buildings

MDPI

Article

Strength Reduction Due to Acid Attack in Cement Mortar Containing Waste Eggshell and Glass: A Machine Learning-Based Modeling Study

Fei Zhu [1,2], Xiangping Wu [3,*], Yijun Lu [4] and Jiandong Huang [4,*]

[1] School of Fine Arts, Suzhou Vocational University, Suzhou 215104, China; zhufeiedu@163.com
[2] School of Civil Engineering, China University of Mining and Technology, Xuzhou 221116, China
[3] Department of Gem Design Engineering, KAYA University, Gimhae 50830, Republic of Korea
[4] School of Civil Engineering, Guangzhou University, Guangzhou 511370, China; yijun.lu@e.gzhu.edu.cn
* Correspondence: xiangping_wu@163.com (X.W.); jiandong.huang@hotmail.com or jh@gzhu.edu.cn (J.H.)

Abstract: The present study utilized machine learning (ML) techniques to investigate the effects of eggshell powder (ESP) and recycled glass powder (RGP) on cement composites subjected to an acidic setting. A dataset acquired from the published literature was employed to develop machine learning-based predictive models for the cement mortar's compressive strength (CS) decrease. Artificial neural network (ANN), K-nearest neighbor (KNN), and linear regression (LR) were chosen for modeling. Also, RreliefF analysis was performed to study the relevance of variables. A total of 234 data points were utilized to train/test ML algorithms. Cement, sand, water, silica fume, superplasticizer, glass powder, eggshell powder, and 90 days of CS were considered as input variables. The outcomes of the research showed that the employed models could be applied to evaluate the reduction percentage of CS in cement composites, including ESP and RGP, after being exposed to acid. Based on the R^2 values (0.87 for the ANN, 0.81 for the KNN, and 0.78 for LR), as well as the assessment of variation between test values and anticipated outcomes and errors (1.32% for ANN, 1.57% for KNN, and 1.69% for LR), it was determined that the accuracy of the ANN model was superior to the KNN and LR. The sieve diagram exhibited a correlation amongst the model predicted and target results. The outcomes of the RreliefF analysis suggested that ESP and RGP significantly influenced the CS loss of samples with RreliefF scores of 0.26 and 0.21, respectively. Based on the outcomes of the research, the ANN approach was determined suitable for predicting the CS loss of mortar subjected to acidic environments, thereby eliminating lab testing trails.

Keywords: cement mortar; eggshell and glass powder; acid attack

Citation: Zhu, F.; Wu, X.; Lu, Y.; Huang, J. Strength Reduction Due to Acid Attack in Cement Mortar Containing Waste Eggshell and Glass: A Machine Learning-Based Modeling Study. *Buildings* **2024**, *14*, 225. https://doi.org/10.3390/buildings14010225

Academic Editor: Abdelhafid Khelidj

Received: 21 December 2023
Revised: 4 January 2024
Accepted: 10 January 2024
Published: 14 January 2024

1. Introduction

Cement-based mixtures have a broad application in the building industry worldwide, second only to water in annual usage [1]. The field of sustainable construction materials has been increasingly recognized and studied. The durability of cement-based materials, documented by various sources, is a key reason for their extensive use in construction [2]. Durable cementitious composites can maintain their reliability, structural performance, and functionality even when subjected to harsh environmental conditions [3]. Durability refers to the inherent ability of cement-based materials to sustain chemical attacks, abrasion, weathering, and various forms of degradation over an extended period [4]. Several hostile environmental factors subsidize the deterioration of concrete compounds. The detrimental mechanism can manifest in several mechanical, chemical, or physical forms and may be initiated by interior or exterior causes. The impact of the physical attack on composite aggregates and pastes is influenced by the specific characteristics of the attack, involving both chemical and physical attacking procedures. The primary concern lies in the functionality of cementitious materials when subjected to severe environmental conditions.

This is because the strength and durability of cementitious composites tend to deteriorate when exposed to various forms of detrimental agents [5]. As a consequence of the swift development of the industrial zone, cement-based compounds are commonly exposed to various forms of degradation caused by salts, acids, sulfate, and other deleterious elements. The current focus of the construction sector is to create resilient and sustainable buildings that possess the potential to endure challenging environmental conditions and adhere to longevity criteria [6].

The longevity of cement-based mixtures is predominantly determined by their resistance to the penetration of aggressive ions. The porosity of cement-based mixtures can be indirectly inferred from their absorption qualities, which are determined by the permeable pore volume and the interconnectivity of pores [7]. Sulfuric acid, chemically represented as H_2SO_4, is a highly corrosive acid that can cause structural corrosion [8]. The corrosive nature of H_2SO_4 may substantially threaten construction materials, leading to accelerated degradation and even structural failure. The alkaline nature of the pore water in cementitious compounds leads to a reaction with mild or strong acids. This reaction involves the calcium-silicate-hydrate (CSH) gel and $Ca(OH)_2$, present in the cement matrix, generating dissolved ions. Consequently, this dissolution process leads to the breakdown of concrete components [5]. H_2SO_4 is widely recognized as a highly hazardous substance when it interacts with cement-based materials due to its corrosive effect, further exacerbated by sulfate ions [8]. The degradation of cementitious composites and its impacts on durability is significantly influenced by sulfate attacks caused by the physical and chemical assault of sulfates and salt crystallization [9]. The durability issues in cement-based compounds can be attributed to the decalcification of CSH and the subsequent formation of other compounds [9]. Hence, it is essential to thoroughly examine the potential degradation of cement-based materials resulting from exposure to acid during their operational lifespan.

Several businesses generate eggshell trash, including restaurants, residences, and bakeries. It has been suggested by multiple sources that prolonged deposition of eggshell waste in landfills without appropriate management may lead to the development of diverse allergic reactions [10]. Improper disposal of eggshells causes persistent problems, including a strong and unpleasant odor [11]. One potential resolution to this matter involves the utilization of eggshell powder (ESP) within construction materials [12]. There needs to be more research conducted to analyze the effects of ESP-modified cement-based materials in a hostile environmental setting. An experiment conducted by Binici et al. [13] involved subjecting cement-based compounds, which contained unprocessed ESP as a replacement for sand, to a 10% Na_2SO_4 solution for 90 days. During this period, CS and flexural strength (FS) changes were recorded. Using ESP proportion as a substitute for sand reduced the CS and FS of submerged samples when exposed to the sulfate solution. The composite exhibited diminished strength performance due to the ESP's worse strength features than sand. In their study, Wei et al. [14] inspected the effects of incorporating unprocessed ESP into cement mortar. Specifically, they examined the changes in weight and CS deterioration after the mortar specimens were exposed to a 5% Na_2SO_4 solution. Using ESP led to increased weight and CS loss, contrary to the control specimen. Moreover, the CS and weight reduction were more pronounced when the ESP proportion increased as a cement substitute. The available research demonstrates that applying unprocessed ESP in cement-based materials diminishes the capacity to withstand acid attack. Nevertheless, reports have indicated that using treated ESP, such as calcinated ESP, along with the incorporation of a pozzolanic compound, could potentially enhance the properties of cement-based compounds [15]. The process of obtaining treated ESP is shown in Figure 1.

Figure 1. Converting eggshell waste to treated eggshell powder [16].

Manufacturing cement demands a substantial amount of energy and results in notable carbon dioxide (CO_2) emissions, making a significant contribution to global warming [17]. Cement companies can reduce costs and CO_2 emissions by implementing waste recycling and reuse practices [18]. Hence, a significant need exists for environmentally sustainable cement-based materials within the building sector [19,20]. In addition, extracting natural aggregates produces a significant amount of CO_2 and contributes to the diminution of natural resources [21]. Recycled glass powder (RGP) modified cement-based composites are increasingly attractive in the building industry, primarily attributed to their cost-effectiveness and wide accessibility. The usage of RGP in cement or sand at a substitution rate of 10–20% has been shown to enhance mechanical qualities, specifically CS and FS [22,23]. The study of Qaidi et al. [24] gathered various investigations that incorporated RGP as a partial or full alternative of cement to assess its hardened and fresh properties. The research of Zeybek, Ö. et al. [25] claimed that substituting 20% of RGP in cement showed optimum results for the mechanical properties of RGP-based concrete. Furthermore, it is concluded by the study that a 10% substitution of RGP with aggregates showed optimum results. Çelik et al. [26] used various proportions of RGP (i.e., 10% to 50%) as an alternative to fine and coarse aggregates. The conclusion of the study suggests that a 20% replacement of RGP showed optimum results. Karalar et al. [27] have studied the flexural behaviors of waste marble-based (WMP) concretes. The results suggest that when the quantity of WMP in the concrete mix is raised from 0% to 40%, it is observed that the nature of the crack shifts from shear crack to flexural crack as the proportion of WMP increases in the mix ratio. The research of Özkılıç, Y.O. et al. [28] utilized WG with fly-ash in geopolymer concrete to assess its fresh and mechanical properties. It is recommended by the study to incorporate 10% glass aggregate with a NaOH molarity of 16 in order to achieve the most sustainable GPC, taking into account both fresh and mechanical attributes. Çelik, A.İ. et al. [29] studied the behavior of incorporating RGP with fly ash in geopolymer concrete. Consequently, substituting sand

with RGP offers potential benefits, such as the conservation of natural resources, improved waste management, and reduced cement consumption and CO_2 emissions [30].

Experts are currently developing forecasting models to evaluate the performance of materials and structures. The primary intent of these techniques is to minimize the need for excessive experimentation in laboratory settings [31–33]. Traditional estimation approaches, such as regression-based models, are commonly employed to assess the properties of construction materials [34,35]. Artificial intelligence (AI) methodologies, such as machine learning (ML), are currently leading the way in the advancement of modeling techniques within this domain [36,37]. ML methods for foreseeing material performance are becoming increasingly prominent [38,39]. While prior research in the field of ML has primarily focused on evaluating the performance of traditional cementitious materials [40,41], there is limited literature that specifically addresses the prediction of properties in cementitious composites modified with ESP and RGP. Wang et al. [42] developed ML models to forecast the CS loss after an acid-attack test on ESP-based mortar. Similarly, Alfaiad et al. [43] claimed to address the CS loss in an acidic environment after performing lab tests on RGP-based mortar. The experimental data were then utilized to develop bagging and random forest ML models. Jin et al. [44] studied the application of genetic programming-based ML models to anticipate the percentage loss in CS when ESP-based mortar samples are subjected to an acidic environment. Amin et al. [45] performed experimental and ML-based analyses of RGP-based concrete to study the flexural behavior of specimens. Khan et al. [46] developed ML models, including AdaBoost and decision tree models, to access the CS of RGP-based concrete after performing lab experimentations. The experimental study followed by the ML models for assessing the impact of ESP on the water absorption of concrete is carried out by Khan et al. [16].

To counter the environmental degradation caused by the continuous use of cement, engineers are focusing on promoting sustainable materials like ESP and RGP-based concrete. Various investigations have been carried out in the past to check the suitability of these materials in harsh environments. Although these studies focused on experimental trials to study the characteristics of ESP and RGP-based materials, limited studies have demonstrated the use of novel ML techniques, i.e., linear regression (LR), artificial neural network (ANN), and K-nearest neighbor (KNN), in this domain. This study aims to assess the fraction loss in CS subjected to acid attack in cement mortar modified with ESP and RGP through a combination of ML techniques. ML prediction models were constructed using a dataset obtained through research articles published in reputed journals. The study employed three ML approaches, including LR, ANN, and KNN, to achieve its objectives. Furthermore, the interaction of input parameters with the target variable is studied via RreliefF analysis. The outcome of the RreliefF analysis is beneficial to study the most prominent parameters that contribute to strength loss in an acidic environment.

2. Research Methodology

Machine learning (ML) techniques are employed across various academic areas for forecasting and understanding the behavior of materials. This study employs KNN, LR, and ANN algorithms to forecast the percentage loss in the compressive strength (CS) of ESP and RGP-based modified cement composites. These techniques are employed due to their extensive usage, reliable prediction abilities demonstrated in related studies, and recognition as the most effective data-mining algorithms. Figure 2 displays the inclusive flowchart in the present study.

2.1. Data

The CS data, after an acid attack, rely on eight input parameters: cement, sand, water, silica fume, superplasticizer, glass powder, eggshell powder, and 90 days of CS. ESP and RGP are often observed ingredients frequently included in the database selection. Furthermore, a deliberate attempt was undertaken to choose publications with comparable components. The relevant data required for determining the proportion loss in the CS

of ESP and RGP-based reformed mortar has been acquired from the existing literature sources [42,43]. Cement, glass, and eggshell types were the same throughout the database. For accurate predictions by ML techniques, it has been suggested that the ratio of inputs to data points should fall above five [47,48]. This investigation employs a dataset comprising 234 samples to calculate the percent loss in CS. Eight distinct input factors are utilized, and the resulting ratios are significantly greater, i.e., 29.25. Therefore, the dataset used in this study met the criteria set by the researchers to apply ML techniques. The researchers purposefully selected studies that followed relevant standards for components strongly influencing the strength and durability of cement mortar. The dataset comprises 234 data points, with 70% allocated for training purposes and 30% reserved for testing the created models.

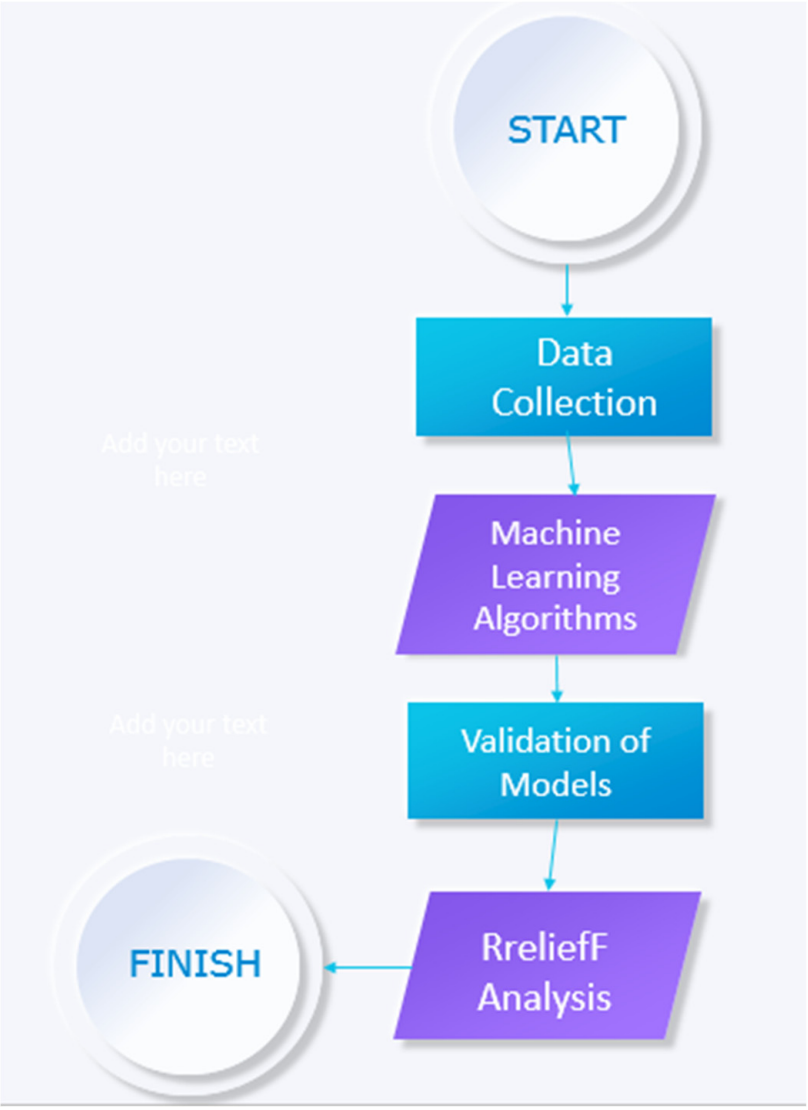


Figure 2. Flowchart representing the study process.

Data preprocessing involves various commonly used operations, such as handling missing data, encoding, detecting and treating outliers, and splitting the data [37,49,50]. The database did not contain any missing data or outliers. Furthermore, all mixes used the same composition and type of ingredients. Although a specific technique for identifying multivariate outliers was not explicitly used in this study, a comprehensive data pretreatment strategy ensured that outliers were absent. Before training the ML models, a process of identifying univariate outliers was carried out for each individual feature. This process

involved identifying and removing data points, if any, that fell outside of predetermined acceptable ranges. Additionally, the dataset underwent a thorough visual and statistical examination to identify and address any potential outliers. Additionally, developing ML algorithms often requires using the same content in multiple combinations or arrangements. A set of coefficients, sometimes referred to as descriptive statistics, provides a feasible judgment that might be applied to either the whole dataset or a sample taken from the entire dataset. Descriptive statistics utilize measurements of central tendency. Various measures of variability are commonly used in statistical analysis. These measurements include the skewness, standard deviation, kurtosis, and maximum and minimum values. Table 1 and Figure 3 presented in this study demonstrate the wide range of data utilized to execute the models. The outcomes of the descriptive study, derived from the data of all input variables, encompass a diverse range of statistics. The table additionally presents data about the ranges of each model variable, including both the maximum and minimum values. This study's other properties, such as standard deviation, mean, mode, and count of all data sets, also demonstrate the significant figures. Standard deviation refers to the extent to which data points are spread out or dispersed. It quantifies the extent to which individual data points differ from the average. Kurtosis quantifies the degree of skewness in the distribution. A high kurtosis suggests a distribution with a greater concentration of values at the tails and a lower concentration in the middle, as compared to a normal distribution. The skewness score determines if the data exhibits symmetry. A skewness value of 0 indicates a distribution that is fully symmetrical. The mean, median, and mode values refer to the central tendency of data. Figure 3 displays the relative frequency distribution of each feature utilized in the dataset. Figure 3a,b showcases multiple bars, which indicate that various proportions of cement and sand have been adopted in the dataset. Similarly, Figure 3c–e shows only three bars, which suggest that the water, silica fume, and superplasticizer utilized in this dataset have only three different proportions. Figure 3f,g exhibits the distribution of glass and eggshell powder, respectively, and Figure 3h shows the same for the 90 days of CS. A relative frequency histogram displays the fraction of all observations that refer to any specific figure or class of figures. Each parameter's selection significantly impacts the cement-based mortar's durability characteristics.

Table 1. Descriptive statistics of input variables [48].

Parameters	**Cement**	**Sand**	**Water**	**Silica Fume**	**Superplasticizer**	**Glass Powder**	**Eggshell Powder**	**90-Day CS**
Units	**kg/m^3**	**kg/m^3**	**kg/m^3**	**kg/m^3**	**kg/m^3**	**kg/m^3**	**kg/m^3**	**MPa**
Mean	732.51	735.62	191.33	151.67	38.17	30.83	30.83	46.56
Median	722.00	729.00	191.00	153.00	38.00	0.00	0.00	45.17
Mode	810.00	810.00	203.00	122.00	40.50	0.00	0.00	41.93
Standard Deviation	53.51	54.39	9.41	23.75	1.84	40.15	40.15	7.02
Kurtosis	−0.61	−0.61	−1.51	−1.51	−1.51	−0.61	−0.61	0.22
Skewness	−0.24	−0.32	0.05	−0.08	0.14	0.93	0.93	0.70
Minimum	612.00	612.00	180.00	122.00	36.00	0.00	0.00	32.32
Maximum	810.00	810.00	203.00	180.00	40.50	121.50	121.50	66.92
Count	234.00	234.00	234.00	234.00	234.00	234.00	234.00	234.00

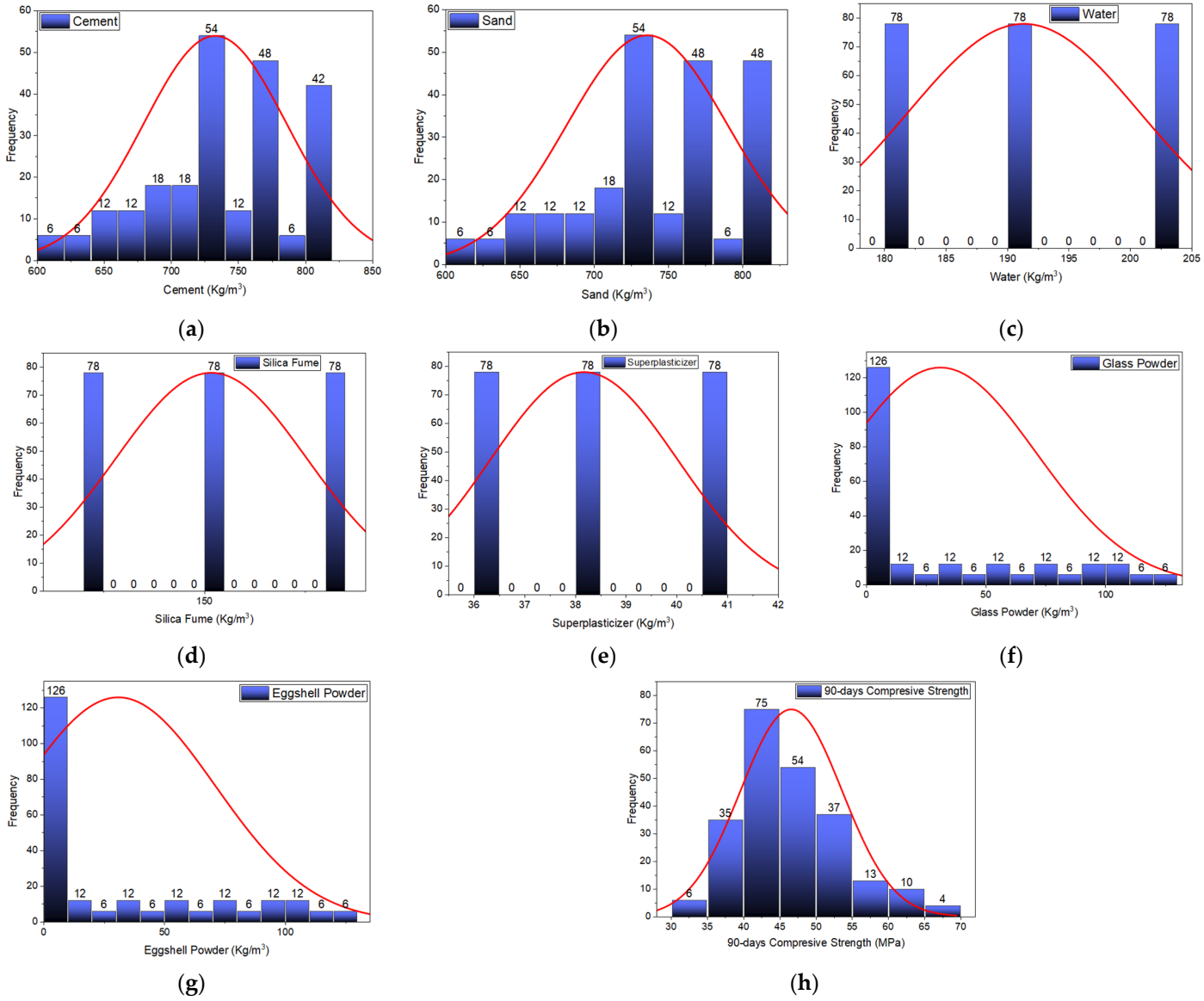


Figure 3. Frequency histograms (**a**) cement, (**b**) sand, (**c**) water, (**d**) silica fume, (**e**) superplasticizer, (**f**) glass powder, (**g**) eggshell powder, and (**h**) 90 days of CS [48].

2.2. Machine Learning Algorithms

Artificial intelligence (AI) models are proving superior to manual methods in accuracy and efficiency [51]. AI is an exceptional tool for addressing complex scenarios and has numerous benefits for navigating intricate problems. AI-based solutions can detect engineering design criteria without physical testing, substantially reducing the necessity for human intervention. Using AI can enhance decision-making, minimize errors, and optimize processing efficiencies [52,53]. The recent surge in the discourse surrounding the utilization of AI in various scientific domains has elicited a diverse array of objectives and aspirations. In recent times, there has been a notable surge in the inclination towards using AI in several scientific domains, leading to the formulation of diverse objectives. The integration of several AI methodologies across multiple domains has yielded a substantial augmentation in the realm of civil engineering. AI techniques, including reinforcement learning (RL), machine learning (ML), and deep learning (DL), are emerging as innovative methodologies to address engineering challenges. The prediction of material behavior is a typical application of ML, a swiftly growing field within the realm of AI [54].

To obtain the desired outcome, it is necessary to have a substantial quantity of input variables for ML approaches [55]. To attain the most favorable results from an ML model, it is crucial to guarantee that the parameters inside the dataset exhibit variation. Employing a constant value or one that exhibits minimal fluctuations can result in outcomes that could be better [45,56]. The current study employed an experimental dataset to assess the percentage of CS loss in cement mortar subjected to acid attack after being modified with ESP and RGP, in contrast to the plain mortar mixture. ML techniques were employed to predict the proportion of CS loss resulting from acid attack, utilizing several input variables, including cement, water, sand, silica fume, superplasticizer, eggshell powder, glass powder, and 90 days of CS. The Python code within the Orange software version 3.35.0, part of the Anaconda package, was utilized to create ML models, such as ANN, KNN, and LR, to accomplish this task. The utilization of these methods can achieve an accurate prediction of material characteristics. Figure 4 represents the employment of ML techniques within the Orange software. The efficiency of ML models was assessed through statistical methodologies, while the significance of input characteristics was ascertained by RreliefF analysis. The subsequent sections elucidate the particular ML algorithms and validation methodologies utilized in the current study.

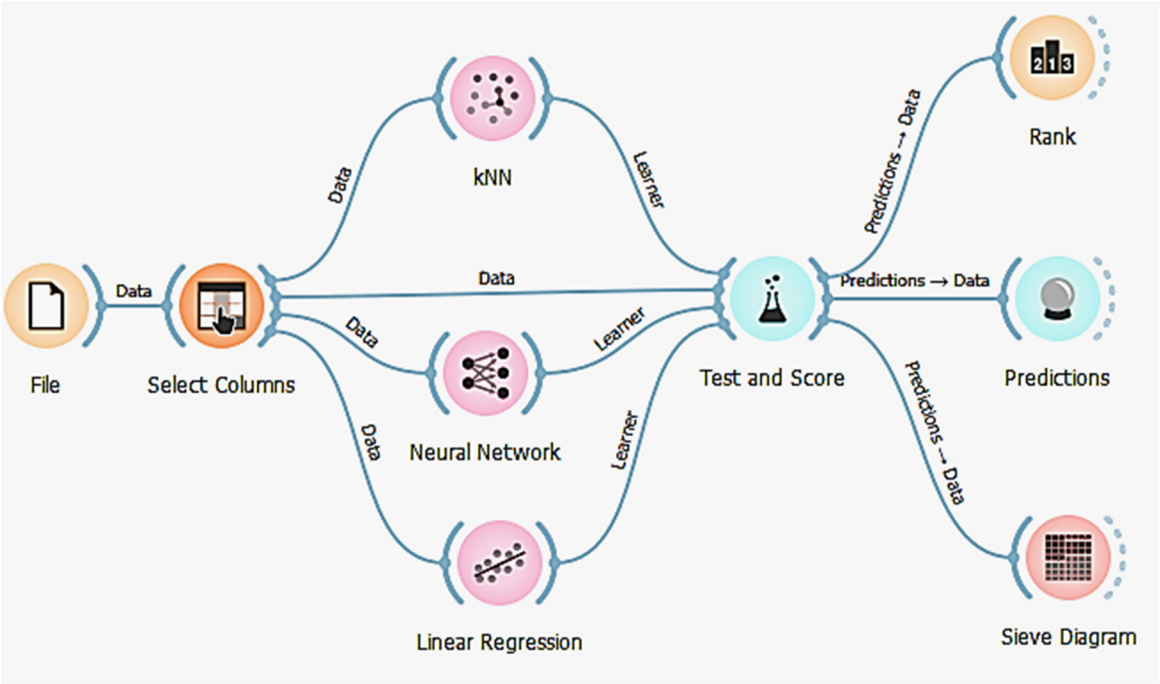


Figure 4. Arrangement of ML techniques in software.

2.2.1. Artificial Neural Network

Artificial neural networks (ANNs) are computational models that seek to replicate the make-up and operation of biotic neural networks seen in the human brain [57]. The primary structure of this entity comprises a series of interconnected units referred to as artificial neurons. ANNs incorporate the fundamental concept of neurons in living organisms, wherein these artificial nodes receive input signals, process them through analysis, and subsequently relay messages to interconnected neurons [58]. Links are commonly referred to as the connections that exist between nodes. A node has the potential to possess multiple forward and backward linkages. In implementing ANNs, the transmission of signals occurs in the form of fundamental values, which are transported from the input node of the network to the output node through one or more hidden layers. In general, a weight is assigned to each link. The input a node receives is determined by the transmission function, which calculates a weighted sum of the targets from its preceding nodes [58,59].

The inclusion of a partiality term is employed to facilitate the propagation of the outcomes. Before generating the outcomes, an activation function is employed to link the weighted sum with the present position of the internal state of this node. Figure 5 illustrates the fundamental design of ANNs. Learning in ANNs is an iterative process of modifying the network's weights to enhance accuracy. The level of precision often demanded is commonly quantified as a cost function, which ANNs are proficient at minimizing. The bulk of adaptive algorithms primarily rely on utilizing optimization approaches and statistical approximation techniques. Backpropagation is a widely employed technique in the field of supervised learning [60]. The backpropagation process entails the calculation of the derivative of the cost function concerning the weights, followed by the utilization of iterative gradient descent to modify the weights. A lower learning rate accelerates the training process, although it also raises the probability of instability and results in convergence towards adjacent minimum instead of a global optimum. Conversely, a higher learning rate diminishes the duration of the training process. Table 2 presents a comprehensive overview of the parameters utilized in constructing the ANN model.

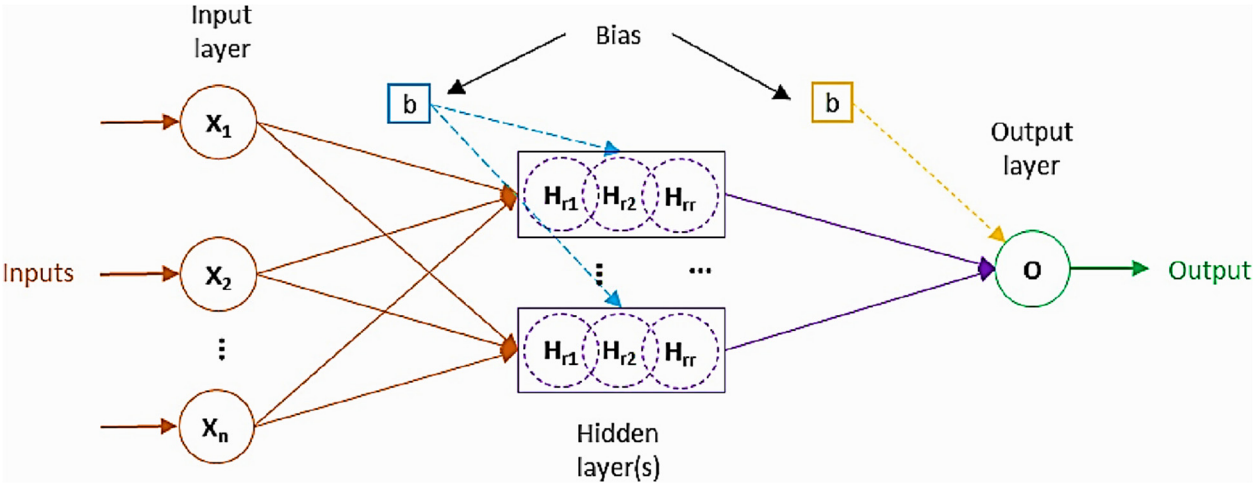


Figure 5. Structure of ANN obtained from [60].

Table 2. Parameters utilized in constructing the ANN model (parameters similar to [61]).

Parameter	Assigned Function
Neurons in hidden layers	500
Activation	ReLu
Solver	SGD
Regularization (α)	0.0001
Maximal no of iteration	1000
Replicable training	yes

2.2.2. K-Nearest Neighbors

The fundamental principle underlying the nearest neighbor arrangement, also called K-nearest neighbors (KNNs), is that the point close to a target array x, for which we seek a description, offers significant labels. The letter "K" in KNN denotes the quantity of nearest neighbors employed for classifying or foreseeing outcomes in a dataset. The classification or estimation of each new observation is determined by calculating how far it is from other values (i.e., the nearest neighbors) and using weighted averages [62]. The class designation of the majority of K-nearest patterns in the dataset is determined by the KNN algorithm [63]. A need for this objective is to utilize a similarity metric inside the data domain. The KNN algorithm is a form of semi-supervised learning that involves training data and a predetermined value for K to identify the K data points most proximate to one

another through distance computation. When a portion of the data exhibits many classes, the algorithm will infer that the unclassified instance will align with the class with the highest frequency within the dataset [64]. Figure 6 illustrates an exemplification of KNN classification. Distance measurements serve as a technique for ascertaining the spatial separation between a novel data point and a pre-existing training dataset. The distance matrices consist of the Manhattan, Mahalanobis, Euclidean, and Chebyshev. Table 3 presents an inclusive overview of the parameters utilized in constructing the KNNs model.

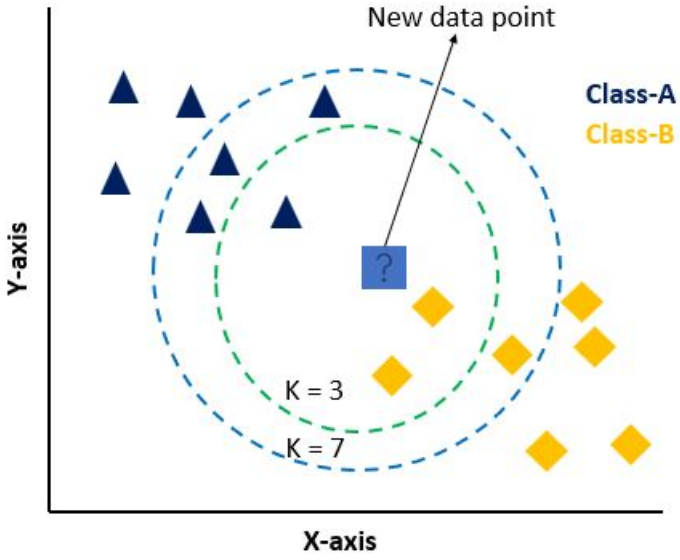


Figure 6. Structural mechanism of KNN model.

Table 3. Parameters utilized in the construction of the KNNs model.

Parameter	Assigned Function
No of neighbors	8
Distance matrix	Chebyshev
Weight	Uniform

2.2.3. Linear Regression

Linear regression (LR) is a deep-rooted method to estimate the relationship among two or more parameters [65,66]. After determining the connection between the target and input, the learning procedure will be executed to reduce the loss function value (i.e., MSE). The factors that decrease the loss function are precisely the optimum factors for the regression. Owing to its straightforwardness, the precision of this technique is low. The standard formula of a multiple LR model is specified in Equation (1) [67].

$$P = x_0 + \sum_{j=1}^{m} x_j y_j \quad (1)$$

In the context of this study, the symbol P represents the predicted outcome. The variables y_j correspond to the features or inputs of the dataset. Additionally, the parameters $x_0, x_1, \ldots, x_m$ are the values that need to be trained to optimize the model.

This work utilizes a model to effectively estimate various linear equations that establish the relationship between percentage loss in CS and the provided features. Based on prior scholarly investigations [68], it has been determined that the relation between attributes and CS has a complicated and nonlinear nature. To enhance the predictive performance of

the LR model, polynomial features are generated by applying various polynomial degrees to the original characteristics.

2.3. Validation of Model

Several frequently used statistical metrics are computed to assess a model's predictive performance on either a training or testing dataset. These metrics comprise the mean square error (MSE), the root mean squared error (RMSE), the mean absolute error (MAE), and the R-square value (R^2). The coefficient of determination (R^2) measures the extent to which a model can accurately predict outcomes [69]. Advancements in AI modeling methodologies have facilitated enhanced accuracy in predicting anisotropic and amorphous material's mechanical characteristics. This study employs a statistical error criterion to compare the ANN, KNN, and LR models.

In addition, the utilization of variance and standard deviation allows for the assessment of the model's efficacy. The utilization of the R^2 holds promise for assessing the validity and precision of the model. Inadequate outcomes are observed when R^2 values fall below 0.50, while outcomes within the range of 0.65 to 0.75 for R^2 values are considered promising. The utilization of Equation (2) enables the determination of R^2. The *MAE* is a metric used to determine the mean errors in the estimated values, irrespective of their direction. The input and output units of the MAE are identical. Despite the fact that a model's MAE may fall within a specific tolerance range, it still can occasionally produce significant errors. The calculation of MAE is derived from Equation (3). The *RMSE* represents the average of the squared differences between estimated values and actual measurements. The calculation of the error square involves the summation of all the squared errors. This novel methodology assigns a higher degree of importance to exceptional instances than previous computations, resulting in substantial disparities squared in specific scenarios while yielding relatively lower disparities in others. The *RMSE* can be computed to ascertain the model's average computational discrepancy when provided with an input. Enhanced models have a reduced *RMSE*. The RMSE is a metric used to evaluate the analytical accuracy of a model concerning future data. The *RMSE* is calculated by Equation (4). An estimator's *MSE* or mean squared deviation quantifies the average of the squared errors, representing the average squared discrepancy between the estimated values and the true value. The calculation of the *MSE* involves the utilization of Equation (5).

$$R^2 = \frac{\sum_{i=1}^{n}\left(O_i - O_i^-\right)\left(P_i - P_i^-\right)}{\sqrt{\sum_{i=1}^{n}(O_i - O_i^-)^2(P_i - P_i^-)^2}} \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|O_i - P_i| \tag{3}$$

$$RMSE = \sqrt[2]{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{N}} \tag{4}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(O_i - P_i\right)^2 \tag{5}$$

2.4. RreliefF Analysis

The algorithms belonging to the Relief family are highly effective in the practice of feature filtering and selection. The selecting features method Relief was presented by Kira and Rendell in 1992 as a solution for binary classification problems. The Relief algorithm has garnered a substantial user base because of its user-friendly interface, efficient processing speed, and consistently favorable results. However, its capacity is limited to processing only two distinct categories of data. To mitigate the challenges posed by noise, missing data, and multiclass scenarios, Kononeill conducted a study on the Relief algorithm and subsequently introduced the ReliefF technique in 1994 [70]. The RreliefF technique was created in 1997 specifically for applying ReliefF in regression issues involving

continuous classes [71]. Several later studies have introduced several Relief-family-based feature selection algorithms [72]. The Relief suite of algorithms, namely Relief, ReliefF, and RreliefF, has gained recognition for its exceptional performance in various domains. These domains encompass signal identification [73], defect localization in transmission lines [74], graphic processing and classifications [75], and gene cataloging [76]. To identify an optimal subspace, scholars have devised techniques for variable selection using structural learning within a particular collection of variables, also known as the learning subspaces [77].

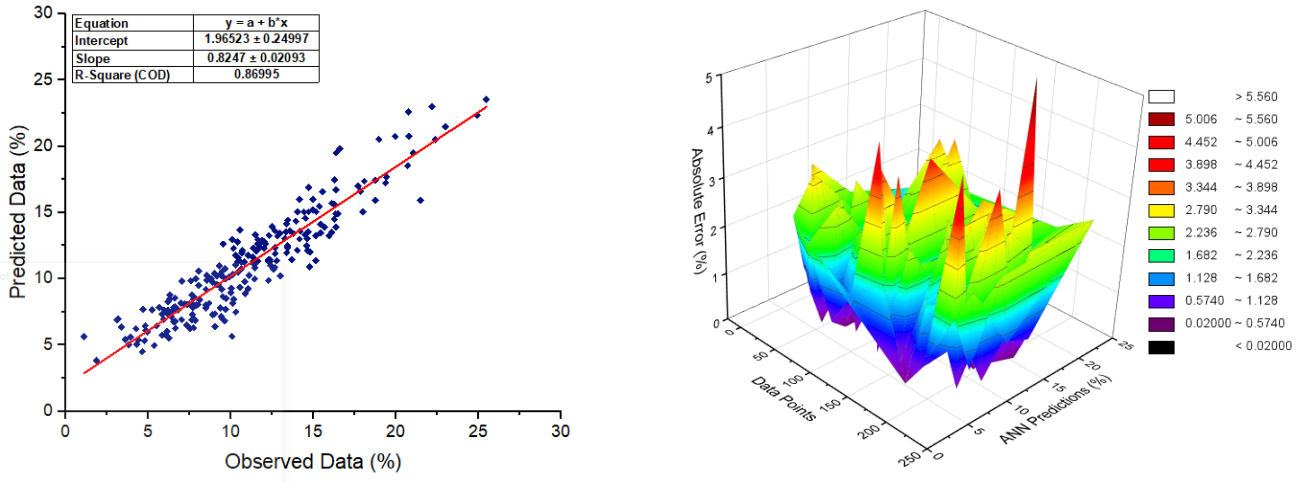
The Relief algorithm employs a variable weighting strategy to effectively tackle the challenges associated with dual classification problems [78]. The assignment of weights to each variable is determined by the strength of its association with each class. Variables that have weights below a certain threshold are excluded from consideration. The discriminatory power of the variable serves as the fundamental basis for the feature-category mapping in the Relief method. The greater the weight assigned to a feature, the more effective it will be in classifying data.

3. Results

This segment introduces the ML methods' outcomes, including the ANN, KNN, and LR models. These models were utilized to forecast the fraction loss in CS subjected to acid-attack tests of cement-based mortar, being modified with ESP and RGP.

3.1. ANN Model Results

The findings from comparing the actual and predicted fraction loss in CS subjected to acid attack using the ANN model are presented in Figure 7a. Figure 7a reveals a significant correlation between the predicted and experimental percentage loss in CS of cement-based mortar cubes containing both ESP and RGP additives. This correlation is indicated by a high R^2 value of 0.87. Furthermore, Figure 7b illustrates the error probability for percentage loss in CS. The error dispersion of ANN reveals that the mean error for the test set is 1.32%. The errors varied from a minimum of 0.02% to a maximum of 5.55%, with a standard deviation of 0.95%, indicating a relatively wide spread of values. Moreover, the error analysis indicates that only 5.6% of the absolute errors exceed 3%, whereas approximately 50% of observation errors were between 1 and 3%. However, a significant proportion of the 44.4% error was computed to be less than 1%.



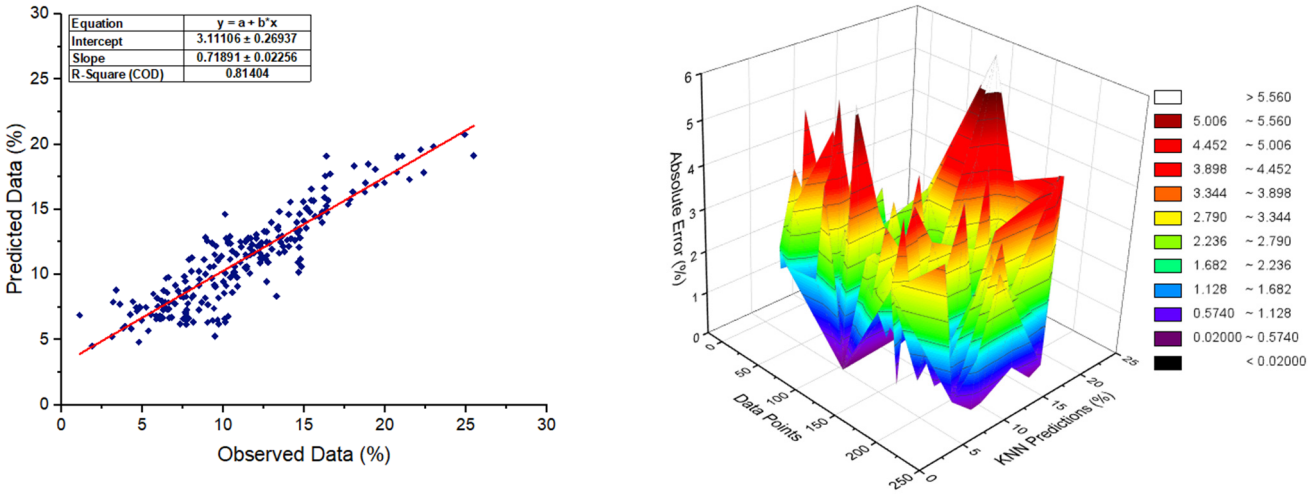
(**a**) (**b**)

Figure 7. (**a**) ANN model outcomes with R^2 value and (**b**) error distribution.

3.2. KNN Model Results

The outcomes obtained through the utilization of the KNN-based ML technique for making predictions are presented in Figure 8. To the best of our current understanding, a

crucial correlation exists between the forecasts generated by various ML techniques and the observed outcomes. Figure 8a illustrates the KNN model outcomes with an R^2 value of 0.81. Figure 8b visually depicts the error scattering of the KNN model between the test group and the predicted group. The model has a mean error of about 1.57%, with a maximum error of approximately 6.34% and a minimum error of approximately 0.01%. It is concluded by error analysis that approximately 12.8% of computed errors lie above 3%, whereas 53.4% of the errors lie within the 1 to 3% range. However, approximately 33.8% proportions of the errors were below 1%.



(**a**) (**b**)

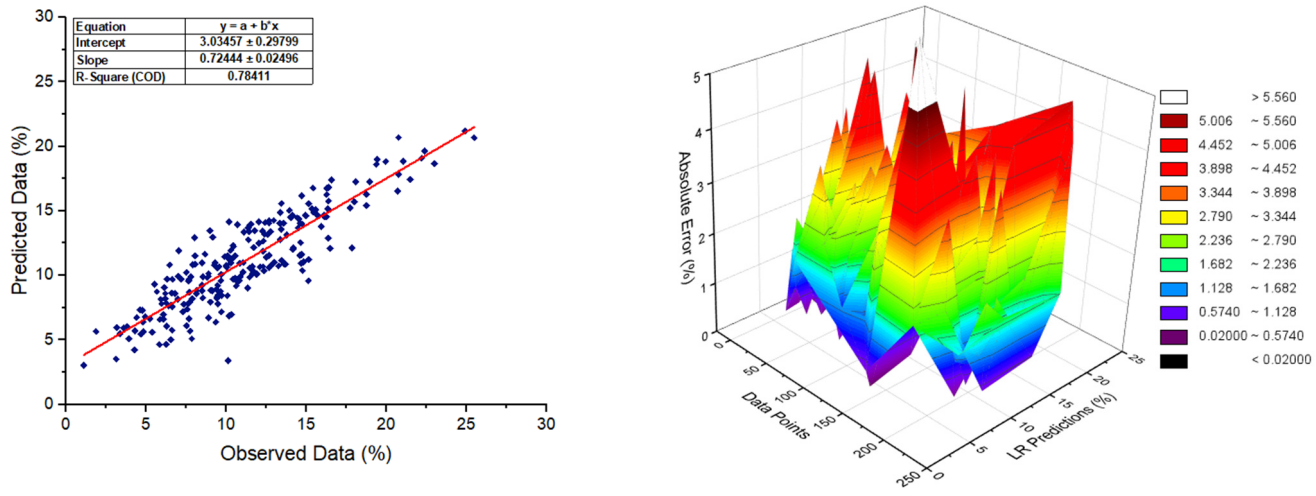
Figure 8. (**a**) KNN model outcomes with R^2 value and (**b**) error distribution.

3.3. LR Model Results

The favorable effect of the predictions obtained by the LR model is demonstrated in Figure 9. The LR model demonstrates a satisfactory correlation of R^2 = 0.78 with the projected output values, as depicted in Figure 9a. Additionally, Figure 9b illustrates the scattering of errors associated with the model. The provided illustration showcases the range of inaccuracies, revealing that the average error is approximately 1.69%. Furthermore, the analysis of error concentration reveals that the standard deviation of the error is 1.24%, with an extreme error of 6.74% and the lowest error of 0.00%. The error analysis revealed that approximately 14.2% of the absolute errors exceed 3%, whereas 51.7% of observation errors were within the range of 1 to 3%. Furthermore, approximately 34.2% of absolute errors were under 1%.

3.4. Validation Results

According to the data presented in Figure 10, it can be observed that ANN models exhibit superior performance compared to other ML models in terms of accurately predicting the fraction loss in CS after an acid attack of ESP and RGP-based modified composites. The ANN model exhibits superior performance compared to other ML models due to its distinctive capability to create a link between complicated input and output features. The ANN model exhibits significantly lower MAE, RMSE, and MSE values, specifically 1.32%, 1.63%, and 5.61%, respectively. In comparison, the LR model demonstrates higher values of 1.69%, 2.07%, and 9.39%, while the KNN model yields values of 1.57%, 1.98%, and 7.17%, respectively.



(**a**) (**b**)

Figure 9. (**a**) LR model outcomes with R^2 value and (**b**) error distribution.

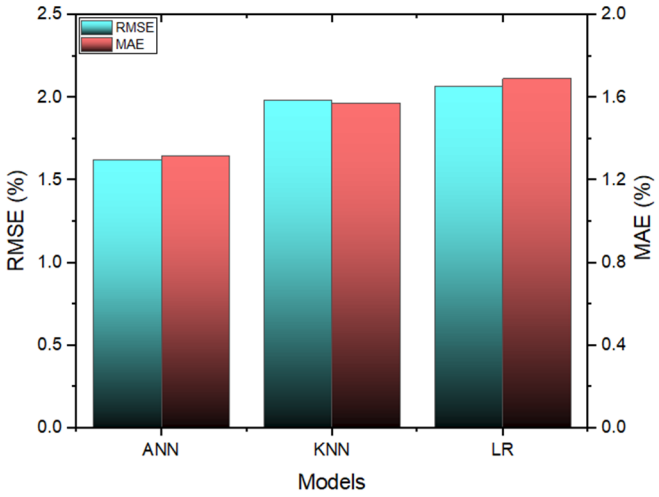


Figure 10. Visual representation of statistical errors.

The investigation resulted in the development of KNN and LR models, which were subsequently employed to forecast the durability behavior of ESP and RGP-based composites. Furthermore, the accuracy of ANN predictions was assessed through these models. Table 4 presents the statistical differences observed concerning the projected values. ANN models provide a higher level of predictability concerning the fraction loss in CS after an acid attack of cement-based mortar cubes being modified with ESP and RGP, as shown by a more robust correlation detected between the experimental and forecasted values in the statistical dataset. The ANN model has exceptional performance in terms of R^2 (0.87) and error rates compared to the LR and KNN machine learning methods.

Table 4. Statistical errors observed for different models.

Errors (%)	ANN	KNN	LR
R^2	0.87	0.81	0.78
MSE	5.61	7.17	9.39
RMSE	1.63	1.98	2.07
MAE	1.32	1.57	1.69

3.5. Results of RreliefF Analysis

The computation of the fraction loss in CS after an acid-attack test of ESP and RGP-based modified composites involves the utilization of eight input variables, namely cement, sand, water, silica fume, superplasticizer, glass powder, eggshell powder, and 90 days of CS. Figure 11 illustrates the respective contributions made by each input parameter to the creation of the ML models.

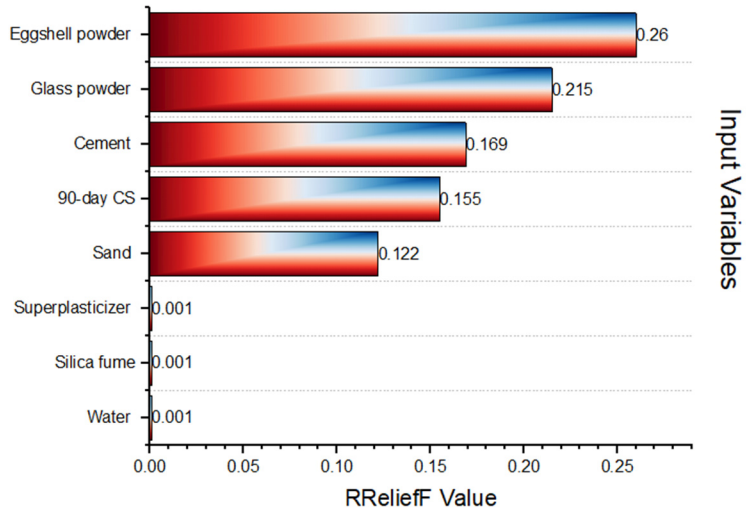


Figure 11. RreliefF analysis graph.

According to the RreliefF analysis graph, it can be observed that eggshell powder, glass powder, and cement exhibit the most substantial influence on the fraction reduction in CS subsequent to an acid-attack test. The inclusion of superplasticizer, water, and silica fume reduces the percentage decrease in CS. The RreliefF value serves as a metric for assessing the significance of individual input parameters. The significance of an input parameter increases as its RreliefF value increases. In this particular instance, it is observed that eggshell powder, glass powder, and cement have the highest RreliefF values of 0.26, 0.215, and 0.169, respectively, hence signifying their significance as input parameters.

The utilization of the RreliefF analysis graph can assist engineers and researchers in the development of acid-resistant concrete designs. Engineers can make informed decisions regarding the selection of materials for optimal protection against acid attack by comprehending the correlation between various constituents in concrete and the extent of CS reduction observed during acid-attack testing.

4. Discussions

The outcomes suggest that the ANN model exhibited superior performance to the alternative models, as evidenced by higher R^2 scores and lesser RMSE and MAE values. The result can be compared with the recent literature tabulated in Table 5. The observed outcome might be attributed to ANN's enhanced robustness compared to other models, such as KNN and LR. The application of ANN enables a precise assessment of the fraction loss in CS of ESP and RGP-based modified composites by effectively capturing the intricate relationships among many factors, including ESP and RGP concentration, water-cement ratio, and age. ANN employs a system of interconnected neurons to acquire knowledge and make predictions, relying on the observed correlation between these variables. During the learning phase, the ANN model incorporates the entire training dataset as an additional precautionary measure for enhanced security. By considering all of these aspects, ANN is capable of generating dependable predictions while disregarding any extraneous data or noise that could undermine the accuracy of its outcomes. Consequently, only the properties that hold significant importance in determining the resilience of ESP and RGP-based com-

posites are considered during the prediction phase. The findings of this analysis indicate that the KNN and LR models exhibit limited reliability in their predictive capabilities. Acid-attack tests encompass multiple components and interactions that may not correspond to conventional linear or nearest-neighbor correlations. KNN relies on the close proximity of data points in the feature space [61]; however, it may not accurately capture the nonlinear relationships that exist in acid-attack data. Due to its linear nature, LR may have difficulties in accurately representing the complex relationships [79] associated with the decrease in CS after acid attack. KNN and LR algorithms necessitate manual feature engineering or encounter difficulties in managing feature spaces with a high number of dimensions, which limits the reliability of these models.

Table 5. Comparison with past relevant studies.

Materials	Properties	Models	Outperformed Model	Reference
Self-healing concrete	CS	ANN, ANFIS	ANN	[80]
Alkali-activated materials	CS	KNN, ANN, DT	ANN	[61]
GGBFS-based concrete	CS	LR, ANN, non-LR, quadratic, full quadratic models	ANN	[81]
Concrete with hooked steel fibers	CS	KNN, ANN	ANN	[82]
Concrete bricks utilizing various industrial wastes like fly ash, rise husk ash, and hydrated lime	Water absorption, CS, Density	Multiple LR, ANN	ANN	[83]
RGP-based concrete	CS	ANN, LR, non-LR	ANN	[79]

Figure 12 illustrates the sieve diagram for ANN, KNN, and LR models. The sieve diagram shows the correlation among the predicted and experimental values of fraction loss in CS. The Pearson χ^2 values for ANN, KNN, and LR are 213.99, 167.47, and 140.08, respectively, suggesting a significant deviation from the independence assumption. In this visual representation, the relative size of each rectangle corresponds to the expected percentage loss in CS. In contrast, the observed percentage loss in CS is depicted by the number of squares within each rectangle. Therefore, the distinction between the experimental and estimated values is visually represented by the intensity of shading, with color utilized to signify whether the departure from independence is positive or negative. In this graph, solid lines with blue color are utilized to represent positive deviations, while red lines are employed to depict negative deviations. As seen in Figure 12 below, the blue shad in KNN and LR is much greater than ANN, which suggests that ANN outperformed other ML techniques.

Although ML models are powerful tools for featuring concrete properties, there are certain limitations associated with these techniques. Various ML algorithms like ANN accurately forecast the target variables, but their interpretation is hard to understand as these predictions are considered a black box by various researchers [61]. Moreover, a comprehensive database is required to train the ML models. ANN, despite its ability to effectively capture intricate nonlinear connections, can be prone to overfitting, particularly when the dataset is small or restricted. Similarly, KNN is dependent on similarity metrics based on proximity and can become computationally difficult, especially when the dataset grows in size. This constraint becomes more noticeable in situations where the number of features is large, which might significantly impact the efficiency and scalability of the model. Thus, a compatible database is essential to execute the ML models.

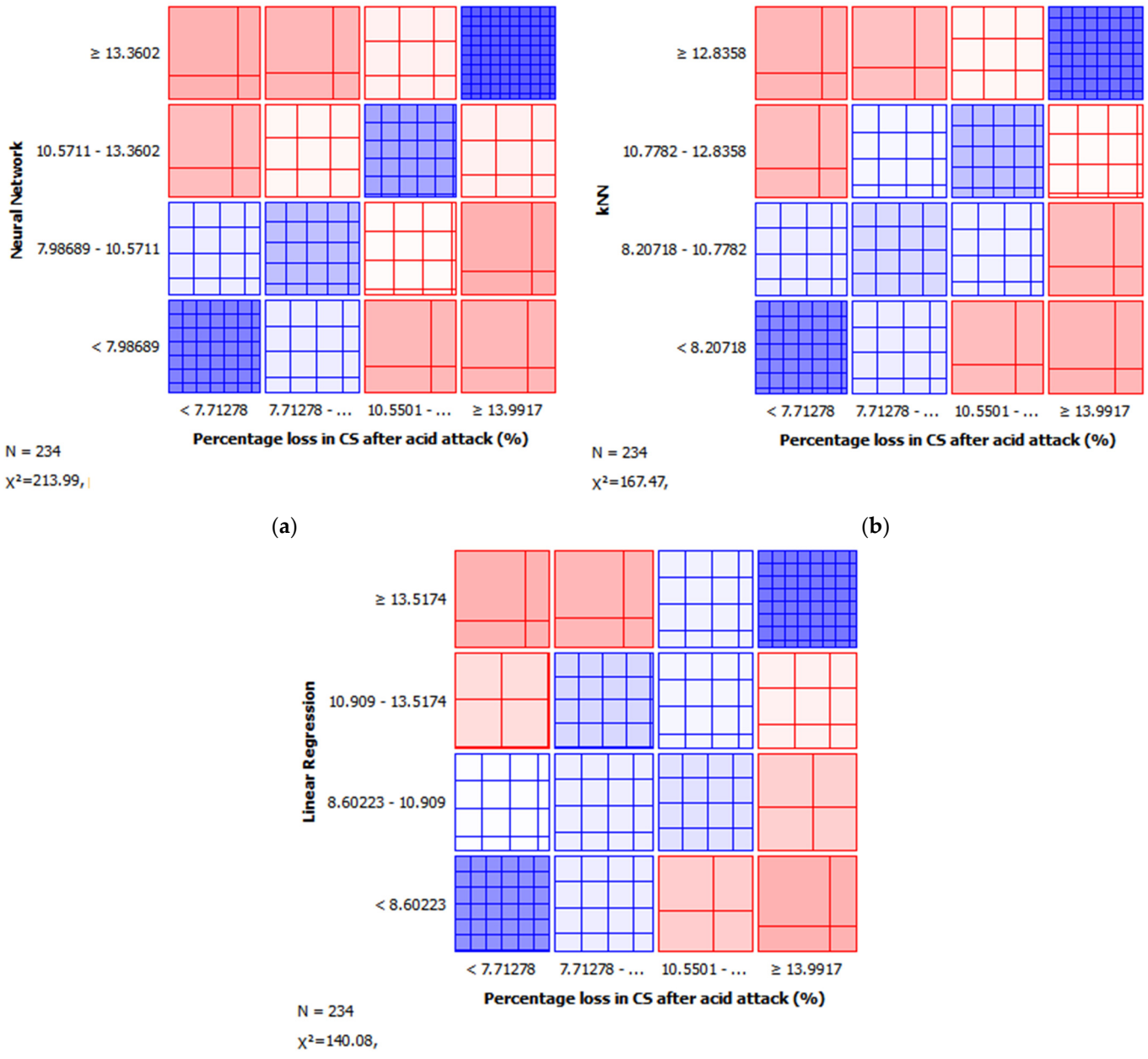


Figure 12. Sieve diagram (**a**) for ANN, (**b**) for KNN, and (**c**) for LR.

5. Conclusions

This study was conducted to investigate the influence of incorporating eggshell powder (ESP) and recycled glass powder (RGP) into cement-based composites on compressive strength (CS) when the specimens are subjected to acidic conditions. The research utilized three machine learning (ML) methods, specifically linear regression (LR), artificial neural network (ANN), and K-nearest neighbor (KNN), to estimate the decline in CS of cement-based composites modified with ESP and RGP. The study's findings are presented as follows:

- A strong correlation was seen between the developed ML models and the results obtained from testing, signifying their probable applicability in estimating the fraction loss in CS of cement-based composites treated with ESP and RGP.

- The ANN model was considered more favorable than the KNN and LR models because of its higher accuracy level, as evidenced by the R^2 values (0.87 for ANN, 0.81 for KNN, and 0.78 for LR).
- The evaluation of errors, such as MSE, MAE, and RMSE, suggested that the ANN model prediction capabilities were more substantial than the KNN and LR models. The ANN model exhibits MSE, MAE, and RMSE values of 6.51%, 1.32%, and 1.63%, respectively.
- The feature importance graph suggested that the concentration of ESP and RGP, as well as the percentage of cement, greatly influenced the decline in CS subjected to acid-attack tests with the RreliefF scores of 0.26, 0.215, and 0.169, respectively.
- The sieve diagram also indicated that the ANN model outperformed other ML methods (KNN and LR), with the χ^2 value of 213.99, suggesting that the two variables were significantly related.

The diversity of ML models depends on the utilized database. This study used 234 data points to assess the performance of ANN, KNN, and LR models; however, future research should be focused on compiling a more comprehensive database for the training and testing of ML models. Additionally, deep learning techniques like deep neural networks can be employed on the same database to contrast the results with machine learning models.

Author Contributions: Conceptualization, F.Z. and J.H.; methodology, J.H.; software, F.Z. and X.W.; formal analysis, Y.L.; data curation, Y.L.; writing—original draft, F.Z. and J.H.; writing—review and editing, X.W., Y.L. and J.H.; supervision, X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

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

References

1. Khan, M.; Cao, M.; Xie, C.; Ali, M. Effectiveness of hybrid steel-basalt fiber reinforced concrete under compression. *Case Stud. Constr. Mater.* **2022**, *16*, e00941. [CrossRef]
2. Khan, M.; Cao, M.; Chu, S.H.; Ali, M. Properties of hybrid steel-basalt fiber reinforced concrete exposed to different surrounding conditions. *Constr. Build. Mater.* **2022**, *322*, 126340. [CrossRef]
3. Bahraq, A.A.; Jose, J.; Shameem, M.; Maslehuddin, M. A review on treatment techniques to improve the durability of recycled aggregate concrete: Enhancement mechanisms, performance and cost analysis. *J. Build. Eng.* **2022**, *55*, 104713. [CrossRef]
4. Baloch, W.L.; Siad, H.; Lachemi, M.; Sahmaran, M. A review on the durability of concrete-to-concrete bond in recent rehabilitated structures. *J. Build. Eng.* **2021**, *44*, 103315. [CrossRef]
5. Sahoo, S.; Mahapatra, T.R.; Priyadarshini, N.; Mahapatra, S.; Naik, S.; Jaypuria, S. Influence of water binder ratio on strength and acid resistance of concrete made up of mineral admixture as supplementary cementitious material. *Mater. Today Proc.* **2020**, *26*, 796–803. [CrossRef]
6. Nnaemeka, O.F.; Singh, N.B. Durability properties of geopolymer concrete made from fly ash in presence of Kaolin. *Mater. Today Proc.* **2020**, *29*, 781–784. [CrossRef]
7. Afroughsabet, V.; Ozbakkaloglu, T. Mechanical and durability properties of high-strength concrete containing steel and polypropylene fibers. *Constr. Build. Mater.* **2015**, *94*, 73–82. [CrossRef]
8. Aliyu, I.; Sulaiman, T.A.; Mohammed, A.; Kaura, J.M. Effect of sulphuric acid on the compressive strength of concrete with quarry dust as partial replacement of fine aggregate. *FUDMA J. Sci.* **2020**, *4*, 553–559.
9. Valencia-Saavedra, W.G.; Mejía de Gutiérrez, R. Resistance to Chemical Attack of Hybrid Fly Ash-Based Alkali-Activated Concretes. *Molecules* **2020**, *25*, 3389. [CrossRef]
10. Kamaruzzaman, W.M.I.W.M.; Shaifudin, M.S.; Nasir, N.A.M.; Hamidi, N.A.S.M.; Yusof, N.; Adnan, A.; Jew, L.O.; Nik, W.M.N.W.; Ghazali, M.S.M. Eggshells biowaste filler for improving the anticorrosive behaviour of waterborne polyurethane coatings on mild steel in artificial seawater. *J. Mater. Res. Technol.* **2022**, *21*, 3815–3827. [CrossRef]
11. Hamada, H.M.; Tayeh, B.A.; Al-Attar, A.; Yahaya, F.M.; Muthusamy, K.; Humada, A.M. The present state of the use of eggshell powder in concrete: A review. *J. Build. Eng.* **2020**, *32*, 101583. [CrossRef]
12. Cree, D.; Pliya, P. Effect of elevated temperature on eggshell, eggshell powder and eggshell powder mortars for masonry applications. *J. Build. Eng.* **2019**, *26*, 100852. [CrossRef]

13. Binici, H.; Aksogan, O.; Sevinc, A.H.; Cinpolat, E. Mechanical and radioactivity shielding performances of mortars made with cement, sand and egg shells. *Constr. Build. Mater.* **2015**, *93*, 1145–1150. [CrossRef]
14. Wei, C.B.; Othman, R.; Ying, C.Y.; Jaya, R.P.; Ing, D.S.; Mangi, S.A. Properties of mortar with fine eggshell powder as partial cement replacement. *Mater. Today Proc.* **2021**, *46*, 1574–1581. [CrossRef]
15. Yang, D.; Zhao, J.; Ahmad, W.; Amin, M.N.; Aslam, F.; Khan, K.; Ahmad, A. Potential use of waste eggshells in cement-based materials: A bibliographic analysis and review of the material properties. *Constr. Build. Mater.* **2022**, *344*, 128143. [CrossRef]
16. Khan, K.; Ahmad, W.; Amin, M.N.; Deifalla, A.F. Investigating the feasibility of using waste eggshells in cement-based materials for sustainable construction. *J. Mater. Res. Technol.* **2023**, *23*, 4059–4074. [CrossRef]
17. Alsharari, F.; Khan, K.; Amin, M.N.; Ahmad, W.; Khan, U.; Mutnbak, M.; Houda, M.; Yosri, A.M. Sustainable use of waste eggshells in cementitious materials: An experimental and modeling-based study. *Case Stud. Constr. Mater.* **2022**, *17*, e01620. [CrossRef]
18. Khan, M.; Cao, M.; Hussain, A.; Chu, S.H. Effect of silica-fume content on performance of $CaCO_3$ whisker and basalt fiber at matrix interface in cement-based composites. *Constr. Build. Mater.* **2021**, *300*, 124046. [CrossRef]
19. Ahmad, M.R.; Khan, M.; Wang, A.; Zhang, Z.; Dai, J.-G. Alkali-activated materials partially activated using flue gas residues: An insight into reaction products. *Constr. Build. Mater.* **2023**, *371*, 130760. [CrossRef]
20. Lao, J.-C.; Xu, L.-Y.; Huang, B.-T.; Zhu, J.-X.; Khan, M.; Dai, J.-G. Utilization of sodium carbonate activator in strain-hardening ultra-high-performance geopolymer concrete (SH-UHPGC). *Front. Mater.* **2023**, *10*, 1142237. [CrossRef]
21. Hamada, H.; Alattar, A.; Tayeh, B.; Yahaya, F.; Adesina, A. Sustainable application of coal bottom ash as fine aggregates in concrete: A comprehensive review. *Case Stud. Constr. Mater.* **2022**, *16*, e01109. [CrossRef]
22. Amin, M.N.; Alkadhim, H.A.; Ahmad, W.; Khan, K.; Alabduljabbar, H.; Mohamed, A. Experimental and machine learning approaches to investigate the effect of waste glass powder on the flexural strength of cement mortar. *PLoS ONE* **2023**, *18*, e0280761. [CrossRef] [PubMed]
23. Qin, D.; Hu, Y.; Li, X. Waste Glass Utilization in Cement-Based Materials for Sustainable Construction: A Review. *Crystals* **2021**, *11*, 710. [CrossRef]
24. Qaidi, S.; Najm, H.M.; Abed, S.M.; Özkılıç, Y.O.; Al Dughaishi, H.; Alosta, M.; Sabri, M.M.; Alkhatib, F.; Milad, A. Concrete Containing Waste Glass as an Environmentally Friendly Aggregate: A Review on Fresh and Mechanical Characteristics. *Materials* **2022**, *15*, 6222. [CrossRef] [PubMed]
25. Zeybek, Ö.; Özkılıç, Y.O.; Karalar, M.; Çelik, A.İ.; Qaidi, S.; Ahmad, J.; Burduhos-Nergis, D.D.; Burduhos-Nergis, D.P. Influence of Replacing Cement with Waste Glass on Mechanical Properties of Concrete. *Materials* **2022**, *15*, 7513. [CrossRef] [PubMed]
26. Çelik, A.İ.; Özkılıç, Y.O.; Zeybek, Ö.; Karalar, M.; Qaidi, S.; Ahmad, J.; Burduhos-Nergis, D.D.; Bejinariu, C. Mechanical behavior of crushed waste glass as replacement of aggregates. *Materials* **2022**, *15*, 8093. [CrossRef]
27. Karalar, M.; Özkılıç, Y.O.; Aksoylu, C.; Sabri Sabri, M.M.; Beskopylny, A.N.; Stel'makh, S.A.; Shcherban, E.M. Flexural behavior of reinforced concrete beams using waste marble powder towards application of sustainable concrete. *Front. Mater.* **2022**, *9*, 1068791. [CrossRef]
28. Özkılıç, Y.O.; Çelik, A.İ.; Tunç, U.; Karalar, M.; Deifalla, A.; Alomayri, T.; Althoey, F. The use of crushed recycled glass for alkali activated fly ash based geopolymer concrete and prediction of its capacity. *J. Mater. Res. Technol.* **2023**, *24*, 8267–8281. [CrossRef]
29. Çelik, A.İ.; Tunç, U.; Bahrami, A.; Karalar, M.; Mydin, M.A.O.; Alomayri, T.; Özkılıç, Y.O. Use of waste glass powder toward more sustainable geopolymer concrete. *J. Mater. Res. Technol.* **2023**, *24*, 8533–8546. [CrossRef]
30. Jiang, Y.; Ling, T.-C.; Mo, K.H.; Shi, C. A critical review of waste glass powder–Multiple roles of utilization in cement-based materials and construction products. *J. Environ. Manag.* **2019**, *242*, 440–449. [CrossRef]
31. Zhang, H.; Geng, Y.; Wang, Y.-Y.; Li, X.-Z. Experimental study and prediction model for bond behaviour of steel-recycled aggregate concrete composite slabs. *J. Build. Eng.* **2022**, *53*, 104585. [CrossRef]
32. Emad, W.; Mohammed, A.S.; Bras, A.; Asteris, P.G.; Kurda, R.; Muhammed, Z.; Hassan, A.M.T.; Qaidi, S.M.A.; Sihag, P. Metamodel techniques to estimate the compressive strength of UHPFRC using various mix proportions and a high range of curing temperatures. *Constr. Build. Mater.* **2022**, *349*, 128737. [CrossRef]
33. Huang, J.; Zhou, M.; Zhang, J.; Ren, J.; Vatin, N.I.; Sabri, M.M.S. Development of a new stacking model to evaluate the strength parameters of concrete samples in laboratory. *Iran. J. Sci. Technol. Trans. Civ. Eng.* **2022**, *46*, 4355–4370. [CrossRef]
34. Deifalla, A. Refining the torsion design of fibered concrete beams reinforced with FRP using multi-variable non-linear regression analysis for experimental results. *Eng. Struct.* **2021**, *226*, 111394. [CrossRef]
35. Al-Tayeb, M.M.; Aisheh, Y.I.A.; Qaidi, S.M.A.; Tayeh, B.A. Experimental and simulation study on the impact resistance of concrete to replace high amounts of fine aggregate with plastic waste. *Case Stud. Constr. Mater.* **2022**, *17*, e01324. [CrossRef]
36. Khan, K.; Ahmad, A.; Amin, M.N.; Ahmad, W.; Nazar, S.; Arab, A.M.A. Comparative study of experimental and modeling of fly ash-based concrete. *Materials* **2022**, *15*, 3762. [CrossRef]
37. Hosseinzadeh, M.; Mousavi, S.S.; Hosseinzadeh, A.; Dehestani, M. An efficient machine learning approach for predicting concrete chloride resistance using a comprehensive dataset. *Sci. Rep.* **2023**, *13*, 15024. [CrossRef]
38. Tran, V.Q.; Dang, V.Q.; Ho, L.S. Evaluating compressive strength of concrete made with recycled concrete aggregates using machine learning approach. *Constr. Build. Mater.* **2022**, *323*, 126578. [CrossRef]

39. Iftikhar, B.; Alih, S.C.; Vafaei, M.; Elkotb, M.A.; Shutaywi, M.; Javed, M.F.; Deebani, W.; Khan, M.I.; Aslam, F. Predictive modeling of compressive strength of sustainable rice husk ash concrete: Ensemble learner optimization and comparison. *J. Clean. Prod.* **2022**, *348*, 131285. [CrossRef]
40. Young, B.A.; Hall, A.; Pilon, L.; Gupta, P.; Sant, G. Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: New insights from statistical analysis and machine learning methods. *Cem. Concr. Res.* **2019**, *115*, 379–388. [CrossRef]
41. de Melo, V.V.; Banzhaf, W. Improving the prediction of material properties of concrete using Kaizen Programming with Simulated Annealing. *Neurocomputing* **2017**, *246*, 25–44. [CrossRef]
42. Wang, N.; Xia, Z.; Amin, M.N.; Ahmad, W.; Khan, K.; Althoey, F.; Alabduljabbar, H. Sustainable strategy of eggshell waste usage in cementitious composites: An integral testing and computational study for compressive behavior in aggressive environment. *Constr. Build. Mater.* **2023**, *386*, 131536. [CrossRef]
43. Alfaiad, M.A.; Khan, K.; Ahmad, W.; Amin, M.N.; Deifalla, A.F.; Ghamry, N.A. Evaluating the compressive strength of glass powder-based cement mortar subjected to the acidic environment using testing and modeling approaches. *PLoS ONE* **2023**, *18*, e0284761. [CrossRef] [PubMed]
44. Jin, C.; Qian, Y.; Khan, K.; Ahmad, A.; Amin, M.N.; Althoey, F.; Nawaz, R. Predicting the damage to cementitious composites due to acid attack and evaluating the effectiveness of eggshell powder using interpretable artificial intelligence models. *Mater. Today Commun.* **2023**, *37*, 107333. [CrossRef]
45. Amin, M.N.; Ahmad, W.; Khan, K.; Al-Hashem, M.N.; Deifalla, A.F.; Ahmad, A. Testing and modeling methods to experiment the flexural performance of cement mortar modified with eggshell powder. *Case Stud. Constr. Mater.* **2023**, *18*, e01759. [CrossRef]
46. Khan, K.; Ahmad, W.; Amin, M.N.; Rafiq, M.I.; Arab, A.M.A.; Alabdullah, I.A.; Alabduljabbar, H.; Mohamed, A. Evaluating the effectiveness of waste glass powder for the compressive strength improvement of cement mortar using experimental and machine learning methods. *Heliyon* **2023**, *9*, e16288. [CrossRef] [PubMed]
47. Tsanas, A.; Xifara, A. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy Build.* **2012**, *49*, 560–567. [CrossRef]
48. Chen, Z.; Amin, M.N.; Iftikhar, B.; Ahmad, W.; Althoey, F.; Alsharari, F. Predictive modelling for the acid resistance of cement-based composites modified with eggshell and glass waste for sustainable and resilient building materials. *J. Build. Eng.* **2023**, *76*, 107325. [CrossRef]
49. Taffese, W.Z.; Espinosa-Leal, L. A machine learning method for predicting the chloride migration coefficient of concrete. *Constr. Build. Mater.* **2022**, *348*, 128566. [CrossRef]
50. Cao, Q.; Yuan, X.; Nasir Amin, M.; Ahmad, W.; Althoey, F.; Alsharari, F. A soft-computing-based modeling approach for predicting acid resistance of waste-derived cementitious composites. *Constr. Build. Mater.* **2023**, *407*, 133540. [CrossRef]
51. Zhou, J.; Shen, X.; Qiu, Y.; Shi, X.; Khandelwal, M. Cross-correlation stacking-based microseismic source location using three metaheuristic optimization algorithms. *Tunn. Undergr. Space Technol.* **2022**, *126*, 104570. [CrossRef]
52. Nica, E.; Stehel, V. Internet of things sensing networks, artificial intelligence-based decision-making algorithms, and real-time process monitoring in sustainable industry 4.0. *J. Self-Gov. Manag. Econ.* **2021**, *9*, 35–47.
53. Qiu, Y.; Zhou, J. Short-term rockburst damage assessment in burst-prone mines: An explainable XGBOOST hybrid model with SCSO algorithm. *Rock Mech. Rock Eng.* **2023**, *56*, 8745–8770. [CrossRef]
54. Zhou, J.; Shen, X.; Qiu, Y.; Shi, X.; Du, K. Microseismic location in hardrock metal mines by machine learning models based on hyperparameter optimization using bayesian optimizer. *Rock Mech. Rock Eng.* **2023**, *56*, 8771–8788. [CrossRef]
55. Sufian, M.; Ullah, S.; Ostrowski, K.A.; Ahmad, A.; Zia, A.; Śliwa-Wieczorek, K.; Siddiq, M.; Awan, A.A. An Experimental and Empirical Study on the Use of Waste Marble Powder in Construction Material. *Materials* **2021**, *14*, 3829. [CrossRef]
56. Khan, M.; Lao, J.; Dai, J.-G. Comparative study of advanced computational techniques for estimating the compressive strength of UHPC. *J. Asian Concr. Fed.* **2022**, *8*, 51–68. [CrossRef]
57. Arifeen, S.U.; Ali, M.; Macioszek, E. Analysis of vehicle pedestrian crash severity using advanced machine learning techniques. *Arch. Transp.* **2023**, *68*, 91–116. [CrossRef]
58. Getahun, M.A.; Shitote, S.M.; Gariy, Z.C.A. Artificial neural network based modelling approach for strength prediction of concrete incorporating agricultural and construction wastes. *Constr. Build. Mater.* **2018**, *190*, 517–525. [CrossRef]
59. Huang, J.; Zhang, J.; Li, X.; Qiao, Y.; Zhang, R.; Kumar, G.S. Investigating the effects of ensemble and weight optimization approaches on neural networks' performance to estimate the dynamic modulus of asphalt concrete. *Road Mater. Pavement Des.* **2023**, *24*, 1939–1959. [CrossRef]
60. Bardhan, A.; Asteris, P.G. Application of hybrid ANN paradigms built with nature inspired meta-heuristics for modelling soil compaction parameters. *Transp. Geotech.* **2023**, *41*, 100995. [CrossRef]
61. Arifeen, S.U.; Amin, M.N.; Ahmad, W.; Althoey, F.; Ali, M.; Alotaibi, B.S.; Abuhussain, M.A. A comparative study of prediction models for alkali-activated materials to promote quick and economical adaptability in the building sector. *Constr. Build. Mater.* **2023**, *407*, 133485. [CrossRef]
62. Abed, M.; Imteaz, M.; Ahmed, A.N.; Huang, Y.F. Application of k-nearest neighbors (KNN) technique for predicting monthly pan evaporation. *AIP Conf. Proc.* **2023**, *2631*, 020020.
63. Kramer, O. K-nearest neighbors. In *Dimensionality Reduction with Unsupervised Nearest Neighbors*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 13–23.

64. Chomboon, K.; Chujai, P.; Teerarassamee, P.; Kerdprasop, K.; Kerdprasop, N. An empirical study of distance metrics for k-nearest neighbor algorithm. In Proceedings of the 3rd International Conference on Industrial Application Engineering 2015, Kitakyushu, Japan, 28–31 March 2015.
65. Weisberg, S. *Applied Linear Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2005; Volume 528.
66. Qiu, Y.; Zhou, J. Short-term rockburst prediction in underground project: Insights from an explainable and interpretable ensemble learning model. *Acta Geotech.* **2023**, *18*, 6655–6685. [CrossRef]
67. Khademi, F.; Akbari, M.; Jamal, S.M.; Nikoo, M. Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete. *Front. Struct. Civ. Eng.* **2017**, *11*, 90–99. [CrossRef]
68. Cook, R.; Lapeyre, J.; Ma, H.; Kumar, A. Prediction of compressive strength of concrete: Critical comparison of performance of a hybrid machine learning model with standalone models. *J. Mater. Civ. Eng.* **2019**, *31*, 04019255. [CrossRef]
69. Bemani, A.; Baghban, A.; Mosavi, A. Estimating CO_2-Brine diffusivity using hybrid models of ANFIS and evolutionary algorithms. *Eng. Appl. Comput. Fluid Mech.* **2020**, *14*, 818–834. [CrossRef]
70. Kononenko, I. On Biases in Estimating Multi-Valued Attributes. 1995. Available online: https://www.ijcai.org/Proceedings/95-2/Papers/003.pdf (accessed on 4 January 2024).
71. Robnik-Šikonja, M.; Kononenko, I. Theoretical and empirical analysis of ReliefF and RReliefF. *Mach. Learn.* **2003**, *53*, 23–69. [CrossRef]
72. Kong, D.; Ding, C.; Huang, H.; Zhao, H. Multi-label relieff and f-statistic feature selections for image annotation. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2352–2359.
73. Yang, Z.; Duan, M. Application of relieff algorithm in radar emitter signal recognition. *J. Chengdu Univ.* **2012**.
74. Farshad, M.; Sadeh, J. Transmission line fault location using hybrid wavelet-Prony method and relief algorithm. *Int. J. Electr. Power Energy Syst.* **2014**, *61*, 127–136. [CrossRef]
75. Jia, J.; Yang, N.; Zhang, C.; Yue, A.; Yang, J.; Zhu, D. Object-oriented feature selection of high spatial resolution images using an improved Relief algorithm. *Math. Comput. Model.* **2013**, *58*, 619–626. [CrossRef]
76. Wang, Y.; Makedon, F. Application of Relief-F feature filtering algorithm to selecting informative genes for cancer classification using microarray data. In Proceedings of the 2004 IEEE Computational Systems Bioinformatics Conference, Stanford, CA, USA, 19 August 2004; pp. 497–498.
77. Zhang, B.; Li, Y.; Chai, Z. A novel random multi-subspace based ReliefF for feature selection. *Knowl. -Based Syst.* **2022**, *252*, 109400. [CrossRef]
78. Kira, K.; Rendell, L.A. The feature selection problem: Traditional methods and a new algorithm. In Proceedings of the AAAI'92: Proceedings of the Tenth National Conference on Artificial Intelligence, San Jose, CA, USA, 12–16 July 1992; pp. 129–134.
79. Ahmad, S.A.; Rafiq, S.K.; Hilmi, H.D.M.; Ahmed, H.U. Mathematical modeling techniques to predict the compressive strength of pervious concrete modified with waste glass powders. *Asian J. Civ. Eng.* **2023**, *25*, 773–785. [CrossRef]
80. Sri, K.S.; Nayaka, R.R.; Kumar, M.V.N.S. Mechanical properties of sustainable self-healing concrete and its performance evaluation using ANN and ANFIS models. *J. Build. Pathol. Rehabil.* **2023**, *8*, 99. [CrossRef]
81. Mohammed, A.K.; Hassan, A.M.T.; Mohammed, A.S. Predicting the Compressive Strength of Green Concrete at Various Temperature Ranges Using Different Soft Computing Techniques. *Sustainability* **2023**, *15*, 11907. [CrossRef]
82. Pakzad, S.S.; Roshan, N.; Ghalehnovi, M. Comparison of various machine learning algorithms used for compressive strength prediction of steel fiber-reinforced concrete. *Sci. Rep.* **2023**, *13*, 3646. [CrossRef]
83. Ganasen, N.; Krishnaraj, L.; Onyelowe, K.C.; Alaneme, G.U.; Otu, O.N. Soft computing techniques for predicting the properties of raw rice husk concrete bricks using regression-based machine learning approaches. *Sci. Rep.* **2023**, *13*, 14503. [CrossRef]