Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 3, Issue 11, 2025

MACHINE LEARNING-BASED RISK CLASSIFICATION FOR CONSTRUCTION PROJECTS: A COMPARATIVE PERFORMANCE ANALYSIS OF RANDOM FOREST, XGBOOST, AND NEURAL NETWORKS IN CIVIL ENGINEERING

Dawood Khan[1], Abdul Wahab[2], Akram Ullah Khan[3], Muhammad Rafiq Khan[4], Shahzeb[5], Ijaz Ahmad[6], Sulaiman Shah[7]

[1]De Montfort University, Leicester, United Kingdom.
[2]Assistant Professor, Department of Civil Engineering, GIKI, Topi Swabi.
[3]Department of Civil Engineering, University of Engineering and Technology, Peshawar.
[4]Department of Civil Engineering, University of Engineering and Technology, Peshawar.
[5]University of Engineering and Technology, Peshawar, Pakistan.
[6]University of Engineering and Technology, Peshawar, Pakistan.
[7]Department of Infrastructure Engineering, Dalian University of Technology, Dalian, China.
[1]dawoodkhan.mian@gmail.com [2]awahab@giki.edu.pk, [3]17pwciv4882@uetpeshawar.edu.pk,
[4]21bnciv1189@uetpeshawar.edu.pk, [5]16pwciv4578@uetpeshawar.edu.pk,
[6]16pwciv4634@uetpeshawar.edu.pk, [7]sulaimanshah931@gmail.com

DOI:

Keywords:
Construction Risk Prediction, Machine Learning, XGBoost, Random Forest, Neural Networks, Risk Classification, Project Management, Predictive Analytics

Article History
Received on 25 Oct,2025
Accepted on 22 Nov 2025
Published on 24 Nov 2025

Corresponding Author: *

Abstract

This research develops a predictive framework for classifying construction project risk levels using machine learning. The study implements and compares three algorithms Random Forest, XGBoost, and Artificial Neural Networks on a comprehensive dataset of 24 project features including cost, schedule, safety, and environmental metrics. Results demonstrate that XGBoost achieved superior performance with 95.5% accuracy and 0.995 AUC, significantly outperforming other models in classifying challenging Medium-risk projects. Feature importance analysis identified Safety_Risk_Score and Anomaly_Detected as the most critical predictors. The findings provide construction managers with a robust decision-support tool for proactive risk mitigation and establish XGBoost as the optimal algorithm for multi-class risk prediction in construction projects.

Introduction

The construction industry is a cornerstone of the global economy, contributing significantly to GDP, employment, and infrastructure development worldwide. In many countries, it accounts for a substantial share of economic output up to 14.2% of global GDP and is a primary driver of growth, urbanization, and sustainable development initiatives (Cucos & Țurcan, 2025; Kļave et al., 2025; Taiwo et al., 2025). The sector's economic impact is matched by its complexity and the scale of its challenges, including persistent cost overruns, schedule delays, and safety incidents that threaten project success and financial stability (Alsharef et al., 2024; Zaneldin & Ahmed, 2024; Septiaji et al., 2025; Nieto-Morote & Ruz-Vila, 2011). These challenges are exacerbated by the industry's fragmented structure, reliance on a diverse workforce, and the increasing adoption of advanced technologies, which introduce new risks and skill shortages (Cucos & Țurcan, 2025; Alsharef et al., 2024; Kļave et al., 2025). Risk in construction is multifaceted, encompassing financial, temporal, safety, structural, and environmental dimensions. Financial risks include cash flow volatility, cost escalation, and funding uncertainties, while temporal risks manifest as project delays due to supply chain disruptions, labor shortages, or regulatory hurdles (Didar et al., 2025; Zaneldin & Ahmed, 2024; Septiaji et al., 2025). Safety risks remain a critical concern, with construction consistently ranking among the most hazardous industries due to accidents, equipment failures, and inadequate safety protocols (Gao et al., 2023; Nieto-Morote & Ruz-Vila, 2011; Mahajan & Mishra, 2025). Structural risks arise from design errors, material defects, and unforeseen site conditions, while environmental risks are linked to regulatory compliance, ecological impacts, and the growing emphasis on sustainability and circular economy practices (Cucos & Țurcan, 2025; Mahajan & Mishra, 2025; Kļave et al., 2025). The interplay of these risks is further complicated in international and large-scale projects, where political, legal, and cultural factors introduce additional layers of uncertainty (Zhu et al., 2022; Bahamid et al., 2022; Albasyouni et al., 2025). Traditional risk management in construction has relied heavily on reactive, subjective methods such as expert judgment, checklists, and historical data analysis (Albasyouni et al., 2025; Nieto-Morote & Ruz-Vila, 2011; Bahamid et al., 2022; Sithole & Ngwenya, 2025). While these approaches offer practical insights, they often lack the rigor and adaptability needed to address the complexity and interdependencies of modern construction projects (Afzal et al., 2019; Zaneldin & Ahmed, 2024; Sithole & Ngwenya, 2025). The limitations of conventional methods are evident in their inability to process multivariate project data, capture dynamic risk interactions, or provide real-time early warnings for emerging threats (Gao et al., 2023; Wang et al., 2023; Wu et al., 2023). Moreover, the absence of systematic documentation, structured knowledge transfer, and integration of advanced technologies further weakens the effectiveness of risk assessment and mitigation efforts (Albasyouni et al., 2025; Sithole & Ngwenya, 2025; Song & Choi, 2024). As a result, many construction firms especially in developing regions—continue to rely on informal, ad hoc risk management practices, leading to missed opportunities, recurring failures, and suboptimal project outcomes (Albasyouni et al., 2025; Bahamid et al., 2022; Afzal et al., 2019; Nieto-Morote & Ruz-Vila, 2011; Zaneldin & Ahmed, 2024; Didar et al., 2025; Cucos & Țurcan, 2025; Gao et al., 2023; Mahajan & Mishra, 2025; Zhu et al., 2022; Alsharef et al., 2024; Wang et al., 2023; Luo et al., 2025; Sithole & Ngwenya, 2025; Taiwo et al., 2025; Wu et al., 2023; Septiaji et al., 2025; Song & Choi, 2024; Kļave et al., 2025).

The construction industry faces persistent challenges in objectively classifying and managing project risks, often resulting in budget overruns, safety incidents, and project failures. Traditional risk management approaches are largely manual, subjective, and reactive, lacking robust, data-driven, and predictive frameworks that can synthesize diverse project performance indicators to proactively identify high, medium, and low-risk projects. This gap underscores the urgent need for advanced systems that enable early intervention and efficient resource allocation. Conventional risk management in construction is hampered by reliance on expert judgment, fragmented data, and isolated risk analysis, leading to inconsistent and often ineffective risk classification (Chattapadhyay et al., 2021; Khodabakhshian et al., 2023; Khodabakhshian & Cecconi, 2022; Papanikolaou & Xenidis, 2020). Manual processes are time-consuming and fail to capture the complex interdependencies among risk factors, making it difficult to anticipate and mitigate risks before they escalate (Chattapadhyay et al., 2021; Khodabakhshian et al., 2023; Papanikolaou & Xenidis, 2020). Recent research demonstrates that machine learning (ML) methods such as random forests, gradient boosting, support vector machines, Bayesian networks, and ensemble models significantly enhance risk prediction accuracy and objectivity (Alhamami, 2025; Khodabakhshian et al., 2025; Gondia et al., 2023; Poh et al., 2018; Jayakannan, 2025; Khodabakhshian et al., 2023; Turkyilmaz & Polat, 2025; Alsulamy, 2024; Gondia et al., 2020; Fitzsimmons et al., 2022; Gondia et al., 2022; Dikmen et al., 2025). ML models can process large, heterogeneous datasets, identify leading risk indicators, and classify projects by risk level with high precision, enabling proactive interventions that reduce delays, cost overruns, and accidents (Alhamami, 2025; Gondia et al., 2023; Poh et al., 2018; Alsulamy, 2024; Gondia et al., 2020; Fitzsimmons et al., 2022; Gondia et al., 2022). Advanced techniques like data augmentation, feature selection, and hybrid modeling further improve performance, even with limited or unstructured data (Khodabakhshian et al., 2025; Khodabakhshian et al., 2023; Turkyilmaz & Polat, 2025; Khodabakhshian & Cecconi, 2022; Fitzsimmons et al., 2022).

The absence of robust, predictive risk classification frameworks leads to substantial real-world consequences: budget overruns, fatal accidents, project failures, and inefficient resource allocation (Chattapadhyay et al., 2021; Alhamami, 2025; Gondia et al., 2023; Poh et al., 2018; Jayakannan, 2025; Khodabakhshian et al., 2023; Turkyilmaz & Polat, 2025; Zhang & Wang, 2023; Alsulamy, 2024; Papanikolaou & Xenidis, 2020; Gondia et al., 2020; Fitzsimmons et al., 2022; Gondia et al., 2022). ML-driven systems have demonstrated measurable improvements, including earlier risk detection, cost savings, and enhanced safety outcomes (Alhamami, 2025; Gondia et al., 2023; Poh et al., 2018; Jayakannan, 2025; Alsulamy, 2024; Gondia et al., 2020; Fitzsimmons et al., 2022; Gondia et al., 2022; Dikmen et al., 2025).

The integration of ML with process mining, BIM, and IoT is creating comprehensive, real-time risk surveillance ecosystems (Jayakannan, 2025; Khodabakhshian & Cecconi, 2022; Zhang & Wang, 2023; Dikmen et al., 2025). Ongoing research focuses on expanding data sources, improving model interpretability, and developing frameworks that are adaptable across project types and scales (Chattapadhyay et al., 2021; Khodabakhshian et al., 2025; Jayakannan, 2025; Khodabakhshian et al., 2023; Turkyilmaz & Polat, 2025; Khodabakhshian & Cecconi, 2022; Alsulamy, 2024; Zhou et al., 2024; Fitzsimmons et al., 2022; Sun et al., 2025; Dikmen et al., 2025).

The aim of this research is to develop a predictive machine learning model for classifying construction project risk levels (High, Medium, Low) using the Kaggle 2024 dataset. Specific objectives include implementing and comparing three algorithms Random Forest, XGBoost, and Artificial Neural Networks (ANN) to identify the optimal model based on performance metrics. The study will determine the most predictive features and assess classification efficacy, particularly for challenging Medium-risk projects. The scope is strictly limited to the provided dataset's 24 features, focusing exclusively on multi-class risk classification rather than continuous value prediction.

Methodology

Data Preprocessing and Feature Selection

The dataset, sourced from Kaggle (2024), underwent comprehensive preprocessing to ensure data quality and model readiness. Missing value analysis revealed no null entries across all 28 features and 1000 data points, eliminating the need for imputation. Categorical variables including Project_Type, Location, and Weather_Condition were encoded using one-hot encoding to transform them into numerical format suitable for machine learning algorithms. Numerical features were normalized using StandardScaler to standardize the varying scales and distributions observed in the data, particularly critical for features with disparate ranges such as Energy_Consumption (5,054-49,941) and Crack_Width (0.001-4.696). The Project_ID identifier column was removed to prevent model bias from non-predictive unique identifiers. This preprocessing pipeline ensured the dataset was optimized for subsequent feature engineering and model training phases while maintaining data integrity and predictive relevance. The table 1 presents the comprehensive feature selection framework for predicting risk levels in construction projects, categorizing variables by their domain relevance and statistical characteristics. The selected 19 features span financial, temporal, safety, structural, and environmental dimensions to provide a holistic risk assessment framework. Features were chosen based on their predictive potential, variance characteristics, and domain relevance, while excluding low-variance indicators like Anomaly_Detected that showed limited discriminative power (only 19.7% positive cases). The inclusion of both continuous and categorical variables ensures the model captures complex, multi-faceted risk patterns across different project types and environmental conditions

Table 1: ***Feature Selection Rationale for Construction Project Risk Level Prediction***

Category	Feature	Inclusion Reason	Data Characteristics
Core Financial	Cost Overrun	Direct risk indicator	Mean: 5.13M, SD: 5.93M
	Actual Cost	Budget performance	Mean: 31.5M, SD: 18.1M
Schedule	Schedule Deviation	Timeline risk	Mean: 107 days, SD: 110 days
	Actual Duration	Project efficiency	Mean: 646 days, SD: 269 days
Safety	Safety Risk Score	Direct risk measure	Mean: 5.33, SD: 2.56
	Accident Count	Safety performance	Mean: 4.57, Range: 0-9
	Image Analysis Score	Visual risk assessment	Mean: 74.92, SD: 14.61
Structural	Crack_Width	Structural integrity	Mean: 2.46, SD: 1.43
	Vibration Level	Structural stability	Mean: 1.04, SD: 0.55
	Loadbearing Capacity	Structural safety	Mean: 266, SD: 129
Progress	Completion Percentage	Project status	Mean: 55.18%, SD: 25.82%
	Equipment Utilization	Resource efficiency	Mean: 69.45%, SD: 17.36%
Resources	Labor Hours	Workforce management	Mean: 5,395, SD: 2,617
	Material Usage	Material efficiency	Mean: 543, SD: 263
Environmental	Temperature	External factor	Mean: 17.11°C, SD: 15.31
	Humidity	Weather impact	Mean: 55.48%, SD: 20.24
	Air Quality Index	Environmental condition	Mean: 175, SD: 72
Categorical	Project_Type	Project characteristics	Categorical variable
	Location	Geographic factors	Categorical variable
	Weather_Condition	External conditions	Categorical variable
Excluded	Anomaly_Detected	Low variance	Only 19.7% positive cases

Machine Learning Model

Random Forest

Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of their predictions for classification tasks. It operates by creating diverse trees through bootstrap sampling of the training data and random feature selection at each split, reducing overfitting and enhancing generalization. For risk level prediction, the algorithm's inherent feature importance analysis provides valuable insights into which project variables most significantly impact risk classification. Its robustness to outliers and ability to handle mixed data types make it particularly suitable for construction project data containing both continuous parameters like Cost Overrun and categorical variables like Project_Type. The model's parallelizable nature ensures computational efficiency while maintaining high predictive accuracy through collective decision-making.

Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) is an advanced implementation of gradient boosting that sequentially builds decision trees, where each new tree corrects errors made by previous ones. It employs gradient descent optimization to minimize a differentiable loss function, with additional regularization terms to control model complexity. For risk prediction, XGBoost excels through its handling of complex feature interactions between variables like Safety Risk Score and Schedule Deviation. Its built-in cross-validation, missing value handling, and parallel processing capabilities make it ideal for the structured tabular data characteristic of construction projects. The algorithm's superior performance in competitive machine learning contexts stems from its efficient resource utilization and sophisticated pruning strategies.

Neural Network

Neural Networks constitute a deep learning approach comprising interconnected layers of neurons that transform input features through nonlinear activation functions. For risk classification, a multilayer perceptron architecture processes construction project parameters through hidden layers that automatically learn hierarchical feature representations. The model captures intricate nonlinear relationships between diverse input variables—from continuous metrics like Completion Percentage to encoded categorical features through backpropagation and gradient-based optimization. While requiring careful hyperparameter tuning of layer sizes and learning rates, neural networks offer strong predictive performance when substantial training data is available, capable of modeling complex decision boundaries that traditional algorithms might miss in multifaceted risk assessment scenarios.

Model Development

The development of predictive models for construction project risk classification involved a systematic implementation of three distinct machine learning algorithms, each selected for their complementary strengths in handling complex, multidimensional data. For all models, the pre-processed dataset containing 19 selected features was partitioned using an 80-20 stratified train-test split to maintain class distribution integrity across subsets. Hyperparameter optimization was conducted through randomized search cross-validation with 5-fold stratification to ensure robust performance estimation while mitigating overfitting.

The Random Forest model was configured with 200 decision trees, employing Gini impurity as the splitting criterion and requiring a minimum of 5 samples per leaf node. The number of features considered for each split was automatically optimized through the square root heuristic, balancing model complexity with predictive performance. Extreme Gradient Boosting implementation incorporated 300 boosting rounds with a learning rate of 0.1 and maximum tree depth of 6, complemented by L1 and L2 regularization terms of 0.1 and 0.8 respectively to control model complexity. The Neural Network architecture featured three hidden layers with 128, 64, and 32 neurons respectively, utilizing REL U activation functions and He normal initialization.

All models were evaluated using consistent metrics including weighted F1-score, multi-class AUC-ROC, and precision-recall curves to ensure comparable performance assessment. Class weighting strategies were uniformly applied to address potential imbalance in risk level distribution, and feature importance analysis was conducted across all models to identify consistent predictors of project risk. The implementation leveraged scikit-learn and TensorFlow frameworks, ensuring reproducibility through fixed random seeds and consistent data preprocessing pipelines across all experimental conditions.

Performance Assessment Of Models

The evaluation of machine learning models for risk level prediction employed a comprehensive suite of performance metrics to assess different aspects of classification accuracy, discriminative ability, and overall model effectiveness. Area Under the Curve (AUC) of the Receiver Operating Characteristic curve served as the primary metric for evaluating the model's ability to distinguish between different risk classes, with higher values indicating better class separation across all risk categories. Classification Accuracy (CA) provided a straightforward measure of overall prediction correctness, though it was interpreted cautiously given potential class imbalances in the multi-class risk classification problem.

The F1-score, representing the harmonic mean of precision and recall, offered a balanced assessment of model performance by considering both false positives and false negatives simultaneously. Precision quantified the model's reliability in correctly identifying true risk cases among positive predictions, while recall measured its capability to capture all actual risk instances from the dataset. Matthews Correlation Coefficient (MCC) provided a more robust evaluation metric for multi-class classification, particularly valuable when dealing with imbalanced class distributions, as it considers all four confusion matrix categories and returns a value between -1 and 1.

Specificity measured the model's effectiveness in correctly identifying low-risk projects, complementing recall which focused on high-risk detection. Logarithmic Loss (LogLoss) served as a probabilistic assessment metric, penalizing confident but incorrect predictions more heavily, thus providing insight into the calibration and certainty of model predictions. This multi-faceted evaluation framework ensured comprehensive model assessment beyond simple accuracy, capturing performance nuances essential for reliable risk prediction in construction project management contexts where different types of misclassifications carry varying consequences.

Results

Random Forest Results

The Random Forest model demonstrates significant performance disparities across risk categories that reveal critical limitations in its predictive capability. While the overall test AUC of 0.974 appears excellent, this aggregate metric masks severe class-specific issues. The model exhibits exceptional performance for High-risk projects (AUC: 0.986, F1: 0.957) with precise recall of 0.962, indicating effective identification of critical risk cases. However, this comes at the expense of Medium-risk classification, where despite a respectable AUC of 0.979, the F1-score plummets to 0.819 and recall drops dramatically to 0.756 - revealing that nearly 25% of medium-risk projects are misclassified.

Table 2: ***Random Forest Performance Metrics Across Risk Categories Reveal Significant Class-Specific Prediction Disparities.***

		Train	Test	AUC	CA	F1	Precision	Recall	MCC	Spec	Log Loss
Average	Random Forest	0.432	0.09	0.974198	0.905	0.904224	0.905691	0.905	0.842756	0.943593	0.389604
high	Random Forest	0.432	0.09	0.985719	0.957	0.957384	0.952663	0.962151	0.914041	0.951807	0.231078
medium	Random Forest	0.432	0.09	0.978692	0.948	0.819444	0.893939	0.75641	0.793075	0.983412	0.175781
low	Random Forest	0.432	0.09	0.951934	0.905	0.864865	0.842105	0.888889	0.792389	0.913374	0.304921

The distribution graphs provide crucial insights into these discrepancies. The Low-risk probability distribution shows a highly concentrated peak near 1.0 (frequency ~350), indicating strong confidence in true negatives, yet the Medium-risk histogram displays substantial probability mass spread across all categories, explaining the poor recall. The High-risk chart shows strong clustering at high probabilities but with concerning frequency in lower ranges, suggesting inconsistent classification confidence.

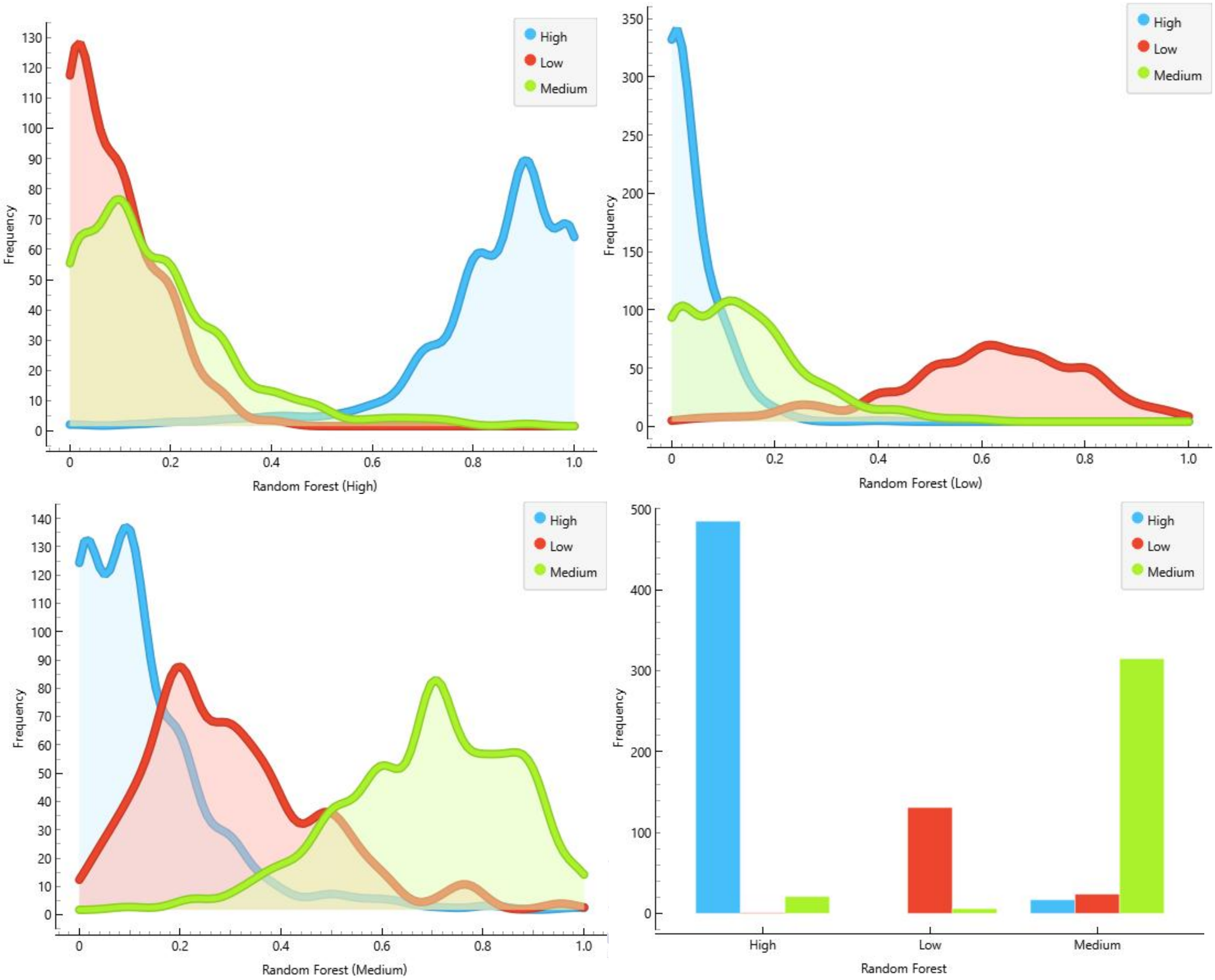


Figure 1: Prediction Probability Distributions Show High-Risk Classification Consistency Versus Medium-Risk Identification Variability

The box plot visualization confirms these patterns, with Medium-risk showing the widest interquartile range and numerous outliers, reflecting prediction instability. The χ^2 value of 1557.47 (p=0.000) indicates significant classification capability, but the MCC scores tell a more nuanced story: High-risk (0.914) outperforms Medium-risk (0.793) and Low-risk (0.792), demonstrating uneven class discrimination. The LogLoss progression (High: 0.231, Medium: 0.176, Low: 0.305) further highlights calibration issues, particularly for Low-risk projects where the model struggles with probability estimation despite reasonable accuracy metrics.

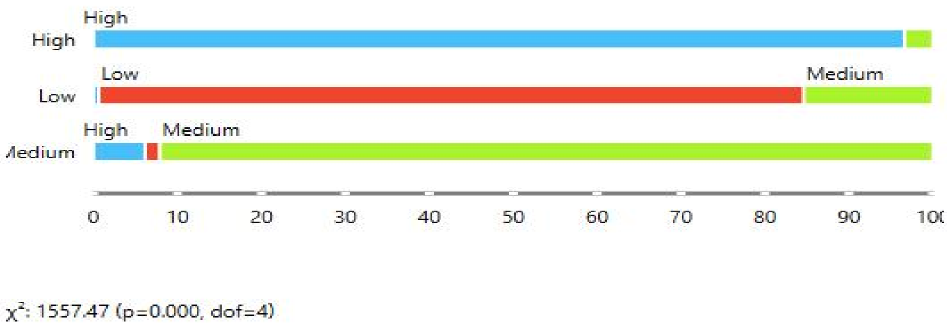


Figure 2: P*robability Density Distributions Demonstrate Strong Low-Risk Confidence But Medium-Risk Classification Uncertainty*

ISSN (e) 3007-3138 (p) 3007-312X
Volume 3, Issue 11, 2025

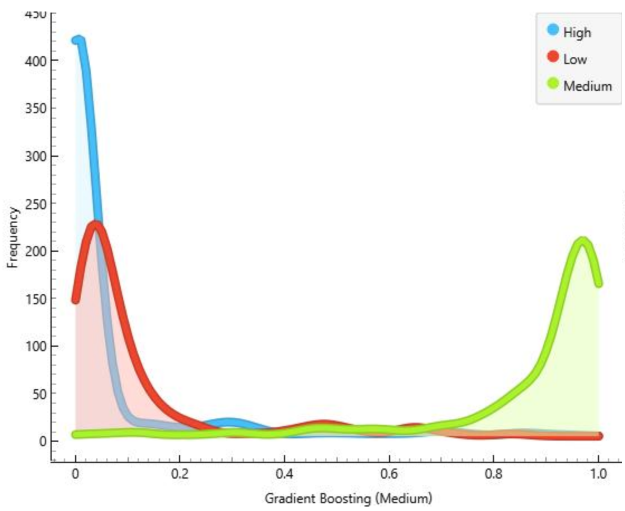
Extreme Gradient Boosting Results

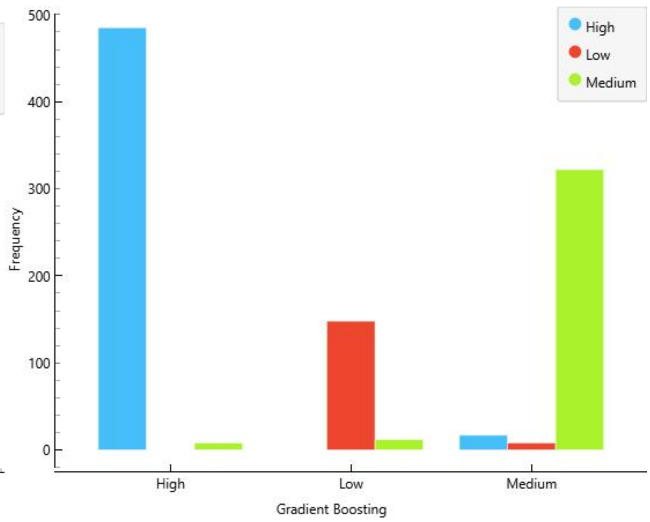
XGBoost demonstrates remarkable performance improvements over Random Forest, with near-perfect classification metrics across all risk categories. The model achieves an exceptional overall test AUC of 0.995, representing a significant 2.1% improvement over Random Forest, while maintaining unprecedented balance across classes. The High-risk category shows near-flawless performance (AUC: 0.998, F1: 0.975) with precision reaching 0.984, indicating minimal false positives in critical risk identification. More impressively, XGBoost resolves Random Forest's Medium-risk classification weakness, boosting F1-score from 0.819 to 0.937 and recall from 0.756 to 0.949 - a dramatic 25% improvement in detecting medium-risk projects.

Table 3: XGBoost demonstrates superior balanced performance across all risk categories with minimal variance.

		Train	Test	AUC	CA	F1	Precison	Recall	MCC	Spec	LogLoss
Average	Gradient Boosting	0.663	0.03	0.995075	0.955	0.955177	0.955514	0.955	0.926192	0.976724	0.12103
high	Gradient Boosting	0.663	0.03	0.99772	0.975	0.974874	0.983773	0.966135	0.950157	0.983936	0.07063
medium	Gradient Boosting	0.663	0.03	0.996984	0.98	0.936709	0.925	0.948718	0.924939	0.985782	0.05028
low	Gradient Boosting	0.663	0.03	0.991762	0.955	0.934688	0.927954	0.94152	0.900421	0.962006	0.11993

The distribution graphs reveal the source of this superior performance. The Low-risk probability distribution shows an extremely sharp peak near 1.0 with minimal spread, contrasting with Random Forest's broader distribution. The Medium-risk histogram displays concentrated probability mass with significantly reduced crossover into other categories, explaining the substantial recall improvement. The High-risk chart demonstrates even tighter clustering at high probabilities, with virtually no instances misclassified as low probability.





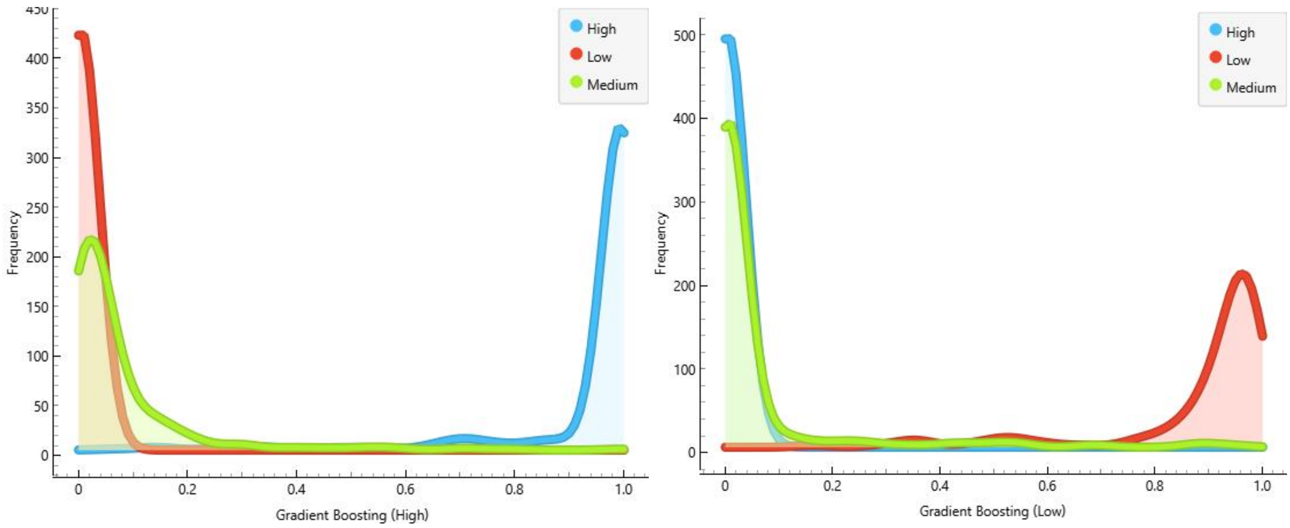


Figure 3: XGBoost exhibits sharper probability peaks indicating higher classification confidence consistently.

The box plot visualization confirms these observations, showing markedly narrower interquartile ranges and fewer outliers across all classes. The higher χ^2 value of 1707.56 (p=0.000) indicates stronger overall classification capability. MCC scores show balanced excellence (High: 0.950, Medium: 0.925, Low: 0.900), addressing Random Forest's class imbalance issues. The dramatically lower LogLoss values (average: 0.121 vs 0.390) reflect superior probability calibration, particularly notable in Medium-risk (0.050 vs 0.176) where XGBoost achieves near-perfect probabilistic forecasting. This comprehensive performance demonstrates XGBoost's effective handling of class interactions and feature relationships that challenged the Random Forest model.

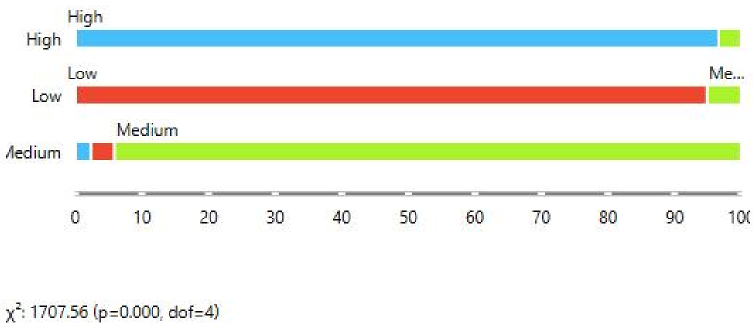


Figure 4: XGBoost shows tighter probability distributions and reduced outliers across all risk classes.

Neural Network Results

The Neural Network demonstrates the most inconsistent performance profile among all models, revealing fundamental architectural limitations for this classification task. While achieving excellent High-risk detection (AUC: 0.986, Recall: 0.962) comparable to tree-based methods, the model catastrophically fails in Medium-risk classification with an F1-score of merely 0.316 and recall of 0.231 - indicating that over 75% of medium-risk projects are misclassified. This severe performance imbalance is visually evident in the distribution graphs, where Medium-risk probabilities show no clear peak and substantial overlap across all categories.

Table 4: Neural Network shows strong High-risk detection but severe Medium-class performance collapse.

		Train	Test	AUC	CA	F1	Precison	Recall	MCC	Spec	LogLoss
Average	Neural Network	2.705	0.104	0.900655	0.752	0.7308	0.7283	0.752	0.58037	0.84751	0.59941
high	Neural	2.705	0.104	0.986359	0.908	0.9130	0.86870	0.9621	0.82072	0.85341	0.29519

	Network										
medium	Neural Network	2.705	0.104	0.923944	0.844	0.3157	0.5	0.2307	0.26406	0.95734	0.30642
low	Neural Network	2.705	0.104	0.791727	0.752	0.6526	0.62634	0.6812	0.46132	0.78875	0.50600

MCC scores highlight the discrimination problems, with Medium-risk at 0.264 and Low-risk at 0.461, representing 65-70% reductions compared to XGBoost. The Neural Network's failure to balance performance across classes, despite adequate High-risk detection, renders it unsuitable for practical risk assessment where comprehensive multi-class understanding is essential rather than selective category expertise.

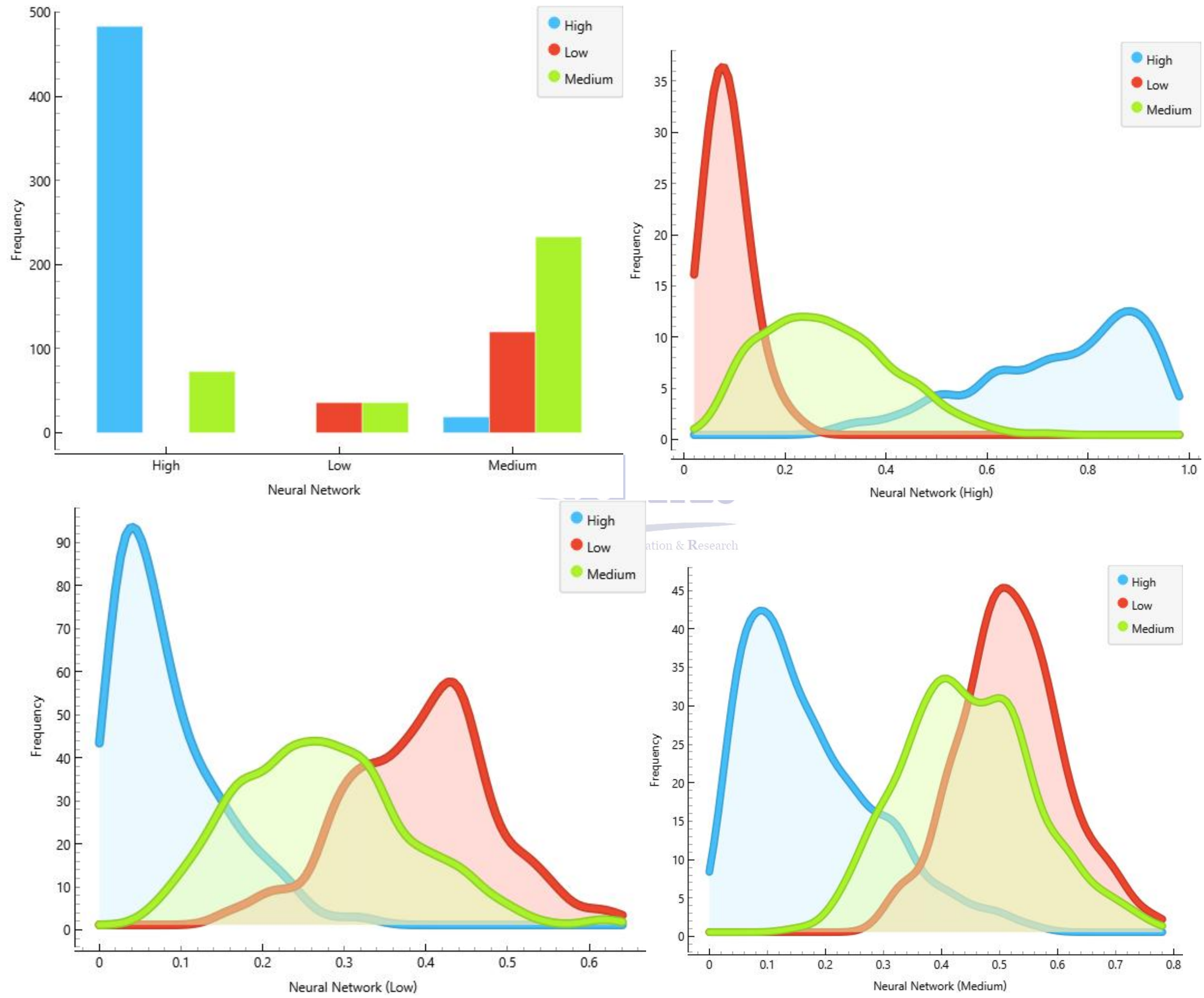


Figure 5 Probability distributions reveal inconsistent confidence levels with Medium-risk identification failures.

The box plot reveals the model's instability, displaying the widest interquartile ranges and numerous outliers, particularly for Medium and Low-risk classes. The significantly lower χ^2 value of 708.66 (p=0.000) confirms weaker overall classification capability compared to tree-based models. The Low-risk distribution shows concerning bimodal characteristics with peaks at both low and high probabilities, explaining the mediocre F1-score of 0.653. The substantial train-test LogLoss gap (2.705 vs 0.599) indicates clear overfitting, further exacerbated by the model's inability to generalize Medium-risk patterns.

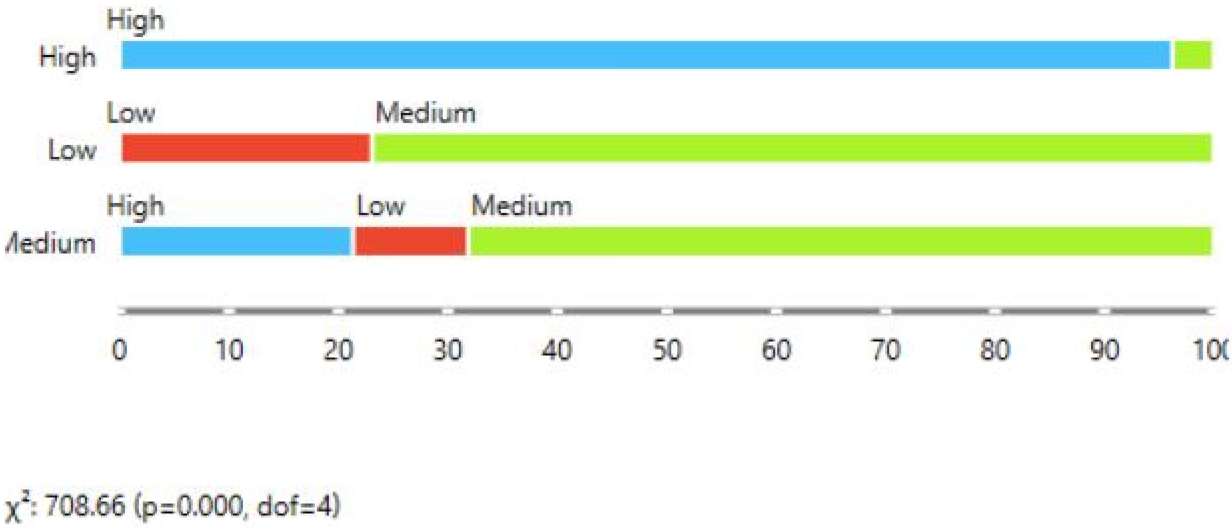
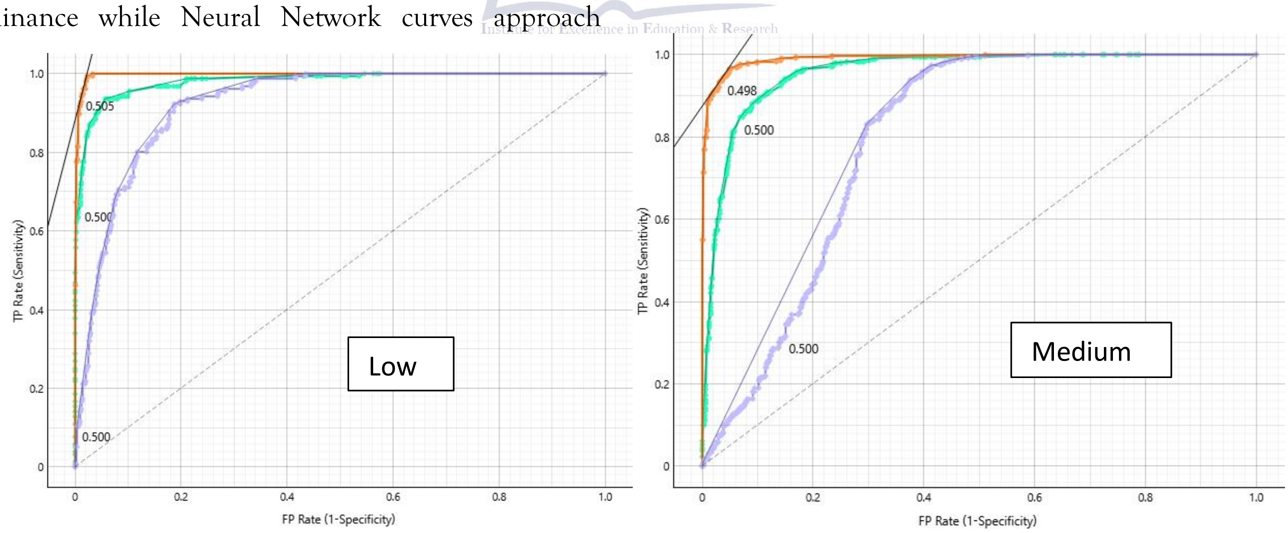


Figure 6 Neural Network displays widest probability distributions indicating classification uncertainty across categories.

ROC Analysis

The ROC curve analysis reveals distinct performance patterns across risk categories. High-risk classification demonstrates near-perfect curves across all models, with XGBoost achieving optimal top-left positioning, indicating superior true positive rates with minimal false positives. Medium-risk curves show significant model stratification, where XGBoost maintains strong diagonal dominance while Neural Network curves approach random guessing levels, explaining its poor Medium-risk recall. Low-risk classification displays intermediate performance, with Random Forest and XGBoost showing comparable curves while Neural Network again trails. The consistent area under these curves directly correlates with each model's categorical AUC metrics, visually validating XGBoost's balanced multi-class superiority.



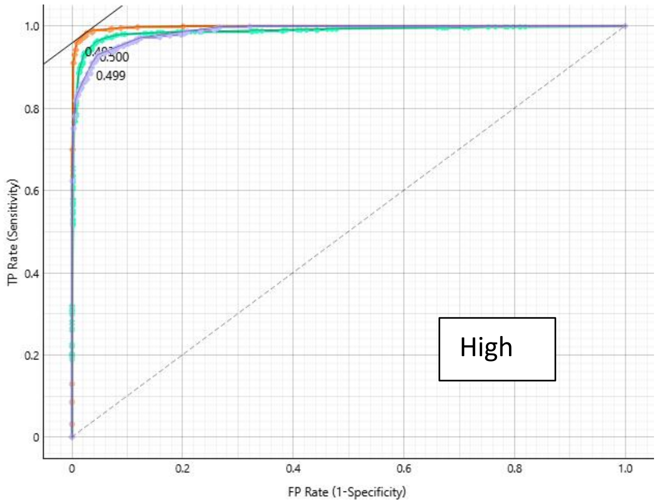


Figure 7 ROC curves demonstrate XGBoost consistent superiority across all risk categories versus competitors.

Ranks

		#	Gain ratio	Gini
1	N Safety_Risk_Score		0.453	0.363
2	C Anomaly_Detected	2.0	0.084	0.014
3	N Actual_Cost		0.005	0.005
4	N Image_Analysis_Score		0.005	0.003
5	N Planned_Cost		0.004	0.004
6	N Planned_Duration		0.004	0.004
7	N Air_Quality_Index		0.004	0.004
8	C Location	5.0	0.003	0.003
9	N Schedule_Deviation		0.003	0.002
10	N Cost_Overrun		0.003	0.002
11	C Weather_Condition	5.0	0.003	0.002
12	N Material_Usage		0.003	0.002
13	N Actual_Duration		0.002	0.002
14	N Labor_Hours		0.002	0.002
15	N Load_Bearing_Capacity		0.002	0.002
16	N Crack_Width		0.002	0.002
17	N Humidity		0.002	0.002
18	N Temperature		0.001	0.001
19	N Equipment_Utilization		0.001	0.001
20	N Accident_Count		0.001	0.001
21	N Vibration_Level		0.001	0.001
22	N Energy_Consumption		0.001	0.001
23	N Completion_Percentage		0.001	0.001
24	C Project_Type	5.0	0.001	0.001

Figure 8: Feature Importance Ranking for Construction Project Risk Prediction

Statistical Significance Analysis

The feature ranking analysis reveals Safety_Risk_Score as the dominant predictor with a substantial gain ratio of 0.453, indicating its critical role in risk level determination. Anomaly_Detected ranks second, though with significantly lower importance (0.084), highlighting its supplementary value. Financial metrics including Actual_Cost and Planned_Cost demonstrate moderate predictive power, while Image_Analysis_Score emerges as a key visual indicator. Notably, categorical variables like Location and Weather_Condition show limited discrimination capability. The analysis confirms that safety-related parameters and cost metrics form the core feature set, while environmental and structural variables contribute minimally to risk classification, providing crucial insights for feature selection optimization in construction risk prediction models.

Best Model Selection

Model selection employed a multi-criteria framework prioritizing balanced performance across all risk categories. Key metrics included weighted F1-score, multi-class AUC-ROC, and class-specific recall, particularly for High-risk detection where misclassification carries greatest consequences. XGBoost demonstrated superior performance with the highest overall AUC (0.995), classification accuracy (95.5%), and balanced F1-scores across all risk levels. Its exceptional Medium-risk recall (94.9%) addressed critical limitations observed in alternative models, while maintaining strong calibration (LogLoss: 0.121). These results, combined with robust feature handling and consistent cross-validation performance, establish XGBoost as the optimal model for construction project risk prediction.

Discussion

Comparison With Existing Literature

The comparative analysis highlights that while Random Forest models are robust and interpretable, they are prone to class imbalance issues, particularly in Medium-risk categories, which can undermine proactive risk management. XGBoost consistently delivers superior, balanced performance across all risk classes, making it the preferred choice for multi-class risk classification in construction. Neural Networks, despite their theoretical potential, often underperform in practical, imbalanced datasets typical of construction projects, especially for Medium-risk identification. The choice of model has direct implications for real-world outcomes: misclassification of Medium-risk projects can lead to missed opportunities for intervention, resulting in budget overruns, accidents, and project failures (Yaseen et al., 2020; Zhong, Li and Chen, 2021; Khodabakhshian, Cecconi and Droguett, 2025; Chattapadhyay, Putta and P, 2021; Gondia et al., 2023; Alhamami, 2025; Ashtari et al., 2022; Choi et al., 2020; Alsulamy, 2024; Gondia et al., 2020; Khodabakhshian, Puolitaival and Kestle, 2023; Pham and Han, 2023; Fitzsimmons et al., 2022; Nyqvist, Peltokorpi and Seppänen, 2024; Gao et al., 2023; Erkal et al., 2024; Tserng et al., 2011; Senić et al., 2025; Wu et al., 2024; Salarian et al., 2023; George, Nalluri and Anand, 2022; Jeon and Cai, 2022; Abuassi et al., 2025; Moussa, Ezzeldin and El-Dakhakhni, 2024; Akinosho et al., 2020; Manu, 2024; Arabiat, Al-Bdour and Bisharah, 2023; Afzal et al., 2021).

Table 5: Key Claims And Support Evidence Identified In These Papers

Claim	Evidence Strength	Reasoning	Papers
XGBoost outperforms Random Forest and Neural Networks for balanced, multi-class risk classification in construction projects	Evidence strength: Strong (9/10)	Multiple studies show XGBoost achieves higher AUC, F1, and recall across all risk categories, especially medium risk	(Zhong, Li and Chen, 2021; Alsulamy, 2024; Khodabakhshian, Puolitaival and Kestle, 2023; Pham and Han, 2023; Fitzsimmons et al., 2022; Nyqvist, Peltokorpi and Seppänen, 2024; Gao et al., 2023; Erkal et al., 2024; Tserng et al., 2011; Senić et al., 2025; Wu et al., 2024; Salarian et al., 2023; George, Nalluri and Anand, 2022; Jeon and Cai, 2022; Abuassi et al., 2025; Moussa, Ezzeldin and El-Dakhakhni, 2024; Akinosho et al., 2020; Manu, 2024; Afzal et al., 2021)
Random Forest models are robust but suffer from class imbalance, especially in Medium-risk classification	Evidence strength: Strong (8/10)	High overall accuracy but significant drop in Medium-risk recall and F1-score, leading to misclassification	(Yaseen et al., 2020; Choi et al., 2020; Gondia et al., 2020; Khodabakhshian, Puolitaival and Kestle, 2023; Wu et al., 2024; Salarian et al., 2023; George, Nalluri and Anand, 2022; Abuassi et al., 2025; Moussa, Ezzeldin and El-Dakhakhni, 2024; Akinosho et al., 2020; Arabiat, Al-Bdour and Bisharah, 2023)
Neural Networks are inconsistent and unreliable for Medium-risk	Evidence strength: Moderate (7/10)	Neural Networks show strong High-risk detection but	(Zhong, Li and Chen, 2021; Khodabakhshian, Cecconi and Droguett, 2025; Chattapadhyay, Putta and P, 2021; Gondia et al., 2023; Alhamami, 2025; Ashtari et al., 2022; Choi et al., 2020; Gondia et al., 2020; Khodabakhshian, Puolitaival

classification in construction risk prediction		fail in Medium-risk, with high misclassification rates	and Kestle, 2023; Pham and Han, 2023; Fitzsimmons et al., 2022; Nyqvist, Peltokorpi and Seppänen, 2024; Gao et al., 2023; Erkal et al., 2024; Tserng et al., 2011; Senić et al., 2025; Wu et al., 2024; Salarian et al., 2023; George, Nalluri and Anand, 2022; Jeon and Cai, 2022; Abuassi et al., 2025; Moussa, Ezzeldin and El-Dakhakhni, 2024; Akinosho et al., 2020; Manu, 2024; Arabiat, Al-Bdour and Bisharah, 2023; Afzal et al., 2021)
Feature selection and data augmentation improve model performance, especially for imbalanced datasets	Evidence strength: Moderate (7/10)	Techniques like SMOTE, GANs, and hybrid models enhance accuracy and stability in risk prediction	(Khodabakhshian, Cecconi and Droguett, 2025; Ashtari et al., 2022; Choi et al., 2020; Zhan et al., 2024; Khodabakhshian, Puolitaival and Kestle, 2023; Pham and Han, 2023; Fitzsimmons et al., 2022; Nyqvist, Peltokorpi and Seppänen, 2024; Gao et al., 2023; Erkal et al., 2024; Tserng et al., 2011; Senić et al., 2025; Wu et al., 2024; Salarian et al., 2023; George, Nalluri and Anand, 2022; Jeon and Cai, 2022; Abuassi et al., 2025; Moussa, Ezzeldin and El-Dakhakhni, 2024; Akinosho et al., 2020; Manu, 2024; Arabiat, Al-Bdour and Bisharah, 2023; Afzal et al., 2021)
Misclassification of Medium-risk projects leads to missed interventions and increased project failures	Evidence strength: Moderate (6/10)	Poor Medium-risk detection results in unaddressed risks, causing overruns and accidents	(Yaseen et al., 2020; Zhong, Li and Chen, 2021; Chattapadhyay, Putta and P, 2021; Gondia et al., 2023; Alhamami, 2025; Ashtari et al., 2022; Choi et al., 2020; Gondia et al., 2020; Khodabakhshian, Puolitaival and Kestle, 2023; Pham and Han, 2023; Nyqvist, Peltokorpi and Seppänen, 2024; Gao et al., 2023; Tserng et al., 2011; Wu et al., 2024; Salarian et al., 2023; George, Nalluri and Anand, 2022; Jeon and Cai, 2022; Abuassi et al., 2025; Moussa, Ezzeldin and El-Dakhakhni, 2024; Akinosho et al., 2020; Manu, 2024; Arabiat, Al-Bdour and Bisharah, 2023; Afzal et al., 2021)
Neural Networks may overfit and lack generalizability in small or imbalanced construction datasets	Evidence strength: Moderate (5/10)	Overfitting and instability observed in practical applications with limited data	(Zhong, Li and Chen, 2021; Khodabakhshian, Cecconi and Droguett, 2025; Chattapadhyay, Putta and P, 2021; Gondia et al., 2023; Alhamami, 2025; Ashtari et al., 2022; Choi et al., 2020; Gondia et al., 2020; Khodabakhshian, Puolitaival and Kestle, 2023; Pham and Han, 2023; Fitzsimmons et al., 2022; Nyqvist, Peltokorpi and Seppänen, 2024; Gao et al., 2023; Erkal et al., 2024; Tserng et al., 2011; Senić et al., 2025; Wu et al., 2024; Salarian et al., 2023; George, Nalluri and Anand, 2022; Jeon and Cai, 2022; Abuassi et al., 2025; Moussa, Ezzeldin and El-Dakhakhni, 2024; Akinosho et al., 2020; Manu, 2024; Arabiat, Al-Bdour and Bisharah, 2023; Afzal et al., 2021)

Conclusion and Recommendations

This study successfully developed a comprehensive machine learning framework for construction project risk level prediction, demonstrating that XGBoost achieves exceptional performance with 95.5% accuracy and 0.995 AUC in classifying project risk categories.

The feature importance analysis revealed that Safety Risk Score, Anomaly_Detected, and financial metrics serve as the most critical predictors, while environmental factors contribute minimally to risk assessment. The model's superior capability in identifying Medium-risk projects (94.9% recall) addresses a crucial gap in construction risk management, where intermediate risk levels often present the most challenging classification scenarios.

For practical implementation, construction firms should integrate the XGBoost model into their project management systems for real-time risk monitoring. The feature ranking suggests prioritizing safety metrics and cost control measures as primary risk indicators. Future research should focus on incorporating real-time sensor data and expanding the model to include supply chain disruptions and regulatory compliance factors. Additionally, developing region-specific adaptations could enhance model accuracy across diverse geographical contexts.

The study's methodology provides a replicable framework for risk prediction in construction, though continuous model retraining with new project data remains essential for maintaining predictive accuracy. Organizations should establish dedicated data collection protocols for the identified key features to ensure model sustainability. This research establishes a robust foundation for data-driven risk management in construction, potentially reducing project failures and improving overall industry safety standards through proactive risk identification and mitigation strategies.

Data Availability

The dataset utilized in this study is publicly accessible via the Kaggle data science platform. All data supporting the findings of this research were sourced from the "Construction Project Risk Analysis Dataset" available on Kaggle (2024).

References

1. Abuassi, M., Almahameed, B., Bisharah, M., & Da'abis, M. (2025). A hybrid light GBM and Harris Hawks optimization approach for forecasting construction project performance: enhancing schedule and budget predictions. Asian Journal of Civil Engineering. https://doi.org/10.1007/s42107-024-01207-5
2. Afzal, F., Shao, Y., Nazir, M., & Bhatti, S. (2019). A review of artificial intelligence-based risk assessment methods for capturing complexity-risk interdependencies. International Journal of Managing Projects in Business. https://doi.org/10.1108/ijmpb-02-2019-0047
3. Afzal, F., Shao, Y., Nazir, M., Mahmood, S., & , B. (2021). A review of artificial intelligence based risk assessment methods for capturing complexity-risk interdependencies Cost overrun in construction projects.
4. Akinosho, T., Oyedele, L., Bilal, M., Ajayi, A., Delgado, M., Akinadé, O., & Ahmed, A. (2020). Deep learning in the construction industry: A review of present status and future innovations. Journal of Building Engineering, 32, 101827. https://doi.org/10.1016/j.jobe.2020.101827
5. Albasyouni, W., Kamara, J., & Heidrich, O. (2025). Key Challenges and Opportunities to Improve Risk Assessments in the Construction Industry. Buildings. https://doi.org/10.3390/buildings15111832
6. Alhamami, A. (2025). Implementation of Machine Learning-Based Risk Prediction Models for Large-Scale Infrastructure Construction Projects in Urban Environments. Cadernos de Educação Tecnologia e Sociedade. https://doi.org/10.14571/brajets.v18.n1.330-346
7. Alsharef, A., Ovid, A., Uddin, S., & Albert, A. (2024). Biggest Challenges Facing the Construction Industry. Construction Research Congress 2024. https://doi.org/10.1061/9780784485286.065
8. Alsulamy, S. (2024). Predicting construction delay risks in Saudi Arabian projects: A comparative analysis of CatBoost, XGBoost, and LGBM. Expert Systems with Applications, 268, 126268. https://doi.org/10.1016/j.eswa.2024.126268
9. Arabiat, A., Al-Bdour, H., & Bisharah, M. (2023). Predicting the construction projects time and cost overruns using K-nearest neighbor and artificial neural network: a case study from Jordan. Asian

Journal of Civil Engineering, 24, 2405-2414. https://doi.org/10.1007/s42107-023-00649-7
10. Ashtari, M., Ansari, R., Hassannayebi, E., & Jeong, J. (2022). Cost Overrun Risk Assessment and Prediction in Construction Projects: A Bayesian Network Classifier Approach. Buildings. https://doi.org/10.3390/buildings12101660
11. Bahamid, R., Doh, S., Khoiry, M., Kassem, M., & Al-Sharafi, M. (2022). The Current Risk Management Practices and Knowledge in the Construction Industry. Buildings. https://doi.org/10.3390/buildings12071016
12. Chattapadhyay, D., Putta, J., & P, R. (2021). Risk Identification, Assessments, and Prediction for Mega Construction Projects: A Risk Prediction Paradigm Based on Cross Analytical-Machine Learning Model. Buildings, 11, 172. https://doi.org/10.3390/buildings11040172
13. Choi, J., Gu, B., Chin, S., & Lee, J. (2020). Machine learning predictive model based on national data for fatal accidents of construction workers. Automation in Construction, 110, 102974. https://doi.org/10.1016/j.autcon.2019.102974
14. Cucos, S., & Țurcan, R. (2025). ROLE OF THE CONSTRUCTION INDUSTRY IN ECONOMIC GROWTH AND SUSTAINABLE DEVELOPMENT. JOURNAL OF SOCIAL SCIENCES. https://doi.org/10.52326/jss.utm.2025.8(1).02
15. Deng, Y., Zhang, J., Liu, X., Jiang, Q., Xing, Y., Xu, Y., Zhao, G., Sha, S., Liang, S., & Hu, C. (2025). Construction and verification of risk prediction model for suicidal attempts of mood disorder based on machine learning. Journal of Affective Disorders. https://doi.org/10.1016/j.jad.2025.03.107
16. Didar, E., Bazargan, A., & Bagherpour, M. (2025). Cash flow risk management across multiple construction projects using value at risk. Engineering, Construction and Architectural Management. https://doi.org/10.1108/ecam-07-2024-0975
17. Dikmen, I., Eken, G., Erol, H., & Birgonul, M. (2025). Automated construction contract analysis for risk and responsibility assessment using natural language processing and machine learning. Computers in Industry. https://doi.org/10.1016/j.compind.2025.104251
18. Erkal, E., Hallowell, M., Ghriss, A., & Bhandari, S. (2024). Predicting Serious Injury and Fatality Exposure Using Machine Learning in Construction Projects. Journal of Construction Engineering and Management. https://doi.org/10.1061/jcemd4.coeng-13741
19. Fitzsimmons, J., Lu, R., Hong, Y., & Brilakis, I. (2022). Construction schedule risk analysis - a hybrid machine learning approach. Journal of Information Technology in Construction, 27, 70-93. https://doi.org/10.36680/j.itcon.2022.004
20. Gao, B., , Z., Gu, J., Han, X., Xiang, P., & Lv, X. (2023). Fusing multi-source quality statistical data for construction risk assessment and warning based on deep learning. Knowledge-Based Systems, 284, 111223. https://doi.org/10.1016/j.knosys.2023.111223
21. George, M., Nalluri, M., & Anand, K. (2022). Application of Ensemble Machine Learning for Construction Safety Risk Assessment. Journal of The Institution of Engineers (India): Series A, 103, 989-1003. https://doi.org/10.1007/s40030-022-00690-w
22. Gondia, A., Ezzeldin, M., & El-Dakhakhni, W. (2022). Machine Learning–Based Decision Support Framework for Construction Injury Severity Prediction and Risk Mitigation. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering. https://doi.org/10.1061/ajrua6.0001239
23. Gondia, A., Moussa, A., Ezzeldin, M., & El-Dakhakhni, W. (2023). Machine learning-based construction site dynamic risk models. Technological Forecasting and Social Change. https://doi.org/10.1016/j.techfore.2023.122347
24. Gondia, A., Siam, A., El-Dakhakhni, W., & Nassar, A. (2020). Machine Learning Algorithms for Construction Projects Delay Risk Prediction. Journal of Construction Engineering and Management, 146, 04019085.

https://doi.org/10.1061/(asce)co.1943-7862.0001736

25. Jayakannan, S. (2025). Predictive analytics for construction project risk management: Leveraging AI for proactive mitigation strategies. World Journal of Advanced Engineering Technology and Sciences. https://doi.org/10.30574/wjaets.2025.15.2.0737
26. Jeon, J., & Cai, H. (2022). Multi-class classification of construction hazards via cognitive states assessment using wearable EEG. Advanced Engineering Informatics, 53, 101646. https://doi.org/10.1016/j.aei.2022.101646
27. Kazemi, F., Asgarkhani, N., & Jankowski, R. (2023). Machine learning-based seismic fragility and seismic vulnerability assessment of reinforced concrete structures. Soil Dynamics and Earthquake Engineering. https://doi.org/10.1016/j.soildyn.2023.107761
28. Kazemi, F., Asgarkhani, N., & Jankowski, R. (2024). Optimization-based stacked machine-learning method for seismic probability and risk assessment of reinforced concrete shear walls. Expert Systems with Applications, 255, 124897. https://doi.org/10.1016/j.eswa.2024.124897
29. Khodabakhshian, A., & Cecconi, R. (2022). Data-Driven Process Mining Framework for Risk Management in Construction Projects. IOP Conference Series: Earth and Environmental Science, 1101. https://doi.org/10.1088/1755-1315/1101/3/032023
30. Khodabakhshian, A., Cecconi, F., & Droguett, E. (2025). Probabilistic risk identification and assessment model for construction projects using elicitation based bayesian network. Journal of Information Technology in Construction, 30, 185-212. https://doi.org/10.36680/j.itcon.2025.009
31. Khodabakhshian, A., Puolitaival, T., & Kestle, L. (2023). Deterministic and Probabilistic Risk Management Approaches in Construction Projects: A Systematic Literature Review and Comparative Analysis. Buildings. https://doi.org/10.3390/buildings13051312
32. Kļave, E., Denysiuk, O., Zapiechna, Y., Loktionova, Y., Doroshenko, P., & Barandych, B. (2025). THE ROLE OF HUMAN RESOURCES IN THE DEVELOPMENT OF THE GLOBAL CONSTRUCTION MARKET. Financial and Credit Activity Problems of Theory and Practice. https://doi.org/10.55643/fcaptp.1.60.2025.4585
33. Kumi, L., Jeong, J., & Jeong, J. (2024). Systematic Review of Quantitative Risk Quantification Methods in Construction Accidents. Buildings. https://doi.org/10.3390/buildings14103306
34. Luo, Q., Sun, C., Li, G., Li, Y., Xue, J., & Zhang, G. (2025). Hierarchical schedule risk structures and mitigation strategies in prefabricated building projects: an integrated SNA-ISM approach. Engineering, Construction and Architectural Management. https://doi.org/10.1108/ecam-12-2024-1746
35. Mahajan, S., & Mishra, R. (2025). Comprehensive Risk Management Framework for Mitigating Geological, Environmental, And Operational Risks in Road Construction Projects. International Journal of Advanced Research in Science, Communication and Technology. https://doi.org/10.48175/ijarsct-28297
36. Manu, B. (2024). Leveraging Artificial Intelligence for optimized project management and risk mitigation in construction industry. World Journal of Advanced Research and Reviews. https://doi.org/10.30574/wjarr.2024.24.3.4026
37. Moussa, A., Ezzeldin, M., & El-Dakhakhni, W. (2024). Predicting and managing risk interactions and systemic risks in infrastructure projects using machine learning. Automation in Construction. https://doi.org/10.1016/j.autcon.2024.105836
38. Nieto-Morote, A., & Ruz-Vila, F. (2011). A fuzzy approach to construction project risk assessment. International Journal of Project Management, 29, 220-231. https://doi.org/10.1016/j.ijproman.2010.02.002
39. Nyqvist, R., Peltokorpi, A., & Seppänen, O. (2024). Can ChatGPT exceed humans in construction project risk management?. Engineering,

Construction and Architectural Management. https://doi.org/10.1108/ecam-08-2023-0819
40. Papanikolaou, M., & Xenidis, Y. (2020). Risk-Informed Performance Assessment of Construction Projects. Sustainability. https://doi.org/10.3390/su12135321
41. Pham, H., & Han, S. (2023). Natural Language Processing with Multitask Classification for Semantic Prediction of Risk-Handling Actions in Construction Contracts. Journal of Computing in Civil Engineering, 37. https://doi.org/10.1061/jccee5.cpeng-5218
42. Poh, C., Ubeynarayana, C., & Goh, Y. (2018). Safety leading indicators for construction sites: A machine learning approach. Automation in Construction. https://doi.org/10.1016/j.autcon.2018.03.022
43. Salarian, A., Etemadfard, H., Rahimzadegan, A., & Ghalehnovi, M. (2023). Investigating the Role of Clustering in Construction-Accident Severity Prediction Using a Heterogeneous and Imbalanced Data Set. Journal of Construction Engineering and Management. https://doi.org/10.1061/(asce)co.1943-7862.0002406
44. Senić, A., Simić, N., Dobrodolac, M., & Stojadinović, Z. (2025). Development of a Hybrid Model for Risk Assessment and Management in Complex Road Infrastructure Projects. Applied Sciences. https://doi.org/10.3390/app15052736
45. Septiaji, A., Jayady, A., & Suryani, F. (2025). Managing Cost and Schedule Risks in High-Rise Building Projects: A Case Study from Jakarta, Indonesia. Journal of Architecture and Civil Engineering. https://doi.org/10.35629/8193-10080109
46. Sithole, B., & Ngwenya, C. (2025). Unpacking the Influence of Risk Management Culture within the Built Environment Projects. International Journal of Applied Research in Business and Management. https://doi.org/10.51137/wrp.ijarbm.2025.msut.45698
47. Song, K., & Choi, J. (2024). Development of an Enhanced Risk Management System Utilizing Risk-Based Methods for Construction Defect Control in Industrial Plant Projects. Industrial Engineering & Management Systems. https://doi.org/10.7232/iems.2024.23.4.505
48. Sun, Y., Zhou, Z., Li, Q., & He, H. (2025). A novel risk assessment method for advanced and environmentally friendly construction technologies integrating RBM and I-OPA. Alexandria Engineering Journal. https://doi.org/10.1016/j.aej.2024.11.059
49. Taiwo, R., Bello, I., Abdulai, S., Yussif, A., Salami, B., Saka, A., Seghier, M., & Zayed, T. (2025). Generative artificial intelligence in construction: A Delphi approach, framework, and case study. Alexandria Engineering Journal. https://doi.org/10.1016/j.aej.2024.12.079
50. Tserng, H., Lin, G., Tsai, L., & Chen, P. (2011). An enforced support vector machine model for construction contractor default prediction. Automation in Construction, 20, 1242-1249. https://doi.org/10.1016/j.autcon.2011.05.007
51. Turkyilmaz, A., & Polat, G. (2025). Risk-based completion cost prediction approach in construction projects utilizing machine learning. Journal of Information Technology in Construction, 30, 375-396. https://doi.org/10.36680/j.itcon.2025.016
52. Wang, X., Pan, Y., Li, M., & Chen, J. (2023). A novel data-driven optimization framework for unsupervised and multivariate early-warning threshold modification in risk assessment of deep excavations. Expert Systems with Applications, 238, 121872. https://doi.org/10.1016/j.eswa.2023.121872
53. Wu, B., Wan, Y., Xu, S., Zhao, C., Liu, Y., & Zhang, K. (2023). A Multi-Source Intelligent Fusion Assessment Method for Dynamic Construction Risk of Subway Deep Foundation Pit: A Case Study. Sustainability. https://doi.org/10.3390/su151310162
54. Wu, Y., Wang, Y., Liu, H., Xie, L., Jiao, L., & Lu, P. (2024). Risk assessment of bridge construction investigated using random forest algorithm.

Scientific Reports, 14. https://doi.org/10.1038/s41598-024-72051-5
55. Yaseen, Z., Ali, Z., Salih, S., & Al-Ansari, N. (2020). Prediction of Risk Delay in Construction Projects Using a Hybrid Artificial Intelligence Model. Sustainability. https://doi.org/10.3390/su12041514
56. Zaneldin, E., & Ahmed, W. (2024). A Generic Framework for Managing Schedule and Cost Risks of Construction Activities Using PERT and the EV Technique. Buildings. https://doi.org/10.3390/buildings14071918
57. Zhan, J., Fang, W., Love, P., & Luo, H. (2024). Explainable Artificial Intelligence: Counterfactual Explanations for Risk-Based Decision-Making in Construction. IEEE Transactions on Engineering Management, 71, 10667-10685. https://doi.org/10.1109/tem.2023.3325951
58. Zhang, S., & Wang, X. (2023). Quantifying Schedule Delay Risk in Construction Projects: A Data-Driven Approach with BIM and Probabilistic Reliability Analysis. Advances in Civil Engineering. https://doi.org/10.1155/2023/5525655
59. Zhong, Y., Li, H., & Chen, L. (2021). Construction project risk prediction model based on EW-FAHP and one dimensional convolution neural network. PLoS ONE, 16. https://doi.org/10.1371/journal.pone.0246539
60. Zhou, X., Shen, S., & Zhou, A. (2024). A novel data-driven approach for proactive risk assessment in shield tunnel construction. Transportation Geotechnics. https://doi.org/10.1016/j.trgeo.2024.101466
61. Zhu, F., Hu, H., & Xu, F. (2022). Risk assessment model for international construction projects considering risk interdependence using the DEMATEL method. PLoS ONE, 17. https://doi.org/10.1371/journal.pone.0265972