International Journal of Technology v(i) pp-pp (YYYY)
Received Month Year / Revised Month Year / Accepted Month Year

International Journal of Technology

http://ijtech.eng.ui.ac.id

Developing a Machine Learning Model to Improve the Accuracy of Owner Estimate Cost (OEC) in the Capital Expenditure Procurement Process

Mohammed Ali Berawi[1,2], Mustika Sari[1,3*], Nur Hayati A. Amiri[2], Suci Indah Susilowati[2], Siti Rahma Utami[2], Anastasia Kulachinskaya[4]

[1]*Department of Civil Engineering, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia*
[2]*Center for Sustainable Infrastructure Development, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia*
[3] *Department of Interdisciplinary Engineering, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, Depok 16424, Indonesia*
[4]*St. Petersburg Polytechnic University, Graduate school of industrial economics Polytechnicheskaya, 29, St. Petersburg, 195251, Russia*

Abstract. Inaccurate Owner Estimate Cost (OEC) calculations often lead to procurement failures in the procurement process, which can affect the success of government capital expenditure projects in Indonesia. As the OEC becomes a critical benchmark for assessing the fairness of bids, and errors in its calculation can cause financial mismanagement and regulatory issues, this study aims to improve its accuracy by developing a Machine Learning (ML) model using Linear Regression (LR) algorithm capable of predicting the price fluctuations in the procurement of government-owned building projects. Data from a state-owned building construction project and the data from Analysis of Work Unit Prices for 2017-2020 were analyzed and used to predict 2021 price amandment for various construction work items in 2021. The developed ML model demonstrated robust accuracy, with Root Mean Squared Error (RMSE) values ranging from 0.012 to 0.037 and Mean Absolute Error (MAE) values between 0.011 and 0.029 across job descriptions, indicating a strong fit. Results of this study highlight the superiority of the developed model over similar studies in terms of precision and interpretability, offering a solution to enhancing procurement decision-making.

Keywords*:* Capital Expenditure Procurement; Machine Learning; Owner Estimate Cost

1. Introduction

Capital expenditure is one of the most important categories in public sector institutions, as it deals with acquiring fixed assets, such as land, buildings, buildings, and equipment (Sutopo & Siddi, 2018). Therefore, capital expenditure procurement has been a key performance indicator and a driving force for economic growth, especially in infrastructure and development projects. Central and local governments both focus on optimizing their revenue generation to fund capital expenditures, which is necessary to develop basic infrastructure, improve public services, and encourage national development (Mukmin et al., 2020).

*Corresponding author's email: mustika.s@ui.ac.id, Tel.: +62-21-7270029; Fax: +62-21-7270028
doi: 10.14716/ijtech.v0i0.0000

However, the realization of capital expenditure in Indonesia has often suffered from a high procurement failure rate, driven by some factors such as non-compliance with regulations, experience among estimators, pricing survey errors, inadequate banking support, and inaccuracies from Owner Estimate Cost (OEC) estimation (Bosio et al., 2023; Strang, 2021). OEC represent the prices estimated for projects or services that incorporate all cost components from the initial design until the project is ready to be delivered to the user (Koo et al., 2010). In government procurement, the OEC includes Value-Added Tax (VAT) once the technical specifications in Terms of Reference (TOR) have been finalized. Inaccurate OEC estimations can create significant risks within the procurement process, potentially leading to financial and legal consequences. OEC preparation should follow a transparent and accountable methodology to mitigate these risks, using up-to-date and reliable data. A comprehensive market analysis before selecting suppliers is crucial to avoid procurement failures (Safa et al., 2014).

In the tendering phase, OEC serves as a benchmark for the determination of reasonable prices considering market conditions. Setting the OEC too low can deter competitive bids, leading to project delays and failures. On the other hand, an overestimated OEC can result in government overspending, which could attract public criticism and raise concerns about financial mismanagement (Astana et al., 2023).

Inaccurate OECs elevate the risk of disputes between contractors and clients. Moreover, significant gaps between the OEC and the final bid can create financial instability for contractors, negatively affecting project timelines and quality (Mohamed et al., 2011; Oo et al., 2022). Contract changes in construction projects often occur due to the combined effects of factors such as scope adjustment, scheduling changes, or cost revisions (Khoso et al., 2019). These changes are referred as Addendums, Contract Change Orders (CCOs), and Variation Orders (VOs). Addendums are additions to the original contract mutually agreed upon by the parties to address inadequate or missing requirements. CCOs are revisions in work volume or project timeline without altering the core contract clauses, while VOs are the changes in scope, specifications, costs, and regulations (Mohammad et al., 2017).

Since procurement makes up a large portion of national capital expenditures, inaccurate OECs directly hinder infrastructure projects' timely completion and cost-efficiency, which can limit government performance (Komakech, 2016). The gap between the winning bid and the OEC is a significant risk factor during the pre-tendering phase, as it can lead to project cancellations, scope reductions, or schedule delays (Almohsen et al., 2023). Economic volatility, incorrect pricing surveys, and estimators' low experience gaps can further complicate these gaps (Liu & Zhu, 2023).

Machine Learning (ML), a branch of Artificial Intelligence (AI), has the potential to enhance the accuracy of the OEC and reduce the risks linked to procurement failures. It uses algorithms to analyze large datasets, uncover patterns, and produce more precise predictions to inform decision-making (Berawi et al., 2019; Ma'ruf et al., 2024; Sari et al., 2023). Cost estimation has the potential to be an alternative to conventional methods dependent on basic statistical analyses, and it may fail to consider market volatility or complex cost structures (Budiono et al., 2014; Creedy et al., 2010; Hashemi et al., 2020).

The application of ML in predicting the OEC during the construction projects' pre-tender phase has been investigated in previous studies. For example, Li et al. (2022) aimed to improve OEC accuracy by employing a feedforward neural network (FFNN) within the framework of Friedman's model that can forecast the lowest submitted bid more precisely. Similarly, Almohsen et al. (2022) created an advanced model combining Artificial Neural Networks (ANN), Deep Neural Networks (DNN), and Time Series (TS) techniques to estimate the ratio of the lowest bid to the OEC in the pre-tender phase across various contract types and sizes. Furthermore, Alsugair et al. (2023) used an ANN model to predict Final Contract Costs (FCC) based on the initial OEC, using Linear Regression (LR) to process data, a square root function for data transformation, as well as Zavadskas and Turskis' logarithmic method for data standardization.

These studies demonstrate the promise of ML for improving OEC accuracy; however, they also reveal key challenges. Neural network (NN)-based approaches often suffer from a lack of transparency in their decision-making process, making it difficult for stakeholders to interpret cost drivers (Ribeiro et al., 2016). This limitation hinders the practical adoption of such models, particularly in public-sector procurement frameworks where interpretability is crucial. This study addresses these challenges by employing a Linear Regression (LR)-based ML model. Unlike NN, LR offers transparency that helps understand the relationships between historical price adjustment as independent variables and the OEC as the dependent variable.

2. Methods

This study aimed to enhance the accuracy of OEC predictions in capital expenditure procurement through the development of an ML model. To achieve this objective, a systematic approach was adopted, combining theoretical insights with practical implementation. The theoretical foundation of the research is grounded in data-driven decision-making theory that emphasizes the use of historical data and statistical modelling to improve predictive accuracy (Montgomery et al., 2012). To operationalize this approach, the Cross Industry Standard Process for Data Mining (CRISP-DM) framework was adopted. It encompasses six iterative phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment (Mandolini et al., 2024; Schröer et al., 2021).

In the business understanding phase, the study identified the need to improve the OEC accuracy in capital expenditure project. Subsequently, during the data understanding phase, procurement contract data for this study was derived from a state-owned building construction project, referred to as the SM project. This data was obtained from a Jakarta-based construction company and included historical records of amended items, price addendums, work volumes, and Work Unit Prices.

The historical price data used spans the period from 2017 to 2020, which was chosen for its completeness and alignment with the SM project's initial contracts and price adjustments, further ensuring the consistent patterns for cost estimation. Economic disruptions from 2020 onwards, such as the COVID-19 pandemic, were excluded to maintain dataset consistency. Exploratory Data Analysis (EDA) was conducted to verify the absence of missing values and to examine relationships between variables.

In the data preparation phase, the data was processed and transformed into a structured format for ML modelling. Work Unit Prices were categorized into independent (YEAR) and dependent (WORK ITEM PRICE) variables. For OEC calculation, the study utilized information on amended items in the SM project, including work volumes and Work Unit Prices, that were analyzed using coefficients from the Unit Price Analysis and Basic Unit Prices for labor, materials, and equipment sourced from historical data published in the Journal of Unit Prices for Building Materials. This analysis was critical to understanding the standards, specifications, and material costs associated with tasks that experienced price adjustments, providing a detailed basis for model development.

A unit price analysis model, developed in accordance with the Indonesian Ministry of Public Works and Housing Regulation No. 28, 2016, was integrated into a database of calculated unit prices for construction works experiencing price addendums. This database formed the basis for identifying attributes to be used in the ML model development.

Several assumptions were made to ensure the reliability of the model, including (1) the representation of pricing trends reflects cost adjustment factors in construction projects; (2) the dataset contained no missing values verified through exploratory data analysis, and (3) pricing trends were linear and followed consistent patterns over the years analyzed. The modelling phase involved developing an LR-based ML, which was implemented in Python-3 within the Jupyter Notebook environment. The LR algorithm was selected for its ability to predict continuous variables based on historical data and its simplicity for interpreting relationships between independent and dependent variables. The model was trained using an 80:20 split of the dataset; hence, the model can learn from existing patterns while testing its accuracy on unseen data. Predictions for the year 2021 were made using the trained model, analyzing pricing trends across various work items.

In the evaluation phase, the model's accuracy was assessed using two metrics: Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). These metrics are valuable because they indicate error in the units of the constituent of interest, which aids in the analysis of results (Hodson, 2022). RMSE is a widely recognized metric for evaluating regression models that refers to the square root of the average of the squares of differences between predicted and observed values over the total number of data points(Chicco et al., 2021; Oke et al., 2020). Mathematically, this is represented with this formula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}{n}} \quad (1)$$

Where:

y_i = observed value

$\bar{y}_i$ = predicted value

i = index of the data in the database

n = total number of data points

MAE is determined by computing the absolute errors between each predicted and the actual value, then finding the mean by evaluating the entire dataset. This is performed by subtracting the mean value from each data point, summing the results, then dividing by the total number of datasets. The formula for MAE is:

$$MAE = \frac{1}{n}\sum xi - x \quad (2)$$

Where:

xi = the actual value

x = the predicted value

n = the total number of values

RMSE and MAE values of 0 indicate a perfect fit (Moriasi et al., 2007). RMSE and MAE values below half the standard deviation of the measured data can be deemed modest, and either metric is suitable for model evaluation. Moreover, RMSE emphasizes larger errors because of the squaring of residuals, whereas MAE offers a direct average error metric.

Though this study did not include full-scale implementation of the deployment phase, a structured and transparent methodology for developing and validating the ML model can be obtained by adopting CRISP-DM framework. Figure 1 illustrates the research workflow.

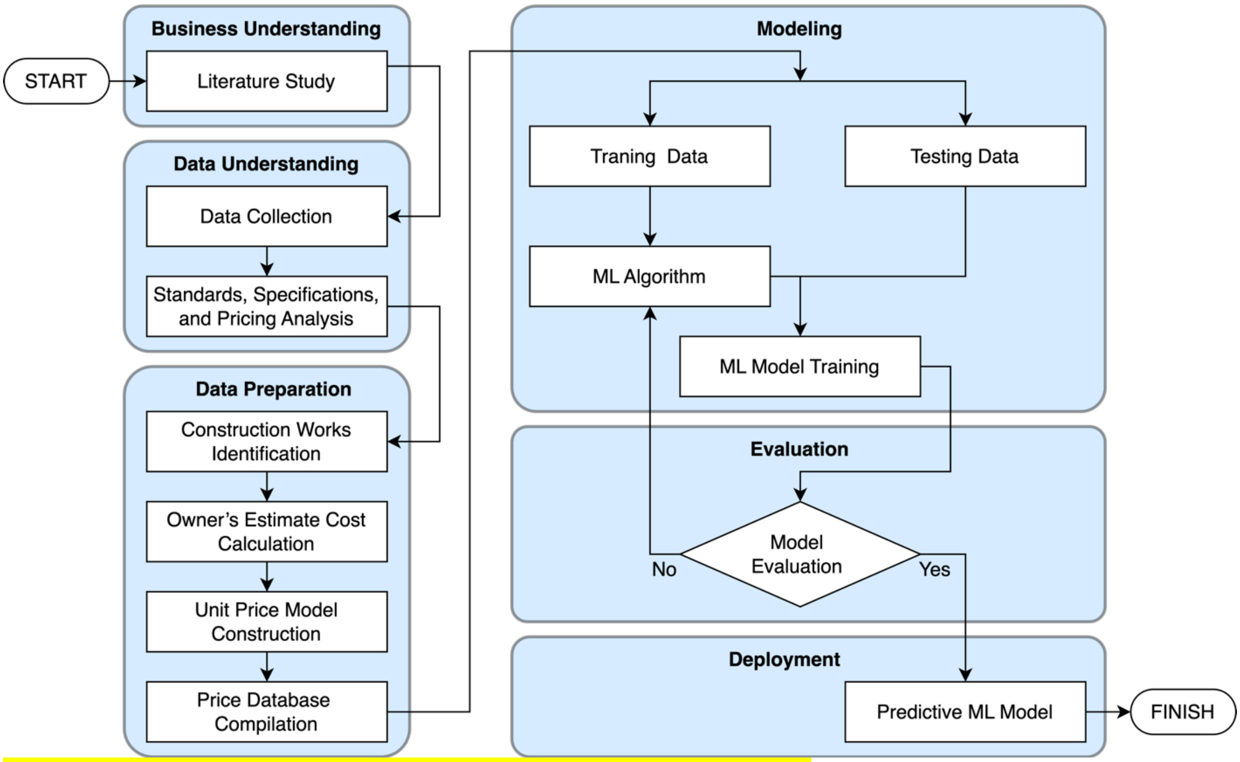

Figure 1. Research Workflow within CRISP-DM framework

3. Results and Discussion

3.1. Identifying the Variables for the ML Model Development

Data from the procurement contract of the SM project, originally established in 2020, contained works whose prices were adjusted with an addendum in 2021. The contract data includes both the original values of OEC and the price revisions that were later affected in one place. Table 1 summarizes in detail the adjustments involved, especially in architectural works. Six work items, which have undergone price changes from the initial contract in 2020 to the 2021 addendum, were selected for further analysis. This analysis is required to develop the ML model that predicts future pricing trends. The adjustments were necessary because the original OEC did not take on post-award signing price increases.

Table 1. Initial OEC and Price Addendums in SM Project

No	Work Description	2020's OEC (USD)	Additional Work (USD)	2021's Amended Contract (USD)
A	Architectural Work			
I	Preparation Work	7,697		7,697
II	Demolition Work			
	A. Main Building	21,704		21,704
	B. Supporting Building	2,007		2,007
	C. Exterior Spaces	4,238		4,238
III	Brick Wall Masonry Works			
	A. Main Building Ground Floor	15,020	1,474	16,218
	B. Main Building Upper Floor	4,115		4,115
	C. Front Supporting Facilities Building	1,391		1,391
	D. Rear Supporting Facilities Building	1,548	84	1,640
	E. Exterior Spaces	4,335		4,335

No	Work Description	2020's OEC (USD)	Additional Work (USD)	2021's Amended Contract (USD)
IV	Gypsum Partition Wall Pairing + Hollow Iron Frame			
	A. Main Building Ground Floor	4,998	264	5,261
	B. Main Building Upper Floor	2,390		2,390
V	Floor & Wall Finishing Work			
	A. Main Building Ground Floor	28,772	2,152	29,310
	B. Main Building Upper Floor	5,104		5,104
	C. Front Supporting Facilities Building	5		4,863
	D. Front Exterior Space	11,140	327	11,343
	E. Rear & Side Space	14,730	511	15,172

The items included in the sub-items of the six work items listed in Table 1 include a number that requires detailed scrutiny. Each sub-item relates to particular contracts with third-party contractors for construction procurement and contains confidential information regarding pricing and contractual terms. To accurately calculate the OEC for each sub-item, it's crucial to have data on the work volumes and unit prices. After pinpointing the work items that experienced price adjustments through addendums, 16 Work Unit Price items were extracted for further analysis using the coefficients from the Analysis of Work Unit Prices from 2017 to 2020. These datasets were run 10 times for each Work Unit Price item, resulting in a total of 64 data points. To ensure robustness, the data was processed 640 times in the ML model, providing sufficient volume and variability for effective ML model training. The trend, represented in Table 2, creates a background understanding of price changes throughout time and helps predict near-future cost trends more effectively.

Table 2. 16 price-adjusted work items and the Analysis of Work Unit Price (2017-2020)

No.	Work Type	Code	2017 (USD)	2018 (USD)	2019 (USD)	2020 (USD)
1	Standard Brick Wall Masonry (1/2 Brick)	WORK1	9.88	11.59	11.55	17.84
2	Standard Plastering + Finishing	WORK2	6.15	6.61	7.08	7.88
3	Door & Windowsills	WORK3	6.15	6.40	6.58	7.04
4	Brick Masonry for Foundation Wall	WORK4	10.16	11.98	12.08	18.42
5	Waterproof Plastering + Finishing	WORK5	6.31	6.83	7.37	8.21
6	Gypsum Partition Wall Masonry	WORK6	4.64	5.09	5.22	5.62
7	Hollow Steel Partition Frame	WORK7	19.10	19.98	23.76	21.81
8	Paving Floor	WORK8	19.16	19.65	19.74	21.55
9	Installation & Procurement of Walls	WORK9	12.27	12.06	12.37	13.30
10	Backfill / Red Soil (Planting Medium)	WORK10	5.60	5.83	5.94	6.72
11	Granite Floor	WORK11	36.47	37.41	39.96	41.66
12	Installation & Supply of Curbstones	WORK12	12.60	12.37	12.74	13.65
13	Granite Tile Polish Floor	WORK13	43.78	42.08	44.22	49.23
14	Rough Motif Ceramic floor	WORK14	22.32	24.41	25.70	37.33
15	Granite Tile Wall	WORK15	19.99	21.04	21.30	26.68
16	Floor Plint	WORK16	6.03	5.63	5.48	5.72

The Analysis of Work Unit Prices for 2017-2020 was used to classify work items, summarized in Table 2. This classification selected the YEAR variable as the input variable (x). At the same time, the price of each work item was set as the dependent response variable (y), labeled from WORK1 to WORK16. An LR-based ML model was employed to predict future price trends for these work items.

3.2 Developing ML Model

The ML model was developed using Python-3 within the web-based Jupyter Notebook application, the first step of which was importing several libraries. Pandas was utilized for data manipulation and structuring, Matplotlib was used for visualization and mathematical functions, and Scikit-learn was used to implement the ML model, specifically the LR model employed in this study.

Subsequently, the data was loaded in two key columns that serve as the variables in the ML development process. The first column, labeled "YEAR," was designated as the independent variable (x), while the second column, "WORK," was the dependent variable (y). Once the data was loaded, the system displayed the data count, types, and memory allocation to ensure everything was correct and consistent before continuing to the next step.

After that, the missing values was checked. There were no missing values if the result was zero (0), meaning the analysis could continue without data imputation. It was then followed by the exploratory data analysis (EDA), which started with a bivariate analysis of the relationship between the YEAR and WORK variables using a scatter plot. This plot helped visualize the relationship between these two factors before proceeding with the modelling process. The correlation was also calculated, which came out to a coefficient of 0.7, indicating a strong positive relationship between YEAR and WORK. This means that as the YEAR variable changed, the WORK variable significantly affected it.

The core ML modeling process began after all those steps were conducted. It included establishing the X and Y variables, splitting the data into training and testing sets using an 80:20 ratio, and applying the LR algorithm to train the model. During this phase, the model learned from the input data and identified the correct slope and coefficients to make accurate predictions. Finally, the model was run through several iterations to fine-tune and optimize the prediction results. This process ensured that the model's predictions were reliable and could improve the accuracy of OEC estimation in building construction projects.

3.3. Evaluating the Accuracy of the ML Model Predictions

The developed ML model generated the optimal prediction results through a repetitive process using variance, where it used slightly different training and testing data sets each time the algorithm ran in a process. Variance refers to how sensitive an algorithm is to the specific data it uses during training; it shows the different results when there are small changes in the data and how the model is trained (Raste et al., 2022).

Several ML algorithms that are not deterministic are stochastic, which implies that the behaviour of the algorithm is influenced by randomness while it is being trained (Barmpalias et al., 2017). On the other hand, being stochastic does not necessarily mean being completely random. Stochastic ML algorithms still learn from the given historical data, but small decisions made during the learning process may vary randomly from one iteration to the next. As a result, each time a stochastic ML algorithm is run on the same data, the model produced may vary slightly, resulting in different predictions for the test data. The performance of such stochastic models can be summarized as a distribution with an expected mean error or accuracy and a standard deviation that reflects the level of randomness in the prediction results.

The final prediction results provided price estimates for each work item, allowing for accurate analysis of future trends. The ML model predictions, after undergoing the processes of data training, data testing, and averaging the distribution of the predictions, are presented in Table 3.

Table 3. Machine Learning Prediction Results with Price Estimates from 2017–2021

No	Work Description	2017 (USD)	2018 (USD)	2019 (USD)	2020 (USD)	2021 (USD)
1	Standard Brick Wall Masonry (1/2 Brick)	9.88	11.59	11.55	17.84	19.29
2	Standard Plastering + Finishing	6.15	6.61	7.08	7.88	8.39
3	Door & Windowsills	6.15	6.40	6.58	7.04	7.28
4	Brick Masonry for Foundation Wall	10.16	11.98	12.08	18.42	19.99
5	Waterproof Plastering + Finishing	6.31	6.83	7.37	8.21	8.78
6	Gypsum Partition Wall Masonry	4.64	5.09	5.22	5.62	5.88
7	Hollow Steel Partition Frame	19.10	19.98	23.76	21.81	23.91
8	Paving Floor	19.16	19.65	19.74	21.55	22.02
9	Installation & Procurement of Walls	12.27	12.06	12.37	13.30	13.46
10	Backfill / Red Soil (Planting Medium)	5.60	5.83	5.94	6.72	6.97
11	Granite Floor	36.47	37.41	39.96	41.66	43.40
12	Installation & Supply of Curbstones	12.60	12.37	12.74	13.65	13.82
13	Granite Tile Polish Floor	43.78	42.08	44.22	49.23	50.00
14	Rough Motif Ceramic floor	22.32	24.41	25.70	37.33	39.02
15	Granite Tile Wall	19.99	21.04	21.30	26.68	27.33
16	Floor Plint	6.03	5.63	5.48	5.72	5.52

The percentage increase in OEC prices was analyzed by multiplying the price database and the predicted price results by the building volume of the SM project. The percentage increase in prices between 2020 and 2021 was then calculated. To evaluate and compare the accuracy of the ML predictions, the percentage price increase was determined for both the predicted data and the actual SM project data for each sub-work item (see Table 4).

Table 4. Comparison of OEC Price Increases Between Actual and Predicted ML Prices

No	Work Description	Initial Contract (USD)	Price Addendum (USD)	Price Increase (%)
III.A	Brick Wall Masonry Work (Main Building)	15,020	16,218	7.98%
III.D	Brick Wall Masonry Work (Rear Supporting Building)	1,548	1,640	5.89%
IV.A	Gypsum Partition Wall Pairing (Main Building)	4,998	5,261	5.28%
V.A	Floor & Wall Finishing (Main Building)	28,772	29,310	1.87%
V.D	Floor & Wall Finishing (Front Exterior Space)	11,140	11,343	1.82%
V.E	Floor & Wall Finishing (Rear & Side Space)	14,730	15,172	3.00%

Accurate cost estimation is critical for project success (Mokoena et al., 2023). The predicted price increases from 2020 to 2021 were compared with the actual price increase data from the SM project to evaluate the accuracy improvement of OEC calculations for capital expenditure procurement implementation using the ML model. The result of RMSE and MAE evaluations indicated that the ML predictions achieved a high level of accuracy and a good fit with the actual data (see Table 5).

Table 5. RMSE and MAE Results compare ML prediction accuracy with actual project data

No.	Job Description	RMSE	MAE
III.A	Brick Wall Masonry Work (Main Building)	0.020	0.015
III.D	Brick Wall Masonry Work (Rear Supporting Building)	0.018	0.016
IV.A	Gypsum Partition Wall Pairing (Main Building)	0.031	0.025
V.A	Floor & Wall Finishing (Main Building)	0.037	0.029
V.D	Floor & Wall Finishing (Front Exterior Space)	0.014	0.012
V.E	Floor & Wall Finishing (Rear & Side Space)	0.012	0.011

The findings of this study demonstrate that the developed ML model for OEC prediction achieved robust performance, with RMSE values between 0.012 and 0.037 and MAE values ranging from 0.015 to 0.029 across several construction job descriptions. In contrast, Zhang et al. (2023)

utilized a combination of extreme gradient boosting (XGBoost) and Bayesian optimization (BO), achieving RMSE and MAE of 0.8690 and 0.4875, respectively. These higher error metrics are likely due to their focus on conceptual cost estimation for diverse infrastructure projects, which involved heterogeneous datasets with greater variability. The LR model developed in this study demonstrates higher precision by focusing on specific procurement datasets with consistent characteristics. Similarly, Sanni-Anibire et al. (2021) employed k-Nearest Neighbors (KNN) to model cost estimations for tall buildings, reporting an RMSE of 6.09. The larger error metrics in their study can be attributed to the broader scope of tall building projects and the variability in cost structures. On the other hand, this paper's LR model benefits from a more specific dataset.

Furthermore, the methodological selections affect the variations. This study has used LR for simplicity and interpretability, other studies have relied on more complicated algorithms. While they may capture non-linear relationships more effectively, they often involve reduced interpretability and increased computational requirements.

4. Conclusions

Inaccurate OEC in the capital expenditure procurement process can cause project delays and financial mismanagement. This study investigated the usage of ML to enhance the accuracy of OEC estimates by developing an ML model with the LR algorithm. Though the dataset used in this study was derived from a specific project, the integration of detailed price addendums and historical adjustments ensures that the model captures key cost-driving trends applicable to similar procurement scenarios. The findings showed that the model significantly improved OEC estimation accuracy, as evidenced by low RMSE values that indicate a strong fit between predicted and actual prices. However, there are some drawbacks in this study, one being its reliance on data from 2017 to 2020 data that restricts the model's relevance to more recent market trends and conditions. Therefore, further research can include more recent data and explore advanced ML techniques, such as deep learning models, to enhance the precision of OEC predictions in capital expenditure procurement.

References

Almohsen, A. S., Alsanabani, N. M., Alsugair, A. M., & Al-Gahtani, K. S. (2023). Integrated artificial and deep neural networks with time series to predict the ratio of the low bid to owner estimate. *Engineering, Construction and Architectural Management*, *31*(13). https://doi.org/10.1108/ECAM-05-2023-0454

Alsugair, A. M., Alsanabani, N. M., & Al-Gahtani, K. S. (2023). Forecasting the Final Contract Cost on the Basis of the Owner's Cost Estimation Using an Artificial Neural Network. *Buildings*, *13*(3). https://doi.org/10.3390/buildings13030786

Astana, I. N. Y., Wiryasa, N. A., & Pinakesty, S. A. P. A. (2023). The Relationship Below 80% of the Owner Estimate Price on Construction Projects to Project Performance. *Journal of Asian Multicultural Research for Economy and Management Study*, *4*(1). https://doi.org/10.47616/jamrems.v4i1.396

Barmpalias, G., Cenzer, D., & Porter, C. P. (2017). Random numbers as probabilities of machine behavior. *Theoretical Computer Science*, *673*. https://doi.org/10.1016/j.tcs.2017.02.001

Berawi, M. A., Leviakangas, P., Muhammad, F., Sari, M., Gunawan, Yatmo, Y. A., & Suryanegara, M. (2019). Optimizing search and rescue personnel allocation in disaster emergency response using fuzzy logic. *International Journal of Technology*. https://doi.org/10.14716/ijtech.v10i7.3709

Bosio, E., Hayman, G., & Dubosse, N. (2023). The Investment Case for E-Government Procurement: A Cost-Benefit Analysis. *Journal of Benefit-Cost Analysis*, *14*. https://doi.org/10.1017/bca.2023.10

Budiono, H. D. S., Kiswanto, G., & Soemardi, T. P. (2014). Method and model development for manufacturing cost estimation during the early design phase related to the complexity of the machining processes. *International Journal of Technology*, *5*(2). https://doi.org/10.14716/ijtech.v5i2.402

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, *7*. https://doi.org/10.7717/PEERJ-CS.623

Creedy, G. D., Skitmore, M., & Wong, J. K. W. (2010). Evaluation of Risk Factors Leading to Cost Overrun in Delivery of Highway Construction Projects. *Journal of Construction Engineering and Management*, *136*(5). https://doi.org/10.1061/(asce)co.1943-7862.0000160

de Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. *Neurocomputing*, *192*. https://doi.org/10.1016/j.neucom.2015.12.114

Hashemi, S. T., Ebadati, O. M., & Kaur, H. (2020). Cost estimation and prediction in construction projects: a systematic review on machine learning techniques. In *SN Applied Sciences* (Vol. 2, Issue 10). https://doi.org/10.1007/s42452-020-03497-1

Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development*, *15*(14), 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

Khoso, A. R., Khan, J. S., Faiz, R. U., & Akhund, M. A. (2019). Assessment of Change Orders Attributes in Preconstruction and Construction Phase. *Civil Engineering Journal (Iran)*, *5*(3). https://doi.org/10.28991/cej-2019-03091273

Komakech, R. (2016). Public Procurement in Developing Countries: Objectives, Principles and Required Professional Skills. *Publicy Policy and Administration Research*, *6*(8).

Koo, C. W., Hong, T. H., Hyun, C. T., Park, S. H., & Seo, J. O. (2010). A study on the development of a cost model based on the owner's decision making at the early stages of a construction project. *International Journal of Strategic Property Management*, *14*(2). https://doi.org/10.3846/ijspm.2010.10

Li, M., Zheng, Q., & Ashuri, B. (2022). Predicting Ratio of Low Bid to Owner's Estimate Using Feedforward Neural Networks for Highway Construction. *Construction Research Congress 2022: Project Management and Delivery, Controls, and Design and Materials - Selected Papers from Construction Research Congress 2022*, *3–C*. https://doi.org/10.1061/9780784483978.036

Liu, M., & Zhu, C. (2023). A Real Option Pricing Decision of Construction Project under Group Bidding Environment. *Applied Sciences (Switzerland)*, *13*(2). https://doi.org/10.3390/app13021130

Mandolini, M., Manuguerra, L., Sartini, M., Presti, G. M. L., & Pescatori, F. (2024). A cost modelling methodology based on machine learning for engineered-to-order products. *Engineering Applications of Artificial Intelligence*, *136A*(108957).

Ma'ruf, A., Nasution, A. A. R., & Leuveano, R. A. C. (2024). Machine Learning Approach for Early Assembly Design Cost Estimation: A Case from Make-to-Order Manufacturing Industry. *International Journal of Technology.*, *15*(4), 1037–1047.

Mohamed, K. A., Khoury, S. S., & Hafez, S. M. (2011). Contractor's decision for bid profit reduction within opportunistic bidding behavior of claims recovery. *International Journal of Project Management*, *29*(1). https://doi.org/10.1016/j.ijproman.2009.12.003

Mohammad, N., Che-Ani, A. I., & Rakmat, R. A. O. K. (2017). Causes and effects of variation orders in the construction of terrace housing projects: A case study in the State of Selangor, Malaysia. *International Journal of Supply Chain Management*, *6*(1).

Mokoena, G. M., Makhanya, B. B. S., & Nel, H. (2023). *Assessing Project Cost Planning in the Construction Industry*. https://doi.org/10.46254/an13.20230144

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to Linear Regression Analysis. In *Sloan Management Review,*. John Wiley & Sons.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3).

Mukmin, M., Kusuma, I., & Susandra, F. (2020). The Governments Capital Expenditure: A View from Central Government and Local Government in Indonesia. *American Based Research Journal*, *9*(10), 18–23.

Oke, E. O., Adeyi, O., Okolo, B. I., Adeyi, J. A., Ayanyemi, J., Osoh, K. A., & Adegoke, T. S. (2020). Phenolic compound extraction from Nigerian Azadirachta Indica leaves: Response surface and neuro-fuzzy modelling performance evaluation with Cuckoo Search multi-objective optimization. *Results in Engineering*, *8*. https://doi.org/10.1016/j.rineng.2020.100160

Oo, B. L., Lim, T. H. B., & Runeson, G. (2022). Critical Factors Affecting Contractors' Decision to Bid: A Global Perspective. *Buildings*, *12*(3). https://doi.org/10.3390/buildings12030379

Raste, S., Singh, R., Vaughan, J., & Nair, V. N. (2022). Quantifying Inherent Randomness in Machine Learning Algorithms. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4146989

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should i trust you?" Explaining the predictions of any classifier. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13-17-August-2016*. https://doi.org/10.1145/2939672.2939778

Safa, M., Shahi, A., Haas, C. T., & Hipel, K. W. (2014). Supplier selection process in an integrated construction materials management model. *Automation in Construction*, *48*. https://doi.org/10.1016/j.autcon.2014.08.008

Sanni-Anibire, M. O., Mohamad Zin, R., & Olatunji, S. O. (2021). Developing a preliminary cost estimation model for tall buildings based on machine learning. *International Journal of Management Science and Engineering Management*, *16*(2). https://doi.org/10.1080/17509653.2021.1905568

Sari, M., Berawi, M. A., Larasati, S. P., Susilowati, S. I., Susantono, B., & Woodhead, R. (2023). Developing Machine Learning Model to Predict HVAC System of Healthy Building: A Case Study in Indonesia. *International Journal of Technology*, *14*(7). https://doi.org/10.14716/ijtech.v14i7.6682

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, *181*(2019), 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Strang, K. D. (2021). Which organizational and individual factors predict success vs. failure in procurement projects. *International Journal of Information Technology Project Management*, *12*(3). https://doi.org/10.4018/IJITPM.2021070102

Sutopo, B., & Siddi, P. (2018). Capital expenditures and performance of local government administration. *Polish Journal of Management Studies*, *17*(1). https://doi.org/10.17512/pjms.2018.17.1.18

Zhang, J., Yuan, J., Mahmoudi, A., Ji, W., & Fang, Q. (2023). A data-driven framework for conceptual cost estimation of infrastructure projects using XGBoost and Bayesian optimization. *Journal of Asian Architecture and Building Engineering*. https://doi.org/10.1080/13467581.2023.2294871