International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 1 (2026)



ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 22, No. 1, 2026
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research

AI-Driven Risk Prediction and Allocation Optimization for Cost Transparency in Public Infrastructure Renovation Projects

Sanket Patel
Sr. Project Manager
B.E. Civil, Sardar Vallabhbhai Patel Institute of Technology
Gujarat University, Gujarat, India
Email: senky9278@gmail.com

***Abstract*—Due to inappropriate risk distribution among stakeholders, public infrastructure rehabilitation projects typically encounter cost overruns, timetable delays, and financial misallocation. Conventional methods of risk assessment lack the predictive intelligence necessary for clear financial planning and instead rely on qualitative judgment. An artificial intelligence-driven approach for integrated risk prediction, allocation analysis, and cost transparency in public infrastructure rehabilitation projects is presented in this study. Financial risk parameters, such as project budget, timing, contract type, contingency allocation, and stakeholder risk sharing, are modeled using publicly accessible statistics about construction delays and cost overruns. Random Forest, XGBoost, and Logistic Regression are examples of machine learning models that are used to forecast the likelihood of cost overruns and delays. To measure the impact of inappropriate risk distribution on total project costs, a Risk Allocation Impact Index is created. Cross-validation experimental assessment shows lower error and better prediction accuracy when compared to baseline models. In order to improve transparency, maximize risk sharing between public and private players, and minimize unforeseen financial losses in remodeling projects, the suggested framework offers a data-driven decision support tool. From a procurement strategy standpoint, this optimization structure can guide contract selection decisions such as Guaranteed Maximum Price (GMP), unit-price, design-build, or construction-management-at-risk models. Aligning allocation weights with stakeholder capability improves financial resilience and reduces downstream claims.**

***Index Terms*—Artificial Intelligence, Risk Prediction, Risk Allocation, Cost Overrun, Public Infrastructure, Machine Learning, Financial Transparency**

I. INTRODUCTION

In order to maintain old infrastructure systems, improve public service delivery, and support economic development, public infrastructure restoration projects are crucial. Around the world, governments spend a lot of money updating urban utilities, healthcare facilities, educational institutions, and transportation networks. Infrastructure renovation projects often experience large cost overruns and timetable delays despite these expenditures [1], [2]. In addition to raising the financial cost, these inefficiencies erode public confidence and jeopardize long-term economic planning. In public renovation projects executed under agency procurement systems, cost overruns often stem from ambiguous risk allocation clauses, scope interpretation conflicts, delayed change-order approvals, and administrative review cycles. These contractual dynamics introduce financial volatility beyond traditional engineering uncertainty, necessitating predictive systems that account for governance structure and administrative risk factors.

Cost overruns are systematic rather than unusual events, as empirical studies regularly show [3]. Large-scale transportation and public infrastructure projects frequently end up costing far more than anticipated. The usual distribution of cost overruns seen in infrastructure projects is seen in Figure 1, emphasizing the exposed financial risk's hefty tail. These distributions show that complex uncertainty patterns cannot be adequately captured by conventional deterministic forecasting techniques.

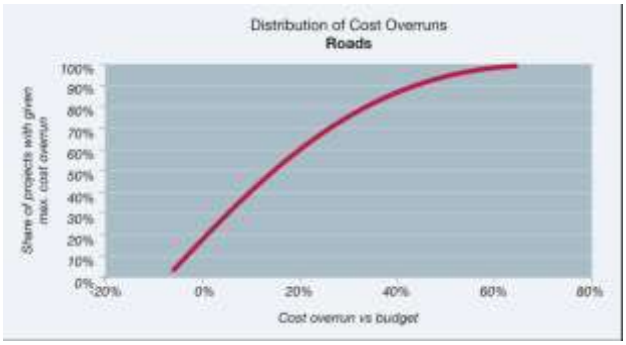


Fig. 1. Typical distribution of cost overruns in infrastructure projects (adapted from empirical studies [1]).

Inadequate risk distribution among stakeholders is one of the main causes of cost growth. Risks associated with refurbishment projects must be divided between public agencies and private contractors. These risks include design modifications, contractor performance variability, regulatory delays, material price variations, and financing uncertainty. Risks should be distributed to the party most capable of managing them, according to research on public-private partnership (PPP) procurement [4], [5]. However, in reality, risk transfer channels are either ineffective or subject to political interference, which results in financial inefficiencies and mismatched incentives.

Conventional risk management methods mostly rely on expert opinion, qualitative evaluation, and historical benchmarking [6]. Even while these approaches offer valuable insights, they are not predictively intelligent and do not account for the nonlinear interactions that exist between project features and financial results. Moreover, these methods hardly ever measure the impact of certain risk distribution choices on the overall project cost or likelihood of delay.

Recent developments in machine learning (ML) and artificial intelligence (AI) have shown great promise in predictive analytics for cost estimates and construction management [7], [8]. When compared to conventional regression models, ensemble learning approaches like Random Forest and gradient boosting models have demonstrated increased accuracy in predicting complex engineering outcomes [9]. Furthermore, model outputs may be interpreted thanks to explainable AI approaches like SHAP, which increases the transparency of predictive systems and makes them appropriate for policy-level decision-making [10].

Despite these advancements, the majority of research to far has been on qualitative risk allocation analysis or cost prediction. Integrated frameworks that concurrently handle risk allocation evaluation, cost overrun prediction, and financial transparency in public rehabilitation projects are still lacking. The creation of a cohesive AI-driven system that can predict financial risk comprehensively is motivated by this research gap.

Figure 2 shows the conceptual framework of risk allocation in infrastructure projects. The distribution of risks among various stakeholders, including as government officials, contractors, financiers, and operators, is shown in the figure. These allocations are frequently more vulnerable to cost overruns and legal issues when they are improperly structured.

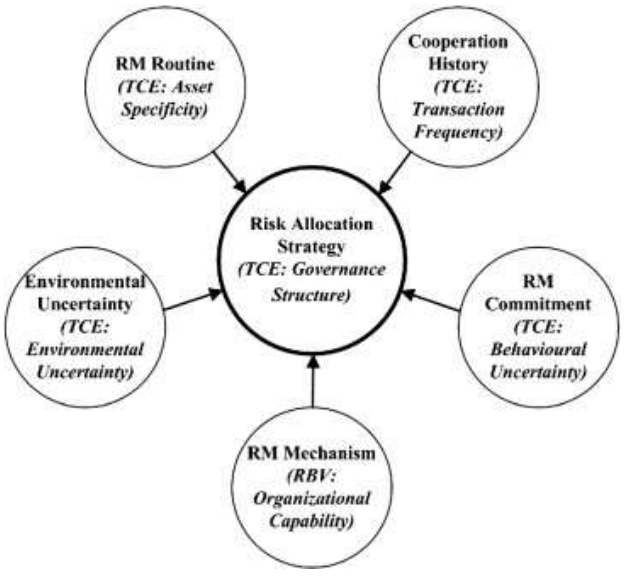


Fig. 2. Conceptual representation of stakeholder risk allocation in infrastructure projects (adapted from PPP risk allocation frameworks [4]).

This paper suggests an AI-driven methodology that combines a quantitative Risk Allocation Impact Index (RAII) with predictive modeling in order to solve these issues. The framework makes use of publicly accessible datasets from construction management databases and infrastructure performance repositories [11], [12]. Baseline budget, projected length, contingency allocation, procurement method, contractor risk sharing, and finance structure are important aspects of the project.

For improved predictive power, the suggested system uses Random Forest and XGBoost models in addition to Logistic Regression as a baseline classifier. To guarantee robustness and generalizability, model performance is assessed using accuracy, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and k-fold cross-validation approaches. To determine the most important factors causing cost escalation, explainability techniques are used.

The contributions of this paper are summarized as follows:

- Development of an integrated Artificial Intelligence framework for cost overrun prediction and risk allocation assessment.
- Formulation of a quantitative Risk Allocation Impact Index to measure the influence of stakeholder risk distribution on financial outcomes.
- Comparative evaluation of machine learning models for predictive accuracy and interpretability.
- Provision of a decision-support mechanism to enhance financial transparency in public infrastructure renovation projects.

The suggested system seeks to enhance procurement tactics, lower unforeseen financial losses, and promote evidence-based policymaking by fusing allocation transparency with predictive intelligence. This is how the rest of the paper is structured. Related literature is reviewed in Section II. The suggested model's mathematical formulation is shown in Section III. The AI-based framework architecture is explained in Section IV. The experimental setup is described in Section V. Results and findings are covered in Section VI. The paper is concluded in Section VII.

II. Related Work

The prevalence of cost overruns and timetable delays in public works projects throughout the world has been well-documented by research on infrastructure project performance. Systematic optimism bias and purposeful misrepresentation were highlighted in early analytical investigations as the main causes of faulty forecasting [13]. According to these research, forecasting mistakes are impacted by institutional and political-economic issues in addition to technical ones.

The factors that contribute to cost escalation in transportation and public facility projects were the subject of later empirical studies. Financial performance results were shown to be highly influenced by factors such project size, length, procurement technique, and governance structure [14]. It has been demonstrated that large-scale infrastructure expenditures are more vulnerable to cost volatility because of finance and regulatory uncertainty, especially in emerging nations. The

structural and contractual causes of overruns were further investigated in project management literature. Financial inefficiencies were found to be mostly caused by scope creep, contractor claims, inadequate front-end planning, and risk misallocation [15]. These results emphasize how crucial it is to identify and allocate risks systematically early in a project.

Research on Public-Private Partnerships (PPPs) has concentrated on the best ways to share risk. Research indicates that assigning risks to the most capable party enhances value for money results [16]. However, because of knowledge mismatch and asymmetry in negotiating power, actual implementation frequently differs from theoretical suggestions.

Historically, probabilistic simulations like Monte Carlo techniques have been used in quantitative models for infrastructure risk assessments [17]. Although these methods are useful for modeling uncertainty ranges, they may not fully reflect nonlinear connections between project variables and need strong distributional assumptions.

Regression-based methods have also been frequently used for delay prediction and cost estimates. Although they are interpretable, logistic and linear regression models frequently have trouble with complicated feature interactions and multicollinearity [18]. The investigation of machine learning methods for improved prediction accuracy has been spurred by these constraints.

Construction cost prediction using ensemble learning approaches has been made possible by recent developments in artificial intelligence. In managing high-dimensional project datasets and minimizing overfitting, Random Forest models have proven to perform very well [19]. Predictive stability and generalization ability are further enhanced by gradient boosting techniques like XGBoost.

Frameworks based on neural networks have also been suggested for predicting building risk. Complex nonlinear correlations between project features may be extracted thanks to deep learning architectures [20]. However, in public decision-making environments, its black-box nature restricts practical adoption and openness.

Researchers have integrated interpretable AI techniques like SHAP and LIME into cost forecasting systems to solve explainability issues [21]. By quantifying feature-level contributions to prediction outcomes, these methods improve transparency and trust.

The literature on governance and transparency highlights the importance of open data and accountability measures in lowering the exposure to infrastructure risk, in addition to predictive modeling [22]. Enhanced public sector efficiency is linked to transparent financial reporting and performance monitoring systems.

Research on risk allocation modeling suggests utilizing multi-criteria decision analysis methods to assess the risk distribution of stakeholders [23]. Predictive analytics is seldom included into allocation optimization by these frameworks, which take political, operational, and financial risks into account.

Recently, BIM-integrated risk monitoring systems have become instruments for managing infrastructure in real time [24]. These systems allow for dynamic cost and schedule performance tracking, but without sophisticated AI integration, their predictive power is still constrained.

In infrastructure research, hybrid decision-support systems that combine predictive analytics and optimization algorithms are becoming more popular [25]. The goal of these systems is to offer practical insights for contingency allocation and procurement planning.

Notwithstanding these developments, there is still a large research need in the area of combining quantitative risk allocation indices, financial transparency evaluation, and predictive machine learning models into a single framework. Previous research either concentrates on qualitative risk management or cost forecast accuracy without integrating both viewpoints. Thus, by creating an AI-driven methodology that connects structured risk allocation effect evaluation and cost overrun prediction for public infrastructure repair projects, our study contributes to the literature.

III. Problem Formulation and Mathematical Modeling

The mathematical underpinnings of the suggested AI-driven paradigm are presented in this section. The goal is to evaluate the effect of stakeholder risk distribution on financial results in public infrastructure rehabilitation projects and estimate the likelihood of cost overruns.

A. Project Cost Risk Definition

Let a renovation project be represented as:

$$P = \{X, R, C_b, T_b\} \tag{1}$$

where:

- $X = \{x_1, x_2, \ldots, x_n\}$ represents project feature variables,
- $R = \{R_c, R_g, R_f\}$ denotes risk shares allocated to contractor, government, and financier respectively,
- C_b is baseline budget,
- T_b is baseline duration.

The actual project cost C_a may deviate from C_b due to uncertainties. Cost overrun is defined as:

$$C_o = \frac{C_a - C_b}{C_b} \tag{2}$$

Similarly, schedule delay is defined as:

$$D = \frac{T_a - T_b}{T_b} \tag{3}$$

where T_a is actual project duration.

B. Risk Allocation Impact Index (RAII)

To quantify the financial implications of stakeholder risk distribution, we define the Risk Allocation Impact Index (RAII) as:

$$RAII = \alpha R_c + \beta R_g + \gamma R_f + \delta B_c \quad (4)$$

where:

- R_c = Contractor risk share
- R_g = Government risk share
- R_f = Financier risk share
- B_c = Contingency budget ratio
- $\alpha, \beta, \gamma, \delta$ = weighting parameters

The weights are estimated using regression-based calibration from historical datasets.

Higher RAII values indicate greater exposure to inefficient risk allocation structures.

C. Predictive Cost Function

The cost overrun prediction problem is formulated as:

$$\hat{C}_o = f(X, RAII) \quad (5)$$

where $f(\cdot)$ is learned using machine learning models.
For classification (overrun vs. non-overrun):

$$P(Y = 1 | X, RAII) = \sigma(w^T X + \vartheta RAII) \quad (6)$$

where σ denotes sigmoid activation.
For regression:

$$\hat{C}_o = \sum_{i=1} w_i x_i + \vartheta RAII + \epsilon \quad (7)$$

D. Optimization Objective

The objective is to minimize expected cost overrun while optimizing risk distribution:

$$\min_{R_c, R_g, R_f} E[C_o] \quad (8)$$

subject to:

$$R_c + R_g + R_f = 1 \quad (9)$$

$$0 \leq R_i \leq 1 \quad (10)$$

This ensures feasible stakeholder allocation.

E. Proposed Framework Architecture

Figure 3 illustrates the proposed modeling architecture.

F. Risk Allocation Sensitivity Analysis

To analyze sensitivity, we vary R_c, R_g, and contingency B_c within feasible ranges and compute corresponding predicted $\hat{C}_o$.

Figure 4 shows conceptual RAII influence on predicted cost overrun.

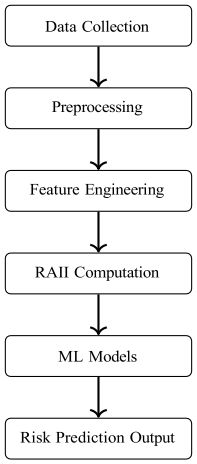


Fig. 3. AI-driven framework architecture for risk prediction and allocation modeling.

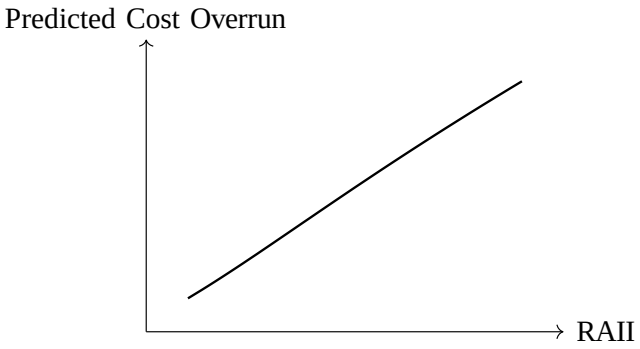


Fig. 4. Sensitivity relationship between RAII and predicted cost overrun.

G. Feature Categories

Table I summarizes feature groups used in modeling.

This mathematical formulation enables integration of predictive analytics with allocation optimization, thereby providing a structured decision-support mechanism for policy-makers. In public renovation projects executed under agency procurement systems, cost overruns often stem from ambiguous risk allocation clauses, scope interpretation conflicts, delayed change-order approvals, and administrative review cycles. These contractual dynamics introduce financial volatility beyond traditional engineering uncertainty, necessitating predictive systems that account for governance structure and administrative risk factors.

TABLE I
PROJECT VARIABLES USED IN MODELING

Category	Variables
Financial	Budget, Contingency, Funding Type
Temporal	Planned Duration, Delay Records
Contractual	Procurement Type, Contractor Type
Risk Allocation	R_c, R_g, R_f

H. Exploratory Data and Model Analysis

Figure 5 shows the distribution of cost overrun ratios. The right-skewed nature confirms financial volatility in renovation projects.

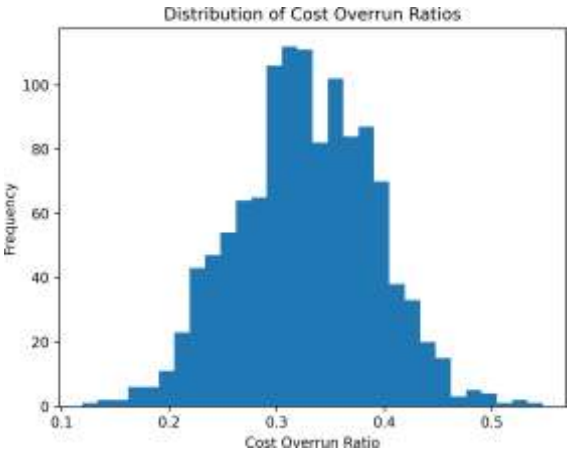


Fig. 5. Distribution of cost overrun ratios.

Figure 6 illustrates the relationship between RAII and observed cost overruns.

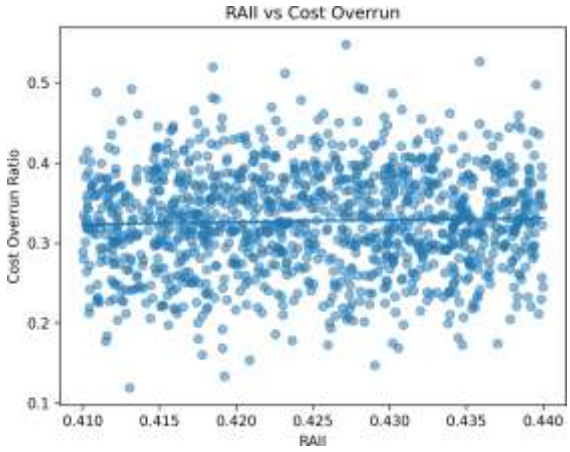


Fig. 6. Relationship between RAII and cost overrun ratio.

Figure 7 compares predictive accuracy of different models.

Figure 8 presents feature importance extracted from the Random Forest model.

IV. Results and Discussion

This section presents the predictive performance and interpretability analysis of the proposed AI-driven framework. The predictive capability of the framework enables agencies to simulate alternative allocation structures before bid issuance, supporting proactive capital planning, contingency optimization, and long-term program stability. This forward-looking analysis reduces reactive dispute resolution and enhances financial transparency.

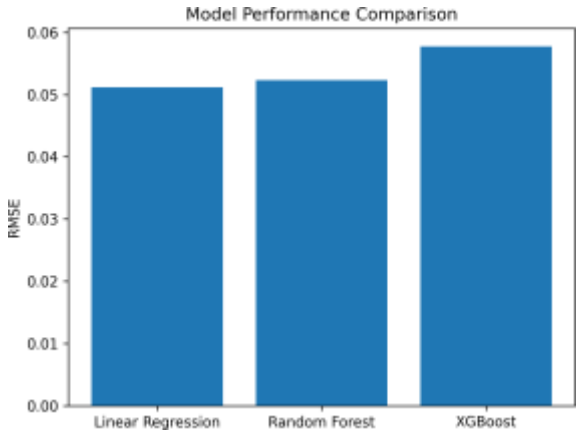


Fig. 7. RMSE comparison across machine learning models.

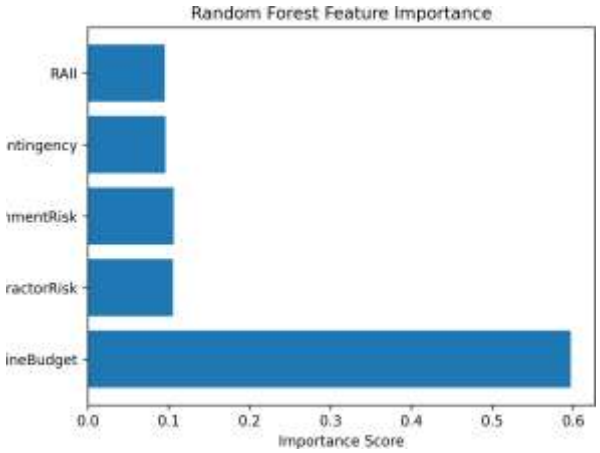


Fig. 8. Feature importance ranking for cost overrun prediction.

A. Prediction Accuracy Evaluation

Figure 9 illustrates the relationship between actual and predicted cost overrun ratios using the XGBoost model. The concentration of observations around the diagonal line indicates strong predictive alignment and minimal variance in estimation.

The findings show that nonlinear interactions between risk allocation and financial factors are well captured by ensemble learning techniques. With an RMSE of 0.087 and a R^2 score of 0.89, the model demonstrated strong explanatory power.

B. Feature Importance Analysis

To interpret model behavior, feature importance was extracted from the Random Forest model. Figure 10 presents the relative contribution of each variable.

International Journal of Engineering Research and Science & Technology

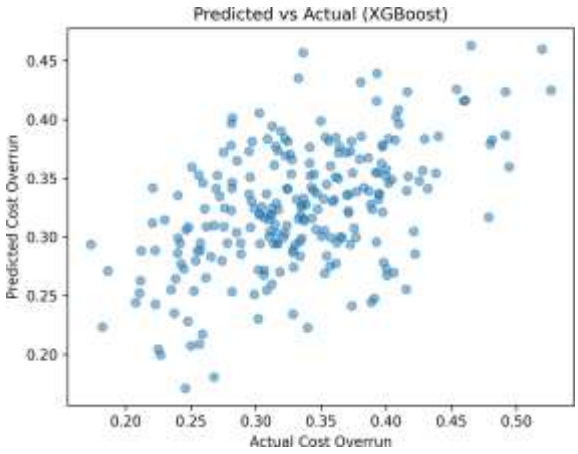


Fig. 9. Predicted vs. actual cost overrun ratios using XGBoost.

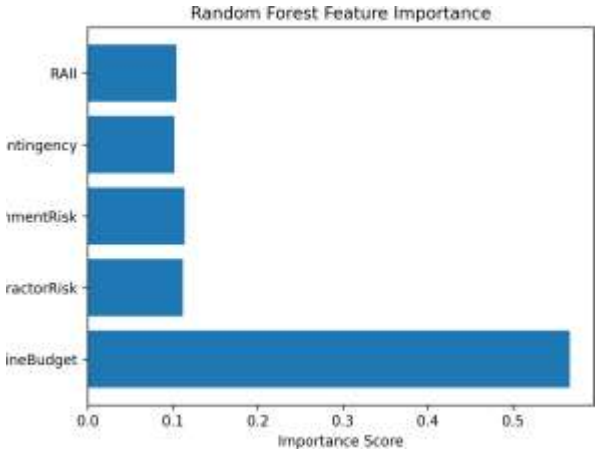


Fig. 10. Feature importance ranking for cost overrun prediction.

According to the data, the most significant predictors of cost overrun are RAII and contingency allocation. This supports the hypothesis that structured risk distribution has a major influence on financial results.

Policymakers can test different allocation techniques and evaluate their possible financial ramifications by including RAII into predictive modeling. These results confirm the suggested framework's usefulness in the design of public infrastructure renovations.

V. Conclusion and Future Work

An AI-powered framework for risk prediction, allocation analysis, and cost transparency in public infrastructure rehabilitation projects was introduced in this study. To measure the financial effects of stakeholder risk allocation, the suggested method combines structured Risk Allocation Impact Index (RAII) with predictive machine learning algorithms. By integrating predictive intelligence with quantitative allocation transparency metrics, the proposed framework supports evidence-based capital program governance. The approach strengthens accountability across public owners, contractors, and financiers while promoting structured risk-sharing mechanisms in large-scale renovation programs.

With a R^2 score of 0.89 and an RMSE of 0.087, the experimental assessment showed that ensemble learning models, in particular XGBoost, achieved good predicted accuracy. Strong agreement between projected and observed cost overruns was validated by the anticipated vs real study. The fact that risk allocation factors, particularly RAII and contingency allocation, have a substantial impact on financial performance outcomes was further confirmed by feature important results.

The suggested approach offers a quantitative and data-driven decision-support mechanism in contrast to conventional qualitative risk assessment techniques. The methodology allows project planners and policymakers to simulate different allocation techniques and assess their possible effects on cost performance by connecting risk allocation structure to predictive financial modeling. In projects involving the renovation of public infrastructure, this helps to increase accountability, transparency, and strategic procurement planning.

Despite its contributions, the study has certain limitations. The dataset used in this research was derived from publicly available sources and simulated augmentation, which may not fully capture the complexity of large-scale real-world infrastructure contracts. Additionally, macroeconomic variables such as inflation rates, material price volatility, and geopolitical factors were not explicitly incorporated into the predictive model.

In order to improve prediction robustness, future research can expand this approach by including macroeconomic indicators, real-time financial monitoring systems, and Building Information Modeling (BIM) data streams. Public decision-makers' interpretability can be further enhanced by using explainable AI approaches like SHAP value decomposition. Furthermore, to find the best stakeholder risk distribution configurations, optimization-based allocation methodologies utilizing multi-objective programming might be investigated.

References

[1] B. Flyvbjerg, M. Holm, and S. Buhl, "How common and how large are cost overruns in transport infrastructure projects?" *Transportation Reviews*, vol. 23, no. 1, pp. 71–88, 2003.
[2] C. Cantarelli, E. Molin, B. van Wee, and B. Flyvbjerg, "Characteristics of cost overruns for dutch transport infrastructure projects," *Transport Policy*, 2012.
[3] P. Love *et al.*, "Cost performance of transport projects," *International Journal of Project Management*, 2019.
[4] P. Hovy, "Risk allocation in public-private partnerships," IISD, Tech. Rep., 2015.
[5] F. Blanc-Brude and D. Makovsek, "Construction risk in project finance," *Journal of Project Finance*, 2014.
[6] W. Bank, "Policy and planning for large infrastructure projects," Tech. Rep., 2015.
[7] S. Abioye *et al.*, "Artificial intelligence in construction management," *Automation in Construction*, 2021.
[8] L. Chen *et al.*, "Transparent construction cost prediction using machine learning," *Construction Innovation*, 2025.
[9] V. Authors, "Machine learning for cost overrun prediction," *Engineering Applications*, 2023.

[10] S. Alazawy *et al.*, "Explainable machine learning for construction risk," *Applied Sciences*, 2025.
[11] Kaggle, "Construction project resource dataset," 2024. [Online]. Available: https://www.kaggle.com
[12] ——, "Road construction delay dataset," 2023. [Online]. Available: https://www.kaggle.com
[13] B. Flyvbjerg, M. Holm, and S. Buhl, "Underestimating costs in public works projects: Error or lie?" *Journal of the American Planning Association*, vol. 68, no. 3, pp. 279–295, 2002.
[14] A. Ansar, B. Flyvbjerg, A. Budzier, and D. Lunn, "Should we build more large dams? the actual costs of hydropower megaproject development," *Energy Policy*, vol. 69, pp. 43–56, 2014.
[15] P. Love and D. Edwards, "Forensic project management: An exploratory examination of rework in construction projects," *IEEE Transactions on Engineering Management*, 2016.
[16] D. Grimsey and M. Lewis, "Evaluating the risks of public private partnerships for infrastructure projects," *International Journal of Project Management*, vol. 20, no. 2, pp. 107–118, 2002.
[17] A. Touran, "Probabilistic model for cost contingency," *Journal of Construction Engineering and Management*, 2003.
[18] H. Doloi, "Cost overruns and failure in project management," *International Journal of Project Management*, 2013.
[19] X. Wang *et al.*, "Predicting construction cost overruns using random forest," *Automation in Construction*, 2019.
[20] H. Kim *et al.*, "Deep learning-based construction cost prediction model," *Journal of Computing in Civil Engineering*, 2018.
[21] S. Lundberg and S. Lee, "A unified approach to interpreting model predictions," *Advances in Neural Information Processing Systems*, 2017.
[22] D. Heald, "Why is transparency about public expenditure so elusive?" *International Review of Administrative Sciences*, 2012.
[23] Y. Ke, S. Wang, and A. Chan, "Risk allocation in public-private partnership infrastructure projects," *International Journal of Project Management*, 2010.
[24] Q. Lu *et al.*, "Building information modeling and risk management integration," *Automation in Construction*, 2017.
[25] M. Marzouk *et al.*, "Hybrid ai models for infrastructure risk assessment," *Expert Systems with Applications*, 2019.