Fırat Üniversitesi Müh. Bil. Dergisi
36(2), 563-580, 2024

Araştırma Makalesi
https://doi.org/10.35234/fumbd.1375026

Comparing Machine Learning Regression Models for Early-Age Compressive Strength Prediction of Recycled Aggregate Concrete

Muhammed ULUCAN[1*], Güngör YILDIRIM[2], Bilal ALATAŞ[3], Kürşat Esat ALYAMAÇ[4]
[1,4] İnşaat Mühendisliği, Mühendislik Fakültesi, Fırat Üniversitesi, Elazığ, Türkiye
[2] Bilgisayar Mühendisliği, Mühendislik Fakültesi, Fırat Üniversitesi, Elazığ, Türkiye
[3] Yazılım Mühendisliği, Mühendislik Fakültesi, Fırat Üniversitesi, Elazığ, Türkiye
[*1] mulucan@firat.edu.tr, [2] gungor.yildirim@firat.edu.tr, [3] balatas@firat.edu.tr, [4] kealyamac@firat.edu.tr

(Geliş/Received: 13/10/2023; Kabul/Accepted: 03/07/2024)

Abstract: A branch of artificial intelligence called machine learning is well-positioned as a prediction method that can take into consideration several influencing factors and complex inter-factor connections. Without being specifically trained to do so, these machine learning models have the ability to generalize, predict, and learn from data. Regression theory is a key topic in statistical modelling and machine learning. The main goal of this study is to compare the performance of several popular machine learning regression models for predicting the early-age compressive strength of concretes made from recycled concrete aggregates from a structure that demolished following the Sivrice-Elazig earthquake on January 24, 2020. Early-age concrete compressive strength is even more crucial due to factors like the fact that there are thousands of newly built structures in the aftermath of the earthquake, the quick manufacturing of these structures, and the completion of the project in the lowest amount of time. Determining the early-age concrete strength with high accuracy and in a useful manner is crucial for this reason. Seven different classical machine learning algorithms were employed in this study to achieve all of these goals. Early-age concrete compressive strength values were considered for 1 and 3 days. The relationship between the experimental results and the predicted outcomes of the employed algorithms was investigated, and a thorough comparison of these intelligent regression algorithms was conducted. Within the scope of sustainable development and circular economy goals, it is thought that this article will make significant contributions to the literature in terms of utilizing these waste materials and determining the early-age compressive strengths of the concretes produced with high accuracy.

Key words: Compressive strength, Construction and demolition waste, Machine learning, Recycled aggregate concrete.

Agrega türlerinin ve farklı oranlarda kullanımının geri dönüştürülmüş agrega beton basınç dayanımına etkisinin makine öğrenmesi regresyon modellemesi uyarlanarak değerlendirilmesi

Öz: Yapay zekanın makine öğrenimi olarak adlandırılan bir dalı, çeşitli etki faktörlerini ve karmaşık faktörler arası bağlantıları dikkate alabilen bir tahmin yöntemi olarak iyi bir konuma sahiptir. Bu makine öğrenimi modelleri, özel olarak eğitilmeksizin verileri genelleştirme, tahmin etme ve onlardan öğrenme becerisine sahiptir. Regresyon teorisi, istatistiksel modelleme ve makine öğreniminde kilit bir konudur. Bu çalışmanın temel amacı, 24 Ocak 2020'deki Sivrice-Elazığ depreminin ardından yıkılan bir binadan elde edilen geri dönüştürülmüş beton agregalarından üretilen betonların erken yaş basınç dayanımını tahmin etmek için birkaç popüler makine öğrenimi regresyon modelinin performansını karşılaştırmaktır. Deprem sonrasında yeni inşa edilen binlerce yapının olması, bu yapıların hızlı bir şekilde imal edilmesi ve projenin en kısa sürede tamamlanması gibi faktörler nedeniyle erken yaş basınç dayanımı daha da büyük önem taşımaktadır. Erken yaş beton dayanımının yüksek doğrulukla ve kullanışlı bir şekilde belirlenmesi bu nedenle çok önemlidir. Bu çalışmada tüm bu hedeflere ulaşmak için yedi farklı klasik makine öğrenimi algoritması kullanılmıştır. Erken yaş basınç dayanımı değerleri 1 ve 3 gün için dikkate alınmıştır. Deneysel sonuçlar ile kullanılan algoritmaların öngördüğü sonuçlar arasındaki ilişki incelenmiş ve bu akıllı regresyon algoritmalarının kapsamlı bir karşılaştırması yapılmıştır. Sürdürülebilir kalkınma ve döngüsel ekonomi hedefleri kapsamında bu atık malzemelerin değerlendirilmesi ve üretilen betonların erken yaş basınç dayanımlarının yüksek doğrulukla belirlenebilmesi açısından makalenin literatüre önemli katkılar sağlayacağı düşünülmektedir.

Anahtar kelimeler: Basınç dayanımı, İnşaat ve yıkıntı atıkları, Makine öğrenme, Geri dönüşüm agregalı beton.

1. Introduction

The construction and industrial sector have a very important market worldwide, therefore, it is one of the sectors most affected by sustainability policies. The construction industry consumes approximately 50% of all natural resources and 40% of all energy, in addition to about 50% of all global waste streams [1,2]. In addition, the human population, the urban population, and the consequent increasing need for shelter bring about rapid urbanization and lead to large amounts of consumption. In rapid urbanization, large amounts of concrete are used

* Sorumlu yazar: mulucan@firat.edu.tr. Yazarların ORCID Numarası: [1] 0000-0001-7629-6846, [2] 0000-0002-4096-4838, [3] 0000-0002-3513-0329, [4] 0000-0002-3226-4073,

to construct new buildings and renew old buildings under the name of urban transformation projects. Concrete is the most widely used building material worldwide and consists mainly of cement, aggregate, and water [3,4]. When the basic components of concrete are examined, two main problems are encountered. The first is the environmental pollution caused by cement worldwide, and the second is the trend toward depletion of natural resources in parallel with the increase in concrete use [5–7].

Aggregates constitute 65-70% of the concrete volume and play a major role in the depletion of natural resources due to the increasing use of concrete [8]. This expanding consumption, and waste materials arising from the construction sector have led researchers to different study subjects [9]. Studies on the evaluation of construction and demolition wastes that are damaged or emerged as a result of earthquakes have an important place in this study area. The large amount of waste generated from demolished buildings and the storage of these wastes pose a great threat to the environment [10,11]. For this reason, evaluating these wastes is very important, and studies on the use of these wastes in concrete are becoming increasingly common. The 24 January 2020 Sivrice-Elazig earthquake that occurred in Turkey in recent years can be shown as a great example of this. Elazig and Malatya provinces were affected after the earthquake and the damage was concentrated in Elazig center. Tens of thousands of buildings across Elazig were examined and classified according to their damage status. After this earthquake, thousands of buildings were destroyed, and millions of tons of waste material emerged. Ulucan and Alyamac analyzed a typical building in detail in their study and estimated the approximate total amount of waste generated after the earthquake, taking into account the total number of demolished buildings [12]. In addition, to evaluate these wastes, a detailed laboratory study was carried out using recycled concrete aggregates (RCAs) obtained from a building destroyed after the earthquake on concrete and evaluated the mechanical, environmental, and economic properties of these aggregates. In another study, Ulucan and Alyamac examined in detail the advantages and disadvantages of using RCA in a high-rise building [13]. In their study, Kül et al. aimed to provide a solution for the conversion of construction and demolition waste (CDW) into building materials suitable for accelerated construction and housing in developing countries and disaster-prone areas [14]. Ilcan et al. evaluated the effects of incorporating industrial wastes into CDW-based geopolymer blends [15]. Ozcelikci et al. were developed non-structural ultra-lightweight circular building insulation materials using significant amounts of CDW [16].

Researchers have offered solutions from different perspectives to different problems in order to reduce the negative effects, provide a sustainable life and obtain high-accuracy results. The increasing use of concrete in recent years makes it necessary to determine the quality and different properties of concrete with high accuracy [17,18]. So, it is very important to determine the compressive strength, which is one of concrete's most important mechanical properties, with high accuracy [19]. For this purpose, machine learning algorithms have been widely used in recent years [20]. In their study, Kandiri et al tried to predict the compressive strength of concretes containing ground granulated blast furnace slag by using hybridized multi-objective ANN and salp swarm algorithm [21]. Zhang et al proposed a method based on machine learning and metaheuristic algorithms to optimize concrete mixing ratios and compared this method with other methods [22]. Golafshani et al. used a multi-layer neural network, a radial basis function neural network and Harris hawks optimization algorithm to develop models that predict the concrete strength of concretes containing supplementary cementitious materials [23]. Tam et al. used artificial neural networks to predict the compressive strength of CO_2 concrete. R square and average error values obtained as a result of this study gave very satisfactory results [24]. Amiri and Hatami used artificial neural networks to predict the mechanical and durability properties of concrete containing slag and RCA [25].

Machine learning, a branch of artificial intelligence, is well-positioned as a prediction method capable of taking into account several determining factors and intricate inter-factor relationships. These machine learning models possess the capacity to generalize, predict, and learn from data without explicitly being programmed to do so. The subject of regression is crucial to statistical modelling and machine learning. It relates to modelling issues when we should predict the values of an additional variable based on the prior values of the predictors, which may contain one or more variables. The main goal of this study is to examine the prediction accuracy of the early-age compressive strength of concretes produced using RCAs obtained from a building that demolished after the 24 January 2020 Sivrice-Elazig earthquake with different machine learning regression models. Early-age concrete strength values of 1 and 3 days, which are of great importance for the rapid construction of the new buildings after the earthquake and the accurate determination of the formwork stripping times, were considered. The relationship between the prediction results of the regression models and the experimental results was examined, and a detailed comparison of these models was made. This study made a comprehensive comparison of different machine learning algorithms that accurately predict the compressive strength of early-age concrete and reduce environmental pollution by using post-earthquake wastes. This paper also contributes to the literature by aiming to utilize waste materials and determine the early-age compressive strength of concrete with high accuracy in the context of sustainable development and circular economy goals. The main contributions of this paper are listed below:

- The effect of aggregate type and its use in different ratios on compressive strength was investigated in detail using intelligent regression models by designing concrete mixtures containing all-natural, all-recyclable and both natural and recycled concrete aggregates.
- The problem of predicting the early age compressive strength of concretes made from RCAs is modelled as a regression problem.
- Instead of a single method, seven different intelligent popular methods were adapted to find solutions for the focused problem. Seven of the most popular machine learning based regression models were applied to achieve better results all together for the first time.

The remaining of this paper is organized as follows. The materials and methods are described in Section 2. The experiments and discussion about the obtained results are presented in Section 3 and the paper is finalized with possible future research directions in Section 4.

2. Materials and method

2.1. Materials

In this study, CEM 1 42.5 R Portland cement obtained from the Elazig Seza cement factory was used. The chemical compositions of cement are given in Table 1. In order to examine the effect of aggregate type on concrete strength in concrete castings, natural concrete aggregates (NCAs) and RCAs were used. RCAs were obtained from concrete masses that emerged after demolishing a damaged building after the 24 January 2020 Sivrice-Elazig earthquake. Natural aggregates were obtained from the Elazig Çemişgezek region. Aggregates were classified as 0-4 (fine), 4-16 (coarse 1), and 16-31.5 (coarse 2). The experiments carried out in the laboratory to determine the physical properties of these two different aggregate types and the results obtained are given in Table 2. CHRYSO Optima 280-SC3 was used as a water-reducing chemical additive in all concrete mixes.

Table 1. Chemical composition of Portland cement

Chemical Properties	CaO	SiO_2	Fe_2O_3	Al_2O_3	SO_3	Na_2O	K_2O	MgO	P_2O_5	Loss of ignition
Cement	62.48	19.15	4.35	5.50	2.55	0.11	0.45	2.05	-	2.09

Table 2. Physical properties of NCA and RCA

Properties	Fine		Coarse 1		Coarse 2	
	NCA	RCA	NCA	RCA	NCA	RCA
Los Angeles abrasion	-	-	-	-	15	23
Specific gravity	2.66	2.48	2.69	2.69	2.71	2.73
Water absorption (%)	1.5	9.8	1.3	4.2	1.2	3.6

2.2. Experimental design and preparation of concrete mixtures

Within the scope of this study, a total of 45 different concrete series were produced to examine the effect of aggregate types on concrete strength in detail. Mixture designs consist of all natural aggregate, all recycled aggregate, and concrete mixtures containing both natural and RCA. In this direction, concrete series containing 3 different aggregate designs in addition to 5 different water-to-cement ratios and 3 different cement dosages were prepared. It was expressed as natural aggregate concretes (NAC), recycled aggregate concretes (RAC), and recycled coarse aggregate concretes (RCAC). The prepared concrete series were subjected to compressive strength tests on the 1st and 3rd days, and early-age concrete strengths were obtained. Mixture amounts of these series are given in Table 3.

Table 3. Mix proportions of concrete mixtures (kg/m^3)

Mixture Code	Cement	Water	NCA			RCA			Chemical Additive
			Fine	Coarse 1	Coarse 2	Fine	Coarse 1	Coarse 2	
NAC-1	300	90	642	649	872	-	-	-	7.8
NAC-2	300	105	630	637	855	-	-	-	6.0
NAC-3	300	120	618	625	839	-	-	-	4.2
NAC-4	300	135	606	613	823	-	-	-	3.3
NAC-5	300	150	594	601	807	-	-	-	1.5
NAC-6	350	105	617	624	838	-	-	-	7.7
NAC-7	350	123	603	610	819	-	-	-	5.3
NAC-8	350	140	589	596	800	-	-	-	3.5
NAC-9	350	158	575	582	781	-	-	-	2.5
NAC-10	350	175	561	568	762	-	-	-	1.8
NAC-11	400	120	592	599	805	-	-	-	9.2
NAC-12	400	140	576	583	783	-	-	-	6.8
NAC-13	400	160	560	567	761	-	-	-	4.0
NAC-14	400	180	544	551	740	-	-	-	2.4
NAC-15	400	200	528	534	718	-	-	-	1.2
RAC-1	300	90	-	-	-	598	649	878	7.8
RAC-2	300	105	-	-	-	587	637	862	6.9
RAC-3	300	120	-	-	-	576	625	845	6.0
RAC-4	300	135	-	-	-	565	613	829	3.3
RAC-5	300	150	-	-	-	554	601	813	1.8
RAC-6	350	105	-	-	-	575	624	844	7.0
RAC-7	350	123	-	-	-	562	610	825	5.6
RAC-8	350	140	-	-	-	549	596	806	2.5
RAC-9	350	158	-	-	-	536	582	787	1.4
RAC-10	350	175	-	-	-	523	568	768	1.1
RAC-11	400	120	-	-	-	552	599	811	5.6
RAC-12	400	140	-	-	-	537	583	789	4.0
RAC-13	400	160	-	-	-	522	567	767	2.4
RAC-14	400	180	-	-	-	508	551	745	1.2
RAC-15	400	200	-	-	-	493	534	723	2.8
RCAC-1	300	90	642	-	-	-	649	878	7.5
RCAC-2	300	105	630	-	-	-	637	862	6.0
RCAC-3	300	120	618	-	-	-	625	845	4.8
RCAC-4	300	135	606	-	-	-	613	829	3.8
RCAC-5	300	150	594	-	-	-	601	813	2.7
RCAC-6	350	105	617	-	-	-	624	844	7.0
RCAC-7	350	123	603	-	-	-	610	825	5.6
RCAC-8	350	140	589	-	-	-	596	806	4.2
RCAC-9	350	158	575	-	-	-	582	787	2.8
RCAC-10	350	175	561	-	-	-	568	768	1.8
RCAC-11	400	120	592	-	-	-	599	811	6.0
RCAC-12	400	140	576	-	-	-	583	789	4.4
RCAC-13	400	160	560	-	-	-	567	767	3.2
RCAC-14	400	180	544	-	-	-	551	745	2.0
RCAC-15	400	200	528	-	-	-	534	723	1.2

2.3. The Machine Learning Methods Used

There are many machine learning algorithms in the literature today. According to the No Free Lunch Theorem [26], no machine learning (ML) algorithm can guarantee always finding the best solution for all problems. Therefore, in this study, the data obtained from the comprehensive laboratory experiments were analyzed with seven different ML regression algorithms that are well-known in the literature. These are Multiple Linear Regression, Support Vector Regression, Decision Tree Regression, Random Forest Regression, K-nearest neighbor regression and Feed Forward Neural Network Regression. The algorithm selections to be used in the study were chosen to take into account the characteristics of the data set and allow possible linear/nonlinear relationships to be modeled. The inclusion of Multiple Linear Regression serves as a basic foundation for capturing linear relationships between variables. Support Vector Regression expands our modeling capabilities by accommodating both linear and nonlinear models. Decision Tree Regression and Random Forest Regression were chosen to capture nonlinear relationships and complex interactions within the dataset. The ensemble structure of Random Forest helps reduce overfitting and improve generalization. K-nearest neighbor regression was chosen for its ability to capture local patterns, which is particularly valuable in scenarios that emphasize spatial dependencies. The inclusion of Feed Forward Neural Network Regression acknowledges the power of deep learning models in extracting complex, hierarchical features. This choice is especially important when dealing with data that exhibits complex relationships that cannot be effectively captured by traditional algorithms. This section contains general information about the machine learning algorithms used in the study.

2.3.1. Multiple Linear Regression (MLR)

Multi Linear Regression is a statistical method that fits a linear equation for data with one dependent (y) and more than one independent variable (x) [27]. The basic principle of MLR is to find the weights (β) of the attributes (independent variables) that minimize the sum of squared errors (SSE) during the training process. During the training process, the estimation ($\hat{y}_i$) made by the MLR for the ith observation data is calculated by Equation (1). The error for this observation data is $\varepsilon_i = (\hat{y}_i - y_i)$. Equation (2) finds the SSE value for N training data. Minimizing SSE means finding the smallest hyperplane with a vertical offset.

$$\hat{y}_i = \beta_0 + \beta_1 x_i^{(1)} + \beta_2 x_i^{(2)} + \cdots + \beta_m x_i^{(m)}, \qquad m{:}\, The\ number\ of\ features \tag{1}$$
$$SSE = \sum_{i=1}^{N} {\varepsilon_i}^2 \tag{2}$$

2.3.2. Multiple Polynomial Regression (MPR)

MPR is a special type of multiple linear regression. It is a frequently used method for data whose distribution is not completely linear [28]. In this method, the aim is to find a polynomial functional relationship between the independent variables and the dependent variable. As shown in Equation (3), the model can include linear, high-order, and interaction forms of the independent variables. Although the dependent variables are not in linear form in the model, the parameters are in linear form. In addition to the coefficients of the independent variables estimated in this method, the degree of the polynomial is one of the main factors that make the SSE minimum.

$$\hat{y}_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1}^2 + \beta_4 x_{i2}^2 + \beta_5 x_{i1} x_{i2} + \cdots \tag{3}$$

2.3.3. Support Vector Regression (SVR)

The general purpose of linear regression methods is to minimize the SSE. Classical linear models are generally not concerned with determining the error level. The SVR method can achieve this using an objective function with a constraint [29]. The main purpose of SVR is to minimize the L2 form of hyperplane coefficients. The constraint of the objective function is that the estimation error remains within a certain margin (ε_{max}, maximum error). However, in practice, margin deviation (δ) can also increase the performance of the model. For this reason, margin deviation is usually added to the objective function. The general objective function and constraint function used in SVR are given in Equation (4) and (5). In SVR, it is tried to ensure that according to the values of the hyperparameters *C* and ε_{max},, the maximum number of training data remains within the margin limits of the hyperplane.

$$\min \left(\frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^{N} |\delta_i| \right) \tag{4}$$
$$|\hat{y}_i - y_i| < \varepsilon_{max} + |\delta_i| \tag{5}$$

For complex data where linear separation is difficult, the SVR method uses kernel functions. Kernel functions increase the dimension of the input space and help find the most suitable hyperplane in the new dimensions to be formed. Although there are different kernel functions in the literature; Linear, RBF, and Polynomial kernel functions given in Equation (6), (7), and (8) are used in this study.

$$K_{lin}(x_i, x_j) = x_i x_j \tag{6}$$
$$K_{RBF}(x_i, x_j) = \exp\left(-\gamma \left|x_i - x_j\right|^2\right) \quad \gamma :\ the\ spread\ of\ the\ kernel, \quad \gamma > 0 \tag{7}$$
$$K_{pol}(x_i, x_j) = \left(\alpha\left(x_i x_j\right) + \tau\right)^d \quad \alpha{:}\, slope \ \ d{:}\, polynom\ degree \ \ \tau{:}\, trade-off\ parameter \tag{8}$$

2.3.4. Decision Tree Regression (DTR)

Decision Tree [30], one of the supervised learning algorithms, is frequently used in both classification and regression problems. The basic principle is to iteratively divide the relevant data set into smaller sub-parts with if–

else rules. In the resulting tree structure, conditions and results are expressed as condition (decision) nodes, and end nodes. The top condition node is called the root node. In DTR, it is important which attribute will be the condition node and how long the tree depth will be. Two basic parameters are taken into account while constructing the regression tree. The first of these is how homogeneous (impurity) the data is in each partition.

In DTR, homogeneity is calculated with the standard deviation. The decision tree is tried to be created to include homogeneous instances of the dependent variable. For this reason, the condition node selection should be made by considering the decrease in standard deviation in each partitioning. For this, the Standard Deviation Reduction (SDR) value, which is dependent on the standard deviation decrease, is taken into account in each segmentation. SDR is the second important parameter in the creation of DTR. The attribute with a high SDR value is selected as the condition node. Equation (9) and (10) show how to calculate the SDR for an attribute. In the equations, y and f represent the dependent variable and an attribute of the training data, respectively. $S(y, f)$, are the sums of the probabilities of each value of the independent variable multiplied by the dependent-variable-standard-deviation-value calculated with respect to the relevant independent variable.

$$SDR(y, f) = Std(y) - S(y, f) \tag{9}$$
$$S(y, f) = \sum_{c \in f} P(c)S(c) \tag{10}$$

2.3.5. Random Forest Regression (RFR)

Overfitting is among the common problems of decision tree algorithms. Random forest algorithms in which more than one decision tree is run can be used to overcome the overfitting problem. With this feature, the random forest algorithm is among a well-known ensemble method in the literature [23]. As shown in Equation (11), for the given feature set *X*, the estimate calculated by a random forest regression (RFR) method containing *T* DTRs is the average of *T* DTRs.

$$RFR_{pred}\,(X) = \frac{1}{T}\sum_{i=1}^{T} DTR_i(X) \tag{11}$$

2.3.6. K-Nearest Neighbours Regression (KNNR)

KNNR [31], a non-parametric technique, uses similarity (distance) information for prediction. The basic principle is to find the similarity (distance) with each data in the data set of the attribute ($X = \{x_1, x_2, .. x_m\}$) to be estimated. The similarity value of X with the jth data can be calculated with the Minkowski distance given in Equation (12). The distance found is called Manhattan distance in case of *p*=1, and Euclidean distance in case of *p*=2. Then, according to the calculated similarity values, the average of the dependent variables (y) of the *K* nearest neighbour data gives the prediction value (Equation (13)).

$$Distance = \left[\sum_{i=1}^{m}\left(\left|x_i - x_i^{(j)}\right|\right)^p\right]^{\frac{1}{p}} \tag{12}$$
$$KNNR_{pred}\,(X) = \frac{1}{K}\sum_{i=1}^{K} Ng_i^y, \quad Ng \in Neighbours\ set \tag{13}$$

2.3.7. Feedforward Neural Network Regression (FFNNR)

Artificial neural networks (ANNs) are a supervised learning method that has been used in the literature for many years. The basic components of ANNs are nodes (neurons) and layers containing nodes. ANNs models may differ according to the type of problem. In other words, different ANNs can contain different numbers of layers and nodes. Nodes can be connected to each other in different ways. ANN models that do not contain a cycle are called Feedforward Neural Networks (FFNNs). An FFNN node multiplies the values from the input attributes or previous layer nodes by weight coefficients (w) and calculates their sum. An intercept (b) value can be added to this total. Initially, w and b values are usually randomly determined. The value obtained as a result of these operations is then transferred to a non-linear function called the activation function (AF, σ). Since the w and b values of the node also determine the output of AF, the behavior of each node in FFNN can be independent and different from each other. This causes the outputs of different nodes to be more dominant for different input values. The purpose of AF is to provide both non-linear and allow the model to be differentiable in order to determine the optimum weight and bias values. The output value (n_{out}) of an FFNN node is found by Equation (14).

$$n_{out} = \sigma(b + \sum_{i=1}^{m} x_i w_i) \tag{14}$$

The main purpose of FFNN is to find the optimum values of w and b for each node that will minimize the cost function (C_f) of the model. For this, a gradient decent-based optimization process is performed for a certain number of epochs. These operations are called back-propagation. The cost function used for regression in FFNN is as in Equation (15).

$$C_f = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{15}$$

2.4. Data Set and its Characteristics

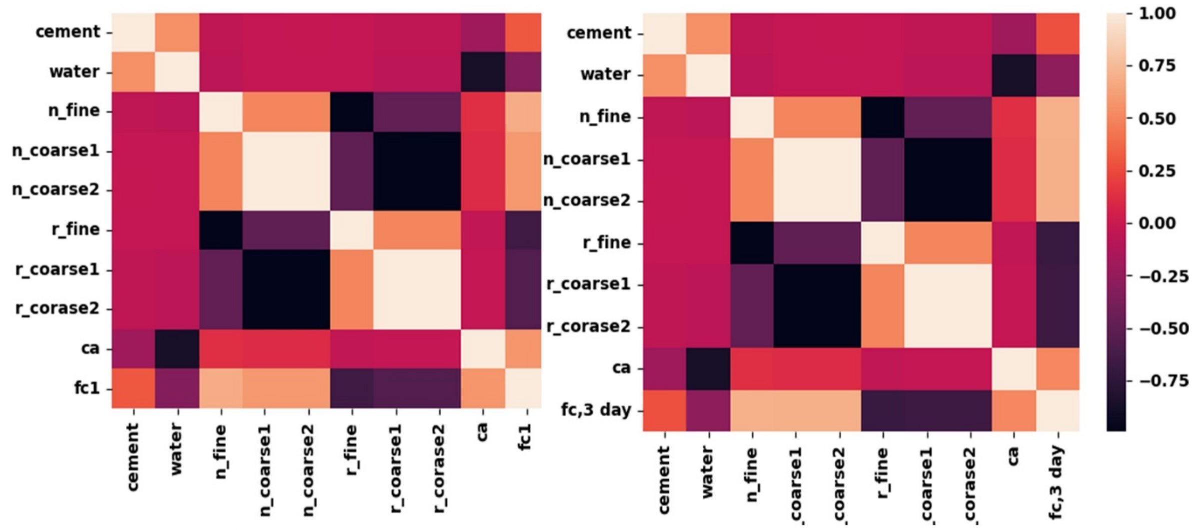
The structure of the data set used for artificial intelligence algorithms and the pre-processing process are important. Numerical, categorical, or ordinal-type values can be found in the data collected in engineering. Most of the time, heterogeneous data types need to be brought into a format that the algorithm to be used will evaluate. In addition, how much all the attributes in a data set affect the regression result is another issue that needs attention. In this section, the details of the data set obtained as a result of the experiments, the pre-processing process of the data set, and the quality analysis will be given. The 9 attributes considered in the data set, and the mixture amounts of these attributes are given in Table 3. Compressive strength tests were applied on the 1st and 3rd days of the concrete series prepared considering these mixing ratios, and the obtained strength values are given in Table 4. The designed laboratory studies contain 135 data for 9 attributes and each response value (f_c,1 day and f_c,3 days). The heatmaps showing the correlation relationship of the considered attributes are presented in Figure 1. Correlation relations of the attributes are given for f_c,1 day in Figure 1.a and f_c,3 days in Figure 1.b. Figures 1.a and 1.b clearly showed similar results. Figure 1.a pointed out that there was a negative correlation between f_c,1 day, and recycled fine, coarse 1, and coarse 2. The main reason for this situation was that RCAs have high water absorption capacity and low specific gravity values. The relationship between aggregates revealed a negative correlation between natural and recycled fine. For the natural fine used in this study, the water absorption was 1.5%, and the specific gravity value was 2.66 g/cm^3, while these values were 9.8% and 2.48 g/cm^3 for recycled fine, respectively. Thus, the study's use of natural fine as fine aggregate showed positive effects on strength values. When the relationship between f_c,3 days, and natural fine, coarse 1, and coarse 2 was examined, it was seen that there was a correlation of approximately 0.75. As the use of natural aggregate increased in the study, the increase in strength values confirmed this result. Again, Figure 1.b showed a negative correlation between natural coarse 1-2 and recycled coarse 1-2. The mortar on the surface of the recycled aggregate caused low mechanical properties. Since all of the basic input attributes used in this study are numeric, no encoding was needed. However, since algorithms such as SVR and FNNR are sensitive to scale differences between attribute values, it is necessary to use normalized values. For this reason, both normalized and real-valued forms of the data set were used in the experiments. 70% of the data set was used for training and 30% for test data.

3. Experiments

In this study, "f_c,1" and "f_c,3" values were tried to be estimated according to the relevant input attributes. For this, seven different machine learning algorithms were used. Although there are different metrics for the evaluation of the regression results, R-squared (R^2) and Mean Squared Error (MSE), which are frequently used in the literature, were taken into account in this study. R^2, given in Equation (16), which is stronger against outliers, gives an idea about the variance ratio for the dependent variable. As can be seen from Equation (17), MSE shows the mean squared difference between the prediction and the true value.

Table 4. The features in the data set and response values

Mixture Code	f_c,1 day			f_c,3 days		
	Sample 1	Sample 2	Sample 3	Sample 1	Sample 2	Sample 3
NAC-1	33.2	33.9	34.2	39.8	40.1	40.2
NAC-2	30.9	30.4	30.0	37.0	36.7	37.0
NAC-3	25.2	25.9	25.7	33.7	34.5	33.8
NAC-4	20.7	20.7	21.1	28.6	28.8	29.1
NAC-5	17.2	17.2	17.4	24.2	24.6	24.8
NAC-6	36.4	36.6	36.7	43.3	44.0	43.3
NAC-7	33.2	33.1	33.5	40.3	40.4	40.6
NAC-8	28.7	29.0	28.7	37.1	37.5	37.3
NAC-9	25.0	24.8	24.9	33.1	34.0	33.5
NAC-10	19.3	19.7	20.7	27.7	28.5	28.7
NAC-11	39.5	40.0	39.8	46.5	46.9	46.9
NAC-12	36.5	36.3	36.1	44.1	43.5	44.0
NAC-13	32.8	32.9	32.4	41.8	41.1	41.5
NAC-14	28.5	28.1	28.4	37.1	37.3	36.6
NAC-15	23.7	24.4	24.3	33.0	32.3	33.2
RAC-1	19.1	18.6	18.3	23.4	23.5	23.5
RAC-2	16.5	16.1	16.9	21.6	21.3	21.2
RAC-3	14.1	14.6	14.9	19.8	19.9	20.3
RAC-4	12.3	12.6	13.0	17.5	17.5	17.9
RAC-5	10.6	10.1	10.4	14.7	14.2	14.8
RAC-6	21.9	21.2	21.1	25.8	25.9	25.4
RAC-7	18.3	17.6	18.3	23.0	22.8	23.3
RAC-8	16.1	15.7	15.6	20.0	19.8	20.4
RAC-9	13.8	13.9	14.3	18.9	18.7	18.4
RAC-10	12.0	12.6	11.7	16.9	17.1	17.1
RAC-11	23.4	23.3	24.0	28.3	28.1	28.6
RAC-12	20.4	20.5	20.0	24.9	24.1	24.2
RAC-13	18.0	17.8	17.9	22.8	22.2	22.9
RAC-14	15.1	15.8	15.6	20.0	20.3	20.7
RAC-15	13.6	13.5	13.9	18.7	18.8	18.1
RCAC-1	25.5	25.3	24.8	31.4	32.1	31.2
RCAC-2	23.3	23.4	23.1	28.9	28.2	28.6
RCAC-3	21.1	21.1	20.9	26.4	26.7	26.0
RCAC-4	19.6	19.6	19.0	24.8	24.2	24.3
RCAC-5	17.7	17.8	17.5	21.7	22.3	21.9
RCAC-6	29.3	29.4	29.2	34.8	34.9	34.1
RCAC-7	26.9	26.2	26.8	31.5	31.1	31.9
RCAC-8	23.9	23.2	23.5	28.6	28.6	29.0
RCAC-9	21.7	21.2	21.3	26.1	26.4	26.1
RCAC-10	19.5	19.3	19.2	23.7	23.8	23.5
RCAC-11	32.5	32.1	32.6	37.5	37.0	37.2
RCAC-12	29.9	29.9	29.1	34.5	34.4	34.4
RCAC-13	26.1	26.8	26.4	31.8	31.2	31.0
RCAC-14	23.3	23.6	23.3	28.3	28.2	27.9
RCAC-15	21.3	21.4	21.2	25.4	25.5	25.2



a) f_c,1 day **b)** f_c,3 day

Figure 1. The heatmap of the data set according to f_c,1 day and f_c,3 days

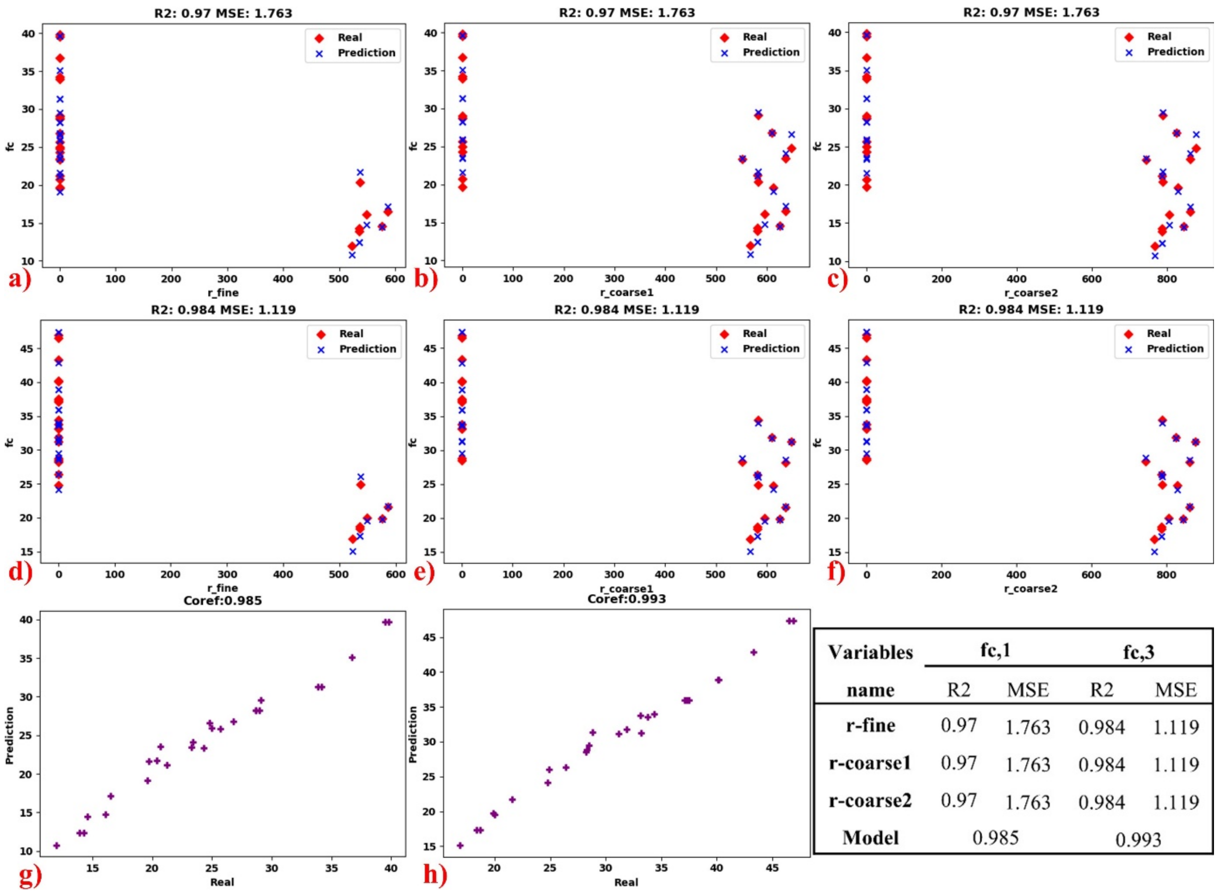
In the equations; $y, \hat{y}, \bar{y}$, and N represent the true value, the prediction, the true value mean, and the total number of data in the data set, respectively. A high (close to 1) R^2 and low MSE value shows high model performance. In general, a good regression model should aim for a high R^2 (indicating a good proportion of explained variance) and a low MSE (indicating accurate predictions). A high (close to 1) R^2 and low (close to 1) MSE value shows high model performance. A higher R^2 suggests a better-fitting model, but a high R^2 alone does not guarantee a model's correctness or practical significance. On the other hand, MSE is a useful metric for assessing the overall accuracy of the model. When these two metrics are evaluated together, a more objective assessment can be made. In addition, the correlation between the prediction results and the actual results is also discussed in the experimental results.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{16}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{17}$$

3.1. MLR Experiments

The first ML algorithm used is MLR. Figure 2 shows the model developed using MLR and the effects of the parameters used in the model on the compressive strength. Since the main purpose of the study is to examine the effects of RCAs, only the effects of RCAs are given as variables in the figures. For example, Figure 2.a shows the effect of fine RCA on the strength values and the relationship between the actual values and the predicted values. Similarly, Figure 2.d shows the effect of fine RCA on f_c,3 days. Figure 2.g and 2.h show the correlation between actual and predicted values in MLR analyses for f_c,1 day, and f_c,3 days. On the lower right side of Figure 2, the R^2 and MSE values of the model and parameters are given in detail. In MLR analyses for f_c,1 day R^2 and MSE were 0.970 and 1.763, respectively, while the values of these metrics for f_c,3 days data were 0.984 and 1.119, respectively. MLR was succeeded in obtaining successful results in both datasets. The correlation value between actual values and estimates is 0.985 for f_c,1 day and 0.993 for f_c,3 days. Detailed evaluation of MLR performance together with other methods will be presented in the discussion section.



Variables	fc,1		fc,3	
name	R2	MSE	R2	MSE
r-fine	0.97	1.763	0.984	1.119
r-coarse1	0.97	1.763	0.984	1.119
r-coarse2	0.97	1.763	0.984	1.119
Model	0.985		0.993	

Figure 2. Prediction results on compressive strength of models and parameters developed using MLR.

3.2. Polynomial Regression Experiments

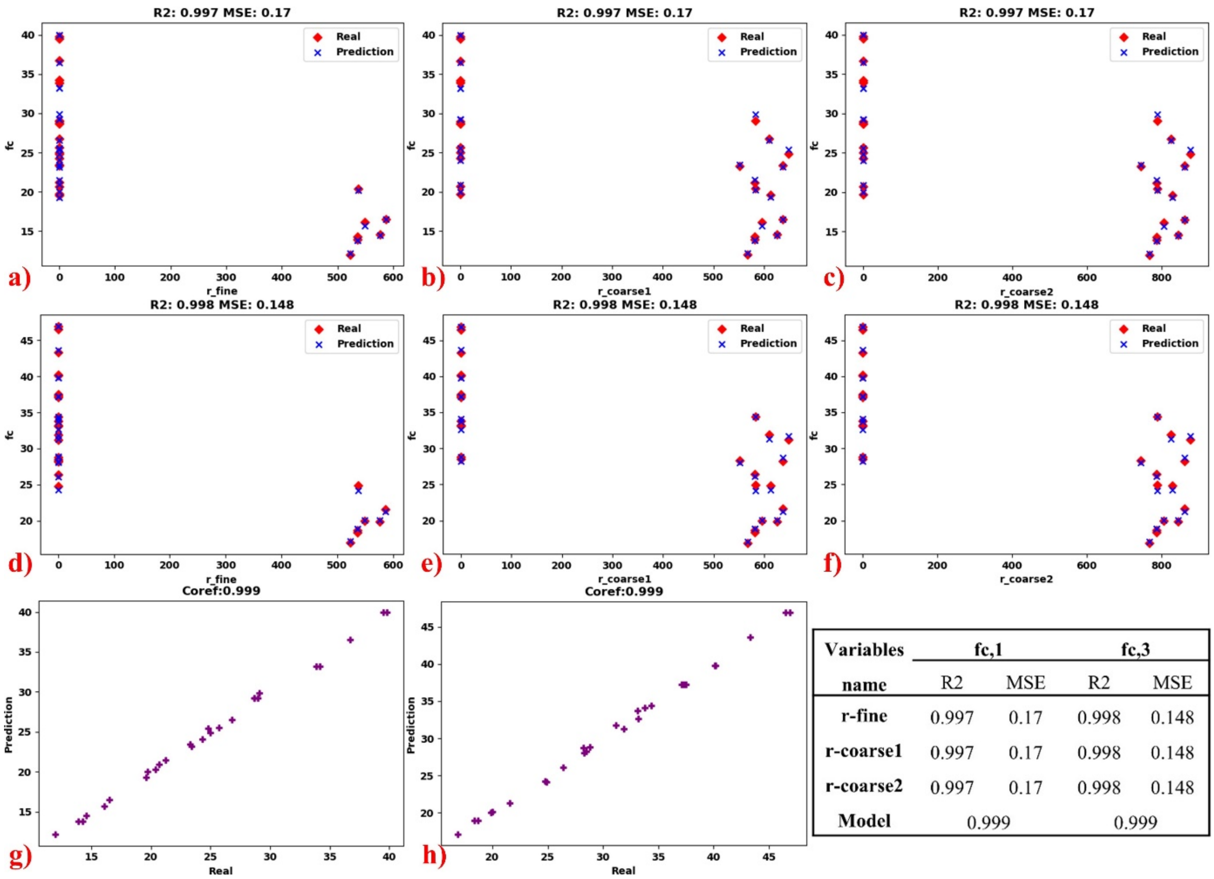
The second ML algorithm used is polynomial regression analysis. While performing the polynomial regression analysis, the best result was tried to be achieved by using different polynomial degrees. R^2 and MSE results obtained for degrees from 1 to 8 are shared in Table 5.

Table 5. Polynomial regression analysis results for different polynomial degrees

		Degree							
Class		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
f_c,1 day	***R*²**	0.970	0.982	0.993	0.997	0.997	>> -1000	>> -1000	>> -1000
	MSE	1.763	1.047	0.429	0.170	0.171	654	929	973
f_c,3 days	***R*²**	0.984	0.998	0.997	0.998	0.998	>> -1000	>> -1000	>> -1000
	MSE	1.119	0.174	0.203	0.148	0.149	>> 1000	>> 1000	>> 1000

3.3. SVR Experiments

The third ML algorithm used is the SVR method. Three types of estimation algorithms, linear, polynomial and RBF regression, were used while performing this analysis. Normalized values of data sets were used in the analysis. The R^2 and MSE values obtained as a result of the analyses are given in Table 6. The best R^2 and MSE values for both datasets were obtained with SVR using the polynomial kernel function, albeit with a small difference. The metric values of f_c for 1 day are 0.995 and 0.005, respectively, while f_c is 0.998 and 0.003 for 3 days. The lowest performance for both datasets was obtained in SVR using the linear kernel. Figure 4 shows the detailed prediction results on the compressive strength of the model developed using the SVR method and the parameters used in the model.



Variables name	fc,1 R2	fc,1 MSE	fc,3 R2	fc,3 MSE
r-fine	0.997	0.17	0.998	0.148
r-coarse1	0.997	0.17	0.998	0.148
r-coarse2	0.997	0.17	0.998	0.148
Model	0.999		0.999	

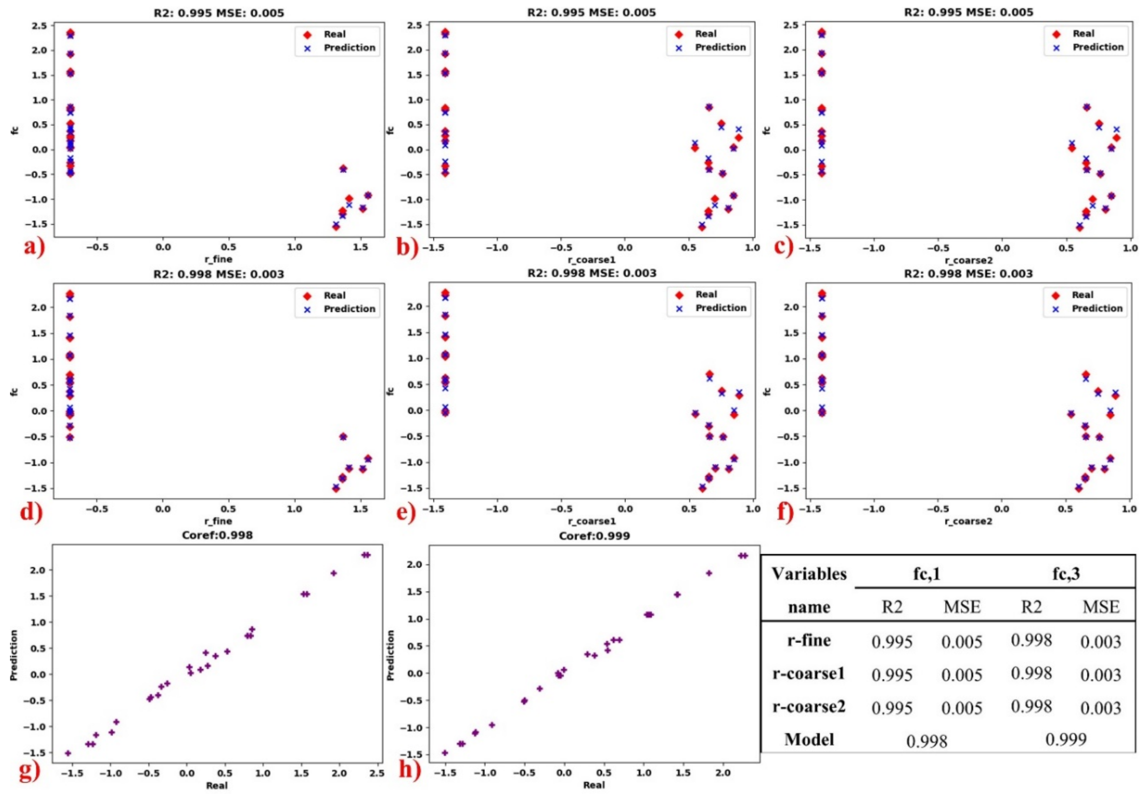
Figure 3. Prediction results on compressive strength of model developed using polynomial regression and the relevant parameters.

Table 6. SVR regression analysis results for different kernels

Class	Kernel	R^2	MSE
f_c,1 day	Linear	0.970	0.034
	Polynomial	0.995	0.005
	RBF	0.995	0.006
f_c,3 days	Linear	0.985	0.003
	Polynomial	0.998	0.003
	RBF	0.996	0.003

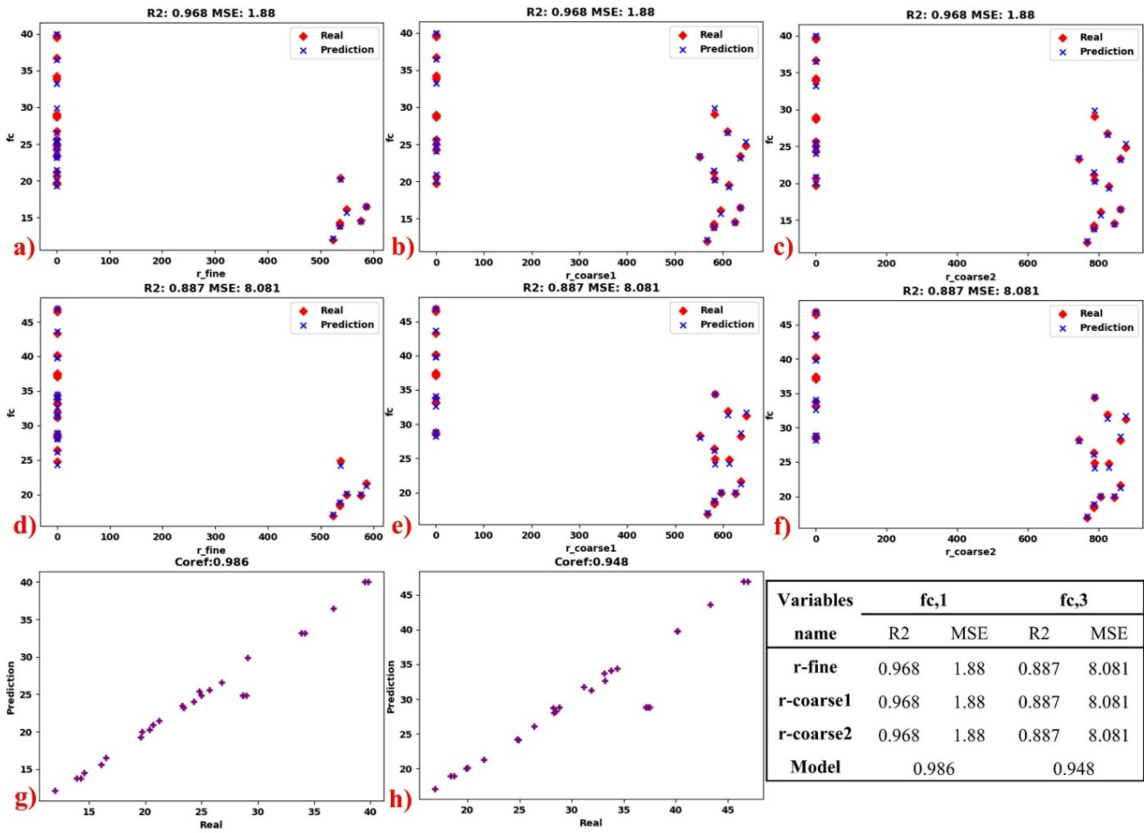
3.4. DTR Experiments

The fourth ML algorithm used is the DTR method. Looking at the graphs of data set attributes and prediction results given in Figure 5, it is seen that the error increases even more in predictions where the true value is greater than 25. In particular, the prediction errors between 25-35 were higher.



Variables name	fc,1 R2	fc,1 MSE	fc,3 R2	fc,3 MSE
r-fine	0.995	0.005	0.998	0.003
r-coarse1	0.995	0.005	0.998	0.003
r-coarse2	0.995	0.005	0.998	0.003
Model	0.998		0.999	

Figure 4. Compressive strength prediction results of models developed with SVR algorithm and its parameters.



Variables name	fc,1 R2	fc,1 MSE	fc,3 R2	fc,3 MSE
r-fine	0.968	1.88	0.887	8.081
r-coarse1	0.968	1.88	0.887	8.081
r-coarse2	0.968	1.88	0.887	8.081
Model	0.986		0.948	

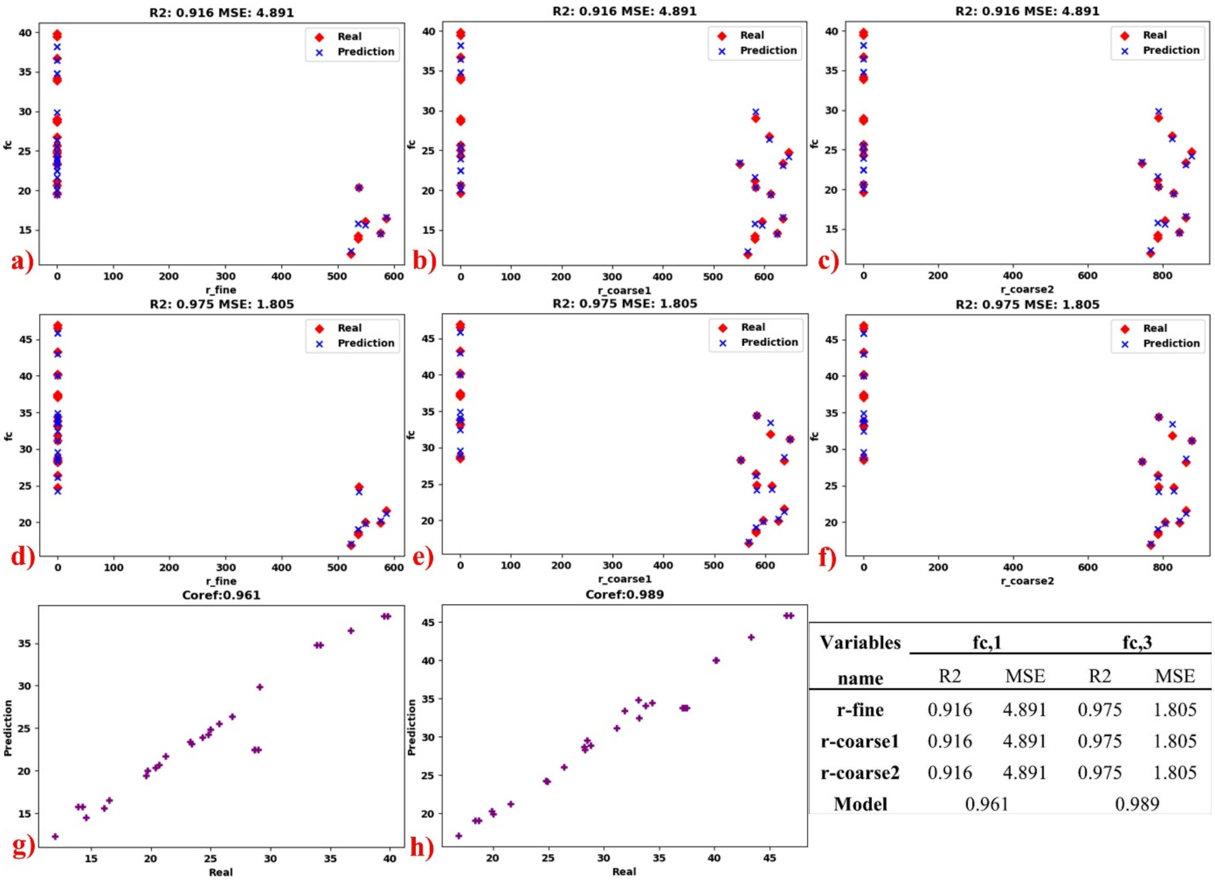
Figure 5. Compressive strength prediction results of models developed with DTR algorithm and its parameters.

3.5. RFR Experiments

In the predictions using the RF method, firstly, experiments were carried out for different tree numbers and the effect of the number of trees on the prediction results was examined. The metric results obtained are shared in Table 7. The increase in the number of trees was not a factor in increasing the performance. The best R^2 and MSE values for f_c, 1-day data were 0.916 and 4.891, respectively, and these values were obtained in the 2-tree RFR model. For these data, the increase in the number of trees affected the performance negatively. The best R^2 and MSE values in experiments for f_c,3 days were 0.975 and 1.805, respectively. These results were captured with the 10-tree RFR model.

Table 7. RFR regression analysis results for different tree numbers

		Tree Number					
Class		**2**	**3**	**5**	**10**	**20**	**30**
f_c,1 day	**R^2**	0.916	0.870	0.908	0.891	0.893	0.910
	MSE	4.891	7.885	5.375	6.355	6.255	5.26
f_c,3 days	**R^2**	0.932	0.911	0.951	0.975	0.937	0.949
	MSE	4.869	6.309	3.474	1.805	4.476	3.66



Variables name	fc,1 R2	fc,1 MSE	fc,3 R2	fc,3 MSE
r-fine	0.916	4.891	0.975	1.805
r-coarse1	0.916	4.891	0.975	1.805
r-coarse2	0.916	4.891	0.975	1.805
Model	0.961		0.989	

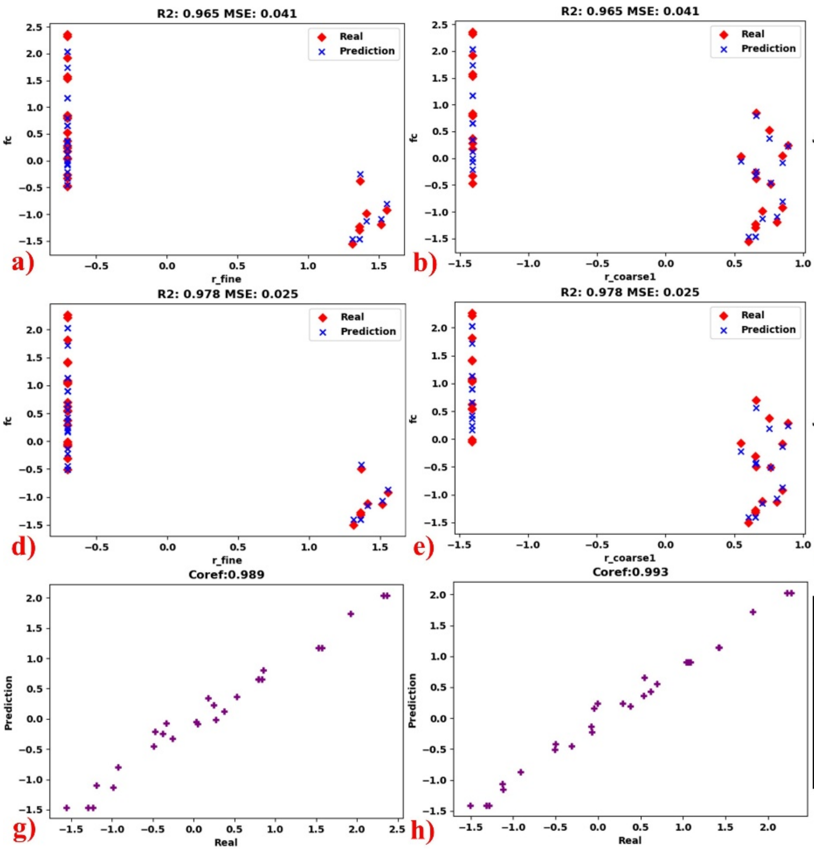
Figure 6. Compressive strength prediction results of models developed with RFR algorithm and its parameters.

3.6. KNN Experiments

The KNN regression experiments were carried out for 2 to 10 neighbours. The metric results of these values are given in Table 8. As can be seen in Figure 7, the prediction errors are relatively noticeable at the values where the actual value is less than 25.

Table 8. KNN regression results for different neighbour numbers

		Neighbour Number								
Class		**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
f_c,**1 day**	R^2	0.957	0.965	0.949	0.942	0.946	0.924	0.914	0.907	0.910
	MSE	0.050	0.041	0.060	0.067	0.063	0.089	0.101	0.108	0.104
f_c,**3 days**	R^2	0.973	0.978	0.965	0.960	0.961	0.945	0.939	0.932	0.937
	MSE	0.031	0.025	0.040	0.045	0.044	0.062	0.068	0.077	0.071

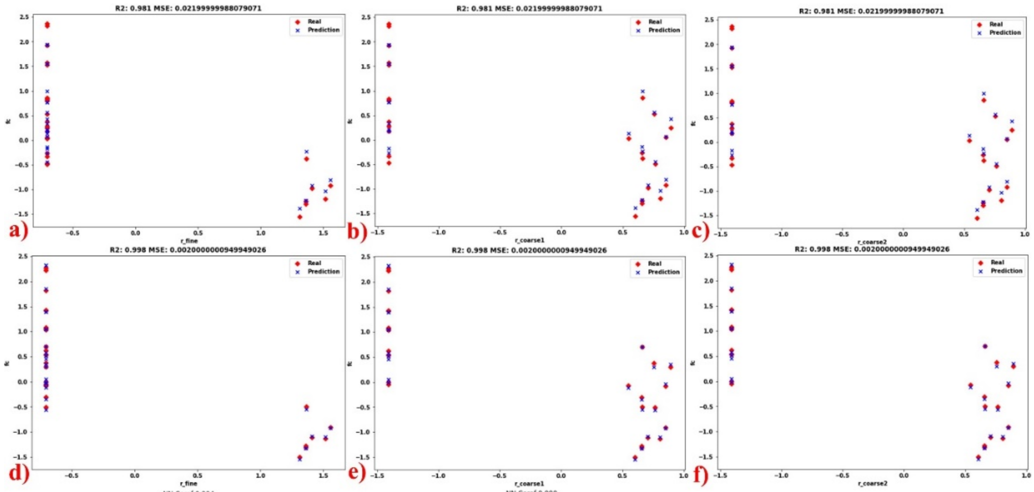


Variables name	fc,1 R2	fc,1 MSE	fc,3 R2	fc,3 MSE
r-fine	0.965	0.041	0.978	0.025
r-coarse1	0.965	0.041	0.978	0.025
r-coarse2	0.965	0.041	0.978	0.025
Model	0.989		0.993	

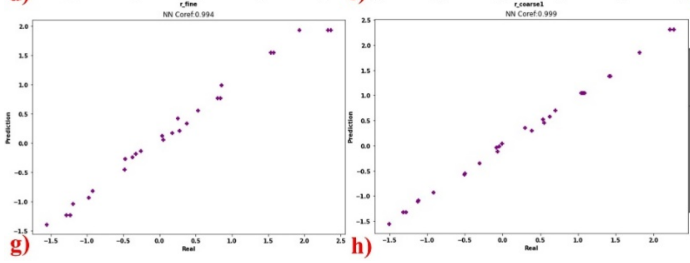
Figure 7. Compressive strength prediction results of models developed with KNN algorithm and its parameters.

3.7. FFNNR Experiments

The last regression algorithm used in this study is FFNNR [32]. It can be seen from Figure 8 that the f_c,3 days predictions were more successful. However, FFNN models have many hyper-parameters that affect the result. Although optimizing these parameters is the subject of another study, we observed the effect of different LR values of the model developed. The metric results obtained in these additional tests are given in Figure 9. In the analyses, R^2 and MSE values were observed according to LR change. The most successful LR values were found to be 0.089 for f_c,1 day, and 0.078 for f_c,3 days. The best R^2 and MSE values obtained were also captured with these LRs.



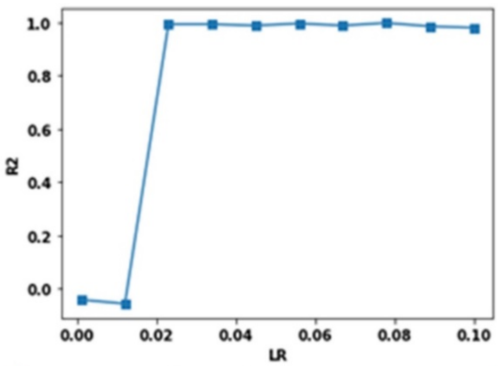
a) b) c) d) e) f)



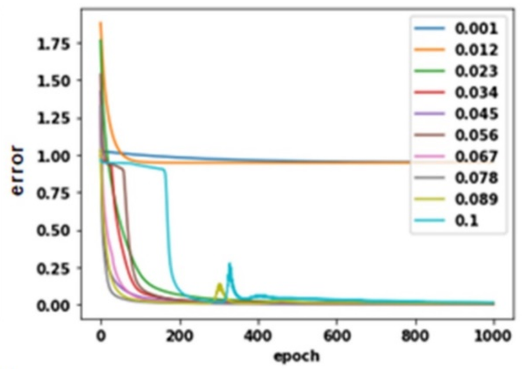
g) h)

Variables name	fc,1 R2	fc,1 MSE	fc,3 R2	fc,3 MSE
r-fine	0.981	0.022	0.998	0.002
r-coarse1	0.981	0.022	0.998	0.002
r-coarse2	0.981	0.022	0.998	0.002
Model	0.994		0.999	

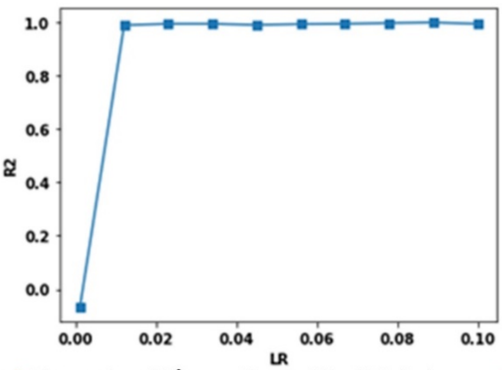
Figure 8. Compressive strength prediction results of models developed with FFNNR algorithm and its parameters.



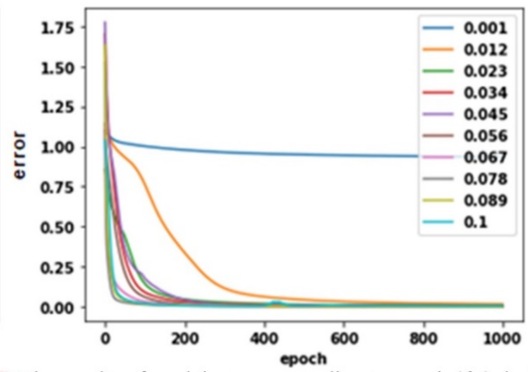
a) The results of R^2 according to LRs (f_c,1 day)



b) The results of model errors according to epoch (f_c,1 day)



c) The results of R^2 according to LRs (f_c,3 day)



d) The results of model errors according to epoch (f_c,3 day)

Figure 9. The results of R^2 and model errors for different LRs.

4. Discussion

The performance of the MLR method is high in both data sets. It is also possible to examine the strength value predicts of RCA in terms of attributes. For example, the actual and predicted values of f_c,1 day were compared according to the r_fine, r_coarse1 and r_coarse2 attributes. As can be seen in these comparisons, the performance was even higher at values greater than zero for these attributes. The prediction errors often occurred on data where these metric values were 0. In both data sets, the prediction performance of all values where the r_fine parameter was different from zero was more successful than the non-zero values of the other parameters. Prediction errors were generally between 25-35. In the polynomial regression experiments, the best R^2 value for f_c, 1 day was 0.997 and the MSE value was 0.170. These values were 0.998 and 0.148 for f_c,3 days, respectively. The most successful value for f_c, 1 day was obtained when the polynomial degree was 4. Increasing the degree of polynomial initially increased the performance, but after the 6th order, the performance deteriorated. A similar situation was observed for f_c, 3. In both f_c,1 day and f_c,3 day experiments, the correlation value reached a very high value of 0.999. In both data sets, the prediction performance of all values where the r_fine parameter was different from zero was more successful than the non-zero values of the other parameters. Prediction errors were generally between 25-35.

In SVR, the correlation value for f_c,1 is 0.998, while f_c,3 is 0.999. f_c performs are better at non-zero r_fine values in the 3-days data set. Prediction errors occurred while f_c,3 days was generally between 25-35, and f_c,1 day was below 25. Besides, in experiments with the DT model, R^2 values of RCAs for f_c,1 day and f_c,3 day were 0.968-0.887, and MSE values were 1.880-8.081, respectively. These results were lower than the previously used models. However, the correlation values were 0.986 for f_c,1 day, and 0.948 for f_c,3 day. In RFR, increasing the number of trees for the data sets did not have a clear effect on the result. Compared to DT, performance decreased for f_c,1 day and increased for f_c,3 day. However, overall performance was poor compared to other methods. According to the prediction results of f_c,1 day and f_c,3 days, the correlation values were 0.961 and 0.989, respectively. Looking at the prediction comparisons, it was seen that the prediction errors are at the points where the true value was between 25-35. The best R^2 and MSE values of K-NN for both f_c,1 day and f_c,3 days were obtained by evaluating the 3 nearest neighbors. As a result of f_c,1 day experiments, the best R^2, and MSE values are 0.965-0.050, while these values for f_c,3 days were 0.978-0.025. The increase in the number of neighbours negatively affected the performance. The correlations between the predicted and actual values of the f_c,1 day, and f_c,3 days experiments were 0.989 and 0.993, respectively. R^2 and MSE values obtained in regression analyses with two-layer FFNNR, each containing 30 nodes, were 0.981-0.022 for f_c,1 day and 0.998-0.020 for f_c,3 days. The correlation results for both data sets were 0.994 and 0.999 respectively. Accordingly, the results of the FFNNR analysis for f_c,3 days were more successful.

Observed variability in performance between different machine learning algorithms on the same dataset can be attributed to a variety of factors that reflect the inherent complexities and nuances associated with the algorithms and dataset. When the results are analyzed, it is seen that there is a non-linear relationship in the data set. In such data spaces, the performance of models such as Polynomial Regression and SVR may come to the forefront. In general, although a certain success was achieved in all models, the nonlinear relationship was captured even better with Polynomial Regression, SVR and FFNNR.

5. Conclusions

This study aimed to adapt seven different well-known machine learning regression algorithms, rather than a single method, to predict with high accuracy the early wet compressive strength of concrete mixtures containing different proportions of recycled concrete aggregates. This paper is also unique in that it simultaneously predicts the early age compressive strength of concrete produced using construction demolition waste using seven different machine methods. The reduction of environmental pollution and the effect of recycled concrete aggregates on strength by enabling waste materials to be reduced, recycled, and reused by using post-earthquake recycled concrete aggregates were examined in detail.

While the best performance for f_c, 1 day was given by Polynomial regression, SVR, and FFNNR, respectively, in the experiments, these three methods achieved approximately the same success in the f_c,3 days experiments. In general, the 4th order Polynomial regression model stands out as the most appropriate model. The lowest performance for f_c,1 day was observed in RFR analysis, and the lowest performance for f_c,3 days was observed in DTR. Especially considering that thousands of new buildings were built after the earthquake in Elazig, it becomes even more important to estimate the early-age compressive strength using different machine learning algorithms. It is thought that evaluating these wastes, which arise in terms of sustainable development and circular economy, will provide significant environmental and economic gains by reducing the consumption of natural resources and reusing these materials. The authors will focus on metaheuristic optimization-based machine learning methods in their future studies in order to obtain better results in terms of different metrics.

Acknowledgement

This research is supported by the Scientific Research Project Fund of Firat University under the project number MF.21.52.

References

[1] Oikonomou ND. Recycled concrete aggregates, Cement and Concrete Composites 2005; 27: 315–318. https://doi.org/10.1016/j.cemconcomp.2004.02.020.

[2] Farina I, Colangelo F, Petrillo A, Ferraro A, Moccia I, Cioffi R. LCA of concrete with construction and demolition waste, in: Advances in Construction and Demolition Waste Recycling, Elsevier 2020; 501–513. https://doi.org/10.1016/B978-0-12-819055-5.00024-3.

[3] de Andrade Salgado F, de Andrade Silva F. Recycled aggregates from construction and demolition waste towards an application on structural concrete: A review, Journal of Building Engineering 2022; 52: 104452. https://doi.org/10.1016/j.jobe.2022.104452.

[4] Abed M, Fořt J, Rashid K. Multicriterial life cycle assessment of eco-efficient self-compacting concrete modified by waste perlite powder and/or recycled concrete aggregate, Construction and Building Materials 2022; 348: 128696. https://doi.org/10.1016/j.conbuildmat.2022.128696.

[5] Qin D, Zong Z, Dong C, Guo Z, Tang L, Chen C, Zhang L. Long-term behavior of sustainable self-compacting concrete with high volume of recycled concrete aggregates and industrial by-products, Structural Concrete 2023; 24: 3385–3404. https://doi.org/10.1002/suco.202200464.

[6] Yan Y, Gao D, Yang L, Pang Y, Zhang Y. Evaluation method of shear toughness for steel fiber-reinforced concrete containing recycled coarse aggregate, Structural Concrete 2023; 24: 2879–2893. https://doi.org/10.1002/suco.202200324.

[7] Kapoor K, Singh SP, Singh B. Improving the durability properties of self-consolidating concrete made with recycled concrete aggregates using blended cements, International Journal of Civil Engineering 2021; 19: 759–775. https://doi.org/10.1007/s40999-020-00584-7.

[8] Marinković S, Radonjanin V, Malešev M, Ignjatović I. Comparative environmental assessment of natural and recycled aggregate concrete, Waste Management 2010; 30: 2255–2264. https://doi.org/10.1016/j.wasman.2010.04.012.

[9] Tao X, Zhang J, Zhang Y, Li X, Zhang M. Bond Behavior Between High-Strength Recycled Aggregate Concrete and UHSSB Using Beam Test, International Journal of Civil Engineering 2022; 20: 1129-1144. https://doi.org/10.1007/s40999-022-00719-y.

[10] Wang J, Xu Q. The combined effect of load and corrosion on the flexural performance of recycled aggregate concrete beams, Structural Concrete 2023; 24: 359–373. https://doi.org/10.1002/suco.202100819.

[11] Yang C, Feng H, Esmaeili-Falak M. Predicting the compressive strength of modified recycled aggregate concrete, Structural Concrete 2022. https://doi.org/10.1002/suco.202100681.

[12] Ulucan M, Alyamac KE. A holistic assessment of the use of emerging recycled concrete aggregates after a destructive earthquake: Mechanical, economic and environmental, Waste Management 2022; 146: 53–65. https://doi.org/10.1016/j.wasman.2022.04.045.

[13] Ulucan M, Alyamac KE. An integrative approach of the use of recycled concrete aggregate in high-rise buildings: example of the Elysium, Structural Concrete 2023; 24: 3329–3350. https://doi.org/10.1002/suco.202200512.

[14] Kul A, Ozel BF, Ozcelikci E, Gunal MF, Ulugol H, Yildirim G, Sahmaran M. Characterization and life cycle assessment of geopolymer mortars with masonry units and recycled concrete aggregates assorted from construction and demolition waste, Journal of Building Engineering 2023; 78: 107546. https://doi.org/10.1016/j.jobe.2023.107546.

[15] Ilcan H, Sahin O, Unsal Z, Ozcelikci E, Kul A, Demiral NC, Ekinci MO, Sahmaran M. Effect of industrial waste-based precursors on the fresh, hardened and environmental performance of construction and demolition wastes-based geopolymers, Construction and Building Materials 2023; 394: 132265. https://doi.org/10.1016/j.conbuildmat.2023.132265.

[16] Özçelikci E, Oskay A, Bayer İR, Şahmaran M. Eco-hybrid cement-based building insulation materials as a circular economy solution to construction and demolition waste, Cement and Concrete Composites 2023; 141: 105149. https://doi.org/10.1016/j.cemconcomp.2023.105149.

[17] Ulucan M, Alyamac KE. A comprehensive assessment of mechanical and environmental properties of green concretes produced using recycled concrete aggregates and supplementary cementitious material, Environmental Science and Pollution Research 2023; 30: 97765–97785. https://doi.org/10.1007/s11356-023-29197-y.

[18] Ulucan M, Yildirim G, Alatas B, Alyamac KE, A new intelligent sunflower optimization based explainable artificial intelligence approach for early-age concrete compressive strength classification and mixture design of RAC, Structural Concrete 2023; 24: 7400-7418. https://doi.org/10.1002/suco.202300138.

[19] Ulucan M, Tas Y, Alyamac KE, Multi-objective optimization and assessment of recycled concrete aggregates for sustainable development: Example of the Kömürhan bridge, Structural Concrete 2023; 24: 5750–5768. https://doi.org/10.1002/suco.202201018.

[20] Han B, Wu Y, Liu L. Prediction and uncertainty quantification of compressive strength of high-strength concrete using optimized machine learning algorithms, Structural Concrete 2022; 23: 3772–3785. https://doi.org/10.1002/suco.202100732.

[21] Kandiri A, Golafshani EM, Behnood A. Estimation of the compressive strength of concretes containing ground granulated blast furnace slag using hybridized multi-objective ANN and salp swarm algorithm, Construction and

Building Materials 2020; 248: 118676. https://doi.org/10.1016/j.conbuildmat.2020.118676.
[22] Zhang J, Huang Y, Wang Y, Ma G. Multi-objective optimization of concrete mixture proportions using machine learning and metaheuristic algorithms, Construction and Building Materials 2020; 253: 119208. https://doi.org/10.1016/j.conbuildmat.2020.119208.
[23] Golafshani EM, Arashpour M, Behnood A. Predicting the compressive strength of green concretes using Harris hawks optimization-based data-driven methods, Construction and Building Materials 2022; 318: 125944. https://doi.org/10.1016/j.conbuildmat.2021.125944.
[24] Tam VWY, Butera A, Le KN, Da Silva LCF, Evangelista ACJ. A prediction model for compressive strength of CO2 concrete using regression analysis and artificial neural networks, Construction and Building Materials 2022; 324: 126689. https://doi.org/10.1016/j.conbuildmat.2022.126689.
[25] Amiri M, Hatami F. Prediction of mechanical and durability characteristics of concrete including slag and recycled aggregate concrete with artificial neural networks (ANNs), Construction and Building Materials 2022; 325: 126839. https://doi.org/10.1016/j.conbuildmat.2022.126839.
[26] Adam SP, Alexandropoulos SAN, Pardalos PM, Vrahatis MN, No free lunch theorem: A review, Approximation and optimization 2019; 57–82. https://doi.org/10.1007/978-3-030-12767-1_5.
[27] Jobson JD. Multiple linear regression, in: Applied Multivariate Data Analysis, Springer 1991: 219–398. https://doi.org/10.1007/978-1-4612-0955-3_4.
[28] Jobson JD. Applied multivariate data analysis: volume II: Categorical and Multivariate Methods, Springer Science and Business Media, 2012. https://doi.org/10.1007/0-387-22753-9_8.
[29] Awad M, Khanna R. Support vector regression, Efficient Learning Machines 2015: 67-80. https://doi.org/10.1007/978-1-4302-5990-9_4.
[30] Ince K, Klawonn F. Decision and regression trees in the context of attributes with different granularity levels, Towards Advanced Data Analysis by Combining Soft Computing and Statistics 2013: 331–342. https://doi.org/10.1007/978-3-642-30278-7_26.
[31] Vanneschi L, Castelli M, Manzoni L, Silva S, Trujillo L. Is k Nearest Neighbours Regression Better Than GP?, European Conference on Genetic Programming 2020: 244–261. https://doi.org/10.1007/978-3-030-44094-7_16.
[32] Caselli M, Trizio L, De Gennaro G, Ielpo P. A simple feedforward neural network for the PM10 forecasting: comparison with a radial basis function network and a multivariate linear regression model, Water Air and Soil Pollution 2009; 201: 365–377. https://doi.org/10.1007/s11270-008-9950-2.