Building Engineering 2024, 2(1), 544.
https://doi.org/10.59400/be.v2i1.544

Review

Machine learning algorithms for safer construction sites: Critical review

Yin Junjia*, Aidi Hizami Alias, Nuzul Azam Haron, Nabilah Abu Bakar

Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang 43400, Malaysia
*** Corresponding author:** Yin Junjia, gs64764@student.upm.edu.my

CITATION

Junjia Y, Alias AH, Haron NA, Abu Bakar N. Machine learning algorithms for safer construction sites: Critical review. Building Engineering. 2024; 2(1): 544. https://doi.org/10.59400/be.v2i1.544

ARTICLE INFO

Received: 18 February 2024
Accepted: 6 April 2024
Available online: 26 April 2024

Abstract: Machine learning, a key thruster of Construction 4.0, has seen exponential publication growth in the last ten years. Many studies have identified ML as the future, but few have critically examined the applications and limitations of various algorithms in construction management. Therefore, this article comprehensively reviewed the top 100 articles from 2018 to 2023 about ML algorithms applied in construction risk management, provided their strengths and limitations, and identified areas for improvement. The study found that integrating various data sources, including historical project data, environmental factors, and stakeholder information, has become a common trend in construction risk. However, the challenges associated with the need for extensive and high-quality datasets, models' interpretability, and construction projects' dynamic nature pose significant barriers. The recommendations presented in this paper can facilitate interdisciplinary collaboration between traditional construction and machine learning, thereby enhancing the development of specialized algorithms for real-world projects.

Keywords: machine learning (ML); construction; risk management; critical review; Litmaps®; Open knowledge maps®

1. Introduction

Occupational safety has always been a headache for many workers in high-risk industries, such as high-voltage electricians, tower crane drivers, and deep-well miners, to name a few [1,2]. Accidents, whether artificial or not, can cause significant loss of life and property, as well as immense psychological grief for families. According to the latest data (11 January 2024) from the International Labor Organization (ILO), many countries such as Costa Rica, Argentina, Chile, France, and Denmark had at least 9421, 3587, 3142, 3043, and 2814 injuries per 100,000 workers, respectively [3]. This has also led to a focus on research around risk. Construction risk is complex and dynamic, arising from inherent uncertainties and variability [4]. Given the involvement of many stakeholders in large construction projects, including owners, financial institutions, project managers, designers, construction crews, manufacturers, suppliers, labor, insurance agencies, legal counsel, and public and regulatory agencies [5,6]. Effective risk management is crucial for their successful execution. Risk assessment is a cornerstone of construction management, involving identifying, analyzing, and mitigating potential risks that may arise during various project phases [7,8]. Historically, construction risk management was regarded as a static process at the project initiation stage [9,10]. However, constant changes in construction methods, coupled with a growing recognition of risk volatility, have led to a paradigm shift in risk management approaches [11]. The contemporary perspective acknowledges that risks are not fixed but evolve over the project lifecycle [12]. Therefore, it is imperative to adopt a dynamic and proactive approach to risk management that accounts for the evolving nature of risks and adapts to changing project circumstances.

Fuzzy set theory (FST) was introduced by scholars in 1989 as a means of linguistic risk assessment for construction [13,14]. This approach enabled analysts to communicate the level of risk associated with individual project elements to stakeholders in easily understandable linguistic terms [15]. On the other hand, OSHA employs the risk matrix, a standard methodology for risk assessment. However, subjective approaches to risk assessment, such as those based on historical accidents, rely heavily on personal knowledge, experience, intuitive judgment, and rules of thumb. Many risk models are still based on expert opinion and emphasize linear causality, making it challenging to incorporate non-linear relationships such as security commitments and organizational culture [16,17]. Recent studies have highlighted the power of ML. As shown in **Figure 1**, according to the most recent data from the Scopus database, articles on ML grew exponentially from 2014 to 2023, reaching 131,489 from 11,264. Many scholars have also recognized that the rapidly changing algorithms have led to the need to regularly review research about the application of ML in construction risk. However, few studies critically examined the ML algorithm in this niche. The collective wisdom of the domain should be continuously constructed and updated, as knowledge is dynamically changing and growing and is contributed by multiple domain experts [18].

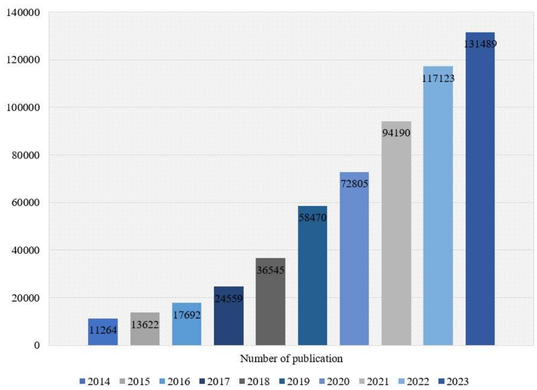

Figure 1. Machine learning annual publications (2014–2023).

Therefore, this article enriches the knowledge system in the following ways: 1) provides scientific and researched empirical evidence for engineers in chaotic working environments by reviewing existing algorithms' purpose, background, and limitations; 2) promotes the iteration of machine learning knowledge in construction risk by critically reviewing previous research; and 3) helps practitioners choose appropriate computational methods to formalize complex engineering knowledge. This study sheds light on the current main ML algorithms in construction risk by selecting 100 powerfully relevant and high-level research articles published between 2018 and 2023. The rest of the paper is structured as follows: Section 2 provides an overview of machine learning. Section 3 introduces the research methodology, and Section 4 introduces each algorithm's application, advantages, and disadvantages in construction risk assessment. Section 5 discusses future improvements in ML. Section 6 summarizes the research findings of this article and gives recommendations.

2. Background

Machine learning (ML) intersects several disciplines, including computer science, statistics, and artificial intelligence developments [19]. It addresses the problem of constructing computer algorithms and models that enable computer systems to improve automatically based on experience and increase performance on specific tasks [20,21]. Fundamentally, the beauty of ML is the ability to analyze, predict, and make decisions based on known data with increasing accuracy if the data sample expands [22]. As a result, it is popular in many data-intensive industries, such as the chemical industry, where active ML has been used to optimize the performance of CO_2 electrocatalysts [23]. In the medical field, Warnat-Herresthal et al. [24] used ML to identify patients with leukemia based on their blood transcriptome.

ML encompasses three fundamental types: supervised, unsupervised, and reinforcement learning [25]. One of the significant challenges that ML models face is overfitting and underfitting. Overfitting occurs when a model matches too closely to the training data, leading to inadequate performance on new and unseen data. On the other hand, underfitting occurs when a model is too simple to capture the underlying data patterns [26]. Supervised learning algorithms are designed to learn from labeled datasets, where each input data point is associated with a corresponding label [27]. The algorithm then learns to map input data to the correct output by adjusting its parameters based on the error between the prediction and the actual label [28]. Where features are the input variables, and labels are the outputs or predictions that the model is trying to learn. For example, algorithms make decisions based on historical data on contractor bidding opportunities, and the individual representations of project characteristics that enable the system to make decisions are called features. Unsupervised learning involves training on unlabeled datasets where the system automatically explores patterns in the data without guidance, such as clustering and lowering dimensionality. Reinforcement learning, on the other hand, trains a model and makes it make decisions by interacting in a scenario. The model is continually modified based on feedback from decisions such as rewards and penalties, aiming to maximize cumulative rewards over time.

3. Methods and materials

Compared to the traditional retrieval methods (from Scopus or WoS), we try to adopt a two-stage meta-analytic paper retrieval method to provide a new reference for future literature retrieval. In the first step, our approach used Open knowledge maps® to sift through the top hundred papers on “construction risk management” and “machine learning” efficiently [29]. By grouping these documents based on their metadata, including title, abstract, author, journal, and subject keywords, we can create a word co-occurrence matrix to determine the relevance of each article. As shown in **Figure 2**, the resulting map represents the textual similarity between each article and the search query. The proximity of circular regions on the map indicates how closely related their subjects are, with more central areas indicating more remarkable similarity to the overall topic. Using this method, we can effectively manage the number of documents to review while exploring a wide range of content.

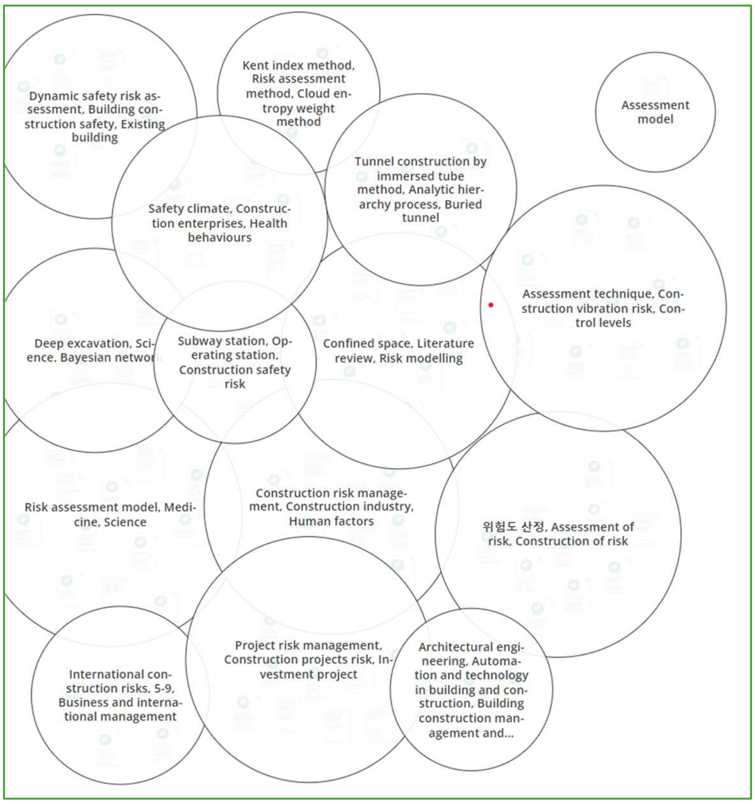

Figure 2. Top 100 Strong-related articles about ML in construction risk.

In the second step, **Figure 3** comprehensively showed the top 20 citations and references associated with the subject matter using Litmaps®. The inner circle of the map represents the input, namely "Construction Risk Assessment" and "Machine Learning," and its corresponding combination of citations and references. The outer circle demonstrates the articles' findings that are most pertinent to the domain. After completing the literature search, we collected data from the selected literature, including key findings, methods, applications of ML techniques, challenges addressed, and innovations introduced [30]. This information was then systematically organized and cataloged for analysis. Identify patterns, trends, and common themes in the selected literature by analyzing the collected data. Finally, we will compare ML algorithms and their effectiveness in construction risk management.

Figure 3. Top 20 high-cited articles about ML in construction risk.

4. Results

The keyword cloud in **Figure 4** is generated based on the frequency statistics of documents. It is intended to facilitate the reader's understanding of the current state of ML in the field. Standard algorithms include SVM, logistic regression, and ANN. They are utilized in almost every risk management process, including risk identification, classification, assessment, diagnosis, and prevention. The study summarizes the main ML algorithms in construction risk, as illustrated in **Table 1**. Since the author's past research has discussed artificial neural networks and Bayesian networks, this article will not discuss them.

Table 1. ML in construction risk management.

Algorithm	Application	Source
Regression	Mapping landslide sensitivity; project delay risk prediction; predicting construction duration; Credit risk assessment models for financial institutions; Analysis of ground settlement during tunnel construction; Predicting variance in construction productivity; Determination of poor compliance with OSH rules of construction workers; long-term probabilistic prediction of rock burst hazard.	Tessema et al. [31]; Zhu et al. [32]; Gariazzo et al. [33]; Hemasinghe et al. [34]; Li and Jimenez [35].
RF	Integrated land carrying capacity assessment; multi-objective optimization of shield construction parameters; Constructing a monitoring model for dam safety; Predicting BIM labor cost; Detecting corporate misconduct; Activity recognition of construction equipment; concrete dam deformation monitoring; Analyzing and adjusting EPB shield steering in real-time; hybrid optimization of seismic performance of mountain buildings.	Xie et al. [36]; Hu et al. [37]; Wang et al. [38]; Wu et al. [39]; Wen et al. [40].
SVM	Predicting project outcomes; Rapid building fire risk assessment; projects delay risk prediction; Hypertension risk assessment for steelworkers in deep foundation pits; Estimation of construction waste generation; Seismic hazard safety evaluation of existing buildings; Early cost estimates of bridges; Estimation of construction waste generation.	Hu et al. [41]; Chen and Lin [42]; Tserng et al. [43]; Fan and Sharma [44]; Fu et al. [45].
GCN	Boring machinery load prediction in tunnel excavation; Interaction Behaviors Identification of Construction Workers; Identification of accident-injury type and body part factors; Action recognition of construction workers under occlusion; Determining construction method patterns to automate and optimize scheduling; Monitoring and prediction of landslide-related deformation.	Mostofi et al. [46]; Khalili et al. [47]; Mostofi et al. [48]; Fu et al. [49]; Li et al. [50]; Zhang et al. [51]
KNN	Projects delay risk prediction; Safety risk evaluations of deep foundation construction schemes; Estimation of management reserve; Assessing worker perceived risk; Analysis of factors influencing rockfall runout distance; Short-term rockburst risk prediction for profound underground works;	Chen et al. [52]; Pandey and Bandhu [53]; Jaber et al. [54]; Lee et al. [55].
Apriori	Analysis of deformation response to landslide disaster; Mining geological disaster sensitivity evaluation indicators; Mining Construction Cross-Operation Risk Association Rules.	Linwei et al. [56]; Chen et al. [57]; Chen et al. [58]
PCA	Extraction of construction accident characteristics; Analysis of crucial behavioral risk factors for construction practitioners; Explore construction settlement data; Identify and remove outliers.	Shao et al. [59]; Xiang et al. [60]; Siddiqui et al. [61]
XGBoost	Handling large datasets; Predicting enterprise financial management risks; Investment Estimates for Assembled Concrete Buildings; Predict construction cost overruns; Investment estimation of prefabricated concrete buildings.	Yan et al. [62]; Cherif and Kortebi [63]; Coffie and Cudjoe [64]; Liu et al. [65]
K-Means	Identifying clusters of projects with similar risk profiles; Early warning of risks in government investment and construction projects; Supplier risk assessment; Risk assessment of integrated pipeline corridor construction projects; BIM performance assessment system; Identifying high frequency-low severity construction safety risks.	Liu and Li [66]; Wang et al. [67]; Ayhan and Tokdemir [68]
ARIMA	Predicting construction cost index; Predicting construction material prices; Forecasting the ratio of a low bid to owner's estimate for highway construction; Effect of dam construction on the lake; Structural health monitoring and identification; Predicting perceived fatigue levels.	Kim et al. [69]; Moon et al. [70]; Ghashghaie and Nozari [71]; Kaloop et al. [72]; Hajifar et al. [73]

Figure 4. Keyword cloud of the documents.

4.1. Linear and logistic regression

Linear regression (LiR) is a statistical technique that models the connection between a dependent variable and one or more independent variables [74]. It aims to create a linear equation that predicts the dependent variable based on the independent variable values [75]. Equation (1) shows that multiple regression analysis integrates numerous independent variables throughout the dataset without considering shared variance, and R^2 assesses reliability. LiR helps analyze the impact of various factors (independent variables) on a specific risk outcome (dependent variable). Huang & Hsieh [76] employed LiR to forecast BIM labor costs. It produces a simple and understandable model that explains each independent variable's influence on predicting risk outcomes. Secondly, LiR is computationally efficient and straightforward, suitable for swift analysis and evaluation. Thirdly, LiR equation coefficients measure the strength and direction of the relationship between the independent and dependent variables. Still, it is unsuitable for capturing complex, variable relationships or interactions. Additionally, linear regression is sensitive to outliers, which can significantly affect model parameters and predictions [77]. Therefore, LiR is not ideal for nonlinear relationships.

$$Y = a + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n \pm e \tag{1}$$

here Y could be the risk factor variable, a is a constant, b_1 through b_n are estimated regression coefficients, x_1 through x_n are predictor or independent variable values, and e is the closure error.

Logistic regression (LoR) is a statistical technique that predicts the probability of a binary outcome (0 or 1) based on one or more predictor variables, as demonstrated in Equation (2). LoR can be utilized to model and predict the probability of a specific risk event happening or not [78]. Levy and Baha [79] employed LoR models to anticipate borrower solvency. Akboğa and Baradan [80] employed LoR modeling to identify factors contributing to construction injury severity. Bhattacharjee et al. [81] used an interval based LoR and FMEA to assess potential failures. Both regression coefficients are interpretable, making it easy to understand how each predictor variable impacts the probability of a risk event occurring. LoR is restricted to binary outcomes, designed for binary outcomes, and needs to be modified or expanded for multi-category outcomes [82].

$$y = \frac{1}{1 + e^{-z}} = \frac{1}{1 + e^{-(w^T x + b)}} \tag{2}$$

where z is a real number, w is a column vector, x is a row vector, and b is a real number. $w^T x$ denotes the inner product of w and x. e is the base of the natural logarithm. The function y has a range of (0,1). When y is greater than 0.5, we consider the input data belong to the positive category; otherwise, we consider the input data to belong to the negative category. In addition, the loss function of logistic regression is the cross-entropy loss function.

4.2. Principal component analysis

Principal Component Analysis (PCA) is a statistical technique used for dimensionality reduction in multivariate data analysis. Its primary purpose is to transform the original variables into a new set of uncorrelated variables, known as principal components, which capture the maximum variance in the data. PCA can identify the data set's most influential patterns or features by arranging these components in descending order of variance. This dimensionality reduction method simplifies the data while preserving its essential features, making it valuable in various fields such as image processing, pattern recognition, and data compression. The specific formula is as follows:

$$x = \frac{1}{n}\sum_{i=1}^{N} x_i \tag{3}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{4}$$

$$\mathrm{Cov}(X, Y) = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - x)(y_i - y) \tag{5}$$

here, x is mean, S^2 isvariance, $\mathrm{Cov}(X, Y)$ is covariance.

Zhang et al. [83] proposed the weight calculation method of Group Analytic Hierarchy Process-Principal Component Analysis to rank the critical construction risk factors. Most studies use PCA as a data preprocessing tool [84,85]. Nevertheless, it also has some disadvantages. First, it is often difficult to directly interpret the specific meaning of the principal components obtained by PCA. Although it can map high-dimensional data to a low-dimensional space, the meaning of the comprehensive evaluation function is unclear when the sign of the factor loading of each principal component is positive or negative. Secondly, PCA is sensitive to outliers, which may cause the extracted principal components to deviate from the actual situation. Finally, PCA assumes that the data follows a Gaussian distribution. If the data distribution does not conform to this assumption, it may result in inaccurate analysis.

4.3. Support vector machines

The Support Vector Machine (SVM) is a powerful binary classification model that utilizes a linear classifier to optimize the feature space [86]. SVM's primary learning strategy involves interval maximization, which can be formulated as a convex quadratic programming problem [87]. This is also equivalent to minimizing a regularized hinge loss function. It is also an optimization algorithm used to solve convex quadratic programming. The fundamental idea behind SVM is to locate a separating hyperplane that accurately separates the training dataset while maximizing the geometric intervals [88]. The basic idea is to solve for a separating hyperplane that correctly divides the training dataset and maximizes the geometric separation [89]. As shown in Equation (6), $\mathrm{w} \cdot \mathrm{x_i} + \mathrm{b} = 0$ is the separating hyperplane, and there are infinitely many such hyperplanes (i.e., perceptual machines) for a linearly divisible dataset. Still, the geometrically maximally spaced separating hyperplane is unique. The algorithmic formulation of the SVM is as follows [90]:

$$\begin{gathered} \mathrm{minw, b}\frac{1}{2} \parallel \mathrm{w} \parallel^2 \\ \text{s.t. } \mathrm{y_i}(\mathrm{w} \cdot \mathrm{x_i} + \mathrm{b}) \geq 1, \mathrm{i} = 1,2, \ldots, \mathrm{n} \end{gathered} \tag{6}$$

here x_i is the feature vector of the ith sample; $\mathrm{y_I}$ is the category labeling of the ith sample, taking the value of +1 or −1.

It can be applied to classify risks such as credit into different categories based on input features [91,92]. Gong et al. [93] used a binary particle swarm optimization algorithm to reduce the redundancy of information in the dataset. Then, they modified the classification algorithm using an Adaboost-enhanced support vector machine classifier, which overcame the difficulties of correctly classifying a small number of samples in an unbalanced dataset. SVM has gained popularity in recent years due to its effectiveness in high-dimensional feature space and its ability to handle complex patterns in data. It is well-suited for scenarios where a clear margin of separation exists between different classes or categories, making it a valuable tool for risk classification tasks such as credit risk assessment.

Several studies have explored the use of SVM in various risk assessment applications. For example, Liu et al. [94] developed an SVM model based on particle swarm optimization to predict the safety risk of metro construction, achieving an average accuracy of 85.26%. Wei et al. [95] proposed a new rapid-fire risk assessment method based on fuzzy mathematics and an SVM algorithm. Additionally, researchers have attempted to enhance the performance of SVM by integrating it with other algorithms, such as the firefly algorithm and Gradient Boosting Decision Tree [96–98]. While SVM has several advantages, it also has some limitations that should be carefully considered. For instance, SVM is computationally intensive, especially for large data sets or complex kernel functions, which may affect scalability. It is also sensitive to noisy data, outliers, and mislabeled instances, which can significantly impact model performance and generalization [99]. Finding the optimal set of hyperparameters for SVM requires careful tuning, which can be challenging.

4.4. Random forest

Random Forest (RF) is an ensemble ML algorithm that has become increasingly popular in various ML applications, including classification and regression tasks [100]. As shown in **Figure 5**, the algorithm constructs many decision trees during training and outputs pattern (classification) or mean (regression) predictions for individual decision trees [101]. In construction risk, RF has been used to predict and classify different types of risks based on relevant input features, such as falls from heights [102,103]. Studies have shown that RF models can effectively estimate the relationship between monitoring values and pit safety risk and predict and prevent occupational accidents [104,105]. RF has also been used to develop risk indicators with high accuracy in various fields, such as supply chain finance risk [106], flood risk [107], and landslide risk assessment [108]. These studies have demonstrated that RF is a robust algorithm that can handle missing data efficiently, including incomplete data sets in the assessment process [109–111]. Scholars have also proposed a fractional RF method with low dependency on a comprehensive training dataset that can predict extensive device activities using a small amount of training data [112,113].

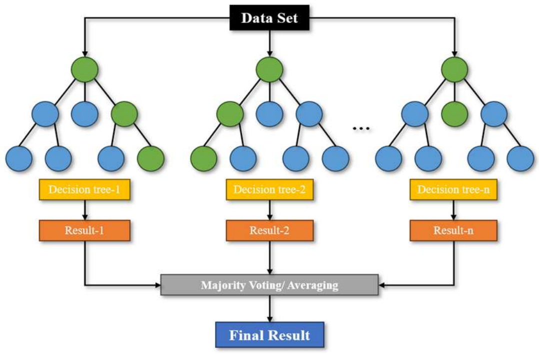

Figure 5. Random forest model.

Furthermore, RF provides a measure of feature importance that can help identify the most influential risk factors [114]. It is also robust to noisy data and outliers, and aggregating predictions from multiple trees eliminates individual errors and outliers [115]. Additionally, training individual trees in an RF can be done in parallel, making it computationally more efficient, especially for large data sets [116]. However, it is essential to note that RF may be biased toward the dominant class in the training data, leading to imbalanced predictions if the class distribution is imbalanced [117]. Moreover, while individual trees can be trained in parallel, the overall construction of an RF can be computationally expensive, particularly for large numbers of trees [118]. The diversity measure between the decision trees improves the model's generalization, but it is still necessary to minimize the number of trees to find the optimal subset [119].

4.5. K-nearest neighbor

The K-nearest neighbor (KNN) algorithm is a powerful instance-based learning method that can be utilized for both classification and regression problems. The fundamental concept behind the KNN algorithm is to locate the k closest instances to a new input instance in the training set and subsequently predict the instance's class based on the classes of these k instances. The Equation (4) of the KNN algorithm is as follows [120]:

$$y = \underset{c_j}{\mathrm{argmax}} \sum_{i=1}^{k} w(i) \cdot I\left(y_i = c_j\right) \tag{7}$$

here y denotes the predicted category, c_j denotes the jth category, $w(i)$ denotes the weight of the distance $d(x, I)$ from the input instance x, and $I\left(y_i = c_j\right)$ is the indicator function, when $y_i = c_j$, $I\left(y_i = c_j\right) = 1$, otherwise $I\left(y_i = c_j\right) = 0$.

In KNN, an object is classified by the majority class of its k nearest neighbors, where "k" is a user-defined parameter. Construction risk can be applied to categorize risks or predict risk outcomes based on the characteristics of similar historical cases. Lee et al. [121] used it to retrieve similar projects and a genetic algorithm to optimize the retrieved cases with an error rate of less than 5%. Kamran et al. [122] reduced the magnification of the original database using the state-of-the-art method of the Isometric Mapping (ISOMAP) algorithm; it then used the Fuzzy c-Mean (FCM) algorithm to classify the datasets obtained from ISOMAP, and thirdly, employed it to predict the short-term rock burst datasets at different levels of accuracy, with an accuracy of 96%. Liu et al. [123] constructed an improved fusion KNN model to evaluate the posture state of workers.

KNN is a straightforward and practical method for quickly assessing risk, mainly when interpretability is crucial. Moreover, it is suitable for analyzing data with an unknown or complicated distribution since it does not rely on making assumptions about the underlying data [124]. KNN can also detect local patterns, making it an effective tool for identifying risks with spatial or temporal clustering. However, it is computationally demanding, mainly when applied to large datasets, as it requires computing the distance between the query and all training instances. KNN is also sensitive to outliers, as extreme values in the dataset can affect the nearest neighbors. In addition, irrelevant or redundant features can introduce noise into distance calculations and compromise the performance of KNN [125,126]. In high-dimensional space, the distance between instances tends to become more uniform, which may reduce the effectiveness of KNN. Finally, the choice of parameter “k” (number of neighbors) can impact KNN’s performance and may need to be adjusted based on the data’s specific characteristics [127]. For instance, Zhang et al. [128] used a weighted k-value to plan deep foundation pits.

4.6. XGBoost

XGBoost is a refined algorithm rooted in GBDT. While sharing the basic concept of GBDT, it incorporates several enhancements, including second-order derivatives for greater loss function accuracy, regularization terms to address tree overfitting, and block storage for parallel computation [129]. Its objective function comprises a loss function and a regularization term. The loss function can be the mean square error (MSE) for regression problems or cross-entropy for classification problems. Qin [130] predicted corporate financial risk and found that the model’s errors were all within 3%, with the maximum prediction error of only 2.68%. In another study, Liu et al. [131] assessed pipeline safety using it and achieved an accuracy of 91.8%. The algorithm analyzes the feature’s importance, which helps prioritize risk factors in decision-making. Including regularization terms in the objective function helps prevent overfitting and improves the model’s generalization [132]. It is designed for parallel and distributed computing, efficiently handling large building datasets [133]. However, it faces the challenge of interpretability, and data preprocessing is necessary to handle missing values and outliers for optimal performance.

In the future, it may be possible to use more superficial ensemble structures to enhance the interpretability of XGBoost models. The development of automatic hyperparameter tuning methods can simplify the model development process and improve the algorithm’s ease of use [134]. Construction projects involve data that changes over time, and improving XGBoost's ability to process time series data directly could enhance its applicability to construction risk management. However, the complexity and dynamic interrelationships of the studied attributes make it difficult for the XGBoost model to predict residual values [135].

4.7. *K*-means

The *k*-means algorithm is a distance-based clustering algorithm. Its steps include [136]: 1) randomly initialize k centers of mass, i.e., the centroids of the k clusters; 2) for each sample, calculate its distance from the k centers of mass and assign it to the cluster with the closest distance; 3) for each cluster, recalculate its center of mass; 4) repeat steps 2 and 3 until the center of mass no longer changes or a preset number of iterations is reached. Distance can be used as Euclidean distance, Manhattan distance, etc. *K*-means can group construction projects based on shared risk characteristics, which allows risk profiles to be created for different projects and can also help identify geographic or project-specific "hot spots" where specific risks are more prevalent [137]. This information is valuable for resource allocation. Many academics use *K*-mean clustering to identify similarities between different construction projects based on risk factors, which can help with benchmarking. Evolving risk patterns are uncovered by regularly updating clusters and reassessing risks.

K-means is computationally efficient and relatively simple to implement, making it suitable for quick analyses and real-time applications. The algorithm scales to large datasets, making it ideal for large-scale complex projects [138]. Being an unsupervised learning algorithm means it does not require labeled data, making it adaptable to situations where comprehensive risk labeling is not readily available. Each item is assigned to a cluster, providing precise categorization and simplifying the interpretation of results. However, the results of *K*-means are sensitive to the initial position of the centroids. Different initializations may lead to other solutions and finding the optimal centroids can be challenging. The algorithm assumes that the clusters are spherical and of equal size, making it difficult to identify irregularly shaped clusters or clusters with different densities, which are common in construction risk datasets [139]. Improvements have also been made to the *K*-means method to deal with non-spherical or irregularly shaped clusters, improving its applicability in various construction risk situations. For example, developing strategies to automatically determine the optimal number of clusters (K) could alleviate the sensitivity to the initial choice of centroids and improve the usability of the algorithm.

4.8. ARIMA

ARIMA (Autoregressive Integrated Moving Average) is a time-series forecasting algorithm that can make time-series forecasts of construction-related variables such as project cost, completion time, or other performance indicators [140]. Equation (8) of the ARIMA model is given below:

$$Y_t = c + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \tag{8}$$

here, Y_t denotes the observed value at time t and c is a constant; $\phi_1, \cdots, \phi_p$ is the autoregressive coefficient; ϵ_t is white noise; $\theta_1, \cdots, \theta_q$ is the moving average coefficient. The parameters p, d, q of the ARIMA model denote the number of autoregressive terms, difference order and moving average terms, respectively.

ARIMA provides a quantitative basis for assessing the likelihood of delays, cost overruns, or other adverse events. It can assist in resource planning by predicting the demand for construction materials, labor, and equipment based on historical usage patterns [141]. ARIMA explicitly accounts for the time dependence of the data and is, therefore, well-suited to modeling construction-related variables that evolve [142]. Secondly, ARIMA is very robust when dealing with noisy time series data and is, therefore, suitable for situations where construction project data may be subject to variability and uncertainty [143]. The parameters and results of the model can usually be interpreted to give an understanding of the impact of past observations on future projections. It is very effective for univariate time series data, which is common in construction risk analysis, where univariate variables (e.g., project duration or cost) are often the focus of the study. However, ARIMA assumes that the underlying data patterns are linear and require the time series data to be static, which is challenging in dynamic yet complex construction systems. In addition, because it focuses primarily on internal time-series patterns, it is easy to overlook external factors or unexpected shocks to a construction project. Therefore, incorporating suitable exogenous variables is a topic worth considering.

4.9. Graph convolutional network

Graph Convolutional Network (GCN) is a deep learning algorithm that operates on graph-structured data. Equation (9) demonstrates the algorithm of GCN [144]:

$$H^{(l+1)} = \sigma\left(\widetilde{D}^{-\frac{1}{2}}\tilde{A}\widetilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right) \tag{9}$$

here $H^{(l)}$ is the node feature matrix of the lth layer, $W^{(l)}$ is the weight matrix of the lth layer, $\tilde{A} = A + I$, I is the unit matrix, and $\widetilde{D}$ is the degree matrix of $\widetilde{A}$.

It can model the complex relationships and dependencies between various risk factors in a construction project and represent them as a graph [145]. Nodes can represent project risk components, while edges represent their relationships [146]. It can help identify critical nodes in a construction project, such as unsafe interactions between people, machines, and materials [147]. Temporal information can be integrated to enable dynamic risk assessment by considering the evolution of risk factors at different stages in the construction. Mostofi and Toğan [148] combined a GCN to account for the dependency information between accidents and predicted the severity outcome of each construction activity with 94% accuracy.

GCN can improve the efficiency of risk assessment by using transfer learning to train models on one construction project that apply to new projects with similar risk structures. However, GCNs require large amounts of labeled data to learn effectively, which may not always be readily available. The raw recorded data may contain noise, which reduces the prediction accuracy of the GCN deep learning model [149]. Integrating external data sources such as weather patterns, economic indicators, or regulatory changes into GCN can enhance its ability to capture external influences on building risk [150].

4.10. Apriori

Apriori algorithm is a classical association rule mining technique used in data mining and machine learning. It aims to discover frequent item sets in transaction databases and extract meaningful associations between items. The algorithm uses a bottom-up approach, starting with a single item and progressively identifying larger item sets through iterative concatenation and pruning based on a predefined support threshold. Support measures how often the itemset appears in the dataset. The strength of association rules can be measured by their support and confidence, as shown in Equations (10) and (11).

$$\text{Support}(X, Y) = \frac{\text{num}(XY)}{\text{num}(\text{allsamples})} \tag{10}$$

$$\text{Confidence}(XY) = \frac{P(X \mid Y)}{P(Y)} \tag{11}$$

Xie et al. [151] used it to mine disaster information and prevent incorrect reference management. Deng et al. [152] analyzed subway operation accident cause association rules based on the Apriori algorithm and network method. This algorithm needs to scan the data set multiple times and calculate frequent item sets, so the calculation complexity is high and the speed is slow. On large data sets, the efficiency may not be high enough. Secondly, it must store many intermediate results, requiring ample memory space. Finally, it can only handle discrete data and is powerless for continuous data.

5. Discussion

Construction projects often involve heterogeneous and incomplete data, leading to inaccurate model predictions. Much of the quantitative data is difficult to collect without the full assistance of site managers and workers, especially in China, where disruptions to the construction schedule can hinder researchers [153]. It has become the norm to model multiple ML tools simultaneously, compare their associated fit parameters, such as F1 and RSMEA, recall rate, and then select the “best performing” tool [154]. Accurately capturing and modeling this complexity is a significant challenge for traditional algorithms. Taking safety risks as an example, the construction industry lacks standardized incident text data formats and reporting practices [155]. Different data sources and formats make integrating information effectively difficult for many ML models. Construction projects change over time, with changing conditions, requirements, and stakeholders, and the many human decisions involved are not so easily quantifiable [156]. However, computer vision attempts to understand human behavior and incorporate it into ML models. Many state-of-the-art ML algorithms lack transparency, interpretability, and complex deep learning models. In risk management, stakeholders often require explanations of predictions, which can hinder adopting specific algorithms. In addition, limited computer literacy and a lack of awareness of potential advantages among construction industry practitioners in most developing countries can similarly hinder the widespread adoption of ML technologies. Data for many projects is subject to confidential contractual terms and regulatory norms, and complying with these regulations while implementing ML models can be challenging, especially if the models are seen as “black box” systems [157].

Improving the use of ML in construction risk management requires a combination of data-driven approaches, advanced algorithms, and integration with existing processes. First, there is data collection and integration. Consider collecting comprehensive, high-quality data from various sources, including project management systems, sensor data, historical project data, weather conditions, etc. Then, data from different departments and systems will be integrated to create a holistic view of the project and its associated risks. Next, fostering collaboration between building technicians and machine learning experts requires promoting mutual understanding of expertise and goals. Launching an interdisciplinary training program can help bridge the knowledge gap, enabling building technicians to grasp basic machine learning concepts and machine learning experts to understand the intricacies of building processes. In addition, joint project planning sessions and interdisciplinary teams can facilitate a holistic approach, allowing building technicians to provide real-world insights and machine learning experts to deliver tailored solutions. Construction companies need to adjust their corporate structure and set up AI represented by ML as a specialized function, which not only creates sustainable returns for the company but is also an inevitable choice not to be eliminated by the times. Besides ML, other methods also made sustained contributions, as shown in **Table 2**. It is also an excellent option to integrate ML with these techniques to form a new approach.

Table 2. Other methods in construction risk evaluation.

Source	Method	Contribution
Zhu et al. [158]	4D simulation	The methodology provides a 4D simulation environment for modeling drone interactions on a dynamic construction site.
Zhong et al. [159]	Finite Element Model (FEM)	The methodology can assess the seismic risk of bridges throughout their life cycle, including construction and use.
Nguyen et al. [160]	Hierarchical regression	The methodology provides a comprehensive list of GB risks, categorized and assessed according to the project life cycle.
Sohrabi and Noorzai [161]	PLS-SEM	The methodology is based on a project life-cycle perspective that considers the link between the risks leading to claims and the main parties involved.
Hatamleh et al. [162]	Factor analysis	The methodology identifies the risks developing countries face and emphasizes how risks can benefit industry practitioners.
Al-Mhdawi et al. [163]	Deductive and inductive reasoning	The methodology proposes new risk models for analyzing the risks associated with extreme situations such as pandemics.
Gashaw and Jilcha [164]	Fuzzy Synthesis Evaluation (FSE) and System Dynamics (SD)	The methodology considers the overall dynamics, interrelationships, and uncertainties of risks to inform the assessment of the impact of project objectives.
Do et al. [165]	Expert scoring	The methodology simultaneously examines the chain of risk factors, the sources of risk, and the scope of influence of risk factors.
He et al. [166]	AHP and Fuzzy Comprehensive Evaluation (FCE)	The method quantitatively evaluates the risk level of pyrotechnic operations and rationally ranks the importance of various risk factors.
Mohandes et al. [167]	AHP	The methodology is based on a fuzzy hybrid multidimensional model that considers the context of construction-related activities that lead to accidents and provides a comprehensive ranking system for project risk.
Badi et al. [168]	Grey Theory	The methodology identifies the risks of construction projects through preliminary research, extensive interviews with construction experts, and site visits.
Zhang and Li [169]	Projection Pursuit Method and Improved Set Pair Analysis	The method is based on the risk decomposition structure matrix, which considers the risk dynamics and establishes the deep foundation pit risk evaluation index system.
Ju et al. [170]	Best worst method (BWM) and game theory and extension cloud	The methodology considers disaster-causing factors when assessing building safety risks.
Sadeghi et al. [171]	Trapezoidal fuzzy ordinal priority approach (OPA-F)	The methodology constructs a new OPA-F using trapezoidal fuzzy numbers, assesses the blockchain risks faced by construction organizations, and develops a framework.

6. Conclusion

The number of machine learning papers is growing exponentially, and regular review is essential to promote the dissemination of interdisciplinary knowledge in the industry. This study's significance is critically reviewing ten machine learning algorithms that have been popular in construction risk over the last five years to inform new researchers. Machine learning has great potential in construction risk management but requires combining technology, data, and domain knowledge to achieve better results. Linear regression is suitable for predicting continuous numerical outcomes, such as project cost or completion time. However, it is assumed that there is a linear relationship between the variables, which may not always be accurate in complex construction projects. Logistic regression is suitable for binary classification problems, such as predicting whether a project will be completed on time or delayed. However, assuming linear decision boundaries may not capture more complex relationships in the data. SVMs are adequate for regression and classification tasks, especially when working with nonlinear and high-dimensional data. However, the choice of kernel and parameters can be sensitive, and different kernel functions, such as the triangular kernel function and the Gaussian kernel function, may perform differently or poorly on large data sets. RF is suitable for classification and regression tasks, robust enough to overfit, and can handle large datasets with many features. However, it lacks interpretability, and the training time for extensive forests can be extended. KNN is simple and effective for classification and regression tasks, especially when dealing with localized patterns. However, it is sensitive to irrelevant or redundant features and computationally expensive to predict for large datasets. PCA in construction risk analysis offers the advantage of reducing dimensionality and aiding in identifying key risk factors and patterns; however, it may oversimplify complex interactions and might not capture non-linear relationships in the data. On the other hand, the Apriori algorithm enables the discovery of association rules among construction risk factors, enhancing understanding; nevertheless, it may face challenges with large datasets and requires careful parameter tuning. XGBoost is very effective for classification and regression tasks and is often used in competitions due to its high predictive performance. However, it is computationally expensive and is easy to overfit if not correctly adjusted. K-Means is suitable for clustering similar construction projects based on characteristics such as project size, location, or complexity. However, it must be assumed that the clusters are spherical, sensitive to the initial cluster center, and may not work well for clusters of uneven size. ARIMA is suitable for time series forecasting in construction risk analysis, such as predicting future project delays or cost overruns. However, it requires assumptions of linearity and stationarity and may not capture complex nonlinear trends in time series data.

The study recommends using a combination of these algorithms to address construction risks. The choice of algorithm should depend on the specific characteristics of the data and the nature of the risk being analyzed. Proper data preprocessing, feature engineering, and hyperparameter tuning are critical to achieving optimal model performance. In addition, there is a distinct lack of standardization across the industry, leading to challenges in actual data collection. Therefore, future research should prioritize standardization efforts and seek consensus on best practices to meet project-specific needs, such as tight deadlines and confidentiality agreements.

Conflict of interest: The authors declare no conflict of interest.

References

1. Liu R, Liu HC, Shi H, et al. Occupational health and safety risk assessment: A systematic literature review of models, methods, and applications. Safety Science. 2023; 160: 106050. doi: 10.1016/j.ssci.2022.106050
2. Pinto A, Nunes IL, Ribeiro RA. Occupational risk assessment in construction industry - Overview and reflection. Safety Science. 2011; 49(5): 616-624. doi: 10.1016/j.ssci.2011.01.003
3. ILOSTAT. ILOSTAT data tools to find and download labor statistics. Available online: https://ilostat.ilo.org/data (accessed on 15 April 2024).
4. Hastak M, Shaked A. ICRAM-1: Model for International Construction Risk Assessment. Journal of Management in Engineering. 2020; 16(1): 59-69.
5. Lee HS, Kim H, Park M, Ai Lin Teo E, Lee KP. Construction Risk Assessment Using Site Influence Factors. Journal of Computing in Civil Engineering. 2012; 26(3): 319-330.
6. Mulholland B, Christian J. Risk Assessment in Construction Schedules. Journal of Construction Engineering and Management. 1999; 125(1): 8-15.
7. Ashtari MA, Ansari R, Hassannayebi E, et al. Cost Overrun Risk Assessment and Prediction in Construction Projects: A Bayesian Network Classifier Approach. Buildings. 2022; 12(10): 1660. doi: 10.3390/buildings12101660
8. Mudiyanselage SE, Nguyen PHD, Rajabi MS, et al. Automated Workers' Ergonomic Risk Assessment in Manual Material Handling Using sEMG Wearable Sensors and Machine Learning. Electronics. 2021; 10(20): 2558. doi: 10.3390/electronics10202558
9. Gondia A, Siam A, El-Dakhakhni W, Nassar AH. Machine Learning Algorithms for Construction Projects Delay Risk Prediction. Journal of Construction Engineering and Management. 2020; 146(1): 04019085.
10. Huang J, Zeng X, Fu J, et al. Safety Risk Assessment Using a BP Neural Network of High Cutting Slope Construction in High-Speed Railway. Buildings. 2022; 12(5): 598. doi: 10.3390/buildings12050598
11. Ni G, Fang Y, Niu M, et al. Spatial differences, dynamic evolution and influencing factors of China's construction industry carbon emission efficiency. Journal of Cleaner Production. 2024; 448: 141593. doi: 10.1016/j.jclepro.2024.141593
12. Verma A, Prakash S, Kumar A. AI-based Building Management and Information System with Multi-agent Topology for an Energy-efficient Building: Towards Occupants Comfort. IETE Journal of Research. 2020; 69(2): 1033-1044. doi: 10.1080/03772063.2020.1847701
13. Kangari R, Riggs LS. Construction risk assessment by linguistics. IEEE Transactions on Engineering Management. 1989; 36(2): 126-131. doi: 10.1109/17.18829
14. Lin SS, Shen SL, Zhou A, et al. Risk assessment and management of excavation system based on fuzzy set theory and machine learning methods. Automation in Construction. 2021; 122: 103490. doi: 10.1016/j.autcon.2020.103490
15. Taroun A. Towards a better modelling and assessment of construction risk: Insights from a literature review. International Journal of Project Management. 2014; 32(1): 101-115. doi: 10.1016/j.ijproman.2013.03.004
16. KarimiAzari A, Mousavi N, Mousavi SF, et al. Risk assessment model selection in construction industry. Expert Systems with Applications. 2011; 38(8): 9105-9111. doi: 10.1016/j.eswa.2010.12.110
17. Subramanyan H, Sawant PH, Bhatt V. Construction Project Risk Assessment: Development of Model Based on Investigation of Opinion of Construction Project Experts from India. Journal of Construction Engineering and Management. 2012; 138(3): 409-421.
18. Hartmann T, Trappey A. Advanced Engineering Informatics - Philosophical and methodological foundations with examples from civil and construction engineering. Developments in the Built Environment. 2020; 4: 100020. doi: 10.1016/j.dibe.2020.100020
19. Sun H, Burton HV, Huang H. Machine learning applications for building structural design and performance assessment: State-of-the-art review. Journal of Building Engineering. 2021; 33: 101816. doi: 10.1016/j.jobe.2020.101816

20. Wang X, Mazumder RK, Salarieh B, et al. Machine Learning for Risk and Resilience Assessment in Structural Engineering: Progress and Future Trends. Journal of Structural Engineering. 2022; 148(8).
21. Zhang J, Zi L, Hou Y, et al. A C-BiLSTM Approach to Classify Construction Accident Reports. Applied Sciences. 2020; 10(17): 5754. doi: 10.3390/app10175754
22. Butler KT, Davies DW, Cartwright H, et al. Machine learning for molecular and materials science. Nature. 2018; 559(7715): 547-555. doi: 10.1038/s41586-018-0337-2
23. Zhong M, Tran K, Min Y, et al. Accelerated discovery of CO_2 electrocatalysts using active machine learning. Nature. 2020; 581(7807): 178-183. doi: 10.1038/s41586-020-2242-8
24. Warnat-Herresthal S, Schultze H, Shastry KL, et al. Swarm Learning for decentralized and confidential clinical machine learning. Nature. 2021; 594(7862): 265-270. doi: 10.1038/s41586-021-03583-3
25. Sun D, Wen H, Wang D, et al. A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. Geomorphology. 2020; 362: 107201. doi: 10.1016/j.geomorph.2020.107201
26. Akinosho TD, Oyedele LO, Bilal M, et al. Deep learning in the construction industry: A review of present status and future innovations. Journal of Building Engineering. 2020; 32: 101827. doi: 10.1016/j.jobe.2020.101827
27. Alpaydın E. Machine Learning. Mit Press; 2021.
28. Hegde J, Rokseth B. Applications of machine learning methods for engineering risk assessment - A review. Safety Science. 2020; 122: 104492. doi: 10.1016/j.ssci.2019.09.015
29. Maps OK. Open Knowledge Maps—A visual interface to the world's scientific knowledge. Open Knowledge Maps; 2023.
30. Litmaps. app.litmaps.com. Available online: https://app.litmaps.com/seed (accessed on 15 April 2024).
31. Tessema AT, Alene GA, Wolelaw NM. Assessment of risk factors on construction projects in gondar city, Ethiopia. Heliyon. 2022; 8(11): e11726. doi: 10.1016/j.heliyon.2022.e11726
32. Zhu Z, Sun J, Li X. An construction method of scorecard using machine learning and logical regression. Procedia Computer Science. 2022; 214: 1541-1545. doi: 10.1016/j.procs.2022.11.341
33. Gariazzo C, Taiano L, Bonafede M, et al. Association between extreme temperature exposure and occupational injuries among construction workers in Italy: An analysis of risk factors. Environment International. 2023; 171: 107677. doi: 10.1016/j.envint.2022.107677
34. Hemasinghe H, Rangali RSS, Deshapriya NL, et al. Landslide susceptibility mapping using logistic regression model (a case study in Badulla District, Sri Lanka). Procedia Engineering. 2018; 212: 1046-1053. doi: 10.1016/j.proeng.2018.01.135
35. Li N, Jimenez R. A logistic regression classifier for long-term probabilistic prediction of rock burst hazard. Natural Hazards. 2017; 90(1): 197-215. doi: 10.1007/s11069-017-3044-7
36. Xie P, Zhang R, Zheng J, et al. Probabilistic analysis of subway station excavation based on BIM-RF integrated technology. Automation in Construction. 2022; 135: 104114. doi: 10.1016/j.autcon.2021.104114
37. Hu W, Zhang S, Fu Y, et al. Objective diagnosis of machine learning method applicability to land comprehensive carrying capacity evaluation: A case study based on integrated RF and DPSIR models. Ecological Indicators. 2023; 151: 110338. doi: 10.1016/j.ecolind.2023.110338
38. Wang R, Asghari V, Hsu SC, et al. Detecting corporate misconduct through random forest in China's construction industry. Journal of Cleaner Production. 2020; 268: 122266. doi: 10.1016/j.jclepro.2020.122266
39. Wu X, Wang L, Chen B, et al. Multi-objective optimization of shield construction parameters based on random forests and NSGA-II. Advanced Engineering Informatics. 2022; 54: 101751. doi: 10.1016/j.aei.2022.101751
40. Wen H, Wu J, Zhang C, et al. Hybrid optimized RF model of seismic resilience of buildings in mountainous region based on hyperparameter tuning and SMOTE. Journal of Building Engineering. 2023; 71: 106488. doi: 10.1016/j.jobe.2023.106488
41. Hu R, Chen K, Chen W, et al. Estimation of construction waste generation based on an improved on-site measurement and SVM-based prediction model: A case of commercial buildings in China. Waste Management. 2021; 126: 791-799. doi: 10.1016/j.wasman.2021.04.012
42. Chen JH, Lin JZ. Developing an SVM based risk hedging prediction model for construction material suppliers. Automation in Construction. 2010; 19(6): 702-708. doi: 10.1016/j.autcon.2010.02.014
43. Tserng H P, Lin G F, Tsai L K, et al. An enforced support vector machine model for construction contractor default prediction. Automation in Construction, 2011; 20(8): 1242-1249. doi: 10.1016/j.autcon.2011.05.007

44. Fan M, Sharma A. Design and implementation of construction cost prediction model based on SVM and LSSVM in industries 4.0. International Journal of Intelligent Computing and Cybernetics. 2021; 14(2): 145-157. doi: 10.1108/ijicc-10-2020-0142
45. Fu W, Zhang H, Huang F. Internet-based supply chain financing-oriented risk assessment using BP neural network and SVM. PLOS ONE. 2022; 17(1): e0262222. doi: 10.1371/journal.pone.0262222
46. Mostofi F, Toğan V, Ayözen YE, et al. Construction Safety Risk Model with Construction Accident Network: A Graph Convolutional Network Approach. Sustainability. 2022; 14(23): 15906. doi: 10.3390/su142315906
47. Khalili M A, Guerriero L, Pouralizadeh M, et al. Monitoring and prediction of landslide-related deformation based on the GCN-LSTM algorithm and SAR imagery. Natural Hazards; 2023; 119(1): 39-68. doi: 10.1007/s11069-023-06121-8
48. Mostofi F, Toğan V. Predicting Construction Accident Outcomes Using Graph Convolutional and Dual-Edge Safety Networks. Arabian Journal for Science and Engineering. 2023; 49: 13315-13312. doi: 10.1007/s13369-023-08609-8
49. Fu X, Pan Y, Zhang L. A causal-temporal graphic convolutional network (CT-GCN) approach for TBM load prediction in tunnel excavation. Expert Systems with Applications. 2024; 238: 121977. doi: 10.1016/j.eswa.2023.121977
50. Li P, Wu F, Xue S, et al. Study on the Interaction Behaviors Identification of Construction Workers Based on ST-GCN and YOLO. Sensors. 2023; 23(14): 6318. doi: 10.3390/s23146318
51. Zhang Y, Du Z, Hu L. A construction method of urban road risky vehicles based on dynamic knowledge graph. Electronic Research Archive. 2023; 31(7): 3776-3790. doi: 10.3934/era.2023192
52. Chen JH, Hsu SC, Luo YH, Skibniewski MJ. Knowledge Management for Risk Hedging by Construction Material Suppliers." Journal of Management in Engineering. 2012; 28(3): 273-280.
53. Pandey P, Bandhu KC. A credit risk assessment on borrowers classification using optimized decision tree and KNN with bayesian optimization. International Journal of Information Technology. 2022; 14(7): 3679-3689. doi: 10.1007/s41870-022-00974-1
54. Jaber F K, Al-Zwainy F M S, Hachem S W. Optimizing of predictive performance for construction projects utilizing support vector machine technique. Cogent Engineering, 2019; 6(1): 1685860. doi: 10.1080/23311916.2019.1685860
55. Chenzhong R, Wenliang K, Taihua Z, et al. Intelligent Generation and Analysis of the Municipal Road Construction Scheme Based on the KNN Algorithm. Mathematical Problems in Engineering. 2022; 10: 1-15. doi: 10.1155/2022/8752870
56. Li L, Wu Y, Huang Y, et al. Optimized Apriori algorithm for deformation response analysis of landslide hazards. Computers & Geosciences. 2023; 170: 105261. doi: 10.1016/j.cageo.2022.105261
57. Chen B, Wei N, Qu T, et al. Research on weighting method of geological hazard susceptibility evaluation index based on Apriori Algorithm. Frontiers in Earth Science. 2023; 11. doi: 10.3389/feart.2023.1127889
58. Chen Q, Tian Z, Lei T, et al. An association rule mining model for evaluating the potential correlation of construction cross operation risk. Engineering, Construction and Architectural Management. 2022; 30(10): 5109-5132. doi: 10.1108/ecam-09-2021-0792
59. Shao B, Hu Z, Liu D. Using Improved Principal Component Analysis to Explore Construction Accident Situations from the Multi-Dimensional Perspective: A Chinese Study. International Journal of Environmental Research and Public Health. 2019; 16(18): 3476-3476.
60. Xiang P, Jia F, Li X. Critical Behavioral Risk Factors among Principal Participants in the Chinese Construction Industry. Sustainability. 2018; 10(9): 3158. doi: 10.3390/su10093158
61. Siddiqui F, Sargent P, Montague G. The use of PCA and signal processing techniques for processing time-based construction settlement data of road embankments. Advanced Engineering Informatics. 2020; 46: 101181. doi: 10.1016/j.aei.2020.101181
62. Yan H, He Z, Gao C, et al. Investment estimation of prefabricated concrete buildings based on XGBoost machine learning algorithm. Advanced Engineering Informatics. 2022; 54: 101789. doi: 10.1016/j.aei.2022.101789
63. Cherif IL, Kortebi A. On using Extreme Gradient Boosting (XGBoost) Machine Learning Algorithm for Home Network Traffic Classification. IEEE Xplore; 2019.
64. Coffie GH, Cudjoe SKF. Using extreme gradient boosting (XGBoost) machine learning to predict construction cost overruns. International Journal of Construction Management. 2023.
65. Liang W, Luo S, Zhao G, et al. Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms. Mathematics, 2020; 8(5): 765. doi: 10.3390/math8050765
66. Liu P, Li Y. An improved failure mode and effect analysis method for multi-criteria group decision-making in green logistics risk assessment. Reliability Engineering & System Safety. 2021; 215: 107826. doi: 10.1016/j.ress.2021.107826

67. Wang G, Liu M, Cao D, et al. Identifying high-frequency-low-severity construction safety risks: an empirical study based on official supervision reports in Shanghai. Engineering, Construction and Architectural Management. 2021; 29(2): 940-960. doi: 10.1108/ecam-07-2020-0581
68. Ayhan BU, Tokdemir OB. Accident Analysis for Construction Safety Using Latent Class Clustering and Artificial Neural Networks. Journal of Construction Engineering and Management. 2020; 146(3): 04019114.
69. Kim S, Choi CY, Shahandashti M, Rok Ryu K. Improving Accuracy in Predicting City-Level Construction Cost Indices by Combining Linear ARIMA and Nonlinear ANNs. Journal of Management in Engineering. 2022; 38(2).
70. Moon S, Chi S, Kim DY. Predicting Construction Cost Index Using the Autoregressive Fractionally Integrated Moving Average Model. Journal of Management in Engineering. 2018; 34(2): 04017063.
71. Ghashghaie M, Nozari H. Effect of Dam Construction on Lake Urmia: Time Series Analysis of Water Level via ARIMA. Journal of Agricultural Science and Technology. 2018; 20(7): 1541-1553.
72. Kaloop MR, Eldiasty M, Hu JW. Safety and reliability evaluations of bridge behaviors under ambient truck loads through structural health monitoring and identification model approaches. Measurement. 2022; 187: 110234. doi: 10.1016/j.measurement.2021.110234
73. Hajifar S, Sun H, Megahed FM, et al. A forecasting framework for predicting perceived fatigue: Using time series methods to forecast ratings of perceived exertion with features from wearable sensors. Applied Ergonomics. 2021; 90: 103262. doi: 10.1016/j.apergo.2020.103262
74. Lam TYM, Siwingwa N. Risk management and contingency sum of construction projects. Journal of Financial Management of Property and Construction. 2017; 22(3): 237-251. doi: 10.1108/jfmpc-10-2016-0047
75. Montgomery DC, Peck EA, Vining GG. Introduction to Linear Regression Analysis. John Wiley & Sons; 2021.
76. Huang CH, Hsieh SH. Predicting BIM labor cost with random forest and simple linear regression. Automation in Construction. 2020; 118: 103280. doi: 10.1016/j.autcon.2020.103280
77. Esmaeili B, Hallowell MR, Rajagopalan B. Attribute-Based Safety Risk Assessment. II: Predicting Safety Outcomes Using Generalized Linear Models. Journal of Construction Engineering and Management. 2015; 141(8): 04015022.
78. Lyu J. Construction of Enterprise Financial Early Warning Model Based on Logistic Regression and BP Neural Network. Computational Intelligence and Neuroscience. 2022; 2022: 1-7. doi: 10.1155/2022/2614226
79. Levy A, Baha R. Credit risk assessment: a comparison of the performances of the linear discriminant analysis and the logistic regression. International Journal of Entrepreneurship and Small Business. 2021; 42(1/2): 169. doi: 10.1504/ijesb.2021.112265
80. Akboğa Kale Ö, Baradan S. Identifying Factors that Contribute to Severity of Construction Injuries using Logistic Regression Model. Teknik Dergi. 2020; 31(2): 9919-9940. doi: 10.18400/tekderg.470633
81. Bhattacharjee P, Dey V, Mandal UK. Risk assessment by failure mode and effects analysis (FMEA) using an interval number based logistic regression model. Safety Science. 2020; 132: 104967. doi: 10.1016/j.ssci.2020.104967
82. Wong CH. Contractor Performance Prediction Model for the United Kingdom Construction Contractor: Study of Logistic Regression Approach. Journal of Construction Engineering and Management. 2004; 130(5): 691-698.
83. Zhang X, Huang S, Yang S, et al. Safety Assessment in Road Construction Work System Based on Group AHP-PCA. Mathematical Problems in Engineering. 2020; 2020: 1-12. doi: 10.1155/2020/6210569
84. Bai L, Song C, Zhou X, et al. Assessing project portfolio risk via an enhanced GA-BPNN combined with PCA. Engineering Applications of Artificial Intelligence. 2023; 126: 106779. doi: 10.1016/j.engappai.2023.106779
85. Shi H, Li W, Deng Y. Applying Principal Component Analysis and Unascertained Method for the Analysis of Construction Accident Risk. Journal of Computers. 2010; 5(8): 1273-1280. doi: 10.4304/jcp.5.8.1273-1280
86. Wang G, Ma J. A hybrid ensemble approach for enterprise credit risk assessment based on Support Vector Machine. Expert Systems with Applications. 2012; 39(5): 5325-5331. 10.1016/j.eswa.2011.11.003
87. Khemakhem S, Ben Said F, Boujelbene Y. Credit risk assessment for unbalanced datasets based on data mining, artificial neural network and support vector machines. Journal of Modelling in Management. 2018; 13(4): 932-951. doi: 10.1108/jm2-01-2017-0002
88. Mangeli M, Shahraki A, Saljooghi FH. Improvement of risk assessment in the FMEA using nonlinear model, revised fuzzy TOPSIS, and support vector machine. International Journal of Industrial Ergonomics. 2019; 69: 209-216. doi: 10.1016/j.ergon.2018.11.004
89. Steinwart I. Support Vector Machines. Springer; 2014.

90. Noble W S. What is a support vector machine?. Nature Biotechnology. 2006; 24(12): 1565-1567. doi: 10.1038/nbt1206-1565
91. Yang R, Feng J, Sun Y. Construction and classification prediction of risk assessment lindicators for water environment treatment PPP projects. Available online: https://www.researchsquare.com/article/rs-2845690/v1 (accessed on 15 April 2024).
92. Zhang L, Hu H, Zhang D. A credit risk assessment model based on SVM for small and medium enterprises in supply chain finance. Financial Innovation. 2015; 1(1). doi: 10.1186/s40854-015-0014-5
93. Gong P, Guo H, Huang Y, et al. Safety risk evaluations of deep foundation construction schemes based on imbalanced data sets. Journal of civil engineering and management. 2020; 26(4): 380-395.
94. Liu P, Xie M, Bian J, et al. A Hybrid PSO-SVM Model Based on Safety Risk Prediction for the Design Process in Metro Station Construction. International Journal of Environmental Research and Public Health. 2020; 17(5): 1714. doi: 10.3390/ijerph17051714
95. Wei Y, Zhang J, Wang J. Research on Building Fire Risk Fast Assessment Method Based on Fuzzy comprehensive evaluation and SVM. Procedia Engineering. 2018; 211: 1141-1150. doi: 10.1016/j.proeng.2017.12.121
96. Chang YC, Chang KH, Wu GJ. Application of Extreme gradient boosting trees in the construction of credit risk assessment models for financial institutions. Applied Soft Computing. 2018; 73: 914-920. doi: 10.1016/j.asoc.2018.09.029
97. Li Z. GBDT-SVM Credit Risk Assessment Model and Empirical Analysis of Peer-to-Peer Borrowers under Consideration of Audit Information. Open Journal of Business and Management. 2018; 6(2): 362-372. doi: 10.4236/ojbm.2018.62026
98. Zhang H, Shi Y, Yang X, et al. A firefly algorithm modified support vector machine for the credit risk assessment of supply chain finance. Research in International Business and Finance. 2021; 58: 101482. doi: 10.1016/j.ribaf.2021.101482
99. Yin Q, Zhou J, Zhou Y, et al. Construction safety risk assessment method of construction engineering based on improved SVM. International Journal of Sustainable Development, 2023; 26(3-4): 329-343. doi: 10.1504/IJSD.2023.10058663155
100. Chen J, Liu L, Pei J, et al. An ensemble risk assessment model for urban rainstorm disasters based on random forest and deep belief nets: a case study of Nanjing, China. Natural Hazards. 2021; 107(3): 2671-2692. doi: 10.1007/s11069-021-04630-y
101. Tang L, Cai F, Ouyang Y. Applying a nonparametric random forest algorithm to assess the credit risk of the energy industry in China. Technological Forecasting and Social Change. 2019; 144: 563-572. doi: 10.1016/j.techfore.2018.03.007
102. Liu W, Zhang Y, Liang Y, et al. Landslide Risk Assessment Using a Combined Approach Based on InSAR and Random Forest. Remote Sensing. 2022; 14(9): 2131. doi: 10.3390/rs14092131
103. Zermane A, Mohd Tohir MZ, Zermane H, et al. Predicting fatal fall from heights accidents using random forest classification machine learning model. Safety Science. 2023; 159: 106023. doi: 10.1016/j.ssci.2022.106023
104. Kang K, Ryu H. Predicting types of occupational accidents at construction sites in Korea using random forest model. Safety Science. 2019; 120: 226-236. doi: 10.1016/j.ssci.2019.06.034
105. Zhou Y, Li S, Zhou C, Luo H. Intelligent Approach Based on Random Forest for Safety Risk Prediction of Deep Foundation Pit in Subway Stations. Journal of Computing in Civil Engineering. 2019; 33(1).
106. Zhang H, Shi Y, Tong J. Online supply chain financial risk assessment based on improved random forest. Journal of Data, Information and Management. 2021; 3(1): 41-48. doi: 10.1007/s42488-021-00042-6
107. Zhu Z, Zhang Y. Flood disaster risk assessment based on random forest algorithm. Neural Computing and Applications. 2021; 34(5): 3443-3455. doi: 10.1007/s00521-021-05757-6
108. Wang Y, Wen H, Sun D, et al. Quantitative Assessment of Landslide Risk Based on Susceptibility Mapping Using Random Forest and GeoDetector. Remote Sensing. 2021; 13(13): 2625. doi: 10.3390/rs13132625
109. Aprillia H, Yang HT, Huang CM. Statistical Load Forecasting Using Optimal Quantile Regression Random Forest and Risk Assessment Index. IEEE Transactions on Smart Grid. 2021; 12(2): 1467-1480. doi: 10.1109/tsg.2020.3034194
110. Armaghani DJ, Mahdiyar A, Hasanipanah M, et al. Risk Assessment and Prediction of Flyrock Distance by Combined Multiple Regression Analysis and Monte Carlo Simulation of Quarry Blasting. Rock Mechanics and Rock Engineering. 2016; 49(9): 3631-3641. doi: 10.1007/s00603-016-1015-z
111. Liu Y, Chen H, Zhang L, et al. Risk prediction and diagnosis of water seepage in operational shield tunnels based on random forest. Journal of civil engineering and management. 2021; 27(7): 539-552.
112. Ghosh S, Das A. Wetland conversion risk assessment of East Kolkata Wetland: A Ramsar site using random forest and support vector machine model. Journal of Cleaner Production. 2020; 275: 123475. doi: 10.1016/j.jclepro.2020.123475
113. Langroodi AK, Vahdatikhaki F, Doree A. Activity recognition of construction equipment using fractional random forest. Automation in Construction. 2021; 122: 103465. doi: 10.1016/j.autcon.2020.103465

114. Junjia Y, Alias AH, Haron NA, et al. A Bibliometrics-Based Systematic Review of Safety Risk Assessment for IBS Hoisting Construction. Buildings. 2023; 13(7): 1853. doi: 10.3390/buildings13071853
115. Han J, Kim J, Park S, et al. Seismic Vulnerability Assessment and Mapping of Gyeongju, South Korea Using Frequency Ratio, Decision Tree, and Random Forest. Sustainability. 2020; 12(18): 7787. doi: 10.3390/su12187787
116. Lee YYR, Samad H, Miang Goh Y. Perceived Importance of Authentic Learning Factors in Designing Construction Safety Simulation Game-Based Assignment: Random Forest Approach. Journal of Construction Engineering and Management. 2020; 146(3): 04020002.
117. Shoar S, Chileshe N, Edwards JD. Machine learning-aided engineering services' cost overruns prediction in high-rise residential building projects: Application of random forest regression. Journal of Building Engineering. 2022; 50: 104102. doi: 10.1016/j.jobe.2022.104102
118. Chen J, Li Q, Wang H, et al. A Machine Learning Ensemble Approach Based on Random Forest and Radial Basis Function Neural Network for Risk Evaluation of Regional Flood Disaster: A Case Study of the Yangtze River Delta, China. International Journal of Environmental Research and Public Health. 2019; 17(1): 49. doi: 10.3390/ijerph17010049
119. Karabadji NEI, Amara Korba A, Assi A, et al. Accuracy and diversity-aware multi-objective approach for random forest construction. Expert Systems with Applications. 2023; 225: 120138. doi: 10.1016/j.eswa.2023.120138
120. Sabry F. K Nearest Neighbor Algorithm. One Billion Knowledgeable; 2023.
121. Lee KP, Lee HS, Park M, Kim DY, Jung M. Management-Reserve Estimation for International Construction Projects Based on Risk-Informed k-NN. Journal of Management in Engineering. 2017; 33(4): 04017002.
122. Kamran M, Ullah B, Ahmad M, et al. Application of KNN-based Isometric Mapping and Fuzzy C-Means Algorithm to Predict Short-term Rockburst Risk in Deep Underground Projects. Frontiers in Public Health. 2022; 10. doi: 10.21203/rs.3.rs-2128698/v1
123. Liu X, Xu F, Zhang Z, et al. Fall-portent detection for construction sites based on computer vision and machine learning. Engineering, Construction and Architectural Management. 2023. doi: 10.1108/ecam-05-2023-0458
124. Sanni-Anibire MO, Zin RM, Olatunji SO. Machine learning model for delay risk assessment in tall building projects. International Journal of Construction Management, 2022; 22(11): 2134-2143. doi: 10.1080/15623599.2020.1768326
125. Burns JJR, Shealy BT, Greer MS, et al. Addressing noise in co-expression network construction. Briefings in Bioinformatics. 2021; 23(1). doi: 10.1093/bib/bbab495
126. Zhong G, Lu G, Liu M, Cui M. A novel risk assessment system for port state control inspection. In: Proceedings of the 2008 IEEE International Conference on Intelligence and Security Informatics; 17-20 June 2008; Taipei. pp. 242-244. doi: 10.1109/ISI.2008.4565068
127. Arabiat A, Al-Bdour H, Bisharah M. Predicting the construction projects time and cost overruns using K-nearest neighbor and artificial neural network: a case study from Jordan. Asian Journal of Civil Engineering. 2023; 24(7): 2405-2414. doi: 10.1007/s42107-023-00649-7
128. Zhang Y, Ding L, Peter ED. Planning of Deep Foundation Construction Technical Specifications Using Improved Case-Based Reasoning with Weighted k-Nearest Neighbors. American Society of Civil Engineers. 2017; 31(5).
129. Brownlee J. XGBoost with Python. Machine Learning Mastery; 2016.
130. Qin R. The Construction of Corporate Financial Management Risk Model Based on XGBoost Algorithm. Chen M, ed. Journal of Mathematics. 2022; 2022: 1-8. doi: 10.1155/2022/2043369
131. Liu W, Chen Z, Hu Y. XGBoost algorithm-based prediction of safety assessment for pipelines. International Journal of Pressure Vessels and Piping. 2022; 197: 104655. doi: 10.1016/j.ijpvp.2022.104655
132. Wang Y, Ni XS. A XGBoost risk model via feature selection and Bayesian hyper-parameter optimization. Available online: https://arxiv.org/abs/1901.08433 (accessed on 15 April 2024).
133. Shi L, Qian C, Guo F. Real-time driving risk assessment using deep learning with XGBoost. Accident Analysis & Prevention. 2022; 178: 106836. doi: 10.1016/j.aap.2022.106836
134. Luo H, Yang Q, Wang W, et al. XGBoost-based assessment method for fire risk levels of transmission lines. Journal of Electric Power Science and Technology. 2024; 38(6): 132-141. doi: 10.19781/j.issn.1673-9140.2023.06.014
135. Shehadeh A, Alshboul O, Al Mamlook RE, et al. Machine learning models for predicting the residual value of heavy construction equipment: An evaluation of modified decision tree, LightGBM, and XGBoost regression. Automation in Construction. 2021; 129: 103827. doi: 10.1016/j.autcon.2021.103827

136. Hamerly G, Elkan C. Learning the k in k-means. Available online: https://proceedings.neurips.cc/paper/2003/hash/234833147b97bb6aed53a8f4f1c7a7d8-Abstract.html (accessed on 15 April 2024).
137. Er Kara M, Oktay Fırat S. Supplier Risk Assessment Based on Best-Worst Method and K-Means Clustering: A Case Study. Sustainability. 2018; 10(4): 1066. doi: 10.3390/su10041066
138. Chattapadhyay DB, Putta J, Rao RMP. Risk Identification, Assessments, and Prediction for Mega Construction Projects: A Risk Prediction Paradigm Based on Cross Analytical-Machine Learning Model. Buildings. 2021; 11(4): 172.
139. Tao Y, Yong X, Yang J, et al. Risk Early-Warning Framework for Government-Invested Construction Project Based on Fuzzy Theory, Improved BPNN, and K-Means. Mathematical Problems in Engineering. 2022; 2022: 1-19. doi: 10.1155/2022/5958472
140. Shumway RH, Stoffer DS, Shumway RH, et al. ARIMA models. Time series analysis and its applications: With R examples, 2017; 75-163. doi: 10.1007/978-3-319-52452-8_3
141. Cao H, Goh YM. Analyzing construction safety through time series methods. Frontiers of Engineering Management. 2019; 6(2): 262-274. doi: 10.1007/s42524-019-0015-6
142. Li M, Baek M, Ashuri B. Forecasting Ratio of Low Bid to Owner's Estimate for Highway Construction. Journal of Construction Engineering and Management. 2021; 147(1).
143. Yi F, Zeng H, Liu T, Wu Y. Research on cement price fluctuation prediction based on EEMD-ARIMA. In: Li J, Lu W, Peng Y, et al. (editors). Proceedings of the 27th International Symposium on Advancement of Construction Management and Real Estate. Springer; 2023.
144. Liu Z, Zhou J. Introduction to Graph Neural Networks. Springer International Publishing; 2020. doi: 10.1007/978-3-031-01587-8
145. Fu X, Wu M, Ponnarasu S, et al. A hybrid deep learning approach for dynamic attitude and position prediction in tunnel construction considering spatio-temporal patterns. Expert Systems with Applications. 2023; 212: 118721. doi: 10.1016/j.eswa.2022.118721
146. Pan X, Zhong B, Wang Y, et al. Identification of accident-injury type and bodypart factors from construction accident reports: A graph-based deep learning framework. Advanced Engineering Informatics. 2022; 54: 101752. doi: 10.1016/j.aei.2022.101752
147. Junjia Y, Alias A H, Haron N A, et al. Identification and analysis of hoisting safety risk factors for IBS construction based on the AcciMap and cases study. Heliyon. 2024; 10(1). doi: 10.1016/j.heliyon.2023.e23587
148. Mostofi F, Toğan V. Construction safety predictions with multi-head attention graph and sparse accident networks. Automation in Construction. 2023; 156: 105102. doi: 10.1016/j.autcon.2023.105102
149. Xue G, Liu S, Ren L, et al. Risk assessment of utility tunnels through risk interaction-based deep learning. Reliability Engineering & System Safety. 2024; 241: 109626. doi: 10.1016/j.ress.2023.109626
150. Zhu W, Shi D, Cheng R, et al. Human risky behaviour recognition during ladder climbing based on multi-modal feature fusion and adaptive graph convolutional network. Signal, Image and Video Processing. 2024; 18: 2473-2483. doi: 10.1007/s11760-023-02923-2
151. Xie X, Fu G, Xue Y, et al. Risk prediction and factors risk analysis based on IFOA-GRNN and apriori algorithms: Application of artificial intelligence in accident prevention. Process Safety and Environmental Protection. 2019; 122: 169-184. doi: 10.1016/j.psep.2018.11.019
152. Deng Y, Zhang Y, Yuan Z, et al. Analyzing Subway Operation Accidents Causations: Apriori Algorithm and Network Approaches. International Journal of Environmental Research and Public Health. 2023; 20(4): 3386. doi: 10.3390/ijerph20043386
153. Junjia Y, Alias A H, Haron N A, et al. A Bibliometric Review on Safety Risk Assessment of Construction Based on CiteSpace Software and WoS Database. Sustainability, 2023; 15(15): 11803. doi: 10.3390/su151511803
154. Sarkar S, Ejaz N, Maiti J, et al. An integrated approach using growing self-organizing map-based genetic K-means clustering and tolerance rough set in occupational risk analysis. Neural Computing and Applications. 2022; 34(12): 9661-9687. doi: 10.1007/s00521-022-06956-5
155. Junjia Y, Alias A H, Haron N A, et al. Trend Analysis of Marine Construction Disaster Prevention Based on Text Mining: Evidence from China. Sustainable Marine Structures, 2024; 6(1): 20-32. doi: 10.36956/sms.v6i1.1026

156. Verma A, Prakash S, Srivastava V, et al. Sensing, Controlling, and IoT Infrastructure in Smart Building: A Review. IEEE Sensors Journal. 2019; 19(20): 9036-9046. doi: 10.1109/jsen.2019.2922409
157. Feng S, He X, Xu H, Armaghani DJ, Sheng D. Applications of Machine Learning in Mechanised Tunnel Construction: A Systematic Review. Eng. 2023; 4(2): 1516-1535.
158. Zhu Z, Jeelani I, Gheisari M. Physical risk assessment of drone integration in construction using 4D simulation. Automation in Construction. 2023; 156: 105099. doi: 10.1016/j.autcon.2023.105099
159. Zhong J, Mao Y, Yuan X. Lifetime seismic risk assessment of bridges with construction and aging considerations. Structures. 2023; 47: 2259-2272. doi: 10.1016/j.istruc.2022.12.035
160. Nguyen HD, Do QNH, Macchion L. Influence of practitioners' characteristics on risk assessment in Green Building projects in emerging economies: a case of Vietnam. Engineering, Construction and Architectural Management. 2021; 30(2): 833-852. doi: 10.1108/ecam-05-2021-0436
161. Sohrabi H, Noorzai E. Risk assessment in Iranian oil and gas construction industry: a process approach. International Journal of Quality & Reliability Management. 2021; 40(1): 124-147. doi: 10.1108/ijqrm-03-2021-0069
162. Hatamleh MT, Moynihan GP, Batson RG, et al. Risk assessment and ranking in the developing countries' construction industry: the case of Jordan. Engineering, Construction and Architectural Management. 2021; 30(4): 1344-1364. doi: 10.1108/ecam-06-2021-0489
163. Al-Mhdawi MKS, Brito M, Onggo BS, et al. COVID-19 emerging risk assessment for the construction industry of developing countries: evidence from Iraq. International Journal of Construction Management. 2024; 24(7): 693-706. doi: 10.1080/15623599.2023.2169301
164. Gashaw T, Jilcha K. Design risk modeling and analysis for railway construction projects. International Journal of Construction Management. 2022; 23(14): 2488-2498. doi: 10.1080/15623599.2022.2070344
165. Do ST, Nguyen VT, Likhitruangsilp V. RSIAM risk profile for managing risk factors of international construction joint ventures. International Journal of Construction Management. 2021; 23(7): 1148-1162. doi: 10.1080/15623599.2021.1957753
166. He S, Xu H, Zhang J, et al. Risk assessment of oil and gas pipelines hot work based on AHP-FCE. Petroleum. 2023; 9(1): 94-100. doi: 10.1016/j.petlm.2022.03.006
167. Mohandes SR, Durdyev S, Sadeghi H, et al. Towards enhancement in reliability and safety of construction projects: developing a hybrid multi-dimensional fuzzy-based approach. Engineering, Construction and Architectural Management. 2022; 30(6): 2255-2279. doi: 10.1108/ecam-09-2021-0817
168. Badi I, Bouraima MB, Jibril ML. Risk Assessment in Construction Projects Using the Grey Theory. Journal of Engineering Management and Systems Engineering. 2022; 1(2): 58-66. doi: 10.56578/jemse010203
169. Zhang L, Li H. Construction Risk Assessment of Deep Foundation Pit Projects Based on the Projection Pursuit Method and Improved Set Pair Analysis. Applied Sciences. 2022; 12(4): 1922. doi: 10.3390/app12041922
170. Ju W, Wu J, Kang Q, et al. A method based on the theories of game and extension cloud for risk assessment of construction safety: A case study considering disaster-inducing factors in the construction process. Journal of Building Engineering. 2022; 62: 105317. doi: 10.1016/j.jobe.2022.105317
171. Sadeghi M, Mahmoudi A, Deng X. Blockchain technology in construction organizations: risk assessment using trapezoidal fuzzy ordinal priority approach. Engineering, Construction and Architectural Management. 2022; 30(7): 2767-2793. doi: 10.1108/ecam-01-2022-0014