Earth Science Informatics (2023) 16:1473–1487
https://doi.org/10.1007/s12145-023-00979-9

RESEARCH

Machine learning for prediction of the uniaxial compressive strength within carbonate rocks

Mohamed Abdelhedi[1] · Rateb Jabbar[2] · Ahmed Ben Said[2] · Noora Fetais[2] · Chedly Abbes[1]

Received: 27 December 2022 / Accepted: 18 February 2023 / Published online: 7 March 2023

Abstract
The Uniaxial Compressive Strength (UCS) is an essential parameter in various fields (e.g., civil engineering, geotechnical engineering, mechanical engineering, and material sciences). Indeed, the determination of UCS in carbonate rocks allows evaluation of its economic value. The relationship between UCS and numerous physical and mechanical parameters has been extensively investigated. However, these models lack accuracy, where as regional and small samples negatively impact these models' reliability. The novelty of this work is the use of state-of-the-art machine learning techniques to predict the Uniaxial Compressive Strength (UCS) of carbonate rocks using data collected from scientific studies conducted in 16 countries. The data reflect the rock properties including Ultrasonic Pulse Velocity, density and effective porosity. Machine learning models including Random Forest, Multi Layer Perceptron, Support Vector Regressor and Extreme Gradient Boosting (XGBoost) are trained and evaluated in terms of prediction performance. Furthermore, hyperparameter optimization is conducted to ensure maximum prediction performance. The results showed that XGBoost performed the best, with the lowest Mean Absolute Error (ranging from 17.22 to 18.79), the lowest Root Mean Square Error (ranging from 438.95 to 590.46), and coefficients of determination (R^2) ranging from 0.91 to 0.94. The aim of this study was to improve the accuracy and reliability of models for predicting the UCS of carbonate rocks.

Keywords Uniaxial compressive strength (UCS) · Carbonate rocks · Machine learning · Ultrasonic pulse velocity (UPV) · Effective porosity · Density

Introduction

Physical and mechanical characteristics of rocks (UCS, porosity, density, abrasion resistance, etc.) affect their areas of application. The economic interest in carbonate rocks is not only associated with the field of civil engineering (e.g., construction materials: marble stones, freestone, aggregates, hydraulic binders) but also with the paper and plastics industries with rubbers, polymers, paints, sealants, adhesives, and pharmaceutical and cosmetic products.

The Uniaxial Compressive Strength (UCS) is one of the most critical mechanical parameters in rocks (Hasanipanah et al. 2022; Hassan & Arman 2022; Moussas & Diamantis 2021). However, in some cases a UCS test cannot be performed because it is costly, time-consuming, and destructive. Therefore, an accurate estimation of this parameter is required (Lai et al. 2016).

Several correlations between mechanical and physical parameters of geomaterials have been established with the UCS. Kurtulus et al. (Kurtulus et al. 2012) determined the mechanical properties of serpentinized ultrabasic rocks through ultrasonic velocity measurements. They found good relationships between UCS and various mechanical parameters (with static elasticity modulus values $R^2=0,7$; with ultrasonic pulse velocity R^2 is more than 0,8 and with Point load index is(50) R^2 is more than 0,9).

Yasar and Erdogan (Yasar & Erdogan 2004) correlated the compressive strength with sound Velocity within carbonate rocks and they found $R^2=0,8$. Within concrete, Del Rıo et al. (Del Río et al. 2004a) reported a exponential

Communicated by: H. Babaie

✉ Mohamed Abdelhedi
mohamed.abdelhedi.etud@fss.usf.tn

[1] Research Laboratory GEOMODELE (LR16ES17), Department of Earth Sciences, Faculty of Sciences, University of Sfax, BP 3000, Sfax, Tunisia

[2] Department of Computer Science and Engineering, College of Engineering, Qatar University, Doha, Qatar

relationship between compressive strength and ultrasonic pulse velocity. Vasconcelos et al. (Vasconcelos et al. 2008) and Chen et al. (X. Chen et al. 2015) reported good relationships within granitic samples and basalt samples. They found determination coefficients 0,7 and 0,8 respectively. Shariati et al. (Shariati et al. 2011) reported a linear relationship between UCS and ultrasonic pulse velocity within concrete samples with $R^2 = 0,9$. A recent Malaysian study established empirical correlations estimating UCS from ultrasonic velocity measurements for granite and schist samples with $R^2 = 0,9$ (Lai et al. 2016).

Moreover, researchers have developed fast and reliable techniques to determine the characteristics of rocks, such as the ultrasonic method, which appears to be a promising technique for experimental laboratory tests (Lai et al. 2016). Numerous research works have developed different predictive models of the UCS in geomaterials. However, they have several drawbacks, such as lack of accuracy (Del Río et al. 2004b) found $R^2 = 0.48$, Abdelhedi et al. (Abdelhedi et al. 2017) found $R^2 = 0,6$; Arman (Arman 2021) found $R^2 = 0,5$), a small sample size. Kumar et al. (Kumar et al. 2020) were studied a Multiple regression model with 7 samples; Xue and Wei (Xue & Wei 2020) were elaborated a hybrid model with 44 data points; Kamaci and Pelin. (Kamaci & Özer 2018) were established empirical models with 9 samples; Abdelhedi et al. (Mohamed Abdelhedi et al. 2020a) were studied artificial neural network models using 66 samples; Sakız et al. (Sakız et al. 2021) were used 37 samples to create fuzzy inference system models predicting drilling rate index from rock strength and cerchar abrasivity index properties), and the study's regional scope (Sharma et al. (Sharma et al. 2017) were Developed a novel models using neural networks and fuzzy systems for the prediction of strength of rocks in India; Ghorbani and Hasanzadehshooiili (Ghorbani & Hasanzadehshooiili 2018) established models to predict UCS and CBR of microsilica-lime stabilized sulfate silty sand using ANN and EPR models in Iran.

Gowida et al. (Gowida et al. 2021) were created models to predict UCS while drilling using artificial intelligence tools in the Eastern province of Saudi Arabia, Barham et al. (Barham et al. 2020) were staudied Artificial Neural Network models to predict UCS in Um-Qais city in Jordan, Assam and Agunwamba (Assam & Agunwamba 2020) were established models to predict CBR and UCS Values of Ntak Clayey Soils in AkwaIbom State, Nigeria**).**

Previous models for predicting the physical and mechanical characteristics of sedimentary rocks, such as carbonate rocks, have been found to have significant limitations and drawbacks (as discussed under the second section of this study). With the ongoing international need for the exploration of new georesources, particularly in the wake of recent economic crises, there is a growing need for new and effective methods of mining exploration. The development of models that can accurately predict the characteristics of sedimentary rocks, such as carbonate rocks, is of paramount importance for the identification and exploration of new georesources. These rocks, which are commonly found in various geological formations and are often used as construction materials, play a crucial role in the building industry (Ammari et al. 2022; Ben Othman et al. 2018; Calvo & Regueiro 2010; Mridekh 2002).

The objective of this study is to evaluate the performance of state-of-the-art machine learning models in predicting the Uniaxial Compressive Strength (UCS) of carbonate rocks using basic physical tests, namely Ultrasonic Pulse Velocity (UPV), density, and effective porosity.

The remainder of this paper is organized as follows. Section 2 presents the dataset and the machine learning techniques employed in this study. In Sect. 3, the results of the computational experiments are presented and analysed. The findings are then discussed and compared to existing literature in Sect. 4. Finally, in Sect. 5, conclusions are presented.

Literature description

The uniaxial compressive strength (UCS) is a critical mechanical property of the rocks used in various engineering projects. It is used to assess the structural stability against the load. To determine the UCS, it is necessary to use high-quality core samples, which are difficult to obtain because of the presence of foliated, fractured, and weak rocks.

Accordingly, several research works have proposed prediction models of the UCS using different tools (new or classic modeling). Table 1 summarizes previous works that established models that predict the UCS.

The previous works presented in Table 1 illustrate defferent limitations, such as the lack of metrics for model evaluation. In other words, some metrics do not reflect the accuracy of models. In addition, the lack of data hinders the creation of good models especially when they are restricted to a specific area or country. This table appears to be summarizing various studies that have used different models and input variables to predict various outputs, such as uniaxial compressive strength (UCS). The studies have used a variety of machine learning techniques, including artificial neural networks (ANN), support vector machines (SVM), extreme learning machines (ELM), multivariate regression, and geostatistical algorithms. The sample sizes for the studies range from 9 to 1771, and the models were trained and tested on samples collected from various locations around the world. The models generally achieved good accuracy, with R-squared values ranging from 0.5 to 0.99 and root mean squared error (RMSE) values ranging from 0.09 to 8.17. However, some of the studies had low sample sizes or were limited to a specific region, which may have reduced

Table 1 Literature review

Reference	Model	samples	Inputs	Output	Accuracy	Disadvantages	Area
Liu et al. (2015)	ELM, SVM and ANN	54	Mineralogical contents, G, ck, n, ne,Id, Vp, Vm	UCS	R^2 from 0.082 to 0.792 MAPE from 3.61 to 37.20%	Low accuracy and low dataset	Trabzon, northeast Turkey
Aboutaleb et al. (2018)	multivariate regression, ANN, SVR	482	ρ, Vp, and Vs		R^2: from 0.5 to 0.8 RMSE: from 7 .35 to 8.17	Regional and Low accuracy	southwest of Iran
Ceryan and Samui (2020)	ELM, MPMR and SVM	47	Wc, n, Id		$R^2=0.9$ RMSE from 0.09 to 0.15	Regional and low dataset	NE Turkey
Nguyen-Sy (2020)	XGB, ANN, SVM	1030	Cement density, BFS, FA, W, Sp, CA, FA, Age		R^2 from 0.91to 0.93 RMSE from 4 to 6	Concrete samples	concrete samples
Gowida et al. 2021	AI:ANN, ANFIS and SVM	1771	ROP, GPM, SPP, RPM, T, and WOB		R = 0.99 AAPE = 3%	Regional	Eastern province of Saudi Arabia
Saedi, B., & Mohammadi, S. D. (2021)	ANN	51	OR, COC, g, Fl% and IS		R = 0.77 to 0.95 MSE 0.003 to 0.023	low dataset and only one method applied	Iran
Mohamed et al. 2018	Simple regression	15	Vp		$R^2=0.93$	Very low dataset, only one method applied and Mortar samples	Tunisia
Arman 2021	Simple regression	48	I_{d2}, ITS		R from 0.53 to 0.63	Evaporitic rocks, low accuracy and low dataset	United Arab Emirates
Kumar et al. 2020	Multiple regression	07	Dbd, PR, SS,DF, BTS, density, abrasivity		VAF = 82.50008, RMSE = 0.1027, MAPE = 0.02728 $R^2=82.50$	low dataset and regional	India
Xue & Wei 2020	LSSVM and ANN	44	BPI, Is(50), SRH and USV		R = 0.979, $R^2=0.959$ and RMSE = 6.58	Granite, schist and sandstone rochks, regional and low dataset	India
Kamani et al., 2020	Simple regression	09	Ep		$R^2=0.8$	Simple regression, Regional and low dataset	Turkey
Abdelhedi et al. 2020a	MULTIPLE REGRESSIONS and ANN	66	Ep, Vp, and density		$R^2=0.83$ $R^2=0.9$	Regional, One metric for model evaluation and low dataset	Tunisia
Teymen & Mengüç, 2020	ANFIS, ANN and GEP	93	Is(50), Vp, BTS, SHH, SSH, UW		$R^2=0.94$ $R^2=0.92$ $R^2=0.95$	Regional and one metric for model evaluation	Turkey
Sharma et al. 2017	ANN, ANFIS	94	SDI, Vp, density		$R^2=0.95$ $R^2=0.98$	Regional, and one metric for model evaluation	India
Ghorbani & Hasanzadehshooiili 2018	ANN and EPR	90	Lime%, Microsilica%, CC, CD and CBR_{10}		$R^2=0.99$ $R^2=0.96$ RMSE = 0.037	Regional and low dataset	Iran
Barham et al. 2020	ANN	56	Density, Vp and SDI		$R^2=0.99$ RMSE = 1.54	Regional and low dataset	Jordan
Onyelowe et al., (2021)	GEP	121	HC Ac γb Ip Cu φ		$R^2=0.99$ Error % = 2.4%	Regional and clayey soil samples	Nigeria

the generalizability of the results. Some studies also only used one evaluation metric or had low accuracy.

Dataset

In this study, 1001 sets of data samples were gathered from references listed in Table 2. These samples came from a variety of countries, as shown in Fig. 1. The data were obtained from scientific articles published in research journals, and they were collected from studies that aimed to create models for predicting physical and mechanical characteristics of carbonate rocks. The data collected for this study were used to train and test the models, and the results of the modeling efforts were used to predict UCS in carbonate rocks.

Machine learning algorithms

•Artificial intelligence (AI) is a rapidly advancing field that encompasses a wide range of computational techniques for clustering, prediction, and classification tasks (Ebid 2020). The development of AI algorithms has led to significant advancements in a variety of fields, including healthcare (Elleuch et al. 2021), agriculture (Ayadi et al. 2020), sustainability (Abulibdeh et al. 2022; R; Jabbar et al. 2021; Zaidan et al. 2022), mines exploration (Mahmoodzadeh et al. 2022a) and transportation (Ben Said & Erradi 2022; Rateb Jabbar et al. 2018; Mirzaei et al. 2022; Mahmoodzadeh et al. 2022b; Mahmoodzadeh et al. 2022c; Mahmoodzadeh et al. 2022d; Mahmoodzadeh et al. 2022e).

The field of geology has seen a significant interest in the application of artificial intelligence (AI) in recent years. AI has been applied to a variety of geoscience-related tasks such as the determination of reservoir rock properties, drilling optimization, and enhanced production facilities (Solanki et al. 2022). Additionally, these techniques have been used in carbonate rock exploration for the prediction of rock and mortar UCS values (M Abdelhedi et al. 2020a, b, c). Furthermore, AI has been applied in mining and geological engineering, including rock mechanics, mining method selection, mining equipment, drilling-blasting, slope stability, and environmental issues (Bui et al. 2021).. These applications of AI in geology demonstrate the potential of this technology to revolutionize the field and provide new insights and solutions to geoscience-related problems. It can lead to more accurate and efficient predictions of geotechnical parameters, understanding of rock properties and ultimately to more efficient and sustainable resource management. The use of AI in geology can also aid in the exploration and discovery of new mineral and energy resources. This highlights the potential of AI to be a valuable tool for geologists and engineers in the field of geology, as it can help to improve our understanding of the earth and its resources.

In this study, four techniques were applied for learning: Random-forest regressor, MLP regressor, support vector machine, and XGB regressor. Cross-validation and the GridSearchCV were used for model optimization. In this study, we focused on supervised machine learning models, which are trained using labeled data and are able to make predictions about new, unseen data. We used the most commonly employed methods for building these models, which involve using algorithms to analyze and learn from the data in order to make accurate predictions. The goal of our study was to compare and evaluate the effectiveness of these methods for predicting geotechnical parameters. By understanding the accuracy and capabilities of these models, we can better understand and predict the behavior of geomaterials, which is important for a variety of engineering applications.

Random-forest Regressor

Over the last decades, random forest (RF) has received considerable attention owing to its reliability and Competitive performance. (Bagherzadeh et al., 2021a; Bagherzadeh

Table 2 Origin of the dataset

	Number of dataset	Country of origin	References
1	15	Tunisia	Abdelhedi et al. 2017
2	40	Hungary	Török & Vásárhelyi 2010
3	16	Italy	Barone et al. 2015
4	31		Pappalardo 2015
5	27	Spain	Gomez-Heras et al. 2020
6	20		Benavente et al. 2006
7	44	Iran	Moradian & Behnia 2009
8	40		Sarkar et al. 2012
9	22		Azimian & Ajalloeian 2015
10	33	India	Madhubabu et al. 2016
11	32		Rahman et al. 2020
12	41	England	Assefa et al. 2003
13	66	Malaysia	Momeni et al. 2015
14	90	Turkey	Çelik 2019
15	13		Yasar & Erdogan 2004
16	55		Ceryan et al. 2013
17	32		Kurtulus et al. 2016
18	12		Sertçelik et al. 2018
19	9		KAMACI et al. 2018
20	13	Chile	González et al. 2019
21	18	Germany	Reyer & Philipp 2014
22	37	Ethiopia	Gudissa et al. 2021
23	30	Jordan	Ahmad 2020
24	190	Thailand	Jaggapan 2017
25	35	Brazil	Silva 2020
26	40	KSA	Al-Osta et al. 2018

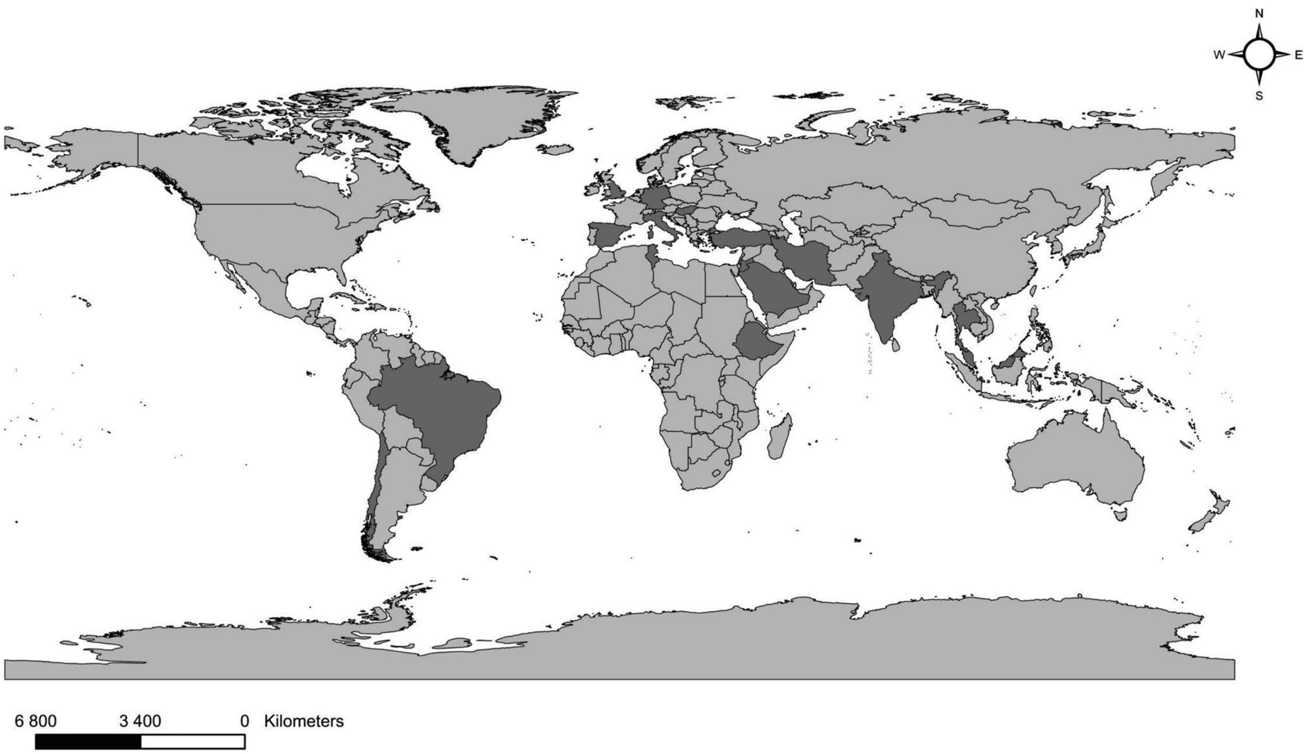
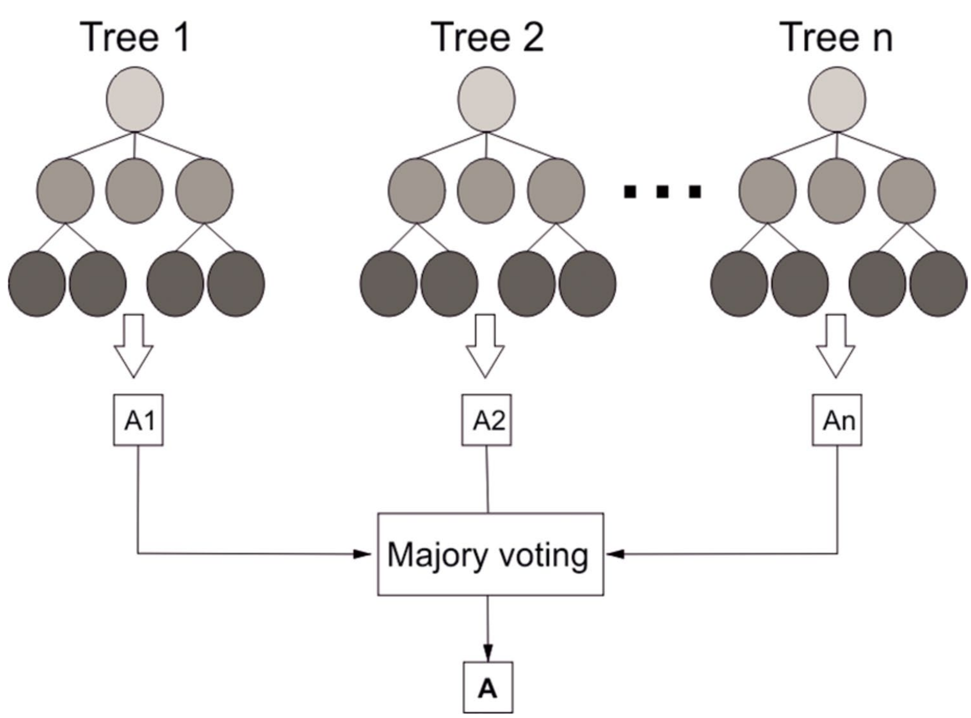


Fig. 1 Samples areas from different countries worldwide

et al., 2021b; Bagherzadeh & Shafighfard 2022; Shafighfard et al. 2022; Tang & Na 2021). Figure 2 illustrates this technique: a supervised ensemble learning algorithm that constructs a "forest" or ensemble of decision trees (DT). Each DT classifies the data instances. The final classification decision is obtained by aggregating the classification results of all the DTs. The common aggregation mechanism is bagging, which attributes the last class based on majority voting.

Fig. 2 Random Forest algorithm



At each node of the decision tree (Fig. 2) entropy is given by:

$$E = -\sum_{i=1}^{c} \text{pixlog(pi)} \quad (1)$$

where E: Entropy.c: The number of unique classes.pi: Prior probability of each given class.

(Schonlau & Zou 2020).

MLPRegressor

Multi Layer Perceptron is a class of neural network that consists of a set of neurons that are connected in a layered fashion. Each neuron at the intermediate layers is fully connected to neurons from the previous layers. At the neuron level, non linear transformation is applied and results are forwarded to the next layer. MLP is trained using backpropagation algorithm with the objective of minimizing a loss function (Okan 2020). The transformation at the neuron is expressed by:

Then, the output can be expressed by:

$$\hat{y}^{(l)} = \int^{(l)} \left(\sum_{i=1}^{q(l-1)} w_{i,j}^{(l)} \hat{y}_i^{(l-1)} + b_j^{(l)} \right) \quad (2)$$

where: $\int^{(l)}$: the activation function of the hidden layer.$\hat{y}_i^{(l-1)}$: the output of the neuron of the (l-1)hidden layer.$w_{i,j}^{(l)}$: the weight between the neuron of the hidden layer and the output layer.b: the bias of the output layer.l: the hidden layer.

(Seo & Cho 2020).

Support vector machine (SVM)

Support vector machine (SVM)(Cortes et al. 1995; Mahmoodzadeh et al. 2022a, b, c, d, e, f) is a traditional machine learning algorithm well known for its simplicity and flexibility in addressing different classification problems. Remarkably, this algorithm has proven its efficiency even for small-scale data sets. The aim of this method is to identify the best position for splitting the data set into a multidimensional space called a hyper plane. A two-dimensional space has a one-dimensional hyper plane, which is just a line. For a three-dimensional space, its hyper plane is a two-dimensional plane that slices the cube, as illustrated in Fig. 3.

Support vector regression (SVR) is a flexible technique not only applicable to linear models but also robust to outliers. Large residuals contribute linearly, whereas the loss function ignores points with small residuals (on the basis of a predefined threshold ε). Using linear kernels, SVR is applicable to linear models. By using radial or polynomial kernels, it becomes suited for non-linear predictions. The expression minimized in SVR is provided below where Le is the loss and c is the cost parameter.

$$c\sum_{i=1}^{n} L_e(yi - \hat{y}i + \sum_{j=1}^{p} \widehat{\beta^2}j \quad (3)$$

$$L_e\left(yi - \hat{y}i\right) = \left\{ \overset{0 \text{ if } yi-\hat{y}i<\varepsilon}{|yi - \hat{y}i| - \varepsilon otherwise} \right\} \quad (4)$$

(Gupta et al. 2019).

XGBRegressor

XGBoost (T. Chen & Guestrin 2016) is another tree-based algorithm that is highly effective and widely used in ML applications. It has successfully solved numerous challenging problems in data science (T. Chen & He 2020; Luckner et al. 2017; Paradkar et al. 2001) and has won several ML competitions (Nielsen 2016). XGBoostis trained using a boosting strategy in which multiple successive weak learners are trained. A weak learner, typically a shallow DT, is generally a lightweight model with several parameters. At each step, another weak learner is added to learn from the error of the previous one, as illustrated in Fig. 4. This algorithm has substantial advantages, including memory efficiency and the specificity of weak learners. More specifically, training vulnerable learners requires less memory than the sequential strategy of RF, where strong learners need to be trained to reach a consensus on an instance

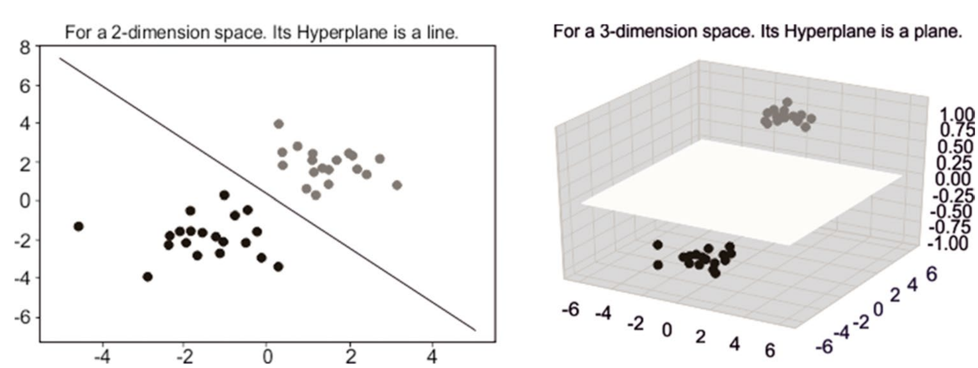


Fig. 3 Support Vector Machine (SVM)

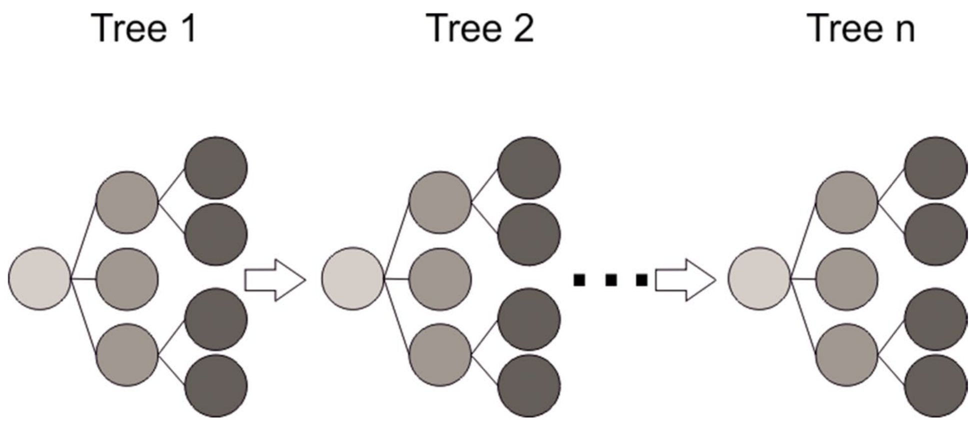


Fig. 4 Gradient Boosting Decision Trees

class. Furthermore, although weak learners do not perform well generally, they perform well in some data instances.

By adding a regularization term into the objective function, the XGBoost algorithm becomes more robust against overfitting. The overall regularized XGBoost loss is expressed as:

$$\mathrm{Obj}^{(r)} = \sum_{i=1}^{n} L\left(yi, \hat{y}_i^{(r)}\right) + \sum_{i=1}^{r} \Omega(g_r) \tag{5}$$

where yi is the real value, $\hat{y}_i^{(r)}$ is the prediction at the r-th round, g_r is the term denoting the structure of the decision tree, $L\left(yi, \hat{y}_i^{(r)}\right)$ is the loss function, n is the number of training examples, and $\Omega(g_r)$ is the regularization term given by:

$$\Omega(g_r) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{6}$$

where T is the number of leaves, w is the weight of the leaves, λ and γ are coefficients whose default values are set at $\lambda = 1$, $\gamma = 0$(Rzychoń et al. 2021).

Cross-validation

Cross validation is a model evaluation method that is better than residuals. The problem with residual evaluations is that they do not give an indication of how well the learner will do when it is asked to make new predictions for data it has not already seen. One way to overcome this problem is to not use the entire data set when training a learner. Some of the data is removed before training begins. Then when training is done, the data that was removed can be used to test the performance of the learned model on ``new'' data (Anderssen et al. 2006; Brereton 2006; Broadhurst & Kell 2006; Westerhuis et al. 2008).

GridSearchCV

Adjustable parameters called hyperparameters can be used to control the training process of a model. To find the best configuration of these hyperparameters, we can use a process called hyperparameter optimization. This involves searching for the combination of hyperparameters that leads to the best model performance. However, this process is often manual and requires significant computational resources.

GridSearchCV is a class established by a scikit-learn framework for parameters adjustment that estimators implement (Müller & Guido 2016).

Model's metrics

In this study, three performance indices, namely the coefficient of determination (R^2), the mean absolute error (MAE) and the root mean square error (RMSE) were used.

The mean absolute error (MAE) and root-mean-square error (RMSE) for evaluating the performance of the established model and the correlation coefficient (R) are defined as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \left| y_i - x_i \right| \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - x_i)^2} \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} (Y_i - \overline{Y}_i)^2} \tag{9}$$

where y_i and x_i denote respectively the preferred output and estimated output;$\overline{y}$ and $\overline{x}$ denote average values, whereas n denotes each sample in the data set (Abdurrahim Akgundogdu 2020; Mahmoodzadeh et al. 2022a, b, c, d, e, f; Tiyasha et al. 2020).

SHapley Additive exPlanations (SHAP)

The Shapley Additive Explanations (SHAP) method was utilized in the analysis of primary factors that influence Uniaxial Compressive Strength (UCS) value of carbonate rocks.

SHAP (Biecek & Burzykowski 2021; Molnar 2022), as a game theoretic method, explains the output of any machine learning model by connecting optimal credit allocation to local explanations through the use of game theory's traditional Shapley values and their related extensions.

In mathematical terms, the Shapley values, denoted by ϕj, provide a way to attribute a "fair" value of a feature j to the prediction of an instance. The Shapley values are defined as the average marginal contribution of a feature j over all possible coalitions of feature values. Mathematically, the Shapley values for a feature j can be defined as:

$$\phi j(f) = (1/|F|!) \sum\_\{S \subseteq F - \{j\}\}|S|!(|F| - |S| - 1)!f(S \cup \{j\}) - f(S) \quad (10)$$

Where F is the set of all features, S is a subset of F, and f is the prediction function.

The SHAP values, denoted by Φj, are a unified measure that combines the Shapley values with local explanations. The SHAP values represent the contribution of a feature j to the prediction of an instance and are defined as:

$$\Phi j(x) = \phi j(f) + E[f(x)] - E[f(x')] \quad (11)$$

Where x is the instance being explained, x' is a reference instance sampled from the background dataset, and E[f(x)] is the expected value of the prediction function over the background dataset.

As conclusion, SHAP uses the Shapley values to attribute a fair value to each feature and combines them with local explanations to provide a unified measure of feature importance called SHAP values.

Carbonate rock tests

Tests must be performed in the laboratory to determine the physical and mechanical parameters of rocks. The uniaxial compressive strength test (UCS), effective porosity, density, and ultrasonic pulse velocity are parameters that determine the mechanical and physical characteristics of rocks.

UCS

We calculated the UCS by dividing in the loaded surface area (MPa) the applied compressive stress applied by the testing machine (Amiri et al. 2022; Y. Liu & Dai 2021; Mohamed et al. 2018).

Ultrasonic velocity test

We used the transmission method to identify the 'P' longitudinal wave velocities. We placed the ultrasonic receiver transducers and the transmitter perpendicularly to the load axis. The ultrasonic device determined the ultrasonic pulse velocity (Mohamed Abdelhedi et al. 2020a; Mohamed Abdelhedi & Abbes 2021).

Effective porosity and density

The volume occupied by the water flow represents the effective porosity. Thus, we saturated the specimens with water to identify the effective porosity (P_e), defined as the following:

$$P_e = V_{pi}/V_t \quad (12)$$

Where V_{pi} and V_t represent respectively the volume of the connected pores and the sample volume (Lafhaj & Goueygou 2009; Peng & Zhang 2007).

The density represents the mass of the specimen contained in a given volume unit, expressed in kN/m^3 or kg/m^3(Mohamed Abdelhedi et al. 2020b; Peng & Zhang 2007).

Results and discussion

We conduct correlation analysis to investigate the relationship between data features. Figure 5 shows relationships between different variables (dependent and independent). It is noted that the coefficient of determination varies between -1 and 1. When the color is very dark or very light, a strong relationship between the two corresponding variables is determined.

This figure shows a strong negative linear relationship between the uniaxial compressive strengthand effective porosity, and in contrast, a strong negative linear relationship between effective porosity and ultrasonic wave velocity. However, there is a strong positive linear relationship between ultrasonic velocity and uniaxial compressive strength.

Nguyen-Sy et al. (Nguyen-Sy et al. 2020) used a similar representation in rating the relationship between cement ratio, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, age, and UCS within concrete. This study found a good relationship between the UCS and age and between the UCS and cement ratio within concrete samples.

Table 3 shows the statistical parameters of the dataset. The range of all variables was enormous: ultrasonic velocity was 6325 m/s, density was 1.91, effective porosity was 42.14%, and MPa of UCS was 179.76. This extensive range allows a good modelling margin, making the model more valuable and the prediction more feasible.

The density of the samples varied between 1.43 and 3.34, where as the effective porosity varied between 0.01%

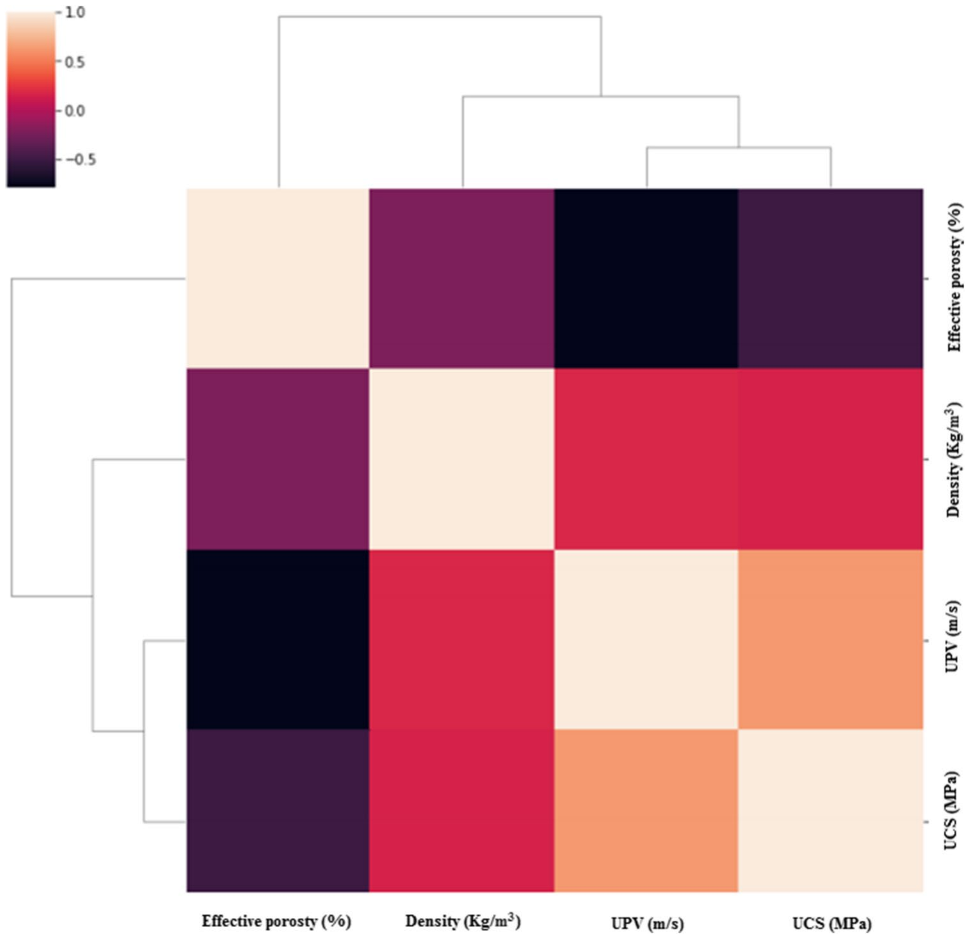


Fig. 5 Variables Heatmap

Table 3 Dataset statistical parameters

	Min	Max	Range	Mean	SD
Vp	1110	7435	6325	4514.9	1268.41
D	1.43	3.34	1.91	2.53	0.23
Pe	0.01	42.15	42.14	8.09	7.73
UCS	0.74	180.5	179.76	63.34	34.77

and more than 40%.Theuniaxial compressive strength varied between less than 1and more than 180 MPa, where as ultrasonic velocity varied between 1110 and 7435 m/s. These results indicate different categories of carbonate rocks such as hard, ductile, and brittle.

Four machine learning algorithms were applied to create four different models predicting this parameter: 'RandomForestRegressor' (Fig. 6), 'MLPRegressor' (Fig. 7), 'support vector machine' (Fig. 8), and 'XgboostRegressor' (Fig. 9).

RandomForestRegressor algorithm (Fig. 6) shows a straight distribution of points, giving a coefficient of determination (R^2) equal to 0.65.

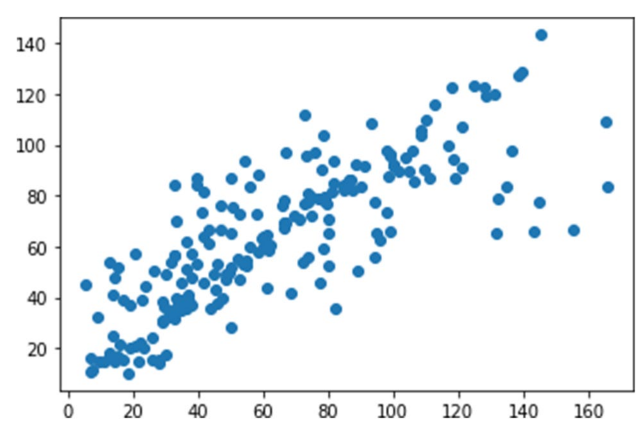


Fig. 6 Expected versus observed UCS values using RandomForestRegressor modeling

The points illustrated in Fig. 7 are more aligned, providing a better coefficient of determination ($R^2 = 0.94$). This figure shows the model created by the MLPRegressor algorithm.

Figure 8 presents a model correlating the predicted UCS values using SVM modelling with the tested values. The

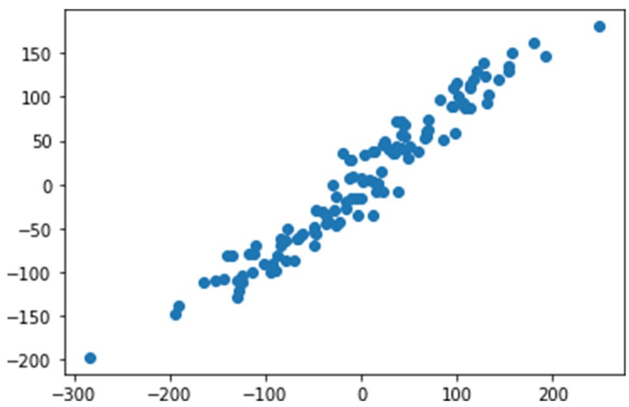


Fig. 7 Expected versus observed UCS values using MLPRegressor modeling

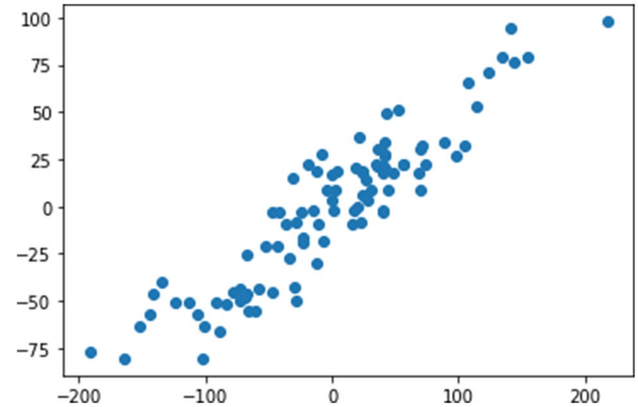


Fig. 9 Expected versus observed UCS values using XgboostRegressor modeling

coefficient of determination of this linear relationship is 0.78.

The XgboostRegressor algorithm's model gives a linear relationship with $R^2 = 0.89$ between predicted UCS values and observed values.

These models were optimized using 'Grid_searchcv' as an optimization algorithm and validated using 'cross-validation.

Table 4 shows the evaluation of the different models using different metrics. In this study, we compared the prediction accuracy of UCS in various machine learning models. We employed three different metrics: mean squared error (MSE), coefficient of determination (R^2), and mean absolute error (MAE).

Before optimization and validation of models, the MLPRegressor algorithm had the lowest mean squared error (584.06), the lowest mean absolute error (20.07), and the best coefficient of determination (0.94). However, with the RandomForestRegressor algorithm, MSE = 5949.55, MAE = 60.35 and $R^2 = 0.65$, with the SVM algorithm, MSE = 3109.17, MAE = 40.99 and $R^2 = 0.78$, and with the XGBRegressor algorithm, MSE = 1753.56, MAE = 32.83 and $R^2 = 0.85$.

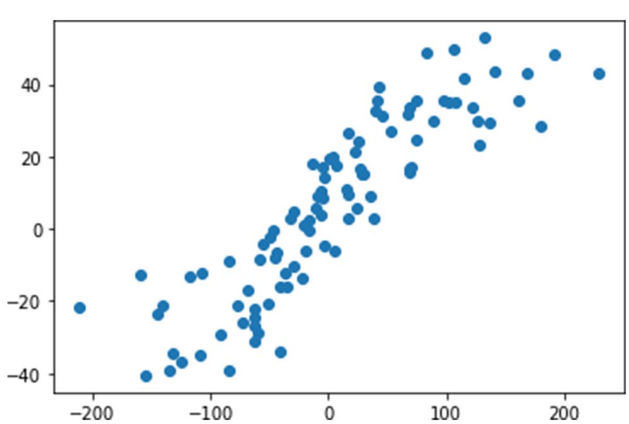


Fig. 8 Expected versus observed UCS values using SVM modeling

After model optimization, the majority of scores improved, and the results show that both SVM and MPL models are the best, with a score equal to 0.91.

After model validation, the cross-validation algorithm divides the data into four parts, and we obtained four very similar scores for each metric, which indicates very good validations (Table 4).

The model that contains the lowest number of errors was created by the XGBRegressor algorithm (MSE is between 438.95 and 590.46, and MAE is between 17.22 and 18.79). However, two models show good coefficients of determinations (R^2 of the MLPRegressor model is between 0.92 and 0.94, and R^2 of the XGBRegressor model is between 0.91 and 0.94).

The model created by MLPRegressor showed, after validation, good coefficients of determination, but it also had vast errors (more than 6000 of MSE).

The results indicated the best model that presents the best coefficients of determinations and fewer errors is the model created by the XGBRegressor algorithm.

Furthermore, a three-fold cross-validation analysis(Schaffer & Edu 1993) was performed to validate the performance of the proposed model and mitigate the potential issue of over fitting. The data was partitioned into three equal segments and each segment was utilized as the validation set while the remaining two were employed as the training set. The results of the validation were then averaged to obtain a comprehensive accuracy score for the model. This procedure was repeated three times, with each segment utilized once as the validation set, thereby ensuring the comprehensive testing of the model on all available data. The results of the analysis confirmed the obtained findings and demonstrate that the best model, in terms of its coefficients of determination and lower error rates, was the model created by the XGBRegressor algorithm.

Liu et al. (Z. Liu et al. 2015) also used MLPRegressor (artificial neural networks using an extreme learning

Table 4 Machine learning models evaluation

	Random Forest Regressor model				MLP Regressor model				SVM model				XGB Regressor model			
MSE	5949.55				**584.06**				3109.17				1753.56			
MAE	60.35				**20.07**				40.99				32.83			
R^2	0.65				**0.94**				0.78				0.85			
Grid_searchcv score	0.90				**0.91**				**0.91**				0.89			
Cross val MSE	530.55	565.73	641.20	614.88	6714.54	6435.86	6581.30	6723.56	4292.67	3936.61	4353.05	4822.44	**438.95**	**516.44**	**562.16**	**590.46**
Cross val MAE	18.83	18.92	20.76	17.63	40.79	39.01	40.02	41.10	53.05	46.74	48.46	49.32	**17.22**	**18.21**	**18.79**	**18.19**
Cross val R^2	0.87	0.89	0.88	0.88	**0.94**	**0.93**	**0.92**	**0.92**	0.39	0.39	0.33	0.33	**0.94**	**0.92**	**0.91**	**0.92**

The entries in bold in the table are interesting values for model evaluation

machine) and found scores of approximately 0.7. However, they employed small data sets to estimate the UCS of carbonate rocks (54 samples).

Aboutaleb et al. (Aboutaleb et al. 2018) used 482 samples to create three models (support vector machine, artificial neural network, and multiple regression analysis) predicting the UCS of carbonate rocks. The authors found three R^2 higher than 0.9. However, they were selected from one place (Iran), and thus, the interpretation of the results was regional.

Ceryan and Samui (Ceryan & Samui 2020) established three models by applying the extreme learning machine (ELM), the minimax probability machine regression (MPMR), and the least square support vector machine (LS-SVM). They found R^2 of approximately 0.9; however, they used just 47 samples, and the study was localized in NE Turkey.

Nguyen-Sy (Nguyen-Sy et al. 2020) established three models by applying the ANN, SVM, and XGBoost methods with 1030 collected concrete datasets. They found R^2 between 0.91 and 0.93.

From the Middle East region in the Eastern Province of Saudi Arabia, a data set of 1771 data points was obtained. To create the models, researchers employed the support vector machine (SVM), the adaptive neuro-fuzzy inference system (ANFIS), and the artificial neural network (ANN).Models were evaluated using the R-value and AAPE as metrics (Gowida et al. 2021).

The Shapley Additive Explanations (SHAP) method was utilized in this study to conduct a comprehensive analysis of the primary factors that impact the Uniaxial Compressive Strength (UCS) of carbonate rocks. The SHAP method, rooted in coalitional game theory, was employed to calculate the Shapley values, which are a fair measure of feature importance among the data instances. The feature values were considered as players in a coalition and the Shapley values determined their relative contributions to the prediction of UCS.

The results of the feature importance analysis, as presented in Fig. 10, indicated that Density and Porosity were the most significant features affecting the Uniaxial Compressive Strength (UCS) of carbonate rocks. In contrast, Ultrasonic Pulse Velocity was found to have limited impact on the prediction of UCS. These results suggest that Density and Porosity play a crucial role in determining the UCS of carbonate rocks.

To further evaluate the generalization capability of the proposed model, a second feature importance method, Permutation Importance, was applied using the Eli5 library (Korobov 2017). The results of this analysis were consistent with the findings obtained through the SHAP method, emphasizing the crucial role of Density and Porosity in the prediction of UCS. The weight values of Density and

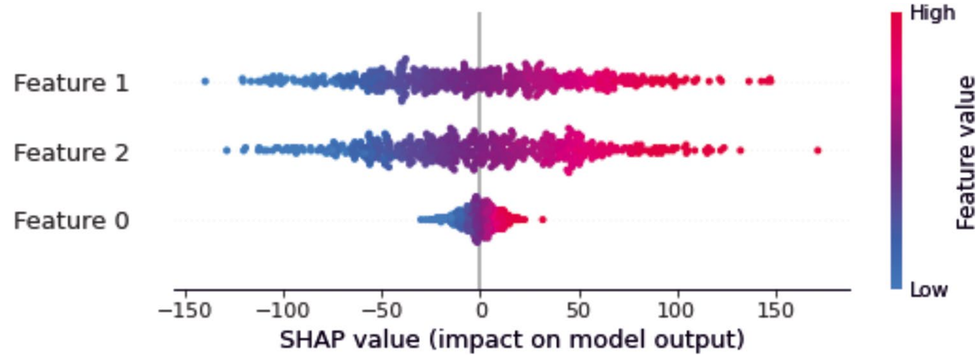


Fig. 10 Feature importance analysis of UCS factors using SHAP

Porosity were 0.9014 ± 0.0876 and 0.7843 ± 0.0822, respectively, while Ultrasonic Pulse Velocity had a weight value of 0.0291. These results reinforce the conclusion that Density and Porosity are the primary determinants of UCS in carbonate rocks.

Conclusions

The goal of this study was to develop an accurate international model for predicting the uniaxial compressive strength (UCS) of carbonate rocks using ultrasonic velocity, effective porosity, and density as input variables. A dataset was compiled from 26 countries worldwide, consisting of data from scientific papers that used these input variables to predict UCS. Four artificial intelligence models were trained and tested using this dataset: random forest regressor, MLPRegressor, SVM, and XGBRegressor.

Initially, the model developed using the MLPRegressor method was found to be the best according to the evaluation metrics used. However, after optimization and validation, both the MLPRegressor and XGBRegressor models were found to have good performance based on the R^2 metric. Upon further evaluation using all three metrics (R^2, MSE, and MAE), the XGBRegressor model was found to be the most accurate, with R^2 values between 0.92 and 0.94, MSE less than 600, and MAE less than 20.

This study represents the first attempt to predict the UCS of carbonate rocks using a model that spans 16 countries and four continents. The results of this study show that the XGBRegressor model developed in this study can be used to accurately estimate the UCS of any carbonate rock found on the earth's surface.

As future work, we plan to investigate the use of advanced machine learning techniques, such as deep learning or transfer learning, to develop and refine models for predicting the uniaxial compressive strength of carbonate rocks. This could potentially improve the accuracy and performance of the models. Additionally, we will continue to study the influence of variables such as grain size, mineral composition, and rock type on the uniaxial compressive strength of carbonate rocks to gain a more comprehensive understanding of the factors that contribute to the strength of these materials.

Author's contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Rateb Jabbar, Ahmed Ben Said, Noora Fetais and Chedly Abbes. The first draft of the manuscript was written by Mohamed Abdelhedi. All authors read and approved the final manuscript.

Funding This work received financial support from the "Ministère de l'Enseignement Supérieur et de la Recherche Scientifique en Tunisie". Experimental assays were performed in the 'Département des Sciences de la Terre' of the 'Faculté des Sciences de Sfax, Université de Sfax-Tunisie'.

Data availability The dataset and associated source codes analyzed during the current study are available in the GitHub repository (https://github.com/RatebJabbar/uniaxial-compressive-strength-within-carbonate-rocks).

Declarations

Conflict of interest There is no financial or personal relationship between the authors of this paper and any other individuals or organizations that could inappropriately influence or bias the content of the paper.

References

Abdelhedi M, Abbes C (2021) Study of physical and mechanical properties of carbonate rocks and their applications on georesources exploration in Tunisia. Carbonates Evaporites 36(2):1–13. https://doi.org/10.1007/S13146-021-00688-8/FIGURES/12

Abdelhedi M, Jabbar R, Mnif T, Abbes C (2020) Prediction of uniaxial compressive strength of carbonate rocks and cement mortar using artificial neural network and multiple linear regressions. Acta Geodynamica Et Geromaterialia 17(3):367–378

Abdelhedi M, Jabbar R, Mnif T, Abbes C (2020) Ultrasonic velocity as a tool for geotechnical parameters prediction within carbonate rocks aggregates. Arab J Geosci 13(4):1–11. https://doi.org/10.1007/S12517-020-5070-0/FIGURES/10

Abdelhedi M, Abbes C, M A, Aloui M, Mnif T (2017) Ultrasonic velocity as a tool for mechanical and physical parameters prediction within carbonate rocks. Res Gate Net 13(3):371-384.https://doi.org/10.12989/gae.2017.13.3.371

Abdelhedi M, Jabbar R, Mnif T, Abbes C(2020a). Prediction of uniaxial compressive strength of carbonate rocks and cement mortar using artificial neural network and multiple linear regressions. Irsm Cas Cz, 17(3):367–377. https://doi.org/10.13168/AGG.2020.0027

Aboutaleb S, Behnia M, Bagherpour R, Bluekian B (2018) Using non-destructive tests for estimating uniaxial compressive strength and static Young's modulus of carbonate rocks via some modeling techniques. Bull Eng Geol Environ 77:4. https://doi.org/10.1007/s10064-017-1043-2

Abulibdeh A, Zaidan E, Jabbar R (2022) The impact of COVID 19 pandemic on electricity consumption and electricity demand forecasting accuracy Empirical evidence from the state of Qatar. Energy Strategy Reviews 44:100980 https://doi.org/10.1016/J.ESR.2022.100980

Abdurrahim A (2020)Comparative Analysis of Regression Learning Methods for Estimation of Energy Performance of Residential Structures Erzincan University. J Sci Technol 13(2):600-608.https://doi.org/10.18185/erzifbed.691398

Amiri M, Lashkaripour GR, Hafezi Moghaddas N, Ghobadi MH, Amiri M (2022) Estimating Uniaxial Compressive Strength of Ilam. Limestones Formation from Index Parameters by Learning Methods

Ammari A, Abbes C, Abida H (2022) Geometric properties and scaling laws of the fracture network of the Ypresian carbonate reservoir in central Tunisia Examples of Jebels Ousselat and Jebil. J Afr Earth Sci 196:104718. https://doi.org/10.1016/j.jafrearsci.2022.104718

Anderssen E, Dyrstad K, Westad F, Martens H (2006) Reducing over-optimism in variable selection by cross-model validation Chemometrics and Intelligent Laboratory Systems 84 1–2 SPEC ISS.https://doi.org/10.1016/j.chemolab.2006.04.021

Arman H (2021) Correlation of Uniaxial Compressive Strength with Indirect Tensile Strength Brazilian and 2nd Cycle of Slake Durability Index for Evaporitic Rocks. Geotechnical and Geological Engineering 39:2.https://doi.org/10.1007/s10706-020-01578-x

Assam SA, Agunwamba JC (2020) Potentials of Processed Palm Kernel Shell Ash Local Stabilizer and Model Prediction of CBR and UCS Values of Ntak Clayey Soils in Akwa Ibom State Nigeria. European Journal of Engineering Research and Science 5:12 https://doi.org/10.24018/ejers.2020.5.12.2143

Ayadi S, Ben Said A, Jabbar R, Aloulou C, Chabbouh A, Achballah AB (2020) Dairy Cow Rumination Detection: A Deep Learning Approach. Communications in Computer and Information Science 1348:123–139. https://doi.org/10.1007/978-3-030-65810-6_7/COVER

Bagherzadeh F, Mehrani MJ, Basirifard M, Roostaei J (2021a) Comparative study on total nitrogen prediction in wastewater treatment plant and effect of various feature selection methods on machine learning algorithms performance. J Wat Proc Eng 41.https://doi.org/10.1016/j.jwpe.2021.102033

Bagherzadeh F, Nouri AS, Mehrani MJ, Thennadil S (2021b) Prediction of energy consumption and evaluation of affecting factors in a full-scale WWTP using a machine learning approach. Process Safety and Environmental Protection 154.https://doi.org/10.1016/j.psep.2021.08.040

Bagherzadeh F, Shafighfard T (2022) Ensemble Machine Learning approach for evaluating the material characterization of carbon nanotube-reinforced cementitious composites. Case Studies in Construction Materials 17:e01537. https://doi.org/10.1016/j.cscm.2022.e01537

Barham WS, Rabab'ah SR, Aldeeky HH, Al Hattamleh OH (2020) Mechanical and Physical Based Artificial Neural Network Models for the Prediction of the Unconfined Compressive Strength of Rock. Geotechnical and Geological Engineering 38:5.https://doi.org/10.1007/s10706-020-01327-0

Ben Othman D, Ayadi I, Abida H, Laignel B (2018) Spatial and inter-annual variability of specific sediment yield: case of hillside reservoirs in Central Tunisia. Bull Eng Geol Environ 77:1. https://doi.org/10.1007/s10064-016-0976-1

Ben Said A, Erradi A (2022) Spatiotemporal Tensor Completion for Improved Urban Traffic Imputation. IEEE Trans Intell Transp Syst 23(7):6836–6849. https://doi.org/10.1109/TITS.2021.3062999

Biecek P, Burzykowski T (2021) Shapley Additive Explanations SHAP for Average Attributions In Explanatory Model Analysis 95–106 https://doi.org/10.1201/9780429027192-10

Brereton RG (2006) Consequences of sample size variable selection, and model validation and optimisation, for predicting classification ability from analytical data. TrAC Trends in Analytical Chemistry 25:11.https://doi.org/10.1016/j.trac.2006.10.005

Broadhurst DI, Kell DB (2006) Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics 2:4. https://doi.org/10.1007/s11306-006-0037-z

Bui XN, Bui HB, Nguyen H (2021) A Review of Artificial Intelligence Applications in Mining and Geological Engineering 109:109–142. https://doi.org/10.1007/978-3-030-60839-2_7/COVER

Calvo JP, Regueiro M (2010) Carbonate rocks in the mediterranean region From classical to innovative uses of building stone. Geological Society Special Publication 331.https://doi.org/10.1144/SP331.3

Ceryan N, Samui P (2020) Application of soft computing methods in predicting uniaxial compressive strength of the volcanic rocks with different weathering degree. Arab J Geosci 13:7. https://doi.org/10.1007/s12517-020-5273-4

Chen X, Schmitt DR, Kessler JA, Evans J, Kofman R (2015) Empirical relations between ultrasonic P-wave velocity porosity and uniaxial compressive strength. CSEG Rec 40(5):24–29

Chen T, Guestrin C (2016) XGBoost A scalable tree boosting system Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 13–17-August-2016 785–794.https://doi.org/10.1145/2939672.2939785

Chen T, He T (2020) xgboost: eXtreme Gradient Boosting

Cortes C, Vapnik V, Saitta L (1995) Support vector networks. Machine Learning 1995 20 3 20 3 273 297 https://doi.org/10.1007/BF00994018

Del Río LM, Jiménez A, López F, Rosa FJ, Rufo MM, Paniagua JM (2004) Characterization and hardening of concrete with ultrasonic testing. Ultrasonics 42(1):9. https://doi.org/10.1016/j.ultras.2004.01.053

Del Río LM, Jiménez A, López F, Rosa FJ, Rufo MM, Paniagua JM (2004) Characterization and hardening of concrete with ultrasonic testing. Ultrasonics 421:9. https://doi.org/10.1016/j.ultras.2004.01.053

Ebid AM (2020) 35 Years of (AI) in Geotechnical Engineering State of the Art. Geotechnical and Geological Engineering 2020 39 2 39 2 637 690 https://doi.org/10.1007/S10706-020-01536-7

Elleuch MA, Hassena ABen, Abdelhedi M, Pinto FS (2021) Real time prediction of COVID 19 patients health situations using Artificial Neural Networks and Fuzzy Interval. Mathematical modeling Applied Soft Computing 110:107643.https://doi.org/10.1016/J.ASOC.2021.107643

Ghorbani A, Hasanzadehshooiili H (2018) Prediction of UCS and CBR of microsilica-lime stabilized sulfate silty sand using ANN and EPR models application to the deep soil mixing. Soils Found 58:1. https://doi.org/10.1016/j.sandf.2017.11.002

Gowida A, Elkatatny S, Gamal H (2021) Unconfined compressive strength UCS prediction in real-time while drilling using artificial intelligence tools. Neural Comput Appl 33:13. https://doi.org/10.1007/s00521-020-05546-7

Gupta I, Devegowda D, Jayaram V, Rai C, Sondergeld C (2019) Machine learning regressors and their metrics to predict synthetic sonic and mechanical properties. Interpretation 7:3. https://doi.org/10.1190/INT-2018-0255.1
Hasanipanah M, Jamei M, Mohammed AS, Amar MN, Hocine O, Khedher KM (2022) Intelligent prediction of rock mass deformation modulus through three optimized cascaded forward neural network models. Earth Sci Inf 15(3):1659–1669. https://doi.org/10.1007/s12145-022-00823-6
Hassan MY, Arman H (2022) Several machine learning techniques comparison for the prediction of the uniaxial compressive strength of carbonate rocks. Sci Rep 12(1):20969.https://doi.org/10.21203/rs.3.rs-1712005/v1
Jabbar R, Al-Khalifa K, Kharbeche M, Alhajyaseen W, Jafari M, Jiang S (2018) Applied Internet of Things IoT Car monitoring system for Modeling of Road Safety and Traffic System in the State of Qatar 2018 3 ICTPP1072 https://doi.org/10.5339/QFARC.2018.ICTPP1072
Jabbar R, Zaidan E, Said B, Ghofrani A, Jabbar R, Zaidan E, Ghofrani A (2021) Reshaping Smart Energy Transition: An analysis of human-building interactions in Qatar Using Machine Learning Techniques
Kamaci Z, Ozer P (2018) Engineering Properties of Egirdir-Kızıldag Harzburgitic Peridotites in Southwestern Turkey. International Journal of Computational and Experimental Science and Engineering 4:2.https://doi.org/10.22399/ijcesen.348339
Korobov M (2017) eli5. https://github.com/eli5-org/eli5/blob/master/docs/source/blackbox/permutation_importance.rst
Kumar V, Vardhan H, Murthy CSN (2020) Multiple regression model for prediction of rock properties using acoustic frequency during core drilling operations Geomechanics and Geoengineering 15 4 https://doi.org/10.1080/17486025.2019.1641631
Kurtulus C, Bozkurt A, Endes H (2012) Physical and mechanical properties of Serpentinized ultrabasic rocks in NW Turkey. Pure Appl Geophys 169:7. https://doi.org/10.1007/s00024-011-0394-z
Lafhaj Z, Goueygou M (2009) Experimental study on sound and damaged mortar: Variation of ultrasonic parameters with porosity. Constr Build Mater 23:2. https://doi.org/10.1016/j.conbuildmat.2008.05.012
Lai GT, Rafek AG, Serasa AS, Hussin A, Ern LK (2016) Use of ultrasonic velocity travel time to estimate uniaxial compressive strength of granite and schist in Malaysia. Sains Malaysiana 45:2
Liu Y, Dai F (2021) A review of experimental and theoretical research on the deformation and failure behavior of rocks subjected to cyclic loading. In Journal of Rock Mechanics and Geotechnical Engineering 13(5):1203–1230. https://doi.org/10.1016/j.jrmge.2021.03.012
Liu Z, Shao J, Xu W, Wu Q (2015) Indirect estimation of unconfined compressive strength of carbonate rocks using extreme learning machine. Acta Geotech 10:5. https://doi.org/10.1007/s11440-014-0316-1
Luckner M, Topolski B, Mazurek M (2017) Application of XGBoost algorithm in fingerprinting localisation task. IFIP International Conference on Computer Information Systems and Industrial Management 661:671
Mahmoodzadeh A, Mohammadi M, Abdulhamid SN, Ali HFH, Ibrahim HH, Rashidi S (2022) Forecasting tunnel path geology using Gaussian process regression. Geomechanics and Engineering 28(4):359–374
Mahmoodzadeh A, Mohammadi M, Abdulhamid SN, Ibrahim HH, Ali HFH, Nejati HR, Rashidi S (2022) Prediction of duration and construction cost of road tunnels using Gaussian process regression. Geomechanics and Engineering 28(1):65–75
Mahmoodzadeh A, Mohammadi M, Abdulhamid SN, Ibrahim HH, Ali HFH, Nejati HR, Rashidi S (2022b) Prediction of duration and construction cost of road tunnels using Gaussian process regression. Geomechanics and Engineering 28(1):65-75.https://doi.org/10.12989/GAE.2022.28.1.065
Mahmoodzadeh A, Nejati HR, Ibrahim HH, Ali HFH, Mohammed A, Rashidi S, Majeed MK (2022c) Several models for tunnel boring machine performance prediction based on machine learning. Geomechanics and Engineering 30(1):75 91.https://doi.org/10.12989/gae.2022.30.1.075
Mahmoodzadeh, A., Nejati, H. R., Mohammadi, Ibrahim, H. H., Rashidi, S., & Mohammed, A., 2022d Meta-heuristic Optimization algorithms for Prediction of Fly-rock in the Blasting Operation of Open-Pit Mines Geomechanics and Engineering 30 6 489 502 https://doi.org/10.12989/gae.2022.30.6.489
Mahmoodzadeh A, Ali HFH, Ibrahim HH, Mohammed A, Rashidi S, Mahmood ML, Ali MS (2022e) Application of Autoregressive Model in the Construction Management of Tunnels Acta Montanistica Slovaca 27(3):581-588. https://doi.org/10.46544/AMS.v27i3.02
Mirzaei, M., Mahmoodzadeh, A., Ibrahim, H., Rashidi, S., Majeed, M. K., Mohammed, A. 2022 Prediction of squeezing phenomenon in tunneling projects Application of Gaussian process regression Geomechanics and Engineering 30 1 1126 https://doi.org/10.12989/gae.2022.30.1.011
Mohamed A, Thameur M, Chedly A (2018) Ultrasonic Velocity as a Tool for Physical and Mechanical Parameters Prediction within Geo-Materials: Application on Cement Mortar. Russ J Nondestr Test 54(5):345–355. https://doi.org/10.1134/S1061830918050091
Molnar, C. 2022 9.6 SHAP SHapley Additive exPlanations | Interpretable Machine Learning https://christophm.github.io/interpretable-ml-book/shap.html
Moussas VC, Diamantis K (2021) Predicting uniaxial compressive strength of serpentinites through physical dynamic and mechanical properties using neural networks. Journal of Rock Mechanics and Geotechnical Engineering 13:1. https://doi.org/10.1016/j.jrmge.2020.10.001
Mridekh, Abdelaziz. 2002 Géodynamique des bassins mésocénozoïques de subsurface de l'offshore d'Agadir Maroc sud-occidental contribution à la reconnaissance de l'histoire atlasique d'un segment de la marge atlantique marocaine
Müller, A. C., & Guido, S. 2016 Introduction to machine learning with Python: a guide for data scientists "O'Reilly Media Inc."
Nguyen-Sy T, WakimJ, ToQD, VuMN, NguyenTD, NguyenTT (2020) Predicting the compressive strength of concrete from its compositions and age using the extreme gradient boosting method Construction and Building Materials 260 https://doi.org/10.1016/j.conbuildmat.2020.119757
Nielsen, D. 2016 Tree boosting with xgboost-why does xgboost win" every" machine learning competition? NTNU
Okan M (2020) AERODYNAMIC FORCE FORECASTING WITH MACHINE LEARNING. Istanbul Technical University, Faculty of Aeronautics and Astronautics
Paradkar, M. M., Singhal, R. S., & Kulkarni, P. R. 2001 An approach to the detection of synthetic milk in dairy milk 4 Effect of the addition of synthetic milk on the flow behaviour of pure cow milk International Journal of Dairy Technology 54 1 36 37 https://doi.org/10.1046/j.1471-0307.2001.00005.x
Peng S, Zhang J (2007) Engineering geology for underground rocks. In Engineering Geology for Underground Rocks. https://doi.org/10.1007/978-3-540-73295-2
Rzychoń, M., Żogała, A., & Róg, L. 2021 Experimental study and extreme gradient boosting XGBoost based prediction of caking ability of coal blends Journal of Analytical and Applied Pyrolysis 156 https://doi.org/10.1016/j.jaap.2021.105020
Sakız U, Kaya GU, Yaralı O (2021) Prediction of drilling rate index from rock strength and cerchar abrasivity index properties using

fuzzy inference system. Arab J Geosci 14:5. https://doi.org/10.1007/s12517-021-06647-w

Schaffer, C., & Edu, S. A. H. C. 1993 Selecting a classification method by cross-validation Machine Learning 1993 13 1 13 1) 135–143 https://doi.org/10.1007/BF00993106

Schonlau M, Zou RY (2020) The random forest algorithm for statistical learning. Stata Journal 20:1. https://doi.org/10.1177/1536867X20909688

Seo, H., & Cho, D. H. 2020 Cancer-Related Gene Signature Selection Based on Boosted Regression for Multilayer Perceptron IEEE Access 8 https://doi.org/10.1109/ACCESS.2020.2985414

Shafighfard T, Bagherzadeh F, Rizi RA, Yoo D-Y (2022) Data-driven compressive strength prediction of steel fiber reinforced concrete SFRC subjected to elevated temperatures using stacked machine learning algorithms. Journal of Materials Research and Technology 21(3777):3794. https://doi.org/10.1016/j.jmrt.2022.10.153

Shariati, M., Ramli-Sulong, N. H., Mohammad Mehdi Arabnejad, K. H., Shafigh, P., & Sinaei, H. 2011 Assessing the strength of reinforced Concrete Structures Through Ultrasonic Pulse Velocity And Schmidt Rebound Hammer tests Scientific Research and Essays 6 1

Sharma, L. K., Vishal, V., & Singh, T. N. 2017 Developing novel models using neural networks and fuzzy systems for the prediction of strength of rocks from key geomechanical properties Measurement Journal of the International Measurement Confederation 102 https://doi.org/10.1016/j.measurement.2017.01.043

Solanki P, Baldaniya D, Jogani D, Chaudhary B, Shah M, Kshirsagar A (2022) Artificial intelligence: New age of transformation in petroleum upstream. Petroleum Research 7(1):106–114. https://doi.org/10.1016/J.PTLRS.2021.07.002

Tang L, Na SH (2021) Comparison of machine learning methods for ground settlement prediction with different tunneling datasets. Journal of Rock Mechanics and Geotechnical Engineering 136. https://doi.org/10.1016/j.jrmge.2021.08.006

Tiyasha, Tung, T. M., & Yaseen, Z. M. 2020 A survey on river water quality modelling using artificial intelligence models 2000–2020 In Journal of Hydrology Vol 585). https://doi.org/10.1016/j.jhydrol.2020.124670

Vasconcelos G, Lourenço PB, Alves CAS, Pamplona J (2008) Ultrasonic evaluation of the physical and mechanical properties of granites. Ultrasonics 48:5. https://doi.org/10.1016/j.ultras.2008.03.008

Westerhuis JA, Hoefsloot HCJ, Smit S, Vis DJ, Smilde AK, Velzen EJJ, Duijnhoven JPM, Dorsten FA (2008) Assessment of PLSDA cross validation Metabolomics 4:1. https://doi.org/10.1007/s11306-007-0099-6

Xue X, Wei Y (2020) A hybrid modelling approach for prediction of UCS of rock materials. CR Mec 348:3. https://doi.org/10.5802/CRMECA.17

Yasar E, Erdogan Y (2004) Correlating sound velocity with the density compressive strength and Young's modulus of carbonate rocks. Int J Rock Mech Min Sci 41:5. https://doi.org/10.1016/j.ijrmms.2004.01.012

Zaidan E, Abulibdeh A, Alban A, Jabbar R (2022) Motivation preference socioeconomic and building features New paradigm of analyzing electricity consumption in residential buildings. Build Environ 219:109177. https://doi.org/10.1016/J.BUILDENV.2022.109177