INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY: APPLIED BUSINESS AND EDUCATION RESEARCH
2024, Vol. 5, No. 7, 2733 – 2745
http://dx.doi.org/10.11594/ijmaber.05.07.27

Research Article

Predictive Models of Construction Project Success Rating Using Regression and Artificial Neural Network

Clyde L. Tamayo[1], Jerome Jordan F. Famadico[2]

[1]Graduate School of Engineering, Adamson University, 1000, Philippines
[2]Civil Engineering Department, Adamson University, 1000, Philippines

Article history:
Submission 30 June 2024
Revised 07 July 2024
Accepted 23 July 2024

**Corresponding author:*
E-mail:

ABSTRACT

This research addresses the gap in comprehensive predictive models for construction project success rating by exploring the potential of regression models to evaluate project success rating. By analyzing 130 datasets from the National Capital Region, the study utilizes Support Vector Regression (SVR), Multiple Linear Regression (MLR), and Artificial Neural Network (ANN) with a 22-30-1 configuration (22 input neurons, 30 neurons in a single hidden layer, and 1 output neuron). The input variables represent critical success factors rated on a scale of 1-5, while the output variable represents the predicted project success percentage rating. Various statistical tools, including ANOVA, Lasso Regression, R^2, MAE, and MSE, are utilized for evaluation. The findings reveal that SVR achieved the highest accuracy (R^2 = 0.881, MAE = 2.172, MSE = 7.054), followed closely by MLR (R^2 = 0.874, MAE = 2.180, MSE = 7.470), while ANN (R^2 = 0.743, MAE = 3.076, MSE = 15.239) may require refinement. Lasso Regression identified 22 critical success factors, with Financial Condition, Effectiveness in Decision-Making, and Compliance to Quality Standards ranking as the top three. This research contributes to the advancement of construction predictive analytics, which can lead to improved decision-making and more efficient, effective, and ethical construction practices.

Keywords: *Artificial neural network, Construction project, Predictive models, Regression, Success rating*

Introduction

A construction business stands as one of the most complex and dynamic sectors, characterized by multifaceted projects that often involve high levels of uncertainty, risk, and resource allocation. While advances in technology have facilitated significant improvements in various aspects of construction, there remains a critical gap in predictively assessing the success rate of construction projects. Traditional methods for gauging the viability and success of construction projects, often reliant on experience-based

How to cite:
Tamayo, C. L. & Famadico, J. J. F. (2024). Predictive Models of Construction Project Success Rating Using Regression and Artificial Neural Network. *International Journal of Multidisciplinary: Applied Business and Education Research. 5*(7), 2733 – 2745. doi: 10.11594/ijmaber.05.07.27

judgment and heuristic evaluations, suffer from a lack of precision and objectivity. This deficiency presents a pressing issue for organizations that need to manage projects within specified budgets, timeframes, and safety parameters, especially given the sector's thin profit margins and high stakes concerning human lives and environmental impact.

Recent years have seen the advent of machine learning technology as a transformative tool across various sectors. Its capability to make accurate, data-driven predictions has been substantiated by several studies in the realm of construction. Previous studies have effectively utilized machine learning to forecast the likelihood of fatalities occurring at construction sites (Choi, J. et. al, 2020). Yet, these models generally focus on specific metrics like safety or productivity, often in isolation from each other. There exists a palpable absence of comprehensive predictive models that take into account a broader range of key performance indicators such as safety metrics, productivity rates, adherence to budgets, and timelines.

This lacuna in research and practical application is especially troubling given the complexity and interdependent risks inherent in construction projects, which have been documented to be significant impediments to their success (Gondia, A., 2019). The aim of this study is to address the notable deficiency by utilizing predictive regression models to assess the critical success rate of construction organizations across multiple success factors. The research takes place in an industry that is in desperate need of predictive capabilities to enhance its management effectiveness, mitigate risks, and ultimately enhance its overall success rate more details about the paper’s rationale, motivation, significance, scope and limitations, and the setting of the study. Both the Abstract and Introduction should be relatively nontechnical yet clear enough for an informed reader to understand the manuscript’s contribution.

Success Factors in Construction Project Management

The literature reviewed in this study indicates that the critical success factors (CSFs) outlined in numerous studies collectively offer a comprehensive insight into the multifaceted nature of achieving success in construction projects. The studies reviewed include those conducted by Dessalegn (2021). A recurring theme across these references is the paramount importance of effective planning and execution, which includes factors such as competence in planning and execution, decision-making effectiveness, and project monitoring and control mechanisms.

The integration of these essential elements for achievement provides a comprehensive framework for understanding and enhancing the success of construction projects in different circumstances. The findings of this study have substantial implications for project managers, stakeholders, and researchers alike. Gaining a more profound comprehension of the interaction between these Critical Success Factors (CSFs) can assist professionals and scholars in pinpointing areas that need enhancement and implementing strategies to augment the success of a project.

Artificial Neural Network (ANN)

Nitin Malik's article, published in June 2005, provides a detailed examination of Artificial Neural Networks (ANN) and their various applications. Artificial Neural Networks (ANNs) are an intriguing amalgamation of biology and computation, drawing inspiration from biological neural structures. Malik highlights that although artificial neural networks (ANNs) aim to replicate the functionality of biological neurons, their real value lies in their ability to address a wide range of practical issues. The utility of Artificial Neural Networks (ANNs) has experienced significant growth over the years, expanding into various domains.

Delving deeper, the versatility of ANNs is worth underscoring. Unlike traditional algorithms that follow a linear trajectory, ANNs, much like the human brain, exhibit adaptive learning. This enables them to discern patterns and nuances in vast and complex datasets, offering solutions that traditional algorithms might overlook. However, this strength also raises pertinent questions. How robust are ANNs against anomalous data? How do they fare in situations where data patterns evolve over time? Malik's exposition, while highlighting the potential of ANNs, inadvertently opens up avenues for discussions around their limitations and areas for future research. The article subtly underscores that while ANNs offer groundbreaking solutions, continuous refinement and contextual applicability remain crucial.

Multiple Linear Regression (ANN)

The research article, "A Model for Business Success Prediction using Machine Learning Algorithms," authored by Ibukun Afolabi et al. in 2019, offers valuable insights into the utilization of machine learning (ML) algorithms for predicting business success. This study is highly pertinent to the ongoing research project that seeks to create predictive models for evaluating the crucial success rate of construction organizations utilizing machine learning algorithms. The study conducted by Afolabi et al. offers valuable methodologies and findings that can be utilized to shape and direct the modeling and data analysis approach in the construction context.

Although the primary focus of the research is on Naïve Bayes classification algorithms, the principle of linear regression, a foundational machine learning technique, is inherent in their methodology. Linear regression is crucial in understanding relationships between variables, which is a key aspect of predictive modeling.

Afolabi et al. used Pearson's correlation coefficient, a fundamental concept in linear regression, to select attributes for prediction. This approach is crucial for determining the magnitude and orientation of relationships between variables, which is essential in any predictive model. In the context of the current study on construction organizations, a similar approach can be used to identify and quantify relationships between various success indicators (like safety, budget adherence, and timelines) and overall project success.

Their research highlights the importance of precision and recall in evaluating the effectiveness of predictive models. This aspect is crucial for the current study as it emphasizes the need for a balance between the accuracy of predictions and the model's ability to generalize across different scenarios in construction projects.

Afolabi et al. highlights the importance of linear regression and other statistical methodologies in creating accurate predictive models in the field of machine learning. The incorporation and modification of these methodologies in construction organizations for the present study have the potential to result in the creation of a comprehensive and precise predictive model. This, in turn, can improve decision-making processes and increase the likelihood of success in construction projects.

Support Vector Regression (SVR)

The research article titled "Binary gravity search algorithm and support vector machine for forecasting and trading stock indices" authored by Haijun Kang et al. and published in the International Review of Economics & Finance in March 2023, offers a detailed investigation of the Support Vector Machine (SVM) algorithm, specifically in relation to its application in financial prediction. This study is notably relevant to the current research aimed at developing predictive models for assessing the critical success rate of construction organizations using machine learning algorithms. The methodologies and findings from Kang et al.'s study can provide valuable insights into the adaptability and efficacy of SVM in complex predictive tasks.

The Support Vector Machine (SVM) is a resilient and adaptable machine learning algorithm employed for tasks involving classification and regression. Kang et al.'s study utilizes Support Vector Machines (SVM) to predict the daily returns of prominent stock indices, showcasing its effectiveness in managing large and intricate datasets.

The novelty of Kang et al.'s research lies in the incorporation of Support Vector Machine (SVM) with the Binary Gravity Search Algorithm (BGSA). By optimizing the parameters and inputs of SVM, this hybrid approach results in improved prediction accuracy. Within the construction industry, a comparable hybrid approach could be utilized. Integrating SVM with additional optimization algorithms could potentially improve the predictive accuracy of the construction project success rate by effectively analyzing complex data, including factors like adherence to budget, project timelines, and safety metrics.

Kang et al. reported an average accuracy of 52.87% across five stock indices using the BGSA-SVM model, outperforming traditional methods. This highlights the potential of SVM to achieve high accuracy in prediction tasks, which is crucial for the development of predictive models in the construction industry. Even small variations in factors such as resource allocation or safety standards can have a significant impact on the overall success of a project.

The successful application of SVM in financial forecasting suggests its potential adaptability to the construction sector. The algorithm's ability to handle diverse and complex datasets makes it suitable for modeling the multifaceted nature of construction projects.

The use of BGSA to optimize SVM parameters in Kang et al.'s study highlights the importance of fine-tuning machine learning algorithms for specific tasks. In construction project prediction, tuning SVM parameters to suit the specific data characteristics of construction projects could significantly enhance model performance.

The study's achievement in creating a lucrative trading strategy using BGSA-SVM predictions suggests that a comparably organized SVM model could assist in decision-making in construction management by offering dependable forecasts on project success factors.

Kang et al.'s research illustrates the effectiveness of Support Vector Machines in handling complex predictive tasks, especially when enhanced by optimization algorithms like the Binary Gravity Search Algorithm. For the current study on construction organizations, leveraging SVM's capabilities, potentially in a hybrid model with other algorithms, could lead to the development of a sophisticated and accurate predictive tool. This tool could significantly aid in strategic planning and decision-making, enhancing success rates in construction projects.

Table 1. Most Cited Factors for Project Success Found in Literature

Success Factor	Frequency of References
Planning and Execution	7
Compliance to Quality Standards	6
Effectiveness in Decision Making	6
Experience	6
Motivation	6
Economic Condition	6
Climatic Condition	6
Selection of Contract	5
Project Management	5
Availability of Resources	5
Good Communication	5
Financial Condition	5
Political Condition	5
Client's Decision Making	5
Technical Capability	4
Risk Management	4
Good Cost Proposal / Budget	4
Procurement Strategy	4
Cost Control & Project Margins	4
Clear Mission, Vision, and Goals	3
Timely Completion	3
Organization Structure	3
Access to IT / Software	3

Success Factor	Frequency of References
Health and Safety Program	3
Stakeholder Management	3
Troubleshooting	2
Leadership	2
Client's Financial System	2
Approval in Project Design	2
Approval of Materials	1

Table 1 presents a comprehensive overview of various success factors and their corresponding frequencies of reference to construction project success. An analysis of this data uncovers significant trends and priorities within this framework.

The success factor most frequently referenced is Planning and Execution, with a frequency of 7. This highlights the critical role of meticulous planning and effective execution in achieving project success, emphasizing the importance of comprehensive planning and efficient implementation.

Compliance to Quality Standards, Effectiveness in Decision Making, Experience, Motivation, and Climatic Conditions each have a frequency of 6, underscoring the importance of adhering to quality benchmarks, making informed decisions, leveraging experience, maintaining motivation, and adapting to environmental conditions.

Selection of Contract, Project Management, Availability of Resources, Good Communication, Financial Condition, Political Condition, and Client's Decision Making each have a frequency of 5, emphasizing the significance of strategic contract selection, robust project management, resource availability, effective communication, financial health, stable political environments, and responsive client decision-making processes.

Technical Capability, Risk Management, Good Cost Proposal / Budget, Procurement Strategy, Cost Control & Project Margins each have a frequency of 4, highlighting the need for technical expertise, risk mitigation, financial acumen, and strategic procurement practices.

Clear Mission, Vision, and Goals, Timely Completion, Organization Structure, Access to IT / Software, Health and Safety Program, and Stakeholder Management each have a frequency of 3, indicating their foundational or supplementary importance.

Troubleshooting, Leadership, Client's Financial System, and Approval in Project Design have a frequency of 2, suggesting their secondary role within the broader context of success factors.

Approval of Materials has a frequency of 1, indicating that it is the least frequently referenced factor, possibly due to established protocols that mitigate its impact on overall success.

All 30 compiled success factors are selected for the study.

LASSO Regression: A Powerful Tool for Feature Selection and Model Regularization

LASSO (Least Absolute Shrinkage and Selection Operator) regression is a well-known statistical method that has gained significant attention in recent years due to its ability to handle high-dimensional data with a large number of features. It is a form of regularized regression that not only estimates the relationship between a dependent variable and independent variables but also performs feature selection, making it a powerful tool for model regularization.

The core concept of LASSO is based on ordinary least squares (OLS) regression, where a penalty term based on the L1 norm (absolute value) of the regression coefficients is introduced. This penalty term, controlled by a hyperparameter called lambda (λ), shrinks the coefficients towards zero. As lambda increases, some coefficients become exactly zero, effectively removing those features from the model. This feature selection characteristic makes LASSO a valuable tool in producing a more interpretable model with fewer, but more informative, features.

LASSO offers several advantages over other regression methods. Firstly, it performs automatic feature selection by driving coefficients of irrelevant or redundant features to zero, simplifying the model and improving interpretability. Secondly, the L1 penalty term acts as a regularizer that prevents the model from becoming overly complex and fitting to noise in the data, reducing overfitting, and improving generalization performance on unseen data. Thirdly, LASSO often yields sparse models, where many coefficients become zero, resulting in a clearer understanding of which features are most influential.

However, implementing LASSO requires careful consideration of the lambda parameter and its impact on model complexity and interpretability. Selecting an appropriate lambda value involves a trade-off between model complexity and accuracy. Techniques like cross-validation can be used to find the optimal lambda value.

It is important to note that LASSO and ridge regression are both regularization techniques that differ in their penalty terms. Ridge regression uses the L2 norm (squared values) of the coefficients, leading to all coefficients shrinking but not necessarily becoming zero. This can result in less interpretable models compared to LASSO.

Finally, LASSO regression is a powerful tool for feature selection and model regularization, particularly when dealing with high-dimensional data. Its ability to reduce overfitting and sparsity in the model makes it a valuable tool in various machine learning applications. However, selecting an appropriate lambda value and understanding its impact on model complexity and interpretability are crucial for successful implementation.

Methods

Regression and machine learning algorithms are revolutionizing construction management by enabling the prediction of critical success factors in this complex, dynamic industry. A well-designed research plan ensured the project's accuracy and usefulness, guiding the exploration of how machine learning can enhance project success, cost efficiency, scheduling, quality, safety, risk management, and stakeholder satisfaction.

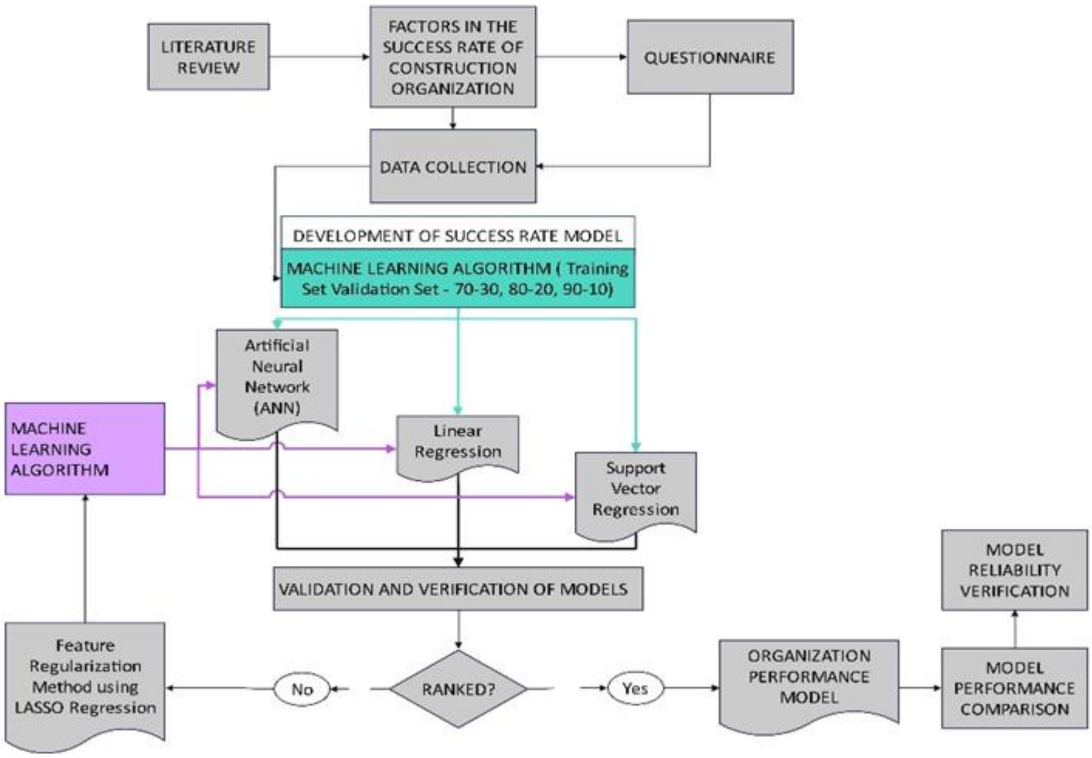

Figure 1. Research Flowchart

The research flowchart outlined a systematic process for developing a regression model to predict the success rate of construction projects. It comprised several critical stages, each of which contributed to the robustness and accuracy of the final predictive model.

First, the literature review phase had examined existing research to identify factors influencing construction project success rates. This foundational step ensured that the model was grounded in established knowledge and that relevant variables were considered.

In the questionnaire development phase, a comprehensive questionnaire had been designed to collect data on the identified factors. This instrument was critical for gathering empirical data from construction professionals, ensuring the model was based on real-world insights and experiences.

The data collection phase involved administering the questionnaire through online surveys or interviews, facilitating the accumulation of data essential for model development. The quality and breadth of this data directly impacted the accuracy of the predictive model.

Developing the project success rating model involved utilizing the collected data to create a predictive model using three regression algorithms: Artificial Neural Network (ANN), Linear Regression, or Support Vector Regression (SVR). Each algorithm offered unique strengths; ANN was adept at modeling complex, non-linear relationships, Linear Regression helped understand linear dependencies, and SVR could handle both linear and non-linear relationships efficiently.

The Machine Learning Algorithm stage involved dividing the data into training and validation sets, commonly split in ratios such as 70-30, 80-20, or 90-10. Techniques like cross-validation and performance metrics such as mean squared error (MSE) or R-squared were employed to ensure the model's robustness and generalizability.

Feature Regularization, often utilizing LASSO regression, had then been applied to minimize model variance and enhance generalizability. This was crucial for refining the model and preventing overfitting.

Subsequently, the models were ranked based on their validation set performance to identify the most effective model. In the Organization Model Performance Comparison phase, the top-performing models were compared to ascertain the best predictive model.

Finally, the model selection phase, a meticulous process, involved choosing the model with the highest performance as the final predictive tool for assessing the success rate of construction projects. This rigorous, multi-step process ensured the development of a reliable and accurate predictive model grounded in empirical data and robust analytical techniques, instilling confidence in the model's reliability.

Research Design

The study garnered comprehensive insights through an integrative mixed-methods approach. To achieve this, it deployed a rating scale survey questionnaire targeting pivotal participants in the construction ecosystem, including project managers, stakeholders, and engineers. These surveys were designed to gather rich qualitative data, essential for understanding the nuanced dynamics of construction projects.

Research Locale

This study focused specifically on construction firms operating within the National Capital Region. The geographical scope included Manila, Quezon City, Caloocan, Las Piñas, Makati, Malabon, Mandaluyong, Marikina, Muntinlupa, Navotas, Parañaque, Pasay, Pasig, Pateros, San Juan, Taguig, and Valenzuela. The research focused on the time frame of 2021 to 2024, a period strategically chosen to guarantee the availability of key stakeholders for effective data collection. This time frame also ensured that the insights derived from the research were both contemporary and pertinent, reflecting the latest trends and developments in the construction sector within these urban locales.

Population and Sampling

The research was designed to capture insights from approximately 130 datasets, comprising responses from contractors and their evaluators on the client side. This was achieved through a qualitative survey method, underpinned by a purposive sampling strategy. In qualitative research, the emphasis was placed on the depth and richness of the data rather than its sheer volume, making it particularly effective for investigating complex phenomena.

Table 2. Criteria for Selection of Respondents

CRITERIA FOR SELECTION OF RESPONDENTS	
Age	24 - 60 years of age
Job Title / Role	Contractor's Evaluator / Representative
Location	Cities in the National Capital Region (NCR)
Years of Experience	At least 3 years
Careers in Construction Management	Construction Manager / Civil Engineer / Architect / Mechanical Engineer / Electrical Engineer / Procurement Officer
Project Size	Category B - AAAA
Certifications	Professional Certifications
Professional Association	Professional Association / Organizations

Respondents were selected based on specific criteria to ensure relevance and expertise. Eligible individuals were aged 24-60 and held positions such as Contractor's Evaluator or Representative within cities in the National Capital Region (NCR). They were required to possess a minimum of three years of experience in roles related to construction management, including Construction Manager, Civil Engineer, Architect, Mechanical Engineer, Electrical Engineer, or Procurement Officer, and be involved in projects classified as Category B - AAAA. Additionally,

Result and Discussion

The Results and Discussion may be combined into a single section or presented separately. They may also be broken into subsections with short, informative headings.

Table 3. Demographic Data Summary

Demographic Data	
Age	24-29 (60, 46%), 30-39 (36,28%), 40-49 (24, 18%), and 50-69 (10, 8%)
Role	Contractor's Evaluator (74, 57%) and Contractor's Representative (56, 43%)
Location	Manila (92, 71%), Taguig (20, 15%), Muntinlupa (12, 9%), and Makati (6, 5%)
Years of Experience	Less than 6 (52, 40%), 6-15 (44, 34%), 16-25 (24, 18%), and more than 25 (10, 8%)
Careers	Construction Managers (56, 43%), CE (31, 24%), Procurement Officer (18, 14%), Architect (11, 8%), Mechanical Engineer (8, 6%), and EE (6, 5%)

This study aimed to assess the predictive capabilities of Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Artificial Neural Network (ANN) models by comparing their accuracy. The evaluation involved ranking the models based on their performance metrics, including R2, Mean Absolute Error (MAE), and Mean Squared Error (MSE). Among the models, SVR with a 90-10 data split showed the highest predictive accuracy with an R2 of 0.881, an MAE of 2.172, and an MSE of 7.054. MLR with a 70-30 data split followed closely, while ANN with a 90-10 data split performed the least accurately with an R2 of 0.743, an MAE of 3.076, and an MSE of 15.239.

In identifying the critical success factors (CSFs) for construction project success, the study highlighted Financial Condition, Effectiveness in Decision Making, and Compliance to Quality Standards as the top three influential factors. Other notable CSFs included clear mission and vision, timely completion, team composition, technical capability, project management, risk management, and various other factors spanning economic conditions and client-related attributes.

Summary of Findings

	OBJECTIVE	FINDINGS			
1	To assess the predictive ability of Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Artificial Neural Network (ANN) models by comparing their accuracy.	Ranking of the Predictive Accuracy of the Models			
		1 - SVR (90-10 Data Split): R2 = 0.881, MAE = 2.172, and MSE = 7.054			
		2 - MLR (70-30 Data Split): R2 = 0.874, MAE = 2.180, and MSE = 7.470			
		3 - ANN (90-10 Data Split): = 0.743, MAE = 3.076, and MSE = 15.239)			
2	To identify the critical success factors (CSFs) and rank the top three most influential factors.	Top 3: Financial Condition, Effectiveness in Decision Making, Compliance to Quality Standards			
		Other Critical Success Factors: Clear Mission, Vision, and Goals, Selection of Contract, Timple Completion, Team Composition, Technical Capibility, Project Management, Troubleshooting, Risk Management, Health and Safety, Experience, Leadership, Motivation, Good Communication, Good Cost Proposal, Procurement Strategy, Economic Condition, Climatic Condition, Client's Financial System, Client's Decision Making.			
3	To quantify the improvement in predictive performance achieved by incorporating the critical success factors into Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Artificial Neural Networks (ANN) models, compared to their performance with all the factors considered.	IMPROVEMENT IN PREDICTION PERFORMANCE			
		SIGNIFICANT	MODERATE TO SIGNIFICANT	MODERATE	SMALL
		R2 - ANN (80-20)	R2 - SVR (90-10)	R2 - ANN (90-10)	R2 - ANN (70-30)
		R2 - MLR (80-20)	R2 - MLR (70-30)	R2 - SVR (80-20)	R2 - SVR (70-30)
		MSE - SVR (80-20)	R2 - MLR (90-10)	MAE - ANN (90-10)	MAE - ANN (70-30)
		MSE - SVR (90-10)	MSE - ANN (90-10)	MAE - SVR (80-20)	MAE - ANN (80-20)
		MSE - MLR (80-20)	MSE - SVR (70-30)	MAE - SVR (90-10)	MAE - SVR (70-30)
			MSE - MLR (70-30)	MAE - MLR (70-30)	MAE - MLR (90-10)
				MAE - MLR (80-20)	MSE - ANN (70-30)
				MSE - ANN (80-20)	MSE - MLR (90-10)
4	To utilize and implement advanced machine learning algorithms that can perform a comprehensive diagnostic analysis of construction projects, accurately quantifying the contribution of each individual success factor towards improving the overall success rate.	Artificial Neural Network (ANN) Final Equation Model			
		Support Vector Regression (SVR)'s Final Equation Model			
		Multiple Linear Regression (MLR) Final Equation Model			
5	To examine and assess the impact of different data split ratios (such as 70-30, 80-20, and 90-10) on the performance and predictive precision of machine learning models	70-30,80-20, and 90-10 Data-Split Ratio are Statistically Similar (P-Value = 0.408 > 0.05)			
6	To determine the significant differences among the Multiple Linear Regression (MLR), Support Vector Regression (SVR), and Artificial Neural Networks (ANN) models.	R2 of ANN & SVR and ANN & MLR are statistically different, while SVR and MLR is not.			
		MAE and MSE of ANN, SVR, and ANN are not statistically different			

Figure 2. Summary of Findings

The study also explored how incorporating these critical success factors into the predictive models improved their performance. It was observed that the ANN model, when incorporating CSFs, demonstrated significant improvements in predictive accuracy, especially in the 80-20 data split scenario. The SVR and MLR models also showed enhancements, though to varying degrees, with significant to moderate improvements noted in several instances.

Advanced machine learning algorithms were developed to provide a detailed diagnostic analysis of construction projects, quantifying the contribution of each success factor to the overall success rate. The study provided a final equation model for ANN, a final equation for SVR, and a final equation model for MLR, which enabled a comprehensive understanding of how each model utilizes different factors to predict project outcomes.

The impact of different data split ratios (70-30, 80-20, and 90-10) on the models' performance was examined. The study found that these data split ratios did not significantly differ in their impact on the models' predictive precision, with a P-Value of 0.408 indicating statistical similarity across these splits.

In terms of comparing the models, significant differences were found in the R2 values between ANN & SVR and ANN & MLR, suggesting that ANN's predictive accuracy was notably different from the other two models. However, no significant differences were observed in the MAE and MSE values among ANN, SVR, and MLR, indicating similar levels of error across these models.

Overall, the study concluded that while SVR exhibited the highest predictive accuracy, integrating critical success factors significantly enhances the performance of predictive models. The findings underscore the importance of selecting appropriate data split ratios and the value of advanced algorithms in providing nuanced insights into project success determinants.

Ranking of the Critical Success Factors in Construction Projects

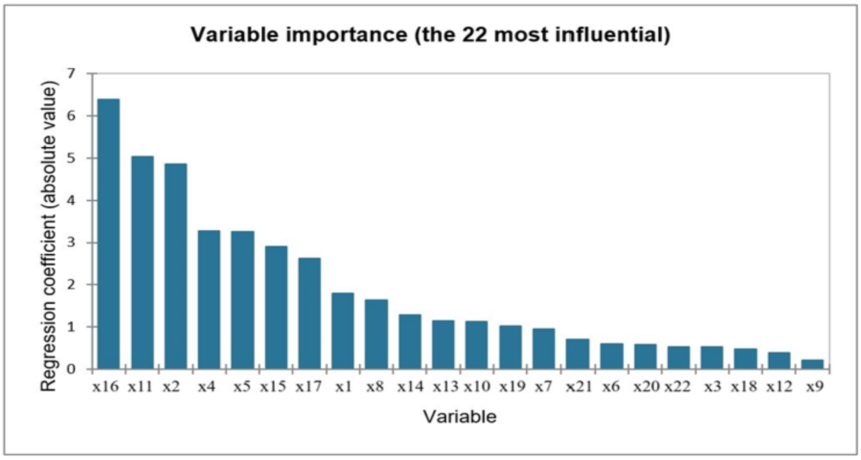

Figure 3. Investigation and Ranking of the Critical Success Factors using Lasso Regression

This study's top three success factors—Financial Condition, Effectiveness in Decision-Making, and Compliance with Quality Standards—were compared with those from a study on the critical success factors of infrastructure construction projects. Yamany, M.S., et al. (2024) highlighted sufficient funding, the project manager's competence, dedication and active participation, and effective communication and coordination among all parties involved as critical success factors, revealing both similarities and differences in contextual importance.

Both studies emphasize the vital role of financial resources and decision-making for project success. The focus on "Financial Condition" in this research parallels "Adequacy of Funding" in Yamany, M.S., et al.'s study, underscoring the need for sufficient and continuous funding. Similarly, "Effectiveness in Decision-Making" aligns with "Project Manager's competence, dedication, and active participation," highlighting the project manager's crucial role. While this study prioritizes "Compliance with Quality Standards" to ensure integrity and satisfaction, Naeem et al. stress "Effective communication and coordination among all parties" for effective organization and timely execution. These insights collectively enhance the understanding of project success, with this study's emphasis complementing Yamany, M.S., et al.'s focus, offering a well-rounded view of project management.

Final Predictive Models

ANN Final General Equation Model

$$y_{pred} = \sum_{1}^{30} w_{(n+660)}\, g_n + b_7$$

(eq. 1)

SVR Final General Equation Model

$$y_{pred} = 59.5910116453101 + 8.96465737182651 * \left[\sum_{1}^{79} (\alpha - \alpha *) K(x_n, x) - 0.00590527542330779\right]$$

(eq. 2)

MLR Final General Equation Model

$$y_{pred} = 58.0157617357846 - 2.18024298973623 * x1 - 4.84743201279304 * x2 - 0.696485280155724 * x3 + 3.2285511679284 * x4 + 3.70715532584847 * x5 - 1.22507351102573 * x6 + 0.101038748075545 * x7 - 2.18771087558717 * x8 + 0.468454052314145 * x9 + 0.672281415991197 * x10 - 5.05016922656888 * x11 + 1.15602144662931 * x12 - 1.0281249671658 * x13 + 2.55908677847786 * x14 + 3.16461876180707 * x15 + 5.69558378971758 * x16 + 2.27439261499998 * x17 + 0.544392441231361 * x18 - 0.902728946515521 * x19 - 0.654529648448829 * x20 + 1.50664184756437 * x21 + 1.08070337911777 * x22 \text{(eq. 3)}$$

LEGEND:

x1 Clear and Realistic Mission, Vision, and Strategic Goals
x2 Compliance to Quality Standards
x3 Selection of Contract
x4 Timely Completion
x5 Team Composition / Organization Structure
x6 Technical Capability
x7 Project Management
x8 Troubleshooting
x9 Risk Management
x10 Maintaining Health and Safety Program
x11 Effectiveness in Decision Making
x12 Experience
x13 Leadership
x14 Motivation
x15 Good Communication
x16 Financial Condition
x17 Good Cost Proposal / Budget
x18 Procurement Strategy
x19 Economic Condition
x20 Climatic Condition
x21 Client's Financial System
x22 Client's Decision Making
y OVERALL CONTRACTOR'S PROJECT PERFORMANCE (%)

Conclusion

1. The study evaluated the accuracy of the model and found that SVR had the highest R-squared (R2) value of 0.881, Mean Absolute Error (MAE) of 2.172, and Mean Squared Error (MSE) of 7.054. MLR followed with an R2 of 0.874, MAE of 2.180, and MSE of 7.470, while ANN ranked lowest with an R2 of 0.743, MAE of 3.076, and MSE of 15.239.
2. 22 Critical Success Factors are identified using LASSO Regression, with Financial Condition, Effectiveness in Decision-Making, and Compliance to Quality Standards ranking as the top three.
3. Increase in predictive performance of ANN, SVR, and MLR are observed with average increased in R2 of 17.45%, and average decrease of MAE and MSE of 0.724 and 6.523, respectively.
4. With R of more than 0.7 and MAE of less than 10% the of the average actual value from the dataset, ANN, SVR, and MLR can perform a comprehensive diagnostic analysis of construction project success rating.
5. With p-value 0.674, the ANN, SVR, and MLR R2 variances across different data splits (70-30, 80-20, and 90-10) are statistically similar. This implies that the performance of these models, as indicated by the R^2 value, is not significantly influenced by the allocation of data-split ratios.
6. While all three models (ANN, SVR, MLR) achieved similar prediction errors (MAE and MSE), ANN exhibited a statistically different fit (R-squared) compared to the other two. This suggests that ANN captured a more complex relationship between the input and output variables, potentially due to its ability to handle non-linearity. Conversely, SVR, although capable of handling non-linearity, might have primarily focused on the strong linear relationships present in the datasets, leading to similar performance to the simpler MLR model. This implies that for datasets with prominent linear features, SVR might not fully utilize its non-linear capabilities, resulting in performance comparable to linear models like MLR

Recommendation

- To enhance the generalizability of models, future researchers should increase dataset size beyond the current limit of 130 for qualitative research surveys.
- For future studies, the researcher recommends limiting the number of input variables to at least 10-20 per output variable in regression modeling. This may limit the existence of overfitting and multicollinearity.
- It is recommended for future studies to apply the ANN, SVR, and MLR models developed in this research to different construction projects to validate their predictive accuracy and robustness. By using these models on a diverse set of projects, researchers can simulate contractor performance ratings and assess the generalizability of the models across various construction management contexts. Conducting simulations with actual performance data will offer valuable feedback for refining these predictive models, ultimately contributing to more accurate and reliable tools for evaluating contractor performance in the construction industry.
- Future researchers should explore alternative feature selection and factor regularization methods other than LASSO regression to identify critical factors effectively within models, enhancing their predictive accuracy and interpretability.

References

Afolabi, I., Ifunaya, C., Ojo, F., & Moses, C. (2019). A model for business success prediction using machine learning algorithms. Journal of Physics: Conference Series 1299 012050. DOI: 10.1088/1742-6596/1299/1/012050

Al Aisri, H. A., Hamad, R. J. A., & Tayeh, B. A. (2021). Critical factors affecting the success of construction projects in Oman. Journal of Sustainable Architecture and Civil Engineering, 29(2). https://doi.org/10.5755/j01.sace.29.2.29269

Alaeos, A. A. A., Aldaud, A. A. A., Chan, D. W. M., Olawumi, T. O., & Sarvari, H. (2021). Critical success factors for managing construction small and medium-sized enterprises in developing countries of the Middle East: Evidence from Iranian construction enterprises (2021). Journal of Building Engineering, 43. https://doi.org/10.1016/j.jobe.2021.103152

Almuajebh, M., & Gunduz, M. (2020). Critical success factors for sustainable construction project management. Sustainability, 12, 1-6. DOI: 10.48550/arXiv.2211.06561

Altarawneh, J. Y., & Samadi, B. (2019). The relationship between critical success factors and success criteria in construction projects in the United Arab Emirates. International Journal of Advanced and Applied Sciences, 6(7), 43-53. https://doi.org/10.21833/ijaas.2019.07.006

Choi, J., Gu, B., Chin, S., & Lee, J. (2020). Machine learning predictive model based on national data for fatal accidents of construction workers. Automation in Construction, 110, 102974. https://doi.org/10.1016/j.autcon.2019.102974

Dessalegn, M. (2021). Assessment of critical success factors in construction projects performance in South West Ethiopia. Civil and Environmental Research, 13(6). DOI: 10.7176/CER/13-6-01

Faten Albtoush, A.M., Doh, S. I., Rahman, R. A., & Al-Molmani. (2022). Critical success factors of construction projects in Jordan: An empirical investigation. Asian Journal of Civil Engineering, 23, 1087–1099. https://doi.org/10.1007/s42107-022-00470-8

Florez-Perez, L.,Song, Z., & Cortissoz, J.C. Using machine learning to analyze and predict construction task productivity. Computer Aided Civil and Infrastructure Engineering, 37(12), 1602-1616. https://doi.org/10.1111/mice.12806

Gondia, A., Siam, A. S., El-Dakhakhni, W., & Nassar, A. H. (2019). Machine learning algorithms for construction projects delay risk prediction. Journal of Construction Engineering and Management, 146(1). https://doi.org/10.1061/(ASCE)CO.1943-7862.0001736

He, Z., Hu, P., Yan, H., & Zhou, C. (2022). Empirical research on the critical success factors of construction program. Hindawi, Computational Intelligence and Neuroscience. https://doi.org/10.1155/2022/9701963

Kang, H., Zong, X., Wang, J., & Chen, H. (2023). Binary gravity search algorithm and support vector machine for forecasting and trading stock indices. International Review of Economics & Finance, 84, 507-526. https://doi.org/10.1016/j.iref.2022.11.009

Kumar, V., Pandey, A., & Singh, R. (2023). Project success and critical success factors of construction projects: Project practitioners' perspectives. Organization, Technology and Management in Construction, 15 (1), 1-22. https://doi.org/10.2478/otmcj-2023-0001

Love, P. E. D., Matthews, J., Fang, W., Porter, S., Luo, H., & Ding, L. (2021). Explainable artificial intelligence in construction: The content, context, process, outcome. Computer Science. https://doi.org/10.48550/arXiv.2211.06561

Management and Strategy Institute. (2022). Understanding the Student's t-Test in statistical analysis. https://www.msicertified.com/students-t-test/

Ross, G., Das, S., Sciro, D., and Raza, H. (2021). CapitalVX: A machine learning model for startup selection and exit prediction. The Journal of Finance and Data Science, 7, 94-114. https://doi.org/10.1016/j.jfds.2021.04.001

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2 (3). https://doi.org/10.1007/s42979-021-00592-x

Shiksha Online. (2023). What is multilayer perceptron (MLP) https://www.shiksha.com/author/Shiksha2101423374

Yamany, M. S., Abdelhameed, A., Elbeltagi, E., Mohamed, H.A.E., (2024). Critical success factors of infrastructure construction projects. Innovative Infrastructure Solutions, 9(95). https://doi.org/10.1007/s41062-024-01394-9