MACHINE LEARNING METHODS IN ANALYSIS OF CONCRETE: A STATE-OF-THE-ART REVIEW

SAIDJON KAMOLOV

ABSTRACT. Machine learning application have permeated many facets of the society including civil engineering. It is used to predict and analyze various scenarios that were hitherto out of the reach of the human experts. In this paper, we aim to review the current state-of-the-art in machine learning applications related to design, compression strength, and fault detection aspect of concrete material. We identify the most relevant works and describe the strengths and weaknesses of each approach as well as provide recommendations for improvement.

1. Introduction

Machine learning (ML) applications have affected many aspects of human endeavor. From natural language processing (NLP) to computer vision (CV), ML and AI tools have improved the productivity across multiple industries. The field of civil engineering has not been left disaffected. There is a significant amount of research that has been undertaken in the effort to improve processes and outcomes in civil engineering. In particular, application of ML with respect to concrete have seen a rise in attention. Researchers have applied ML to several aspects of concrete including its design, compression strength, and fault detection. The plethora of research in this field necessitates a review of the current literature. To this end, the goal of the present paper is to survey the literature related to the applications of ML to concrete, identify and describe major approaches, provide feedback on the strengths and weakness of the current state-of-the-art.

Concrete plays a crucial role in civil engineering. Concrete is the second-most-used substance in the world after water [3], and is the most widely used building material [2]. Its usage worldwide is twice that of steel, wood, plastics, and aluminum combined. It is a primary material in construction and forms the backbone of modern physical infrastructures. While concrete is a well-known material there remains room for better understanding of its properties under different conditions.

Date: Received: Sep 2, 2023; Accepted: Sep 29, 2023.
* Corresponding author.

Key words and phrases. machine learning, deep learning, concrete, civil engineering, construction, compression strength, concrete design, fault detection.
The application of ML techniques in the study of concrete has progressed along three primary routes: design, compression strength, and fault detection. The design of concrete structures often involves nonlinear systems that are hard to solve mathematically. In this case, soft computing methods including machine learning have potential to be useful. Several studies have attempted to apply ML methods to produce improved concrete designs. Compression strength analysis is another avenue of modern research. Researchers have attempted to build intelligent systems that are able to forecast the compression strength of concrete under different conditions. Finally, fault detection is another active area of research. Since concrete is susceptible to corrosion due to weather, early detection of fault signs is vital to the maintenance of concrete structures including buildings and bridges. This study aims to analyze the existing literature along the above directions and evaluate the current state-of-the-art.

The paper is structured as follows. Section 2 provides a brief overview popular ML techniques. Section 3 discusses ML applications in concrete design. Section 4 analyzes the literature related to ML applications in compression strength. Section 5 is concerned with the use of ML in fault detection. We end with concluding remarks in Section 6.

2. Machine Learning

Recent advances in machine learning algorithms, increased computing power, and big data have led to significant improvements in the performance of machine learning algorithms. Breakthroughs in natural language processing and computer vision as well as robotics demonstrate the potential of AI in the society. Generative AI has been the newest form of machine learning algorithms that has shifted the paradigm. While significant advances have been made over the last few years, the gap between the human and digital mind remains abysmal.

Machine learning applications have permeated many facets of science and engineering [14, 23] This section provides a brief overview of several core algorithms, namely MLP, CNN, RNN, LSTM, and Autoencoders. In addition, the discussion extends to address the importance of feature selection and the challenges associated with imbalanced datasets.

2.1. Multilayer Perceptron. A Multilayer Perceptron (MLP) is a type of feed-forward artificial neural network modeled after the structure of human brain. It consists of multiple layers: an input layer, one or more hidden layers, and an output layer. Each layer is made up of neurons, where every neuron in one layer connects to every neuron in the subsequent layer. The operation in each neuron is a weighted sum followed by an activation function [16]. This can be mathematically expressed as:

\[ y = f \left( \sum_i w_i x_i + b \right), \]  

(2.1)
where: $x_i$ are the input values, $w_i$ are the associated weights, $b$ is a bias term, and $f$ is the activation function, such as the sigmoid or ReLU. MLP is used in a variety of applications including finance, healthcare, and others [15, 68].

2.2. **Convolutional Neural Networks.** Convolutional Neural Networks (CNN) are particularly tailored for grid-structured data, notably images. These networks have a unique architectural design to automatically and adaptively learn spatial hierarchies of features from input images. The fundamental operation in the convolutional layer is the convolution operation:

$$(I * K)(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n)$$  \hspace{1cm} (2.2)

where: - $I$ represents the input image or feature map, - $K$ is the kernel or filter, and - '$*$' denotes the convolution operation.

CNNs are designed to work on 2-dimensional data such as images. A sliding window is applied across the 2-dimensional input and the convolution operation is carried out at each sliding step. CNNs are feature preserving and efficient. In addition, CNN can be used in 1-dimensional data such as time-series. The applications of CNN include image processing, finance, healthcare, and others [17, 18].

2.3. **Recurrent Neural Networks.** Recurrent Neural Networks (RNN) are quintessential for sequential data, capturing temporal dependencies. Unlike traditional neural networks, RNNs retain a memory, or hidden state, from one step to the next:

$$h_t = f(W_{hh}h_{t-1} + W_{xh}x_t + b),$$  \hspace{1cm} (2.3)

where: - $h_t$ is the hidden state at time $t$, - $x_t$ denotes the input at that time, - $W_{hh}$ and $W_{xh}$ are weight matrices, and - $b$ is the bias.

In RNN, the output in the current time step depends on the input in the current time together with the output in the previous step. In this way, RNNs preserve the memory from the history of the network. Unfortunately, RNNs do not work well with long sequences resulting in exploding or vanishing gradients. RNNs are used in a variety of time-series applications related to forecasting [28, 29].

2.4. **Long Short-Term Memory.** Long Short-Term Memory (LSTM), a special kind of RNN, are adept at capturing long-term dependencies in data. They incorporate a cell state and utilize various gates to regulate the flow of information:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$  \hspace{1cm} (2.4)

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$  \hspace{1cm} (2.5)

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C)$$  \hspace{1cm} (2.6)

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t$$  \hspace{1cm} (2.7)
LSTMs are an extension of RNN but with a more advanced gate structure that allows the control of the gradient flow. It is designed to deal with the main issue in RNNs which the exploding or vanishing gradient. By controlling the rate of gradient flow LSTMs ensure that the signal is not lost as it travels downstream the neural network. Thus, LSTM perform better than RNN on longer sequences where there is a greater chance of problems related to the gradient. LSTMs are used in a variety of forecasting applications [45, 46, 47].

2.5. **Autoencoders.** Autoencoders are unsupervised neural network models aimed at data compression and decompression. They consist of an encoder that compresses the input into a latent-space representation and a decoder that reconstructs the input from this representation. The objective is to minimize the difference between the input and the reconstructed output.

Autoencoders belong to the family of generative AI. By altering the input in the latent space one can generate different outputs in smooth manner. Since the latent space representation is often narrower than the original input space, autoencoders act as feature extraction modules. Autoencoders are used in a variety applications including healthcare, education, and imbalanced data [48, 49, 50].

2.6. **Feature Selection.** Feature selection is pivotal in machine learning to enhance the model’s performance by choosing only the most relevant features. It aids in simplifying models, expediting training, and curtailing the risk of overfitting. Various techniques, both statistical and algorithmic, can be employed for this purpose, such as mutual information, recursive feature elimination, and feature importance from tree-based algorithms.

Feature selection algorithms can be grouped into three main categories: filters, wrappers, and embedded methods. Filter methods employ a univariate statistic to measure the importance of an individual feature with respect to the target variable. The popular means of evaluating individual features using filters include mutual information and $\chi^2$-statistic. Wrapper methods employ a classifier to determine the value of a feature. A popular example of a wrapper is recursive feature elimination where a classifier is fitted on the data and the feature with the lowest coefficient value in the fitted model is eliminated. Embedded methods perform feature selection as part of the model fitting. Lasso regression has a regularization structure that is conducive to small number of features. Feature selection is used in a variety of applications.

3. **Concrete Design**

The mechanical properties of high-strength concrete are significantly influenced by the structure of its constituent materials and the reactions occurring between them. Consequently, concrete engineers and designers strive to develop a detailed understanding of how different types and quantities of materials influence concrete construction [37]. The strength of concrete, fundamentally a blend of water, cement, and other components, hinges largely on the chemical interplay between the cement and water [35]. It is imperative for high-strength concrete to
maintain both resistance and flexibility to a variety of forces and environmental conditions. The judicious selection and combination of raw materials, therefore, play a pivotal role in enhancing the robustness and quality of the concrete. Multiple factors including mixing properties and methods, conditions of mixture, and transportation affect the compressive strength of the concrete [33]. While design engineers may have an initial understanding of the composition and potential material proportions, precise estimation remains crucial. The quality of the concrete can be truly ascertained only after it has been efficiently produced and subjected to specialized testing after a certain period, often extending to a month [36]. This current methodology, leaning heavily on trial and error, sometimes results in substantial waste of raw materials. Monitoring the decreasing slope in the concrete mixture, however, can offer valuable insights into its performance [43].

Over recent decades, the surge in global construction activities has prompted extensive research to enhance concrete quality. Artificial intelligence (AI), particularly artificial neural networks (ANNs), has emerged as a prominent tool in analyzing data in this domain, aiding in the accurate estimation of requisite raw material quantities [54]. Alongside ANNs, adaptive neuro-fuzzy inference systems (ANFIS) have also been employed for clustering methods [56]. The integration of rapid learning with Cascade-forward neural networks (CfNN) has been noted for reducing the processing time of concrete constituents, thereby facilitating the achievement of desired results more efficiently [52]. Efforts to boost the accuracy of predictions have seen the coupling of neural networks with evolutionary algorithms or metaheuristic techniques (MH) [44, 42]. These strategies seek to rectify the inherent shortcomings of ANNs and other regression-based estimation approaches, with genetic algorithms (GA) [41, 53], particle swarm optimization (PSO) [55], and simulated annealing (SA) [40] finding widespread application in training ANNs for precise error determination. Despite potential challenges such as occasional entrapment in local optimums, genetic algorithms are favored for their quick convergence [41, 53]. Recently, there has been an uptick in studies utilizing ANN [53, 55, 70] and SVR [40] for regression network estimation. Furthermore, deep learning methods have been gaining traction in predicting concrete quality, encouraging scholars to explore various deep learning models to uplift the standard of concrete [71, 60, 61]. The authors in [72] delineate an evolutionary-based model to forecast the 28-day compressive strength of high-performance concrete enriched with cementitious elements.

4. Compressive strength

The widespread use of concrete in construction owes to its beneficial properties, which are derived from this composition. Besides the conventional type, high-strength concrete is another variant enriched with special additives to enhance its properties. Yet, accurately predicting the attributes of hardened concrete remains a significant challenge due to a range of predictable and unpredictable parameters influencing its characteristics [10, 9].

One pivotal property to discern is the compressive strength (CS), which plays a critical role in the design of engineering structures and directly influences other
vital attributes like water tightness and elastic modulus. Currently, determining CS involves testing a plethora of cylindrical or cubic samples at differing ages, a procedure that is both time-consuming and costly. Alterations in concrete mix proportions often necessitate repeated tests to avoid undesired outcomes, a situation partly mitigated by replacing a portion of the cement with pozzolan powders [1, 4, 8].

The prediction of concrete’s CS has spurred considerable research interest, with machine learning (ML) becoming a central tool in this endeavor, demonstrating prowess in solving linear and nonlinear problems where traditional mathematical models fall short [5]. A remarkable initiative by the authors in [7] led to the development of two smart computation (SC) systems, namely Artificial Neural Networks (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS), showing promising results in predicting CS, although the latter displayed a slight edge in accuracy.

A string of investigations by several groups of researchers [19, 12, 20] echoed the potential of leveraging artificial intelligence (AI) models such as Support Vector Regression (SVR) and ANN in predicting the 28-day CS of concrete, with SVR presenting itself as a faster and more precise alternative. Encouraging results emerged from integrating SVR with cross-validation approaches, showcasing enhanced accuracy over other AI models. Simultaneously, ANN displayed competency in predicting high-performance and self-compacting concretes [22]. SVR continued to showcase its utility in predicting the early age (7-day) CS of lightweight foamed concrete [70].

Comparative studies highlighting the efficacy of ANN and ANFIS and others involving ANN and Multiple Linear Regression (MLR) approaches have further enriched the understanding of CS estimation, albeit with varying degrees of success [6, 13]. The exploration continued with the team of researchers in [21] offering a higher prediction accuracy through ANN, a trend echoed in a novel proposal utilizing a hybrid ML approach in handling industrially wasteful materials in concrete.

To augment standard models, scholars have introduced hybrid AI-based metaheuristic algorithms, optimizing ANN, SVR, and ANFIS frameworks to procure better predictions [26, 24, 27, 30, 39, 32]. Significant strides were made by the authors in [34] through utilizing adaptive neuro-fuzzy inference systems (ANFIS) and the Group Method of Data Handling (GMDH), enhancing the latter with genetic algorithms and singular value decomposition methods to accurately estimate CS at various ages. Subsequent studies have continued to explore the synergies of ANN with genetic algorithms and Particle Swarm Optimization (PSO), opening new frontiers in CS prediction of diverse concrete materials with improved accuracy [38, 31].

5. Concrete Crack Detection

Cracks in concrete happen due to a variety of reasons but mainly due to weather. Heat, water, wind, and other meteorological factors affect the condition of concrete [51, 25]. Cracks compromise the mechanical function, endurance,
serviceability [69], and durability of concrete and its structures [58]. Traditional methods of crack detection tend to be subjective [59], time-consuming, and costly, often requiring human inspectors [69, 73]. As a result, several deep learning (DL) applications have been developed for the detection of cracks in concrete structures.

With the advancement of computer vision technology, convolutional layers have taken a pivotal role in image recognition. The authors in [73] employed R-CNN to identify multiple cracks on concrete surfaces. However, the detection was sub-optimal when two cracks were interconnected. To precisely categorize transverse, longitudinal, and crocodile cracks, the researchers in [77] incorporated deformation rules within the convolutional and pooling layers of R-CNN. The authors in [69] concentrated on the classification and localization of cracks on concrete surfaces, especially those with noises resembling cracks. To minimize the noise impact, researchers in [78] employed both a fully convolutional network and a naive Bayes data fusion (NB-FCN) model to detect cracks on a concrete bridge. In [79], the team presented a deeply supervised object detector to pinpoint the initial position of fatigue cracks. The use of the Voronoi diagram was highlighted as a means to forecast crack patterns and directions based on existing cracks [57], thereby offering a chance for preventative action. In [76], the authors integrated a weight share in CNN to expedite the training process for crack identification and localization. Yet, their method struggled to recognize cracks that shared a similar color with the background.

6. Conclusion

Concrete plays a crucial role in construction and civil engineering. Recently there has been an increased push towards the application of machine learning algorithms in the study of concrete. In particular, research efforts have been directed primarily in 3 directions: concrete design, compression strength, and fault detection. In this paper, we reviewed the major approaches along each axis of investigation. Strengths and weakness of the existing approaches have been identified. We conclude that while a lot progress has been made, there remains a substantial room for improving the existing machine learning methods in the study of concrete and its properties.

References

MACHINE LEARNING METHODS IN CONCRETE


Email address: said.kamolov@ttu.tj