




## Article

# Machine Learning-Based Intelligent Prediction of Elastic Modulus of Rocks at Thar Coalfield

Niaz Muhammad Shahani <sup>1,2</sup> , Xigui Zheng <sup>1,2,3,4,\*</sup>, Xiaowei Guo <sup>1,2</sup>  and Xin Wei <sup>1,2</sup> 

- <sup>1</sup> School of Mines, China University of Mining and Technology, Xuzhou 221116, China; shahani.niaz@cumt.edu.cn (N.M.S.); ckgwx@cumt.edu.cn (X.G.); ts21020059a31@cumt.edu.cn (X.W.)
- <sup>2</sup> The State Key Laboratory for Geo Mechanics and Deep Underground Engineering, China University of Mining & Technology, Xuzhou 221116, China
- <sup>3</sup> School of Mines and Civil Engineering, Liupanshui Normal University, Liupanshui 553004, China
- <sup>4</sup> Guizhou Guineng Investment Co., Ltd., Liupanshui 553600, China
- \* Correspondence: 3774@cumt.edu.cn; Tel.: +86-139-1204-1768

**Abstract:** Elastic modulus (E) is a key parameter in predicting the ability of a material to withstand pressure and plays a critical role in the design of rock engineering projects. E has broad applications in the stability of structures in mining, petroleum, geotechnical engineering, etc. E can be determined directly by conducting laboratory tests, which are time consuming, and require high-quality core samples and costly modern instruments. Thus, devising an indirect estimation method of E has promising prospects. In this study, six novel machine learning (ML)-based intelligent regression models, namely, light gradient boosting machine (LightGBM), support vector machine (SVM), Catboost, gradient boosted tree regressor (GBRT), random forest (RF), and extreme gradient boosting (XGBoost), were developed to predict the impacts of four input parameters, namely, wet density ( $\rho_{\text{wet}}$ ) in  $\text{gm}/\text{cm}^3$ , moisture (%), dry density ( $\rho_d$ ) in  $\text{gm}/\text{cm}^3$ , and Brazilian tensile strength (BTS) in MPa on output E (GPa). The associated strengths of every input and output were systematically measured employing a series of fundamental statistical investigation tools to categorize the most dominant and important input parameters. The actual dataset of E was split as 70% for the training and 30% for the testing for each model. In order to enhance the performance of each developed model, an iterative 5-fold cross-validation method was used. Therefore, based on the results of the study, the XGBoost model outperformed the other developed models with a higher accuracy, coefficient of determination ( $R^2 = 0.999$ ), mean absolute error (MAE = 0.0015), mean square error (MSE = 0.0008), root mean square error (RMSE = 0.0089), and a20-index = 0.996 of the test data. In addition, GBRT and RF have also shown high accuracy in predicting E with  $R^2$  values of 0.988 and 0.989, respectively, but they can be used conditionally. Based on sensitivity analysis, all parameters were positively correlated, while BTS was the most influential parameter in predicting E. Using an ML-based intelligent approach, this study was able to provide alternative elucidations for predicting E with appropriate accuracy and run time at Thar coalfield, Pakistan.

**Keywords:** elastic modulus; K-fold cross-validation; mining; rock engineering; sensitivity analysis; XGBoost



**Citation:** Shahani, N.M.; Zheng, X.; Guo, X.; Wei, X. Machine Learning-Based Intelligent Prediction of Elastic Modulus of Rocks at Thar Coalfield. *Sustainability* **2022**, *14*, 3689. <https://doi.org/10.3390/su14063689>

**Academic Editors:**  
Mahdi Hasanipanah, Danial Jahed Armaghani and Jian Zhou

Received: 11 February 2022

Accepted: 13 March 2022

Published: 21 March 2022

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



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

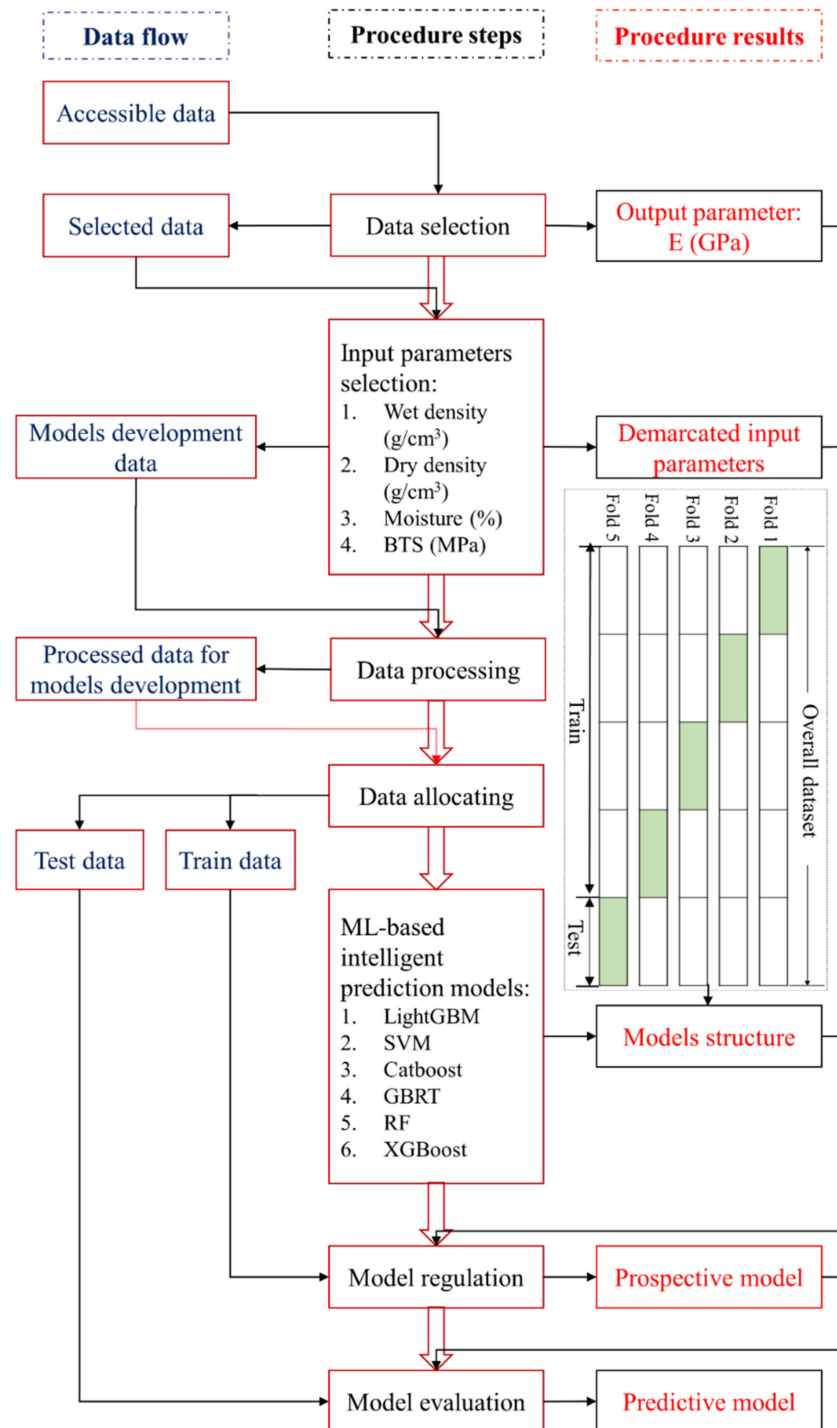
Elastic modulus (E) is a key parameter in predicting the ability of a material to withstand pressure and plays a critical role in the design process of rock-related projects. E has broad applications in the stability of structures in mining, petroleum, geotechnical engineering, etc. Accurate estimation of deformation properties of rocks, such as E, is very important for the design process of any underground rock excavation project. Intelligent indirect techniques for designing and excavating underground structures make use of a limited amount of data for design, saving time and money while ensuring the stability of the structures. This study has economic and even social implications, which

are integral elements of sustainability. Moreover, this paper aims to determine the stability of underground mine excavation, which may otherwise result in a disturbed overlying aquifer and earth surface profile, adversely affecting the environment.  $E$  provides insight into the magnitude and characteristics of the rock mass deformation due to changes in the stress field. Deformation and behavior of different types of rocks have been examined by different scholars [1–4]. Usually, there are two common methods, namely, direct (destructive) and indirect (non-destructive), to calculate the strength and deformation of rocks. Based on the principles suggested by ISRM (International Society for Rock Mechanics) and the ASTM (American Society for Testing Materials), direct evaluation of  $E$  in the laboratory is a complex, laborious, and costly process. Simultaneously, in the case of fragile, internally broken, thin, and highly foliated rocks, the preparation of a sample is very challenging [5]. Therefore, attention should be given to evaluate  $E$  indirectly by the use of rock index tests.

Several authors have developed prediction frameworks to overcome these limitations by using machine learning (ML)-based intelligent approaches such as multiple regression analysis (MRA), artificial neural network (ANN), and other ML methods [6–21]. Advances in ML have so far been driven by the development of new learning algorithms and theories, as well as by the continued explosion of online data and inexpensive computing [22]. Similarly, Waqas et al. used linear and nonlinear regression, regularization and ANFIS (using a neuro-fuzzy inference system) to predict the dynamic  $E$  of thermally treated sedimentary rocks [23]. Abdi et al. developed ANN and MRA (linear) models, including porosity (%), dry density ( $\gamma_d$ ) ( $\text{g}/\text{cm}^3$ ), P-wave velocity ( $V_p$ ) ( $\text{km}/\text{s}$ ), and water absorption ( $A_b$ ) (%) as input features to predict the rock  $E$ . According to their results, the ANN model showed high accuracy in predicting  $E$  compared to the MRA [10]. Ghasemi et al. evaluated the UCS and  $E$  of carbonate rocks by developing a model tree-based approach. According to their findings, the applied method revealed highly accurate results [24]. Shahani et al. developed a first-time XGBoost regression model in combination with MLR and ANN for predicting  $E$  of intact sedimentary rock and achieved high accuracy in their results [25]. Ceryan applied the minimax probability machine regression (MPMR), relevance vector machine (RVM), and generalized regression neural network (GRNN) models to predict the  $E$  of weathered igneous rocks [26]. Umrao et al. determined strength and  $E$  of heterogeneous sedimentary rocks using ANFIS based on porosity,  $V_p$ , and density. Thus, the proposed ANFIS models showed superb predictability [27]. Davarpanah et al. established robust correlations between static and dynamic deformation properties of different rock types by proposing linear and nonlinear relationships [28]. Aboutaleb et al. conducted non-destructive experiments with SRA (simple regression analysis), MRA, ANN, and SVR (support vector regression) and found that ANN and SVR models were more accurate in predicting dynamic  $E$  [29]. Mahmoud et al. employed an ANN model for predicting sandstone  $E$ . In that study, 409 datasets were used for training and 183 datasets were used for model testing. The established ANN model exposed highly accurate results (coefficient of determination ( $R^2$ ) = 0.999) and the lowest mean absolute percentage error ((AAPE) = 0.98) in predicting  $E$  [30]. Roy et al. used ANN, ANFIS, and multiple regression (MR) to predict the  $E$  of  $\text{CO}_2$  saturated coals. Thus, ANN and ANFIS outperformed the MR models [31]. Armaghani et al. predicted  $E$  of 45 main range granite samples by applying the ANFIS model in comparison with MRA and ANN. Based on their results, ANFIS proved to be an ideal model against MRA and ANN [32]. Singh et al. proposed an ANFIS framework for predicting  $E$  of rocks [33]. Köken predicted the deformation properties of rocks, i.e., tangential  $E$  ( $E_{ti}$ ) and tangential Poisson's ratio ( $\nu_{ti}$ ) of coal-bedded sandstones located in the Zonguldak Hard Coal Basin (ZHB), northwestern Turkey, using various statistical and soft computing methods such as different regression and ANN evaluations including the physicomaterial, mineralogical, and textural properties of the rocks. According to this analysis, the remarkable results were that the mineralogical characteristics of the rock have a significant influence on the deformation properties. In addition to comparative analysis, ANN was considered as a more effective tool than regression analysis in predicting  $E_{ti}$

and  $v_{ti}$  of coal-bed sandstones [34]. Yesiloglu-Gultekin et al. used the different ML-based regression models such as NLMR, ANN, and ANFIS, and 137 datasets using unit weight, porosity, and sonic velocity to indirectly determine E of basalt. Based on the results and comparisons of various performance matrices such as  $R^2$ , RMSE, VAF, and a20-index, ANN was successful in predicting E over NLMR and ANFIS [35]. Rashid et al. used non-destructive tests, i.e., MLR and ANN, to estimate the Q-factor and E for intact sandstone samples collected from the Salt Range region of Pakistan. The ANN model predicted Q-factor ( $R^2 = 0.86$ ) and E ( $R^2 = 0.91$ ) more accurately than MLR regression for Q-factor ( $R^2 = 0.30$ ) and E ( $R^2 = 0.36$ ) [36]. E was predicted using RF by Matin et al. For comparison, multivariate regression (MVR) and generalized regression neural network (GRNN) were used for the prediction of E. The input  $V_p$ - $R_n$  was used for E. According to their results, RF yielded more satisfactory conclusions than MVR and GRNN [37]. Cao et al. used an extreme gradient boosting (XGBoost) integrated with the firefly algorithm (FA) model for predicting E. consequently, the proposed model was appropriate for predicting E [17]. Yang et al. developed the Bayesian model to predict the E of intact granite rocks; thus, the model performed with satisfactory predicted results [38]. Ren et al. developed several ML algorithms, namely, k-nearest neighbors (KNN), naive Bayes, RF, ANN, and SVM, to predict rock compressive strength by ANN and SVM with high accuracy [39]. Ge et al. determined rock joint shear failures using scanning and AI techniques. Thus, the developed SVM and BPNN were considered as sound determination methods [40]. Xu et al. developed several ML algorithms, namely, SVR, nearest neighbor regression (NNR), Bayesian ridge regression (BRR), RF, and gradient tree boosting regression (GTBR), to predict microparameters of rocks by RF with high accuracy [41].

Based on the above literature and the limitations of the conventional predictive methods, a single model has low robustness, cannot achieve ideal solutions for all complex situations, and its performance varies with the input features. Therefore, authors have endeavored to use ML-based intelligent models that integrate multiple models to overcome the drawbacks of individual models and play a key role in determining the accuracy of the corresponding data for tests performed in the laboratory. However, there are few studies in predicting E. In addition, there are no comprehensive studies on the selection and application of such models in E prediction. To address this gap, this study developed six models based on an intelligent prediction approach, namely, light gradient boosting machine (LightGBM), support vector machine (SVM), Catboost, gradient boosted tree regressor (GBRT), random forest (RF), and extreme gradient boosting (XGBoost) to predict E, including wet density ( $\rho_{wet}$ ) in  $gm/cm^3$ , moisture in %, dry density ( $\rho_d$ ) in  $gm/cm^3$ , and Brazilian tensile strength (BTS) in MPa as input features under intricate and unsteady engineering situations. Next, 70% of the actual dataset of 106 is used for training and 30% for testing each model. To enhance the performance of the developed models, a repetitive 5-fold cross-validation approach is used. Intelligent prediction of E of sedimentary rocks from Block-IX of Thar coalfield has been applied for the first time. To the best of the author's knowledge, application of intelligent prediction techniques in this scenario is lacking. Figure 1 depicts a systematic ML-based intelligent approach for predicting E.



**Figure 1.** Systematic ML-based intelligent approach for predicting E.

## 2. A Brief Summary of the Study Area

The Thar coalfield is located in Sindh Province of Pakistan and is the seventh largest coal mine around the world in terms of coal potential [42]. Thar coal is classified as 175.5 billion tons of lignite, which can be used for fuel and power generation. The Thar coalfield is distributed in twelve different blocks as shown in Figure 2. The Thar coalfield is enclosed by dune sand that spreads to a normal distance of 80 m and rests upon an essential stand in the eastern portion of the desert. The general stratigraphic arrangement in the Thar coalfield encompasses the Basement Compound, coal posture Bara Formation, alluvial deposits, and dune sand. For coal mining in the region, both open-pit and

underground mining methods can be preferred. Particularly, Sindh Engro Coal Mining Company (SECMC) has fully developed Block-II of the twelve blocks using an open-pit mining method, whereas Block-1 is under a development stage by Sino-Sindh Resources Ltd. (SSRL) in partnership with China and the Sindh government of Pakistan. Block-IX has been recommended for the underground mining method. The thickness of the coal seam of Block-IX of the Thar coalfield is approximately 12 m, the dip angle is  $0^\circ$  to  $7^\circ$ , and the top-bottom plate is siltstone–claystone to claystone. Shahani et al. proposed the use of the mechanized longwall caving (LTCC) method at Block-IX of the Thar coalfield in Pakistan for the first time [42,43]. In addition, Shahani et al. developed various gradient boosting machine learning algorithms to predict UCS of sedimentary rocks of Block-IX of the Thar coalfield [44]. Similarly, correct determination of the mechanical properties of Block-IX of the Thar coalfield, particularly E, plays an important role in fully understanding the behavior of the roof and ground prior to mining operations.

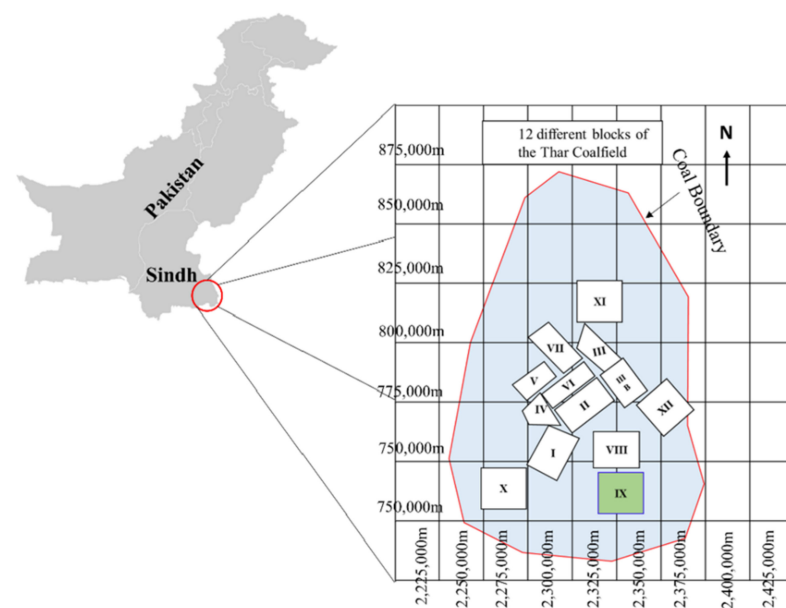
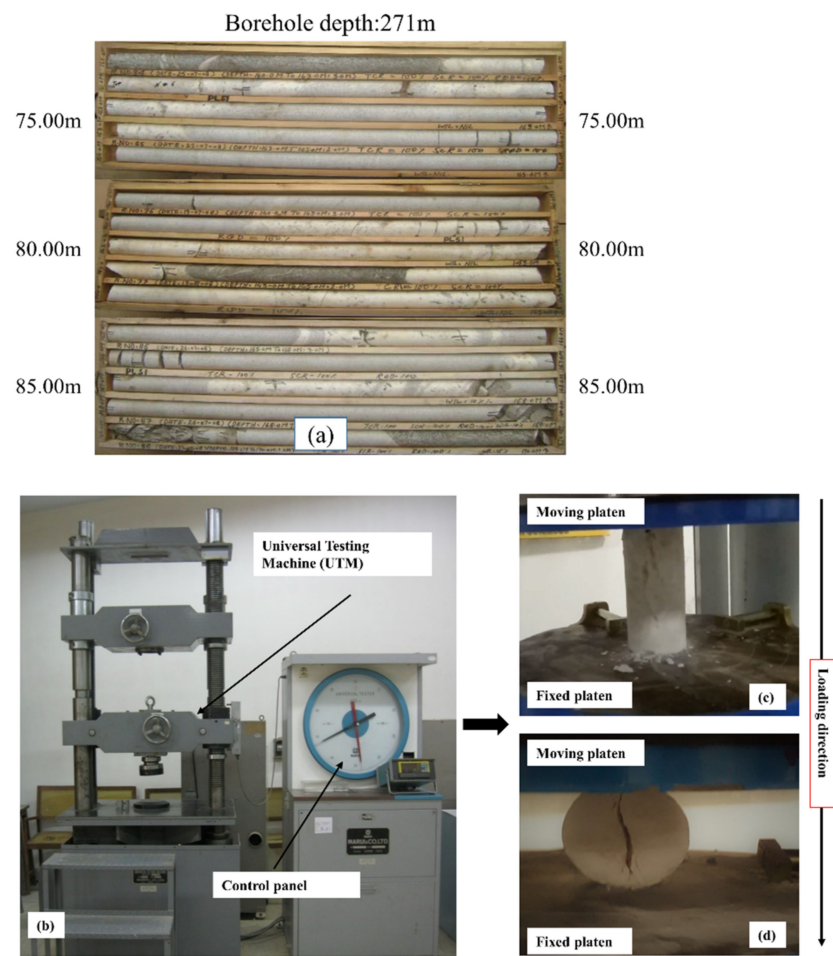


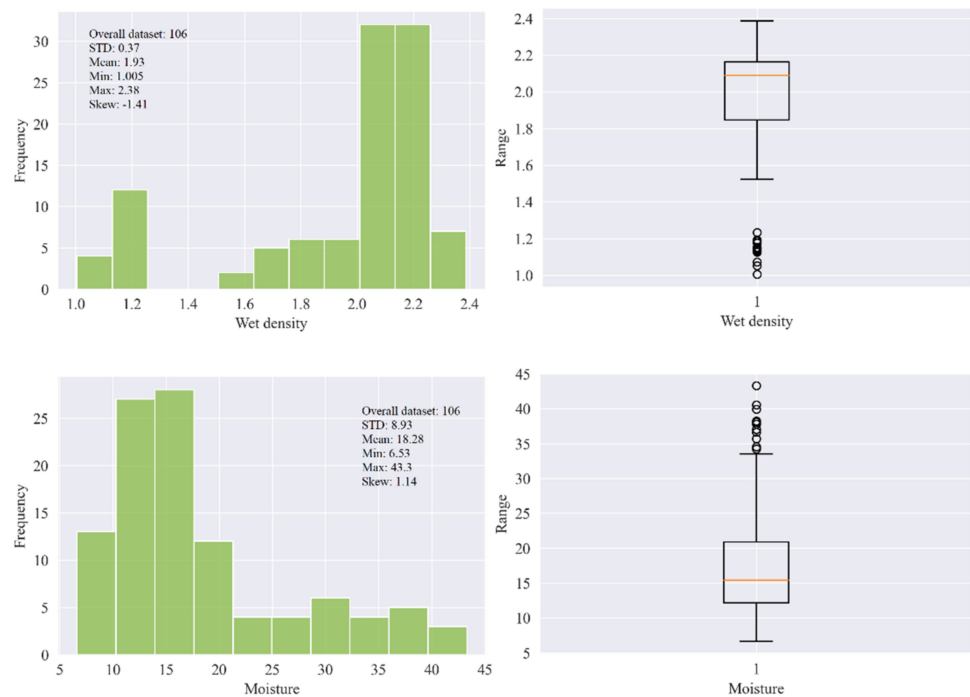
Figure 2. Location map of the study area.

### 3. Data Curation

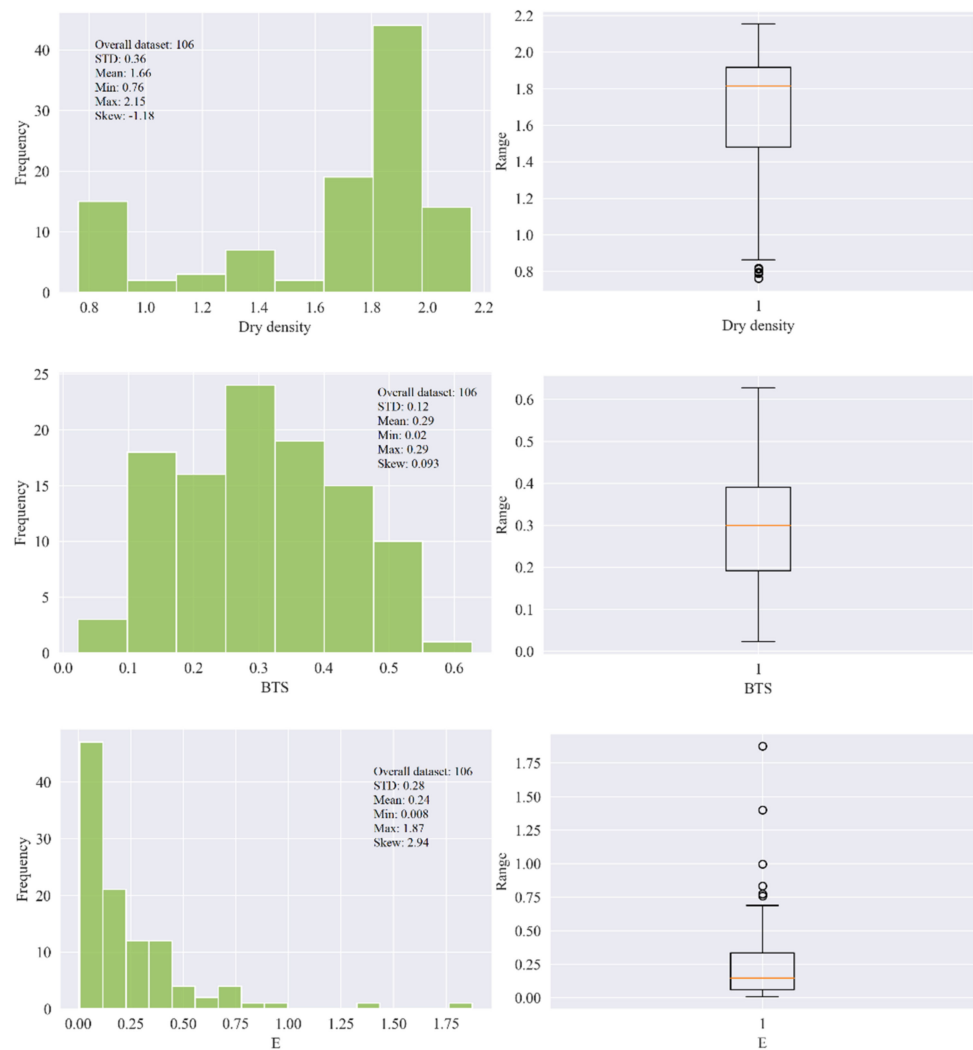
In this research, 106 samples of soft sedimentary rocks, i.e., siltstone, claystone, and sandstone were collected from Block-IX of the Thar coalfield, as shown Figure 2, with the location map in the green. Then, the rock samples were prepared and partitioned according to the principles suggested by ISRM [45] and the ASTM [46] to maintain the same core size, and geological and geometric characteristics. In the laboratory of the Mining Engineering Department of Mehran University of Engineering and Technology (MUET), the experimental work was conducted on the studied rock samples to determine the physical and mechanical properties such as wet density ( $\rho_{wet}$ ) in  $g/cm^3$ , moisture (%), dry density ( $\rho_d$ ) in  $g/cm^3$ , Brazilian tensile strength (BTS) in (MPa), and elastic modulus (E) in (GPa). Figure 3 shows (a) collected core samples, (b) universal testing machine (UTM), (c) deformed core sample under compression for E test, and (d) deformed core sample for BTS test. The purpose of the UCS test was conducted on the standard core samples of NX size 54 mm in diameter with an applied load of 0.5 MPa/s using UTM according to the recommended ISRM standard to find the E of the rocks. Similarly, in order to find the tensile strength of the rock samples indirectly, we performed the Brazilian test using UTM. Figure 4 illustrates the statistical distribution of the input features and output in the original dataset used in this study. In Figure 4, the legend of boxplots can be explained as:  $\square$  25~75%,  $\perp$  Range within 1.5 IQR,  $-$  Median line, and  $\circ$  Outliers.



**Figure 3.** (a) Rock core samples for test, (b) uniaxial testing machine, (c) deformed rock core specimen under compression, and (d) deformed core sample for BTS test.

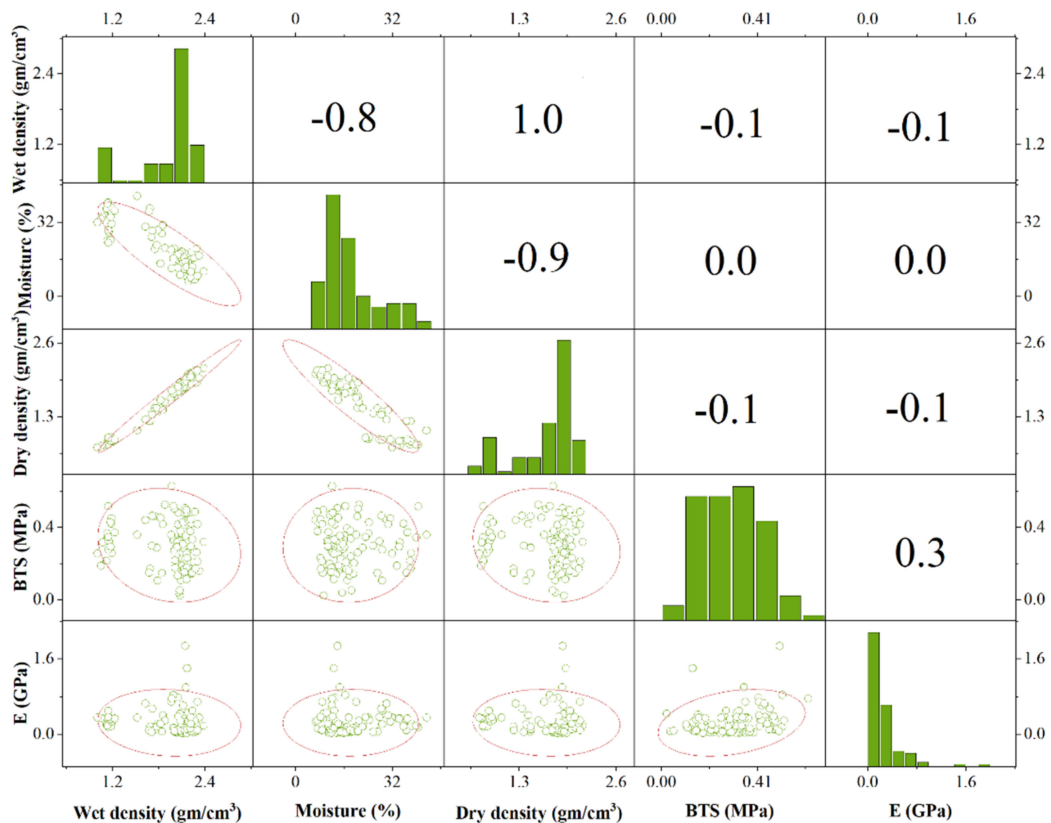


**Figure 4.** Cont.



**Figure 4.** The statistical distribution of the input features and output in the original dataset.

In order to visualize the original dataset of E, the seaborn module in Python was employed in this study, and Figure 5 demonstrates the pairwise correlation matrix and distribution of different input features and output E. It can be seen that BTS is moderately correlated to the E, whereas  $\rho_{wet}$  and  $\rho_d$  are negatively correlated to the E. Moisture representation does not correlate with E. It is worth mentioning that each feature cannot be well correlated with E independently, so all features are evaluated together to predict E.

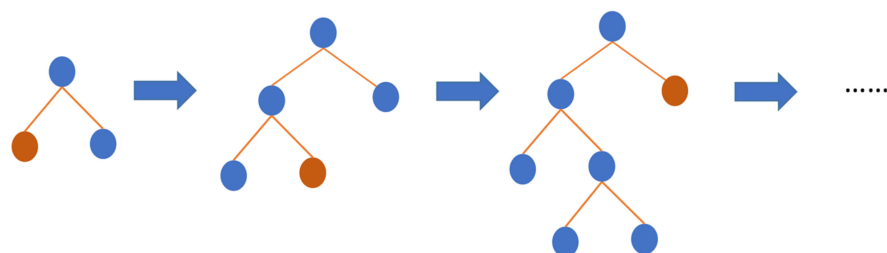


**Figure 5.** Pairwise correlation matrix and distribution of different input features and output E.

## 4. Developing ML-Based Intelligent Prediction Models

### 4.1. Light Gradient Boosting Machine

Light gradient boosting machine abbreviated as LightGBM, an open-source gradient boosting ML model from Microsoft, uses decision trees as the base training algorithm [47]. LightGBM puts continuous buckets of elemental values into separate bins with greater adeptness and a fast speed of training. It uses a histogram-based algorithm [48,49] to improve the learning phase, reduce consumption of memory, and integrate updated communication networks to enhance the regularity of training and is known as a parallel voting decision tree ML algorithm. The data for learning were partitioned into several trees, and local voting techniques were executed in each iteration to select top-k elements and gain globing voting techniques. As shown in Figure 6, LightGBM operates the leaf-wise approach to identify the leaf with the maximum splitter gain. LightGBM is best adopted for regression, classification, sorting, and several ML schemes. It builds a more complex tree than the level-wise distribution method through the leaf-wise distribution method, which can be considered as the main component of the execution algorithm with greater effectiveness. For all that, it can cause overfitting; however, by using the maximum depth element in LightGBM, it can be disabled.



**Figure 6.** The general structure of LightGBM.



LightGBM [47] is a widespread library for performing gradient boosting, with some modifications intended. The implementation of gradient boosting is mainly focused on algorithms for building a computational system. The library includes tenfold training hyperparameters to validate the implementation of the framework in different scenarios. The implementation of LightGBM also demonstrates advanced capabilities on CPUs and GPUs, which can work like gradient boosting with multifold integrations, comprising column randomization, bootstrap subsampling, and so on. The main features of LightGBM are gradient-based one-sided sampling and unique attribute bundling. Gradient-based one-sided sampling is a sub-sampling technique used to construct the base tree of learning data as an ensemble. In the AdaBoost ML algorithm, the purpose of this technique is to increase the significance of samples with greater likelihood that are connected with samples with higher gradients. When gradient-based one-sided sampling is executed, the base learner's learning data are articulated based on the top portion of samples with greater gradients (a) plus the portion of arbitrary orders (b) recouped from samples with lower gradients. To compensate for changes in measurement propagation, samples from the lesser gradient class are organized together and weighted by  $(1 - x)/y$ , and at the same time, computing the data gain. In contrast, the unique attribute bundling technique accrues meager elements into an individual element. This can be ended in the absence of impeding any information when these elements do not contain a non-zero number of coincidences. Both mechanisms predict a gain in the complementary learning rate.

#### 4.2. Support Vector Machine

In 1997, Vapnik et al. originally proposed support vector machines (SVMs), which are a type of supervised learning [50]. SVMs can be widely used for regression analysis and for classification using hyperplane classifiers. The ideal hyperplane enhances the boundary between the two classes in which the support vector is positioned [51]. The SVM utilizes a high-extent feature space to develop the forecast function by proposing kernel functions and Vapnik's  $\varepsilon$ -insensitive loss function [52].

For a dataset  $P = \{(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)\}$ , where  $x_i \in R^n$  is the input and  $y_i \in R^n$  is the output, the SVM employs a kernel function to plot the nonlinear input data in a high-extent feature space, and attempts to discover the best hyperplane to disperse them. This permits the narration of the original input to the output by a linear regression function [53–55] characterized as follows in Equation (1).

$$f(x) = M_v \cdot \varphi(x) + l_b \quad (1)$$

where  $\varphi(x)$  shows the kernel function,  $M_v$  and  $l_b$  show the weight vector and bias term, respectively. In order to obtain  $M_v$  and  $l_b$ , the cost function proposed by Cortes and Vapnik [56] is required to be reduced as follows in Equation (2).

$$\text{cost function} = \frac{1}{2} M_v^2 + C \sum_{i=1}^k (\xi_i^- + \xi_i^+) \quad (2)$$

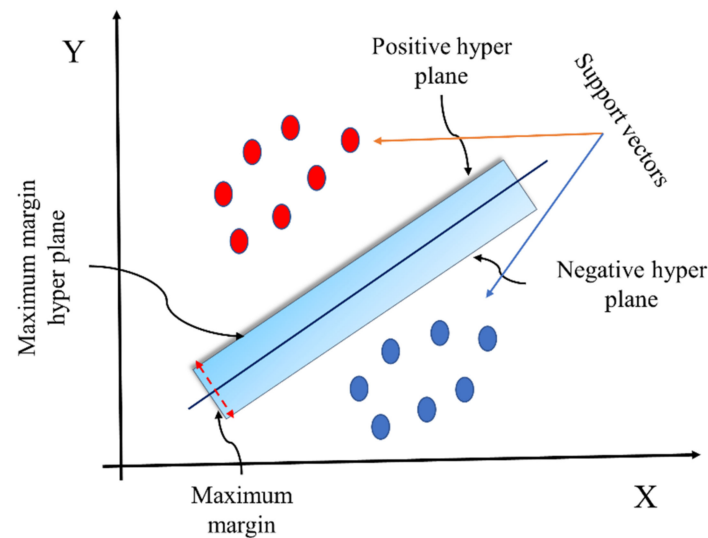
$$\text{Subject to : } \begin{cases} y_i - (M_v \cdot \varphi(x_1) + l_b) \leq \varepsilon_0 + \xi_i^+ \\ (M_v \cdot \varphi(x_1) + l_b) - y_i \leq \varepsilon_0 + \xi_i^- \\ \xi_i^-, \xi_i^+ \geq 0, i = 1, 2, \dots, n \end{cases}$$

When converted to the dual space using the Lagrange multiplier method, Equation (2) can be reduced to obtain the following solution in Equation (3).

$$f(x) = \sum_{i=1}^n (\infty_i - \infty'_i) \varphi(x_i, x_j) + l_b \quad (3)$$

where  $\infty_i$  and  $\infty'_i$  are Lagrange multipliers with  $0 \leq \infty_i$  and  $\infty'_i \leq C$ , though  $\varphi(x_i, x_j)$  is the kernel function. The choice of the latter is important to the accomplishment of SVR. A

large number of kernel functions have been studied in SVM, such as linear, polynomial, sigmoid, gaussian, radial basis, and exponential radial basis [54]. Figure 7 illustrates the basic structure of the SVM model.



**Figure 7.** The basic structure of the SVM model.

#### 4.3. Catboost

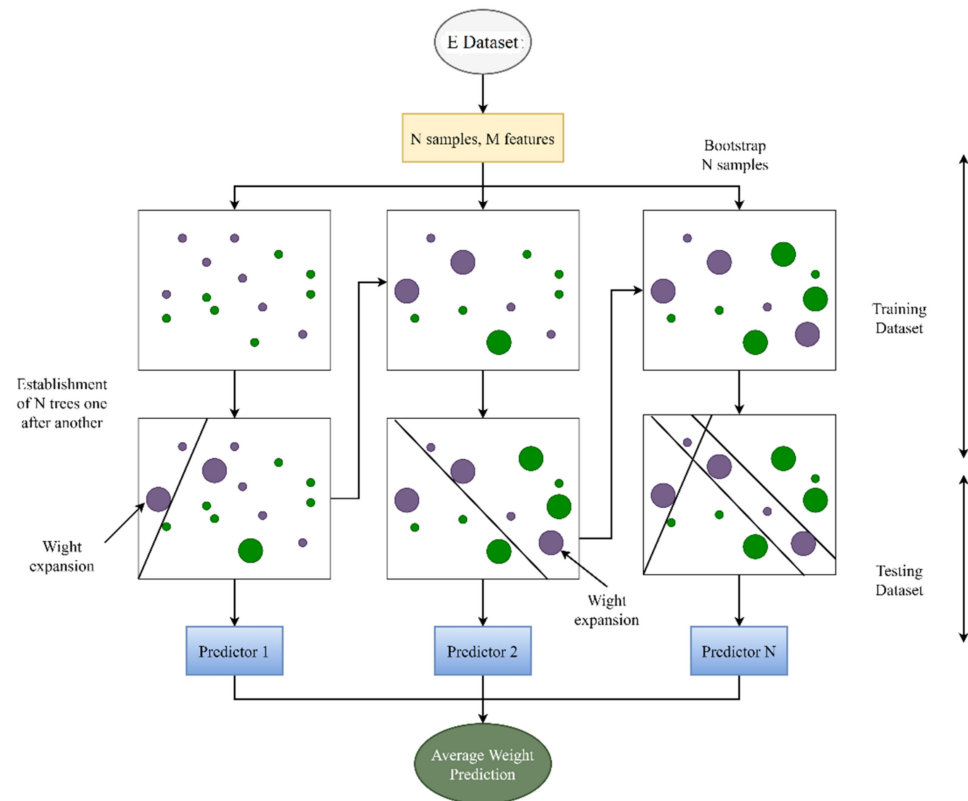
Catboost is a gradient boosting algorithm recently inherited by Dorogush et al. [57]. Catboost solves the complex problem of regression and classification simultaneously and is publicly available in an open-source multi-platform gradient boosting library [57,58]. In the Catboost algorithm, the decision tree is used as the underlying weak learner and the gradient boosting is successively fitted to the decision tree. To improve the implementation of the Catboost algorithm and to avoid overfitting, an inconsistent arrangement of gradient learning information is used [57].

The purpose of the Catboost algorithm is to reduce the predictive movement that occurs in the learning phase. Propagation movement is the deletion of  $F(y) | (y_i)$  in the case that  $y_i$  is the learning sample, related to  $F(y) | (y_i)$  of the test sample  $y$ . In the learning phase, gradient boosting uses the same samples to compute the gradient and the model for lowering that gradient. The Catboost's concept is to initiate  $j \dots n$ , the framework underlying the repetition of separate  $P$  enhancing. The  $m$ th recurrence's  $i$ th framework is learned from the permutation of the initial  $i$ th sample and is suitable for computing the  $j + 1$  sample's gradient of the  $p + 1$  recurrence. Subsequently, in order not to be limited by the starting arbitrary permutation, the technique employs an arbitrary permutation of  $s$  reciprocals. Each repetition constructs a distinguished framework that achieves all combinations and frameworks. Symmetric trees are used as the basis of the framework. These trees are prolonged by using the same partitioning criterion so that all leaf nodes grow level-wise.

In the Catboost algorithm, the mechanism proposed is to compute the identical up-to-date characters as the ones imitated when the network was built. Thus, for any specified samples' permutation, data samples  $<i$  are utilized to computing the character values for each sample  $i$ . Then, different combinations are implemented, and the character values obtained for each sample are averaged. Catboost is a large-scale comprehensive library consisting of several elements such as GPU learning, standard boosting, and including fivefold hyperparameter optimization to amend to various practical examination situations. Standard gradient boosting is to be considered as part of the Catboost algorithm also. Figure 8 shows an explanation of the Catboost algorithm.

It is very important to note that the Catboost algorithm's training ability is managed by its framework hyperparameters, i.e., iterations number, rate of learning, maximum depth,

etc. Determining the optimal hyperparameters of a model is a challenging, laborious, and tedious task that depends on the user's skills and expertise.



**Figure 8.** Explanation of the Catboost.

#### 4.4. Gradient Boosted Regressor Tree

The gradient boosted regressor tree (GBRT) regression integrates the weak learner, i.e., the learner algorithms with average performance compared to random algorithms into a robust learner with an iterative method [59]. Contradictory of the bagging method, the boosted algorithm continuously generates the underlying framework. The soundness of the predictive framework is improved by prioritizing this hard-to-evaluate learning information to generate several frameworks in a series. In the boosting algorithm, underlying frameworks that were previously not suitable for estimation are frequently established in the training dataset compared to those frameworks that have been accurately evaluated. Each complementary underlying framework is designed to correct inaccuracies arising from its prior underlying framework. The occurrence of the boosting mechanism comes from reaction of Schapire's feedback to Kearns' investigation [60,61] (Kearns): Is the aggregation of weak learners a substitute for distinguishing strong learners? Weak learners are explained as the algorithms that work well compared to random approximation; strong underlying frameworks are a more realistic classification or regression algorithms that are incoherent with their effective counterparts to the problem. The reaction to such inquiry is extremely noteworthy. Assessments of weak frameworks tend to be less challenging than strong frameworks, and Schapire's establishment of a "yes" response to Kearns' inquiry, as evidenced by the combination of several weak frameworks into an upgraded and independent sound framework. The key dissimilarity between the boosting and bagging mechanism is that in the boosting approach, the training dataset is analytically investigated in order to predict the most appropriate instructions for each subsequent framework. In every phase of training, the modified propagation is dependent on the inaccuracies raised by the previous frameworks. On the contrary, in the bagging mechanism, every trial is constantly specified to produce a training dataset, and for the boosting mechanism, the vagueness of specifying

an independent trial is conflicting. Trials which were erroneously assessed were more likely to set higher weights. Accordingly, each newly evolved framework underscores trials that are inaccurately assessed by subsequent frameworks.

Boosting assembles the auxiliary frameworks that decrease a specific loss function averaged over the learning dataset, i.e., the MAE or the MSE. The loss function computes the total number of predicted values that differ from the investigated values. The advanced staged modeling method is one of the assessed elucidations to the problem. This modeling method continuously attaches the new underlying framework without substituting the coefficients and specifications of the previously connected model. Referring to the regression model, the boosting mechanism is a “function gradient descent” configuration. Functional gradient descent is an optimization mechanism that minimizes the loss function by connecting the underlying framework to each stage to reduce the loss function by a certain amount. Figure 9 demonstrates the schematic diagram of GBTR (after [62,63]).

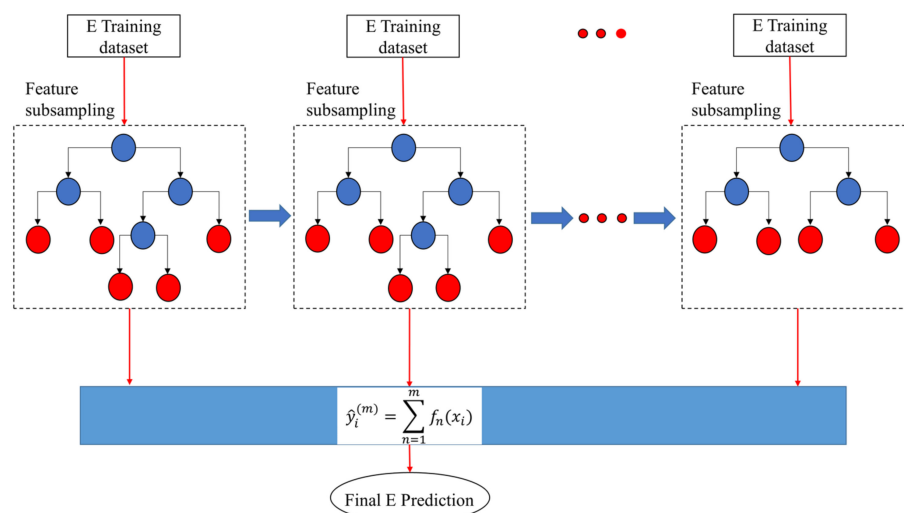


Figure 9. The schematic diagram of the GBRT model.

Friedman recommended improving the gradient boosting regression model by using pre-established regression trees for the underlying framework. The improved framework amplifies the performance of Friedman’s model [64]. For predicting E, the improved version of gradient boosted regression was used. Considering that the number of leaves is  $l$ , each tree divides the input space into  $l$  independent territory  $T_{1p}, T_{2p} \dots \dots T_{lp}$  and a perpetual value  $k_{lp}$  is predicted for territory  $T_{lp}$ . Equation (4) represents the gradient boosting regression tree as follows:

$$f_p(a) = \sum_{l=1}^L k_{lp}F(a \in T_{lp}) \tag{4}$$

where  $F(a \in T_{lp}) = \begin{cases} 1, & \text{if } a \in T_p \\ 0, & \text{otherwise} \end{cases}$ .

By using a regression tree to recover  $f_p(a)$  in the generic gradient boosting mechanism, the framework gradient descent stage size and updating equation are given by Equations (5) and (6), respectively.

$$f_p(a) = f_{p-1}(a) + \rho_p g_p(a) \tag{5}$$

$$\rho_p = \operatorname{argmin}_{\rho} \sum_{l=1}^L M(b_l, f_{p-1}(a_l) + \rho g_p(a_l)) \tag{6}$$

Hence, Equations (5) and (6) fit as Equations (7) and (8).

$$f_p(a) = f_{p-1}(a) + \sum_{l=1}^L \rho_p k_{lp}F(a \in T_{lp}) \tag{7}$$

$$\rho_p = \operatorname{argmin}_\rho \sum_{l=1}^L M(b_l, f_{p-1}(a_l)) + \sum_{l=1}^L \rho_p k_{lp} F(a \in T_{lp}) \quad (8)$$

By applying a discrete ideal  $\rho_{lp}$  for each territory  $T_{lp}$   $k_{lp}$  should be separated. The simplified framework Equations (9) and (10) are given by

$$f_p(a) = f_{p-1}(a) + \sum_{l=1}^L \rho_{lp} F(a \in T_{lp}) \quad (9)$$

$$\rho_p = \operatorname{argmin}_\rho \sum_{l=1}^L M(b_l, f_{p-1}(a_l)) + \sum_{l=1}^L \rho_{lp} F(a \in T_{lp}) \quad (10)$$

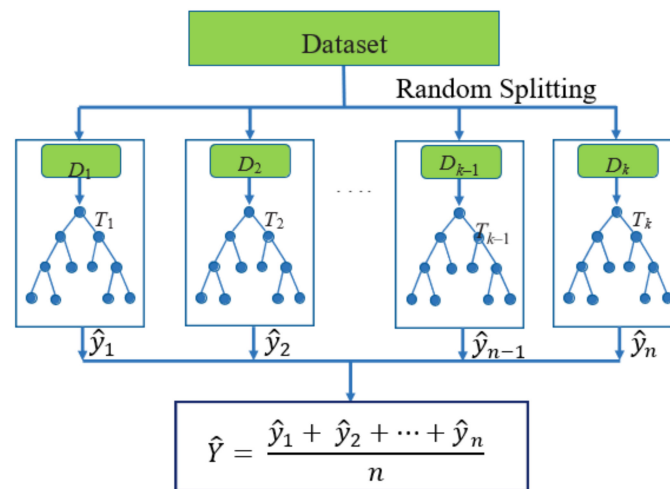
The overfitting of the framework can be limited by managing the number of iterations of gradient boosting or more competently by evaluating the degree of benefit of each tree by  $J \in (0, 1)$ . Thus, the simplified model is given by Equation (11).

$$f_p(a) = f_{p-1}(a) + L \sum_{l=1}^L \rho_{lp} F(a \in T_{lp}) \quad (11)$$

#### 4.5. Random Forest

In 2001, Breiman originally proposed random forest (RF), a type of ensemble machine learning algorithm [65]. RF can be widely used for regression analysis and classification. RF is a state-of-the-art approach to bootstrap aggregating or bagging. RF perceives a one-of-a-kind relationship of model embodiment and predictive accuracy among alternative recognized AI computing [66].

To calculate the performance of the model, RF of 100 trees with a range of default settings was chosen for the study. Figure 10 shows the basic structure of the RF model.

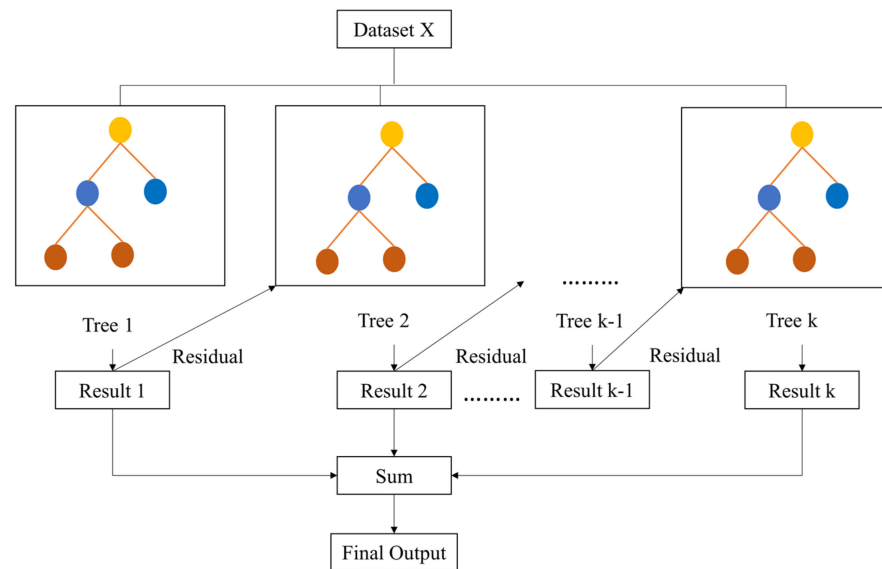


**Figure 10.** The basic structures of RF model.

#### 4.6. Extreme Gradient Boosting

Extreme gradient boosting (XGBoost) is an important type of ensemble learning algorithm in ML approaches [67]. XGBoost consists of usual regression and classification trees with the addition of analytical boosting methods. The boosting method improves the accurate framework assessment by constructing different trees as alternatives to develop an addressed tree and then connecting them to estimate a systematic predictive algorithm [68]. It instigates the tree by consecutively holding the residuals of the historical trees as the effect of the resultant tree. Because of this, the resultant tree builds a full prediction by generating errors in the past tree. In the loss function reduction stage, the consecutive framework structural relationship can be subdivided into gradient descent types, which advances the forecast by connecting a supplementary tree at every stage to lessen the depletion [69]. Tree development stops at the time of the most unprecedented tree's predetermined number is obtained, or at the time of the training stage error when it cannot be amplified to the predicted number of sequential trees. By attaching an arbitrary survey, the performance

timeliness and estimation accuracy of gradient boosting can be greatly improved. In particular, for all symmetric trees, a random subsample of training information from the whole training dataset is considered, without substitutions. This arbitrarily described subsample replaces the whole sample, which is later used to adapt the tree and is identified as an improved framework. XGBoost is a state-of-the-art rearranged gradient boosting ML algorithm that manages and implements the latest prediction demonstrations [49]. The loss function's subsequent assessment is used in the XGBoost and is fast and rapidly matched to the usual gradient boosting algorithms. XGBoost has widely been used to mine the features of gene coupling. Figure 11 shows the general structure of XGBoost models.



**Figure 11.** The general structure of XGBoost model.

Consider  $\bar{u}_i$  is the predicted outcome of the  $i$ th data, where the feature vector is  $V_i$ ;  $E$  denotes the number of estimators and for every estimator  $f_k$  (with  $k$  from 1 to  $E$ ) analogous to the analysis of a single tree.  $u_i^0$  describes the initial hypothesis and is the mean of the examined features in the data for learning. Equation (12) executes different extension functions to predict the results.

$$\bar{u}_i = u_i^0 + \eta \sum_{k=1}^E f_k(V_i) \quad (12)$$

Additionally, the  $\eta$  parameter is the learning rate, which is contiguous to the implementation of the improved model to enhance the model, perform rhythmically when connecting the latest trees, and combat overfitting.

With respect to Equation (12), the  $k$ th character is attached to the model in the  $k$ th state, the  $k$ th prediction  $u_i^{-k}$  is realized by the prediction  $u_i^{-(k-1)}$  of the previous state, and the additional  $k$ th character augmentation  $f_k$  is described as in Equation (13).

$$u_i^{-k} = u_i^{-(k-1)} + \eta f_k \quad (13)$$

where  $f_k$  denotes the weight of the leaves established by decreasing the  $k$ th tree's objective function signified by Equation (14).

$$obj = \gamma N + \sum_{a=1}^N [U_a \omega_a + \frac{1}{2} (V_a + \lambda) \omega_a^2] \quad (14)$$

where  $N$  represents the  $k$ th tree's leaves,  $\omega_a$  denotes the leaves' weights from 1 to  $N$ .  $\gamma$  and  $\lambda$  are the regularity attributes to achieve anatomical consistency to avoid the model's

overfitting. The  $V_a$  and  $U_a$  parameters are the sets of whole data connected with the prior data's leaf and the gradient of the posterior loss function, respectively.

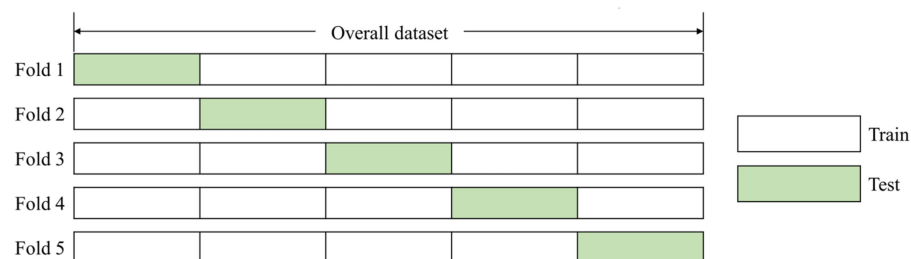
In the process of building the  $k$ th tree, individual leaves are partitioned into a different number of leaves. Equation (15) represents the dissection using the gain parameter. Consider that  $U_R$  and  $V_R$  describe inter-reliant right leaves and  $U_L$  and  $V_L$  are inter-reliant left leaves for divergence. At this point, the gain parameter is near to zero, which is traditionally considered as the benchmark for divergence.  $\gamma$  and  $\lambda$  are uniform features that affect the gain features, such as the gain parameter being reduced by a higher regularization parameter and thus avoiding leaf convolution. However, it reduces the adaptability of the framework to the training dataset.

$$gain = \frac{1}{2} \left[ \frac{U_L^2}{V_L + \lambda} + \frac{U_R^2}{V_R + \lambda} + \frac{(U_L + U_R)^2}{V_L + V_R + \lambda} \right] \quad (15)$$

XGBoost is a broadly adopted ML algorithm that brings together articulation and logical achievements of gradient boosting ML algorithms. A numerical value causes the problem with the prediction regression model. XGBoost can be accomplished in a timely manner in a probabilistic regression framework. The ensemble is built from a decision tree model. Ensembles are continuously connected trees that can be adjusted to predict imprecise models. These ensemble-type ML methods are called boosting. These frameworks are built by executing any random gradient descent optimization method with a unique loss function. When the model is executed, the gradient loss function is reduced, and so this technique is known as "gradient boosting". Compared with LightGBM, SVM, Catboost, GBRT, and RF, the XGBoost model performed well on the E dataset with the identical parameters  $n\_splits = 5$ ,  $n\_repeats = 3$ , and  $random\_state = 1$  (all remaining parameters were used as default parameters in Python). Although in this study,  $gbm\_param\_grid$  was further implemented in order to further improve the XGBoost model's performance.

#### 4.7. K-Fold Cross-Validation

The K-fold cross-validation is a technique employed to regulate the hyperparameters [70]. The technique accepts a search within a demarcated hyperparameters' range and describes the predictable outcomes leading to the best outcome for calculation criteria such as  $R^2$ , MAE, MSE and RMSE. In the scikit-learn Python programming language, K-fold Cross-Validation has been implemented to handle this approach. This method simply calculates the score of CV for all hyperparameters integrated with a specific range. In this study, a 5-fold iterated arbitrary arrangement practice was integrated into the CV command as illustrated in Figure 12. `GridSearchCV()` permits not only the calculation of the anticipated hyperparameters, but also the evaluation of the metric values to their anticipated results.



**Figure 12.** The diagram of 5-fold cross-validation used in this study.

#### 4.8. Models Performance Evaluation

To accurately and approximately evaluate the performance of ML-based intelligent models, different authors have used different estimation criteria, namely, coefficient of determination ( $R^2$ ) [71], mean absolute error (MAE), mean square error (MSE), root mean

square error (RMSE) [72], and a20-index [71]. Performance criteria are the main metrics used to assist in the highly accurate model evaluation, with the highest  $R^2$ , minimum MAE, MSE, RMSE, and appropriate a20-index. The following performance indices are employed to evaluate the performance of each model in E prediction.

$$R^2 = 1 - \frac{\sum_{i=1}^n (E_i - \hat{E}_i)}{E_i - \bar{E}_i} \quad (16)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^n |E_i - \hat{E}_i| \quad (17)$$

$$\text{MSE} = \frac{\sum_{i=1}^n (E_i - \hat{E}_i)^2}{N} \quad (18)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (E_i - \hat{E}_i)^2}{N}} \quad (19)$$

$$a20 - index = \frac{m20}{N} \quad (20)$$

where  $\bar{E}_i$  and  $\hat{E}_i$  are the mean values of the measured and predicted values of E, and  $E_i$  are measured values of E, respectively. m20 represents the datasets with a value of rate original/estimated values between 0.80 and 1.20 and N denotes the number of datasets.

## 5. Analysis of Results and Discussion

This study aims to examine the capability of various ML-based intelligent prediction models, namely, LightGBM, SVM, Catboost, GBRT, RF, and XGBoost, for predicting a substantial E using Python programming. In order to propose the most suitable prediction model to predict E, the selection of appropriate input features can be considered as one of the most important tasks. In this study, wet density ( $\rho_{wet}$ ) in  $\text{gm/cm}^3$ , moisture (%), dry density ( $\rho_d$ ) in  $\text{gm/cm}^3$ , and Brazilian tensile strength (BTS) in (MPa) were taken as the input features for all developed models.

Later, the measured and predicted output values were organized and plotted to facilitate the performance analysis and correlation of the developed models. The final output was examined using various analytical indices such as  $R^2$ , MAE, MSE, RMSE, and a20-index as performance criteria to analyze and compare the anticipated models and to evaluate the ideal model in terms of data prediction. The 106 data points of the overall dataset were allocated as 70% (74 data points) for training and 30% (32 data points) for testing the model.

Figure 13 illustrates the scatter plots of predicted E of the test data by LightGBM, SVM, Catboost, GBRT, RF, and XGBoost models. The  $R^2$  value of each model is determined according to the test prediction. The  $R^2$  value of LightGBM, SVM, Catboost, GBRT, RF, and XGBoost models is 0.281, 0.32, 0.577, 0.988, 0.989, and 0.999, respectively.

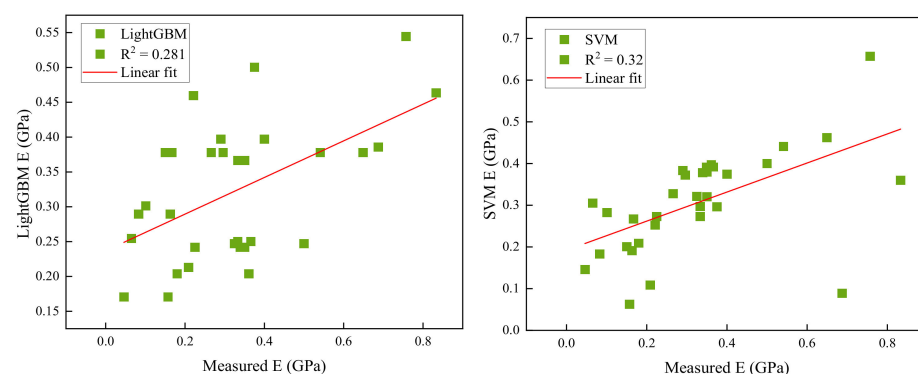
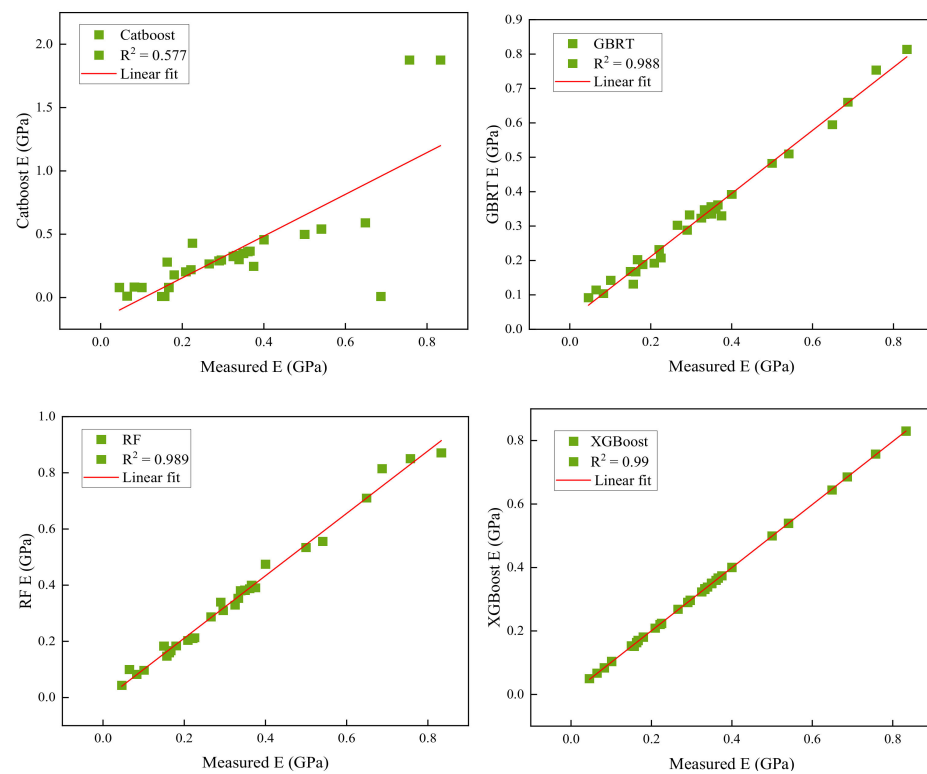


Figure 13. Cont.





**Figure 13.** Scatter plots of E prediction by LightGBM, SVM, Catboost, GBRT, RF, and XGBoost models at the test data.

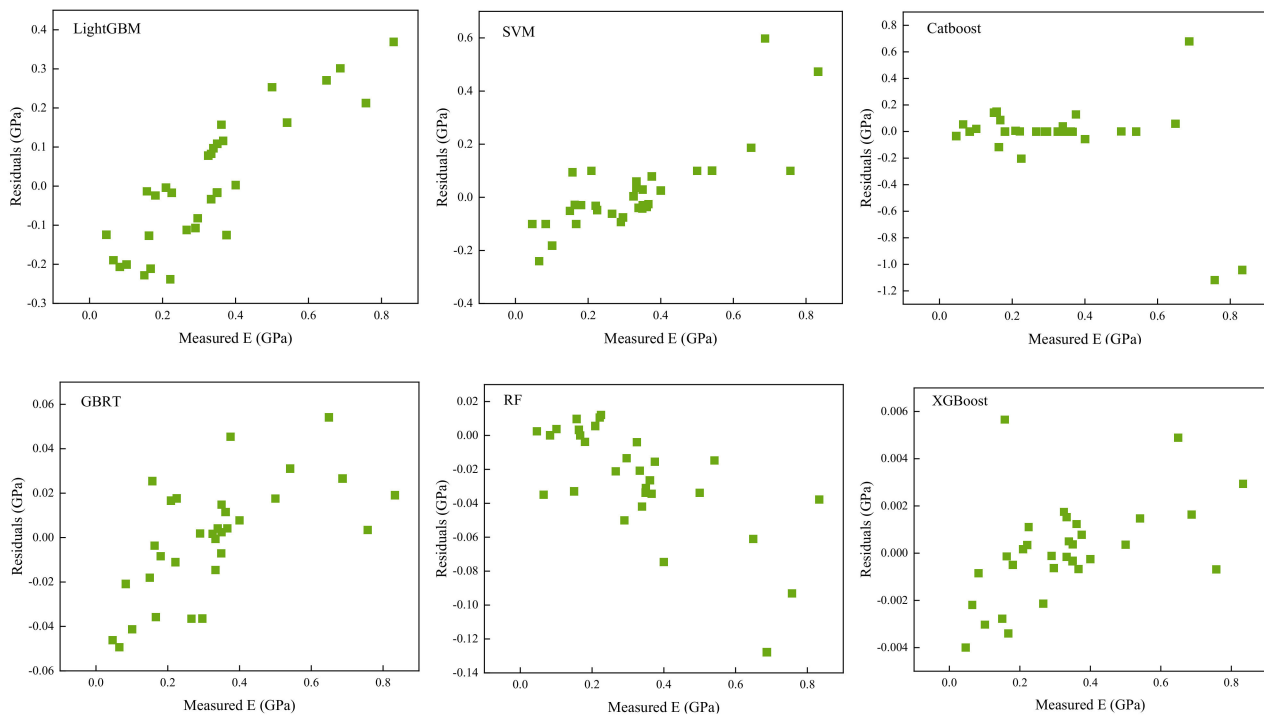
Furthermore, to further understand the performance of the predicted E, it will be interesting to study the prediction rules of six developed ML-based intelligent models due to the wide dispersion of the range of values of E in the established dataset of the test data. The residuals (GPa) and percentage errors (%) of six models were utilized to view the predicting results. The residuals allow for the observation of the contrast between the predicted E and the measured E for each data point, and the percentage error shows the percentage by which the predicted E surpasses the measured E. They are expressed as Equations (21) and (22).

$$r = E_m - E_p \quad (21)$$

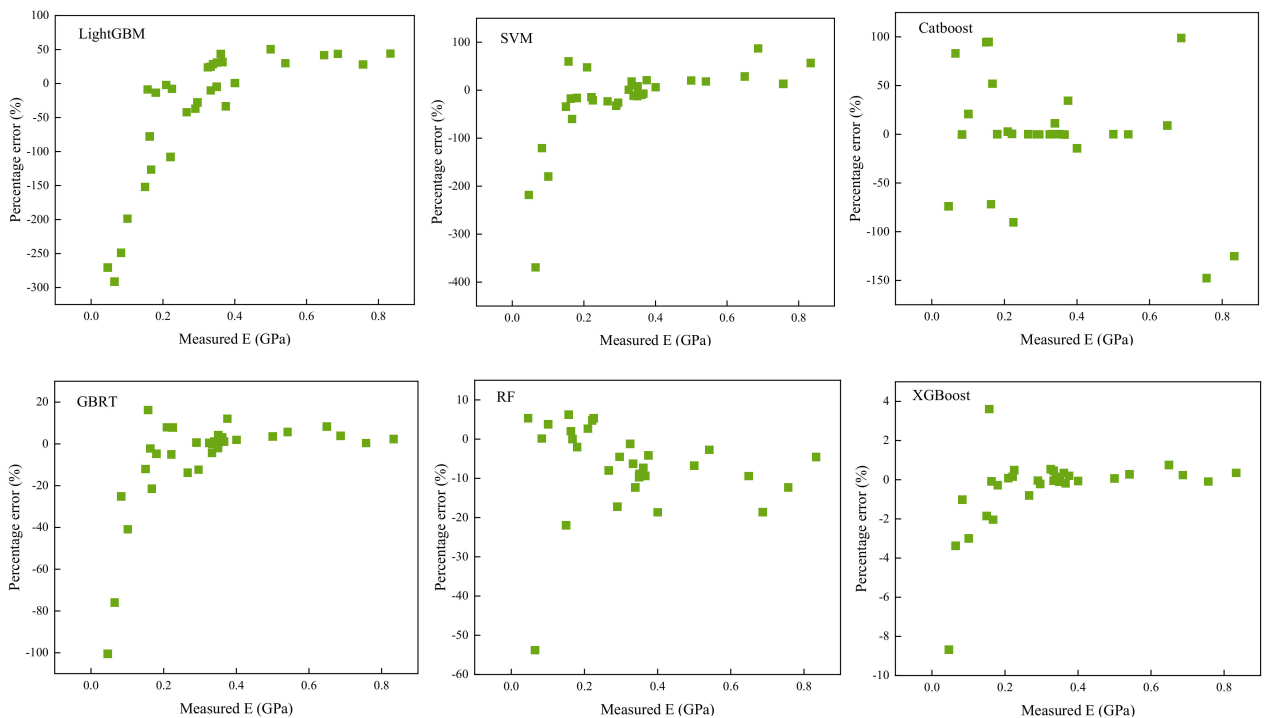
$$p_{error} = \frac{r}{E_m} * 100 \quad (22)$$

where  $r$  = residual in GPa;  $E_m$  and  $E_p$  are the measured and predicted E, respectively; and  $p_{error}$  is the percentage error in %.

In Figure 14, the residuals indicate a direct relationship with the E, since the corresponding residuals can increase as the E increases. In contrast, in Figure 15, the percentage error shows an inverse relationship with the E, because it decreases as the E increases. Some models show negative residuals and percentage errors for smaller E measures and positive values for larger E measures. It revealed that these ML-based intelligent models seem to tend to have a predicted E higher than the measured E when the measured E was small, and tend to have a predicted E smaller than the measured E when the measured E is higher.



**Figure 14.** Residual plots of E prediction by LightGBM, SVM, Catboost, GBRT, RF, and XGBoost models of the test data.



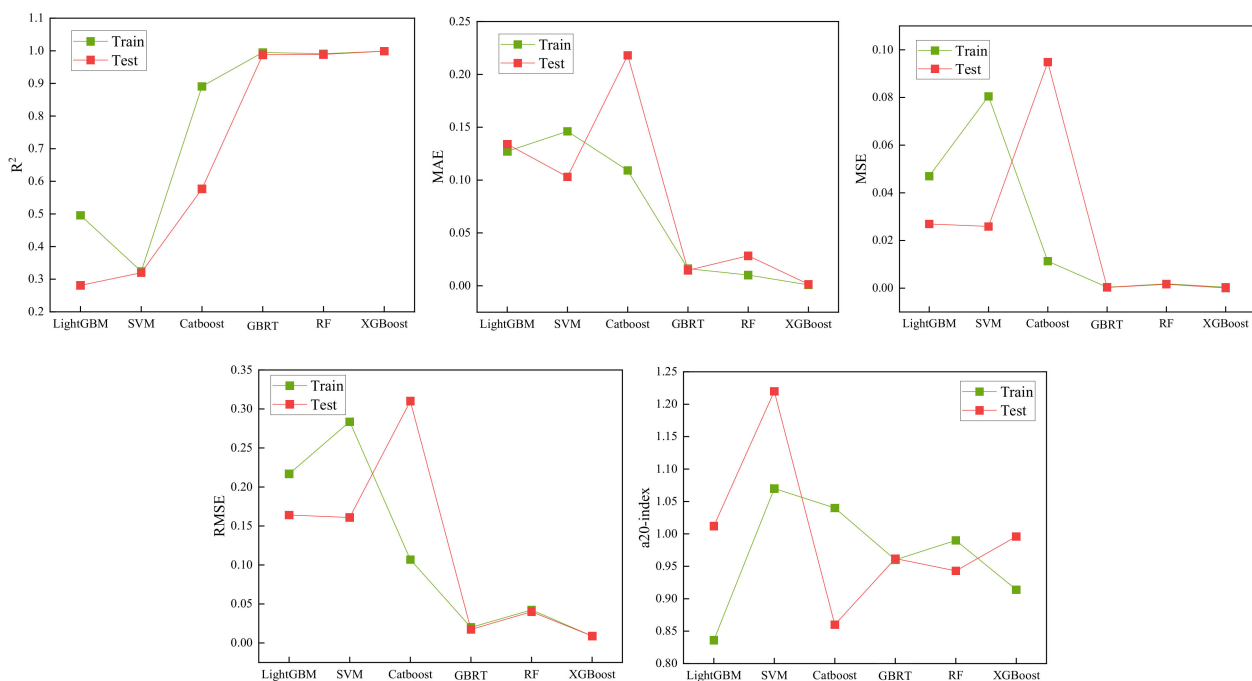
**Figure 15.** Percentage error plots of E prediction by LightGBM, SVM, Catboost, GBRT, RF, and XGBoost models of the test data.

Table 1 exhibits the performance indices of the developed LightGBM, SVM, Catboost, GBRT, RF, and XGBoost models computed by Equation (16) to Equation (20). In this study, according to the proposed LightGBM, SVM, Catboost, GBRT, RF, and XGBoost models, the XGBoost outperformed when at the test data with  $R^2$  of 0.999, MAE of 0.0015, MSE of 0.0008,

RMSE of 0.0089, and a20-index of 0.996, for E prediction. In addition, GBRT and RF have also shown high accuracy and achieved second place next to XGBoost in predicting E, but they can be used conditionally. Therefore, XGBoost is an applicable ML-based intelligent approach that can be applied to accurately predict E, as shown in Figure 16.

**Table 1.** Performance indices of the developed ML-based intelligent models in this study.

Model	Training					Testing				
	R <sup>2</sup>	MAE	MSE	RMSE	a20-Index	R <sup>2</sup>	MAE	MSE	RMSE	a20-Index
LightGBM	0.496	0.1272	0.0470	0.2168	0.836	0.281	0.1340	0.0269	0.1640	1.012
SVM	0.324	0.1461	0.0805	0.2837	1.07	0.32	0.1031	0.0259	0.1609	1.22
Catboost	0.891	0.1091	0.0113	0.1069	1.04	0.577	0.218	0.0948	0.3101	0.86
GBRT	0.995	0.0162	0.0004	0.0200	0.96	0.988	0.0147	0.0003	0.0173	0.962
RF	0.991	0.0102	0.0018	0.0424	0.99	0.989	0.0284	0.0016	0.0400	0.943
XGBoost	0.999	0.0008	0.0004	0.0089	0.914	0.999	0.0015	0.0008	0.0089	0.996



**Figure 16.** Performance indices of the developed LightGBM, SVM, Catboost, GBRT, RF, and XGBoost models of the test data.

The Taylor diagram explains a brief qualitative depiction of the best fit of the model to standard deviations and correlations. The expression for the Taylor diagram is given in Equation (23) [73].

$$R = \frac{\frac{1}{P} \sum_p^P (r_n - \bar{r}) (f_n - \bar{f})}{\sigma_r \sigma_f} \tag{23}$$

where  $R$  denotes a correlation,  $P$  shows the discrete point number,  $r_n$  and  $f_n$  show two vectors,  $\sigma_r$  and  $\sigma_f$  show  $r$  and  $f$  standard deviation, and  $\bar{r}_n$  and  $\bar{f}_n$  illustrate the mean value of vectors  $r_n$  and  $f_n$ , respectively.

Figure 17 represents the correlation between the predicted E and the measured E for the LightGBM, SVM, Catboost, GBRT, RF, and XGBoost models from Figure 16 in terms of standard deviation (STD), RMSE, and  $R^2$ . Based on the consequences, the XGBoost model was highly correlated with measured E than the other models developed in this study in predicting E.

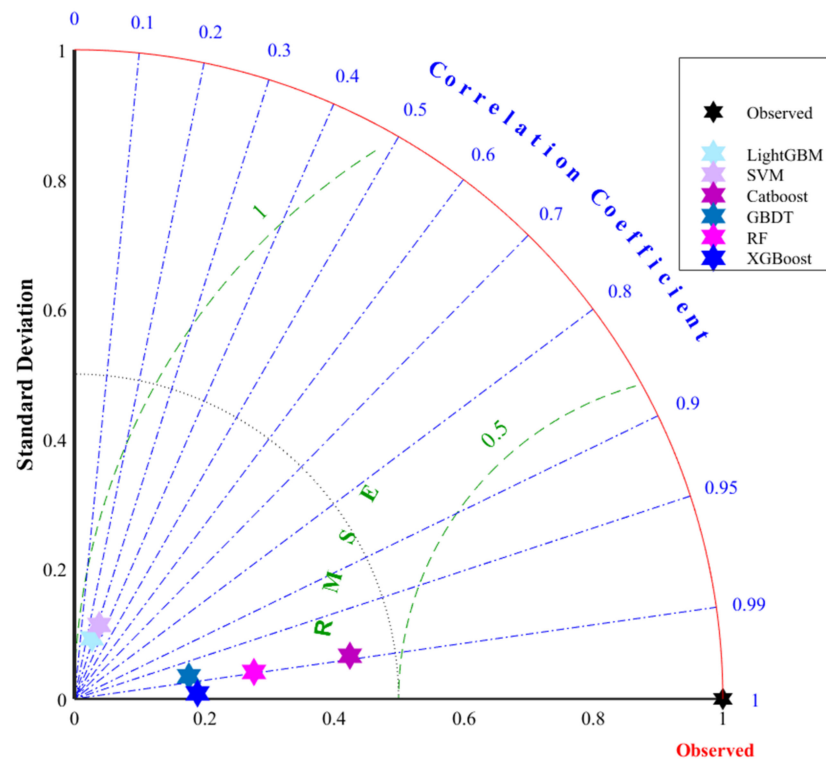


Figure 17. Taylor diagram of the developed LightGBM, SVM, Catboost, GBDT, RF, and XGBoost models of the test data.

Furthermore, the standard deviation (STD) of XGBoost was closest to the measured STD. Thus, compared to the existing published literature [8,74–76], XGBoost exhibits high accuracy and proved to be a highly accurate model for predicting E. The STD of GBTR and RF was also close to the measured STD but indicates the lowest R<sup>2</sup> values. Meanwhile, LightGBM, SVM, and Catboost showed the least correlation and were far from the measured STD.

### 6. Sensitivity Analysis

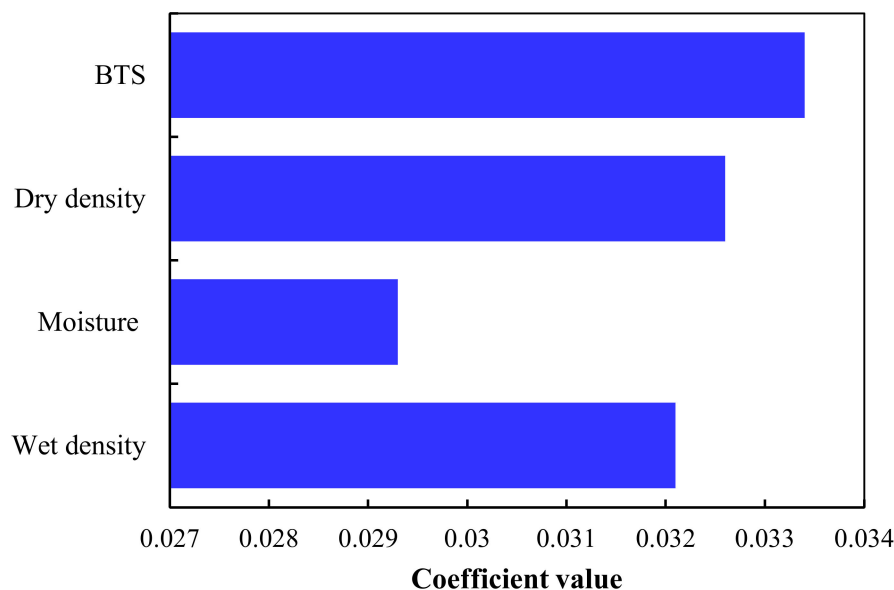
It is very important to correctly evaluate the essential parameters that have a large impact on the E of rock, which is undoubtedly a challenge in the design of rock structures. Thus, in this study, the cosine amplitude method [77,78] was adopted to investigate the relative impact of the inputs over the output. The general formulation of the adopted method is shown in Equation (24).

$$XGBoost_{ij} = \frac{\sum_{k=1}^n (E_{ik}E_{jk})}{\sqrt{\sum_{k=1}^n E_{ik}^2 \sum_{k=1}^n E_{jk}^2}} \tag{24}$$

where E<sub>i</sub> and E<sub>j</sub> are the input and output values, respectively, and n is the number of datasets in the test phase. Finally, the range of XGBoost<sub>ij</sub> is between 0 and 1, additionally proving the precision between each variable and the target. According to Equation (24), if XGBoost<sub>ij</sub> of any parameter has a value of 0, it shows that there is no significant relationship between that parameter and the target. On the contrary, when XGBoost<sub>ij</sub> is equal to 1 or nearly 1, it can be considered a significant relationship that has a large effect on the E of the rock.

Because of the high accuracy of the XGBoost model in predicting E, only a sensitivity analysis was performed on it at the testing level. Figure 18 shows the relationship between each input parameter of the developed model and output. Therefore, it can be seen from the figure that all parameters are positively correlated, while BTS is the most influential

parameter in predicting E. The feature importance of each input parameter is given as  $\rho_{wet} = 0.0321$ , moisture = 0.0293,  $\rho_d = 0.0326$ , and BTS = 0.0334.



**Figure 18.** The effect of input variables on the result of the established XGBoost model.

## 7. Conclusions

Elastic modulus (E) plays a key role in the designing of any rock engineering project. Therefore, an accurate determination of E is a prerequisite. In this study, six novel ML-based intelligent models, namely, LightGBM, SVM, Catboost, GBRT, RF, and XGBoost, were developed to predict E, including four input features, namely,  $\rho_{wet}$ , moisture,  $\rho_d$ , and BTS. To avoid overfitting of these models, the original dataset was distributed into 70% for the training and 30% for the testing of 106 data points. The study concludes that the XGBoost model performed more accurately than the other developed models, such as LightGBM, SVM, Catboost, GBRT, and RF, in predicting E with  $R^2$ , MAE, MSE, RMSE, and a20-index values of 0.999, 0.0015, 0.0008, 0.0089, and 0.996 of the test data, respectively. By employing the ML-based intelligent approach, this study was able to provide alternative elucidations for predicting E with appropriate accuracy and run time.

In future rock engineering projects, it is highly recommended to undertake proper field investigations prior to decision making. The XGBoost ML-based intelligent model performed well to predict the E. The conclusions of GBRT and RF are also applicable for the prediction of E; however, these methods can be used conditionally. Thus, for a large-scale study, this study recommends an adequate dataset to overcome the above limitation. In order to undertake other projects, the model proposed in this study should be considered as a foundation and the result should be reanalyzed, reevaluated, and even re-addressed.

**Author Contributions:** Conceptualization, N.M.S.; methodology, N.M.S.; software, X.G.; validation, X.W.; formal analysis, X.G.; investigation, N.M.S., X.Z.; resources, X.Z.; data curation, N.M.S., X.G.; writing—original draft preparation, N.M.S.; writing—review and editing, N.M.S., X.Z.; visualization, N.M.S., X.G.; supervision, X.Z.; project administration, X.Z.; funding acquisition, X.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Science and Technology Innovation Project of Guizhou Province (Qiankehe Platform Talent (2019) 5620 to X.Z.). No additional external funding was received for this study.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data will be available on request by the corresponding author.

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

## References

1. Davarpanah, M.; Somodi, G.; Kovács, L.; Vásárhelyi, B. Complex analysis of uniaxial compressive tests of the Mórág granitic rock formation (Hungary). *Stud. Geotech. Mech.* **2019**, *41*, 21–32. [\[CrossRef\]](#)
2. Xiong, L.X.; Xu, Z.Y.; Li, T.B.; Zhang, Y. Bonded-particle discrete element modeling of mechanical behaviors of interlayered rock mass under loading and unloading conditions. *Geomech. Geophys. Geo-Energy Geo-Resour.* **2019**, *5*, 1–16. [\[CrossRef\]](#)
3. Rahimi, R.; Nygaard, R. Effect of rock strength variation on the estimated borehole breakout using shear failure criteria. *Geomech. Geophys. Geo-Energy Geo-Resour.* **2008**, *4*, 369–382. [\[CrossRef\]](#)
4. Zhao, Y.S.; Wan, Z.J.; Feng, Z.J.; Xu, Z.H.; Liang, W.G. Evolution of mechanical properties of granite at high temperature and high pressure. *Geomech. Geophys. Geo-Energy Geo-Resour.* **2017**, *3*, 199–210. [\[CrossRef\]](#)
5. Jing, H.; Rad, H.N.; Hasanipanah, M.; Armaghani, D.J.; Qasem, S.N. Design and implementation of a new tuned hybrid intelligent model to predict the uniaxial compressive strength of the rock using SFS-ANFIS. *Eng. Comput.* **2021**, *37*, 2717–2734. [\[CrossRef\]](#)
6. Lindquist, E.S.; Goodman, R.E. Strength and deformation properties of a physical model melange. In Proceedings of the 1st North American Rock Mechanics Symposium, Austin, TX, USA, 1–3 June 1994; Nelson, P.P., Laubach, S.E., Eds.; Balkema: Rotterdam, The Netherlands, 1994.
7. Singh, T.N.; Dubey, R.K. A study of transmission velocity of primary wave (P-Wave) in Coal Measures sandstone. *J. Sci. Ind. Res.* **2000**, *59*, 482–486.
8. Tiryaki, B. Predicting intact rock strength for mechanical excavation using multivariate statistics, artificial neural networks and regression trees. *Eng. Geol.* **2008**, *99*, 51–60. [\[CrossRef\]](#)
9. Ozcelik, Y.; Bayram, F.; Yasitli, N.E. Prediction of engineering properties of rocks from microscopic data. *Arab. J. Geosci.* **2013**, *6*, 3651–3668. [\[CrossRef\]](#)
10. Abdi, Y.; Garavand, A.T.; Sahamieh, R.Z. Prediction of strength parameters of sedimentary rocks using artificial neural networks and regression analysis. *Arab. J. Geosci.* **2018**, *11*, 587. [\[CrossRef\]](#)
11. Teymen, A.; Mengüç, E.C. Comparative evaluation of different statistical tools for the prediction of uniaxial compressive strength of rocks. *Int. J. Min. Sci. Technol.* **2020**, *30*, 785–797. [\[CrossRef\]](#)
12. Li, C.; Zhou, J.; Armaghani, D.J.; Li, X. Stability analysis of underground mine hard rock pillars via combination of finite difference methods, neural networks, and Monte Carlo simulation techniques. *Undergr. Space* **2021**, *6*, 379–395. [\[CrossRef\]](#)
13. Momeni, E.; Yarivand, A.; Dowlatshahi, M.B.; Armaghani, D.J. An efficient optimal neural network based on gravitational search algorithm in predicting the deformation of geogrid-reinforced soil structures. *Transp. Geotech.* **2021**, *26*, 100446. [\[CrossRef\]](#)
14. Parsajoo, M.; Armaghani, D.J.; Mohammed, A.S.; Khari, M.; Jahandari, S. Tensile strength prediction of rock material using non-destructive tests: A comparative intelligent study. *Transp. Geotech.* **2021**, *31*, 100652. [\[CrossRef\]](#)
15. Armaghani, D.J.; Harandizadeh, H.; Momeni, E.; Maizir, H.; Zhou, J. An optimized system of GMDH-ANFIS predictive model by ICA for estimating pile bearing capacity. *Artif. Intell. Rev.* **2021**, *55*, 2313–2350. [\[CrossRef\]](#)
16. Harandizadeh, H.; Armaghani, D.J. Prediction of air-overpressure induced by blasting using an ANFIS-PNN model optimized by GA. *Appl. Soft Comput.* **2021**, *99*, 106904. [\[CrossRef\]](#)
17. Cao, J.; Gao, J.; Rad, H.N.; Mohammed, A.S.; Hasanipanah, M.; Zhou, J. A novel systematic and evolved approach based on XGBoost-firefly algorithm to predict Young's modulus and unconfined compressive strength of rock. *Eng. Comput.* **2021**, 1–17. [\[CrossRef\]](#)
18. Yang, F.; Li, Z.; Wang, Q.; Jiang, B.; Yan, B.; Zhang, P.; Xu, W.; Dong, C.; Liaw, P.K. Cluster-formula-embedded machine learning for design of multicomponent  $\beta$ -Ti alloys with low Young's modulus. *npj Comput. Mater.* **2020**, *6*, 1–11. [\[CrossRef\]](#)
19. Duan, J.; Asteris, P.G.; Nguyen, H.; Bui, X.N.; Moayedi, H. A novel artificial intelligence technique to predict compressive strength of recycled aggregate concrete using ICA-XGBoost model. *Eng. Comput.* **2020**, *37*, 3329–3346. [\[CrossRef\]](#)
20. Pham, B.T.; Nguyen, M.D.; Nguyen-Thoi, T.; Ho, L.S.; Koopialipoor, M.; Quoc, N.K.; Armaghani, D.J.; Van Le, H. A novel approach for classification of soils based on laboratory tests using Adaboost, Tree and ANN modeling. *Transp. Geotech.* **2021**, *27*, 100508. [\[CrossRef\]](#)
21. Asteris, P.G.; Mamou, A.; Hajihassani, M.; Hasanipanah, M.; Koopialipoor, M.; Le, T.T.; Kardani, N.; Armaghani, D.J. Soft computing based closed form equations correlating L and N-type Schmidt hammer rebound numbers of rocks. *Transp. Geotech.* **2021**, *29*, 100588. [\[CrossRef\]](#)
22. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Waqas, U.; Ahmed, M.F. Prediction Modeling for the Estimation of Dynamic Elastic Young's Modulus of Thermally Treated Sedimentary Rocks Using Linear-Nonlinear Regression Analysis, Regularization, and ANFIS. *Rock Mech. Rock Eng.* **2020**, *53*, 5411–5428. [\[CrossRef\]](#)
24. Ghasemi, E.; Kalhori, H.; Bagherpour, R.; Yagiz, S. Model tree approach for predicting uniaxial compressive strength and Young's modulus of carbonate rocks. *Bull. Eng. Geol. Environ.* **2018**, *77*, 331–343. [\[CrossRef\]](#)

25. Shahani, N.M.; Zheng, X.; Liu, C.; Hassan, F.U.; Li, P. Developing an XGBoost Regression Model for Predicting Young's Modulus of Intact Sedimentary Rocks for the Stability of Surface and Subsurface Structures. *Front. Earth Sci.* **2021**, *9*, 761990. [[CrossRef](#)]
26. Ceryan, N. Prediction of Young's modulus of weathered igneous rocks using GRNN, RVM, and MPMR models with a new index. *J. Mt. Sci.* **2021**, *18*, 233–251. [[CrossRef](#)]
27. Umrao, R.K.; Sharma, L.K.; Singh, R.; Singh, T.N. Determination of strength and modulus of elasticity of heterogenous sedimentary rocks: An ANFIS predictive technique. *Measurement* **2018**, *126*, 194–201. [[CrossRef](#)]
28. Davarpanah, S.M.; Ván, P.; Vásárhelyi, B. Investigation of the relationship between dynamic and static deformation moduli of rocks. *Geomech. Geophys. Geo-Energy Geo-Resour.* **2020**, *6*, 29. [[CrossRef](#)]
29. Aboutaleb, S.; Behnia, M.; Bagherpour, R.; Bluekian, B. 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.* **2018**, *77*, 1717–1728. [[CrossRef](#)]
30. Mahmoud, A.A.; Elkatatny, S.; Ali, A.; Moussa, T. Estimation of static young's modulus for sandstone formation using artificial neural networks. *Energies* **2019**, *12*, 2125. [[CrossRef](#)]
31. Roy, D.G.; Singh, T.N. Regression and soft computing models to estimate young's modulus of CO<sub>2</sub> saturated coals. *Measurement* **2018**, *129*, 91–101.
32. Armaghani, D.J.; Mohamad, E.T.; Momeni, E.; Narayanasamy, M.S. An adaptive neuro-fuzzy inference system for predicting unconfined compressive strength and Young's modulus: A study on Main Range granite. *Bull. Eng. Geol. Environ.* **2015**, *74*, 1301–1319. [[CrossRef](#)]
33. Singh, R.; Kainthola, A.; Singh, T.N. Estimation of elastic constant of rocks using an ANFIS approach. *Appl. Soft Comput.* **2012**, *12*, 40–45. [[CrossRef](#)]
34. Köken, E. Assessment of Deformation Properties of Coal Measure Sandstones through Regression Analyses and Artificial Neural Networks. *Arch. Min. Sci.* **2021**, *66*, 523–542.
35. Yesiloglu-Gultekin, N.; Gokceoglu, C. A Comparison Among Some Non-linear Prediction Tools on Indirect Determination of Uniaxial Compressive Strength and Modulus of Elasticity of Basalt. *J. Nondestruct. Eval.* **2022**, *41*, 10. [[CrossRef](#)]
36. Awais Rashid, H.M.; Ghazzali, M.; Waqas, U.; Malik, A.A.; Abubakar, M.Z. Artificial Intelligence-Based Modeling for the Estimation of Q-Factor and Elastic Young's Modulus of Sandstones Deteriorated by a Wetting-Drying Cyclic Process. *Arch. Min. Sci.* **2021**, *66*, 635–658.
37. Matin, S.S.; Farahzadi, L.; Makaremi, S.; Chelgani, S.C.; Sattari, G. Variable selection and prediction of uniaxial compressive strength and modulus of elasticity by random forest. *Appl. Soft Comput.* **2018**, *70*, 980–987. [[CrossRef](#)]
38. Yang, L.; Feng, X.; Sun, Y. Predicting the Young's Modulus of granites using the Bayesian model selection approach. *Bull. Eng. Geol. Environ.* **2019**, *78*, 3413–3423. [[CrossRef](#)]
39. Ren, Q.; Wang, G.; Li, M.; Han, S. Prediction of rock compressive strength using machine learning algorithms based on spectrum analysis of geological hammer. *Geotech. Geol. Eng.* **2019**, *37*, 475–489. [[CrossRef](#)]
40. Ge, Y.; Xie, Z.; Tang, H.; Du, B.; Cao, B. Determination of the shear failure areas of rock joints using a laser scanning technique and artificial intelligence algorithms. *Eng. Geol.* **2021**, *293*, 106320. [[CrossRef](#)]
41. Xu, C.; Liu, X.; Wang, E.; Wang, S. Calibration of the microparameters of rock specimens by using various machine learning algorithms. *Int. J. Geomech.* **2021**, *21*, 04021060. [[CrossRef](#)]
42. Shahani, N.M.; Wan, Z.; Guichen, L.; Siddiqui, F.I.; Pathan, A.G.; Yang, P.; Liu, S. Numerical analysis of top coal recovery ratio by using discrete element method. *Pak. J. Eng. Appl. Sci.* **2019**, *24*, 26–35.
43. Shahani, N.M.; Wan, Z.; Zheng, X.; Guichen, L.; Liu, C.; Siddiqui, F.I.; Bin, G. Numerical modeling of longwall top coal caving method at thar coalfield. *J. Met. Mater. Miner.* **2020**, *30*, 57–72.
44. Shahani, N.M.; Kamran, M.; Zheng, X.; Liu, C.; Guo, X. Application of Gradient Boosting Machine Learning Algorithms to Predict Uniaxial Compressive Strength of Soft Sedimentary Rocks at Thar Coalfield. *Adv. Civ. Eng.* **2021**, *2021*, 2565488. [[CrossRef](#)]
45. Brown, E.T. *Rock Characterization Testing & Monitoring—ISRM Suggested Methods, ISRM—International Society for Rock Mechanics*; Pergamon Press: London, UK, 2007; Volume 211.
46. *D4543-85*; Standard Practices for Preparing Rock Core as Cylindrical Test Specimens and Verifying Conformance to Dimensional and Shape Tolerances. ASTM—American Society for Testing and Materials: West Conshohocken, PA, USA, 2013.
47. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. Lightgbm: A highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3146–3154.
48. Zeng, H.; Yang, C.; Zhang, H.; Wu, Z.H.; Zhang, M.; Dai, G.J.; Babiloni, F.; Kong, W.Z. A lightGBM-based EEG analysis method for driver mental states classification. *Comput. Intell. Neurosci.* **2019**, *2019*, 3761203. [[CrossRef](#)]
49. Liang, W.; Luo, S.; Zhao, G.; Wu, H. Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms. *Mathematics* **2020**, *8*, 765. [[CrossRef](#)]
50. Vapnik, V.; Golowich, S.E.; Smola, A. Support vector method for function approximation, regression estimation, and signal processing. *Adv. Neural Inf. Process. Syst.* **1997**, *9*, 281–287.
51. Sun, J.; Zhang, J.; Gu, Y.; Huang, Y.; Sun, Y.; Ma, G. Prediction of permeability and unconfined compressive strength of pervious concrete using evolved support vector regression. *Constr. Build. Mater.* **2019**, *207*, 440–449. [[CrossRef](#)]

52. Negara, A.; Ali, S.; AlDhamen, A.; Kesserwan, H.; Jin, G. Unconfined compressive strength prediction from petrophysical properties and elemental spectroscopy using support-vector regression. In Proceedings of the SPE Kingdom of Saudi Arabia Annual Technical Symposium and Exhibition, Dammam, Saudi Arabia, 24–27 April 2017.
53. Xu, C.; Amar, M.N.; Ghriga, M.A.; Ouaer, H.; Zhang, X.; Hasanipanah, M. Evolving support vector regression using Grey Wolf optimization; forecasting the geomechanical properties of rock. *Eng. Comput.* **2020**, *1*–15. [[CrossRef](#)]
54. Barzegar, R.; Sattarpour, M.; Nikudel, M.R.; Moghaddam, A.A. Comparative evaluation of artificial intelligence models for prediction of uniaxial compressive strength of travertine rocks, case study: Azarshahr area, NW Iran. *Model. Earth Syst. Environ.* **2016**, *2*, 76. [[CrossRef](#)]
55. Dong, L.; Li, X.; Xu, M.; Li, Q. Comparisons of random forest and support vector machine for predicting blasting vibration characteristic parameters. *Procedia Eng.* **2011**, *26*, 1772–1781.
56. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
57. Dorogush, A.V.; Ershov, V.; Gulin, A. CatBoost: Gradient boosting with categorical features support. *arXiv* **2018**, arXiv:1810.11363.
58. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. *Adv. Neural Inf. Process. Syst.* **2018**, *31*, 6638–6648.
59. Freund, Y.; Schapire, R.; Abe, N. A short introduction to boosting. *J. Jpn. Soc. Artif. Intell.* **1999**, *14*, 771–780.
60. Schapire, R.E. The strength of weak learnability. *Mach. Learn.* **1990**, *5*, 197–227. [[CrossRef](#)]
61. Kearns, M. Thoughts on Hypothesis Boosting. Mach. Learn. Class Proj.. 1988, pp. 1–9. Available online: <https://www.cis.upenn.edu/~jmkerns/papers/boostnote.pdf> (accessed on 10 February 2022).
62. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [[CrossRef](#)]
63. Friedman, J.; Hastie, T.; Robert, T. The elements of statistical learning. In *Statistics*; Springer: New York, NY, USA, 2001; Volume 1.
64. Friedman, J.H. Stochastic gradient boosting. *Comput. Stat. Data Anal.* **2002**, *38*, 367–378. [[CrossRef](#)]
65. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
66. Yang, P.; Hwa, Y.; Zhou, B.; Zomaya, A.Y. A review of ensemble methods in bioinformatics. *Curr. Bioinform.* **2010**, *5*, 296–308. [[CrossRef](#)]
67. Meng, Q.; Ke, G.; Wang, T.; Chen, W.; Ye, Q.; Ma, Z.M.; Liu, T.Y. A communication-efficient parallel algorithm for decision tree. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 1271–1279.
68. Ranka, S.; Singh, V. Clouds: A decision tree classifier for large datasets. In Proceedings of the 4th Knowledge Discovery and Data Mining Conference, New York, NY, USA, 27–31 August 1998; pp. 2–8.
69. Jin, R.; Agrawal, G. Communication and memory efficient parallel decision tree construction. In Proceedings of the 2003 SIAM International Conference on Data Mining, San Francisco, CA, USA, 1–3 May 2003; pp. 119–129.
70. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
71. Shahani, N.M.; Kamran, M.; Zheng, X.; Liu, C. Predictive modeling of drilling rate index using machine learning approaches: LSTM, simple RNN, and RFA. *Pet. Sci. Technol.* **2022**, *40*, 534–555. [[CrossRef](#)]
72. Willmott, C.J. Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.* **1982**, *63*, 1309–1313. [[CrossRef](#)]
73. Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.* **2001**, *106*, 7183–7192. [[CrossRef](#)]
74. Zhong, R.; Tsang, M.; Makusha, G.; Yang, B.; Chen, Z. Improving rock mechanical properties estimation using machine learning. In *Proceedings of the 2021 Resource Operators Conference, Wollongong, Australia, 10–12 February 2021*; University of Wollongong-Mining Engineering: Wollongong, Australia, 2021.
75. Ghose, A.K.; Chakraborti, S. Empirical strength indices of Indian coals. In Proceedings of the 27th U.S. Symposium on Rock Mechanics, Tuscaloosa, AL, USA, 23–25 June 1986.
76. Katz, O.; Reches, Z.; Roegiers, J.C. Evaluation of mechanical rock properties using a Schmidt Hammer. *Int. J. Rock Mech. Min. Sci.* **2000**, *37*, 723–728. [[CrossRef](#)]
77. Momeni, E.; Nazir, R.; Armaghani, D.J.; Maizir, H. Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN. *Measurement* **2014**, *57*, 122–131. [[CrossRef](#)]
78. Ji, X.; Liang, S.Y. Model-based sensitivity analysis of machining-induced residual stress under minimum quantity lubrication. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2007**, *231*, 1528–1541. [[CrossRef](#)]