

Article

Artificial Intelligence for Prediction of Physical and Mechanical Properties of Stabilized Soil for Affordable Housing

Woubishet Zewdu Taffese ^{1,*}  and Kassahun Admassu Abegaz ²

¹ Department of Civil Engineering, Aalto University, 02150 Espoo, Finland

² Department of Construction Engineering and Management, Institute of Technology, University of Gondar, Gondar P.O. Box 385, Ethiopia; Kassahun.Admassu@uog.edu.et

* Correspondence: woubishet.taffese@aalto.fi

Abstract: Soil stabilization is the alteration of physicomechanical properties of soils to meet specific engineering requirements of problematic soils. Laboratory examination of soils is well recognized as appropriate for examining the engineering properties of stabilized soils; however, they are labor-intensive, time-consuming, and expensive. In this work, four artificial intelligence based models (OMC-EM, MDD-EM, UCS-EM⁺, and UCS-EM⁻) to predict the optimum moisture content (OMC), maximum dry density (MDD), and unconfined compressive strength (UCS) are developed. Experimental data covering a wide range of stabilized soils were collected from previously published works. The OMC-EM, MDD-EM, and UCS-EM⁻ models employed seven features that describe the proportion and types of stabilized soils, Atterberg limits, and classification groups of soils. The UCS-EM⁺ model, besides the seven features, employs two more features describing the compaction properties (OMC and MDD). An optimizable ensemble method is used to fit the data. The model evaluation confirms that the developed three models (OMC-EM, MDD-EM, and UCS-EM⁺) perform reasonably well. The weak performance of UCS-EM⁻ model validates that the features OMC and MDD have substantial significance in predicting the UCS. The performance comparison of all the developed ensemble models with the artificial neural network ones confirmed the prediction superiority of the ensemble models.



Citation: Taffese, W.Z.; Abegaz, K.A. Artificial Intelligence for Prediction of Physical and Mechanical Properties of Stabilized Soil for Affordable Housing. *Appl. Sci.* **2021**, *11*, 7503.

<https://doi.org/10.3390/app11167503>

Academic Editor: Raffaele Zinno

Received: 5 July 2021

Accepted: 13 August 2021

Published: 16 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The use of natural soils for habitat development is not a new concept. As of yet, it has continued to be a topic of interest to equilibrate the imbalance between the use of natural resources for human settlement development and the excessive exploitation of nonrenewable basic industrial raw materials. The industrial production of building materials, in the long run, has become a frontier of environmental calamity, be it in natural resource degradation or excessive use of energy coupled with incalculable CO₂ emission [1,2]. That being its ecofactor, the economical aspect of the products is also mostly targeting a minute segment of the global population that can afford to expend to that scale at the dearly living cost of the majority. Offsetting such disparities requires a focus on relevant researches to ensure global welfare and improved human livelihood sustainability in the long run.

The use of stabilized natural soils as a sustainable alternative construction material could offer important economic and environmental benefits to society. Soil stabilization is the alteration of physicomechanical properties of soils to meet specific engineering requirements. There are different methods that can be selected for soil improvement such as chemical, dynamical, hydraulic, physical, and mechanical methods [3]. Chemical soil stabilization is the most-used technique that involves the addition of minerals to the natural soil such as lime, cement, silica fume, natural pozzolana, slag, and fly ash, or a combination of them [3,4]. The incorporated minerals chemically react with the soil constituents and

result in the enhancement of strength and durability. Chemical stabilization of soil usually leads to savings in construction costs of civil engineering applications such as earth wall construction, foundation, and other earthwork purposes.

Compaction is another important technique used to improve the engineering properties of earth-based construction where shrinkage can be controlled, leading to a reduction in permeability for a more stable structure [5]. The effectiveness of the compaction is usually measured by the soil's optimum moisture content (OMC) and maximum dry density (MDD). Compaction tests in the laboratory are performed to determine the OMC at which the dry densities of soils are the highest. Similar to compaction, the strength gain of soils has a paramount role in the design, construction, and long-term stability of structures of soil materials. The unconfined compressive strength (UCS) of compacted soil is usually determined by laboratory tests using an advanced machine. Though laboratory tests of OMC, MDD, and UCS are well recognized as appropriate for examining the engineering properties of stabilized soils, these tests are labor intensive, time-consuming, and costly [6]. Due to this, the proper selection of a chemical stabilizer or a combination of stabilizers and their proportion to meet natural soil, a specific engineering characteristic is a challenging activity [7].

To mitigate extensive and cumbersome laboratory testing of the OMC, MDD, and UCS of stabilized soils, it is practical to develop models that predict these values from the basic properties of natural soils and stabilizers such as plasticity, types, and quantities of stabilizers. Indeed, the properties of natural soils and stabilizers exhibit varied and uncertain behavior due to the complex and indistinct physical processes related to the formation of these materials [4]. The complexity of the behavior of the natural soil coupled with their spatial variability and the addition of stabilizers makes the development of reliable OMC, MDD, and UCS physics-based prediction models challenging. Unlike a physics-based system that performs a task by following explicit rules, such a complex problem requires intelligent systems that learn from experience. Learning a complex behavior using an artificial intelligence (AI) method is thus the best alternative.

In recent years, various artificial intelligence techniques have been applied to predict OMC, MDD, and UCS of natural and stabilized soils. For example, Das et al. [8] adopted support vector machine (SVM) and three different types of ANN, which are Bayesian regularization method (BRNN), Levenberg–Marquardt algorithm (LMNN) and differential evolution algorithm (DENN) to predict the MDD and UCS of cement stabilized soil. Based on the statistical performance measure, the authors claimed that SVM outperforms ANN models. Alavi et al. [9] proposed artificial neural networks for the prediction of MDD and OMC for stabilized soil. Multilayer perceptron (MLP) architecture was adopted to develop the ANN model. The results of the tests confirm that the proposed models are satisfactory. Suman et al. [10] developed AI models to determine the MDD and the UCS of cement stabilized soil. The employed algorithms were functional networks (FN) and multivariate adaptive regression splines (MARS) and compared their performance with four models presented in [8], which are BRNN, LMNN, DENN, and SVM. Based on statistical performance measurement, the authors concluded that the adopted algorithms perform better compared to SVM and ANN based models. Bahmed et al. [3] attempted to predict OMC and MDD of lime-stabilized clayey soils using ANN. The authors claimed that the developed ANN model can be effectively utilized to predict MDD and OMC properties of stabilized clayey soils. Le et al. [11] also developed a model to predict the UCS of soils using ANN. Based on the performance evaluation, the authors concluded that ANN can accurately predict UCS. ANN and adaptive neuro fuzzy inference system (ANFIS) have been applied to predict UCS of compacted soils by Kalkan et al. [12]. They compared the performance of both models and found that the ANFIS model is superior to the ANN. Saadat and Bayat [13] also adopted ANFIS to predict the UCS of stabilized soil. The authors compared the prediction performance of ANFIS with nonlinear regression (NLR) and claimed that ANFIS outperforms NLR. Chore and Magar [14] reported that ANN can be utilized for predicting the UCS fiber-reinforced cement stabilized fly ash mixes

accurately. They also compared its performance with the conventional multiple linear regression model (MLR) and concluded that ANN is superior to MLR.

The earlier works' attempt to predict OMC, MDD, and UCS of stabilized soils using AI approaches are encouraging. Though several types of AI algorithms can solve complex nonlinear regression problems efficiently, most of the works mainly utilized ANN. Indeed, ANN is one of the commonly applied approaches to solve several civil engineering problems, and some practical examples can be found in [15–18]. However, ANN has several limitations. For instance, a reasonable interpretation of the overall structure of the network is often challenging, and they do not provide information about the relative importance of the predictors. ANN is also not always superior to other AI models to solve problems, and it is impossible to comprehend which algorithm would surpass for a given problem all the time. As a show case, the work of Das et al. [8] demonstrated that SVM outperforms ANN models in predicting the MDD and UCS of cement-stabilized soil. In addition, the majority of the previous works employed a low number of data obtained mainly from a single experimental study.

In this work, models based on an ensemble of regression trees are developed to predict OMC, MDD, and UCS of stabilized soils. The main contribution of this work is the development of OMC, MDD, and UCS prediction models using ensemble methods and employing a wide range of stabilized soils acquired from different countries around the globe.

The structure of the remainder of the paper is as follows: In Section 2, the fundamental principle about regression tree ensembles is presented. In this section, the most commonly applied methods to form ensembles of regression trees are discussed in detail. In Section 3, the applied materials and methods are elaborated. As a data-driven method, the utilized data and modeling approach are presented in detail. The results and discussion of the findings are discussed in Section 4. In this section, the performance comparison of all the developed ensemble models with the artificial neural network ones is also presented. Finally, conclusions are provided in Section 5.

2. Regression Tree Ensembles

The fundamental principle of any ensemble method is to establish a powerful predictive model by aggregating multiple base models, each of which solves the same problem [19–23]. Though there are several ensemble models in the literature, there are two models that use regression tree learners as a base model and have proven to be powerful in solving complex regression problems for a wide range of datasets. These ensemble models are: bagging and boosting regression trees [23–25]. Unlike ANN, models based on the ensemble of regression trees are interpretable and have the potential to provide relative importance of the predictors. Use of regression tree models in the field of civil engineering is not new. There are several successful practical applications in this domain [26–29]. In this work, an optimizable ensemble method that selects either the bagging or boosting method to create an ensemble of regression trees is adopted.

2.1. Bagging Regression Tree

In bagging regression tree, the base models are formed using multiple randomly drawn bootstrapped samples from the original dataset. This procedure is conducted numerous times until a substantial subset of training datasets are made, and the same samples can be gathered more than one time. On average, each formed bootstrapped training dataset holds $N\left(1 - \frac{1}{e}\right) \approx 0.63N$ instances, where N is the total number of samples in the original dataset. The left-out instances are known as out-of-bag observations, and it is used to assess the performance of the model. The final output of the bagging regression tree model is the average of the predicted output of the individual base models, which in turn is diminishing its variance and produces better stability [21,22,24,30]. In bagging regression tree, the base model fits the training dataset $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, attaining the tree's prediction $\hat{f}(x)$ at input vector x . Bagging averages this prediction

over a collection of bootstrap samples. For each bootstrap sample D^{*t} , $t = 1, 2, \dots, T$, the model delivers prediction $\hat{f}^{*t}(x)$. The bagged estimate is the mean prediction at x from T trees as presented by Equation (1).

$$\hat{f}_{\text{bag}}(x) = \frac{1}{T} \sum_{t=1}^T \hat{f}^{*t}(x) \quad (1)$$

2.2. Boosting Regression Tree

Boosting can be characterized as an enhancement of bagging that involves multiple base models by shifting the focus toward cases that encounter challenges in performing well [24,25,30]. In contrast to bagging, boosting regression trees serially construct simple tree models with improvement from one tree model to the other and fuse them to boost the model performance. Each tree is grown from a training dataset $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, utilizing knowledge from previously grown trees. A relevant algorithm is applied to fit the training datasets $D^{(t)}$, $t = 1, 2, \dots, T$, employing a sequence of varying weights $w^{(1)}, w^{(2)}, \dots, w^{(T)}$, returning the tree's predictions $\hat{f}^{(1)}(x), \hat{f}^{(2)}(x), \dots, \hat{f}^{(T)}(x)$ for each input vector x and their corresponding weight vector w . The weight vector is usually originated by implementing an initial weight $w^{(1)}$ and continuously adjusted in each created base model depending on perceived errors. The weight is increased for cases in which the base model generates big errors and reduced for situations in which the model produces low errors. The final output of the models is a weighted sum of the individual model outputs, with the weights being larger for the better models.

One of the best-known boosting approaches is the LSBoost (least-squares boosting) algorithm, which is adopted in this work. It begins from the null model with residuals $r_i = y_i$ for all i in the training dataset $D = \{(x_i, y_i), i = 1, \dots, N\}$, where $y \in \mathbb{R}$. Then, it fits a decision tree to the residuals from the model instead of the outcome y . Sequentially, the algorithm updates the residuals by adding the newly generated decision tree into the fitted function. Each of these trees can be small to a certain extent by controlling the parameter d (number of splits) in the algorithm. By fitting small trees to the residuals, the \hat{f} is slowly boosted in areas where the performance is weak. The shrinkage parameter (learning rate) λ slows down the learning process further. For a small value of λ , the iteration number needed to attain a certain training error increases. The output of the boosted model is presented by Equation (2).

$$\hat{f}_{\text{LSBoost}}(x) = \hat{f}(x) = \sum_{t=1}^T \lambda \hat{f}^t(x). \quad (2)$$

3. Materials and Methods

3.1. Raw Data

To establish a reliable database for the development of OMC, MDD, and UCS prediction models, experimental data of natural and stabilized soils were collected from the published literature with certain selection criteria. This is because a data-driven model is not a common practice in this research area and there is no readily available large enough dataset. To acquire experimental data, databases of Scopus and Web of Science were largely used. Both are abstract and citation databases of the peer-reviewed literature, delivering a complete citation search by giving access to numerous databases. A set of queries that comprise the manuscript's title, topic, abstract, and keywords were carried out on every database to choose works reporting "stabilized soils". Any duplicated records from the databases were removed and then manually checked to select suitable ones. A total of 79 scientific papers which have performed experiments in stabilized soil were found and used to gather the data. Indeed, gathering data from the works of the literature was not a straightforward activity. The typical challenges were (i) the use of different measuring units, (ii) some results presented in the form of charts, and (iii) the use of different soil

classification methods. Translating all the data in preferred units and format to obtain well-sounded data required great attention and was a time-demanding task.

Data consisting of 408 observations and 13 features that comprise information regarding the quantity of soils and stabilizers (cement, lime, pozzolans, and fly ash), Atterberg limits (LL, PL, and PI), compaction properties (OMC and MDD), soil classification, and unconfined compressive strength of soils at different ages were collected. The dataset covers a wide range of soils from 12 countries in Africa, America, the Middle East, South Asia, and Oceania.

3.2. Modelling Process

The workflow of the OMC, MDD, and UCS prediction models is illustrated in Figure 1. The initial task is retrieving the gathered experimental data that comprise information regarding the proportion of soils and stabilizers, Atterberg limits, compaction properties, soil classification, curing ages, and unconfined compressive strength of the stabilized soils. Then, data preparation follows, in which data preprocessing, feature engineering, feature selection, and scaling are performed. The next major step is to fit the data using the ensemble method (either bagging or boosting regression trees). The performance of the models is then evaluated using a previously unseen dataset. The model training and validation process is iterated by adjusting the hyperparameters till the best result is obtained. The detail of the major activities is discussed below.

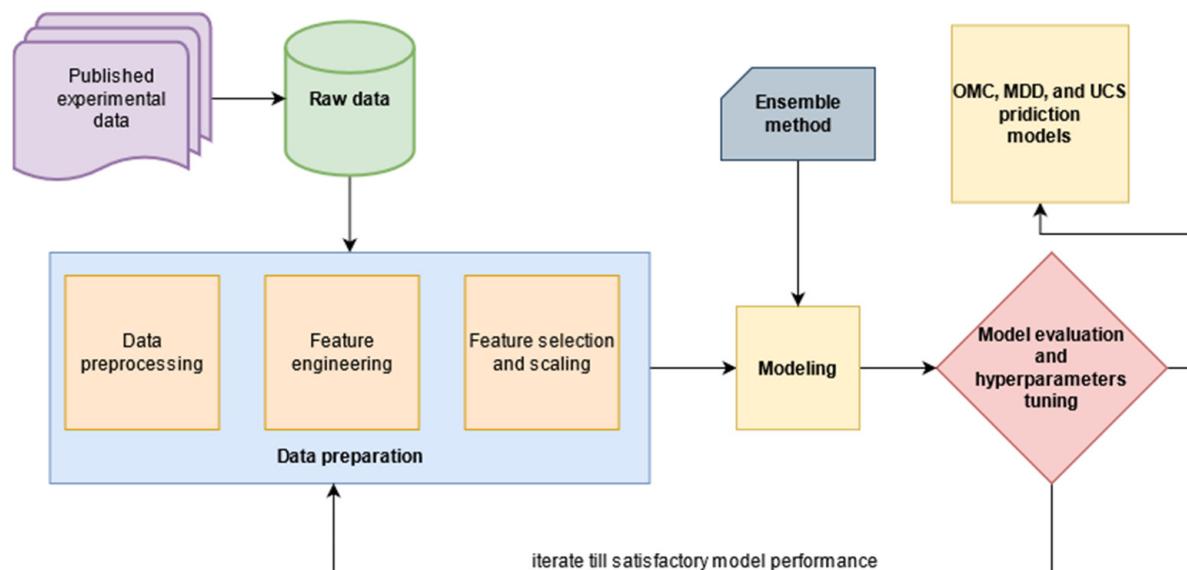


Figure 1. The development process of OMC, MDD, and UCS prediction models.

3.2.1. Data Preparation

Data Preprocessing and Feature Engineering

In the modeling process, feature engineering is the process of representing the data appropriately. It is one of the key components as it considerably influences the performance of a model. No AI algorithm is able to predict data for which it has no appropriate information. Feature engineering is often performed by the domain expert, and most of the required feature engineering activities have already been carried during the collection of the data from previously published works. Here, after careful examination of the distribution of each feature, three features (pozzolans, fly ash, and curing age) which do not have representative enough data are excluded from the database. For instance, fly ash and natural pozzolans are utilized only in 36 and 48 cases, respectively. This is a very small number of cases compared with other types of stabilizers. For comparison, the number of observations in the case of cement and lime is 153 and 250, respectively. Thus, retaining

pozzolans and fly ash in the database causes data imbalance, which ultimately affects the performance of the model. Similarly, there are unconfined compressive strength tests performed at the age of 1, 7, 14, 28, 32, 64, and 90 days in the database, but except UCS carried out at the age of 28 days, the other age groups represent only between 1% and 15%. Thus, unconfined compressive strength tests carried out only at the age of 28 days are considered. Finally, nine features with 190 instances are selected. The data are presented in Table S1 in Supplementary Materials, and description of the features is presented in Table 1.

Table 1. Description of the features considered in the dataset.

Category	No.	Variable Subcategory	Range	Unit	Data Type
Amended soils	1	Soil	70–100	[% by mass]	Numerical
	2	Cement	0–30	[% by mass]	Numerical
	3	Lime	0–30	[% by mass]	Numerical
Atterberg limits	4	Liquid limit	18–102	[%]	Numerical
	5	Plastic limit	12–58.24	[%]	Numerical
	6	Plasticity index	0–70	[%]	Numerical
Soil classification	7	USCS	CH	—	Categorical
			CL		
			CL-ML		
Compaction properties	8	Optimum moisture content	5.40–36.80	[%]	Numerical
Strength	9	Maximum dry density	1200–2210	[kg/m ³]	Numerical
	10	Unified comprehensive strength	55.31–5400	[kN/m ²]	Numerical

The soils are stabilized by cement and/or lime. Though there are different types of cement and lime, the literature where the data gathered from does not explicitly mention their types. All the stabilized soils were classified based on the Unified Soil Classification System (USCS). A total of five kinds of soil types are employed in the dataset. These are clay of low plasticity (CL), clay of high plasticity (CH), silt (ML), silt of high plasticity (MH), and clayey silt (CL-ML). The soil types of CL and ML represent 50% and 25%, respectively. The remaining 25% are represented by CH, CL-ML, and MH. The unconfined compressive strength of the soils considered in this dataset was those taken at the age of 28 days and their values range from 55.31–5400 kN/m². Descriptive statistics of the data are illustrated in Table 2.

Table 2. Descriptive statistics of the dataset.

	Soil	Cement	Lime	LL	PL	PI	MDD	OMC	UCS
Count	190	190	190	190	190	190	190	190	190
mean	93.71	3.75	2.55	39.53	22.74	16.79	1778.69	13.98	2302.36
std	4.67	4.31	4.07	16.69	9.35	12.63	226.22	7.33	1170.42
Min	70.00	0.00	0.00	18.00	12.00	0.00	1200.00	5.40	55.31
1st quantile	94.00	0.00	0.00	29.00	17.00	7.00	1626.50	8.80	1610.00
Median	94.00	4.00	2.00	35.00	20.00	15.50	1820.00	11.30	2260.00
3rd quantile	95.00	6.00	4.00	45.25	26.00	21.00	1960.00	16.50	3000.00
Max	100.00	30.00	30.00	102.00	58.24	70.00	2210.00	36.80	5400.00

Feature Selection and Scaling

Feature selection is the process of selecting the most relevant features from the data. This is because some features may be highly correlated and thus redundant to a certain degree or even they may be irrelevant. Generally, feature selection methods are categorized into three groups: filter, wrapper, and embedded [31,32]. The filter method is independent of the learning algorithms and relies only on the inherent nature of the data. The wrapper

method demands a prespecified algorithm and is based on the utilized algorithm. The performance of each feature is adopted as the criterion for defining the final subset of features. This technique, compared to the filter, is computationally expensive but produces better accuracy. The embedded method includes the feature selection process as a component of model development. This approach is computationally inexpensive and improves the prediction performance of the predictors. The adopted ensemble method performs embedded feature selection so that it internally selects relevant features, thus enhancing the prediction performance of the model.

Numerous AI algorithms demand that the selected features are on the same scale for optimal performance, which is usually achieved by transforming the features in the range $[0, 1]$ or a standard normal distribution with zero mean and unit variance. However, the adopted ensemble methods are based on regression trees, which do not require scaling.

3.2.2. Model Training

The data were divided into training and test subsets that represented 80% and 20% of the data, respectively. The training dataset is used to fit the predictors, whereas test dataset is applied to evaluate the predictive performance of the fitted or developed model. In ensemble method, one way to achieve differences between base models is to train each model on a different subset of the available training data. Models are trained on different subsets of the training data naturally through the use of resampling methods such as cross-validation and the bootstrap, which are designed to estimate the average performance of the model generally on unseen data. Though bagging regression tree forms the training and validation set based on the embedded sampling procedure, cross-validation resampling technique is applied for both methods. There are diverse types of cross-validation procedures. In the case of a limited dataset, K -fold cross-validation technique is the most desirable choice to achieve an unbiased prediction, which in turn enhances the generalization ability of the model without overfitting [21]. In K -fold cross-validation, the training data are arbitrarily partitioned into K subsets with approximately the same sizes. Each of the K subsets is employed as a validation dataset for assessing the performance of the model and the remaining $(K - 1)$ subsets as a training dataset. In total, K models are fit, and K validation statistics are obtained. The performance evaluations from the K -folds are averaged to measure the overall performance of the model. In this work, 10-fold cross-validation was applied.

Regression trees are fitted on bootstrap samples and aggregated to create bagging regression trees. Boosting regression trees are also formed by fitting multiple regression trees iteratively in such a way that the model training at a given step depends on the models fitted at the previous steps. Every new model focuses its efforts on the most problematic instances, ultimately forming a strong learner. Using optimizable ensemble method, either bagging or boosting regression trees, four models to predict compaction and strength of stabilized soils are developed. Seven features under the category of amended soils, Atterberg limits, and soil classification are employed to predict OMC and MDD of soils. Based on the input variables, two UCS prediction models (UCS-EM⁺ and UCS-EM⁻) are developed. UCS-EM⁺, besides the seven features, employed two more features describing the compaction properties (OMC and MDD), whereas UCS-EM⁻ employed the same features considered in the OMC and MDD models. Classification of the models is presented in Table 3.

Table 3. Description of features considered in each model.

Properties to Be Examined	Models	Number of Features	Feature Details
Compaction	OMC-EM	7	All the variables under the category of amended soils (soil, cement, and lime), Atterberg limits (LL, PL, and PI), and soil classification.
	MDD-EM	7	All the variables under the category of amended soils (soil, cement, and lime), Atterberg limits (LL, PL, and PI), soil classification, and the compaction properties (OMC and MDD).
	UCS-EM ⁺	9	All the variables under the category of amended soils (soil, cement, and lime), Atterberg limits (LL, PL, and PI), and soil classification.
Strength	UCS-EM ⁻	7	All the variables under the category of amended soils (soil, cement, and lime), Atterberg limits (LL, PL, and PI), and soil classification.

Tuning the hyperparameters of any AI-based models is essential to optimize their performance. In the present work, with the combination of 10-fold crossvalidation, hyperparameter tuning is carried out using Bayesian optimization to improve the performance of the adopted models by finding the optimal combination of hyperparameter values that minimizes the loss function (mean square error). The considered hyperparameters and the search range of each hyperparameter are: (i) Ensemble method—Bag and LSBoost; (ii) Minimum leaf size—integers log-scaled in the range [1, max (2, floor (n/2))], where n is the number of observations; (iii) Number of learners—integers log-scaled in the range [10, 500]; (iv) Learning rate—real values log-scaled in the range [0.001, 1]; and (v) Number of predictors to sample—integers in the range [1, max (2, p)], where p is the number of predictor variables.

3.2.3. Model Evaluation

Once the optimum hyperparameters have been obtained for each model, their performance is assessed by measuring errors of mean square error (*MSE*), root mean square error (*RMSE*), and mean absolute error (*MAE*) on test dataset. *MSE* is the average of the squared difference between the actual and the predicted value. It is the most commonly applied loss function for regression models. *MSE* is calculated using Equation (3).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (3)$$

RMSE is just the square root of the *MSE*. It has the same unit as the target variable. Irregularly, *RMSE* is preferable than *MSE* because understanding the error values of *MSE* is tricky due to the squaring effect, particularly if the target variable describes quantities in a unit of measurements. The formula of *RMSE* is described by Equation (4).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

The *MAE*, also known as the absolute loss, is the mean of the absolute errors (the deviation between the actual and the predicted value). Similar to *RMSE*, *MAE* is measured in the same units as the target variable. It is mathematically denoted by Equation (5).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

Coefficient of determination, R^2 , is another valuable quantitative measure of goodness of a prediction for a regression model. It is the fraction of response variance that is achieved by the model, and it can be described as a standardized version of the *MSE*, for greater interpretability of the model's performance. The value of R^2 can be computed by the mathematical relation expressed by Equation (6). The value of R^2 is bounded between 0

and 1 in case of model training, but it can become negative in case of testing. If $R^2 = 1$, the model fits the data perfectly with a corresponding $MSE = 0$.

$$R^2 = 1 - \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{MSE}{Var(y)} \quad (6)$$

where n is the number of observations, y_i is the actual target value, \hat{y}_i is the predicted output value, \bar{y} the mean value of the actual target, and Var is the variance of the target variable.

4. Results and Discussion

In this section, the performance of the developed OMC, MDD, and UCS models is presented. Examining the generalization performance, how well the models can make predictions for data that was not observed during training, is very essential compared to how well they fit the training set. The coefficient of determination, R^2 , is one of the statistical measures applied to examine the generalization performance of all the models. The R^2 score indicates how well the developed OMC, MDD, and UCS models explain and predict future outcomes. It yields a score between 0 and 1. The training performance of all models is illustrated in Figure 2. Regression plot for predicted vs. measured OMC, MDD, and UCSs with their corresponding validation R-square scores, showing the true versus the predicted response. It can be perceived from Figure 2 that the score of OMC-EM, MDD-EM, and UCS-EM⁺ models exceed 0.50. Observably, OMC-EM performs best ($R^2 = 0.76$), followed by UCS-EM⁺ ($R^2 = 0.69$), and MDD-EM ($R^2 = 0.59$). This validates that these models reasonably track their corresponding target features during the training stage. Among all, UCS-EM⁻ perform least with ($R^2 = 0.49$). This model entails all features under the category of amended soils (soil, cement, and lime), Atterberg (LL, PL, and PI), and soil classification, but does not entail features that describe the compaction properties (OMC and MDD). Hence, the weak performance of UCS-EM⁻ model corroborates that the features OMC and MDD have significant importance in predicting UCS.

It can be noticed that all the models exhibit a slight tendency to underestimate or overestimate their corresponding target variables in which the number of observations is very limited. For example, measurements of OMC > 25% and UCS > 3500 kN/m² represent only about 12% and 13% of the total observation, respectively. This insufficient number of observations in the data might be the cause for the underestimation in those ranges. This is a normal phenomenon as the algorithm did not obtain adequate observations to learn and generalize. Incorporating more observations could enhance the performance of the models.

Scores of other statistical performance indicators ($RMSE$, MSE , and MAE) on the test dataset for all the developed models are given in Table 4. Indeed, it is impossible to compare one model with another by considering these indicators, except UCS-EM⁺ and UCS-EM⁻. This is because they are predicting different variables. The lower the statistical errors are the superior the performance of the model. The $RMSE$, MSE , and MAE of UCS-EM⁺ model are considerably lower than the UCS-EM⁻, confirming its superiority. It can also be noticed from MAE results that the average prediction errors of OMC-EM, MDD-EM, UCS-EM⁺, and UCS-EM⁻ are 2.68%, 110.06 kg/m³, 472.33 kN/m², and 622.97 kN/m², respectively. This confirms that the three ensemble models (OMC-EM, MDD-EM, and UCS-EM⁺) performed rationally well on previously unseen data, considering the fact that their corresponding median values are 11.30%, 1820 kg/m³, and 2260 kN/m². All the results are valid only for the employed dataset. The performance of each model could be enhanced if more data are utilized. The optimized hyperparameters are also presented in the Table 4. As presented in Section 3.2.2, five hyperparameters were considered to optimize the performance of the models. Bagged regression trees yielded optimal performance for all models. The minimum leaf size, number of learners, and number of predictors for the samples of each model are mostly different.

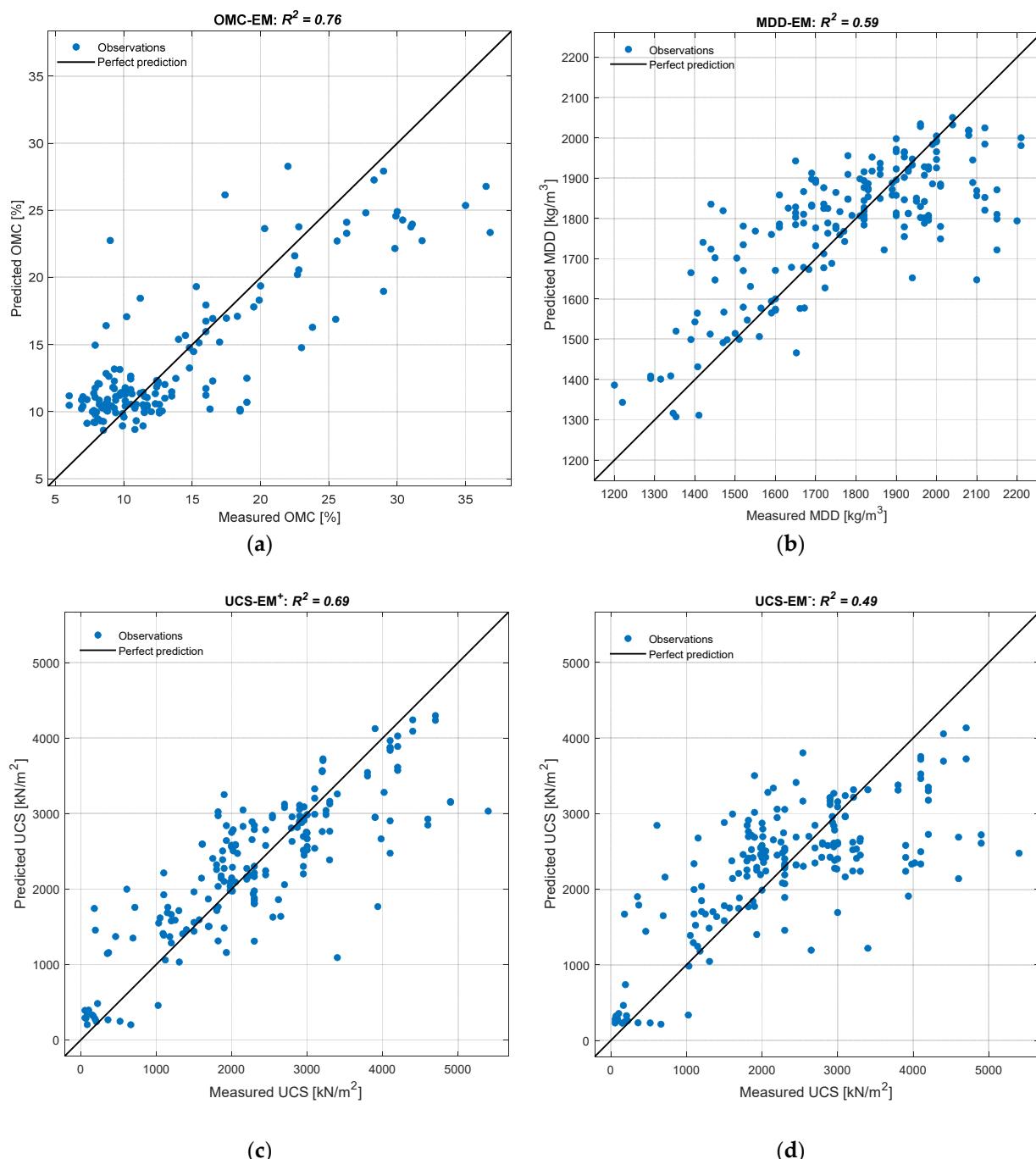


Figure 2. Regression plot and validation R-square scores of: (a) OMC-EM, (b) MDD-EM, (c) UCS-EM⁺ and (d) UCS-EM⁻.

Table 4. Statistical performance measurement of the developed models on unseen data and the optimized hyperparameters.

Models	Validation Results				Ensemble Method	Optimized Hyperparameters		
	RMSE	R^2	MSE	MAE		Minimum Leaf Size	No of Learners	No. of Predictors
OMC	3.64	0.76	13.23	2.68	Bag	1	475	7
MDD	145.60	0.59	21,201	110.06	Bag	1	264	5
UCS-EM ⁺	649.73	0.69	422,149	472.33	Bag	4	148	7
UCS-EM ⁻	833.41	0.49	694,752	622.98	Bag	2	38	6

To compare the performance of the developed four ensemble models, four other models employing ANN algorithms are developed. These models are OMC-ANN, MDD-ANN, UCS-ANN⁺, and UCS-ANN⁻. All the ANN models have three layers: an input, a hidden, and an output layer. The number of the input neurons for each model corresponds to the number of the predictors, which is seven neurons for OMC-ANN, MDD-ANN, and UCS-ANN⁻ and nine neurons for UCS-ANN⁺. The number of neurons utilized in the hidden layer was ten, which was determined based on the generalization error after performing several trainings. The number of neurons at the output layer is one for all models, which is the feature going to be predicted. The data were also randomly divided into three clusters: training, validation, and test datasets, which hold 70%, 15%, and 15% of the dataset, respectively. Once the datasets are ready, the network was trained using a Levenberg–Marquardt algorithm.

The difference between the actual and the predicted OMC, MDD, and UCSs by the developed four ANN models are calculated, and their distributions are visualized with a boxplot in Figure 3. The median of the errors is designated by a red line within the blue box, embracing the middle 50% (25th–75th percentiles) of the errors. It can be seen from Figure 3 that the medians of errors of all the models are closer to either the first or the third quartile. For instance, the medians of the errors of OMC-ANN and UCS-ANN⁺ models are closer to the first quartile. This means that the distributions of the errors are slightly skewed to the right. The whiskers stretch from the ends of the box to the smallest and largest error values. Errors greater than 1.5 box lengths above the whiskers are outliers and characterized by a red plus sign. It can be observed that there are a significant number of outliers in all models.

Statistical performance indicators (*MSE* and R^2) of all the ensemble and ANN models are given in Table 5. The lower the *MSE* and the higher the R^2 are, the more superior the model. It can be observed from Table 5 that the *MSE* errors of all models which employed ensemble of regression trees are lower than the corresponding models based on ANN. For instance, the *MSE* of the OMC-EM compared to OMC-ANN is lower by 34%. The R^2 of all the ensemble models is higher than ANN models. For example, the R^2 of OMC-EM is higher than OMC-ANN by 31%. All these confirm that the ensemble methods outperformed the ANN models on the utilized dataset.

Table 5. Statistical performance measurements for models based on ensemble methods and ANN on unseen data.

Algorithm	Model Names	MSE	R^2
Ensemble of regression trees	OMC-EM	13.23	0.76
	MDD-EM	21,201	0.59
	UCS-EM ⁺	422,149	0.69
	UCS-EM ⁻	694,752	0.49
	OMC-ANN	20.10	0.58
Artificial neural network	MDD-ANN	31,011	0.50
	UCS-ANN ⁺	519,445	0.62
	UCS-ANN ⁻	874,933	0.45

The developed ensemble models are a promising approach to evaluate the performance of compaction and strength properties of stabilized soils. Housing blocks based on stabilized soils have been already gaining recognition in many emerging countries, especially in the global south as the weather condition favors their wide application. Indeed, the stabilizers type varies depending on their availability. For instance, a recent research work by Admassu K. [6,7] demonstrated the possibility of stabilizing soils using locally available raw lime and raw natural pozzolan on three types of soils. Across the considered soil range, varying improvement effects were recorded in OMC, MDD, and UCS. In general, the soils were effectively stabilized to certainly induce physical and mechanical property changes on the main ingredients, making them fit for the production of wall building blocks and

jointing mortar. The trend shades promising limelight on the initiated make for a futuristic affordable, sustainable, and ecofriendly alternative earth-based built environment. With the availability of more and more such experiments (data), the same models with new training dataset could be utilized to predict the physical and mechanical properties of stabilized soils.

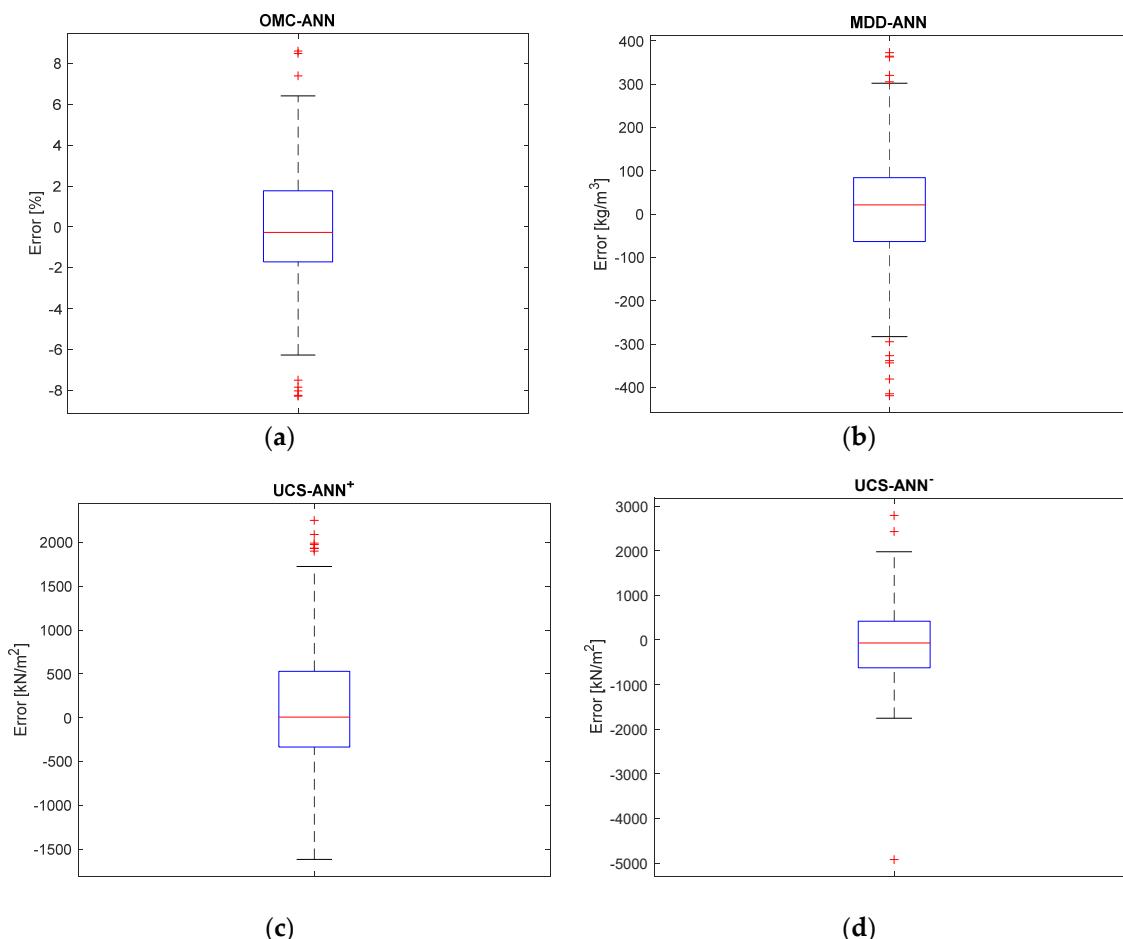


Figure 3. Distribution of errors in models of: (a) OMC-ANN, (b) MDD-ANN, (c) UCS-ANN⁺, and (d) UCS-ANN⁻.

Apart from that, incorporating additional features which describe the soil and the stabilizers such as soil grading, soil compaction state, or describing stabilizers explicitly (e.g., type of cement and lime) could enhance the prediction capability of the models. Both compaction and grading play a significant role in keeping the integrity of the stabilized soil. The type and proportion of the stabilizers content also significantly influence the development of the unconfined compressive strength of the stabilized soil. The inclusion of experimental data in the database in which the proportion of the utilized stabilizers in each range such as 0–10%, 10–20%, and 20–30% is sufficient enough. The models can be adopted to formulate the proportion of stabilizers that could meet the desired compaction properties (OMC and MDD) or strength of stabilized soils. It is important to note that the prediction accuracy of the current model is not high enough due to limited data. With sufficient data (both observations and number of predictors), the model performance could be enhanced, and it can substitute the cumbersome and expensive laboratory tests of stabilized soils. Moreover, it could assist the scientific community to obtain a better insight into how different complex soil types and stabilizers in relative proportions could be estimated to predict the possibly achievable compaction and strength properties of stabilized soils.

5. Conclusions

A total of four artificial intelligence based models to predict the compression strength properties of stabilized soils were developed. These are OMC-EM, MDD-EM, UCS-EM⁺, and UCS-EM⁻. An optimizable ensemble method which selects either bagging or boosting regression trees to create an ensemble was adopted. To establish a reliable database for the development of the models, experimental data of a wide range of stabilized soils were collected from previously published works. The OMC-EM, MDD-EM, and UCS-EM⁻ models employed seven features which describe the proportion and types of stabilized soils (soil, cement, and lime), Atterberg limits (LL, PL, and PI) and classification group of soils. The UCS-EM⁺ model utilized a total of nine input features. These are all the seven features as well as the OMC and MDD properties of the stabilized soils. The performance of all the developed models demonstrated their promising application in the prediction of OMC, MDD, and UCS. The weak performance of UCS-EM⁻ model corroborates that the features OMC and MDD play a significant role in predicting UCS. The performances of all the developed ensemble models were compared with the artificial neural network based models. The comparison corroborated that the ensemble models outperform the ANN models. The performance of all the models could be improved further with more data and can be applied to determine the optimal proportion of stabilizers that could meet the desired OMC, MDD, and UCS.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/app11167503/s1>, Table S1: Dataset. References [33–42] refer to the supplementary material.

Author Contributions: Conceptualization, W.Z.T. and K.A.A.; methodology, W.Z.T. and K.A.A.; software, W.Z.T.; validation, W.Z.T.; formal analysis, W.Z.T.; investigation, W.Z.T. and K.A.A.; resources, W.Z.T. and K.A.A.; data curation, W.Z.T. and K.A.A.; writing—original draft preparation, W.Z.T. and K.A.A.; writing—review and editing, W.Z.T. and K.A.A.; visualization, W.Z.T.; project administration, K.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in Supplementary Material here.

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

References

1. UNEP. 2018 *Global Status Report: Towards a Zero-Emission, Efficient and Resilient Buildings and Construction Sector*; UNEP: Nairobi, Kenya, 2018.
2. Taffese, W.Z.; Abegaz, K.A. Embodied Energy and CO₂ Emissions of Widely Used Building Materials: The Ethiopian Context. *Buildings* **2019**, *9*, 136. [[CrossRef](#)]
3. Bahmed, I.T.; Harichane, K.; Ghrici, M.; Boukhatem, B.; Rebouh, R.; Gadouri, H. Prediction of geotechnical properties of clayey soils stabilised with lime using artificial neural networks (ANNs). *Int. J. Geotech. Eng.* **2019**, *13*, 191–203. [[CrossRef](#)]
4. Alavi, A.H.; Gandomi, A.H.; Gandomi, M.; Hosseini, S.S.S. Prediction of maximum dry density and optimum moisture content of stabilised soil using RBF neural networks. *IES J. Part A Civ. Struct. Eng.* **2009**, *2*, 98–106. [[CrossRef](#)]
5. Günaydin, O.; Özbeyaz, A.; Söylemez, M. Regression Analysis of Soil Compaction Parameters Using Support Vector Method. *Celal Bayar Univ. J. Sci.* **2018**, *14*, 443–447. [[CrossRef](#)]
6. Admassu, K. Engineered soil and the need for lime-natural pozzolan mixture percentage approximation. *SINET Ethiop. J. Sci.* **2018**, *41*, 70–79.
7. Admassu, K. Method of Amended Soils for Compressed Block and Mortar in Earthen Construction. *Zede J.* **2019**, *37*, 69–84.
8. Das, S.K.; Samui, P.; Sabat, A.K. Application of Artificial Intelligence to Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilized Soil. *Geotech. Geol. Eng.* **2011**, *29*, 329–342. [[CrossRef](#)]
9. Alavi, A.H.; Gandomi, A.H.; Mollahassani, A.; Heshmati, A.A.; Rashed, A. Modeling of maximum dry density and optimum moisture content of stabilized soil using artificial neural networks. *J. Plant Nutr. Soil Sci.* **2010**, *173*, 368–379. [[CrossRef](#)]
10. Suman, S.; Mahamaya, M.; Das, S.K. Prediction of Maximum Dry Density and Unconfined Compressive Strength of Cement Stabilised Soil Using Artificial Intelligence Techniques. *Int. J. Geosynth. Ground Eng.* **2016**, *2*, 11. [[CrossRef](#)]

11. Le, H.A.; Nguyen, T.A.; Nguyen, D.D.; Prakash, I. Prediction of soil unconfined compressive strength using artificial neural network model. *Vietnam J. Earth Sci.* **2020**, *42*, 255–264. [[CrossRef](#)]
12. Kalkan, E.; Akbulut, S.; Tortum, A.; Celik, S. Prediction of the unconfined compressive strength of compacted granular soils by using inference systems. *Environ. Geol.* **2009**, *58*, 1429–1440. [[CrossRef](#)]
13. Saadat, M.; Bayat, M. Prediction of the unconfined compressive strength of stabilised soil by Adaptive Neuro Fuzzy Inference System (ANFIS) and Non-Linear Regression (NLR). *Geomech. Geoengin.* **2019**, *1*–12. [[CrossRef](#)]
14. Chore, H.S.; Magar, R.B. Prediction of unconfined compressive and brazilian tensile strength of fiber reinforced cement stabilized fly ash mixes using multiple linear regression and artificial neural network. *Adv. Comput. Des.* **2017**, *2*, 225–240. [[CrossRef](#)]
15. Rachedi, M.; Matallah, M.; Kotronis, P. Seismic behavior & risk assessment of an existing bridge considering soil-structure interaction using artificial neural networks. *Eng. Struct.* **2021**, *232*, 111800. [[CrossRef](#)]
16. Taffese, W.Z.; Al-Neshawy, F.; Sistonen, E.; Ferreira, M. Optimized neural network based carbonation prediction model. In Proceedings of the International Symposium Non-Destructive Testing in Civil Engineering (NDT-CE), Berlin, Germany, 15–17 September 2015; pp. 1074–1083.
17. de Souza, C.P.G.; Kurka, P.R.G.; Lins, R.G.; de Araújo, J.M. Performance comparison of non-adaptive and adaptive optimization algorithms for artificial neural network training applied to damage diagnosis in civil structures. *Appl. Soft Comput.* **2021**, *104*, 107254. [[CrossRef](#)]
18. Taffese, W.Z.; Sistonen, E. Neural network based hygrothermal prediction for deterioration risk analysis of surface-protected concrete façade element. *Constr. Build. Mater.* **2016**, *113*, 34–48. [[CrossRef](#)]
19. Rokach, L.; Maimon, O. *Data Mining with Decision Trees: Theory and Applications*; World Scientific Publishing Co., Ltd.: Singapore, 2008.
20. Seni, G.; Elder, J.F. *Ensemble Methods in Data Mining: Improving Accuracy through Combining Predictions*; Morgan & Claypool: San Rafael, CA, USA, 2010. [[CrossRef](#)]
21. Alpaydin, E. *Introduction to Machine Learning*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2010. [[CrossRef](#)]
22. Sutton, C.D. Classification and regression trees, bagging, and boosting, In *Handbook of Statistics: Data Mining and Data Visualization*; Rao, C.R., Wegman, E.J., Solka, J.L., Eds.; Elsevier B.V.: Amsterdam, The Netherlands, 2005; pp. 303–330. [[CrossRef](#)]
23. Witten, I.H.; Frank, E.; Hall, M.A. *Data Mining: Practical Machine Learning Tools and Techniques*; Morgan Kaufmann: Burlington, MA, USA, 2011.
24. Hsieh, W.W. *Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels*; Cambridge University Press: Cambridge, UK, 2009. [[CrossRef](#)]
25. Cichosz, P. *Data Mining Algorithms: Explained Using R*; John Wiley & Sons, Ltd.: Chichester, UK, 2015. [[CrossRef](#)]
26. Taffese, W.Z.; Sistonen, E.; Puttonen, J. Prediction of Concrete Carbonation Depth using Decision Trees. In Proceedings of the 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine, Bruges, Belgium, 22–24 April 2015; pp. 415–420.
27. Malinowska, A. Classification and regression tree theory application for assessment of building damage caused by surface deformation. *Nat. Hazards* **2014**, *73*, 317–334. [[CrossRef](#)]
28. Mariniello, G.; Pastore, T.; Menna, C.; Festa, P.; Asprone, D. Structural damage detection and localization using decision tree ensemble and vibration data. *Comput. Civ. Infrastruct. Eng.* **2020**, *36*, 1129–1149. [[CrossRef](#)]
29. Taffese, W.Z.; Sistonen, E. Significance of chloride penetration controlling parameters in concrete: Ensemble methods. *Constr. Build. Mater.* **2017**, *139*, 9–23. [[CrossRef](#)]
30. Berthold, M.R.; Borgelt, C.; Höppner, F.; Klawonn, F. *Guide to Intelligent Data Analysis: How to Intelligently Make Sense of Real Data*; Springer: London, UK, 2010. [[CrossRef](#)]
31. Stańczyk, U. Feature evaluation by filter, wrapper, and embedded approaches. In *Feature Selection for Data and Pattern Recognition*; Stańczyk, U., Jain, L.C., Eds.; Springer: Berlin, Germany, 2015; pp. 29–44. [[CrossRef](#)]
32. Tuv, E.; Borisov, A.; Runger, G.; Torkkola, K. Feature selection with ensembles, artificial variables, and redundancy elimination. *J. Mach. Learn. Res.* **2009**, *10*, 1341–1366. [[CrossRef](#)]
33. Aytekin, M.; Nas, E. Soil Stabilization with Lime and Cement. *Teknik Dergi/Tech. Journal Turk. Chamb. Civil Eng.* **1998**, *9*, 1573–1585.
34. Bell, F.G. Lime stabilisation of clay soils. *Bull. Int. Assoc. Eng. Geol.* **1989**, *39*, 67–74. [[CrossRef](#)]
35. Burroughs, V.S. Quantitative Criteria for the Selection and Stabilisation of Soils for Rammed Earth Wall Construction. Doctoral Dissertation, University of New South Wales, Sydney, New South Wales, Australia, 2001.
36. Harichane, K.; Ghrici, M.; Kenai, S.; Grine, K. Use of natural pozzolana and lime for stabilization of cohesive soils. *Geotech. Eng.* **2011**, *29*, 759–769. [[CrossRef](#)]
37. Harichane, K.; Mohamed, G.; Wiem, K.; Hanifi, M. Effect of the Combination of Lime and Natural Pozzolana on the Durability of Clayey Soils. *Electron. J. Geotech. Eng.* **2010**, *15*, 1194–1209.
38. Hasan, H.; Dang, L.; Khabbaz, H.; Fatahi, B.; Terzaghi, S. Remediation of Expansive Soils Using Agricultural Waste Bagasse Ash. *Procedia Eng.* **2016**, *143*, 1368–1375. [[CrossRef](#)]
39. Phanikumar, B.R.; Raju, E.R. Compaction and Strength Characteristics of an Expansive Clay Stabilised with Lime Sludge and Cement. *Soils Found.* **2020**, *60*, 129–138. [[CrossRef](#)]
40. Siddique, A.; Hossain, M.A. Effects of Lime Stabilisation on Engineering Properties of an Expansive Soil for Use in Road Construction. *J. Soc. Transp. Traffic Stud.* **2011**, *2*, 1–8.

41. Solanki, P.; Khoury, N.N.; Zaman, M.M. *Engineering Properties of Stabilized Subgrade Soils for Implementation of the AASHTO 2002 Pavement Design Guide*; Department of Transportation: Norman, OK, USA, 2009.
42. Workie, A.G.; Alam, S. Effect of Lime and Fly Ash on Load Bearing Capacity of Expansive Clay Soil. *Int. J. Innov. Technol. Explor. Eng.* **2019**, *8*, 554–563. [[CrossRef](#)]