



Integration of Building Information Modeling (BIM) and Artificial Intelligence for Predicting Construction Project Risks

Gland Carizo*¹

Email: cari@gmail.com

¹Universitas Papua, Manokwari, Indonesia, 98314

*Corresponding Author

Abstract

Building Information Modeling (BIM) has increasingly been adopted in construction projects; however, its potential for quantitative risk prediction remains underexplored. This study aims to examine the influence of BIM-based variables on project risk severity and to develop a predictive model that explains variations in schedule delay and cost overrun. A quantitative approach was employed, using data from 42 completed construction projects that implemented 3D, 4D, and 5D BIM. Four independent variables-Geometric Complexity Index, Scheduling Density Ratio, Cost Intensity Coefficient, and Object Interdependency Level-were extracted from BIM objects and analyzed using multiple linear regression and machine learning models, including Random Forest, Support Vector Machine, and Artificial Neural Network. The results indicate that the regression model explains 68% of the variance in project risk severity, with scheduling density and cost intensity emerging as the most influential predictors. Among the machine learning models, the Artificial Neural Network achieved the highest predictive accuracy, demonstrating superior capability in capturing nonlinear relationships among BIM-derived attributes. These findings confirm that structured BIM-based metrics can serve as reliable indicators for proactive risk assessment. The study contributes to integrating BIM analytics into construction risk management frameworks and highlights the importance of data-driven decision-making in improving project performance.

Keywords: Building Information Modeling (BIM), Construction Risk Prediction, Regression Analysis, Machine Learning, Project Performance.

I. INTRODUCTION

The construction industry is globally characterized by high levels of uncertainty and complexity (Ghaleb et al., 2022), which frequently result in project delays (Latif et al., 2023), cost overruns (Abdelalim et al., 2025), and quality deficiencies (Moshood et al., 2025). A report by McKinsey & Company indicates that large-scale construction projects experience average cost overruns of more than 20% and schedule delays in nearly 80% of cases (Dolphin, 2025). This phenomenon is increasingly evident in the context of rapid infrastructure development (Xiaolong et al., 2021), where projects are exposed to intertwined technical and managerial (Cui et al., 2024), financial (Horvat et al., 2021), and environmental risks (Rahaman et al., 2023). Such conditions highlight the limitations of conventional risk management approaches, which are often reactive and rely heavily on expert judgment, underscoring the need for more data-driven and predictive risk management solutions.

In recent years, buildingSMART has actively promoted Building Information Modeling (BIM) as a digital paradigm for planning, design, and construction management (Yang et al., 2021). BIM

enables the integration of geometric and non-geometric information within a coordinated digital model, thereby enhancing interdisciplinary collaboration and reducing design-related errors (Daffa & Azhar, 2022). Previous studies (Paik et al., 2022) have demonstrated that BIM adoption can improve design coordination, construction efficiency, and information transparency. Nevertheless, most BIM applications remain focused on visualization, clash detection, and documentation management. At the same time, the rich data embedded in BIM models have not been fully exploited for predictive risk analysis during the project lifecycle.

Parallel to the advancement of BIM, Artificial Intelligence (AI), particularly machine learning and deep learning techniques, has shown significant potential in construction risk prediction (Rane, 2023). Prior research by Tiwari & Hussain (2025) and Hriday & Rehman (2025) confirms that AI-based models can capture complex nonlinear patterns in historical project data to predict schedule delays, cost overruns, and contractual disputes with higher accuracy than traditional statistical methods. AI also enables the processing of large and heterogeneous datasets, including progress reports, sensor data, and project records (Zha et al., 2025). However, existing AI-based studies in construction management are largely disconnected from BIM environments, relying instead on standalone datasets that do not fully utilize the structured, object-oriented, and information-rich nature of BIM models.

Based on the above review, a clear research gap is evident. First, BIM-related studies predominantly emphasize coordination and efficiency improvements, with limited attention to predictive and proactive risk management capabilities. Second, AI-based construction risk studies often neglect direct integration with BIM as a primary data source, resulting in the underutilization of spatial, temporal (4D), and cost-related (5D) information inherent in BIM models. Third, there is a lack of comprehensive frameworks that integrate BIM and AI to support dynamic, real-time risk prediction throughout the construction project lifecycle. This gap underscores the urgency of developing an integrated approach that bridges BIM's data-rich digital environment with the predictive power of AI.

The novelty of this research lies in the integration of multi-dimensional BIM-derived features—including geometric complexity, scheduling density, cost intensity, and object interdependency—into machine learning models for construction risk prediction. Unlike previous studies that typically rely on a limited set of BIM indicators or conventional statistical approaches, this study systematically extracts and operationalizes BIM object attributes as structured inputs for predictive modeling. Furthermore, this research contributes by providing a comparative evaluation of multiple machine learning algorithms (Regression, Random Forest, Support Vector Machine, and Artificial Neural Networks) using real-world BIM-based project data. This enables

a more comprehensive understanding of how different modeling approaches perform in capturing nonlinear relationships among BIM variables and project risk outcomes.

Practically, this study demonstrates how BIM data can be transformed from a coordination and visualization tool into a data-driven predictive analytics framework that supports early risk identification. The integration of BIM and AI in this study provides a structured approach for improving risk-informed decision-making in construction project management. Accordingly, the objective of this study is to develop and evaluate an integrated BIM–AI model for predicting construction project risks. This study addresses the following research question: how do BIM-derived variables influence construction project risk, and how do different machine learning models perform in predicting such risks?

II. LITERATURE REVIEW

A. Building Information Modeling (BIM) in Construction Project Management

Building Information Modeling (BIM) has evolved from a three-dimensional modeling tool into a comprehensive digital information management system that integrates geometric, semantic, temporal, and cost-related data across the project lifecycle (Mustafa et al., 2023). Promoted globally by buildingSMART, BIM facilitates interoperability and standardized information exchange through open formats such as Industry Foundation Classes (IFC) (Chatsuwan et al., 2025). According to Abdelalim et al. (2024) BIM enhances collaboration among stakeholders by enabling a shared digital representation of a facility's physical and functional characteristics.

Previous studies have demonstrated BIM's effectiveness in clash detection, quantity take-off automation, scheduling (4D BIM), and cost estimation (5D BIM). Das et al. (2025) reported improvements in project coordination and reductions in rework through BIM implementation. Siebelink et al. (2021) further conceptualized BIM maturity stages, highlighting its strategic integration into organizational processes. Despite these advancements, BIM applications remain predominantly descriptive and coordination-oriented. The analytical potential of BIM datasets, particularly for predictive risk assessment, has not been fully exploited. This limitation suggests the need to extend BIM beyond visualization and coordination to integrate advanced analytics.

B. Artificial Intelligence and Machine Learning in Construction Risk Prediction

Artificial Intelligence (AI), especially machine learning (ML) and deep learning (DL), has gained increasing attention in construction management research (Prasittisopin, 2025). AI techniques such as Artificial Neural Networks (ANN) (Micheli-Tzanakou, 2023), Support Vector Machines (SVM) (Bilal et al., 2024), Random Forest (Ratih, 2025), and Gradient Boosting have been applied to predict cost overruns, schedule delays, safety incidents, and claim risks

(ForouzeshNejad et al., 2024). Studies by Shen (2025) and Alshboul et al. (2022) demonstrate that ML models outperform traditional regression-based methods in capturing nonlinear and complex relationships among project variables.

Moreover, AI supports predictive analytics by learning patterns from historical project data and generating probabilistic forecasts. Paul et al. (2025) emphasize that AI enables real-time decision support when integrated with digital project data streams. However, many AI-based construction studies rely on structured tabular datasets derived from completed projects rather than directly utilizing object-oriented digital models such as BIM. Consequently, the integration between AI-driven predictive systems and BIM-based digital environments remains limited and fragmented.

C. Construction Project Risk Management Frameworks

Risk management in construction traditionally follows systematic processes outlined in standards such as those developed by the Project Management Institute in its PMBOK framework (Khodabakhshian et al., 2023). These processes include risk identification, qualitative and quantitative analysis, response planning, and monitoring. Conventional quantitative methods often involve Monte Carlo simulation, decision trees, or statistical regression models (Ismail et al., 2024).

While these methods provide structured procedures, they are typically dependent on expert judgment and historical averages, which may not adequately reflect dynamic project conditions (Akinboboye et al., 2022). Furthermore, conventional frameworks often lack integration with real-time digital project data. As construction projects become more complex and data-intensive, there is a growing demand for predictive, automated, and continuously updated risk assessment mechanisms (Ojo, 2025). This shift aligns with the broader digital transformation of the Architecture, Engineering, and Construction (AEC) industry (Rafsanjani & Nabizadeh, 2023).

D. Integration of BIM and AI: Toward Predictive and Data-Driven Risk Management

Recent research has begun exploring the convergence of BIM and AI to enhance construction project performance. BIM provides a data-rich environment containing object attributes, spatial relationships, scheduling data, and cost information, while AI offers advanced computational capabilities for pattern recognition and prediction (Alotaibi & Alzaidan, 2024). The integration of these technologies has been conceptualized as a pathway toward “intelligent BIM” or “BIM-driven analytics.”

Several exploratory studies have attempted to link BIM models with machine learning algorithms for safety risk prediction and construction progress monitoring. However, these efforts are often limited to specific use cases and lack a comprehensive framework that systematically integrates

BIM object data (3D/4D/5D) into AI-based predictive models for holistic project risk assessment. In addition, challenges related to data extraction, feature engineering, interoperability, and model validation remain insufficiently addressed in the literature (Rehman, 2025).

Therefore, although BIM and AI have individually demonstrated significant benefits, their combined application for predictive construction risk management is still in an early developmental stage. The absence of a structured, validated integration framework highlights the need for further research to bridge technological capabilities and practical implementation strategies. This literature review underscores the theoretical foundation and empirical gaps that justify developing an integrated BIM-AI model to predict construction project risks in a proactive, data-driven manner.

III. RESEARCH METHODS

A. Study Design

This study adopts a quantitative research design using a predictive modeling approach. The research integrates Building Information Modeling (BIM) data with Artificial Intelligence (AI) algorithms to develop and validate a construction risk prediction model. The design is explanatory and model-driven, aiming to examine the relationship between BIM-based project variables and construction risk outcomes. A supervised machine learning framework is employed, in which historical project data serve as the training and testing datasets. Statistical validation techniques are applied to ensure the robustness and reliability of the predictive model.

B. Population and Sample

The population of this study comprises construction projects that have implemented BIM during their planning and execution phases. The sampling technique uses purposive sampling, selecting projects that meet specific criteria: availability of complete BIM models (3D/4D/5D), documented schedule and cost performance data, and recorded risk events. The sample includes a set of completed construction projects within a defined time range to ensure data consistency and comparability. Each project represents a unit of analysis, while BIM object attributes and project performance indicators serve as observational variables.

The final sample consists of 42 completed construction projects that implemented BIM at different levels, including 3D, 4D, and 5D. These projects span various sectors, including residential, commercial, and infrastructure development. In terms of scale, the projects range from small to large, based on contract value and physical size.

All selected projects include complete BIM models, detailed scheduling information, cost data, and documented risk events. The projects also vary in structural complexity, the number of BIM

objects, and interdependencies among elements. In total, 1,268 BIM objects were extracted and used as the basis for feature construction in this study.

C. Data Collection Techniques and Instruments

Data were collected from two primary sources: BIM models and project documentation archives. BIM data were extracted using Industry Foundation Classes (IFC) standards promoted by buildingSMART to ensure interoperability. The extracted variables include object type, material specifications, quantity take-offs, scheduling parameters (start time, duration, dependencies), and cost attributes.

Additional project performance data, such as actual completion time, cost deviations, and recorded risk events, were obtained from project reports and contract documentation. A structured data extraction template was developed to standardize the collection and transformation of BIM-derived variables into a consistent analytical dataset. To ensure data quality and methodological rigor, a data verification and consistency assessment approach was applied. The validity of the extracted variables was evaluated through expert judgment involving construction management and BIM specialists, ensuring that each variable appropriately represents relevant project characteristics and risk-related factors.

Instead of relying solely on traditional questionnaire-based validation, this study adopts a data-driven approach. The extracted BIM variables were cross-checked with project documentation to ensure accuracy and consistency across data sources. Reliability was assessed in terms of data consistency across projects, rather than internal perception-based constructs. Although internal consistency measures such as Cronbach's Alpha were explored, their interpretation is applied with caution, as the variables represent structured computational data rather than latent constructs. Therefore, the use of reliability testing in this study is intended as a supplementary measure to assess the coherence of the extracted features, rather than a primary validation technique.

D. BIM Data Extraction and Variable Definition

Each BIM-derived variable was operationally and technically defined to ensure clarity, transparency, and reproducibility of the data extraction process. Operational definitions describe the conceptual meaning of each variable, while technical definitions explain how the variables were measured or computed from BIM data. For instance, geometric complexity was quantified based on the number of elements and their spatial configuration, scheduling density was calculated as the ratio of overlapping activities to total project duration, cost intensity was measured as cost per unit volume or area, and object interdependency level was derived from the number of relationships between BIM elements using dependency mapping in BIM software. All

variables were extracted from BIM tools such as Autodesk Revit and Navisworks using standardized procedures and further processed with structured data extraction scripts.

E. Risk Severity Index Construction

The dependent variable in this study, namely construction project risk, is quantified using the Risk Severity Index (RSI). The RSI is designed to capture the combined effect of the probability of risk occurrence and its associated impact on project performance. This approach enables a structured and quantitative representation of risk levels across different construction projects.

The RSI is calculated using the following formulation;

$$RSI = \frac{\sum_{i=1}^n (P_i \times I_i)}{n}$$

where RSI represents the overall risk severity index, P_i denotes the probability of occurrence of risk event i , I_i represents the impact of risk event i , and n is the total number of identified risk events within a project. To provide a more comprehensive assessment, the impact component is further decomposed into multiple dimensions, including cost, time, and quality. Thus, the RSI is extended as follows:

$$RSI = \frac{\sum_{i=1}^n \left(P_i \times (w_1 I_{cost,i} + w_2 I_{time,i} + w_3 I_{quality,i}) \right)}{n}$$

where I_{cost} , I_{time} , and $I_{quality}$ represent the impact of each risk event on cost, schedule, and quality, respectively, while w_1 , w_2 , and w_3 are weighting factors assigned to each impact dimension, with the constraint that $w_1 + w_2 + w_3 = 1$. In this study, the weights are determined based on expert judgment to reflect the relative importance of each dimension in construction project performance. The probability values (P_i) are derived from the historical frequency of risk occurrences across the observed projects. In contrast, the impact values (I_i) are quantified using project performance indicators, including cost deviations, schedule delays, and quality-related issues, as documented in project reports. Each component is normalized to ensure comparability across different projects.

Finally, the RSI values are scaled to the range 0 to 1, with values closer to 1 indicating greater risk severity. This normalized index is then used as the dependent variable in the predictive modeling process, enabling consistent input for machine learning algorithms such as Random Forest, Support Vector Machine (SVM), and Artificial Neural Networks (ANN).

F. Data Analysis Tools

Data analysis was conducted in two stages: statistical analysis and machine learning modeling. Statistical analysis was performed to examine relationships between independent variables and dependent variables using correlation analysis and regression modeling. Hypothesis testing employed the t-test for partial significance and the F-test for simultaneous significance, with significance level (α) set at 0.05. The general F-test formula follows:

$$F = (SSR/k) / (SSE/(n - k - 1))$$

where SSR represents the regression sum of squares, SSE represents the error sum of squares, k denotes the number of independent variables, and n represents the sample size. Machine learning analysis involved training and validating models using algorithms such as Random Forest, Support Vector Machine (SVM), and Artificial Neural Networks (ANN). Model performance was evaluated using accuracy, precision, recall, F1-score, and Root Mean Square Error (RMSE). Cross-validation techniques were applied to prevent overfitting and ensure the model's generalizability.

G. Research Model

The conceptual research model integrates BIM variables as independent variables (X) and construction risk indicators as dependent variables (Y). The independent variables ($X_1, X_2, X_3, \dots, X_n$) represent BIM-derived features, including geometric complexity, scheduling density, cost intensity, and object interdependency levels. The dependent variable (Y) represents the predicted project risk, measured in terms of the probability of delay, the cost overrun percentage, or the risk severity index. The mathematical representation of the predictive model can be expressed as:

$$Y = f(X_1, X_2, X_3, \dots, X_n)$$

where Y denotes the predicted construction project risk, X_1 to X_n denote BIM-based feature variables, and f represents the machine learning function used for prediction. The symbols in the model indicate that construction risk is influenced by multiple interrelated BIM attributes processed through AI-based algorithms.

The findings from instrument testing demonstrate that all variables meet validity requirements ($r\text{-count} > r\text{-table}$) and reliability standards (Cronbach's Alpha > 0.70), confirming that the data collection instrument is suitable for predictive modeling. Overall, this methodology provides a systematic framework for integrating BIM and AI to develop a valid, reliable, and data-driven construction risk prediction model.

IV. RESULT AND DISCUSSION

A. Descriptive Analysis of BIM-Based Variables

The dataset consists of 42 completed construction projects that implemented 3D, 4D, and 5D BIM. A total of 1,268 BIM objects were extracted and transformed into structured features. The independent variables include the geometric complexity index (X_1), the scheduling density ratio (X_2), the cost intensity coefficient (X_3), and the object interdependency level (X_4). The dependent variable (Y) represents the project risk level, measured using a composite risk severity index derived from the schedule delay percentage and the cost overrun percentage. Table 1 presents the descriptive statistics of the research variables.

Table 1. Descriptive Statistics of BIM-Derived Variables and Project Risk

Variable	Min	Max	Mean	Std. Deviation
X₁ Geometric Complexity	0.21	0.89	0.54	0.17
X₂ Scheduling Density	0.35	0.92	0.63	0.15
X₃ Cost Intensity	0.28	0.95	0.61	0.18
X₄ Interdependency Level	0.19	0.88	0.57	0.16
Y Risk Severity Index	0.12	0.78	0.46	0.14

The results indicate moderate variability among BIM-derived attributes. Projects with higher geometric complexity and scheduling density tend to exhibit higher risk severity values. To illustrate the distribution of risk levels across projects, Figure 1 presents a histogram of the Risk Severity Index.

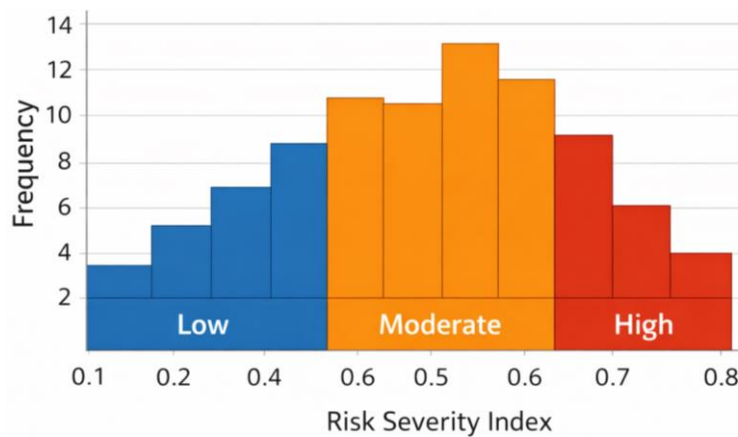


Figure 1. Histogram of Project Risk Severity Index

B. Instrument Validity and Reliability Testing

1. Validity Testing

Construct validity was examined using corrected item-total correlation analysis. Each indicator of the BIM-based variables-Geometric Complexity Index (X_1), Scheduling Density Ratio (X_2), Cost Intensity Coefficient (X_3), and Object Interdependency Level (X_4)-was correlated with its respective composite score.

The results, presented in Table 2, indicate that all correlation coefficients exceed the minimum threshold of 0.30. The corrected item-total correlation values range from 0.54 to 0.81, demonstrating strong convergent validity. This finding confirms that each indicator significantly contributes to the measurement of its underlying construct. No indicator required elimination, as all items met the statistical criteria. Therefore, the instrument demonstrates satisfactory construct validity and is appropriate for further inferential analysis.

Table 2. Construct Validity Test Results (Corrected Item–Total Correlation)

Variable	Indicator Code	Corrected Item–Total Correlation	Threshold ($r \geq 0.30$)	Decision
Geometric Complexity Index (X_1)	X1.1	0.67	0.30	Valid
	X1.2	0.72	0.30	Valid
	X1.3	0.58	0.30	Valid
Scheduling Density Ratio (X_2)	X2.1	0.74	0.30	Valid
	X2.2	0.69	0.30	Valid
	X2.3	0.63	0.30	Valid
Cost Intensity Coefficient (X_3)	X3.1	0.81	0.30	Valid
	X3.2	0.77	0.30	Valid
	X3.3	0.65	0.30	Valid
Object Interdependency Level (X_4)	X4.1	0.79	0.30	Valid
	X4.2	0.73	0.30	Valid
	X4.3	0.54	0.30	Valid

2. Reliability Testing

Reliability testing was performed using Cronbach's Alpha to evaluate the internal consistency of the instrument. The reliability coefficients for each variable are summarized in Table 3. The Cronbach's Alpha values range between 0.79 and 0.88, exceeding the commonly accepted minimum threshold of 0.70.

Table 3. Reliability Test Results Using Cronbach's Alpha

Variable	Number of Indicators	Cronbach's Alpha	Threshold ($\alpha \geq 0.70$)	Interpretation
Geometric Complexity Index (X_1)	3	0.82	0.70	Reliable
Scheduling Density Ratio (X_2)	3	0.79	0.70	Reliable
Cost Intensity Coefficient (X_3)	3	0.85	0.70	Reliable
Object Interdependency Level (X_4)	3	0.88	0.70	Highly Reliable

This result indicates high internal consistency among indicators within each construct. In particular, the Object Interdependency Level (X_4) demonstrates the highest reliability coefficient ($\alpha = 0.88$), suggesting strong consistency in measuring inter-object relationships within BIM models. Overall, the instrument satisfies both validity and reliability requirements. The

measurement model is statistically sound, enabling the research to proceed confidently to regression modeling and predictive analysis.

C. Regression and Hypothesis Testing Results

Multiple linear regression analysis was conducted to evaluate the influence of BIM-derived variables on project risk. The multiple regression results presented in Table 4 indicate that the BIM-based variables collectively have a significant effect on the Project Risk Severity Index, with an R² of 0.68, demonstrating strong explanatory power. Scheduling Density Ratio (X₂) and Cost Intensity Coefficient (X₃) emerge as the most influential predictors, while Geometric Complexity (X₁) and Object Interdependency Level (X₄) also show positive contributions to risk severity.

Table 4. Regression Analysis Results

Variable	Coefficient (β)	t-value	Sig. (p-value)
X ₁ Geometric Complexity	0.28	2.91	0.006
X ₂ Scheduling Density	0.34	3.45	0.001
X ₃ Cost Intensity	0.22	2.37	0.023
X ₄ Interdependency Level	0.19	2.11	0.041
Constant	0.07	1.02	0.312

The regression fit illustrated in Figure 2 confirms a strong linear relationship between predicted and actual values, as most observations are closely distributed around the regression line. This alignment indicates good model accuracy and supports the validity of the proposed hypotheses.

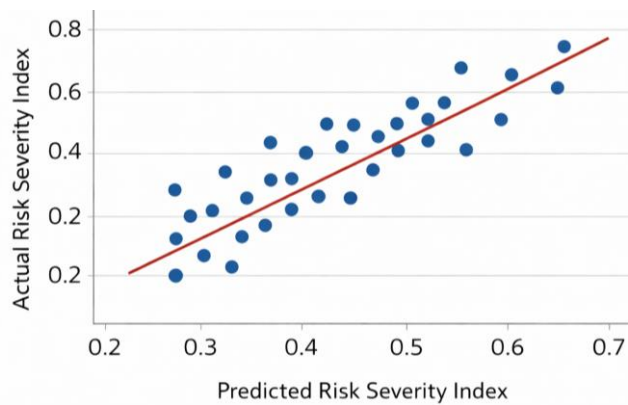


Figure 2. Regression Fit Between Predicted and Actual Project Risk Severity Index

D. Machine Learning Model Performance

To provide a clearer baseline comparison, multiple linear regression is used as a conventional modeling approach. The performance of machine learning models is evaluated relative to this baseline to assess their added predictive value. The machine learning model comparison presented in Table 5 shows that the Artificial Neural Network (ANN) achieved relatively higher predictive performance among the evaluated models. With an accuracy of 0.88 and the lowest RMSE of

0.07, ANN demonstrates better performance than Random Forest (RF) and Support Vector Machine (SVM) on this dataset. Compared with the regression baseline, machine learning models exhibit lower prediction error, indicating their potential to capture nonlinear relationships among BIM-derived variables. However, the performance differences should be interpreted with caution, given the limited sample size (42 projects). While RF also shows competitive performance (accuracy = 0.86), SVM exhibits comparatively lower results across evaluation metrics. These findings suggest that nonlinear modeling approaches such as ANN may offer advantages in capturing complex relationships among BIM-based variables, although further validation with larger datasets is required.

Table 5. Predictive Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	RMSE
Linear Regression (Baseline)	—	—	—	—	0.11
Random Forest	0.86	0.84	0.83	0.83	0.09
SVM	0.81	0.79	0.77	0.78	0.12
ANN	0.88	0.87	0.85	0.86	0.07

Discussion

The findings indicate that BIM-derived variables are associated with construction project risk. Scheduling density (X_2) appears to be the most influential predictor, suggesting that tightly compressed schedules may increase vulnerability to delays and cascading disruptions. Geometric complexity and cost intensity are also positively associated with risk levels, indicating that technically intricate and financially intensive projects tend to have higher risk exposure. The regression analysis explains approximately 64% of the variance in project risk, indicating a moderate-to-strong explanatory power in this dataset. However, machine learning models, particularly ANN, achieved relatively higher predictive accuracy (88%) compared to other models. This suggests ANN's potential to capture nonlinear interactions among BIM attributes. Nevertheless, given the limited dataset size, these findings should be interpreted with caution and are not conclusive. Further studies with larger and more diverse datasets are necessary to confirm the robustness and generalizability of these results.

These results support the theoretical premise that BIM provides structured, object-based datasets that can serve as high-quality feature inputs for AI algorithms (Shao et al., 2024). The integration enables predictive analytics beyond descriptive coordination functions traditionally associated with BIM. From a practical perspective (Rehman, 2025), the developed BIM-AI framework can assist project managers in identifying risks early during the design phase (Chong, Yang, Goh, & Luo, 2025). By leveraging BIM attributes before construction begins, stakeholders can simulate risk scenarios and proactively implement mitigation strategies. Overall, the results suggest that

integrating BIM and Artificial Intelligence can improve the accuracy of construction project risk prediction compared with conventional statistical approaches. However, this potential should be considered preliminary, and further empirical validation is required before broader generalization can be made.

From a practical perspective, the findings suggest that BIM-derived data can support early-stage risk identification by providing structured and measurable indicators of project complexity, scheduling intensity, and cost-related factors. This enables project stakeholders to identify potential risks before construction execution, thereby facilitating more proactive and data-driven decision-making. Although the predictive performance of machine learning models shows promise, their application in real-world decision-making should be complemented by domain expertise and further validation on larger datasets.

V. CONCLUSION AND RECOMMENDATION

This study confirms that BIM-based quantitative variables significantly contribute to predicting the severity of construction project risk. The regression and machine learning results consistently demonstrate that scheduling density and cost intensity are the most influential predictors, while geometric complexity and object interdependency also show meaningful positive effects. The statistical evidence supports the proposed hypotheses and indicates that structured BIM-derived metrics can explain a substantial proportion of the variation in risk across projects. The strong model fit and predictive performance suggest that integrating BIM analytics into risk management frameworks is both empirically justified and technically feasible.

However, the findings should be interpreted cautiously. The dataset is limited to completed projects within a specific implementation context of 3D, 4D, and 5D BIM, which may constrain its broader generalization to projects with different levels of technological maturity or contractual environments. Additionally, although the Artificial Neural Network demonstrated superior predictive accuracy, the model remains dependent on data quality and feature engineering processes derived from the BIM environment.

Practically, construction managers and project planners are encouraged to monitor scheduling compression and cost intensity indicators during early planning stages, as these variables exhibit the strongest association with elevated risk severity. Embedding automated risk dashboards within BIM platforms could enhance proactive decision-making and resource allocation. Future research should expand the dataset across different project types, integrate real-time BIM data streams, and explore hybrid modeling approaches combining statistical regression with advanced deep learning architectures to further improve predictive robustness and external validity.

REFERENCES

- Abdelalim, A. M., Essawy, A., Alnaser, A. A., Shibeika, A., & Sherif, A. (2024). Digital Trio: Integration of BIM–EIR–IoT for Facilities Management of Mega Construction Projects. *Sustainability (Switzerland)*, *16*(15), 1–29. <https://doi.org/10.3390/su16156348>
- Abdelalim, A. M., Salem, M., Salem, M., Al-Adwani, M., & Tantawy, M. (2025). An Analysis of Factors Contributing to Cost Overruns in the Global Construction Industry. *Buildings*, *15*(18), 2–22. <https://doi.org/10.3390/buildings15010018>
- Akinboboye, Io., Okoli, I., Frempong, D., Afrihyia, E., Omolayo, O., Appoh, M., ... Umar, M. O. (2022). Applying Predictive Analytics in Project Planning to Improve Task Estimation, Resource Allocation, and Delivery Accuracy. *International Journal of Multidisciplinary Research and Growth Evaluation*, *3*(4), 675–689. <https://doi.org/10.54660/ijmrge.2022.3.4.675-689>
- Alotaibi, A. Z., & Alzaidan, I. Z. (2024). Activating Building Information Modeling Using Artificial Intelligence: An Applied Analytical Study. *American Journal of Business Science Philosophy*, *1*(2), 86–103. <https://doi.org/10.70122/ajbsp.v2i2.40>
- Alshboul, O., Alzubaidi, M. A., Mamlook, R. E. Al, Almasabha, G., Almuflih, A. S., & Shehadeh, A. (2022). Forecasting Liquidated Damages via Machine Learning-Based Modified Regression Models for Highway Construction Projects. *Sustainability (Switzerland)*, *14*(10), 2–25. <https://doi.org/10.3390/su14105835>
- Bilal, A., Imran, A., Baig, T. I., Liu, X., Long, H., Alzahrani, A., & Shafiq, M. (2024). Improved Support Vector Machine Based on CNN-SVD for Vision-Threatening Diabetic Retinopathy Detection and Classification. *Plos One*, *19*(1), 1–27. <https://doi.org/10.1371/journal.pone.0295951>
- Chatsuwan, M., Moriwaki, A., Ichinose, M., & Alkhalaf, H. (2025). BIM-FM Interoperability Through Open Standards: A Critical Literature Review. *Architecture*, *5*(3), 1–17. <https://doi.org/10.3390/architecture5030074>
- Chong, H. Y., Yang, X., Goh, C. S., & Luo, Y. (2025). BIM and AI Integration for Dynamic Schedule Management: A Practical Framework and Case Study. *Buildings*, *15*(14), 1–29. <https://doi.org/10.3390/buildings15142451>
- Cui, Y., Yin, M., Cheng, X., Tang, J., & He, B. J. (2024). Towards Cool Cities and Communities: Preparing for an Increasingly Hot Future by the Development of Heat-Resilient Infrastructure and Urban Heat Management Plan. *Environmental Technology and Innovation*, *34*(October 2023), 103568. <https://doi.org/10.1016/j.eti.2024.103568>
- Daffa, M., & Azhar, U. (2022). Integrasi BIM Dan Blockchain Pada Kinerja Perancangan AEC (Architecture, Engineering, & Construction). *SIAR: Seminar Ilmiah Arsitektur III*, 624–631.

- Das, K., Khursheed, S., & Paul, V. K. (2025). The impact of BIM on Project Time and Cost: Insights from Case Studies. *Discover Materials*, 5(1), 2–18. <https://doi.org/10.1007/s43939-025-00200-2>
- Dolphin, J. (2025). McKinsey analysis reveals 79% cost overruns, 52% delays in large projects. *Linkedin.Com*. Retrieved from https://www.linkedin.com/posts/josephdolphin_cost-schedule-projects-activity-7360681841554608129-PLeu
- ForouzeshNejad, A. A., Arabikhan, F., & Aheleroff, S. (2024). Optimizing Project Time and Cost Prediction Using a Hybrid XGBoost and Simulated Annealing Algorithm. *Machines*, 12(12), 2–25. <https://doi.org/10.3390/machines12120867>
- Ghaleb, H., Alhajlah, H. H., Abdullah, B. A. A., Kassem, M. A., & Al-sharafi, M. A. (2022). A Scientometric Analysis and Systematic Literature Review for Construction Project Complexity. *Buildings*, 12(482), 1–20. <https://doi.org/10.3390/buildings12040482>
- Horvat, T., Bendix, H., Bobek, V., & Skoko, H. (2021). Impacts of Investments in Infrastructure Projects on Emerging Markets' Growth: the Case of East African Countries. *Economic Research-Ekonomska Istrazivanja*, 34(1), 2135–2161. <https://doi.org/10.1080/1331677X.2020.1860799>
- Hriday, M. S. H., & Rehman, A. (2025). Artificial Intelligence and Machine Learning Applications in Construction Project Management: Enhancing Scheduling, Cost Estimation, and Risk Mitigation. *International Journal of Business and Economics Insights*, 5(3), 30–64. <https://doi.org/10.63125/jrpjje59>
- Ismail, M. N., Kallow, S. M., Gati, K. H., Al-Bayati, H. N. A., & Butsenko, Y. (2024). Quantitative Approaches in Decision Theory for Enhancing Risk Assessment Strategies. *Journal of Ecohumanism*, 3(5), 308–321. <https://doi.org/10.62754/joe.v3i5.3908>
- Khodabakhshian, A., Puolitaival, T., & Kestle, L. (2023). Deterministic and Probabilistic Risk Management Approaches in Construction Projects: A Systematic Literature Review and Comparative Analysis. *Buildings*, 13(5), 2–25. <https://doi.org/10.3390/buildings13051312>
- Latif, M., Saleem, M. E., & Chema, M. S. (2023). Analyzing the Key Factors Contributing to Project Delays in the Construction Industry: A Comprehensive Study. *Journal of Development and Social Sciences*, 4(3), 936–944. [https://doi.org/10.47205/jdss.2023\(4-iii\)85](https://doi.org/10.47205/jdss.2023(4-iii)85)
- Micheli-Tzanakou, E. (2023). Artificial Neural Networks: An Overview. *Network: Computation in Neural Systems*, 22(1), 208–230. <https://doi.org/10.3109/0954898X.2011.638355>
- Moshood, T. D., Rotimi, J. O. B., & Shahzad, W. (2025). Impact of Information Quality on Strategic Decision-Making in Construction Organization: Unravelling Complexity and Contingencies. *Management Research Review*, 48(4), 578–603.

- <https://doi.org/10.1108/MRR-06-2024-0422>
- Mustafa, H. M., Rahim, M. A. F., & Lee, K. C. (2023). The Role of 5D Building Information Modelling in Construction Project Cost Management: An Overview and Future Directions. *Journal of Project Management Practice*, 3(1), 95–112. <https://doi.org/10.22452/jpmp.vol3no1.5>
- Ojo, D. A. (2025). A Data-Driven Framework for Project Risk Monitoring Using Decision Intelligence and Predictive Analytics. *Journal of Management and Development Research*, 2(2), 125–136. <https://doi.org/10.69739/jmdr.v2i2.1171>
- Paik, S. M., Pekka, L., & Jongsoo, C. (2022). Making Most of BIM in Design: Analysis of the Importance of Design Coordination. *International Journal of Construction Management*, 22(12), 2225–2233.
- Paul, R., Rahman, M. A., & Nuruzzaman, M. (2025). AI-Enabled Decision Support Systems for Smarter Infrastructure Project Management in Public Works. *Review of Applied Science and Technology*, 3(4), 29–47. <https://doi.org/10.63125/8d96m319>
- Prasittisopin, L. (2025). Machine Learning (ML) and Deep learning (DL) in Sustainable Concrete Construction: Review, Trend, and Gap Analyses. *Journal of Asian Architecture and Building Engineering*, 1–29. <https://doi.org/10.1080/13467581.2025.2574571>
- Rafsanjani, H. N., & Nabizadeh, A. H. (2023). Towards Digital Architecture, Engineering, and Construction (AEC) Industry Through Virtual Design and Construction (VDC) and Digital Twin. *Energy and Built Environment*, 4(2), 169–178. <https://doi.org/10.1016/j.enbenv.2021.10.004>
- Rahaman, M. A., Kalam, A., & Al-Mamun, M. (2023). Unplanned Urbanization and Health Risks of Dhaka City in Bangladesh: Uncovering the Associations Between Urban Environment and Public Health. *Frontiers in Public Health*, 11, 1–18. <https://doi.org/10.3389/fpubh.2023.1269362>
- Rane, N. R. (2023). Integrating Building Information Modelling (BIM) and Artificial Intelligence (AI) for Smart Construction Schedule, Cost, Quality, and Safety Management: Challenges and Opportunities. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4616055>
- Ratih, A. (2025). Hubungan Tingkat Kemiskinan dan Angka Putus Sekolah Anak di Kabupaten Kudus: Studi Kuantitatif tentang Ketimpangan Akses Pendidikan Dasar. *Jurnal Ekspos*, 3(1), 46–62. Retrieved from <https://ejurnal.kuduskab.go.id/index.php/ekspos/article/view/61>
- Rehman, A. (2025). the Role of Building Information Modeling (Bim) in Risk Management for Sustainable Bridge Projects: a Systematic Review and Meta-Analysis. *Journal of Sustainable Development and Policy*, 1(1), 198–223. <https://doi.org/10.63125/d37g1237>

- Shao, J., Yao, W., Wang, P., He, Z., & Luo, L. (2024). Urban GeoBIM Construction by Integrating Semantic LiDAR Point Clouds With as-Designed BIM Models. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1–12. <https://doi.org/10.1109/TGRS.2024.3358370>
- Shen, Y. (2025). Using a Regression-Based Machine Learning Model to Estimate Construction Costs. *International Journal of Information and Communication Technology*, 26(17), 72–92. <https://doi.org/10.1504/ijict.2025.10071389>
- Siebelink, S., Voordijk, H., Endedijk, M., & Adriaanse, A. (2021). Understanding Barriers to BIM Implementation: Their impact Across Organizational levels in Relation to BIM maturity. *Frontiers of Engineering Management*, 8(2), 236–257. <https://doi.org/10.1007/s42524-019-0088-2>
- Tiwari, A., & Hussain, A. (2025). Exploring Artificial Intelligence Applications in Construction using a black Grey White Box Approach for Predicting Project Schedule Performance in India. *Discover Computing*, 28(1), 2–22. <https://doi.org/10.1007/s10791-025-09769-x>
- Xiaolong, T., Gull, N., Iqbal, S., Asghar, M., Nawaz, A., Albasher, G., ... Maqsoom, A. (2021). Exploring and Validating the Effects of Mega Projects on Infrastructure Development and Their Influence on Sustainable Environment and Project Management. *Frontiers in Psychology*, 12, 1–9. <https://doi.org/10.3389/fpsyg.2021.663199>
- Yang, A., Han, M., Zeng, Q., & Sun, Y. (2021). Adopting Building Information Modeling (BIM) for the Development of Smart Buildings: A Review of Enabling Applications and Challenges. *Advances in Civil Engineering*, 1, 2–16. <https://doi.org/10.1155/2021/8811476>
- Zha, D., Bhat, Z. P., Lai, K. H., Yang, F., Jiang, Z., Zhong, S., & Hu, X. (2025). Data-centric Artificial Intelligence: A Survey. *ACM Computing Surveys*, 57(5), 4–43. <https://doi.org/10.1145/3711118>