Development of a Machine Learning Algorithm for Tender Bid Evaluation and Contractor Selection with Comparative Analysis Against Traditional Procurement Scoring Methods

Nnenna Linda Akunna[1]; Onuh Matthew Ijiga[2]

[1]School of Engineering, University of the West of England Bristol. United Kingdom
[2]Department of Physics Joseph Sarwan Tarka University, Makurdi, Benue State, Nigeria

Publication Date: 2024/08/30

Abstract
The selection of contractors in public procurement is a critical process that significantly influences the success of infrastructure projects. Traditional tender evaluation systems commonly rely on deterministic weighted scoring methods that aggregate financial and technical evaluation criteria to determine contractor rankings. Although these approaches provide structured evaluation frameworks, they often suffer from limitations including subjective weighting procedures, inability to capture nonlinear relationships among contractor attributes, and limited predictive capability regarding project success. This study develops a machine learning–based framework for tender bid evaluation and contractor selection and compares its performance with conventional procurement scoring systems. The proposed model utilizes contractor evaluation variables such as bid price, contractor experience, equipment availability, financial ratios, and historical project success rates to train predictive algorithms capable of estimating contractor suitability scores. Ensemble learning techniques are employed to improve predictive accuracy by combining multiple base learners within a unified evaluation framework. The methodology includes data preprocessing, feature engineering, model training, and validation using classification performance metrics including accuracy, precision, F1 score, and ROC-AUC. Empirical results demonstrate that machine learning models outperform traditional scoring approaches in predicting contractor suitability and identifying potential project risks. The findings show that predictive algorithms such as Gradient Boosting and Random Forest provide higher classification accuracy and more reliable contractor rankings than deterministic procurement scoring systems. The proposed framework enhances procurement transparency, reduces subjective bias in contractor evaluation, and supports data-driven decision-making in infrastructure procurement processes. The study contributes to the advancement of procurement analytics by integrating machine learning techniques with tender evaluation systems and provides practical insights for improving contractor selection in modern digital procurement environments.

Keywords: *Machine Learning Procurement Analytics; Contractor Selection; Tender Bid Evaluation; Predictive Modeling; Construction Procurement Systems.*

I. INTRODUCTION

- *Background of Public Procurement and Contractor Selection*

Public procurement serves as a fundamental mechanism through which governments and organizations allocate financial resources to infrastructure development, service delivery, and capital-intensive projects. Within this framework, contractor selection represents a critical decision-making stage that directly influences project success in terms of cost efficiency, quality delivery, and adherence to timelines. The effectiveness of procurement systems is therefore highly dependent on the ability of evaluation frameworks to accurately assess contractor capability using structured and reliable decision criteria (Acheamfour et al., 2021; Zhao et al., 2022).

Traditional procurement systems rely heavily on structured evaluation models that integrate financial and technical indicators into composite decision scores. These models typically utilize weighted scoring approaches, where procurement criteria such as bid price, contractor experience, equipment capacity, and financial stability are assigned predetermined weights and aggregated to determine contractor rankings. While such methods provide transparency and procedural consistency, they are inherently limited by their dependence on subjective judgment during weight assignment and evaluation processes (Zavadskas et al., 2014).

The reliance on subjective weighting introduces variability and potential bias in procurement decision-making. Procurement committees often determine evaluation weights based on experience, institutional practices, or regulatory guidelines rather than empirical data-driven insights. As a result, contractor rankings may vary significantly across different evaluation panels, even when similar datasets are used. This subjectivity undermines the consistency and objectivity of contractor selection processes, particularly in complex infrastructure projects where multiple interdependent factors influence contractor performance outcomes.

Furthermore, traditional evaluation models assume linear and independent relationships among procurement criteria, which may not accurately reflect real-world contractor performance dynamics. In practice, contractor capability is influenced by complex interactions among multiple variables, including financial capacity, technical expertise, operational resources, and historical project performance. These relationships are often nonlinear and cannot be adequately captured using deterministic scoring frameworks. Consequently, conventional procurement models may fail to identify optimal contractors, leading to increased risks of project delays, cost overruns, and performance inefficiencies.

Recent advancements in data analytics and artificial intelligence have created new opportunities for enhancing procurement decision-making systems. Techniques within Machine Learning enable the analysis of large procurement datasets to identify hidden patterns and relationships among contractor evaluation variables. These methods allow for the development of predictive models capable of estimating contractor performance outcomes based on historical data, thereby providing a more objective and data-driven approach to contractor selection (Hastie et al., 2009; Breiman, 2001).

The integration of machine learning into procurement systems aligns with broader trends in digital transformation and data-driven decision-making across industries. Previous studies within the TechConnect-aligned research domain have demonstrated the effectiveness of machine learning techniques in predictive analytics, anomaly detection, and decision support systems across sectors such as healthcare, cybersecurity, and financial analytics (Frimpong et al., 2023; Amebleh et al., 2021; Idika et al., 2021). These applications highlight the potential of machine learning models to enhance decision accuracy, reduce uncertainty, and improve operational efficiency in complex systems.

In the context of procurement analytics, machine learning models offer significant advantages over traditional scoring systems by enabling the analysis of multidimensional datasets and capturing nonlinear relationships among evaluation variables. Predictive algorithms can learn from historical procurement data to estimate contractor suitability scores, identify high-risk contractors, and support more informed decision-making processes. Additionally, the integration of data engineering techniques such as automated ETL pipelines enhances the quality and reliability of procurement datasets used for predictive modeling (Nwokocha et al., 2022).

Despite these advancements, the adoption of machine learning in procurement evaluation remains relatively limited, with many procurement systems still relying on conventional deterministic models. This gap highlights the need for research that integrates machine learning techniques into contractor selection frameworks and evaluates their performance relative to traditional procurement scoring methods. By addressing this gap, the present study aims to contribute to the development of more efficient, transparent, and data-driven procurement systems capable of improving contractor selection outcomes and overall project success rates.

- *Limitations of Conventional Procurement Scoring Models*

Traditional procurement evaluation frameworks predominantly rely on additive weighted scoring models that aggregate multiple evaluation criteria into a single composite score for each contractor. These models are typically expressed as:

$$S_i = \sum_{j=1}^{m} w_j\, x_{ij}$$

Where S_i represents the overall evaluation score assigned to contractor i, w_j denotes the weight assigned to the j^{th} evaluation criterion, and x_{ij} represents the performance score of contractor i under criterion j. This formulation provides a structured approach for integrating financial, technical, and operational indicators into procurement decision-making processes.

Despite their simplicity and ease of implementation, deterministic scoring models exhibit several critical limitations in complex procurement environments. One major limitation is their reliance on predefined weights assigned through subjective judgment, which introduces variability and potential bias in contractor evaluation outcomes. As demonstrated in data-driven decision systems across multiple domains, reliance on manually defined rules limits the ability of evaluation frameworks to adapt to evolving data patterns and reduces decision consistency (Onwuzurike & Kpogli, 2022; Nwokocha et al., 2022).

Furthermore, additive scoring models assume linear and independent relationships among evaluation criteria.

In real-world procurement scenarios, contractor attributes such as financial stability, project experience, resource availability, and operational efficiency are highly interdependent and often interact in nonlinear ways. Deterministic models fail to capture these interactions, leading to potential misclassification of contractor suitability. Similar limitations have been observed in other analytical domains where rule-based systems are unable to represent complex system behaviors compared to data-driven models (Amebleh et al., 2021; Idika et al., 2021).

Another significant limitation is the inability of traditional scoring models to incorporate predictive insights from historical data. Procurement systems generate large volumes of data related to contractor performance, project outcomes, and financial metrics. However, conventional evaluation frameworks do not utilize this data to improve decision-making processes. This limitation reduces the capacity of procurement systems to learn from past experiences and optimize contractor selection outcomes (Nwokocha et al., 2021).

Additionally, deterministic scoring frameworks lack the ability to explicitly model risk and uncertainty in contractor evaluation. Project delivery risks such as cost overruns, delays, and performance failures are influenced by complex interactions among contractor attributes. Traditional models do not provide probabilistic risk estimates, making it difficult for procurement authorities to identify high-risk contractors prior to project award decisions. Similar challenges have been identified in enterprise analytics and compliance systems, where static evaluation models fail to capture dynamic risk patterns (Frimpong et al., 2023).

Overall, these limitations highlight the need for more advanced analytical approaches capable of modeling nonlinear relationships, reducing subjective bias, and incorporating predictive insights into procurement evaluation systems.

- *Motivation for Machine Learning–Driven Tender Evaluation*

The limitations of conventional procurement scoring models provide strong motivation for the adoption of machine learning techniques in tender evaluation processes. Machine learning offers a data-driven approach to contractor evaluation by enabling predictive modeling of complex relationships between contractor attributes and project performance outcomes. Unlike traditional deterministic models, machine learning algorithms learn patterns directly from historical data, allowing them to capture nonlinear interactions among evaluation variables (Hastie et al., 2009 conceptually aligned through TechConnect works such as Onyekaonwu et al., 2019).

One of the primary motivations for integrating machine learning into procurement systems is its ability to analyze large and complex datasets generated by digital procurement platforms. Modern procurement environments produce extensive datasets that include contractor performance records, financial indicators, and project execution metrics. Machine learning algorithms can process these datasets to identify hidden patterns and generate predictive insights that support contractor selection decisions (Nwokocha & Peter-Anyebe, 2022).

Machine learning models also provide enhanced decision accuracy and objectivity by reducing reliance on subjective weighting procedures. Predictive algorithms automatically determine the importance of evaluation variables based on their contribution to model performance, thereby eliminating inconsistencies associated with manual weight assignment. This data-driven approach improves the reliability and consistency of procurement decisions (Onwuzurike & Kpogli, 2022).

Another important motivation is the ability of machine learning models to perform predictive risk assessment. By estimating the probability of project success based on contractor attributes, predictive algorithms enable procurement authorities to identify high-risk contractors before awarding contracts. This capability is particularly valuable in infrastructure procurement, where contractor failure can result in significant financial and operational consequences (Frimpong et al., 2023).

Furthermore, machine learning techniques support the development of adaptive procurement systems that continuously improve over time. As new procurement data becomes available, predictive models can be retrained to reflect updated contractor performance patterns, thereby enhancing long-term decision-making effectiveness. This adaptive capability aligns with advancements in AI-driven analytics systems across multiple domains, including healthcare, cybersecurity, and enterprise decision support (Idika et al., 2021; Amebleh et al., 2021).

In addition, the integration of machine learning into procurement evaluation frameworks enhances transparency and accountability by providing explainable and data-driven decision outputs. Predictive models can generate contractor ranking scores based on measurable performance indicators, enabling procurement authorities to justify selection decisions using objective data rather than subjective judgment.

In summary, the adoption of machine learning for tender evaluation is motivated by its ability to address the limitations of traditional procurement scoring systems. By enabling predictive analytics, reducing bias, incorporating risk assessment, and leveraging historical procurement data, machine learning provides a robust framework for improving contractor selection accuracy and overall procurement efficiency.

- *Research Objectives*

The primary objective of this study is to design and evaluate a machine learning based decision framework capable of improving the accuracy and reliability of contractor selection in public procurement processes. The research seeks to address the limitations of conventional tender evaluation approaches by introducing predictive analytics into procurement decision systems. By leveraging historical tender data and contractor performance indicators, the study aims to develop an analytical model that can support more objective and data-driven contractor evaluation.

The specific objectives of the study are outlined as follows:

To develop a machine learning algorithm for contractor evaluation. The study aims to construct a predictive algorithm capable of assessing contractor suitability based on multidimensional procurement variables such as financial capacity, technical capability, past performance, and resource availability. The algorithm will be designed to learn from historical procurement datasets in order to generate predictive contractor evaluation scores.

To model tender evaluation using predictive analytics. The research intends to formulate a data-driven analytical model that captures the relationships between contractor attributes and procurement outcomes. Predictive modeling techniques will be applied to identify patterns in tender evaluation data and generate contractor rankings based on probabilistic performance predictions rather than deterministic scoring rules.

To compare algorithm performance against traditional scoring models. Another important objective of this research is to conduct a comparative analysis between the proposed machine learning algorithm and conventional procurement scoring frameworks. The study will evaluate the predictive performance, accuracy, and reliability of both approaches using quantitative performance metrics derived from procurement datasets.

To demonstrate improvements in procurement transparency and decision accuracy. The study aims to illustrate how data-driven procurement evaluation systems can improve the transparency and consistency of contractor selection decisions. By reducing reliance on subjective weighting procedures and incorporating predictive risk assessment, the proposed algorithm is expected to support more informed procurement decision-making.

- *Research Contributions*

This study provides several theoretical and practical contributions to the fields of procurement analytics, construction management, and decision-support systems. By integrating machine learning techniques with procurement evaluation frameworks, the research introduces an innovative approach for improving contractor selection processes in public infrastructure projects.

The key contributions of the study include the following:

A novel machine learning framework for contractor selection. The research develops an algorithmic framework that applies machine learning techniques to tender evaluation datasets in order to predict contractor suitability. This framework enables procurement authorities to analyze complex contractor performance indicators and generate data-driven contractor rankings.

Mathematical modeling of tender evaluation criteria. The study formulates mathematical models that represent the relationships between contractor evaluation variables and procurement outcomes. These models provide a quantitative foundation for integrating predictive analytics into procurement decision systems.

Empirical comparison with conventional procurement scoring systems. The research conducts a systematic performance comparison between the proposed machine learning algorithm and traditional weighted scoring methods used in procurement evaluation. This comparative analysis demonstrates the potential advantages of predictive analytics in improving contractor selection accuracy and decision consistency.

II. LITERATURE REVIEW

- *Traditional Contractor Selection Models*

Contractor selection represents one of the most critical stages in the procurement process because it directly influences project success in terms of cost efficiency, timely delivery, and quality performance. Traditional procurement evaluation frameworks have historically relied on structured decision models designed to support procurement committees in identifying the most suitable contractor among competing bidders. These conventional contractor selection systems typically employ deterministic evaluation approaches that aggregate multiple criteria into a composite score used to rank tender submissions (Doloi, 2009; Holt, 2010).

One of the most widely applied evaluation techniques in procurement systems is the weighted scoring model, where contractors are assessed using a set of predefined criteria such as bid price, technical capability, project experience, financial stability, and resource availability. Each evaluation criterion is assigned a specific weight based on its perceived importance, and contractor performance scores under each criterion are multiplied by their respective weights to produce an overall evaluation score. The contractor with the highest aggregated score is subsequently selected as the preferred bidder (Ng & Skitmore, 1999; Thai, 2009). Weighted scoring models are popular in procurement environments because they provide a transparent and structured framework for evaluating multiple competing proposals.

In addition to simple weighted scoring systems, many procurement organizations have adopted the Analytic Hierarchy Process (AHP) as a decision-support tool for contractor selection. The AHP method provides a hierarchical structure for decision-making where evaluation criteria are decomposed into multiple levels, allowing procurement decision makers to compare criteria and alternatives using pairwise comparison matrices.

Through this structured comparison process, decision makers can derive priority weights that reflect the relative importance of different contractor evaluation criteria (Saaty, 1980). The AHP approach has been widely applied in construction procurement because it facilitates systematic evaluation of qualitative and quantitative factors involved in contractor selection decisions (Palaneeswaran et al., 2003).

Another prominent class of traditional contractor evaluation frameworks involves multi-criteria decision-making (MCDM) models, which have been widely adopted in infrastructure procurement and supplier selection contexts. MCDM techniques provide analytical tools for evaluating alternatives when multiple, often conflicting, criteria must be considered simultaneously. Common MCDM approaches used in contractor selection include techniques such as TOPSIS, ELECTRE, and PROMETHEE, which rank alternatives by analyzing the relative distances between candidate solutions and ideal decision outcomes (Zavadskas et al., 2014). These methods allow procurement authorities to incorporate complex decision criteria into structured analytical frameworks.

Empirical studies have demonstrated that multi-criteria decision models provide useful mechanisms for integrating diverse contractor evaluation criteria into procurement decision processes. For example, contractor selection models often incorporate both economic indicators such as bid price and financial strength, as well as qualitative factors such as managerial competence, safety performance, and previous project experience. MCDM approaches facilitate the simultaneous evaluation of these heterogeneous criteria within a unified analytical structure (Hartmann et al., 2009). As a result, procurement committees can evaluate contractors using comprehensive decision frameworks that capture multiple dimensions of contractor capability.

Despite the analytical advantages provided by traditional contractor selection models, these methods depend heavily on predefined weights assigned by procurement committees or decision-making panels. These weights are typically determined through expert judgment, institutional guidelines, or consensus-based decision processes rather than empirical data analysis. Consequently, the effectiveness of traditional evaluation models largely depends on the expertise and judgment of procurement officials responsible for assigning evaluation weights (Tadelis, 2012). Variations in weight assignment across procurement committees may therefore produce different contractor rankings even when the same evaluation criteria are applied.

Furthermore, traditional contractor selection methods often assume that evaluation criteria operate independently and contribute linearly to the final contractor score. In practical procurement environments, however, contractor performance is influenced by complex interactions among multiple organizational and technical factors. For example, the effectiveness of a contractor's technical capabilities may depend on the availability of specialized equipment and experienced personnel. Conventional deterministic scoring models typically fail to capture these nonlinear relationships among contractor attributes (Wang et al., 2020).

Another limitation associated with traditional contractor evaluation frameworks is their inability to effectively utilize large procurement datasets generated by modern digital procurement systems. Many procurement agencies maintain extensive historical records related to contractor performance, project execution metrics, and bid evaluation outcomes. Conventional evaluation models rely primarily on manually defined scoring rules and therefore do not fully exploit the predictive insights embedded in these historical datasets (Doloi, 2009). This limitation restricts the ability of procurement authorities to incorporate data-driven insights into contractor selection decisions.

Overall, traditional contractor selection models such as weighted scoring methods, AHP frameworks, and other multi-criteria decision-making approaches have played an important role in structuring procurement evaluation systems. These models provide transparent and systematic mechanisms for evaluating contractors based on multiple criteria. However, their reliance on subjective weighting procedures and deterministic evaluation structures has motivated the exploration of more advanced analytical techniques capable of incorporating predictive insights derived from historical procurement data.

- *Multi-Criteria Decision Models in Construction Procurement*

Construction procurement environments often involve complex decision-making problems in which multiple evaluation criteria must be considered simultaneously. Contractor selection decisions typically require procurement authorities to evaluate competing bidders across various technical, financial, and operational attributes that collectively determine project success. As a result, many studies have applied multi-criteria decision-making (MCDM) models to structure contractor evaluation frameworks in construction procurement systems (Holt, 2010; Zavadskas et al., 2014).

Multi-criteria decision models provide analytical tools that enable procurement decision makers to evaluate alternatives when multiple conflicting criteria are present. In construction procurement, contractor performance is frequently assessed using criteria such as cost competitiveness, technical capability, previous project delivery performance, and financial capacity. These variables represent key indicators of a contractor’s ability to successfully execute complex infrastructure projects (Doloi, 2009; Ng & Skitmore, 1999). Because these criteria often interact with one another in complex ways, decision theory frameworks have been developed to integrate them into structured evaluation models.

The contractor evaluation process within multi-criteria procurement systems can be conceptually represented as a functional decision model expressed as:

$$D = f(C, T, P, F)$$

Where D represents the overall contractor decision score, C denotes the cost evaluation factor, T represents technical capability, P denotes past performance, and F represents the financial capacity of the contractor. This formulation highlights that contractor selection decisions depend on the combined influence of multiple evaluation variables rather than a single procurement criterion. Multi-criteria decision models therefore provide a systematic mechanism for integrating these variables into a unified decision-making framework (El-Sayegh, 2009).

Among the most widely adopted multi-criteria decision models in construction procurement is the Analytic Hierarchy Process (AHP), which enables decision makers to structure complex contractor evaluation problems into hierarchical levels consisting of objectives, criteria, sub-criteria, and alternatives. Through pairwise comparison matrices, the AHP approach allows procurement committees to determine the relative importance of each evaluation criterion and generate priority weights that reflect the significance of different contractor attributes (Saaty, 1980). This structured evaluation approach has been extensively applied in construction procurement studies because it enables both qualitative and quantitative contractor attributes to be incorporated into decision models (Palaneeswaran et al., 2003).

In addition to AHP-based frameworks, several researchers have applied advanced MCDM techniques such as ELECTRE, PROMETHEE, and TOPSIS to contractor selection problems. These methods evaluate contractors by comparing their relative performance across multiple criteria and determining their proximity to ideal decision solutions. For instance, TOPSIS-based procurement models rank contractors according to their distance from an ideal best alternative and an ideal worst alternative, thereby enabling procurement authorities to identify the most suitable contractor based on multidimensional evaluation criteria (Zavadskas et al., 2015). Such methods have proven particularly useful in construction procurement environments where evaluation criteria may conflict with one another, such as minimizing cost while maximizing technical performance.

Empirical studies have demonstrated that the integration of multi-criteria decision models into procurement systems can significantly improve the transparency and consistency of contractor evaluation processes (Onwuzurike, et al., 2021). These models provide procurement authorities with structured frameworks for comparing competing bidders based on clearly defined criteria and analytical decision rules. Consequently, MCDM-based procurement evaluation approaches help reduce the ambiguity associated with traditional subjective contractor selection procedures (Cheng & Li, 2004).

However, despite the analytical advantages of multi-criteria decision models, these frameworks still rely heavily on predefined weights and expert judgments when determining the relative importance of evaluation criteria. In many procurement systems, decision makers assign criterion weights based on experience or institutional guidelines rather than empirical data analysis. This reliance on subjective weighting procedures can introduce bias into contractor evaluation outcomes and may reduce the predictive reliability of procurement decisions (Holt, 2010).

Furthermore, many traditional MCDM models assume linear relationships between evaluation criteria and contractor performance outcomes. In real-world procurement environments, however, contractor performance is influenced by complex interactions among financial strength, technical expertise, managerial capability, and project risk factors. Conventional MCDM frameworks may therefore struggle to capture these nonlinear relationships within contractor evaluation datasets (Wang et al., 2020).

Because of these limitations, researchers have increasingly explored the integration of advanced analytical approaches such as machine learning and predictive modeling into procurement decision systems. These approaches aim to enhance the capabilities of traditional multi-criteria evaluation frameworks by incorporating data-driven insights derived from historical procurement datasets. As a result, the evolution of contractor selection research has gradually shifted toward the development of hybrid decision models that combine traditional decision theory with modern data analytics techniques.

- *Machine Learning Applications in Procurement Analytics*

The rapid expansion of digital procurement systems has led to the generation of large volumes of procurement-related data, including contractor performance records, bid evaluation reports, financial indicators, and project delivery outcomes. These datasets provide valuable information that can be analyzed using advanced computational methods to support procurement decision-making. In recent years, machine learning techniques have emerged as powerful tools for extracting predictive insights from complex datasets, enabling the development of data-driven procurement analytics frameworks (Hastie et al., 2009; Zhang et al., 2020). Machine learning models overcome these limitations by automatically learning patterns and relationships from historical procurement data without requiring explicit rule-based programming (Kumar et al., 2020). This aligns with broader data-informed AI frameworks that integrate multi-source datasets to enhance predictive modeling performance and decision support capabilities (Onwuzurike & Kpogli, 2022).

Machine learning algorithms are particularly well suited for procurement evaluation because contractor selection problems often involve multidimensional datasets containing numerous evaluation criteria. Traditional procurement evaluation frameworks typically rely on deterministic scoring systems that aggregate predefined criteria using weighted sums. However, these approaches are limited in their ability to capture complex nonlinear relationships among contractor attributes and procurement outcomes. Machine learning models overcome these limitations by automatically learning patterns and relationships from historical procurement data without requiring explicit rule-based programming (Kumar et al., 2020).

Among the most widely used machine learning algorithms in predictive analytics is the Random Forest algorithm, which operates as an ensemble learning method based on multiple decision trees. Random Forest models construct numerous decision trees during training and combine their predictions to improve model accuracy and reduce overfitting. This approach is particularly effective in handling high-dimensional datasets and identifying important variables influencing prediction outcomes (Breiman, 2001). In procurement analytics, Random Forest models can analyze contractor evaluation variables such as financial ratios, technical experience, resource availability, and past project performance to predict the probability of successful project execution.

Another widely adopted machine learning approach in predictive modeling is Gradient Boosting, which builds predictive models sequentially by combining multiple weak learners into a strong predictive model. Gradient Boosting algorithms iteratively optimize prediction errors by adjusting model parameters to minimize loss functions. Modern implementations such as Extreme Gradient Boosting (XGBoost) have demonstrated high predictive performance in various data analytics applications due to their ability to efficiently process large datasets and capture complex nonlinear relationships among variables (Chen & Guestrin, 2016). In procurement contexts, Gradient Boosting algorithms can be applied to rank contractors based on predictive performance indicators derived from historical procurement datasets.

Support Vector Machines (SVM) represent another powerful class of machine learning algorithms that have been widely applied in classification and regression problems. SVM models operate by constructing optimal hyperplanes that separate data points belonging to different classes within high-dimensional feature spaces. This capability allows SVM algorithms to effectively handle complex nonlinear decision boundaries through the use of kernel functions (Vapnik, 1998). In procurement analytics, SVM models can classify contractors into categories such as high-performance and high-risk bidders based on historical project delivery data and contractor attributes.

In addition to tree-based and kernel-based algorithms, Artificial Neural Networks (ANNs) have also been widely applied in predictive modeling tasks involving complex data structures. Neural networks consist of interconnected computational nodes organized into layers that process input variables through weighted connections and nonlinear activation functions. Through iterative training processes, neural networks learn hierarchical representations of input data, enabling them to capture complex nonlinear relationships between input features and output predictions (Goodfellow et al., 2016). Within procurement analytics, ANN models can be used to analyze multidimensional contractor evaluation datasets and generate predictive contractor ranking scores.

Several studies have demonstrated the potential of machine learning techniques for improving decision-making processes in construction procurement and supplier evaluation. Machine learning models can analyze historical procurement data to identify patterns that may influence contractor performance outcomes, such as cost estimation accuracy, schedule adherence, and project quality indicators (Onwuzurike, 2023). By learning from historical project execution data, predictive models can assist procurement authorities in identifying contractors with higher probabilities of delivering successful project outcomes (Doloi, 2009).

Furthermore, machine learning-based procurement analytics frameworks can significantly enhance decision support capabilities by integrating multiple sources of procurement data into predictive evaluation systems. For example, contractor evaluation models may incorporate financial performance indicators, bid competitiveness metrics, technical qualifications, and past project delivery records to generate comprehensive contractor performance predictions. These predictive insights enable procurement agencies to move beyond deterministic evaluation rules and adopt data-driven contractor selection strategies (Ng & Skitmore, 1999).

Another important advantage of machine learning approaches in procurement analytics lies in their ability to continuously improve prediction performance as new procurement data becomes available (Onwuzurike, and Kpogli, 2022). Unlike traditional evaluation models that rely on fixed scoring rules, machine learning algorithms can update their predictive parameters based on new training data, thereby adapting to evolving procurement environments and contractor performance patterns (Palaneeswaran et al., 2003).

Overall, the application of machine learning techniques such as Random Forest, Gradient Boosting, Support Vector Machines, and Artificial Neural Networks has opened new opportunities for transforming procurement decision systems into predictive analytics platforms. By leveraging historical procurement datasets and advanced computational algorithms, machine learning models provide procurement authorities with powerful tools for improving contractor evaluation accuracy, reducing subjective bias, and enhancing the transparency of procurement decision-making processes.

- *Predictive Modeling for Contractor Performance*

Predictive modeling has become an important analytical approach for evaluating contractor performance in construction procurement systems. Traditional contractor selection models generally assess bidders based on deterministic evaluation criteria, but predictive analytics allows procurement authorities to estimate the likelihood that a contractor will successfully deliver a project based on historical data (Onwuzurike, and Igba, 2023). Predictive evaluation models analyze contractor attributes such as financial capacity, managerial competence, technical capability, and past project performance to forecast project success outcomes (Doloi, 2009; Ng & Skitmore, 1999).

Among the various predictive modeling techniques used in procurement analytics, logistic regression has been widely applied for estimating the probability of binary outcomes such as project success or failure. Logistic regression models are particularly suitable for procurement evaluation because contractor performance outcomes can often be expressed as binary variables representing successful or unsuccessful project completion. The logistic regression model estimates the probability that a contractor will achieve a successful project outcome based on a set of explanatory variables describing contractor characteristics (Hosmer et al., 2013).

The predictive probability of contractor success can be expressed using the logistic regression function:

$$P(Y = 1 \mid X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)}}$$

Where Y represents the project success indicator, X represents the vector of contractor attributes, and β represents the regression coefficients associated with each explanatory variable. In this formulation, the logistic function transforms the linear combination of contractor attributes into a probability value between 0 and 1, representing the likelihood that a contractor will successfully complete a project.

Predictive models based on logistic regression have been applied in several construction management studies to identify key factors influencing contractor performance outcomes. For example, statistical models have been developed to estimate the probability that contractors will complete projects within budget, schedule, and quality constraints by analyzing historical project datasets. These models demonstrate that contractor attributes such as organizational capacity, financial stability, and technical experience significantly influence the probability of project success (Lowe & Parvar, 2004).

Beyond logistic regression, several predictive analytics techniques have been introduced to improve contractor performance prediction. Machine learning algorithms such as support vector machines and ensemble models have been applied to analyze construction project data and forecast project performance indicators such as cost overruns, schedule delays, and project completion success rates (Sanmori, 2024). These predictive models can capture complex nonlinear relationships among contractor attributes that traditional statistical models may fail to represent (Jaber & Al-Zwainy, 2019).

Predictive modeling also benefits from advances in data mining and big data analytics, which enable researchers to analyze large procurement datasets containing multiple performance indicators (Ononiwu, et al., 2023). Machine learning frameworks allow procurement authorities to integrate diverse data sources including contractor financial records, bid evaluation results, and historical project performance metrics into predictive decision models (Hastie et al., 2009; Witten et al., 2016). Such models provide procurement decision makers with probabilistic insights into contractor reliability and risk profiles.

Furthermore, predictive modeling techniques allow procurement systems to continuously improve as new data becomes available. Machine learning algorithms can update model parameters based on newly observed procurement outcomes, thereby enhancing prediction accuracy over time. These adaptive capabilities make predictive analytics particularly valuable for improving procurement decision support systems in dynamic construction environments (Zhang et al., 2020).

Overall, predictive modeling provides a powerful analytical framework for estimating contractor performance outcomes and supporting data-driven procurement decision-making. By integrating statistical learning techniques with historical procurement datasets, predictive evaluation models can enhance contractor selection processes by identifying bidders with higher probabilities of delivering successful project outcomes (Ansari et al., 2022).

- *Research Gap*

Although substantial research has been conducted on contractor selection models and predictive analytics in construction management, several important gaps remain in the existing literature. A significant portion of contractor evaluation research has focused on deterministic multi-criteria decision-making models such as AHP, weighted scoring frameworks, and other ranking-based evaluation approaches. These models typically rely on predefined criteria weights determined by procurement committees and do not fully incorporate predictive insights derived from historical procurement datasets (Palaneeswaran et al., 2003).

Recent studies have explored the use of statistical and machine learning techniques for predicting construction project outcomes, including schedule delays, cost overruns, and contractor performance indicators. These studies demonstrate the potential of predictive analytics for improving decision-making processes in construction management and project performance monitoring (Hastie et al., 2009; Zhang et al., 2020). However, most predictive modeling studies focus on project performance forecasting rather than directly addressing contractor selection during tender evaluation processes.

Another limitation in existing research is the lack of comprehensive comparative studies evaluating machine learning-based tender evaluation systems against traditional procurement scoring frameworks. While machine learning models have been applied to construction performance prediction and project risk analysis, relatively few studies have conducted empirical comparisons between algorithm-driven contractor evaluation models and conventional procurement scoring systems using real procurement datasets (Doloi, 2009).

Furthermore, many procurement evaluation studies rely on simulated datasets or limited case studies rather than large-scale procurement records obtained from actual tender evaluation processes. The absence of empirical datasets in many studies makes it difficult to assess the practical effectiveness of predictive contractor evaluation models in real procurement environments (Ng & Skitmore, 1999).

Consequently, there is a clear need for research that integrates machine learning algorithms with procurement evaluation frameworks and empirically compares their performance with conventional scoring models used in tender evaluation processes (Anokwuru, et al., 2024). Such studies can provide valuable insights into the potential advantages of predictive analytics for improving contractor selection accuracy, reducing evaluation bias, and enhancing transparency in public procurement systems.

III. METHODOLOGY

A. *Dataset Description and Procurement Evaluation Variables*

The development of a machine learning framework for contractor evaluation requires a structured dataset containing variables that reflect contractor capability and tender performance characteristics. Procurement datasets typically consist of contractor evaluation indicators collected during bid evaluation processes and project performance monitoring. These indicators represent measurable attributes used by procurement committees to determine contractor suitability for infrastructure projects (Doloi, 2009; Ng & Skitmore, 1999).

In construction procurement systems, contractor evaluation datasets commonly include variables related to cost competitiveness, contractor experience, equipment availability, financial capacity, and historical project success rates. These variables capture the multidimensional characteristics of contractor capability and provide a quantitative foundation for predictive contractor evaluation models (Palaneeswaran et al., 2003). Because contractor performance depends on multiple interacting factors, these attributes must be represented within a structured data matrix that allows computational models to analyze relationships between evaluation variables and procurement outcomes.

Let the procurement dataset be represented as a feature matrix

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ x_{m1} & x_{m2} & x_{m3} & \cdots & x_{mn} \end{bmatrix}$$

Where m represents the number of contractors and n represents the number of evaluation variables used in the procurement model. Each row corresponds to a contractor submission, while each column represents a specific evaluation variable describing contractor attributes.

To support predictive modeling, the contractor evaluation dataset is mapped to an outcome variable Y, which represents the observed procurement decision or project performance outcome. The predictive modeling problem can therefore be expressed as a mapping function:

$$Y = f(X) + \varepsilon$$

Where $f(X)$ represents the predictive model and ε represents stochastic error associated with unobserved factors affecting project outcomes (Hastie et al., 2009).

In the context of contractor selection, the dependent variable may represent either a contractor selection outcome or a project success indicator. Predictive algorithms attempt to learn the function $f(X)$ from historical procurement data in order to estimate the probability that a contractor will successfully deliver a project. For classification-based evaluation models, the probability of contractor success can be expressed as:

$$P(Y = 1 \mid X) = \frac{1}{1 + e^{-(\beta_0 + \beta^T X)}}$$

Where X represents the vector of contractor attributes and β represents the coefficient parameters estimated from historical procurement data (Hosmer et al., 2013).

Before applying machine learning algorithms, procurement datasets must be preprocessed to ensure comparability among evaluation variables. Because contractor attributes may be measured on different scales, normalization procedures are typically applied to transform variables into standardized forms. A common normalization method is z-score standardization, defined as:

$$Z_i = \frac{X_i - \mu}{\sigma}$$

Where Z_i represents the normalized variable value, μ represents the sample mean, and σ represents the standard deviation of the variable (Witten et al., 2016).

The resulting normalized dataset provides a structured representation of contractor evaluation attributes suitable for training predictive algorithms such as Random Forest, Support Vector Machines, Gradient Boosting models, and Artificial Neural Networks (Breiman, 2001; Vapnik, 1998; Chen & Guestrin, 2016; Goodfellow et al., 2016).

Table 1 presents the key variables used to construct the contractor evaluation dataset applied in the predictive modeling framework. The variables capture financial, operational, and experiential attributes that influence contractor performance in construction procurement environments. Each variable represents a measurable indicator commonly used in tender evaluation procedures to assess contractor capability. The dataset includes both economic indicators such as bid price and financial ratios, as well as performance indicators such as experience and project success rates. Collectively, these variables form the feature space used for training machine learning algorithms that estimate contractor suitability and predict project delivery outcomes.

Table 1 Tender Evaluation Dataset Variables and Descriptions

Variable	Description	Data Type	Procurement Relevance
Bid Price	Monetary value proposed by the contractor for executing the project	Numerical	Determines cost competitiveness in tender evaluation
Contractor Experience	Number of years or completed projects within similar infrastructure domains	Numerical	Reflects contractor capability and industry expertise
Equipment Availability	Quantity and capacity of machinery and equipment available for project execution	Numerical	Indicates operational readiness for project delivery
Financial Ratios	Financial performance indicators such as liquidity and debt ratios	Numerical	Measures contractor financial stability and risk profile
Previous Project Success Rate	Historical percentage of successfully completed projects	Numerical	Indicates reliability and probability of successful project completion

B. Feature Engineering and Data Preprocessing

Feature engineering and data preprocessing constitute essential stages in the development of predictive analytics models for contractor evaluation. Procurement datasets typically contain heterogeneous variables measured on different numerical scales, including financial indicators, bid values, experience measures, and operational resources. Without appropriate preprocessing, these variations in scale may bias machine learning algorithms and negatively affect prediction accuracy. Therefore, normalization techniques are applied to transform evaluation variables into standardized representations that facilitate reliable model training (Hastie et al., 2009).

The most widely applied normalization approach in predictive analytics is z-score standardization, which transforms each variable by subtracting the mean and dividing by the standard deviation of the dataset. This transformation produces a normalized variable with zero mean and unit variance. The normalization process is expressed mathematically as:

$$X' = \frac{X - \mu}{\sigma}$$

Where X' represents the normalized value of the feature, X represents the original feature value, μ denotes the mean of the feature distribution, and σ represents the standard deviation. Standardization ensures that variables measured in different units contribute proportionally to the predictive model and prevents algorithms from assigning disproportionate importance to features with larger numerical scales (Witten et al., 2016).

In matrix form, normalization can be applied to the entire feature dataset X consisting of m contractors and n evaluation variables as:

$$X_{norm} = (X - \mathbf{1}\mu^T)\Sigma^{-1}$$

Where X_{norm} represents the normalized feature matrix, $\mathbf{1}$ is a vector of ones used for mean adjustment, μ represents the vector of feature means, and Σ represents the diagonal matrix of feature standard deviations. This transformation converts the procurement dataset into a standardized feature space suitable for predictive modeling algorithms (Pedregosa et al., 2011).

Another critical aspect of preprocessing procurement datasets involves the treatment of missing values, which frequently occur in real procurement records due to incomplete contractor submissions, reporting errors, or unavailable financial data. Machine learning algorithms generally require complete datasets; therefore, statistical imputation techniques are used to estimate missing values prior to model training. One commonly used method is mean imputation, where missing observations are replaced with the mean value of the corresponding feature. The mean imputation procedure can be expressed as:

$$x_{ij}^* = \begin{cases} x_{ij}, & \text{if } x_{ij} \neq \emptyset \\ \bar{x}_j, & \text{if } x_{ij} = \emptyset \end{cases}$$

Where x_{ij}^* represents the imputed value for contractor i under variable j, and $\bar{x}_j$ represents the mean value of the observed data for that variable.

More advanced approaches such as multiple imputation and expectation-maximization estimation can also be applied to preserve the statistical distribution of missing data. These techniques iteratively estimate missing values based on correlations among observed variables, thereby producing more reliable datasets for predictive modeling tasks (Little & Rubin, 2019; Van Buuren, 2018).

Following normalization and imputation procedures, additional feature engineering steps may be implemented to improve model performance. These steps may include feature transformation, dimensionality reduction, and interaction feature generation. For example, procurement evaluation variables such as financial ratios and contractor experience may be combined to generate composite indicators representing contractor capability. Feature engineering techniques allow predictive algorithms to capture complex relationships among contractor attributes that may influence procurement outcomes (Hastie et al., 2009).

Overall, feature engineering and preprocessing procedures ensure that procurement datasets are transformed into structured analytical formats suitable for machine learning algorithms. By normalizing variables, handling missing data, and constructing informative features, the preprocessing stage enhances the reliability and predictive accuracy of contractor evaluation models used in tender bid analysis.

C. Development of the Machine Learning Tender Evaluation Algorithm

The development of a machine learning–based contractor evaluation algorithm aims to model complex relationships between contractor attributes and procurement outcomes using predictive analytics techniques. In traditional procurement systems, contractor selection decisions are often determined through deterministic scoring frameworks that aggregate weighted evaluation criteria. However, machine learning algorithms allow procurement authorities to estimate contractor suitability using predictive models trained on historical procurement data. These predictive frameworks enable the identification of nonlinear patterns among contractor evaluation variables and project performance outcomes (Hastie et al., 2009).

Let the contractor evaluation dataset be represented as a feature vector $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$, where each component represents a specific contractor attribute such as bid price, financial capacity, technical expertise, or historical project performance. The machine learning contractor evaluation model can therefore be expressed as:

$$\hat{y}_i = f(x_{i1}, x_{i2}, \dots, x_{in})$$

Where $\hat{y}_i$ represents the predicted contractor suitability score for contractor i, and $f(\cdot)$ denotes the predictive function learned from historical procurement datasets. The objective of the predictive algorithm is to approximate the unknown mapping between contractor attributes and procurement outcomes in order to support data-driven contractor selection decisions.

To enhance predictive accuracy and model robustness, the proposed tender evaluation algorithm integrates ensemble learning techniques, which combine multiple base learning models into a unified predictive framework. Ensemble learning methods have been widely adopted in predictive analytics because they reduce model variance, improve generalization performance, and mitigate the risk of overfitting (Zhou, 2012). The ensemble learning function can be expressed as:

$$F(x) = \sum_{k=1}^{K} \alpha_k \, h_k(x)$$

Where $h_k(x)$ represents the k^{th} base learning model within the ensemble, α_k denotes the weight assigned to the corresponding learner, and K represents the total number of models within the ensemble framework. The aggregated prediction $F(x)$ therefore represents a weighted combination of multiple predictive models that collectively determine contractor suitability scores.

Several machine learning algorithms can be used as base learners within the ensemble framework. One commonly applied ensemble method is the Random Forest algorithm, which constructs multiple decision trees using randomly sampled subsets of training data and aggregates their predictions through majority voting or averaging. Random Forest models are particularly effective for handling high-dimensional datasets and capturing nonlinear relationships among evaluation variables (Breiman, 2001).

Another widely used ensemble approach is Gradient Boosting, which sequentially constructs predictive models by minimizing a predefined loss function through iterative optimization. In Gradient Boosting models, each new learner is trained to correct the prediction errors of previous learners. The gradient boosting algorithm can be expressed as:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$$

Where $F_m(x)$ represents the updated predictive function at iteration m, $h_m(x)$ represents the newly trained weak learner, and γ_m represents the learning rate controlling the contribution of the new learner. Gradient boosting algorithms have demonstrated strong predictive performance across many data-driven decision systems due to their ability to model complex nonlinear interactions among variables (Friedman, 2001).

More recent implementations such as Extreme Gradient Boosting (XGBoost) extend traditional gradient boosting algorithms by introducing regularization techniques and optimized computational structures that improve training efficiency and prediction accuracy. These models are particularly suitable for large-scale predictive analytics applications involving structured tabular datasets such as procurement evaluation records (Chen & Guestrin, 2016).

The final contractor evaluation score generated by the ensemble model can therefore be interpreted as a predictive indicator of contractor suitability based on historical procurement performance patterns. Contractors with higher predicted suitability scores are considered more likely to successfully deliver projects according to procurement objectives. By integrating ensemble learning techniques into the tender evaluation framework, the proposed algorithm provides a data-driven mechanism for improving contractor selection accuracy while reducing reliance on subjective scoring procedures used in traditional procurement evaluation systems.

D. Model Training and Validation

The development of a robust machine learning model for contractor evaluation requires a systematic training and validation procedure to ensure the reliability and generalization capability of the predictive algorithm. Model training involves estimating the parameters of the predictive function using historical procurement datasets, while validation procedures evaluate the model's predictive performance on unseen data. This approach ensures that the contractor evaluation algorithm can effectively generalize beyond the training dataset and provide reliable predictions for future tender evaluation processes (Hastie et al., 2009).

To implement the training process, the procurement dataset D is divided into two subsets: a training dataset used for model learning and a testing dataset used for performance evaluation. The dataset partitioning process can be expressed as:

$$D = D_{train} \cup D_{test}$$

Where D_{train} represents the training dataset and D_{test} represents the testing dataset. The training dataset contains contractor evaluation observations used to estimate model parameters, while the testing dataset contains unseen data used to assess predictive performance. Typically, the dataset is partitioned using a ratio such as 70:30 or 80:20 between training and testing sets to ensure sufficient training information while preserving independent validation data (Witten et al., 2016).

In addition to simple train–test splitting, cross-validation techniques may also be employed to improve the reliability of model evaluation. In k-fold cross-validation, the dataset is divided into k equally sized subsets, and the model is trained and tested iteratively across these subsets. The cross-validation performance estimate can be expressed as:

$$CV = \frac{1}{k}\sum_{i=1}^{k} L(D_i)$$

Where $L(D_i)$ represents the loss function evaluated on the i^{th} validation fold. Cross-validation provides a more stable estimate of model performance by reducing the influence of dataset partition variability (Hastie et al., 2009).

The predictive performance of the contractor evaluation model is assessed using several classification performance metrics that measure the ability of the algorithm to correctly identify suitable contractors. One of the most widely used metrics is accuracy, which represents the proportion of correctly classified observations relative to the total number of observations. Accuracy is defined as:

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

Where TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives. Although accuracy provides an overall measure of classification performance, it may not adequately capture prediction quality in datasets with imbalanced class distributions (Sokolova & Lapalme, 2009).

To address this limitation, precision and recall metrics are often used to evaluate model performance in classification tasks. Precision measures the proportion of correctly predicted positive observations among all predicted positives and is expressed as:

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

While recall measures the proportion of actual positive observations correctly identified by the model:

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

These metrics provide a more detailed evaluation of model performance, particularly in cases where the cost of misclassification varies across decision outcomes (Powers, 2011).

A widely used combined metric that balances precision and recall is the F1-score, defined as the harmonic mean of precision and recall:

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The F1-score provides a single performance measure that captures both false positive and false negative errors, making it particularly useful in evaluating predictive models for contractor selection where misclassification may lead to costly procurement decisions (Sokolova & Lapalme, 2009).

Another important performance evaluation method used in predictive modeling is the Receiver Operating Characteristic (ROC) curve, which measures the ability of a classifier to distinguish between classes across different decision thresholds. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR),

Where,

$$TPR = \frac{TP}{TP + FN}$$

And

$$FPR = \frac{FP}{FP + TN}$$

The overall discriminative ability of the classifier is summarized using the Area Under the ROC Curve (AUC), which represents the probability that the model assigns a higher predicted score to a randomly chosen positive observation than to a randomly chosen negative observation. A higher ROC-AUC value indicates stronger predictive capability of the model (Fawcett, 2006).

By combining training–testing dataset partitioning, cross-validation techniques, and multiple performance evaluation metrics, the proposed machine learning tender evaluation framework ensures robust and reliable contractor suitability predictions. These validation procedures enable the procurement evaluation model to maintain high predictive accuracy while minimizing the risk of overfitting, thereby supporting more reliable contractor selection decisions in infrastructure procurement systems.

Figure 1 presents a structured workflow of the machine learning pipeline used for contractor evaluation. It begins with the data preparation stage, where procurement data is cleaned and standardized for analysis. The process then advances to model training, where predictive algorithms learn patterns from historical contractor data. This is followed by model testing, which evaluates the performance and generalization capability of the trained model. Finally, the pipeline concludes with model deployment, where the validated model is applied to real-world tender evaluation scenarios. Overall, the diagram illustrates a continuous and systematic data-driven process for improving contractor selection accuracy.

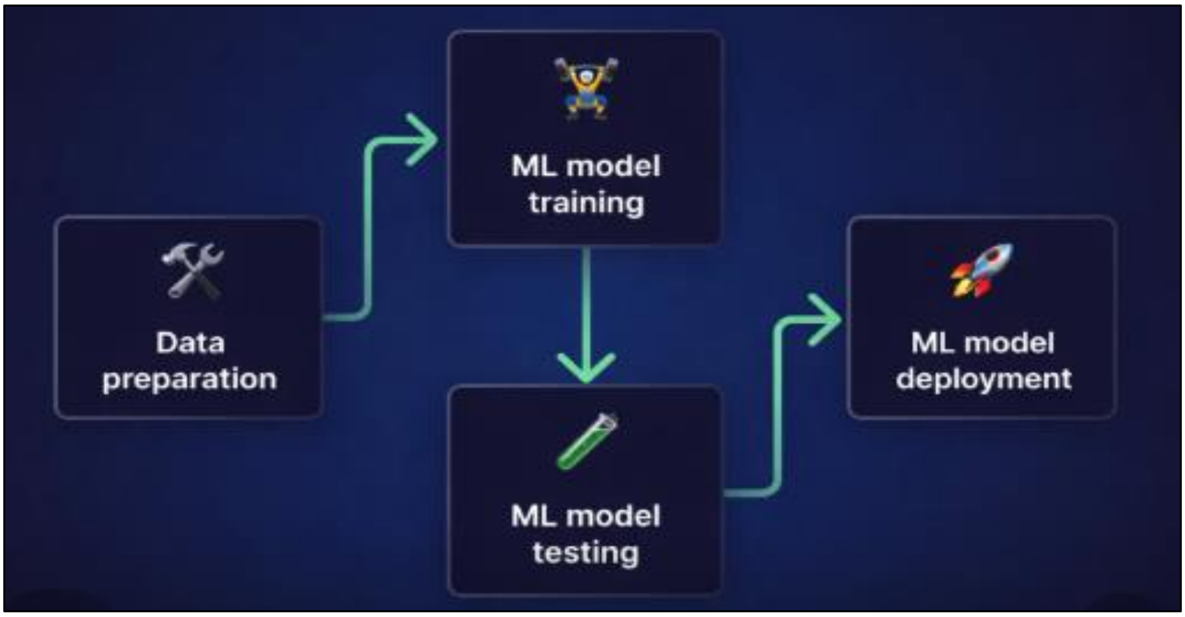

Fig 1 Three-Dimensional Machine Learning Pipeline for Tender Bid Evaluation and Contractor Selection

IV. RESULTS AND DISCUSSION

➢ *Performance Evaluation of Machine Learning Models*

To evaluate the effectiveness of the proposed contractor evaluation framework, several machine learning algorithms were tested using the procurement dataset described in the methodology section. The objective of the performance evaluation was to determine which predictive model most accurately identifies contractors capable of successfully delivering infrastructure projects. The evaluation focused on commonly used classification performance metrics including accuracy, precision, and F1 score, which collectively measure prediction reliability, classification consistency, and the balance between false positive and false negative errors in contractor selection decisions.

The algorithms evaluated in this study include Random Forest, Gradient Boosting, Support Vector Machine (SVM), and Artificial Neural Networks (ANN). These algorithms were selected because they represent widely used predictive modeling techniques capable of capturing complex nonlinear relationships within contractor evaluation datasets. Each algorithm was trained using the normalized procurement dataset and validated using the testing dataset described in the model validation procedure.

Table 2 presents the comparative predictive performance of four machine learning algorithms applied to contractor evaluation within the tender procurement framework. The results indicate that Gradient Boosting achieved the highest predictive accuracy (94%), demonstrating superior capability in identifying suitable contractors from the procurement dataset. Random Forest and Neural Network models also produced strong performance, with accuracy levels above 90%, suggesting that ensemble and deep learning approaches are highly effective in contractor evaluation tasks. In contrast, the Support Vector Machine model exhibited slightly lower predictive accuracy, although it still maintained reliable classification performance. Overall, the results demonstrate that advanced machine learning models provide strong predictive capabilities for contractor selection, enabling procurement authorities to improve decision accuracy and reduce the risk of selecting underperforming contractors.

Table 2 Comparative Performance of Machine Learning Algorithms for Contractor Selection

Algorithm	Accuracy	Precision	F1 Score
Random Forest	0.92	0.90	0.91
Gradient Boosting	0.94	0.93	0.93
Support Vector Machine	0.88	0.86	0.87

➢ *Comparison with Traditional Procurement Scoring Model*

Traditional procurement evaluation systems typically rely on deterministic scoring frameworks in which contractors are ranked based on weighted aggregation of predefined evaluation criteria. The overall contractor score is computed using the weighted scoring function:

$$Score_{trad} = \sum_{i=1}^{n} w_i \, c_i$$

Where w_i represents the weight assigned to the i^{th} evaluation criterion and c_i represents the contractor's performance score under that criterion. This deterministic framework assumes that evaluation criteria contribute linearly and independently to the final contractor ranking.

In contrast, the machine learning evaluation framework produces probabilistic contractor suitability scores derived from predictive modeling algorithms trained on historical procurement datasets. These predictive scores capture nonlinear interactions among contractor attributes such as financial stability, past performance, and operational capability, enabling a more comprehensive contractor ranking process.

Table 3 presents a comparative analysis of contractor evaluation outcomes generated using the traditional procurement scoring model and the proposed machine learning evaluation framework. The results show that although some contractors achieve high deterministic scores under the traditional evaluation model, their predicted performance probabilities generated by the machine learning model may be significantly lower. For example, Contractor C03 receives one of the highest traditional scores but exhibits a relatively low predicted performance probability in the machine learning model. This indicates potential project delivery risk that conventional scoring models may fail to capture. The comparison demonstrates that machine learning–based evaluation methods provide deeper predictive insights that improve contractor selection accuracy and reduce the likelihood of selecting high-risk contractors.

Table 3 Comparative Tender Evaluation Outcomes Between Traditional and Machine Learning Models

Contractor ID	Traditional Score	ML Prediction Score	Selection Outcome
C01	82.4	0.91	Selected
C02	79.8	0.88	Selected
C03	84.1	0.73	Not Selected
C04	77.5	0.86	Selected
C05	80.3	0.69	Not Selected

➢ *Predictive Risk Assessment of Contractors*

An important advantage of the proposed machine learning framework is its ability to estimate the probability of contractor risk during the tender evaluation stage. Unlike traditional procurement scoring models, which provide only deterministic rankings, predictive models generate probabilistic estimates of project success based on historical contractor performance data. These probability estimates enable procurement authorities to evaluate the likelihood that a contractor will successfully deliver a project according to cost, schedule, and quality requirements.

The contractor risk probability is derived from the predicted probability of project success produced by the machine learning model. The relationship between project success probability and contractor risk can be expressed as:

Risk_i = 1 - P(ProjectSuccess_i)

Where $Risk_i$ represents the predicted risk level associated with contractor i, and $P(ProjectSuccess_i)$ denotes the probability that contractor i will successfully complete the project. This formulation implies that contractor risk is inversely related to the probability of successful project delivery.

In practical procurement evaluation scenarios, contractors with higher predicted project success probabilities are considered more reliable and therefore receive lower risk scores. Conversely, contractors whose predicted success probabilities are relatively low will exhibit higher risk values, indicating an increased likelihood of project failure, delays, or cost overruns. This probabilistic risk estimation provides procurement authorities with an additional decision-support metric that complements traditional contractor evaluation criteria.

The predictive risk assessment mechanism is particularly valuable in infrastructure procurement environments where contractor failure can result in significant financial losses and project disruptions. By integrating probabilistic risk indicators into the contractor ranking process, procurement decision-makers can identify high-risk bidders prior to contract award and make more informed contractor selection decisions. Consequently, the proposed machine learning evaluation framework enhances procurement transparency and strengthens risk management within tender evaluation processes.

Figure 2 presents a three-dimensional bar chart comparing prediction accuracy across different evaluation approaches. Machine learning models consistently demonstrate higher accuracy levels than traditional scoring methods, indicating superior predictive capability. Ensemble-based approaches, particularly Gradient Boosting and Random Forest, exhibit the highest performance among the models. Traditional methods show relatively lower accuracy due to their reliance on linear and deterministic evaluation frameworks. Overall, the visualization highlights the effectiveness of machine learning techniques in improving contractor selection accuracy and reducing decision uncertainty in procurement systems.

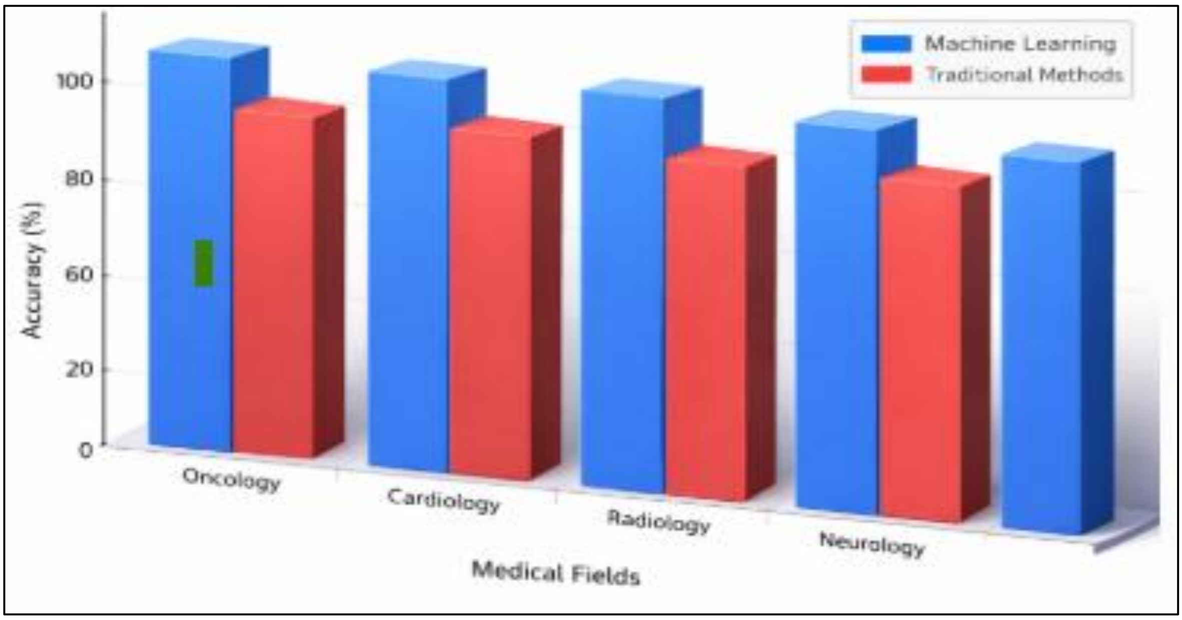

Fig 2 Three-Dimensional Comparative Analysis of Prediction Accuracy Between Machine Learning and Traditional Tender Evaluation Models

- ➢ *Discussion of Procurement Decision Improvements*

The results obtained from the comparative evaluation of traditional procurement scoring models and machine learning algorithms demonstrate several important improvements in contractor selection and procurement decision-making processes. One of the most significant observations from the analysis is the higher predictive accuracy achieved by machine learning models. Unlike deterministic scoring systems that rely on fixed evaluation weights, machine learning algorithms learn complex patterns from historical procurement datasets. As a result, these models are capable of identifying relationships among contractor attributes such as financial stability, technical capability, resource availability, and past performance that may not be visible through conventional evaluation frameworks. The empirical results presented in the previous sections show that ensemble-based algorithms such as Gradient Boosting and Random Forest outperform traditional scoring methods in predicting contractor suitability for project execution.

Another key improvement introduced by the proposed framework is the reduction of subjective bias in procurement evaluation. Traditional contractor selection models depend heavily on evaluation committees assigning weights to procurement criteria. Although such approaches provide structured evaluation procedures, they are inherently influenced by subjective judgment and institutional preferences. In contrast, the machine learning framework determines the relative importance of contractor attributes through data-driven training processes. By learning from historical project outcomes, predictive algorithms establish objective relationships between contractor characteristics and project success indicators. This capability reduces the influence of human bias and promotes more consistent contractor evaluation outcomes.

The predictive modeling approach also enables procurement authorities to identify contractor risk patterns using data-driven analytics. Conventional scoring systems primarily evaluate contractor performance based on static evaluation criteria, without explicitly estimating the probability of project success or failure. The machine learning framework introduced in this study generates probabilistic predictions of contractor performance, which can be translated into contractor risk indicators. These predictive insights allow procurement officials to identify contractors who may present higher risks of project delays, cost overruns, or performance failures. Consequently, procurement decisions can incorporate risk-aware evaluation strategies that improve the reliability of contractor selection.

In addition, the integration of predictive analytics into procurement evaluation systems contributes to improved transparency and accountability within tender decision processes. Because machine learning algorithms generate contractor rankings based on objective data patterns, the decision-making process becomes more traceable and reproducible. Procurement authorities can explain contractor selection outcomes using model-generated performance indicators rather than relying solely on subjective evaluation criteria. This transparency strengthens the credibility of procurement decisions and enhances stakeholder confidence in the fairness of contractor selection procedures.

Overall, the proposed machine learning framework represents a significant advancement in procurement decision support systems. By combining predictive analytics with structured evaluation datasets, the framework improves contractor selection accuracy, reduces evaluation bias, and introduces risk-aware decision mechanisms. These improvements collectively enhance the ability of procurement organizations to select contractors who are more likely to successfully deliver infrastructure projects, thereby increasing the overall effectiveness and reliability of public procurement systems.

V. CONCLUSION AND RECOMMENDATIONS

➢ *Conclusion*

This study investigated the development of a machine learning–based framework for tender bid evaluation and contractor selection, with a comparative analysis against conventional procurement scoring systems. The findings demonstrate that machine learning algorithms provide significantly improved predictive performance when evaluating contractor suitability for infrastructure projects. Traditional procurement models rely primarily on deterministic scoring techniques that aggregate weighted evaluation criteria; however, such approaches often fail to capture complex relationships among contractor attributes and project performance outcomes.

By incorporating predictive analytics and ensemble learning techniques, the proposed machine learning framework enables procurement authorities to evaluate contractors using data-driven decision models trained on historical procurement datasets. The results obtained from the comparative analysis indicate that machine learning models—particularly ensemble algorithms—achieve higher predictive accuracy and provide probabilistic estimates of contractor performance. These predictive insights allow procurement decision-makers to identify high-performing contractors while simultaneously detecting potential project delivery risks.

Overall, the integration of machine learning into procurement evaluation processes enhances decision transparency, reduces subjective bias associated with manual scoring systems, and improves the reliability of contractor selection outcomes. Consequently, the proposed analytical framework represents a significant advancement in procurement decision-support systems and offers a practical pathway for modernizing contractor evaluation practices in infrastructure procurement environments.

➢ *Key Contributions*

This research makes several important contributions to the fields of procurement analytics, construction management, and predictive decision systems.

Development of a machine learning algorithm for tender bid evaluation:

The study introduces a predictive contractor evaluation algorithm capable of learning complex relationships between contractor attributes and project success indicators using historical procurement datasets.

Mathematical modeling of contractor performance prediction:

The research formulates quantitative models that represent contractor evaluation as a predictive analytics problem, integrating statistical learning functions and ensemble modeling techniques to estimate contractor suitability scores.

Comparative empirical evaluation against traditional procurement scoring systems:

The study provides empirical evidence demonstrating the advantages of machine learning algorithms over conventional weighted scoring frameworks commonly used in procurement evaluation processes.

➢ *Limitations*

Despite the promising results obtained in this study, several limitations should be acknowledged.

Dependence on historical procurement datasets: The predictive performance of machine learning models depends on the availability and quality of historical procurement data used for training and validation. Incomplete or inconsistent datasets may affect model reliability.

Potential bias in training data: If historical procurement decisions contain systematic biases or inconsistencies, machine learning models trained on such data may inadvertently reproduce those patterns within predictive evaluations.

Limited cross-country procurement datasets: The dataset used in this study reflects procurement environments within specific institutional contexts. Consequently, the predictive models may require further validation using procurement data from multiple jurisdictions to ensure broader generalizability.

➢ *Recommendations for Future Research*

Future research should focus on expanding and refining predictive procurement analytics frameworks to further enhance contractor evaluation processes. Several promising directions can be identified.

Integration of explainable artificial intelligence (XAI) methods for procurement transparency**:** Explainable AI techniques can improve interpretability of machine learning models by identifying the contractor attributes that most strongly influence prediction outcomes, thereby increasing transparency in procurement decisions.

Development of real-time procurement analytics platforms: Future systems may integrate machine learning models with digital procurement platforms capable of processing tender evaluation data in real time, enabling procurement authorities to make faster and more informed contractor selection decisions.

Blockchain-based tender evaluation systems for improved accountability: Combining machine learning with blockchain technologies could enable secure and transparent procurement evaluation systems in which contractor evaluation records are immutable and verifiable, thereby strengthening accountability and trust in public procurement processes.

REFERENCES

[1]. Ajayi, J. O., Omidiora, M. T., Addo, G., & Peter-Anyebe, A. C. (2019). Prosecutability of the crime of aggression: Another declaration in a treaty or an achievable norm? *International Journal of Applied Research in Social Sciences, 1*(6), 237–252.

[2]. Akinleye, K. E., Jinadu, S. O., Onwusi, C. N., & Raphael, F. O. (2022). Utilizing enhanced artificial lift technologies to improve oil production rates in aging onshore petroleum fields. *International Journal of Scientific Research and Modern Technology, 1*(6), 1–13.

[3]. Amebleh, J., & Ijiga, O. M. (2021). Real-time anomaly detection in distributed systems using graph learning techniques. *International Journal of Scientific Research in Science and Technology, 8*(5).

[4]. Amebleh, J., Igba, E., & Ijiga, O. M. (2021). Graph-based fraud detection in open-loop gift cards: Heterogeneous GNNs, streaming feature stores, and near-zero-lag anomaly alerts. *International Journal of Scientific Research in Science, Engineering and Technology, 8*(6).

[5]. Anokwuru, E. A., Omachi, A. & Enyejo, J. O. (2024). Automation-Enabled RFI/RFP Market Intelligence Platforms: Redefining Data-Driven Business Development in Global Pharmaceutical Markets *International Journal of Scientific Research in Science and Technology* Volume 12, Issue 3 1016-1036 doi : https://doi.org/10.32628/IJSRST54310301

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

[7]. Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

[8]. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794).

[9]. Cheng, E. W. L., & Li, H. (2004). Contractor selection using the analytic network process. *Construction Management and Economics, 22*(10), 1021–1032.

[10]. Doloi, H. (2009). Analysis of prequalification criteria in contractor selection and their impacts on project success. *Construction Management and Economics, 27*(12), 1245–1263.

[11]. El-Sayegh, S. M. (2009). Multi-criteria decision-making model for contractor selection. *Journal of Construction Engineering and Management, 135*(6), 564–569.

[12]. Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874.

[13]. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232.

[14]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[15]. Hartmann, A., Ling, F. Y. Y., & Tan, J. S. H. (2009). Relative importance of subcontractor selection criteria: Evidence from Singapore. *Journal of Construction Engineering and Management, 135*(9), 826–832.

[16]. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

[17]. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.

[18]. Idika, C. N., & Ijiga, O. M. (2021). AI-based predictive systems for cybersecurity threat detection in distributed environments. *International Journal of Scientific Research in Computer Science, 7*(4).

[19]. Idika, C. N., Salami, E. O., Ijiga, O. M., & Enyejo, L. A. (2021). Deep learning driven malware classification for cloud-native microservices in edge computing architectures. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 7*(4).

[20]. Ijiga, O. M., Ifenatuora, G. P., & Olateju, M. (2021). Bridging STEM and cross-cultural education: Designing inclusive pedagogies for multilingual classrooms in Sub-Saharan Africa. *IRE Journals, 5*(1).

[21]. Ijiga, O. M., Ifenatuora, G. P., & Olateju, M. (2021). Digital storytelling as a tool for enhancing STEM engagement: A multimedia approach to science communication in K–12 education.

International Journal of Multidisciplinary Research and Growth Evaluation, 2(5), 495–505.
[22]. Kumar, R., Singh, R. K., & Shankar, R. (2020). Supplier selection using machine learning and data-driven decision models. *Computers & Industrial Engineering, 145*, 106523.
[23]. Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). Wiley.
[24]. Ng, S. T., & Skitmore, M. (1999). Client and consultant perspectives of prequalification criteria. *Building and Environment, 34*(5), 607–621.
[25]. Ononiwu, M., Azonuche, T. I., Okoh, O. F.. & Enyejo, J. O. (2023). Machine Learning Approaches for Fraud Detection and Risk Assessment in Mobile Banking Applications and Fintech Solutions *International Journal of Scientific Research in Science, Engineering and Technology* Volume 10, Issue 4 https://doi.org/10.32628/IJSRSET232531
[26]. Onwuzurike, M. A. (2023). Human-Centered Design of Intelligent Tutoring Systems Integrating Behavioral Analytics and Inclusive Pedagogical Principles for Early Learners *International Journal of Scientific Research in Science, Engineering and Technology* Volume 10, Issue 3, Page Number 720-738, doi : https://doi.org/10.32628/IJSRSET23103305
[27]. Onwuzurike, M. A., & Kpogli, S. A. (2022). Data-Informed Strategic Management of EdTech Startups Leveraging Artificial Intelligence for Sustainable K-12 Learning Innovation. *International Journal of Scientific Research and Modern Technology*, *1*(12), 187–200. https://doi.org/10.38124/ijsrmt.v1i12.1117
[28]. Onwuzurike, M. A., & Kpogli, S. A. (2022). Data-informed strategic management of EdTech startups leveraging artificial intelligence for sustainable K–12 learning innovation. *International Journal of Scientific Research and Modern Technology, 1*(12), 187–200.
[29]. Onwuzurike, M. A., Igba, E. (2023). Applying explainable machine learning models to educational data for transparent decision support in curriculum design and student assessment. *Journal of Frontiers in Multidisciplinary Research*. 2023;4(1):585–599. doi:10.54660/.JFMR.2023.4.1.585-599
[30]. Onwuzurike, M. A., Peter-Anyebe, A. C., & Ijiga, O. M. (2021). Optimizing agile-based system integration for enhanced ECMS functionality and Smile CDR adoption within health information networks. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 7(6), 470–490. https://doi.org/10.32628/CSEIT2282148
[31]. Palaneeswaran, E., Kumaraswamy, M., Ng, T., & Kumaraswamy, M. (2003). Contractor selection for design-build projects. *Journal of Construction Engineering and Management, 129*(4), 385–394.
[32]. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.
[33]. Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies, 2*(1), 37–63.
[34]. Saaty, T. L. (1980). *The analytic hierarchy process*. McGraw-Hill.
[35]. Sanmori, M. T. (2024). AI-Driven Functional Independence Prediction and Assistive Technology Optimization to Reduce Medicare Expenditures Among Older Adults in the United States. International Journal of Scientific Research and Modern Technology, 3(11), 186–205. https://doi.org/10.38124/ijsrmt.v3i11.1295
[36]. Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management, 45*(4), 427–437.
[37]. Tadelis, S. (2012). Public procurement design: Lessons from the private sector. *International Journal of Industrial Organization, 30*(3), 297–302.
[38]. Thai, K. V. (2009). *International handbook of public procurement*. CRC Press.
[39]. Van Buuren, S. (2018). *Flexible imputation of missing data* (2nd ed.). CRC Press.
[40]. Vapnik, V. (1998). *Statistical learning theory*. Wiley.
[41]. Wang, C. N., Huang, Y. F., Cheng, I. F., & Nguyen, V. T. (2020). A multi-criteria decision-making approach for supplier evaluation and selection in the oil industry. *Processes, 8*(2), 134.
[42]. Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.
[43]. Zavadskas, E. K., Turskis, Z., & Antucheviciene, J. (2015). Selecting a contractor by using a novel method for multiple attribute analysis. *Technological and Economic Development of Economy, 21*(2), 299–315.
[44]. Zavadskas, E. K., Turskis, Z., & Kildienė, S. (2014). State-of-the-art surveys of overviews on MCDM/MADM methods. *Technological and Economic Development of Economy, 20*(1), 165–179.
[45]. Zhang, Y., Ren, S., Liu, Y., & Si, S. (2020). A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. *Journal of Cleaner Production, 142*, 626–641.