Random Forest and Path Diagram Taxonomies of Risks Influencing Higher Education Construction Projects

Olufisayo Adedokun [1], Temitope Egbelakin [1], Temitope Omotayo[2]

School of Architecture and Built Environment. The University of Newcastle, Australia NSW 2308.

School of Built Environment, Engineering and Computing, Leeds Beckett University, LS2 8AG. United Kingdom

Abstract

While risk factors are sine qua non for construction projects' non-performance, the research efforts are directed toward the likelihood of risks at the detriment of their level of influence on higher education building projects. This study assessed the perceptions of construction key stakeholders about the influence of risk factors on higher education building projects using machine learning-based random forest classification. A questionnaire survey was administered to four hundred and sixty-five (465) respondents comprising clients' representatives, consultants, and contractors across five (5) higher education institutions in Nigeria. Of 465 questionnaires, 295 retrieved were suitable for the analysis implying a 63.44% response rate. The Random Forest (RF) classification used 295 samples, out of which 189 (64%) formed the training dataset, while the validation and testing data sets are 47 (16%) and 59 (20%), respectively. The RF model accuracy conducted shows the optimized model with the test accuracy and out-of-bag accuracy (OOB). The study clustered 58 risk factors into four comprising (i) security, access, health, and safety risks, (ii) construction dispute resolution risks, (iii) construction planning and contract documentation risks, and (iv) construction cost and management risks. Further, the proposed recommendations could help enhance the performance of higher education building projects.

Keywords: Construction projects, Nigeria, random forest classification, performance, risk factors.

1. Introduction

The construction sector is essential in any nation's economic settings. For instance, it is used as a policy driver in Australia because it contributes about 8% to the GDP (Infrastructure Australia, 2019; Kadir *et al.*, 2022). Similarly, the industry plays a formidable role in driving economic advancement in developing nations by contributing significantly to the improved performance of the economy (Chileshe & Boadua Yirenkyi-Fianko, 2012; Shehu *et al.*, 2014). Furthermore, the Nigerian construction industry contributes to the GDP, having both forward and backward linkages with other sectors as it brings about infrastructures that improve the quality of life (Adedokun *et al.*, 2019; Kadir *et al.*, 2022; Oke *et al.*, 2016). However, despite the contributions of this sector to the economy, its shortcoming cannot be overemphasized. First, one of the most notable qualities of construction projects is the long implementation period and various project phases involved in the delivery of the project which could result in changes of circumstances. This raises the level of uncertainty and the probability of risks, which could have a negative influence on the construction projects (Alklkali, 2022). Second, the success of these projects becomes evident when they are completed to the terms of the contract such as the completion period, budget, quality, and end-users satisfaction among others. However, evidence shows that it is somewhat difficult to see projects meeting these performance expectations. More often, construction projects are susceptible to budget and schedule overruns (Lapidus *et al.*, 2022a; Sepasgozar *et al.*, 2022).

The delivery of construction projects earlier than the contract period is challenging (Adedokun *et al.*, 2021a; Shendurkar *et al.*, 2021). However, the challenges are occasioned by the risks that plagued the projects during both the design and construction stages (Nketekete *et al.*, 2017). This explains why the construction projects are far behind the expected level of success. Risk is the potential cause of problems and complications concerning completing and achieving a project goal. Further, delayed payments, inflation, variation, high competition bids, and progress delay are some of the risks identified (Adedokun *et al.*, 2021a). Moreover, in the African countries, the construction projects are currently witnessing a significant failure rate up to 90% consequent upon the inherent risks (Nketekete *et al.*, 2017). Similarly, the indian's construction sector recorded a growth of 5.6% between 2016 – 2020 when compared with the 2.9% growth during the 2011 – 2015 period. Further, 215 of 762 construction projects in 2017 experienced delays ranging between 1 to 261 months, leading to an increased estimated cost (Shendurkar *et al.*, 2021). Therefore, cost and time overruns remains the major risks that impact on projects (Sepasgozar *et al.*, 2022).

While the building projects in higher education institutions are not exempted from the influence of risks; there exist higher discrepancies between the initial tender sum and final contract sum including the initial and final contract period than in other building projects (Sepasgozar *et al.*, 2022). Further, most of the HEBP are delivered an at increased contract sum, beyond the contract completion period, and with a poor quality of work done (Raamkumar & Indhu, 2021; Shendurkar *et al.*, 2021). According to Aladağ and Işık (2019), there are several sources from which risks could occur while executing these construction projects; however, the extents of these risks are not defined at times. Regardless of the undefined extent of risks, the ultimate goal of stakeholders on HEBP is about delivering the project within the established period (time), cost (budget), and quality (Laryea & Hughes, 2011; Ofori, 2007). While the poor performance of HEBP is occasioned by the risks, risk occurrence does not translate to its influence on projects. So, the influence of these risks on HEBP becomes imperative. Therefore, the study aims to evaluate the perceptions of practitioners on the influence of risks on the performance of higher education building projects (HEBP) through machine learning-based random forest classification.

2. Literature review

Several risk factors are associated with building projects. These risks manifest in various phases of the projects life cycle commencing from the feasibility stage through completion, handing over and maintenance (Bakri *et al.*, 2021). Further, in recent years of active construction of building projects, the scope of work, the timing and limited finance paved way for the emergence of risks influencing the budget and project's duration (Lapidus *et al.*, 2022b). For instance, construction risks and financial risks are the most frequently occurring while servicing relocation, design changes, additional works, and price fluctuation influence cost risks in the Malaysian construction industry. Similarly, the main factors influencing the schedule of the project are permit approval, the condition of the weather, reduced workers' productivity, untimely preparation and approval of drawings, changes to the design, and worker shortage (Bakri *et al.*, 2021). Moreover, the construction industry is plagued by risks thereby subjecting it to an age-long poor track record in comparison with other economy sectors (Mishra & Mishra, 2016). Furthermore, risks cannot be ignored but could be minimised, shared, managed, accepted or transferred to a third party who has the capacity for it (Renuka *et al.*, 2014). While acknowledging the extensive studies on the construction project risk management, they are number of significant risks impacting the realization of these projects up to date (Hanafi *et al.*, 2021; Mahmoud *et al.*, 2021a). Therefore, successful management of these risks becomes important, and this necessitated the need to evaluate the influence of risks on HEBP.

The performance of construction projects has been affected by construction risks leading to seemingly low performance, and resultant negative effect on the national development plan (Hanafi *et al.*, 2021). Moreover, some of the reasons advanced in lthe iterature include complexity of the construction projects, huge initial capital outlay and diverse resources includingthe labour (Hanafi *et al.*, 2021). The common construction risks during the construction phase include land acquisition and compensation from the Malaysian context (Bakri *et al.*, 2021). However, this varies from Szymański (2017) where progress delay, changes in the design, and discrepancies between the bill and actual quantities of work as built on site are construction risks. While risks are location-specific based on the aforementioned (Banaitis & Banaitiene, 2012), these risks influence the construction period and cost of the project. Another important risk factor is financial risk factors which are derived from the inflation and interest rate (Zhang *et al.*, 2021). The risk of interest rate could affect the construction project if the anticipated cost is lower when compared with the actual cost of financing (Bakri *et al.*, 2021). Further, the operational stage of the construction projects is not uncommon with inflation risk which mostly affects the materials and equipment. Therefore, the resultant effect of financial risks becomes evident on the project in terms of budget overrun. The financial risks increase when the capital investment of the project increases (Mahmoud *et al.*, 2021b). So, the chances of failure could increase as a result of higher risks that hinder project performance. In addition, the financial risk is an unsystematic risk that is germane to insurance and construction industries because it’s organisation-specific and as such, could be controlled (Ghaffari, 2013).

The survey undertaken by Nguyen *et al.* (2021) shows that the critical risks influencing building projects during the execution stage are inseparable from the projects. However, while the factors fell into financial, management, schedule, construction, environmental risks, the most influential factors are material cutting, construction ground problems, and design errors. In contrast, following a systematic literature review of PPP projects, out of 37 articles that were analyzed, 16 articles reported political risks as the critical risk factor affecting PPP contract (Tetteh *et al.*, 2020). Ling and Hoang (2010) described changes in the business environment occasioned by political changes as political risks. However, the effect of these political risks could either be macro or micro in nature. The effect of political risks is macro when it affects all businesses, whereas if it is affecting only a selected industry, for instance construction sector, firms or projects it’s said to be a micro (Ling & Hoang, 2010). Therefore, changes in the business environment could affect project performance to constitute a political risk and varies from one company to the other (Adedokun *et al.*, 2021a). Some potential risks - management risks, influence construction projects such that without the top management’s approval, an informed decision could not be made. Most organisations usually face a dilemma when it comes to making a bid or no bid decision (Odimabo & Oduoza, 2013). Moreover, it’s a decision about increasing the markup high enough to make more profit and not getting the job or reducing the markup to win the contract in a competitive tender while risking the possible loss from the contract (Laryea, 2011; Stanley, 2011). Others include managing the site restriction, restriction in hours of work, and the expectations from owners regarding higher quality above the standard which could be potential post-contract risks awaiting the contractor as the work progresses.

The delays or poor performance of construction projects are also traceable to environmental risks in the form of disruptions and pollution (Rahman & Esa, 2014). According to Akinbile *et al.* (2018), environmental risks become imminent when construction projects could negatively impact or influence the immediate environment outside the construction sites thereby giving rise to financial and legal issues. However, construction management should mitigate risk through risk assessment of environmental factors to safeguard the likelihood of problems during the delivery phase of the projects (Aladağ & Işık, 2019; Ansah *et al.*, 2016). Further, the sources from which external risk variables could impact construction projects are not within the control of project management including social-cultural, political, economic, legal and environmental (Ansah *et al.*, 2016). So, considering the dangers of the physical environment i.e external environment, risk analysis and management are essential. In addition, several challenges like poor construction, scope creep, and problems in determining the main risk sources are often faced by the design professionals (Sutrisna & Goulding, 2019), this could lead to design risks when the design is not well coordinated, lowering work quality or rush design (Odimabo & Oduoza, 2013). With the varying risks affecting projects in the construction sector, there is a need to evaluate the influence of risks on higher education building projects using a more robust methodology.

3. Research Method

The quantitative method, underpinned by a positivist epistemology, was used in assessing the risk influence on higher education building projects (HEBP). The method ensures objectivity and generalisation of the study's findings. However, a structured questionnaire was adopted to not only elicit information from the respondents but also enable a uniform basis of response to the questions. The question in the instrument was framed as "kindly rate your perceived level of influence of the underlisted risks on HEBP" using a 5-point Likert scale, where 1 is the lowest and 5 represents the highest value. The participants included were the stakeholders involved in the completed building projects from the higher education institutions. The lists for these participants were sourced from each of the institution's physical planning units comprising the consultants, contractors, and clients' representatives. The study relied on the HEBP that was procured using the traditional procurement method of a contract between 2000 and 2022. While the stakeholders' information earlier than 2000 was not available from some institutions, therefore, the study was limited to the year 2000. Further, the data set used for the study accounted for stakeholders on completed HEBP as of second quarter of 2022. Moreover, respondents that were commissioned for more than one project in the same institutions were carefully identified. After identification, the excess appearances were removed from the population of 512. This was done to ensure a uniform basis of representation thereby giving rise to a target population of 465. In addition, the study used the random forest classification method, a machine learning algorithm to categorise the predictive variables of risk management in HEBP. The ranking shows the order in which the identified risk factors that influence higher education building projects.

3.1 Profile of respondents

Four hundred and sixty-five questionnaires were administered to the participants comprising consultants, contractors, and client representatives. Of the administered questionnaires, 295 questionnaires (63.44%) received were deemed fit for further analysis and considered sufficient (Moser & Kalton, 2017). Table 1 shows the respondents' background information, indicating the organization where they belong, their profession and years of experience. Most respondents are from consulting firms accounting for 42.70% (Table 1). Moreover, 33.90% of the respondents are from contracting firms, while the remaining group representing 23.40% are from the client's organisation. Further analysis shows that 34.60% of the respondents are Quantity Surveyors, 14.60% are Architects, and 18% represent Builders. The Engineers account for 32.90% of the total comprising Structural/Civil Engineers (16.60%), Electrical Engineers 9.20%, and Mechanical Engineers (5.10%). The respondents had 13years of work experience on average. The high years of experience possessed by these respondents substantiated the adequacy and reliability of the information supplied to them.

Table 1: Background information of the respondent

Category	Classification	Frequency	Per cent
Type of Organization	Client organisation	69	23.40
	Contracting firm	100	33.90
	Consulting firms	126	42.70
	Total	**295**	**100.00**
Profession Of Respondents	Quantity Surveying	102	34.60
	Architecture	43	14.60
	Building	53	18.00
	Structural/Civil Engineering	55	18.60
	Electrical Engineering	27	9.20
	Mechanical Engineering	15	5.10
	Total	**295**	**100.00**
Years Of Working Experience	1 – 5	46	15.60
	6 – 10	72	24.40
	11 – 15	68	23.10
	16 – 20	66	22.40
	Above 21	43	14.60
Mean	**12.51** **Total**	**295**	**100.00**

Table 2: List of abbreviations

S/N	Acronyms	Full meaning
1	LeF1	Site accidents
2	LeF2	Supplies of defective materials
3	LeF3	Shortage of labour
4	LeF4	Environmental factors (flood, earthquake etc.)
5	LeF5	Difficulty to access the site (very far, settlements)
6	LeF6	Adverse weather conditions
7	LeF7	Pollution
8	LeF8	Defective design (incorrect)
9	LeF9	Not coordinated design (structural, mechanical, electrical etc.)
10	LeF10	Inaccurate quantities

11	LeF11	Discrepancies between BOQ, drawings and specifications
12	LeF12	Rush design
13	LeF13	Unavailable labour, materials, and equipment
14	LeF14	Undefined scope of working
15	LeF15	High competition in bids
16	LeF16	Inaccurate project program
17	LeF17	Poor communication between the home and field officers (contractor side)
18	LeF18	Materials monopoly
19	LeF19	Inflation
20	LeF20	Delayed payments in contracts
21	LeF21	The financial failure of the contractor
22	LeF22	Unmanaged cashflow
23	LeF23	Exchange rate fluctuation
24	LeF24	Rate of interest
25	LeF25	Invoices delay
26	LeF26	Difficulty to permit/licenses
27	LeF27	Ambiguity to work legislations
28	LeF28	Legal disputes during the construction phase among the parties of the contract
29	LeF29	Delayed dispute resolutions
30	LeF30	No special arbitrators to help settle fast
31	LeF31	Rush bidding
32	LeF32	Gaps between the implementation and the specification due to misunderstanding and specification
33	LeF33	Undocumented change orders
34	LeF34	Lower quality work
35	LeF35	Design changes
36	LeF36	Actual quantities differ from the contract quantities
37	LeF37	Working hours restrictions
38	LeF38	On-site congestion
39	LeF39	Deficient and insufficient safety rules
40	LeF40	Differing site conditions
41	LeF41	Owner’s high expectations for quality beyond standards
42	LeF42	Delay or inability of the owner to give full possession
43	LeF43	Delay in progress payments
44	LeF44	Delay in the start of the project
45	LeF45	Segmentation of construction process
46	LeF46	Working in hot (dangerous) areas
47	LeF47	New governmental acts or legislations
48	LeF48	Unstable security circumstances (invasion)
49	LeF49	Closure
50	LeF50	Bribery and corruption
51	LeF51	Wars and revolutions

52	LeF52	Import and export restrictions
53	LeF53	Ambiguous planning due to project complexity
54	LeF54	Resource management
55	LeF55	Change in management ways
56	LeF56	Information unavailability (including uncertainty)
57	LeF57	Poor communication between involved parties
58	LeF58	Occurrence of variations

4. Analysis and discussion of findings

4.1 Random Forest Classification

Machine learning applications such as random forest (RF) classification in construction and risk management have proved to produce more accurate results compared to other statistical analyses such as factor analysis, structural equation modelling and correlation (Omotayo, Awuzie and Ayokunle, 2020). Higher levels of uncertainty and decisions are better captured with the use of machine learning decision tree classification (Roman *et al.*, 2020). RF classification produces better decision measures and applications in construction risk management research is imperative. Gislason *et al.* (2006); (Petkovic *et al.*, 2018; Xu *et al.*, 2012)corroborated RF as being a generic terminology for the ensemble method of classification using a tree-type algorithm which is trained. Hence, RF classification is a form of machine learning analysis used for classification. The algorithm of RF classifiers can be expressed as:

$$\{h(x,Hk),k=1,...,\} \qquad (1)$$

Where the {Hk} independent identically distributed random vectors and "x" is an input pattern (Gislason *et al.*, 2006), further, Azar (2014) noted that the computational time for the RF classification model is represented as:

$$\frac{T\sqrt{MN\log(N)}}{1} \qquad (2)$$

Where T is the number of trees, M is the number of variables used in each split, and N is the number of training samples (Belgiu & Dra, 2016). RF classification is widely used in remote sensing and physical sciences. In applying RF for built environment research, the classification of factors in terms of their relative significance, predictive relevance, and accuracy can be attained through the model accuracy, receiver operating curves (ROC) curves, and the total increase in node purity. Each of these features of RF classification was applied to analyse the dataset abbreviated in Table 2. The purpose of the RF classification is to categorise the risks associated with higher education construction projects under broader umbrellas of risk management typologies.

The RF classification used 295 samples, out of which 188 formed the training dataset, which accounts for 64.16% of the sample. The validation and testing data sets are 48 and 59, respectively. The three main categories of the samples viewed the respondents from their years of experience, respondent's profession, and the type of organisation. As presented in Tables 3 to 5, the RF model accuracy shows the optimised model with the test accuracy and out-of-bag accuracy (OOB).

Table 3: Random Forest Classification based on years on experience

Trees	Features per split	n(Train)	n(Validation)	n(Test)	Validation Accuracy	Test Accuracy	OOB Accuracy
98	7	188	48	59	0.896	0.847	0.897

Table 4: Random Forest Classification and the respondent's profession

Trees	Features per split	n(Train)	n(Validation)	n(Test)	Validation Accuracy	Test Accuracy	OOB Accuracy
55	7	188	48	59	0.708	0.763	0.900

Table 5: Random Forest Classification and type of organisation

Trees	Features per split	n(Train)	n(Validation)	n(Test)	Validation Accuracy	Test Accuracy	OOB Accuracy
47	7	188	48	59	0.854	0.814	0.863

The OOB measures the RF machine learning model (Azar, 2014; Gislason *et al.*, 2006; Pal, 2005). A good out of OOB should be less than 1%. Thus, when comparing the OOB accuracy of Tables 3, 4 and 5, Table 5, RF and type of organisation produces an OOB of 0.863 based on a test accuracy of 0.927. Compared with the OOB or Tables 3, years of experience, 0.897 and Table 4, respondent's profession, 0.900. The RF model of Table 3 produces more significant outcomes with 98 trees and the highest test accuracy of 0.847 (84.7%). Thus, the years of experience of the respondents forms a significant stake in the overall analysis. Further analysis using the RF classification applied the ROC plots to further classify the specific important predictive variables from the trio mentioned above of years of experience, respondent's profession, and type of organisation of the participants.

4.2 Receiver operating curves (ROC) Plot

The receiver operating curves (ROC) plot is a graphical representation of the outcomes between true positive rate (TPR) and false-positive rate (FPR) (Carter *et al.*, 2016). ROC curves are created from a diagnostic test of a dual classifier system known as a discriminatory threshold (Fan *et al.*, 2006). The TPR is also known as the sensitivity or recall, and it is denoted as:

$$TPR = \frac{TP}{\mathrm{P}} = \frac{TP}{\mathrm{TP+FN}} = 1 - FN \qquad (3)$$

TP is truly positive, FN is a false negative, and P is positive. The TPR measures the positive outcomes in a sample after dividing the addition of true positives and false negatives. The FPR, also called the fall-out, is represented as:

$$FPR = \frac{FP}{\mathrm{P}} = \frac{FP}{\mathrm{FP+TN}} = 1 - TFN \quad (4)$$

FP is the number of false positives, and TFN is the true false negatives. The FPR is used to test the level of significance based on inference. The FPR is calculated by checking the number of negative results categorised as positive and the number of negative results (Fan *et al.*, 2006). The ROC plot produced in this study used the binary comparison of TPR against FPR for the respondents' years of experience, respondent's profession, and type of organisation of each participant. The area under the curve (AUC) values are presented in the Tables included in Appendix A, B and C.

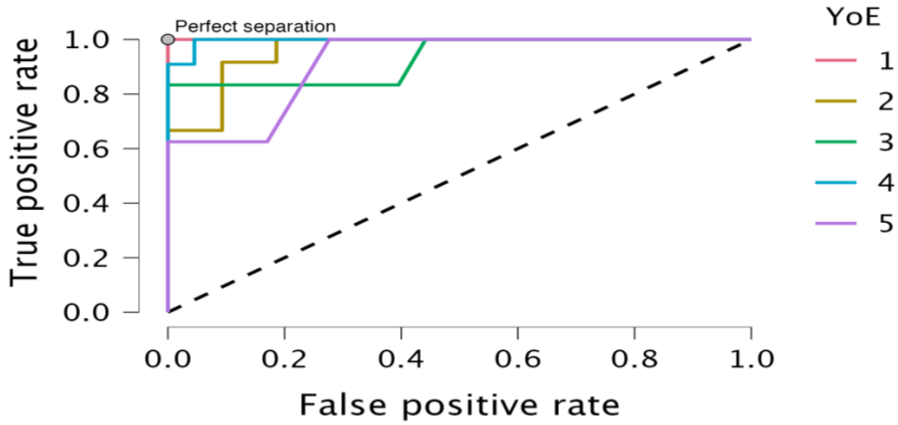

Figure 1. ROC plot for years of experience of the participants.

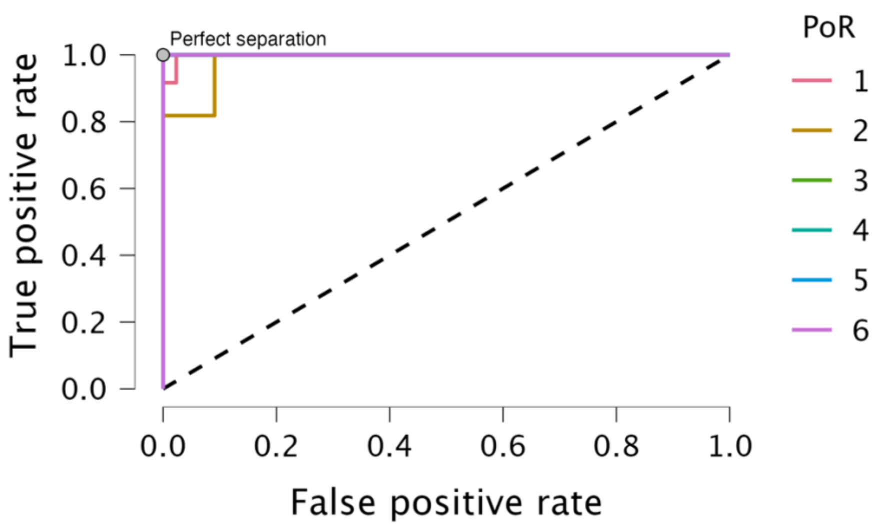

Figure 2. ROC plot for respondent's profession

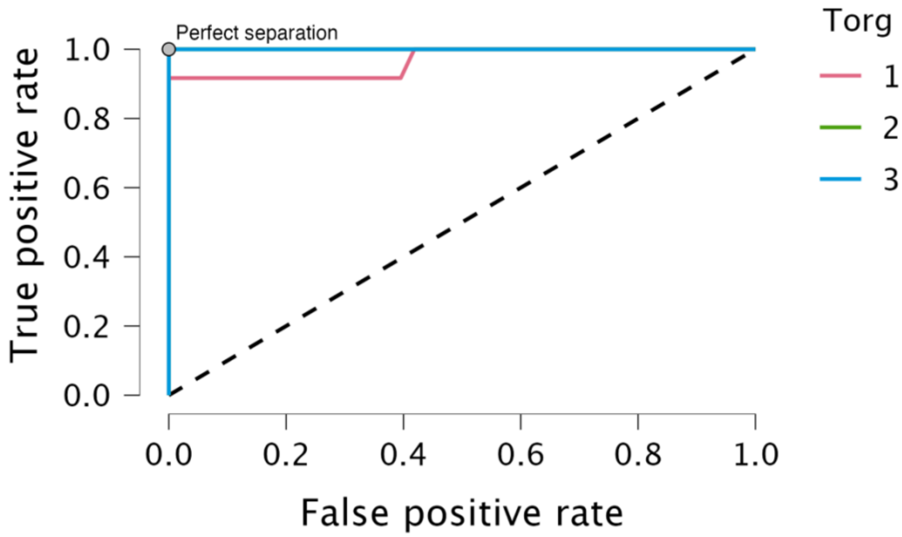

Figure 3. ROC plot for the type of organisation

From Figure 1, respondents who fall into the category of 0-5 years (denoted as 1 in red) produced more risk predictions than occur in the sample. This may be due to their low experience in higher education construction projects. 0-5 years produced a perfect ROC TPR of 1.0. and FPR of 0. The implication of this result compared to 4 (16-20 years) and 5 (Above 20 years) with TPR of 1.0 but FPR of 0.1 and above shows that the factors elucidated in Table 2 will depend more on respondents with 0.5 years of experience. Figure 2 illustrates the ROC plot for the participants' professions involved in this study. The perfect ROC with an AUC of 1.0 is 6, with mechanical engineers predicting more risks due to collaboration, external works, and MEP experience level in higher education construction projects. The other categories, such as 3 (building), 4 (structural/civil engineers) and 5 (electrical engineers), are not represented in the analysis because the AUC for 1.0 was produced in the ROC plot. Hence, 3, 4, 5 and 6 predict the most important risk variables. Figure 3 shows number 3, consulting firms, as the perfect AUC with a value of 0.992. This implies that consulting firms contribute more risks through the planning phases of HEB projects in Nigeria. The risk contributing variables highlighted in Table 2 were analysed using the total increase in node purity of the RF classification trees. This output is explained in the next section.

4.3 Total increase in node purity

Variable importance in RF classification can be calculated using the node purity of each variable. An increase in node purity is calculated using the reduction in the sum of squares errors when the variable is split. This total increase in node purity is similar to Gini-based and can rank variables as indicated in Table 6. There are 58 variables under the categories of years of experience, profession, and type of organisation, all contained in 2a 79 sample size. The mean total increase in node purity was computed for each variable and ranked in Table 6.

Table 6: Variable importance according to the average increase in node purity

Variables	**The total increase in node purity**			Mean	Rank
	Years of exp.	Profession	Type of org.		
LeF22	0.054	0.017	0.013	0.028	3rd
LeF10	0.040	0.025	0.015	0.027	4th

LeF4	0.039	0.007	0.009	0.018	8th
LeF51	0.038	0.009	0.017	0.021	5th
LeF12	0.036	0.006	0.015	0.019	7th
LeF31	0.034	0.036	0.025	0.032	2nd
LeF13	0.027	0.008	0.007	0.014	20th
LeF16	0.027	0.006	0.009	0.014	21st
LeF55	0.026	0.011	0.014	0.017	12th
LeF38	0.025	0.009	0.007	0.014	22nd
LeF9	0.022	0.002	0.018	0.014	23rd
LeF50	0.022	0.006	0.022	0.017	13th
LeF48	0.022	0.010	0.017	0.016	16th
LeF44	0.021	0.008	0.021	0.017	14th
LeF17	0.021	0.013	0.009	0.014	24th
LeF52	0.021	0.014	0.012	0.016	18th
LeF29	0.019	0.012	0.006	0.012	29th
LeF40	0.019	0.014	0.021	0.018	10th
LeF56	0.019	0.023	0.003	0.015	19th
LeF36	0.018	0.019	0.026	0.021	6th
LeF7	0.017	0.012	0.012	0.014	25th
LeF1	0.016	0.006	0.009	0.010	39th
LeF26	0.016	0.012	0.007	0.012	31st
LeF11	0.016	0.008	0.007	0.010	41st
LeF14	0.016	0.013	0.011	0.013	28th
LeF42	0.016	0.014	0.004	0.011	35th
LeF19	0.015	0.017	0.018	0.017	15th
LeF46	0.014	0.011	0.013	0.013	30th
LeF8	0.013	0.01	0.018	0.014	26th
LeF57	0.013	0.023	0.009	0.015	16th
LeF21	0.013	0.005	0.009	0.009	45th
LeF27	0.013	0.009	0.015	0.012	32nd
LeF2	0.012	0.008	0.003	0.008	51st
LeF39	0.012	0.14	0.026	0.059	1st
LeF53	0.012	0.010	0.005	0.009	46th
LeF23	0.011	0.007	0.007	0.008	53rd
LeF47	0.011	0.006	0.017	0.011	36th
LeF34	0.011	0.014	0.028	0.018	9th
LeF32	0.011	0.013	0.010	0.011	37th
LeF30	0.010	0.007	0.020	0.012	33rd

LeF3	0.010	0.013	0.009	0.011	38th
LeF24	0.010	0.008	0.013	0.010	43rd
LeF25	0.009	0.012	0.009	0.010	44th
LeF37	0.009	0.010	0.007	0.009	47th
LeF49	0.009	0.014	0.013	0.012	34th
LeF54	0.009	0.006	0.014	0.010	43rd
LeF58	0.008	0.006	0.018	0.011	40th
LeF43	0.008	0.010	0.010	0.009	48th
LeF33	0.008	0.008	0.004	0.007	53rd
LeF18	0.006	0.026	0.013	0.015	17th
LeF6	0.006	0.006	0.008	0.007	55th
LeF5	0.005	0.019	0.007	0.010	49th
LeF41	0.005	0.005	0.016	0.009	50th
LeF20	0.004	0.01	0.006	0.007	56th
LeF45	0.002	0.008	0.007	0.006	57th
LeF28	0.001	0.009	0.006	0.005	58th
LeF35	0.001	0.014	0.026	0.014	27th
LeF15	-0.003	0.016	0.013	0.009	52nd

The mean total increase in node purity produced LeF39 (Deficient and insufficient safety rules) as the highest risk variable with a value of 0.059, LeF31 (Rush bidding) with a 0.032 as the 2nd most important variable, and the 3rd, LeF22 (Unmanaged cashflow) (0.028). The lowest rank variable is LeF28 (Legal disputes during the construction phase among the contract parties) produced a value of 0.005. The final phase of the RF classification will lead to a path diagram showing the classification of the variable into four (4) main risk categories. The path diagram applied the component loading (as indicated in Table 7) after taking 64% of the variable importance from Table 6 above. 64% of the variables were selected because this represented the training dataset, and it amounted to 37 variables out of 58. In machine learning applications, the training datasets usually constitute a larger proportion of the dataset, in most cases 60% and above while the testing and validation samples are divided within the 40% (Omotayo et al., 2020; Carter *et al.*, 2016; Azar, 2014). The variables were compressed into RC1 to RC4 as shown in appendix F, Figure 4 and Tables 7 and 8. Further explanations of the risk categories are explained in the discussed section after a brief explanation after Table 8.

Table 7: Component Loadings

	RC1	**RC2**	**RC3**	**RC4**	**Uniqueness**
LeF51	0.907				0.276
LeF49	0.833				0.339
LeF52	0.825				0.373
LeF46	0.800				0.304

Table 7: Component Loadings

	RC1	RC2	RC3	RC4	Uniqueness
LeF48	0.799				0.399
LeF56	0.691				0.342
LeF50	0.682	-0.532			0.294
LeF47	0.554				0.370
LeF30		0.764			0.326
LeF29		0.759			0.306
LeF7		0.651			0.595
LeF26		0.591			0.405
LeF13		0.589			0.312
LeF4		0.577			0.522
LeF38		0.561			0.459
LeF27		0.525		0.617	0.245
LeF12			0.824		0.356
LeF10			0.810		0.270
LeF16			0.771		0.371
LeF9			0.644		0.305
LeF8			0.546		0.366
LeF36				0.756	0.368
LeF39				0.715	0.497
LeF31				0.593	0.389
LeF32				0.571	0.345
LeF22					0.590
LeF55					0.418
LeF44					0.651
LeF17					0.550
LeF40					0.531
LeF14					0.498
LeF19					0.764
LeF57					0.317
LeF34					0.438
LeF18					0.593
LeF42					0.462

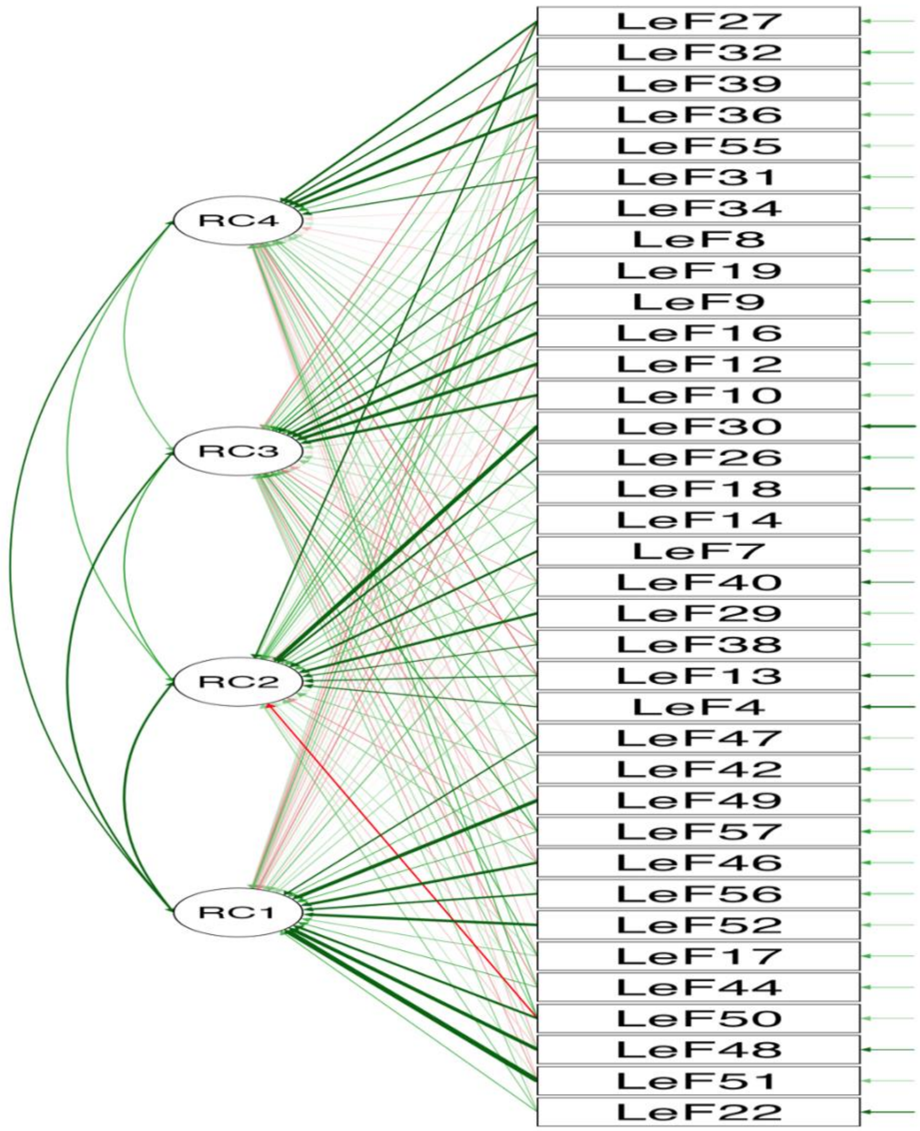

Figure 4. Path diagram of risk variables.

Table 8: Risk factors

Nr	Factor	Variables
RC1	Security, access, health and safety risks	Closure (LeF49); Working at hot (dangerous) areas (LeF46); Information unavailability (including uncertainty) (LeF56); Import and export restrictions (LeF52); Unstable security circumstances (invasion) (LeF48); Wars and revolutions (LeF51).
RC2	Construction dispute resolution risks	No special arbitrators to help settle fast (LeF30); Difficulty to permit/licenses (LeF26); Pollution (LeF7); Delayed dispute resolutions (LeF29).
RC3	Construction planning and contract documentation risks	Not coordinated design (structural, mechanical, electrical etc) (LeF9); Inaccurate project program (LeF16); Rush design (LeF12); Inaccurate quantities (LeF10)
RC4	Construction cost and management risks	Ambiguity to work legislations (LeF27); Gaps between the implementation and the specification due to misunderstanding and specification (LeF32); Deficient and insufficient safety rules (LeF39); Actual quantities differ from the contract quantities (LeF36)

The categorisation of the risks associated with HEBP in Nigeria used the dimension reduction method to produce four (4) broad themes are explained below.

4.4 Security, access, health and safety risks (RC1)

The variables contributing to security, access, health and safety risks are Closure (LeF49) with a loading; Working at hot (dangerous) areas (LeF46); Information unavailability (including uncertainty) (LeF56); Import and export restrictions (LeF52); Unstable security circumstances (invasion) (LeF48); Wars and revolutions (LeF51). The issue of security, access to construction sites and health and safety in Nigeria can culminate in increased project cost, delays and abandonment. Nigeria is plagued with kidnapping, militancy, and terrorism challenges (Badiora, 2015; Habila, 2017). Furthermore, most HEBP construction projects in Nigeria are outside the urban and suburban environment. Therefore, access roads are usually constructed to the site's location, and project security is considered an issue of construction workers' health and safety.

4.5 Construction dispute resolution risks (RC2)

Construction dispute resolution in HEBP in Nigeria is essential for forming construction project plans. Usually, the standard forms of contract used for building construction projects in Nigeria are the Joint Contract Tribunal (JCT) and the International Federation of Consulting Engineers (FIDIC). The aforementioned standard forms of the contract contain elements of dispute resolution that may be in mediation, adjudication or arbitration. This risk category was developed after the combination of factors such as no special arbitrators to help settle fast (LeF30); Difficulty to permit/licenses (LeF26); Pollution (LeF7); Delayed dispute resolutions (LeF29). Femi (2014) opined that some of the causations of disputes in the Nigerian construction industry mostly emanates from diverse perspectives, assumptions, withdrawal and miscommunication. Risks associated with dispute resolutions in Nigerian HEBP can come from the project's performance. For instance, all forms of pollution from construction sites may lead to complaints and further delays. Delays in obtaining building permits and dispute resolutions are concomitant risk elements that influence the performance of HEBP in Nigeria.

4.6 Construction planning and contract documentation risks (RC3)

In every construction project, construction planning and contract documentation risks are sacrosanct elements of risk management. The dimension reduction analysis combined the following risk variables: not collaborative design (structural, mechanical, electrical etc.) (LeF9); Inaccurate project program (LeF16); Rush design (LeF12); Inaccurate quantities (LeF10). HEBP in Nigeria is usually large and unique. Hence, mistakes in design, bills of quantities and programme of works may lead to extensive time overruns and relevant matters of additional expenses. Jayasudha and Vidivelli (2016) observed that most building construction projects with planning and documentation risks have limited awareness and application of digital construction planning tools. The use of digital tools such as Microsoft office projects for scheduling, Cost X or BlueBeam applications for measurement and bills of quantities preparation are contributory factors that may decide the influence of construction planning and contract documentation risks.

4.7 Construction cost and management risks (RC4)

Construction cost and management risks are derived from the variables of ambiguity to work legislations (LeF27); Gaps between the implementation and the specification due to misunderstanding and specification (LeF32); Deficient or insufficient safety rules (LeF39); Actual quantities differ from the contract quantities (LeF36). Construction project management depends on cost, schedule, quality, the scope of work, and available labour resources. In HEBP, the construction cost may escalate when there are deficiencies in the construction project management. The execution phase of construction projects depends on the skills of the construction project management in interpreting the contract documents. Any gap between execution and project plans will produce a major risk in the form of major or minor variations (Oladapo, 2007; Balbaa *et al.*, 2019). Construction variations, either minor or major, may be evidence of poor project management and cost planning techniques.

The nexus between the four broad themes of risks influencing HEBP in Nigeria was explored in Appendix D, where the component correlations were produced. Security, access, health and safety risks (RC1) are central to RC2, RC3, and RC4 with a correlation coefficient of 0.569, 0.548, and 0.519 individually. Consequently, risks to security, access to construction sites and health and safety must be prioritised in risk management documentation for HEBP in Nigeria.

5. Implications of findings

Regarding the risk factors influencing HEBP, the most evident are deficient or insufficient safety rules, rush bidding, and unmanaged cash flow. Nevertheless, the outcome of this study is at variance with some empirical evidence from the developing economies. For instance, Mahmoud *et al.* (2021b) advanced financial risks as influencing construction projects to comprise inflation, financial failure and delay in payment. Similarly, Bakri *et al.* (2021); Musa *et al.* (2015); (Zhang *et al.*, 2021) regarded inflation and interest rate as economic factors affecting building projects especially if the anticipated cost is lower when compared with actual cost of financing; this could be referred to as external factors because they are not within the resolve of project participants. According to Babalola *et al.* (2015), the management of external factors is indispensable in achieving successful construction projects. However, the unmanaged cashflow found in this study compares favourably with other studies where clients' cashflow-related problems, the contractor's financial incapacity, and budget overruns occasioned by delay are the most critical risk factors that affect the financial phase of project (Akinsiku & Akinsulire, 2012; Zhang *et al.*, 2021). The differential in the study's findings could be because of the methodology, random forest classification, adopted to harness the factors' relative significance, predictive relevance and accuracy. It could also be consequent upon the differences in the location of the projects as risk factors are location specific. Furthermore, the significant risk factor triggering delay along the stages of construction projects is financial risks which could emanate from the cashflow problems of the contractors (Nguyen *et al.*, 2021; Shehu *et al.*, 2014; Zhang *et al.*, 2021). While the fifty-eight risk factors were clustered into four categories using random forest classification, this is at variance with studies that clustered similar risks into eight groups (Adedokun & Egbelakin, 2022; Adedokun *et al.*, 2021b). This study focused on higher education building projects (HEBP) in a developing economy based on random forest classification of risk factors, unlike other studies like Adedokun and Egbelakin (2022); (Adedokun *et al.*, 2021b), used the structural equation modelling technique for similar building projects.

6. Conclusion

The inherent risk factors are catalysts that trigger the construction project's nonperformance. Therefore, this study evaluated the influence of these risks on higher education building projects (HEBP). Consequently, the study concludes that beyond the occurrence of risks, the risks also influence HEBP, with the consulting firms in Nigeria contributing more risks than the contracting firms and client organisations. While the assessment and clustering of risks could enhance the risk management of projects, the top three risks influencing HEBP are deficient or insufficient safety rules, rush bidding, and unmanaged cash flow using the average increase in node purity. Further, the study also compressed and categorised the fifty-eight risk factors into four clusters. The clusters were named security, access, health and safety risks, construction dispute resolution plans, construction planning and contract documentation risks, and construction cost and management risks. Regarding the higher education building projects, the consulting firms contribute more risks than the contracting firms or client organisations in Nigeria. Therefore, the consultants on the project should finalise design decisions before awarding the contract to guard against a variation that could further increase the construction cost above the budgeted amount. In addition, deficient or insufficient safety rules could lead to site accidents with cascading effects on the projects. So, the contracting organisations should be mandated to present detailed and functional safety rules, including insurance for work and workers on the projects. In addition, rush bidding on contracting organisations could lead to underpricing some work items, eroding the contractor's markup. To avert shoddy jobs, performance and bid bonds should form part of the contract documents. The higher education institutions, the client, should set aside a contingency fund to absorb and cushion the effect of inflation on building projects. The fund set aside will ensure that the project runs smoothly without hitches regarding funding the HEBP. Besides, there should be a priority for payments due to the contractor. The prompt payment could prevent work delay resulting from lack of funds and financial failure of the contractor. However, the findings are limited to HEBP that were procured through the traditional procurement method and could make the results differ while considering different procurement method. Moreover, the study adopted classification based on organization for project participants and not on ownership status based. Therefore, further research should focus risk appraisal in terms of the ownership status of the project as the order of risk influence could change due to project ownership. Finally, further studies could consider the relationship between the risks occurrence and the corresponding influence on the HEBP.

Appendices

Appendix A: Evaluation Metrics for years of experience

	Precision	Recall	F1 Score	Support	AUC
1	1.000	1.000	1.000	12	1.000
2	0.647	0.917	0.759	12	0.912
3	1.000	0.833	0.909	12	0.851
4	0.909	0.909	0.909	11	0.991
5	1.000	0.625	0.769	8	0.840
Average / Total	0.905	0.873	0.876	55	0.919

Note. Area Under Curve (AUC) is calculated for every class against all other classes.

Appendix B: Evaluation Metrics for respondent's profession

	Precision	Recall	F1 Score	Support	AUC
1	0.800	1.000	0.889	12	0.998
2	0.900	0.818	0.857	11	0.959
3	1.000	0.917	0.957	12	1.000
4	1.000	0.909	0.952	11	1.000
5	1.000	1.000	1.000	4	1.000
6	1.000	1.000	1.000	5	1.000
Average / Total	0.936	0.927	0.928	55	0.993

Note. Area Under Curve (AUC) is calculated for every class against all other classes.

Appendix C: Evaluation Metrics for the type of organisation

	Precision	Recall	F1 Score	Support	AUC
1	1.000	0.917	0.957	12	0.950
2	1.000	0.813	0.897	16	0.980
3	0.871	1.000	0.931	27	0.992
Average / Total	0.937	0.927	0.927	55	0.974

Note. Area Under Curve (AUC) is calculated for every class against all other classes.

Appendix D: Component Correlations

	RC1	RC2	RC3	RC4
RC1	1.000	0.569	0.548	0.519

Appendix D: Component Correlations

	RC1	RC2	RC3	RC4
RC2	0.569	1.000	0.458	0.389
RC3	0.548	0.458	1.000	0.326
RC4	0.519	0.389	0.326	1.000

References

Adedokun, O., & Egbelakin, T. (2022). Structural equation modelling of risk factors influencing the success of building projects. *Journal of Facilities Management*. doi:10.1108/jfm-01-2022-0002

Adedokun, O. A., Aje, I. O., Awodele, O. A., & Egbelakin, T. (2021a). Are the Public Tertiary Education Building Projects Devoid of Risks in the Nigerian Construction Context? *Journal of Financial Management of Property and Construction*. doi:10.1108/JFMPC-01-2021-0002

Adedokun, O. A., Akinmusire, A. O., & Abiola-Ogedengbe, D. (2019). Budget Overruns Experienced on Tertiary Institutional Building Projects – Recourse to the Contractors' Related Factors. *Journal of Building Performance, 10*(1), 79 - 84. Retrieved from http://spaj.ukm.my/jsb/index.php/jbp/index

Adedokun, O. A., Egbelakin, T., Adedokun, D. O., & Adafin, J. K. (2021b). Success Criteria-Based Impacts of Risk Factors on Education Building Projects in Southwestern Nigeria. *Journal of Engineering, Design and Technology*. doi:10.1108/JEDT-09-2021-0458

Akinbile, B. F., Ofuyatano, M., Oni, O. Z., & Agboola, O. D. (2018). Risk Management and Its Influence on Construction Project in Nigeria. *International Journal of Engineering, XVI*, 189-194.

Akinsiku, O. E., & Akinsulire, A. (2012). Stakeholders' Perception of the Causes and Effects of Construction Delays on Project Delivery. *Journal of Construction Engineering and Project Management*, *2*(4), 25-31. doi:10.6106/jcepm.2012.2.4.025

Aladağ, H., & Işık, Z. (2019). Design and construction risks in BOT type mega transportation projects. *Engineering, Construction and Architectural Management, 26*(10), 2223-2242. doi:10.1108/ecam-08-2018-0351

Alklkali, A. M. T. (2022). Risk assessment in construction projects of Iraq. *Journal of Algebraic Statistics, 13*(1), 517-525. Retrieved from <Go to ISI>://WOS:000813510700014

Ansah, R. H., Sorooshian, S., Mustapha, S. B., & Duvvuru, G. (2016). *Assessment of Environmental Risks in Construction Projects: A Case of Malaysia*. Paper presented at the International Conference on Industrial Engineering and Operations Management, Detroit, Michigan, USA, September 23 - 25.

Azar, A. T. e. a. (2014). A random forest classifier for lymph diseases. *Computer Methods and Programs in Biomedicine, 113*(2), 465-473. doi:10.1016/j.cmpb.2013.11.004

Babalola, I. H., Oluwatuyi, O. E., Akinloye, L., & Aiyewalehinmi, E. (2015). Factors Influencing the Performance of Construction Projects in Akure, Nigeria. *International Journal of Civil Engineering, Construction and Estate Management, 3*(4), 57-67.

Bakri, A. S., Ab Razak, M. A., & Abd Shukor, A. S. (2021). Identification of Factors Influencing Time and Cost Risks in Highway Construction Projects. *International Journal of Sustainable Construction Engineering and Technology, 12*(3), 280-288. doi:10.30880/ijscet.2021.12.03.027

Banaitis, A., & Banaitiene, N. (2012). Risk Management in Construction Projects. 429-448. doi:10.5772/51460

Belgiu, M., & Dra, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing, 114*, 24-31. doi:10.1016/j.isprsjprs.2016.01.011

Carter, J. V., Pan, J., Rai, S. N., & Galandiuk, S. (2016). ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery*, 1638-1645. doi:10.1016/j.surg.2015.12.029

Chileshe, N., & Boadua Yirenkyi-Fianko, A. (2012). An evaluation of risk factors impacting construction projects in Ghana. *Journal of Engineering, Design and Technology, 10*(3), 306-329. doi:10.1108/17260531211274693

Fan, J., Upadhye, S., & Worster, A. (2006). Understanding receiver operating characteristic (ROC) curves. *Canadian Journal of Emergency Medicine, 8*(1), 19-20.

Ghaffari, A. (2013). Concepts of Risk In Construction Projects. In B. Gan, Y. Gan, & Y. Yu (Eds.), *Advances in Applied Materials and Electronics Engineering* (Vol. 684, pp. 644-649).

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random Forests for land cover classification. *Pattern Recognition Letters, 27*(4), 294-300. doi:10.1016/j.patrec.2005.08.011

Hanafi, M. H., Zhen, O. M., & Razak, A. A. (2021). Contractors' Perspective on the Main Factors Influencing On-Site Labour Productivity: A Focus on Malaysian Infrastructure Projects. *International Journal of Sustainable Construction Engineering and Technology, 12*(1), 68-78. doi:10.30880/ijscet.2021.12.01.007

Infrastructure Australia. (2019). An Assessment of Australia’s Future Infrastructure Needs: The Australian Infrastructure Audit 2019. Retrieved from https://www.infrastructureaustralia.gov.au/sites/default/files/2019-08/Australian%20Infrastructure%20Audit%202019.pdf, accessed 16th November, 2022

Kadir, A., Lestari, F., Sunindijo, R. Y., Erwandi, D., Kusminanti, Y., Modjo, R., . . . Ramadhan, N. A. (2022). Safety Climate in the Indonesian Construction Industry: Strengths, Weaknesses, and Influential Demographic Characteristics. *Buildings, 12*(5). doi:10.3390/buildings12050639

Lapidus, A., Topchiy, D., Kuzmina, T., & Chapidze, O. (2022a). Influence of the Construction Risks on the Cost and Duration of a Project. *Buildings, 12*(4). doi:10.3390/buildings12040484

Lapidus, A., Topchiy, D., Kuzmina, T., & Chapidze, O. (2022b). Influence of the Construction Risks on the Cost and Duration of a Project. *Buildings, 12*(4), 19. doi:10.3390/buildings12040484

Laryea, S. (2011). Quality of tender documents: case studies from the UK. *Construction Management and Economics, 29*(3), 275-286. doi:10.1080/01446193.2010.540019

Laryea, S., & Hughes, W. (2011). Risk and Price in the Bidding Process of Contractors. *Journal of Construction Engineering and Management, 137*(4), 248-258. doi:10.1061/(Asce)Co.1943-7862.0000293

Ling, F. Y. Y., & Hoang, V. T. P. (2010). Political, Economic, and Legal Risks Faced in International Projects: Case Study of Vietnam. *Journal of Professional Issues in Engineering Education and Practice, 136*(3), 156 - 164. doi:10.1061/共 ASCE 兲 EI.1943-5541.0000015

Mahmoud, H., Ahmed, V., & Beheiry, S. (2021a). Construction Cash Flow Risk Index. *Journal of Risk and Financial Management, 14*(6), 17. doi:10.3390/jrfm14060269

Mahmoud, H., Ahmed, V., & Beheiry, S. (2021b). Construction Cash Flow Risk Index. *Journal of Risk and Financial Management, 14*(6). doi:10.3390/jrfm14060269

Mishra, S., & Mishra, B. (2016). A Study on Risk Factors Involved in the Construction Projects. *International Journal of Innovative Research in Science, Engineering and Technolog, 5*(2). doi:10.15680/IJIRSET.2016.0502005

Moser, C. A., & Kalton, G. (2017). Survey Methods in Social Investigation. Retrieved from https://doi.org/10.4324/9781315241999, accessed 16th Novermber, 2022

Musa, M. M., Bin Amirudin, R., Sofield, T., & Musa, M. A. (2015). Influence of External Environmental Factors on the Success of Public Housing Projects in Developing Countries. *Construction Economics and Building, 15*(4), 30-44. doi:10.5130/AJCEB.v15i4.4514

Nguyen, P. T., Pham, C. P., Phan, P. T., Vu, N. B., Duong, M. T. H., & Nguyen, Q. (2021). Exploring Critical Risk Factors of Office Building Projects. *Journal of Asian Finance Economics and Business, 8*(2), 309-315. doi:10.13106/jafeb.2021.vol8.no2.0309

Nketekete, M., Emuze, F., & Smallwood, J. (2017). Risk management in public sector construction projects: Case studies in Lesotho. *Acta Structilia 23(2), 23*(1). doi:10.18820/24150487/as23i2.1

Odimabo, O. O., & Oduoza, C. F. (2013). Risk Assessment Framework for Building Construction Projects' in Developing Countries. *International Journal of Construction Engineering and Management, 2*(5), 143-154. doi:10.5923/j.ijcem.20130205.02

Ofori, G. (2007). Construction in Developing Countries. *Construction Management and Economics, 25*(1), 1-6. doi:10.1080/01446190601114134

Oke, A., Ogungbile, A., Oyewobi, L., & Tengan, C. (2016). Economic Development as a Function of Construction Project Performance. *1447-1459, 6*, 1447-1459.

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing, 26*(1), 217-222. doi:10.1080/01431160412331269698

Petkovic, D., Altman, R., Wong, M., & Vigil, A. (2018). Improving the explainability of Random Forest classifier - user centered approach. *Pacific Symposium Biocomputing, 23*, 204-215.

Raamkumar, J. S., & Indhu, B. (2021, Mar 25-26). *A Study on the Assessment of Risk Management in High-Rise Buildings.* Paper presented at the International Conference on Advances in Construction Materials and Management (ACMM), Kattankulathur, INDIA.

Rahman, N. N. N. A., & Esa, N. (2014). Managing Construction Development Risks to the Environment. In *Sustainable Living with Environmental Risks* (pp. 193-202).

Renuka, S. M., Umarani, C., & Kamal, S. (2014). A Review on Critical Risk Factors in the Life Cycle of Construction Projects. *Journal of Civil Engineering Research, 4*(2A), 31-36. doi:10.5923/c.jce.201401.07

Sepasgozar, S. M. E., Costin, A. M., Karimi, R., Shirowzhan, S., Abbasian, E., & Li, J. Y. (2022). BIM and Digital Tools for State-of-the-Art Construction Cost Management. *Buildings, 12*(4). doi:10.3390/buildings12040396

Shehu, Z., Endut, I. R., Akintoye, A., & Holt, G. D. (2014). Cost overrun in the Malaysian construction industry projects: A deeper insight. *International Journal of Project Management, 32*(8), 1471-1480. doi:10.1016/j.ijproman.2014.04.004

Shendurkar, P. S., Jain, M. S., & Sudarsan, J. S. (2021, Mar 25-26). *Impact of Risk Assessment in Project Execution and Its Mitigation Strategies Using Modern Automation.* Paper presented at the International Conference on Advances in Construction Materials and Management (ACMM), Kattankulathur, INDIA.

Stanley, M. B. (2011). Seven Reasons Competitive Tendering Fails (And What You Can Do About It). Retrieved from https://www.processexcellencenetwork.com/innovation/columns/seven-reasons-competitive-tendering-fails-and-what, accessed 16th November, 2022.

Sutrisna, M., & Goulding, J. (2019). Managing information flow and design processes to reduce design risks in offsite construction projects. *Engineering, Construction and Architectural Management, 26*(2), 267-284. doi:10.1108/ecam-11-2017-0250

Szymański, P. (2017). Risk management in construction projects. *Procedia Engineering, 208*, 174-182. doi:10.1016/j.proeng.2017.11.036

Tetteh, M. O., Chan, A. P. C., Darko, A., & Nani, G. (2020). Factors affecting international construction joint ventures: a systematic literature review. *International Journal of Construction Management*, 1 - 16. doi:10.1080/15623599.2020.1850203

Xu, B., Guo, X., Ye, Y., & Cheng, J. (2012). An Improved Random Forest Classifier for Text Categorization. *J. Comput., 7*(12), 2913-2920.

Zhang, Q. M., Guo, X. M., & Li, H. C. (2021). The Impact of Financial Risks on Financial Investment in Infrastructure: Based on a Two-Factor Stochastic Differential Equation. *Discrete Dynamics in Nature and Society, 2021*. doi:10.1155/2021/9112739