International Journal of Construction Management

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjcm20

Machine learning model for delay risk assessment in tall building projects

Muizz O. Sanni-Anibire, Rosli Mohamad Zin & Sunday Olusanya Olatunji

To cite this article: Muizz O. Sanni-Anibire, Rosli Mohamad Zin & Sunday Olusanya Olatunji (2022) Machine learning model for delay risk assessment in tall building projects, International Journal of Construction Management, 22:11, 2134-2143, DOI: 10.1080/15623599.2020.1768326

To link to this article: https://doi.org/10.1080/15623599.2020.1768326

Published online: 25 May 2020.

INTERNATIONAL JOURNAL OF CONSTRUCTION MANAGEMENT
2022, VOL. 22, NO. 11, 2134–2143
https://doi.org/10.1080/15623599.2020.1768326

Taylor & Francis
Taylor & Francis Group

Check for updates

Machine learning model for delay risk assessment in tall building projects

Muizz O. Sanni-Anibire[a], Rosli Mohamad Zin[b] and Sunday Olusanya Olatunji[c]

[a]Dammam Community College, King Fahd University of Petroleum and Minerals, Dhahran, Kingdom of Saudi Arabia; [b]Faculty of Engineering, School of Civil Engineering, Universiti Teknologi Malaysia (UTM), Johor, Malaysia; [c]Department of Computer Science, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam, Kingdom of Saudi Arabia

ABSTRACT

Risky projects such as tall buildings have suffered an alarming rate of increase in delays and total abandonment. Though numerous delay studies predominate, what is lacking is constructive research to develop tools and techniques to wrestle the inherent problem. Consequently, this paper presents the development of a machine learning model for delay risk assessment in tall building projects. Initially, 36 delay risk factors were identified from previous literature, and subsequently developed into surveys to determine the likelihood and consequence of the risk factors. Forty-eight useable responses obtained from subject matter experts were used to develop a dataset suitable for machine learning application. K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), Support Vector Machines (SVM) and Ensemble methods were considered. Feature subset selection revealed that the most relevant independent variables include "slowness in decision making"; "delay in sub-contractors work"; "architects'/structural engineers' late issuance of instruction"; and "waiting for approval of shop drawings and material samples". The final results showed that the best model for predicting the risk of delay was based on ANN with a classification accuracy of 93.75%. Ultimately, the model developed in this study could support construction professionals in project risk management of tall buildings.

KEYWORDS
Tall buildings; delays; risk assessment; k nearest neighbor; neural networks; support vector machines; ensemble methods

Introduction

The global construction industry has been afflicted for many decades by underperformance issues, significant amongst which includes the incessant occurrence of delays (Mahamid et al. 2012; Sanni-Anibire et al. 2020). The negative effects of delays may include lawsuits, cost and time overruns, loss of productivity and revenue, and contract termination (Sambasivan and Soon 2007; Tumi et al. 2009). Interestingly, the construction industry continues to soar higher with the astronomical growth of tall buildings, notably in Asian and Middle Eastern skies (Moon 2015). Despite its exponential rate of growth, tall buildings have suffered from delayed completion times. According to the Council on Tall Buildings and Urban Habitat (CTBUH), there is an alarming rate of increase in never competed tall building projects across the globe (CTBUH 2014). Although, numerous studies exist relevant to delay in construction projects (Assaf et al. 1995; Ogunlana et al. 1996; Kaming et al. 1997; Abd El-Razek et al. 2008; Muneeswaran et al. 2018; Sanni-Anibire et al. 2020), these studies are at best exploratory in nature and do not provide a pragmatic solution to the inherent problem. Remarkably, AlSehaimi et al. (2013) suggest that the problem of delays in the construction industry could be mitigated through alternative research approaches. The much desired alternative will be geared towards the development of innovative tools that could better transfer research findings to practical use, and ultimately tackle the problem at hand.

On this note, previous research efforts have sought to develop risk assessment tools as a viable strategy to mitigate construction delays. ISO 31000: 2018 define risk as "the effect of uncertainty on an organization's ability to meet its objectives". Hence, risk assessment entails identifying events that have the likelihood to cause deviation in a project's objectives. A study by Kim et al. (2009) developed a Bayesian Belief Network (BBN) model to quantify the probability of construction project delays in Vietnam. Hossen et al. (2015) proposed a method that combines the Analytical Hierarchy Process (AHP) and the Relative Importance Index (RII) for assessing construction delay risk in nuclear power plants. Muneeswaran et al. (2018) also utilized the RII and fuzzy ranking to assess schedule delay and risk in the Indian construction industry. Similarly, Gunduz et al. (2015) proposed a system to quantify the probability of delay in construction projects in Turkey based on a combination of the RII methodology and fuzzy logic. Budayan et al. (2018) developed a fuzzy-based computerized method for the risk assessment of construction delay relying on users' previous experience, expertise and judgement. Despite evident research progress that has been made, the construction industry is undergoing a technological shift driven by the fourth industrial revolution (Industry 4.0). The current mantra of the construction industry is "Construction 4.0" – aimed at digitalization and automation in the industry for improved productivity (García de Soto et al. 2019). Machine learning-a subset of artificial intelligence-is considered one of the top ten technologies driving Industry 4.0 (PricewaterhouseCoopers 2017). Despite the industry's ambition manifest through Construction 4.0, it is still lagging behind other industries in adopting smart technologies. Strikingly, a random literature search reveals that numerous studies to apply machine learning to delay risk prediction in other industries predominate (Yaghini et al. 2013; Choi et al. 2016; Kim et al. 2016; Takeichi et al. 2017). However, there has been little research effort aimed

CONTACT Muizz O. Sanni-Anibire muizzanibire@kfupm.edu.sa

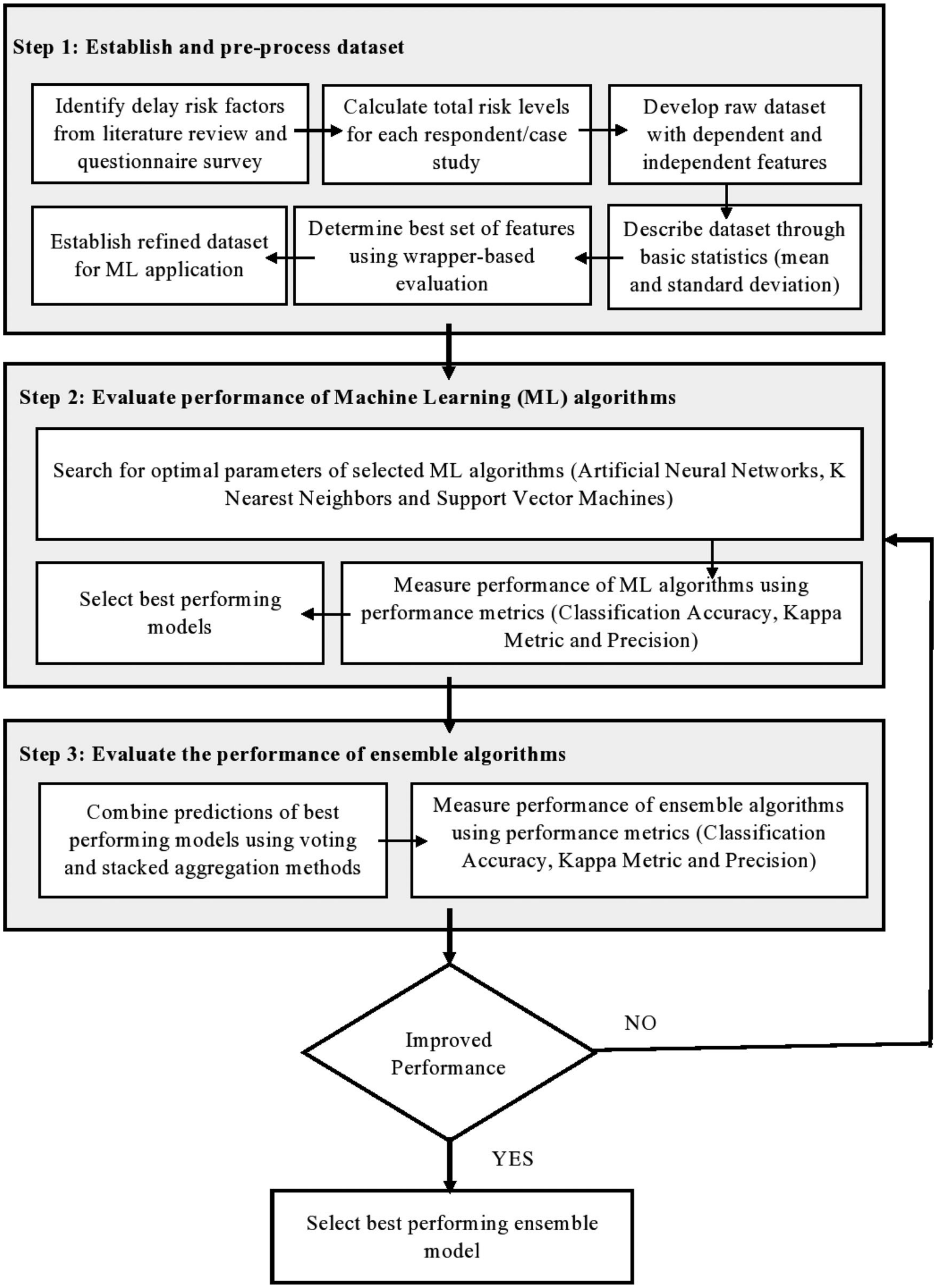

Figure 1. Methodology for developing the proposed delay risk prediction model.

at the application of machine learning to mitigate delays in construction projects. A recent study by Gondia et al. (2020) sought to apply machine learning algorithms to predict construction project delay risk. The study cited its main value proposition to be the use of objective data extracted from real case studies. Regrettably, such data are hard to obtain as the construction industry is still largely deficient in recording and publishing data suitable for machine learning applications. The implication of such deficiency is to resort to subjective data which can be obtained from subject matter experts. Notably, a study by El-Kholy (2019) utilized subjective data and proposed a modular neural network model to predict the percentage delay in highway projects in Egypt. The existence of subjective data, though unsuited to generalization of results, has the potential to promote

Table 1. Descriptive statistics of the questionnaire response.

S/N		Causes of delay	Likelihood Mean	Likelihood Standard Deviation	Consequence Mean	Consequence Standard Deviation
Causes related to material						
1	Mat. 1	Shortage in construction materials/unforeseen material damages	2.96	1.18	3.57	1.06
2	Mat. 2	Slow delivery of materials	3.23	0.99	3.77	0.94
3	Mat. 3	Waiting for approval of shop drawings and material samples	3.28	1.16	3.72	1.06
Causes related to manpower						
4	Man. 1	Shortage in manpower (skilled, semi-skilled, unskilled labor)	3.28	1.19	3.89	0.98
5	Man. 2	Poor labor productivity	3.5	0.96	3.89	0.92
6	Man. 3	Labor disputes and strikes	2.47	1.25	3.35	1.39
Causes related to equipment						
7	Equip. 1	Poor equipment productivity (breakdown/maintenance problem)	2.79	1.02	3.57	1.08
8	Equip. 2	Shortage in equipment	2.84	1.15	3.63	0.99
Causes related to contractual relations						
9	Cont. 1	Inappropriate construction/contractual management/ construction methods	3.51	1.1	4.19	0.79
10	Cont. 2	Slowness in decision making	3.45	1.12	3.98	1.04
11	Cont. 3	Delay in mobilization	3.34	1.15	3.59	1.02
12	Cont. 4	Excessive bureaucracy/interference by the owner	3.62	0.95	3.8	1.07
13	Cont. 5	Delay in approval of completed work	3.38	1.05	3.82	0.89
14	Cont. 6	Delay in sub-contractors work	3.71	0.79	4.05	0.65
Causes related to government						
15	Gov. 1	Slow permits from municipality/government	3.79	1.13	4.07	0.98
16	Gov. 2	Government regulations	3.31	1.13	3.67	1.08
Causes related to financing						
17	Fin. 1	Contractor's financial difficulties	3.92	1.01	4.30	0.84
18	Fin. 2	Client's cash flow problems/Delays in contractor's payment	4.1	0.88	4.44	0.69
19	Fin. 3	Price escalation/fluctuations	3.29	0.98	3.56	0.89
Causes related to environmental factors						
20	Env. 1	Weather condition	2.71	1.13	2.96	1.03
21	Env. 2	Civil disturbances/Hostile political conditions	2.15	1.13	2.87	1.13
Causes related to changes						
22	Chng. 1	Design errors/incomplete made by designers (Architects and structural drawing)	3.48	1.24	3.98	0.95
23	Chng. 2	Design variations/change orders/increase in scope of work	3.85	0.89	4.09	0.69
24	Chng. 3	Errors committed due to lack of experience	3.42	1.16	3.96	0.94
25	Chng. 4	Unexpected foundation conditions encountered in the field	2.79	1.15	3.5	1.05
26	Chng. 5	Changes in materials types and specifications during construction	3.1	1.02	3.67	0.82
27	Chng. 6	Inaccurate site/soil investigation	2.96	1.16	3.52	1.07
28	Chng. 7	Frequent change of sub-contractor	3.1	0.95	3.61	0.93
Causes related to scheduling and controlling techniques						
29	Sch. 1	Poor site organization and coordination between various parties	3.75	1.19	4.24	0.89
30	Sch. 2	Poor planning of resources and duration estimation/scheduling	3.61	1.18	3.93	1.01
31	Sch. 3	Inadequate supervision, inspection and testing procedures	3.4	1.16	3.73	0.99
32	Sch. 4	Accidents during construction/lack of safety measures	3.06	1.09	3.67	1.15
33	Sch. 5	Poor communication/documentation and detailed procedures	3.13	0.97	3.59	0.87
34	Sch. 6	Unrealistic time schedule imposed in contract	3.53	0.95	3.69	1.16
35	Sch. 7	Poor qualification of the contractor or consultant	3.78	0.87	4.18	0.72
36	Sch. 8	Architects'/structural engineers' late issuance of instruction	3.21	0.91	3.62	0.81

rapid application and adoption of machine learning techniques to solve construction related problems.

In light of the foregoing, the goal of this study is to develop a machine learning model in order to facilitate delay risk assessment in tall building projects. The study has relied on subjective data obtained from subject matter experts working on tall building projects in the Gulf Cooperating Council (GCC) countries. To begin with, the study identified delay risk factors from the extant literature, and further assessed the identified delay risks through questionnaire surveys in two domains of likelihood and consequence. The collected data was used to estimate the risk of delay for the project under consideration according to the expert's judgement. Subsequently, the data obtained was formatted to be suitable for machine learning application. The choice of machine learning algorithms used in this study was based on two factors: (1) methods commonly used in the problem's domain i.e., construction research; (2) based on Benzécri (1973)'s idea to "let the data speak for itself", i.e., experimenting with a suite of algorithms and determining what works best for the dataset. This study considered Artificial Neural Networks (ANN), k Nearest Neighbors (KNN), Support Vector Machines (SVM) and Ensemble techniques. These methods have been adopted in similar construction research (Attal 2010; Czarnigowska and Sobotka 2014; Bayram 2017; Peško et al. 2017). Finally, the algorithms used were compared based on relevant performance metrics such as the classification accuracy, precision, Cohen's kappa statistic, true positive rate, and false positive rate. The ensuing sections describe the machine learning techniques considered, methodology of the study, results and findings, discussion and conclusion sections.

Overview of machine learning techniques

Machine learning is a burgeoning field of artificial intelligence used for data modelling i.e., developing mathematical abstractions of data that can be used by computers to provide accurate prediction. Supervised classification is a main category of machine learning problems. Supervised classification contain three fundamental components: (1) the instance space containing a set of independent variables; (2) the label space for each instance consisting of the dependent variable; (3) the machine learning algorithm (Larrañaga et al. 2018). In this section, a

Table 2. Respondents' demographics.

Demographics	Category	Frequency (No.)	Percentage (%)
Professional organisation	Contractors	14	30
	Contactors	17	35
	Clients' representatives/facility managers	17	35
Professional role	Executive director	6	13
	Director	7	15
	Senior project manager	15	32
	Civil engineer	9	19
	Facility manager	11	23
Years of experience	5 to 10	13	28
	10 to 15	12	24
	> 15	23	48
Location	United Arab Emirates	20	42
	Saudi Arabia	15	32
	Kuwait	7	14
	Bahrain	3	6
	Oman	2	4
	Qatar	1	2

summary of common state-of-the-art machine learning algorithms used in this study (ANN, KNN and SVM) is presented. Detailed mathematical descriptions of these techniques may be found in relevant references (Cortes and Vapnik 1995; Witten et al. 2011; Olatunji 2017; Sethi et al. 2017; Wauters and Vanhoucke 2017).

K-nearest neighbors (KNN)

The k-nearest neighbor (KNN) classifier uses a simple majority decision rule to make new class predictions based on classes associated with the k-instances of the training set (Fix and Hodges 1951). KNN is a simple and widespread method that employs historical data to identify the nearest neighbors of a given data point (Yu et al. 2016; Wauters and Vanhoucke 2017). In order to compute a prediction, the training instances closest to the new observation are utilized (Wauters and Vanhoucke 2017). The prediction accuracy based on the k-NN model is highly contingent on the value of k (Sethi et al. 2017). For instance, with $k = 3$ neighbors, the instance is assigned to the majority class of the three nearest neighbors (Larrañaga et al. 2018).

Artificial neural networks (ANN)

Artificial Neural Networks are computational models designed to mimic the behavior of biological neural networks (McCulloch and Pitts 1943). ANNs are represented as adaptive systems of interconnected "neurons" arranged in a multi-layered network. It is capable of capturing, representing and simulating complex non-linear relationships between inputs and outputs by performing multiple parallel computations (Lek and Guégan 1999; Sethi et al. 2017). The layers are mainly composed of an input layer, an output layer, as well as hidden layers. The learning process is based on repeatedly adjusting the numerical weights associated with the interconnecting edges between different artificial neurons (Sethi et al. 2017).

Support vector machines (SVM)

Support vector machines is a blend of linear modeling and instance-based learning (Cortes and Vapnik 1995; Witten et al. 2011). Currently, SVM is one of the most widely used machine learning techniques for both classification and regression problems (Olatunji 2017). The theoretical basis of SVM is based on structural risk minimization and statistical learning theory with the aim of determining the hyperplane (decision boundaries) that produce the efficient separation of classes (Sethi et al. 2017). The boundaries are formed through the selection of a small number of critical boundary instances called support vectors. This is utilized to build a linear discriminant function that separates them as widely as possible. This instance-based approach used by SVM allows it to use extra nonlinear terms in the function, making it possible to form quadratic, cubic, and higher-order decision boundaries (Witten et al. 2011). This phenomenon is referred to as the "kernel trick", which is simply an approach to transform the input into high dimensional feature spaces where linear classification can then be carried out (Cortes and Vapnik 1995; Sethi et al. 2017). The ability to translate predictors into a higher feature space gives SVM the unique quality of handling complex relationships between predictors and outcomes (Olatunji 2017; Wauters and Vanhoucke 2017).

Methodology

The methodology adopted to achieve the objectives of this study is illustrated in Figure 1, and described in the following sections.

Dataset establishment

The establishment of the dataset suitable for the study was developed in three phases. Initially, a thorough review of the extant literature was made to identify the most common delay risk factors in the construction industry. The review of literature revealed that there is no consensus on the number of construction delay risk factors, while some researchers have presented a list of over 200 delay risk factors (Aziz and Abdel-Hakam 2016). Thus this study identified 36 frequently investigated delay risk factors extracted from high impact studies published in the last 15 years (Aibinu and Odeyinka 2006; Faridi and El-Sayegh 2006; Lo et al. 2006; Sambasivan and Soon 2007; Abd El-Razek et al. 2008; Sweis et al. 2008; Toor and Ogunlana 2008; Enshassi et al. 2009; Fugar and Agyakwah-Baah 2010; Doloi et al. 2012; Gündüz et al. 2013). These studies have been selected due to their influential status in the construction research domain reflected in their high number of citations. Further coding and categorization of the risk factors was made as presented in Table 1. Nine risk categories were considered in line with previous studies (Assaf et al. 1995; Fugar and Agyakwah-Baah 2010;

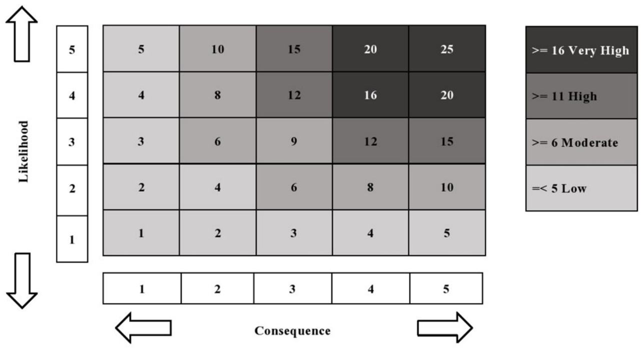

Figure 2. Discretization matrix adopted for delay risk assessment.

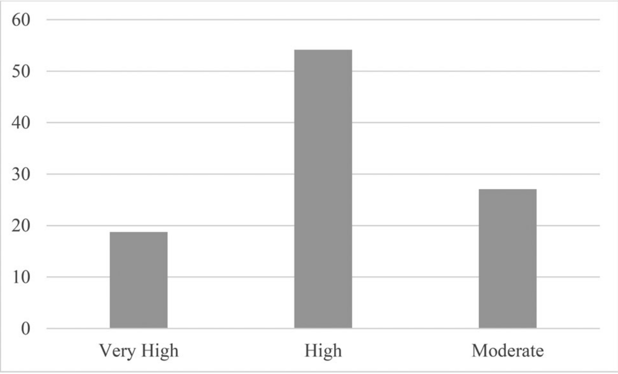

Figure 3. Imbalanced distribution of output class for data set.

Gündüz et al. 2013). This risk categories include material, manpower, equipment, contractual relations, government, financing, environmental factors, changes and scheduling and controlling techniques.

The second phase entailed the development of a questionnaire survey to obtain the feedback of construction experts on the consequence and likelihood of the identified risk factors as it pertains to their respective tall building project. The questionnaire survey contained two domains (consequence and likelihood), each designed with a Likert scale from (1) to (5). Where (1) represents: Very Low, (2): Low 3: Moderate, (4): High, and (5): Very High. The respondents included professionals active across the life cycle of tall building projects in three major categories (i.e., consultants, contractors, and clients' representatives/facility managers). The demographic profile of the respondents is presented in Table 2. Various strategies were used in distributing the questionnaire survey, this involved hand delivered hard copies to managers at tall building construction sites, as well as a web-based format sent by emails to tall building professionals in the GCC countries. These countries include Saudi Arabia, United Arab Emirates (UAE), Bahrain, Kuwait, Oman and Qatar. The sample collection was based on Fellows and Liu (2015)'s suggestion to obtain at least 32 responses for "large number" statistics. Therefore, a total of 62 responses were received, while 14 responses were removed due to numerous missing cells which may influence the performance of machine learning algorithms. Thus, a total of 48 responses were considered suitable for further establishment of the dataset. The mean values and standard deviation of the questionnaire results for all 36 risk factors are presented in Table 1.

The third phase was to develop the input and output variables of the dataset. The input variables of the dataset were based on the likelihood rating assigned by the experts in the questionnaire survey. This is based on the premise that risk is a function of likelihood. As given in ISO 31000:2018, the risk level is the product of an events' consequence and its likelihood. Therefore, the output variable was calculated based on Equations (1) and (2). Where (RL_1) is the risk level calculated for a single delay risk factor for each respondent, and (RL_2) is the average risk level of all risk factors for each respondent.

$$Risk\ Level\ (RL_1) = Consequence\ (C) * Likelihood\ (L) \tag{1}$$

$$Risk\ Level\ (RL_2) = \sum (RL_1)/n \tag{2}$$

where n is the total number of delay risk factors which is 36.

The risk levels were further interpreted according to the risk discretization matrix as shown in Figure 2. Thus, the entire dataset had three imbalanced classes as presented in Figure 3, where very high risks represent 19%, high risks 54%, and moderate risk 27%. The final structure of the compiled dataset is as presented in Table 3.

Data pre-processing

Machine learning problems require the identification of the best set of features that improves the predictive performance of a model. This helps to resolve a common phenomenon known as the "curse of dimensionality". In this study, the Waikato Environment for Knowledge Analysis (WEKA 3.8.3) has been used. This is an open source machine learning software written in Java, and developed at the University of Waikato, New Zealand (Witten et al. 2011). WEKA is one of the most popular tools used in the machine learning community due to its user friendliness and the wide range of algorithms offered on the platform (Larrañaga et al. 2018). The ML algorithms used in this study are described in the WEKA environment as follows: k-NN: *"IBk"*, ANN: *"Multilayer Perceptron"*, SVM: *"SMO"*, *Ensemble voting: "meta.Vote", and Ensemble stacking: "meta.stacking"*. To execute feature selection, the "*CorrelationAttributeEval*" technique was used to determine the correlation and rank of various features in the dataset to the prediction output. Subsequently, attribute selection was made based on the Recursive Feature Elimination (RFE) procedure explained by (Akande et al. 2015). In RFE, the entire feature set (V) ranked according to the correlation co-efficient is split in half to derive the best V/2 features, and the worst V/2 features are eliminated. The splitting process continues recursively until only one best feature is left. Thereafter, the feature subset that achieved the best accuracy/or the best performance measure is finally chosen as the best subset to be used. The results of the RFE approach are presented in Table 4. The next step was to determine the best set of features to be used in the model development process. Wrapper-based feature evaluation was employed. This is a situation where a powerful classifier is used to compare all the feature sets, and the best set of features according to relevant performance metrics is selected. In this study, SVM (*SMO*) was adopted as the wrapper-based classifier. It can be observed from Table 6 that the feature set with the highest accuracy was the best V/8 features (described in Table 4). In essence, the refined dataset to be used for developing machine learning models contained 4 independent variables ("slowness in decision making"; "delay in sub-contractors work"; "architects'/structural engineers' late issuance of instruction"; and "waiting for approval of shop drawings and material samples"), and three classes as dependent variables (described in Figure 3). The dataset was split into a train-test ratio of 66% to 34%.

Table 3. Compiled data set structure.

Project Case	Mat. 1	Mat. 2	Mat. 3	–	–	–	Sch. 6	Sch. 7	Sch. 8	Class
Case 1	Moderate	Low	Very Low	–	–	–	Low	Low	Very Low	High
Case 2	Low	Moderate	High	–	–	–	Very Low	Moderate	High	Moderate
Case 3	Very Low	High	Moderate	–	–	–	High	High	Moderate	Very High
–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–
Case 47	High	Moderate	Very Low	–	–	–	High	Very Low	Low	Moderate
Case 48	Moderate	Low	Very Low	–	–	–	Low	Very Low	High	Very High

Table 4. Description of feature sets based on *CorrelationAttributeEval*.

RFE process	No. of features	Description
V features	36	All features
Best V/2	18	Cont. 2; Cont. 6; Sch. 8; Mat. 3; Sch. 7; Cont. 1; Fin. 2; Cont. 4; Equip. 1; Chng. 6; Chng. 1; Chng. 3; Sch. 4; Equip. 2; Man. 1; Sch. 3; Man. 2; Gov. 2
Best V/4	9	Cont. 2; Cont. 6; Sch. 8; Mat. 3; Sch. 7; Cont. 1; Fin. 2; Cont. 4; Equip. 1
Best V/8	4	Cont. 2; Cont. 6; Sch. 8; Mat. 3
Best V/16	2	Cont. 2; Cont. 6
Best feature	1	Cont. 2

Evaluate machine learning algorithms

The performance of ML algorithms is dependent on the tuning of optimization hyperparameters. In this study, a systematic search was employed, where various sets of hyperparameters were used to train the model progressively until satisfactory results were obtained. The optimization hyperparameters for the algorithms considered in this study are described in Table 5.

Performance measurement

In measuring the performance of the algorithms employed, the confusion matrices were used. The confusion matrix provides values for the True Positives (TP), False Positives (FP), False Negatives (FN) and True Negatives (TN). In binary classification (2×2 matrices), TP and TN are the number of instances correctly classified as positive and negative, respectively, while FP and FN are the falsely classified positives and negatives, respectively. As for multi-class confusion matrices, the TPs, FPs, FNs and TNs are computed separately for each class and the weighted average determined. In this case, TPs are the correctly classified instances along the diagonal of the matrix. TNs for each class is the sum of all classified instances outside the class in consideration. FNs for each class are the instances from that class incorrectly classified as other classes. Finally, FPs of a class are the sums of samples from all other classes (Tharwat 2018). An illustrative example has been provided in a later section of this study. The classification accuracy is the ratio of correct classifications to the total number of classifications by a model. The Cohen's kappa statistic is also considered very useful in situations of class imbalance as is the case with this study (see Figure 3). The performance metrics used in this study are computed as described in Equations (3)–(7).

$$Classification\ Accuracy = \frac{TP + TN}{N} \tag{3}$$

where $N =$ total number of samples

And $Misclassification\ Error\ (1 - Classification\ Accuracy)$

$$Cohen's\ kappa\ statistic = \frac{\frac{TP}{N} + \frac{TN}{N} - A}{1 - A} \tag{4}$$

$$A = \left(\frac{FN + TP}{N}\right)\left(\frac{FP + TP}{N}\right) + \left(\frac{FP + TN}{N}\right)\left(\frac{FN + TN}{N}\right)$$

$$Precision\ (Positive\ Preditive\ Value) = \frac{TP}{TP + FP} \tag{5}$$

$$TP\ Rate\ (Sensitivity) = \frac{TP}{TP + FN} \tag{6}$$

$$FP\ Rate\ (1 - Specificity) = \frac{FP}{FP + TN} \tag{7}$$

Evaluate ensemble algorithms

In machine learning, the performance of base classifiers may be improved through ensemble techniques. This is an approach where the predictions of base classifiers are combined by some standard procedure. In this study, voting and stacking ensemble methods have been considered (Xia et al. 2011; Kuncheva and Rodríguez 2014). Voting as the name implies, involves the combination of two or more sub-models by some system such as the mean or mode. Thus, each sub-model votes on what the outcome should be. Stacking allows another algorithm to learn how best to combine the predictions of other sub-models (Brownlee 2018).

Results and findings

As noted earlier, feature subset selection identifies and removes irrelevant and redundant features in the dataset. Additionally, data dimensionality is reduced, which helps learning algorithms induce faster and more effective models (Larrañaga et al. 2018). Firstly, "*CorrelationAttributeEval*" was used to rank the 36 delay risk factors base on the correlation (Pearson) with the output class. The ranked delay risk factors were further split according to the Recursive Feature Elimination (RFE) process as explained in a previous section. Finally, a wrapper-based feature evaluation process was employed to evaluate the performance of the various feature sets in Table 4. In this study, SVM was adopted as the wrapper-based classifier, and the results are presented in Table 6. It can be perceived that there was an increase in performance directly proportional to the decrease in dimensionality of the dataset. However, performance started to diminish after the "Best V/8" features, which had a classification accuracy of 87.5%. In other words, the best set of delay risk factors include "slowness in decision making"; "delay in sub-contractors work"; "architects'/structural engineers' late issuance of instruction"; and "waiting for approval of shop drawings and material samples". Therefore, the "Best V/8" features were selected for further model development.

The next phase involved the evaluation of the selected ML algorithms (ANN, KNN & SVM). To deploy the ML algorithms, a systematic search for the optimization hyperparameters was made as presented in Table 7. Basically, three initial models were developed and labelled MOD1, MOD2 and MOD3. The

Table 5. Description of ML algorithm's optimization parameters (Larrañaga et al. 2018).

ML Algorithm	Optimization parameters	
ANN (Multilayer Perceptron)	Learning rate	A value in the range [0,1] that changes the speed at which the weights of each connection between neurons is updated.
	Momentum	A value in the range [0,1] that uses the direction of previous weight updates to adjust the weight change speed.
	Network topology	Number of neurons in each hidden layer.
	Transfer function	It determines the output of each neuron given the input.
KNN (IBk)	k value	Number of nearest neighbors considered by the algorithm.
	Search algorithm	The manner in which the algorithm should find the nearest neighbors.
	Distance function	The distance in the feature space. The most common include Euclidean, Manhattan and Minkowski distances.
SVM (SMO)	Cost function, C	Also known as complexity constant, controls the trade-offs between errors and the margin size.
	Kernel function	A symmetric function of two arguments that returns the value of the inner product of the two mapped arguments.
	Tolerance parameter	Controls the amount of permissible SVM optimization problem-solving error
	Filter	Data transformation of the input variables before training.

Table 6. Performance of SVM wrapper-based evaluation for various feature sets.

ML Algorithm	Performance measure	All features	Best V/2	Best V/4	Best V/8	Best V/16	Best feature
SVM (*SMO*)	Classification Accuracy (%)	62.5	68.75	81.3	**87.5**	56.25	56.25
	Kappa Statistic	0.35	0.48	0.68	**0.77**	0.043	0
	Precision	0.67	0.75	0.89	**0.82**	0.6	0.56

Table 7. Parameters selected for ML models.

Model	ML Algorithm	Optimization parameters	
MOD1	ANN (*Multilayer Perceptron*)	Learning rate	0.3
		Momentum	0.2
		Network topology	One hidden layer with 3 neurons
		Transfer function	Sigmoid
MOD2	KNN (*IBk*)	*k* value	7
		Search algorithm	Linear search
		Distance function	Euclidean distance
MOD3	SVM (*SMO*)	Cost function, *C*	5
		Kernel function	Polynomial kernel
		Tolerance parameter	0.001
		Filter type	No normalization/standardization
MOD4	Voting (Combine MOD1, MOD2 and MOD3 by majority voting)	NA	NA
MOD5	Stacking (Combine MOD1 and MOD3 with ANN as combining classifier)	Learning rate	0.2
		Momentum	0.2
		Network topology	One hidden layer with 3 neurons
		Transfer function	Sigmoid
MOD6	Stacking (Combine MOD1 and MOD3 with SVM as combining classifier)	Cost function, *C*	17
		Kernel function	Polynomial kernel
		Tolerance parameter	0.001
		Filter type	No normalization/standardization

performance results of these models are computed according to Equations (3)–(7). For illustration, a sample calculation for MOD1 is presented in Table 8, while the results for the other models are presented in Table 9. It can be observed from the table that MOD1 and MOD3 had similar performance with classification accuracy of 93.75% each, while MOD2 had the least level of performance with classification accuracy of 75%. Though, the performance of MOD1 and MOD3 were highly impressive, further investigation was made to determine if an improvement in performance could be achieved through ensemble methods. Thus, another model (MOD4) was developed based on a majority vote of the achieved initial models (i.e., MOD1, MOD2 and MOD3). Also, two models based on stacking using ANN and SVM as the combining classifier were developed, and labelled as MOD5 and MOD6. The results, as shown in Table 9 showed that no further improvement in performance was achieved, and the best performing model was selected to be MOD1. The confusion matrix of MOD1 is presented in Table 10, and it can be seen that only one misclassified instance occurred, where the risk of delay was predicted to be "moderate"

Table 8. Illustrative example to compute model's performance measure.

Given:
Class = "Very High" TP (2), FN (0), TN (13), FP (1)
Class = "High" TP (9), FN (0), TN (7), FP (0)
Class = "Moderate" TP (4), FN (1), TN (11), FP (0)
Where N = total number of samples = 16

Performance measure	Very High	High	Moderate	Weighted average
Classification Accuracy (%) - Eq. (3)	–	–	–	$\frac{2+9+4}{16} = \mathbf{0.9375}$
Misclassification Error (%) - Eq. (3)	–	–	–	$1 - 0.9375 = \mathbf{0.0625}$
Cohen's Kappa Statistic - Eq. (4)	$A = \left(\frac{0+2}{16}\right)\left(\frac{1+2}{16}\right) + \left(\frac{1+13}{16}\right)\left(\frac{0+13}{16}\right) = 0.734$	$A = \left(\frac{0+9}{16}\right)\left(\frac{0+9}{16}\right) + \left(\frac{0+7}{16}\right)\left(\frac{0+7}{16}\right) = 0.508$	$A = \left(\frac{1+4}{16}\right)\left(\frac{0+4}{16}\right) + \left(\frac{0+11}{16}\right)\left(\frac{1+11}{16}\right) = 0.723$	–
	$\frac{\frac{2}{16}+\frac{13}{16}-0.734}{1-0.734} = 0.765$	$\frac{\frac{9}{16}+\frac{7}{16}-0.508}{1-0.508} = 1.000$	$\frac{\frac{4}{16}+\frac{11}{16}-0.723}{1-0.723} = 0.775$	$\frac{(0.765*3)+(1*9)+(0.775*4)}{16} = \mathbf{0.899}$
Precision - Eq. (5)	$\frac{2}{2+1} = 0.8$	$\frac{9}{9+0} = 1.0$	$\frac{4}{4+0} = 1.0$	$\frac{(0.8*3)+(1*9)+(1*4)}{16} = \mathbf{0.963}$
TP Rate - Eq. (6)	$\frac{2}{2+0} = 1.0$	$\frac{9}{9+0} = 1.0$	$\frac{4}{4+1} = 0.8$	$\frac{(1*3)+(1*9)+(0.8*4)}{16} = \mathbf{0.95}$
FP Rate	$\frac{1}{1+13} = 0.071$	$\frac{0}{0+7} = 0$	$\frac{0}{0+11} = 0$	$\frac{(0.071*3)+(0*9)+(0*4)}{16} = \mathbf{0.013}$

Table 9. Performance of initial ML models.

Performance measure	MOD1	MOD2	MOD3	MOD4	MOD5	MOD6
Classification Accuracy (%)	**93.75**	75	93.75	93.75	93.75	87.5
Misclassification Error (%)	**6.25**	25	6.25	6.25	6.25	12.5
Cohen's Kappa Statistic	**0.89**	0.50	0.89	0.89	0.89	0.77
Precision	**0.96**	0.83	0.94	0.95	0.94	0.89
TP Rate	**0.95**	0.75	0.94	0.94	0.94	0.88
FP Rate	**0.013**	0.32	0.08	0.021	0.08	0.16

Table 10. Confusion matrix for best performing model (MOD1).

	Predicted class			
Actual class	Very High	High	Moderate	Totals
Very High	2	0	1	3
High	0	9	0	9
Moderate	0	0	4	4

instead of "very high". This is represented in the misclassification error computed to be 6.25%, an indication of the excellent accuracy achieved. Overall, the performance of the models are impressive, considering that the classification accuracy are all above 75%.

Discussion

The construction industry is still at its early stages of adopting modern technology in solving its inherent under productivity issues. Interestingly, the industry's commitment towards Construction 4.0 has paved the way for an increase in research in the adoption of machine learning in construction research (Darko et al. 2020). More specifically, the construction industry has been struggling with the problem of persistent delay. Researchers have sought to tackle the problem through exploratory and investigative studies to identify the main causes of construction delay (Assaf et al. 1995; Ogunlana et al. 1996; Kaming et al. 1997; Abd El-Razek et al. 2008; Sanni-Anibire et al. 2020). Experts are of the opinion that, although identifying the sources of construction delay is the first step in providing potential solutions, it is required that research should go further to provide prescriptive tools to mitigate delay (AlSehaimi et al. 2013). While the industry, in general is in need of such tools to be developed by research, mega projects such as tall buildings have exhibited dire need for such research. For instance, CTBUH in its report "Dream Deferred: Unfinished Tall Buildings" listed over 50 "never completed" tall buildings across the globe. It is in light of the foregoing, that the objective of this study was formed.

This study aimed at developing a machine learning model to predict the risk of delay in tall building projects. The first stage of the study was to identify delay risk factors from the extant literature. Eleven influential studies in the research domain were identified, and subsequently, 36 delay risk factors were extracted. These delay risk factors were further formatted in the form of questionnaire surveys to obtain feedback on the consequence and likelihood of occurrence from tall building experts across the GCC countries. Mid-level to senior level managers were contacted to participate in the exercise, and 48 useable feedback for the machine learning problem was obtained. The reliability of the feedback is reflected in that responses were only obtained from the experts that were responsible for the day-to-day monitoring and control of the project. The risk class for each of the 48 case studies were computed as discussed in the "dataset establishment" section. In establishing the dataset, the likelihood of the delay risk factors represented the input, and the computed risk class represented the output. This is due to the fact that risk is an expression of the likelihood of occurrence. Once the final dataset was established, pre-processing was carried out, where the relevant features that yield optimum performance of the machine learning algorithms were determined. The results showed that the most relevant features were "slowness in decision making"; "delay in sub-contractors work"; "architects'/structural engineers' late issuance of instruction"; and "waiting for approval of shop drawings and material samples". Essentially, this means that project managers could determine the level of risk in their construction project by assessing the likelihood of these four factors when using the machine learning model developed in this study.

The next phase entailed the algorithm evaluation process. In this study, three common machine learning algorithms have been employed including ANN, KNN and SVM. Generally, the best algorithm to be used to solve an ML problem is usually not known beforehand. Though, there are rules of thumb based on the statistical properties of the dataset, experts have suggested that common ML algorithms should firstly be explored-especially those common in the field of the ML problem at hand (Wu et al. 2008; Brownlee 2018). The next step involved a search for the optimal hyperparameters of the selected algorithms using a modified systematic approach i.e., a range of randomly spaced values are searched first, and then the range that performs best is zoomed in for further investigation. The hyperparameters to be tuned for the selected algorithms were described in Table 5. Thus, three initial models were developed MOD1, MOD2 and MOD3. The performance evaluation of these models based on performance metrics such as the classification accuracy, misclassification error, Cohen's kappa statistic, precision, TP rate and FP rate were employed. Cohen's kappa statistic is particularly better as a performance criteria in case of an imbalanced output class illustrated in Figure 3 (Brownlee 2018). Though, an excellent performance level of 93.75% classification accuracy was achieved for MOD1 and MOD3, the study sought to further investigate the possibility of an improved performance through ensemble techniques. Majority voting and stacking was employed. It was concluded that the best model was MOD1, where ANN was used to build the model. The confusion matrix in Table 10 also showed that only one predicted instance was misclassified indicating an excellent performance of the model selected. In comparison of the model in this study with a similar study by Gondia et al. (2020), the model in this study exhibited superior performance. This is since Gondia et al. (2020)'s best performing model based on Naïve Bayes algorithm had a 51.2% classification accuracy.

The limitations of this study may be reflected in the source of the dataset used. The dataset in this study was obtained from construction professionals in the GCC region. This limits the potential to generalize the results of such a study. However, the approach and idea may be generalized and applied in other construction climates. The implication of this study is reflected in the fact that it addressed a significant problem in the construction industry, which is the increasing rate of delay in tall building projects. Furthermore, this study is in line with the current research trend in the construction industry to leverage the capabilities of machine learning-a subset of artificial intelligence, in solving some of the age long problems of the industry. Finally, the model developed in this study could help construction professionals to carry out delay risk assessments of tall building projects. This has the advantage to foster a proactive project risk management at the start of the project, and dynamic risk management through the life cycle of the project. Practically, the likelihood of delay risk factors should be determined and monitored in the course of a tall building project, and then utilized as inputs in the machine learning model.

Conclusion

The construction industry is currently witnessing a technological shift for enhanced productivity. A significant under-productivity issue is that of project delays and abandonment, which has become a menace in tall building projects. Numerous studies to investigate the cause of construction delay already predominate, however, the current challenge is in developing prescriptive tools to tackle the problem. This study makes a significant contribution towards this objective by developing a model that could facilitate delay risk assessments in tall building projects. In developing the delay risk model, 36 delay risk factors were considered, and further analysis revealed that the most influential factors included "slowness in decision making"; "delay in sub-contractors work"; "architects'/structural engineers' late issuance of instruction"; and "waiting for approval of shop drawings and material samples". Subjective data on these factors obtained from industry professionals was therefore used to develop a classification model using ANN, and the results revealed an excellent performance level of 93.75% classification accuracy. This study shows that machine learning could be considered a suitable strategy in developing models that can effectively assess the risk of delays in tall building projects. It could be concluded that the model developed in this study can actively support risk-based decision making in tall building projects. Although, the model is based on data gathered from tall building professionals in the GCC region, the approach used can be adapted to other construction climates globally. Future work may seek to incorporate machine learning-based delay risk assessment models into state of the art construction risk management systems.

Data availability statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Abd El-Razek ME, Bassioni HA, Mobarak AM. 2008. Causes of delay in building construction projects in Egypt. J Constr Eng Manage. 134(11):831–841.

Aibinu AA, Odeyinka HA. 2006. Construction delays and their causative factors in Nigeria. J Constr Eng Manage. 132(7):667–677.

Akande KO, Owolabi TO, Olatunji SO. 2015. Investigating the effect of correlation-based feature selection on the performance of support vector machines in reservoir characterization. J Nat Gas Sci Eng. 22:515–522.

AlSehaimi A, Koskela L, Tzortzopoulos P. 2013. Need for alternative research approaches in construction management: case of delay studies. J Manage Eng. 29(4):407–413.

Assaf SA, Al-Khalil M, Al-Hazmi M. 1995. Causes of delay in large building construction projects. J Manage Eng. 11(2):45–50.

Attal A. 2010. Development of neural network models for prediction of highway construction cost and project duration [Doctoral dissertation]. Ohio University.

Aziz RF, Abdel-Hakam AA. 2016. Exploring delay causes of road construction projects in Egypt. Alexandria Eng J. 55(2):1515–1539.
Bayram S. 2017. Duration prediction models for construction projects: In terms of cost or physical characteristics? KSCE J Civ Eng. 21(6): 2049–2060.
Benzécri JP. 1973. L'analyse des données. Vol. 2. Paris: Dunod; p. l.
Brownlee J. 2018. Machine learning mastery with Weka. https://machinelearningmastery.com/machinelearning-mastery-weka/.
Budayan C, Dikmen I, Talat Birgonul M, Ghaziani A. 2018. A computerized method for delay risk assessment based on fuzzy set theory using MS Project™. KSCE J Civ Eng. 22(8):2714–2712.
Choi S, Kim YJ, Briceno S, Mavris D. 2016. Prediction of weather-induced airline delays based on machine learning algorithms. In 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC). IEEE. p. 1–6.
Cortes C, Vapnik V. 1995. Support-vector networks. Mach Learn. 20(3): 273–297.
CTBUH. 2014. Dreams deferred: unfinished tall buildings. CTBUH J. 4:46–47.
Czarnigowska A, Sobotka A. 2014. Estimating construction duration for public roads during the preplanning phase. EPPM-J. 4(1):26–35.
Darko A, Chan AP, Adabre MA, Edwards DJ, Hosseini MR, Ameyaw EE. 2020. Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities. Autom Constr. 112:103081.
Doloi H, Sawhney A, Iyer KC, Rentala S. 2012. Analysing factors affecting delays in Indian construction projects. Int J Project Manage. 30(4): 479–489.
El-Kholy AM. 2019. Exploring the best ANN model based on four paradigms to predict delay and cost overrun percentages of highway projects. Int J Constr Manage. :1–19.
Enshassi A, Al-Najjar J, Kumaraswamy M. 2009. Delays and cost overruns in the construction projects in the Gaza Strip. J Fin Man Prop Cons. 14(2): 126–151.
Faridi AS, El-Sayegh SM. 2006. Significant factors causing delay in the UAE construction industry. Constr Manage Econ. 24(11):1167–1176.
Fellows RF, Liu AM. 2015. Research methods for construction. West Sussex (UK): Wiley.
Fix E, Hodges J. 1951. Discriminatory analysis, nonparametric discrimination: Consistency properties. Texas: USAF School of Aviation Medicine, Randolph Field. Technical Report 4.
Fugar FD, Agyakwah-Baah AB. 2010. Delays in building construction projects in Ghana. CEB. 10(1/2):103.
García de Soto B, Agustí-Juan I, Joss S, Hunhevicz J. 2019. Implications of Construction 4.0 to the workforce and organizational structures. Int J Constr Manage. :1–13.
Gondia A, Siam A, El-Dakhakhni W, Nassar AH. 2020. Machine learning algorithms for construction projects delay risk prediction. J Constr Eng Manage. 146(1):04019085.
Gündüz M, Nielsen Y, Özdemir M. 2013. Quantification of delay factors using the relative importance index method for construction projects in Turkey. J Manage Eng. 29(2):133–139.
Gunduz M, Nielsen Y, Ozdemir M. 2015. Fuzzy assessment model to estimate the probability of delay in Turkish construction projects. J Manage Eng. 31(4):04014055.
Hossen MM, Kang S, Kim J. 2015. Construction schedule delay risk assessment by using combined AHP-RII methodology for an international NPP project. Nucl Eng Technol. 47(3):362–379.
International Organization for Standardization Technical Committee. 2018. Risk management-guidelines (Standard No. ISO 31000: 2018). Washington, DC: International Organization for Standardization.
Kaming PF, Olomolaiye PO, Holt GD, Harris FC. 1997. Factors influencing construction time and cost overruns on high-rise projects in Indonesia. Constr Manage Econ. 15(1):83–94.
Kim S-Y, Tuan NV, Ogunlana SO. 2009. Quantifying schedule risk in construction projects using Bayesian belief networks. Int J Project Manage. 27(1):39–50.
Kim YJ, Choi S, Briceno S, Mavris D. 2016. A deep learning approach to flight delay prediction. In 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC). IEEE. p. 1–6.
Kuncheva LI, Rodríguez JJ. 2014. A weighted voting framework for classifiers ensembles. Knowl Inform Syst. 38(2):259–275.
Larrañaga P, Atienza D, Diaz-Rozo J, Ogbechie A, Puerto-Santana CE, Bielza C. 2018. Industrial applications of machine learning. Boca Raton (FL): CRC Press.
Lek S, Guégan JF. 1999. Artificial neural networks as a tool in ecological modelling, an introduction. Ecol Model. 120(2–3):65–73.
Lo TY, Fung IW, Tung KC. 2006. Construction delays in Hong Kong civil engineering projects. J Constr Eng Manage. 132(6):636–649.
Mahamid I, Bruland A, Dmaidi N. 2012. Causes of delay in road construction projects. J Manage Eng. 28(3):300–310.
McCulloch WS, Pitts W. 1943. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys. 5(4):115–133.
Moon KS. 2015. Supertall Asia/Middle East: Technological responses and contextual impacts. Buildings. 5(3):814–833.
Muneeswaran G, Manoharan P, Awoyera PO, Adesina A. 2018. A statistical approach to assess the schedule delays and risks in Indian construction industry. Int J Constr Manage. :1–12.
Ogunlana SO, Promkuntong K, Jearkjirm V. 1996. Construction delays in a fast-growing economy: comparing Thailand with other economies. Int J Project Manage. 14(1):37–45.
Olatunji SO. 2017. Extreme learning machines and support vector machines models for email spam detection. In Canadian Conference on Electrical and Computer Engineering. IEEE. p. 1–6.
Peško I, Mučenski V, Šešlija M, Radović N, Vujkov A, Bibić D, Krklješ M. 2017. Estimation of costs and durations of construction of urban roads using ANN and SVM. Complexity. 2017:1–13.
PricewaterhouseCoopers. 2017. Innovation for the earth. Davos: World Economic Forum. Technical Report 161222-113251-LA-OS.
Sambasivan M, Soon YW. 2007. Causes and effects of delays in Malaysian construction industry. Int J Project Manage. 25(5):517–526.
Sanni-Anibire MO, Mohamad Zin R, Olatunji SO. 2020. Causes of delay in the global construction industry: a meta analytical review. Int J Constr Manage. 1–13.
Sethi H, Goraya A, Sharma V. 2017. Artificial intelligence based ensemble model for diagnosis of diabetes. Int J Adv Res Comp Sci. 8(5):1540–1548.
Sweis G, Sweis R, Abu Hammad A, Shboul A. 2008. Delays in construction projects: The case of Jordan. Int J Project Manage. 26(6):665–674.
Takeichi N, Kaida R, Shimomura A, Yamauchi T. 2017. Prediction of delay due to air traffic control by machine learning. In AIAA Modeling and Simulation Technologies Conference. p. 1323.
Tharwat A. 2018. Classification assessment methods. Appl Comput Inf. 1–13.
Toor S-U-R, Ogunlana S. 2008. Problems causing delays in major construction projects in Thailand. Constr Man Econ. 26(4):395–408.
Tumi SAH, Omran A, Pakir AHK. 2009. Causes of delay in construction industry in Libya. In: The International Conference on Economics and Administration; p. 265–272.
Wauters M, Vanhoucke M. 2017. A nearest neighbour extension to project duration forecasting with artificial intelligence. Eur J Oper Res. 259(3): 1097–1111.
Witten IH, Frank E, Hall MA. 2011. Data mining: practical machine learning tools and techniques. Burlington, Massachusetts, USA: Morgan Kaufmann Publishers.
Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, McLachlan GJ, Ng A, Liu B, Yu PS, et al. 2008. Top 10 algorithms in data mining. Knowl Inform Syst. 14(1):1–37.
Xia R, Zong C, Li S. 2011. Ensemble of feature sets and classification algorithms for sentiment classification. Inf Sci. 181(6):1138–1152.
Yaghini M, Khoshraftar MM, Seyedabadi M. 2013. Railway passenger train delay prediction via neural network model. J Adv Transp. 47(3):355–368.
Yu B, Song X, Guan F, Yang Z, Yao B. 2016. k-Nearest neighbor model for multiple-time-step prediction of short-term traffic condition. J Transp Eng. 142(6):04016018.