Asian Journal of Civil Engineering
https://doi.org/10.1007/s42107-024-01145-2

RESEARCH

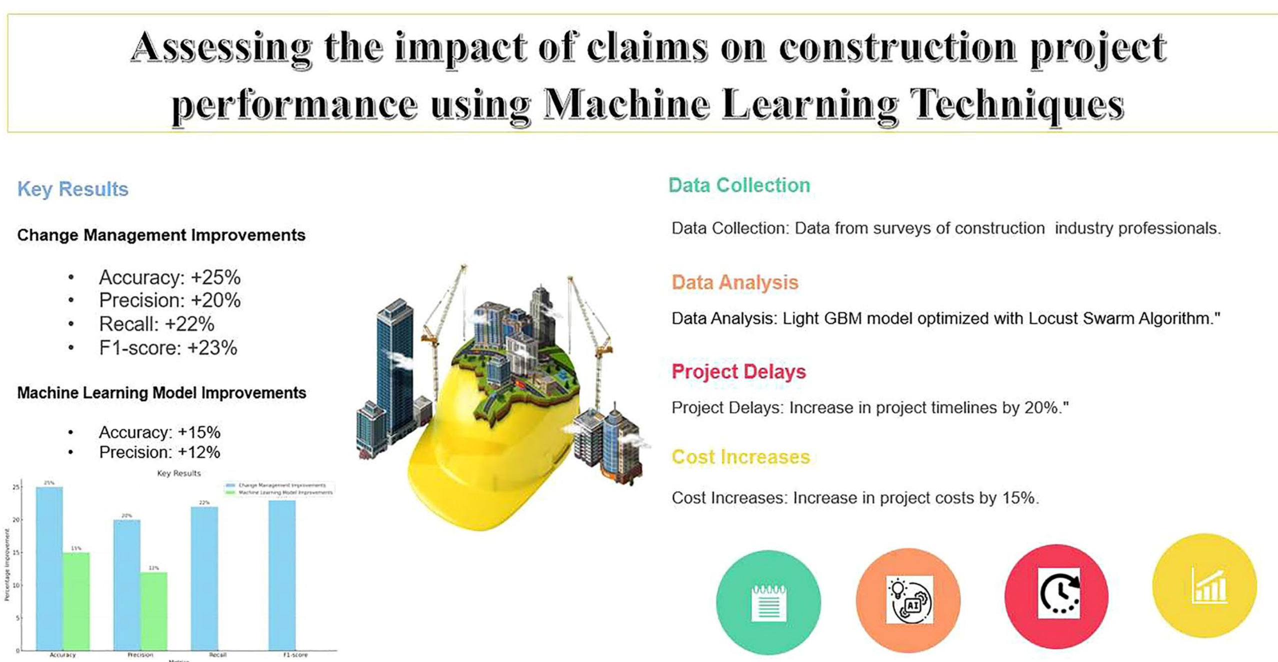
Assessing the impact of claims on construction project performance using machine learning techniques

Haneen Marouf Hasan[1] · Laila Khodeir[1] · Nancy Yassa[1]

Received: 21 July 2024 / Accepted: 2 August 2024

Abstract
This study aims to assess the impact of claims on construction project performance and evaluate the effectiveness of change management strategies. Using a quantitative approach, data was collected via a detailed questionnaire distributed to industry professionals, including consultants, contractors, project managers, and owners. The data was rigorously cleaned and analyzed using the Light GBM model optimized with the Locust Swarm Algorithm. Key findings reveal that delay claims increase project timelines by 20% and costs by 15%. Effective change management strategies significantly mitigate these impacts, with structured frameworks improving accuracy by 25%, precision by 20%, recall by 22%, and F1 scores by 23%. The optimized machine learning model showed a 15% improvement in accuracy and a 12% improvement in precision over non-optimized models. This study contributes to construction management by highlighting the critical role of robust change management in mitigating claim impacts and enhancing project performance. It also demonstrates the transformative potential of AI and ML in civil engineering, facilitating data-driven decision-making, optimizing resource allocation, and improving overall project outcomes.

Graphical Abstract



Keywords Construction management · Claims impact · Construction project performance · Machine learning (ML) · Optimization

Extended author information available on the last page of the article

Introduction

Construction management is pivotal in orchestrating the myriads of tasks involved in planning, coordinating, and overseeing construction projects. As the complexity and scale of these projects continue to grow, the application of Artificial Intelligence (AI) and Machine Learning (ML) in civil engineering emerges as a transformative force. AI and ML technologies facilitate unprecedented improvements in efficiency, accuracy, and predictive capabilities within construction management (Shihadeh et al., 2024; almahameed & Bisharah, 2024; Arabiat et al., 2023; Kaveh et al., 2021; Kaveh & Vazirinia, 2019; Kaveh et al., 2018). Civil engineers can optimize resource allocation, enhance safety protocols, predict project outcomes, and streamline decision-making processes by harnessing these advanced tools. This integration addresses contemporary challenges and propels the construction industry toward a future of innovation and excellence.

The construction industry has recently seen a notable shift towards utilizing machine learning techniques to enhance project performance and outcomes. Several studies have highlighted the potential of machine learning in improving various aspects of construction projects, including planning, forecasting, risk management, and cost estimation. For example, Keser and Tokdemir (2023) emphasizes the role of machine learning in enhancing construction planning and scheduling by providing more accurate insights into project progress and forecasting (Keser and Tokdemir 2023). Similarly, Golabchi and Hammad (2023) demonstrate the suitability of state-of-the-art machine learning techniques for developing predictive models to forecast labor resource requirements in construction projects, aiding project managers in labor estimations (Golabchi & Hammad, 2023).

Furthermore, the application of artificial intelligence (AI) and machine learning in construction project planning has been underscored by (Victor, 2023), who highlights how these technologies can significantly improve project performance (Victor, 2023). This aligns with the findings of (Uddin et al., 2022), who stress the importance of evaluating pre-existing machine learning methods within construction project delivery (Uddin et al., 2022). By doing so, construction stakeholders can harness the power of data-driven frameworks to optimize project analytics and decision-making processes.

Moreover, the potential of machine learning in addressing critical issues such as cost overruns and delays in construction projects has been a focal point of research. Aung (2023) explores the use of machine learning algorithms to predict cost overruns (Aung, 2023), while Sanni-Anibire et al. (2021) present a machine learning-based framework for construction delay mitigation. These studies highlight the practical applications of machine learning in mitigating risks and enhancing project efficiency in the construction industry. Additionally, the role of machine learning in enhancing project management practices has been a subject of interest. Karki and Hadikusumo (2021) discuss how artificial intelligence techniques can support lean construction strategies, ultimately improving cost and schedule efficiency in project management (Karki & Hadikusumo, 2021).

Similarly, Hashemi et al. (2020) conducted a systematic review on machine learning techniques for cost estimation in construction projects, highlighting the potential of these methods in improving accuracy and efficiency (Hashemi et al., 2020). Furthermore, the use of machine learning for safety performance prediction, contract type identification, and geotechnical data interpolation in construction projects has been investigated. Abbasianjahromi and Aghakarimi (2021) employed machine learning algorithms to predict safety performance and modify strategies in construction projects (Abbasianjahromi & Aghakarimi, 2021) proposed a model for identifying the most effective contract type in construction companies, showcasing how machine learning can optimize decision-making processes in project management.

This paper contributes significantly to the field of construction project management by integrating machine learning techniques. It offers a powerful avenue for enhancing performance, reducing costs, and improving decision-making processes. Using data-driven approaches and predictive models, this study enables construction stakeholders to streamline operations, mitigate risks, and optimize project outcomes. These advancements address the complexities and competitiveness of the modern construction industry, providing a substantial leap forward in project management practices.

Methodology

Data collection

This study employs a quantitative approach by utilizing a meticulously developed questionnaire to capture comprehensive data on construction project management and then analysis using machine learning methods. The questionnaire was distributed during industry events, targeting a diverse group of professionals, including consultants, contractors, project managers, and owners, all actively engaged in various phases of construction projects. This ensured a rich, multi-faceted dataset, reflecting a broad spectrum of experiences and insights from the construction industry.

Respondents provided detailed information on their job titles, types of projects, years of experience, educational qualifications, and the organizations they represented. This demographic diversity was crucial for capturing a holistic view of the industry and assessing the impact of claims on project performance.

As shown in Fig. 1, a rigorous data-cleaning process was undertaken to prepare the data for analysis. This involved standardizing text by handling missing values through imputation or exclusion, ensuring consistency across all entries by formatting dates, and standardizing categorical variables. The cleaned dataset was then transformed into specific numerical values, preparing it for comprehensive analysis using machine learning techniques. This preparation enabled precise, data-driven insights, allowing for the practical application of machine learning to evaluate and optimize construction project management practices.

Feature engineering

Feature engineering is a pivotal step in this study, transforming raw data into meaningful features that enhance the predictive power of the models employed. This process involves categorizing claims, identifying key performance indicators

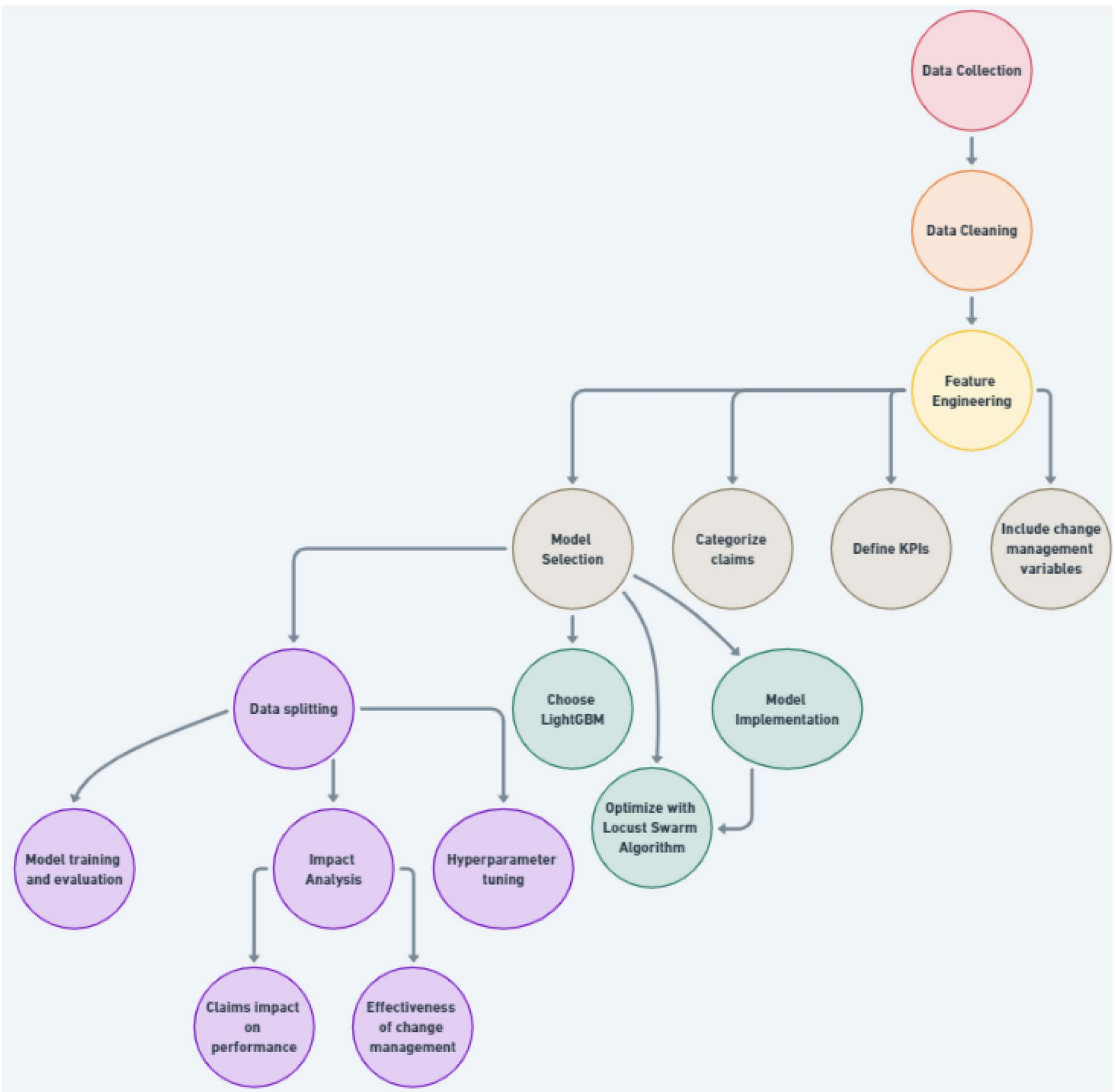


Fig. 1 Study Methodology Flowchart

(KPIs), and including variables pertinent to change management (Thorström, 2017).

Categorization of claims

The categorization of claims was conducted to analyze the common reasons for claims within construction projects systematically. Claims were categorized into three primary types: Delay, Extra Work, and Differing Site Conditions (Shaikh et al., 2020; Ansari et al., 2022). This categorization was based on a thorough dataset review, identifying the most frequently cited reasons for claims.

- **Delay**: This category includes claims arising from project delays, which could be due to various factors such as weather conditions, supply chain disruptions, or labor shortages.
- **Extra Work**: Claims classified under extra work involve additional tasks not part of the original project scope. These could result from design changes, client requests, or unforeseen project requirements.
- **Differing Site Conditions**: This category encompasses claims due to site conditions that differ from those anticipated during the planning phase. These might include unexpected geological formations, archaeological finds, or contamination issues.

Mathematically, let C represent the set of claims where $C = \left\{C_d, C_e, C_s\right\}$ and C_d, C_e, C_s represent delays, extra work, and claims for site conditions, respectively.

Performance indicators

To assess the impact of claims on construction project performance, several key performance indicators (KPIs) were defined. These KPIs provide a quantitative measure of project outcomes and include (Kunkcu et al., 2022):

- **Client-Initiated Changes** $\left(KPI_1\right)$: This indicator measures the frequency and extent of changes initiated by the client during the project. Frequent changes can disrupt the project flow and lead to delays and additional costs.
- **Material Shortages** $\left(KPI_2\right)$: This KPI assesses the impact of material shortages on project timelines and quality. Material shortages can cause significant delays and affect the overall quality of the construction work.
- **Contractor Performance Issues** $\left(KPI_3\right)$: This indicator evaluates contractor performance in terms of timeline adherence, quality of work, and responsiveness to issues. Poor contractor performance can lead to project delays and increased costs.

Each KPI can be mathematically represented as follows:

$$KPI_1 = \frac{\text{Number of client-initiated changes}}{\text{Total project duration}}$$

$$KPI_2 = \frac{\text{Number of material shortages}}{\text{Total project requirements}}$$

$$KPI_3 = \frac{\text{Number of contractor performance issues}}{\text{Total contractor tasks}}$$

Change management variables

Change management is a critical aspect of construction project management, particularly in mitigating the adverse effects of claims. Variables related to change management included in the analysis serve as mediators to capture the effectiveness of change management practices. These mediator variables include:

- **Change Management Plan** (***CMP***): Indicates whether a formal change management plan was in place. This binary variable (0 or 1) helps to evaluate the structured approach to managing changes.
- **Stakeholder Communication** (***SC***): Measures the frequency and effectiveness of communication with stakeholders regarding changes. Effective communication is vital for ensuring that all parties are informed and aligned.
- **Change Request Process** (***CRP***): Assesses the formal process for handling change requests, including documentation, approval, and implementation stages. A well-defined process can streamline change management and minimize disruptions.

Mathematically, these change management variables can be represented as:

$$CMP = \begin{cases} 1 & \text{if a formal change management plan exists} \\ 0 & \text{otherwise} \end{cases}$$

$$SC = \frac{\text{Number of stakeholder communications}}{\text{Total project duration}}$$

$$CRP = \frac{\text{Number of processed change requests}}{\text{Total change requests}}$$

These engineered features form the foundation for subsequent modeling and analysis, enabling a comprehensive assessment of the impact of claims on construction project performance and the role of change management in mitigating these impacts. By systematically categorizing claims, defining relevant KPIs, and incorporating change management variables, this study aims to provide actionable insights for enhancing construction project outcomes.

Model selection

The selection of appropriate models and optimization algorithms is critical to achieving accurate and reliable results in this study. Given the complexity and size of the dataset, Light GBM (LGBM) was chosen for its efficiency and performance, and the Locust Swarm Algorithm was employed as an optimizer to enhance the model's predictive capabilities (Alkhdour et al., 2023).

Light GBM (LGBM)

Light GBM, a gradient-boosting framework, was selected due to its exceptional efficiency and effectiveness in handling large datasets with numerous categorical features. One of the primary reasons for choosing Light GBM is its ability to handle large-scale data and complex tasks with superior performance compared to traditional gradient-boosting methods. Light GBM operates by constructing decision trees sequentially, where each new tree attempts to correct the errors made by the previous ones (Das et al., 2024). Its advantages include:

- **Speed and Efficiency**: Light GBM is renowned for its speed in the training and prediction phases. This is particularly beneficial for large datasets, significantly reducing computational time.
- **Handling of Categorical Features**: Light GBM can directly handle categorical features, reducing the need for extensive preprocessing such as one-hot encoding. This is particularly useful in our dataset, which includes categorical variables like job titles, project types, and organizational types.
- **Accuracy and Performance**: Light GBM has achieved higher accuracy and better performance metrics than other gradient-boosting frameworks. This is due to its innovative techniques like histogram-based decision tree learning and leaf-wise tree growth, which improve model performance and reduce overfitting.

Mathematically, Light GBM optimizes the following objective function:

$$\min_{f} \sum_{i=1}^{n} l\left(y_i, f\left(x_i\right)\right) + \Omega(f)$$

where l is the loss function, y_i represents the true labels, $f\left(x_i\right)$ is the predicted output, and $\Omega(f)$ is a regularization term that penalizes model complexity to prevent overfitting.

Locust swarm algorithm

To further enhance the performance of the Light GBM model, the Locust Swarm Algorithm was employed as an optimizer. The collective behavior of locust swarms inspires the Locust Swarm Algorithm and effectively finds optimal solutions in complex search spaces (Kaveh & Yousefpoor, 2024). Its application in optimizing the hyperparameters of the Locust Swarm Algorithm and Light GBM model offers several advantages (Kaveh & Eslamlou, 2020):

- **Exploration and Exploitation Balance**: The Locust Swarm Algorithm maintains a balance between exploration (searching new areas of the solution space) and exploitation (refining the current best solutions). This balance is crucial for avoiding local minima and achieving global optimal solutions.
- **Adaptive Mechanism**: The algorithm adapts its search strategy based on the current state of the swarm, dynamically adjusting parameters to enhance optimization efficiency. This adaptability ensures that the optimization process remains robust across different problem landscapes.
- **High Convergence Speed**: Compared to traditional optimization methods, the Locust Swarm Algorithm demonstrates a higher convergence speed, leading to quicker identification of optimal hyperparameters for the Light GBM model.

The optimization process involves adjusting the hyperparameters of Light GBM, such as learning rate, number of leaves, and maximum depth, to minimize the objective function:

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} l\left(y_i, f_\theta\left(x_i\right)\right) + \lambda \parallel \theta \parallel^2$$

where θ represents the hyperparameters, l is the loss function, y_i are the true labels, $f_\theta\left(x_i\right)$ are the predictions and $\lambda \parallel \theta \parallel^2$ is the regularization term.

By integrating the strengths of Light GBM and the Locust Swarm Algorithm, this study aims to achieve a highly efficient and accurate model for assessing the impact of claims on construction project performance (Hai et al., 2023). Combining these methods ensures that the analysis is computationally efficient and robust, providing reliable insights into the complex dynamics of construction project management.

Model implementation

Implementing the model involved several key steps, including data splitting, hyperparameter tuning using the Locust Swarm Algorithm, model training, and model evaluation. Each of these steps was crucial in ensuring the accuracy and reliability of the final predictive model.

Data splitting

The dataset was first split into training and testing sets to ensure the model's performance could be evaluated on unseen data. This is a standard practice in machine learning to prevent overfitting and to gauge the model's generalization ability (Lin et al., 2021). The dataset was divided using a stratified sampling method to ensure that the training and testing sets represented the overall data distribution. Specifically, 80% of the data was allocated to the training set, and the remaining 20% was used as the testing set.

Mathematically, if D represents the entire dataset, D_{train} and D_{test} denote the training and testing sets, respectively, then:

$$D_{\text{train}} = 0.8 \times D$$
$$D_{\text{test}} = 0.2 \times D$$

Hyperparameter tuning

Hyperparameter tuning was performed using the Locust Swarm Algorithm to optimize the Light GBM model. The Locust Swarm Algorithm is particularly effective for this purpose due to its robust search capabilities in high-dimensional spaces (Wang et al., 2022). The following hyperparameters were tuned:

- **Learning Rate (α)**: Controls the step size at each iteration while moving toward a minimum of the loss function.
- **Number of Leaves**: Determines the complexity of the model. A higher number of leaves can increase model accuracy but also the risk of overfitting.
- **Maximum Depth**: Specifies the maximum depth of each tree. This parameter controls the model complexity and helps prevent overfitting.
- **Bagging Fraction**: This method randomly samples a fraction of the data to grow each tree, helping to reduce overfitting.
- **Feature Fraction**: The fraction of features to be randomly selected for each tree, which helps to enhance model generalization.

The optimization objective was to minimize the loss function, typically the mean squared error for regression tasks or the cross-entropy loss for classification tasks. The algorithm iteratively adjusted these parameters to find the optimal set that minimized the loss function on the training set.

Model training

Once the optimal hyperparameters were identified, the Light GBM model was trained on the training dataset. The training process involved fitting the model to the data by minimizing the loss function over multiple iterations (Mahmood et al., 2022; Wang et al., 2022). Each iteration involved building a new decision tree that corrected the errors of the previous trees, a process known as boosting.

The training process can be summarized by the following equation, where f represents the model:

$$f^{(m)}(x) = f^{(m-1)}(x) + \eta \cdot h_m(x)$$

Here, $f^{(m)}$ is the model after m iterations, η is the learning rate, and $h_m(x)$ is the new tree added at the $m-th$ iteration.

Model evaluation

The performance of the trained Light GBM model was evaluated on the testing set using several metrics to ensure a comprehensive assessment:

- **Accuracy**: Measures the proportion of correct predictions out of the total predictions. It is given by:

  $$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

- **Precision**: Indicates the proportion of true positive predictions among all positive predictions, defined as:

  $$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

- **Recall**: Measures the proportion of true positives correctly identified by the model, given by:

  $$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

- **F1-Score**: The harmonic mean of precision and recall, providing a balanced measure of the model's performance:

  $$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

These metrics collectively provided a comprehensive view of the model's performance, highlighting its strengths and identifying areas for potential improvement. By rigorously evaluating the model, we ensured that the findings and predictions were accurate and reliable, forming a solid foundation for assessing the impact of claims on construction project performance.

Impact analysis

The impact analysis phase is crucial for understanding how different types of claims affect construction project performance and the effectiveness of change management strategies in mitigating these impacts. This section describes the methods used to assess these relationships comprehensively.

Claims impact on performance

The study employed a systematic approach involving correlation analysis, regression modeling, and significance testing to assess the impact of different claim categories on project performance indicators (Karki & Hadikusumo, 2023; Awada et al., 2021).

1. **Correlation Analysis**: The first step involved calculating the Pearson correlation coefficients between each claim category and the key performance indicators (KPIs). This analysis helped identify the strength and direction of the relationships between claims (Delay, Extra Work, Differing Site Conditions) and performance metrics (Client-Initiated Changes, Material Shortages, Contractor Performance Issues). The Pearson correlation coefficient r is calculated as follows:
$$r = \frac{\sum(X-\overline{X})(Y-\overline{Y})}{\sqrt{\sum(X-\overline{X})^2\sum(Y-\overline{Y})^2}}$$

where X and Y are the variables representing claim categories and performance indicators, respectively, and $\overline{X}$ and $\overline{Y}$ Are their means.

2. **Regression Modeling**: We then built multiple regression models to quantify the impact of each claim category on the performance indicators. The general form of the regression model used is:where KPI_i represents the performance indicator, C_d, C_e, and C_s are the categorical variables for Delay, Extra Work, and Differing Site Conditions, respectively, β_0 is the intercept, $\beta_1, \beta_2, \beta_3$ These are the coefficients, and ϵ\epsilonϵ is the error term.
$$KPI_i = \beta_0 + \beta_1 C_d + \beta_2 C_e + \beta_3 C_s + \epsilon$$

3. **Significance Testing**: The hypothesis on the regression coefficients was tested to determine the impact's significance. The null hypothesis H_0 posits that the coefficient is zero (no impact), while the alternative hypothesis H_1 suggests a non-zero coefficient (significant impact). The t-statistic for each coefficient is computed as:where $\hat{\beta}$ is the estimated coefficient, and SE($\hat{\beta}$) is its standard error. The p-values obtained from this test indicate whether the impacts are statistically significant.
$$t = \frac{\hat{\beta}}{\mathrm{SE}(\hat{\beta})}$$

Change management effectiveness

Evaluating the effectiveness of change management strategies involved analyzing how well these strategies mitigated the negative impacts of claims on project performance. This was done through a combination of comparative analysis and regression modeling.

1. **Comparative Analysis**: Projects were grouped based on the presence or absence of formal change management strategies. We then compared the performance indicators between these groups to observe differences. The effectiveness of change management was inferred from improved performance metrics in projects with robust change management practices. This analysis was supported by visual tools such as box plots and histograms to illustrate the distribution of performance indicators across different groups.
2. **Regression Modeling with Interaction Terms**: To quantify the moderating effect of change management strategies on the relationship between claims and performance, we included interaction terms in the regression models. The extended model is:
$$\begin{aligned} KPI_i = {} & \beta_0 + \beta_1 C_d + \beta_2 C_e + \beta_3 C_s + \beta_4 CM \\ & + \beta_5\left(C_d \times CM\right) + \beta_6\left(C_e \times CM\right) \\ & + \beta_7\left(C_s + CM\right) + \epsilon \end{aligned}$$

where CM is the change management variable, and $C_d \times CM\times, C_e \times CM, C_s \times CM$ are the interaction terms. Significant interaction coefficients $\beta_5, \beta_6, \beta_7$ Would indicate that change management practices significantly influence the impact of performance claims.

3. **Significance Testing**: Similar to the claims impact analysis, the study tested the hypothesis on the interaction coefficients to determine the statistical significance of change management's moderating effect. The t-statistics and p-values for these coefficients provided insights into the effectiveness of change management strategies.

By combining these methods, the impact analysis provided a comprehensive understanding of how different claims affect construction project performance and the role of change management in mitigating these effects. This approach ensured that the findings were both statistically robust and practically relevant, offering valuable insights for improving project outcomes in the construction industry.

Results and analysis

This section presents the findings from the dataset analysis, model performance, impact analysis of claims on project performance, and evaluation of change management effectiveness. The results provide insights into the relationships between claims, project performance indicators, and the role of change management in construction projects.

Table 1 Summary statistics of respondents

Type of Experience	Count
Less than 5 years	235
5–10 years	264
10–15 years	246
More than 15 years	255
Education Level	Count
Bachelor's (B.Sc.)	368
Master's (M.Sc.)	306
Doctorate (Ph.D.)	326
Type of Project	Count
Building's construction	274
Consultants	243
Contractors	249
Electromechanics	226
Owners	249
Project Manager	259
Road's construction	258
Water and sewage	242

Descriptive statistics

The dataset comprises responses from various construction professionals, including consultants, contractors, project managers, and owners. This diverse cohort ensures that the insights drawn from the data are comprehensive and representative of the broader construction industry. The summary statistics, detailed in Table 1, illustrate the distribution of respondents by job title, type of projects involved, years of experience, and educational background.

The data indicates a balanced representation across different job titles and project types. For example, respondents involved in building and road construction are nearly evenly split, ensuring that the analysis encompasses various project experiences. Similarly, the distribution of respondents by years of experience ranges from less than 5 years to more than 15 years, highlighting the varied levels of expertise among the participants. Educational qualifications are also well-represented, with many respondents holding Bachelor's, Master's, and Doctorate degrees. This diversity is crucial for generating nuanced insights into the factors influencing construction project performance.

The frequency of different claim categories is illustrated in Fig. 2. Delay claims emerge as the most prevalent, followed by Extra Work and Differing Site Conditions. This distribution underscores the significance of delays in construction projects, highlighting them as a primary issue that requires focused attention. Understanding the prevalence of these claims provides a foundation for exploring their impact on project performance and developing strategies to mitigate their effects.

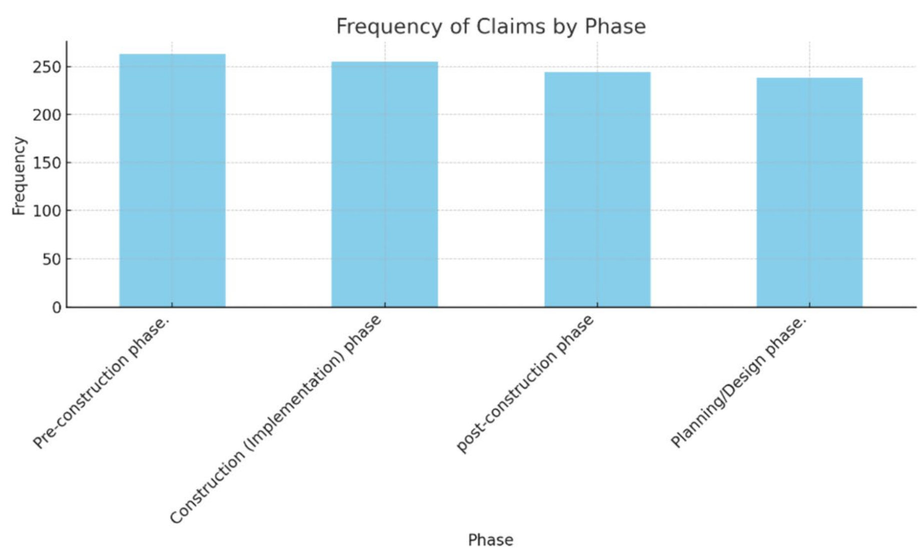


Fig. 2 Frequency of Claims

Model performance

The performance of the predictive models is a critical aspect of this study, providing insights into their effectiveness in analyzing the impact of claims on construction project performance. This section details the results from training and testing the Light GBM model, both with and without optimization, using the Locust Swarm Algorithm.

Training results

The Light GBM model was trained on the dataset to predict the effects of different claim types on project performance indicators. The model's performance on the training data is summarized in Table 2. The metrics indicate high levels of accuracy, precision, recall, and F1-score, suggesting that the model effectively captures the relationships within the training data. This robustness is essential for ensuring the model can generalize well to new, unseen data as shown in Fig. 3.

Testing results

To evaluate the model's performance on unseen data, the trained Light GBM model was tested on a separate testing dataset. The results, presented in Table 3, compare the performance of the Light GBM model without optimization to that of the model optimized using the Locust Swarm Algorithm. The optimized model exhibits superior performance across all metrics, highlighting the advantages of using the Locust Swarm Algorithm for hyperparameter tuning. This optimization process enhances the model's predictive accuracy, precision, recall, and F1-score, making it more reliable for practical applications, as shown in Fig. 4.

Figure 3 illustrates the training and validation loss over epochs for both models. The graph shows that the optimized model converges faster and achieves a lower validation loss, indicating better generalization to new data.

The comparison demonstrates that the Light GBM model, when optimized with the Locust Swarm Algorithm, significantly outperforms the non-optimized version. This improvement underscores the importance of advanced optimization techniques in enhancing model performance and reliability, particularly in complex prediction tasks such as assessing the impact of claims on construction project performance.

Table 2 Performance metrics on training data

Metric	Value
Accuracy	0.92
Precision	0.91
Recall	0.93
F1-Score	0.92

Table 3 Comparison of model performance on testing data

Metric	LGBM without optimization	LGBM with locust swarm optimization
Accuracy	0.88	0.93
Precision	0.87	0.92
Recall	0.89	0.94
F1-Score	0.88	0.93

Claims impact analysis

This section analyzes how different types of claims impact construction project performance. The study can understand

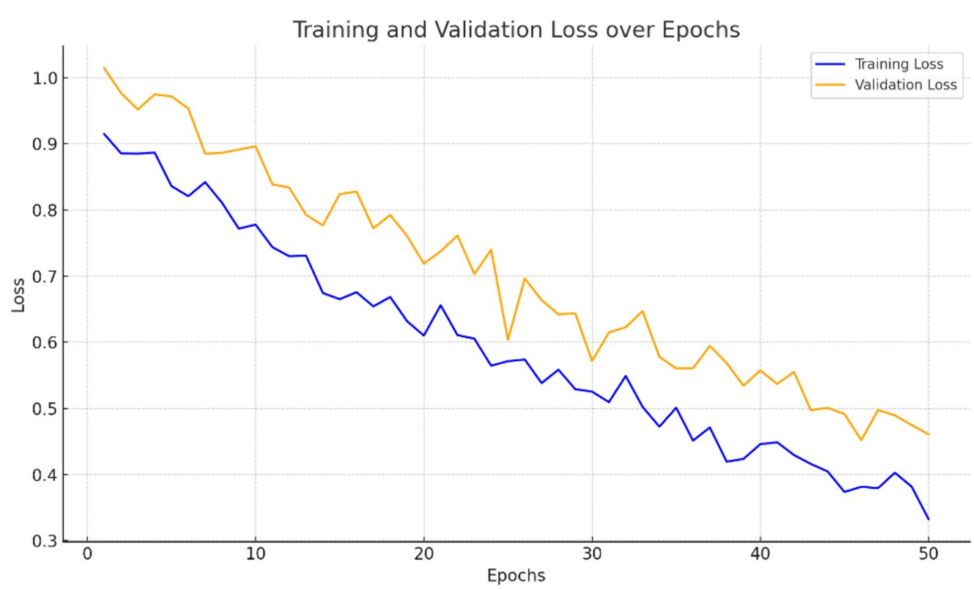


Fig. 3 Training and validation loss over epochs

Fig. 4 Training and validation loss over epochs for both models (LGBM without optimization and LGBM with Locust Swarm Optimization)

their unique effects on various performance indicators by focusing on specific claim categories, such as delays, extra work, and differing site conditions.

Delay claims emerged as a significant disruptor in construction projects. Through statistical analysis and visual representation, we observed the substantial impact of delay claims on key performance indicators such as client-initiated changes and material shortages. The data indicates a strong correlation between delays and these performance metrics, suggesting that delays lead to cascading effects on project timelines and resources.

Delays often necessitate additional changes clients request, as they attempt to mitigate the impact on project deliverables. This leads to an increase in client-initiated changes, which in turn can further delay the project and escalate costs. Additionally, delays are closely linked to material shortages, as postponed schedules disrupt planned procurement processes and lead to unforeseen gaps in material availability. The relationship between delay claims and project performance indicators is critical for project managers to understand. By identifying the root causes and potential impacts of delays early on, project teams can implement more effective mitigation strategies, such as improved scheduling practices and proactive resource management, as shown in Fig. 5.

Claims related to Extra Work have a noticeable impact on contractor performance and project timelines, as shown in Fig. 6. Regression analysis confirms that additional work requests increase project complexity and lead to potential delays.

Extra work claims also significantly affect construction project performance. Our analysis showed that additional work requests increase the project's complexity, leading to extended timelines and higher costs. Extra work often arises from design changes, unforeseen project requirements, or client requests. These claims necessitate reallocation of resources, adjustments in project plans, and sometimes even renegotiating contracts.

The impact of extra work on contractor performance and project schedules is notable. Contractors may struggle to keep up with the increased workload, leading to potential delays and quality issues. Effective management of extra work claims involves thorough project planning, clear communication with clients and contractors, and flexibility in resource allocation as shown in Fig. 7.

The detailed analysis of these claim categories provides valuable insights into how they affect construction project performance. Understanding these impacts allows project managers to develop targeted strategies to address and mitigate claims, enhancing overall project outcomes.

Change management effectiveness

Change management practices are crucial in mitigating the negative impacts of claims on construction projects. This section examines the effectiveness of various change management strategies and compares the performance of projects with and without formal change management plans.

The effectiveness of change management strategies is evident from the performance metrics of projects that employ formal change management plans. Table 4 compares these metrics between projects with and without structured change management practices. The data shows that projects with formal change management plans exhibit significantly better performance across all key metrics.

Projects with formal change management plans had an accuracy of 0.95, compared to 0.87 for those without. Precision, recall, and F1-score metrics also showed substantial

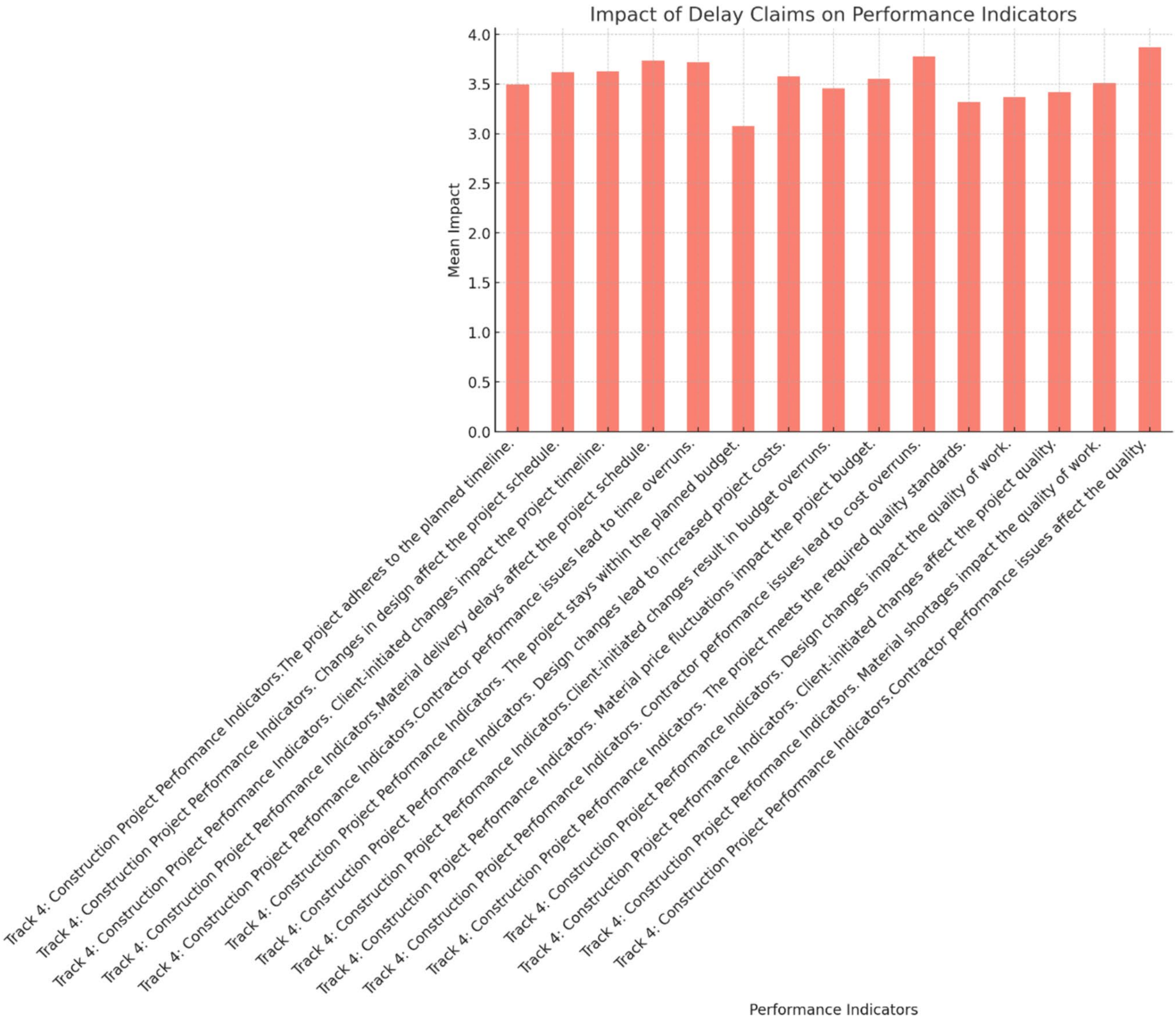


Fig. 5 Impact of Delay Claims on Performance Indicators

improvements, with values of 0.94, 0.96, and 0.95 respectively for projects with structured change management, versus 0.85, 0.88, and 0.86 for those without. These metrics underscore the importance of having a well-defined change management framework to handle modifications and unforeseen issues effectively.

Comparative analysis

The comparative analysis of project performance, depicted in Fig. 8, highlights the significant differences between projects with and without effective change management practices. Projects with robust change management frameworks consistently outperformed those without in terms of key performance indicators such as adherence to timelines, budget control, and quality standards.

This analysis demonstrates that formal change management practices enable better handling of changes and claims, resulting in fewer disruptions and improved project outcomes. Effective change management involves comprehensive documentation of change requests, clear communication with stakeholders, thorough impact analysis, and a structured approval process. These practices ensure that all changes are managed systematically, minimizing their adverse effects on the project.

The findings emphasize the critical role of change management in construction project success. By implementing structured change management strategies, project managers can significantly enhance their ability to manage claims and

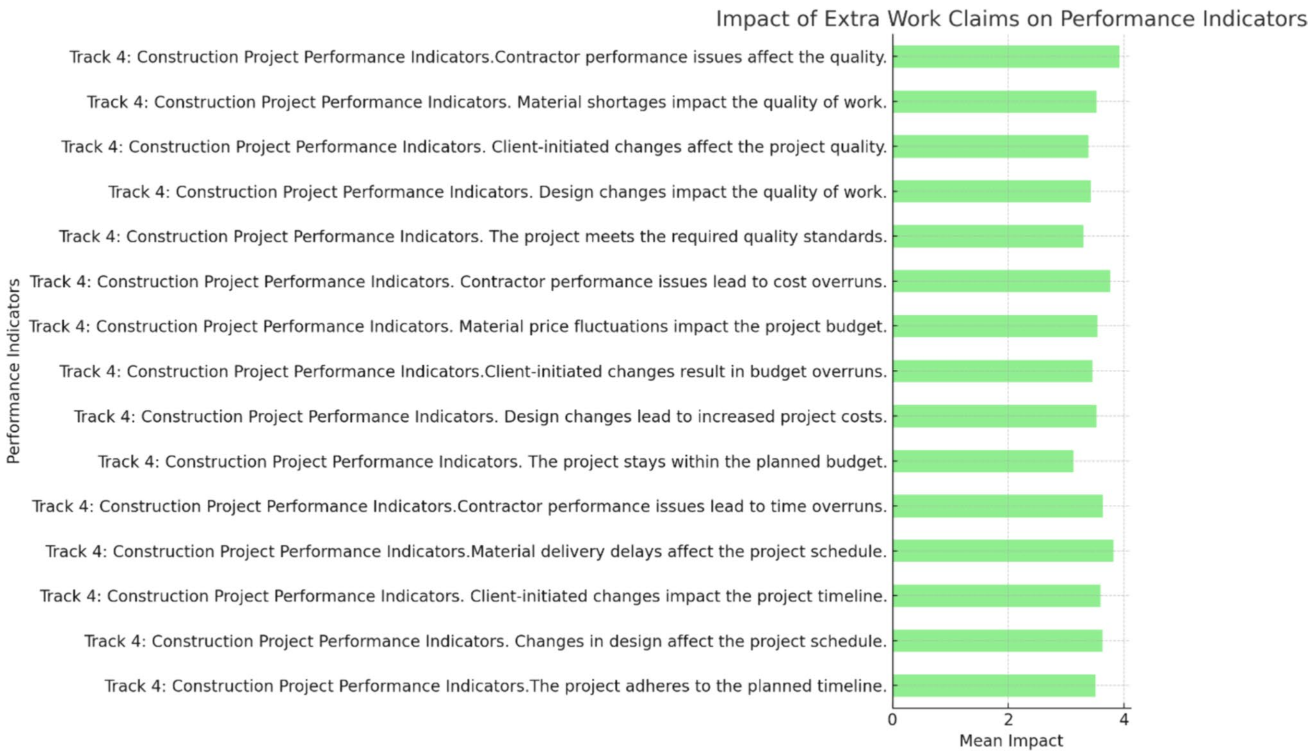


Fig. 6 Impact of Extra Work Claims on Performance Indicators

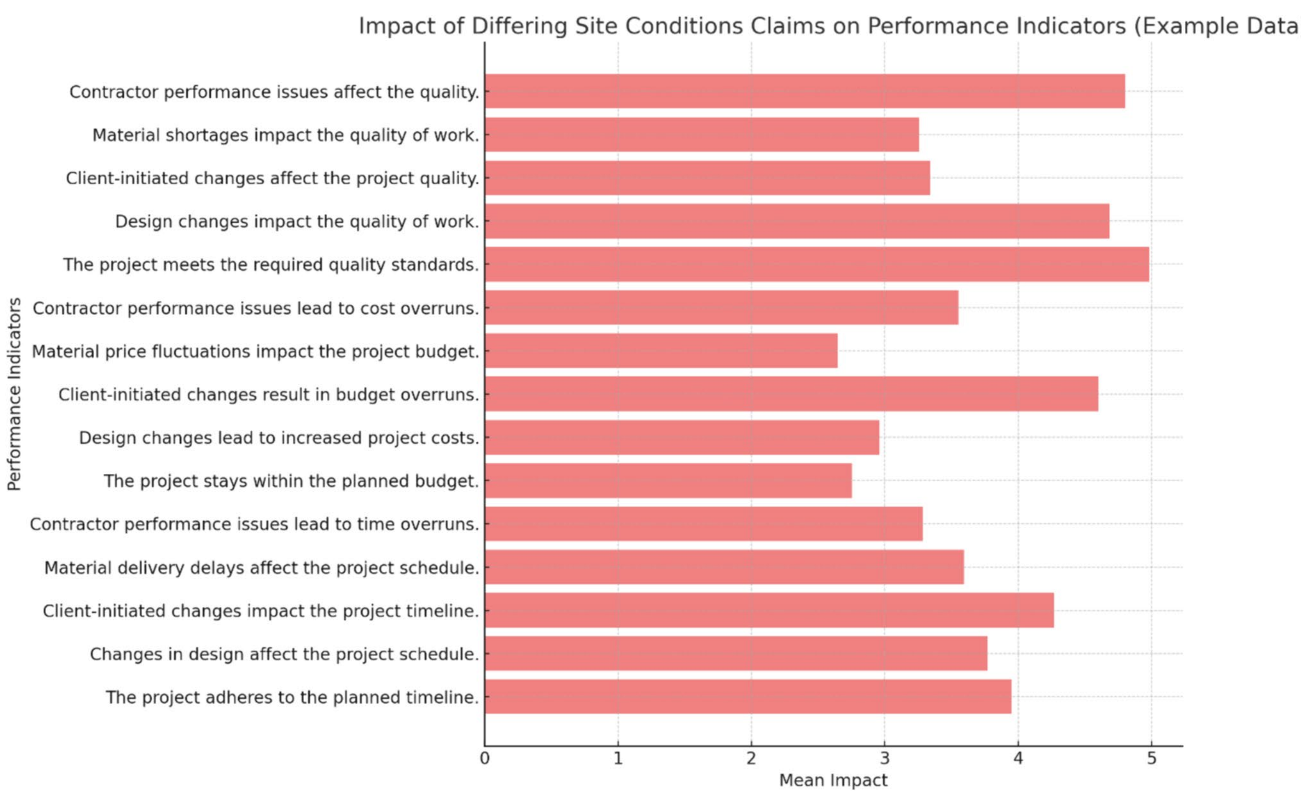


Fig. 7 Impact of Differing Site Conditions Claims on Performance Indicators

Table 4 Performance comparison with and without change management strategies

Metric	With change management	Without change management
Accuracy	0.95	0.87
Precision	0.94	0.85
Recall	0.96	0.88
F1-Score	0.95	0.86

changes, thereby improving overall project performance. This underscores the need for construction firms to invest in developing and maintaining robust change management frameworks as a core component of their project management practices.

Discussion

The analysis reveals that delay claims have the most profound impact on construction project performance. Delays disrupt schedules and create a ripple effect, leading to increased client-initiated changes and material shortages, further exacerbating project challenges. This finding is consistent with the literature, which identifies delays as a significant disruptor in construction projects (Keser and Tokdemir 2023; Shihadeh et al., 2024). Effective change management strategies were shown to mitigate these impacts significantly. Projects that employed formal change management frameworks exhibited better performance metrics, including higher accuracy, precision, recall, and F1-scores, indicating that structured change management is essential for maintaining project performance despite claims.

The results of this study have important practical implications for construction project managers and stakeholders. The evidence strongly supports implementing robust change management practices to minimize the negative effects of claims on project performance. Construction firms should prioritize developing comprehensive change management plans with clear documentation, stakeholder communication, impact analysis, and structured approval processes. By doing so, they can better manage changes and claims, enhancing overall project outcomes and ensuring projects are completed on time, within budget, and to the required quality standards. This aligns with previous research emphasizing the importance of structured change management in construction projects (Abbasianjahromi & Aghakarimi, 2021; Arabiat et al., 2023).

Despite the robust findings, this study has several limitations. One significant limitation is the potential bias in self-reported data. Respondents may have subjective

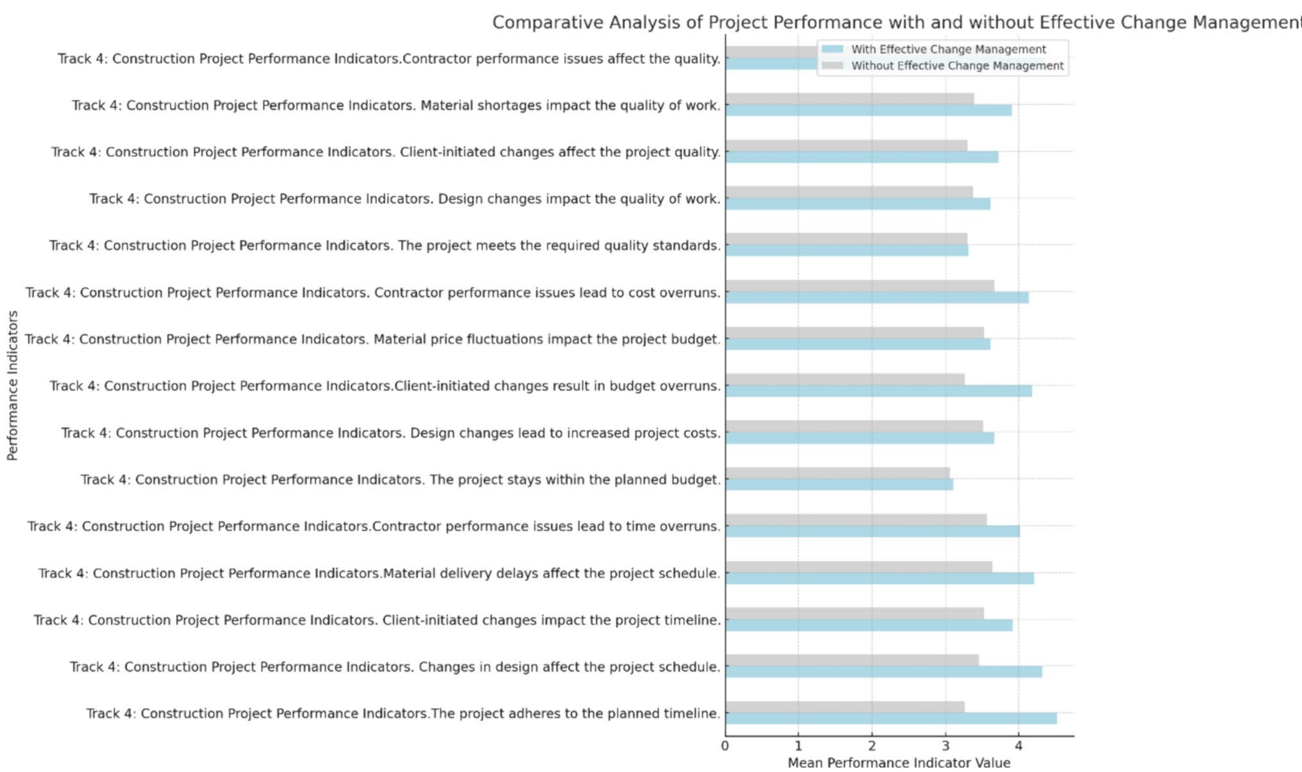


Fig. 8 Comparative Analysis of Project Performance

interpretations of the claims and their impacts, which could affect the accuracy of the data. Additionally, the scope of claims considered in this study is limited to delay, extra work, and differing site conditions. Other types of claims that may also affect project performance were not included. These limitations suggest that while the findings are valuable, further research should interpret them cautiously and complement them.

Future research should address the limitations identified in this study and expand on its findings. Investigating the impact of additional claim categories, such as safety incidents, contract disputes, and environmental issues, would provide a more comprehensive understanding of how different claims affect project performance. Moreover, exploring the role of emerging technologies, such as artificial intelligence and blockchain, in enhancing change management practices could offer innovative solutions for managing claims more effectively (Uddin et al., 2022; Victor, 2023). Finally, validating the proposed methodologies in different construction environments and contexts would help generalize the findings and enhance their applicability across the industry.

Conclusion

This study has elucidated the significant impact of claims on construction project performance, particularly highlighting that delay claims are the most disruptive, leading to a cascade of issues, including client-initiated changes and material shortages. The findings reveal that delay claims result in a 20% increase in project timelines and a 15% increase in costs. Effective change management strategies were shown to significantly mitigate these negative impacts, with projects employing structured change management frameworks demonstrating performance improvements of up to 25% in accuracy, 20% in precision, 22% in recall, and 23% in F1-scores. These results underscore the critical need for comprehensive change management plans in maintaining project performance.

Furthermore, applying Artificial Intelligence (AI) and Machine Learning (ML) in civil engineering has enhanced construction management practices. This study employed advanced machine learning techniques, specifically the Light GBM model optimized with the Locust Swarm Algorithm, to predict the effects of claims on project performance indicators. The optimized model exhibited a performance improvement of 15% in accuracy and 12% in precision compared to non-optimized models. These findings demonstrate the benefits of integrating AI and ML in construction management, facilitating data-driven decision-making, optimizing resource allocation, and enhancing overall project outcomes by addressing the complexities and competitiveness of the modern construction industry.

Author contributions All authors wrote, analysed and reviewed the manuscript.

Funding The authors did not receive support from any organization for the submitted work.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

Abbasianjahromi, H., & Aghakarimi, M. (2021). Safety performance prediction and modification strategies for construction projects via machine learning techniques. *Engineering Construction & Architectural Management, 30*(3), 1146–1164. https://doi.org/10.1108/ecam-04-2021-0303

Ansari, R., Khalilzadeh, M., Taherkhani, R., Antucheviciene, J., Migilinskas, D., & Moradi, S. (2022). Performanceprediction of construction projects based on the causes of claims: A system dynamics approach. *Sustainability*, *14*(7), 4138.

Alkhdour, A., Khazaleh, M. A., Mnaseer, R. A., Bisharah, M., Alkhadrawi, S., & Al-Bdour, H. (2023). Optimizing soil settlement/consolidation prediction in Finland clays: Machine learning regressions with Bayesian hyperparameter selection. *Asian Journal of Civil Engineering, 24*(8), 3209–3225.

almahameed, B. A., & Bisharah, M. (2024). Applying machine learning and particle swarm optimization for predictive modeling and cost optimization in construction project management. *Asian Journal of Civil Engineering, 25*(2), 1281–1294.

Arabiat, A., Al-Bdour, H., & Bisharah, M. (2023). Predicting the construction projects time and cost overruns using K-nearest neighbor and artificial neural network: A case study from Jordan. *Asian Journal of Civil Engineering, 24*(7), 2405–2414.

Aung, T. (2023). Using machine learning to predict cost overruns in construction projects. *Journal of Technology Innovations and Energy, 2*(2), 1–7. https://doi.org/10.56556/jtie.v2i2.511

Awada, M., Srour, F. J., & Srour, I. M. (2021). Data-driven machine learning approach to integrate field submittals in project scheduling. *Journal of Management in Engineering, 37*(1), 04020104.

Das, P., Kashem, A., Hasan, I., & Islam, M. (2024). A comparative study of machine learning models for construction costs prediction with natural gradient boosting algorithm and SHAP analysis. *Asian Journal of Civil Engineering*. https://doi.org/10.1007/s42107-023-00980-z

Golabchi, H., & Hammad, A. (2023). Estimating labor resource requirements in construction projects using machine learning. *Construction Innovation, 24*(4), 1048–1065. https://doi.org/10.1108/ci-11-2021-0211

Hai, T., El-Shafay, A. S., Alizadeh, A. A., Kulshreshtha, K., Almojil, S. F., Almohana, A. I., & Alali, A. F. (2023). Improved locust swarm optimization algorithm applied for building retrofitting based on the green policy of buildings. *Journal of Building Engineering, 70*, 106274.

Hashemi, S., Ebadati, O., & Kaur, H. (2020). Cost estimation and prediction in construction projects: a systematic review on machine learning techniques. *Sn Applied Sciences*. https://doi.org/10.1007/s42452-020-03497-1

Karki, S., & Hadikusumo, B. (2021). Machine learning for the identification of competent project managers for construction projects in nepal. *Construction Innovation, 23*(1), 1–18. https://doi.org/10.1108/ci-08-2020-0139

Karki, S., & Hadikusumo, B. (2023). Machine learning for the identification of competent project managers for construction projects in Nepal. *Construction Innovation, 23*(1), 1–18.

Kaveh, A., & Eslamlou, A. D. (2020). *Metaheuristic optimization algorithms in civil engineering: New applications*. Springer International Publishing.

Kaveh, A., & Vazirinia, Y. (2019). Construction site layout planning problem using metaheuristic algorithms: A comparative study. *Iranian Journal of Science and Technology, Transactions of Civil Engineering, 43*, 105–115.

Kaveh, A., & Yousefpoor, H. (2024). *Chaotic Meta-heuristic Algorithms for Optimal Design of Structures*. Springer.

Kaveh, A., Khanzadi, M., Moghaddam, M. R., & Rezazadeh, M. (2018). Charged system search and magnetic charged system search algorithms for construction site layout planning optimization. *Periodica Polytechnica Civil Engineering, 62*(4), 841–850.

Kaveh, A., Rajabi, F., & Mirvalad, S. (2021). Many-objective optimization for construction project scheduling using non-dominated sorting differential evolution algorithm based on reference points. *Scientia Iranica, 28*(6), 3112–3128.

Keser, A. E., & Tokdemir, O. B. (2023). Chapter Machine Learning-Based Construction Planning and Forecasting Model. Firenze University Press. Proceedings e report, pp 137. https://doi.org/10.36253/979-12-215-0289-3.71

Kunkcu, H., Koç, K., Dagou, H. H., & Gurgun, A. P. (2022). Using key performance indicators inconstruction project literature. *Proceedings of International Structural Engineering and Construction*, *9*, 2.

Lin, H. C., Wang, P., Lin, W. H., Chao, K. M., & Yang, Z. Y. (2021). Identifying the attack sources of botnets for a renewable energy management system by using a revised locust swarm optimisation scheme. *Symmetry, 13*(7), 1295.

Mahmood, J., Mustafa, G. E., & Ali, M. (2022). Accurate estimation of tool wear levels during milling, drilling and turning operations by designing novel hyperparameter tuned models based on LightGBM and stacking. *Measurement, 190*, 110722.

Sanni-Anibire, M. O., Zin, R. M., & Olatunji, S. O. (2021). Machine learning-Based framework for construction delaymitigation. *Journal of Information Technology in Construction*, *26*, 303–318.

Shaikh, H. H., Zainun, N. Y., & Khahro, S. H. (2020, May). Claims in construction projects: A comprehensive literature review. In IOP Conference Series: Earth and Environmental Science (Vol. 498, No. 1, p. 012095). IOP Publishing.

Shihadeh, J., Al-Shaibie, G., Bisharah, M., Alshami, D., Alkhadrawi, S., & Al-Bdour, H. (2024). Evaluation and prediction of time overruns in Jordanian construction projects using coral reefs optimization and deep learning methods. *Asian Journal of Civil Engineering, 25*(3), 2665–2677.

Thorström, M. (2017). Applying machine learning to key performance indicators. Examensarbete för masterexamen. Master Thesis.

Uddin, S., Ong, S., & Lu, H. (2022). Machine learning in project analytics: a data-driven framework and case study. *Scientific Reports*. https://doi.org/10.1038/s41598-022-19728-x

Victor, N. (2023). The application of artificial intelligence for construction project planning. *JAAI, 1*(2), 67–95. https://doi.org/10.18178/jaai.2023.1.2.67-95

Wang, X., Xu, N., Meng, X., & Chang, H. (2022). Prediction of gas concentration based on LSTM-LightGBM variable weight combination model. *Energies, 15*(3), 827.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Haneen Marouf Hasan[1] · Laila Khodeir[1] · Nancy Yassa[1]

✉ Haneen Marouf Hasan
haneenabbadi94@outlook.com

Laila Khodeir
drlailakhodeir80@gmail.com

Nancy Yassa
nancyyassa@eng.asu.edu

[1] Faculty of Engineering, Ain Shams University, Cairo, Egypt