UTS ePRESS

Construction Economics and Building

Vol. 24, No. 1/2
July 2024

Citation: Wood, X., Ghimire, P., Kim, S., Barutha, P., Jeong, H. D. 2024. Framework for Evaluating the Success of Integrated Project Delivery in the Industrial Construction Sector: A Mixed Methods Approach & Machine Learning Application. *Construction Economics and Building*, 24:1/2, 94–118. https://doi.org/10.5130/AJCEB.v24i1/2.8783

ISSN 2204-9029 | Published by UTS ePRESS | https://epress.lib.uts.edu.au/journals/index.php/AJCEB

RESEARCH ARTICLE

Framework for Evaluating the Success of Integrated Project Delivery in the Industrial Construction Sector: A Mixed Methods Approach & Machine Learning Application

Xavier Wood[1], Prashnna Ghimire[1], Suryeon Kim[2], Philip Barutha[3], H. David Jeong[2]

[1] Durham School of Architectural Engineering & Construction, University of Nebraska-Lincoln, USA

[2] Department of Construction Science, Texas A&M University, USA

[3] Civil and Environmental Engineering and Earth Sciences, University of Notre Dame, USA

Corresponding author:

DOI: https://doi.org/10.5130/AJCEB.v24i1/2.8783

Article History: Received 25/08/2023; Revised 21/01/2024; Accepted 02/03/2024; Published 01/07/2024

Abstract

Integrated project delivery (IPD) has gained traction as a collaborative approach to managing complexity and uncertainty in large industrial capital projects. While IPD emphasizes team integration and process alignment to drive better outcomes, the lack of standardized benchmarks to evaluate its performance relative to traditional methods persists as a barrier. To bridge this gap, this study developed a practical, and unbiased Project Success Framework (PSF) for IPD on industrial projects. A mixed methods research approach including subject matter experts' survey, research charrette, and validation survey was conducted to build and validate the PSF. In addition, this study proposed a machine learning (ML)-based application tool embedding PSF to enhance the practicality and applicability of PSF. The machine learning-based application tool was validated by comparing the results with the PSF suggested in this research. The PSF developed in this study allows researchers and practitioners to empirically evaluate the integrated project delivery's efficacy on key industrial project outcomes. In addition, it offers a method to compare project delivery methods across diverse projects, aiding

DECLARATION OF CONFLICTING INTEREST The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. **FUNDING** The author(s) received no financial support for the research, authorship, and/or publication of this article.

organizations in precise selection using empirical evidence for optimal results. Moreover, this framework aids clients in crafting shared risk/reward models that foster successful outcomes by encouraging desirable behaviours.

Keywords

Project Success Framework; IPD; Industrial Construction; Machine Learning; Mixed Methods Research

Introduction

Industrial projects are challenging due to their high complexity, multiple stakeholders with varying business objectives, evolving technology, construction in adverse environments, and other complications(Yun and Jung, 2017; Barutha, et al., 2021). Due to their complexities, industrial projects often fail to meet their objectives(Merrow, 2011). These failures can lead to substantial financial losses, missed market opportunities, strained business relationships, and even legal disputes. One study found that 70% of industrial megaprojects experience cost overruns or fail to meet project objectives entirely (Flyvbjerg, 2014). The traditional approach to delivering industrial projects has normally followed either design-bid-build or design-build methods (Ahmed and El-Sayegh, 2021). Under a traditional framework, owners approve linear phasing from initial design through separate contracting and finally construction in sequence. Researchers point out downsides to conventional industrial project structures, including lack of early contractor involvement in design choices, misalignment among stakeholder priorities, and adversarial business affairs (Walker and Rowlinson, 2019). This fragmented arrangement tends to provide few incentives for collaboration, leaving projects exposed to delays, cost escalations, and friction between parties (Jones, 2014; Ahmed and El-Sayegh, 2021).

The American Institute of Architects (AIA) introduced the IPD technique to enhance construction project delivery systems by promoting smooth integration and collaboration among project participants. (AIA, 2007; Whang, Park and Kim, 2019). The core of IPD is fostering collaborative, unified, and efficient project teams, where all parties involved jointly agree on goals and commit to decisions that benefit the project's outcomes, rather than focusing on individual objectives (Rodrigues and Lindhard, 2021). Several studies (El Asmar, Hanna and Loh, 2013; González-Boubeta and Prado-Prado, 2020) indicated that (IPD) has seen the greatest adoption in healthcare and commercial construction projects and concluded that IPD initiatives demonstrated statistically significant improvements in key performances. While industrial construction has lagged other sectors in employing IPD methods, industrial project leaders are showing rising levels of interest amid challenges in delivering facilities on a budget (Al Subaih, 2015). There are several studies aimed at evaluating the impact of integrated project delivery (IPD) on efficiency, cost, schedule, quality, safety, and other outcomes across construction sectors. Metrics have been developed to measure IPD performance in areas including healthcare (El Asmar, Hanna and Loh, 2013), commercial construction (Hanna, 2016), infrastructure projects (Bapat, Sarkar and Gujar, 2021), and trade contractors (Iwanski, 2013). However, there remains minimal research specifically concentrated on assessing IPD effectiveness for industrial facilities.

As industrial sector clients are increasingly interested in integrated project delivery methods, it becomes imperative to assess whether such methods offer superior outcomes as compared to traditional delivery methods (Barutha, et al., 2021). However, at present, there is a lack of comprehensive frameworks that allow for a comparison of the performance of industrial projects executed under different delivery methods. Without a dedicated framework to evaluate industrial IPD, owners lose visibility into whether this collaborative approach is achieving better outcomes compared to traditional project execution methods in terms of schedule, cost savings, quality, safety, sustainability, client satisfaction, and other dimensions prioritized by the industry. Therefore, the overarching goal of this research is to establish a project success framework (PSF) that will allow industrial sector clients to evaluate the performance of their conventional projects, thereby enabling them to make more informed decisions regarding the effectiveness of collaborative delivery methods in this domain. This study aims to not only establish a PSF but also to develop a suitable machine-learning model that can accurately identify the degree of collaboration and integration (C&I) exhibited in industrial projects. The machine learning (ML) model is designed to operate based on the project performance in each of the success factors. The goal is to create a model that can provide insights into the level of collaboration and integration within the project, ultimately aiding in the improvement of project outcomes.

Literature Review on Project Success

The choice of project delivery approach is sometimes determined by the project's nature and past experiences. A recent study synthesized the research trends and gaps in IPD within the construction industry, highlighting the growing trend in IPD research and the need for a more comprehensive understanding on project success criteria (Arar and Poirier, 2022). During the early phases of a project, owners are tasked with making numerous critical project delivery decisions that can significantly impact project performance. Consequently, there is a strong interest among project owners to gain a better understanding of the intricate relationships that exist between these key decisions and their overall impact on project success (Esmaeili, et al., 2013). There exist two distinct research streams on the topic of project success, namely project success criteria and project success factors (Bannerman, 2008; Han, et al., 2012; Bapat, Sarkar and Gujar, 2021). The former aims to understand how project success is evaluated by focusing on the information necessary to determine whether a project was successful or not. On the other hand, the latter examines the factors that contribute to a project's success and aims to elucidate why different projects may have varying levels of success. As the objective of this paper is to develop a framework for evaluating the success of industrial projects, the current literature review primarily focuses on the first research stream. Specifically, the review commences by exploring conceptual models of project success and subsequently delves into practical approaches that have been employed to assess the success of industrial construction projects. Another study focused on critical success factors for implementing IPD in the construction industry missed providing an analysis of critical success factors specifically for industrial IPD, which have unique requirements compared to other construction sectors (Whang, Park and Kim, 2019). A study evaluating the performance of IPD on complex building projects missed providing a specific performance analysis related to industrial projects (Mesa, Molenaar and Alarcón, 2016).

While existing research has explored integrated project delivery success factors across construction domains, there remain gaps in defining appropriate evaluation criteria tailored to industrial sector projects. Recent studies (Bilbo, et al., 2015; Mesa, Molenaar and Alarcón, 2016; 2019; Yu, et al., 2019; Elghaish, et al., 2020) have worked to identify key IPD performance indicators and success factors from various lenses. However, these analyses centred on buildings and infrastructure, without a specific focus on the unique scale, complexity, and stakeholder dynamics related to large industrial manufacturing, production, and processing facilities. The "iron triangle" concept, which encompasses cost, time, and quality, has been recognized as the fundamental criteria for evaluating project success by several researchers (Pinto and Slevin, 1988; Atkinson, 1999; Chua, Kog and Loh, 1999; Lim and Zain, 1999; Al-Tmeemy, Abdul-Rahman and Harun, 2011; Müller and Jugdev, 2012). Atkinson suggested the "iron triangle" approach to evaluating project success reduces success to just cost, time, and quality is not entirely accurate and can lead to Type 2 errors. Instead, he suggested a two-dimensional approach to success, which includes both the delivery stage and the post-delivery stage. The delivery stage relates to traditional outcomes of meeting cost, time, and quality targets, while the post-delivery stage dimension focuses on providing benefits to stakeholders, user satisfaction, and meeting customer needs. Pinto and Slevin's influential work on project success has emphasized the importance of evaluating project success. They have proposed a conceptual model that has two primary dimensions: the project and the client. The project dimension focuses on the project meeting cost, time, and quality targets, while the client dimension evaluates whether the project meets user satisfaction, addresses problems, and leads to effective decision-making. According to Pinto and Slevin (1988), project managers often have their bonuses, promotions, and career progression determined by how successful they are in delivering projects. (Shrnhur, Levy and Dvir, 1997) proposed four dimensions for project success: Project Efficiency, Impact on the Customer, Business and Direct Success, and Preparing for the Future. Lim and Zain (1999) developed a model that divides project success into micro and macro perspectives. The micro perspective assesses the results of the construction phase while the macro perspective examines whether the project fulfilled user or stakeholder needs.

Baccarini (1999) categorized project success into two dimensions: project management success and product success. The former relates to meeting the project's budget, schedule, and quality objectives, while the latter corresponds to satisfying the stakeholders' needs and achieving the project's original goals. Bannerman (2008) stated that the challenge in determining project success lies in deciding if it is a means to an end or an end in itself, with success being measured accordingly based on strategic goals, end-user needs, or traditional characteristics such as time, cost, and quality. The above models offer insights into what outcomes a project should achieve to be successful but failed to provide specific measurements for certain criteria. Some researchers demonstrated more practical approaches to measuring construction project success. The KPI Working Group used a two-step framework to measure civil construction project success with Key Result Areas (KRAs) and Key Performance Indicators (KPIs) (Caldwell, 2007). Six KRAs are defined in Table 1.

Table 1. The KPI working group's KRAs and their definitions

KRAs	Definition
Client Satisfaction	Measures how satisfied the client was with the quality of the finished product and the service (of the whole project team). Usually measured at or shortly after completion and handover.
Defects	Measures the degree to which the completed facility was free from defects that impacted the client. Usually measured at the point the project is offered for handover.
Cost	Measures how well out-turn costs compared with original estimates.
Time	Measures how closely the project was delivered to the original timetable.
Safety	A measure of the number of Lost Time Incidents per 200,000 hours worked. Equivalent to 100 Full Time Equivalent (FTE) employees.
Profitability	Measures company profit before tax and interest as a percentage of sales.

El Asmar, Hanna and Loh (2016) created the project quarterback rating (PQR) to assess building project success. The PQR is a weighted sum of KRAs and KPIs, with seven KRAs and multiple KPIs in each. The Construction Industry Institute (CII) has also developed success frameworks, including a linear weighted sum of two KRAs and four KPIs by their Pre-Project Planning Research Team (RT-039) in 1994 (Gibson and Hamilton, 1994) shown in Table 2.

Table 2. CII RT-039's project success index

KRA	KPI	Measurement
Project Success	Budget Achievement	(Actual cost – Estimated cost) / Estimated cost
	Schedule Achievement	(Actual duration – Estimated cost) / Estimated duration
Operating Success	Design Capacity Attained	(Actual output rate – Estimated output rate) / Estimated output rate
	Plant Utilization	Number of plants produces products in six months / 182

A comprehensive evaluation framework for project success is necessary to provide an objective and quantifiable measure of success. Although some studies have developed project success evaluation frameworks, they have primarily focused on the commercial building sector, where the IPD originated from. Because of the differences between commercial buildings and industrial projects, project success measurement needs to be adjusted. The unique characteristics of industrial construction projects, such as their complexity, scale, and technological requirements, suggest that they require a specific project success evaluation framework of IPD in this sector. The lack of standardized benchmarks to evaluate the performance of IPD in industrial projects relative to traditional methods persists as a barrier. To bridge this gap, this study developed a practical PSF for IPD on industrial projects. The study's goal is to establish a PSF that will support industrial sector clients in evaluating the performance of their conventional projects, thereby enabling them to make more informed decisions regarding the effectiveness of collaborative delivery methods in this domain. The PSF developed in this study allows researchers and practitioners to empirically evaluate the integrated project delivery's efficacy on key industrial project outcomes.

Methodology

This research implemented a mixed methods design to build and validate the PSF for industrial projects, as depicted in Figure1. This methodology was chosen for several reasons. First, the concept of project success is a complex phenomenon, and the exploration of complex phenomena is well suited to qualitative methods (Creswell and Plano Clark, 2018). Second, survey-dominant research methods have several drawbacks in construction research, such as low response rates, long response times, and long development periods (Gibson and Whittington, 2010). The explanatory sequential design facilitates Gibson data analysis and integration at various stages of the study. It involves the collection and analysis of quantitative survey data, followed by the collection and analysis of qualitative data. By using the qualitative results to complement and clarify the quantitative findings, the researcher can develop a more thorough understanding of the research topic. The integration of qualitative results with quantitative findings enables the researcher to enhance and elucidate their understanding of the research topic in a more comprehensive manner (Creswell and Plano Clark, 2018). This study was approved by the Institutional Review Board (IRB) of the University of Nebraska-Lincoln with the IRB number 20191219809EX.

PHASE A: SUBJECT MATTER EXPERTS SURVEY

In the first phase of this study, a survey was conducted to investigate the success factors used to measure project success. The survey aimed to identify the success factors, measure their level of importance, evaluate the project performance, and identify the collaboration and integration level for the most collaborative and integrated project compared to a typical project. The survey consisted of 13 items based on success factors identified in the study (CII RT-341, 2019) dedicated to IPD for industrial projects. A web-

Phase	Procedure	Product
Phase A. Subject matter experts survey	• Web-based survey was developed, and 87 complete responses were collected based on the purposive sampling.	• Identified & measured the importance of success factors • Developed , compared, and identified the best ML model to predict C&I level based on performance of success factors.
Phase B. Research Charrette	• Total 12 participants including five client representative, four consultants, three academicians were joined for the three-hour workshop. • 14 KRAs identified RT-341 was provided to the participants in the workshop. • Each KRA was discussed as it relates to success evaluation of industrial projects.	• Finalized and defined 11 KRAs • 27 KPIs and formula for each KPI were developed.
Phase C. Validation survey	• Web-based survey was developed to validate the relationship between the KPIs and KRAs defined from the research charrette. • 41 responses were collected based on the purposive sampling. • Each respondent categorized each KPI into one of the KRAs.	• Each of 27 KPIs was categorized under one of the 11 KRAs.
Integration of results	• Integration and explanation of results.	• Project Success Framework (PSF) for industrial projects was developed.

Figure 1. Research Methodology

based questionnaire was developed in Qualtrics and distributed to a purposive sample of 87 construction professionals, including project managers, owners, project engineers, architects, and executives. Participants were asked to pick the most collaborative and integrated project they have experienced recently as compared to the most typical projects they have experienced. Also, participants were asked to "rate the level of project performance for each success factor on the most collaborative & integrated project you previously identified as compared to a typical project" on a 5-point Likert-type scale (1=very little, 5=very high) if it applied to the project, to ensure that it measured the envisioned success factors. In addition, participants were asked to list the strategies that were used in the collaborative and integrated project. The intensity level of each strategy used on the project was collected, and the average intensity was calculated using the formula in Equation 1. It is calculated by summing the intensity of these strategies and then dividing by the number of strategies used. This measure provides insight into the overall level of collaboration and integration implemented.

$$\textit{Average Intensity} = \frac{\textit{Sum of the Intensity of Collaboration \& Integraion Strategies}}{\textit{Numer of Strategies Used}} \qquad \text{(Equation 1)}$$

The machine learning (ML) model, an emerging area of artificial intelligence (AI), was proposed to prioritize result prediction over variable relationship determination, in contrast to traditional statistical methods predictions (Rajula, et al., 2020; Sanni-Anibire, Zin and Olatunji, 2022; Ghimire, Kim and Acharya, 2024). This ML-based application tool embedding the PSF was designed to enhance the practicality and applicability of the framework, ultimately empowering industrial sector clients to make more informed decisions regarding the effectiveness of collaborative delivery methods in this domain. Therefore, the survey and the subsequent development of the ML-based application tool were integral to the study's goal of establishing a comprehensive framework for assessing project success in the industrial sector.

In this study, we aimed to develop a machine learning classification system that can predict whether the project demands high vs low (1 vs 0) collaboration and integration based on categorical input variables. The use of machine learning techniques for the survey type of data is commonly used in the construction field and has been utilized by various studies (Ayhan, Dikmen and Talat Birgonul, 2021; Sanni-Anibire, Zin and Olatunji, 2022).To investigate the relationship between the performance of project success factors and the level of collaboration and integration, machine learning models were deployed. It involves four steps in the machine learning application: (1) data collection through a subject matter expert survey, (2) data pre-processing, (3) development of machine learning models, and (4) model performance and evaluation.

The choice of machine learning models, including Decision Tree (DT), Naïve Bayes (NB), Logistic Regression (LR), and k-Nearest Neighbors (k-NN), was based on their suitability for the specific objectives of the study. These models were selected due to their proven effectiveness in handling classification tasks and their potential to provide accurate predictions. The tree-based models, DT, were chosen for their ability to handle complex, non-linear relationships within data and to identify patterns and similarities within the data, respectively. On the other hand, the non-tree-based models were selected for their efficiency in handling large datasets and their ability to provide probabilistic predictions, which can be valuable in assessing the collaboration and integration based on various input variables.

Decision Tree (DT) is a prevalent supervised learning algorithm that recursively partitions data on feature values to construct a predictive model represented as a tree structure (Janikow, 1998; Charbuty and Abdulazeez, 2021; Ghimire, et al., 2023; Pokharel and Ghimire, 2023). NB and LR models are well-established for their efficiency in handling large datasets and their ability to provide probabilistic predictions, which can be valuable in assessing the likelihood of project success based on various input variables. The NB classifier applies Bayes' theorem to machine learning by assuming conditional feature independence to predict target class probabilities (Rish, 2001; Leung, 2007). Mathematically, Bayes' theorem estimates the likelihood of an event based on its prior probability and the conditional probabilities associated with related evidence. LR, on the other hand, applies sigmoid functions to regression analysis to predict the probability of categorical outcomes based on input data features (Tien Bui, et al., 2019). In contrast to linear regression which forecasts continuous values, LR handles classification tasks by estimating discrete class membership likelihoods given evidence conditions. Additionally, the k-Nearest Neighbors model was chosen for its capability to identify patterns and similarities within the data, which is essential for understanding the interrelationships between different project success factors in industrial contexts. The k-NN algorithm is a simple, non-parametric classifier that categorizes new cases based on a majority vote of the nearest examples in the training feature space (Guo, et al., 2003; Lee, et al., 2017). k-NN relies on distance metrics like Euclidean distance to find the k most similar instances, with new data points classified to the dominant class among the nearest neighbours. The Synthetic Minority Oversampling Technique (SMOTE) applied in our model development, is commonly used to balance class distributions by oversampling minority cases in imbalanced datasets (Fernandez, et al., 2018). Therefore, the selection of these machine learning models was driven by their specific strengths in addressing the complexities and nuances of industrial project success evaluation, as well as their potential to provide valuable insights for decision-making in this domain.

PHASE B: RESEARCH CHARRETTE

A structured workshop or "research charrette" has become a popular research method in construction studies. According to (Gibson and Whittington, 2010), research charrettes are an effective way to explore a topic as they combine the best aspects of surveys, interviews, and focus groups within a short timeframe. This method has been utilized in several construction studies, including Esmaeili, et al.'s (2013) research on project success for building projects and Cho and Gibson's (2012) development of the PDRI for building projects. Due to the complex and multidimensional nature of evaluating industrial project success, a research charrette was considered a suitable approach for this study. Gibson and Whittington explain that the construction industry typically does not permit probabilistic sampling because it is logistically impossible to define a sample frame. The research charrette used a purposive sample of 12 participants, including five client representatives, four consultants, and three academics. A three-hour research charrette was conducted in Lincoln, Nebraska. The workshop began with a review of the objective and the relationship between KRAs, KPIs, and project success. To identify common KRAs used to define construction project success, CII's RT-341 conducted a literature review (CII RT-341, 2019), which was used as a starting point for the group discussion. Participants received a printed list of the 14 KRAs identified by RT-341 (CII RT-341, 2019) and discussed each KRA in relation to evaluating industrial project success. RT-341 identified 14 KRAs from a comprehensive literature review, incorporating articles that addressed construction project performance outcomes and multiple key performance indicators in the industrial sector. A group consensus was necessary to include KRAs in the final framework. Upon completing the KRA task, the group proceeded to discuss KPIs that represent each of the KRAs. Each participant was provided with a list of previously used KPIs to aid in the discussion. The KPIs needed to be measurable and applicable to all industrial projects. Similar to the KRAs, a group consensus was required for the inclusion of any KPIs in the final framework. Notes taken during the research charrette were used to create the framework, which was later reviewed by the research team.

PHASE C: VALIDATION SURVEY

A targeted survey was conducted to validate the relationship between KPIs and KRAs and ensure their measurability across projects. Respondents were asked to categorize each KPI into one of the KRAs to determine if the KPI is representative of the key areas of project success. The survey was web-based and developed using Qualtrics. Each research team member, including 9 members- 5 client representatives and 4 consultants, was instructed to send the survey to 10 people in their professional network. To prevent order bias, the list of KPIs was randomly displayed to each participant, and definitions were provided as downloadable PDFs to assist in their responses.

Analysis and Result

SUBJECT MATTER EXPERT SURVEY RESULTS

The subject matter expert survey results indicated that safety, quality, overall client satisfaction, and early schedule certainty are the four most selected success factors (Table 3). Out of the 87 respondents, the majority of the participants identified safety (85 reports, 98% of the sample), quality (80 reports, 92% of the sample), client satisfaction (59 reports, 68% of the sample), and early schedule certainty (50 reports, 57% of the sample) as the four success factors with the highest level of importance to overall project success.

Respondents were also asked to rate the level of project performance, on a 5-point Likert-type scale (1=very little, 5=very high) if it applied to the project, for each success factor listed in Table 3, on the most collaborative and integrated project as compared to a typical project. In addition, participants were asked to list the strategies that were used in the collaborative and integrated project. The list of the seven most important strategies mentioned in Table 4 was developed based on the literature review (AIA, 2007; Barutha, et al., 2021, Government of Western Australia, 2010).

MACHINE LEARNING MODELS DEVELOPMENT

Data pre-processing

This stage of the research involved the utilization of data obtained from a survey of subject matter experts. The final data set comprised 87 projects and 14 variables, 13 of which were ordinal features listed in Table 1,

Table 3. The rank of Project Success Factors

Rank	Success Factor	Coded as	Percentage
1	Safety	safe	98%
2	Quality	qual	92%
3	Overall client satisfaction	stfn	68%
4	Early schedule certainty	sche	57%
5	Early cost certainty	cost	44%
6	Profitability (contractor) / ROI (owner)	prft	45%
7	Speed to market (owner)	speed	32%
8	Ease of start-up	stup	31%
9	Facility production (revenue generation)	fclt	30%
10	Environmental	envr	28%
11	Lowest life cycle cost	lfcs	26%
12	Flexibility/adaptability of the facility	flex	25%
13	Lowest initial cost (owner)	incs	18%

Table 4. Collaboration and Integration Strategies

	Collaboration and Integration Strategies
1	Early involvement of key participants
2	Shared cost and shared reward (Pain Share Gain Share)
3	Collaborative and equitable decision-making
4	Jointly developed and validated targets
5	Negotiated risk distribution (e.g. Mutual Liability Waivers)
6	Non-traditional Owner Engineer Contractor Relational Contracting (e.g. multi-party agreement, IFOA, alliance)
7	Continuous team building and conflict resolution

while the other variable was the target variable representing the intensity of collaboration and integration. Python was chosen as the primary programming language due to its open-source nature and the availability of rich libraries such as pandas, numpy, scikit-learn, and matplotlib, which are useful for data preprocessing and developing machine learning classifiers. The ordinal variables had five levels: 0, 1, 2, 3, 4, and 5. The target variable was transformed into a binary categorical variable such that when the value exceeded the mean (17), it was coded as high collaboration and integration intensity (1), while a value less than or equal to the mean (17) was coded as low collaboration and integration intensity (0). It was noted that the data set was imbalanced, with a ratio of 1 to 0 being 59:41. Imbalanced data sets pose a challenge to traditional training criteria, as classifiers may have good accuracy for the majority class but poor accuracy for the minority class (Ganganwar, 2012). Therefore, several resampling approaches have been developed in recent years to address this issue. The Synthetic Minority Oversampling Technique (SMOTE) is a widely used and well-known oversampling algorithm that creates synthetic instances of the minority class by bootstrapping using the nearest neighbours of the sample. The SMOTE technique was applied to balance the data set, resulting in a 50:50 ratio for both classes. The data set was balanced only on the training data set to prevent data leakage (Schlögl, et al., 2019). As a result, a final data set of (102, 14) was chosen for the deployment of machine learning algorithms, with a training data shape of (84, 14) and a test data shape of (18, 14).

Splitting data into a training set and a test set is a widely used approach in machine learning. In our study, the complete data set was randomly divided into two sets; the training set, which comprised 80% of the data, and the test set, which contained the remaining 20%. The variables were separated into two categories: feature variables (X) and the target variable (Y). The feature variables were as follows: 'safe', 'qual', 'cost', 'sche', 'stfn', 'prft', 'speed', 'incs', 'lfcs', 'flex', 'fclt', 'stup,' and 'envr'. These features represented the attributes of the project performance and were used as inputs to the machine learning models. The target variable, denoted as 'totic', represented the intensity of collaboration and integration, and was the variable that the models aimed to predict. We deployed four machine learning models namely Decision Tree (DT), Naïve Bayes (NB), Logistic Regression (LR), k -Nearest Neighbors (k-NN)

Hyperparameters Tuning

Optimizing the hyperparameters of machine learning models is essential for achieving robust performance results. Default hyperparameter settings may not be optimal and require additional attention during this critical step in the machine-learning process(Schratz, et al., 2019). There are various search strategies available to find the best and most robust parameter or set of parameters for an algorithm on a given problem, with GridSearchCV being a commonly used technique (*Sklearn.Model_selection. GridSearch CV*). Grid search is a systematic approach to parameter tuning, where a model is built and evaluated for every combination of algorithm parameters specified in a grid (Ranjan, Kumar Verma and Radhika, 2019). The optimal parameters identified using GridSearchCV are presented in Table#.

Model Performance Evaluation

Model performance evaluation is an important process to understand how well a model is able to learn from the training data and make accurate predictions on new, unseen data. One can assess the accuracy of classification by calculating the number of class examples that were correctly identified as such (true positives-TP), the number of examples that were correctly identified as not belonging to the class (true negatives-TN), and the number of examples that were either wrongly assigned to the class (false positives-FP) or not identified as class examples (false negatives-FN) (Sokolova and Lapalme, 2009). There are various metrics that can be used to evaluate the performance of a classifier, depending on the specific problem and the requirements of the application. Accuracy, Precision, recall, F1-Score, and Area under the Curve (AUC) - Reception Operating Characteristics (ROC) curve are the most common evaluation metrics for classification models. And the metrics were calculated as follows (Wang, Shao and Tiong, 2021; Sanni-Anibire, Zin and Olatunji, 2022):

$$\text{Classification Accuracy} = \frac{TP + TN}{N}$$

Where, N= total number of samples

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

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

F1-score is the weighted harmonic mean of precision and recall. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) at different classification thresholds. The area under the curve (AUC) is measured in AUC-ROC, and a classifier with an AUC-ROC of 1 is considered perfect.

Four machine learning classifiers were evaluated, and their performance is summarized in Table 5. The decision tree (DT) model achieved the highest accuracy of 0.72, followed by the Naive Bayes (NB) model with an accuracy of 0.67. This means that the DT model was able to correctly classify 72% of the instances in the dataset. The logistic regression (LR) and k-nearest neighbour (kNN) models achieved an accuracy of 0.56. The precision, recall, and F1-score for the DT model were 0.75, 0.72, and 0.71, respectively, which were the best score among four models. In the context of the DT model, precision of 0.75 means that out of all the instances predicted as positive by the DT model, 75% of them were positive. Recall, on the other hand, measures the proportion of true positive predictions among all actual positive instances in the dataset. In the context of the DT model, a recall of 0.72 means that the DT model was able to identify 72% of all the actual positive instances in the dataset. F1-score is the harmonic mean of precision and recall and provides a balanced measure of both metrics. In the context of the DT model, an F1 score of 0.71 indicates that the model had a relatively balanced performance in terms of precision and recall.

Table 5. ML Model Performance Evaluation Metrics

ML Model	Best Hyperparameters	Accuracy	Precision	Recall	F1-score	AUC-ROC
DT	{'criterion': 'entropy', 'max_depth': 10, 'max_features': 'sqrt', 'min_samples_leaf': 2, 'min_samples_split': 2}	0.72	0.75	0.72	0.71	0.80
NB	{'var_smoothing': 1e-09}	0.67	0.68	0.67	0.67	0.75
LR	{'C': 0.01, 'class_weight': None, 'max_iter': 100, 'penalty': 'l2', 'solver': 'liblinear'}	0.56	0.56	0.56	0.56	0.65
k-NN	{'algorithm': 'auto', 'n_neighbors': 1, 'p': 2, 'weights': 'uniform'}	0.56	0.64	0.56	0.52	0.59

The AUC-ROC curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) for different threshold values of the model's output (Figure 2). The AUC-ROC score is then calculated as the area under this curve. In our case, the DT model had the highest AUC-ROC score of 0.80, indicating that it had a better ability to distinguish between positive and negative instances compared to the other models. The NB model had an AUC-ROC score of 0.75, which is also relatively high. The LR model had a lower AUC-ROC score of 0.65, while the k-NN model had the lowest AUC-ROC score of 0.59, indicating that these models had relatively poorer discrimination ability.

There can be several reasons why the decision tree (DT) model performed best in the given scenario. One reason is that the DT model is a non-parametric and interpretable algorithm that can handle both categorical and numerical data, making it suitable for the dataset. Furthermore, decision trees can handle non-linear relationships between features, therefore the DT model has been able to identify and utilize these relationships to better classify instances, resulting in its superior performance compared to other models. The DT model, which was developed to predict collaboration and integration intensity during

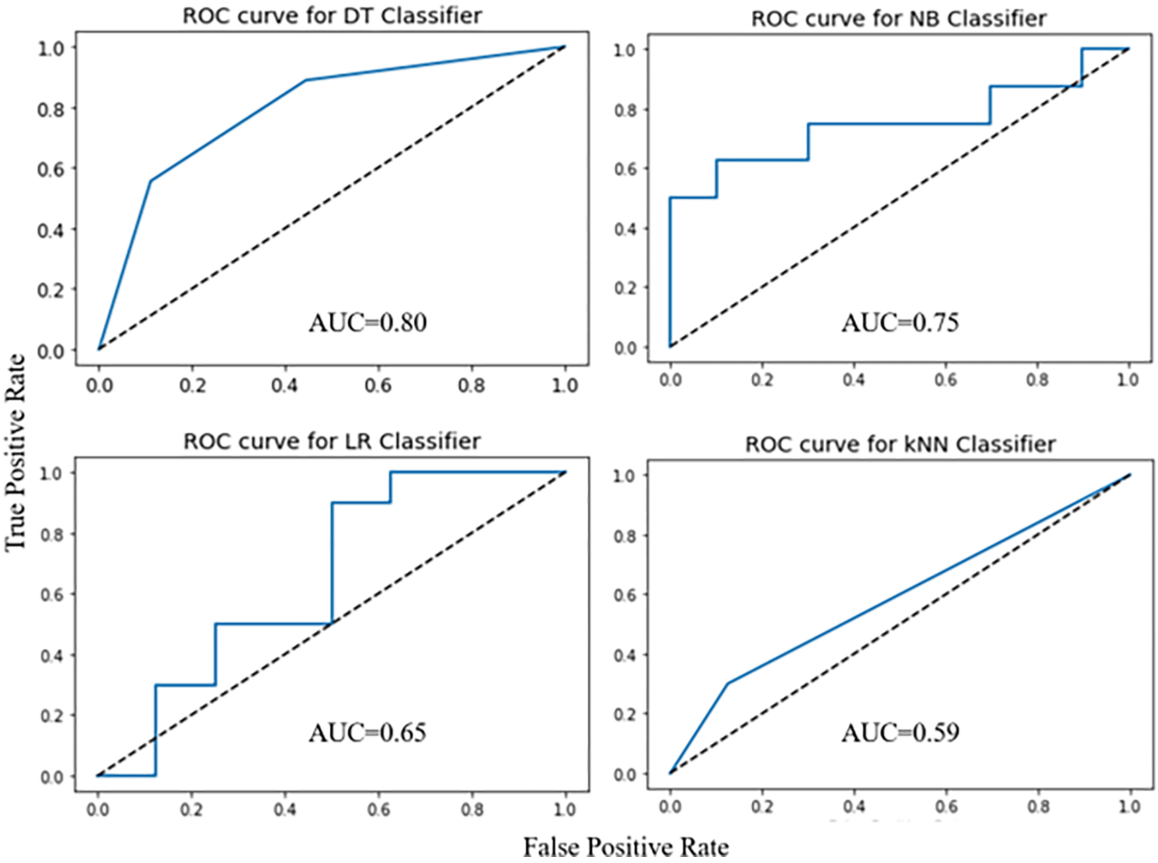

Figure 2. AUC-ROC of ML Models

the early phase of industrial construction projects, has been identified as the most effective model for enhancing the efficiency of the decision-making process. It is believed by the authors that the proposed DT model can be employed by various project stakeholders such as capital project owners, project management organizations, contractors, designers, and project participants to accurately forecast the collaboration and integration intensity based on projected performance on 13 project success factors.

RESEARCH CHARRETTE RESULTS

The first objective of the research charrette was to finalize a comprehensive list of KRAs that can be used to compare the success of an industrial project executed under different delivery methods. Figure 3 shows the 11 KRAs that participants from the research charrette decided would be important to compare.

Definitions for each of the final 11 KRAs developed during the research charrette are presented in Table 6.

VALIDATION OF THE KPIs

To validate that the KPIs are representative of their higher-order KRA, a targeted survey was distributed to professionals in the industrial sector asking them to categorize each KPI into its respective KRA. After data cleaning, a total of 41 complete responses were received. The respondents' average number of years of experience was 21.5 years, with a standard deviation of 11.3 years. The sample population represented 21 clients, 10 contractors, and 10 engineers/consultants. Out of 41 respondents, 32 respondents primarily worked in the heavy industrial sector, 5 in the light industrial sector, 4 in power, utilities, and infrastructure.

Figure 3. The 11 KRAs that shape the success of industrial projects.

Each respondent was asked to categorize the list of KPIs into one of the 11 KRAs. Table 7 presents each KPI, the KRA it was categorized into during the research charrette (expected), and the KRA that was categorized most frequently in the targeted survey. Table 7 also indicates if there was a match between the expected and found KRA categorization.

Table 6. Definition of the 11 KRAs for industrial projects

KRA	Definition
Cost Competitiveness	This construct represents a measure of how competitively the project was priced compared to the typical market conditions at the time the project was delivered. The objective of this KRA can be thought of as capturing the "value for money" that the owner receives.
Cost Certainty	Cost certainty represents a measure of how well the project's actual costs met the project's early cost targets.
Schedule Competitiveness	Schedule competitiveness is a measure of how competitively the project's schedule was compared to typical market conditions at the time the project was delivered.
Schedule Certainty	Schedule certainty represents a measure of how well the project's actual schedule met the project's early schedule targets.
Quality	Quality represents a measure of how well the project's products and services complied with its plans and specifications. This is not to be confused with the quality of the finishes used on the project.
Safety	Safety represents a measure of the frequency of recordable safety incidents that occurred on the project.
Project Functional Objectives	A project's functional objectives are a measure of how well the project achieved the client's functional objectives as defined in the client's business case that was used to justify the project's funding.
Project Financial Objectives	Financial objectives are a measure of how well the project achieved the financial objectives of all of the major participants in the project (typically the client, contractor, and engineer).

Table 6. continued

KRA	Definition
External Stakeholder Impacts	External stakeholder impacts are a measure of how much the execution of the project impacted external stakeholders. External stakeholders can include the public, the client's end users, or the client's internal operations.
Environmental Impacts	Environmental impacts are a measure of the frequency and magnitude of recordable environmental events that occurred on the project.
Change Management	The change management construct represents a measure of the frequency, size, and duration of changes that occurred on the project.

Table 7. KPI categorization from the targeted survey

KPIs	KRAs		Match
	Expected	Top Survey Categorization	
Contingency Index	Cost Competitiveness	Cost Certainty	×
Cost Efficiency	Cost Competitiveness	Cost Competitiveness	●
Direct Work Rate	Cost Competitiveness	-	×
Productivity	Cost Competitiveness	Cost Competitiveness	●
Contingency Used %	Cost Certainty	Cost Certainty	●
Cost Variation	Cost Certainty	Cost Certainty	●
Buffer Index	Schedule Competitiveness	Schedule Certainty	×
Time per Unit	Schedule Competitiveness	Schedule Competitiveness	●
Schedule Variation	Schedule Certainty	Schedule Certainty	●
Construction Defects	Quality	Quality	●
Design Defects	Quality	Quality	●
Non-conformance reports	Quality	Quality	●
Quality Performance Rating	Quality	Quality	●
Commissioning Time	Quality	Schedule Certainty	×
DART Rate	Safety	Safety	●
TRIR	Safety	Safety	●
Goal Achievement	Project Functional Objectives	Project Functional Objectives	●

Table 7. continued

KPIs	KRAs		Match
	Expected	Top Survey Categorization	
Contractor Financial Objective Realization	Project Financial Objectives	Project Financial Objectives	●
Owner Financial Objective Realization	Project Financial Objectives	Project Financial Objectives	●
Complaints	External Stakeholder Impact	External Stakeholder Impact	●
External Stakeholder Impact	External Stakeholder Impact	External Stakeholder Impact	●
Notice of Violation	Environmental	-	×
Recordable Environmental Events	Environmental	Environmental	●
Change Cost Index	Change Management	Cost Certainty	×
Change Time Index	Change Management	Change Management	●
Non-owner Initiated Changes	Change Management	Change Management	●
Owner Initiated Changes	Change Management	Change Management	●
Speed of Change Approval	Change Management	Change Management	●

Two of the KPIs (direct work rate and notice of violation) were categorized as "not important" by more than 15% of the respondents. Upon recommendations from the research team members, these two KPIs were removed from the PSF. The final PSF for industrial projects is presented in Table 8.

Table 8. Project success framework for industrial projects

KRAs	KPIs	Formula
Cost Competitiveness	Cost Efficiency	Total project cost / Capacity of facility[a]
	Contingency Index	Project contingency / Total project budget[b]
Cost Certainty	Cost Variation	(Actual project cost – Total project budget[b]) / Total project budget[b]
Schedule Competitiveness	Schedule Efficiency	Total project duration / Capacity of facility[a]
	Buffer Index	Project schedule buffer / Predicted project duration[c]

Table 8. continued

KRAs	KPIs	Formula
Schedule Certainty	Schedule Variation	(Actual project duration - Predicted project duration[c]) / Predicted project duration[c]
Quality	QPR	(Σ N × w / Number of employee labour hours) × 200,000 N = number of unplanned quality events (variation, defect, or failure) w = weighted severity level of each event
	Design Defects	Total number of design errors / Total number of drawings
	Construction Defects	Total number of punch items at mechanical completion
	NCRs	Total number of non-conformance reports
	Commissioning Duration	Substantial completion – Mechanical completion
Safety	TRIR	(Number of OSHA recordable cases / Number of employee labour hours) ✕ 200,000
	DART	(Number of DART incidents / Number of employee Labor hours) ✕ 200,000
Project Functional Objectives	Goal Achievement	1-7 Likert scale (1 = This project achieved none of the functional objectives as set out in the projects business case, 7 = This project achieved all of the functional objectives as set out in the projects business case)
Project Financial Objectives	Owner Financial Objective Realization	1-7 Likert scale (1 = This project achieved none of the financial objectives as set out in the projects business case, 7 = This project achieved all of the financial objectives as set out in the projects business case)
	Contractor Financial Objective Realization	1-7 Likert scale (1 = This project achieved none of the financial objectives as set out in the projects business case, 7 = This project achieved all of the financial objectives as set out in the projects business case)
External Stakeholder Impacts	Complaints	Total number of complaints received during the execution of the project
	External Stakeholder Impact	1-7 Likert scale (1 = This project had a maximum impact on its external stakeholders, 7 = This project had minimal impact on its external stakeholders) (External stakeholders could include: local businesses in the surrounding area, local residents, and other divisions in a facility during a renovation.)

Table 8. continued

KRAs	KPIs	Formula
Environmental Impacts	Recordable Environmental Events	Total number of recordable environmental events
Change Management	Change Cost Index	Approved project development change costs / Total project budget[a]
	Change time index	Approved project development change duration / Predicted project duration[c]
	Number of Non-owner Changes	Total number of changes initiated by non-owner parties
	Number of Owner Initiated Changes	Total number of changes initiated by the owner
	Speed of Change Approval	Average duration that RFIs are open in weeks

[a] The capacity of the facility should be a comparable industry metric for the facility (e.g. kWh, tonnes per day, bpd, etc.)

[b] The total project budget must include project contingency and be adjusted for scope changes.

[c] The predicted project duration must include project buffers and be adjusted for scope changes.

Discussions

The objective of this paper was to develop a framework that can evaluate the effectiveness of integrated project delivery methods for industrial projects. 11 KRAs were defined as essential outcomes that can be used to define the success of an industrial project. Specific KPIs were also identified that can be used to consistently measure the performance of a project in each of the 11 KRAs. The PSF was designed to be flexible, so the KPIs that measure each KRA can change, enabling clients to replace them with metrics that are important to their specific project, sector, or business. The KPIs also provide a specific way for clients to develop their non-cost incentive mechanisms in the shared risk/reward commercial model. One of the time-consuming activities associated with developing a collaborative contract is to identify suitable metrics to incentivize performance. The PSF provides clients with a "menu" of KPIs that can be used to develop these agreements.

KEY RESULT AREAS

Two dimensions have consistently emerged from existing conceptual models of project success. The first dimension pertains to the success of the management of the project against predetermined targets: usually regarding cost, time, and quality. These outcomes can be evaluated immediately after the conclusion of the project. In Baccarini's terms, this is "project management success." The second dimension of success relates to how well the project fulfills the client's original need for the project. This type of success must be evaluated at some time after the delivery of the project and is referred to as "product success." The 11 KRAs presented in the PSF can be separated into the project management and product success dimensions. The PSF contains 9 KRAs that relate to project management success and 2 KRAs that relate to product success as shown in Figure 4.

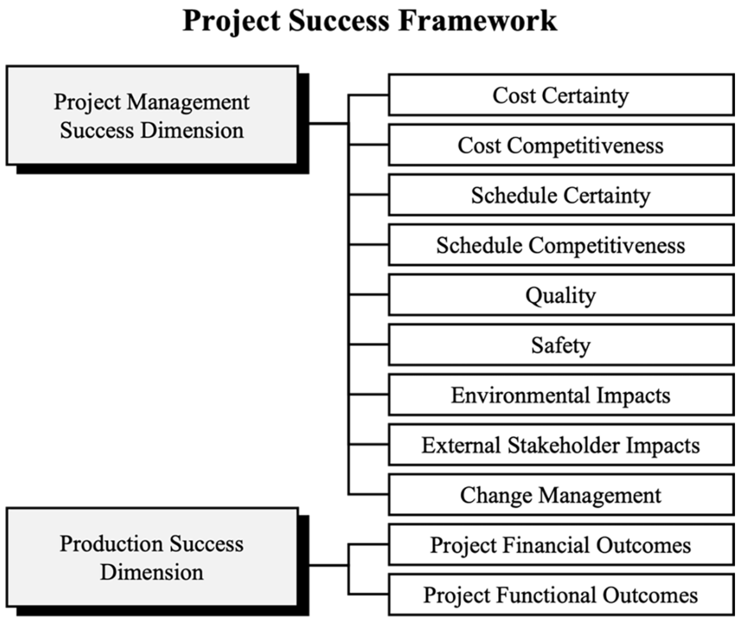

Figure 4. PSF KRAs separated into project management and product success dimensions.

The PSF places greater emphasis on outcomes in the project management dimension of overall project success. The PSF was developed with the purpose of comparing the success of projects delivered under different delivery methods. This indicates that project delivery methods have a greater influence over project management outcomes, rather than how well a project achieves its client's organizational needs. This makes intuitive sense, as a project delivery method is, in essence, the process by which a client realizes a project, not a way for clients to identify projects that will fulfill their organizational needs. Baccarini states that a project's product success will trump its project management success (Baccarini, 1999). In practical terms, this means that projects that meet all their cost, schedule, and quality targets may still be considered a failure if they fail to meet their client's long-term organizational needs. Therefore, clients may be inclined to evaluate the performance of a delivery method based on a project's product success. Doing so would be a mistake, as the delivery method has limited influence over the project product success. This shows that clients need to consciously separate project management and product outcomes when evaluating the performance of a delivery method, and comparing delivery methods should focus on comparing project management outcomes.

The PSF presents an alternative approach for evaluating a project's cost and schedule performance. Typically, cost and schedule performance relate to a project's adherence to its target budget and target finish date. This success framework separates cost and schedule into their respective competitiveness and certainty. Practitioners provided this recommendation because they said a project's adherence to its targets is largely dictated by how competitive those targets are. For example, a project could be significantly under budget because of an inflated target budget. By separating a project's cost certainty from its cost competitiveness, clients gain a better understanding of how successful their project was. Another reason for separating cost and schedule certainty from competitiveness is because it addresses an important criticism of collaborative delivery models. Under the shared risk/reward commercial model, non-owner parties share in the underruns of the project. This creates an implicit incentive to inflate the target cost (Henneveld, 2006; Wood and Duffield, 2009; Thomsen, et al., 2016). By separating cost certainty from cost competitiveness, clients can increase their visibility of the performance of a project and properly evaluate the effectiveness of the delivery method.

KEY PROJECT INDICATORS

The value of any comparison is dependent on the quality of information inputted into the analysis. One of the issues present with existing project success frameworks is that there is a lack of detailed instruction provided with their KPIs. For example, the traditional cost variation KPI is present in many existing project success frameworks (Chan, 2001; KPI Working Group, 2007; El Asmar, Hanna and Loh, 2013; Franz, et al., 2014; Hanna, 2016). In these frameworks, the cost variation KPI is typically defined as:

$$Cost\ Variation = \frac{Final\ Project\ Cost - Initial\ Budget}{Initial\ Budget}$$

The issue with this formula is that it does not inform the individual providing information on how to handle project change orders. Changes are a part of construction and will appear on every project. Changes will occur for a number of reasons, including differing ground conditions, inaccurate specifications, owner-initiated design changes, or errors and omissions in the drawings. The initial budget in this formula needs to reflect the cost of changes to the project. CII separates changes into two categories: project development changes and scope changes. Project development changes are defined as: "changes required to execute the original scope of work or obtain original process basis." Scope changes are defined as: "changes in the base scope of work or process basis." The PSF requires individuals to correct the initial budget for scope changes but not project development-related changes. Doing so will better reflect the performance of the project team without distorting the cost information because of changes. Additionally, this will ensure that the project information being collected is consistent and thus will improve the accuracy of comparisons that can be generated from the framework. The financial profitability indicator has been used in a variety of existing frameworks; however, its previous appearances usually only refer to the profitability of the client (Chan, 2001; Nassar and AbouRizk, 2014). The PSF includes a KPI for the financial performance of the contractor. This indicator helps to identify if a project was more competitive because the work was delivered more efficiently, or if it was simply the result of contractors reducing their profit margins.

MACHINE LEARNING APPLICATION

The machine learning models were deployed to investigate the relationship between the performance of project success factors and the level of collaboration and integration. Four machine learning models were selected due to their proven effectiveness in handling classification tasks and their potential to provide accurate predictions. These models were used to develop a machine learning classification system that can predict whether the project demands high vs low (1 vs 0) collaboration and integration based on categorical input variables. The research results indicate that the Decision Tree model achieved the highest accuracy, followed by the Naive Bayes model. The Decision Tree model was able to correctly classify 72% of the instances in the dataset, making it the most suitable model to support the decision-making process.

Conclusion

This study developed and validated a Project Success Framework that enables researchers and practitioners to empirically evaluate the effectiveness of integrated project delivery methods on important project outcomes for industrial capital projects. The benefits of the PSF are threefold. First, it provides clients with a comprehensive list of KRAs and KPIs to compare the performance of a collaboratively delivered project with one that is delivered under traditional methods. This will enable clients to accurately evaluate the effectiveness of collaborative delivery methods for industrial projects. Second, the PSF will also enable clients to make more informed decisions about the application of all project delivery methods. Several efforts have been made to develop project delivery method selection tools, such as the CII's

"Project Delivery and Contracting Strategy" tool, and the United States Federal Highway "Contracting Alternatives Suitability Evaluator," but these tools continue to rely on judgment and subjectivity to make their evaluations. The PSF provides clients with a structured approach to evaluating the performance of their projects. And third, the PSF provides clients with a "menu" of KPIs that they can use to develop their shared risk/reward commercial models. Integrated delivery methods provide clients with an opportunity to incentivize the achievement of their non-financial objectives through KPIs. However, it can be challenging to measure performance in outcomes other than cost. The PSF provides clients with a range of KPIs as well as the KRA that they will each incentivize. This will help reduce the time spent negotiating the metrics within the commercial model and ensure that their commercial models incentivize behaviour that promotes project success.

While our study provides validated PSF for industrial projects, certain limitations exist and present opportunities for refinement in future investigations. This framework has been developed for evaluating IPD effectiveness on industrial projects. Additional analysis and modification would be required prior to implementing this IPD assessment framework. Additionally, the Decision Tree model developed in this study can actively support performance-based decision-making for collaboration and integration strategies in industrial construction projects. Although the model was created on a limited dataset in the industrial construction sector, this approach based on artificial intelligence possesses the potential to be adapted and applied to other construction categories including commercial and infrastructure sectors. Future research can explore other machine learning models as well as deep neural networks-based decision support tools to identify and compare the efficacy and computational complexities in this type of data with a larger dataset.

Data Availability Statement

All data that supports the findings of this study will be available upon request.

Acknowledgment

This study was supported by a grant from the Construction Industry Institute (CII) under Grant No. RT-383. The authors sincerely thank the CII and the professional industry members who generously contributed their expertise, insights, assistance, and direction throughout this study.

References

Ahmed, S. and El-Sayegh, S., 2021. Critical Review of the Evolution of Project Delivery Methods in the Construction Industry. *Buildings*, [e-journal] 11(1), p.11. https://doi.org/10.3390/buildings11010011

AIA, 2007. *Integrated project delivery: A guide. American Institute of Architects, California.* [online] Available at: https://info.aia.org/SiteObjects/files/IPD_Guide_2007.pdf.

Al Subaih, A., 2015. Integrated Project Delivery: A Paradigm Shift for Oil and Gas Projects in the UAE and the Middle East Region. *Oil and Gas Facilities*, [e-journal] 4(04), pp.064–077. https://doi.org/10.2118/171722-PA

Al-Tmeemy, S.M.H.M., Abdul-Rahman, H. and Harun, Z., 2011. Future criteria for success of building projects in Malaysia. *International Journal of Project Management*, 29(3), pp.337–48. https://doi.org/10.1016/j.ijproman.2010.03.003

Arar, A.J. and Poirier, E., 2022. The Next Era of IPD Research: A Systematic Literature Review of The IPD Research Trends 2017-2020. In: *IOP Conference Series: Materials Science and Engineering*, 1218(1), p.012039. https://doi.org/10.1088/1757-899X/1218/1/012039

Atkinson, R., 1999. Project management: cost, time and quality, two best guesses and a phenomenon, it's time to accept other success criteria. *International Journal of Project Management*, [e-journal] 17(6), pp.337–42. https://doi.org/10.1016/S0263-7863(98)00069-6

Ayhan, M., Dikmen, I. and Talat Birgonul, M., 2021. Predicting the Occurrence of Construction Disputes Using Machine Learning Techniques. *Journal of Construction Engineering and Management*, [e-journal] 147(4), p.04021022. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002027

Baccarini, D., 1999. The Logical Framework Method for Defining Project Success. *Project Management Journal*, [e-journal] 30(4), pp.25–32. https://doi.org/10.1177/875697289903000405

Bannerman, P., 2008. Defining Project Success: A Multi-Level Framework. In: *Proceedings of the Project Management Institute Research Conference*, 5(7), pp.1-14.

Bapat, H., Sarkar, D. and Gujar, R., 2021. Evaluation of Key Performance Indicators of Integrated Project Delivery and BIM Model for Infrastructure Transportation Project in Ahmedabad, India through Decision-Making Approach. *Journal of The Institution of Engineers (India): Series A*, [e-journal] 102(4), pp.995–1011. https://doi.org/10.1007/s40030-021-00568-3

Barutha, P.J., Jeong, H.D., Gransberg, D.D. and Touran, A., 2021. Evaluation of the Impact of Collaboration and Integration on Performance of Industrial Projects. *Journal of Management in Engineering*, [e-journal] 37(4), p.04021037. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000921

Bilbo, D., Bigelow, B.F., Escamilla, E. and Lockwood, C., 2015. Comparison of Construction Manager at Risk and Integrated Project Delivery Performance on Healthcare Projects: A Comparative Case Study. *International Journal of Construction Education and Research*, [e-journal] 11(1), pp.40–53. https://doi.org/10.1080/15578771.2013.872734

Caldwell, S., 2007. *The New Zealand Construction National Key Performance Indicators: Handbook & Industry Measures (2006 data).* [e-book] New Zealand: Centre for Advanced Engineering, University of Canterbury. Available at: https://ir.canterbury.ac.nz/server/api/core/bitstreams/50f1b6ee-0170-432c-84d9-35d6f11d59bc/content.

Chan, A.P., 2001. *Framework For Measuring Success Of Construction Projects.* [pdf] Brisbane: QUT CRC for Construction Innovation. Available at: https://eprints.qut.edu.au/216001/1/2001-003-C-01_Framework_for_Measuring_Success.pdf.

Charbuty, B. and Abdulazeez, A., 2021. Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, [e-journal] 2(01), pp.20–28. https://doi.org/10.38094/jastt20165

Cho, C.-S. and Gibson, J., 2012. Development of a Project Definition Rating Index (PDRI) for General Building Projects. In: *Proceedings of Construction Congress VI: Building Together for a Better Tomorrow in an Increasingly Complex World.* Orlando Florida, 20-22 February 2000. pp.343–52. https://doi.org/10.1061/40475(278)38

Chua, D.K.H., Kog, Y.C. and Loh, P.K., 1999. Critical Success Factors for Different Project Objectives. *Journal of Construction Engineering and Management-asce*, 125(3), pp.142–50. https://doi.org/10.1061/(ASCE)0733-9364(1999)125:3(142)

CII RT-341, 2019. *Integrated Project Delivery for Industrial Projects. Construction Industry Institute. University of Texas, Austin.* [online] Available at: https://www.construction-institute.org/resources/knowledgebase/knowledge-areas/project-program-management/topics/rt-341.

Creswell, J.W. and Plano Clark, V.L., 2018. *Designing and conducting mixed methods research*. 3rd ed. Thousand Oaks, CA: Sage.

El Asmar, M., Hanna, A.S. and Loh, W.Y., 2013. Quantifying Performance for the Integrated Project Delivery System as Compared to Established Delivery Systems. *Journal of Construction Engineering and Management*, [e-journal] 139(11), p.04013012. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000744

El Asmar, M., Hanna, A.S. and Loh, W.Y., 2016. Evaluating integrated project delivery using the project quarterback rating. *Journal of Construction Engineering and Management*, 142(1), p.04015046. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001015

Elghaish, F., Hosseini, M.R., Talebi, S., Abrishami, S., Martek, I. and Kagioglou, M., 2020. Factors Driving Success of Cost Management Practices in Integrated Project Delivery (IPD). *Sustainability*, [e-journal] 12(22), p.9539. https://doi.org/10.3390/su12229539

Esmaeili, B., Molenaar, K.R., Leicht, R.M. and Messner, J., 2013. A Review of Critical Success Factors and Performance Metrics on Construction Projects. In: *Proceedings of 4th Construction Specialty Conference*, Montreal, Quebec, 29 May to 1 June. p.10.

Fernandez, A., Garcia, S., Herrera, F. and Chawla, N.V., 2018. SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary. *Journal of Artificial Intelligence Research*, [e-journal] 61, pp.863–905. https://doi.org/10.1613/jair.1.11192

Flyvbjerg, B., 2014. What you Should Know about Megaprojects and Why: An Overview. *Project Management Journal*, [e-journal] 45(2), pp.6–19. https://doi.org/10.1002/pmj.21409

Franz, B., Esmaeili, B., Leicht, R., Molenaar, K. and Messner, J., 2014. Exploring the Role of the Team Environment in Building Project Performance. In: *Proceedings of Construction Research Congress 2014: Construction in a Global Network*. Atlanta, Georgia, 19-21 May 2014. pp.1997–2010. https://doi.org/10.1061/9780784413517.204

Ganganwar, V., 2012. An overview of classification algorithms for imbalanced datasets. *International Journal of Emerging Technology and Advanced Engineering*, 2, pp.42–47.

Ghimire, P., Kim, K. and Acharya, M., 2024. Opportunities and Challenges of Generative AI in Construction Industry: Focusing on Adoption of Text-Based Models. *Buildings*, [e-journal] 14(1), p.220. https://doi.org/10.3390/buildings14010220

Ghimire, P., Pokharel, S., Kim, K. and Barutha, P., 2023. Machine learning-based prediction models for budget forecast in capital construction. In: *Proceedings of the 2nd International Conference on Construction, Energy, Environment & Sustainability,* Funchal, Portugal, 27-30 June. [pdf] pp.27–30. Available at: https://www.cees2023.uc.pt/projectos/cees2023/atas/pdfs/ID311.pdf.

Gibson, G.E. and Hamilton, M.R., 1994. *Analysis of Pre-Project Planning Effort and Success Variables for Capital Facility Projects–Source Document 105*. Austin, Texas: Construction Industry Institute, The University of Texas at Austin.

Gibson, G.E. and Whittington, D.A., 2010. Charrettes as a Method for Engaging Industry in Best Practices Research. *Journal of Construction Engineering and Management*, [e-journal] 136(1), pp.66–75. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000079

González-Boubeta, I. and Prado-Prado, J.C., 2020. Integrated Project Delivery: A Literature Review and Research Agenda. In: R. de Castro and G. Giménez, eds. *Advances in Engineering Networks*, *Lecture Notes in Management and Industrial Engineering*. Cham: Springer International Publishing. pp.11–18. https://doi.org/10.1007/978-3-030-44530-0_2

Government of Western Australia, 2010. *Infrastructure procurement options guide*. Perth, WA: Centre for Excellence and Innovation in Infrastructure Delivery. [online] Available at: http://www.ceiid.wa.gov.au/Docs/IPOG.pdf.

Guo, G., Wang, H., Bell, D., Bi, Y. and Greer, K., 2003. *KNN Model-Based Approach in Classification*. *Lecture Notes Computer Science*, p.996. https://doi.org/10.1007/978-3-540-39964-3_62

Han, W.S., Yusof, A.M., Ismail, S. and Aun, N.C., 2012. Reviewing the notions of construction project success. *International Journal of Business and Management*, 7(1), p.90. https://doi.org/10.5539/ijbm.v7n1p90

Hanna, A.S., 2016. Benchmark Performance Metrics for Integrated Project Delivery. *Journal of Construction Engineering and Management*, [e-journal] 142(9), p.04016040. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001151

Henneveld, M., 2006. Alliance Contracting--Removing the Boundaries for Infrastructure Delivery. In: *Annual Conference & Exhibition of the Transportation Association of Canada, 2006. Congres et exposition annuels de l'Association des transport du Canada, 2006*. Charlottetown, Prince Edward Island , Canada, 17-20 September. Transportation Association of Canada (TAC). Available at: https://trid.trb.org/view/794946 [Accessed 6 April 2023].

Iwanski, M., 2013. *Performance of Integrated Project Delivery (IPD) For Mechanical and Electrical Contractors*. MSc. Civil and Environmental Engineering. University of Wisconsin, Madison. [pdf] Available at: https://minds.wisconsin.edu/handle/1793/66680 [Accessed 19 December 2023].

Janikow, C.Z., 1998. Fuzzy decision trees: issues and methods. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, [e-journal] 28(1), pp.1–14. https://doi.org/10.1109/3477.658573

Jones, B., 2014. Integrated Project Delivery (IPD) for Maximizing Design and Construction Considerations Regarding Sustainability. *Procedia Engineering*, [e-journal] 95, pp.528–38. https://doi.org/10.1016/j.proeng.2014.12.214

KPI Working Group, 2007. *The New Zealand Construction Industry National Key Performance Indicators: Handbook and Industry Measures*. [online] Available at: https://ir.canterbury.ac.nz/handle/10092/11589.

Lee, K.-P., Lee, H.-S., Park, M., Kim, D.Y. and Jung, M., 2017. Management-Reserve Estimation for International Construction Projects Based on Risk-Informed k-NN. *Journal of Management in Engineering*, [e-journal] 33(4), p.04017002. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000510

Leung, K.M., 2007. *Naive Bayesian Classifier* [pdf online] Polytechnic University Department of Computer Science/ Finance and Risk Engineering. Available through: https://cse.engineering.nyu.edu/~mleung/FRE7851/f07/naiveBayesianClassifier.pdf.

Lim, C.S. and Zain, M., 1999. Criteria of project success: An exploratory re-examination. *International Journal of Project Management*, [e-journal] 17, pp.243–48. https://doi.org/10.1016/S0263-7863(98)00040-4

Merrow, E.W., 2011. *Industrial Megaprojects: Concepts, Strategies, and Practices for Success*. Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781119201045

Mesa, H., Molenaar, K.R. and Alarcón, L.F., 2016. Exploring performance of the integrated project delivery process on complex building projects. *International Journal of Project Management*, [e-journal] 34(7), pp.1089–1101. https://doi.org/10.1016/j.ijproman.2016.05.007

Mesa, H.A., Molenaar, K.R. and Alarcón, L.F., 2019. Comparative analysis between integrated project delivery and lean project delivery. *International Journal of Project Management*, [e-journal] 37(3), pp.395–409. https://doi.org/10.1016/j.ijproman.2019.01.012

Müller, R. and Jugdev, K., 2012. Critical success factors in projects: Pinto, Slevin, and Prescott – the elucidation of project success. *International Journal of Managing Projects in Business*, [e-journal] 5(4), pp.757–75. https://doi.org/10.1108/17538371211269040

Nassar, N. and AbouRizk, S., 2014. Practical Application for Integrated Performance Measurement of Construction Projects. *Journal of Management in Engineering*, [e-journal] 30(6), p.04014027. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000287

Pinto, J.K. and Slevin, D.P., 1988. Project success: definitions and measurement techniques. Project Management Journal, [online pdf] XIX(1), pp.67-72. Available at: https://www.researchgate.net/publication/242530015_Project_success_Definitions_and_measurement_techniques.

Pokharel, S. and Ghimire, P., 2023. Data-driven MLmodels for accurate prediction of energy consumption in a low-energy house: A comparative study of XGBoost, Random Forest, Decision Tree, and Support Vector Machine. *Journal of Innovations in Engineering Education*, [e-journal] 6(1), pp.12–20. https://doi.org/10.3126/jiee.v6i1.54965

Rajula, H.S.R., Verlato, G., Manchia, M., Antonucci, N. and Fanos, V., 2020. Comparison of Conventional Statistical Methods with Machine Learning in Medicine: Diagnosis, Drug Development, and Treatment. *Medicina*, [e-journal] 56(9), p.455. https://doi.org/10.3390/medicina56090455

Ranjan, G.S.K., Kumar Verma, A. and Radhika, S., 2019. K-Nearest Neighbors and Grid Search CV Based Real Time Fault Monitoring System for Industries. In: *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*. Bombay, India, 29-31 March. pp.1–5. https://doi.org/10.1109/I2CT45611.2019.9033691

Rish, I., 2001. An empirical study of the naive Bayes classifier. In: *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*. Seattle, Washington, USA, 4-10 August 2001, 3(22), pp.41-46.

Rodrigues, M.R. and Lindhard, S.M., 2021. Benefits and challenges to applying IPD: experiences from a Norwegian mega-project. *Construction Innovation*, [e-journal] 23(2), pp.287–305. https://doi.org/10.1108/CI-03-2021-0042

Sanni-Anibire, M.O., Zin, R.M. and Olatunji, S.O., 2022. Machine learning model for delay risk assessment in tall building projects. *International Journal of Construction Management*, [e-journal] 22(11), pp.2134–43. https://doi.org/10.1080/15623599.2020.1768326

Schlögl, M., Stütz, R., Laaha, G. and Melcher, M., 2019. A comparison of statistical learning methods for deriving determining factors of accident occurrence from an imbalanced high resolution dataset. *Accident Analysis & Prevention*, [e-journal] 127, pp.134–49. https://doi.org/10.1016/j.aap.2019.02.008

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J. and Brenning, A., 2019. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, 406, [e-journal] pp.109–20. https://doi.org/10.1016/j.ecolmodel.2019.06.002

Scikit Learn, 2023. *sklearn.model_selection.GridSearchCV*. [online] scikit-learn. Available at: https://scikit-learn/stable/modules/generated/sklearn.model_selection.GridSearchCV.html [Accessed 5 April 2023].

Shrnhur, A.J., Levy, O. and Dvir, D., 1997. Mapping the dimensions of project success. *Project management journal*, 28(2), pp.5–13.

Sokolova, M. and Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, [e-journal] 45(4), pp.427–37. https://doi.org/10.1016/j.ipm.2009.03.002

Thomsen, C., Darrington, J., Dunne, D. and Lichtig, W., 2016. *Managing Integrated Project Delivery*. [online] Available at: https://lean-construction-gcs.storage.googleapis.com/wp-content/uploads/2022/08/08161051/CMAA_Managing_Integrated_Project_Delivery_1-1.pdf.

Tien Bui, D., Khosravi, K., Shahabi, H., Daggupati, P., Adamowski, J.F., Melesse, A.M., Thai Pham, B., Pourghasemi, H.R., Mahmoudi, M., Bahrami, S., Pradhan, B., Shirzadi, A., Chapi, K. and Lee, S., 2019. Flood Spatial Modeling in Northern Iran Using Remote Sensing and GIS: A Comparison between Evidential Belief Functions and Its Ensemble with a Multivariate Logistic Regression Model. *Remote Sensing*, [e-journal] 11(13), p.1589. https://doi.org/10.3390/rs11131589

Walker, D. and Rowlinson, S., 2019. *Routledge Handbook of Integrated Project Delivery*. [e-book] Devon, UK: Routledge. https://doi.org/10.1201/9781315185774

Wang, Y., Shao, Z. and Tiong, R.L.K., 2021. Data-Driven Prediction of Contract Failure of Public-Private Partnership Projects. *Journal of Construction Engineering and Management*, [e-journal] 147(8), p.04021089. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002124

Whang, S.-W., Park, K.S. and Kim, S., 2019. Critical Success Factors for Implementing Integrated Construction Project Delivery. *Engineering, Construction and Architectural Management*, [e-journal] 26(10), pp.2432–46. https://doi.org/10.1108/ECAM-02-2019-0073

Wood, P. and Duffield, C., 2009. *In Pursuit of Additional Value: A benchmarking study into alliancing in the Australian Public Sector*. Melbourne, VIC: Department of Treasury and Finance, State of Victoria.

Yu, J.H., Yoo, S.E., Kim, J.I. and Kim, T.W., 2019. Exploring the Factor-Performance Relationship of Integrated Project Delivery Projects: A Qualitative Comparative Analysis. *Project Management Journal*, [e-journal] 50(3), pp.335–45. https://doi.org/10.1177/8756972819832206

Yun, S. and Jung, W., 2017. Benchmarking Sustainability Practices Use throughout Industrial Construction Project Delivery. *Sustainability*, [e-journal] 9(6), p.1007. https://doi.org/10.3390/su9061007