

Article

Mixture Optimization of Recycled Aggregate Concrete Using Hybrid Machine Learning Model

Itzel Nunez , Afshin Marani  and Moncef L. Nehdi * 

Department of Civil and Environmental Engineering, Western University, London, ON N6G 1G8, Canada; inezvarg@uwo.ca (I.N.); amarani@uwo.ca (A.M.)

* Correspondence: mnehdi@uwo.ca

Received: 26 August 2020; Accepted: 24 September 2020; Published: 29 September 2020



Abstract: Recycled aggregate concrete (RAC) contributes to mitigating the depletion of natural aggregates, alleviating the carbon footprint of concrete construction, and averting the landfilling of colossal amounts of construction and demolition waste. However, complexities in the mixture optimization of RAC due to the variability of recycled aggregates and lack of accuracy in estimating its compressive strength require novel and sophisticated techniques. This paper aims at developing state-of-the-art machine learning models to predict the RAC compressive strength and optimize its mixture design. Results show that the developed models including Gaussian processes, deep learning, and gradient boosting regression achieved robust predictive performance, with the gradient boosting regression trees yielding highest prediction accuracy. Furthermore, a particle swarm optimization coupled with gradient boosting regression trees model was developed to optimize the mixture design of RAC for various compressive strength classes. The hybrid model achieved cost-saving RAC mixture designs with lower environmental footprint for different target compressive strength classes. The model could be further harvested to achieve sustainable concrete with optimal recycled aggregate content, least cost, and least environmental footprint.

Keywords: recycled aggregate concrete; machine learning; model; Gaussian process; deep learning; gradient boosting; regression trees; gated recurrent unit

1. Introduction

Portland cement concrete is the most widely used construction material and the most consumed industrial product in the world. However, the more frequent local shortages of natural aggregates (NA) and the enormous environmental footprint caused by its extraction are global concerns regarding the production of Portland cement concrete. The global annual consumption of NA is estimated at 8 to 12 billion tons [1–3].

Moreover, there has been a growing worldwide shortage of available landfilling sites to dispose of construction and demolition waste (CDW). In Canada, it is estimated that 9 million tons of CDW are produced every year. Despite the vast country, its biggest cities are encountering stern CDW disposal challenges [4]. Likewise, several reports are forecasting that landfills in Hong Kong will be overfilled in about eight years [2]. The use of recycled aggregate (RA) offers a potential solution to overcome the drawbacks of concrete production. Among the most promising advantages of RA are reductions in CO₂ emissions and beneficiation of CDW in applications with added value. About 75% of construction waste, including concrete and masonry, can be reused in concrete production [5].

However, the inclusion of RA in concrete has been reported to reduce the compressive strength [6]. Several researchers have explored the influential factors on recycled aggregate concrete (RAC) compressive strength [7–10]. The moisture content of RA, replacement level of NA, and the water-to-cement (w/c) ratio were found as the parameters with the highest effect on RAC compressive

strength [10]. The higher absorption capacity of RA compared to that of NA, along with the relatively weaker interfacial bond between RA and the cementitious matrix are major explanations for such a behavior [9,11,12].

Although there have been multiple studies on the mechanical properties of RAC, more research needs to be devoted to investigating the effects of various parameters on RAC compressive strength, including the variability of RA, the effect of old cement mortar attached to the RA surface, and the crushing process of RA [7,13]. Considering the fundamental knowledge gaps in the mechanical, durability, and structural performance of RAC, its application has generally been limited to road bases and non-structural concrete [6,8]. Hence, it is of paramount importance to conduct comprehensive experimental studies on RAC, develop pertinent standards and guidelines, and deploy advanced practical frameworks to promote its wider utilization.

The lack of understanding of RAC has resulted in developing new modeling techniques, such as machine learning (ML) algorithms, to predict its mechanical properties. A major advantage of ML methods is that they can capture the underlying mechanisms, despite the lack of information regarding specific parameters. Data driven ML techniques have proven to be successful in the prediction of RAC mechanical properties including the modulus of elasticity and compressive strength [11,14–16]. Nevertheless, most research studies employed small datasets, which compromises the training of such models and the ability to generalize their predictions to new data unseen to the model. Thus, creating reliable and more comprehensive data has been regarded as a key research need, and several studies aimed at deploying larger datasets for better generalization and robustness of the RAC-ML models.

ML techniques have also been employed for mixture design and optimization. Concrete mixture design is the process of selecting the appropriate quantitative proportions of concrete ingredients [17]. From a computational point of view, mixture optimization is the process of minimizing a prior defined objective function [18]. A common practice in concrete mixture optimization procedures is to consider the concrete production cost function as the objective function [19–21]. Moreover, the current stringent mechanical requirements for concrete should be met in the optimization process. Hence, in this study, the particle swarm optimization (PSO) algorithm was used to execute the mixture optimization. Subsequently, to assure that the compressive strength was met, the best performing ML model was used to predict the compressive strength of the RAC.

Accordingly, the present study aims at creating a large and comprehensive experimental dataset from the available studies in the open literature to develop powerful state-of-the-art ML models to predict the compressive strength of RAC. For this purpose, a dataset consisting of 1134 experimental examples of RAC mixture designs with 10 attributes was developed. Moreover, three different novel ML models were developed, and their performance was compared. Gaussian processes (GP), deep learning (DL) and gradient boosting regression trees (GBRT) techniques were employed for the first time to model the compressive strength of RAC. Eventually, an optimization of the RAC mixture design was performed by coupling PSO with the best proposed ML model to develop a hybrid powerful model for optimizing RAC mixture composition for different target ranges of compressive strength at 28 days. The superior accuracy of the proposed models should assist various stakeholders in optimal use of recycled concrete in diverse construction applications.

1.1. Related Work

Other studies have employed ML to predict the compressive strength of RAC. For instance, Khademi et al. [15] used three different approaches to model the compressive strength of RAC: artificial neural network (ANN), adaptive neuro-fuzzy inference systems (ANFIS), and multiple linear regression. They used 14 different input parameters, including the dosage of concrete ingredients and non-dimensional parameters such as the water-to-cement ratio and aggregate-to-cement ratio. It was concluded that multiple linear regression might be inaccurate to determine the compressive strength of RAC due to the highly non-linear relationships between the concrete ingredients and its

strength. However, both ANN and ANFIS models proved to be powerful in modeling the compressive strength of RAC, with a coefficient of determination of 0.9185 and 0.9075 for ANN and ANFIS, respectively. Furthermore, Khademi et al. [15] performed a sensitivity analysis in which they concluded that the inclusion of more input features resulted in higher model predictive accuracy. Likewise, Naderpour et al. [1] developed an ANN model to predict the compressive strength of RAC with a coefficient of determination of 0.829 for the testing dataset. They also performed a sensitivity analysis via the weights of the input features. Accordingly, it was found that the water absorption of aggregates and the water-to-total materials mass ratio had the highest importance. In another study, Deng et al. [16] built a convolutional neural network to predict the compressive strength of RAC. Experimental work was carried out along with the development of the deep learning model. The authors compared the convolutional neural network with a support vector machine and a back propagation neural network concluding that the convolutional neural network has superior capability to predict the compressive strength of RAC. They used the relative error percentage to measure the performance of the models, and thus the error for the convolutional neural network, back propagation neural network and support vector machine was 3.65%, 6.63%, and 4.35%, respectively. Deshpande et al. [11] compared three different techniques: ANN, model tree, and non-linear regression. They studied the influence of adding non-dimensional parameters as input features. To accomplish such analysis, they created 10 different models for each algorithm and added a different non-dimensional input feature to the parameters corresponding to the ingredients content. The accuracy of the ANN model was at least 2% higher than that of the other techniques, even when the non-dimensional parameters were considered. Using a larger dataset, Gholampour et al. [22] predicted the compressive strength and other mechanical properties of RAC employing three types of algorithms including multivariate adaptive regression splines, M5 model tree, and least squares support vector regression. They created two different models for each algorithm corresponding to the cube and cylinder compressive strengths, respectively. For these models, results on 332 cube-specimens and 318 cylinder-specimens were collected from the open literature. It was found that least squares support vector regression achieved higher performance than the remaining models, with at least 12.6% better mean absolute percentage error. Duan et al. [2] proposed using the characteristics of the recycled aggregate as input parameters, including the saturated surface dry mass, water absorption, and volume fraction of coarse aggregate. They concluded that the inclusion of these features had a positive effect on model accuracy. Moreover, Topcu and Saridemir [6] found that ANN had better predictive accuracy than of the RAC compressive and splitting tensile strengths than fuzzy logic. The ANN model demonstrated to be a powerful tool to determine the mechanical properties of RAC, achieving a coefficient of determination of 0.9984 and 0.9979 for the prediction of compressive strength and splitting tensile strength, respectively. Dantas et al. [23] gathered the largest dataset and used ANN to develop an equation to describe the compressive strength of RAC. Their model included 17 input features, from which, the ratio of recycled concrete, absorption rate of fine recycled aggregate, content of dry aggregate, and finesses modulus of aggregates were the parameters with highest effect on the compressive strength of RAC. The reported accuracy for the training and testing sets were 0.928, and 0.971, respectively.

In summary, Khademi et al. [15], Naderpour et al. [1], Deng et al. [16], Deshpande et al. [11], Gholampour et al. [22], Duan et al. [2], Topcu and Saridemir [6], and Dantas et al. [23] used 257, 139, 74, 257, 650, 168, 210, and 1178 data points, respectively, to predict the compressive strength of RAC. In addition to the quality and size of existing datasets, the advent of new and more powerful ML algorithms has stimulated researchers to explore the ability of state-of-the-art methods to enhance the accuracy and robustness of predictive models. Among various ML techniques to predict the compressive strength of RAC, artificial neural networks (ANNs) and fuzzy logic have been the most widely applied methods (Table 1).

Table 1. Studies on using ML techniques for prediction of RAC compressive strength.

Machine Learning Technique	No. of Samples	Ref.
Artificial neural networks, adaptive neuro-fuzzy inference system and multiple linear regression	257	[15]
Artificial neural networks	168	[2]
Artificial neural networks, model tree and non-linear regression model	257	[11]
Artificial neural networks and fuzzy logic	210	[6]
Convolutional neural networks	74	[16]
Artificial neural networks	139	[1]
Artificial neural networks	1178	[23]
Multivariate adaptive regression splines, M5 model tree and least support vector regression	650	[22]

1.2. Research Significance

As elaborated on above, there have been various studies on the application of traditional ML techniques to predict the compressive strength of RAC. The present study aims at creating a large and more comprehensive dataset and deploy it with state-of-the-art ML techniques that have not yet been explored for RAC in the open literature. The models presented herein are executed using the Python programming language. Therefore, to utilize these models, the user can simply apply the development steps along with hyperparameters reported in this study. Furthermore, the compressive strength predictive tools developed in this study are further complemented with optimization of the mixture proportions using a coupled PSO-GBRT model. The proposed mixture proportions can be used as a reference guideline for designing eco-friendlier and more economical RAC mixtures in practice.

2. Machine Learning Basis

ML refers to a computer's capacity of analyzing data and learning complex patterns within the data without being rigorously programmed [24]. Depending on the nature of the data, ML algorithms are categorized into supervised learning, unsupervised learning, and reinforcement learning [25]. Supervised learning aims at capturing underlying patterns in the data with known outputs. Depending on the type of the output, it can be further categorized as classification for discrete outputs, and regression for continuous outputs. Unsupervised learning, on the other hand, is associated with the data with unknown outputs and thus, clusters the data by finding relationships within the observations [26]. The third type of machine learning, reinforcement learning, bridges the gap between supervised and unsupervised learning since it clusters similar data given the correct answers [27]. Three powerful ML models were developed herein to forecast the compressive strength of RAC: GP, recurrent neural networks (RNNs), and GBRT. The three algorithms have different approaches for data analysis. Whilst GBRT is an ensemble of decision trees, GP uses the Gaussian distribution, and finally, RNNs are an advanced type of neural networks. The diverse nature of these algorithms is considered to explore the robustness of ML algorithms. The fundamentals of GP, RNNs, and GBRT are discussed below.

2.1. Gaussian Processes

GP are stochastic processes that generalize the Gaussian probability distribution [28]. In contrast to single- or multi-variable probability distribution in which a scalar or a vector is mapped, a process describes the properties of functions [29]. Therefore, a GP is defined as a probability distribution of functions, $P(f)$, where $P(f)$ has a Gaussian distribution [30]. GPs are parametrized with mean and covariance by analogy with Gaussian distribution, whereas mean and covariance for GP are functions [31]. The purpose of training a supervised learning algorithm using the available training dataset is to develop a model capable of predicting the output for unseen input data. There are two common approaches to determine the appropriate function that fits a set of data with promising accuracy [29]. In the first approach, the model is generated by considering only certain types of functions, e.g., exponential functions [29]. However, the prediction accuracy of such models strongly depends on the performance of the given functions. Conversely, the second approach considers

pre-assigned probabilities of several types of functions such that higher probability is assigned to those that are more likely to predict with higher accuracy [32]. The complexity of the first approach is limited to the selected functions. By contrast, the second approach is computationally more costly since there is an infinite number of possible functions to consider [29]. GPs are based on the second approach. The probabilistic formulation of GPs gives rise to a phenomenon called computational tractability in which the properties of functions are inferred even when some of the functions are ignored [33].

2.2. Recurrent Neural Networks

DL models are multiple-level computational algorithms able to learn complex underlying structures within a database [34]. DL models have proven successful in diverse applications such as image recognition, language understanding, and deoxyribonucleic acid (DNA) biological processes prediction [34]. However, applications of recent DL algorithms in civil engineering, including convolutional neural networks (CNNs) and RNNs have been more common in structural health monitoring and crack detection owing to the large data sets available in these fields [35,36]. CNNs and RNNs are among the most popular DL algorithms. In the present study, a novel type of RNN is deployed to predict the compressive strength of RAC.

RNN is a class of neural networks with an internal loop that allows the algorithm to keep memories from past information, commonly referred to as hidden state [37,38]. In RNNs, the output of a certain step, t , is used as input for the next step, $t+1$, emphasizing that every single step is based on the previous one, a process referred to as long-term dependencies; see Figure 1 [37]. Simple RNNs have a limitation regarding the contribution of earlier steps to the later ones, known as vanishing gradients [37]. Two main variants of layers have been proposed for RNN to overcome vanishing gradients: long-short-term memory (LSTM) and gated recurrent unit (GRU) [38]. The main difference of these RNN algorithms is the inclusion of gates for computing data. For example, LSTM layers incorporate a third gate, named the forget gate, in addition to the input and output gates in the simple RNN [37]. The forget gate maintains information and includes it in a non-consecutive step [38]. Conversely, GRU layers have only two types of gates: reset gate and update gate. In the reset gate, the previous information is combined with the most recent information, whereas in the update gate, it is decided how much information is to be passed to the following step. Figure 2 displays the structure of the first GRU layer used in this study [37]. Like LSTMs, GRUs are not affected by vanishing gradients. Nonetheless, GRU is considered a more efficient algorithm due to its simpler structure and formulation [27]. The formulation of GRU is summarized in the following:

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (1)$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (2)$$

$$\bar{h}_t = \text{ReLU}(W_h x_t + U_h (r_t \times h_{t-1})) \quad (3)$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \bar{h}_t \quad (4)$$

where r_t and z_t are the reset and update gate, respectively, \bar{h}_t is the candidate output, and h_t is the corresponding output of the cell for the time step t . Accordingly, W_r , W_z , W_h , U_r , U_z , and U_h are the weight matrices that operate the input vector x_t and the previous state h_{t-1} , and ReLU is the rectified linear unit activation function [39,40].

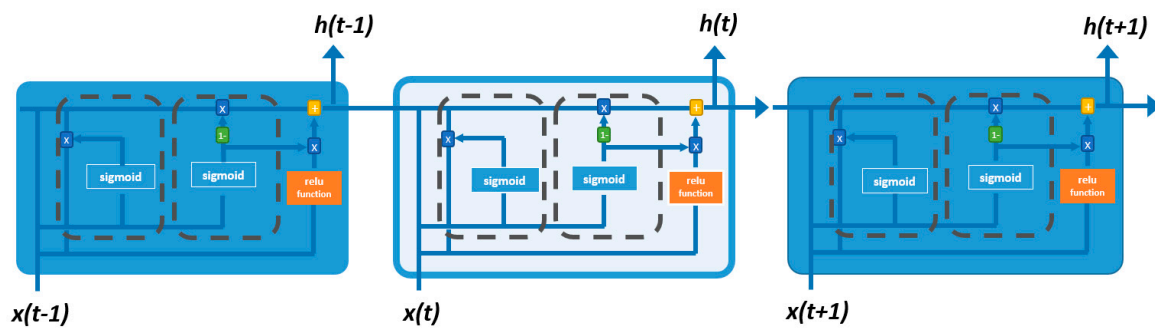


Figure 1. RNN structure using one GRU hidden layer.

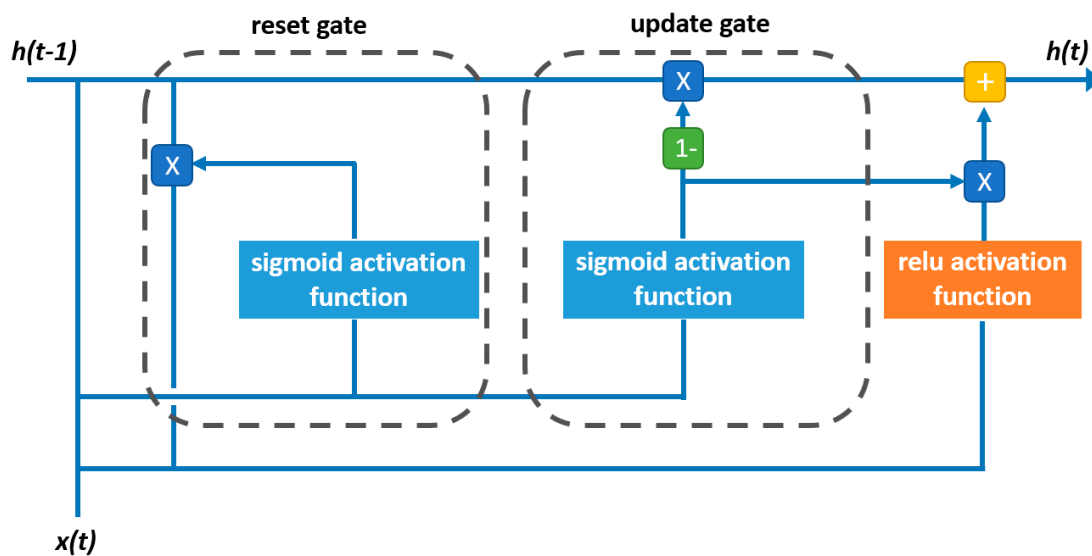


Figure 2. GRU hidden state computation, corresponding to the first layer of the developed deep learning model.

2.3. Gradient Boosting Decision Trees

GBRT algorithm integrates multiple weak learners using a boosting approach in which additional trees are appended in sequence without model parameters being changed. The objective of the gradient boosting is to find the function $\mathcal{F}(X)$ that minimizes the loss function $\mathcal{L}(\mathcal{F}(X), \mathcal{Y})$ (e.g., mean squared error or mean absolute error) using a given dataset, $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ [41,42]. The predictions of the GBRT model, y_t for a given input data can be expressed as:

$$y_t = \mathcal{F}_m(X_t) = \sum_{m=1}^M \mathcal{h}_m(x_t) \tag{5}$$

where the \mathcal{h}_m are referred to as weak learners. The constant M represents the number of weak learners which is known as the `n_estimators` hyperparameter. The loss function represents to what extent the predicted value is close to the output in the dataset using a specific metric. GBRT approaches the best function using the weighting of weak learner models, $\mathcal{h}(x_t)$, which is the basic decision tree fit by the input variables and the negative gradient of the last model's loss function. GBRT develops the model in a greedy manner considering a constant initial function $F_0(X)$ as follows [41–44]:

$$F_0(X) = \operatorname{argmin} \sum_{t=1}^N \mathcal{L}(y_t, \gamma) \tag{6}$$

$$\mathcal{F}_m(\mathcal{X}) = \mathcal{F}_{m-1}(\mathcal{X}) + \gamma_m \mathcal{h}_m(x) \quad (7)$$

where $\mathcal{h}_m(x)$ is the m th regression tree and γ_m is its weighting coefficient, also called learning rate. In a GBRT model, the number of trees, the learning rate, and the max depth of the tree are amongst the most essential hyperparameters that noticeably affect the predictive performance of the model. Larger number of trees increases the prediction accuracy of the model; however, excessive trees could result in an over-fitted model with lack of generalization for new unseen data. On the other hand, the learning rate controls the contribution of each tree to the predictions, while the max depth indicates the complexity of each tree. Immoderate values of such hyperparameters could bring about either over-fitted or erroneous models [41–44]. Other parameters of the GBRT model, such as subsample, maximum number of features, etc., also have noticeable effects on the model output and should be considered. Hence, tuning the GBRT hyperparameters is essential to propound robust and reliable performance.

3. Dataset Creation and Model Development

3.1. Data Collection and Preprocessing

The experimental data used in this study was collected from 55 peer-reviewed publications (Table 2). The collected data consists of 1134 RAC mixture design examples, with nine input features and one output. Statistical characteristics of the dataset are given in Table 3. Figure 3 illustrates the Pearson correlation coefficient between different attributes of the dataset. The Pearson correlation coefficient is an indicator of linear dependencies within two random variables; i.e., a coefficient of correlation close to one within two variables indicates that an increase in one of those variables will result in a proportional increment of the other [45]. Accordingly, the w/c ratio and superplasticizer dosage were the features with the highest correlation to the compressive strength. Conversely, aggregates (sand, natural gravel and RCA) did not have significant linear correlation to the compressive strength. Furthermore, since gravel is an ingredient replaced by recycled coarse aggregate (RCA), there was a high correlation between these two features.

Table 2. Sources of experimental data used in this study.

Authors	No. of Samples	Ref.	Authors	No. of Samples	Ref.
M. C. Limbachiya et al.	12	[46]	S. Manzi et al.	10	[47]
A. Ajdukiewicz and A. Kliszczewicz	117	[48]	A. B. Ajdukiewicz & A. T. Kliszczewicz	16	[49]
J. M. V. Gómez-Soberón	15	[50]	Y.-N. Sheen et al.	27	[51]
Y. H. Lin et al.	24	[52]	C. Thomas et al.	72	[53]
C. S. Poon et al.	36	[9]	V. A. Ulloa et al.	18	[54]
D. Matias et al.	9	[55]	W. Z. Taffese	10	[56]
M. Etxeberria et al.	4	[57]	G. Andreu and E. Miren	30	[58]
M. Etxeberria et al.	12	[59]	M. G. Beltrán et al.	9	[60]
S. C. Kou et al.	40	[61]	M. G. Beltrán et al.	8	[62]
C. S. Poon et al.	8	[63]	Ö. Çakır and Ö. Ö. Sofyanlı	27	[64]
K. Rahal	70	[65]	J. A. Carneiro et al.	2	[66]
R. Sato et al.	11	[67]	H. Dilbas et al.	12	[68]
M. Casuccio et al.	9	[69]	Z. H. Duan and C. S. Poon	26	[8]
S. C. Kou et al.	24	[70]	P. Folino and H. Xargay	4	[71]
K.-H. Yang et al.	42	[72]	F. López Gayarre et al.	14	[73]
A. Domingo-Cabo et al.	8	[74]	C. Medina et al.	16	[75]
V. Corinaldesi	10	[76]	D. Pedro et al.	18	[7]
R. Kumutha and K. Vijai	12	[77]	M. Pepe et al.	15	[78]
M. Malešev et al.	9	[79]	G. Wardeh et al.	16	[80]
G. F. Belén et al.	16	[81]	Y. Haitao and T. Shizhu	20	[82]
G. Fathifazl et al.	6	[83]	V. W. Y. Tam et al.	24	[84]
M. Chakradhara Rao et al.	16	[85]	A. S. Abdel-Hay	4	[86]
R. Somna	18	[87]	C. Zheng et al.	36	[88]
A. Abd Elhakam et al.	30	[89]	M. C. S. Nepomuceno et al.	15	[90]
A. Barbudo et al.	36	[91]	N. Mohammed et al.	12	[92]
L. Butler et al.	8	[93]	C. Thomas et al.	23	[94]
S. Ismail and M. Ramli	12	[95]			
K. Kim et al.	18	[97]	K. H. Younis and K. Pilakoutas	18	[96]

Table 3. Statistical characteristics of the dataset.

Input Feature	Units	Min.	Max.	Mean	Standard Deviation
Water-to-cement ratio	-	0.24	1.02	0.49	0.12
Cement content	kg/m ³	210.00	650.00	387.60	71.36
Sand content	kg/m ³	419.52	1010.00	691.71	131.65
Recycled aggregate content	kg/m ³	0.00	1358.00	527.83	444.75
Gavel content	kg/m ³	0.00	1524.00	542.94	470.19
Superplasticizer	kg/m ³	0.00	45.00	2.63	4.53
Silica fume content	kg/m ³	0.00	50.00	3.47	11.60
Age	Days	2.00	365.00	44.57	70.69
Specimen type	Type	1.00	5.00	2.79	1.15
Output	Units	Min.	Max.	Mean	Standard Deviation
Compressive strength	MPa	4.30	108.51	43.57	17.72

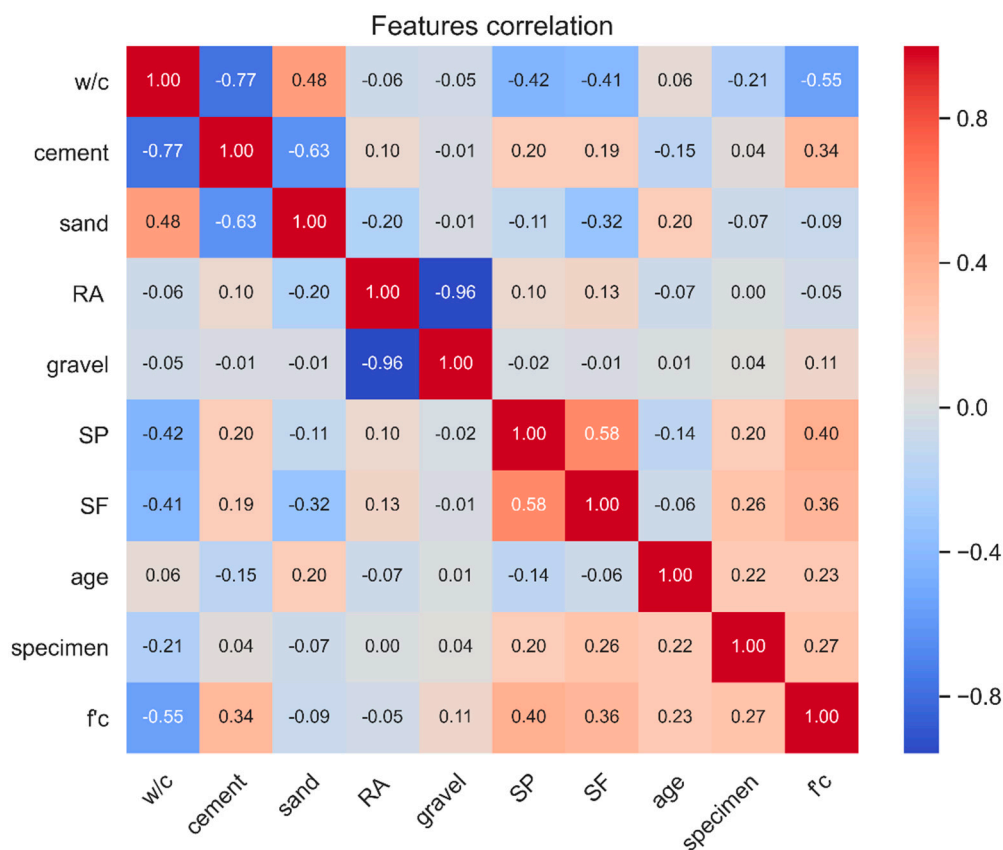


Figure 3. Pearson correlation coefficient for the dataset attributes.

Feature normalization is commonly applied prior to modeling. Although normalization is not required for all ML algorithms, it has been proven to improve the model performance [27]. Linear transformation and statistical standardization are among the most popular normalization techniques [98]. In the linear transformation, values are ranged within the domain of [0, 1], whereas in the statistical standardization, the mean and the standard deviation values of the data are set equal to 0 and 1, respectively [98]. In this study, statistical standardization was used prior to GP, GBRT and RNNs modeling. The data was then randomly divided into training and testing sets using 70% (793 samples) for training and the remaining (341 samples) for model testing.

A common practice to assess the performance of ML models is to divide the entire dataset into three different subsets: training, validation and testing. Whilst the learning process is accomplished

with the training set, the validation set is used to track the performance of the model, while the testing set serves to assess the extrapolation capabilities of the model by performing it over unseen samples that are not known to the model [27]. However, the partition of data into three subsets leads to a reduction of the training samples, which consequently can result in an insufficiently trained model [27]. Thus, cross-validation is a common technique to prevent the over reduction of the training set, especially for small datasets. There are several techniques to perform cross-validation, most of which consist in leaving out random data to validate the model [99]. In this study, K -fold cross-validation was used. K -fold cross-validation is a resampling method that splits the data into k number of subsets and keeps one subset for validation, while the other $k - 1$ subsets are used for training [100]. The 5-fold cross-validation employed for hyperparameter selection in this study is schematically depicted in Figure 4.

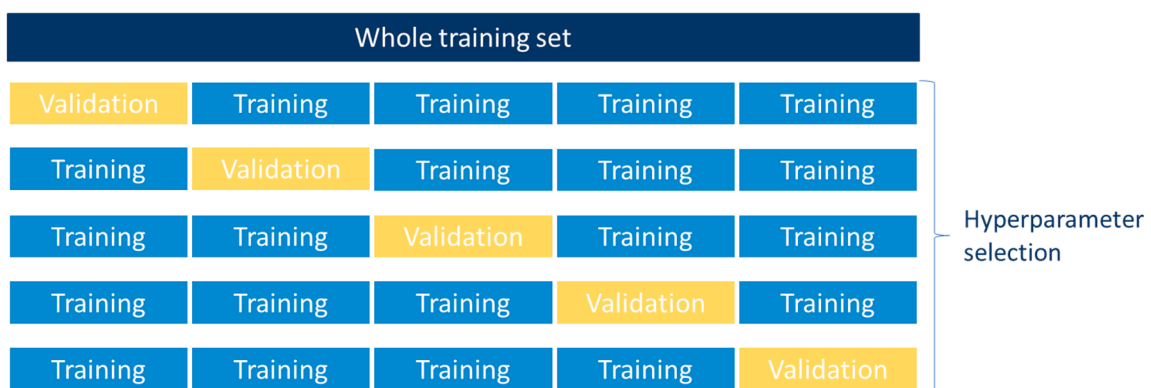


Figure 4. 5-fold cross validation for hyperparameter tuning.

3.2. Hyperparameter Tuning

Hyperparameter tuning is a crucial step in developing robust ML models. Tuning of ML model mitigates over-fitting and thus, enhances the versatility of the model to unseen data [101]. The selection of optimum hyperparameters is also a determinant factor in increasing the model accuracy [102]. Aiming to avoid manual tuning, there have been different approaches proposed to automatize the selection of hyperparameters, such as grid search and random search hyperparameter optimization [103]. These approaches are distinct from each other by the domain of the potential values considered in the search attempt. Whilst grid search explores all possible values in a pre-defined domain for hyperparameters, random search algorithms select the different hyperparameter values in a random manner for a specific number of iterations [103]. A randomized search procedure with 5-fold cross-validation was used herein for exploring possible values of hyperparameters using the Scikit-learn package in Python [104].

3.3. Model Development

3.3.1. GP Model

GP is a non-parametric model [29] and thus, the selection of hyperparameters is less challenging, especially compared to DL models. The hyperparameters of GP models are those required for the kernel function. Therefore, the kernel function, also known as the covariance function, is key to creating robust GP models [29]. In this study, a linear combination of several default kernel functions was implemented as defined in Equation (8). This kernel function includes the periodic kernel, matern kernel, and dot-product kernel. It is worth mentioning that all available kernels, such as the

periodic kernel, the rational quadratic kernel, white kernel, matern kernel, and dot-product kernel, were tested for tuning the GP model.

$$k(x_i, x_j) = \sigma_0^2 + x_i \cdot x_j + 2^2 * \exp\left(-\frac{2 \sin^2\left(\pi \frac{d(x_i, x_j)}{p}\right)}{l_1^2}\right) + \frac{1}{\Gamma(\nu)2^{\nu-1}} \left(\frac{\sqrt{2\nu}}{l_2} d(x_i, x_j)\right)^\nu K_\nu\left(\frac{\sqrt{2\nu}}{l_2} d(x_i, x_j)\right) \quad (8)$$

According to the former equation, parameters associated with the considered kernels were tuned as the hyperparameters of the GP model, including the length scale 1 (l_1) and periodicity (p) corresponding to the periodic kernel; ν and length scale 2 (l_2) corresponding to the matern kernel; and σ_0 of the dot-product kernel. The optimizing of hyperparameters was carried out using 5-fold cross-validation (CV) as described earlier. The tuned values of the hyperparameters are listed in Table 4. Scikit-learn library in Python was employed for tuning and executing the GP model [104].

Table 4. Hyperparameters for Gaussian processes model.

Hyperparameter	Assigned Value
Length scale 1, l_1	0.6
Periodicity, p	16.0
Sigma naught, σ_0	1.9
Length scale 2, l_2	1
Nu, ν	0.5

3.3.2. RNN Model

The developed architecture of the RNN model consists of 3-GRU layers and 1 dense layer with 239, 238, 217, and 1 hidden neuron, respectively. In the first layer, ReLU activation function and sigmoid recurrent activation function were utilized (Figure 2). In the second layer, the activation function and recurrent activation function were the sigmoid and ReLU, respectively. In the third layer, the scaled exponential linear unit (SELU) and softsign were used as activation and recurrent activation functions, respectively. For the dense layer, only the softplus activation function was used. Moreover, the kernel initializer and recurrent initializer were tuned for GRU layers. The kernel initializer was fixed as random uniform for the first and second layers, whereas a constant initializer was used for the third layer. The recurrent initializer was set as constant for the first layer, and zero recurrent initializer for the second and third layers. Mean squared error (MSE) was used as the model loss function, whereas the Adam optimization algorithm was employed as the model optimizer, with a learning rate of 0.0002. Ultimately, the number of epochs and batch size were set to 360 and 11, respectively. According to Whang and Matsukawa [105], the performance of GRU models was improved when batch normalization was applied. Batch normalization mitigates the so-called internal covariate shift [105]. Internal covariate shift is a frequent problem in the training step of deep neural networks in which the distribution of inputs at each layer is changed and thus, finer tuning for models along with smaller learning rates are required [106]. Hence, batch normalization was implemented in the developed RNN model because it improves the performance of GRU networks [105]. Momentum and epsilon are the parameters associated with batch normalization. The optimum momentum and epsilon were found to be 0.95 and 0.0001, respectively. Table 5 summarizes the tuned hyperparameters of the RNN model. The hyperparameter selection for the DL models was performed using a randomized search approach along with 5-fold CV. Keras API and Scikit-learn packages in Python were utilized for building and tuning the RNN model [104,107].

Table 5. Hyperparameters for deep learning model.

Layer	Units	Activation Function	Recurrent Activation Function	Kernel_INITIALIZER	Recurrent_INITIALIZER
Gated recurrent unit	239	ReLU	Sigmoid	Random Uniform	Constant
Gated recurrent unit	238	Sigmoid	ReLU	Random Uniform	Zeros
Gated recurrent unit	217	SELU	Softsign	Constant	Zeros
Dense	1	Softplus	-	-	-

3.3.3. GBRT Model

GBRT has multiple hyperparameters that need tuning prior to model training. In the current study, a randomized search procedure alongside 5-fold CV was used to obtain optimum hyperparameters of the GBRT model. Generally, *n_estimators* and *learning_rate*, which indicate the number of the weak learners in the model and the weighting of each estimator, respectively, are the most influential hyperparameters of the GBRT model that are essential to be tuned. Additionally, *max_depth*, *max_features*, and *subsample* can greatly affect the prediction performance of the GBRT model [108]. Table 6 presents the tuned values of the seven hyperparameters considered. The mean absolute error (MAE) was monitored as the statistical error to achieve optimum hyperparameters yielding highest accuracy while mitigating over-fitting. The Scikit-learn package was implemented to perform GBRT modeling and tuning [104].

Table 6. Hyperparameters for the GBRT model.

Hyperparameter	Number of Estimators	Learning Rate	Min Samples Split	Min Samples Leaf	Max Depth	Max Features	Subsample
Value	315	0.44	33	17	5	7	0.98

3.4. RAC Mixture Optimization

This section presents the framework adopted for optimizing the mixture design of RAC using the ML model with best predictive performance. The objective of the optimization is to propose the most economic mixture proportions of RAC considering different classes of compressive strength. The PSO algorithm, which is a metaheuristic method that mimics the social interactions of birds or insects (particles) in the search of an optimal solution, was adopted [109]. The particles modify their position in every iteration based on the individual velocity vector of each particle, which in turn is dependent on both the best-found particle and swarm positions [110]. The PSO minimizes an objective function while limiting the domain of the solution. According to the optimization procedure proposed by Yeh [19], the function that is to be optimized herein is the cost to produce a batch of RAC as defined in Equation (9). The considered unit costs, which are averages of values retrieved from multiple material suppliers across Canada, are presented in Table 7. These values can easily be replaced by costs corresponding to other locations. The unit cost of RCA was considered equal to that of NA as recommended in ref. [111].

$$P = C_1 I_1 + C_2 I_2 + \dots + C_i I_i \quad (9)$$

where C_i represents the unit cost of the i th ingredient of the mixture and I_i is the i th ingredient dosage in kg/m^3 . To limit the domain of the solution, two bounder vectors were defined: upper limit and lower limit. The bounder vectors (Table 8) were strategically defined based on a real experiment from the dataset with certain compressive strength to draw meaningful comparison and thus, better validate the performance of the algorithm. In other words, for sand, cement, and water, the upper and lower bounder limits were defined in average 20% up and down the values given for the base mixture. To promote the use of recycled aggregate, the assigned values to the lower and upper bounder vectors were kept high, and the corresponding values for gravel were maintained low. Also, due to the high cost of superplasticizer, the assigned values for the bounder vectors were kept as low as

possible. The 28-day compressive strength of a standard 15 × 30 cm cylinder specimen was considered for the sake of comparison. The results of the optimized mixture proportions are given in Table 9. The optimized mixture was subsequently tested using the GBRT (being the best predictive model in this study) and compared to the real concrete sample extracted from the dataset to ensure the required compressive strength criteria, as shown in Table 10.

Table 7. Unit price of ingredients of concrete mixtures.

Ingredient	Units	Currency	Unit Price
Water	\$/kg	Canadian dollar	0.004
Cement	\$/kg	Canadian dollar	0.43
Sand	\$/kg	Canadian dollar	0.28
Recycled coarse aggregate	\$/kg	Canadian dollar	0.20
Gavel	\$/kg	Canadian dollar	0.20
Superplasticizer	\$/kg	Canadian dollar	71.07
Silica fume	\$/kg	Canadian dollar	2.85

Table 8. Bounder vectors for mixture optimization.

Input Feature	Unit	25 MPa		30 MPa		35 MPa		40 MPa	
		Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit
Water	kg/m ³	350	200	350	190	230	160	230	160
Cement	kg/m ³	424	290	424	292	424	323	424	280
Sand	kg/m ³	942	650	942	650	942	720	942	750
RCA ^a	kg/m ³	1080	700	1080	750	1080	550	900	750
Gavel	kg/m ³	511	50	511	50	511	100	750	220
SP ^b	kg/m ³	0	0	0	0	0	0	2	0.9
Age	Days	28	28	28	28	28	28	28	28
Specimen	Type	1	1	1	1	1	1	1	1

^a recycled coarse aggregate, ^b superplasticizer.

Table 9. Optimized mixtures.

Optimized Mix	Water	Cement	Sand	RCA ^a	Gravel	SP ^b	Age	ST ^c
	(kg/m ³)	(kg/m ³)	(kg/m ³)	(kg/m ³)	(kg/m ³)	(kg/m ³)	Days	Type
25 MPa	246.46	296.62	701.67	711.90	155.23	0.00	28	1
30 MPa	239.56	298.52	701.67	760.33	155.23	0.00	28	1
35 MPa	181.68	327.99	759.29	566.60	193.82	0.00	28	1
40 MPa	178.83	310.45	767.23	768.92	313.78	1.23	28	1

^a recycled coarse aggregate, ^b superplasticizer, ^c specimen type.

Table 10. Comparison of optimized mixture with base mixture.

Input Feature	Units	25 MPa		30 MPa		35 MPa		40 MPa	
		Base	Opt.	Base	Opt.	Base	Opt.	Base	Opt.
Water	kg/m ³	234.10	246.46	190.00	239.56	175.00	181.68	187.00	178.83
Cement	kg/m ³	390.16	296.62	380.00	298.52	350.00	327.99	311.00	310.45
Sand	kg/m ³	702.30	701.67	637.00	701.67	730.00	759.29	840.00	767.23
RCA ^a	kg/m ³	1053.45	711.90	1123.00	760.33	989.00	566.60	0.00	768.92
Gravel	kg/m ³	0.00	155.23	0.00	155.23	0.00	193.82	935.00	313.78
SP ^b	kg/m ³	0.00	0.00	0.00	0.00	1.68	0.00	1.56	1.23
Age	Days	28	28	28	28	28	28	28	28
ST ^c	Type	1	1	1	1	1	1	1	1
F'c	MPa	25.3	25.5	30.1	29.6	36.0	35.5	40.0	39.9
Price	CAD	577.21	499.44	568.49	510.02	673.94	507.12	668.66	654.35

^a recycled coarse aggregate; ^b superplasticizer; ^d specimen type.

4. Results, Discussion, and Recommendations

This section presents the results of ML modeling of RAC. The three different models outlined earlier were implemented and their prediction performance is discussed herein. Purposefully, the root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination R^2 were used for assessing the performance of each model. Moreover, the best acquired ML model was employed to perform RAC mixture design optimization for different ranges of 28-day compressive strength. The optimization results along with mixture proportion recommendations are discussed below.

4.1. Prediction Performance of ML Models

GP, GRU, and GBRT models were trained using 793 data examples and tested with the remaining 341 data. The final tuned models were executed over five different seed numbers of data split to assess the robustness of the models trained with randomized split of the data for training and the testing sets. The predictive performance of the GP model for the five random seed numbers is summarized in Table 11. The model predicted the output with average RMSE, MAE and R^2 of 7.087 MPa, 4.911 MPa, and 0.844, respectively for the test dataset. However, the model performance was greatly superior for the training dataset with average RMSE, MAE and R^2 of 0.735, 0.138, and 0.998, respectively. This trend can be further observed in the residual plot of the GP model shown in Figure 5. The residuals for the training data were less than 10 MPa, while they were as high as 40 MPa for some data points in the testing set. The actual versus predicted output of the GP model is illustrated in Figure 6.

The GRU model attained better performance compared to that of the GP model (see Table 12). The difference between the GRU statistical errors of train and test data were less than that of the GP model. The RMSE, MAE, and R^2 values for the test dataset were 6.502 MPa, 4.364 MPa, and 0.868, respectively, while the corresponding values were 3.183 MPa, 2.285 MPa, and 0.968, respectively, for the train dataset. This demonstrates more robust predictive performance along with higher accuracy compared to the GP model. The residuals of the predictions varied in a narrower range compared to that in the GP model, as depicted in Figure 7. The residuals for both testing and training datasets had similar normal distribution, indicating more robust predictive performance. Figure 8 shows the actual versus predicted compressive strength of the test data for the GRU model.

Table 11. Measured performance of Gaussian process model.

Random Seed and Global Performance	Set	RMSE ^b	MAE ^c	R^2
RS ^a = 59	Test	7.468	5.157	0.827
	Train	0.556	0.111	0.999
RS ^a = 1718	Test	7.589	5.197	0.834
	Train	0.789	0.144	0.998
RS ^a = 1009	Test	6.582	4.762	0.854
	Train	0.595	0.103	0.999
RS ^a = 3097	Test	7.492	4.875	0.841
	Train	0.680	0.135	0.998
RS ^a = 7	Test	6.305	4.566	0.862
	Train	1.055	0.197	0.997
Average	Test	7.087	4.911	0.844
	Train	0.735	0.138	0.998
Standard Deviation	Test	0.597	0.267	0.014
	Train	0.200	0.037	0.001

^a random seed; ^b root mean squared error; ^c mean absolute error.

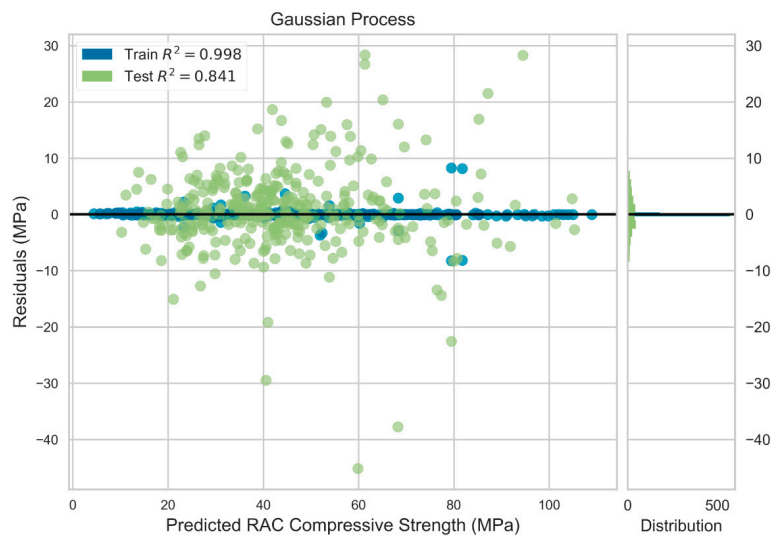


Figure 5. Residuals plot for Gaussian process model.

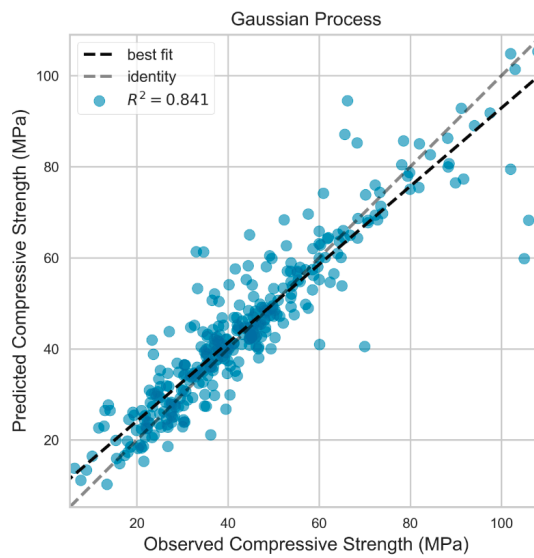


Figure 6. Actual vs. predicted values for testing set in Gaussian process model.

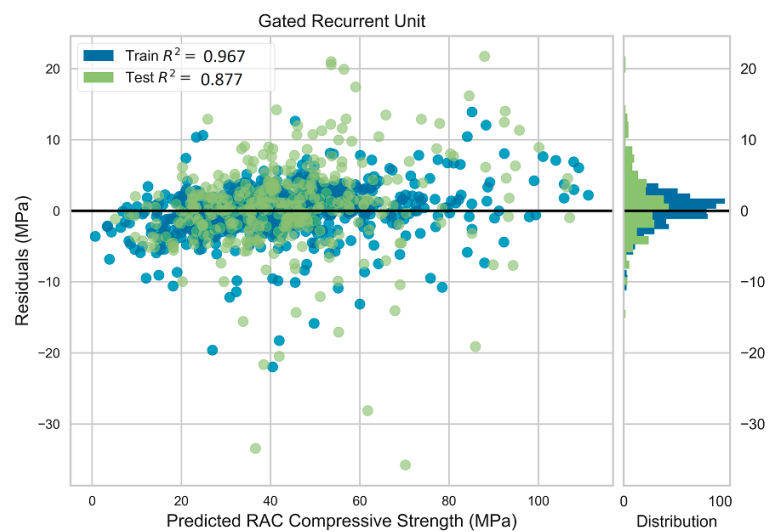
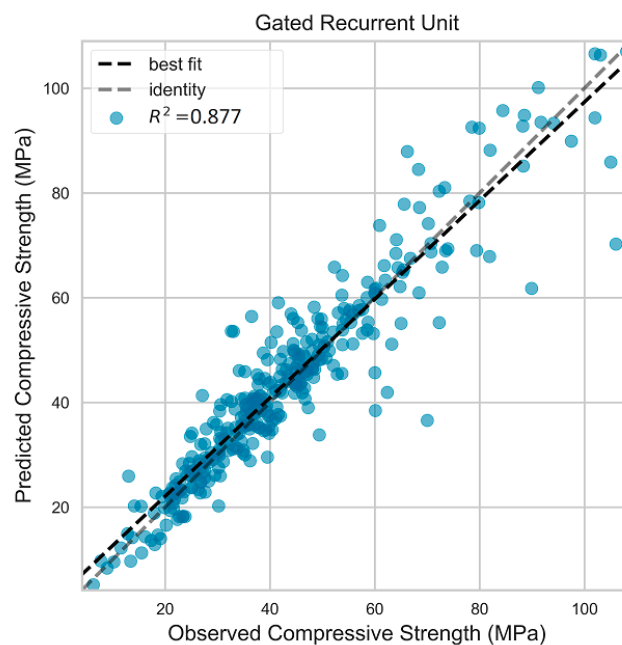


Figure 7. Residuals plot for deep learning model.

Table 12. Measured performance of deep learning model.

Random Seed and Global Performance	Set	RMSE ^b	MAE ^c	R ²
RS ^a = 59	Test	7.298	4.663	0.835
	Train	3.064	2.16	0.97
RS ^a = 1718	Test	6.927	4.567	0.861
	Train	3.140	2.274	0.968
RS ^a = 1009	Test	5.778	4.106	0.888
	Train	3.172	2.316	0.969
RS ^a = 3097	Test	6.589	4.312	0.877
	Train	3.144	2.251	0.967
RS ^a = 7	Test	5.918	4.172	0.878
	Train	3.394	2.422	0.965
Average	Test	6.502	4.364	0.868
	Train	3.183	2.285	0.968
Standard Deviation	Test	0.649	0.243	0.021
	Train	0.125	0.096	0.002

^a random seed; ^b root mean squared error; ^c mean absolute error.

**Figure 8.** Actual vs. predicted values for testing set in deep learning model.

The GBRT model scored superior predictive execution, as indicated in Table 13, with lowest RMSE and MAE values for the test dataset, along with the highest coefficient of determination compared to that of the GP and GRU models. RMSE and MAE were 5.074 and 3.396 MPa, respectively for the GBRT model. Figure 9 depicts the residuals of the predicted compressive strength for the training and testing datasets of the GBRT model. It can be observed that the model captured the trend in the data and demonstrated powerful performance on both the train and test datasets. The model achieved R² value of 0.997 and 0.925 for training and testing data, respectively. Furthermore, less scatter of the GBRT predicted values of the test dataset was accomplished compared to the GRU and GP models. The actual versus GBRT predicted compressive strength of the test data is displayed in Figure 10.

Table 13. Measured performance of GBRT model.

Random Seed and Global Performance	Set	RMSE ^b	MAE ^c	R ²
RS ^a = 59	Test	5.124	3.354	0.918
	Train	1.102	0.743	0.996
RS ^a = 1718	Test	5.359	3.698	0.917
	Train	1.008	0.710	0.996
RS ^a = 1009	Test	4.640	3.196	0.927
	Train	0.965	0.683	0.997
RS ^a = 3097	Test	5.168	3.335	0.924
	Train	0.970	0.704	0.996
RS ^a = 7	Test	5.087	3.398	0.911
	Train	1.052	0.748	0.996
Mean	Test	5.076	3.396	0.919
	Train	1.019	0.718	0.996
Standard Deviation	Test	0.236	0.165	0.005
	Train	0.051	0.024	0.0003

^a random seed; ^b root mean squared error; ^c mean absolute error.

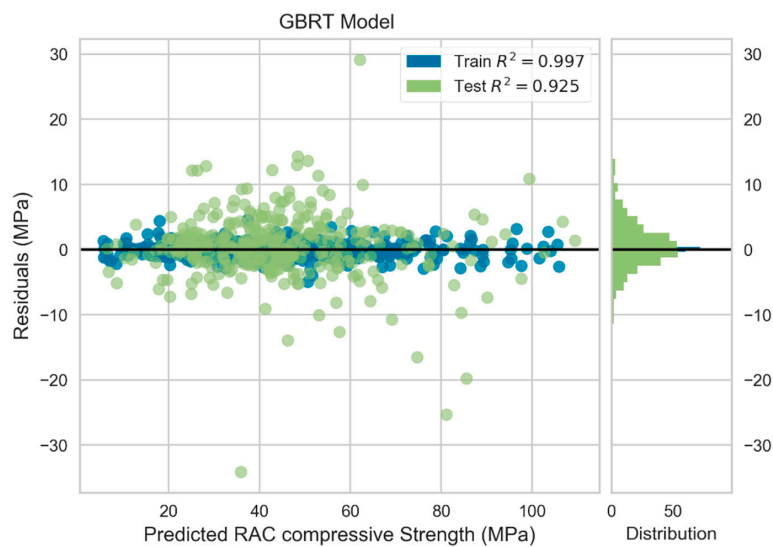


Figure 9. Residuals plot for GBRT model.

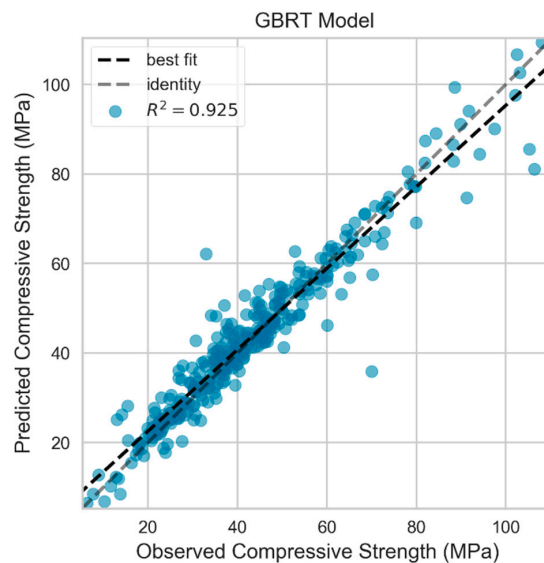


Figure 10. Actual vs. predicted values for testing set in GBRT model.

4.2. Comparison of Model Performance

Based on the results discussed above, all developed ML models could predict the compressive strength of RAC with reasonable accuracy. However, the GRU and GBRT models demonstrated higher generalization capacity since their prediction errors for training and testing datasets were highly analogous, in contrast to the GP model. The prediction accuracy for the training dataset in the GP model was very high, while it was quite low for the testing dataset. Thus, the GP model suffers from over-fitting and lack of generalization to new unseen data. Although DL models are recognized to be more accurate on large datasets, the finely tuned GRU model, despite the relatively small dataset, reached outstanding prediction performance with high generalization capacity.

Figure 11 illustrates the Taylor diagram of the GP, GRU and GBRT models using the RMSE, Pearson correlation and standard deviation of the predictions. The Taylor diagram suggests that the GBRT model had superior performance in terms of RMSE, whereas the GRU model provided predictions of the output with a highly correlated standard deviation to the actual observations. It is worth mentioning that the GBRT model required considerably shorter execution time for training compared to the GRU model. This comparison was performed using the same computer without mounting or connecting it to a hosted GPU. Ultimately, it was concluded that the GBRT model had the best performance and will be considered for the mixture optimization process.

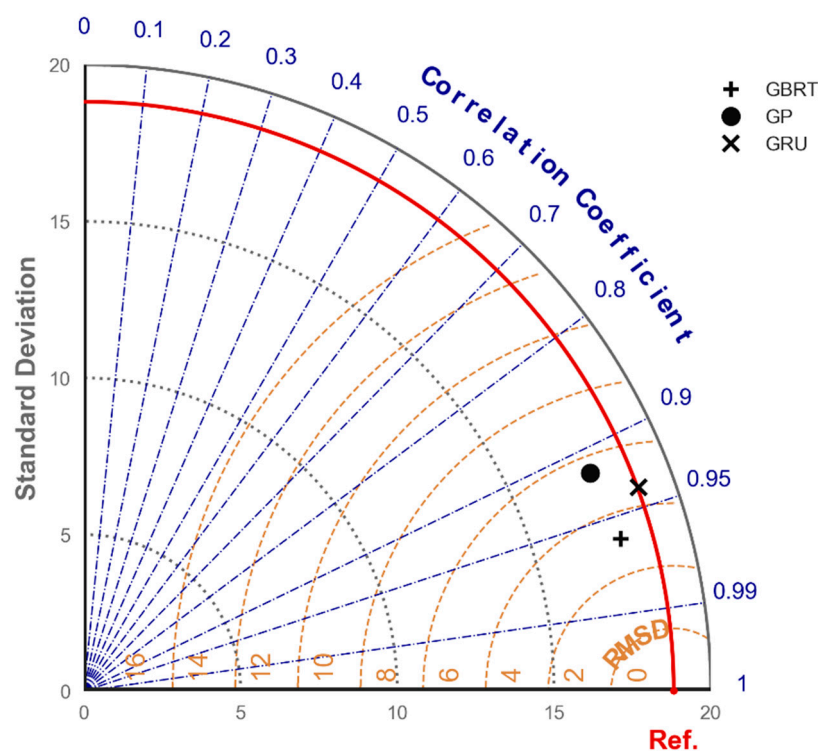


Figure 11. Taylor diagram comparing performance of three developed ML models.

4.3. Comparison with Previous Studies

A prime goal in ML is to create models that can accurately predict the output for new unseen data never presented to the model, i.e., achieving models that can generalize [38]. ML models generalize a phenomenon through learning the underlying principles within the training data. Hence, they are capable of generalizing when predicting sensible outputs from inputs different than those of the training dataset [27]. Testing the model on a high number of unseen data samples is the rational way to determine whether the model is generalizing or not, thus the importance of having large datasets [27]. The models proposed in the present study have demonstrated better generalization capability than

those in former studies. A major reason for this superior performance is that the test dataset used in this study has more data samples than the entire datasets used in developing previous models, including Khademi et al. [15], Duan et al. [2], Deshpande et al. [11], Topçu and M. Saridemir [6], Deng et al. [16], and Naderpour et al. [1] (see Table 14). It is important to mention that the Deng et al. [16] model was not included in Table 14 because they neither reported the coefficient of determination nor the root mean squared error. However, they reported the relative error percentage, which corresponded to 6.63, 4.35, and 3.65 for the back propagation neural network, support vector machine, and convolutional neural network, respectively.

Table 14. Comparison of statistical measurements with previous studies.

Machine Learning Technique	R^2	RMSE	No. of Samples	Ref.
Multiple linear regression	0.609	9.975		
Artificial neural networks	0.919	4.446	257	[15]
Adaptive neuro-fuzzy inference system	0.908	5.045		
Artificial neural networks	0.995	3.6804	168	[2]
Artificial neural networks	0.903	-		
Model tree	0.757	-	257	[11]
Non-linear regression model	0.740	-		
Artificial neural networks	0.998	2.395		
Fuzzy logic	0.996	3.866	210	[6]
Artificial neural networks	0.688	-	139	[1]
Artificial neural networks	0.971	-	1178	[23]
Multivariate adaptive regression splines	-	8.750		
M5 model tree	-	8.250	650	[22]
Least support vector regression	-	7.550		
Gradient Boosting ^a	0.919	5.076		
Deep Learning ^a	0.868	6.502	1134	-

^a model of the present study.

Table 14 shows the coefficient of determination and the root mean squared error of models in previous studies that predicted the compressive strength of RAC. It can be observed that models in the present study achieved better accuracy than that of Gholampour et al. [22] and Deshpande et al. [11] who used relatively large data samples. As expected, studies that reported smaller databases reached higher accuracy. For instance, Duan et al. [2] and Khademi et al. [15] used 168 and 257 samples, respectively. The reported accuracy was 0.995 for Duan et al. [2] and 0.919 for Khademi et al. [15], both studies using ANNs. This indicates that although higher number of samples might result in a better generalized model, the accuracy can decrease, and thus accuracy metrics alone may not be enough to assess predictive models. The ability to generalize predictions beyond a limited dataset is a more important feature. Also, several models which used smaller data sets than that in the present study, including Khademi et al. [15], Duan et al. [2], Deshpande et al. [11], Topçu and Saridemir [6], Deng et al. [16], and Naderpour et al. [1], have compromised generalization capability. Furthermore, in the case of Gholampour et al. [22], the authors decided to split the available data and create two different models to predict the compressive strength of those samples corresponding to cylindrical specimens and those corresponding to cube specimens. Conversely, the present study considered the specimen type as an input feature, resulting in higher accuracy. Generally, the present study along with Dantas et al. [23] used the highest number of data. However, Dantas et al. [23] reported a coefficient of determination higher for the testing set than that for the training set, 0.971 and 0.928, respectively. This is a sign that their model was not sufficiently trained, as suggested by Gulli and Pal [37].

4.4. RAC Mixture Proportioning and Optimization

A PSO was coupled with the GBRT model to optimize the mixture design and predict the compressive strength of RAC, such that the most economic mixture proportion is obtained for a given compressive strength class. The optimization was performed considering the unit costs of materials presented in Table 7. Not only does the optimization process reduce the higher unit cost ingredients, but it also reduces cement in the mixture, providing both economic benefit and sustainable mixture designs with less CO₂ emission. High upper limit of RCA was considered in the optimization to ensure maximum replacement of RCA as presented in Table 8. Although using higher portions of RCA may contradict with compressive strength requirements, the optimization was carried out to maintain highest possible recycle aggregate content along with the desired compressive strength class.

Table 9 presents the optimized mixture designs of RAC for different compressive strength classes as obtained by the PSO model. The mixture proportions were then used to predict the compressive strength using the GBRT model. Silica fume was not considered in the optimization process, and thus was set to zero when predicting the compressive strength with the GBRT model. Ultimately, considerable reduction of cost in all cases, especially for the lower compressive strength range, was achieved as outlined in Table 10. For instance, there was 25% reduction in the cost of the RAC mixture without affecting its compressive strength when the target compressive strength was 35 MPa. The optimization process demonstrated the outstanding capability of the PSO-GBRT model in capturing complex relationships within the data to select the best mixture proportions, while maintaining a similar water-to-cement ratio to that of the base mixture. This can be observed for instance when considering the 25 and 30 MPa compressive strength classes in which high water-to-cement ratio was proposed with high RCA content with high water absorption capacity, as observed in experimental studies [9].

5. Conclusions

The present study explores deploying state-of-the-art machine learning models to predict the compressive strength of RAC and optimize its mixture design. For this purpose, one of the largest existing experimental datasets, including 1134 mixture design examples and featuring 10 attributes was built from studies in the open literature. Three advanced machine learning models, including Gaussian processes (GP), deep learning (DL), and gradient boosting regression trees (GBRT) were tuned, trained, and tested using the dataset. To guarantee that the developed models were able to generalize the compressive strength of RAC, *K*-fold cross-validation was used during the tuning process. The results show that the three models successfully captured the underlying principles contributing to the compressive strength of RAC. Furthermore, the diverse nature of the algorithms used herein proves the robustness of ML algorithms for data analysis despite the complexity within the dataset. The comparison of the models' performance revealed that the GBRT and DL (recurrent neural network) models had superior predictive performance compared to the GP model in terms of different statistical metrics and performance indicators. Accordingly, the obtained coefficient of determination of the testing set for GBRT, DL, and GP was 0.919, 0.868, and 0.844, respectively. Furthermore, the GBRT model was coupled with PSO to create a hybrid model for optimizing the mixture design of RAC with various compressive strength classes. Accordingly, the GBRT-PSO hybrid model successfully proposed economic mixture designs that fulfill the compressive strength requirement, reduce cost, and mitigate the environmental footprint of concrete production. To further the high potential of the developed ML models, it is proposed to integrate supplementary cementitious materials, such as fly ash and blast furnace slag in the dataset, and to extend the models to also capture durability requirements of RAC in future work.

Author Contributions: Conceptualization I.N. and M.L.N.; data curation I.N.; formal analysis I.N. and A.M.; investigation I.N.; methodology I.N. and A.M.; project administration M.L.N.; resources M.L.N.; software I.N. and A.M.; supervision M.L.N.; validation I.N., A.M. and M.L.N.; visualization I.N. and A.M.; writing—original draft, I.N. and A.M.; writing—review & editing, M.L.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors acknowledge the scholarship support from Consejo Nacional de Ciencia y Tecnologia (CONACYT) to the first author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Naderpour, H.; Rafiean, A.H.; Fakharian, P. Compressive strength prediction of environmentally friendly concrete using artificial neural networks. *J. Build. Eng.* **2018**, *16*, 213–219. [[CrossRef](#)]
2. Duan, Z.H.; Kou, S.C.; Poon, C.S. Prediction of compressive strength of recycled aggregate concrete using artificial neural networks. *Constr. Build. Mater.* **2013**, *40*, 1200–1206. [[CrossRef](#)]
3. Gholampour, A.; Gandomi, A.H.; Ozbakkaloglu, T. New formulations for mechanical properties of recycled aggregate concrete using gene expression programming. *Constr. Build. Mater.* **2017**, *130*, 122–145. [[CrossRef](#)]
4. Yeheyis, M.; Hewage, K.; Alam, M.; Eskicioglu, C.; Sadiq, R. An overview of construction and demolition waste management in Canada: A lifecycle analysis approach to sustainability. *Clean Technol. Environ. Policy* **2012**, *15*, 81–91. [[CrossRef](#)]
5. González-Fontboa, B.; Martínez-Abella, F. Concretes with Aggregates from Demolition and Construction Waste and Silica Fume. Materials and Mechanical Properties. *Build. Environ.* **2008**, *43*, 429–437. [[CrossRef](#)]
6. Topçu, I.B.; Saridemir, M.; Saridemir, M. Prediction of mechanical properties of recycled aggregate concretes containing silica fume using artificial neural networks and fuzzy logic. *Comput. Mater. Sci.* **2008**, *42*, 74–82. [[CrossRef](#)]
7. Pedro, D.; De Brito, J.; Evangelista, L. Performance of concrete made with aggregates recycled from precasting industry waste: Influence of the crushing process. *Mater. Struct.* **2014**, *48*, 3965–3978. [[CrossRef](#)]
8. Duan, Z.H.; Poon, C.S. Properties of recycled aggregate concrete made with recycled aggregates with different amounts of old adhered mortars. *Mater. Des.* **2014**, *58*, 19–29. [[CrossRef](#)]
9. Poon, C.S.; Shui, Z.H.; Lam, L.; Fok, H.; Kou, S.C. Influence of moisture states of natural and recycled aggregates on the slump and compressive strength of concrete. *Cem. Concr. Res.* **2004**, *34*, 31–36. [[CrossRef](#)]
10. Silva, R.V.; de Brito, J.; Dhir, R.K. The influence of the use of recycled aggregates on the compressive strength of concrete: A review. *Eur. J. Environ. Civ. Eng.* **2014**, *19*, 825–849. [[CrossRef](#)]
11. Deshpande, N.; Londhe, S.; Kulkarni, S. Modeling Compressive Strength of Recycled Aggregate Concrete by Artificial Neural Network, Model Tree and Non-Linear Regression. *Int. J. Sustain. Built Environ.* **2014**, *3*, 187–198. [[CrossRef](#)]
12. Xu, J.; Chen, Y.; Xie, T.; Zhao, X.; Xiong, B.; Chen, Z. Prediction of triaxial behavior of recycled aggregate concrete using multivariable regression and artificial neural network techniques. *Constr. Build. Mater.* **2019**, *226*, 534–554. [[CrossRef](#)]
13. Xu, J.; Zhao, X.; Yu, Y.; Xie, T.; Yang, G.; Xue, J. Parametric sensitivity analysis and modelling of mechanical properties of normal- and high-strength recycled aggregate concrete using grey theory, multiple nonlinear regression and artificial neural networks. *Constr. Build. Mater.* **2019**, *211*, 479–491. [[CrossRef](#)]
14. Behnood, A.; Olek, J.; Glinicki, M.A. Predicting modulus elasticity of recycled aggregate concrete using M5' model tree algorithm. *Constr. Build. Mater.* **2015**, *94*, 137–147. [[CrossRef](#)]
15. Khademi, F.; Jamal, S.M.; Deshpande, N.; Londhe, S. Predicting strength of recycled aggregate concrete using Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System and Multiple Linear Regression. *Int. J. Sustain. Built Environ.* **2016**, *5*, 355–369. [[CrossRef](#)]
16. Deng, F.; He, Y.; Zhou, S.; Yu, Y.; Cheng, H.; Wu, X. Compressive Strength Prediction of Recycled Concrete Based on Deep Learning. *Constr. Build. Mater.* **2018**, *175*, 562–569. [[CrossRef](#)]
17. Ziolkowski, P.; Niedostatkiwicz, M. Machine Learning Techniques in Concrete Mix Design. *Materials* **2019**, *12*, 1256. [[CrossRef](#)]
18. Simon, M.J. *Concrete Mixture Optimization Using Statistical Methods: Final Report*; Office of Infrastructure Research and Development: McLean, VA, USA, 2003.
19. Yeh, I.C. Computer-Aided Design for Optimum Concrete Mixtures. *Cem. Concr. Compos.* **2007**, *29*, 193–202. [[CrossRef](#)]
20. Cheng, M.-Y.; Prayogo, D.; Wu, Y.-W. Novel Genetic Algorithm-Based Evolutionary Support Vector Machine for Optimizing High-Performance Concrete Mixture. *J. Comput. Civ. Eng.* **2014**, *28*, 06014003. [[CrossRef](#)]

21. Golafshani, E.M.; Behnood, A. Estimating the optimal mix design of silica fume concrete using biogeography-based programming. *Cem. Concr. Compos.* **2019**, *96*, 95–105. [[CrossRef](#)]
22. Gholampour, A.; Mansouri, I.; Kisi, O.; Ozbakkaloglu, T. Evaluation of mechanical properties of concretes containing coarse recycled concrete aggregates using multivariate adaptive regression splines (MARS), M5 model tree (M5Tree), and least squares support vector regression (LSSVR) models. *Neural Comput. Appl.* **2018**, *32*, 295–308. [[CrossRef](#)]
23. Dantas, A.T.A.; Leite, M.B.; De Jesus Nagahama, K. Prediction of compressive strength of concrete containing construction and demolition waste using artificial neural networks. *Constr. Build. Mater.* **2013**, *38*, 717–722. [[CrossRef](#)]
24. Salehi, H.; Burgueño, R. Emerging artificial intelligence methods in structural engineering. *Eng. Struct.* **2018**, *171*, 170–189. [[CrossRef](#)]
25. Mahdavinejad, M.S.; Rezvan, M.; Barekatain, M.; Adibi, P.; Barnaghi, P.; Sheth, A.P. Machine learning for internet of things data analysis: A survey. *Digit. Commun. Netw.* **2018**, *4*, 161–175. [[CrossRef](#)]
26. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, MA, USA, 2012.
27. Marsland, S. *Machine Learning: An Algorithmic Perspective*; Taylor & Francis: New York, NY, USA, 2015.
28. Noori, M.; Hassani, H.; Javaherian, A.; Amindavar, H.; Torabi, S. Automatic fault detection in seismic data using Gaussian process regression. *J. Appl. Geophys.* **2019**, *163*, 117–131. [[CrossRef](#)]
29. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.
30. Omran, B.A.; Chen, Q.; Jin, R. Comparison of Data Mining Techniques for Predicting Compressive Strength of Environmentally Friendly Concrete. *J. Comput. Civ. Eng.* **2016**, *30*, 04016029. [[CrossRef](#)]
31. Lawrence, N. Probabilistic Non-linear Principal Component Analysis with Gaussian Process Latent Variable Model. *J. Mach. Learn. Res.* **2005**, *6*, 1783–1816.
32. Williams, C.K.I.; Barber, D. Bayesian classification with Gaussian processes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 1342–1351. [[CrossRef](#)]
33. Tobar, F.; Bui, T.D.; Turner, R.E. Learning stationary time series using Gaussian processes with nonparametric kernels. *Adv. Neural. Inf. Process. Syst.* **2015**, *2015*, 3501–3509.
34. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
35. Ye, X.W.; Jin, T.; Chen, P.Y. Structural crack detection using deep learning-based fully convolutional networks. *Adv. Struct. Eng.* **2019**, *22*, 3412–3419. [[CrossRef](#)]
36. Toh, G.; Park, J. Review of vibration-based structural health monitoring using deep learning. *Appl. Sci.* **2020**, *10*, 1680. [[CrossRef](#)]
37. Gulli, A.; Pal, S. *Deep Learning with Keras*; Packt Publishing: Birmingham, UK, 2017.
38. Chollet, F. *Deep Learning with Python*; Manning Publications Co.: New York, NY, USA, 2018.
39. Zhao, H.; Chen, Z.; Jiang, H.; Jing, W.; Sun, L.; Feng, M. Evaluation of three deep learning models for early crop classification using Sentinel-1A imagery time series—a case study in Zhanjiang, China. *Remote Sens.* **2019**, *11*, 2673. [[CrossRef](#)]
40. Yao, K.; Cohn, T.; Vylomova, K.; Duh, K.; Dyer, C. Depth-Gated Recurrent Neural Networks. *arXiv* **2015**, arXiv:1508.03790.
41. Zhan, X.; Zhang, S.; Szeto, W.Y.; Chen, X. Multi-step-ahead traffic speed forecasting using multi-output gradient boosting regression tree. *J. Intell. Transp. Syst. Technol. Plan. Oper.* **2020**, *24*, 125–141. [[CrossRef](#)]
42. Friedman, J. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
43. Persson, C.; Bacher, P.; Shiga, T.; Madsen, H. Multi-site solar power forecasting using gradient boosted regression trees. *Sol. Energy* **2017**, *150*, 423–436. [[CrossRef](#)]
44. Friedman, J.H. Stochastic gradient boosting. *Comput. Stat. Data Anal.* **2002**, *38*, 367–378. [[CrossRef](#)]
45. Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Noise Reduction in Speech Processing. *Nat. Comput. Ser.* **2009**, *2*, 37–40.
46. Limbachiya, M.C.; Leelawat, T.; Dhir, R.K. Use of recycled concrete aggregate in high-strength concrete. *Mater. Struct. Constr.* **2000**, *33*, 574–580. [[CrossRef](#)]
47. Manzi, S.; Mazzotti, C.; Bignozzi, M.C. Short and long-term behavior of structural concrete with recycled concrete aggregate. *Cem. Concr. Compos.* **2013**, *37*, 312–318. [[CrossRef](#)]
48. Ajdukiewicz, A.; Kliszczewicz, A. Influence of recycled aggregates on mechanical properties of HS/HPC. *Cem. Concr. Compos.* **2002**, *24*, 269–279. [[CrossRef](#)]

49. Ajdukiewicz, A.B.; Kliszczewicz, A.T. Comparative Tests of Beams and Columns Made of Recycled Aggregate Concrete and Natural Aggregate Concrete. *J. Adv. Concr. Technol.* **2007**, *5*, 259–273. [[CrossRef](#)]
50. Gómez-Soberón, J.M.V. Porosity of recycled concrete with substitution of recycled concrete aggregate: An experimental study. *Cem. Concr. Res.* **2002**, *32*, 1301–1311. [[CrossRef](#)]
51. Sheen, Y.-N.; Wang, H.-Y.; Juang, Y.-P.; Le, D.-H. Assessment on the engineering properties of ready-mixed concrete using recycled aggregates. *Constr. Build. Mater.* **2013**, *45*, 298–305. [[CrossRef](#)]
52. Lin, Y.H.; Tyan, Y.Y.; Chang, T.P.; Chang, C.Y. An assessment of optimal mixture for concrete made with recycled concrete aggregates. *Cem. Concr. Res.* **2004**, *34*, 1373–1380. [[CrossRef](#)]
53. Thomas, C.; Setién, J.; Polanco, J.A.; Alaejos, P.; de Juan, M.S. Durability of recycled aggregate concrete. *Constr. Build. Mater.* **2013**, *40*, 1054–1065. [[CrossRef](#)]
54. Ulloa, V.A.; García-Taengua, E.; Pelufo, M.J.; Domingo, A.; Serna, P. New views on effect of recycled aggregates on concrete compressive strength. *ACI Mater. J.* **2013**, *110*, 687–696.
55. Matias, D.; de Brito, J.; Rosa, A.; Pedro, D. Mechanical properties of concrete produced with recycled coarse aggregates—Influence of the use of superplasticizers. *Constr. Build. Mater.* **2013**, *44*, 101–109. [[CrossRef](#)]
56. Taffese, W.Z. Suitability Investigation of Recycled Concrete Aggregates for Concrete Production: An Experimental Case Study. *Adv. Civ. Eng.* **2018**, *2018*, 8368351. [[CrossRef](#)]
57. Etxeberria, M.; Marí, A.R.; Vázquez, E. Recycled aggregate concrete as structural material. *Mater. Struct. Constr.* **2007**, *40*, 529–541. [[CrossRef](#)]
58. Andreu, G.; Miren, E. Experimental analysis of properties of high performance recycled aggregate concrete. *Constr. Build. Mater.* **2014**, *52*, 227–235. [[CrossRef](#)]
59. Etxeberria, M.; Vázquez, E.; Marí, A.; Barra, M. Influence of amount of recycled coarse aggregates and production process on properties of recycled aggregate concrete. *Cem. Concr. Res.* **2007**, *37*, 735–742. [[CrossRef](#)]
60. Beltrán, M.G.; Agrela, F.; Barbudo, A.; Ayuso, J.; Ramírez, A. Mechanical and durability properties of concretes manufactured with biomass bottom ash and recycled coarse aggregates. *Constr. Build. Mater.* **2014**, *72*, 231–238. [[CrossRef](#)]
61. Kou, S.C.; Poon, C.S.; Chan, D. Influence of Fly Ash as Cement Replacement on the Properties of Recycled Aggregate Concrete. *J. Mater. Civ. Eng.* **2007**, *19*, 709–717. [[CrossRef](#)]
62. Beltrán, M.G.; Barbudo, A.; Agrela, F.; Galvín, A.P.; Jiménez, J.R. Effect of cement addition on the properties of recycled concretes to reach control concretes strengths. *J. Clean. Prod.* **2014**, *79*, 124–133. [[CrossRef](#)]
63. Poon, C.S.; Kou, S.C.; Lam, L. Influence of recycled aggregate on slump and bleeding of fresh concrete. *Mater. Struct. Constr.* **2007**, *40*, 981–988. [[CrossRef](#)]
64. Çakır, Ö.; Sofyanlı, Ö.Ö. Influence of silica fume on mechanical and physical properties of recycled aggregate concrete. *HBRC J.* **2015**, *11*, 157–166. [[CrossRef](#)]
65. Rahal, K. Mechanical properties of concrete with recycled coarse aggregate. *Build. Environ.* **2007**, *42*, 407–415. [[CrossRef](#)]
66. Carneiro, J.A.; Lima, P.R.L.; Leite, M.B.; Filho, R.D.T. Compressive stress-strain behavior of steel fiber reinforced-recycled aggregate concrete. *Cem. Concr. Compos.* **2014**, *46*, 65–72. [[CrossRef](#)]
67. Sato, R.; Maruyama, I.; Sogabe, T.; Sogo, M. Flexural behavior of reinforced recycled concrete beams. *J. Adv. Concr. Technol.* **2007**, *5*, 43–61. [[CrossRef](#)]
68. Dilbas, H.; Şimşek, M.; Çakır, Ö. An investigation on mechanical and physical properties of recycled aggregate concrete (RAC) with and without silica fume. *Constr. Build. Mater.* **2014**, *61*, 50–59. [[CrossRef](#)]
69. Casuccio, M.; Torrijos, M.C.; Giaccio, G.; Zerbino, R. Failure mechanism of recycled aggregate concrete. *Constr. Build. Mater.* **2008**, *22*, 1500–1506. [[CrossRef](#)]
70. Kou, S.C.; Poon, C.S.; Chan, D. Influence of fly ash as a cement addition on the hardened properties of recycled aggregate concrete. *Mater. Struct. Constr.* **2008**, *41*, 1191–1201. [[CrossRef](#)]
71. Folino, P.; Xargay, H. Recycled aggregate concrete—Mechanical behavior under uniaxial and triaxial compression. *Constr. Build. Mater.* **2014**, *56*, 21–31. [[CrossRef](#)]
72. Yang, K.-H.; Chung, H.-S.; Ashou, A.F. Influence of Type and Replacement Level of Recycled Aggregates on Concrete Properties. *ACI Mater. J.* **2008**, *105*, 289–296.
73. Gayarre, F.L.; Pérez, C.L.C.; López, M.A.S.; Cabo, A.D. The effect of curing conditions on the compressive strength of recycled aggregate concrete. *Constr. Build. Mater.* **2014**, *53*, 260–266. [[CrossRef](#)]

74. Domingo-Cabo, A.; Lázaro, C.; López-Gayarre, F.; Serrano-López, M.A.; Serna, P.; Castaño-Tabares, J.O. Creep and shrinkage of recycled aggregate concrete. *Constr. Build. Mater.* **2009**, *23*, 2545–2553. [[CrossRef](#)]
75. Medina, C.; Zhu, W.; Howind, T.; de Rojas, M.I.S.; Frías, M. Influence of mixed recycled aggregate on the physical-mechanical properties of recycled concrete. *J. Clean. Prod.* **2014**, *68*, 216–225. [[CrossRef](#)]
76. Corinaldesi, V. Mechanical and elastic behaviour of concretes made of recycled-concrete coarse aggregates. *Constr. Build. Mater.* **2010**, *24*, 1616–1620. [[CrossRef](#)]
77. Kumutha, R.; Vijai, K. Strength of concrete incorporating aggregates recycled from demolition waste. *J. Eng. Appl. Sci.* **2010**, *5*, 64–71.
78. Pepe, M.; Filho, R.D.T.; Koenders, E.A.B.; Martinelli, E. Alternative processing procedures for recycled aggregates in structural concrete. *Constr. Build. Mater.* **2014**, *69*, 124–132. [[CrossRef](#)]
79. Malešev, M.; Radonjanin, V.; Marinković, S. Recycled concrete as aggregate for structural concrete production. *Sustainability* **2010**, *2*, 1204–1225. [[CrossRef](#)]
80. Wardeh, G.; Ghorbel, E.; Gomart, H. Mix Design and Properties of Recycled Aggregate Concretes: Applicability of Eurocode 2. *Int. J. Concr. Struct. Mater.* **2015**, *9*, 1–20. [[CrossRef](#)]
81. Belén, G.F.; Fernando, M.A.; Diego, C.L.; Sindy, S.P. Stress-strain relationship in axial compression for concrete using recycled saturated coarse aggregate. *Constr. Build. Mater.* **2011**, *25*, 2335–2342. [[CrossRef](#)]
82. Haitao, Y.; Shizhu, T. Preparation and properties of high-strength recycled concrete in cold areas. *Mater. Constr.* **2015**, *65*, 2–9. [[CrossRef](#)]
83. Fathifazl, G.; Razaqpur, A.G.; Isgor, O.B.; Abbas, A.; Fournier, B.; Foo, S. Creep and drying shrinkage characteristics of concrete produced with coarse recycled concrete aggregate. *Cem. Concr. Compos.* **2011**, *33*, 1026–1037. [[CrossRef](#)]
84. Tam, V.W.Y.; Kotrayothar, D.; Xiao, J. Long-term deformation behaviour of recycled aggregate concrete. *Constr. Build. Mater.* **2015**, *100*, 262–272. [[CrossRef](#)]
85. Rao, M.C.; Bhattacharyya, S.K.; Barai, S.V. Influence of field recycled coarse aggregate on properties of concrete. *Mater. Struct. Constr.* **2011**, *44*, 205–220.
86. Abdel-Hay, A.S. Properties of recycled concrete aggregate under different curing conditions. *HBRC J.* **2017**, *13*, 271–276. [[CrossRef](#)]
87. Somna, R.; Jaturapitakkul, C.; Chalee, W.; Rattanachu, P. Effect of the Water to Binder Ratio and Ground Fly Ash on Properties of Recycled Aggregate Concrete. *J. Mater. Civ. Eng.* **2012**, *24*, 16–22. [[CrossRef](#)]
88. Zheng, C.; Lou, C.; Du, G.; Li, X.; Liu, Z.; Li, L. Mechanical properties of recycled concrete with demolished waste concrete aggregate and clay brick aggregate. *Results Phys.* **2018**, *9*, 1317–1322. [[CrossRef](#)]
89. Elhakam, A.A.; Mohamed, A.E.; Awad, E. Influence of self-healing, mixing method and adding silica fume on mechanical properties of recycled aggregates concrete. *Constr. Build. Mater.* **2012**, *35*, 421–427. [[CrossRef](#)]
90. Nepomuceno, M.C.S.; Isidoro, R.A.S.; Catarino, J.P.G. Mechanical performance evaluation of concrete made with recycled ceramic coarse aggregates from industrial brick waste. *Constr. Build. Mater.* **2018**, *165*, 284–294. [[CrossRef](#)]
91. Barbudo, A.; de Brito, J.; Evangelista, L.; Bravo, M.; Agrela, F. Influence of water-reducing admixtures on the mechanical performance of recycled concrete. *J. Clean. Prod.* **2013**, *59*, 93–98. [[CrossRef](#)]
92. Mohammed, N.; Sarsam, K.; Hussien, M. The influence of recycled concrete aggregate on the properties of concrete. In *MATEC Web of Conferences*; EDP Sciences: Ulis, France, 2018; Volume 162, pp. 1–7.
93. Butler, L.; West, J.S.; Tighe, S.L. Effect of recycled concrete coarse aggregate from multiple sources on the hardened properties of concrete with equivalent compressive strength. *Constr. Build. Mater.* **2013**, *47*, 1292–1301. [[CrossRef](#)]
94. Thomas, C.; Setién, J.; Polanco, J.A.; Cimentada, A.I.; Medina, C. Influence of curing conditions on recycled aggregate concrete. *Constr. Build. Mater.* **2018**, *172*, 618–625. [[CrossRef](#)]
95. Ismail, S.; Ramli, M. Engineering properties of treated recycled concrete aggregate (RCA) for structural applications. *Constr. Build. Mater.* **2013**, *44*, 464–476. [[CrossRef](#)]
96. Younis, K.H.; Pilakoutas, K. Strength prediction model and methods for improving recycled aggregate concrete. *Constr. Build. Mater.* **2013**, *49*, 688–701. [[CrossRef](#)]
97. Kim, K.; Shin, M.; Cha, S. Combined effects of recycled aggregate and fly ash towards concrete sustainability. *Constr. Build. Mater.* **2013**, *48*, 499–507. [[CrossRef](#)]
98. Shanker, M.; Hu, M.Y.; Hung, M.S. Effect of Data Standardization on Neural Network Training. *Omega Int. J.* **1996**, *24*, 385–397. [[CrossRef](#)]

99. Nilsen, V.; Pham, L.T.; Hibbard, M.; Klager, A.; Cramer, S.M.; Morgan, D. Prediction of concrete coefficient of thermal expansion and other properties using machine learning. *Constr. Build. Mater.* **2019**, *220*, 587–595. [[CrossRef](#)]
100. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2008.
101. Bardenet, R.; Brendel, M.; Kégl, B.; Sebag, M. Collaborative hyperparameter tuning. In Proceedings of the 30th International Conference on Machine Learning (ICML 2013), Atlanta, GA, USA, 16–21 June 2013; Volume 2, pp. 858–866.
102. Bergstra, J.; Yamins, D.; Cox, D.D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In Proceedings of the 30th International Conference on Machine Learning (ICML 2013), Atlanta, GA, USA, 16–21 June 2013; Volume 1, pp. 115–123.
103. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
104. Varoquaux, G.; Buitinck, L.; Louppe, G.; Grisel, O.; Pedregosa, F.; Mueller, A. Scikit-learn. *GetMob. Mob. Comput. Commun.* **2015**, *19*, 29–33. [[CrossRef](#)]
105. Whang, J.; Matsukawa, A. Exploring Batch Normalization in Recurrent Neural Networks. Stanford Center for Professional Development. Available online: https://jaywhang.com/assets/batchnorm_rnn.pdf (accessed on 28 September 2020).
106. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning (ICML 2015), Lille, France, 6–11 July 2015; Volume 1, pp. 448–456.
107. Chollet, F. Keras. GitHub Repository. 2015. Available online: <https://github.com/fchollet/keras> (accessed on 28 September 2020).
108. Marani, A.; Nehdi, M.L. Machine learning prediction of compressive strength for phase change materials integrated cementitious composites. *Constr. Build. Mater.* **2020**, *265*, 120286. [[CrossRef](#)]
109. Penadés-Plà, V.; García-Segura, T.; Martí, J.V.; Yepes, V. A review of multi-criteria decision-making methods applied to the sustainable bridge design. *Sustainability* **2016**, *8*, 1295. [[CrossRef](#)]
110. Lu, P.; Chen, S.; Zheng, Y. Artificial intelligence in civil engineering. *Math. Probl. Eng.* **2012**, *2012*, 1–22. [[CrossRef](#)]
111. Wijayasundara, M.; Mendis, P.; Zhang, L.; Sofi, M. Financial assessment of manufacturing recycled aggregate concrete in ready-mix concrete plants. *Resour. Conserv. Recycl.* **2016**, *109*, 187–201. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).