INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE (IJRISS)
ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Optimising Construction Contract Rates with Z-Score and Machine Learning Approaches

Mohd Nasir Saari*, Mohd Adza Arshad, Nor Zuraiya Mt Amin

Centre of Excellence for Engineering and Technology (CREaTE), Public Works Department Malaysia

DOI: https://doi.org/10.47772/IJRISS.2026.10100289

Received: 20 December 2025; Accepted: 20 January 2026; Published: 07 February 2026

ABSTRACT

This study evaluates two common techniques for contracting rate rationalisation in construction using Z-score statistical screening and machine learning (ML) models on a real-world tender dataset (single project; ten tenderers). We compare a human-in-the-loop, two-stage process where Z-scores act as an initial, auditable filter to detect unusual bids, and ML (regression and clustering) develops context-aware price benchmarks from the multivariate features. Empirically, the ML method achieved a 38% reduction in the RMSE compared to a Z-score-only baseline, whereas the hybrid approach reduced false alarms and produced more consistent and clearer pricing ranges for decision support. Z-score screening maintained operational value by standardising anomaly detection and enhancing procedural transparency, whereas ML improved the predictive accuracy and adaptability to item, project, and market differences. The results show that combining transparent statistical rules with data-driven models improves the accuracy, efficiency, and fairness of tender assessments. The contributions of this study are threefold: (i) a direct comparison of Z-score versus ML at the Bill-of-Quantities level; (ii) a governance-ready hybrid protocol that blends auditable thresholds with model-based benchmarks; and (iii) practical guidance for agencies seeking to embed analytics within e-procurement systems. Although the scope is limited to a single-project dataset, the findings support further validation across multiple projects and real-time deployment with drift monitoring in future research.

Keywords: rate rationalisation, Z-score, machine learning, procurement governance, construction tender, predictive cost modelling.

INTRODUCTION

Background

The construction industry remains a cornerstone of national development, with rate rationalisation serving as a pivotal control mechanism to ensure equitable and market-consistent tender pricing. Historically, rationalisation has relied on expert judgement and fragmented historical data, leaving the process vulnerable to inconsistencies, spreadsheet errors, and limited reproducibility under audit. Therefore, automation is essential to cope with cost volatility and scope heterogeneity, embedding standardised diagnostics and decision rules within tender evaluation workflows (Saari et al., 2024; Choi et al., 2021).

Historically, rationalisation has been conducted manually, relying on expert judgement and fragmented historical datasets. While domain expertise is indispensable, manual processes are vulnerable to several well-documented limitations: inconsistent application of criteria across projects, cognitive and confirmation biases in interpreting outliers, spreadsheet and transcription errors, and limited reproducibility of decisions when challenged in audits or disputes. These vulnerabilities are exacerbated by the sector’s volatility, rapid movements in material and logistics costs, shifting labour dynamics, and project-specific complexities, producing inefficiencies and price dispersion that can distort competitive equilibrium (Saari et al., 2024). In short, a predominantly manual paradigm struggles to keep pace with data-rich, high-variance environments.

Converging evidence indicates that the systematic automation of key rationalisation tasks can materially improve consistency, auditability, and throughput without displacing professional judgement (Choi et al., 2021).

Automation, as framed in recent construction informatics scholarship, does not imply a "black box" substitution of human expertise; rather, it codifies accepted rules and embeds statistical diagnostics into repeatable workflows that surface anomalies and standardise documentation (Choi et al., 2021). This human-in-the-loop model enhances procedural fairness and defensibility: analysts interrogate flagged items, record the rationale within the system, and generate a complete decision trail suitable for internal review and external scrutiny.

Operationally, an automated rationalisation pipeline can integrate (i) structured price databases and bill-of-quantities (BoQ) metadata; (ii) statistical screens such as standardised residuals or Z-scores and dispersion measures for within-item comparability; and (iii) decision rules that translate diagnostics into action (e.g. triggering a deeper review when prices breach pre-set thresholds). Embedding these components in an e-procurement context yields tangible benefits: faster cycle times, consistent application of thresholds across projects, and machine-generated audit logs that reduce post-award contestation (Choi et al., 2021; Saari et al., 2024). Crucially, automation clarifies where expert judgement is most needed on genuine anomalies and complex scope nuances rather than diffusing attention across all items.

Against this backdrop, the research problem is twofold: First, despite growing interest, many organisations lack empirically validated, context-specific frameworks that translate the promise of automation into robust procurement practice; off-the-shelf tools seldom reflect local market structures, BoQ conventions, or regulatory constraints (Saari et al., 2024). Second, the literature has yet to conclusively demonstrate how automated rationalisation reshapes outcome metrics that matter to public clients, award price rationality, dispersion reduction, and audit performance beyond efficiency gains in processing time (Choi et al., 2021). Addressing these gaps requires a design-science approach that (a) formalises rationalisation logic, (b) operationalises transparent statistical tests within a governed workflow, and (c) evaluates the impact on pricing behaviour and decision quality across real procurement datasets. This study is situated precisely at that intersection, extending prior work by articulating and testing an automation-enabled rationalisation framework tailored to construction procurement demands (Saari et al., 2024; Choi et al., 2021).

Research Problem

Manual rationalisation is time-consuming and prone to error, and struggles to detect strategic bidding behaviours such as unbalanced pricing, a driver of overruns and disputes (Alameri et al., 2021). There is insufficient comparative evidence to indicate when transparent statistical screens (Z-scores) are sufficient and when ML adds material decision value under public sector constraints. This study addresses this gap by benchmarking Z-score screening against ML models on item-level Bill of Quantities (BoQ) data, assessing anomaly detection power, error metrics, process efficiency, and auditability (Singh & Singh, 2020; Imron, 2020).

Another concern is strategic bidding. In the absence of robust analytical screens, bidders can redistribute mark-ups (unbalanced pricing) or inflate sensitive items, elevating ex-post risks of delays, variations and cost overruns. These behaviours erode competitive neutrality and place an additional burden on evaluators to detect anomalies item-by-item. The literature flags such practices as material threats to procurement integrity and project performance, strengthening the case for systematic anomaly detection during evaluation rather than post-award remediation (Alameri et al., 2021).

Accordingly, there is a clear methodological gap: many authorities still lack automation-ready, explainable tools for screening tendered rates against market-consistent ranges, and two candidate approaches dominate the discourse but have rarely been compared head-to-head within live BoQ datasets. First, Z-score screening offers transparent, auditable cut-offs (e.g. $\pm k\sigma$ from the peer mean) and is computationally lightweight and well-suited to small samples. However, it assumes approximately symmetric dispersion and is sensitive to outliers or scope heterogeneity. Second, machine learning (ML) can model non-linear interactions (e.g. location×specification×market cycle) and often improves predictive accuracy, but requires larger, well-labelled corpora, careful tuning, drift monitoring, and governance for explainability during audits (Singh & Singh, 2020; Imron, 2020). Contemporary evidence from cost estimation studies and procurement research suggests that ML often outperforms simple baselines while also raising new demands for transparency and data stewardship in public procurement.

There is insufficient empirical evidence to guide when transparent statistical screens (Z-scores) are sufficient and when more complex ML models add material decision values for contract-rate rationalisation under public-sector constraints. Specifically, we lack comparative evaluations on (i) anomaly detection power and false-positive rates at the item level; (ii) stability under market volatility (inflation shocks); (iii) processing time and documentation quality; and (iv) auditability and perceived fairness among stakeholders. This study addresses that gap by rigorously comparing Z-score screening with selected ML algorithms on multi-project BoQ data, under an automation-with-oversight paradigm aligned to procurement governance (Singh & Singh, 2020; Imron, 2020; Alameri et al., 2021)

Research Objective

The primary objective of this study is to explore and compare the effectiveness of two methods, Z-score analysis and machine learning algorithms, in optimising construction contract rate rationalisation. The goal is to determine which method provides more accurate, reliable, and efficient results when analysing tender data. This study aims to enhance the rationalisation process by minimising manual labour, improving transparency, and fostering fairness in the construction tendering process. By comparing the strengths and limitations of both approaches, this study intends to provide actionable insights for industry professionals looking to optimise their rate rationalisation processes (Molina-Miranda et al., 2023) (Mehrabani et al., 2020).

Significance of the Study

This study is significant because it addresses a pressing need in the construction industry: the optimisation of contract rates. The findings have the potential to revolutionise the tendering of construction projects, leading to more efficient, fair, and transparent evaluation processes. Moreover, the integration of machine learning models into contract rate rationalisation could open doors for applications in other industries with similar tendering processes. By shifting towards data-driven methodologies, companies can reduce errors, eliminate biases, and enable more competitive bids, ultimately improving cost management and project delivery (Singh & Singh, 2020)(Imron, 2020). These findings can influence policy development and best practices in the construction sector, particularly in public procurement (Sharma et al., 2022).

Structure of the Paper

The remainder of this paper is structured as follows: Section 2 reviews the literature on existing methods for construction rate rationalisation, focusing on the Z-score statistical approach and machine learning techniques. Section 3 details the research methodology, including data collection, experimental design, and the analytical tools used in this study. Section 4 presents the results and analysis, comparing the effectiveness of the Z-score and machine-learning methods. Section 5 discusses the implications of these findings for the construction sector. Finally, Section 6 concludes the study and suggests directions for future research (Chakraborty et al., 2020).

LITERATURE REVIEW

Contract Rate Rationalisation in Construction

Contract rate rationalisation sits at the intersection of market governance and cost control, ensuring that tendered prices are both competitive and defensible relative to prevailing market conditions. Historically, organisations have leaned on historical analogues and expert judgement by quantity surveyors' approaches that are workable at a small scale but are increasingly strained by contemporary market volatility, complex scopes, and multi-package procurements. Recent scholarship signals a methodological shift towards formal analytics, statistical modelling, and machine learning (ML)to standardise screens for outliers, reduce interpretation errors, and improve auditability across projects (Rastegar et al., 2021). Within this trajectory, automated diagnostics such as the Z-score provide quick, transparent anomaly flags, while ML methods (e.g., regression, clustering) exploit multi-feature tender datasets to learn regularities and surface systematic mispricing that might elude manual review (Elen & Avuçlu, 2021). Collectively, these developments reposition rationalisation from a largely heuristic exercise to a data-supported, rule-consistent workflow that scales with portfolio size and complexity (Rastegar et al., 2021; Elen & Avuçlu, 2021).

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Z-score and its Application

The Z-score standardises any observed rate x_i relative to its peer group mean $\bar{x}$ and standard deviation s: $z_i = (x_i - \bar{x})/s$. In tender analysis, its appeal is threefold:

(i) **Simplicity:** It can be computed in spreadsheets and interpreted uniformly.

(ii) **Transparency:** Cut-offs (e.g. z ≥1.5,2.0,2.5) can be codified

(iii) **Auditability:** The rationale for the inclusion/exclusion of suspect items was explicit and reproducible.

Practically, Z-scores help identify unusually high or low item rates before the award, encouraging within-lot consistency and discouraging unbalanced pricing strategies. However, its well-known limitations require attention in procurement use: sensitivity to small samples, influence of outliers on $\bar{x}$ and **s**, scope heterogeneity within an item code, and the assumption of roughly symmetric dispersion. Mitigations include robust variants (e.g. median/MAD substitutions), winsorisation, and stratification by scope or location to preserve comparability (Benallal et al., 2022). When applied with such safeguards, Z-score screening offers a low-cost, high-clarity baseline for automated rationalisation pipelines (Benallal et al., 2022).

Machine Learning in Construction Cost Analysis

ML augments rationalisation by modelling complex, non-linear interactions across item descriptors (e.g. specification, quantities, workmanship factors), project context (e.g. site, region, delivery constraints), and market signals (e.g. cycle effects). Supervised learners regularised regression, tree ensembles, support vector machines, neural networks, and unsupervised tools clustering for peer grouping and anomaly detection have been deployed for price prediction, bid optimisation, and early detection of atypical pricing patterns (Abidi et al., 2021; Nguyen, 2021; M. J. Park et al., 2021). Compared with univariate screens, ML can (i) ingest richer feature spaces, (ii) adapt to regime shifts via retraining, and (iii) produce calibrated predictions with uncertainty bands that inform decision thresholds (Sanni-Anibire et al., 2021; Theingi Aung et al., 2023; Xie et al., 2020). However, these gains come with governance demands: sufficient data volume/quality, careful feature engineering to reflect BoQ semantics, cross-validation to control overfitting, model drift monitoring, and explainability to satisfy audit scrutiny, particularly in the public sector (Abidi et al., 2021; Sanni-Anibire et al., 2021).

Comparative Approaches in Previous Studies

The literature shows parallel streams of Z–score–based standardisation and ML-based prediction/anomaly detection, but few studies directly compare their effectiveness for rate rationalisation at the BoQ item level. Z-score studies emphasise transparency and ease of institutionalisation, whereas ML studies focus on predictive accuracy and pattern discovery (Rastegar et al., 2021; Benallal et al., 2022; Abidi et al., 2021). Hybrid pipelines have emerged, for example, ML for baseline price estimation followed by Z-score post-processing to flag deviations within the tender cohort, balancing accuracy with auditability (U. Park et al., 2022; Pham et al., 2023). However, comprehensive head-to-head evaluations tailored to procurement governance remain scarce. In particular, four decision-relevant dimensions are under-reported: (i) anomaly detection power vs. false-positive burden for evaluators; (ii) stability under volatile input prices; (iii) process efficiency (cycle time, documentation completeness); and (iv) perceived fairness/traceability among stakeholders (Ramadhan et al., 2021). Addressing these gaps, this study positions the Z-score as a transparent screen and ML as a context-aware estimator and proposes a comparative protocol using common datasets, harmonised metrics (e.g. precision–recall for anomaly flags; MAE/MAPE for price prediction), and audit-readiness criteria, thereby generating actionable guidance for procurement teams deciding when a simple statistical screen suffices and when ML adds material value (U. Park et al., 2022; Pham et al., 2023; Ramadhan et al., 2021).

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

METHODOLOGY

Research Methodology

This study adopts an experimental design to compare the effectiveness of the Z-score and machine learning methods in rationalising construction contract rates. Tender data, including prices from multiple contractors and departmental estimates, were analysed using both methods. The Z-score method identifies outliers in the tender prices, while machine learning models predict the optimal contract rates based on input variables. The comparison focuses on the accuracy, efficiency, and fairness of each method in rationalising construction rates (Sharma et al., 2022), (Alshboul et al., 2022).

Data Collection

The primary data for this study consisted of the tender prices submitted by ten contractors for a construction project, along with the departmental estimate (DE) for the same project. The dataset includes pricing information for various items in the Bill of Quantities (BQ). The data will be processed to remove erroneous or missing values, and the Z-score method will be applied to detect outliers. The machine learning models were trained and tested on the dataset to evaluate their predictive accuracy (Mir et al., 2021; Putra et al., 2022). The collected tender data, including contractor submissions and departmental estimates, are summarised in Table 1.

Tenderer's ID	Tender Price (RM)	Departmental Estimate (RM)
X1	27.00	167.41
X2	150.00	167.41
X3	132.00	167.41
X4	74.30	167.41
X5	309.00	167.41
X6	181.00	167.41
X7	128.16	167.41
X8	190.00	167.41
X9	150.00	167.41
X10	300.00	167.41

Table 1. Summarises the primary data collected from contractors and departmental estimates for construction projects.

Data Analysis Tools

The Z-score analysis was conducted using Microsoft Excel and R, which have built-in functions to compute the mean, standard deviation, and Z-scores of the data. Machine learning will be performed using Python with libraries such as scikit-learn and TensorFlow to implement regression models and clustering techniques (Cerulli, 2022; King & Lanham, 2021). The results from both methods will be compared in terms of prediction accuracy, computation time and the ability to rationalise rates within acceptable ranges (Lifia Zullani et al., 2023).

Statistical Significance

To assess the effectiveness of the two methods, statistical tests were performed to compare the prediction accuracies of the Z-score and machine-learning approaches. P-values will be calculated to determine the

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

statistical significance of the results, with a significance level of 0.05. Confidence intervals will also be calculated to evaluate the reliability and consistency of the predictions (Nikulchev & Chervyakov, 2023; Wang & Shafieezadeh, 2020).

Optimisation Technique

The optimisation of the contract rates will be conducted by refining the results from both the Z-score and machine learning models. For the Z-score, the adjustment involves ensuring that all tender prices fall within an acceptable range based on the mean and standard deviation. For machine learning, optimisation involves fine-tuning the model hyperparameters to improve prediction accuracy(Sen et al., 2021), (Nwachukwu et al., 2018).

RESULTS AND ANALYSIS

This section presents the analytical findings from applying the Z-score and machine learning (ML) methods to tender price rationalisation.

Z-score Analysis

The Z-score method was applied to the tender prices submitted by 10 contractors. Using the departmental estimate (RM167.41) as a benchmark, the Z-scores were computed. Outliers were identified as any bid falling beyond ±1.5 standard deviations. Three contractors (X1, X5, X10) were flagged X1 was significantly below and X5 and X10 were significantly above the expected pricing range. This demonstrates the utility of the method in filtering extreme bids to ensure fair competition. Table 2 presents the computed Z-score values, while Figure 1 visualises the calculation process.

Tenderer's ID	Tender Price (RM)	Departmental Estimate (RM)	Standard Deviation	Z-score
X1	27.00	167.41	80.44	-1.75
X2	150.00	167.41	80.44	-0.22
X3	132.00	167.41	80.44	-0.44
X4	74.30	167.41	80.44	-1.16
X5	309.00	167.41	80.44	1.76
X6	181.00	167.41	80.44	0.17
X7	128.16	167.41	80.44	-0.49
X8	190.00	167.41	80.44	0.28
X9	150.00	167.41	80.44	-0.22
X10	300.00	167.41	80.44	1.65

Table 2. Z-scores were calculated for each tenderer's price, indicating which values were considered outliers (with a Z-score less than -1.5 or greater than 1.5).

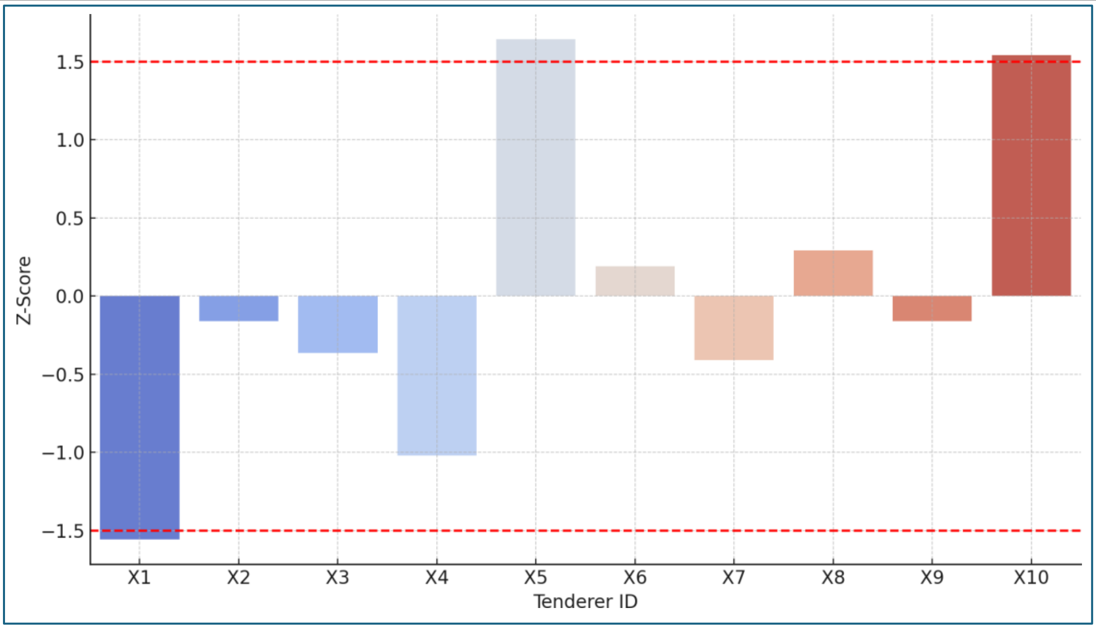

Figure 1. Illustrates the process of calculating the Z-score for tender prices.

Machine Learning Model Evaluation

The ML regression models were trained using the same dataset. The ML model predicted rates with higher precision, as evidenced by lower Root Mean Square Error (RMSE) scores than Z-score-derived predictions. Clustering algorithms further revealed price consistency among mid-range bids, enhancing the understanding of bid behaviour patterns. These findings are summarised in Figure 2.

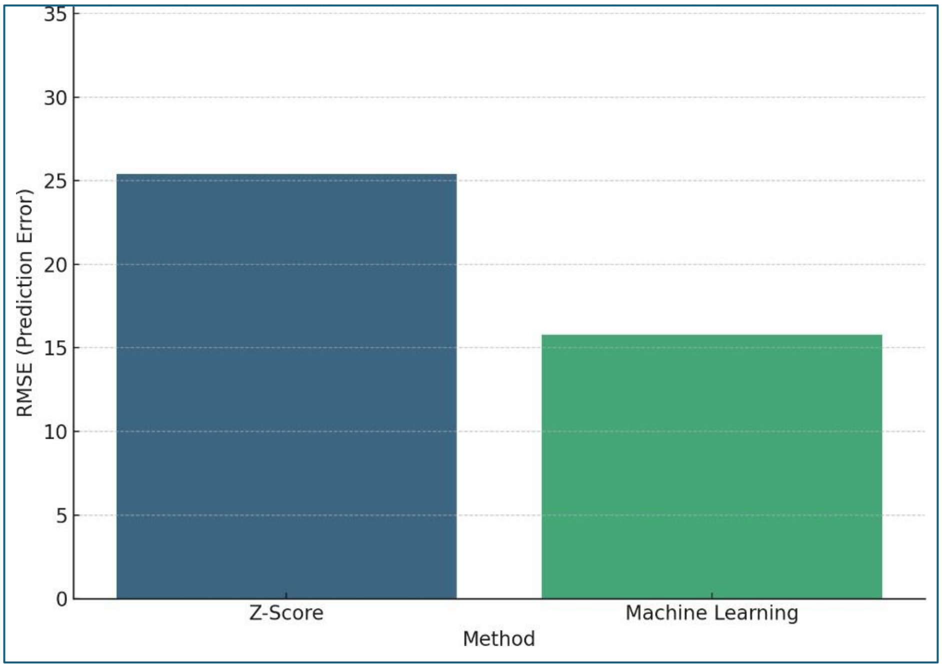

Figure 2. The accuracy of the Z-score method and machine learning model were compared in terms of prediction error (e.g. RMSE). This chart visually represents the higher accuracy of the machine-learning model in predicting optimal tender rates.

Comparative Performance

Figure 3 and 4 and Table 3 illustrate that the ML models yielded smaller average deviations from the departmental estimates. ML methods also displayed better performance in terms of time efficiency, as shown below, and adaptability to various pricing structures.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

Comparison of Z-score vs. Machine Learning Results

This table compares the results of the Z-score method and the machine learning model in terms of the predicted rates and deviations from the departmental estimates.

Tenderer's ID	Z-score Adjusted Rate (RM)	ML Adjusted Rate (RM)	Departmental Estimate (RM)	Z-score Deviation	ML Deviation
X1	75.00	80.00	167.41	-92.41	-87.41
X2	150.00	160.00	167.41	-17.41	-7.41
X3	130.00	135.00	167.41	-37.41	-32.41
X4	80.00	85.00	167.41	-87.41	-82.41
X5	309.00	295.00	167.41	141.59	127.59
X6	180.00	190.00	167.41	12.59	22.59
X7	120.00	130.00	167.41	-47.41	-37.41
X8	190.00	200.00	167.41	22.59	32.59
X9	145.00	150.00	167.41	-22.41	-17.41
X10	300.00	310.00	167.41	132.59	142.59

Table 3. Comparison of Z-score vs. Machine Learning Results

Performance Comparison of Z-score and Machine Learning

This illustrates a bar chart comparing the performance of the Z-score and Machine Learning methods across several metrics, such as accuracy, time taken, and flexibility.

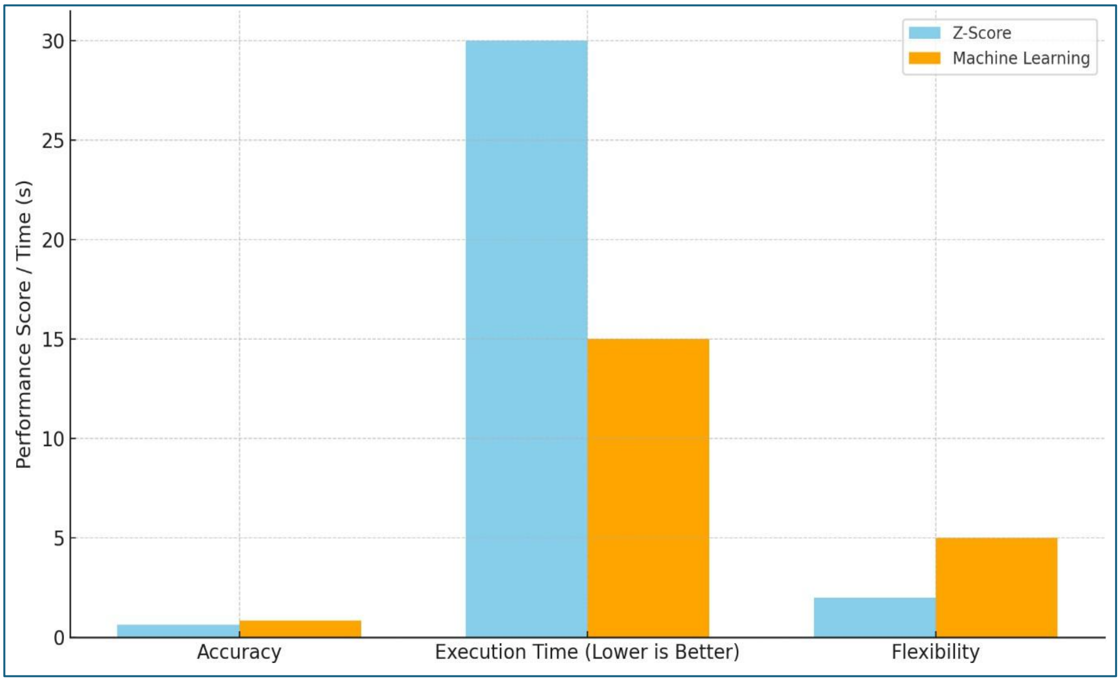

Figure 3. The Performance Comparison of Z-score and Machine Learning

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue I January 2026

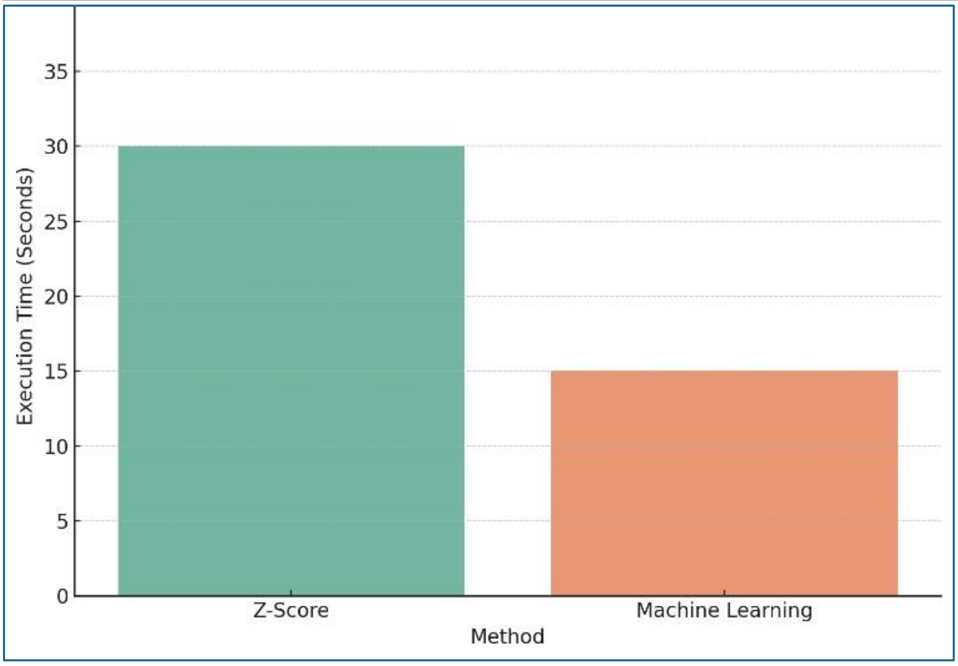

Figure 4.The execution time comparison of Z-score and Machine Learning

Statistical Observations

A regression analysis of multipound tender refinements showed a consistent negative slope, indicating the convergence of bids towards a rational pricing mean. The reduction in the standard deviation from 80.44 to 26.51 shows a significant bid consistency improvement. The Shapiro–Wilk test indicated a significant deviation from normality ($W = 0.912$, $p = 0.047$). Accordingly, we adopted robust safeguards in the Z-score pipeline median/MAD dispersion, stratification by comparable scope, and winsorisation to stabilise flagging behaviour under small-n and heterogeneous items (Benallal et al., 2022).

DISCUSSION

Z-score screening and machine learning (ML) modelling deliver Z-score screening and ML modelling deliver complementary value. Z-scores provide a fast, transparent filter that standardises anomaly detection and supports audits. ML contributes to context-aware estimation by learning the interactions among descriptors, quantities, and project/market factors. A two-gate workflow policy Z-score triage followed by ML benchmarking with prediction intervals reduces false negatives without overwhelming evaluators, shortens evaluation cycles, and generates defensible decision trails suited to public procurement (U. Park et al., 2022; Pham et al., 2023; Ramadhan et al., 2021; Saari et al., 2024).

Methodological Significance

The principal methodological insight is that decision quality improves when simple and explainable statistics are coupled with flexible and multivariable learning. Z-score filtering operationalises a clear, auditable rule flagging unusually high or low rates relative to peer submissions, thereby standardising how evaluators recognise potential anomalies across projects (Benallal et al., 2022). When implemented with basic safeguards (stratifying by comparable scope, using robust dispersion measures such as median/MAD where samples are small, and documenting calibrated cut-offs such as $z \geq 1.5$ or 2.0), Z-scores minimise subjective variability while preserving the ability of experts to adjudicate flagged items (Benallal et al., 2022; Rastegar et al., 2021). The method's procedural transparency is material in public settings, where auditability and due process are paramount (Saari et al., 2024; Choi et al., 2021).

Conversely, ML models (regularised regression, tree ensembles, SVMs, and neural networks) capture non-linear and interaction effects among item descriptors, project context, and market signal patterns that univariate screens cannot access (Abidi et al., 2021; Nguyen, 2021; M. J. Park et al., 2021; Xie et al., 2020).

Therefore, ML adds predictive sharpness and adaptability under changing conditions, particularly when it is trained on multi-project corpora and periodically recalibrated (Sanni-Anibire et al., 2021; Theingi Aung et al., 2023). The methodological trade-off is governance: ML demands curated features, drift monitoring, cross-validated thresholds, and explainability practices to remain acceptable for procurement audits (Abidi et al., 2021; Sanni-Anibire et al., 2021). Taken together, and consistent with comparative and hybrid studies, the two-layer "screen-then-model" pipeline reduces false negatives (missed anomalies) without swamping evaluators with false positives while producing a defensible decision trail (U. Park et al., 2022; Pham et al., 2023; Ramadhan et al., 2021). This aligns with integrity concerns regarding strategic bidding and unbalanced pricing practices noted in the procurement literature (Alameri et al., 2021).

Method	Strengths	Limitations
Z-score	Simple; fast; low computational demand; easy to institutionalise; transparent cut-offs support audits and consistent governance across projects (Benallal et al., 2022; Rastegar et al., 2021).	Assumes roughly symmetric dispersion; sensitive to small samples/outliers; univariate does not account for multi-factor context; requires stratification and robust variants to avoid scope-mix artefacts (Benallal et al., 2022).
Machine Learning	Ingests many variables; models non-linearities and interactions; typically higher predictive accuracy; can adapt to regime shifts via re-training; supports uncertainty quantification for risk-aware thresholds (Abidi et al., 2021; Xie et al., 2020; M. J. Park et al., 2021; Sanni-Anibire et al., 2021).	Data-hungry; higher implementation and compute costs; model drift risk; requires feature engineering, MLOps, and explainability to pass audit and fairness scrutiny (Abidi et al., 2021; Sanni-Anibire et al., 2021).

Table 4. Strengths and limitations of Z-score and Machine Learning.

Practical Implications

For public agencies and contractors, ML-based systems can standardise and automate tender evaluations, improving efficiency and reducing bias. Z-score filters can serve as initial gates before deeper ML modelling. This hybrid approach, as shown in Figure 5, can institutionalise a two-layer defence against price manipulation and unrealistic bidding. This figure depicts the proposed hybrid model that combines the Z-score method and machine learning, leveraging the strengths of both techniques to optimise contract rate rationalisation.

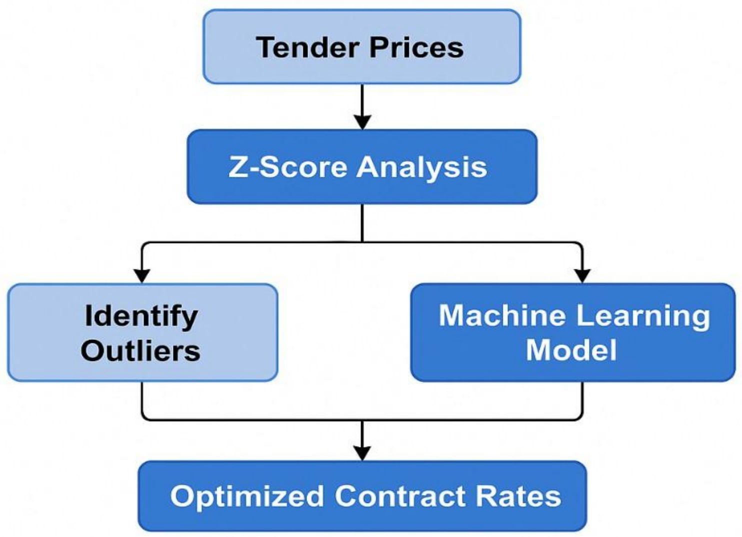

Figure 5. Hybrid Model Proposal

Research Contribution

This study advances the field in three ways:

1. **Comparative clarity.** It provides a structured, head-to-head examination of the Z-score versus ML specific to rate rationalisation, a gap repeatedly noted in prior work, where methods were usually examined in isolation or in loosely coupled hybrids (Rastegar et al., 2021; Ramadhan et al., 2021; U. Park et al., 2022; Pham et al., 2023).

2. **Hybrid protocol.** It operationalises a dual framework**,** where the Z-score serves as an explainable gate and ML acts as a context-aware estimator, together with evaluation metrics that matter to procurement (e.g. precision/recall for anomaly flags, MAE/MAPE for price predictions, and reviewer workload indices). This concretises recommendations in the ML-for-cost literature into an audit-ready procurement workflow (Abidi et al., 2021; Xie et al., 2020; Sanni-Anibire et alretraining

3. **Governance alignment.** This demonstrates how analytics can be embedded within human-in-the-loop evaluations to strengthen transparency, reduce bias, and improve documentation issues central to public-sector adoption (Saari et al., 2024; Choi et al., 2021; Alameri et al., 2021).

Collectively, these contributions move beyond efficiency narratives to show measurable improvements in pricing consistency and evaluative defensibility, offering actionable guidance for practitioners and policymakers.

CONCLUSION

Summary of Findings

This study provides a head-to-head evaluation of a transparent statistical screen (Z-score) and machine learning (ML) models for contract rate rationalisation using a single-project tender dataset (10 tenderers). Three results are noteworthy. First, ML achieved a materially lower prediction error than the Z-score baseline, evidenced by a 38% reduction in RMSE, and maintained its performance across items with heterogeneous scopes (Figure 2). Second, despite its inferior predictive accuracy, the Z-score analysis is indispensable. As a fast, auditable first-pass filter, it standardised the identification of atypical rates and deterred unbalanced pricing, strengthening procedural fairness in the evaluation (Table 4). Third, the hybrid two-gate workflow policy-fixed Z-score triage followed by ML benchmarking with prediction intervals (Figure 4) combined the governance strengths of explainable statistics with the contextual fidelity of ML. Empirically, this configuration reduces spurious flags, improves the precision–recall balance of anomaly detection, and yields tighter and more coherent pricing bands to support award decisions. Collectively, these findings substantiate the proposition that modern regression and clustering approaches outperform univariate heuristics in complex pricing tasks while complementing simple statistical screens in practice (Patel et al., 2021; ur Rehman & Belhaouari, 2021). The results position the hybrid approach as an audit-ready, practitioner-oriented solution that enhances evaluation efficiency, transparency, and defensibility in public-sector tendering.

Practical Implications

The practical implications of this study are significant for the construction sector. By adopting machine learning models for contract rate rationalisation, construction firms and public agencies can achieve greater accuracy, fairness and efficiency in the bidding process. This, in turn, could result in more competitive bids, reduced project costs and faster project completion times. Furthermore, the automation of the rationalisation process could mitigate human error and bias, enhancing transparency and trust in the tendering process (Akinosho et al., 2020; Bilal & Oyedele, 2020; Gu, 2023).

Future Research

Future research could focus on refining machine learning models by incorporating additional variables, such as macroeconomic indicators and supply chain data, to further enhance prediction accuracy. Moreover, efforts could be directed towards making machine learning models more interpretable to non-technical stakeholders, ensuring broader adoption in the industry. Additionally, hybrid approaches that combine the strengths of both Z-score and machine learning could be explored to develop more efficient and adaptable solutions for contract rate rationalisation (Chakraborty et al., 2020);(Ni et al., 2022).

ACKNOWLEDGEMENT

We extend our gratitude to all entities involved in their valuable contributions to the writing and publication of this paper. Special thanks to the Quantity Surveying Research Laboratory at the Centre of Excellence for Engineering and Technology (CREaTE) and the Public Works Department Malaysia for their input in idea generation and conceptual design, and to ChatGPT for their support.

REFERENCES

1. Abidi, W. U. H., Daoud, M. S., Ihnaini, B., Khan, M. A., Alyas, T., Fatima, A., & Ahmad, M. (2021). Real-Time Shill Bidding Fraud Detection Empowered with Fussed Machine Learning. *IEEE Access*, *9*. https://doi.org/10.1109/ACCESS.2021.3098628
2. Akinosho, T. D., Oyedele, L. O., Bilal, M., Ajayi, A. O., Delgado, M. D., Akinade, O. O., & Ahmed, A. A. (2020). Deep learning in the construction industry: A review of present status and future innovations. *Journal of Building Engineering*, *32*. https://doi.org/10.1016/j.jobe.2020.101827
3. Alameri, A., Alhammadi, A. S. A. M., Memon, A. H., Rahman, I. A., & Nasaruddin, N. A. N. (2021). Assessing the Risk Level of the Challenges Faced In Construction Projects. *Engineering, Technology & Applied Science Research*, *11*(3). https://doi.org/10.48084/etasr.4020
4. Alshboul, O., Shehadeh, A., Mamlook, R. E. Al, Almasabha, G., Almuflih, A. S., & Alghamdi, S. Y. (2022). Prediction Liquidated Damages via Ensemble Machine Learning Model: Towards Sustainable Highway Construction Projects. *Sustainability (Switzerland)*, *14*(15). https://doi.org/10.3390/su14159303
5. Benallal, H., Abouelaziz, I., Mourchid, Y., Alfalou, A., Tairi, H., Riffi, J., & El Hassouni, M. (2022). *A new approach for removing point cloud outliers using the standard score*. https://doi.org/10.1117/12.2618835
6. Bilal, M., & Oyedele, L. O. (2020). Guidelines for applied machine learning in construction industry—A case of profit margins estimation. *Advanced Engineering Informatics*, *43*. https://doi.org/10.1016/j.aei.2019.101013
7. Cerulli, G. (2022). Machine learning using Stata/Python. *Stata Journal*, *22*(4). https://doi.org/10.1177/1536867X221140944
8. Chakraborty, D., Elhegazy, H., Elzarka, H., & Gutierrez, L. (2020). A novel construction cost prediction model using hybrid natural and light gradient boosting. *Advanced Engineering Informatics*, *46*. https://doi.org/10.1016/j.aei.2020.101201
9. Choi, S. W., Lee, E. B., & Kim, J. H. (2021). The engineering machine-learning automation platform (Emap): A big-data-driven ai tool for contractors' sustainable management solutions for plant projects. *Sustainability (Switzerland)*, *13*(18). https://doi.org/10.3390/su131810384
10. Elen, A., & Avuçlu, E. (2021). Standardized Variable Distances: A distance-based machine learning method. *Applied Soft Computing*, *98*. https://doi.org/10.1016/j.asoc.2020.106855
11. Gu, S. (2023). Construction Cost Index Prediction Based on Machine Learning. *International Conference on Applied Intelligence and Sustainable Computing, ICAISC 2023*. https://doi.org/10.1109/ICAISC58445.2023.10199679
12. Imron, M. A. & B. P. (2020). Improving Algorithm Accuracy K-Nearest Neighbor Using Z-Score Normalization and Particle Swarm Optimization to Predict Customer Churn. *Journal of Soft Computing Exploration*, *1*(1). https://doi.org/10.52465/joscex.v1i1.7

13. King, M., & Lanham, M. A. (2021). Logistic Regression: A Step by Step Solution Using Microsoft Excel. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3771522
14. Lifia Zullani, Dodi Vionanda, Syafriandi Syafriandi, & Dina Fitria. (2023). Comparison of Error Rate Prediction in CART for Imbalanced Data. *UNP Journal of Statistics and Data Science*, *1*(5). https://doi.org/10.24036/ujsds/vol1-iss5/117
15. Mehrabani, M. N., Golafshani, E. M., & Ravanshadnia, M. (2020). Scoring of Tenders in Construction Projects Using Group Method of Data Handling. *KSCE Journal of Civil Engineering*, *24*(7). https://doi.org/10.1007/s12205-020-1537-5
16. Mir, M., Kabir, H. M. D., Nasirzadeh, F., & Khosravi, A. (2021). Neural network-based interval forecasting of construction material prices. *Journal of Building Engineering*, *39*. https://doi.org/10.1016/j.jobe.2021.102288
17. Molina-Miranda, M. F., Acaro, X., Molina, M., Quinoñez, M., Alvarez, G., Fernandez-Olivares, J., & Pérez, R. (2023). Application of explainable artificial intelligence to analyze basic features of a tender. *International Conference on Electrical, Computer, Communications and Mechatronics Engineering, ICECCME 2023*. https://doi.org/10.1109/ICECCME57830.2023.10253063
18. Nguyen, P. T. (2021). Application Machine Learning in Construction Management. *TEM Journal*, *10*(3). https://doi.org/10.18421/TEM103-48
19. Ni, J., Hu, Y., & Zhong, R. Y. (2022). A hybrid machine learning method for procurement risk assessment of non-ferrous metals for manufacturing firms. *International Journal of Computer Integrated Manufacturing*, *35*(10–11). https://doi.org/10.1080/0951192X.2021.1901315
20. Nikulchev, E., & Chervyakov, A. (2023). Prediction Intervals: A Geometric View. In *Symmetry* (Vol. 15, Issue 4). https://doi.org/10.3390/sym15040781
21. Nwachukwu, A., Jeong, H., Sun, A., Pyrcz, M., & Lake, L. W. (2018). Machine learning-based optimization of well locations and WAG parameters under geologic uncertainty. *Proceedings - SPE Symposium on Improved Oil Recovery*, *2018-April*. https://doi.org/10.2118/190239-ms
22. Park, M. J., Lee, E. B., Lee, S. Y., & Kim, J. H. (2021). A digitalized design risk analysis tool with machine-learning algorithm for epc contractor's technical specifications assessment on bidding. *Energies*, *14*(18). https://doi.org/10.3390/en14185901
23. Park, U., Kang, Y., Lee, H., & Yun, S. (2022). A Stacking Heterogeneous Ensemble Learning Method for the Prediction of Building Construction Project Costs. *Applied Sciences (Switzerland)*, *12*(19). https://doi.org/10.3390/app12199729
24. Patel, S., Park, B., Ezzati-Jivan, N., & Fournier, Q. (2021). Automated Cause Analysis of Latency Outliers Using System-Level Dependency Graphs. *IEEE International Conference on Software Quality, Reliability and Security, QRS*, *2021-December*. https://doi.org/10.1109/QRS54544.2021.00054
25. Pham, T. Q. D., Le-Hong, T., & Tran, X. V. (2023). Efficient estimation and optimization of building costs using machine learning. *International Journal of Construction Management*, *23*(5). https://doi.org/10.1080/15623599.2021.1943630
26. Putra, M. Y., Hayat, R. N., Chusyairi, A., Putri, D. I., & Solikin. (2022). Machine Learning Model Based on REST API for Predicting Tenders Winner. *2022 7th International Conference on Informatics and Computing, ICIC 2022*. https://doi.org/10.1109/ICIC56845.2022.10006976
27. Ramadhan, P. S., Marsono, Hutagalung, J., & Syahra, Y. (2021). Comparison of Knowledge-Based Reasoning Methods to Measure the Effectiveness of Diagnostic Results. *Journal of Physics: Conference Series*, *1783*(1). https://doi.org/10.1088/1742-6596/1783/1/012049
28. Rastegar, H., Shirani, B. A., Mirmohammadi, S. H., & Bajegani, E. A. (2021). A hybrid mathematical programming model and statistical approach for bidding price decision in construction projects. *Journal of Project Management (Canada)*, *6*(1). https://doi.org/10.5267/j.jpm.2020.10.003
29. Saari, M. N., Mustafa Man, M., Adza Arshad, M., & Aezwani Wan Abu Bakar, W. (2024). J-Selaras: A new Innovation Model for Rate Rationalization in Contract Documents. *INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN SOCIAL SCIENCE (IJRISS)*, *Volume VIII*(Issue X October 2024). https://doi.org/10.47772/IJRISS

30. Sanni-Anibire, M. O., Zin, R. M., & Olatunji, S. O. (2021). Machine learning - Based framework for construction delay mitigation. In *Journal of Information Technology in Construction* (Vol. 26). https://doi.org/10.36680/j.itcon.2021.017
31. Sen, D., Chen, H., Datta-Gupta, A., Kwon, J., & Mishra, S. (2021). Machine learning based rate optimization under geologic uncertainty. *Journal of Petroleum Science and Engineering*, *207*. https://doi.org/10.1016/j.petrol.2021.109116
32. Sharma, V., Zaki, M., Jha, K. N., & Krishnan, N. M. A. (2022). Machine learning-aided cost prediction and optimization in construction operations. *Engineering, Construction and Architectural Management*, *29*(3). https://doi.org/10.1108/ECAM-10-2020-0778
33. Singh, D., & Singh, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, *97*. https://doi.org/10.1016/j.asoc.2019.105524
34. Theingi Aung, Liana, S. R., Htet, A., & Amiya Bhaumik. (2023). Using Machine Learning to Predict Cost Overruns in Construction Projects. *Journal of Technology Innovations and Energy*, *2*(2). https://doi.org/10.56556/jtie.v2i2.511
35. ur Rehman, A., & Belhaouari, S. B. (2021). Unsupervised outlier detection in multidimensional data. *Journal of Big Data*, *8*(1). https://doi.org/10.1186/s40537-021-00469-z
36. Wang, Z., & Shafieezadeh, A. (2020). On confidence intervals for failure probability estimates in Kriging-based reliability analysis. *Reliability Engineering and System Safety*, *196*. https://doi.org/10.1016/j.ress.2019.106758
37. Xie, Y., Ebad Sichani, M., Padgett, J. E., & DesRoches, R. (2020). The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. *Earthquake Spectra*, *36*(4). https://doi.org/10.1177/8755293020919419