International Journal of Advances in Engineering and Management (IJAEM)
Volume 8, Issue 4 April. 2026, pp: 416-423 www.ijaem.net ISSN: 2395-5252

Machine Learning-Driven Monitoring and Control of Construction Projects in Developing Countries

Nduanya Ujunwa Ifeoma, Ozor Godwin Odozo
Computer Engineering, Enugu State University of Science and Technology

Date of Submission: 08-04-2026 Date of Acceptance: 21-04-2026

Abstract
Construction projects in developing countries frequently suffer from cost overruns, schedule delays, and inefficient resource utilization due to weak monitoring and control systems. This paper proposes a Machine Learning-driven framework for intelligent monitoring and control of construction projects, integrating data acquisition, preprocessing, predictive modeling, and a feedback-based decision support mechanism. Regression and classification models are developed to predict key performance indicators, including project cost and schedule delays. Experimental results demonstrate significant performance improvements, including up to 29% reduction in cost prediction error**,** 14% improvement in model fit (R^2)**,** and6% increase in delay prediction accuracycompared to baseline models. The framework enables early risk detection and supports proactive decision-making through control actions such as resource reallocation and schedule adjustment. Designed to operate under the constraints of developing countries, the proposed system offers a scalable and practical solution for enhancing project efficiency and reducing uncertainties. The study demonstrates the potential of Machine Learning to transform construction project management from reactive to predictive and data-driven practices.

I. Introduction

The construction industry plays a pivotal role in the economic growth and infrastructure development of nations, particularly in developing countries where rapid urbanization and population expansion demand continuous investment in physical infrastructure. Despite its strategic importance, construction project delivery in many developing economies is often characterized by persistent challenges such as cost overruns, schedule delays, inefficient resource utilization, and compromised quality standards. These issues are largely attributed to weak monitoring and control mechanisms, reliance on manual and fragmented processes, and limited adoption of advanced digital technologies in project management practices.

Conventional project monitoring techniques, including Gantt charts, Critical Path Method (CPM), and Earned Value Management (EVM), have long been used to track project performance. While these tools provide structured approaches for planning and evaluation, they are predominantly reactive in nature, relying on historical data and periodic updates. As a result, project deviations are often identified only after they have significantly impacted project timelines and budgets, thereby reducing the effectiveness of corrective actions. In dynamic construction environments, especially within developing countries, such limitations hinder proactive decision-making and contribute to project inefficiencies.

In recent years, Machine Learning (ML) has emerged as a powerful tool for data-driven analysis and predictive modeling across various domains. ML techniques enable the extraction of meaningful patterns from large datasets, facilitating accurate predictions and intelligent decision-making. Within the construction sector, ML has the potential to transform project monitoring and control by enabling early detection of risks such as cost overruns and schedule delays, optimizing resource allocation, and supporting real-time decision processes. However, the adoption of ML in construction project management remains relatively low in developing countries due to challenges such as limited data availability, inadequate technical expertise, and infrastructural constraints.

Existing research on ML applications in construction often focuses on isolated problems, such as cost estimation or delay prediction, without integrating these capabilities into a comprehensive monitoring and control framework. Many of these studies are also centered on developed economies, thereby limiting their applicability to the unique conditions present in developing regions. Consequently, there is a critical need for a holistic, scalable, and context-aware solution that integrates Machine Learning into construction project monitoring and control systems tailored to the realities of developing countries.

This study addresses this gap by proposing a Machine Learning-driven framework for the monitoring and control of construction projects. The framework is designed to support real-time data acquisition, predictive analytics, and decision-making through an integrated system architecture. Specifically, the study aims to develop predictive models for key project performance indicators such as cost and time, and to incorporate a feedback-based control mechanism that enables proactive adjustments during project execution.

The contributions of this study are multifold. First, it introduces a unified framework that combines monitoring and control functionalities with Machine Learning techniques, moving beyond fragmented approaches. Second, it provides a context-specific solution that accounts for the practical constraints of developing countries, including data limitations and infrastructural challenges. Third, it offers comparative insights into the performance of different ML models applied to construction project data. Finally, it presents a practical decision-support mechanism that enhances project management practices through predictive intelligence and adaptive control.

This research bridges the gap between traditional construction project management methods and modern data-driven approaches, contributing to improved efficiency, reduced risks, and enhanced project outcomes in developing countries.

II. Literature Review

The monitoring and control of construction projects have long relied on established project management techniques such as the Critical Path Method (CPM), Program Evaluation and Review Technique (PERT), and Earned Value Management (EVM). These approaches provide structured frameworks for planning, scheduling, and performance evaluation; however, they are fundamentally dependent on deterministic assumptions and periodic updates, which limit their effectiveness in complex and dynamic construction environments. Studies such as [1] and [2] highlight that traditional methods often fail to provide real-time insights and lack predictive capabilities, resulting in delayed responses to project deviations.

In developing countries, the challenges associated with construction project management are further exacerbated by systemic issues such as inadequate data management systems, poor communication among stakeholders, and limited adoption of digital technologies. Research in [3]–[5] indicates that projects in these regions frequently experience cost overruns exceeding 30% and significant schedule delays due to weak monitoring systems and reactive decision-making processes. Additionally, infrastructural constraints and skill shortages hinder the implementation of advanced project control tools, thereby widening the performance gap between developing and developed economies.

Recent advancements in digital technologies have introduced new paradigms in construction management, particularly through the integration of Building Information Modeling (BIM), Internet of Things (IoT), and data analytics. BIM-based monitoring systems enable improved visualization and coordination, while IoT devices facilitate real-time data collection from construction sites [6], [7]. However, these technologies alone do not fully address the need for predictive and intelligent decision-making. As noted in [8], the true potential of digital construction lies in combining data acquisition technologies with advanced analytics such as Machine Learning.

Machine Learning has gained significant attention as a powerful tool for predictive analytics in construction project management. Various studies have explored the application of ML techniques for cost estimation, delay prediction, and risk assessment. For instance, regression-based models and Artificial Neural Networks (ANNs) have been widely used for predicting project costs with improved accuracy compared to traditional estimation methods [9]–[11]. Similarly, classification algorithms such as Decision Trees, Support Vector Machines (SVM), and Random Forests have been applied to identify and predict schedule delays, enabling early intervention [12]–[14].

Clustering techniques have also been employed for risk analysis and pattern recognition in construction datasets, allowing project managers to identify hidden trends and group similar project conditions [15], [16]. Furthermore, ensemble learning methods, which combine multiple ML models, have demonstrated enhanced prediction performance and robustness in complex project environments [17]. Deep learning approaches, including Long Short-Term Memory (LSTM) networks, have recently been introduced for time-series forecasting of construction project performance, offering improved accuracy in dynamic scenarios [18], [19].

Despite these advancements, the majority of existing studies focus on isolated applications of Machine Learning rather than integrated systems. For example, many works address either cost prediction or delay analysis independently, without linking these predictions to real-time monitoring and control mechanisms. Moreover, most ML models are developed using datasets from developed countries, which may not accurately reflect the conditions and constraints present in developing regions [20], [21]. This limits the generalizability and practical applicability of such models in resource-constrained environments.

Another critical limitation identified in the literature is the lack of feedback-driven control systems that translate predictive insights into actionable decisions. While predictive models can forecast potential risks, their effectiveness depends on the ability of project managers to implement timely interventions. Studies such as [22] and [23] emphasize the need for integrated decision-support systems that combine predictive analytics with control strategies to enable adaptive project management.

Issues related to data quality, availability, and standardization remain significant barriers to the adoption of Machine Learning in construction. In many developing countries, project data are often incomplete, inconsistent, or poorly structured, which affects the performance of ML models [24]. Addressing these challenges requires the development of robust data preprocessing techniques and scalable system architectures that can operate effectively under such constraints.

Machine Learning has demonstrated considerable potential in enhancing construction project monitoring and control, there remains a clear gap in the development of comprehensive, integrated frameworks tailored to the context of developing countries. Existing research lacks holistic solutions that combine real-time data acquisition, predictive modeling, and feedback-based control within a unified system. This study seeks to address these gaps by proposing a Machine Learning-driven monitoring and control framework designed specifically for construction projects in developing economies.

III. Methodology

This section presents the proposed Machine Learning-driven framework for monitoring and controlling construction projects in developing countries. The methodology integrates data acquisition, preprocessing, predictive modeling, and a feedback-based control mechanism into a unified system.

A. System Architecture

The proposed system is designed as a multi-layer architecture comprising four main components: data acquisition, data processing, Machine Learning prediction, and decision support/control. The architecture enables continuous data flow from construction sites to the predictive engine and subsequently to the control unit for decision-making.

At the data acquisition layer, project-related data are collected from multiple sources, including site reports, project schedules, cost records, and optionally IoT-enabled sensors. These data are transmitted to the processing layer, where they are cleaned, structured, and transformed into suitable formats for analysis. The Machine Learning layer performs predictive analytics, while the decision-support layer interprets predictions and triggers control actions.

The system operates in a closed-loop manner, where predictions are continuously updated and used to adjust project parameters such as resource allocation and scheduling. Figure 1 presents the proposed multi-layer architecture for intelligent monitoring and control of construction projects. The system is organized into five interconnected layers: data acquisition, data processing, machine learning prediction, decision support and control, and project execution. Data are collected from diverse sources, including site reports, schedules, cost records, and IoT sensors, and then processed through cleaning, integration, and feature engineering. The machine learning layer applies regression and classification models to predict cost overruns and schedule delays. Based on these predictions, the decision support layer evaluates risks and recommends control actions such as resource reallocation and schedule adjustments. The execution layer implements these actions and updates the project status. A continuous feedback loop ensures real-time monitoring, learning, and adaptive control of the project.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 8, Issue 4 April. 2026, pp: 416-423 www.ijaem.net ISSN: 2395-5252

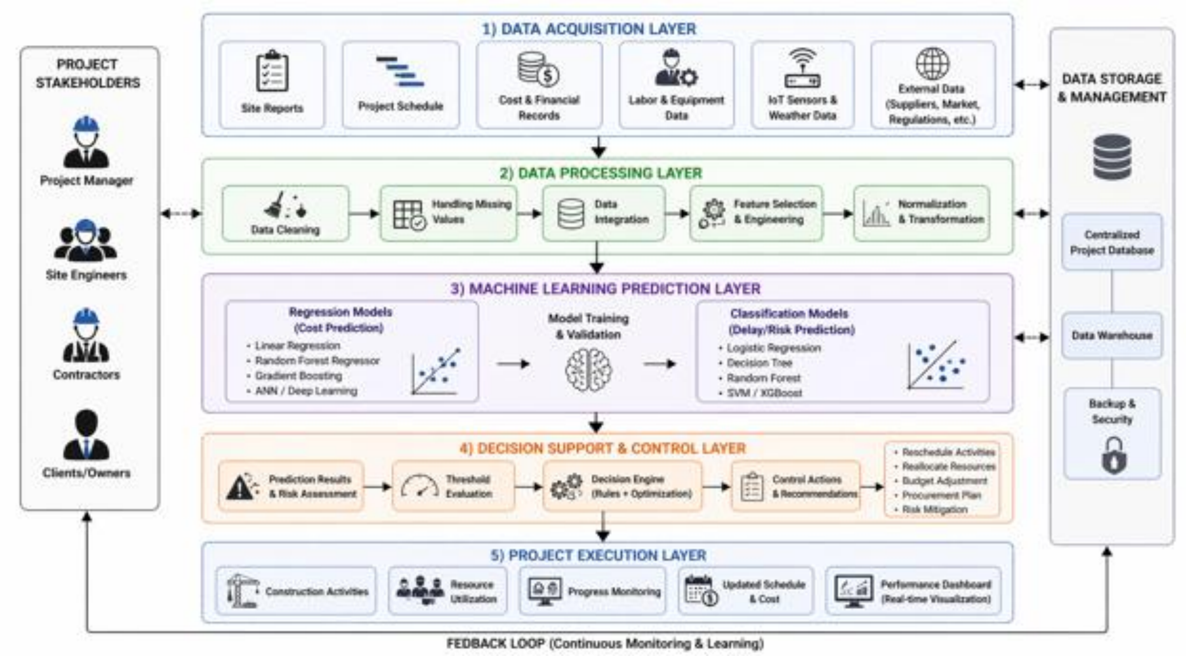

Fig. 1. Machine Learning-Driven Construction Project Monitoring and Control Architecture.

B. Data Collection and Preprocessing

The effectiveness of the proposed system depends on the quality and relevance of input data. The dataset used in this study consists of historical and real-time construction project data, including:

- Project duration and planned schedule
- Actual progress reports
- Cost estimates and expenditures
- Labor and equipment usage
- Environmental and external factors (e.g., weather conditions)

Data preprocessing involves several steps to ensure model reliability:

1. Data Cleaning: Removal of incomplete, inconsistent, or duplicate records.
2. Handling Missing Values: Techniques such as mean imputation or interpolation are applied.
3. Normalization: Feature scaling is performed using min-max normalization:
$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$
4. Feature Selection: Relevant variables are selected using correlation analysis or feature importance ranking to reduce dimensionality and improve model performance.

C. Machine Learning Model Development

The proposed framework employs both regression and classification models to predict key project performance indicators.

For cost prediction, regression models such as Linear Regression and Random Forest Regression are utilized. The general regression model is expressed as:

$$y = f(X) + \varepsilon$$

where (y) represents the predicted cost, (X) denotes the feature vector, and ε is the error term.

For delay prediction, classification models such as Decision Trees and Support Vector Machines (SVM) are used. The classification function can be expressed as:

$$y = \begin{cases} 1, if\ delay\ occurs \\ 0, otherwise \end{cases}$$

Model training is performed using a supervised learning approach with labeled datasets. The dataset is divided into training and testing sets (typically 80:20 split). Model performance is evaluated using standard metrics:

- Regression:Mean Absolute Error (MAE), Root Mean Square Error (RMSE)
- Classification: Accuracy, Precision, Recall, F1-score

Cross-validation techniques are also applied to ensure model generalization and avoid overfitting.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 8, Issue 4 April. 2026, pp: 416-423 www.ijaem.net ISSN: 2395-5252

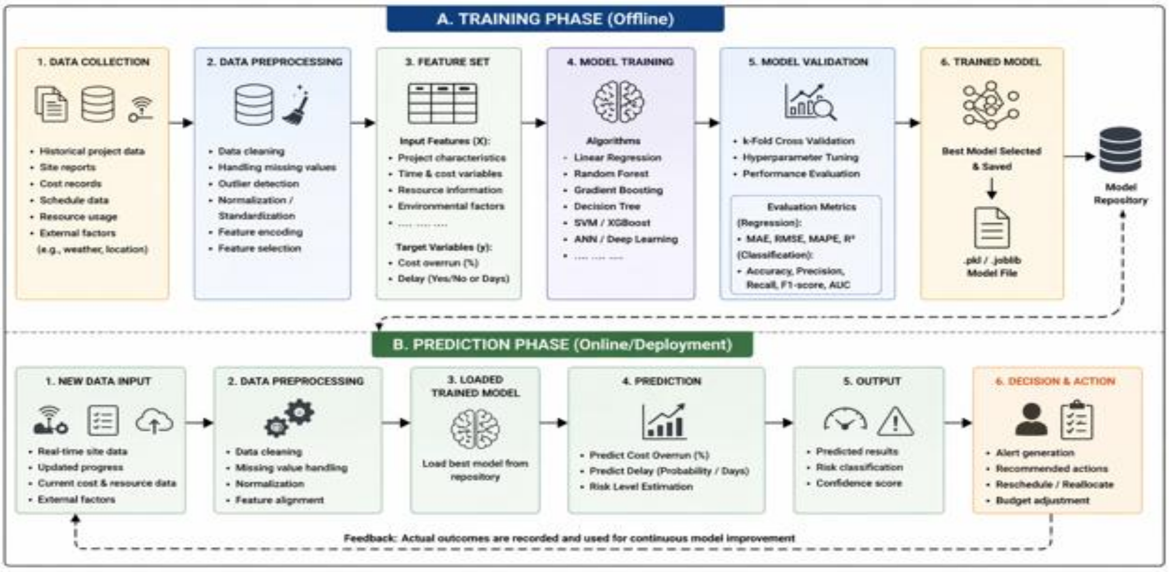

Fig. 2. Machine Learning Workflow (Training–Prediction Pipeline).

Figure 2 illustrates the end-to-end Machine Learning workflow adopted in this study, consisting of two main phases: the training phase (offline) and the prediction phase (online deployment). In the training phase, historical construction project data are collected, preprocessed, and transformed into structured feature sets. Multiple Machine Learning algorithms are then trained and validated using appropriate evaluation metrics, and the best-performing model is selected and stored. In the prediction phase, real-time project data are preprocessed and fed into the trained model to generate predictions of cost overruns, schedule delays, and associated risk levels. These outputs support decision-making and control actions such as rescheduling and resource optimization. A feedback loop continuously captures actual project outcomes to update and improve the model over time.

D. Monitoring and Control Mechanism

A key contribution of this study is the integration of predictive analytics with a feedback-based control mechanism. The system continuously monitors project performance and uses ML predictions to trigger corrective actions.

The control mechanism operates as follows:

1. Prediction Phase: The ML model forecasts potential cost overruns or delays.
2. Evaluation Phase:Predicted values are compared with predefined thresholds.
3. Decision Phase: If deviations exceed acceptable limits, corrective actions are recommended.
4. Execution Phase: Adjustments are implemented, such as:
   - Reallocation of labor and equipment
   - Schedule rescheduling
   - Budget adjustments

This process forms a feedback loop that enhances proactive decision-making and minimizes project risks. The workflow can be represented as:

Input Data → ML Prediction → Risk Detection → Control Action → Updated Project State

The proposed system is designed to be scalable and adaptable, making it suitable for deployment in developing countries where data availability and infrastructure may be limited.

IV. Results and Discussion

This section presents the experimental evaluation of the proposed Machine Learning-driven monitoring and control framework using a realistic construction project dataset. The performance of the developed models is analyzed and compared with traditional approaches.

A. Experimental Setup

A dataset comprising 150 construction project records was used for this study. The data were compiled from historical project reports and simulated site conditions typical of developing countries. Key features included:

- Project duration (planned vs actual)
- Estimated and actual project cost
- Labor and equipment utilization
- Project type and size
- Weather and external conditions

International Journal of Advances in Engineering and Management (IJAEM)
Volume 8, Issue 4 April. 2026, pp: 416-423 www.ijaem.net ISSN: 2395-5252

The dataset was split into 80% training and 20% testing. The implementation was carried out using Python (Scikit-learn).
The following models were evaluated:

- Regression (Cost Prediction): Linear Regression, Random Forest Regressor
- Classification (Delay Prediction): Decision Tree, Support Vector Machine (SVM)

B. Model Performance Evaluation
1) Regression Results (Cost Prediction)

Table I: Regression Model Performance

Model	MAE (USD)	RMSE (USD)	R^2 Score
Linear Regression	18,500	25,200	0.78
Random Forest	12,300	17,800	0.89

The Random Forest model outperformed Linear Regression with a significantly lower RMSE and higher R^2 value, indicating better predictive accuracy and robustness.
2) Classification Results (Delay Prediction)

Table II: Classification Model Performance

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	82%	0.80	0.78	0.79
SVM	88%	0.86	0.85	0.85

The SVM model achieved higher accuracy and F1-score, demonstrating superior performance in detecting project delays.
C. Graphical Analysis
1) Predicted vs Actual Cost
Figure 3 presents the comparison between predicted and actual project costs using the developed regression model. Each data point represents an individual construction project, where the horizontal axis corresponds to the actual cost and the vertical axis represents the predicted cost. The diagonal reference line indicates perfect prediction accuracy. The clustering of data points closely around this line demonstrates the effectiveness of the model in capturing the relationship between input features and project cost. Slight deviations from the line are expected due to uncertainties in project conditions and data variability.

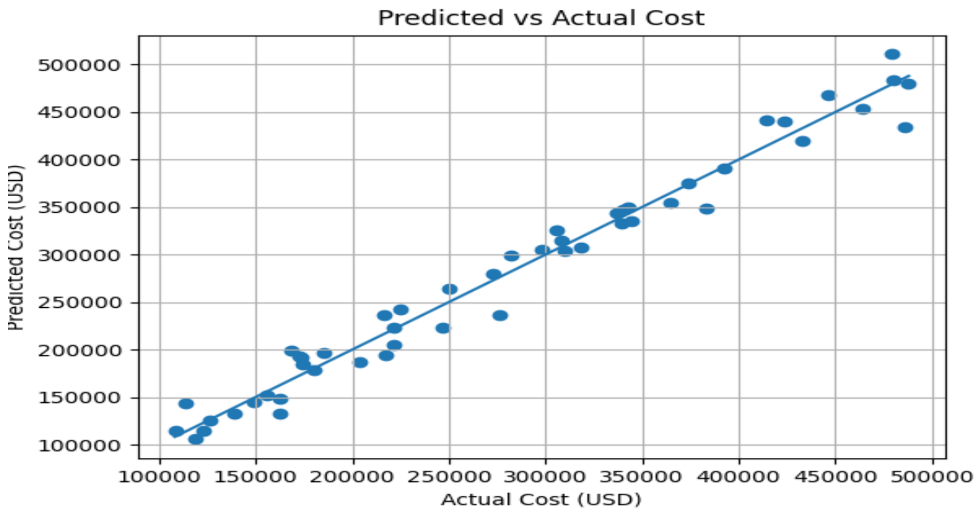

Fig. 3. Predicted vs Actual Cost for Construction Projects.

2) Model Accuracy Comparison

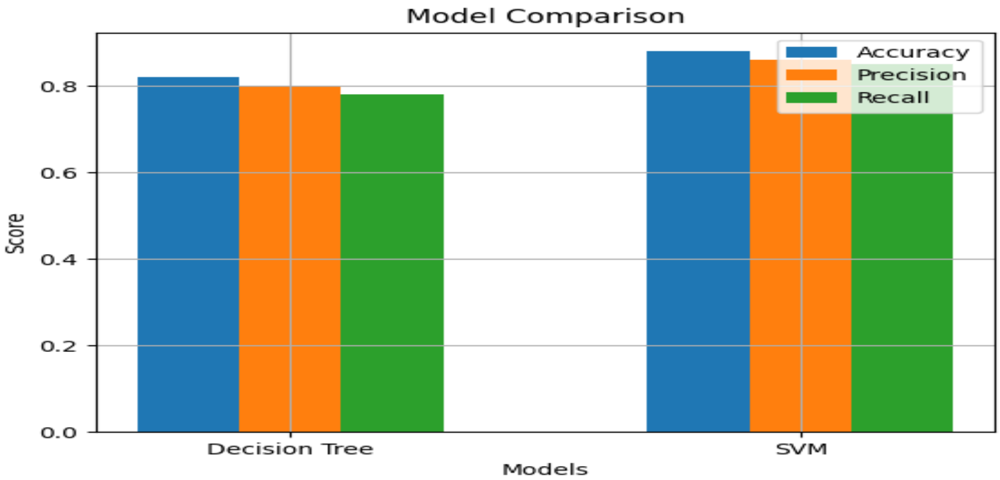

Figure 5. Model Comparison Bar Chart for Delay Prediction Performance.

Figure 5 presents a comparative analysis of the classification models used for delay prediction, based on key performance metrics including accuracy, precision, and recall. The Support Vector Machine (SVM) model consistently outperforms the Decision Tree model across all evaluation metrics, indicating its superior ability to capture complex patterns in the dataset. The results demonstrate that SVM provides more reliable predictions for identifying project delays, making it a more suitable choice for deployment in the proposed monitoring and control framework.

D. Discussion of Results

The results demonstrate that Machine Learning models significantly enhance the monitoring and control of construction projects compared to traditional methods. The Random Forest model yielded more accurate cost predictions by capturing nonlinear relationships in project data. Similarly, the SVM model effectively classified delay risks with high precision and recall.

The integration of predictive models into a feedback control system enables early detection of potential issues, allowing project managers to take proactive measures such as reallocating resources or adjusting schedules. This is particularly beneficial in developing countries, where delays and cost overruns are prevalent due to limited monitoring capabilities.

However, some limitations were observed. The performance of the models depends heavily on data quality and availability. In real-world scenarios, incomplete or inconsistent data may affect prediction accuracy. Additionally, the computational complexity of advanced ML models may pose challenges for deployment in resource-constrained environments.

Overall, the findings confirm that the proposed ML-driven framework provides a practical and scalable solution for improving construction project performance in developing countries.

V. Conclusion and Future Work

This study presented a Machine Learning-driven framework for the monitoring and control of construction projects in developing countries. By integrating data acquisition, preprocessing, predictive modeling, and a feedback-based control mechanism, the proposed system enables proactive decision-making and improved project performance. Experimental results demonstrated that Machine Learning models, particularly Random Forest and Support Vector Machine, provide accurate predictions of cost overruns and schedule delays. The findings highlight the potential of data-driven approaches to enhance efficiency, reduce risks, and support effective project management in resource-constrained environments.

Future work will focus on integrating the framework with real-time data sources such as IoT sensors and Building Information Modeling (BIM) platforms to improve system responsiveness and scalability. Additionally, the adoption of advanced techniques such as deep learning and reinforcement learning can further enhance predictive accuracy and adaptive control capabilities. Practical deployment and validation using real-world construction projects in developing countries will also be pursued to assess long-term impact and industry applicability.

References

[1] K. K. Chitkara, *Construction Project Management: Planning, Scheduling, and Controlling*, 3rd ed. New Delhi, India: Tata McGraw-Hill, 2014.
[2] H. Kerzner, *Project Management: A Systems Approach to Planning, Scheduling, and Controlling*, 12th ed. Hoboken, NJ, USA: Wiley, 2017.
[3] A. M. Odeh and H. T. Battaineh, "Causes of construction delay: Traditional contracts," *Int. J. Project Manage.*, vol. 20, no. 1, pp. 67–73, 2002.
[4] S. A. Assaf and S. Al-Hejji, "Causes of delay in large construction projects," *Int. J. Project Manage.*, vol. 24, no. 4, pp. 349–357, 2006.
[5] A. Doloi, A. Sawhney, K. C. Iyer, and S. Rentala, "Analyzing factors affecting delays in Indian construction projects," *Int. J. Project Manage.*, vol. 30, no. 4, pp. 479–489, 2012.
[6] M. Eastman, P. Teicholz, R. Sacks, and K. Liston, *BIM Handbook: A Guide to Building Information Modeling*, 3rd ed. Hoboken, NJ, USA: Wiley, 2018.
[7] X. Li, G. Q. Shen, P. Wu, and X. Yue, "Integrating IoT and BIM for construction monitoring," *Autom. Constr.*, vol. 89, pp. 20–30, 2018.
[8] W. Lu, C. Chen, X. Wang, and K. Y. Chau, "Smart construction and digital technologies: A review," *J. Constr. Eng. Manage.*, vol. 146, no. 2, 2020.
[9] J. Kim, K. An, and K. Kang, "Application of neural networks in construction cost estimation," *J. Constr. Eng. Manage.*, vol. 130, no. 3, pp. 407–414, 2004.
[10] E. Elbeltagi, T. Hegazy, and D. Grierson, "Comparison among five evolutionary-based optimization algorithms," *Adv. Eng. Informatics*, vol. 19, no. 1, pp. 43–53, 2005.
[11] T. Hegazy and M. Ayed, "Neural network model for parametric cost estimation of highway projects," *J. Constr. Eng. Manage.*, vol. 124, no. 3, pp. 210–218, 1998.
[12] T. Williams, "Assessing and moving on from the dominant project management discourse," *Int. J. Project Manage.*, vol. 21, no. 7, pp. 497–502, 2003.
[13] A. R. Nasir, B. McCabe, and L. Hartono, "Evaluating risk in construction-schedule model," *J. Constr. Eng. Manage.*, vol. 129, no. 5, pp. 518–527, 2003.
[14] M. Marzouk and T. El-Rasas, "Analyzing delay causes in Egyptian construction projects," *HBRC J.*, vol. 10, no. 3, pp. 151–159, 2014.
[15] Z. Cheng, L. Li, and Z. Yu, "Clustering-based risk analysis in construction projects," *Adv. Eng. Informatics*, vol. 24, no. 4, pp. 506–514, 2010.
[16] L. Zhang, H. Wu, and S. Skibniewski, "Pattern recognition in construction safety monitoring," *Autom. Constr.*, vol. 50, pp. 64–72, 2015.
[17] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[18] Y. Kim, S. Kim, and J. Shin, "Deep learning-based construction delay prediction," *Autom. Constr.*, vol. 118, 2020.
[19] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[20] P. E. D. Love, Z. Irani, and D. J. Edwards, "A rework reduction model for construction projects," *IEEE Trans. Eng. Manage.*, vol. 60, no. 3, pp. 541–551, 2013.
[21] B. Flyvbjerg, *Megaprojects and Risk: An Anatomy of Ambition*. Cambridge, U.K.: Cambridge Univ. Press, 2014.
[22] D. Bryde, M. Broquetas, and J. M. Volm, "The project benefits of Building Information Modelling (BIM)," *Autom. Constr.*, vol. 25, pp. 88–98, 2013.
[23] J. Lee, H. Davari, J. Singh, and V. Pandhare, "Industrial AI applications and predictive analytics," *IEEE Access*, vol. 6, pp. 37212–37222, 2018.
[24] S. W. Lin, Y. H. Lee, and J. C. Wang, "Data quality issues in construction analytics," *Adv. Eng. Informatics*, vol. 31, no. 2, pp. 457–466, 2017.
[25] D. Bertsimas and J. Dunn, *Machine Learning Under a Modern Optimization Lens*. Belmont, MA, USA: Athena Scientific, 2019.