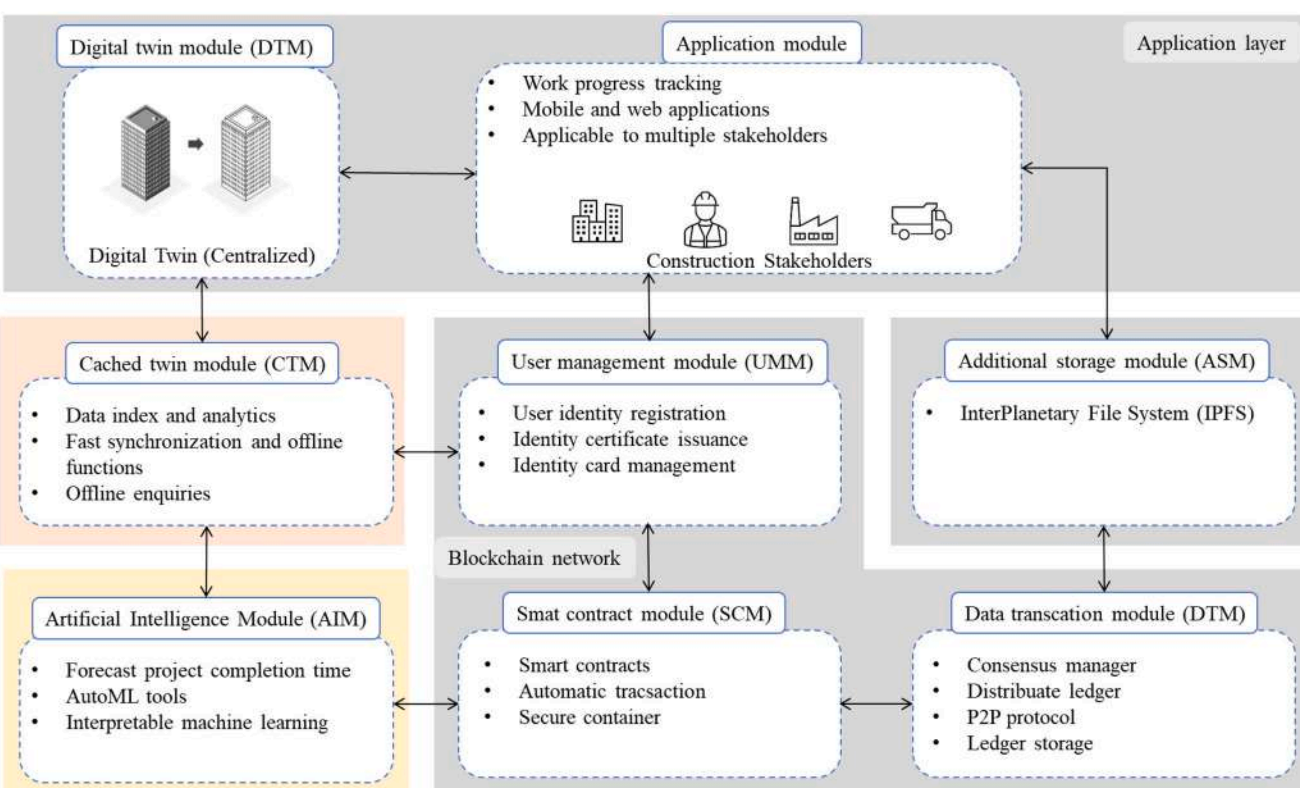

Fig. 2. The framework of the BC4DT.

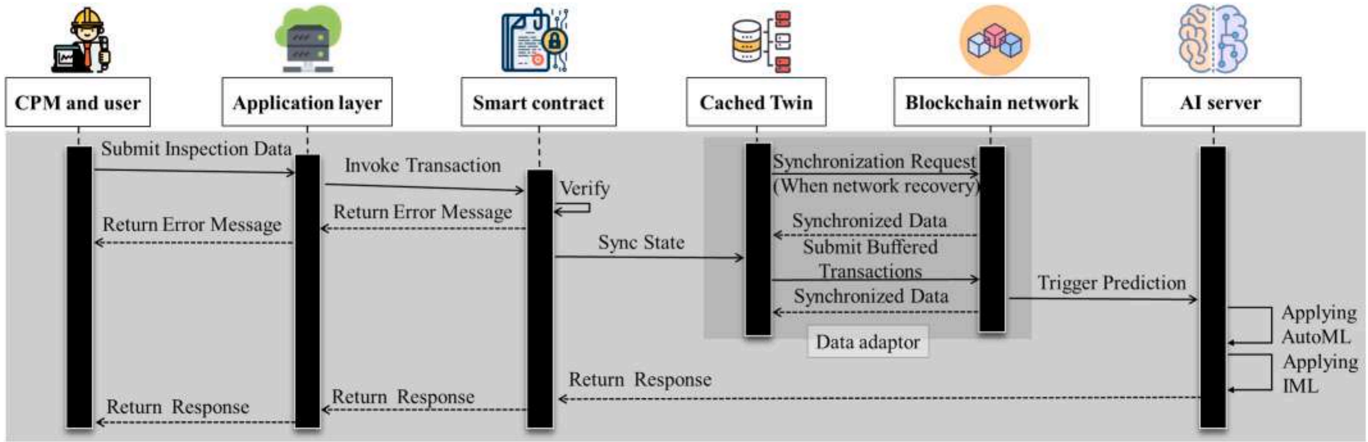

Fig. 3. The workflow of the BC4DT.

application layer, which is then validated by smart contracts. If errors are detected during validation, the system returns error messages to the users.

After verification is complete, the transaction is forwarded to the Cached Twin Module for processing. The Cached Twin Module temporarily stores data and manages synchronization with the blockchain network when connectivity is restored. Synchronized data is then sent to the blockchain network for further processing. In the BC4DT, smart contracts primarily automate the execution of tasks that meet defined conditions. For instance, if uploaded inspection data fails to meet requirements, the smart contract automatically updates the working status and notifies relevant stakeholders. Finally, automated ML and interpretable ML would analyze and process data within the blockchain network. CPM stakeholders can leverage AI services to predict project progress and analyze factors influencing construction project development.

3.4. ML pipeline and algorithm pool

3.4.1. ML pipeline for household-level prediction

Fig. 4 illustrates the BC4DT pipeline proposed in this study. Data acquisition primarily focuses on collecting data generated during the MiC module production process. The subsequent data preprocessing stage includes data visualization, missing data processing and data type standardization. Through preprocessing, the integrity and consistency of the data format are ensured. Secondly, a ML model library is established, which mainly selects the ML algorithms commonly used in the construction field. These algorithms include traditional algorithms such as linear regression, random forest, decision tree, k-nearest neighbor, and Automated machine learning (AutoML) tools such as H2O and Auto-sklearn. Then, taking Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2) and Mean Absolute Percentage Error (MAPE) as the core evaluation indicators. In addition, interpretable ML algorithm is introduced, SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) are used. Finally, the study identified and validated the most influential CPM metrics and verified the effectiveness of the solution through interviews.

3.4.2. ML algorithm pool

Table 2 lists the algorithm selection spans of four methodological categories. The first category comprises automated machine learning frameworks. Typical representatives include TPOT, H2O, and Auto-sklearn. The second category encompasses mature ensemble methods. These include random forest and gradient boosting algorithms, which represent the cutting-edge application level of machine learning in the construction cycle prediction (Datta et al., 2024). The third category comprises distance-based analysis methods, exemplified by the KNN algorithm. The fourth category encompasses classical regression models, including linear regression.

In the hyperparameter tuning phase, most construction industry practitioners lack specialized knowledge in hyperparameter optimization and selection. Furthermore, this study primarily aims to identify plug-and-play models suitable for the construction sector. Therefore, default parameter values were uniformly applied across all algorithms in this research.

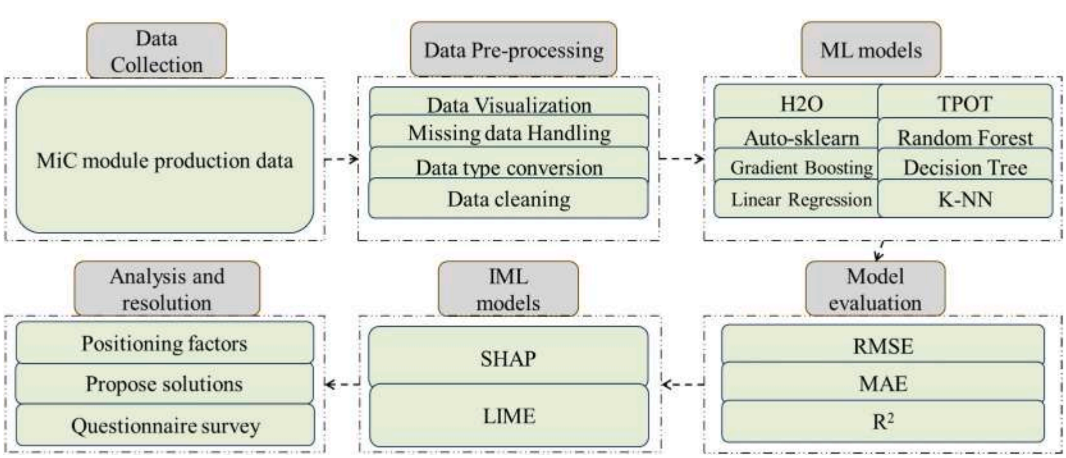

Fig. 4. The ML and interpretable ML pipeline.

Table 2
The details of the ML algorithm pool.

Category	Name	Default params	Source
Automated ML	H2O	*max_runtime_secs*=3600, *max_models*=None, *stopping_metric*='AUC'	LeDell & Poirier (2020)
	TPOT	*generations* = 5, *population_size* = 50, *random_state* = 42	Olson & Moore (2019)
	Auto-sklearn	*time_left_for_this_task*=3600, *per_run_time_limit*=300	Feurer et al. (2015)
Tree-based modeling	Random Forest	*n_estimators*=100, *max_features*=auto	Breiman (2001)
	Gradient Boosting	*n_estimators*=100, *learning_rate*=0.1, *max_depth*=6	Chen & Carlos (2016)
	Decision Tree	*max_depth*=None, *min_samples_split*=2	Zhao & Zhang (2008)
Example-based learning	KNN	*n_neighbors*=5	Zouhal & Denoeux (1998)
Regression	Linear Regression	*fit_intercept*=True, *normalize*=False	Lin & Wei (1990)

Fig. 5 details the training and evaluation pipeline of the BC4DT prediction model. The entire process can be highly summarized into three core stages. Firstly, in the unbiased preprocessing and feature engineering stages, we introduced a feature indicator strategy to address the issue of missing records after initializing the TPOT hyperparameters. For delayed variables with a value of 0, we abandoned the mean imputation method that introduces global statistical bias and instead generated corresponding binary missing value labels. This not only preserves the unobserved true state in actual business, but also fundamentally cuts off the information backflow from the test set to the training set. Secondly, in the entity-based data isolation stage, we use the *GroupShuffleSplit* tool to divide the data into 8-2 parts, and all inspection records of the same module will be packaged together, fundamentally eliminating the possibility of records from the same module appearing simultaneously in the training and testing stages. 15,318 samples were ultimately integrated into 562 modules' data. Finally, in the stage of model optimization and quantitative evaluation, the training data that has undergone strict physical isolation is sent to the automated machine learning pipeline for dynamic search and adaptation of the optimal regression model structure and hyperparameters. The final output of the optimal model will be blindly tested on an independent test set, and comprehensive quantitative validation of prediction accuracy will be completed through evaluation criteria such as MAE.

Algorithm 1 Training and Evaluation Algorithm of BC4DT

```
Require: Dataset D containing historical inspection records, including
    module identifiers (moduleID)
Ensure: Optimized regression pipeline (best_pipeline), Evaluation metrics
Initialize TPOT parameters: generations = 5, population_size = 50,
    random_state = 42
/* Feature Engineering: Handling Missing Data without Leakage */
for each feature column c containing 'Lag' in D do
    Create a new indicator column c_missing_flag
    if c(value) == 0 then
        c_missing_flag = 1
    else
        c_missing_flag = 0
    end if
end for
/* Data Splitting: Preventing Data Leakage */
Extract moduleID from D as the grouping array G
Remove moduleID column from D
Split D into training set D_train and testing set D_test with ratio 8 to 2
    using GroupShuffleSplit based on G
/* Model Training and Evaluation */
best_pipeline ← TPOTRegressor.fit(D_train.X, D_train.Y)
predictions ← best_pipeline.predict(D_test.X)
metrics ← calculate(MAE, RMSE, R^2, predictions, D_test.Y)
Return best_pipeline, metrics
```

Fig. 5. Training and evaluation algorithm of BC4DT.

3.5. Interpretable analysis of trained ML models

3.5.1. SHAP analysis

SHAP derives from cooperative game theory. SHAP computes feature importance using Shapley values, which equitably distribute predictive contributions across input variables. The framework quantifies global and local interpretability by averaging marginal feature contributions over all possible combinatorial subsets. For a given feature *i*, its SHAP value (φ) is defined as:

$$\varphi_i = \sum_{S\subseteq N\backslash\{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!}[f(S\cup\{i\})-f(S)]$$

where:

- φ_i is the Shapley value for feature *i*.
- *N* is the set of all features.
- *S* is a subset of features not including *i*.
- $f(S)$ is the model's prediction based on the subset *i*.

Fig. 6 illustrates the model prediction and SHAP explanation for an example dataset that contains diabetes information. As shown in Fig. 6 (a), the model takes the input features Age, Sex, Blood Pressure (BP), and Body Mass Index (BMI) to predict an output. Using SHAP's interpretability, the individual contributions of Age, Sex, BP, and BMI to the final prediction of 0.4 can be observed.

Fig. 6(b) details the SHAP value calculation, showing each feature's positive and negative contributions. Red indicates a positive correlation, while blue indicates a negative correlation. Age contributes +0.4, Sex contributes −0.3, BP contributes +0.1, and BMI contributes +0.1, which cumulatively explain the model's output, starting from the base rate of 0.1.

3.5.2. LIME analysis

LIME is a widely used approach for interpreting ML models and has gained significant attention recently. In comparison with SHAP, LIME provides locally interpretable explanations for the predictions of black-box models, making it valuable in the real estate industry and other domains where transparency and interpretability are critical.

Fig. 7 illustrates the LIME algorithm's sampling process and local model training. The procedure is demonstrated using a classification example with tabular data:

ML prediction results using Features 1 and 2, with classes A (green) and B (red).
Sampling points around an instance of interest (yellow cross), with weights assigned to closer points and a local decision boundary (dashed line).
Feature importance, with weights for Feature 1 and Feature 2.

4. Case study

4.1. Case introduction

The case selected for this study is a high-rise building project developed for a client in Hong Kong. The chosen architectural structure consists of a three-story podium base and two 17-story towers, as shown in Fig. 8. The core walls, elevator shafts, and other major structural components of the podium and towers are constructed using the traditional cast-in-place method. MiC is primarily used in the bedroom and

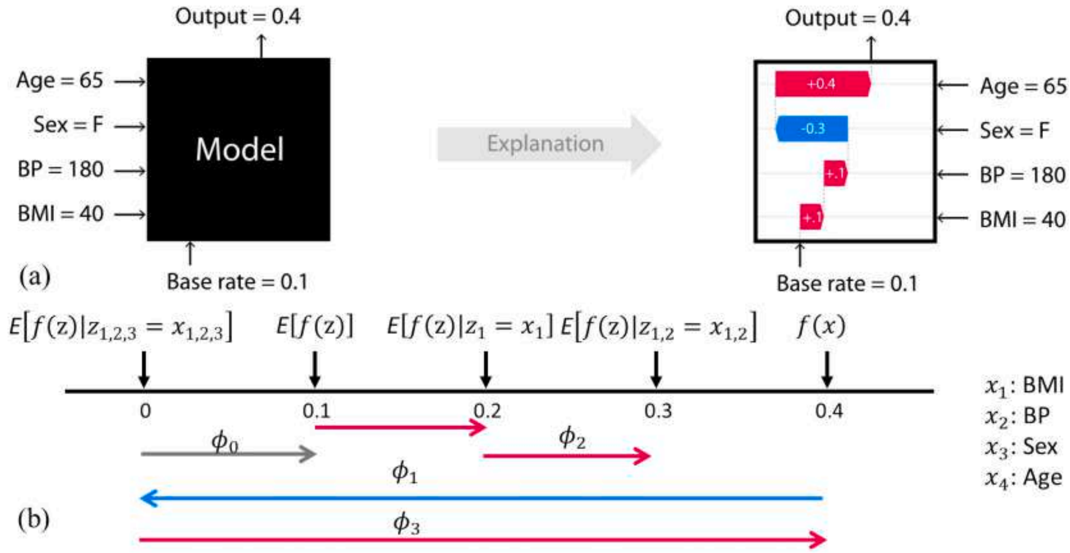

Fig. 6. Example of SHAP prediction. (a) SHAP prediction results; (b) detailed calculation of SHAP values. [Adapted from (Lundberg & Lee 2017)].

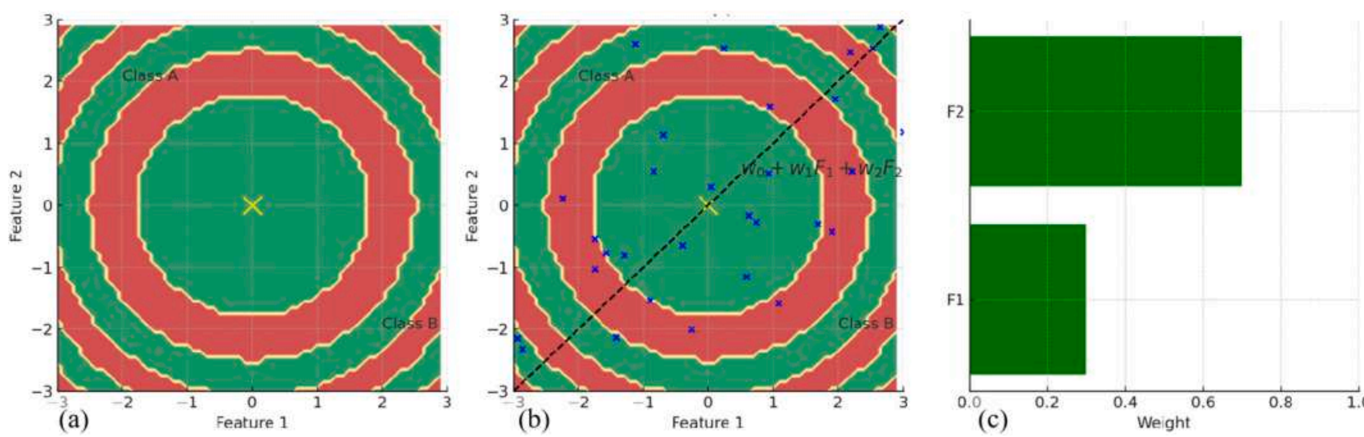

Fig. 7. LIME algorithm. (a) LIME prediction results; (b) Local linear approximation; (c) LIME feature weights [Adapted from (Ribeiro et al., 2016)].

Fig. 8. The overall case project BIM model.

bathroom areas. The MiC modules involved in the project were designed in Hong Kong, China, manufactured in factories in the Greater Bay Area of China, and finally transported to the construction site in Hong Kong, China, for final installation. The selected case study collected 23,858 inspection data points from January 11, 2022, to January 30, 2023. The collected data included inspection records from the selected case project process, such as photos, videos, and timestamps for each inspection task. All the data were converted into a structured format and stored in a Microsoft Excel file.

As shown in Fig. 9, during the on-site execution phase of the mobile terminal, user registration and authentication, project creation, and inspection task generation are all completed on the mobile device. After the construction task is completed, relevant personnel record the construction data on the mobile device. After the data is collected in the field stage, it enters the back-end blockchain processing stage. The remote inspection and digital signature can be completed across devices. Data encryption and block generation are executed on the client side and ultimately enable data storage, requests, and sharing across devices. This process realizes the seamless connection between on-site operation and trusted certificate storage through phased division of labor and equipment coordination, and ensures the full link traceability of construction data.

4.2. Data preprocessing

4.2.1. Data preparation

Table 3 presents the dataset samples collected and organized in this study. During the project, each module was assigned a unique module ID. Statistical features aim to clarify the correlations between different data points through quantitative methods. Module feature types detail

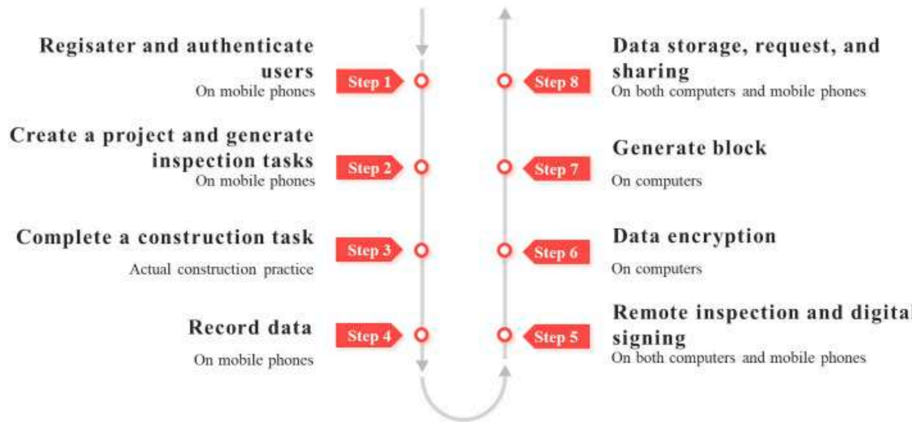

Fig. 9. Flowchart of the inspection tasks.

Table 3
Sample the top 5 rows of the collected and compiled dataset.

ModuleID	Statistical features *Hold point rate* (MC)	…	90% CI *Lag High* (BSI)	Module features *Block Type*	…	*Floor Type*	Inspection tasks *T12* … by MC	…	*T128* … by MC	*Duration* (Days)
A-10F-01-M2-M	0.43	…	19.19	A	…	hostel room	NA	…	0	129
A-10F-02-M3-M	0.55	…	23.095	A	…	toilet	1	…	NA	154
A-10F-03-M3-M	0.5	…	19.11	A	…	toilet	0	…	NA	154
A-10F-04-M2-M	0.43	…	24.76	A	…	hostel room	NA	…	0	134

information about each module, such as floor type and building type. Table 3 also records the results of specific inspection tasks. Finally, the production cycle time of the MiC module is used as a target column for training the machine learning model.

Table 4
Additional statistical features.

Feature	Description	Objective
$\bar{L}^{C}_{BSI}$	Average lag difference between the Client and MC inspection	Evaluates the average impact of Client and BSI lag differences on the module production period
δL^{C}_{BSI}	The standard deviation of lag differences between BSI's inspection and the most recent MC inspection	Evaluates the variability of lag differences between BSI and MC on the module production period
$minL^{C}_{BSI}$	Minimum lag difference between BSI and the most recent MC's inspection	Evaluates the impact of the shortest lag difference on the module production period
$maxL^{C}_{BSI}$	Maximum lag difference between BSI and the most recent MC's inspection	Evaluates the impact of the longest lag difference on the module production period
$IQRL^{MC}_{BSI}$	Interquartile range of lag differences between BSI's inspection and the most recent MC inspection	Evaluates the mid-range impact of lag differences on module production period
$90\%C.IL^{MC}_{BSI}$	90% confidence interval for lag differences between BSI's inspection and the most recent MC inspection	Evaluates the certainty range of lag difference impacts on the module production period
$F\left(\frac{N_C}{N_{MC}}\right)$	The ratio of Client to MC inspections in each module	Evaluate how the Client's participation frequency affects the module production period
$F\left(\frac{N_{BSI}}{N_{MC}}\right)$	Ratio of BSI to MC inspections in each module	Evaluates the effect of BSI's participation frequency on the module production period
α_i^{Wright}	Learning rate in each module based on Wright's Model	Evaluates MC's learning efficiency over time in the module production period
$\alpha_i^{Stanford\text{-}B}$	Learning rate in each module based on the Stanford-B Model	Evaluates MC's learning efficiency over time in the module production period

4.2.2. *Feature engineering*

Table 4 lists other statistical characteristics of the characteristic engineering phase. These characteristics are used to assess lag variance, participation frequency, and learning efficiency. The first set of characteristics focuses on lag variance. By measuring the time interval between inspections conducted by the owner or BSI and the general contractor's most recent inspection, a large lag variance may indicate delays in communication or coordination between the parties.

Another critical feature set is the participation rate ratio. This ratio measures the relative frequency of client or BSI involvement in inspections compared to the main contractor. A higher participation rate among the client or BSI may indicate faster project progress. The final set comprises learning efficiency characteristics. Derived from the Wright model and Stanford-B model, these characteristics evaluate the efficiency gains inspectors achieve over time. Such proficiency may lead to more efficient project execution and accelerated project completion.

4.3. *Implementation*

This section details the technical implementation of the proposed BC4DT paradigm. The BC4DT blockchain component's network is built upon Hyperledger Fabric v2.2. Each organization in the BC4DT blockchain network maintains a peer node. Ledger consistency is guaranteed by centralized ordering nodes using the Raft consensus algorithm. The Cached Twin Module is built by SQLite (v3.31.1).

Fig. 10 illustrates how blockchain technology is integrated into the BC4DT paradigm. In Fig. 10(a), transaction data containing construction inspection task information is constructed and submitted, and the node configuration information on the right is used to define the organizational identifier and access address of the network nodes. In Fig. 10(b) key details of the blockchain blocks (such as block hash and generation time) are displayed to verify the credible notarization of construction inspection data. And in Fig. 10(c) the network configuration file of the blockchain consortium is presented, and the organizational structure and access policies of multiple stakeholders are clearly defined. This set of interfaces and configurations is used to verify the feasibility of blockchain technology in the scenario of notarization of construction project inspection data.

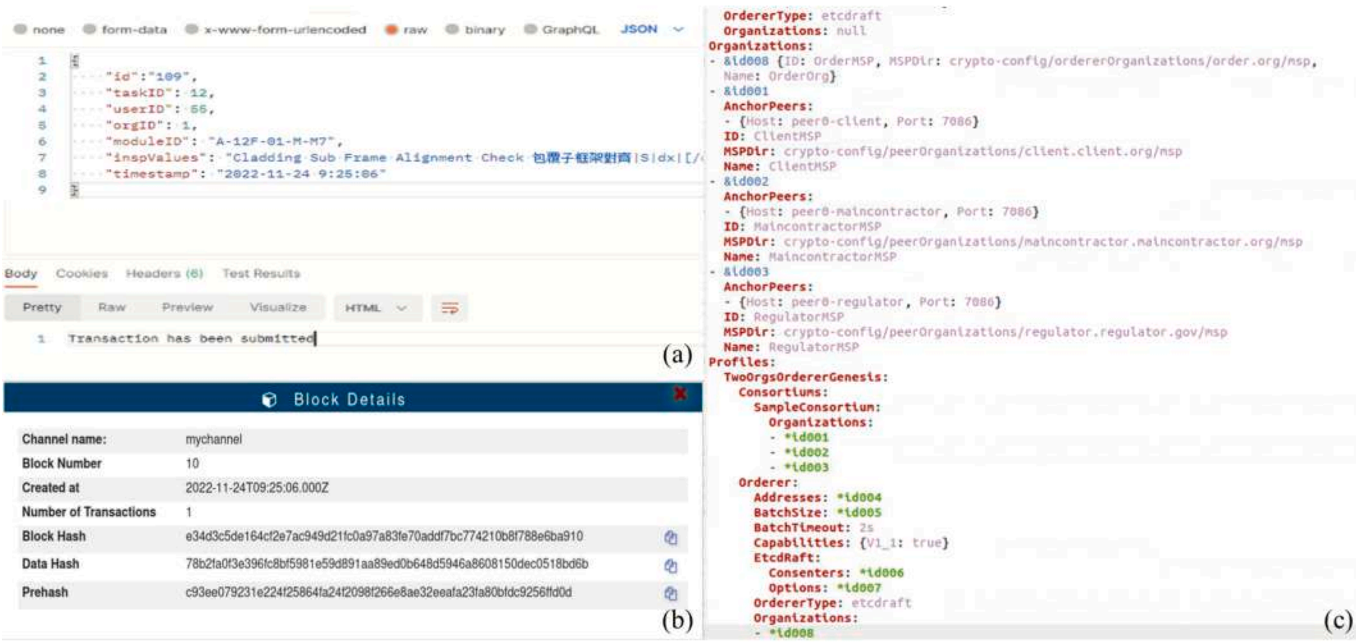

Fig. 10. Blockchain is used in BC4DT. (a) API format; (b) one example block; and (c) organization configuration.

In order to provide the performance of blockchain in BC4DT, we refer to the rigorous infrastructure evaluation conducted in previous supporting studies (Zhao et al., 2023). BC4DT based on cached twin module can realize the average response time of upload operation is only 3.53 milliseconds, while the response time of query operation is only 1.10 milliseconds. It is particularly important that the system can still maintain this millisecond response speed in the offline state, thus avoiding the common problem that the traditional blockchain cannot respond in the offline state.

4.4. Interviews with industry experts

To validate the BC4DT results, we conducted a series of interviews. In October 2024, we conducted semi-structured interviews with four participants. Table 5 presents the basic information of the interviewees in this study, systematically organizing and presenting their roles, core responsibilities, and professional experience. Four key stakeholders from the construction industry were selected as interviewees, including a project manager with 5 to 10 years of experience, whose core responsibility is to coordinate project planning and execution; a technical lead with 3 to 5 years of experience, responsible for leading the implementation of digital innovation; a senior manager from the owner's side with over 10 years of experience, responsible for ensuring the owner's needs and compliance requirements are met; and a quality inspector from a general contractor with over 10 years of experience, primarily responsible for conducting quality inspections and implementing industry standards. This personnel selection method was used to ensure that the research covers the multi-dimensional pain points in CPM practices.

This study's questionnaire comprises four sections. The first section collects users' background information. This section aims to clarify the responsibilities respondents undertake within the project and their specific work experience. The second section investigates challenges associated with existing Blockchain 2.0 technologies in case studies. The questions in this section assess users' experiences regarding network instability and processing delays. The third section focuses on the outcomes of Interpretable ML applications. It centers on whether Interpretable ML accurately identifies key factors affecting module production cycles. The questionnaire concludes with open-ended questions designed to gather additional suggestions and improvement proposals from interviewees.

Table 5
Basic information on interviews.

No.	Role	Duty	Experience
1	Project manager	Oversee project planning and execution	5–10 years
2	Technology leader	Lead implementation of digital innovation	3–5 years
3	Senior officer of the client	Ensure client requirements and compliance	More than 10 years
4	Inspector of the main contractor	Conduct quality inspections and ensure standards	More than 10 years

5. Results

5.1. Prediction results

The BC4DT's predictive performance was evaluated using three key metrics: R2, RMSE, and MAE. The three metrics provide a comprehensive and unbiased view of each model's accuracy and reliability in predicting the production cycle for MiC projects. Table 6 compares the experimental results for eight ML algorithms tested.

Among the assessed ML algorithms, TPOT achieved the most favorable predictive performance. Specifically, TPOT yielded the highest R^2 score of 0.972, coupled with the lowest RMSE of 5.386 days and the

Table 6
Performance Evaluation of ML Models (10-fold, 10 runs, random seed ∈ {2020, 2021, …, 2029}; ↑: Higher is better; ↓ Lower is better; Best in each column in bold).

Category	Algorithms	R^2 Score ↑	RMSE (day) ↓	MAE (day) ↓
Automated ML	H2O	0.888	10.003	8.042
	TPOT	**0.972**	**5.386**	**3.853**
	Auto-sklearn	0.946	7.506	5.139
Tree	Decision Tree	0.949	7.302	4.667
Bagging	Random Forest	0.952	7.503	4.494
Boosting	Gradient Boosting	0.895	10.450	5.225
Meta	KNN	0.663	18.782	8.991
Linear	Linear Regression	0.933	8.872	5.313

lowest MAE of 3.853 days. The performance indicators of the remaining two AutoML algorithms, i.e., H2O and Auto-sklearn, were very close, with R^2 scores of 0.888∼0.946 and RMSE values of 7.506∼10.003 days. Thus, the satisfactory results confirm that AutoML achieves the designated BC4DT's prediction goal.

Other algorithms, such as Random Forest, also demonstrated strong performance with an R^2 score of 0.952, an RMSE of 7.503 days, and an MAE of 4.494 days. In contrast, some non-learning algorithms, like KNN, exhibited comparatively low performance, with an R^2 score of 0.663, an RMSE of 18.782 days, and an MAE of 8.991 days; similar lower performance was observed in another non-learning algorithm, Linear Regression. The TPOT results, highlighting its superior performance metrics, were bolded in the table. Since TPOT model achieves the highest prediction accuracy, it is important to understand its complex decision-making process. For this reason, this study applies the analytical methods of SHAP and LIME to the TPOT model.

5.2. Interpretable ML results

The interpretability of Decision Tree models is key to understanding how predictions are made when evaluating ML models for predicting module production period timelines. The Decision Tree model shown in Fig. 11 was used to analyze the influencing factors of owner lag in construction projects. The model uses hierarchical splitting to screen key variables. The root node uses "T25 Visual Inspection by MC ≤ 0.5" as the splitting condition, and relevant data from 449 samples were recorded. The left branch uses "Block Type ≤ 0.5" as the condition, and the samples are further divided into sub nodes with spatial features such as "Floor Type ≤ 0.5" and "Floor ≤ 5.5". The sample size, squared error, and other data of each node are used to quantify the impact of different factors on owner lag, providing data support for project management optimization.

The Fig. 12 showing the SHAP value feature importance is used to quantify the average impact of each input feature on the output of the prediction model related to building projects. The average absolute SHAP value of each feature is presented visually. Among them, "Block Type" is identified as the single feature with the greatest impact, with an average SHAP value of +13.63. Features related to main contractor inspections (such as "T25 Visual Inspection by MC" with +6.47) and spatial attributes (such as "Floor" with +5.69) also play critical roles. Furthermore, a series of lag indicators (e.g., 90% CI Lag Low for BSI and Client, Min Lag) have surfaced as significant factors. This objective machine learning finding highlights that coordination and inspection delays among different stakeholders (Client, BSI, and MC) form a major bottleneck affecting the overall module production cycle, which often overshadows the traditional emphasis on simple worker learning curves.

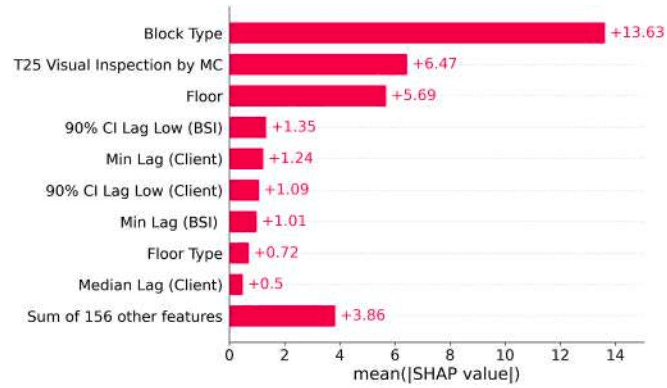

Fig. 12. SHAP summary plot of feature importance.

Fig. 13 presents the LIME analysis of specific prediction results for the production cycle of a MiC module. The analysis reveals the contribution levels of each feature variable to this prediction. Fig. 13(a) displays feature contributions ranked by their influence on the predicted value. Block Type and Floor Type were identified as the most influential characteristics, with contributions of minus 13.57 and 0.84. Other key factors included Median Lag (Client) and 90% CI Lag High (Client), both contributed over 0.3 units to the prediction.

Fig. 13(b) presents the predicted range for this module's production cycle. The forecasted minimum is 107.05 days, maximum is 326.04 days, with the actual prediction at 170.82 days. And Fig. 13(c) lists the values of relevant influencing factors, such as Block Type, Floor, and specific numerical values for other indicators.

5.3. Interview results from industrial experts

Interviewees were asked about their network-related difficulties to ensure the network issues targeted by the BC4DT reflect common challenges in construction projects. Their feedback provided insights into typical connectivity problems faced during project execution. In our interviews, all interviewees stated:

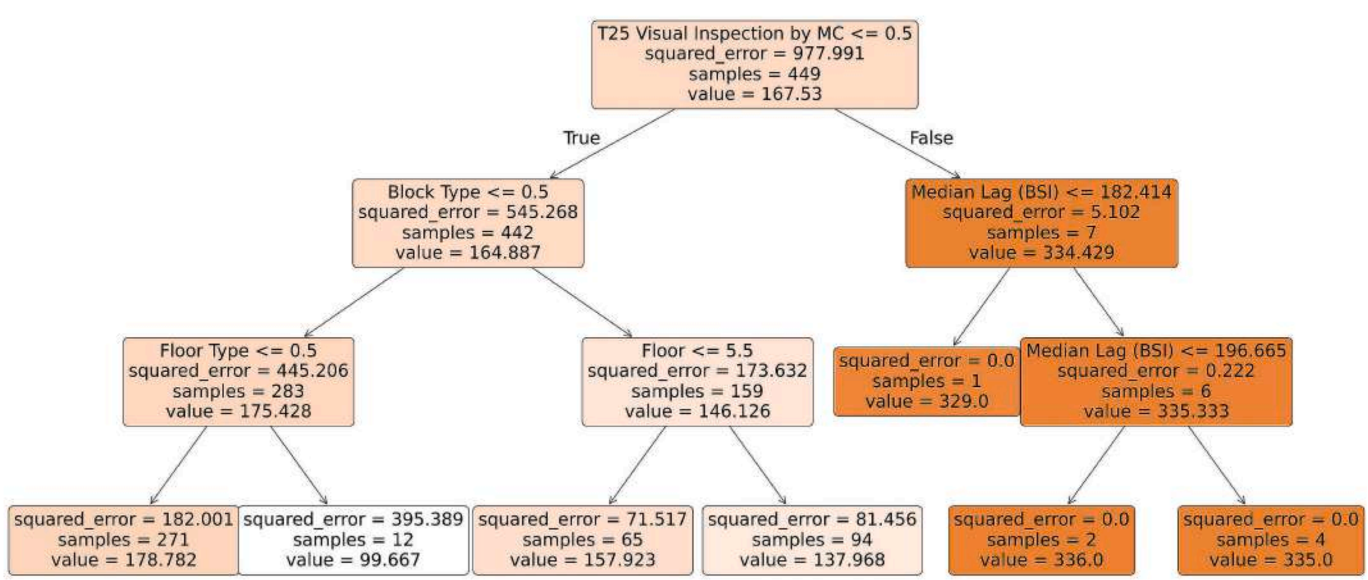

Fig. 11. Decision tree visualization for production period prediction (Darker color indicates higher value).

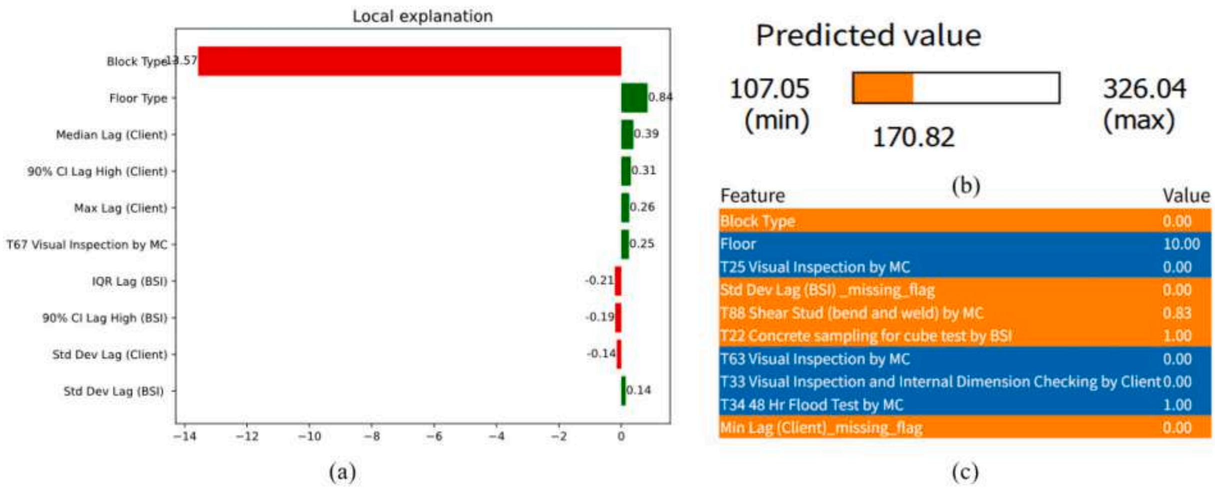

Fig. 13. LIME explanation (Instance No. 1).

"Both the unstable network environment at the construction site and high delays after submitting inspection records, I have encountered."

Their responses confirmed that unstable network issues were a significant challenge in construction projects. Furthermore, to assess the potential impact of the Blockchain 4.0 paradigm, we asked interviewees how satisfied they would be if the paradigm could effectively address these network challenges. Most of the interviewees said that:

"Very satisfied if BC4DT could be implemented to solve the network problem."

The responses confirm that the proposed BC4DT aligns well with the actual needs of project participants. Furthermore, when the data analysis function of the BC4DT was introduced. Interviewee 3 emphasized:

"Undoubtedly, data analysis for construction projects is one of the most important aspects. Data analysis can better identify which parts of the entire MiC production process have the highest failure rate and which workflows can be optimized to speed up construction efficiency."

The next questions focused on validating the Interpretable ML results. First, interviewees were asked whether they observed that spatial and structural characteristics strongly correlated with project timelines, given that the ML model identified "Block Type" as the most critical feature (SHAP value +13.63) followed by "Floor" (+5.69). The interviewees confirmed:

"Specifically, there are only two types of modules in this MiC project, toilets and hostel rooms, so the biggest difference is the difference in production time. Since the lower floors are produced first, the inspectors have more time to familiarize themselves with the inspection process. Therefore, the *completion time of* the *inspection tasks for* the *lower floors is relatively faster."*

When "we presented this finding regarding the high importance of coordination delays (such as Client and BSI lag) to the experts, they strongly acknowledged this hidden impact. Interviewees identified the following key factors in their responses:

"Timeliness of communication is crucial, and reminders can be added to the app. Not only can reminders be sent through the app, but also via WhatsApp or WeChat to minimize these inter organizational lags."

Finally, the open discussion allowed respondents to share their expectations for improvements to BC4DT from their own perspectives. The key responses are summarized below:

"We hope to add AI recognition capabilities to detect images and reduce duplicate, erroneous, and blurry image uploads. We also hope to improve stability and provide notification features to enhance user collaboration."

To sum up, based on the results of the questionnaire survey, there are some common problems in construction projects. These common problems include network instability and delayed data submission. The respondents confirmed that BC4DT can address these challenges and expressed high satisfaction with its potential. Additionally, the data intelligence analysis function of BC4DT was widely recognized by the respondents. This function helps to more accurately identify the key factors for improving production efficiency.

6. Discussion

The BC4DT paradigm proposed in this paper has the following three main advantages. First, the researchers integrated the automatic machine learning technology into the BC4DT framework. In the selected mic scenario, the BC4DT prediction accuracy reached 0.972 (R^2 = 0.972). Secondly, this study combined with the actual case to verify the offline capability of BC4DT in the real scene. Thirdly, the interpretability of ML algorithm is proved, which alleviates the common black box problem in artificial intelligence. SHAP and LIME technology are applied in the system, and industry experts are invited to offer practical insights and qualitative support.

Although BC4DT has addressed some challenges related to CPM, several limitations remain in the current research. The first limitation concerns the scope and scalability of the research. The existing validation work is limited to specific MiC projects, and the number of nodes involved is limited. Future research should expand the application scenarios for BC4DT. The focus of the evaluation is on the scalability and latency performance of BC4DT under high data loads (Liang et al., 2024). The second limitation concerns the scope of BC4DT's implementation. The current research implementation of BC4DT mainly focuses on the production stage of MiC. Future research needs to expand the application scenarios of CPM, including cross-border transportation and on-site assembly.

BC4DT can assist the CPM model in transforming it from passive data monitoring to intelligent data management. This enables professionals in the construction industry to make better use of the massive data generated by construction projects through BC4DT. Construction project managers can leverage BC4DT's tamper-proof nature for passive dispute recording and apply its predictive capabilities to analyze and manage project progress.

In practical application, it should be noted that at the beginning of a new project, the parameters such as learning rate and various detection ratios are often unknown. Therefore, at the beginning of the project, BC4DT mainly used its blockchain technology to store data and record the project process. In the later stage, one can use the interpretable machine learning to summarize the bottlenecks in the current project. Due to the universality of the challenges faced in the process of modular construction, the project manager can use the historical data of the previous project as the benchmark of the new project, so as to more effectively optimize the workflow and resource allocation and then improve the construction efficiency of subsequent modular construction projects.

7. Conclusion

Traditional CPM DT suffers from fragmented data integration, insecure access to information, and inefficient stakeholder coordination. Although the Blockchain 3.0 framework avoids reliance on a stable internet connection, it fails to meet the need to analyze and efficiently manage the massive data from construction sites. In addition, the existing DT is more focused on digital replicas and lacks application value in data analysis.

This paper presents two key innovations. First, unlike traditional ML algorithms, the AI implemented in the BC4DT is based on AutoML and interpretable ML, enabling it to accurately and interpretably identify the most significant barriers in CPMs. Second, the proposed BC4DT has been validated through real-world case studies and interviews with several industry experts to effectively reveal the most significant barriers in CPM, which are more conducive to stakeholder training, planning, simulation, and thus the promise of DT.

The experimental results showed that BC4DT achieved an R^2 score of 0.972 and an RMSE of 5.386 days when predicting the production time of MiC modules. At the same time, the experts interviewed said that BC4DT, combined with ML, could be more effective at making practical, efficient recommendations for construction project production processes. The proposed BC4DT enables construction projects to use blockchain technology without relying on a continuous network and extends its capabilities by incorporating advanced artificial intelligence techniques. Future research can be directed towards (i) conducting large-scale experiments on real projects, (ii) exploring more areas involved in construction projects for data analysis and prediction, and (iii) exploring smarter approaches by combining multi-chain technologies and integrating more intelligent agents.

Declaration of generative AI in writing

The authors declare the following ways of interactions with generative AI which may be considered as potential uses in writing or editing the content in this paper: GenAI-assisted proofreading. A university self-hosted LLM GenAI was used to assist with proofreading the work to eliminate grammatical and connection errors in the main text, and to improve readability of Abstract, Introduction, and Conclusion sections. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

CRediT authorship contribution statement

Rui Zhao: Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Lingming Kong:** Writing – original draft, Software. **Meng Sun:** Visualization, Validation, Data curation. **Fan Xue:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgments

The work presented in this paper was supported by Hong Kong Research Grants Council (RGC) (No. 17200221). Portions of this article are reproduced from the University of Hong Kong's doctoral dissertation, "A blockchain 4.0 paradigm for coordinating centralized construction digital twin and decentralized project management" (2025), available at http://hdl.handle.net/10722/360625, with permission.

Data availability

Data will be made available on request.

References

Alsakka, F., Yu, H., El-Chami, I., Hamzeh, F., & Al-Hussein, M. (2024). Digital twin for production estimation, scheduling and real-time monitoring in offsite construction. *Computers & Industrial Engineering, 191*, Article 110173. https://doi.org/10.1016/j.cie.2024.110173

Alshboul, O. (2024). Empirical exploration of predictive maintenance in concrete manufacturing: Harnessing machine learning for enhanced equipment reliability in construction project management. *Computers & Industrial Engineering, 190*, Article 110046. https://doi.org/10.1016/j.cie.2024.110046

Li, X., Lu, W., Xue, F., Wu, L., Zhao, R., Lou, J., & Xu, J. (2022). Blockchain-enabled IoT-BIM platform for supply chain management in modular construction. *Journal of Construction Engineering and Management, 148*(2), 04021195. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002229

Wu, L., Lu, W., Xue, F., Li, X., Zhao, R., & Tang, M. (2022). Blockchain and smart contracts: A solution for payment issues in construction supply chains. *Computers in Industry, 135*, Article 103573. https://doi.org/10.1016/j.compind.2021.103573

Zhao, R., Chen, Z., & Xue, F. (2023). A blockchain 3.0 paradigm for digital twins in construction project management. *Automation in Construction, 145*, Article 104645. https://doi.org/10.1016/j.autcon.2022.104645

Lu, W., Wu, L., & Zhao, R. (2023). Rebuilding trust in the construction industry: a blockchain-based deployment framework. *International Journal of Construction Management, 23*(8), 1405. https://doi.org/10.1080/15623599.2021.1974683

Kong, L., Zhao, R., Anumba, C. J., Lu, W., & Xue, F. (2025). Open BIM exchange on Blockchain 3.0 virtual disk: A traceable semantic differential transaction approach. *Frontiers of Engineering Management*. https://doi.org/10.1007/s42524-024-4006-x

Atlam, H., Azad, M., Alzahrani, A., & Wills, G. (2020). A review of blockchain in Internet of Things and AI. *Big Data and Cognitive Computing, 4*(4), 28. https://doi.org/10.3390/bdcc4040028

Azhar, S., Khalfan, M., & Maqsood, T. (2012). Building information modeling (BIM): Now and beyond. *Construction Economics and Building, 12*(4), 15–28. https://doi.org/10.5130/ajceb.v12i4.3032

Bodkhe, U., Tanwar, S., Parekh, K., Khanpara, P., Tyagi, S., Kumar, N., & Alazab, M. (2020). Blockchain for industry 4.0: A comprehensive review. *IEEE Access, 8*, 79764–79800. https://doi.org/10.1109/ACCESS.2020.2988579

Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic construction digital twin: Directions for future research. *Automation in Construction, 114*, Article 103179. https://doi.org/10.1016/j.autcon.2020.103179

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324

Bryde, D., Martí, B., & Volm, J. (2013). The project benefits of building information modelling (BIM). *International Journal of Project Management, 31*(7), 971–980. https://doi.org/10.1016/j.ijproman.2012.12.001

Chen, C., Fu, H., Zheng, Y., Tao, F., & Liu, Y. (2023). The advance of digital twin for predictive maintenance: The role and function of machine learning. *Journal of Manufacturing Systems, 71*, 581–594. https://doi.org/10.1016/j.jmsy.2023.10.010

Chen, K., Zhou, X., Bao, Z., Skibniewski, M. J., & Fang, W. (2025). Artificial intelligence in infrastructure construction: A critical review. *Frontiers of Engineering Management, 12*(1), 24–38. https://doi.org/10.1007/s42524-024-3128-5

Daboun, O., Aminah, M. Y., & Ali, R. K. (2023). Relationship management in construction projects: Systematic literature review. *Engineering Management Journal, 35*(2), 120–143. https://doi.org/10.1080/10429247.2022.2041962

Dangana, M., Ansari, S., Asad, S. M., Hussain, S., & Imran, M. A. (2022). Towards the digital twin (DT) of narrow-band Internet of Things (NBIoT) wireless communication in industrial indoor environment. *Sensors, 22*(23), 9039. https://doi.org/10.3390/s22239039

Datta, S., Islam, M., Ahmed, S., Kar, K., & Kar, M. (2024). Artificial intelligence and machine learning applications in the project lifecycle of the construction industry: A comprehensive review. *Heliyon, 10*(5). https://doi.org/10.1016/j.heliyon.2024.e26888

Feurer, M., Aaron, K., Katharina, E., Jost, S., Manuel, B. & Frank, H. (2015). Efficient and robust automated machine learning. *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, (pp. 2962–2970).

Galloway, P. D. (2006). Survey of the construction industry relative to the use of CPM scheduling for construction projects. *Journal of Construction Engineering and Management, 132*(7), 697–711. https://doi.org/10.1061/(ASCE)0733-9364(2006)132:7(697)

Grieves, M., & Vickers, J. (2017). Digital Twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. *Transdisciplinary Perspectives on Complex Systems, 85–113*. https://doi.org/10.1007/978-3-319-38756-7_4

Harris, F., McCaffer, R., Baldwin, A. & Edum-Fotwe, F. (2021). *Modern construction management.* Wiley-Blackwell.

Kerzner, H. (1998). *Project management: A systems approach to planning, scheduling, and controlling* (6th ed.). John Wiley & Sons.

Knebel, F. P., Trevisan, R., do Nascimento, G. S., Abel, M., & Wickboldt, J. A. (2023). A study on cloud and edge computing for the implementation of digital twins in the oil & gas industries. *Computers & Industrial Engineering, 182*, Article 109363. https://doi.org/10.1016/j.cie.2023.109363

LeDell, E., & Poirier, S. (2020). H2O AutoML: Scalable automatic machine learning. *7th ICML workshop on automated machine learning*.

Li, W., Duan, P., & Su, J. (2024). The effectiveness of project management construction with data mining and blockchain consensus. *Journal of Ambient Intelligence and Humanized Computing, 15*, 175. https://doi.org/10.1007/s12652-020-02668-7

Li, X., Wu, L., Zhao, R., Lu, W., & Xue, F. (2021). Two-layer Adaptive Blockchain-based Supervision model for off-site modular housing production. *Computers in Industry, 128*, 103437. https://doi.org/10.1016/j.compind.2021.103437

Li, Z., Barenji, A. V., & Huang, G. Q. (2018). Toward a blockchain cloud manufacturing system as a peer to peer distributed network platform. *Robotics and Computer-Integrated Manufacturing, 54*, 133–144. https://doi.org/10.1016/j.rcim.2018.05.011

Liang, D., Chen, S.-H., Chen, Z., Wu, Y., Chu, L. Y. L., & Xue, F. (2024). 4D point cloud-based spatial-temporal semantic registration for monitoring mobile crane construction activities. *Automation in Construction, 165*, 105576. https://doi.org/10.1016/j.autcon.2024.105576

Lin, J., & Wei, L. (1990). Linear regression analysis for multivariate failure time observations. *Journal of the American Statistical Association, 87*(420), 1091–1097. https://doi.org/10.1080/01621459.1992.10476264

Liu, R., Zhao, C., & Tan, C. (2024). Vaccine supply chain decision and coordination with manufacturers competition under blockchain technology. *Computers & Industrial Engineering, 190*, Article 110031. https://doi.org/10.1016/j.cie.2024.110031

Liu, Z., Zhang, A., & Wang, W. (2020). A framework for an indoor safety management system based on digital twin. *Sensors, 20*(20), 5771. https://doi.org/10.3390/s20205771

Lu, Q., Xie, X., Heaton, J., Parlikad, A. K., & Schooling, J. (2019). *From BIM towards digital twin: Strategy and future development for smart asset management. International workshop on service orientation in holonic and multi-agent manufacturing*, 10.1007/978-3-030-27477-1_30 (pp. 392–404). Valencia: Springer.

Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*, 4765–4774.

Luntovskyy, A., & Guetter, D. (2019). Cryptographic technology blockchain and its applications. In *2018 International conference on information and telecommunication technologies and radio electronics. Springer*. https://doi.org/10.1007/978-3-030-16770-7_2

Ma, A. (2020). Emerging legal issues in blockchain for construction supply chains. Proceedings of the 2020 4th international conference on vision (pp. 1-7). New York: Association for Computing Machinery. doi:10.1145/3448823.3448880.

Mahmudnia, D., Arashpour, M., & Yang, R. (2022). Blockchain in construction management: Applications, advantages and limitations. *Automation in Construction, 140*, Article 104379. https://doi.org/10.1016/j.autcon.2022.104379

Mostofi, F., Bahadır, Ü., Tokdemir, O., Toğan, V., & Yepes, V. (2025). Enhancing strategic investment in construction engineering projects: A novel graph attention network decision-support model. *Computers & Industrial Engineering, 203*, Article 111033. https://doi.org/10.1016/j.cie.2025.111033

Mutesi, E. T., & Kyakula, M. (2011). Application of ICT in the construction industry in Kampala. In *Second international conference on advances in engineering and technology* (pp. 263–269).

Natanelov, V., Cao, S., Foth, M., & Dulleck, U. (2022). Blockchain smart contracts for supply chain finance: Mapping the innovation potential in Australia-China beef supply chains. *Journal of Industrial Information Integration, 30*, Article 100389. https://doi.org/10.1016/j.jii.2022.100389

Oesterreich, T. D., & Teuteberg, F. (2016). Understanding the implications of digitisation and automation in the context of Industry 4.0: A triangulation approach and elements of a research agenda for the construction industry. *Computers in Industry, 83*, 121–139. https://doi.org/10.1016/j.compind.2016.09.006

Olawumi, T. O., Chan, D. W., Ojo, S., & Yam, M. C. (2022). Automating the modular construction process: A review of digital technologies and future directions with blockchain technology. *Journal of Building Engineering, 46*(1), Article 103720. https://doi.org/10.1016/j.jobe.2021.103720

Olson, R. & Moore, J. (2019). TPOT: A tree-based pipeline optimization tool for automating machine learning. Proceedings of the 2016 Workshop on Automatic Machine Learning (pp. 151–160). New York: Springer, Cham. doi:10.1007/978-3-030-05318-5_8.

Owusu, E. K., Chan, A. P., DeGraft, O.-M., Ameyaw, E. E., & Robert, O.-K. (2019). Contemporary review of anti-corruption measures in construction project management. *Project Management Journal, 50*, 40–56. https://doi.org/10.1177/8756972818808983

Perera, G., Tennakoon, T., Kulatunga, U., Jayasena, H. S., & Wijewickrama, M. (2021). Selecting suitable procurement system for steel building construction. *Built Environment Project and Asset Management, 11*(4), 611. https://doi.org/10.1108/BEPAM-03-2020-0056

Perera, S., Nanayakkara, S., Rodrigo, M., Senaratne, S., & Weinand, R. (2020). Blockchain technology: Is it hype or real in the construction industry? *Journal of Industrial Information Integration, 17*, Article 100125. https://doi.org/10.1016/j.jii.2020.100125

Qi, Q., & Tao, F. (2018). Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. *IEEE Access, 6*, 3585. https://doi.org/10.1109/ACCESS.2018.2793265

Ribeiro, M., Singh, S. & Guestrin, C. (2016). "Why should i trust you?": Explaining the predictions of any classifier. KDD '16: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mini (pp. 1135 - 1144). ACM. doi:10.1145/2939672.2939778.

Sadeghi, M., Mahmoudi, A., & Deng, X. (2023). Blockchain technology in construction organizations: Risk assessment using trapezoidal fuzzy ordinal priority approach. *Engineering, Construction and Architectural Management, 30*(7), 2767. https://doi.org/10.1108/ECAM-01-2022-0014

Sears, S., Sears, G., Clough, R., Rounds, J., & Segner, R. (2015). *Construction Project Management* (6th Edition). Wiley.

Soori, M., Dastres, R., & Arezoo, B. (2023). AI-powered blockchain technology in industry 4.0, a review. *Journal of Economy and Technology, 1*, 222–241. https://doi.org/10.1016/j.ject.2024.01.001

Sreedharan, S., Ramachandran, M., & Ramesh, D. (2025). Harnessing digital twins and industrial-IoT for cutting-edge mining automation: A methodological and technology assessment prototype. *Computers & Industrial Engineering, 201*, Article 110871. https://doi.org/10.1016/j.cie.2025.110871

Stanitsas, M., Konstantinos, K., & Vrassidas, L. (2021). Integrating sustainability indicators into project management: The case of construction industry. *Journal of Cleaner Production, 279*, Article 123774. https://doi.org/10.1016/j.jclepro.2020.123774

Taylan, O., Bafail, A., Abdulaal, R., & Kabli, M. (2014). Construction projects selection and risk assessment by fuzzy AHP and fuzzy TOPSIS methodologies. *Applied Soft Computing, 17*, 105–116. https://doi.org/10.1016/j.asoc.2014.01.003

Teizer, J., Wolf, M., Golovina, O., Perschewski, M., Propach, M., Neges, M., & König, M. (2017). Internet of Things (IoT) for integrating environmental and localization data in Building Information Modeling (BIM). In *34th international symposium on automation and robotics in construction* (pp. 603–609). https://doi.org/10.22260/ISARC2017/0084

UKNIC. (2017). Data for the public good. London, UK: United Kingdom National Infrastructure Commission. Retrieved from https://nic.org.uk/app/uploads/Data-for-the-Public-Good-NIC-Report.pdf.

Vrchota, J., Petr, Ř., Monika, M., & Martin, P. (2021). Critical success factors of the project management in relation to industry 4.0 for sustainability of projects. *Sustainability, 13*(1), 281. https://doi.org/10.3390/su13010281

Wang, W., Hoang, D. T., Hu, P., Xiong, Z., Niyato, D., Wang, P., Wen, Y., & Kim, D. I. (2019). A survey on consensus mechanisms and mining strategy management in blockchain networks. *IEEE Access, 7*, 22328. https://doi.org/10.1109/ACCESS.2019.2896108

Wang, Z., Wang, T., Hu, H., Gong, J., Ren, X., & Xiao, Q. (2020). Blockchain-based framework for improving supply chain traceability and information sharing in precast construction. *Automation in Construction, 111*, Article 103063. https://doi.org/10.1016/j.autcon.2019.103063

Wong, S., Zheng, C., Su, X., & Tang, Y. (2024). Construction contract risk identification based on knowledge-augmented language models. *Computers in Industry, 157–158*, Article 104082. https://doi.org/10.1016/j.compind.2024.104082

Wu, H., Zhang, P., Li, H., Zhong, B., Fung, I. W., & Lee, Y. Y. (2022). Blockchain technology in the construction industry: Current status, challenges, and future directions. *Journal of Construction Engineering and Management, 148*(10), Article 03122007. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002380

Xue, F., Chen, K., Lu, W., Niu, Y., & Huang, G. (2018). Linking radio-frequency identification to Building Information Modeling: Status quo, development trajectory and guidelines for practitioners. *Automation in Construction, 93*, 241–251. https://doi.org/10.1016/j.autcon.2018.05.023

Yaqoob, I., Salah, K., Uddin, M., Jayaraman, R., Omar, M., & Imran, M. (2020). Blockchain for digital twins: Recent advances and future research challenges. *IEEE Network, 34*(5), 290–298. https://doi.org/10.1109/MNET.001.1900661

Zhang, X., Wang, X., Yao, H., Luo, W., Qiang, Z., & Luo, D. (2025). Multi-similarity and gradient fusion digital twins for fault detection and diagnosis of rolling bearings. *Computers in Industry, 168*, Article 104273. https://doi.org/10.1016/j.compind.2025.104273

Zhong, B., Pan, X., Ding, L., Chen, Q., & Hu, X. (2023). Blockchain-driven integration technology for the AEC industry. *Automation in Construction, 150*, Article 104791. https://doi.org/10.1016/j.autcon.2023.104791