Predicting compressed earth blocks compressive strength by means of machine learning models

Sarah Nassar [a,c,*], Cédric Baudrit [b,c], Jacqueline Saliba [d], Nadia Saiyouri [a,c]

[a] *Univ. Bordeaux, CNRS, Bordeaux INP, I2M, UMR 5295, Talence F-33400, France*
[b] *Univ. Bordeaux, CNRS, INRAE, Bordeaux INP, I2M, UMR 5295, USC 1368, Talence F-33400, France*
[c] *Arts et Metiers Institute of Technology, CNRS, Bordeaux INP, Hesam Universite, I2M, UMR 5295, Talence F-33400, France*
[d] *College of Engineering and Technology, American University of the Middle East, Egaila 54200, Kuwait*

Keywords:
Compressed Earth Blocks
Compressive Strength
Machine Learning
Classification
SHapley Additive exPlanations (SHAP)

A B S T R A C T

Compressed Earth Blocks (CEB) are an interesting alternative to conventional masonry units. They are unfired and provide a thermal comfort in the constructions. However, their compressive strength needs to be assessed to ensure a mechanical stability. The latter depends on the soil variability, as well as several manufacturing parameters such as water content, compaction pressure, stabilizer type and proportion. This is challenging as it requires time and effort to adapt the parameters to achieve satisfactory results. In this study, machine learning classification models were trained using historical data for predicting CEB compressive strength. Voting Classifier (VC) provided the highest performance with an accuracy of 78 %. SHapley Additive exPlanations (SHAP) were used to identify and prioritize the features in the model's decision-making process. The compaction pressure and soil granularity were the most decisive parameters. VC was also tested to assess the compressive strength of sediment-based CEB manufactured at the laboratory scale.

1. Introduction

In the context of the ecological transition, the construction sector is facing an environmental challenge and seeks alternatives to cementitious materials that have an important carbon footprint. Earth constructions are an interesting option that existed since ancient times and more than a third of the world's population live today in such constructions [1].

Among different types of earth constructions, CEB emerged in the 20th century [2] and are currently gaining momentum. Their fabrication process is simple and does not require firing of the blocks at elevated temperatures. They are recyclable [3] and provide interesting hygrothermal properties. CEB absorb ten times more humidity than fired blocks [4] and reduce the CO_2 emissions as demonstrated by Morton et al. [5]. They are a mixture of soil, water, and in some cases, stabilizers compacted into a mold. Soil extracted locally or dredged sediments can be used and thus can have very variable properties depending on the location of extraction [6,7]. Dredged sediments could present spatial and temporal variability since they are deposits of transported insoluble materials issued from the natural environment and surrounding human activity and discharges. Therefore, the reuse of such materials is challenging and requires testing as their properties are variable and the resulting CEB performance is not promised. Nassar et al. [6] considered eight locations on the Arcachon Bay in France to investigate the sediments' properties for their reuse in CEB and found different textures for the same bay. Belayali et al. [8] also used dredged sediments for the fabrication of CEB. Serbah et al. [9] valorized sediments in CEB but had to add sand to adapt their granulometry to the recommended ranges [10]. These ranges are for guidance and not an obligation. Nagaraj et al. [11] explored using the soil in its natural texture and adapting the stabilization with lime and cement to reach the required performance. Experimental testing requires time, costs money and is a waste of material. In addition, the optimization is not evident due to numerous influencing parameters and complex relations.

Compressed Earth Blocks may be viewed as complex systems with intra/inter interactions between individual components (sediments, sand, stabilizing …) at various levels. The uncertain environments such as soil variability, lack of knowledge about physicochemical interactions, etc. determine the emergent functionalities and properties of CEB. Tackling this complexity through interdisciplinary fields (artificial intelligence, civil engineering) could improve our grasp of the system, and provide a more robust and efficient way to analyze the mechanical

* Corresponding author at: Univ. Bordeaux, CNRS, Bordeaux INP, I2M, UMR 5295, Talence F-33400, France
E-mail address: sarah.nassar@u-bordeaux.fr (S. Nassar).

https://doi.org/10.1016/j.conbuildmat.2024.138135

performance of CEB. To this end, supervised machine learning (a subset of artificial intelligence) proposes a set of data-driven methods and algorithms for making predictions or inferences on new data from historical data. It involves training the machine-learning model using a known dataset with inputs A (e.g. the quantities of components) and outcomes B (e.g. compressive strength). The model identifies an optimized pathway between the points A and B. This approach has been adopted by several researchers in the construction and materials fields using several machine learning models adapted to their database. In fact, many models were developed by artificial intelligence (AI) and the choice depends on the type of problem and its complexity as demonstrated by Mohtasham Moein et al. [12] who conducted a review of various models for predicting the mechanical performance of concrete. Tran [13] predicted the compressive strength of stabilized sediments using artificial neural networks (ANN) and Moghrabi et al. [14] assessed the mechanical performance of sediments by using the regression model. As for Ozsagir et al. [15], they used seven different algorithms to predict the soil liquefaction. In their case, the decision tree outperformed the other models used such as ANN, support vector machines (SVM), logistic regression and random forest (RF). The comparison of several models could help select the best performing one. Hilloulin and Tran [16] also employed this approach, comparing four machine learning algorithms to estimate the autogenous shrinkage of concrete. They emphasized the impact of the hyperparameters assigned to the models, which therefore need to be optimized for a better performance.

Machine learning algorithms may be often complex and the need for transparency and interpretability in these models becomes paramount. Explainability in machine learning refers to the ability to understand and interpret the decisions made by a model. Unlike sensitivity analysis that evaluates the robustness and reliability of models by examining the effects of input variations on outputs, explainability focuses on explaining predictions in machine learning models. Explainability thus allows stakeholders, including data scientists, domain experts, and end-users, to comprehend why a particular decision or prediction was made by the model. It provides insights into the features or factors that influenced the model's output, thus building trust and facilitating informed decision-making. There are various techniques and approaches to achieve explainability in machine learning models. These include feature importance analysis [17], which identifies the most influential features in the model's decision-making process, and model-agnostic methods which generate local explanations for individual predictions [18].

In the frame of this study, the objective is to optimize the CEB mix design by understanding the relations between the fabrication parameters and predicting the mechanical performance of the block, particularly its dry compressive strength. Despite the advancements of research, there remains a gap in comprehending this interplay between the parameters and estimating the resulting strength.

The primary objective of the paper is to explore supervised machine learning techniques to predict the dry compressive strength of CEB from a dataset collected from the literature. The machine learning model will be a useful tool to combine all the affecting parameters and help in decision making. Classification models have been considered in this study and the predicted class is an indicator of the suitability of the CEB for the intended use. The secondary objective is to evaluate the extent to which the predictive model can assess the compressive strength of earth brick mixes made of sediments from Arcachon Bay, in the southwest of France, that were also mechanically tested in the laboratory. While the predictive performance of the model is crucial, the ultimate objective is to provide explanations for the predictions made by the best-performing machine learning model. This entails understanding and identifying the most important and influential features of the model that contribute to accurate predictions. The goal is to be able to assist and optimize the mixture without the necessity of conducting a large-scale experimental campaign to achieve the desired CEB properties.

Compared to existing research, this work aims to provide a generic approach that covers numerous mix designs, such as stabilized and unstabilized CEB. It doesn't focus on an isolated aspect nor merely test a single algorithm for validation; rather it aims to comprehend the relations between the parameters and test the reliability through laboratory experimentation. The choice of the parameters considers the different aspects of the fabrication process unlike some previous works that focus commonly on the raw materials as features, or even on other testing methods (e.g. nondestructive testing) that would require additional experiments. In this study, importance is given to the understanding of the results which may be briefly or not discussed in previous works.

2. Materials and methods

2.1. Dataset

2.1.1. Feature selection and data collection

When selecting a construction material, various characteristics are examined to guarantee the safety and comfort of the occupants. For instance, the material needs to be assessed on its physical, mechanical, thermal performance and durability. Fig. 1 shows some of the main properties studied for CEB and their prevalence in research. A literature review identified scientific articles investigating the experimental characterization of CEB by using the following query in databases, ScienceDirect and Web of Science: *Title-abs-key: Compressed earth blocks*. The performed experiments were noted and it emerges that the compressive strength is the most studied property, with 80 % of the papers focusing on it.

This was observed as well by Turco et al. [19] in their review that also included other characterizations of the material. It seems relevant since the compressive strength is an indicator of the quality of CEB (Morel et al., 2007) that are also intended to be used for masonry walls where mechanical stability is needed. Therefore, as a first approach, focus will be given to the mechanical performance of the block, particularly the dry compressive strength at 28 days.

This strength is affected not only by the composition of the CEB but also by the manufacturing process and storage conditions. The software VOS Viewer was used to analyze the keywords occurrence in all articles related to compressed earth blocks. The compressive strength was indeed a recurrent term, as the keywords visualized with bigger nodes are more frequently found. Fig. 2 shows the resulting network with a focus on the keywords connected strictly to the compressive strength which included some mix parameters such as the soil classification, the microstructure, the stabilization, and stabilizers such as cement. Indeed, Houben et Boubekeur [10] provide recommendations on the particle size distribution and plasticity of the soil to achieve a satisfactory mechanical performance. Moreover, increasing the compaction pressure and stabilizer content can improve strength [20,21,22]. In addition, the amount of water is optimized, by Proctor tests [23] in case of dynamic compaction or through static compaction [1,24], to increase the block's density and therefore its performance [23].

The relationship between these parameters is complex, and it is difficult to predict their final impact on the compressive strength. Therefore, experimental results from the literature were collected to build a database.

In particular, the focus is on the manufacturing conditions and their impact on the CEB compressive strength (denoted Rc). The selected parameters include the soil properties, the mixture, and the manufacturing conditions, i.e., sand, silt, and clay percentages in the soil, plasticity index (denoted PI) and liquid limit (denoted LL) of the fine fraction, the water content of the mixture, the type and percentage of stabilizer, and the compaction pressure (denoted CP). The hydraulic binders were considered in this study. The most commonly used are cement and lime. Other parameters, such as organic matter content, methylene blue value, or density, which also have an impact on the performance, can be added, as well as the effect of other types of

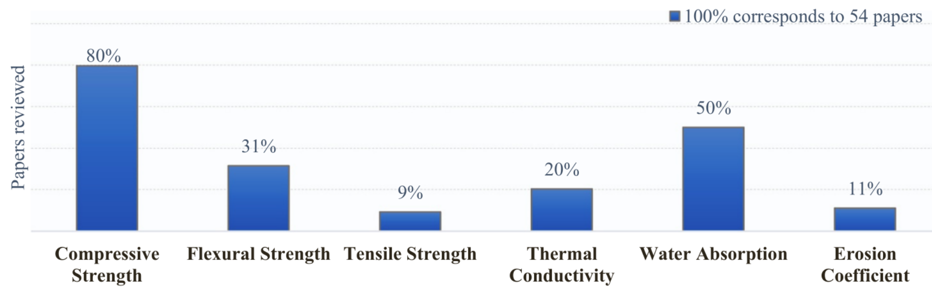

Fig. 1. Commonly studied CEB properties in selected papers.

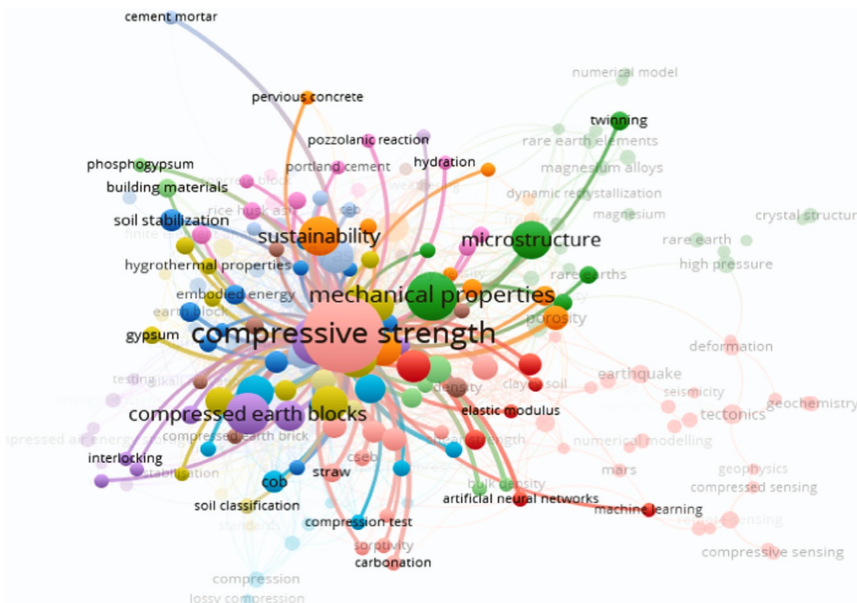

Fig. 2. Relationship of keywords in CEB related documents of Scopus and ScienceDirect.

additives and binders. However, these were not considered in this study due to data limitations. The collected database comprises a total of 276 data points with the above specified nine parameters in addition to the resulting compressive strength of the block. Table 1 shows an extract of the dataset.

2.1.2. Data discretization

The compressive strength was discretized into representative intervals using expert judgement. Classification of values was chosen as it was deemed more appropriate for the prediction of the compressive strength, which can exhibit slight variability for identical mixes. Thus, the intervals are more judicious to account for this variability and the potential standard deviation. Furthermore, it is common to label blocks with a strength class since the standard XP P13–901 [29] (relative to CEB) also categorized CEB into six classes based on compressive strength intervals.

The chosen intervals in this study are reported in Table 2. The values, varying between 0 and 8 MPa, were divided into four intervals enabling the evolution of the strength to be tracked as accurately as possible. According to Van Damm & Houben, the average compressive strength of unstabilized CEB vary between 1 and 7 MPa [30]. Note that, CEB belonging to the interval [0;2[may not present a sufficient strength according to the minimum strength requirements of some normative documents [31].

Table 2
Discretized Intervals of compressive strength.

Rc (MPa)	[0;2[	[2;4[	[4;6[	[6;8]

2.1.3. Data pre-processing

Data pre-processing is a step that precedes the feeding of the created model with the collected dataset. This step is crucial as it enhances the quality of the data and therefore improves the performance of the model itself [32]. The data was reviewed to remove entries with missing values associated with the parameters, in addition to outliers by using the

Table 1
5 out of 276 mixes of the dataset.

Authors	Sand (%)	Silt (%)	Clay (%)	PI (%)	LL (%)	Water (%)	Cement (%)	Lime (%)	CP (MPa)	Rc (MPa)
[25]	48.0	3.8	26.2	10.1	27	14.2	0	3	4.4	3.25
[26]	48.4	10.1	5.4	16.7	50.5	27.9	0	0	3.5	1.36
[27]	25.0	50.0	25.0	16	38	20.0	0	0	3.5	2
[1]	40.0	20.0	25.0	8	81	47.0	0	0	3	1.00
[28]	53.3	28.2	17.0	12	44.3	18.2	4	0	4	4.24

interpercentile range method. It was also normalized as explained below.

In fact, data standardization has been performed. It consists of scaling the data of the different parameters in order to ensure that features are comparable and have similar scales. The importance of this step has been studied by several authors and for different algorithms [32,33]. In this paper, each data point is subtracted by the mean of its respective variable, and then divided by the standard deviation of that variable.

This standardization process may improve the performance and convergence of many machine learning algorithms.

2.2. Classification models

Eight different supervised machine learning models were tested for the prediction of the compressive strength of CEB.

2.2.1. K-nearest neighbors (KNN)

KNN is a distance-based classifier well suited for various applications, especially when the dataset is small. It was used by Beskopylny et al. [34] for the prediction of concrete strength and for material classification by Fernando and Marshall [35]. It assumes greater similarity between two points as their distance decreases, indicating that similar data points typically belong to the same class. KNN differs from other classifiers in that during the "fit" step, KNN stores all the training data and corresponding labels without calculating distances. Instead, all computations are performed during the "predict" phase. For each point where a class prediction is conducted, KNN identifies its K closest neighbors in the training set based on a chosen distance metric, such as Euclidean distance. Then, KNN assigns to this data point the class that has the most frequent occurrence among these K neighbors. The selection of K may affect the model's performance, and there isn't a specific method for determining its optimal value. However, a general recommendation is to avoid extremely low values that could lead to inaccuracies due to noise and outliers. Instead, consider selecting an odd number for K to break ties when needed. It's also widespread practice to try different values within a reasonable range and evaluate their effect on the model's accuracy through cross-validation techniques.

2.2.2. Random forest (RF)

Random Forest [36] is an extension of decision trees. It consists in (1) building many decision trees from many smaller datasets randomly sampled from the original training dataset; (2) getting predictions from each tree and (3) obtaining the most frequent predicted class winning the vote. The random forest (RF) algorithm relies on various hyperparameters that require user specification. These include for instance the number of observations randomly selected for each tree, the number of variables chosen randomly for each split, the specific splitting rule employed, the minimum number of samples required within a node, and the total number of trees in the forest, *etc*. According to the literature [37], the most important influence parameter concerns the number of variables considered as candidate splitting variables at each split. A high number of trees may generally improve performance avoiding overfitting [36]. Although the number of used data for building each tree and the minimum number of data associated with leaf nodes have less influence, they are worth tuning in many cases. RF is very user-friendly because it is mainly controlled by two parameters namely the number of variables in the random subset at each node and the number of trees in the forest. Anysz et al. [17] used RF in the prediction of the compressive strength of rammed earth. RF also showed good accuracy in predicting compressive strength of concrete according to Shaqadan [38].

2.2.3. Support vector machine (SVM)

The SVM algorithm [39,40] tries to find a separating optimal hyperplane dividing the input space between classes that maximizes the distance between this hyperplane and the closest data from each class. SVM can handle non-linear datasets by changing its dimensionality to higher dimensionality by using a kernel function as radial basis function, quadratic, sigmoid kernel and Polynomial kernel. In the case of non-linearly separable cases, the notion of soft margin is introduced through a degree of tolerance which consists in finding the optimal decision boundary by tolerating misclassified data. This degree is known as the regularization parameter controlling the trade-off between maximizing the distance between both classes and minimizing training errors. That means that if the regularization parameter is low, the SVM algorithm focuses more on finding a simple decision boundary that separates classes reasonably well, even if it means making some mistakes on the training data. On the other hand, if the regularization parameter is high the SVM algorithm focus on each individual training data, being able to lead a more complex decision boundary that may lead to over fitting. The performance of SVM algorithm is mainly determined by these two parameters. SVM was used by Tanyildizi [41] to estimate the compressive and flexural strength of carbon fiber-reinforced lightweight concrete. Omran et al. [42] also used SVM, among other algorithms, to predict the compressive strength of an alternative type of concrete.

2.2.4. Artificial neural network (ANN)

There are mainly three types of Artificial Neural Networks: feedforward, feedback, and graph. A specific type of feed-forward neural network is considered in this paper known as Multi-layer Perceptron. Its structure includes three types of layers: (1) the input layer containing neurons which receive the training data; (2) the hidden layer composed of one or more sub-layers containing neurons performing mathematical transformations on the data and (3) the output layer given the results of the classification. Data flows in the forward direction from input layers, through hidden layer and to the output layer. Each neuron in one layer is directly connected to the neurons of the previous layer through an activation function. This function computes a weighted sum of neurons from the previous layer, adds a bias, and applies a non-linear function to generate the output signal. Training the MLP consists in finding the set of weights and biases for the network that minimizes a loss function using optimization solvers. MLP training is mainly controlled by the choice of (1) the number of hidden layers, (2) the number of neurons in each hidden layer, (3) the activation functions (e.g. rectified linear, sigmoid, hyperbolic, …) and (4) the optimizer algorithms (e.g. stochastic gradient descent, swarm particle optimization, evolutionary algorithms, …) [43,44,45]. For instance, a high number of hidden layers and neurons per layer may lead to a risk of overfitting. ANN model was widely used in the literature, for example for a soil classification [46] or the prediction of the compressive strength of concrete [47,48] and others applications.

2.2.5. Gaussian naïve Bayes (GNB)

The gaussian naïve bayes model is a statistical approach that is based on the Bayes' theorem and the "naïve" assumption of conditional independence [49]. Naïve Bayes assumption assumes that all features are conditionally independent given the class label. GNB model assumes that each feature within each class follows a Gaussian distribution. Unlike previous models, GNB model doesn't involve iteratively adjusting parameters to minimize an error function. Training a GNB model consists in calculating (1) the probability distribution associated with class labels, (2) the means and the standard deviations of Gaussian distributions associated with features within each class from the dataset. By using the Bayes' theorem and the naive independence assumption, GNB model estimates the most likelihood class label given the features of the new data point. Among other applications, it was used for damage detection in engineering materials [50] or even to predict the compressive strength of ultra-high-performance concrete [51].

2.2.6. Boosting algorithms (Adaptive boosting and extreme gradient boosting)

Boosting algorithms work by combining multiple weak learning models to create a strong learning model that makes more accurate predictions. Adaptative boosting, Gradient Boosting and XGBoost algorithms [52,53] are the most popular boosting algorithms. Adaptative boosting algorithms construct a sequence of weak classifiers. Each classifier attempts to correct the errors of the previous classifiers by focusing on the misclassified examples and increasing their weights in the training set. This allows the next classifiers to focus more on these misclassified points. Each weak classifier is assigned a weight that depends on its performance on the training dataset. Finally, these weak classifiers are combined in a weighted manner to form a strong model. By sequentially combining weak classifiers in this way, AdaBoost forms a strong model capable of making accurate predictions on new examples, with particular attention to difficult examples. Gradient boosting relies on the same principle using only decision trees as base learners, but it focuses on minimizing a loss function of the previous ensemble. Each new learner is trained to predict the residuals (errors) of the ensemble, effectively correcting the mistakes made by the previous models. XGBoost is an extension of Gradient boosting. These boosting algorithms were previously used in the construction field to predict, for example, the autogenous shrinkage of concrete [16] or the compressive strength of self-compacting concrete [54].

2.2.7. Voting classifier (VC)

A voting classifier combines the predictions of multiple individual learning models to make a final prediction. It is mainly controlled by the number and the choice of learning models and the voting strategy. The voting strategy may either consist of choosing the class label that receives the most votes (hard voting) among the learning models or weighting the votes of different models based on their performance (soft voting). For example, it was used by Son et al. [55] to detect construction materials and showed good accuracy compared to single classifiers.

As depicted, each model has its own approach to assess the data and make the predictions. However, they might not be performant on all types of data [56] and they have their advantages and disadvantages [57]. For example, the KNN approach is non-parametric; it is based on identifying the closest neighbors without an assumption on the data distribution. Hence, it will dominate other algorithms when the data is nonlinear [58]. Furthermore, GNB assumes that the features are independent and can perform poorly when the data is highly correlated. As for the SVM, it can be slow when dealing with complex data [57]. Therefore, the performance of an algorithm is data dependent, and the eight proposed models are tested in this study to find the most accurate and efficient one for the collected dataset.

2.3. Performance metrics

There are many global and per-class metrics to assess the performance of multiclass classification models [59,60]. Table 3 summarizes the different metrics. Accuracy is frequently assessed, and it is the ratio of the correct predictions to the total number of predictions. Precision concentrates on positive predictions, serving as an indicator to minimize false positive estimates. Recall emphasizes false negatives that were inaccurately classified as belonging to another class. Furthermore, the F1-score is the harmonic mean of the precision and the recall. F1-score allows quantifying the model's ability to correctly classify observations while minimizing the errors of incorrectly predicting that an observation (1) belongs to a given class when, in reality, it does not belong to that class and (2) does not belong to a given class when, in reality, it does belong to that class. A higher F1-score signifies better overall model performance [60].

Table 3
Assessment metrics.

Assessment Metrics	Meaning
Global Accuracy	Proportion of correctly classified data points
Macro F1-score	Averages of the F1-scores calculated for each class
Micro F1-score	A weighted average of the F1 scores calculated for each class according to the number of data points belonging to each class
Precision per class	Proportion of correctly classified data points per class
Recall per class	Proportion of correctly identified data point by the model per class
F1-score per class	The harmonic mean of precision and recall per class

3. Results and discussion

The Scikit-learn [61]. environment was used to generate classification models. The hyper-parameters were tuned by using the GridSearch method [61], in which all possible combinations of given discrete parameter spaces are evaluated. Table 4 summarizes the main hyper-parameters of each training model. The performance of the models was assessed through a "leave-one-out" cross-validation [62]. which is beneficial and efficient for small dataset. It consists of removing one data point at a time from the database to test the model prediction for the excluded point.

3.1. Validation

Table 5 shows the results of the assessment metrics of the eight models. These metrics provide insights of various aspects of the performance, and some might be more relevant in specific scenarios. Accuracy is commonly used but is not very representative when the data is imbalanced. As for the F1-score, it balances the precision and recall values that focus on false positive and false negative, respectively. The voting classifier achieved the best accuracy and F1-score and relatively high precision and recall. VC thus provided the highest performance among model and predicted compressive strength with a classification accuracy of 78 %.

Table 6 displays the confusion matrix proposing a visualization tool displaying how often the model correctly classified and incorrectly classified data points for each class. For example, out of 45 predicted values of the class [4,6[, 28 actually come from the class [4,6[, 14 come from the class [2,4[and 3 from the class [6,8[, which means that the model is 62 % reliable when it predicts the class [4,6[(see the grey column in Table 6).

3.2. Feature importance analysis and local explanations

Both feature importance analysis and model-agnostic methods for local explanations are complementary useful approaches to better understand machine learning models. While feature importance analysis focus on the understanding of the global behavior of models helping to identify the most influential variables; model-agnostic methods focus on

Table 4
Optimized hyper-parameters by the GridSearch method.

Models	Optimized Parameters
KNN	# neighbors = 6
RF	# trees = 56
	# splitting variables = 3
SVM	Regularization parameter = 87
	Kernel = Radial basis function
MLP	# hidden layers = 3
	#number of neurons in each hidden layer =(15,10,5)
	activation functions = rectified linear
	optimizer algorithms = stochastic gradient
GNB	e.g. (Sand|Rc=[0,2[) = Normal(32.42,23.86)
Adaboost	Classifier = Decision tree
	# boosting rounds = 50
XGBoost	# boosting rounds = 35

Table 5
Results of performance measures.

Models	Accuracy	Average Precision	Average Recall	Macro F1-score
KNN	0.73	0.76	0.59	0.64
RF	0.78	0.73	0.69	0.71
SVM	0.71	0.64	0.62	0.63
MLP	0.72	0.58	0.53	0.54
GNB	0.55	0.42	0.54	0.37
Adaboost	0.76	0.69	0.71	0.7
XGBoost	0.76	0.68	0.65	0.66
VC	0.78	0.75	0.7	0.72

individual predictions offering justifications for the model's decisions on specific instances. SHapley Additive exPlanation (SHAP) values may be both used in these two contexts regardless of the model type to explain how much each variable contributes to a prediction and thus to verify the consistency with the expert knowledge of the domain [63]. The SHapley Additive exPlanations (SHAP) Python library ([64]; version 0.44.1) was used to interpret the voting classifier.

Fig. 3 displays the variables ordered by how much they influenced the model's prediction according to the average of the absolute SHAP value of each variable. They are divided into categories according to the colors (red: raw material-dependent; blue: processing variables; green: treatment variables). Globally, it can be seen that the three categories are influential and are represented in the top three. This proves their relevant influence and that the compressive strength indeed depends on various decisive parameters, not only on the raw material-dependent variables, as discussed previously.

The rate of silt is the most important feature, followed by the compaction pressure, cement, sand, *etc*. Indeed, the soil texture is an important factor. The sand and silt are among the most decisive parameters and serve as the structural framework of the soil. The silt can provide cohesion and sand adds strength to the CEB. Furthermore, the cementation of the soil is more efficient in the presence of sand. As for the lime, it has better action on clayey soils with high water content [10].

However, the soil used for the fabrication of CEB tends to have low clay fractions. In this analysis the percentage of clay is not playing an important role. It is definitely decisive and affects the soil texture as stated before, in addition to being an active phase. Other related parameters such as the blue methylene value and the activity need to be specified when dealing with clay fractions. Moreover, the compaction pressure turned out to be an important factor. In fact, it is considered as a mechanical stabilization as it reduces porosity and results in more resistant CEB [10].

Fig. 4 displays the SHAP values for each prediction and variable. While red point means higher value of a variable, blue point means lower value. The magnitude of the SHAP value represents the strength of the effect the variable has on the prediction. A value closer to zero indicates a weaker influence on the prediction, regardless of its positive or negative sign. Higher distribution of SHAP values (i.e concentrated dots) implies a more consistent impact of the variable on the predictions of model. For instance, low percentage of Cement generates negative SHAP values on the compressive strength of CEB meaning that a lower percentage of Cement leads to a lower predicted compressive strength. Fig. 4 shows that a higher (*resp*. a lower) compaction pressure leads to a higher (*resp*. a lower) predicted compressive strength. This is expected since a higher compression strength leads to reducing the porosity and increasing the density of the blocks which results in higher compressive strength [22].

The concentration of the CP's negative SHAP values contrary to the CP's scattered positive SHAP values suggests that the model gives more robust predictions for lower CP than higher CP. The same tendencies are observed for the cement. Indeed, increasing the cement ratio helps achieve better compressive strength results. However, the predictions

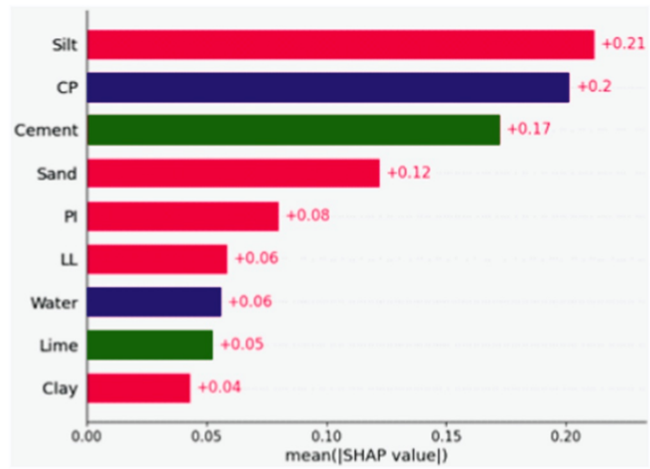

Fig. 3. Global importance of variables averaging the absolute value of the Shap values (red: raw material-dependent; blue: processing variables; green: treatment variables).

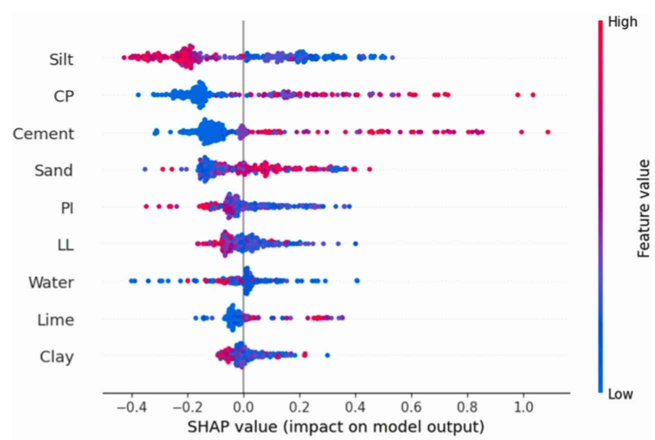

Fig. 4. SHAP values for each prediction in x axis associated with variables importance along y axis.

Table 6
Confusion Matrix of VC model.

		Predicted compressive strength (MPa)			
		[0 ;2[	[2 ;4[	[4 ;6[	[6 ;8]
Raw compressive strength (MPa)	[0 ;2[	**136**	10	0	0
	[2 ;4[	11	**45**	14	0
	[4 ;6[	2	18	**28**	1
	[6 ;8]	0	1	3	**7**

were more accurate for lower cement addition as the concentrated blue points correspond to negative SHAP values.

The concentration of PL's, LL's, Water's, Lime's and Clay's SHAP values around zero indicates that these variables have a weak influence on the model's prediction.

Fig. 5 displays the decision plots for two samples of the dataset. The x-axis represents the model's output (i.e. predicted compressive strength classes). Each line represents the cumulative SHAP value (i.e. the cumulative impact of variables) for a single data point in the dataset. Starting from the average prediction of the model (which is not relevant in this case as we're only focusing on classes) without considering any variable, each line moves on the left (*resp.* on the right) when variable negatively (resp. positively) influences the compressive strength prediction. The end point of the line represents the final prediction made by the model. Each value in parentheses represents the values taken by the variables for the sample.

For instance, Fig. 5, on the left, shows that the model determined that a silt rate of 55 % (corresponding to a high concentration), a compaction pressure of 0.6 MPa (corresponding to a low pressure) and a sand rate of 28.25 % (corresponding to a low rate) negatively influence the prediction of compressive strength. This sample confirms the previous global result where high silt concentration and low compaction pressure lead to low compressive strength, i.e. the class [0,2[. Conversely, Fig. 5 (on the right) shows that the model determined a water rate of 10 % (corresponding to a low concentration), a compaction pressure of 8.3 MPa (corresponding to a high pressure) and a sand rate of 60 % (corresponding to a high rate) positively influence the prediction of compressive strength, i.e. the class [6,8[. It can be noted that the silt rate (26 %) has a weak influence, which contradicts the previous global results. This can be explained by the fact that for low silt rate, the model is not very influenced, as suggested by the blue dot scatter plot in Fig. 4. These results suggest that the predictive model will be more robust for high silt rate and low compaction pressure which may be explained by the fact that the class [0,2[is overrepresented.

3.3. *Reliability and versatility of the voting classifier*

The aim of this section is to assess the robustness and the versatility of the previous voting classifier faced with new data of CEB manufactured and tested at the laboratory. The proposed mix design is an example of many other possible mixes.

The CEB are not stabilized and are made from dredged sediments collected from the Arcachon Bay in southwest France. The sediments are characterized [65] by geotechnical tests and the following properties were obtained: Sand (51 %), Silt (23 %), Clay (23 %), LL (36 %), PI (10 %). The water content for the mixes is 18 %, determined by a static compaction test [1,66]; it corresponds to the percentage achieving the maximum CEB dry density in static compaction conditions.

The blocks were manufactured by using a press and applying a static compaction pressure varying between 2 and 6 MPa. Hence, the mix design of the CEB is the same and only the compaction pressure was varied. A summary of the input parameters for the mixes is presented in Table 7.

Three specimens were tested for each CP as shown in Table 8 along with the results of the compressive strength test. As discussed in Section 2.1.2, a slight difference in the results is observed for the same mixtures and it is actually more likely to have a variability when dealing with sediments.

The VC classifier model is used to predict Rc for the mixes above. Table 9 displays the prediction performance of the model when faced with new mixes described in Table 7 without integrating them into the training dataset versus a "leave-one-out" cross-validation that integrates these experimental trials into the training data set. The total accuracy, whether or not experimental trials are integrated into the training dataset, remains unchanged at about 78 %. Table 9 shows that the class [0;2[is accurately predicted, contrary to the mixes leading to the Rc class [2;4[which are not predicted correctly.

In fact, the only varying parameter in the mixes is the compaction pressure, which is being increased. As discussed in Section 3.2, the scattered positive SHAP values of CP in Fig. 4 may indicate instability or inconsistency in the model's predictions, potentially reducing trust in the model's reliability and performance. That also means that the configurations of the experimental trail mixes are poorly representative in the training dataset. To improve this shortcoming, a "leave-one-out" cross-validation is performed by integrating experimental trials into the training data set. Table 9 shows the improvement in the prediction of the compressive strength associated with the experimental trials.

The VC classifier, developed in this work seems to be promising to assess the compressive strength of CEB made from any collected soil. The data needs to be representative of all the classes of the different parameters to ensure a greater accuracy as seen in Table 9 for the compaction pressure. This classifier appears to be an efficient tool and data could be continuously enriched by the experimental results to cover a wider range of values.

The AI model can provide a prediction of the compressive strength class after precising all the input parameters specifying the soil properties, mix proportions and compaction pressure. Therefore, these predictions are a tool to guide the experiments and select the most promising mixes. Paul et al. [67] reported Rc requirements according to international standards and the most severe value is around 2 MPa.

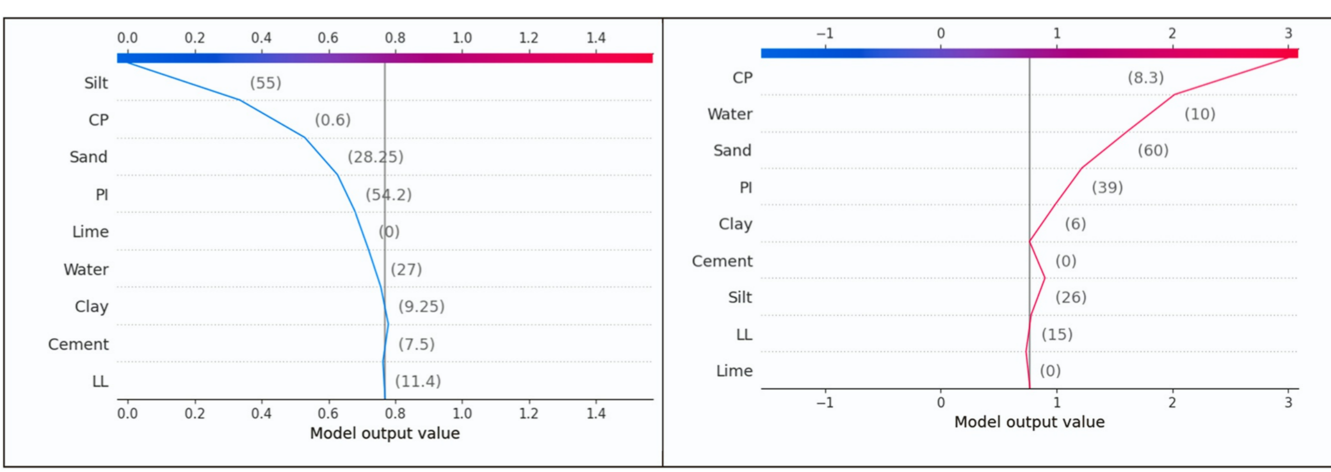

Fig. 5. SHAP decision plot: Explanation of the impact of variables for two samples of the dataset leading to the predicted compressive strength class [0,2[on the left (reps. [6,8] on the right) where {0, 1, 2, 3} along x axis corresponds to the classes {[0,2[, [2,4[, [4,6[, [6,8]}.

Table 7
Input parameters of the mixes of CEB manufactured at the laboratory.

Sand	Silt	Clay	LL	PI	Water	Cement	Lime	CP
51 %	23 %	23 %	36 %	10 %	18 %	0 %	0 %	2; 4 and 6 MPa

Table 8
Results of the compressive strength test for the manufactured samples.

CP	Sample	Rc (MPa)	Raw Rc Class
2 MPa	1	1.68	[0;2[
	2	1.72	[0;2[
	3	1.77	[0;2[
4 MPa	1	1.57	[0;2[
	2	1.56	[0;2[
	3	2.31	[2;4[
6 MPa	1	3.25	[2;4[
	2	3.51	[2;4[
	3	2.21	[2;4[

Table 9
Confusion Matrix of VC classifier for experimental data.

		Predicted compressive strength (MPa) with only literature training data set		Predicted compressive strength (MPa) with the training dataset enriched with experimental data	
		[0;2[	[2;4[	[0;2[	[2;4[
Raw compressive strength (MPa)	[0;2[	**5**	0	5	0
	[2;4[	4	**0**	1	3

Hence, the parameters can be fine-turned to reach at least an acceptable compressive strength. Only the preselected CEB will therefore be manufactured and tested.

4. Conclusion

CEB are masonry units made from a mixture of soil and water, with or without stabilizers. They are unfired and their fabrication consists of applying compaction pressure to the mix. These units are intended for wall construction, so it is crucial to assess their compressive strength. However, the variability of the soil can be challenging, as it implies the adaptation of the rest of the parameters accordingly, to achieve satisfactory results. In this article, a machine learning approach is proposed to predict the compressive strength of CEB. Eight classification models are used, and the Voting Classifier turned out to be the most performant one for the database collected from the literature. The input parameters are the soil's granularity and plasticity, the water content, the percentage of hydraulic binders and the compaction pressure. A feature importance study was conducted, revealing that the granularity and the compaction pressure are the most influential factors. This VC model can be used to predict the compressive strength of any CEB mix and serves as a tool to help in decision-making. Furthermore, CEB were manufactured in the laboratory and their compressive strength was predicted by the VC algorithm. These predictions were compared to the results of the compression destructive test. Some limitations are observed for high compaction pressure that are not quite present in the database. Therefore, the latter will be extended to cover more mixes. Also, more features shall be added to improve the accuracy such as methylene blue value, clay activity, organic matter content, and other stabilizers such as activators for geopolymerization.

CRediT authorship contribution statement

Sarah Nassar: Writing – original; **Cédric Baudrit:** Writing – review & editing, Validation, Conceptualization; **Jacqueline Saliba:**Writing – review & editing; **Nadia Saiyouri:** Writing – review & editing, Validation, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

References

[1] M. Abbou, A. Semcha, F. Kazi-Aoual, Stabilization of compressed earth block clayey materials from Adrar (Algeria) by lime and crushed sand, J. Build. Mater. Struct. (2020), https://doi.org/10.34118/jbms.v7i1.137.
[2] A. Fabbri, et al., An overview of the remaining challenges of the RILEM TC 274-TCE, testing and characterisation of earth-based building materials and elements, RILEM Tech. Lett. (2022), https://doi.org/10.21809/rilemtechlett.2021.149.
[3] A.W. Bruno, B. Scott, Y. D'Offay-Mancienne, C. Perlot, Recyclability, durability and water vapour adsorption of unstabilised and stabilised compressed earth bricks, Mater. Struct. (2020), https://doi.org/10.1617/s11527-020-01585-7.
[4] F. Pacheco-Torgal, S. Jalali, Earth construction: lessons from the past for future eco-efficient construction (avr), Constr. Build. Mater. vol. 29 (2012) 512–519, https://doi.org/10.1016/j.conbuildmat.2011.10.054.
[5] T. Morton, F. Stevenson, B. Taylor, N.C. Smith, « Low Cost Earth Brick Construction: 2 Kirk Park, Dalguise - Monitoring and Evaluation ». Consulté le: 18 novembre 2022. [En ligne]. Disponible sur: ⟨https://www.abebooks.com/9780955058004/Low-Cost-Earth-Brick-Construction-0955058007/plp⟩.
[6] S. Nassar, J. Saliba, N. Saiyouri, Investigation of the possible valorization of dredged sediments in compressed earth blocks, *Mater. Today Proc.*, juill (2023), https://doi.org/10.1016/j.matpr.2023.06.164.
[7] L.W. Ean, M.A. Malek, B.S. Mohammed, C.-W. Tang, P.X.H. Bong, A review on characterization of sediments for green bricks production, Int. J. Eng. Technol. vol. 7 (n° 4.35) (nov. 2018) 41, https://doi.org/10.14419/ijet.v7i4.35.22319.
[8] F. Belayali, W. Maherzi, M. Benzerzour, N.-E. Abriak, A. Senouci, Compressed earth blocks using sediments and alkali-activated byproducts (mars), Sustainability vol. 14 (n° 6) (2022) 3158, https://doi.org/10.3390/su14063158.
[9] B. Serbah, N. Abou-Bekr, S. Bouchemella, J. Eid, S. Taibi, « Dredged sediments valorisation in compressed earth blocks: Suction and water content effect on their mechanical properties », 2018, doi: 10.1016/j.conbuildmat.2017.10.043.
[10] H. Houben S. Boubekeur, Compressed earth blocks: standards (Guide Séries Technologies No. 11). 1998.
[11] H.B. Nagaraj, A. Rajesh, M.V. Sravan, Influence of soil gradation, proportion and combination of admixtures on the properties and durability of CSEBs, Constr. Build. Mater. (2016), https://doi.org/10.1016/j.conbuildmat.2016.02.023.
[12] M. Mohtasham Moein, et al., Predictive models for concrete properties using machine learning and deep learning approaches: a review (oct), J. Build. Eng. vol. 63 (2022) 105444, https://doi.org/10.1016/j.jobe.2022.105444.
[13] V.Q. Tran, Compressive strength prediction of stabilized dredged sediments using artificial neural network, 2021, Adv. Civ. Eng. vol (1-8, mars) (2021), https://doi.org/10.1155/2021/6656084.
[14] I. Moghrabi, H. Ranaivomanana, F. Bendahmane, O. Amiri, D. Levacher, Modelling the mechanical strength development of treated fine sediments: a statistical approach, Environ. Technol. (2019), https://doi.org/10.1080/09593330.2018.1432697.
[15] M. Ozsagir, C. Erden, E. Bol, S. Sert, A. Özocak, Machine learning approaches for prediction of fine-grained soils liquefaction, déc, Comput. Geotech. vol. 152 (2022) 105014, https://doi.org/10.1016/j.compgeo.2022.105014.
[16] B. Hilloulin, V.Q. Tran, Using machine learning techniques for predicting autogenous shrinkage of concrete incorporating superabsorbent polymers and supplementary cementitious materials (mai), J. Build. Eng. vol. 49 (2022) 104086, https://doi.org/10.1016/j.jobe.2022.104086.
[17] H. Anysz, W. Brzozowski, P. Kretowicz, Narloch, Feature importance of stabilised rammed earth components affecting the compressive strength calculated with

explainable artificial intelligence tools, Art. n° 10, janv, Materials vol. 13 (n° 10) (2020), https://doi.org/10.3390/ma13102317.
[18] Y. Zhao, et al., Assessing the influencing factors of soil susceptibility to wind erosion: A wind tunnel experiment with a machine learning and model-agnostic interpretation approach, août, CATENA vol. 215 (2022) 106324, https://doi.org/10.1016/j.catena.2022.106324.
[19] C. Turco, A.C. Paula Junior, E.R. Teixeira, R. Mateus, Optimisation of Compressed Earth Blocks (CEBs) using natural origin materials: a systematic literature review, Constr. Build. Mater. (2021), https://doi.org/10.1016/j.conbuildmat.2021.125140.
[20] S. Mkaouar, W. Maherzi, P. Pizette, H. Zaitan, M. Benzina, A comparative study of natural Tunisian clay types in the formulation of compacted earth blocks, déc, J. Afr. Earth Sci. vol. 160 (2019) 103620, https://doi.org/10.1016/j.jafrearsci.2019.103620.
[21] J.-C. Morel, A. Pkla, P. Walker, Compressive strength testing of compressed earth blocks, févr, Constr. Build. Mater. vol. 21 (n° 2) (2007) 303–309, https://doi.org/10.1016/j.conbuildmat.2005.08.021.
[22] B. Taallah, A. Guettala, S. Guettala, A. Kriker, Mechanical properties and hygroscopicity behavior of compressed earth block filled by date palm fibers (mai), Constr. Build. Mater. vol. 59 (2014) 161–168, https://doi.org/10.1016/j.conbuildmat.2014.02.058.
[23] Afnor, « NF P94-093 Soils: investigation and testing - Determination of the compaction reference values of a soil type - Standard proctor test - Modified proctor test ». [En ligne]. Disponible sur: ⟨https://viewerbdc.afnor.org/pdf/viewer/LGm05JFnXHA1?proxy=true⟩.
[24] P. Nshimiyimana, S. Omar Sore, C. Hema, O. Zoungrana, A. Messan, L. Courard, A discussion of "optimisation of compressed earth blocks (CEBs) using natural origin materials: a systematic literature review, Constr. Build. Mater. (2022), https://doi.org/10.1016/j.conbuildmat.2022.126887.
[25] M.M. Barbero-Barrera, F. Jové-Sandoval, S. González Iglesias, Assessment of the effect of natural hydraulic lime on the stabilisation of compressed earth blocks, Constr. Build. Mater. vol. 260 (nov. 2020) 119877, https://doi.org/10.1016/j.conbuildmat.2020.119877.
[26] S. Omar Sore, A. Messan, E. Prud'homme, G. Escadeillas, F. Tsobnang, Stabilization of compressed earth blocks (CEBs) by geopolymer binder based on local materials from Burkina Faso, Constr. Build. Mater. vol. 165 (2018) 333–345, https://doi.org/10.1016/j.conbuildmat.2018.01.051.
[27] P. Nshimiyimana, C. Hema, O. Zoungrana, A. Messan, L. Courard, Thermophysical and mechanical properties of compressed earth blocks containing fibres: by-product of okra plant and polymer waste », présenté à, WIT Trans. Built Environ. (2020) 149–161, https://doi.org/10.2495/ARC200121.
[28] J. Goutsaya, G.E. Ntamack, B. Kenmeugne, S. Charif d'Ouazzane, Mechanical characteristics of compressed earth blocks, compressed stabilized earth blocks and stabilized adobe bricks with cement in the town of Ngaoundere - Cameroon, déc, J. Build. Mater. Struct. vol. 8 (n° 2) (2021) 139–159, https://doi.org/10.34118/jbms.v8i2.1441.
[29] XP P13-901 Earth bricks and earth blocks for walls and partitions - Definitions - Specifications - Test methods - Delivery acceptance conditions ».
[30] H. Van Damme, H. Houben, Earth concrete. Stabilization revisited, déc, Rep. UNEP SBCI Work. GROUP LOW-CO2 ECO-Effic. Cem. -BASED Mater. vol. 114 (2018) 90–102, https://doi.org/10.1016/j.cemconres.2017.02.035.
[31] M.C. Jiménez Delgado, I.C. Guerrero, The selection of soils for unstabilised earth building: A normative review, févr, Constr. Build. Mater. vol. 21 (n° 2) (2007) 237–251, https://doi.org/10.1016/j.conbuildmat.2005.08.006.
[32] D. Singh, B. Singh, Investigating the impact of data normalization on classification performance, déc, Appl. Soft Comput. vol. 97 (2020) 105524, https://doi.org/10.1016/j.asoc.2019.105524.
[33] J. Sola, J. Sevilla, Importance of input data normalization for the application of neural networks to complex industrial problems (juin), IEEE Trans. Nucl. Sci. vol. 44 (n° 3) (1997) 1464–1468, https://doi.org/10.1109/23.589532.
[34] A.N. Beskopylny, et al., Concrete strength prediction using machine learning methods CatBoost, k-nearest neighbors, support vector regression, Appl. Sci. vol. 12 (n° 21) (oct. 2022) 10864, https://doi.org/10.3390/app122110864.
[35] H. Fernando, J. Marshall, What lies beneath: Material classification for autonomous excavators using proprioceptive force sensing and machine learning, », *Autom. Constr.* vol. 119 (nov. 2020) 103374, https://doi.org/10.1016/j.autcon.2020.103374.
[36] L. Breiman, Random Forests, Mach. Learn. vol. 45 (n° 1) (oct. 2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[37] P. Probst, M. Wright, A.-L. Boulesteix, Hyperparameters and Tuning Strategies for Random Forest, e1301, mai, WIREs Data Min. Knowl. Discov. vol. 9 (n° 3) (2019), https://doi.org/10.1002/widm.1301.
[38] A. Shaqadan, « Prediction of concrete mix strength using random forest model », vol. 11, p. 11024-11029, janv. 2016.
[39] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. vol. 20 (n° 3) (sept. 1995) 273–297, https://doi.org/10.1007/BF00994018.
[40] A. Roy, S. Chakraborty, Support vector machine in structural reliability analysis: A review (mai), Reliab. Eng. Syst. Saf. vol. 233 (2023) 109126, https://doi.org/10.1016/j.ress.2023.109126.
[41] H. Tanyildizi, Prediction of the strength properties of carbon fiber-reinforced lightweight concrete exposed to the high temperature using artificial neural network and support vector machine, *Adv. Civ. Eng.*, Vol. 2018 (n° 1) (2018) 5140610, https://doi.org/10.1155/2018/5140610.
[42] B.A. Omran, Q. Chen, R. Jin, Comparison of data mining techniques for predicting compressive strength of environmentally friendly concrete, J. Comput. Civ. Eng. vol. 30 (n° 6) (nov. 2016) 04016029, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000596.
[43] M. Rojas, A. Olivera, P. Vidal, A genetic operators-based Ant Lion Optimiser for training a medical multi-layer perceptron (janv), Appl. Soft Comput. vol. 151 (2024) 111192, https://doi.org/10.1016/j.asoc.2023.111192.
[44] A. Al Bataineh, S. Manacek, MLP-PSO hybrid algorithm for heart disease prediction (juill), J. Pers. Med. vol. 12 (n° 8) (2022) 1208, https://doi.org/10.3390/jpm12081208.
[45] S. Mirjalili, « Evolutionary multi-layer perceptron », in *Studies in Computational Intelligence*, in Studies in Computational Intelligence., Springer Verlag, 2019, p. 87-104. doi: 10.1007/978-3-319-93025-1_7.
[46] J.D. Sitton, Y. Zeinali, B.A. Story, Rapid soil classification using artificial neural networks for use in constructing compressed earth blocks (mai), Constr. Build. Mater. vol. 138 (2017) 214–221, https://doi.org/10.1016/j.conbuildmat.2017.02.006.
[47] M. Hadzima-Nyarko, E.K. Nyarko, N. Ademović, I. Miličević, T. Kalman Šipoš, Modelling the influence of waste rubber on compressive strength of concrete by artificial neural networks, Art. n° 4, janv, Materials vol. 12 (n° 4) (2019), https://doi.org/10.3390/ma12040561.
[48] M.A. Getahun, S.M. Shitote, Z.C. Abiero Gariy, Artificial neural network based modelling approach for strength prediction of concrete incorporating agricultural and construction wastes, Constr. Build. Mater. vol. 190 (nov. 2018) 517–525, https://doi.org/10.1016/j.conbuildmat.2018.09.097.
[49] K.P. Murphy, Naive Bayes classifiers, Univ. Br. Columbia *1860* (2006).
[50] O. Addin, S.M. Sapuan, E. Mahdi, M. Othman, A Naïve-Bayes classifier for damage detection in engineering materials (janv), Mater. Des. vol. 28 (n° 8) (2007) 2379–2386, https://doi.org/10.1016/j.matdes.2006.07.018.
[51] Z. Zhao, Predicting compressive strength of ultra-high-performance concrete using Naive Bayes regression in novel approaches (juin), Multiscale Multidiscip. Model. Exp. Des. (2024), https://doi.org/10.1007/s41939-024-00511-6.
[52] T. Chen C. Guestrin, « XGBoost: A Scalable Tree Boosting System », in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD '16. New York, NY, USA: Association for Computing Machinery, août 2016, p. 785-794. doi: 10.1145/2939672.2939785.
[53] J. Zhu, S. Rosset, H. Zou, T. Hastie, Multi-class AdaBoost, févr, Stat. Interface vol. 2 (2006), https://doi.org/10.4310/SII.2009.v2.n3.a8.
[54] J. de-Prado-Gil, C. Palencia, N. Silva-Monteiro, R. Martínez-García, To predict the compressive strength of self compacting concrete with recycled aggregates utilizing ensemble machine learning models, e01046, juin, Case Stud. Constr. Mater. vol. 16 (2022), https://doi.org/10.1016/j.cscm.2022.e01046.
[55] H. Son, C. Kim, N. Hwang, C. Kim, Y. Kang, Classification of major construction materials in construction environments using ensemble classifiers (janv), Adv. Eng. Inform. vol. 28 (n° 1) (2014) 1–10, https://doi.org/10.1016/j.aei.2013.10.001.
[56] A. Gajurel, B. Chittoori, P.S. Mukherjee, M. Sadegh, Machine learning methods to map stabilizer effectiveness based on common soil properties (mars), Transp. Geotech. vol. 27 (2021) 100506, https://doi.org/10.1016/j.trgeo.2020.100506.
[57] B.A. Hamed, O.A.S. Ibrahim, T. Abd El-Hafeez, Optimizing classification efficiency with machine learning techniques for pattern matching (juill), J. Big Data vol. 10 (n° 1) (2023) 124, https://doi.org/10.1186/s40537-023-00804-6.
[58] An Introduction to Statistical Learning. Consulté le: 30 avril 2024. [En ligne]. Disponible sur: ⟨https://link-springer-com.docelec.u-bordeaux.fr/book/10.1007/978-1-4614-7138-7⟩.
[59] I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis, N. Doulamis, Multiclass confusion matrix reduction method and its application on net promoter score classification problem, Art. n° 4, déc, Technologies vol. 9 (n° 4) (2021), https://doi.org/10.3390/technologies9040081.
[60] A. Tharwat, Classification assessment methods (janv), Appl. Comput. Inform. vol. 17 (n° 1) (2020) 168–192, https://doi.org/10.1016/j.aci.2018.08.003.
[61] F. Pedregosa, et al., Scikit-learn: machine learning in python, J. Mach. Learn. Res. vol. 12 (n° 85) (2011) 2825–2830.
[62] V. Ghate, S. Hemalatha C, A comprehensive comparison of machine learning approaches with hyper-parameter tuning for smartphone sensor-based human activity recognition, déc, Meas. Sens. vol. 30 (2023) 100925, https://doi.org/10.1016/j.measen.2023.100925.
[63] S.M. Lundberg S.-I. Lee, « A Unified Approach to Interpreting Model Predictions », in Advances in Neural Information Processing Systems, Curran Associates, Inc., 2017. Consulté le: 3 mai 2024. [En ligne]. Disponible sur: ⟨https://papers.nips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html⟩.
[64] shap/shap. (3 juin 2024). Jupyter Notebook. shap. [En ligne]. Disponible sur: ⟨https://github.com/shap/shap⟩.
[65] S. Nassar, J. Saliba, N. Saiyouri, « Investigation of the possible valorization of Arcachon Bay dredged sediments in earth constructions », présenté à CMSS23, 2023.
[66] M. Olivier, M.A. Mesbah, Le matériau terre: Essai de compactage statique pour la fabrication de briques de terre compressées (janv), Bull. Liaison Lab. P Ch vol. 146 (1986) 37–43 (janv).
[67] S. Paul, M.S. Islam, M.I. Hossain, Suitability of Vetiver straw fibers in improving the engineering characteristics of compressed earth blocks, déc, Constr. Build. Mater. vol. 409 (2023) 134224, https://doi.org/10.1016/j.conbuildmat.2023.134224.