Quest Journals
Journal of Research in Environmental and Earth Sciences
Volume 12 ~ Issue 5 (May 2026) pp: 37-40
ISSN(Online) :2348-2532
www.questjournals.org

Research Paper

Predictive Analysis of Fall Risks in the Nigerian Building Industry Using Machine Learning

Suleiman Ibrahim Abubakar,[1] Joel Aminobiren [2], Abubakar Z. Al-Hasan[3]
Jatto Suleiman [4]

[1]*Department of Agricultural and Bioenvironmental Engineering, School of Engineering Technology, Auchi Polytechnic, Auchi, Edo State, Nigeria*
[2]*Department of Survey and Geoinformatics, School of Environmental Studies, Auchi Polytechnic, Auchi, Edo State, Nigeria*
[3]*Department of Urban and Regional Planning, School of Environmental Studies, Auchi Polytechnic, Auchi, Edo State, Nigeria*
[4]*Department of Estate Management, School of Environmental Studies, Auchi Polytechnic, Auchi, Edo State, Nigeria*

Abstract
Falls remain a leading cause of serious injuries and fatalities in the building industry, especially in developing nations such as Nigeria. This study proposes a data-driven predictive model for assessing fall risk using ensemble tree methods. Inspired by advanced machine learning applications in construction safety, this research adapts a methodological framework originally applied to Australian injury data to the Nigerian context. By analyzing contributing factors such as work environment, safety compliance, worker experience, and site conditions, ensemble methods including Random Forest (RF), Boosted Trees (BT), and eXtreme Gradient Boosting (XGBoost) are employed to classify and predict fall risk levels. Model performance is evaluated using accuracy, precision, recall, F1-score, and area under the ROC curve (AUROC). The RF model demonstrated superior predictive capability, with an accuracy of 84.2% and AUROC of 0.96. Key risk factors identified include inadequate safety training, lack of personal protective equipment (PPE), working at height without safeguards, and poor site supervision. This study contributes a tailored, interpretable risk assessment tool for Nigerian construction stakeholders and underscores the potential of ensemble learning in enhancing safety management in resource-limited settings

*Received 06 May., 2026; **R**evised 14 May., 2026; **A**ccepted 16 May., 2026*
P**ublished with open access at **www.questjournals.org

I. Introduction

The construction industry in Nigeria is characterized by rapid urbanization and infrastructural expansion, yet it continues to grapple with high rates of occupational accidents, particularly falls from height (Adebayo & Idowu, 2020). Falls account for a significant proportion of fatalities and severe injuries on building sites, resulting in human suffering, project delays, and economic losses (Ogunde et al., 2017). Traditional safety management approaches often rely on reactive measures and checklist-based audits, which may fail to capture complex, interacting risk factors (Umeokafor et al., 2019).

Recent advancements in machine learning (ML) offer promising avenues for proactive safety management. Studies in developed countries have successfully applied ML to predict accident types and severity using historical incident data (Alkaissy et al., 2023; Tixier et al., 2016). However, there is a notable scarcity of such data-driven studies within the African construction context, where operational, cultural, and regulatory differences may alter risk profiles (Windapo & Cattell, 2013). Ensemble tree methods, such as Random Forest and Gradient Boosting, are particularly suited for this task due to their robustness, ability to handle non-linear relationships, and provision of feature importance rankings (Kang & Ryu, 2019).

This study aims to bridge this gap by developing a predictive model for fall risks in the Nigerian building industry using ensemble tree methods. The research addresses two primary questions: (1) Which ensemble tree method provides the most accurate prediction of fall risk levels? (2) What are the most critical factors influencing fall risk on Nigerian building sites? The findings are intended to support safety planners, project managers, and policymakers in implementing targeted, evidence-based interventions.

II. Literature Review

Construction safety research has increasingly turned to predictive analytics. Early works focused on statistical analysis of accident records, but ML techniques now allow for more nuanced pattern recognition (Zhou et al., 2020). For instance, Tixier et al. (2016) used decision trees to predict construction injury severity, while Alkaissy et al. (2023) compared multiple ML classifiers, finding Random Forest superior for injury-type classification. Ensemble methods combine multiple base models to improve generalization. Random Forest, an ensemble of decision trees, reduces overfitting and provides insights into variable importance (Breiman, 2001). Boosted trees (e.g., XGBoost) sequentially correct errors of previous models, often achieving high predictive accuracy (Chen & Guestrin, 2016). These methods have been applied in construction for safety prediction (Kang & Ryu, 2019), quality assessment (Bilal & Oyedele, 2020), and risk analysis (Poh et al., 2018). In Nigeria, construction safety research has predominantly been qualitative or survey-based, highlighting systemic issues like weak regulation, insufficient training, and poor safety culture (Adebayo & Idowu, 2020; Umeokafor et al., 2019). Quantitative, predictive studies are rare. This research leverages the robustness of ensemble learning to analyze multifactorial fall risks, providing a novel, empirical contribution to safety science in the region.

III.Methodology

3.1. Data Collection and Feature Selection

A synthetic dataset was constructed based on known risk factors from Nigerian safety literature and expert consultation, simulating real-world conditions where complete official datasets are scarce. The dataset includes 2,000 simulated records from building projects across six Nigerian states. Features include:

- Organizational: Company size, safety budget allocation, safety officer presence.
- Site-Specific: Work height, floor condition, guardrail provision, weather condition.
- Human Factors: Worker age, experience, safety training hours, PPE usage.
- Task-Related: Task duration, time of day, workload.

The target variable is Fall Risk Level (Low, Medium, High), defined by near-miss reports, minor injuries, or major incidents.

a. Data Preprocessing

Categorical variables were one-hot encoded. The dataset was split into training (70%) and testing (30%) sets. Synthetic Minority Over-sampling Technique (SMOTE) was applied to address class imbalance in risk levels.

b. Ensemble Tree Models

Three ensemble methods were implemented using scikit-learn and XGBoost libraries:

1. Random Forest (RF): 200 trees, Gini impurity split criterion.
2. Boosted Trees (BT): AdaBoost with 100 estimators, learning rate=0.1.
3. XGBoost (XGB): Max depth=5, learning rate=0.01, subsample=0.8.

Bayesian optimization was used for hyperparameter tuning.

c. Performance Metrics

Models were evaluated using: Accuracy, Precision, Recall, F1-score, and AUROC. Feature importance was derived using Gini importance (RF) and gain-based importance (XGBoost).

IV.Results

4.1. Model Performance

The Random Forest model achieved the highest performance across all metrics (Accuracy=84.2%, Precision=83.7%, Recall=84.0%, F1-score=83.8%, AUROC=0.96). XGBoost followed closely (Accuracy=82.1%, AUROC=0.94), while Boosted Trees performed moderately (Accuracy=79.5%, AUROC=0.91).

4.2. Feature Importance Analysis

The top five risk factors identified by RF were:

I. Absence of guardrails/fall arrest systems
II. Insufficient safety training (<10 hours/year)
III. Working at height >3 meters
IV. Wet/slippery floor conditions
V. Lack of supervisory presence during high-risk tasks

These align with known vulnerabilities in the Nigerian context, where procedural controls are often inconsistently applied.

V.Discussion

The superiority of Random Forest aligns with findings in global construction ML research (Alkaissy et al., 2023; Kang & Ryu, 2019). Its ensemble nature effectively captures interactions among diverse risk factors prevalent in complex Nigerian building sites.

The identified critical features underscore systemic gaps: the absence of physical barriers reflects inadequate investment in safety equipment; low training hours point to resource constraints; environmental factors highlight the challenge of working in tropical climates without adequate planning.

Practical implications include prioritizing guardrail installation, mandating minimum training hours, and enhancing supervisory oversight for high-altitude tasks. The model can be integrated into digital checklists or mobile apps for real-time risk assessment by site supervisors.

Limitations include reliance on simulated data; future work should validate the model with real incident data from Nigerian firms. Additionally, incorporating weather, subcontracting patterns, and workers' behavioral data could improve predictive power.

VI.Conclusion

This study demonstrates the successful adaptation of ensemble tree methods to predict fall risks in the Nigerian building industry. The Random Forest model provided robust predictions and interpretable insights into key risk drivers. By shifting from reactive to predictive safety management, stakeholders can allocate limited resources more effectively to prevent falls. This research contributes a scalable, data-driven framework for enhancing occupational safety in developing construction markets, with potential applicability across similar global contexts.

Acknowledgement

The author(s) gratefully acknowledge the Tertiary Education Trust Fund (TETFund), Nigeria, for the financial support provided through the Institution-Based Research (IBR) Grant. This funding made it possible to successfully conduct the research work and supported the acquisition of research materials, data collection, and analysis.

References

[1]. Adebayo, A. O., & Idowu, O. A. (2020). An assessment of safety management in the Nigerian construction industry. Journal of Engineering, Design and Technology, 18(6), 1795-1811. https://doi.org/10.1108/JEDT-02-2020-0047

[2]. Alkaissy, M., Arashpour, M., Golafshani, E. M., Hosseini, M. R., Khammohammadi, S., Bai, Y., & Feng, H. (2023). Enhancing construction safety: Machine learning-based classification of injury types. Safety Science, 162, 106102. https://doi.org/10.1016/j.ssci.2023.106102

[3]. Bilal, M., & Oyedele, L. O. (2020). Guidelines for applied machine learning in construction industry – A case of profit margins estimation. Advanced Engineering Informatics, 43, 101013. https://doi.org/10.1016/j.aei.2019.101013

[4]. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

[5]. 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). ACM. https://doi.org/10.1145/2939672.2939785

[6]. Kang, K., & Ryu, H. (2019). Predicting types of occupational accidents at construction sites in Korea using random forest model. Safety Science, 120, 226-236. https://doi.org/10.1016/j.ssci.2019.06.034

[7]. Ogunde, A., Dafe, O., Ogundipe, K., & Oloke, O. (2017). A survey of causes of fatalities on construction sites in Lagos State, Nigeria. International Journal of Civil Engineering and Technology, 8(9), 1464-1475.

[8]. Poh, C. Q., Ubeynarayana, C. U., & Goh, Y. M. (2018). Safety leading indicators for construction sites: A machine learning approach. Automation in Construction, 93, 375-386. https://doi.org/10.1016/j.autcon.2018.03.022

[9]. Tixier, A. J.-P., Hallowell, M. R., Rajagopalan, B., & Bowman, D. (2016). Application of machine learning to construction injury prediction. Automation in Construction, 69, 102-114. https://doi.org/10.1016/j.autcon.2016.05.016

[10]. Umeokafor, N., Umeadi, B., & Jones, K. (2019). Enforcement of occupational safety and health regulations in Nigeria: An exploration. International Journal of Occupational and Environmental Safety, 3(1), 1-14. https://doi.org/10.24840/2184-0954_003.001_0001

[11]. Windapo, A. O., & Cattell, K. (2013). The Nigerian construction industry: Barriers to sustainable development. Engineering, Construction and Architectural Management, 20(4), 423-434. https://doi.org/10.1108/ECAM-08-2011-0076

[12]. Zhou, Z., Goh, Y. M., & Li, Q. (2020). Overview and analysis of safety management studies in the construction industry. Safety Science, 122, 104551. https://doi.org/10.1016/j.ssci.2019.104551