Results in Engineering 28 (2025) 107882

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

Review article

State-of-the-art artificial intelligence techniques in structural engineering: A review of applications and prospects

Md. Tarif Aziz [a], Dave Montellano Osabel [b], Youngju Kim [c], Sanghoon Kim [d,*], Jaehoon Bae [b,*], Konstantinos Daniel Tsavdaridis [e,*]

[a] Department of Sustainable Architecture ICT Convergence, Chonnam National University, Yeosu City, Jeollanam-do 59626, Republic of Korea
[b] Department of Architectural Design, Chonnam National University, Yeosu City, Jeollanam-do 59626, Republic of Korea
[c] Korea Institute of Structural Engineering & Consulting (KISEC), Busan, Republic of Korea
[d] Department of Mechanical Design Engineering, Chonnam National University, Yeosu City, Jeollanam-do 59626, Republic of Korea
[e] Department of Engineering, School of Science & Technology, City St George's, University of London, Northampton Square, London EC1V 0HB, UK

ARTICLE INFO

Keywords:
Artificial intelligence
Structural engineering
Reinforced concrete structures
Steel structures
Structural safety

ABSTRACT

Artificial intelligence (AI) has emerged as a key driver of modern technological development, with widespread applications across various domains, including civil engineering. Structural engineering, a subdiscipline of civil engineering, requires the evaluation of the suitability of different structural components before final construction phase and during recycling processes. Traditionally, this evaluation relies on laboratory experiments and highly complex numerical simulations, which are often impractical due to space and time constraints, equipment complexity, and high costs. To address these challenges, researchers worldwide have developed AI-based solutions for applications such as structural damage detection and the prediction of failure loads and patterns. These solutions offer predictive accuracy comparable to that of experimental and numerical analyses. This review presents a detailed analysis of around 100 AI-integrated studies in structural engineering conducted between 2020 and 2024, with a focus on concrete, steel, and composite structures, particularly building frames. The study summarizes the performance benchmarking of commonly used AI algorithms, such as neural networks, genetic algorithms, tree-based algorithms, and boosting methods, reporting accuracy scores above 0.80 (out of 1.00), and highlights average accuracy values of 0.90 for optimized and hybrid AI approaches. Additionally, the review explores emerging AI applications, including retrofitting technologies, buckling-restrained braces, dampers, column-beam connections, and life-cycle assessment. Critical analysis identifies key limitations of recent AI-based research, especially those implemented regionally, and proposes novel solutions to overcome existing challenges.

1. Introduction

Artificial Intelligence (AI) was first conceptualized by a group of scientists at a conference at Dartmouth College in 1956 [1] to develop intelligent systems capable of reasoning and exhibiting human-like intelligence [2]. Significant interest and growth in AI emerged through U. S. Defense Advanced Research Projects Agency (DARPA) funding from 1962 [3] (Fig. 1). However, between 1970 and 1980, AI research stagnated due to limited high-performance computing resources, and DARPA funding was discontinued following the critical "Lighthill Report", which reported AI to have failed to achieve its purpose [4]. From the 1980s, AI experienced a resurgence driven by the evolution of some early methods (e.g., expert systems and cybernetics) and practical industrial applications [2]. Nevertheless, between 1987 and 1993, AI faced another decline caused by unrealistic expectations and limited computational power [5]. Later, subsequent advances in Information Technology (IT) and the industrial revolution enabled significant progress in AI [6]. Today, the availability of faster, cost-effective, and more powerful processing systems has facilitated the widespread adoption of AI [7].

AI encompasses a broad range of methods, including—but not limited to—machine learning (ML), neural networks (NN), deep

* Corresponding authors.
E-mail addresses: tarifctg98@gmail.com (Md.T. Aziz), osabel.dm@chonnam.ac.kr (D.M. Osabel), shkim83@chonnam.ac.kr (S. Kim), skycity-bjh@jnu.ac.kr (J. Bae), konstantinsos.tsavdaridis@city.ac.uk (K.D. Tsavdaridis).

https://doi.org/10.1016/j.rineng.2025.107882
Received 26 April 2025; Received in revised form 12 October 2025; Accepted 21 October 2025
Available online 22 October 2025

learning (DL), data mining, knowledge discovery and advanced analytics, rule-based modeling and decision making, fuzzy logic, knowledge representation, reasoning under uncertainty, expert systems, case-based reasoning, text mining and natural language processing, visual analytics, computer vision and pattern recognition, hybrid approaches, and optimization techniques [8]. These techniques are closely associated with disciplines such as computer science, information theory, cybernetics, linguistics, and neurophysiology [9]. By integrating the capabilities of these methods, AI can mimic human intelligence and apply human-inspired reasoning and algorithms to solve complex engineering problems [10]. Researchers worldwide are actively developing innovative AI approaches that are cost-effective, rapid, robust, and highly accurate [11]. Sarker [8] provided a comprehensive review of AI-based modeling in real-world applications.

In recent years, ML, DL, and NN have been extensively applied in civil engineering subfields, including structural, geotechnical, transportation, water supply, and hydraulic engineering. Recent studies have reviewed AI developments and applications in these areas. Pan and Zhang [12] conducted a scientometric analysis of AI-related publications from 1997 to 2020, highlighting AI's potential in automation and construction engineering and management. Manzoor et al. [13] reviewed 105 studies from 1995 to 2021, focusing on AI's role in sustainable development. Xu et al. [11] presented a systematic review on intelligent architectural design, structural health monitoring, and disaster prevention, emphasizing computer-vision-based advancements. Vadyala et al. [14] investigated the integration of ML methods with physics-based models to address data shift problems in supervised learning and proposed a physics-informed ML approach. Rezania et al. [15] discussed pioneering software, AI-related terminology, and parameters affecting progressive structural collapse. More recently, Harle [16] provided an overview of AI applications across some areas of civil engineering, including analysis and design, construction management, geotechnical engineering, and transportation planning, with a focus on ML and genetic algorithms.

AI has been applied in structural engineering for decades, particularly in the design of structural systems that account for critical factors such as load application characteristics, service life expectancy, durability against environmental effects, and fire-induced issues [17–20]. It serves as a powerful tool for generating efficient and accurate preliminary structural design predictions, reducing the reliance on cumbersome experimental setups, and enhancing safety measures during laboratory testing. Moreover, AI reduces the demand for high-precision instruments, which are often unavailable in many institutions and industries. For example, large-scale fire tests on structural frames cannot typically be conducted in laboratory settings, forcing researchers to rely on small-scale experiments and assumptions. AI can overcome such limitations by processing large-scale variable inputs and producing highly accurate predictions. However, as emphasized in this review, AI-based results must be validated against experimental and code-based outcomes, particularly in light of challenges such as data shift, domain shift, and extrapolation risk [18,21]. As illustrated in Fig. 1, AI was first applied to structural engineering in the early 1990s through expert systems, particularly for concrete, steel, and composite structures. In subsequent years, advanced methods such as ML, DL with NN were increasingly adopted in structural engineering. Following the first major AI revolution in 2012, DL has become increasingly prevalent in structural health monitoring (SHM) and structural damage detection.

In terms of AI's development and application within structural engineering, a review study examined four novel ML algorithms in structural system identification, SHM, structural vibration control, and structural design and prediction between 2017 and 2020 [22]. Another review [23] on fundamental ML techniques addressed a wide range of applications, including structural analysis and design, SHM, damage detection, fire resistance assessment, evaluation of mechanical properties, and concrete mix design. A more recent review [24] focused exclusively on ML applications in SHM. A comprehensive literature review [25] covering AI, ML, and DL discussed commonly used algorithms in structural engineering across >200 sources. Although the study conducted a scientometric analysis to map the best practices from several scholarly works, its primary focus was on supervised learning

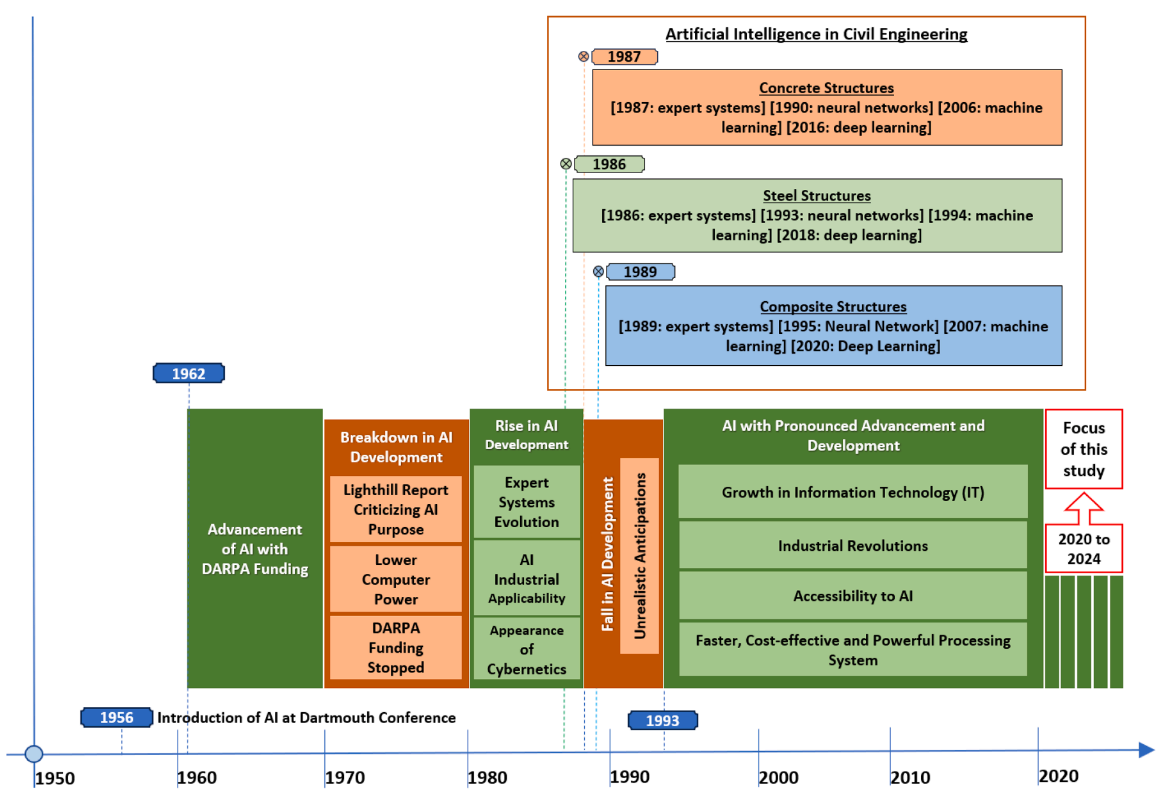

Fig. 1. Historical timeline and evolution of artificial intelligence (AI) in structural engineering.

methods up to the year 2021. However, these reviews only partially addressed recent trends of AI in structural engineering for material- and component-level analysis. They did not provide an in-depth discussion of structural components such as reinforced concrete, steel, and composite frames exposed to fire or seismic effects for the published studies during 2020–2024. Similarly, they did not explicitly consider innovative approaches such as AI integration with buckling-restrained braces (BRBs), viscoelastic dampers, retrofitting technologies (e.g., fiber-reinforced jacketing), beam–column joints, or life-cycle assessment (LCA) of buildings.

As presented in Fig. 2a, based on Scopus search results, published articles on prominent AI algorithms (NN, GA, tree-based algorithms, and boosting algorithms) in structural engineering have increased significantly from 2015 to 2024. The number of published articles in the last five years (2020–2024) has increased by 2.72 times in comparison to 2015–2019. Fig. 2b illustrates a similar upward trend for each prominent group of algorithms. Research on NN has grown most dramatically, increasing by approximately 4.8 times during 2020–2024 compared with 2015–2019. Tree-based algorithms also expanded substantially, with a 3.2-fold increase, while GA showed only modest growth of about 1.3 times. Boosting methods, which were rarely applied before 2019, experienced the strongest relative growth, rising nearly 11-fold over the last five years. These findings highlight a clear trend: NN and boosting methods are being adopted at accelerating rates, GA continues to attract steady interest, and tree-based approaches are strengthening as complementary techniques. Since previous review articles only covered trends up to 2020 [22], with partial updates through 2021 [25], this present study focuses specifically on 2020–2024, when the number of AI-related publications in structural engineering increased most sharply (Fig. 2).

Following the PRISMA flow diagram (Fig. 3), five steps are followed as the formal review protocol of this study (i.e., preliminary, identification, screening, eligibility and inclusion). Firstly, the preliminary stage involved analyzing recent review articles on AI in civil and structural engineering to identify research gaps. The review timeline was set to 2020–2024, given the sharp rise in AI applications during this period (Fig. 2) and the absence of prior reviews covering this interval. Identification of relevant studies was carried out using keywords such as AI in *reinforced concrete structures, steel structural frames, composite structural frames*, and *LCA*. Additional topic-specific keywords—such as *concrete mix design, concrete mechanical properties, durability, fire resistance of materials, structural response under lateral loads*, and *life cycle impact analysis*—were also employed, yielding 150 candidate papers. The screening process involved removing duplicates (n = 10) and excluding irrelevant topics (n = 20). This process includes works focusing on *frost durability resistance, carbonation depth, compression on composite columns, shear capacity of composite slabs*, and *behavior of beam–column joints*. Next, eligibility was confirmed through full-text evaluation, resulting in the exclusion of an additional 20 articles. Ultimately, this review selected around 100 papers, prioritizing studies that compared AI predictive results against laboratory-scale experimental and numerical datasets.

The application of AI in structural engineering is diversifying through innovative approaches such as its integration with BRBs, viscoelastic dampers, retrofitting technologies (e.g., fiber-reinforced jacketing), beam–column joints, and LCA of buildings. For example, AI has been applied in seismic protection systems to enhance energy dissipation and resilience using dampers and BRBs. In retrofitting, AI-driven models assist in identifying structural weaknesses and recommending cost-effective strengthening strategies. For beam–column connections, AI improves the prediction of joint behavior under cyclic loading. AI has also been introduced in small-scale LCA studies to evaluate sustainability and long-term structural performance. In this context, the present review highlights the limitations, challenges, and potential solutions for such emerging applications. It also addresses practical implications of AI, barriers to its adoption, and prospects for industry implementation. Specifically, this study evaluates prediction accuracy in terms of the coefficient of determination (R^2), the number of databases used, and key input parameters or governing factors across various applications (e.g., mechanical and durability properties of concrete members, fire-induced effects on structural components, seismic impact-based design, and LCA of structures). In addition, it examines accuracy levels achieved by optimized versions of traditional AI models. This review does not cover AI research trends for 2025, nor does it address structural health monitoring, remote sensing, or construction automation. Although R^2 is adopted as the primary performance benchmark, other evaluation metrics inconsistently applied across the literature (e.g., RMSE, MAPE) are not considered for sole benchmarking. Likewise, variations in dataset size and in characteristics among studies (e.g., for durability assessment of concrete or fire-induced effects) are not discussed in detail, as the focus remains on comparative evaluation of prominent algorithms with key input parameters.

In view of these considerations, this paper provides a comprehensive analysis of effective AI methodologies in structural engineering, emphasizing advancements between 2020 and 2024. Section 2 presents a critical overview of widely adopted AI techniques. Section 3 evaluates the predictive accuracy of these models relative to traditional numerical simulations, experimental data, and design codes, with applications in

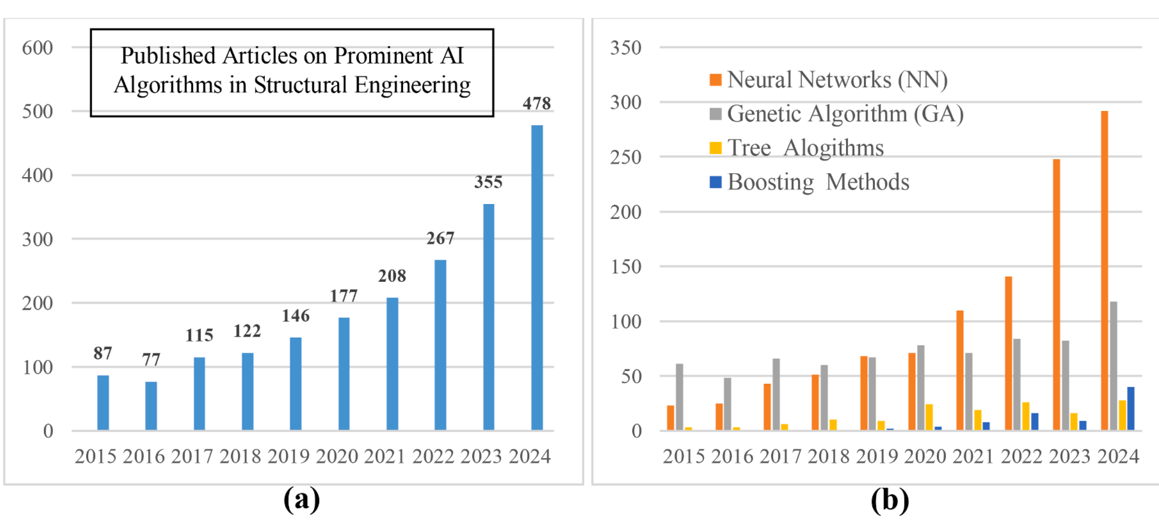

Fig. 2. Published article trends in Structural Engineering (Scopus-sourced, 2015–2024): (a) overall use of prominent AI algorithms, (b) categorized by specific methods (NN, GA, tree-based, and boosting algorithms).

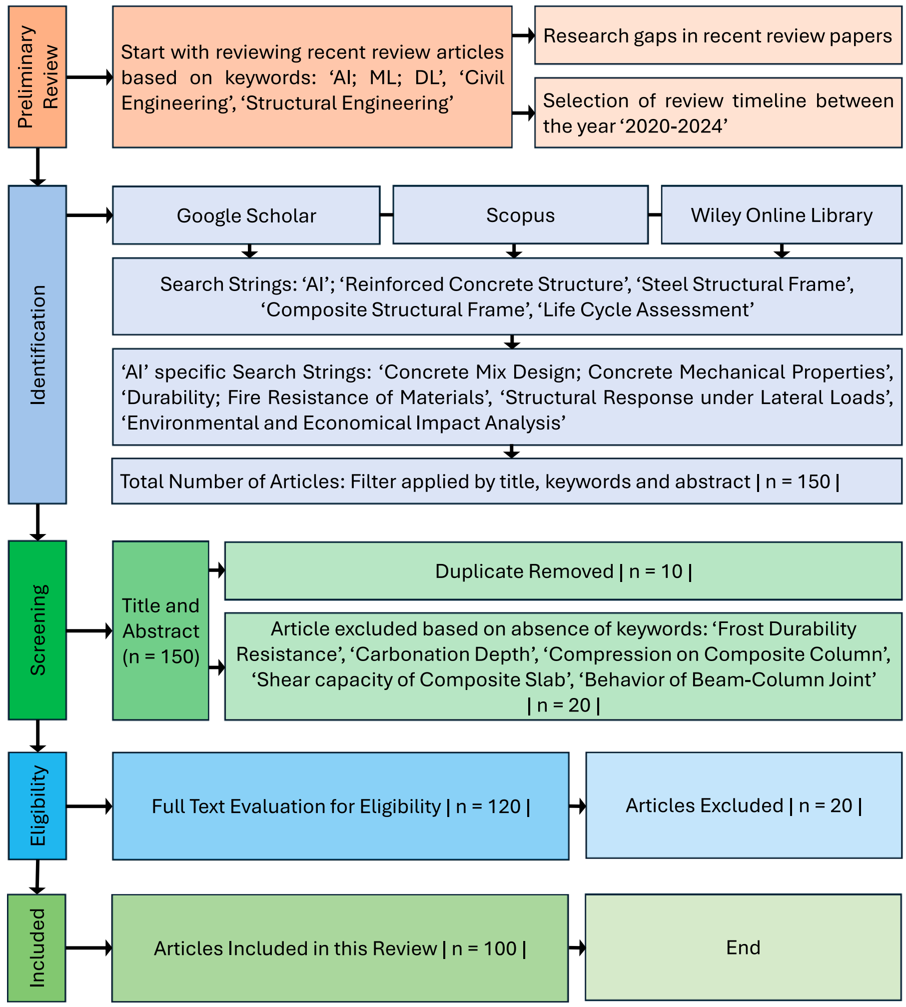

Fig. 3. PRISMA flow diagram for the reviewed articles.

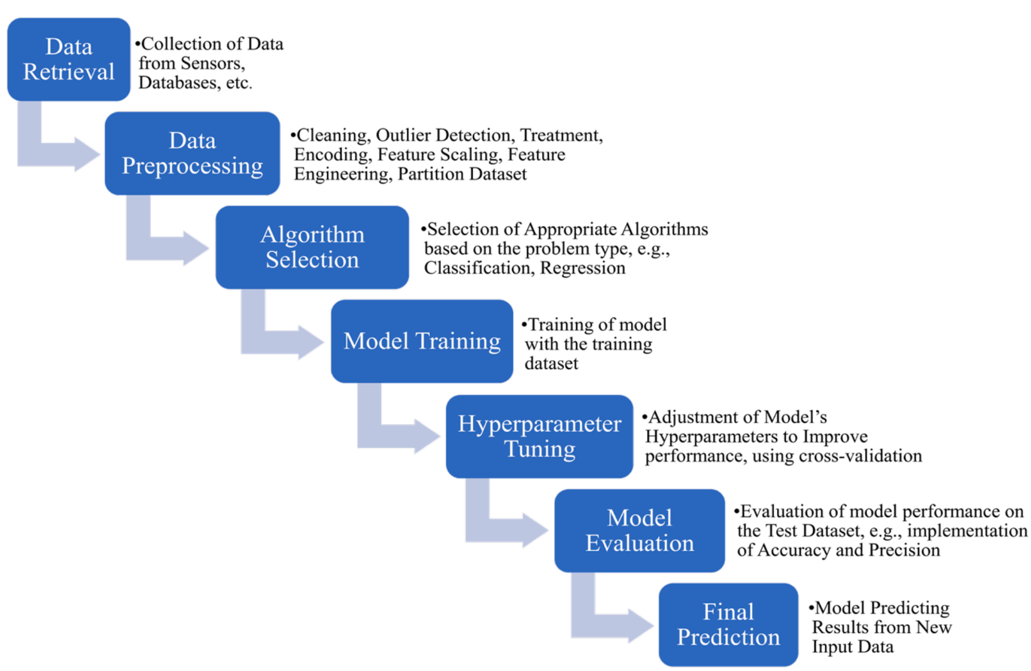

Fig. 4. Artificial intelligence (AI) model development in structural engineering.

concrete, reinforced and composite concretes, and steel structures, as well as in structural response to lateral loads and LCA. A summary of comparative structural analyses using AI methods is also included. Section 4 discusses limitations of current AI-driven research and the potential implications for industry. For the first time, it highlights future research directions such as AI applications in retrofitting technologies, BRBs, dampers, easy-to-dismantle beam–column connections, and LCA-driven sustainability assessments. By bridging the gap between AI advancements and structural engineering challenges, this review aims to serve as a practical guide for researchers and engineers seeking to integrate AI into structural analysis, design, and sustainability practices.

2. Overview of basic AI techniques relevant to structural engineering

This section introduces the AI techniques most commonly applied in structural engineering and discusses their applications in detail. It begins with the general process of AI model development (Section 2.1), followed by brief introductions to the widely used AI techniques in structural engineering (Sections 2.2–2.8). Finally, the evaluation of model accuracy and precision is addressed (Section 2.9).

2.1. Model development

A well-structured model development process (Fig. 4) is essential for generating reliable predictions. The process begins with data retrieval and preprocessing, which are critical because model performance largely depends on data quality. Typical preprocessing tasks include outlier detection and treatment, data encoding, feature scaling, feature engineering, and partitioning the dataset into training and testing subsets to ensure suitability for modeling. The next stages involve algorithm selection and model training using the training data. Training is usually performed iteratively, with hyperparameter tuning to optimize performance until satisfactory cross-validation (CV) results are achieved. Finally, the model's performance is evaluated using a separate test dataset, and predictions are generated for comparison with the observed outcomes [26]. This development framework is common to all AI models and serves as the foundation for the evaluation methods described in subsequent sections.

2.2. Neural networks (NN) architecture

To handle complex data relationships, NN consists of artificial neurons interconnected in a specific topology, designed to mimic the behavior of the human nervous system and adopting a structure analogous to the human brain. Artificial neural networks (ANNs), a basic form of NN, leverage their self-learning capability to produce highly accurate results as the amount of experimental data increases. By managing high dimensional data, ANN can solve highly nonlinear classification and regression problems, as well as complex relationships. Essentially, an ANN can be considered an information-processing system with specialized neurons for receiving external input and generating output [27].

As illustrated in Fig. 5, an ANN model consists of an input layer, one or more hidden layers, and an output layer, interconnected with randomly assigned weights and biases. For the input layer, the number of neurons (nodes) are equal to the number of variables of the specific problem to be solved. The product of the inputs and their respective weights is added to the deviation (bias). For the hidden layer that lies between the input and output layers, a predefined activation function is applied to the nodes to process the inputs. The optimal weights and biases are determined through training to minimize the error between the outputs and targets. The training is considered complete when the model achieves the desired performance with the smallest errors. The output layer produces the final response from the network when the model is considered suitable for generating predictions from unknown data [19].

ANNs can be classified based on the number of hidden layers, i.e., single layer perceptron for one hidden layer or multilayered perceptron (MLP) for two or more hidden layers [29]. Since an MLP is a deep ANN that has excellent ability of function approximation, it can be used in diverse engineering applications. It implements nonlinear transformations to convert input variables to an expected output. As shown in Fig. 6, similar to ANN, each layer of an MLP network contains several nodes (neurons) or processing elements that may be partially or fully connected. A feedforward process is executed between nodes of different layers, with each neuron processing an input and generating an output, which is then used as the input for the next neuron. The connection strength or weight between nodes includes independent values that are modified throughout the training stage in a process known as back-propagation. This combined process of forward signal propagation and backward error correction is referred to as feedforward backpropagation

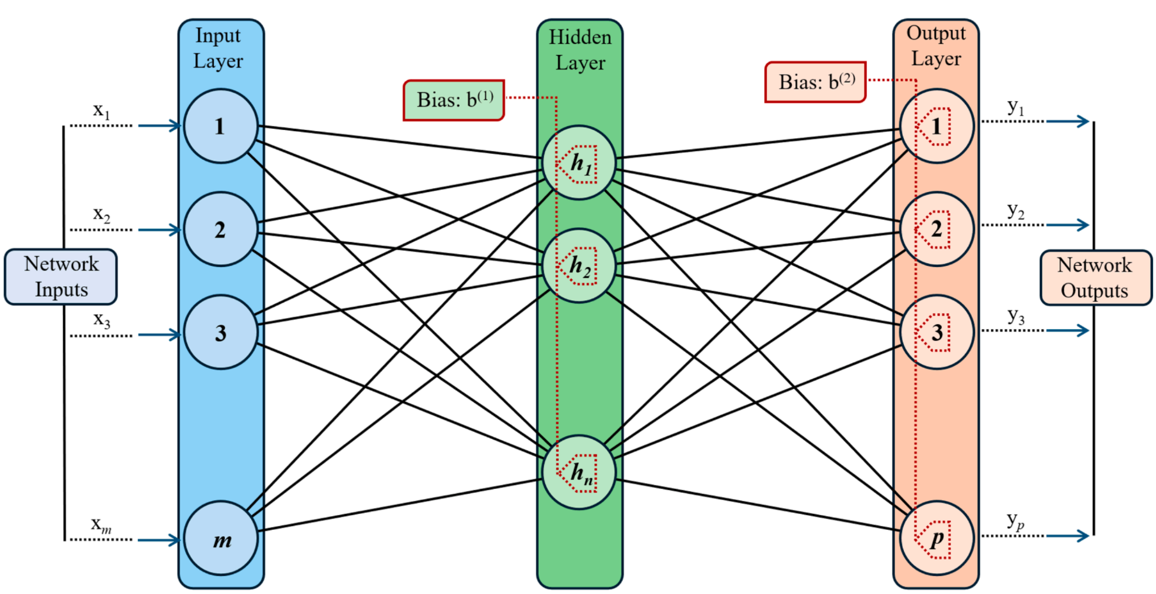

Fig. 5. Artificial neural networks (ANN) structure (adapted from [28]).

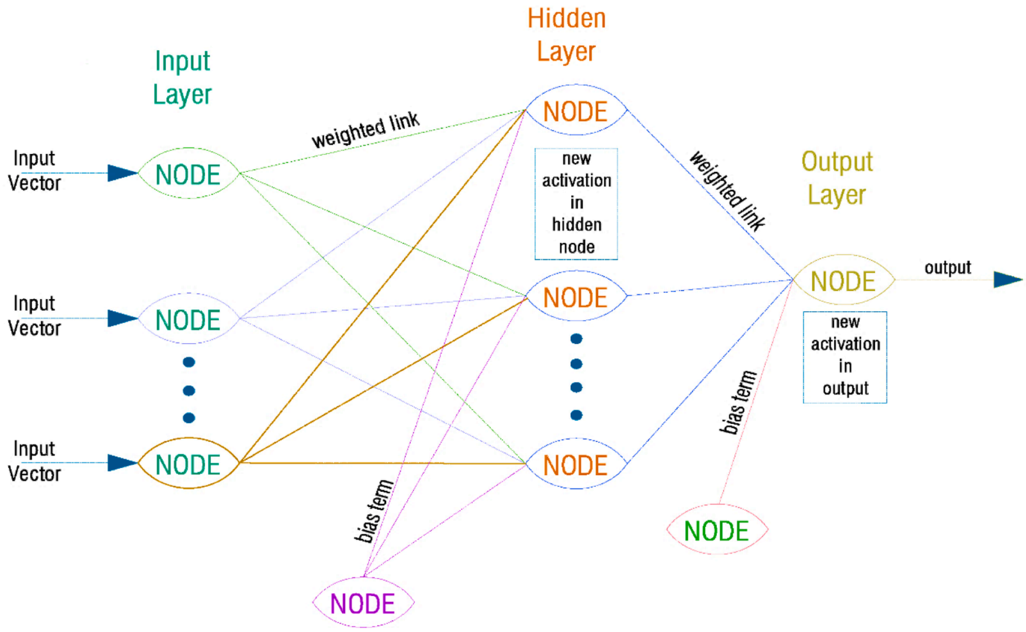

Fig. 6. Representative multilayered perceptron–artificial neural networks (MLP-ANN) framework (adapted from [31]).

(FFBP). The optimum set of weights, yielding the smallest errors, is subsequently used to perform predictions with new data [30].

Nonlinear autoregressive exogenous (NARX) is a dynamic recurrent ANN which can capture inherent intricate relationships between inputs in its memory and effectively predict the outputs. Especially for nonlinear discrete time series, NARX ANN can serve as a dynamic modeling tool that encloses multiple layers with feedback connections [32]. As shown in Fig. 7, two primary configurations exist: (a) the open-loop or series-parallel mode (NARX-SP) represents a stable network in which actual target values from previous tests are fed back during training process, and (b) the closed-loop or parallel mode (NARX-P) represents a network where predicted outputs are fed back as inputs for the feedforward neural network and incorporated in the output regressor due to the absence of true outputs for new data. Thus, the NARX-P mode is an FFBP network with a feedback connection from output to input. As such, it can generate final predictions using the training and test data from the NARX-SP mode [31].

Long short-term memory recurrent neural network (LSTM-RNNs) is a variant of the recurrent neural network (RNNs), and its basic modeling pattern is the same as that of ANNs. However, LSTM-RNNs consist of several decisive hidden layers apart from the input and output layers, as well as a group of LSTM cells with four interrelated units, i.e., an internal cell, an input gate, a forget gate, and an output gate (Fig. 8). By utilizing a self-recurrent connection, the internal cell remembers the cell state at the former time step, while the input gate regulates the input activation flow into the internal cell state. The forget gate allows the LSTM cell to forget or reset the cell memory, as necessary. The tangent gate or tanh (hyperbolic tangent) function transforms values (compressing between −1 and 1) before the values are read from cell state. The output gate normalizes the flow of output activation into the LSTM cell output [33].

2.3. Machine learning (ML) and deep learning (DL) algorithms

ML encompasses four types of learning methods, i.e., supervised, unsupervised, semi-supervised and reinforcement learnings [7]. In supervised learning that accounts for approximately 70% of ML applications, the algorithm is trained on an experimental dataset containing both inputs and outputs. The model predictions are compared with the true outputs to identify the errors, and the learning process is refined accordingly. Patterns are assessed to predict labeled information for additional unlabeled data. In contrast, unsupervised learning involves exploration of data for pattern identification in the absence of historical labels. This approach is well-suited for transactional data. Popular unsupervised learning algorithms (e.g., self-organizing maps, nearest-neighbor mapping, k-means clustering, and single-value decomposition) have been used to segment textual topics, recommend

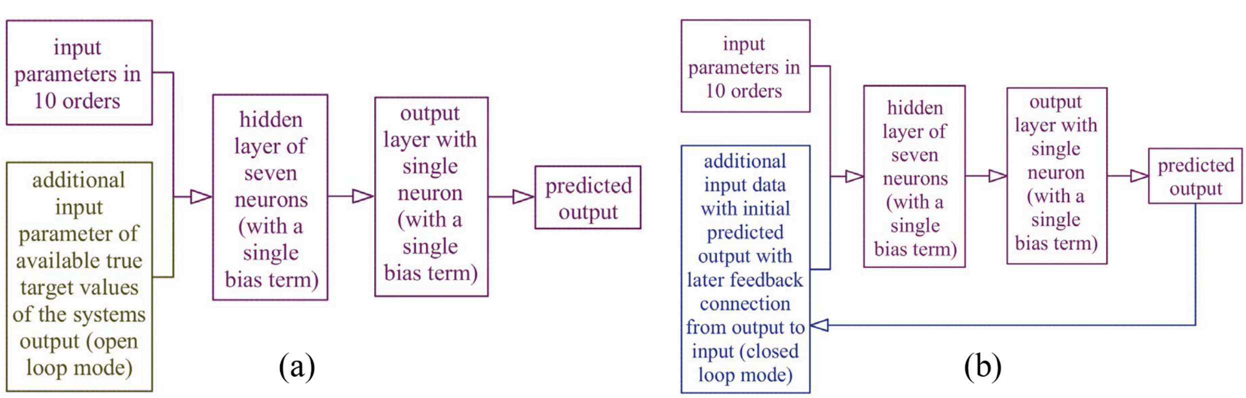

Fig. 7. Nonlinear autoregressive exogenous–artificial neural networks (NARX-ANN) frameworks: (a) NARX-SP; (b) NARX-P (adapted from [31]).

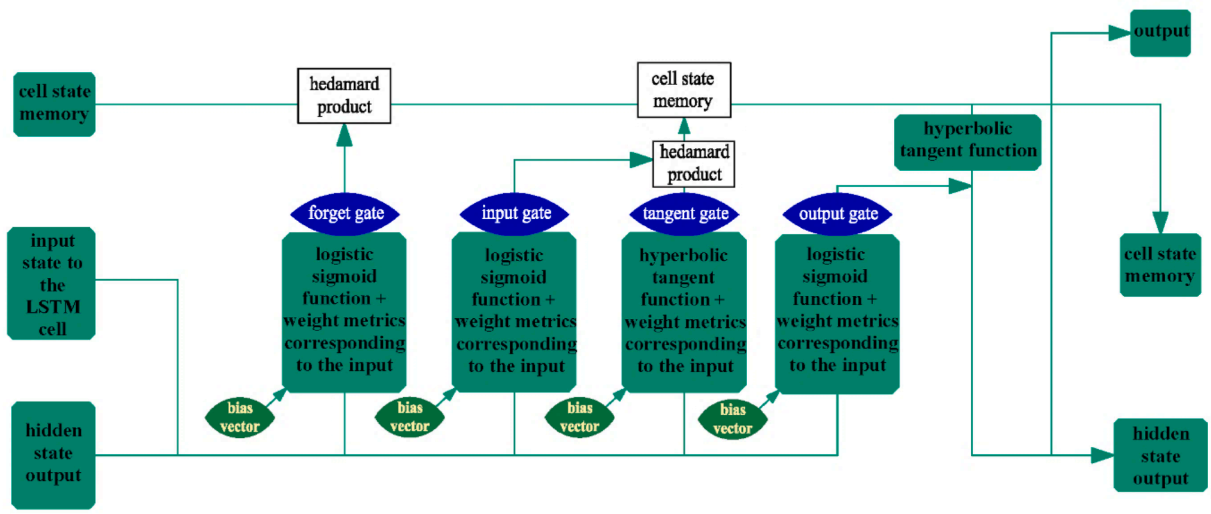

Fig. 8. Long short-term memory (LSTM) cell diagram (adapted from [33]).

items, and identify data outliers. Semi-supervised learning follows a pattern similar to, but its amount of unlabeled information is much higher than the supervised learning. Classification, regression, and prediction models are trained using this learning method. Reinforcement learning aims to learn the best options available by adapting a trial-and-error process involving three primary mechanisms: the learner or decision maker, environmental components, and actions. The goal of the learner is to adopt the best actions available to produce the expected result within a pre-scheduled time, following the most suitable pattern [7].

DL is a branch of ML that employs unsupervised networks to learn from unstructured or unlabeled data [34]. It starts with the input layers that are connected to a series of hidden layers through nonlinear activation functions. The activation functions generate approximation forms that allow gradient-based optimization. Results from the optimization process are displayed as final output. The main objective of DL architecture is to learn the feature illustration of input data and achieve implicit representation that best generates an output $Y = f\left(\sum_{i=1}^{n} ln_i w_{ij} + b_j\right)$, where f is the activation function, ln_i represents ith input signal, b_j represents bias value of jth neuron, w_{ij} represents connecting weights between ln_i and b_j [35].

Multiple hidden layers create deep neural networks (DNN) and more hidden layers result in deeper networks [34]. A variation of DNN is convolutional neural networks (CNN) which is specialized for image recognition. CNNs mimic the visual cortex to distinguish and classify images. The architecture consists of two main sections: feature learning and classification (Fig. 9). Initially, input images pass through the feature extraction network, where convolutional layers transform the images and pooling layers reduce dimensionality. The resulting feature maps are then fed into classification layers to generate predictions. In classification layers, flatten layer converts the 2D feature maps into a 1D vector allowing fully connected layers to process them as input. Soft-max layer produces probabilities for each class such as no visible cracks (connections), micro-cracks (partial degradation) and significant cracks (full separation) are processed as rigid, semi-rigid and damaged joints, respectively [36].

2.4. Naïve bayes (NB) and K-nearest neighbors (KNN)

NB is a simple multiclass linear classification algorithm that is based on Bayes' theorem [32]. Its learning process can be simplified using generative assumptions and parameter estimations. By using Bayes optimal classifier, the required number of equations of NB increases exponentially with an increase in the number of features. Hence, by simplifying Bayes classifier through appropriate assumptions in equations, the number of features can be significantly reduced. However, modifying one feature does not alter other features, as this method neglects possible correlations between features [26]. In contrast, KNN is a nonparametric ML algorithm that is used for both classification and regression, and it does not incorporate assumptions regarding the decision on boundaries [37]. For each test instance, the algorithm identifies the K most relevant data points (nearest neighbors) within the training dataset and predicts outcomes based on the most frequently occurring class among these neighbors. This approach, often referred to as the majority rule, is conceptually similar to the probabilistic reasoning used in Bayesian methods [38].

2.5. Genetic algorithm (GA) and particle swarm optimization (PSO)

GA and PSO are another type of AI methods that are broadly used for

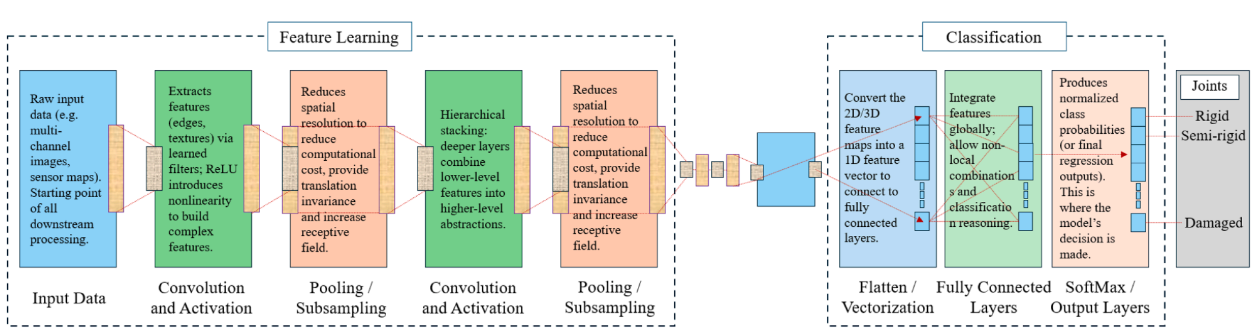

Fig. 9. Basic convolutional neural networks (CNN) model architecture (adapted from [36]).

optimization and searching. GA is a strategic model based on the principles of genetic evolution [39], focusing on the principles of survival of the fittest and adaptation. GA continuously produces new groups of genes (populations of chromosomes or strings) to perform a task by formulating old groups of genes. A GA contains three parts: (i) coding and decoding of variables into strings, (ii) evaluating the fitness of each solution string, and (iii) applying genetic operators (crossover, mutation) to generate the generations of next solution strings [40]. To derive accurate solutions (Fig. 10), an appropriate number of chromosomes (strings) must be selected, which are obtained in multiple phases. First, the necessity of reproducing a string is assessed based on its fitness function. If the optimal solution is not reached, a crossover operation creates modified offspring by combining parent genes, and mutation introduces additional variability. This process is repeated until an optimal solution is obtained [41].

Genetic programming (GP) is an extension of GA in which solutions are represented as computer programs, whereas GA typically produces numeric strings as solutions. The classical version, known as tree-based GP, constructs models as trees consisting of functions and terminals with a root node (Fig. 11). After generating an initial set of random models, successive generations are created using mutation, crossover, and reproduction, and the best program across all generations is selected as the output [43]. Gene expression programming (GEP) is a developed version of GP first invented by Ferreira [44] where new generations of models created by GP are represented as linear strings that are decoded and expressed as nonlinear entities (trees) [45]. Multi-expression programming (MEP) is a more advanced linear GP approach, where a single chromosome can encode multiple programs. Fitness values are evaluated across these programs to identify the optimal solution [46].

PSO resembles GA and is inspired by communal behavior of animals with five main features [41]: (i) Proximity: Simple calculations are performed in definite time and space; (ii) Stability: The system does not react to every environmental change; (iii) Quality: Significant changes in the environment are detected to ensure solution quality; (iv) Diverse response: No singular limitation exists in system response to environmental changes; (v) Adaptability: Changes in the environment are considered during optimization.

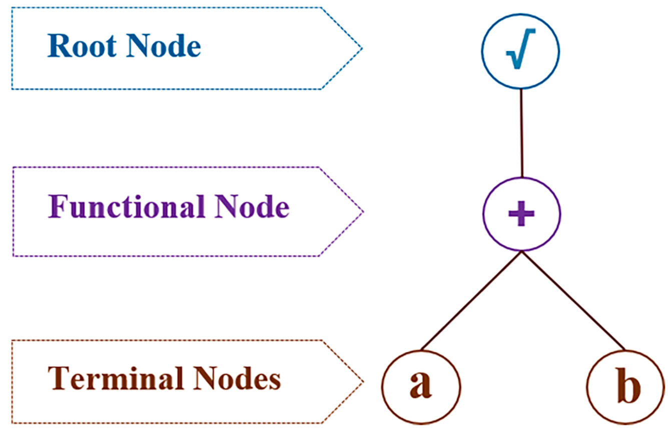

Fig. 11. A tree-structured genetic programming (GP) model (adapted from [43]).

2.6. *Gaussian process regression (GPR) and multivariate adaptive regression spline (MARS)*

GPR is a nonparametric model that systematically quantifies the prediction uncertainty of nonlinear high-dimensional problems with small simplistic samples. It has a simple training process by selecting appropriate functions according to the pattern in the training data. By setting the initial values and optimizing the hyperparameters using the input training data, prior distributions are determined, and prior model is transformed into posterior model, respectively. Thus, GPR offers adaptability in hyperparameter selection with flexible nonparametric inference. Finally, it performs its prediction using the regression prediction model [19]. Another method, MARS, is suitable for generating solutions to problems with continuous numerical outcomes and high input dimensions. Similar to GPR, it can perform nonparametric and nonlinear regression. It partitions the input space into subgroups and fits piecewise regression models within each subgroup using basic functions. This process enables MARS to capture complex data structures and identify potential interactions among input variables across all degrees [19].

2.7. *Tree algorithm-based models and boosting methods*

Decision trees (DT), also referred to as regression trees (RT), are nonparametric models that solve classification and regression problems by recursively splitting data into a hierarchy of simple decisions based on one or more input features. A typical DT structure is shown in Fig. 12 and involves two key steps: (1) *tree-building*: training dataset is

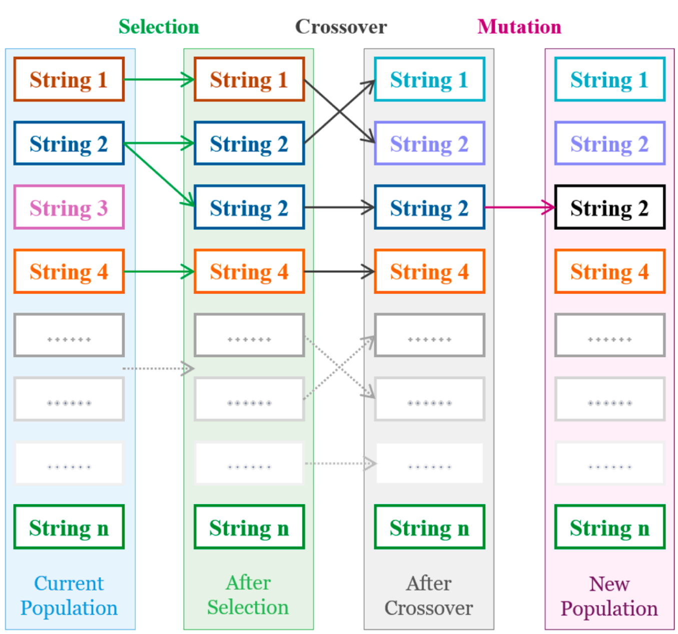

Fig. 10. Operational structure of genetic algorithm (GA) (adapted from [42]).

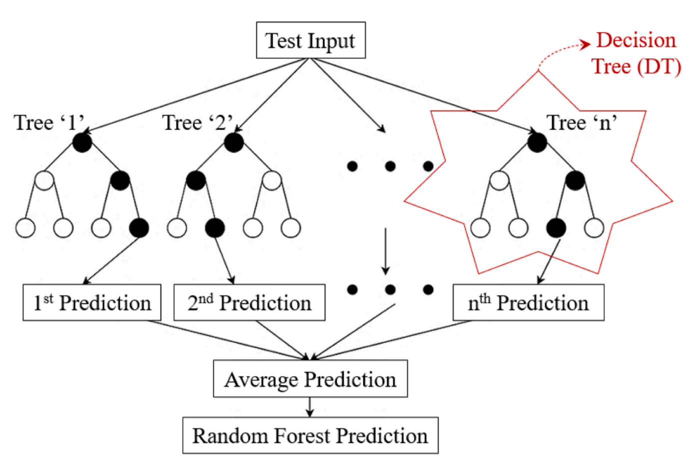

Fig. 12. Tree pattern of random forest (RF) model (adapted from [47]).

partitioned into non-overlapping regions, and (2) *tree-pruning*: reduces overfitting by trimming unnecessary branches of the tree [37]. Random forest (RF), also called an ensemble of decision trees, consists of multiple DTs operating together (Fig. 12).

Each tree generates predictions on new data, and the final output is obtained via majority voting for classification or averaging for regression. Overfitting in individual trees is mitigated because multiple trees contribute to the final result [47]. The performance and robustness of DTs and other single predictive models can be significantly enhanced using ensemble ML methods. One such method is the bagging regressor (BR), which primarily relies on bootstrap aggregating. In this process, multiple copies of the original dataset are generated through resampling (bagging) [48]. Data points are randomly selected from the original dataset with replacement to create bootstrap samples, suggesting some original data may not appear in certain samples. Finally, base models (e.g., DTs) are then trained on these samples, and predictions on new data are combined, typically using majority voting for classification or averaging for regression [49]. Bagging reduces variance, variability and noise in predicted output by training multiple models in slightly different data [50].

Apart from BR, some boosting techniques can enhance the performance of DT by merging a set of weak classifiers to form a strong classifier. Among them are adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), light gradient boosting method (LGBoost), natural gradient boosting (NGBoost), gradient boosting regressor (GBR), categorized boosting (CatBoost), and histogram gradient boosting (HGBoost) [51]. In AdaBoost, all observations are weighted equally, and the model is retained by correctly classifying the incorrectly classified observations with higher weights than usual. In this manner, the learners are trained using the weighted classification accuracy of the previous learners [52]. In contrast, the remaining methods mentioned above are variants of gradient boosting (GB) framework which performs gradient optimization on the contribution of each weak learner to reduce the overall error of the strong learner (Fig. 13) [53]. XGBoost leverages the misclassification error of the prior model, although the need for successive model training leads to slow processing. LGBoost operates leaf-wise rather than depth-wise, thus providing more precise but complex trees aiming at computational efficiency. It poses enhanced training speed, greater efficiency, improved precision, lower memory consumption and competence to process large datasets [54]. NGBoost generalizes GB to estimate the parameters of a conditional probability distribution as target for a multiparameter boosting algortihm [55] and GBR deals with regression problems [56]. CatBoost can control the categorical features of the input parameters during the training phase by operating in the preprocessing stage [37]. HGBoost employs histrogram based methods interpreted by DT to handle efficient bulk dataset [51].

2.8. Support vector machine (SVM)

SVM was developed by Vapnik [57] that uses optimal separating hyperplane to separate positive and negative classes of datapoints with the farthest possible two marginal boundary lines (Fig. 14). Support vectors are derived from datapoints representing the separating hyperplanes in a transformed space [26]. Fig. 14(a) shows several possible classifiers separating the datapoints, while one optimal separating hyperplane separates the data with maximum margin as further shown in Fig. 14(b). SVM can also be operated through regression applications as support vector regressor (SVR) [56]. As an extension of SVM, SVR aims to find a hyperplane with a specific number of dimensional space (input parameters) that classifies the training datasets in different classes. It differs from SVM as it targets a flat type of hyperplane that accepts the data points within or on the margins and rejects data points outside the margins [58].

2.9. Key performance metrics and model validation techniques

2.9.1. Performance evaluation

There is no predefined method for determining the optimal architecture of a model. However, the accuracy of a model layout can be evaluated using several performance criteria, such as root-mean-square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). RMSE quantifies the magnitude of errors and is sensitive to outliers. MAE utilizes a scale similar to that of the data to compute the variance between predicted and target values. MAPE measures the range of errors in percentages. R^2 represents the proportion of the difference in predicted values that can be explained by the model. These metrics can be mathematically expressed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{n=1}^{m}(e_n - \hat{e}_n)^2} \tag{1}$$

$$\mathrm{MAE} = \frac{1}{m}\sum_{n=1}^{m}|e_n - \hat{e}_n| \tag{2}$$

$$\mathrm{MAPE} = \frac{1}{m}\sum_{n=1}^{m}\left|\frac{e_n - \hat{e}_n}{e_n}\right| \tag{3}$$

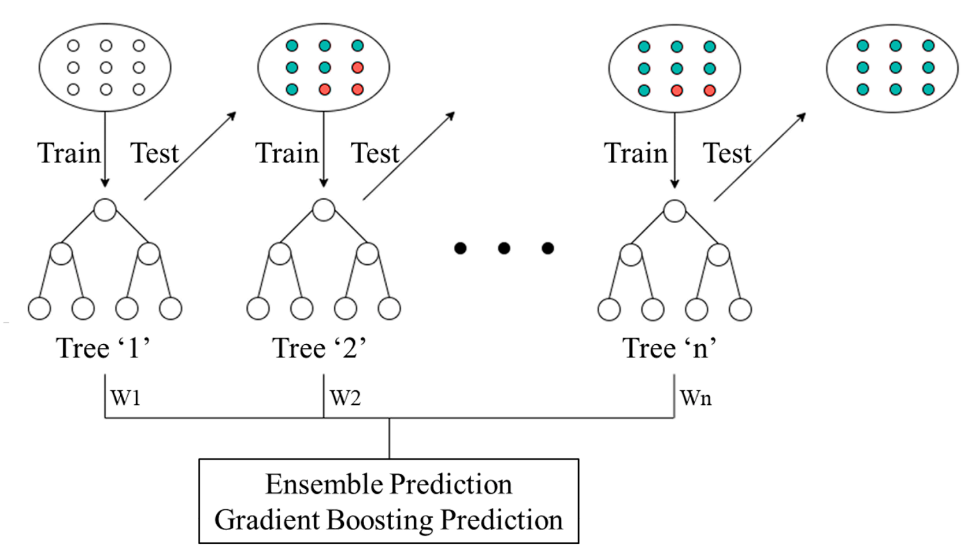

Fig. 13. Gradient boosting (GB) pattern (adapted from [47]).

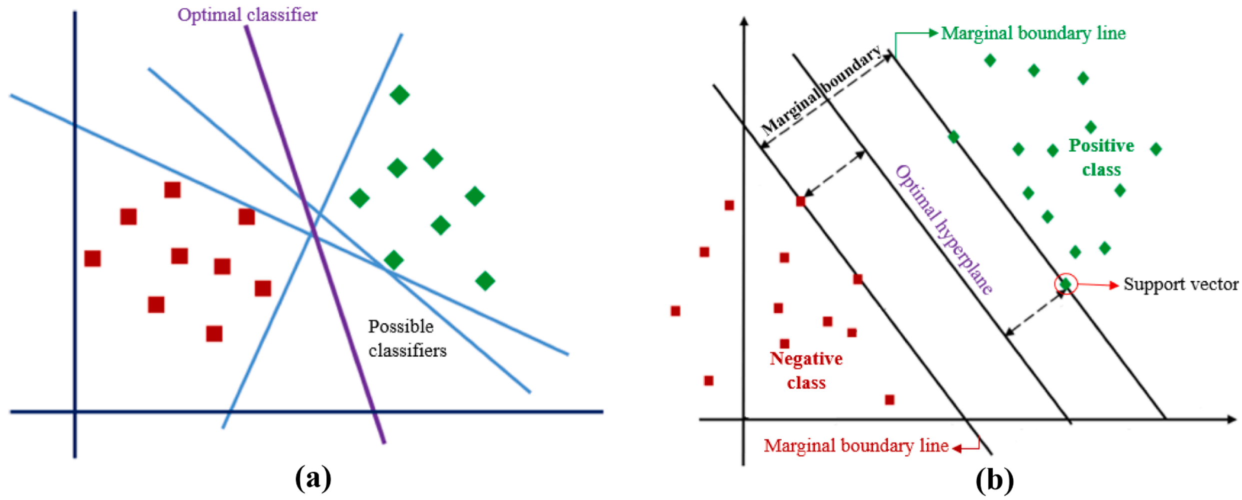

Fig. 14. Decision boundary from support vector machine (SVM) (adapted from [59]).

$$R^2 = 1 - \frac{\sum_{n=1}^{m}(e_n - \hat{e}_n)^2}{\sum_{n=1}^{m}(e_n - <\mathrm{ct}> \bar{e} <\mathrm{ot}>_n)^2}\bar{e} = \frac{1}{m}\sum_{n=1}^{m} e_n \tag{4}$$

where m, n, $\bar{e} - \bar{e}_n$, e_n, and $\hat{e}_n$ denote the number of data points, the data sample index (runs from 1 to m), the difference in mean of observed values, the target value, and the predicted value, respectively. Although various evaluation metrics have been identified [189], RMSE, MAE, MAPE, and R^2 remain the most widely used for AI-predicted results in engineering applications [190,191]. However, these metrics have limitations. For example, RMSE is not always appropriate for comparing accuracy across time series [192], while MAPE can be unreliable and misleading [193]. R^2 is dimensionless, allowing comparison across heterogeneous target variables (e.g., MPa, mm/m, %) and enabling cross-dataset benchmarking. Nevertheless, earlier surveys reported that, although R^2 is the most commonly used metric in engineering, it can sometimes be biased, insufficient, and misleading; meanwhile, other metrics also present challenges [194]. In fact, no single metric can be considered universally superior [195]. In this study, most structural engineering papers published between 2020 and 2024 consistently reported R^2 as the primary benchmark for evaluating predictive accuracy, typically alongside complementary metrics such as RMSE, MAE, and MAPE. R^2 is unitless, with values ranging from 0.0 to 1.0 [196]. For RMSE, the normalized values below about 5% are generally regarded as excellent predictive accuracy and for MAE, lower values (e.g., < 0.1 for displacement prediction) demonstrate strong alignment with engineering tolerances [227]. For general cases, MAPE values under 10–15% are frequently considered acceptable, while errors above 20–25% indicate weaker model performance [228].

2.9.2. Model robustness and accuracy

Although R^2 is a popular and intuitive measure of accuracy in AI-based structural engineering models, it is important to recognize its limitations to ensure proper interpretation. First, R^2 values are scale-dependent, meaning that the same R^2 can correspond to very different levels of absolute error across datasets, which limits cross-study comparisons [208]. Second, it naturally increases with model complexity, even when added predictors are not meaningful, making it sensitive to overfitting unless paired with adjusted R^2 or error-based metrics [209]. Third, it is insensitive to systematic bias: a model can consistently underpredict or overpredict and still yield a high R^2 if variance trends are well captured [210]. Fourth, it does not provide information about the distribution of errors, meaning outliers may be masked, while metrics like RMSE penalize such deviations more strongly [210]. Finally, because R^2 is bounded above by 1 but unbounded below ($-\infty$), negative values do not clearly quantify the degree of model inadequacy [208]. To address these issues, structural engineering studies often report heterogeneous metrics such as RMSE, MAE, and MAPE to complement R^2; for example, Yang and Liu [211] demonstrated that their model's high R^2 was further substantiated by substantially reduced RMSE and MAE compared to code predictions. Similarly, normalized or relative error measures help standardize error magnitudes across studies (e.g., RMSE or errors normalized by mean or range) so that results are more comparable across datasets with different scales [212]. Thus, while R^2 remains a valuable benchmark for variance explanation, pairing it with error- and scale-sensitive metrics provides a more holistic and trustworthy evaluation framework for AI applications in structural engineering.

Monte Carlo simulations (MCS) provide a robust method for evaluating model performance and reliability. This sampling-based methodology involves performing many simulations of the same process to converge the average of large samples to an anticipated value for infinite samples. This approach is suitable for randomizing the sampling method of training and testing dataset for the selected models. Subsequently, a specific number of simulations are performed with different dataset splits for training and testing. By employing MCS, random sampling helps enhance model accuracy by reducing errors in AI predictions. For example, if 10 samples are selected from among 100 samples that contain 50% each of two different types of data, the correct proportion of the two types of data may not be achieved. Therefore, certain variations may appear as sampling errors in the predictions. The random sampling process can eliminate bias in parameter selection, thereby maximizing accuracy [60].

2.9.3. Data minimization

The availability of input parameters generally improves the accuracy of model predictions. However, unnecessary data, with no influence on the final prediction, can introduce system noise, mislead the training programs, and compromise the model's performance. The influence of such input parameters can be regulated through a sensitivity analysis. The most widely used methods for sensitivity analysis include the cosine amplitude method, Milne method, and generalized cross-validation (GCV) method, as discussed in the following sections [19]:

2.9.3.1. Cosine amplitude method. This method quantifies sensitivity by computing the correlation between the input and output data. The correlation factor (R_i) is calculated as:

$$R_i = \left(\sum_{k=1}^{m} x_{ik} y_k\right) \Big/ \sqrt{\sum_{k=1}^{m} x_{ik}^2 \sum_{k=1}^{m} y_k^2} \tag{5}$$

where x_{ik} is the value of the i-th input parameter corresponding to the k-th output, y_k is the value of the k-th output, and m is the total number of samples. A higher value of R_i represents a stronger correlation between the input and output parameters [19].

2.9.3.2. Milne method. This method analyzes the effects of inputs on outputs based on the weights of the connections between nodes, represented by a static weight matrix:

$$\text{Influence of input parameter} = \frac{\sum_{m=1}^{N_h} \frac{w_{ml}}{\sum_{l=1}^{N_i} |w_{ml}|} w_{pm}}{\sum_{k=1}^{li} \left(\sum_{m=1}^{N_h} \left| \frac{w_{jk}}{\sum_{l=1}^{N_i} |w_{ml}|} \right| \right)}, \tag{6}$$

where N_i represents the number of input parameters (i), N_h represents the number of hidden neurons, w_{ml} represents the weight of the connection between the input neuron l and the hidden neuron m, w_{jk} represents the weight of the connections between node j and node k, and w_{pm} represents the weight of the connection between hidden neuron m and output neuron p [19].

2.9.3.3. GCV method. The significance of an input parameter can be defined as the square root of the GCV of the model with all basic functions containing the eliminated parameter, minus the square root of the GCV of the full model. The GCV can be defined as follows:

$$GCV = MSE_{train}/(1\text{-enp}/n)^2 \tag{7}$$

MSE_{train} represents mean square error of the training data, n represents number of samples in the training data, and *enp* represents effective number of parameters/variables [19].

2.9.4. Data interpretability

As mentioned in Section 2.9.1, RMSE, MAE, MAPE, and R^2 are commonly used in regression problems to evaluate model performance. Beyond accuracy, SHapley Additive exPlanations (SHAP) is an interpretability method that provides insights into feature contributions in regression tasks. For example, SHAP values quantify the influence of each feature on the predicted continuous outcome. SHAP is employed to interpret the decision-making process of complex AI models by providing post-hoc explanations of model predictions [61]. Similar to parametric analysis, SHAP isolates the individual contribution of each input parameter to the model's output. This facilitates a clear understanding of the inherent reasoning behind predictions and allows the relative influence of each parameter on the predicted results to be distinguished [51]. Feature importance refers to analytical techniques that quantify the relative contribution of each input variable to the predictive performance of a machine learning model, typically through model-based or permutation-based measures. Recent engineering studies show that such methods enable researchers to identify which design or material parameters most strongly influence target responses, offering insights that align with established physical or experimental understanding. For example, Nguyen et al. [222] used SHAP on a boosting model for concrete-filled steel tube columns and found that the most important features were exactly the expected mechanical loads (bending moment and axial force). In other words, SHAP highlighted that these domain-critical forces drive the model's predictions, validating the AI with known structural physics. Similarly, an RF-based framework for multi-distress prediction in continuously reinforced concrete pavements highlighted the dominance of structural and environmental variables, further emphasizing the interpretability of feature importance in linking AI outcomes to engineering behavior [223].

However, SHAP and related interpretability methods should be applied with caution, as correlated features may distribute importance scores misleadingly. Additionally, issues such as data leakage during preprocessing or feature engineering can inflate interpretability results, limiting their generalizability. On the other hand, precision and recall are frequently employed evaluation metrics for classification problems due to their ability to capture the trade-off between false positives and false negatives [62]. Precision represents the percentile of successful predictions for each predicted result, while recall measures the fraction of relevant instances. Both metrics are considered more effective when their values approach 1 [37]. Generally, regression metrics (RMSE, MAE, MAPE, and R^2) are considered separately from classification metrics (precision, recall, and F1), with precision and recall applied only when the prediction target is categorical. Only a small proportion of published machine-learning models undergo true external validation – testing on wholly independent data – despite its recognized importance for unbiased performance assessment [224,225]. This gap often reflects practical constraints, but it means many models are evaluated only on internally held-out or cross-validated samples. Dataset sizes are frequently inadequate [226], and most ML studies fail to justify sample-size calculations [226]. Proper validation protocols (e.g. k-fold or nested CV) are therefore essential to mitigate overfitting and bias [224,225]. In addition, performance reports should include measures of variability – such as the standard deviation (SD) or coefficient of variation (CoV) of repeated runs, and 95 % confidence intervals (CI) for key metrics – rather than only point estimates [225]. Advanced techniques like Bayesian hyperparameter optimization can help tune models efficiently, and explicit uncertainty quantification (via bootstrapped CIs or Bayesian credible intervals) provide critical insight into model robustness. Finally, assessment should follow established external-validation criteria and report overall performance indices (e.g., accuracy and calibration metrics) with their uncertainties to fully characterize model generalizability [225].

3. Application of AI in structural engineering

AI is increasingly being applied to predict the mechanical properties, environmental impacts, durability, life cycle, and service life of structural materials such as concrete, reinforced concrete (RC), composites, and structural steel. Large-scale laboratory experiments can be complemented with AI-based predictions to minimize testing costs and time. The following sections provide an overview of recent studies that have employed AI models in structural engineering.

3.1. Materials performance prediction

3.1.1. AI in concrete mix design and mechanical properties

Recent studies have applied various AI and ML methods—such as KNN, ANN, BR, GPR, SVM, DT, RF, MLP, GEP, boosting, and stacking techniques—to optimize concrete mix design and predict its mechanical properties. These approaches have primarily focused on forecasting the compressive strength of concrete incorporating recycled aggregates and supplementary cementitious materials (e.g., slag, silica fume (SF), fly ash, and ground granulated blast-furnace slag (GGBFS)), using datasets ranging from 1000 to 3600 samples. Among these, XGBoost [63,64] and stacking methods (an ensemble technique combining multiple models) [65] achieved the highest accuracy, with an R^2 value exceeding 0.950. The most influential parameters were concrete testing age, cement content, and the replacement ratio of recycled coarse aggregates (CA). Similarly, high R^2 values of 0.960 and 0.970 were reported using GEP for SF-concrete [66] and BR for geopolymer concrete [67], respectively. Generic CV (not specified) was applied to reduce overfitting, but no CI or SD values were reported. The influencing parameters were cement and water, identified via sensitivity analysis over a dataset of 283 compressive and 149 tensile samples. The use of multiple ML models and CV improved reliability, but the absence of explicit k-fold details and external datasets limits generalization [66]. Onyelowe et al. [68] examined the mix design of fly ash–incorporated concrete using statistical analysis, linear regression, and AI algorithms, where ANN achieved the best performance (R^2 = 0.92) in predicting 28-d compressive strength. The study used 112 mix samples with binder ratios as inputs, and models were validated by statistical comparisons, with uncertainty of 15–20 MPa (SD for compressive strength models) and 2–3 points (environmental impact models). The governing parameters were the fly ash-to-binder ratio and aggregate-to-binder ratio and all methods. The study demonstrated the capability of ANN to develop a robust mix design tool for sustainable concrete with comprehensive parametric considerations considering.

GB and XGBoost also outperformed other methods in a study by Kang et al. [69], where the water–cement ratio and SF content were identified as the most critical parameters affecting the compressive and flexural strength of steel fiber–reinforced concrete, based on a dataset of 220 samples. Similarly, GB was shown to be more accurate and robust in determining the flexural strength of fiber-reinforced concrete beams, achieving a higher slope validation ratio (0.83/1). The depth of the beam was the most influential factor, followed by the flexural reinforcement area [70]. In contrast, Khan et al. [71] reported ANN to be superior to RF, reaching an R^2 of 0.990, while using ~120 FRP beam samples with geometric, reinforcement, and material inputs. For validation, dataset was split into training/test sets with error indices RMSE of 7.37 kN-m. This result was further validated by Zhang et al. [72] over 134 data points, where ANN ($R^2 = 0.979$) outperformed GEP and existing ACI 440.11–22 [188] equations. Li et al. [73] recommended XGBoost ($R^2 = 0.93$) for predicting the flexural strength of concrete with cementitious materials. Using SHAP analysis, the water–cement ratio and curing age were identified as the key factors. Khan et al. [46] used GEP and MEP to determine the flexural capacity in FRP-strengthened beams using 200 samples. Validation was based on holdout sets, not k-fold CV. No uncertainty intervals were provided, though strong R^2 values (0.96–0.98) and low MAE demonstrated accuracy. SHAP analysis identified beam width, depth, and reinforcement ratio as key predictors. Overfitting was managed with evolutionary learning techniques, but the study was limited to smaller dataset size and lack of external validation.

The prediction of split tensile strength for concrete containing different cementitious materials (e.g., GGBFS and Portland slag), sand replacements, and recycled aggregates was examined using up to twelve AI methods on 168, 310, and 381 dataset points. The peak R^2 values achieved were 0.98 (XGBoost) [74], 0.892 (extra tree regressor) [75], and 0.842 (XGBoost) [76], respectively. The common influencing factors were water-cement ratio, curing age and ratio of cementitious materials. Nguyen et al. [56] implemented four ML methods, i.e., SVR, MLP, GBR, and XGBoost, to determine the strength characteristics of high-performance concrete. Among these, SVR and XGBoost offered the most accurate results (R^2 of 0.96–0.98) with reduced computational effort via random-search-tuned train/test validation. Key governing factors were cement contents, blast furnace slag, fly ash, water–cement ratio, superplasticizer, coarse and fine aggregates (FA), and curing period. Overfitting was reduced by efficient parameter tuning and data imputation strategies, though the absence of k-fold CV or external test sets limits robustness.

Earlier, Gholampour et al. [77] developed empirical models using GEP to predict the 28-d compressive strength, elastic modulus, flexural strength, and splitting tensile strength of recycled aggregate concrete (RAC) samples. A comprehensive database with 650 compressive strength values, 421 elastic modulus values, 346 splitting tensile strength values, and 152 flexural strength values from previous reports was compiled to compare the performance of 34 mechanical-property models for RAC, developed in 21 other studies. The proposed GEP model provided more accurate results than other models on large datasets and exhibited consistency with existing code expressions. Similarly, a hybrid GP model was developed in a study to predict the triaxial compression loading based on 330 tests on concrete samples: comparisons with earlier studies across several statistical criteria confirmed the model's accuracy and reliability. The GEP-based approach using only the mix-design properties as predictors achieved $R^2 = 0.81$ [17]. In another study [47], four models (RF, NN, GB, and AdaBoost) were applied to predict the fatigue life of concrete under uniaxial compression. The dataset, containing 1300 sets of experimental data, was split into 90:10 for training and testing. Instead of CV, the authors focused on dataset cleaning and feature analysis. Six key input variables, related to the material and dimensional properties (compressive strength of concrete, height-to-width ratio, and shape of test specimen) and the loading conditions (maximum stress level, minimum stress–to–maximum stress ratio, and loading frequency), were adopted. Maximum stress level and frequency were the most influential features. The GB model yielded the lowest error and high predictive accuracy ($R^2 \approx 0.915$) across the three datasets. Overfitting risk was reduced by outlier filtering, though the absence of CV and uncertainty measures limits confidence in generalization. Cascardi et al. [78] proposed an ANN-based analytical model to predict the compressive strength of circular concrete columns wrapped with fiber-reinforced polymer (FRP). The compressive strength of the FRP-confined concrete was influenced by the column diameter, compressive strength and Young's modulus of the unconfined concrete, and thickness of the FRP jacket. A total of 465 samples were included in the database, and the key parameters were column diameter, unconfined concretes' compressive strength, thickness and Young's modulus of FRP jacket. The model achieved high consistency and accuracy (R^2 of 0.83) when compared with results from laboratory tests and equations derived from international codes and scientific literature.

Few researchers have used ML to assess the self-healing ability of cementitious materials. Rajczakowska et al. [20] compiled a detailed database with 197 records by extracting results from 12 experimental studies to predict the compressive strength recovery of concrete using four interpretable ML methods: SVM, RT, ANN, and ensemble of RT. The 12 input variables were water-cement ratio, concrete age, cement content, fine and CA, peak loading temperature and its duration, cooling regime and duration, curing regime and duration, and volume of samples. The stability of the models was verified through Monte Carlo analysis. The Ensembled RT achieved the highest accuracy ($R^2 > 0.900$) and robustness. The most influential parameters were temperature, curing regime, curing time, and aggregate amounts. Huang et al. [79] assessed 797 bacterial self-healing concrete test results with 22 features to determine self healing properties. ML models including GBR, SVR, RF, and DNN were compared, with GBR performing best ($R^2 = 0.956$). 10-fold CV and grid search optimization were applied to reduce overfitting. RMSE was used for error estimation, but no CI was reported. The key parameters were bacteria type, healing time, crack width, and environment. Overfitting risk was explicitly addressed with CV and sensitivity analysis, improving robustness. Some other studies also reported promising accuracy such as GEP with an R^2 of 0.938 for admixture-based concrete using 619 data points [80], and BR with an R^2 of 0.974 for engineered cementitious composite based concrete using 617 crack data samples [81]. Most of the contributing variables were associated with FA, cementitious materials (e.g., fly ash, SF, and limestone powder), water–binder ratio, and crack width before self-healing.

3.1.2. Durability and fire resistance of materials

Predicting the long-term serviceability of structural materials is crucial for effective structural design. Several studies have explored optimal performing methods for examining key durability factors (DF). For example, ANN achieved an R^2 exceeding 0.950 in predicting moisture exposure using 429 observations. Regression models were tuned via Bayesian optimisation and evaluated on a held-out test set; classification used stratified 10-fold CV. The study does not report SD/CI for model outputs; it reports standard predictive metrics (MSE, RMSE, and MAE). Important influential variables include exposure duration and environmental factors (relative humidity and temperature), alongside geometrical and material properties. The use of stratified CV and Bayesian hyperparameter tuning reduces overfitting risk but the study provides limited formal uncertainty quantification [82]. In similar manner, MEP achieved R^2 of 0.921 and 0.977 in modeling concrete corrosion using

256 experimental records (chemical and biological tests). Models were trained on a 50/50 split, and the performances were reported using MSE and R^2 (MEP performed best). As the inputs were very small (time ± pH), the influential parameters were limited to exposure time and pH. The study did not report formal uncertainty bounds (SD/CI) and external test set comparisons — this limits quantified generalization/overfitting analysis beyond train/test errors [83]. In other studies, SVM (88–89% accuracy; 204 datasets) and back propagation NN (85% accuracy; 159 specimens) were used to predict chloride resistance [26,84]. Khan et al. [50] found peak accuracy from BR (an R^2 of 0.999) in predicting depth of wear with 216 datapoints, and SHAP analysis identified testing time and specimen age as the dominant features. The study addressed overfitting by using ensembles, objective function minimization, and external validation metrics — although formal SD/CI intervals were not reported for predictions. External validation criteria and performance index were used to support model generalization claims. In predicting frost resistance and impermeability, peak accuracies were observed for RF (R^2 = 0.950) using 100 groups of orthogonal–experimental data samples and for optimized ANN (R^2 = 0.926) using 417 sets of experimental data, respectively [85,86]. In another study, hybrid ANN achieved an R^2 exceeding 0.990 in modeling carbonation penetration with 532 data records where all models were trained over 70/15/15 (training/test/validation) split and ten-fold CV. Uncertainty is presented via SD and fold-by-fold MAE/RMSE (no classical CI). The top influential parameters were exposure time (≈27%), CO_2 concentration (≈22%), and water–binder ratio (≈18%). Overfitting was controlled by using CV and validation partition, and the identified errors were reduced by applying hybrid ANN across training, validation, test, and CV folds. However, the authors still recommended enlarging the database to further lower overfitting risk [87]. Across these studies, the governing parameters were duration of exposure, volume fraction of CA, cement content, water–binder ratios, FA, supplementary cementitious material (SCM) contents, thickness of protective layer, and ratio of environmental to relative humidity. Liu et al. [19] predicted the frost durability of RAC based on the DF using three soft computing models: ANN, GPR, and MARS. The database contained experimentally measured DF values of 142 samples from 23 published studies. The ANN model (with 19 neurons) achieved the highest accuracy (R^2 = 0.951) followed by GPR, with the lowest RMSE and MAPE. Sensitivity analysis identified air-entrainment as the critical factor influencing frost durability. Parametric analysis further showed that frost resistance improved with reduced recycled aggregate replacement, higher air-entrainment, lower water–cement ratio, and an optimized sand–RAC ratio.

At elevated temperatures due to fire exposure, structural materials often lose strength and durability. Because full-scale fire tests on structural prototypes are difficult to conduct, there is a growing demand for numerical and AI-based predictive models. Numerous studies have explored AI for predicting fire-induced effects on structural components. Concrete is widely used by fire engineers, but all concrete types may fail under extreme fire exposure [88]. Liu et al. [89] examined the thermal spalling of steel and polypropylene fiber-reinforced concrete. Among six tested AI models, XGBoost achieved the highest accuracy (R^2 = 0.972), with polypropylene fiber content identified as the key parameter for preventing spalling. Habib et al. [90] evaluated six classification models on fire-exposed fiber-reinforced concrete beams using 50 experimental tests, with AdaBoost demonstrating reasonable accuracy (R^2 = 0.90), followed by GB. In another study, SVR, RF, and DNN were employed using a compiled database to predict the fire resistance of FRP-strengthened RC beams. DNN performed best (R^2 = 0.910), with critical parameters including geometrical features of the beam section, applied loading, and thermal properties of fire insulation [29]. Similarly, ensemble models (XGBoost, CatBoost, LGBoost, HGBoost, GB, and RF) achieved accuracy exceeding 0.90, outperforming traditional ML models (ANN, DT, PR, and SVM) when trained on 21,000 data points from numerical simulations. SHAP analysis revealed the most significant negative factors as loading ratio, FRP area, and total applied load, while the positive factors were total area of steel reinforcement, thickness of insulation on beam sides, and steel reinforcement cover depth [51].

Composite structures are widely used in industrial buildings and commercial spaces. As steel is more prone to failure than concrete in fire exposure, construction with composite structures is often preferred, and related AI-based research has received significant interest. Moradi et al. [91] implemented an ANN model using 300 experimental data points to evaluate the fire resistance and strength behavior of concrete-filled steel tubes. The model achieved R^2 values of 0.967 and 0.970, a more accurate model than the existing empirical relationships. However, these tubes require protective fire coatings, making concrete encased steel columns superior in fire resistance. Naser [92,21] studied the fire behavior of RC columns using a combination of AI methods, including intelligent PR, GP, DL, and traditional multi-linear regression. The study analyzed 112 test observations under standard fire conditions for exposure durations of up to 5 h. The governing factors associated with concrete performance were concrete mix proportion, components (aggregate type and water–cement ratio), and supplementary additives (e.g., superplasticizers, fibers, and SF). In predicting fire-induced spalling, GP achieved the highest accuracy (R^2 = 0.940), followed by DL. Moreover, DL also performed better in predicting the relationship between the governing factors and the fire resistance of concrete columns. Li et al. [93] developed ANN and analytical models to predict the buckling resistance of axially loaded concrete-encased steel columns exposed to fire conditions, considering 15,200 specimens. The ANN and analytical models achieved R^2 values of 0.990 and 0.950, respectively, and showed lower dispersion (smaller SD) than the analytical equations. Validation was conducted by an 80/20–train/test–split and by comparison with experimental fire tests. The temperature of concrete and steel sections were affected by concrete grade, heating time, section factor, and thickness of concrete cover. The very large synthetic dataset and direct comparison to experimental tests reduce overfitting risk, though explicit regularization or nested-CV details were not found in the examined parts of the study. Naser and Kodur [35] used a dataset including 494 observations to develop a systematic ML (combining ensemble of RF, XGBoost and DL) approach that enabled explainable and rapid assessment of fire resistance and fire-induced spalling of normal- and high-strength RC columns. This ensemble could analyze 5000 reinforced columns within 60 s by incorporating a wide range of geometric characteristics and material properties and it achieved an R^2 of 0.86. Although tunnel fires are relatively infrequent [94], the growing number of tunnels has resulted in catastrophic incidents worldwide [95]. Wu et al. [96,97] investigated fire source behavior, hazards, and critical temperature fields in tunnels using LSTM-RNN and transpose-CNN under 100 simulated tunnel fire scenarios. The LSTM-RNN model achieved an accuracy of 0.90 with recommended 20 m sensor spacing, while the transpose-CNN model achieved ∼0.97 accuracy with 32 sensors placed at 5 m intervals. Both models effectively identified critical temperature fields, providing valuable insights for safe evacuation, emergency response, and firefighting strategies.

Globally, the structural frames of most tall buildings, older buildings, and large-span warehouses are either composed of or supported by steel components. Fire remains a critical hazard with potentially catastrophic consequences for steel and steel–concrete interfaces. Engineers must therefore accurately assess design parameters to ensure fire safety, a process increasingly supported by AI-based predictive models. Fu [60] developed an ML framework incorporating DT, KNN, and NN to rapidly predict the failure patterns of simple steel-framed buildings subjected to fire and to assess the potential for subsequent progressive collapse. Failure patterns were defined using the critical temperature method, and MCS and random sampling were performed to develop a sufficiently large dataset for training and testing. The KNN and NN models provided satisfactory predictions of the failure pattern and collapse potential of a two-story, two-bay steel-framed building. Data driven ML models such as ANN, RF, GB and KNN have also been used to explore the performance of the concrete-steel bond under high temperatures. Al Hamd and

Warren [98] analyzed 316 data points from previous laboratory-based studies and found GB to be the most accurate model ($R^2 = 0.970$), with the other models also yielding consistent results. The key input features were concrete compressive strength, testing age, concrete surface temperature at failure, thermal saturation ratio, bond length-to-diameter ratio, cover-to-diameter ratio, and fiber volume. The dataset was validated with train-test split, and uncertainty was reported with RMSE of 1.08–2.62 MPa and CoVs ranging from 18 to 74%.

3.2. Structural behavior analysis

3.2.1. AI in reinforced concrete and steel structural elements

A recent extension in GP, a GEP-based nonlinear model, was developed by Gandomi et al. [99] to assess the shear resistance of RC beams with shear steel. The database comprised 466 experimental measurements for both high- and normal-strength concrete beams. The proposed model outperformed existing design-code models, achieving an R^2 value approximately 0.89. Sensitivity analysis revealed that concrete compressive strength, web width, and effective depth were the key factors controlling the variations in the shear resistance of RC beams with stirrups. Cascardi [100] developed an analytical ANN model to predict the in-plane shear strength of masonry panels retrofitted with fiber-reinforced mortar. The study considered different varieties of masonry types (by material and texture) and reinforcement, in terms of both the fiber (glass, carbon, steel, basalt, and phenylene-benzobisoxazole) and matrices (cement, lime, and hydraulic mortars). Despite the large diversity in the input parameters, the model demonstrated high precision and accuracy ($R^2 = 0.91$), demonstrating robustness and sensitivity, with predictions consistent with results obtained using international design codes. To assess rapid damage and seismic risks, and determine appropriate retrofitting strategies, Mangalathu et al. [37] developed a comprehensive database of 393 one-story, one-bay RC shear walls with both rectangular and non-rectangular sections. The dataset included 152 flexural failure, 96 flexure–shear failure, 122 shear failure, and 23 sliding failure samples. The model performances were evaluated using three metrics: global accuracy, precision, and recall. Among the eight ML algorithms, RF achieved the highest accuracy (0.86), with a recall of 70% and precision of 84% in identifying the flexure–shear failure mode on the test set. The aspect ratio of the shear wall, boundary element reinforcement indices, and wall length–to–thickness ratio were the critical factors governing the failure mode. Retaining walls provide permanent lateral support for vertical soil slopes in infrastructure such as roads and bridges. In one study, a modified SVM model outperformed alternatives in determining the safety criteria of cantilever-type retaining walls [101]. Key parameters, including cohesion, angle of shearing resistance, angle of wall friction, and reliability index, were computed using the first-order–second-moment method. The modified SVM predictions deviated from synthetic reference values by <2%.

An ANN model was developed to predict the ultimate compressive load of rectangular concrete filled steel tube columns in both concentric and eccentric loadings. The dataset included 1224 test results for both long and short specimens. The model showed improved accuracy compared with available design codes (50% reduction in RMSE), and the most influencing parameters were steel tube dimensions, thickness, and material strength [102]. For a similar case, Asteris et al. [103] developed three alternative models using optimized ANN with a hybrid database of 1857 specimens. These models outperformed code-based methodologies, with reduced RMSE by 34%. Lemonis et al. [104] developed an ANN model to predict the ultimate axial compressive capacity of square and rectangular concrete filled tubes. The database included experimental results of 1193 long, thin-walled and high-strength specimens. The model offered satisfactory results, with a 20% error margin for 92% of the specimens. Sensitivity analysis revealed that the influencing factors were the tube dimensions and steel yield limit. Ferreira et al. [105] built a finite element (FE)-based database, and trained five ML models to predict global shear capacity of a steel-concrete composite down-stand cellular beams with precast hollow-core units. Among the models, the Catboost regressor algorithm showed optimal performance (R^2 of 0.982), followed by GEP (R^2 of 0.953), using >6 geometrical features (e.g., opening diameter, web opening spacing, tee-section height, concrete topping thickness, interaction degree, and number of shear studs above web opening). The FE-based database and the reliability analysis were used to quantify prediction uncertainty at the design level, but per-prediction SD/CI numbers were not presented in the examined portions of the study.

Lateral torsional buckling resistance, including web-post buckling and web distortional buckling of slender cellular beams were accurately predicted (R^2 of 0.99) by developing ANN formula on 768 training models [106], validated with a 70/15/15 (training/validation/testing) hold-out split. A 7-neuron model was chosen for stability and practicality as overfitting risk increased with more neurons. The key input parameters were beam dimensions, eccentricity from shear center and moment gradient factor, and uncertainty was quantified via RMSE (1.2–2.2), MAE(0.6–1.5) and SD. Degtyarev [54] proposed an interactive notebook to predict elastic buckling (3645 FE datasets) and ultimate loads (78,390 FE datasets) of steel cellular beams using FE method optimized with seven ML models (DT, KNN, RF, XGBoost, GBR, LGBoost, and CatBoost). The ML models were in remarkable agreement with the numerical data and surpassed design codes (GBR with an R^2 of 0.997). The key influencing factors were beam span length, flange width, and web thickness. The study used 10-fold CV for validation, but did not report uncertainty intervals. Overfitting was addressed with CV and model comparisons, while relying solely on FE-generated data. Shamass et al. [107] implemented a MATLAB-based graphical interface design tool by utilizing ANN with an overall accuracy of 0.932. This tool integrated data generation from FE analysis, web-post buckling resistance predictions, and failure mode classification of perforated steel beams with elliptical web openings. For similar case, Rabi et al. [108] used a total of 10,764 web-post FE models on high strength steel beams. The dataset was further employed to train and validate different ML methods (ANN, SVR, and GEP), achieving R^2 values of 0.998, 0.999, 0.977, respectively. These methods were compared with analytical model ($R^2 = 0.982$), and a novel design model was proposed. The study used 10-fold CV and grid search for tuning, and reported detailed statistics including SD and CoV to quantify uncertainty and compare generalisation. SVR showed the lowest CoV and smallest RMSE in the reported comparisons. Overfitting was explicitly assessed via CV and training-validation splits. Degtyarev et al. [61] applied the NGBoost model to predict the probabilistic load-bearing capacities of laterally restrained cellular beams subjected to uniformly distributed loads, considering all possible failure modes and their interactions. A database with 14,094 numerical simulation results was considered, and the model was further interpreted with SHAP method. The dataset was validated with 10-fold CV (80/20 train–test split) and uncertainty was reported with CoV (≈ 0.014 across test data). The model significantly outperformed the existing design provisions with an R^2 of 0.999 while offering probabilistic predictions.

Le et al. [41] investigated the prediction capability of two hybrid AI models, GA and PSO, combination with a modified ANN to determine the buckling loads of 420-MPa high-strength steel Y-section columns with slenderness ratios of 30–80. The dataset included 57 buckling test results from previous studies. The input variables were column length, cross-sectional geometry, and initial geometrical deviation in the horizontal and vertical directions. Both models performed well, but the PSO combined with the modified ANN model achieved a higher R^2 value of 0.929. Gandomi et al. [17] used GEP to construct an accurate empirical prediction model that could relate the load capacity of castellated steel beams (CSB) to their geometrical and mechanical properties. Considering the nonlinear collapsible characteristics of CSBs, the GEP model and derived equation outperformed a multivariable linear regression and conventional constitutive models based on first-principle investigations (e.g., elasticity and plasticity theories). Because of the repetitive nature of beam–column connections, compact and efficient designs are required to reduce fabrication costs while maintaining quality. However, the vast number of connection types and loading scenarios makes obtaining sufficient experimental data from laboratory setups impractical. Abdollahzadeh and Shabanian [109] addressed this by using both mechanical modeling and an NN-based approach to simulate the complex hysteresis behavior of beam–column connections with flange plates. The combined NN approach accurately captured the narrowed hysteresis behavior, with RMSE = 0.712 and MAPE = 0.9166. Paral et al. [36] developed a DL-based nonparametric approach that used continuous wavelet transforms of acceleration signal and 2D CNN for image recognition to facilitate condition assessment of structural connections. Updated FE models were used to train the CNN model, which successfully identified damaged locations and measured the stiffness loss in the damaged beam-column joint. The study considered 80% training and 20% testing data split for hold-out validation method.

3.2.2. *Structural response under lateral loads*

Inelastic dynamic analysis based on modern building codes is widely used to accurately determine the seismic response of building frames. However, for large-scale structures, such analysis becomes computationally intensive and difficult to implement. To address this challenge, predictive models can be employed to achieve sufficiently accurate results with significantly reduced computational demand, thereby facilitating seismic analysis and optimization of large structures. Moreover, these models support the design and installation of structural solutions such as retrofitting, BRBs, and dampers. In a recent study, an AI-enhanced computational method was proposed by integrating AI with a shear building model, to determine the nonlinear seismic response of RC frames under displacement controlled quasi-static cyclic loading and dynamic earthquake ground motions. The database included test results of 272 RC columns. Numerical results showed reductions of 60% and 62% in RMSE and MAE, respectively, indicating that the proposed method outperformed existing physics-based and classical fiber-based models. In particular, the AI technique accurately used real-world experimental data to evaluate the lateral stiffness, and the shear model efficiently formulated the structural stiffness matrix [110]. Another study highlighted the capability of wavelet-weighted least squares-SVM and an FFBP-ANN to predict the inelastic force- and displacement-based seismic responses of an 18-story RC frame. The model was trained with design–basis and maximum ground earthquake motions. The study showed how training sample size (75/150/225) and choice of inputs (first three natural period combinations) affect accuracy (assessed by MAPE, NRMSE, R^2). Uncertainty was expressed via these error metrics and performance sensitivity, emphasizing robustness across sample sizes to address overfitting risk. The ANN model achieved slightly higher accuracy ($R^2 = 0.999$ with 225 samples) while exhibiting lower sensitivity compared with integrated SVM model [111]. Gondaliya et al. [112] applied a probabilistic framework combining classification models (ANN) and regression models (LASSO regression, RF, and GB) to assess the seismic response of a four-story RC building frame under epistemic uncertainty. The models achieved high accuracy, ranging from 0.87 to 0.97. To investigate the ultimate load-bearing capacity of inadequate RC frames, six ML models were developed and validated against experimental and numerical analyses of the load–displacement behavior of a one-story frame. Among these, RF performed best, achieving an R^2 of 0.870. The most influential input parameters were axial load, rebar diameter, and concrete strength [113]. As a solution to such inadequate capacity of columns, additional confining pressure can be provided by utilizing FRP-retrofitting jacketing with internal grouting, which prevents failure under extreme seismic and blast conditions. Shin et al. [114] proposed a rapid decision-making tool for multi-hazard assessment and mitigation using ANN models capable of rapidly generating large datasets. The ANN-based models achieved $R^2 = 0.98$ over 78 samples under seismic loading and $R^2 = 0.99$ over 83 samples under blast loading. In a follow-up study, Shin et al. [115] developed hybrid ML models which could optimize the retrofit details within desired performance by controlling the confinement and stiffness ratios. First, ANN was used to rapidly generate seismic and blast responses, and then GA was employed to optimize the retrofit details within multiple objective functions. The ANN model achieved a high regression value of 0.994 using dataset from FE simulation-based ML models. For the model, validation conducted against full-scale dynamic seismic tests and blast tests of RC frames, and reliability reported via small simulation variations (<12% for seismic and <3% for blast) compared to experimental tests.

Few studies have incorporated AI methods to determine seismic response of structural steel building frames. In a recent study, a portal frame was analyzed through four machine learning models (RF, GB, XGBoost, and DNN) to determine its top floor displacement under lateral load. RF outperformed others ($R^2 = 0.987$) in predicting displacement, and XGBoost also demonstrated satisfactory performance in determining failure probability. While not explicitly stating CV, the study used a holdout test set to validate models. Uncertainty was not reported via SD or CI; instead, the study presented a series of performance metrics (e.g., RMSE, MAE, and MAPE) [116]. Later, RF was hybridized into three variants: RF dragonfly optimization algorithm (RF-DOA), RF sparrow search algorithm (RF-SSA), and RF whale optimization algorithm (RF-WOA). RF-WOA outperformed RF-DOA and RF-SSA, offering engineers a valuable tool for designing portal frames with enhanced features. The study adopted a train/test split validation strategy and compared among the hybrid models through rank analysis and regression line performance. Uncertainty was not expressed via SD/CI, but through reliability indices and error/rank metrics. Influential features include structural and loading variables incorporated into the RF models. Overfitting risk was mitigated by comparing multiple hybrid configurations and leveraging reliability analysis, though no explicit CV was declared [117]. Seismic fragility analysis traditionally requires sophisticated numerical models and significant computational resources. By contrast, ML models can efficiently identify high-dimensional input variables and capture complex nonlinear relationships. In line with this, four ML models—RF, AdaBoost, GB regression tree (GBRT), and XGBoost—were employed to construct fragility curves based on nonlinear time-history analyses of 616 steel moment frames subjected to 240 ground motions. The models were trained on 56,479 datapoints, and a graphical user interface was developed using best performing models (GBRT and XGBoost, both achieving $R^2 = 0.999$). The inputs consisted of structural descriptors and the first three natural periods, capturing the essential dynamic properties of the frames. Model training and evaluation used a 70/30 holdout split, with hyperparameter tuning performed (e.g., number of trees, and learning rate) to enhance generalization. The very large dataset helped minimize variance and overfitting risk, and the ensemble models were chosen specifically for their robustness. Feature-importance analysis highlighted the natural periods as particularly influential in predicting fragility parameters. No independent experimental test dataset was used, since all data came from simulation. However, the size and diversity of the generated dataset strengthen external validity. The models were implemented using Scikit-learn v0.22.2 in Python [118]. Automatic seismic design was explored by Guan et al. [18], who developed a nonlinear structural model to simulate the static–dynamic response of steel moment-resisting frames (SMRF) using a Python-based end-to-end modular platform. Automatic seismic design and analysis (AutoSDA) was implemented as the first module to generate SMRF designs (such as section sizes and detailing) for beams, columns, and beam–column connections. The input design parameters included building characteristics (e.g., number of stories, number of lateral-force-resisting systems, and building dimensions), applicable loads (i.e., dead and live loads on each floor), and site conditions (mapped spectral acceleration). OpenSees was then used to create two-dimensional nonlinear structural models based on these designs. This module performs nonlinear static and dynamic analyses to comprehensively evaluate seismic performance. The model-based framework and object-oriented programming structure made the platform easily adaptable, efficient, reliable, and accurate. Zhang et al. [33] developed an LSTM RNN-based DL approach to model and predict data-driven nonlinear structural seismic responses. Specifically, two schemes were developed: LSTM-f (full sequence-to-sequence mapping) and LSTM-s (stacked sequence-to-sequence mapping), both incorporating multiple LSTM layers and fully connected layers to create time-dependent and causal input–output sequence models. The approach was validated through three case studies: a nonlinear hysteretic system (100 data samples), a six-story instrumented building with field sensing recordings (23 earthquake records), and a three-story nonlinear SMRF (548 datasets generated via incremental dynamic analysis). Among the models, LSTM-s demonstrated superior precision ($R^2 = 0.99$), reliability, computational efficiency, robustness, and scalability compared with LSTM-f and a classical ANN (MLP).

In seismic design, BRBs and supplemental dampers (e.g., steel plate, viscous and viscoelastic dampers) are important devices that provide high stiffness, ductility, and energy dissipation to lateral-force-resisting systems. BRBs, with a stable yielding core and an outer restraining member, exhibit symmetrical hysteresis and absorb large inelastic deformations, significantly enhancing a structure's energy dissipation capacity. Steel dampers likewise provide additional damping by stiffening the frame, absorbing vibration energy, and reducing seismic loads, which enhances overall dynamic response and structural resilience. Moreover, the use of BRBs and dampers together can reduce damage during earthquakes in a synergistic manner. BRBs are focused on yielding in sacrificial braces, while dampers are responsible for dissipating energy and limiting displacement. In a study on BRB frames, four ML methods (RF, ANN, XGBoost, and AdaBoost) were applied across six brace–frame configurations using 79,500 FE-based pushover analyses in OpenSeesPy. Inputs included frame geometry, BRB core area, section properties, and loads. The dataset was divided into 80/20 train-test splits, repeated across different configurations and for combined data. No uncertainty intervals were reported; performance was assessed using R^2, RMSE, MAE, and MSE. A graphical interface with the most accurate model (XGBoost, R^2 of 0.983–0.993) was developed, and feature importance analysis showed the base-shear to be most significantly governed by number of stories, followed by BRBs core area. While AdaBoost achieved perfect R^2 in training, its testing accuracy dropped, indicating overfitting. XGBoost provided the most balanced performance, reducing overfitting risk [121]. Conventional concentric braces face several limitations, including low ductility, asymmetric behavior under tension and compression, strength deterioration, and stiffness degradation. To address these limitations, AlHamaydeh et al. [31] combined an FFBP with NARX-ANN to predict the nonlinear hysteric behavior of BRBs under cyclic loading with 4 full-scale BRB specimens. Normalized brace forces during load reversals, as a response to normalized time-delayed inputs to the NARX-ANN, were denormalized using the auxiliary FFBP-ANN. Brace deformation was used as the input variable, while brace forces were set as the output variable in the proposed model. The model captured both linear deformations with corresponding linear forces and nonlinear deflections with corresponding nonlinear forces, with predictions closely matching experimental results (accuracy between 0.969 to 0.981). The ANN-based model outperformed the traditional FE modeling approach for the following reasons: (i) it established closed-form relationships between the input and response data; (ii) it learned and adapted to different types of data, (iii) it formed a simple structure, which facilitated reconfiguration and ensured significantly faster simulation runs. Sun et al. [213] applied ML methods to perform seismic fragility analysis of large-scale steel BRBs. A database of 9000 nonlinear time-history simulations was created to generate training and testing data. Input features included ground motion intensity measures, BRB design parameters and structural responses. Models were validated with 10-fold CV, achieving good generalization without severe overfitting. Uncertainty was quantified through fragility curves with confidence bounds, while peak values for predictive metrics such as R^2 (0.986) and RMSE (0.056) were also reported for XGBoost model. The large synthetic dataset and CV approach reduced overfitting risk, though no real-world external testing was performed. Tamimi et al. [214] combined FE modeling, ANN, and MCS to evaluate the seismic reliability of BRBs. Experimental tests (5 specimens) validated the FE model, which was then used to generate simulation data. Sensitivity indices revealed that gap size, friction coefficient, and steel core thickness were the most influential parameters. A bias factor distribution (mean = 0.99, SD = 0.038) quantified prediction uncertainty. Overfitting was mitigated by filtering non-influential variables before ANN training and using Monte Carlo for robustness. Although k-fold CV was not performed, the study ensured generalization by integrating ANN with reliability-based simulations. Anand et al. [215] developed ML models to predict seismic engineering demand parameters (maximum inter story drift, residual drift, and maximum and cumulative ductility) of BRBs. A database of 16,694 nonlinear time-history analyses records of BRB frames was generated from OpenSees models. Nine ML algorithms were tested with hyperparameter tuning via 10-fold CV on the training set, followed by evaluation on the test set. XGBoost emerged as the best-performing model, with peak R^2 values of 0.963 (maximum inter story drift), 0.928 (residual drift), 0.968 (maximum ductility), and 0.983 (cumulative ductility). Influential parameters identified by SHAP included spectral accelerations at 1–5 s, Arias intensity, and peak ground velocity. Overfitting was minimized through CV and the large dataset, but external validation was not performed. Sagheer et al. [216] developed a deep learning framework that combines ResNet for classifying BRB specimen types and LSTM for predicting their cyclic hysteretic responses. The study used experimental data from six specimens (threaded, shaved, and notched core bars), expanded into thousands of training sequences by resampling and segmentation. An 80/20 split was used for validation, supplemented by dynamic hyperparameter tuning. ResNet classification reached up to 99–100% accuracy with $R^2 \approx 0.993$, while LSTM achieved force prediction–index of agreement values ranging from 0.979 to 0.999, demonstrating very high fidelity. Overfitting was mitigated using dropout, pooling layers, and sequence augmentation, though the study lacked external test data beyond the same experimental campaign. The results showed that deep learning models provided accurate and efficient alternatives to computationally expensive non-linear FE analyses simulations. Mohammadi et al. [217] investigated ANN-based models to estimate seismic demands of BRBs subjected to pulse-like ground motions. Using over hundreds of nonlinear dynamic analyses, the study trained ANNs to predict maximum inter-storey and global drift ratios as key seismic demand parameters. Validation relied on an 80/20 train–test split, with no k-fold CV. The best ANN models achieved peak R^2 values of 0.96 for maximum inter-storey drift ratio and 0.95 for global drift ratio in training, with corresponding test values of 0.94 and 0.93. Although no uncertainty intervals were provided, the performance metrics (R^2, RMSE, and MAE) indicated strong predictive accuracy. Overfitting was reduced by optimizing ANN architecture and testing on separate holdout datasets, but lack of external data remained a limitation.

For dampers, a recent study proposed a fast-forward approach to analyze seismic vulnerability through FE analysis, structural design software, and ANN, with fluid viscous dampers in varied locations of a building frame [119]. In another study, two crucial properties of a steel plate damper, stiffness and slenderness factor, were predicted using response surface methodology (RSM), ANN, and evolutionary polynomial regression (EPR). The study considered elastic-inelastic-plastic buckling modes and flexural, shear, and flexural–shear failure mechanism of concentrically braced frames, with 33 geometric property entries. EPR showed the best performance, with R^2 of 0.999 for slenderness and 1 for stiffness. Validation was conducted by splitting dataset into training and validation sets with multiple error metrics, while uncertainty was expressed through descriptive statistics and error values [120]. Onyelowe et al. [218] presented a hybrid framework combining response surface methodology and ML models to predict the seismic performance of steel plate dampers in concentrically braced frames. Input features included geometric and material properties of dampers. The study used a train–test validation approach, where ANN outperformed response surface methodology and other ML methods, achieving R^2 values up to 0.99 with low RMSE and MAE. Although uncertainty intervals were not explicitly reported, results highlighted ANN's superior predictive ability. Overfitting was addressed by benchmarking different methods, though dataset size and lack of broad external testing limited generalization. Chen and Chien [219] trained MLP and auto-regressive model with exogenous controllers for seismic response control and validated their performance through both numerical simulations and shake-table experiments on a single-degree-of-freedom specimen. Validation combined offline train/validation splits from excitation data and real-time experimental closed-loop tests. The study reported objective metrics averaged across records (objective functions, RMSE), but it did not report R^2 for accuracy and SD/CIs for uncertainty. Experimental validation was a strong mitigation against overfitting; auto-regressive model with exogenous controllers performed faster and with similar accuracy to MLP in tests. Shao and Andrawes [220] trained ANNs to predict the maximum displacement of a single-degree-of-freedom reinforced concrete structure with super-elastic dampers using a large simulated dataset generated in OpenSees, consisting of approximately 109,000 samples derived from 200 ground motions. The validation method was a hold-out split (70% training, 15% validation, and 15% testing), and generalization was further tested on separate ground motions whose parameter values differed from those in the training set. Reported uncertainty was given in terms of RMSE and average error (best performance: RMSE $\approx$ 0.1012 and average error $\approx$ 6.55 % for the 200-ground motion case). The most influential parameters were spectral acceleration and peak ground acceleration. Hu et al. [221] built an explainable probabilistic buckling-stress predictor by training ML models (ANN performing best: RMSE $\approx$ 0.0094 and $R^2 \approx 0.9952$) on an FE-generated database of $\sim$32, 400 cases. They validated FE against experiments and used Latin hypercube sampling to propagate input uncertainties. They also produced probability densities and global sensitivity analysis indices which showed yield-stress and initial-imperfection to be the dominant uncertainty drivers. Validation relied on held-out testing and distributional comparisons; ensembles (RF and XGBoost) and the large database helped reduce overfitting risk. Bae et al. [122] investigated a double-coke damper with multiple strips based on a modified radius-cut section. In this configuration, increasing numbers of plastic hinges on a single strip increased the ductility of the entire damper, producing a stable hysteresis diagram. Computations based on the proposed equation (for damage index determination using parameters such as maximum strength and effective stiffness), experimental results, and ML-derived predictions were found to be in close agreement. The fatigue performance of the damper was assessed through a constant cyclic loading test on a specimen. The analyses revealed a stable load–displacement hysteresis graph, a shear resistance exceeding the theoretical value, and an increase in ductility or fractural strength. A low-cycle fatigue model was developed using a linear regression algorithm based on ML to estimate the damage index. The damage point was estimated based on the maximum strain and effective stiffness variation. The number of periodic failures was found to be in excellent agreement with the experimental results. The model achieved over 0.90 accuracy and RMSE of 0.1 over six different specimens in predicting damage points compared with test data.

3.3. Environmental and economic impact analysis

Conventional LCA and lifecycle cost (LCC) analysis are the two primary approaches for evaluating the environmental and economic feasibility of building construction. However, these methods often rely on assumptions—such as a building lifespan of >50 years and the exclusion of maintenance costs—that may result in inaccuracies in practical applications. ML techniques offer a more reliable alternative by predicting effective lifespans and estimating costs while accounting for variability in environmental factors (e.g., material manufacturing, transportation, construction, operation and maintenance, demolition, and waste disposal) and economic conditions (e.g., operational and maintenance costs). Ji and Yi [123] collected 1812,700 records related to construction and demolition processes to analyze the lifespan of buildings using modern prediction models, including linear regression, XGBoost, LGBoost, and DNN. For the study area, the average lifespans of RC-structured and brick-structured buildings were found to be 22.8- and 29.3-y, respectively, significantly lower than the assumed span of 50-y. The DNN model achieved the highest accuracy ($R^2 = 0.955$). Onyelowe et al. [68] extensively explored the mix design of fly ash-incorporated concrete using statistical analysis, linear regression, and AI to predict the environmental impact. The database included 112 concrete samples, with three input variables: fly ash-to-binder ratio, FA (sand)-to-binder ratio, and CA-to-binder ratio. ANN achieved the best accuracy ($R^2 = 0.991$), identifying the aggregate-binder ratio as the most influential parameter. An increase in both the fly ash-to-binder ratio and the aggregate-to-binder ratio was found to reduce the carbon footprint. Validation was conducted using a holdout split (90/22– training/validation); no k-fold or nested CV was reported. Uncertainty was presented via SDs of inputs/outputs and residual diagnostics — the study reported residual SD bands and average error percentages; explicit CIs were not provided. Koyamparambath et al. [204] processed 980 datapoints (784 for training) from environmental production declarations data to predict environmental impacts for construction products with 7 vital information (e.g., name/description, location, 3 classification levels, functional unit, and values of selected impacts category). The study employed RF to predict environmental impacts such as photochemical ozone creation potential ($R^2 = 0.70$), abiotic depletion potential for fossil resources ($R^2 = 0.77$), global warming potential ($R^2 = 0.81$), and acidification potential ($R^2 = 0.68$). Sharif et al. [205] developed surrogate models to predict energy consumption using ANN with 463 renovation scenarios (325 training and 138 testing datasets) generated from simulation-based multi-objective optimization. The models achieved strong predictive accuracy with lower error rates (MSE from 0.016 to 0.124), confirming their reliability in forecasting total energy consumption, LCC, and LCA. Another recent study from Baehr et al. [206] predicted life cycle environmental impacts filtering 5251 datasets (60% for training, 20% for validation, and 20% for testing) integrating ML methods (ANN, residual GPR, and ANN-residual GPR). ANN-GPR hybrid models produced most accurate results ($R^2 = 0.95$) with input parameters (e.g., environmental production declarations' attributes, product class, functional unit/reference flow, embodied fossil/renewable energy, and recycled contents). Askarinejad and Behnia [207] implemented ML algorithms (DT, polynomial regression, SVR, and elastic-net) as early design tools using several high-rise buildings (varying heights up to 100 floors), four different types of construction materials (concrete, steel, hybrid and timber) and concrete with varied strength (32 to 90 MPa). DT outperformed other models with an accuracy of 0.99 (MAE of 13 and MSE of 452). In this study, validation was carried out via hold-out splits, and uncertainty was not quantified beyond overfitting concerns.

3.4. Summary on comparative structural analysis with AI models

AI techniques, such as ANN, GEP, XGBoost, GB, RF, BR, CNN, and emerging methods like MEP, SVR, and LGBoost have demonstrated high efficiency and robustness in predicting concrete mix design–driven mechanical properties, durability, structural seismic response, and fire-induced effects. As shown in Table 1, NN and their optimized variants are among the most widely applied models across these domains because they can capture highly complex nonlinear structural behavior, predict strength and durability properties with high accuracy, and enable rapid post-earthquake assessments that improve structural safety

Table 1
Accuracy of prominently used AI-algorithms in structural engineering based on the coefficient of determination (R^2).

AI Applications in the determination of Properties		Section 2.2	Section 2.3	Section 2.5	Section 2.7		Section 2.8	References
		NN: ANN, RNN, KNN	**DL: DNN, CNN**	**GA: GP, GEP, MEP**	**Tree-based Algorithms**	**Boosting Methods**	**SVM, SVR**	
Section 3.1.1 **Concrete Strength Properties**	Mix Design-Compressive Strength	0.92	0.89	0.96	0.97	0.95		[63,66–68,131]
	Tensile Strength	0.95				0.84–0.98	0.98	[56,74–76]
	Flexural Strength	0.98		0.81–0.98	0.99			[71,72,132,133]
	Uniaxial-Triaxial Compression			0.99		0.92–0.97		[17,47,134]
	Self-healing Ability			0.94	0.9–0.97	0.96		[20,79–81]
Section 3.1.2 **Durability of Concrete**	Moisture Exposure	0.95						[82]
	Corrosion of Concrete			0.98				[83]
	Chloride Resistance	0.85			0.94*	0.96	0.89	[26,84,135,136]
	Depth of Wear	0.99		0.97–0.99	0.99			[50,137–139]
	Frost Durability/Resistance	0.96			0.95	0.96	0.98	[19,85,140–142]
	Impermeability	0.93*			0.95	0.97–0.99*	0.97	[86,143,144]
	Carbonation Penetration	0.98–0.99*		0.88		0.98		[87,145–147]
Section 3.1.2 **Fire-Induced Effect**	Thermal Spalling of FRP-based Concrete	1.00	0.91		0.90	0.90–0.97		[29,51,89,90]
	Fire-induced Spalling of Reinforced Concrete Member	0.99	0.86	0.94		0.96		[21,52,148–150]
	Buckling & Thermal Spalling of Composite Member	0.97–0.99		0.82		0.91–0.99		[91,93,97,151,152]
	Buckling & Progressive Collapse of Steel Frame	1.00	0.96	0.90			0.99	[58,60,153,154]
	Concrete-Steel Bond Strength				0.95	0.92–97		[98,155,156]
Section 3.2.1 **Mechanical Properties of Sections**	Shear Resistance of RC Member	0.89		0.89	0.95		0.95*	[99,157–159]
	Capacity of Masonry and RC Wall	0.95–0.99			0.8–0.94	0.97		[37,100,160–164]
	Safety Criteria of Retaining Wall	0.97				0.99	1.00	[101,165,166]
	Compression on Composite Column	0.80–0.99				0.98–0.99*		[103,104,167–169]
	Shear Capacity of Composite Beam	0.93–0.99		0.95				[105,170]
	Shear Capacity of Composite Slab	0.89				0.96–0.99	0.96	[172–175]
	Buckling of Steel Beam	0.93–0.99		0.97			0.99	[106–108,176]
	Buckling of Steel Column	0.93*-0.99						[41,[171],177,178]
	Behavior of Column-Beam Joint	0.99	1.00		0.87	0.91		[36,62,179,180]
Section 3.2.2 **Seismic Response**	Load-Deflection Response of RC Frame	0.87–0.99	0.92		0.8–0.97	0.98	0.98	[111–113,181,182]
	Deformation-based Fragility of Steel Frame	0.96–0.98			0.99	0.96–0.99	0.96–0.98*	[116,118,178,183–185]
	Nonlinear Hysteretic Behavior of Retrofitting Systems	0.94–0.95			0.96	0.99		[27,120,186,187]
Section 3.3 **Analysis**		0.97	0.96					[68,123]

* Optimized or hybridized versions of an AI method.
Note: Peak accuracy values are considered screening R^2 values between 0.80 to 1.00 from Section 3.

[124]. They also learn directly from experimental data, offer computational efficiency by generalizing new fire scenarios and structural configurations with diverse training datasets [125], and, in their optimized versions, provide quantifiable and transparent insights that improve the reliability of predictions [126]. In addition to NNs, boosting methods are also widely adopted because of their ability to sequentially correct errors from weaker models. These methods effectively capture complex relationships between seismic parameters and structural response, efficiently handle data variability, reduce errors in predicting damage states compared to standalone models, avoid bias to generate more stable and generalizable predictions, and improve precision and recall in classification tasks [127]. For durability-related studies, tree-based algorithms have been particularly effective. In the context of seismic response, they can manage complex nonlinear interactions among structural parameters, provide interpretable results through visualizations, and process large input datasets with relatively low computational demand [128]. For durability analyses, they can capture interactions between environmental conditions and material properties, identify the most critical factors contributing to concrete deterioration, and handle diverse datasets with limited sensitivity to outliers [129]. However, they remain underutilized for evaluating the mechanical properties of structural members. GA and its extensions (GP and GEP) have been more commonly applied to predicting the strength properties of concrete. GA offers a simple yet robust encoding process, GP improves interpretability through flexible expression trees, and GEP combines the strengths of both by capturing complex nonlinear relationships while maintaining

interpretable formulations that closely align with experimental data [130]. Among the AI methods in Table 1, deep learning–based algorithms remain less explored across most areas of structural engineering, although they show considerable promise in applications such as carbon footprint estimation and economic assessment. SVM and its variants (e.g., SVR) have also demonstrated consistent robustness and accuracy across different applications. Table 2 presents the comparative performance of these AI techniques along with their field-specific advantages and limitations. On average, accuracy values of about 0.80 were observed for NN, GA, tree-based and boosting algorithms, while optimized or hybridized versions of AI algorithms achieved notably higher prediction accuracy, averaging around 0.90.

As a continuation of Tables 1 and 2, Table 3 presents comparative metadata with an overview of dataset size, features, validation methods, metrics, and external test indications for each AI integrated field. From the reviewed studies, the most used feature sets included material composition and mix proportions (cement, water, aggregates, admixtures, and age) for concrete strength and durability, extended by geometric and loading parameters in fire-induced and mechanical property analyses, while seismic response models predominantly relied on natural periods and structural descriptors. In terms of validation methods, the dominant approaches were holdout splits (typically training/testing of 70/30, 75/25, or 80/20) and 10-fold CV. The most frequently reported metrics were R^2, RMSE, and MAE, with occasional use of MAPE, and classification indices (basic introduction in Section 2.9.1). For external test indication, most studies validated models only on internal datasets, with comparatively fewer works employing independent experimental databases or literature-based test comparisons for external validation. According to field-based data shown in Tables 3, 4 further presents commonly used software packages and libraries in the reviewed AI integrated fields. Across the reviewed studies, MATLAB, Python, and Scikit-learn are found to be the most frequently used tools, often serving as core environments for model development and data analysis. TensorFlow/Keras, XGBoost, and SHAP are also common for deep learning, boosting, and explainability tasks, while specialized tools like OpenSees, GeneXproTools, and EPR Toolkits appear in domain-specific applications.

Table 2
Summary of different AI methods across operations comparing performances, advantages, and limitations.

Sections	Reviewed AI Techniques	Performance based on the coefficient of determination (R^2)	Advantages	Limitations
2.2	NN: ANN, RNN, KNN	Accuracy > 0.80 across all study fields	• Models highly complex nonlinear structural behavior • Learns directly from experimental data. • Computationally efficient in predicting new fire scenarios and structural configurations using diverse datasets • Optimized ANN provides quantifiable and transparent insights, improving reliability	• Struggles to generalize across varied conditions, particularly for concrete durability and seismic response
2.3	DL: DNN, CNN	Accuracy > 0.85 for mix design, fire-induced effects, beam–column joints, seismic response of RC frames, and lifecycle analysis	• Captures complex nonlinear relationships, enabling accurate prediction of fire-induced effects, seismic response, and beam–column joint behavior • Supports life-cycle analysis by identifying hidden patterns in RC frame performance and durability.	• Sensitive to noise and computationally intensive, limiting practical deployment in structural assessment
2.5	GA: GP, GEP, MEP	R^2 = 0.81–0.99 (concrete strength), 0.97–0.992 (durability), 0.823–0.94 (fire-induced effects), 0.89–0.97 (shear resistance and buckling of RC members)	• Widely used for predicting strength properties of concrete. • GA offers a simple yet robust encoding process • GP improves GA with interpretable, flexible expression trees • GEP combines GA and GP, capturing nonlinear relationships with interpretable formulas closely matching experimental data	• Less commonly applied to seismic response and concrete durability
2.7	Tree-based Algorithms	R^2 > 0.90 for concrete strength; > 0.94 for durability; 0.87–0.99 for seismic response	• Manages intricate non-linear interactions, provides transparent visualization, and processes wide input data influencing seismic performance. • Handles large and diverse input datasets for seismic performance, identifies key factors influencing concrete durability while remaining robust to outliers	• Underexplored for evaluating the mechanical properties of structural members
	Boosting Methods	Accuracy > 0.91 across most domains	• Sequentially corrects errors from weaker models. • Captures complex relationships between seismic parameters and structural response. • Handles variability efficiently while reducing errors in damage prediction • Avoids bias, improving generalization and stability • Enhances precision and recall in classification tasks	• High computational cost and sensitivity to input features, particularly for predicting mechanical, fire-induced, and durability-related concrete responses
2.8	SVM, SVR	Specifically, better results in determining concrete tensile strength: R^2 = 0.98 (concrete tensile strength), 0.89–0.98 (durability), 0.95–0.96 (RC shear resistance), 0.99 (steel beam buckling), 0.98 (RC seismic response), 0.96–0.98 (steel seismic response)	• Demonstrates high accuracy and robustness across multiple areas	• Limited application to concrete properties, fire effects, and lifecycle analysis due to challenges in handling complex nonlinearities and interpretability issues

Table 3
Comparative metadata (field-based tasks, data size, features, validation methods, metrics, external test validation) across reviewed AI integrated fields.

Fields	Tasks	Dataset Size	Features	Validation Methods	Metrics	External Test Indication	References
Section 3.1.1 **Concrete Strength Properties**	Mix Design-Compressive Strength	432	6 (cement, FA, CA, water, superplasticizer, SF)	*NS*	MAE, RMSE, R^2	No external dataset	Nafees et al. [66]
	Tensile Strength	*NS*	8–9 (cement, slag, fly ash, water, superplasticizer, CA, FA, age)	Hyperparameter tuning (random search), no explicit CV	RMSE, R^2	No external dataset	Nguyen et al. [56]
	Flexural Strength	200	~5–7 (beam width, depth, reinforcement ratio, FRP parameters)	Holdout (train/validation split)	R^2, MAE	No external dataset	Khan et al. [46]
	Uniaxial-Triaxial Compression	1298	6 (crushing strength, height–width ratio, shape, Pearson correlation coefficient, stress ratio, loading frequency)	Data cleaning/averaging, no explicit CV	R^2 (0.75–0.915)	No external dataset (literature-based)	Son & Yang [47]
	Self-healing Ability	797	22 (e.g., bacteria type, healing environment, cement type, crack width, healing time)	10-fold CV + grid search	R^2 (0.956), RMSE	No external dataset	Huang et al. [79]
Section 3.1.2 **Durability of Concrete**	Moisture Exposure	429	8–10 inputs (Geometric, mechanical, and environmental)	Train/test split (regression), Stratified 10-fold CV (classification)	R^2, MSE, RMSE, MAE; classification metrics	Yes – held-out test set (no independent external dataset)	Baghaei and Hadigheh [82]
	Corrosion of Concrete	256	Chemical: time, pH; Biological: time	50/50 split (train/test)	MSE, R^2	No external dataset (internal only)	Sabour et al. [83]
	Chloride Resistance	30+/study	Water/cement, thickness, aggregate fraction, temperature/humidity ratios, exposure time ratio	Train/validation; scikit-learn DT defaults; k-fold mentioned	Accuracy %, RMSE, errors	Yes – compared with external test results from literature	XuanRui et al. [84]
	Depth of Wear	216	Cement, fly ash, water, aggregates, plasticizer, age/time, curing/test	Train/validation split + external validation framework	R^2, MAE, RMSE	Yes – external validation criteria applied	Khan et al. [50]
	Frost Durability/Resistance	94	10 inputs (cement, water, sand, natural and recycled CA, fly ash, recycled CA replacement ratio, water absorption, RCA treatment method, air-entraining type)	75/25 train/test split	R^2, RMSE, MAE	No (train/test only)	Esmaeili & Sarkhani [141]
	Impermeability	417	≈10 (water, water–cement ratio, cement, aggregates, rubber size, cycles)	347 train / 70 test split	R^2, RMSE, MAE	No independent dataset (literature split only)	Huang et al. [86]
	Carbonation Penetration	532	6 inputs (cement, FA, water–binder ratio, CO_2 %, relative humidity, exposure time)	70/15/15 split + 10-fold CV	MAE, RMSE, R^2	Yes – held-out test set and CV folds	Kazemi [87]
Section 3.1.2 **Fire–Induced Effect**	Thermal Spalling of FRP-based Concrete	531	17 inputs (e.g., mix proportions, moisture, specimen size, temperature, heating rate, fibers, silica fume)	K-fold CV + supplementary test	Accuracy, F1, precision/recall	Yes – 36 experimental tests + expanded dataset	Liu et al. [89]
	Fire-induced Spalling of Reinforced Concrete Member	100+	Concrete material and mix proportions, geometric and configuration/size features, and those relating to applied loading, intensity, heating rate, and exposure duration	Database validation vs test series (not k-fold)	Comparisons with experimental outcomes	Yes – multiple independent fire test campaigns	Naser & Seitllari [21]
	Buckling & Thermal Spalling of Composite Member	15,200	7 inputs (cross-sectional dimensions, thicknesses of concrete cover for steel section and rebars, steel area ratio, effective length, concrete grade, steel grade, and heating time)	80/20 train-test split	R^2, MAE, SD vs. analytical eqns	15,200 specimens (synthetic FD model)	Li et al. [93]
	Buckling & Progressive Collapse of Steel Frame	*NS*	Thermal/mechanical variables (fire temperature, maximum steel temperature, load Ratio, critical temperature based on the Eurocode)	80/20 train-test split	Accuracy, classifier comparison	Monte Carlo + random sampling (case study 2 × 2 building)	Fu [60]
	Concrete-Steel Bond Strength	316	7 inputs (compressive strength under elevated temperature, testing age,	Train-test split (no explicit k-fold)	MAE, RMSE, R^2, CoV	316 experimental tests	Al Hamd & Warren [98]

(*continued on next page*)

Table 3 (*continued*)

Fields	Tasks	Dataset Size	Features	Validation Methods	Metrics	External Test Indication	References
			surface temperature at failure, thermal saturation ratio Δ, length–diameter ratio, cover–diameter ratio, total volume of fiber if used)				
Section 3.2.1 **Mechanical Properties of Sections**	Shear Resistance of RC Member	466	Geometrical and mechanical variables (beam/stirrup properties)	Train/validation within dataset, sensitivity analysis	Error metrics vs. codes	Validated against experimental database	Gandomi et al. [99]
	Capacity of Masonry and RC Wall	*NS*	Masonry & FRM mechanical/ geometrical variables	Trained on lab results (no CV specified)	Accuracy/ precision qualitatively	Yes – calibrated vs lab tests	Cascardi et al. [100]
	Safety Criteria of Retaining Wall	*NS*	Cohesion, angle of shearing resistance, angle of wall friction, and unit weight	10-fold CV	Reliability index (first-order second moment method)	Compared to reference reliability values	Mishra et al. [101]
	Compression on Composite Column	*NS*	Geometrical and material properties	ANN training/ validation (no CV info in snippet)	Comparative performance (R^2, RMSE likely)	Yes – experimental DB	Lemonis et al. [104]
	Shear Capacity of Composite Beam	*NS*	6+ geometrical/interaction vars (e.g., opening diameter, web opening spacing, tee-section height, concrete topping thickness, interaction degree, number of shear studs above web opening)	Comparative training, reliability analysis	Performance metrics, safety factors (1.25–1.26)	FE-based, compared with literature	Ferreira et al. [105]
	Shear Capacity of Composite Slab	273	≈8–10 (e.g., slab depth, slab side length, flexural reinforcement ratio, FRP type/properties, concrete compressive strength, loading type)	10-fold CV	R^2, RMSE, MAE	External: dataset compiled from multiple independent experiments	Shen and Liang [172]
	Buckling of Steel Beam	9744	≈10 geometric features (e.g., height, web thickness, opening height/width/ radius)	Holdout (train/ validation/test splits)	R^2, RMSE, MAE, SD/Variation, a20-index	No external experiments (FE model vs ANN)	Shamass et al. [177]
	Buckling of Steel Column	10,764	Multiple geometric & material grades (e.g., web thickness, web-post width, opening height, steel grade)	10-fold CV and train/validation/test splits	R^2, RMSE, MAE, SD, CoV	FE model vs Euro-Code 3 (no experimental set)	Rabi et al. [108]
	Behavior of Column-Beam Joint	387	11 (cross-section dimensions (top/bottom flange widths & thicknesses, max/min section heights, web thickness), elastic modulus, column height and corrosion time for corroded cases)	Holdout (train/ validation/test)	R^2, RMSE	No external experiments (analytical dataset)	Nguyen et al. [179]
Section 3.2.2 **Seismic Response**	Load-Deflection Response of RC Frame	300	First three natural periods (T1-T3) and combinations thereof; natural periods derived from generated frames	Holdout (varying train/test sizes)	MAPE, NRMSE, R^2	Synthetic (OpenSees) – no external experiments	Gharehbaghi et al. [111]
	Deformation-based Fragility of Steel Frame	56,479	First three natural periods and other structural descriptors; inputs chosen to build Probabilistic Seismic Demand Models (PSDMs)	Holdout 70/30	R^2, RMSE, MAPE	Large synthetic dataset built from extensive nonlinear analyses (616 frames × 240 motions)	Nguyen et al. [118]
	Nonlinear Hysteretic Behavior of Retrofitting Systems	33	4 geometric properties of damper (e.g., plate thickness, plate dimensions, and the number of plates used.)	Holdout (train/val splits)	R^2, RMSE	No external validation	Onyelowe et al. [120]
Section 3.3 **Life Cycle Carbon Assessment**		*NS*	Encoded text/categorical environmental production declaration features	Holdout (80/20)	R^2	20 % environmental production declaration as external holdout	Koyamparambath et al. [204]

Note: *NS* refers to *not specified value* in a certain study.

4. Challenges and future directions in AI methods

Building on the future opportunities for industrial integration of AI, the prediction accuracy of AI models at the laboratory scale can be observed from the previous section and the demonstrations in Tables 1 and 2. Extensive research has applied AI methods to evaluate the properties of concrete specimens as well as RC, composite, and steel structural members and frames. However, only limited studies have addressed areas such as self-healing ability and concrete–steel bonding under severe fire conditions. Similarly, seismic response analysis of steel

Table 4
Software packages and libraries commonly used in the reviewed studies.

AI Integrated Reviewed Study Fields	Software Packages / Libraries Mentioned
Section 3.1.1 | Concrete Strength Properties	MATLAB [66,46], Python [66,46,47], sci-kit learn [56,47,79], XGBoost [56], SHAP for explainability [46], ML ensemble methods [47], DNN frameworks (TensorFlow/Keras/Pytorch) [79], Grid Search Algorithm (GSA) [79]
Section 3.1.2 | Durability of Concrete	MATLAB [82], Python [82], sci-kit learn [50,84], SHAP for explainability [50], Bayesian Optimisation [82], custom C++ for GP/MEP [83], decision tree defaults [84], custom SVR with metaheuristic implementations [86,141], Alyuda NeuroIntelligence [87], hybrid ANN [87]
Section 3.1.2 | Fire–Induced Effect	MATLAB [93], Python [60,89], sci-kit learn [89,98], XGBoost [89], TensorFlow [60,98], Keras [60]
Section 3.2.1 | Mechanical Properties of Sections	MATLAB [172,177], Python, sci-kit learn [108], GeneXproTools [108]
Section 3.2.2 | Seismic Response	Scikit-learn [118], FE Modeling [111], OpenSees [111], EPR Toolkits [120]
Section 3.3 | Life Cycle Carbon Assessment	Python (Selenium, SQLite) [204]

frames with bracings, shear capacity of RC members, and failure modes of column–beam connections have received comparatively less attention than other AI applications. Fire-induced effects on BRBs, column–beam connections, and LCA also remain underexplored. Nevertheless, significant potential exists for optimized and hybridized variants of widely used algorithms, such as NN and GA-based, tree-based models, and boosting techniques. Sections 4.1, 4.2, and 4.3 discuss the industrial applications of AI, the limitations observed in laboratory scale studies, and future recommendations for AI integration in structural engineering.

4.1. Industrial implications, barriers to adoption and potential for implementations

For construction projects that rely on accurately evaluating environmental factors (e.g., seismic events), risks, and costs, AI offers significant practical advantages [197]. ML is increasingly applied in big data analytics for risk detection and assessment, and ML models are also used in robotics and automation. For instance, aerial drones and robotic vehicles are frequently deployed on survey sites to generate 3D models of building structures. AI algorithms further support on-site problem identification and provide strategic solutions that enhance efficiency. In construction automation, AI is also applied to improve workers safety through smart wearable technologies that monitor movement, activities, and posture, helping to prevent collisions between workers and heavy equipment [198]. Despite these benefits, several barriers hinder AI adoption in the construction industry. These include the fragmented nature of the industry, challenging environmental conditions on-site, and non-standardized building designs, all of which complicate data collection, integration, and standardization [199]. Additional constraints include limited technical skills, inappropriate business processes, and insufficient knowledge, making AI adoption time-consuming, costly, and prone to errors [200,201]. Moreover, many large firms continue to rely on traditional processes rather than automation, and subcontractors often follow the same outdated practices [203]. On a positive side, the construction industry has been investing heavily in AI, with an estimated USD 8 billion allocated in the five years leading up to 2019 [202]. This investment paves the way for AI-enabled technologies such as digital twins and 3D printing, which can significantly reduce repetitive and labor-intensive tasks. Looking ahead, future AI integration should also target innovative fields such as fire-induced effects, seismic impact analysis, and LCA. This study has reviewed AI-based findings in these areas using laboratory scale experimental and numerical databases; the associated limitations and prospects are discussed in the following sections.

4.2. Limitations in AI applications in structural engineering

Limited access to diverse and representative datasets, high costs of data collection, and data scarcity due to legal restrictions often result in inadequate data availability. Missing data, model bias, data drift, and errors further affect the reliability of AI predictions. These challenges, arising from limited datasets, difficulties in maintaining data quality, and research gaps, can be summarized as follows:

1. In studies on the mix design of sustainable concrete, only limited scale experimental datasets have been used in recent research. The generalization performance of AI models for sustainable concrete preparation can only be suggested from these limited studies, and future work should involve extensive tests on various SCMs used in concrete. Additionally, some studies have reported missing information (e.g., inappropriate or incomplete input variables) from experimental data which lessens the prediction accuracy and reliability.
2. For predicting the strength and durability of concrete materials, variations are observed in the optimal algorithms, such as XGBoost, GEP, BR, ANN, GB, MEP, SVM, and RF. The selection of a specific AI method is often subjective and depends on researcher's expertise. In some cases, even the most preferred methods fail to outperform the existing design codes. Apart from reliance of quality datasets, another issue is the time-consuming process of parameter tuning.
3. AI-based predictions generally require large volumes of experimental data to ensure accuracy and precision. However, data availability for specific problems is often limited owing to laboratory constraints (e.g., fire testing facilities for fire resistance analysis). In such cases, AI models may suffer from numerical complexities, including overfitting training data without comparable practical test data for validation.
4. Notably, most reviewed studies compiled their datasets from prior works conducted in different regions. However, these prior studies often differ significantly in environmental conditions, material characteristics, and experimental setups, which may limit the reliability and generalizability of AI models.

A few studies have incorporated random sampling and SHAP analysis, which are important for selecting appropriate data and providing detailed explanations of model accuracy. Some researchers have also resorted to using synthetic data to train and validate ML models due to the lack of real-world data. However, this approach can yield unrealistic results. Emerging structural concepts, such as carbon-neutral and easy-to-dismantle beam–column joints, offer promising solutions to reduce the carbon footprint across the structural lifecycle. Yet, no AI methods have been developed to evaluate their fatigue performance, load-carrying capacity, or associated carbon footprint.

4.3. Recommendations for future research in structural engineering

This section addresses the previously discussed limitations related to data availability and quality, while also providing insights into potential solutions and highlighting new research opportunities:

1. Extensive research has been conducted on AI-based sustainable concrete mix design using SCMs, byproducts, and waste materials. However, further studies are needed to assess the applicability of single AI-based mix design approaches across different concrete types. The use of locally available experimental data is recommended to obtain more accurate predictions for optimum mix proportions of concrete materials. ML approaches combined with heuristic methods, such as PSO, can further enhance prediction accuracy in mix design.

2. The inertia in selecting AI techniques for similar problems can be reduced through collaboration among researchers to identify the most suitable methods for specific, concrete-related challenges. Issues related to parameter tuning can be addressed by incorporating optimization algorithms such as GA, which can streamline the tuning process.
3. To address numerical complexities caused by limited real-world data, AI models should be regularly updated, supported by open-access databases that facilitate information sharing. More diverse data points and comprehensive experimental datasets are required to capture a wide range of scenarios. Training models with homogeneous data obtained under specific environmental and material conditions can improve prediction accuracy. Additionally, random sampling can be employed to refine results after training. Raw experimental data should be prioritized over synthetic data to ensure practical relevance. Normalization of input variables into uniform ranges can also minimize dataset bias and enhance model performance.
4. Existing models for fiber-reinforced concrete elements should be extended to account for the confinement effects of different fiber-reinforced polymer configurations. For instance, incorporating input variables such as the placement and orientation of polymer wrappings could improve the evaluation of structural integrity and failure potential under fire exposure.

In this study, adaptive explainable AI methods (e.g., SHAP and local explainable model-agnostic explanations), are not discussed broadly due to limited existing research. Future studies should explore more of these techniques. Some studies have used a single AI technique for multiple test specimens with different criteria. Further studies can compare different AI methods to identify the simplest and most accurate models for a specific problem. At present, most construction projects utilize structural steel components, and future studies should focus on AI-based analysis of the LCA and LCC of steel structures. Additionally, a promising innovative research direction can be represented with AI-based predictions for evaluating the self-healing property, post-fire conditions and fire-induced effects of concrete structures, and seismic response analyses of steel building frames. Easy-to-dismantle beam-column multiple connections are crucial for reducing construction time and labor costs, as they can minimize the requirement of bolts and rivets at working sites. These prefabricated joints also contribute to lower carbon emissions in the lifecycle of structures. AI-based research can be suggested for such connections. Moreover, experimental, theoretical, numerical, and prediction-based analysis must be conducted, and the results should be compared to determine effective and efficient designs. Moreover, different AI methods can be used to determine the fire effects, shear capacity, and failure modes of steel beam-column connections and BRBs.

5. Conclusions

AI has demonstrated exceptional accuracy in structural engineering research, producing predictions that are comparable to, and in many cases superior to, experimental tests, numerical simulations, and design codes. In recent years (2020–2024), ANN, boosting methods, tree-based algorithms and SVM models have been widely adopted for their strong predictive capabilities, while MEP, SVR, BR, LGBoost, and deep learning models have also proven robust and reliable for capturing complex structural behavior. For instance, NN and boosting methods in particular exhibit high predictive accuracy ($R^2 > 0.80$ and $R^2 > 0.90$, respectively) across diverse applications, making them the most widely applied approaches. Deep learning methods are particularly effective in mix design, strength prediction, fire-induced effects, beam–column joints, seismic response of RC frames, and LCA, achieving accuracy levels above 0.85. Meanwhile, GA, SVM, SVR, and tree-based models have shown strong performance in specialized tasks, including concrete durability, fire resistance, shear behavior, and seismic buckling, with reported accuracy ranging from 0.80 to 0.99. Based on studies published between 2020 and 2024, the main findings can be summarized as follows:

1. **Concrete mix design and strength prediction:** Tree-based algorithms are prominent, with XGBoost, ANN, GEP, GB, BR, ensemble DT, and stacking methods achieving R^2 values between 0.91 and 0.99. Key influencing parameters include curing age, cement content, recycled aggregate replacement ratio, SCM-to-binder ratio, aggregate-to-binder ratio, and specimen dimensions. Integration of AI is highly significant for accurately predicting mixing parameters that influence the strength properties of concrete.
2. **Durability prediction**: ANN demonstrates strong accuracy ($R^2 =$ 0.85–0.99) for most durability aspects. Specific properties such as corrosion resistance, chloride permeability, depth of wear, frost resistance, impermeability, and carbonation depth are best predicted by MEP, XGBoost, BR, SVR, and optimized ANN models. Governing parameters include exposure duration, aggregate fractions, cement–SCM ratios, water–binder ratios, FA, protective layer thickness, and environmental conditions. According to these findings, AI prediction provides reliable insights into concretes durability while highlighting the critical role of mix parameters for long-term performance.
3. **Fire-induced effects:** Neural networks and boosting algorithms accurately predict spalling in RC, composite, and steel structures (R^2 = 0.90–1.00), as well as concrete–steel bond strength (accuracy of 0.97). Critical parameters include member geometry, applied load and load ratios, thermal properties, fire insulation depth, and reinforcement area. These findings show that AI-based approaches have the capacity to effectively capture the critical factors that influence fire-induced spalling, providing a strong framework for predictive assessment and design.
4. **Structural behavior and failure modes:** In determining mechanical properties of steel structural beam-column joints, only a few studies have implemented CNN. Apart from this, different algorithms show peak accuracy across different fields such as GEP, SVR and tree algorithms for shear resistance of RC members, ANN and RF for shear strength of masonry and RC walls, SVM and boosting methods for safety criteria of retaining walls, optimized ANN and boosting methods for ultimate load capacity of composite beams and columns, and ANN-SVR for buckling failure modes of steel beam-column. Across these applications, the most influential parameters include strength properties, section geometry, and member aspect ratios.
5. **Seismic response:** These have been accurately predicted using classification-based, hybrid, and optimized ANN variants for RC frames (with accuracy above 0.87), and hybridized boosting methods for steel frames (with accuracy above 0.96). Key factors include axial load, concrete strength, reinforcement dimensions, building characteristics, applied loads, and site conditions. AI integration in seismic response analysis is particularly significant, as it supports the development of innovative designs incorporating BRBs, viscoelastic dampers, beam–column joints, and advanced retrofitting technologies.
6. **LCA and economic analysis:** LCA of carbon emissions and lifecycle cost analysis for optimized building construction solutions can be further enhanced using high-performing AI methods such as ANN (accuracy above 0.97) and DNN (accuracy above 0.96). By integrating AI into LCA, it becomes possible to identify complex interdependencies among materials, energy use, and costs, enabling more precise and strategic sustainability decisions in construction.

With continued advancements, AI-based predictions have a strong potential to be integrated into updated structural design codes, provided results are rigorously validated through experimental and real-world applications.

Abbreviations

AI, artificial intelligence; CV, cross-validation; ANN, artificial neural network; PR, pattern recognition; ML, machine learning; DL, deep learning; NN, neural networks; SHM, structural health monitoring; BRBs, buckling restrained braces; LCA, lifecycle assessment; MLP, multilayer perceptron; FFBP, feedforward backpropagation; NARX, nonlinear autoregressive exogenous; NARX-SP, NARX series parallel; NARX-P, NARX parallel; LSTM, long short term memory; RNN, recurrent neural network; DNN, deep neural network; CNN, convolutional neural network; NB, naïve Bayes; KNN, K-nearest neighbors; GA, genetic algorithm; PSO, particle swarm optimization; GP, genetic programming; GEP, gene expression programming; MEP, multi expression programming; GPR, Gaussian process regression; MARS, multivariate adaptive regression spline; DT, decision tree; RT, regression tree; RF, random forest; BR, bagging regressor; AdaBoost, adaptive boosting; XGBoost, extreme gradient boosting; LGBoost, light gradient boosting; NGBoost, natural gradient boosting; GBR, gradient boosting regressor; CatBoost, categorized boosting; HGBoost, histogram gradient boosting; GB, gradient boosting; SVR, support vector regressor; SVM, support vector machine; RMSE, root-mean-square error; MAE, mean absolute error; MAPE, mean absolute percentage error; R^2, coefficient of determination; MCS, Monte Carlo simulation; GCV, generalized cross-validation; SHAP, shapley additive explanations; SD, standard deviation; CoV, coefficient of variation; CI, confidence intervals; RC, reinforced concrete; FE, finite element; FA, fine aggregates; CA. coarse aggregates; SF, silica fume; GGBS, ground granulated blast furnace slag; RAC, recycled aggregate concrete; FRP, fiber-reinforced polymer; DF, durability factor; CSB, castellated steel beam; SMRF, steel moment-resisting frame; EPR, evolutionary polynomial regression; LCC, lifecycle cost; MPa, megapascal.

CRediT authorship contribution statement

Md. Tarif Aziz: Writing – original draft. **Dave Montellano Osabel:** Writing – review & editing. **Youngju Kim:** Writing – review & editing. **Sanghoon Kim:** Validation, Supervision. **Jaehoon Bae:** Writing – review & editing, Supervision, Funding acquisition. **Konstantinos Daniel Tsavdaridis:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors hereby declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government (MSIT) [grant number 2022R1C1C1003594], and by the Chonnam National University Smart Plant Reliability Center grant funded by the Ministry of Education [grant number RS-2020-NF000321].

Data availability

The data that support the findings of this study are available from the corresponding author, [Jaehoon Bae], upon reasonable request.

References

[1] S.J. Russell, P. Norvig, Instructor's Solution Manual For Artificial intelligence: a Modern Approach, Pearson, 2003.
[2] K. Warwick, Artificial intelligence: the Basics, Routledge, 2013.
[3] S. Fouse, S. Cross, Z. Lapin, DARPA's impact on artificial intelligence, AI Mag. 41 (2) (2020) 3–8.
[4] J.O.N. Agar, What is science for? The Lighthill report on artificial intelligence reinterpreted, Br. J. Hist. Sci. 53 (3) (2020) 289–310.
[5] T. Menzies, 21st-century AI: proud, not smug, IEEE Intell Syst 18 (3) (2003) 18–24.
[6] P.P. Groumpos, A critical historic overview of artificial intelligence: issues, challenges, opportunities, and threats, In: Artif. Intell. Appl. 1 (4) (2023) 197–213.
[7] P. Ongsulee, Artificial intelligence, machine learning and deep learning, in: 2017 15th International Conference on ICT and Knowledge Engineering (ICT&KE), IEEE, 2017, pp. 1–6.
[8] I.H. Sarker, AI-based modeling: techniques, applications and research issues towards automation, intelligent and smart systems, SN Comput. Sci. 3 (2) (2022) 158.
[9] Y. Wang, et al., Brain-inspired systems: a transdisciplinary exploration on cognitive cybernetics, humanity, and systems science toward autonomous artificial intelligence, IEEE Syst. Man Cybern. Mag. 6 (1) (2020) 6–13, https://doi.org/10.1109/MSMC.2018.2889502.
[10] M. Flasiński, Introduction to Artificial Intelligence, Springer, 2016.
[11] Y. Xu, W. Qian, N. Li, H. Li, Typical advances of artificial intelligence in civil engineering, Adv. Struct. Eng. 25 (16) (2022) 3405–3424.
[12] Y. Pan, L. Zhang, Roles of artificial intelligence in construction engineering and management: a critical review and future trends, Autom. Constr. 122 (2021) 103517.
[13] B. Manzoor, I. Othman, S. Durdyev, S. Ismail, M.H. Wahab, Influence of artificial intelligence in civil engineering toward sustainable development—A systematic literature review, Appl. Syst. Innov. 4 (3) (2021) 52.
[14] S.R. Vadyala, S.N. Betgeri, J.C. Matthews, E. Matthews, A review of physics-based machine learning in civil engineering, Results Eng. 13 (2022) 100316.
[15] J. Rezania, M. Hamian, A. Rasekhi, Investigating the application of artificial intelligence in civil engineering and progressive collapse, Civ. Proj. 5 (9) (2023) 11–22.
[16] S.M. Harle, Advancements and challenges in the application of artificial intelligence in civil engineering: a comprehensive review, Asian J. Civ. Eng. 25 (1) (2024) 1061–1078.
[17] A.H. Gandomi, A.H. Alavi, M. Mousavi, S.M. Tabatabaei, A hybrid computational approach to derive new ground-motion prediction equations, Eng Appl Artif Intell 24 (2011) 717–732, https://doi.org/10.1016/j.engappai.2011.01.005.
[18] X. Guan, H. Burton, T. Sabol, Python-based computational platform to automate seismic design, nonlinear structural model construction and analysis of steel moment resisting frames, Eng. Struct. 224 (2020) 111199.
[19] K. Liu, C. Zou, X. Zhang, J. Yan, Innovative prediction models for the frost durability of recycled aggregate concrete using soft computing methods, J. Build. Eng. 34 (2021) 101822.
[20] M. Rajczakowska, M. Szeląg, K. Habermehl-Cwirzen, H. Hedlund, A. Cwirzen, Interpretable machine learning for prediction of post-fire self-healing of concrete, Mater. (Basel) 16 (3) (2023) 1273.
[21] M.Z. Naser, A. Seitllari, Concrete under fire: an assessment through intelligent pattern recognition, Eng Comput 36 (4) (2020) 1915–1928.
[22] J.P. Amezquita-Sanchez, M. Valtierra-Rodriguez, H. Adeli, Machine learning in structural engineering, Sci. Iran. 27 (6) (2020) 2645–2656.
[23] H.T. Thai, Machine learning for structural engineering: a state-of-the-art review, in: Structures, 38, Elsevier, 2022, pp. 448–491.
[24] P. Kumar, S.R. Kota, Machine learning models in structural engineering research and a secured framework for structural health monitoring, Multimed Tools Appl 83 (3) (2024) 7721–7759.
[25] A.T.G. Tapeh, M.Z. Naser, Artificial intelligence, machine learning, and deep learning in structural engineering: a scientometrics review of trends and best practices, Arch. Comput. Methods Eng. 30 (1) (2023) 115–159.
[26] W.Z. Taffese, L. Espinosa-Leal, Prediction of chloride resistance level of concrete using machine learning for durability and service life assessment of building structures, J. Build. Eng. 60 (2022) 105146.
[27] A. Mohammadi, S. Karimzadeh, S. Yaghmaei-Sabegh, M. Ranjbari, P.B. Lourenço, Utilising artificial neural networks for assessing seismic demands of buckling restrained braces due to pulse-like motions, Buildings 13 (10) (2023) 2542.
[28] A. Kaveh, A. Eskandari, M. Movasat, Buckling resistance prediction of high-strength steel columns using metaheuristic-trained artificial neural networks, in: Structures, 56, Elsevier, 2023 104853.
[29] P.P. Bhatt, N. Sharma, V.K.R. Kodur, M.Z. Naser, Machine learning approach for predicting fire resistance of FRP-strengthened concrete beams, Struct. Concr. (2024).
[30] M.W. Gardner, S.R. Dorling, Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences, Atmos, Env. 32 (14–15) (1998) 2627–2636.
[31] M. AlHamaydeh, I. Choudhary, K. Assaleh, Virtual testing of buckling-restrained braces via nonlinear autoregressive exogenous neural networks, J. Comput. Civ. Eng. 27 (2013) 755–768, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000247.
[32] E. Diaconescu, The use of NARX neural networks to predict chaotic time series, WSEAS Trans. Comput. Res. 3 (3) (2008) 182–191.
[33] R. Zhang, Z. Chen, S. Chen, J. Zheng, O. Büyüköztürk, H. Sun, Deep long short-term memory networks for nonlinear structural seismic response prediction, Comput. Struct. 220 (2019) 55–68.
[34] H. Salehi, R. Burgueño, Emerging artificial intelligence methods in structural engineering, Eng. Struct. 171 (2018) 170–189.

[35] M.Z. Naser, V.K. Kodur, Explainable machine learning using real, synthetic and augmented fire tests to predict fire resistance and spalling of RC columns, Eng. Struct. 253 (2022) 113824.
[36] A. Paral, D.K.S. Roy, A.K. Samanta, A deep learning-based approach for condition assessment of semi-rigid joint of steel frame, J. Build. Eng. 34 (2021) 101946.
[37] S. Mangalathu, H. Jang, S.H. Hwang, J.S. Jeon, Data-driven machine-learning-based seismic failure mode identification of reinforced concrete shear walls, Eng. Struct. 208 (2020) 110331.
[38] S. Zhang, Challenges in KNN classification, IEEE Trans Knowl Data Eng 34 (10) (2021) 4663–4675.
[39] J.H. Holland, Adaptation in Natural and Artificial systems: an Introductory Analysis With Applications to biology, control, and Artificial Intelligence, MIT Press, 1992.
[40] C. Camp, S. Pezeshk, G. Cao, Optimized design of two-dimensional structures using a genetic algorithm, J. Struct. Eng. 124 (5) (1998) 551–559.
[41] L.M. Le, H.B. Ly, B.T. Pham, V.M. Le, T.A. Pham, D.H. Nguyen, X.T. Tran, T.T. Le, Hybrid artificial intelligence approaches for predicting buckling damage of steel columns under axial compression, Mater. (Basel) 12 (10) (2019) 1670.
[42] T.V. Mathew, Genetic algorithm, Rep. submitt. IIT Bombay 53 (2012) 18–19.
[43] S.M. Mousavi, P. Aminian, A.H. Gandomi, A.H. Alavi, H. Bolandi, A new predictive model for compressive strength of HPC using gene expression programming, Adv. Eng. Softw. 45 (1) (2012) 105–114.
[44] Ferreira, C., 2001. Gene expression programming: a new adaptive algorithm for solving problems. arXiv preprint cs/0102027, 10.48550/arXiv.cs/0102027.
[45] M. Oltean, C. Grosan, A comparison of several linear genetic programming techniques, Complex Syst. 14 (4) (2003) 285–314.
[46] M. Khan, A. Khan, A.U. Khan, M. Shakeel, K. Khan, H. Alabduljabbar, T. Najeh, Y. Gamil, Intelligent prediction modeling for flexural capacity of FRP-strengthened reinforced concrete beams using machine learning algorithms, Heliyon 10 (1) (2024).
[47] J. Son, S. Yang, A new approach to machine learning model development for prediction of concrete fatigue life under uniaxial compression, Appl. Sci. 12 (19) (2022) 9766.
[48] S. Borra, A. Di Ciaccio, Improving nonparametric regression methods by bagging and boosting, Comput. Stat. Data Anal. 38 (4) (2002) 407–420.
[49] A. Nafees, S. Khan, M.F. Javed, R. Alrowais, A.M. Mohamed, A. Mohamed, N. I. Vatin, Forecasting the mechanical properties of plastic concrete employing experimental data using machine learning algorithms: DT, MLPNN, SVM, and RF, Polym. (Basel) 14 (8) (2022) 1583.
[50] M. Khan, A.U. Khan, M. Houda, C. El Hachem, M. Rasheed, W. Anwar, Optimizing durability assessment: machine learning models for depth of wear of environmentally-friendly concrete, Results Eng. 20 (2023) 101625.
[51] H. Kumarawadu, P. Weerasinghe, J.S. Perera, Evaluating the performance of ensemble machine learning algorithms over traditional machine learning algorithms for predicting fire resistance in FRP-strengthened concrete beams, Electron. J. Struct. Eng. 24 (3) (2024) 47–53.
[52] T.N.T. Ho, T.P. Nguyen, G.T. Truong, Concrete spalling identification and fire resistance prediction for fired RC columns using machine learning-based approaches, Fire Technol (2024) 1–44.
[53] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. (2001) 1189–1232.
[54] V.V. Degtyarev, K.D. Tsavdaridis, Buckling and ultimate load prediction models for perforated steel beams using machine learning algorithms, J. Build. Eng. 51 (2022) 104316.
[55] T. Duan, A. Anand, D.Y. Ding, K.K. Thai, S. Basu, A. Ng, A. Schuler, NGBoost: natural gradient boosting for probabilistic prediction, in: International Conference on Machine Learning, PMLR, 2020, pp. 2690–2700.
[56] H. Nguyen, T. Vu, T.P. Vo, H.T. Thai, Efficient machine learning models for prediction of concrete strengths, Constr. Build. Mater. 266 (2021) 120950.
[57] V. Vapnik, The Nature of Statistical Learning Theory, Springer Science & Business Media, 2013.
[58] Q. Tong, C. Couto, T. Gernay, Machine learning models for predicting the resistance of axially loaded slender steel columns at elevated temperatures, Eng. Struct. 266 (2022) 114620.
[59] A. Roy, S. Chakraborty, Support vector machine in structural reliability analysis: a review, Reliab. Eng. Syst. Saf. 233 (2023) 109126.
[60] F. Fu, Fire induced progressive collapse potential assessment of steel framed buildings using machine learning, J. Constr. Steel Res. 166 (2020) 105918, https://doi.org/10.1016/j.jcsr.2019.105918.
[61] V.V. Degtyarev, S.J. Hicks, F.P.V. Ferreira, K.D. Tsavdaridis, Probabilistic resistance predictions of laterally restrained cellular steel beams by natural gradient boosting, Thin-Walled Struct. 205 (2024) 112367.
[62] X. Gao, C. Lin, Prediction model of the failure mode of beam-column joints using machine learning methods, Eng Fail Anal 120 (2021) 105072.
[63] K. Liu, J. Zheng, S. Dong, W. Xie, X. Zhang, Mixture optimization of mechanical, economical, and environmental objectives for sustainable recycled aggregate concrete based on machine learning and metaheuristic algorithms, J. Build. Eng. 63 (2023) 105570.
[64] P. Zandifaez, E.A. Shamsabadi, A.A. Nezhad, H. Zhou, D. Dias-da-Costa, AI-assisted optimisation of green concrete mixes incorporating recycled concrete aggregates, Constr. Build. Mater. 391 (2023) 131851.
[65] E.M. Golafshani, T. Kim, A. Behnood, T. Ngo, A. Kashani, Sustainable mix design of recycled aggregate concrete using artificial intelligence, J. Clean. Prod. 442 (2024) 140994.
[66] A. Nafees, M.F. Javed, S. Khan, K. Nazir, F. Farooq, F. Aslam, M.A. Musarat, N. I. Vatin, Predictive modeling of mechanical properties of silica fume-based green concrete using artificial intelligence approaches: MLPNN, ANFIS, and GEP, Mater. (Basel) 14 (24) (2021) 7531.
[67] Y. Zou, C. Zheng, A.M. Alzahrani, W. Ahmad, A. Ahmad, A.M. Mohamed, R. Khallaf, S. Elattar, Evaluation of artificial intelligence methods to estimate the compressive strength of geopolymers, Gels 8 (5) (2022) 271.
[68] K.C. Onyelowe, A.M. Ebid, H.A. Mahdi, F.K. Onyelowe, Y. Shafieyoon, M. E. Onyia, H.N. Onah, AI mix design of fly ash admixed concrete based on mechanical and environmental impact considerations, Civ. Eng. J. 9 (2023) 27–45.
[69] M.C. Kang, D.Y. Yoo, R. Gupta, Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete, Constr. Build. Mater. 266 (2021) 121117.
[70] M.N. Amin, M. Iqbal, K. Khan, M.G. Qadir, F.I. Shalabi, A. Jamal, Ensemble tree-based approach towards flexural strength prediction of FRP-reinforced concrete beams, Polym. (Basel) 14 (7) (2022) 1303.
[71] K. Khan, M. Iqbal, B.A. Salami, M.N. Amin, I. Ahmad, A.A. Alabdullah, A.M. A. Arab, F.E. Jalal, Estimating flexural strength of FRP reinforced beam using artificial neural network and random forest prediction models, Polym. (Basel) 14 (11) (2022) 2270.
[72] T. Zhang, D. Gao, C. Xue, Flexural strength prediction of concrete beams reinforced with hybrid FRP and steel bars based on machine learning, in: Structures, 65, Elsevier, 2024 106652.
[73] Y. Li, Y. Liu, H. Lin, C. Jin, Study of flexural strength of concrete containing mineral admixtures based on machine learning, Sci. Rep. 13 (1) (2023) 18061.
[74] R.K. Paswan, A. Gogineni, S. Sharma, P. Kumar, Predicting split tensile strength in Portland and geopolymer concretes using machine learning algorithms: a comparative study, J. Build. Pathol. Rehabil. 9 (2) (2024) 129.
[75] I. Albaijan, A. Mahmoodzadeh, L.R. Flaih, H.H. Ibrahim, Y. Alashker, A. H. Mohammed, Evaluating the tensile strength of reinforced concrete using optimized machine learning techniques, Eng Fract Mech 292 (2023) 109677.
[76] J. de-Prado-Gil, C. Palencia, P. Jagadesh, R. Martínez-García, A comparison of machine learning tools that model the splitting tensile strength of self-compacting recycled aggregate concrete, Mater. (Basel) 15 (12) (2022) 4164.
[77] A. Gholampour, A.H. Gandomi, T. Ozbakkaloglu, New formulations for mechanical properties of recycled aggregate concrete using gene expression programming, Constr. Build. Mater. 130 (2017) 122–145.
[78] A. Cascardi, F. Micelli, M.A. Aiello, An artificial neural networks model for the prediction of the compressive strength of FRP-confined concrete circular columns, Eng. Struct. 140 (2017) 199–208, https://doi.org/10.1016/j.engstruct.2017.02.047.
[79] X. Huang, J. Sresakoolchai, X. Qin, Y.F. Ho, S. Kaewunruen, Self-healing performance assessment of bacterial-based concrete using machine learning approaches, Mater. (Basel) 15 (13) (2022) 4436.
[80] F. Althoey, M.N. Amin, K. Khan, M.M. Usman, M.A. Khan, M.F. Javed, M.M. S. Sabri, R. Alrowais, A.M. Maglad, Machine learning-based computational approach for crack width detection of self-healing concrete, Case Stud. Constr. Mater. 17 (2022) e01610.
[81] H. Alabduljabbar, K. Khan, H.H. Awan, R. Alyousef, A.M. Mohamed, S.M. Eldin, Modeling the capacity of engineered cementitious composites for self-healing using AI-based ensemble techniques, Case Stud. Constr. Mater. 18 (2023) e01805.
[82] K.A. Baghaei, S.A. Hadigheh, Durability assessment of FRP-to-concrete bonded connections under moisture condition using data-driven machine learning-based approaches, Compos Struct (2021) 114576.
[83] M.R. Sabour, G.A. Dezvareh, K.P. Niavol, Application of artificial intelligence methods in modeling corrosion of cement and sulfur concrete in sewer systems, Environ. Process. 8 (2021) 1601–1618.
[84] Y. XuanRui, Developing an artificial neural network model to predict the durability of the RC beam by machine learning approaches, Case Stud. Constr. Mater. 17 (2022) e01382.
[85] H. Chen, Y. Cao, Y. Liu, Y. Qin, L. Xia, Enhancing the durability of concrete in severely cold regions: mix proportion optimization based on machine learning, Constr. Build. Mater. 371 (2023) 130644.
[86] X. Huang, S. Wang, T. Lu, K. Wu, H. Li, W. Deng, J. Shi, Frost durability prediction of rubber concrete based on improved machine learning models, Constr. Build. Mater. 429 (2024) 136201.
[87] R. Kazemi, A hybrid artificial intelligence approach for modeling the carbonation depth of sustainable concrete containing fly ash, Sci Rep 14 (1) (2024) 11948.
[88] R. Jansson, L. Boström, Factors influencing fire spalling of self-compacting concrete, Mater. Struct. 46 (2013) 1683–1694.
[89] J.C. Liu, L. Huang, Z. Chen, H. Ye, A comparative study of artificial intelligent methods for explosive spalling diagnosis of hybrid fiber-reinforced ultra-high-performance concrete, Int. J. Civ. Eng. 20 (6) (2022) 639–660.
[90] A. Habib, S. Barakat, S. Al-Toubat, M.T. Junaid, M. Maalej, Developing machine learning models for identifying the failure potential of fire-exposed FRP-strengthened concrete beams, Arab. J. Sci. Eng. (2024) 1–16.
[91] M.J. Moradi, K. Daneshvar, D. Ghazi-Nader, H. Hajiloo, The prediction of fire performance of concrete-filled steel tubes (CFST) using artificial neural network, Thin-Walled Struct. 161 (2021) 107499.
[92] M.Z. Naser, Heuristic machine cognition to predict fire-induced spalling and fire resistance of concrete structures, Autom. Constr. 106 (2019) 102916.
[93] S. Li, J.R. Liew, M.X. Xiong, Prediction of fire resistance of concrete encased steel composite columns using artificial neural network, Eng. Struct. 245 (2021) 112877.
[94] R.O. Carvel, A.N. Beard, P.W. Jowitt, Fire spread between vehicles in tunnels: effects of tunnel size, longitudinal ventilation and vehicle spacing, Fire Technol 41 (2005) 271–304.

[95] Y.Z. Li, H. Ingason, Overview of research on fire safety in underground road and railway tunnels, Tunn. Undergr. Space Technol. 81 (2018) 568–589.
[96] X. Wu, Y. Park, A. Li, X. Huang, F. Xiao, A. Usmani, Smart detection of fire source in tunnel based on the numerical database and artificial intelligence, Fire Technol 57 (2021) 657–682.
[97] X. Wu, X. Zhang, X. Huang, F. Xiao, A. Usmani, A real-time forecast of tunnel fire based on numerical database and artificial intelligence. Building Simulation, Tsinghua University Press, 2022, pp. 1–14.
[98] R. Al Hamd, H. Warren, Predicting concrete-steel bond performance at high temperatures: a data-driven approach using AI modelling, in: 4th International Conference on Structural Safety Under Fire & Blast Loading, 2024. September 2024.
[99] A.H. Gandomi, A.H. Alavi, M. Gandomi, S. Kazemi, Formulation of shear strength of slender RC beams using gene expression programming, part II: with shear reinforcement, Measurement 95 (2017) 367–376.
[100] A. Cascardi, F. Micelli, M.A. Aiello, Analytical model based on artificial neural network for masonry shear walls strengthened with FRM systems, Compos. B: Eng. 95 (2016) 252–263.
[101] P. Mishra, P. Samui, E. Mahmoudi, Probabilistic design of retaining wall using machine learning methods, Appl. Sci. 11 (12) (2021) 5411.
[102] P.G. Asteris, M.E. Lemonis, T.T. Le, K.D. Tsavdaridis, Evaluation of the ultimate eccentric load of rectangular CFSTs using advanced neural network modeling, Eng. Struct. 248 (2021) 113297.
[103] P.G. Asteris, K.D. Tsavdaridis, M.E. Lemonis, F.P.V. Ferreira, T.T. Le, C.J. Gantes, A. Formisano, AI-powered GUI for prediction of axial compression capacity in concrete-filled steel tube columns, Neural Comput. Appl. 36 (35) (2024) 22429–22459.
[104] M. Lemonis, A. Daramara, A. Georgiadou, V. Siorikis, K.D. Tsavdaridis, P. Asteris, Ultimate axial load of rectangular concrete-filled steel tubes using multiple ANN activation functions. Steel and Composite structures, Int. J. 42 (4) (2022) 459–475.
[105] F.P.V. Ferreira, S.H. Jeong, E. Mansouri, R. Shamass, K. Tsavdaridis, C.H. Martins, S. De Nardin, Five machine learning models predicting the global shear capacity of composite cellular beams with hollow-core units. https://doi.org/10.3390/buildings14072256.
[106] F.P.V. Ferreira, R. Shamass, V. Limbachiya, K.D. Tsavdaridis, C.H. Martins, Lateral–torsional buckling resistance prediction model for steel cellular beams generated by artificial neural networks (ANN), Thin-Walled Struct. 170 (2022) 108592.
[107] R. Shamass, F.P.V. Ferreira, V. Limbachiya, L.F.P. Santos, K.D. Tsavdaridis, Web-post buckling prediction resistance of steel beams with elliptically-based web openings using artificial neural networks (ANN), Thin-Walled Struct. 180 (2022) 109959.
[108] M. Rabi, Y.S. Jweihan, I. Abarkan, F.P.V. Ferreira, R. Shamass, V. Limbachiya, K. D. Tsavdaridis, L.F.P. Santos, Machine learning-driven web-post buckling resistance prediction for high-strength steel beams with elliptically-based web openings, Results Eng. 21 (2024) 101749.
[109] G. Abdollahzadeh, S.M. Shabanian, Experimental and numerical analysis of beam to column joints in steel structures, Front. Struct. Civ. Eng. 12 (2018) 642–661, https://doi.org/10.1007/s11709-017-0457-z.
[110] H. Luo, S.G. Paal, Artificial intelligence-enhanced seismic response prediction of reinforced concrete frames, Adv. Eng. Inform. 52 (2022) 101568.
[111] S. Gharehbaghi, H. Yazdani, M. Khatibinia, Estimating inelastic seismic response of reinforced concrete frame structures using a wavelet support vector machine and an artificial neural network, Neural Comput. Appl. 32 (8) (2020) 2975–2988.
[112] K.M. Gondaliya, S.A. Vasanwala, A.K. Desai, J.A. Amin, V. Bhaiya, Machine learning-based approach for assessing the seismic vulnerability of reinforced concrete frame buildings, J. Build. Eng. 97 (2024) 110785.
[113] T.Y. Altıok, B. Üstüner, A. Özyüksel Çiftçioğlu, A. Demir, Enhancing structural evaluation: machine learning approaches for inadequate reinforced concrete frames, Iran. J. Sci. Technol. Trans. Civ. Eng. (2024) 1–21.
[114] J. Shin, D.W. Scott, L.K. Stewart, J.S. Jeon, Multi-hazard assessment and mitigation for seismically-deficient RC building frames using artificial neural network models, Eng. Struct. 207 (2020) 110204.
[115] J. Shin, S. Park, Optimum retrofit strategy of FRP column jacketing system for non-ductile RC building frames using artificial neural network and genetic algorithm hybrid approach, J. Build. Eng. 57 (2022) 104919.
[116] M.S. Sufyan, P. Samui, S.S. Mishra, Reliability analysis of frame structures under top-floor lateral load using artificial intelligence, Asian J. Civ. Eng. 24 (8) (2023) 3653–3665.
[117] M.S. Sufyan, P. Samui, S.S. Mishra, Reliability analysis of portal frame subjected to varied lateral loads using machine learning, Asian J. Civ. Eng. 25 (2) (2024) 2045–2058.
[118] H.D. Nguyen, Y.J. Lee, J.M. LaFave, M. Shin, Seismic fragility analysis of steel moment frames using machine learning models, Eng. Appl. Artif. Intell. 126 (2023) 106976.
[119] R.J. Kudari, L. Geetha, A. Satyanarayana, Assessing seismic vulnerability of structures with damper using an ANN-based approach, Asian J. Civ. Eng. 25 (7) (2024) 5335–5347.
[120] K.C. Onyelowe, J.L. Yaulema Castañeda, A.F.H. Adam, D.R. Ñacato Estrella, N. Ganasen, Prediction of steel plate-based damper for improving the behavior of concentrically braced frames based on RSM and ML approaches for sustainable structures, Sci Rep 14 (1) (2024) 4065.
[121] H.H.M. Al-Ghabawi, M.M. Khattab, I.A. Zahid, B. Al-Oubaidi, The prediction of the ultimate base shear of BRB frames under push-over using ensemble methods and artificial neural networks, Asian J. Civ. Eng. 25 (2) (2024) 1467–1485.
[122] J. Bae, C.-H. Lee, M. Park, R.W. Alemayehu, J. Ryu, Y.K. Ju, Modified low-cycle fatigue estimation using machine learning for radius-cut coke-shaped metallic damper subjected to cyclic loading, Int. J. Steel Struct. 20 (2020) 1849–1858, https://doi.org/10.1007/s13296-020-00377-7.
[123] S. Ji, B. Lee, M.Y. Yi, Building life-span prediction for life cycle assessment and life cycle cost using machine learning: a big data approach, Build Env. 205 (2021) 108267.
[124] P.C. Lazaridis, I.E. Kavvadias, K. Demertzis, L. Iliadis, L.K. Vasiliadis, Structural damage prediction of a reinforced concrete frame under single and multiple seismic events using machine learning algorithms, Appl. Sci. 12 (8) (2022) 3845.
[125] H. Bartsch, J. Voelkel, M. Feldmann, Developing artificial neural networks to estimate the fatigue strength of structural steel details using the new European database, Steel Construction (2024), https://doi.org/10.1002/stco.202400029.
[126] M.Z. Naser, V.K. Kodur, Explainable machine learning using real, synthetic and augmented fire tests to predict fire resistance and spalling of RC columns, Engineering Structures 253 (2022) 113824.
[127] K.M. Gondaliya, S.A. Vasanwala, A.K. Desai, J.A. Amin, V. Bhaiya, Machine learning-based approach for assessing the seismic vulnerability of reinforced concrete frame buildings, J. Build. Eng. 97 (2024) 110785.
[128] A.E. Özsoy Özbay, A decision tree-based damage estimation approach for preliminary seismic assessment of reinforced concrete buildings, Rev. Constr. 22 (1) (2023) 5–15.
[129] Taffese, W.Z., 2020. Data-driven method for enhanced corrosion assessment of reinforced concrete structures. arXiv preprint arXiv:2007.01164. https://doi.org/10.48550/arXiv.2007.01164.
[130] A. Rauf, U. Asif, K. Onyelowe, M.F. Javed, H. Alabduljabbar, Experimental analysis and gene expression programming optimization of sustainable concrete containing mineral fillers, Sci Rep 14 (1) (2024) 29280.
[131] R. Biswas, M. Kumar, D.R. Kumar, P. Samui, M.K. Rajak, D.J. Armaghani, S. Singh, Application of Novel Deep Neural Network On Prediction of Compressive Strength of Fly Ash-Based Concrete, Nondestructive Testing and Evaluation, 2024, pp. 1–31.
[132] Y. Murad, A. Tarawneh, F. Arar, A. Al-Zu'bi, A. Al-Ghwairi, A. Al-Jaafreh, M. Tarawneh, Flexural strength prediction for concrete beams reinforced with FRP bars using gene expression programming, in: Structures, 33, Elsevier, 2021, pp. 3163–3172.
[133] N. Sharma, M.S. Thakur, A. Upadhya, P. Sihag, Assessment of flexural strength of concrete with marble powder applying soft computing techniques, J. Build. Pathol. Rehabil. 8 (1) (2023) 4.
[134] W. Zhu, L. Huang, L. Mao, M. Esmaeili-Falak, Predicting the uniaxial compressive strength of oil palm shell lightweight aggregate concrete using artificial intelligence-based algorithms, Struct. Concr. 23 (6) (2022) 3631–3650.
[135] Y. Liu, Y. Cao, L. Wang, Z.S. Chen, Y. Qin, Prediction of the durability of high-performance concrete using an integrated RF-LSSVM model, Constr. Build. Mater. 356 (2022) 129232.
[136] P. Gao, Y. Song, J. Wang, Z. Yang, K. Wang, Y. Yuan, Prediction model for the chloride ion permeability resistance of recycled aggregate concrete based on machine learning, Buildings 14 (11) (2024) 3608.
[137] S. Malazdrewicz, Ł. Sadowski, Neural modelling of the depth of wear determined using the rotating-cutter method for concrete with a high volume of high-calcium fly ash, Wear 477 (2021) 203791.
[138] M.A. Khan, F. Farooq, M.F. Javed, A. Zafar, K.A. Ostrowski, F. Aslam, S. Malazdrewicz, M. Maślak, Simulation of depth of wear of eco-friendly concrete using machine learning based computational approaches, Mater. (Basel) 15 (1) (2021) 58.
[139] A. Khan, M. Khan, M. Ali, M. Khan, A.U. Khan, M. Shakeel, M. Fawad, T. Najeh, Y. Gamil, Predictive modeling for depth of wear of concrete modified with fly ash: a comparative analysis of genetic programming-based algorithms, Case Stud. Constr. Mater. 20 (2024) e02744.
[140] X. Gao, J. Yang, H. Zhu, J. Xu, Estimation of rubberized concrete frost resistance using machine learning techniques, Constr. Build. Mater. 371 (2023) 130778.
[141] M. Esmaeili-Falak, R. Sarkhani Benemaran, Application of optimization-based regression analysis for evaluation of frost durability of recycled aggregate concrete, Struct. Concr. 25 (1) (2024) 716–737.
[142] J. Bae, A. Jang, M.J. Park, J. Lee, Y.K. Ju, Assessment of concrete macrocrack depth using infrared thermography. Steel and Composite structures, Int. J. 43 (4) (2022) 501–509.
[143] Lalitha, G. and Reddy, C.R., 2023. Impermeability evaluation of concrete with fly ash aggregate and prediction with modelling.
[144] S. Alsubai, A. Alqahtani, S. Hashim Muhodir, A. Alanazi, M. Ahmed, D.J. Jasim, S. Palani, The remarkable potential of machine learning algorithms in estimating water permeability of concrete incorporating nano natural pozzolana, Sci. Rep. 14 (1) (2024) 12532.
[145] S.N. Londhe, P.S. Kulkarni, P.R. Dixit, A. Silva, R. Neves, J. De Brito, Predicting carbonation coefficient using artificial neural networks and genetic programming, J. Build. Eng. 39 (2021) 102258.
[146] V.Q. Tran, H.V.T. Mai, Q.T. To, M.H. Nguyen, Machine learning approach in investigating carbonation depth of concrete containing fly ash, Struct. Concr. 24 (2) (2023) 2145–2169.
[147] Z. Huo, L. Wang, Y. Huang, Predicting carbonation depth of concrete using a hybrid ensemble model, J. Build. Eng. 76 (2023) 107320.
[148] B. Cai, B. Zhang, F. Fu, Post-fire reliability analysis of concrete beams retrofitted with CFRPs: a new approach, Proc. Inst. Civ. Eng. – Struct. Build. 173 (11) (2020) 888–902.

[149] M. Hisham, G.A. Hamdy, O.O. El-Mahdy, Prediction of temperature variation in FRP-wrapped RC columns exposed to fire using artificial neural networks, Eng. Struct. 238 (2021) 112219.
[150] S.M. Kang, J.K. Kim, Prediction of the moment capacity of FRP-strengthened RC beams exposed to fire using ANNs, KSCE J. Civ. Eng. 27 (8) (2023) 3471–3483.
[151] Q.V. Vu, V.H. Truong, H.T. Thai, Machine learning-based prediction of CFST columns using gradient tree boosting algorithm, Compos. Struct. 259 (2021) 113505.
[152] X.Y. Zhao, J.X. Chen, B. Wu, An interpretable ensemble-learning-based open source model for evaluating the fire resistance of concrete-filled steel tubular columns, Eng. Struct. 270 (2022) 114886.
[153] Y.F. Zhu, Y. Yao, Y. Huang, C.H. Chen, H.Y. Zhang, Z. Huang, Machine learning applications for assessment of dynamic progressive collapse of steel moment frames, in: Structures, 36, Elsevier, 2022, pp. 927–934.
[154] J. Qiu, L. Jiang, Development of modular and reusable AI models for fast predicting fire behaviour of steel columns in structural systems, Eng. Struct. 297 (2023) 116994.
[155] Y. Mei, Y. Sun, F. Li, X. Xu, A. Zhang, J. Shen, Probabilistic prediction model of steel to concrete bond failure under high temperature by machine learning, Eng. Fail. Anal. 142 (2022) 106786.
[156] I.A. Reshi, A.H. Shah, A. Jan, Z. Tariq, S. Sholla, S. Rashid, M.U. Wani, Machine learning enhanced modeling of steel–concrete bond strength under elevated temperature exposure, Struct. Concr. 25 (6) (2024) 4609–4622.
[157] O.R. Abuodeh, J.A. Abdalla, R.A. Hawileh, Prediction of shear strength and behavior of RC beams strengthened with externally bonded FRP sheets using machine learning techniques, Compos. Struct. 234 (2020) 111698.
[158] G. Zhang, Z.H. Ali, M.S. Aldlemy, M.H. Mussa, S.Q. Salih, M.M. Hameed, Z.S. Al-Khafaji, Z.M. Yaseen, Reinforced concrete deep beam shear strength capacity modelling using an integrative bio-inspired algorithm with an artificial intelligence model, Eng. Comput. 38 (Suppl 1) (2022) 15–28.
[159] Y. Yu, X.Y. Zhao, J.J. Xu, S.C. Wang, T.Y. Xie, Evaluation of shear capacity of steel fiber reinforced concrete beams without stirrups using artificial intelligence models, Mater. (Basel) 15 (7) (2022) 2407.
[160] D.D. Nguyen, V.L. Tran, D.H. Ha, V.Q. Nguyen, T.H. Lee, A machine learning-based formulation for predicting shear capacity of squat flanged RC walls, in: Structures, 29, Elsevier, 2021, pp. 1734–1747.
[161] M. Khaleghi, J. Salimi, V. Farhangi, M.J. Moradi, M. Karakouzian, Application of artificial neural network to predict load bearing capacity and stiffness of perforated masonry walls, Civil Eng. 2 (1) (2021) 48–67.
[162] B. Keshtegar, M.L. Nehdi, R. Kolahchi, N.T. Trung, M. Bagheri, Novel hybrid machine learning model for predicting shear strength of reinforced concrete shear walls, Eng. Comput. (2022) 1–12.
[163] H. Zhang, X. Cheng, Y. Li, X. Du, Prediction of failure modes, strength, and deformation capacity of RC shear walls through machine learning, J. Build. Eng. 50 (2022) 104145.
[164] A. Tabrizikahou, G. Pavić, Y. Shahsavani, M. Hadzima-Nyarko, Prediction of reinforced concrete walls shear strength based on soft computing-based techniques, Soft Comput. 28 (15) (2024) 8731–8747.
[165] M. Koopialipoor, B.R. Murlidhar, A. Hedayat, D.J. Armaghani, B. Gordan, E. T. Mohamad, The use of new intelligent techniques in designing retaining walls, Eng Comput 36 (2020) 283–294.
[166] C. Cakiroglu, K. Islam, G. Bekdaş, M.L. Nehdi, Data-driven ensemble learning approach for optimal design of cantilever soldier pile retaining walls, in: Structures, 51, Elsevier, 2023, pp. 1268–1280.
[167] H.Q. Nguyen, H.B. Ly, V.Q. Tran, T.A. Nguyen, T.T. Le, B.T. Pham, Optimization of artificial intelligence system by evolutionary algorithm for prediction of axial capacity of rectangular concrete filled steel tubes under compression, Mater. (Basel) 13 (5) (2020) 1205.
[168] C. Cakiroglu, K. Islam, G. Bekdaş, U. Isikdag, S. Mangalathu, Explainable machine learning models for predicting the axial compression capacity of concrete filled steel tubular columns, Constr. Build. Mater. 356 (2022) 129227.
[169] T.A. Nguyen, H.B. Ly, Predicting axial compression capacity of CFDST columns and design optimization using advanced machine learning techniques, in: Structures, 59, Elsevier, 2024 105724.
[170] A.N. Hanoon, A.W. Al Zand, Z.M. Yaseen, Designing new hybrid artificial intelligence model for CFST beam flexural performance prediction, Eng Comput 38 (4) (2022) 3109–3135.
[171] T.A. Ebid Nguyen, H.B. Ly, H.V.T. Mai, V.Q. Tran, Using ANN to estimate the critical buckling load of Y-shaped cross-section steel columns, Sci Program 2021 (2021) 5530702.
[172] Y. Shen, J. Sun, S. Liang, Interpretable machine learning models for punching shear strength estimation of FRP reinforced concrete slabs, Crystals 12 (2) (2022) 259.
[173] G. Doğan, M.H. Arslan, Determination of punching shear capacity of concrete slabs reinforced with FRP bars using machine learning, Arab. J. Sci. Eng. 47 (10) (2022) 13111–13137.
[174] H. Liu, H. Wang, Y. Zhang, X. Liu, Shear resistance of UHPC connection for prefabricated reinforced concrete slabs with shear grooves and dowel rebars, Constr. Build. Mater. 454 (2024) 139153.
[175] H. Yan, N. Xie, D. Shen, Hybrid machine learning algorithms for prediction of failure modes and punching resistance in slab-column connections with shear reinforcement, Buildings 14 (5) (2024) 1247.
[176] M. Hosseinpour, Y. Sharifi, H. Sharifi, Neural network application for distortional buckling capacity assessment of castellated steel beams, in: Structures, 27, Elsevier, 2020, pp. 1174–1183.
[177] R. Shamass, F.P.V. Ferreira, V. Limbachiya, L.F.P. Santos, K.D. Tsavdaridis, Web-post buckling prediction resistance of steel beams with elliptically-based web openings using artificial neural networks (ANN), Thin-Walled Struct. 180 (2022) 109959.
[178] H.D. Nguyen, N.D. Dao, M. Shin, Prediction of seismic drift responses of planar steel moment frames using artificial neural network and extreme gradient boosting, Eng. Struct. 242 (2021) 112518.
[179] T.H. Nguyen, V.T. Phan, D.D. Nguyen, Practical ANN model for estimating buckling load capacity of corroded web-tapered steel I-section columns, Int. J. Steel Struct. 23 (6) (2023) 1459–1475.
[180] H. Dabiri, K. Rahimzadeh, A. Kheyroddin, A comparison of machine learning- and regression-based models for predicting ductility ratio of RC beam-column joints, in: Structures, 37, Elsevier, 2022, pp. 69–81.
[181] S. Ramavath, S.R. Suryawanshi, Optimal prediction of shear properties in beam-column joints using machine learning approach, Int. J. Eng. 37 (1) (2024) 67–82.
[182] W. Wen, C. Zhang, C. Zhai, Rapid seismic response prediction of RC frames based on deep learning and limited building information, Eng. Struct. 267 (2022) 114638.
[183] F. Kazemi, R. Jankowski, Machine learning-based prediction of seismic limit-state capacity of steel moment-resisting frames considering soil-structure interaction, Comput Struct 274 (2023) 106886.
[184] S.H. Hwang, S. Mangalathu, J. Shin, J.S. Jeon, Estimation of economic seismic loss of steel moment-frame buildings using a machine learning algorithm, Eng. Struct. 254 (2022) 113877.
[185] A. Su, J. Cheng, Y. Wang, Y. Pan, Machine learning-based processes with active learning strategies for the automatic rapid assessment of seismic resistance of steel frames, in: Structures, 72, Elsevier, 2025 108227.
[186] N. Asgarkhani, F. Kazemi, A. Jakubczyk-Gałczyńska, B. Mohebi, R. Jankowski, Seismic response and performance prediction of steel buckling-restrained braced frames using machine-learning methods, Eng Appl Artif Intell 128 (2024) 107388.
[187] H.H.M. Al-Ghabawi, M.M. Khattab, I.A. Zahid, B. Al-Oubaidi, The prediction of the ultimate base shear of BRB frames under push-over using ensemble methods and artificial neural networks, Asian J. Civ. Eng. 25 (2) (2024) 1467–1485.
[188] M.L. Murphy, et al., Building Code Requirements For Structural Concrete Reinforced with Glass Fiber-Reinforced Polymer (GFRP) Bars – Code and Commentary: An ACI Standard, American Concrete Institute, 2022.
[189] Makridakis, S. and Hibon, M., 1995. Evaluating accuracy (or error) measures [online].
[190] A. Botchkarev, A new typology design of performance metrics to measure errors in machine learning regression algorithms, Interdiscip. J. Inf. Knowl. Manag. 14 (2019) 45–76.
[191] M.Z. Naser, A.H. Alavi, Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences, Archit. Struct. Constr. 3 (4) (2023) 499–517.
[192] J.S. Armstrong, F. Collopy, Error measures for generalizing about forecasting methods: empirical comparisons, Int J Forecast 8 (1) (1992) 69–80.
[193] T. Foss, E. Stensrud, B. Kitchenham, I. Myrtveit, A simulation study of the model evaluation criterion MMRE, IEEE Trans. Softw. Eng. 29 (11) (2003) 985–995.
[194] J. Li, Assessing the accuracy of predictive models for numerical data: not r nor r^2, why not? Then what? PLOS ONE 12 (8) (2017) e0183250.
[195] E.A. Silver, D.F. Pyke, D.J. Thomas, Inventory and Production Management in Supply Chains, CRC Press, 2016.
[196] S.W. Jarantow, E.D. Pisors, M.L. Chiu, Introduction to the use of linear and nonlinear regression analysis in quantitative biological assays, Curr. Protoc. 3 (6) (2023) e801.
[197] M. Regona, T. Yigitcanlar, B. Xia, R.Y.M. Li, Opportunities and adoption challenges of AI in the construction industry: a PRISMA review, J. Open Innov.: Technol. Mark. Complex. 8 (1) (2022) 45.
[198] M.Z. Naser, AI-based cognitive framework for evaluating response of concrete structures in extreme conditions, Eng Appl Artif Intell 81 (2019) 437–449.
[199] T. Yigitcanlar, Greening the artificial intelligence for a sustainable planet: an editorial commentary, Sustainability 13 (24) (2021) 13508.
[200] J.J. Yun, D. Lee, H. Ahn, K. Park, T. Yigitcanlar, Not deep learning but autonomous learning of open innovation for sustainable artificial intelligence, Sustainability 8 (8) (2016) 797.
[201] S. Na, S. Heo, S. Han, Y. Shin, Y. Roh, Acceptance model of artificial intelligence (AI)-based technologies in construction firms: applying the Technology Acceptance Model (TAM) in combination with the Technology–Organisation–Environment (TOE) framework, Buildings 12 (2) (2022) 90.
[202] D. Young, K. Panthi, O. Noor, Challenges involved in adopting BIM on the construction jobsite, EPiC Ser. Built Environ. 2 (3) (2021) 302–310.
[203] A. Gondia, A. Siam, W. El-Dakhakhni, A.H. Nassar, Machine learning algorithms for construction projects delay risk prediction, J Constr Eng Manag 146 (1) (2020) 04019085.
[204] A. Koyamparambath, N. Adibi, C. Szablewski, S.A. Adibi, G. Sonnemann, Implementing artificial intelligence techniques to predict environmental impacts: case of construction products, Sustainability 14 (6) (2022) 3699.
[205] S.A. Sharif, A. Hammad, Developing surrogate ANN for selecting near-optimal building energy renovation methods considering energy consumption, LCC and LCA, J. Build. Eng. 25 (2019) 100790.
[206] J. Baehr, A. Koyamparambath, E. Dos Reis, S. Weyand, C. Binnig, L. Schebek, G. Sonnemann, Predicting Product Life Cycle Environmental Impacts With Machine learning: Uncertainties and Implications For Future Reporting Requirements, 52, Sustainable Production and Consumption, 2024, pp. 511–526.

[207] Askarinejad, P. and Behnia, B., 2024. Decarbonizing tall building structures: implementing machine learning At the early-stage of design process.
[208] D. Chicco, M.J. Warrens, G. Jurman, The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, Peerj comput. sci. 7 (2021) e623.
[209] B.H.Z. Sami, B.F.Z. Sami, P. Kumar, A.N. Ahmed, G.E. Amieghemen, M.M. Sherif, A. El-Shafie, Feasibility analysis for predicting the compressive and tensile strength of concrete using machine learning algorithms, Case Stud. Constr. Mater. 18 (2023) e01893.
[210] T.O. Hodson, Root mean square error (RMSE) or mean absolute error (MAE): when to use them or not, Geosci. Model Dev. Discuss. 2022 (2022) 1–10.
[211] Y. Yang, G. Liu, Data-driven shear strength prediction of FRP-reinforced concrete beams without stirrups based on machine learning methods, Buildings 13 (2) (2023) 313.
[212] L.R. Kalabarige, J. Sridhar, S. Subbaram, P. Prasath, R. Gobinath, Machine learning modeling integrating experimental analysis for predicting compressive strength of concrete containing different industrial byproducts, Adv. Civ. Eng. 2024 (1) (2024) 7844854.
[213] B. Sun, Y. Zhang, C. Huang, Machine learning-based seismic fragility analysis of large-scale steel buckling restrained brace frames, Comput. Model. Eng. Sci. 125 (2) (2020) 755–776.
[214] M.F. Tamimi, A.A. Alshannaq, I. Mu'ath, Sensitivity and reliability assessment of buckling restrained braces using machine learning assisted-simulation, J. Constr. Steel Res. 211 (2023) 108187.
[215] T.P. Anand, M.S. Pandikkadavath, S. Mangalathu, D.R. Sahoo, Machine learning models for seismic analysis of buckling-restrained braced frames, J. Build. Eng. 98 (2024) 111398.
[216] A.M. Sagheer, M. Alhamaydeh, J. Fayaz, Z.A. Al-Sadoon, Deep learning-based modeling of the cyclic behavior of replaceable fuse buckling-restrained braces (BRBs), in: Structures, 63, Elsevier, 2024 106484.
[217] A. Mohammadi, S. Karimzadeh, S. Yaghmaei-Sabegh, M. Ranjbari, P.B. Lourenço, Utilising artificial neural networks for assessing seismic demands of buckling restrained braces due to pulse-like motions, Buildings 13 (10) (2023) 2542.
[218] K.C. Onyelowe, J.L. Yaulema Castañeda, A.F.H. Adam, D.R. Ñacato Estrella, N. Ganasen, Prediction of steel plate-based damper for improving the behavior of concentrically braced frames based on RSM and ML approaches for sustainable structures, Sci Rep 14 (1) (2024) 4065.
[219] P.C. Chen, K.Y. Chien, Machine-learning based optimal seismic control of structure with active mass damper, Appl. Sci. 10 (15) (2020) 5342.
[220] T. Shao, B. Andrawes, Using machine learning to predict the seismic response of an SDOF RC structure with superelastic dampers, Int. J. Civ. Eng. 20 (10) (2022) 1165–1180.
[221] S. Hu, W. Wang, Y. Lu, Explainable machine learning models for probabilistic buckling stress prediction of steel shear panel dampers, Eng. Struct. 288 (2023) 116235.
[222] T.A. Nguyen, K. Le Nguyen, H.B. Ly, Universal boosting ML approaches to predict the ultimate load capacity of CFST columns, Struct. Des. Tall Spec. Build. 33 (2) (2024) e2071.
[223] A. Alnaqbi, G.G. Al-Khateeb, W. Zeiada, M. Abuzwidah, Random forest-based frame work for multi-distress prediction in CRCP: a feature importance approach, Discov. Civ. Eng. 2 (1) (2025) 140.
[224] G. Gallitto, R. Englert, B. Kincses, R. Kotikalapudi, J. Li, K. Hoffschlag, U. Bingel, T. Spisak, External validation of machine learning models—Registered models and adaptive sample splitting, GigaScience 14 (2025) giaf036.
[225] G. Varoquaux, O. Colliot, Evaluating machine learning models and their diagnostic value, Mach. Learn. Brain Disord. (2023) 601–630.
[226] R.D. Riley, J. Ensor, K.I. Snell, L. Archer, R. Whittle, P. Dhiman, J. Alderman, X. Liu, L. Kirton, J. Manson-Whitton, M. van Smeden, Importance of sample size on the quality and utility of AI-based prediction models for healthcare, Lancet Digit. Health (2025).
[227] A. Ghaffari, Y. Shahbazi, M. Mokhtari Kashavar, M. Fotouhi, S. Pedrammehr, Advanced predictive structural health monitoring in high-rise buildings using recurrent neural networks, Buildings 14 (10) (2024) 3261.
[228] B. Panfeng, Z. Songlin, C. Hongyu, L. Caiwei, W. Pengtao, Q. Lichang, Structural monitoring data repair based on a long short-term memory neural network, Sci Rep 14 (1) (2024) 9974.